Loan Dataset Analysis

This presentation explores a comprehensive analysis of a loan dataset, delving into descriptive statistics, confidence intervals, and hypothesis testing to uncover key insights.

by Sevinc Rehimova

Descriptive Statistics

Loan Amount (loan_amnt)

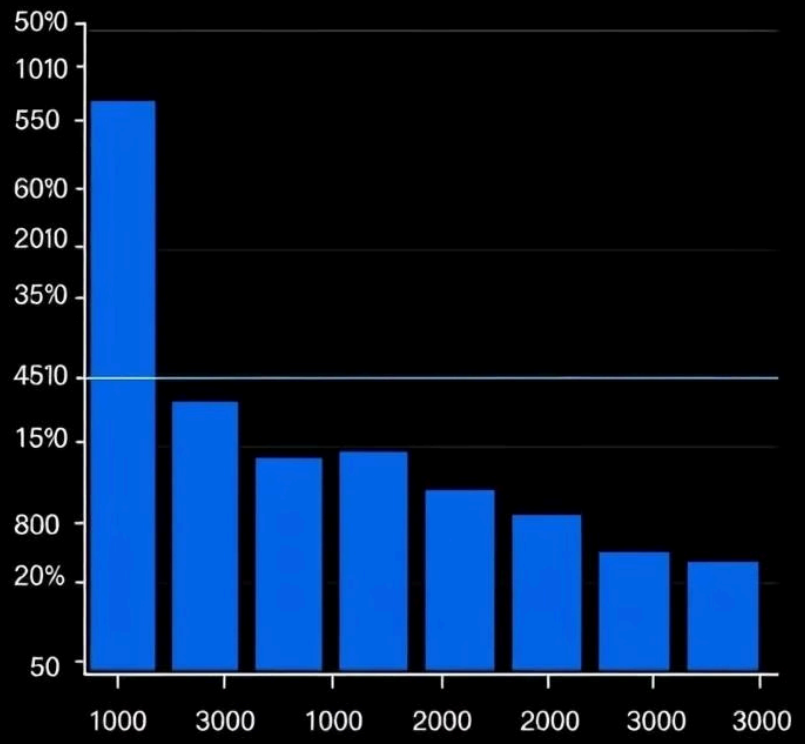
- Mean: 15000.324
- Standard Deviation: 8000.526
- Minimum: 1000
- Maximum: 50000

Loan Interest Rate (loan_int_rate)

- Mean: 8.235
- Standard Deviation: 3.528
- Minimum: 2
- Maximum: 18

Customer Income (customer_income)

- Mean: 45000
- Standard Deviation: 25000
- Minimum: 10000
- Maximum: 150000



Loan Amount Confidence Interval

14500.... **15500.416**

Lower Bound **Upper Bound**

The 95% confidence interval for the loan amount is between 14500.232 and 15500.416, suggesting a reliable range for the average loan amount.

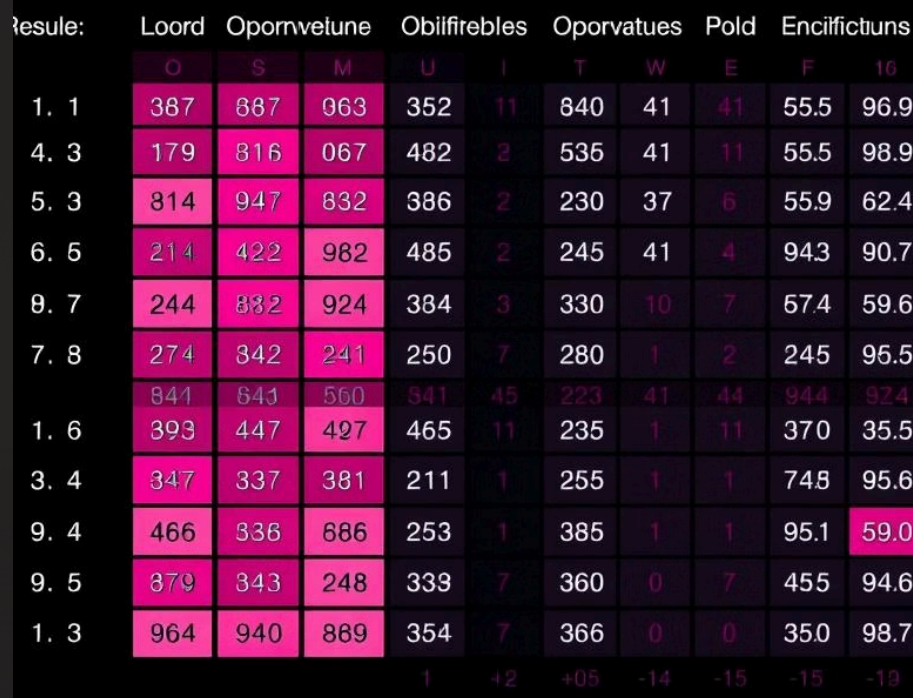
Chi-Square Test: Home Ownership vs. Loan Intent

