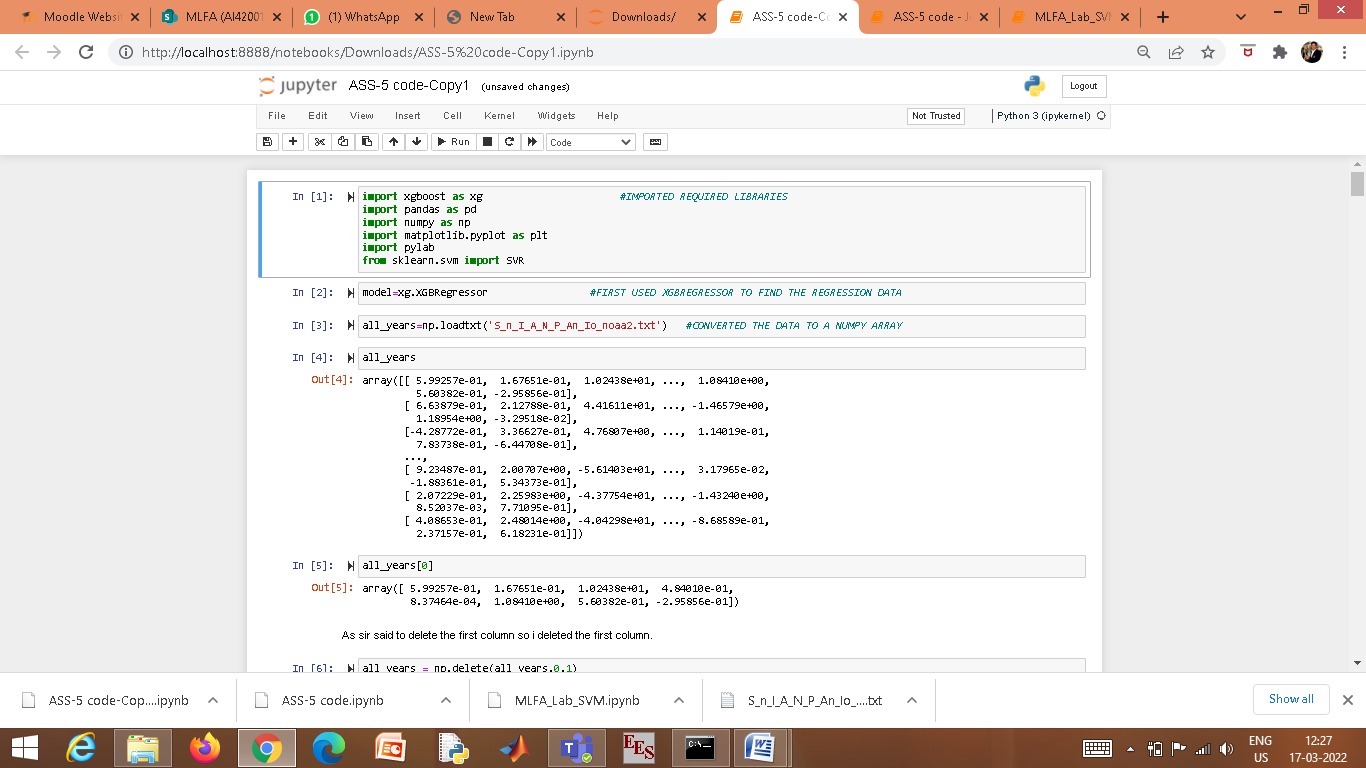
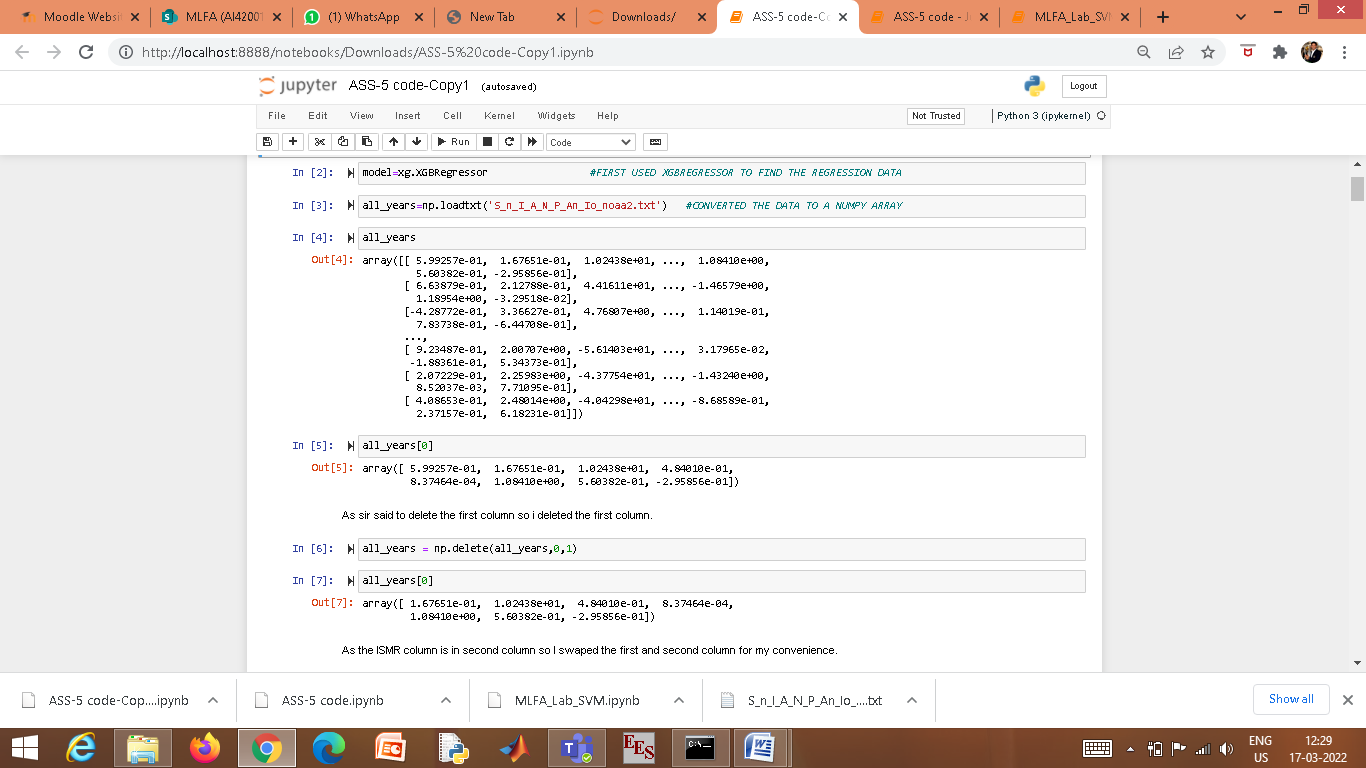
**MLFA ASSIGNMENT 5 OUTPUT FILE**

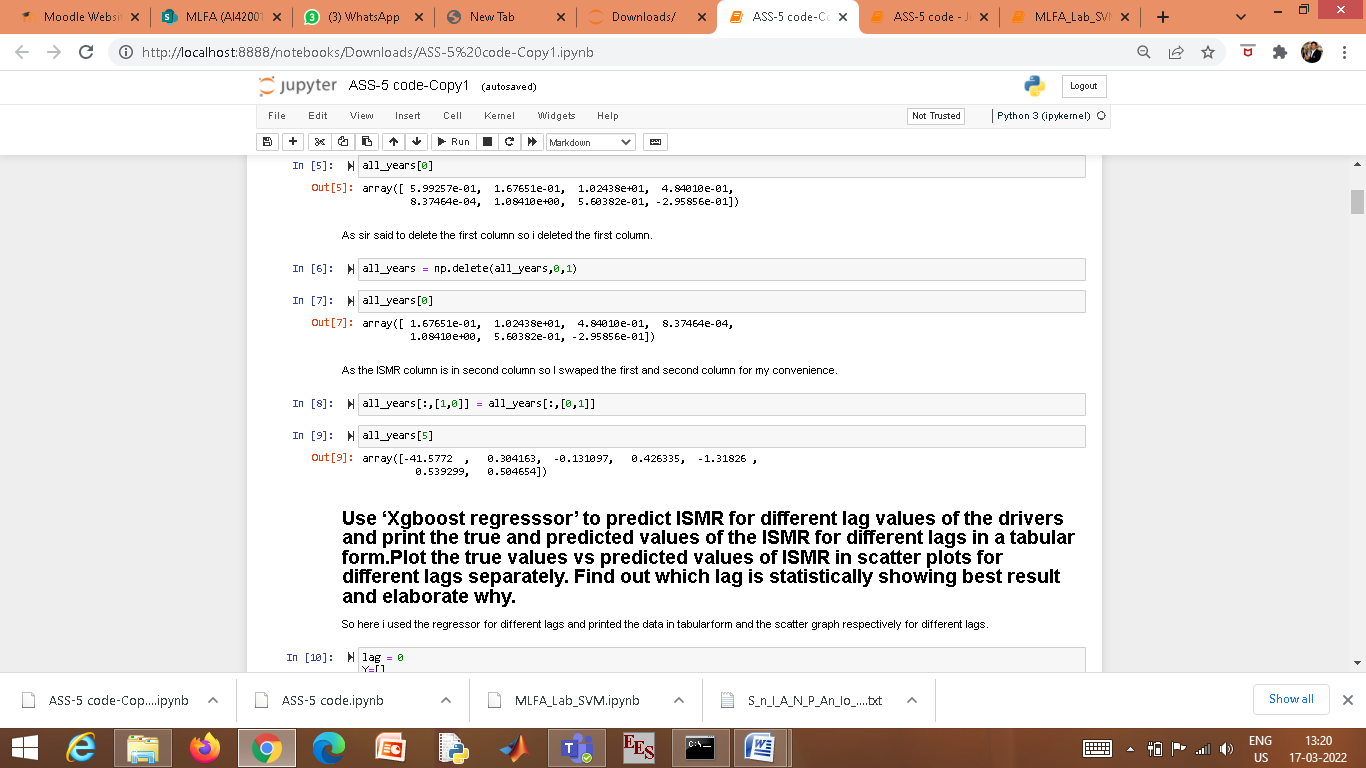
* Importing required libraries:



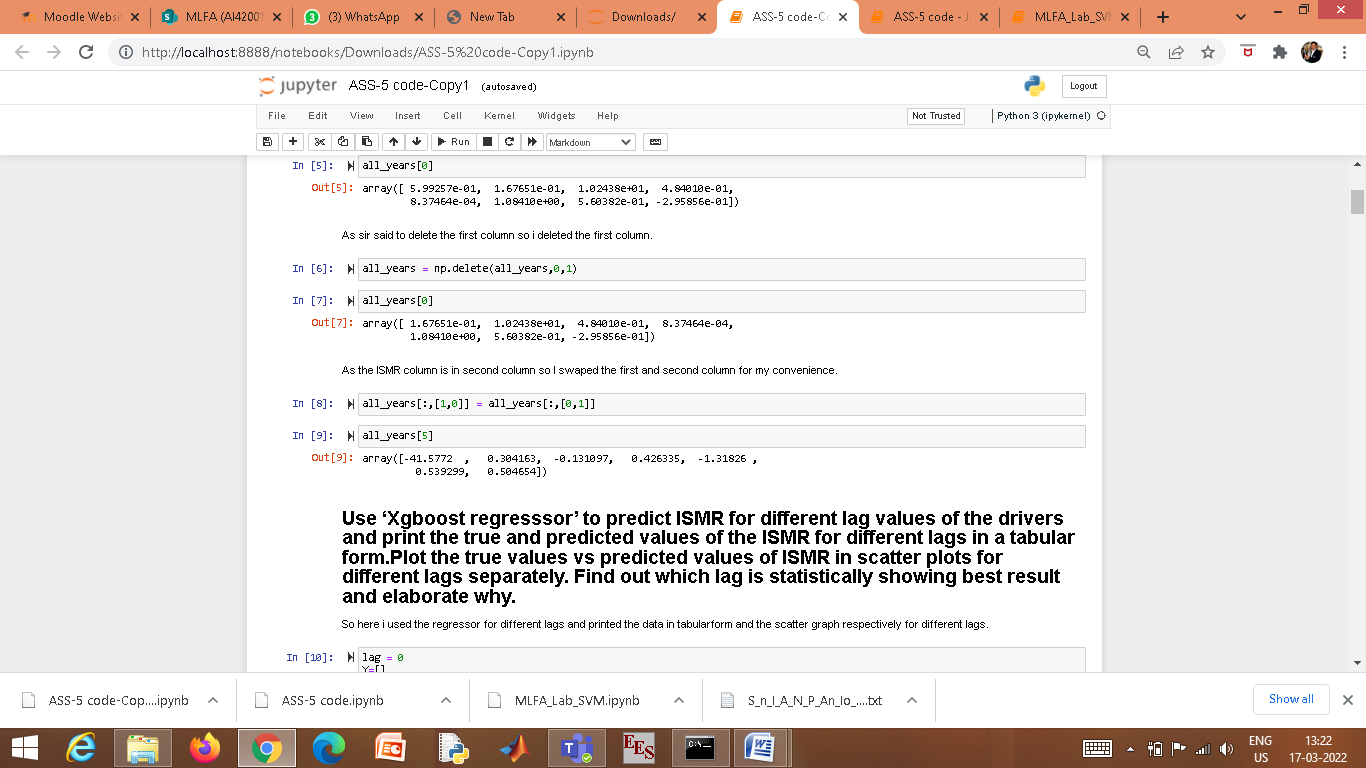
* Use of XG Boost Regressor and conversion of given data to a numpy array



