



# Today's Agenda

1 Introduction

2 The brief

3 Data set analysis

4 Models

5 Next Steps

6 Learnings

# Introduction

A bank designs a focused marketing study, with 18,000 current bank customers. This focused approach allows the bank to know who does and does not respond to the offer, and to use existing demographic data that is already available on each customer.

The task is to build a model that will provide insight into why some bank customers accept credit card offers. There are also other potential areas of opportunities that the bank wants to understand from the data.



Risk Analyst Team :  
Léa, Vikas & Danny

Documents provided

# The brief

1

Project  
Details

2

SQL  
Questions

3

Tableau  
Classification

## Project steps

Trello



Github Repository

Python Cleaning and  
Exploration

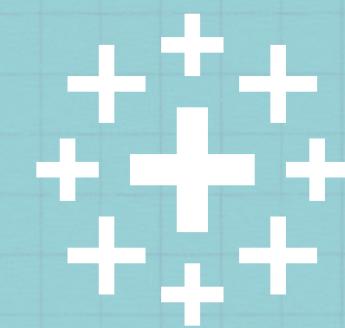
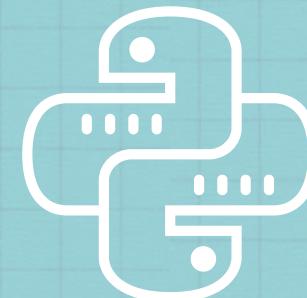
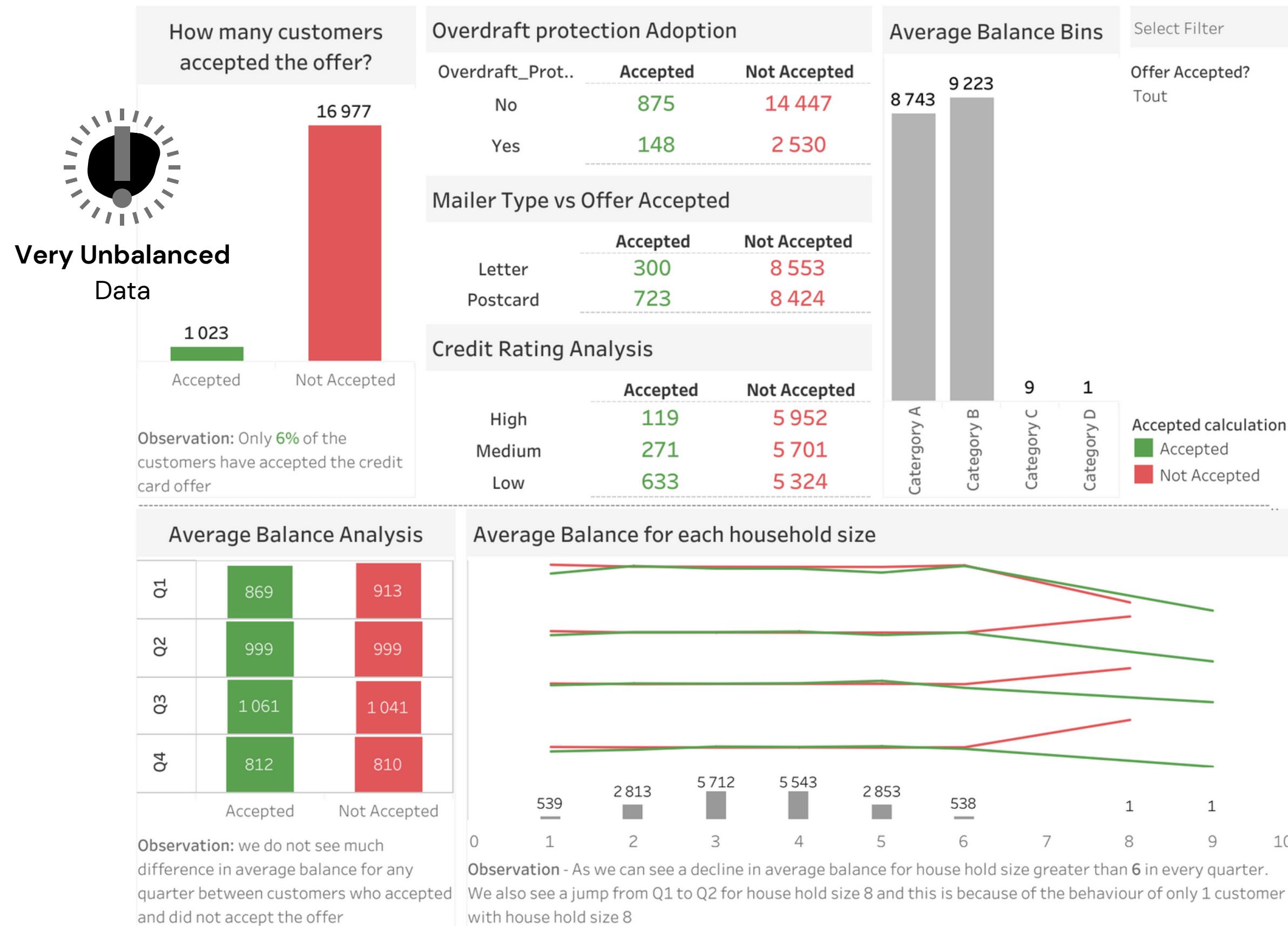


