Lending Club Case Study

Prepared By: Shishir Khanal

Table of Contents:

Introduction:

- The consumer finance industry is highly competitive, and lending companies face various challenges.
- Identifying risky loan applicants and minimizing credit losses are major concerns for lending companies.
- This case study aims to help a consumer finance company understand and mitigate the risk of loan defaults.
- We will use Exploratory Data Analysis (EDA) to gain insights into the factors that influence loan default.

Background:

- Lending Club is a consumer finance marketplace that connects borrowers with investors.
- Company specializes in offering various loans to urban customers.
- Loan types include personal loans, business loans, and medical procedure financing etc.
- Lending Club is largest online loan marketplace with competitive interest rates.

Figure:1

Figure:2

Problem Statement:

- For lending companies, the largest source of financial loss is credit loss
- Which mainly occurs when borrowers refuse to pay or default on their loans
- By utilizing EDA techniques, the company intends to understand the driving factors behind loan default, which are strong indicators of potential credit losses
- We will use Exploratory Data Analysis (EDA) to gain insights into the factors that influence loan default
- The company faces two primary risks in loan decision-making
  - The risk of not approving a potentially repayable loan
  - The risk of approving a loan that may default

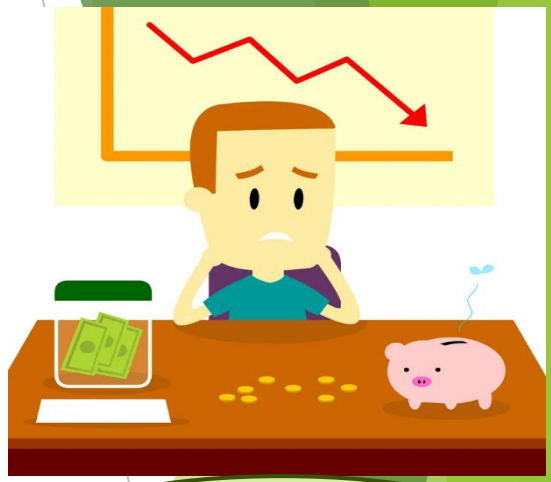


Figure:3

Figure:4

Objectives:

- The goal of this case study is to identify risky loan applicants by using exploratory data analysis (EDA)
- Primary objective is to identify key leading indicators (driver variables) in the dataset that contribute to loan defaults
- The main target is to help in decision making of approval or rejection of loan applications
- This knowledge can be applied to risk assessment, portfolio management, and strategic decision-making

Figure:5

Data Overview:

- Dataset Description:
  - The dataset contains complete loan data for all loans issued between 2007 and 2011.
  - It includes information about loan applicants and the outcome of their loans, specifically whether they defaulted or not. It has three cases of loan which is fully Paid, current and charged-off.

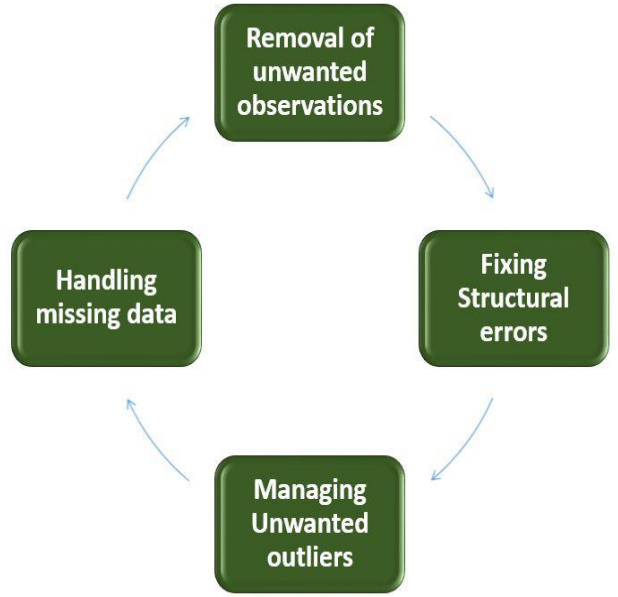
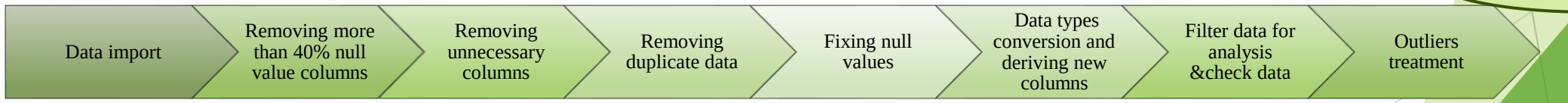


Figure:6

- Data Cleaning and Manipulation:



Analysis Approach:

- Data Preprocessing
- Exploratory Data Analysis (EDA)
- Data Visualizations
- Feature Importance
- Insights and Recommendations

