



TRONCO COMÚN

LIBRO DE TEXTO ANÁLISIS



AUTORES:

NATALY PAZ

VICTOR REYES

FREDDY LLULLUNA

PAQUITA CUADROS

JUSTO ROJAS

XENIA PEDRAZA

MARCO SALINAS

ALEX LOZADA

LIBRO DE TEXTO

TRONCO COMÚN

Instituto Superior Tecnológico Consulting Group Ecuador – Esculapio (ISTCGE)

Av. 10 de Agosto e Ignacio de San María

<https://web.istcge.edu.ec/>

Rector: Mgs. Ramón Pineda

Directora Académica: Ph.D. Yemala Castillo

Tel. 0987067892

direccionacademica@istcge.edu.ec

Título original: **LIBRO DE TEXTO ANÁLISIS DE DATOS**

Primera Edición, septiembre 2024

© Autores, Nataly Alejandra Paz Salas, Víctor Manuel Reyes, Freddy Rafael LLulluna, Paquita Alejandra Cuadros, Justo Antonio Rojas, Xenia Pedraza González, Marco Salinas Copo, Alex Lozada Paccha

© Editorial ISTCGE

I.S.B.N.: 978-9942-7301-9-0

PUBLICADO DIGITALMENTE POR:

Editorial Instituto Superior Tecnológico Consulting Group Ecuador – Esculapio



DERECHOS RESERVADOS

Todos los derechos reservados. Queda prohibida, sin la autorización escrita del titular de los derechos de autor, la reproducción total o parcial de esta obra por cualquier medio, ya sea impreso, digital, electrónico, o cualquier otro formato conocido o por conocer. Cualquier uso no autorizado, incluyendo su distribución, comunicación pública, transformación o cualquier otra forma de explotación, estará sujeto a las sanciones civiles y penales establecidas en la ley vigente.

CITACIÓN:

Paz, N, et al. (2024). Libro de texto análisis de datos. Quito: ISTCGE.

Publicación arbitrada por el Instituto Superior Tecnológico Consulting Group Ecuador – Esculapio. Pares de revisión Mgs. Pamela Pasato y Lic. Elizabeth Girón.

ÍNDICE

| | |
|--|----|
| INTRODUCCIÓN | 6 |
| DATOS GENERALES DE LA CARRERA..... | 7 |
| DATOS GENERALES DE LA ASIGNATURA..... | 7 |
| ORIENTACIONES GENERALES PARA EL ESTUDIO DE LA ASIGNATURA | 9 |
| DESARROLLO DE LOS CONTENIDOS | 10 |
| REFERENCIAS BIBLIOGRÁFICAS..... | 99 |

ÍNDICE DE ILUSTRACIONES

| | |
|---|----|
| ILUSTRACIÓN 1 – QUE ES LA ESTADÍSTICA. FUENTE: ELABORACIÓN PROPIA..... | 11 |
| ILUSTRACIÓN 2 – PARA QUÉ SIRVE LA ESTADÍSTICA. FUENTE: ELABORACIÓN PROPIA..... | 12 |
| ILUSTRACIÓN 3 – DATO, INDIVIDUO, MUESTRA, MUESTREO. FUENTE: ELABORACIÓN PROPIA..... | 13 |
| ILUSTRACIÓN 4 – HERRAMIENTAS DE ANÁLISIS. FUENTE: ELABORACIÓN PROPIA..... | 18 |
| ILUSTRACIÓN 5 –ANÁLISIS DE DATOS E INTERPRETACIÓN. FUENTE: ELABORACIÓN PROPIA..... | 19 |
| ILUSTRACIÓN 6 – DIFERENCIACIÓN ENTRE LA ESTADÍSTICA DESCRIPTIVA Y LA INFERENCIAL.. FUENTE: ELABORACIÓN PROPIA..... | 22 |
| ILUSTRACIÓN 7 –ESTADÍSTICA DESCRIPTIVA. FUENTE: ELABORACIÓN PROPIA..... | 23 |
| ILUSTRACIÓN 8 –ESTADÍSTICA INFERENCIAL. FUENTE: ELABORACIÓN PROPIA..... | 24 |
| ILUSTRACIÓN 9 –ANÁLISIS DE ESTADÍSTICA INFERENCIAL. FUENTE: ELABORACIÓN PROPIA..... | 25 |
| ILUSTRACIÓN 10 –APLICACIONES EN EL CAMPO ISTCGE. FUENTE: ELABORACIÓN PROPIA..... | 25 |
| ILUSTRACIÓN 11 –APLICACIONES EN EL CAMPO EMERGENCIAS MÉDICAS FUENTE: ELABORACIÓN PROPIA..... | 26 |
| ILUSTRACIÓN 12 –CLASIFICACIONES DE LA VARIABLE- FUENTE: ELABORACIÓN PROPIA..... | 34 |
| ILUSTRACIÓN 13 –SEGÚN SU PAPEL EN LA PREGUNTA DE INVESTIGACIÓN- FUENTE: ELABORACIÓN PROPIA..... | 36 |
| ILUSTRACIÓN 14 –CLASIFICACIONES DE LA VARIABLE SEGÚN LA NATURALEZA- FUENTE: ELABORACIÓN PROPIA..... | 38 |
| ILUSTRACIÓN 15 –CLASIFICACIONES DE LA VARIABLE SEGÚN EL NÚMERO DE VALORES- FUENTE: ELABORACIÓN PROPIA..... | 38 |
| ILUSTRACIÓN 16 –CLASIFICACIONES DE LA VARIABLE SEGÚN EL NÚMERO DE VALORES POSIBLES- FUENTE: ELABORACIÓN PROPIA..... | 39 |
| ILUSTRACIÓN 17 –CLASIFICACIONES DE LA VARIABLE SEGÚN EL NÚMERO DE VALORES POSIBLES- FUENTE: ELABORACIÓN PROPIA..... | 41 |

| | |
|--|----|
| ILUSTRACIÓN 18 –DISTRIBUCIÓN DE FRECUENCIAS- FUENTE: ELABORACIÓN PROPIA..... | 46 |
| ILUSTRACIÓN 19 –DISTRIBUCIÓN DE FRECUENCIAS PARA ESTUDIO- FUENTE: ELABORACIÓN PROPIA..... | 47 |
| ILUSTRACIÓN 20 –GRÁFICO DE BARRAS- FUENTE: ELABORACIÓN PROPIA..... | 52 |
| ILUSTRACIÓN 21 –GRÁFICO DE BARRAS- FUENTE: ELABORACIÓN PROPIA..... | 53 |
| ILUSTRACIÓN 22 –GRÁFICO DE LINEAL- FUENTE: ELABORACIÓN PROPIA..... | 53 |
| ILUSTRACIÓN 23 –GRÁFICO DE LINEAL- FUENTE: ELABORACIÓN PROPIA..... | 54 |
| ILUSTRACIÓN 24 –VENTAJAS DE USO- FUENTE: ELABORACIÓN PROPIA..... | 55 |
| ILUSTRACIÓN 25 – SPSS (STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES)- FUENTE: ELABORACIÓN PROPIA..... | 56 |
| ILUSTRACIÓN 26 – VENTANA VISTA DE VARIABLES DE SPSS. FUENTE: ELABORACIÓN PROPIA..... | 57 |
| ILUSTRACIÓN 27 – VISTA DE DATOS SPSS. FUENTE: ELABORACIÓN PROPIA..... | 58 |
| ILUSTRACIÓN 28 – VENTANA DE SINTAXIS. FUENTE: ELABORACIÓN PROPIA..... | 58 |
| ILUSTRACIÓN 29 – VENTANA DE DIALOGO. FUENTE: ELABORACIÓN PROPIA..... | 59 |
| ILUSTRACIÓN 30 – IMPORTACIÓN DE DATOS. FUENTE: ELABORACIÓN PROPIA..... | 60 |
| ILUSTRACIÓN 31 – HERRAMIENTAS DE ANÁLISIS DESCRIPTIVO. FUENTE: ELABORACIÓN PROPIA..... | 61 |
| ILUSTRACIÓN 32 – SELECCIÓN DE VARIABLES. FUENTE: ELABORACIÓN PROPIA..... | 61 |
| ILUSTRACIÓN 33 – OPCIONES ADICIONALES. FUENTE: ELABORACIÓN PROPIA..... | 62 |
| ILUSTRACIÓN 34 – EJECUCIÓN DE ANÁLISIS. FUENTE: ELABORACIÓN PROPIA..... | 63 |
| ILUSTRACIÓN 35 – INTERPRETACIÓN DE RESULTADOS. FUENTE: ELABORACIÓN PROPIA..... | 63 |
| ILUSTRACIÓN 36 – MEDIDAS DE TENDENCIA CENTRAL, CONCEPTOS. FUENTE: ELABORACIÓN PROPIA..... | 64 |
| ILUSTRACIÓN 37 – CRITERIOS DE USO DE MEDIDAS DE TENDENCIA CENTRAL. FUENTE: ELABORACIÓN PROPIA..... | 65 |
| ILUSTRACIÓN 38 – MEDIA ARITMÉTICA. FUENTE: ELABORACIÓN PROPIA..... | 66 |
| ILUSTRACIÓN 39 – MEDIA ARITMÉTICA, EJEMPLOS. FUENTE: ELABORACIÓN PROPIA..... | 67 |
| ILUSTRACIÓN 40 – MEDIANA EJEMPLOS. FUENTE: ELABORACIÓN PROPIA..... | 68 |
| ILUSTRACIÓN 41 – MODA Y EJEMPLOS. FUENTE: ELABORACIÓN PROPIA..... | 71 |
| ILUSTRACIÓN 42 – EJEMPLOS DE MEDIDA DE TENDENCIA CENTRAL. FUENTE: ELABORACIÓN PROPIA..... | 72 |
| ILUSTRACIÓN 43 – TIPOS DE VARIABLES. FUENTE: ELABORACIÓN PROPIA..... | 73 |
| ILUSTRACIÓN 44 –CUARTILES. FUENTE: ELABORACIÓN PROPIA..... | 75 |
| ILUSTRACIÓN 45 –MEDIDAS DE POSICIÓN. FUENTE: ELABORACIÓN PROPIA..... | 80 |
| ILUSTRACIÓN 46 –RANGO INTERCUARTIL. FUENTE: ELABORACIÓN PROPIA..... | 82 |



| | |
|---|----|
| ILUSTRACIÓN 47 –CARACTERÍSTICAS RANGO INTERCUARTIL. FUENTE: ELABORACIÓN PROPIA | 83 |
| ILUSTRACIÓN 48 – DIFERENCIAS ENTRE LA VARIANZA Y LA DESVIACIÓN ESTÁNDAR. FUENTE: (CÁRDENAS ANTÚNEZ, 2014)..... | 85 |
| ILUSTRACIÓN 49 – EJEMPLO DE DESVIACIÓN ESTÁNDAR. FUENTE: (CÁRDENAS ANTÚNEZ, 2014)..... | 87 |
| ILUSTRACIÓN 50 – VARIANZA COMO MEDIDA. FUENTE: ELABORACIÓN PROPIA | 88 |
| ILUSTRACIÓN 51 – DEFINICIONES CLAVE DE LO QUE ES EL COEFICIENTE DE VARIACIÓN. FUENTE: (CÁRDENAS ANTÚNEZ, 2014). | 89 |
| ILUSTRACIÓN 52 –EJEMPLO DE DISTRIBUCIÓN NORMAL (HERNÁNDEZ, ET. AL. 2014) | 91 |
| ILUSTRACIÓN 53 – EJEMPLOS DE HISTOGRAMA. FUENTE: ELABORACIÓN PROPIA | 91 |

INTRODUCCIÓN

La asignatura Análisis de Datos en Investigación tiene como objetivo fundamental el desarrollo de competencias técnicas esenciales para el análisis de datos en investigaciones e informes de naturaleza cuantitativa. Este enfoque se basa en los lineamientos establecidos por los protocolos de investigación del Instituto Superior Técnico de Ciencias de la Gestión Empresarial (ISTCGE), garantizando que los estudiantes adquieran un conocimiento sólido y aplicable en el campo de la investigación. A través de esta asignatura, los estudiantes no solo aprenderán los principios y fundamentos de la investigación cuantitativa, sino que también se familiarizarán con las diversas metodologías y técnicas de análisis de datos. Esto incluye la selección de herramientas adecuadas, la interpretación de resultados y la presentación de hallazgos de manera clara y efectiva. Además, se enfatiza la importancia de alinear el análisis de datos con un reporte de investigación, asegurando que los resultados obtenidos sean relevantes y estén directamente relacionados con las preguntas de investigación planteadas.

La capacidad de analizar datos de manera crítica y sistemática es fundamental en el mundo actual, donde la toma de decisiones informadas se basa en evidencia cuantitativa. Por lo tanto, esta asignatura se convierte en una herramienta clave para preparar a los estudiantes para enfrentar desafíos en diversos ámbitos, incluyendo la investigación académica, el sector empresarial y la formulación de políticas. A lo largo de este curso, se fomentará un ambiente de aprendizaje activo y participativo, donde los estudiantes tendrán la oportunidad de aplicar sus conocimientos en situaciones prácticas, desarrollar habilidades analíticas y contribuir al avance del conocimiento en sus respectivas áreas de estudio. Una introducción es la sección inicial de la guía didáctica, diseñada para presentar al lector, la asignatura y sus propósitos, buscando captar la atención de los estudiantes.

DATOS GENERALES DE LA CARRERA

- Nombre de la carrera
- Modalidad
- Descripción de la carrera
- Titulación
- Perfil de egreso

DATOS GENERALES DE LA ASIGNATURA

- **Nivel (semestre):** Tercer Semestre
- **Objetivo:** Desarrollar de competencias técnicas en cuanto al análisis de datos en investigaciones e informes de naturaleza cuantitativa, de acuerdo a los lineamientos que se desprenden de los protocolos de investigación del ISTCGE.
- **Requisitos previos:** Enfoques y Métodos de Investigación
- **Resultados de aprendizaje:**
 - ✚ El estudiante aplica los fundamentos de la investigación cuantitativa.
 - ✚ El estudiante desarrolla el criterio técnico para diferenciar variables y definir indicadores medibles que permitan elaborar instrumentos de recolección de datos.
 - ✚ El estudiante procesa datos cuantitativos de forma descriptiva.
 - ✚ El estudiante desarrolla competencias en el uso de programas computarizados para el análisis de datos desde la perspectiva cuantitativa (EXCEL – SPSS).
 - ✚ El estudiante aplica las estrategias y pasos para la recolección de datos cuantitativos.
 - ✚ El estudiante elabora un reporte de investigación mediante el análisis de datos.

- **Bibliografía básica y complementaria:**

Básica: obligatoria

- Schettini, P. & Cortazzo, I. (2015). Análisis de datos cualitativos en la investigación social: procedimientos y herramientas para la interpretación de información cualitativa: (ed.). D - Editorial de la Universidad de La Plata. <https://elibro.net/es/lc/istcge/titulos/66451>

Complementaria:

- Borjas García, J. E. (2020). Validez y confiabilidad en la recolección y análisis de datos bajo un enfoque cualitativo. *Trascender, contabilidad y gestión*, 5(15), 79-97. Disponible en: <https://www.scielo.org.mx/pdf/tcg/v5n15/2448-6388-tracender-15-79.pdf>

- Rebière, Olivier/ Rebière, Cristina (2019). Cómo usar Google Forms para evaluar: Formularios Google como herramientas efectivas de enseñanza. Disponible en: <https://www.iberlibro.com/9781098810788/C%C3%B3mo-Google-Forms-evaluar-Formularios-1098810783/plp>

- Rivadeneira Pacheco, J. L., Barrera Argüello, M. V., & De La Hoz Suárez, A. I. (2020). Análisis general del spss y su utilidad en la estadística. *E-IDEA Journal of Business Sciences*, 2(4), 17-25. Disponible en: <https://revista.estudioidea.org/ojs/index.php/eidea/article/view/19>

ORIENTACIONES GENERALES PARA EL ESTUDIO DE LA ASIGNATURA

1. Lee y comprende la guía de la asignatura: dar indicaciones para que el estudiante se familiarice con la guía didáctica.
2. Organice su tiempo: orientar sobre la creación de un horario de estudio que se adapte a sus responsabilidades y preferencias, esto permite establecer metas semanales y distribuye el tiempo de estudio de manera equitativa entre las distintas asignaturas.
3. Identificar recursos disponibles: recomendar fuentes de información y recursos disponibles, como libros de texto, material en línea, bibliotecas, laboratorios, y cualquier otro recurso que la asignatura pueda requerir
4. Participar en clases y actividades: se recomienda asistir y participar activamente en las clases presenciales, virtuales o actividades programadas. Se sugiere provechar para hacer preguntas, discutir temas y colaborar con compañeros.
5. Consultar con el/la profesor/a: se sugiere comunicarse con su profesor para aclarar dudas o pedir ayuda.
6. Tomar apuntes y organizar la información: Tomar apuntes durante las clases y organizar la información de manera clara, esto facilitará la revisión y estudio posterior.
7. Establecer metas de estudio: Definir metas de aprendizaje específicas para cada sesión de estudio. Esto te ayudará a mantenerse enfocado y medir el progreso.
8. Utilizar técnicas de estudio efectivas: sugerir técnicas de estudio, como resúmenes, mapas conceptuales, flashcards, para encontrar la que mejor se adapte al estilo de aprendizaje de los estudiantes.
9. Revisar regularmente: se les debe recomendar no dejar la revisión para el último momento. Repasar regularmente los conceptos aprendidos para fortalecer la comprensión de la asignatura a largo plazo.
10. Colaborar con compañeros: sugerir participar en grupos de estudio cuando es posible. Discutir temas con compañeros puede proporcionar diferentes perspectivas y mejorar la comprensión.

DESARROLLO DE LOS CONTENIDOS

UNIDAD 1: FUNDAMENTOS DE ESTADÍSTICA



Estadística: definición. Objeto.

Estadística descriptiva e inferencial.

Dato estadístico.

Aplicaciones de la estadística en el campo laboral.

RESULTADO DE APRENDIZAJE:

El estudiante aplica los fundamentos de la investigación cuantitativa.

¿QUÉ ES ESO QUE LLAMAMOS ESTADÍSTICA?

La Estadística es una ciencia (Barreto-Villanueva, 2012), que se ocupa de la recolección, análisis, interpretación, de la presentación y resumen de los datos, de manera que se facilite la toma de decisiones en un determinado campo del saber (Escudero, 2006). El propósito principal justamente se centra en extraer información relevante de los datos crudos para la toma de decisiones relacionada, por ejemplo, con la descripción de las características de sujetos, tratamientos, intervenciones, factores sociodemográficos; hasta llegar a las relaciones, comparaciones, la identificación de las relaciones causa-efecto, y la evolución en el tiempo de determinados fenómenos. Si se toma una ruta de investigación cualitativa, la Estadística facilita la toma de decisiones con base en la evidencia fáctica y empírica, es decir, en los hechos y los datos. Para esto, la estadística se basa en métodos matemáticos, que van desde la aritmética, hasta modelos probabilísticos que permiten inferir conclusiones sobre los fenómenos sujetos a observación.

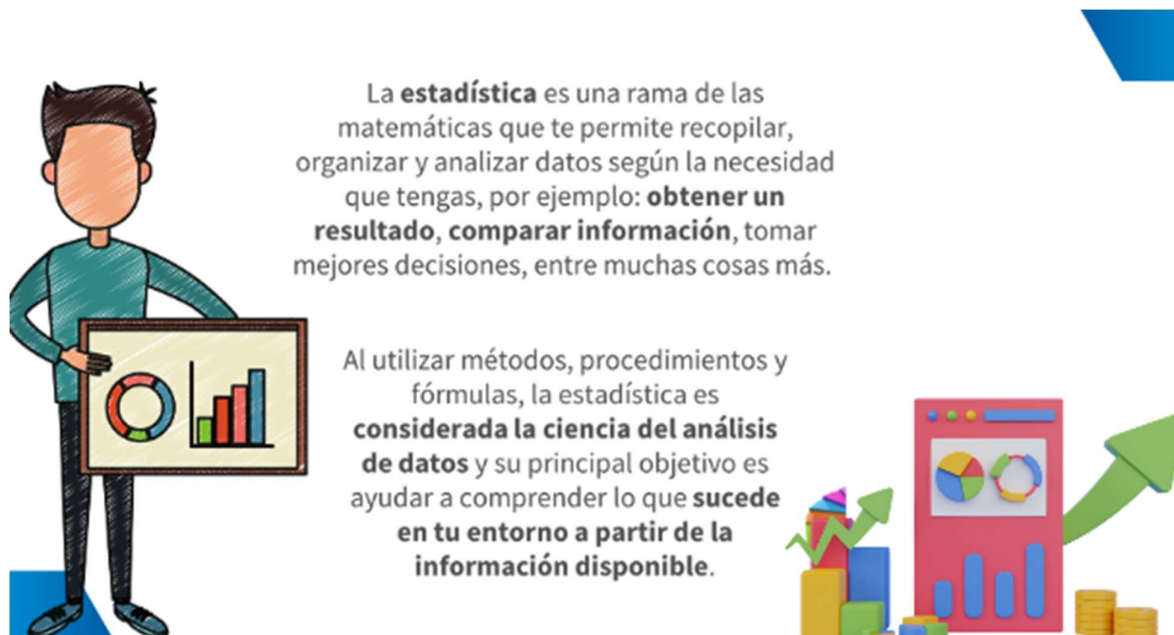


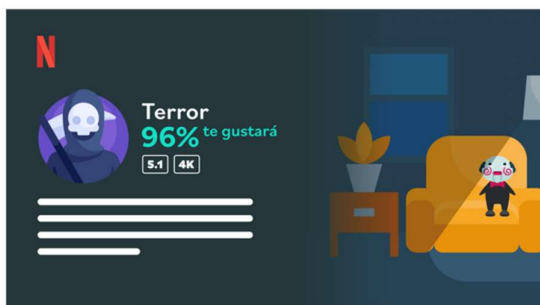
Ilustración 1 – Que es la estadística. Fuente: Elaboración propia.

Hoy en día, no es necesario que Usted como estudiante sea un matemático o estadístico titulado para utilizar la Estadística. Ese es un enfoque obsoleto de principios del siglo XX. Por un lado, la Estadística se ha posicionado como piedra angular del lenguaje científico, y la misma se ha repositionado como un eje transversal de todo lo que pretende ser científico, al menos desde la postura cuantitativa. Desde las Ciencias Naturales y la Medicina tenemos demandas sistemáticas sobre gestar conclusiones “basadas en la evidencia y los números”. Por otra parte, el desarrollo de las tecnologías de información ha favorecido el desarrollo de softwares con interfaces cada vez más amigables, que permiten al neófito a adentrarse en la Estadística de una manera amigable, aprovechando las bondades tanto de la informática como de la Inteligencia Artificial (IA).

¿Para qué sirve la estadística?

La estadística sirve para **conocer comportamientos**, aspectos de la industria y tendencias en el mercado.

Ese es el caso de **Netflix**, la plataforma de películas y series que **reúne datos de sus usuarios** para saber cuál es el contenido que **más les gusta**. Así, pueden decidir cuál será la película del momento o qué **serie lanzarán el próximo mes**.



Google también procesa **millones de búsquedas a diario** y las analiza para **mejorar la experiencia de sus usuarios**



Además, la estadística se utiliza en **investigaciones médicas** para comprobar la eficacia de un medicamento, en la economía para saber cuáles **productos son los más consumidos**.

Ilustración 2 – Para qué sirve la estadística. Fuente: Elaboración propia.

OBJETO DE ESTUDIO DE LA ESTADÍSTICA

El objeto de la estadística se corresponde con el análisis de datos que provienen de observaciones y de la ejecución de experimentos, así como de la aplicación de cuestionarios medidos en distintos tipos de escalas que pueden medir la presencia o ausencia de atributos o apreciaciones de percepciones generalmente medidas con escalar ordinales (Listas de cotejo, escalas tipo Likert, entre otras).

El abordaje de este objeto requiere como:

manejar conceptos básicos

- **Población:** Conjunto total de elementos.
- **Muestra:** Subgrupo representativo de la población.
- **Variables:** Características o atributos medidos en los elementos.



Ilustración 3 – Dato, Individuo, Muestra, Muestreo. Fuente: Elaboración propia.

EL DATO ESTADÍSTICO

Los datos son expresiones cuantitativas o potencialmente cuantificables. Para ello se emplean procesos de codificación (transformación de un texto en dato numérico), o simplemente se obtienen métricas que surgen de aparatos de observación. Dependiendo del enfoque asumido, los datos pueden ser clasificados con base en distintos criterios. Según el método de recopilación pueden ser:

Datos primarios: obtenidos directamente de la realidad mediante observaciones, encuestas o con el uso de aparatos de observación (termómetros, aparatos meteorológicos, medidores automotrices, multímetros, entre otros).

Datos secundarios: son datos obtenidos de fuentes secundarias como historias clínicas, reportes financieros o contables, informes técnicos, bases de datos oficiales, entre otros.

¿Qué es un dato?

Un dato es la representación de una variable que puede ser cuantitativa o cualitativa que indica un valor que se le asigna a las cosas y se representa a través de una secuencia de símbolos, números o letras.

