

# EDA CASE STUDY

---

By MOHIT RAJ



# PROBLEM STATEMENT – UNDERSTAND DRIVING FACTORS BEHIND LOAN DEFAULT

---



# AIM – ENSURE COMPANY LOSES NEITHER MONEY NOR BUSINESS

---

- Can't deny loan to all as that is major source of income
- Can't approve all applications, customers may default (result in loss)
- Can lend at higher rates
- Can lend lesser amount

# ANALYSIS

---

- Ever wondered why your manager rejects loan application of your customers and not your colleague's customer?
- Ever wondered why you get loans easily while your friend doesn't?
- Or why your EMI amount is less than your friend's for the same amount and tenure?

We got you covered. By the end of this presentation, you will know the answers

# APPROACH

---

- Divide customers into two segments and perform analyses:
  1. Those who defaulted
  2. Those who didn't default
- Divide customers into 4 categories for application success rate:
  1. Approved
  2. Refused
  3. Canceled
  4. Unused



# APPROACH

---

- First division helps us in knowing who are less likely to default
- For more profit, we can focus more on them and avoid customers who can default
- Second division helps us target people who have high chance of loan approval
- Helps us save manforce and by avoiding those who have less chance of approval or those who cancel or don't use loan

# LET'S BEGIN ANALYSIS

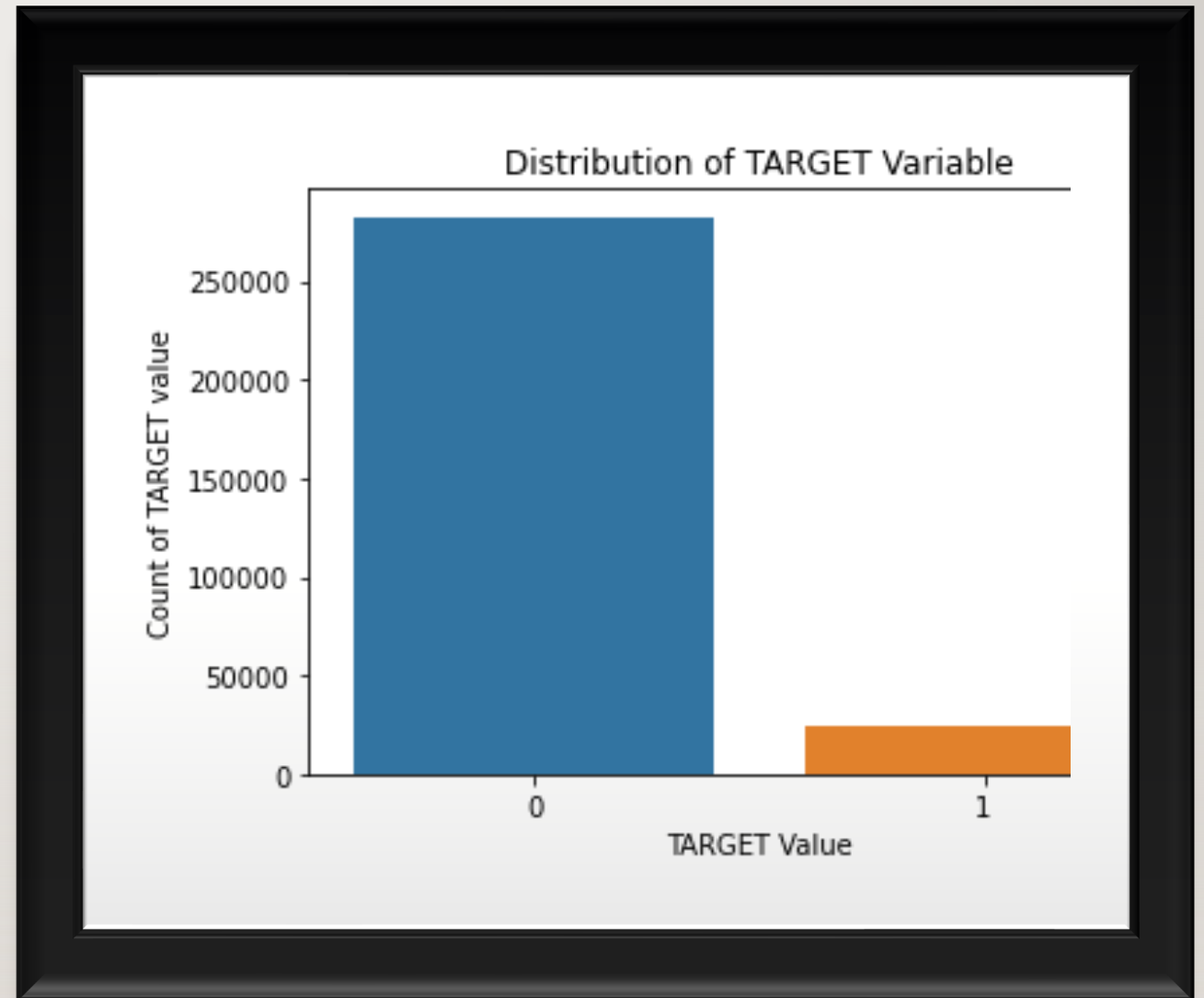
---



# DATA IMBALANCE

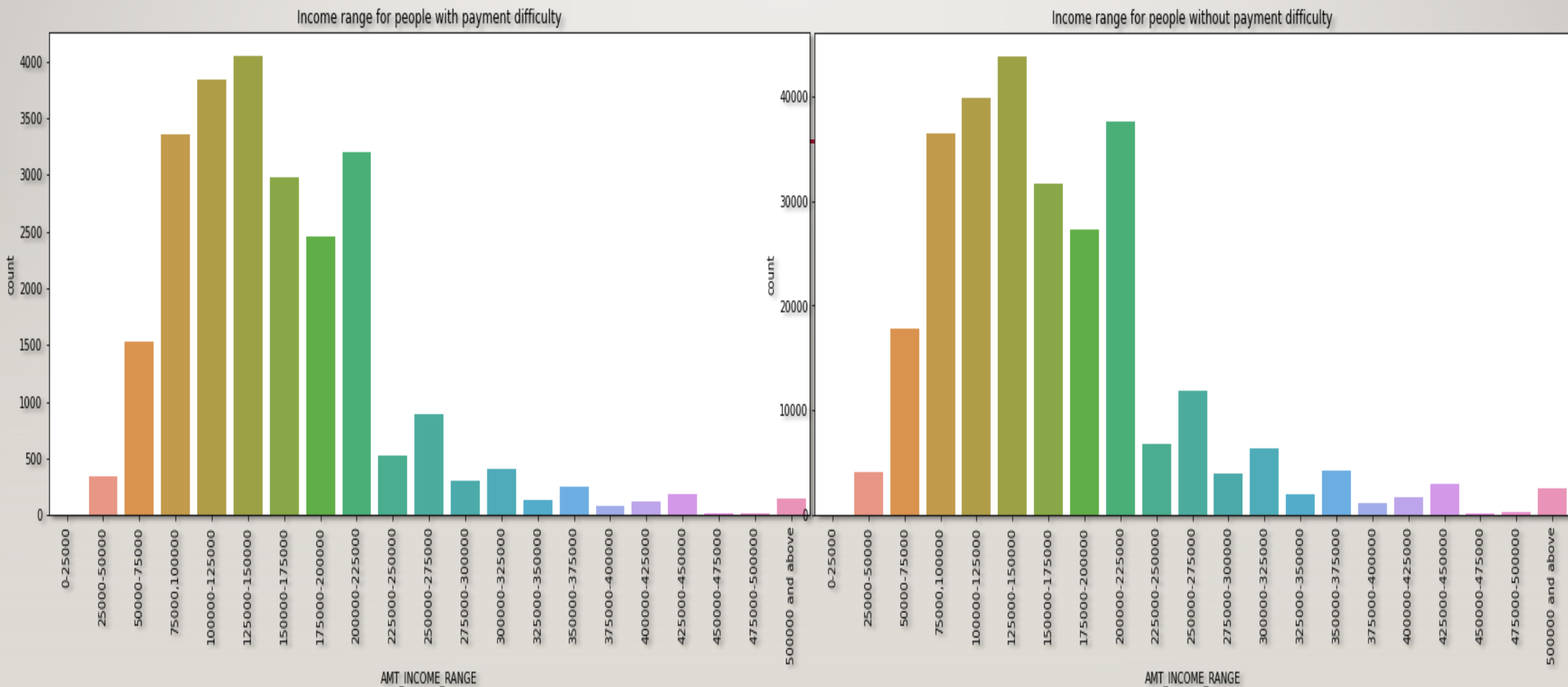
---

- 0 shows people with on-time payment
- 1 shows people who defaulted
- Ratio Data Imbalance = 11.39

