LENDING CLUB CASE STUDY

AI and ML

Group Member:
- Ankit Shah
- Anand Shah

TABLE OF CONTENTS

PROBLEM STATEMENT

Lending Club Case Study

The case study is prepared for a Financial Institution called "Lending Club" based in USA which provides loan to the end consumers. As it always happens in most lending institution, Many are good loans whereas some are the bad loans. Few people tends to default their loan and not make timely payment or no payment at all to the organization.

PROBLEM STATEMENT

Lending Club Case Study

This Case study is being prepared by putting own self on the shoes of the person who is responsible for the approval and disbursement of the loan to the consumer. When a consumers comes and asks for a loan, we need to analyze past data and decide accordingly whether the person will be willing to pay back the loan in the future.

PROBLEM STATEMENT

Lending Club Case Study

We should also not decline all the loan as it may cause the loss of business whereas accepting all the loans will increase probability of having a greater number of defaulters. To make the decision more accurate and precise, we will be analyzing almost 40000 rows and 111 columns of the data.

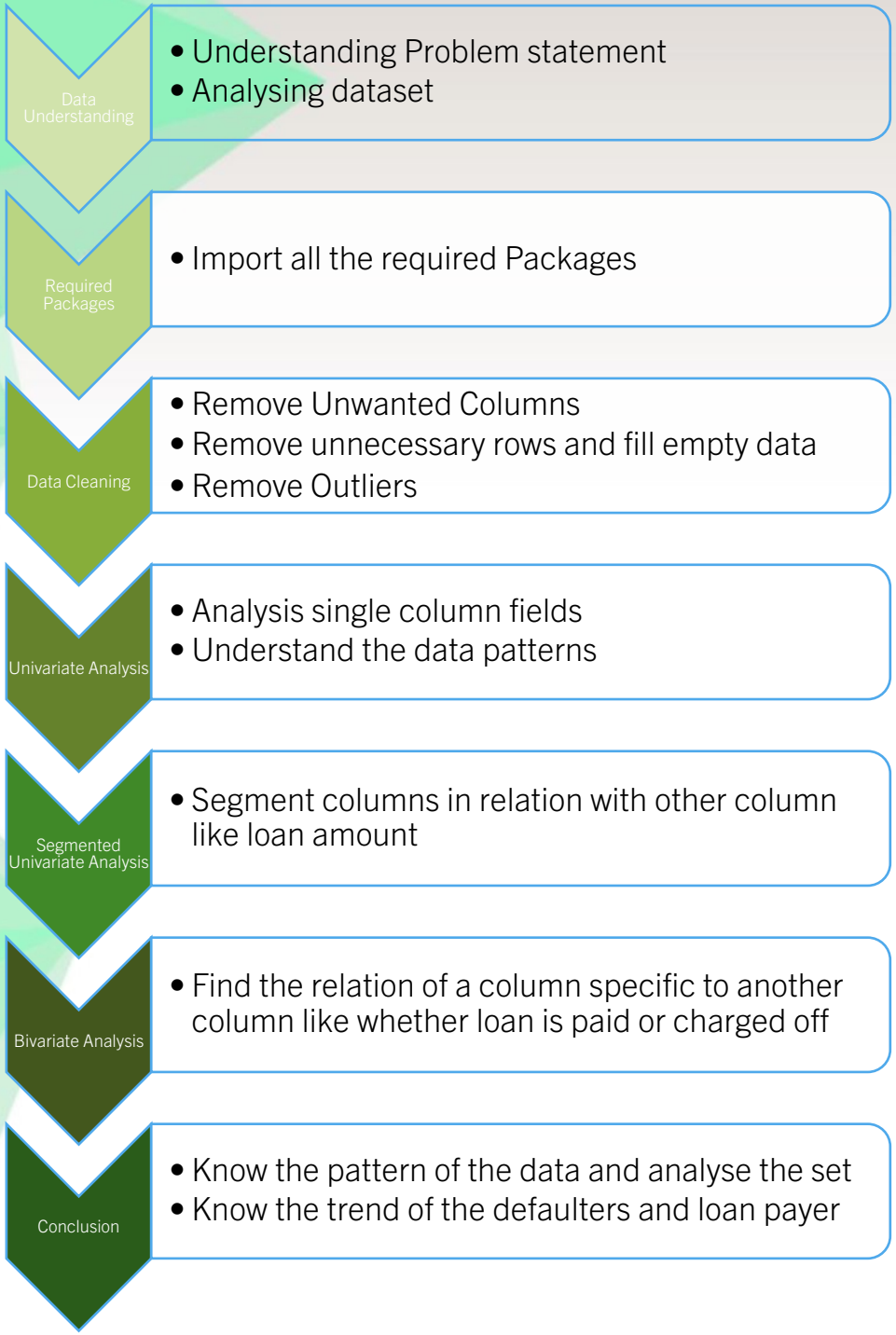
PROBLEM STATEMENT

Lending Club Case Study

All the rows and columns will not be required and may result in irrelevancy in the decision; thus, we will find out such data and remove them accordingly.

INTRODUCTION TO PROCESS INVOLVED

Lending Club Case Study



DATA CLEANING

Lending Club Case Study

Data cleaning is the first phase of data analysis, this phase involves cleaning data according to specific rules, criteria and requirements. It helps in reducing the number of columns and rows from a dataset which might not be necessary for analysis and create issue in reporting it

DATA CLEANING

Lending Club Case Study

BENEFITS OF DATA CLEANING

1. Removes anomalies & noise.
2. Improves quality of data
3. Removes outliers and unnecessary data

DATA CLEANING

Lending Club Case Study

Removes anomalies & noise

Cleaning data doesn't work in everyone's favor which means some rows in data set may come in special cases like special characters, color, wrong formatting and more which can't be included in data analytics projects.

DATA CLEANING

Lending Club Case Study

Improves Quality of Data

It is preferred to clean data with no missing values as business transactions are processed by standard SQL queries or a set of analysis program, without any manipulation of data.

UNIVARIATE ANALYSIS

Lending Club Case Study

Univariate Analysis

It is simplest form of analyzing data, as “Uni” means “one”, so we can say our data has only one variable. Its major task is to describe as it takes data, summarizes that data and finds patterns in the data.

SEGMENTED UNIVARIATE ANALYSIS

Lending Club Case Study

Segmented Univariate Analysis

Univariate analysis is very useful in many cases, the real strength lies here in Segmented Univariate Analysis.

Basic of Segmentation

SEGMENTED UNIVARIATE ANALYSIS

Lending Club Case Study

1. It takes raw data.
2. Group by dimensions.
3. Summaries using a relevant metric like mean, median, etc.
4. Compare the aggregated metric across groups/categories.

BIVARIATE ANALYSIS

Lending Club Case Study

Bivariate Analysis

It is used to find if there is relationship between two different variables. It is something as simple as creating a scatterplot by plotting one variable against another on a cartesian plane (think X and Y axis) can give us a picture. If data seems to fit a line on a curve, then there is correlation between two variables.