Loan Data Report

1. Introduction:

Dataset Overview:

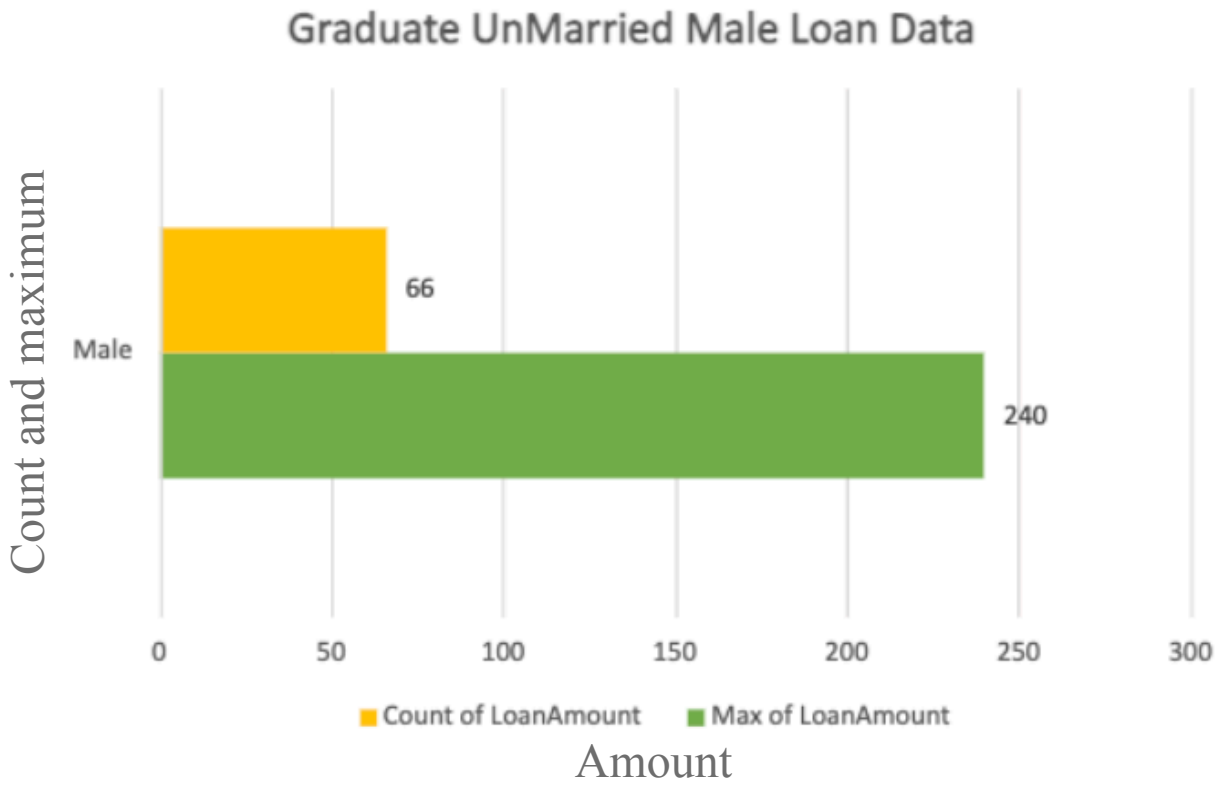
Our dataset comprises a diverse array of variables, each providing insights into the intricate dynamics of loan applications. From fundamental applicant details such as Gender, Marital Status, and Education to more nuanced factors like Employment Status, Loan Amount, and Residential Type, every aspect has been meticulously documented.

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.

