# LENDING CLUB CASE STUDY

- By Ruchika Pande

# PROBLEM STATEMENT

- Lending Club is a financial institution that focuses on providing urban customers with different kinds of loans. The bank's choice to accept or refuse the loan entails two different kinds of risks:

  The bank's choice is subject to two different kinds of risks:

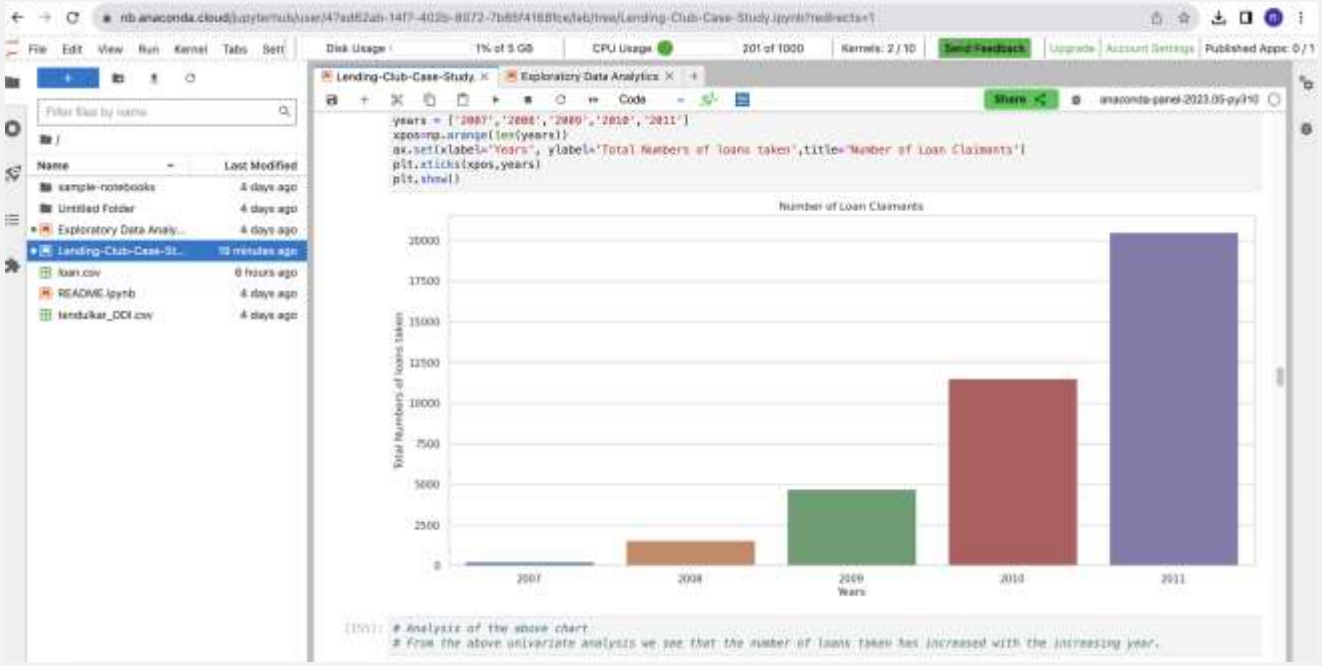
- If the borrower is likely to repay the loan, refusing to grant it results in the company losing business.
- If the borrower is unlikely to pay back the loan, or is likely to default, authorising the loan could result in a loss for the business.

# OBJECTIVE

- Utilise the data offered to pinpoint the causes (or cause variables) of loan default so that the business can use this information to assess risk and manage its portfolio.

