



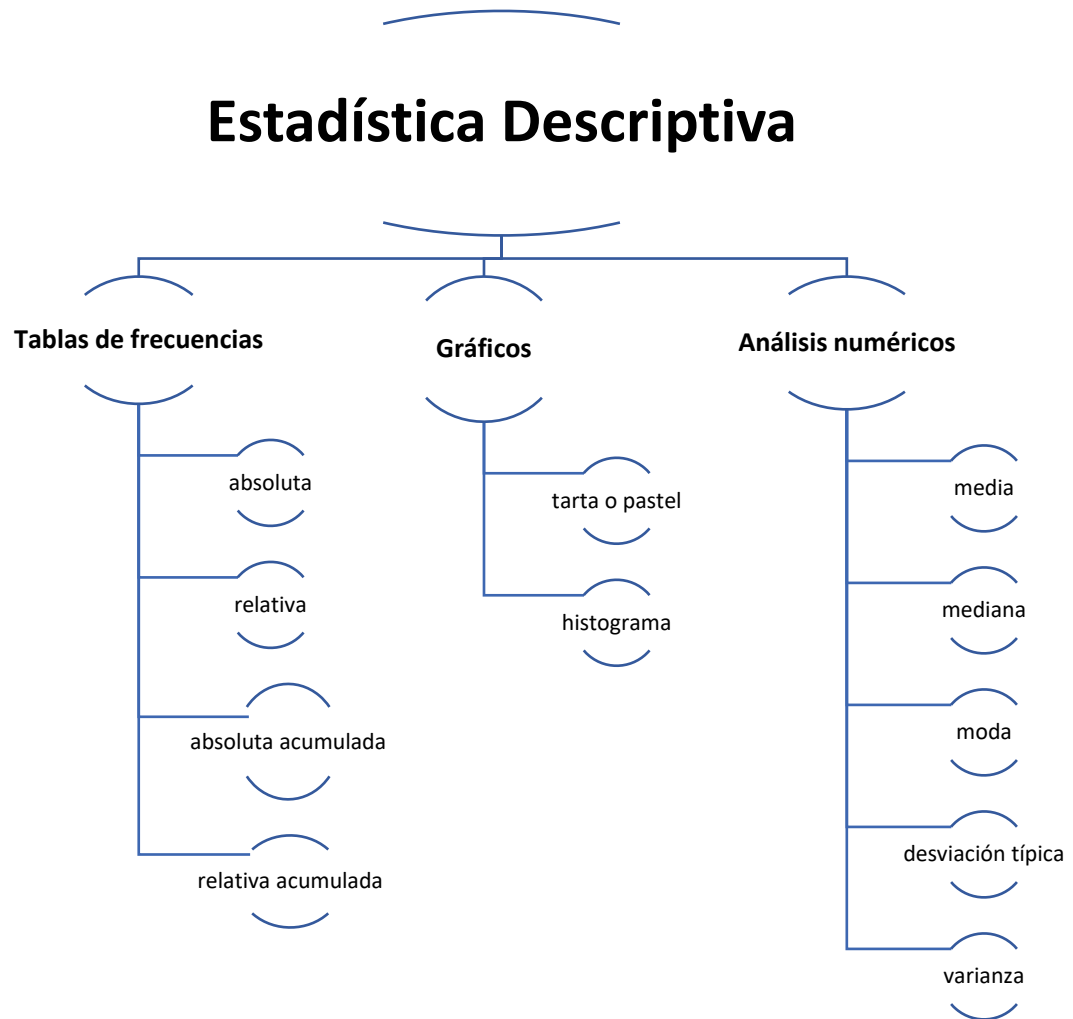
APRENDE CON ELI

ESTADÍSTICA

Estadística Descriptiva

CONCEPTOS

Estos son los conceptos básicos que vamos a aprender en la parte de Estadística Descriptiva:



La **Estadística descriptiva** es sencillamente la parte de la Estadística que se encarga de **describir** como su nombre lo dice una variable aleatoria es decir describir los datos que tengamos a mano.

