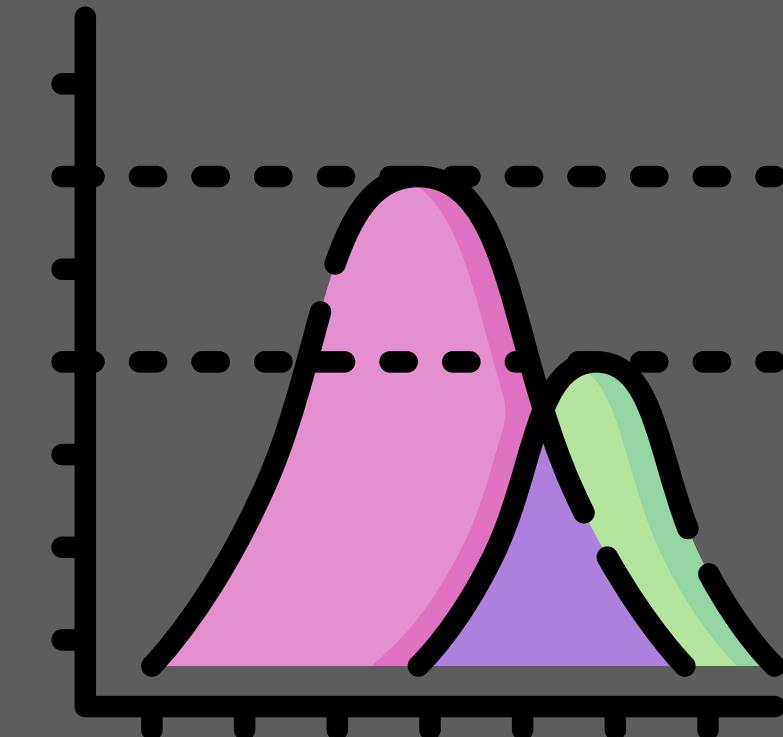


Análisis Predictivo

Unidad II



Introducción

La estadística es el estudio de los **fenómenos aleatorios**.

La mayor parte de la información viene expresada en forma de datos por lo que es necesario el uso de **estadística** para la correcta interpretación de los mismos.



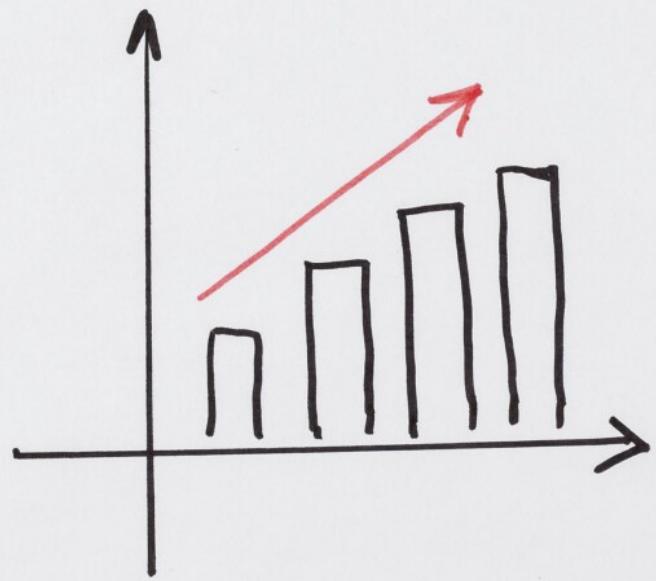
Datos históricos

Los datos históricos son la clave para predecir el futuro. Aprende cómo **analizar** y **visualizar** tus datos para identificar patrones y tendencias que te ayuden a prever eventos futuros.



Descriptiva

Recolección de **datos**, resumen, presentación y descripción de los mismos, por lo que resulta útil para realizar análisis exploratorios (**EDA**).



Inferencial

Proceso de utilizar los datos experimentales de una **muestra** proveniente de una **población** para tomar decisiones. Es fundamental para hacer **análisis predictivos**.



