TAREAS APOYADAS EN EXCEL PARA PROMOVER LA INTERPRETACIÓN DE LAS MEDIDAS DE DISPERSIÓN

GERALDINE GISSEL VARGAS DAZA

SERGIO DAVID MELO CASTAÑEDA

UNIVERSIDAD PEDAGÓGICA NACIONAL

FACULTAD DE CIENCIA Y TECNOLOGÍA

DEPARTAMENTO DE MATEMÁTICAS

MAESTRÍA EN DOCENCIA DE LA MATEMÁTICA

BOGOTÁ, D.C. 2024

TAREAS APOYADAS EN EXCEL PARA PROMOVER LA INTERPRETACIÓN DE LAS MEDIDAS DE DISPERSIÓN

GERALDINE GISSEL VARGAS DAZA

SERGIO DAVID MELO CASTAÑEDA

Trabajo de Grado como requisito para optar al título de

Magíster en Docencia de la Matemática

Director

Mg. CÉSAR GUILLERMO RENDÓN MAYORGA

UNIVERSIDAD PEDAGÓGICA NACIONAL

FACULTAD DE CIENCIA Y TECNOLOGÍA

DEPARTAMENTO DE MATEMÁTICAS

MAESTRÍA EN DOCENCIA DE LA MATEMÁTICA

BOGOTÁ, D.C. 2024

DEDICATORIA

A mi hija, el motor de mi vida.
A mi familia, mi apoyo incondicional.
A mi pareja, mi futuro.
A mí, por perseverar y creer ciegamente en mis capacidades.

AGRADECIMIENTOS

A mi hija, gracias por soportar mis días de ausencia, los juegos no jugados y las películas dormidas, por siempre sentirte orgullosa de mí y por ese amor incondicional que me brindas.

A mis padres y hermanos, por su constante apoyo y por conformar la familia más polivalente, fuerte y unida del mundo. Su presencia en mi vida ha sido un pilar fundamental en este camino.

A Valentina, mi pareja, mi compañera, mi amiga, mi vida, por su amor incondicional. Gracias por soportar mis días de ausencia y por estar siempre a mi lado, brindándome fortaleza y comprensión en cada paso.

A mi amiga y colega Geraldine, por ser mi compañera y casi hermana. Gracias por siempre estar presente en mi proceso académico y por ese apoyo constante, disposición y confianza que han sido claves para sacar adelante este proyecto.

Al profesor César, por su valiosa orientación y dedicación. Gracias por su guía y compromiso, los cuales fueron esenciales para culminar este trabajo.

Finalmente, a la Universidad Pedagógica Nacional, mi alma mater, por brindarme las herramientas y el espacio para hacer realidad este proyecto de siete años, y, sobre todo, por cambiarme la vida.

Sergio David Melo Castañeda

DEDICATORIA

A mi familia, por la unión y el amor incondicional.
A Santiago, este amor verdadero es sempiterno.
A mí, por seguir soñando y cumpliendo metas.

AGRADECIMIENTOS

En este hermoso proceso de formación académica, quiero agradecer a mi familia por su apoyo incondicional, por su amor, y por alentarme cada día para que no me rindiera. Todo lo que un día soñamos se está cumpliendo poco a poco.

A mis padres, por enseñarme desde muy pequeña la importancia de estudiar y ser profesional, por darme palabras de cariño y confianza cuando los días son soleados, pero también cuando los días son grises. Todos mis logros siempre serán una muestra del amor que les tengo, espero siempre se sientan muy orgullosos de la mujer en la que me he convertido.

A mis hermanos y a mi sobrino, por hacerme saber todo el tiempo lo orgullosos que están de mí, por permitirme construir momentos que me hacen recordar que vale la pena seguir luchando por todo aquello que deseamos, y por permitir fortalecer todos los días nuestro vinculo de hermandad, pero en especial de amistad.

A mi compañero de investigación, por brindarme su amistad durante tantos años y por confiar en mi trabajo para conformar este equipo de trabajo por segunda vez. Tal y como lo imaginamos hace algunos años, nuevamente logramos dar este paso profesional apoyándonos uno al otro.

A mi director de trabajo de grado, por estar siempre dispuesto a darnos su apoyo durante este proceso, por su paciencia y por todo el tiempo que se tomó para ayudarnos a construir esta investigación.

Gracias también a mi gloriosa Universidad Pedagógica Nacional, porque es en este lugar que me he formado profesionalmente y fue aquí donde empezó mi sueño de ser docente.

Geraldine Gissel Vargas Daza

Para todos los efectos, declaramos que el presente trabajo es original y de nuestra total autoría: en aquellos casos en los cuales hemos requerido del trabajo de otros autores o investigadores, hemos dado los respectivos créditos. (Acuerdo 031 de 2007. Artículo 42. Parágrafo 2)

Tabla de Contenido

1. ASPECTOS GENERALES 1

1.1. Introducción 1

1.2. Antecedentes 2

1.3. Delimitación del problema 5

1.4. Objetivos 9

1.4.1. Objetivo general 10

1.4.2. Objetivos específicos 10

2. MARCO DE REFERENCIA 11

2.1. Competencia interpretativa 11

2.1.1. Competencia 12

2.1.2. Competencias matemáticas 15

2.1.3. Competencia interpretativa 17

2.2. Medidas de dispersión 24

2.3. Tecnologías digitales 50

2.3.1. Microsoft Excel 51

2.3.2. Bases de datos 53

2.3.3. Génesis instrumental 55

2.4. Tareas de aprendizaje 59

3. ASPECTOS METODOLÓGICOS 61

3.1. Revisión documental 61

3.2. Experimento de enseñanza 64

3.3. Diseño de las tareas 68

3.3.1. Tarea 1 68

3.3.2. Tarea 2 76

3.3.3. Tarea 3 82

3.4. Categorías de análisis 89

3.5. Implementación de las tareas 92

4. RESULTADOS 95

4.1. Tarea 1 95

4.2. Tarea 2 112

4.3. Tarea 3 123

5. CONCLUSIONES 140

6. REFERENCIAS 143

7. ANEXOS 149

7.1. Anexo 1 149

7.2. Anexo 2 161

7.3. Anexo 3 166

Figuras

Figura 1. Respuesta de estudiante en relación con la interpretación de la media 6
Figura 2. Gráfico estadístico en relación con las respuestas correctas e incorrectas obtenidas para la pregunta de la Figura 1. 7
Figura 3. Enunciado de ejercicio propuesto a estudiantes en relación con medidas estadísticas 7
Figura 4. Respuesta de estudiante a ejercicio ilustrado en la figura 3 (inciso c). 8
Figura 5. Concepción de interpretación a través de concepto de probabilidad. 18
Figura 6. Forma de una distribución normal (Campana) 44
Figura 7. Representación gráfica de la distribución de datos bajo la interpretación de la desviación estándar. 48
Figura 8. Ilustración de una base de datos. 54
Figura 9. Cálculo de media y varianza usando las funciones denominadas operaciones matemáticas 57
Figura 10. Gráfico de dispersión elaborados en Excel en relación con dos conjuntos de datos 57
Figura 11. Cálculo de medidas estadísticas en Excel a partir de funciones denominadas operaciones matemáticas 58
Figura 12. Plan de ejecución para la revisión documental. 62
Figura 13. Plan de ejecución para el experimento de enseñanza. 66
Figura 14. Ejemplar de la Muestra 1 correspondiente a la base de datos Licores Risaralda. 71
Figura 15. Ejemplar de la base de datos Velocidad y tiempo de viaje GT. 78
Figura 16. Ejemplar de la base de datos Notas de los Exámenes 2023-II. 84
Figura 17. Ejemplar de la base de datos Calidad del Agua. 85
Figura 18. Respuesta de estudiante para la tabla T1- 3. 95
Figura 19. Respuestas de estudiantes para el ítem T1 – 3A. 96
Figura 20 . Respuestas de estudiantes para el ítem T1 – 3c en correspondencia con C6. 98
Figura 21. Respuestas de estudiantes para el ítem T1 – 3c en correspondencia con C2. 98
Figura 22. Respuesta de estudiante para la tabla T1- 4. 98
Figura 23. Respuestas de estudiantes para el ítem T1 – 4b que cumplen C9 y C13. 100
Figura 24. Respuesta de estudiante para la tabla T1- 5. 101
Figura 25. Respuesta de estudiante para la tabla T1- 6. 101
Figura 26. Respuestas de estudiantes para el ítem T1 – 6b que cumplen C10 102
Figura 27. Respuestas de estudiantes para el ítem T1 – 6b que cumplen C12 102
Figura 28. Respuesta de estudiante para el ítem T1 – 6c. 103
Figura 29. Respuestas de estudiantes para el ítem T1 – 7b. 107
Figura 30. Respuestas de estudiantes para el ítem T1 – 9 108
Figura 31. Respuestas de estudiantes para el ítem T1 – 11. 108
Figura 32. Respuestas de estudiantes para el ítem T1 – 13. 111
Figura 33. Respuestas de estudiantes para el ítem T1 – 13. 111
Figura 34. Respuestas de los estudiantes T2-3. 113
Figura 35. Respuestas de los estudiantes T2-11 119
Figura 36. Respuesta de los estudiantes T2-13. 121
Figura 37. Repuesta de estudiante a T2 -15. 121
Figura 38. Respuesta de estudiante #2 a T2 -15. 122
Figura 39. Desviación estándar muestral vs poblacional. 122
Figura 40. Medidas estadísticas a partir de funciones en Excel 124
Figura 41. Respuestas de los estudiantes T3-3**.** 125
Figura 42. Respuestas de los estudiantes T3-6. 127
Figura 43. Respuestas de los estudiantes T3-7. 128
Figura 44. Respuestas de los estudiantes T3-13C. 133

Tablas

Tabla 1. Definiciones de competencia……..……...……12
Tabla 2. Niveles del componente interpretación de resultados de la GAISE…….……...……20
Tabla 3. Definiciones de estadística……..……..24
Tabla 4. Tabla de frecuencia para datos no agrupados……..……..32
Tabla 5. Ejemplo de tabla de frecuencia para datos no agrupados……....32
Tabla 6. Tabla de frecuencia para datos agrupados……..……..34
Tabla 7. Definiciones de gráficos estadísticos……..……..…..35
Tabla 8. Fórmulas de medidas estadísticas para datos agrupados……..……...50
Tabla 9. Funciones de Microsoft Excel……..……..52
Tabla 10. Preguntas de la tarea 1 que no se relacionan con los procesos de la génesis instrumental……..72
Tabla 11. Preguntas de la tarea 1 que se relacionan con el proceso de instrumentación…….....74
Tabla 12. Clasificación de las preguntas de la tarea 1 en relación con características de la competencia interpretativa ……...75
Tabla 13. Clasificación de las preguntas de la tarea 1 en relación con procesos de la génesis instrumental ….....76
Tabla 14. Preguntas de la tarea 2 que no se relacionan con los procesos de la génesis instrumental……..79
Tabla 15. Preguntas de la tarea 2 que se relacionan con el proceso de instrumentación……80
Tabla 16. Preguntas de la tarea 2 que se relacionan con el proceso de instrumentalización……..80
Tabla 17. Clasificación de las preguntas de la tarea 2 en relación con características de la competencia interpretativa ……...81
Tabla 18. Clasificación de las preguntas de la tarea 2 en relación con procesos de la génesis instrumental ….....82
Tabla 19. Preguntas de la tarea 3 que no se relacionan con los procesos de la génesis instrumental……..85
Tabla 20. Preguntas de la tarea 3 que se relacionan con el proceso de instrumentación……87
Tabla 21. Clasificación de las preguntas de la tarea 3 en relación con características de la competencia interpretativa ……...88
Tabla 22. Clasificación de las preguntas de la tarea 3 en relación con procesos de la génesis instrumental ….....88
Tabla 23. Categorías de análisis diseñadas para las tareas……90
Tabla 24. Sesiones de clase y tiempos para realizar las tareas……...93
Tabla 25. Relación entre características y respuestas de los estudiantes para el ítem T1 – 3B……..97
Tabla 26. Relación entre características y respuestas de los estudiantes para el ítem T1 – 4A……..99
Tabla 27. Respuestas de los estudiantes para el ítem T1 – 4C……100
Tabla 28. Respuestas de los estudiantes para el ítem T1 – 6A……101
Tabla 29. Relación entre características y respuestas de los estudiantes para el ítem T1 – 7C……104
Tabla 30. Relación entre ítems y respuestas de los estudiantes para las preguntas de T1 -1. ……..105
Tabla 31. Relación entre ítems y respuestas de los estudiantes para las preguntas de T1 -7……106
Tabla 32. Relación entre ítems y respuestas de los estudiantes para las preguntas de T1-7……107
Tabla 33. Relación entre ítems y respuestas de los estudiantes para las preguntas de 10A y 12A de T1……..110
Tabla 34. Características de la Tarea 1. ……112
Tabla 35. Análisis de las preguntas 7, 12 y 16 de T2……114
Tabla 36. Análisis de la pregunta T2 – 5……116
Tabla 37. Análisis de la pregunta T2 – 6……116
Tabla 38. Análisis de la pregunta T2 – 10……..117
Tabla 39. Análisis de las preguntas 4 y 9 de T2……...119
Tabla 40. Características de la Tarea 2……122
Tabla 41. Respuestas de los estudiantes T3-2……124
Tabla 42. Análisis de preguntas 4 y 5 de T3……...125
Tabla 43. Análisis de las preguntas 9, 10 y 11 de T3……...128
Tabla 44. Análisis de las preguntas 13d, 13e y 13f de T3……133
Tabla 45. Análisis de las preguntas 14a, 14b y 14c de T3. ……..136
Tabla 46. Características de la Tarea 3……138

1. ASPECTOS GENERALES

1.1. Introducción

Este trabajo de grado surge como resultado del proceso formativo en la Maestría en Docencia de la Matemática del Departamento de Matemáticas de la Universidad Pedagógica Nacional. Se tiene como propósito plantear una secuencia de tareas mediada por Excel y con el uso de bases de datos reales, para estudiantes de grado noveno, que busca promover la interpretación de los significados de las medidas de dispersión estadística varianza y desviación estándar.

El documento se organiza en seis capítulos que abordan los aspectos generales de la investigación, los fundamentos teóricos (marco de referencia), las metodologías empleadas (aspectos metodológicos), el análisis de los resultados obtenidos durante la implementación de las tareas (resultados), las conclusiones de la investigación y, finalmente, las referencias del documento.

El primer capítulo presenta la problemática vinculada a la inadecuada o ausente interpretación de las medidas de dispersión, identificada en estudiantes de grado noveno de la institución educativa Tandem - Apoyo en procesos escolares. Asimismo, se exponen los antecedentes relevantes que respaldan la pertinencia del trabajo, así como los objetivos que orientan la propuesta.

En el segundo capítulo se expone el marco de referencia, en el que se describen cuatro enfoques claves para la investigación. El primero se refiere a la competencia interpretativa, la cual incluye una recopilación de diversas definiciones y concepciones, y a partir de esto, una propuesta de definición de competencia y de competencia interpretativa. El segundo enfoque consiste en una conceptualización sobre las medidas de dispersión estadística.. El tercer enfoque se centra en los conceptos asociados a Microsoft Excel, bases de datos y algunos elementos de la teoría de la génesis instrumental.

El tercer capítulo detalla los enfoques metodológicos implementados en la investigación, abarcando principalmente la revisión documental y los experimentos de enseñanza. Asimismo, se describe el diseño de las tres tareas, fundamentado en la teoría de la génesis instrumental y las categorías de análisis empleadas, las cuales están enmarcadas dentro de los lineamientos teóricos definidos (marco de referencia). En el cuarto capítulo se realiza el análisis de los resultados obtenidos teniendo en cuenta las categorías de análisis correspondientes a la competencia interpretativa.

Finalmente, en el quinto capítulo, se presentan las conclusiones de la investigación, destacando: (1) Se cumplió con los objetivos de la investigación, destacándose la revisión documental que clarificó elementos teóricos sobre la interpretación de medidas de dispersión; (2) las tareas diseñadas con el apoyo de Microsoft Excel y bases de datos reales fomentaron habilidades relacionadas con la competencia interpretativa en los estudiantes; (3) aunque se observó un progreso significativo, persiste una dificultad en la transición de lo contextual a lo matemático, lo que sugiere áreas de mejora; (4) es recomendable explorar otros modelos de tareas para enriquecer las intervenciones pedagógicas; y (5) los resultados de la investigación aportaron nuevas perspectivas a las prácticas docentes, resaltando la importancia de diversificar las herramientas y estrategias de enseñanza, además de haber sido compartidos con la comunidad académica en varios eventos especializados.

1.2. Antecedentes

A continuación, se presenta una breve descripción de investigaciones que abordan de forma parcial aspectos relacionados con la interpretación de las medidas de dispersión mediante el uso de herramientas tecnológicas y que permiten evidenciar los avances realizados frente a dicha problemática. Con el fin de reportar algunos de estos hallazgos, se clasifican dichas investigaciones teniendo en cuenta dos aspectos: en primer lugar, investigaciones que involucren el estudio de la interpretación de medidas estadísticas, en particular de las medidas de dispersión y, en segundo lugar, investigaciones que aborden el uso de la tecnología y las hojas de cálculo en la enseñanza y aprendizaje de la estadística.

En relación con la interpretación de las medidas estadísticas, Batanero et al. (1994) resaltan la importancia de reconocer las dificultades y errores que se pueden observar en los estudiantes a la hora de abordar tareas que pongan en juego la aplicación de conceptos estadísticos, teniendo en cuenta que algunas dificultades están asociadas a la falta de comprensión correcta del concepto o procedimiento dado. Puntualmente, al abordar las medidas de dispersión, se hace énfasis en que el análisis de un conjunto de datos no puede limitarse a determinar su promedio, puesto que se está dejando a un lado la variabilidad, de allí la importancia de estudiar dichas medidas.

Es por ello que, con el fin de evidenciar las dificultades relacionadas al interpretar las medidas de dispersión, realizan un experimento con 154 estudiantes de un curso de psicología. Se proporciona

a dichos estudiantes dos conjuntos de datos, con el fin de que puedan determinar cuál de los dos conjuntos presentan mayor dispersión; al hacer una revisión de los resultados se encuentra que los estudiantes determinan la dispersión de los datos comparando cómo varían unos datos con respecto a los otros y no cómo varían con respecto a un número fijo (la media). Finalmente, los autores resaltan que lo anterior es el resultado de que, en la enseñanza, los conceptos se presentan aislados de las aplicaciones originales, pero cada una de estas aplicaciones aporta una parte del significado global de los mismos.

Por otra parte, Rojas y Fúneme (2021) proponen como estrategia el abordar las medidas de dispersión a través del aprendizaje basado en proyectos [ABP] con el fin de aportar al desarrollo de la cultura estadística y fomentar el uso, la interpretación y la argumentación crítica de los estudiantes en la temática mencionada. Para ello, desarrollaron un proyecto con seis estudiantes del grado undécimo de la institución educativa San Isidro ubicada en Úmbita – Boyacá. Inicialmente, se realiza un cuestionario diagnóstico en el cual debían usar las medidas de dispersión en diversas situaciones contextualizadas, allí se identifica que independientemente de los aciertos al realizar los cálculos correspondientes a las medidas estadísticas, los estudiantes no logran articular la relación entre el valor numérico y la situación contextualizada.

Posteriormente, se desarrolla el proyecto en torno a un fenómeno social cercano al contexto de los estudiantes: el alcoholismo, con el fin de establecer el nivel de cultura estadística en que se encuentran los estudiantes. De ello, se pudo determinar que el trabajo con datos reales despierta el interés de los estudiantes (en particular, porque es una situación que ocurre realmente en su diario vivir), y de esta manera resulta para ellos más sencillo el interpretar la información puesto que pueden usar sus argumentos e ideas empíricas de lo que conocen del contexto en el que se enmarca el proyecto y también las medidas de dispersión calculadas, generando ideas entre estas dos perspectivas para así establecer una posición crítica frente a la situación planteada. Sin embargo, los autores hacen hincapié en que la correcta interpretación de las medidas estadísticas en cuestión podría estar sujeta al reconocimiento de la problemática trabajada, de lo contrario en otros contextos no podría ser tan sencillo para los estudiantes argumentar las interpretaciones realizadas.

Por otro lado, el estudio de Díaz y Landázuri (2008) aborda las concepciones sobre la desviación estándar en estudiantes de educación media del Colegio San Francisco I La Casona, utilizando la metodología fenomenográfica, empleada para identificar las diversas formas en que los estudiantes

conciben, entienden o conceptualizan un mismo fenómeno o concepto, en este caso específico el de desviación estándar. La investigación utilizó textos escolares, un cuestionario y una entrevista como instrumentos de recolección de datos. De la revisión de los textos escolares se identifican dos concepciones principales: la *concepción algorítmica*, que relaciona la desviación estándar con el cálculo de fórmulas matemáticas sin profundizar en el significado de los resultados, y la *concepción de medidas de dispersión*, que aborda comparaciones numéricas entre distribuciones, pero no fomenta la toma de decisiones basadas en estos valores.

Mediante la aplicación del cuestionario y la entrevista a los estudiantes de la institución mencionada, se identifican cuatro concepciones principales: la *concepción algorítmica*, que entiende la desviación estándar como una medida de variación calculada a través de una formula algebraica; la *concepción de cambio en las frecuencias* , que vincula a la desviación estándar como una medida de variación de tamaño de las barras de un histograma; la *concepción de distancia*, que interpreta la desviación estándar como una medida de variación donde se mide la distancia entre los valores de la variable y un punto de referencia; y la *concepción de función*, que entiende la desviación estándar como una función en la que los datos son la variable independiente.

Al contrastar las concepciones de la aplicación con los enfoques presentes en textos escolares, concluyeron que algunas de las concepciones de la desviación estándar identificadas en los estudiantes, tienen características equivalentes a las encontradas en los textos. En general, el aprendizaje de los estudiantes se centra en un enfoque algorítmico. En cuanto a la concepción de dispersión, los textos muestran criterios limitados para comparar la desviación estándar en gráficos estadísticos, omitiendo elementos clave como la identificación de variables en los ejes, la escala de valores y la ubicación de la media aritmética. A pesar de esto, los estudiantes suelen tomar decisiones adecuadas en diversas situaciones, aunque en los textos estudiados no se planteen escenarios en los cuales haya que tomar decisiones con base en lo que representa la desviación estándar. Aunque un alto porcentaje de las tareas en los textos consisten en ejercicios simples que requieren aplicar la fórmula de desviación estándar, solo una pequeña proporción de los estudiantes la utilizan correctamente.

En cuanto al componente tecnológico, López et al. (2009) realizan una prueba con dos grupos de estudiantes homogéneos en términos de conocimientos estadísticos, que cursaban el espacio académico de Estadística en un programa de medicina. Es importante mencionar que el uso de una

herramienta digital no estaba considerado en el desarrollo del curso, pero la investigación precisamente apuntaba a evaluar si el uso de una tecnología digital, en este caso Excel, contribuía al aprendizaje de los estudiantes, teniendo en cuenta que al culminar el curso los alumnos deben realizar cálculos estadísticos y representar datos por medio de gráficos y tablas. Posterior a la prueba piloto, se concluye que el grupo en el cual se implementó el uso de Excel tiene avances significativos en cuanto al procesamiento de los datos, desarrollo de habilidades y adicionalmente se genera un mayor interés por parte de los alumnos al abordar las situaciones planteadas mediante el uso de tecnologías digitales.

1.3. Delimitación del problema

Teniendo en cuenta que esta investigación se inscribe en la línea de Tecnología y Educación Matemática de la Maestría en Docencia de la Matemática de la Universidad Pedagógica Nacional; se propone realizar un trabajo asociado al desarrollo de la interpretación de las medidas de dispersión estadística, varianza y desviación estándar, mediante el desarrollo de tareas mediadas por Microsoft Excel en el contexto escolar.

La elección del asunto a abordar se debe en un inicio a que, en el ejercicio docente de los autores, se ha identificado una problemática con respecto a la interpretación de las medidas de dispersión estadística por parte de los estudiantes. Aunque los alumnos suelen ser capaces de realizar cálculos algorítmicos para determinar medidas como la varianza y la desviación estándar, enfrentan dificultades al interpretar adecuadamente dichas medidas. A continuación, se describen los aspectos que permiten delimitar y establecer la problemática en cuestión.

La problemática es detectada en el colegio Tandem – Apoyo en procesos escolares*,* institución educativa por ciclos semestrales que cuenta con grupos de estudio de máximo cuatro estudiantes. Esto tiene como finalidad garantizar que sean atendidos de acuerdo con sus necesidades cognitivas y disciplinares; también permite clasificarlos teniendo en cuenta el nivel de sus habilidades en grupos altos, medios o bajos. En particular, los estudiantes de noveno grado, reciben cuatro horas semanales para Matemáticas y dos horas para Estadística.

Dentro de la malla curricular de matemáticas propuesta por la institución Tandem - Apoyo en procesos escolares, se establece abordar durante un semestre, que a su vez se divide en tres cortes

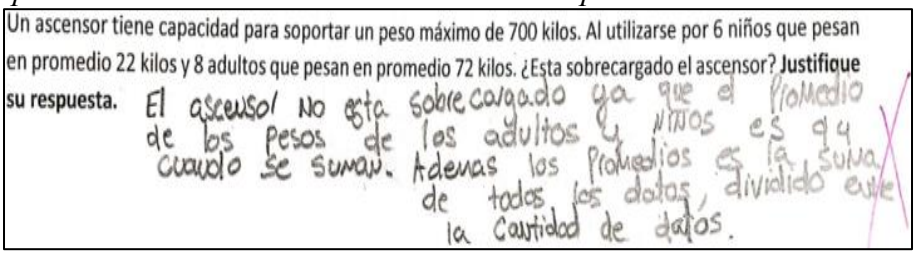
(periodos de tiempo de aproximadamente 7 semanas), las siguientes temáticas en lo que respecta a estadística:

- Primer corte: Población, muestra, variables estadísticas, gráficos y estimadores estadísticos (medidas de tendencia central, posición y dispersión), todo lo anterior para datos no agrupados (6 semanas).
- Segundo corte: Población, muestra, variables estadísticas, gráficos y medidas de tendencia central, todo lo anterior para datos agrupados (6 semanas).
- Tercer corte: Temáticas estudiadas en los dos primeros cortes empleando las herramientas de Excel (6 semanas).

Los tiempos asignados para el desarrollo de procesos y conceptos asociados a las temáticas son muy reducidos, por ello es habitual que no se profundice en algunos procesos matemáticos, como la interpretación y resolución de problemas. Por ende, cuando se logra proponer actividades que apuntan a dichos procesos, los estudiantes no logran alcanzar los objetivos de aprendizaje planteados.

Durante una tarea propuesta en relación con la interpretación de diferentes situaciones estadísticas, se evidencia que el estudiante reconoce el procedimiento para determinar la media aritmética e incluso lo describe en su explicación; sin embargo, no emplea el algoritmo de manera concreta, ni para llevar a cabo el cálculo convencional, ni mucho menos para emplearlo de otras formas como sustituir valores y resolver ecuaciones (ver Figura 1). A pesar de que el estudiante intenta establecer una relación entre las medidas presentadas en el enunciado, su interpretación, que es inadecuada, y la falta de comprensión del proceso algorítmico descrito, resulta en respuestas incorrectas.

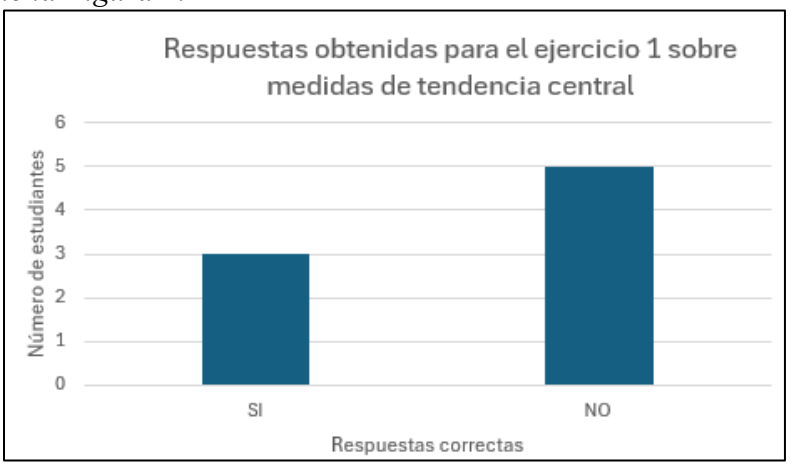
Figura 1. *Respuesta de estudiante en relación con la interpretación de la media*



El ejercicio propuesto en la Figura 1 fue desarrollado por un total de ocho estudiantes, que se distribuyen en dos grupos de noveno grado, quienes estaban cursando el espacio académico de Estadística durante el semestre 2023-I. Mas de la mitad del grupo no logró resolver correctamente

el ejercicio (ver Figura 2). En general, en la mayoría de las preguntas que involucraban interpretación de medidas de tendencia central, se repetía una inadecuada interpretación y falta de comprensión del proceso algorítmico que llevaban a respuestas erróneas.

Figura 2. *Gráfico estadístico en relación con las respuestas correctas e incorrectas obtenidas para la pregunta de la Figura 1.*



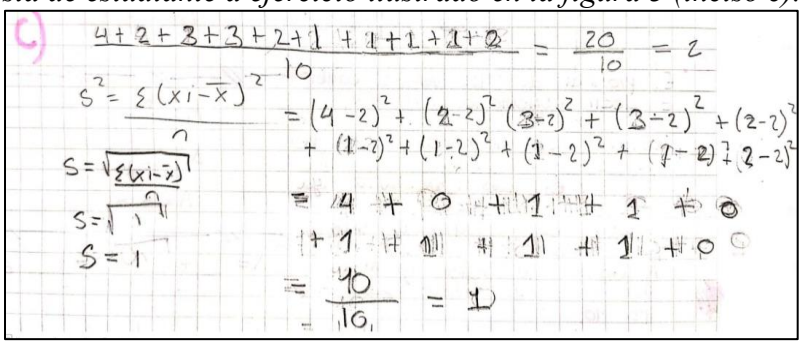
Por otro lado, en preguntas que implican la interpretación de medidas de dispersión, se observa un escenario similar, por ejemplo, ninguno de los estudiantes logra proporcionar una solución completamente correcta a la situación presentada en la Figura 3. No obstante, en la Figura 4, la cual muestra una solución particular al ejercicio planteado en la Figura 3, se puede apreciar que el estudiante reconoce y aplica correctamente el proceso algorítmico para determinar la media aritmética, la varianza y la desviación estándar, puesto que posteriormente utiliza un método de ensayo y error con el fin de encontrar valores que cumplan con las medidas estadísticas dadas. Aunque los estudiantes logran encontrar un conjunto de datos que satisface la media aritmética propuesta en el enunciado, no consiguen que el mismo conjunto cumpla con la desviación estándar indicada.

Figura 3. *Enunciado de ejercicio propuesto a estudiantes en relación con medidas estadísticas*

Para cada una de las siguientes medidas, determine un conjunto de 10 datos.

a. $\bar{x} = 5$ y $S = 3$
b. $\bar{x} = 12{,}5$ y $S = 0{,}5$
c. $\bar{x} = 2$ y $S = 0$
d. $\bar{x} = 25$ y $S = 1$
e. $\bar{x} = 13$ y $S = 10$

Figura 4. *Respuesta de estudiante a ejercicio ilustrado en la figura 3 (inciso c).*



Teniendo en cuenta los dos ejercicios presentados junto con una respuesta que logró capturar el panorama general de lo que fue el desarrollo de esta tarea, se logra evidenciar que existe un conocimiento por parte de los estudiantes en cuanto a los algoritmos de las medidas estadísticas en cuestión, sin embargo, no sucede lo mismo si se considera lo referente a la interpretación, puesto que parece ser no adecuada o inexistente. Lo anterior tiene una estrecha correspondencia con lo señalado por Rojas y Fúneme (2021), quienes destacan que la dificultad en el aprendizaje de la desviación estándar radica en que los estudiantes la perciben como el resultado de un algoritmo, sin llegar a comprender su verdadero significado. Cabe aclarar que en esta indagación no se limitó la problemática únicamente a la desviación estándar.

Por otra parte, y con el fin de establecer el foco de la investigación, es importante reconocer que desde los antecedentes se evidenció que ya existen investigaciones que han abordado la problemática descrita desde diversos ámbitos. Rojas y Fúneme (2021) abordan la interpretación de las medidas de dispersión a través de contextos reales; sin embargo, hacen hincapié en la posibilidad de que las interpretaciones reportadas por los estudiantes, puedan estar sesgadas por el contexto en el que se enmarcó la investigación, Batanero et al. (1994) estudian los posibles errores y dificultades que presentan los estudiantes al trabajar con medidas estadísticas, y afirman que lo anterior está asociado a la falta de compresión de los conceptos y procesos relativos a dichas medidas, y finalmente López et al. (2009) trabaja con conjuntos de datos a través de herramientas digitales como Microsoft Excel, dicha interacción le permite afirmar a los autores que se genera mayor interés por parte de los estudiantes para abordar situaciones estadísticas y se evidencia una mejora en las habilidades relacionadas con el procesamiento de datos.

Además, en relación con lo mencionado por Rojas y Fúneme (2021), se pone en consideración lo referente al uso de datos en contexto reales. Es por ello que autores como Makar y Confrey (2005,

citado en Estepa et al., 2013) señalan que los datos no son solo números, son números con un contexto; por lo que resulta útil apoyar el proyecto de investigación en el uso de bases de datos, pues la información que allí se proporciona corresponde a datos reales.

Otros autores como Pérez (2006), mencionan la potencialidad que se puede establecer al usar la dupla tecnología digital y bases de datos, como se muestra en la siguiente afirmación:

> En la actualidad se requiere de la utilización de herramientas potentes para facilitar el uso de los datos, agilizar la obtención de los resultados y para obtener una mejor precisión de los mismos. Una de esas potentes herramientas que pueden utilizarse en nuestro medio es el paquete de office de Windows y dentro de este al programa Microsoft Excel, el cual tiene una amplia gama de utilidades. (p. 68)

Es por lo anterior que, al recoger todos los elementos mencionados, tanto teóricos como empíricos, en la presente investigación se propone trabajar la interpretación de las medidas de dispersión a través del uso de tecnologías digitales que a su vez permita el tratamiento de datos reales, lo cual apunta a integrar todos los elementos con los cuales otros investigadores ya han experimentado. La finalidad de este trabajo es determinar la respuesta a la siguiente pregunta de investigación:

¿De qué forma se promueve la interpretación de la varianza y la desviación estándar a través de tareas mediadas por Microsoft Excel y el uso de bases de datos reales en estudiantes de grado noveno?

Es importante especificar que para abordar dicha pregunta se propone el diseño e implementación de una secuencia de tareas que empleen la herramienta Microsoft Excel y el uso de bases de datos reales. Las tareas estarán dirigidas a los estudiantes de noveno grado de la institución educativa Tandem – Apoyo en procesos escolares y se enfocarán en la interpretación de la varianza y la desviación estándar.

1.4. Objetivos

A partir de la problemática delineada en la sección anterior y de la pregunta de investigación derivada de esta, se plantean los siguientes objetivos con el propósito de dar respuesta a dichos cuestionamientos.

1.4.1. Objetivo general

Promover la interpretación de las medidas de dispersión, varianza y desviación estándar, a través de una secuencia de tareas mediada por Excel y con el uso de bases de datos reales para estudiantes de grado noveno.

1.4.2. Objetivos específicos

1. Examinar, a través de una revisión documental, diferentes enfoques y concepciones sobre el desarrollo de la competencia interpretativa en la escuela, reconociendo su papel dentro del aprendizaje de las medidas de dispersión estadística.
2. Diseñar e implementar una secuencia de tareas que promuevan la interpretación de la varianza y desviación estándar, involucrando Microsoft Excel y bases de datos reales.
3. Evaluar el impacto de las tareas en la interpretación de la varianza y la desviación estándar por parte de los estudiantes.

2. MARCO DE REFERENCIA

Es importante destacar que, desde las disposiciones educativas nacionales, y en particular en los Estándares Básicos de Competencias Matemáticas [EBCM] (Ministerio de Educación Nacional [MEN], 2006) se señala que:

> Los métodos estadísticos desarrollados durante los siglos XIX y XX se han refinado y potenciado en los últimos decenios con los avances de la computación electrónica y, por ello, hoy día ya no es tan importante para los estudiantes el recuerdo de las fórmulas y la habilidad para calcular sus valores, como sí lo es el desarrollo del pensamiento aleatorio, que les permitirá interpretar, analizar y utilizar los resultados que se publiquen en periódicos y revistas. (p. 65)

Lo mencionado anteriormente, junto con lo descrito en el capítulo 1, permite establecer los aspectos que se desarrollarán con detalle durante el siguiente capítulo: 1) el papel que juegan los procesos matemáticos como la interpretación en el pensamiento aleatorio, 2) la relevancia de los métodos estadísticos, 3) el uso de las tecnologías digitales en el aula de clases y 4) la importancia de usar contextos reales en el desarrollo de conceptos y procesos estadísticos.

2.1. Competencia interpretativa

Uno de los propósitos de la presente investigación es estudiar acerca de la interpretación de las medidas de dispersión, en específico, la varianza y la desviación estándar en el contexto escolar. En este sentido, resulta esencial sumergirse en aspectos vinculados a la competencia interpretativa, por consiguiente, se describirá el proceso seguido para establecer conexiones entre conceptos fundamentales como competencia, competencia matemática y competencia interpretativa. Es relevante resaltar que se asume la existencia de una relación de contenencia entre estos conceptos, donde la competencia interpretativa está inmersa en las competencias matemáticas (aunque no necesariamente de manera exclusiva, en tanto se reconoce que la interpretación subyace a muchas otras áreas de la escolaridad), las cuales a su vez forman parte de las competencias educativas en general.

Es importante partir de que las matemáticas forman parte de la vida real de los niños como instrumento que les permite desarrollarse mejor en su entorno, así como lo señalan Cardoso y

Cerecedo (2008), quienes afirman que es importante que desde la infancia se desarrolle el pensamiento lógico matemático basado en la construcción de un conjunto de competencias que le posibiliten utilizarlas en cualquier situación que se le presente, ya sea escolar o no.

Adicionalmente, se debe tener presente que la legislación educativa actual de múltiples países apoya un currículo o disposiciones curriculares organizados por competencias (Pastells, 2004); es por lo anterior que, inicialmente se aborda el concepto de competencia, puesto que se trata de un término ampliamente utilizado en el contexto educativo y su caracterización depende en gran medida de la época y del autor que la define, es por ello que la idea de competencia cuenta con múltiples definiciones y clasificaciones de acuerdo con las intenciones que se tenga al utilizarla.

2.1.1. Competencia

Con el propósito de indagar sobre este concepto, se llevó a cabo una revisión documental enfocada en analizar las definiciones propuestas por diversos autores. Aunque esta revisión no fue exhaustiva, la recopilación y análisis de alrededor 12 fuentes permitieron identificar definiciones en distintos contextos, como el ámbito laboral, el desempeño humano en actividades cotidianas y el ámbito educativo. Fue así como, durante la exploración de los Estándares Básicos de Competencias en Matemáticas [EBCM] (MEN, 2006) y los Lineamientos Curriculares de Matemáticas [LCM] (MEN, 1998), se observó que la definición de competencia no está explícita en estos documentos, y un análisis más profundo de diversas fuentes reveló cambios en los conceptos presentados en estos referentes de calidad que tienen más de 20 años desde su concepción, lo cual sugiere la pertinencia de profundizar en la comprensión del concepto de competencia.

Por consiguiente, se realiza una recopilación de algunas definiciones, entre las cuales se tiene en cuenta lo indagado por Contreras (2011), con el propósito de identificar los elementos que se consideran relevantes y transversales al contexto educativo. A continuación, en la Tabla 1, se presentan los conceptos encontrados:

Tabla 1. *Definiciones de competencia*

Autor(es)	Definición
Graham. et al., (1991)	Las capacidades para transferir destrezas y conocimientos a nuevas situaciones dentro de un área de ocupación laboral; abarca la organización y la planificación del trabajo, la innovación y la capacidad para abordar actividades no rutinarias; incluye las cualidades de eficacia

	personal que se necesitan en el puesto de trabajo para relacionarse con los compañeros, los directivos y los clientes.
Perrenoud (2000)	Una capacidad de movilizar diversos recursos cognitivos para enfrentar un tipo de situaciones.
Le Boterf (2001)	Una estructura basada en recursos personales (conocimientos, habilidades, cualidades o aptitudes) y recursos ambientales (relaciones, documentos, información) que se movilizan para lograr un desempeño.
Irigoin y Vargas (2002)	Competencia es transformar el conocimiento en acción. Otra definición que propondríamos: "competencia es la combinación integrada de un saber, un saber hacer, un saber ser y un saber ser con los demás; que se ponen en acción para un desempeño adecuado en un contexto dado".
OCDE (2005)	La capacidad de los estudiantes para extrapolar lo que han aprendido y aplicar sus conocimientos ante nuevas circunstancias, su relevancia para el aprendizaje a lo largo de la vida y su regularidad.
ICFES (2023)	Las habilidades necesarias para aplicar los conocimientos en diferentes contextos, lo cual significa que, no implica solamente conocer conceptos o datos, sino que involucra saber cómo emplear dichos conceptos para resolver problemas en situaciones de la vida cotidiana.
Tardif (2008)	Un saber actuar complejo que se apoya en la movilización y combinación eficaces de una variedad de recursos internos y externos al interior de una familia de situaciones.
Gallart y Jacinto (s.f)	Conjunto de saberes puestos en juego por los trabajadores para resolver situaciones concretas de trabajo, configurando una disociación entre competencia y acción, con una exigencia de determinado conocimiento que oriente esta acción.
Tejada (1999)	El conjunto de saberes (saber, saber hacer, saber estar y saber ser conocimientos, procedimientos y actitudes) combinados, coordinados e integrados en el ejercicio profesional.
Spencer y Spencer (s.f)	Característica subyacente en el individuo que está causalmente relacionada con un estándar de efectividad y/o a una performance superior en un trabajo o situación.
Pinto (s.f)	La capacidad para actuar con eficiencia, eficacia y satisfacción sobre algún aspecto de la realidad personal, social, natural o simbólica.
Rychen y Hersh (s.f)	La capacidad de responder a demandas complejas y llevar a cabo tareas diversas de forma adecuada. Supone una combinación de habilidades prácticas, conocimientos, motivación, valores éticos, actitudes, emociones y otros componentes sociales y de comportamiento que se movilizan conjuntamente para lograr una acción eficaz.
La Oficina Internacional del Trabajo (s.f)	Capacidad de articular y movilizar condiciones intelectuales y emocionales en términos de conocimientos, habilidades, actitudes y prácticas, necesarias para el desempeño de una determinada función o actividad, de manera eficiente, eficaz y creativa, conforme a la naturaleza del trabajo. Capacidad productiva de un individuo que se define y mide en términos de desempeño real y demostrando en determinado contexto de trabajo y que no resulta sólo de la instrucción,

	sino que, de la experiencia en situaciones concretas de ejercicio ocupacional.
Valdiri y Rincón (2016).	La capacidad compleja que integra conocimientos, potencialidades, habilidades, destrezas, prácticas y acciones que se manifiestan en el desempeño en situaciones concretas, en contextos específicos (saber hacer en forma pertinente). Las competencias se construyen, se desarrollan y evolucionan permanentemente. (ICFES, 2015)

Nota. La tabla fue tomada de Contreras (2011). Se añaden definiciones tomadas de otras fuentes.

Posterior a la revisión documental presentada, se establece la posibilidad de construir una definición que integre la mayor parte de los elementos necesarios para su posterior uso en la comprensión del concepto de competencia dentro del contexto de las matemáticas. Teniendo en cuenta la naturaleza de la investigación que se va a desarrollar, se propone la siguiente definición para el concepto de competencia: *Capacidad que integra y moviliza saberes, conocimientos, potencialidades, habilidades, destrezas, prácticas, acciones y actitudes para que se manifiesten en el abordaje de actividades nuevas o no rutinarias y en el desempeño en situaciones concretas asociadas a los contextos personal, social, natural o simbólico del individuo. Las competencias se construyen, se desarrollan y evolucionan permanentemente*.

Para clarificar la definición propuesta, es crucial precisar algunos términos para evitar ambigüedades en su uso, en este sentido. En este sentido, Tejada (1999) emplea el término “saberes” en plural para referirse al conjunto que abarca el saber (conocimiento), saber-hacer (procedimiento) y, saber estar y saber ser (actitud).

Por otro lado, la definición propuesta incluye diversos contextos (personal, social, natural y simbólico), los cuales fueron tomados de la definición de Pinto (s.f. citado en Contreras, 2011) presentada en la Tabla 1. Para comprender a qué hacía referencia este autor con "realidades" (lo que acá se entiende como contextos) se consultaron diversas fuentes, incluyendo Contreras (2011) y Pinto (1996, citados en Sánchez y Aguilar, 2009) quienes citan al autor de dicha definición, sin embargo, en ninguno de estos documentos se encontró evidencia explícita que aclarara un significado específico para estos contextos o realidades. Por lo tanto, estos contextos se interpretan de la siguiente manera: el primero refiere a las experiencias, creencias y circunstancias individuales que moldean la perspectiva y la interacción de una persona (contexto personal); el segundo representa el entorno colectivo, incluyendo normas, valores, estructuras y dinámicas

interpersonales (contexto social); el tercero, aquel que comprende el entorno físico, biológico y geográfico que rodea a los seres (contexto natural) y el último, aquel que alude a un entorno abstracto que en particular puede ser las matemáticas (contexto simbólico).

Con esta definición de competencia, se reconoce la existencia de diversos tipos de competencias relacionadas con diferentes áreas del conocimiento. En este caso, se ha llevado a cabo una dinámica similar a la realizada para revisar las definiciones de competencia, aunque de manera menos exhaustiva. Así, se revisaron diferentes documentos para obtener información específica sobre lo que son las competencias matemáticas.

2.1.2. Competencias matemáticas

Según Cardoso y Cerecedo (2008) una competencia matemática se vincula con el ser capaz de hacer, relacionado con el cuándo, cómo y por qué utilizar determinado conocimiento como una herramienta. Por su parte, Alsina (2023) menciona que los procesos y las competencias matemáticas enfatizan una misma idea: la capacidad de usar de forma comprensiva y eficaz las matemáticas que se aprenden en la escuela en una variedad de contextos. Finalmente, en los EBCM (MEN, 2006) se menciona que las competencias matemáticas utilizan los conceptos, proposiciones, sistemas y estructuras matemáticas como herramientas eficaces mediante las cuales se lleva a la práctica determinados tipos de pensamiento lógico y matemático dentro y fuera de la institución educativa.

