Loan Data Report:

1. Introduction:

Dataset Overview:

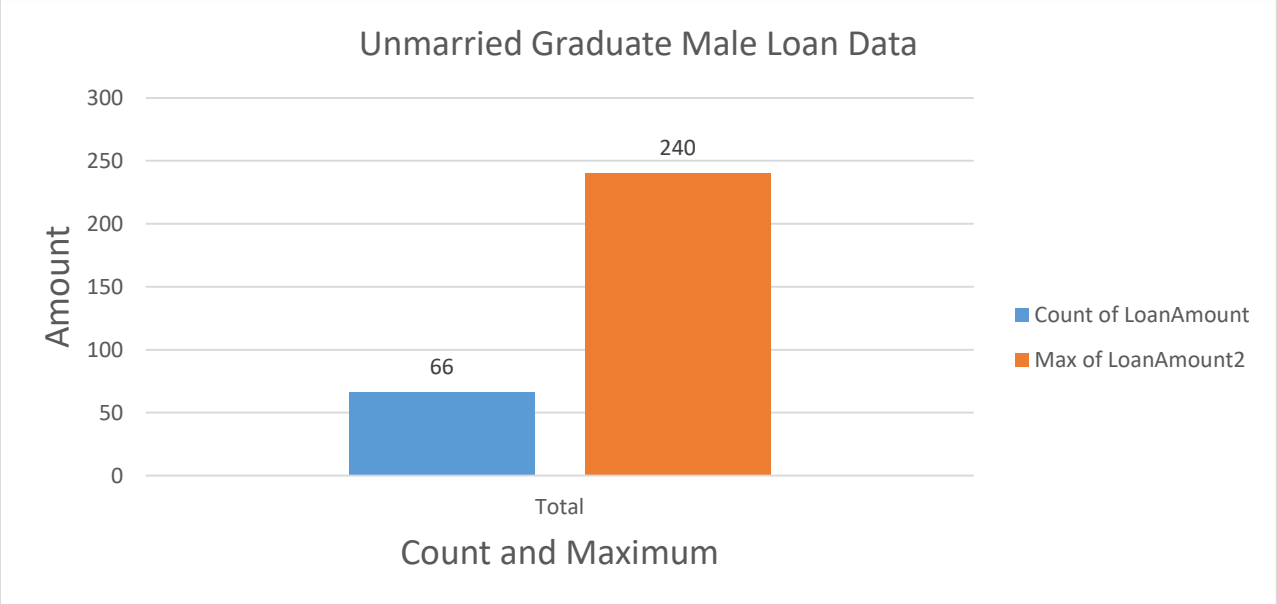
Our dataset encompasses a diverse range of variables, each shedding light on the intricate dynamics of loan applications. From fundamental applicant details such as Gender, Marital Status, and Education to more nuanced factors like Employment Status, Loan Amount, and Residential Type, every aspect has been meticulously recorded.

2. Questionnaire:

Q1. How many male graduates who are not married applied for Loan? What was the highest amount?
Q2. How many female graduates who are not married applied for Loan? What was the highest amount?
Q3. How many male non-graduates who are not married applied for Loan? What was the highest amount?
Q4. How many female graduates who are married applied for Loan? What was the highest amount?
Q5. How many male and female who are not married applied for Loan? Compare Urban, Semi-urban and rural on the basis of amount.

