**GLOBAL LAYOFFS**

# SQL Data Cleaning & EDA Project

Presented By **ANJALI BAJAJ**

# Project Objectives:

## Data Cleaning:

- **Created a Staging Table:** Established a staging table to preserve raw data while performing data cleaning operations.
- **Identified and Removed Duplicates:** Detected and eliminated duplicate records to ensure data uniqueness and integrity.
- **Standardized Data Entries:** Corrected inconsistencies and standardized values in key fields (e.g., industry names, country names) for uniformity.
- **Managed Missing Values:** Addressed and managed missing or NULL values to maintain data quality and usability.
- **Trimmed Unnecessary Data:** Removed irrelevant or redundant columns and rows to streamline the dataset.

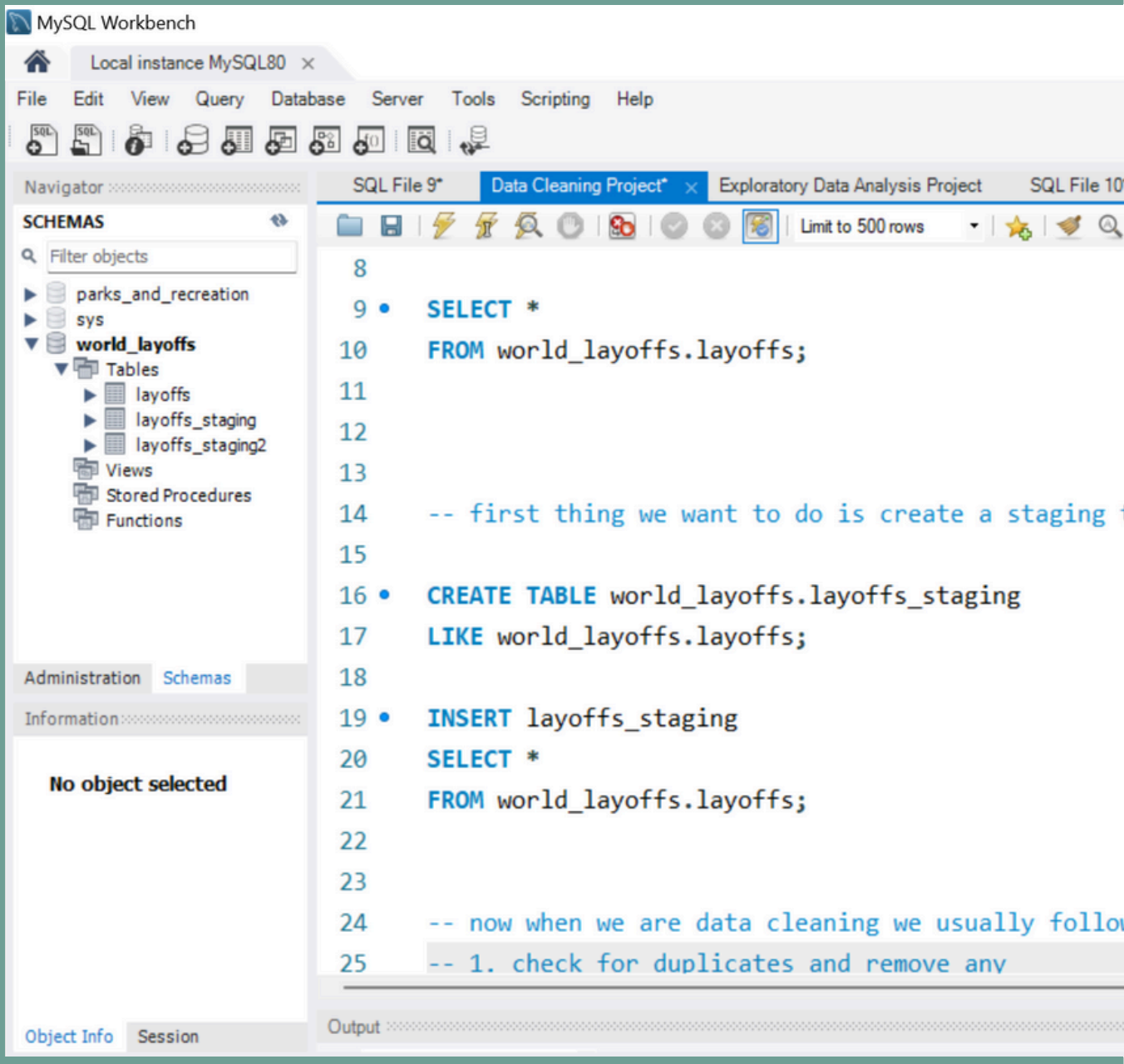
# Project Objectives:

## Exploratory Data Analysis:

- **Trend Identification:** Analyzed the dataset to uncover significant insights, patterns, and outliers related to global layoffs.
- **Analyzed Layoffs by Company:** Examined and ranked companies based on the total number of layoffs and significant single-day layoffs.
- **Geographic and Industry Analysis:** Assessed layoffs by location and industry to understand regional and sector-specific impacts.
- **Temporal Trends:** Conducted a rolling analysis of layoffs over time to detect fluctuations and patterns.
- **Economic Correlation:** Analyzed how funds raised by companies relate to the percentage of layoffs to understand economic impacts.

