## **Conceptos básicos de estadística descriptiva**

#### **¿Qué son las estadísticas y qué importancia tienen?**

La estadística es un área de las matemáticas aplicadas que tiene dos ramas principales: descriptiva e inferencial.

* **La estadística descriptiva** intenta proporcionar **un resumen conciso del conjunto de datos** en cuestión. Independientemente del tipo de conjunto de datos que tengamos o de dónde provienen nuestros datos, explorar las cualidades descriptivas del conjunto de datos es un primer paso esencial para comprender lo que los datos intentan decirle.
* **La estadística inferencial** estudia las **observaciones del mundo real** **o** las ocurrencias de **eventos pasados** ​​para **inferir las proporciones del futuro**.

#### **Introducción a la estadística descriptiva**

Una estadística descriptiva es una estadística que **resume los datos de manera concisa** y nos brinda una visión general de los datos. Mirar las estadísticas descriptivas de nuestros datos es uno de los pasos más útiles e importantes de nuestro análisis. Pero, ¿qué son las estadísticas descriptivas? Algunas de las estadísticas descriptivas más populares son medidas de centralidad. Estas estadísticas nos brindan diferentes formas de definir cuál es el medio (o centro) de los datos. Aquí hablaremos de las dos medidas más importantes que son:

1. Medida de tendencia central.
2. Medidas de dispersión.

Nuestro objetivo aquí es aprender a usar la **Statistics** biblioteca en **Python** para calcular las estadísticas descriptivas discutidas anteriormente.

Algunas otras estadísticas descriptivas incluyen:

* Max
* Min
* Count
* Frequency
* Quartiles

#### **Tipos de datos: numéricos (cuantitativos) frente a categóricos (cualitativos)**

* **Los datos cuantitativos** son el tipo de datos que **se pueden describir mediante números**. Por ejemplo, el *promedio de horas* que los estudiantes en un bootcamp trabajan por semana es un ejemplo de este tipo de datos: se puede expresar mediante un número. A veces, esto se denomina simplemente datos "numéricos".
* Por otro lado, **los datos cualitativos** son un tipo de datos que **NO se pueden describir con números**. Un ejemplo de este tipo de datos sería *la formación educativa* más común para todas aquellas personas que han asistido y han asistido a bootcamps: no podemos asignar un número a “literatura” o “ingeniería” y mantener su relevancia. Esto también se denomina a menudo datos "categóricos".

Dentro de los datos numéricos o cuantitativos, existen **dos subtipos**: datos discretos y continuos.

* **Datos discretos** se refieren a datos que **se CUENTAN** y **solo pueden tomar un cierto rango o tipo de valores (Nº DEFINIDO DE VALORES)**. Por ejemplo, *el número de estudiantes* en cada bootcamp es un número discreto: ¡no vamos a contar medio estudiante! En este caso, los datos sólo pueden tomar la forma de números enteros.
* **Datos continuos**, o datos que **se MIDEN** y **pueden tomar una cantidad INFINITA DE VALORES**. Un buen ejemplo de esto es la *altura de cada estudiante en la clase:* este número podría ser cualquier número, entero o decimal, dentro de un rango de valores (0 cm hasta la altura más alta jamás registrada, o infinito, si estamos siendo aspiracionales) .