



Manual de Estadística



SEGUNDO SEMESTRE 2025

Introducción

La estadística es una herramienta esencial para comprender fenómenos, tomar decisiones informadas y analizar información con rigor. En un entorno cada vez más orientado al uso de datos, su dominio resulta indispensable tanto en el ámbito académico como en la vida cotidiana y profesional. Aprender estadística no se limita a realizar cálculos; implica desarrollar una forma de pensar basada en la evidencia y la interpretación crítica de los resultados.

Este manual de prácticas tiene como propósito guiar al estudiante en el aprendizaje progresivo de la estadística descriptiva. A través de ejercicios estructurados y casos contextualizados, se busca que el estudiante adquiera habilidades técnicas y comprenda el sentido práctico de cada procedimiento. Las prácticas abordan temas fundamentales como la clasificación de variables, el cálculo de medidas de tendencia central y dispersión, el uso de medidas de posición y el análisis de datos agrupados.

El enfoque adoptado privilegia la claridad conceptual, la aplicación práctica y la reflexión sobre los resultados. Se parte de situaciones reales o cercanas al entorno del estudiante para favorecer un aprendizaje significativo. Asimismo, se promueve el desarrollo de competencias como el razonamiento lógico, la organización de información y la toma de decisiones fundamentadas.

Cada sección incluye objetivos específicos, un marco conceptual accesible, ejemplos resueltos y una hoja de trabajo que permite aplicar lo aprendido. El propósito es que el estudiante no solo sepa cómo se calcula una medida estadística, sino que pueda interpretar su significado y aplicarla con criterio.

Este manual busca formar estudiantes capaces de analizar datos con responsabilidad, pensamiento crítico y sentido práctico.

INSTRUCCIONES PARA REALIZAR LA PRÁCTICA

Para la realización adecuada de las prácticas deberán atenderse las siguientes indicaciones:

1. Presentarse puntualmente a la hora del inicio del laboratorio y permanecer durante la duración de este.
2. Realizar las actividades y hojas de trabajo planteadas durante la práctica.
3. Participación y cuidado de cada uno de los integrantes del grupo en todo momento de la práctica.
4. Conocer la teoría, (leer el manual antes de presentarse a cada práctica).
- 5. No se permite el uso de teléfono celular dentro del laboratorio,**
6. Si tiene llamadas laborales deberá atender las mismas únicamente en el horario de receso. Si sale del salón de clases sin la autorización del docente perderá el valor de la práctica.
7. No puede atender visitas durante la realización de la práctica.
8. El horario de receso es únicamente de 15 minutos.
- 9. Respeto dentro del laboratorio hacia los catedráticos o compañeros (as).**

La falta a cualquiera de los incisos anteriores será motivo de una inasistencia.

Considere que se prohíbe terminantemente comer, beber y fumar. Éstos también serán motivos para ser retirado de la práctica.

Recuerde que para tener derecho al punteo y aprobar el curso deberá presentarse a las prácticas y realizar las evaluaciones en línea, las cuales estarán habilitadas del 27 de octubre 2025 a las 8:00 al 31 de octubre 2025 a las 18:00.

INFORME DE PRÁCTICA

Las secciones de las cuales consta un informe, el punteo de cada una y el orden en el cual deben aparecer son las siguientes:

- a) Resultados
- b) Resumen de la práctica
- c) Conclusiones

Si se encuentran dos informes parcial o totalmente parecidos se anularán automáticamente dichos reportes. a.

- a) **RESULTADOS:** Es la sección en la que se presentan de manera clara y objetiva los datos obtenidos a partir de la práctica realizada.
- b) **RESUMEN DE LA PRÁCTICA:** Esta sección corresponde al contenido del informe, aquello que se ha encargado realizar según las condiciones del laboratorio.
- c) **CONCLUSIONES:** Constituyen la parte más importante del informe. Son las decisiones tomadas, respuestas a interrogantes o soluciones propuestas a las actividades planteadas durante la práctica.

DETALLES FÍSICOS DEL INFORME

El informe debe presentarse en hojas de papel bond **tamaño carta**.

Cada sección descrita anteriormente, debe estar debidamente identificada y en el orden establecido. Todas las partes del informe deben estar escritas a mano CON LETRA CLARA Y LEGIBLE, a menos que se indique lo contrario. Se deben utilizar ambos lados de la hoja. No debe traer folder ni gancho, simplemente engrapado.

IMPORTANTE: Los informes se entregarán al día siguiente de la realización de la práctica al entrar al laboratorio **SIN EXCEPCIONES**. Todos los implementos que se utilizarán en la práctica se tengan listos antes de entrar al laboratorio pues el tiempo es muy limitado. Todos

los trabajos y reportes se deben de entregar en la semana de laboratorio no se aceptará que se entregue una semana después.

Práctica 1

Tipos de datos, variables, niveles de medición y graficas

Objetivos de la práctica

1. Comprender el concepto de estadística, su utilidad en la vida cotidiana y su división en descriptiva e inferencial.
2. Identificar y clasificar variables según su tipo (cualitativa o cuantitativa) y su nivel de medición (nominal, ordinal, intervalo o razón).
3. Organizar datos mediante tablas de frecuencia y representarlos gráficamente para facilitar su análisis e interpretación.

Marco Teórico

La estadística es una herramienta que te permite entender mejor el mundo a través de los datos. Cada vez que cuentas cuántas personas llegaron a clase, revisas cuántos puntos obtuviste en un examen o comparas precios en el supermercado, estás usando estadística sin darte cuenta.

En términos simples, la estadística es la ciencia que se encarga de:

- ✓ Recolectar datos
- ✓ Organizar datos
- ✓ Analizar datos
- ✓ Interpretar resultados

Todo esto con el fin de tomar decisiones más informadas.

¿Para qué sirve la estadística?

- ✓ La estadística no es solo para matemáticos o científicos. La usas a diario, aunque no lo notes. Sirve para:

- ✓ Tomar decisiones personales (¿cuánto gastas al mes?, ¿cuál es tu promedio?)
- ✓ Entender fenómenos sociales (¿cuántas personas están desempleadas?, ¿qué enfermedades son más comunes?)
- ✓ Mejorar procesos en empresas (¿qué producto se vende más?, ¿cuál es el horario con más clientes?)
- ✓ Evaluar resultados en educación, salud, economía, política y más

En resumen, la estadística te ayuda a ver patrones, identificar problemas y encontrar soluciones basadas en hechos, no en suposiciones.

