

# **HDSC Spring'23 Premiere Project**

## **Documentation by Team Neural Network**

### **Topic: Technology company layoff (2022–2023)**

#### **Project Description:**

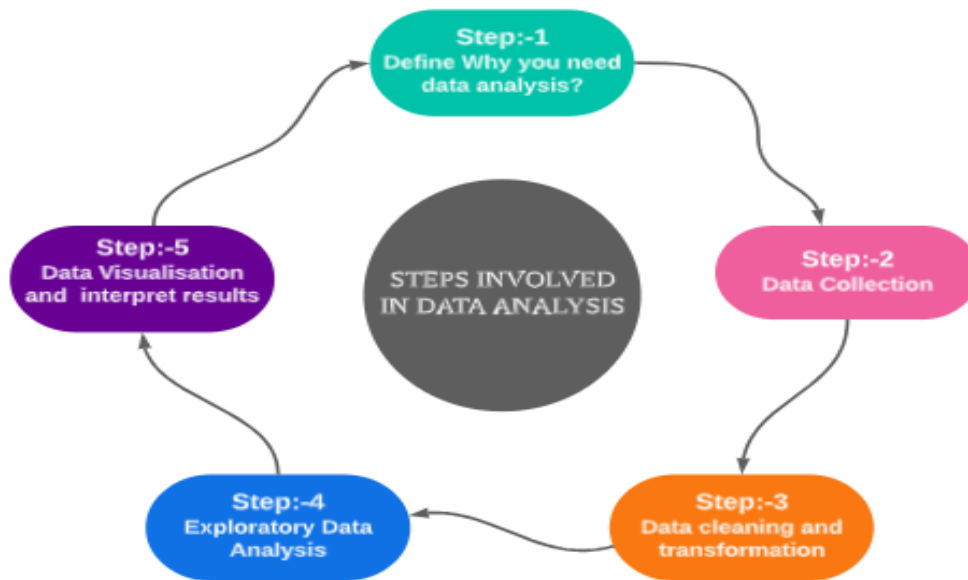
From Amazon, Microsoft, Google, to Wayfair, the technology industry has been shaken by massive layoffs since mid-2022. Use this data to gain insights on technology industry trends and make informed decisions for your career or business.

#### **Aim and Objectives:**

The aim of the project is to analyze and gain insights on technology company layoff trends with a focus on Data Analysis. The objectives are to address some questions as follows:

- 1) Total layoffs per month
- 2) Total layoffs per year
- 3) Average layoffs per months
- 4) How the IPO company status affected the layoffs
- 5) Which industries were most affected by the layoffs?
- 6) Which industries were least affected by the layoffs?
- 7) Which companies were most affected by the layoffs?
- 8) Which companies were least affected by the layoffs?
- 9) Which locations had the most layoffs?
- 10) Which locations had the fewest layoffs?

## Process Flow:



## Data Collection:

We used two datasets:

- a) Kaggle Dataset which we had originally been given

Kaggle-Link:

<https://www.kaggle.com/datasets/salimwid/technology-company-layoffs-20222023-data>

- b) Open dataset from the Layoff.fyi website, which was downloaded as a PDF file and then converted to csv

