











Librerías para manipulación y análisis de datos



Ingesta y Procesamiento de Archivos

Pandas

Pandas es la librería por excelencia para cargar conjuntos de imágenes.

Permite leer archivos de distinto formato como imágenes, o videos.

Excelente integración con archivos con valores separados por comas CSV

Pickle

Permite serializar cualquier objeto

ldeal para guardar datos no estandarizados

Shutil

Permite mover, copiar, eliminar carpetas y archivos

Permite comprimir archivos en el formato ZIP

Librerías para manipulación y análisis de datos





Numpy

Librería que permite trabajar con imágenes en formato de matrices

Acelera los cálculos numericos

Scipy

Librería orientada al desarrollo científico

Contiene herramientas matemáticas de optimización, distribuciones de probabilidad

Contiene también funciones para el algebra lineal y el procesamiento de señales

OpenCV

Permite procesar imágenes y videos

Contiene implementaciones a los algoritmos mas populares en el área de la visión artificial

Librerías para manipulación y análisis de datos





Introducción a Pandas

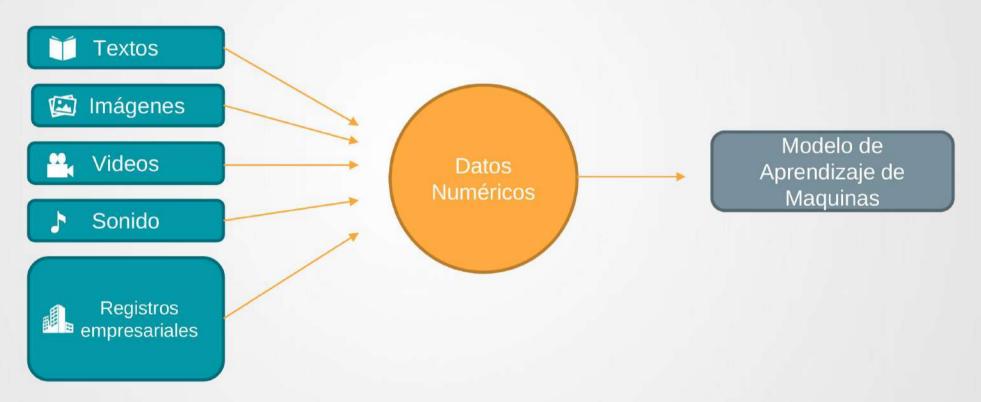
Datos en distintas formas





Datos en distintas formas





Datos en distintas formas



Numérico

Continuo (float): altura, peso, salario, valor del dólar.

Entero (int): unidades vendidas, clientes en una hora, autos en un minuto

Categórico

Ordinal: dominio, nivel de ingreso (bajo, medio, alto), o rango de edad (joven, adulto, mayor)

No ordinal: comuna, profesión, color

Temporal

Una serie de tiempo es una secuencia de dato s indexados y ordenados en el tiempo (día, hora, año, etc).

Algunos ejemplos son: precios de acciones, datos de sensores, etc.

Booleano

Variables que toman dos posibles valores: verdadero o falso.

Algunos ejemplos:

- ¿Es mayor a 65 años?
- ¿Es deudor?
- ¿Contiene cierto objeto?