



Layoffs Data Analysis

IE6600 Computation and Visualization of Data

Prof. Sivarit Sultornsanee

Group 3:

- 1. Atharva Mahajan**
- 2. Ruchirkanth Gandikota**
- 3. Prabhat Chanda**
- 4. Shashank Farkase**

Index

No.	Title	Page no.
1	Introduction	3
2	Summary	4
3	Data source	4
4	Data dictionary	5
5	Data collection and Data cleaning	5
6	Data Visualization	7
7	Limitation and future scope	12

1. Introduction

In March 2020, with the initial reports of COVID-19 cases, it was impossible to predict the extensive disturbance that ensued, resulting in a global standstill. This had a widespread effect on various industries and economies, resulting in a surge of job cuts. According to the U.S Bureau of Labor Statistics (BLS) Current Employment Statistics (CES) survey, employment in the USA witnessed a decline of 9.4 million in 2020. The significant loss of jobs caused a major upheaval.

The goal of our project is to conduct a comprehensive analysis of worldwide job layoffs that transpired from 2019 to 2022. Although Covid-19 was a major contributor to these layoffs, we intend to investigate further to uncover the industries and company scales that experienced the most significant impact. Additionally, we plan to scrutinize the layoff trends across diverse global regions and pinpoint the regions with the highest frequency of job cuts. Our ultimate objective is to identify the key determinants that increase a company's susceptibility to layoffs.

The dataset comprises 1056 tech companies listed as rows which have seen 164511 employees laid off. In 2022 alone 930 companies laid off 146407 employees.

To uncover the devastating impact of recent layoffs on industries we would investigate into the following questions:

- The number of individuals affected by layoffs in various sectors and countries?
- The trend of layoffs in different aspects over recent years?
- The country hit hardest by job loss?
- The month-by-month breakdown of layoffs to reveal patterns and potential future risks?
- What was the layoff trend from 2020-2023 for the various industries?

We used the Pandas library of Python to manipulate and edit our dataset and we used Tableau to visualize the data and create dashboards.

2. Summary

During the data visualization phase of our analysis, we noticed that the two countries most severely impacted by layoffs were the United States of America and India. These two nations had the highest recorded number of layoffs across the globe. Additionally, it was observed that the peak in the layoff rate occurred around the beginning of February 2020, which coincided with the outbreak of the Covid-19 pandemic.

Furthermore, upon conducting an in-depth industry-wise analysis of the data, we discovered that the retail and transportation industries were the most adversely affected industries across the world, with the highest number of layoffs reported in these two sectors.

Lastly, we delved into the company-wise layoff data to identify the corporations that were most affected by this phenomenon. Our analysis revealed that Amazon had the highest number of layoffs amongst all the companies that were included in our dataset.

3. Data Source

We acquired our dataset from Kaggle – Layoff 2020 and the dataset link is provided ([link](#)). The dataset that we had at our disposal during the analysis phase comprised of a staggering number of rows, almost 31861 to be precise, and contained a wide range of data fields, numbering 12 in total. The fields of information included, but were not limited to, the name of each company that had gone through a layoff, the country where the company is headquartered, the total number of employees who were laid off, the specific date on which the layoffs occurred, the percentage of the company's workforce that was affected, the industry sector to which the company belonged, the developmental stage of the company at the time when the layoffs happened, the total amount of funds that the company had raised, a comprehensive database of all the employees who were laid off, and lastly, a source link for all the layoff data that was compiled for this dataset.

