LENDING CASE-STUDY

BY:- APOORV JAIN & ASHISH RASTOGI

AGENDA

- ✓ The aim is to find the identifiers that can help reduce the risks when lending to different types of customers.
- ✓ This is the data cleaning and data analysis project (EDA), aimed at applying the concepts learned as part of Upgrad AI/ML program

GENERAL INFORMATION

1. **What is the background of your project?**

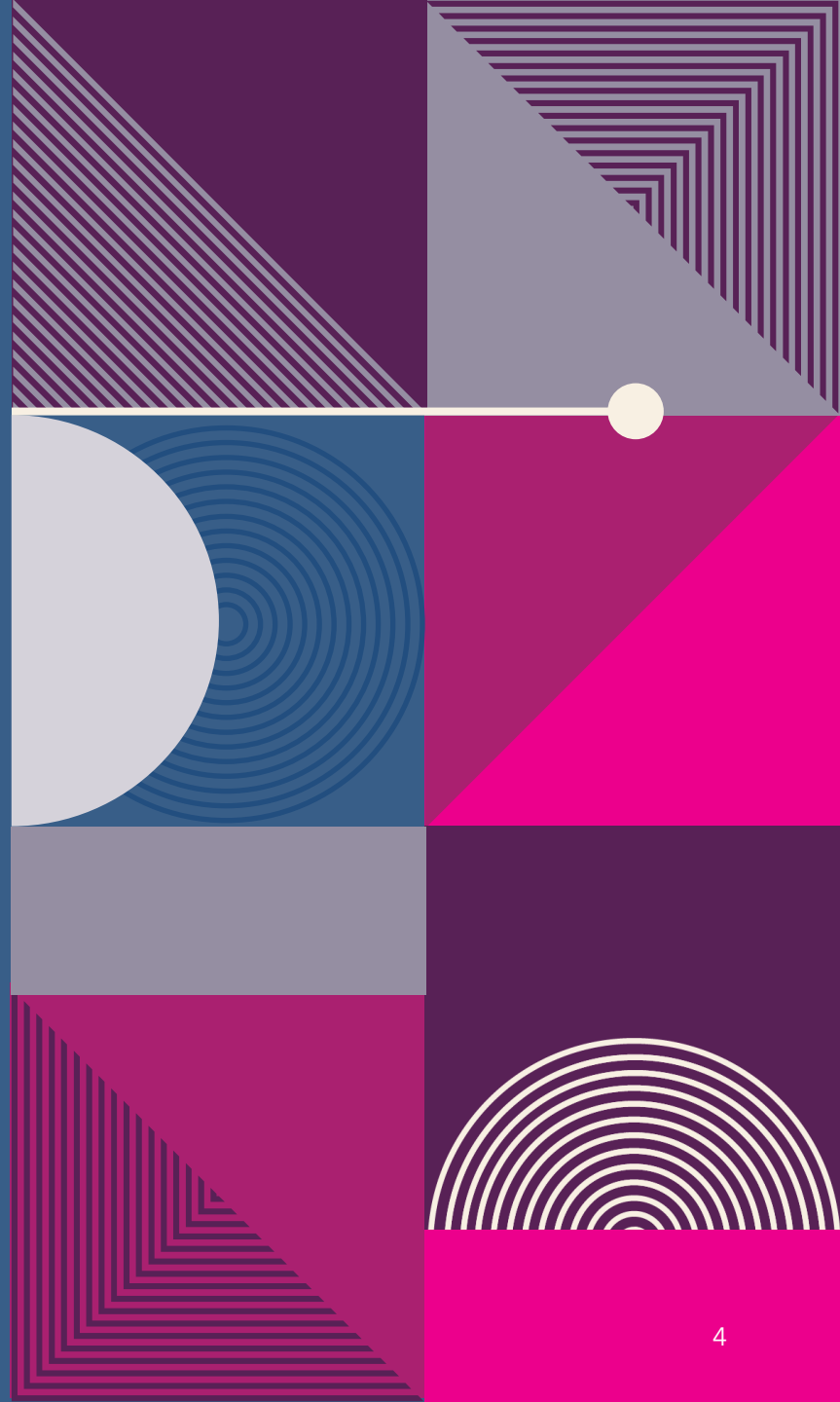
   This project deals with exploratory data analysis of lending data from the banking and financial sectors.

2. **What is the business problem that your project is trying to solve?**

   - In the lending industry, there are different types of risks involved when dealing with different customers.
   - To figure out these risks, we need to analyze existing data and figure out identifiers that can help in filtering customers who are more likely to default.

3. **What is the dataset that is being used?**

   The dataset being used is a "Loan Data Set" which contains the complete loan data for all loans issued through the time period 2007 to 2011, as provided by Upgrad, the copy of this data set has been added to this repo as well by the name "loan.csv"



TECHNOLOGIES USED

1. jupyter notebook - version 7.0.8
2. pandas - version 2.2.1
3. numpy - version 1.26.4
4. matplotlib - version 3.8.0
5. seaborn - version 0.13.2
6. python - version 3.11.8

CONCLUSION

- The loans with 1 or more publicly recorded bankruptcies are more likely to be "Charged Off" or default on their loan
- The loans where the percent of the total amount of loan paid or promised by an investor is less, the loans are more likely to be defaulted
- Loans are more likely to default when the interest rates are higher
- Long-term loans (60 months) are more likely to be charged off
- For loans where homes are "MORTGAGE" and the loan amount is high, it is more likely to default

THANK YOU

Ashish Rastogi – Githup Id – ashish123007

Apoorv Jain – Github Id – apoorvjain1