Problem Statement

Analyze loan application at the time of application to identify applicants who are capable of repaying loan are not rejected

- The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it as their advantage by becoming a defaulter. Consumer finance company which specializes in lending various types of loans to urban customers have to use EDA to analyze the patterns present in the data. This will ensure that the applicants are capable of repaying the loan are not rejected.
- EDA to understand how consumer attributes and loan attributes influence the tendency to default.

Analyze Approach

— Data Imported and analyzed to ignore/delete unimportant columns
- — Columns with very high percentage of NA values
- — Outliers for different columns observed
- — Columns with similar trends observed using correlation
- — Columns not having any significance with regards to the analysis carried out

— Target variable set as:-
- — 1: Difficulty in repayment of loan
- — 0: All other cases

— Derived metrics created to get better analytic understanding

— Univariate and segment analysis performed to get more data insights

— Bivariate analysis performed to identify possible trends

— Identify variables/columns that have which are drivers of identifying clients facing difficulty in loan repayments

Univariate Analysis

— 90% loans are cash loans, and only a few revolving loans

— Most loans for more than Rs.10lacs

— 65% of applicants are female

— Most applicants living in house or apartments and are working

— Laborers and applicants with no children make up the largest demographic

— Most applicant aged between 30-50 with income of around 1 to 2 lacs

— All applicants have mobile phones but only a small number have emails which could indicate that most applicants don't have regular access to the computer

— Most applicants from regions with moderate living conditions indicated by region rating of two

Bivariate Analysis

— Analysis shows no significant trend between amount of loan and difficulty in repayment or the income of the applicant

— No specific trend to be seen in demographics such gender, employment or type of housing

— Higher percentage of applicants with high yearly annuity to total credit amount face difficulty in repayment

— Younger applicant (less than 30) have higher percentage and older applicants (more than 60) have lower percentage of applicant facing difficulty in repayment

— High scores from external data sources indicate less likelihood of facing difficulty in repayment

— High number of previous refusal have a strong correlation with difficulty faced in repayment

Annuity vs Target

— Derived metric created:

