import pandas as pd

df = pd.read\_csv('https://archive.ics.uci.edu/ml/machine-learning-databases/housing/housing.data',header=None,sep='\s+')

df.columns=['CRIM','ZN','INDUS','CHAS','NOX','RM','AGE','DIS','RAD','TAX','PTRATIO','B','LSTAT','MEDV']

df.head()

VERIFIQUE CORRELACION DE 5 VARIABLES, EL TARGET EL MEDV

import seaborn as sns

import matplotlib.pyplot as plt

|  |  |
| --- | --- |
| sns.set(style='whitegrid',context='notebook')  cols = ['DIS','INDUS','CRIM','RM','MEDV']  sns.pairplot(df[cols], height=2.5) | cm =df[cols].corr()  sns.heatmap(cm, cbar=True,              annot=True,              square=True,              fmt='.1f',              annot\_kws={'size':15},              yticklabels=cols,  xticklabels=cols);  Se pone ‘;’ para eliminar comentarios |
|  | A picture containing chart  Description automatically generated |

**INTERPRETAR**

RM : PROMEDIO DE CASAS, TIENE UNA ALTA CORRELACION CON LOS PRECIOS DE LAS CASAS 0.7

UNA CORRELACION NEGATIVA, DIS CON INDUS, -0.71