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

"""

Created on Sun Jan 19 18:19:43 2020

@author: Dell 5530

"""

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

# Importing the libraries

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

# Importing the dataset

dataset = pd.read\_csv('global\_co2.csv')

X = dataset.iloc[220:,:1].values

Y = dataset.iloc[220:,1:2].values

# Splitting the dataset into the Training set and Test set

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

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size = 1/3, random\_state = 0)

# Fitting Simple Linear Regression to the Training set

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression()

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

# Predicting the Test set results

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

# Fitting Polynomial Regression to the dataset

from sklearn.preprocessing import PolynomialFeatures

poly\_reg = PolynomialFeatures(degree = 6)

X\_poly = poly\_reg.fit\_transform(X)

poly\_reg.fit(X\_poly, Y)

lin\_reg\_2 = LinearRegression()

lin\_reg\_2.fit(X\_poly, Y)

# Visualising the Data set to see the relation between X and Y

plt.scatter(X\_train, Y\_train, color = 'brown')

plt.plot(X\_train, regressor.predict(X\_train), color = 'black')

plt.title(' CO2 production from 1970s (Training set)')

plt.xlabel('Years ')

plt.ylabel('CO2 Production')

plt.show()

# Visualising the Test set results

plt.scatter(X\_test, Y\_test, color = 'brown')

plt.plot(X\_train, regressor.predict(X\_train), color = 'black')

plt.title(' CO2 production from 1970s (Training set)')

plt.xlabel('Years ')

plt.ylabel('CO2 Production')

plt.show()

#since the data is non linear, Polynomial egression will be applied

# Visualising the Polynomial Regression results

plt.scatter(X, Y, color = 'red')

plt.plot(X, lin\_reg\_2.predict(poly\_reg.fit\_transform(X)), color = 'black')

plt.title('CO2 production from 1970s')

plt.xlabel('Years')

plt.ylabel('CO2 Production')

plt.show()

print('CO2 production in 2011 is')

print(regressor.predict([[2011]]))

print('CO2 production in 2012 is')

print(regressor.predict([[2012]]))

print('CO2 production in 2013 is')

print(regressor.predict([[2013]]))

checking accuracy

from sklearn import metrics

print('Mean Absolute Error:', metrics.mean\_absolute\_error(Y\_test, Y\_pred))

print('Mean Squared Error:', metrics.mean\_squared\_error(Y\_test, Y\_pred))

print('Root Mean Squared Error:', np.sqrt(metrics.mean\_squared\_error(Y\_test, Y\_pred)))