Figure:7

Loan Status Vs Loan Amount

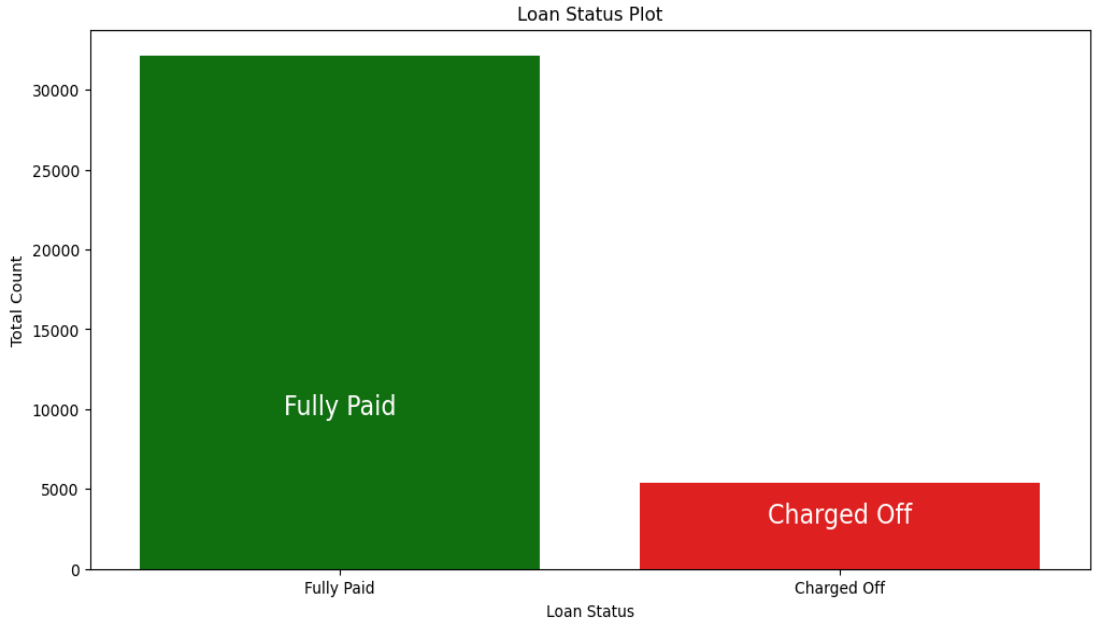


Figure:8

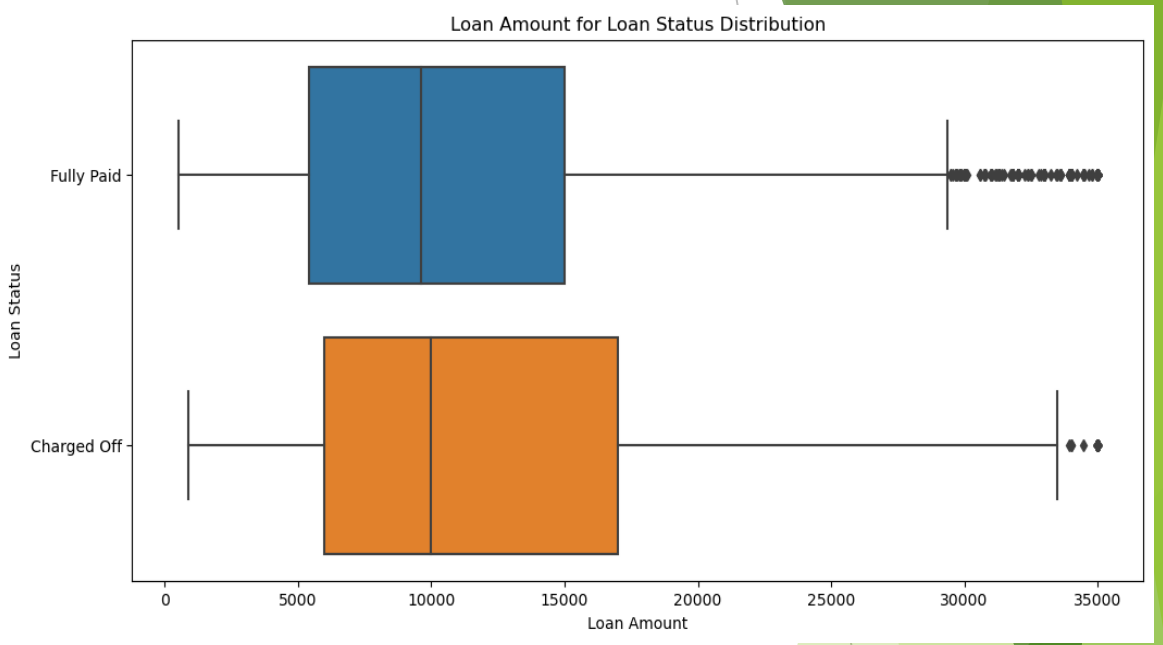


Figure:9

Observations 1:

- The proportion of defaulted loans is less than fully paid.

Observations 2:

- The mean and 25th percentile values are similar for both fully paid and defaulted loans, but 75th percentile in defaulted loans is noticeably higher. This discrepancy implies that larger loan amounts have a higher likelihood of defaulting.

