



## UNIDAD 1: ANALISIS DE VARIABLES

### CONCEPTOS GENERALES

- **Conceptos de ESTADÍSTICA:**

Es la ciencia que se ocupa del estudio de los fenómenos de tipo genérico, en el ámbito social y económico, normalmente complejos y enmarcados en un contexto variable.

Es la rama de las matemáticas que se encarga del análisis e interpretación de la información.

La estadística se puede clasificar en dos grandes ramas: descriptiva e inferencial

La **estadística descriptiva** se encarga de la recolección, procesamiento, análisis y presentación de un conjunto de datos.

La **estadística inferencial** se refiere al método para lograr generalizaciones acerca de las propiedades del todo.

- **POBLACIÓN Y MUESTRA**

**POBLACIÓN:** es cualquier conjunto de unidades o elementos claramente definido en el espacio y el tiempo. Los elementos pueden ser personas, hogares, manzanas, juguetes, comidas, escuelas, hospitales y empresas, entre otras.

**MUESTRA:** es un subconjunto representativo de la población a partir del cual se pretende realizar inferencias para dicha población.

- **VARIABLES**

**Variable:** es una característica de la población o de la muestra cuya medida puede cambiar de valor. Según su naturaleza puede ser cualitativa o cuantitativa.

**Variable cualitativa:** es aquella que representa cualidades, atributos o características numéricas.

**Variable cuantitativa:** es aquella característica de la población o de la muestra que es posible representar numéricamente.

- **MÉTODOS DE RECOLECCIÓN**

Para hacer la recolección de datos se usan principalmente dos métodos:

La **encuesta:** es un método de recolección de datos que se lleva a cabo generalmente por medio de algún cuestionario que puede o no ser diligenciado por la persona encuestada.

La **entrevista:** consiste en una serie de preguntas realizadas por quien entrevista personalmente, a cada uno de los entrevistados.

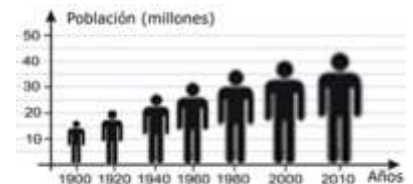
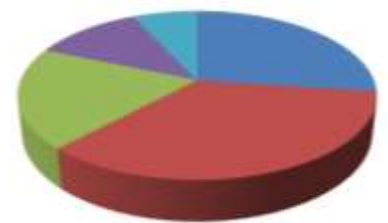
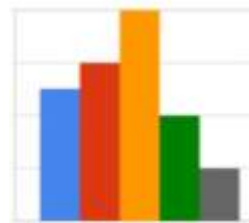
A partir de estos dos métodos, se reúne información de la muestra que constituye a lo que en estadística se conoce como **base de datos**.

El termino **dato** también es conocido como información y es el valor de la variable asociada a un elemento de una población o una muestra.

Es posible representar la información de una base de datos en diferentes formas; la más usada son los diagramas.

Los diagramas se usan dependiendo del contexto y de la información que se desea mostrar. Ofrecen una presentación sencilla y objetiva de la información.

Los diagramas más usados en estadística son de diferentes tipos: de barras o histogramas, de líneas, pictogramas, circulares.





## ANÁLISIS DE UNA VARIABLE CUALITATIVA

### CARACTERIZACIÓN DE VARIABLES CUALITATIVAS.

Para analizar una variable cualitativa se hace una caracterización de ella. **Caracterizar una variable** tiene como objetivo presentar tablas de frecuencias que brinden información resumida; además, presentar diagramas en los cuales se pueda interpretar dicha información y determinar de manera general el o los datos de mayor frecuencia.

### DISTRIBUCIÓN DE FRECUENCIAS

Una distribución de frecuencias es un resumen del conjunto que muestra el número (frecuencia) de artículos de cada una de las clases de la variable estudiada. En una distribución de frecuencias se identifican las siguientes columnas.

**Clases:** corresponde a opiniones, gustos, preferencias u otras características de la variable estudiada.

**Frecuencia:** es el número de datos de cada clase. Se representa con la letra  $f$ .

**Frecuencia relativa:** es el cociente entre la frecuencia y el número total de datos, se simboliza  $f_r$ .

**Porcentaje:** La frecuencia relativa se puede expresar como un porcentaje que se halla multiplicando por 100 el anterior cociente.

