Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
RAHUL SINGH WALDIA rahulwaldia2@gmail.ciom 1. CHECK THE DATA 2. DATA EXPLORATION 3. PREDICTIVE MODEL 4. CONCLUSION
Please paste the GitHub Repo link.
Github Link:- https://github.com/RahulSinghWaldia/CREDIT-CARD-DEFAULT-PREDICTION.git
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

In case you're unclear on what defaulting on a credit card means, here's the gist: After you've failed to make a payment on your credit card for 180 days (or as decided by your credit card company), your issuer assumes you're probably never going to. At this point, the issuer can (and usually does) close your card, write off what you owe as bad debt and sell your account to a collections agency. Now your credit card issuer is out of the picture. Your debt belongs to a collector, and you'll start getting flooded with phone calls asking you to pay the bill. Although you can send written notice asking them to stop calling you and there are laws limiting what debt collectors can say and do when they contact you, you'll still owe the debt. And if you don't deal with it one way or another, you could get sued. Therefore, this project aims to bridge this gap of uncertainty by utilizing a data-driven approach by using past data of credit card customers in ***conjunction with machine learning*** to predict whether or not a consumer will default on their credit cards. In this case, categorical columns like `PAY_0` , …, `PAY_6`,`MARRIAGE`,`EDUCATION` may not have been represented the dataset in the best way. A better way could have been to do one-hot encoding (creating dummy variables) instead. Classifiers like RandomForests are great at segregating columns like these and could have resulted in a model that leads to better prediction of credit card defaulters.