    — Annuity Percentage = Amount of Annuity/Goods Price %

— Percentage of applicants facing difficulty in repayment have higher annuity

— Chance of facing difficulty in repayment increase as percentage of annuity goes above 5%
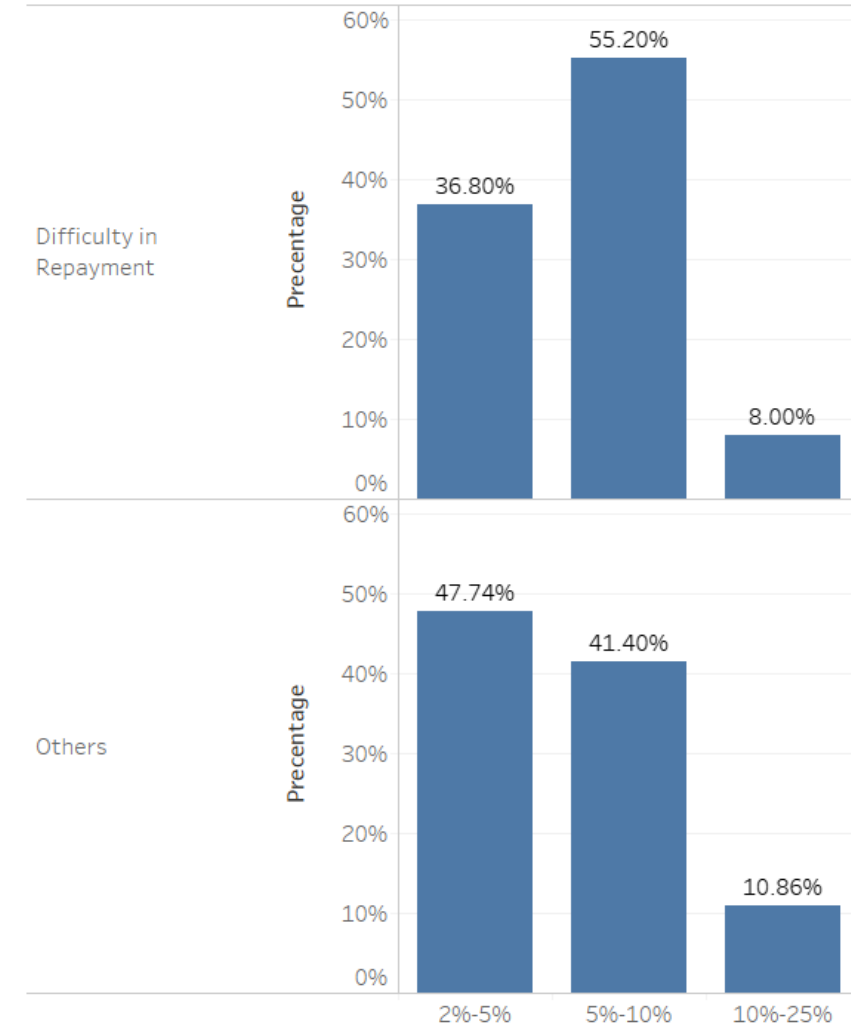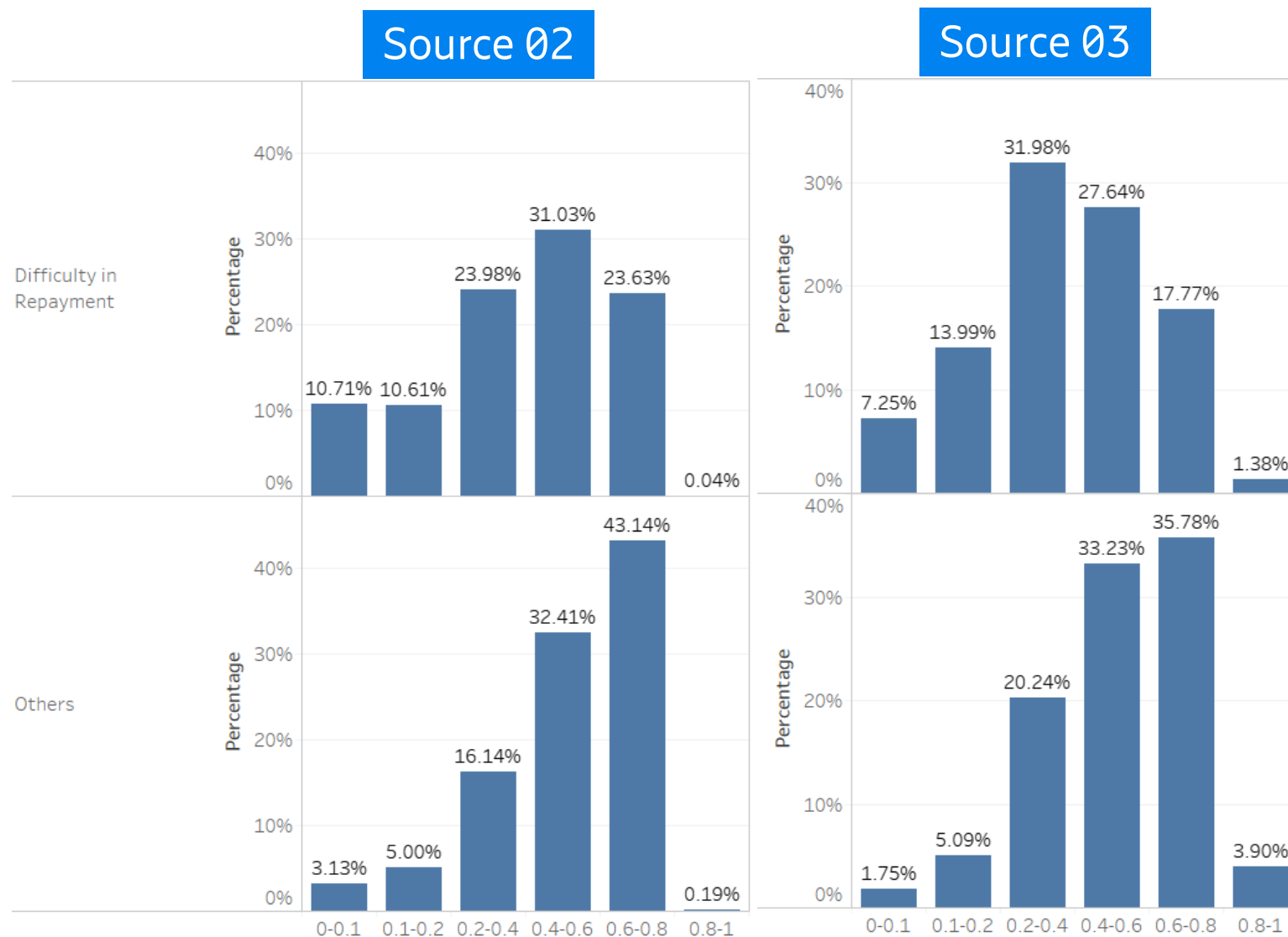
AGE vs Target

