Statistical Testing Results – Bank Loan Analysis

Executive Summary Report

Prepared by Robes Fokoueng – April 2024

- **Project Overview:** In this part of the project, we will conduct a hypothesis test to analyze the relationship between loan_status (Good Loans/Bad Loans) and the Debt_to-Income (DTI).

- **Target Goal:** Develop a two-sample hypothesis test to analyze and determine whether there is a statistically significant difference in the average Debt-to-Income between loan applications(Customers) classified as Good loans and loan applications(Customers) categorized as Bad loans.

Key Insights

The p-value is significantly smaller than the significance level of 5%.

Thus, there is a statistically significant difference in the average Debt-to-Income between loan applications (Customers) classified as Good loans and loan applications (Customers) categorized as Bad loans.

As a result, these findings suggest there might be fundamental behavioral differences between these two groups of customers: Good borrowers and bad borrowers.

Details

We considered the relationship between Debt-to-Income and Loan_status.

One approach conducted was to examine the mean values of Debt-to-Income for each group of Loan_status (Good Loans [0]– Bad Loans[1]) in the data.

```
loan_status
0    0.132188
1    0.140047
Name: dti, dtype: float64
```

It appears that customers with Good loan status are those who have smaller DTI.

Next Steps

We suggest moving forward and building a **regression model** on Loan_status.

A regression model for Loan_status can help analyze customers behavior in this group of bad loans (Bad borrowers).