



Universidad
Industrial de
Santander



A CALCULAR...

Lina Vanessa Girón Lozano
Grupo Halley

CONTENIDO

- ▶ Introducción a la estadística
 - ¿Qué es la estadística?
 - Conceptos Básicos (población, muestra)
 - Clasificación de la estadística.
- ▶ Datos individuales y datos estadísticos
- ▶ Estructura del dato
- ▶ Variables estadísticas
- ▶ Tipos de variables estadísticas
- ▶ La medición
- ▶ Tabla de frecuencia
- ▶ Media, mediana y moda
- ▶ Varianza, desviación estándar y coeficiente de variación.

INTRODUCCIÓN A LA ESTADÍSTICA

- ▶ La **estadística** (la forma femenina del término alemán *Statistik*, derivado a su vez del italiano *statista*, "hombre de Estado") es una rama de las matemáticas.

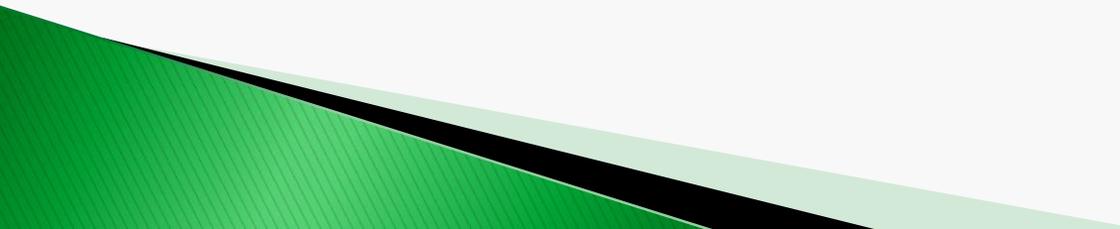


¿QUÉ ES LA ESTADÍSTICA?

- ▶ Es una herramienta que estudia usos y análisis provenientes de una muestra representativa de datos, que busca explicar las correlaciones y dependencias de un fenómeno físico o natural, de ocurrencia en forma aleatoria o condicional, mediante procedimientos que permiten recoger datos, presentarlos, ordenarlos y analizarlos, de manera que a partir de ellos se puedan inferir conclusiones.

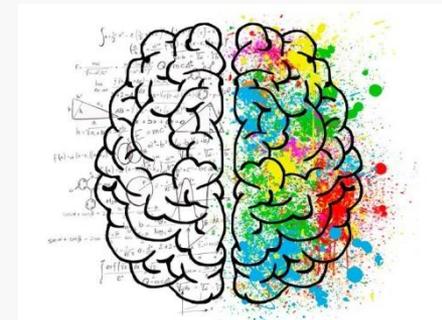
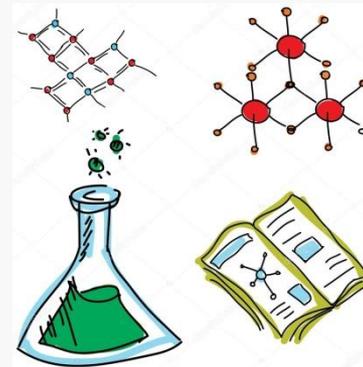
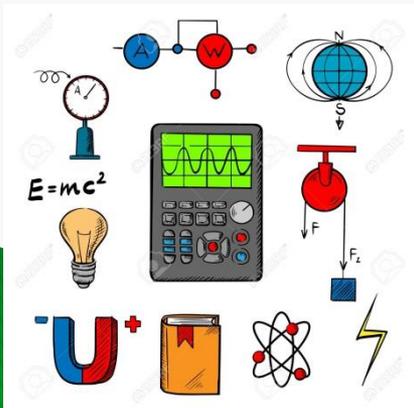


¿QUÉ ES LA ESTADÍSTICA?

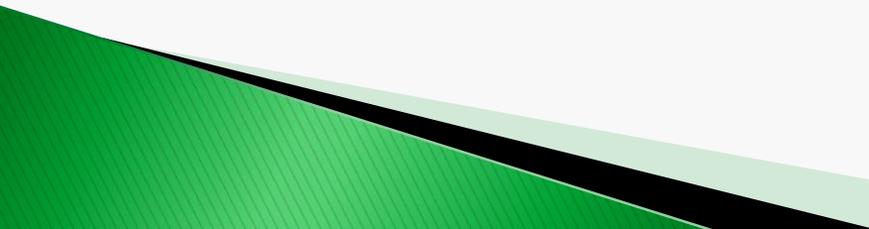
- ▶ Por ejemplo, la estadística interviene cuando se quiere conocer el estado sanitario de un país, a través de ciertos parámetros como la tasa de morbilidad o mortalidad de la población.
 - ▶ En este caso la estadística describe la muestra en términos de datos organizados y resumidos, y luego infiere conclusiones respecto de la población.
- 

¿QUÉ ES LA ESTADÍSTICA?

- ▶ Aplicada a la investigación científica, también infiere cuando provee los medios matemáticos para establecer si una hipótesis debe o no ser rechazada.
- ▶ La estadística puede aplicarse a cualquier ámbito de la realidad, y por ello es utilizada en física, química, biología, medicina, astronomía, psicología, sociología, lingüística, demografía, etc.



CONCEPTOS BÁSICOS

- ▶ Puesto que la estadística se ocupa de una gran cantidad de datos, debe primeramente definir de cuáles datos se va a ocupar. El conjunto de datos de los cuales se ocupa un determinado estudio estadístico se llama **población**.
 - ▶ No debe confundirse la población en sentido demográfico y la población en sentido estadístico.
- 

CONCEPTOS BÁSICOS

- ▶ La población en sentido demográfico es un conjunto de individuos (todos los habitantes de un país, todas las ratas de una ciudad), mientras que una población en sentido estadístico es un conjunto de datos referidos a determinada característica o atributo de los individuos (las edades de todos los individuos de un país, el color de todas las ratas de una ciudad).

