

UNICORN COMPANIES ANALYSIS USING SAS STUDIO

CIS 5250 – Visual Analytics

Mitali Purohit

Table of Contents

Α.	Introduction
В.	Data Descriptions
C.	Data Cleaning5
1	. Data Category Name: Split column5
2	. Data Category Name: Replacing empty values7
3	. Data Category Name: Converting all values into one unit9
D.	Analysis & Visualizations11
1	. Which are the top 10 highest valued Unicorn Companies globally?11
2	. Which industry (sector/field) has the highest valuation (\$)? 12
3	. Which country has the largest shares in Unicorn Companies?14
E.	Statistical Summary
1	. Statical Summary for column "Valuation (\$)"
2	. Statistical Summary for column "Date Joined" 16
F.	Statical Tests
1	. One-way frequency 18
2	. Correlation Analysis 19
3	. T-Test
G.	References

A. Introduction

Private businesses worth more than \$1 billion are known as Unicorn Companies or Unicorn Startups. The 21st century is the era of startups. Numerous Unicorn Companies have been involved in the creation of new technology, space exploration, educational breakthroughs, and other things over the past 20 years. Based on the top of 348 in 2018, there will be 354 active unicorn firms worldwide in 2019. There were 439 active Unicorn Companies, including 139 new startups, at the end of 2019. By 2020, there would be 538 active Unicorns, with 162 new businesses across the globe achieving unicorn status. A surprising 355 new unicorn firms were started in 2021, bringing the total to 537, roughly on par with the previous year. There are 1,000 unicorns in the world as of March 2022. In addition to Airbnb, Facebook, and Google are well-known ex-unicorns. Ten kernels, worth over \$10 billion, and Hectocorn seeds, valued over \$100 billion, are examples of variations. Due to the tremendous expansion of unicorn companies and the growing combination of the private and public markets, many unicorn companies are created. Companies have frequently used initial public offerings (IPOs) to raise money to expand their operations. However, businesses can now attract more private capital at an earlier stage and attain valuations of billions without going public. [1]

Unicorns are startups that are believed to exhibit a unique form of entrepreneurial magic. They are a privately held business with a \$1 billion market value. Spotify and Uber were unicorns. That is still how DoorDash, and Airbnb are described. However, as investment in Silicon Valley has soared in recent years, there are now more businesses valued at \$1 billion or more than ever before. A surge in businesses and investors has led to an increase in the number of unicorns. Many new enterprises, many of them startups, entered the market after the recession triggered a steep decrease

in startups and other new businesses. Today's businesses have more early opportunities to secure sizeable private investments, which enables them to accumulate billion-dollar valuations without going public. [2]

Five of the top 10 new unicorns in the US were blockchain or crypto-currency related, accounting for 60% of the FinTech start-ups. Given the success of these top fintech unicorns, there appears to be a lot of interest in this field, which could lead to the emergence of new unicorn businesses. E-commerce, fintech, internet software, healthcare, travel technology, and education technology are the seven industries where unicorns are most prevalent. Silicon Valley in San Francisco is still associated with startups, but more unicorn businesses are appearing elsewhere. Overall, the number of startups is skyrocketing, therefore over time, you'll see an increase in the number of unicorns. Technology advancements enable firms to expand more quickly. Unicorn firms can reach their customers more quickly and cut down on the time needed for mass production thanks to new technology.

In this project, we are examining the data set of unicorn firms to comprehend the growth based on value, location, industry, total funding, and investors. Data analysis and visualization will be done for the startup company that becomes a unicorn business. Python will be used to build a visualization that will show how these businesses have changed over time. Insights regarding numerous unicorn firms' growth depending on various industries, values, and nations were offered through these analysis and visualizations.

B. Data Descriptions

Dataset(s)/ Data source URL:

https://www.kaggle.com/datasets/deepcontractor/unicorn-companies-dataset

The dataset is regarding the Unicorn Companies taken from Kaggle website. This data set contains 1038 rows and 13 columns. It has various details of all Unicorn Companies over the world.

