**CREDIT RISK ANALYSIS PROJECT**

**Project Overview**

This project aims to analyse credit risk using the Universidad de Santiago de Chile credit risk dataset. The goal is to identify drivers of default risk among loan applicants, visualize patterns and trends, and provide recommendations to reduce the risk of default. The analysis is presented through a dashboard designed with Excel.

**Objectives**

1. **Default Risk Assessment:**  To uncover the pattern among loan applicants with a positive default. The default column is the target variable for the default risk assessment.
2. **Visualize Trends:** Create an intuitive dashboard to show default patterns across columns such as savings balance, checking balance, amount range, age bracket, loan duration and percent of income.
3. **Provide Recommendations:** Offer actionable solutions to minimize loss from credit defaults based on insights from the dataset.

**Stakeholders**

**Primary Stakeholders**

**Lending Institutions (Banks, Credit Unions, Financial Companies)**

**Role:** Primary providers of loans to applicants represented in the dataset.

**Interest:**

* Default risk assessment
* Loan approval policies
* Profitability

**Secondary Stakeholders**

**Loan Applicants (Individuals)**

**Role:** Individuals applying for loans, represented by the dataset’s rows.

**Interest:** Loan approval

**Dataset**

The dataset contains 1000 rows and 17 columns of loan applicant records.

**Data Transformation**

Data in columns such as age, loan duration and amount were applied in creating new columns to categorize the data provided in those columns.

**Tools and Technologies**

**Microsoft Excel:** Used for Data Preparation, Exploratory Data Analysis, and Data Visualization.

**Microsoft Word:** Used for Documentation and Report Findings.

**Breakdown of Analysis**

**Key Metrics**

**- Total Applicants: 1000** (Count of applicants)

**- Default Rate: 30%** (Count of default (NO)/ Total default multiplied by 100)

**Target Variable:** The default column is the target variable for the default risk assessment. With a 30% default (YES) against 70% default (NO). The default column enabled analysis of feature columns such as savings balance, checking balance, amount range, age bracket, loan duration and percent of income.

**Detailed Findings**

* **Default Risk by Savings Balance:** Out of the **1000** total applicants, **817** applicants had a verified savings balance. The analysis showed a lower default percentage for **savings balance between 500 – 100 DM** and **savings balance** **above 1000 DM**, with the former at **17%** default, and the latter at **12.5%**. **Savings balance less than 100 DM** and **savings balance between 100 – 500 DM** carried a greater percentage of default, standing at over **30%** for both category of savings balance.
* **Default Risk by Checking Balance:** The sum of applicants with a verified checking balance amounted to **606** out of the **1000** total applicants. **Checking balance below zero** revealed the highest default percentage, which is about **49.3%** default. This is followed by a decrease in default percentage of **39%** and **22.2%** for **checking balance between** **1 – 200 DM** and **checking balance above 200DM**.
* **Default Risk by Amount Range:** The **10K – 20K amount range** indicates a high number of defaulting applicants, out of the **40** applicants in that amount range, **24** applicants defaulted, which is **60%** of the **40** applicants. This contrast the other amount range of **5K – 10K**, **1K – 5K** and **250 – 1K**, where the **default (NO**) is greater than the **default (YES)**.
* **Default Risk by Loan Duration:** The **2 -5 years loan duration** has the highest default percentage, out of the **230** applicants in the loan duration, **44.35%** of the applicants defaulted. The default percentage for the **1- 2 years loan duration** is **28.98%** of the **590** applicants. The lowest default percentage of **15%** is the **less** **than 1 year loan duration**, which has **180** applicants.
* **Default Risk by Percent of Income:** Applicants with **4% income to interest** **payment** has **33.40%** rate of default. The **income to interest payment of 1%, 2% and 3%** has a default rate between **25%** - **27%**.
* **Default Risk by House Ownership: Owners** have a low default percentage of **26.01%** out of the **713** applicants. The applicants in the **rent** and **othe**r have a default percentage of **39.11%** and **40.74%** respectively.
* **Default Risk by Age Bracket:** The **age bracket of 18 – 30** presents a high default percentage of **36.01%**, this contrast the **age bracket of 30 – 55** and **age bracket of 55-75**, which stands at **25.68%** and **26.76%**.

**Recommendations**

1. **Set a minimum balance requirement for loan eligibility:** Applicants with a higher savings and checking balance had a low percentage of default compared to applicants with less savings and checking balance. Filter out financially unstable applicants, especially those with negative savings and checking balance.
2. **Cap loan amounts and durations:** For high risk applicants, especially applicants with low savings and checking balance, the amount range of 10K – 20K should be made ineligible. A period of less than 24 months should be the enforced as the benchmark duration for loans.
3. **Require additional safeguard for younger applicants:** Applicants between 18 – 25 years have the highest default rate (36.01%) compared to other age bracket at around 26% default rate. Younger applicants may lack financial stability or experience. Co-signers or collateral should be a requirement to help mitigate risk and the loans should be set for smaller amounts.

**Limitations**

The savings balance and checking balance columns has a large number of “unknown” as values. In my analysis of default risk by saving and checking balance columns, I ignored all unknown values and worked only with applicants with verified balance for both the savings and checking balance.