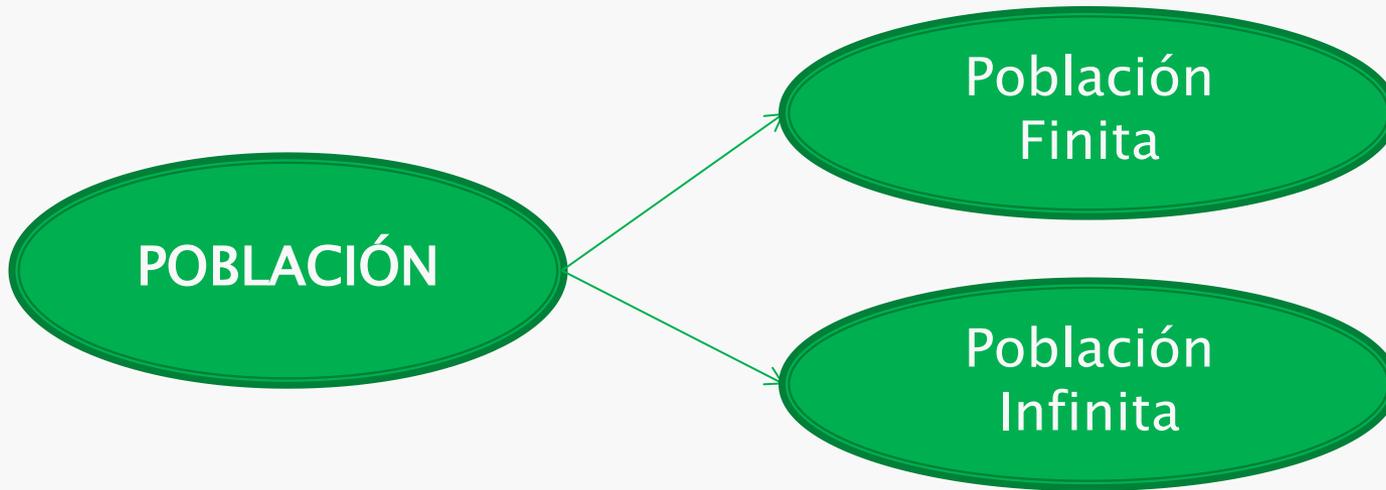


CONCEPTOS BÁSICOS

- ▶ Incluso una población en sentido estadístico no tiene por qué referirse a muchos individuos. Una población estadística puede ser también el conjunto de calificaciones obtenidas por un individuo a lo largo de sus estudios universitarios.



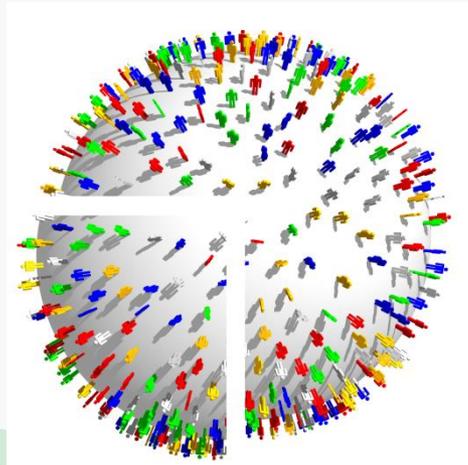
POBLACIÓN



- ▶ **Población:** Es el conjunto de elementos (cosas), individuos (personas) o entes sujetos a estudio (situaciones) y de los cuales queremos obtener un resultado.

POBLACIÓN

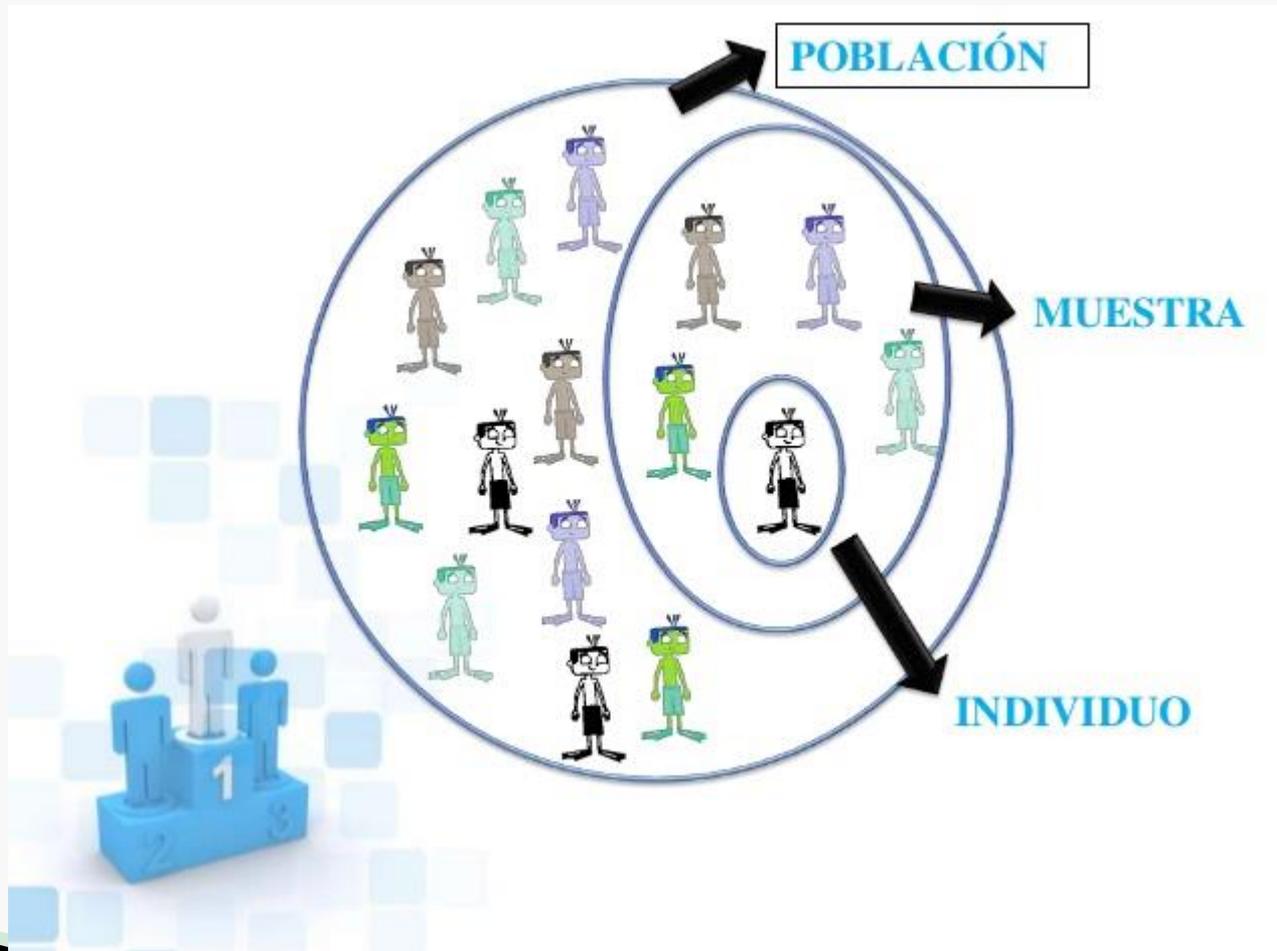
- ▶ **Población finita:** Cuando el número de elementos que la forman es finito, por ejemplo el número de alumnos de un centro de enseñanza, o grupo clase.
- ▶ **Población infinita:** Cuando el número de elementos que la forman es infinito, o tan grande que pudiesen considerarse infinitos.



