Financial Palm Reading: Credit Card Customer Segmentation and Churn Prediction

Chelsea Zhang, Haomin Fan, Vasu Sharma (Group 34)
Lucan Yan, Xuhan Tong, Bowen Zhang (Group 39)

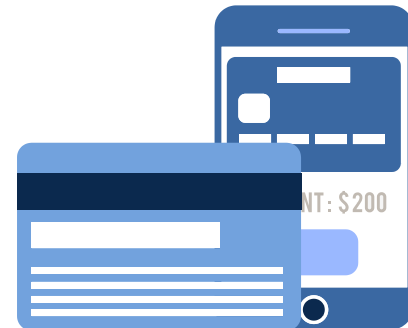# The Dataset

- **Source**

  Found on Kaggle.com

- **Rows**
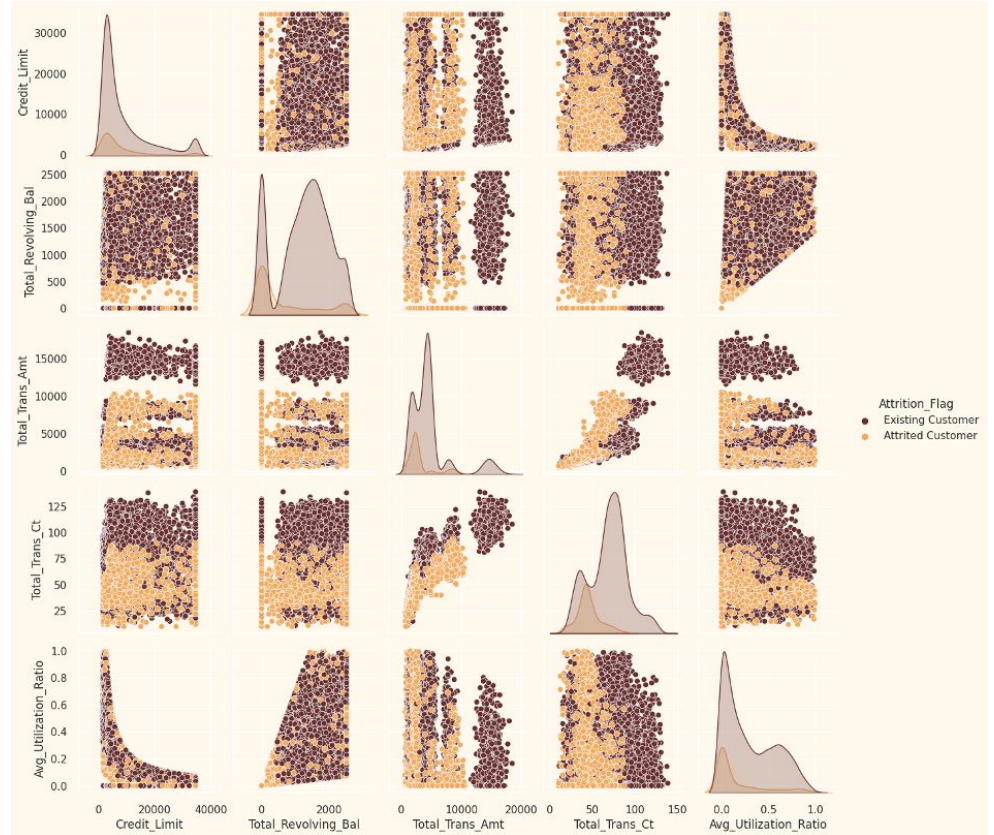
  10,127

- **Variables**

  19

- **Outcome**

  Attrition_Flag(Existing/ Attrited Customer)

