

# El conjunto de datos California Housing

En esta actividad se trabajará con el conjunto de datos California Housing. Con el siguiente código se puede leer los datos.

```
import matplotlib.pyplot as plt
from sklearn.datasets import fetch_california_housing
import umap
```

```
california_housing = fetch_california_housing()
X = california_housing.data
y = california_housing.target
X.shape, y.shape
```

((20640, 8), (20640,))

Que contiene las siguientes características:

```
print(california_housing.feature_names)
```

['MedInc', 'HouseAge', 'AveRooms', 'AveBedrms', 'Population', 'AveOccup', 'Latitude', 'Longitude']

Con este conjunto de datos se deberán proyectar los datos en dos dimensiones usando el precio y como color. Se usarán los métodos

- PCA
- UMAP
- t-SNE
- Autoencoders

Comparar los resultados