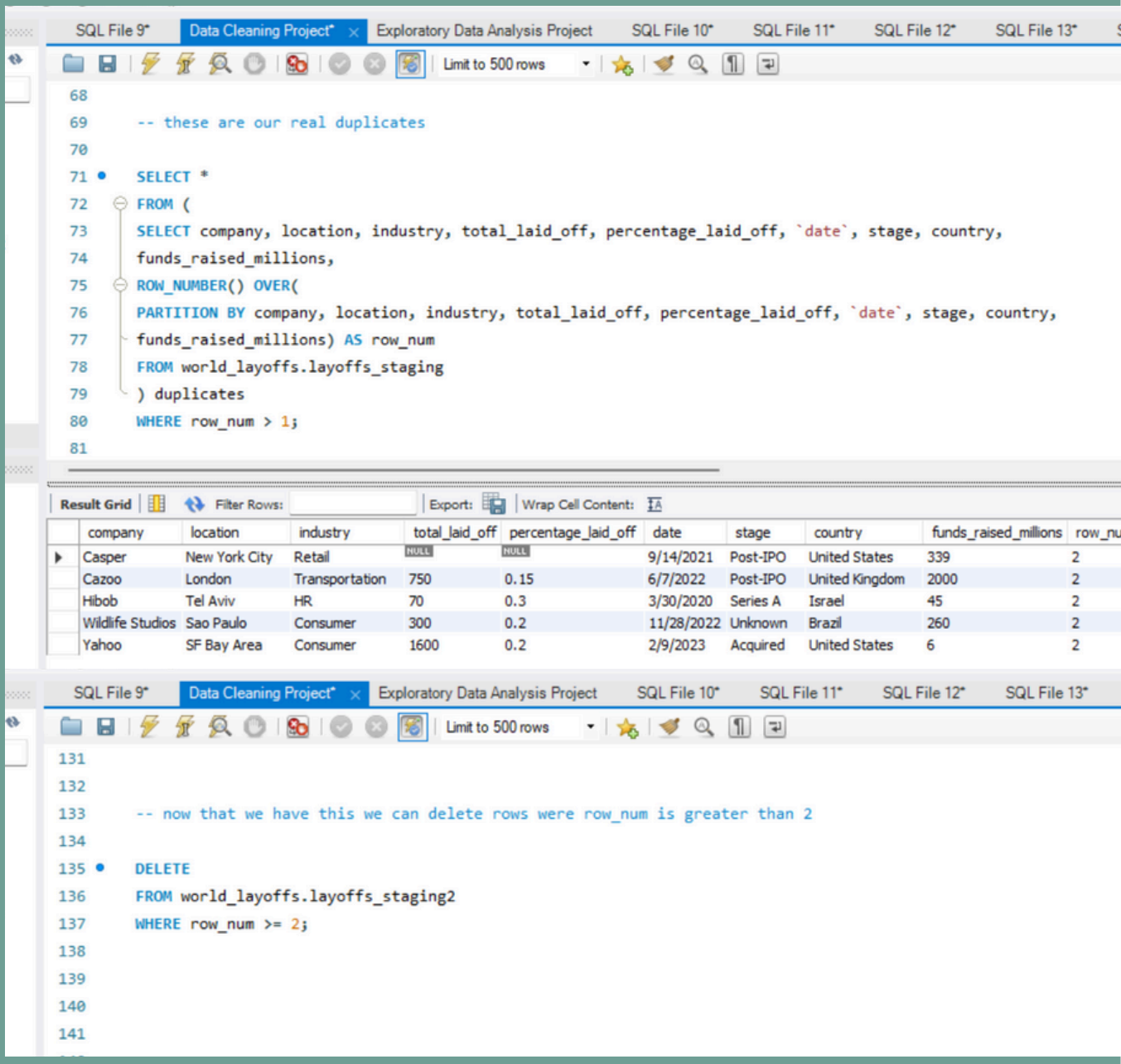
# Created a Staging Table `layoffs_staging`

