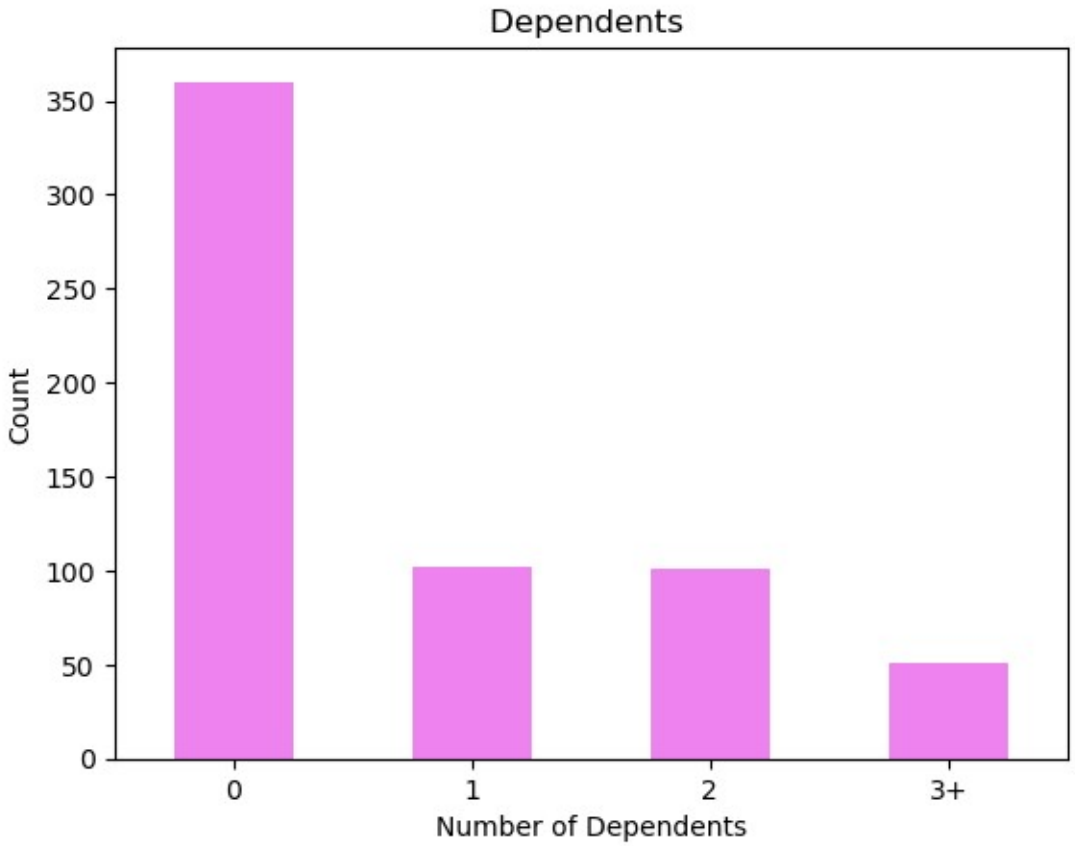
REPORT FOR LOAN DATASET

Introduction

This report dives into a dataset filled with info about people applying for loans. We're going to explore this data using different types of graphs. Our goal is to learn more about the data and tell a clear story using these graphs.

Graph 1: Dependents Count (Bar Chart)

- **Purpose**: This chart helps us see how many people have different numbers of dependents (like family members they're responsible for).
- **Data**: We're looking at categories like '0', '1', '2', or '3+' dependents.
- **Reason**: This chart is like a picture that shows us how many people fall into each category. It helps us understand the family situations of loan applicants.