Base de datos

Las bases de datos están formadas por un conjunto de datos clasificados según un criterio y almacenados en un soporte (digital o no) con el fin de ser consultados y acceder a ellos de manera simple y rápida. Los datos pueden ser generados de forma automática y acumulativa con programas informáticos o ingresados de forma manual

Otra clasificación posible es según el nivel de medición de la variable. Allí se puede precisar datos categóricos [(Escala nominal: sexo, tipo de suelo, la raza de los animales, el tipo de accidentes, lugar de nacimiento y la religión / ordinal: nivel de satisfacción, grado de estudio, nivel de ingreso económico, entre otros)] y datos cuantitativos propiamente dichos (Escala de razón: edad, altura, masa corporal; y Escala de intervalo: ratios de rentabilidad financiera, temperatura en grados celsius). Las precisiones sobre estos aspectos se desarrollarán en la unidad siguiente. Es un tema relevante porque este aspecto condiciona las pruebas a utilizar.

EL MÉTODO ESTADÍSTICO EN LA INVESTIGACIÓN CUANTITATIVA

Exploraremos las diferentes etapas del método estadístico, siendo esto un proceso fundamental en la investigación y el análisis de datos. El método estadístico consta de varias etapas, desde la recolección de datos hasta la interpretación de los resultados. Cada etapa desempeña un papel fundamental en el proceso de toma de decisiones, la obtención de conclusiones significativas y la formulación de recomendaciones. Al comprender y aplicar estas etapas de manera efectiva, los investigadores pueden obtener información valiosa que contribuya al avance del conocimiento en sus respectivos campos.

Por razones prácticas vinculadas a la naturaleza del curso “Análisis de datos”, reduciremos el proceso a cuatro etapas:

1. **Recolección de Datos:** recopilación de información relevante.
2. **Análisis de Datos:** aplicación de métodos matemáticos.
3. **Interpretación:** discusión de los resultados, conclusiones y recomendaciones.
4. **Reporte de los resultados:** corresponde al informe final.

1. Recolección de Datos

Esta primera etapa del método implica la recopilación de los datos, cuyo procesamiento se traducirá en información relevante para el estudio. Pero, ¿cómo se recolecta información? Básicamente se realiza en función del manejo de tres técnicas de recolección, con los debidos instrumentos:

| Técnica de recolección | Instrumento |
|--|--|
| Observación (aplicadas a sujetos, con base en casos clínicos, fenómenos sociales, conductas, entre otros). | Guía de observación / Ficha de observación y registro |
| Encuesta (aplicadas a sujetos). | Cuestionarios estructurados: Listas de cotejo: respuestas dicotómicas (SI / NO). Respuestas del tipo nominal-politómica, donde el consultado tiene más de dos opciones sin orden de |

| | |
|---|---|
| | <p>prevalencia (Lugar de residencia: Guayas, Pichincha, El Oro, Santo Domingo...)</p> <p>Escalas tipo Likert: con opciones de respuesta ordinal-politómica, donde el sujeto tiene más de dos opciones de respuesta.</p> |
| <p>Análisis de documentos (análisis de datos existentes, ya recolectados)</p> | <p>Ficha de observación y registro, donde se observan documentos como:</p> <p>Historias clínicas.</p> <p>Reportes de siniestros vehiculares.</p> <p>Evaluaciones de revisión vehicular.</p> <p>Sentencias judiciales.</p> <p>Reportes financieros.</p> <p>Memorias de sostenibilidad.</p> <p>Informes contables, entre otros.</p> |

Dependiendo del objetivo y la naturaleza del dato a obtener, se selecciona una u otra técnica, pudiendo también existir una utilización conjunta de varias de ellas. Se aclara que estos procesos de recolección deben cumplir con condiciones de validez y confiabilidad, para así evitar sesgos y problemas de objetividad en la obtención de resultados y la posterior toma de decisiones.

2. Análisis de Datos

Una vez recolectados los datos, la etapa siguiente corresponde al análisis de los datos, que se centra en la identificación de características comunes (patrones), tendencias centrales (media, moda, mediana), homogeneidad de los datos (variabilidad o dispersión), relaciones significativas en los datos (asociaciones o correlaciones), comparaciones (identificación de diferencias entre dos o más grupos) y la aplicación de técnicas que identifican la causalidad, permiten reducir en dimensiones o componentes todos los datos, y realizar predicciones. También es posible identificar patrones en series de tiempo. A nivel de educación de tercer nivel, los estudios estadísticos generalmente se restringen a un nivel descriptivo, alcanzándose hasta la identificación de relaciones y la precisión de diferencias entre grupos. Los demás métodos tienden a ser trabajados a nivel de Maestría y Doctorado. Estos procesos analíticos incluyen el uso de softwares especializados que permiten facilitar la complejidad de los cálculos y generar visualizaciones y gráficos que ayudan a comprender mejor los datos. Entre ellos tenemos:

| Software | Tipo de Software | Link |
|-----------------|----------------------|--|
| IBM SPSS | Software propietario | https://www.ibm.com/es-es/products/spss-statistics |
| MatLab | Software propietario | https://la.mathworks.com/products/matlab-online.html |
| Minitab | Software propietario | https://www.minitab.com/es-mx/ |
| Jamovi | Software libre | https://www.jamovi.org/ |
| Jasp | Software libre | https://jasp-stats.org/ |
| RStudio Desktop | Software libre | Instalar R: http://cran.r-project.org/bin/windows/base/ Instalar Rstudio: http://www.rstudio.com/products/rstudio/download/ |

También es posible utilizar plataformas como PowerBI y Tableau, que resultan muy útiles para el resumen de datos y la visualización de los mismos de una manera dinámica. Se aclara que el uso de RStudio Desktop requiere de competencias en programación.





Ilustración 4 – Herramientas de análisis. Fuente: Elaboración propia.

3. Interpretación

Después del análisis de los datos, se llega a la etapa de interpretación de los mismos, donde se extraen los aspectos más relevantes que conducen a establecer conclusiones y a formular recomendaciones, fundadas en los resultados obtenidos. Se centra la etapa en responder a preguntas como: qué me dicen los datos, cómo se caracteriza la muestra, qué atributos resaltan sobre los demás, cómo se relacionan los datos y variables, qué diferencia existe entre los grupos, entre otras. Esta etapa es fundamental para comprender los hallazgos en función del contexto del estudio, y se considera esto crucial para proceder a la toma de decisiones.

4. Reporte de los resultados: el reporte es un aspecto crucial para la difusión de la intencionalidad, la metodología, los resultados, la discusión y las conclusiones y recomendaciones que se desprenden del trabajo. Por razones didácticas ampliaremos esta sección seguidamente. El informe de datos o reporte de resultados, sirve para la comunicación de la información cuantitativa y sus implicaciones prácticas y decisorias. Para garantizar la claridad y la eficacia, dicho informe debe abarcar varios elementos esenciales, resumidos en la estructura: RE-I-M-R-D-CR-R.



Ilustración 5 –Análisis de datos e interpretación. Fuente: Elaboración propia.

RE-I-M-R-D-CR-R.

Resumen Ejecutivo (RE)

En primer lugar, un Resumen Ejecutivo es vital. Esta sección proporciona una visión general concisa de los principales hallazgos y conclusiones del informe, lo que permite a los lectores comprender rápidamente las ideas principales. No supera las 300 palabras.

Introducción (I)

Consiste en el preámbulo del estudio, donde se precisa el problema y la pregunta de investigación, la intencionalidad (los objetivos) y la importancia del trabajo.

Metodología (M)

Acá se describen las técnicas de recopilación de datos y los instrumentos, la selección de la muestra y los métodos analíticos empleados (pruebas estadísticas), lo que garantiza la transparencia de los datos y la reproducibilidad de los métodos. Esto ayuda a los lectores a evaluar la validez y fiabilidad de los hallazgos.

Resultados (R)

La sección de resultados, es donde se resume la presentación de los datos, lo que generalmente se presenta en forma de tablas y gráficos, que inicialmente reflejan las estadísticas descriptivas (frecuencias, porcentajes, promedios, medianas, modas, desviaciones, percentiles, rangos, intervalos de confianza, entre otros).

Discusión (D)

Se analizan acá las implicaciones de los hallazgos, comparándolos con la literatura existente y los antecedentes; destacando las bondades y las limitaciones inherentes del estudio. Tal reflexión fomenta una comprensión profunda de los resultados y su significado más amplio.

Conclusiones y recomendaciones (CR) Por último, se elabora una conclusión, donde se resumen los puntos clave y se proponen recomendaciones basadas en los resultados.

Un informe de datos estadísticos bien estructurado refleja estos elementos, lo que da garantía sobre la información presentada, de manera que esta sea accesible, significativa y procesable para el público objetivo.

Referencias: Se presentan acá las referencias bibliográficas en el formato solicitado por la carrera (APA, Vancouver, entre otros).



ESTADÍSTICA DESCRIPTIVA E INFERENCIAL

Existen dos grandes ramas de la Estadística: la Estadística descriptiva y la Estadística inferencial.

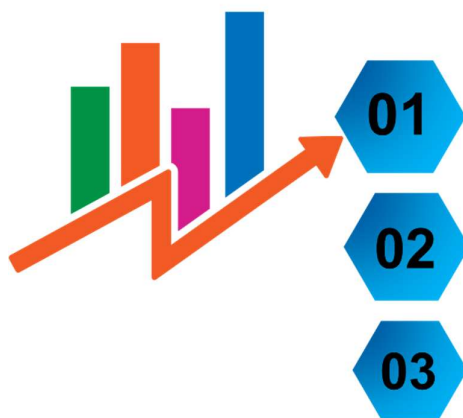
La primera se centra en describir y resumir los datos, solo basándose en la información que surge de una muestra o grupo de estudio. Para ellos utiliza medidas como frecuencias y porcentajes, medidas de tendencia central (media, moda y mediana) y medidas de dispersión o variabilidad (varianza, rangos, desviación estándar). La pregunta pertinente es: ¿cuáles son las características de la población (N) o muestra (n) observada?

ESTADÍSTICA DESCRIPTIVA

La estadística descriptiva se centra en resumir y describir las características de un conjunto de datos. Utiliza medidas como la media, la mediana y la moda para proporcionar una visión general del comportamiento de los datos. También incluye la variabilidad a través de medidas como el rango, la desviación estándar y la varianza. Además, emplea representaciones gráficas, como histogramas y gráficos de barras, que ayudan a visualizar la distribución de los datos. Su objetivo principal es ofrecer una comprensión clara y concisa de los datos recolectados, sin realizar inferencias más allá de ellos.

ESTADÍSTICA DESCRIPTIVA

La **estadística descriptiva** es el conjunto de métodos estadísticos que describen y/o caracterizan un grupo de datos.



- Determinar la tendencia central de una variable: promedio o media aritmética, mediana o moda.
- Determinar la variabilidad de una variable: desviación estándar, varianza, rangos.
- Determinar cómo es la distribución de una variable: histograma de frecuencias, distribución normal





Ilustración 6 – Diferenciación entre la Estadística descriptiva y la inferencial.. Fuente: Elaboración propia.

EJEMPLO

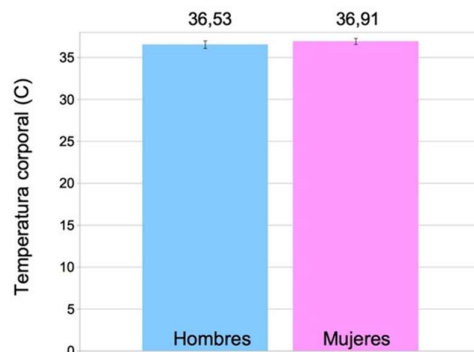
Ejemplos de estadística descriptiva

Cuando se quiere caracterizar un grupo de individuos, se usa la estadística descriptiva.

Por ejemplo, tenemos los siguientes datos de temperatura corporal en un grupo de hombres y mujeres:

| Temperatura corporal en hombres (°C) |
|--------------------------------------|
| 36,1 |
| 35,9 |
| 36,0 |
| 36,4 |
| 36,3 |
| 36,7 |
| 36,9 |
| 36,8 |

| Temperatura corporal en mujeres (°C) |
|--------------------------------------|
| 36,2 |
| 37,2 |
| 37,3 |
| 37,1 |
| 37,0 |
| 37,2 |
| 36,9 |
| 36,8 |



ANÁLISIS ESTADÍSTICO

Los hombres en este grupo tienen una temperatura promedio de 36,53 °C con una desviación estándar de 0,45;

Las mujeres en este grupo tienen una temperatura promedio de 36,91 °C, con una desviación estándar de 0,36.

ESTADÍSTICA DESCRIPTIVA

Parámetro de ubicación

- Media
- Mediana
- Valor modal
- Suma

Parámetros de dispersión

- Desviación estándar
- Varianza
- Gama

Tablas



Gráficos



Se encarga de recoger, almacenar, ordenar, realizar tablas o gráficos y calcular parámetros básicos sobre el conjunto de datos.



Es junto a la inferencia estadística una de las dos grandes ramas de la estadística.



Su propio nombre lo indica, trata de describir algo. Pero no describirlo de cualquiera forma, sino de manera cuantitativa.



Responde a preguntas como por ejemplo: ¿cuál es la altura promedio en centímetros de una población?

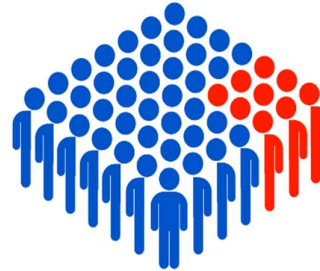
Ilustración 4 –Estadística Descriptiva. Fuente: Elaboración propia.

ESTADÍSTICA INFERENCIAL

Por otro lado, la estadística inferencial permite realizar conclusiones o generalizaciones sobre una población basándose en una muestra representativa. Utiliza técnicas como la estimación de parámetros, pruebas de hipótesis y intervalos de confianza para hacer afirmaciones sobre la población de interés. A través de métodos estadísticos, se puede determinar la probabilidad de que los resultados observados en la muestra sean representativos de la población. La estadística inferencial es esencial en investigaciones donde no es factible recolectar datos de toda la población, y su uso correcto puede guiar decisiones basadas en evidencias.

ESTADÍSTICA INFERENCIAL

La **estadística inferencial** es una rama de la estadística que se enfoca en hacer **conclusiones y generalizaciones** sobre una **población** a partir de la información obtenida de una **muestra** de la misma.



POBLACIÓN



MUESTRA

Ilustración 5 – Estadística Inferencial. Fuente: Elaboración propia.

Pero, ¿qué es inferencia? Se refiere al proceso de inferir las características de una población a partir de las métricas obtenidas para una muestra. En tal sentido, es un ejercicio de generalización, mediada por modelos de probabilidad.

Un ejemplo podría ser planteado desde la dinámica de unas elecciones gubernamentales de alcance nacional. Si se aplicase una encuesta sobre preferencias de voto de los ecuatorianos, se contaría solo con la intención de una muestra. Sin embargo, desde el punto de vista estadístico, podríamos decir que existe una tendencia al voto en favor de un determinado candidato, solo explorando los resultados de esa muestra de sujetos. En función de eso, los políticos pueden utilizar las estadísticas inferenciales para extraer conclusiones sobre el comportamiento electoral de los electores ecuatorianos basándose en esa encuesta realizada a una muestra de los electores. De esta forma, la Estadística inferencial procura dar respuestas a: ¿qué puede pasar en esa población? en lugar de ¿qué pasó solamente con la muestra?

MÉTODOS DE ANÁLISIS

Las más conocidas son las **pruebas de hipótesis** con las que se puede comprobar una diferencia de grupo, **como la prueba t, la prueba chi-cuadrado o el análisis de la varianza.**



Procedimientos sencillos

- Prueba t
- Prueba binominal
- Prueba chi-cuadrado
- Prueba U de Mann-Whitney
- Prueba de Wilcoxon
- ...

Análisis de regresión

- Regresión lineal simple
- Regresión múltiple
- Regresión logística
- ...



Análisis de correlación

- Análisis de correlación según Pearson
- Correlación de rangos de Spearman
- ...

Análisis de la varianza

- ANOVA factorial
- ANOVA bifactorial
- ANOVA con medidas repetidas
- ...

Ilustración 9 –Análisis de estadística inferencial. Fuente: Elaboración propia.

APLICACIONES DE LA ESTADÍSTICA EN EL CAMPO LABORAL.

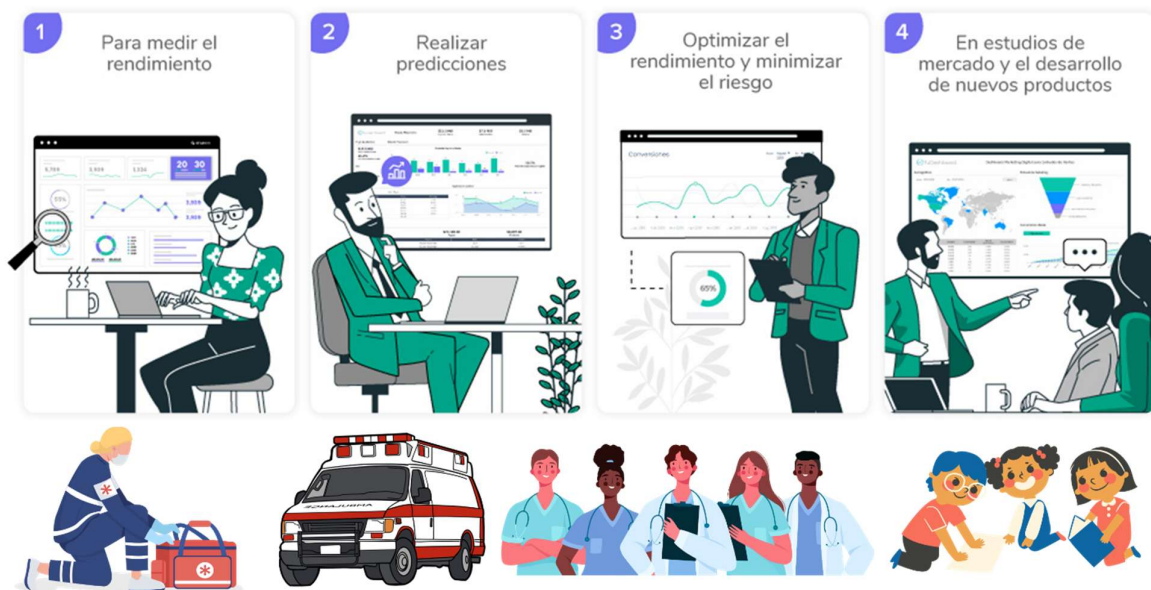


Ilustración 10 –Aplicaciones en el campo ISTCGE. Fuente: Elaboración propia.

EMERGENCIAS MÉDICAS

En el campo de las emergencias médicas, la estadística juega un papel crucial en la evaluación de la efectividad de los protocolos de atención y el tiempo de respuesta. Los profesionales utilizan datos estadísticos para analizar patrones en las llamadas de emergencia y determinar la incidencia de ciertos tipos de emergencias en diferentes regiones. Además, se emplean estudios estadísticos para evaluar la eficacia de intervenciones específicas, como el uso de desfibriladores o la administración de medicamentos, lo que contribuye a mejorar los resultados en situaciones críticas.

Ejemplo 1: Estimación del Tiempo Promedio de Respuesta

Contexto: Un hospital desea estimar el *tiempo promedio de respuesta de sus ambulancias a las llamadas de emergencia* en una gran ciudad. Debido a restricciones de tiempo y recursos, no pueden medir todos los tiempos de respuesta y deciden trabajar con una muestra.

Población y Muestra:

- **Población:** Todos los tiempos de respuesta de las ambulancias a las llamadas de emergencia en la ciudad durante un año.
- **Muestra:** Tiempos de respuesta de las ambulancias a 50 llamadas de emergencia seleccionadas aleatoriamente durante el último mes.

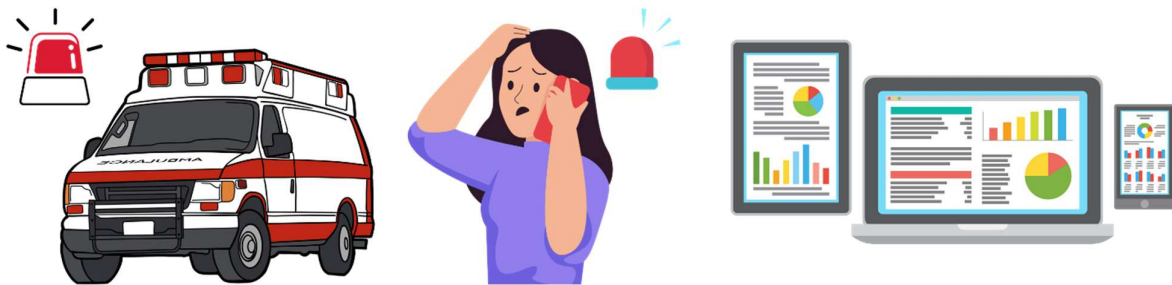


Ilustración 6 –Aplicaciones en el campo emergencias médicas Fuente: Elaboración propia.

ENFERMERÍA

La estadística es fundamental en la enfermería para la investigación clínica y la mejora de la calidad del cuidado. Los enfermeros utilizan métodos estadísticos para analizar datos sobre la salud de los pacientes, evaluar la efectividad de tratamientos y medir la satisfacción del paciente. La recopilación y análisis de datos permiten identificar tendencias en la salud de la población y desarrollar protocolos basados en evidencia, lo que mejora la atención y promueve la seguridad del paciente.

EDUCACIÓN INICIAL

En la educación inicial, la estadística se aplica para evaluar el progreso de los estudiantes y la efectividad de los programas educativos. Los educadores utilizan herramientas estadísticas para analizar resultados de pruebas y encuestas que miden el desarrollo cognitivo y emocional de los niños. Esto permite ajustar estrategias pedagógicas y programas de enseñanza, garantizando que se satisfagan las necesidades de todos los estudiantes. Además, la estadística ayuda a identificar tendencias en el aprendizaje que pueden influir en políticas educativas.

REHABILITACIÓN

En el campo de la rehabilitación, la estadística es esencial para evaluar la efectividad de diferentes terapias y tratamientos. Los profesionales analizan datos sobre el progreso de los pacientes a lo largo del tiempo, utilizando métricas como la mejora en la movilidad o la reducción del dolor. Los estudios estadísticos permiten comparar distintos enfoques terapéuticos y ajustar los planes de tratamiento para maximizar los resultados. Además, la estadística ayuda a establecer protocolos basados en evidencia para la rehabilitación, asegurando un enfoque centrado en el paciente.

ADMINISTRACIÓN

La estadística es fundamental en la administración para la toma de decisiones informadas. Los gerentes utilizan análisis estadísticos para interpretar datos financieros, realizar proyecciones de ventas y evaluar el rendimiento de los empleados. Herramientas como gráficos y análisis de regresión permiten identificar tendencias y patrones que pueden influir en la estrategia empresarial. Además, la estadística ayuda en la investigación de mercado, facilitando la segmentación de clientes y la evaluación de la efectividad de campañas publicitarias.

NATUROPATÍA

En la naturopatía, la estadística se utiliza para evaluar la efectividad de tratamientos alternativos y enfoques integrativos. Los profesionales analizan datos sobre la respuesta de los pacientes a diversas terapias naturales, ayudando a establecer la validez científica de estas prácticas. A través de estudios estadísticos, es posible identificar tendencias en la salud de los pacientes y desarrollar protocolos de tratamiento personalizados. Esto permite a los naturópatas ofrecer cuidados más fundamentados y basados en evidencia.

MARKETING

En el marketing, la estadística es crucial para la segmentación de mercados, el análisis del comportamiento del consumidor y la evaluación de campañas publicitarias. Los especialistas en marketing utilizan análisis estadísticos para interpretar datos de ventas, encuestas y estudios de mercado, lo que les permite identificar tendencias y oportunidades. La estadística también ayuda a medir el retorno de la inversión (ROI) de diferentes estrategias, optimizando así la asignación de recursos y mejorando la efectividad de las campañas. A través de la analítica predictiva, las empresas pueden anticipar cambios en el comportamiento del consumidor y adaptar sus enfoques para maximizar su impacto en el mercado.



ESTRATEGIA DE APRENDIZAJE:

Se empleará el Flipped classroom o aula invertida para promover el aprendizaje (Reyes et al., 2023, 2024). Esta estrategia consiste en Aula Invertida enfocarse en tres momentos: pre-clase, clase y postclase.

Momentos de una clase invertida

Pre-Clase:

- Se asignarán lecturas sobre Estadística y se colocarán tutoriales de instalación y uso de técnicas descriptivas con el software Jamovi.
- Se creará el foro: ¿por qué es importante la Estadística en tu futuro campo laboral? ¿Qué ventajas te genera el conocer de Estadística?

Clase:

Actividades de Inicio: Presentación de temas a abordar en clases y horas de estudio autónomo.

Actividades de desarrollo: Clase magistral y aprendizaje basado en problemas.

Rol del estudiante: rol activo y trabajando de manera grupal para estimular el aprendizaje colaborativo.

