



Bank **EDA & Visualization**

Bank Loan Applications
(Risk Analytics)



Table Of Contents

1. Problem Statement & Objective |
2. Analysis Approach |
3. Univariate Analysis |
4. Contract Type
5. Age Groups
6. Income Levels
7. Income Type
8. Education Type
9. Occupation Type
10. Last Info Changed
11. Bi & Multivariate Analysis |
12. Income vs Credit Request
13. Income vs Age
14. Income Type vs Housing Type
15. Income vs Last Phone Change
16. Income vs Region Rating
17. Contract Status



Problem Statement & Objective

Problem Statement

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 to their advantage by becoming a defaulter.

When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
2. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Objective

- Using EDA to analyse the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected. And reduces the risk of approving the loan to defaulters.

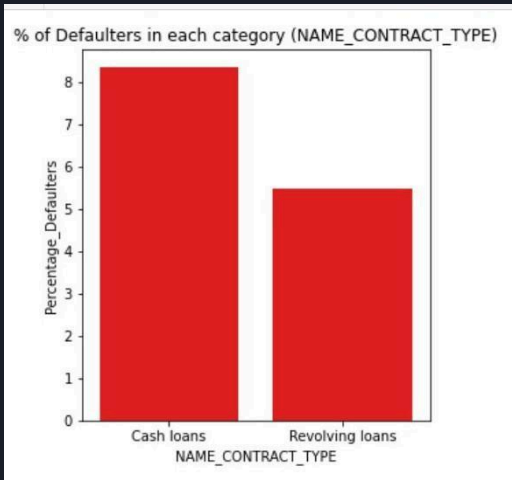


Analysis Approach

1. Domain Understanding
2. Business Problem Understanding
3. Loading Data into Python
4. Data Understanding
5. Handling Missing Values
6. Structuring Data to Right Format
7. Outliers Analysis
8. Data Imbalance Check
9. Data Analysis (Univariate, Multivariate, Correlation)

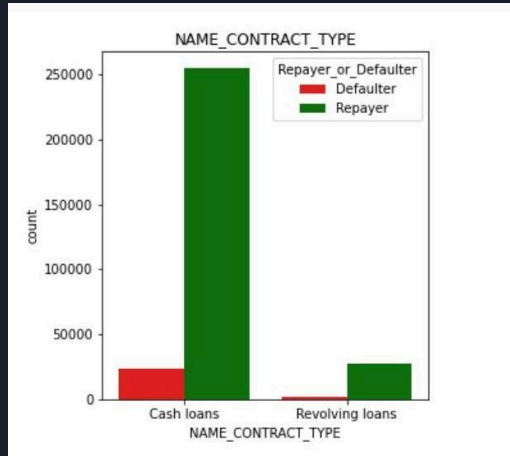


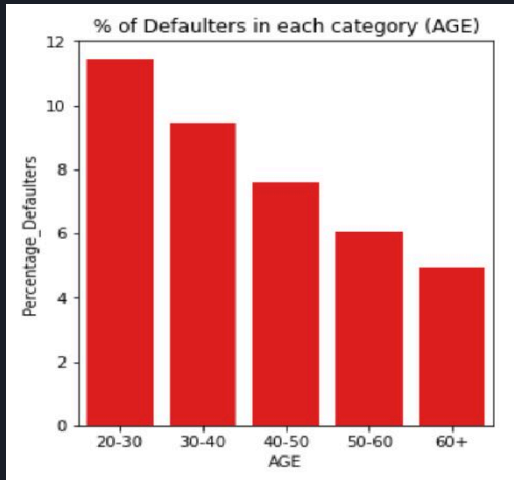
Univariate Analysis



Contract Type (Cash & Revolving Loans)