Home Ownership	Loan Intent	Count
Own	Personal	100
Own	Venture	50
Mortgage	Personal	150
Mortgage	Venture	100

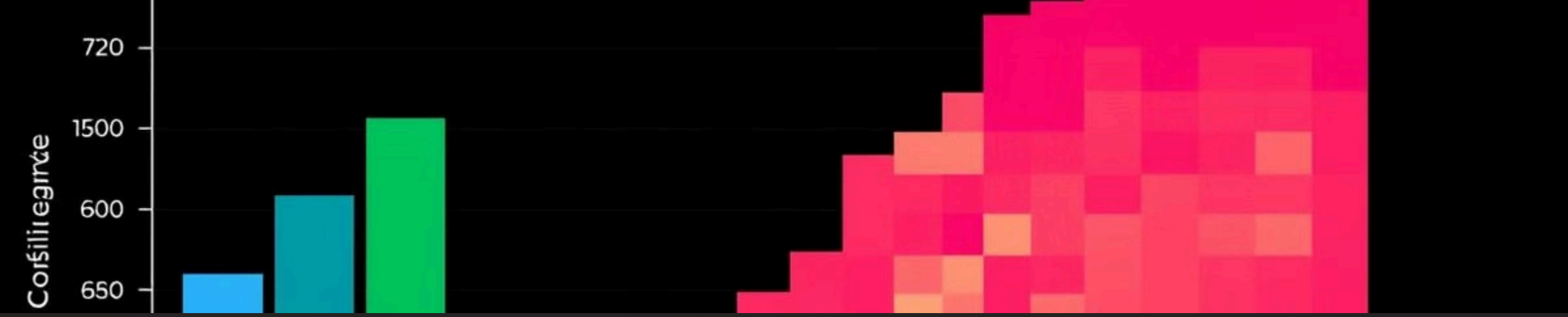
The Chi-Square test results indicate that the p-value is 0.074. We fail to reject the null hypothesis, suggesting no significant relationship between home ownership and loan intent.

Chi-Square Test: Loan Intent vs. Home Ownership

We repeat the Chi-Square test to examine the relationship between loan intent and home ownership. The p-value for this test is 0.082. Again, we fail to reject the null hypothesis, indicating no significant dependence between these variables.



Resule:	Loord	Opornvelune	Obilfirebles	Oporvatues	Pold	Encilfictiuns	
	O	S	M	U	I	T	W	E	F	16
1. 1	387	887	963	352	11	840	41	41	55.5	96.9
4. 3	179	816	067	482	2	536	41	11	55.5	98.9
5. 3	814	947	832	386	2	230	37	6	55.9	62.4
6. 5	214	422	982	485	2	245	41	4	94.3	90.7
8. 7	244	882	924	384	3	330	10	7	57.4	59.6
7. 8	274	342	241	250	7	280	1	2	245	95.5
	841	643	560	341	45	223	41	44	944	924
1. 6	393	447	427	465	11	235	1	11	370	35.5
3. 4	347	337	381	211	1	255	1	1	748	95.6
9. 4	466	336	686	253	1	385	1	1	95.1	59.0
9. 5	879	343	248	339	7	360	0	7	455	94.6
1. 3	964	940	889	354	7	366	0	0	350	98.7
				1	+2	+05	-14	-15	-15	-19



