Loan Analysis Report

Date:- 11/05/2023

Author:- Shuprotik Roy

Objective:- The report is used to summarize the insights from a statistical analysis conducted on loan applicants data by Dream Housing Finance.

1. Males v/s Females(Loan Amount Acquired Comparison)
   To ascertain this, a hypothesis test was conducted. Since the sample sizes were different, Welch T-test was conducted.
   - Stating Null And Alternative Hypothesis:-
     H0: Males Loan amount<=Females
     H1: Males Loan Amount>Females
     (Critical Value is assumed to be 0.05)

   - Test Results:-
     P value is calculated to be 0.040386239437698215
     As p value<0.05 we can reject the null hypothesis and conclude that Males acquire more loan amount than Females.

   - Visualizing The Data
     We visualized the above findings with the help of Bar graph to better understand the insights with the help of a visual medium.

## Loan Amount Acquired By Males Vs Females
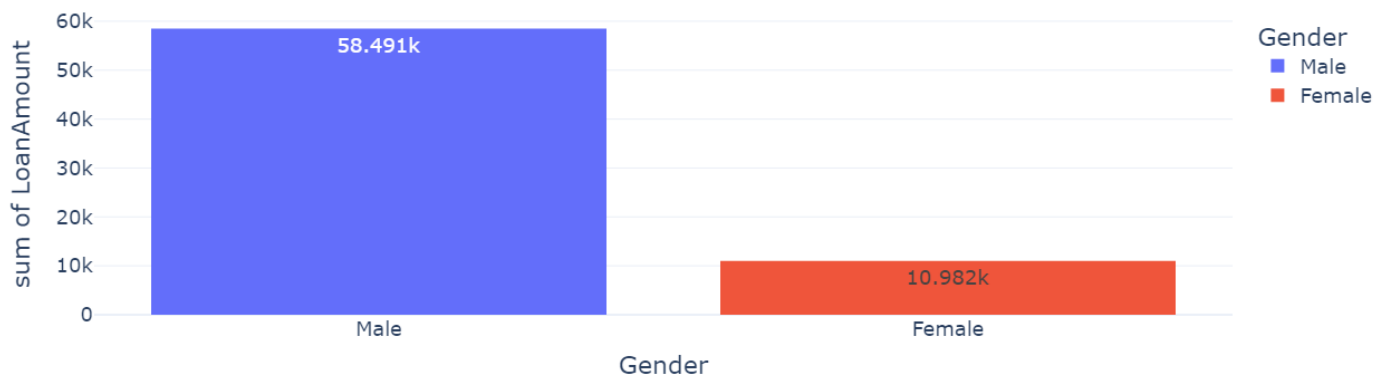
- Conclusion

    Males do acquire more loan amount than females on
    a total but this can also be due to the fact that
    there are significantly more male applicants than
    female applicants.

    Male applicants are 394 vs Female applicants are
    merely 86 i.e. almost 4x lesser than males.

2. Graduates v/s Non Graduates(Loan Approval Chances
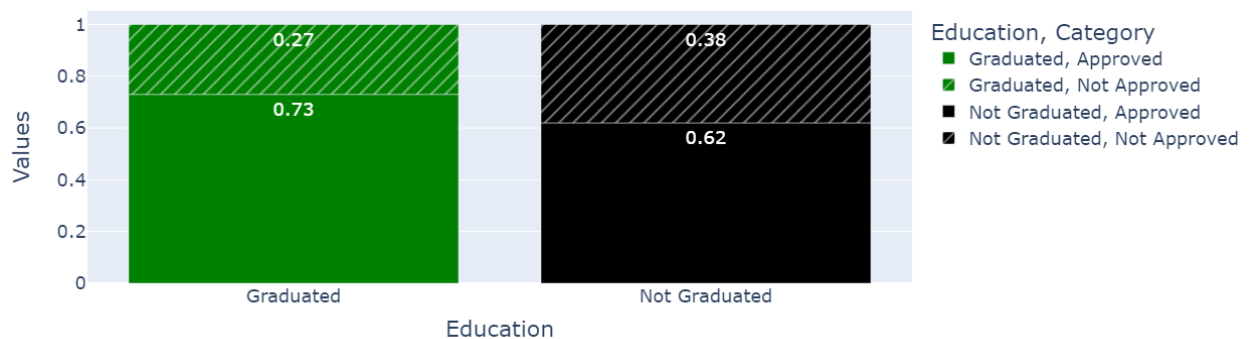   Comparison)
    - Comparison on a whole