Estadística

Resumen

- Análisis descriptivo y exploratorio
- Análisis predictivo

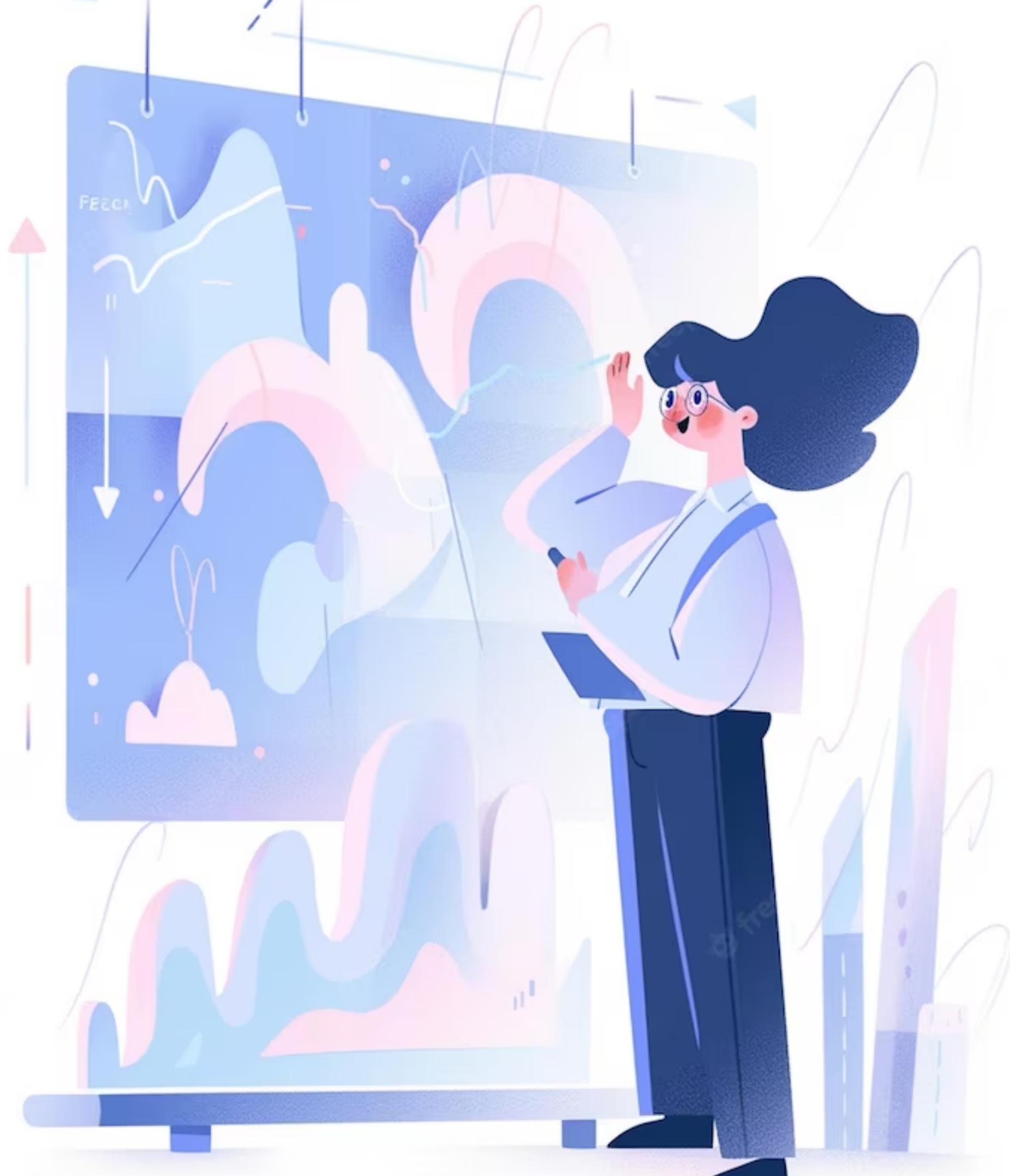


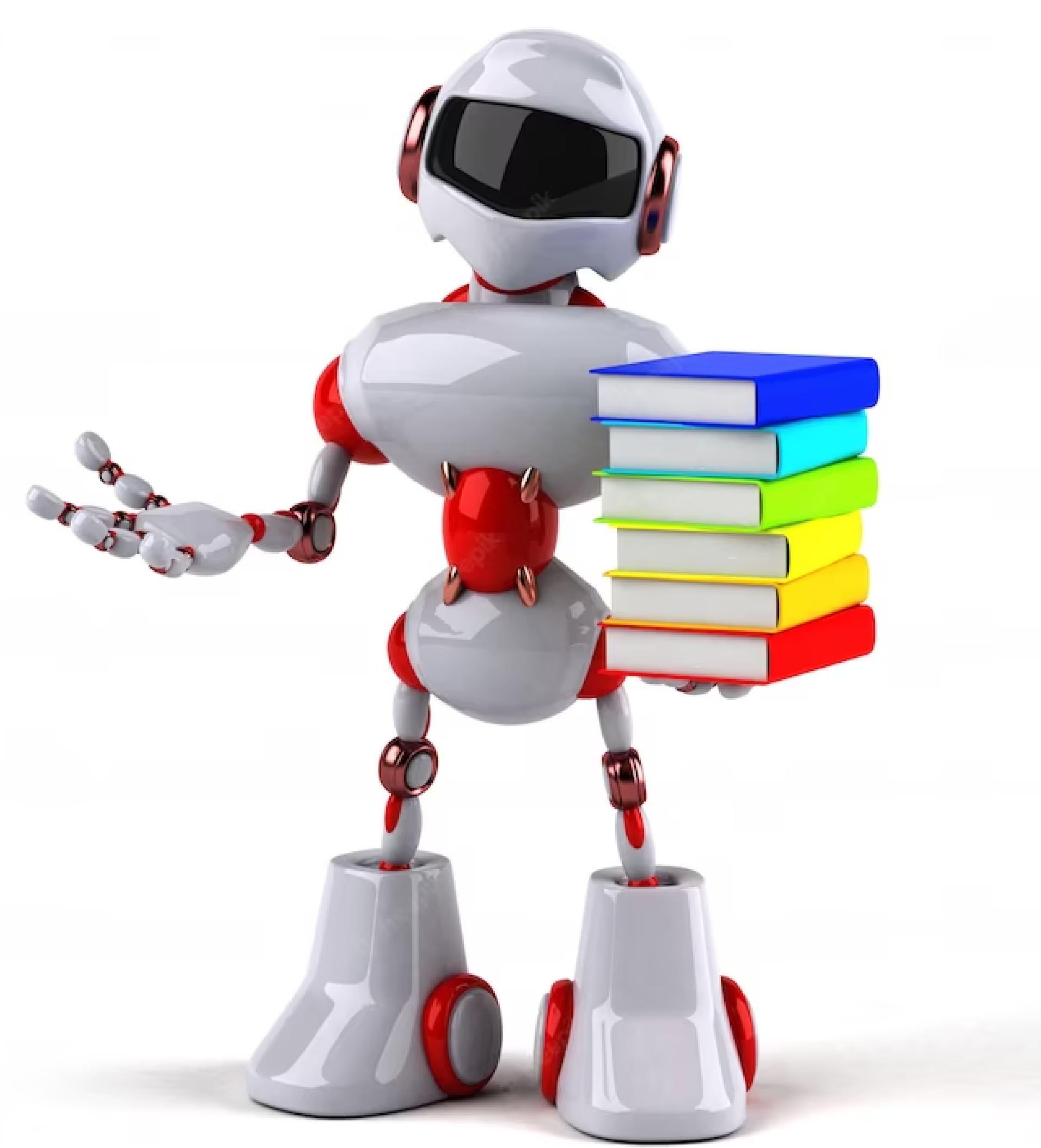
Análisis predictivo

Técnica estadística y de aprendizaje automático para hacer una **predicción** sobre el futuro.

Crea **modelos estadísticos** para predecir un evento o un resultado.

Generar un **modelo analítico** de datos históricos para predecir el comportamiento de la población de la cual proceden los datos muestrales.





Caso de uso

Un ejemplo podría ser con el **clasificador** de mails en spam/no spam: la clasificación de los mails se realiza gracias al modelo predictivo que se construye para **reconocer patrones**.

Estos patrones pueden bien ser la **probabilidad** o **frecuencia de ocurrencia** de cada palabra en los textos de los mails de spam y en los textos de no-spam.