División de la estadística

La estadística se divide en dos grandes ramas:

1. Estadística descriptiva

Es la parte que se encarga de resumir y organizar los datos. Aquí se usan:

- ✓ Tablas
- ✓ Gráficos
- ✓ Promedios
- ✓ Porcentajes

Su objetivo es mostrar lo que los datos dicen, sin hacer predicciones.

Ejemplo: Si encuestas a 50 personas sobre su sabor de helado favorito y haces una tabla con los resultados, estás usando estadística descriptiva.

2. Estadística inferencial

Va un paso más allá. Toma una parte de los datos (una muestra) y saca conclusiones sobre un grupo más grande (la población).

Ejemplo: Si encuestas a 100 estudiantes y con esos datos concluyes qué opina toda la universidad, estás usando estadística inferencial.

¿Por qué es importante aprender estadística?

Porque te da herramientas para pensar con lógica, analizar información y tomar decisiones con base en evidencia. En un mundo lleno de datos, quien sabe interpretarlos tiene una gran ventaja.

Variables

Cuando tomas decisiones simples como qué ropa usar, cuánto dinero llevar al salir, o cuántas horas dormir, estás manejando **datos personales**. Muchos de esos datos cambian de persona en persona o de situación en situación. A esos elementos que pueden cambiar, la estadística los llama **variables**. Imagina que estás en un parque y observas a las personas que pasan. Algunas vienen solas, otras con mascotas, algunas están sonriendo, otras hablando por celular. Todo eso es información variable: la edad, el estado de ánimo, el color de la ropa, el tipo de transporte. ¡Tu entorno está lleno de variables sin que te des cuenta!

¿Qué es una variable?

Una *variable* es cualquier característica que puede tomar valores distintos entre individuos o situaciones. Por ejemplo:

- ✓ En una clase: edad, género, calificación.
- ✓ En tu familia: número de hermanos, nivel educativo, ingresos.
- ✓ En la calle: tipo de vehículo, color del semáforo, número de personas en fila.

¿Cuáles son los tipos de variable?

En estadística descriptiva, las variables se dividen en dos grandes grupos:

- ✓ **Cuantitativas:** se expresan en números. Se pueden contar o medir. Ejemplos: edad, número de hijos, ingresos mensuales, peso, estatura.
- ✓ **Cualitativas:** se expresan con palabras o categorías. Describen cualidades. Ejemplos: color de ropa, estado civil, profesión, sabor favorito, nacionalidad.

Ejercicio guiado

Instrucción: Observa tu entorno (aula, hogar, barrio) y lista **10 datos que podrías recolectar**. Luego completa la siguiente tabla de acuerdo con el tipo de variable:

Dato observado	Tipo de variable	¿Por qué?
Edad	Cuantitativa	Se mide con números (años)
Color de mochila	Cuantitativa	Es una palabra, no número
Número de hermanos	Cuantitativa	Se puede contar en unidades
Tipo de música favorita	Cualitativa	Describe una preferencia (ej. rock, salsa)
Horas de sueño por noche	Cuantitativa	Se mide en horas (ej. 6, 8, 10)
Marca de celular	Cualitativa	Es un número asignado (ej. 85 puntos)
Tipo de transporte usado	Cualitativa	Se cuenta fácilmente (ej. 3 tareas)

El estudiante puede agregar otros ejemplos del aula, hogar o calle.

¿Hay datos que usas a diario sin darte cuenta que son variables?

Sí. Desde revisar la temperatura en el celular, calcular el tiempo que tomarás en llegar a clase, hasta elegir tu comida favorita, estás usando variables todo el tiempo. Identificarlas te permite entender el mundo con más orden y preparar tus decisiones de forma informada.

Clasificación por nivel de medición

Cuando recolectas datos, no todos pueden usarse de la misma forma. Algunos sólo te dicen "qué" sin permitir cálculos, mientras otros te dejan medir, ordenar y hasta comparar. Esto depende del **nivel de medición** del dato.

Conocer el nivel te ayuda a saber **qué tipo de análisis es válido**. No puedes calcular un promedio de colores favoritos, pero sí puedes hacerlo con edades. Esta clasificación ordena los datos según qué tan "medibles" o "operables" son.

Nivel Nominal

Es el más básico. Se usa para **nombrar o etiquetar** cosas sin que exista ningún tipo de orden entre ellas. No puedes decir que una es "más" o "menos" que otra, porque simplemente son distintas categorías.

Ejemplos cotidianos:

- ✓ Marcas de celular: Samsung, Motorola, iPhone.
- ✓ Género: masculino, femenino.
- ✓ Sabor de helado: chocolate, vainilla, fresa.

En este nivel no puedes hacer operaciones matemáticas ni comparaciones. Solo te sirve para identificar o agrupar.

Nivel Ordinal

Aquí ya puedes establecer un orden, pero no sabes cuánto más o cuánto menos hay entre los datos. Sirve para clasificar por niveles, rangos o escalas. Ejemplos cotidianos:

- ✓ Opinión sobre el clima: malo < regular < bueno.
- ✓ Tamaño de ropa: S, M, L, XL.
- ✓ Lugar obtenido en una carrera: 1°, 2°, 3°.

No puedes hacer sumas ni promedios, pero sí puedes decir quién está por encima o debajo en la clasificación.

Nivel de Intervalo

En este nivel los datos se pueden ordenar y medir, y además conoces la diferencia entre ellos. Sin embargo, el cero no significa la ausencia total.

Ejemplos cotidianos:

- ✓ Temperatura en °C: 25°C, 30°C, 35°C (el 0 no significa “sin calor”).
- ✓ Fechas del calendario: 1990, 2000, 2024.
- ✓ Estatura de un salón de estudiantes en centímetros 165, 178, 180

Puedes calcular diferencias, pero no usar el cero para decir “nada” o “inexistente”.

Nivel de Razón

Es el más completo. Los datos se pueden ordenar, medir, comparar y tienen un cero verdadero, lo que significa que podés hacer operaciones matemáticas como dividir o calcular porcentajes.

Ejemplos cotidianos:

- ✓ Edad: si alguien tiene 0 años, literalmente acaba de nacer.
- ✓ Distancia recorrida: 0 metros = no se ha movido.
- ✓ Dinero en la billetera: Q0 = no hay dinero.

Este nivel permite **todo tipo de análisis estadístico** porque tiene sentido matemático completo.

Los cuatro niveles de medición

Nivel	Lo que permite	Ejemplo real
Nominal	Solo identificar y categorizar	Marca de celular, género, sabor favorito
Ordinal	Ordenar, pero sin conocer la magnitud	Opinión del clima, rangos laborales
Intervalo	Medir diferencias, sin cero absoluto	Temperatura (°C), calendario (años)
Razón	Medir con operaciones matemáticas	Edad, distancia, ingresos

El estudiante puede agregar otros ejemplos del aula, hogar o calle.