<https://layoffs.fyi/>

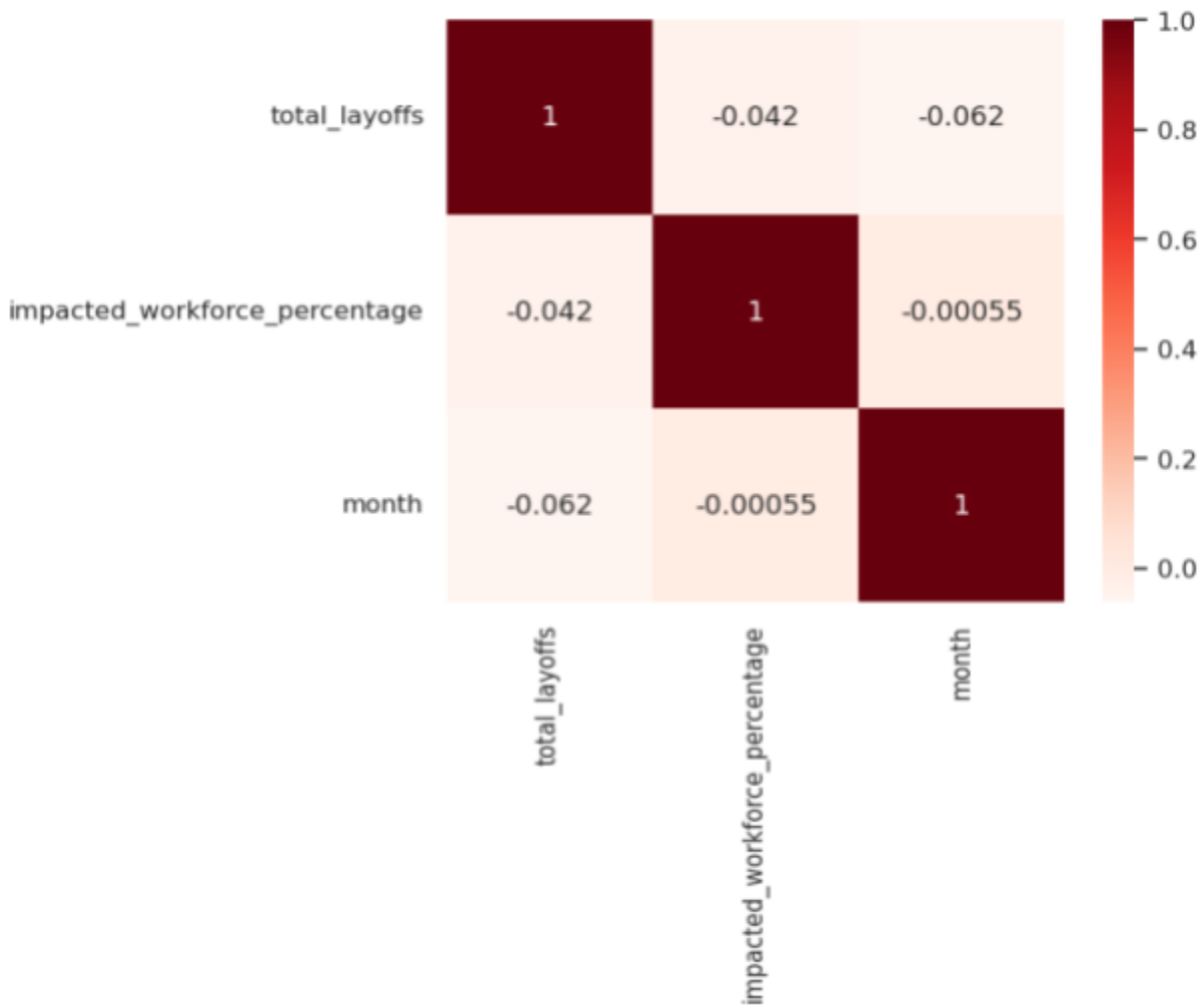
## Data Preprocessing:

The initial step in data analysis is to curate and organize the data in a suitable format to facilitate exploration and analysis. This involves preparing the data by cleaning, transforming, and structuring it in a manner that enables effective investigation of patterns and relationships within the data.

- **Data Collection-** gathering data from Kaggle and Layoffs.fyi websites
- **Data Cleaning-** starts with the Kaggle tabular dataset containing information on 477 technology companies that have announced layoffs or are known to have laid off their employees between mid-2022 to 2023. The Dataset had very few entries and many of the attributes were found empty. Apart from this the data was inconsistent and extra information attributes had to be removed as nothing could be inferred from it.
- **Data Transformation-** few attributes were converted to Datetime and integer from Object type. Missing values were dropped, and to supplement the remaining records, Web scraping was performed. The dataset was converted from PDF to a CSV (comma separated values) file. There is a possibility of data loss from the file conversion. After cleaning the second dataset, it was concatenated with the first dataset.
- **Data Labeling-** We added labels to a few features to understand the data in depth and gain better analysis and results. E.g:- Status Label was categorized into private and public to understand which type of company laid off more people.

