Understanding a Unicorn

Daniel gaza

1/10/2022

```
What Is a Unicorn?
```

A unicorn is a privately held startup company with a valuation of more than $1 billion. It's popular in the venture capital industry. Aileen Lee, a venture capitalist, popularized the term in 2013. Unicorns are extremely rare and necessitate creativity. Because of their size, unicorn investors are typically private investors or venture capitalists, making them out of reach for retail investors. Although it is not required, many unicorns work their way up to becoming public.

This project will involve working with Unicorn Companies data to answer a series of questions about Unicorn companies. This project's analysis will be carried out using R.

```
#Loading Libraries we'll be using for this project
library(tidyverse)

## Warning in as.POSIXlt.POSIXct(Sys.time()): unable to identify current
timezone 'C':
## please set environment variable 'TZ'

## ── Attaching packages ─────────────────────────────────────── tidyverse
1.3.1 ──

## ✓ ggplot2 3.3.6     ✓ purrr   0.3.4
## ✓ tibble  3.1.7     ✓ dplyr   1.0.9
## ✓ tidyr   1.2.0     ✓ stringr 1.4.0
## ✓ readr   2.1.2     ✓ forcats 0.5.1

## ── Conflicts ──────────────────────────────────────────
tidyverse_conflicts() ──
## ✕ dplyr::filter() masks stats::filter()
## ✕ dplyr::lag()    masks stats::lag()

library(lubridate)

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
library(ggplot2)
library(readr)
library(dplyr)
library(tidyr)
```

With Polarity Groups launching a new venture, Polarity Ventures, investing in and funding start-ups, small businesses with impressive growth and high growth potential will be the best bet for breaking into the venture capital world. Unicorn Companies, startups, and small businesses with the potential to become Unicorns are the target market.

This open data set from datacamp.com was discovered during one of our research projects. This data set will be used to determine what kind of market we are entering and what to look for and off when evaluating a business.

The following are the questions that will be addressed in this project:

• Which Unicorn Companies have had the biggest return on investments?

• How long does it usually take for a company to become a unicorn? Has it always been this way?

• Which countries have the most unicorns? Are there any cities that appear to be industry hubs?

• Which investors have funded the most unicorns?

The dataset contains a record of private companies with a valuation of more than $1 billion as of March 2022, including each company's current valuation, funding, country of origin, industry, select investors, and the years it was founded and became a unicorn.

```r
#Loading dataset a csv file, and assigning date to date datatype
Unicorn_Companies <- data.frame(read_csv("Unicorn_Companies.csv", #setting
date in date datatype
                              col_types = cols(Date_Joined = col_date(format
= "%m/%d/%Y"))))
head(Unicorn_Companies)
```

```
##     Company Valuation Date_Joined                        Industry
City
## 1 Bytedance     $180B  2017-04-07         Artificial intelligence
Beijing
## 2    SpaceX     $100B  2012-12-01                           Other
Hawthorne
## 3     SHEIN     $100B  2018-07-03 E-commerce & direct-to-consumer
Shenzhen
## 4    Stripe      $95B  2014-01-23                         Fintech San
Francisco
## 5    Klarna      $46B  2011-12-12                         Fintech
Stockholm
## 6     Canva      $40B  2018-01-08    Internet software & services   Surry
Hills
```

```
##         Country     Continent Year_Founded Funding
## 1         China          Asia         2012     $8B
## 2 United States North America         2002     $7B
## 3         China          Asia         2008     $2B
## 4 United States North America         2010     $2B
## 5        Sweden        Europe         2005     $4B
## 6     Australia       Oceania         2012   $572M
##                                                            Select_Investors
## 1  Sequoia Capital China, SIG Asia Investments, Sina Weibo, Softbank Group
## 2              Founders Fund, Draper Fisher Jurvetson, Rothenberg Ventures
## 3 Tiger Global Management, Sequoia Capital China, Shunwei Capital Partners
## 4                              Khosla Ventures, LowercaseCapital, capitalG
## 5        Institutional Venture Partners, Sequoia Capital, General Atlantic
## 6             Sequoia Capital China, Blackbird Ventures, Matrix Partners
```

