Lending Club Case Study

By Abhijit Ray
Class of ML60 NLP

The business problem

- Our client, the largest online loan marketplace, is in the business of lending to retail customers.
- The current default % of loans is 14.17%, which means that 14 loans out of 100 are defaulted. This results in quite a lot of credit loss for the company.
- The company needs to identify risky loan applicants so that they can significantly bring down the risk of default and hence reduce the credit loss.
- In other words, the company wants to understand the **driving factors (or driver variables)** behind loan default

The method

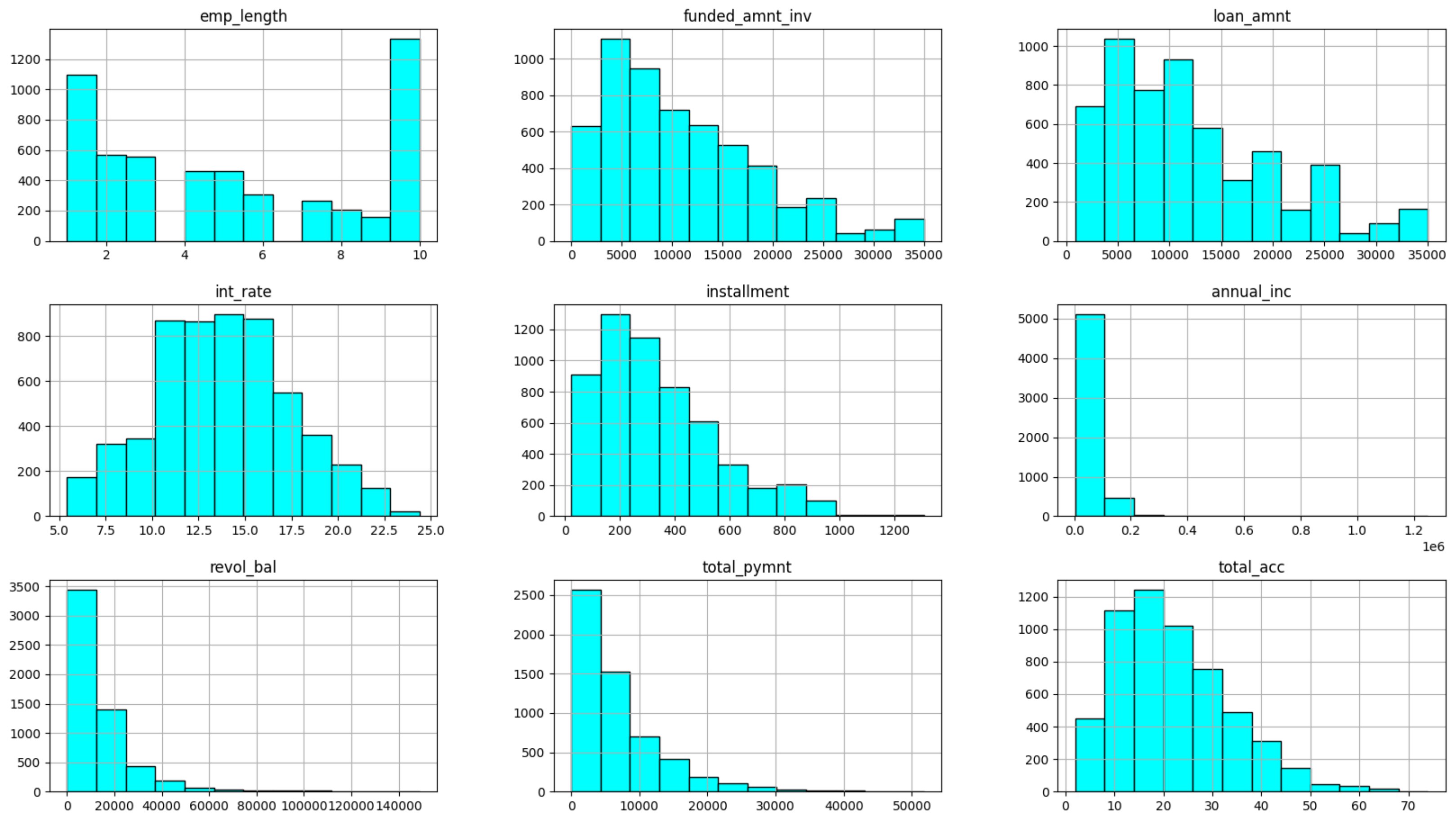
The What

- The company has shared loan details of all the loans with us.
- The data contains genuine borrowers along with defaulters.