Relative Plot Of Some Selected Features

Total Revolving Balance, Total

Transaction Amount, Total

Trans Count, Average

Utilization_Ratio

can be used as important
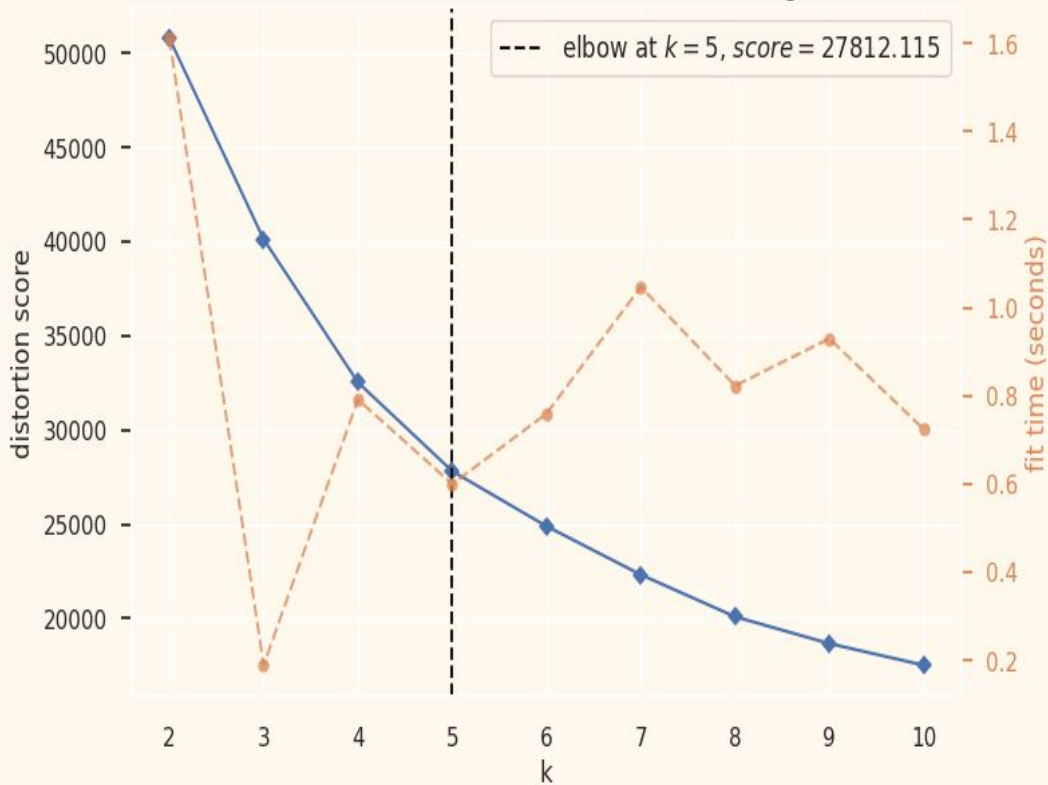
predictors for churning.