Si bien no todos los organismos y autores coinciden en la concepción de competencia matemática, sí parece que todos consideran de un modo u otro que es imprescindible que, como menciona Alsina (2023), los alumnos aprendan a:

- Plantearse y resolver problemas: comprender el enunciado, generar preguntas, planificar y desarrollar estrategias de resolución y verificar la validez de las soluciones;
- Razonar y probar: realizar inducciones o deducciones, particularizar y generalizar; argumentar las decisiones, los procesos seguidos y las técnicas usadas;
- Comunicar: expresar oralmente y por escrito las ideas matemáticas, usando lenguaje matemático;
- Conectar: vincular saberes dentro y fuera de las matemáticas;
- Representar: usar materiales, palabras, dibujos, signos y símbolos;

- Crear modelos: analizar, explicar y comprender la realidad, a partir de un proceso de reflexión entre los contextos reales y las matemáticas.

Además, desde una perspectiva más contemporánea, Alsina (2023) señala que los currículos de matemáticas han comenzado a incorporar una competencia que no se contempló en las caracterizaciones de inicios del siglo XXI, la cual corresponde a:

- Pensar computacionalmente: resolver problemas aplicando algunos componentes fundamentales de la informática como los patrones, los algoritmos, la descomposición o la abstracción.

Debido a las similitudes halladas entre las concepciones de competencias matemáticas de distintos autores, no se ahondó de más en dicho concepto, dado que se percibe que la definición previamente propuesta en los EBCM (MEN, 2006) reúne la mayoría de las características. Además, este proyecto no tiene como prioridad explorar las competencias matemáticas en su conjunto, ya que se ha enfocado de manera específica en la competencia interpretativa.

Teniendo en cuenta lo anterior, es importante reconocer que, dentro de las competencias matemáticas que establece el MEN (2006) se consideran la argumentación, la representación, la comunicación, entre otras; pero para este estudio la competencia interpretativa es de especial interés. Con el fin de definirla, se llevó a cabo una revisión documental en la cual se realizó una exploración inicial, en la que se establecieron criterios de búsqueda para los documentos, en particular, palabras clave como "interpretar", "interpretación", "interpretación matemática" y "competencia interpretativa". Los resultados de esta búsqueda, combinados con las definiciones previamente consideradas sobre competencia y competencia matemática, permitieron comprender y evidenciar que la competencia interpretativa va más allá de la mera acción de interpretar[1]. Con el fin de detallar un poco mejor esta revisión, a continuación, se presentan algunas reseñas de diversos documentos en las cuales se establecen definiciones o características de la competencia interpretativa.

[1] Entendido interpretar de la manera más sencilla como «explicar acciones, dichos o sucesos que pueden ser entendidos de diferentes modos» según la RAE.

2.1.3. Competencia interpretativa

Algunos autores por medio de artículos o investigaciones dentro y fuera del campo de las matemáticas, logran establecer una noción de la competencia interpretativa. Un ejemplo de ello es la propuesta de Batanero y Godino (2004, citados en Moreno y Villanueva, 2013), quienes describen la competencia interpretativa como "la capacidad del estudiante para leer los datos, dentro de los datos, más allá de los datos y detrás de los datos y así logra identificar, analizar, comprender e interpretar una información matemática y lingüística a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural" (p.14). Adicionalmente, Moreno y Villanueva (2013) puntualizan que, particularmente, en el campo de la Estadística entienden dicha competencia como "la capacidad que desarrolla el estudiante para leer los datos, es decir, escudriñar los datos para generar las conclusiones que sean posibles sobre la situación que se esté desarrollando; logrando identificar tendencias, conjeturas, construir hipótesis, argumentar y tomar posición crítica frente a una determinada información" (p.8).

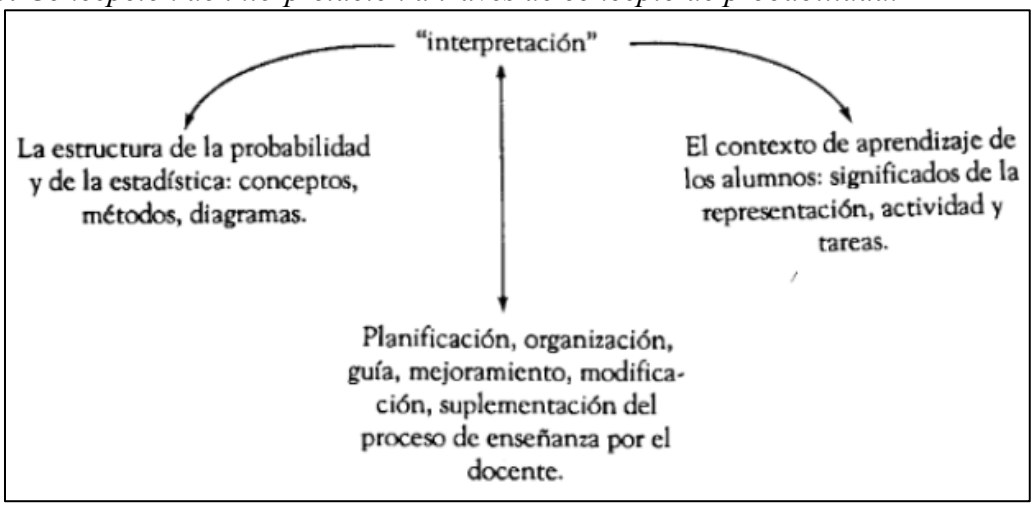
Por otro lado, la definición de Moreno y Montañez (2015) propone que en general la competencia interpretativa se define como "la capacidad de comprender textos, gráficas, esquemas, mapas, ambientes, medios informáticos, comunicativos, y situaciones problemáticas" (p. 25). Sin embargo, al situarse en el área de matemáticas, los autores afirman que se entiende como "la capacidad de leer datos, entender los datos, comprender datos y explicar datos, estos datos pueden presentarse de diferentes formas" (p. 25). Otra definición que aborda algunas características distintas a las mencionadas hasta el momento es la de Contreras (2018), que describe la competencia interpretativa como la capacidad de lograr acceder a la información que se desea transmitir, a través de elementos matemáticos, por lo tanto, exige al estudiante observar y discernir la información presentada para extraer lo que se quiere expresar.

Si bien es cierto que en las definiciones presentadas se logró identificar características similares entre lo propuesto por los autores para caracterizar la competencia interpretativa, se quiso profundizar en el concepto en cuestión, revisando documentos curriculares nacionales e internacionales, teniendo en cuenta que la finalidad de este ejercicio es tomar los elementos suficientes y necesarios para así establecer una definición propia.

En los LCM (MEN, 1998), acerca de la interpretación, se cita un modelo para la enseñanza de la probabilidad propuesto por Steinbring (1989) quien en su artículo denominado "La interacción

entre la práctica de la enseñanza y las concepciones teóricas", presenta un modelo basado en un análisis epistemológico de la naturaleza de la probabilidad, el cual considera tres niveles. El primero tiene que ver con la estructura del contenido, el segundo tiene en cuenta el estudiante que aprende significativamente y el tercero considera al docente quien planifica, organiza, apoya y desarrolla esta forma de aprendizaje. La Figura 5 muestra cómo se interrelacionan estos tres niveles mencionados anteriormente:

Figura 5. *Concepción de interpretación a través de concepto de probabilidad.*



Nota. Ilustración tomada de Steinbring (1989). Modelo de niveles en un análisis epistemológico de la naturaleza de la probabilidad.

Es importante precisar que en dicho documento no se encuentra una concepción puntual de interpretación ni mucho menos de competencia interpretativa, sino que se limita a presentar un caso particular de interpretación dentro del campo de la estadística.

En el caso de los EBCM (MEN,2006) se evidencia una particularidad, pues en el documento no se especifica una definición de competencia interpretativa; sin embargo, muchos de los procesos matemáticos allí planteados sí la mencionan de una u otra manera, de hecho, parece que la competencia interpretativa es transversal a los cinco procesos matemáticos descritos en el documento. Resulta curioso que 18 de los 172 estándares, que corresponde a más del 10%, contemplan la interpretación de manera explícita sin antes haber establecido una definición de dicho concepto.

Por otra parte, el *Programme for International Student Assessment* [PISA] (ICFES, 2020) evalúa tres procesos en relación con las competencias matemáticas, en este caso, únicamente es de interés al que denominan "Interpretar, aplicar y evaluar los resultados matemáticos". Esta competencia indica que la interpretación (y evaluación) se refiere a la eficacia con la que los estudiantes pueden reflexionar sobre soluciones o conclusiones matemáticas, interpretarlas en el contexto del problema en el mundo real y determinar si los resultados o conclusiones son razonables o útiles. Se puede requerir que las personas involucradas en este proceso construyan y comuniquen explicaciones y argumentos en el contexto del problema, reflexionando tanto sobre el proceso de modelado como sobre sus resultados.

Adicionalmente se especifican algunas de las actividades que se pueden realizar con el fin de desarrollar dicha competencia:

- Interpretar información presentada en forma gráfica o diagramas
- Evaluar un resultado matemático en términos del contexto
- Interpretar un resultado matemático en el contexto del mundo real
- Evaluar la razonabilidad de una solución matemática en el contexto de un problema del mundo real.
- Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados.
- Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.
- Comprender el alcance y los límites de los conceptos y soluciones matemáticas
- Criticar e identificar los límites del modelo utilizado para resolver un problema
- Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas.

En uno de los documentos curriculares internacionales como son los estándares elaborados por el National Council of Teachers of Mathematics [NCTM] (2000), denominados *Principles and standards for school Mathematics*, no se establece una definición puntual sobre la competencia interpretativa o la interpretación, es importante destacar que los procesos matemáticos que se pueden desarrollar están sujetos a las representaciones que el estudiante realiza al interactuar con

un objeto matemático; por lo cual, se puede obtener información valiosa sobre las formas en que los estudiantes interpretan y piensan las matemáticas al observar sus representaciones.

Por su parte, en *Guidelines for Assessment and Instruction in Statistics Education* [GAISE] (Franklin et al., 2020) se proponen 4 componentes para el proceso de resolución de problemas: Formular preguntas de investigación estadística, recopilar datos/considerar datos, analizar los datos e interpretar los resultados. Este último componente es el foco de interés en este caso. A continuación, se presenta un ejemplo de interpretación de resultados asociados a la variabilidad:

> Las interpretaciones estadísticas se realizan en presencia de variabilidad y deben tener en cuenta la variabilidad. Por ejemplo, debemos interpretar el resultado de una encuesta electoral como una estimación que puede variar de la muestra a la muestra de los votantes encuestados. Al interpretar los resultados de un experimento médico comparativo aleatorio, debemos recordar que hay dos fuentes importantes de variabilidad: la aleatorización al grupo de tratamiento y la variabilidad de un individuo a otro. Cuando generalizamos los resultados y miramos más allá de los datos del estudio recopilados, debemos tener en cuenta estas fuentes de variabilidad. (p. 23)

Aunque el documento no ofrece una definición formal del componente de interpretación de resultados, presenta una tabla que detalla ciertos elementos esenciales que se espera que los estudiantes alcancen al desarrollar este componente. Es importante mencionar que cada componente comprende tres niveles que dan cuenta y razón de la evolución en la interpretación a partir de la resolución de problemas, donde el trabajo en el Nivel B adopta y desarrolla aún más los conceptos del Nivel A; de manera similar, el trabajo en el Nivel C adopta y utiliza conceptos de los niveles inferiores. A continuación, se presenta la Tabla 2 correspondiente al componente de interpretación de resultados:

Tabla 2. *Niveles del componente interpretación de resultados de la GAISE*

	Niveles del componente analizar los datos e interpretar los resultados
Nivel A	• Utilice evidencia estadística de los análisis para responder las preguntas de investigación estadística y comunicar los resultados a través de respuestas estructuradas con la orientación del maestro. • Describir la diferencia entre dos grupos con diferentes condiciones. • Hacer declaraciones sobre el grupo o población de la cual se recopilaron los datos, reconociendo que las conclusiones se limitan a estos grupos y no se pueden generalizar a otros grupos.

Nivel B	• Utilice evidencia estadística de los análisis para responder las preguntas de investigación estadística y comunicar los resultados con respuestas integrales y cierta orientación del maestro. • Reconocer que mirar más allá de los datos es factible. • Reconocer la incertidumbre causada por la variabilidad de una muestra a otra. • Comparar resultados para diferentes condiciones en un experimento. • Generalizar más allá de la muestra proporcionando evidencia estadística para la generalización e incluyendo una declaración de incertidumbre y plausibilidad cuando sea necesario. • Establecer las limitaciones de la información de la muestra (por ejemplo, una muestra puede o no ser representativa de la población más grande, variabilidad de las mediciones).
Nivel C	• Utilice evidencia estadística de los análisis para responder las preguntas de investigación estadística y comunicar los resultados a través de informes y presentaciones más formales. • Evaluar e interpretar el impacto de los valores atípicos en los resultados. • Comprender lo que significa que un resultado o una estimación de una característica de una población sea plausible o no plausible en comparación con la variación aleatoria. • Interpretar el margen de error asociado con una estimación de una característica de la población. • Reconozca la presencia de valores faltantes y comprenda cómo los valores faltantes pueden agregarse sesgo a un análisis. • Utilice el pensamiento multivariado para comprender cómo las variables se impactan entre sí. • Comunicar el razonamiento estadístico y los resultados a otros en una variedad de formatos (verbal, escrito, visual). • Comprender cómo interpretar adecuadamente los valores simulados.

Nota. Información tomada de Franklin et al. (2020). Tabla de elementos correspondientes al componente de interpretación de resultados por niveles.

Finalmente, el Instituto Colombiano para la Evaluación de la Educación [ICFES] (2023) evalúa tres competencias especificas en el área de matemáticas: Interpretación y representación, argumentación, formulación y ejecución. En este caso, es de interés al que denominan interpretación y representación, el cual es definido como

> La habilidad para comprender y transformar la información presentada en distintos formatos como tablas, gráficas, conjuntos de datos, diagramas, esquemas, etc., así como la capacidad de utilizar estas representaciones para extraer información relevante que permita, entre otras, establecer relaciones matemáticas e identificar tendencias y patrones.

> Con el desarrollo de esta competencia se espera que un estudiante utilice coherentemente registros como el simbólico, el natural, el gráfico y todos aquellos que se dan en situaciones que involucran las matemáticas. (p. 17)

Al analizar las diversas definiciones consultadas, se observó que ninguna integraba de manera concreta todas las características transversales comunes a la mayoría de las definiciones revisadas. Esto se debió, en parte, a que algunas de las definiciones eran de hace varios años o únicamente abordaban aspectos relacionados con el término "interpretar" y no como competencia. En consecuencia, se decidió desarrollar una definición propia que reflejara de manera más precisa y actualizada las perspectivas del estudio sobre la competencia interpretativa.

Para lograr esto, se construyó una matriz que incluía todas las definiciones recolectadas, con el objetivo de comparar y distinguir las principales características presentes en cada una. Se identificaron similitudes, a pesar de las variaciones en el lenguaje y la forma de expresión, y se destacaron características únicas de algunas definiciones que aportaron claridad a los elementos considerados relevantes. A partir de este análisis, se seleccionó la definición del ICFES (2023), ya que esta incluía la mayor cantidad de características relevantes y posteriormente, se realizaron ajustes y modificaciones a dicha definición para asegurar que incorporara todas las características principales identificadas durante el análisis, logrando así una definición más completa y adecuada para el contexto del estudio.

Fue así como se estableció la siguiente definición: Esta competencia consiste en la habilidad para comprender la información presentada en distintos formatos como ecuaciones, formulas, textos estadísticos, tablas, gráficas, conjuntos de datos, diagramas, esquemas, y poderla transformar en distintos registros de representación, así como la capacidad de utilizar estas representaciones para extraer información relevante que permita, entre otras, reflexionar sobre soluciones o conclusiones matemáticas en el contexto del problema, determinar si los resultados o conclusiones son razonables o útiles, establecer relaciones matemáticas e identificar tendencias y patrones.

Adicionalmente, se listan y codifican las características destacables correspondientes a la competencia interpretativa que se lograron identificar al realizar la revisión documental. Es importante destacar que estas no se especifican exclusivamente para las medidas de dispersión, ya que la revisión abarcó de manera general las matemáticas y la estadística. Sin embargo, dichas características (y su codificación) permitirán, posteriormente, configurar las categorías de análisis

del trabajo y, en consecuencia, evaluar la implementación de las tareas relacionadas con las medidas de dispersión.

C1. Observar, discernir y extraer información a través de elementos matemáticos. (Contreras, 2018)

C2. Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos. (Moreno y Villanueva, 2013)

C3. Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013)

C4. Identificar tendencias, conjeturas, construir hipótesis, argumentar y tomar posición crítica frente a una determinada información. (Moreno y Villanueva, 2013)

C5. Interpretar información presentada en forma gráfica o diagramas.

C6. Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.

C7. Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados.

C8. Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.

C9. Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado.

C10. Criticar e identificar los límites del modelo utilizado para resolver un problema.

C11. Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas.

C12. Establecer las limitaciones de la información de la muestra (por ejemplo, una muestra puede o no ser representativa, variabilidad de las mediciones). (Franklin et al., 2020)

C13. Comparar resultados para diferentes condiciones en un experimento. (Franklin et al., 2020)

C14. Evaluar e interpretar el impacto de los valores atípicos en los resultados. (Franklin et al., 2020)

C15. Reconozca la presencia de valores faltantes y comprenda cómo los valores faltantes pueden agregarse sesgo a un análisis. (Franklin et al., 2020)

2.2. Medidas de dispersión

En esta sección se abordan los principales conceptos de las medidas de dispersión estadística (y algunos otros que, si bien, no corresponden a las medidas de dispersión, son necesarios para darle sentido a estas), proporcionando una base matemática sólida y comprensible para el desarrollo del trabajo. Aunque no sea el objetivo principal profundizar en todos los conceptos estadísticos, se pretende ofrecer una introducción clara que permita al lector familiarizarse con los elementos clave y comprender su importancia en el análisis e interpretación de las medidas de dispersión.

Según Manzano y Sánchez (2005), la definición de Estadística es una de las cuestiones claves en la formación de los propios estadísticos y de los investigadores en general. Sin embargo, resulta sorprendente que en los estudios sobre Didáctica de esta materia se hacen pocas referencias a su definición. En los manuales y en los diccionarios el concepto de Estadística suele resolverse con alguna cita normalmente descriptiva sobre el objeto y los métodos que la ocupan, realizándose una clasificación en distintas categorías en función de los instrumentos utilizados, pero no suele hacerse referencia a una definición previa, propiamente dicha, que establezca los límites de la materia de una forma clara y que pueda ser utilizada para mejorar la comprensión sobre la misma. Por ende, a continuación, se presenta una corta recopilación de algunas definiciones de Estadística:

Tabla 3. *Definiciones de estadística*

Autor (es)	Definición
Olivera (2013)	Es el conjunto de métodos o técnicas que estudia y analiza los datos que son susceptibles de expresión numérica para llegar a conclusiones que permitan tomar decisiones y pronosticar las consecuencias de estas.
Contreras y Molina (2019)	Es un conjunto de herramientas que permiten ofrecer argumentos sólidos, basados en la evidencia, para evaluar críticamente las afirmaciones basadas en datos.
Pérez (2010)	La estadística estudia un conjunto de métodos (métodos estadísticos) en relación con la obtención y tratamiento de la información. Los fines que persigue, principalmente son: la toma de decisiones, la contrastación de hipótesis, la estimación de parámetros y la predicción de resultados.
Barreto (2012)	Es la ciencia cuyo objetivo es reunir información cuantitativa concerniente a individuos, grupos, series de hechos, etc., y deducir de ello gracias al análisis de estos datos unos significados precisos o unas previsiones para el futuro.

	La estadística, en general, es la ciencia que trata de la recopilación, organización, presentación, análisis e interpretación de datos numéricos con el fin de realizar una toma de decisiones más efectiva.
Benitez y Arrondo (2014)	Ciencia que se ocupa del estudio de fenómenos de tipo genérico, normalmente complejos y enmarcados en un universo variable, mediante el empleo de modelos de reducción de la información y de análisis de validación de los resultados en términos de representatividad
Ross (2018)	Es el arte de aprender a partir de los datos. Está relacionada con la recopilación de datos, su descripción subsiguiente y su análisis, lo que nos lleva a extraer conclusiones.

Como se puede evidenciar, no hay una uniformidad evidente entre las definiciones, en cambio, se pueden identificar ciertas características comunes en algunas de estas, como su naturaleza cuantitativa, su aplicación en la predicción, la conclusión o la toma de decisiones, entre otras.

Es curioso notar que una misma definición puede tener diversas concepciones sobre su naturaleza, lo cual revela su complejidad epistemológica. En los casos mencionados, la estadística se concibe como un conjunto de métodos, un conjunto de herramientas, un arte o una ciencia. Por lo tanto, a continuación, se presenta una definición que busca proporcionar una visión clara y precisa del concepto de Estadística, destacando sus aspectos más relevantes y especificando su naturaleza. De esta forma, la estadística, en el marco de este trabajo, se define como:

La ciencia que se encarga de la recopilación, organización, presentación, análisis e interpretación de los datos de un fenómeno, empleando métodos, técnicas y herramientas para llegar a conclusiones que faciliten la toma de decisiones, contrastación de hipótesis, estimación de parámetros y la predicción de resultados.

Según Olivera (2013) la Estadística se presta a diversas formas de clasificación, por ejemplo, la Estadística paramétrica y no paramétrica, la Estadística univariada y multivariada, la Estadística experimental y observacional, entre otras. No obstante, en este trabajo se centra en la clasificación propuesta por Sabadías (1995) entre la Estadística descriptiva e inferencial: la primera tiene como objetivo resumir y describir numéricamente un conjunto de datos con el fin de facilitar su interpretación, mientras que la segunda permite realizar inferencias acerca de las características de los individuos de la población a partir de las características de una muestra de esta. En particular, debido al enfoque del trabajo, las tareas que se proponen se encuentran dentro del campo de la Estadística descriptiva.

En el contexto de la estadística es esencial comprender la distinción entre población y muestra. Según Canavos (1988), la **población** es la colección de toda la posible información que caracteriza un fenómeno y se representa por Ω. Es decir, una población es cualquier colección ya sea de un numero finito de mediciones o una colección grande, virtualmente infinita[2], de mediciones u observaciones. Por su parte, una **muestra** es un subconjunto de una población; y según Marín, (1987, citado en Argibay, 2015) se habla de una **muestra representativa** cuando "los estadísticos que se obtienen en la muestra corresponden y representan, sin error sistemático, a los parámetros de la población".

Según Otzen y Manterola (2017) una muestra puede ser obtenida de dos tipos: probabilística y no probabilística. Las **técnicas de muestreo probabilísticas** permiten conocer la probabilidad de que un individuo sea incluido en la muestra mediante una selección al azar. En cambio, en las **técnicas de muestreo de tipo no probabilísticas**, la selección de los sujetos a estudiar dependerá de ciertas características, criterios, etc., que los investigadores consideren relevantes en ese momento; por lo que pueden ser poco válidos y confiables o reproducibles, debido a que este tipo de muestras no se ajustan a un fundamento probabilístico.

Desde una perspectiva más teórica, Heumann y Schomaker (2016) señalan que las unidades con las que se miden los datos (personas, automóviles, animales, plantas) se denominan observaciones, y cada observación se representa con $\omega_i (i \in \{1, \dots, n\})$. Si se considera una selección de observaciones $\omega_1, \omega_2, \omega_3, \dots, \omega_n$ entonces estas observaciones se llaman **muestra**, siendo n la cantidad de datos de la muestra. Cuando se hace referencia a que $\omega \in \Omega$ quiere decir que se elige una sola observación entre todo el conjunto de las observaciones, por ejemplo, si la población fuera un conjunto de personas, entonces se elige una persona entre todas las personas de interés.

Por lo tanto, si se especifica la población de interés para una pregunta de investigación específica, se puede pensar en aquello que es de interés acerca de las observaciones. Una característica concreta de estas observaciones se denomina una **variable estadística**. Cualquier información que sea de interés puede representarse en una variable estadística; por ejemplo, si las observaciones se refieren a seres humanos, la variable X puede describir una característica de la población, como el

[2] En algunas situaciones, como en estudios de microbiología o medicina, se asume que una población es infinita como una simplificación para facilitar cálculos matemáticos; aunque la población en la vida real puede ser muy grande, no necesariamente es infinita, pero considerarla como tal puede simplificar los análisis.

estado civil, la identidad sexual, la edad o cualquier otro aspecto que pueda relacionarse con una persona. Características de interés pueden ser diferentes, cada una de ellas se representa en una variable diferente X_i , $i = 1, 2, \ldots, p$, donde p es la cantidad de variables estadísticas que se puede establecer a un grupo de observaciones. Por ejemplo, cada observación ω_i toma un valor particular para X. Si X se refiere al estado civil, cada observación, es decir, cada persona, tiene un valor particular x_i que se refiere, por ejemplo, a "soltero" o "casado".

Asimismo, es relevante señalar que las variables estadísticas se pueden clasificar según su naturaleza o uso. Entre las muchas clasificaciones, a continuación, seis tipos mencionados por Ávila (1996), a saber:

1. **Según la naturaleza de la variable.**

   a) **Variables cualitativas o de atributos**: cuando expresan una cualidad, característica o atributo. Por ejemplo: la calidad de un servicio según la opinión de los consumidores, estado de una solicitud, clasificación de los animales según su alimentación, el sector al que pertenecen los bienes y servicios de un país, entre otras. A su vez, las variables cualitativas reciben una categorización más que se comenta enseguida.

      i. **Variables nominales:** son aquellas variables que establecen la distinción de los elementos en diversas categorías, sin implicar algún orden entre ellas, distribuye a la unidad de análisis en dos o más categorías. Ejemplos: estado civil, profesiones, lugar de nacimiento, productos que exporta un país, causas de deserción escolar, entre otras.

      ii. **Variables ordinales:** aquellas variables que implican orden entre sus categorías, pero no grados de distancia iguales entre ellas; están referidas a un orden o jerarquía, en el que las categorías expresan una posición de orden. Ejemplo: clases sociales, grado de simpatía, rango de agresividad, orden de mérito, etc.

   b) **Variables cuantitativas**: cuando el valor de la variable se expresa a través de una cantidad o magnitud que es de carácter numérico. El dato o valor puede resultar de la operación de contar o medir. Por ejemplo: número de bacterias presentes en una muestra de agua, intensidad de una señal de radio en micro vatios, tamaño de partículas en un experimento de laboratorio, etc. Las variables cuantitativas pueden ser a su vez:

i) **Variable discreta:** cuando los valores de la variable son susceptibles de contarse; es decir, se representa por números naturales. Ejemplos: cantidad de accidentes de tráfico en una intersección por mes, cantidad de defectos en un lote de productos fabricados, etc.

ii) **Variable continua:** cuando los valores de la variable son susceptibles de medirse. Es toda variable cuyo valor se obtiene por medición o comparación con una unidad o patrón de medida. Las variables continuas pueden tener cualquier valor dentro de su rango o recorrido, por tanto, se expresa con números racionales. Ejemplos: nivel de ruido ambiental, consumo de energía eléctrica de una casa, velocidad de crecimiento de una planta, temperatura, etc.

2. Según el orden de las observaciones.

a) **Datos atemporales o no ordinales:** cuando se prescinde del orden en que se realiza las observaciones. El orden de la observación no altera el valor de la variable. Por ejemplo, averiguar las edades de los alumnos de un aula, en este caso, cualquiera sea el orden que elijan los alumnos, la edad no va a variar.

b) **Series de tiempo, cronológicas o históricas:** cuando en la obtención de los datos se tiene en cuenta el orden cronológico de la observación, o al momento o tiempo que pertenece, son variables analizadas en función del tiempo. Por ejemplo: población demográfica en la década 1985-1995, variación mensual de precios, volumen anual de las exportaciones, evolución y tendencia de la matrícula en educación superior en el periodo 1990-1995, etc.

3. Según la relación entre las variables:

a) **Variables dependientes**: son aquellas que se explican por otras variables, son los efectos o resultados respecto a los cuales hay que buscar su motivo, causas o razón de ser. Es la variable que traduce la consecuencia del efecto de una o varias razones o causas, de otras variables.

b) **Variables independientes**: son las variables explicativas o predictivas, cuya asociación, relación o influencia en la variable dependiente se pretende descubrir en la investigación. A continuación, se presentan ejemplos en los que las variables

independientes (VI) son las que traducen o explican las causas o razones de las variaciones en la variable dependiente (VD); simplificando las causas o antecedentes serían las variables independientes (VI) y el efecto o consecuente es la variable dependiente (VD):

- ✓ El presupuesto familiar (VD) depende de los ingresos (VI).
- ✓ El volumen de ventas (VD) se explica por la inversión en propaganda (VI).
- ✓ El número de hijos por familia (VD) tiene relación con el nivel educativo de los padres (VI).
- ✓ El analfabetismo (VD) tiene relación con el lugar de residencia (VI) y la expansión del servicio educativo (VI).

4. **Según la escala de medición:** aunque las variables ordinales, que establecen categorías y jerarquías, han sido previamente explicadas, es importante recordar que ellas también forman parte de esta clasificación. A continuación, se detallan las variables de intervalo y razón:

   a) **Variables de intervalo:** son aquellas que suponen a la vez orden y grados de distancia iguales entre las diversas categorías, pero no tienen un origen natural, sino convencional, tiene un cero relativo. Ejemplo: coeficiente de inteligencia, temperatura, puntuación obtenida en una escala, etc.

   b) **Variables de razón:** estas variables comprenden a la vez a todos los casos anteriores, distinción, orden, distancia y origen único natural; el valor se expresa con un número real, tiene un cero absoluto. Ejemplo: edad, peso, ingresos, número de hijos, producción, accidentes de tránsito, etc.

5. **Según la amplitud de las unidades de observación:**

   a) **Variables individuales:** referida a características de individuos o personas, una empresa, centro educativo. Son variables para estudio de casos, donde se pueden subdividir en variables públicas y privadas.

      i. **Variable pública**: aquella en que los valores individuales son conocidos por otras personas y se sabe que son conocidos. Ejemplos: Edad, sexo, la ocupación, estado civil, deporte que practica, etc.

ii. **Variable privada**: cuando los valores individuales pueden ser conocidos por otros, una vez averiguados. Ejemplos, el coeficiente de inteligencia, opiniones frente a la política económica, grado de agresividad, conductas de consumo, etc.

b) **Variables colectivas**: aquellas que se refieren a características de las unidades cuando estas son colectivas, conjuntos o grupos (empresas, ciudades, escuelas, etc.). Ejemplos: tasa de mortalidad, urbanización, nivel educativo promedio, tasa de crecimiento demográfico, etc.

Por otro lado, los conjuntos de datos se pueden describir, esencialmente, a través de su representación gráfica y sus medidas numéricas. Las representaciones gráficas son aquellas que permiten resumir la información de una variable estadística (o más) de forma visual; lo anterior, se puede llevar usualmente a cabo a través de los gráficos estadísticos y las tablas de frecuencia.

Tablas de frecuencia: según Montero (2007) toda la información obtenida acerca de los diferentes valores o modalidades que pueda tomar una variable se puede ordenar y presentar en tablas de frecuencias. Antes de proceder al estudio de estas tablas, es necesario conocer los diferentes tipos de frecuencias que existen.

Supóngase que, de una población o muestra determinada, se quiere analizar la variable X, que toma los valores $x_1, x_2, x_3, \dots, x_p$, donde cada uno de ellos puede repetirse una o varias veces. Se define la *frecuencia total (N)* como el número total de observaciones, valores o datos que se tienen de la variable X. El número de veces que se repite el valor x_i $(1 \leq i \leq p)$ se denomina *frecuencia absoluta* o, simplemente *frecuencia*. Si la variable X representa p valores distintos, cada uno de los cuales se repite una o varias veces, la suma de las respectivas frecuencias de cada valor será igual a la frecuencia total.

$$\sum_{i=1}^{p} n_i = n_1 + n_2 + n_3 + \cdots + n_p = N$$

Para conocer la importancia relativa que tiene cada valor dentro del conjunto total de valores observados se utiliza la *frecuencia relativa*, que se define como el cociente entre la frecuencia absoluta del valor considerado y la frecuencia total.

$$f_i = \frac{n_i}{N}$$

La expresión anterior multiplicada por 100, refleja la frecuencia absoluta en porcentaje respecto al total de datos. La suma de las frecuencias relativas es igual a la unidad o a 100 si se trabaja con porcentajes.

$$\sum_{i=1}^{p} f_i = \frac{\sum_{i=1}^{p} n_i}{N} = \frac{n_1 + n_2 + n_3 + \cdots + n_p}{N} = \frac{N}{N} = 1$$

La *frecuencia absoluta acumulada* hasta el valor x_i (N_i) indica la suma del número de frecuencias que hay hasta el valor x_i, incluido este y supuestos los valores ordenados de menor a mayor.

$$N_i = \sum_{x \leq x_i} n_i$$

La última frecuencia absoluta acumulada es igual a la frecuencia total N:

$$N_1 = n_1$$

$$N_2 = n_1 + n_2$$

$$N_3 = n_1 + n_2 + n_3$$

$$(\ldots)$$

$$N_p = n_1 + n_2 + n_3 + \cdots + n_p = N$$

De forma análoga, la *frecuencia relativa acumulada (F_i)* se define como el cociente entre la frecuencia absoluta acumulada y la frecuencia total.

$$F_i = \frac{N_i}{N}$$

Esta expresión multiplicada por 100 es equivalente a la frecuencia acumulada en porcentaje respecto al total. La última frecuencia relativa acumulada es igual a la unidad o a 100 si se trabaja con porcentajes.

$$F_1 = f_1$$

$$F_2 = f_1 + f_2$$

$$F_3 = f_1 + f_2 + f_3$$

$$(\ldots)$$

$$F_p = f_1 + f_2 + f_3 + \cdots + f_p = 1$$

Todo lo expuesto anteriormente queda reflejado en la siguiente tabla de frecuencias (Tabla 4), en la que la primera columna representa los valores de la variable (x_i), la segunda las frecuencias

absolutas (n_i), la tercera las frecuencias relativas (f_i), la cuarta las frecuencias absolutas acumuladas (N_i) y en la quinta las frecuencias relativas acumuladas (F_i).

Tabla 4. *Tabla de frecuencia para datos no agrupados*

x_i	n_i	$f_i = \frac{n_i}{N}$	N_i	$F_i = \frac{N_i}{N}$
x_1	n_1	$f_1 = \frac{n_1}{N}$	N_1	$F_1 = \frac{N_1}{N}$
x_2	n_2	$f_2 = \frac{n_2}{N}$	N_2	$F_2 = \frac{N_2}{N}$
⋮	⋮	⋮	⋮	⋮
x_i	n_i	$f_i = \frac{n_i}{N}$	N_i	$F_i = \frac{N_i}{N}$
⋮	⋮	⋮	⋮	⋮
x_p	n_p	$f_p = \frac{n_p}{N}$	N_p	$F_p = \frac{N_p}{N}$
	N	1		

Nota. La tabla fue tomada de Montero (2007).

Tabla 5. *Ejemplo de tabla de frecuencia para datos no agrupados.*

Tipo de producto	n	N	F
Snacks	87	87	0,087
Cárnicos	84	171	0,084
Cereales	95	266	0,095
Frutas	86	352	0,086
Alimento infantil	73	425	0,073
Bebida	87	512	0,087
Cosméticos	74	586	0,074
Cuidado personal	84	670	0,084
Material de oficina	80	750	0,08
Verduras	89	839	0,089
Ropa	78	917	0,078
Doméstico	83	1000	0,083
	1000		1

Cuando el número de valores que toma la variable no es demasiado grande (teniendo en cuenta que esto es subjetivo y está sujeto a juicio del investigador), toda información disponible aparece recogida en una tabla como la anterior, donde cada uno de los valores se encuentra asociado con su correspondiente frecuencia. En este caso, se dice que la distribución se encuentra *no agrupada*.

Sin embargo, cuando una variable presenta un gran número de valores, como suele ocurrir con las variables continuas (aunque no es exclusivo de estas), es recomendable agrupar estos valores en intervalos, esto facilita su presentación y simplifica la realización de ciertos cálculos, haciéndolos relativamente más fáciles de determinar. Cuando se da esta situación, la distribución de frecuencias se denomina *agrupada en intervalos*, siendo la frecuencia (absoluta) de cada intervalo igual a la suma de las frecuencias de todos los valores incluidos en él.

Para realizar una distribución agrupada en intervalos, es fundamental primero construir dichos intervalos. Para esto, se debe considerar el *recorrido* o *rango* de la variable, el cual se obtiene al calcular la diferencia entre el máximo y el mínimo valor de la variable.

$$R = Máx(x_i) - Mín\ (x_i)$$

Los *intervalos* se representan como $(L_{i-1} - L_i)$, donde L_{i-1} es el extremo inferior del intervalo y L_i el extremo superior. La diferencia entre ambos extremos es la *amplitud* del intervalo, que se denota c_i, y se determina en función del rango y el número de intervalos (k) que el investigador define, según la fórmula:

$$c_i = \frac{R}{k}$$

Normalmente se establece que los intervalos son abiertos por la izquierda y cerrados por la derecha $(L_{i-1} - L_i]$, con lo cual la frecuencia incluye los valores comprendidos entre L_{i-1} y L_i, incluido L_i y excluido L_{i-1}. En el caso particular del primer intervalo ambos extremos se consideran cerrados, por lo que la frecuencia incluye ambos extremos del intervalo.

Cuando se trabaja con distribuciones agrupadas en intervalos, no es posible asignar a cada intervalo su frecuencia (ya que ésta corresponde a un conjunto de valores y no a uno solo), ni tampoco se puede operar matemáticamente para determinar ciertos cálculos de interés. Ante esta situación, se elige la marca de clase como el valor representativo del intervalo, la cual se define como el punto medio del mismo.

$$x_i{}' = \frac{L_i + L_{i-1}}{2}$$

Esta manera de proceder lleva implícita la suposición de que los valores se encuentran uniformemente distribuidos a lo largo del intervalo. Por lo tanto, esto conlleva una pérdida de información, ya que esta suposición implica que se omite la variabilidad específica dentro de los intervalos.

La Tabla 6 ilustra una distribución de datos agrupados en intervalos; esta presenta en la primera columna los intervalos $(L_{i-1} - L_i]$, en la segunda la marca de clase $x_i{}'$, en la tercera las frecuencias absolutas (n_i), en la cuarta las frecuencias relativas (f_i), en la quinta las frecuencias absolutas acumuladas (N_i) y en la sexta las frecuencias relativas acumuladas (F_i).

Tabla 6. *Tabla de frecuencia para datos agrupados*

$(\boldsymbol{L_{i-1} - L_i}]$	$\boldsymbol{x_i}'$	$\boldsymbol{n_i}$	$\boldsymbol{f_i}$	$\boldsymbol{N_i}$	$\boldsymbol{F_i}$
$(L_0 - L_1]$	$x_i' = \frac{L_0 + L_1}{2}$	n_1	f_1	N_1	F_1
$(L_1 - L_2]$	$x_i' = \frac{L_1 + L_2}{2}$	n_2	f_2	N_2	F_2
⋮	⋮	⋮	⋮	⋮	⋮
$(L_{i-1} - L_i]$	$x_i' = \frac{L_{i-1} + L_i}{2}$	n_i	f_i	N_i	F_i
⋮	⋮	⋮	⋮	⋮	⋮
$(L_{p-1} - L_p]$	$x_i' = \frac{L_{p-1} + L_p}{2}$	n_p	f_p	N_p	F_p
		N	1		

Es importante aclarar que, aunque la agrupación en intervalos hace que la distribución sea más manejable, puede llevar a una pérdida u omisión de información, por ejemplo:

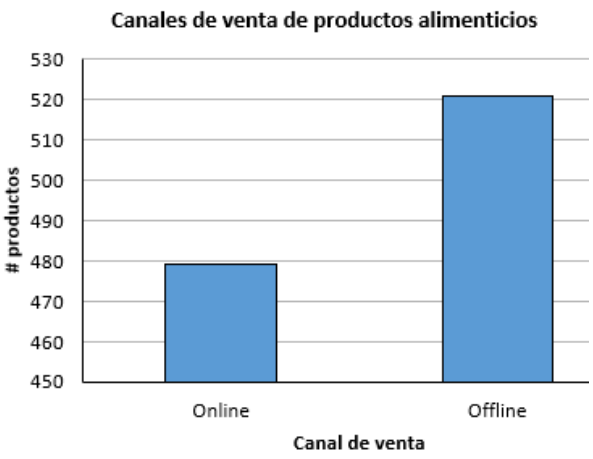
1. Como se indicó anteriormente, en una tabla de frecuencias agrupadas no se puede determinar la distribución exacta de los datos dentro de cada intervalo.
2. La agrupación de datos puede simplificar demasiado la información, lo que puede ocultar patrones o tendencias que serían visibles en los datos originales. Esto es especialmente relevante si hay variaciones significativas dentro de los intervalos.
3. En algunas ocasiones, se utiliza la marca de clase para representar intervalos en las tablas de frecuencias agrupadas. Esto puede llevar a malinterpretaciones, ya que la representación de un intervalo con un solo valor no refleja la variabilidad de los datos.
4. La elección de cómo agrupar los datos puede influir en la interpretación de los resultados. Diferentes agrupaciones pueden llevar a conclusiones diferentes, y no hay un criterio único para determinar la mejor forma de agrupar.
5. Ciertas medidas estadísticas, como la media o la desviación estándar, son menos precisas cuando se basan en datos agrupados en lugar de en datos individuales, ya estas utilizan la marca de clase, como representación de los datos de un intervalo, para ser calculadas.
6. Algunas medidas estadísticas, como la media y la desviación estándar, suelen ser menos precisas cuando se calculan a partir de datos agrupados. Esto ocurre porque, en su cálculo, se utiliza el valor de la marca de clase para representar los datos dentro del intervalo.

Gráficos estadísticos: según Ávila (1996) son representaciones pictóricas (figuras geométricas, de superficie o volumen) que tienen por finalidad ilustrar los cambios o dimensión de una variable, para comparar visualmente dos o más variables similares o relacionadas. Para una rápida comprensión de situaciones o variaciones en cantidades es muy útil traducir los números en gráficos o imágenes. Todo gráfico es superior al texto escrito porque transmite de manera casi instantánea, hechos, cantidades y comportamientos de variables.

Las principales partes de un gráfico estadístico son: 1) Titulo, es una descripción del contenido; debe indicar claramente la naturaleza del fenómeno representado. 2) Los diagramas, está dado por el propio dibujo del gráfico, y como en el cuerpo del gráfico, están representados los datos indicados en el título. 3) Escalas o leyendas, son indicaciones donde se precisa la correspondencia entre los elementos del gráfico y la naturaleza de las medidas representadas. 4) Fuente de la que se obtienen los datos estadísticos representados.

A continuación, se presenta una breve descripción de algunos gráficos estadísticos usuales; es importante puntualizar que no se abordarán aspectos ni paso a paso relacionados con la construcción de estos, teniendo en cuenta que esta tarea se puede llevar a cabo a través de diversos softwares. Por otro lado, no es propósito de este trabajo ahondar en estos detalles en este capítulo, pero tampoco en el diseño mismo de las tareas.

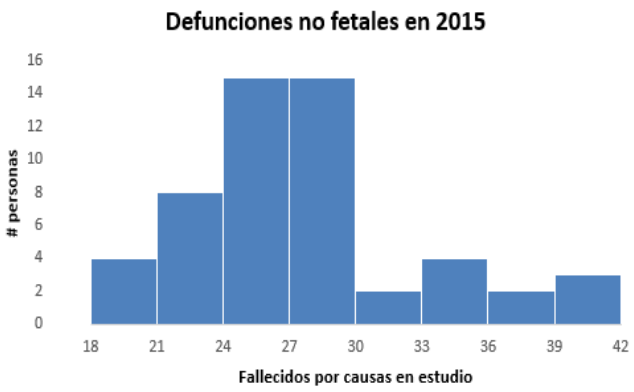
Tabla 7. *Definiciones de gráficos estadísticos*[3]

a) Gráfico de barras Dicho gráfico permite visualizar las frecuencias relativas o absolutas de las observaciones de una variable. Un gráfico de barras puede utilizarse para variables cualitativas y cuantitativas discretas, siempre que el número de categorías no sea muy grande. Consta de una barra para cada categoría. La altura de cada barra viene determinada por la frecuencia absoluta o la frecuencia relativa de la categoría respectiva y se muestra en el eje y. Se disponen las barras en el eje x según sus rangos o valores.	

[3] Los datos utilizados para elaborar los gráficos de la tabla son meramente ilustrativos y no corresponden a datos reales o contextualizados. Las definiciones fueron tomadas de Ávila (1996).

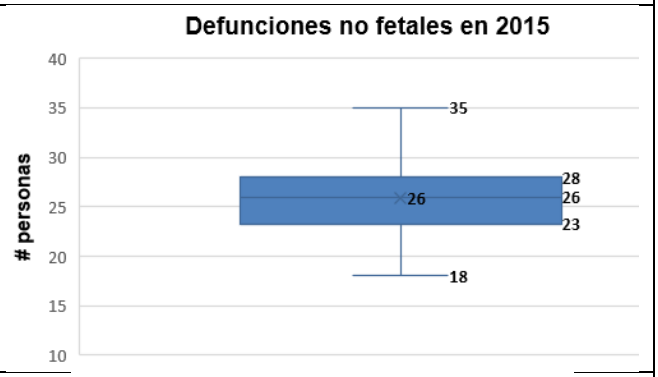
b) Histograma

Un histograma permite representar la distribución de observaciones de variables continuas. Se basa en la idea de categorizar los datos en diferentes grupos (o intervalos) y trazar las barras para cada categoría donde la altura representa su frecuencia. Una consideración importante para este concepto es que el área de las barras (altura × anchura) es proporcional a la frecuencia relativa.



c) Diagrama de caja y bigotes

Este gráfico resume la distribución de variables continúas, utilizando su mediana, cuartiles y valores extremos. El gráfico de caja en su conjunto nos informa sobre la distribución de los datos, el rango y la variabilidad de las observaciones, así como de la presencia de datos atípicos.

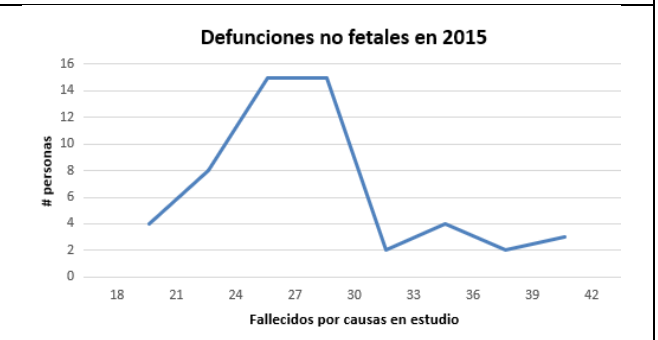


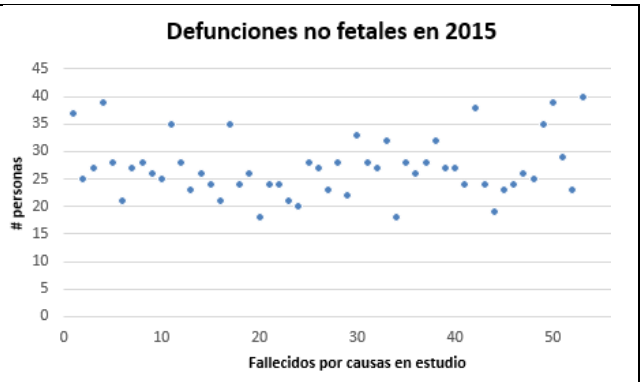
d) Gráfico circular

Son otra opción para visualizar las frecuencias absolutas y relativas de variables cualitativas y cuantitativas discretas. Un gráfico circular es un círculo dividido en sectores donde cada uno de los sectores representa una observación y el tamaño de un sector determinado representa la frecuencia relativa o absoluta de la observación asociada al sector.

e) Polígono de frecuencia

En cada punto medio o marca de clase de los diferentes grupos de datos (o intervalos), se construye un segmento de altura igual a la respectiva frecuencia, luego se unen los extremos con un segmento, formando una línea poligonal. Este tipo de gráfico estadístico permite representar distribuciones de observaciones de variables continuas.



f) Diagrama de dispersión Según Aroca et al. (2009) el diagrama de dispersión es un tipo de gráfico que se utiliza para mostrar los valores de dos variables para un conjunto de datos, son útiles para mostrar la relación entre diferentes puntos de datos, y utiliza valores numéricos para ambos ejes. Los datos se muestran como un conjunto de puntos, cada uno con el valor de una variable en el eje de abscisas y el valor de la otra variable situado en el de ordenadas.	