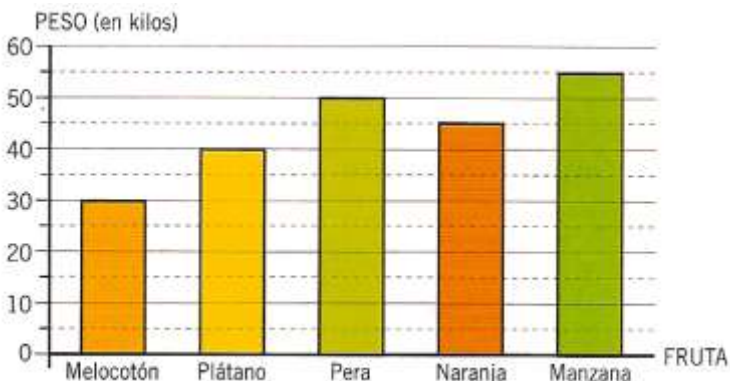
Distribución de frecuencia – Tabla de frecuencia

Clase	f	$f_r$	%

### GRAFICAS

Un **grafico estadístico** es un resumen visual de la distribución de frecuencias. Para las variables cualitativas se plantean tres tipos de graficas: los diagramas de barras, los diagramas circulares y los pictogramas.

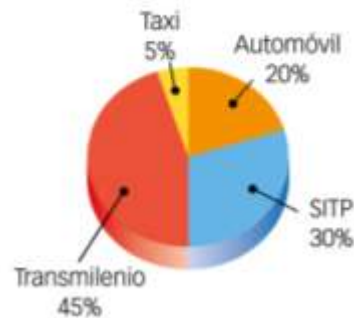
**Diagramas de barras:** es una representación de los datos estadísticos asociados a una variable.



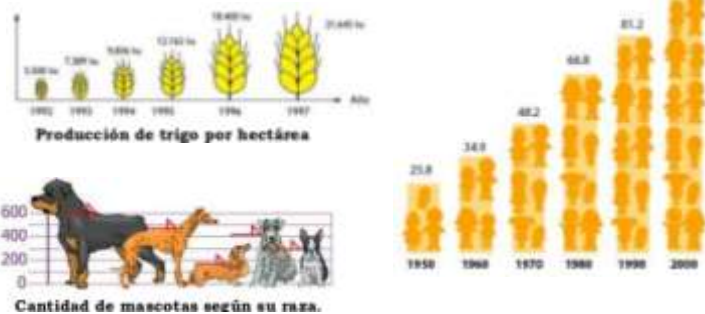
**Diagrama circular:** es la representación de datos en un círculo. Se usa para representar porcentajes correspondientes a cada clase.

En el diagrama circular, la información corresponde a cada clase de la variable queda representada por un sector circular. El sector circular es proporcional a un ángulo dentro del círculo. Para construir un diagrama circular se usan las fracciones como operador. Para hallar los ángulos correspondientes se parte de la relación en la cual el total de la población representa  $360^\circ$  del círculo.

Uso de medios de transporte



**Pictogramas:** es un gráfico vistoso, similar a un gráfico de barras; emplea un dibujo alusivo al tema que representa, en una determinada escala, para expresar la unidad de medida de los datos correspondientes a cada clase. Para la construcción de pictogramas se tiene en cuenta que su formato es libre, se emplea una secuencia de símbolos para representar frecuencias y se utilizan para el tratamiento de datos tanto cualitativos como cuantitativo



**Moda:** es una medida de tendencia central que describe el valor de la variable que tiene mayor frecuencia, es decir, la que más se repite.

Es la única medida de tendencia central que tiene sentido estudiar en una variable cualitativa, pues no precisa la realización de ningún cálculo. Por su propia definición la moda no es única, pues pueden haber dos o más valores de la variable que tengan la misma frecuencia siendo esta la máxima.



**ACTIVIDAD N° 1**

Una empresa de telecomunicaciones quiere conocer los medios de comunicación que más se usan en cierta población para informarse sobre las noticias. A continuación, se presentan los resultados obtenidos al aplicar una encuesta a 50 personas.

Televisión	Radio	Radio	Internet	Radio
Internet	Internet	Televisión	Televisión	Internet
Periódicos	Radio	Radio	Radio	Televisión
Radio	Periódicos	Periódicos	Internet	Internet
Televisión	Radio	Televisión	Televisión	Radio
Televisión	Radio	Radio	Internet	Radio
Radio	Televisión	Televisión	Televisión	Internet
Periódicos	Radio	Internet	Periódicos	Periódicos
Internet	Periódicos	Radio	Radio	Internet
Televisión	Internet	Periódicos	Televisión	Periódicos



Caracterizar la variable medios de comunicación para la información de noticias.