Metodología: El docente hará una exposición general del tema; luego, los estudiantes desarrollarán lo pertinente a su estudio de caso, lo presentarán por escrito y expondrán oralmente. Las actividades prácticas se centrarán en estudios de casos. Se promoverá el uso de herramientas tecnológicas como apoyo a la actividad académica.

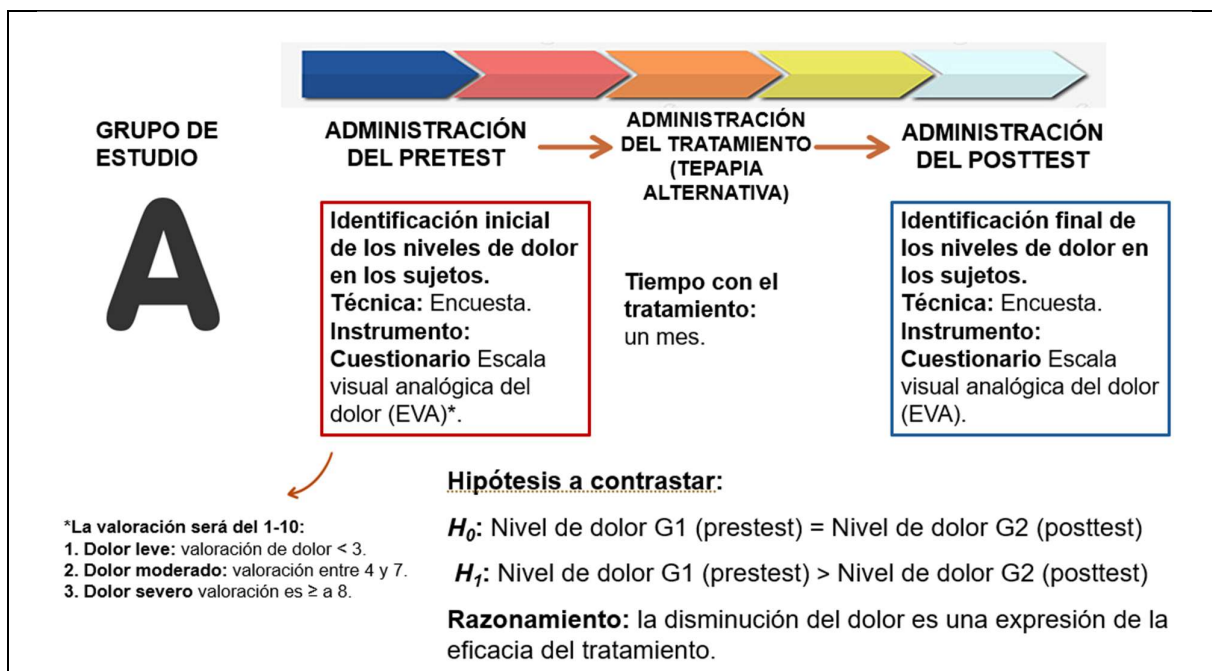
 Actividades de cierre: Análisis de casos y uso de los softwares Excel y Jamovi.

 Recursos Necesarios

 Aula virtual, Computador, base de datos en Excel, Software Jamovi.

 Aplicaciones de la Estadística en el campo laboral.

Hablar de aplicaciones en el campo laboral es abordar un campo amplio. Veamos, si Usted se firma en el campo de la Naturopatía, la Estadística le permite dar un paso al frente y asumir el marco empirista de la evidencia para respaldar los hallazgos, generalmente influidos por creencias y percepciones. Experiencias con nuestros estudiantes de ese campo del saber, nos han permitido validar la utilidad del método estadístico para conocer los efectos de terapias alternativas en sujetos. Un enfoque clásico empleado es el de asumir una estrategia basada en un diseño preexperimental (pretest vs posttest) con un solo grupo. En cada momento es necesario calcular los estadísticos descriptivos pertinentes (frecuencias y porcentajes), los que serán comparados con un test para identificar diferencias significativas. Seguidamente se ilustra la estrategia:



Asumir un diseño de este tipo evidencia un dominio de los aspectos descriptivos e inferenciales. Es un resultado de aprendizaje del curso de Análisis de Datos, mediado por el uso de softwares especializados. En los estudios de casos a abordar en esta unidad, se desarrollarán tres estrategias simples que alcanzan un nivel descriptivo y que le orientarán en su proceso de aprendizaje.

EJERCICIOS PROPUESTOS

En esta práctica, nos centraremos en comprender y comunicar datos de investigación de una manera clara y concisa. Aprenderemos a organizar datos de variables, precizando el accionar preliminar con variables nominales y ordinales, el cálculo de frecuencias y porcentajes. Se presentarán los resultados de manera resumida, utilizando tablas de frecuencias y porcentajes, gráficos de barras y gráficos tipo torta.

Campo amplio Salud:

Se desconocen los atributos de estrés académico en los estudiantes del ISTCGE. Para identificar esto se aplicará un cuestionario [Cuestionario de Estrés Académico – ECEA, elaborado por Cabanach et al., 2010 (<https://www.redalyc.org/pdf/2451/245116411005.pdf>)]. El mismo debe ser transcrito en formato virtual mediante GoogleForm (Rodríguez Romero et al., 2020), o plataforma similar. Se aplicará a tres secciones de carreras correspondientes al campo de la salud. Posteriormente, la base de datos en Excel será trasladada a Jamovi, y se obtendrá el resumen de frecuencias y porcentajes por dimensiones, y el nivel de estrés de los estudiantes. El reporte se realizará en tres páginas bajo la estructura: RE-I-M-R-D-CR-R.

Campo amplio Ciencias Sociales:

Se desconoce el nivel de Inteligencia Emocional de los estudiantes del ISTCGE. Para identificar esto se aplicará un cuestionario [Escala Rasgo de Metaconocimientos sobre Estados Emocionales (TMMS-24), Salovey et al., 1995; adaptada al español por Fernández Berrocal & Extremera Pacheco, 2005; (<https://espectroautista.info/TMMS24-es.html>)]. El mismo debe ser transcrito en formato virtual mediante GoogleForm o plataforma similar. Se aplicará a tres secciones de carreras correspondientes al campo de la salud. Posteriormente, la base de datos en Excel será trasladada a Jamovi, y se obtendrá el resumen de frecuencias y porcentajes por dimensiones: (Atención emocional: estoy en capacidad de sentir y expresar los sentimientos de forma apropiada; Claridad de sentimientos: comprendo muy bien mis estados emocionales; Reparación emocional: estoy en capacidad de regular los estados emocionales correctamente).

El reporte se realizará en tres páginas bajo la estructura: RE-I-M-R-D-CR-R.

Campo amplio Mecánica:

Se desconocen las fallas más frecuentes en vehículos Chevrolet Aveo. Se solicita visitar talleres mecánicos y llenar una ficha de registro con base en los indicadores:

- 📌 Año del vehículo: año en número (2010, 2018, etc.)
- 📌 Kilometraje: medida en km.
- 📌 Falla más frecuente: tren delantero, frenos, suspensión, sistema de inyección, entre otros.

Se recolectarán 30 impresiones sobre este modelo de vehículo. Posteriormente, se construirá la base de datos en Excel, y finalmente será trasladada a Jamovi, donde se obtendrá el resumen de frecuencias y porcentajes por indicadores. El reporte se realizará en tres páginas bajo la estructura: RE-I-M-R-D-CR-R.

EVALUACIÓN:

Los estudios de caso se evaluarán con el reporte técnico (Trabajo autónomo: 3 puntos), el cual se realizará en tres páginas bajo la estructura: RE-I-M-R-D-CR-R. De igual forma, se hará una sustentación oral del mismo (20%). Ambos aspectos se valorarán con una rúbrica.

Resumen evaluativo de la unidad:

| La evaluación de los estudiantes se realiza por el producto de la actividad de aprendizaje de los estudiantes (actividad en contacto con el docente, incluye el trabajo colaborativo), autónomo y/o de práctica) | |
|--|--------------|
| Tipo de actividad | Calificación |
| Actividad en contacto con el docente, incluye el trabajo colaborativo | 3 puntos |
| Actividad autónoma | 2 puntos |
| Actividad Práctica | 2 Puntos |
| Producto final | 3 puntos |
| Nota: El estudiante para aprobar debe obtener una nota mínima de 7/10 | |

UNIDAD 2: VARIABLES E INDICADORES ESTADÍSTICOS



Variables.

Tipos de variables según su escala.

Variables categóricas: Escala nominal y escala Ordinal

Identificación de la escala de una variable en un reporte de investigación

RESULTADO DE APRENDIZAJE

El estudiante desarrolla el criterio técnico para diferenciar variables y definir indicadores medibles que permitan elaborar instrumentos de recolección de datos.

VARIABLES

LAS VARIABLES

Todo fenómeno de la realidad en tanto proceso, está en constante transformación. Su organización y dinámica varía a través del espacio y del tiempo. En consecuencia, la variable es la variación pero sometida a estudio. (Breilh,1985, p.69).

Para identificar las variables de estudio debemos revisar la **pregunta de investigación**.

EJEMPLO:

¿Cuáles son las características de la creatividad de los niños zurdos?

VARIABLE: Creatividad, porque puede ser medida, varía en presencia, frecuencia, nivel y tipos de expresión y adopta diferentes valores o escalas.

Se llama variable estadística o carácter a cada una de las características que pueden estudiarse de la población.

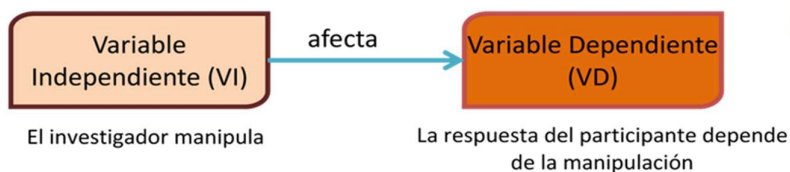
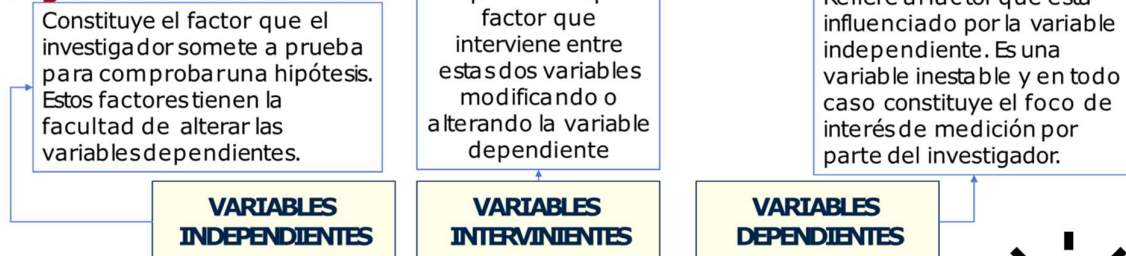
| | |
|--|--|
| Según su papel en la pregunta de investigación | Dependiente (respuesta) |
| | Independiente (explicativa o exposición) |
| Según el número de valores posibles: | Continuas |
| | Discretas |
| Según su escala de medición: | Nominal (dicotómica, politómicas) |
| | Ordinal |
| | De intervalos (continuas o de escala) |
| | De razones (continuas o de escala) |

Ilustración 7 –Clasificaciones de la variable- Fuente: Elaboración propia.

VARIABLES SEGÚN SU PAPEL EN LA PREGUNTA DE INVESTIGACIÓN

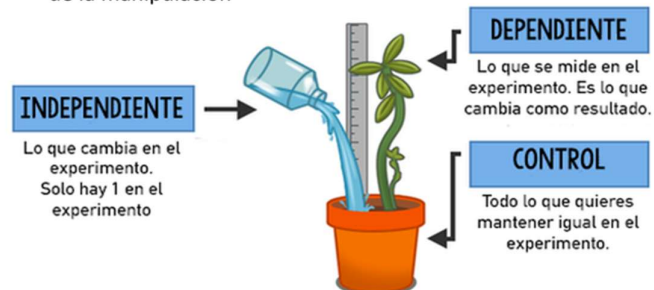
Las variables se clasifican en dependientes e independientes según su rol en una investigación. La variable independiente es aquella que se manipula o se controla para observar su efecto en otra variable, mientras que la variable dependiente es la que se mide o se evalúa como resultado de esa manipulación. Por ejemplo, en un estudio sobre el efecto de un nuevo medicamento (variable independiente) en la reducción de síntomas (variable dependiente), el investigador puede observar cómo los cambios en la dosificación del medicamento influyen en los síntomas de los pacientes.

Según su función o relación:



Una variable dependiente representa una cantidad cuyo valor depende de cómo se modifica la variable independiente.

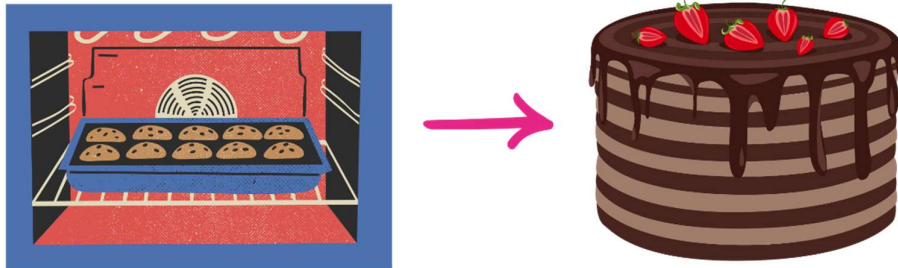
Un variable independiente es una variable que representa una cantidad que se modifica en un experimento.



Ejemplo 1:**Situación:** Hornear un pastel.

Variable independiente: Tiempo de horneado (en minutos).

Variable dependiente: Textura del pastel (por ejemplo, esponjoso o seco)

*Ilustración 8 –Según su papel en la pregunta de investigación– Fuente: Elaboración propia.*

VARIABLES SEGÚN SU NATURALEZA

Variables Cualitativas

Las variables cualitativas, también conocidas como categóricas, describen cualidades o características que no pueden medirse numéricamente. Se dividen en dos tipos: nominales y ordinales. Las **variables nominales** representan categorías sin un orden específico, como el color de los ojos, el tipo de mascota o la nacionalidad. Por otro lado, las **variables ordinales** tienen un orden jerárquico, lo que significa que se pueden clasificar en niveles de satisfacción, como bajo, medio y alto. Esta clasificación permite a los investigadores agrupar y analizar datos de manera efectiva, aunque no se pueden realizar operaciones matemáticas en estas categorías.

Variables Cuantitativas

Las variables cuantitativas, en contraste, representan cantidades y pueden medirse numéricamente. Se subdividen en **discretas** y **continuas**. Las **variables discretas** toman valores específicos y contables; por ejemplo, el número de estudiantes en un aula o la cantidad de productos vendidos en un mes. Estas variables son útiles para realizar conteos y análisis que requieren datos enteros. Las **variables continuas**, por su parte, pueden asumir cualquier valor dentro de un rango determinado, como la altura, el peso o la temperatura. Estas variables

permiten mediciones más precisas y el uso de técnicas estadísticas avanzadas que requieren un rango de valores.

Variables Mixtas

Existen también variables mixtas, que combinan aspectos cualitativos y cuantitativos, aunque son menos comunes. Un ejemplo podría ser una variable que clasifica los niveles de ingresos (bajo, medio, alto) y que al mismo tiempo permite medir la cantidad exacta de ingresos en cifras. Esta mezcla de cualidades permite una comprensión más completa de ciertos fenómenos, integrando tanto la categorización como la cuantificación en el análisis de datos. La clasificación de las variables según su naturaleza es esencial para elegir los métodos de análisis adecuados y para interpretar correctamente los resultados en cualquier investigación.

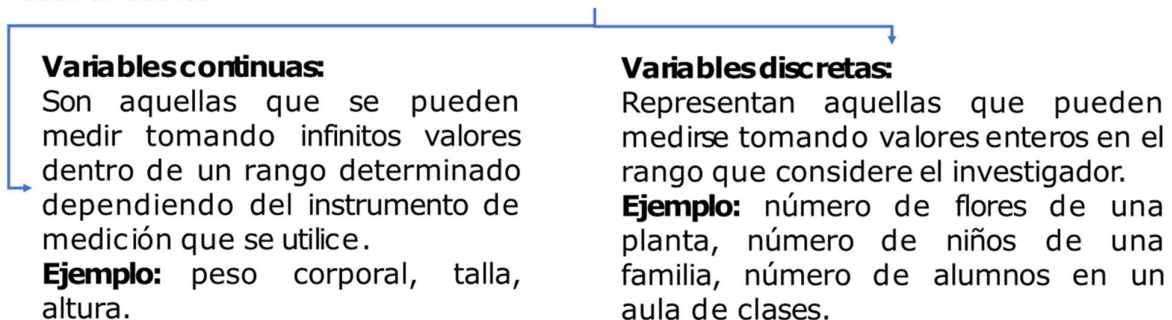
1. Según su naturaleza pueden ser:

A) Variables cualitativas. Son procesos cuyas variaciones aluden a propiedades o atributos de las variables. Para valorar estas variables se utilizan las escalas nominales y valores numéricos.



1. Según su naturaleza:

B) Variables cuantitativas. Aquellas que refieren a propiedades medibles o cuantificables.



Las variables estadísticas según su naturaleza pueden ser de dos tipos:

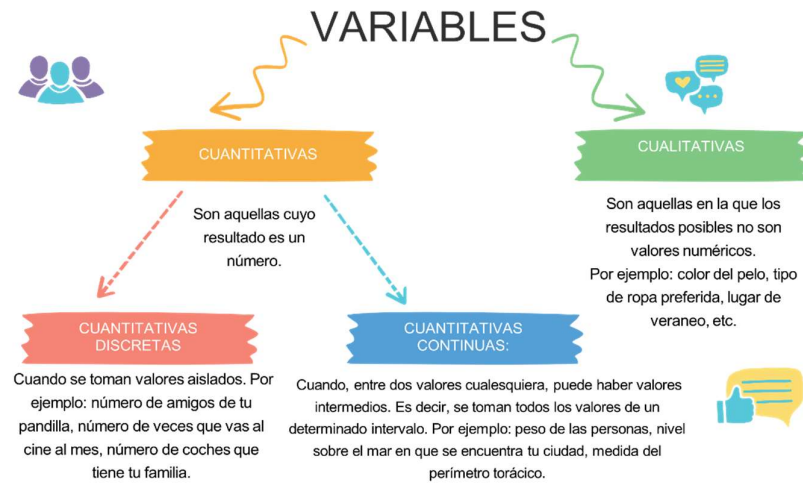


Ilustración 9 –Clasificaciones de la variable según la naturaleza- Fuente: Elaboración propia.

VARIABLES SEGÚN EL NÚMERO DE VALORES POSIBLES

Las variables también se pueden clasificar según el número de valores posibles: continuas y discretas. Las variables continuas pueden tomar cualquier valor dentro de un rango determinado, como la altura o el peso, lo que permite mediciones en fracciones y decimales. Por otro lado, las variables discretas solo pueden tomar valores específicos y contables, como el número de hijos o el número de estudiantes en un aula. Esta clasificación es importante para elegir las técnicas estadísticas adecuadas para el análisis de datos.

Variables discretas:

Son aquellas que no aceptan un valor entre dos números consecutivos. Si tienes los datos 1, 2, 3, 10, 11 y 15, entre el 1 y 2 no puede aparecer el 1.48, porque del 1 salta directamente al 2. Generalmente, las variables discretas son resultado de un conteo y no permiten los números decimales. Por ejemplo: número de pacientes, número de alumnos, número de motos por modelo.



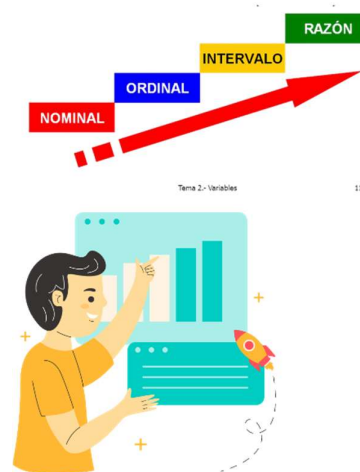
Variables continuas:

Son aquellas que pueden tomar cualquier valor entre dos intervalos o números. Por ejemplo, si necesitas escribir la estatura de un grupo de basketbolistas, seguramente, no podrás utilizar los números 1 y 2, pero si las variables 1.78, 1.65, 1.45, porque la altura suele expresarse de esa manera.

Ilustración 10 –Clasificaciones de la variable según el número de valores- Fuente: Elaboración propia.

VARIABLES SEGÚN SU ESCALA DE MEDICIÓN

- Variables nominales son aquellas que categorizan sin un orden inherente, como el género o el tipo de sangre.
- Variables ordinales, en cambio, permiten un orden o jerarquía, como las clasificaciones de satisfacción (bajo, medio, alto), pero no indican la magnitud de la diferencia entre categorías.
- Las variables de intervalo tienen un orden y una distancia constante entre los valores, como la temperatura en grados Celsius, pero carecen de unos cero absolutos.



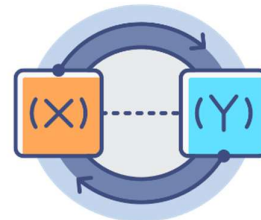
| TIPO DE VARIABLE: CUALITATIVA | | | | TIPO DE VARIABLE: CUANTITATIVA | | | |
|--|-----------------------|--|---|---|---------------------------------------|--|--------------------------------|
| NOMINAL | | ORDINAL | | INTERVALO | | RAZÓN | |
| Solo se puede etiquetar al elemento para poder identificarlo | | Además de etiquetar al elemento para poder identificarlo se puede establecer ORDEN en las categorías | | no tiene punto de partida, puede asumir valores negativos. Tiene un CERO RELATIVO Este cero no indica ausencia de la variable | | posee un punto de partida Tiene un CERO ABSOLUTO Este cero si indica ausencia de la variable | |
| Distrito de residencia | Género | Nivel de satisfacción | Grado de instrucción | Temperatura | Año | Estatura | Nro de defectos de un producto |
| Surco Barranco Miraflores | Masculino Femenino | Muy Bueno Bueno Regular Malo Muy Malo | Inicial Primaria Secundaria Superior | 10 °C 5 °C 0 °C -5 °C -10°C | 500 AC 100 AC 200 DC 1290 DC | 1,56 cm 1,58 cm 1,66 cm 1,51 cm | 0 1 2 3 4 |
| Número de celular DNI Código de AFP | | Rango militar Nivel socioeconómico Tamaño de una botella | | Año de fabricación Calificación en un examen Coeficiente de inteligencia | | Tiempo de servicio Monto de préstamo Vida útil de un USB | |

Ilustración 11 –Clasificaciones de la variable según el número de valores posibles- Fuente: Elaboración propia.

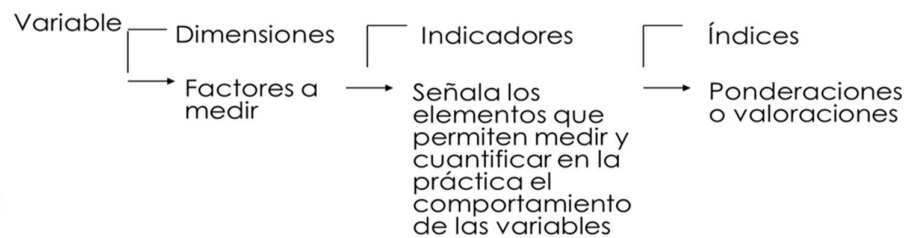
OPERACIONALIZACIÓN DE LAS VARIABLES

La operacionalización de las variables es el proceso a través del cual se transforma la variable de conceptos abstractos a términos concretos, observables y medibles; es decir, en dimensiones e indicadores.

Var(X)



En términos simples, **operacionalizar una variable** es hacerla **medible**.

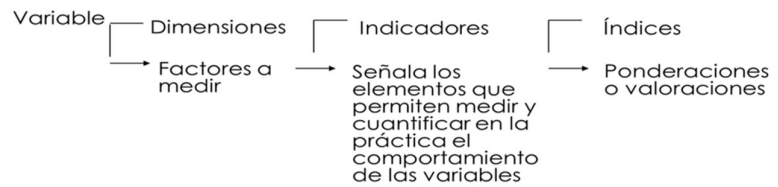


La dimensión : es el factor rasgo de la variable que debe medirse y que permite establecer indicadores; se apoyan en el marco teórico, al igual que la variable operacional. Es un elemento que resulta del análisis y/o descomposición de la misma.

| Variable | Dimensiones |
|----------------------|--------------------------|
| Clima organizacional | Involucramiento laboral, |
| | Supervisión |
| | Comunicación, |
| | Condiciones laborales |
| | Autorrealización |

Los indicadores son **aquellos valores** que permiten observar claramente la variable, **surgen de las dimensiones** y pueden ser expresadas en palabras, frases o números.

Algunos ejemplos de indicadores en palabras serían desde la variable simple, estado civil: casado, soltero, viudo, divorciado.



| Variables | Indicadores |
|------------------------------------|---|
| Entorno urbano | Banquetas, zonas peatonales, estacionamientos y espacios para personas con movilidad reducida (PMR), señalización |
| Entradas restaurantes | Accesos libres, anchura, contrastes de color. |
| Entradas hotel | Accesos |
| Movilidad interior de restaurantes | Espacio físico, pisos, puertas, señalización, mobiliario, ascensores |
| Sanitarios | Señalización, iluminación, baño adaptado, agarres, espacios amplios |

Ilustración 17—Clasificaciones de la variable según el número de valores posibles— Fuente: Elaboración propia.