The How

- We looked into data shared and utilised EDA methods to analyse the data and identify the likely parameters leading to defaults
- We focused on the defaulters to identify the factors that have most likely led to defaults.

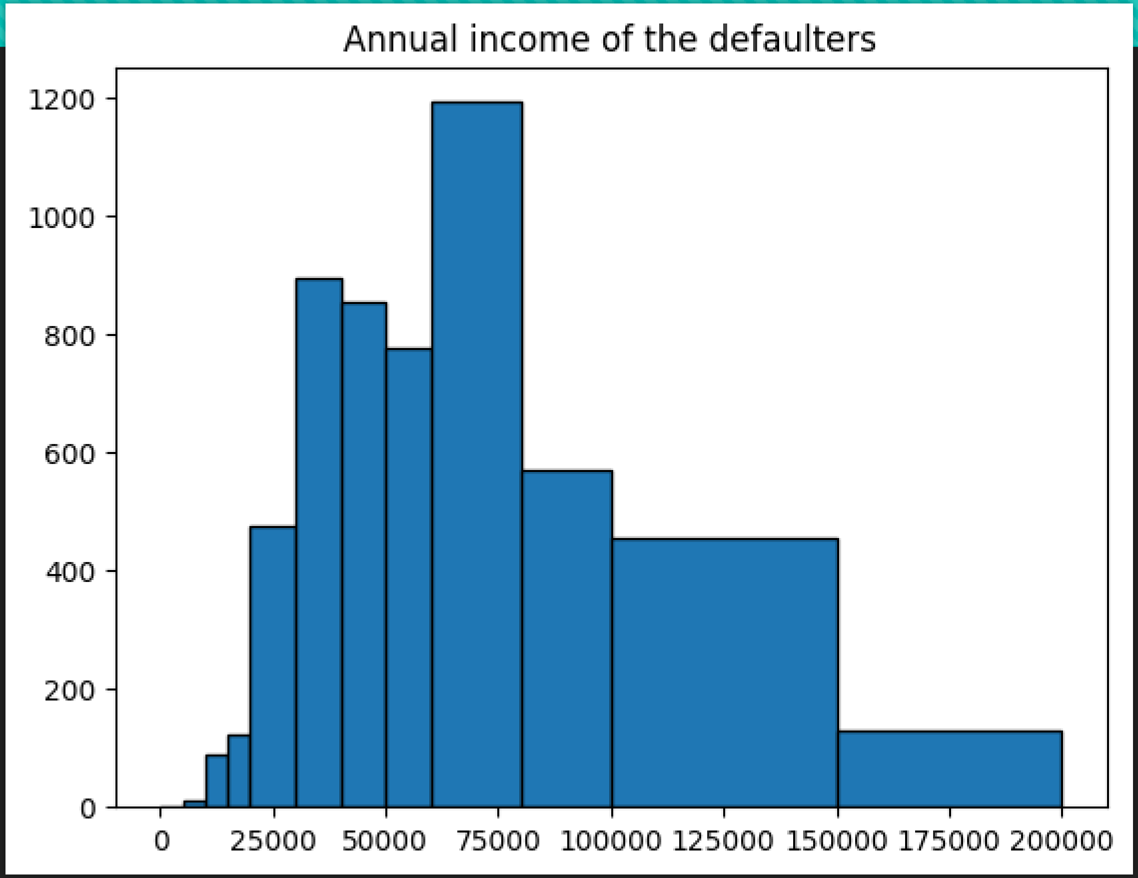
The defaulter data



Indicators

- Defaulters usually work long i.e. 10 years or under 1 year.
- Defaulters' interest rates are mainly 10-16%
- Defaulters' mostly pay installments from 0 to 500
- The annual incomes are on the lower side of almost all the defaulters