En cuanto a las medidas numéricas de interés para describir un conjunto de datos, es importante precisar que no todas las que se van a describir a continuación, se pueden calcular en cualquier conjunto de datos, esto depende de la variable estadística de estudio de la muestra o población en cuestión. Teniendo en cuenta lo anterior, se especificará, en la descripción de cada una de las medidas numéricas, su aplicabilidad.

Las primeras medidas numéricas corresponden a aquellas que describen lo que sucede en el centro de los datos y se denominan *medidas de tendencia central.* Las segundas medidas a aquellas que arrojan información en relación con la variabilidad, llamadas *medidas de dispersión*. Es importante mencionar que la conceptualización que se presenta a continuación corresponde a lo definido para conjuntos de datos no agrupados.

Medidas de tendencia central: según Heumman y Schomaker (2016) se denominan medidas de tendencia central a aquellas que describen el promedio o el centro de un conjunto de datos; adicionalmente, estas medidas dan una idea acerca de la ubicación donde se concentra la mayoría de los datos. Las medidas de tendencia central más usuales son:

Media aritmética: dadas $x_1, x_2, x_3, \dots x_n$, observaciones de un conjunto (muestra o población) de tamaño n la media aritmética o promedio de estos datos se define como:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Es importante precisar que la media solo se puede determinar en un conjunto de datos de variables cuantitativas. El valor de la media puede verse afectado de manera desproporcionada por la existencia de valores atípicos.

La media aritmética cumple las siguientes propiedades:

- *Propiedad 1:* si los datos se transforman linealmente como $y_i = a + bx_i$, donde a y b son constantes, se cumple que

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(a + bx_i) = \frac{1}{n}\sum_{i=1}^{n} a + \frac{b}{n}\sum_{i=1}^{n} x_i = a + b\bar{x}$$

En otras palabras, cuando se suman o multiplican constantes a cada dato, la media responde de manera proporcional a esa misma constante, por ejemplo, si se multiplican todos los valores en un conjunto de datos por 3 y, además, se les suma 2, la media también se triplicará y se incrementará en 2.

Teniendo en cuenta que la desviación se define como la diferencia entre una observación x_i y un valor fijo, pero no arbitrario A, que puede corresponder a la media aritmética o puede tomar el valor de una de las observaciones del conjunto de datos $(x_i - A)$, entonces:

- *Propiedad 2:* la suma de las desviaciones de cada variable alrededor de la media aritmética es cero.

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

  Lo cual corresponde a mostrar lo siguiente

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = \sum_{i=1}^{n} x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0$$

Mediana: dadas $x_1, x_2, x_3, \ldots x_n$, observaciones de un conjunto y ordenadas de tal manera que $x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$. La mediana se define como:

$$Me = \begin{cases} x_{\left[\frac{n+1}{2}\right]} & si\ n\ es\ impar \\ \frac{1}{2}\left[x_{(n/2)} + x_{(n/2)+1}\right] & si\ n\ es\ par \end{cases}$$

Es importante aclarar que la expresión definida para el subíndice corresponde a la posición de la observación. Así, esta medida es el valor que divide el conjunto de datos en dos partes iguales, de modo que al menos el 50% de los valores son mayores o iguales a la mediana y al menos el 50% de los valores son menores o iguales a la mediana.

Es importante señalar que la mediana puede calcularse en ciertos conjuntos de datos con variables cualitativas ordinales (aunque su interpretación es limitada), así como con variables cuantitativas; además, al basarse en la secuencia ordenada de los valores de la variable de un conjunto de datos, la existencia de algunos valores extremos no afecta su valor.

Moda: dados n valores $x_1, x_2, x_3, \dots x_n$, la moda se define como el valor de la variable que ocurre con mayor frecuencia en el conjunto. Es importante precisar que es posible que en un conjunto haya más de una moda; los conjuntos de datos con una moda se denominan unimodal; con dos modas, bimodales y con tres o más modas, multimodales. Además, la moda a diferencia de la media y la mediana, sí se puede determinar en un conjunto de datos de variables cualitativas o cuantitativas.

La medida numérica que se presenta a continuación corresponde a una generalización de la idea de mediana, y es conocida como una medida de localización o posición.

Cuantiles: los cuantiles son una generalización de la idea de mediana. La mediana es el valor que divide los datos en dos partes iguales. De manera similar, un cuantil divide los datos en otras proporciones. En general, sea α un número entre cero y uno. El $(\alpha * 100)\%$ - cuantil, se define como el valor que divide los datos en proporciones de $(\alpha * 100)\%$ y $(1 - \alpha) * 100$ % tal que al menos $\alpha * 100$ % de los valores sean menores o iguales al cuantil y al menos $(1 - \alpha) * 100$ % de los valores son mayores o iguales al cuantil.

De acuerdo con los valores que tome α se pueden determinar cuantiles usuales como: Deciles (10 partes), percentiles (100 partes), quintiles (5 partes) y cuartiles (4 partes).

Dados n valores conformados por $x_1, x_2, x_3, \dots x_n$ y ordenados de tal manera que $x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$, el $(\alpha * 100)\%$ - cuantil se calcula como:

$$x_\alpha = \begin{cases} x_{(k)} & si\ n\alpha \notin \mathbb{Z} \Rightarrow k > n\alpha \ con \ k \in \mathbb{Z} \\ \frac{1}{2}\left[x_{(n\alpha)} + x_{(n\alpha+1)}\right] & si\ n\alpha \in \mathbb{Z} \end{cases}$$

Los cuantiles se pueden determinar en variables cuantitativas.

Medidas de dispersión: se denominan medidas de dispersión a aquellas que cuantifican la concentración o dispersión de valores alrededor de cualquier valor particular (Heumman y Schomaker, 2016). Se distinguen de forma usual las siguientes medidas numéricas:

Rango: dados n valores $x_1, x_2, x_3, \dots x_n$ ordenados de tal manera que $x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$; el rango se define como la diferencia entre el valor máximo y mínimo de los datos, es decir:

$$R = x_n - x_1$$

El rango solo se puede determinar en un conjunto de datos con variables cuantitativas. Además, por su simplicidad, el rango proporciona una rápida indicación de la variabilidad existente entre los valores del conjunto de datos. Sin embargo, como principal defecto, se debe mencionar que es una medida muy sensible a datos extremos o atípicos del conjunto de datos y, por tanto, su análisis debe acompañarse del estudio de otras medidas de dispersión (Canavos, 1988).

Rango intercuartil: se define como la diferencia entre los cuartiles que representan el 75% y 25% del conjunto de datos, como se muestra a continuación

$$d_Q = x_{0.75} - x_{0.25}$$

Esta medida incluye el centro del conjunto de datos y contiene el 50 % de los valores de la variable. El rango intercuartil solo se puede determinar en un conjunto de datos con variables cuantitativas.

Con el fin de definir la medida numérica denominada *varianza*, es importante establecer algunos conceptos previos, es por ello por lo que se define enseguida la desviación de los datos. Al considerar el promedio de las desviaciones de n observaciones, se obtiene la siguiente expresión (es importante tener en cuenta lo definido como desviación):

$$D = \frac{1}{n}\sum_{i=1}^{n}(x_i - A)$$

Heumman y Schomaker (2016) señalan que esta medida tiene como inconveniente que su resultado puede ser positivo o negativo y, en consecuencia, la sumatoria puede ser muy pequeña o incluso cero. Por lo tanto, utilizar D como medida de variabilidad no es una buena idea ya que D puede ser pequeño incluso para una gran variabilidad en los datos. Dicho inconveniente se resuelve usando el valor absoluto; sin embargo, también se puede considerar el uso de los cuadrados de las desviaciones.

En particular, si se toma el cuadrado de las desviaciones, se obtiene la expresión:

$$D = \frac{1}{n}\sum_{i=1}^{n}(x_i - A)^2$$

Y es de interés poder encontrar el valor de A que minimiza los resultados de la expresión, es decir, de la desviación construida; para ello se procede a partir de optimizar la expresión D. En primer lugar, se realiza la derivada respectiva:

$$\frac{dD}{dA} = \frac{1}{n}\sum_{i=1}^{n} 2(x_i - A)(-1)$$

$$= -\frac{2}{n}\sum_{i=1}^{n} x_i + \frac{2}{n}(nA)$$

$$= -\frac{2}{n}\sum_{i=1}^{n} x_i + 2A$$

Expresión que se iguala a cero y se despeja respecto a A para conocer el valor crítico correspondiente:

$$-\frac{2}{n}\sum_{i=1}^{n} x_i + 2A = 0$$

$$-\frac{1}{n}\sum_{i=1}^{n} x_i + A = 0$$

$$A = \bar{x}$$

Para confirmar que este punto crítico corresponde a un mínimo, se puede evaluar la segunda derivada de D de la siguiente manera:

$$\frac{d^2D}{dA^2} = 2$$

Dado que la segunda derivada es positiva, se tiene que la mínima desviación entre los datos se da cuando A toma el valor de la media aritmética $\bar{x}$, es así como:

$$D = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Teniendo en cuenta la definición de desviación y las consideraciones hechas, se define la siguiente medida:

Varianza: dados n valores $x_1, x_2, x_3, \dots x_n$, se define la varianza como el promedio del cuadrado de las distancias entre cada valor de la variable y la media de dichos valores entre N o $n-1$, esta distinción se especifica en el siguiente apartado. Así, la varianza se denota como:

Varianza muestral	***Varianza poblacional***
$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$	$\sigma^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N}$
s^2 = Varianza muestral	σ^2 = Varianza poblacional
n = Tamaño de la muestra	N = Tamaño de la población
$\bar{x}$ = Media muestral	μ = Media poblacional

La varianza es una medida razonablemente buena para la variabilidad debido a que si muchas de las diferencias son grandes (o pequeñas) entonces el valor de las varianzas s^2 será grande (o pequeño), esta es precisamente una de las ventajas de emplear el cuadrado en cambio del valor absoluto. El valor de la varianza puede sufrir un cambio muy desproporcionado por la existencia de valores extremos en el conjunto. Asimismo, la varianza se puede calcular solamente en un conjunto de datos con variables cuantitativas y el resultado obtenido quedará en términos de las unidades cuadradas de las observaciones, lo cual es un defecto que se soluciona con la desviación estándar.

Una de las principales interrogantes al comparar la varianza muestral con la poblacional es por qué en la varianza muestral se divide entre $n-1$. A primera vista, puede parecer que no existe una justificación clara para esta diferencia, sin embargo, a continuación, se explica la razón detrás de esta corrección en la varianza muestral. Según lo expuesto por Triola (2009) se dice que la varianza muestral s^2 es un **estimador sin sesgo** de la varianza poblacional σ^2, lo que significa que los valores de s^2 tienden a igualar el valor de σ^2, en lugar de tender, de manera sistemática, a sobreestimar o subestimar σ^2.

Según Mendenhall y Beaver (2010) un *estimador* es una regla, generalmente expresada como fórmula, que nos dice cómo calcular una estimación basada en información de la muestra. Para

estimar el valor de un parámetro poblacional, se puede usar información de la muestra en la forma de un **estimador**. Los estimadores se calculan usando información de las observaciones muestrales y, en consecuencia, por definición son también *estadísticas*[4]. Por otro lado, un *estimador sin sesgo o insesgado* es aquel cuya esperanza matemática (valor esperado) es igual al verdadero valor del parámetro que se intenta estimar, es decir que si θ un parámetro poblacional que se desea estimar y $\hat{\theta}$ es el estimador del parámetro basado en una muestra, entonces $\hat{\theta}$ se considera insesgado si:

$$\mathbb{E}(\hat{\theta}) = \theta$$

Según Spiegel (1977), para una variable aleatoria discreta X que puede tomar los valores $x_1, x_2, \dots, x_n$ con probabilidades asociadas $P(X = x_i)$, la *esperanza matemática (o valor esperado)* se define como:

$$\mathbb{E}(X) = x_1 P(X = x_1) + x_2 P(X = x_2) + \cdots + x_n P(X = x_n)$$

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i P(X = x_i)$$

En el caso de la varianza, es de interés que la esperanza matemática de la varianza muestral sea igual a la varianza poblacional. Por lo tanto, se espera que:

$$\mathbb{E}(s^2) = \sigma^2$$

Remplazando la definición de varianza muestral en $\mathbb{E}(s^2)$, se tiene que

$$\mathbb{E}\left(\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}\right)$$

Expandiendo el cuadrado y distribuyendo la sumatoria, entonces

$$\mathbb{E}\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}^2\right)\right]$$

$$\mathbb{E}\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\right]$$

[4] Las mediciones descriptivas numéricas asociadas con una población de mediciones se llaman *parámetros*; las calculadas a partir de mediciones muestrales reciben el nombre de *estadísticas*.

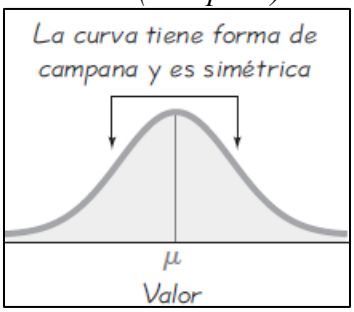
Dado que el valor esperado es un operador lineal sobre variables aleatorias, entonces se cumple que $E(kX) = kE(X)\ (con\ k \in \mathcal{R})$ y $E(X+Y) = E(X) + E(Y)$ con X y Y variables aleatorias. Por tanto, la expresión anterior es igual a

$$\frac{1}{n-1}\left(\mathbb{E}\left[\sum_{i=1}^{n} x_i^2\right] - n\mathbb{E}[\bar{x}^2]\right)$$

Para continuar con la demostración, es esencial definir algunos conceptos. En primer lugar, una *distribución normal* de una variable aleatoria continua se caracteriza por una gráfica simétrica en forma de campana, como se ilustra en la Figura 6 (Triola, 2009). Esta distribución tiene la siguiente función de distribución de probabilidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Figura 6. *Forma de una distribución normal (Campana)*



Nota. Ilustración tomada de Triola (2019). La distribución normal.

En segundo lugar, se define un momento de una variable aleatoria X alrededor de la media μ, también conocido como *momento central* de orden r (m_r). Según Spiegel (1977) se expresa como:

$$m_r = \sum_{i=1}^{n}(x_i - \mu)^r f(x_i) = \sum_{i=1}^{n}(x-\mu)^r f(x) \qquad \text{(Variables discretas)}$$

$$m_r = \int_{-\infty}^{\infty}(x-\mu)^r f(x)\,dx \qquad \text{(Variables continuas)}$$

Particularmente cuando $\mu = 0$, se le denomina *momento de X alrededor del origen,* y según Spiegel (1977), se define como:

$$m_r = \sum_{i=1}^{n} x_i^r f(x_i) = \sum_{i=1}^{n} x^r f(x) \qquad \text{(Variables discretas)}$$

$$m_r = \int_{-\infty}^{\infty} x^r f(x)\, dx \qquad \text{(Variables continuas)}$$

Donde $f(x)$ representa la *función de densidad de probabilidad* de X, que, según Triola (2009), es una representación gráfica de una distribución de probabilidad continua que debe cumplir con las siguientes propiedades:

1. El área total bajo la curva debe ser igual a 1.
2. Cada punto de la curva debe tener una altura vertical igual o mayor que 0. (Es decir, la curva no puede estar por debajo del eje x).

Retomando la demostración, tenemos que

$$\frac{1}{n-1}\left(\mathbb{E}\left[\sum_{i=1}^{n} x_i^2\right] - n\mathbb{E}[\bar{x}^2]\right)$$

Al calcular por separado los dos valores esperados, utilizando las propiedades de esperanza matemática, y suponiendo que cada x_i tiene la misma distribución, se deduce que $E[x_i^2]$ es la misma para todos i. Así se obtiene para la primera:

$$\mathbb{E}\left[\sum_{i=1}^{n} x_i^2\right] = n\mathbb{E}[x_i^2]$$

Al aplicar la definición de momento alrededor del origen, se tiene que:

$$\mathbb{E}[x_i^2] = m_2 = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

Suponiendo que $f(x)$ es la función de densidad de probabilidad normal, entonces

$$m_2 = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Al resolver la integral impropia[5]:

$$m_2 = \sigma^2 + \mu^2$$

[5] Para resolver esta integral, se debe aplicar el método de sustitución con $z = \frac{x-\mu}{\sigma}$.

Así, el resultado de la primera expresión del valor esperado es:

$$n\sum_{i=1}^{n}[x_i^2]f(x_i) = n(\sigma^2 + \mu^2)$$

Del mismo modo, para la segunda expresión se tiene:

$$\mathbb{E}[\bar{x}^2] = m_2 = \int_{-\infty}^{\infty} \bar{x}^2 f(\bar{x})\, dx$$

Al emplear la función de densidad para $\bar{x}$ con media μ, desviación típica $\sigma/\sqrt{n}$ [6] y resolviendo la integral[7], se obtiene:

$$m_2 = \int_{-\infty}^{\infty} \bar{x}^2 \frac{1}{\frac{\sigma}{\sqrt{n}}\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2} dx$$

$$m_2 = \int_{-\infty}^{\infty} \bar{x}^2 \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}\right)^2} dx$$

$$m_2 = \frac{\sigma^2}{n} + \mu^2$$

$$\mathbb{E}[\bar{x}^2] = \frac{\sigma^2}{n} + \mu^2$$

Remplazando y reduciendo

$$\frac{1}{n-1}\left(n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right)$$

$$\frac{1}{n-1}(n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2)$$

$$\frac{n-1}{n-1}(\sigma^2)$$

[6] Según Alvarado y Batanero (2008) una muestra aleatoria simple de tamaño n de cualquier población de media μy desviación típica finita σ. Cuando n es grande, la distribución de la media muestral $\bar{x}$ se aproxima mucho a la distribución normal $N(\mu, \sigma/\sqrt{n})$ con media μ u desviación típica $\sigma/\sqrt{n}$.

[7] Para resolver esta integral, se debe aplicar el método de sustitución con $z = \frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}$.

$$\mathbb{E}(s^2) = \sigma^2$$

Nótese que si el denominador de la varianza muestral no fuese $n-1$, no hubiera sido posible obtener la igualdad buscada. Es decir, s^2 no sería un estimador insesgado de σ^2.

Desviación estándar: se define como la raíz cuadrada positiva de la varianza y se denota como

Desviación muestral	***Desviación poblacional***
$s = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}}$	$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N}}$

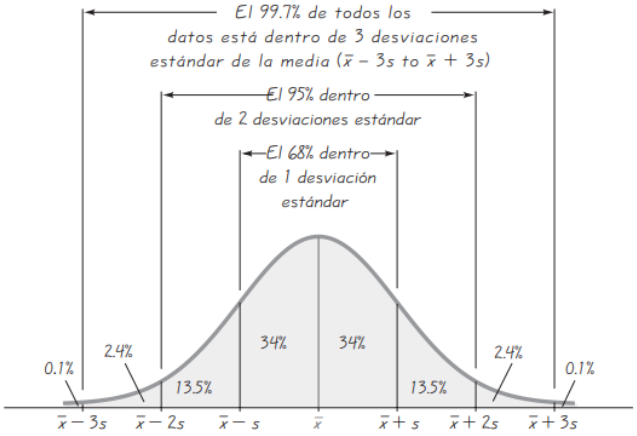
La desviación estándar mide cuánto varían los valores de la variable alrededor de la media aritmética. Un valor bajo de la desviación estándar indica que los valores están muy concentrados alrededor de la media; mientras que, un alto valor de la desviación estándar indica una menor concentración de los valores de la variable alrededor de la media. Además, a menudo se prefiere la desviación estándar sobre la varianza, porque se expresa en las mismas unidades físicas de dichos valores. Es importante recordar que esta medida numérica solo se puede determinar en un conjunto de datos con variables cuantitativas.

Según Triola (2009), una herramienta básica pero efectiva para comprender la desviación estándar es la regla práctica del intervalo. Esta se basa en el principio de que, en muchos conjuntos de datos, la mayoría (alrededor del 95%) de los valores muestrales se encuentran dentro de dos desviaciones estándar desde la media. Aunque se podría mejorar la precisión de esta regla considerando factores como el tamaño de la muestra y la naturaleza de la distribución, se prefiere mantenerla simple a costa de algo de exactitud. También se podría optar por usar tres o incluso cuatro desviaciones estándar en lugar de dos, lo cual es una elección algo arbitraria. Sin embargo, se busca una regla simple que facilite la interpretación de las desviaciones estándar, dejando métodos más precisos para análisis posteriores.

Otra herramienta valiosa para interpretar los valores de la desviación estándar es lo que Triola (2009) nombra como la “regla empírica”. Esta regla se aplica a conjuntos de datos con una distribución aproximadamente normal y establece lo siguiente (ver Figura 7):

- Alrededor del 68% de los valores se encuentran dentro de una desviación estándar de la media.
- Aproximadamente el 95% de los valores están dentro de dos desviaciones estándar de la media.
- Cerca del 99.7% de los valores se sitúan dentro de tres desviaciones estándar de la media.

***Figura** 7. Representación gráfica de la distribución de datos bajo la interpretación de la desviación estándar.*



Nota. Ilustración tomada de Triola (2019). Regla empírica para el cálculo e interpretación de la desviación estándar.

Esta regla permite usar la desviación estándar como una herramienta para comprender qué tan dispersos o concentrados están los datos en una distribución normal, facilitando la identificación de valores atípicos o extremos que se encuentran más allá de estos intervalos. Esta regla indica que el 68% de los datos se encuentran dentro de una desviación estándar de la media, es decir, entre $-1s$ y $+1s$, lo que significa que la mayoría de los datos están cerca de esta, con poca variabilidad. Si ampliamos el rango a dos desviaciones estándar, es decir, entre $-2s$ y $+2s$, el 95% de los datos caerán dentro de ese intervalo, lo que muestra que los datos están más dispersos, pero todavía concentrados alrededor de la media. Finalmente, el 99,7% de los datos se ubican dentro de tres desviaciones estándar, entre $-3s$ y $+3s$, lo que abarca casi toda la distribución.

Desviación media: la desviación media es el promedio de los valores absolutos de las diferencias entre cada valor de la variable y la media de dichos valores. La desviación media está dada por

$$D.M. = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Es una medida de variación útil debido a que en muchas ocasiones el interés se centra en las desviaciones y no es los signos de estas. Esta medida es menos sensible a los efectos inducidos por los valores de las variables extremas del conjunto de datos.

Coeficiente de variación: según Gómez (2012) en el análisis de datos, a menudo surge la necesidad de comparar la variabilidad entre dos o más conjuntos. Si los datos se presentan en las mismas unidades y los promedios de los conjuntos son relativamente similares, la desviación estándar resulta ser una herramienta adecuada para esta comparación. Sin embargo, si alguna de estas condiciones no se cumple, la desviación estándar y otras medidas absolutas de dispersión pierden su utilidad.

Cuando los datos están en unidades diferentes, la comparación de la variabilidad utilizando la desviación estándar se vuelve inapropiada, dado que no tiene sentido comparar desviaciones estándar en unidades distintas, como kilogramos y centímetros. Además, incluso si los datos están en la misma unidad, una diferencia significativa entre los promedios de los conjuntos puede hacer que la comparación directa de las desviaciones estándar sea poco adecuada. Por ejemplo, la variabilidad en el peso de conejos comparada con la de caballos revela que una variación de un kilogramo tiene implicaciones muy distintas dependiendo del tipo de animal.

Por ende, es crucial emplear medidas que sean independientes tanto de las unidades de medida como de la magnitud general de los datos para llevar a cabo comparaciones precisas. Según Triola (2009), una de las medidas frecuentemente utilizadas para estos fines es el coeficiente de variación [CV], el cual se expresa como un porcentaje y refleja la desviación estándar en relación con la media. El coeficiente de variación se define de la siguiente manera:

CV muestral	***CV poblacional***
$CV = \frac{s}{\lvert\bar{x}\rvert}$	$CV = \frac{\sigma}{\lvert\mu\rvert}$

Al simplificar las unidades de medida y proporcionar un valor generalmente expresado en un rango entre 0 y 1, el coeficiente de variación facilita una comparación directa entre variables con diferentes unidades o escalas.

Aunque este trabajo no tiene como objetivo principal abordar este tema, se incluyen las fórmulas correspondientes a las medidas de tendencia central y dispersión para conjuntos de datos agrupados

en la Tabla 8 a modo de referencia. Esta inclusión tiene el propósito de ofrecer una base de comprensión y contexto que puede ser útil para aquellos interesados en profundizar en el análisis de datos agrupados en futuros estudios o aplicaciones relacionadas.

Tabla 8. *Fórmulas de medidas estadísticas para datos agrupados.*

Medidas	**Expresión**	
Media	$\bar{x} = \sum_{i=1}^{k} \frac{m_i f_i}{n}$	Siendo m la marca de clase y f la frecuencia absoluta de cada intervalo. n es total de datos y k la cantidad de intervalos
Mediana	$M_e = e_{m-1} + \frac{\frac{n}{2} + \sum_{j=1}^{m-1} f_j}{f_m} * d_m$	Siendo e_{m-1} el límite inferior de la clase donde está ubicada la mediana d_m la amplitud y f_m la frecuencia absoluta de la clase. n total de datos
Moda	$M_o = e_{m-1} + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})}\ d_m$	Siendo e_{m-1} el límite inferior de la clase modal d_m la amplitud y f_m la frecuencia absoluta de la clase.
Rango	$R = x_n - x_1$	$x_n =$ límite superior de la última clase $x_1 =$ límite inferior de la primera clase
Varianza	$s^2 = \sum_{i=1}^{n} \frac{(m_i - \bar{x})^2}{n-1}$	Siendo m la marca de clase, n total de datos, $\bar{x}$ la media aritmética.
Desviación estándar	$s = \sqrt{\sum_{i=1}^{n} \frac{(m_i - \bar{x})^2}{n-1}}$	Siendo m la marca de clase, n total de datos, $\bar{x}$ la media aritmética

2.3. Tecnologías digitales

Teniendo en cuenta que en los apartados anteriores se abordaron aspectos relacionados con el objeto y el proceso matemático correspondientes al foco de estudio de esta investigación (*i. e*. medidas de dispersión y competencia interpretativa), resulta importante caracterizar el papel de las

tecnologías digitales en el proceso de interpretación de las medidas de dispersión teniendo en cuenta que se pretende trabajar dicho proceso matemático a través del uso de software que posibilite el tratamiento de bases de datos reales.

Al respecto, Jiménez (2019) afirma que una metodología que contenga el recurso tecnológico lleva al educando a la comprensión, transformación y acercamiento del conocimiento matemático, partiendo de situaciones reales e interactuando con ella por medio de herramientas digitales. Esta idea ya había sido expresada por Chance et al. (2007) cuando señalaban que el uso de las herramientas tecnológicas puede contribuir a la compresión de conceptos estadísticos y también permite dar mayor atención a la exploración, análisis e interpretación de los datos, puesto que pasa a un segundo plano el desarrollar la habilidad de usar fórmulas en busca de valores numéricos. Los autores señalan, adicionalmente, que otras de las bondades que ofrece el uso de herramientas digitales en clase, es el tratamiento de grandes conjuntos de datos reales, con los que la aplicación de algunos conceptos y sus procesos correspondientes resulta ser una tarea compleja de ejecutar; en este caso, la tecnología facilita y simplifica el desarrollo de algunos procedimientos y se convierte en un puente para visualizar el conjunto de datos de tal manera que surjan nuevos métodos de análisis de la información presentada.

En correspondencia con lo anterior, surge la necesidad de establecer una herramienta tecnológica que sea de fácil acceso para los estudiantes y que permita trabajar con un alto volumen de datos, por lo que se determina trabajar con Microsoft Excel, pues autores como López et al. (2009) afirman que el uso de paquetes estadísticos especializados resulta de especial ayuda para el procesamiento de los datos, sin embargo, no siempre se tiene fácil acceso a tales herramientas informáticas, por lo que se deben identificar otras alternativas más accesibles en el procesamiento de los datos, como Microsoft Excel.

2.3.1. Microsoft Excel

Esta aplicación está integrada en los programas ofimáticos de Microsoft Office. Luna (2011) lo describe como un paquete especializado que permite realizar cálculos y gráficos estadísticos, además menciona que se creó para desarrollar ejercicios como: nominas, inventarios, bases de datos, entre otros.

En cuanto a sus características, Excel es una hoja de cálculo, a sus archivos se les da el nombre de libros y cada libro puede contener varias hojas. Cada hoja está formada por columnas y filas, ordenadas con el alfabeto y con números, respectivamente, y la intersección de estas se llama celda. La celda se identifica por ser un rectángulo y sirve para guardar información, en las celdas se puede almacenar hasta 255 caracteres y su ancho puede ser modificado. Por otro lado, es importante precisar que en las celdas se puede almacenar valores de texto que incluye letras, números o símbolos, valores numéricos, valores de fecha, fórmulas, funciones, y otros. El nombre de una celda se identifica por la columna y la fila donde se encuentra; por ejemplo, la celda determinada por la columna A y la fila 1, se denomina A1.

Por otro lado, resulta importante definir el concepto función en Excel, según Luna (2011) este término hace referencia a una fórmula predefinida que opera sobre uno o más valores (argumentos) en un orden determinado (estructura) y el resultado se muestra en la celda donde se introduce la fórmula en cuestión. El argumento que utiliza una función es particular en cada caso, puede corresponder a números, texto, valores lógicos, matrices, valores de error o referencias de celda; además, un argumento puede ser una constante, una fórmula o incluso otra función.

Excel tiene varias funciones según el tipo de operación o cálculo que se requiera realizar. Estas funciones pueden ser matemáticas y trigonométricas, estadísticas, financieras, de texto, de fecha y hora, lógicas, de base de datos, de búsqueda y referencia y de información. Sin embargo, en la sintaxis de cualquier función se reconocen 5 elementos: Signo igual (=), nombre de la función, paréntesis de apertura, argumentos de la función separados por puntos y comas, y paréntesis de cierre, es decir:

$$= nombre_funcion(argumento1; argumento2; \dots; argumentoN)$$

Finalmente, en la Tabla 9 se listan las principales funciones que se usarán durante el desarrollo de las tareas propuestas.

Tabla 9. *Funciones de Microsoft Excel*

Funciones	Estructuras
Operadores aritméticos	• + (suma) • - (resta) • / (división) • * (multiplicación)

	• ^ (potencia)
Operaciones matemáticas	• suma (número1; número2;...) Suma los componentes de determinado rango. • raiz(número) Devuelve la raíz cuadrada de un número. • contar(rango) Cuenta cuántos números hay en un rango de celdas. • max (número1; número2;...) Devuelve el mayor valor de la lista de valores indicados. • min (número1; número2;...) Devuelve el menor valor de la lista de valores indicados.
Funciones estadísticas	• promedio(número1; número2;...) Calcula el promedio o media aritmética de los valores que recibe como parámetros. • var.s (número1; número2;...) Devuelve la varianza de la muestra. • var.p (número1; número2;...) Devuelve la varianza sobre una población completa. • desvest.p (número1; número2;...) Calcula la desviación estándar de la población • desvest.s (número1; número2;...) Calcula la desviación estándar de la muestra

2.3.2. Bases de datos

Aunque las bases de datos no corresponden a una tecnología digital por sí mismas, se desarrolla en este subcapítulo la descripción correspondiente teniendo en cuenta que el tratamiento a la información presentada en este formato se llevará a cabo por medio de Microsoft Excel.

Según Garmendia (2018), una base de datos es la más simple forma de almacenar y organizar información en formato digital, lo anterior, permite el procesamiento y el análisis de la información de manera más rápida y eficaz. En cuanto a su estructura el autor menciona que:

> Una base de datos de manera estándar está formada por campos y registros. Los campos son las columnas de una hoja de datos y representa a las variables tanto categóricas como numéricas. Por lo general, se le asigna un nombre a cada columna que represente el tipo de

información desplegada en ella. Los registros son las filas de una hoja de datos y representan la información que entra a la base de datos. (p. 21)

A continuación, en la Figura 8, se ejemplifica lo mencionado con anterioridad a través de una base de datos, formada por 4 campos y 10 registros.

Figura 8. *Ilustración de una base de datos.*

		Columna 1	Columna 2	Columna 3	Columna 4
		A	B	C	D
Fila 1	1	LOCALIDAD	PUNTO	USO	HR (%)
Fila 2	2	Santa Teresa	Punto 1	Pasto	45.8
Fila 3	3	Santa Teresa	Punto 2	Bosque	60.5
Fila 4	4	Santa Teresa	Punto 3	Agroforestal	85.6
Fila 5	5	La Trinidad	Punto 1	Granos Básicos	67.2
Fila 6	6	La Trinidad	Punto 2	Bosque	85.3
Fila 7	7	La Trinidad	Punto 3	Pasto	91.4
Fila 8	8	Diriamba	Punto 1	Granos Básicos	56.1
Fila 9	9	Diriamba	Punto 2	Agroforestal	76.9
Fila 10	10	Diriamba	Punto 3	Café y Sombra	87.4

Nota. La imagen fue tomada de Garmendia (2018, p. 21). Esta es una ilustración de una base de datos estándar en Microsoft Excel bajo lo definido por el autor.

Por otra lado, es importante mencionar que para el diseño de las tareas de esta investigación se tomaron bases de datos de una plataforma nacional llamada Datos Abiertos Colombia[8], en la cual el Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC] (2019) menciona que dispone de información pública (derivada de entidades de carácter público o privado) en formatos que permiten su uso y reutilización bajo licencia abierta y sin restricciones legales para el aprovechamiento por parte de cualquier ciudadano, de forma libre y sin restricciones.

Los beneficios del uso de dichos datos pueden estar asociados a la realización de investigaciones sobre transparencia y control social, el mejoramiento o creación de productos, servicios y negocios innovadores, pronosticar y predecir fenómenos, y generar nuevos conocimientos. En este caso, el propósito de usar dichos datos es propiciar que los estudiantes tengan un acercamiento a información en contextos reales de situaciones o investigaciones que se llevan a cabo en el país, a

[8] Datos abiertos Colombia https://www.datos.gov.co/browse?sortBy=newest&utf8=%E2%9C%93

través de las bases de datos que nos ofrece la plataforma y de esta manera generar un tratamiento a la información allí presentada.

2.3.3. Génesis instrumental

Aun cuando la teoría que se describe a continuación no corresponde exclusivamente al ámbito de las tecnologías digitales, resulta importante tener en consideración lo mencionado en relación con las bondades que ofrecen las tecnologías digitales y, por tanto, entender las posibles interacciones que se dan, en general, entre los individuos y las herramientas tecnológicas. Para ello, se presentan algunos elementos de la génesis instrumental, teniendo en cuenta que, como menciona Pérez (2014), dicha teoría se centra en analizar la relación entre una herramienta tecnológica por parte del sujeto en un contexto educativo desde la actividad de uso. Es importante destacar que en esta investigación la teoría de la génesis instrumental se empleó únicamente como un referente conceptual para la apoyar construcción de las tareas, por cuanto contribuyó a establecer los roles del estudiante y la herramienta tecnológica utilizada. No obstante, este aspecto se aborda con mayor detalle en la sección 3.3 (Diseño de las tareas).

La génesis instrumental tuvo sus inicios luego del estudio de los resultados de un proyecto de investigación en el cual se realizó un contraste entre el discurso teórico de los expertos y la realidad observada en las aulas de clase, del cual surgió y se identificó la necesidad de una nueva perspectiva que considere las relaciones dialécticas entre conceptos y técnicas, que abordara cuestiones sistémicas y reconociera la dimensión instrumental de los procesos de aprendizaje. (Artigue, 2011, como se citó en Pérez, 2014).

Lo mencionado antes, se estableció teóricamente a partir de dos investigaciones fundamentales: la ergonomía cognitiva de Rabardel (2011) que resalta el papel que juegan las herramientas en el proceso de aprendizaje de los estudiantes y la teoría antropológica de lo didáctico de Chevallard (1999) que, por su parte, abre una puerta al análisis del sujeto que aprende a partir del desarrollo conceptual y las herramientas con las que interactúa.

Con el fin de profundizar en los aspectos básicos de la génesis instrumental, es importante abordar previamente algunos términos que están relacionados con la teoría, a saber: artefacto, instrumento, esquemas de utilización, instrumentación e instrumentalización.

✓ **Artefacto**

Según Rabardel (2011) se llama artefacto a un objeto susceptible de uso y que se propone o se construye con un ideal o para algún efecto en la actividad matemática. Por otra parte, Trouche (2005) lo define como una herramienta con la cual un sujeto puede llevar a cabo cierta actividad, esta puede carecer de sentido para el usuario cuando no identifica potencialidades o formas de resolver un problema con dicho objeto. Es por ello que, teniendo en cuenta lo reportado por los autores, se considerará artefacto a aquella herramienta creada con un propósito específico para desarrollar una actividad. En el caso particular de esta investigación se trabajará con Microsoft Excel, es importante precisar que esta tecnología digital no es usada netamente en el campo de la educación matemática; sin embargo, permite el trabajo con elementos que aportan al desarrollo de procedimientos algorítmicos y generan diversas representaciones de un conjunto de datos.

✓ **Instrumento**

Rabardel (2011) concibe como instrumento a un artefacto junto con la habilidad del sujeto en su utilización, es decir, un instrumento se presenta como una amalgama en la que convergen tanto el artefacto como la forma en que el sujeto lo utiliza, con el objetivo determinado que este último le atribuye, teniendo en consideración las particularidades propias del artefacto. Lo anterior, destaca la relación significativa entre el artefacto y las tareas que el sujeto desarrolla a través de este (al desarrollar los procesos de instrumentalización e instrumentación, comentados más adelante), es por ello que en el caso de esta investigación Microsoft Excel permitirá a los estudiantes desarrollar las tareas propuestas por medio del uso de las herramientas como lo son comandos, representaciones gráficas, entre otras. Es importante señalar que Excel es considerado un artefacto antes de su manipulación por parte del sujeto y se transforma en un instrumento una vez que los estudiantes (sujetos) lo utilizan en el desarrollo de las tareas.

✓ **Esquemas de utilización**

Según Leung et al. (2000) estos esquemas de utilización (o de uso) corresponden a un procedimiento sistemático sobre el uso de un determinado instrumento para lograr un fin específico. Por su parte, Rabardel (2011) afirma que la elaboración y la producción de los usos se da después de la concepción del artefacto y se materializan como acciones privadas y sociales que ejecuta el sujeto. Las definiciones permiten establecer que los esquemas de utilización se conciben

a partir de las acciones ejecutadas por los sujetos (que en esta investigación corresponden a los estudiantes), al desarrollar una tarea haciendo uso del artefacto que se convierte en instrumento a través de este proceso. Con el fin de ejemplificar este concepto, para el caso particular en el cual se usen las tecnologías digitales en la interpretación de medidas de dispersión, se proponen ejemplos de posibles esquemas de utilización a propiciar por medio de las tareas mediadas por Microsoft Excel.

- El estudiante realiza simultáneamente cálculos asociados a dos o más conjuntos de datos (que tengan las mismas unidades de medida) con el fin de realizar un paralelo y evidenciar las diferencias entre los valores encontrados (Figura 9).

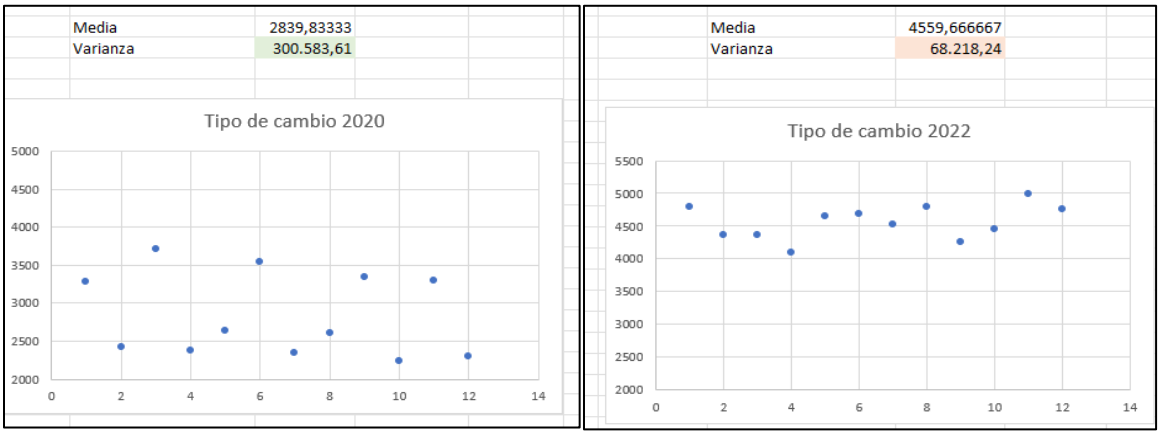
Figura 9. *Cálculo de media y varianza usando las funciones denominadas operaciones matemáticas.*

MES	TIPO DE CAMBIO DE 2020	xi-X	(xi-X)^2
Enero	3288	$ 448,17	$ 200.853,36
Febrero	2415	-$ 424,83	$ 180.483,36
Marzo	3711	$ 871,17	$ 758.931,36
Abril	2373	-$ 466,83	$ 217.933,36
Mayo	2630	-$ 209,83	$ 44.030,03
Junio	3544	$ 704,17	$ 495.850,69
Julio	2340	-$ 499,83	$ 249.833,36
Agosto	2612	-$ 227,83	$ 51.908,03
Septiembre	3335	$ 495,17	$ 245.190,03
Octubre	2244	-$ 595,83	$ 355.017,36
Noviembre	3294	$ 454,17	$ 206.267,36
Diciembre	2292	-$ 547,83	$ 300.121,36
			$ 3.306.419,67
	Media	2839,83333	
	Varianza	300.583,61	

MES	TIPO DE CAMBIO DE 2022	xi-X	(xi-X)^2
Enero	4791	$ 231,33	$ 53.515,11
Febrero	4359	-$ 200,67	$ 40.267,11
Marzo	4362	-$ 197,67	$ 39.072,11
Abril	4092	-$ 467,67	$ 218.712,11
Mayo	4649	$ 89,33	$ 7.980,44
Junio	4682	$ 122,33	$ 14.965,44
Julio	4523	-$ 36,67	$ 1.344,44
Agosto	4787	$ 227,33	$ 51.680,44
Septiembre	4262	-$ 297,67	$ 88.605,44
Octubre	4458	-$ 101,67	$ 10.336,11
Noviembre	4986	$ 426,33	$ 181.760,11
Diciembre	4765	$ 205,33	$ 42.161,78
			$ 750.400,67
	Media	4559,666667	
	Varianza	68.218,24	

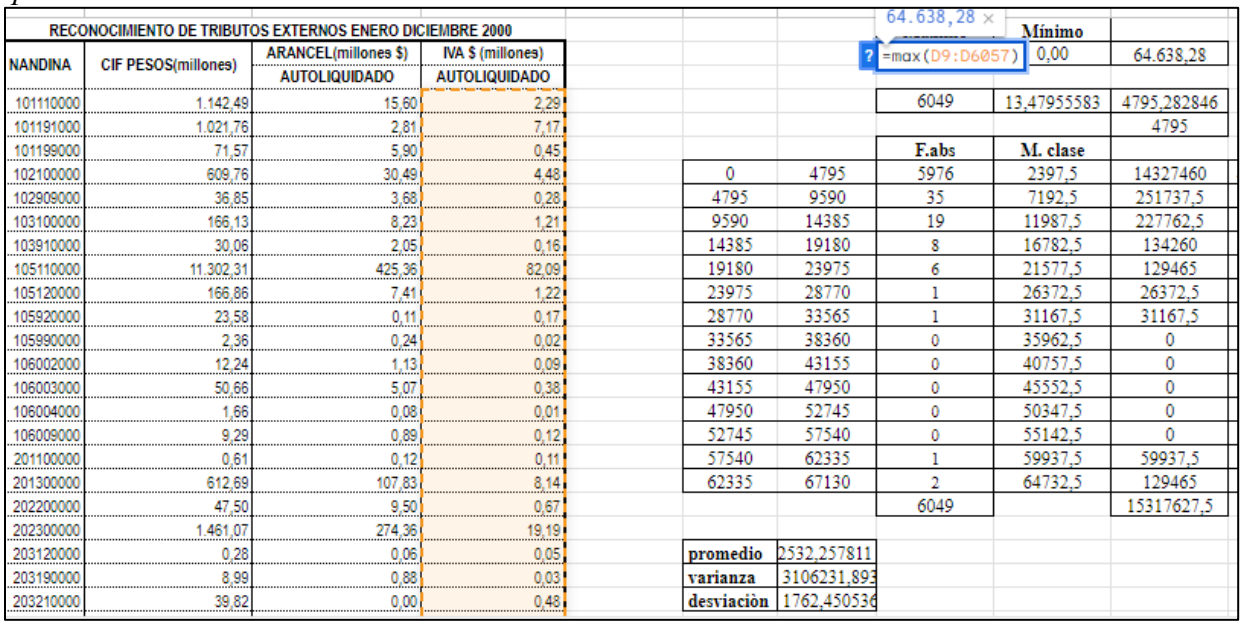
- El estudiante usa la herramienta que le permite generar un gráfico de dispersión para visualizar el conjunto de datos a partir de otro sistema de representación (Figura 10).

. ***Figura 10***. *Gráfico de dispersión elaborados en Excel en relación con dos conjuntos de datos.*



- El estudiante usa los comandos de Microsoft Excel para ejecutar los cálculos correspondientes que le permiten hallar el valor numérico de las medidas de dispersión cuando se propone una situación contextualizada que presenta un alto volumen de datos (Figura 11).