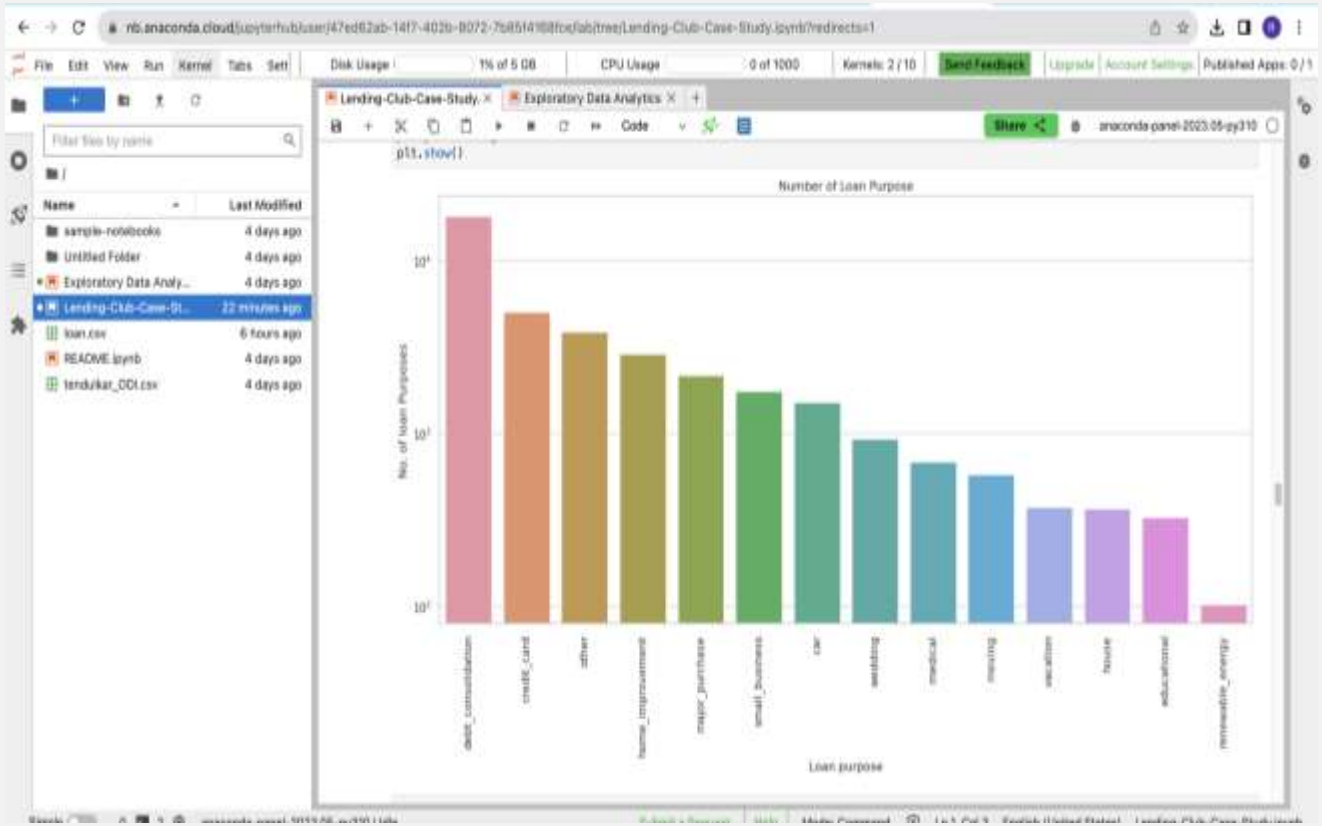
# DATA CLEANING

- Columns with more than 50% of the data missing were eliminated.
- Because they don't factor into the decision-making process for loan_status, which must be done before granting the loan, columns like customer behaviour factors like last_pymnt_amnt etc. were excluded from our analysis.
- Entries in columns with a single unique entry were removed.
- Additional columns that did not contribute to our study or columns that were duplicated in other columns were also eliminated.
- The relevant values from the strings were extracted using manipulations on columns like term, int_rate, and revol_util.
- Eliminating rows with loan_status 'Current' as our goal is to find patterns that show whether a person is going to default or not.