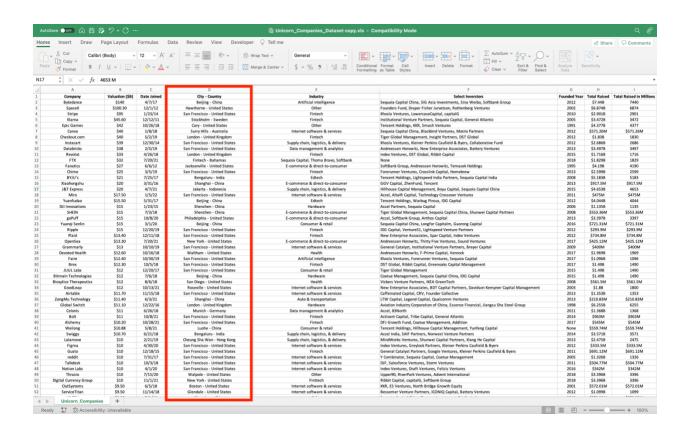
<u>Field Name</u>	Field Description	<u>Example</u>	
Company	Name of the Company	SpaceX	
Valuation (\$ B)	Company's Estimated Value in Billions	100.3	
Date Joined	Date when the company was founded	12/1/12	
Country	Country of Company located	United States	
City	City of the Company located	Hawthorne	
Industry	Type of Industry	Other	
Select Investors	Companies invested in different Companies	Rothenberg Ventures	
Founded Year	Year in which the company was founded	2002	
Total Raised	Total fund raised by the company in dollars	6874 M	
Financial Stage	Financial Stage of the Company	IPO	
Investors Count	Numbers of Investors in the Company	29	
Deal Terms	Numbers of deals with the Company	12	
Portfolio Exits	Portfolio of the company exists or not	3	

C. Data Cleaning

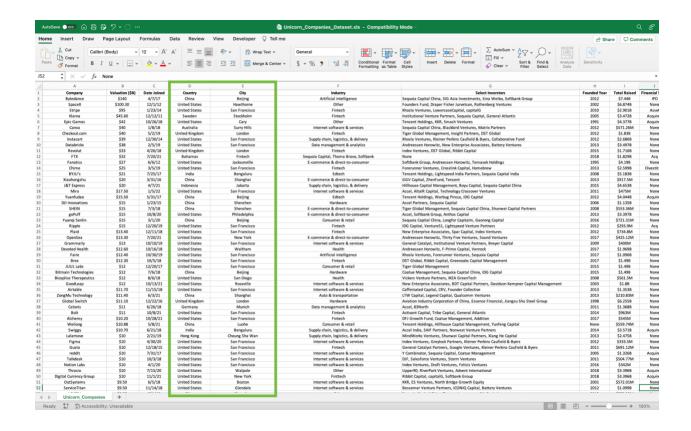
1. Data Category Name: Split column

Action: The column name "City-Country" were together as shown in the pre-cleaning image and difficult to understand the specific country/region. Split them into two different columns "City" and "Country" using "Text-to-column" formula.

Pre-cleaning screenshot



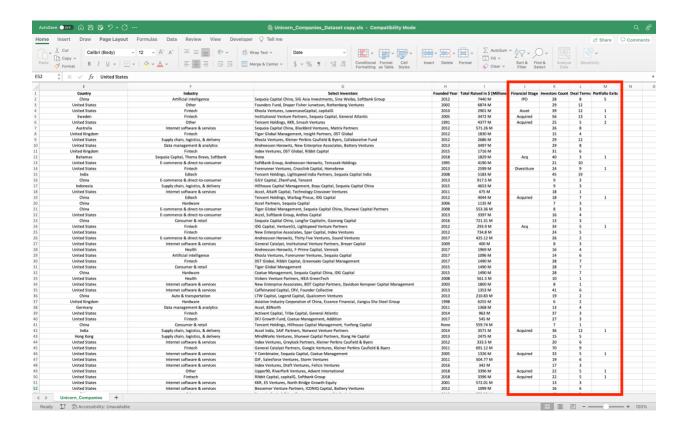
Post-cleaning screenshot



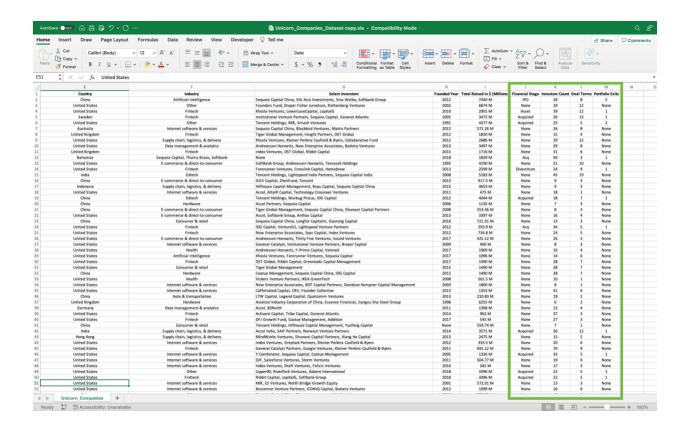
2. Data Category Name: Replacing empty values

