



# CREDIT EDA



## EDA Key Steps

- Importing and Cleaning Data
- Univariate Analysis
- Bivariate Analysis
- Multivariate Analysis
- Data Transformation



# Problem Definition

- The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it as their advantage by becoming a defaulter.
- To Perform EDA to Analysis the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected.
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, if he/she is likely to default, then approving the loan may lead to a financial loss for the company.



## Steps

- Import Dataset
- Read Application CSV
- Data Inspection on Application dataset
- Data quality Check
- Impute Missing values
- Check / Validate the data types of the columns

A decorative background image on the left side of the slide, showing a rolled-up document. The document features a line chart at the top with two series: a blue line with circular markers and a black line with circular markers. The x-axis is labeled with months: Jun, Jul, Aug, Sep, Oct. A legend in the top left corner indicates '2017/18' for the blue line and '2016/17' for the black line. Below the line chart is a bar chart with five bars, each divided into two segments (blue and black). The x-axis is labeled 1, 2, 3, 4, 5. A legend to the right of the bar chart shows a blue square and a black square. At the bottom of the document, there is a caption 'Graph / Statistic' followed by some placeholder text: 'Sed ut eliam sit amet, consectetur adipiscing elit. Sed ut eliam sit amet, consectetur adipiscing elit. Sed ut eliam sit amet, consectetur adipiscing elit.' and 'according to 2017 survey'. At the very bottom, it says 'NEXT PAGE Analysis of h'.

## Steps

- Check / Validate the data types of the columns
- Binning of continuous variables
- Data Imbalance check
- Univariate Analysis
- Plot on Categorical columns
- Plot on Numerical columns



## Steps

- Bivariate & Multivariate Analysis
- Co-relation between Numerical Columns
- Read Previous Application CSV and repeat all the previous steps
- Merge the Application and Previous Application Dataframes



**LET' S GET CODING**