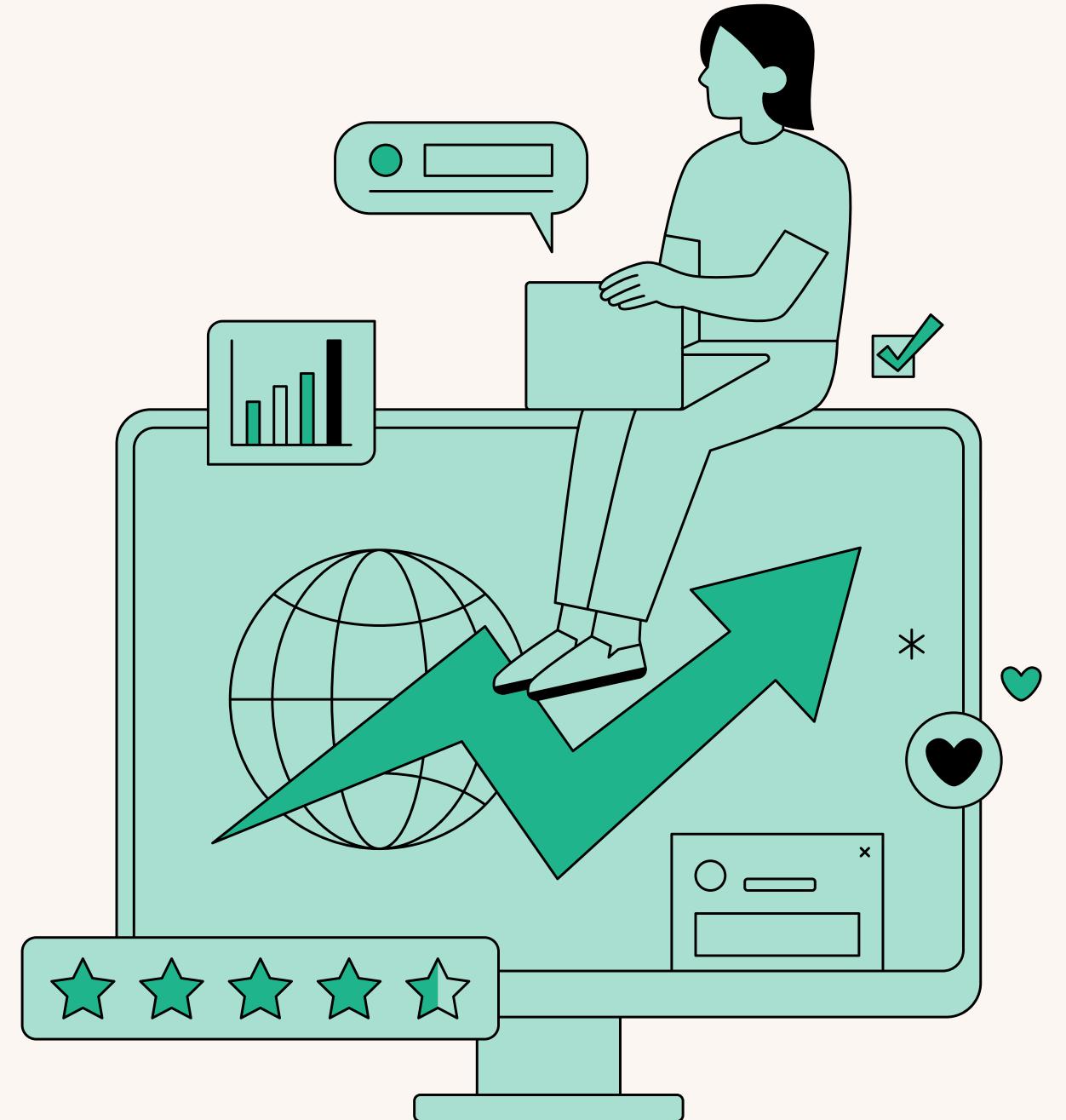


Bank Credit Score Classification

Gustana Satiawan