Dimensionality Reduction Using PCA

- Reduced to 3 dimensions

- By reducing the number of features, we are easily to identify and visualize the patterns and relationships within the data.
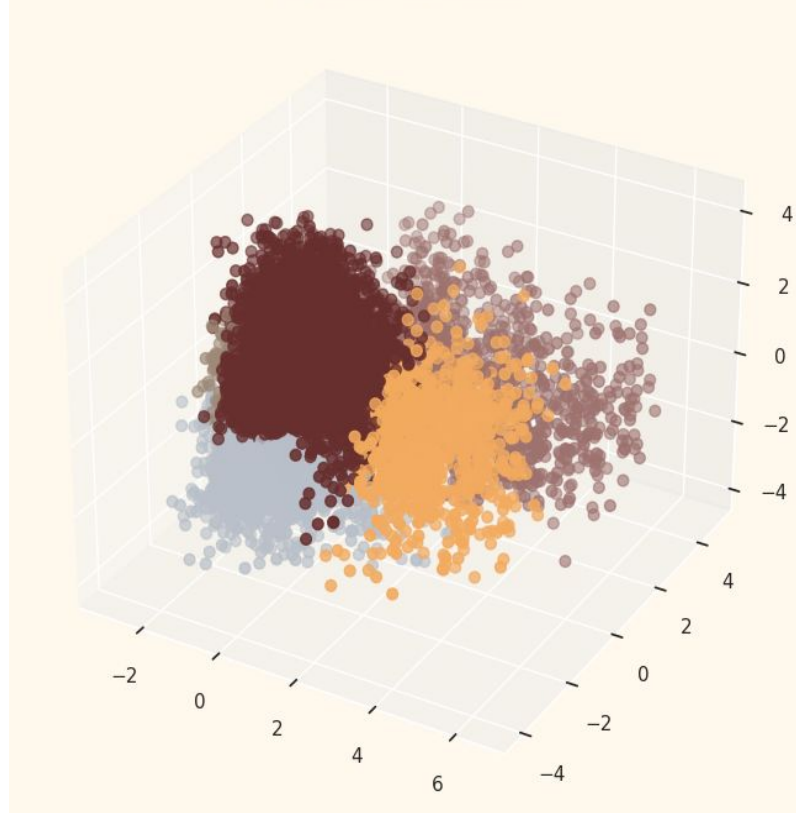
A 3D Projection Of Data In The Reduced Dimension
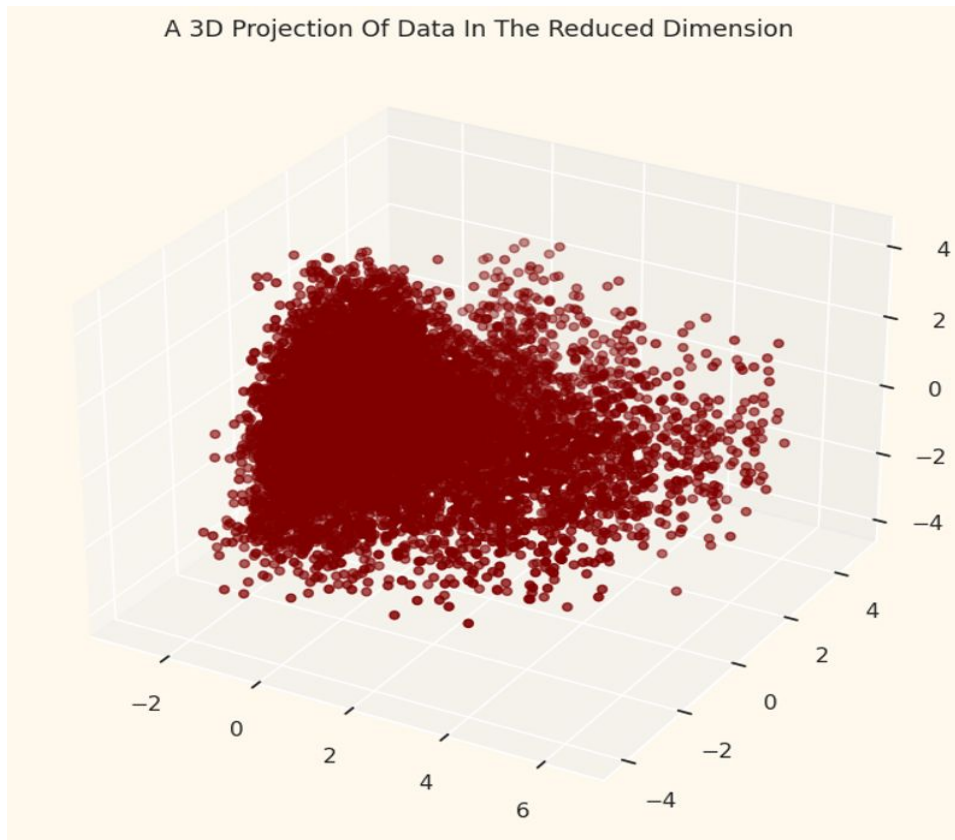
Clustering with Elbow rule

Who are those five clusters?

