José Eduardo Zárate Aranda

Analysis of the crisis and decline of the technology industry during the post-pandemic phase. (PowerBI Dashboard Documentation): November 17, 2023

Hypothesis

Conducting a comprehensive data analysis on the mass layoffs in the technology industry in 2019-2023 is vital. This analysis will allow us to understand these layoffs' magnitude and distribution better. With the results of this analysis, it will be possible to make informed decisions, and updated on the sector's economic situation. It will be determined that the behavior of layoffs is not trivial, and in turn, describing this behavior will allow other analysts or scientists to anticipate the future possibilities of the technology industry.

Principal Objectives

Implement an interactive dashboard encompassing a comprehensive analysis since the few graphics are included on the Layoffs.FYI, the site itself needs to include more information present. This will be done to complement and improve data analysis on layoffs in the information technology industry. The door will be open for new work focused on the diagnosis of the situation, prediction, and prescription or possible decisions to overcome the current crisis.

Project Scope

The analysis of data on layoffs in the technology industry has a scope that covers different levels of application, from the operational level to the public. These application levels allow the information collected and analyzed to make informed decisions effectively in various contexts.

Scope: end user and application level of the analysis

Operative Level

At the operational level, data analysis addresses immediate, day-to-day issues related to layoffs. This involves collecting and processing data to understand when and where layoffs occur most frequently. End users at this level can be human resources managers, department heads, or base-level employees.

They use the information to manage the workforce more efficiently, make decisions about hiring and redeploying resources, and ensure that layoffs do not negatively affect daily operations.

General Public Level

In addition to the operational level, analyzing data on layoffs in the technology industry can also significantly impact the general public. This refers to the audience outside the company and the community as a whole that may benefit or be affected by the analysis results. In summary, descriptive analysis of data on layoffs in the technology industry can serve multiple levels of application, from day-to-day operational decisions to professionals who can make the most of the information available and make informed decisions that allow them to adapt to the challenges they face in a constantly changing business environment. Below, Figure 1 shows a complete list of the tables and their specific details, and Figure 2 shows the relationships between the entities.

Scope: types of analyzes performed

- **Descriptive**: Visualize the periods in which layoffs have been observed the most and those most affected by sector. Analyze the density of layoffs in large companies compared to small ones and study the frequency with which layoffs are carried out in large and small companies. Implement an interactive dashboard to understand better which companies and geographic regions have been most affected and observe the evolution of the crisis since the beginning of the post-pandemic phase **(Covered)**.
- **Diagnosis:** Determine the general and specific causes of the underlying problem regarding layoffs. Relate the causes to the possible consequences, such as the increase in discrimination and the setback in issues of inclusion and diversity in work teams in companies **(Not covered, future work)**.
- **Predictive:** A model could be added, or an analysis could be made to find hidden patterns behind layoffs and the trend over time **(Not covered, future work)**.
- **Prescriptive:** We can identify the patterns and trends behind the layoffs that have devastated the technology industry with the information obtained. This will be very useful for professionals who seek to stay informed, make assertive decisions when looking for a stable job, and at the same time identify the companies with the most instability, being essential information for the development and decision-making about the professional career of many engineers, even for future investors, who look for companies with healthy finances and at the same time socially responsible **(Not covered, future work)**.

Transformations: October 31, 2023, 9:16 p.m.

Minimal transformations were implemented using Power BI to develop the descriptive dashboard focused on 'Tech Layoffs.' It is important to note that all tables were adequately cleaned and normalized in Excel beforehand. Additionally, Power BI identified and established the relationships between the tables precisely as they were initially defined. A thorough review was performed to confirm that the delete empty rows feature did not affect the tables, which were free of any anomalies requiring additional adjustments.

Database description; Main layoffs table

The database is made up of 19 tables, of which one is the main table, two are complementary tables, and another two are bridge tables that join the main table with the complementary ones through the **company** dimensional table, belonging to a total of 14 tables—dimensions resulting from normalization.

This section will detail the content of the **layoffs** table, which is bridged to the other two main tables, migrations and **top50stats** columns, that hold different information about a company. Here is a more detailed description:

Qualitative Information (5 columns):

- **Company Name**: The official company name.
- **Country:** The country where the company is located.
- **Office's Location**: The specific address of the company's offices.
- **Industry**: The economic sector to which the company belongs (for example, technology, health, etc.).
- **Level or Stage of the Company:** The phase the company is in (e.g. startup, growing, established, etc.).

Quantitative Information (4 columns):

- **Layoff Lifting Date:** The date on which the layoffs were recorded in the company.
- **Number of Layoff:** Amount of laid off employees.
- **Layoffs Percentage**: The percentage of laid-off employees relative to the company's total number of active employees.
- **Raised funds:** The amount of money the company has raised through different means (e.g., investments, loans, etc.).