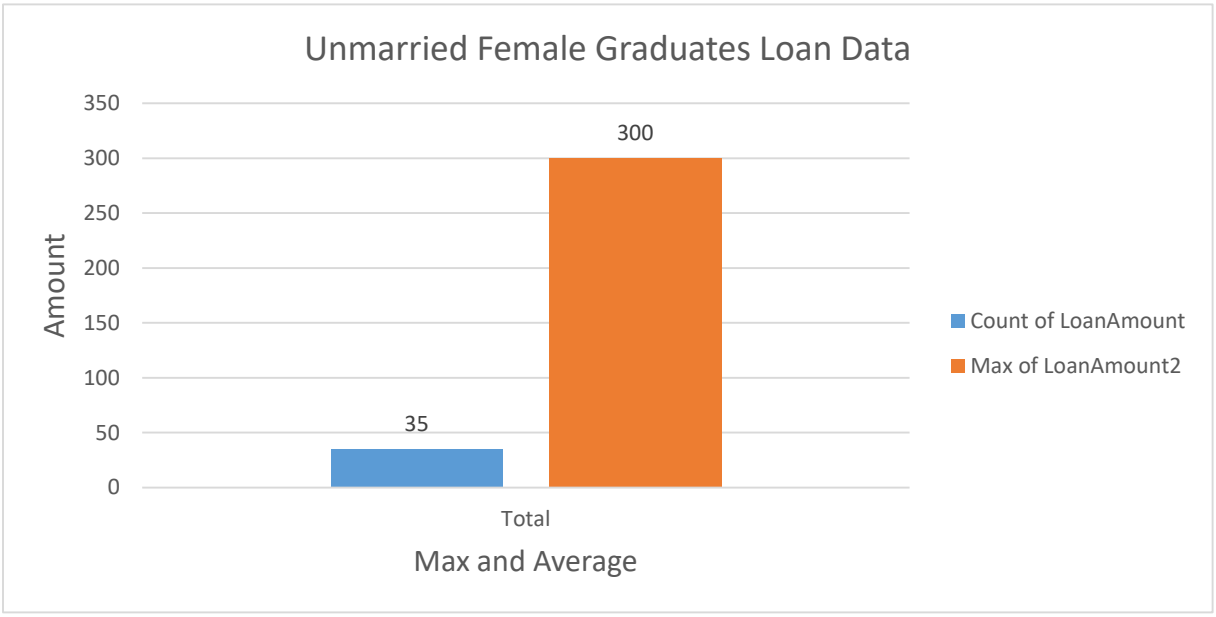
3. Analytics:

Q1. How many male graduates who are not married applied for Loan? What was the highest amount?



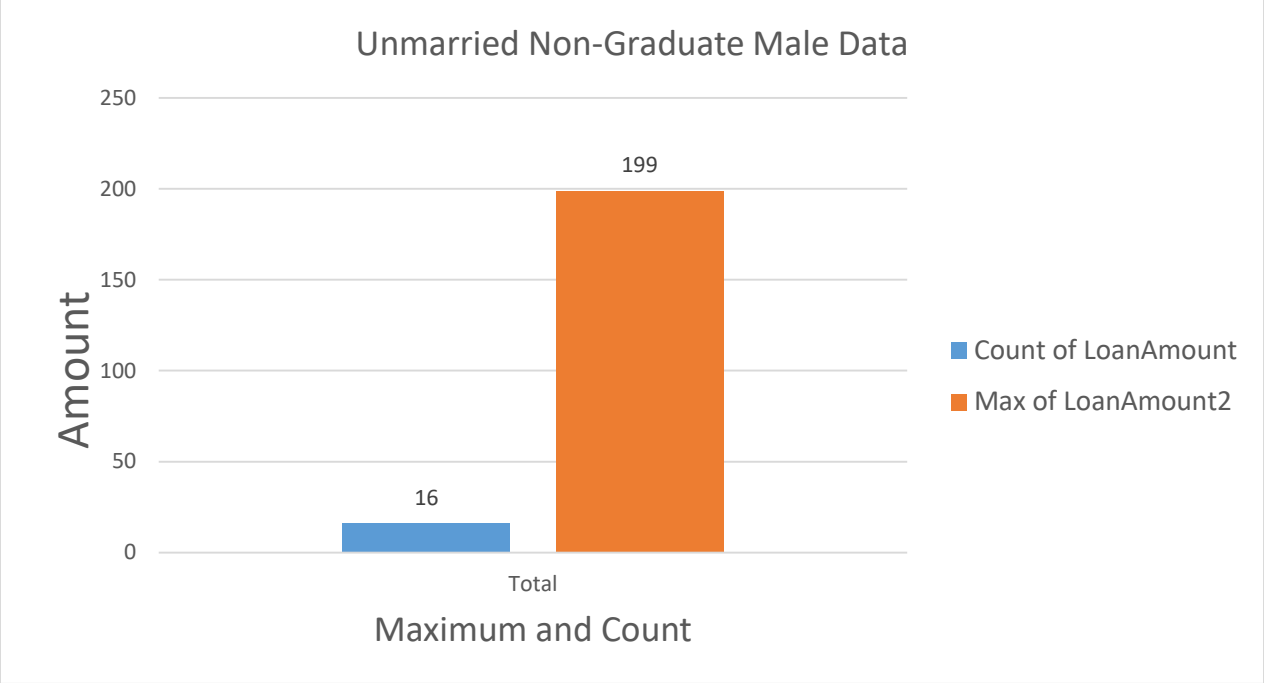
Answer:
240 Unmarried Graduate Males applied for the loan out of which the maximum amount of loan applied was 66.