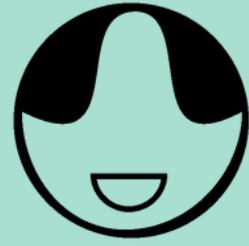
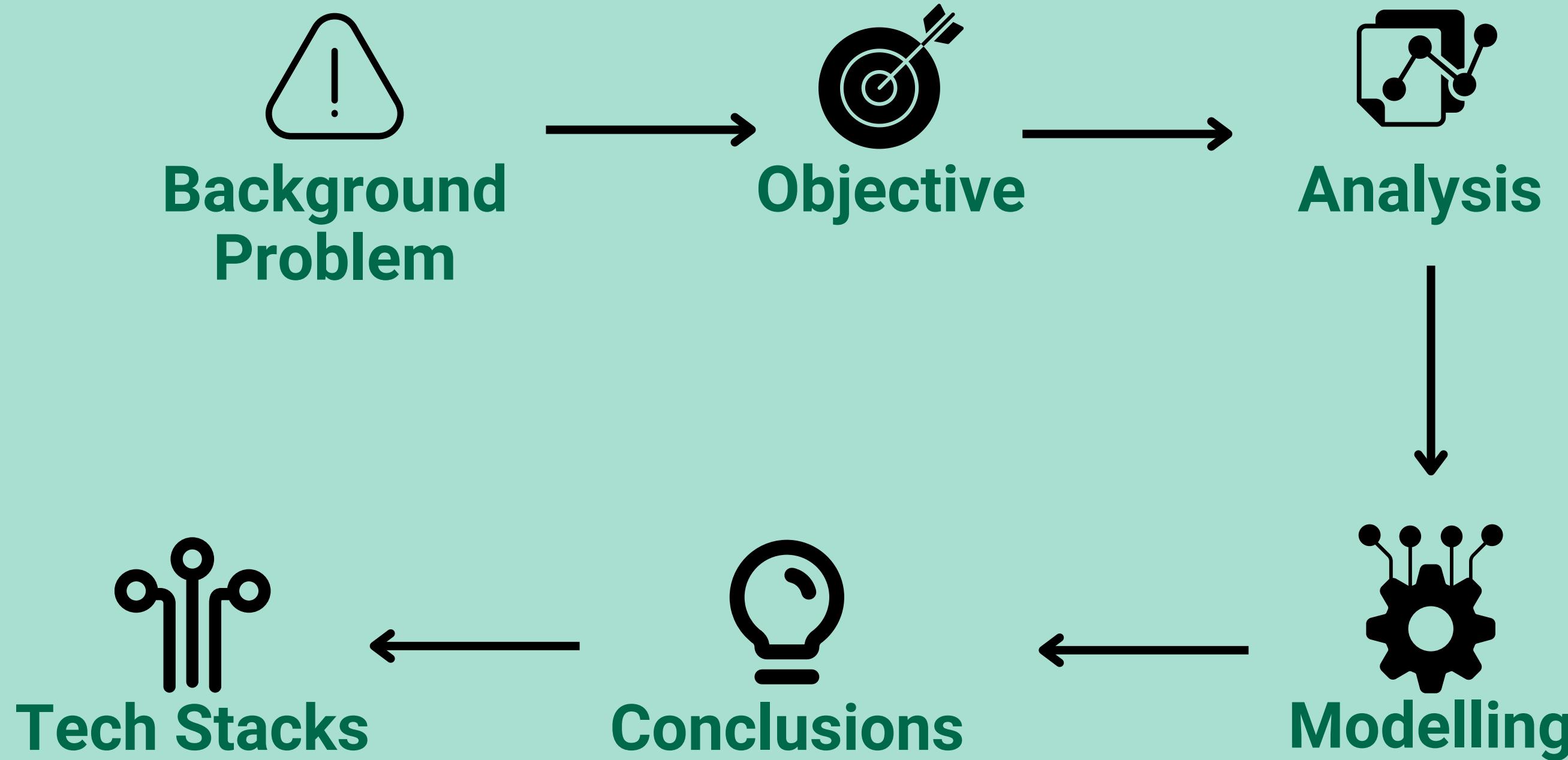
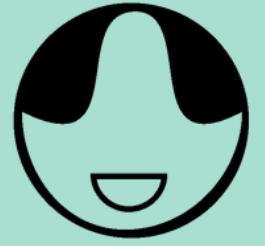