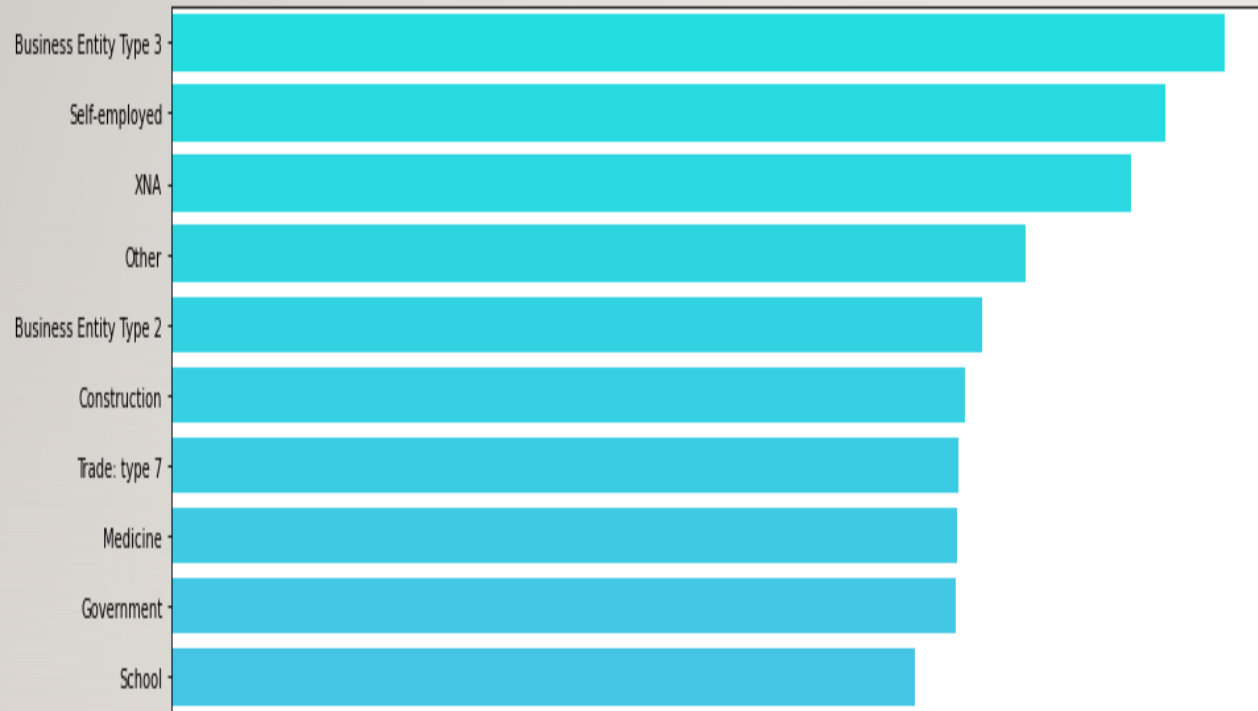




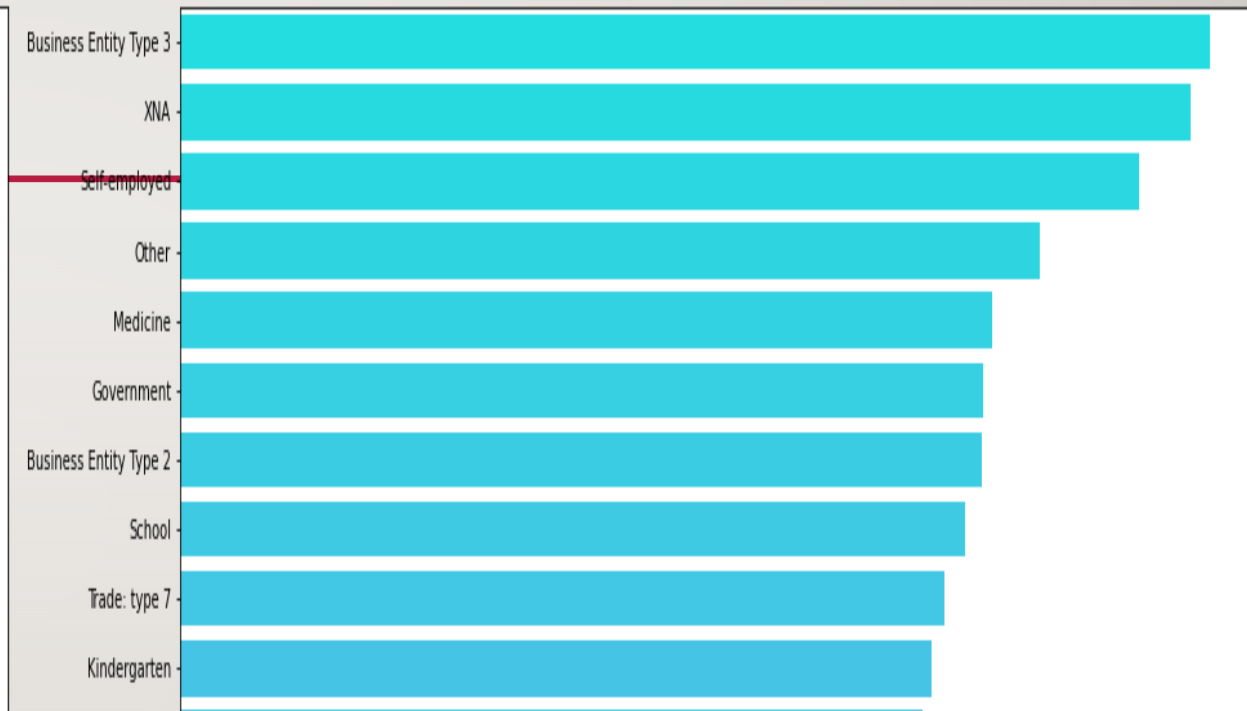
Not much difference in terms of income range