ESTRATEGIA DE APRENDIZAJE

Fomentar la comprensión y aplicación de la clasificación de variables en el contexto de una investigación científica, así como la capacidad de operacionalizarlas.

Descripción del Ejercicio

Título: Análisis de Variables en un Proyecto de Investigación Real

Duración: 60-90 minutos

Pasos a seguir

1. Selección del Proyecto:

- Cada estudiante o grupo de estudiantes seleccionará un artículo de investigación científica real que aborde un tema de su interés. Este artículo debe ser accesible y relevante, preferentemente relacionado con su campo de estudio.

2. Lectura y Comprensión:

- Los estudiantes deberán leer el artículo cuidadosamente, prestando atención a la metodología y los resultados presentados. Es esencial que identifiquen las variables que se discuten en el estudio.
-

3. **Identificación y Clasificación de Variables:**

- En una hoja de trabajo proporcionada, los estudiantes deberán:
 - **Identificar las variables dependientes e independientes** mencionadas en el estudio.
 - **Clasificar las variables** como cualitativas (nominales u ordinales) o cuantitativas (discretas o continuas).
 - **Definir cada variable**, explicando su papel en la investigación y cómo se relaciona con la pregunta de investigación.

4. **Operacionalización de Variables:**

- A continuación, los estudiantes describirán cómo se operacionalizan las variables en el estudio. Deberán responder a las siguientes preguntas:
 - ¿Cómo se mide cada variable?
 - ¿Qué indicadores se utilizan para evaluar las variables?
 - ¿Qué instrumentos se emplean para la recolección de datos?

5. **Presentación:**

- Al final del ejercicio, cada grupo o estudiante presentará sus hallazgos al resto de la clase. La presentación debe incluir:
 - Una breve descripción del artículo seleccionado.
 - Las variables identificadas y su clasificación.
 - Ejemplos de la operacionalización de las variables.

6. **Discusión y Retroalimentación:**

- Después de las presentaciones, se abrirá un espacio para preguntas y comentarios, fomentando una discusión grupal sobre las diferentes investigaciones analizadas. Los estudiantes podrán reflexionar sobre cómo las variables afectan los resultados y la validez de los estudios.

RECURSOS

1. **Materiales Didácticos:**

- Presentación en PowerPoint que contenga definiciones, ejemplos y clasificaciones de variables (dependientes, independientes, cualitativas, cuantitativas, mixtas, continuas, discretas, nominales y ordinales).

- Infografías que ilustren la clasificación de variables con ejemplos concretos.
- Videos cortos que expliquen la operacionalización de variables y su importancia en la investigación.

2. Herramientas Interactivas:

- Simuladores en línea donde los estudiantes puedan practicar la identificación y clasificación de diferentes tipos de variables en escenarios de investigación.
- Foros de discusión en plataformas educativas para fomentar el intercambio de ideas sobre cómo se utilizan las variables en investigaciones reales.

3. Ejercicios Prácticos:

- Fichas de trabajo donde los estudiantes deban clasificar ejemplos de variables según su naturaleza y rol en la investigación.
- Actividades en grupo para crear preguntas de investigación y definir las variables involucradas, promoviendo el trabajo colaborativo

EVALUACIÓN

1. Criterios de Evaluación:

- Claridad y precisión en la identificación y clasificación de variables (30%).
- Calidad y profundidad de la explicación sobre la operacionalización de variables (40%).
- Efectividad en la presentación y participación en la discusión (30%).

2. Método de Evaluación:

- Se utilizará una rúbrica para evaluar cada uno de los criterios, proporcionando retroalimentación específica sobre el desempeño de cada estudiante o grupo.

UNIDAD 3: MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN

Estadística descriptiva.

Frecuencias y porcentajes

Gráficos de barras y tipo torta.

Uso de softwares estadístico. El SPSS.

Análisis descriptivo de datos en SPSS.

Medidas de tendencia central y resumen de los datos

Media, moda y mediana.

Medidas de posición. Percentiles, cuartiles, quintiles y deciles.

Medidas de forma. Asimetría y curtosis.

Medidas de dispersión o variabilidad.

Valores máximos y mínimos

El rango. Rango intercuartil.

Varianza y desviación estándar

Qué aprenderemos
en la Unidad 3



¿QUÉ ES LA ESTADÍSTICA DESCRIPTIVA?

La estadística descriptiva, como indica Fernández y Guitart (2018), es una rama de la estadística que se encarga de recolectar, organizar, presentar y describir un conjunto de datos. Su objetivo principal es resumir y sintetizar la información contenida en esos datos, facilitando su comprensión e interpretación.

¿PARA QUÉ SIRVE LA ESTADÍSTICA DESCRIPTIVA?

Imagina tener una gran cantidad de datos, como las edades de todos los estudiantes de una universidad. ¿Cómo podrías entender rápidamente cuál es la edad promedio, cuál es la edad más común o cómo se distribuyen las edades? La estadística descriptiva te proporciona las herramientas para responder a estas preguntas y muchas más.

¿CÓMO FUNCIONA LA ESTADÍSTICA DESCRIPTIVA?

La estadística descriptiva utiliza diversas técnicas y herramientas, entre las que destacan:

Tablas de frecuencia: Organizan los datos en categorías y muestran la frecuencia con la que cada categoría aparece.

GRÁFICOS: Representan los datos de forma visual, facilitando su interpretación. Algunos ejemplos son los histogramas, los diagramas de barras, los diagramas de caja y los gráficos circulares.

MEDIDAS DE TENDENCIA CENTRAL: Indican el valor "típico" de un conjunto de datos (media, mediana, moda).

MEDIDAS DE DISPERSIÓN: Miden la variabilidad de los datos (rango, varianza, desviación estándar).

TABLA DE FRECUENCIAS

En estadística descriptiva, una tabla de frecuencias es una herramienta fundamental que nos permite organizar y visualizar la distribución de los datos. Esta tabla indica cuántas veces aparece cada valor en un conjunto de datos, facilitando su análisis e interpretación. Las tablas de frecuencias se pueden usar para ordenar variables cuantitativas o cualitativas.

DISTRIBUCIÓN DE FRECUENCIAS

¿Qué es una tabla de frecuencias?

Una **tabla de frecuencias** muestra de forma ordenada un conjunto de datos estadísticos y a cada uno de ellos le asigna una **frecuencia** que, en pocas palabras, son las veces que se repite un número o dato.

* Puedes usar las tablas de frecuencias para ordenar **variables cuantitativas** o **cualitativas**.



Ilustración 12 –Distribución de frecuencias- Fuente: Elaboración propia.

TIPOS DE FRECUENCIAS

Frecuencias absolutas: son el número de veces que se repite un número en un conjunto de datos.

Frecuencia relativa: corresponde a las veces que se repite un número en un conjunto de datos respecto al total, pero se expresa en porcentajes (%).

Frecuencia relativa acumulada: es la suma de las frecuencias relativas.

TIPOS DE FRECUENCIAS



- ▶ **Frecuencias absolutas:** son el número de veces que se repite un número en un conjunto de datos.
- ▶ **Frecuencias absolutas acumuladas:** es la suma de las frecuencias absolutas.
- ▶ **Frecuencia relativa:** corresponde a las veces que se repite un número en un conjunto de datos respecto al total, pero se expresa en porcentajes (%).
- ▶ **Frecuencia relativa acumulada:** es la suma de las frecuencias relativas.

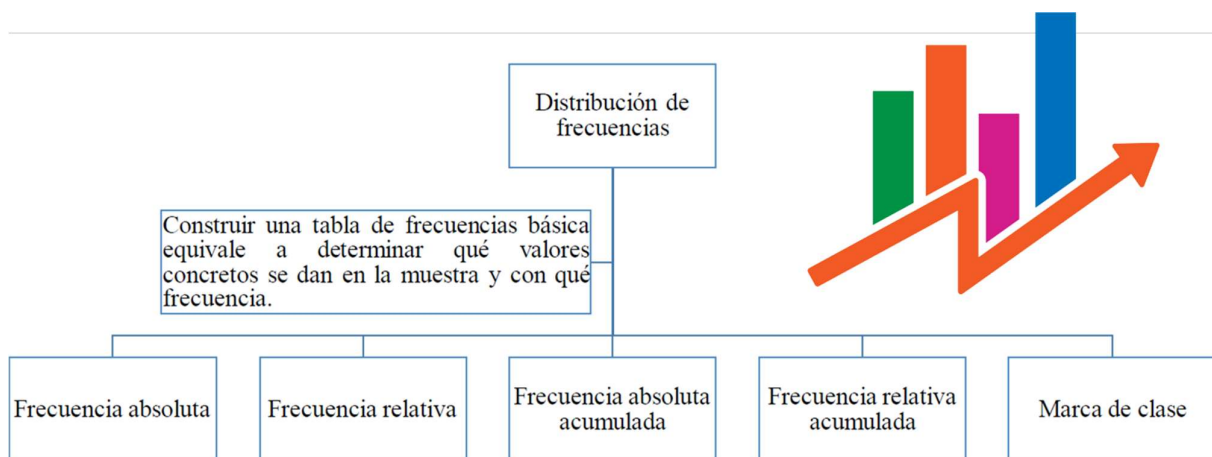
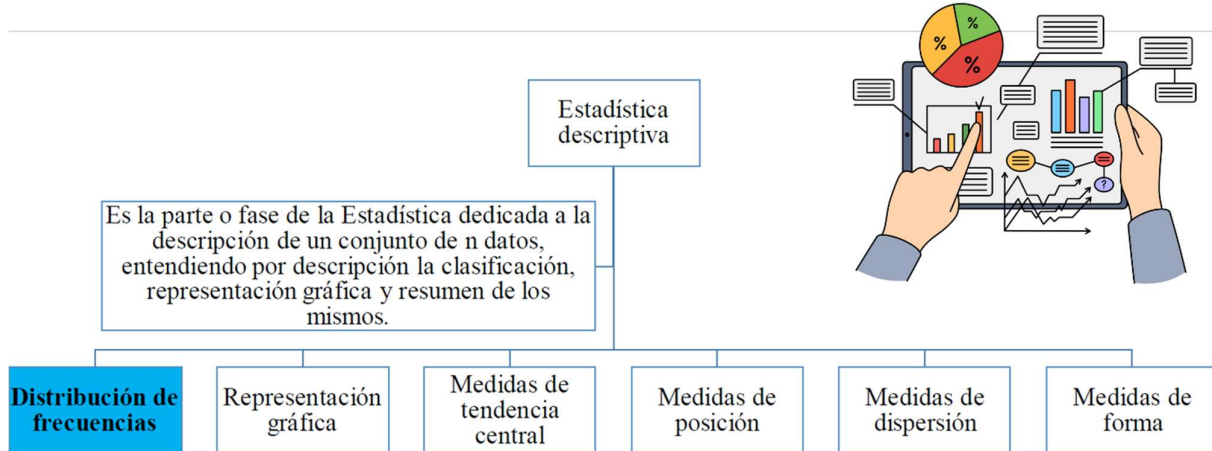


Ilustración 13 – Distribución de frecuencias para estudio- Fuente: Elaboración propia.

Ejercicio práctico:

Los siguientes datos representar las estaturas de un grupo de estudiantes se secundaria.

| Estatura de estudiantes de secundaria | | | | | | | | | |
|---------------------------------------|----|-----|-----|-----|-----|-----|-----|-----|-----|
| 15 | 16 | | | | | | | | |
| | | 161 | 160 | 160 | 172 | 162 | 160 | 172 | 151 |
| | 17 | | | | | | | | |
| 161 | | 160 | 169 | 169 | 176 | 160 | 173 | 184 | 172 |
| 16 | 17 | | | | | | | | |
| | | 153 | 167 | 167 | 175 | 166 | 173 | 169 | 178 |
| | 17 | | | | | | | | |
| 170 | | 175 | 174 | 160 | 174 | 149 | 162 | 161 | 168 |
| | 17 | | | | | | | | |
| 170 | | 156 | 159 | 154 | 156 | 160 | 166 | 170 | 169 |
| | 16 | | | | | | | | |
| 163 | | 171 | 178 | 179 | 164 | 176 | 163 | 182 | 162 |

Para realizar la tabla de frecuencias debemos realizar los siguientes pasos:

- Organizar la información de menor a mayor con ayuda software especializado.
- Establecer los intervalos
- Elaborar la tabla de frecuencias
- Analizar los resultados



| Tabla de frecuencias de datos no agrupados | | | | | | | |
|---|---------------------|-------------------------------|---------------------|-------------------------------|--|----------------------|---------------------------------|
| Categorías | Frecuencia Absoluta | Frecuencia Absoluta Acumulada | Frecuencia relativa | Frecuencia relativa acumulada | | Frecuencia relativa% | Frecuencia relativa acumulada % |
| 149 | 1 | 1 | 0,02 | 0,02 | | 2% | 2% |
| 150 | 1 | 2 | 0,02 | 0,03 | | 2% | 3% |
| 151 | 1 | 3 | 0,02 | 0,05 | | 2% | 5% |
| 153 | 1 | 4 | 0,02 | 0,07 | | 2% | 7% |
| 154 | 1 | 5 | 0,02 | 0,08 | | 2% | 8% |
| 156 | 2 | 7 | 0,03 | 0,12 | | 3% | 12% |
| 159 | 1 | 8 | 0,02 | 0,13 | | 2% | 13% |
| 160 | 9 | 17 | 0,15 | 0,28 | | 15% | 28% |
| 161 | 3 | 20 | 0,05 | 0,33 | | 5% | 33% |
| 162 | 3 | 23 | 0,05 | 0,38 | | 5% | 38% |
| 163 | 2 | 25 | 0,03 | 0,42 | | 3% | 42% |
| 164 | 1 | 26 | 0,02 | 0,43 | | 2% | 43% |
| 166 | 2 | 28 | 0,03 | 0,47 | | 3% | 47% |
| 167 | 2 | 30 | 0,03 | 0,50 | | 3% | 50% |
| 168 | 2 | 32 | 0,03 | 0,53 | | 3% | 53% |
| 169 | 4 | 36 | 0,07 | 0,60 | | 7% | 60% |
| 170 | 4 | 40 | 0,07 | 0,67 | | 7% | 67% |
| 171 | 1 | 41 | 0,02 | 0,68 | | 2% | 68% |
| 174 | 2 | 43 | 0,03 | 0,72 | | 3% | 72% |
| 173 | 7 | 50 | 0,12 | 0,83 | | 12% | 83% |
| 174 | 0 | 50 | 0,00 | 0,83 | | 0% | 83% |

| | | | | | | |
|-----|----|----|------|------|------|------|
| 175 | 2 | 52 | 0,03 | 0,87 | 3% | 87% |
| 176 | 2 | 54 | 0,03 | 0,90 | 3% | 90% |
| 178 | 2 | 56 | 0,03 | 0,93 | 3% | 93% |
| 179 | 2 | 58 | 0,03 | 0,97 | 3% | 97% |
| 182 | 1 | 59 | 0,02 | 0,98 | 2% | 98% |
| 184 | 1 | 60 | 0,02 | 1,00 | 2% | 100% |
| | 60 | | 1,00 | | 100% | |

TABLA DE DATOS AGRUPADOS

Para determinar la tabla de frecuencias por datos agrupados se deben seguir los siguientes pasos:

- Organizar la información de menor a mayor con ayuda software especializado.
- Establecer los intervalos por datos agrupados
- Elaborar la tabla de frecuencias
- Analizar los resultados
- Ejemplo de cálculo de intervalos
- Determinar tamaño de muestra
- Identificar el max y el min.
- Calcular la cantidad de intervalos
- Calcular la longitud de cada intervalo
- Armar tabla de frecuencia en función de las arcas de clase.

| Categorías | Intervalos | | | Frecuencia Absoluta | Frecuencia Absoluta Acumulada | Frecuencia relativa | Frecuencia relativa acumulada | Frecuencia relativa% | Frecuencia relativa acumulada % |
|------------|------------|---------|---------|---------------------|-------------------------------|---------------------|-------------------------------|----------------------|---------------------------------|
| | Xi | Lim Inf | Lim sup | | | | | | |
| 1 | 1 | 15 | 151,5 | 4 | 4 | 0,067 | 0,067 | 6,67% | 6,67% |
| 2 | 1 | 15 | 156,5 | 3 | 7 | 0,050 | 0,117 | 5,00% | 11,67% |
| 2 | 1 | 16 | 161,5 | 18 | 25 | 0,300 | 0,417 | 30,00% | 41,67% |
| 3 | 1 | 16 | 166,5 | 7 | 32 | 0,117 | 0,533 | 11,67% | 53,33% |
| 4 | 1 | 17 | 171,5 | 16 | 48 | 0,267 | 0,800 | 26,67% | 80,00% |
| 5 | 1 | 17 | 176,5 | 8 | 56 | 0,133 | 0,933 | 13,33% | 93,33% |
| 7 | 1 | 18 | 181,5 | 4 | 60 | 0,067 | 1,000 | 6,67% | 100,00% |

GRÁFICOS DE BARRAS

Un gráfico de barras es una representación visual de datos que utiliza barras rectangulares para mostrar la frecuencia o magnitud de diferentes categorías. Es una herramienta muy útil para comparar cantidades y hacer que los datos sean más fáciles de entender.

¿Qué es un gráfico de barras?

Un diagrama de barras es un gráfico usado para mostrar de forma resumida un grupo de datos que puede incluir variables cualitativas y cuantitativas.

En este video te enseñamos como construir este tipo de diagramas:



Ilustración 14 –Gráfico de barras- Fuente: Elaboración propia.

Características de un diagrama de barras

- Se compone de columnas o barras de diferentes alturas, estas pueden ser horizontales o verticales.
- Tiene un eje horizontal o eje x, donde se ubica una variable, por lo general, cualitativa.
- Tiene un eje vertical o eje y, donde se ponen los valores que determinan la altura de las barras. A estos números se les conoce como frecuencia.
- El ancho de las barras y el espacio entre cada una debe ser el mismo.
- Las barras también sirven para comparar valores.

¿CÓMO FUNCIONA?

Eje X: Generalmente representa las categorías o grupos que estamos comparando (por ejemplo, meses del año, marcas de productos, etc.).

Eje Y: Representa la frecuencia o cantidad de cada categoría. La altura de cada barra corresponde al valor de esa categoría.

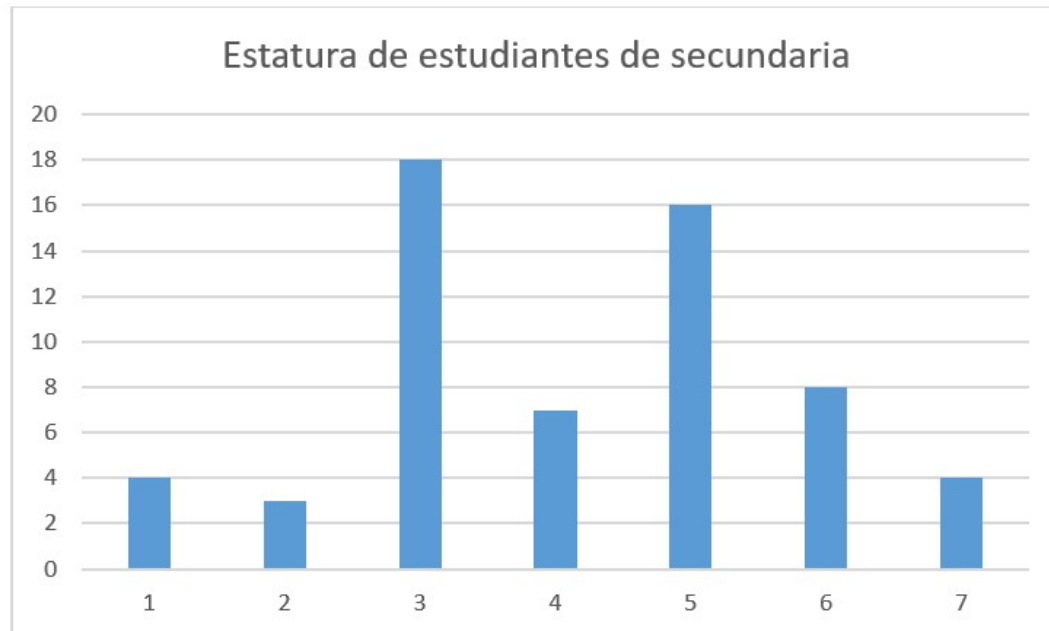


Ilustración 15 –Gráfico de Barras- Fuente: Elaboración propia.

¿QUÉ ES UN GRÁFICO DE TIPO TORTA?

Un gráfico de torta, también conocido como gráfico circular o de pastel, es una representación visual de datos que utiliza un círculo dividido en secciones. Cada sección representa una categoría de datos y su tamaño es proporcional a la frecuencia o proporción de esa categoría en relación al total.

Gráfico de torta o diagrama circular

Los diagramas circulares también son llamados gráficos de torta, tarta, pizza o 360 grados, porque se dividen en porciones.

Características de un diagrama circular:

- ▶ Estos gráficos sólo **representan** y **comparan** una variable.
- ▶ A cada cualidad de la variable le pertenece una **porción del gráfico**.
- ▶ Las cifras que se anotan sobre cada trozo se expresan en **porcentajes**, por eso son útiles cuando la información que tienes es en su mayoría numérica y son pocos datos.
- ▶ La unión de todos los sectores forman un círculo completo.
- ▶ Se usa para expresar proporcionalmente los **distintos valores de la frecuencia** de las variables.

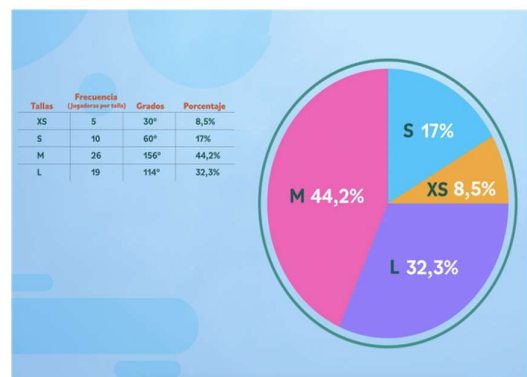


Ilustración 16 –Gráfico de Lineal- Fuente: Elaboración propia.

PARA QUÉ SE UTILIZAN?

- **Mostrar partes de un todo:** Son ideales para visualizar cómo se compone un conjunto de datos en diferentes categorías.
- **Comparar proporciones:** Permiten comparar fácilmente la proporción de cada categoría en relación al total.
- **Representar porcentajes:** Cada sección del gráfico representa un porcentaje del total.

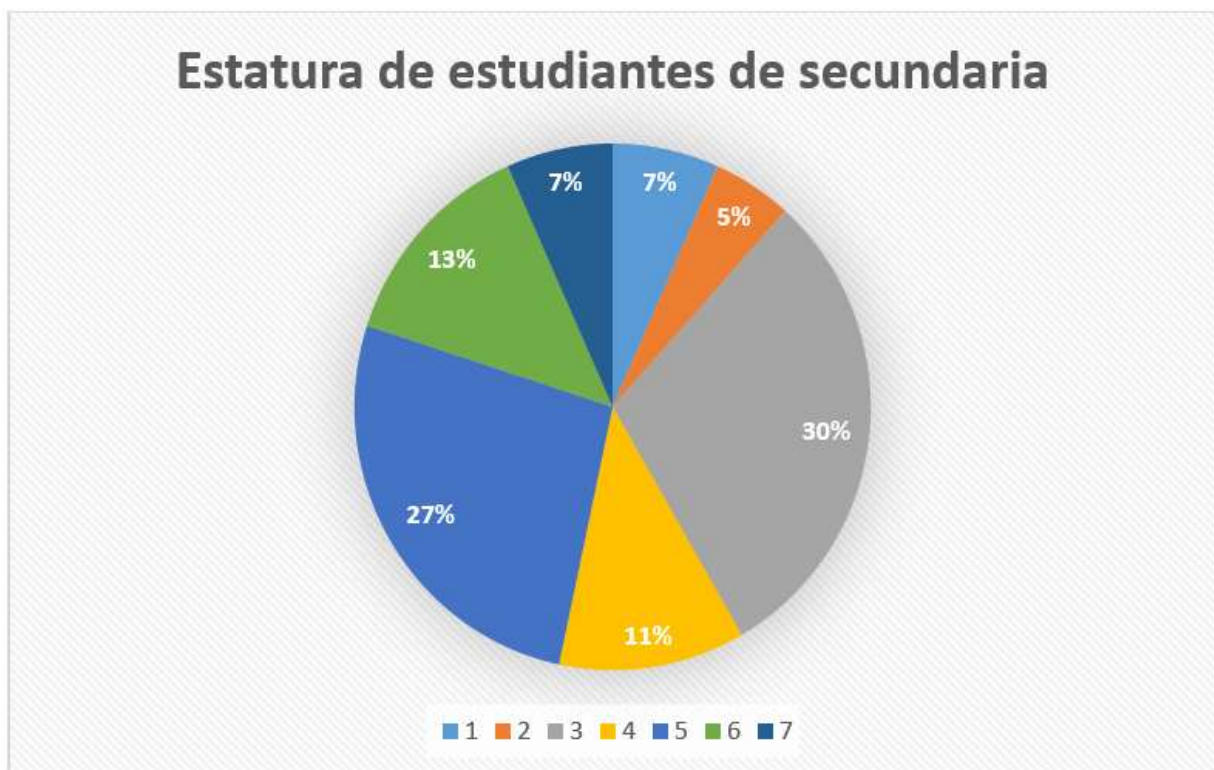


Ilustración 17 –Gráfico de Lineal- Fuente: Elaboración propia.

