



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Facultad de Ingeniería Asignatura Técnicas de Análisis Estadístico

Unidad de aprendizaje 1: Estadística Descriptiva

Introducción

Hola, estamos en la primera unidad de la asignatura denominada: **“Introducción a la estadística descriptiva”**. En esta unidad aprenderemos qué es la estadística, por qué es importante estudiarla y cuáles son sus ramas. Asimismo, se identificarán algunos conceptos importantes de la estadística, los tipos de gráficos utilizados para la presentación de datos, las medidas utilizadas y el coeficiente de correlación.

¿QUÉ ES LA ESTADÍSTICA?



La **estadística** es una disciplina que se encarga de recolectar, organizar, analizar e interpretar datos numéricos para obtener información significativa.

Se utiliza en diversos campos, como la investigación científica, la economía, la medicina y muchas otras áreas.

La **estadística** nos permite resumir grandes cantidades de información en medidas más manejables, como promedios, porcentajes y gráficos.

También nos ayuda a tomar decisiones informadas y a comprender patrones, tendencias y variaciones en los datos. En resumen, la estadística nos ayuda a entender y sacar conclusiones basadas en datos, lo que contribuye al conocimiento y la toma de decisiones en diferentes contextos.

Glosario de términos educativos de © www.proferrecursos.com | Material con derechos de autor



Fuente: <https://www.proferrecursos.com/que-es-la-estadistica/>

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.1 ¿Qué es la estadística?

La estadística puede definirse como una disciplina matemática y científica que se dedica a la recolección, organización, análisis, interpretación y presentación de datos. Su propósito primordial consiste en extraer información de valor a partir de conjuntos de datos, lo que posibilita comprender patrones, tendencias y relaciones en una diversidad de fenómenos. Con aplicaciones que abarcan desde la investigación científica hasta la toma de decisiones empresariales, pasando por la evaluación de políticas públicas y el análisis de resultados experimentales. Antes de abordar los diferentes temas que están relacionados con la estadística, es necesario entender por qué es importante su estudio y cuáles son sus ramas.

1.2 ¿Por qué es importante estudiar la estadística?

Es importante estudiar la estadística porque los datos y las conclusiones que se obtienen a partir de ella y sus diferentes técnicas o métodos, tienen una influencia importante en diversos campos y área de conocimiento, algunas razones son:

- **Toma de Decisiones Informadas:** La estadística proporciona herramientas para analizar y comprender datos complejos, lo que permite tomar decisiones basadas en evidencia en vez de depender de intuiciones o suposiciones.
- **Investigación Científica:** En la investigación, la estadística es crucial para diseñar experimentos adecuados, analizar resultados y extraer conclusiones válidas y confiables.
- **Negocios y Economía:** En el mundo empresarial, la estadística ayuda a identificar tendencias del mercado, pronosticar demanda, tomar decisiones financieras y evaluar el éxito de estrategias comerciales.
- **Medicina y Salud:** En el ámbito de la salud, la estadística es esencial para analizar ensayos clínicos, entender la eficacia de tratamientos, evaluar riesgos y contribuir a la toma de decisiones médicas.
- **Políticas Públicas:** En el gobierno y las políticas públicas, la estadística ayuda a analizar datos demográficos, evaluar impactos de programas sociales y fundamentar decisiones en áreas como la educación y la salud pública.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

- **Investigación Social:** En sociología, psicología y otras disciplinas sociales, la estadística permite analizar datos de encuestas, estudios de población y experimentos para comprender comportamientos y tendencias.
- **Avances Tecnológicos:** En la era de los datos, la estadística es crucial para la inteligencia artificial, el aprendizaje automático y la minería de datos, impulsando avances tecnológicos en diversas áreas.
- **Comprensión del Mundo:** La estadística ayuda a comprender la variabilidad inherente en muchos fenómenos y a hacer generalizaciones precisas a partir de muestras representativas.
- **Comunicación Efectiva:** Saber interpretar y presentar datos estadísticos de manera clara y precisa es crucial para comunicar hallazgos de manera efectiva a audiencias diversas.
- **Pensamiento Crítico:** El estudio de la estadística fomenta el pensamiento crítico al analizar y evaluar la validez de argumentos basados en datos, evitando caer en conclusiones erróneas.

Estudiar estadística brinda habilidades analíticas valiosas y una perspectiva sólida para abordar problemas complejos en una amplia gama de campos, ayudando a tomar decisiones informadas y respaldadas por datos en un mundo cada vez más orientado a la información.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.3 Ramas de la estadística

La estadística se divide fundamentalmente en dos ramas: la **estadística descriptiva**, enfocada en resumir y visualizar datos; y la **estadística inferencial**, que se centra en realizar inferencias y deducciones sobre poblaciones más amplias basándose en muestras de datos.

Estadística Descriptiva

Los Estadística Descriptiva define un papel fundamental en el análisis exploratorio de los datos, permite identificar la estructura y características de los datos, para su representación con gráficos que contribuyen a minimizar la pérdida de información. Algunos expertos recomiendan que este sea el primer paso en el análisis de datos, porque se pueden identificar registros erróneos o no apropiados que requieren de su respectivo análisis, para evitar que sean procesados de manera inadecuada.

Video Resumen: Estadística descriptiva

<https://www.youtube.com/watch?v=buKDQZG5phM>

Estadística Inferencial

La estadística inferencial está relacionada con métodos que permiten generar predicciones a partir de características o fenómenos, para tomar decisiones sobre una población a partir de muestras de datos observados. Su enfoque se basa en la generalización de los resultados obtenidos de una muestra a toda la población, utilizando métodos probabilísticos y técnicas de estimación e inferencia.

En el proceso se utilizan pruebas de hipótesis, intervalos de confianza, análisis de regresión y otros métodos que permiten tomar decisiones informadas en diversos campos, desde la investigación científica hasta la toma de decisiones en negocios.

En esta unidad de aprendizaje trabajaremos la estadística descriptiva, que se relaciona con los eventos observados y las propiedades que conforman los datos. Asimismo, utilizaremos algunos conceptos que permiten identificar las características más relevantes de nuestro conjunto de datos, para la toma de decisiones en diferentes áreas de conocimiento.

Video Resumen: Estadística Inferencial

https://www.youtube.com/watch?v=N_Bnk9Wq7E4

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

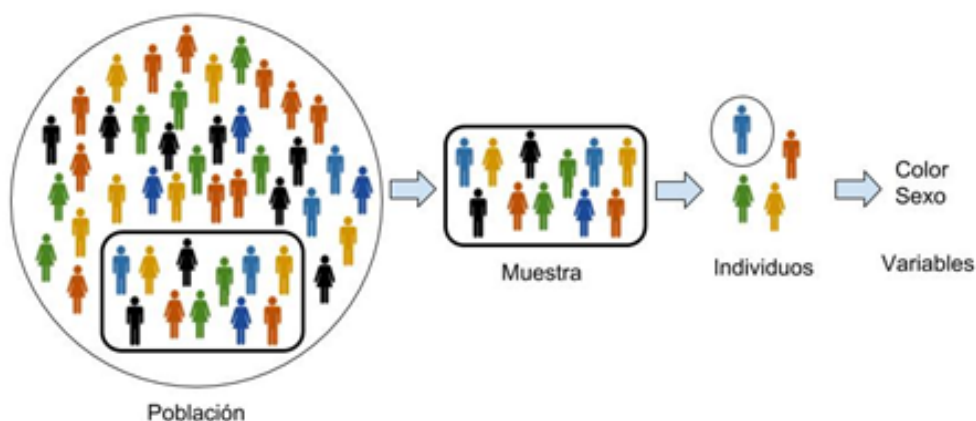
INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.4 Población, muestra y variables

En estadística llamamos **población** al grupo de sujetos o elementos que son de interés para el estudio, tienen características similares para el análisis y permiten generar información. El número de elementos o tamaño de la población generalmente se representa con la letra N.

En algunas ocasiones el tamaño de la población resulta ser demasiado grande para el análisis, por este motivo es necesario seleccionar una pequeña parte de los sujetos, a esta porción se le llama **muestra** y en términos generales debe reflejar el comportamiento de la población que se desea estudiar.

A los sujetos de la población o muestra se le observan características que hacen parte del estudio, pero estas pueden ser de diferente naturaleza y por ende pueden tomar diferentes valores, a esto se le denomina **variable**. En otras palabras, una variable es una característica que puede tener diferentes valores, por ejemplo, edad, género y peso. Las variables pueden ser de dos tipos, la primera se denominada cuantitativa, aquí las características se pueden contar o medir. El segundo tipo se denomina cualitativa y se caracteriza porque la característica se puede clasificar en categorías.



