**Tech Layoffs Data**

**Exploratory Analysis**

Kerry Effinger, [keffinger2@bellarmine.edu](mailto:keffinger2@bellarmine.edu)

1. **INTRODUCTION**

<https://www.federalreserve.gov/monetarypolicy/openmarket.htm>

layoffs.fyi

The website, [layoffs.fyi](https://layoffs.fyi/), is a website created by startup founder, Roger Lee, to track companies that have laid off employees (in small numbers or very large) since the beginning of the COVID-19 pandemic in March of 2020. While the layoffs in the past several months are likely not a result of the initial shocks of the pandemic, the website still tracks the layoffs on a regular basis, with new entries nearly every day. For this project, only layoffs since the beginning of 2023 are considered and continually added to the dataset as more data becomes available.

1. **DATA SET DESCRIPTION**

This set contains 309 samples and 11 columns. Table 1 shows the data types of each column and the percentage of missing data in each column. The original data from layoffs.fyi has 12 columns, but only 9 of the columns were taken to be used in this analysis. The three removed columns were the source from which the information comes (URL), a link to a list of laid off employees, and the country. While the source is very important, URLs cannot be used for statistical analyses like categorical or numerical data can. For the list of employees laid off, it also cannot be analyzed like other data types can. Finally, country was omitted as only companies from the United States were considered for this analysis. However, despite the spreadsheet from layoffs.fyi being filtered with the condition that country is United States, a few entries still showed up with their country as the United States but the city in which they are located is a non-U.S. city. These will remain in the sample since it is unclear whether they are American companies or not. In addition to the 9 columns from layoffs.fyi, an additional column for the Symbol or “ticker” was added after the name of the company. Many of the companies in this dataset were not publicly traded however, so many entries in this column are “private.” Finally, a column which contains the Federal Interest rate at the time of the layoff event is included to give an indication of the state of the economy.

**Table 1: Data Types and Missing Data**

|  |  |  |
| --- | --- | --- |
| *Variable Name* | *Data Type* | *Missing Data (%)* |
| Company | Nominal/Object | 0% |
| Symbol | Nominal/Object | 69.26% |
| Location | Nominal/Object | 0% |
| Number\_laid\_off | Ratio/Float64 | 37.54% |
| Layoff\_date | Interval/Datetime64 | .32% |
| Percent | Ratio/Float64 | 34.30% |
| Industry | Nominal/Object | .65% |
| Stage | Nominal/Object | .32% |
| Money\_raised\_mm | Ratio/Float64 | 12.30% |
| Date\_reported | Interval/Datetime64 | .32% |
| Fed\_int\_rate | Ratio/Float64 | .32% |

1. **Data Set Summary Statistics**

The only continuous variables in the dataset are the number of employees laid off, the percentage of employees laid off, the money raised from the layoffs (in millions of dollars), and the federal interest rate at the time of the layoff. Using pandas, the summary statistics were calculated and are summarized in the table below.

**Table 2: Summary Statistics for Layoffs Data:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Variable Name* | *Count* | *Mean* | *Standard Deviation* | *Min* | *25th* | *50th* | *75th* | *Max* |
| *Number Laid Off* | *193* | *462.54* | *1475.37* | *12* | *55* | *104* | *250* | *12000* |
| *Percent* | *203* | *19.56* | *21.50* | *2* | *7* | *12* | *20* | *100* |
| *Money Raised (mm)* | *271* | *679.13* | *1934.97* | *1* | *101* | *245* | *556.5* | *22200* |
| *Federal Int Rate* | *308* | *4.60* | *.123* | *4.5* | *4.5* | *4.5* | *4.75* | *4.75* |

*Table

Description automatically generated*

Table 3: Proportions for Location(cities) (n=309)