VENTAJAS DEL USO DE SOFTWARES ESTADÍSTICO



Ilustración 18 –Ventajas de Uso- Fuente: Elaboración propia.

1

Automatización y precisión: Los softwares estadísticos permiten automatizar cálculos complejos, reduciendo el riesgo de errores humanos y aumentando la precisión de los resultados. Herramientas como R, SPSS y SAS son ampliamente utilizadas por su capacidad de manejar grandes conjuntos de datos y realizar análisis avanzados (Fernández, 2020).

2

Visualización de datos: Estos programas facilitan la creación de gráficos y visualizaciones que ayudan a interpretar los resultados de manera más intuitiva y comprensible. Programas como Tableau y Power BI son destacados por su capacidad de transformar datos en visualizaciones dinámicas e interactivas (Martínez, 2021).

3

Análisis avanzado: Los softwares estadísticos permiten realizar análisis de regresión, análisis multivariante, minería de datos y modelos predictivos, lo cual es esencial en campos como la investigación científica, la economía y la salud (Gómez, 2019).

DESAFIOS EN EL USO DE SOFTWARES ESTADÍSTICO

Curva de aprendizaje: Algunos softwares estadísticos, como R, tienen una curva de aprendizaje pronunciada debido a su complejidad y la necesidad de conocimientos en programación. Sin embargo, la inversión en tiempo de aprendizaje se ve recompensada por la flexibilidad y el poder analítico que ofrecen (Hernández, 2019).

Costo: Mientras que algunos softwares, como R, son de código abierto y gratuitos, otros como SPSS y SAS requieren licencias pagadas, lo que puede ser una limitación para algunas instituciones y usuarios (Pérez, 2020).

Actualización constante: Los usuarios deben mantenerse actualizados con las últimas versiones y paquetes para aprovechar al máximo las capacidades de los softwares estadísticos. Esto implica un compromiso continuo con el aprendizaje y la adaptación (Alvarez, 2021).

PRINCIPALES SOFTWARES ESTADÍSTICOS



Excel es una herramienta ampliamente utilizada en el análisis de datos debido a su accesibilidad y versatilidad. Aunque no es un software estadístico especializado, ofrece una variedad de funciones que permiten realizar análisis cuantitativos básicos y avanzados.



SPSS es un software estadístico especializado que es ampliamente utilizado en las ciencias sociales, la investigación de mercado y otras disciplinas. Su interfaz intuitiva y sus poderosas capacidades analíticas lo convierten en una herramienta esencial para el análisis cuantitativo avanzado.

Ilustración 25 – SPSS (Statistical Package for the Social Sciences)- Fuente: Elaboración propia.

SPSS (Statistical Package for the Social Sciences) es uno de los softwares más utilizados en la investigación social y de la salud. Según Morales (2012), "SPSS facilita el procesamiento y análisis de datos mediante una interfaz amigable y herramientas robustas para el análisis descriptivo, inferencial y multivariante" (p. 35). Su facilidad de uso lo hace accesible tanto para investigadores novatos como experimentados. Ofrece una interfaz amigable y herramientas poderosas para análisis descriptivo y avanzado.

DESCRIPCIÓN DE LA INTERFAZ DE SPSS

En esta vista, los usuarios pueden definir y modificar las propiedades de las variables, tales como el nombre, tipo, ancho, decimales, etiquetas, valores, y medidas. Morales (2012) menciona que "la vista de variables proporciona una manera estructurada de gestionar las propiedades de cada variable, asegurando que los datos se codifiquen correctamente para el análisis" (p. 45).

Ventana Vista de variables de SPSS.

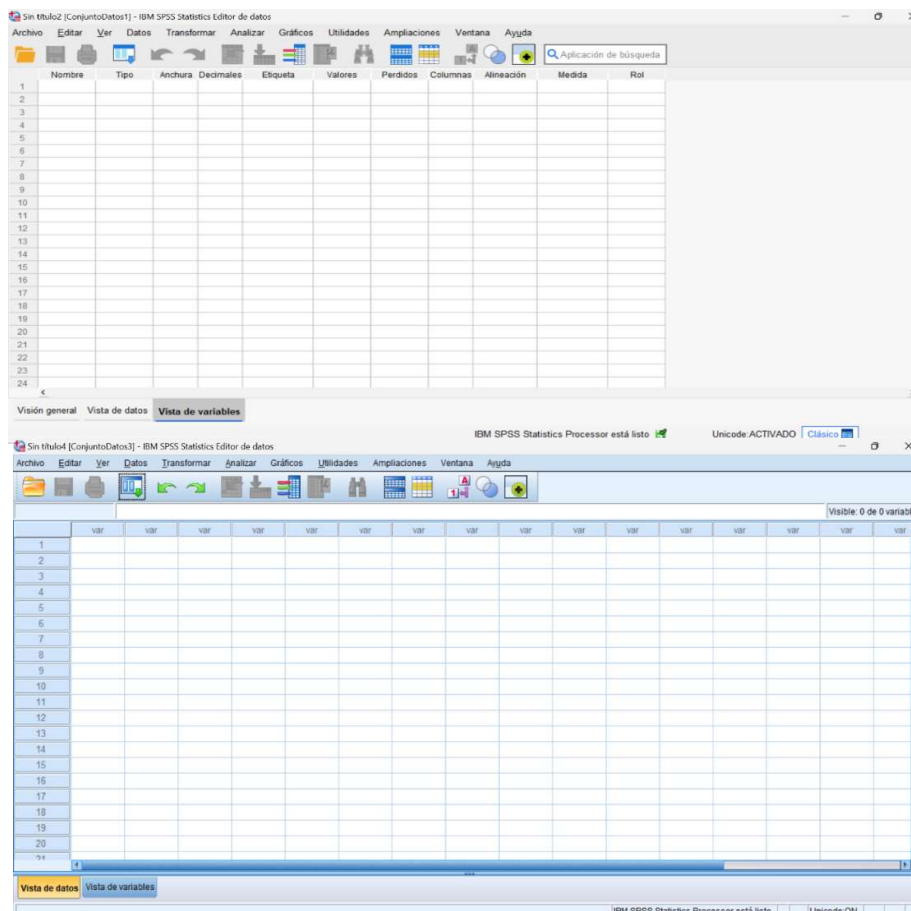


Ilustración 19 – Ventana Vista de variables de SPSS. Fuente: Elaboración propia.

Esta vista presenta los datos en un formato de tabla donde las filas representan casos individuales y las columnas representan variables. Cada celda de la tabla contiene un dato específico. Según Pardo y Ruiz (2012), "la vista de datos de SPSS permite una entrada y visualización directa de los datos, facilitando su manipulación y análisis" (p. 29).

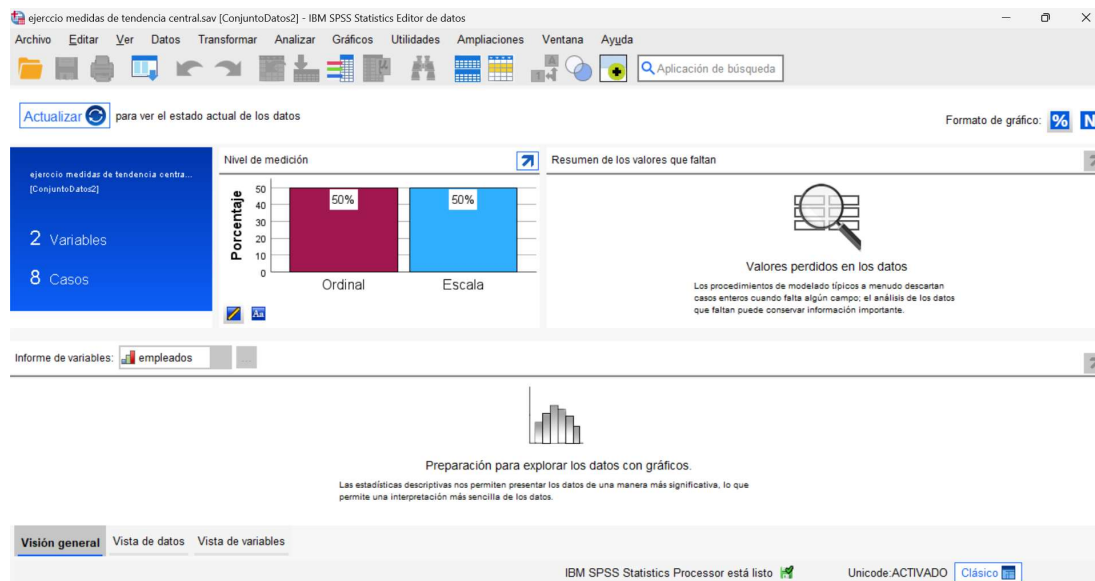


Ilustración 20 – Vista de datos SPSS. Fuente: Elaboración propia.

Esta ventana muestra los resultados de los análisis estadísticos en forma de tablas, gráficos y textos. García Martínez (2018) destaca que "el editor de salidas de SPSS organiza los resultados de manera clara y estructurada, lo que facilita la interpretación y presentación de los hallazgos" (p. 53).

Ventana de Sintaxis:

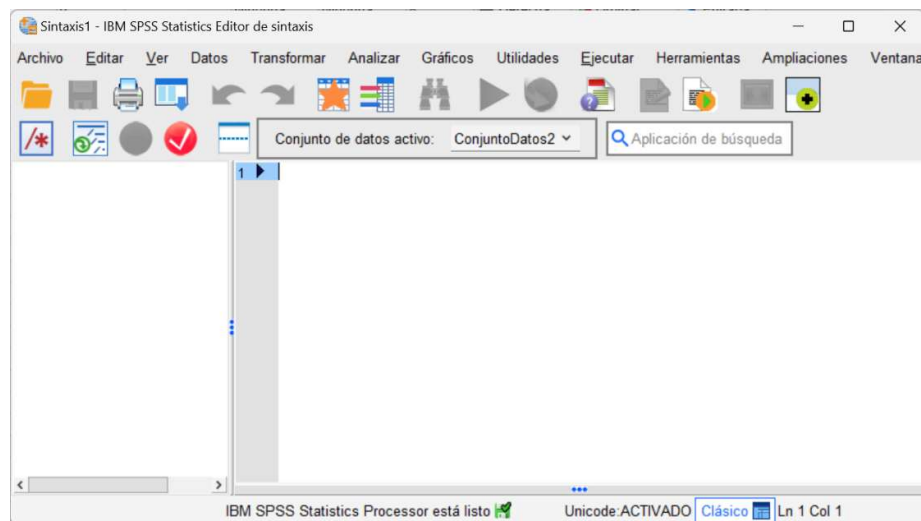


Ilustración 21 – Ventana de sintaxis. Fuente: Elaboración propia.

Permite a los usuarios escribir y ejecutar comandos en lenguaje de sintaxis de SPSS para realizar análisis y manipulaciones de datos más avanzadas. Según Bausela Herreras (2014), "el uso de sintaxis en SPSS es esencial para la replicación de análisis y la automatización de procesos complejos" (p. 64).

Menú Principal

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda

Contiene opciones como Archivo, Editar, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, y Ventana, cada uno con submenús específicos. Hernández Sampieri et al. (2014) señalan que "el menú principal de SPSS organiza todas las funciones necesarias para el manejo de datos y análisis estadísticos en un formato accesible" (p. 77).

Barra de Herramientas.



Incluye iconos para accesos rápidos a funciones comunes como abrir, guardar, imprimir, cortar, copiar, pegar, y ejecutar análisis estadísticos frecuentes.

Ventanas de Diálogo:

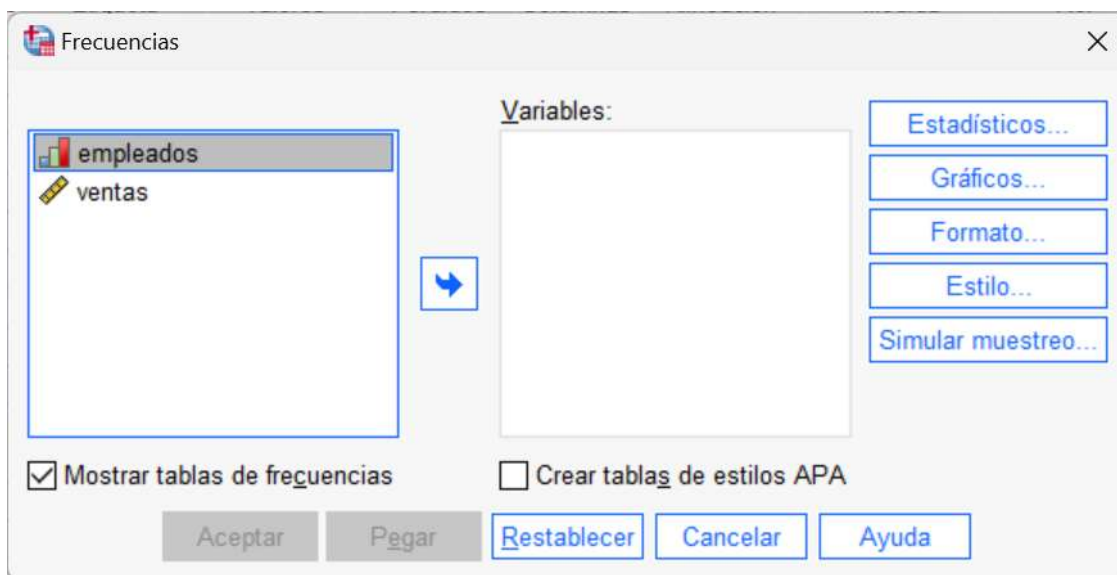


Ilustración 22 – Ventana de diálogo. Fuente: Elaboración propia.

Aparecen al seleccionar una opción de análisis o transformación de datos del menú. Estas ventanas permiten a los usuarios especificar detalles del análisis que desean realizar, seleccionando variables y definiendo parámetros. López y Gómez (2015) mencionan que "las ventanas de diálogo en SPSS guían

al usuario a través de los pasos necesarios para configurar y ejecutar análisis, lo que minimiza errores y facilita el uso del software" (p. 39).

Análisis descriptivo de datos en SPSS y Organización de los resultados.

El análisis descriptivo en SPSS versión 29 permite a los investigadores resumir y describir las características básicas de un conjunto de datos, proporcionando medidas como la media, la mediana, la moda, la desviación estándar y distribuciones de frecuencia. A continuación, se describe detalladamente cómo realizar este análisis, referenciando fuentes bibliográficas en español.

PASOS PARA REALIZAR UN ANÁLISIS DESCRIPTIVO EN SPSS VERSIÓN 29

Importación de Datos

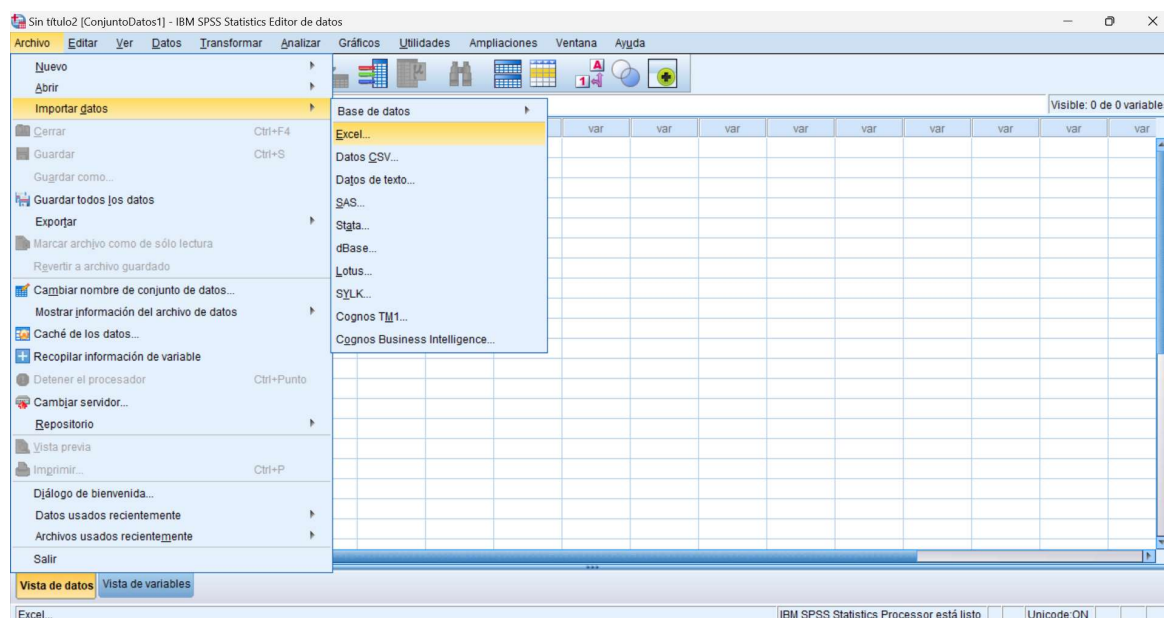


Ilustración 23 – Importación de datos. Fuente: Elaboración propia.

Los datos pueden ser ingresados manualmente en la ventana del editor de datos de SPSS o importados desde otros formatos como Excel, CSV, o bases de datos.

Según Pardo y Ruiz (2012), "la capacidad de SPSS para importar datos de diversas fuentes facilita la preparación inicial para el análisis descriptivo" (p. 31).

Acceso a las Herramientas de Análisis Descriptivo

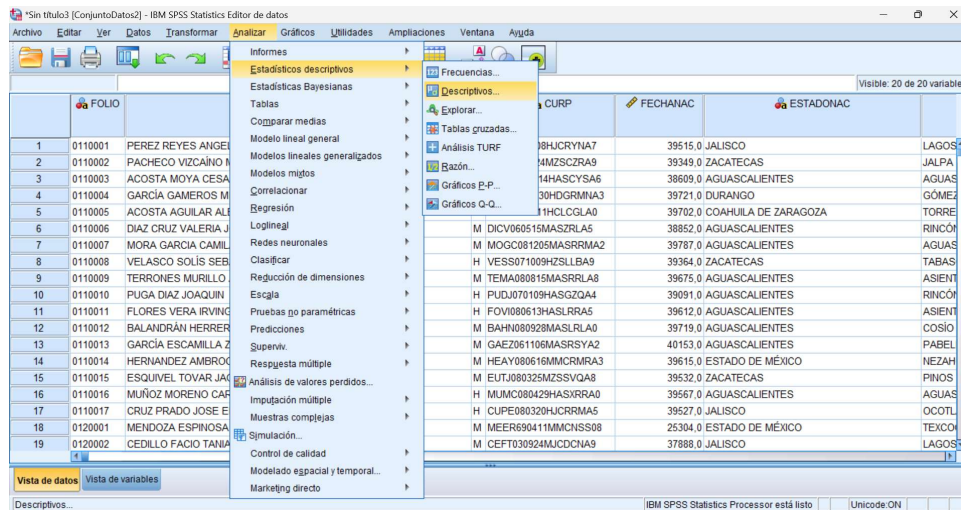


Ilustración 24 – Herramientas de análisis descriptivo. Fuente: Elaboración propia.

En la barra de menú principal, selecciona la opción "Análizar" y luego elige "Estadísticos descriptivos". Dentro de este submenú, selecciona "Descriptivos..."

Morales (2012) menciona que "el menú de estadísticos descriptivos en SPSS organiza todas las herramientas necesarias para resumir y describir los datos de manera eficiente" (p. 50).

Selección de Variables



Ilustración 25 – Selección de variables. Fuente: Elaboración propia.

En el cuadro de diálogo que aparece, selecciona las variables que deseas analizar. Puedes mover las variables desde la lista de la izquierda a la caja de la derecha utilizando los botones de flecha. García Martínez (2018) destaca que "la interfaz intuitiva de selección de variables en SPSS permite a los usuarios especificar fácilmente qué datos se incluirán en el análisis descriptivo" (p. 55).

Configuración de Opciones Adicionales

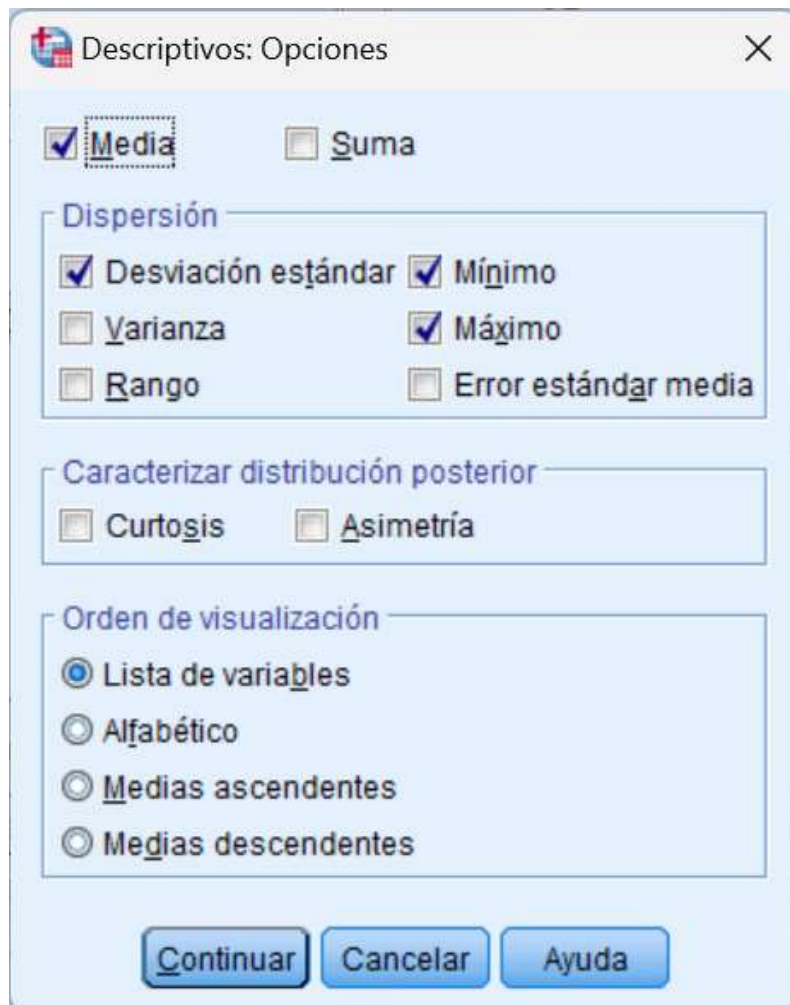


Ilustración 26 – Opciones adicionales. Fuente: Elaboración propia.

Botón "Opciones...": Al hacer clic en este botón, puedes seleccionar las estadísticas específicas que deseas calcular, tales como la media, mediana, moda, desviación estándar, varianza, mínimo, máximo y rango.

Bausela Herreras (2014) señala que "SPSS ofrece una amplia gama de opciones estadísticas que permiten personalizar el análisis descriptivo según las necesidades específicas de la investigación" (p. 67).

Ejecución del Análisis



Ilustración 27 – Ejecución de análisis. Fuente: Elaboración propia.

Una vez seleccionadas las variables y opciones deseadas, haz clic en "Aceptar". SPSS procesará los datos y generará un informe en la ventana de salidas.

Hernández Sampieri et al. (2014) indican que "la ejecución del análisis descriptivo en SPSS es rápida y eficiente, proporcionando resultados inmediatos y fácilmente interpretables" (p. 80).

Interpretación de Resultados

NEW FILE.
DATASET NAME ConjuntoDatos3 WINDOW=FRONT.
DATASET ACTIVATE ConjuntoDatos2.
DESCRIPTIVES VARIABLES=FECHANAC
/STATISTICS=MEAN STDDEV MIN MAX.

→ **Descriptivos**

Estadísticos descriptivos

| | N | Mínimo | Máximo | Media | Desv. Desviación |
|----------------------|-------|---------|---------|-----------|------------------|
| FECHA NAC | 25276 | 22357,0 | 40178,0 | 38541,176 | 1141,0914 |
| N válido (por lista) | 25276 | | | | |

Ilustración 28 – Interpretación de resultados. Fuente: Elaboración propia.

Los resultados del análisis descriptivo aparecerán en la ventana del editor de salidas. Estos resultados incluyen tablas con las estadísticas descriptivas seleccionadas y, opcionalmente, gráficos que ilustran la distribución de los datos.

López y Gómez (2015) afirman que "los resultados del análisis descriptivo en SPSS se presentan de manera clara y organizada, facilitando su interpretación y utilización en informes y presentaciones" (p. 42).