Figura 11. *Cálculo de medidas estadísticas en Excel a partir de funciones denominadas operaciones matemáticas.*



RECONOCIMIENTO DE TRIBUTOS EXTERNOS ENERO DICIEMBRE 2000

NANDINA	CIF PESOS(millones)	ARANCEL(millones $) AUTOLIQUIDADO	IVA $ (millones) AUTOLIQUIDADO
101110000	1.142,49	15,60	2,29
101191000	1.021,76	2,81	7,17
101199000	71,57	5,90	0,45
102100000	609,76	30,49	4,48
102909000	36,85	3,68	0,28
103100000	166,13	8,23	1,21
103910000	30,06	2,05	0,16
105110000	11.302,31	425,36	82,09
105120000	166,86	7,41	1,22
105920000	23,58	0,11	0,17
105990000	2,36	0,24	0,02
106002000	12,24	1,13	0,09
106003000	50,66	5,07	0,38
106004000	1,66	0,08	0,01
106009000	9,29	0,89	0,12
201100000	0,61	0,12	0,11
201300000	612,69	107,83	8,14
202200000	47,50	9,50	0,67
202300000	1.461,07	274,36	19,19
203120000	0,28	0,06	0,05
203190000	8,99	0,88	0,03
203210000	39,82	0,00	0,48

64.638,28 ×

=max(D9:D6057)

		Mínimo	
		0,00	64.638,28
	6049	13,47955583	4795,282846
			4795
	F.abs	M. clase	
0 4795	5976	2397,5	14327460
4795 9590	35	7192,5	251737,5
9590 14385	19	11987,5	227762,5
14385 19180	8	16782,5	134260
19180 23975	6	21577,5	129465
23975 28770	1	26372,5	26372,5
28770 33565	1	31167,5	31167,5
33565 38360	0	35962,5	0
38360 43155	0	40757,5	0
43155 47950	0	45552,5	0
47950 52745	0	50347,5	0
52745 57540	0	55142,5	0
57540 62335	1	59937,5	59937,5
62335 67130	2	64732,5	129465
	6049		15317627,5

promedio	2532,257811
varianza	3106231,893
desviación	1762,450536

✓ Instrumentalización

La instrumentalización está relacionada con el artefacto y se materializa en cuanto a este se le asignan atributos, transformaciones, propiedades y se personaliza tanto su estructura como su funcionamiento (Pedreros y Mejía, 2013). En este proceso se evidencian las limitaciones y virtudes que tiene el artefacto para resolver una tarea, y si estas cualidades le permiten desarrollar o no algo en específico. Es importante en este momento la interacción entre el individuo y el artefacto para caracterizarlo y transformarlo en aquello que sea pertinente y útil en determinada tarea.

✓ Instrumentación

La instrumentación está ligada al sujeto que interactúa con el artefacto en cuestión, y se materializa en cuanto se establece la evolución de los esquemas de utilización (Pedreros y Mejía, 2013). En este proceso resulta importante el determinar cómo se va a incorporar el artefacto en el desarrollo de determinada tarea, de esta manera se empezarán a identificar las adaptaciones y desarrollarán los esquemas de uso con los cuales se quiere dotar el artefacto.

Es así como se denomina génesis instrumental a un proceso de gran generalidad, a partir del cual se da la elaboración de los instrumentos por parte del sujeto a través de la utilización del artefacto Este proceso incluye las formas como el sujeto atribuye funciones al artefacto que está usando en su actividad y por lo tanto las habilidades que desarrolla en este uso; además, involucra la evolución de los artefactos en relación con la actividad del sujeto y el surgimiento de los usos. Como se señaló, este proceso de génesis tiene dos dimensiones producto de la relación bilateral que se establece entre el artefacto y el usuario, la instrumentalización y la instrumentación.

Finalmente, como mencionan Pedreros y Mejía (2013), la génesis instrumental es un proceso de construcción de un instrumento por un sujeto, que va desde la utilización de un artefacto a la construcción de esquemas de utilización para realizar un tipo de tarea. Es importante mencionar que, para los autores, el artefacto puede ser material o abstracto y ayuda o sustenta toda actividad humana al hacer un tipo de tarea; mientras que, el instrumento se constituye como aquello que el sujeto construye desde el artefacto.

2.4. Tareas de aprendizaje

Aunque, según Gómez et al. (2018), el término tarea puede tener diferentes significados dependiendo del contexto educativo en el que se use, en este caso la definen como "una demanda estructurada de actuación, con un contenido matemático y un propósito de aprendizaje, que el profesor propone a los estudiantes con carácter intencional". Adicionalmente, los autores mencionan que particularmente las tareas de aprendizaje (que se ajustan a la definición mencionada), tienen como propósito el contribuir a que los estudiantes logren alcanzar las expectativas que se hayan propuesto con dicha tarea y que superen sus limitaciones de aprendizaje. Lo anterior, se encuentra en correspondencia con lo que se espera alcanzar al desarrollar esta investigación, teniendo en cuenta lo reportado en el Capítulo 1 en la delimitación del problema.

Según Gómez, et al. (2018), una tarea de aprendizaje se compone de los siguientes 7 elementos:

- ✓ *Requisitos:* hace referencia a los conocimientos y destrezas previos que el profesor supone que debería tener un estudiante que va a desarrollar determinada tarea, esto se relaciona directamente con las metas y el contenido matemático.

- ✓ *Meta:* describen el propósito que el profesor tiene al proponer la tarea, haciendo referencia a las expectativas de aprendizaje a las que se pretende contribuir y también a los errores y dificultades que se quieren superar.

- ✓ *Formulación:* corresponde a las instrucciones que el profesor proporciona a los estudiantes, en ellas se establece un punto de partida, lo que se espera que los estudiantes realicen y produzcan como solución teniendo en cuenta el contexto descrito.

- ✓ *Materiales y recursos:* resulta importante establecer la distinción entre los dos términos. Los recursos corresponden a cualquier medio que permita el aprendizaje de un concepto y un proceso matemático, aunque no se haya diseñado con ese fin; mientras que los materiales son diseñados con fines didácticos. Se destaca que el uso de dichos elementos puede ser una posible solución a la superación de dificultades en el proceso de aprendizaje.

- ✓ *Agrupamiento:* aunque es usual plantear que el estudiante trabaje de forma individual, es importante mencionar que también favorece el proceso de aprendizaje el trabajo en parejas o grupal, teniendo en cuenta que esto da lugar a contrastar y validar soluciones, negociar significados, superar dificultades, realizar puestas en común, entre otras bondades.

- ✓ *Formas de interacción:* la interacción se da entre diferentes actores: profesor, estudiante, pareja, grupo pequeño y grupo de clase, y esto a su vez da lugar a diversas formas de comunicación. El profesor es el encargado de planificar y prever las posibles interacciones que se pueden dar al desarrollar la tarea, puesto que resulta importante el generar espacios en los que el estudiante comunique su solución, argumente su validez y llegue a acuerdos con sus pares y profesor.

- ✓ *Temporalidad:* el profesor debe determinar las etapas en las que se desarrollará la tarea estableciendo que parte de la formulación se trabajará, los materiales, el agrupamiento y las interacciones; posteriormente, debe definir el tiempo que tomará a los estudiantes culminar cada una de las etapas y el orden en que deben ejecutarlas.

3. ASPECTOS METODOLÓGICOS

En este trabajo se emplean dos enfoques metodológicos que permitieron abordar los objetivos propuestos. En primer lugar, se llevó a cabo una revisión documental, la cual fue fundamental para la construcción de algunas secciones del marco teórico. En segundo lugar, se implementaron elementos de un experimento de enseñanza, diseñado específicamente para el planteamiento y aplicación de las tareas y posterior análisis de los resultados. A continuación, se presenta la descripción de las estrategias metodológicas empleadas y la manera en que se relacionan con la investigación realizada.

3.1. Revisión documental

Durante la construcción del marco teórico, surgieron diversas problemáticas en relación con el rastreo y la precisión de las definiciones de los conceptos de interpretación y Estadística. En principio, podría parecer que la interpretación es un concepto suficientemente estudiado en el campo de la Educación Matemática, lo cierto es que rastrear sus conceptualizaciones y características no es una tarea tan sencilla, lo cual concitó una búsqueda más amplia y sistemática con el propósito de dilucidar sus particularidades.

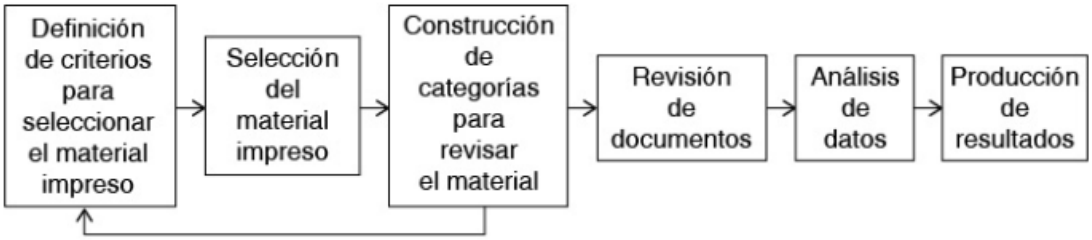
Por otro lado, aunque la definición de Estadística ha sido ampliamente abordada, la existencia de numerosas concepciones sobre este concepto ha generado una falta de consenso en la comunidad científica, lo que complicó aún más la elaboración del marco teórico. Por lo tanto, se decidió llevar a cabo una revisión documental con el fin de aclarar y especificar las definiciones y conceptos relevantes que se presentan en la sección 2.1 del marco teórico, donde se aborda la Competencia interpretativa, y en la sección 2.2, que se refiere a la definición de Estadística.

Antes de detallar el proceso de revisión documental, es fundamental mencionar que esta estrategia se define como una indagación sistemática de material escrito, ya sea en formato impreso o digital, destinada a describir e interpretar el contenido y la perspectiva de sus autores, a partir de consideraciones conceptuales específicas. El objetivo principal de la revisión documental es generar clasificaciones o tipificaciones de propuestas, conceptualizaciones y enfoques que se reflejan en los materiales analizados. Asimismo, es posible realizar conteos de frecuencia de ciertos contenidos y ofrecer explicaciones sobre los hallazgos al comparar datos provenientes de diversas fuentes (Camargo, 2021).

Según Camargo (2021) en la revisión documental toman parte un investigador o grupo de investigadores que seleccionan un conjunto de materiales escritos a partir de una inquietud investigativa y de unos criterios que tienen que ver con el tipo de documentos, el rango de fechas para la revisión, el ámbito geográfico, etc. Cada criterio puede derivar en un conjunto independiente de documentos que serán revisados. Además de criterios para seleccionar el material, los investigadores también elaboran categorías preliminares que determinan lo que van a buscar en los documentos.

En la Figura 12 se presenta el esquema para adelantar una revisión documental propuesto por Camargo (2021), una vez que se han definido los objetivos del estudio.

Figura 10. *Plan de ejecución para la revisión documental.*



Nota. La imagen fue tomada de Camargo (2021, p. 53).

En el contexto específico de este trabajo, el objetivo de la revisión documental fue recopilar las principales características, procesos y componentes relacionados con los conceptos y definiciones de competencia, competencia interpretativa y Estadística (para los dos primeros términos se realizó una revisión y, para la Estadística, otra), para, posteriormente, apropiarse de alguna de las definiciones encontradas o evaluar la posibilidad de formular una propuesta de definición que incorporara los elementos que los autores consideraban relevantes para la investigación.

Es importante señalar que, aunque no se llevó a cabo una construcción formal de categorías para la revisión del material, se consideraron parámetros para la selección, organización y clasificación de documentos. Estos incluyen, por ejemplo, artículos, tesis de maestría o doctorado en el ámbito de la Educación y la Educación Matemática, así como documentos curriculares nacionales (EBCM y LCM) o que evalúan el sistema educativo a nivel nacional (ICFES) e internacional (NCTM, PISA, entre otros).

A continuación, se presenta una descripción concisa de las etapas del proceso de revisión documental, sin entrar en detalles, dado que esto ya se abordó previamente en las secciones 2.1 para “Competencia” y “Competencia Interpretativa”, y en la sección 2.2 para el concepto de “Estadística”. Las etapas son las siguientes:

- **Etapa 1: identificación, depuración y selección de la literatura**
  Durante la fase inicial, se realizó la búsqueda y selección de documentos pertinentes que abordaran las definiciones de competencia, competencia interpretativa o Estadística. Este proceso incluyó la depuración de los textos identificados, utilizando criterios específicos para su selección, organización y clasificación.

- **Etapa 2: extracción de las definiciones o concepciones.**
  En esta etapa, se extrajeron las definiciones y concepciones más relevantes de los documentos seleccionados. Se recopiló información sobre el tratamiento que diversos autores y organismos dan a los conceptos, lo que permitió obtener una perspectiva amplia y variada sobre las nociones de competencia, competencia interpretativa y Estadística.

- **Etapa 3: organización matricial de las definiciones o concepciones.**
  Las definiciones extraídas se organizaron de manera sistemática en una matriz, lo que permitió comparar y analizar cómo distintos autores y organismos abordan los conceptos. Esta estructura facilitó la identificación de similitudes, diferencias y patrones comunes entre las diversas concepciones, proporcionando una visión particular y detallada de las nociones de competencia, competencia interpretativa y estadística.

- **Etapa 4: identificación de los elementos distintivos, transversales o significativos.**
  En esta etapa, se realizó un análisis de las definiciones previamente organizadas con el propósito de identificar los elementos que caracterizan cada concepción. Este se enfocó, por un lado, en los aspectos distintivos de cada definición y, por otro lado, en los elementos que eran transversales a varias definiciones, contribuyendo así a esclarecer las similitudes y divergencias presentes en las concepciones recolectadas.

- **Etapa 5: propuesta de definición utilizando los elementos relevantes identificados en la etapa anterior.**

  Con base en el análisis anterior, se construyó una propuesta de definición para competencia, competencia interpretativa y estadística, que incorpora los elementos más relevantes y significativos identificados.

3.2. Experimento de enseñanza

En el proceso de determinar la metodología, resultó importante definir dos conceptos relaciones: el enfoque y la aproximación ya que de acuerdo con Camargo (2021) "la metodología de investigación corresponde al conjunto formado por un enfoque, una o varias aproximaciones, una o varias estrategias investigativas, un conjunto de recursos y una fundamentación teórica especifica que genere la base racional para orientar la investigación".

Teniendo en cuenta lo mencionado, la metodología se enmarca en un enfoque fenomenológico el cual Camargo (2021) lo define como:

> El conjunto de supuestos, visiones y prácticas de acercamiento a la investigación que, desde una aproximación antropológica, hermenéutica, o colaborativa social, asumen quienes consideran que los fenómenos de indagación sistemática pueden ser descritos, interpretados, explicados (a fin de construir significados sobre las acciones y discursos humanos) y cuestionados (a fin de buscar alternativas para su transformación). (p.16)

Por otro lado, y en correspondencia con la definición presentada, se asume la aproximación hermenéutica la cual, como menciona Camargo (2021) consiste es rastrear, sin juzgar, las diversas "capas" de significado que subyacen a las acciones, las interacciones y los discursos de las personas (gestuales, verbales o escritos), vistos como conjuntos de signos. Esto va en correspondencia al propósito de observar e identificar si las tareas mediadas por la herramienta tecnológica Excel, contribuye al desarrollo de habilidades propias de la competencia interpretativa.

Atendiendo a lo anterior, se define el uso de la estrategia investigativa denominada *Experimento de enseñanza* la cual consiste en el diseño, implementación y evaluación de una secuencia de enseñanza organizada con la meta de poner en funcionamiento una conjetura sobre un aprendizaje específico, la cual surge de comprometerse con un marco conceptual sobre el aprendizaje en general y sobre la forma de promover el desarrollo específico de un contenido matemático en un

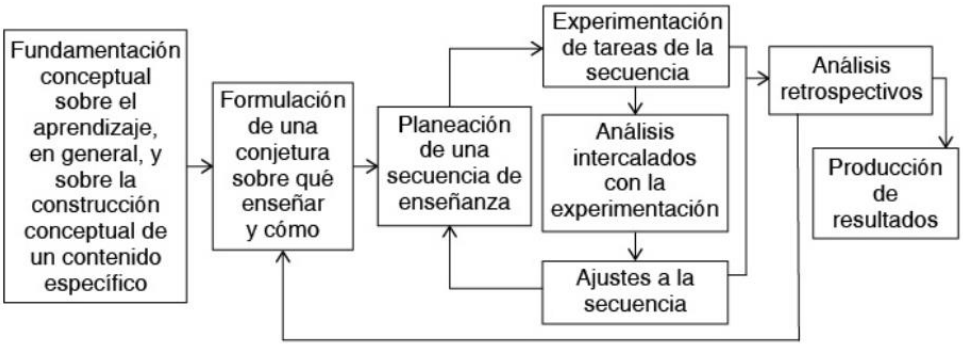
escenario social determinado. Además, por medio de la estrategia se busca determinar si ocurre un desarrollo relevante en contextos enriquecidos mediante diseño y estudio sistemáticos de formas y oportunidades de aprendizaje, estrategias de enseñanza, artefactos mediadores, formas de interacción, entre otros ambientes de clase (Camargo, 2021).

Por otro lado, es importante mencionar que los participantes en un experimento de enseñanza como lo menciona Camargo (2021) son: los investigadores responsables de proponer la secuencia de enseñanza y recopilar la información para analizarla; un profesor que hace parte del equipo de investigación y es el encargado de la gestión de la secuencia; los investigadores que actúan como observadores no participantes en la clase; y los estudiantes del aula en donde tiene lugar la clase, que son considerados colaboradores en el estudio. Para este caso en particular, los estudiantes o colaboradores en el estudio corresponde a 11 alumnos de noveno grado que cursan el espacio de Estadística durante el semestre 2024-1 en la institución educativa Tandem – Apoyo en procesos escolares, mientras que los autores de este documento, también docentes en la misma institución, desempeñan el rol de investigadores que proponen la secuencia de enseñanza, la gestionan y actúan como observadores.

Adicionalmente, Camargo (2021) también hace precisiones en relación con el escenario de un experimento de enseñanza, el cual suele ser el aula regular de clase, aunque también puede suceder en aulas acondicionadas para tal efecto con menos estudiantes o incluso con grupos de estudiantes que se involucran voluntariamente en la enseñanza en situaciones extracurriculares. Por otra parte, el profesor está acompañado de uno o más observadores no participantes que hacen parte del equipo de investigación, ellos contribuyen a interpretar lo que los estudiantes dicen o hacen y apoyan la toma de decisiones sobre los pasos a dar en la clase, su posición de observadores les permite captar información que puede habérsele escapado al profesor y ampliar las evidencias de cara a la interpretación que se busca.

Finalmente, Camargo (2021) sugiere un esquema general de un experimento de enseñanza (Figura 13). Este se lleva a cabo por medio de ciclos repetidos de diseño, puesta a prueba experimental y refinamiento de la conjetura que supone una vía para aprender un contenido:

Figura 11. *Plan de ejecución para el experimento de enseñanza.*



Nota. La imagen fue tomada de Camargo (2021, p. 66).

En la ejecución de los experimentos de enseñanza, Cobb y Gravemeijer (2008, citados en Molina et al., 2011), distinguen tres fases:

- **Fase 1: preparación del experimento:**

En esta fase se consideran inicialmente las siguientes acciones: Definir el problema y los objetivos de la investigación, ubicar el experimento dentro de un contexto teórico más amplio en el que se enmarque el modelo teórico emergente, detallar el contexto en el que se va a realizar el estudio y las características de los participantes. Los aspectos mencionados anteriormente, se describen con mayor detalle en el capítulo 1, en el cual se presenta y se delimita la problemática, especificando la población y el contexto en el que dichos individuos se encuentran inmersos. Adicionalmente, se realiza la descripción de los elementos teóricos relacionados con el objeto matemático alrededor del cual se desarrolla la investigación (ver Capítulo 2).

Por otro lado, hay algunas otras acciones que se especifican en el diseño de las tareas, como lo son: evaluar el conocimiento inicial de los estudiantes, identificar las metodologías de enseñanza adecuadas para los contenidos elegidos en función de los objetivos planteados y los conocimientos previos de los alumnos, diseñar de forma justificada la secuencia de intervenciones en el aula y su temporalización, diseñar la recolección de datos, delinear una trayectoria hipotética de aprendizaje que describa el resultado esperado del proceso de aprendizaje y el modo en que se va a promover y alcanzar dicho aprendizaje. Las acciones mencionadas, se describen en el apartado denominado "Diseño de tareas" (sección 3.3), en el cual se especifican todos los detalles en relación con la intencionalidad de las tareas y las especificaciones correspondientes a cada uno de los ítems

planteados en donde se establecen las características que se pretenden abordar en relación con la competencia interpretativa y los procesos de la génesis instrumental. Además, en los anexos se establecen los elementos reportados por Gómez et al. (2018) en los cuales se mencionan los requisitos y la temporalidad de cada una de las tareas, entre otros aspectos; esto resalta la importancia de elegir un modelo de tarea que permita aportar elementos al experimento de enseñanza que se está construyendo.

Lo anterior tiene como finalidad el poder abordar todas las acciones mencionadas en esta primera fase, teniendo en cuenta que la profundidad al aclarar cada uno de los asuntos, va en correspondencia con el propósito establecido en cada tarea.

- **Fase 2: experimentación para promover el aprendizaje**

En esta fase tienen lugar las intervenciones en el aula y las sucesivas iteraciones del ciclo de tres pasos:

1) *Diseño y formulación de hipótesis (antes de cada intervención)*

Camargo (2021) afirma que la conjetura es el motor que dirige el experimento y determina el curso de los acontecimientos, por ello algunos autores sugieren hacer un ejercicio preliminar que permita evidenciar su potencial funcionamiento. Por otro lado, dicha conjetura abarca dos dimensiones:

- ✓ Una, referida al contenido matemático en juego, que responde al interrogante sobre qué debería ser enseñado; en el caso de esta investigación, corresponde a las medidas de dispersión (varianza y desviación estándar).
- ✓ La otra dimensión está referida al aprendizaje, que responde al cómo debería ser enseñado. Para el caso de esta investigación, se hace alusión a las tareas mediadas por Microsoft Excel y las bases de datos, que tienen como propósito el promover las habilidades de la competencia interpretativa de las medidas estadísticas mencionadas.

Es así como se define que la conjetura que dirige este experimento de enseñanza es la siguiente:

El diseño e implementación de una secuencia de tareas de aprendizaje mediadas por Excel, con el uso de bases de datos reales, dirigida a estudiantes de grado noveno, promueve la interpretación de los conceptos de dispersión (varianza y desviación estándar).

2) *Intervención en el aula y recolección de datos (durante cada intervención)*

La propuesta consiste en diseñar e implementar una secuencia de tres tareas las cuales apuntan a promover diferentes asuntos y características en relación con la competencia interpretativa (lo cual se detalla en el apartado Diseño de tareas (sección 3.3)). Para realizar la recolección de la información, los estudiantes contarán en cada tarea con dos archivos: Uno en Microsoft Word con las respectivas indicaciones y preguntas que deben desarrollarse y otro en Microsoft Excel en el cual se relacionará la información correspondiente a una base de datos. Los archivos serán compartidos con el profesor al terminar la sesión o al culminar la tarea correspondiente.

3) *Análisis de los datos y revisión y reformulación de hipótesis (posterior a cada intervención)*

Posterior a la intervención realizada para la primera tarea, se tendrá en cuenta lo acontecido durante las sesiones de clase para realizar las modificaciones correspondientes a la tarea #2 y este mismo ejercicio se realizará previo a la implementación de la tarea #3. Pese a que no se realice una modificación a la conjetura, se espera que las modificaciones realizadas a las tareas permitan comprobar la efectividad de esta.

- **Fase 3: ejecución del análisis retrospectivo de los datos**

Molina et al. (2011) destacan que en esta etapa se tiene como finalidad recopilar y organizar toda la información recogida y analizar el conjunto de los datos. Para llevar a cabo lo dispuesto en esta última fase, se tendrá en cuenta la caracterización realizada en las categorías de análisis que están descritas en el subcapítulo 3.4. Dicho análisis se reporta con detalles en el Capítulo 4 denominado *Resultados*, en el que se especifica y presenta evidencia de las producciones de los estudiantes para cada una de las tareas.

Enseguida se comenta cómo fue el proceso para elaborar las tareas de aprendizaje diseñado, se reconoce en este elemento un componente sustancial de este capítulo relacionado con lo metodológico.

3.3. Diseño de las tareas

3.3.1. Tarea 1

A continuación, se describe el diseño de la Tarea 1, cuya versión completa se encuentra en el Anexo 1. Para su elaboración, se determinó inicialmente que el objetivo fuera que los estudiantes

interpretaran el valor numérico de las medidas de tendencia central y dispersión, las cuales ya sabían calcular debido a su dominio previo de los algoritmos. Es decir, el propósito fundamental radicaba en analizar la forma en que los estudiantes interpretaban dichas medidas, más allá de su aplicación mecánica.

Considerando el objetivo de la tarea, para el diseño se tomaron en cuenta algunas características de la interpretación definidas previamente en el marco teórico, entre las cuales se destacan las siguientes:

C2. Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013).
C6. Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.
C8. Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.
C9. Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado.
C11. Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas.
C13. Comparar resultados para diferentes condiciones en un experimento. (Franklin et al., 2020)
C14. Evaluar e interpretar el impacto de los valores atípicos en los resultados. (Franklin et al., 2020)
C15. Reconozca la presencia de valores faltantes y comprenda cómo los valores faltantes pueden agregarse sesgo a un análisis. (Franklin et al., 2020).

Aunque estas no fueron las únicas características trabajadas en la tarea, sí son las que, en la generalidad, guiaron en gran medida su diseño.

Esta Tarea 1 se estructuró en dos partes. La primera se enfocó en la interpretación de las medidas de dispersión tanto de una población como de una muestra, analizando además algunas características de estas medidas frente a la adición o modificación sutil de datos. La segunda parte está centrada en la elaboración de conjeturas sobre los valores numéricos de algunas medidas estadísticas cuando los datos tienen errores respecto a la toma, recolección o manipulación de estos.

En la primera parte, se plantearon cuatro subtareas: la primera se orientó hacia la interpretación que los estudiantes realizaban al valor numérico de las medidas, la segunda se centró en la interpretación de la variación de estas medidas al incorporar un dato "atípico", la tercera tuvo como objetivo comparar las medidas de dispersión en una población con los parámetros de una muestra relativamente pequeña y, finalmente, la cuarta subtarea, similar a la segunda, promovía la interpretación de la variación de los parámetros ante el impacto de un dato "atípico" en la muestra.

La segunda parte planteaba dos subtareas: la primera se enfocó en fomentar los procesos de interpretación y elaboración de conjeturas, considerando la interpretación de datos que presentaban errores debido a una manipulación incorrecta. La segunda, con un objetivo similar, buscaba promover los mismos procesos, pero tomando en cuenta la pérdida de datos ocasionada por factores externos, en este caso, por errores del software.

Con estos aspectos definidos, se procedió a buscar una base de datos que se ajustara a las características y propósitos de las subtareas y que estuviera relacionada con una temática o problemática del contexto de los estudiantes. Aunque inicialmente resultó difícil encontrar una base de datos adecuada, se identificó en la página gubernamental "Datos Abiertos Colombia" la base de datos titulada "Productos de licor registrados en el Departamento de Risaralda para su distribución"[9]. Esta base con 24.618 datos incluye información sobre los productos de licor registrados en el Departamento de Risaralda, con detalles sobre el registro INVIMA, la vigencia del registro, el origen (Nacional o Importado) y el grado de alcohol de los productos.

Aunque esta base de datos no está asociada en sí misma con el contexto social de los estudiantes, se decide adaptar la tarea de modo que se enfoque en la problemática del consumo de bebidas alcohólicas en adolescentes. Por otra parte, respecto a los fines de la actividad, la base de datos era especialmente conveniente pues tenía una variable cuantitativa (grado de alcohol) que permitía moldear los datos a las subtareas previamente definidas.

Para adaptar la base de datos, se organizaron y definieron las variables que podían ser útiles para el desarrollo de las tareas, buscando evitar que fuera demasiado extensa. Se decidió eliminar cuatro de las variables que no eran relevantes para la tarea (origen, registro sanitario, vigencia del registro y productor), dejando únicamente la variable principal (grado de alcohol) y aquellas variables que

[9] Base de datos original disponible en https://www.datos.gov.co/Hacienda-y-Cr-dito-P-blico/Productos-de-licor-registrados-en-el-Departamento-/he89-pbj8/about_data

se consideraba aportaban información al contexto o a la interpretación de los datos (Resolución de registro sanitario, Nombre empresa distribuidora y NIT). Luego, se eliminaron todos los registros con 0 grados de alcohol, con dos objetivos: el primero, evitar que el rango fuera igual al dato mayor, y el segundo, impedir que la media se viera excesivamente influenciada por alrededor de 1000 datos que presentaban esta condición.

Posteriormente, como en la parte uno de la tarea se pretendía trabajar con muestras de la base de datos, se toman dos muestras tomando como referencia todos los datos de dos distribuidoras de licores (ver Figura 14). No era necesario que las muestras fueran aleatorias ya que se quería que la cantidad de datos no fuera demasiado extensa para promover la caracterización e interpretación de la variación de las medidas de dispersión de una población y muestra considerando la inclusión de un dato atípico.

Figura 12. *Ejemplar de la Muestra 1 correspondiente a la base de datos Licores Risaralda.*

	A	B	C	D	E
1	**PRODUCTO**	**GRADO DE ALCOHOL**	**RESOLUCIÓN DE REGISTRO SANITARIO**	**NOMBRE EMPRESA DISTRIBUIDORA**	**NIT**
2	Ap. Crema De Whisky Casa Del Rhin	14	19,037,470	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
3	Ap. Crema Sab. A Pistacho	14	16,034,714	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
4	Ap. Crema Sabor A Whisky Kingston	14	19,016,829	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
5	Ap. Guarilaque Alicante Sab Hinojo 15°	15	20,046,389	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
6	Ap. No Vin. De Cafe Cafesillo	15	17,013,704	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
7	Ap. No Vin. Quinado Campeon	12	20,012,618	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
8	Ap. No. Vin. Cerezano De Cerezas	9,2	19,038,991	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
9	Ap. No. Vin. Cerezano De Cerezas	9,2	19,038,991	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
10	Ap. Sab. A Br. Arthur Alicante 8°	8	15,008,067	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
11	Ap. Sab. A Ron Borinquen 8°	8	20,020,826	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
12	Ap. V. Casa Del Rhin Blanco Para Misa	10,3	19,038,989	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
13	Ap. V. Gran Concilio Blanco Para Misa	9,3	19,038,990	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
14	Ap. Vin Gedeon Con Sabor A Quina	13	21,028,706	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
15	Ap. Vin. Gran Ducado Moscatel Tinto	8,5	19,036,270	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
16	Ap. Vin. Moscatel De Pasas	10,1	19,038,993	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
17	Ap. Vin. Moscatel De Pasas	10,1	19,038,993	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
18	Ap. Vin. Moscatel De Pasas	10,1	19,038,993	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
19	Beb. Emb. De Cereza Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
20	Beb. Emb. De Manzana Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
21	Beb. Emb. De Manzana Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
22	Beb. Emb. De Maracuya Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
23	Beb. Emb. De Maracuya Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250
24	Beb. Emb. De Melocoton Alicante	2	19,024,245	IMPORTACIONES Y ASESORIAS TROPI S.A.S	8,001,570,250

Licores Risaralda | Muestra 1 | Muestra 2 | Notas - Parte II

Para la segunda parte de la tarea, cuyo objetivo era fomentar los procesos de interpretación y elaboración de conjeturas ante datos que presentaban errores de manipulación o pérdida, y con el fin de justificar que en el manejo de datos pueden ocurrir errores, en este caso humanos por parte de un profesor, se generó una base de datos de 110 números aleatorios que buscaba simular el contexto (considerando que este facilitaría la interpretación de los datos) de las notas obtenidas por los estudiantes de la institución Tandem – Apoyo en Procesos Escolares en el examen final del tercer corte de matemáticas, correspondiente al semestre 2023-II. Aunque era factible construir la

segunda parte de la tarea utilizando la base de datos de la primera parte, siempre que se realizara una muestra, se optó por no hacerlo con el propósito de ofrecer un contexto más cercano y relevante para los estudiantes.

Por otra parte, teniendo en cuenta lo establecido para los procesos de instrumentación e instrumentalización en el marco de la génesis instrumental, enseguida se presentan las relaciones establecidas entre algunas de las preguntas planteadas en esta tarea y dicho marco teórico (Tabla 10),. Es importante mencionar que esto se fue contemplando durante la construcción de los enunciados e instrucciones de la tarea.

En el caso de las preguntas 1, 2, 7, 9, 11 y 13 no se pretende evidenciar ninguno de los procesos asociados a la génesis instrumental, puesto que dichas preguntas están relacionadas con las características que se busca abordar en relación con la competencia interpretativa y en las cuales no está involucrada la interacción con Microsoft Excel, sino la observación y el análisis del conjunto de datos.

Tabla 10. *Preguntas de la tarea 1 que no se relacionan con los procesos de la génesis instrumental*

ITEMS	DESCRIPCIÓN
1. Responda las siguientes preguntas: **a)** ¿Qué piensa usted sobre el consumo excesivo de alcohol? **b)** ¿Cuáles problemas de salud conoce usted que se pueden generar por el consumo excesivo de alcohol? **c)** ¿Cuáles problemáticas sociales considera usted que se generan por el consumo excesivo de alcohol? **d)** En promedio, ¿cuál cree usted que es el grado de alcohol de las bebidas alcohólicas que conoce?, ¿en qué considera que afecta el grado de alcohol a los licores y a las personas que los consumen?, ¿por qué?	Preguntas orientadas al reconocimiento de la problemática a través de las opiniones personales de los estudiantes.
2. Descargue la base de datos a través del enlace anterior y ubíquese en la hoja denominada "Licores Risaralda". Responda las siguientes preguntas. **a)** ¿Cuál es la población de estudio? **b)** ¿Cuántas variables estadísticas identifica?, ¿cuáles? **c)** Clasifique las variables que identificó en el literal anterior. **d)** ¿Con cuáles de las variables estadísticas tiene sentido (dentro del contexto de la base de datos) calcular las medidas de tendencia central y dispersión?, ¿con cuáles no? Justifique su respuesta.	Preguntas que van en correspondencia con los conceptos previos que se plantearon en los requisitos de la tareas (Anexo 1).
7. Responda las siguientes preguntas: **a)** Teniendo en cuenta toda la información suministrada, ¿cómo se relaciona la problemática del consumo de alcohol con la producción de las bebidas alcohólicas? **b)** ¿Qué significado puede dar a las medidas halladas para la variable "Grado de alcohol" (punto 2), en relación con el contexto de los licores y los problemas existentes frente al consumo de estas bebidas?	Preguntas enfocadas a la interpretación de las medidas de dispersión en relación con la problemática y la base de datos.

c) ¿Qué implicaciones tuvo el agregar el dato del licor adulterado (dato atípico) al conjunto de datos dado inicialmente (Pregunta 2)? ¿Qué sucedió con las medidas calculadas cuando se agregó el mismo dato a la muestra (Pregunta 4)? Escriba una conclusión.

d) Realice la lectura de la información suministrada por el Ministerio de Salud en relación con la problemática del consumo de bebidas alcohólicas en Colombia (Anexo). Teniendo en cuenta la información de la base de datos y la lectura realizada, ¿qué conclusiones puede establecer?

e) ¿Cuál es su postura frente a esta problemática?

9. Suponga que el profesor al transcribir las calificaciones al archivo Excel, cometió un error y escribió un cero adicional en 10 de las notas. Escriba una conjetura, en la primera fila de la Tabla 6, sobre los resultados que podría obtener al calcular las medidas allí descritas tras modificar los 10 datos.

Tabla 6

Eliminación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura				
Medida				

11. Suponga que se eliminan 20 datos de forma aleatoria. Escriba una conjetura, en la primera fila de la Tabla 7, en relación con los resultados que podría obtener tras dicha eliminación.

Tabla 7

Modificación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura				
Medida				

13. ¿Qué puede concluir en relación con las medidas, al hacer la modificación y eliminación al conjunto de datos suministrado?

Preguntas enfocadas en el proceso de realizar un ejercicio de conjeturacion y verificación de conjeturas en relación con las medidas de dispersión, bajo la condición de que se agregue y elimine un determinado números de datos del conjunto dado.

En los ítems 3, 4, 5, 6, 8, 10 y 12 de la Tarea 1 se busca promover el proceso de instrumentación, teniendo en cuenta que se requiere mayormente del uso de las funciones estadísticas en Excel para determinar las medidas solicitadas en cada caso (Tabla 11). Además, se espera que los estudiantes logren desarrollar esquemas de utilización relacionados con la sistematización de las funciones mencionadas al aplicarlas a los conjuntos de datos; es por ello que, en cada uno de los ítems que se presenta en la Tabla 13, se solicita a los estudiantes diligenciar un recuadro con algunas medidas estadísticas, lo cual tiene como implicación el uso de determinadas funciones de Excel que solo difieren en su estructura cuando se usa para calcular valores poblacionales o muestrales.

Tabla 11. *Preguntas de la tarea 1 que se relacionan con el proceso de instrumentación.*

ITEMS	DESCRIPCIÓN
3. La variable "Grado de alcohol" mide el porcentaje de alcohol que contiene cada bebida registrada en la base de datos. Calcule la media, el rango, la varianza y la desviación estándar de la variable Grado de alcohol, haciendo uso de los comandos de Excel y consigne las respuestas en la Tabla 1 (utilice las funciones que permiten calcular estas medidas para las poblaciones). Tabla 1 | | **Media** | **Rango** | **Varianza** | **Desviación estándar** | | *Población* | | | | | **a)** Según el Decreto 162 de 2021 una bebida alcohólica es un "producto apto para consumo humano que contiene una concentración no inferior a 2.5 grados alcoholimétricos". ¿Hay bebidas no alcohólicas en la base de datos? Justifique su respuesta. **b)** Teniendo en cuenta la media y la desviación estándar calculadas para los datos de la variable "Grado de alcohol", realice una descripción de dicho conjunto de datos. **c)** ¿Qué puede decir respecto a la dispersión de la variable "Grado de alcohol" de las bebidas presentadas en la base de datos?	Las preguntas 3, 4, 5 y 6 tienen una estructura similar y un objetivo en común, el cual corresponde a lograr afianzar y diferenciar las funciones que permiten calcular las medidas indicadas. Aunque los ítems cuentan con algunas preguntas que buscan fortalecer las características de la competencia interpretativa, el punto 3a permite promover otro tipo de interacción entre el estudiante y Excel relacionada con la exploración e identificación de un valor específico dentro de la base de datos que podría darse a través de la observación o el uso de nuevas funciones.
4. "Las intoxicaciones por bebidas alcohólicas adulteradas por metanol son un evento de interés en salud pública con vigilancia epidemiológica durante todo el año e intensificada durante la temporada decembrina" (Ministerio de Salud y Protección social, 2024). Teniendo en cuenta esta afirmación, suponga que existe un dato adicional en la base de datos que corresponde a un licor adulterado que se prepara a base de alcohol etílico (con un 96% de alcohol) y un saborizante artificial. Agregue este nuevo dato al final de las columnas Nombre y Grado de alcohol y calcule nuevamente las medidas de tendencia central y dispersión solicitadas considerando dicho dato. Consigne las respuestas en la Tabla 2. Tabla 2 | | **Media** | **Rango** | **Varianza** | **Desviación estándar** | | *Población con dato adicional* | | | | | **a)** ¿Qué representa el cambio en el rango de los datos al comparar las Tablas 1 y 2? **b)** Al comparar la información de la población y la población con el dato adicional de las Tablas 1 y 2, ¿qué diferencias o similitudes puede enunciar para cada medida calculada? **c)** ¿Por qué cree usted que las medidas de la población no cambian demasiado aun cuando se agrega un dato adicional mucho mayor a los demás?	
5. Ubíquese en la hoja denominada Muestra 1 o 2 (según indique el profesor), las cuales corresponden, como su nombre lo indica, a dos muestras distintas de la base de datos original obtenidas de forma aleatoria. Enseguida, calcule las medidas: media, rango, varianza y desviación estándar de la variable Grado de alcohol (utilice de nuevo las funciones para muestra en Excel). Consigne las respuestas en la Tabla 3. Tabla 3 | | **Media** | **Rango** | **Varianza** | **Desviación estándar** | | *Muestra* | | | | |	
6. Según el proyecto de Acuerdo 95 del 2014 presentado por Liliana Graciela Guaqueta De Diago, concejal de Bogotá: "Un estudio sobre consumo de alcohol en menores de edad en Colombia", elaborado por la Corporación Nuevos Rumbos con 9.276 alumnos de colegios públicos y privados del país, indica que muchos niños empiezan a ingerir licor a los 10 años y las niñas a los 11 años. No es para menos, si se tiene en cuenta que, de acuerdo con la misma investigación, niños y adolescentes, en general, ven normal y aceptable tomar trago para celebrar, socializar, vencer la timidez, lograr aceptación, conquistar a alguien y pasarla bien. Para la elaboración de un coctel peligroso cuyo consumo crece entre los jóvenes y alarma a las autoridades, son usados algunos ingredientes como: alcohol antiséptico, agua, refrescos en polvo, gaseosas, saborizantes y hasta leche condensada. Teniendo en cuenta lo anterior, suponga que existe un dato adicional, en la hoja Muestra 1 o 2 según haya sido asignado, que corresponde a un licor adulterado que se prepara a base de alcohol etílico (concentrado al 96%) y un saborizante artificial. Agregue este nuevo dato al final de las columnas Nombre y Grado de alcohol, y calcule nuevamente las medidas del numeral 4. Consigne las respuestas en la Tabla 4.	

Tabla 4

	Media	Rango	Varianza	Desviación estándar
Muestra con dato adicional				

a) ¿Qué diferencias o similitudes puede enunciar para cada medida, teniendo en cuenta la información de las tablas 3 y 4?

b) De acuerdo con lo observado, ¿de qué dependería poder asumir que la media y las medidas de dispersión obtenidas en la muestra sean aproximadamente iguales a las de la población? Explique.

8. Ubíquese en la hoja del archivo Excel denominado Notas - Parte II y teniendo en cuenta la lista de datos calcule la media, rango, varianza y desviación estándar. Registre los resultados obtenidos en la Tabla 5.

Tabla 5

Media	Rango	Varianza	Desviación estándar

10. Modifique aleatoriamente los 10 datos teniendo en cuenta el error cometido por el profesor y calcule la media, rango, varianza, desviación estándar. Registre los resultados en la segunda fila de la Tabla 6.

a) ¿Qué diferencias o similitudes existe entre la conjetura y las medidas obtenidas? Comente por qué considera que surgen estas diferencias o similitudes.

12. Proceda a eliminar 20 datos de forma aleatoria, a continuación, calcule la media, rango, varianza, desviación estándar y registre los resultados en la segunda fila de la Tabla 7.

a) ¿Qué diferencias o similitudes existe entre las conjeturas y las medidas obtenidas? Describa y explique por qué surgen estas diferencias o similitudes.

Las preguntas 8, 10 y 12 nuevamente se caracterizan por propiciar el uso de las funciones estadísticas en Excel.
Adicionalmente, la modificación y eliminación de datos permite que el estudiante trabaje en la adaptación y trasformación del conjunto de datos a partir del uso de las herramientas de Excel (borrar un grupo de valores e ingresar nuevos datos).

Por otra parte, se considera en principio que el proceso de instrumentalización no está presente en el planteamiento de esta tarea, teniendo en cuenta que el papel principal de la herramienta digital no corresponde directamente a la asignación de nuevos atributos o transformaciones a las funciones de Microsoft Excel. Sin embargo, esto no implica que durante la implementación de la tarea no pueda emerger el proceso de instrumentalización, ni que algunos ítems no puedan reflejar el proceso esperado.

Teniendo en cuenta la información presentada anteriormente, a continuación, se sintetiza en las Tablas 12 y 13 la relación entre cada ítem de la tarea 1, las características de las categorías de análisis y los procesos de la génesis instrumental, respectivamente.

Tabla 12. *Clasificación de las preguntas de la tarea 1 en relación con características de la competencia interpretativa..*

Ítem	N/A	C2	C6	C8	C9	C11	C13	C14	C15
1	X								
2	X								
3	X								
3(A, B, C)		X	X		X				
4	X								
4(A, B, C)			X		X		X		

5	X							
6	X							
6(A, B, C)				X		X		
7	X							
7(A, B, C)				X		X		
8	X							
9			X				X	X
10	X							
10A			X		X		X	
11			X				X	X
12	X							
12A					X			X
13		X				X		

Tabla 13. *Clasificación de las preguntas de la tarea 1en relación con procesos de la génesis instrumental.*

Ítem	Ninguna	Instrumentación	Instrumentalización
1	X		
2	X		
3		X	
4		X	
5		X	
6		X	
7	X		
8		X	
9	X		
10		X	
11	X		
12		X	
13	X		

3.3.2. Tarea 2

A continuación, se describe el diseño de la Tarea 2, cuya versión completa se encuentra en el Anexo 2. Para su elaboración, teniendo en cuenta que su objetivo es la construcción de la fórmula de la desviación estándar a partir de una situación contextualizada, se tomaron en cuenta algunas características de la interpretación definidas previamente en el marco teórico, entre las cuales se destacan las siguientes:

C3. Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013)

C6. Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.
C7. Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados.
C8. Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.

Si bien es cierto que estas no fueron las únicas características trabajadas, sí son las que, en la generalidad, guiaron en gran medida el diseño de la tarea.

En primer lugar, se llevó a cabo la búsqueda de una base de datos que incluyera una variable cuya unidad de medida resultara familiar para los estudiantes. Asimismo, se consideró que, al elevar dicha variable al cuadrado, su unidad de medida no fuera interpretable en términos físicos, con el propósito de justificar el proceso que permite la transición de la varianza a la desviación estándar.

Con estos aspectos definidos, se procedió a buscar una base de datos que se ajustara a las características y propósitos de la tarea, y al igual que en la tarea anterior, que estuviera en un contexto cercano al de los estudiantes. En este caso se identificó en la página gubernamental "Datos Abiertos" la base de datos titulada "Velocidad y tiempo de viaje GT"[10]. Esta base de 682.503 datos incluye información corresponde a los datos recopilados por el Sistema Inteligente de Movilidad de Medellín (SIMM), mediante los dispositivos de Cámaras de Circuito cerrado de Televisión (CCTV) y Apoyo a la Red Semafórica (ARS), por lo que los datos corresponden a velocidades puntuales en el momento de paso por las cámaras.

La elección de esta base de datos se fundamentó, en primer lugar, en el contenido de variables cuantitativas con unidades de medida familiares para los estudiantes, particularmente la variable "VELOCIDAD_KM/H" que cumplía con los criterios previamente establecidos y fue la que se utilizó para el diseño de la tarea. En segundo lugar, esta base de datos ofrecía la oportunidad de generar discusiones en torno a los problemas de movilidad, específicamente en Bogotá, ciudad en la que residen los estudiantes.

[10] Base de datos original disponible en https://www.datos.gov.co/Transporte/Velocidad-y-tiempo-de-viaje-GT/bb66-tm7q/about_data

Para adaptar la base de datos, se decidió eliminar ciertas variables que se consideraron de poca relevancia y que hacían la base de datos demasiado extensa. Asimismo, se seleccionaron las variables que resultaban más convenientes y menos propensas a generar confusión entre los estudiantes, como resultado, se eliminaron las siguientes variables: TV_[S], CORREDOR, DÍA, MES, AÑO, HORA, DIASEM y NOMBRE DÍA. Las variables que quedaron finalmente en la base de datos fueron: Flujo_Tipo, SENTIDO, INICIO, FIN, LONGITUD_KM, TV_MINUTOS, VELOCIDAD_KM/H, TV_MIN/KM, lo anterior se puede observar en la Figura 15.