    Graduates have higher loan approval chances(70.76%)

    than Non Graduates(62.89)
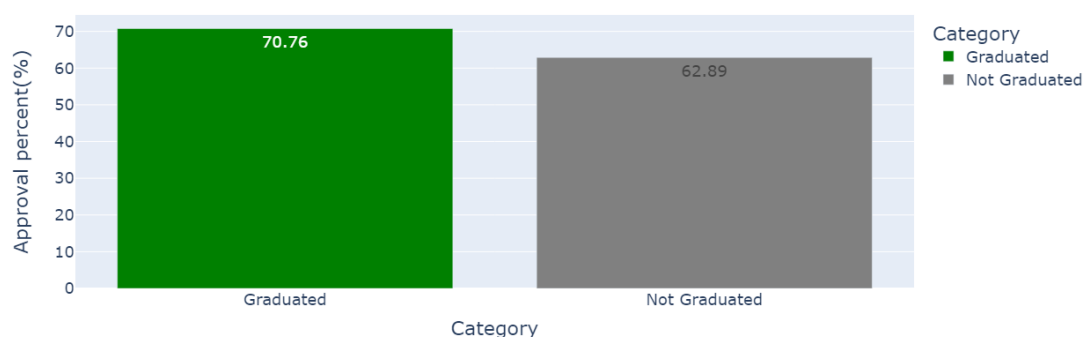
- Comparison(Genderwise)(Males)
  - Graduated Males have a 0.73 i.e. 73% chance of their loans getting which is 11% more than Non Graduates chances of getting loan approved which is 0.62 i.e. 62%.

  - Non Graduated Males have a 0.38 i.e. 38% chances of their loan not being approved which is significantly higher than chances of graduated males loan being disapproved which stands to 0.27 i.e. 27%.

Probability Chances Of Loan Approval Among Male Applicants

- Comparison(Genderwise)(Females)
  - Female Graduates suprisingly have a lower chances of loan getting approved of 0.62 i.e. 62% whereas Non Graduates have a higher chance of getting loan approved of 0.69 i.e. 69%.

  - Female Graduates face higher risk of their loans being disapproved of 0.38 i.e. 38% whereas Non Graduates have lower chances of loans being disapproved (0.31 i.e 31%).

Probability Chances Of Loan Approval Among Female Applicants
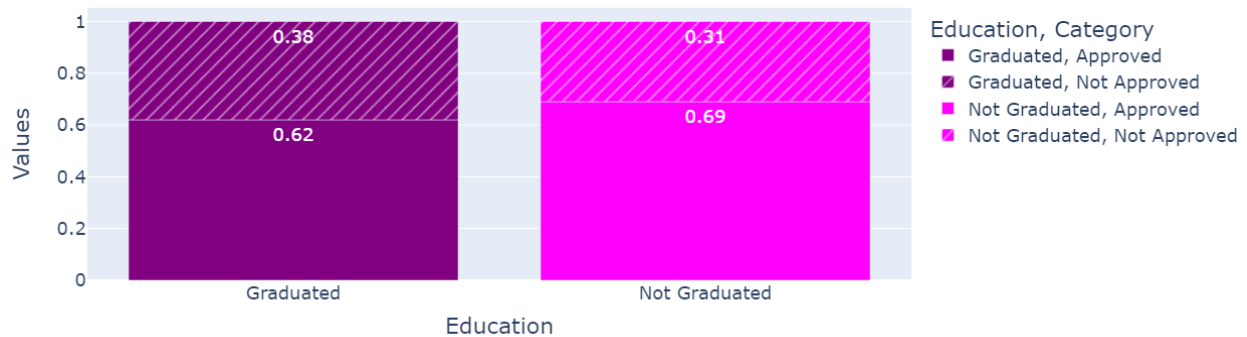
3. Applicant Income Distribution Analysis