Grade, Sub-grade and Term

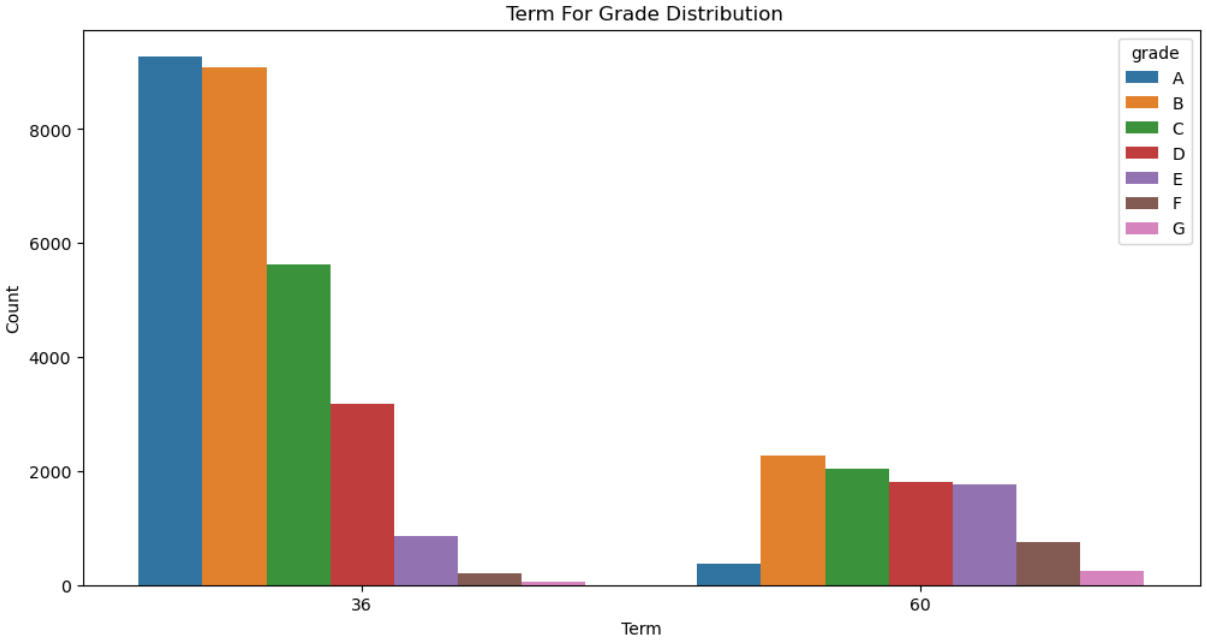


Figure:10

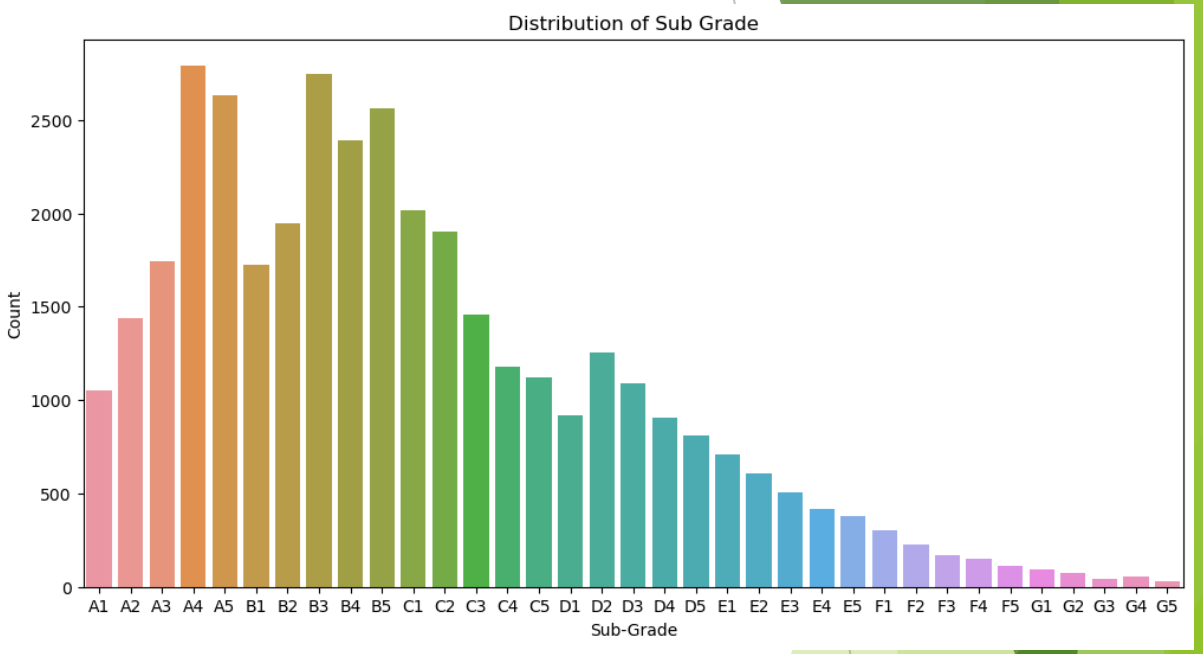


Figure:11

Observations 3:

- The majority of loans with a 36 months term are in grade A and B, while those with a 60 months term are primarily in grade B, C, and D.

Observations 4:

- Most loans belong to high-grade categories 'A' and 'B.'