Q2. How many female graduates who are not married applied for Loan? What was the highest amount?



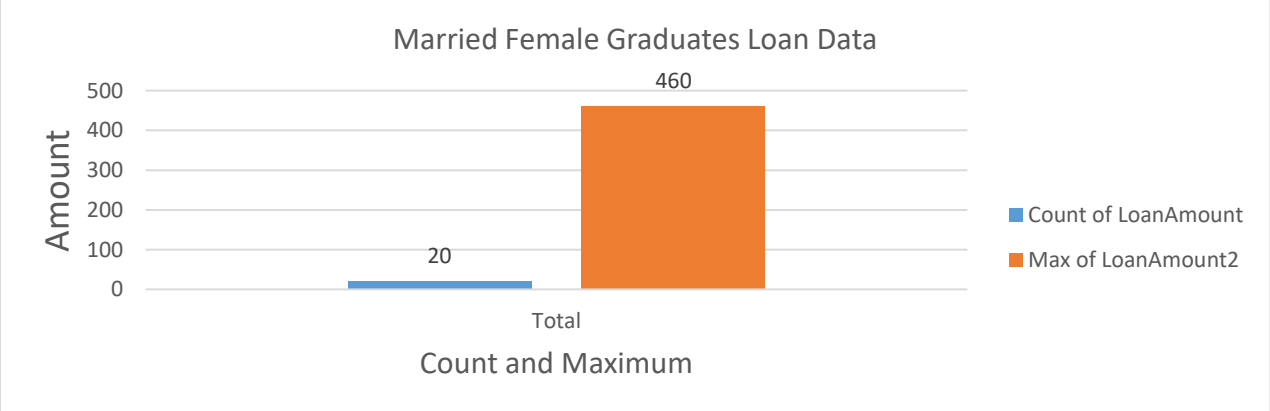
Answer:
300 Unmarried Graduate Females applied for the loan out of which the maximum amount of loan applied was 35

Q3. How many male non-graduates who are not married applied for Loan? What was the highest amount?