Table of Contents

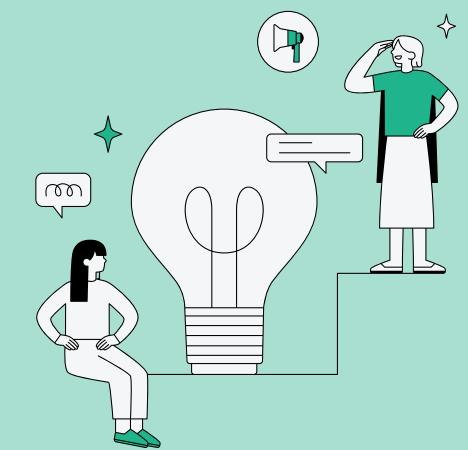


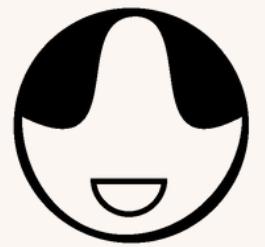


Background Problem

Credit score is a metric used to determine credit worthiness of a customer. In other words, this score reveals how reliable a customer is in their ability to pay off their debt.

For a banking business, one of the key benefits of assessing its customers' credit scores is reducing the risk of non-performing loans (NPLs), thereby lowering overall financial risk.

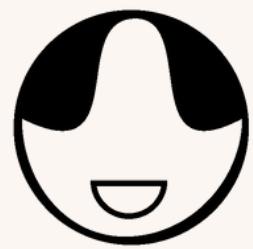




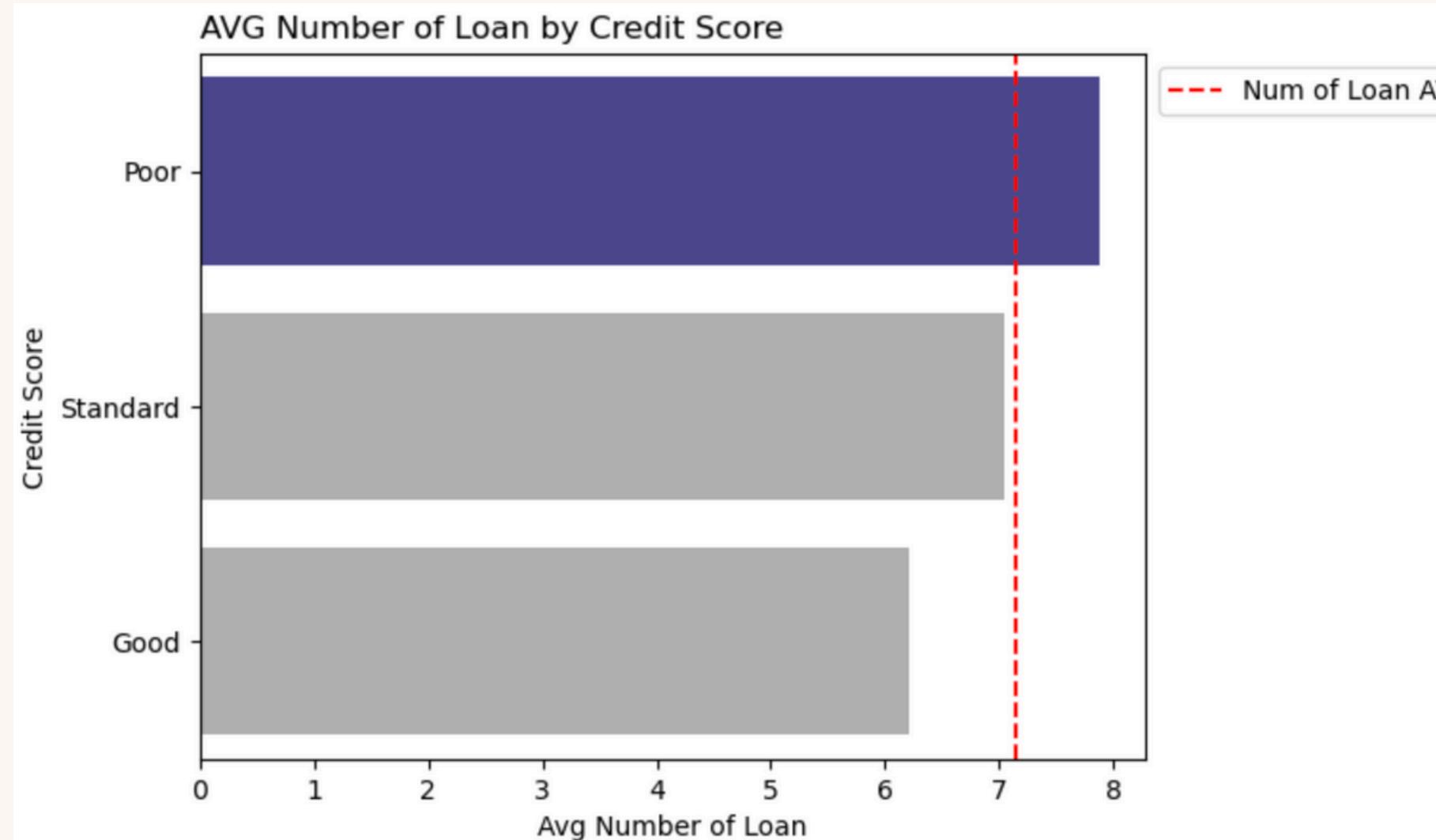
Objective

- Creating a classification model to predict customer's credit score based on their characteristics.
- The model should be able to achieve minimum precision of 90% using the following criterias:
 - Positive class: good credit score
 - Negative class: standard or bad credit score
 - FN: customers with good credit scores classified as bad or standard credit scores
 - FP: customers with bad or standard credit scores classified as good credit scores