Term, Interest Rate and Loan Status

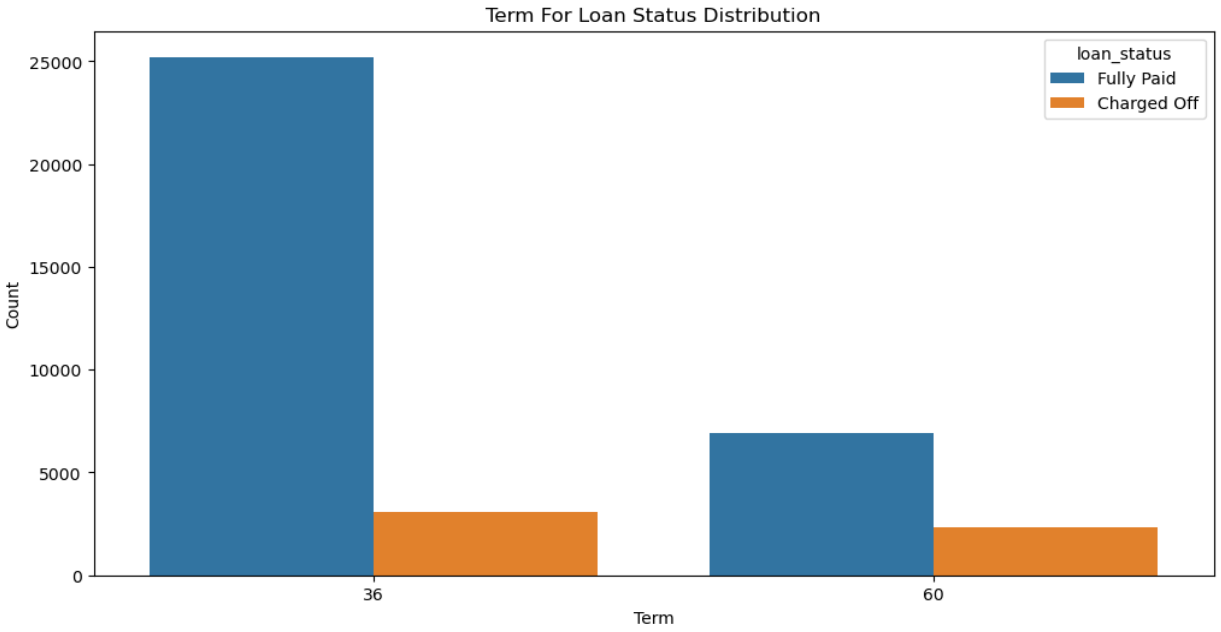


Figure:12

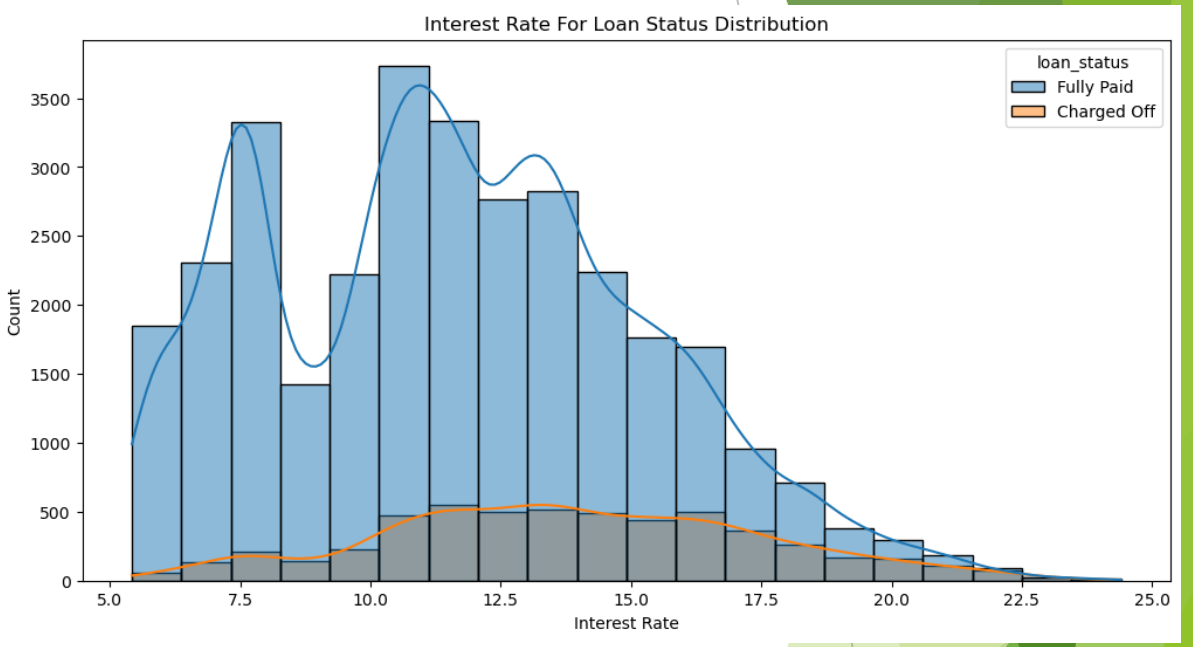


Figure:13

Observations 5:

- A majority of loans are for a 36-month term compared to 60-month term.

Observations 6:

- The default loan amount tends to rise with increasing interest rates, but it shows a decline after reaching a peak around 17.5%.