These records add up to 3022 observations/rows, and each day, they accumulate in the link Layoffs.fyi - Tech Layoff Tracker and Startup Layoff Lists. However, for the preparation of this analysis, the data compiled in a file separated by commas was obtained through the website of Layoffs Dataset | Kaggle,

made by Roger Lee with username SWAPTR, who sifted the Airtable table published on the Layoffs.fyi site. In this way, the CSV file is available on Kaggle with the license for open domain data http://opendatacommons.org/licenses/odbl/1.0/. However, it does not receive the daily update as is the case with the primary source on Layoffs.fyi, but is updated weekly. This work will analyze the data screened until September 7, 2023. Although the first records began on March 11, 2023, the date on which the start of the pandemic was declared.

Database description; complementary tables

To complement the main data set, 2 data sets related to the topic were added, thus completing the database for the project; however, the same data will only be used to build an introduction to the main analysis and, in turn, complete it. The complementary data sets are:

Complementary table *migrations*

1. The first set of complementary data consists of 5 columns with 259 observations, the majority of which are qualitative information that describes recent changes or technology transfers involving migration to new programming languages or complete tools used in particular companies worldwide. Below is a more detailed description:
2. **Qualitative Information** (5 columns):
    - **Company Name**: The official company name.
    - **URL**: Web link to the source of information.
    - **From:** The column mentions the name of the technology, tool, and programming language used before the migration.
    - **To:** The column mentions the name of the technology, tool, and programming language used now after the migration.
3. **Quantitative Information** (4 columns):
    - **Year of migration**: Year in which the technology change was made.

      Below is the link to the dataset, updated through September 5, 2023, covering data from web screening through March 2, 2023.
    - **Migration or Technology Transfer in Companies:**

Tech Migrations Dataset | Kaggle

Complementary table *top5ostats*

1. The second set of complementary data consists of 10 columns with 50 observations applicable to the U.S. and leading companies in the business technology sector. Here is a more detailed description:

Qualitative Information (5 columns):

- **Company Name**: The official company name.
- **Industry or Sector**: The economic sector to which the company belongs (for example, technology, health, etc.).
- **Geographic Status** The state of the U.S. in which the bases are located.
- **Name of Stock:** The name of the company's shares.

2. **Quantitative Information** (4 columnas):
   - **Foundation Year:** The year the company was founded.
   - **Annual Profit:** The annual profit between 2022 and 2023 in billions of dollars
   - **Laid Off Percentage**: The percentage of laid-off employees to the company's total number of active employees.
   - **Market Value**: The trillions of dollars capitalized in the public company to the stock market.
   - **Annual Revenue**: The amount in billions of dollars that I earned during 2022 to 2023 to the company.
   - **Company Size**: Whole number of employees that the company has.

The link to the dataset is provided below, covering data up to March 23, 2023, updated six months ago as of this writing. Furthermore, the data set comes from a web screening with Python programming methodologies.

Statistics and Relevant Data of Technology Companies:

USA Tech Companies Stats | Kaggle

Description of the database, bridge, and dimensional tables

The ***layoffs*** primary key ***index*** was added to identify each record uniquely; dimensional tables were also added for the aforementioned qualitative fields, where there is the relationship between ***'qualitative_column'*** and ***'identifier_of_qualitative_column'***, which consists of ascending integers. Thus, in the main table ***layoffs*** the columns were replaced by their respective identifiers that refer to 5-dimensional tables.In the same case for the table ***migrations***, the primary key ***index_m*** was added to identify each record uniquely. In the same way, four dimensional tables were generated, and the qualitative fields of the **migrations** table were replaced by their identifiers.

For the table ***top50stats***, adding a primary key was unnecessary since the column ***company_t*** had exemplary characteristics of a candidate key, which became the primary key. Their identifiers replaced The qualitative fields, including the primary key, which became ***idcompany_t***.

Regarding the bridge tables ***layoffs_migrations_bridge*** and ***layoffs_top_bridge***, companies were connected in common between the main table ***layoffs*** and the complementary tables ***migrations*** and ***top50stats***; it was performed with the dimensional tables of ***company_m*** y ***company_t*** that connect with ***idcompany*** which corresponds in its entirety to the main table ***layoffs***.

Description of tables, columns and data types in the database. Diagram of entity relationship, and relational type.