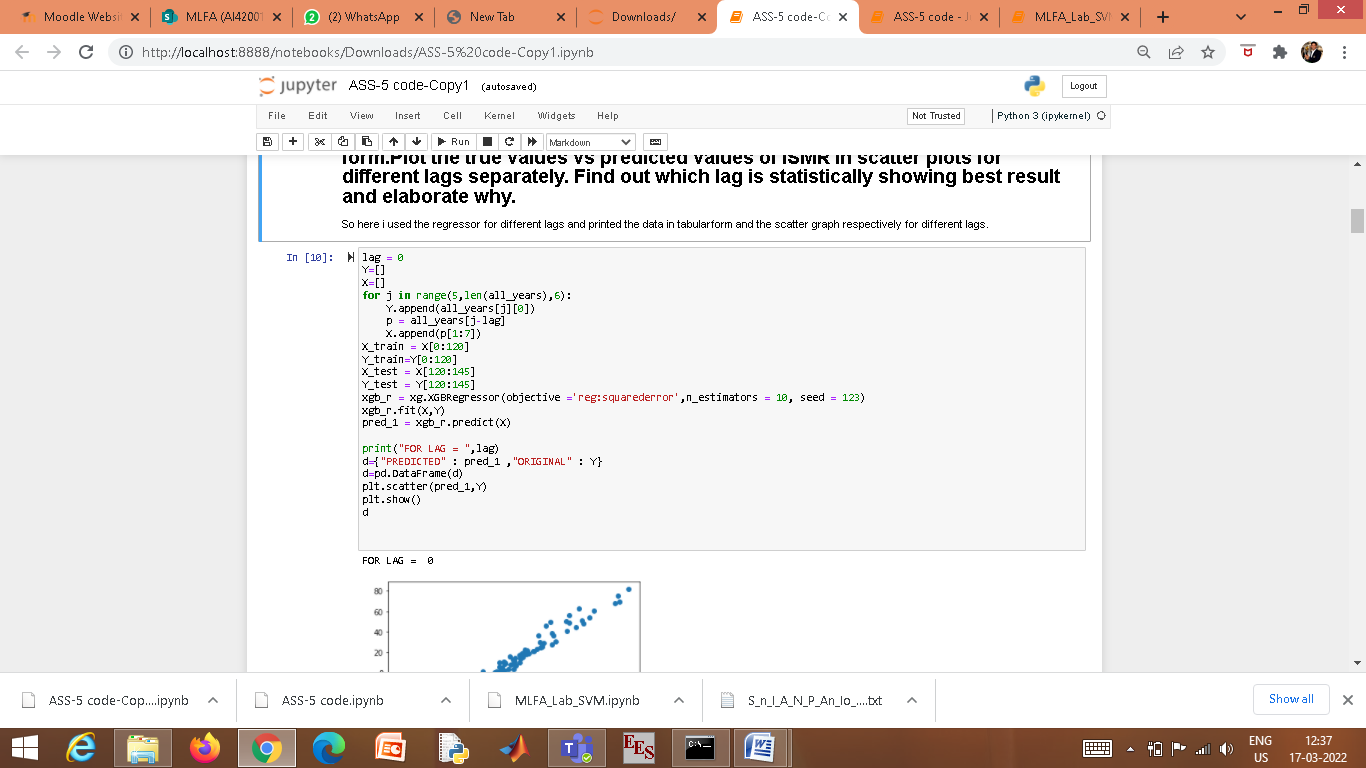
* Deleting the first column

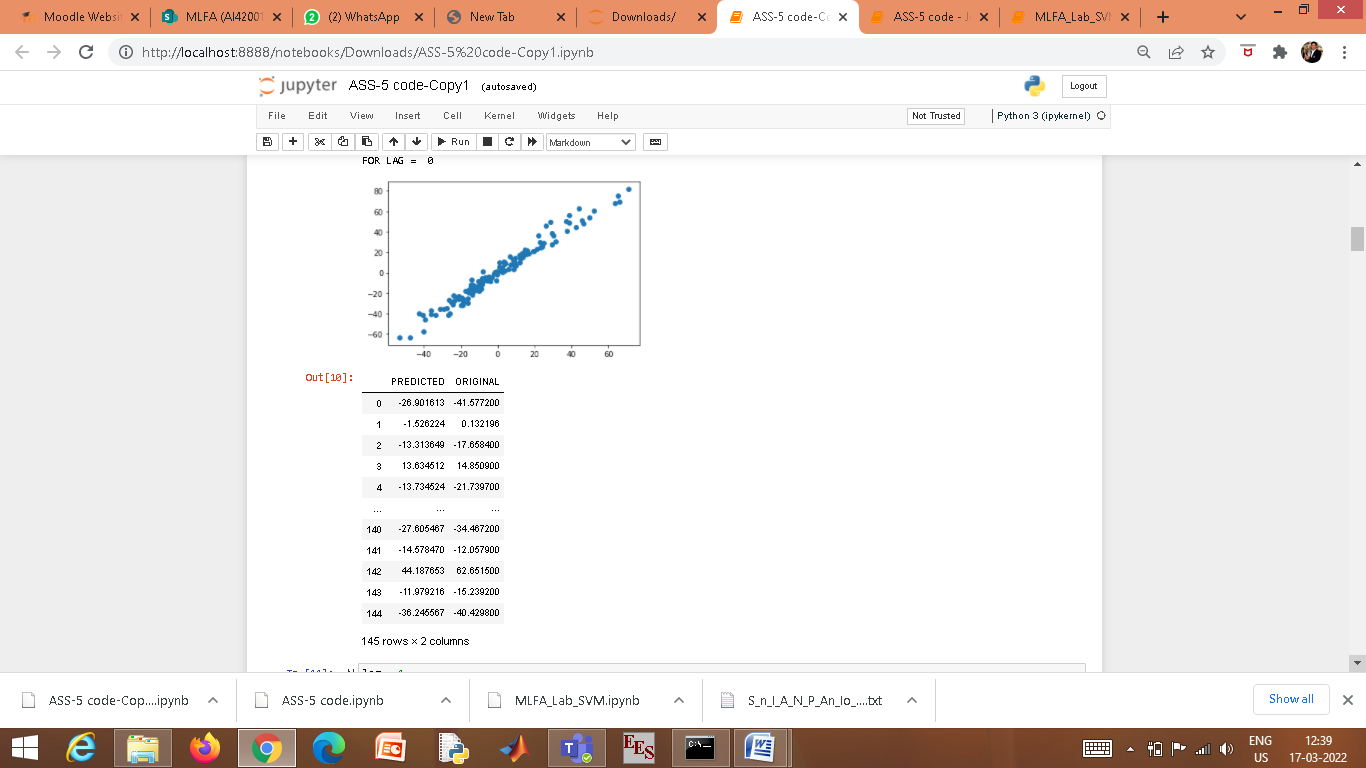


* Swapping the first two columns for convenience

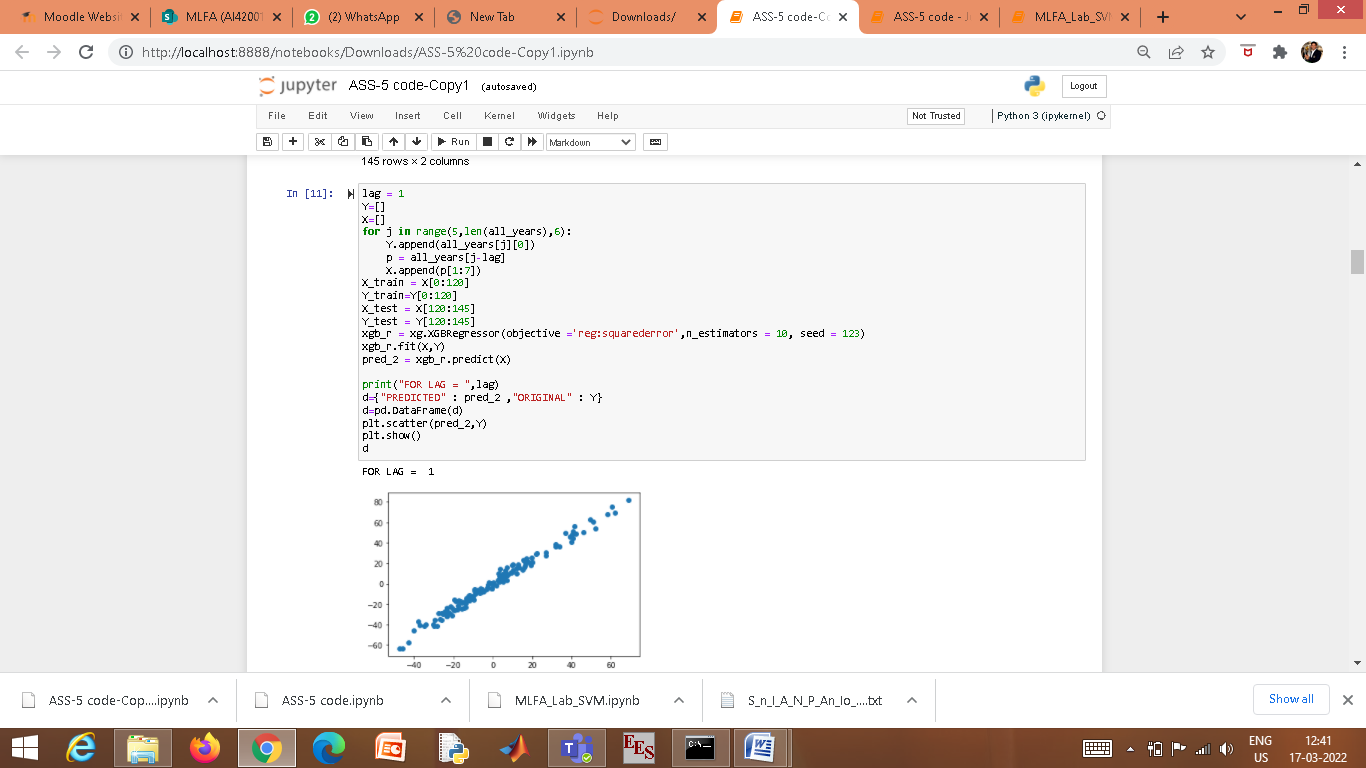


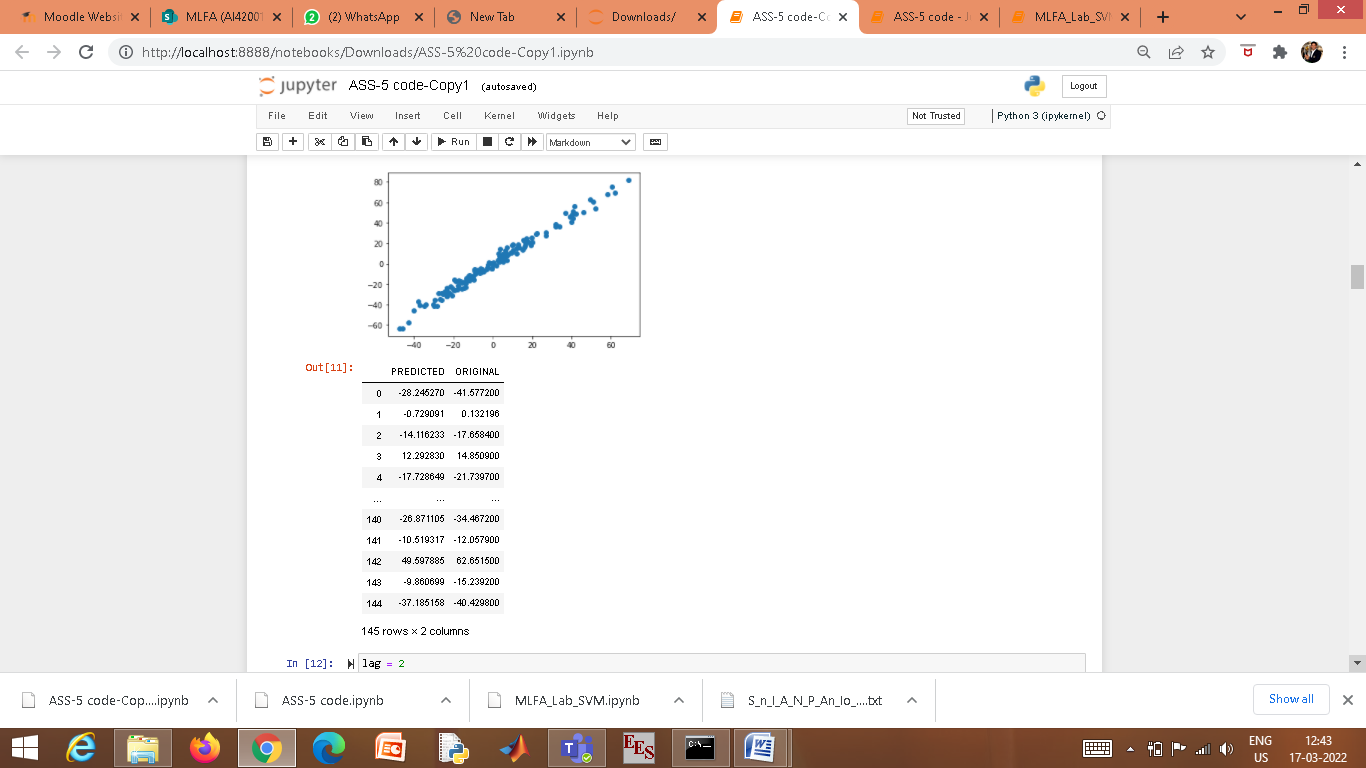
* Use ‘Xgboost regressor’ to predict ISMR for different lag values of the drivers and print the true and predicted values of the ISMR for different lags in a tabular form. Plot the true values vs predicted values of ISMR in scatter plots for different lags separately. Find out which lag is statistically showing best result and elaborate why.
* For lag=0