Tableau Exploration  
and Dashboard

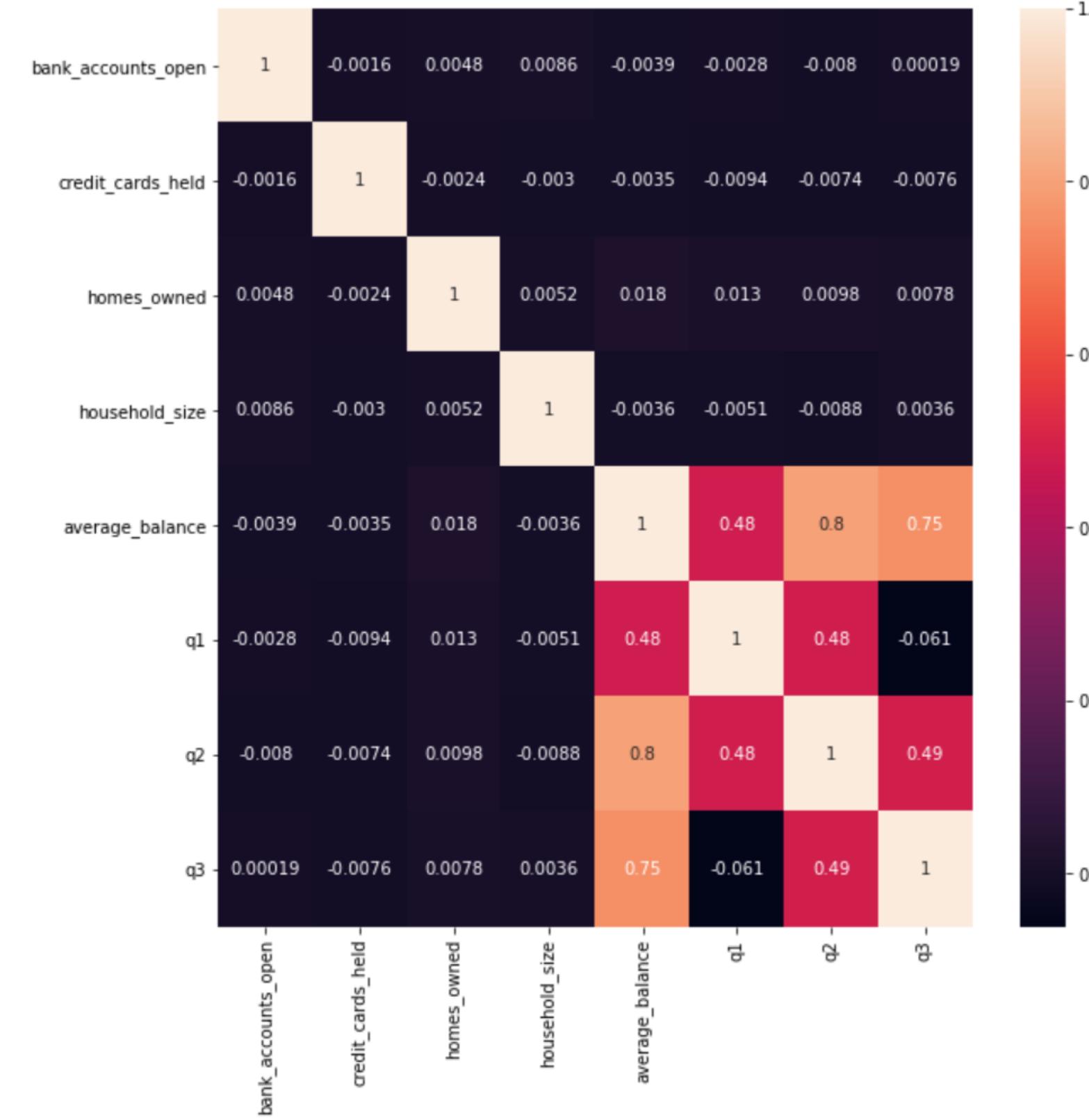