Company	Location_HQ	Industry	Laid_Off_Count	Percentage	Date	Funds_Raised	Stage	Country
Yggon	Bengaluru	Crypto	100	0.2	2023-02-2	451	Unknown	India
Gate	Bengaluru	Other	200	0.3	2023-02-2	79	Series B	India
iblocks	New York City	Crypto	30	0.05	2023-02-2	1000	Series E	United States
de	Sydney	Other	8	0.28	2023-02-2	10	Seed	Australia
ola	Chennai	Food		1	2023-02-2	9	Series A	United States
pper Cash	SF Bay Area	Finance	100	0.33	2023-02-1	302	Series C	United States
uSign	SF Bay Area	Sales	680	0.1	2023-02-1	536	Post-IPO	United States
Interactive	SF Bay Area	Other	400	0.2	2023-02-1	62	Acquired	United States
RealReal	SF Bay Area	Retail	230	0.07	2023-02-1	356	Post-IPO	United States
artsheet	Seattle	Other	85	0.03	2023-02-1	152	Post-IPO	United States
c	Tel Aviv	Marketing	370	0.06	2023-02-1	58	Post-IPO	Israel
viceTitan	Los Angeles	Sales	221	0.08	2023-02-1	1100	Series G	United States
zn	Sao Paulo	Finance	210	0.09	2023-02-1	720	Series D	Brazil
italOcean	New York City	Infrastructure	200	0.11	2023-02-1	491	Post-IPO	United States
inkl	New York City	Support	100	0.04	2023-02-1	429	Post-IPO	United States
krun	Sydney	Food		0.2	2023-02-1	86	Series A	Australia
mentive	SF Bay Area	Marketing		0.14	2023-02-1	1100	Post-IPO	United States
kle	Boise	Other		0.15	2023-02-1	148	Series C	United States
arious Surgical	Boston	Healthcare		0.14	2023-02-1	185	Post-IPO	United States
nmmerceHub	Albany	Retail	371	0.31	2023-02-14 00:00:00		Acquired	United States
kerEarth	SF Bay Area	HR		0.08	2023-02-1	11	Series B	United States
bleCare	Bengaluru	Healthcare		0.7	2023-02-1	40	Series B	India
amy	SF Bay Area	Education		0.1	2023-02-1	311	Post-IPO	United States
ilio	SF Bay Area	Other	1500	0.17	2023-02-1	614	Post-IPO	United States
ctric	New York City	Other	141	0.25	2023-02-1	212	Series D	United States
bot	Boston	Consumer	85	0.07	2023-02-1	30	Acquired	United States
iavo	Milan	Real Estate		0.3	2023-02-1	708	Unknown	Italy
ladin	Jakarta	Transportation	360	0.11	2023-02-1	138	Series B	Indonesia
leLift	New York City	Marketing	100	0.2	2023-02-1	16	Acquired	United States
etti Computing	SF Bay Area	Hardware		0.28	2023-02-1	298	Post-IPO	United States
oo	SF Bay Area	Consumer	1600	0.2	2023-02-0	6	Acquired	United States
fits Market	Philadelphia	Food	649	0.33	2023-02-0	526	Series C	United States
ivern	London	Food	350	0.09	2023-02-0	1700	Post-IPO	United Kingdom

4. Data Dictionary

In addition to our primary analysis, we took the initiative of generating a data dictionary to provide a detailed and comprehensive overview of the various data elements that we were utilizing or capturing. The data dictionary was essentially a compilation of names, definitions, and attributes that described each data element in detail, thus facilitating a better understanding of the data that was being used. The purpose of the data dictionary was to ensure that all team members involved in the project had a clear and consistent understanding of the data, which was essential for accurate analysis and interpretation of results. In summary, the data dictionary served as a valuable reference guide that ensured consistency and accuracy throughout the project.

Column	Content	Column	Content
Company	Name of the company	Location	Location of company, city
Industry	Categorical Variable	Percentage	Percent of total employees laid off
Date	Date of layoff	Source	The source from which data is generated
Funds Raised	Total funds raised	Stage	Categorical variable, indicates stage of company
Date added	The date on which the data is added to database	Country	Country where company is based
Laid off count	Number of employees laid off	List of employees laid off	Names of employees

5. Data collection and data cleaning

Data Collection: -

The data for the project was acquired from the data source- [Link](#). The data which was collected was not cleaned and needed the data cleaning steps.

Data Cleaning and Processing: -

The columns which are not useful for analysis and visualization have been dropped.

