Attrition assignment

* Import packages and show 5 rows value

import pandas as pd

dataset=pd.read\_csv("general\_data.csv")

dataset.head()

Age Attrition ... YearsSinceLastPromotion YearsWithCurrManager

0 51 No ... 0 0

1 31 Yes ... 1 4

2 32 No ... 0 3

3 38 No ... 7 5

4 32 No ... 0 4

* Droping “Nan” null values

Data=dataset.dropna()

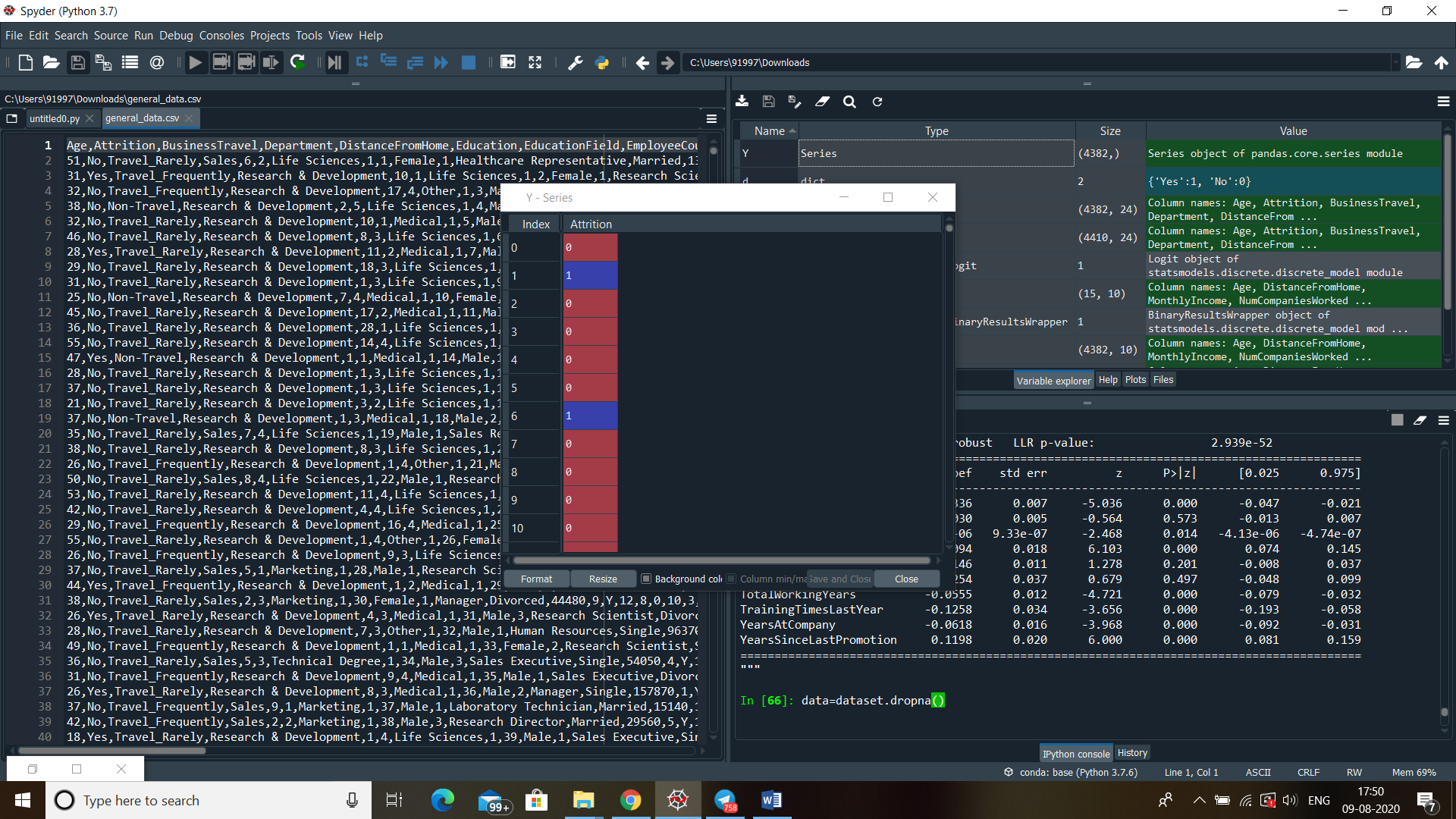
* Getting dependent variable

y=dataset.attrition

now setting values at it will give an error in the multiple regression model (cause string creating problem while performing)

d={“Yes”: 1, “No”: 0}

Y=y.replace(d)