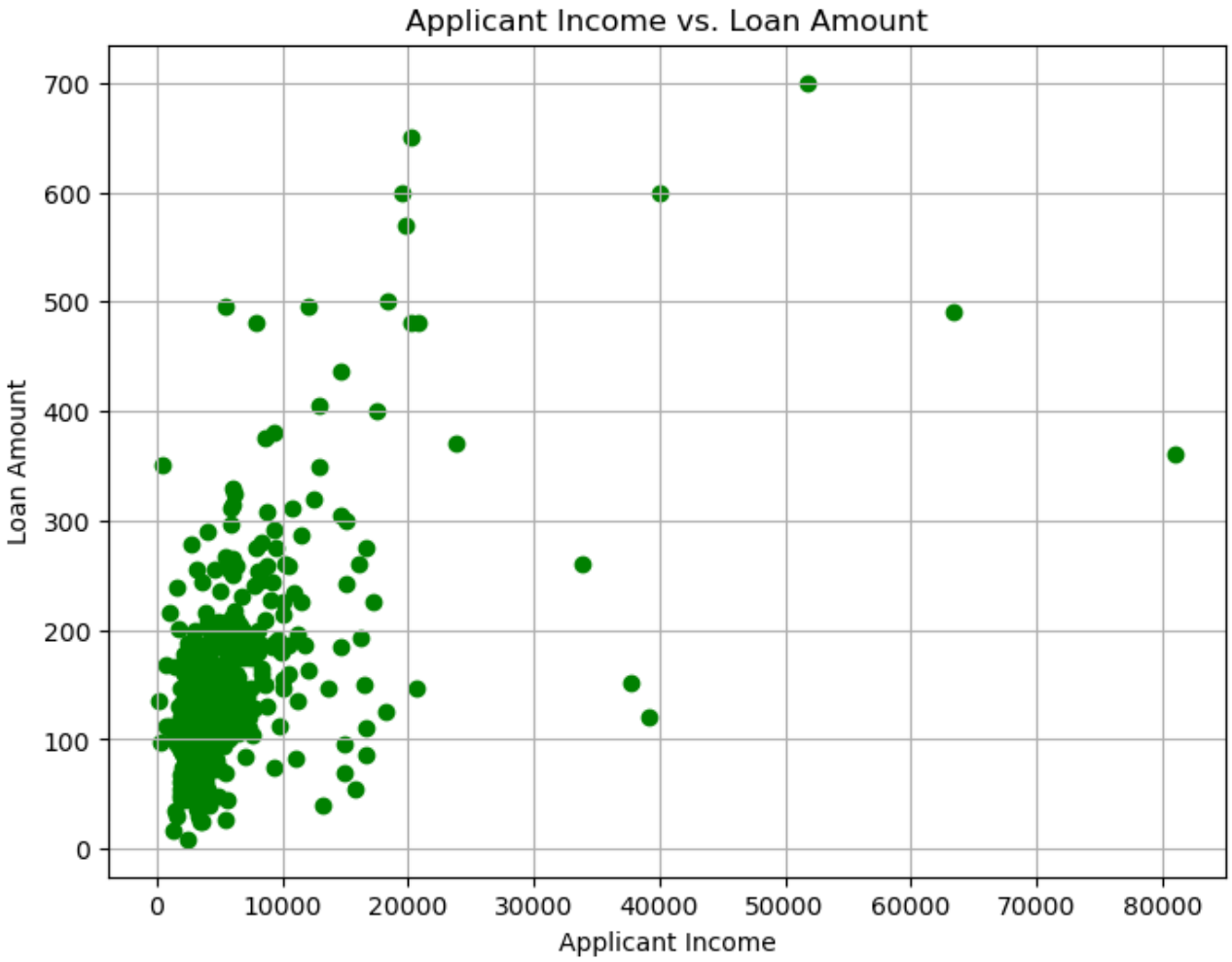


Graph 3: Scatter Plot (Applicant Income vs. Loan Amount)

- **Purpose**: This scatter plot helps us see if there's a connection between how much money an applicant earns and how much they're asking to borrow.
- **Data**: We're comparing two continuous things: how much someone earns (income) and how much they want to borrow (loan amount).
- **Reason**: It's like a map of dots that helps us spot trends. We can figure out if people with higher incomes tend to ask for bigger loans.

Observations:

- **Positive Correlation**: As applicant income increases, there's a trend of higher requested loan amounts, suggesting that individuals with higher incomes often seek larger loans.
- **Scattered Variation**: While the overall trend is positive, there's variability among the data points, indicating that factors beyond income influence loan requests.
- **Concentration in Lower Range**: Many applicants request smaller loan amounts, possibly for immediate needs aligned with their current income.

