SHIVAJI COLLEGE, UNIVERSITY OF DELHI
DEPARTMENT OF ECONOMICS
INTERNAL ~~ASSIGNMENT~~/PROJECTS/~~PRESENTATIONS~~
(Academic Year 2023-24)

Name of the Course	: B.A. program	Semester	: V
Name of the Paper	: Data Analysis	Maximum Marks	: 10
Faculty Name	: Ajay	Last Date of Submission: 28th December	

Project Description:

In this project, you will conduct a regression analysis and interpret the results using a dataset provided to you. This project is designed to assess your understanding of regression analysis and your ability to apply it to real-world data.

Dataset: You are free to choose any dataset.

Project Tasks:

1. Data Preparation (15 points):

   a) Clean the dataset by handling missing values, outliers, and any data quality issues.

   b) Describe the dataset, including the number of observations, variables, and their types (categorical or numerical).

2. Exploratory Data Analysis (15 points):

   a) Perform a summary statistics analysis of all variables.

   b) Create appropriate visualizations (e.g., histograms, scatterplots, box plots) to explore the relationships between variables.

   c) Identify potential multicollinearity issues among the independent variables.

3. Regression Analysis (25 points):

   a) Select one of the independent variables as the dependent variable and the remaining independent variables as predictors.

   b) Choose an appropriate regression model (linear, multiple linear, or another regression type if justified).

c) Conduct the regression analysis using a statistical software of your choice (e.g., R or Excel).

d) Interpret the regression coefficients, including the intercept and slopes.

e) Assess the goodness-of-fit of the model using relevant statistics (e.g., R-squared, adjusted R-squared).

f) Check the assumptions of regression, such as linearity, homoscedasticity, and normality of residuals. Discuss any violations and their implications.

4. Interpretation and Conclusion (25 points):

a) Interpret the coefficients of the independent variables. What does each coefficient signify in the context of the problem?

b) Provide insights into the relationships between the independent and dependent variables.

c) Discuss the statistical significance of the coefficients and any potential practical significance.

d) Summarize the overall fit of the regression model and its predictive power.

e) Draw conclusions regarding the research questions or hypotheses, if applicable.

Project Report:

Prepare a well-organized project report that includes all the above tasks. Clearly present your findings, interpretations, and any visualizations. Include code and statistical output as appendices if necessary.