Fresh Wallets

Big Spenders

Balanced Buyers

Selective Shoppers

Silver Savers
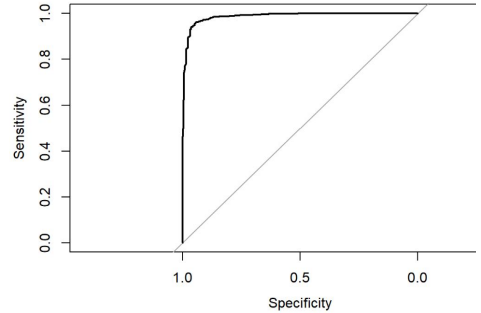
Distribution Of The Clusters
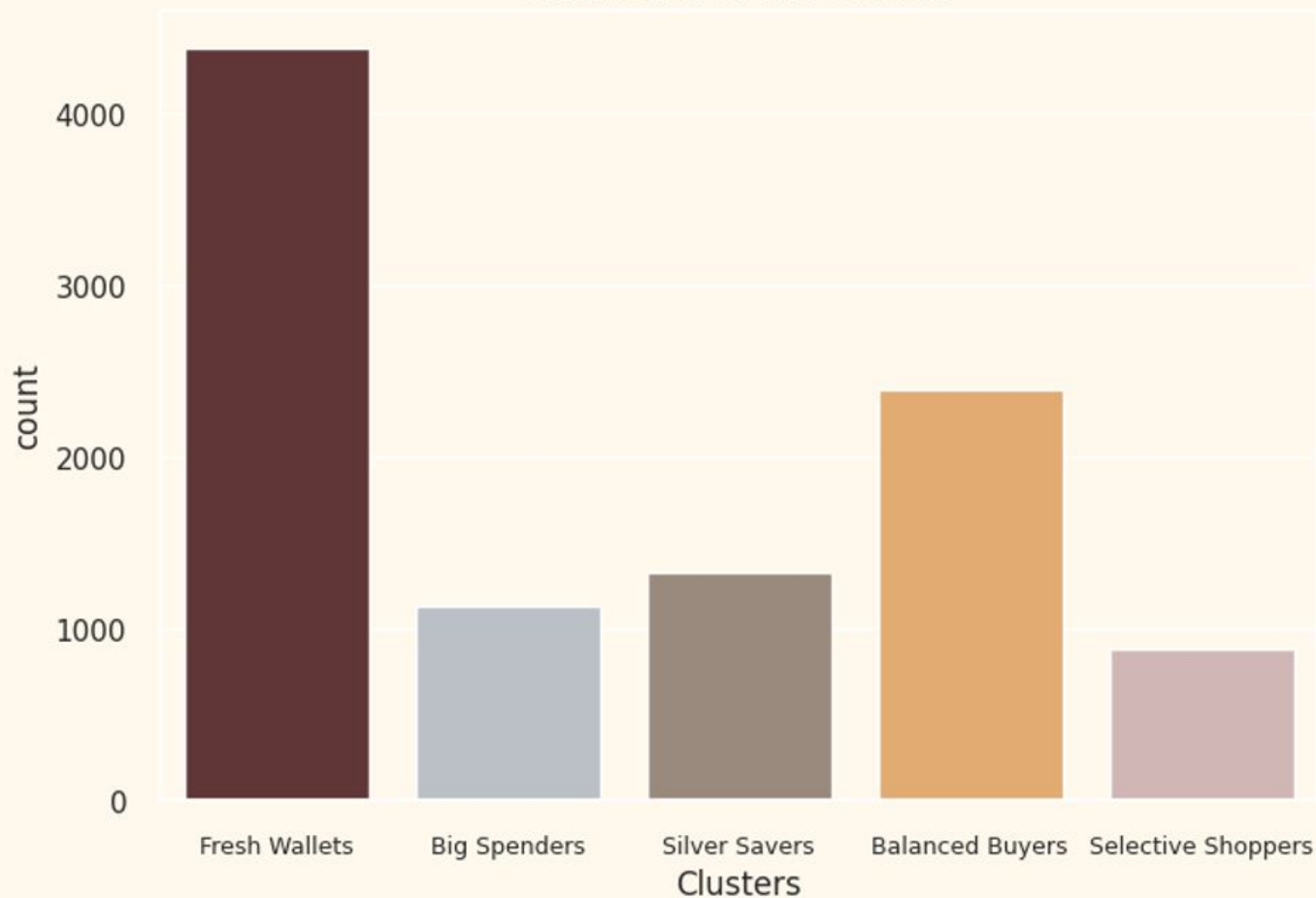
Winning model: Random Forest

