

Chapter 4: Data Preprocessing and Exploration

Your Name

Your University

Chapter 4: Data Preprocessing and Exploration

Descriptive Statistics

Table ?? presents descriptive statistics comparing fraudulent and non-fraudulent PPP loans. Key financial indicators, employee counts, and forgiveness patterns are summarized for both classes.

Metric	Non-Fraud	Fraud
Loan Amount (Mean)	530486.47	958492.36
Loan Amount (SD)	737515.30	1458612.38
Loan Amount (Median)	295000.00	420552.00
Loan Amount (Min)	150000.00	150000.00
Loan Amount (Max)	10000000.00	10000000.00
Jobs Reported (Mean)	51.88	61.79
Jobs Reported (SD)	67.55	86.97
Jobs Reported (Median)	30.00	26.00
Jobs Reported (Min)	0.00	3.00
Jobs Reported (Max)	500.00	490.00
Forgiveness Ratio (Mean)	0.97	0.32
Forgiveness Ratio (SD)	0.17	0.47
Forgiveness Ratio (Median)	1.01	0.00
Forgiveness Ratio (Min)	0.00	0.00
Forgiveness Ratio (Max)	2.02	1.03

Table 1

Summary Statistics by Fraud Label

Note. This table presents the mean, standard deviation, median, minimum, and maximum values for loan amount, number of jobs reported, and forgiveness ratio. Non-fraudulent and fraudulent loans are compared to highlight observed differences.