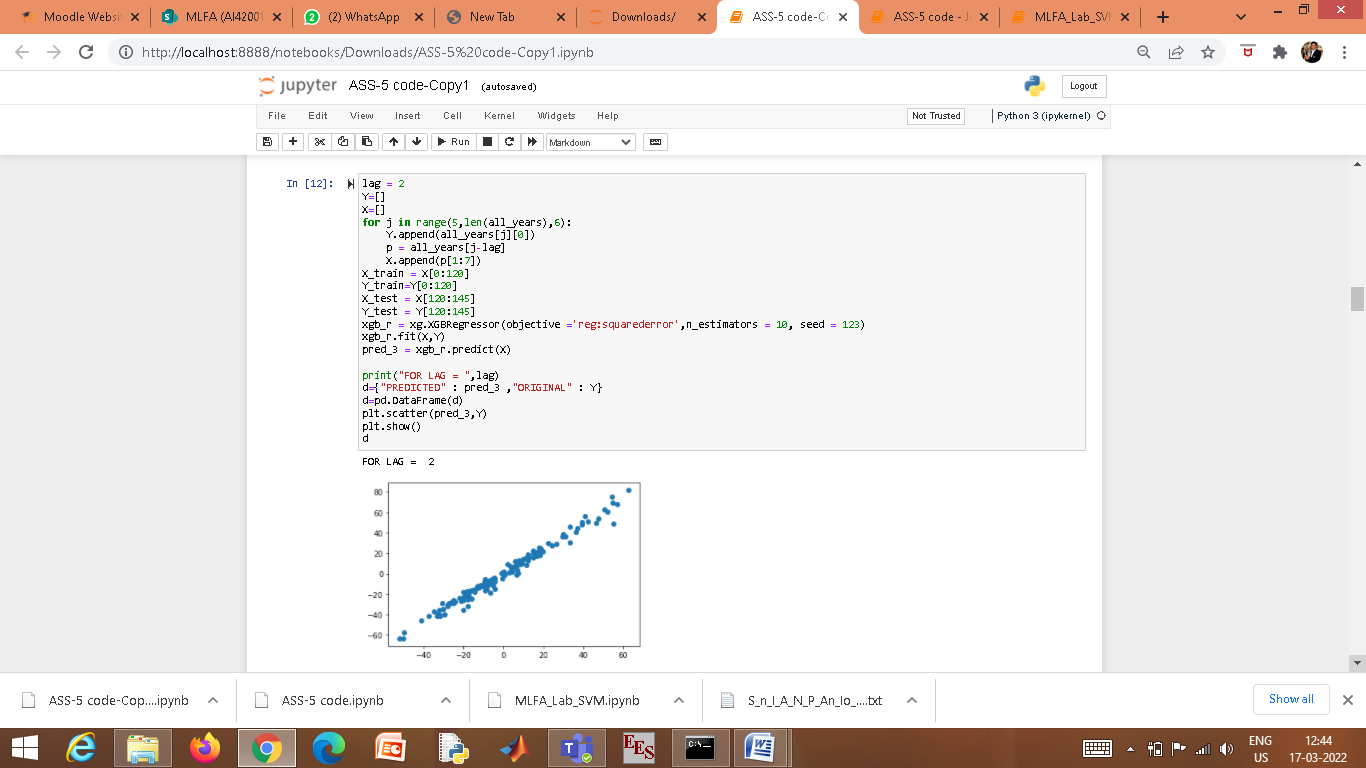


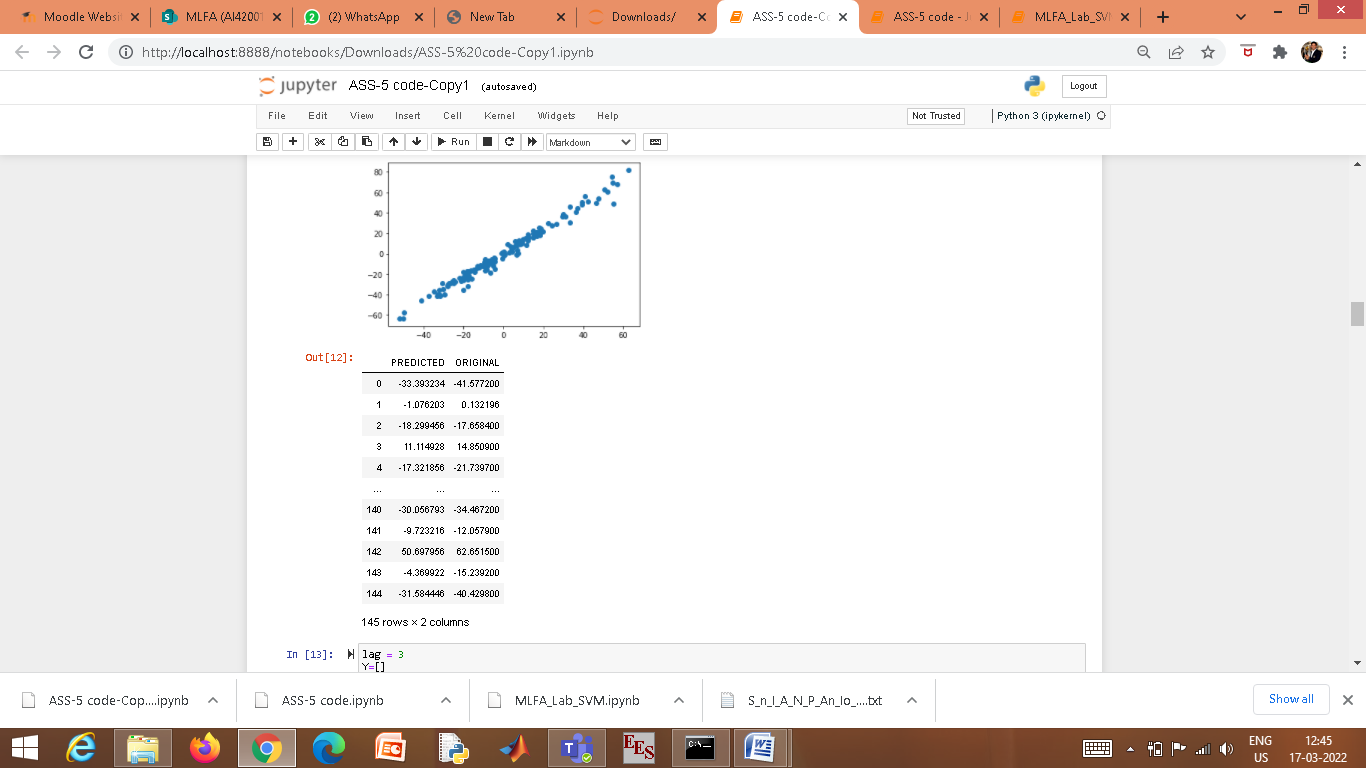
* For lag=1



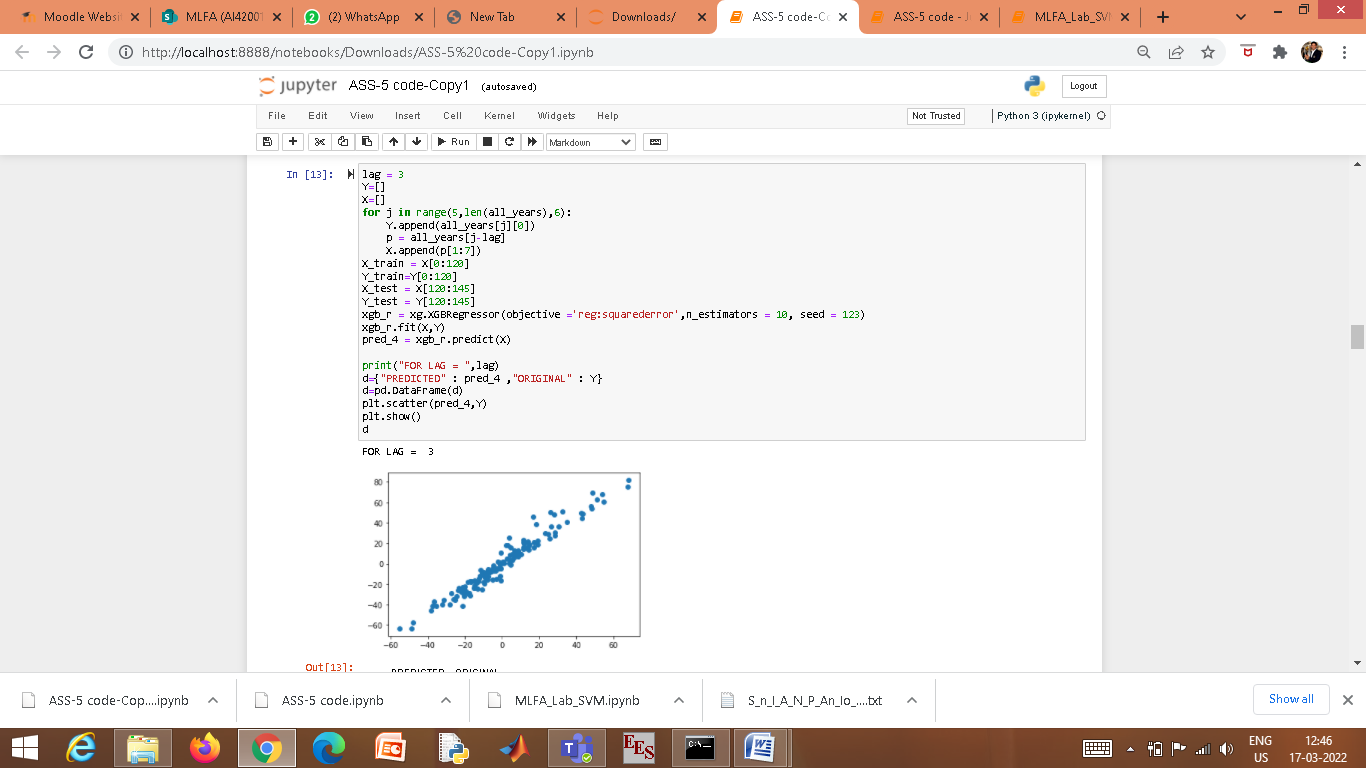


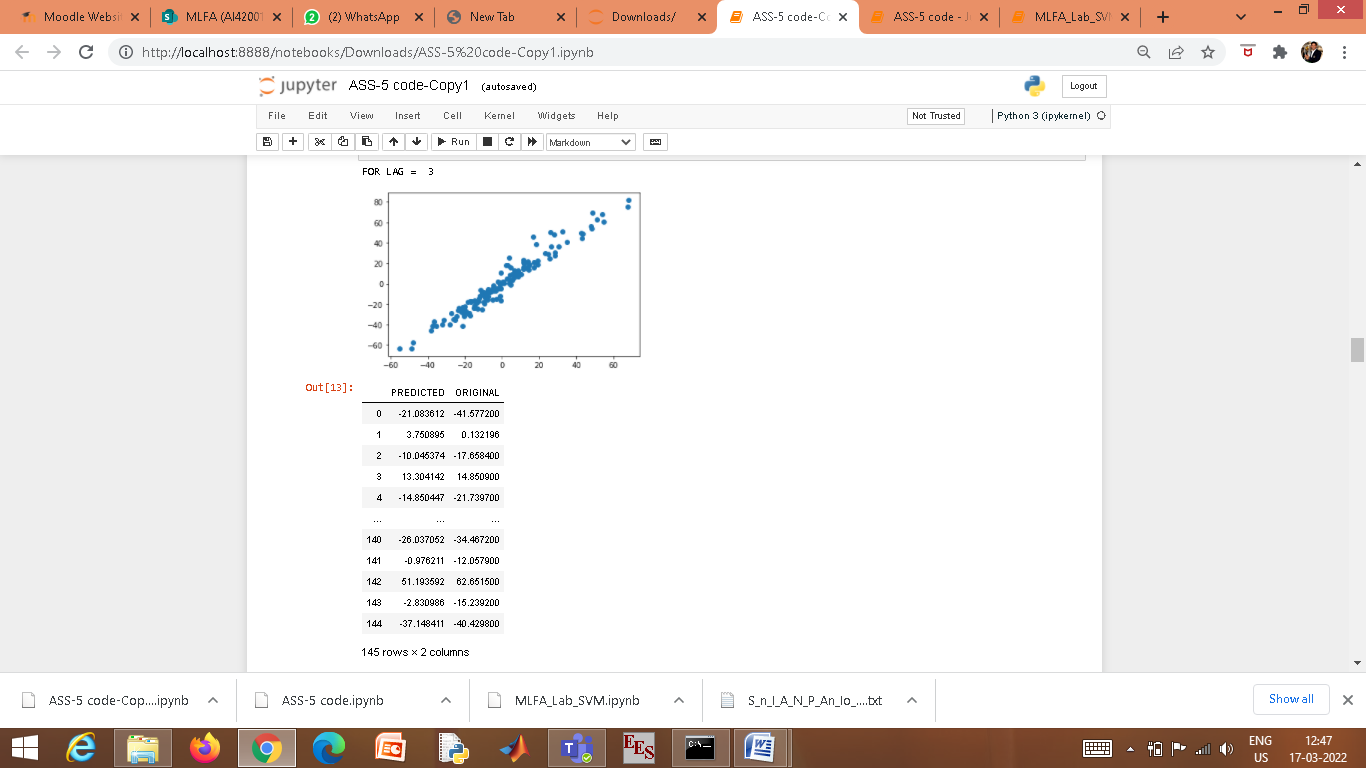
* For lag=2