Fuente: <https://www.datalaria.com/post/fundamentos/2018-10-07-estadistica-descriptiva/>

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



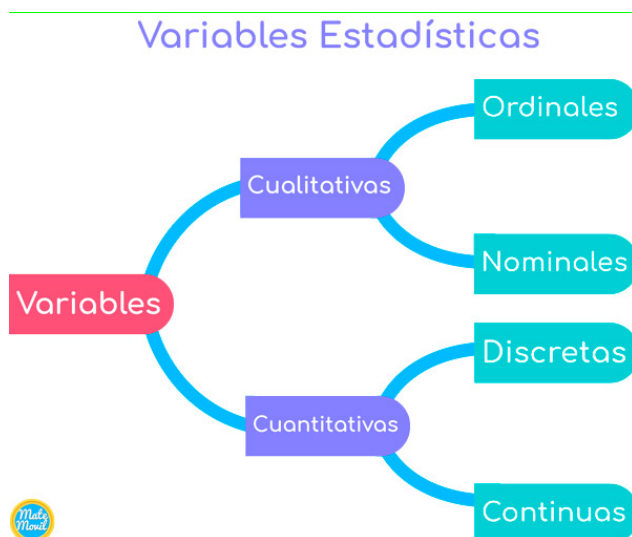
CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.5 Tipos de datos

Las variables pueden tener diferentes tipos de datos según su naturaleza o tipo de escala. Las variables de tipo cualitativo pueden ser ordinales o nominales, mientras que las cuantitativas pueden ser discretas o continuas.



Fuente: <https://matemovil.com/variables-cualitativas-y-cuantitativas/>

Variables Ordinales

Una variable ordinal permite identificar características que se clasifican en grupos con un respectivo orden. Representan valores o categorías que tienen un orden o una jerarquía, pero la distancia que existe entre los valores no es necesariamente cuantificable o uniforme. Esto quiere decir, que tienen una relación, pero la diferencia entre los valores no es significativa ni constante. Algunos ejemplos de este tipo de variables pueden ser: resultados en una prueba de natación (medalla de oro, medalla de plata o medalla de bronce); satisfacción en un servicio (muy satisfecho, satisfecho, ni satisfecho ni insatisfecho, insatisfecho y muy insatisfecho).

Las variables ordinales son comunes en diversas áreas, como en escalas de calificación, niveles de acuerdo o desacuerdo, rangos socioeconómicos y clasificaciones de nivel educativo. En el análisis estadístico, las variables ordinales se manejan con métodos específicos que respetan su naturaleza jerárquica sin hacer suposiciones sobre las distancias absolutas entre las categorías.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Variables Nominales

En las variables nominales las características se pueden categorizar en grupos sin ningún orden. Las categorías representan diferentes valores, pero no existe una relación ni numérica ni jerárquica entre ellas. En otras palabras, las variables nominales simplemente etiquetan o identifican diferentes grupos, sin indicar ninguna relación de magnitud, orden o distancia entre ellos. Por ejemplo, género (masculino, femenino, otros), estado civil (soltero, casado, divorciado).

Cuando se trabaja con variables nominales en estadística, se utilizan técnicas específicas para describir y analizar la frecuencia de ocurrencia de cada categoría y para identificar posibles relaciones o asociaciones entre diferentes variables nominales. Estas técnicas a menudo involucran tablas de contingencia, pruebas y herramientas que están diseñadas para evaluar la dependencia entre las variables.

Variables Discretas

Este tipo de variables generalmente representan conteos, toman un número finito de valores y se representan con números enteros que no se pueden dividir en partes más pequeñas o fracciones dentro del contexto. Algunos ejemplos de este tipo de variable pueden ser: número de hijos (0, 1, 2, ...), cantidad de autos estacionados en un momento dado en un parqueadero, número de libros en un estante.

En análisis estadístico, las variables discretas se manejan con técnicas específicas para describir su distribución de frecuencia, calcular medidas de resumen y analizar patrones de ocurrencia. Histogramas, gráficos de barras y cálculos de probabilidad discreta son herramientas comunes utilizadas para comprender y analizar este tipo de variables.

Variables Continuas

Las variables continuas son un tipo de variable que pueden tomar un rango infinito de valores dentro de un intervalo específico. A diferencia de las variables discretas, las variables continuas pueden asumir valores en cualquier punto dentro de un intervalo, incluso valores fraccionales o decimales. Estas variables se caracterizan por en su naturaleza existe una cantidad infinita de posibles valores entre cualquier par de valores.

Ejemplos comunes de este tipo de variables pueden ser: altura de una persona (1.65, 1.80, 1.73), el peso, la temperatura. En las variables continuas, se utilizan técnicas específicas para describir su distribución, calcular medidas de tendencia central y dispersión, y analizar patrones de comportamiento. Histogramas, gráficos de densidad y cálculos de probabilidad continua son herramientas comunes para abordar este tipo de variables y comprender sus propiedades estadísticas.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

En resumen, tenemos:

Los valores que toma la variable:		
Cualitativa	Ordinal	Se pueden clasificar en grupos ordenados.
	Nominal	Se pueden clasificar en grupos, pero no hay un orden entre ellos.
Cuantitativa	Discreta	Son finitos o contables.
	Continua	Son infinitos en cualquier intervalo razonable.

Fuente: Libro Estadística Descriptiva y Probabilidad de Leandro González Támara (página 12)

En estadística también existen otros tipos de datos, estos son: porcentajes, escalas analógicas visuales y score. A continuación, se describen:

Porcentajes

Los datos en forma de porcentaje representan una forma de expresar proporciones con relación a 100 unidades. Los valores en porcentaje son una manera de comparar diferentes partes o categorías dentro de un conjunto en términos relativos. Los datos en forma de porcentaje son útiles cuando se desea comparar partes relativas dentro de un conjunto y cuando se busca comunicar información de manera más intuitiva y comprensible. En estadística los porcentajes se utilizan frecuentemente en gráficos circulares, barras y tablas para resaltar la distribución relativa de diferentes categorías.

En ocasiones se piensa que los porcentajes son variables continuas, sin embargo, pueden causar problemas en el análisis, cuando se toman valores mayores y menores al 100% o en valores negativos.

Escalas Analógicas Visuales

Las escalas analógicas visuales son una forma de medir y representar datos que utiliza una línea continua o un gráfico para indicar la magnitud de una característica específica. Estas escalas permiten a las personas hacer estimaciones o valoraciones visuales basadas en la posición en la línea o en el gráfico, en lugar de usar valores numéricos exactos.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



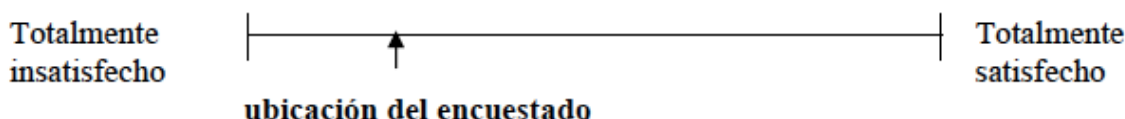
CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Estas escalas son particularmente útiles cuando se necesita obtener respuestas rápidas y aproximadas, y no es necesario medir con precisión valores numéricos específicos. A menudo, se utilizan en encuestas y evaluaciones subjetivas donde se busca capturar percepciones de manera intuitiva.

Un ejemplo común de este tipo de datos es el grado de satisfacción con un tratamiento, en donde se utiliza este tipo de escala.



Fuente: Libro Estadística Descriptiva de Liliana Orellana (página 7)

Scores

En el contexto de la estadística un "score" se refiere a una puntuación o valor asignado a un individuo o elemento específico en función de ciertas medidas o criterios. Los scores son utilizados para cuantificar o calificar características, rendimientos o evaluaciones en una escala numérica.

Los scores pueden variar en su naturaleza y propósito. Algunos ejemplos comunes incluyen:

- **Puntajes de Evaluación:** En exámenes, pruebas o evaluaciones educativas, los scores se utilizan para medir el rendimiento académico. Por ejemplo, una puntuación en una prueba podría indicar la cantidad de respuestas correctas.
- **Calificaciones de Riesgo:** En análisis financieros o de crédito, los scores se emplean para evaluar el riesgo crediticio de una persona o empresa. Esto ayuda a los prestamistas a tomar decisiones informadas sobre préstamos.
- **Evaluaciones Subjetivas:** En encuestas de satisfacción o evaluaciones de productos, los scores pueden reflejar la opinión o nivel de satisfacción de los participantes en una escala, como de 1 a 5.
- **Indicadores de Rendimiento:** En deportes, juegos y competencias, los scores miden el desempeño y el resultado de los participantes, como puntos anotados en un partido de fútbol.
- **Puntuaciones de Salud:** En escalas de medición de salud, como el Índice de Masa Corporal (IMC), se utilizan scores numéricos para evaluar el estado de salud o la condición física.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.6 Gráficos para datos cualitativos y cuantitativos

La estadística descriptiva permite generar diferentes gráficas, con el fin de presentar los datos o características más relevantes. En este apartado vamos a presentar los tipos de gráficas más utilizadas para datos cualitativos y cuantitativos.

