You need to submit one Ipython notebook which clearly explains the thought process behind your analysis (either in comments or markdown text), code and relevant plots.

Important Note: Please make sure to rename your Python notebook "Group_Facilitator_Name.ipynb".

You need to submit a GitHub repository link that contains the files above . You need to write a README.md for the GitHub repository (Official README.md template is given). You can follow the video below to understand how to create the GitHub repository. Make sure your GitHub link is public.

Evaluation Rubric

Criteria	Meets Expectations	Does Not Meet Expectations
Data understanding (10%)	All data quality issues are correctly identified and reported. Wherever required, the meanings of the variables are correctly interpreted and written either in the comments or text.	Data quality issues are overlooked or are not identified correctly such as outliers, missing values and other data quality issues. The variables are interpreted incorrectly or the meaning of variables is not mentioned.
Data Cleaning and Manipulation (20%)	Data quality issues are addressed in the right way (missing value imputation, outlier treatment and other kinds of data redundancies, etc.). If applicable, data is converted to a suitable and convenient format to work with using the right methods. Manipulation of strings and dates is done correctly wherever required.	Data quality issues are not addressed correctly. The variables are not converted to an appropriate format for analysis. String and date manipulation is not done correctly or is done using complex methods.
Data analysis (40%)	The right problem is solved which is coherent with the needs of the business. The analysis has a clear structure and the flow is easy to understand. Univariate and segmented univariate analysis is done correctly and appropriate realistic assumptions are made wherever required. The analyses successfully identify at least the 5 important driver variables (i.e. variables which are strong indicators of default).	The analyses do not address the right problem or deviate from the business objectives. The analysis lacks a clear structure and is not easy to follow. The univariate and bivariate analysis is not performed in sufficient detail and thus some crucial insights are missed out. The analyses are not able to identify enough important driver variables. New metrics are not derived wherever appropriate. The explanation for creating the

	Business-driven, type-driven and data-driven metrics are created for the important variables and utilised for analysis. The explanation for creating the derived metrics is mentioned and is reasonable. Bivariate analysis is performed correctly and is able to identify the important combinations of driver variables. The combinations of variables are chosen such that they make business or analytical sense. The most useful insights are explained correctly in the comments. Appropriate plots are created to present the results of the analysis. The choice of plots for respective cases is correct. The plots should clearly present the relevant insights and should be easy to read. The axes and important data points are labelled correctly.	derived metrics is either not mentioned or the metrics are not reasonable. Derived metrics are not analysed correctly/are insufficiently utilised. Important insights are not mentioned in the report or the Python file. Relevant plots are not created. The choice of plots is not ideal and the plots are either difficult to interpret or lack clarity or neatness. Relevant insights are not clearly presented by the plots. The axes and important data points are not labelled correctly/neatly.
Presentation and Recommendations (20%)	The presentation has a clear structure, is not too long, and explains the most important results concisely in simple language. The recommendations to solve the problems are realistic, actionable and coherent with the analysis. The GitHub repository link contains a Python file, presentation file and a README.md file. README.md file should describe the project briefly.	The presentation lacks structure, is too long or does not put emphasis on the important observations. The language used is complicated for business people to understand. The recommendations to solve the problems are either unrealistic, non-actionable or incoherent with the analysis. The GitHub repository should not be private. The GitHub repository link which does not contain a