MEDIDAS DE TENDENCIA CENTRAL Y RESUMEN DE LOS DATOS (MEDIA, MODA Y MEDIANA)

Las medidas de tendencia central son herramientas fundamentales en el análisis de datos descriptivo, ya que nos permiten resumir y describir grandes conjuntos de datos en valores representativos. Las tres principales medidas de tendencia central son la media (promedio), la mediana (el valor central) y la moda (el valor que más se repite). Cada una de estas medidas proporciona una perspectiva distinta sobre el comportamiento de los datos, y su elección depende de la naturaleza y distribución de los mismos. En esta sección, aprenderemos a calcular e interpretar estas medidas y cómo aplicarlas adecuadamente según los datos disponibles (Hurtado, 2010; Hernández et al., 2014, Rivadeneira et al., 2020).

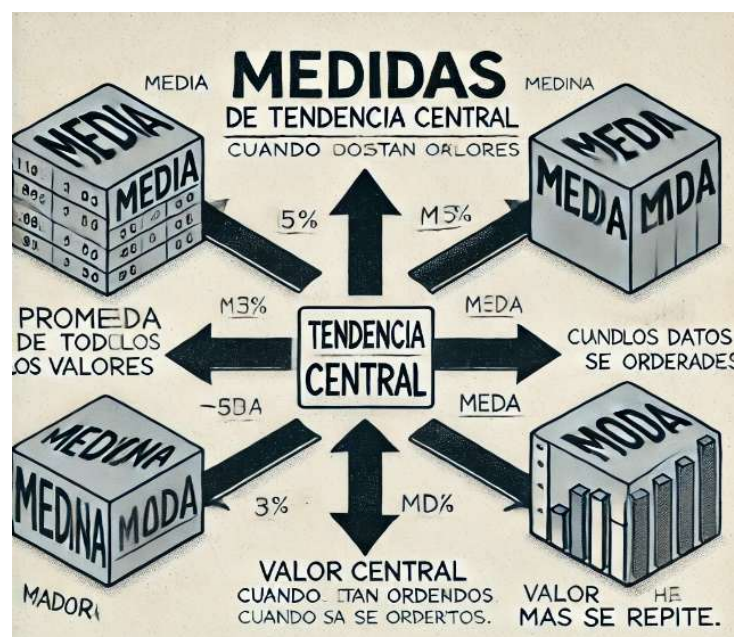


Ilustración 29 – Medidas de tendencia central, conceptos. Fuente: Elaboración propia.

CRITERIO DE USO DE LAS MEDIDAS DE TENDENCIA CENTRAL

No todas las medidas de tendencia central son igualmente útiles en todos los contextos. La elección de la media, moda o mediana depende de la distribución de los datos y su sensibilidad a valores extremos o atípicos. Por ejemplo, la media es altamente afectada por estos valores atípicos, mientras que la mediana es más robusta frente a tales casos. Este contenido se centrará en analizar criterios prácticos y teóricos para seleccionar la medida de tendencia central más adecuada según el contexto del análisis, permitiendo una interpretación precisa y adecuada de los resultados. A decir de León y Pérez (2019) “La finalidad del uso de medidas de tendencia central se dirige a darle significado a las puntuaciones individuales de los casos de estudio comparándolas con las medidas centrales” (p. 80).

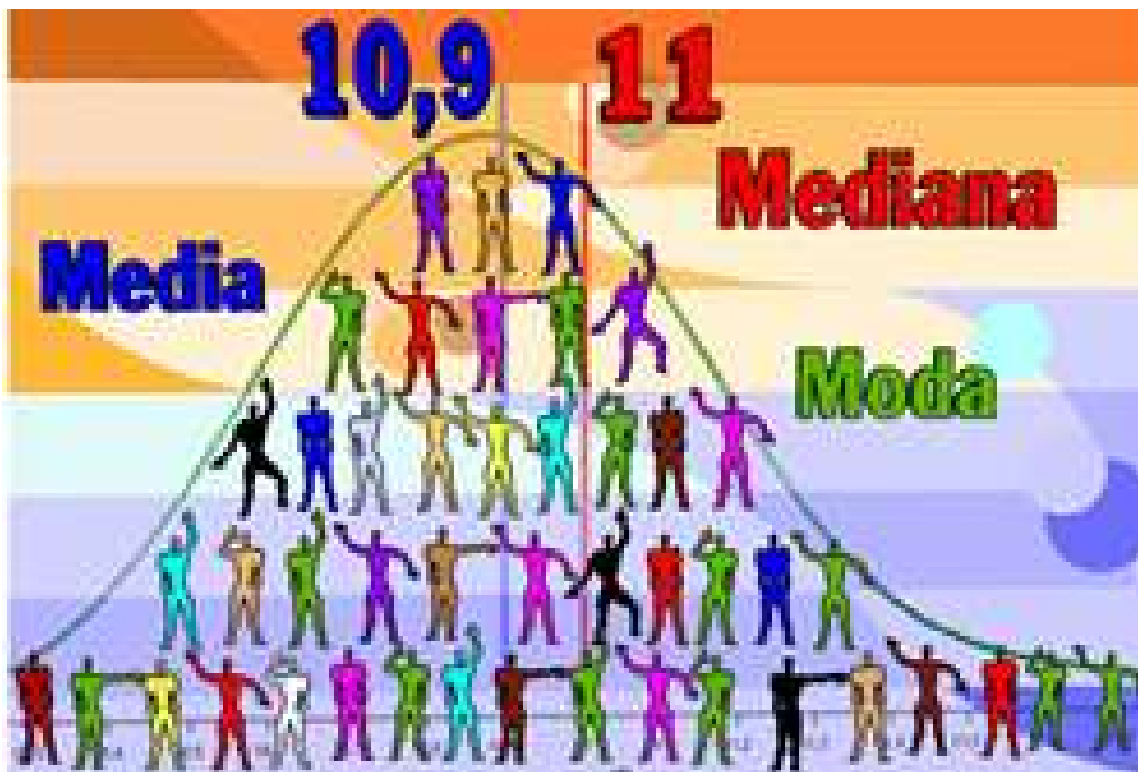
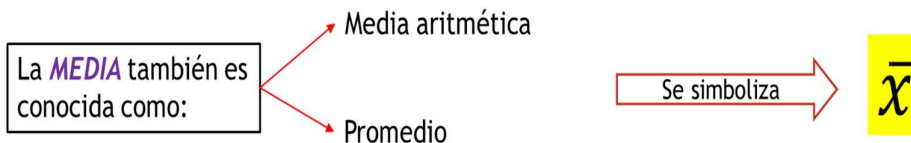


Ilustración 30 – Criterios de uso de medidas de tendencia central. Fuente: Elaboración propia.

MEDIA.

La media, también conocida como promedio, es una medida de tendencia central que se calcula sumando todos los valores de un conjunto de datos y dividiendo el resultado entre el número total de valores. Es útil para tener una idea general de la tendencia del conjunto, pero puede verse influenciada por valores extremos o atípicos. Por ejemplo, en un conjunto de ingresos donde la mayoría gana entre 30,000 y 50,000, pero hay una persona que gana 1,000,000, la media podría no reflejar con precisión la situación económica de la mayoría.

MEDIA



La definimos como el valor característico de la serie de dato.
O como el valor que representa a un conjunto de datos.

Si se trata de los datos de una muestra, estaremos hablando de la *media muestral*.
Si el conjunto de datos es toda la población, se llama *media poblacional*.

La media se puede utilizar solo para *variables cuantitativas*.

Se calcula sumando todos los datos (sin importar el orden de éstos), y luego dividiendo la suma entre la cantidad total de datos

Si son datos agrupados en intervalos se debe calcular la “marca de clase” antes de cualquier cosa.

Ilustración 31 – Media aritmética. Fuente: Elaboración propia.

Ejemplos:

1

Se tienen las siguientes notas de una evaluación de matemática de un curso de 3ro medio (notas redondeadas al entero)

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 5 | 6 | 7 | 4 | 2 | 5 | 5 | 5 | 2 | 3 | 5 | 1 | 5 |
| 4 | 3 | 2 | 3 | 5 | 5 | 5 | 6 | 7 | 5 | 4 | 5 | 3 | 4 | 4 |
| 5 | 5 | 6 | 5 | 3 | 2 | 5 | 7 | 5 | 6 | 4 | 3 | 5 | 6 | 4 |

| Notas | Frecuencia absoluta (f_i) | Dato $\cdot f_i$ |
|--------------|-------------------------------|-------------------|
| 1 | 1 | $1 \cdot 1 = 1$ |
| 2 | 6 | $2 \cdot 6 = 12$ |
| 3 | 6 | $3 \cdot 6 = 18$ |
| 4 | 7 | $4 \cdot 7 = 28$ |
| 5 | 17 | $5 \cdot 17 = 85$ |
| 6 | 5 | $6 \cdot 5 = 30$ |
| 7 | 3 | $7 \cdot 3 = 21$ |
| Total | 45 | |

Media (\bar{x}): 4,3

¿Cómo podemos interpretar esta información?

La media o promedio de las notas del 3ro medio es de 4,3.

Lo que permite deducir que, en general el curso aprobó esta evaluación.

$$\bar{x} = \frac{1 + 12 + 18 + 28 + 85 + 30 + 21}{45} = \frac{195}{45} = 4,3 \approx 4,3$$

Cálculo de la media aritmética cuando los datos se repiten.

- 1°. Se multiplican los datos por sus frecuencias absolutas respectivas, y se suman.
- 2°. El resultado se divide por el total de datos.

Ilustración 32 – Media aritmética, ejemplos. Fuente: Elaboración propia.

MEDIANA

La mediana es el valor que se encuentra en el medio de un conjunto de datos cuando estos se ordenan de menor a mayor. Si el número de valores es impar, la mediana es el valor central; si es par, se calcula como el promedio de los dos valores centrales. A diferencia de la media, la mediana no se ve afectada por valores extremos, lo que la convierte en una medida más robusta en situaciones donde los datos tienen una distribución sesgada. Por ejemplo, en el mismo conjunto de ingresos, la mediana proporcionaría una mejor representación de la mayoría de los ingresos.

MEDIANA

La mediana es un estadístico de posición central que parte la distribución en dos, es decir, deja la misma cantidad de valores a un lado que a otro.

Se simboliza

M_e

Partiendo del latín encontramos la forma más directa de definir este concepto matemático que proviene de *medianus*, que significa “del medio”.

La mediana es el punto medio del conjunto de datos. El valor de este punto medio es el punto en el cual la mitad de las observaciones está por encima del valor y la otra mitad está por debajo del valor.

La mediana se puede utilizar solo para *variables cuantitativas*.

Se calcula ordenando los datos en forma ascendente o descendente. La mediana será el dato del centro.

Si hubiese más de un dato central, se calcula el promedio de éstos

Ejemplos:

1

Se tienen las siguientes notas de una evaluación de matemática de un curso de 3ro medio (notas redondeadas al entero)

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 5 | 6 | 7 | 4 | 2 | 5 | 5 | 5 | 2 | 3 | 5 | 1 | 5 |
| 4 | 3 | 2 | 3 | 5 | 5 | 5 | 6 | 7 | 5 | 4 | 5 | 3 | 4 | 4 |
| 5 | 5 | 6 | 5 | 3 | 2 | 5 | 7 | 5 | 6 | 4 | 3 | 5 | 6 | 4 |

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 |

Son 45 datos, entonces la técnica es dividir 45 en 2, para saber exactamente dónde se encuentra el dato central

$$45:2 = 22,5 \approx 23$$

Este valor va a indicar una **POSICIÓN** en el orden que le dimos a nuestros datos. Como la posición es un número exacto, entonces debemos aproximar nuestro resultado si da decimal.

∴ Nuestra mediana se encuentra en el dato de la posición n° 23.

Mediana (Me): 5

¿Cómo podemos interpretar esta información?

La mediana de notas es 5, por lo que podemos deducir que en general, el curso obtuvo buen rendimiento, ya que la mitad derecha de la mediana es un valor mayor o igual a 5.

Ilustración 33 – Mediana ejemplos. Fuente: Elaboración propia.

- 2 Se estudió el ingreso de 500 familias en cierto sector de la RM. La siguiente tabla muestra el ingreso familiar en intervalos.

| Ingresos (en miles de pesos) | Frecuencia absoluta (f_i) | Frecuencias acumuladas (F_i) |
|---------------------------------|----------------------------------|-------------------------------------|
| [0 - 200] | 11 | 11 |
| [201 - 400] | 84 | 95 |
| [401 - 600] | 257 | 352 |
| [601 - 800] | 79 | 431 |
| [801 - 1000] | 41 | 472 |
| [1001 - 1200] | 15 | 487 |
| [1201 - 1400] | 8 | 495 |
| [1401 - 1600] | 5 | 500 |
| Total | 500 | |

Son 500 datos, entonces la técnica es dividir 500 en 2, para saber exactamente dónde se encuentra el dato central

$$500:2 = 250$$

Las frecuencias acumuladas me permiten obtener la posición de los datos hasta cierto punto.

Como es una cantidad par de datos, se obtienen dos datos centrales

Es decir, la mediana se encuentra entre el dato n° 250 y 251

Mediana (Me): [401,600]

¿Cómo podemos interpretar esta información?

La mediana de ingreso familiar es de \$401.000 a \$600.000. Lo que quiere decir que de la mitad de las familias percibe un ingreso menor o igual a \$600.000.

MODA

La moda es la medida de tendencia central que representa el valor o los valores más frecuentes en un conjunto de datos. Un conjunto puede tener una moda (unimodal), más de una moda (multimodal) o ninguna (si todos los valores son únicos). La moda es especialmente útil en datos categóricos donde se desea conocer el elemento más común. Por ejemplo, en una encuesta sobre preferencias de colores, el color más elegido sería la moda, proporcionando información sobre las preferencias del grupo.

MODA

La moda es el valor que tiene mayor frecuencia absoluta.

Se simboliza

M_o

Si en un grupo de datos, existen dos datos que tengan la misma frecuencia absoluta, entonces el grupo de datos es **Bimodal**

Si en un grupo de datos, existen más de dos datos que tengan la misma frecuencia absoluta, entonces el grupo de datos es **Multimodal**.

Si en un grupo de datos, todos los datos tienen la misma frecuencia absoluta, entonces el grupo de datos es **Amodal**, es decir, no hay moda.

Se puede determinar la moda para variables cualitativas y cuantitativas.

Cuando se tienen datos agrupados en intervalos, entonces se habla de “intervalo modal”

La **moda** de un conjunto de datos es el dato que más se repite.

Ejemplo. Una zapatería ha vendido en una semana los zapatos que se reflejan en la tabla:

| | | | | | | | | |
|----------------|----|----|----|----|----|----|----|----|
| Nº de calzado | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
| Nº de personas | 16 | 21 | 30 | 35 | 29 | 18 | 10 | 7 |

El número de zapato más vendido, el dato con mayor frecuencia absoluta, es el 41.

Lo compran 35 personas

La moda es 41.

Ejemplos:

- 1 En un colegio, se pregunta a estudiantes de 3ro medio cuál es la red social que más utiliza. Los resultados son los siguientes:



| Red social | Frecuencia absoluta (f_i) |
|------------|-------------------------------|
| Instagram | 64 |
| TikTok | 49 |
| Facebook | 8 |
| Snapchat | 9 |
| Otra | 5 |
| Total | 135 |

Moda (Mo): Instagram

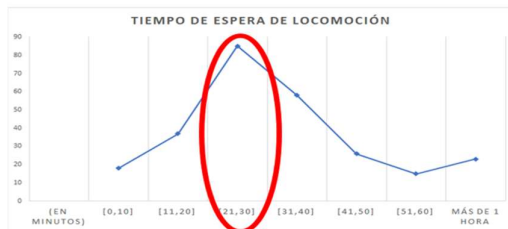
¿Cómo podemos interpretar esta información?

La red social de moda entre los estudiantes de 3ro medio es Instagram.

Ilustración 34 – Moda y ejemplos. Fuente: Elaboración propia.

2

Se ha preguntado a un grupo de estudiantes: ¿cuánto tiempo han tenido que esperar para tomar el transporte público para llegar a sus colegios?



| Tiempo de espera (en minutos) | Frecuencia absoluta (f_i) |
|-------------------------------|-------------------------------|
| [0,10] | 18 |
| [11,20] | 37 |
| [21,30] | 85 |
| [31,40] | 58 |
| [41,50] | 26 |
| [51,60] | 15 |
| Más de 1 hora | 23 |
| Total | 262 |

Intervalo Modal (Mo): [21,30]

¿Cómo podemos interpretar esta información?

El tiempo más frecuente que deben esperar los estudiantes es entre 21 y 30 minutos para tomar locomoción pública para llegar a sus colegios.

OBSERVACIONES



En ocasiones la Moda, Mediana y media coinciden en su valor numérico, pero no siempre es así.

Hay ocasiones, dependiendo del contexto, y el tipo de datos, en que estas medidas de tendencia central no son suficientes para realizar un análisis que refleje la fiel realidad de los datos.

Para ello vamos a estudiar, más adelante, las **MEDIDAS DE DISPERSIÓN**.

Ilustración 35 – Ejemplos de medida de tendencia central. Fuente: Elaboración propia.

MEDIDAS APROPIADAS DE TENDENCIA CENTRAL SEGÚN EL TIPO DE VARIABLE

Cada tipo de variable, ya sea cualitativa o cuantitativa, exige el uso de una medida de tendencia central apropiada. Mientras que la media solo es aplicable a variables cuantitativas, la moda puede emplearse tanto para variables cualitativas como cuantitativas. La mediana, por su parte, es especialmente útil en variables ordinales o cuantitativas cuando los datos están distribuidos de manera asimétrica. Este apartado permitirá entender la relación entre los tipos de variables y la selección de la medida de tendencia central más adecuada, ayudando a los estudiantes a aplicar correctamente estas medidas en investigaciones y análisis cuantitativos (De la Puente, 2018; Gamboa, 2018; Pardo, 2020; Cortés et al., 2023).

TIPOLOGÍA DE VARIABLES

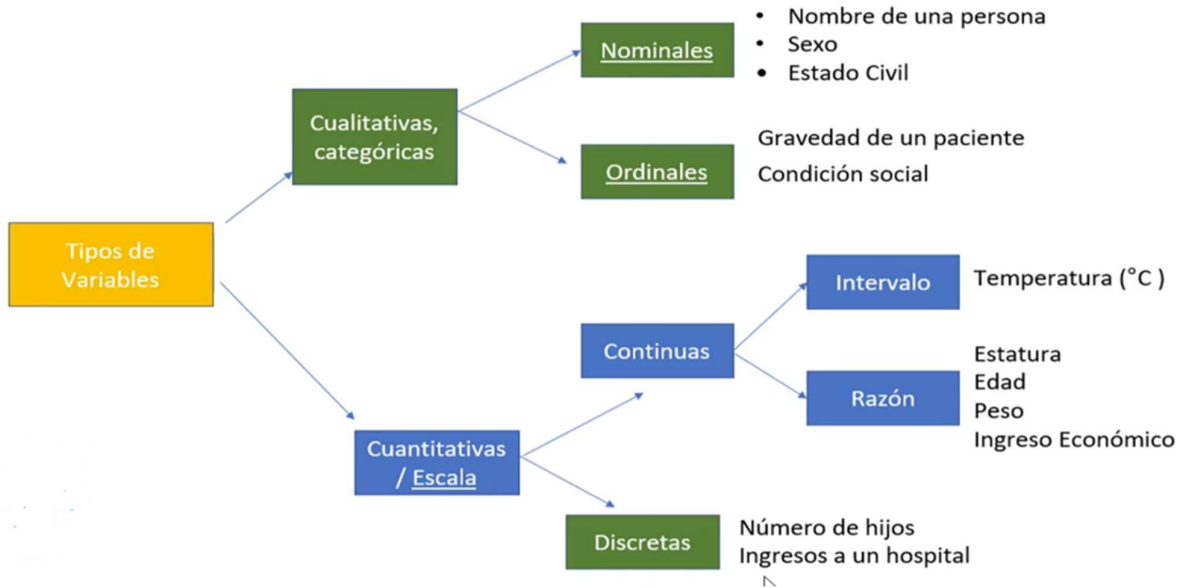
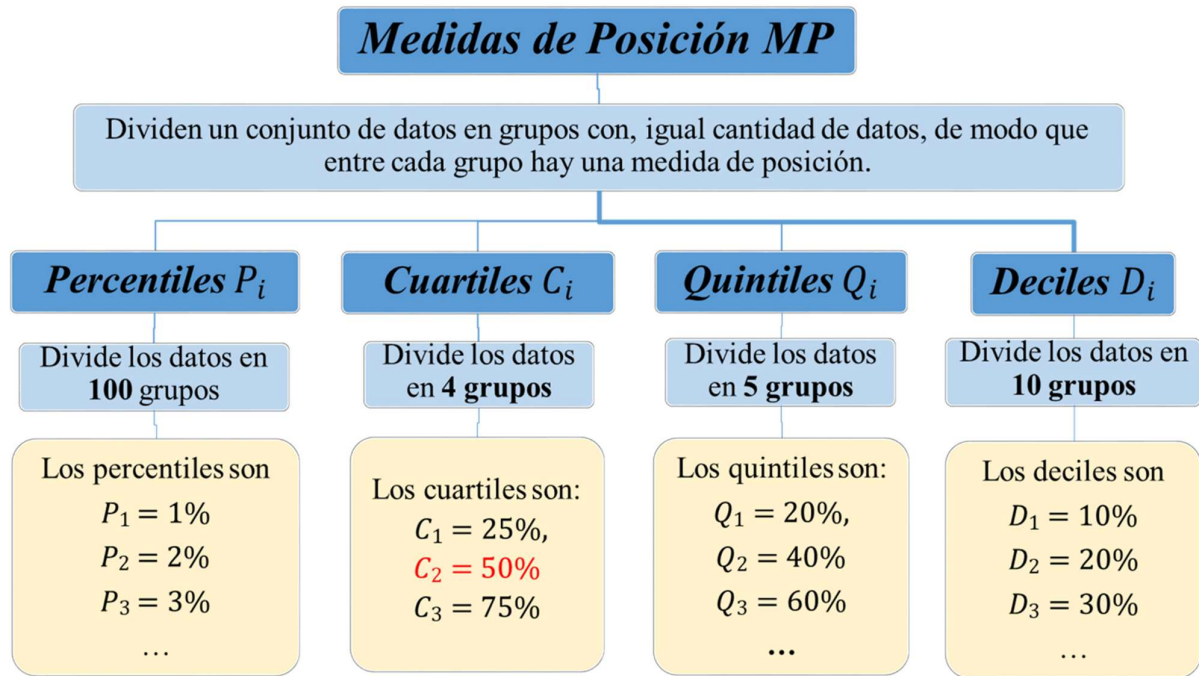


Ilustración 36 –Tipos de variables. Fuente: Elaboración propia.

MEDIDAS DE POSICIÓN: PERCENTILES, CUARTILES, QUINTILES Y DECILES.



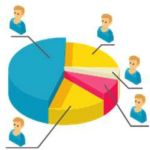
Definición, conceptos básicos y procedimiento de cálculos de medidas de posición.

Las medidas de posición, como percentiles, cuartiles, quintiles y deciles, son herramientas estadísticas utilizadas para dividir un conjunto de datos en partes iguales, facilitando la interpretación y comparación de los datos en diferentes puntos de la distribución. Estas medidas son fundamentales en la estadística descriptiva porque proporcionan puntos de referencia que describen la localización relativa de un dato dentro del conjunto. Permiten identificar la centralidad y la dispersión de los datos, así como realizar comparaciones entre diferentes conjuntos de datos. En la investigación cuantitativa, estas medidas ayudan a identificar tendencias y patrones dentro de los datos, lo que es esencial para la toma de decisiones basadas en evidencia (Triola, 2018).

CUARTILES

Los cuartiles son valores que dividen un conjunto de datos en cuatro partes iguales. Se utilizan para entender la dispersión y la posición relativa de los datos. El primer cuartil (Q_1) es el valor que separa el 25% inferior de los datos, el segundo cuartil (Q_2) es la mediana, que divide el conjunto en dos mitades, y el tercer cuartil (Q_3) separa el 75% inferior del 25% superior. Los cuartiles son especialmente útiles para identificar rangos de datos y detectar valores atípicos.

CUARTILES



Los cuartiles dividen en cuatro partes las observaciones. El primer cuartil Q_1 es un valor que deje por debajo de él 25% de las y por encima 75% de las observaciones. El Q_2 es la mediana (50%) y Q_3 deja por debajo 75% y por encima 25% de las observaciones

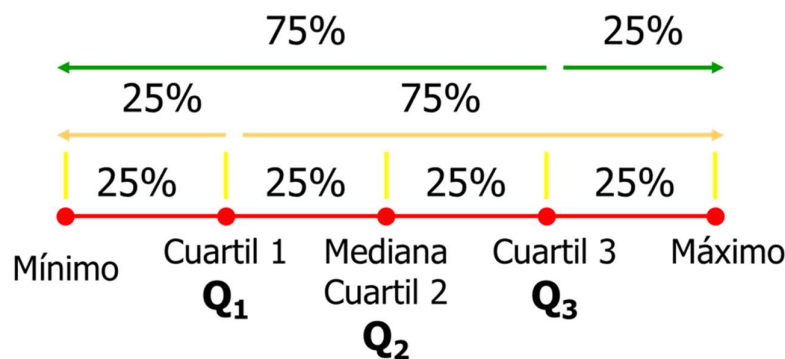


Ilustración 37 –Cuartiles. Fuente: Elaboración propia.

