Lab 2: Analyzing the Impact of Company Funding on 2023 Layoffs

Jinney Hong, John Bang, Steven Au

12-13-2024

Abstract

In 2023, economic instability following COVID-19 led to mass layoffs across tech companies in various industries, significantly impacting the U.S. workforce. This study investigates the role of funding availability, company growth stage, and industry-specific factors in shaping corporate responses to financial uncertainty. Using regression analysis on a dataset of U.S.-based tech companies, this research tests the hypothesis that companies with lower funding have the same number of layoffs as companies with higher funding in 2023. Results indicate that while funding is statistically significant, company growth stage and industry-specific dynamics account for a larger proportion of variation in layoffs when included as additional variables. This highlights that layoffs are driven by a combination of financial pressures and company-specific factors. Ultimately, this report will help businesses identify the key factors driving workforce reductions and develop strategies that align more effectively with their financial and industry contexts during economic downturns.

Introduction

In 2023, more than 191,000 workers in U.S.-based tech companies alone faced layoffs as organizations responded to financial and market uncertainties. "The Crunchbase Tech Layoffs Tracker" (2024) These layoffs not only affected the lives of employees but also reflected corporate financial health and resilience in navigating market challenges. Companies with limited financial reserves may resort to workforce reductions, but even those with sufficient funding may implement layoffs to meet market expectations or to make operations more efficient.

This study examines the role of funding, growth stage, and industry as factors that could influence layoffs. Specifically, we will address the research question: How does a company's funding, stage, and industry within the United States impact the number of employees laid off in 2023? We hypothesize that companies with lower funding will experience the same number of layoffs as those with higher funding. This reflects the idea that layoffs may not only depend on financial health but also on strategic decisions influenced by market conditions. The findings aim to describe the effect of financial and company-specific dynamics on layoffs.

Data Source

Our dataset is from Kaggle's 2024 layoffs data, which is sourced from a public database called layoffs.fyi created to track layoffs during the pandemic. The dataset includes companies from various countries, ranging from the United States to India, spanning from March 10, 2020 to June 6, 2024. Our analysis focuses on companies in the United States and layoffs during 2023 (January 1st to December 31st). This data is collected from publicly available sources, including news reports, input from former employees, and company press releases with each entry representing a unique layoff event.

The dataset contains the following columns: Company, Location_HQ, Industry, Laid_Off_Count, Date, Source, Funds_Raised, Stage, Date_Added, Country, Percentage, List_Of_Employees_Laid_Off. Our analysis focuses on the Laid_Off_Count as the response variable, which represents the number of employees laid off. The predictor variables include Funds_Raised, Stage, and Industry.

Operationalization

Table 1

	Response Variable	Predictor Variables	Other Variables for filtering and relevance
Feature used	Laid_Off_Count (metric)	Funds_Raised (metric), Stage, Industry	Date, Country, Company
Definition	The total number of employees laid off by a company.	**Funds_Raised**: Total funding secured by a company, in millions of dollars. **Stage**: Company growth phase, from early-stage (ex. seed) to later-stage (ex. post-IPO). **Industry**: Sector the company operates in such as technology, healthcare, or finance.	**Date**: The date on which the layoffs were reported. **Country**: The country that the company is headquartered. **Company**: The name of the organization conducting the layoffs.

	Response Variable	Predictor Variables	Other Variables for filtering and relevance
Rationale	This variable directly captures the absolute magnitude of workforce reductions as a quantifiable measure of layoffs that avoids potential bias from company size variations inherent in the variable, *Percentage*.	**Funds_Raised**: Indicates a company's financial health and ability to navigate economic challenges, influencing layoff decisions; missing values were excluded for data quality. **Stage**: Reflects business strategies and economic vulnerability, making it relevant to layoffs. **Industry**: Captures industry-specific trends and pressures that impact layoff likelihood.	**Date**: Constrains the dataset to 2023 to avoid temporal bias and focus on that year's layoffs. **Country**: Excludes non-U.S. observations to align with the study's focus on U.S. market dynamics. **Company**: Ensures unique observations by removing duplicate layoff records for distinct occurrences.

Data Wrangling

Table 2

Cause	Observations Remaining	Dropped Observations
Start	3642	0
Filter Date range 01-01-2023 to 12-31-2023	1383	-2259
Limit Country to United States	881	-502
Drop all Nulls	460	-421
Group by companies to remove temporal bias	406	-54
Remove duplicated Companies that changed stage or industry	400	-6
Remove unknown entries and apply data munging changes	362	-38

We have wrangled the dataset per the causes down from 3642 to 362 observations. We then randomly sampled 30% of the results by industry to ensure a reasonable distribution 119 for the training and the remainder 243 for the confirmation datasets. However, stage was not used as part of the distribution split since there were fewer categories that would cause an imbalance.