## To work on data cleaning while preserving the raw data.



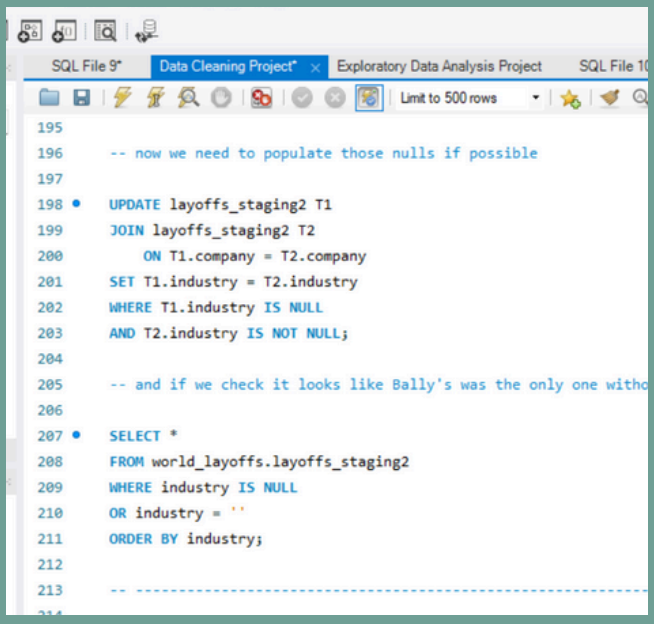
# Removed Duplicates

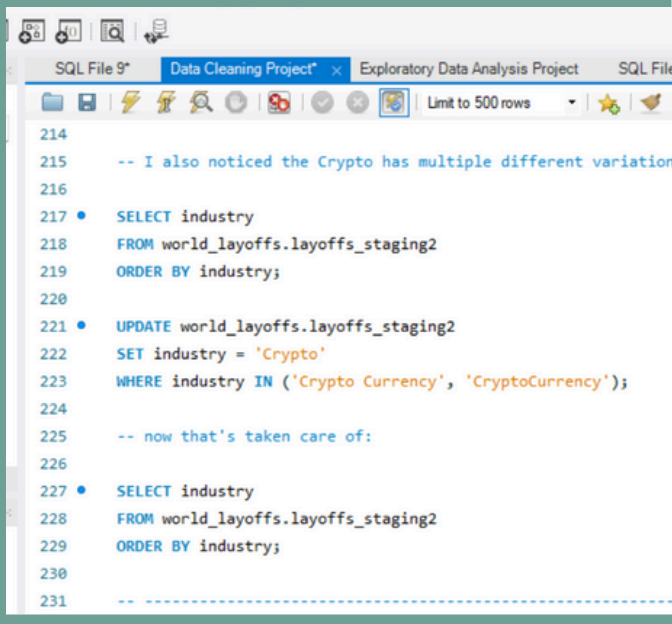
## By checking them using the `ROW_NUMBER()` function.

