EDA Assignment-Loan Approvals

This assignment is to produce insights on data that's collected on its applicants, for a bank to increase its business in the loans department

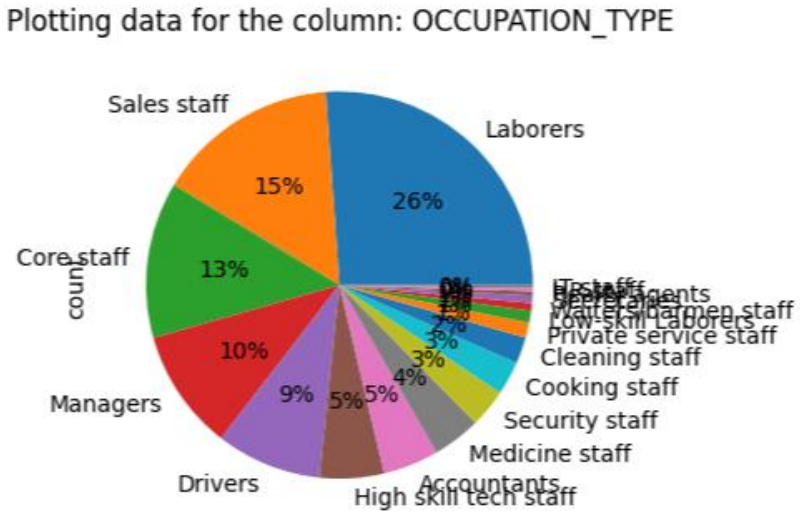
Introduction

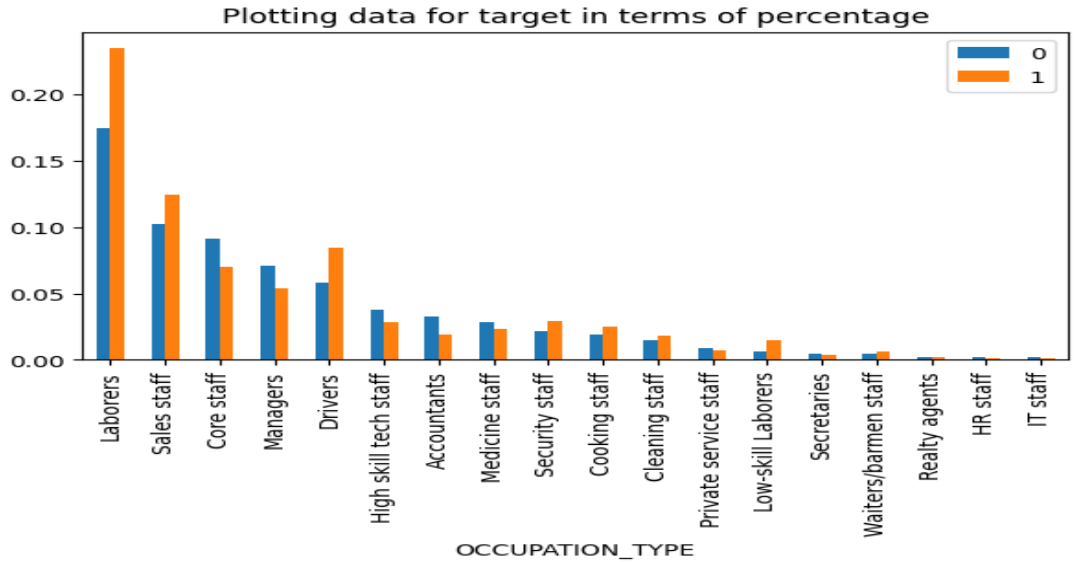
- We have two data sets with current and previous loan applications.
- Applicants with insufficient to no credit history are tough to gauge for a loan approval.
- The business gains if loans are approved/denied to the right applicants and loses if not.
- Here in these data sets we have data about the client, we need to find a relation between these traits and the how likely is the applicant able to pay their loans.