— Younger applicant of age below 40 find it more difficult to repay loans

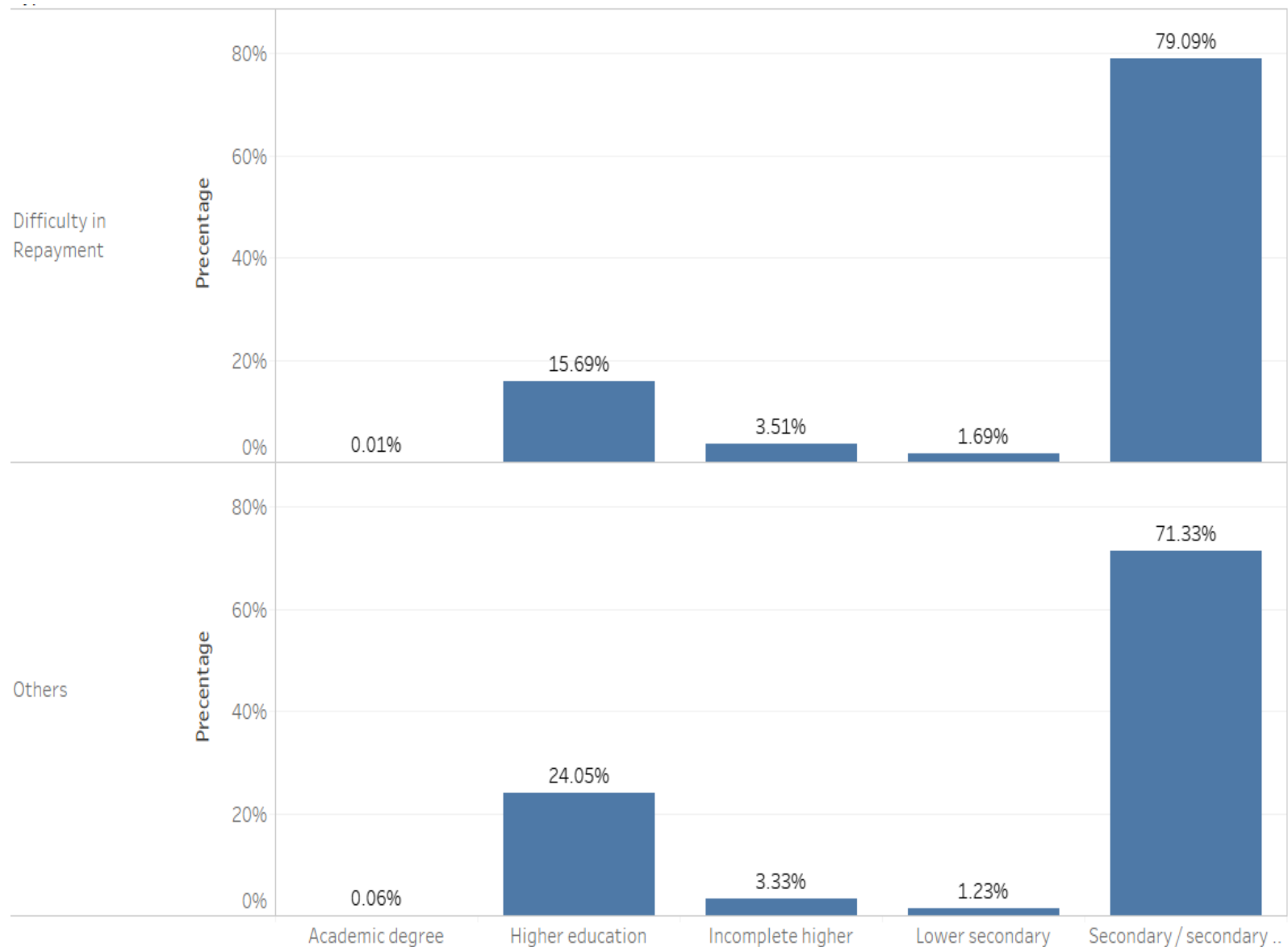— Older