Employment Length, Home Ownership & Loan Status

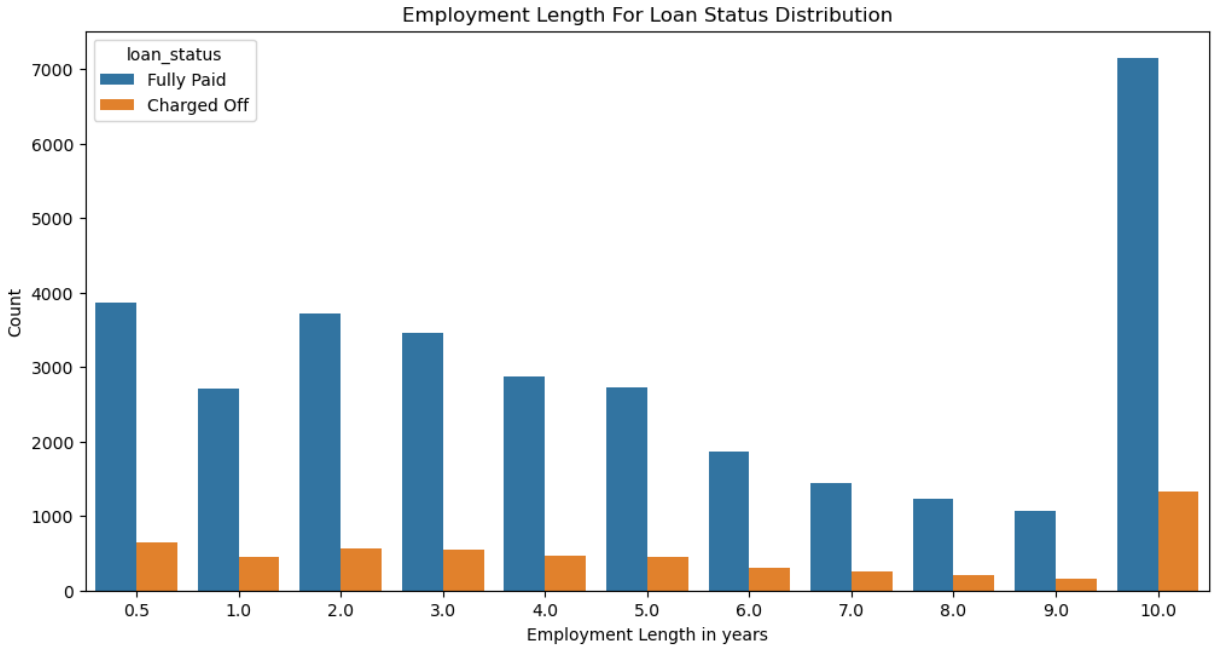


Figure:14

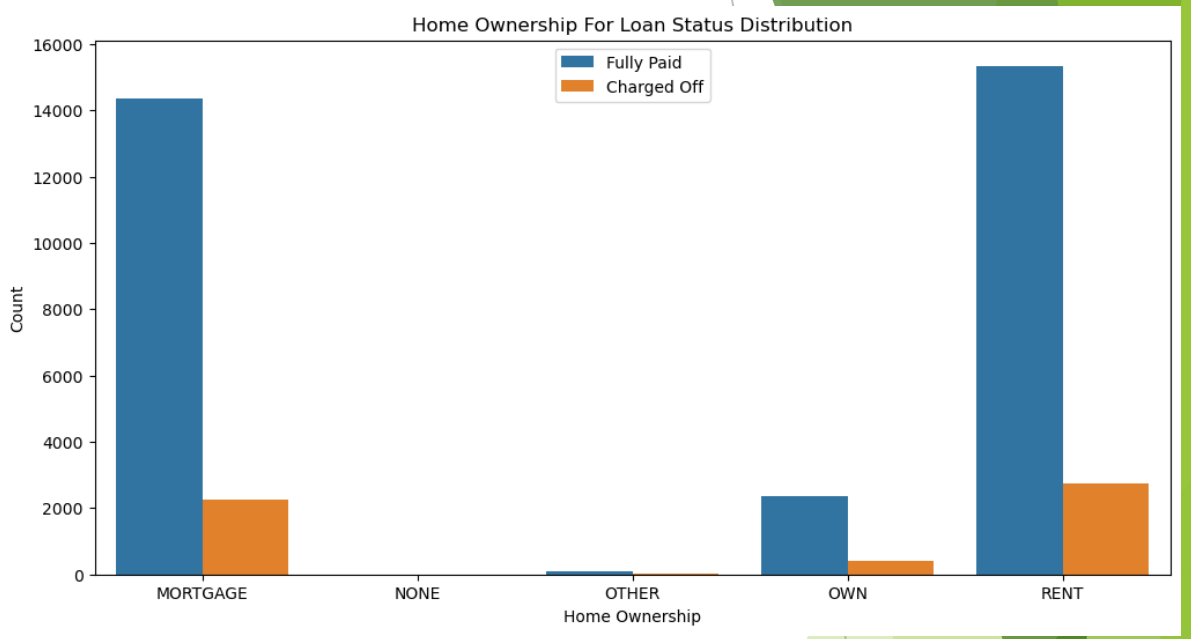


Figure:15

Observations 7:

- Employees with over 10 years of experience are more likely to default on their loans, but they also have a higher probability of fully repaying them.

Observations 8:

- Most borrowers do not own property instead, they are either in the mortgage or rental category.