The columns that are dropped are:

- 'Source'
- 'Date_Added'
- 'List_of_Employees_Laid_off'

The number of null values in the data set were checked and below are the variables that had the null values:

- Laid off count
- Percentage
- Funds_Raised

The first approach to handling the null values was to drop all the records that had null values in all the three above mentioned columns. The data will have no meaning if these three columns are missing.

The rest of the null values were replaced with 0 signifying that those companies did not have any layoffs.

The finalized dataset has been saved as a new excel file which can now be used for effective visualization.

Cleaned Data Set

	Company	Location	Industry	Laid_Off_Count	Percentage	Date	Funds_Raised	Stage	Country
0	Swiggy	Bengaluru	Food	600	0.1	19-01-2023 00:00	3600	Unknown	India
1	WeWork	New York City	Real Estate	300	0	19-01-2023 00:00	22200	IPO	United States
2	Hubilo	SF Bay Area	Other	115	0.35	19-01-2023 00:00	153	Series B	United States
3	Riot Games	Los Angeles	Consumer	46	0	19-01-2023 00:00	21	Acquired	United States
4	Magnite	Los Angeles	Marketing	0	0.06	19-01-2023 00:00	400	IPO	United States
5	Personalis	SF Bay Area	Healthcare	0	0.3	19-01-2023 00:00	225	IPO	United States
6	Wallbox	Barcelona	Energy	0	0.15	19-01-2023 00:00	167	IPO	Spain
7	Microsoft	Seattle	Other	10000	0.05	18-01-2023 00:00	1	IPO	United States
8	Sophos	Oxford	Security	450	0.1	18-01-2023 00:00	125	Acquired	United States
9	Teladoc Health	New York City	Healthcare	300	0.06	18-01-2023 00:00	172	IPO	United States
10	Vroom	New York City	Transportation	275	0.2	18-01-2023 00:00	1300	IPO	United States
11	8x8	SF Bay Area	Support	155	0.07	18-01-2023 00:00	253	IPO	United States
12	Benevity	Calgary	Other	137	0.14	18-01-2023 00:00	69	Unknown	United States
13	Inspirato	Denver	Travel	109	0.12	18-01-2023 00:00	179	IPO	United States
14	Jumpcloud	Boulder	Security	100	0.12	18-01-2023 00:00	416	Series F	United States
15	nCino	Wilmington	Finance	100	0.07	18-01-2023 00:00	1100	IPO	United States
16	Starry	Boston	Other	100	0.24	18-01-2023 00:00	260	IPO	United States
17	Hootsuite	Vancouver	Marketing	70	0.07	18-01-2023 00:00	300	Series C	Canada
18	Addepar	SF Bay Area	Finance	20	0.03	18-01-2023 00:00	491	Series F	United States
19	80 Acres Farms	Cincinnati	Food	0	0.1	18-01-2023 00:00	275	Unknown	United States
20	Aiven	Helsinki	Infrastructure	0	0.2	18-01-2023 00:00	420	Series D	Finland
21	Betterfly	Santiago	Healthcare	0	0.3	18-01-2023 00:00	204	Series C	Chile
22	Pagaya	New York City	Finance	0	0.2	18-01-2023 00:00	571	IPO	United States
23	Redbubble	Melbourne	Retail	0	0.14	18-01-2023 00:00	55	IPO	Australia
24	Rappi	Buenos Aires	Food	400	0	17-01-2023 00:00	2300	Unknown	Argentina
25	Lightspeed Commerce	Montreal	Retail	300	0.1	17-01-2023 00:00	1200	IPO	Canada
26	Unity	SF Bay Area	Other	284	0.03	17-01-2023 00:00	1300	IPO	United States
27	Britishvolt	London	Transportation	206	1	17-01-2023 00:00	2400	Unknown	United Kingdom
28	Clutch	Toronto	Transportation	150	0	17-01-2023 00:00	253	Unknown	Canada
29	Exotel	Bengaluru	Support	142	0.15	17-01-2023 00:00	87	Series D	India
30	Unico	Sao Paulo	Other	110	0.1	17-01-2023 00:00	336	Series D	Brazil
31	Tul	Bogota	Construction	100	0	17-01-2023 00:00	218	Series B	Colombia
32	American Robotics	Boston	Other	50	0.65	17-01-2023 00:00	92	Acquired	United States
33	Luxury Presence	Los Angeles	Real Estate	44	0	17-01-2023 00:00	31	Series B	United States
34	RingCentral	SF Bay Area	Other	30	0	17-01-2023 00:00	44	IPO	United States
35	GoMechanic	Gurugram	Transportation	0	0.7	17-01-2023 00:00	54	Series C	India
36	PagBank	Sao Paulo	Finance	500	0.07	16-01-2023 00:00	0	IPO	Brazil
37	ShareChat	Bengaluru	Consumer	500	0.2	16-01-2023 00:00	1700	Series H	India
38	Gramophone	Indore	Food	75	0	16-01-2023 00:00	17	Series B	India
39	ClearCo	Toronto	Finance	50	0.3	16-01-2023 00:00	698	Series C	Canada
40	Dunzo	Bengaluru	Food	0	0.03	16-01-2023 00:00	382	Unknown	India