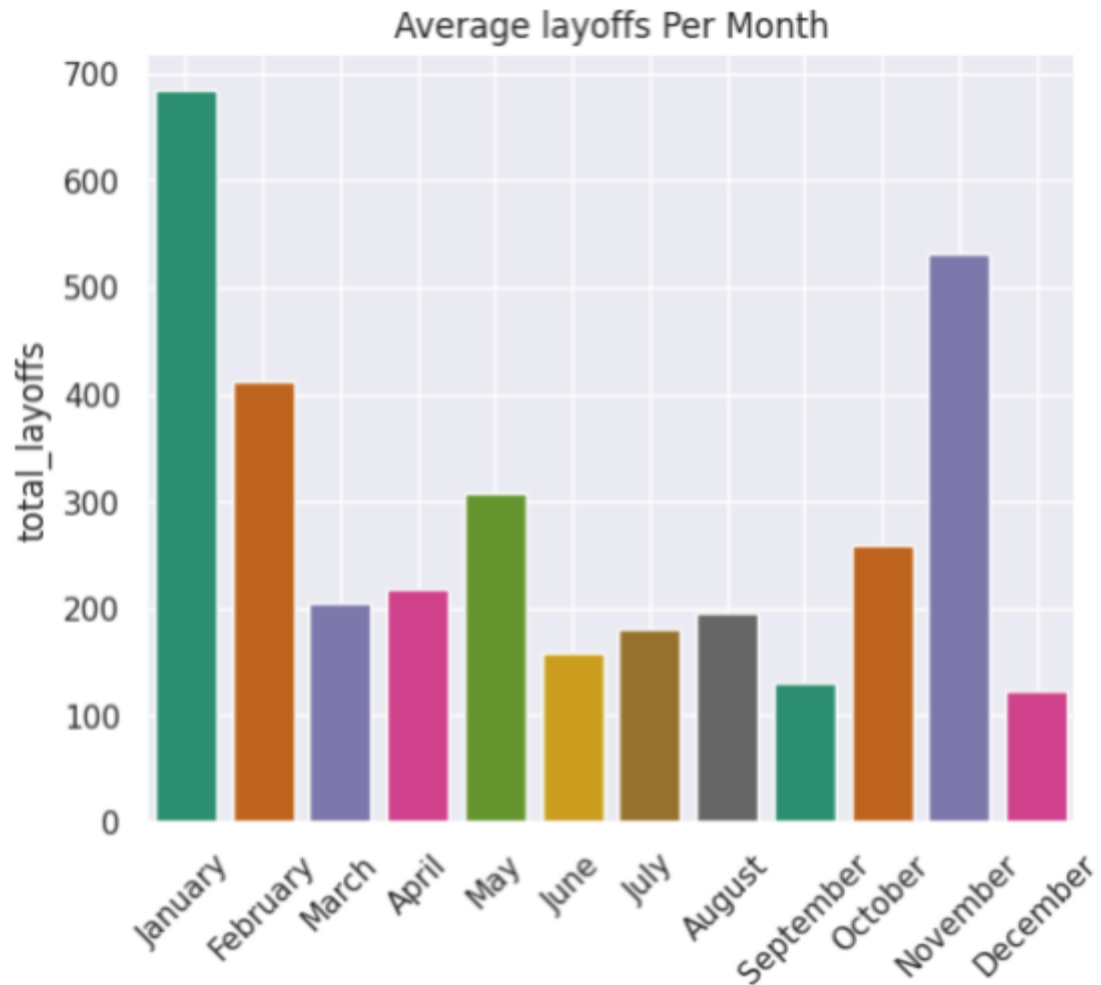
## **Exploratory Data Analysis:**

The visualization was done to better understand and gain insights from the data. A heat map was used to determine the correlation between the variables. From the heatmap below, it can be seen that there is neither a positive nor a negative correlation between the variables.



**Figure1:-** Correlation Heatmap

An analysis of the average layoffs per month was done, and it was seen that January is the month with the most layoffs.



**Figure 2:- Average Layoffs per month**

The rest of the visualizations that meet the objectives are in the dashboard found [here](#):

## Results

From our analysis, we found the following insights across all four years:

1) The total number of layoffs per month

- Jan - 114,000
- Feb - 39,000
- Mar - 13,000
- Apr - 46,000
- May - 48,000

- Jun - 32,000
- Jul - 19,000
- Aug - 18,000
- Sep - 6,000
- Oct - 21,000
- Nov - 75,000
- Dec - 11,000

2) Total layoffs per year

- 2020 - 60,000
- 2021 - 7,000
- 2022 - 183,000
- 2023 - 192,000

3) Average layoffs per month

- Jan - 684
- Feb - 412
- Mar - 205
- Apr - 218
- May - 309
- Jun - 169
- Jul - 180
- Aug - 197
- Sep - 130
- Oct - 259
- Nov - 532
- Dec - 123

4) How the IPO company status affected the layoffs

- Public- 217,000
- Private- 159,000
- Acquired- 33,000
- Unknown- 32,000
- Subsidiary- 1,000

5) Which five industries were most affected by the layoffs?

- Retail 42,000
- Consumer 38,000
- Other 37,000
- Transportation 31,000
- Finance 26,000

6) The industries that were least affected by the layoffs are

- Artificial intelligence- 5
- Privacy Tech, Marketing Tech- 7
- Beauty Tech- 11
- Saas, Analytics- 13
- Media, Content Creators- 13

7) The top five companies that were most affected by the layoffs are

- Amazon- 36,000
- Meta- 22,000
- Google-12,000
- Microsoft-10,000
- Philips- 10,000

8) Which companies were least affected by the layoffs?

- Branch- 3
- TutorMundi- 4
- Spyce- 4
- Woven- 5
- Synergysuite- 5

9) Which locations had the most layoffs?

- SF Bay Area- 121,000
- Seattle- 51,000
- New York City- 21,000
- San Francisco-
- Amsterdam- 16,000

10) Which locations had the fewest layoffs?

- Boise- 7,000
- Chicago- 7,000
- Sao Paulo- 8,000
- Austin- 10,000
- Boston-

## **Conclusion and Recommendation**

The conclusion from our analysis across all the years from 1062 companies analyzed is:

- ➔ January is the month with the most layoffs, with a total of 114,000
- ➔ The year with the highest layoff is 2023, with a total of 192,000
- ➔ The average layoff per month is between 120 - 700, depending on the month
- ➔ Companies whose IPO status is Public had the highest number of layoffs, with a record of 217,000
- ➔ The Retail industry is most affected by the layoffs, with a record of 42,000
- ➔ The Artificial Intelligence industry is least affected, with only 5 layoffs for all the years
- ➔ Amazon is the company with the most layoffs of 36,000
- ➔ Branch is the company with the fewest layoffs of 3
- ➔ SF Bay Area is the location with the highest number of layoffs of 121,000
- ➔ Boise is the location with the lowest number of layoffs of 7,000

We recommend that further data collection be done so as to later on be able to have a target variable for prediction purposes.