Figura 13. *Ejemplar de la base de datos Velocidad y tiempo de viaje GT.*

Flujo_tipo	SENTIDO	INICIO	FIN	LONGITUD_KM	TV_MINUTOS	VELOCIDAD_KM/H	TV_MIN/KM
Continuo	SN	Tugo	San Juan	6,6	6,3	63	1,0
Continuo	NS	CTI	Calle 10	6,5	8,3	47	1,3
Continuo	NS	Solla	Punto Cero	5,3	8,3	39	1,6
Continuo	SN	Punto Cero	Solla	5,3	6,7	48	1,3
Interrumpido	NS	Calle 65	Calle 30	5,1	11,6	26	2,3
Interrumpido	SN	Calle 30	Calle 65	5,1	10,7	29	2,1
Continuo	NS	Sandiego	Chuscalito	5,9	8,3	43	1,4
Continuo	SN	Chuscalito	Sandiego	6	7,4	49	1,2
Interrumpido	NS	Calle 30	Calle 12Sur	3,8	8,7	26	2,3
Interrumpido	SN	Calle 12Sur	Calle 30	3,8	7,6	30	2,0
Interrumpido	NS	Calle 30	Calle 12Sur	4,1	6,5	38	1,6
Interrumpido	SN	Calle 12Sur	Calle 30	4,1	6,9	35	1,7
Interrumpido	EW	Autopista Sur	Avenida 80	2,5	4,3	35	1,7
Interrumpido	WE	Avenida 80	Autopista Sur	2,6	4,8	33	1,8
Interrumpido	WE	Avenida 80	Industriales	2,9	7,1	24	2,5
Interrumpido	EW	Industriales	Avenida 80	2,9	7,3	24	2,5
Interrumpido	WE	Avenida 80	Avenida Regiona	2,5	5,8	26	2,3
Interrumpido	EW	Avenida Regional	Avenida 80	2,5	5,8	26	2,3

Velocidad y tiempo de viaje GT +

Posteriormente, debido a la extensión de la base de datos, se optó por tomar una muestra aleatoria de 607 datos. Para lograrlo, se utilizó una función de Excel para asignar un número aleatorio a cada dato, luego se seleccionaron únicamente aquellos datos cuyo número aleatorio correspondía a un número determinado, mientras que el resto de los datos fueron eliminados.

Al igual que como se presentó en la Tarea 1, también se estableció la relación entre cada uno de los puntos de esta segunda tarea con los procesos de instrumentación e instrumentalización. Es importante indicar que en la construcción de las preguntas se pretendía favorecer el proceso de instrumentalización, dado que no se abordó en la primera tarea.

Nuevamente, se identifica que algunas preguntas no están relacionadas en principio con ninguno de los dos procesos en cuestión, puesto que el objetivo de los ítems 1, 2, 5, 6, 7, 10, 12, 13, 15 y 16 es fortalecer lo correspondiente a las habilidades propias de la competencia interpretativa como se muestra en la Tabla 14.

Tabla 14. *Preguntas de la tarea 2 que no se relacionan con los procesos de la génesis instrumental*

ITEMS	DESCRIPCIÓN
1. Responda las siguientes preguntas: **a)** ¿Qué piensa usted acerca de la movilidad en Bogotá? **b)** ¿Cree pertinente y oportuno el planteamiento e implementación de estrategias que permitan mejorar la movilidad en Bogotá? ¿Por qué? **c)** ¿Cuáles estrategias implementaría usted para mejorar la movilidad en Bogotá? ¿Por qué?	Preguntas orientadas al reconocimiento de la problemática a través de las opiniones personales de los estudiantes.
2. Descargue la base de datos a través del enlace anterior y determine cuál es la población, identifique y reporte todas las variables estadísticas que se presentan y luego clasifíquelas.	Preguntas que van en correspondencia con los conceptos previos que se plantearon en los requisitos de la tareas (Anexo 2).
5. ¿Cuáles son las unidades de medida correspondientes a los datos obtenidos en el punto anterior? **6.** Es posible establecer una interpretación para las magnitudes obtenidas en el punto 4, ¿Qué podría establecer al respecto? ¿cuál es el significado de estos valores numéricos teniendo en cuenta el contexto dado? **7.** Teniendo en cuenta las respuestas dadas en los ítems 5 y 6, ¿qué interpretación les da a los signos de los valores obtenidos dentro del contexto establecido? ¿Por qué? Justifique su respuesta. **10.** ¿Cuáles son las unidades de medidas correspondientes al resultado obtenidos en el punto anterior? Mencione si es coherente respecto a la situación planteada y justifique. **12.** ¿Qué representa el valor obtenido en relación con la situación propuesta?	Preguntas enfocadas a la interpretación de las medidas de dispersión en relación con la problemática, la base de datos y las unidades de medida.
13. El Sistema Inteligente de Movilidad de Medellín (SIMM) planea llevar a cabo un estudio similar el próximo año, sin embargo, con el objetivo de simplificar el proceso, se busca desarrollar una fórmula que sea aplicable independientemente de los datos recopilados. Los pasos para realizar este trabajo se muestran a continuación, completa la tabla para finalizar la deducción de la fórmula deseada por el Sistema Inteligente de Movilidad de Medellín (SIMM). **Pasos** / **Formula planteada** ***a)*** Si se tiene en cuenta que la media es $\bar{x}$, se resta la media a cada valor (x_i) del conjunto de datos. / $x_i - \bar{x}$ ***b)*** Se eleva al cuadrado cada diferencia. / ***c)*** Se suman todas las diferencias al cuadrado. / $\sum_{i=1}^{n}(x_i - \bar{x})^2$ ***d)*** Se divide la suma obtenida entre el número total de datos (n). / ***e)*** Por último, se calcula la raíz cuadrada del paso anterior. /	Pregunta relacionada con el uso del lenguaje matemático en la formulación de la varianza y la desviación estándar (Aunque este ítem no está directamente vinculado a los procesos de la génesis instrumental, existe una conexión con algunos ítems anteriores (los cuales son necesarios para resolver este ítem) en los que se emplean dichos procesos).
15. Al usar los comandos de varianza y desviación, ¿los valores numéricos obtenidos son iguales a los calculados con la fórmula?	Pregunta relacionada con la validación de resultados.
16. ¿Qué representa en el contexto de la situación los valores obtenidos en ítem 13?	Pregunta de interpretación de medidas estadísticas.

Por otra parte, en los ítems 3 y 14 se hace presente el proceso de instrumentación (ver Tabla 15), teniendo en cuenta que en la Tarea 1 se profundizó en el esquema de utilización relacionado con

sistematizar el uso de las funciones de Excel, estos dos ítems permitirán evidenciar si efectivamente los estudiantes lograron desarrollar el esquema mencionado. Por otra parte, se propicia el uso de funciones que involucran operaciones matemáticas para el desarrollo de los cálculos correspondiente a la varianza y desviación estándar sin usar directamente las funciones estadísticas, lo cual tiene como finalidad promover otro esquema de utilización, que corresponde a aplicar los algoritmos e identificar la rapidez y eficiencia de Excel a la hora de realizar cálculos que involucran un alto volumen de datos.

Tabla 15. *Preguntas de la tarea 2 que se relacionan con el proceso de instrumentación.*

ITEMS	DESCRIPCIÓN
3. Considerando los datos de la variable estadística "*VELOCIDAD KM/H*", calcule su media y responda: ¿cuáles son las unidades de medida de la media? Para usted, ¿qué significado tiene la media calculada dentro del contexto de la base de datos?	Preguntas que pretende fortalecer el esquema de utilización relacionado con el uso de las funciones en Excel (estadísticas y matemáticas).
14. Calcule la varianza y la desviación estándar de la variable estadística "LONGITUD_KM" usando la fórmula establecida en el ítem 12 (desarrolle el procedimiento en Excel).	

En los ítems 4, 8, 9 y 11 se apunta al proceso de instrumentalización, puesto que la propuesta de estas preguntas es propiciar que los estudiantes interactúen con Microsoft Excel de tal manera que puedan descubrir atributos a partir del uso de las funciones que ofrece dicha herramienta digital e identificar cuáles de estas les permite ejecutar las operaciones que dan solución a las cuestiones que se proponen en la tarea (ver Tabla 16).

Es importante precisar que para el caso de los ítems 4, 8 y 11, se espera que los estudiantes realicen una resta y eleven los valores al cuadrado, respectivamente; sin embargo, podría ocurrir que ejecuten operaciones diferentes a las mencionadas. Para los casos en los que esto se dé, se realizarán algunas preguntas orientadoras que den algunas pistas o indicios que permitan direccionar sus resultados, puesto que esto permitirá la correcta construcción de las fórmulas establecidas para la varianza y la desviación estándar.

Tabla 16. *Preguntas de la tarea 2 que se relacionan con el proceso de instrumentalización.*

ITEMS	DESCRIPCIÓN
4. El Sistema Inteligente de Movilidad de Medellín (SIMM) está interesado en determinar si los conductores mantienen velocidades cercanas al promedio, lo que sugeriría una ausencia de cambios bruscos en la velocidad, o si, por el contrario, hay fluctuaciones de velocidad muy marcadas. ¿Cómo se puede determinar si **cada conductor** mantiene velocidades cercanas al promedio? Proponga y justifique un procedimiento matemático.	

8. Si hay datos para los cuales el signo de los valores numéricos no permita una interpretación adecuada y acorde al contexto, utilice para el conjunto de datos alguna operación o propiedad matemática para asegurarse de que los signos estén en coherencia con la situación planteada. Mencione que propiedad u operación matemática utilizó.	Preguntas que pretenden relacionar los conocimientos previos de los estudiantes en relación con el lenguaje y las operaciones matemáticas, y las funciones de Excel.
11. A partir del valor numérico obtenido (punto 9) junto con las unidades de medida reportadas en el punto 10 ¿Es posible establecer una interpretación adecuada y que corresponda al contexto planteada? De no ser así, utilice alguna operación o propiedad matemática para asegurar que las unidades de medida se puedan interpretar en correspondencia con la situación planteada. Mencione que propiedad u operación matemática utilizó y porqué.	
9. El Sistema Inteligente de Movilidad de Medellín (SIMM) quiere presentar un informe para los habitantes de Medellín y, dentro de este, desea sintetizar la información obtenida sobre la variable km/h en el punto 6, para presentarla como **un solo valor numérico**. ¿Qué procedimiento realizaría para poder sintetizar esta información? Realice este proceso en Excel.	Pregunta en la cual el estudiante tiene la posibilidad de ejecutar dos procedimientos a través de las funciones de Excel: uso de la función media o uso de fórmula para calcular la media.

Teniendo en cuenta la información presentada anteriormente, a continuación, se sintetiza en las Tablas 17 y 18 la relación entre cada ítem de la tarea 2, las características de las categorías de análisis y los procesos de la génesis instrumental, respectivamente.

Tabla 17. *Clasificación de las preguntas de la tarea 2 en relación con características de la competencia interpretativa.*

Ítem	N/A	C3	C6	C7	C8
1	X				
2	X				
3			X		
4				X	
5		X	X		
6				X	
7			X		X
8	X				
9				X	X
10			X		X
11		X			X
12			X		
13		X			
14	X				
15			X	X	
16			X		

Tabla 18. *Clasificación de las preguntas de la tarea 2 en relación con procesos de la génesis instrumental.*

Ítem	Ninguna	Instrumentación	Instrumentalización
1	X		
2	X		
3		X	
4			X
5	X		
6	X		
7	X		
8			X
9			X
10	X		
11			X
12	X		
13	X		
14		X	
15	X		
16	X		

3.3.3. Tarea 3

A continuación, se presenta el diseño de la tarea, cuya versión completa se encuentra en el Anexo 3. Para su elaboración inicial, se definieron dos objetivos principales: el primero, que los estudiantes analizaran la dispersión en gráficos estadísticos y realizaran comparaciones entre sus potencialidades y limitaciones; el segundo, que trabajaran en el reconocimiento e identificación de ciertas características de las medidas de dispersión, permitiéndoles comparar la dispersión entre variables con diferentes unidades de medida, en particular mediante el uso del coeficiente de variación.

Considerando el objetivo de la tarea, para el diseño se tomaron en cuenta algunas características de la interpretación definidas previamente en el marco teórico, entre las cuales se destacan las siguientes:

C2. Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013)

C3. Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013)

C5. Interpretar información presentada en forma gráfica o diagramas

C6. Interpretar un resultado matemático en el contexto del mundo real

C8. Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.

C13. Comparar resultados para diferentes condiciones en un experimento. (Franklin et al., 2020)

C14. Evaluar e interpretar el impacto de los valores atípicos en los resultados. (Franklin et al., 2020)

Aunque estas no fueron las únicas características abordadas, sí fueron las que, en términos generales, orientaron en gran medida el diseño de la tarea.

En función de los objetivos de la tarea, esta se estructuró en dos secciones, la primera se centró en la interpretación de las medidas de dispersión en gráficos estadísticos, por lo que se decidió generar una base de datos que reflejara características específicas de dispersión y la presencia de datos atípicos, utilizando un contexto familiar para los estudiantes, sus calificaciones. Se optó por generar diagramas de cajas y gráficos de dispersión, pues se considera que son particularmente adecuados para interpretar las medidas de dispersión. Estos gráficos presentan diversas limitaciones y posibilidades que, en conjunto, pueden maximizar la capacidad de interpretación, análisis y comprensión de los datos.

Para generar la base de datos, se elaboraron tres conjuntos de datos, cada uno con características distintas para cumplir con los objetivos de la tarea. El primer conjunto de datos se generó con valores aleatorios y una alta dispersión a lo largo de todo el intervalo posible de calificaciones, comprendido entre 1,0 y 10,0. En el segundo conjunto, los datos se distribuyeron en un rango más reducido, ajustando algunos valores para que los extremos del diagrama de cajas y el gráfico de dispersión evidenciaran la presencia de observaciones que se desviaban significativamente de la tendencia central. Finalmente, el tercer conjunto de datos se generó con una variabilidad considerablemente menor en comparación con los dos conjuntos anteriores, y posteriormente se introdujeron observaciones atípicas para aumentar la heterogeneidad de los datos y, en particular, para que dichas observaciones atípicas fueran claramente evidentes en los diagramas de cajas y dispersión. Lo descrito anteriormente se puede observar en la Figura 16.

Figura 14. *Ejemplar de la base de datos Notas de los Exámenes 2023-II.*

NOTAS DE LOS EXÁMENES 2023-II DE LOS ESTUDIANTES DE TANDEM		
NOTAS EXÁMENES FINALES MATH CORTE I	NOTAS EXÁMENES FINALES MATH CORTE II	NOTAS EXÁMENES FINALES MATH CORTE III
8,8	7,5	4,5
6,8	4,2	5
7,7	7,5	4,2
8,6	6,1	3,4
3,6	8	4,5
4	3,7	4,9
5,4	8,6	4,1
5,5	8,4	3,4
4,9	8,6	4,7
2	6,2	4,9
1,8	8,9	3,1
9,6	8	3,3
5,2	5,7	4,5
4,7	1	4,5

Notas Calidad del agua +

La segunda sección se centró en la definición, aplicación e interpretación del coeficiente de variación. Para ello, se buscó una base de datos que incluyera diferentes variables con unidades de medida distintas pero pertenecientes a una misma población. En el momento del planteamiento de la tarea, el contexto del racionamiento de agua en Bogotá y sus alrededores fue muy relevante, por lo que se decidió buscar una base de datos relacionada con la calidad o características del agua. Tras una exhaustiva búsqueda, se encontró en la plataforma gubernamental "Datos Abiertos" la base de datos titulada "Calidad del agua del Río Cauca"[11]. Esta base de datos de 2.255 datos proporciona información sobre la calidad del agua en el Río Cauca, recopilada en 19 estaciones de muestreo e incluye aproximadamente 50 parámetros de calidad del agua, tanto fisicoquímicos, microbiológicos como biológicos y abarca datos desde el año 1990.

Para adaptar la base de datos, teniendo en cuenta que era excesivamente grande y la cantidad de variables, se decidió eliminar aquellas que no tenían muchas mediciones, es decir que tenían casillas en blanco dentro de la base de datos, eliminando así 25 de 58 variables. Posteriormente, debido a la extensión de la base de datos se optó por tomar una muestra de 1363 datos (ver Figura 17), para lograrlo, se realizó el mismo procedimiento que con la base de datos de la primera tarea, garantizando que la muestra fuera aleatoria.

[11] Base de datos original en https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/Calidad-del-agua-del-Rio-Cauca/d3ft-wu2b/about_data

Figura 15. *Ejemplar de la base de datos Calidad del Agua.*

FECHA DE MUESTREO	ESTACIONES	pH	TEMPERATURA (°C)	COLOR (UPC)	TURBIEDAD (UNT)	SOLIDOS TOTALES (mg SST/l)	SOLIDOS SUSPENDIDOS TOTALES	SOLIDOS DISUELTOS (mg SD/l)	DEMANDA BIOQUIMICA DE OXIGENO (mg O2/l)	DEMANDA QUIMICA DE OXIGENO	OXIGENO DISUELTO (mg O2/l)
1993-06-03T00:00:00,000	ANTES SUAREZ	6,6	20	100	40	84	28	56	0,6	19,8	5,7
1993-07-03T00:00:00,000	ANTES RIO OVEJAS	6,6	20,5	100	33	63	15,5	47,5	0,8	14,3	6
1993-06-02T00:00:00,000	ANTES RIO TIMBA	6,4	19	100	85	234	140	94	3,7	6,8	6,5
1993-07-02T00:00:00,000	PASO DE LA BALSA	7,1	20	100	48	128	72	56	2	11	6,9
1993-07-03T00:00:00,000	PASO DE LA BOLSA	6,8	23	100	47	126	46	80	3,5	26,4	6,2
1993-07-02T00:00:00,000	PUENTE HORMIGUERO	6,9	20	100	75	225	169	56	2,5	13,2	6,1
1993-08-18T00:00:00,000	ANTES INTERCEPTOR SUR	6,6	0	150	65	248	75	173	15,1	86,8	N,D,
1993-06-10T00:00:00,000	JUANCHITO	7	21	200	100	242	119	123	1,35	40	5,3
1993-06-10T00:00:00,000	PASO DEL COMERCIO	6,8	21	100	30	202	113	89	1,52	46,1	4,8
1993-06-09T00:00:00,000	PASO DE LA TORRE	7,2	23,8	20	43	174	58	116	3	19,8	2,3
1993-06-10T00:00:00,000	MEDIACANOA	7	24	0	70	336	230	106	1,5	108,2	0,5
1993-06-23T00:00:00,000	PUENTE GUAYABAL	6,8	22	0	60	301	118	183	3,9	61,6	2,1
1993-06-23T00:00:00,000	LA VICTORIA	6,9	23	0	60	237	116	121	4,6	84,7	2
1993-06-23T00:00:00,000	ANACARO	6,4	21	0	60	364	256	108	13,9	88,5	2,1
1993-06-23T00:00:00,000	PUENTE LA VIRGINIA	6,9	22	0	57	228	114	114	2,4	65,4	2,9
1993-09-02T00:00:00,000	ANTES SUAREZ	6,8	20,2	25	2,3	95	2,6	92,4	2,7	12,4	2,8

Notas Calidad del agua +

Siguiendo la misma estructura utilizada en el diseño de las Tareas 1 y 2, a continuación, se expone la relación entre cada uno de los puntos de la tarea y los procesos de la génesis instrumental. Sin embargo, es importante mencionar que, aunque para la construcción de esta tarea se usó Microsoft Excel, no se saca provecho de todas las funciones que ofrece, puesto que su papel protagónico está enfocado en permitir mostrar a los estudiantes diversas representaciones gráficas de un conjunto de datos. Lo anterior, deja abierta la posibilidad de que futuros trabajos puedan incluir en esta tarea elementos que den mayor relación entre los ítems planteados y los procesos de la génesis instrumental.

Nuevamente se tiene un conjunto de preguntas, que corresponden a los ítems 1 al 12, que no tienen como propósito inicial el desarrollo de los procesos de instrumentación e instrumentalización, más bien tienen como finalidad el poder abordar las características de la competencia interpretativa (ver Tabla 19). Es importante mencionar que en esta ocasión se incluyen elementos relacionados con gráficos estadísticos y el coeficiente de variación, que no se habían abordado en las tareas previas, con esto se pretende dar otra perspectiva al uso e interpretación de las medidas de dispersión.

Tabla 19. *Preguntas de la Tarea 3 que no se relacionan con los procesos de la génesis instrumental*

ITEMS	DESCRIPCIÓN
	Pregunta en correspondencia con los conceptos previos que se plantearon en los requisitos de la tareas (Anexo 3).

1. Con base en los datos que aparecen en la hoja denominada "Notas" del archivo de Excel descargado, responda las siguientes preguntas.

a) ¿Cuál es la población de estudio?

b) ¿Cuáles variables estadísticas se están representando?

c) Clasifique las variables que identificó en el literal anterior.

2. Considerando el diagrama de caja y el diagrama de dispersión presentados en la base de datos, ¿cómo describiría el desempeño de los estudiantes durante cada uno de los tres cortes?

Corte I	Corte II	Corte III

3. ¿Durante cuál de los cortes los estudiantes fueron más constantes con sus calificaciones? Justifique su respuesta.

4. ¿Cómo varió la dispersión de las calificaciones entre el corte 1 y el corte 2? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.

5. ¿Cómo varió la dispersión de las calificaciones entre el corte 2 y el corte 3? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.

6. ¿Cuántas y cuáles variables considera usted que influyen en las calificaciones de los estudiantes? ¿Estas variables cambian durante los tres cortes o se mantienen constantes? ¿Por qué piensa eso?

Preguntas relacionadas con la interpretación de dispersión a partir de gráficos estadísticos. (barras y de dispersión)

7. Teniendo en cuenta el contexto, que representaría un dato atípico dentro de este.

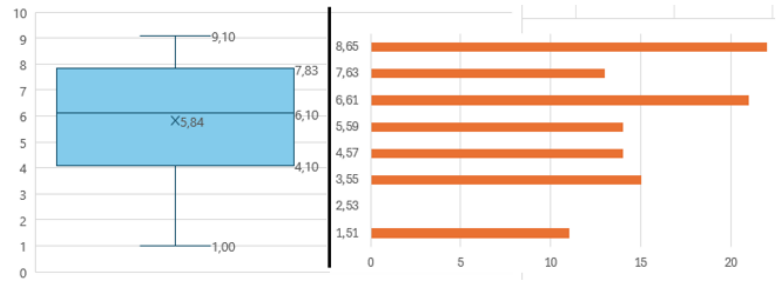
8. Según el diagrama de cajas, ¿hay datos atípicos en alguna de las cajas? ¿Para cuál o cuáles cortes?

9. ¿Como interpreta los datos atípicos encontrados respecto al corte correspondiente?

10. ¿Como interpreta estos mismos datos respecto a las calificaciones de los tres cortes?

Preguntas relacionadas con la interpretación del concepto dato atípico presentado en un gráfico estadístico.

11. El siguiente grafico muestra un paralelo entre el diagrama de caja y el histograma correspondientes a las calificaciones del corte II.



a. ¿Qué información se puede visualizar en el diagrama de cajas que no es posible ver en el histograma? ¿Cree usted que al observar e interpretar únicamente el diagrama de cajas se está ignorando alguna información del conjunto de datos?

b. ¿Qué información se puede visualizar en el histograma que no es posible ver en el diagrama de caja? ¿Cree usted que al observar e interpretar únicamente el histograma se está ignorando alguna información del conjunto de datos?

c. ¿Qué beneficios o dificultades puede traer observar simultáneamente e interpretar varios diagramas del mismo conjunto de datos?

Preguntas relacionadas con la interpretación de dispersión a partir de gráficos estadísticos (barras y diagrama de caja).

12. A continuación, se aborda una problemática que se ha venido presentando en los últimos años debido a diversos factores ambientales, políticos, etc. Teniendo en cuenta noticias, comentarios o información a la que haya tenido acceso, **sobre la calidad del agua potable**, responda las siguientes preguntas:

a. Según sus concepciones e ideas, ¿a qué hace referencia el concepto "calidad del agua"?

b. ¿Por qué cree que es importante atender los temas relacionados con la "calidad del agua"?

c. ¿Cuáles son las principales fuentes de contaminación de los embalses y cómo han afectado los parámetros de calidad del agua en los últimos diez años?

d. ¿Cómo podría comprobar la veracidad de la información suministrada en la respuesta a la pregunta anterior?

Preguntas orientadas al reconocimiento de la problemática a través de las opiniones personales de los estudiantes.

Se destaca que de manera parcial se fortalecen aspectos relacionados con el esquema de utilización que pretende promover el uso de los gráficos estadísticos para la visualización de la información aun cuando no son los estudiantes quieren los generan con las herramientas de Excel. Teniendo en cuenta lo mencionado anteriormente, no se consideró la pertinencia de categorizarlos en el siguiente apartado.

En los ítems 13 y 14 se busca apuntar al proceso de instrumentación y se tiene por objetivo seguir fortaleciendo el uso de las funciones estadísticas, que es uno de los esquemas de utilización que se propuso en la presente investigación (ver Tabla 20); adicionalmente, los estudiantes deben hacer uso de la función matemática para hacer divisiones (en los casos en que debe calcular los coeficientes de variación).

Tabla 20. *Preguntas de la Tarea 3 que se relacionan con el proceso de instrumentación*

ITEMS	DESCRIPCIÓN
13. La secretaría de medio ambiente desea realizar una comparación entre las variables PH y TEMPERATURA. Calcule la media y la desviación estándar para las variables estadísticas mencionadas y registre la información en la siguiente tabla: **Nota:** Escriba el valor numérico junto con las unidades de medida. **a.** La secretaría de medio ambiente encuentra que no es coherente comparar la desviación estándar entre PH y temperatura ¿Por qué cree que la secretaria del medio ambiente afirma esto? **b.** Calcule el coeficiente de variación para las variables estadísticas. **c.** ¿Cuál es la unidad de medida obtenida en ambos casos al aplicar la fórmula correspondiente? Justifique su respuesta. **d.** En general, ¿cuál es la unidad de medida correspondiente al coeficiente de variación? **e.** Es posible establecer una comparación entre las dos variables estadísticas a partir del valor numérico obtenido correspondiente al coeficiente de variación? ¿por qué? **f.** ¿Teniendo en cuenta la respuesta anterior, podría establecer cuál de las variables estadísticas presenta una mayor dispersión?	Las preguntas se caracterizan por propiciar el uso de las funciones estadísticas en Excel (para todos los casos en que se debe completar la información de la tabla). Adicionalmente, los incisos de cada ítem buscan generar interpretaciones de las medidas obtenidas, tal y como se propuso en algunos de los ejercicios de las Tareas 1 y 2.
14. Calcule los coeficientes de variación de siguientes variables estadísticas **a.** Realice una conjetura sobre cuál es el mínimo y el máximo valor que puede tomar el coeficiente de variación **b.** ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el mínimo valor? **c.** ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el máximo valor?	

	Media	Desviación estándar
PH		
Temperatura		

	Coeficiente de variación
PH	
Temperatura	

Variables	Coeficiente de variación
Solidos Totales	
Solidos Disueltos	
Demanda Bioquímica de Oxigeno	
Demanda química de Oxigeno	
Oxígeno Disuelto	
Dureza Total	
Dureza Cálcica	
Dureza Magnésica	
Conductividad Eléctrica	

Por otra parte, al igual que en la Tarea 1, en esta tarea no está presente en principio el proceso de instrumentalización, puesto que el papel principal de la herramienta Microsoft Excel es permitir al estudiante visualizar los diferentes gráficos estadísticos (los cuales están previamente construidos); sin embargo, no se descarta del todo que pueda darse este proceso durante la resolución de la tarea, puesto que el estudiante podría usar la exploración (borrar o agregar datos) con el fin de caracterizar el conjunto de datos en relación con los cambios que podría representar el gráfico estadístico suministrado.

Teniendo en cuenta la información presentada anteriormente, a continuación, se sintetiza en las Tablas 21 y 22 la relación entre cada ítem de la tarea 3, las características de las categorías de análisis y los procesos de la génesis instrumental, respectivamente.

Tabla 21. *Clasificación de las preguntas de la tarea 3 en relación con las características de la competencia interpretativa.*

Ítems	N/A	C2	C3	C5	C6	C8	C13	C14
1	X							
2			X	X				
3			X	X	X			
4			X	X				
5			X	X				
6						X		
7								X
8			X	X				
9			X	X				
10			X	X				
11 (A, B, C, D)			X	X				
12	X							
13	X							
13(C, D, E, F)		X			X		X	
14 (A, B, C)					X			

Tabla 22. *Clasificación de las preguntas de la tarea 3 en relación con los procesos de la génesis instrumental.*

Ítem	Ninguna	Instrumentación	Instrumentalización
1	X		
2	X		
3	X		
4	X		
5	X		
6	X		
7	X		

8	X		
9	X		
10	X		
11	X		
12	X		
13		X	
14		X	

3.4. Categorías de análisis

Al realizar una revisión de la caracterización establecida para la competencia interpretativa (presentada en el subcapítulo 2.1), se logran establecer las categorías que permitirán organizar y codificar la información que será recolectada a través de la implementación de las tareas diseñadas. Es por ello por lo que, a continuación, se describe el proceso de construcción de las categorías a la luz del marco teórico mencionado y teniendo en cuenta las tareas diseñadas.

Inicialmente se establecieron tres categorías pensadas en aquellos componentes teóricos que parecían ser mayormente representativos a la hora de reconocer la competencia interpretativa en las producciones de los estudiantes, a saber:

- Plantea interpretaciones de la información de una base de datos a partir del contexto.
- Reconoce la necesidad de determinar las medidas numéricas de una base de datos para generar interpretaciones de la información.
- Explica si un resultado o conclusión tiene sentido en el contexto de la base de datos.

Sin embargo, posterior al diseño de las tareas, se evidenció que las categorías que se habían propuesto anteriormente no abarcaban en su totalidad las posibles respuestas que podrían dar los estudiantes. Tras esta situación, resultó pertinente hacer una comparación más exhaustiva entre las características de la competencia interpretativa, lo cual permitió reconocer que entre ellas existen similitudes en relación con aspectos específicos, por ejemplo, un grupo de estas hacen referencia a los resultados matemáticos (como C9) y otras a la interpretación de información sin tener en cuenta aspectos matemáticos sino únicamente el conjunto de datos (como C2). Adicionalmente, algunas características apuntan a aspectos relacionados con establecer predicciones y conjeturas a partir de la información (como C4) y otras tantas a los sistemas de representación (como C5).

El ejercicio mencionado anteriormente, se concretó en la construcción de tres categorías que logran abarcar los elementos identificados y clasificar las 15 características de la competencia

interpretativa. En la Tabla 23, se presentan las categorías de análisis planteadas junto con las características asociadas de la competencia interpretativa.

Tabla 23. *Categorías de análisis diseñadas para las tareas.*

Categorías	Características de la competencia interpretativa
Relación entre lo matemático y lo contextual	C2. Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013). C6. Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real. C7. Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados. C8. Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema. C9. Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado. C13. Comparar resultados para diferentes condiciones en un experimento. (Franklin et al., 2020)
Tendencias y patrones (Predicción o conjeturas y Sesgos)	C4. Identificar tendencias, conjeturas, construir hipótesis, argumentar y tomar posición crítica frente a una determinada información (Moreno y Villanueva, 2013). C10. Criticar e identificar los límites del modelo utilizado para resolver un problema. C11. Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas. C12. Establecer las limitaciones de la información de la muestra (por ejemplo, una muestra puede o no ser representativa, variabilidad de las mediciones) (Franklin et al., 2020) C14. Evaluar e interpretar el impacto de los valores atípicos en los resultados. (Franklin et al., 2020)

	C15. Reconozca la presencia de valores faltantes y comprenda cómo los valores faltantes pueden agregarse sesgo a un análisis. (Franklin et al., 2020)
Sistemas de representación	C1. Observar, discernir y extraer información a través de elementos matemáticos.[12] (Contreras, 2018) C3. Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013) C5. Interpretar información presentada en forma gráfica o diagramas.

En la categoría denominada *Relación entre lo matemático y lo contextual* se agrupan las características que relacionan elementos referentes a los conceptos y procesos matemáticos, que para el caso particular de la investigación se asocian con el cálculo de medidas estadísticas; por otro lado, esta categoría también se asocia a la información que no ha tenido ningún tratamiento numérico, es decir al conjunto de datos dado, que en el caso de esta investigación corresponderá a la información en la base de datos. Lo anterior, no solo tiene la intencionalidad de dar una interpretación al número por sí solo, sino que las conclusiones se enmarquen en un contexto real y que a partir de allí los estudiantes sean capaces de identificar la relación entre lo numérico y lo contextual, determinando entre otras cosas las limitaciones y la coherencia entre los elementos mencionados.

En la categoría denominada *Tendencias y patrones (Predicción o conjeturas y sesgos)* se agruparon todas las características que tienen por finalidad generar interpretaciones que estén asociadas a la interacción con un conjunto de datos del cual se puedan establecer relaciones que involucren identificar el impacto de valores faltantes o atípicos, asumir una posición crítica frente a determinada modificación del conjunto de datos y establecer los limitantes que se pueden presentar cuando una información dada tiene determinadas características o particularidades. Esta categoría también da lugar a la comparación de poblaciones o muestras de datos y al uso de

[12] En esta investigación, los elementos matemáticos se entenderán como las diversas formas de representación o síntesis de los datos.

herramientas digitales en el proceso de comparar y determinar interpretaciones asociadas al conjunto de datos en cuestión.

Por último, en la categoría *Sistemas de representación* se incluyen las características en las cuales se evidencia un cambio en la forma de representación del conjunto de datos y los procesos asociados a dicha transformación de la información, como lo son: la identificación, extracción, interpretación, análisis y comprensión de lo que se quiere evidenciar a través del conjunto de datos.

3.5. Implementación de las tareas

La implementación de las tareas se llevó a cabo en la institución educativa *Tandem – Apoyo en procesos escolares* durante el semestre 2024-1, en el marco del curso de Estadística para estudiantes de grado noveno. La muestra, compuesta por 11 estudiantes que representaban el total de inscritos en este espacio académico (debido a las características particulares de la institución), se distribuyó en tres grupos de la siguiente forma:

- Grupo 1: Cuatro estudiantes
- Grupo 2: Tres estudiantes
- Grupo 3: Cuatro estudiantes

Es importante señalar que la distribución de los estudiantes en estos grupos no fue determinada por los autores, sino que correspondió a la organización establecida por la institución educativa.

La aplicación de las tareas fue realizada por los dos autores, el primero tuvo a su cargo la implementación de las tres tareas en los grupos 1 y 2, mientras que el segundo investigador se encargó del grupo 3. Debido a actividades institucionales que llevaron a la reducción de horarios, el grupo 3 no pudo completar la Tarea 3. Esta situación es significativa, ya que, de los 11 estudiantes inicialmente participantes, cinco no lograron realizar dicha tarea, lo que incluye a los cuatro integrantes del grupo 3 y un estudiante del grupo 2. Además, otros dos estudiantes completaron la tarea de manera parcial, limitándose a responder solo hasta las preguntas 9 y 10, lo que evidencia las dificultades que se presentaron durante el proceso de aplicación de esta última tarea.

Las tareas se realizaron utilizando las herramientas Google Docs y Microsoft Excel en computadoras que contaban con licencia para su uso; algunas de estas fueron proporcionadas por la institución, mientras que otras pertenecían a los estudiantes. La mecánica de envío y recolección

de resultados fue la misma para las tres tareas, independientemente del investigador o del grupo al que se aplicó. Cada estudiante tenía una carpeta compartida en Google Drive, etiquetada con su nombre, donde el docente subía las tareas en formato Google Docs, esto permitiría a los estudiantes acceder a la carpeta y manipular el archivo correspondiente.

Es importante mencionar que cada una de las bases de datos utilizadas para el desarrollo de las tareas se podía descargar (en formato .xls) a través de un enlace incluido en el archivo de la tarea (ver Anexos). Una vez que el estudiante finalizaba la tarea, debía subir el archivo de Microsoft Excel de la base de datos a la carpeta compartida, lo que permitía a los investigadores recolectar la información directamente en la nube de Google Drive.

Al disponer de respuestas digitalizadas, se eliminó la necesidad de realizar un proceso manual de transcripción de datos, lo que simplificó considerablemente la recolección y organización de la información. Por lo tanto, el único material analizado fue el de las producciones de los estudiantes, cuyo detalle se expondrá en el siguiente capítulo.

Durante la implementación de las tareas los investigadores intervinieron exclusivamente para aclarar dudas relacionadas con la forma, como el significado de ciertas palabras o la aclaración de algunas preguntas. Esta intervención fue deliberadamente limitada para evitar cualquier tipo de sesgo en la información, ya que el objetivo era que las tareas, por sí mismas, promovieran la competencia interpretativa de los estudiantes.

Finalmente, en la Tabla 24 se presenta el tiempo y las sesiones de clase que cada grupo empleó para completar las tareas:

Tabla 24. *Sesiones de clase y tiempos para realizar las tareas.*

	Grupo 1	**Grupo 2**	**Grupo 3**
Tarea 1	Dos sesiones de 45 minutos cada una y 15 minutos en la tercera sesión.	Dos sesiones de 45 minutos cada una y 15 minutos en la tercera sesión.	Dos sesiones de 45 minutos cada una y 15 minutos en la tercera sesión.
Tarea 2	30 minutos en la tercera sesión y una sesión completa de 45 minutos.	30 minutos en la tercera sesión y una sesión completa de 45 minutos.	30 minutos en la tercera sesión y completada como trabajo extraclase.

Tarea 3	Dos sesiones de 45 minutos.	Una sesión de 45 minutos y completada como trabajo extraclase.	N/A

Como se mencionó anteriormente, la imposibilidad de implementar ciertas tareas, así como el hecho de que algunas se completen como trabajo extraclase, se debió a una serie de cambios en los horarios institucionales. Estos cambios impactaron de manera significativa los tiempos de clase disponibles para varios grupos, lo que generó dificultades en la planificación y ejecución de las actividades programadas.

4. RESULTADOS

Resulta importante mencionar que algunos ítems de las tareas no hacen parte del análisis que se presenta a continuación, puesto que las preguntas que allí se plantean no caracterizan ninguna habilidad de la competencia interpretativa, sino que más bien su finalidad fue abordar asuntos relacionados con la contextualización de las problemáticas trabajadas, verificar conocimientos previos o propiciar los procesos de instrumentación e instrumentalización. A continuación, se presenta para cada tarea un análisis de su implementación a la luz de las categorías presentadas en el subcapítulo 3.4.

4.1. Tarea 1

En relación con la categoría denominada *Relación entre lo matemático y lo contextual* se evidencia que, en 13 de 18 ítems planteados, las respuestas de los estudiantes están asociadas a las características allí relacionadas; y, por otro lado, 5 de 18 ítems corresponden a la categoría denominada *Tendencias y patrones (Predicción o conjeturas y Sesgos).* Además, se evidencia que en esta tarea ninguna de las respuestas de los estudiantes está presente en la categoría *sistemas de representación.*

Con el fin de detallar y presentar el proceso de análisis de las producciones de los estudiantes, se presenta a continuación cada uno de los ítems con algunas respuestas y su clasificación en las categorías mencionadas anteriormente. Es importante mencionar que en el desarrollo de esta tarea participaron 11 estudiantes.

Para los ítems 3A, 3B y 3C (se presentan enseguida), los estudiantes tenían en cuenta las medidas estadísticas calculadas para todo el conjunto de datos correspondiente a los grados de alcohol como se muestra en la Figura 18.

Figura 16. *Respuesta de estudiante para la tabla T1- 3.*

	Media	Rango	Varianza	Desviación estándar
Población	12.0582641	54	115.056035	10.7264176

Nota. Todos los estudiantes obtuvieron los mismos valores numéricos en este punto.

Pregunta 3A: Según el Decreto 162 de 2021 una bebida alcohólica es un "producto apto para consumo humano que contiene una concentración no inferior a 2.5 grados alcoholimétricos". ¿Hay bebidas no alcohólicas en la base de datos? Justifique su respuesta.

En las Figura 19 se muestran las respuestas dadas por los 11 estudiantes que desarrollaron esta primera tarea. Se evidencia que 10 de los estudiantes respondieron acertadamente esta pregunta y solo un estudiante no logra llegar a la conclusión esperada en el proceso de identificar si existen bebidas no alcohólicas en la base de datos (probablemente sesgado por el contexto en el que se encuentran inmersos los datos). Sin embargo, es posible establecer que los estudiantes se ubican en la característica C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos correspondiente a la categoría *Relación entre lo matemático y lo contextual*, puesto que solo se requería realizar una observación de los datos para dar respuesta a este ítem.

Figura 1917*. Respuestas de estudiantes para el ítem T1 – 3A.*

Sí, hay bebidas que tienen un grado de alcohol menor a 2,5.
No hay bebidas no alcohólicas, ya que es una gráfica de licores con alcohol.
Si hay, gracias a el comando =COUNTIF pude verificar que hay 624 bebidas con concentración menor a 2.5 grados
Si hay bebidas no alcohólicas en la base de datos, como por ejemplo la Beb. Emb. Sab. Uva Moscatel Gran Viña. La cual tiene un grado de alcohol del 1%.
Si hay ya que hay bebidas hasta con un 1% de alcohol
Si debido a que la mayoría de los vinos son menos a 2.5
Si, hay bebidas como el vino, o la champaña que tienen menos de 2.5 grados de alcohol.
con esta información se podría decir que si hay bebidas no alcohólicas con grados de alcohol menores de 2,5
Si hay bebidas no alcohólicas en la base de datos, hay varias que tienen menos de 2.5 grados de alcohol.
Si hay bebidas no alcohólicas en la base de datos, porque hay bebidas con un grado menor a 2.5, siendo 2 o 1.
Si hay varias bebidas que para el estado no son bebidas alcohólicas ya que su grado de alcohol es inferior a 2.5, incluso en los datos tenemos bebidas que tienen tan solo 1% de alcohol. Un ejemplo de esto es Beb Emb Sab A Fresa Casa La V.

Pregunta 3B: Teniendo en cuenta la media y la desviación estándar calculadas para los datos de la variable "Grado de alcohol", realice una descripción de dicho conjunto de datos.

Esta pregunta fue respondida por 8 de 11 estudiantes. Teniendo en cuenta las respuestas dadas, se identifica que las interpretaciones apuntan a tres características correspondientes a la categoría *Relación entre lo matemático y lo contextual* como se muestra en la Tabla 25.

La característica C6 es la que se hace presente en mayor medida, exactamente en las respuestas de 7 estudiantes, puesto que se evidencia que logran establecer conclusiones que relacionan la media y la desviación estándar con el contexto de la base de datos; sin embargo, también se identificó

que por lo menos tres estudiantes logran proponer interpretaciones que involucran las limitaciones que tiene la desviación estándar y lo extrapolan al contexto de las bebidas alcohólicas, lo cual va en correspondencia con la característica C9. Por otro lado, se evidencia que uno de los estudiantes no propone una respuesta que relacione las medidas calculadas, por lo tanto, su interpretación se categoriza en la característica C2, teniendo en cuenta que solo se evidencia una lectura de la información presente en la base de datos.

Tabla 25. *Relación entre características y respuestas de los estudiantes para el ítem T1 - 3b.*

CARACTERÍSTICA	C6. Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.	C9. Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado.	C2. Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos
RESPUESTA ESTUDIANTE	La mayoria de las bebidas de la base de datos esta por encima de los 10 grados de alcohol, pero en promedio las bebidas estan por los 12 grados de alcohol	La media de grados de alcohol 12,1 eso hace que las bebidas alcohólicas están entre el 1,4% y el 22,8%	Que tiene un grado bastante alto de alcohol, las bebidas que se venden en el departamento son de grado de alcohol bastante alto

Pregunta 3C: ¿Qué puede decir respecto a la dispersión de la variable "Grado de alcohol" de las bebidas presentadas en la base de datos?

Las respuestas a dicho ítem se relacionan con dos características correspondientes a la categoría *Relación entre lo matemático y lo contextual.* Se identifica que una parte de los estudiantes establece una interpretación que relaciona las medidas estadísticas con el contexto de los datos, mientras que la otra parte solo establecen una conclusión con las medidas calculadas. En términos de la caracterización realizada para la competencia interpretativa, lo mencionado anteriormente se relaciona con C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real, como se evidencia en la Figura 20 y C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos, como se presenta en la Figura 21.

Figura 20 18. *Respuestas de estudiantes para el ítem T1 – 3c en correspondencia con C6.*

Hay poca, ya que el rango es 54 y la desviación es 10,7264176, entonces esto significa que hay poca dispersión basado en el rango de este.
Podemos decir que hay un rango de dispersión de 54, teniendo en cuenta que el maximo dato es 55 y el mínimo es 10
la dispersión del grado de alcohol varía entre 22 % y 2%
Que la dispersión es medianamente alta por que estan entre el 2% al 22,8%
Que tienen una dispercion muy alta te puedes encontrar desde 1% hasta 55%.

Figura 2119*. Respuestas de estudiantes para el ítem T1 – 3c en correspondencia con C2.*

Yo creo que los datos están, mayormente muy separados, ya que hay muchos tipos de grados diferentes.
La dispercion es bastante grande por que los datos estan algunos muy alejados unos entre otros entonces la dispercion es grande.
Que el grado de alcohol varia bastante, los espacios entre los grados son grandes.

Para los ítems 4A, 4B y 4C (se presentan enseguida), los estudiantes tenían en cuenta las medidas estadísticas calculadas para el conjunto de datos correspondiente a los grados de alcohol con un dato adicional como se muestra en la Figura 22, junto la tabla de la Figura 18.

Figura 2220*. Respuesta de estudiante para la tabla T1-4.*

	Media	Rango	Varianza	Desviación estándar
Población con dato adicional	12,06179984	95	115,3479715	10,74001729

Nota. La mayoría de los estudiantes obtuvo los mismos valores numéricos en este punto.

Pregunta 4A: ¿Qué representa el cambio en el rango de los datos al comparar las Tablas 1 y 2?

Para este ítem se obtuvo un total de 10 respuestas, todas correspondiente a la categoría denominada *Relación entre lo matemático y lo contextual.* La característica que se hizo presente en todas las producciones corresponde a C13: Comparar resultados para diferentes condiciones en un experimento. Sin embargo, se evidenció que en algunos casos se presentó simultáneamente la característica C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real. A continuación, se ejemplifica lo mencionado anteriormente (Tabla 26):

Tabla 26. *Relación entre características y respuestas de los estudiantes para el ítem T1 – 4A.*

	DESCRIPCIÓN
C6	Lo que representa el cambio es que el máximo porcentaje de alcohol cambiará de 55 a 96 por ciento. El cambio en los datos se debe a que el otro producto que metimos era mucho mayor en volumen de alcohol y por esta razon el rango aumento tanto.
C13	Como son muchos datos solo se nota en el rango el cual pasa de ser 55 a ser 95 por este ultimo dato. Ahí un cambio es que el rango es mayor pero además de eso son casi iguales

Se observa que en todos los casos los estudiantes logran identificar que al comparar la medida estadística de los dos conjuntos de datos, el cambio en el rango corresponde a un aumento, lo cual permite evidenciar la característica C13. Sin embargo, la diferencia entre una interpretación y otra radica en que no todos los estudiantes contextualizan sus respuestas y por lo tanto no mencionan los grados de alcohol. La particularidad mencionada anteriormente solo se evidenció en las respuestas de cuatro estudiantes, los cuales presentaron respuestas similares a la del ejemplo presentado en la Tabla 24 para la característica C6.

