Laboratorio 1: Análisis Exploratorio, Clustering, PCA y A

- Eric Mendoza 15002
- Marlon Fuentes 15240

Preparación de entorno de ejecución

Carga de librerías

```
import os
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from scipy import stats
import copy
```

```
from sklearn.decomposition import PCA
from sklearn import preprocessing
from sklearn.preprocessing import StandardScaler,Normalizer,MinMaxScaler, RobustScaler
from sklearn.manifold import TSNE
from sklearn.cluster import KMeans
```

Carga de data

Cargar test.csv desde github.

```
url = 'https://raw.githubusercontent.com/fuentesmarlon/CLUSTERING-PCA-APRIORI/master/trai
```

1. Análisis exploratorio

Se tienen las siguientes columnas.

```
Index(['Id', 'MSSubClass', 'MSZoning', 'LotFrontage', 'LotArea', 'Street',
       'Alley', 'LotShape', 'LandContour', 'Utilities', 'LotConfig',
       'LandSlope', 'Neighborhood', 'Condition1', 'Condition2', 'BldgType',
       'HouseStyle', 'OverallQual', 'OverallCond', 'YearBuilt', 'YearRemodAdd'
       'RoofStyle', 'RoofMatl', 'Exterior1st', 'Exterior2nd', 'MasVnrType',
       'MasVnrArea', 'ExterQual', 'ExterCond', 'Foundation', 'BsmtQual',
       'BsmtCond', 'BsmtExposure', 'BsmtFinType1', 'BsmtFinSF1',
       'BsmtFinType2', 'BsmtFinSF2', 'BsmtUnfSF', 'TotalBsmtSF', 'Heating',
       'HeatingQC', 'CentralAir', 'Electrical', '1stFlrSF', '2ndFlrSF',
       'LowQualFinSF', 'GrLivArea', 'BsmtFullBath', 'BsmtHalfBath', 'FullBath'
       'HalfBath', 'BedroomAbvGr', 'KitchenAbvGr', 'KitchenQual',
       'TotRmsAbvGrd', 'Functional', 'Fireplaces', 'FireplaceQu', 'GarageType'
       'GarageYrBlt', 'GarageFinish', 'GarageCars', 'GarageArea', 'GarageQual'
```