Defaulters' annual incomes



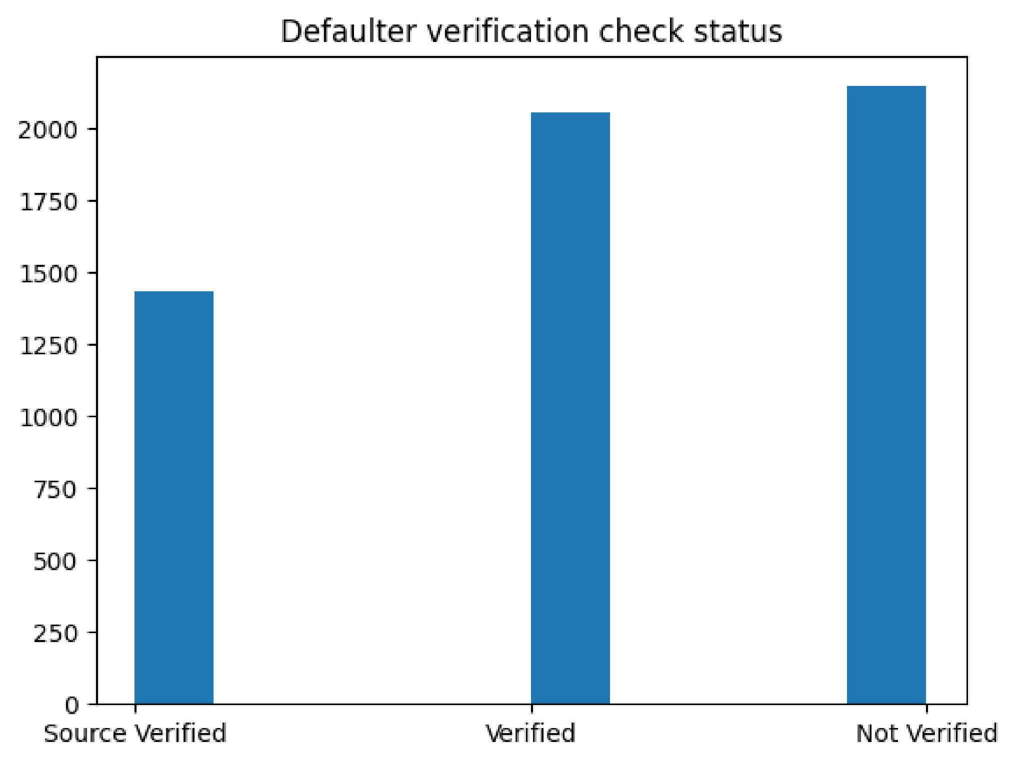
Indicators

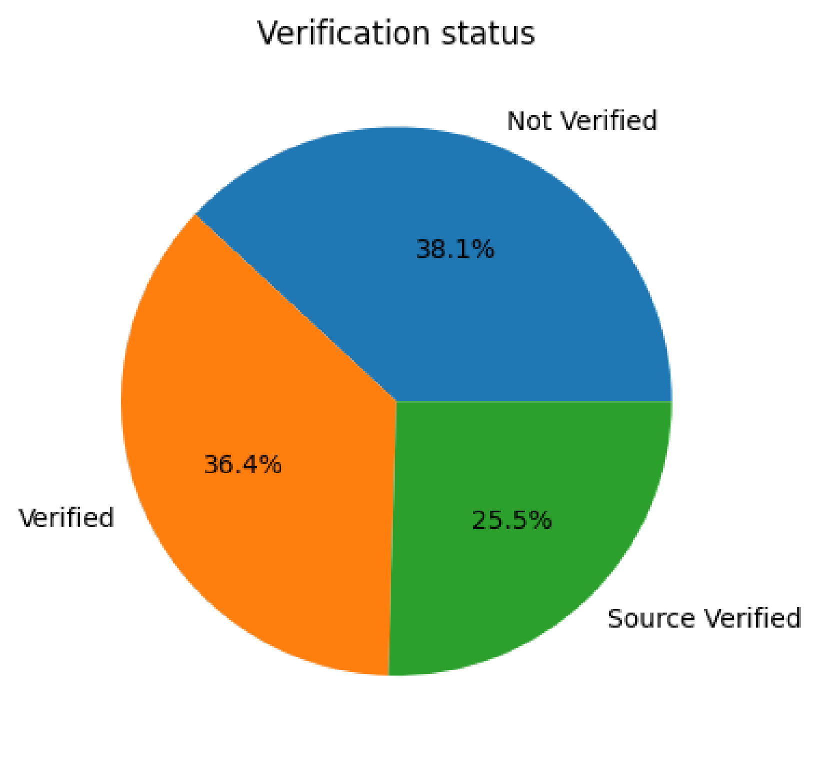
- A closer look reveals that the maximum defaulters earn between 60K and 80K. However we need to also consider those earning between 30K to 90K.

Conclusion