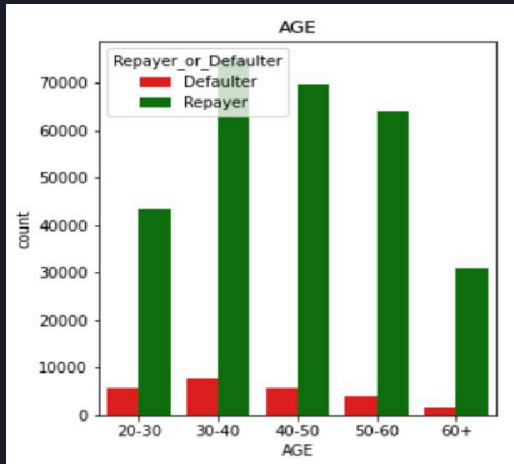
- **More than 8% of the clients under cash loans are defaulters, whereas $>5\%$ of the clients under revolving loans are defaulters**
- **Less risk is associated with revolving loans type**
- **Revolving loans type can be prioritised first**

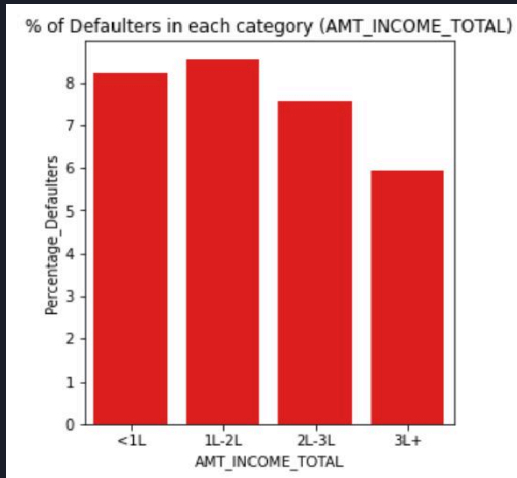




Age Groups (% Defaulters)

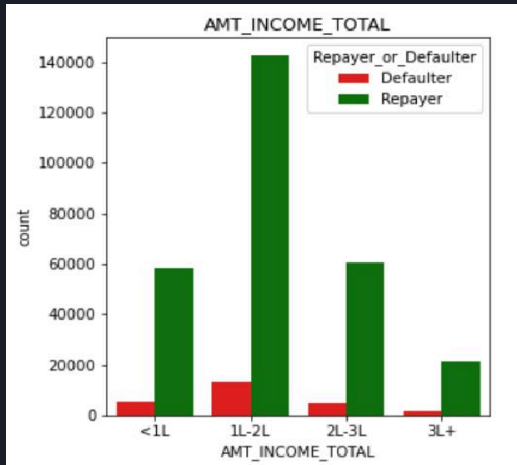
- **10% of the 20-40 age group clients are defaulters, 7% of the 40-60 age group clients are defaulters and 5% of 60+ age groups are defaulters**
- **More the age of the client, less is the defaulting rate**
- **Less risk is associated with older people**

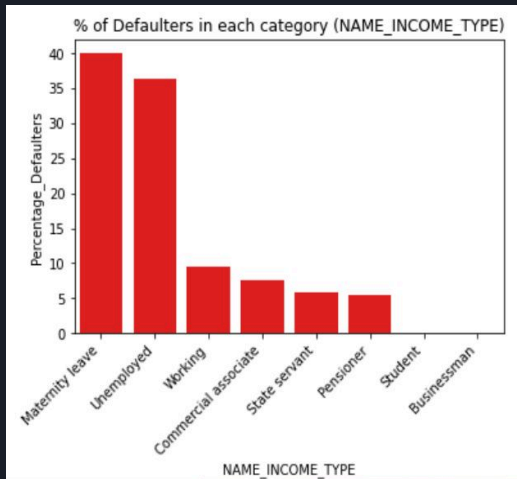




Income Levels

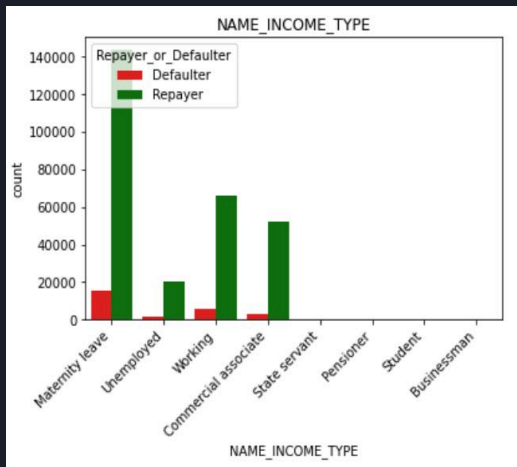
- Clients having total income more than 3 Lakhs have less defaulting rate (less risk associated)
- Clients having total income less than 2 Lakhs have more defaulting rate
- More the income less is the defaulting rate of the clients