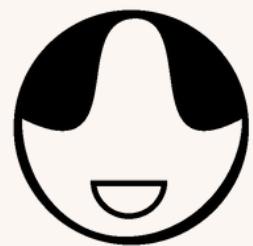




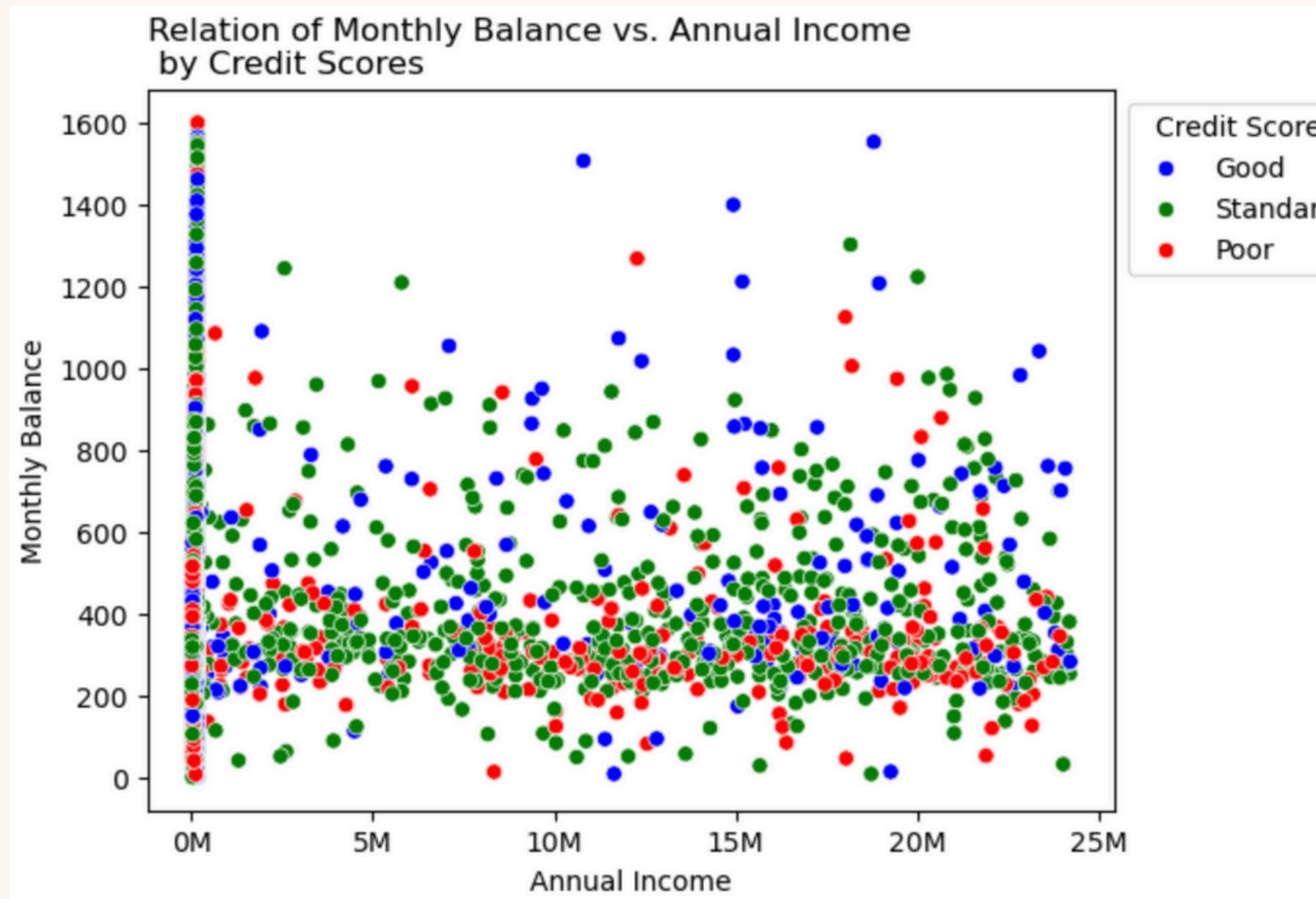
Analysis



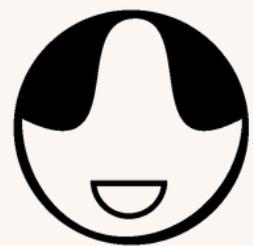
Customers with poor credit scores have a slightly higher number of credit cards on average, but the difference compared to those with standard and good credit scores is not significant