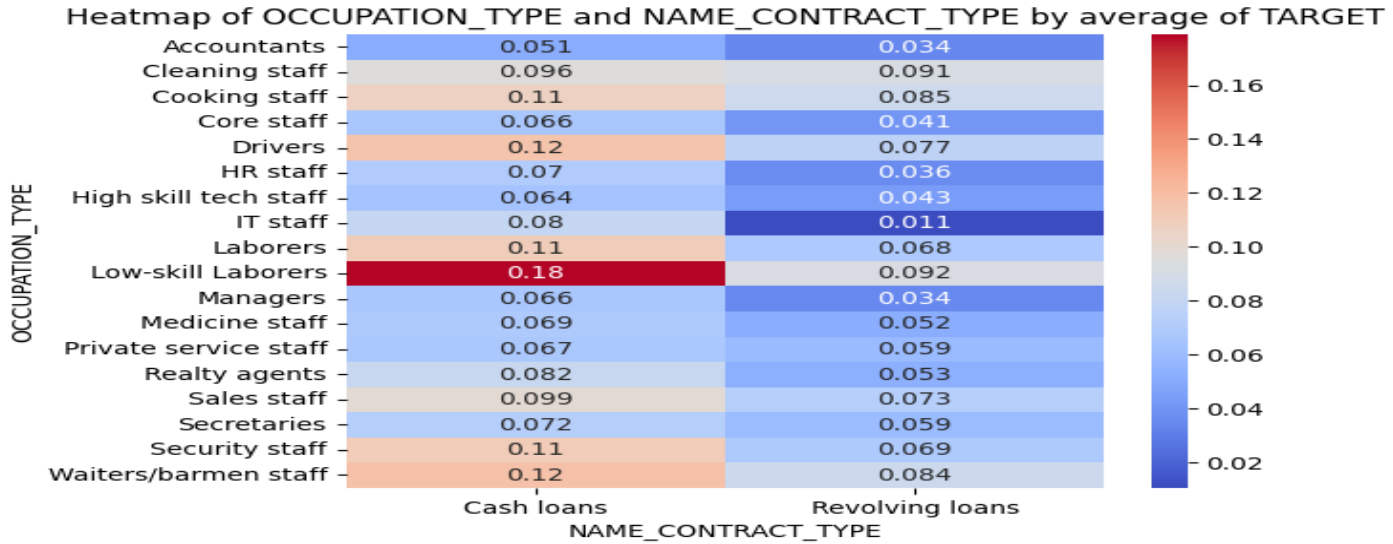
Assumptions

- People with consistent flow of money are more likely to pay off their loans.
- People with more assets are a safer choice for the bank to approve secured loans.
- People with more liabilities are less likely to pay their loans.
- People with consistent flow of money are perfect candidates for unsecured loans, as the bank can predict a trend of their applicant’s income.
- Applicants are likely to take loans proportionate to their incomes.
- People with an asset are more likely to pay their loan, especially cash loans, as they wouldn’t want to lose it.