Null Hypothesis

Companies with lower funding have the same number of layoffs than companies with higher funding in 2023.

Model Specification

To explore our research question, we created several linear regression models. Our first model focuses on the relationship between the dependent variable, funding(Funds_Raised), and the independent variable, layoffs (Laid_Off_Count). Both variables resulted in a positive skew, so we applied log transformations to normalize the distributions. The following model tests if there is a significant relationship between funding levels and the number of logged layoffs:

$$Log(\text{Layoffs}) = \beta 0 + \beta 1 * Log(\text{Funds\_Raised}) + \epsilon$$

Next, we take into account the industry, as an additional categorical variable to see if a company's industry influences layoffs:

$$Log(\text{Layoffs}) = \beta 0 + \beta 1 * Log(\text{Funds\_Raised}) + \beta 2 * Industry + \epsilon$$

To improve the explanatory power, another factor (stage) was added to further explain the variation and to analyze the effect of the stage on logged layoffs.

$$Log(\text{Layoffs}) = \beta 0 + \beta 1 * Log(\text{Funds\_Raised}) + \beta 2 * Industry + Stage + \epsilon$$

Model Assumptions

We evaluated the large sample assumptions for the existence of the best linear predictor (BLP) and independent and identically distributed (IID) data for our exploratory model.

For the BLP assumption, because we applied a log transformation for Laid_Off_Count and Funds_Raised, this resulted in less heavy tails (Appendix 1 - skew and kurtosis subtitled) and a normalized distribution (Figure 1). We also calculated the variance inflation factor (VIF) to confirm that the predictor and response variables have no perfect collinearity if the value is below 5. The computed VIF value is 1.65. As a result, we verified that a unique BLP exists for our model.

For the IID assumption, we evaluated that the dataset's sampling method is gathered independently as mentioned in the Data Source section. Although we understand that not all companies per industry have similar business operations and structures leading to a possible bias in our results, we can confirm that our data does not have particular industry and stage clustering for our analysis per our scatterplots. (Appendix 2)

Visualization

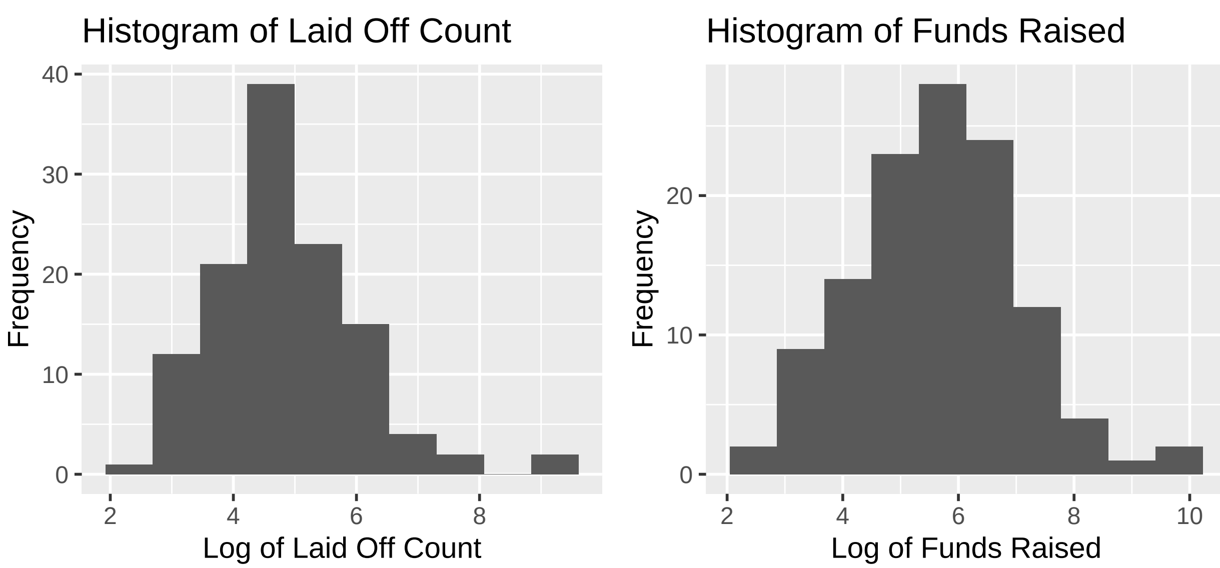


Figure 1: A log transformation applied to both the Funds_Raised and Laid_Off_Count metric variables results in normalized distribution histogram plots for our model fit assumptions.