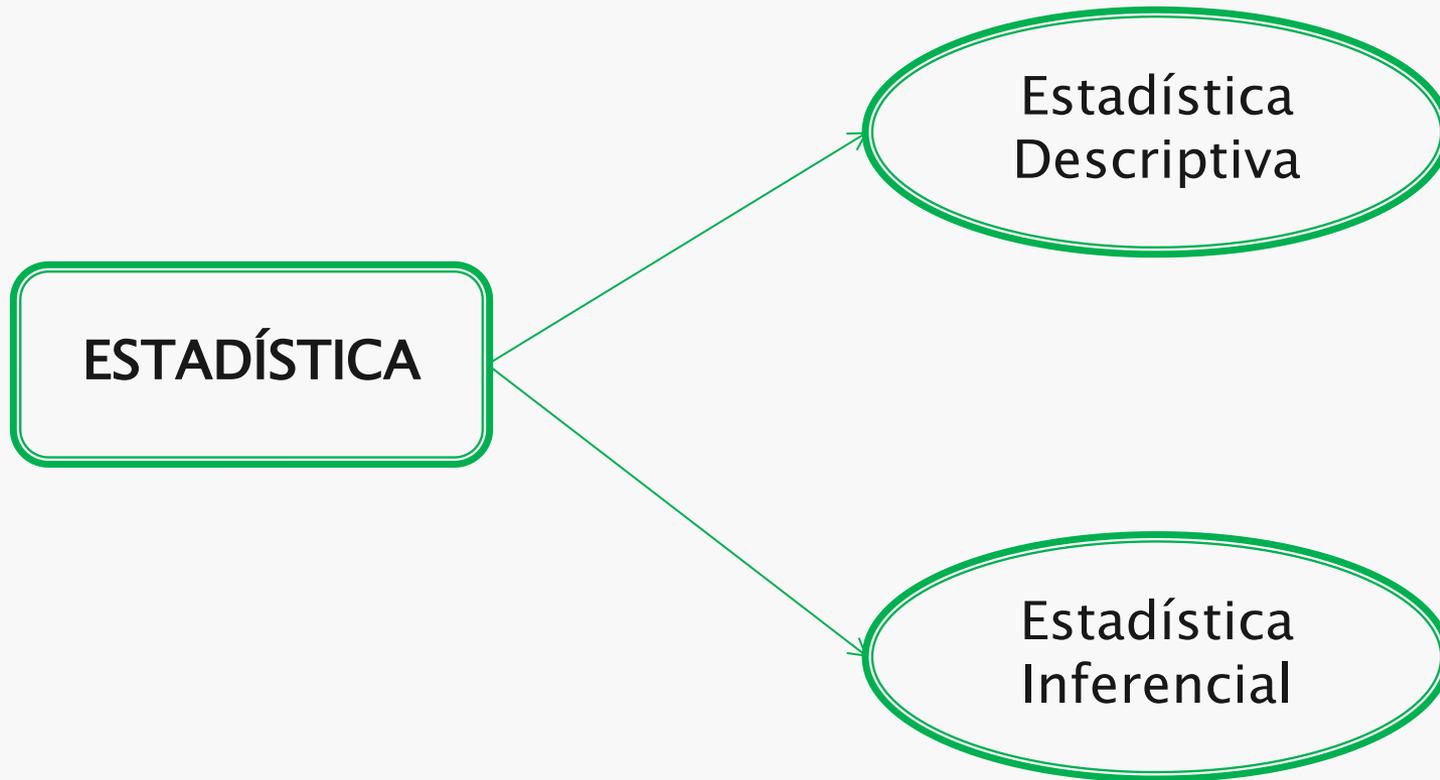
MUESTRA E INDIVIDUO

- ▶ **Muestra:** Es un subconjunto de una determinada población, la parte más representativa, elegida al azar.
 - ▶ **Individuo:** Cada uno de los elementos que componen la población estadística en estudio. Es un ser observable que no tiene por qué ser una persona, puede ser un objeto, un ser vivo, etc...
- 

POBLACIÓN, MUESTRA E INDIVIDUO



CLASIFICACIÓN DE LA ESTADÍSTICA



CLASIFICACIÓN DE LA ESTADÍSTICA

- ▶ **Estadística descriptiva:** Realiza el estudio sobre la **población completa**, observando una característica de la misma y calculando unos parámetros que den información global de toda la población.
 - ▶ **Estadística inferencial:** Realiza el estudio descriptivo sobre un **subconjunto** de la población llamado **muestra** y posteriormente, extiende los resultados obtenidos a toda la población.
- 

DATOS INDIVIDUALES Y ESTADÍSTICOS

- ▶ Un **dato individual** es un dato de un solo individuo, mientras que un **dato estadístico** es un dato de una muestra o de una población en su conjunto. Por ejemplo, la edad de Juan es un dato individual, mientras que el promedio de edades de una muestra o población de personas es un dato estadístico.