The data arrived in a format that would have made analysis extremely difficult; a restructuring was required to make the analysis possible; changes such as substituting B (for billions) for actual zeros (000..) and setting appropriate datatypes were required. Symbols and letters must be removed in order to assign an appropriate datatype. NAs in funding are funds invested by investors that are not recorded in the dataset and cannot be accessed from any sources, so NAs were converted to zeros for ease of computation. Data Manipulation

```
#Resigning Valuation column to appropriate Datatype
#For Valuation
Unicorn_Companies<-mutate(Unicorn_Companies,Valuation = substring(Valuation, 
2))
Unicorn_Companies<-
mutate(Unicorn_Companies,Valuation=gsub("B","000000000",Unicorn_Companies$Val
uation))
Unicorn_Companies<- mutate(Unicorn_Companies,Valuation = 
as.double(Valuation))

#For Funding
Unicorn_Companies<-
mutate(Unicorn_Companies,Funding=gsub("B","000000000",Unicorn_Companies$Fundi
ng))
Unicorn_Companies<-
mutate(Unicorn_Companies,Funding=gsub("M","000000",Unicorn_Companies$Funding)
)
Unicorn_Companies<-mutate(Unicorn_Companies,Funding = substring(Funding, 2))
Unicorn_Companies<- mutate(Unicorn_Companies,Funding = as.double(Funding))
```

```
## Warning in mask$eval_all_mutate(quo): NAs introduced by coercion
```

```
#Substituting "NA" to 0, for easy computation.
Unicorn_Companies<-mutate(Unicorn_Companies, Funding = 
replace(Unicorn_Companies$Funding,is.na(Unicorn_Companies$Funding),0))
head(Unicorn_Companies)
```

```
##     Company Valuation Date_Joined                        Industry
City
## 1 Bytedance   1.8e+11  2017-04-07         Artificial intelligence
Beijing
## 2    SpaceX   1.0e+11  2012-12-01                           Other
Hawthorne
## 3     SHEIN   1.0e+11  2018-07-03 E-commerce & direct-to-consumer
Shenzhen
## 4    Stripe   9.5e+10  2014-01-23                         Fintech San
Francisco
## 5    Klarna   4.6e+10  2011-12-12                         Fintech
Stockholm
## 6     Canva   4.0e+10  2018-01-08    Internet software & services   Surry
Hills
##         Country     Continent Year_Founded  Funding
## 1         China          Asia         2012 8.00e+09
## 2 United States North America         2002 7.00e+09
## 3         China          Asia         2008 2.00e+09
## 4 United States North America         2010 2.00e+09
## 5        Sweden        Europe         2005 4.00e+09
## 6     Australia       Oceania         2012 5.72e+08
##                                                          Select_Investors
## 1  Sequoia Capital China, SIG Asia Investments, Sina Weibo, Softbank Group
## 2              Founders Fund, Draper Fisher Jurvetson, Rothenberg Ventures
## 3 Tiger Global Management, Sequoia Capital China, Shunwei Capital Partners
## 4                                Khosla Ventures, LowercaseCapital, capitalG
## 5        Institutional Venture Partners, Sequoia Capital, General Atlantic
## 6               Sequoia Capital China, Blackbird Ventures, Matrix Partners
```

The data is now ready for analysis after all necessary cleaning and formatting. Now we will investigate the questions that have been posed for this project; the results are only as good as the questions posed. Data always tells you what you ask it, so no matter how much potential a data has, if you don't ask the right questions, you won't get anything useful out of it. To achieve the desired results, four (4) questions were developed for this analysis.

