# -\*- coding: utf-8 -\*-

"""

Created on Mon Jul 2 15:27:32 2018

@author: shubham b thorat

"""

# -\*- coding: utf-8 -\*-

"""

Created on Thu Jun 28 16:51:21 2018

Indian-Institutes\_ranking

@author: shubham b thorat

"""

# importing the libraries

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

# importing the data set

dataset=pd.read\_csv('Indian\_college\_ranking.csv', sep = "\t")

dataset2=pd.read\_csv('VIT\_data.csv')

X = dataset.iloc[:,1:5].values

Y = dataset.iloc[:,0].values

X\_VIT = dataset2.iloc[:,1:5].values

# dividing dataset into test and training dataset

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X,Y, test\_size = 1/5, random\_state = 1)

# feature scaling

from sklearn.preprocessing import StandardScaler

sc\_X = StandardScaler()

X\_train = sc\_X.fit\_transform(X\_train)

X\_test = sc\_X.fit\_transform(X\_test)

X\_VIT = sc\_X.fit\_transform(X\_VIT)

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression()

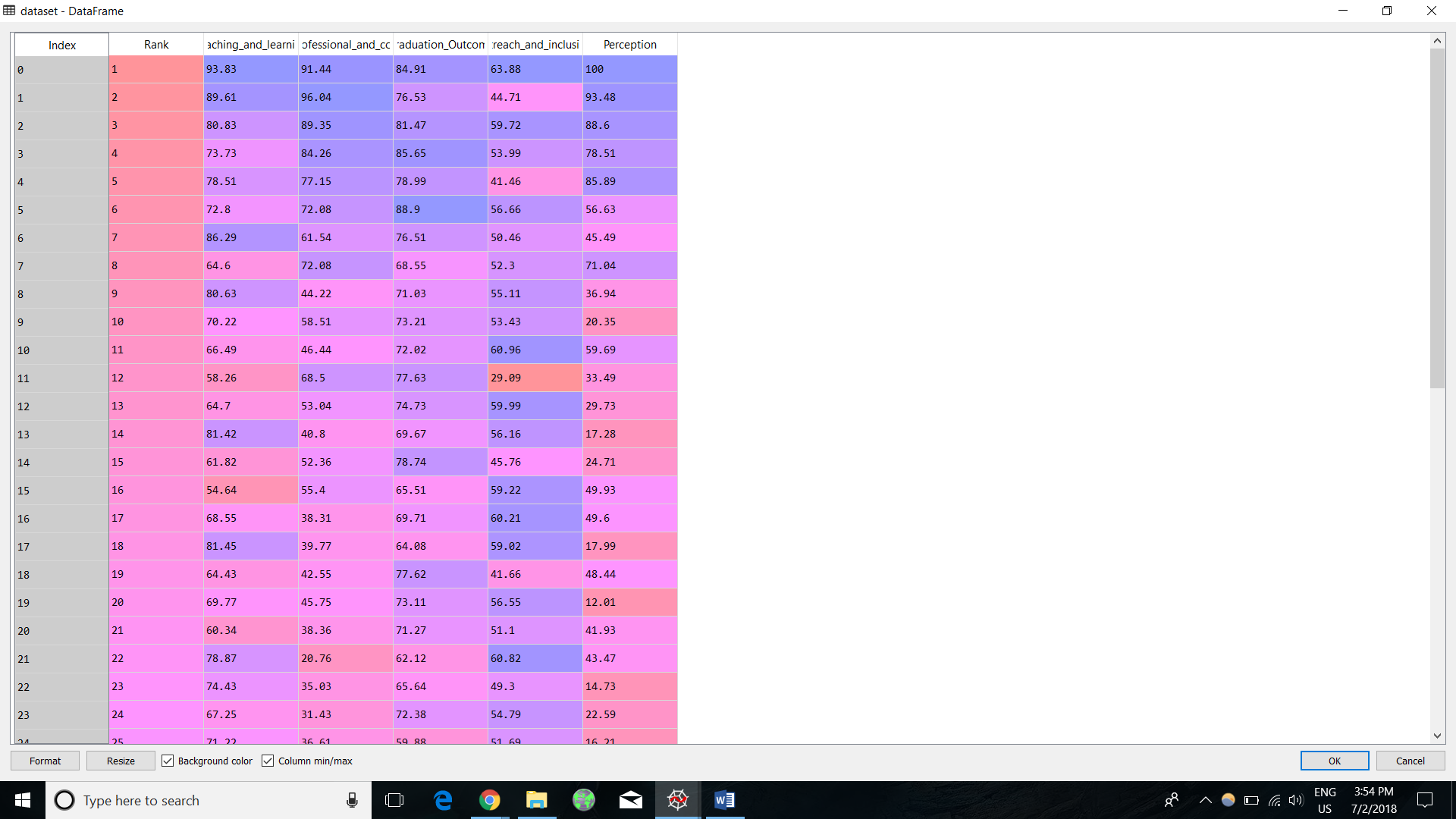
regressor.fit(X\_train,Y\_train)

