FIRST-ETL-PIPELINE-PROJECT

Created by	sanagavaram badrinath
Created time	@May 18, 2025 8:43 PM
Category	Project Description
Last edited by	sanagavaram badrinath
Last updated time	@May 18, 2025 8:54 PM

Background

Background

This project analyses global tech industry layoffs using publicly available data. We uncovered geographic and industry-specific layoff patterns through data cleaning, transformation, and visualisation. Our analysis helps stakeholders and policymakers understand which regions and sectors face the greatest workforce reductions, enabling more informed decision-making.

The scope included:

- Importing and preparing raw data from a CSV file
- Cleaning and transforming the dataset using SQL
- Performing exploratory and aggregate analysis
- Visualising global layoffs in Tableau

Analysis

The data analysis consisted of the following operations and tasks:

1. **Data Cleaning** (`Data_cleaning.sql`)
    - Removed duplicates using `ROW_NUMBER()` and `GROUP BY + HAVING`
    - Standardised inconsistent entries (e.g., "CryptoCurrency" to "Crypto")
    - Cleaned country and industry columns by removing excess spaces and punctuation
    - Converted `date` field from TEXT to DATE format
    - Populated missing values in the `industry` column using self-joins
    - Removed rows with completely NULL layoff metrics
2. **Data Transformation** (`ETL_Loading.ipynb`)
    - Imported cleaned SQL data into a Jupyter Notebook
    - Aggregated layoffs by `company`, `country`, `industry`, `month`, and `year`
    - Calculated rolling totals using SQL window functions
3. **Insights via SQL** (`Data_analysis.sql`)
    - Identified countries and companies with the highest layoff numbers
    - Calculated total workforce size using `total_laid_off` and `percentage_laid_off`
    - Examined layoff trends across years and months
    - Identified most affected industries (e.g., Tech, Crypto)
4. **Visualization**
    - Exported country-level data to Tableau
    - Created a choropleth world map showing layoff intensity
    - Highlighted most impacted countries: USA, Japan, and Canada

Recommendations

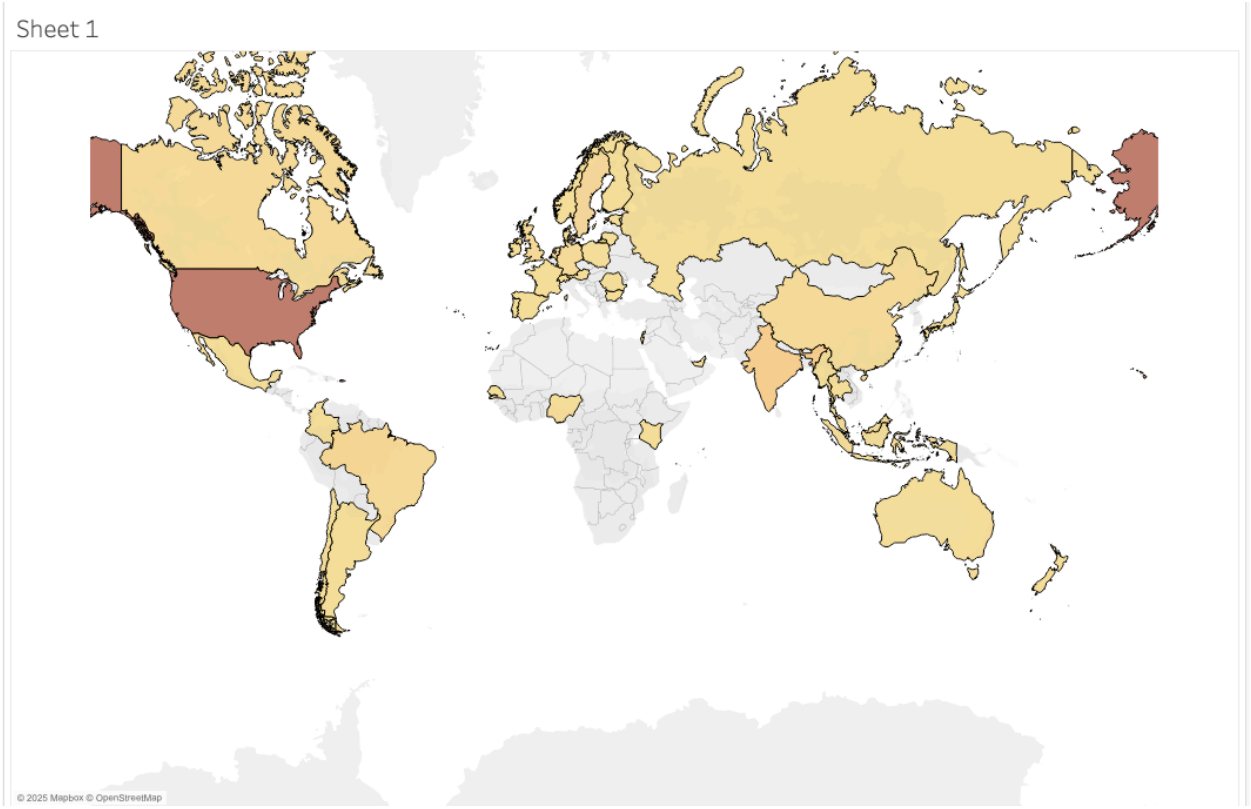
Based on the analysis, we recommend:

- **For Companies**: Organisations in heavily affected sectors should develop proactive workforce plans and diversify their business strategies.

- **For Governments**: Direct support programs—including skills training and unemployment benefits—toward regions experiencing the highest concentration of layoffs.
- **For Analysts**: Leverage this dataset to develop detailed company-level risk assessments and predict broader economic impacts.

Implementation

Step	Task	Tools Used
1	Load CSV data to MySQL database	MySQL Workbench
2	Clean and standardize data	SQL
3	Transform and aggregate for analysis	Python (Jupyter), SQL
4	Visualize key insights on a global map	Tableau
5	Export reports and automate insights	Tableau dashboards, Python scripting (optional)


Sheet 1
© 2025 Mapbox © OpenStreetMap