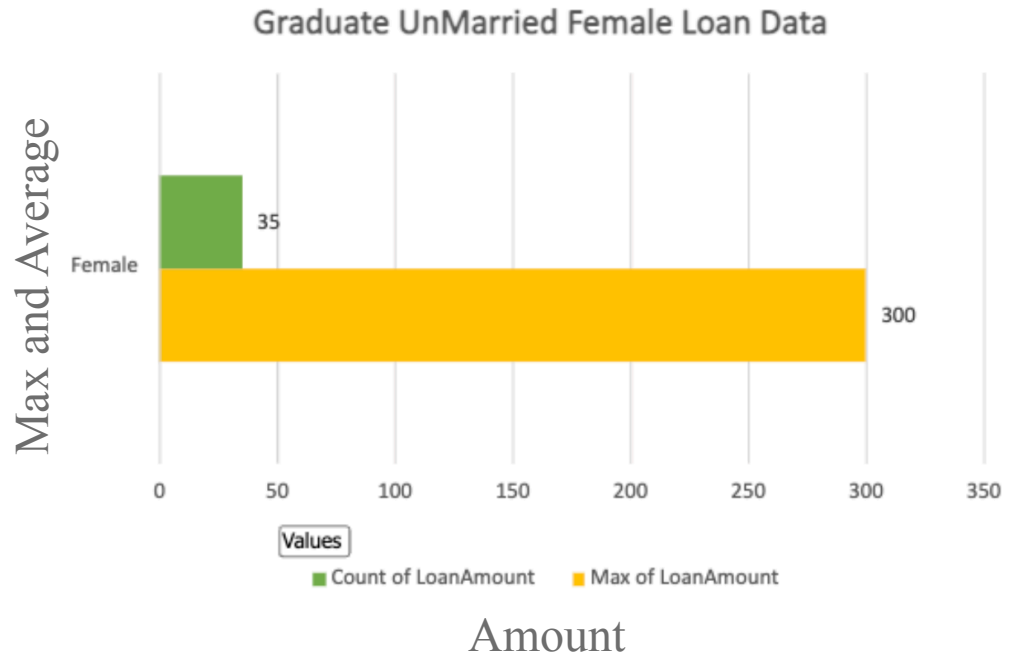
3. Analytics:

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



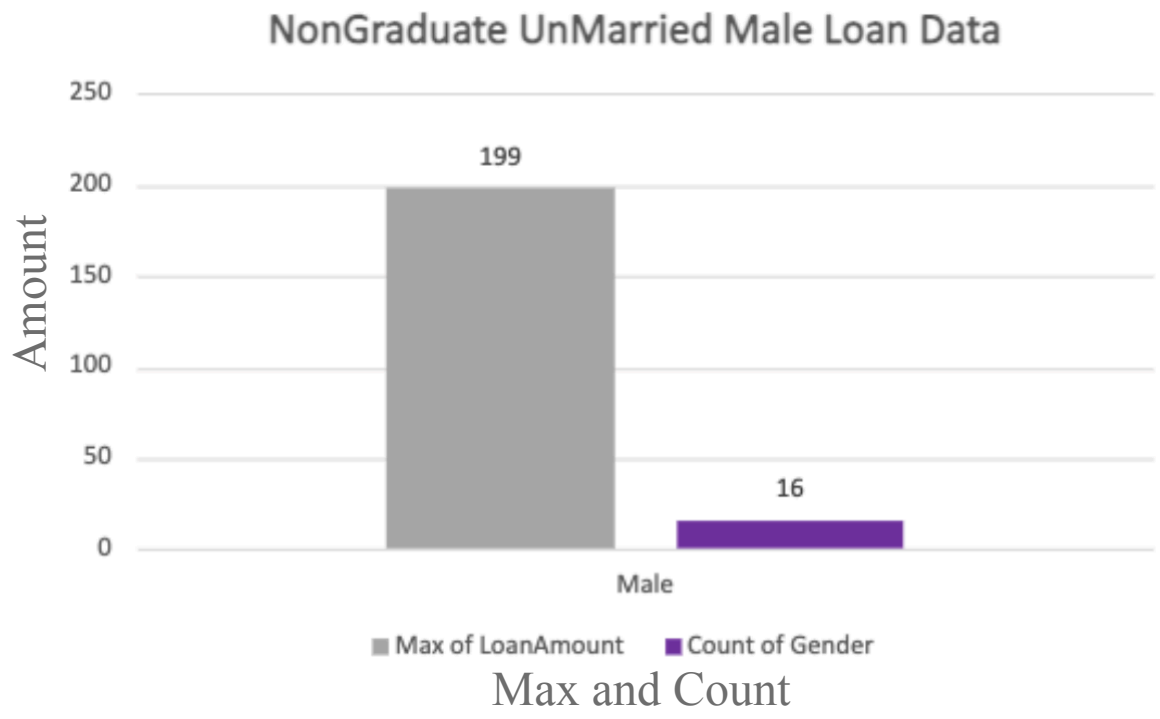
Answer:
Out of 240 loan applicants who were unmarried graduates and males, the highest loan amount applied for was $66.