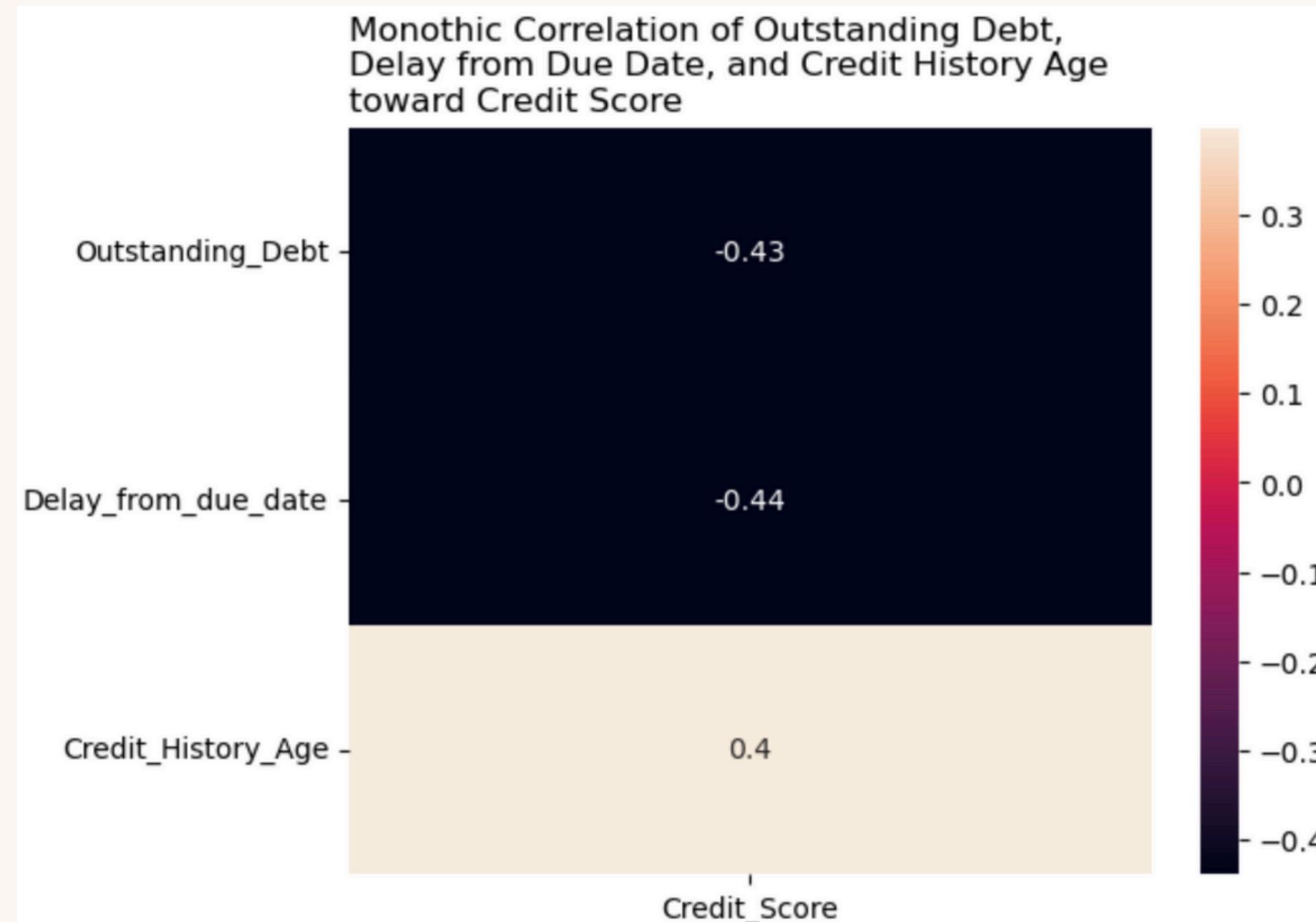
Analysis



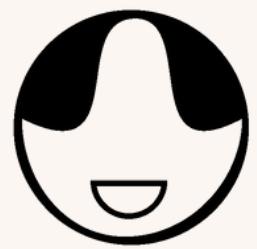
- Data points that grouped by credit scores are not forming any trend.
- This concludes that **there's no relation among monthly balance, annual income, and credit score**