Folder Structure: -

The folder provided contains the following files:

1. Initial_Dataset.csv

This file contains data which is not cleaned and has raw data with null values.

2.Cleaned_Dataset.csv

This file contains the final data set.

3.CompVis-Project.twbx

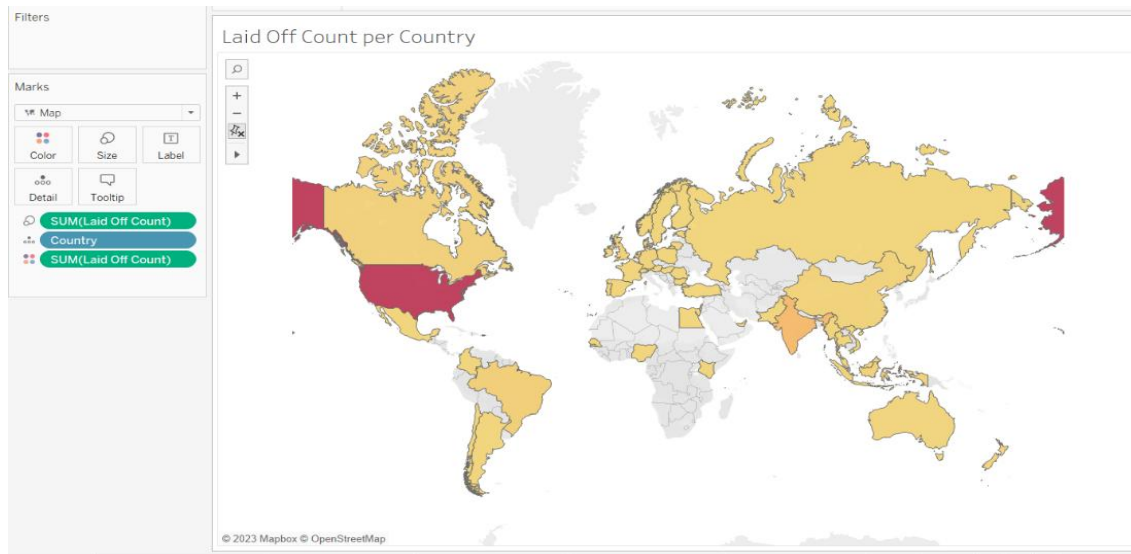
This file contains the vizulaziation and dashboard.

6. Visualizations

During the data analysis phase, we employed the data visualization tool, Tableau, to create a series of visual representations that would help us in understanding the dataset better. For this purpose, we utilized the Cleaned_Dataset.csv file as our primary data source, which contained all the relevant information that we needed to answer our research questions. Using the functionalities provided by Tableau, we created various visualizations that helped us gain insights into the dataset. We further organized these visualizations into multiple workbooks, dashboards, and stories, which provided us with a comprehensive overview of the entire dataset.