#predicting the test set result

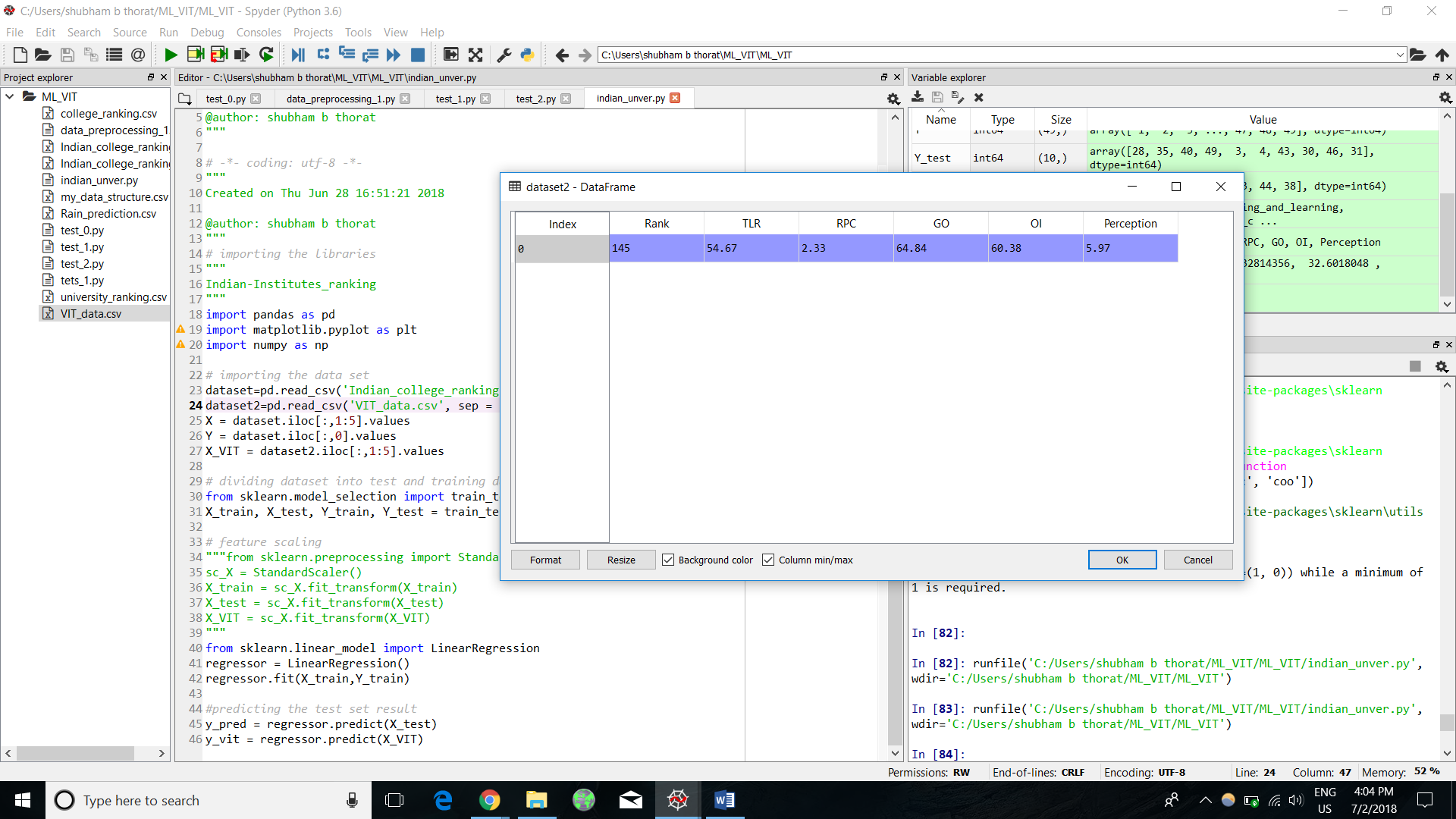
y\_pred = regressor.predict(X\_test)