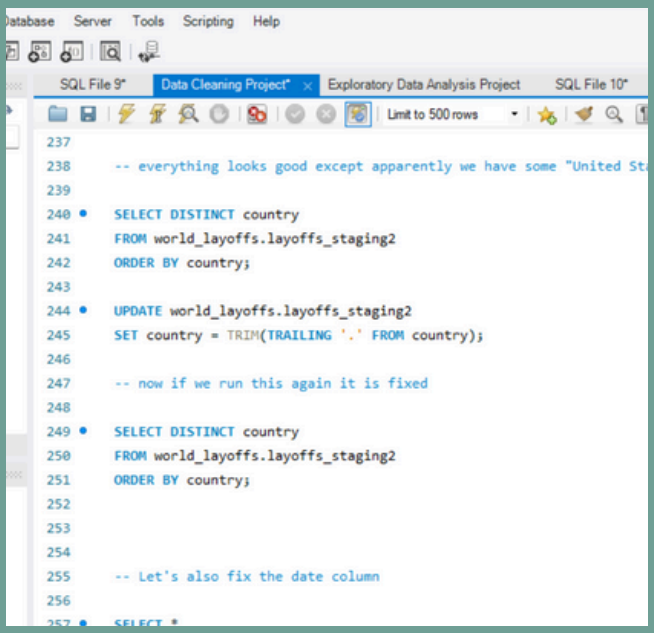


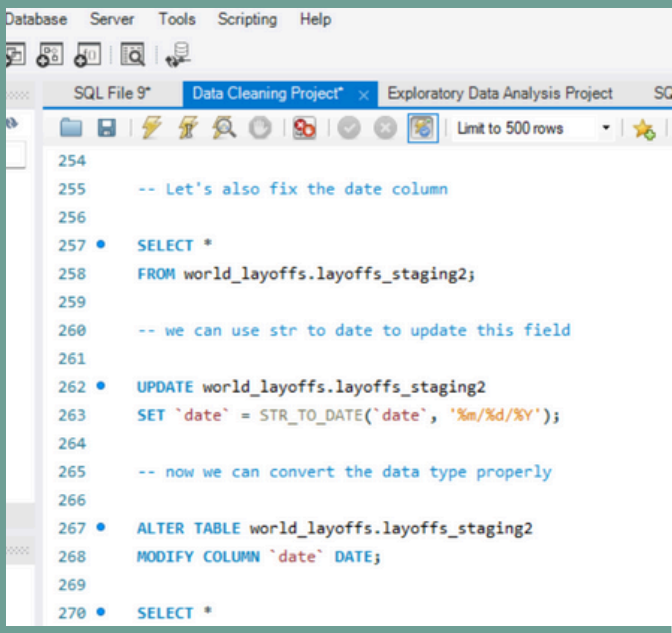
# Standardized Data

- Industry - Populated NULLs from non-null values and standardized "Crypto".
- Country - Fixed errors by trimming trailing periods.
- Date - Standardized date format using `STR_TO_DATE()`.