DTI Ratio, Bankruptcy and Loan Status

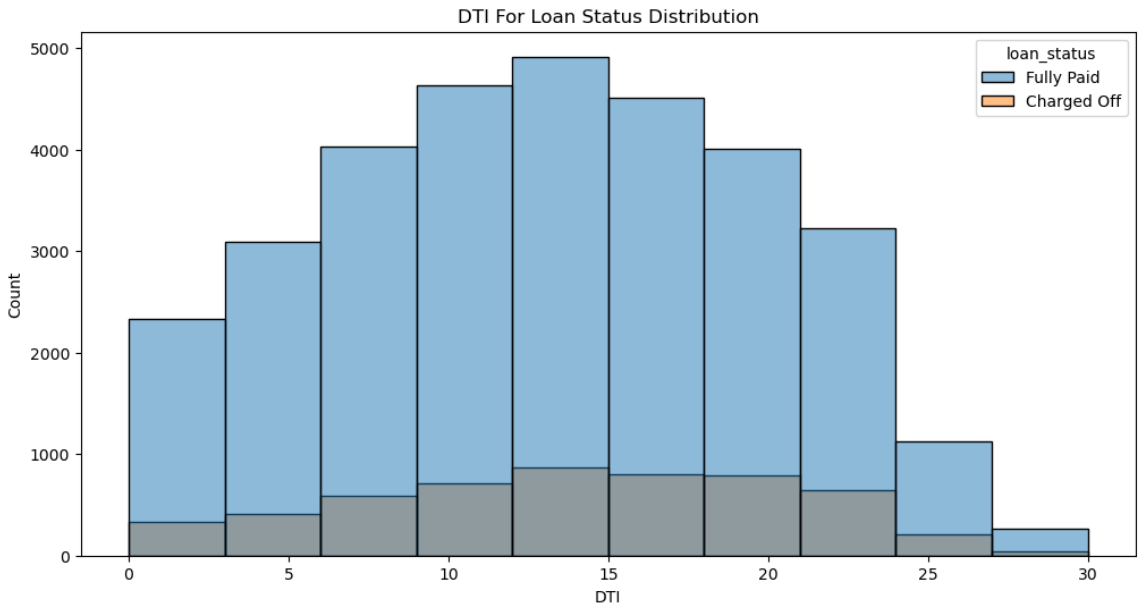


Figure:16

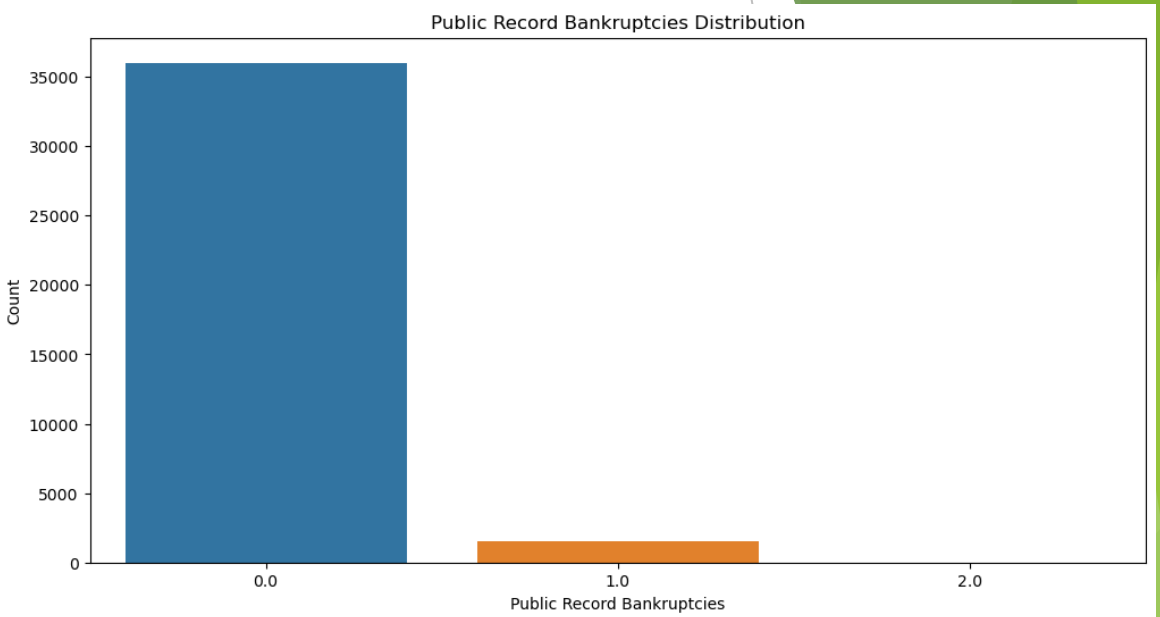


Figure:17

Observations 9:

- A significant percentage of clients have a high Debt-to-Income ratio, indicating that lending to such clients carries a high level of risk.

Observations 10:

- The majority of borrowers do not have any recorded history of bankruptcy in the public records.