Those in the income range of 30K to 90K must be scrutinized and verified when they apply for a loan.

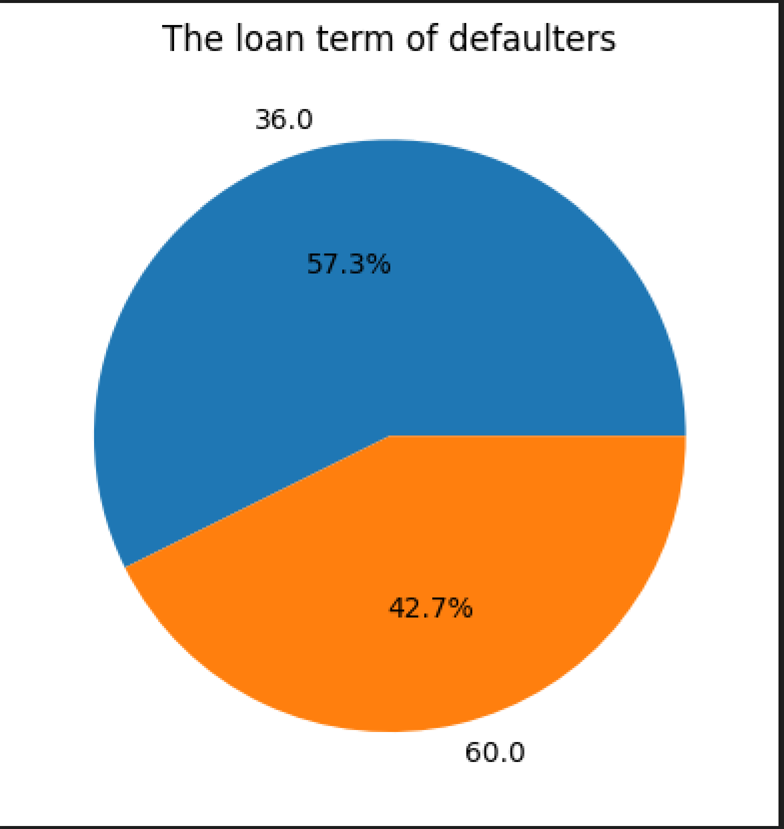
Defaulter: Verification status check





38.1% of the defaulters, i.e. more than a 3rd of the defaulters have not been verified.