Action: There are several columns with missing values/empty cells as shown in the pre-cleaning screenshot. Replaced them with "None" to make the graph more precise.

Pre-cleaning screenshot



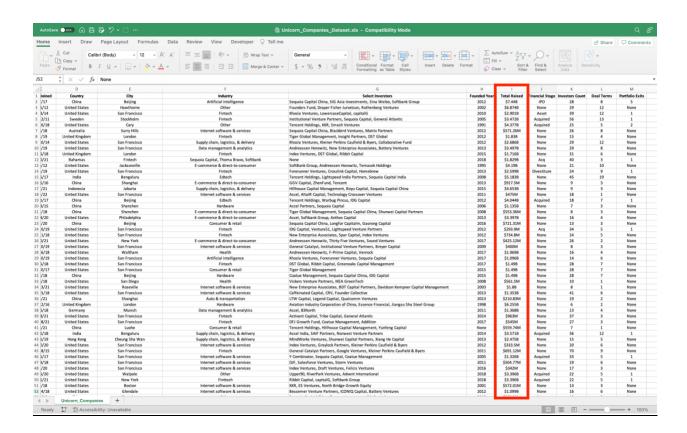
Post-cleaning screenshot



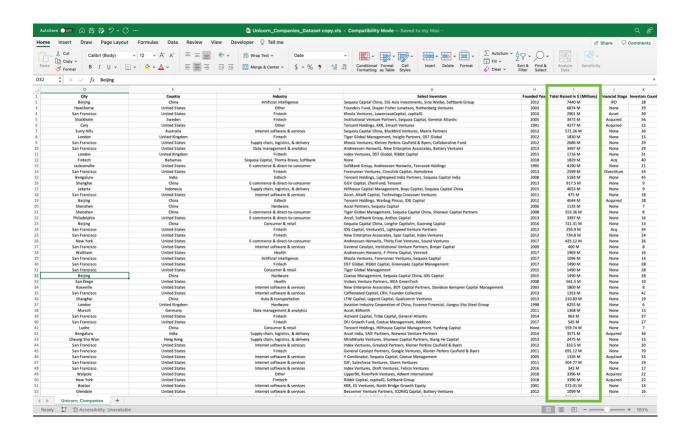
3. Data Category Name: Converting all values into one unit

Action: The column name "Total Raised in \$" had two different unit values i.e., "millions" and "billions". Converted all the values into one single unit "millions (M)" to understand much better in the visuals.