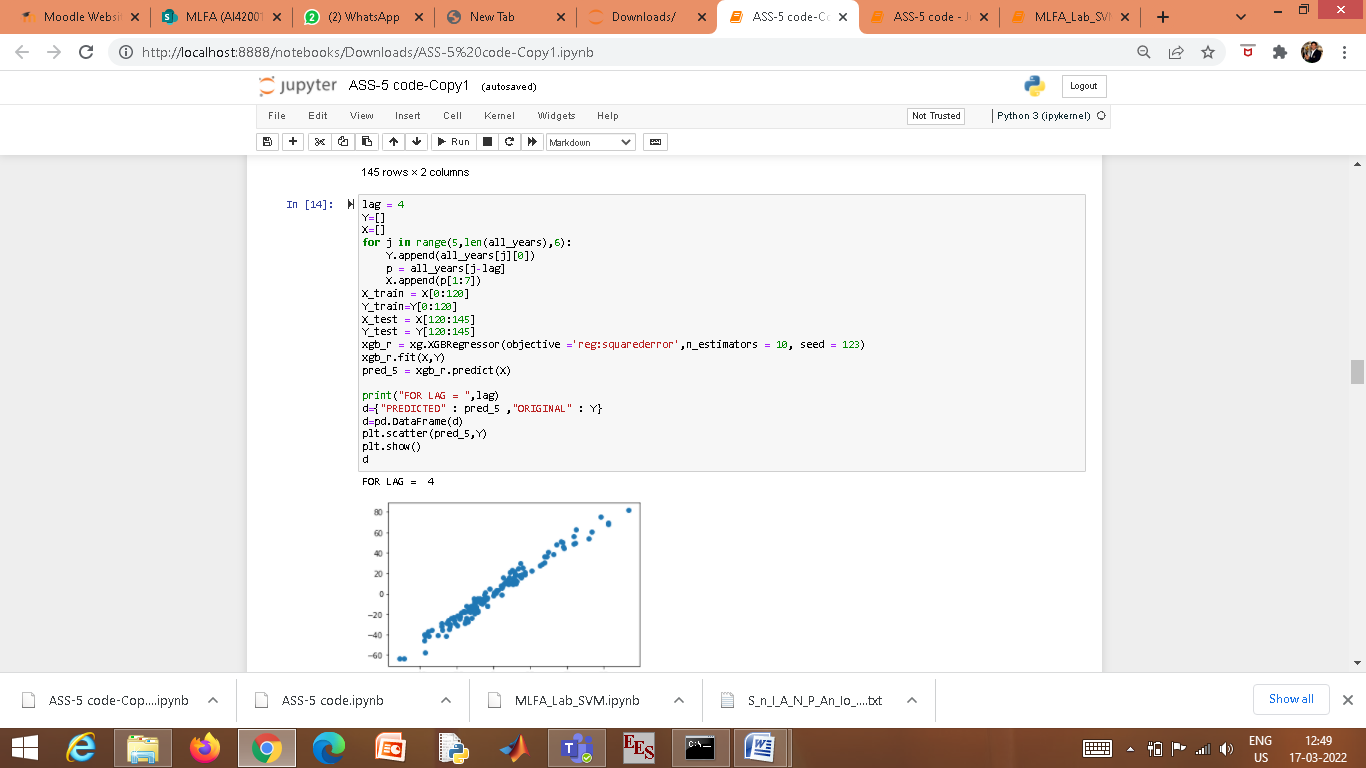


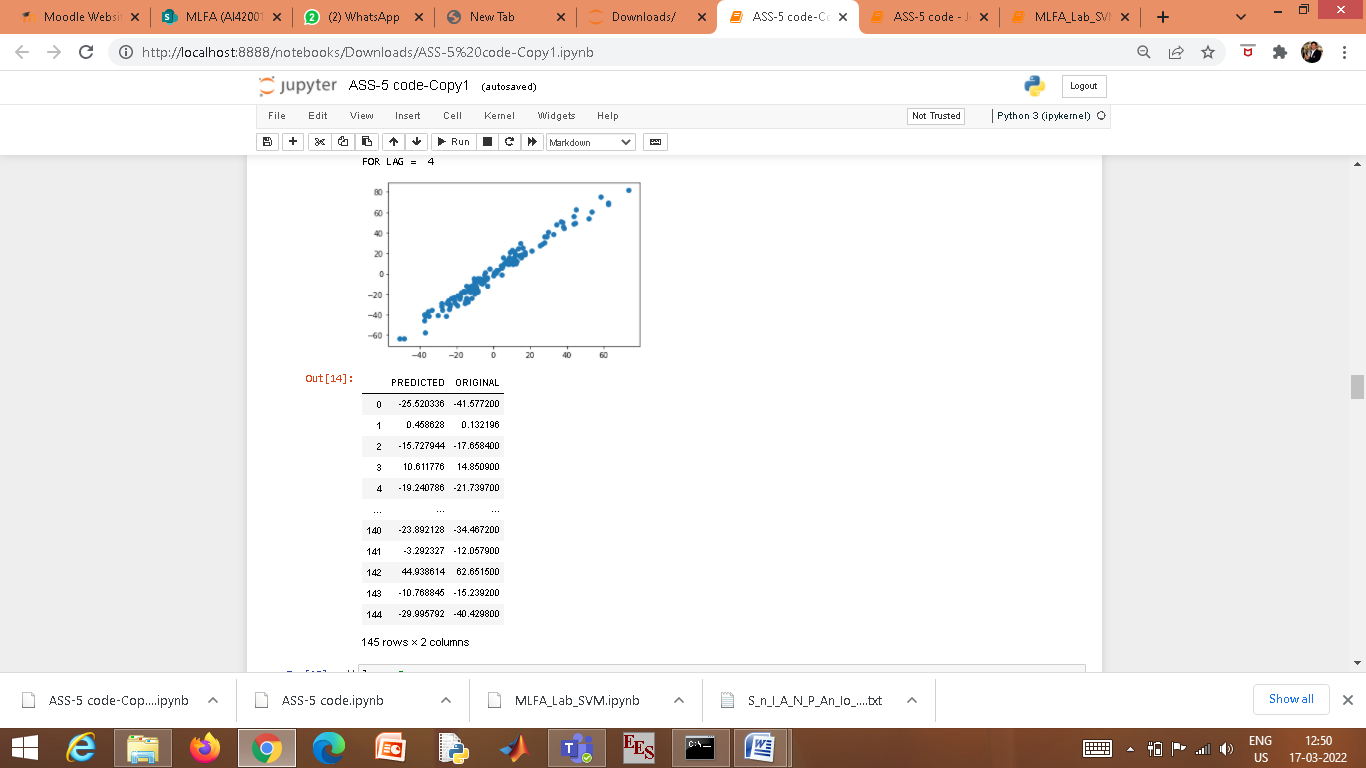
* For lag=3



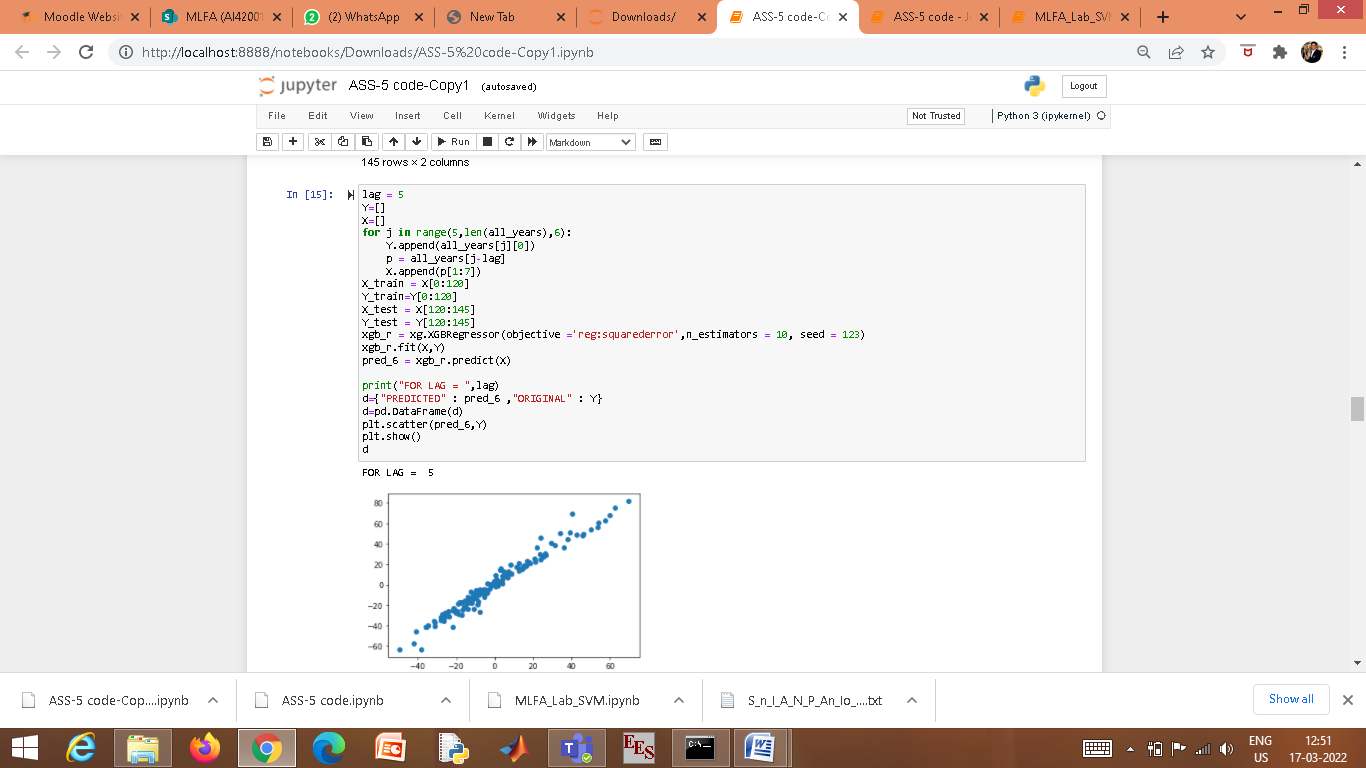


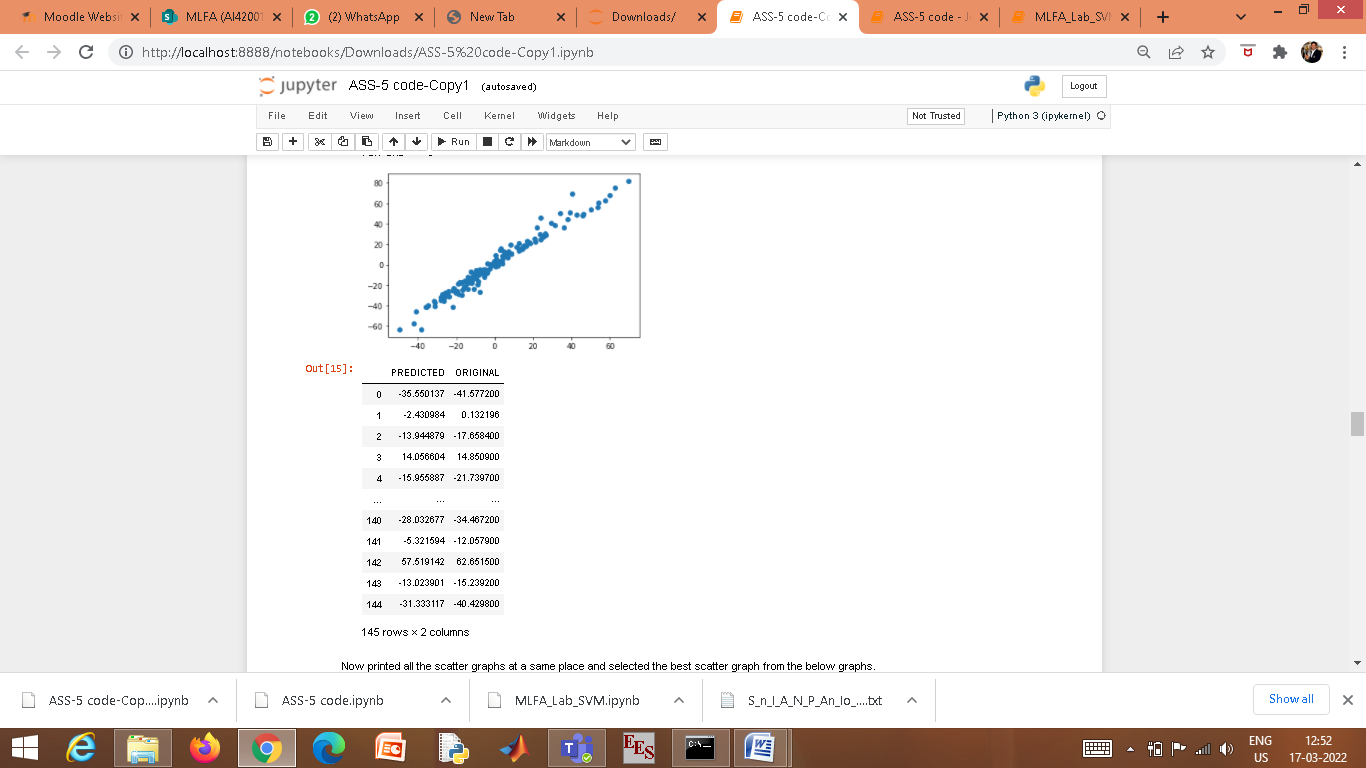
* For lag=4