Organization type for people with payment difficulty

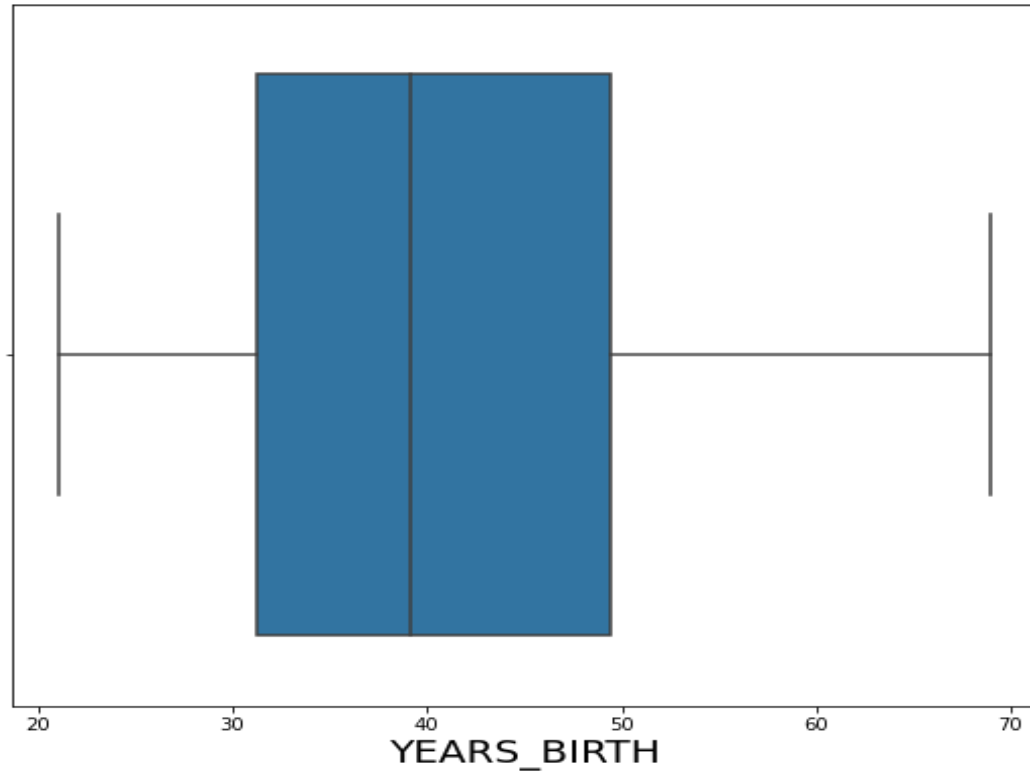


Organization type for people without payment difficulties

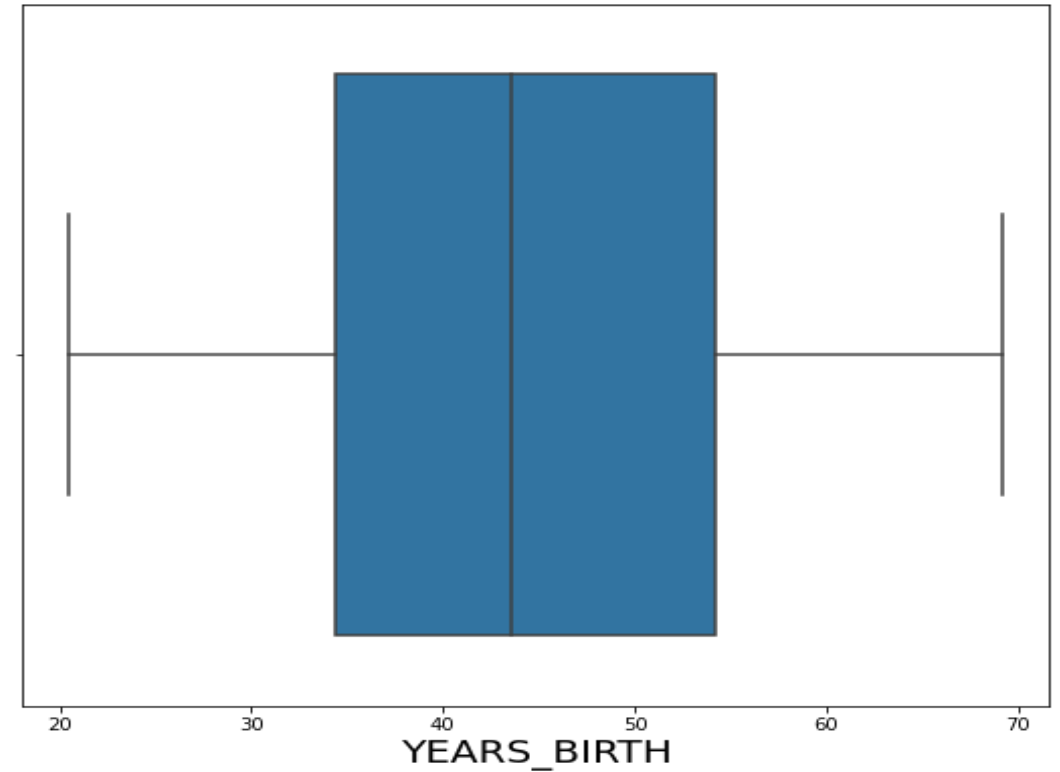


**PEOPLE WORKING IN BUSINESS ENTITY TYPE 2 AND  
CONSTRUCTION HAVE PROBLEMS REPAYING**

Age of people with payment difficulties

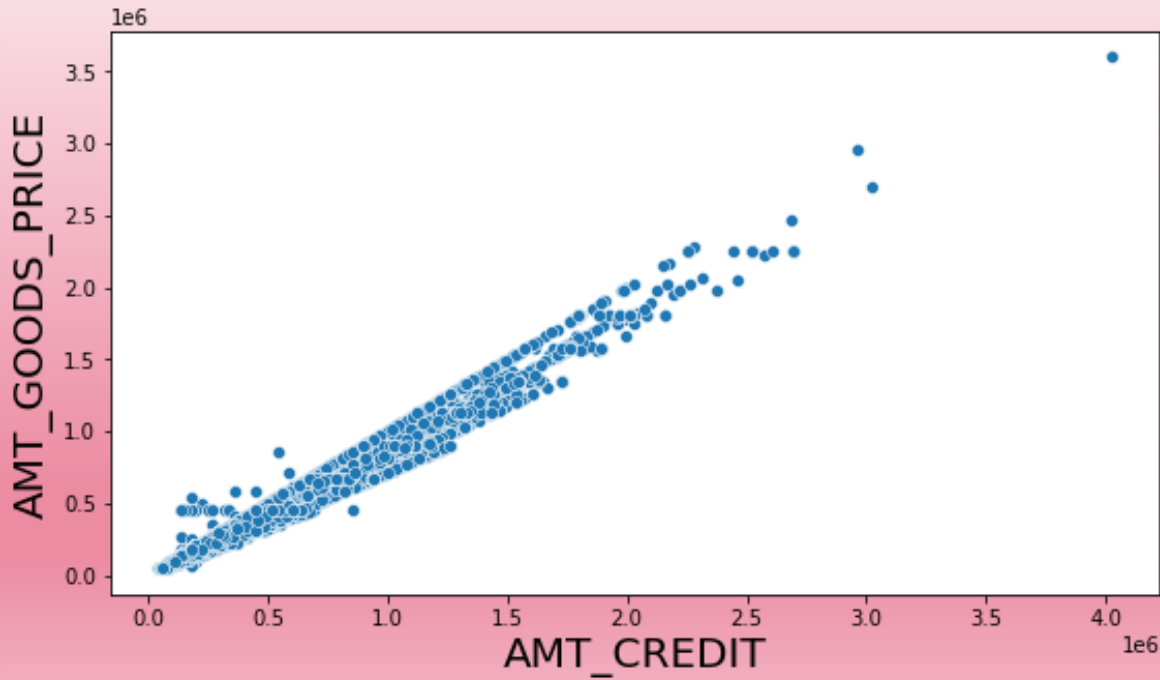