Occupation type vs Target

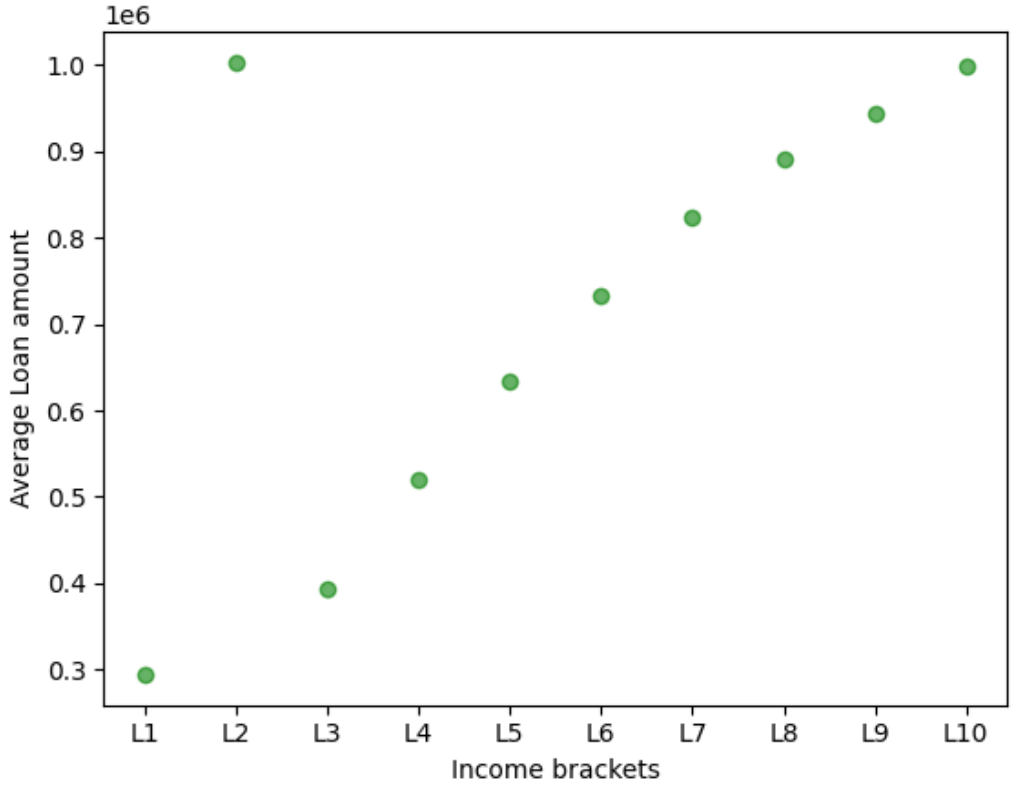






- We see that laborers, sales staff and drivers are less likely to pay off a loan.
- This holds true with the assumption that a consistent supply of income is more likely to pay off a loan.
- As the jobs that we have noted are usually contractual, seasonal or have variable pays this could be a reason for more defaults of loans people working in these sectors.
- The chart shows us that most of the clients of the bank are labourers, this is not a good sign as seen from the trends mentioned above.
- We can also see jobs that have a consistent cash flow are more likely to pay their debts, like managers, core staff, highly skilled tech staff.
- We don’t have enough data on bar staff, low skill labourers, realty agents, secretaries, IT staff and HR staff to conclude anything from them.
- Although from the heat map, it contradicts our assumption that a revolving loan would be better for a consistent income sector, there a lot of difference.
- But IT staff is least likely to default on the loan and low skilled labourers are most likely to default on one.