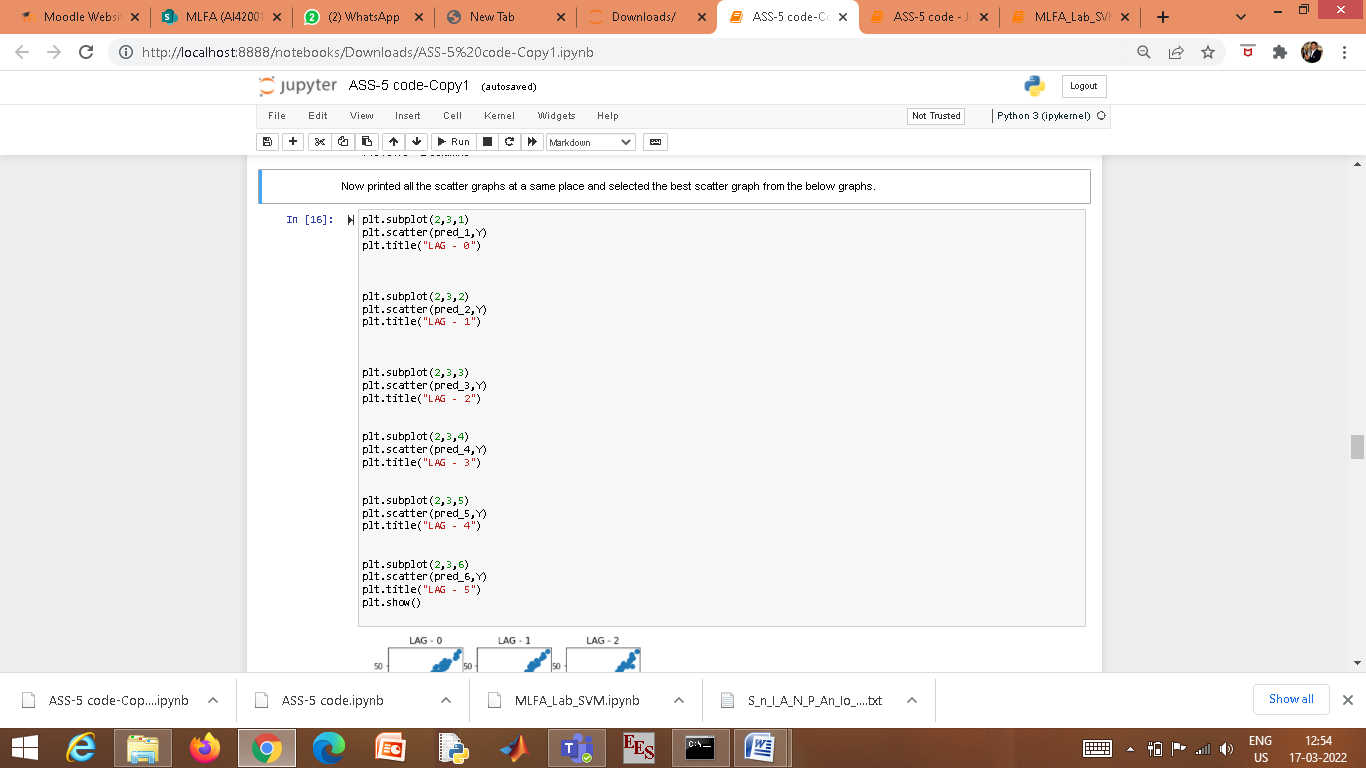


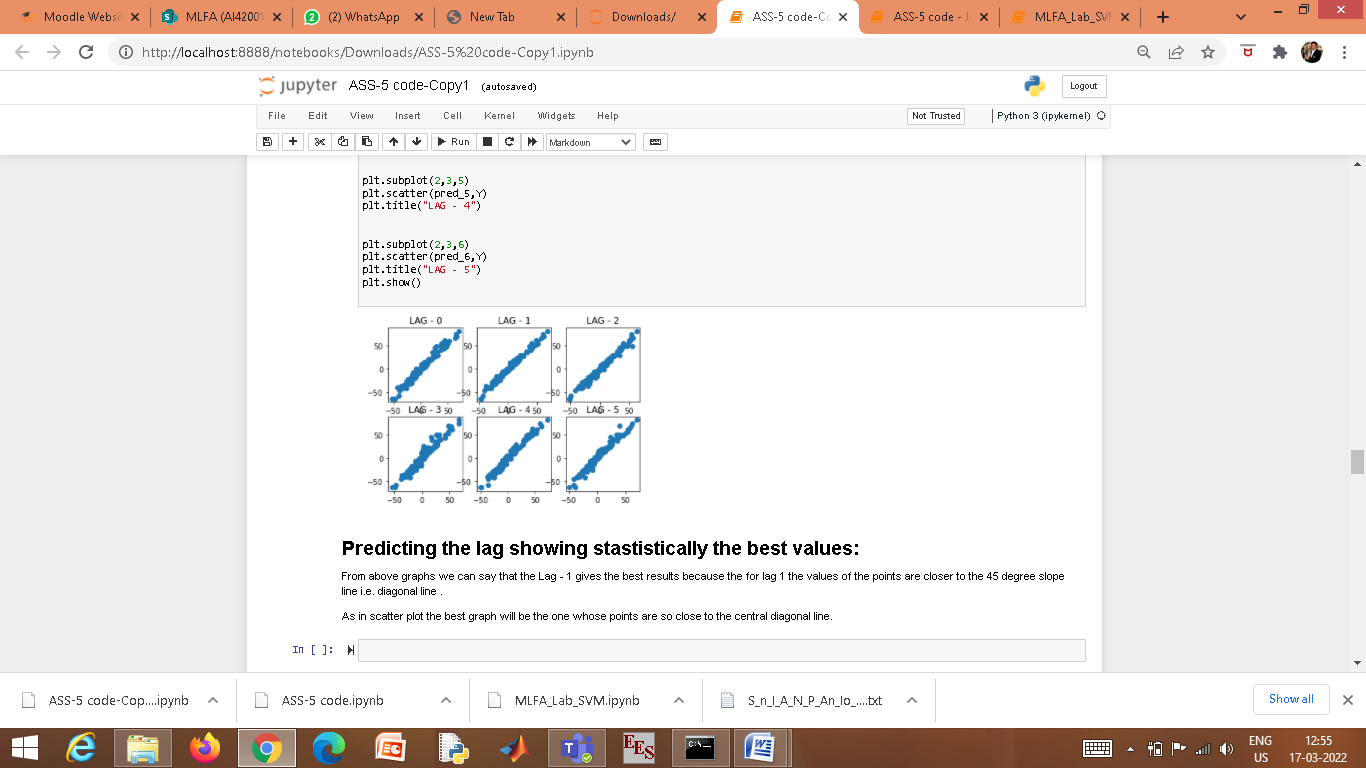
* For Lag=5





* Printing all the scatter graphs at the same place and selection of the best scatter graph from the below graphs.