Pre-cleaning screenshot

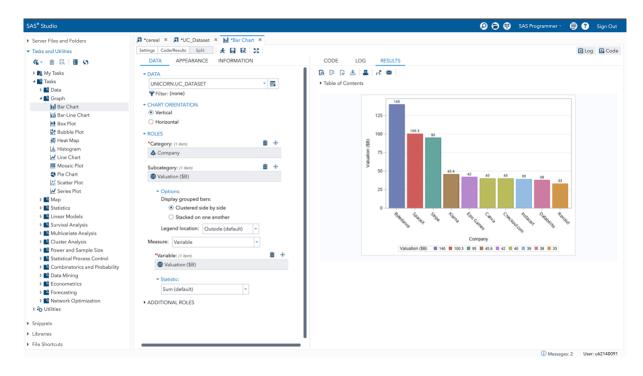


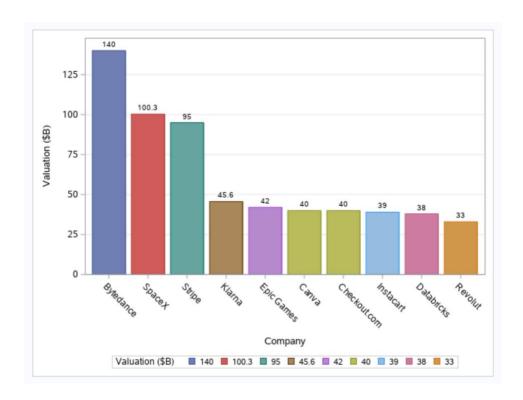
Post-cleaning screenshot



D. Analysis & Visualizations

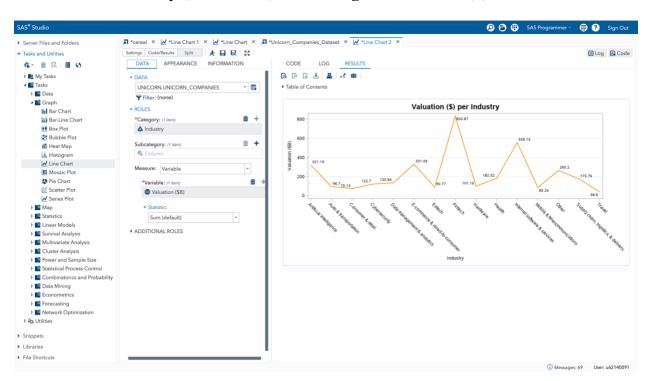
1. Which are the top 10 highest valued Unicorn Companies globally?

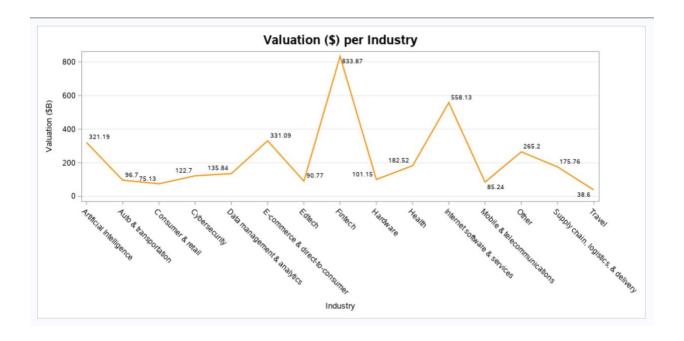




The bar graph shows the top 10 highest valued Unicorn companies globally. All the bars are with different colors to understand the graph easily with the labels on each top to show the total valuation of the company. The highest valued Unicorn company is Bytedance from Beijing, China with approx. valuation of \$140 billion followed by SpaceX from Unites States (\$100.3 billion) and the third is Stripe from United States (\$95 billion). The Unicorn companies with the ranks from 4th to 9th are all valued between \$38 billion to \$46 billion. The lowest among the top 10 companies is Revolut with \$33 billion valuation.

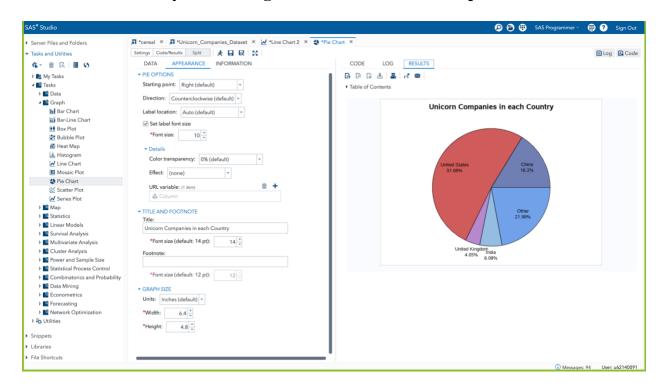
2. Which industry (sector/field) has the highest valuation (\$)?

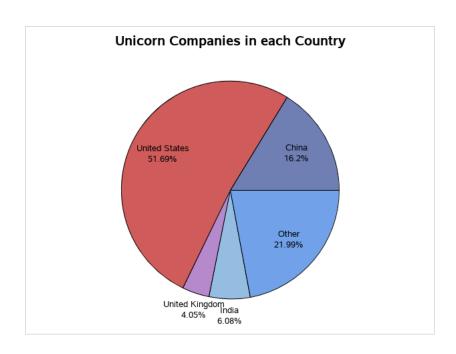




This line chart displays the summed valuations of all the Unicorn companies respective to each Industry they fall within. As we can see from the graph that the industry(sector) with the highest valuations are the Fintech companies. These fintech companies includes some giant Unicorn companies like Stripe, Klarna, Checkout.com, Revolut, FTX and many more. The top 5 highest valuations industry followed by Fintech are Internet Software and Services, E-commerce & direct-to-consumer, Artificial Intelligence, and Others respectively which includes widely known companies like SpaceX, Bytedance, Epic Games, Instacart, BYJU's, SHEIN, Ripple, OpenSea, Grammarly and more. The lowest industry with the least valuation of the Unicorn companies is the Travel sector. Also, we can see that there are lots of industry sectors that averages between \$85-\$185 billions of valuations.