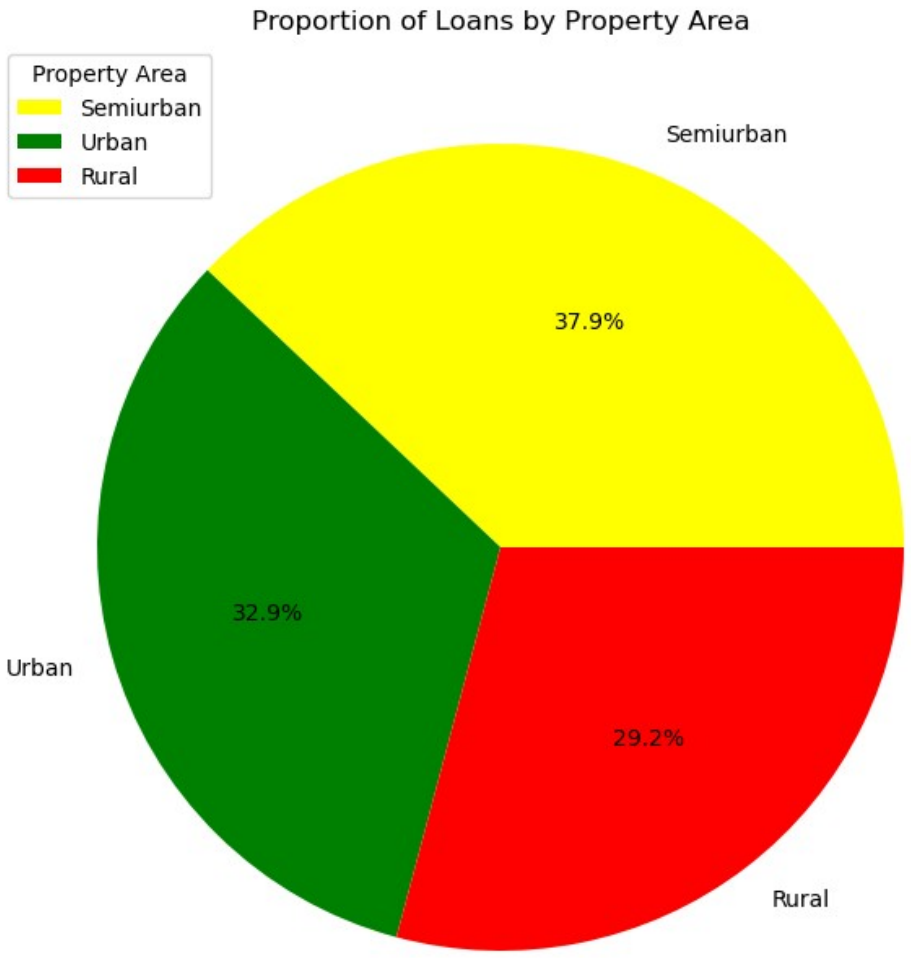


Graph 4: Pie chart (Proportion of Loans by Property Area)

- **Purpose**: The pie chart illustrates the proportion of loans attributed to each property area, offering a clear and concise snapshot of their relative contributions to the overall loan dataset.
- **Data**: The data consists of categorical values indicating property areas, and the pie chart displays the percentage distribution of loans within each property area.
- **Reason**: A pie chart is a great choice when we want to understand the share of each category within a whole. The use of colors, percentages, and the circular format makes it easy to visualize the relative significance of property areas in the loan dataset.

Observations:

- We notice that the "Urban" property area, represented in yellow, accounts for the largest share of loans, suggesting a higher demand for loans in urban locales.
- The "Semiurban" property area, depicted in green, follows closely, indicating a significant proportion of loans are associated with semiurban regions.
- The "Rural" property area, shown in red, has the smallest share of loans but still contributes to the overall loan distribution.