SQL Queries

# Dataset Analysis with Tableau



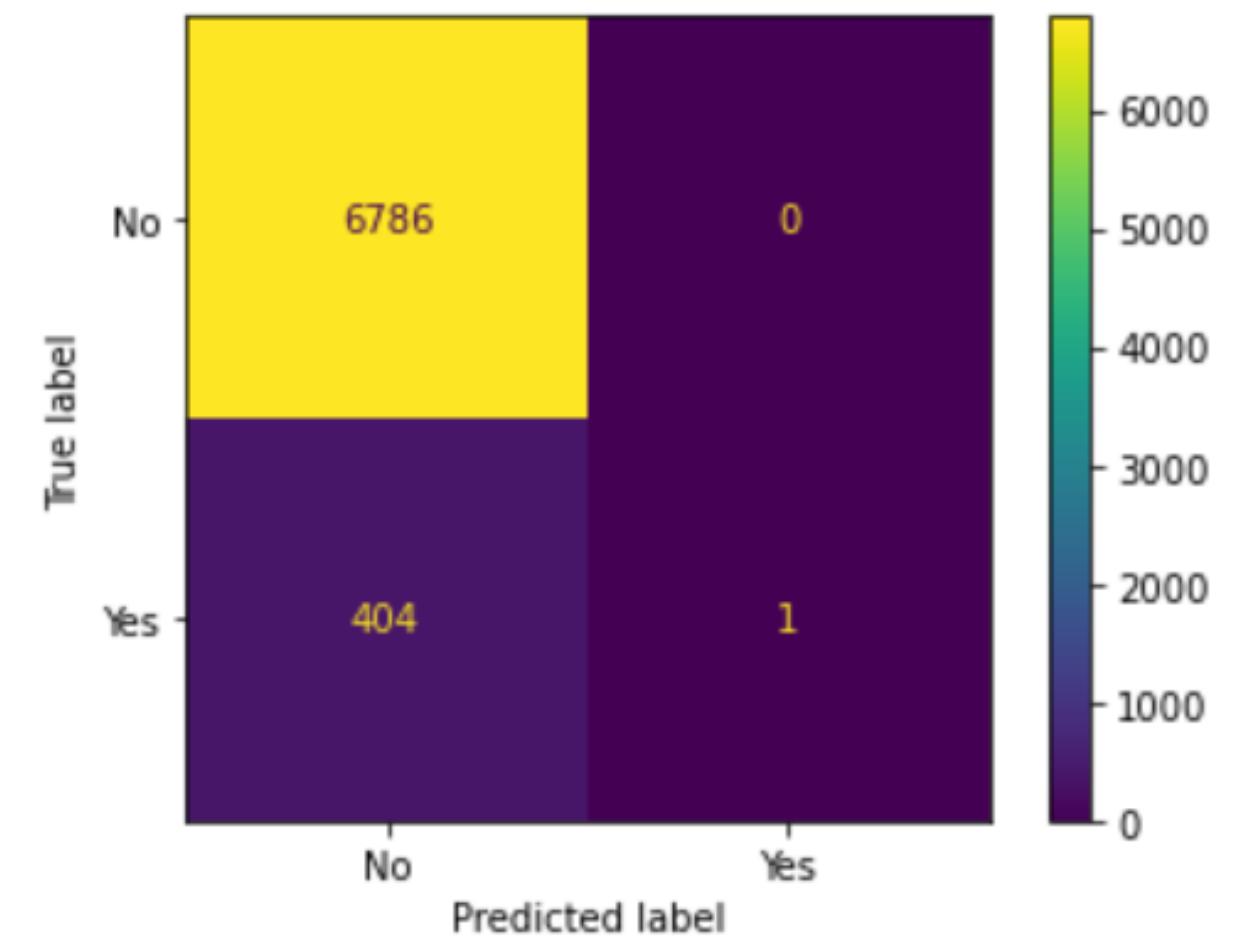