3. Which country has the largest shares in Unicorn Companies?





This pie chart shows the countries with the most Unicorn companies. The chart has been displayed with different colors which easily shows the percentage of shares of Unicorn companies each country has. As we see from the graph, United States shares the highest number of Unicorn countries with more than half of the total Unicorn companies from the data. Followed by China, India, UK, and Others shares the other half of the Unicorn companies.

E. Statistical Summary

1. Statical Summary for column "Valuation (\$)"

Analysis Variable : Valuation (\$B) Valuation (\$B)								
Mean	Std Dev	Minimum	Maximum	N				
3.2920829	7.3097808	1.0000000	140.0000000	1037				

Statical Summary	<u>Explanation</u>
Mean	Average value of Valuation \$ = \$3.29 billion
Standard Deviation	The high value of std dev: 7.31 which means that the data points are distributed over a wide range of values.
Minimum	The minimum value is \$1 billion, which is the least valuation of a Unicorn company
Maximum	The maximum value is \$140 billion, which is the highest valued Unicorn company
N	1037 is the number of observations for the analysis

2. Statistical Summary for column "Date Joined"

Analysis Variable : Date Joined Date Joined									
Mean	Std Dev	Minimum	Maximum	N					
22030.28	721.3772297	17367.00	22704.00	1037					

Statical Summary	Explanation
Mean	Average value = 22030.28
Standard Deviation	Value of standard deviation = 721.37
Minimum	Minimum value = 17367
Maximum	Maximum value = 22704
N	1037 is the number of observations for the analysis

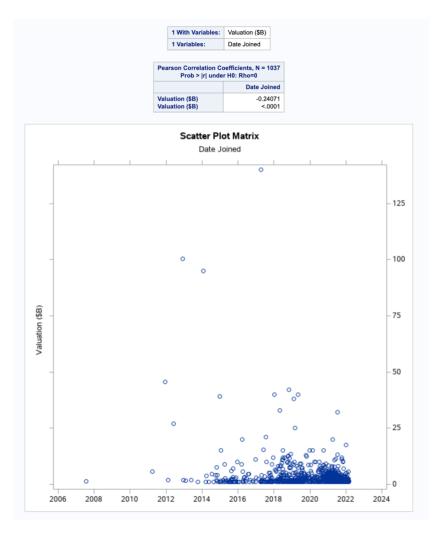
F. Statical Tests

1. One-way frequency

	The FREQ P	rocedure		
To	otal Raised in	\$ (Millions)	
Total Raised in \$ (Millions)	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 M	1	0.10	1	0.10
1.42 M	1	0.10	2	0.19
100 M	5	0.48	7	0.68
100.12 M	1	0.10	8	0.77
100.25 M	1	0.10	9	0.87
1003 M	1	0.10	10	0.96
101.62 M	1	0.10	11	1.06
101.97 M	1	0.10	12	1.16
1013 M	2	0.19	14	1.35
1021 M	1	0.10	15	1.45
1024 M	1	0.10	16	1.54
1026 M	1	0.10	17	1.64
1027 M	1	0.10	18	1.74
1039 M	1	0.10	19	1.83
1040 M	2	0.19	21	2.03
1042 M	1	0.10	22	2.12
1044 M	1	0.10	23	2.22
1050 M	1	0.10	24	2.31
1057 M	1	0.10	25	2.41
106.7 M	1	0.10	26	2.51
1061 M	1	0.10	27	2.60
1070 M	1	0.10	28	2.70
1072 M	1	0.10	29	2.80
1077 M	1	0.10	30	2.89
108.03 M	1	0.10	31	2.99
108.5 M	1	0.10	32	3.09
1089 M	1	0.10	33	3.18
109.4 M 1093 M	1	0.10	34	3.28
1093 M	1	0.10	35 36	3.38
1090 M	1	0.10	37	3.57
11.3 M	1	0.10	38	3.66
11.54 M	1	0.10	39	3.76
11.54 W	2	0.10	41	3.95
110.59 M	1	0.10	42	4.05
1100 M	1	0.10	43	4.15
1103 M	1	0.10	44	4.13
111.1 M	1	0.10	45	4.34
111.5 M	1	0.10	46	4.44
1127 M	1	0.10	47	4.53
1132 M	1	0.10	48	4.63
1135 M	1	0.10	49	4.73
115 M	2	0.19	51	4.92