Los gráficos más utilizados para representar datos cualitativos son: barras y pastel. Mientras que los gráficos utilizados para datos cuantitativos son: líneas, histograma, puntos y diagrama de caja.

Gráfico de Barras

Una gráfica de barras es una representación visual que utiliza barras rectangulares para comparar diferentes categorías o valores. Cada barra representa una categoría específica y su altura está proporcionalmente relacionada con el valor que representa. Este tipo de gráfico es ampliamente utilizado para visualizar datos categóricos o discretos, como frecuencias, recuentos o porcentajes.

En esta gráfica el eje horizontal muestra las categorías o etiquetas que se están comparando y el eje vertical representa la magnitud o frecuencia del valor en cada categoría. El gráfico de barra se utiliza cuando se quiere mostrar una comparación de forma clara y fácil de entender entre diferentes categorías o cuando se busca resaltar diferencias significativas entre valores. Pueden ser utilizadas para representar datos simples, como conteos de elementos en categorías, o para datos más complejos como porcentajes.

Ejemplo en Python

El gráfico de barras que se presenta a continuación, muestra el precio promedio de la gasolina por estado, en el eje X se encuentran los estados: Hawaii, Florida, Missouri, Washington y California, mientras que en el eje Y está el precio promedio. Se puede observar que el estado con el precio promedio más alto es Hawaii, con un valor de 3.6. Sin embargo, Missouri resulta tener el precio promedio más bajo con un valor de 3.0.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

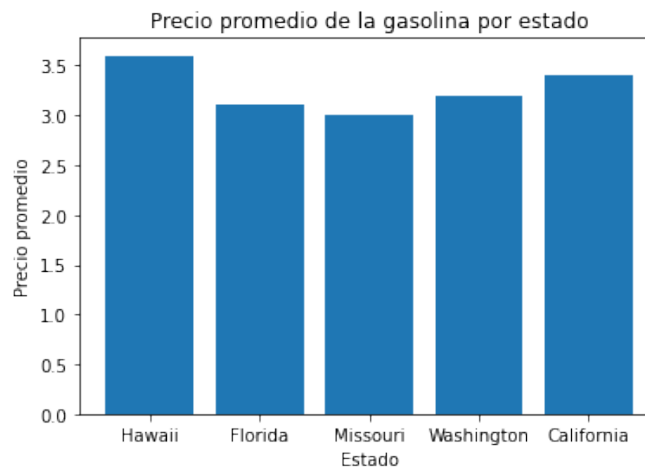
CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 x = ['Hawaii', 'Florida', 'Missouri', 'Washington', 'California']
7 y = [3.6, 3.1, 3.0, 3.2, 3.4]
8
9 # Paso 3: Generamos la gráfica
10 plt.bar(x, y)
11 plt.title("Precio promedio de la gasolina por estado")
12 plt.xlabel("Estado")
13 plt.ylabel("Precio promedio")
14 plt.show()
```

[2] ✓ 1.5s



Los datos de este ejemplo se tomaron de:

<https://flexbooks.ck12.org/cbook/ck-12-conceptos-de-matem%C3%A1ticas-de-la-escuela-secundaria-grado-8-en-espa%C3%B1ol/section/1.1/primary/lesson/interpretaci%C3%B3n-de-gr%C3%A1ficos-de-barras/>

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Gráfico de Pastel

Una gráfica de pastel o circular es un tipo de representación visual utilizada para mostrar la proporción o distribución de diferentes categorías en relación con un total. Este tipo de gráfico utiliza una forma circular dividida en segmentos, donde cada segmento representa una categoría en el conjunto de datos. Este tipo de gráfica se utiliza cuando se desea resaltar la contribución relativa de cada categoría en un conjunto y son efectivas para visualizar proporciones de manera intuitiva y fácil de comprender.

Ejemplo en Python

Se realizó una encuesta a diferentes usuarios que utilizan una página de contenidos, para identificar cuál es el tipo de contenido más visitado. Las opciones de contenido son: deportes, cultura, sociedad, tecnología y otros. Después de aplicar el instrumento los resultados indican que el contenido más visitado en la página fue deportes con un 22.1%, mientras que el menos visitado fue tecnología con un valor del 17.9%.

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 x = ['Deportes', 'Cultura', 'Sociedad', 'Tecnología', 'Otros']
7 y = [22.1, 20.3, 19.9, 17.9, 19.8]
8
9 # Paso 3: Generamos la gráfica de Pie
10 plt.pie(y, labels=x, autopct="%0.1f %%")
11 plt.title("Visitas a Contenidos")
12 plt.show()
```

[6] ✓ 0.3s

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828



Los datos de este ejemplo se tomaron de:

Los datos de este ejemplo se tomaron de:

<https://www.merkle.com/es/es/blog/representacion-datos-graficas-tarta>

Gráfico de Líneas

Una gráfica de líneas es una representación que utiliza líneas para conectar puntos de datos y mostrar cómo cambian los valores a lo largo de un eje, generalmente el tiempo u otra variable continua. Este tipo de gráfico es especialmente útil para visualizar y analizar tendencias, patrones y fluctuaciones en datos numéricos a lo largo de una secuencia de tiempo u otra variable independiente.

En una gráfica de líneas el eje horizontal generalmente representa la variable independiente y en el eje vertical representa la variable dependiente, es decir, los valores que se están observando y registrando. Cada punto en el gráfico de líneas representa una observación en un momento específico, las líneas conectan estos puntos para mostrar la relación y la dirección del cambio a lo largo de la variable independiente.

Las gráficas de líneas son muy utilizadas en campos como la economía, la ciencia, la salud y el análisis financiero para visualizar tendencias y patrones a lo largo del tiempo. También son útiles para mostrar la relación entre dos variables continuas y cómo interactúan entre sí.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"

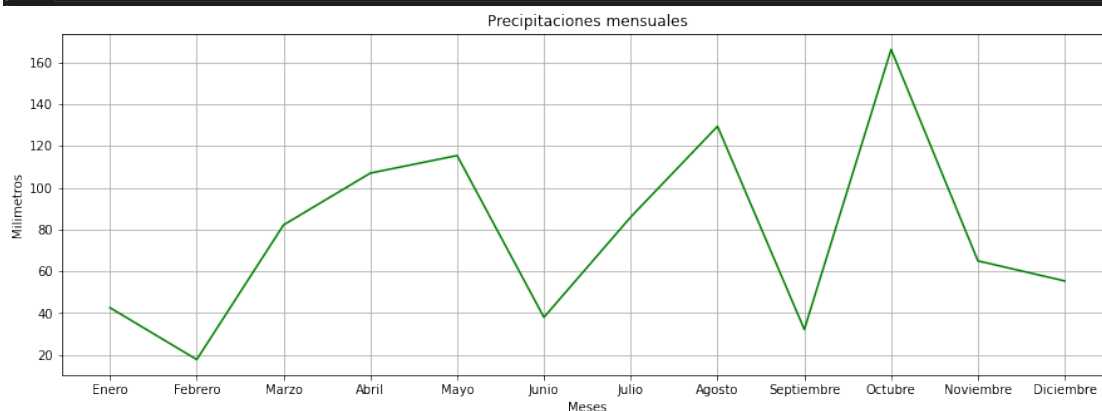


Ejemplo en Python

Se solicitó a un grupo de estudiantes que realizará un gráfico de líneas para representar las precipitaciones mensuales de lluvia en un país durante el año 2022. En el eje X se muestran los meses del año y en el eje Y la precipitación mensual en Milímetros. Se puede observar que el mes con más lluvias fue octubre con una precipitación de 166 milímetros (aproximadamente). Asimismo, se puede observar que el mes donde menos hubo lluvias fue febrero con un valor de 18 milímetros en la precipitación (aproximadamente).