To understand the income distribution of our applicants, we conducted a zscore analysis since our data was normally distributed.
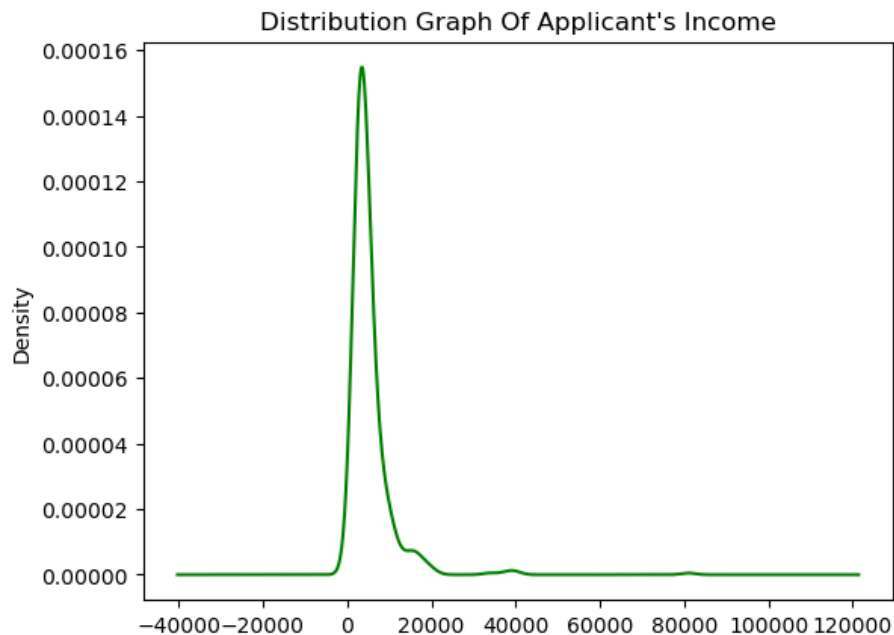
Although it's a positively skewed data, we can use zscore analysis to get an approximate percentile distribution.

Upon calculation of percentile from zscore, the applicant income could be divided in two broad categories i.e first being income

scale of majority applicants i.e. (50-90%) and minority applicants i.e (10-30%).

- Income scale/range for 50-90% applicants

- 50-90% applicants fall within the income range of 5391-12.5k with 6875 being the median income.

- Income scale/range of 10-30% applicants

- 10-30% applicants fall within the income range of 1299-2383 with 2031 being the median income

- Conclusion
- Loan approval facilities should have exclusive approval systems and support for applicants who fall in the income category of 5391-12.5k as this constitutes the majority of loan applicants of the bank.

4. Which Income Range has the highest/lowest Loan Approval/Rejection Chances?
    - (30-60%) of our total applicants have the highest Approval and highest rejection chances which constitutes an income range of 2400-6875.

    - (60-90%) of our total applicants have the lowest rejection chances and lowest approval chances which constitutes an income range of 6822-12.5k

    - (10-30%) of our total applicants have median approval and rejection chances constituting an income range of 1299-2526.

Loan Approval vs Rejection Chances (Income Distribution Wise)

