import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

data\_set = pd.read\_csv("50\_Startups.csv")

---------------------------------------------------------------------------------------------------------------------

X = data\_set.iloc[:, :-1].values

y = data\_set.iloc[:, -1].values

------------------------------------------------------------------------------------------------------------------------------------------

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

label\_encoder = LabelEncoder()

X[:,3] = label\_encoder.fit\_transform(X[:,3])

------------------------------------------------------------------------------------------------------------------------------------------

one\_hot\_encoder = OneHotEncoder(categorical\_features = [3])

X = one\_hot\_encoder.fit\_transform(X).toarray()

------------------------------------------------------------------------------------------------------------------------------------------

# avoid dummy trap variable

X = X[:, 1:]

------------------------------------------------------------------------------------------------------------------------------------------

# split data to train and test data

from sklearn.cross\_validation import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state = 0)

------------------------------------------------------------------------------------------------------------------------------------------

**Multiple Linear Regression**

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression()

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

# predicting test result

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

------------------------------------------------------------------------------------------------------------------------------------------

**Backward elimination**

import statsmodels.forumula.api as sm

------------------------------------------------------------------------------------------------------------------------------------------

Append a column of ones to a matrics

X = np.append(X, values = np.ones((50, 1).astype(int))