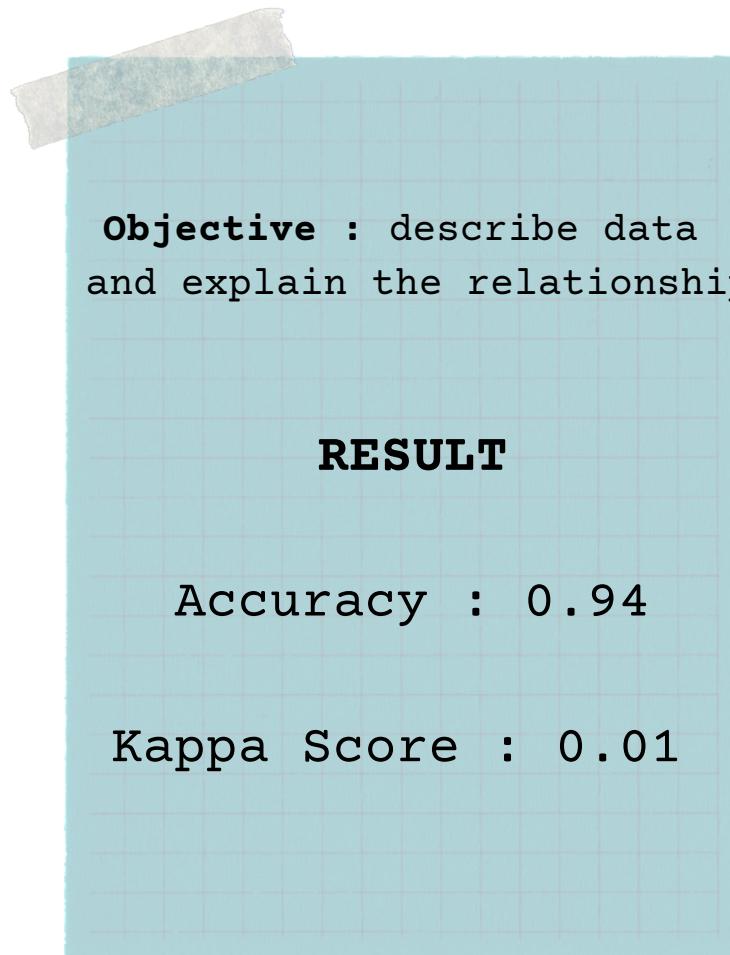
# Correlation Analysis with Python