Frecuencias

Cuando tienes muchos datos sin orden, como los resultados de una encuesta, puede ser difícil entender qué está pasando. ¿Qué opción fue la más elegida? ¿Cuántas veces se repite cada respuesta? Para organizar esa información, en estadística se utiliza una **tabla de frecuencia**, que resume los datos y permite visualizarlos de manera sencilla.

Una tabla de frecuencia muestra cuántas veces aparece cada valor en tu conjunto de datos. También puedes incluir porcentajes y frecuencias acumuladas para tener una idea más completa del comportamiento de los datos.

Datos del caso:

Se aplicó una encuesta a 50 personas para conocer su sabor favorito de helado. Las respuestas fueron las siguientes:

Sabor de helado	Frecuencia
Chocolate	20
Vainilla	15
Fresa	10
Mango	5

Paso a paso:

A partir de estos datos, construye una **tabla de frecuencia ampliada** que contenga:

1. **Frecuencia absoluta:** cuántas personas eligieron cada sabor.
2. **Frecuencia relativa (%):** qué porcentaje representa cada grupo sobre el total de encuestados.
3. **Frecuencia acumulada (%):** suma progresiva de los porcentajes.

Instrucciones:

1. Calcula la **frecuencia relativa** con esta fórmula: $(\text{Frecuencia del sabor} \div \text{Total de respuestas}) \times 100$
2. Suma los porcentajes de cada fila para obtener la **frecuencia acumulada**.

Tabla modelo:

Sabor de helado	Frecuencia absoluta (%)	Frecuencia relativa (%)
Chocolate	20 40%	40%
Vainilla	15 30%	70%
Fresa	10 20%	90%
Mango	5 10%	100%

Elaboración de tres tipos de gráfico

Una tabla de datos es útil, pero muchas veces no basta para entender rápidamente qué está ocurriendo. Los gráficos permiten transformar los números en imágenes que muestran patrones, comparaciones y tendencias de forma clara. En esta actividad vas a construir tres tipos de gráficos distintos para representar los datos de los sabores de helado y el consumo diario.

Datos a utilizar:

Tabla de sabores favoritos (frecuencia):

Sabor de helado	Frecuencia	Porcentaje
Chocolate	20 40%	40%
Vainilla	15 30%	70%
Fresa	10 20%	90%
Mango	5 10%	100%

Instrucciones:

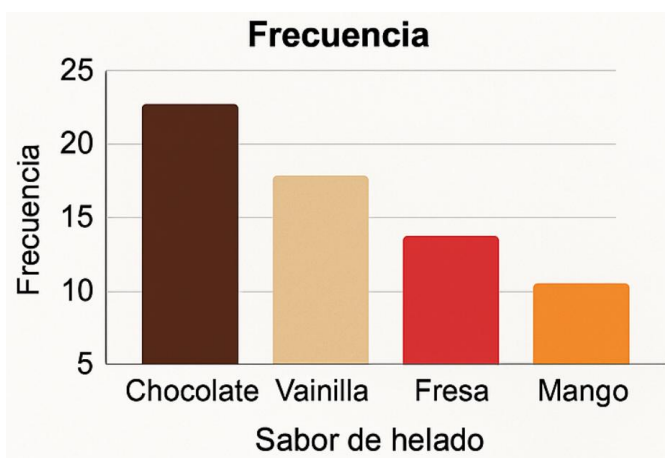
- Realiza los tres gráficos en hoja bond o en Excel.
- Usa colores distintos para cada categoría.
- Coloca títulos, etiquetas y leyendas claras.
- Al finalizar, responde la pregunta: ¿Qué gráfico muestra mejor la información?

1. Histograma (barras verticales)

Este gráfico muestra la frecuencia de cada sabor. Cada barra representa un sabor y su altura indica cuántas personas lo eligieron.

Pasos:

- En el eje horizontal coloca los sabores: Chocolate, Vainilla, Fresa, Mango.
- En el eje vertical coloca la escala de frecuencia (de 0 a 25).
- Dibuja una barra para cada sabor según su frecuencia.

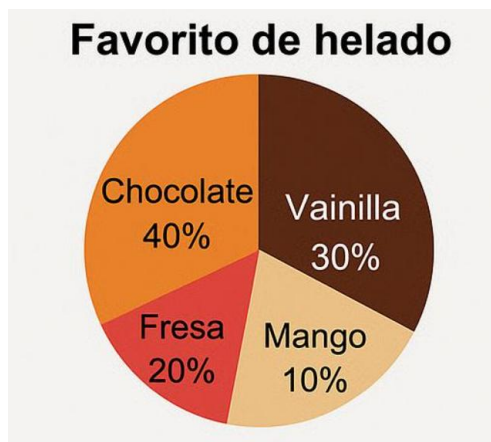


2. Gráfico de pastel (circular)

Este gráfico muestra la proporción que representa cada sabor respecto al total. Es útil para ver qué parte del “todo” representa cada categoría.

Pasos:

- Dibuja un círculo y divídelo en secciones según el porcentaje de cada sabor.
- Chocolate ocupará el 40% del círculo, vainilla el 30%, fresa el 20% y mango el 10%.
- Colorea cada sección y coloca etiquetas con el nombre y porcentaje.

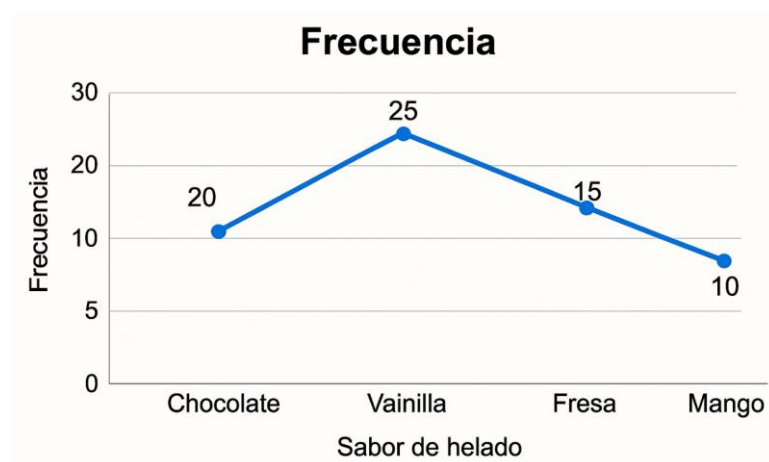


3. Gráfico de línea

Este gráfico muestra cómo varía la cantidad de helados. Es útil para observar tendencias a lo largo del tiempo.

