

## Proyecto 1 - Datos y visualización

---

- 1) Importar el siguiente dataset:

<https://drive.google.com/file/d/1e2Qn8mskSm6EizdAp5gPs1WIYB3mP58W/view?usp=sharing>

Presenta las ganancias de un negocio para cada mes, además de las ventas de cada categoría de productos.

- Graficar las ganancias según el mes e indicar cuál fue el mes de mayor ganancia con un punto.
- Graficar luego las ventas para cada categoría a lo largo del año para visualizar si hay alguna estación en la cual un producto es más rentable.

Usar subplots para mostrar en grupo estos gráficos.

- 2) Para el dataset del ejercicio 1, sumar todas las ventas para cada producto durante el año. Luego hacer un gráfico de torta mostrando el porcentaje de ventas de cada categoría respecto al total.

- 3) Descargar el siguiente dataset:

<https://drive.google.com/file/d/1F8pwUJBAP32ZmUDN5wY-tkF1Ds3aAm8M/view?usp=sharing>

Este contiene información sobre la vacunación cada 100 personas en países alrededor del mundo desde diciembre de 2020.

- Filtrar los datos que se correspondan con Argentina.
- Extraer la información de fecha y vacunación, y graficar una curva que muestre el progreso de la campaña de vacunación.

- 4) Importar el dataset:

<https://raw.githubusercontent.com/selva86/datasets/master/BostonHousing.csv>  
En el mismo hay información sobre viviendas en Boston en 1978. Hay datos como el precio, y otros datos de interés. Más información al respecto puede encontrarse en: [Boston Dataset \(toronto.edu\)](https://www.toronto.edu/boston/dataset.html)

Buscaremos graficar los datos de forma que puedan apreciarse relaciones en los mismos. Probar con distintos gráficos e identificar el que brinda mayor información.

Los datos a graficar son:

- precio (medv) vs crimen (crim)
- precio (medv) vs porcentaje de población con bajos ingresos (lstat)
- precio (medv) vs impuestos (tax)
- precio (medv) vs cantidad de maestros por alumno (ptratio)

Utilizar subplots para mostrar todos los gráficos.

5) La siguiente carpeta:

<https://drive.google.com/drive/folders/1o8QVBnmW1ZrpbkWnaKvDyJxgNjPkDBqr?usp=sharing> contiene información sobre los mundiales de fútbol desde 1930 hasta 2014.

El conjunto de datos WorldCups muestra toda la información sobre todas las Copas del Mundo en la historia, WorldCupMatches muestra todos los resultados de los partidos disputados como parte de las copas, y WorldCupPlayers los jugadores de los diferentes equipos.

- a) Mostrar en un gráfico los países que ganaron cada mundial y el número total de goles que hizo.
- b) Crear un sub-set de datos (dataframe) con todos los partidos que se jugaron en la ciudad de Río de Janeiro. Con este, hacer un gráfico que muestre por cada partido (fecha), el total de goles (sumando los goles del equipo local y el visitante).
- c) Hacer un filtro con los jugadores con el número 10 de camiseta, y de estos, filtrar aquellos que son capitanes (Position = C).

Más info en:

<https://www.kaggle.com/datasets/abecklas/fifa-world-cup?select=WorldCupPlayers.csv>

Todos los gráficos deben tener título.

Elegir el tipo de gráfico que aporte mayor información visual.

Elegir algunos de los gráficos usados en el proyecto y:

- a) Cambiar forma de línea
- b) Cambiar estilo de puntos
- c) Cambiar colores
- d) Añadir grilla en el gráfico