* Getting independent variables

x=dataset[['Age','DistanceFromHome','MonthlyIncome','NumCompaniesWorked','PercentSalaryHike','StandardHours','TotalWorkingYears','TrainingTimesLastYear','YearsAtCompany','YearsSinceLastPromotion']]

* Import statsmodal api

Import statsmodels.api as sm

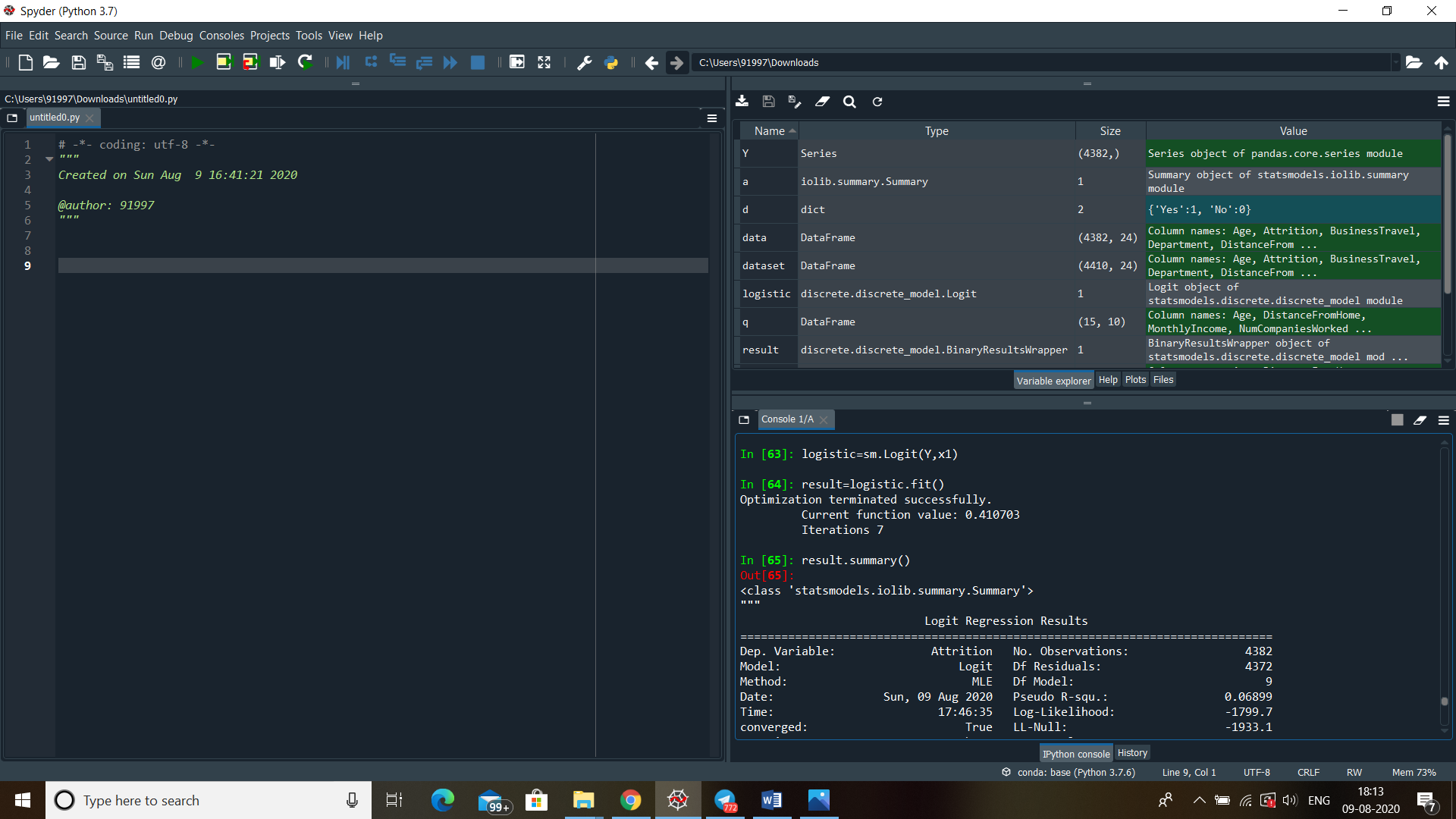
* Add constants

X1=sm.add\_constant(x)