Loan Approval Chances(in %)   Loan Rejection Chances(in %)
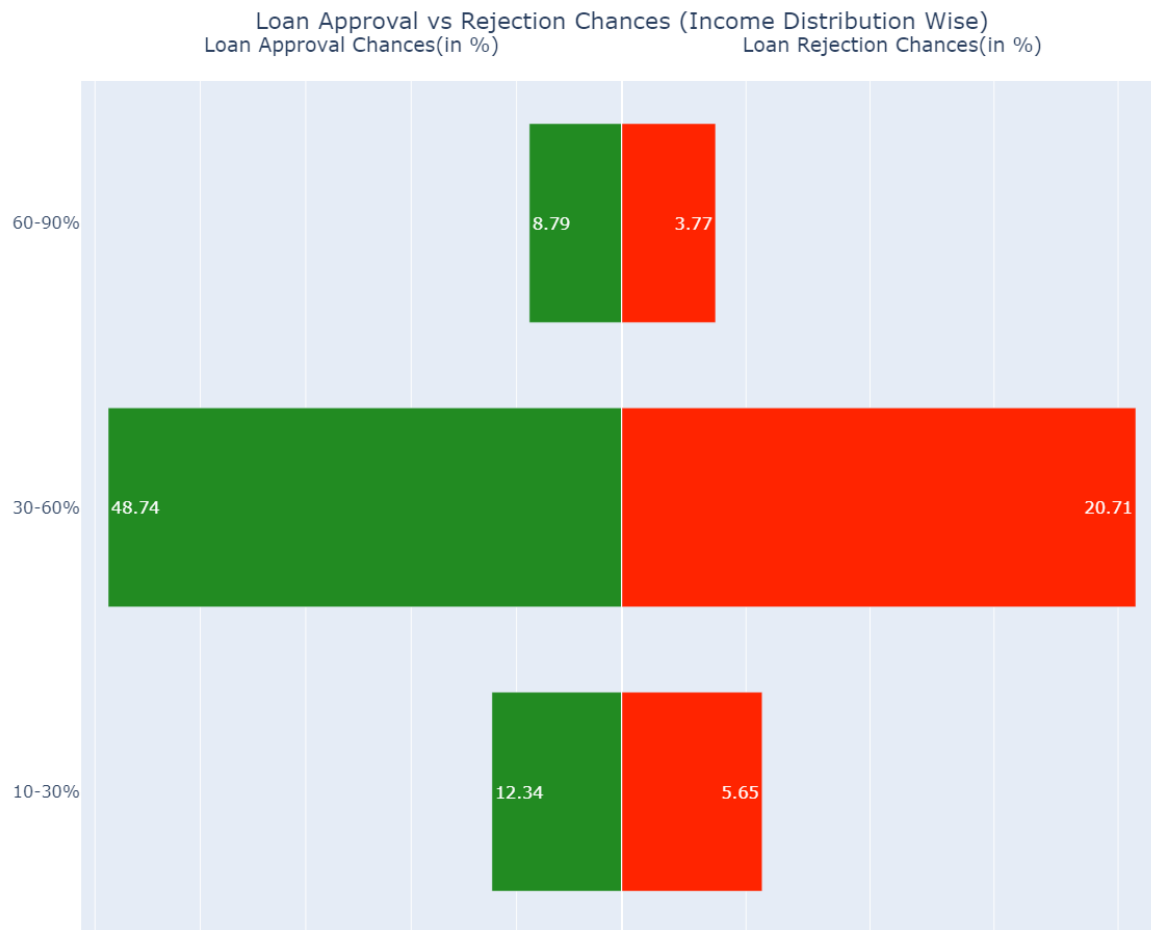
5. Loan Applicant's Location Distribution
    - 29% are from Rural Areas, 31.3% are from Urban Areas and 39.8% are from Semiurban Areas.
    - Semiurban Applicants and Areas need to be exclusively focused on as they constitute the majority of applicants
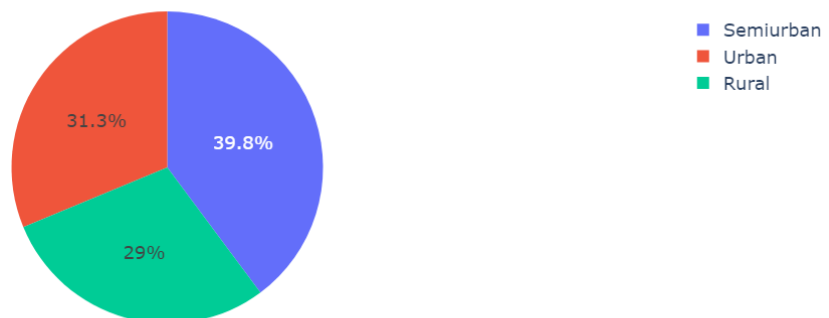
6. Loan Amount Distribution And Analysis

   - Z score analysis was conducted to better understand the Loan Amount Distribution.