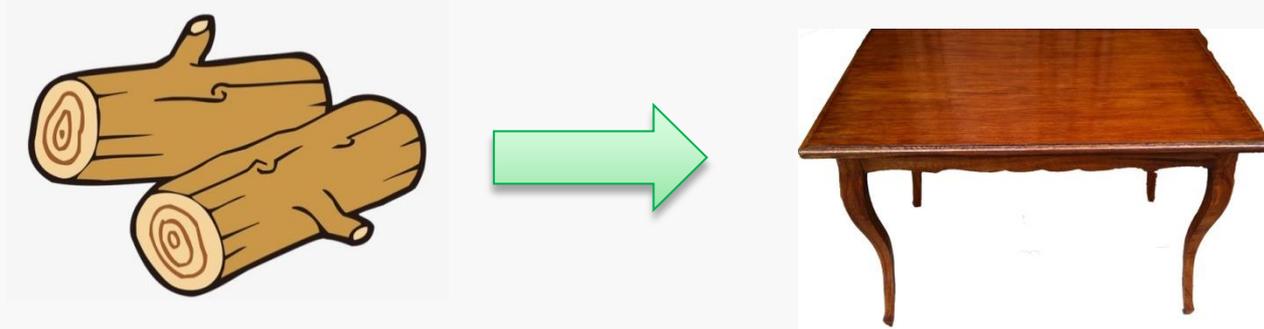
DATOS INDIVIDUALES Y ESTADÍSTICOS

- ▶ Los datos que describen una muestra suelen llamarse **estadísticos** (por ejemplo, el promedio de ingresos mensuales de las personas de una muestra), mientras que los datos estadísticos descriptores de una población suelen llamarse **parámetros** (por ejemplo, el promedio de ingresos mensuales de las personas de una población).



ESTRUCTURA DEL DATO

- ▶ Los datos son la materia prima con que trabaja la estadística, del mismo modo que la madera es la materia prima con que trabaja el carpintero.



- ▶ Tanto los datos como la madera no se inventan: se extraen de la realidad; en todo caso el secreto está en recoger la madera o los datos más adecuados a los objetivos del trabajo a realizar.

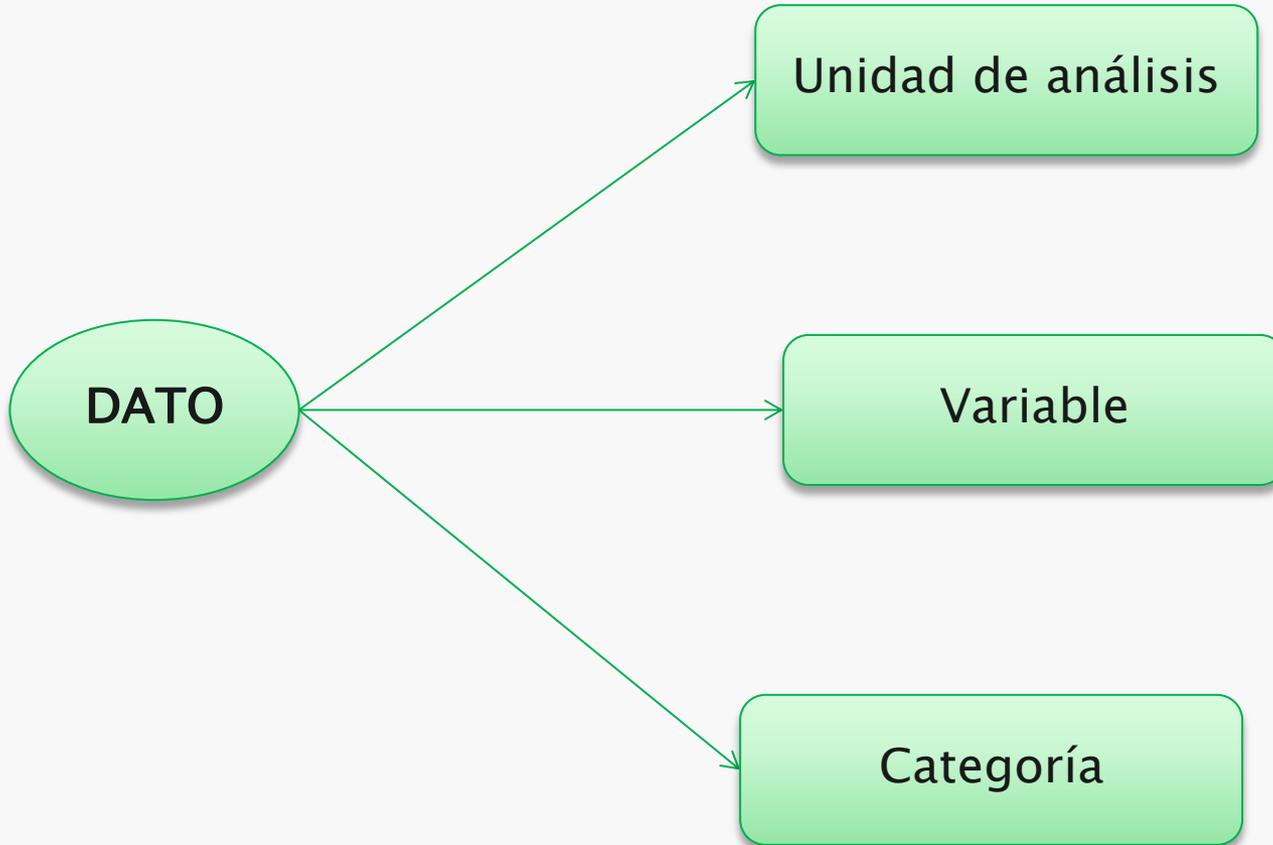
ESTRUCTURA DEL DATO

- ▶ De una manera general, puede definirse técnicamente **dato** como una categoría asignada a una variable de una unidad de análisis.