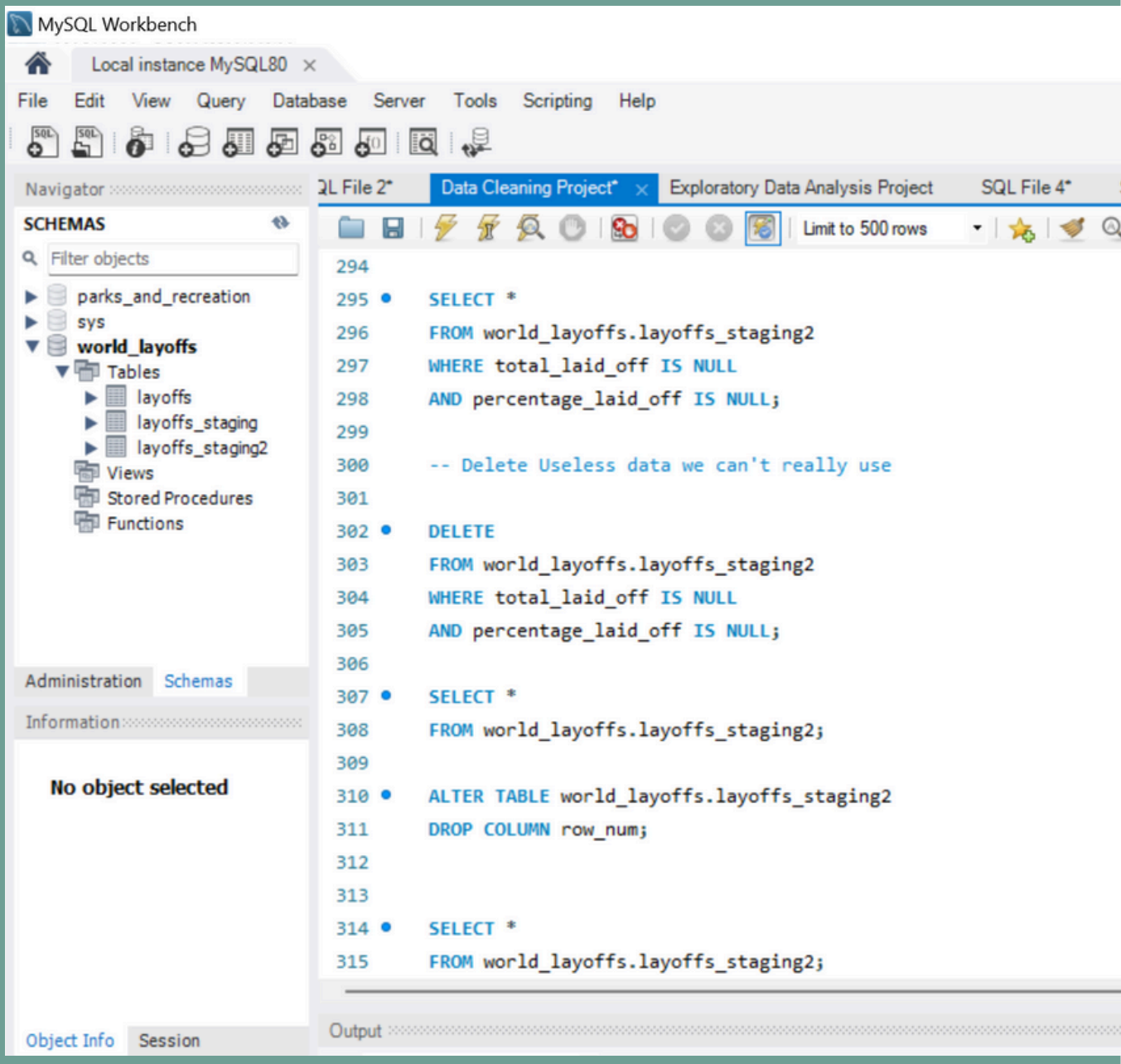



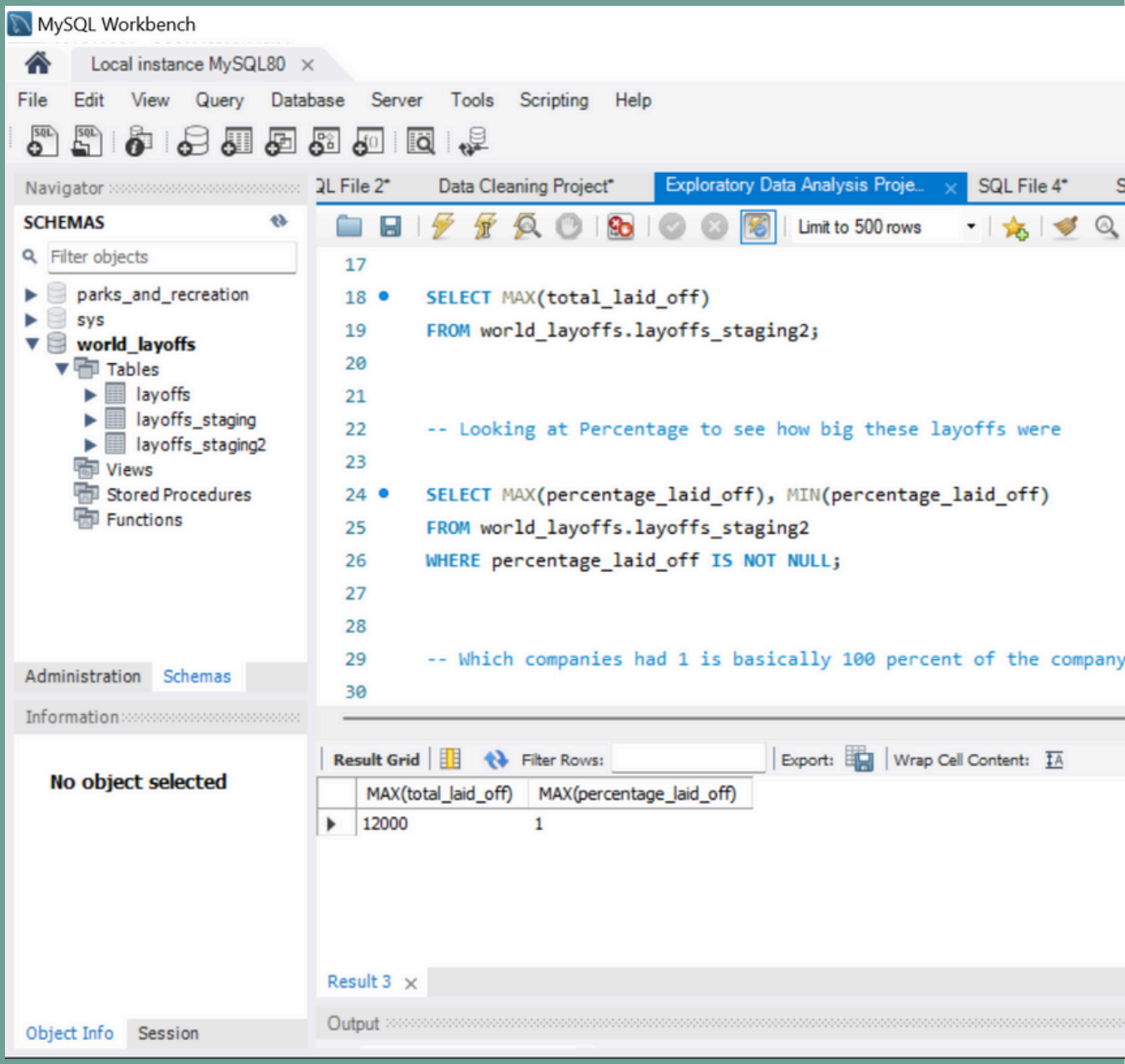
# Removed:

- Rows with NULL values in both `total_laid_off` and `percentage_laid_off` columns, as they were deemed unusable.
- Column `row_num` after completing the deduplication process.