Pasos:

- En el eje horizontal coloca el sabor de helado.
- En el eje vertical coloca la cantidad de helados
- Marca un punto por cantidad de helados y une los puntos con líneas.
- Observa si hay aumento, descenso o estabilidad.



Hoja de trabajo

Ejercicio 1: Comprensión del concepto de estadística

Lee el siguiente caso y responde:

Una asociación juvenil quiere conocer los hábitos de lectura de los estudiantes de su comunidad. Para ello, aplican una encuesta a 80 jóvenes con preguntas como: ¿cuántos libros leíste el mes pasado?, ¿qué género literario prefieres?, ¿cuánto tiempo dedicas a la lectura diaria?, ¿cómo calificarías tu experiencia lectora?

a) ¿Qué tipo de datos se están recolectando: ¿cualitativos, cuantitativos o ambos? Justifica tu respuesta. b) ¿Este estudio corresponde a estadística descriptiva o inferencial? Explica por qué. c) ¿Qué decisiones podría tomar la asociación con base en los resultados?

Ejercicio 2: Clasificación de variables

A partir del caso anterior, completa la siguiente tabla:

Variable observada	Tipo de variable	Nivel de medición	Justificación breve
Número de libros leídos en un mes			
Género literario preferido			
Tiempo diario dedicado a la lectura			
Opinión sobre la experiencia lectora			

Ejercicio 3: Clasificación de variables en otros contextos

Observa las siguientes situaciones y completa la tabla:

Situación observada	Variable involucrada	Tipo de variable	Nivel de medición	Justificación breve
Tiempo que tarda una persona en llegar al trabajo				
Marca de zapatos preferida por adolescentes				
Número de veces que una persona revisa su celular al día				
Nivel de satisfacción con el servicio de transporte				

Ejercicio 4: Construcción de tabla de frecuencia

Se registró la cantidad de veces que 40 personas visitaron una tienda en línea durante una semana. Estos fueron los resultados:

2, 3, 4, 2, 5, 3, 3, 4, 2, 3, 5, 4, 3, 2, 4, 5, 3, 2, 4, 3, 5, 3, 2, 4, 3, 5, 2, 3, 4, 3, 5, 2, 3, 4, 2, 5, 3, 2, 4, 3

a) Agrupa los datos y completa la tabla:

Visitas a la tienda	Frecuencia absoluta	Frecuencia relativa (%)	Frecuencia acumulada (%)
2			
3			
4			
5			

b) ¿Cuál es la frecuencia más común?

c) ¿Qué porcentaje representa el grupo que visitó la tienda 4 veces o más

Ejercicio 5: Elaboración de gráficos

Con base en la tabla anterior, realiza los siguientes gráficos en hoja bond o en Excel:

a) Histograma (barras verticales)

b) Gráfico de pastel (circular)

c) Gráfico de línea (usa los siguientes datos de visitas por día):

Día	Visitas registradas
Lunes	10
Martes	12
Miércoles	8
Jueves	5
Viernes	5

Incluye título, etiquetas, escala y colores.

Ejercicio 6: Análisis gráfico

Con base en los gráficos que realizaste, responde:

a) ¿Qué gráfico te pareció más claro para interpretar los datos? ¿Por qué? b) ¿Qué día tuvo más visitas y qué explicación podrías dar? c) ¿Qué conclusiones puedes sacar sobre el comportamiento de los usuarios?

Ejercicio 7: Reflexión final

Responde con tus propias palabras:

- a) ¿Qué aprendiste sobre el uso de variables y datos en situaciones reales?
- b) ¿Cómo puede ayudarte la estadística a tomar mejores decisiones en tu vida diaria?
- c) ¿Qué habilidades crees que desarrollaste con esta práctica?

Practica 2

Medidas de tendencia central y dispersión

Objetivos de la práctica

1. Comprender el significado y la utilidad de las medidas de tendencia central y dispersión.
2. Calcular e interpretar media, mediana, moda, rango, varianza, desviación estándar, coeficiente de variación y coeficiente de sesgo.
3. Comparar dos conjuntos de datos para analizar cuál es más representativo y homogéneo.

Marco Teórico

Cuando se recolectan datos numéricos, como los gastos, las calificaciones o los tiempos de espera, es importante encontrar un valor que represente al grupo. Ese valor central puede ayudarte a resumir la información. Pero también es importante saber si los datos están muy dispersos o si todos están cerca de ese valor central. Por eso, en esta práctica vas a trabajar con dos tipos de medidas:

- ✓ Las medidas de tendencia central, que indican dónde se concentra la mayoría de los datos. (media, mediana, moda)
- ✓ Las que muestran qué tanto se alejan los datos entre sí (rango, varianza, desviación estándar)

Estas herramientas permiten comparar grupos, identificar comportamientos extremos y tomar decisiones con base en evidencia.

Medidas de tendencia central

Las medidas de tendencia central son valores que resumen un conjunto de datos con un solo número representativo. Las más comunes son:

- Media: el promedio. Se obtiene sumando todos los datos y dividiendo entre la cantidad de datos.
- Mediana: el valor que queda en el centro cuando los datos están ordenados.

- Moda: el valor que más se repite.

Caso guía: Gastos semanales en transporte

Se encuestó a dos grupos de estudiantes sobre cuánto gastan en transporte durante una semana.

Grupo A	Grupo B
Q10	Q5
Q15	Q5
Q20	Q25
Q25	Q45
Q30	Q60

Cálculo de la media

Fórmula: Media = (Suma de los datos) ÷ (Cantidad de datos)

Grupo A: $(10 + 15 + 20 + 25 + 30) \div 5 = 100 \div 5 = Q20$

Grupo B: $(5 + 5 + 25 + 45 + 60) \div 5 = 140 \div 5 = Q28$

Análisis: El promedio de gastos es mayor en el Grupo B, pero eso no significa que todos gasten más. Hay que ver cómo están distribuidos los datos.

Cálculo de la mediana

Paso 1: Ordenar los datos

Grupo A: 10, 15, 20, 25, 30 → Mediana = 20

Grupo B: 5, 5, 25, 45, 60 → Mediana = 25

Análisis: Aunque el promedio del Grupo B es más alto, su mediana está más cerca del centro real de los datos. Esto sugiere que hay valores extremos que están influyendo en la media.

Cálculo de la moda

Grupo A: No hay ningún valor que se repita → Sin moda

Grupo B: El valor 5 se repite dos veces → Moda = 5

Análisis: La moda del Grupo B muestra que hay estudiantes que gastan muy poco, aunque otros gasten mucho más.