To answer our research questions, we closely examined each visualization and drew conclusions based on the insights we gained. We then displayed these visualizations in a clear and concise manner, making it easy for others to understand our findings. In summary, we used Tableau as a powerful data visualization tool to extract meaningful insights from the Cleaned_Dataset.csv file, and we presented our findings in a visually appealing and informative manner.

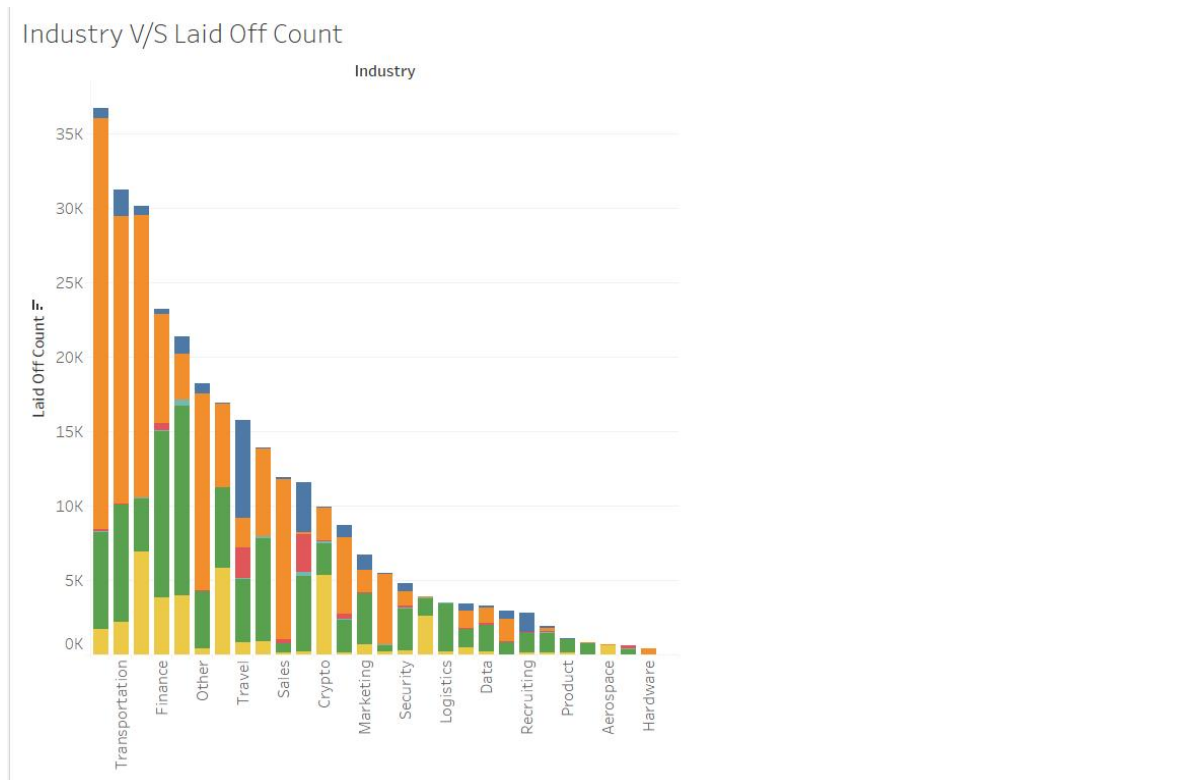
a. Which countries were worst affected by layoffs?



As a part of our data analysis, we used visualization tools to gain insights into the global layoff patterns during the Covid-19 pandemic. Our observations from the visualization revealed that the United States of America (USA) had the highest number of layoffs among all the countries that were analyzed. This means that a significant number of employees in the USA were laid off from their jobs during the Covid-19 pandemic.