Below is a general list of tables that summarizes the general structure:

layoffs

It is a table that contains records of layoffs during and after the pandemic; it also contains relevant data such as the company, location, industry, stage of the company, and country where the mass layoff occurred.

company

The dimensional table of companies defines the equivalence between identifier-element.

location

The dimensional location table defines the equivalence between the identifier-element

industry

The dimensional table of industries defines the equivalence between identifier-element

stage

A dimensional table of stages defines the equivalence between identifier-element

country

The country dimensional table defines the equivalence between the identifier-element

layoffs_migrations_bridge

It is a bridge table that unifies the companies in common with the **layoffs** and **migrations** table, it bridges the companies that laid off and migrated some type of technology or software tool to a different one. Through the dimensional tables of both main tables in the company's field.

migrations

It is a table that contains records of technological migrations in terms of software, it includes the company, as well as the native technology and the one that was updated.

company_m

The company dimensional table defines the equivalence between identifier-element.

url

The company dimensional table defines the equivalence between identifier-element.

from

Dimensional table of technology/tool to which it is migrated, element-identifier.

to

Native technology/tool dimensional table, element-identifier

layoffs_top_bridge

It is a bridge table that unifies the companies that are part of the top 50 in the technology sector with those that have laid off people during and after the pandemic, **layoffs**, and **top50stats** through the dimensional tables of both main tables in the companies field.

top50stats

company_t

A dimensional table of companies, defines the equivalence between identifier-element

industry_t

A dimensional table of industries, defines the equivalence between identifier-element

sector

A dimensional table of sectors, defines the equivalence between identifier-element

hqstate

A dimensional table of states, defines the equivalence between identifier-element

stockname

A dimensional table of public name, element-identifier equivalence

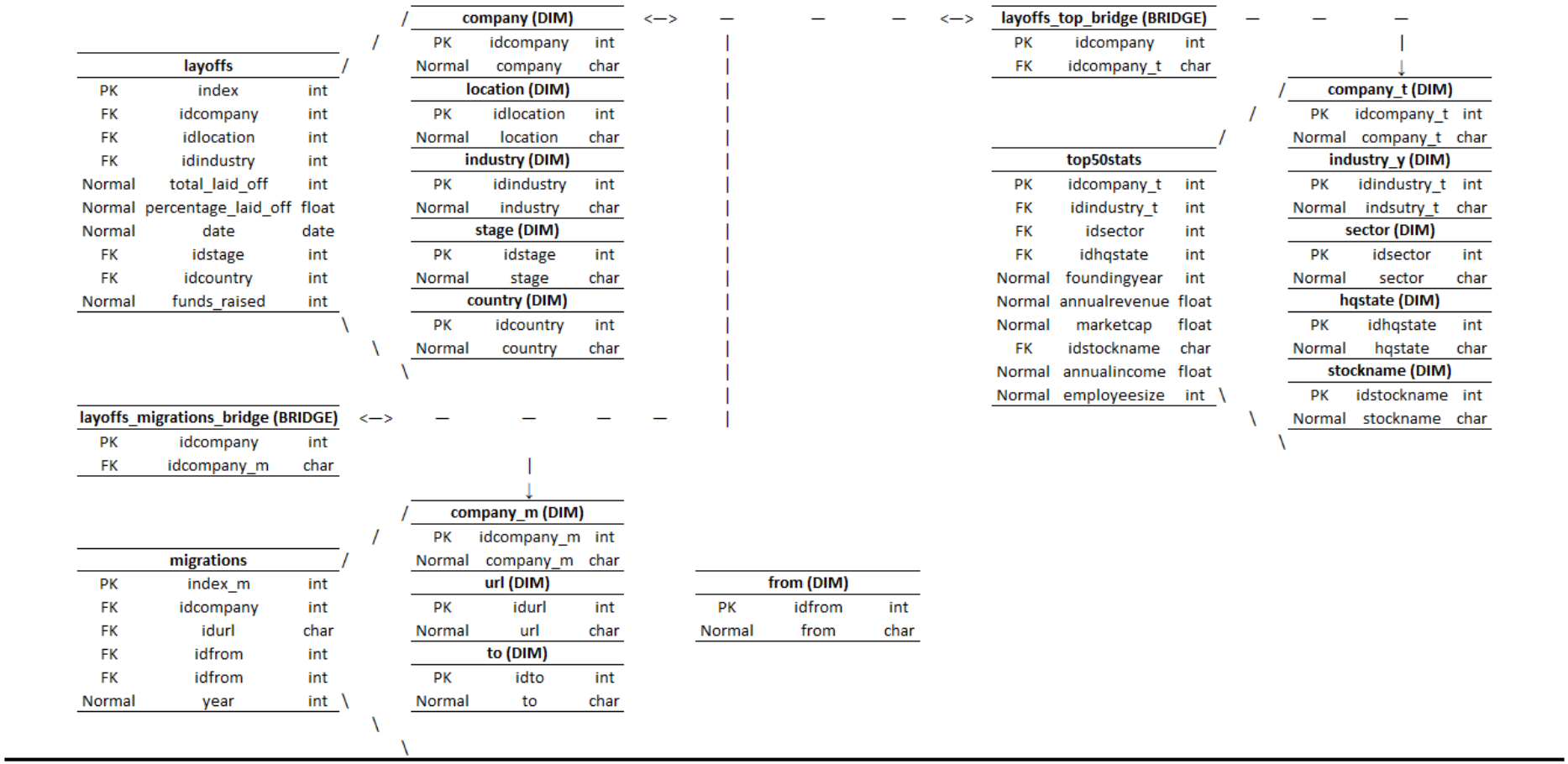

Figure 1. Descriptive diagram where the tables are listed with their own definition of primary and foreign keys, as well as their relationship, and the use of bridge tables that join data sets