Age of people without payment difficulties

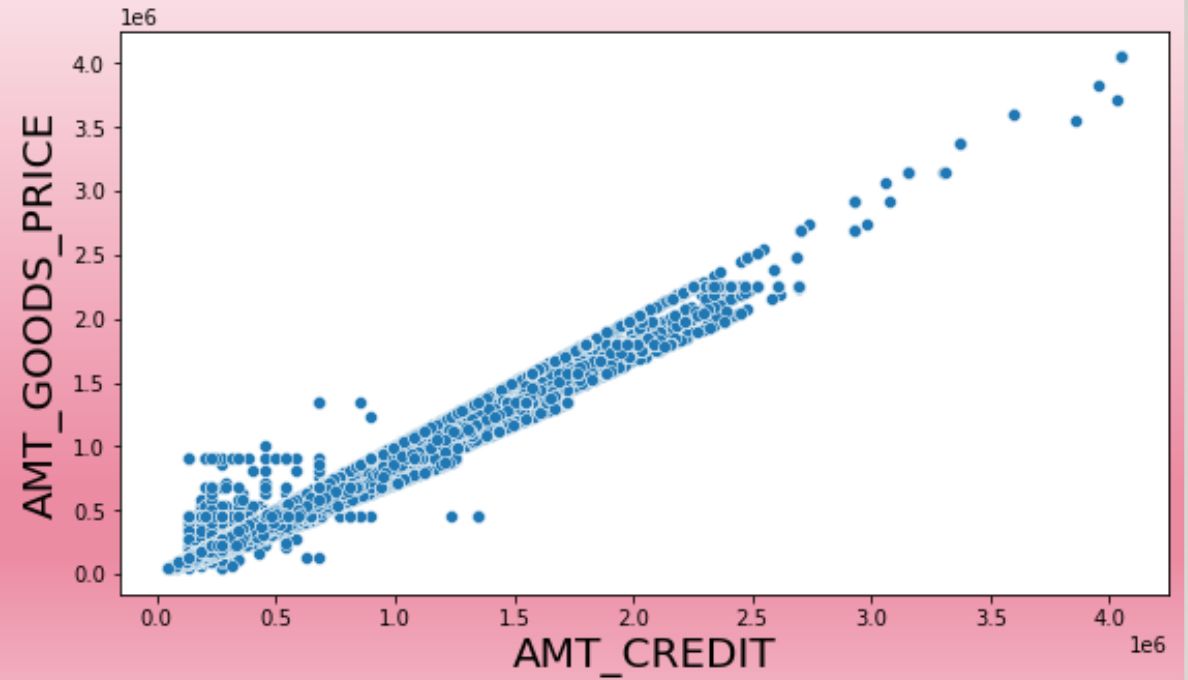


People defaulting aged 31-49, not defaulting in 35-55

People With Payment Difficulties  
Previous amount applied vs Amt credited



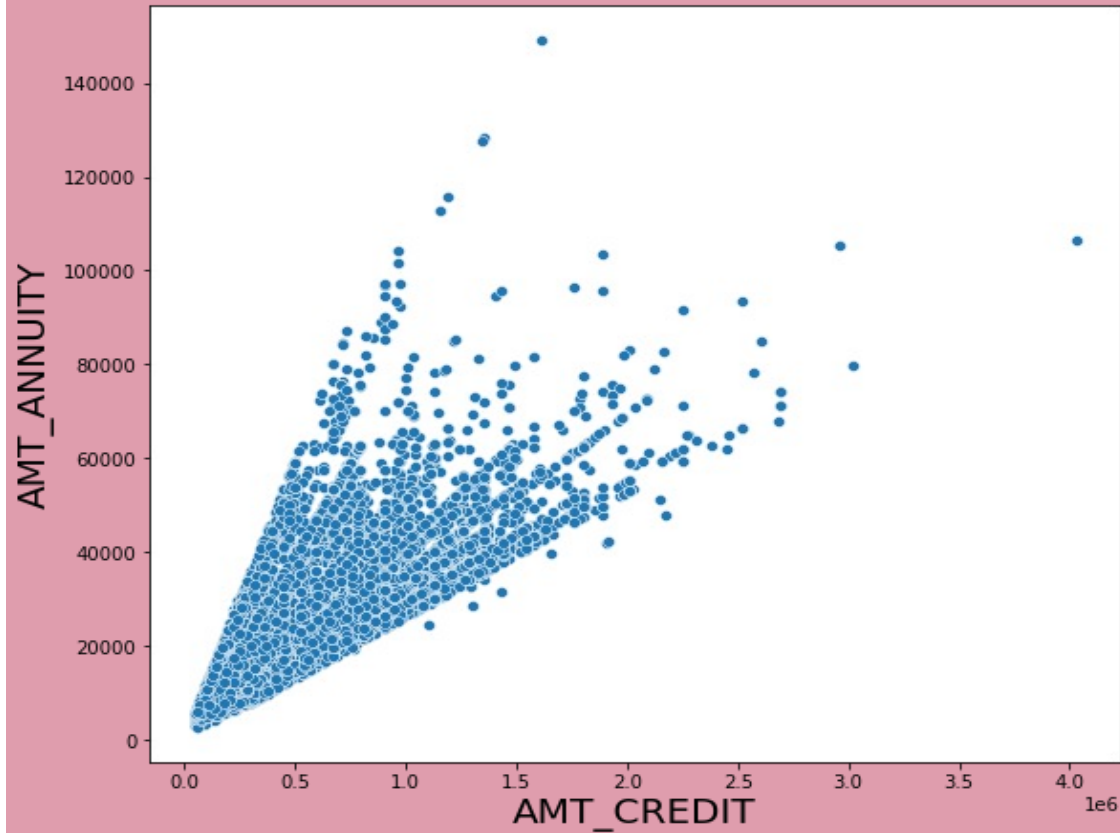
People Without Payment Difficulties  
Previous amount applied vs Amt credited