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 x = ['Enero', 'Febrero', 'Marzo', 'Abril', 'Mayo', 'Junio', 'Julio', 'Agosto', 'Septiembre', 'Octubre', 'Noviembre', 'Diciembre']
7 y = [42.6, 17.8, 82.2, 107.0, 115.4, 38.0, 86.0, 129.4, 32.2, 166.2, 65.0, 55.4]
8
9 # Generar la gráfica de Línea
10 plt.figure(figsize=(15, 5))
11 plt.plot(x, y, color="green", linestyle="--")
12 plt.title("Precipitaciones mensuales")
13 plt.xlabel("Meses")
14 plt.ylabel("Milímetros")
15 plt.grid()
16 plt.show()
```



Los datos de este ejemplo se tomaron de:
<https://ordenadorpractico.es/mod/assign/view.php?id=274>



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Histograma

Los histogramas se utilizan para visualizar la distribución de datos numéricos continuos o discretos en intervalos. Este tipo de gráfico permite comprender cómo se distribuyen los valores dentro de un conjunto y proporciona información sobre la frecuencia o densidad de observaciones en diferentes rangos. En este tipo de gráficas, el eje horizontal se divide en intervalos, que son segmentos contiguos en el rango de los datos y el eje vertical representa la frecuencia o densidad de observaciones dentro de cada intervalo.

Cada barra en el histograma representa un intervalo y su altura está relacionada con la cantidad de observaciones en ese intervalo. Cuanto más alta sea la barra, mayor es la concentración de valores en ese rango. Este tipo de gráficas se utilizan para identificar patrones de distribución, como simetría, asimetría, concentraciones o dispersión de datos. También permiten detectar posibles valores atípicos o picos en la distribución.

Ejemplo en Python

En la imagen se presentan algunas muestras de los niveles de presión sonora generados en decibeles (dB), para algunas estaciones de ruido ambiental de la ciudad de Cali (Colombia) en el año 2007 en horario diurno. Con base a estos datos se solicita construir un histograma.

63.7	75.0	70.5	72.1	67.2	65.1	59.6	64.1	61.1	62.0
66.9	76.3	73.7	74.1	62.3	55.3	70.6	53.3	65.9	64.0
66.8	71.4	71.0	76.5	69.4	71.3	65.3	62.5	62.6	58.7
75.3	77.4	56.1	57.3	60.5	72.3	74.0	62.3	50.2	68.2
70.8	71.6	69.0	71.6	75.0	64.6	74.9	75.4	50.9	61.6

Fuente: Libro Tratamiento Estadístico de Datos con Aplicaciones en R de Luis Alberto Díaz Chávez y Jairo Rafael Rosado Vega (página 29)

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



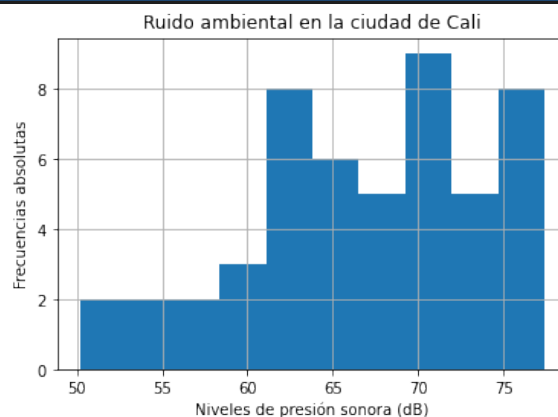
CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 presion_sonora = [63.7, 75.0, 70.5, 72.1, 67.2, 65.1, 59.6, 64.1, 61.1, 62.0,
7                  66.9, 76.3, 73.7, 74.1, 62.3, 55.3, 70.6, 53.3, 65.9, 64.0,
8                  66.8, 71.4, 71.0, 76.5, 69.4, 71.3, 65.3, 62.5, 62.6, 58.7,
9                  75.3, 77.4, 56.1, 57.3, 60.5, 72.3, 74.0, 62.3, 50.2, 68.2,
10                 70.8, 71.6, 69.0, 71.6, 75.0, 64.6, 74.9, 75.4, 50.9, 61.6]
11
12 # Generar la gráfica de Histograma
13 plt.hist(x=presion_sonora)
14 plt.title('Ruido ambiental en la ciudad de Cali')
15 plt.xlabel('Niveles de presión sonora (dB)')
16 plt.ylabel('Frecuencias absolutas')
17 plt.grid()
18 plt.show()
```



En la gráfica se observa que al agrupar los valores se definen los intervalos que se encuentran en el eje X denominados niveles de presión sonora en dB y en el eje Y se presentan las frecuencias para cada uno de los niveles definidos. Visualmente se puede observar que entre el intervalo 50 – 55 la frecuencia del ruido es de aproximadamente 2, pero en el intervalo entre 70 – 75 esta frecuencia aumenta a 9 aproximadamente.

Los datos de este ejemplo se tomaron de:
Tratamiento Estadístico de Datos con Aplicaciones en R de Luis Alberto Díaz Chávez y Jairo Rafael Rosado Vega (página 29 y 44)



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Gráfico de Puntos (Scatter Plot)

Una gráfica de puntos, también conocida como gráfico de dispersión o scatter plot en inglés, es una representación que utiliza puntos individuales para mostrar la relación entre dos variables numéricas. Cada punto en la gráfica de puntos representa una observación en el conjunto de datos y se coloca en un plano cartesiano con ejes horizontal y vertical que representan las dos variables en estudio.

En este tipo de gráfica el eje horizontal representa una variable independiente y el eje vertical representa la variable dependiente. Cada punto en la gráfica de puntos se coloca de acuerdo con las coordenadas que corresponden a los valores de las dos variables para esa observación. La posición de los puntos en el plano revela la relación entre las variables, como si hay alguna correlación, tendencia o agrupamiento.

Los gráficos de puntos son especialmente útiles cuando se desea explorar la relación entre dos variables y entender cómo cambian juntas. Pueden ayudar a identificar patrones, tendencias, valores atípicos y proporcionar una visión general de la distribución de los datos. También, se utilizan ampliamente en diversos campos, desde estadísticas y ciencia hasta economía y ciencias sociales, para analizar la asociación entre variables y hacer observaciones visuales sobre sus interacciones.

Ejemplo en Python

Se recopiló el salario y los gastos mensuales de 30 personas con edad de 35 años, para analizar si entre las variables existe una relación. Con los datos se solicita realizar una gráfica de puntos, en el eje X debe ir el salario mensual y en el eje Y los gastos mensuales.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

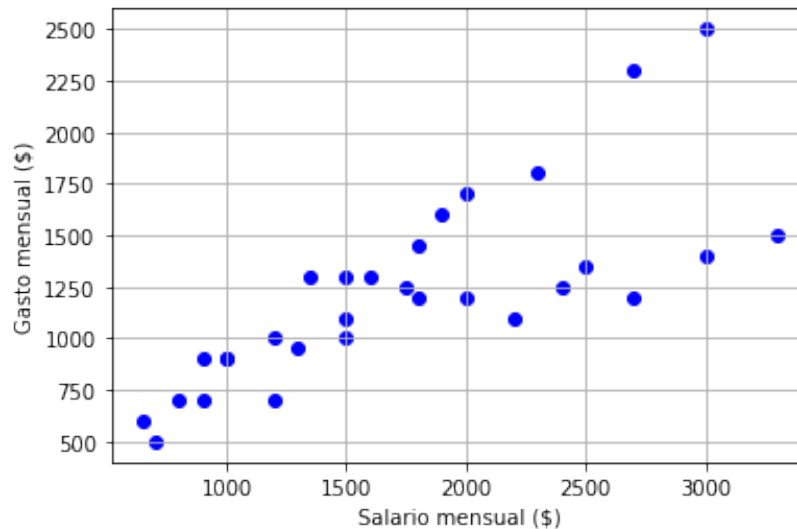
CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 salarios = [1500, 1200, 650, 1000, 3000, 2300, 800, 1800, 700, 2000, 900, 1350, 2700, 1900, 1750,
7             1500, 2000, 2500, 900, 1300, 2400, 1600, 3300, 2200, 3000, 1000, 1500, 2700, 1800, 1200]
8 gastos = [1300, 700, 600, 900, 2500, 1800, 700, 1450, 500, 1700, 700, 1300, 2300, 1600, 1250,
9           1000, 1200, 1350, 900, 950, 1250, 1300, 1500, 1100, 1400, 900, 1100, 1200, 1200, 1000]
10
11 # Generar la gráfica de Puntos
12 plt.scatter(salarios, gastos, label="", color="blue")
13 plt.xlabel('Salario mensual ($)')
14 plt.ylabel('Gasto mensual ($)')
15 plt.grid()
16 plt.show()
```

[1] ✓ 2.3s



A simple vista se puede observar que las variables tienen una relación lineal, donde el salario mensual aumenta y los gastos también. También se puede apreciar que el salario más bajo es de 650, con un gasto de 600. Mientras que el salario más alto es de 3300 con un gasto de 1500.

Los datos de este ejemplo se tomaron de:

https://www.probabilidadyestadistica.net/grafica-de-puntos/#google_vignette

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Diagrama de Caja (Boxplot)

Un diagrama de caja, también conocido como diagrama de caja y bigotes o box plot en inglés, es una representación utilizada para mostrar la distribución y la dispersión de un conjunto de datos numéricos. Este tipo de gráfico proporciona información sobre los valores centrales, los valores atípicos y la variabilidad en los datos.

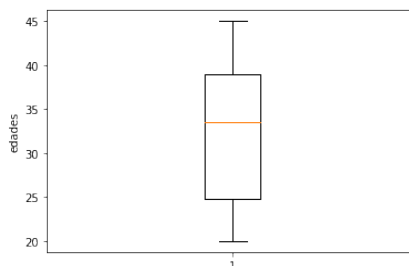
En un diagrama de caja, una caja rectangular se dibuja a lo largo de un eje, generalmente vertical, y se extiende desde el primer cuartil (Q1) al tercer cuartil (Q3) del conjunto de datos. Una línea en el interior de la caja representa la mediana, que es el valor que divide los datos en dos partes iguales y líneas o "bigotes" se extienden desde los bordes de la caja hasta los valores más extremos dentro de un rango específico. A menudo, se consideran valores atípicos aquellos que se encuentran más allá de los bigotes.

Los diagramas de caja son útiles para identificar la dispersión y la concentración de los datos, así como para detectar valores atípicos que pueden estar fuera del rango típico. Estos gráficos proporcionan una imagen rápida de la distribución de los datos sin necesidad de cálculos detallados y se utilizan en ciencia de datos, investigación científica y análisis financiero para resumir y visualizar propiedades clave de un conjunto de datos, facilitando la comparación y la identificación de patrones.

Ejemplo en Python

Se recolecta la edad de 20 personas y con ello se solicita generar una gráfica de caja para ver la distribución de los datos. Escriba y ejecute el siguiente código Python en un Notebook:

Esta gráfica permite identificar la distribución de las edades, en el primer cuartil (Q1) tenemos el valor de 24.5; en la mediana su valor es 33.5 y en tercer cuartil (Q3) su valor es 39.