Defaulter: Loan term



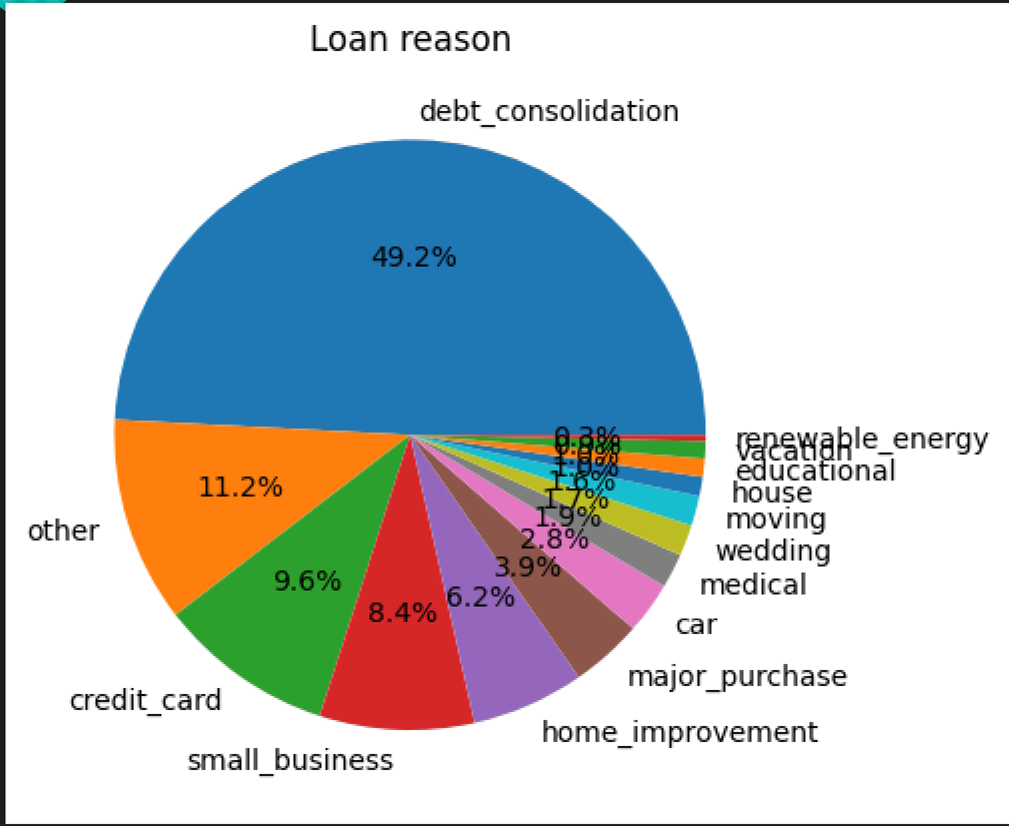
Analysis

- 57.3 of all the defaulters were on a 36 month loan

Conclusion

- **This is not a very conclusive evidence that the loan term has any direct correlation with default.**

