



BANK LOAN CASE STUDY

Dency Claris Thomas

PROJECT DESCRIPTION

This case study aims to give us an idea of applying EDA in a real business scenario. In this case study, apart from applying the techniques of Exploratory Data Analysis (EDA), we will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimise the risk of losing money while lending to customers.

LETS GET STARTED



APPROACH



I began by exploring the application_data.csv and previous_application.csv datasets using various exploratory data analysis techniques. I checked for missing data, cleaned and preprocessed the data, and removed any outliers that could skew the analysis.

Next, I conducted univariate, bivariate, and multivariate analyses to identify patterns, correlations, and relationships between variables in the datasets. I also visualized the data using various graphs and charts to help communicate my findings.

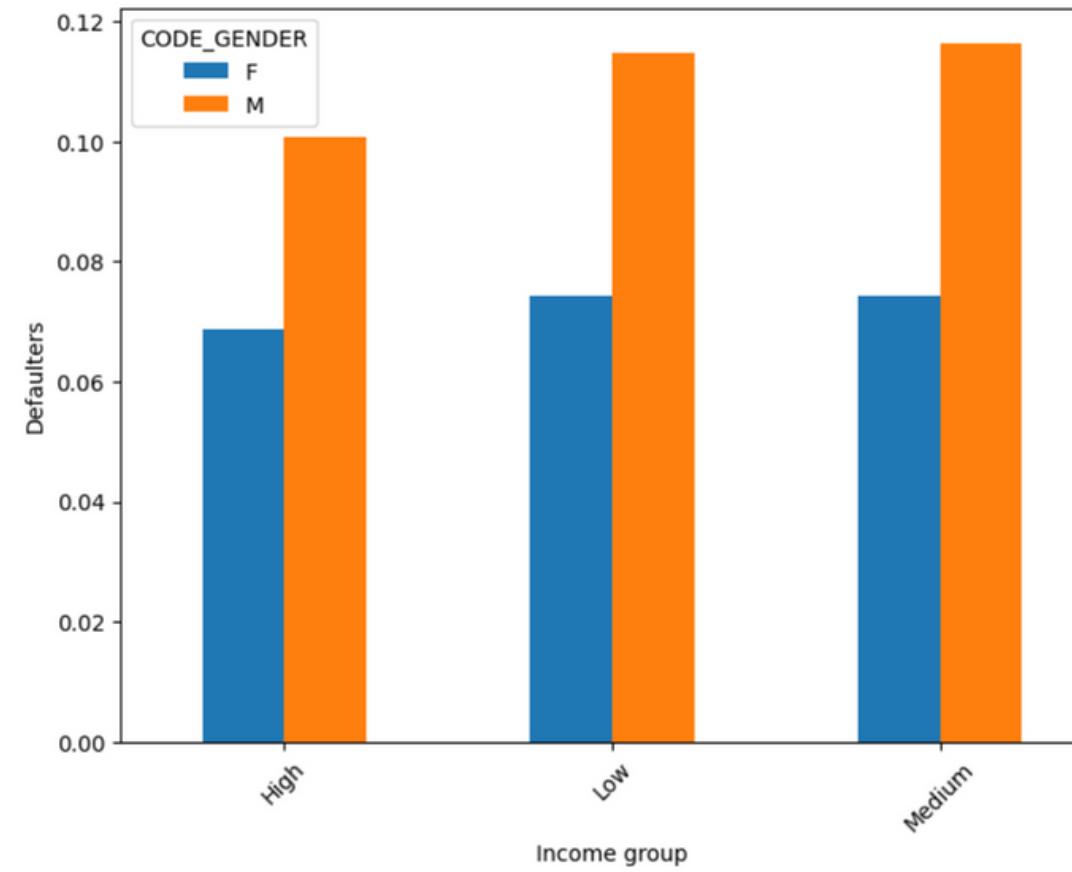
Finally, I drew conclusions from my analyses and presented them in a clear and concise manner. Overall, my goal was to gain insights into the loan application and approval process and to identify any factors that may be influencing approval decisions.