applicants of age above 50 find it easier to repay loans
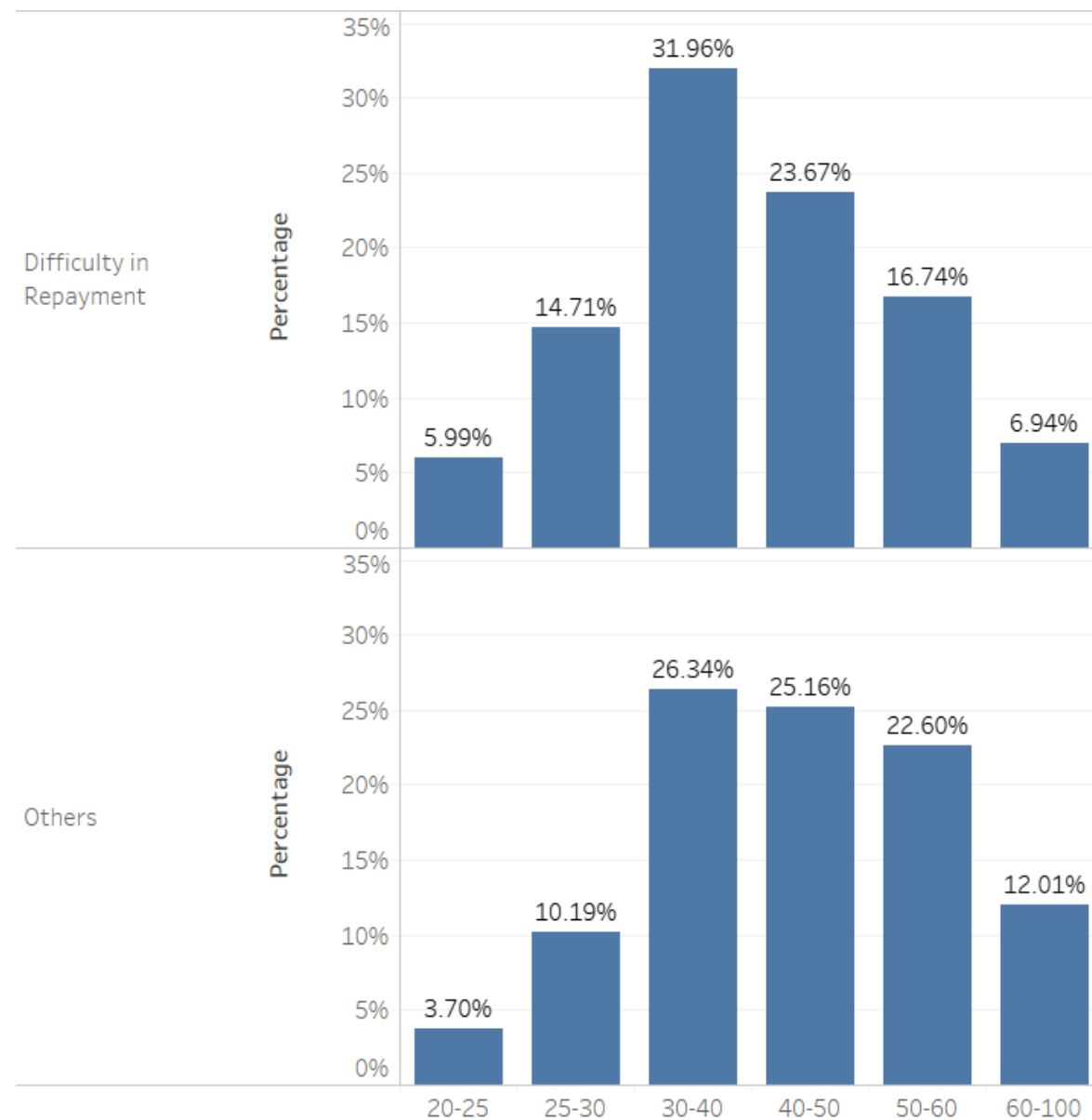
External Scores vs Target