**Correlation between Average Balance and Q1, Q2 and Q3.  
→ Drop the columns**

# Output : Models tried

## Logistic Regression



Confusion Matrix

# Output : Models tried

## Undersampling

**Objective** : undersample  
the majority class

### RESULT

Accuracy : 0.67

Kappa Score : 0.34

## Oversampling

**Objective** : oversample  
the minority

### RESULT

Accuracy : 0.83

Kappa Score : 0.65

## Random Forest

**Objective** : get prediction  
from each tree and selects the  
best solution by means of voting

### RESULT

Accuracy : 0.93

Kappa Score : 0.86

## KNN

**Objective** : calculate the distance  
of a data point to all other  
training data points

### RESULT

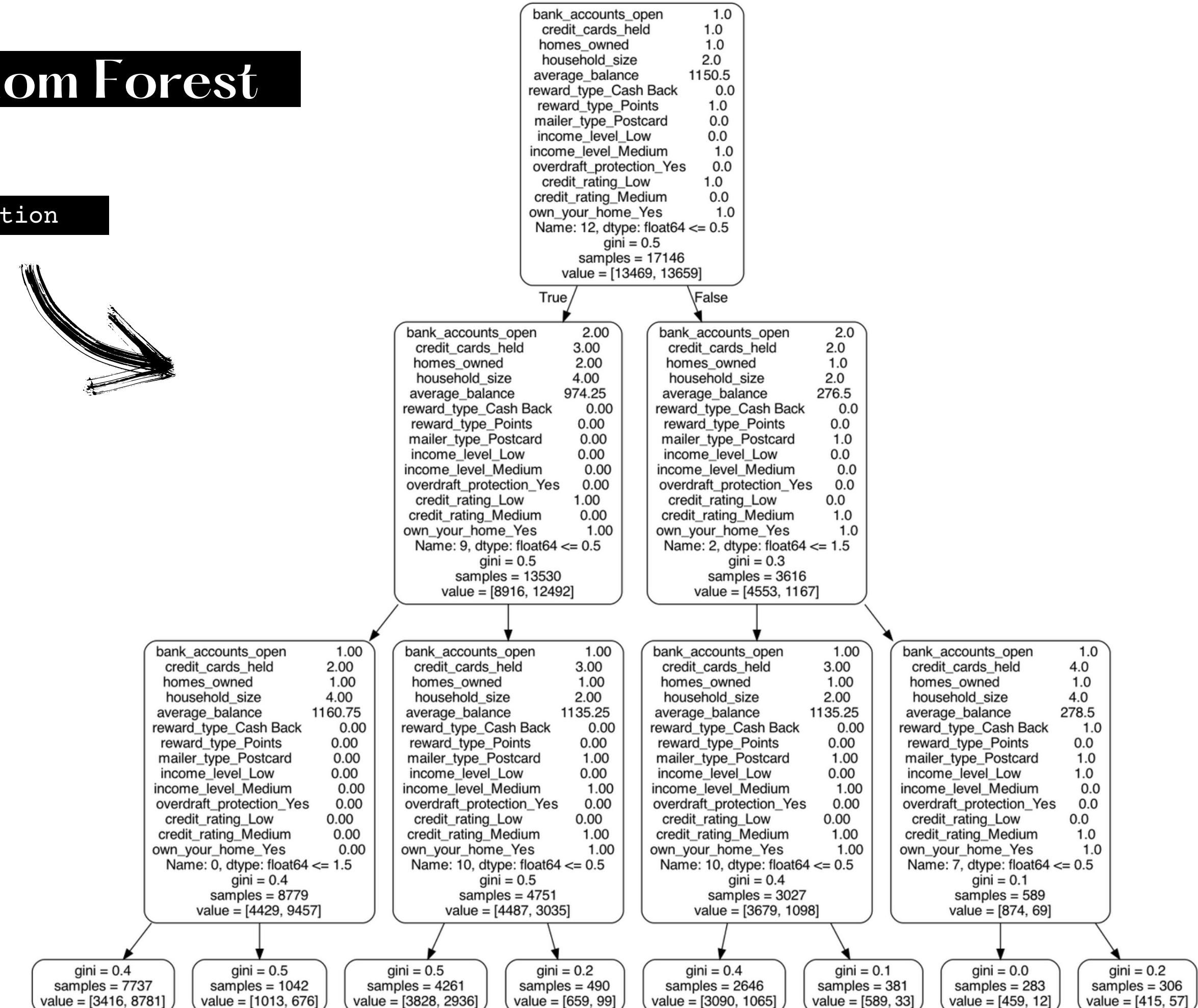