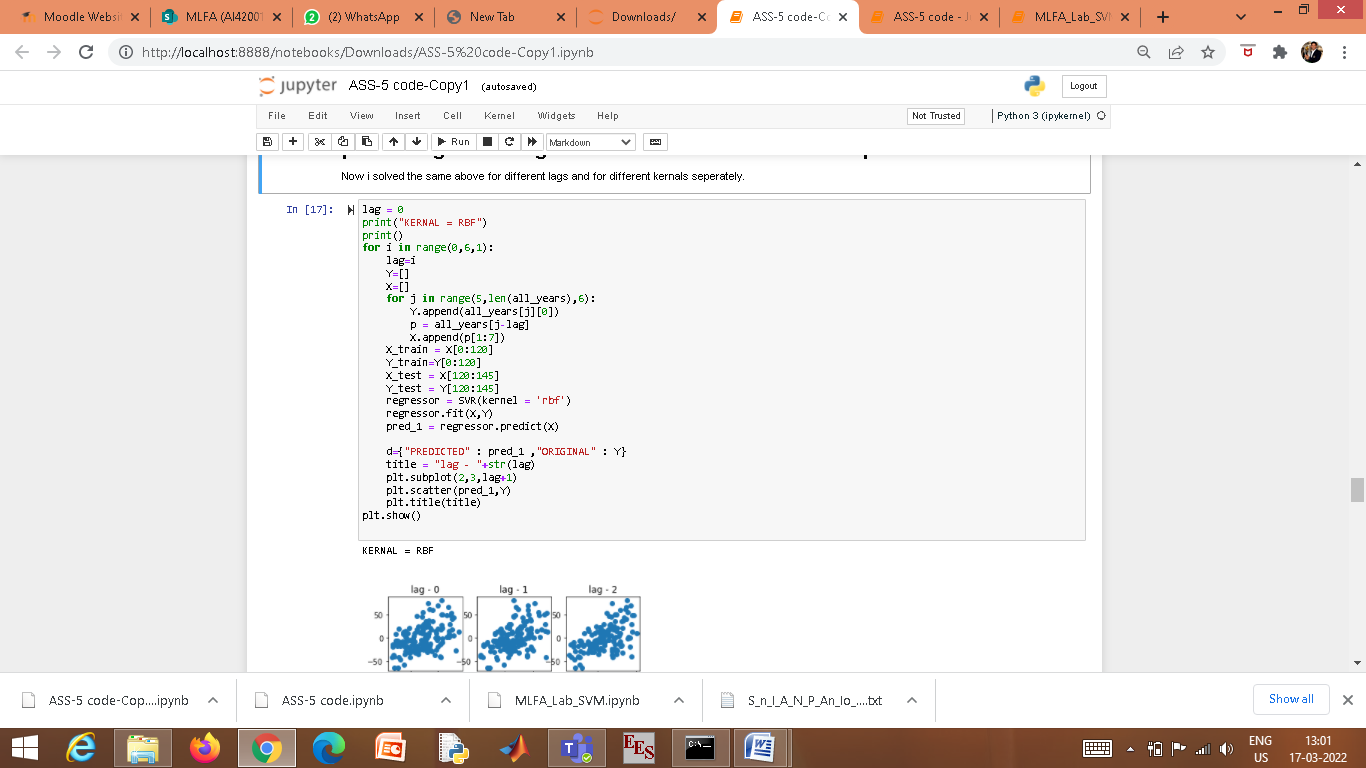


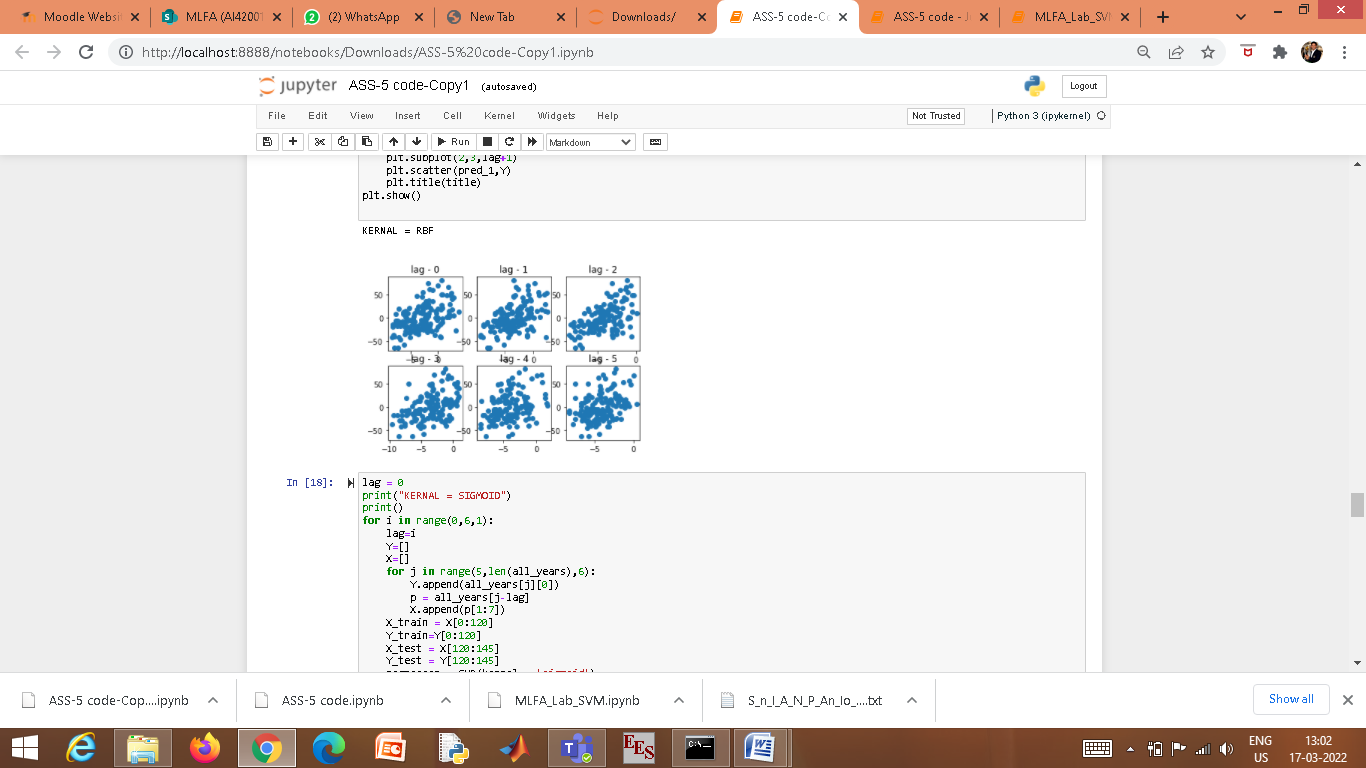
We can conclude that Lag=1 gives the best results as for lag=1 the values of the points are closer to the 45 degree slope line i.e. diagonal line.

# Repeat the experiment (2,3) for SVM with different kernel values to find the best performing one. Use grid search method to find the best performance.

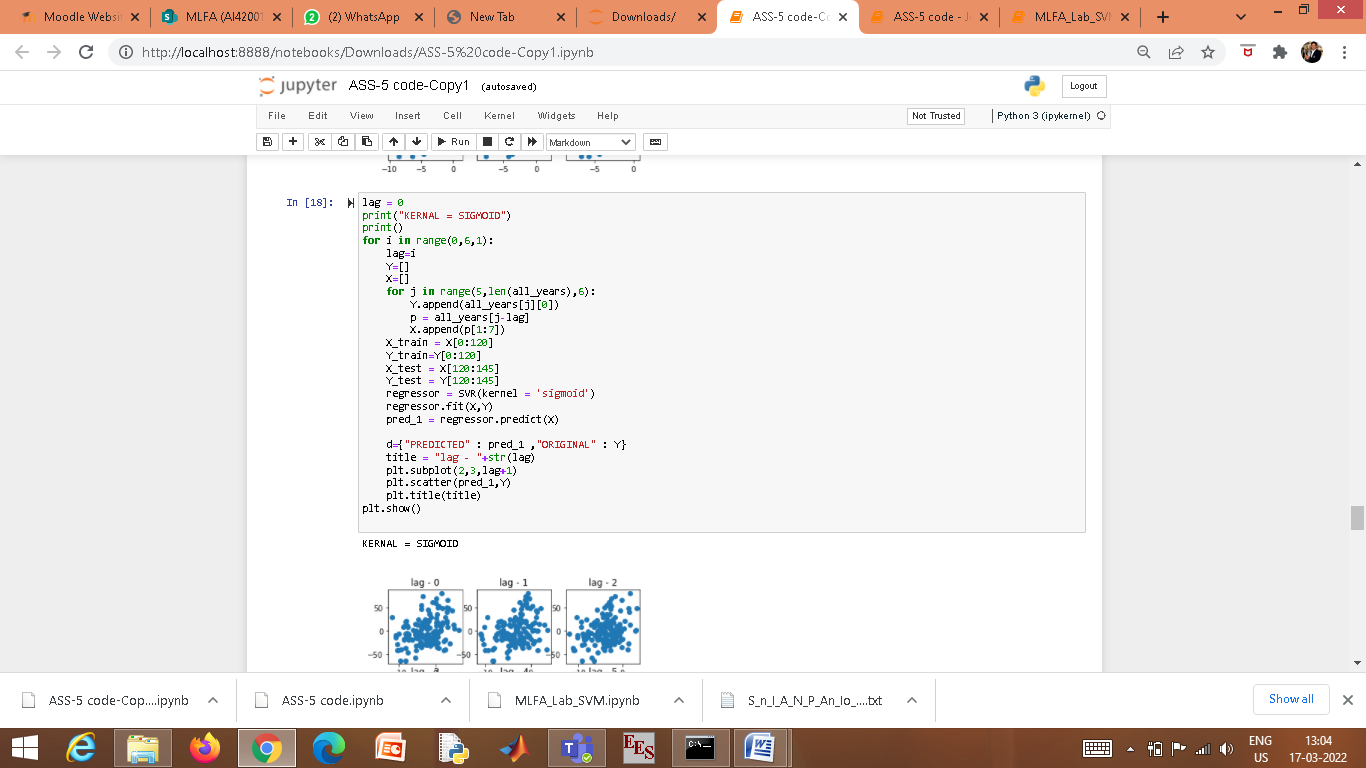
* Solving for different lags and different kernals separately.