TECH STACK USED

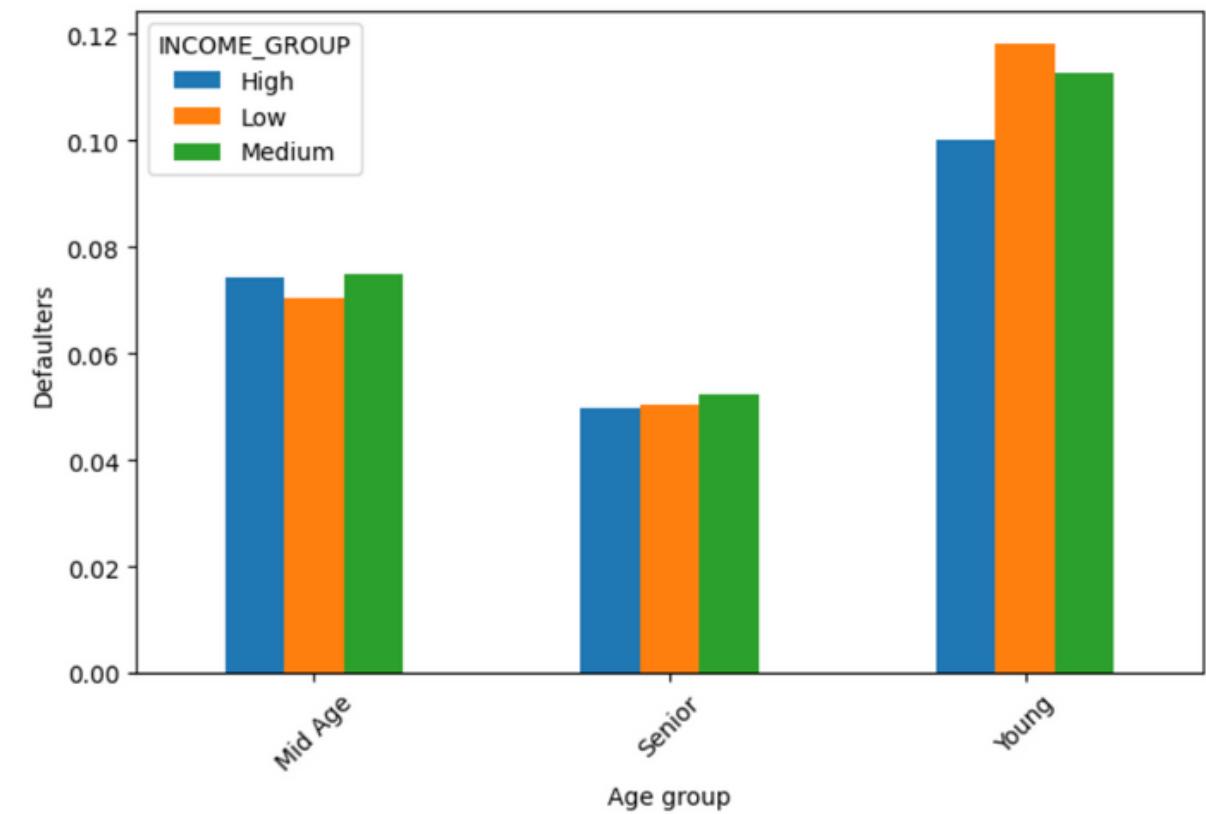
EDA using Python in JupyterLab

CURRENT APPLICATION

Income groups and gender



Income groups and age groups



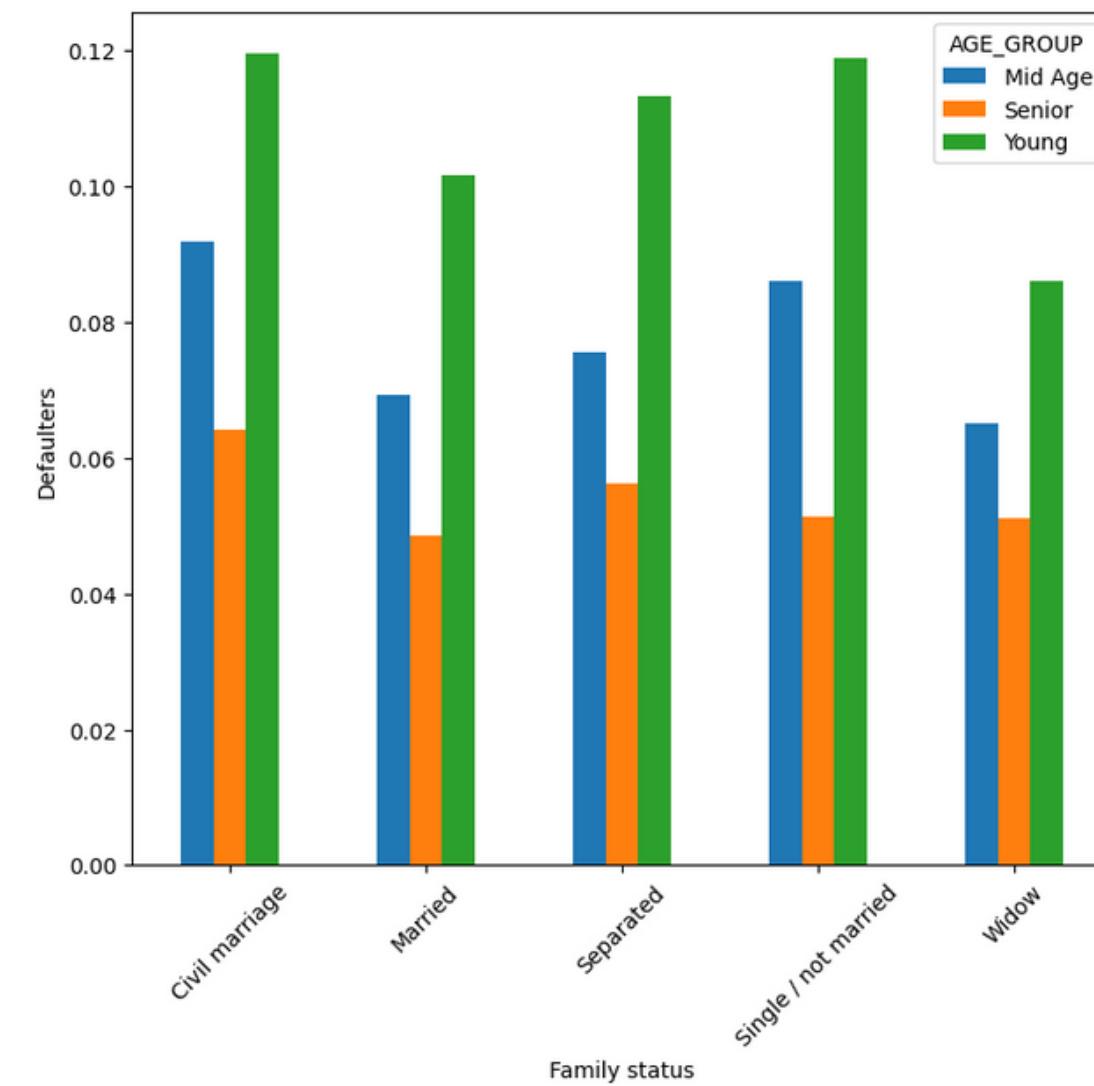
OBSERVATION:

1. High-income groups are fewer defaulters than comparatively lower-income group
2. Mid-age and senior people with all income groups are less defaulted.

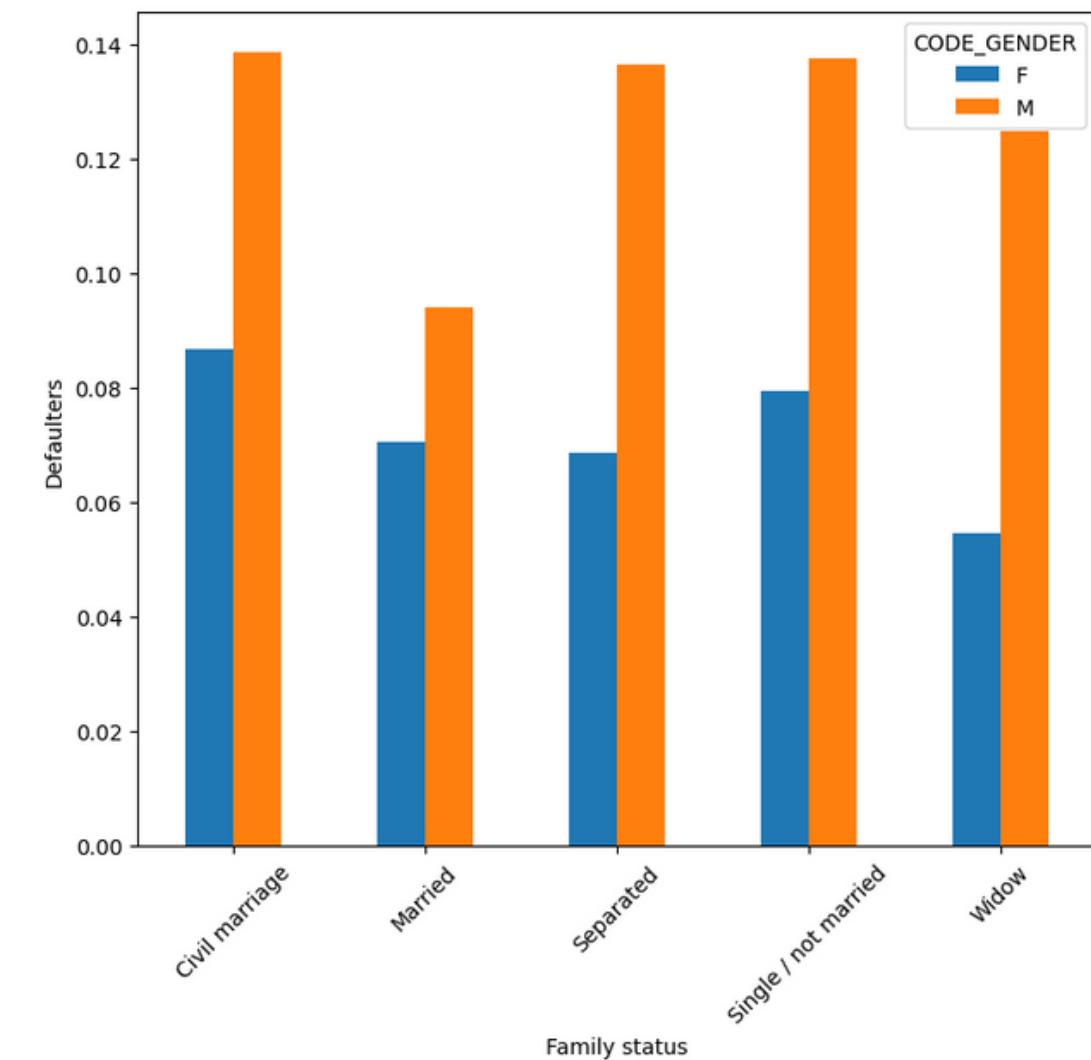
RECOMMENDATION:

1. Safer to grant loans to mid-age and senior citizen clients with high income.
2. Risky to grant loans to young people with low income.

FAMILY STATUS AND AGE GROUP



FAMILY STATUS AND GENDER

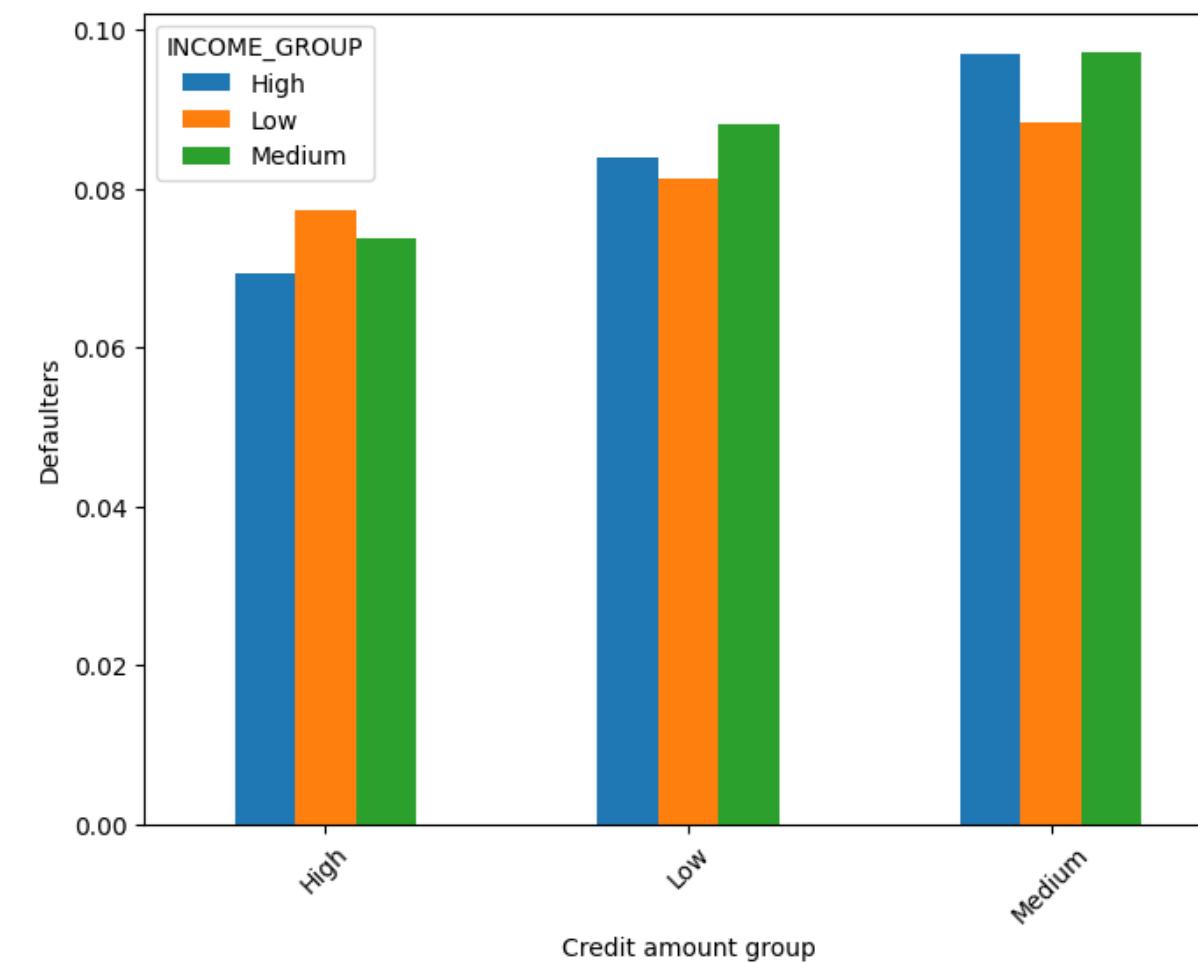
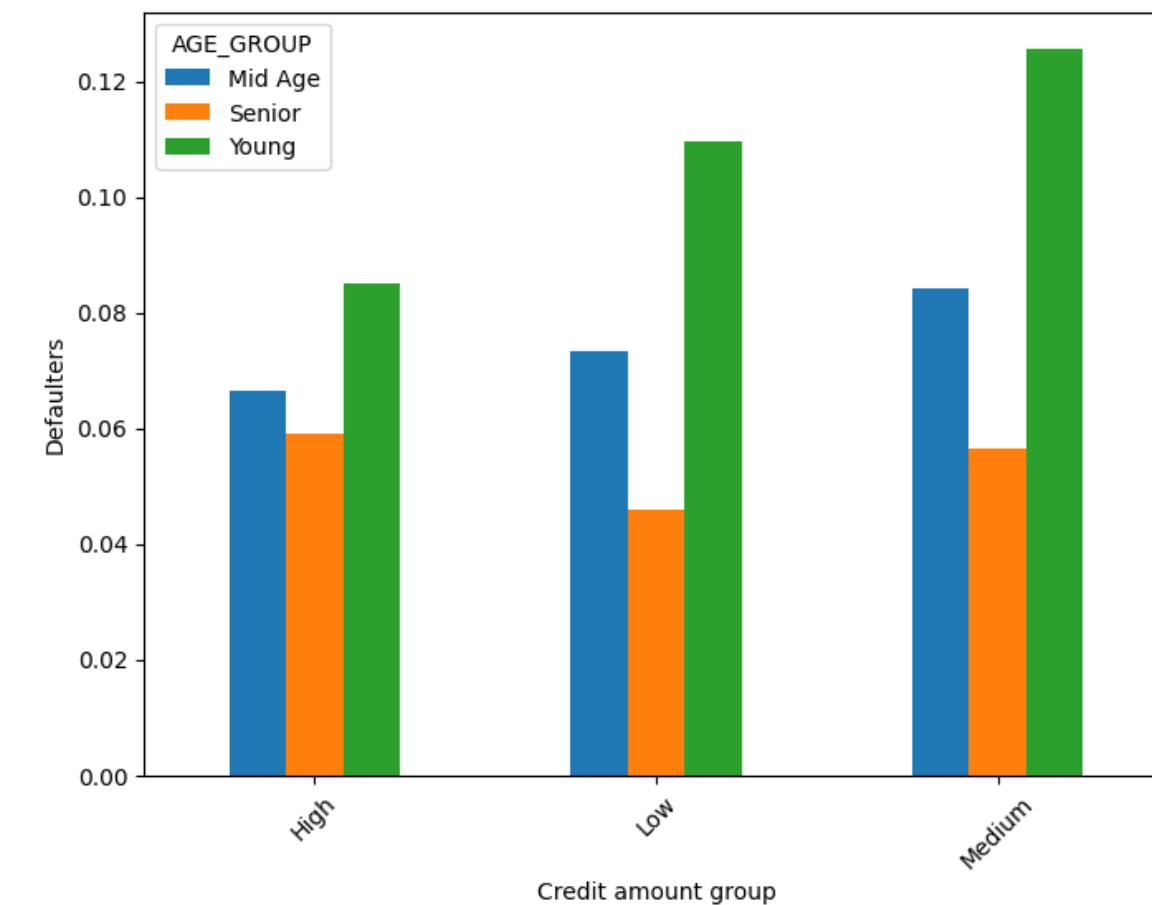