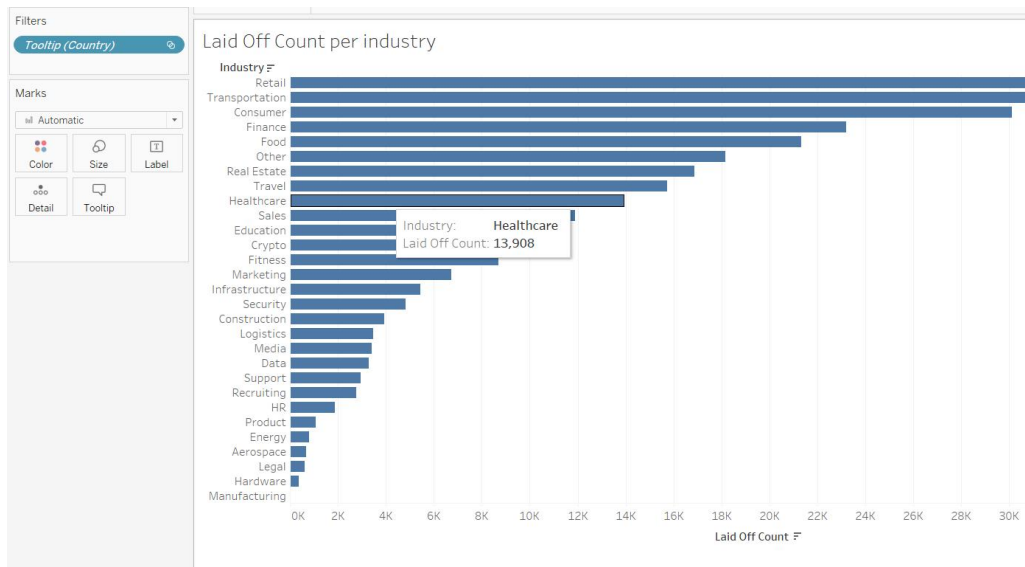
b. Which were the major organizations or sectors that were worst affected by layoffs?

During our data analysis, we examined various visualizations and observed that the largest number of layoffs occurred in Transportation sectors, primarily because they tend to have a larger workforce. However, upon further analysis, we discovered that the finance sector was also affected by layoffs and also the travel sector which was due to widespread lockdown across the world.