**A. Realizar la tabla de distribución de frecuencias**

Medios de comunicación	f	f <sub>r</sub>	F	%
Televisión				
Internet				
Periódicos				
Radio				
Internet				

- B. Construir el diagrama de barras con la frecuencia de cada medio.
- C. Construir el diagrama circular con los porcentajes de cada una de las categorías.
- D. Escriba 2 conclusiones con base en la información planeada en la tabla y en las gráficas.

2. En una importante empresa exportadora de máquinas se aplicó un estudio para determinar qué tipo de bebida consumían sus empleados en las horas laborales. El objetivo de dicho estudio era determinar de qué bebidas deberían surtirse el dispensador automático de tal forma que el aprovechamiento del espacio fuera máximo. A continuación se presentan los resultados obtenidos al tomar una muestra de 50 empleados.

Jugo	Jugo	Malta	Limonada	Jugo
Agua	Malta	Jugo	Malta	Malta
Malta	Jugo	Jugo	Jugo	Agua
Agua	Limonada	Jugo	Malta	Jugo
Uva	Jugo	Malta	Limonada	Limonada
Uva	Jugo	Limonada	Jugo	Malta
Jugo	Malta	Jugo	Limonada	Agua
Uva	Agua	Malta	Uva	Jugo
Jugo	Malta	Jugo	Malta	Uva
Limonada	Jugo	Malta	Jugo	Limonada

- A. Realizar la tabla de distribución de frecuencia, diagrama de barras y escribir 3 conclusiones de acuerdo a la información dada.



### ACTIVIDADES DE AVANCE – TAREA

1. **Determina cuatro opciones de respuesta para cada una de las siguientes variables cualitativas.**

- Calidad del servicio en un restaurante
- Profesiones que estudiarán los estudiantes cuando terminen el colegio
- Deportes que practican los estudiantes del colegio usualmente.

2. **Observa la información. Luego, resuelve.**

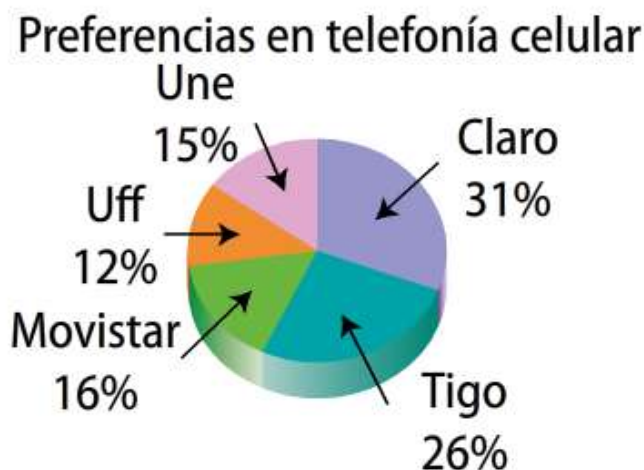
En una compañía de seguros se entrevistó a 60 clientes para determinar cuál es el tipo de seguros que más compran y poder dar mejores ofertas en seguros.

Vida	Vida	Vida	Vehículos	Salud
Salud	Salud	Salud	Vida	Desempleo
Vida	Salud	Vida	Salud	Hogar
Hogar	Salud	Hogar	Vida	Vida
Hogar	Vida	Vida	Salud	Vida
Salud	Salud	Desempleo	Vida	Hogar
Vida	Vida	Salud	Hogar	Vehículos
Hogar	Desempleo	Vida	Desempleo	Salud
Vida	Vida	Vehículos	Salud	Hogar
Salud	Vehículos	Desempleo	Hogar	Desempleo
Vida	Salud	Hogar	Vehículos	Vehículos
Hogar	Salud	Hogar	Vida	Salud

- ¿Cuál es la variable de estudio?
- Construye la tabla de frecuencias.
- Escribe tres conclusiones del análisis realizado.
- Determinar la moda.
- Realizar el diagrama de barras correspondiente.

3. **Extrae información del diagrama circular.**

El diagrama circular muestra el porcentaje de preferencia de 200 personas por algún operador de telefonía celular.



- Construye la tabla de distribución de frecuencias a partir de la información presentada en el diagrama circular.
- Realizar el diagrama de barras.
- Escribe 2 conclusiones