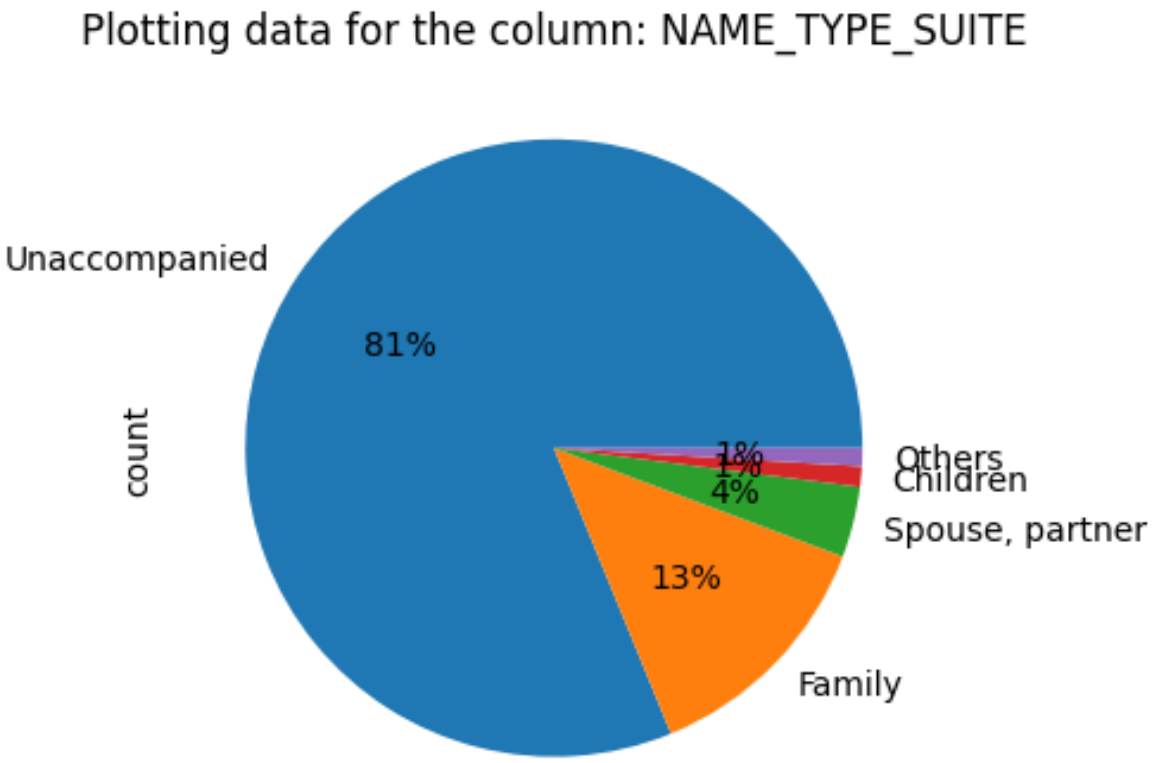
Loan requirements

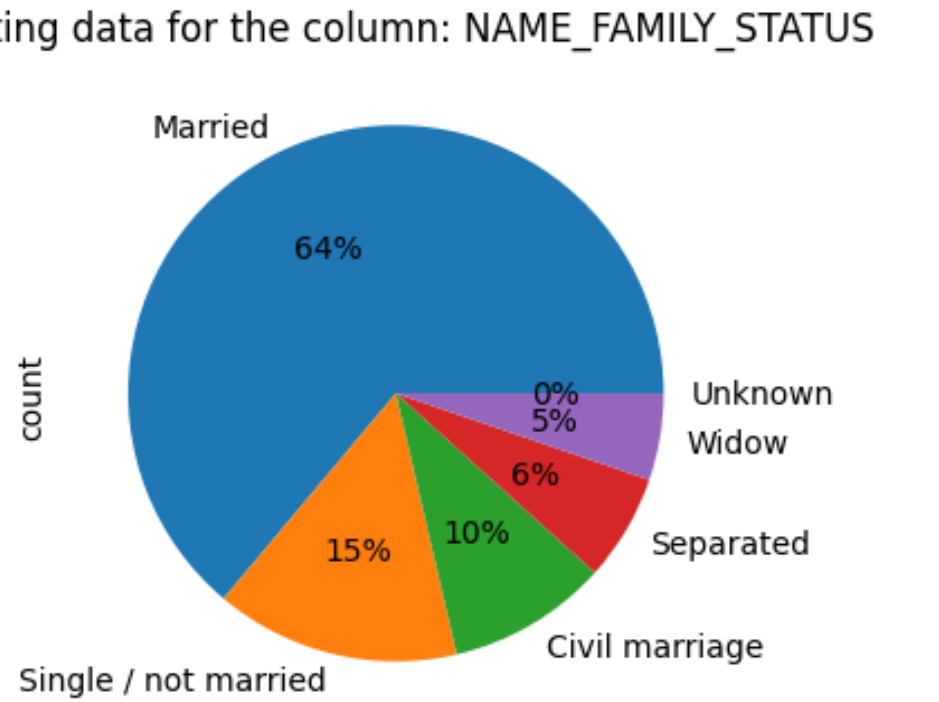


Category	Income range
L1	<50000
L2	50001-100000
L3	100001-150000
L4	150001-200000
L5	200001-250000
L6	250001-300000
L7	30001-350000
L8	350001-400000
L9	40001-450000
L10	450001-500000

- We can see that if we plot the average loan requirements against the income brackets, they increase linearly.
- This makes in sense in both the banks and the clients perspective.
- This way the bank will be in a more comfortable position to provide the client with a loan.
- Although, there is an anomaly in the L2 category of the income.
- The client can pay off these loans according to their salary and also, obviously won’t take loans for smaller amounts, since they already have that amount because of a high income scale.
- This also agrees with our assumptions that people take loans proportionate to their income.