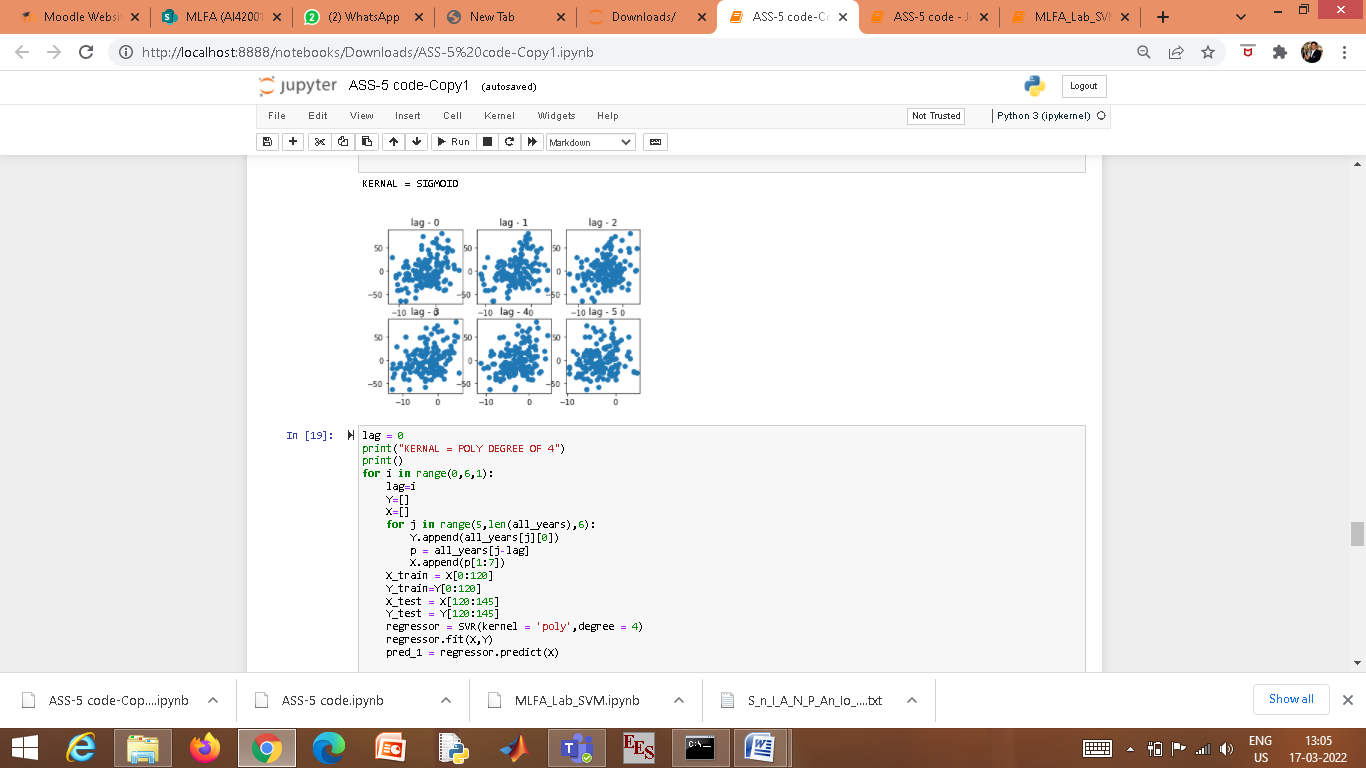
For Kernal=RBF



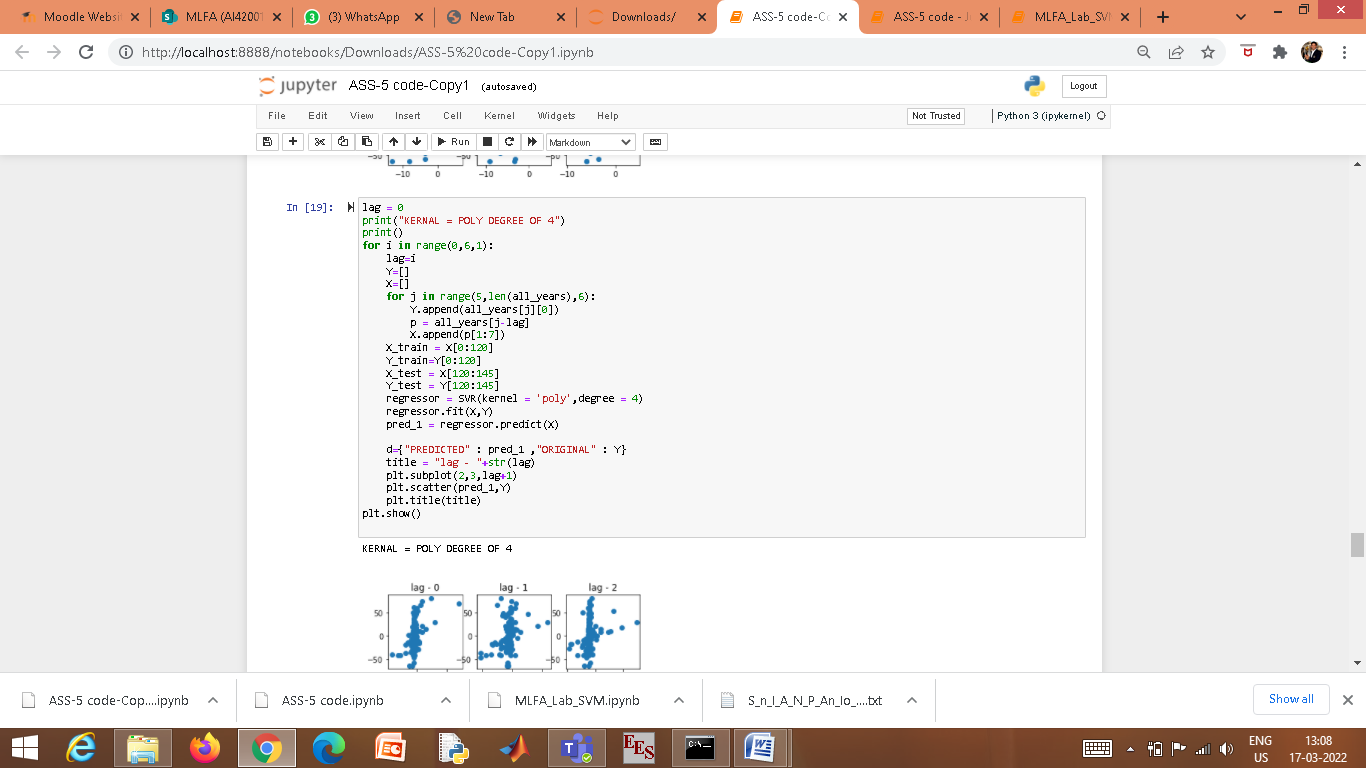


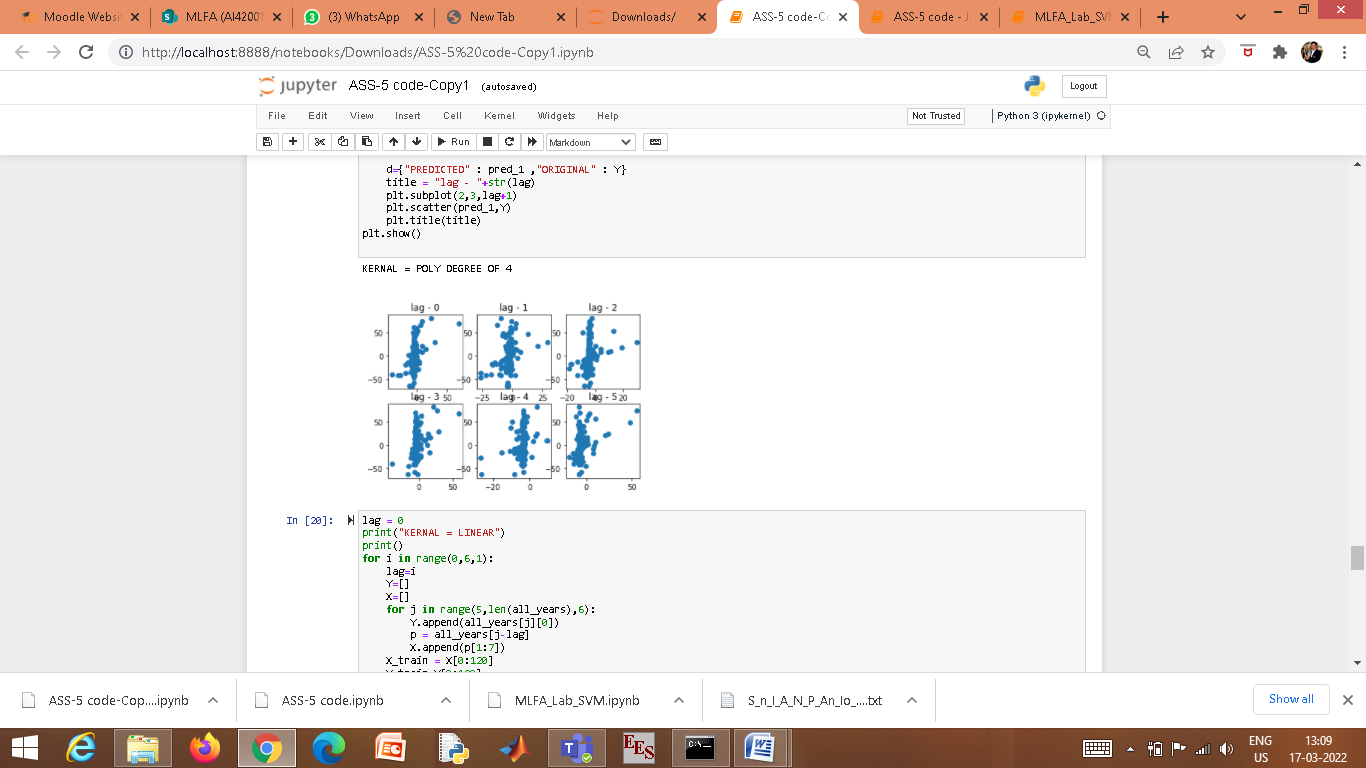
* For Kernal=SIGMOID



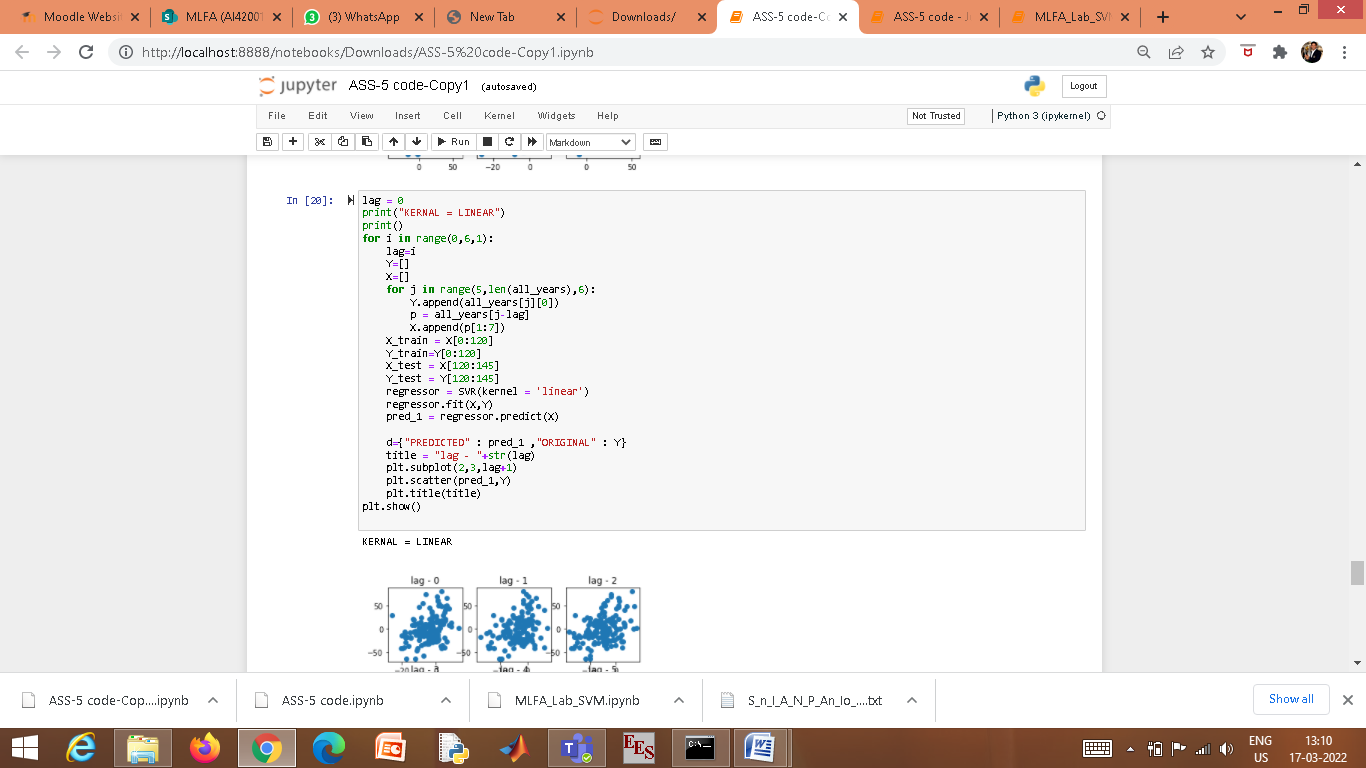


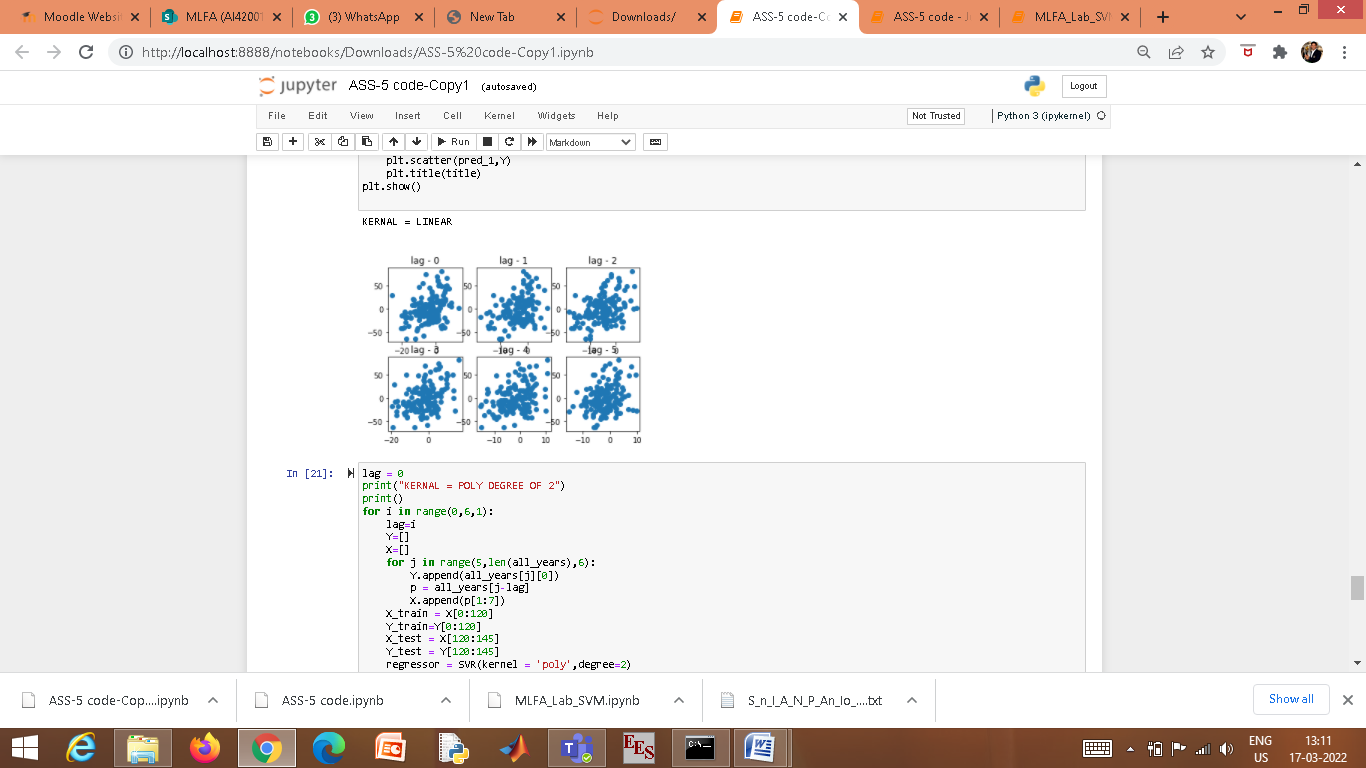
* For Kernal=Polynomial of degree 4



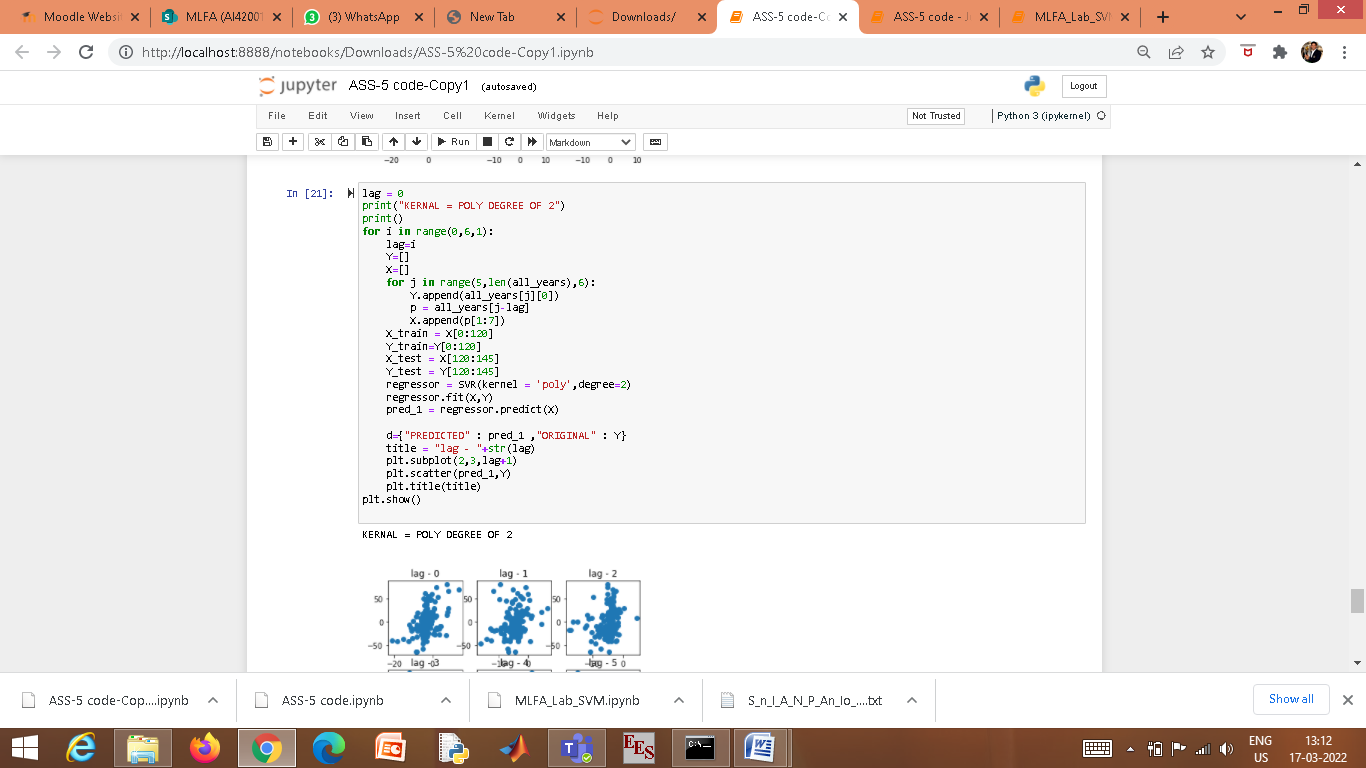


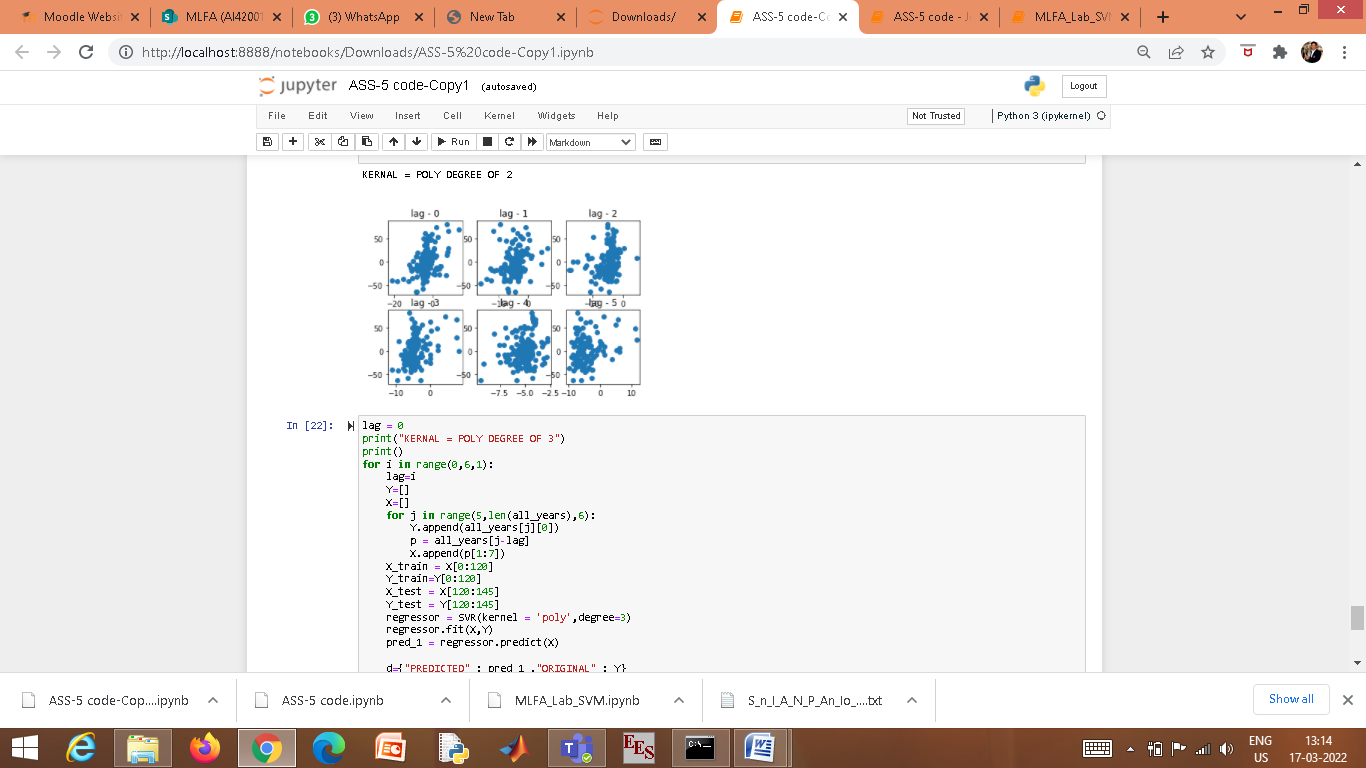
* For kernel=linear



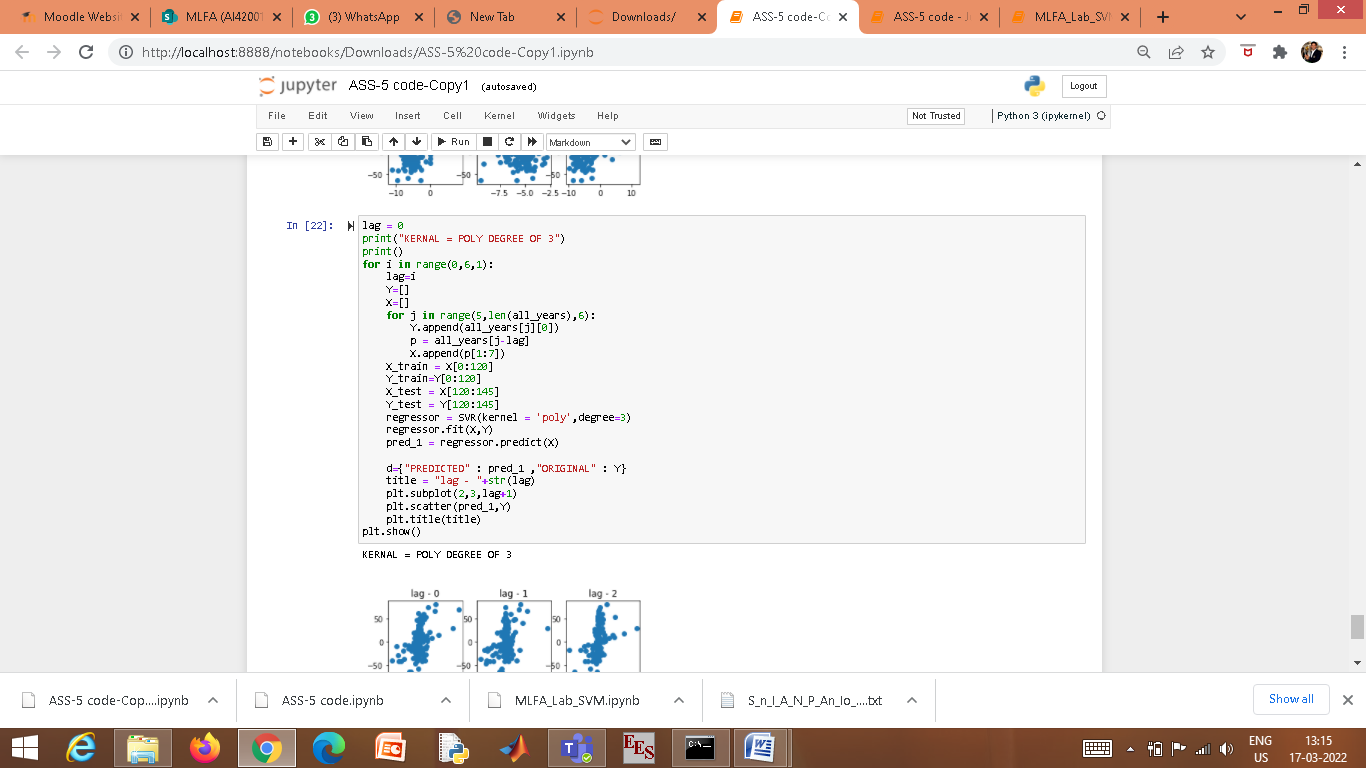