Output Summary

Final Tree with Rpart
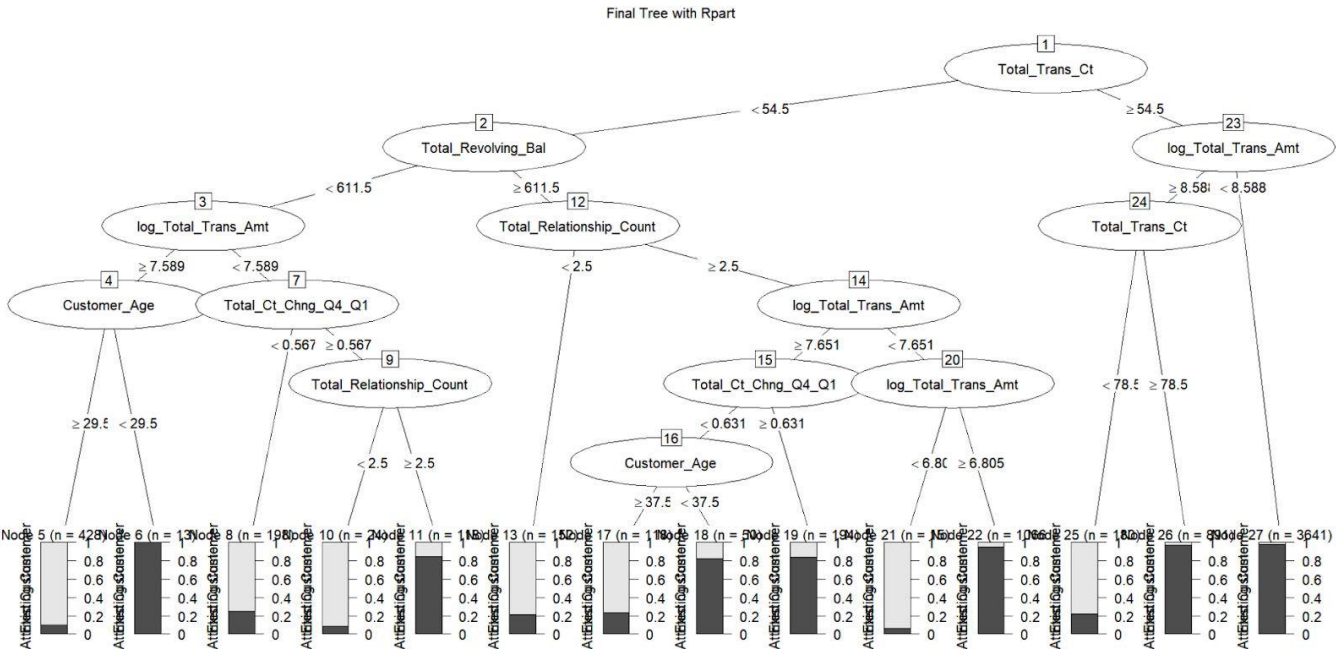
Model Performance

- AUC = 0.9877, which is the best performance so far.

Findings

- Best predictors:
  - total revolving balance
  - total transaction amount
  - total transaction count
  - average utilization ratio
  - user's total number of financial products in this bank
  - change in transaction amount from Quarter 4 over Quarter 1

- THESE ARE THE SAME FROM CLUSTERING

Conclusion

- "You are what you spent…But don't worry, there is always room for improvement."

- "All models are wrong, but some are useful." - George E. P. Box