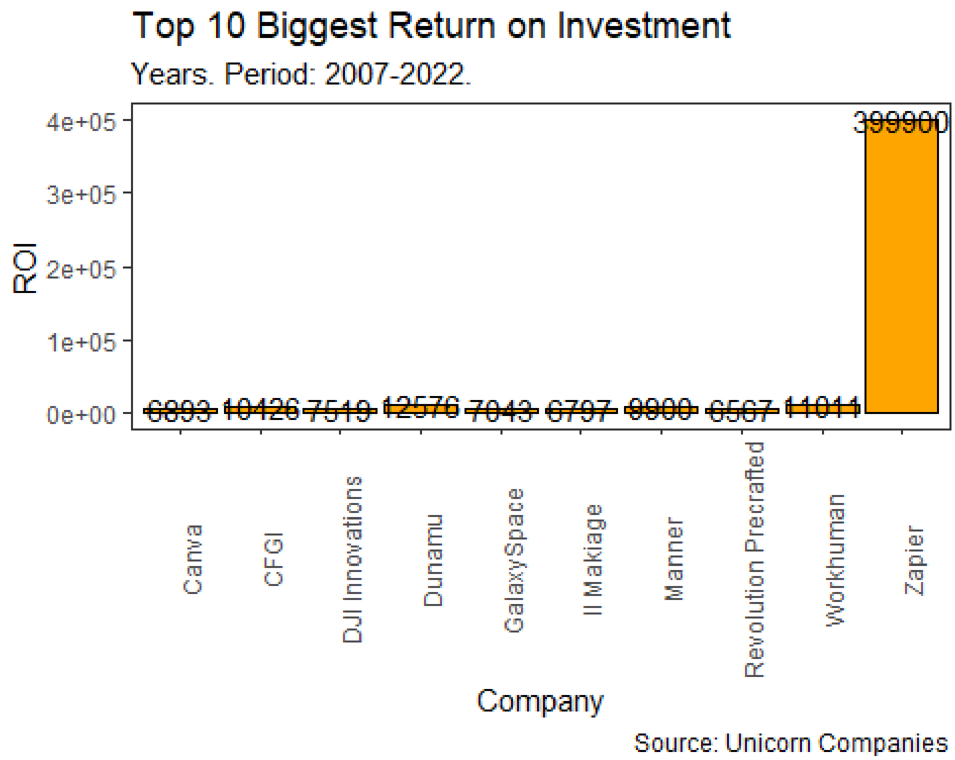
1. Which Unicorn Companies have had the biggest return on investments?

To determine which companies have the highest return on investment, we must first calculate the return on investment (ROI) for all of the companies in the data. By calculating the ROI, we can correctly and accurately determine which companies have the highest return on investment from all investments made on this data set. Naming the column with the results as ROI (Return On investment).

```
#Creating a table to house return On Investment. Keeping Calculated columns
separate from original data
return_on_investment<-Unicorn_Companies%>%
  select(Company,Select_Investors, Funding,Valuation)

#Calculating return on investment (ROI)
ROI<-((Unicorn_Companies$Valuation -
```

```
Unicorn_Companies$Funding)/Unicorn_Companies$Funding)*100

#Fixing the result into our table
return_on_investment<-mutate(return_on_investment,ROI = round(ROI))#removing
decimal point

#filtering to remove Companies without funding,after the calculation their
ROI returned as Inf
return_on_investment<-return_on_investment%>%
 filter(ROI != "Inf")%>%
   arrange(desc(ROI))
head(return_on_investment)

##            Company
## 1           Zapier
## 2           Dunamu
## 3        Workhuman
## 4             CFGI
## 5           Manner
## 6 DJI Innovations
##                                                        Select_Investors
Funding
## 1      Sequoia Capital, Bessemer Venture Partners, Threshold Ventures
1.00e+06
## 2 Qualcomm Ventures, Woori Investment, Hanwha Investment & Securities
7.10e+07
## 3                                                                  ICG
9.00e+06
## 4                                 The Carlyle Group, CVC Capital Partners
1.90e+07
## 5                             Coatue Management, H Capital, Capital Today
1.00e+07
## 6                                         Accel Partners, Sequoia Capital
1.05e+08
##   Valuation    ROI
## 1     4e+09 399900
## 2     9e+09  12576
## 3     1e+09  11011
## 4     2e+09  10426
## 5     1e+09   9900
## 6     8e+09   7519

# Limiting to Top 10 companies for better presentation
Top.10.ROI<-return_on_investment%>%
  top_n(10)

## Selecting by ROI

# A bar chart showing what all this code means
ggplot(Top.10.ROI, aes(x=Company, y=ROI))+
  geom_bar(color="black",fill="orange",stat="identity")+
```

```
  labs(title = "Top 10 Biggest Return on Investment",
       subtitle="Years. Period: 2007-2022.",
       caption="Source: Unicorn Companies")+
  geom_text(aes(label=signif(ROI)),nudge_y = 1)+
  theme_test()+
  theme(axis.text.x = element_text(angle = 90))
```



Companies with a high valuation do not necessarily represent the best investments, because a large amount of money could have been invested in the company, resulting in a high valuation. Looks can be deceiving when it comes to a company that received a small amount of funding and has done exceptionally well, but its valuation cannot be compared to companies that received much larger amounts of funding. According to the bar chart above, these companies are not even in the top 20 when it comes to valuation.