Goods price positively correlated to  
credit amount, but in both cases

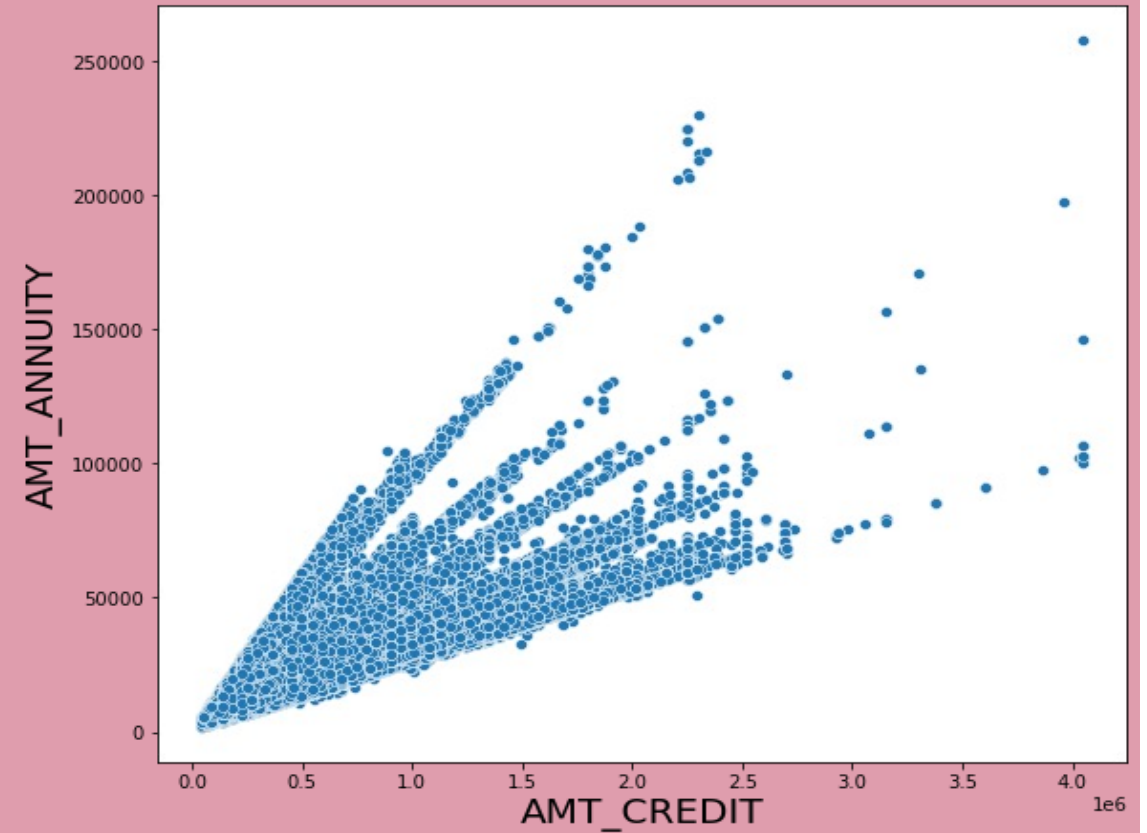
### People With Payment Difficulties

Previous amount credited vs Loan annuity



### People Without Payment Difficulties

Previous amount credited vs Loan annuity

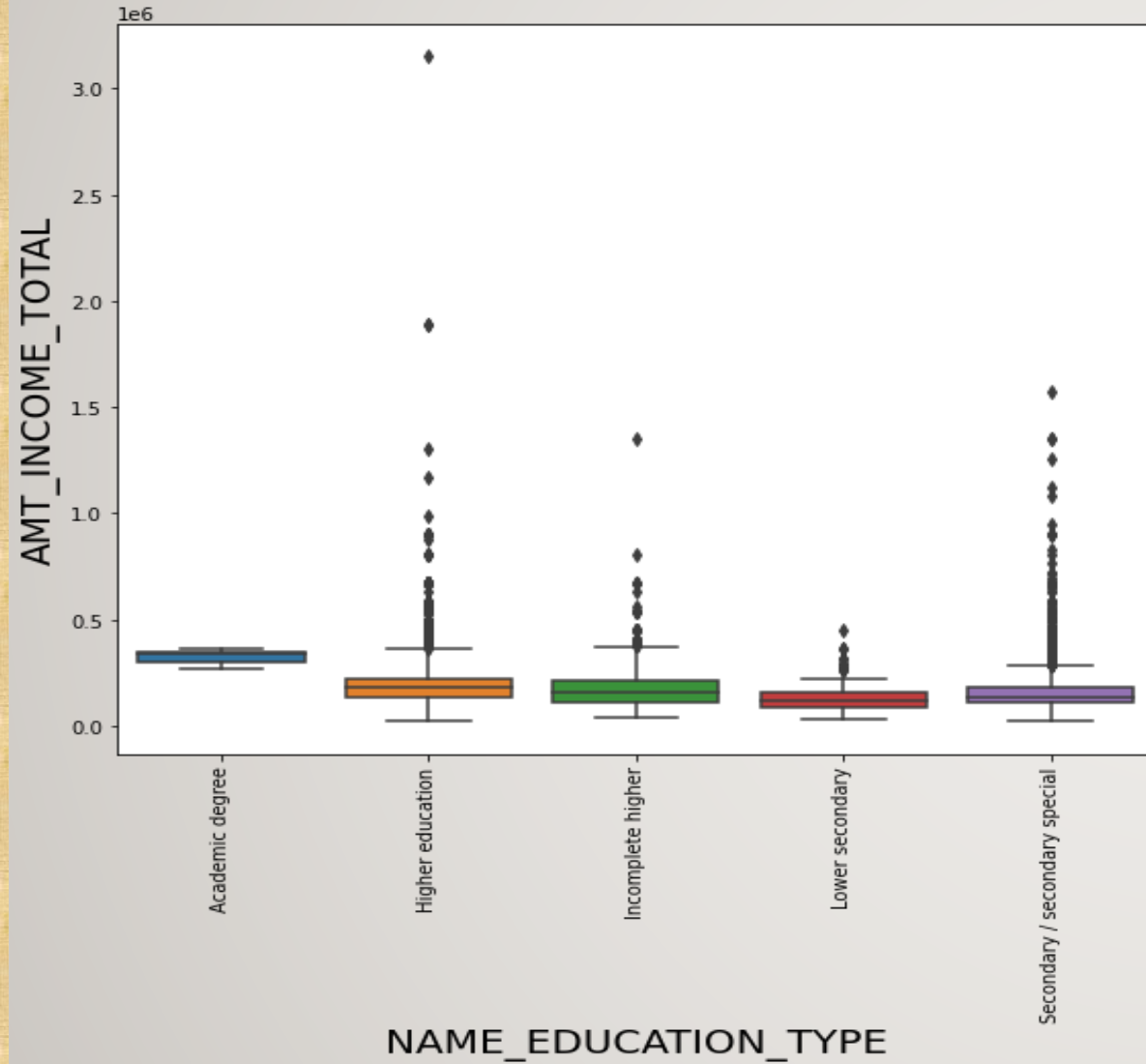


People with payment difficulties take more credit for the annuity they have