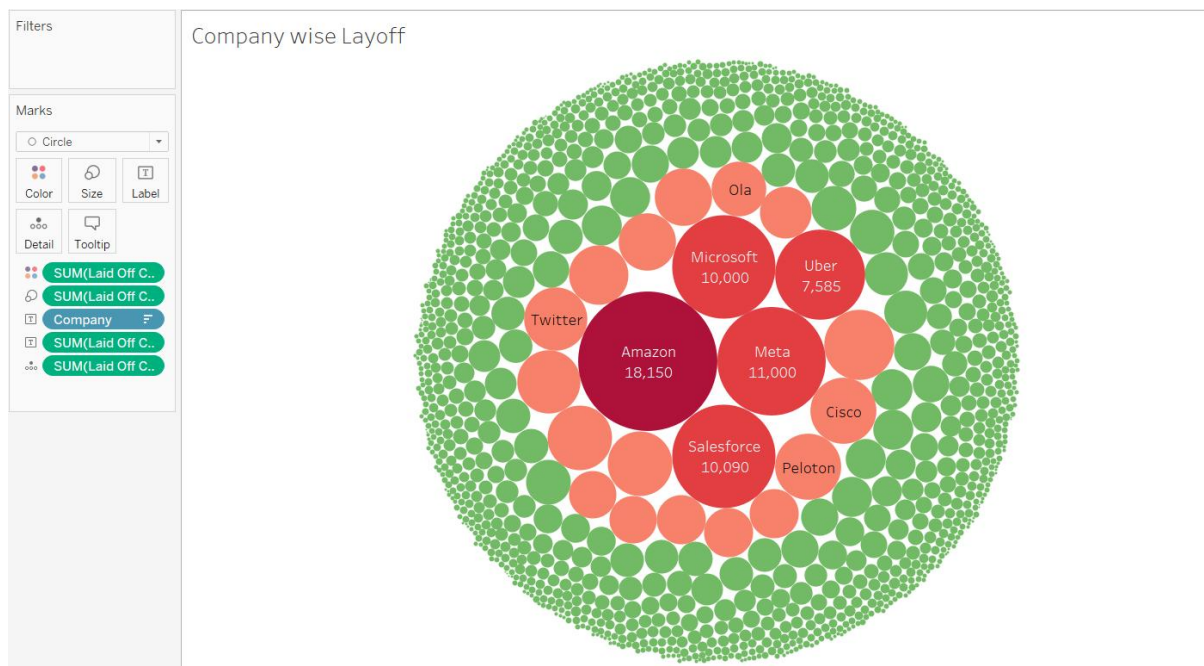


c. Which were the worst affected industries?

After conducting our data analysis and examining various visualizations, we have observed that the retail and transportation industries were the most severely affected by the Covid-19 pandemic-induced layoffs. These visualizations provided us with a clear and concise overview of the layoff trends across different industries, enabling us to draw conclusions about the relative impacts of Covid-19 on different sectors.

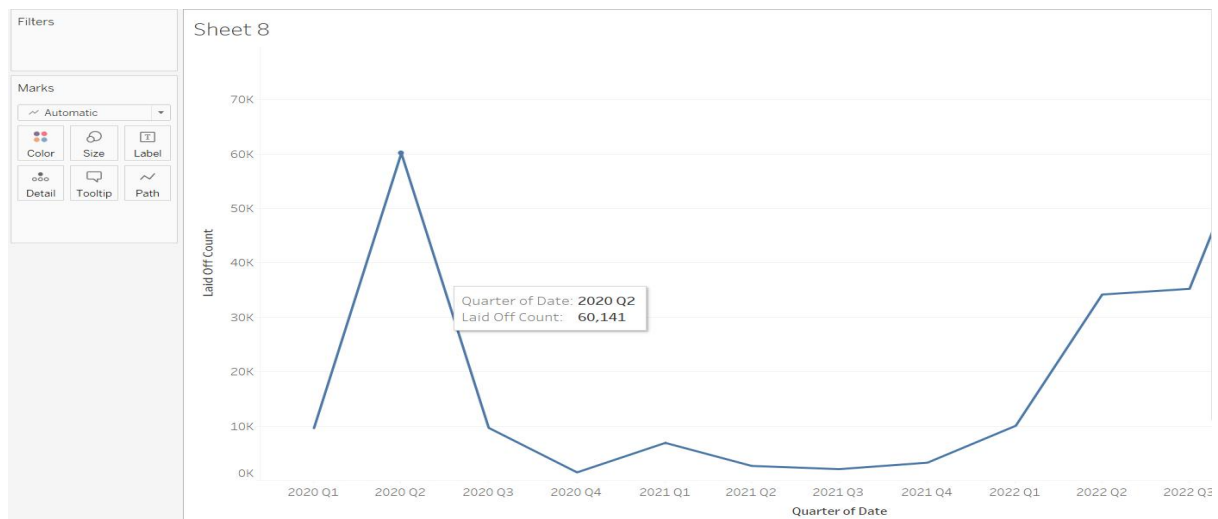


d. Which companies had the largest number of layoffs?