Medidas de dispersión

Las medidas de dispersión indican qué tan separados están los datos entre sí. Permiten saber si los valores están concentrados cerca del promedio o si hay mucha variabilidad.

- ✓ Rango: Diferencia entre el valor más alto y el más bajo.
- ✓ Varianza: Promedio de las desviaciones cuadradas respecto a la media.
- ✓ Desviación estándar: Raíz cuadrada de la varianza. Mide cuánto se alejan los datos del promedio.
- ✓ Coeficiente de variación (CV): Relación entre la desviación estándar y la media. Se expresa en porcentaje. Sirve para comparar la dispersión entre grupos con medias distintas.
- ✓ Coeficiente de sesgo (asimetría): Indica si los datos están distribuidos de forma simétrica o si hay inclinación hacia un lado.

Cálculo del rango

Fórmula: $\text{Rango} = \text{Valor máximo} - \text{Valor mínimo}$

Grupo A: $30 - 10 = 20$

Grupo B: $60 - 5 = 55$

Análisis: El Grupo B tiene un rango mucho mayor, lo que indica que hay más diferencia entre quienes gastan poco y quienes gastan mucho.

Cálculo de la varianza y desviación estándar

Fórmulas:

- ✓ Varianza = Suma de los cuadrados de las desviaciones $\div$ n
- ✓ Desviación estándar = $\sqrt{\text{Varianza}}$

Paso 1: Calcular la desviación de cada dato respecto a la media

(Desviación = dato – media)

Paso 2: Elevar al cuadrado cada desviación

Paso 3: Sumar los cuadrados

Paso 4: Dividir entre n

Paso 5: Sacar la raíz cuadrada

Cálculo para Grupo A (media = 20)

Dato	Dato – Media	(Dato – Media) ²
10	-10	100
15	-5	25
20	0	0
25	5	25
30	10	100
		250

Varianza = $250 \div 5 = 50$

Desviación estándar = $\sqrt{50} \approx 7.07$

Cálculo para Grupo B (media = 28)

Dato	Dato – Media	(Dato – Media) ²
5	-23	529
5	-23	529
25	-3	9
45	17	289
60	32	1024
		2380

Varianza = $2380 \div 5 = 476$

Desviación estándar = $\sqrt{476} \approx 21.82$

Análisis comparativo:

- ✓ El Grupo A tiene una desviación estándar baja (7.07), lo que indica que sus gastos están concentrados cerca del promedio.
- ✓ El Grupo B tiene una desviación estándar alta (21.82), lo que muestra que hay mucha variación entre los estudiantes.

Cálculo del coeficiente de variación (CV)

$CV = (\text{Desviación estándar} \div \text{Media}) \times 100$

- ✓ Grupo A: $(7.07 \div 20) \times 100 \approx 35.35\%$
- ✓ Grupo B: $(21.82 \div 28) \times 100 \approx 77.93\%$

El Grupo B tiene más del doble de variabilidad relativa que el Grupo A.

Cálculo del coeficiente de sesgo (asimetría)

Una forma sencilla de estimar el sesgo es comparar la relación entre media, mediana y moda:

- Si Media > Mediana > Moda → Sesgo positivo (asimetría a la derecha)
- Si Moda > Mediana > Media → Sesgo negativo (asimetría a la izquierda)
- Si Media ≈ Mediana ≈ Moda → Distribución simétrica
- Grupo A: Media = Mediana = 20 → Distribución simétrica
- Grupo B: Media (28) > Mediana (25) > Moda (5) → Sesgo positivo

El Grupo B tiene una distribución asimétrica hacia la derecha, con valores extremos altos.

Tabla comparativa final

Medida	Grupo A	Grupo B
Media	Q20	Q28
Mediana	Q20	Q25
Moda	—	Q5
Rango	Q20	Q55
Varianza	50	476
Desviación estándar	7.07	21.82
Coefficiente de variación	35.35%	77.93%
Coefficiente de sesgo	Simétrico	Positivo

Conclusión

Aunque el Grupo B tiene un promedio más alto, sus datos están mucho más dispersos y asimétricos. Esto significa que no todos gastan lo mismo: algunos gastan muy poco y otros mucho más. En cambio, el Grupo A es más homogéneo: todos gastan cantidades similares, y su distribución es equilibrada.

Hoja de trabajo

Ejercicio 1: Gasto semanal en transporte

Se desea comparar el gasto semanal en transporte de estudiantes de dos carreras universitarias: Derecho y Enfermería. Se seleccionaron cinco estudiantes de cada carrera y se registró cuánto gastan en promedio por semana (en quetzales):

Derecho (Q)	Enfermería (Q)
20	10
25	10
30	25
35	50
40	60

1. Calcula para cada grupo:

- ✓ Media
- ✓ Mediana
- ✓ Moda
- ✓ Rango
- ✓ Varianza
- ✓ Desviación estándar
- ✓ Coeficiente de variación
- ✓ Tipo de sesgo (positivo, negativo o simétrico)

2. Interpreta: a) ¿Cuál grupo tiene un gasto promedio más alto? b) ¿Cuál grupo presenta mayor desigualdad en los gastos? c) ¿Qué medida central representa mejor a cada grupo: la media o la mediana? d) ¿Cuál grupo es más homogéneo en su comportamiento? Justifica.

Ejercicio 2: Calificaciones en dos secciones

Se comparan las calificaciones obtenidas por estudiantes de dos secciones distintas en un examen de estadística:

Sección A	Sección B
65	50
70	55
75	60
80	90
85	95

1. Calcula todas las medidas estadísticas para ambas secciones.
2. Interpreta: a) ¿Cuál sección tiene mejor rendimiento promedio? b) ¿Cuál sección es más consistente en sus resultados? c) ¿Qué sección tiene una distribución más equilibrada? d) ¿Qué conclusiones podrías dar al docente sobre el desempeño de cada grupo?

Ejercicio 3: Tiempo de llegada a clase

Se registró el tiempo (en minutos) que tardan dos estudiantes en llegar a clase durante cinco días:

Estudiante A	Estudiante B
15	10
20	10
25	20
30	40
35	50

1. Calcula todas las medidas estadísticas para ambos estudiantes.
2. Interpreta: a) ¿Quién tiene un tiempo promedio de llegada más alto? b) ¿Quién es más puntual y constante? c) ¿Qué estudiante presenta mayor variabilidad? d) ¿Qué medida central representa mejor a cada uno?