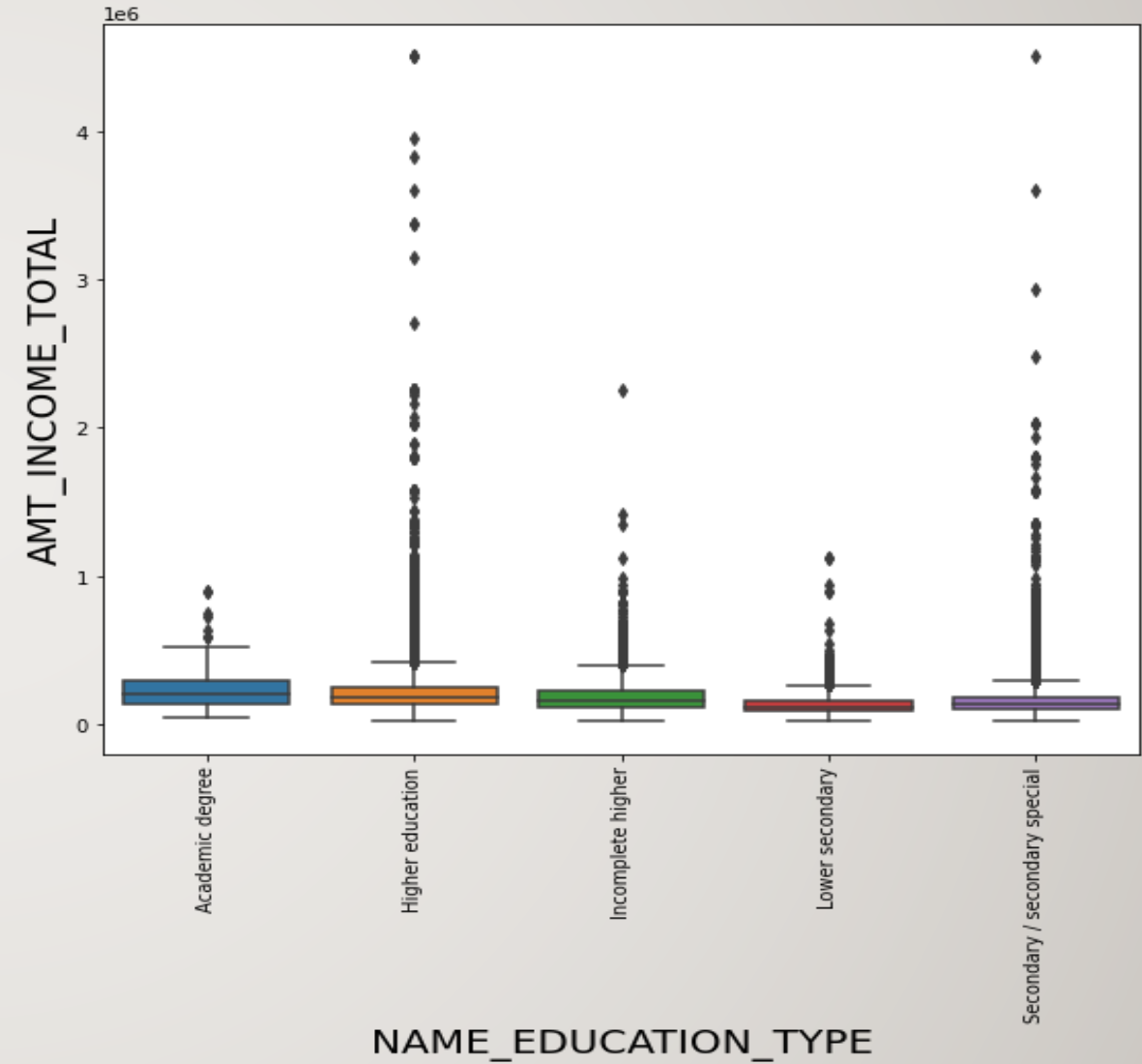
## People With Payment Difficulties

### Education type vs Income



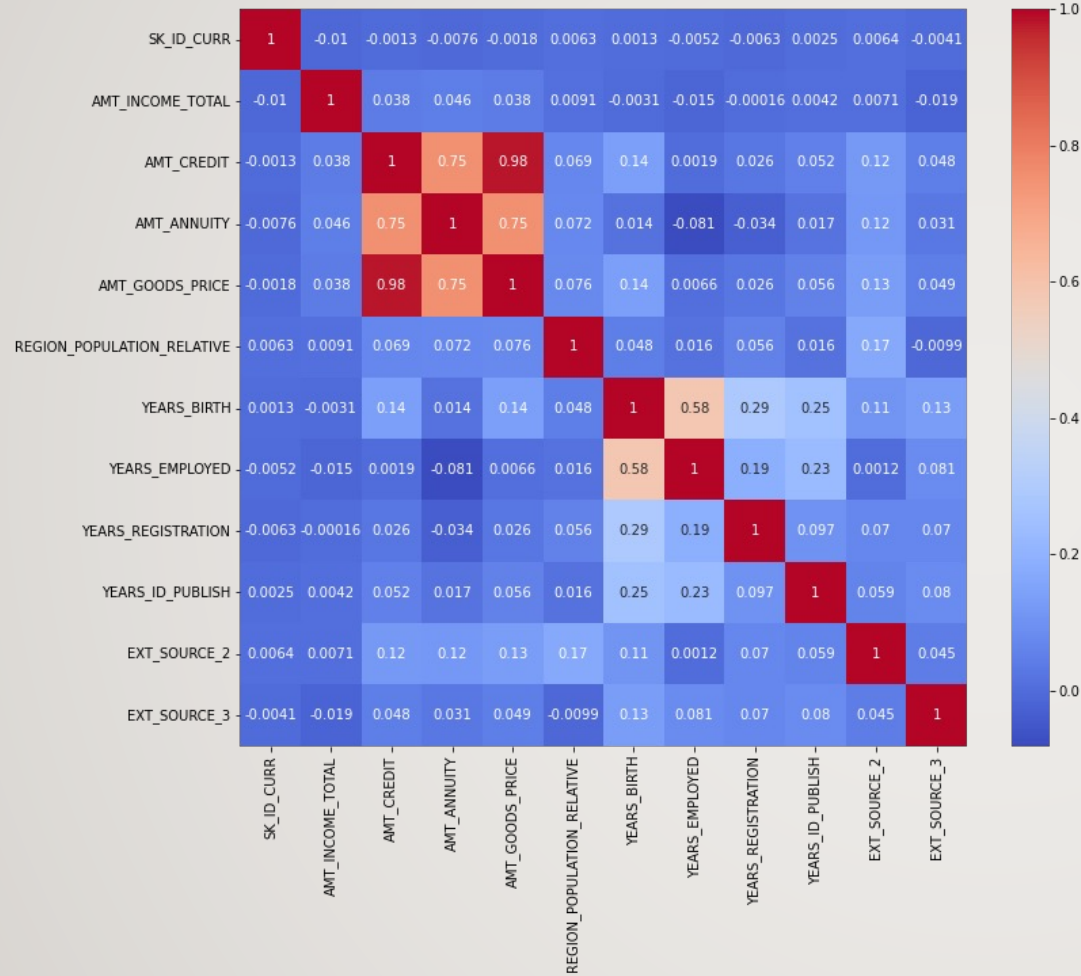
## People Without Payment Difficulties

### Education type vs Income

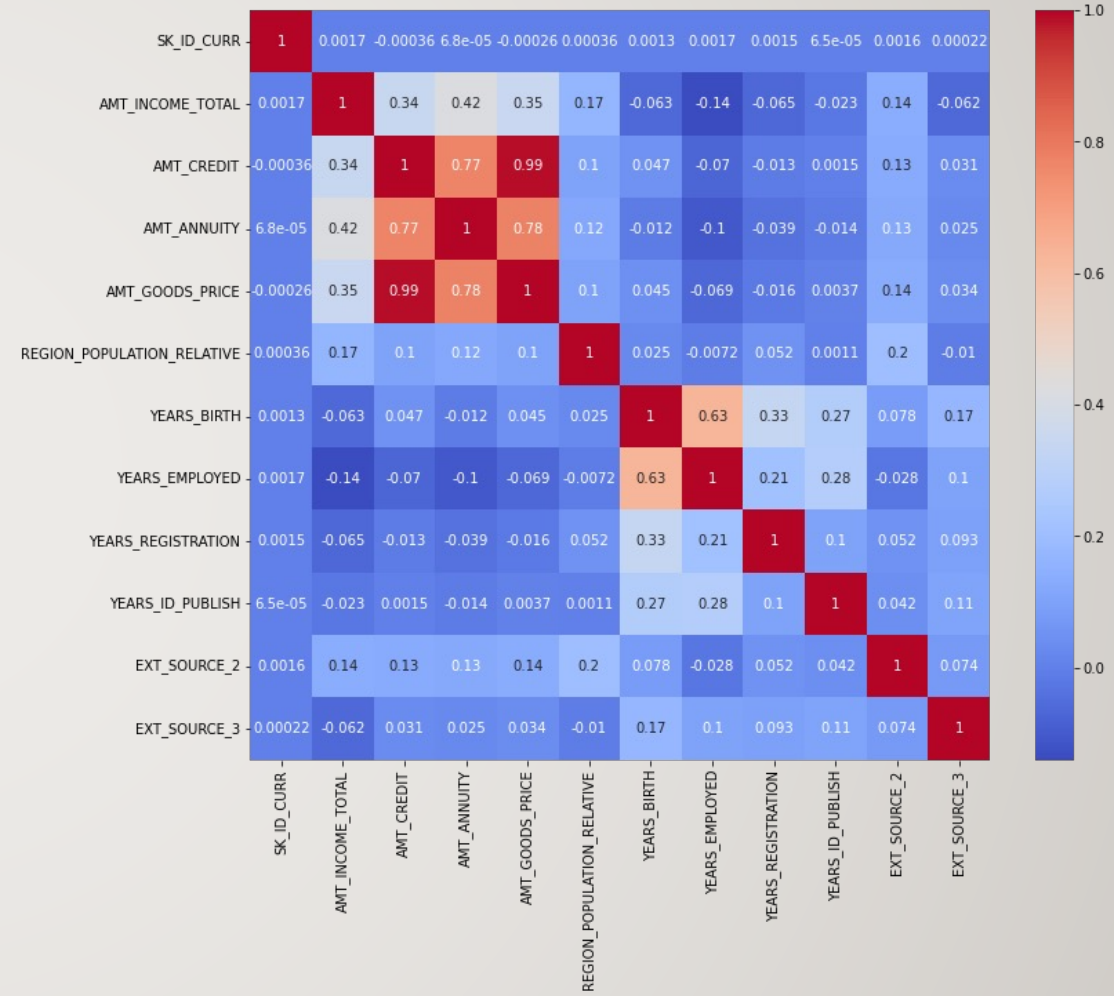


People with 'Academic degree' find it more difficult to pay

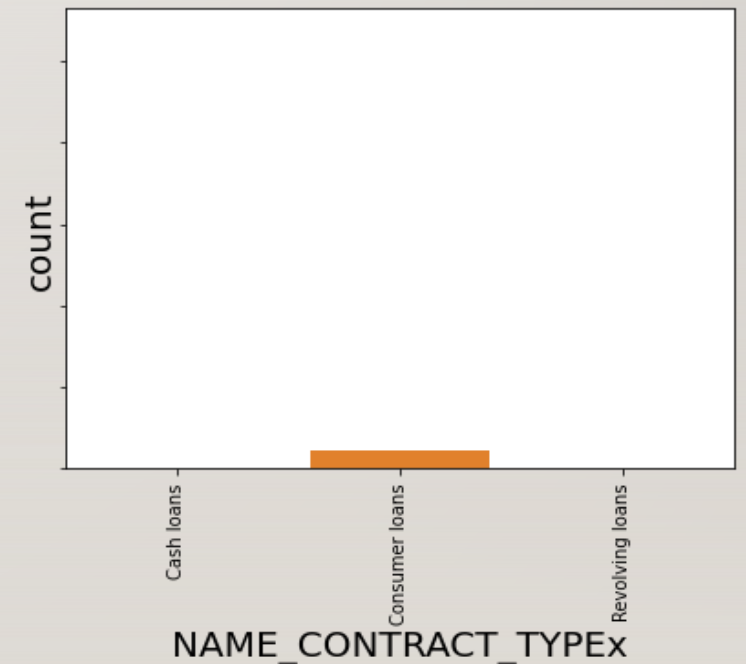
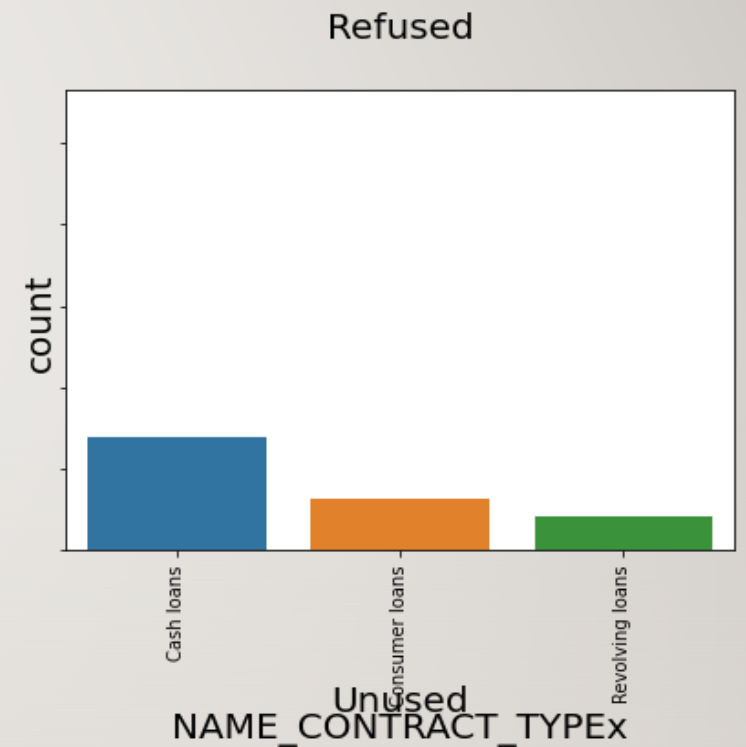