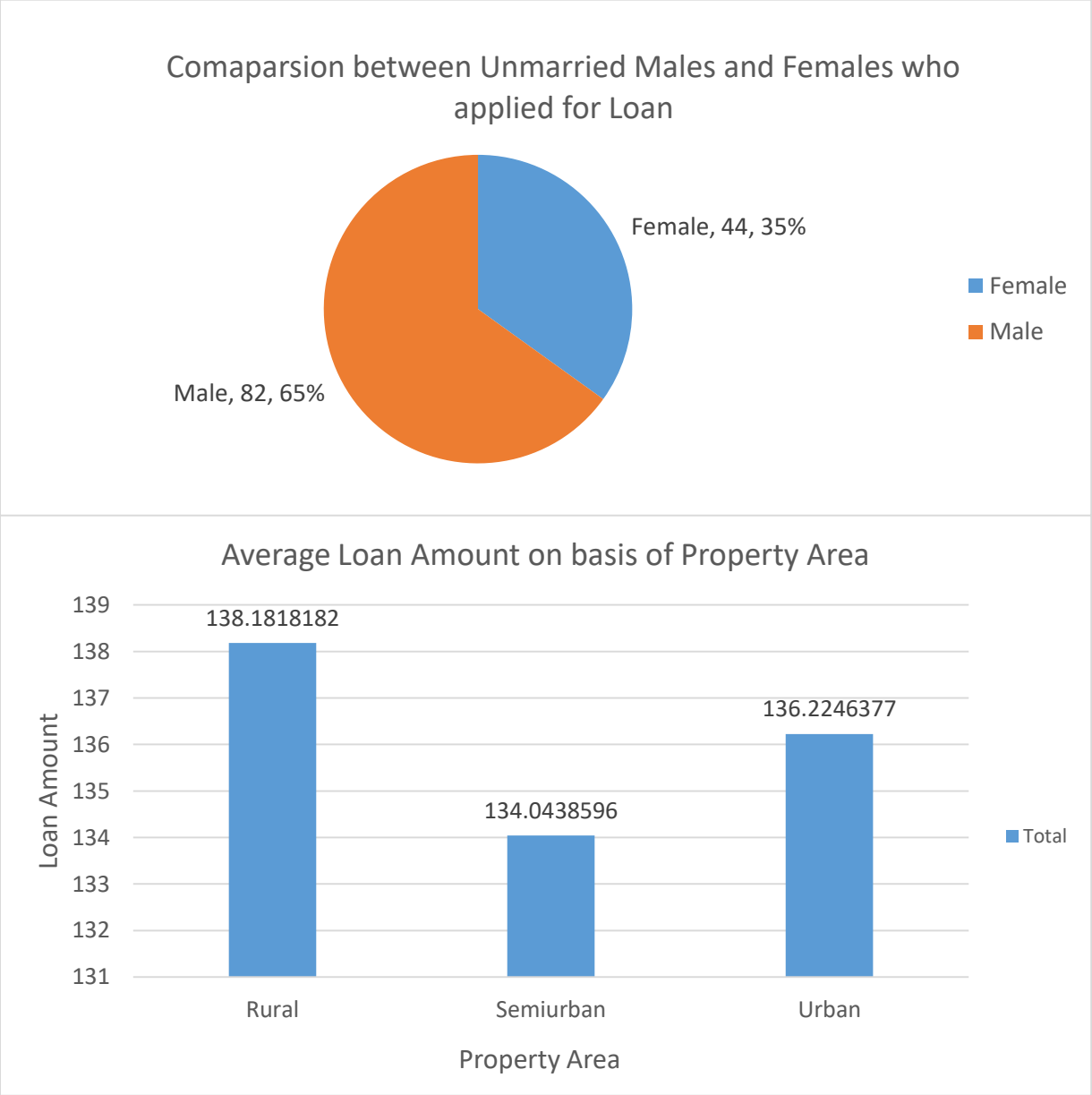
Answer:
199 Non Graduate Unmarried Males applied for loan out of which 16 was the highest amount of loan applied.

Q4. How many female graduates who are married applied for Loan? What was the highest amount?



Answer:
460 Married Graduate Females applied for the loan out of which 20 was the highest amount for which loan was applied

Q5. How many male and female who are not married applied for Loan? Compare Urban, Semi-urban and rural on the basis of amount.



Answer:

Out of Unmarried Males and Females, Males applied 38 more requests than females. The average amount of Loan Amount in Rural Area = 131.182, Semi Urban Area = 134.04, Urban Area = 136.22.

Regression:

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.45908096							
R Square	0.21075532							
Adjusted R Square	0.20858707							
Standard Error	56.0766111							
Observations	366							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	305655.205	305655.205	97.2004502	1.7676E-20			
Residual	364	1144629.42	3144.58631					
Total	365	1450284.62						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	106.07753	4.10024098	25.8710478	1.7585E-84	98.014396	114.140665	98.014396	114.140665
5720	0.0058851	0.00059692	9.85902887	1.7676E-20	0.00471125	0.00705895	0.00471125	0.00705895

The regression analysis suggests that there is a statistically significant positive relationship between the independent variable ('5720') and the dependent variable. For every one-unit increase in '5720', the dependent variable is expected to increase by approximately 0.0059 units. However, it's important to note that the model only accounts for about 21.1% of the total variance in the dependent variable.