Accuracy : 0.87

Kappa Score : 0.75

# Models – Focus on Random Forest

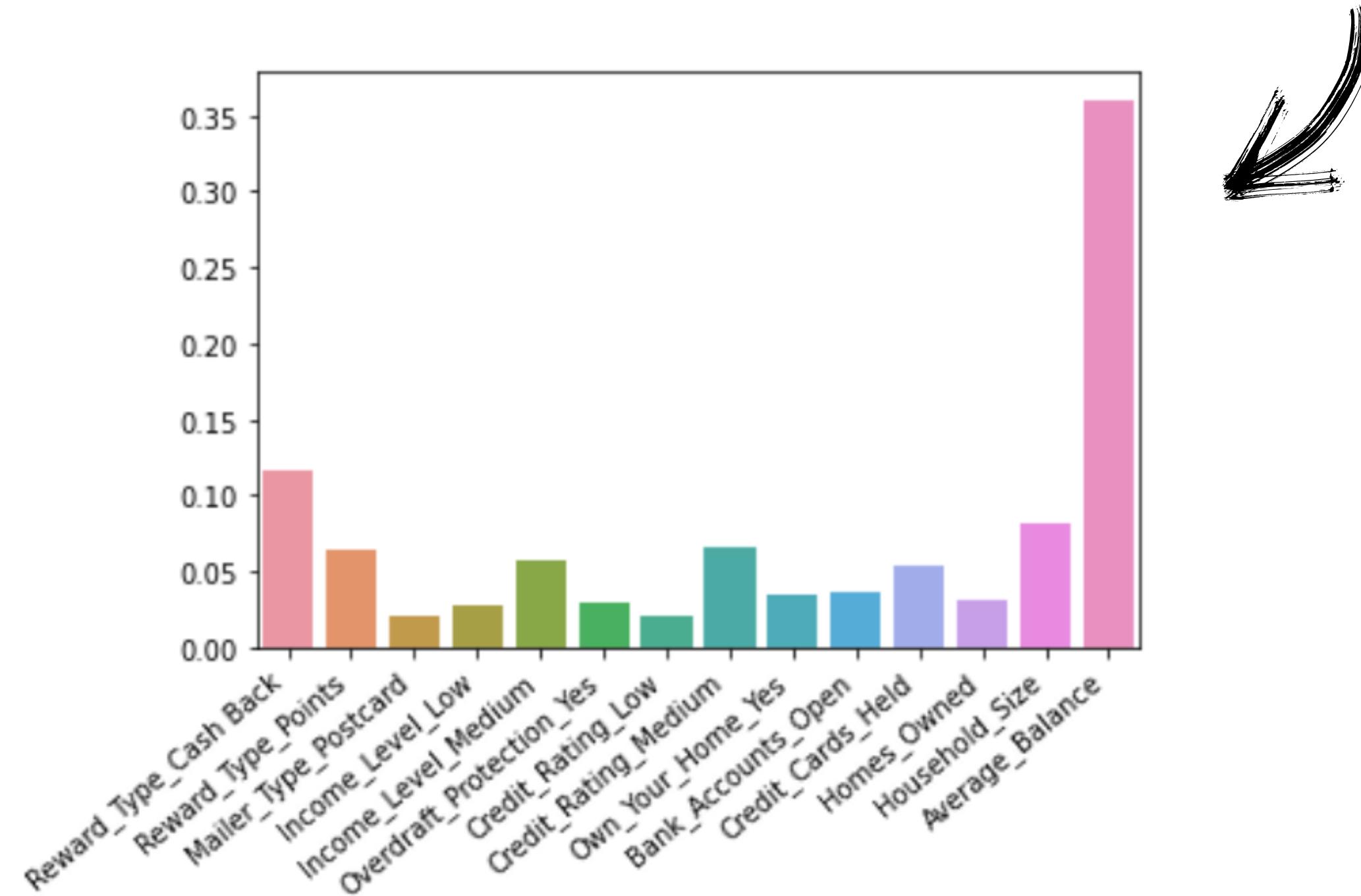
## Random Forest tree visualization

From Random Forest  
model to the  
visualization  
of a multitude of  
decision trees



# Models – Result Exploration

Show how much each feature affects the model



# Next Steps

1

Remove Outliers

2

Try with bins

3

Drop more columns

4

Transformation

# Learnings



Work in groups

Help each other and share our  
points of view

In-depth knowledge of the  
different models used

Conduct a data analysis project  
from exploration until presentation  
of results

Thank  
you!

The Risk Analyst Team :  
Léa, Vikas & Danny