Tiene como objetivo **resumir, analizar y sacar conclusiones** sobre ese conjunto de datos y va a estar dividida o se puede agrupar en tres estudios fundamentales. Uno es el estudio con **tablas de frecuencias**, otro es mediante **gráficos** y otro es mediante **resúmenes numéricos**.

En las tablas de frecuencia, la primera columna, si los datos no están agrupados, será la **clase**. Y para cada clase, vamos a medir ciertas **frecuencias**. Una de ellas es la **frecuencia absoluta**, que es simplemente la cantidad de veces que se repite esa clase.

Si sumamos todas las frecuencias absolutas tenemos que obtener el total de datos que tenemos.

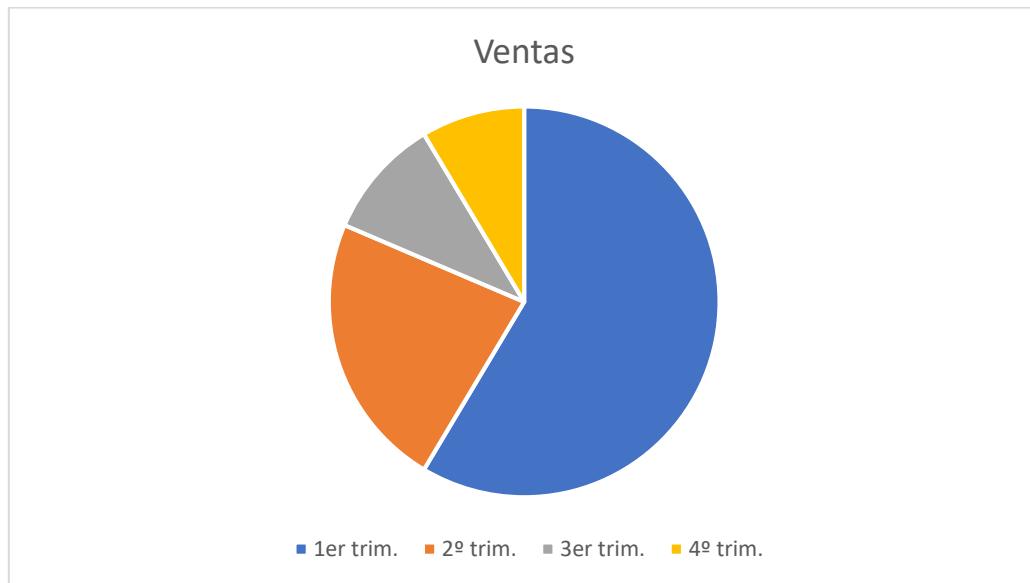
Luego, la siguiente columna será la **frecuencia absoluta acumulada** que como su nombre lo dice lo que vamos a hacer es acumular las frecuencias absolutas. La primera frecuencia absoluta acumulada tiene que ser igual a la primera frecuencia absoluta sin acumular, y luego vamos sumando. Otra cosa que siempre se tiene que cumplir es que la última frecuencia absoluta acumulada tiene que coincidir con el total de datos que tenemos, porque ya hemos acumulado todo.

Luego viene la **frecuencia relativa**, que como su nombre lo dice va a ser la frecuencia absoluta dividida entre el total. Eso nos va a dar una idea de qué porcentaje del total representa cada clase. Además, siempre se va a cumplir es que si sumamos todas las frecuencias relativas me tiene que dar 1.

