

ADVANCED STATISTICS PROJECT

Task-01

- Collect the data of **Dependent variable** (assigned to your group) and **Independent Variables** (keeping in mind literature review).
- Collect the data of variables from some authentic Source like WDI, IFS, PBS, SBP etc.

Task-02

Discuss Summary Statistics (Mean, Median, Mode, Quartiles) of each variable included in the Model using R output of summary statistics.

Task-03

- Construct Box and Whisker Plots for all the variables in one diagram with different colors and identify outliers if any.

Task-04

- Construct Scatter Plot for all the variables on one grid. And interpret each part of the grid separately

Task-05

- There should be at least five independent variables (at least 35 observations) in the model.
- We shall consider only two countries for Inflation Forecasting i.e. Pakistan and USA

For **Pakistan** run the following models **ARIMA, Ridge, LASSO and Elastic Net**

For **USA** run the following models **ARIMA, SVM and Random Forest**

- In case you find that Regression output shows that more independent variables are insignificant then replace that insignificant variable with a new one which is significant.
- In conclusion section compare the predicted values of the models with original values using graph and MSE.
- Project Contents
 - Introduction (1-2) pages
 - Literature Review (at least 10 research papers)
 - Description of Variables and Data Description
 - Estimation of the Models (Model Introduction, Equation etc.)
 - Results and Conclusion
 - References

NOTE: Submit the following files on Submission Day

- Excel File of Data Collection
- R File of Programming and LaTeX File, **Use only R.**
- Pdf File of Final Project

SUBMISSION DATE: 4th May 2025