Annual Income, Purpose and Loan Status

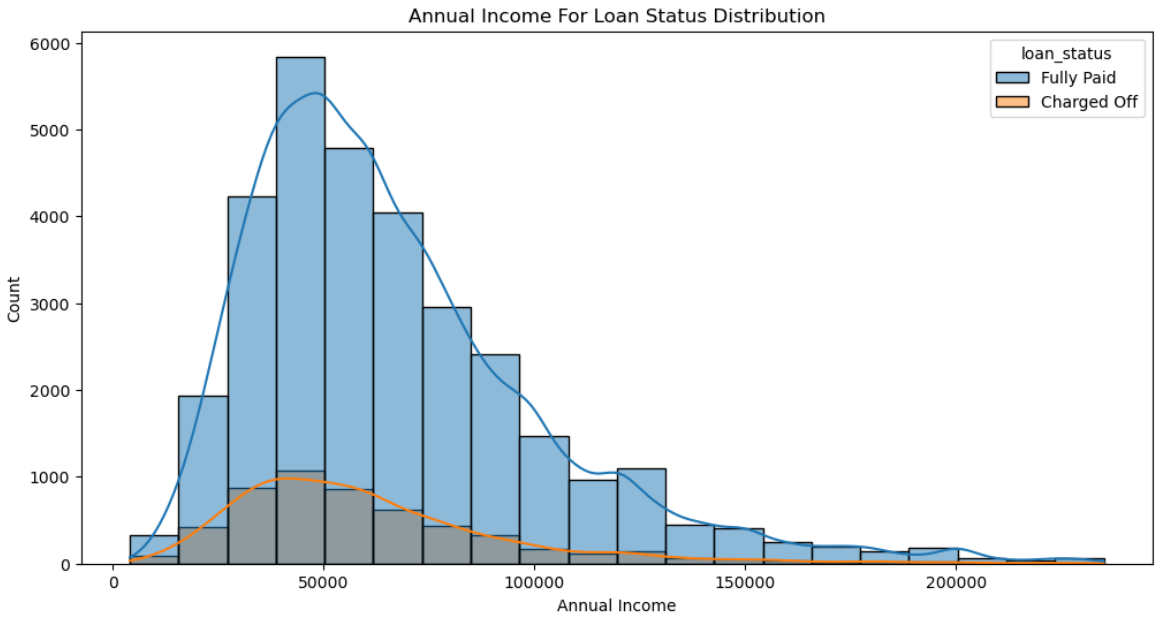


Figure:18

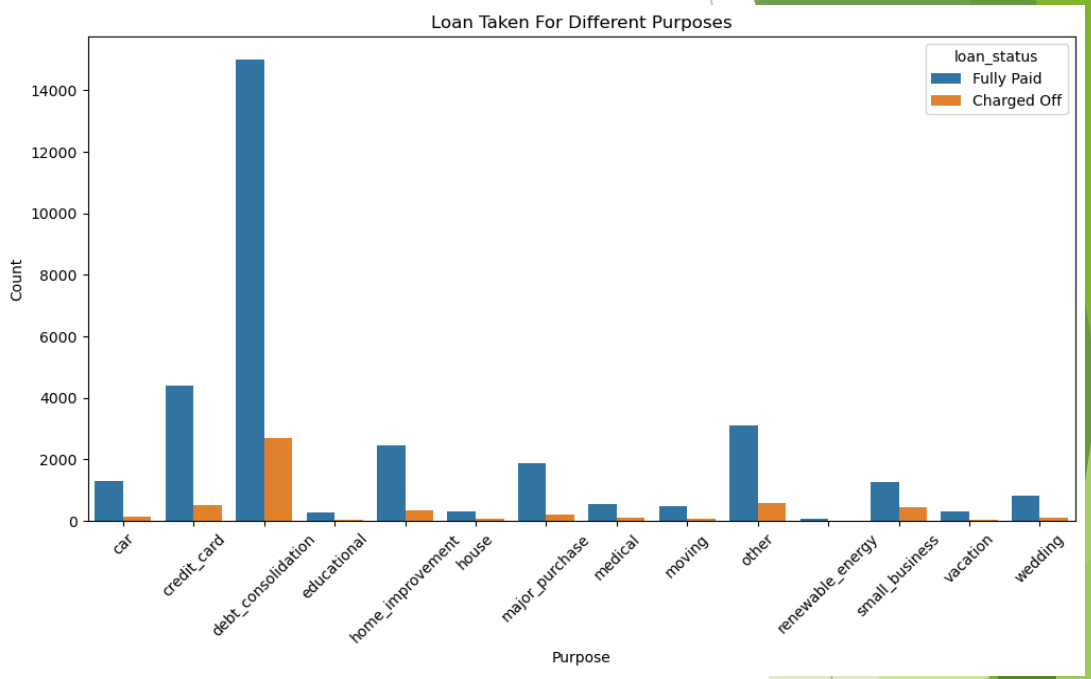


Figure:19

Observations 11:

- Borrowers with annual incomes below 50,000 are more likely to default, while those with higher annual incomes are less likely to default.

Observations 12:

- The majority of loans are used for debt consolidation, with credit card-related loans as the second most common purpose.

State

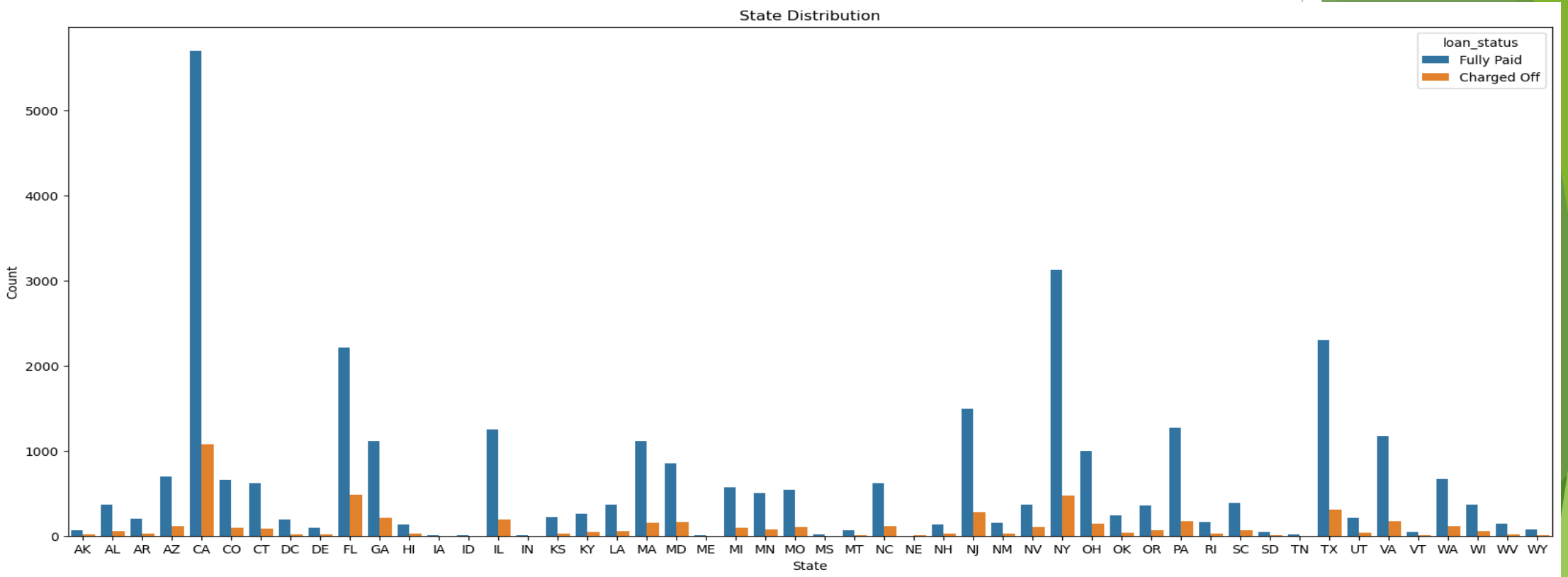


Figure:20

Observations 13:

- Most borrowers originate from major urban areas such as California, New York, Texas, Florida, and others.