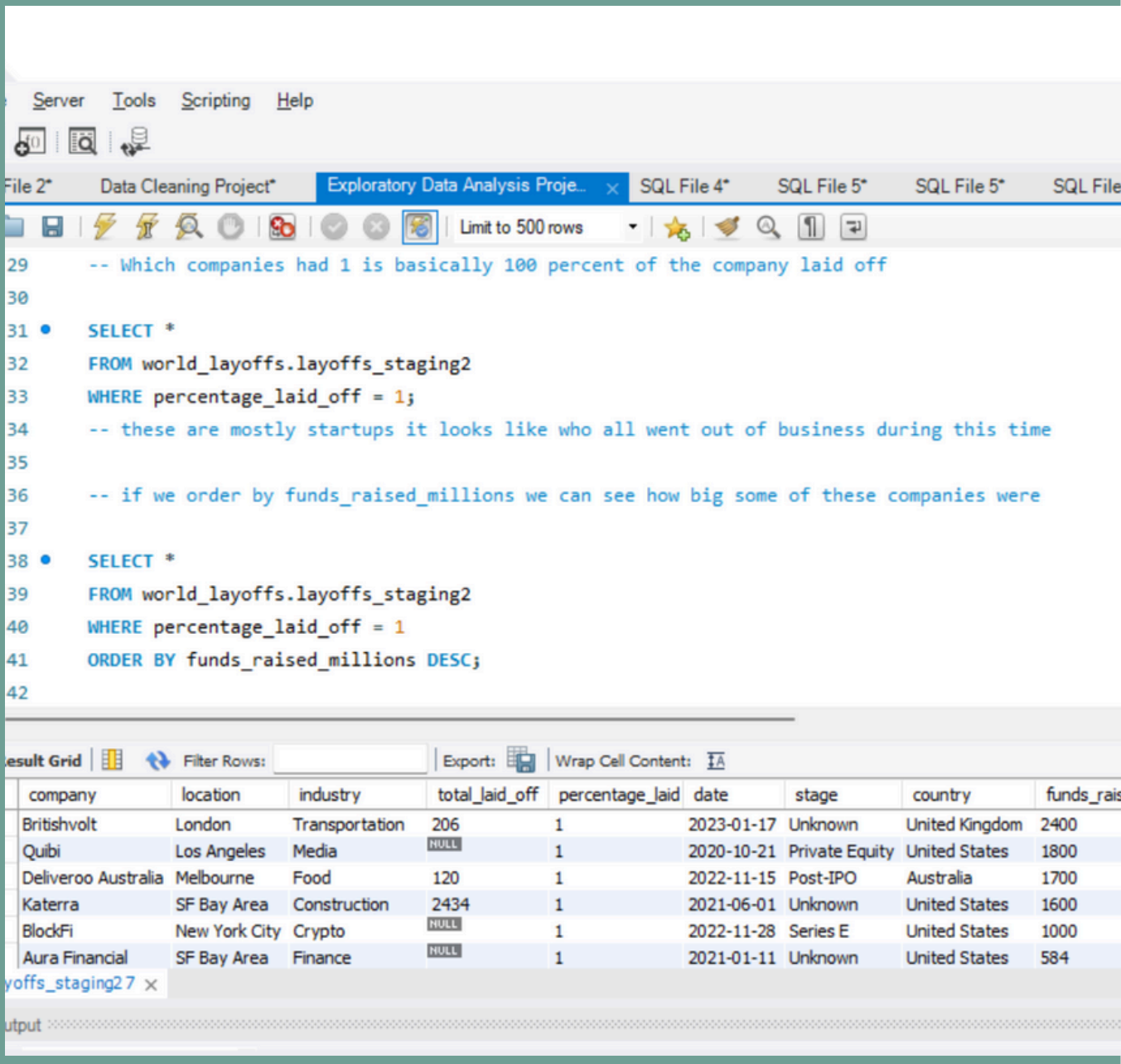
# Descriptive Analysis

Queried the maximum values of `total_laid_off` and `percentage_laid_off` to identify the highest layoffs and their impact.