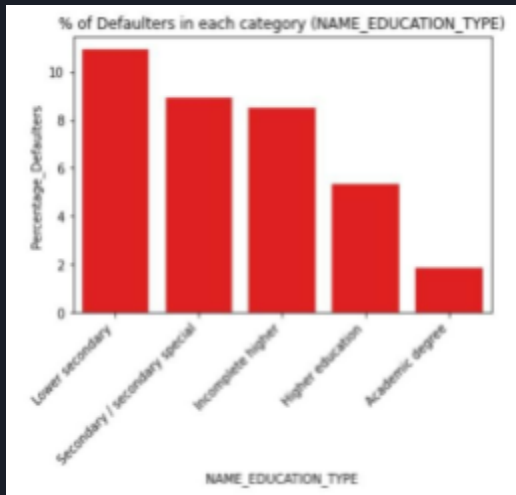




Income Type

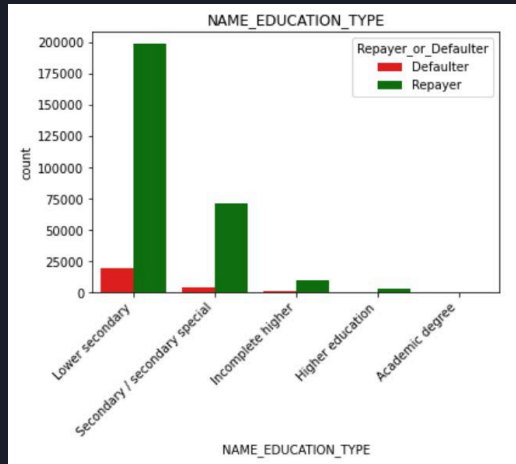
- **Very high % of defaulters in 'Maternity Leave' & 'Unemployed' category. But the counts are very less, so can be ignored.**
- **Working professionals and commercial associates have a higher default rate**
- **Pensioners and State Servants have a less default rate comparatively (Less risk with them)**

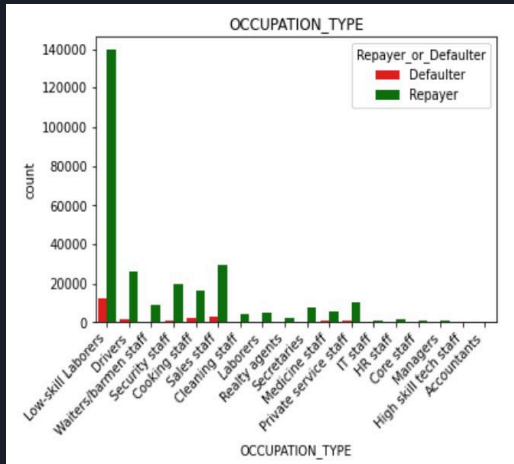
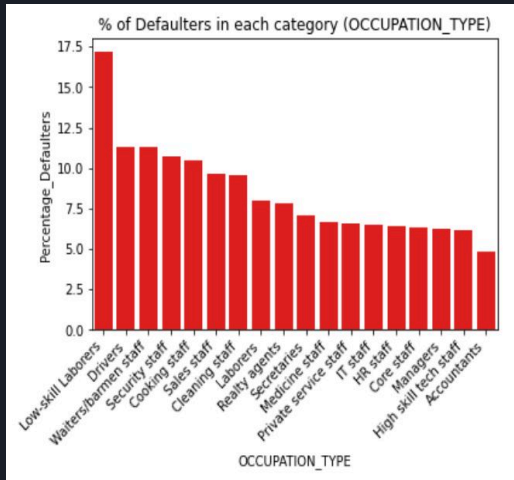