Por ejemplo, "Luis tiene 1.70 metros de estatura" es un dato, donde 'Luis' es la unidad de análisis, 'estatura' es la variable, y '1.70 metros' es la categoría asignada.



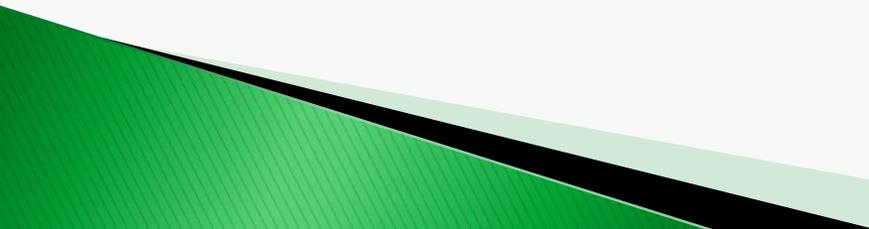
ESTRUCTURA DEL DATO



ESTRUCTURA DEL DATO

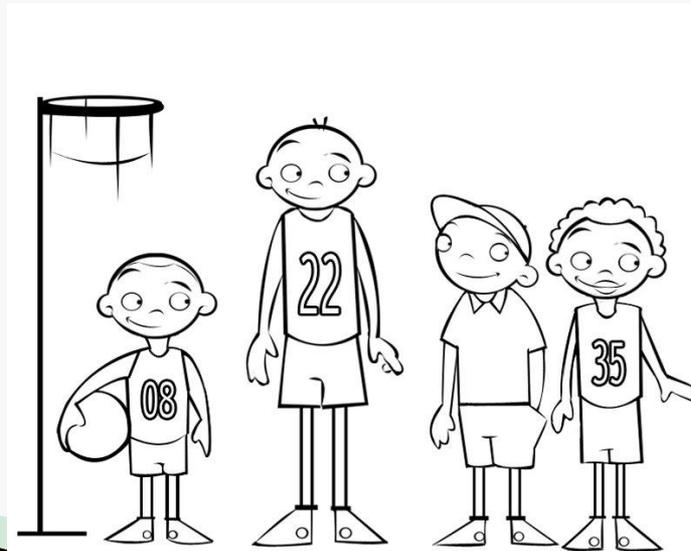
- ▶ La **unidad de análisis** es el elemento del cual se predica una propiedad y característica. Puede ser una persona, una familia, un animal, una sustancia química, o un objeto como una dentadura o una mesa.
- ▶ La **variable** es la característica, propiedad o atributo que se predica de la unidad de análisis.
 - Por ejemplo puede ser la edad para una persona, el nivel de aprendizaje alcanzado para un animal, el peso específico para una sustancia química, el nivel de 'salud' para una dentadura, y el tamaño para una mesa.

ESTRUCTURA DEL DATO

- ▶ Pueden entonces también definirse **población estadística** (o simplemente población) como el conjunto de datos acerca de unidades de análisis (individuos, objetos) en relación a una misma característica, propiedad o atributo (variable).
 - ▶ Sobre una misma población demográfica pueden definirse varias poblaciones de datos, una para cada variable.
- 

ESTRUCTURA DEL DATO

- ▶ La **categoría** es cada una de las posibles variaciones de una variable. Categorías de la variable sexo son masculino y femenino, de la variable ocupación pueden ser arquitecto, médico, etc, y de la variable edad pueden ser 10 años, 11 años, etc.

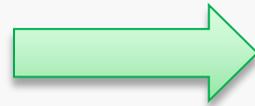


VARIABLES ESTADÍSTICAS

- ▶ Al hacer un estudio de una determinada población, observamos una característica o propiedad de sus elementos o individuos. Cada una de estas características o cualidades estudiadas se llama **variable estadística**.