An understanding of one variable for analysis is obtained by a "One-Way Frequency" analysis. The analysis for "*Total Raised in \$ (Million)*" is displayed in this table. The graph only shows the certain rows out of 1037 records because doing so would be not possible here. The column "*Frequency*" shows the number of values repeated in the "*Total Raised in \$ (Million)*". However, as we can also see from the complete table, that many amounts are unique.

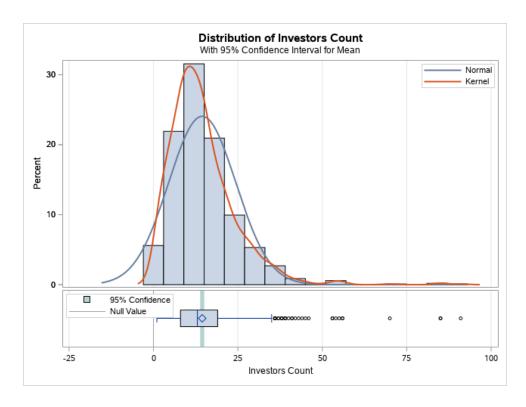
2. Correlation Analysis

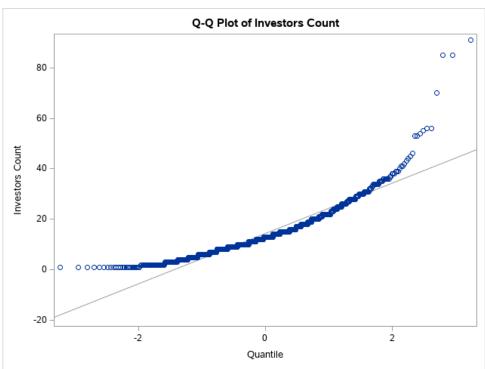


We aim to calculate the Pearson Correlation Coefficient in this correlation analysis. Generally, the sample correlation coefficient ranges from -1 to 1. The correlation between the two variables is quantified by this value. The above example shows a moderate correlation between "*Valuation* (\$)" and "*Date Joined*" based on the Coefficient value of -0.24071. Because all data points had too many values overlaid over one another, I only used a sample of about 60% of the data for this data frame.

3. T-Test

Variable: Investors Count										
N		Mean	St	d Dev	St	d Eı	r	Min	nimum	Maximum
1036	14	.4324	9	9.9465	0	.309	0	•	1.0000	91.0000
Me	an	95% C		CL Mean		an Std		Dev 95% C		L Std Dev
14.43	24	13.82	60	15.0	388	9	.946	65	9.5359	10.3944
				DF	t Val	ue	Pr	> t	1	
		10	1035 46		70	<.0	0001			





Description	Explanation					
Null Hypothesis	Ho = There is no significance of the Investors Count from the					
	mean					
Alternative Hypothesis	Ha = There is a significance of the Investors Count from the mean					
Significant Value	Alpha (α) = 0.05					
Comparison	0.01 < 0.05					
	We reject the Ho. This means there is a significant gap in the Investors count from the mean					

G. References

- 10, D., | by Dave Gilson, & Samp; Gilson, D. (2021, December 10). What makes Unicorns Special? these numbers may hold the answers. Stanford Graduate School of Business.
 Retrieved October 30, 2022, from https://www.gsb.stanford.edu/insights/what-makes-unicorns-special-these-numbers-may-hold-answers
- 2. SoFi. (2021, November 8). What are unicorn companies? SoFi. Retrieved October 30, 2022, from https://www.sofi.com/learn/content/what-is-a-unicorn-company
- 3. What is a unicorn company? PitchBook. (n.d.). Retrieved October 30, 2022, from https://pitchbook.com/blog/what-is-a-unicorn