Pregunta 4B: Al comparar la información de la población y la población con el dato adicional de las Tablas 1 y 2, ¿qué diferencias o similitudes puede enunciar para cada medida calculada?

Este ítem lo respondieron 10 de 11 estudiantes. Nuevamente, al igual que en la pregunta 4A, se evidencia que la característica C13 se hace presente en todas las respuestas dadas por los estudiantes. Adicionalmente, también se identificó en dos de las respuestas la característica C9: Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado. Las dos características mencionadas hacen parte de la categoría *Relación entre lo matemático y lo contextual.*

En la Figura 23 se presentan algunas de las interpretaciones presentadas por los estudiantes, en las cuales se observan las apreciaciones realizadas en relación con las medidas calculadas para los dos conjuntos de datos. La mayoría de los estudiantes establecen la comparación entre las medidas identificando la relación de igualdad o de aumento, pero solo en los dos casos presentados, los

estudiantes establecen adicionalmente una interpretación que involucra la amplitud del rango y lo que representa en el conjunto de datos este modelo matemático (diferencia entre el dato mayor y el dato menor).

Figura 2321. *Respuestas de estudiantes para el ítem T1 – 4b que cumplen C9 y C13.*

Media: Muy poca diferencia al ser tantos datos Rango: Cambia drásticamente al tener un dato 74.55% mayor que en la primera tabla Varianza: Lo mismo que media Desviación: Lo mismo que media
Las medidas calculadas cambiaron por el dato adicional. el rango aumento significativamente pero la varianza y la desviación solo cambiaron por decimales. y claro que el maximo cambio al 96 por que la bebida alterada es la mas alcohólica de todas.

Pregunta 4C: ¿Por qué cree usted que las medidas de la población no cambian demasiado aun cuando se agrega un dato adicional mucho mayor a los demás?

Para este ítem se observó que todas las respuestas se pueden caracterizar como C9: Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado y C13: Comparar resultados para diferentes condiciones en un experimento, las cuales corresponden a la categoría *Relación entre lo matemático y lo contextual*, a continuación, en la Tabla 27 se presentan algunas de las interpretaciones realizadas por los estudiantes:

Tabla 27. *Respuestas de los estudiantes para el ítem T1 – 4C.*

RESPUESTAS DE ESTUDIANTES
Las medidas de población no cambian demasiado ya que el solo agregamos un solo dato que a pesar de que sea más alto que las demás no cambia demasiado ya que solo añadimos solo uno de esto.
Porque aunque hay un valor mucho más grande que todos, al calcular los valores no va a haber cambio porque solo es un valor más que no tendrá tanto peso en los valores.

Al observar las respuestas, es posible establecer similitudes en relación con la identificación de los alcances y limitantes que se podrían presentar al calcular las medidas numéricas cuando se agrega un dato atípico a un conjunto con un alto volumen de datos, lo anterior permite comprender por qué no se presenta un cambio significativo entre los dos conjuntos de datos.

Por otra parte, para el ítem 6A, los estudiantes tenían en cuenta las medidas estadísticas calculadas para las muestras derivadas del conjunto de datos correspondiente a los grados de alcohol y también para el mismo conjunto con el dato adicional. Además, puntualmente en los ítems 6B, 6C

y 7C era importante considerar la tabla correspondiente a la Figura 14. A continuación, en las Figuras 24 y 25, se presentan algunas de las tablas diligenciadas por los estudiantes correspondientes a la muestra y la muestra con el dato adicional, respectivamente.

Figura 2422. *Respuesta de estudiante para la tabla T1- 5.*

	Media	Rango	Varianza	Desviación estándar
Muestra	8,546666667	13	14,62348889	3,856338241

Figura 2523. *Respuesta de estudiante para la tabla T1- 6.*

	Media	Rango	Varianza	Desviación estándar
Muestra con dato adicional	9,980327869	94	140,0019399	11,83224154

Pregunta 6A: ¿Qué diferencias o similitudes puede enunciar para cada medida, teniendo en cuenta la información de las tablas 3 y 4?

Para este ítem se compararon las respuestas de 10 estudiantes los cuales describen la interpretación de las medidas teniendo en cuenta lo registrado en dos tablas como las presentadas en las Figuras 24 y 25. La característica que se hace presente en todas las respuestas corresponde a C13: Comparar resultados para diferentes condiciones en un experimento. Teniendo en cuenta que estaban comparando una muestra de la base de datos y esta misma, pero con un dato atípico, los estudiantes logran identificar que hay un aumento significativo en los valores, particularmente en rango, varianza y desviación estándar. Adicionalmente, cuatro de los estudiantes añaden a su interpretación otra característica que se relaciona con C9: Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado, puesto que reconocen el impacto que tiene el agregar un dato atípico a un conjunto que tiene un bajo volumen de datos y cómo esto tiene implicaciones en el cálculo de las medidas estadísticas.

Las características mencionadas anteriormente corresponden a la categoría *Relación entre lo matemático y lo contextual* y se ejemplifican con las siguientes respuestas (Tabla 28).

Tabla 28. *Respuestas de los estudiantes para el ítem T1 – 6A.*

	C13	C9 y C13
RESPUESTAS ESTUDIANTES	La media es muy parecida, pero si sube casi un 1.5%, El rango es totalmente distinto pues ahora el dato max es de 96, asi que el rango es de 94, La Varianza cambia totalmente pasa de un 14% a un 140% y por ultimo la desviacion estandar, tambien cambia bastante pasando de casi un 4% a un 12%	Entre la tabla 3 y 4 no hay similitudes solo hay diferencias ya que como son tan poquitos datos con números muy bajos en la tabla 3, los datos de tabla 4 cambian por culpa del 96 ya que es un número muy alto y como no hay tantos datos en la tabla cambian drásticamente

Pregunta 6B: De acuerdo con lo observado, ¿de qué dependería poder asumir que la media y las medidas de dispersión obtenidas en la muestra sean aproximadamente iguales a las de la población? Explique.

Para este ítem, las respuestas obtenidas por 6 de los estudiantes corresponden a la categoría denominada *Tendencias y patrones (Predicción o conjeturas y Sesgos),* en la cual se identificó que predominan (para este punto en particular) dos características: C10. Criticar e identificar los límites del modelo utilizado para resolver un problema y C12. Establecer las limitaciones de la información de la muestra (por ejemplo, una muestra puede o no ser representativa, variabilidad de las mediciones).

Se evidencia que las interpretaciones de 4 de los 6 estudiantes presentan en su estructura rasgos de la característica C12, puesto que la caracterización de la muestra que cumple las condiciones solicitadas en el enunciado se establece a partir de un caso particular tal y como se ejemplifica en la Figura 27. Por otro lado, los 2 estudiantes restantes, proponen las características que se deberían presentar en una muestra, pero haciendo énfasis en las variaciones presentadas particularmente en el rango y las implicaciones que tienen esas diferencias observadas en cuanto a lograr establecer medidas estadísticas iguales o similares entre una población y una muestra que tiene un bajo volumen de datos, lo cual va en correspondencia con C10 (ver Figura 26).

A continuación, se presentan algunas de las respuestas dadas por los estudiantes para este ítem y que ilustran de forma precisa lo mencionado anteriormente.

Figura 2624. *Respuestas de estudiantes para el ítem T1 – 6b que cumplen C10.*

En la media sigue siendo el mismo promedio, el rango cambia ya que este mira los números y el 96 extra es muy alto comparado con los otros números. En la desviación y la varianza al ser el número alto y se tiene que restar por el promedio y luego elevarlo al cuadrado cambia el número a comparación que en la primera muestra
Que la poblacion inclyuera mas datos del alcohol etilico de 96% pues solo con uno para casi 3000 datos no cambiara casi nada

Figura 2725. *Respuestas de estudiantes para el ítem T1 – 6b que cumplen C12.*

depende de los datos que se toman en la muestra
para poder asumir esto tendríamos que tener una muestra más grande porque al tener una muestra más grande habria mucha mas probabilidad de que las medidas de dispersión fueran más parecidas a las de la población

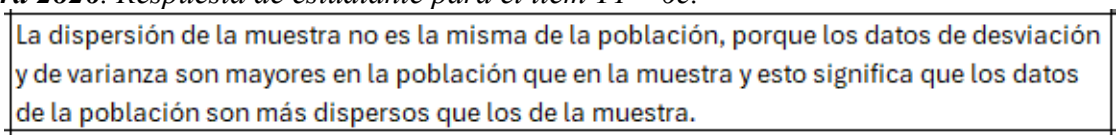
Resulta importante mencionar que otros 3 estudiantes respondieron el ítem, pero sus respuestas no fueron categorizadas porque en ellas no se realizó una comparación entre los conjuntos de datos ni se incluyeron elementos relacionados con la población, la muestra o las medidas estadísticas; por lo cual, no se evidenció la coherencia ni la relación con lo preguntado.

Pregunta 6C: ¿La dispersión observada en la muestra es equivalente a la de la población? Justifique su respuesta.

En este ítem, al igual que en los presentados anteriormente en el análisis de esta primera tarea, se identifica que las respuestas de los estudiantes corresponden a la categoría *Relación entre lo matemático y lo contextual.* Las 10 respuestas obtenidas se caracterizan por la correspondencia con C13: Comparar resultados para diferentes condiciones en un experimento. Es por ello que, entre las interpretaciones presentadas por los estudiantes, se establece una comparación entre las medidas estadísticas de la población y la muestra, entre estas se presenta una característica en común, la cual corresponde a observar que las medidas cambiaron y que hay un aumento de los valores consignados en las tablas.

Se destaca que uno de los estudiantes reconoce y reporta en su conclusión el alcance que tiene la desviación estándar en el conjunto de datos, lo cual va en correspondencia con la característica C9: Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado. Lo anterior se evidencia en la siguiente respuesta (ver Figura 28):

Figura 2826*. Respuesta de estudiante para el ítem T1 – 6c.*


La dispersión de la muestra no es la misma de la población, porque los datos de desviación y de varianza son mayores en la población que en la muestra y esto significa que los datos de la población son más dispersos que los de la muestra.

Pregunta 7C: ¿Qué implicaciones tuvo el agregar el dato del licor adulterado (dato atípico) al conjunto de datos dado inicialmente (Pregunta 2)? ¿Qué sucedió con las medidas calculadas cuando se agregó el mismo dato a la muestra (Pregunta 4)? Escriba una conclusión.

Para dar respuesta a este ítem, los estudiantes debían considerar las tablas correspondientes a la población y la muestra con el dato adicional (Figuras 22 y 25, respectivamente). En esta

categorización, se tuvo en cuenta un total de 10 respuestas las cuales en su totalidad hacen parte de la categoría *Relación entre lo matemático y lo contextual.*

Aunque una gran parte de las interpretaciones se agrupan en la característica C13 (exactamente la de 9 estudiantes), simultáneamente 5 de ellos logran establecer en su conclusión elementos que involucran la característica C9. Además, en una de las respuestas se evidencian elementos de C7: Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados. A continuación, en la Tabla 29, se presenta un ejemplo que representa lo descrito anteriormente.

Tabla 29. *Relación entre características y respuestas de los estudiantes para el ítem T1 – 7C.*

CATEGORIA	RESPUESTAS ESTUDIANTES	ANÁLISIS
C9	las medidas de dispersión cambiaron mucho debido a que este dato es mucho más significativo ya que solo hay 44 licores en esta muestra, entonces para las medidas de dispersión afecta mucho más este licor adulterado	Se identifica que el estudiante evidencia el cambio en las medidas estadísticas porque reconoce las implicaciones de agregar un dato atípico a una muestra que tiene un bajo volumen de datos en comparación a una población con alto volumen de datos.
C13	Ya que el dato era enorme, el maximo y minimo cambia	Aunque la respuesta no es del todo correcta, pues no hay un cambio en el dato menor del conjunto de datos, se evidencia una comparación entre las medidas estadística de los dos conjuntos (población y muestra).
C9, C13	En el conjunto inicial, no se noto la diferencia en los datos pues eran mas de 20000 datos por lo cual un dato mucho mas arriba de grado de alcohol no afecto. En cambio cuando se aplica en una muestra, la media aumenta, pero lo que mas aumenta drasticamente es la desviacion que en este caso pasa de 3 a 11. En conclusion entre menos datos 1 muy por encima afectara mucho mas al resultado final.	En la respuesta se propone simultáneamente un análisis que relaciona las características que tienen los dos conjuntos en cuanto a la cantidad de datos y también, relaciona lo anterior con las medidas estadísticas calculadas que a su vez cambian al agregar un dato atípico.
C7, C9, C13	en conclusión este licor adulterado tuvo cierto impacto en las medidas, cambió la media, las de dispersión por ser un licor con muchisimo mas alcohol que los demás y todo esto terminó dándonos valores que no son muy buenos representando la distribución de los datos que no tenían el adulterado	Aunque la respuesta reúne las características mencionadas en el análisis de la respuesta anterior, se incluye un elemento no presentado hasta ahora en este análisis, correspondiente a identificar como las medidas estadísticas calculadas repercuten en el contexto de las bases de datos y el impacto negativo que tendría en un contexto probablemente social.

A continuación, se presenta el análisis correspondiente a los ítems 7A, 7C y 7D, esto se debe a que en ellos se tenía un propósito común, el cual era propiciar que las respuestas de los estudiantes se relacionaran con la característica C7: Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados. Sin embargo, al verificar las respuestas de los estudiantes, se logró evidenciar que la relación entre el contexto propuesto sobre la problemática del consumo de las bebidas alcohólicas y las medidas de dispersión junto con la base de datos no era totalmente visible para los estudiantes, por lo cual no les fue sencillo plantear interpretaciones que vincularan ambos elementos y en la mayoría de los casos sus conclusiones apuntaron casi que en su mayoría a destacar elementos de la lectura que se propuso junto en esta tarea (Tabla 31).

No obstante, lo que sí se logró a través del planteamiento de estos ítems, fue propiciar que los estudiantes determinaran algunas posibles causas de las altas ventas de las bebidas alcohólicas y de las repercusiones sociales que inciden en el consumo de estas; para ello, se hace una revisión en paralelo con las respuestas obtenidas en los ítems 1A, 1B, 1C y 1D (Tabla 30).

Tabla 30. *Relación entre ítems y respuestas de los estudiantes para las preguntas de T1-1.*

ÍTEMS PLANTEADOS PREVIO A LA INTERACCIÓN CON LA BASE DE DATOS	
Pregunta 1A: ¿Qué piensa usted sobre el consumo excesivo de alcohol?	**Respuesta de estudiante:** Creo que el consumo excesivo de alcohol lleva a que las personas pierdan el control y no sean conscientes ni de su entorno, ni de lo que hacen. Por lo que la mayoría de veces puede terminar en accidentes, enfermedades, problemas y abusos.
Pregunta 1B: ¿Cuáles problemas de salud conoce usted que se pueden generar por el consumo excesivo de alcohol?	**Respuesta de estudiante:** Genera, pérdida de neuronas, muchos problemas al hígado como la cirrosis o cáncer y en casos de mucha dependencia puede generar hasta la muerte tanto por sobreconsumo como por abstinencia.
Pregunta 1C: ¿Cuáles problemáticas sociales considera usted que se generan por el consumo excesivo de alcohol?	**Respuesta de estudiante:** Hay varias problemáticas que relacionan el consumo excesivo de alcohol con problemas sociales, un ejemplo muy común sobre el tema es la gran facilidad que tiene la persona a la hora de tomar decisiones, lo cual en la mayoría de los casos lleva a la violencia física y verbal.

Pregunta 1D: En promedio, ¿cuál cree usted que es el grado de alcohol de las bebidas alcohólicas que conoce?, ¿en qué considera que afecta el grado de alcohol a los licores y a las personas que los consumen?, ¿por qué?	**Respuesta de estudiante:** El promedio del porcentaje de las bebidas alcohólicas son de 29%. Entre más alcohol alla, mas rapido te emborrachas y es muy pesado, ya que esta muy concentrado y hay mucha cantidad

Tabla 31. *Relación entre ítems y respuestas de los estudiantes para las preguntas de T1 -7.*

ÍTEMS PLANTEADOS POSTERIOMENTE A LA INTERACCIÓN CON LA BASE DE DATOS	
Pregunta 7A: Teniendo en cuenta toda la información suministrada, ¿cómo se relaciona la problemática del consumo de alcohol con la producción de las bebidas alcohólicas?	**Respuestas de estudiantes:** Las empresas producirán más bebidas alcohólicas dependiendo de la cantidad del nivel de alcohol de las bebidas, y de qué bebidas se venden más Las empresas producirán más bebidas alcohólicas dependiendo de la cantidad del nivel de alcohol de las bebidas, y de qué bebidas se venden más La excesiva venta de alcohol con altos porcentajes de alcohol, claramente impulsan mas el consumo de alcohol.
Pregunta 7B: ¿Qué significado puede dar a las medidas halladas para la variable "¿Grado de alcohol" (punto 2), en relación con el contexto de los licores y los problemas existentes frente al consumo de estas bebidas?	**Respuestas de estudiantes:** Esto indica que los porcentajes de alcohol son muy altos y eso que es solo en el departamento de Risaralda. el significado que pueden dar es que dependiendo del nivel de adicción hay ciertas bebidas y para los que necesitan más alcohol pueden tomar una bebida con mayor porcentaje
Pregunta 7D: Realice la lectura de la información suministrada por el Ministerio de Salud en relación con la problemática del consumo de bebidas alcohólicas en Colombia (Anexo). Teniendo en cuenta la información de la base de datos y la lectura realizada, ¿qué conclusiones puede establecer?	**Respuestas de estudiantes:** Que hay que tener mucho cuidado con el alcohol, y mucho mas con las bebidas no controladas por la ley las cuales cuentan con porcentajes de alcohol mucho mas alto. Las conclusiones que puedo establecer sobre el tema es que las bebidas alcohólicas en Colombia tienen un porcentaje de alcohol muy alto que afecta la salud de las personas.

Al comparar las respuestas dadas por los estudiantes en las Tablas 30 y 31, se evidencia que las conclusiones dadas para los ítems del punto 1 muestran la opinión de los estudiantes en cuanto a la problemática que se les plantea (lo cual va en correspondencia con el propósito de dichas preguntas). Sin embargo, las respuestas del ítem 7, no se limitan únicamente a una perspectiva personal del contexto, más bien en estas conclusiones se evidencia la intención de establecer una interpretación que vincule la base de datos y la problemática. Sin embargo, las conclusiones se ven mayormente sesgadas por la lectura y aunque algunos elementos son comunes a lo que se presenta en la base de datos, no es posible asegurar que la conclusión se haya planteado tomando elementos de las dos fuentes de información, de allí la no categorización de las interpretaciones presentadas.

Por otro lado, también se identificó que los estudiantes no establecieron conclusiones en las cuales se relacionaran o mencionaran las medidas de dispersión, solo se logró observar dos respuestas en las que se intentó vincular la media para dar respuesta a uno de los ítems mencionados en la Tabla 30, en la Figura 29 se muestran las respuestas:

Figura 2927. *Respuestas de estudiantes para el ítem T1 – 7b.*

Las bebidas por arriba del promedio de grado de alcohol el cual es mayor a 12, ayudan aun mas a la problematria del alcohol.
Las bebidas alcohólicas, pueden afectar la salud de las personas ya que la media de alcohol es muy alta.

Las preguntas de la 9 a la 13 corresponden a la segunda parte de la Tarea 1, en ellas el contexto propuesto se relaciona con las calificaciones obtenidas en un examen de matemática por parte de los estudiantes de la institución donde se implementaron las tareas (Tabla 32).

Tabla 32. *Relación entre ítems y respuestas de los estudiantes para las preguntas de T1-7.*

ITEM	ANÁLISIS
Pregunta 9: Suponga que el profesor al transcribir las calificaciones al archivo Excel, cometió un error y escribió un cero adicional en 10 de las notas. Escriba una conjetura, en la primera fila de la Tabla 6, sobre los resultados que podría obtener al calcular las medidas allí descritas tras modificar los 10 datos.	Los dos ítems fueron respondidos por 10 estudiantes. Las respuestas en ambos casos se caracterizan por tener elementos correspondientes a C4: Identificar tendencias, conjeturas, construir hipótesis, argumentar y tomar posición crítica frente a una determinada información, que corresponde a la categoría denominada *Tendencias y patrones (Predicción o conjeturas y Sesgos)*. Como se

Pregunta 11: Suponga que se eliminan 20 datos de forma aleatoria. Escriba una conjetura, en la primera fila de la Tabla 7, en relación con los resultados que podría obtener tras dicha eliminación.	describe en C4, los estudiantes llevan a cabo un proceso de elaborar conjeturas a partir de la identificación y lo ya estudiado acerca de cómo cambia un conjunto de datos al agregar datos atípicos (en el caso de la pregunta 9) y las variaciones que tienen las medidas estadísticas cuando tomo una muestra de datos y la comparo con su respectiva población (en el caso de la pregunta 11).

A continuación, se presentan algunas de las respuestas dadas por los estudiantes para cada uno de los ítems, es importante aclarar que en este análisis no se da un peso mayor a los cambios que hay entre los valores numéricos hallados, puesto que el foco de interés está puesto en las conjeturas planteadas para casa caso.

Figura 3028. *Respuestas de estudiantes para el ítem T1 – 9.*

Modificación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura	disminuye	se mantiene	aumento	aumenta un poco
Medida	5,3	9,0	5,74496402	2,396865457

Modificación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura	9.7	79.8	225.47798	15.015924
Medida	8.8	79.8	194.87033	13.959597

Figura 3129. *Respuestas de estudiantes para el ítem T1 – 11.*

Eliminación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura	10	99	100	10
Medida	8.9	89.4	181,8799815	13,48628865

Eliminación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura	Creo que la media al ser el promedio aumentará mucho ya que agregando un 0 todos los datos suben una décima y los 10 suben incluso una centésima	El rango al demostrar la diferencia también cambiará ya que ahora la resta entre el dato mayor y menor es mucho más grande.	La varianza también aumenta ya que la dispersión entre los números será mucho mayor	también aumenta ya que la dispersión entre los números será mucho mayor
Medida	9,8	79,8	208,2470045	14,4307659

Como se puede apreciar en las Figuras 30 y 31, se observaron dos posibles respuestas, las primeras corresponde a escribir por medio de lenguaje natural lo que podría ocurrir con las medidas estadísticas, estableciendo si los valores se mantienen iguales, aumentan o disminuyen. Por otro lado, algunos estudiantes proponen un valor numérico que simbolice lo que podría obtenerse al calcular las medidas estadísticas; para esta segunda opción, resulta importante mencionar que los estudiantes pueden estar sesgados al dar sus repuestas, porque antes de realizar las conjeturas calcularon las medidas y lo con lo reportado en la tabla buscaban dar un valor aproximado y cercano al "correcto".

Lo anterior, hace visible que algunos estudiantes parcialmente elaboran o desarrollan aspectos de la característica C4, puesto que logran construir conjeturas que den ideas de lo que podría pasar en el conjunto de datos cuando se realizan la eliminación o modificación de algunos de sus datos.

Al verificar los ítems 10A y 12A se identificó que las respuestas dadas por los estudiantes correspondían a la categoría *Tendencias y patrones (Predicción o conjeturas y Sesgos).* Adicionalmente, la mayoría de las conclusiones se caracterizaban a través de C11: Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas. Únicamente 2 respuestas del ítem 10A presentan otro tipo de caracterización, por lo cual se relacionaban con C4: Identificar tendencias, conjeturas, construir hipótesis, argumentar y tomar posición crítica frente a una determinada información. A continuación, en la Tabla 33, se presentan los enunciados de las preguntas, algunas respuestas de los estudiantes y el análisis correspondiente.

Tabla 33. *Relación entre ítems y respuestas de los estudiantes para las preguntas de 10A y 12A de T1.*

ITEMS	**Pregunta 10A:** Modifique aleatoriamente los 10 datos teniendo en cuenta el error cometido por el profesor y calcule la media, rango, varianza, desviación estándar. Registre los resultados en la segunda fila de la Tabla 6. ¿Qué diferencias o similitudes existe entre la conjetura y las medidas obtenidas? Comente por qué considera que surgen estas diferencias o similitudes.	**Pregunta 12A:** Proceda a eliminar 20 datos de forma aleatoria, a continuación, calcule la media, rango, varianza, desviación estándar y registre los resultados en la segunda fila de la Tabla 7. ¿Qué diferencias o similitudes existe entre las conjeturas y las medidas obtenidas? Describa y explique por qué surgen estas diferencias o similitudes.
RESPUESTAS	Mi hipótesis de la conjetura se hizo realidad en la medida ya que como dije agregar un 0 hace que todos los datos aumentan en gran cantidad La conjetura es simplemente agregar un cero mientras que en la medida ya se agregan datos aleatorios con un cero mas por lo cual hay mas posibilidades de cambiar los numeros. La varianza es muy diferente, de resto mas o menos son similares. Esto gracias a que nosotros podemos podemos predecir los datos en base de un resultado inicial	Como dije al eliminar 20 de los datos las medidas disminuyeron El promedio dsiminuyo, el rango se mantuvo, la varianza disminuyo y la desviación disminuyo. Al quitar datos de una base los valores calculados cambián dependiendo de cuantos quitaste Las diferencias que tenemos gracias a esta tabla es que como los datos fueron modificados y la conjetura no se parece para nada en la tabla ya que la media el rango la varianza y la desviación es diferente
ANÁLISIS	En los dos casos, los estudiantes hacen uso de las medidas estadísticas calculadas para poder determinar la veracidad de su conjetura, las soluciones obtenida para estos ítems variaron de forma significativa entre los estudiantes; sin embargo, todos lograron abordar de forma asertiva la situación plateada, observando que su predicción junto con el resultados real va a estar determinado por las características de la población o la muestra que se esté analizando, por lo cual se evidencia que logran identificar dichas características como lo son: aumento de medidas ante un dato atípico y disminución de estás al eliminar determinada cantidad de datos. Lo mencionado anteriormente, también permite observar que los estudiantes reconocer 2 elementos importante presentados en las características C14: Evaluar e interpretar el impacto de los valores atípicos en los resultados y C15: Reconozca la presencia de valores faltantes y comprenda cómo los valores faltantes pueden agregarse sesgo a un análisis, respectivamente.	

Pregunta 13: ¿Qué puede concluir en relación con las medidas, al hacer la modificación y eliminación al conjunto de datos suministrado?

En esta última pregunta de la Tarea 1, se identificó que las respuestas proporcionadas por los estudiantes corresponden a la categoría *Relación entre lo matemático y lo contextual* y la

característica predominante es C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos. Algunas de las respuestas se relacionan en la Figura 32.

Figura 3230. *Respuestas de estudiantes para el ítem T1 – 13.*

Que dependiendo de los datos, todo va a cambiar, incluyendo las medidas y asi mismo el promedio
Que entre mayor número de datos mayores las respuestas en las medidas
A la conclusión que se puede llegar es que los datos cambian mucho debido a que dependiendo del dato cambiado o eliminado cambian los resultados
Que los valores cambiaorn drásticamente cuando haces un cambio o un error pequeño en la base de datos

Como se observa en la Figura 32, las interpretaciones de 6 de los estudiantes únicamente se enfocan en las medidas estadísticas y no involucran el contexto en el que se encuentran inmensos los datos, lo anterior permite relacionar las respuestas con los elementos mencionados en la característica C2. Por otro lado, uno de los estudiantes realiza una descripción de lo que sucede con las medidas de los datos haciendo énfasis en la comparación entre los conjuntos de datos (el dado y el correspondiente a la modificación o eliminación) como se muestra en la Figura 33, lo anterior se evidencia en la característica C13: Comparar resultados para diferentes condiciones en un experimento.

Figura 3331. *Respuestas de estudiantes para el ítem T1 – 13.*

Al hacer modificaciones o borrar partes de los datos suministrados dependiendo de qué tan extrema sea la borrada o la modificación pueden o no cambiar significativamente las medidas de dispersión y de tendencia central. pueden aumentar o disminuir dependiendo de los cambios que se hagan pero al final de cuentas sigue siendo el mismo procedimiento para sacarlas. hay que tener cuidado al equivocarse porque eso sí puede dar alteraciones en nuestros resultados y no es lo ideal. pero en general las medidas normalmente con lo que hicimos se mantuvieron parecidas con pocas fluctuaciones menos cuando se le agregaron los 0 extras que eso si cambio las medidas extremadamente.

Con el fin de concretar la información analizada en la Tarea 1, en la Tabla 34 se presenta un resumen de las características asociadas a las respuestas de los estudiantes en cada una de las preguntas:

Tabla 34. *Características de la Tarea 1.*

Pregunta \ Categoría	N/A	C2	C4	C6	C7	C9	C10	C11	C12	C13	C14	C15
1	X											
2	X											
3A		X										
3B		X		X		X						
3C		X		X								
4	X											
4A				X						X		
4B						X				X		
4C						X				X		
5	X											
6	X											
6A						X				X		
6B							X		X			
6C						X				X		
7	X											
7A	X											
7B	X											
7C					X	X				X		
8	X											
9			X									
10	X											
10A			X					X			X	
11			X									
12	X											
12A			X					X				X
13		X								X		

4.2. Tarea 2

En analogía con la Tarea 1, es relevante destacar que algunas preguntas de la Tarea 2 no fueron diseñadas para abordar directamente las categorías de análisis relacionadas con la competencia interpretativa. En particular, el ítem 1 tenía como objetivo principal contextualizar al estudiante sobre la problemática de movilidad en las ciudades colombianas, mientras que el ítem 2 solicitaba a los individuos identificar y categorizar las variables estadísticas presentes en la base de datos, con el fin de familiarizarlos con esta.

A continuación, se lleva a cabo un análisis considerando las categorías de la competencia interpretativa.

Pregunta 3: Considerando los datos de la variable estadística "VELOCIDAD KM/H", calcule su media y responda: ¿cuáles son las unidades de medida de la media? Para usted, ¿qué significado tiene la media calculada dentro del contexto de la base de datos?

Al respecto, la Figura 34 corresponde a las respuestas de la pregunta. En esta figura se evidencia que todos respondieron correctamente a la pregunta y llevaron a cabo una interpretación adecuada de la medida en función del contexto, excepto uno. Aunque este último estudiante respondió correctamente en cuanto a las unidades de medida, al expresar "la media calculada es calcular la media", no se puede asegurar que haya interpretado la medida en el contexto requerido, por lo cual esta respuesta no se categoriza.

***Figura 3432**. Respuestas de los estudiantes T2-3.*

La unidad de medida de la media sería de km/h, ya que al encontrar la media no cambiaría nada. La media es el promedio de las velocidades
27,25041186. la medida de esta media es km/h. Pues es el promedio de velocidad a la que van los habitantes de medellin.
El promedio calculado en los datos de 27,25 kilómetros por hora indica que esta es la velocidad promedio registrada en el conjunto de datos.
la unidad de medida es km/h ya que esa medida esta en cada dato, lo que significa es que el promedio de las velocidades.
La media es 24,38 aproximadamente el significa que tiene calcular la media es que con la media podemos saber qué velocidad pasan las personas mas o menos.
27,25041186 en Km/H. Significa que la velocidad promedio que tiene un carro cuando viaja en Medellin es de 27.25 Km/H
las unidades de medida son km/h la media calculada es calcular la media
La unidad de medida sería km/h, significa que ese número es la velocidad promedio que se mantiene
Las unidades para la media son Km/h, porque es la velocidad. Significa cual es la velocidad promedio de la movilidad en Bogotá
El significado que tiene la medida calculada dentro del contexto es que el tráfico se estanca mucho en la ciudad ya que la media de velocidad dio 27km/h en vias que estan hechas para un promedio de 50km/h

Estas 10 respuestas han sido clasificadas dentro de la categoría *Relación entre lo matemático y lo contextual*, desarrollando la característica C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real. Esto se debe a que las respuestas cumplieron con los objetivos planteados, que, en términos generales, consistían en identificar las unidades de medida tras realizar un proceso matemático y en interpretar el valor obtenido dentro del contexto.

Al igual que esta pregunta, las preguntas 7, 12 y 16 desarrollan la misma característica C6 dentro de la misma categoría. En la Tabla 35, se presentan las particularidades y el análisis de algunas de las respuestas a cada pregunta.

Tabla 35. *Análisis de las preguntas 7, 12 y 16 de T2*

ITEMS	RESPUESTAS	ANÁLISIS
Pregunta 12. ¿Qué representa el valor obtenido en relación con la situación propuesta?	**Respuesta 1.** Representa la velocidad promedio de los vehículos. **Respuesta 2.** Representa la desviación de los datos.	A pesar de que algunas respuestas presentaron interpretaciones inadecuadas, como la respuesta 1, y otras mostraron interpretaciones adecuadas, como la respuesta 2, por lo cual, independientemente de la calidad de la interpretación evidenciaron el desarrollo de la característica C6. Aunque en la respuesta 2 no se explicita claramente la relación con el contexto, considerando las preguntas previas, la variable utilizada en todas ellas y la mención a "los datos", se asume que el estudiante se refiere a la desviación de las velocidades, lo que sugiere una relación implícita con el contexto trabajado. En esta pregunta, 2 de los estudiantes no la respondieron, 3 proporcionaron respuestas correctas y 5 dieron respuestas incorrectas.
Pregunta 16. ¿Qué representa en el contexto de la situación los valores obtenidos en ítem 13?	**Respuesta 1.** Representa los espacios entre las velocidades. **Respuesta 2.** Representan el cambio de las velocidades. **Respuesta 3.** Representan las velocidades de los métodos de transporte en Medellín	Al igual que en las respuestas de la pregunta anterior, algunas presentan interpretaciones inadecuadas, como es el caso de la respuesta 3. Otras respuestas evidencian interpretaciones más adecuadas, como la respuesta 2, sin embargo, en particular, la respuesta 1 muestra interpretaciones que, aunque no son del todo correctas, intentan justificar una comprensión de la desviación estándar dentro del contexto. Aunque esta interpretación no es la óptima, sugiere una utilización del contexto de las preguntas previas, específicamente, se infiere que esta respuesta se basa en la

		pregunta 4[13], cuyo objetivo es que los estudiantes determinen la "distancia" entre la media y cada uno de los datos.
Pregunta 7. Teniendo en cuenta las respuestas dadas en los ítems 5 y 6, ¿qué interpretación les da a los signos de los valores obtenidos dentro del contexto establecido? ¿Por qué? Justifique su respuesta.	**Respuesta 1.** Los datos positivos indican cuanto aceleran, y los negativos cuanto desaceleran en velocidad **Respuesta 2.** Que hay datos muy variados con respecto al promedio y que los signos negativos son carros que van por debajo del promedio de velocidad y los positivos para lo que van por arriba de este promedio. **Respuesta 3.** La interpretación es que esa es la diferencia que hay entre esas velocidades ya que si no, no tendría sentido porque una velocidad no puede ser negativa **Respuesta 4.** la interpretación que le doy es que esa es la diferencia que hay entre el promedio y el dato	Las respuestas a esta pregunta mostraron una considerable variabilidad, con interpretaciones adecuadas en algunos casos como la respuesta 1 e interpretaciones inadecuadas en otros, como la respuesta 2. A excepción de un estudiante que no respondió la pregunta, todas las respuestas, independientemente de la calidad de la interpretación, evidenciaron el desarrollo de la característica C6, no obstante, la respuesta 3 presentó una particularidad, ya que, además de cumplir con la característica mencionada, también abordó la característica C8: *Explicar por qué un resultado matemático o una conclusión tiene o no sentido en el contexto de un problema*. En esta respuesta, el estudiante no solo interpretó el resultado en su contexto original, sino que también realizó un análisis de los valores negativos fuera de dicho contexto, adoptando un enfoque más matemático-físico al afirmar que "una velocidad no puede ser negativa". A pesar de ello, mantuvo la coherencia de su análisis con el contexto inicial, logrando identificar que ciertos resultados carecen de sentido en dicho marco. En esta pregunta, 2 estudiantes no la respondieron, 4 ofrecieron respuestas correctas, 3 presentaron respuestas incorrectas y una respuesta (Respuesta 4) no fue clasificada en ninguna de las categorías, dado que no se evidencia la relación con el contexto.

[13] El Sistema Inteligente de Movilidad de Medellín (SIMM) está interesado en determinar si los conductores mantienen velocidades cercanas al promedio, lo que sugeriría una ausencia de cambios bruscos en la velocidad, o si, por el contrario, hay fluctuaciones de velocidad muy marcadas. ¿Cómo se puede determinar si cada conductor mantiene velocidades cercanas al promedio? Proponga y justifique un procedimiento matemático.

Al igual que una de las respuestas de la pregunta 7 que abordó de manera natural dos características de la competencia interpretativa, algunas de las preguntas se diseñaron con el objetivo de explorar más de una característica, como es el caso de las preguntas 5, 6 y 10. Estas tres preguntas comparten la característica C6, además de incluir otra característica adicional que también pertenece a la categoría *Relación entre lo matemático y lo contextual*. En las Tablas 36, 37 y 38 se presenta un análisis de cada una de las preguntas, junto con ejemplos de las respuestas que evidencian las respectivas características.

Tabla 36. *Análisis de la pregunta T2 - 5*

Pregunta 5:	¿Cuáles son las unidades de medida correspondientes a los datos obtenidos en el punto anterior?	
Características:	**C6:** Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.	**C2:** Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013).
Ejemplo de respuesta:	Las unidades de medidas son Km/H	
Análisis:	En esta pregunta, se evidencian ambas características en todas las respuestas presentadas. En particular, en el ejemplo mencionado, es claro que, para que el individuo determine la unidad de medida de los datos es imprescindible realizar una lectura de estos, lo que resalta la característica C2. Asimismo, tras llevar a cabo el procedimiento solicitado en la pregunta anterior (pregunta 4), el estudiante debía en esta pregunta, analizar e interpretar el resultado obtenido para verificar que la unidad de medida permaneciera consistente con la previamente establecida durante la lectura de los datos, lo que pone de manifiesto la característica C6. En este caso, uno de los estudiantes no respondió a la pregunta, mientras que las demás respuestas fueron correctas.	

Tabla 37. *Análisis de la pregunta T2 - 6*

Pregunta 6:	Es posible establecer una interpretación para las magnitudes obtenidas en el punto 4, ¿Qué podría establecer al respecto? ¿cuál es el significado de estos valores numéricos teniendo en cuenta el contexto dado?	
Características:	**C6:** Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.	**C7:** Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados.

Respuesta:	Se puede establecer que el resultado es la diferencia entre los datos y el promedio, con respecto al contexto se puede decir que esa es la diferencia entre las velocidades que hay entre la velocidad del conductor y el promedio de las velocidades.
Análisis:	En esta pregunta, aunque no todas las respuestas son correctas ambas características se evidencian en la mayoría de las respuestas analizadas. En particular, en el ejemplo citado, es evidente que el individuo interpreta el resultado de una pregunta anterior dentro del contexto específico en el que se desarrolla, lo que resalta la característica C6. Asimismo, la característica C7 se manifiesta cuando el individuo establece un paralelismo entre la interpretación matemática de los resultados y su interpretación contextual. Esto se debe a que, en una de las preguntas previas, el resultado matemático se deriva de una problemática planteada en dicho contexto, lo que pone de relieve cómo la problemática influye en los resultados obtenidos. En este caso, 1 estudiante no respondió, 5 proporcionaron respuestas correctas y 4 presentaron respuestas incorrectas. De estas últimas, dos únicamente evidenciaron la característica C6, ya que, a pesar de su interpretación en el contexto, no lograron identificar la relación de implicación entre el contexto y los resultados matemáticos. Un ejemplo de esto es la siguiente respuesta: "Que todas están relacionadas con la velocidad."; en esta se observa una relación entre el resultado y un aspecto del contexto, sin embargo, más allá de un nivel de interpretación superficial, carece de los elementos necesarios para evidenciar la característica C7.

Tabla 38. *Análisis de la pregunta T2 - 10*

Pregunta 10:	¿Cuáles son las unidades de medidas correspondientes al resultado obtenidos en el punto anterior? Mencione si es coherente respecto a la situación planteada y justifique.	
Características:	**C6:** Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real .	**C8:** Explicar por qué un resultado matemático o conclusión tiene o no sentido dado el contexto de un problema.
Ejemplo de respuesta:	**Respuesta 1**. La unidades de medida serian KM^2/H^2, no es coherente ya que serían medidas de área **Respuesta 2.** Km/h, porque solo se usó la función de valor absoluto y son unidades de velocidad.	
Análisis:	En esta pregunta, aunque no todas las respuestas son correctas ambas características se evidencian en la mayoría de las respuestas analizadas. En	

particular, en la *Respuesta 1*, es evidente que el individuo interpreta el resultado de una pregunta anterior dentro del contexto específico en el que se desarrolla, lo que resalta la característica C6. Asimismo, la característica C8 se manifiesta cuando el individuo expresa y justifica la incoherencia de las unidades de medida del resultado.

Por su parte la *Respuesta 2*, al igual que la *Respuesta 1* contiene los elementos necesarios para clasificarla en las características C6 y C8, no obstante, destaca por una particularidad: en el ítem 8[14], es el único individuo que recurre al uso del valor absoluto, lo cual diferencia su respuesta del resto. Este aspecto fue previsto durante la formulación de las tareas, ya que, las características de la interpretación previstas fueron abordadas.

En este caso, 1 estudiante no respondió, 6 proporcionaron respuestas correctas y 3 presentaron respuestas incorrectas. En estas últimas, solo se evidenció la característica C6, ya que, a pesar de la interpretación contextual, no lograron expresar ni justificar la incoherencia en las unidades de medida del resultado, lo que impidió evidenciar la característica C8. Un ejemplo de ello es la respuesta: "Km/h cuadrados".

Otra de las preguntas que también aborda la característica C8 es la **Pregunta 11: A partir del valor numérico obtenido (punto 9) junto con las unidades de medida reportadas en el punto 10 ¿Es posible establecer una interpretación adecuada y que corresponda al contexto planteada? De no ser así, utilice alguna operación o propiedad matemática para asegurar que las unidades de medida se puedan interpretar en correspondencia con la situación planteada. Mencione que propiedad u operación matemática utilizó y porqué.**

Exceptuando un estudiante que no proporcionó respuesta a la pregunta, ocho de los restantes contestaron de manera adecuada, mencionando y justificando el uso de operaciones o propiedades matemáticas que explicaban por qué el resultado obtenido carecía de sentido en el contexto planteado, lo que evidencia la competencia C8: Explicar por qué un resultado matemático o una conclusión tiene o no sentido en el contexto de un problema. En la Figura 35, en color verde se presentan las ocho respuestas que muestran lo expuesto.

[14] Ítem 8: Si hay datos para los cuales el signo de los valores numéricos no permita una interpretación adecuada y acorde al contexto, utilice para el conjunto de datos alguna operación o propiedad matemática para asegurarse de que los signos estén en coherencia con la situación planteada. Mencione que propiedad u operación matemática utilizó.

Figura 3533. *Respuestas de los estudiantes T2-11.*

No, no tiene sentido. Para que tenga sentido es necesario sacarle la raíz cuadrada. Sería 13,28
le sacó raíz cuadrada para despejar el cuadrado
No, no tiene sentido. Para que tenga sentido es necesario sacarle la raíz cuadrada. Sería 13,28
tocaria sacar raiz para eliminar el a la 2
La propiedades que aproveche fueron la de los exponentes y raíces utilice el exponente para ponerlos positivos y las raíces para deshacerme del exponente.
El numero elevado al cuadrado de dif con el promedio, pero este dato va a ser mas claro pues no va a estar con numeros negativos
N/A
No es posible ya que son medidas de área, al número resultante se le saca raíz y resulta en km/H
Si, porque no se hizo muchos cambios en el valor ni las medidas de velocidad entonces si se puede establecer
No es posible ya que el contexto pide otra cosa por que hubieron medidas que dieron negativas.

Particularmente, la respuesta de color morado que se muestra en la Figura 35 se distingue de las demás al mencionar la posibilidad de interpretar de manera adecuada las unidades de medida, lo cual es coherente con la respuesta proporcionada por el mismo individuo en la pregunta 10, como se evidencia en la *Respuesta 2* de la Tabla 38. En dicha respuesta, se hacía referencia al uso del valor absoluto, lo que implica que las unidades de medida no se modifican (km/h), siendo claramente interpretables. Esto justifica de manera adecuada la respuesta a la pregunta 11 manteniendo la coherencia entre sus respuestas, además de mantener los elementos necesarios para clasificarla bajo la característica C8.

Las dos últimas preguntas, clasificadas en la categoría *Relación entre lo matemático y lo contextual*, corresponden a las preguntas 4 y 9, cuyas respuestas se asociaron exclusivamente a la competencia C7: Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para realizar juicios contextuales sobre cómo deben ajustarse o aplicar los resultados. La Tabla 39 expone el análisis de las respuestas correspondientes, incluyendo una de las respuestas de la pregunta evaluada.

Tabla 39. *Análisis de las preguntas 4 y 9 de T2.*

Pregunta 4. El Sistema Inteligente de Movilidad de Medellín (SIMM) está interesado en determinar si los conductores mantienen velocidades cercanas al promedio, lo que sugeriría una ausencia de cambios bruscos en la velocidad, o si, por el contrario, hay fluctuaciones de velocidad muy marcadas.	**Pregunta 9.** El Sistema Inteligente de Movilidad de Medellín (SIMM) quiere presentar un informe para los habitantes de Medellín y, dentro de este, desea sintetizar la información obtenida sobre la variable km/h en el punto 6, para presentarla como un solo valor numérico. ¿Qué procedimiento realizaría para

¿Cómo se puede determinar si cada conductor mantiene velocidades cercanas al promedio? Proponga y justifique un procedimiento matemático.	poder sintetizar esta información? Realice este proceso en Excel.
Respuesta: Restando el promedio con el dato de cada conductor, asi se vera la diferencia de Velocidad que tienen con respecto al promedio.	**Respuesta:** Usaría la Media, o promedio. Pondría la función en excel y calcular el promedio de los números del punto 8.
Análisis: Para la pregunta 4, con excepción de una, las respuestas evidencian los elementos necesarios para ser clasificadas bajo la competencia C7. De manera similar, en la pregunta 9, tres estudiantes no respondieron y una de las respuestas no es clasificable debido a problemas de redacción; el resto de las respuestas pueden ser clasificadas bajo esta misma categoría. Esto se evidencia en las respuestas previamente expuestas, donde se justifica el uso de un proceso matemático a partir de una problemática contextual, comprendiendo que el propio contexto influye y condiciona los resultados matemáticos o los cálculos de un procedimiento.	