```
1 # Paso 1: Cargar las librerías
2 # Librería para gráficas
3 import matplotlib.pyplot as plt
4
5 # Paso 2: Cargamos los datos
6 datos = [36, 25, 37, 24, 39, 20, 36, 45, 31, 31,
7          39, 24, 29, 23, 41, 40, 33, 24, 34, 40]
8
9 # Generar la gráfica de Caja
10 plt.boxplot(datos)
11 plt.ylabel('edades')
12 plt.show()
```

Los datos de este ejemplo se tomaron de:

<https://www.estadisticaparatodos.es/taller/graficas/cajas.html>

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

1.7 Medidas

Media

La media, también conocida como promedio, es una medida estadística utilizada para calcular el valor central de un conjunto de datos numéricos. Se calcula sumando todos los valores en el conjunto y luego dividiendo esta suma por la cantidad de valores presentes. La media es una medida de tendencia central que busca representar un valor típico o característico del conjunto.

La fórmula para calcular la media es:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

Fuente: Libro Tratamiento Estadístico de Datos con Aplicaciones en R de Luis Alberto Díaz Chávez y Jairo Rafael Rosado Vega (página 45)

Donde:

- \bar{X} es la media.
- X_1, X_2, \dots, X_n son los valores en el conjunto de datos.
- n es la cantidad de valores en el conjunto.

La media es ampliamente utilizada para resumir un conjunto de datos y proporcionar una medida de su valor central. Sin embargo, puede verse afectada por valores atípicos o extremos, lo que puede distorsionar su interpretación. En algunos casos, otras medidas de tendencia central, como la mediana o la moda, pueden ser más apropiadas si se busca evitar la influencia de valores atípicos en la representación del conjunto de datos.

Ejemplo:

Imagine que está analizando las edades de un grupo de personas en una muestra. Tenemos las edades de 10 personas: 25, 32, 28, 45, 19, 40, 36, 22, 30 y 27 años.

Para calcular la media de estas edades, sumamos todos los valores y luego dividimos entre la cantidad de valores:

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

$$\bar{X} = (25 + 32 + 28 + 45 + 19 + 40 + 36 + 22 + 30 + 27) / 10 = 30.4$$

La media de las edades en este caso es 30.4 años. Esto significa que, en promedio, la edad de las personas en esta muestra es de 30.4 años. La media es una forma de resumir los datos y proporciona una idea general del valor central de las edades en el grupo.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import numpy as np
4
5 # Paso 2: Cargar los datos
6 datos = [25, 32, 28, 45, 19, 40, 36, 22, 30, 27]
7
8 # Calcular la Media
9 media = np.mean(datos)
10 print("Media = ", media)
```

[1] ✓ 0.4s

... Media = 30.4

Mediana

La mediana es una medida estadística utilizada para representar el valor central de un conjunto de datos. Es el valor que se encuentra exactamente en el medio de un conjunto ordenado de datos, de manera que la mitad de los valores se encuentran por encima de la mediana y la otra mitad por debajo. La mediana es una medida de tendencia central que es menos sensible a los valores atípicos o extremos que la media.

Para calcular la mediana en un conjunto de datos:

- Ordena los valores de menor a mayor.
- Si la cantidad de valores es impar, la mediana es el valor que se encuentra

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

en el centro del conjunto ordenado.

- Si la cantidad de valores es par, la mediana es el promedio de los dos valores centrales.

Esta medida se utiliza cuando se trabaja con conjuntos de datos que contienen valores atípicos o cuando la distribución de los datos no es simétrica. A diferencia de la media, la mediana no se ve afectada por valores extremos, lo que la convierte en una medida robusta de la tendencia central.

Ejemplo:

Supongamos que tenemos un conjunto de puntuaciones en un examen de diez estudiantes: **85, 92, 78, 60, 95, 87, 75, 82, 90 y 88**. Donde 50 es la calificación más baja y 100 la más alta.

Para encontrar la mediana, primero ordenamos las puntuaciones de menor a mayor:

60, 75, 78, 82, 85, 87, 88, 90, 92, 95

Dado que tenemos un total de 10 puntuaciones, la mediana estará en la posición central, es decir, entre la quinta y sexta posición (para este caso). La puntuación está entre 85 y 87. Por lo tanto, la mediana de este conjunto de puntuaciones es 86.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import numpy as np
4
5 # Paso 2: Cargar los datos
6 datos = [85, 92, 78, 60, 95, 87, 75, 82, 90, 88]
7
8 # Paso 3: Calcular la Mediana
9 mediana = np.median(datos)
10 print("Mediana = ", mediana)
```

[1] ✓ 0.4s

... Mediana = 86.0

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Moda

La moda es una medida estadística que representa el valor o los valores en un conjunto de datos que ocurren con mayor frecuencia. En otras palabras, la moda es el valor que aparece con mayor frecuencia en un conjunto de datos. Si hay más de un valor que se repite con la misma frecuencia máxima, el conjunto se considera multimodal y tiene múltiples modas.

La moda es especialmente útil para identificar los valores más comunes o típicos en un conjunto de datos. A diferencia de la media y la mediana, la moda no se ve influenciada por los valores numéricos de los datos, sino por la frecuencia de ocurrencia de esos valores. Puede ser una medida útil en situaciones donde se busca identificar los valores más frecuentes en una distribución.

Ejemplo

Imagine que estamos observando las alturas de los participantes en una competencia deportiva. Tenemos las siguientes alturas en centímetros: 165, 172, **160**, 172, 175, **160**, 168, 178, **160** y 175.

En este conjunto de alturas, podemos ver que la altura 160 cm aparece tres veces, mientras que las otras alturas aparecen una o dos veces. La moda es el valor que aparece con mayor frecuencia en un conjunto de datos. En este caso, la moda sería 160 cm, ya que es la altura que se repite más veces en el conjunto.

La moda nos indica cuál es el valor más común en el conjunto de datos. En este ejemplo, 160 cm es la altura que más participantes tienen, lo que puede ser útil para comprender la distribución de alturas en el grupo.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import statistics as stat
4
5 # Paso 2: Cargar los datos
6 datos = [165, 172, 160, 172, 175, 160, 168, 178, 160, 175]
7
8 # Paso 3: Calcular la Moda
9 moda = stat.mode(datos)
10 print("Moda = ", moda)
```

[3] ✓ 0.0s

... Moda = 160

- 📍 Sede Quirinal: Calle 21 No. 6 - 01
- 📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220
- 📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699
- ✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA
"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Percentiles

El percentil es una medida estadística utilizada para describir la posición relativa de un valor específico dentro de un conjunto de datos ordenados. Representa el porcentaje de valores que son iguales o menores que el valor en cuestión. En otras palabras, el percentil indica la proporción de valores que se encuentran por debajo de cierto valor dado en un conjunto de datos.