Issue Month, Year and Loan Status

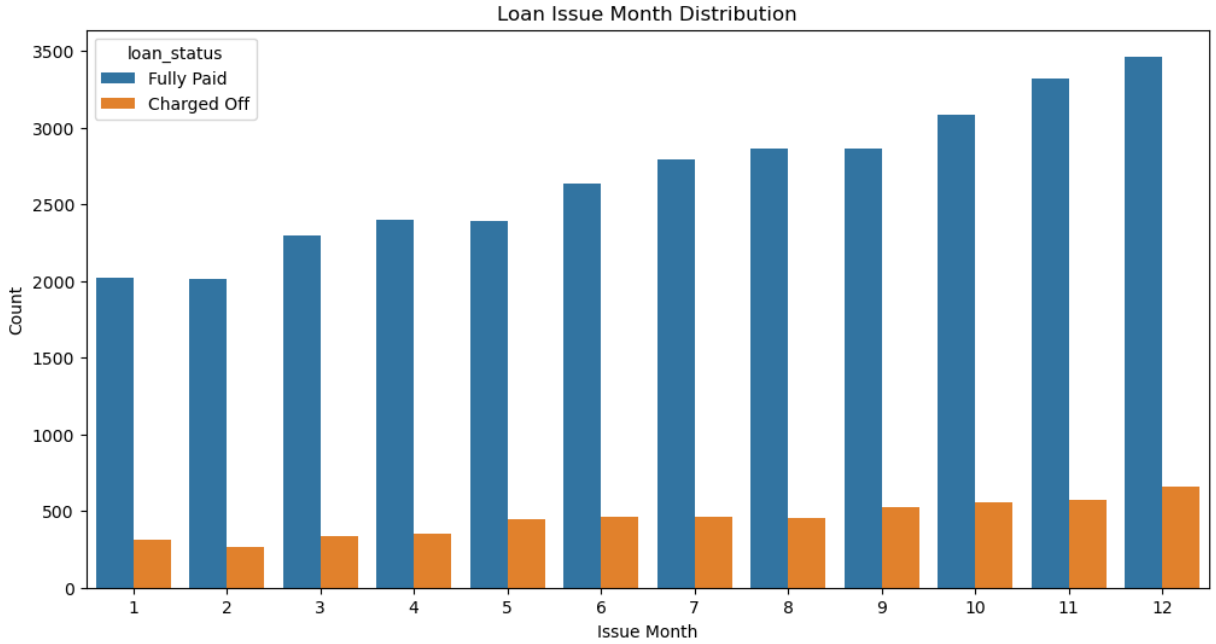


Figure:21

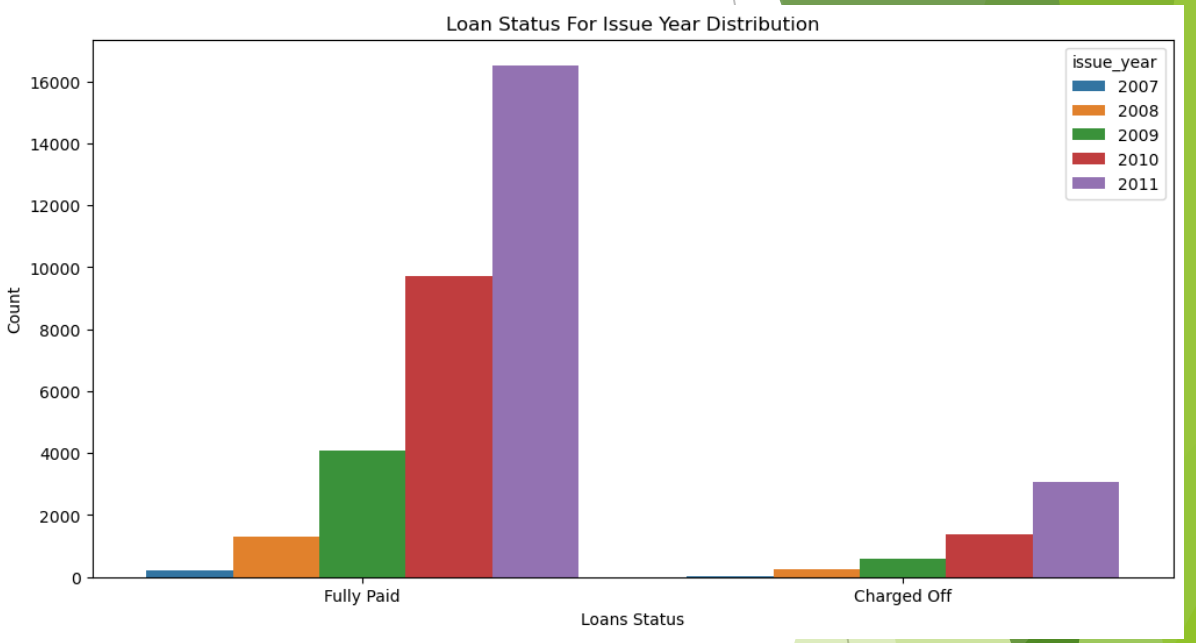


Figure:22

Observations 14:

- Most loans are distributed during the last quarter of the year.

Observations 15:

- As each year passes, there is an exponential increase in the number of loans taken, which suggests a significant rise in the Debt-to-Income (DTI) ratio and a decrease in the default rate.

Conclusion:

- The key factors which can be used for predicting default and minimizing credit losses in lending include grades, Debt-to-Income ratio (DTI), verification status, annual income, and public record bankruptcies.

- Additionally, borrowers with over a decade of work experience, lower annual incomes, a history of public-recorded bankruptcies, high Debt-to-Income ratios, lower grades (E, F, G), and those located in non-urban regions are more likely to default on their loans.

Recommendations:

To minimize credit losses and improve lending practices, it is recommended to:

- Focus on borrowers with lower DTI ratios, as they tend to have a lower risk of default
- Prioritize borrowers with higher grades, as they are less likely to default
- Exercise caution when lending to borrowers with over a decade of work experience
- Pay special attention to borrowers with lower credit grades (E, F, G) due to their increased risk
- Evaluate loans in non-urban regions more carefully
- Be cautious when dealing with borrowers who have a history of public-record bankruptcies

Technologies Used:

- Programming Language:
  - Python, version 3

- Libraries and Tools:
  - NumPy for numerical computations
  - Matplotlib and Seaborn for data visualization
  - Pandas for data manipulation
  - Jupyter Notebook for interactive analysis

Figure:23

Acknowledgements:

- Acknowledgements:
  - I acknowledge and appreciate the valuable course materials from upGrad and IIIT-B which enhanced my understanding of data analysis and EDA.

- References:
  - Python documentations,
  - Exploratory Data Analysis
  - Stack Overflow

Figure:24

Thank You!!