Por otro lado, las únicas dos preguntas que no se enmarcan en la categoría *Relación entre lo matemático y lo contextual* fueron las preguntas 13 y 15, las cuales se clasificaron en las categorías *Sistemas de representación* y *Tendencias y patrones (Predicción o conjeturas y sesgos)*, respectivamente. A continuación, se realiza un análisis similar al de las preguntas previas.

Pregunta 13. El Sistema Inteligente de Movilidad de Medellín (SIMM) planea llevar a cabo un estudio similar el próximo año, sin embargo, con el objetivo de simplificar el proceso, se busca desarrollar una fórmula que sea aplicable independientemente de los datos recopilados. Los pasos para realizar este trabajo se muestran a continuación, completa la tabla para finalizar la deducción de la fórmula deseada por el Sistema Inteligente de Movilidad de Medellín (SIMM).

Como se mencionó anteriormente, las respuestas a esta pregunta están enmarcadas en la categoría *Sistemas de representación,* siendo un claro ejemplo de la característica C3: Identificar, analizar, comprender e interpretar una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013). Esto se puede evidenciar en la Figura 36 correspondiente a una de las respuestas a esta pregunta, en esta se evidencia análisis, interpretación y transformación entre dos sistemas de representación, la correspondiente al lenguaje literal (columna “Pasos”) y el sistema simbólico. Dos de los estudiantes no respondieron esta pregunta, el resto de ellos la realizó correctamente.

Figura 3634. *Respuesta de los estudiantes T2-13.*

Pasos	Formula planteada
a) Si se tiene en cuenta que la media es $\underline{x}$, se resta la media a cada valor (x_i) del conjunto de datos.	$x_i - \underline{x}$
b) Se eleva al cuadrado cada diferencia.	$(X_i - \underline{x})^{\square^2}$
c) Se suman todas las diferencias al cuadrado.	$\sum_{i=1}^{n} \square (x_i - \underline{x})^2$
d) Se divide la suma obtenida entre el número total de datos (n).	$\frac{\sum_{i=1}^{n} \square (x_i - \underline{x})^2}{n}$
e) Por último, se calcula la raíz cuadrada del paso anterior.	$\sqrt{\frac{\sum_{i=1}^{n} \square (x_i - \underline{x})^2}{n}}$

Pregunta 15. Al usar los comandos de varianza y desviación, ¿los valores numéricos obtenidos son iguales a los calculados con la fórmula?

Como se indicó previamente, las respuestas a esta pregunta se enmarcan en la categoría *Tendencias y patrones (Predicción o conjeturas y sesgos)*, constituyendo un claro ejemplo de la competencia C11: Usar el pensamiento matemático y el pensamiento computacional para hacer predicciones, proporcionar evidencia para argumentos, probar y comparar las soluciones propuestas.

Las respuestas a esta pregunta fueron variadas; dos estudiantes omitieron responder, otros dos proporcionaron respuestas incorrectas y seis de ellos respondieron de manera correcta. Los dos estudiantes cuyas respuestas fueron incorrectas cumplen con los criterios de la característica C11, ya que utilizaron el instrumento para probar y comparar soluciones. No obstante, este aspecto no será objeto de mayor análisis, ya que el interés principal radica en las respuestas correctas, como se evidencia en la Figura 37

Figura 3735. *Repuesta de estudiante a T2 -15.*

Si. Al calcular con formula en excel fue más largo que con la función pero al comparar resultados, los resultados fueron los mismos

Esta respuesta está en correspondencia con lo esperado, dado que presenta una afirmación coherente con el propósito de la tarea, que es la construcción de la fórmula de la desviación

estándar, en consecuencia, los dos resultados solicitados para comparación en la pregunta deben ser iguales. Asimismo, esta respuesta cumple con los criterios para ser clasificada dentro de la característica C11, por las mismas razones mencionadas en relación con la respuesta incorrecta.

Otra de las respuestas que en este caso no es afirmativa, pero de igual modo correcta es (Figura 38):

Figura 3836. *Respuesta de estudiante #2 a T2 -15.*

No da lo mismo ya que dependiendo del comando que se use sale un número un tanto diferente

Esta resulta particularmente interesante, ya que lo esperado era una afirmación clara, sin embargo, la explicación se fundamenta en la distinción entre los comandos para calcular la desviación estándar de la población y de la muestra, como se ve en la Figura 39 (tomada de la base de datos del estudiante que dio esta respuesta), lo que lleva a una conclusión que, aunque no es afirmativa, sigue siendo válida y precisa. Al igual que las dos respuestas anteriores, esta también se puede categorizar en la característica C11.

Figura 3937. *Desviación estándar muestral vs poblacional.*

Desvest.p	Desvest m
7,47348604	7,479639591

En la Tabla 40, se presenta un resumen de las características asociadas a las respuestas de los estudiantes para cada una de las preguntas planteadas en la Tarea #2.

Tabla 40. *Características de la Tarea 2*

Categoría / Pregunta	N/A	C2	C3	C6	C7	C8	C11
1	X						
2	X						
3				X			
4					X		
5		X		X			
6				X	X		
7				X		X	
8	X						
9					X		
10				X		X	
11						X	
12				X			

13			X				
14	X						
15							X
16				X			

4.3. Tarea 3

En analogía con las Tarea 1 y 2, es relevante destacar que algunas preguntas de la Tarea 3 no fueron diseñadas para abordar directamente las categorías de análisis relacionadas con la competencia interpretativa

Es relevante señalar que, de los 11 estudiantes participantes inicialmente, 5 no realizaron la actividad y 2 la completaron parcialmente, respondiendo únicamente hasta las preguntas 9 y 10 respectivamente. Todo esto estuvo condicionado por las diversas restricciones y eventualidades de los tiempos institucionales, lo cual influye directamente en el análisis que se desarrolla a continuación, donde se considerarán las categorías de la competencia interpretativa.

Pregunta 2. Considerando el diagrama de caja y el diagrama de dispersión presentados en la base de datos, ¿cómo describiría el desempeño de los estudiantes durante cada uno de los tres cortes?

Considerando que la pregunta estaba diseñada para que el estudiante realizara una interpretación de los gráficos estadísticos presentados, se pudo categorizar de manera precisa las respuestas en la categoría de *Sistemas de Representación*. En particular, se ubicaron en las características C3. Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013) y C5. Interpretar información presentada en forma gráfica o diagramas.

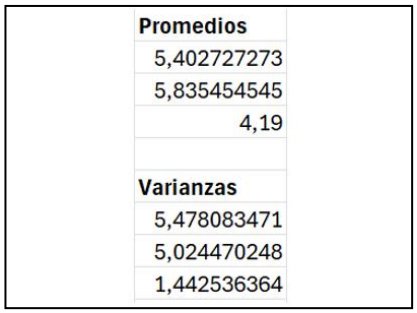
En la Respuesta 1 de la Tabla 41, se evidencia una interpretación adecuada y relativamente detallada de la situación, teniendo en cuenta los gráficos estadísticos, haciendo referencia explícita a las características estadísticas de los mismos. El estudiante utiliza las medidas estadísticas presentes en los gráficos, así como el contexto en el que se desarrollan, para emitir una opinión crítica sobre el rendimiento de los estudiantes. En contraste, la Respuesta 2 de la misma tabla refleja una interpretación más superficial, limitada a una lectura básica de la dispersión de los datos, sin integrar elementos adicionales que sustenten o amplíen su análisis.

Tabla 41. *Respuestas de los estudiantes T3-2*

	Corte I	Corte II	Corte III
Respuesta 1:	Muy Mal desempeño el 75% de las notas no pasan de 7	Mejorar el desempeño con respecto al Corte 1 el Q2 ya está en 6 y se acerca más a pasar, igual la mitad o más de estudiantes perdieron.	Es el peor de todos, solo 4 Estudiantes pasaron este corte, el resto tuvo notas de 5 e inferiores y el promedio pasa hasta 4,19.
Respuesta 2:	Las notas de los estudiantes son muy diferentes, es decir tuvieron algunos desempeño bajo otros medio y otros alto está muy variado.	Las notas están menos dispersas a comparación del corte 1, sin embargo siguen estando diferentes las unas de las otras	Las notas del tercer corte están más unificadas es decir los estudiantes tuvieron un desempeño parecido y constante no hay gran diferencia entre unos y otros a excepción de unos pocos que sí están más alejados

Pregunta 3. ¿Durante cuál de los cortes los estudiantes fueron más constantes con sus calificaciones? Justifique su respuesta.

Figura 4038. *Medidas estadísticas a partir de funciones en Excel.*



Promedios
5,402727273
5,835454545
4,19
Varianzas
5,478083471
5,024470248
1,442536364

Como se evidencia en las respuestas marcadas en color verde presentadas en la Figura 41, estas se pueden clasificar dentro de las mismas características C3 y C5, dado que los estudiantes efectúan una interpretación de los gráficos estadísticos incluidos en la base de datos. Sin embargo, destaca una respuesta en particular, aquella sombreada de color morado que a diferencia de las demás, no se enfoca en la interpretación de los gráficos estadísticos, sino que parece fundamentarse en la experiencia previa del estudiante al medir la dispersión de los datos mediante el cálculo de medidas de dispersión utilizando funciones de Microsoft Excel, tal como se ilustra en la Figura 40.

Posteriormente, el estudiante ofrece una respuesta sustentada en la interpretación de medida estadística obtenida, por lo tanto, esta respuesta no se puede clasificar junto con las restantes, sino que corresponde a la categoría de *Relación entre lo Matemático y lo Contextual*, desarrollando así la característica C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.

Cabe aclarar que esta respuesta surgió de manera natural, sin una intervención por parte del docente investigador, por el contrario, parece que este estudiante utiliza el esquema de uso que se relaciona con la interpretación de conjunto de datos a través de la representación de estos por medio de gráficos estadísticos correspondiente al proceso de instrumentación.

Figura 4139. *Respuestas de los estudiantes T3-3.*

Las notas más constantes fueron las del 3er corte ya que todos tienen notas muy bajas, solo muy pocos tienen buenas notas.
Durante el corte 3 debido a que fue donde sacaron las notas mas cercanas entre todos
Durante el III Corte porque aunque fueron las peores notas, la mayoría de estudiantes sacaron notas entre 3,40 y 4,60-.
El corte donde las notas fueron más constantes fueron en el corte III y que en el gráfico se ve como los puntos están unificados y que las notas no tienen un promedio de notas muy alejado entre ellos.
En el tercer corte los estudiantes fueron muchísimo más constante con sus calificaciones, sacando la varianza de cada uno de los cortes el menor valor de varianza sale del 3 corte osea que los datos varían menos y los estudiantes fueron más constantes.
En el tercer corte, aunque no les fue muy bien fueron constantes con las entregas ya que en el diagrama de dispersión las notas están bastante pegadas

Otras preguntas que por su parte desarrollan la misma característica y por ende clasificados dentro de la misma categoría, son las presentadas en la Tabla 42:

Tabla 42. *Análisis de preguntas 4 y 5 de T3*

Pregunta 4. ¿Cómo varió la dispersión de las calificaciones entre el corte 1 y el corte 2? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.	**Pregunta 5.** ¿Cómo varió la dispersión de las calificaciones entre el corte 2 y el corte 3? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.
Respuesta a la pregunta 4: Segun el grafico del corte 1 yo pienso que la varianza es mayor eso hace que las notas y los datos atipos sean menos probables pero al mismo tiempo terminan siento menos contantes y con mas notas extremistas, encambio en el grafico del corte 2 la varianza es muy parecida pero terminan habiendo menos datos atipicos y extremistas.	**Respuesta a la pregunta 5:** La dispersión de las calificaciones entre el corte 2 y 3 disminuyo bastante, los estudiantes fueron mucho más constantes con sus entregas durante el tercer corte
Análisis: Las dos preguntas estaban orientadas a comparar la dispersión de dos conjuntos de datos, considerando tanto la gráfica de dispersión como el diagrama de caja correspondientes. Todos los estudiantes respondieron de manera adecuada a ambas preguntas, llevando a cabo un análisis e interpretación de la dispersión de cada uno de los gráficos antes de proceder a la	

comparación, no obstante, las justificaciones respecto a las variaciones en la dispersión entre los cortes mostraron diferencias significativas. Por ejemplo, en la respuesta a la pregunta 4, el estudiante fundamenta su razonamiento a partir de lo que considera datos atípicos en el contexto y la homogeneidad de los datos, en contraste, la respuesta a la pregunta 5 justifica el cambio de la variabilidad entre los cortes, teniendo en cuenta variables externas que podrían haber influido en el rendimiento de los estudiantes. Independientemente de la exactitud de las respuestas o de la justificación empleada, se confirma que todas las respuestas a estas dos preguntas se pueden clasificar dentro de las características C3 y C5 de la categoría *Sistemas de Representación*.

Como era previsible, algunos estudiantes podrían haber considerado diferentes variables para justificar la variabilidad observada. Por esta razón, la siguiente pregunta aborda específicamente esta situación particular.

Pregunta 6. ¿Cuántas y cuáles variables considera usted que influyen en las calificaciones de los estudiantes? ¿Estas variables cambian durante los tres cortes o se mantienen constantes? ¿Por qué piensa eso?

Las respuestas señaladas en color verde presentadas en la Figura 42, se clasifican en la categoría de *Relación entre lo Matemático y lo Contextual*, desarrollando la característica *C7. Comprender cómo el mundo real impacta los resultados y los cálculos de un procedimiento o modelo matemático, para hacer juicios contextuales sobre cómo deben ajustarse o aplicarse los resultados*. Esto debido a que los individuos identifican y mencionan claramente algunas variables que pueden afectar el rendimiento de los estudiantes y responden de una forma explícita o implícita si estas se mantienen constantes o cambian al pasar determinado tiempo.

Por otro lado, las respuestas marcadas en color naranja en la Figura 42, no se pueden categorizar dentro de ninguna de las características establecidas. La primera, " considero que hay 4 variables, horario del sueño, concentración, conocimiento del tema y necesidad de buena nota.", si bien reconoce la existencia de otras variables, no es explícita en cuanto a cómo estas impactan el rendimiento del estudiante, ni aclara si se mantienen constantes a lo largo de los distintos periodos evaluativos. La segunda respuesta, "2 variables, una serían los exámenes y la otra el acumulado para el año. Se mantienen constantes durante los 3 cortes, porque los estudiantes quieren pasar el año y la materia, ya que esto les beneficia para las universidades o para entrar al siguiente curso en otro colegio." no evidencia una verdadera identificación de variables, en su lugar, parece referirse a los componentes que forman parte del sistema de evaluación del estudiante, los cuales,

aunque podrían interpretarse como variables en un sentido amplio, no son explicitados en la respuesta.

***Figura 4240**. Respuestas de los estudiantes T3-6.*

considero que hay 4 variables, horario del sueño, concentración, conocimiento del tema y necesidad de buena nota.
N/A
Hay varias, cansancio pues ya las notas al final bajan mucho, esto también puede estar relacionado con que el colegio está exigiendo más temas en el último corte que en el primero pues se ve que al final ya no se acuerdan mucho de estos.
●El nivel de dedicación que el estudiante tome frente a la clase ●la entrega de tareas ●el desempeño de los quizzes las tres variables podrían variar a medida de los cortes ya que es algo personal del estudiante y algo que puede cambiar para que su desempeño mejore o empeore. pienso eso ya que creo que las notas no solo dependen de qué tan inteligente sea un estudiante, si no que también importa la dedicación y el tiempo que le invierta a estudiar y hacer los trabajos bien, de esta forma si los habitos del estudiante cambian a medida de que los cortes pasan sus notas también se verían afectadas junto a su rendimiento.
Las variables que creo que influyen son 1.Nivel de estudio (osea si estudia para el examen o no) 2.Comprensión de los temas 3.Atentividad a clase 4.Dificultad de los exámenes 5.Estado de salud mental Estas variables no se mantienen del todo constantes en los tres cortes, como podemos ver de el primer al segundo corte se ve cómo estas variables “mejoran” así dando paso a un mejor desempeño académico para los estudiantes, pero para el tercer corte estas variables pueden estar peores por que el desempeño decae un montón entonces en los tres cortes cambia la situación de los estudiantes.
2 variables, una serían los examénes y la otra el acumulado para el año.Se mantienen constantes durante los 3 cortes, porque los estudiantes quieren pasar el año y la materia, ya que esto les beneficia para las universidades o para entrar al siguiente curso en otro colegio.

Pregunta 7. Teniendo en cuenta el contexto, que representaría un dato atípico dentro de este.

Las respuestas señaladas en color verde presentadas en la Figura 43, se clasifican en la categoría de *Tendencias y patrones (Predicción o Conjeturación y Sesgos)*, desarrollando la característica C14: Evaluar e interpretar el impacto de los valores atípicos en los resultados. (Franklin et al., 2020). Esto debido a que los individuos identifican y mencionan claramente algunas variables que pueden afectar el rendimiento de los estudiantes y responden de una forma explícita o implícita si estas se mantienen constantes o cambian al pasar determinado tiempo.

Por otro lado, las respuestas marcadas en color naranja en la Figura 43 no se pueden categorizar dentro de ninguna de las características establecidas. La primera, "considero que hay 4 variables, horario del sueño, concentración, conocimiento del tema y necesidad de buena nota.”, si bien reconoce la existencia de otras variables, no es explícita en cuanto a cómo estas impactan el rendimiento del estudiante, ni aclara si se mantienen constantes a lo largo de los distintos periodos

evaluativos. La segunda respuesta, "2 variables, una sería los exámenes y la otra el acumulado para el año. Se mantienen constantes durante los 3 cortes, porque los estudiantes quieren pasar el año y la materia, ya que esto les beneficia para las universidades o para entrar al siguiente curso en otro colegio.", no evidencia una verdadera identificación de variables, en su lugar, parece referirse a los componentes que forman parte del sistema de evaluación del estudiante, los cuales, aunque podrían interpretarse como variables en un sentido amplio, no son explicitados en la respuesta.

Figura 4341. *Respuestas de los estudiantes T3-7.*

un dato atípico representaría una nota que está muy alejada de las otras como un diez en sólo unos.
Un dato atipico podría ser un dato que sobre pase o sea mas bajo que Q3 o Q1
Un dato que no está dentro de las líneas de arriba como en este caso el 1,5 7 8 8,5 9 10.
Un dato atípico dentro de esta tabla podría ser que el estudiante como nota final obtenga un cero ya que la nota mínima es 1. o en su defecto algo mayor a diez o incluso eln tener un 10 ya es algo atípico, sin embargo estadísticamente si las cuentas están bien hechas esto no podría ser posible
Un dato atípico en este contexto serían los datos más altos, ósea en los tres cortes los datos atípicos serían los más "escasos" en este caso sería la nota más altas que serían de un estudiante que se preparó y tiene un mejor desempeño que los demás de cada corte, o tambien podria ser el dato más bajo de cada corte osea el dato más bajo de todos y en este caso seria un estudiante con menor desempeño positivo en lo académico osea que le debe faltar preparación para los examen sy por eso su nota es considerablemente más baja.
La dispersión en las notas del corte 3

Las preguntas 8, 9, 10 y 11 (ítems a, b y c) se clasificaron dentro de la categoría de *Sistemas de Representación*, específicamente se ubicaron en las características C3: identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural. (Moreno y Villanueva, 2013) y en C5: interpretar información presentada en forma gráfica o diagramas. A continuación, se muestra la Tabla 43 con el análisis de estas preguntas.

Tabla 43. *Análisis de las preguntas 9, 10 y 11 de T3.*

ITEMS	RESPUESTAS Y ANÁLISIS
Pregunta 8. Según el diagrama de cajas, ¿hay datos atípicos en alguna de las cajas? ¿Para cuál o cuáles cortes?	**Respuesta 1.** En la mayoría ahi debido a los datos tan pequeños que ahí. **Respuesta 2.** En la última caja hay 5 datos atípicos **Respuesta 3.** Logró identificar dos datos que podría considerar atípicos los cuales son algunos estudiantes con promedio de 10 y de 1 ya que para obtener 10 tendría que haber perfección del estudiante por parte de tareas quizzes el examen final y el comportamiento, y para obtener 1 sería todo lo contrario es decir no hacer nada y tener un mal comportamiento.. Hay dentro de los 3 cortes notas de 1 y de 10 en el primero y el segundo.

	Análisis: La respuesta 1 no es clasificable, ya que, aunque se puede inferir que el estudiante comprende el concepto de dato atípico, no existe evidencia clara de que haya realizado una lectura adecuada del gráfico. Por el contrario, las respuestas 2 y 3 se clasifican en las categorías C3 y C5, dado que ambas evidencian una interpretación de los gráficos, sin embargo, solo la respuesta 2 es correcta, aunque carece de una justificación explícita. En esta, se identifica y verifica de manera precisa la cantidad de datos atípicos, reconociéndolos como un componente característico del diagrama de caja. En cuanto a la respuesta 3, no logra identificar con claridad los datos atípicos, limitándose a extraer conclusiones de los gráficos estadísticos que podrían estar o no relacionadas con dichos datos, por lo tanto, se puede inferir que esta respuesta es incorrecta, no por una interpretación inadecuada del gráfico, sino debido a un desconocimiento de la estructura y los componentes fundamentales del mismo. Este desconocimiento puede haber llevado a que el estudiante no identificara correctamente los datos atípicos y no comprendiera cómo estos se representan en el diagrama de caja, lo que afectó su capacidad para realizar una lectura precisa del gráfico. Se registraron tres respuestas correctas, dos respuestas incorrectas y una respuesta que no pudo clasificarse.
Pregunta 9. ¿Como interpreta los datos atípicos encontrados respecto al corte correspondiente?	**Respuesta 1.** Que 5 estudiantes se separando muchísimo de los promedios del curso **Respuesta 2.** Por cada caja hay dos datos atípicos, uno que es la nota máxima de los exámenes de ese corte y el 2 que es la nota mínima de los exámenes en el corte. En el 1 corte la máxima nota es de 10.0 y la mínima nota es de 1.0 y estos son los datos atípicos de este corte. En el 2 corte los datos atípicos son 9,1 como nota máxima y 1.0 como la nota más baja obtenida en los exámenes de ese corte. y por último en el 3 corte la nota más alta obtenida que es un dato atípico es 5.0 y la más baja es de 1.5 que también es un dato atípico.
	Análisis: En esta pregunta, todas las respuestas pudieron clasificarse dentro de las características C3 y C5; sin embargo, no todas fueron correctas, por ejemplo, en la respuesta 1, el estudiante demuestra una comprensión adecuada de los datos atípicos al interpretar correctamente los gráficos estadísticos. En contraste, aunque la

	respuesta 2 también muestra una interpretación de los gráficos, el estudiante confunde los datos atípicos con los valores máximos y mínimos de cada conjunto de datos, lo que lleva a una respuesta incorrecta. Un estudiante no respondió la pregunta; de los restantes, tres respondieron de manera correcta, mientras que dos presentaron respuestas incorrectas.
Pregunta 10. ¿Como interpreta estos mismos datos respecto a las calificaciones de los tres cortes?	**Respuesta 1.** Datos que igual estaban afuera de la caja, pero más cercanos al promedio del grupo **Respuesta 2.** Se podría interpretar de que la mayoría podría interpretar de que los datos de ciertos cortes son mas desiguales que las de los otros **Análisis:** A partir de esta pregunta, dos estudiantes no continuaron con el resto de la tarea, por lo que el análisis se centra en las respuestas de los cuatro estudiantes restantes. Tres de ellos no lograron proporcionar respuestas correctas, como se observa en la respuesta 1, donde se realizó una interpretación de la dispersión de los datos a partir del gráfico, aunque no se puede afirmar con certeza que dicha interpretación haya sido basada en los datos atípicos identificados. En contraste, uno de los estudiantes respondió de manera precisa, como se refleja en la respuesta 1, donde realiza una interpretación clara de los datos atípicos en relación con los resultados de los otros cortes académicos. Es preciso aclarar que todas las respuestas, independientemente de su veracidad, se pudieron clasificar dentro de las categorías C3 y C5.
Pregunta 11a. ¿Qué información se puede visualizar en el diagrama de cajas que no es posible ver en el histograma? ¿Cree usted que al observar e interpretar únicamente el diagrama de cajas se está ignorando alguna información del conjunto de datos?	**Respuesta 1.** En el diagrama de caja se puede ver el promedio del corte, en cambio en el histograma no. Si, se estarían ignorando la cantidad de alumnos que pasarón el corte. **Respuesta 2.** El promedio y la Dispersión **Análisis:** En este caso, tres de los cuatro estudiantes respondieron correctamente a la pregunta, ya que lograron identificar elementos visualizables en el diagrama de cajas que no son evidentes en el histograma. Asimismo, también reconocieron aspectos que pueden pasarse por alto al observar únicamente el diagrama de cajas; un ejemplo de esto se encuentra en la respuesta 1. Por otro lado, la respuesta 2 presenta una respuesta parcial a la pregunta, al identificar que el diagrama de cajas permite observar el promedio

	y la dispersión, pero no responde a la parte de la pregunta que requiere identificar aspectos que pueden ignorarse al visualizar únicamente el diagrama de cajas. Las cuatro respuestas proporcionadas por los estudiantes pueden clasificarse dentro de las categorías C3 y C5, ya que se evidencia una interpretación de los gráficos estadísticos en relación con el reconocimiento de las posibilidades y limitaciones de los diagramas de caja e histogramas.
Pregunta 11b. ¿Qué información se puede visualizar en el histograma que no es posible ver en el diagrama de caja? ¿Cree usted que al observar e interpretar únicamente el histograma se está ignorando alguna información del conjunto de datos?	**Respuesta 1.** El hisotograma muestra la distribución de las frecuencias de los datos en intervalos específicos y esto ni nos lo muestra el diagrama de caja y de nuevo, si solo observamos este diagrama no es que estes ignorando al de cajas sino que depende mucho de lo que queiras analizar a profundidad. Entocnes si quieres un analisis mas preciso mira los dos diagramas pero con el hisotograma ya puedes ver varias cosas que podrian ser suficientes. **Respuesta 2.** La cantidad de alumnos que pasarón los examénes en el corte. Si, se estaría ignorando el promedio de los estudiantes.
	Análisis: En este caso, solo uno de los cuatro estudiantes respondió correctamente a la pregunta (respuesta 2), al lograr identificar los elementos visibles en el histograma que no son perceptibles en el diagrama de cajas, además de reconocer aquellos aspectos que pueden omitirse al analizar únicamente el histograma. Por otro lado, la respuesta 2 ofrece una respuesta parcial, al señalar que el histograma permite visualizar la distribución de frecuencias de los datos en intervalos específicos, pero no aborda la parte de la pregunta que requiere identificar los aspectos que pueden pasarse por alto al observar únicamente el histograma. No obstante, el estudiante menciona que, para un análisis más preciso, es necesario considerar ambos gráficos (diagrama de cajas e histograma), lo cual, de manera implícita, refleja el propósito de la pregunta. Las cuatro respuestas proporcionadas por los estudiantes pueden clasificarse dentro de las categorías C3 y C5, ya que se evidencia una interpretación de los gráficos estadísticos en relación con el reconocimiento de las posibilidades y limitaciones de los diagramas de caja e histogramas.
Pregunta 11c. ¿Qué beneficios o dificultades puede traer observar simultáneamente e	**Respuesta 1.** Trae solo beneficios pues se pueden interpretar distintas cosas como frecuencias o dispersión.

interpretar varios diagramas del mismo conjunto de datos?	**Respuesta 2.** Los beneficios son que puedes ver toda la información que necesitas y no tienes que buscar los datos que necesitas para poder escribir una conclusión.
	Análisis: Tal como Ase evidencia en las respuestas 1 y 2, los cuatro estudiantes coinciden en que la utilización de múltiples gráficos aporta un mayor beneficio, dado que incrementa la cantidad y diversidad de la información disponible sobre los datos. Análogo a las respuestas de las dos preguntas anteriores, las respuestas proporcionadas por los estudiantes se pueden clasificar en las categorías C3 y C5, ya que reflejan una interpretación de los gráficos estadísticos mencionados.

Los cuatro ítems (a, b, c y d) del punto 12 no fueron concebidos específicamente para abordar las categorías de análisis asociadas a la competencia interpretativa, su función principal es introducir el contexto sobre la calidad del agua, que sirve como base para el desarrollo de la segunda parte de la Tarea 3. Por otra parte, los ítems 13, 13b y 14 no fueron diseñados para ser analizados bajo las categorías de la competencia interpretativa, su objetivo es llevar a cabo el cálculo de determinadas medidas estadísticas, sin requerir ningún tipo de análisis o razonamiento adicional respecto a los resultados obtenidos.

Pregunta 13C. ¿Cuál es la unidad de medida obtenida en ambos casos al aplicar la fórmula correspondiente? Justifique su respuesta.

Las respuestas destacadas en naranja y verde de la Figura 44, corresponden a dos de las respuestas a la pregunta, las cuales se clasifican dentro de la categoría *Relación entre lo matemático y lo contextual*, específicamente en la característica C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real. La respuesta resaltada en verde es correcta, ya que el estudiante, en primer lugar, interpreta adecuadamente las unidades de medida de la media y la varianza, para luego analizar e interpretar la unidad de medida resultante de la división entre ambas magnitudes. Esto evidencia una interpretación precisa de los resultados matemáticos en términos del contexto, cumpliendo con los criterios para ser clasificada en la característica mencionada.

Por otro lado, la respuesta resaltada en naranja no es correcta, aunque interpreta y menciona de manera adecuada las unidades de medida para la media y la desviación estándar, no consigue

hacerlo de forma precisa para el coeficiente de variación (división entre ambas magnitudes). No obstante, el estudiante logra realizar una interpretación de los resultados matemáticos en términos del contexto, lo que permite clasificar esta respuesta dentro de la característica C6.

Finalmente, la respuesta destacada en morado en la Figura 44 no se clasifica dentro de la misma característica que las dos respuestas anteriores, dado que no realiza un análisis de las unidades de medida de la media, la desviación estándar ni del coeficiente de variación. En su lugar, se limita a mencionar las unidades de medida de la variable, lo que sugiere que el estudiante ha interpretado únicamente los datos, sin profundizar en el análisis de las medidas estadísticas calculadas previamente en el ítem 13B, por ende, esta respuesta se clasifica en la característica C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013).

Figura 4442. *Respuestas de los estudiantes T3-13C.*

Porcentaje pues es *100
N/A
La unidad de medida para el ph es complicada por que la verdad el ph no tiene unidades, eso se mide con una escala de 0 a 14 donde 7 es neutro, menor a 7 es ácido y mayor es alcalino. pero la temperatura se mide en fahrenheit y en celsius. entonces tanto la varianza como el promedio y la desviación son en las mismas unidades del ph o temperatura
En la del pH no tendrémos medida ya que en el pH no hay medida, en cambio en la temperatura se mediría en °C porque son medidas de temperatura

Otros ítems vinculados a la categoría *Relación entre lo matemático y lo contextual*, pero asociados a las características C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos (Moreno y Villanueva, 2013), C6: Evaluar e interpretar un resultado matemático en función del contexto de un problema del mundo real, y C13: Comparar resultados bajo diferentes condiciones en un experimento (Franklin et al., 2020), corresponden a los tres ítems posteriores a la pregunta anterior, cuyo análisis se detalla en la Tabla 44.

Tabla 44. *Análisis de las preguntas 13d, 13e y 13f de T3.*

PREGUNTA	RESPUESTA	ANALISIS	CATEGORÍA
13d. En general, ¿cuál es la unidad de medida correspondiente al coeficiente de variación?	**1.** porcentaje	Si bien el porcentaje no se considera una unidad de medida, la respuesta hace explícito que las unidades de medida correspondientes a las magnitudes previas al cálculo del	C6

		coeficiente de variación desaparecen en el proceso.	
	2. Para la temperatura sería en este caso en celsius y el ph son solo números que se ubican en una escala	Esta respuesta no evidencia una interpretación de las unidades de medida correspondientes a las medidas estadísticas calculadas, por el contrario, parece que únicamente se interpreta la unidad de medida asociada a la variable estadística y, por lo tanto, a los datos iniciales.	C2
	3. Porcentaje	Si bien el porcentaje no se considera una unidad de medida, la respuesta hace explícito que las unidades de medida correspondientes a las magnitudes previas al cálculo del coeficiente de variación desaparecen en el proceso.	C6
13e. ¿Es posible establecer una comparación entre las dos variables estadísticas a partir del valor numérico obtenido correspondiente al coeficiente de variación? ¿por qué?	**1.** ahora si pues por el coeficiente de variación ya se pueden comparar bien	Esta respuesta refleja que el estudiante posee conocimiento, o al menos una intuición, de que el coeficiente de variación permite comparar dos variables independientemente de sus unidades de medida. En consecuencia, realiza una interpretación adecuada y establece comparaciones de los resultados matemáticos en función del contexto, lo que permite su clasificación dentro de la categoría correspondiente.	C6
	2. Comparar los coeficientes de variaciones de distintas variables nos deja ver cual tiene una dispersión más relativa creo que se diría, y pues las unidades no importan.	Siguiendo los mismos argumentos que en la respuesta anterior, el estudiante realiza una interpretación adecuada y establece comparaciones de los resultados matemáticos en función del contexto, lo que justifica su clasificación dentro de la categoría correspondiente.	C6
	3.	Si se asume que al referirse a "datos diferentes" el individuo está aludiendo a las distintas variables estadísticas, entonces	C2

	No, porque son 2 datos diferentes que se miden de manera diferente.	su interpretación se limita a la unidad de medida correspondiente a la variable estadística y en consecuencia a los datos originales, por lo tanto, esta interpretación justifica su clasificación en la categoría C2.	
13f. ¿Teniendo en cuenta la respuesta anterior, podría establecer cuál de las variables estadísticas presenta una mayor dispersión?	**1.** La Desviacion seria en el caso del PH el 7.72% de la medi	El individuo compara los dos coeficientes de variación y concluye que la variable PH presenta menor variabilidad, lo cual ha sido verificado en la base de datos generada por el propio estudiante. En consecuencia, la respuesta cumple con los criterios necesarios para ser clasificada dentro de la categoría indicada.	C13
	2. El ph tiene una dispersión mayor a la temperatura, esto esta muy claro al mirar la varianza de cada uno y su desviación, los valores del ph son mucho mas elevados, osea que hay mas variabilidad en el ph	Aunque la respuesta no es correcta, dado que compara la varianza y la desviación estándar de dos variables estadísticas con unidades de medida distintas, y mucho menos utiliza el coeficiente de variación, sigue evidenciando un proceso de interpretación de las medidas estadísticas y una posterior comparación entre ellas. Por lo tanto, la respuesta cumple con los criterios para ser clasificada dentro de la categoría indicada.	C13
	3. El pH es la variable que presenta mayor dispersión.	El individuo compara los dos coeficientes de variación y concluye que la variable PH presenta menor variabilidad, lo cual ha sido verificado en la base de datos generada por el propio estudiante. En consecuencia, la respuesta cumple con los criterios necesarios para ser clasificada dentro de la categoría indicada.	C13

Finalmente, se presentan los tres ítems de la última pregunta de la tarea, cuyas respuestas fueron clasificadas dentro de la categoría *Tendencias y Patrones (Predicción o Conjeturas y Sesgos)*, específicamente en la característica C11: Aplicar el razonamiento matemático y computacional para realizar predicciones, sustentar argumentos, validar y comparar las soluciones propuestas. Es relevante destacar que, al igual que en las preguntas anteriores, únicamente tres estudiantes respondieron a esta pregunta. La Tabla 45 muestra el análisis de las respuestas a la pregunta.

Tabla 45. *Análisis de las preguntas 14a, 14b y 14c de T3.*

PREGUNTA	RESPUESTA	ANALISIS	CATEGORÍA
14a. Realice una conjetura sobre cuál es el mínimo y el máximo valor que puede tomar el coeficiente de variación	**1.** El min seria 0% y el max 100%	Esta respuesta demuestra que el individuo es capaz de estimar y conjeturar los valores máximos y mínimos del coeficiente de variación, lo que permite clasificarla dentro de la característica previamente señalada.	C11
	2. El minimo valor que puede tomar la variacion es si todos los datos son iguales por que en este caso la variación nos daria 0 por que no habria dispercion. Mientras que el maximo valor que podria tomar seria dependiendo de que tan grande sea la media. Si la media es grande la varianza no sera tsn grande pero si la media es pequeña la varianza puede alcanzar un numero mucho mas alto. Por que habria mas dispercion entre los datos y la media	Esta respuesta también identifica de manera precisa el valor mínimo, aportando además una justificación. Sin embargo, respecto al valor máximo, no presenta una respuesta concreta, sino que desarrolla una explicación que resulta incorrecta. A pesar de ello, dado que se evidencia un intento de predicción sobre los valores extremos del coeficiente de variación, aunque no sea del todo correcta, la respuesta puede clasificarse en la característica C11.	C11
	3. El mínimo valor es el de la conductividad eléctrica y el máximo el el de la demanda química de óxigeno	A diferencia de las dos respuestas anteriores, este individuo no formula una conjetura sobre el tema; en cambio, menciona explícitamente cuáles son los coeficientes de variación más grande y pequeño de la	N/A

		respuesta de la pregunta 14, por lo tanto, esta respuesta no se clasifica en ninguna de las categorías establecidas, ya que no aborda adecuadamente la pregunta formulada.	
14b. ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el mínimo valor?	**1.** Que ninguna parte de la desviacion haga parte del promedio	Esta respuesta carece de claridad en su formulación, lo que impide su clasificación en cualquiera de las categorías establecidas.	N/A
	2. En El conjunto de datos todos los datos tienen que ser iguales por que de esta forma no habria ninguna dispersión o variación entre todos los datos	Esta respuesta expone de forma precisa las condiciones requeridas para que el coeficiente de variación alcance su valor mínimo, acompañada de una justificación adecuada, por lo tanto, se puede clasificar plenamente en la categoría C11.	C11
	3. Todos los valores del conjunto deben valer lo mismo, y con estás condiciones el coeficiente es 0.	Siguiendo la misma lógica que se utilizó en la respuesta anterior, esta respuesta puede ser clasificada dentro de la característica C11 en el marco de las categorías de análisis de la competencia interpretativa.	C11
14c. ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el máximo valor?	**1.** Que toda la desviacion sea parte del Promedio	Esta respuesta carece de claridad en su formulación, lo que impide su clasificación en cualquiera de las categorías establecidas.	N/A
	2. La media tiene que ser pequeña asi que yo creo que tiene que ser cerca a 0 y la desviación tiene que ser considerablemente mas grande que la media. Por que en conclusión si los datos tienen una dispersión alta y una media muy chiquita la variación va a tomat el maximo valor que se pueda.	Si se considera que la expresión " la variación va a tomar el máximo valor que se pueda." se refiere a que el coeficiente de variación alcanzará su valor máximo, se puede concluir que esta respuesta describe con precisión las condiciones necesarias para que el coeficiente de variación logre dicho valor, respaldada por una justificación adecuada. En	C11

		consecuencia, se clasifica plenamente dentro de la categoría C11.	
3.	La media tiene que ser muy pequeña y la desviación tiene que ser muy grande.	De acuerdo con la misma lógica aplicada en la respuesta anterior, esta respuesta puede clasificarse dentro de la característica C11.	C11

En la Tabla 46, se presenta un resumen de las características asociadas a las respuestas los estudiantes en cada una de las preguntas diseñadas para la tarea#3.

Tabla 46. *Características de la Tarea 3*

Categoría / **Pregunta**	**N/A**	**C2**	**C3**	**C5**	**C6**	**C7**	**C11**	**C13**	**C14**
1	X								
2			X	X					
3			X	X	X				
4			X	X					
5			X	X					
6						X			
7									X
8			X	X					
9			X	X					
10			X	X					
11 (A, B, C, D)			X	X					
12	X								
13	X								
13C		X			X				
13D		X			X				
13E		X			X				
13F								X	
14 (A, B, C)							X		

En relación con lo reportado para estas tres tareas y teniendo en cuenta que la conjetura planteada para este experimento de enseñanza corresponde a: El diseño e implementación de una secuencia de tareas de aprendizaje mediadas por Excel, con el uso de bases de datos reales, dirigida a estudiantes de grado noveno, promueve la interpretación de los conceptos de dispersión (varianza y desviación estándar), se hace explícito y evidente que se diseñaron e implementaron las tres tareas proyectadas en la secuencia y que en todos los casos se usó Microsoft Excel tanto en el

diseño como en el desarrollo de la tarea por parte de los estudiantes, aun cuando no se trabajara con la herramienta al mismo nivel de profundidad.

Adicionalmente, y en relación con el propósito principal de la conjetura, en todos los casos se dispuso de una base de datos y a través de esta se abordaron diferentes preguntas que tenían por finalidad promover algunas de las 15 características clasificadas en la categoría de análisis. Sin embargo, aunque en una gran parte de las respuestas se lograron evidenciar elementos y particularidades de las características de la competencia interpretativa, no se puede desconocer que algunas tuvieron una mayor presencia o protagonismo durante el desarrollo de estas tareas como lo fueron:

- ✓ C2: Leer los datos, dentro de los datos, más allá de los datos y detrás de los datos
- ✓ C3: Identificar, analizar y comprender una información matemática a partir de una tabla, una gráfica, una expresión simbólica o una situación descrita en lenguaje natural.
- ✓ C5: Interpretar información presentada en forma gráfica o diagramas.
- ✓ C6: Evaluar e interpretar un resultado matemático en términos del contexto de un problema del mundo real.
- ✓ C9: Comprender el alcance y los límites de los conceptos y soluciones matemáticas que son el resultado del modelo matemático utilizado.
- ✓ C13: Comparar resultados para diferentes condiciones en un experimento.

Esto no quiere decir que las otras características no se identificaron en las respuestas de los estudiantes, pero a diferencia de las mencionadas anteriormente, se hicieron presentes con menor frecuencia y solo en algunos ítems. Adicionalmente, también sucedió que una característica no estuviera relacionada en ninguno de los casos con las respuestas como sucedió con C1: Observar, discernir y extraer información a través de elementos matemáticos.

Lo anterior permite reconocer que, efectivamente, se logra promover la interpretación a través de la caracterización planteada, y se destaca que se privilegiaron aquellas interpretaciones que relacionaban el componente matemático (medidas estadísticas) y lo contextual, así como aquellas que surgen de los cambios de representación de los conjuntos de datos.

5. CONCLUSIONES

En principio, es importante exponer el alcance que se logró con los objetivos específicos, el objetivo general y si se dio respuesta a la pregunta problema planteada. Se identifica que efectivamente, se llevó a cabo una detallada revisión documental con la cual se pretendía clarificar cada uno de los elementos teóricos que se iban a abordar en la investigación, en particular en correspondencia con el primer objetivo específico, se presentaron en el capítulo 2 todos los hallazgos relacionados con la competencia interpretativa, allí se hace énfasis en aspectos relacionados con la indagación que se realizó y que generó la construcción de la definición y posterior caracterización de todo lo referente a la interpretación. Este ejercicio tuvo como propósito el poder extrapolar desde diferentes autores elementos que se pudieran aplicar en el campo de la educación matemática y, más específicamente, en las intervenciones que se proyectaron realizar en un aula de clase de la institución Tandem – Apoyo en procesos escolares; lo anterior nos permite afirmar que se dio cumplimiento a este primer objetivo.

En relación con el segundo objetivo específico, se puede identificar que, desde el diseño de las tareas, cada uno de los ítems nació con un propósito que está asociado con tres de los siguientes aspectos: Identificar la claridad frente a los temas previos o acercar al estudiante al contexto que se quiere estudiar, generar una interacción entre los estudiantes y Microsoft Excel y, por último, promover y generar las habilidades relacionadas con la competencia interpretativa. En todos los casos, al realizar la implementación, se evidencia que los ítems cumplieron con su objetivo, pues hasta este punto el interés estaba en lograr determinar la efectividad de las producciones realizadas.

Por último, en relación con el tercer objetivo específico, se establece una conexión directa con la pregunta problema: ¿De qué forma se promueve la interpretación de la varianza y la desviación estándar a través de tareas mediadas por Microsoft Excel y el uso de bases de datos reales en estudiantes de grado noveno? Lo anterior se puede afirmar debido a que se pudo evidenciar por medio del análisis de las tareas propuestas que en las respuestas de los estudiantes se identifican elementos relacionados con la caracterización establecida para la competencia interpretativa. Esto quiere decir que los ítems planteados en las tareas al parecer sí promueven en los estudiantes el desarrollo de estas habilidades, y que los medios usados para tal finalidad resultaron ser efectivos (Microsoft Excel y las bases de datos). Sin embargo, se reconoce la posibilidad de mejorar un

aspecto que en el análisis se hizo presente, y corresponde a la dificultad que aún parece estar presente en los estudiantes cuando se propone establecer una interpretación que vaya de lo contextual a lo matemático (puesto que en caso contrario los estudiantes evidenciaron un dominio mayor).

Adicionalmente, lo anterior también se encuentra en una correspondencia muy estrecha con lo reportado en relación con la conjetura planteada para el experimento de enseñanza y la eventual verificación que se dio a esta (como se describió al final del capítulo anterior). No obstante, es pertinente que para futuras investigaciones se retome lo correspondiente a proponer tareas en las que también se privilegie y se profundice lo correspondiente a las características de la categoría denominada *Tendencias y Patrones (Predicción o conjeturas y Sesgos)*, teniendo en cuenta que estas no se destacaron entre las características que se lograron promover al analizar las producciones presentadas por los estudiantes.

Otro asunto de interés que se quiere poner en consideración en este apartado está relacionado con aquellos aspectos que se podrían profundizar, mejorar o reformular. Por ejemplo, a partir del trabajo desarrollado en el cual se tomaron los elementos propuestos por Gómez et al. (2018) para el diseño de las tareas de aprendizaje y reconociendo que esto aportó significativamente a la primera fase del experimento de enseñanza (en el que se requería hacer una preparación previa a la implementación), queda abierta la posibilidad de diseñar las intervenciones de clase tomando otro modelo de tareas de aprendizaje matemático escolar, esto con el fin de estudiar qué nuevos hallazgos o cambios se darían. Adicionalmente, también se puede reformular la investigación cambiando la herramienta tecnológica, teniendo en cuenta que en este caso las tareas se desarrollaron a través de Microsoft Excel, se podría considerar la opción de explorar otras herramientas como GeoGebra. También sería posible profundizar en el planteamiento, la implementación y el análisis de las tareas con el objetivo de promover los procesos vinculados a la génesis instrumental.

En cuanto al ámbito personal y laboral, el desarrollo de este proyecto permitió dar una nueva perspectiva a las prácticas docentes de los autores, puesto que se reconoce que ocasionalmente se prioriza (debido a las dinámicas institucionales) el trabajo con los procesos algorítmicos, y tareas como las propuestas en esta investigación permiten identificar la importancia de abordar conceptos

matemáticos o estadísticos desde la diversidad de contextos reales y relevantes para los estudiantes, usando herramientas diferentes al tablero y el marcador y, por supuesto, dando un papel protagónico a otros procesos como lo es la interpretación. Lo anterior también es un llamado a reflexionar sobre cómo se pueden generar nuevas estrategias de enseñanza que permitan acceder al conocimiento y que a su vez este sea más significativo para los estudiantes.