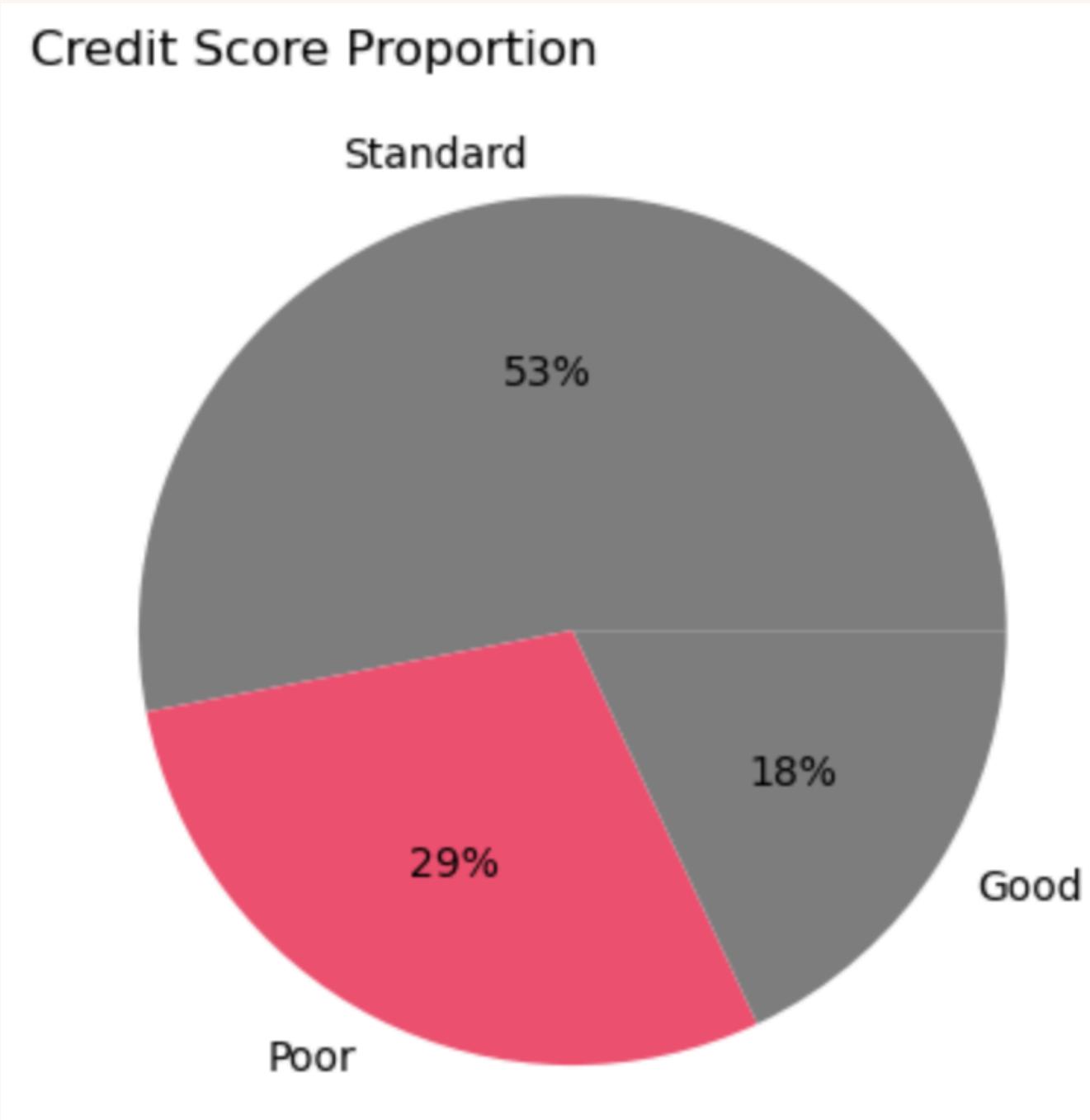
Analysis



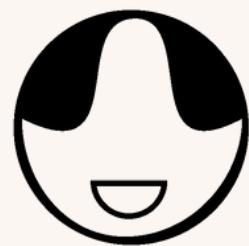
- **As the outstanding debt and delay from due date increases, the credit scores decreases.** Implying customer with high outstanding debt and/or delay from due date will have low credit score
- **As the credit card go older, the credit score increases.** This might happen because they become a loyal customer and able to manage their credit score. Consequently, they can utilize the credit card for their own benefit.