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



Answer:
Among the 300 unmarried female loan applicants who were graduates, the highest loan amount applied for was $35.

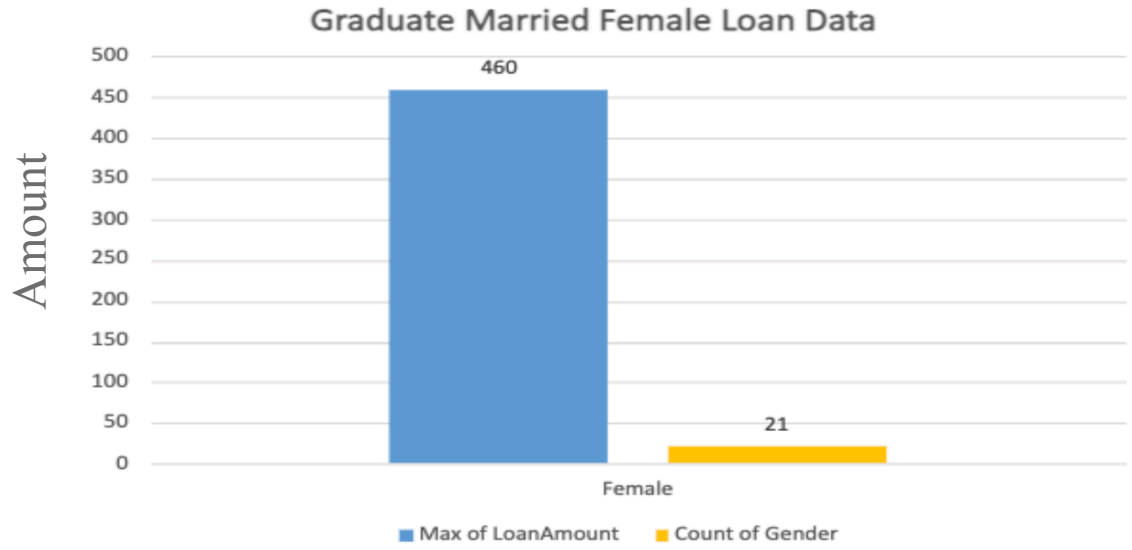
Q3. How many male non-graduates who are not married applied for Loan? What was the highest amount?



Max and Count

Answer:
Out of the 199 loan applicants who were unmarried males and non-graduates, the highest loan amount applied for was $16.

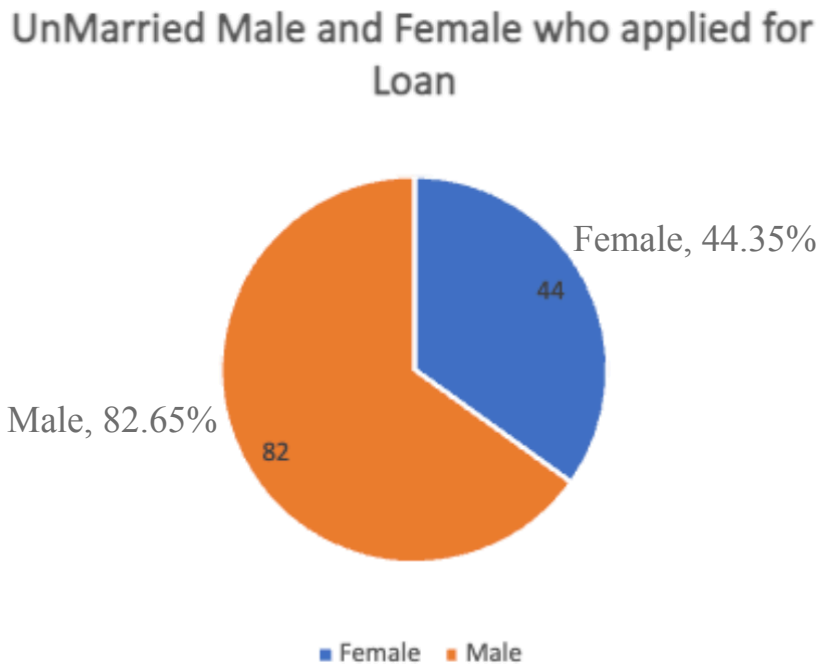
Q4. How many female graduates who are married applied for Loan? What was the highest amount?

