

MINISTERIO DEL TRABAJO

Ciencia de Datos (Estadística)

Centro de Servicios y Gestión Empresarial SENA Regional Antioquia





Cuando programas leyendo documentación





Cuando programas con ChatGPT





Estadística

Estadística



La estadística es la ciencia que se encarga de recopilar, organizar, procesar, analizar e interpretar datos con el fin de deducir las características de un grupo o población objetivo, su importancia radica en que es una fuente de información altamente confiable para la toma de decisiones. Ocaña-Riola, R. (2017)

Estadística



Para el estudio de la estadística, la podemos dividir en tres secciones:

- La obtención de datos proporciona métodos para obtener datos que permiten dar respuestas claras a preguntas concretas.
- El análisis de datos se ocupa de los métodos y las ideas necesarias para organizar y describir datos utilizando gráficos y resúmenes numéricos.
- La inferencia estadística va más allá de los datos disponibles y obtiene conclusiones sobre un universo más amplio teniendo en cuenta la variabilidad y la incertidumbre de las conclusiones.

Ciencia de Datos

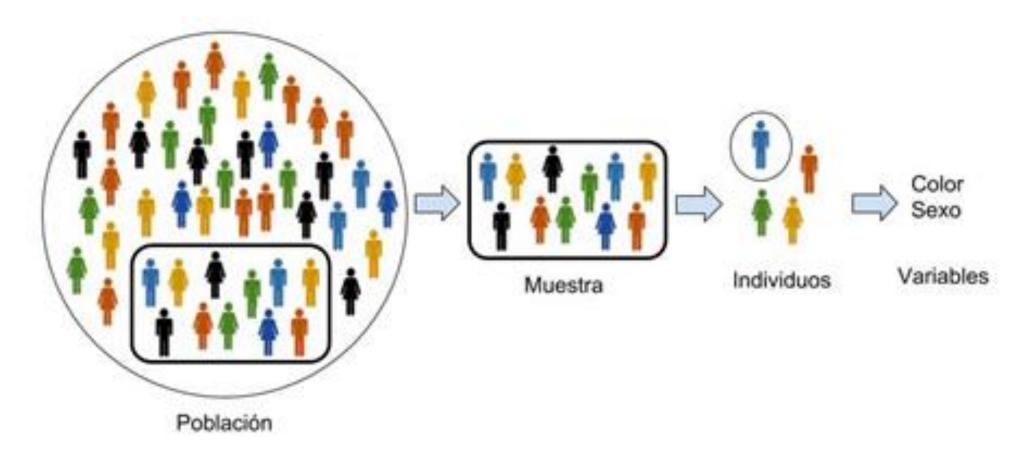


Se relacionan los datos con un propósito (**objetivos**) que permite proponer soluciones o metodologías y así plantear recomendaciones a un problema en particular, a partir de información estructurada en una base de datos, conformada por **variables** (columnas) y **registros** (filas).

El número de registros por lo regular de gran tamaño, hace necesaria la utilización de herramientas computacionales para su análisis y procesamiento (**Python**).

Conceptos





Parámetro: Una característica específica de una población (fijo)

Estadístico: Una característica específica de una muestra (varía de muestra en muestra)

Variable: Una característica de un individuo

Tipos de estadística





Tipos de estadística



Descriptiva

Recolección de datos Crítica y clasificación de datos Presentación de información Tablas, gráficos y medidas estadísticas

Análisis descriptivo

Inferencial

Población objetivo

Características desconocidas Muestra probabilística

Seleccionada con leyes de probabilidad **Datos**

Obtenidos de cada unidad de la muestra Estimaciones

Calculadas con los datos de la muestra Conclusiones

Válidas para la población con un error y confiabilidad tolerada

Tipos de variables



Nominales: categorías no ordenables (sexo, grupo sanguíneo, origen de un producto)

Ordinales: categorías con orden natural (nivel de estudios, nivel de satisfacción)

Discretas: conjunto de valores aislados, finito o numerable (número de llamadas, número de hijos)

Continuas: valores en un intervalo de la recta real de forma continua (temperatura, concentración de CO2)

Categóricas o cualitativas

Tipos de Variables

Variable: Son características de interés de una población sobre las que se quiere hacer un estudio estadístico

Cuantitativas o numéricas

Indicadores y graficas



Moda

Categóricas o cualitativas

Diagrama de torta o diagrama de barras

Indicadores

Cuantitativas o numéricas

Media, mediana, mínimo, máximo, rango, varianza, desviación estándar, coeficiente de variación, coeficiente de curtosis, coeficiente de asimetría, percentiles, deciles y cuartiles.

Histograma, diagrama de cajas, diagrama de densidad, diagrama de puntos, diagrama de líneas o mapa de calor



GRACIAS

Presentó: Alvaro Pérez Niño Instructor Técnico

Correo: aperezn@misena.edu.co

http://centrodeserviciosygestionempresarial.blogspot.com/

Línea de atención al ciudadano: 01 8000 910270 Línea de atención al empresario: 01 8000 910682



www.sena.edu.co