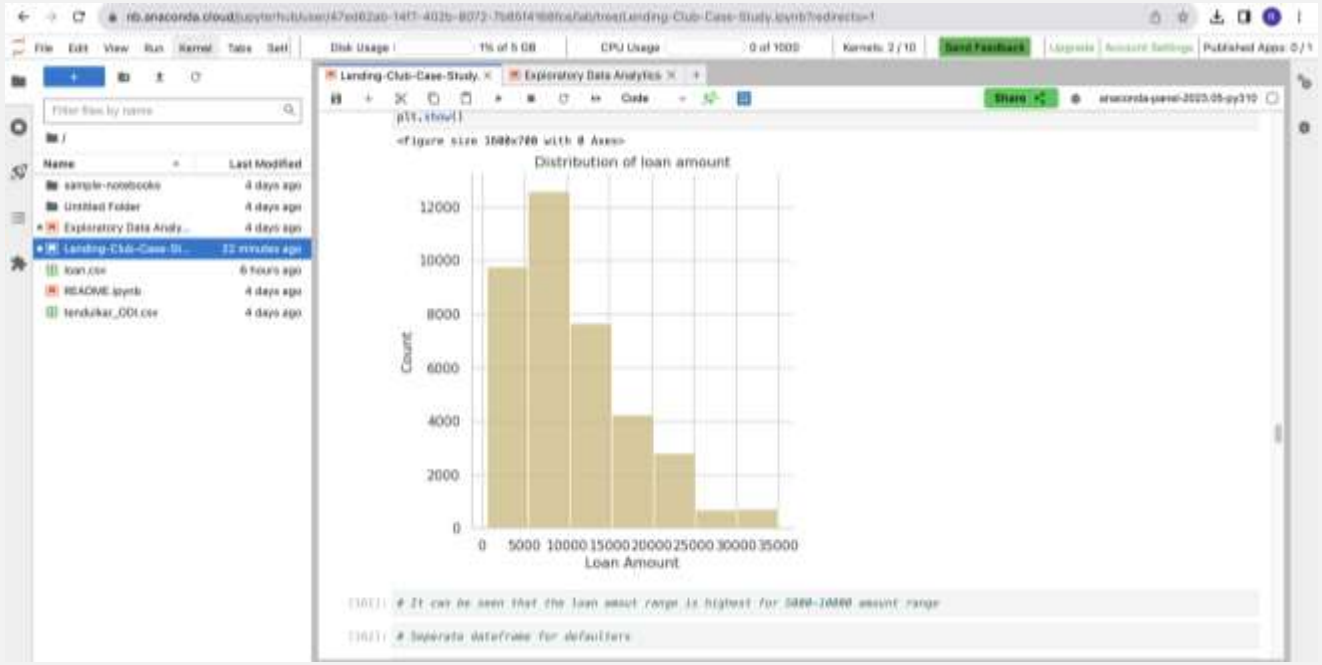

# OBSERVATION 1

- Total number of loans increased with the increasing year.



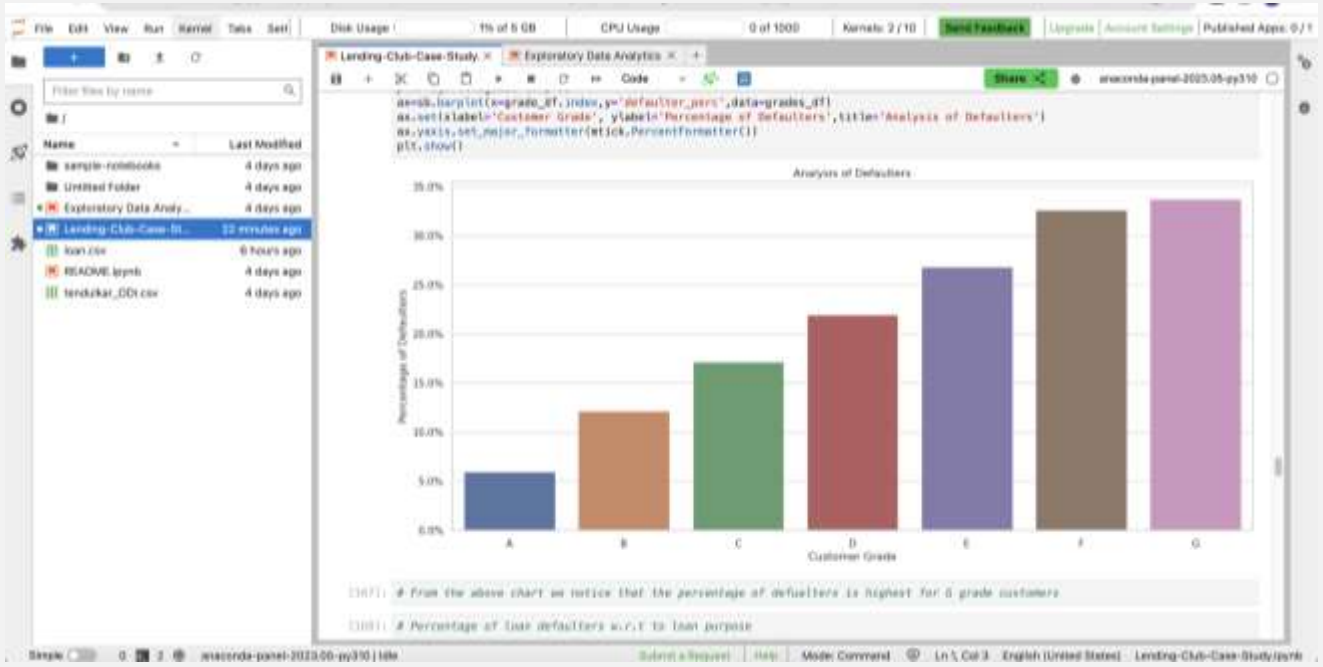

# OBSERVATION 2

- The top 3 loan purposes are 'debt_consolidation', 'credit_card' and 'other' category