Moving on.

2. How long does it usually take for a company to become a unicorn? Has it always been this way?

PPolarity Ventures wants to know how long it will take their principal to reach maturity, which is defined as the company being valued at $1 billion while remaining privately held. We can calculate the difference in years by using the year each company was founded and the year they achieved Unicorn status. We can then draw conclusions and look for patterns in the differences.

```
# Extracting the year companies reached unicorn status as Year Reached
Unicorn_Companies<-mutate(Unicorn_Companies,Year_Reached = year(Date_Joined))
```

```r
# Difference in years will give us duration in years
Unicorn_Companies<-mutate(Unicorn_Companies,diffInYear = Year_Reached - 
Year_Founded)

# DiffInYear, grouping the duration (diffInYear) to see how many companies 
fall into each duration of years
diffInYearCount<-Unicorn_Companies%>%
  count(diffInYear,name = "CountofdiffInYear", sort = TRUE)%>%
  arrange(desc(CountofdiffInYear))
diffInYearCount
```

```
##    diffInYear CountofdiffInYear
## 1           6               138
## 2           5               129
## 3           4               125
## 4           7               107
## 5           3                98
## 6           8                88
## 7           2                68
## 8          10                62
## 9           9                52
## 10          1                36
## 11         11                36
## 12         14                20
## 13         13                19
## 14         12                17
## 15         16                10
## 16         17                10
## 17          0                 9
## 18         15                 9
## 19         19                 7
## 20         18                 6
## 21         20                 6
## 22         21                 6
## 23         22                 5
## 24         25                 2
## 25         27                 2
## 26         37                 2
## 27         -4                 1
## 28         24                 1
## 29         26                 1
## 30         28                 1
## 31         98                 1
```

The created diffInYearCount table shows how many companies fall into each category, as well as how long it took each company to become a unicorn. In order to accurately answer the question of how long it takes for a company to become a unicorn and if it has already done so, some criteria must be considered, such as the year the company was founded, the continent of the company, and so on.

```r
#how long it takes
average_diffInYear<-round(mean(Unicorn_Companies$diffInYear),2)
paste(average_diffInYear,"years",sep = "")

## [1] "7years"
```

It takes an average of 7 years for a company to become a unicorn. There have been companies that have taken 40 years or more to become unicorns. The year the company was founded will be grouped. The first group (group 1) will be for companies founded between 1919 and 2000, and the second group (group 2) will be for companies founded between 2001 and 2021.

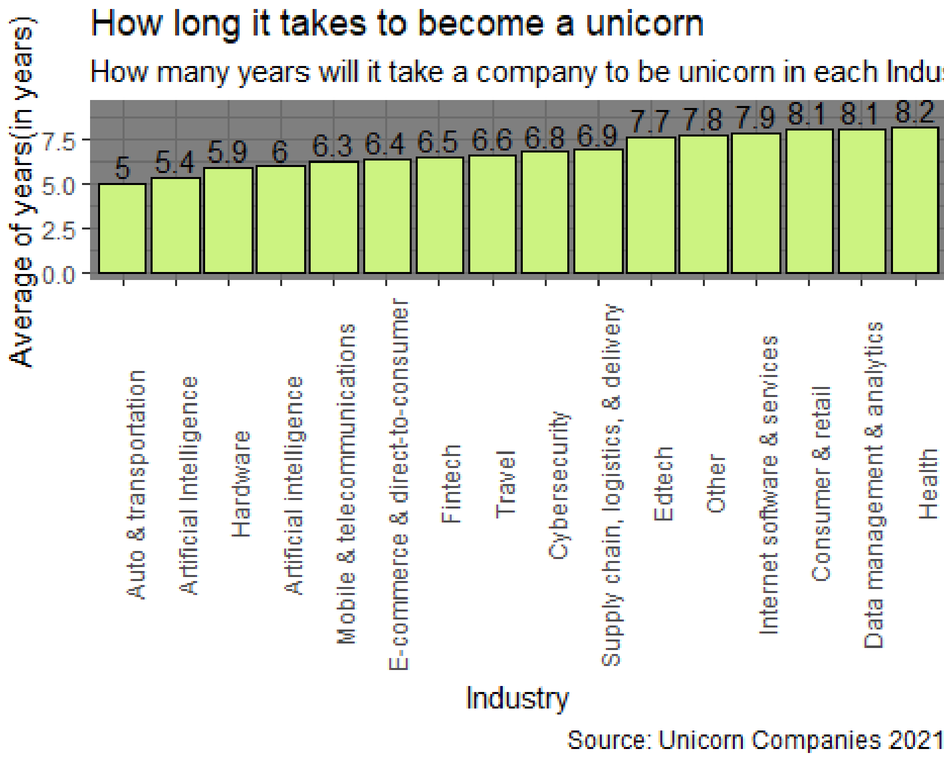
```r
#Creating filter criteria with year founded
start<-1919
end<-2000
start1<-2001
end1<-2021

#Getting the average duration for companies created from 1919 to 2000(group1)
group1<-Unicorn_Companies%>%
  filter(between(Year_Founded,start,end))%>%
  summarise(Avg.year = mean(diffInYear,na.rm = TRUE))%>%
  mutate(Avg.year= round(Avg.year,1))
group1

##   Avg.year
## 1     23.7

#Getting the average duration for companies created from 2001 to 2021(group2)
group2<-Unicorn_Companies%>%
  filter(between(Year_Founded,start1,end1))%>%
  summarise(Avg.year = mean(diffInYear,na.rm = TRUE))%>%
  mutate(Avg.year= round(Avg.year,1))
group2

##   Avg.year
## 1      6.4

# Average year to become a unicorn
Avg.year<-data.frame(group= c("group1","group2"),
             Avg.year=  c(23.7,6.4),
             Timelime= c("1919 to 2000","2001 to 2021"))
Avg.year

##    group Avg.year     Timelime
## 1 group1     23.7 1919 to 2000
## 2 group2      6.4 2001 to 2021
```