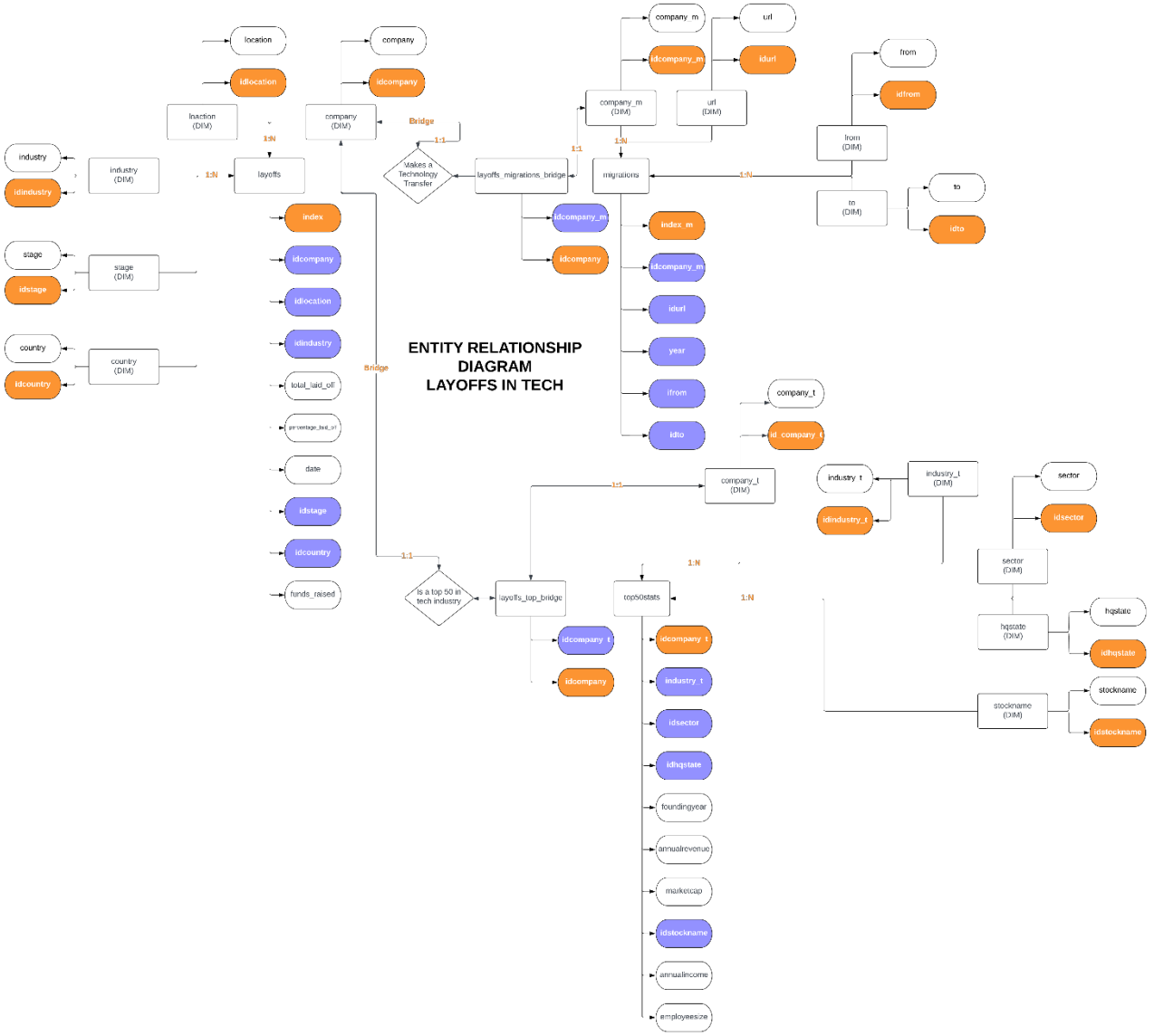

Figure 2. Entity relationship diagram where the tables are listed with their own definition of primary and foreign keys, as well as their relationship, and the use of bridge tables that join data sets