Education Type

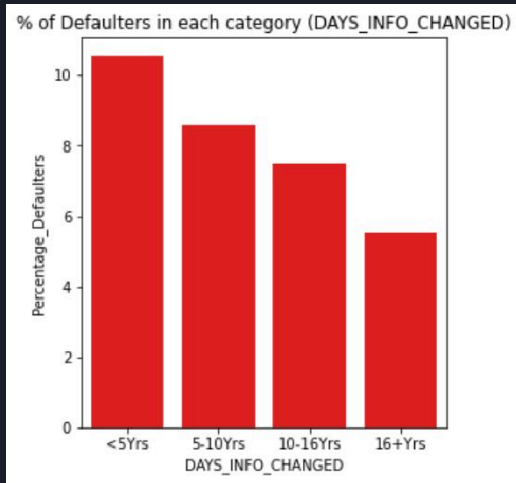
- Clients have higher default rate who have lower secondary, secondary special or Incomplete higher education type
- Clients with Academic degree or Higher education tend to default less
- Less risk associated is with Academic degree and Higher education holders





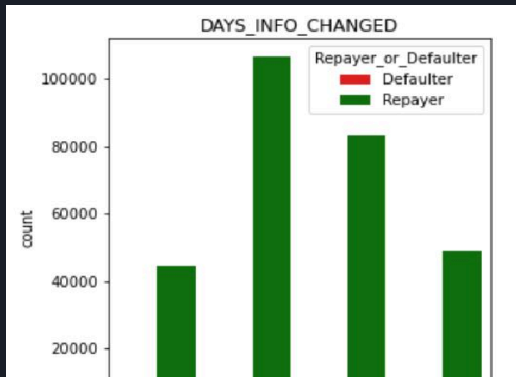
Occupation Type

- **Sales staffs, Drivers, Security staffs, Cooking staffs, low-skill Laborers (Very high), Waiters/barmen staffs have high defaulting rate**
- **Core staffs, Managers, High skill tech staffs, Accountants, Medicines, Private service, Seceretaries, HR & IT staffs have less defaulting rate (Less risk associated)**
- **Others are in the acceptable range (7.5-9)**



Last Info Changed (Reg. & ID)

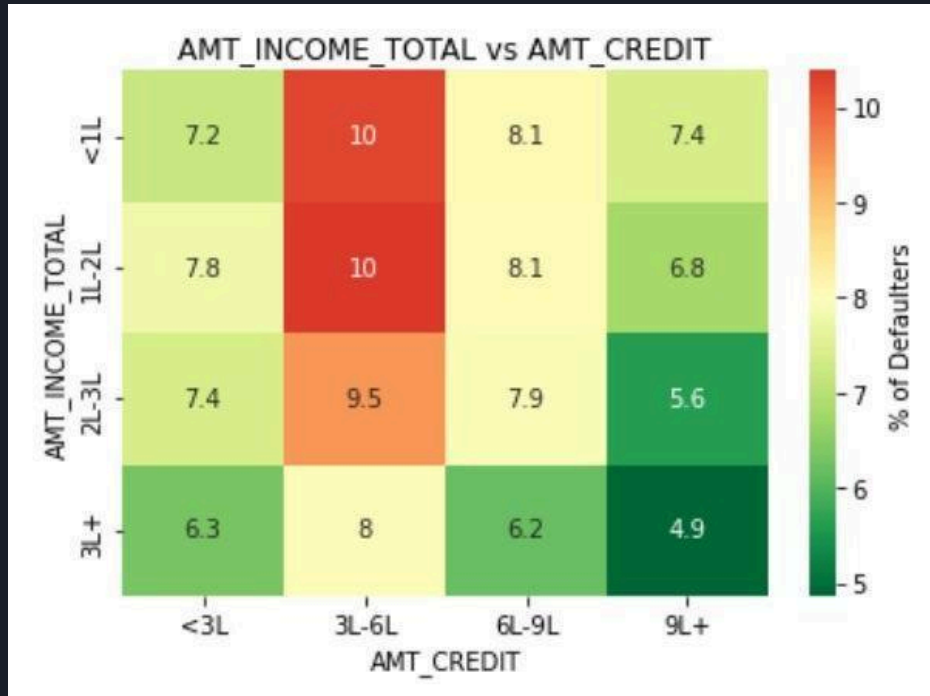
- Clients who have changed their information <5 years have more defaulting rate
- Clients who have changed their information >5 years have less defaulting rate
- Clients who have changed their information >16 years age have very less defaulting rate (Less risk associated)