Model Results

The regression results provide insights into the relationship between funding and layoffs, with additional factors included to improve our model.

First trained model (raw data with no transformations): In the first trained model with unlogged variables, we have a p-value of 0.855 and an R-squared value of 2.9×10^{-4}, which captures very little of the variation, meaning the independent variable (layoffs) is not well explained by the funding level.

Second trained model: We then logged the data because both distributions are heavily skewed positively to show a normalized distribution. (Figure 1). The p-value is 0.023 and the R-squared explains only 4.349% of the variation in logged laid off count.

Since our dependent variable (funds raised) and our independent variable (laid off count) were both skewed, we performed a log transformation on the laid off count and the funds raised.

Third trained model with industry as a factor: We then added Industry as a factor in our model, hoping to explain more of the variation. This addition significantly improved the model's fit, with the p-value for Log(Funds_Raised) as 0.022 and

R-squared increasing to 0.29629. The coefficient for Log(Funds_Raised) is 0.173, meaning a 1% unit increase in funding is associated with a 0.173% increase in layoffs, but this is not statistically significant. See appendix 3 for the residual plot.

Fourth trained model with stage as a factor: We then added stage as a factor in our model, hoping to explain more of the variation. The R-squared increased to 0.247 and the p-value for Log(Funds_Raised) was 0.39381, which is closer to the significance level when industry and stage are taken into account. As for practical significance, the coefficient for Log(Funds_Raised) is 0.173, meaning a 1% increase in funding is associated with a 0.173% increase in layoffs, but this is not statistically significant. See appendix 3 for the residual plot.

A partial linear regression summary of the second, third, and fourth models is below in table 3.

The unlogged summary is available in table 4 and a full version of the below table 3 is on Appendix 4.

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Dec 12, 2024 - 10:22:36 % Requires LaTeX packages: dcolumn

Table 3: Linear Regression Summary

	Dependent variable:		
	Log_Laid_Off_Count		
	explanatory model	explanatory model industry	explanatory model industry stage
	(1)	(2)	(3)
factor(Industry)Data		−0.957	−0.743
		(0.671)	(0.672)
factor(Industry)Finance		−0.885	−0.495
		(0.548)	(0.563)
factor(Industry)Healthcare		−0.983*	−0.832
		(0.547)	(0.547)
Observations	119	119	119
R^2	0.043	0.296	0.394
Adjusted R^2	0.035	0.097	0.138
Residual Std. Error	1.145 (df = 117)	1.108 (df = 92)	1.083 (df = 83)
F Statistic	5.320** (df = 1; 117)	1.490* (df = 26; 92)	1.541* (df = 35; 83)
Note:			*p<0.1; **p<0.05; ***p<0.01

The unlogged summary

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Dec 12, 2024 - 10:22:36

Table 4: Linear Regression Summary

	Dependent variable:
	explanatory model unlogged
Funds_Raised	−0.007
	(0.038)
Constant	375.187***
	(125.060)
Observations	119
R^2	0.0003
Adjusted R^2	−0.008
Residual Std. Error	1,301.362 (df = 117)
F Statistic	0.034 (df = 1; 117)
Note:	*p<0.1; **p<0.05; ***p<0.01