Table

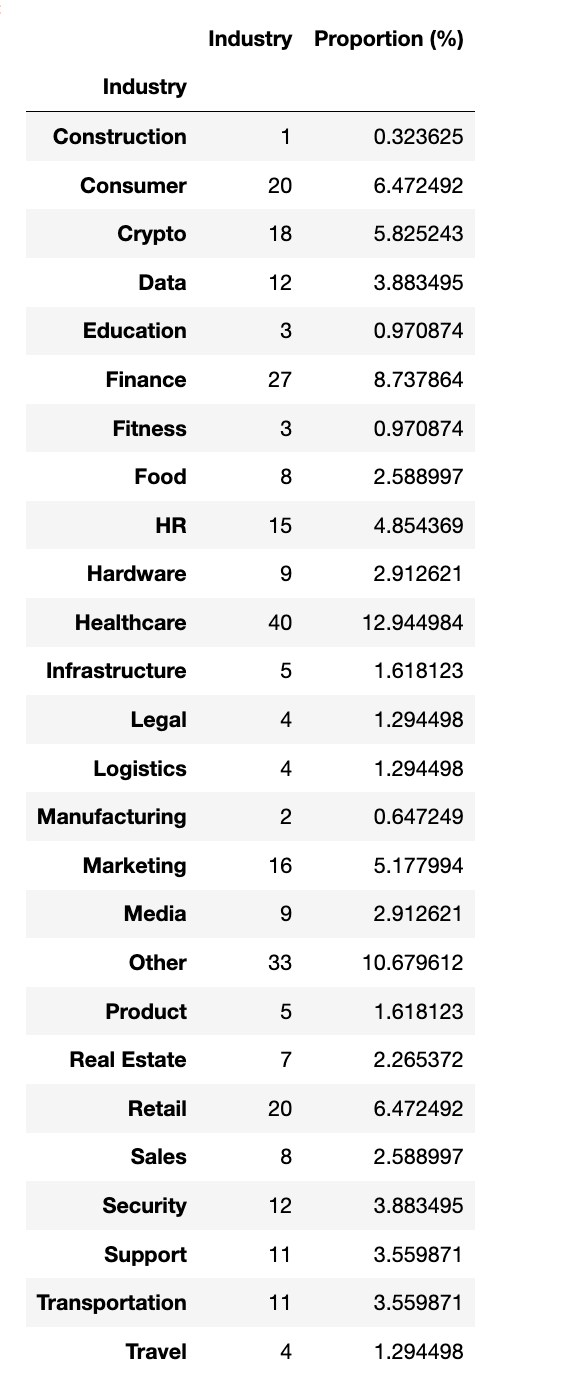
Description automatically generated

**Bar graph for distribution of cities in dataset**

*Chart

Description automatically generated*

Table 4: Proportions for Industry (n=309)



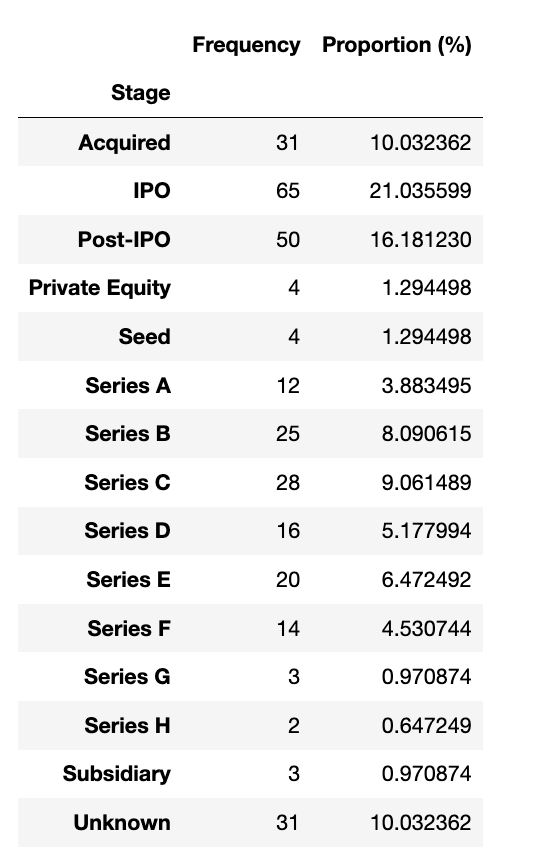
**Bar Graph for distribution of industries**

Chart

Description automatically generated

Table 5: Proportions for Stage (n=309)

After many datapoints were already collected, the way in which the business stages are named and reported was altered slightly. For example, “Post-IPO” is not longer used and “IPO” is now used. At the time of this analysis, these were not yet changed, so this particular variable includes a mix of both the former business stage naming as well as the updated one. It will later be changed to be one uniform naming system.



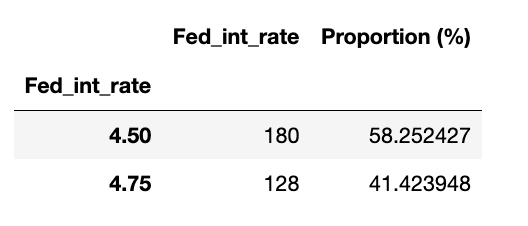
**Bar graph for distribution of business stage for each company**

Chart, bar chart

Description automatically generated

Table 6: Proportions for Fed\_int\_rate (n=308)

At the time of this analysis, only two different federal interest rates were implemented from January 1, 2023 to March 5, 2023. The two rates are 4.5% and 4.75%.