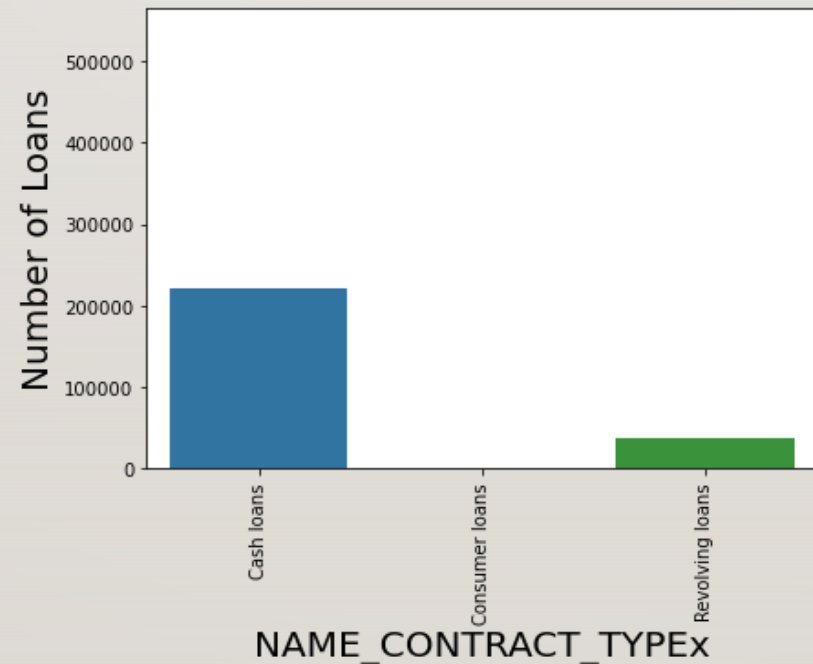
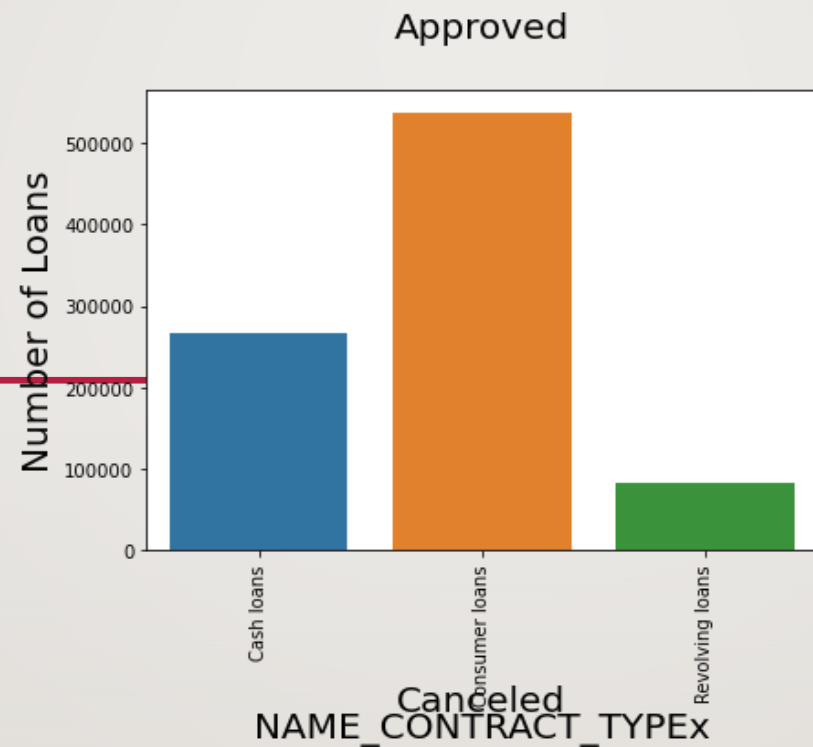
Correlation Matrix for People With Difficulty paying



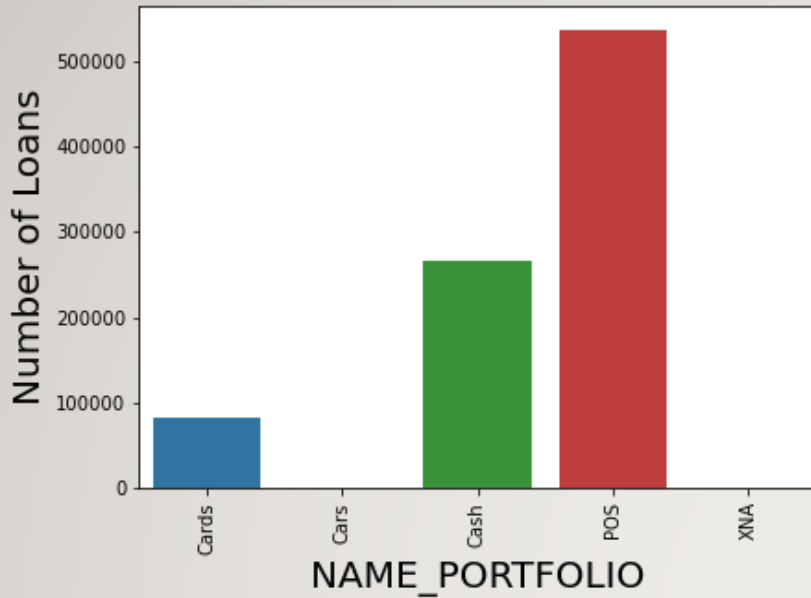
Correlation Matrix for People Without Difficulty paying



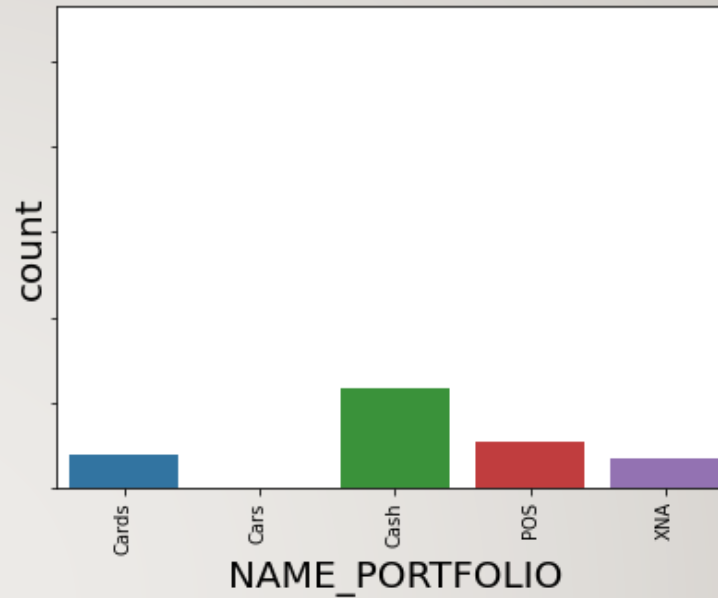
- **Consumer loans most likely to be approved**
- **Consumer loans most likely to be unused too**
- **Cash loans most refused and canceled**