OBSERVATION:

1. Senior people irrespective of family status are less likely to be defaulted.
2. Young people are more likely to default in all family statuses
3. Males are more likely to be defaulted than Females.

RECOMMENDATION:

1. Better to grant loans to senior citizens with all family statuses.
2. It is risky to grant loans to single, civil married, and separated young men.

CREDIT AMOUNT GROUP & INCOME GROUP**CREDIT AMOUNT GROUP & AGE GROUP**

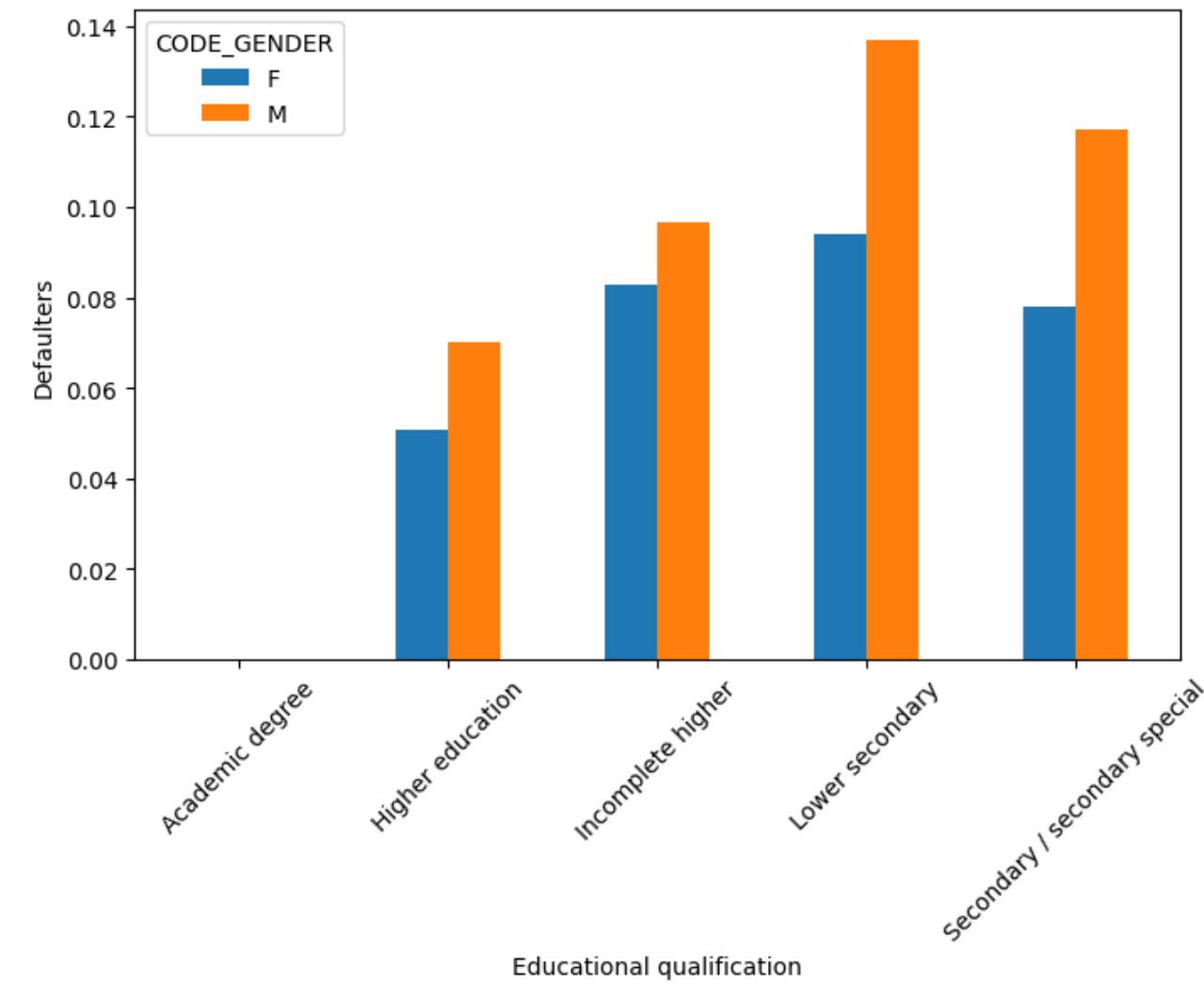
OBSERVATION:

1. Across all income groups clients with the medium amount credited is highly defaulted followed by the low and high credit amount.
2. Young clients with the medium and low amounts credited are most likely to default.