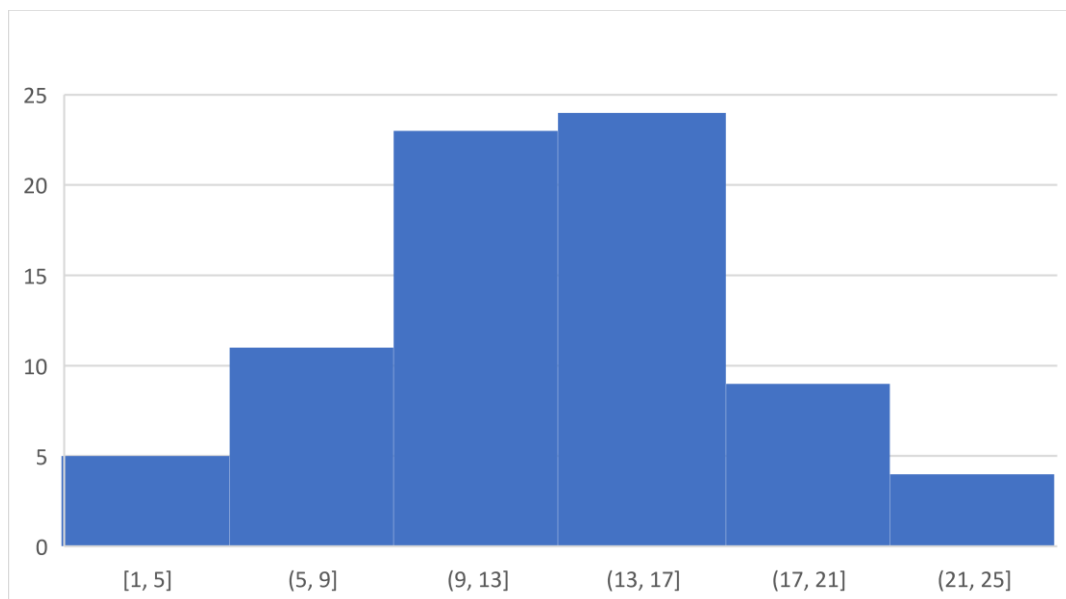
Luego estaría la **frecuencia relativa acumulada**, que igual que antes es acumular las frecuencias relativas. Por lo tanto, la primera frecuencia relativa acumulada va a ser igual a la primera frecuencia relativa sin acumular y luego iremos sumando, por tanto, la última siempre será 1.

Ahora bien, estos resultados de frecuencias se pueden representar con **gráficos**. Por ejemplo, el gráfico de tarta o de pastel representa una circunferencia y está dividida por sectores y cada sector representa la

frecuencia absoluta o relativa de cada clase. El sector con mayor área es el sector más frecuente. Un ejemplo de este gráfico sería:



Otro gráfico es el **histograma**. En el histograma, abajo se representan las clases. Y en el eje vertical, las frecuencias, que serán las alturas de cada barra. La utilidad del histograma frente al gráfico de tarta o de pastel es que podemos ver la forma de la distribución de la variable aleatoria, si hay simetría, dónde se concentran más los valores de mi variable, etc. Un ejemplo sería:



Finalmente, están los **resúmenes numéricos**, que son simplemente medidas numéricas que me caractericen a los datos. Por ejemplo, la media, la mediana, la moda, que son medidas de centralización para ver dónde está el centro de los datos.

Por ejemplo, tenemos unas alturas de las personas, entonces la altura **media** es hacer el promedio de esas alturas.

La **mediana** la podremos hallar si ordenamos primero los datos de menor a mayor y luego cogemos el que está en el medio justo. Una vez que ordenemos los datos, si quedan dos en el medio lo que vamos a hacer con esos dos datos es hacer un promedio de ellos dos y ese promedio va a ser la mediana. Es la única dificultad que tiene en caso de que tengamos un número par de datos. Porque en caso que tengamos un número impar simplemente es coger el del medio.

La **moda** es el valor que más se repite (mayor frecuencia).

Ahora, las medidas de centralización son la media, la moda y la mediana. Pero cómo podemos ver si los datos están muy concentrados alrededor de la medida representativa, por ejemplo, alrededor de la media. Pues esas son las medidas de **dispersión** que son la **varianza** y la **desviación típica** que miden cuán alejados están los datos de la media.