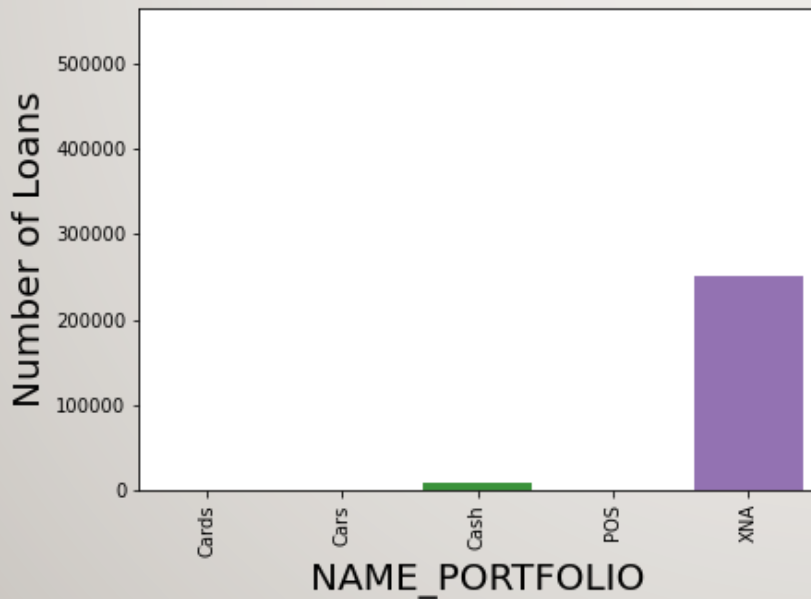
Approved



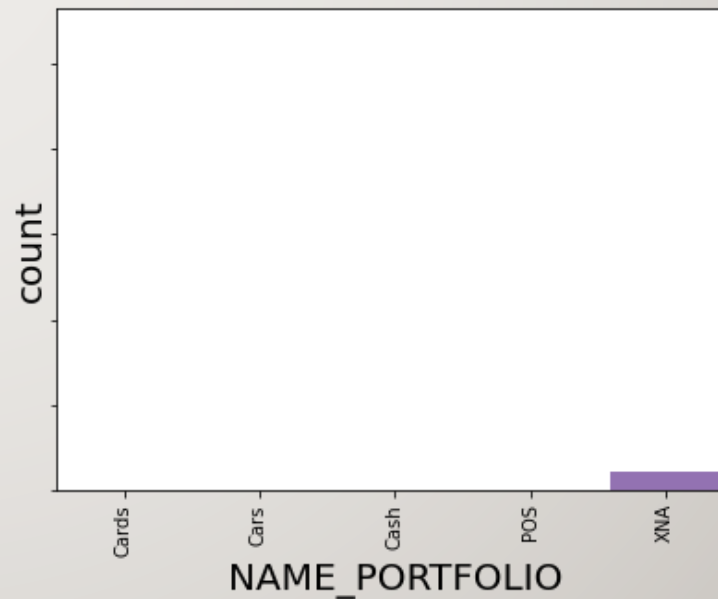
Refused



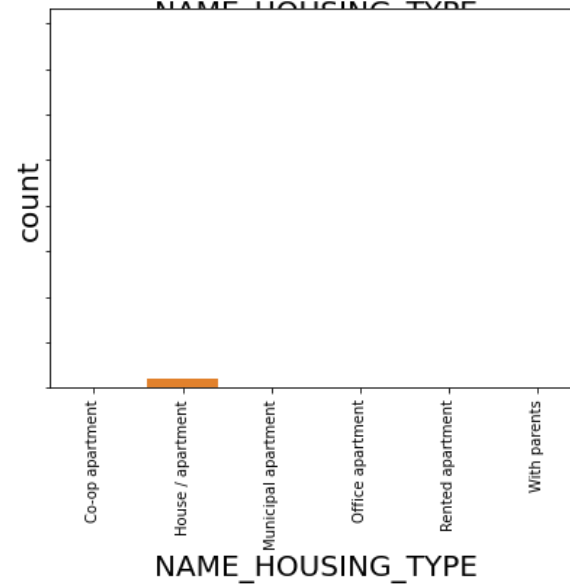
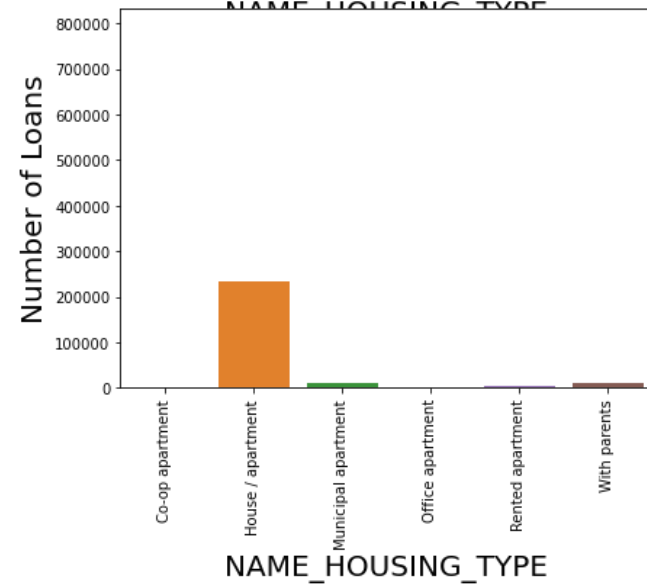
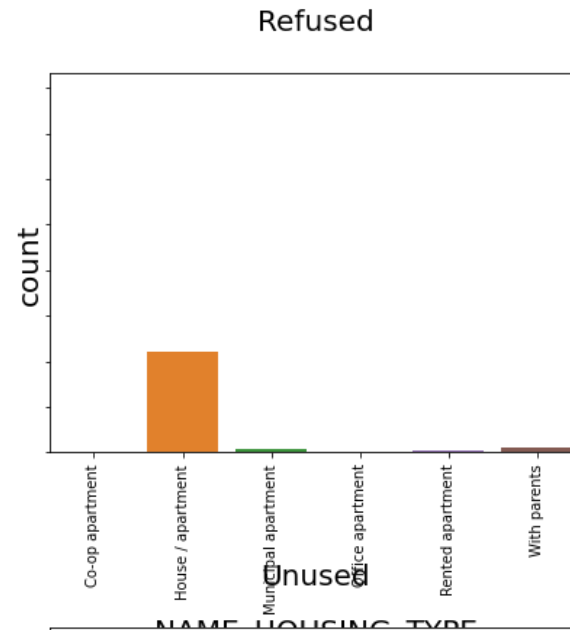
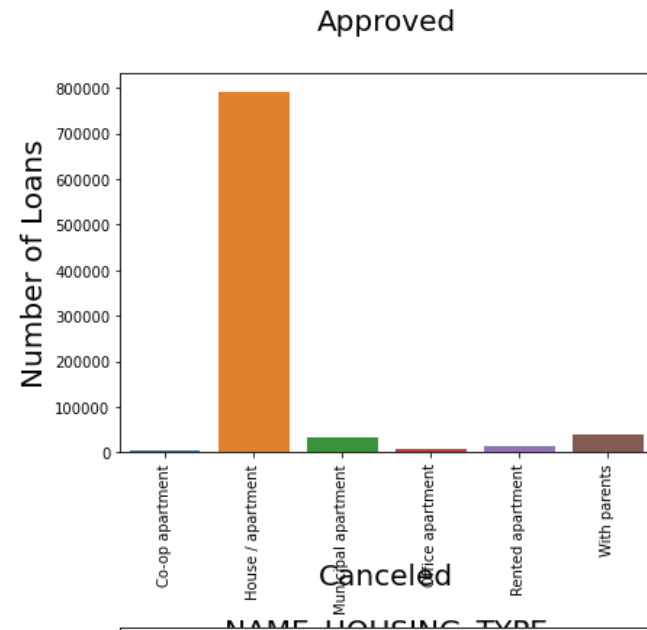
Canceled



Unused



- **POS get most approvals**
- **Cash application most refused**



Those who stayed in house/apartment got more approvals



# CONCLUSION

## WHO NOT TO TARGET ✖

---

- People working in Business entity type 2 and Construction
- People aged 31-35
- Don't give more credit than annuity
- People with 'Academic Degree'
- Avoid cash loan (most refused and canceled)

# CONCLUSION

## WHO TO TARGET

---

- People aged 50-55
- People working in 'Kindergarten' type organization
- People who stay in house/apartment
- POS applications
- Consumer loans

---

THANK YOU