Ejercicio 4: Consumo de datos móviles

Se comparó el consumo semanal de datos móviles (en GB) de dos jóvenes durante cinco semanas:

Persona A	Persona B
2.5	1.0
3.0	1.0
3.5	2.0
4.0	5.0
4.5	6.0

1. Calcula todas las medidas estadísticas para ambos casos.
2. Interpreta: a) ¿Quién consume más datos en promedio? b) ¿Quién tiene un patrón de consumo más estable? c) ¿Qué medida central representa mejor a cada uno? d) ¿Qué conclusiones podrías dar a una empresa de telefonía para ofrecer un plan adecuado?

Ejercicio 5: Ingresos mensuales

Se comparan los ingresos mensuales (en quetzales) de dos trabajadores informales durante cinco meses:

Trabajador A	Trabajador B
2000	1000
2100	1000
2200	1500
2300	3000
2400	4000

1. Calcula todas las medidas estadísticas para ambos trabajadores.
2. Interpreta: a) ¿Quién tiene un ingreso promedio más alto? b) ¿Quién tiene ingresos más estables? c) ¿Qué trabajador enfrenta mayor incertidumbre económica? d) ¿Qué medida usarías para representar mejor su situación?

Practica 3

Medidas de posición

Objetivos de la practica

1. Comprender el significado y la utilidad de las medidas de posición en el análisis de datos.
2. Calcular cuartiles, deciles y percentiles en un conjunto de datos ordenados.
3. Interpretar los resultados para ubicar valores dentro de una distribución y tomar decisiones informadas.

Marco teórico

Cuando se analiza un conjunto de datos, no siempre basta con conocer el promedio o la dispersión. A veces es necesario saber en qué lugar se encuentra un valor dentro del grupo. Por ejemplo, si un estudiante obtuvo 72 puntos en una prueba, ¿eso es alto o bajo? ¿Está entre los mejores o en la mitad?

Las medidas de posición permiten responder esas preguntas. Estas herramientas dividen los datos en partes iguales y ubican un valor dentro de la distribución. Son muy útiles para clasificar, comparar y tomar decisiones basadas en rangos. Se utilizan en educación, salud, economía, deportes y muchas otras áreas.

Las medidas de posición son valores que dividen un conjunto de datos ordenados en partes iguales. Permiten ubicar un dato dentro de la distribución y saber qué porcentaje de observaciones están por debajo o por encima de él.

Tipos de medidas de posición

- ✓ Cuartiles (Q): Dividen los datos en cuatro partes iguales.
 - Q1 (primer cuartil): el 25% de los datos está por debajo.
 - Q2 (segundo cuartil): equivale a la mediana (50%).
 - Q3 (tercer cuartil): el 75% de los datos está por debajo.

- ✓ Deciles (D): Dividen los datos en diez partes iguales.
 - $D1 = 10\%$, $D2 = 20\%$, ..., $D9 = 90\%$.
- ✓ Percentiles (P): Dividen los datos en cien partes iguales.
 - $P25 = Q1$, $P50 = \text{mediana}$, $P75 = Q3$.
 - $P80$, por ejemplo, indica que el 80% de los datos están por debajo de ese valor.

Estas medidas permiten responder preguntas como:

- ✓ ¿Qué puntaje supera al 75% de los estudiantes?
- ✓ ¿En qué lugar se encuentra un valor dentro del grupo?
- ✓ ¿Qué porcentaje de observaciones están por debajo de cierto dato?

Caso guía: Puntajes en una prueba

Se aplicó una prueba de razonamiento lógico a 20 estudiantes. Estos fueron los puntajes (ordenados de menor a mayor):

45, 48, 50, 52, 55, 57, 58, 60, 61, 62, 63, 65, 66, 68, 70, 72, 75, 78, 80, 85

Cuartiles

Los cuartiles dividen el conjunto de datos en cuatro partes iguales. Cada cuartil representa un 25% del total. Son útiles para identificar valores bajos, medios y altos dentro de una distribución.

Paso 1: Identificar la posición

Fórmula para la posición del cuartil: $Q_k = (k \times n) \div 4$ Donde:

- ✓ k = número del cuartil (1, 2 o 3)
- ✓ n = número total de datos

Para este caso ($n = 20$):

- ✓ $Q1 = (1 \times 20) \div 4 = \text{posición } 5$
- ✓ $Q2 = (2 \times 20) \div 4 = \text{posición } 10$
- ✓ $Q3 = (3 \times 20) \div 4 = \text{posición } 15$

Paso 2: Ubicar los valores

- ✓ $Q1 = 5.^\circ \text{ dato} = 55$
- ✓ $Q2 = 10.^\circ \text{ dato} = 62$
- ✓ $Q3 = 15.^\circ \text{ dato} = 70$

Interpretación:

- ✓ El 25% de los estudiantes obtuvo menos de 55 puntos.
- ✓ El 50% obtuvo menos de 62 puntos.
- ✓ El 75% obtuvo menos de 70 puntos.

Cálculo de deciles

¿Qué son los deciles?

Los deciles dividen los datos en diez partes iguales. Cada decil representa un 10% del total. Son útiles para hacer clasificaciones más detalladas que los cuartiles.

Fórmula para la posición del decil:

$$D_k = (k \times n) \div 10$$

Ejemplo:

- ✓ $D3 = (3 \times 20) \div 10 = \text{posición } 6 \rightarrow 57$
- ✓ $D7 = (7 \times 20) \div 10 = \text{posición } 14 \rightarrow 68$

Interpretación:

- ✓ El 30% de los estudiantes obtuvo menos de 57 puntos.
- ✓ El 70% obtuvo menos de 68 puntos.

Cálculo de percentiles

¿Qué son los percentiles?

Los percentiles dividen los datos en cien partes iguales. Cada percentil representa un 1% del total. Son útiles para ubicar con precisión un valor dentro de una población.

Fórmula para la posición del percentil:

$$Pk = (k \times n) \div 100$$

Ejemplo:

- ✓ $P25 = (25 \times 20) \div 100 = \text{posición } 5 \rightarrow \mathbf{55}$
- ✓ $P50 = \text{posición } 10 \rightarrow \mathbf{62}$
- ✓ $P80 = (80 \times 20) \div 100 = \text{posición } 16 \rightarrow \mathbf{72}$

Interpretación:

- ✓ $P80 = 72 \rightarrow$ El 80% de los estudiantes obtuvo menos de 72 puntos.