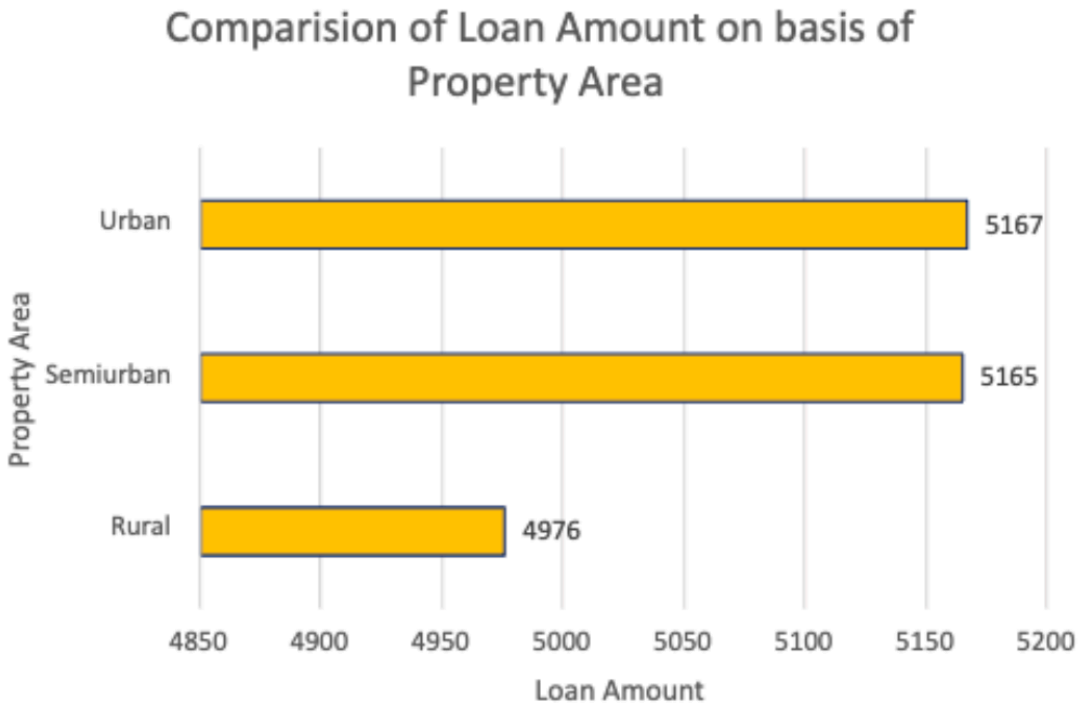


Max and Count

Answer:
Among the 460 married female loan applicants who were graduates, the highest loan amount applied for was $20.

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:
The number of loan applications from unmarried males exceeded those from females by 38 requests.
The average loan amount in rural areas is $131.182, in semi-urban areas is $134.04, and in urban areas is $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 indicates 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 essential to note that the model only explains 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 exhibits a weak negative correlation (-0.11) between Applicant-Income and Co-applicant-Income, a moderate positive correlation (0.46) between Applicant-Income and Loan-Amount, and a weaker positive correlation (0.14) between Co-applicant-Income and Loan-Amount.

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					

Columns: The p-value (1.47092E-27) is extremely small, indicating a highly significant difference among the column categories. Thus, variations observed between columns are not due to random chance but are likely influenced by the factor being studied.

The ANOVA results reveal significant variation both within rows ($p = 0.441$) and between columns ($p < 0.001$). This indicates meaningful differences among the row categories and column categories in the dataset, necessitating 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 comprises information on Applicant-Income and Loan-Amount. The highest recorded Applicant-Income is $72,529, and the lowest is $0. Similarly, the Loan-Amount ranges from a maximum of $550 to a minimum of $0.

4. CONCLUSION:

Our analysis, using varied visualization techniques, revealed valuable insights, enhancing comprehension and decision-making. Visualizing data clarified complex findings, facilitating actionable strategies. This highlights the pivotal role of data visualization in extracting meaningful insights and informing decisions effectively.