Chart, bar chart

Description automatically generated

Table 7: Correlation Table/Tables

Table

Description automatically generated

**Heat Map:**

Square

Description automatically generated with medium confidence

1. **DATA SET GRAPHICAL EXPLORATION**

As mentioned previously, there are only four categorical variables. The distributions for these will be discussed in this section, though they all have very similar patterns. One continuous variable is the Federal interest rate. For analyses, this can also be treated as a categorical variable since there are only two interest rates in this dataset. There will be some visualizations in this section that investigate the impact that federal interest rate could have on the other variables.

**Figure 1: Distribution for Number Laid Off**

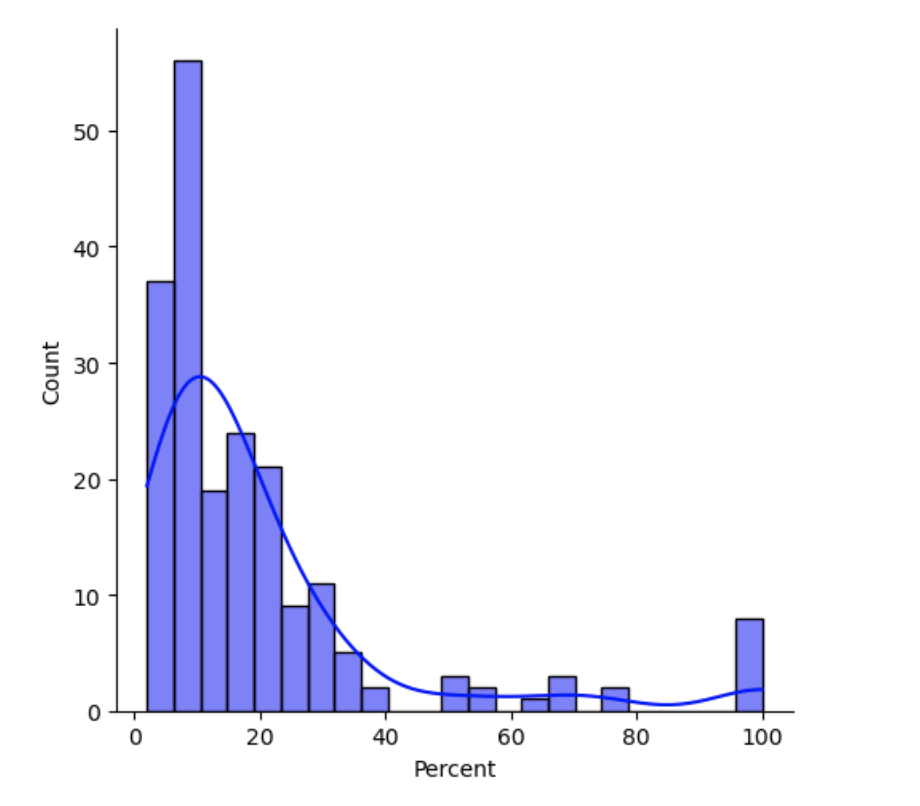
The spread for the number of employees laid off is very skewed right with the vast majority of the data clustered around the lower number of employees.

**Chart

Description automatically generated**

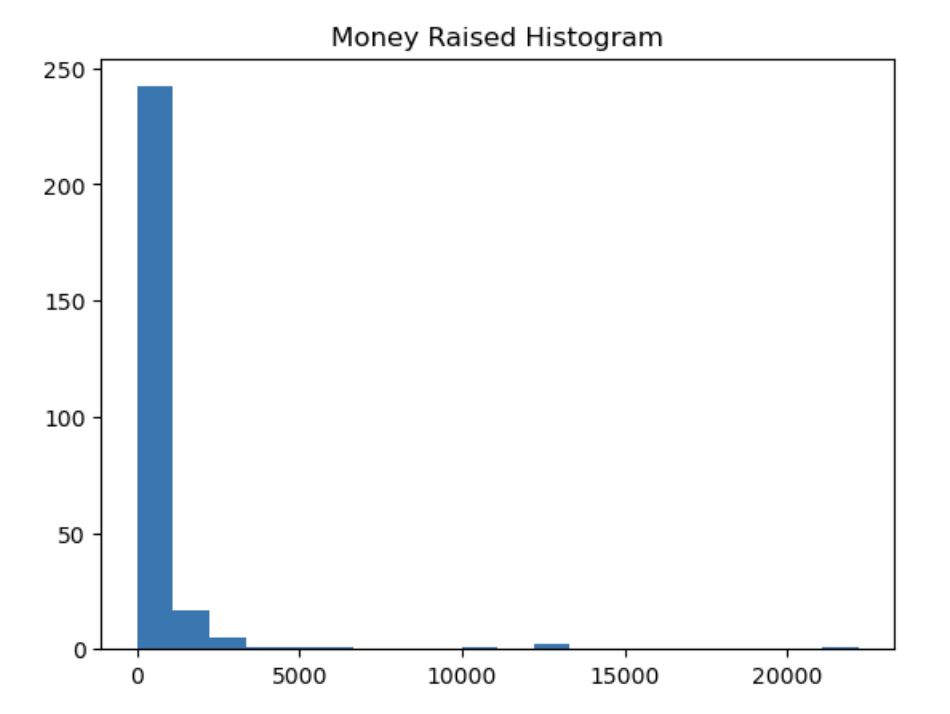
**Figure 2: Distribution for Percent Laid Off**