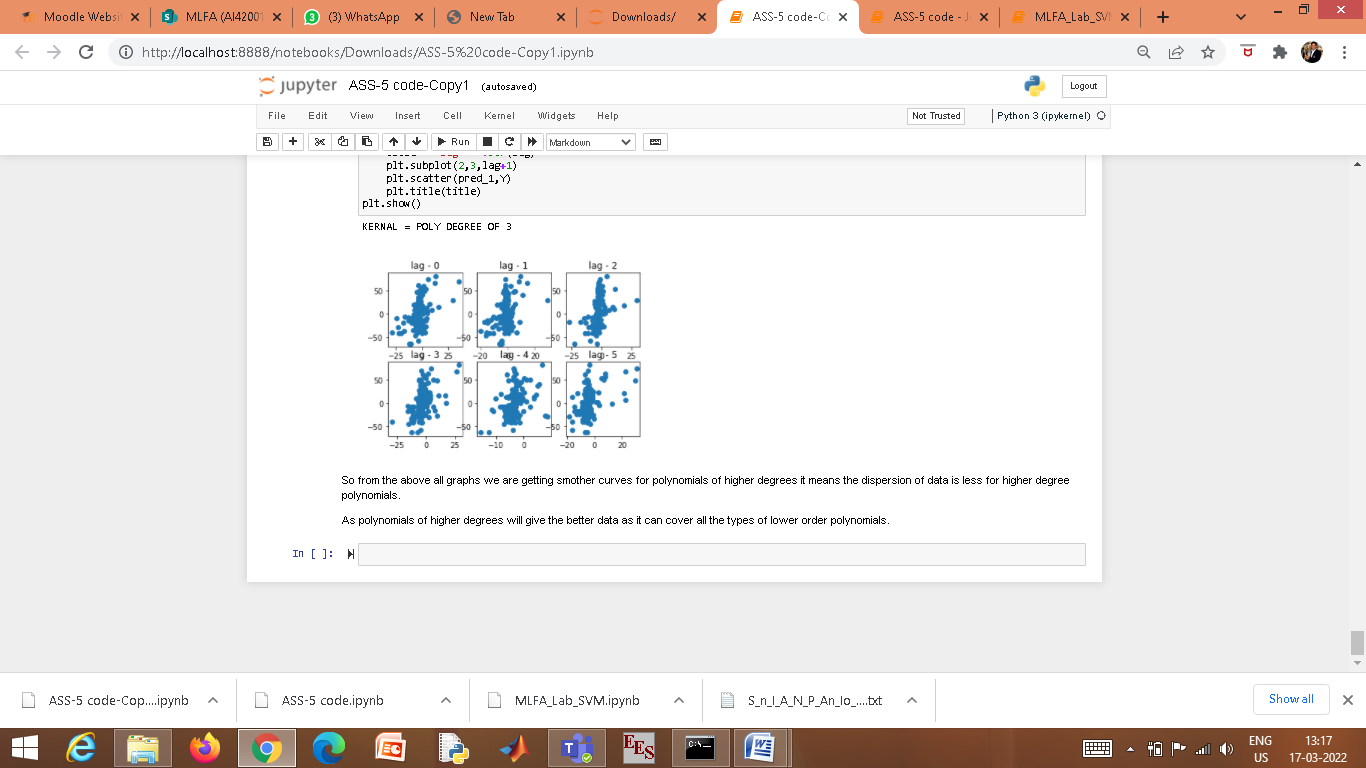
* For kernel=Polynomial of degree 2





* For kernel=Polynomial of degree 3





Therefore, from the above graphs we are getting relatively more smooth curves for polynomials of higher degrees which implies that the dispersion of data is less for higher degree polynomials.

Polynomials of higher degrees will give the better data as it can cover all the types of lower order terms.