Los percentiles son especialmente útiles para comprender cómo se ubica un valor en relación con el resto de los valores del conjunto. Por ejemplo, el percentil 25 (también conocido como primer cuartil) indica que el 25% de los valores en el conjunto son iguales o menores que ese valor.

Algunos percentiles comunes incluyen:

- **Primer cuartil (Q1):** El valor que deja el 25% de los datos por encima.
- **Mediana (Q2 o percentil 50):** El valor que divide el conjunto en dos partes iguales.
- **Tercer cuartil (Q3):** El valor que deja el 75% de los datos por debajo.

Los percentiles pueden ser útiles para identificar valores atípicos o entender la distribución de un conjunto de datos. Por ejemplo, si un valor está en el percentil 90, significa que es mayor que el 90% de los valores en el conjunto.

Ejemplo

Supongamos que estás analizando los tiempos de finalización de una carrera de maratón. Tienes los tiempos en minutos de 100 corredores:

145, 152, 160, 168, 173, 180, 185, 190, ... (y así sucesivamente para los otros corredores).

Ahora, imagina que deseas calcular el percentil 75 en esta distribución de tiempos. Esto significa que estás buscando el valor que separa el 75% más bajo de los tiempos del 25% más alto.

Primero, calculamos el rango en el que caerá el percentil 75. Dado que tenemos 100 tiempos, el 75% de los tiempos corresponde a los primeros 75 valores. Ahora, ordenamos los tiempos y encontramos el valor en la posición 75:

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA
"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

145, 152, 160, 168, 173, 180, 185, ... (continuando con los valores)

Imagine que el tiempo en la posición 75 es 185 minutos. Esto significa que el 75% de los corredores completaron la carrera en 185 minutos o menos. El percentil 75 en este caso es 185 minutos, lo que indica que tres cuartos de los corredores finalizaron la carrera en ese tiempo o menos.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import numpy as np
4
5 # Paso 2: Cargar los datos
6 datos = [145, 152, 160, 168, 173, 180, 185, 190]
7
8 # Paso 3: Calcular Percentiles
9 percentil_25= np.percentile(datos, 25)
10 percentil_50= np.percentile(datos, 50)
11 percentil_75= np.percentile(datos, 75)
12
13 print("Percentil 25% = ", percentil_25)
14 print("Percentil 50% = ", percentil_50)
15 print("Percentil 75% = ", percentil_75)
```

[2] ✓ 0.0s

```
... Percentil 25% = 158.0
Percentil 50% = 170.5
Percentil 75% = 181.25
```

Los resultados muestran para los datos seleccionados que el valor correspondiente para el percentil 25% es 158, la mediana o percentil 50% es 170.5 y el percentil 75% es 181.2.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Rango Muestral

El rango muestral, también conocido como amplitud muestral, es una medida estadística que describe la diferencia entre el valor más alto y el valor más bajo en un conjunto de datos. En otras palabras, es la distancia entre el valor máximo y el valor mínimo presentes en el conjunto de observaciones.

El cálculo del rango muestral es bastante simple. Si tienes un conjunto de datos $x_1, x_2, x_3, \dots, x_n$, el rango muestral (R) se calcula como:

$$R = \max(X_i) - \min(X_i)$$

Donde:

- R es el rango muestral.
- \max representa el valor máximo.
- \min representa el valor mínimo.

El rango muestral proporciona una idea de la variabilidad de los datos en el sentido de cuánto se extienden desde el valor mínimo hasta el valor máximo. Sin embargo, como medida de dispersión, el rango puede ser influenciado en gran medida por valores extremos o atípicos, lo que puede hacer que no sea una medida de variabilidad muy robusta en algunos casos.

Ejemplo

Supongamos que estamos analizando las temperaturas máximas registradas en una ciudad durante una semana de verano. Estos son los valores en grados Celsius para cada día: 32, 34, 30, 33, 35, 29 y 36.

Para calcular el rango muestral, simplemente restamos el valor más alto del valor más bajo:

$$\text{Rango muestral} = 36 - 29 = 7$$

En este caso, el rango muestral es 7 grados Celsius. Esto indica que las temperaturas máximas variaron en 7 grados durante esa semana en la ciudad. El rango muestral nos da una idea de la amplitud de la variabilidad en los datos, mostrando cuánto se extienden desde el valor más bajo hasta el valor más alto.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Desviación estándar

La desviación estándar es una medida estadística que cuantifica la dispersión o variabilidad de un conjunto de datos con respecto a su media. En otras palabras, mide qué tan dispersos o alejados están los valores individuales del valor promedio. Una desviación estándar más grande indica que los valores individuales tienden a estar más alejados de la media, mientras que una desviación estándar más pequeña indica que los valores están más cerca de la media.

La fórmula para una población implica los siguientes pasos:

- Calcular la media del conjunto de datos.
- Restar la media de cada valor individual.
- Elevar al cuadrado cada diferencia.
- Calcular la media de las diferencias al cuadrado.
- Tomar la raíz cuadrada de la media de las diferencias al cuadrado.

Y está dada por:

$$\sigma^2 = \frac{(X_1 - \mu)^2 + \dots + (X_n - \mu)^2}{N} = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

Fuente: Libro Estadística Descriptiva de Liliana Orellana (página 41)

Y para una muestra es:

$$s^2 = \frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Fuente: Libro Estadística Descriptiva de Liliana Orellana (página 41)

Donde:

- N es el tamaño de la población.
- n es el tamaño de la muestra.
- Xi es el valor individual en el conjunto de datos.
- μ o \bar{X} es la media del conjunto de datos.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

La desviación estándar es ampliamente utilizada para describir la variabilidad en diferentes campos, como la investigación científica, el análisis financiero, la ciencia de datos y más. Proporciona una medida cuantitativa de cuánto se extienden los valores individuales con respecto a la media, lo que ayuda a comprender la dispersión de los datos y a identificar posibles valores atípicos.

Ejemplo

Supongamos que estamos analizando los puntajes de un grupo de estudiantes en un examen final. Aquí están los puntajes de los 10 estudiantes: 85, 92, 78, 60, 95, 87, 75, 82, 90 y 88.

Primero, calculamos la media de los puntajes:

$$\text{Media} = (85 + 92 + 78 + 60 + 95 + 87 + 75 + 82 + 90 + 88) / 10 = 84.2$$

Luego, calculamos las diferencias entre cada puntaje y la media, las elevamos al cuadrado y calculamos la media de estas diferencias al cuadrado. Finalmente, tomamos la raíz cuadrada de la media de diferencias al cuadrado para obtener la desviación estándar:

$$\text{Desviación estándar} = 9.68$$

La desviación estándar en este caso es 9.68. Esto nos indica cuánto varían los puntajes individuales con respecto a la media. Una desviación estándar más grande implica que los puntajes están más dispersos alrededor de la media, mientras que una desviación estándar más pequeña indica una dispersión más cercana a la media.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import numpy as np
4
5 # Paso 2: Cargar los datos
6 datos = [85, 92, 78, 60, 95, 87, 75, 82, 90, 88]
7
8 # Paso 3: Calcular la Desviación Estándar
9 desviacion_estandar = np.std(datos)
10 print("Desviación Estándar = ", desviacion_estandar)
11
```

[3] ✓ 0.0s

... Desviación Estándar = 9.682974749528164

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Rango Intercuartil

El rango intercuartil (RIC), también conocido como rango intercuartílico, es una medida estadística que describe la amplitud de los valores ubicados en el rango intermedio de un conjunto de datos. Se calcula como la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1), lo que significa que el RIC abarca el 50% central de los datos ordenados.

El cálculo del rango intercuartil (*RIC*) en un conjunto de datos es:

$$RIC = Q3 - Q1$$

Donde:

- Q3 es el tercer cuartil.
- Q1 es el primer cuartil.

El rango intercuartil es útil para describir la dispersión de los datos en el rango medio, excluyendo los valores extremos o atípicos que podrían afectar el rango muestral. Debido a su naturaleza, el RIC es una medida de dispersión que es menos sensible a valores atípicos y proporciona una visión más robusta de la variabilidad central en el conjunto de datos.