Los Cuartiles dividen un conjunto de datos en cuatro partes iguales, lo que significa que cada cuartil representa el 25% de los datos.

- ✚ **Primer Cuartil (Q_1):** Es el valor por debajo del cual se encuentra el 25% de los datos.
- ✚ **Segundo Cuartil (Q_2):** Corresponde a la mediana y divide los datos en dos partes iguales, dejando el 50% por debajo de él.
- ✚ **Tercer Cuartil (Q_3):** Es el valor que deja por debajo al 75% de los datos.

CUARTILES

Cálculo de cuartiles

Dependiendo del tamaño de la muestra n (*cantidad total de datos*), la fórmula y el proceso de resolución cambia. Las fórmulas entregan la **posición de la medida** en la distribución.

Tamaño de muestra **Impar**

$$C_1 = x_{\frac{n+1}{4}} \quad C_2 = x_{\frac{n+1}{2}}$$

$$C_3 = x_{\frac{3(n+1)}{4}}$$

Dependiendo del tamaño:

Caso 1: Las posiciones corresponde a los cuartiles.

Caso 2: Los datos de promedian.

Tamaño de muestra **Par**

$$C_1 = \frac{x_{\frac{n}{4}} + x_{\frac{n}{4}+1}}{2} \quad C_2 = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

$$C_3 = \frac{x_{\frac{3n}{4}} + x_{\frac{3n}{4}+1}}{2}$$

Los datos siempre se van a promediar.

Caso 1: Se deben promediar todos los datos.

Caso 2: Se deben promediar algunos datos.

Los cuartiles se calculan de forma similar a los percentiles, pero en lugar de dividir los datos en 100 partes, se dividen en 4 partes:

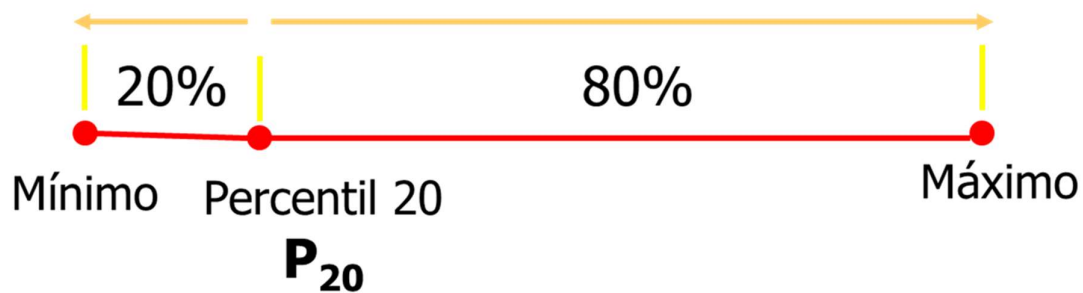
$$Q_1 = P_{25} \quad Q_2 = P_{50} \quad Q_3 = P_{75}$$

PERCENTILES

Los percentiles dividen un conjunto de datos en 100 partes iguales, proporcionando una visión más detallada de la distribución de los datos. El percentil P_k es el valor debajo del cual se encuentra el k -ésimo porcentaje de los datos. Por ejemplo, el percentil 50 es la mediana, el percentil 25 corresponde al primer cuartil (Q_1) y el percentil 75 al tercer cuartil (Q_3). Los percentiles son útiles en diversas aplicaciones, como en la evaluación de resultados académicos o en estudios de salud, donde se quiere entender cómo se posiciona un individuo en relación con un grupo.

PERCENTILES

Los percentiles dividen en dos partes las observaciones. Por ejemplo, el percentil 20, P20, es el valor que deja por debajo un 20% y por encima un 80% de las observaciones



Los **percentiles** dividen un conjunto de datos en 100 partes iguales, es decir, cada percentil representa el 1% de los datos. Por ejemplo, el percentil 50 (P50) corresponde al valor que deja por debajo de él al 50% de los datos, coincidiendo con la mediana en una distribución simétrica.

Cálculo: Si se tiene un conjunto de datos ordenado de menor a mayor, el percentil k (P_k) se calcula utilizando la fórmula:

$$P_{K=} \frac{K}{100} * (n + 1)$$

donde k es el número del percentil que se desea encontrar y n es el número total de datos.

QUINTILES

Los **quintiles** dividen un conjunto de datos en cinco partes iguales, con cada quintil representando el 20% de los datos. Son menos comunes que los cuartiles y percentiles, pero son útiles en algunos análisis económicos y sociodemográficos.

El primer **quintil (Q1)** corresponde al percentil 20, el segundo quintil (Q2) al percentil 40, y así sucesivamente.

$$Q_1 = P_{20} \quad Q_2 = P_{40} \quad Q_3 = P_{60} \quad Q_4 = P_{80}$$

DECILES

Los **deciles** dividen un conjunto de datos en diez partes iguales, con cada decil representando el 10% de los datos. Son útiles para análisis detallados donde se requiere mayor precisión que con cuartiles.

Los **deciles** se calculan dividiendo los datos en 10 partes, por lo que el primer decil (D1) es el percentil 10, el segundo decil (D2) es el percentil 20, y así sucesivamente:

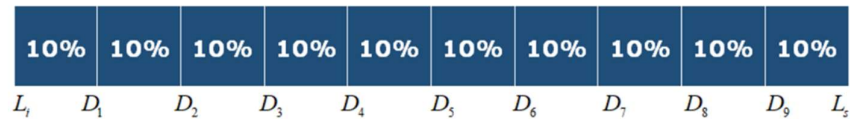
$$D_1 = P_{10} \quad D_2 = P_{20}, \dots \dots \dots D_9 = P_{90}$$

El mapa de contenidos ilustra cómo las diferentes medidas de posición dividen un conjunto de datos en partes iguales. El concepto central de "División de Datos" se desglosa en cuatro ramas principales: Percentiles, Cuartiles, Quintiles y Deciles. Cada una de estas medidas divide los datos de manera específica: los Percentiles en 100 partes, los Cuartiles en 4 partes, los Quintiles en 5 partes y los Deciles en 10 partes.

DECILES

Son valores de la variable que dividen los datos ordenados en diez partes iguales (9 divisiones).

Datos clasificados en orden ascendente



Datos no agrupados y n es par

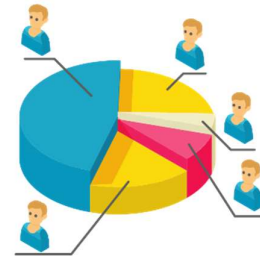
$$D_k = \frac{k \cdot n}{10}$$

Datos no agrupados y n es impar

$$D_k = \frac{k(n+1)}{10}$$

k= Decil deseado (1, 2,3,4,5,6,7,8 y 9)

n= Total de datos



DECILES

Dados los siguientes datos ordenados de menor a mayor, halla el primer, el tercer y el octavo decil de la muestra.

| Datos ordenados | | | | | |
|-----------------|-----|-----|-----|-----|-----|
| 81 | 82 | 85 | 89 | 90 | 91 |
| 92 | 94 | 97 | 99 | 101 | 105 |
| 108 | 109 | 111 | 114 | 119 | 120 |
| 121 | 123 | 125 | 126 | 128 | 131 |
| 135 | 139 | 146 | 148 | 150 | |

El resultado de la fórmula es un 3, por lo que el primer decil estará en la tercera posición de la lista ordenada, y ese valor corresponde a 85.

$$\frac{k \cdot (n + 1)}{10} \quad k = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

El tamaño de la muestra de este ejercicio es de 29 observaciones, por lo tanto, para calcular la posición del primer decil debemos sustituir la n por 29 y la k por 1:

$$\frac{1 \cdot (29 + 1)}{10} = 3 \quad \rightarrow \quad D_1 = 85$$

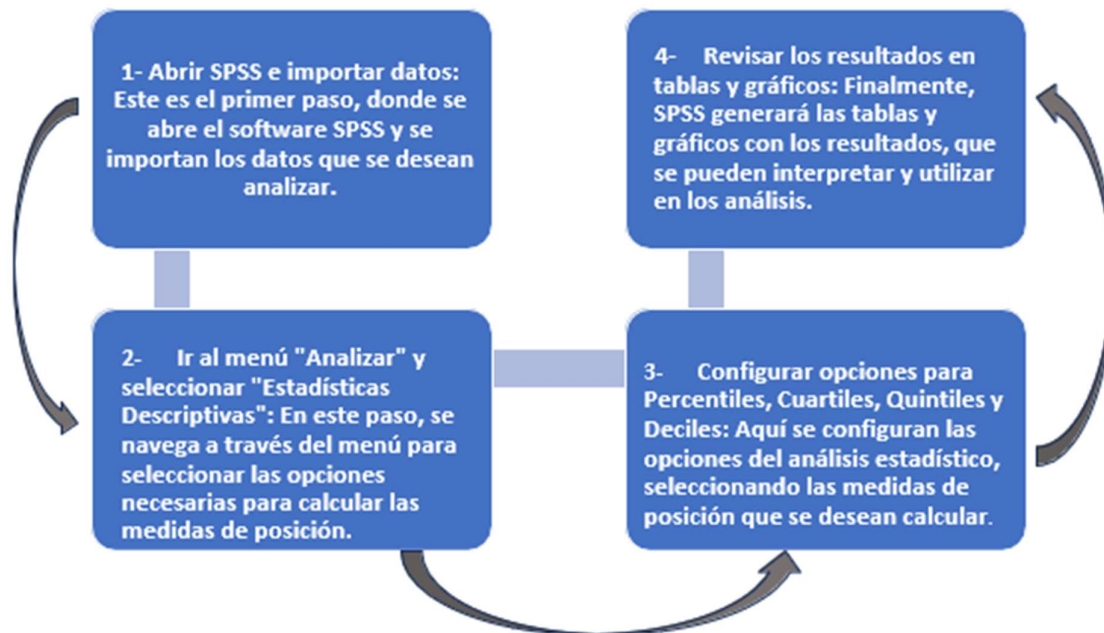


Ilustración 45 –Medidas de Posición. Fuente: Elaboración propia.

APLICACIÓN PRÁCTICA DE LAS MEDIDAS DE POSICIÓN EN ESTUDIOS CUANTITATIVOS

Aplicación de los Percentiles: son ampliamente utilizados en estudios educativos y psicológicos para comparar el rendimiento de individuos o grupos en relación con una población de referencia (Agresti & Finlay, 2017)

Ejemplo: Un examen estandarizado que utiliza percentiles para determinar cómo se desempeñó un estudiante en comparación con otros estudiantes.

Un estudiante que está en el percentil 85 en un examen significa que se desempeñó mejor que el 85% de los estudiantes que tomaron el examen. Este tipo de medida es útil para identificar a los estudiantes que están por encima o por debajo de ciertos umbrales de rendimiento.

Aplicación de los Cuartiles: son útiles en la evaluación de la distribución de ingresos, tiempo de reacción en experimentos, o cualquier otra métrica que se beneficie de ser dividida en cuatro

partes. Son comunes en la creación de gráficos de caja y bigotes (boxplots) para visualizar la dispersión de los datos.

- 📊 El primer cuartil (Q1) representa el valor por debajo del cual se encuentra el 25% de los datos.
- 📊 El segundo cuartil (Q2) es la mediana, y el tercer cuartil (Q3) indica el valor por debajo del cual se encuentra el 75% de los datos. Este análisis ayuda a identificar la tendencia central y la dispersión de los datos.

Aplicación de Quintiles: se utilizan comúnmente en análisis económicos y sociodemográficos para dividir una población en cinco grupos iguales. Por ejemplo, en estudios de distribución de riqueza o ingresos, los quintiles permiten analizar cómo se distribuye la riqueza entre diferentes segmentos de la población.

Ejemplo: El primer quintil agrupa el 20% más bajo de la población en términos de ingresos, mientras que el quinto quintil agrupa el 20% más alto. Esta segmentación ayuda a los investigadores a entender las diferencias entre los grupos socioeconómicos y a formular políticas dirigidas a grupos específicos.

Aplicación de Deciles: se utilizan en estudios financieros y de mercado para analizar la distribución de variables como ingresos, puntuaciones de riesgo o rendimiento de inversiones. También son útiles para la segmentación de clientes en marketing.

Ejemplo: un análisis de deciles podría mostrar que el 10% superior de los clientes (décimo decil) genera el 50% de los ingresos de una empresa. Esto permite a las empresas identificar y enfocarse en los clientes más valiosos.

INTERPRETACIÓN DE LOS ANÁLISIS CON MEDIDAS DE POSICIÓN

Distribución y Desigualdad: Las medidas de posición ayudan a interpretar cómo se distribuyen los datos dentro de un conjunto y a identificar la desigualdad o concentración en la distribución. Por ejemplo, una concentración alta de ingresos en los quintiles superiores podría indicar una desigualdad económica significativa (Sullivan, 2019).

Identificación de Outliers: Los cuartiles y percentiles son útiles para identificar valores atípicos (outliers) en los datos. Un valor que está significativamente por encima del tercer cuartil (Q3) o por debajo del primer cuartil (Q1) podría considerarse un outlier.

Comparación entre Grupos: Al utilizar medidas de posición, los investigadores pueden comparar diferentes grupos dentro de un estudio. Por ejemplo, pueden comparar los percentiles de rendimiento académico entre diferentes escuelas para identificar diferencias significativas en la calidad educativa.

Tendencias Centrales y Dispersión: Estas medidas permiten una comprensión más profunda de la tendencia central (como la mediana) y la dispersión (rango intercuartil) dentro de un conjunto de datos. Esto es crucial para interpretar la consistencia y la variabilidad de los datos en estudios cuantitativos.

RANGO

Es la diferencia entre el valor máximo y el valor mínimo de un conjunto de datos. El rango, en términos generales, representa la amplitud o extensión de un conjunto de valores. Es decir, nos indica la diferencia entre el valor más alto y el valor más bajo dentro de un determinado conjunto de datos. El rango nos proporciona una idea rápida y sencilla sobre cuán dispersos están los datos, aunque es una medida que puede verse afectada por valores extremos (Lacourly, 2012).

- **Fórmula:** Rango = Valor máximo - Valor mínimo
- **Interpretación:** Nos proporciona una idea rápida de la amplitud de los datos, pero es sensible a valores extremos (Cárdenas Antúñez, 2014).

Rango Intercuartílico

Es la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1). El rango intercuartílico es una medida de dispersión estadística que nos indica la amplitud de la mitad central de un conjunto de datos. Se calcula restando el primer cuartil (Q1) al tercer cuartil (Q3). A diferencia del rango, que considera todos los datos, el rango intercuartílico se enfoca en la dispersión de los valores centrales, lo que lo hace menos susceptible a la influencia de valores atípicos. Es una medida más robusta y se utiliza comúnmente en la construcción de diagramas de caja y bigotes.

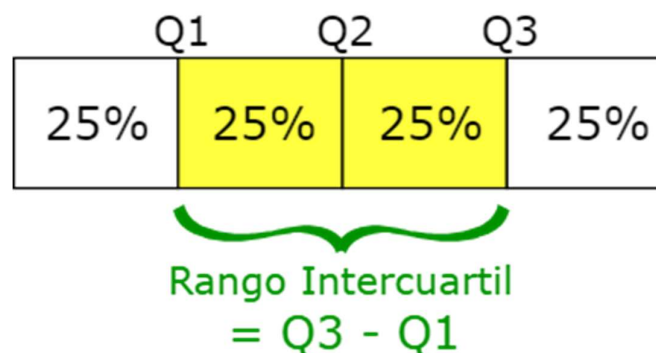


Ilustración 38 –Rango Intercuartil. Fuente: Elaboración propia.

Interpretación: Representa el rango de los valores que contienen el 50% central de los datos, siendo menos sensible a valores atípicos que el rango (Lacourly, 2012).

El siguiente gráfico ilustra los aspectos más sobresalientes relacionados al rango:

| Concepto | Explicación | Ejemplo |
|---------------------------|---|--|
| Rango | Es la diferencia entre el valor más grande y el valor más pequeño de un conjunto de datos. | Imagina que tenemos las siguientes edades de un grupo de amigos: 15, 18, 20, 22, 25. El rango sería: 25 (edad mayor) - 15 (edad menor) = 10 años. |
| ¿Qué nos dice el rango? | Nos indica cuán dispersos están los datos. Un rango grande significa que los datos están muy dispersos, mientras que un rango pequeño indica que los datos están más agrupados. | Si el rango de las edades es pequeño (por ejemplo, 2 años), significa que todos los amigos tienen edades muy similares. Si el rango es grande (por ejemplo, 10 años), significa que hay una gran diferencia de edad entre los más jóvenes y los mayores. |
| ¿Para qué sirve el rango? | * Obtener una idea rápida de la dispersión: Nos da una primera impresión de cuán variados son los datos. * Comparar conjuntos de datos: Podemos comparar el rango de diferentes conjuntos para ver cuál tiene mayor variabilidad. | * Si comparamos el rango de las alturas de dos equipos de baloncesto, podemos ver cuál equipo tiene jugadores de alturas más variadas. |
| Limitaciones del rango | * Sensibilidad a valores extremos: Un solo valor muy grande o muy pequeño puede afectar significativamente el rango. * No considera la distribución de los datos: El rango solo nos dice la diferencia entre los valores extremos, pero no nos informa sobre cómo se distribuyen los demás valores. | * Si en un grupo de edades incluimos a una persona de 80 años, el rango aumentará mucho, aunque la mayoría de las personas tengan edades similares. |

Ilustración 39 –Características Rango Intercuartil. Fuente: Elaboración propia.

VARIANZA

Es el promedio de las desviaciones cuadráticas de cada dato respecto a la media. La varianza es una medida estadística que nos indica cuán dispersos están los datos con respecto a su media. Es decir, nos muestra qué tan alejados están, en promedio, los valores de un conjunto de datos de su valor central (la media). A mayor varianza, mayor dispersión de los datos (Ruiz Muñoz, 2012).

Fórmula de la varianza: $(s^2) = \sum(x_i - \bar{x})^2 / (n-1)$

Donde:

- x_i : cada valor de los datos
- \bar{x} : media de los datos
- n : número de datos (Cárdenas Antúnez, 2014).

Interpretación: Indica qué tan dispersos están los datos en relación a la media. Sin embargo, su unidad de medida es el cuadrado de la unidad de los datos originales, lo que puede dificultar su interpretación directa.

Desviación Estándar

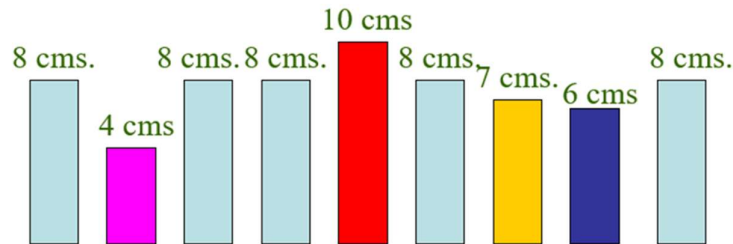
Es la raíz cuadrada de la varianza. La desviación estándar es una medida estadística que nos indica qué tan dispersos están los datos de un conjunto con respecto a su media. Es decir, nos muestra cuán alejados están, en promedio, los valores individuales de la media aritmética (Martínez Ruiz, 2012).

- **Fórmula de la desviación estándar:** $(s) = \sqrt{\text{Varianza}}$
- **Interpretación:** Expresa la dispersión de los datos en las mismas unidades que los datos originales, facilitando su interpretación. Es la medida de dispersión más utilizada (Martínez, 2020).

| Característica | Varianza | Desviación Estándar |
|------------------|---|--|
| Definición | Es el promedio de las desviaciones cuadráticas de cada dato respecto a la media. | Es la raíz cuadrada positiva de la varianza. |
| Fórmula | $\Sigma(x_i - \bar{x})^2 / (n-1)$ | $\sqrt{\Sigma(x_i - \bar{x})^2 / (n-1)}$ |
| Unidad de medida | Unidades al cuadrado de los datos originales. | Misma unidad que los datos originales. |
| Interpretación | Indica qué tan dispersos están los datos con respecto a la media, pero su unidad no es intuitiva. | Representa la dispersión típica de los datos en las mismas unidades que los datos originales, lo que facilita su interpretación. |
| Relación | La desviación estándar es la raíz cuadrada de la varianza. | La varianza es el cuadrado de la desviación estándar. |
| Uso | Se utiliza principalmente en cálculos estadísticos posteriores, como en el análisis de varianza. | Se utiliza para comparar la dispersión de diferentes conjuntos de datos y para realizar pruebas estadísticas. |
| Ventajas | Proporciona una medida precisa de la dispersión. | Fácil de interpretar y comparar, ya que está en las mismas unidades que los datos originales. |
| Desventajas | La unidad de medida no es intuitiva, lo que dificulta su interpretación directa. | Puede ser sensible a valores atípicos. |

Ilustración 48 – Diferencias entre la varianza y la desviación estándar. Fuente: (Cárdenas Antúnez, 2014).

DESVIACIÓN ESTÁNDAR

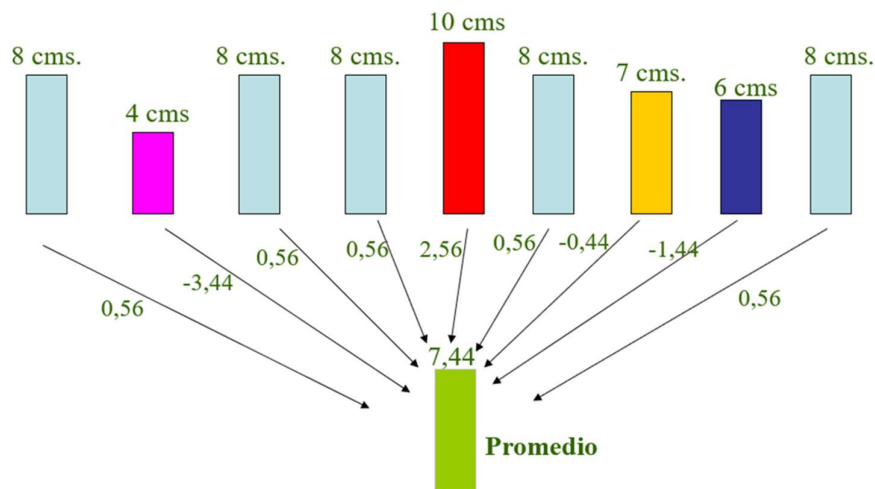


¿Cuál es la varianza y la desviación estándar de las alturas de los rectángulos?

En primer lugar debemos calcular el promedio

$$\frac{8 + 4 + 8 + 8 + 10 + 8 + 7 + 6 + 8}{9} = 7,44$$

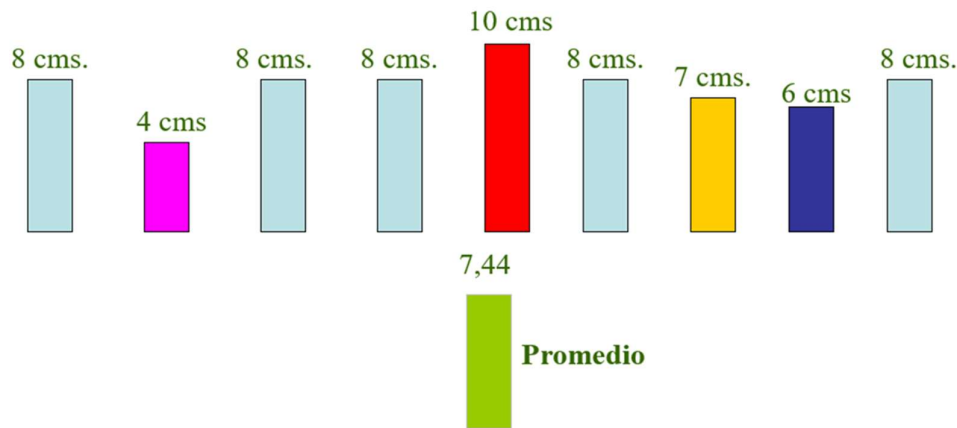
Luego debemos calcular la varianza



$$\frac{0,56^2 + (-3,44)^2 + 0,56^2 + 0,56^2 + 2,56^2 + 0,56^2 + (-0,44)^2 + (-1,44)^2 + 0,56^2}{9} = \frac{22,2224}{9}$$

Este es el valor de la varianza → = 2,469

DESVIACIÓN ESTÁNDAR



Si la varianza fue de 2,469, entonces la desviación estándar es de...

$$\sqrt{2,469} = 1,57$$

Lo que significa que, en promedio, los rectángulos se desviaron más o menos (más arriba o más abajo) en 1,57 centímetros.

Ilustración 40 – Ejemplo de desviación estándar. Fuente: (Cárdenas Antúnez, 2014).

COEFICIENTE DE VARIACIÓN

El coeficiente de variación (CV) es una medida estadística que nos permite comparar la dispersión relativa de diferentes conjuntos de datos, incluso si están en diferentes escalas. En otras palabras, nos indica qué tan grande es la desviación estándar en comparación con la media. Es una herramienta útil para determinar si un conjunto de datos es homogéneo o heterogéneo (Lacourly, 2012).