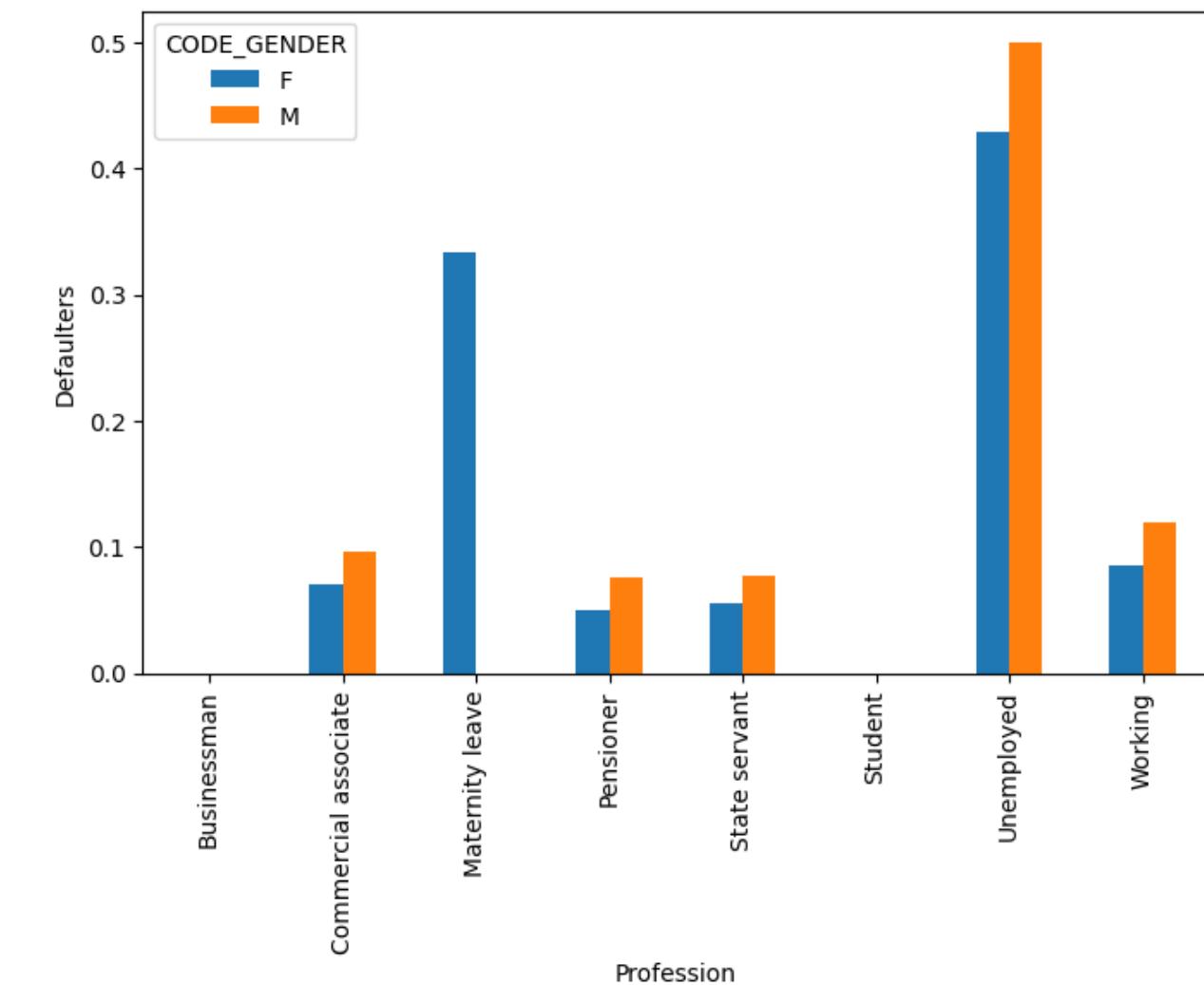
RECOMMENDATION:

1. Recommended granting a slightly higher amount of loan to all income groups.
2. It is very risky to grant medium and low amounts of loans to young clients.

EDUCATION AND GENDER



PROFESSION AND GENDER



OBSERVATION:

1. Higher educated people are less defaulted and lower secondary educated people are more.
2. Unemployed clients along with clients with maternity leave are heavily defaulted.

RECOMMENDATION:

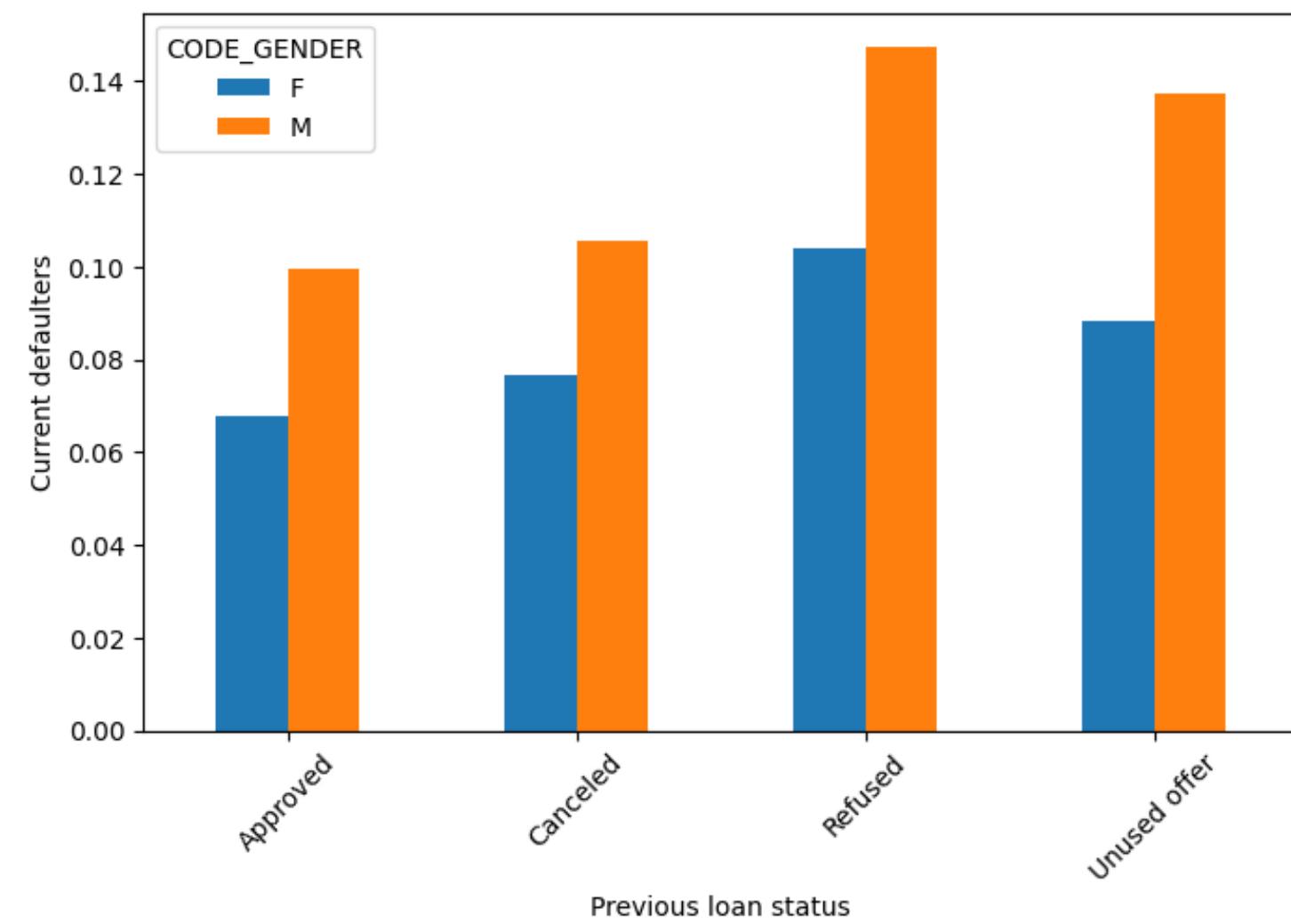
1. Safe to grant loans to higher educated clients across all professions except unemployed and women with maternity.

