**Final Data Analysis Report**

**Introduction**

This project aims to analyse the closure of banks over time, focusing on the distribution of closures by state and acquiring institution. The primary goal is to understand the trends and patterns in bank closures, which can inform policy decisions and strategic planning in the banking sector. The original data source for this analysis is a collection of bank closure records, available at [link](https://www.kaggle.com/datasets/sahirmaharajj/failed-banks-dataset?resource=download). This project was selected due to the relevance of understanding the dynamics of bank closures, which can impact financial stability and consumer trust.

**ETL Script Steps**

1. **Data Extraction**: The data was extracted from CSV files using custom extraction functions provided by the extract module. These functions load the data into pandas DataFrames for further processing.
2. **Data Transformation**: The 'Closing Date' column was converted to datetime format to facilitate time-based analysis. The data was then grouped by year to analyse the number of closures over time.
3. **Data Loading**: The transformed data was loaded into memory for analysis and visualization.

**Visualizations and Interpretations**

1. **Number of Closed Banks Over Time**: This line plot shows the trend in the number of bank closures over the years. The visualization indicates a fluctuating pattern, suggesting that bank closures may be influenced by various factors, including economic conditions and regulatory changes.

A graph with a line graph

Description automatically generated

1. **Distribution of Closed Banks by State**: This pie chart illustrates the proportion of bank closures by state. The visualization highlights the states with the highest number of closures, providing insights into regional disparities in bank health and stability.

A pie chart with different colors and numbers

Description automatically generated

1. **Top Acquiring Banks by Number of Closed Banks**: This bar chart displays the top 10 acquiring banks by the number of closed banks. The visualization reveals which banks were most affected by closures, potentially indicating areas of concern or systemic issues within the banking industry.

A graph of a financial report

Description automatically generated with medium confidence

**Conclusion**

The analysis of bank closures provides valuable insights into the dynamics of the banking sector. The data reveals fluctuations in closures over time, significant regional disparities, and identifies key acquiring banks most affected by closures. These findings underscore the importance of continuous monitoring and analysis of bank closures to ensure financial stability and consumer protection. Further research could explore the underlying causes of these trends and evaluate the effectiveness of current policies and regulations in mitigating bank closures.