   - Percentile was calculated on basis of the z scores and the calculated percentile was divided in two categories:-

     - >80% (Highest Range Of Loan Amount Applied)

     - <10% (Least Range Of Loan Amount Applied)

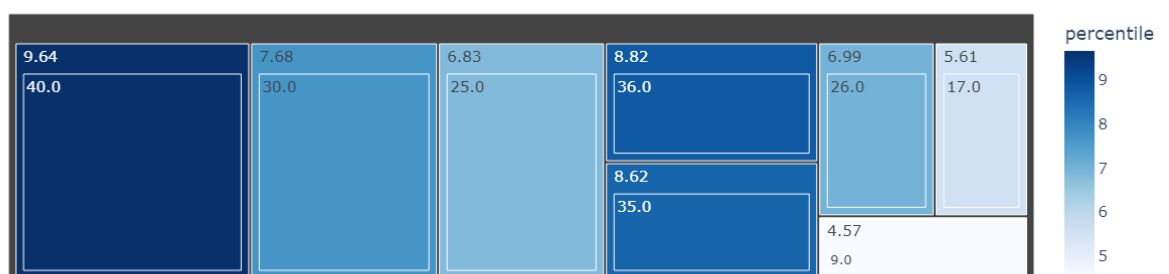   • >80% Distribution

Percentile Distribution for >80(%) Loan Amount

   - Mini boxes contain the Loan Amount.
   - Labels contain the percentile numbers.
   - 480-600 is the maximum loan amount applied.
   - The highest range of loan amount applied is 214-600.

   • <10% Distribution

Percentile Distribution for <10(%) Loan Amount

- Mini boxes contain the Loan Amount.
- Labels contain the percentile numbers,
- 9.0 is the minimum loan amount applied.
- The least range of loan amount applied is 9-40.

7. Dependent v/s Non-Dependent Analysis

-The given dataset was split in dependents and non dependents to draw comparisons between the both.

- Do Dependents apply for more loan amount on an average when compared to average loan amount applied by all the loan applicants?

  Conducting a hypothesis test to get an answer:-
  - Null Hypothesis: Average of Loan Amount applied by dependents<=Average Loan Amount applied by total customers

  - Alternative Hypothesis: Average of Loan Amount applied by dependents>Average Loan Amount applied by total customers
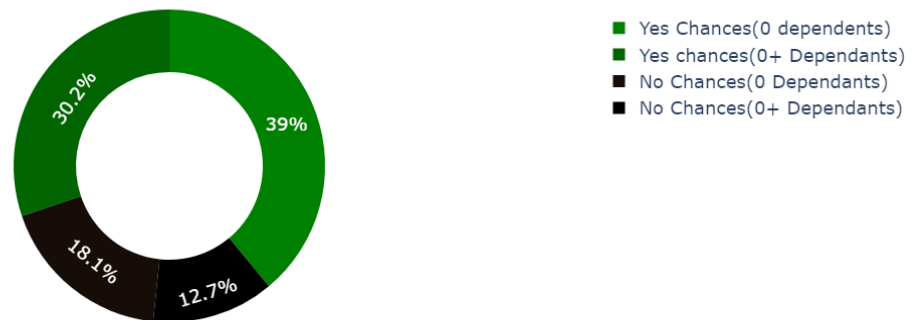
  - critical value is assumed to be 0.05 (1.645 in zscore)

  - Upon calculation, the zscore came out to be -2.95, as z score<critical value, alternative hypothesis can be rejected and we can conclude Average of Loan Amount Applied By Dependents<Average of Loan Amount Applied By Total Customers.

  - Rejection/Approval Chances For Dependants/Non Dependants

- Probability scores for dependants & Non dependants were calculated and converted to percentage to ascertain a probable estimate of rejection/acceptance rates

- Dependants(0+) have a lower risk of rejection than non dependants(0) but at the same time non dependants have a higher chance of approval

Legend:
- Yes Chances(0 dependents)
- Yes chances(0+ Dependants)
- No Chances(0 Dependants)
- No Chances(0+ Dependants)

Chart values: 39%, 30.2%, 18.1%, 12.7%

8. Do Non Dependents apply for lesser loan amount on an average when compared to average loan amount applied by all the loan applicants?

   Conducting a hypothesis test to get an answer
   - Null Hypothesis: Average of Loan Amount applied by Non Dependents>=Average Loan Amount applied by total customers

   - Alternative Hypothesis: Average of Loan Amount applied by Non Dependents<Average Loan Amount applied by total customers
   - critical value is assumed to be 0.05 (1.645 in zscore)
   - z score was calculated as 2.55.

   As z test score is greater than z score value therefore we can reject the null hypothesis and conclude that:-

- Average of Loan Amount Applied By Non Dependents<Average of Loan Amount Applied By Total Customers