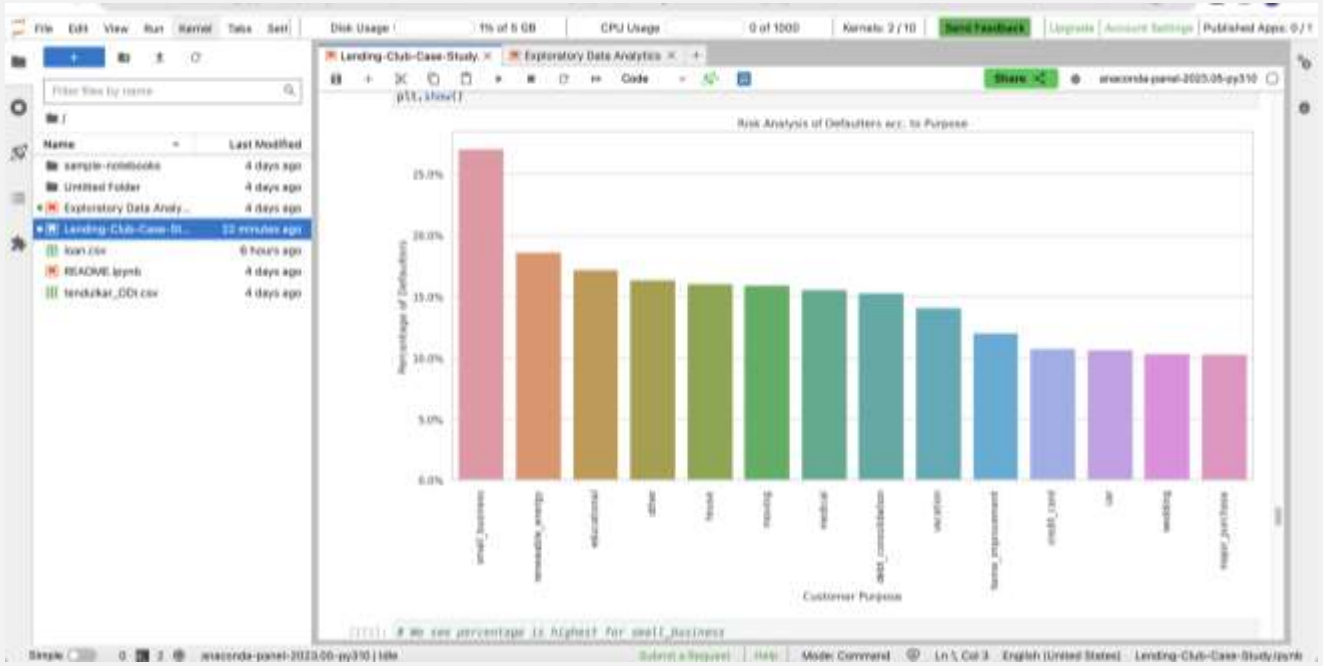

# OBSERVATION 3

- The loan amout range is highest for 5000-10000 amount range



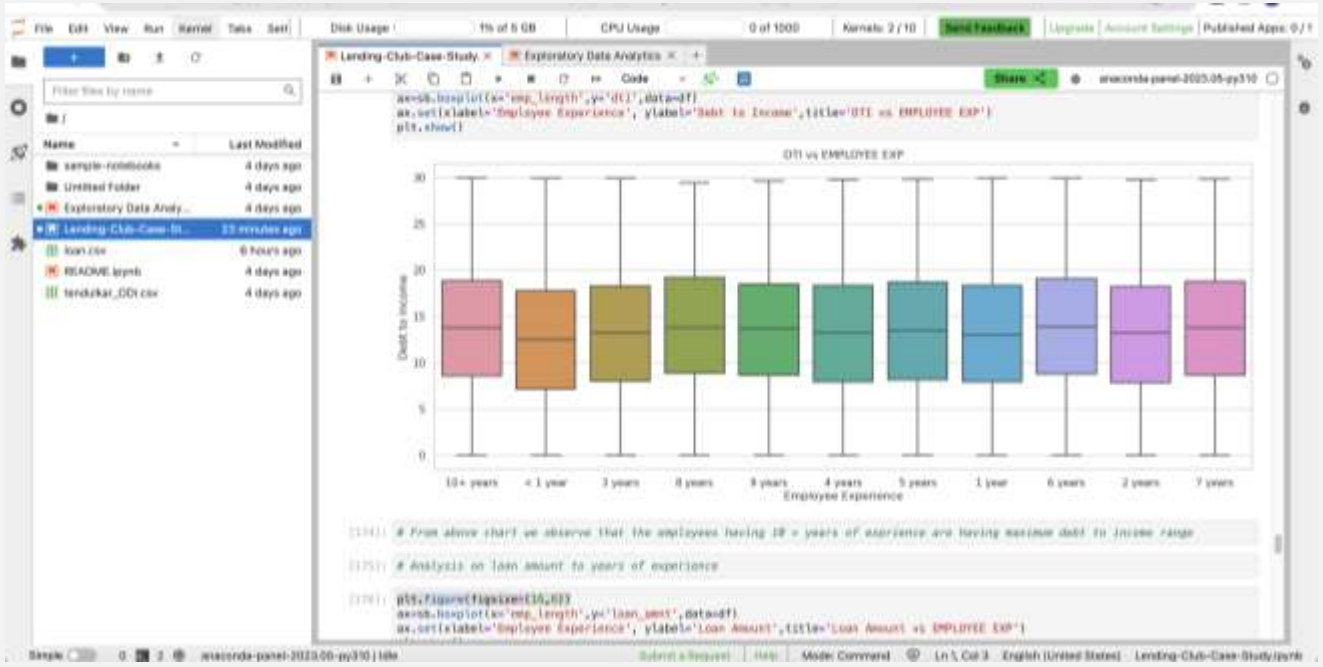

# OBSERVATION 4

- The percentage of defaulters is highest for G grade customers




# OBSERVATION 5

- We see percentage is highest for small_business.

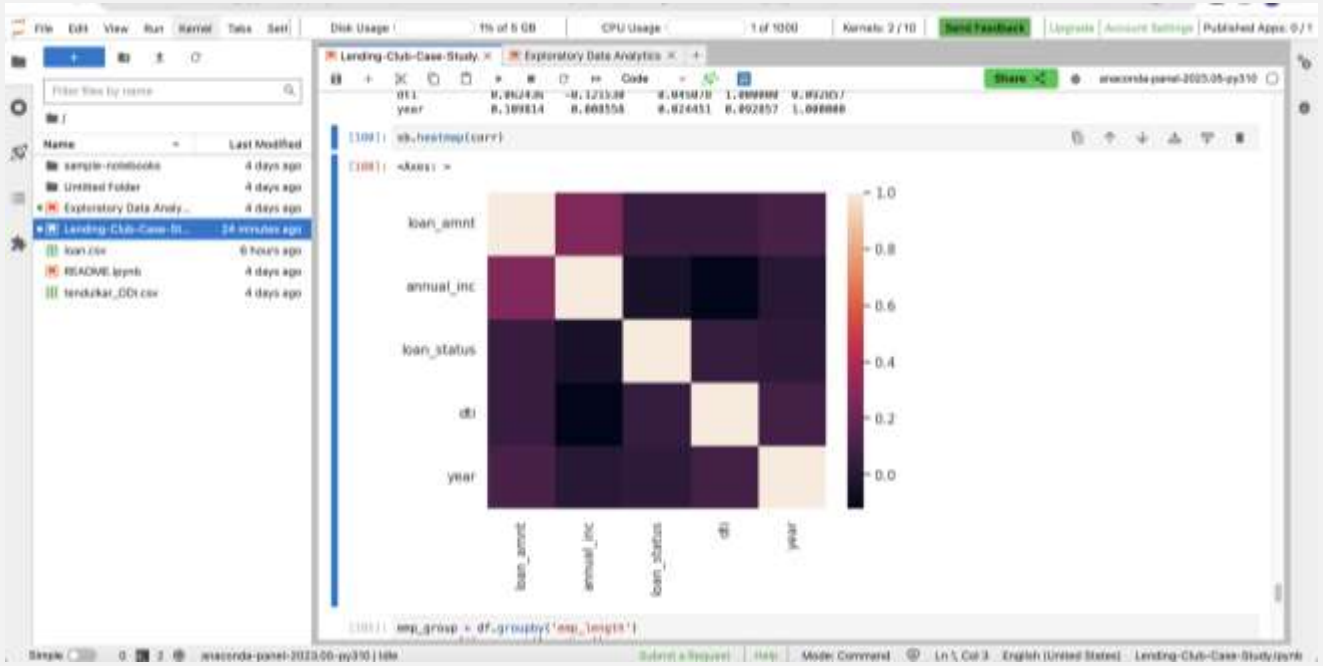



# OBSERVATION 6

- The employees having 10 + years of exprience are having maximum debt to income range


# OBSERVATION 7

- The employees having 10 + years experience are taking more amount of loan.




# OBSERVATION 8

- Correlation matrix for the dataframe which shows highly correlated variables and less correlated variables.

# THANK YOU!

- Ruchika Pande