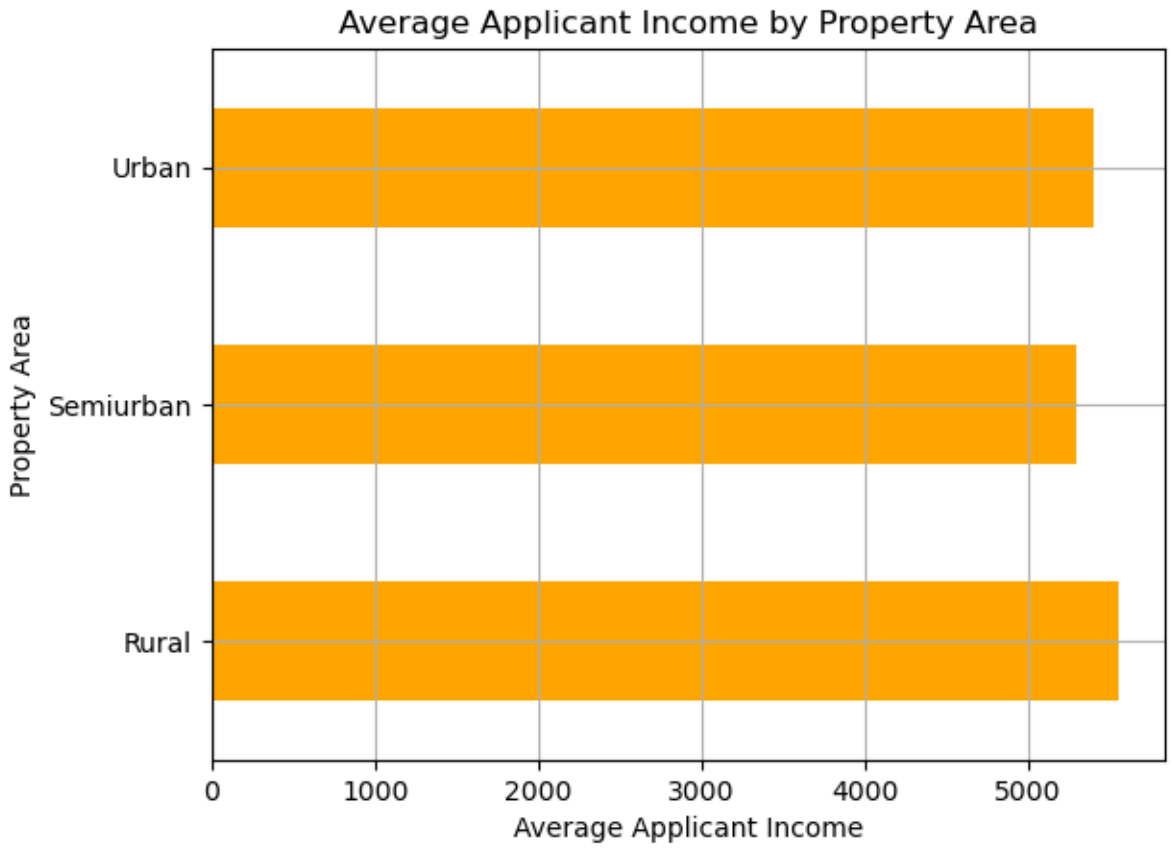


Graph 5: Horizontal Bar Chart (Average Applicant Income by Property Area)

- **Purpose**: The horizontal bar chart displays the average income of loan applicants in each property area, allowing us to directly compare income levels among the property areas.
- **Data**: The data includes the average income for each property area category, and the horizontal bar chart emphasizes the differences in income.
- **Reason**: A horizontal bar chart is well-suited for comparing values across categories. The use of bars that extend horizontally makes it easy to compare income levels in each property area.

Observations:

- We observe that the "Semiurban" property area has the highest average income among loan applicants. This suggests that individuals from semiurban areas, on average, tend to have a higher income compared to other property areas.
- The "Urban" property area follows, with a moderately high average income, indicating that urban regions also have relatively well-off loan applicants.
- The "Rural" property area, while having the lowest average income among the three, still contributes to the overall applicant pool.