* Performing Logit() method

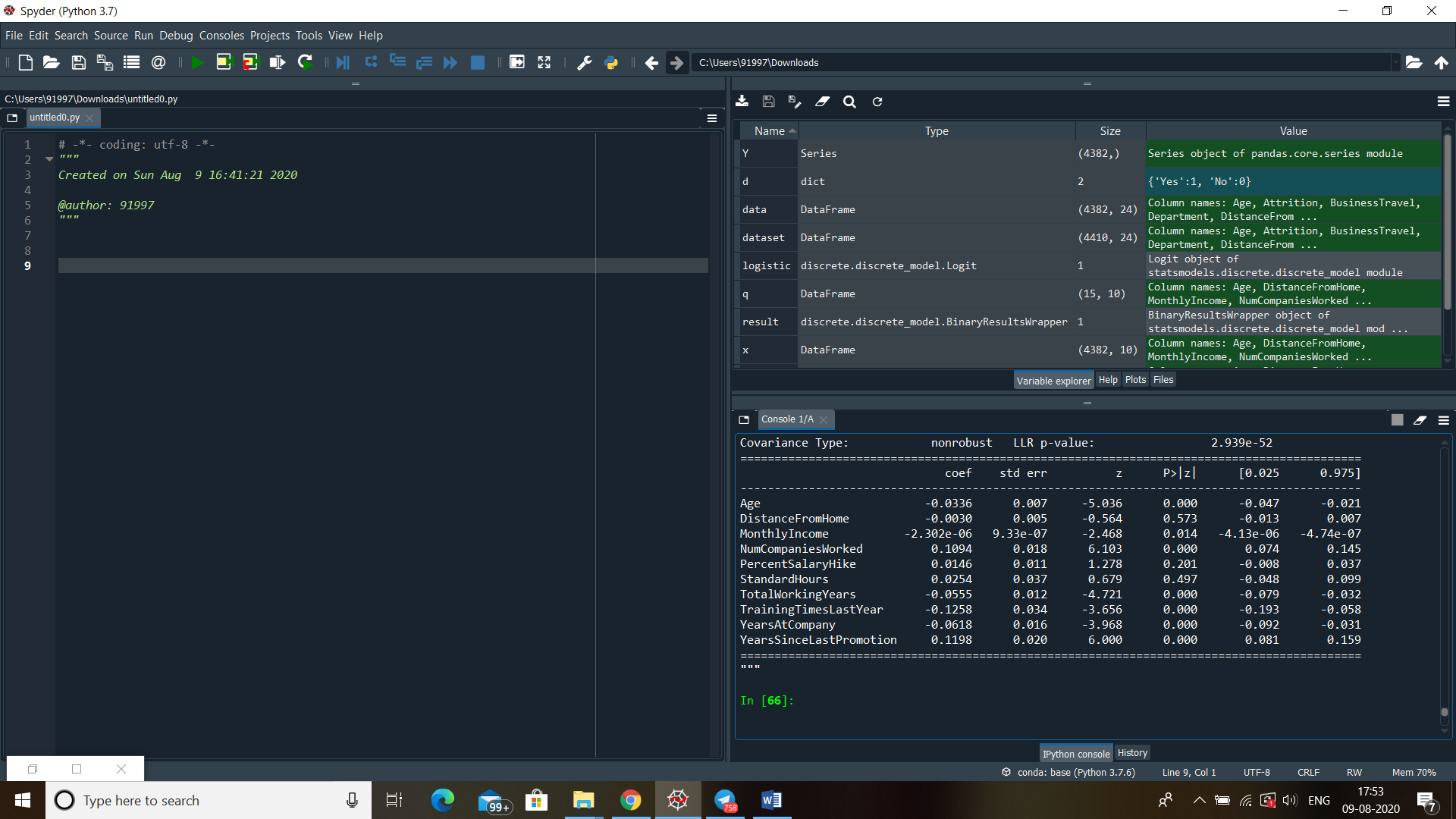
Logistic =sm.logit(Y,X1)

result=logistic.fit()



* Displaying summary

result.summary()



* Significant variables use form equations are

Age , MonthlyIncome , NumCompaniesWorked , PercentSalaryHike , StandardHours

TotalWorkingYears , TrainingTimesLastYear YearsAtCompany YearsSinceLastPromotion

* Now use coefficient of variables to form an equation (equation image)