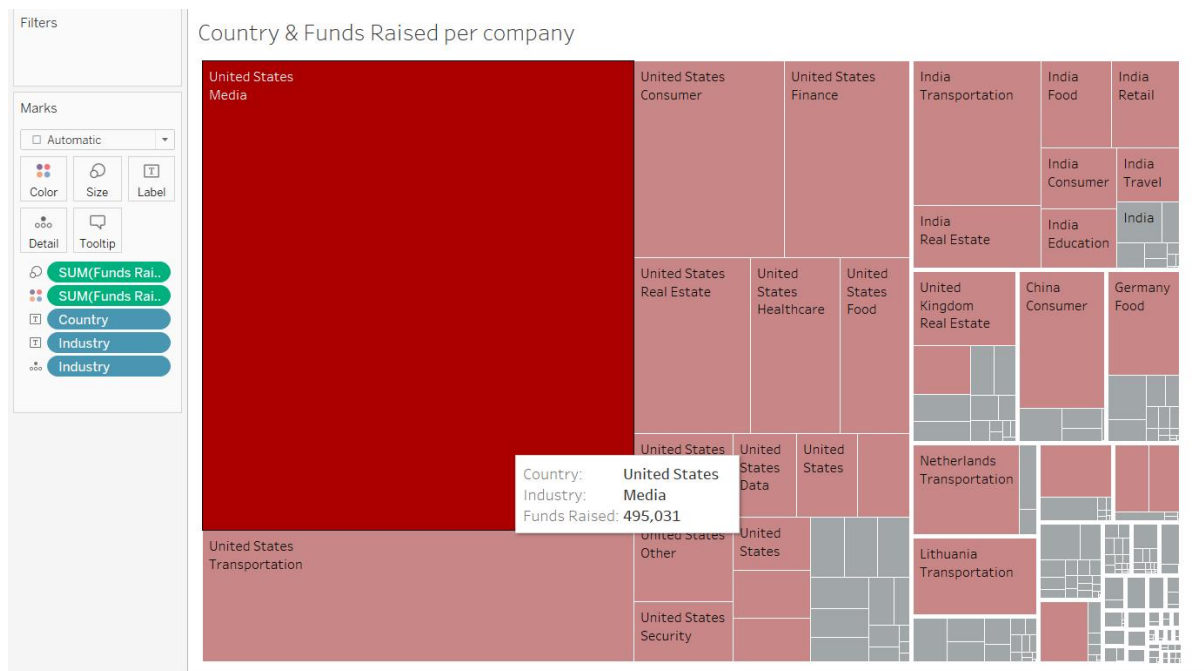
Upon analyzing our visualizations, we observed that the company with the highest number of layoffs was Amazon, followed by Meta (formerly known as Facebook), Microsoft, and Salesforce. These companies were identified based on the number of employees they laid off during the specified time period, and the ranking was determined accordingly.

e. What was the layoff trend from 2020-2022?



Upon analyzing the data visualization, we observed that the highest number of layoffs occurred during the first quarter of 2020. Additionally, there was a smaller peak in layoffs during the first quarter of 2021, followed by a substantial increase in the number of layoffs in the more recent years. These observations indicate that there have been significant fluctuations in the rate of layoffs over time, with some periods experiencing more significant job losses than others.

f. Country and Funds raised by layoff per company.



In our data visualization, we have presented information regarding the total funds raised by companies through layoffs. This visualization allows us to understand the amount of funding that companies have been able to secure by laying off employees. This information can be useful in assessing the financial health of companies and in understanding the impact of layoffs on their overall profitability.

7. Limitations and future scope

- The dataset we worked with was dynamic in nature, meaning that it was regularly updated with new rows of data. As a result, we continuously monitored the dataset to identify any new layoff patterns that emerged and to assess whether our initial findings still held true.
- To improve our predictions, we decided to expand our dataset by incorporating data from a greater number of countries, industries, and years. By doing so, we hoped to gain a more comprehensive understanding of the factors that contribute to layoffs and to identify potential solutions that could prevent or mitigate future layoffs.
- One limitation of our dataset was that it did not include information on the hiring of employees after layoffs. This information would have been valuable in helping us to understand the recovery of various industries after layoffs and to determine whether these industries were able to regain their previous levels of productivity and profitability.
- Spanning from 2020 to 2023, our project would paint a comprehensive picture of the industries hit hardest by job loss.
- Utilizing Tableau software, the project would provide key insights and trend predictions to anticipate which industries may be at risk in the future.