LOAN APPLICATION STATUS RELATIONS

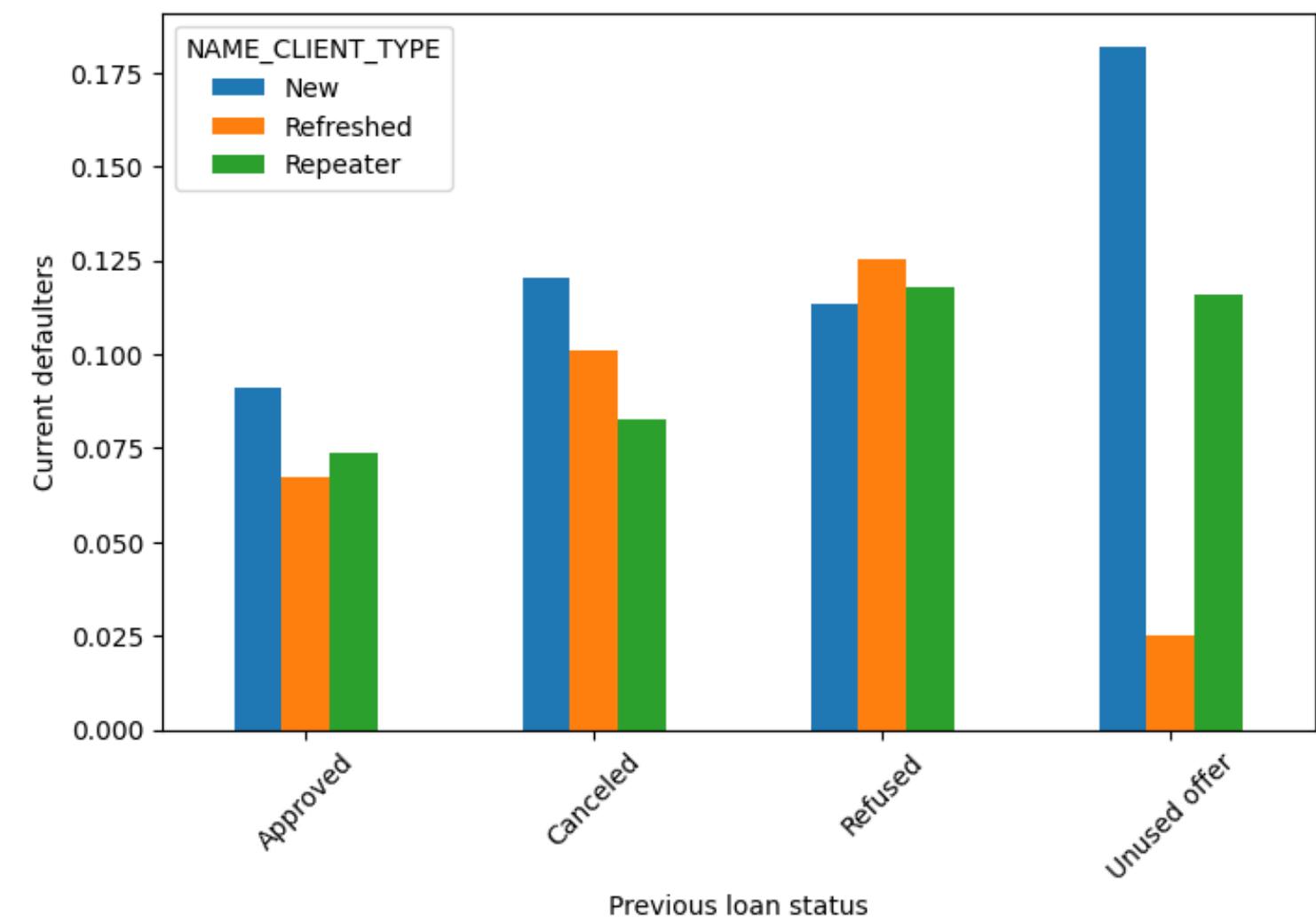
CURRENT AND PREVIOUS



PREVIOUS LOAN STATUS AND GENDER



PREVIOUS LOAN STATUS AND CLIENT TYPE



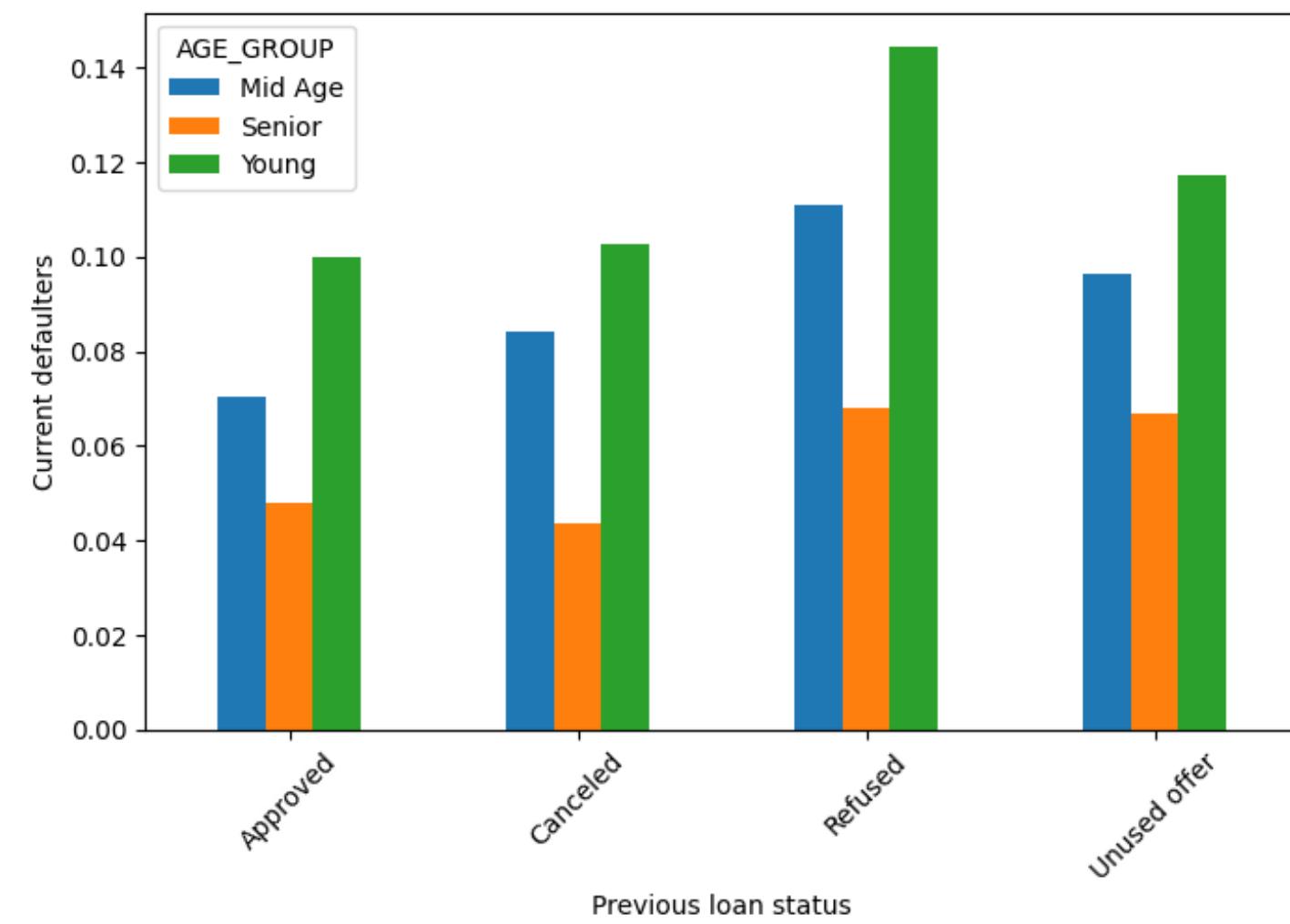
OBSERVATION:

1. Previously refused and unused offer applications were more defaulted in males.
2. New clients with previously unused offer are more defaulted.