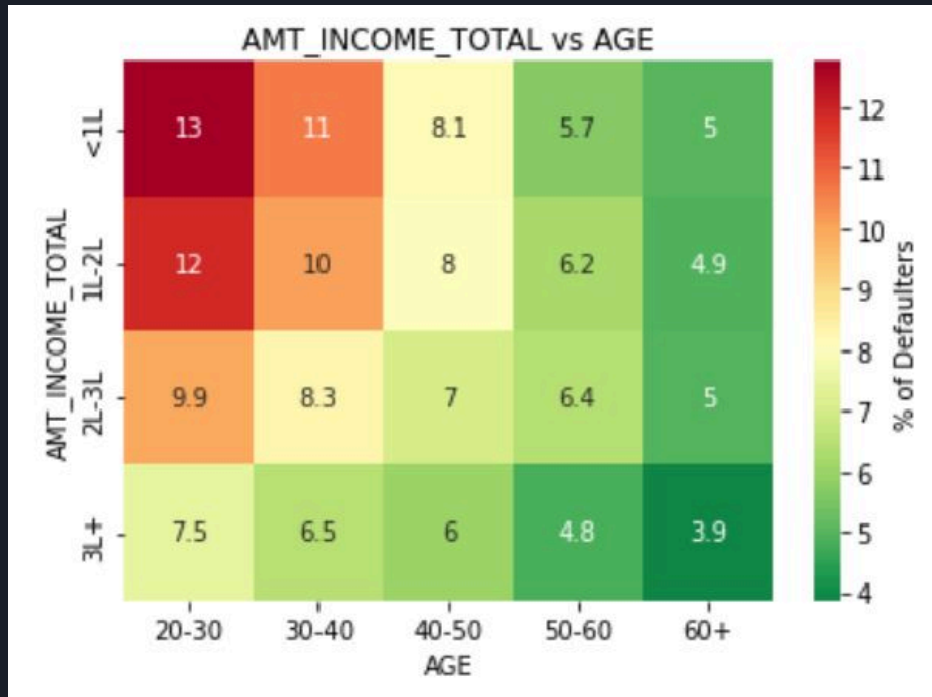
Bi & Multivariate Analysis

Income vs Credit Request



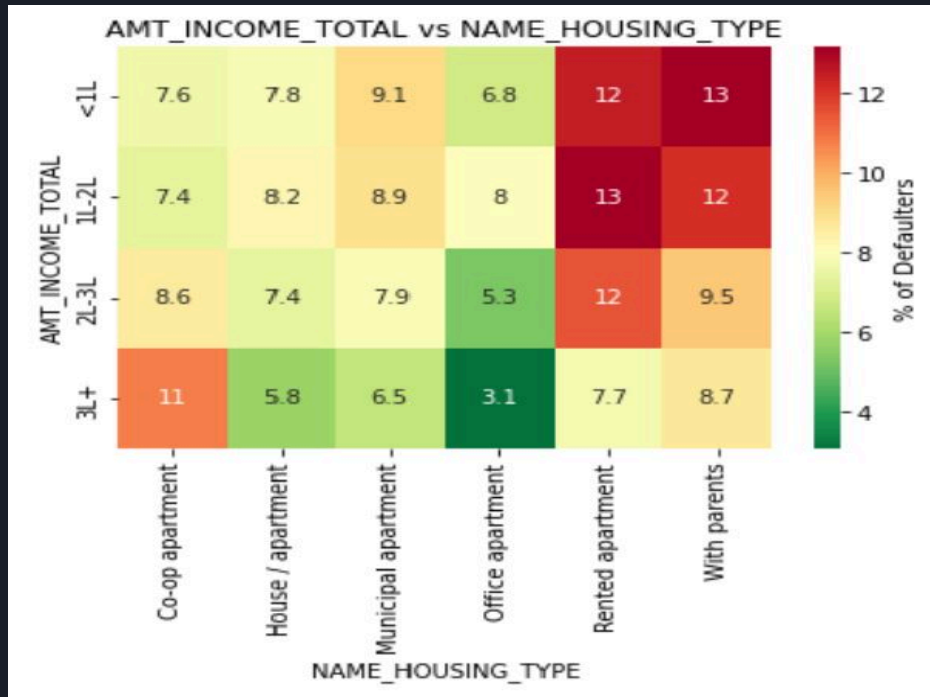
- **Clients with less income and credit amount requested higher than their income have high defaulting rate (10%)**
- **Clients with more salary and equivalent credit amount requested have less defaulting rate (5-6%) (Less risk)**

