**Appendix**

**Data Cleaning and Transformation**

**Observations on Data Cleaning and Transformation**

1. **Duplicate Values Removal**:
   * Removed duplicate records based on company, location, industry, total laid off, percentage laid off, date, stage, country, and funds raised.
   * Used a Common Table Expression (CTE) to identify duplicates and delete them.
2. **Standardizing Formats**:
   * Trimmed unnecessary whitespace in company names.
   * Standardized industry names (e.g., consolidated different variations of 'Crypto' into 'Crypto').
   * Cleaned up country names by removing trailing periods (e.g., 'United States.' → 'United States').
3. **Date Format Conversion**:
   * Changed the date column to an appropriate date format using SET datestyle TO MDY; and altering the column type.
4. **Handling Missing Data**:
   * Replaced missing industry values using self-join where a company's industry was available in another row.
   * Manually assigned missing industry values for specific companies (e.g., Airbnb → Travel, Carvana → Transportation, Juul → Consumer).
   * Removed records where both total\_laid\_off and percentage\_laid\_off were NULL.
   * Ensured that no essential fields contained NULL values.

**Columns Used for Analysis and Visualization**

1. **Company**: Identifies the organization affected by layoffs.
2. **Location**: Indicates where the layoffs occurred.
3. **Industry**: Helps analyze layoff trends across different industries.
4. **Total Laid Off**: Measures the impact of layoffs in absolute numbers.
5. **Percentage Laid Off**: Assesses the severity of layoffs in relation to company size.
6. **Date**: Enables time-series analysis of layoff trends.
7. **Stage**: Represents the growth stage of affected companies (e.g., Startup, Public, etc.).
8. **Country**: Helps analyze the geographical distribution of layoffs.
9. **Funds Raised (Millions)**: Determines if well-funded companies lay off more employees.

**Data Analysis Queries**

1. **Industries with the Highest Total Layoffs**:
   * Summed layoffs per industry and ranked them in descending order.
2. **Industries with the Highest Layoff Percentage**:
   * Averaged layoff percentages per industry to determine relative impact.
3. **Layoff Trends Over Time (Monthly Layoffs)**:
   * Aggregated layoffs per month to analyze time-series trends.
4. **Layoffs Over Time by Industry**:
   * Analyzed how layoffs evolved across industries over time.
5. **Companies with the Largest Single Layoffs**:
   * Identified firms with the highest individual layoff events.
6. **Countries Most Affected by Layoffs**:
   * Summed layoffs per country to determine the most impacted regions.
7. **Funding vs. Layoffs**:
   * Compared total funds raised against total layoffs per industry.
8. **Rolling Total of Layoffs Per Month**:
   * Computed a cumulative sum of layoffs to track long-term trends.
9. **Companies with 100% Workforce Laid Off**:
   * Identified organizations that completely shut down.
10. **Highly Funded Companies with 100% Layoffs**:

* Filtered companies that raised significant funding but still laid off all employees.

1. **Ranking Industries by Layoffs Per Year**:

* Used window functions to rank industries by layoffs for each year.

This appendix provides a comprehensive overview of the data preparation and analysis methods.