EXPLANATION OF THE PYTHON CODE USED FOR THE PREDICTION.

This is python code to import the necessary libaries for analysis

%matplotlib inline

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

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

from sklearn.linear\_model import LinearRegression

import os

import pandas

import matplotlib.pyplot as plt

import numpy as np

......................................................................

DATA = pandas.read\_csv("cw1data.csv")

This is python code used to read csv data

..................................................................

print(DATA.shape) # Display the total rows and colums of cw1data.csv

# Getting the scatter plot for DATA(x2) in respect to DATA(y)

plt.scatter(DATA['x2'], DATA['y'])

plt.xlabel('x2')

plt.ylabel('Focus point(y)')

plt.title('Visualizing the Data of y in respect to x2')

plt.show()

....................................................................................

x = DATA.drop(['y'],axis=1) # use others as inputs for prediction

y = DATA['x2'] # set cw1data as output for prediction

............................................................................................

python code to split test dataset and train dataset

x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size=0.2, random\_state=42)

........................................................................................................

from sklearn.linear\_model import Lasso # lasso model prediction

lasso\_model = Lasso(alpha = 1.0) # set penalty value

lasso = lasso\_model.fit(X\_train , Y\_train)

lasso\_predict = lasso.predict(X\_test)

.........................................................................................

import numpy as np

error = np.mean(np.abs((Y\_test - lasso\_predict)))

print('error is', error)

python code to predict the error in the predicted model

..............................................................................................

# Mean absolute percentage error

import numpy as np

def MAPE(Y\_actual,Y\_Predicted):

mape = np.mean(np.abs((Y\_actual - Y\_Predicted)/Y\_actual))\*100

return mape

....................................................................................

Accuracy = 100 - Lasso\_MAPE

print('Accuracy of Lasso Regression: {:0.2f}%.'.format(Accuracy))

python code to determine the percentage of accuracy

...................................................................................................

ANALYSIS OF THE OUTCOME OF THE PREDICTION.

From my observation of the dataset of (CW1DATA.CSV), the numbers of rows and colums are 135 and 14.

I was able to visualize the data using the scatter plot linear regression model using y as my input variables for prediction and x4 as output variable, i observed the increase nature of x4 in respect to y using the scatter plot that i derived.

I finaly used lasso regression model to calculate the errors in my prediction and my error was , firstly i split the data into training and testing dataset ( x4, y ). then the rows and colums reduce to 108 rows 13 columns which was used for the prediction.

Using Lasso regression model to calculate the errors in the prediction i got

(error = 3.4369404975237923) , then considering the level of accuracy of my prediction i got Accuracy of Lasso Regression: 93.64%.

So because of my level of accuracy i believe that my predicted data is a strong prediction because of the level of accuracy.