y\_vit = regressor.predict(X\_VIT)

data\_sheet =

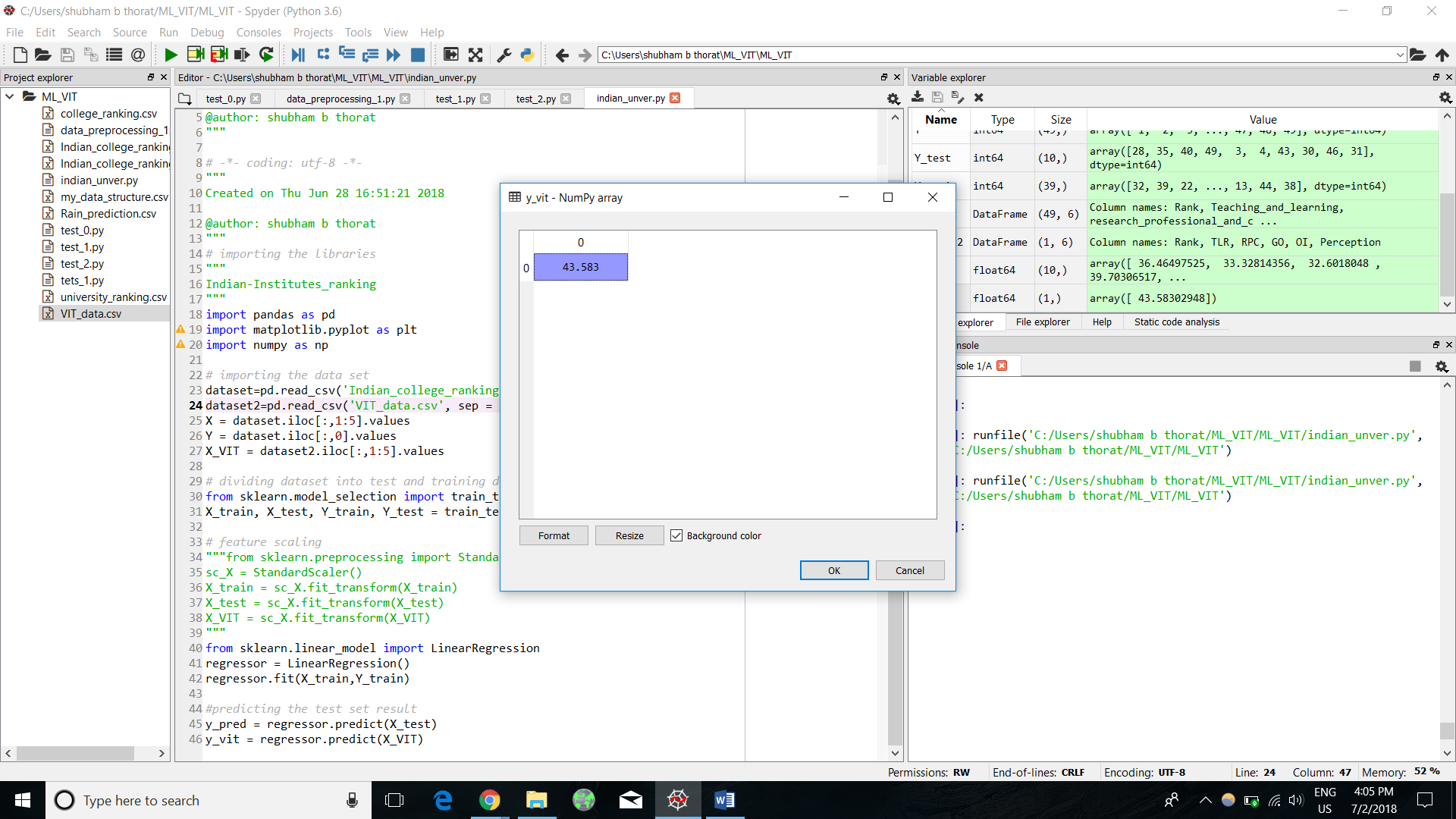


Datasheet2 = (Info Of VIT)