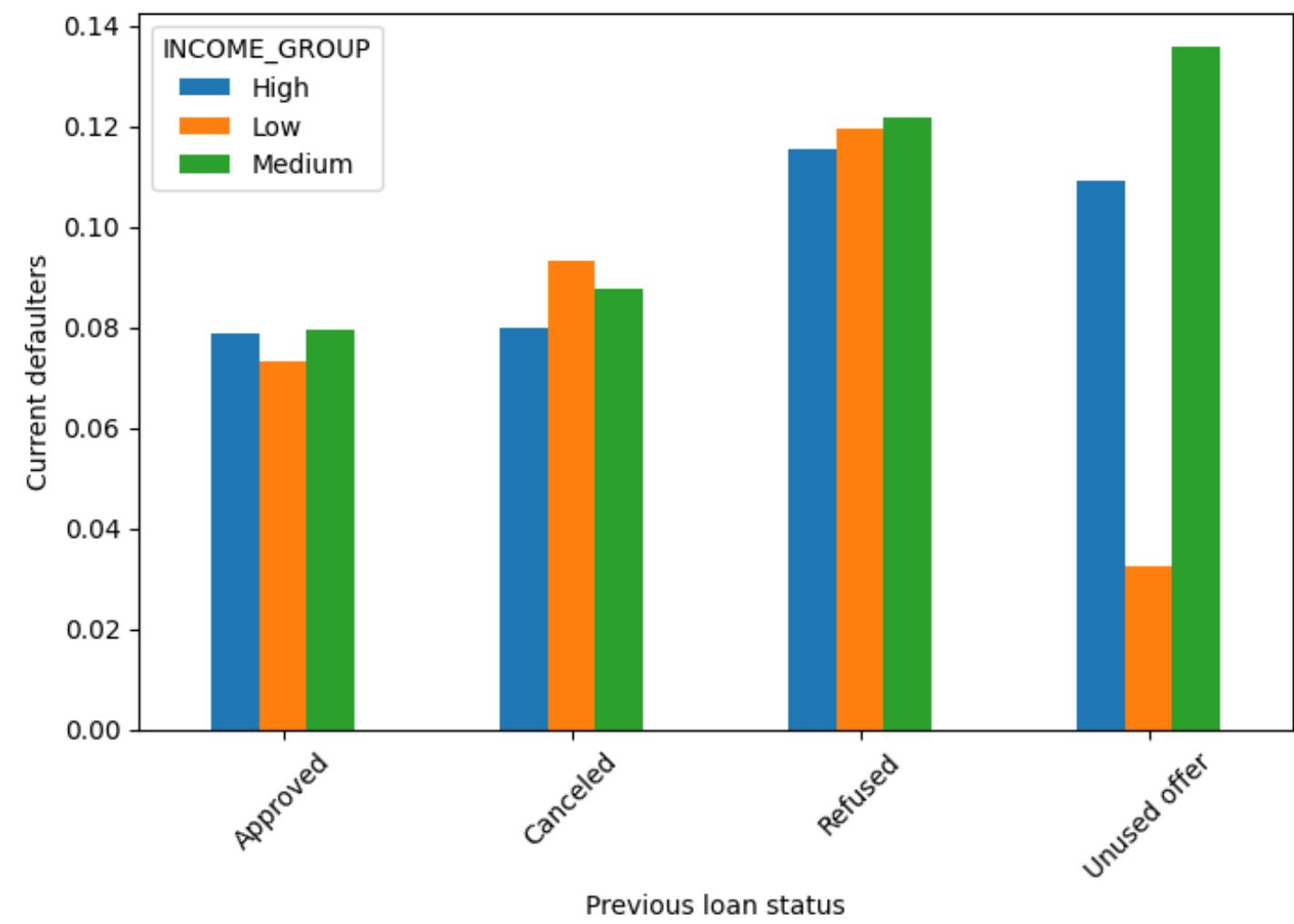
RECOMMENDATION:

1. It is recommended to provide loans to previously approved females.
2. There is a risk to grant loans for clients, whose applications were refused or unused previously.

AGE GROUP AND PREVIOUS LOAN STATUS



INCOME GROUP & PREVIOUS LOAN STATUS



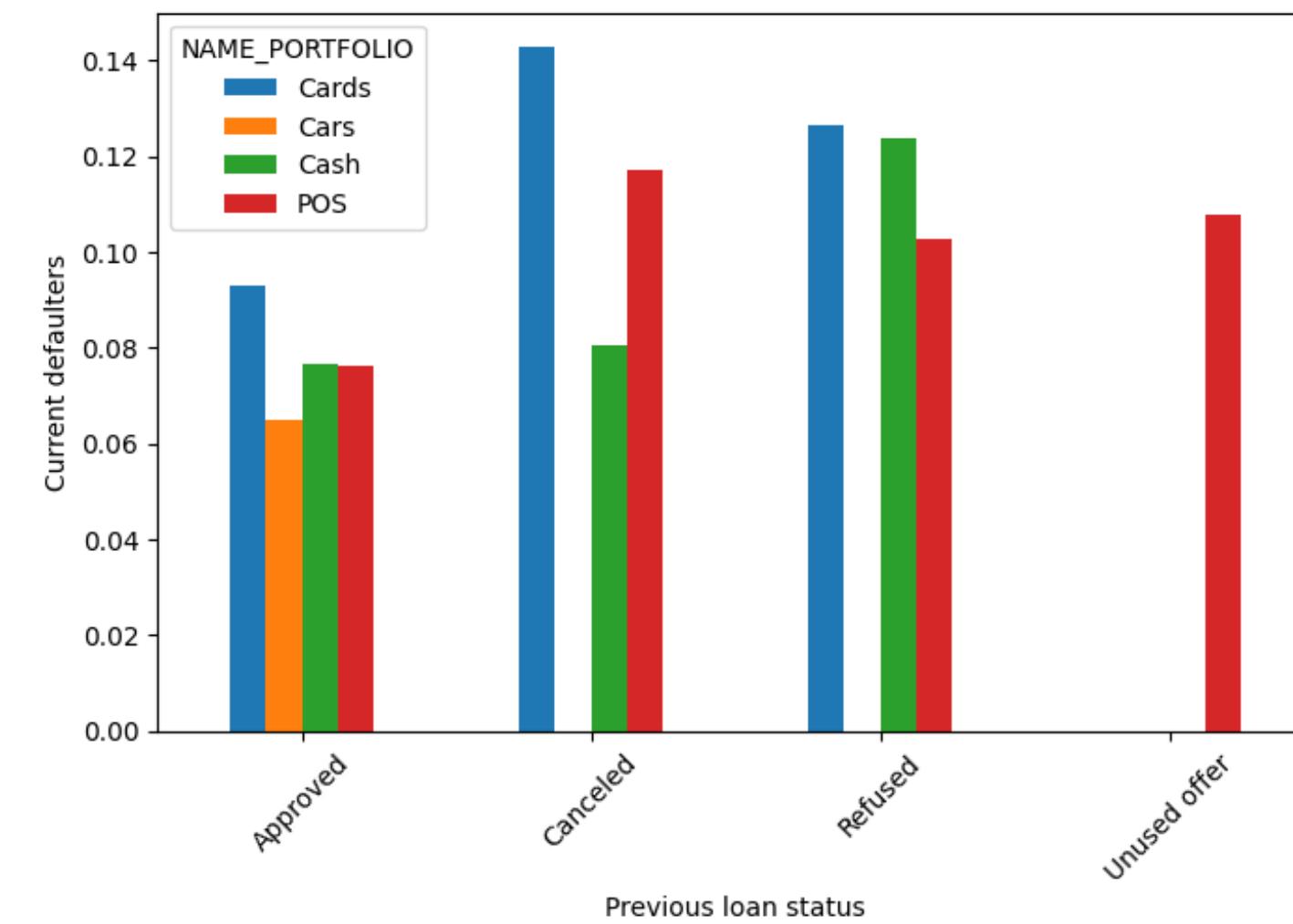
OBSERVATION:

