CASE STUDY :

INSURANCE CLAIM PREDICTION:

1. Read the Data Definition and Ensure to give proper Interpretation
2. Import the Following Libraries
3. Set Options for the following Output
4. Data Analysis and Preparations:
   - Understand the Dataset
   - Data Dimensions
   - Check for Data Types
   - Change the Incorrect Data Types if it's necessary
   - Sanity check for the Incorrect Data Types
   - Summary Statistics for (both Numerical and Categorical Columns)
   - Missing Value Treatment (showcase in Coding and Visualization – Choose Appropriate plotting)
   - Sanity Check for Missing Values (Showcase in Coding and Visualization – Choose Appropriate Plotting)
   - IF missing Values are exists, Deal with Appropriate Treatment.
   - Calculate Pearson's Correlation for Both Categorical and Continuous Data.
   - Set of Analysing Categorical Columns with Appropriate Plotting, with inferences
   - Analyse the Relationship B/w Target and Categorical Variables.
5. Feature with loaded Information about the Dataset, The insights should be meaningful.
6. Discover Outliers, and add up necessary extraction insights with proper treatment.
7. Check for Multicollinearity with Adaptive Method as user knows.
8. Prepare the Data:
   - Check the Normality
   - Plot a Histogram and also perform the Jarque-Bera Test
     (It is a test to check the Normality of the Target variable, consider with p-value)
   - If the data is not normally distributed, use appropriate way to get near normally distributed data.
   - Recheck the normality by plotting Histogram, and Perform Jarque-Bera Test
   - Perform One way Anova
9. Encoding for Categorical Data
10. Scaling for Ensured Method
11. Split the Dataset for further analysis
12. Impute the Linear Regression Analysis (OLS Summary)

----------xxxxxxxxxxxxxxxxxxxxxxxxxxx----------

Prepared By,

Bhuvanesh Chellapandian

IPCS GLOBAL SOLUTIONS PVT LTD