Correlation:

	ApplicantIncome	*CoapplicantIncome*	*LoanAmount*
ApplicantIncome	1		
CoapplicantIncome	-0.110334799	1	
LoanAmount	0.458768926	0.144787815	1

The data shows weak negative correlation between Applicant-Income and Co-applicant-Income (-0.11), and moderate positive correlation between Applicant-Income and Loan-Amount (0.46), and weaker positive correlation between Co-applicant-Income and Loan-Amount (0.14).

Anova (Single Factor):

SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
ApplicantIncome	367	1763655	4805.599455	24114831.09		
CoapplicantIncome	367	576035	1569.577657	5448639.491		
LoanAmount	367	49280	134.2779292	3964.141124		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4202537452	2	2101268726	213.2009841	5.87569E-79	3.003920577
Within Groups	10821681107	1098	9855811.573			
Total	1502421856	1100				

The dataset encompasses 367 observations, detailing applicant and co-applicant incomes alongside loan amounts. On average, applicants possess a higher income, averaging around $4805.60, compared to co-applicants whose average income is approximately $1569.58. Loan amounts vary widely, averaging $134.28. ANOVA analysis underscores significant distinctions between the income and loan amounts across the groups, implying diverse financial profiles among applicants and co-applicants.

Anova two factor without Replication:

ANOVA							
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*	
Rows	1004340909	365	2751618.93	1.015674698	0.440986529	1.1881716	
Columns	379216841.8	1	379216841.8	139.9761235	1.47092E-27	3.867061668	
Error	988841123.7	365	2709153.763				
Total	2372398875	731					

The ANOVA results indicate significant variation both within rows ($p = 0.441$) and between columns ($p < 0.001$). This suggests that there are meaningful differences among the row categories and column categories in the dataset, warranting further investigation into the factors influencing these variations.

Descriptive Statistics:

LoanAmount	
Mean	136.1326
Standard Error	3.22536
Median	125
Mode	150
Standard Deviation	61.36665
Sample Variance	3765.866
Kurtosis	9.407853
Skewness	2.223512
Range	522
Minimum	28
Maximum	550
Sum	49280
Count	362

ApplicantIncome	
Mean	4805.599
Standard Error	256.3357
Median	3786
Mode	5000
Standard Deviation	4910.685
Sample Variance	24114831
Kurtosis	103.1275
Skewness	8.441375
Range	72529
Minimum	0
Maximum	72529
Sum	1763655
Count	367

The dataset includes information on Applicant-Income and Loan-Amount. The largest Applicant-Income recorded is $72,529, while the smallest is $0. Additionally, the Loan-Amount ranges from a maximum of $550 to a minimum of $0.

4. Conclusion and Reviews

Conclusion:

Our loan data analysis reveals several key insights into the demographics and financial behaviors of loan applicants. We observed notable trends across various demographic groups, such as unmarried male and female graduates and non-graduates, as well as married female graduates. Additionally, our regression analysis indicates a statistically significant positive relationship between certain variables, though the model accounts for only about 21.1% of the total variance in loan amounts. Moreover, correlation analysis underscores relationships between applicant incomes, co-applicant incomes, and loan amounts, providing valuable context for understanding financial dynamics.

Reviews:

The comprehensive nature of the dataset allowed for in-depth exploration of loan application dynamics, providing valuable insights for stakeholders in the lending industry. However, while the analysis sheds light on important trends and relationships, there are areas for further exploration and refinement. Enhancing the model's predictive power and understanding additional factors influencing loan amounts could enrich future analyses. Moreover, incorporating qualitative data or borrower feedback could offer deeper insights into the loan application process and borrower motivations. Overall, this report serves as a solid foundation for future research and decision-making in the lending sector.