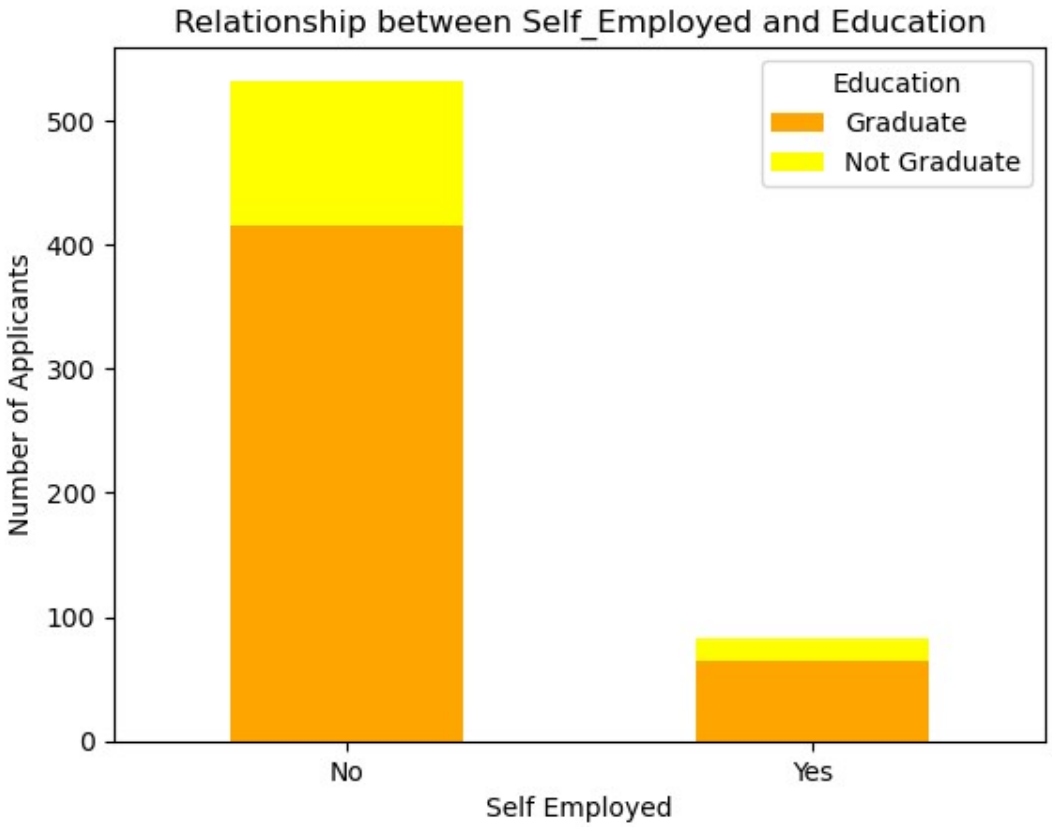


Graph 5: Stacked Bar Chart (Relationship between Self_Employed and Education)

- **Purpose**: This stacked bar chart illustrates the count of loan applicants by their self-employment status, further segmented by their education level.
- **Data**: The data consists of categorical variables, with 'Self_Employed' and 'Education' serving as the two key factors. The chart visually demonstrates the composition of loan applicants within each combination of these factors.
- **Reason**: Stacked bar charts are ideal for showcasing the distribution of a categorical variable within multiple categories. The use of color and the stacked format enables us to analyze the relative proportions of applicants within each education level for both self-employed and non-self-employed groups.

Observations:

- The stacked bar chart provides a clear picture of the distribution of loan applicants based on their self-employment status and education level.
- Among those who are **not self-employed** ('Not Self Employed'), the majority are **graduates**, indicating a higher proportion of educated applicants in this group.
- Among those who are **self-employed** ('Self Employed'), we observe a relatively balanced distribution of both **graduate** and **not graduate** applicants, suggesting that self-employment doesn't seem to be heavily skewed towards a particular education level.