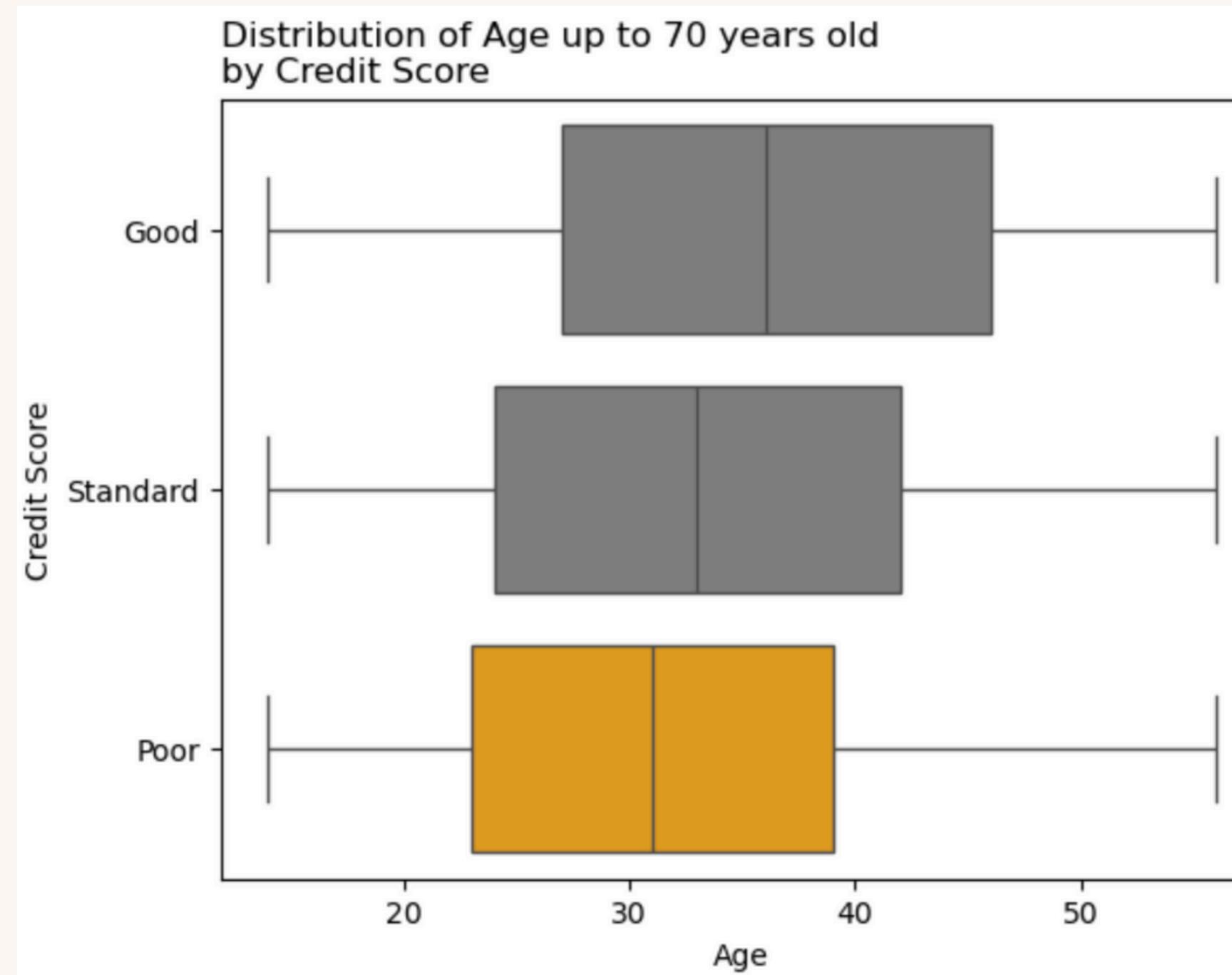
Analysis



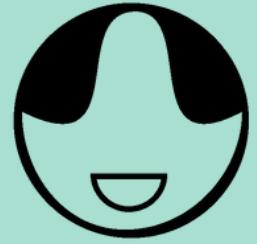
- There are 29% of poor credit scores in the dataset.
- It's quite a high number and might increase the number of non-performing loans, which can lead to an increase in the bank's financial risk.



Analysis



- Box plot shown that the median age of customers with poor credit scores are the lowest, but not significant.
- The finding indicate that these distribution appear by a coincidence. In other words, **younger customers won't automatically have bad credit scores.**



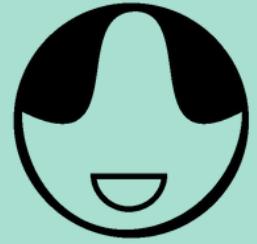
Modeling

The machine learning algorithm used in this project to build a classifier model is random forest.

The model's performance is shown below.

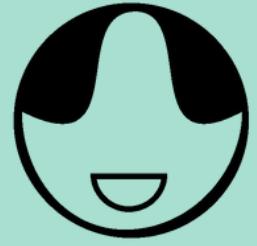
