**Employee Turnover Analysis**

**Changelog**

All notable changes to this project will be documented in this file.

**27th May, 2024**

Added

* Checked for duplicates
  + Created subquery to find duplicates
  + Specifically checked the suspicious records to confirm that they were duplicates
  + Found duplicates that matched exactly in all fields and removed them using a DELETE\_CTE
  + Created a row\_num field that displays a duplicate if it’s value is 2 or greater

Changed

* Data Standardization
  + Checked for null or empty string values for each field
  + Found that some field values were missing for industry field for companies that had more than 1 record in data
  + Fixed this by self joining the table and filling those empty cells
  + Standardized industry value ‘Crypto’ as there were other values for it
  + Fixed country field as there were periods in some values with ‘United States.’
  + Changed date format from text to date.

Removed

* Deleted rows with row\_num >= 2

**29th May, 2024**

Added

* Look at null values:
  + Observed null values for quantitative fields, but left them as they would prove useful for aggregations later on in the analysis section.

Changed

* n/a

Removed

* Dropping any columns and rows we don’t need
  + Removing rows that had all quantitative fields as null, they won’t provide much information in analysis later on.
  + Dropping row\_num column as it is no longer needed.

*Exported SQL table to .csv for analysis*

**30th May, 2024**

Added

* Added conditional formatting to all fields to check if data types and values were preserved through import process
* Created pivot table and chart for the top 10 industries with the highest percentage of laid off employees
* Created pivot table and chart for funds raised in millions by industry
* Created pivot table and chart for sum of employees laid off by companies, the top 10
* Created pivot table and chart for median percentage laid off by company stage of business life cycle
* Created pivot table and chart for funds raised in millions by each country
* Created a scoreboard chart for displaying the total number of laid off employees in the dataset

Changed

* Data Preservation
  + Changed 1 value from company field to text
  + Left 1 value in industry field as empty due to no clear information on how to proceed
  + Total\_laid\_off and percentage\_laid\_off is unchanged, empty values indicate no entries
  + Changed date cell values of 187 records to date format for future aggregations
  + 5 empty values in stage field, but we left it as such since it refers to small companies or potential startups
  + No issues with country field
  + Empty values in funds\_raised\_millions is left unchanged
* Chart Modifications
  + Modified titles, subtitles, axes titles and styling of each chart for the dashboard
  + Modified background color to light grey
  + Modified title color to dark blue
  + Highlights key data points using darker tinted colours and less important points with lighter colors
* Dashboard
  + Added All 6 visualizations

PROJECT COMPLETED