Rank\_of\_VIT

Y\_VIT = 43.583



# checking P values

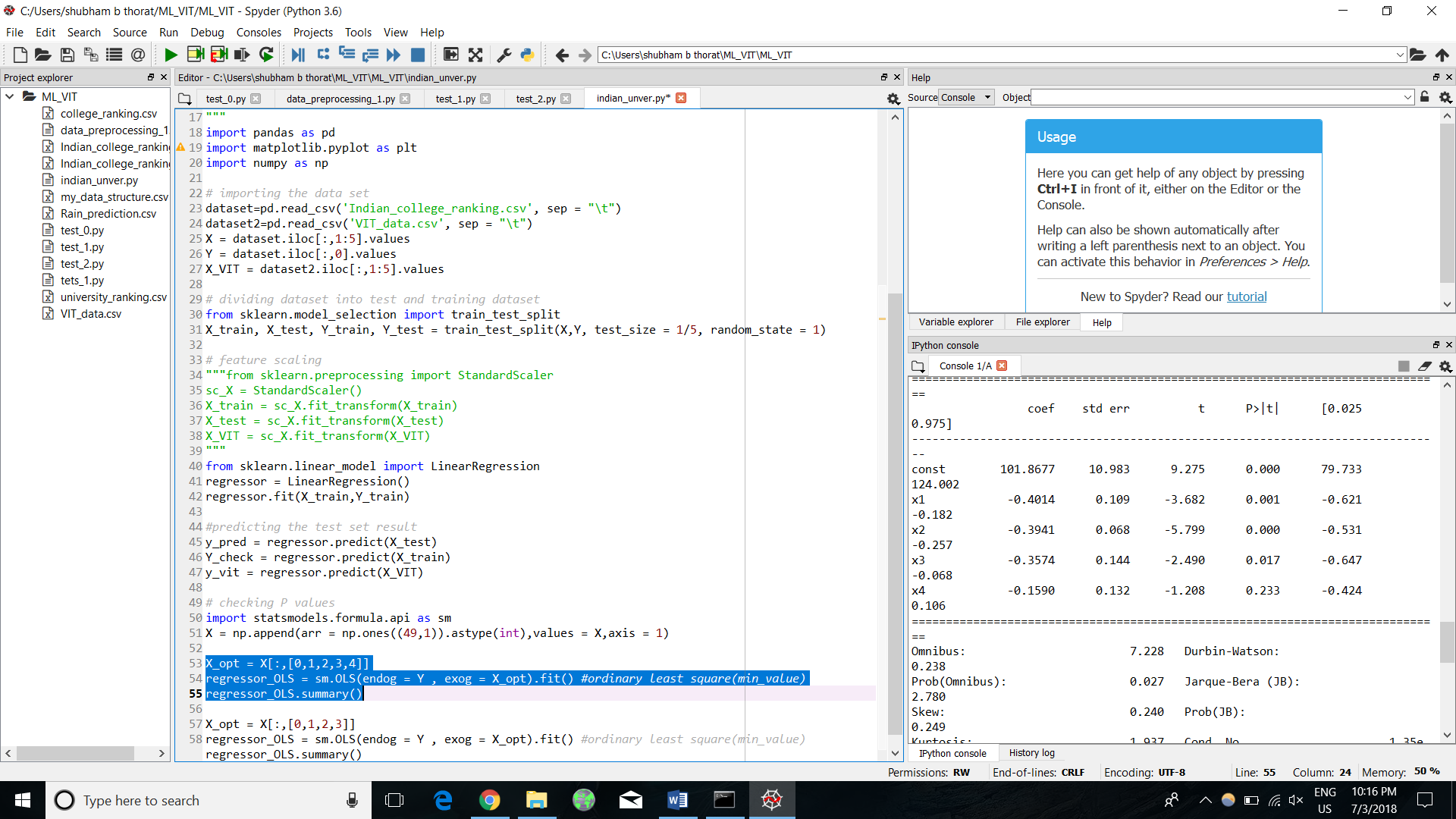
import statsmodels.formula.api as sm

X = np.append(arr = np.ones((49,1)).astype(int),values = X,axis = 1)

X\_opt = X[:,[0,1,2,3,4]]

regressor\_OLS = sm.OLS(endog = Y , exog = X\_opt).fit() #ordinary least square(min\_value)

regressor\_OLS.summary()



X\_opt = X[:,[0,1,2,3]]

regressor\_OLS = sm.OLS(endog = Y , exog = X\_opt).fit() #ordinary least square(min\_value)

regressor\_OLS.summary()