The distribution for percent laid off is also very skewed right with the majority of datapoints clustered below 40 percent. There is slightly more spread than the other two continuous variables, but it is still skewed.

****

**Figure 3: Distribution for Money Raised**

The spread of money raised in millions of dollars is also very skewed to the right as most datapoints are for a small amount of money raised.

****

**Figure 4: Pivot table with Interest Rates as categories**

The two different interest rates are compared with the total employees laid off, the average percent of employees laid off, and the average number of employees laid off during the period of time with that interest rate. When the rate was 4.5%, the sum of employees laid off was significantly higher than when the rate was 4.75%. This could be due to larger companies conducting layoffs in this time period. The average percent of employees laid off was relatively close between each time period, with the 4.5% interest rate layoffs having a slightly larger percent than during the 4.75% time. The average number of employees laid off during the 4.5% interest rate era was much higher than when the rate was 4.75%. The difference in these sums could be explained by larger companies conducting layoffs in January (when the interest rate was 4.5%) and smaller companies primarily having layoffs in February (when the rate was 4.75%) just by chance. As more data continues to be collected, this relationship may appear to have a stronger or weaker correlation.

**Table

Description automatically generated**

**Figure 5: Pair plot comparing Number Laid Off, Percent and Money Raised**

No comparisons between any two variables appear to have any sort of linear relationship. For the most part, the data points are very clustered around the low numbers. So, the majority of layoffs in this dataset were both small numbers of employees laid off and a small percent of the company’s workforce. There are some clear outliers as some companies laid off a small number but very large percent of employees (these are seen in the Tableau dashboard with several small companies laying off 100% of employees). There are also very large companies that laid off thousands of employees, such as Google, Microsoft, Salesforce and Amazon, but the percent of the company was still very small. These are all displayed on the Tableau Dashboard as well and can be seen as the points along the bottom of the scatterplots below.

**Chart

Description automatically generated**

**Figure 6: Scatter plot comparing Number Laid Off and Percent Laid off**

Although this scatterplot is displayed above in the pair plot, this once again highlights that the majority of tech layoffs were very small in number but with varying percentages skewed towards the lower end.

Chart, scatter chart

Description automatically generated

**Figure 4: Categorical Scatter plot with Federal Interest Rate and Percent Laid off**

This scatterplot separated by the two different interest rates in this dataset once again appears to show that the distributions of percent of employees laid off is very similar for both the time periods when the interest rate was lower at 4.5 and after it was raised to 4.75.

*Chart, scatter chart

Description automatically generated*

1. **SUMMARY OF FINDINGS**

Fortunately, layoffs.fyi was able to provide me with great data on layoffs all in one place. I supplemented this data with an additional column containing “symbol” which is the stock symbol for publicly traded companies. I wanted to add this column as I was collecting data in the chance than I am able to figure out how to systematically collect financial data about each company from Yahoo Finance using their tickers. I quickly discovered, however, that many of the companies in the dataset are private and do not have a public ticker. So, while I have found a way to return a few financial ratios about each company using their ticker, there are not that many companies with this data. It may still be included in a later version of the project. The only other column I added also was the federal interest rate column. That variable may show some interesting relationships in the long-term, but for the short term data collected now, there does not appear to be a strong relationship between interest rate and percent of employees laid off from a given company. Another small issue with the data is that ther is a decent amount of missing data in both the number laid off and percent laid off columns. Each have over 30% of the column missing. Many companies may not release the exact numerical data for their layoffs, or they only report either a count of employees or the percent but not both. This may cause some issues when creating the model if the dataset is too small. However, with new data being added every day, the model should still have a large number of datapoints. The rest of the data is relatively complete and has been instrumental in understanding the nature of layoffs in tech jobs to complete this project.

**Screenshot of Tableau Dashboard:**

Application

Description automatically generated with low confidence