Tabla resumen del caso guía

Medida	Valor	Interpretación breve
Q1	55	25% obtuvo menos de 55
Q2 (mediana)	62	50% obtuvo menos de 62
Q3	70	75% obtuvo menos de 70
D3	57	30% obtuvo menos de 57
D7	68	70% obtuvo menos de 68
P25	55	Coincide con Q1
P50	62	Coincide con Q2
P80	72	80% obtuvo menos de 72

Hoja de trabajo

Caso 1: Cuartiles, deciles y percentiles

Contexto: Tiempo de uso diario del celular. Un grupo de 20 estudiantes registró cuántos minutos al día usan su celular para redes sociales. Los datos (en minutos) ya están ordenados de menor a mayor:

45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140

1. Calcula los siguientes valores:

- ✓ Primer cuartil (Q1)
- ✓ Segundo cuartil (Q2 o mediana)
- ✓ Tercer cuartil (Q3)

2. Calcula los siguientes deciles:

- ✓ D2
- ✓ D5
- ✓ D9

3. Calcula los siguientes percentiles:

- ✓ P20
- ✓ P50
- ✓ P75

4. Interpreta los resultados: a) ¿Qué porcentaje de estudiantes usa el celular menos de 90 minutos al día? b) ¿Qué valor representa el límite superior del 25% de los estudiantes con mayor uso? c) Si un estudiante usa el celular 125 minutos al día, ¿en qué posición se encuentra dentro del grupo?

Caso 2: Consumo diario de agua en litros

Contexto: Hábitos de consumo en el hogar. Se midió el consumo diario de agua (en litros) de 20 hogares durante un día. Los datos están ordenados:

110, 115, 118, 120, 122, 125, 127, 130, 132, 135, 138, 140, 142, 145, 148, 150, 152, 155, 158, 160

1. Calcula:

- ✓ Q1, Q2, Q3
- ✓ D2 y D9
- ✓ P10, P50, P90

2. Interpreta: a) ¿Qué consumo representa el 90% de los hogares con menor gasto? b) ¿Qué porcentaje de hogares consume más de 150 litros al día? c) Si un hogar consume 135 litros, ¿está por encima o por debajo del promedio?
3. Reflexiona: ¿Cómo podrían usarse estas medidas para diseñar una campaña de ahorro de agua?

Caso 3: Calificaciones en una prueba de matemáticas.

Contexto: Evaluación académica. En una prueba de matemáticas aplicada a 20 estudiantes, se obtuvieron las siguientes calificaciones (ordenadas de menor a mayor):

45, 48, 50, 52, 55, 57, 58, 60, 61, 62, 63, 65, 66, 68, 70, 72, 75, 78, 80, 85

1. Calcula:

- ✓ Q1, Q2 (mediana), Q3
- ✓ D3 y D7
- ✓ P25, P50, P80

2. Interpreta: a) ¿Qué porcentaje de estudiantes obtuvo menos de 68 puntos? b) ¿Qué calificación representa el límite inferior del 25% con mejor rendimiento? c) Si un estudiante obtuvo 75, ¿en qué rango de percentil se encuentra?
3. Reflexiona: ¿Por qué puede ser más útil conocer el percentil de un estudiante que solo su nota

Caso 4 – Tema: Tiempo de traslado al centro educativo.

Contexto: Movilidad estudiantil. Se registró el tiempo (en minutos) que tardan 20 estudiantes en llegar a su centro educativo. Los datos están ordenados:

15, 18, 20, 22, 25, 27, 28, 30, 32, 33, 35, 36, 38, 40, 42, 45, 48, 50, 52, 55

1. Calcula:

- ✓ Q1, Q2, Q3
- ✓ D4 y D8
- ✓ P25, P50, P85

2. Interpreta: a) ¿Qué porcentaje de estudiantes tarda más de 45 minutos? b) ¿Qué tiempo representa el límite superior del 25% más rápidos? c) Si un estudiante tarda 38 minutos, ¿en qué decil se encuentra?
3. Reflexiona: ¿Cómo podrían usarse estas medidas para planificar rutas de transporte escolar?

Caso 5 – Tema: Gasto semanal en recargas telefónicas

Contexto: Consumo personal. Se registró cuánto gastan semanalmente 20 jóvenes en recargas telefónicas (en quetzales). Los datos están ordenados:

5, 8, 10, 12, 15, 18, 20, 22, 25, 28, 30, 32, 35, 38, 40, 42, 45, 48, 50, 55

1. Calcula:

- ✓ Q1, Q2, Q3
- ✓ D2 y D9
- ✓ P10, P50, P75

2. Interpreta: a) ¿Qué gasto representa el 75% de los jóvenes con menor consumo? b) ¿Qué porcentaje gasta más de Q40 a la semana? c) Si una persona gasta Q22, ¿está por encima o por debajo del percentil 50?
3. Reflexiona: ¿Cómo podrían estas medidas ayudar a tomar decisiones sobre promociones o planes de datos?

Práctica 4

Análisis de datos agrupados en intervalos

Objetivos de la práctica

1. Calcular la media, mediana y moda a partir de una tabla de frecuencia con intervalos de clase.
2. Determinar el rango, la varianza, la desviación estándar y el coeficiente de variación en distribuciones agrupadas.
3. Interpretar los resultados obtenidos para evaluar la representatividad y homogeneidad de los datos en contextos reales.

Marco teórico

Cuando se trabaja con grandes cantidades de datos numéricos, especialmente continuos (como tiempo, peso o ingresos), es más práctico agruparlos en intervalos. Esto permite observar patrones, concentraciones y tendencias sin perder claridad. A partir de estos intervalos, se pueden calcular medidas representativas como la media, la mediana o la moda, y construir gráficos que revelen la forma de la distribución.

Caso guía: Tiempo de espera en una clínica

Se registró el tiempo (en minutos) que 50 personas tardaron en ser atendidas en una clínica. Los datos fueron agrupados en los siguientes intervalos:

Intervalo (min)	Frecuencia
0 – 9	4
10 – 19	6
20 – 29	10
30 – 39	14
40 – 49	9
50 – 59	5
60 – 69	2

Construcción de la tabla de frecuencia completa

Intervalo	f	Marca de clase (x)	F acumulada	f (%)	x·f
0-9	4	4.5	4	8%	18.0
10-19	6	14.5	10	12%	87.0
20-29	10	24.5	20	20%	245.0
30-39	14	34.5	34	28%	483.0
40-49	9	44.5	43	18%	400.5
50-59	5	54.5	48	10%	272.5
60-69	2	64.5	50	4%	129.0
Total	50			100%	1635

Medidas de tendencia central para datos agrupados

1. Cálculo de la media

Fórmula:

$$\bar{x} = \frac{\sum (x \cdot f)}{n}$$

$$\bar{x} = \frac{1635}{50} = 32.7 \text{ minutos}$$

Interpretación: En promedio, una persona espera 32.7 minutos para ser atendida.

