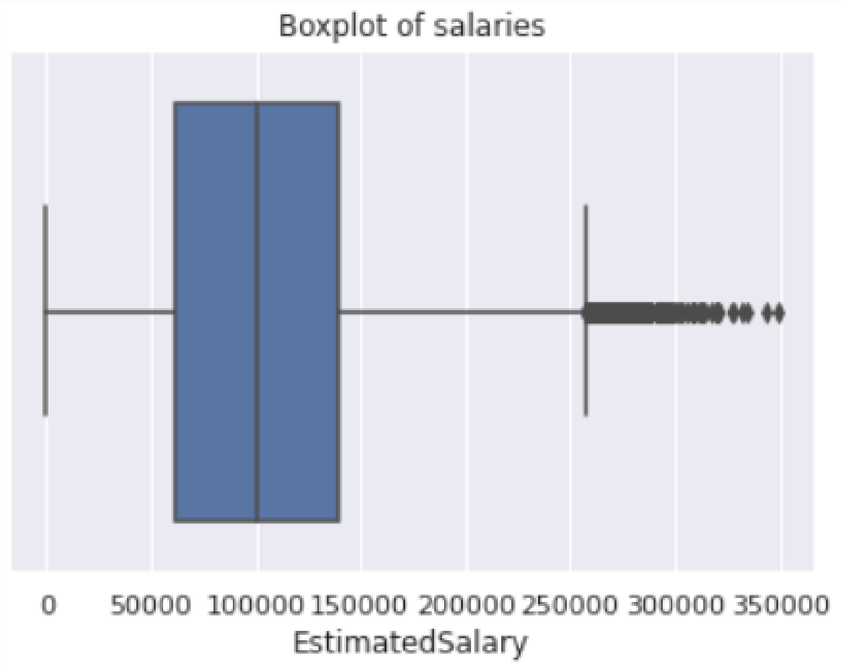
Who will stay 2 years or more?

Kin Security assumes 80% of costs. Clients must stay 2 years or more in order to profit. The goal is to identify those people based on their history, credit score and bank transactions.

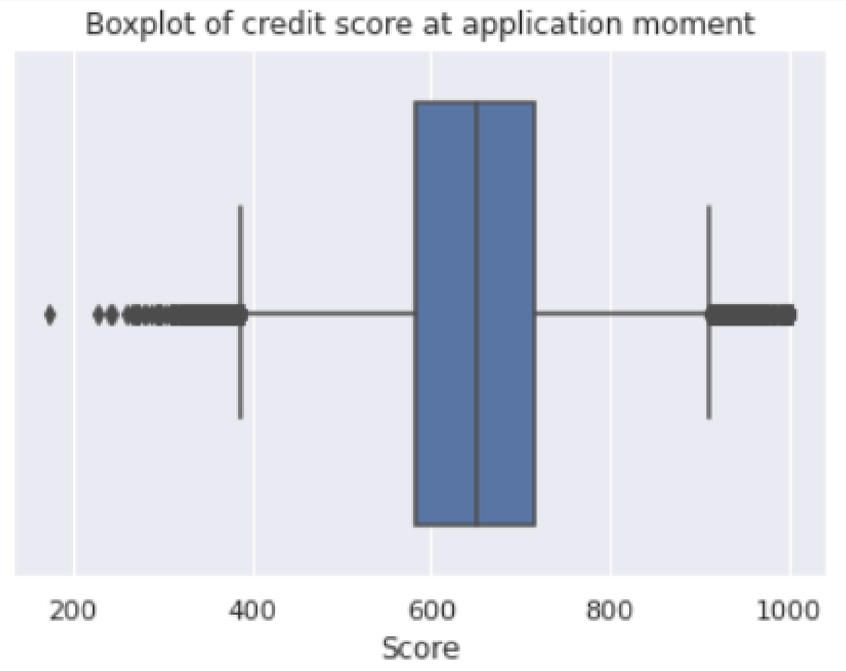
Desired population.
The filters.