Defaulter: Loan reason



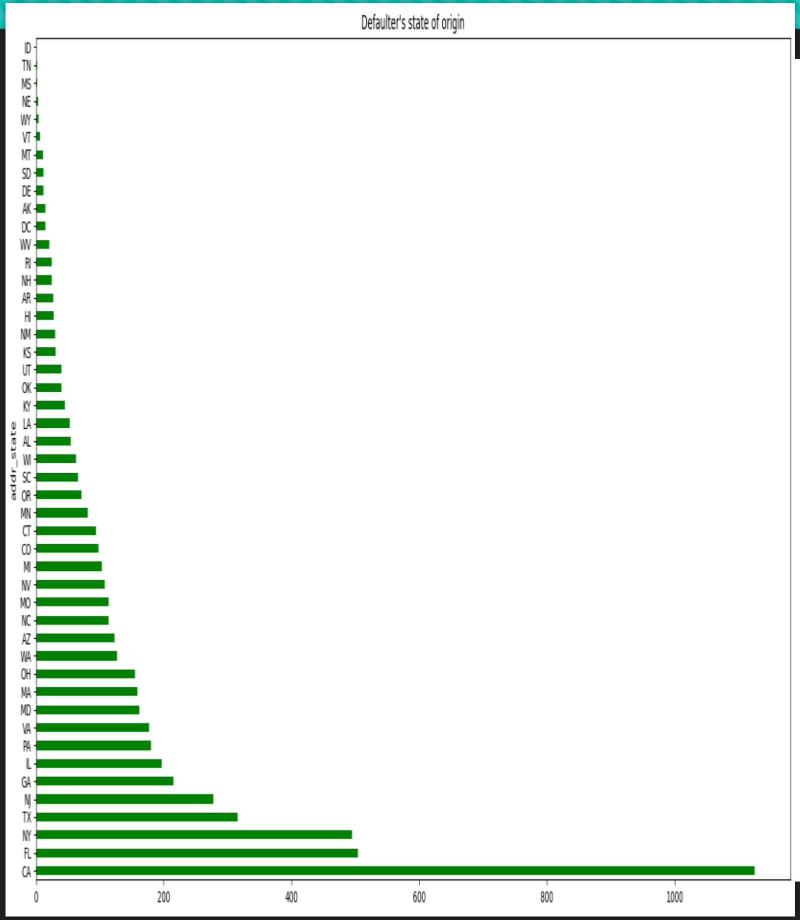
Analysis

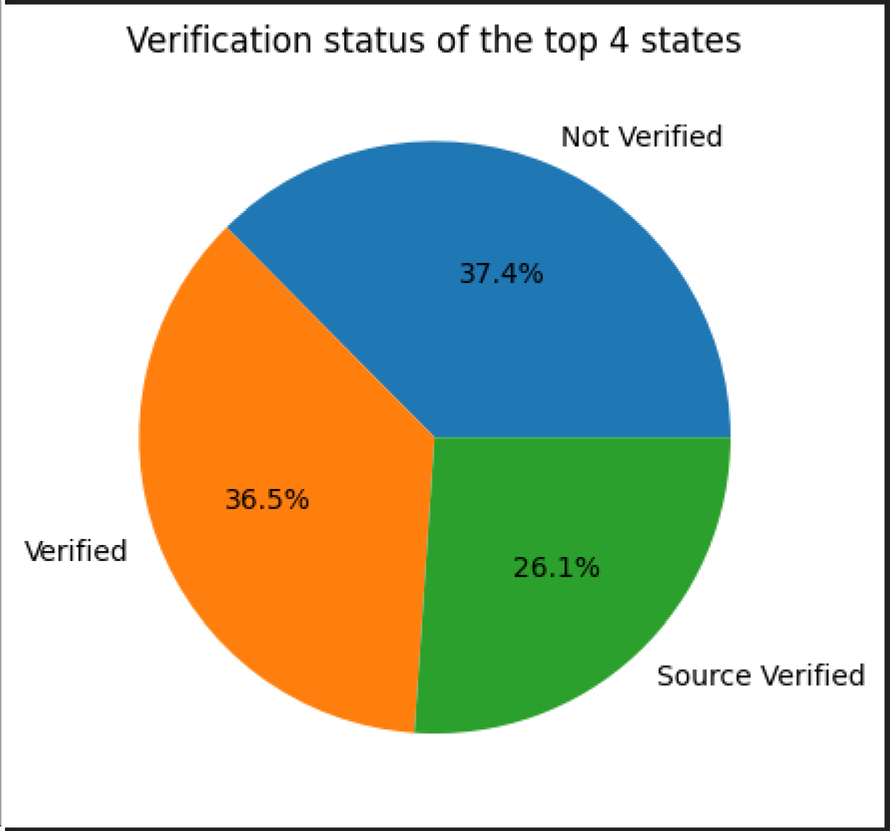
- **The data tells us that 49.2% have taken the loan for debt consolidation.**
- **This shows that applicants taking loans for repaying other loans are very likely to default.**

Conclusion

Any applicant with the need to consolidate loans or repay other loans need to be scrutinized closely.

Defaulter: State of origin





Analysis

- We find that most of defaulters reside in the states of,
- CA
- FL
- NY
- TX

- The total default % from the top 4 states is 43.36%

- Only 36.5% applicants from these 4 states have been verified

Conclusion

Loan applications from these states have to be scrutinized very closely.

Conclusion

All applicants have to be verified to before approving a loan

Application verification procedures have to be more robust with special emphasis on

Annual income	Applicant State	Reason for loan application