2. Cálculo de la mediana

Fórmula:

$$\text{Mediana} = L_i + \left(\frac{\frac{n}{2} - F_{i-1}}{f_i} \right) \cdot a$$

Donde:

- ✓ L_i = límite inferior del intervalo mediano
- ✓ F_{i-1} = frecuencia acumulada anterior
- ✓ f_i = frecuencia del intervalo mediano
- ✓ a = amplitud del intervalo

$$\frac{n}{2} = \frac{50}{2} = 25$$

El intervalo que contiene el dato 25 es 30–39

- ✓ $L_i = 29.5$
- ✓ $F_{i-1} = 20$
- ✓ $f_i = 14$
- ✓ $a = 10$

$$\text{Mediana} = 29.5 + \left(\frac{25 - 20}{14} \right) \cdot 10 = 29.5 + 3.57 = 33.07 \text{ minutos}$$

3. Cálculo de la moda

Fórmula:

$$\text{Moda} = L + \left(\frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \right) \cdot a$$

Donde:

- ✓ L = límite inferior del intervalo modal
- ✓ f_m = frecuencia modal
- ✓ f_1 = frecuencia anterior
- ✓ f_2 = frecuencia posterior
- ✓ a = amplitud

Intervalo modal: 30–39

$L=29.5,$ $f_m=14,$ $f_1=10,$ $f_2=9,$ $a=10$

$$\text{Moda} = 29.5 + \left(\frac{14 - 10}{(14 - 10) + (14 - 9)} \right) \cdot 10 = 29.5 + \left(\frac{4}{9} \right) \cdot 10 = 29.5 + 4.44 = 33.94 \text{ minutos}$$

Medidas de dispersión para datos agrupados

1. Rango

Fórmula: Rango=Límite superior del último intervalo–Límite inferior del primer intervalo

$$\text{Rango}=69.5-(-0.5)=70 \text{ minutos}$$

Interpretación: El tiempo de espera varía en un rango de 70 minutos entre el más rápido y el más lento.

2. Varianza y desviación estándar

Fórmulas:

$$\text{Varianza} = \frac{\sum f \cdot (x - \bar{x})^2}{n} \quad \text{y} \quad \text{Desviación estándar} = \sqrt{\text{Varianza}}$$

Ya tenemos: $\bar{x} = 32.7$

Marcas de clase y frecuencias

x (marca de clase)	f	$x - \bar{x}$	$(x - \bar{x})^2$	$f \cdot (x - \bar{x})^2$
4.5	4	-28.2	795.24	3180.96
14.5	6	-18.2	331.24	1987.44
24.5	10	-8.2	67.24	672.40
34.5	14	1.8	3.24	45.36
44.5	9	11.8	139.24	1253.16
54.5	5	21.8	475.24	2376.20
64.5	2	31.8	1011.24	2022.48
Total	50			11538.00

$$\text{Varianza} = \frac{11538}{50} = 230.76 \Rightarrow \text{Desviación estándar} = \sqrt{230.76} \approx 15.19 \text{ minutos}$$

Interpretación: En promedio, los tiempos de espera se desvían 15.19 minutos respecto a la media

3. Coeficiente de variación (CV)

Fórmula:

$$CV = \left(\frac{\text{Desviación estándar}}{\text{Media}} \right) \cdot 100$$

$$CV = \left(\frac{15.19}{32.7} \right) \cdot 100 \approx 46.45\%$$

Interpretación: La variabilidad relativa es moderada. Casi la mitad del valor promedio representa dispersión.

Conclusión del caso

- ✓ Aunque la media es de 32.7 minutos, la desviación estándar de 15.19 indica que hay una dispersión considerable.
- ✓ El coeficiente de variación del 46.45% sugiere que los tiempos de espera no son muy consistentes.
- ✓ La clínica podría usar esta información para reducir la variabilidad y mejorar la experiencia del paciente.

Hoja de trabajo

Caso 1: Tiempo de espera en una clínica

Se registró el tiempo (en minutos) que 50 personas tardaron en ser atendidas en una clínica. Los datos fueron agrupados así:

Intervalo (min)	Frecuencia
0 – 9	4
10 – 19	6
20 – 29	10
30 – 39	14
40 – 49	9
50 – 59	5
60 – 69	2

1. Construcción de tabla de frecuencia

Completa la siguiente tabla:

Intervalo	f	Marca de clase (x)	f·x	f acumulada	$(x - \bar{x})^2 \cdot f$
0-9					
10-19					
20-29					
30-39					
40-49					
50-59					
60-69					
Total	50				

2. Cálculo de medidas de tendencia central

a) Calcula la media; b) Calcula la mediana; c) Calcula la moda

3. Cálculo de medidas de dispersión

a) Calcula el rango; b) Calcula la varianza; c) Calcula la desviación estándar y d) Calcula el coeficiente de variación:

4. Responde con base en los resultados obtenidos:

a) ¿Cuál es el tiempo promedio de espera? ¿Qué indica este valor sobre la atención en la clínica?

b) ¿Qué tan homogéneos son los tiempos de espera? Justifica tu respuesta utilizando la desviación estándar y el coeficiente de variación.

c) ¿Qué medida central representa mejor el comportamiento general de los datos: ¿la media, la mediana o la moda? Explica por qué.

d) ¿Qué conclusiones podrías dar a la administración de la clínica sobre la eficiencia del servicio? ¿Qué recomendarías mejorar?

Caso 2: Consumo mensual de agua en hogares

Una empresa municipal de servicios recolectó datos sobre el consumo mensual de agua (en metros cúbicos) en 50 hogares. Los datos fueron agrupados en los siguientes intervalos:

Intervalo (m³)	Frecuencia
0 – 9	3
10 – 19	5
20 – 29	9
30 – 39	15
40 – 49	10
50 – 59	6
60 – 69	2

1. Construcción de tabla de frecuencia

Completa la siguiente tabla:

Intervalo	f	Marca de clase (x)	f·x	f acumulada	$(x - \bar{x})^2 \cdot f$
0-9					
10-19					
20-29					
30-39					
40-49					
50-59					
60-69					
Total	50				

2. Cálculo de medidas de tendencia central

a) Calcula la media; b) Calcula la mediana; c) Calcula la moda

3. Cálculo de medidas de dispersión

a) Rango b) Varianza c) Desviación estándar d) Coeficiente de variación

4. Análisis e interpretación

a) ¿Cuál es el consumo promedio mensual de agua? b) ¿Qué tan homogéneo es el consumo entre los hogares? c) ¿Qué medida representa mejor el comportamiento general? d) ¿Qué recomendaciones podrías dar a la empresa para mejorar la eficiencia del servicio?