Income vs Age



- Young clients whose age is 20-40 with income level <2-3 Lakh have high defaulting rates
- Clients whose age is more than 50 with all income levels have less defaulting rate (Less risk)
- More the age and more the income, lesser is the defaulting rate

Income vs Housing Type



- Clients with income <2 Lakhs who have either rented apartment or with their parents have very high defaulting rate (10-13%)
- Other clients who have owned apartment / office apartments have average defaulting rate (6-8%)
- Clients with office apartment or owned house/apartment with high income have very less defaulting rate (3-6%)

Income vs Last Phone Change

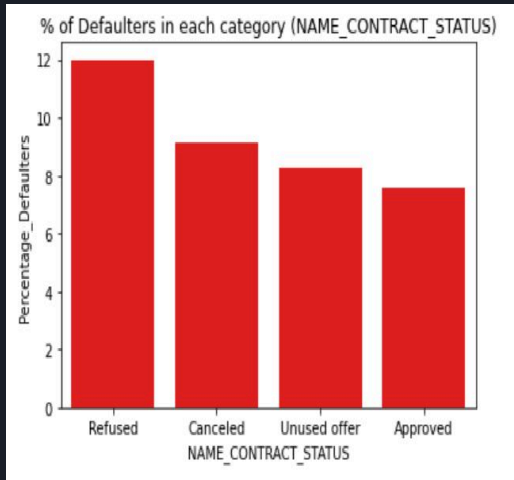


- Clients with income <2-3 lakhs & who have changed their phone number in <2-3 years have high defaulting rates
- Clients holding their phone number for longer duration (>5-6 years) have less defaulting rate (Less risk associated)

Income vs Region Rating

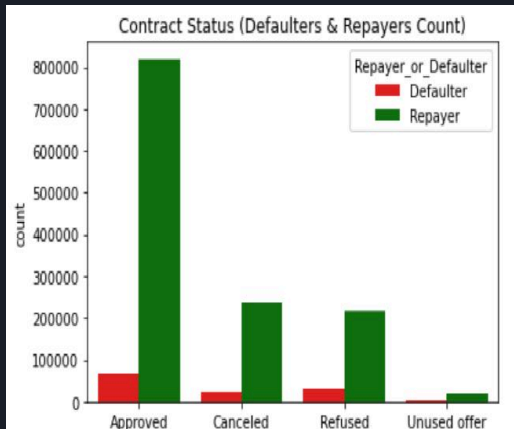


- Clients whose region rating is <1.5 with all levels of income have less defaulting rate
- Clients whose region rating is between 2-2.5 have acceptable defaulting rate
- Clients from region rated with 3 (of all income levels) have high defaulting rates



Contract Status - Previous (% of Defaulters)

- Clients whose application had been rejected before have very high defaulting rate (~12%)
- Clients who cancelled their application previously have moderate defaulting rate of ~9%
- Clients whose application had been approved before have less defaulting rate ~7.5%
- Seems like bank has rejected the right people before who have high defaulting rate





Thank you!