- Contracts from 2015 onwards → 623242 records
- Operations in Italy were closed → 487424 records
- Clients with 75% of missing info removed → 464075 records
- Clients with one contract only → 118897 records

Variables and their stats

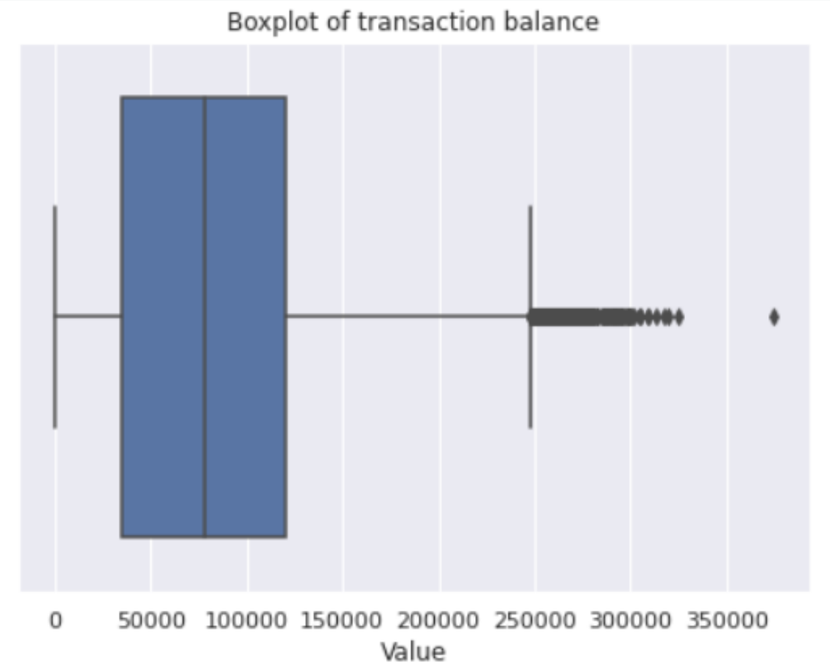


Mean: 101169
Std: 55698
Min: 0
Max: 349080

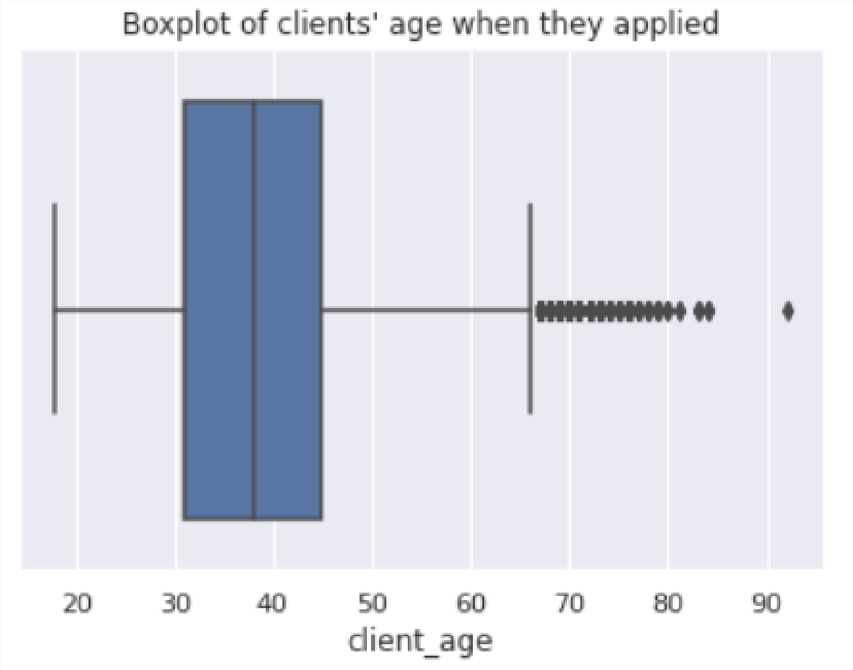


Mean: 650
Std: 97
Min: 174
Max: 1000

Variables and their stats



Mean: 80724
Std: 56654
Min: 0
Max: 374634



Mean: 39
Std: 10
Min: 18
Max: 92

Who will stay: Logistic Regression

Logistic regression is the model chosen to predict who will stay. Numerical variables were standardized and categorical variables were encoded. This way, the model performs better.

Why logistic regression?

As the goal is a binary classification and we have enough variables, this model fits the context. Furthermore, in this case, it's not computationally expensive.

Having trained the training set and then compared to the testing set, we got an accuracy of 79.24%.

Thus, commercial campaigns can be based on this model.

See the full analysis on GitHub

Repository on GitHub

Created by Axel Yaguana