Esta medida es especialmente valiosa en situaciones donde se busca comprender la variación en los valores del medio y se desea evitar la influencia de valores extremos. El rango intercuartil es una herramienta comúnmente utilizada en estadísticas, análisis de datos y otros campos donde se busca caracterizar la dispersión en los datos de manera más robusta.

Ejemplo

Suponga que está analizando los sueldos mensuales de los empleados en una empresa. Aquí están los sueldos de un grupo de empleados en miles de dólares: 28, 30, 32, 35, 40, 42, 45, 50, 55 y 60.

Primero, calculamos el primer cuartil (Q1), y el tercer cuartil (Q3)

$$Q1 = 32.75$$

$$Q3 = 48.75$$

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Ahora, calculamos el rango intercuartil (RIC):

$$RIC = Q3 - Q1$$

$$RIC = 16$$

En este caso, el rango intercuartil es 16 mil dólares. Esto indica que la diferencia entre el tercer cuartil y el primer cuartil es de este valor. El rango intercuartil nos proporciona información sobre la variabilidad en el centro de los datos, excluyendo los valores extremos que podrían afectar el rango total.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para operaciones estadísticas
3 import numpy as np
4
5 # Paso 2: Cargar los datos
6 datos = [28, 30, 32, 35, 40, 42, 45, 50, 55, 60]
7
8 # Paso 3: Calcular el Rango Intercuartil
9 Q1 = np.percentile(datos, 25)
10 Q3 = np.percentile(datos, 75)
11
12 RIC = Q3 - Q1
13
14 print("Q1 = ", Q1)
15 print("Q3 = ", Q3)
16 print("El rango intercuartil es = ", RIC)
17
18
19
20
21
```

[2] ✓ 0.0s

... Q1 = 32.75
Q3 = 48.75
El rango intercuartil es = 16.0

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"

1.8 Coeficiente de correlación

El coeficiente de correlación es una medida estadística que cuantifica la fuerza y dirección de la relación lineal entre dos variables numéricas. En otras palabras, indica en qué medida los cambios en una variable están asociados con los cambios en otra variable. El coeficiente de correlación proporciona información sobre cómo se comportan conjuntamente dos variables y si existe una relación y en qué grado. Existen varias pruebas para determinar el coeficiente de correlación, sin embargo, las más utilizadas son: Pearson, Spearman y Kendall.

Pearson

El coeficiente de Pearson es una medida estadística que evalúa la fuerza y dirección de la relación lineal entre dos variables numéricas. Es ampliamente utilizado para cuantificar cómo los cambios en una variable están relacionados con los cambios en otra variable, y proporciona información sobre la forma en que estas variables se comportan juntas.

El coeficiente de correlación de Pearson se denota como r y se calcula utilizando la fórmula:

x_i son los valores de una variable, por ejemplo, la edad.

y_i son los valores de la otra variable, por ejemplo, el salario.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

r es el coeficiente de correlación de Pearson

\bar{x} y \bar{y} son respectivamente los valores medios de las dos variables.

Fuente: <https://datatab.es/tutorial/pearson-correlation>

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

El coeficiente de Pearson varía entre -1 y 1:

- **$r = 1$** indica una correlación positiva perfecta, lo que significa que a medida que una variable aumenta, la otra también aumenta en una relación lineal perfecta.
- **$r = -1$** indica una correlación negativa perfecta, lo que significa que a medida que una variable aumenta, la otra disminuye en una relación lineal perfecta.
- **$r = 0$** indica que no hay correlación lineal entre las variables.

Un valor de r cercano a 1 o -1 indica una correlación fuerte, mientras que un valor cercano a 0 indica una correlación débil o nula.

Es importante destacar que el coeficiente de Pearson mide específicamente relaciones lineales entre variables y no captura relaciones no lineales. Además, la correlación no implica causalidad; es decir, una alta correlación no necesariamente significa que un cambio en una variable causa un cambio en la otra.

Ejemplo

Imagine que está analizando la relación entre las horas de estudio y las calificaciones obtenidas en un grupo de estudiantes. Aquí tienes los datos de cinco estudiantes:

Horas de Estudio	Calificación
5	85
10	90
15	92
8	88
12	91

Para calcular la correlación de Pearson entre las horas de estudio y las calificaciones, seguimos estos pasos:

1. Calcular la media de las horas de estudio y las calificaciones.
2. Calcular las diferencias entre cada valor y la media correspondiente.
3. Calcular el producto de las diferencias para cada par de valores y sumar los productos.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

4. Calcular la suma de los cuadrados de las diferencias para horas de estudio y calificaciones.

5. Aplicar la fórmula del coeficiente de Pearson, finalmente tenemos:

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Librería para calcular el coeficiente de correlación
3 from scipy import stats
4
5 # Paso 2: Cargar los datos
6 horas_estudio = [5, 10, 15, 8, 12]
7 calificacion = [85, 90, 92, 88, 91]
8
9 # Paso 3: Calcular la correlación de Pearson
10 r = stats.pearsonr(horas_estudio, calificacion)
11
12 print("Coeficiente Pearson = ", r[0])
13
14
15
16
17
18
```

[3] ✓ 0.0s

... Coeficiente Pearson = 0.970050852541194

En este ejemplo, el coeficiente de correlación de Pearson r es aproximadamente **0.97**. Esto indica una correlación positiva bastante fuerte entre las horas de estudio y las calificaciones de los estudiantes. Un valor cercano a **1** sugiere una relación positiva lineal, lo que significa que a medida que las horas de estudio aumentan, las calificaciones también tienden a aumentar.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Spearman

El coeficiente de correlación de Spearman, también conocido como coeficiente de correlación por rangos de Spearman, es una medida estadística que evalúa la relación entre dos variables, pero en lugar de medir una correlación lineal, mide la correlación de los rangos o posiciones relativos de los valores en ambas variables. Es especialmente útil cuando las relaciones entre variables no son lineales, pero aún existe una relación monotónica (en la misma dirección) entre ellas.

El coeficiente de Spearman se denota como r_s y se calcula de la siguiente manera:

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

Fuente: <https://www.questionpro.com/es/help/correlation-analysis-spearman-rho.html>

Donde:

- D es la diferencia entre los rangos de los valores correspondientes en las dos variables.
- n es el número de observaciones en el conjunto de datos.

El coeficiente de Spearman varía entre **-1** y **1**:

- $r_s = 1$ indica una correlación positiva perfecta, lo que significa que a medida que aumenta una variable, la otra también aumenta de manera monotónica.
- $r_s = -1$ indica una correlación negativa perfecta, lo que significa que a medida que aumenta una variable, la otra disminuye de manera monotónica.
- $r_s = 0$ indica que no hay una correlación monotónica entre las variables.

El coeficiente de Spearman es robusto ante valores atípicos y no requiere que las variables estén distribuidas normalmente o que la relación sea lineal. Por lo tanto, es una herramienta útil cuando se desea evaluar la relación general entre variables que no cumplen con las suposiciones de la correlación de Pearson.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Ejemplo

Suponga que está investigando si hay una relación entre la cantidad de horas de estudio y las calificaciones de un grupo de 10 estudiantes. Tienes los siguientes datos:

Estudiante	Horas de Estudio	Calificación
1	5	85
2	7	78
3	3	90
4	6	82
5	4	88
6	2	92
7	8	75
8	9	70
9	1	96
10	7	79

Pasos:

1. Ordene los datos de cada variable por separado

Horas de estudio: 1, 2, 3, 4, 5, 6, 7, 7, 8, 9

Calificaciones: 70, 75, 78, 79, 82, 85, 88, 90, 92, 96

2. Asigne los rangos a los valores de ambas variables. Si existen valores repetidos, se calcula el promedio de los rangos correspondiente.