VARIABLES ESTADÍSTICAS

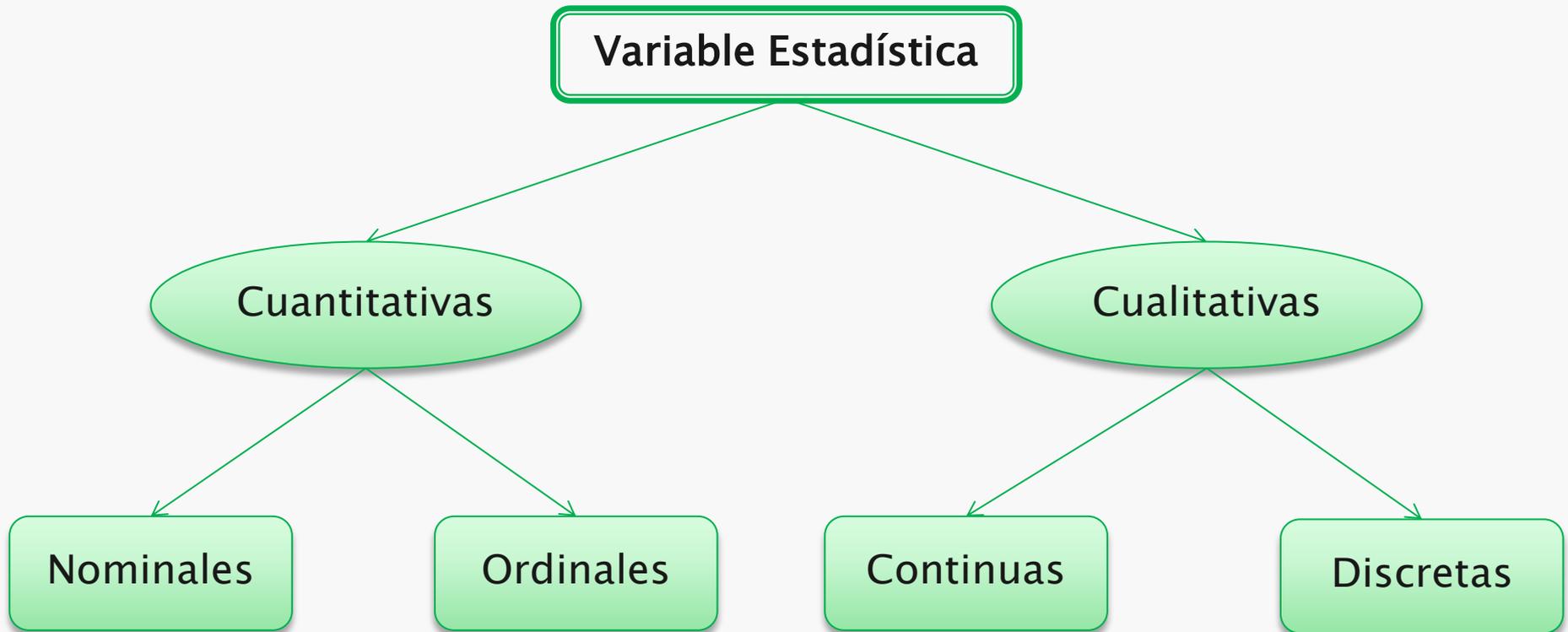


- Sector
- Tamaño
- Número de empleados
- Producción
- Ventas



- Edad
- Estatura
- Peso
- Color de ojos
- Escolaridad

TIPOS DE VARIABLES ESTADÍSTICAS



TIPOS DE VARIABLES ESTADÍSTICAS

- ▶ **Variables cualitativas:** Es aquella característica que no podemos expresar con números y hay que expresarla con palabras. Por ejemplo, el lugar de residencia, comida favorita, profesión que te gusta....
 - **Nominales:** Aquellas que sólo admiten una mera ordenación alfabética, pero no establece orden por su naturaleza, por ejemplo, el color del cabello, sexo, estado civil, etc...
 - **Ordinales:** Aquellas que sugieren una ordenación, por ejemplo la graduación militar, el nivel de estudios, etc....

TIPOS DE VARIABLES ESTADÍSTICAS

- ▶ **Variables cuantitativas:** Es cualquier característica que se puede expresar con números. Por ejemplo, el número de hermanos, la estatura, ´ números de alumnos en tu colegio, etc....
 - **Continua:** Es aquella variable que puede tomar **cualquier valor dentro de un intervalo real**. Por ejemplo, la estatura.
 - **Discreta:** Es aquella variable que puede tomar únicamente un **número finito de valores**. Por ejemplo, el número de hermanos.
- 

LA MEDICIÓN

- ▶ Los datos se obtienen a través un proceso llamado medición. Desde este punto de vista, puede definirse **medición** como el proceso por el cual asignamos una categoría (o un valor) a una variable, para determinada unidad de análisis.
 - ▶ Ejemplo: cuando decimos que Martín es varón, estamos haciendo una medición, porque estamos asignando una categoría (varón) a una variable (sexo) para una unidad de análisis (Martín).
- 

LA MEDICIÓN

- ▶ Se pueden hacer mediciones con mayor o menor grado de precisión.
 - ▶ Cuanto más precisa sea la medición, más información nos suministra sobre la variable y, por tanto, sobre la unidad de análisis. No es lo mismo decir que una persona es alta, a decir que mide 1,83 metros.
- 

LA MEDICIÓN

- ▶ Los diferentes grados de precisión o de contenido informativo de una medición se suelen caracterizar como **niveles de medición**. Típicamente se definen cuatro niveles de medición, y en cada uno de ellos la obtención del dato o resultado de la medición será diferente:

Nivel de medición	Nivel nominal	Nivel ordinal	Nivel cuantitativo discreto	Nivel cuantitativo continuo
DATO	Martín es electricista	Elena terminó la secundaria	Juan tiene 32 dientes	María tiene 70 pulsaciones por minuto
Unidad de análisis	Martín	Elena	Juan	María
Variable	Oficio	Nivel de instrucción	Cantidad de piezas dentarias	Frecuencia cardíaca
Categoría o valor	Electricista	Secundaria completa	32	70
Unidad de medida	-----	-----	Diente	Pulsaciones por minuto

TABLA DE FRECUENCIA

22	19	16	13	18	15	20	14	15	16
15	16	20	13	15	18	15	13	18	15

- ▶ Rango $R = x_{m\acute{a}x} - x_{m\acute{i}n}$
- ▶ Número de intervalos K : Número de filas.
 - Regla de Sturges:

$$K = 1 + 3.322 \text{Log}(N)$$

Donde N es el número de datos.