Finalmente, es importante mencionar que a medida que el proyecto fue avanzando, se fueron compartiendo resultados parciales con la comunidad de Educación Matemática a través de tres eventos académicos: el Encuentro Internacional de Investigación en Educación Matemática (EIEM 7) organizando por la Universidad del Atlántico, el Décimo Encuentro Distrital de Educación Matemática (EDEM 10) organizado por la Universidad Distrital Francisco José de Caldas, y el XIV Simposio de Matemática y Educación Matemática organizando por la Universidad Antonio Nariño. Cabe enfatizar que la aprobación del trabajo presentado a estos eventos implicó la evaluación de pares externos, asunto que, por un lado, permitió recibir realimentación de aspectos por mejorar y, por otro lado, pone de manifiesto una valoración positiva (en cuanto a la pertinencia, por ejemplo) por parte de la comunidad académica respecto al trabajo que se desarrolló.

6. REFERENCIAS

Alsina, Á. (2023). Conocimientos esenciales sobre los procesos, habilidades o competencias matemáticas: orientaciones para implementar situaciones de aprendizaje. Edma 0-6: Educación Matemática en la Infancia, 12(2),65-108. DOI: https://doi.org/10.24197/edmain.2.2023.65-108

Alvarado, H., y Batanero, C. (2008). Significado del teorema central del límite en textos universitarios de probabilidad y estadística. *Estudios pedagógicos (Valdivia)*, *34*(2), 7-28.

Argibay, J. (2015). Muestra en Investigación Cuantitativa. Subjetividad y Procesos Cognitivos, vol. 13, núm. 1, 2009, pp. 13-29. Universidad de Ciencias Empresariales y Sociales. Buenos Aires, Argentina

Aroca, P., García, C., y López, J. (2009). Estadística descriptiva e inferencial. *Revista el auge de la estadística en el siglo XX, 22*, 165-176.

Ávila, R. (1996). *Estadística elemental. Novena Edición*. Estudios y ediciones R.A. Lima, Perú.

Barreto, A. (2012). El progreso de la Estadística y su utilidad en la evaluación del desarrollo. Papeles de población, 18(73), 241-271.

Batanero, C., Godino, J., Vallecillos, A., Green, D. R., y Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematical Education in Science and Technology*, *25*(4), 527–547.

Benítez, G. & Arrondo, V. (2014). Sobre la definición de estadística. Brasil: DataGamaZero– Revista de Ciências da Informação, 6(4).

Canavos, G. (1988). *Probabilidad y estadística, aplicaciones y métodos*. McGrawHill/Interamericana de México, SA. México.

Camargo, L. (2021). *Estrategias cualitativas de investigación en Educación Matemática*. Bogotá: Fondo Editorial Universidad Pedagógica Nacional.

Cardoso, E. & Cerecedo, M. (2008). El desarrollo de las competencias matemáticas en la primera infancia. *Revista iberoamericana de educación*, *47*(5), 1-11.

Chance, B., Ben-Zvi, D., Garfield, J., y Medina, E. (2007). The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education*, 1(1). http://dx.doi.org/10.5070/T511000026

Chevallard, Y. (1999). El análisis de las prácticas docentes en la teoría antropológica de lo didáctico. *Recherches en Didactique des Mathématiques, 19*(2), 221-266.

Contreras, J. (2011). Formación de competencias: tendencias y desafíos en el siglo XXI. *Universitas*, (15), 109-138.

Contreras, A. (2018). Fortalecer la competencia de interpretación matemática a través de la implementación de una estrategia pedagógica. *Actualidades Pedagógicas, (71*), 13-34. doi: http://dx.doi.org/10.19052/ap.4341

Contreras, J, & Molina, E. (2019). Elementos clave de la cultura estadística en el análisis de la información basada en datos. Actas del Tercer Congreso Internacional Virtual de Educación Estadística. Disponible en www.ugr.es/local/fqm126/civeest.html

Díaz, L. & Landázuri, M. (2008). C*oncepciones de desviación estándar en estudiantes de educación media* [Tesis de maestría]. Universidad Pedagógica Nacional.

Estepa, A. & Del Pino, J. (2013). Elementos de interés en la investigación didáctica y enseñanza de la dispersión estadística. *Números. Revista de Didáctica de las Matemáticas, 83*, 43-63.

Franklin, C., Bargagliotti, A., Arnould, P., Gould, R., Jhonson, S., Pérez, L. y Spangler, D. (2020). *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II). A Framework for Statistics and Data Science Education*.

Garmendia, M. (2018). *Bases de datos en Microsoft Excel. Diseño y administración*. Universidad Nacional Agraria. Mangua, Nicaragua.

Gómez, M. (2012). *Elementos de estadística descriptiva*. Editorial Universidad Estatal a Distancia.

Gómez, P., Mora, M., y Velasco, C. (2018). Formación de profesores de matemáticas y práctica de aula: conceptos y técnicas curriculares. Ediciones Uniandes.

Heumann, C., Schomaker, M., y Shalabh (2016). *Introduction to statistics and data analysis*. Springer International Publishing Switzerland.

Instituto Colombiano para la Evaluación de la Educación [ICFES] (s.f). Glosario ICFES. Recuperado de https://www.icfes.gov.co/glosario

Instituto Colombiano para la Evaluación de la Educación [ICFES]. (2023). Guía de orientación Saber 11.

Instituto Colombiano para la Evaluación de la Educación [ICFES]. (2020). Marco para la prueba de Matemáticas PISA 2021.

Jiménez, D. (2019). Herramientas digitales para la enseñanza de las matemáticas en la educación básica. Tesis de especialización en multimedia para la docencia. Universidad Cooperativa de Colombia.

Leung, A., Chan, Y., y López, F. (2000). Instrumental Genesis in Dynamin Geometry Environments.

López, A, Cruañas, J, Salgado A, Lastayo. L y Rodríguez V. (2009). Microsoft Excel y la estadística. *Revista Habanera de Ciencias Médicas, 8(5*), 20-25. Recuperado de: https://www.redalyc.org/articulo.oa?id=180414046005

Luna, C. (2011). *Excel: Fórmulas y funciones*. [Taller de conferencia]. 3er congreso anual Ecorfan. https://www.ecorfan.org/congresos/3eranual/Introduccion%20a%20EXCEL%20Formulas%20y%20Funciones.pdf

Manzano, V. & Sánchez, G. (2005). Sobre la definición de estadística. DataGramaZero - Revista de Ciência da Informação - Artigo02, 6 (4), 1-6.

Mendenhall, W. & Beaver, R. (2010). *Introducción a la probabilidad y estadística*. Editorial Thomson.

Ministerio de Educación Nacional de Colombia [MEN]. (1998). Lineamientos Curriculares para Matemáticas. Bogotá: Magisterio.

Ministerio de Educación Nacional de Colombia [MEN]. (2006). Estándares Básicos de Competencias en Matemáticas. Bogotá: Magisterio.

Ministerio de Tecnologías de la Información y las Comunicaciones [MinTIC] (2019). *Guía para el uso y el aprovechamiento de Datos Abiertos en Colombia*. Bogotá: Magisterio.

Molina, M., Castro, E., Molina, J. L., y Castro, E. (2011). Un acercamiento a la investigación de diseño a través de los experimentos de enseñanza. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 75-88.

Montero, J. (2007). *Estadística descriptiva.* Thomson Ediciones Paraninfo S.A.

Moreno, L. & Montañez, J. (2015). Desarrollo de la competencia interpretativa en matemáticas con mediaciones de las TIC, en estudiantes de grado sexto de la IED República de Estados Unidos de América. [Tesis de Especialización, Universidad de San Buenaventura]. Repositorio institucional Universidad de San Buenaventura.

Moreno, M. & Villanueva, H. (2013). Propuesta metodológica para el desarrollo de competencia interpretativa en estadística descriptiva desde las medidas de tendencia central. *Amazonia Investiga*, 6–31.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: autor.

Olivera, B. (2013). *Estadística y probabilidad*. Pearson Educación.

Otzen, T. & Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *International Journal of Morphology, 35(*1), 227-232.

Pastells, Á. (2004). *Desarrollo de competencias matemáticas con recursos lúdico-manipulativos: para niños y niñas de 6 a 12 años* (Vol. 2). Narcea Ediciones.

Pedreros, M. & Mejía, M. (2013). Relaciones e integraciones de la Teoría Antropológica de lo Didáctico y la Génesis Instrumental. *Educación científica y tecnológica. Edición especial. Bogotá D.C.* 397–401.

Pérez, R. (2010). Nociones básicas de estadística. Editorial Rigoberto Pérez.

Pérez, C. (2014). Enfoques teóricos en investigación para la integración de la tecnología digital en la Educación Matemática. *Perspectiva educacional*, *53*(2).

Pérez, L. (2006). Microsoft Excel: Una herramienta para la investigación. *Revista Electrónica de las Ciencias Médicas en Cienfuegos. 4(3*), 68 – 71.

Pinto, P (1996). Currículo por competencias: Desafío educativo. *Propuestas Pedagógicas, 33*, 44 - 50.

Rabardel, P. (2011). *Los hombres y las tecnologías. Visión cognitiva de los instrumentos contemporáneos* (Trad. M. Acosta). Colombia: Ediciones Universidad Industrial de Santander. (Trabajo original publicado en 1995).

Rojas, A. & Fúneme, C. (2021). Aprendizaje Basado en Proyectos de las medidas de dispersión en busca de la cultura estadística en un contexto rural. *Innovaciones educativas*, *23*(35), 73–87. https://doi.org/10.22458/ie.v23i35.3811

Ross, S. (2018). Introducción a la estadística. Segunda Edición. Editorial Reverté.

Sabadías, A. (1995). *Estadística descriptiva e inferencial* (Vol. 8). Universidad de Castilla La Mancha.

Sánchez, L. & Aguilar, G. (2009). Experiencia educativa: Taller de Habilidades de Pensamiento Crítico y Creativo. Área de Formación Básica General. Universidad Veracruzana. México.

Spiegel, M. (1977). *Probabilidad y Estadística. Teoría y 760 problemas resueltos*. McGraw Hill.

Steinbring, H. (1989). La Interacción entre la práctica de la enseñanza y las concepciones teóricas: un modelo cooperativo de capacitación en servicio en estocástica para profesores de matemática. *Estudios en educación matemática (7)*, 202-214.

Sua, C. & Camargo, L. (2019). Geometría dinámica y razonamiento científico: dúo para resolver problemas. *Educación Matemática, 31*(1), 7-37.

Tardif, J. (2008). Desarrollo de un programa por competencias: De la intención a su implementación. *PROFESORADO*. 12.

Tejada, J. (1999). Acerca de las competencias profesionales I. *Revista Herramientas, 56*, 20-30.

Triola, M (2009). *Estadística*. Décima Edición. Pearson.

Trouche, L. (2005). Instrumental genesis, individual and social aspects. In D. Guin, K. Ruthven, & L. Trouche (Eds.), The didactical challenge of symbolic calculators: Turning a computational device into a mathematical instrument (pp. 197–230). New York: Springer

Valdiri. L. & Rincón, A. (2016). Análisis de las competencias en razonamiento cuantitativo evaluadas por las pruebas saber-pro en los contenidos programáticos de ciencias básicas en la UMNG. Encuentro Internacional de Educación en Ingeniería ACOFI. Universidad Militar Nueva Granada. Bogotá, Colombia.

7. ANEXOS

A continuación, se presenta una descripción detallada de los elementos propuestos por Gómez, Mora y Velazco (2018) en el proceso de construcción y diseño de las 3 tareas implementadas.

7.1. Anexo 1

TAREA 1: Sobre el volumen de datos
(*Rango, varianza y desviación)*

- ✓ **Requisitos:** Para el desarrollo de la tarea es importante precisar que los estudiantes ya habrán estudiado previamente conceptos estadísticos como: población, muestra, variables estadísticas y su clasificación en nominales, ordinales, discretas y continuas. Además, ya estarán familiarizados con las definiciones de las medidas de tendencia central y de dispersión para datos no agrupados, así como con los algoritmos para calcularlas dichas medidas de forma manual. Por otro lado, previo al desarrollo de esta tarea, también se habrá realizado un acercamiento al uso de Microsoft Excel, con el fin de que los estudiantes reconozcan las funciones que les permiten calcular medidas como la media, el rango, la varianza y la desviación estándar.
- ✓ **Metas:** Se espera que los estudiantes logren reconocer las potencialidades de Microsoft Excel a la hora de realizar algún tratamiento a la base de datos suministrada en la tarea y que, con ello, puedan establecer interpretaciones y proponer conjeturas sobre el conjunto de datos teniendo en cuenta el contexto en el que se encuentran inmersos. Para ello, los ítems propuestos en esta tarea se encuentran en correspondencia con las características descritas en el capítulo 2.1 para la competencia interpretativa.
- ✓ **Formulación:** Se presentan los enunciados juntos con las instrucciones y preguntas que deben resolver los estudiantes.

PARTE I:

1. **Responda las siguientes preguntas:**

a) ¿Qué piensa usted sobre el consumo excesivo de alcohol?

b) ¿Cuáles problemas de salud conoce usted que se pueden generar por el consumo excesivo de alcohol?

c) ¿Cuáles problemáticas sociales considera usted que se generan por el consumo excesivo de alcohol?

d) En promedio, ¿cuál cree usted que es el grado de alcohol de las bebidas alcohólicas que conoce?, ¿en qué considera que afecta el grado de alcohol a los licores y a las personas que los consumen?, ¿por qué?

El consumo excesivo de alcohol acarrea graves consecuencias para la salud de las personas, principalmente asociadas a enfermedades como el cáncer. Esto se evidencia al indagar sobre el impacto que se tiene en relación con los años perdidos de vida por muerte prematura y los años perdidos de vida saludable por el padecimiento de enfermedades; también, conlleva daños y consecuencias sociales devastadoras en familias y comunidades, especialmente asociadas a siniestralidad vial y a diferentes tipos de violencia como lesiones personales, riñas y homicidios (Ministerio de Salud y Protección social, 2024)[15].

Para realizar un análisis estadístico relacionado con la problemática descrita, a continuación, se presenta un enlace que contiene una base de datos correspondiente a productos de licor registrados en el departamento de Risaralda, con información del nombre

[15] Ministerio de Salud y Protección social (2024). Minsalud entrega recomendaciones frente al consumo de alcohol en festividades. Boletín de prensa. Tomado de https://www.minsalud.gov.co/Paginas/Minsalud-entrega-recomendaciones-frente-al-consumo-de-alcohol-en-festividades.aspx

de la empresa distribuidora, su respectivo NIT, resolución del registro INVIMA y el grado de alcohol que contiene cada bebida reportada.

https://drive.google.com/uc?export=download&id=1uzB3ZO8HNnzeosF1psNff5NDl_S4r0kH

2. Descargue la base de datos a través del enlace anterior y ubíquese en la hoja denominada "Licores Risaralda". Responda las siguientes preguntas.

a) ¿Cuál es la población de estudio?

b) ¿Cuántas variables estadísticas identifica?, ¿cuáles?

c) Clasifique las variables que identificó en el literal anterior.

d) ¿Con cuáles de las variables estadísticas tiene sentido (dentro del contexto de la base de datos) calcular las medidas de tendencia central y dispersión?, ¿con cuáles no? Justifique su respuesta.

3. La variable "Grado de alcohol" mide el porcentaje de alcohol que contiene cada bebida registrada en la base de datos. Calcule la media, el rango, la varianza y la desviación estándar de la variable Grado de alcohol, haciendo uso de los comandos de Excel y consigne las respuestas en la Tabla 1 (utilice las funciones que permiten calcular estas medidas para las poblaciones).

Tabla 1

	Media	**Rango**	**Varianza**	**Desviación estándar**
Población				

a) Según el Decreto 162 de 2021 una bebida alcohólica es un "producto apto para consumo humano que contiene una concentración no inferior a 2.5 grados alcoholimétricos". ¿Hay bebidas no alcohólicas en la base de datos? Justifique su respuesta.

b) Teniendo en cuenta la media y la desviación estándar calculadas para los datos de la variable "Grado de alcohol", realice una descripción de dicho conjunto de datos.

c) ¿Qué puede decir respecto a la dispersión de la variable “Grado de alcohol” de las bebidas presentadas en la base de datos?

4. "Las intoxicaciones por bebidas alcohólicas adulteradas por metanol son un evento de interés en salud pública con vigilancia epidemiológica durante todo el año e intensificada durante la temporada decembrina" (Ministerio de Salud y Protección social, 2024). Teniendo en cuenta esta afirmación, suponga que existe un dato adicional en la base de datos que corresponde a un licor adulterado que se prepara a base de alcohol etílico (con un 96% de alcohol) y un saborizante artificial. Agregue este nuevo dato al final de las columnas Nombre y Grado de alcohol y calcule nuevamente las medidas de tendencia central y dispersión solicitadas considerando dicho dato. Consigne las respuestas en la Tabla 2.

Tabla 2

	Media	**Rango**	**Varianza**	**Desviación estándar**
Población con dato adicional				

a) ¿Qué representa el cambio en el rango de los datos al comparar las Tablas 1 y 2?

b) Al comparar la información de la población y la población con el dato adicional de las Tablas 1 y 2, ¿qué diferencias o similitudes puede enunciar para cada medida calculada?

c) ¿Por qué cree usted que las medidas de la población no cambian demasiado aun cuando se agrega un dato adicional mucho mayor a los demás?

5. Ubíquese en la hoja denominada Muestra 1 o 2 (según indique el profesor), las cuales corresponden, como su nombre lo indica, a dos muestras distintas de la base de datos original obtenidas de forma aleatoria. Enseguida, calcule las medidas: media, rango, varianza y desviación estándar de la variable Grado de alcohol (utilice de nuevo las funciones para muestra en Excel). Consigne las respuestas en la Tabla 3.

Tabla 3

	Media	**Rango**	**Varianza**	**Desviación estándar**
Muestra				

6. Según el proyecto de Acuerdo 95 del 2014 presentado por Liliana Graciela Guaqueta De Diago, concejal de Bogotá: "Un estudio sobre consumo de alcohol en menores de edad en Colombia", elaborado por la Corporación Nuevos Rumbos con 9.276 alumnos de colegios públicos y privados del país, indica que muchos niños empiezan a ingerir licor a los 10 años y las niñas a los 11 años. No es para menos, si se tiene en cuenta que, de acuerdo con la misma investigación, niños y adolescentes, en general, ven normal y aceptable tomar trago para celebrar, socializar, vencer la timidez, lograr aceptación, conquistar a alguien y pasarla bien. Para la elaboración de un coctel peligroso cuyo consumo crece entre los jóvenes y alarma a las autoridades, son usados algunos ingredientes como: alcohol antiséptico, agua, refrescos en polvo, gaseosas, saborizantes y hasta leche condensada. Teniendo en cuenta lo anterior, suponga que existe un dato adicional, en la hoja Muestra 1 o 2 según haya sido asignado, que corresponde a un licor adulterado que se prepara a base de alcohol etílico (concentrado al 96%) y un saborizante artificial. Agregue este nuevo dato al final de las columnas Nombre y Grado de alcohol, y calcule nuevamente las medidas del numeral 4. Consigne las respuestas en la Tabla 4.

Tabla 4

	Media	**Rango**	**Varianza**	**Desviación estándar**
Muestra con dato adicional				

a) ¿Qué diferencias o similitudes puede enunciar para cada medida, teniendo en cuenta la información de las tablas 3 y 4?

b) De acuerdo con lo observado, ¿de qué dependería poder asumir que la media y las medidas de dispersión obtenidas en la muestra sean aproximadamente iguales a las de la población? Explique.

c) ¿La dispersión observada en la muestra es equivalente a la de la población? Justifique su respuesta.

7. Responda las siguientes preguntas:

a) Teniendo en cuenta toda la información suministrada, ¿cómo se relaciona la problemática del consumo de alcohol con la producción de las bebidas alcohólicas?

b) ¿Qué significado puede dar a las medidas halladas para la variable "Grado de alcohol" (punto 2), en relación con el contexto de los licores y los problemas existentes frente al consumo de estas bebidas?

c) ¿Qué implicaciones tuvo el agregar el dato del licor adulterado (dato atípico) al conjunto de datos dado inicialmente (Pregunta 2)? ¿Qué sucedió con las medidas calculadas cuando se agregó el mismo dato a la muestra (Pregunta 4)? Escriba una conclusión.

d) Realice la lectura de la información suministrada por el Ministerio de Salud en relación con la problemática del consumo de bebidas alcohólicas en Colombia (Anexo). Teniendo en cuenta la información de la base de datos y la lectura realizada, ¿qué conclusiones puede establecer?

e) ¿Cuál es su postura frente a esta problemática?

PARTE II:

En la hoja de Excel *"*Notas - Parte II*",* se presentan las notas obtenidas por los estudiantes de la institución Tandem – Apoyo en procesos escolares en el examen final del tercer corte de matemáticas, correspondiente al semestre 2023 – II.

8. Ubíquese en la hoja del archivo Excel denominado Notas - Parte II y teniendo en cuenta la lista de datos calcule la media, rango, varianza y desviación estándar. Registre los resultados obtenidos en la Tabla 5.

Tabla 5

Media	Rango	Varianza	Desviación estándar

9. Suponga que se modifican 10 datos de forma aleatoria. Escriba una conjetura, en la primera fila de la Tabla 6, sobre los resultados que podría obtener al calcular las medidas allí descritas tras modificar los 10 datos.

Tabla 6

Eliminación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura				
Medida				

10. Modifique 10 datos aleatoriamente y calcule la media, rango, varianza, desviación estándar. Registre los resultados en la segunda fila de la Tabla 6.

a) ¿Qué diferencias o similitudes existe entre la conjetura y las medidas obtenidas? Comente por qué considera que surgen estas diferencias o similitudes.

11. Suponga que se eliminan 20 datos de forma aleatoria. Escriba una conjetura, en la primera fila de la Tabla 7, en relación con los resultados que podría obtener tras dicha eliminación.

Tabla 7

Modificación de datos	Media	Rango	Varianza	Desviación estándar
Conjetura				
Medida				

12. Proceda a eliminar 20 datos de forma aleatoria, a continuación, calcule la media, rango, varianza, desviación estándar y registre los resultados en la segunda fila de la Tabla 7.

a) ¿Qué diferencias o similitudes existe entre las conjeturas y las medidas obtenidas? Describa y explique por qué surgen estas diferencias o similitudes.

13. ¿Qué puede concluir en relación con las medidas, al hacer la modificación y eliminación al conjunto de datos suministrado?

Texto: Minsalud entrega recomendaciones frente al consumo de alcohol en festividades

24/12/2021

Boletín de Prensa No 1272 de 2021

El consumo nocivo de alcohol trae consecuencias a la salud de la persona que consume, principalmente asociados a enfermedades no transmisibles como el cáncer, lo cual contribuye a la carga de enfermedad del país, pero también daños y consecuencias sociales devastadoras en familias y comunidades, especialmente asociadas a siniestralidad vial y a diferentes tipos de violencia como lesiones personales, riñas y homicidios.

En las festividades decembrinas y de la llegada del nuevo año 2022, se busca concientizar a la población de las consecuencias que puede llegar a tener el consumo de alcohol, en tanto, se ha documentado el aumento del consumo de bebidas alcohólicas y los riesgos derivados del consumo excesivo.

"Las intoxicaciones por bebidas alcohólicas adulteradas por metanol son un evento de interés en salud pública con vigilancia epidemiológica durante todo el año e intensificada durante esta temporada", explicó Ana María Peñuela, coordinadora del grupo Convivencia Social y Ciudadanía.

De acuerdo con la Organización Mundial de la Salud, el alcohol es la sustancia psicoactiva más consumida en el mundo y mata a más de 3 millones de personas al año, en su mayoría hombres.

Situación en Colombia, algunas cifras:

En Colombia en el año 2019 el DANE realizó la Encuesta Nacional de Consumo de Sustancias Psicoactivas, con respecto al consumo de alcohol o bebidas alcohólicas, el 84% de las personas de 12 a 65 años informaron haber consumido alguna vez en su vida, 54.5% consumieron el último

año y el 30.1% en el último mes. Los departamentos de Boyacá y Risaralda registran las prevalencias más altas.

Por otra parte estima que en el país existen 1.461.000 personas con consumo de riesgo de alcohol y 995.000 persona con patrones de dependencia los cuales se benefician de recibir intervenciones de prevención y tratamiento, sin embargo, solo 348.000 personas refirieron que en los últimos 12 meses sintió la necesidad de recibir ayuda para dejar de consumir alguna sustancia psicoactivas, finalmente por primera vez se logró identificar que alrededor de 578.000 mujeres de 12 a 65 años informaron haber consumido alguna sustancia psicoactiva durante el embarazo.

Por su parte, el estudio nacional de consumo de sustancias psicoactivas en población escolar Colombia – 2016, refiere que el 69,2% de los escolares declaran haber consumido alguna bebida alcohólica en su vida, el 70,4% de las mujeres y el 68,1% de los hombres. En el uso actual de bebidas alcohólicas se encuentran 1.200.622 estudiantes, con un significativo mayor uso entre las mujeres respecto de los hombres, 37,9% y 36,1%, respectivamente.

Recomendaciones, prácticas y comportamientos generales

- Recuerde se puede festejar sin ingesta de bebidas alcohólicas. Genere prácticas de disfrute de las diferentes actividades familiares y en estas festividades que no dependan del consumo de bebidas alcohólicas busque alternativas como cocteles a base de frutas y otras preparaciones que no requieran de esta sustancia.
- El consumo de alcohol como estrategia para afrontar el estrés, es una medida desacertada; se sabe que pueden aumentan los síntomas de ansiedad y angustia, depresión y otros trastornos mentales, además aumenta el riesgo de violencia doméstica e intrafamiliar.
- La exposición a alcohol durante la gestación puede afectar al desarrollo del bebé, con posibles consecuencias como malformaciones y alteraciones físicas y a nivel del neurodesarrollo. Por lo tanto, consumir cualquier cantidad de alcohol durante el embarazo presenta alto riesgo.
- El alcohol aumenta el riesgo, la frecuencia y la gravedad de la victimización y todo tipo de actos de violencia interpersonal, como la violencia de pareja, la violencia sexual, maltrato a las personas mayores y la violencia contra los niños, las niñas, los jóvenes, los adolescentes y las mujeres.

- El alcohol se asocia alta accidentalidad vial y comportamientos de riesgo, así mismo aumenta el riesgo de muerte y traumatismos causados por accidente de tránsito, así como por ahogamiento y caídas.
- Respecto a la embriaguez por alcohol las manifestaciones clínicas pueden variar acorde a la cantidad ingerida, la concentración de alcohol presente en la bebida y características propias del individuo como cronicidad del consumo. Sin embargo; si alguien refiere o presenta dolor de cabeza de gran intensidad, visión borrosa, vómito persistente o dolor abdominal o compromiso del estado de consciencia (no responde al llamado) debe acudir al centro médico más cercano.
- Adicionalmente el consumo de alcohol puede ser un factor de riesgo o activador de conductas violentas en personas de por sí agresoras, es importante que ante los casos anteriores busque ayuda, a través de la comunidad o líneas locales.

Tomado de https://www.minsalud.gov.co/Paginas/Minsalud-entrega-recomendaciones-frente-al-consumo-de-alcohol-en-festividades.aspx

✓ **Materiales y recursos:** Para esta tarea se dispone de materiales como: Una guía elaborada en Word que cuenta con 13 puntos, en los cuales se establecen indicaciones y preguntas como los presentados en el apartado de formulación. También, un archivo en hoja de cálculo de Excel en el cual se encuentra la base de datos con la cual los estudiantes deben interactuar y dar respuesta a lo propuesto en el archivo Word. Los estudiantes tienen a la mano computadores propios o de la institución para el desarrollo de esta tarea.

✓ **Agrupamiento:** Los estudiantes desarrollarán la tarea de forma individual, se plantea de esta manera teniendo en cuenta que en la implementación se contará con pocos estudiantes y que se quiere obtener diferentes perspectivas en cuanto al planteamiento de conjeturas e interpretaciones de las situaciones planteadas en el contexto de la base de datos.

✓ **Interacción:** La comunicación se dará principalmente entre el profesor y los estudiantes, es importante precisar que las preguntas que se resolverán no estarán relacionadas con el contenido sino con la forma (es decir, dudas sobre los enunciados, con las indicaciones dadas, entre otros); lo anterior tiene como finalidad obtener producciones propias de cada estudiante

que no estén sesgadas por la opinión del profesor debido a que la tarea propuesta es de interpretación y conjeturación a partir de la exploración con el conjunto de datos.

- ✓ **Temporalidad:** Dado que la tarea consta de dos partes: la primera, dirigida a identificar e interpretar las similitudes y diferencias entre el cálculo de la media, la varianza y la desviación estándar de una población en comparación con una muestra de esta; y la segunda, centrada en realizar estimaciones de los valores de la media, la varianza y la desviación estándar en una población al modificar el conjunto de datos. El tiempo asignado para el desarrollo de la tarea en su totalidad será 2 horas: 1 hora y media para la primera parte (que consta del ítem 1 al 8) y el tiempo restante para la segunda parte (del ítem 9 al 13).

7.2. Anexo 2

TAREA 2: Unidades de las medidas de dispersión

Media, varianza y desviación

- ✓ **Requisitos:** Para el desarrollo de la tarea es importante precisar que los estudiantes ya habrán estudiado previamente los conceptos estadísticos. Además, ya se encuentran familiarizados con las definiciones de las medidas de tendencia central y de dispersión para datos no agrupados, y con los algoritmos para calcularlas de forma manual.

  Por otro lado, previo al desarrollo de esta tarea, ya se habrá desarrollado una tarea usando Microsoft Excel, por lo cual los estudiantes ya reconocen las funciones que les permiten calcular medidas como la media, el rango, la varianza y la desviación; así como realizar operaciones aritméticas entre valores numéricos.

- ✓ **Metas:** Se espera que los estudiantes logren establecer la fórmula que permite calcular la varianza y la desviación estándar para un conjunto de datos no agrupados, realizando la interpretación correspondiente a cada una de las operaciones aritméticas que se realizan en estas, y que en general permiten establecer conclusiones sobre el conjunto de datos en determinado contexto.

- ✓ **Formulación:**

1. Responda las siguientes preguntas:

a) ¿Qué piensa usted acerca de la movilidad en Bogotá?

b) ¿Cree pertinente y oportuno el planteamiento e implementación de estrategias que permitan mejorar la movilidad en Bogotá? ¿Por qué?

c) ¿Cuáles estrategias implementaría usted para mejorar la movilidad en Bogotá? ¿Por qué?

El **Sistema Inteligente de Movilidad de Medellín** (SIMM) nace en 2011 como respuesta a la necesidad de optimización de la red vial de la ciudad. Se trata de un proyecto pionero en Colombia, en el uso de **tecnologías de la información y la comunicación**, para mejorar las condiciones en que se movilizan las personas por las calles de la ciudad a través de la integración de siete grandes componentes tecnológicos. El **SIMM también s**urge como una **respuesta** a la necesidad de optimización de la red vial de la ciudad, que en los últimos años ha desbordado su capacidad ante la creciente demanda del parque automotor, actualmente cerca de un millón de vehículos circula por las vías de Medellín. (Secretaría de Movilidad de Medellín, 2011)[16]

Para realizar un análisis estadístico relacionado con la problemática descrita, a continuación, se presenta un enlace que contiene una base de datos correspondiente a una muestra de los datos recopilados por el Sistema Inteligente de Movilidad de Medellín - SIMM, mediante los dispositivos de Cámaras de Circuito cerrado de Televisión (CCTV) y Apoyo a la Red Semafórica (ARS).

https://drive.google.com/uc?export=dowload&id=1CkrfyswP9IOzQD6_k3avqNJj4_SNlyjP

2. Descargue la base de datos a través del enlace anterior y determine cuál es la población, identifique y reporte todas las variables estadísticas que se presentan y luego clasifíquelas.

3. Considerando los datos de la variable estadística "*VELOCIDAD KM/H*", calcule su media y responda: ¿cuáles son las unidades de medida de la media? Para usted, ¿qué significado tiene la media calculada dentro del contexto de la base de datos?

4. El Sistema Inteligente de Movilidad de Medellín (SIMM) está interesado en determinar si los conductores mantienen velocidades cercanas al promedio, lo que sugeriría una ausencia de cambios bruscos en la velocidad, o si, por el contrario, hay fluctuaciones de velocidad muy marcadas. ¿Cómo se puede determinar si **cada conductor** mantiene velocidades cercanas al promedio? Proponga y justifique un procedimiento matemático.

[16] Secretaría de Movilidad de Medellín (2011). Sistema Inteligente de Movilidad de Medellín (SIMM). Tomado de https://www.medellin.gov.co/es/secretaria-de-movilidad/sistema-inteligente-de-movilidad-de-medellin/

Determine si su planteamiento es correcto haciendo uso de la base de datos en Excel.

5. ¿Cuáles son las unidades de medida correspondientes a los datos obtenidos en el punto anterior?

6. Es posible establecer una interpretación para las magnitudes obtenidas en el punto 4, ¿Qué podría establecer al respecto? ¿cuál es el significado de estos valores numéricos teniendo en cuenta el contexto dado?

7. Teniendo en cuenta las respuestas dadas en los ítems 5 y 6, ¿qué interpretación les da a los signos de los valores obtenidos dentro del contexto establecido? ¿Por qué? Justifique su respuesta.

8. Si hay datos para los cuales el signo de los valores numéricos no permita una interpretación adecuada y acorde al contexto, utilice para el conjunto de datos alguna operación o propiedad matemática para asegurarse de que los signos estén en coherencia con la situación planteada. Mencione que propiedad u operación matemática utilizó.

9. El Sistema Inteligente de Movilidad de Medellín (SIMM) quiere presentar un informe para los habitantes de Medellín y, dentro de este, desea sintetizar la información obtenida sobre la variable km/h en el punto 6, para presentarla como **un solo valor numérico**. ¿Qué procedimiento realizaría para poder sintetizar esta información? Realice este proceso en Excel.

10. ¿Cuáles son las unidades de medidas correspondientes al resultado obtenidos en el punto anterior? Mencione si es coherente respecto a la situación planteada y justifique.

11. A partir del valor numérico obtenido (punto 9) junto con las unidades de medida reportadas en el punto 10 ¿Es posible establecer una interpretación adecuada y que corresponda al contexto planteada? De no ser así, utilice alguna operación o propiedad matemática para asegurar que las unidades de medida se puedan interpretar en correspondencia con la situación planteada. Mencione que propiedad u operación matemática utilizó y porqué.

12. ¿Qué representa el valor obtenido en relación con la situación propuesta?

13. El Sistema Inteligente de Movilidad de Medellín (SIMM) planea llevar a cabo un estudio similar el próximo año, sin embargo, con el objetivo de simplificar el proceso, se busca desarrollar una fórmula que sea aplicable independientemente de los datos recopilados. Los pasos para realizar este trabajo se muestran a continuación, completa la tabla para finalizar la deducción de la fórmula deseada por el Sistema Inteligente de Movilidad de Medellín (SIMM).

Pasos	**Formula planteada**
a) Si se tiene en cuenta que la media es $\bar{x}$, se resta la media a cada valor (x_i) del conjunto de datos.	$x_i - \bar{x}$
b) Se eleva al cuadrado cada diferencia.	
c) Se suman todas las diferencias al cuadrado.	$\sum_{i=1}^{n}(x_i - \bar{x})^2$
d) Se divide la suma obtenida entre el número total de datos (n).	
e) Por último, se calcula la raíz cuadrada del paso anterior.	

14. Calcule la varianza y la desviación estándar de la variable estadística "LONGITUD_KM" usando la fórmula establecida en el ítem 12 (desarrolle el procedimiento en Excel).

15. Al usar los comandos de varianza y desviación, ¿los valores numéricos obtenidos son iguales a los calculados con la fórmula?

16. ¿Qué representa en el contexto de la situación los valores obtenidos en ítem 13?

- ✓ **Materiales y recursos:** Para esta tarea se dispone de materiales como: Una guía elaborada en Word que cuenta con 16 puntos, en los cuales se establecen indicaciones y preguntas como los presentados en el apartado de formulación. También, un archivo en hoja de cálculo de Excel en el cual se encuentra la base de datos con la cual los estudiantes deben interactuar y dar respuesta a lo propuesto en el archivo Word. Los estudiantes tienen a la mano computadores propios o de la institución para el desarrollo de esta tarea.

- ✓ **Agrupamiento:** Los estudiantes desarrollarán la tarea de forma individual, se plantea de esta manera teniendo en cuenta que en la implementación se contará con pocos estudiantes.

- ✓ **Interacción:** La comunicación se dará principalmente entre el profesor y los estudiantes, es importante precisar que las preguntas que se resolverán no estarán relacionadas con el contenido sino con la forma (es decir, dudas sobre los enunciados, con las indicaciones dadas, entre otros) y también se realizarán algunas preguntas orientadoras particularmente en los ítems en los que deben elegir y efectuar una operación matemática; lo anterior tiene como finalidad obtener producciones propias de cada estudiante que no estén sesgadas por la opinión del profesor.

- ✓ **Temporalidad:** El tiempo asignado para el desarrollo de la tarea corresponde a 1 hora y 30 minutos.

7.3. Anexo 3

TAREA 3: Coeficiente de variación

Media y desviación

- ✓ **Requisitos:** Para el desarrollo de la tarea es importante precisar que los estudiantes habrán estudiado previamente gráficos estadísticos y lo correspondiente a datos atípicos. Además, los estudiantes ya se encuentran familiarizados con las definiciones de las medidas de tendencia central y de dispersión para datos no agrupados, y con los algoritmos para calcularlas de forma manual (como lo han hecho previamente en las tareas 1 y 2)

- ✓ **Metas:** Se espera que los estudiantes logren establecer interpretación en relación con la dispersión a partir de gráficos estadísticos, y que adicionalmente evidencien como los datos atípicos puede cambiar la variabilidad de un conjunto de datos. También, se espera que reconozcan el uso que se puede dar al coeficiente de variación para comprar la dispersión de dos conjuntos de datos con diferentes unidades de medida.

- ✓ **Formulación:**

A continuación, se presenta el enlace para descargar un archivo de Excel que contiene la base de datos de las calificaciones de los exámenes finales de matemáticas obtenidas por los estudiantes de Tandem durante los tres cortes del semestre 2023-II:

https://drive.google.com/uc?export=download&id=1kFbfaQjK9cMx7ZlSo0dlONSNSaz1LL2

1. Con base en los datos que aparecen en la hoja denominada “Notas” del archivo de Excel descargado, responda las siguientes preguntas.
   - **a)** ¿Cuál es la población de estudio?
   - **b)** ¿Cuáles variables estadísticas se están representando?
   - **c)** Clasifique las variables que identificó en el literal anterior.

2. Considerando el diagrama de caja y el diagrama de dispersión presentados en la base de datos, ¿cómo describiría el desempeño de los estudiantes durante cada uno de los tres cortes?

Corte I	Corte II	Corte III

3. ¿Durante cuál de los cortes los estudiantes fueron más constantes con sus calificaciones? Justifique su respuesta.

4. ¿Cómo varió la dispersión de las calificaciones entre el corte 1 y el corte 2? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.

5. ¿Cómo varió la dispersión de las calificaciones entre el corte 2 y el corte 3? ¿Qué comportamiento de los estudiantes pudo haber causado esta variación? Mencione posibles razones por las cuales las notas entre cortes presentaron dicho cambio.

6. ¿Cuántas y cuáles variables considera usted que influyen en las calificaciones de los estudiantes? ¿Estas variables cambian durante los tres cortes o se mantienen constantes? ¿Por qué piensa eso?

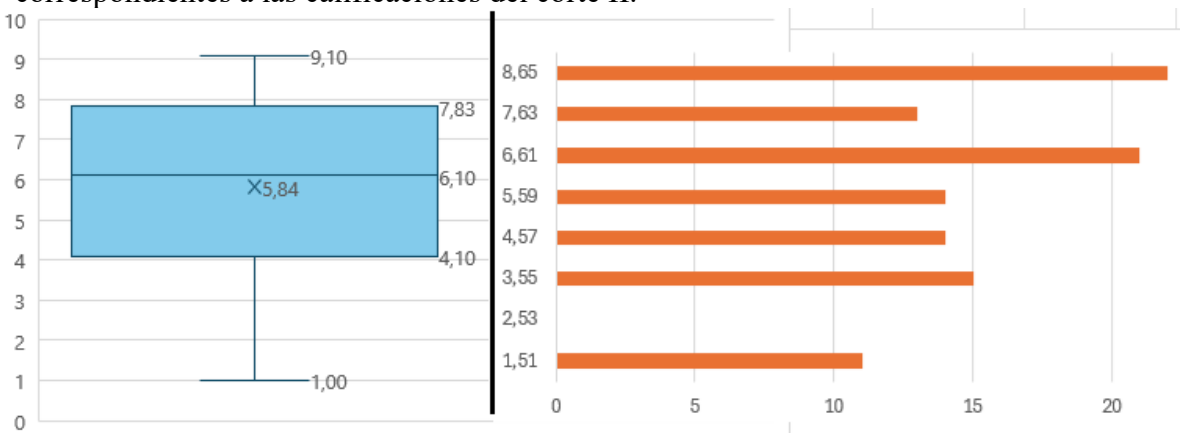
7. Teniendo en cuenta el contexto, que representaría un dato atípico dentro de este.

8. Según el diagrama de cajas, ¿hay datos atípicos en alguna de las cajas? ¿Para cuál o cuáles cortes?

9. ¿Como interpreta los datos atípicos encontrados respecto al corte correspondiente?

10. ¿Como interpreta estos mismos datos respecto a las calificaciones de los tres cortes?

11. El siguiente grafico muestra un paralelo entre el diagrama de caja y el histograma correspondientes a las calificaciones del corte II.



a. ¿Qué información se puede visualizar en el diagrama de cajas que no es posible ver en el histograma? ¿Cree usted que al observar e interpretar únicamente el diagrama de cajas se está ignorando alguna información del conjunto de datos?

b. ¿Qué información se puede visualizar en el histograma que no es posible ver en el diagrama de caja? ¿Cree usted que al observar e interpretar únicamente el histograma se está ignorando alguna información del conjunto de datos?

c. ¿Qué beneficios o dificultades puede traer observar simultáneamente e interpretar varios diagramas del mismo conjunto de datos?

12. A continuación, se aborda una problemática que se ha venido presentando en los últimos años debido a diversos factores ambientales, políticos, etc. Teniendo en cuenta noticias, comentarios o información a la que haya tenido acceso, **sobre la calidad del agua potable**, responda las siguientes preguntas:

a. Según sus concepciones e ideas, ¿a qué hace referencia el concepto "calidad del agua"?

b. ¿Por qué cree que es importante atender los temas relacionados con la "calidad del agua"?

c. ¿Cuáles son las principales fuentes de contaminación de los embalses y cómo han afectado los parámetros de calidad del agua en los últimos diez años?

d. ¿Cómo podría comprobar la veracidad de la información suministrada en la respuesta a la pregunta anterior?

La calidad del agua, de acuerdo con la Organización Mundial de la Salud (OMS) y otros organismos internacionales, se entiende como las condiciones en que se encuentra el agua respecto

a características físicas, químicas y biológicas, en su estado natural o después de ser alteradas por el accionar humano. El deterioro de la calidad del agua se ha convertido en motivo de preocupación a nivel mundial, debido al crecimiento de la población humana, la expansión de la actividad industrial y agrícola y la amenaza del cambio climático como causa de importantes alteraciones en el ciclo hidrológico.[17]

Para realizar un análisis estadístico relacionado con la problemática descrita, acceda a la pestaña "Calidad del agua" de la base de datos trabajada hasta el momento, la cual contiene información correspondiente a la calidad de agua en el Rio Cauca, en 19 estaciones de muestreo, donde se miden algunos parámetros de calidad en el agua, Fisicoquímicos, Microbiológicos y Biológicos.

En Estadística, otra forma de medir la dispersión es con el coeficiente de variación que se define como:

$CV = \frac{s}{\bar{x}}$ Donde s representa la desviación estándar y $\bar{x}$ representa la media

13. La secretaría de medio ambiente desea realizar una comparación entre las variables PH y TEMPERATURA. Calcule la media y la desviación estándar para las variables estadísticas mencionadas y registre la información en la siguiente tabla:

Nota: Escriba el valor numérico junto con las unidades de medida.

	Media	**Desviación estándar**
PH		
Temperatura		

a. La secretaría de medio ambiente encuentra que no es coherente comparar la desviación estándar entre PH y temperatura ¿Por qué cree que la secretaria del medio ambiente afirma esto?

b. Calcule el coeficiente de variación para las variables estadísticas.

[17] Baeza, E. (2016). Calidad del agua. Biblioteca del Congreso Nacional de Chile. Tomado de https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/23747/2/Calidad%20del%20Agua%20Final.pdf

	Coeficiente de variación
PH	
Temperatura	

c. ¿Cuál es la unidad de medida obtenida en ambos casos al aplicar la fórmula correspondiente? Justifique su respuesta.

d. En general, ¿cuál es la unidad de medida correspondiente al coeficiente de variación?

e. Es posible establecer una comparación entre las dos variables estadísticas a partir del valor numérico obtenido correspondiente al coeficiente de variación? ¿por qué?

f. ¿Teniendo en cuenta la respuesta anterior, podría establecer cuál de las variables estadísticas presenta una mayor dispersión?

14. Calcule los coeficientes de variación de siguientes variables estadísticas

Variables	Coeficiente de variación
Solidos Totales	
Solidos Disueltos	
Demanda Bioquímica de Oxigeno	
Demanda química de Oxigeno	
Oxígeno Disuelto	
Dureza Total	
Dureza Cálcica	
Dureza Magnésica	
Conductividad Eléctrica	

a. Realice una conjetura sobre cuál es el mínimo y el máximo valor que puede tomar el coeficiente de variación.

b. ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el mínimo valor?

c. ¿Qué condiciones debe cumplir un conjunto de datos para que el coeficiente de variación tome el máximo valor?

- ✓ **Materiales y recursos:** Para esta tarea se dispone de materiales como: Una guía elaborada en Word que cuenta con 14 puntos, en los cuales se establecen indicaciones y preguntas como los presentados en el apartado de formulación. También, un archivo en hoja de cálculo de Excel en el cual se encuentra la base de datos con la cual los estudiantes deben interactuar y dar respuesta a lo propuesto en el archivo Word. Los estudiantes tienen a la mano computadores propios o de la institución para el desarrollo de esta tarea.

- ✓ **Agrupamiento:** Los estudiantes desarrollarán la tarea de forma individual, se plantea de esta manera teniendo en cuenta que en la implementación se contará con pocos estudiantes.

- ✓ **Interacción:** La comunicación se dará principalmente entre el profesor y los estudiantes, es importante precisar que las preguntas que se resolverán no estarán relacionadas con el contenido sino con la forma (es decir, dudas sobre los enunciados, con las indicaciones dadas, entre otros); lo anterior tiene como finalidad obtener producciones propias de cada estudiante que no estén sesgadas por la opinión del profesor.

- ✓ **Temporalidad:** El tiempo estimado para el desarrollo de esta tarea corresponderá a 1 hora.