Practical Significance

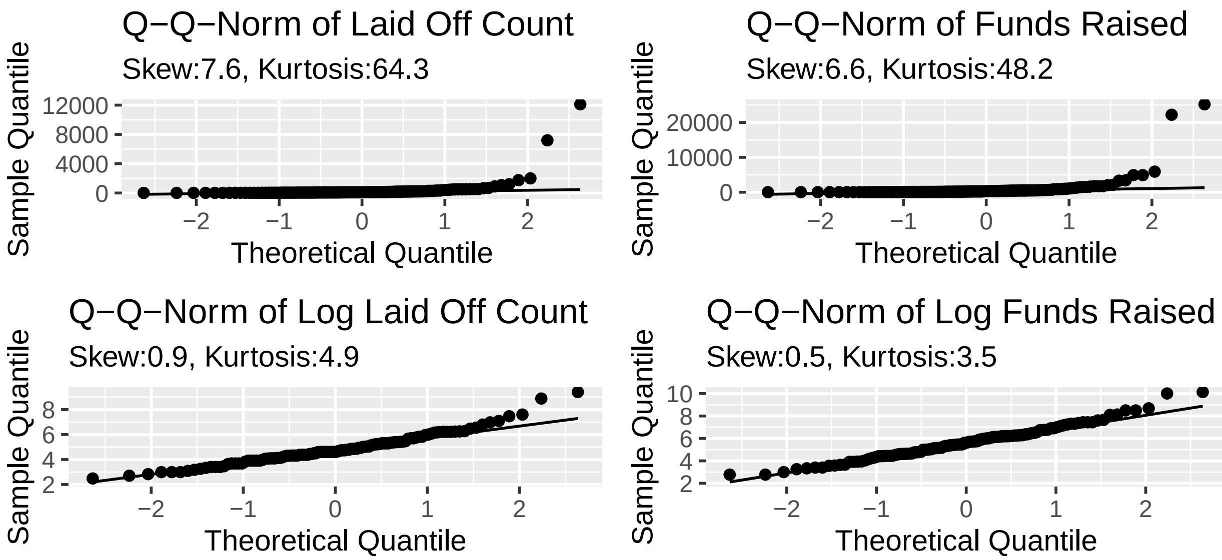
In conclusion, our study fails to reject the null hypothesis. Although our findings indicate a slight increase in layoffs associated with company funding, the relationship was determined to not be statistically significant given the dataset used. However, adding industry and growth stage into the model improved the model's R-squared value, suggesting that these variables play a meaningful role, alongside funding, in influencing layoff numbers.

For future model improvements and research, we could include additional variables, such as company size or geographic location, which could yield additional insights while building upon existing studies exploring drivers of workforce reductions. We hope to provide organizations with data-driven insights for better workforce management during economic challenges.

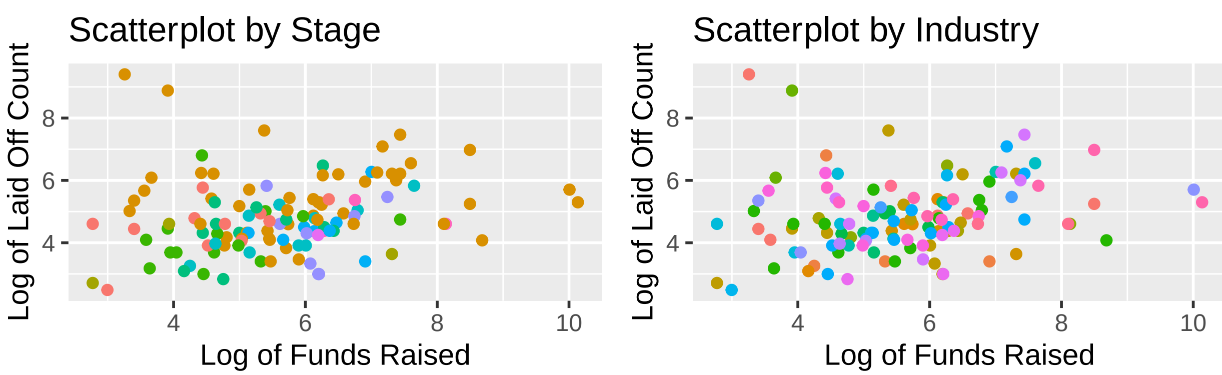
Datasource https://www.kaggle.com/datasets/theakhilb/layoffs-data-2022/data

Appendix

Appendix 1 - QQNorm plots from exploratory data showing the decrease heavy tail distributions after logging

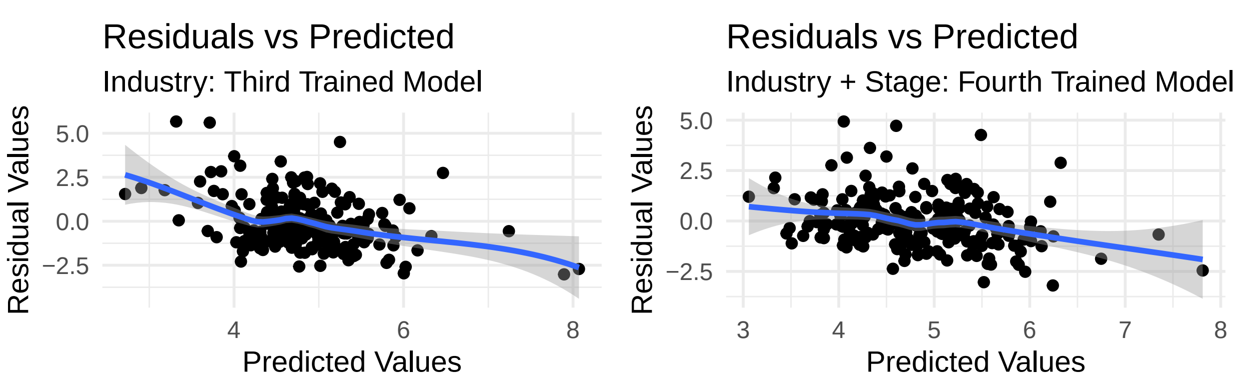


Appendix 2 - Scatterplots from exploratory data showing no clusters - Legend removed due to space constraints