- ▶ Amplitud A : amplitud del rango.

$$A = \frac{R}{K}$$

TABLA DE FRECUENCIA

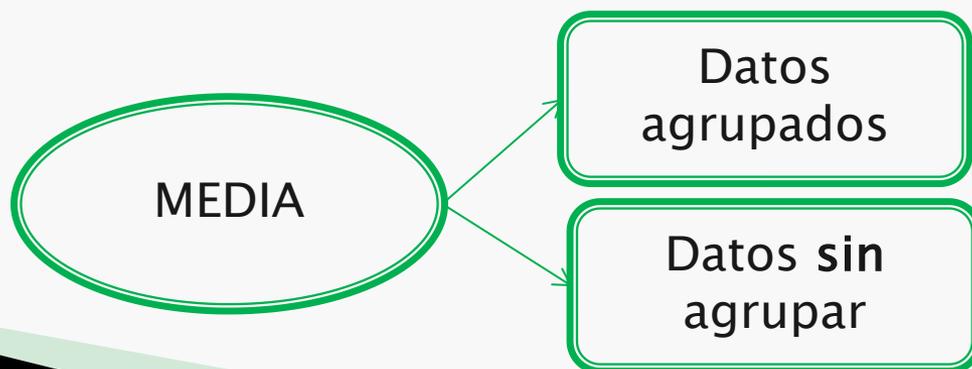
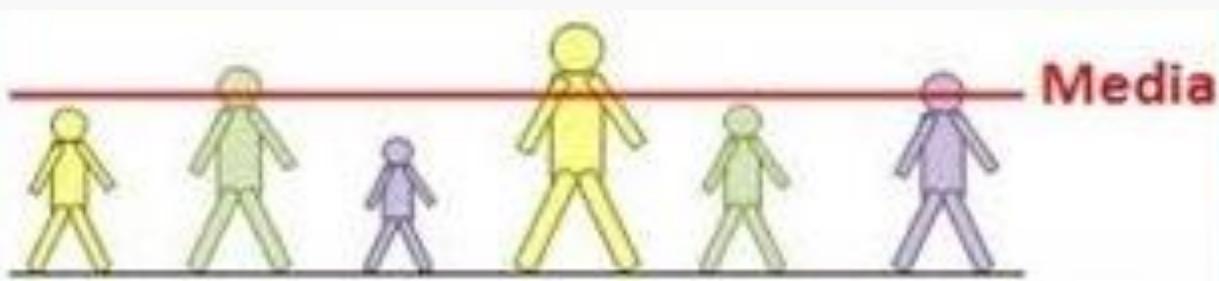
<i>Clases</i>	<i>x</i>	<i>f</i>	<i>fr</i>	<i>F</i>
[13-15)				
[15-17)				
[17-19)				
[19-21)				
[21-23)				

- ▶ x es la marca de clase: promedio entre los límites de clase.
- ▶ f es la frecuencia con que se repiten los datos.
- ▶ F es la frecuencia acumulativa.
- ▶ fr es la frecuencia relativa.

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

MEDIA

- ▶ En matemáticas y estadística una **media** o **promedio** es una medida de tendencia central. Resulta al efectuar una serie determinada de operaciones con un conjunto de números. Existen diferentes tipos de medias, pero en matemáticas el término más común es la media aritmética.



MEDIA

- ▶ La *media aritmética* es un promedio estándar que a menudo se denomina "promedio".

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

Sin agrupar

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

Agrupados

Sin agrupar → x_i representa cada dato.

Agrupados → x_i representa la marca de clase.

- ▶ La media se confunde a veces con la mediana o moda. La media aritmética es el promedio de un conjunto de valores, o su distribución.

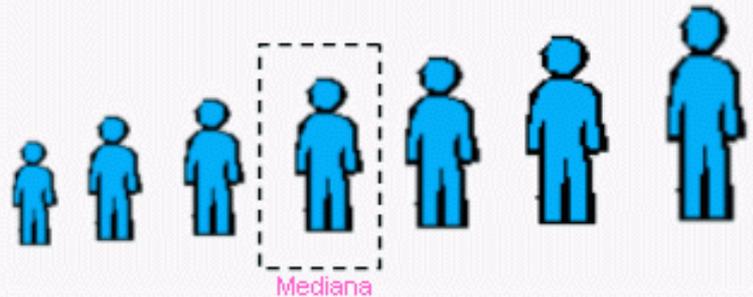
MEDIANA

- ▶ En el ámbito de la estadística, la **mediana** (del latín *mediānus* 'del medio') representa el valor de la variable de posición central en un conjunto de datos ordenados.
- ▶ La mediana no siempre es un valor que pertenece a la serie de datos.

MEDIANA

Datos
agrupados

Datos sin
agrupar



MEDIANA

- ▶ Datos sin agrupar.
- ▶ En este caso los datos se ordenan de forma creciente o decreciente, teniendo un total de datos llamado N.
- ▶ Si N es impar, la mediana es el valor que ocupa la posición $(N + 1)/2$, este es el valor central.

$$M_e = x_{(N+1)/2}$$

- ▶ Si N es par, la mediana es la media aritmética de los valores centrales.

$$M_e = \frac{\left(x_{N/2} + x_{\left(\frac{N}{2}\right)+1} \right)}{2}$$

MEDIANA

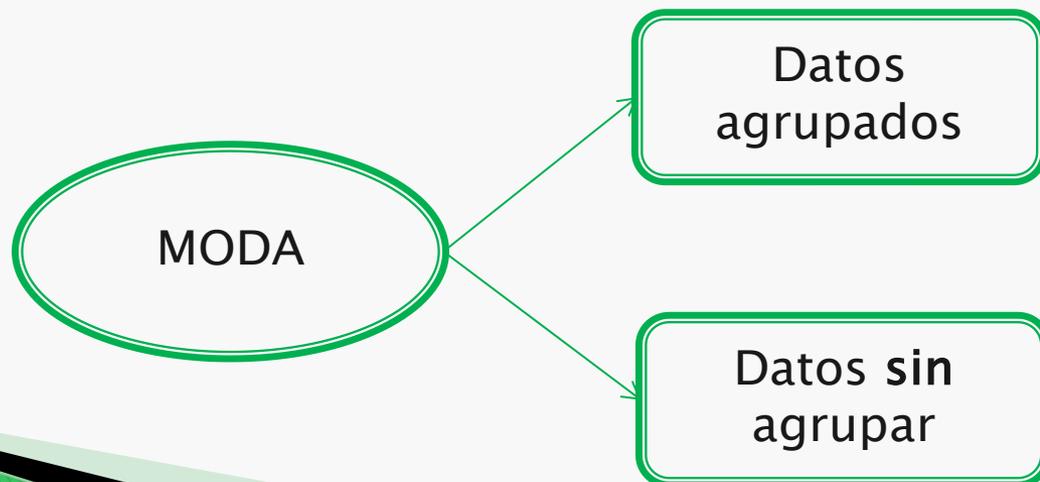
- ▶ **Datos agrupados.**
- ▶ Lo primero que debemos hacer para poder calcular la mediana es identificar la **clase mediana**. Para esto tenemos que buscar el intervalo en el que se encuentre $N/2$.
- ▶ Ahora se aplicará la siguiente fórmula:

$$M_e = L_i + \left(\frac{\frac{N}{2} - F_{i-1}}{f_i} \right) t_i$$

- ▶ L_i es el límite inferior de la clase donde se encuentra la mediana.
- ▶ F_{i-1} es la frecuencia acumulada anterior a la clase mediana.
- ▶ f_i es la frecuencia absoluta del intervalo mediano.
- ▶ t_i es la amplitud de los intervalos. Se calcula restando el extremo superior menos el inferior del intervalo.

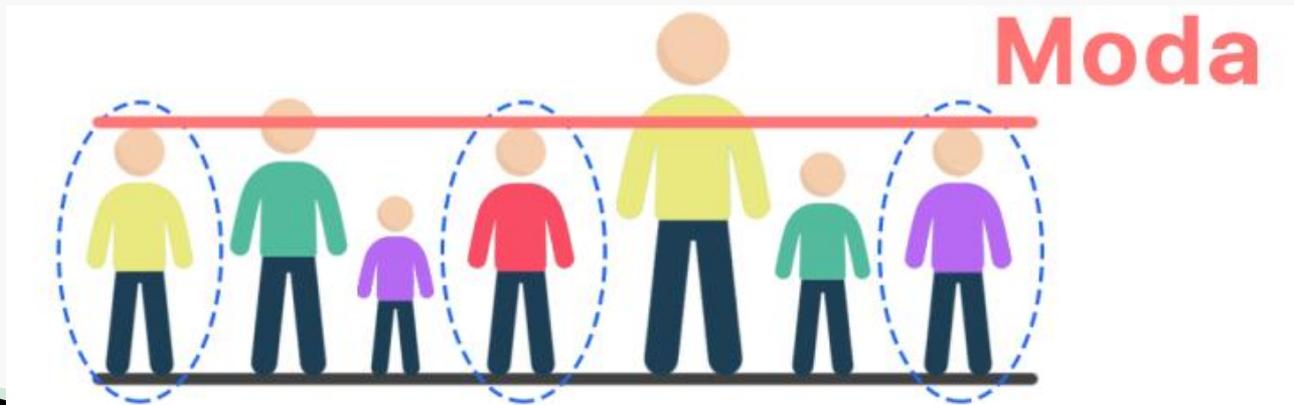
MODA

- ▶ La moda se conoce como el dígito o individuo que más se repite. En estadística, la moda es el valor con mayor frecuencia en una distribución.
- ▶ Al igual que la mediana la moda se calcula para dos tipos de datos.



MODA

- ▶ Datos sin agrupar.
- ▶ En este caso los datos se ordenan de forma creciente o decreciente, teniendo un total de datos llamado N.
- ▶ La moda es simplemente el valor o individuo que más se repite.



MODA

- ▶ Datos agrupados.
- ▶ En este caso, cuando los datos se encuentran agrupados, lo primero que se debe hacer es identificar el intervalo modal, o sea, el intervalo con mayor frecuencia absoluta.
- ▶ Luego, aplicamos la fórmula:

$$M_o = L_i + \left(\frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \right) t_i$$

VARIANZA

- ▶ La varianza o variancia (que suele representarse como σ^2) de una variable aleatoria es una medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variables respecto a su media.
- ▶ Si tenemos un conjunto de datos de una misma variable, la varianza se calcula de la siguiente forma:

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

Sin agrupar

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

Agrupados

DESVIACIÓN ESTÁNDAR

- ▶ La desviación típica o desviación estándar (denotada como σ) es una medida de dispersión para variables de razón (variables cuantitativas o cantidades racionales) y de intervalo. Se define como la raíz cuadrada de la varianza.
 - ▶ Es básicamente una medida del grado de dispersión de los datos con respecto al valor promedio.
 - ▶ La desviación estándar puede ser interpretada como una medida de incertidumbre.
- 

DESVIACIÓN ESTÁNDAR

- ▶ Por ejemplo, las tres poblaciones (0,0,14,14), (0,6,8,14) y (6,6,8,8) cada una tiene una media 7. Sus desviaciones estándar poblacionales son 7, 5 y 1, respectivamente. La tercera población tiene una desviación mucho menor que las otras dos porque sus valores están más cerca de 7.
- ▶ Se calcula usando la siguiente fórmula:

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

Sin agrupar

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

Agrupando

COEFICIENTE DE VARIACIÓN

- ▶ En estadística, cuando se desea hacer referencia a la relación entre el tamaño de la media y la variabilidad de la variable, se utiliza el **coeficiente de variación**.
- ▶ Su fórmula expresa la desviación estándar como porcentaje de la media aritmética, mostrando una mejor interpretación porcentual del grado de variabilidad de la desviación típica o estándar.

$$C_v = \frac{\sigma}{|\bar{x}|}$$

$$C_v = \frac{\sigma}{|\bar{x}|} * 100$$

EJERCICIO

22	19	16	13	18	15	20	14	15	16
15	16	20	13	15	18	15	13	18	15

1. Construir la tabla de frecuencias.
2. Hallar la media, mediana y moda, agrupando los datos o sin agrupar los datos.
3. Calcular la varianza, desviación estándar y el coeficiente de variación

Gracias!!!