Supuestos

Los modelos **esconden** supuestos

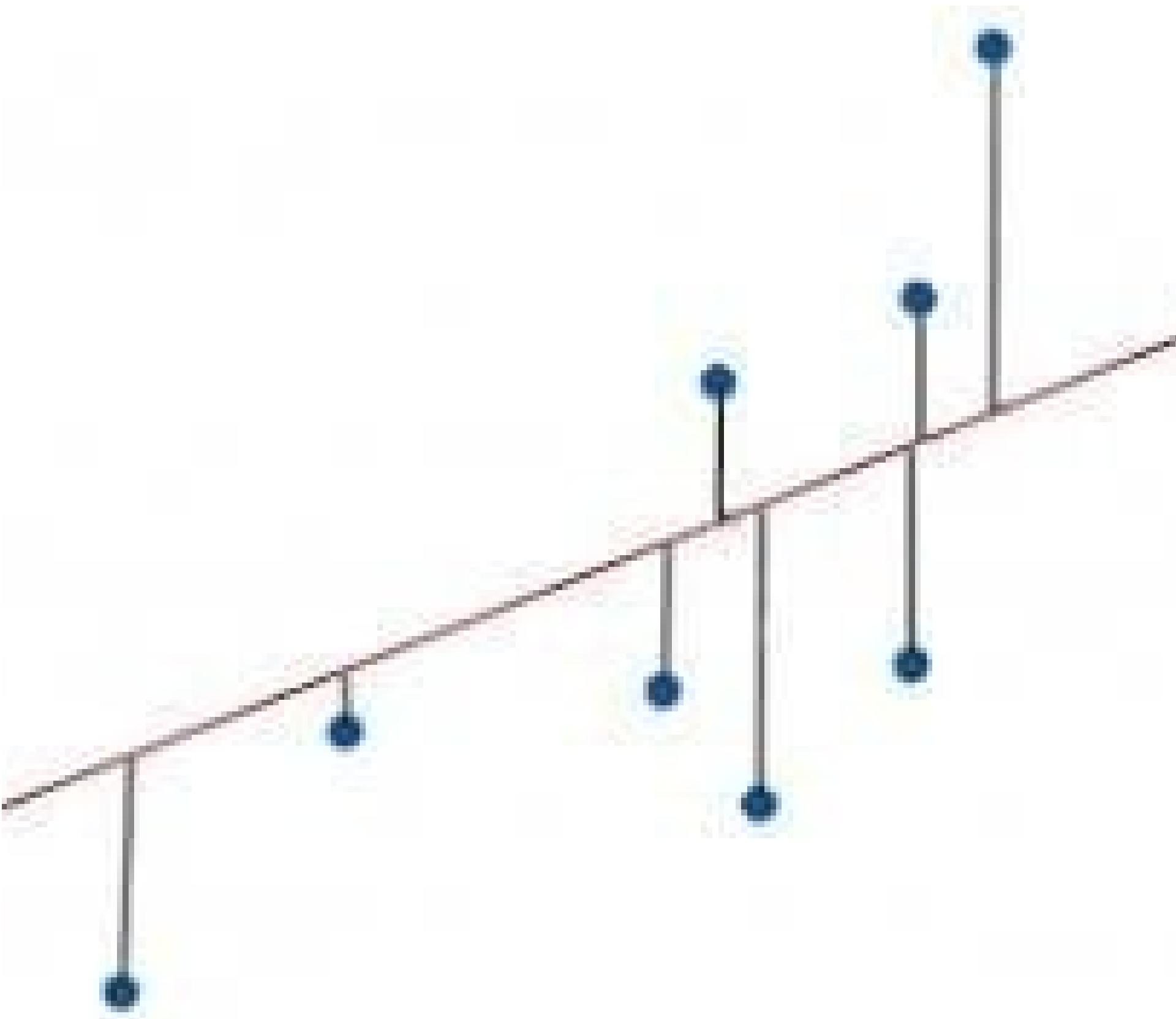
Modelizar es representar la **realidad**
con una cantidad **menor** de
información



Errores

Resulta imposible una eliminación total, ya que implicaría una perfecta identificación del **modelo** con el objeto real. Pero dejaríamos de tener un modelo y tendríamos la **realidad** con toda su complejidad.

$$E = | \text{Valor Real} - \text{Valor Predicho} |$$



Gracias!