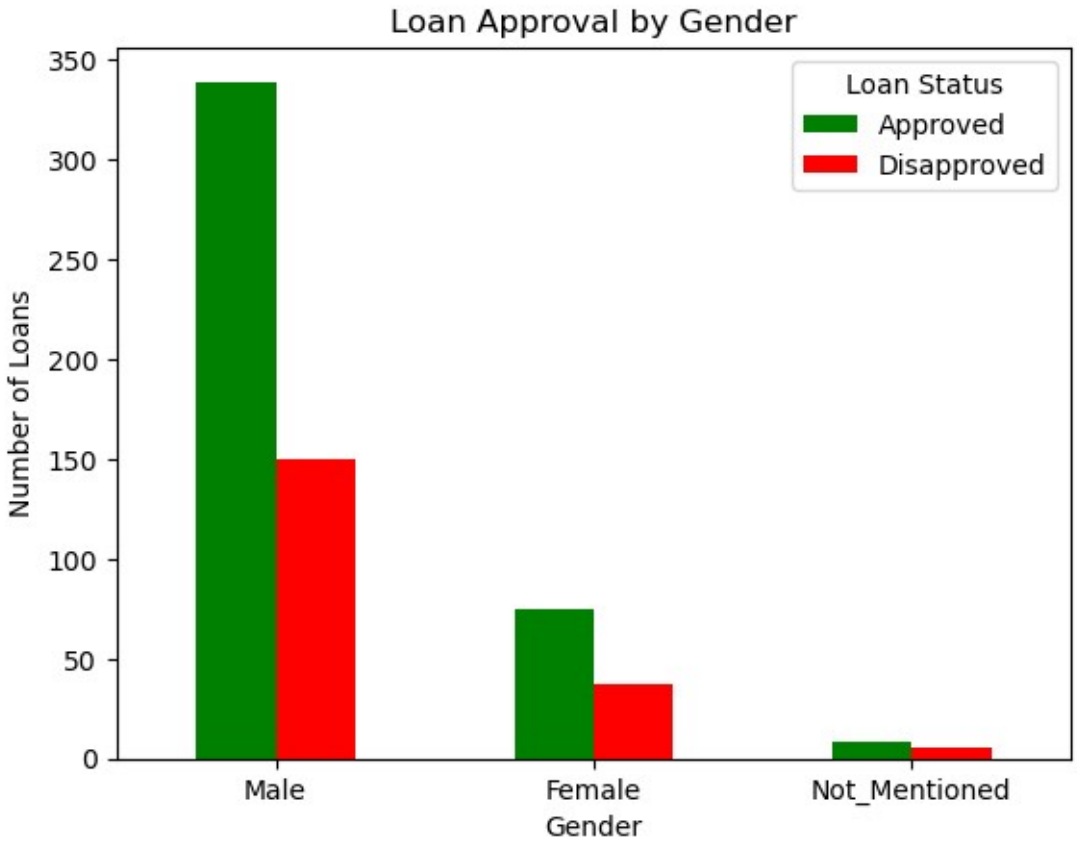


Graph 6 : Grouped Bar Chart (Loan Approval by Gender)

- **Purpose**: This grouped bar chart showcases the count of loans approved and disapproved, highlighting the gender-based distribution.
- **Data**: The data is structured in a way that tracks the loan status ('Approved' or 'Disapproved') for each gender category.
- **Reason**: Grouped bar charts are ideal for comparing the values of different categories side by side. The use of color and separate bars for each gender makes it easy to compare the loan approval outcomes for each gender.

Observations:

- The grouped bar chart vividly presents the loan approval outcome by gender.
- We observe that **males** have a higher count of approved loans compared to disapproved loans, indicating a favorable approval rate for male applicants.
- Similarly, **females** also have a higher count of approved loans, but the difference between approved and disapproved loans is relatively smaller compared to males.



Conclusion:

The

Conclusion:

comprehensive analysis of the loan dataset, visualized through diverse graphs, provides invaluable insights. We observed that most loan applicants have fewer dependents ('0' or '1') and higher income individuals tend to request larger loan amounts. The proportion of dependents and self-employment status varies across education levels, suggesting nuanced applicant profiles. Both genders have higher counts of approved loans, while semiurban areas display the highest average income among applicants. These findings enable better-informed lending decisions, highlighting opportunities for tailored loan products and fairer lending practices.