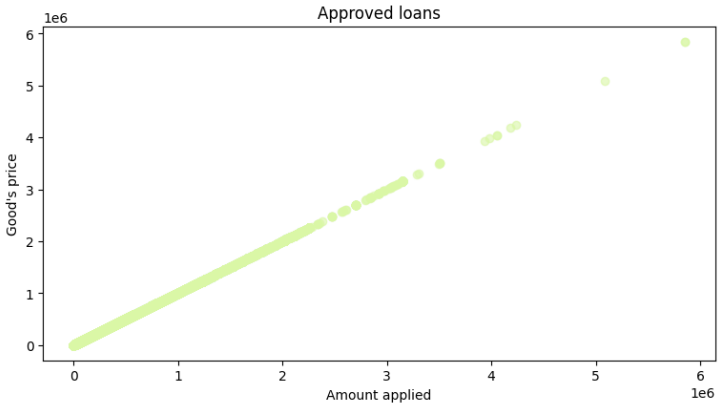
Applicant preference for tagging along

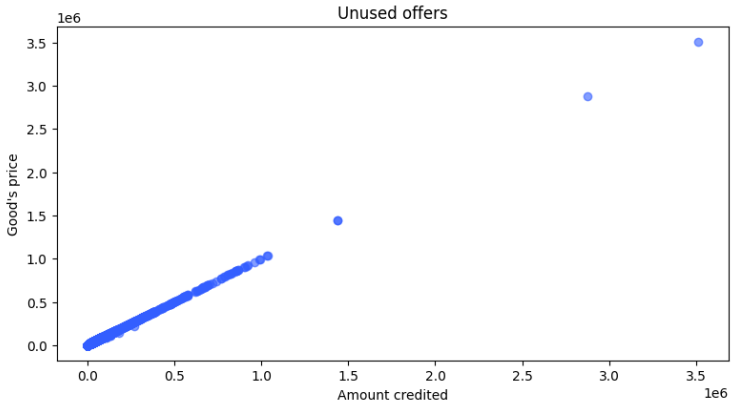


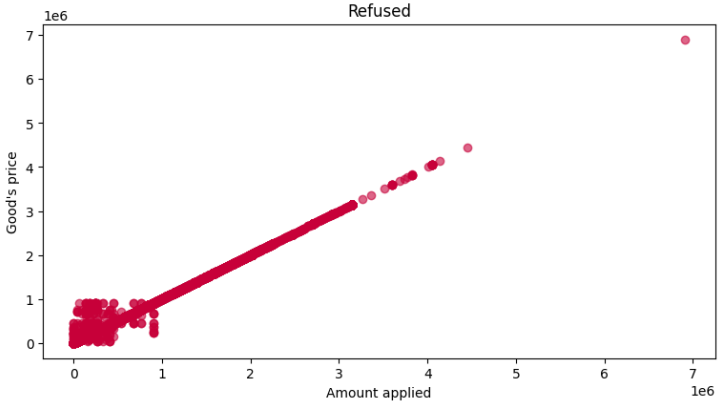


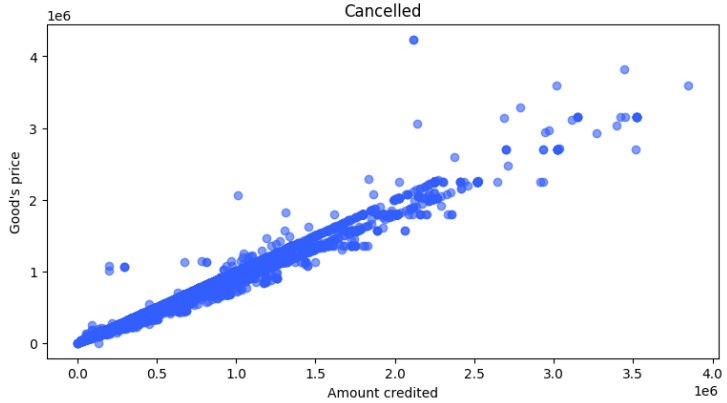
- NAME_TYPE_SUITE has very less nan, we can impute them, because the count is high enough to not want to drop the row, but low enough to not have to worry about any changes made by imputing it.
- We'll use mode to impute it since this is a categorical column. Most people were unaccompanied, so, we'll replace it with that.
- We see that 3 categories, other_B, other_A, group of people are very less to have any affect on the analysis, so we'll make them into a new category called 'Others'
- Most people seem to come unaccompanied, this might be because most people might not want to disclose their financial affairs with anyone else.
- This is contrary to a usual expectation that most people would be accompanied by a spouse. As we have seen that most people are married.
- We can't draw any conclusions about how this affects the credibility of the client as more than 80% of our data consists on unaccompanied people.