— Both external scores indicate that scores below 0.4 have high correlation with person facing difficulty in repayment

— Source 03 better at identifying applicants with repayment difficulty

— Source 02 at identifying other types of applicants

Source 02 — Difficulty in Repayment
- 0-0.1: 10.71%
- 0.1-0.2: 10.61%
- 0.2-0.4: 23.98%
- 0.4-0.6: 31.03%
- 0.6-0.8: 23.63%
- 0.8-1: 0.04%

Source 02 — Others
- 0-0.1: 3.13%
- 0.1-0.2: 5.00%
- 0.2-0.4: 16.14%
- 0.4-0.6: 32.41%
- 0.6-0.8: 43.14%
- 0.8-1: 0.19%

Source 03 — Difficulty in Repayment
- 0-0.1: 7.25%
- 0.1-0.2: 13.99%
- 0.2-0.4: 31.98%
- 0.4-0.6: 27.64%
- 0.6-0.8: 17.77%
- 0.8-1: 1.38%

Source 03 — Others
- 0-0.1: 1.75%
- 0.1-0.2: 5.09%
- 0.2-0.4: 20.24%
- 0.4-0.6: 33.23%
- 0.6-0.8: 35.78%
- 0.8-1: 3.90%

Education vs Target

— Applicants with Higher education less likely to face difficulty with loan repayment
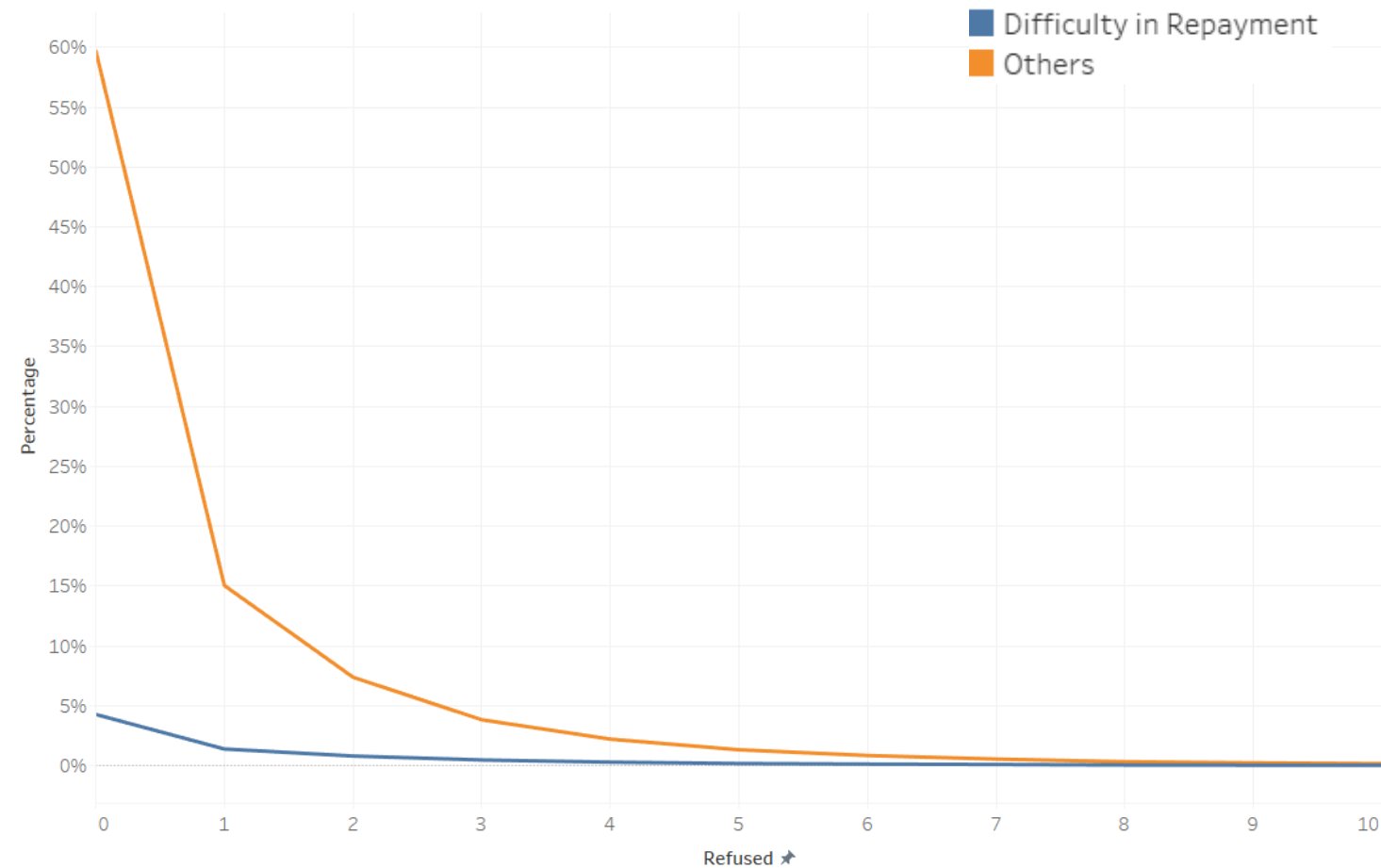
Previous Refusals vs Target

— As number of previous refusal increase less likely is the applicant not to face difficulty in repayment

— If 2 or less loans previously refused then likely of the clients not facing difficulty in repayment are high

Precentage of previouly refused loan requests

Conclusions

- Chance of facing difficulty in repayment increase as percentage of annuity goes above 5%

- Younger applicant of age below 40 find it more difficult to repay loans than older applicants

- Older applicants of age above 50 find it easier to repay loans than younger applicants

- Both external scores indicate that scores below 0.4 have high correlation with person facing difficulty in loan repayment

- Applicants with Higher education less likely to face difficulty with loan repayment

- If two or less loans previously refused then likelihood of the clients not facing difficulty in repayment are high