Horas de Estudio	Rango de Horas	Calificación	Rango de Calificación
1	1	96	10
2	2	92	9
3	3	90	8
4	4	88	7
5	5	85	6
6	6	82	5
7	7.5	79	4
7	7.5	78	3
8	9	75	2
9	10	70	1

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

3. Calcule la diferencia entre los rangos de las dos variables y súmelas al cuadrado

Horas de Estudio	Rango de Horas	Calificación	Rango de Calificación	Diferencia de Rangos	Diferencia al cuadrado
1	1	96	10	-9	81
2	2	92	9	-7	49
3	3	90	8	-5	25
4	4	88	7	-3	9
5	5	85	6	-1	1
6	6	82	5	1	1
7	7.5	79	4	3.5	12.25
7	7.5	78	3	4.5	20.25
8	9	75	2	7	49
9	10	70	1	9	81

4. Calcule el coeficiente de correlación de Spearman utilizando la fórmula

Donde:

- D^2 = Es la diferencia entre los rangos al cuadrado
- n = número de registros, para este ejemplo 10 estudiantes
- $r_s = 1 - (6 \times 328.5) / (10 \times (10^2 - 1))$
- $r_s \sim -0.99$

En este ejemplo, el coeficiente de correlación de Spearman es aproximadamente -0.99, lo que sugiere una correlación negativa muy fuerte entre las horas de estudio y las calificaciones. Lo que significa que a medida que aumenta una variable, la otra disminuye de manera monotónica.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA
"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

```
1 # Paso 1: Cargar las librerías
2 # Libreria para calcular el coeficiente de correlación
3 from scipy import stats
4
5 # Paso 2: Cargar los datos
6 horas_estudio = [5, 7, 3, 6, 4, 2, 8, 9, 1, 7]
7 calificacion = [85, 78, 90, 82, 88, 92, 75, 70, 96, 79]
8
9 # Paso 3: Calcular la correlación de Spearman
10 rs = stats.spearmanr(horas_estudio, calificacion)
11
12 print("Coeficiente Spearman = ", rs[0])
13
```

[1] ✓ 3.1s

... Coeficiente Spearman = -0.9969650916353059

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Kendall

La correlación de Kendall, también conocida como coeficiente de concordancia de Kendall, es una medida estadística utilizada para evaluar la similitud o concordancia entre dos conjuntos de datos ordenados. A diferencia de otras medidas de correlación, como Pearson o Spearman, la correlación de Kendall se basa en comparar las relaciones de orden entre los valores en lugar de los valores numéricos reales.

Esta medida es especialmente útil cuando se trata de datos que no siguen una distribución normal o cuando la relación no es lineal. La correlación de Kendall se concentra en evaluar si las observaciones en dos variables se mueven en la misma dirección o en direcciones opuestas.

El coeficiente de correlación de Kendall se denota como τ (tau) y se calcula utilizando la fórmula:

$$\tau = ((\text{número de pares concordantes} - \text{número de pares discordantes}) / (\text{número de pares concordantes} + \text{número de pares discordantes}))$$

Donde:

- Los "pares concordantes" son pares de observaciones que mantienen su orden en ambas variables.
- Los "pares discordantes" son pares de observaciones que cambian su orden en ambas variables.

El valor de τ varía entre -1 y 1:

- $\tau = 1$ indica una correlación perfecta positiva, lo que significa que todos los pares de observaciones están en el mismo orden en ambas variables.
- $\tau = -1$ indica una correlación perfecta negativa, lo que significa que todos los pares de observaciones están en órdenes opuestos en ambas variables.
- $\tau = 0$ indica que no hay correlación entre las variables en términos de su orden.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

La correlación de Kendall es útil en situaciones donde se busca medir la concordancia en el ordenamiento de observaciones, como en comparaciones de rankings, preferencias o evaluaciones. Es menos sensible a valores atípicos y adecuada para datos no paramétricos o cuando no se cumplen las suposiciones de la correlación lineal.

Ejemplo

Suponga que dos entrenadores de voleibol clasifican a 12 de sus jugadores según su desempeño con el equipo. La siguiente tabla muestra la clasificación:

Jugador	Clasificación Entrenador 1	Clasificación Entrenador 2
1	1	1
2	2	2
3	3	3
4	4	5
5	5	4
6	6	7
7	7	6
8	8	8
9	9	10
10	10	9
11	11	11
12	12	12

Como estamos trabajando con datos clasificado, lo ideal es aplicar Kendall para la correlación entre la clasificación de los entrenadores.

Pasos:

1. Encuentre los pares concordantes

Para este paso observe solo la clasificación del entrenador 2, comenzando en el jugador 1 y cuente cuantos valores debajo de él son más grandes. Por ejemplo, en 1 hay 11 valores por debajo de él que son más grandes, en 2 hay 10 valores por debajo que son mayores y así sucesivamente hasta finalizar. La tabla resultante es:

📍 Sede Quirinal: Calle 21 No. 6 - 01
📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220
📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699
✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA
"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Jugador	Clasificación Entrenador 1	Clasificación Entrenador 2	Concordantes
1	1	1	11
2	2	2	10
3	3	3	9
4	4	5	7
5	5	4	7
6	6	7	5
7	7	6	5
8	8	8	4
9	9	10	2
10	10	9	2
11	11	11	1
12	12	12	

2. Encuentre los pares discordantes

Realice el mismo proceso del paso 1, pero ahora cuente los valores que están por debajo de la clasificación del entrenador 2. Por ejemplo, en 1 no hay jugadores por debajo con valor menor, entonces se asigna 0; en 2 tampoco hay valores menores debajo de él, entonces se asigna 0. Finalice el proceso, la tabla resultante es:

Jugador	Clasificación Entrenador 1	Clasificación Entrenador 2	Concordante s	Discordante s
1	1	1	11	0
2	2	2	10	0
3	3	3	9	0
4	4	5	7	1
5	5	4	7	0
6	6	7	5	1
7	7	6	5	0
8	8	8	4	0
9	9	10	2	1
10	10	9	2	0
11	11	11	1	0
12	12	12		

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

3. Sume las columnas de concordantes y discordantes

$$\text{concordantes} = 11+10+9+7+7+5+5+4+2+2+1 = 63$$

$$\text{discordantes} = 0+0+0+1+0+1+0+0+1+0+0 = 3$$

4. Calcule Kendall

$$\tau = (\text{concordantes} - \text{discordantes}) / (\text{concordantes} + \text{discordantes})$$

$$\tau = (63 - 3) / (63 + 3) = 0.90$$

Como el valor es cercano a 1, indica una correlación positiva buena (por ser cercana a 1), lo que significa que todos los pares de observaciones están en el mismo orden en ambas variables.

Ejemplo en Python

Escriba y ejecute el siguiente código Python en un Notebook:

```
1 # Paso 1: Cargar las librerías
2 # Libreria para calcular el coeficiente de correlación
3 from scipy import stats
4
5 # Paso 2: Cargar los datos
6 entrenador_1 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
7 entrenador_2 = [1, 2, 3, 5, 4, 7, 6, 8, 10, 9, 11, 12]
8
9 # Paso 3: Calcular la correlación de Kendall
10 t = stats.kendalltau(entrenador_1, entrenador_2)
11
12 print("Coeficiente Kendall = ", t[0])
13
```

[1] ✓ 2.9s

... Coeficiente Kendall = 0.909090909090909

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989

NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"



CORHUILA

CORPORACIÓN UNIVERSITARIA DEL HUILA
Vigilada Mineducación

INSTITUCIÓN DE EDUCACIÓN SUPERIOR SUJETA A INSPECCIÓN
Y VIGILANCIA POR EL MINISTERIO DE EDUCACIÓN NACIONAL - SNIES 2828

Bibliografía

- Agresti, A., & Finlay, B. (2009). Statistical Methods for the Social Sciences (4th ed.). Pearson.
- Babbie, E. R. (2015). The Basics of Social Research (7th ed.). Cengage Learning.
- Bluman, A. G. (2019). Elementary Statistics: A Step by Step Approach (10th ed.). McGraw-Hill Education.
- Cramer, D., & Howitt, D. L. (2004). The SAGE Dictionary of Statistics: A Practical Resource for Students in the Social Sciences. Sage Publications.
- Dancey, C. P., & Reidy, J. (2017). Statistics Without Maths for Psychology (7th ed.). Pearson.
- Levin, R. I., Fox, J. A., & Forde, D. R. (2018). Elementary Statistics in Social Research (13th ed.). Pearson.
- Moore, D. S., & McCabe, G. P. (2018). Introduction to the Practice of Statistics (9th ed.). W. H. Freeman.
- Pagano, R. R., & Gauvreau, K. (2018). Principles of Biostatistics (2nd ed.). Cengage Learning.
- Rosner, B. (2015). Fundamentals of Biostatistics (8th ed.). Cengage Learning.
- Triola, M. F. (2017). Elementary Statistics (13th ed.). Pearson.

📍 Sede Quirinal: Calle 21 No. 6 - 01

📍 Sede Prado Alto: Calle 8 No. 32 - 49 PBX: (608) 8754220

📍 Sede Pitalito: Carrera 2 No. 1 - 27 - PBX: (608) 8360699

✉ Email: contacto@corhuila.edu.co - www.corhuila.edu.co

Personería Jurídica Res. Ministerio de Educación No. 21000 de Diciembre 22 de 1989
NIT. 800.107584-2



CORPORACIÓN UNIVERSITARIA DEL HUILA - CORHUILA

"Diseño y prestación de servicios de docencia, investigación y extensión de programas de pregrado, aplicando todos los requisitos de las normas ISO implementadas en sus sedes Neiva y Pitalito"