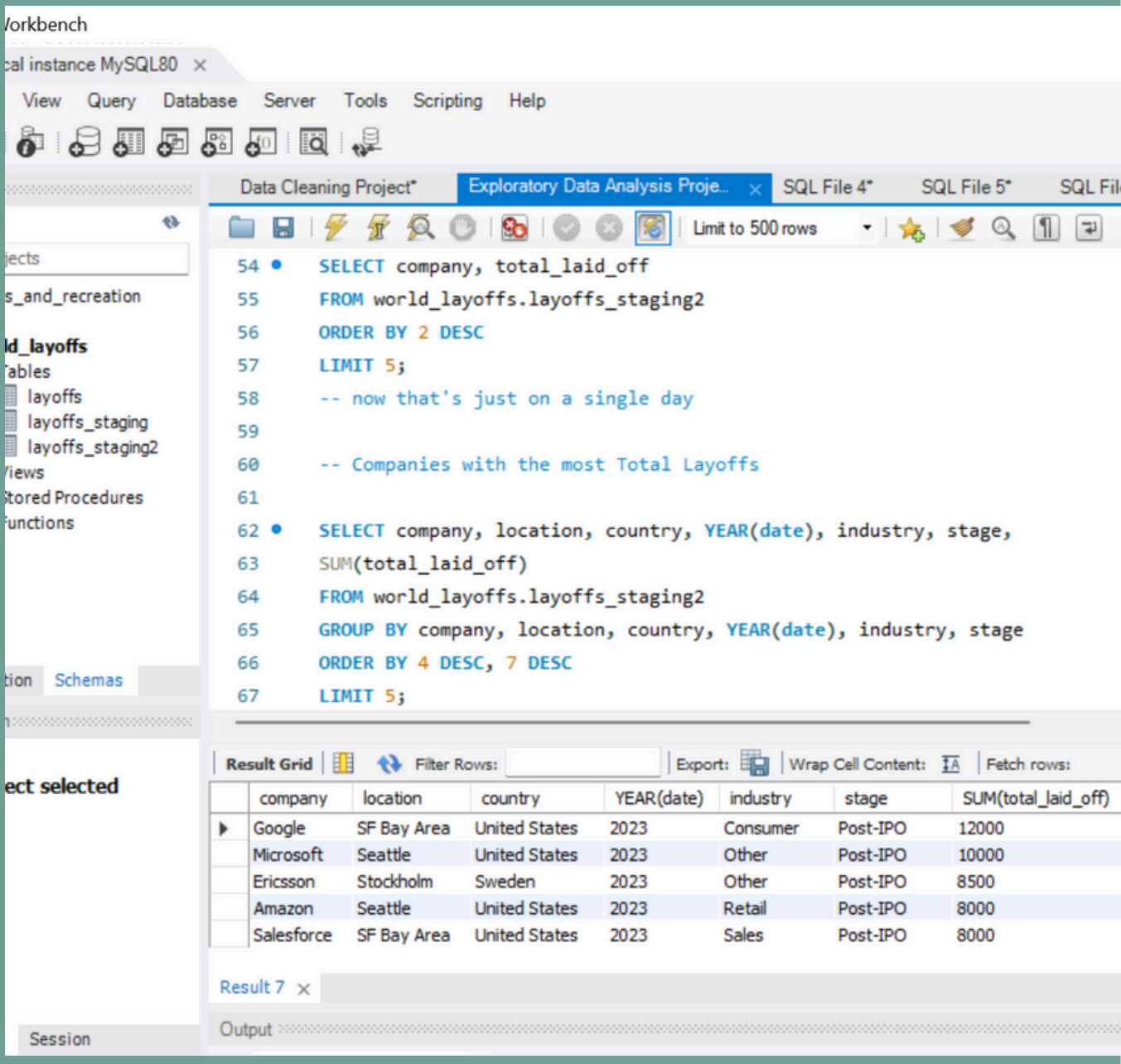
# Layoffs by Company

- Identified companies with the highest percentage of layoffs (close to 100%).
- Ordered companies by `funds_raised_millions` to see the financial impact of the layoffs.