Residuals vs fitted plots

Appendix 3 - Residual vs Fitted Plots for Explanatory Model



Appendix 4 - Full Logged Models 2, 3, 4 Stargazer

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Dec 12, 2024 - 10:22:36 % Requires LaTeX packages: dcolumn

Table 5: Appendix 4 - Full Linear Regression Summary

	Dependent variable:		
		Log_Laid_Off_Count	
	explanatory model unlogged	explanatory model	explanatory model industry
	(1)	(2)	(3)
Log_Funds_Raised	0.173^{**}	0.195^{**}	0.109
	(0.075)	(0.084)	(0.093)
factor(Industry)Crypto		-1.261^{*}	-0.682
		(0.715)	(0.736)
factor(Industry)Data		-0.957	-0.743
		(0.671)	(0.672)
factor(Industry)Education		-1.563^{*}	-1.125
		(0.788)	(0.825)
factor(Industry)Finance		-0.885	-0.495
		(0.548)	(0.563)
factor(Industry)Fitness		-0.936	-0.520
		(0.905)	(0.925)
factor(Industry)Food		-0.637	-0.373
		(0.789)	(0.819)
factor(Industry)Hardware		2.260^{**}	1.853^{**}
		(0.913)	(0.907)
factor(Industry)Healthcare		-0.983^{*}	-0.832
		(0.547)	(0.547)
factor(Industry)HR		-0.829	-0.255
		(0.715)	(0.729)
factor(Industry)Infrastructure		-1.097	-0.883
		(0.784)	(0.792)
factor(Industry)Legal		-1.007	-0.927
		(0.905)	(0.908)
factor(Industry)Logistics		-0.528	-0.424
		(0.906)	(0.914)
factor(Industry)Manufacturing		0.588	0.508
		(1.213)	(1.187)
factor(Industry)Marketing		-0.476	-0.261
		(0.723)	(0.715)
factor(Industry)Media		-1.066	-1.168
		(0.784)	(0.770)
factor(Industry)Other		-0.904	-0.691
		(0.549)	(0.542)
factor(Industry)Product		-0.395	0.251
		(0.908)	(0.950)
factor(Industry)Real Estate		-0.694	-0.581
		(0.785)	(0.771)
factor(Industry)Recruiting		-1.393	-1.024
		(0.905)	(0.919)
factor(Industry)Retail		-0.147	-0.239
		(0.645)	(0.632)
factor(Industry)Sales		-1.982^{***}	-1.621^{*}
		(0.717)	(0.821)
factor(Industry)Security		-0.647	-0.452
		(0.616)	(0.617)
factor(Industry)Support		-0.202	-0.039
		(0.784)	(0.783)
factor(Industry)Transportation		-0.289	0.015
		(0.664)	(0.661)
factor(Industry)Travel		-0.779	-0.611
		(0.793)	(0.789)
factor(Stage)Post-IPO			0.657
			(0.408)
factor(Stage)Series A			-0.570
			(0.778)
factor(Stage)Series B			-0.379
			(0.468)
factor(Stage)Series C			-0.062
			(0.501)
factor(Stage)Series D			-0.059
			(0.533)
factor(Stage)Series E			-0.006
			(0.582)
factor(Stage)Series F			-0.308
			(0.579)
factor(Stage)Series G			0.649
			(1.332)
factor(Stage)Series H			0.033
			(0.812)
Constant	3.868^{***}	4.482^{***}	4.560^{***}
	(0.437)	(0.632)	(0.673)
Observations	119	119	119
R^2	0.043	0.296	0.394
Adjusted R^2	0.035	0.097	0.138
Residual Std. Error	1.145 (df = 117)	1.108 (df = 92)	1.083 (df = 83)
F Statistic	5.320^{**} (df = 1; 117)	1.490^{*} (df = 26; 92)	1.541^{*} (df = 35; 83)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

References

"The Crunchbase Tech Layoffs Tracker." 2024. *LinkedIn And AppLovin Rejoin Tech Layoffs Tracker In A Somewhat Subdued Year-Ending Quarter*. https://news.crunchbase.com/startups/tech-layoffs/.