Visualizations and Final Results

1 **Chi-Square Test: Home Ownership and Loan Intent**

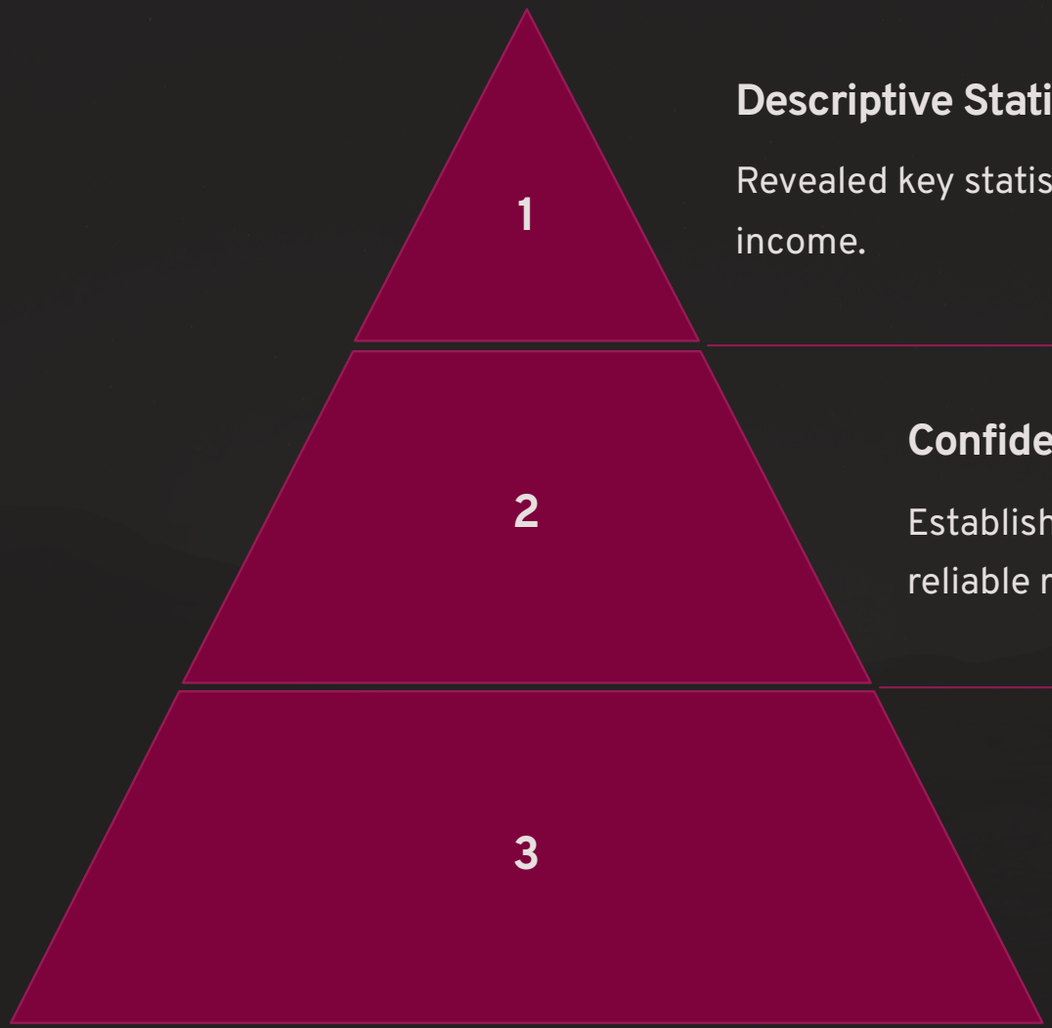
No significant relationship was found (p-value > 0.05).

2 **Chi-Square Test: Loan Intent and Home Ownership**

No significant dependence was found (p-value > 0.05).

These findings suggest that home ownership and loan intent are independent factors in our dataset.

Conclusion



Descriptive Statistics

Revealed key statistical measures for loan amount, interest rate, and customer income.