91%
Precision on
Train Set

74%
Precision on
Test Set



Conclusion

- **Customers with poor credit score have slightly higher number of credit card on average.**
 - Outstanding Debt, Delay from Due Date, and Credit Card Age are top 3 variables with a high correlation to Credit Score.
 - There are 29% customers with poor credit score.
-
- The Model is highly overfitted, thus it's not reliable for the task
 - Generalization can be used to improve it's performance
 - Improving the quality of the dataset can also greatly improve the model's performance



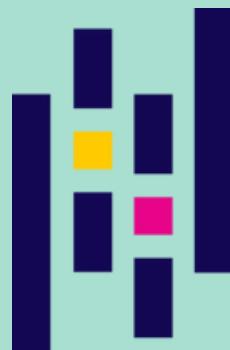
Tech Stack



Scikit Learn



Numpy



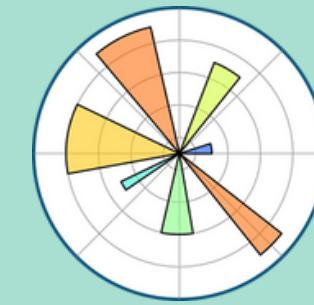
Pandas



Imbalanced-learn



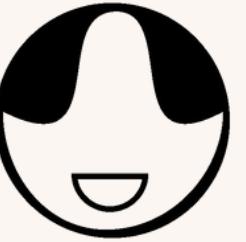
Feature-engine



Matplotlib



Python



Thank You