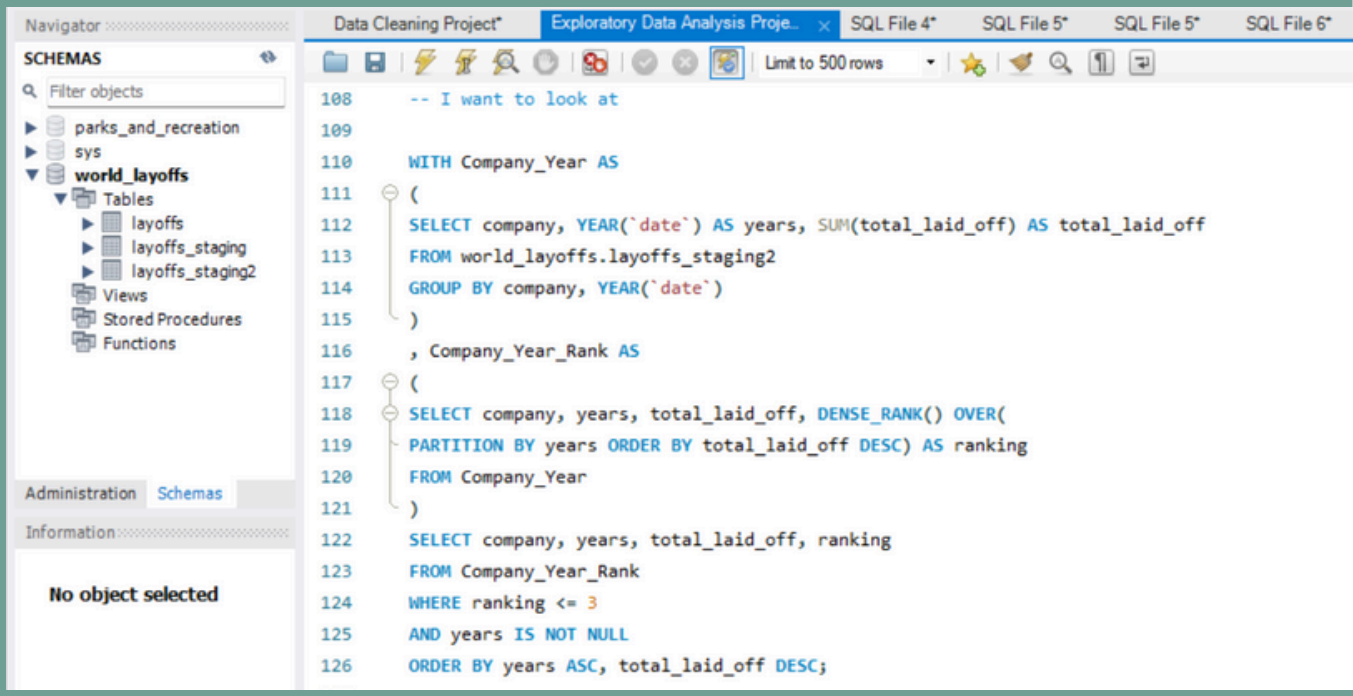
# Aggregate Analysis

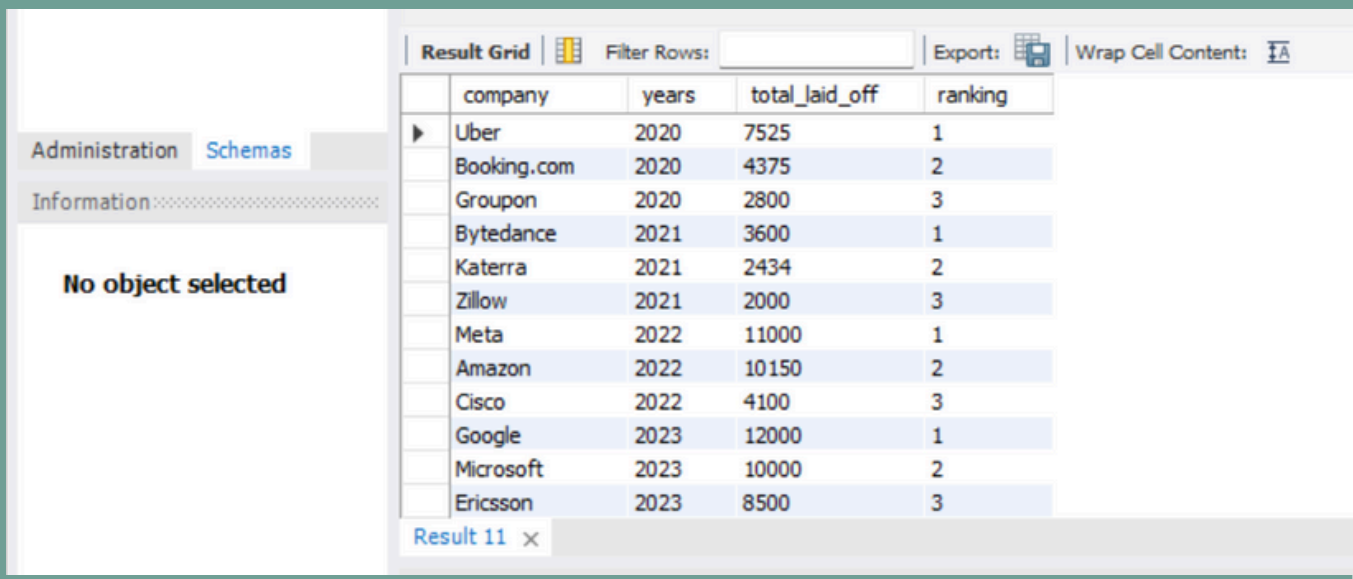
- Calculated top 5 total layoffs by company, location, country, year, industry, and company stage.
- Summarized total layoffs by grouping data by various attributes.



# Rank Analysis:

Ranked companies based on yearly layoffs using Common Table Expressions (CTEs) and tracked changes over time.




| company | years | total_laid_off | ranking |
|---|---|---|---|
| Uber | 2020 | 7525 | 1 |
| Booking.com | 2020 | 4375 | 2 |
| Groupon | 2020 | 2800 | 3 |
| Bytedance | 2021 | 3600 | 1 |
| Katerra | 2021 | 2434 | 2 |
| Zillow | 2021 | 2000 | 3 |
| Meta | 2022 | 11000 | 1 |
| Amazon | 2022 | 10150 | 2 |
| Cisco | 2022 | 4100 | 3 |
| Google | 2023 | 12000 | 1 |
| Microsoft | 2023 | 10000 | 2 |
| Ericsson | 2023 | 8500 | 3 |

# Rolling Total Analysis:

Calculated the cumulative total of layoffs per month using a CTE to observe trends and fluctuations over time.