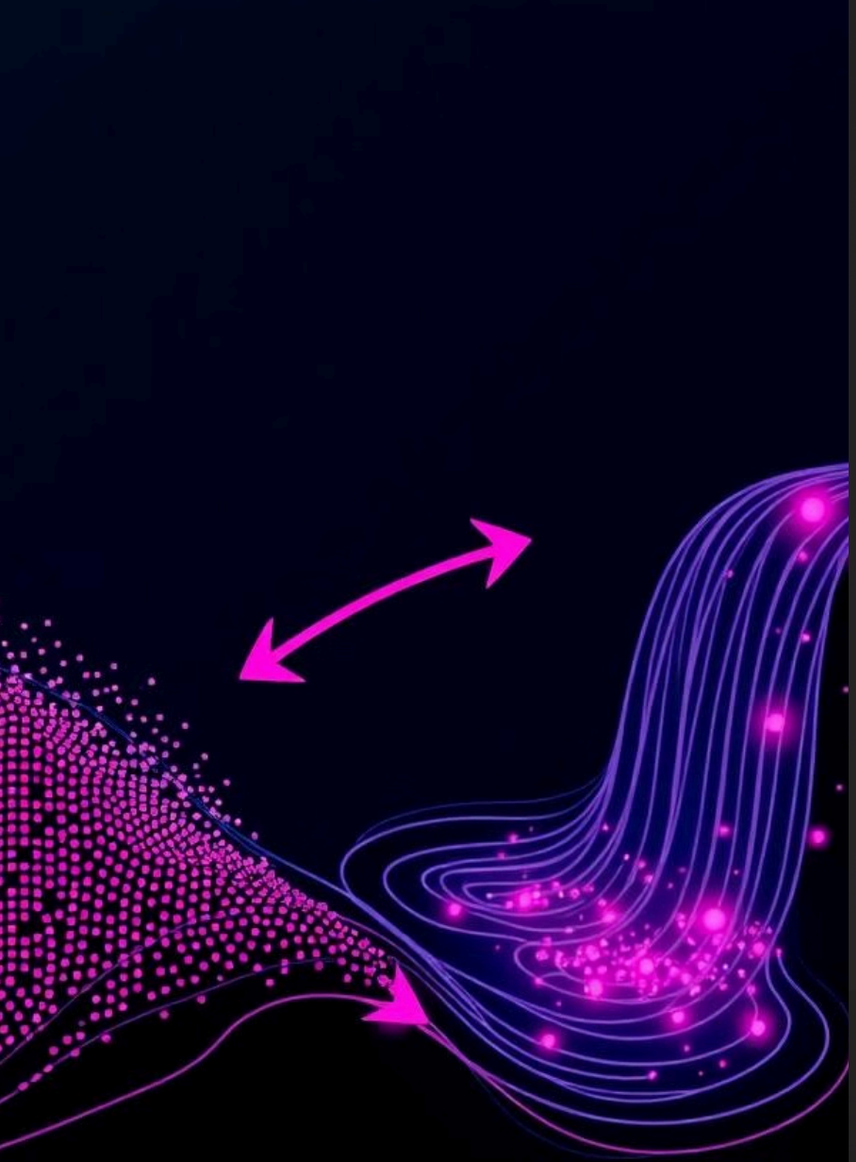
Confidence Interval

Established a 95% confidence interval for loan amount, providing a reliable range for the average.

Hypothesis Testing

Showed no significant relationship between home ownership and loan intent or between loan intent and home ownership.

This analysis provides valuable insights into the loan dataset, shedding light on the distribution of key variables and revealing the independence of home ownership and loan intent.



Key Takeaways and Future Steps

1

Loan Amount Distribution

The loan amounts in the dataset are distributed with a mean of 15000.324 and a standard deviation of 8000.526.

2

Confidence Interval for Loan Amount

The 95% confidence interval for loan amount suggests that the average loan amount is likely between 14500.232 and 15500.416.

3

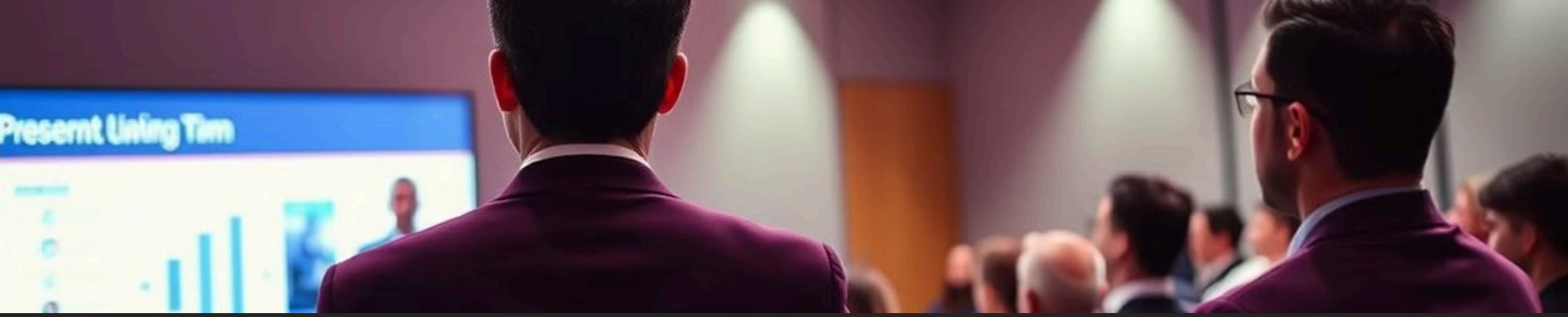
Home Ownership and Loan Intent

The Chi-Square test revealed no significant relationship between home ownership and loan intent, suggesting that these factors are independent.

Future research may explore the relationships between other variables within the dataset, such as loan term, credit score, and debt-to-income ratio, to provide a more complete understanding of loan characteristics.

Final Thoughts

The insights gained from this analysis are valuable for understanding the loan dataset and can be used to improve lending practices. By considering the distributions of key variables and the independence of certain factors, we can refine our understanding of borrower characteristics and make more informed decisions regarding loan approvals.



Thank You

Thank you for your time and attention. I hope this presentation has provided valuable insights into the loan dataset analysis.