**Bank Loan Defaulters**

**Aim:**

 To give an idea of applying EDA in a real business scenario. In this case study, we will develop a basic understanding of business analytics in banking and financial services and understand how data is used to minimise the risk of losing money while lending to customers.

**Business Problem:**

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.

Our objective is to provide a complete analysis of the behaviour of the customers and understand who or which category/background people having what attributes has more defaulters.

**Learning Outcomes:**

1. Understand the KBRs and KPIs of a banking business
2. Cleaning data and feature selection.
3. Learn EDA through Python, to find all the relevant metrics that would help us to track our KPIs.
4. Dynamic Charts creation and Dashboarding through Tableau using various metrics to provide a clear analysis of the defaulters.

**Initial Skill Requirement:**

Proficiency in Python libraries: pandas, numpy, matplotlib and seaborn.

Along with that Tableau dynamic charts creation and Dashboarding skills.

**Data Dictionary:**

As the dataset contains 122 columns, please refer to the excel file containing the details of each column.

**Road map:**

Phase: 1

Initially we need to have a good understanding of the project requirement and domain knowledge about how banking data is processed, and the terms used for tracking its several metrices.

1. Step 1: Explanation of all the metrices and dimensions recorded in this dataset.
2. Step 2: What are business insights and how are they best detected.
3. Step 3: Understanding the data with the help of excel and then loading it into Python for further analysis and EDA
4. Creation of KBRs and KPIs and maintaining proper notes for it.
5. Step 4: Required Installations – Python, MS Excel, Tableau

Phase: 2

After planning and studying our business requirement, it’s time to perform further data cleaning and EDA in Python using various Python libraries.

1. Step 1: Load the dataset in python as pandas’ data frame.
2. Step 2: Perform statistical analysis for finding the insights about the features of the dataset.
3. Step 3: Using those insights to perform feature selection and imputation.
4. Step 4: Visualise the metrics that are related to our KPIs using matplot and seaborn library in python. To understand how they are related to our cause.

Phase 3:

In this phase our focus would be to visualize the data through Tableau as, dashboards are the end product that needs to be delivered to the clients as the most important part of our analytics services. All the insights related to banking defaulters business needs to be presented in the best way possible.

1. Step 1: Connect the data in Tableau.
2. Step 2: Create the required charts, according to the features and metrics selected during EDA.
3. Step 3: Add dynamic features to the charts to make it more insightful and user friendly with the help of parameters and calculated fields.
4. Step 4: Design a Dashboard that perfectly presents all the business insights with the dynamic chats created above in step 3.