Fórmula del Coeficiente de variación: $(CV) = (s / \bar{x}) * 100\%$

Interpretación: Permite comparar la dispersión de diferentes conjuntos de datos, incluso si están en diferentes escalas (Arribas & Barbut, 2014).

Varianza

La Varianza es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística.

Datos no Agrupados:

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

Datos Agrupados:

$$(\sigma^2) = \frac{\sum_{i=1}^n (x_{mci} - \bar{x})^2 \cdot f_i}{n}$$

Características de la Varianza

- Es siempre un valor no negativo, que puede ser igual o distinta de 0.
- La Varianza es la medida de dispersión cuadrática optima por ser la menor de todas.
- Si a todos los valores de la variable se le suma una constante la varianza no se modifica.
- Si todos los valores de la variable se multiplican por una constante la varianza queda multiplicada por el cuadrado de dicha constante.

Ilustración 41 – Varianza como medida. Fuente: Elaboración Propia

El coeficiente de variación bajo (CV bajo), indica una baja dispersión relativa, es decir, los datos están relativamente cerca de la media. Mientras tanto que, el coeficiente de variación alto (CV alto) por su parte, indica una alta dispersión relativa, es decir, los datos están muy dispersos alrededor de la media. Dentro de sus usos más frecuentes se destacan los siguientes (Martínez, 2020):

- **Comparación entre conjuntos de datos:** Permite comparar la dispersión de conjuntos de datos con diferentes unidades de medida o escalas. Por ejemplo, podemos comparar la dispersión de los salarios en dos países diferentes, aunque las monedas sean distintas (Martínez, 2020).
- **Análisis de la homogeneidad:** Un CV bajo indica que los datos están relativamente cerca de la media, mientras que un CV alto sugiere una mayor dispersión (Martínez, 2020).
- **Identificación de valores atípicos:** Un CV elevado puede indicar la presencia de valores atípicos que están influyendo en la dispersión de los datos (Martínez, 2020). A continuación, se detalla un gráfico que ilustra mejor lo que es el coeficiente de variación.

| Característica | Descripción | Fórmula | Interpretación |
|------------------------------|--|---|--|
| Definición | Mide la dispersión relativa de un conjunto de datos, es decir, compara la desviación estándar con la media. | $CV = \left(\frac{s}{\bar{x}} \right) 100\%$ | Indica qué tan grande es la desviación estándar en relación a la media. Un CV alto indica una alta dispersión relativa, mientras que un CV bajo indica una baja dispersión relativa. |
| Componentes | * s: Desviación estándar: Mide la dispersión típica de los datos. * \bar{x} : Media: Representa el valor central de los datos. | | |
| Unidad de medida | Porcentaje (%) | | |
| Ventajas | * Comparación entre conjuntos de datos: Permite comparar la dispersión de conjuntos de datos con diferentes unidades de medida o escalas. * Independencia de la escala: El CV es una medida adimensional, por lo que puede utilizarse para comparar conjuntos de datos con diferentes unidades de medida. * Identificación de valores atípicos: Un CV alto puede indicar la presencia de valores atípicos. | | |
| Desventajas | * No se puede calcular si la media es cero: Si la media es cero, el CV no está definido. * Sensibilidad a valores extremos: Al igual que la desviación estándar, el CV puede ser influenciado por valores atípicos. | | |
| Aplicaciones | * Análisis financiero: Comparar la volatilidad de diferentes activos. * Control de calidad: Evaluar la consistencia de un proceso de producción. * Investigación científica: Comparar la variabilidad de diferentes grupos de estudio. * Análisis de datos en general: Siempre que se quiera comparar la dispersión de diferentes conjuntos de datos. | | |
| Interpretación del resultado | * CV bajo: Los datos están relativamente cerca de la media (baja dispersión). * CV alto: Los datos están muy dispersos alrededor de la media (alta dispersión). | | |

Ilustración 51 – Definiciones clave de lo que es el coeficiente de variación. Fuente: (Cárdenas Antúnez, 2014).

INTERACCIÓN ENTRE MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN

Las medidas de tendencia central (media, mediana, moda) y las medidas de dispersión (rango, varianza, desviación estándar) proporcionan una imagen más completa de un conjunto de datos variación (Cárdenas Antúnez, 2014).

Media y desviación estándar: La media indica el centro de los datos, mientras que la desviación estándar indica qué tan dispersos están alrededor de ese centro (Hernández, et. al. 2014).

Mediana y rango intercuartílico: La mediana representa el valor central de los datos cuando están ordenados, y el rango intercuartílico indica la dispersión de la mitad central de los datos (Cárdenas Antúnez, 2014).

Ejemplo:

A continuación, dos grupos de estudiantes que sacaron las siguientes notas en un examen:

- **Grupo A:** 70, 75, 80, 85, 90 (media = 80, desviación estándar = 7.07)
- **Grupo B:** 60, 70, 80, 90, 100 (media = 80, desviación estándar = 14.14)

Ambos grupos tienen la misma media (80), pero el grupo B tiene una mayor desviación estándar, lo que indica que las calificaciones del grupo B están más dispersas alrededor de la media (Lacourly, 2012).

Representación Gráfica

- **Histograma:** Muestra la distribución de frecuencias de los datos en intervalos (Arribas & Barbut, 2014).
- **Diagrama de caja:** Representa la distribución de los datos a través de cinco valores: mínimo, primer cuartil, mediana, tercer cuartil y máximo (Lacourly, 2012).
- **Gráfico de dispersión:** Muestra la relación entre dos variables numéricas (Hernández, et. al. 2014). Los gráficos a continuación ilustran mejor el tema de los diagramas:

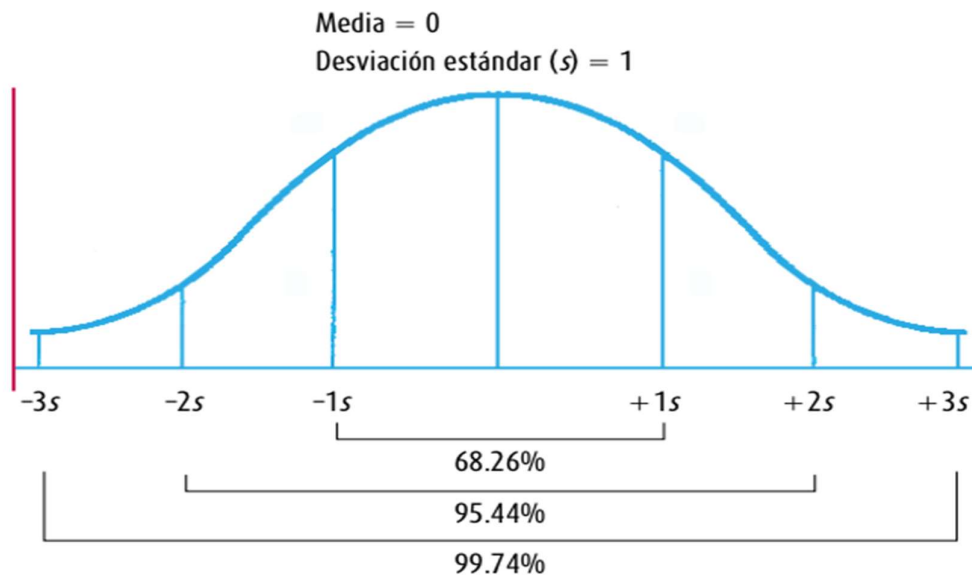


Ilustración 42 –Ejemplo de distribución normal (Hernández, et. al. 2014)

Opinión acerca del actual alcalde del municipio de San Martín Aurelio

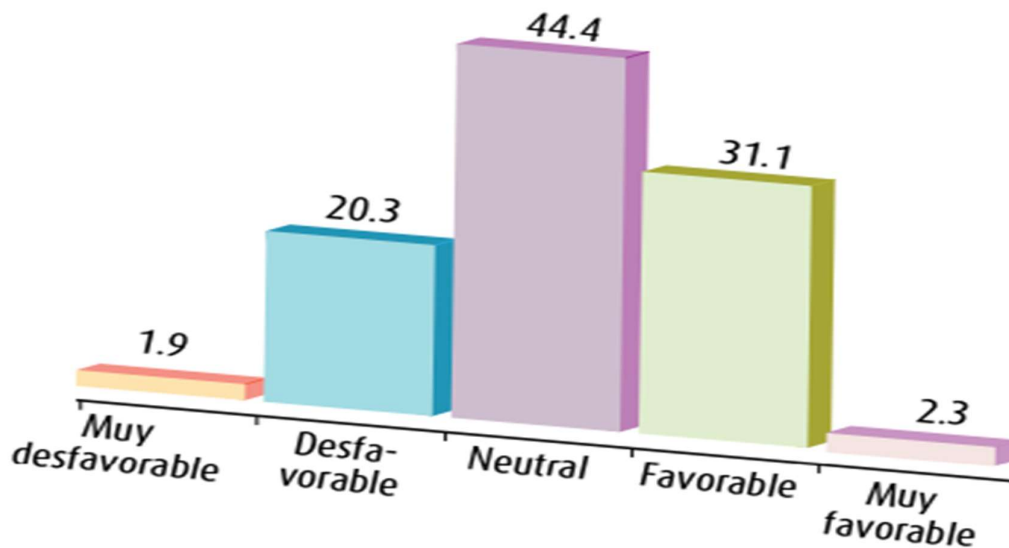


Ilustración 43 – Ejemplos de histograma. Fuente: Elaboración Propia

¿CÓMO ENTENDER LA VARIANZA Y LA DESVIACIÓN ESTÁNDAR?

Conocer los conceptos y aplicación de varianza y desviación estándar son esenciales en un estudio estadístico, se emplean para evaluar la dispersión de los datos. Por ejemplo, en el campo de las ciencias médicas nos ayudan a comprender y analizar información clínica, mejorar el cuidado de los pacientes y tomar decisiones fundamentadas en evidencia. Y así, en las diversas áreas del conocimiento.

¿QUÉ ES LA VARIANZA?

Se emplea en la evaluación de la distribución de los datos en relación con la media. Se obtiene al sumar las diferencias al cuadrado entre cada valor y la media; y luego, dividir este total por el número total de observaciones. La fórmula establecida para calcular la varianza es:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

Nomenclatura:

s^2 = Varianza de la muestra.

x_i = Valor en el conjunto de datos.

\bar{x} = Media de la muestra.

\sum = Sumatoria.

n = Tamaño de la muestra.

EL PROCEDIMIENTO PARA CALCULAR LA VARIANZA ES EL SIGUIENTE:

1. Deducir la media de los datos.
2. Halla la diferencia de cada punto con respecto al valor de la media.
3. Elevar al cuadrado cada valor.
4. Sumar los valores y elevarlos al cuadrado.
5. Dividir esta suma de cuadrados entre $n - 1$ (muestra) o N (población)

¿QUÉ ES LA DESVIACIÓN ESTÁNDAR O DESVIACIÓN DE LA MEDIA?

La raíz cuadrada de la varianza, se conoce como desviación estándar, es utilizada para medir la dispersión de los datos en las mismas unidades que los datos originales. En el caso del tamaño de la muestra, la fórmula para calcular la desviación estándar es:

$$S = \sqrt{s^2}$$

Ejemplo.

Evaluación de tiempos de respuesta en minutos del personal médico para la atención a pacientes en el área de emergencia de un hospital.

| Atenciones | Tiempos de respuesta |
|------------|----------------------|
| 1 | 5' |
| 2 | 7' |
| 3 | 8' |
| 4 | 9' |
| 5 | 12' |

Calculando la Media.

$$\bar{x} = \frac{5 + 7 + 8 + 9 + 12}{5} = 8.2'$$

Calculando la Varianza.

$$s^2 = \frac{(5 - 8.2)^2 + (7 - 8.2)^2 + (8 - 8.2)^2 + (9 - 8.2)^2 + (12 - 8.2)^2}{5 - 1}$$

$$(5 - 8.2)^2 = (-3.2)^2 = 10.24$$

$$(7 - 8.2)^2 = (-1.2)^2 = 1.44$$

$$(8 - 8.2)^2 = (-0.2)^2 = 0.04$$

$$(9 - 8.2)^2 = (-0.8)^2 = 0.64$$

$$(12 - 8.2)^2 = (-3.8)^2 = 14.44$$

$$10.24 + 1.44 + 0.4 + 0.64 + 14.44 = 26.8$$

Dividiendo por n-1; donde en este caso, n=5

$$s^2 = \frac{26.8}{5 - 1} = \frac{26.8}{4} = 6.7$$

Varianza de la muestra.

$$s^2 = 6.7$$

Calculando la Desviación Estándar.

$$S = \sqrt{6.7}$$

Desviación Estándar de la muestra

$$S \approx 2.58$$

En conclusión: el tiempo de respuesta del personal médico para la atención en el área de emergencia de un hospital, tiene una desviación estándar de aproximadamente 2.58 minutos.

¿CUÁL ES LA INTERACCIÓN ENTRE MEDIDAS DE DISPERSIÓN EN EL REPORTE DE DATOS?

La relación entre el rango, la variabilidad y la desviación estándar puede ayudarnos a entender mejor la diversidad de los datos.

- 📊 **El Rango,** da una idea rápida de la amplitud de los datos, aunque puede resultar engañoso si hay valores inusuales.
- 📊 **La Varianza:** Considera todas las observaciones y proporciona una medida más sólida de la dispersión, pero su unidad es el cuadrado de la unidad original, lo que puede complicar su comprensión.
- 📊 **La Desviación Estándar:** Al ser la raíz cuadrada de la varianza, ofrece una medida de dispersión en las mismas unidades que los datos originales, facilitando la comparación directa con la media y otras medidas.

Ejemplo:

Evaluación de los tiempos de espera en minutos para la toma de signos vitales por parte del personal de enfermería, previo a la atención médica.

A) RECOLECCIÓN DE DATOS

| Atenciones | Tiempos de respuesta |
|------------|----------------------|
| 1 | 5' |
| 2 | 7' |
| 3 | 8' |
| 4 | 9' |
| 5 | 12' |

B) MEDIDAS DE TENDENCIA CENTRAL

1. Calculando la Media (promedio).

$$\bar{x} = \frac{5 + 7 + 8 + 9 + 12}{5} = 8.2'$$

En promedio, los pacientes aguardan alrededor de 8.2 minutos antes de que les tomen los signos vitales.

2. Calculando la mediana.

Entendiendo que los datos se encuentran organizados y hay una cantidad impar de observaciones, la mediana correspondiente al valor ubicado en el centro es:

$$\text{Mediana} = 8'$$

Esto es un indicativo que la mitad de los pacientes esperó 8 minutos o menos, y la otra mitad aguardó más de 8 minutos.

3. Moda.

Dado que cada valor se presenta solo una vez, no existe un valor modal en este conjunto de datos.

C) MEDIDAS DE DISPERSIÓN

Las medidas de dispersión son útiles para comprender cómo se distribuyen los tiempos de espera en relación con el promedio.

4. Calculando el rango.

$$\text{Rango} = X_{\max} - X_{\min}$$

$$\text{Rango} = 12 - 5 = 7$$

El intervalo de siete minutos señala una disparidad importante entre la espera mínima de 5 minutos y la máxima de 12 minutos, lo que evidencia una variabilidad en los minutos de espera.

5. Calculando la Varianza.

$$s^2 = \frac{(5 - 8.2)^2 + (7 - 8.2)^2 + (8 - 8.2)^2 + (9 - 8.2)^2 + (12 - 8.2)^2}{5 - 1}$$

$$(5 - 8.2)^2 = (-3.2)^2 = 10.24$$

$$(7 - 8.2)^2 = (-1.2)^2 = 1.44$$

$$(8 - 8.2)^2 = (-0.2)^2 = 0.04$$

$$(9 - 8.2)^2 = (-0.8)^2 = 0.64$$

$$(12 - 8.2)^2 = (-3.8)^2 = 14.44$$

$$10.24 + 1.44 + 0.4 + 0.64 + 14.44 = 26.8$$

a) Dividiendo por $n-1$; donde en este caso, $n=5$

$$s^2 = \frac{26.8}{5 - 1} = \frac{26.8}{4} = 6.7$$

b) Varianza de la muestra.

$$s^2 = 6.7$$

La varianza de la muestra equivale de 6.7 minutos al cuadrado, indica cómo se distribuyen los tiempos de espera en relación con la media. Una mayor varianza sugiere que los tiempos de espera difieren notablemente entre pacientes.

6. Calculando la Desviación Estándar.

$$S = \sqrt{6.7}$$

Desviación Estándar de la muestra

$$S \approx 2.58$$

El cálculo de la desviación estándar indica un tiempo 2.58 minutos, lo que sugiere que generalmente los lapsos de espera varían alrededor de 2.58 minutos por encima o por debajo del promedio de 8.2 minutos. Esto nos proporciona una forma concreta de entender la variabilidad en los tiempos de espera de los pacientes para la toma de signos vitales.

Interpretación de resultados

El intervalo de 7 minutos señala que los tiempos de espera para la toma de signos vitales pueden variar ampliamente, desde tan solo 5 minutos hasta un máximo de 12 minutos. En cuanto a la varianza y desviación estándar, con una varianza de 6.7 minutos al cuadrado y una desviación estándar de 2.58 minutos se indica que hay una dispersión moderada en los tiempos de espera. Esto significa que no se

concentran alrededor de un único valor, sino que se distribuyen a lo largo de un intervalo relativamente amplio. En relación con la media y la mediana, donde la media es de 8.2 minutos y es cercana a la mediana de 8 minutos, esto sugiere que los tiempos de espera no están sesgados y que la mayoría de los pacientes esperan tiempos alrededor de este valor medio.

Conclusión

La evaluación de los tiempos de espera para la toma de signos vitales muestra cierta variabilidad. La mayoría de los pacientes experimentan periodos de espera de alrededor de 8 minutos, sin embargo, otros pacientes pueden esperar más o menos tiempo, como se evidencia en el rango y la desviación estándar. Estos hallazgos podrían ser útiles para identificar áreas de mejora en el proceso de toma de signos vitales, con el objetivo de reducir la variabilidad y garantizar tiempos de espera más consistentes para todos los usuarios.

ESTRATEGIA DE ENSEÑANZA

1: Aprendizaje Basado en Problemas (ABP) y simulación (Role – Playing)

Use casos cotidianos o supuestos para recolección y análisis de datos, ejercitando de manera práctica medidas de tendencia central, esto fortalecerá la identificación de patrones y áreas de mejora.

2: Utilización de Software Estadístico

Mediante el manejo de software estadístico como SPSS o Excel, calcular medidas de tendencia central.

RECURSOS:

- Datos reales o simulados sobre tiempos de respuesta, tasas de infección, etc.
- Software Estadístico: SPSS, Excel u otro software equiparable.
- Guía de aprendizaje y tutoriales de Análisis de Datos Consulting Group Ecuador, para realizar procedimientos y cálculos estadísticos.

EJERCICIO PROPUESTO:

1. ¿Qué representa la media en un conjunto de datos?

- a) El valor más frecuente.

- b) El valor central.
- c) El promedio de todos los valores.
- d) La variabilidad de los valores.

Respuesta: c) El promedio de los valores.

2. ¿Cómo se determina la mediana?

- a) Sumando todos los valores y dividiéndolos por el número total de observaciones.
- b) Encontrando el valor que divide el conjunto de datos en dos partes iguales.
- c) Identificando el valor que se repite con mayor frecuencia.
- d) Calculando la raíz cuadrada de la varianza.

Respuesta: b) Encontrando el valor que divide el conjunto en dos partes iguales.

3. Si las presiones arteriales de los pacientes son 120, 125, 130, 135, 140, ¿cuál es la mediana?

- a) 125
- b) 130
- c) 135
- d) 140

Respuesta: b) 130

REFERENCIAS BIBLIOGRÁFICAS

- Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2018). Estadística para negocios y economía. Cengage Learning.
- Cadena Lozano, J.B. (2024). ESTADÍSTICA DESCRIPTIVA Y PROBABILIDAD PARA LOS NEGOCIOS CON APLICACIONES EN EXCEL Y JAMOVI. CESA.
- Escudero, L. (2006). La Estadística Aplicada a la Empresa. Tomo I: Fundamentos prácticos de la Estadística de Empresa. Boletín de Estudios Económicos. <https://search.proquest.com/openview/1a44cf7f0b9a27a50db4c89afae6cf3f/1?pq-origsite=gscholar&cbl=1818612>
- Grisales Aguirre, A. M. (2019). Estadística descriptiva y probabilidad. Con aplicaciones en Excel y SPSS. Ecoe ediciones.
- González Támara, L. (2013). Estadística descriptiva y probabilidad. Editorial de la Universidad de Bogotá Jorge Tadeo Lozano. Facultad de Ciencias Naturales e Ingeniería. https://www.utadeo.edu.co/files/node/publication/field_attached_file/pdf-_estadistica_descriptiva.pdf
- López, J. (2019). Introducción a la estadística: conceptos y aplicaciones. Ediciones Paraninfo.
- Mesa G., J.A. & Caicedo Z., S.J. (2020). Introducción a la Estadística Descriptiva. Editorial Universitaria de la Universidad de Nariño. <https://sired.udenar.edu.co/6671/1/Introducci%C3%B3n%20a%20la%20Estad%C3%ADstica%20Descriptiva.pdf>
- Rustom J., A. (2012). ESTADÍSTICA DESCRIPTIVA, PROBABILIDAD E INFERENCIA. Una visión conceptual y aplicada. Editorial de la Universidad de Chile. https://repositorio.uchile.cl/bitstream/handle/2250/120284/Rustom_Antonio_Estadistica_descriptiva.pdf
- Barreto-Villanueva, A. (2012). El progreso de la Estadística y su utilidad en la evaluación del desarrollo. Papeles de población, 18(73), 241-271. http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S1405-74252012000300010&lng=es&nrm=iso&tlng=es
- Crespo Antepará, D. N., Reyes, V. M., & Sequera Morales, A. G. (2022). Physical activity and body mass in college students during the "new normal" associated with COVID-19. Research,

- Society and Development, 11(7), e10011729729.
<https://rsdjournal.org/index.php/rsd/article/view/29729>
- Crespo Antepará, D. N., Reyes, V. M., & Sequera Morales, A. G. (2022). La nueva normalidad y los niveles de actividad física y sedentarismo en estudiantes universitarios. *UCV Hacer*, 11(2), 31–38. <https://doi.org/10.18050/revucvhacer.v11i2.2164>
- Fernández Berrocal, P., & Extremera Pacheco, N. E. (2005). La Inteligencia Emocional y la educación de las emociones desde el Modelo de Mayer y Salovey. *Revista Interuniversitaria de Formación del Profesorado*, 19(3), 63–93. <https://www.redalyc.org/articulo.oa?id=27411927005>
- Reyes, V. M., Bustillos, J. K. L., & Morales, A. G. S. (2024). FLIPPED CLASSROOM Y APRENDIZAJE: FACTORES DETERMINANTES DEL APRENDIZAJE EN POSTGRADO. *YACHAQ*, 7(1), Article 1. <https://doi.org/10.46363/yachaq.v7i1.2>
- Reyes, V. M., Luján, V. W. R., Rodríguez, Ó. F. S., Jiménez, J. R. R., Antepará, D. N. C., Mendoza, G. R. G., Morales, A. G. S., Bustillos, J. K. L., Farías, W. B., & Varela, R. E. P. (2023). Student Perspective of Learning in Research Courses in Law Under the Flipped Classroom Modality. *Journal of Law and Sustainable Development*, 11(11), e1441-e1441. <https://doi.org/10.55908/sdgs.v11i11.1441>
- Rodríguez Romero, B., Ramos Gómez, F., & Quintela-del-Río, A. (2020). Formularios Google como recurso educativo. *Contextos universitarios transformadores: boas prácticas no marco dos GID*, 2020, ISBN 978-84-9749-775-6, págs. 575-578, 575-578. <https://dialnet.unirioja.es/servlet/articulo?codigo=7615341>
- Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In *Emotion, disclosure, & health* (pp. 125-154). American Psychological Association. <https://doi.org/10.1037/10182-006>
- Creswell, J. W. (2014). *Research Design; Qualitative, Quantitative and Mixed Methods Approaches* (ed.). SAGE Publications.
- Polit, D. F., & Beck C. T. (2021). *Nursing Research; Generating and Assessing Evidence, for Nursing Practice* (11th ed.). Wolters Kluwer.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (ed.). SAGE Publications.
- Hernández Sampieri, R. Fernández Collado, C., & Baptista Lucio, P. (2014). *Research Methodology* (6th ed.). McGraw Hill Education.



Álvarez Gayou, J. L. (2018). How to Conduct Qualitative Research; Fundamentals and Methodology (3rd ed.). Paidós.

Pagano, R. R. (2013). Understanding Statistics in the Behavioral Sciences (10 th ed.). Cengage Learning.

Gordillo, J.J.. Navarro, I.(2016). Biostatistics, for Health Sciences (2nd ed.). Editorial Síntesis.



*¡Revolucionando
tu manera de **Aprender!***