1. Young people, who were previously refused are mostly defaulted.
2. The senior citizens are less defaulted irrespective of their previous loan status.
3. In all income groups previously refused applicants are more defaulted.

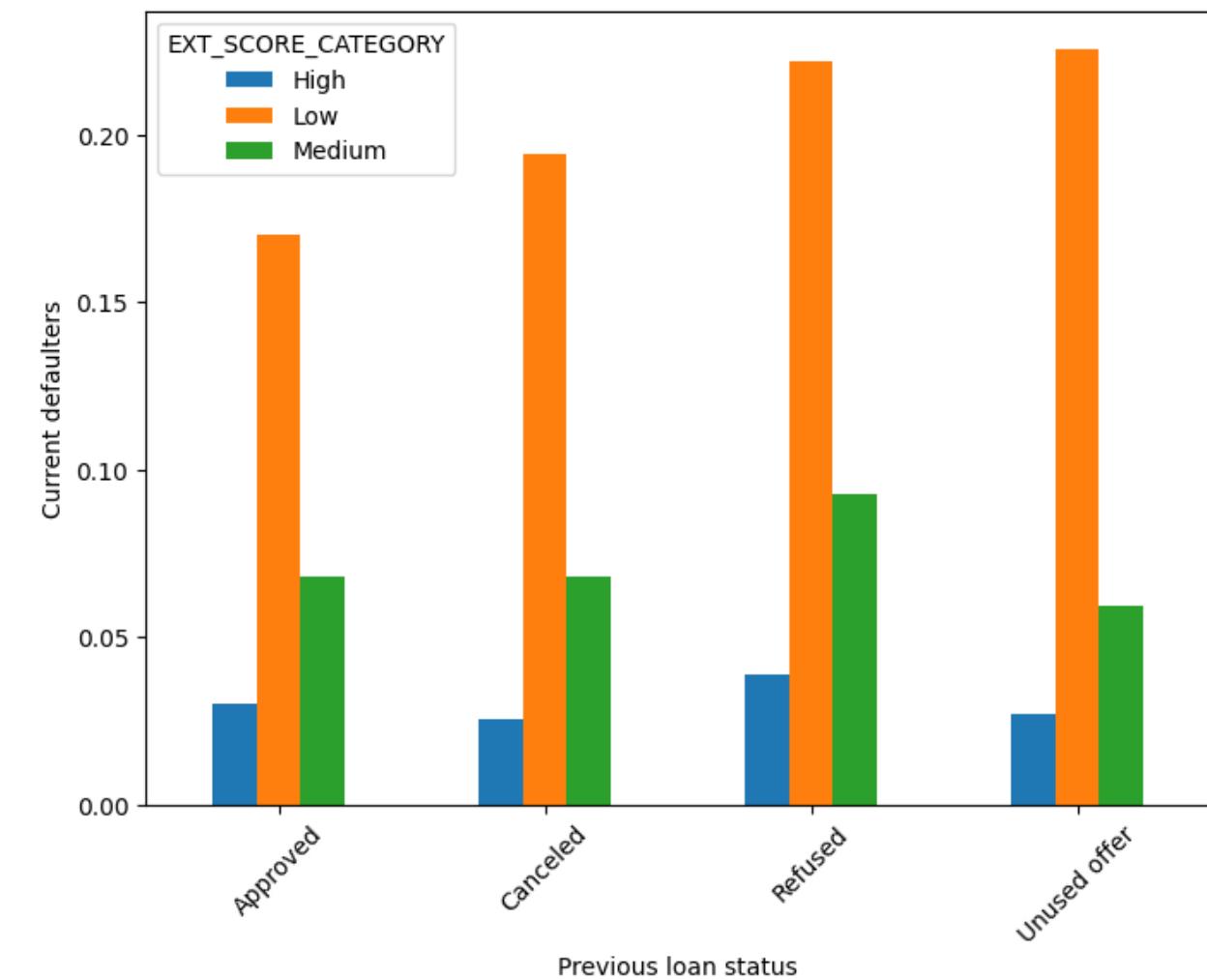
RECOMMENDATION:

1. Safer to grant loans to senior citizens.
2. Lesser risk granting loans for approved applicants to all income groups.

PORTFOLIO AND PREVIOUS LOAN STATUS



EXTERNAL SOURCE SCORE & PREVIOUS LOAN STATUS



OBSERVATION:

1. The previous applications for portfolio Cards and POS have mostly defaulted.
2. Previously refused applications for Cash are also defaulted at a higher rate.
3. Low external source scorers are highly defaulted irrespective of their previous loan status.

RECOMMENDATION:

1. It is safer to grant loans for any portfolio for previously approved applicants.
2. It is a high risk to grant loans to applicants, who have poor external source scores, especially those whose loans were previously refused, unused, or canceled.

RESULTS

HIGHLY RECOMMENDED GROUPS:

1. Approved clients in their previous applications
2. Highly educated clients with higher incomes.
3. Clients with higher external source scores.
4. Senior citizens in all categories.
5. Married clients compared to other family statuses.
6. Females are comparatively more favorable than males.

HIGH RISK GROUP

1. Previously refused, canceled or unused offer clients.
2. Low-income groups with previously refused status.
3. Unemployed clients.
4. Poor external source scorer.
5. Young clients are comparatively riskier than mid-age and senior citizens.
6. Lower secondary and secondary educated clients.

Link to github repository with the jupyterLab file: <https://github.com/dency-claris/Bank-Loan-Case-Study>

THANK YOU!

