

The dataset belongs to a leading life insurance company. The company wants to predict the bonus for its agents so that it may design appropriate engagement activity for their high performing agents and upskill programs for low performing agents.

The business report submitted for this assignment must describe the data and be aligned to the analytical business problem above. It should also include the rationale of the project (why is the problem important and what value will it add to the company). The submission should be of 7 to 10 pages (but not limited to) containing a simple and straightforward depiction of the objective of the project, description of data, Exploratory Data Analysis (EDA) report and the insights from EDA.

Below is the rubric for this project:

| Review Parameters | Points |
|--|---------------|
| 1) Introduction of the business problem | 20 |
| a) Defining problem statement | |
| b) Need of the study/project | |
| c) Understanding business/social opportunity | |
| 2) Data Report | 10 |
| a) Understanding how data was collected in terms of time, frequency and methodology | |
| b) Visual inspection of data (rows, columns, descriptive details) | |
| c) Understanding of attributes (variable info, renaming if required) | |
| 3) Exploratory data analysis | 50 |
| a) Univariate analysis (distribution and spread for every continuous attribute, distribution of data in categories for categorical ones) | |
| b) Bivariate analysis (relationship between different variables , correlations) | |
| a) Removal of unwanted variables (if applicable) | |
| b) Missing Value treatment (if applicable) | |
| d) Outlier treatment (if required) | |
| e) Variable transformation (if applicable) | |
| f) Addition of new variables (if required) | |
| 4) Business insights from EDA | 20 |
| a) Is the data unbalanced? If so, what can be done? Please explain in the context of the business | |
| b) Any business insights using clustering (if applicable) | |
| c) Any other business insights | |
| Total | 100 |