Import a dataset from **http://www.ats.ucla.edu/stat/data/binary.csv**. Do the Logistic Regression to find out relation between variables that are affecting the admission of a student in a institute based on his or her GRE score, GPA obtained and rank of the student. Also check the model is fit or not. Apply regression Model techniques to predict the data on above dataset

<https://notebook.community/rayjustinhuang/DataAnalysisandMachineLearning/Logistic%20Regression>

import numpy as np

import pandas as pd

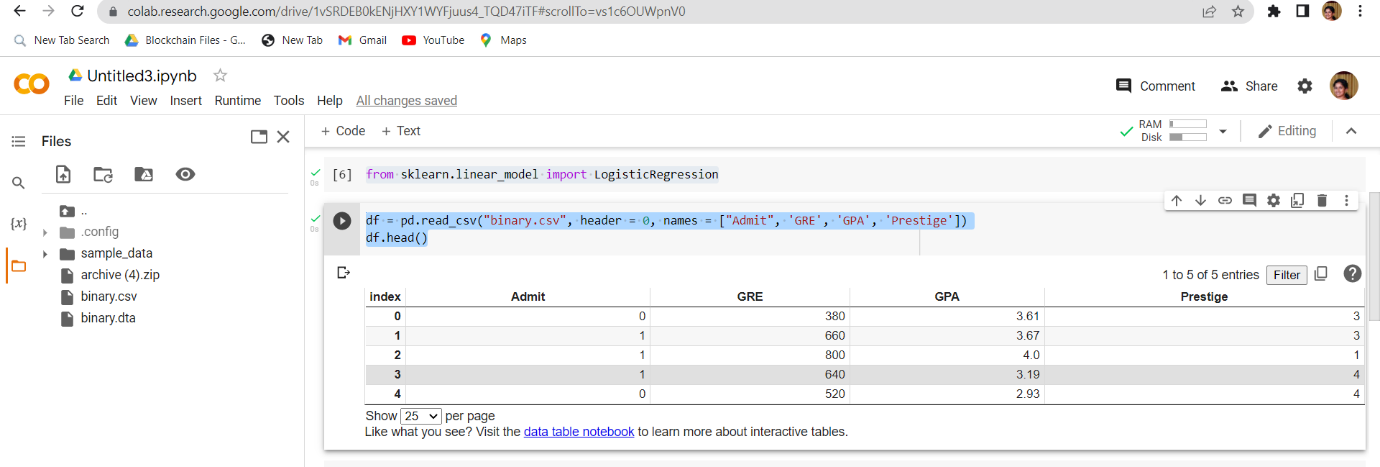
import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.linear\_model import LogisticRegression

df = pd.read\_csv("binary.csv", header = 0, names = ["Admit", 'GRE', 'GPA', 'Prestige'])

df.head()

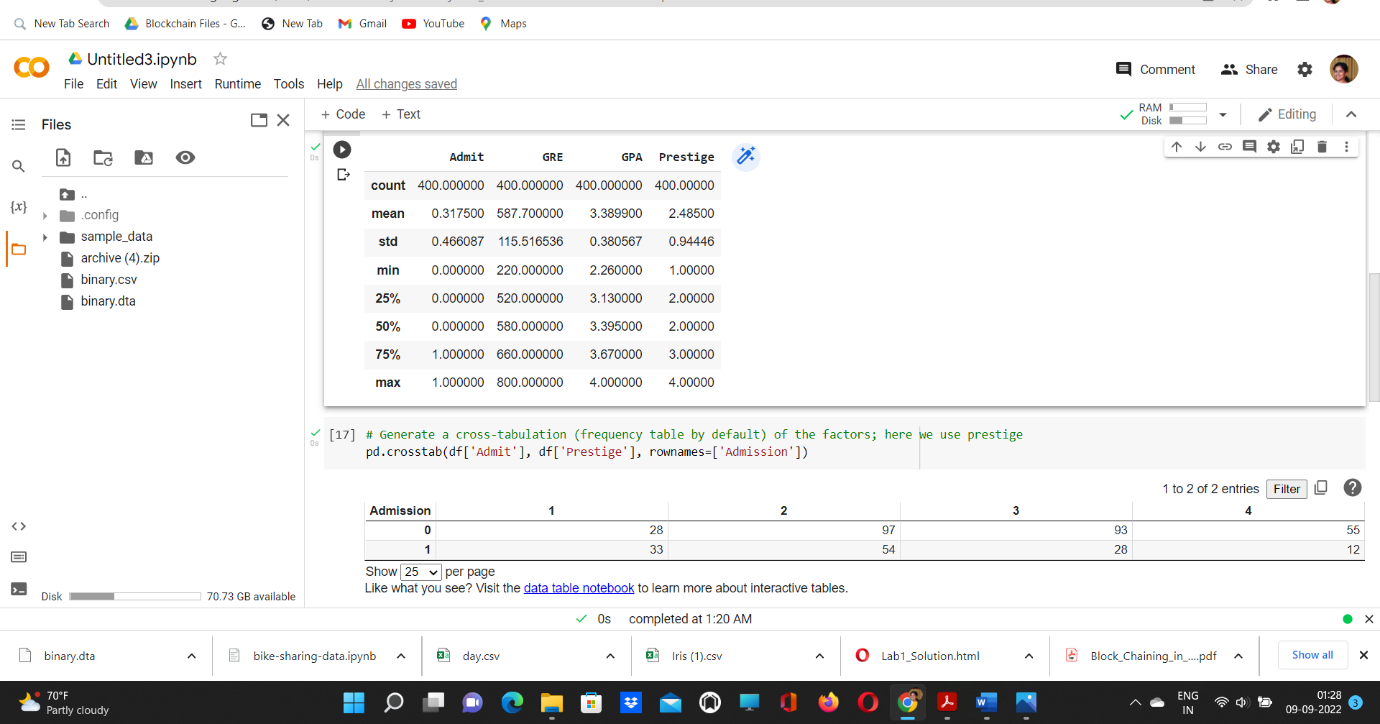


# Basic summary of the data

df.describe()

# Generate a cross-tabulation (frequency table by default) of the factors; here we use prestige

pd.crosstab(df['Admit'], df['Prestige'], rownames=['Admission'])



# Generate histograms

sns.set\_color\_codes('muted')

df.hist(color='g')

plt.show()

