Workshop Regresión Lineal

# Consigna:

**Iniciaremos realizando un breve EDA del dataset:**

1. Importar las librerías necesarias para la realización del desafío (Ej: pandas, numpy, matplotlib, etc) y el dataset propuesto para la ejercitación.

2. Imprimir las 5 primeras filas.

3. Imprimir las 5 últimas filas.

4. Imprimir las dimensiones del dataset.

5. Obtener la totalidad de registros por columnas.

6. Analizar el tipo de dato de cada columna.

7. Obtener el tipo de estructura del dataset. Ej: si es o no un DataFrame.

8. Listar los nombres de las columnas.

**Machine Learning - Modelo de Regresión Lineal, Árboles de Decisión, KNN:**

1. Filtrar el dataset por la columna property\_type con los valores de: "PH", "Casa","Departamento” y el feature 12 == "Capital Federal" (Dato: Nos debe dar un shape de 91485 rows × 19 columns)
2. Quedarnos con aquellas propiedades cuya superficie total es menor igual a 1000 m2 y mayor igual a 15 m2. (Dato: Nos debe dar un shape 82405, 19).
3. Seleccionar aquellas propiedades cuyo precio es menor igual a 4000000 dólares. (Dato: Nos debe dar un shape (82373, 19))
4. Ahora deberemos quedarnos con las columnas: rooms, bedrooms, bathrooms, surface\_total, surface\_covered y price. (Dato: Nos debe dar un shape (82373, 6))
5. Eliminar los valores missings del dataset.
6. Verificar el shape del dataset luego de la eliminación. (Dato: Nos debe dar un shape (81019, 6)).
7. Realizar el análisis de correlación del dataset.
8. Generar los features de X e Y.
9. Separar los datos en Train (70 %) y Test (30 %) con un random\_state=42.
10. Crear un objeto de Regresión Lineal.
11. Entrenar el modelo creado.
12. Realizar la predicción del modelo en Train y Test y calcular el RMSE.
13. Importar las librerías para un modelo de DecisionTreeRegressor y KNeighborsRegressor.
14. Crear un modelo de DecisionTreeRegressor (max\_depth=3, random\_state=42) y KNeighborsRegressor(n\_neighbors=5).
15. Entrenar el modelo de DecisionTreeRegressor.
16. Entrenar el modelo de KNeighborsRegressor.
17. Realizar la predicción del modelo en Train y Test y calcular el RMSE para ambos modelos DecisionTreeRegressor y KNeighborsRegressor.
18. Para pensar: ¿Qué podemos hacer para mejorar nuestro modelo?. Plantear alguna modificación a realizarse para mejorar la performance del modelo.