Matching interests



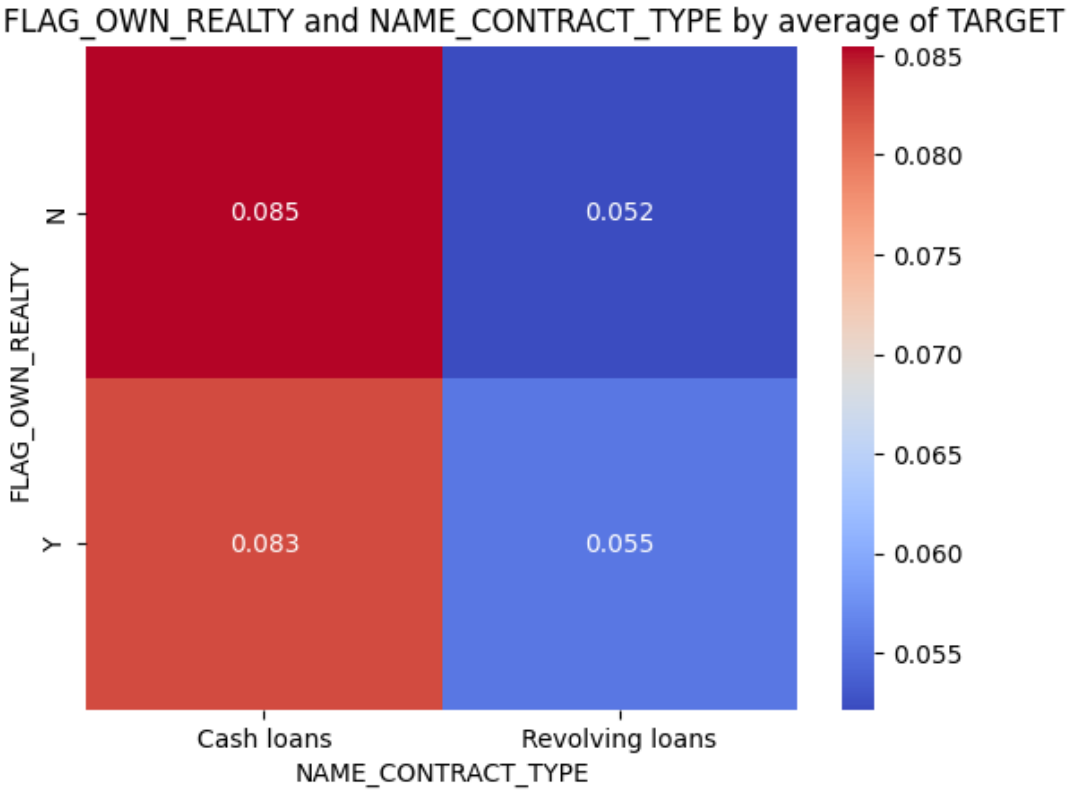






- This observation tells us that the loan is most likely to be approved if there is a match of price of goods and the applied loan.
- If there is a mismatch from any of the parties the deal may not work out.
- Of course, we can see that most of the points in a deal breaker is still linear, but it is a very clean graph when the loans are approved.
- There are other factors that can be weighed in to offer a loan to the client, as shown in the previous slides.

Assets and credibility



- This contradicts our assumptions.
- From the heat map people with a land are less likely to default on a revolving loan than on a cash loan.
- Also people without lands are most likely to default on cash loans.
- But people without lands are also least like to default on revolving loans.
- Although this might not have anything to do with if they do have a land or not and more about how much they earn and how much they spend.
- People without lands are usually younger in age, this is an assumption, and a revolving loan is more favourable to their spending habits.
- We can see a glimpse of this assumption here:
  - Age of most of the people with land: 39.43835616438356
  - Age of most of the people without land: 28.405479452054795

Conclusion

- Most of the people prefer applying for the loan alone.
- Most of the applicants are married.
- We are on a safer side if we approve the loans for people with more consistent cash flow, like managers, core staff, etc.
- Most of our applicants are laborers and sales staff, this is bad for business as per our previous observations.
- Applicants are more likely to apply for loans proportionate to their incomes, this make it easier on the bank as well.
- The loan is most likely to be approved if the goods and the amount applied for are a match, straying away from this reduces the chances for approvals.
- Revolving loans are better suited for people lesser in age.

Future scope

- As seen that people like to apply for loans alone, it means we can increase customer satisfaction rating if we were to provide them with private cabins for discussions.
- We can make offer that are more appealing to a married couple, like good offers on jewelleries, home loans etc.
- Try to create more variation in occupations of the customer, since people with consistent cash flow are good way to make profits with unsecured loans as well, they often have high interest rates and short tenures, this is quick money for the bank.
- Offer better deals on things like credit cards and paperless loans for the younger applicants, usually under 30, as these are revolving loans and produce high profits for the bank.

THANK YOU