According to the table above, it took group1 23.7 years on average to become a unicorn, while group2 took 6.41 years. Taking into account other factors, if Polarity Ventures wishes to invest in a company, we can now easily access or evaluate the time it will take an investment to mature. We can see how long it takes companies in a specific industry to become a unicorn, and we can further narrow it down by country, industry, and continent.

```
#Average years for companies in a Industry to become unicorn
duration_Industry<-Unicorn_Companies%>%
  group_by(Industry)%>%
  summarise(Avg.year =mean(diffInYear), count_of_company = n())%>%
  mutate(Avg.year = round(Avg.year,1))%>%
  arrange(desc(count_of_company))
duration_Industry
```

```
## # A tibble: 16 × 3
##    Industry                            Avg.year count_of_company
##    <chr>                                  <dbl>            <int>
##  1 Fintech                                  6.5              224
##  2 Internet software & services             7.9              205
##  3 E-commerce & direct-to-consumer          6.4              111
##  4 Health                                   8.2               74
##  5 Artificial intelligence                  6                 73
##  6 Other                                    7.8               58
##  7 Supply chain, logistics, & delivery      6.9               57
##  8 Cybersecurity                            6.8               50
##  9 Data management & analytics              8.1               41
## 10 Mobile & telecommunications              6.3               38
## 11 Hardware                                 5.9               34
## 12 Auto & transportation                    5                 31
## 13 Edtech                                   7.7               28
## 14 Consumer & retail                        8.1               25
## 15 Travel                                   6.6               14
## 16 Artificial Intelligence                  5.4               11
```

```
# Getting a clear picture
ggplot(duration_Industry, aes(x=reorder(as.character(Industry),Avg.year),
y=Avg.year)) +
  geom_col(fill="#CCF381",color = "black")+
  labs(title="How long it takes to become a unicorn",
       subtitle = "How many years will it take a company to be unicorn in
each Industry",
       caption="Source: Unicorn Companies 2021",
       x="Industry",
       y="Average of years(in years)")+
  geom_text(aes(label=signif(Avg.year)),nudge_y = 1)+
  theme_dark()+
  theme(axis.text.x = element_text(angle = 90))
```



The chart above shows how long it takes for a company in a specific industry to become a unicorn.

Is this how it has always been? No, according to our above analysis, the average time it takes for a company to become a unicorn has decreased to 6.41 years since 2001, from over 23 years previously. To go deeper and gain a better understanding of countries, industries, and cities, simply substitute the criteria you want in the groupby statement, and the mean (average) of the duration (diffInYear) can be determined. When evaluating a potential investment, this information can help predict how long the investment will be profitable.

3. Which countries have the most unicorns? Are there any cities that appear to be industry hubs?

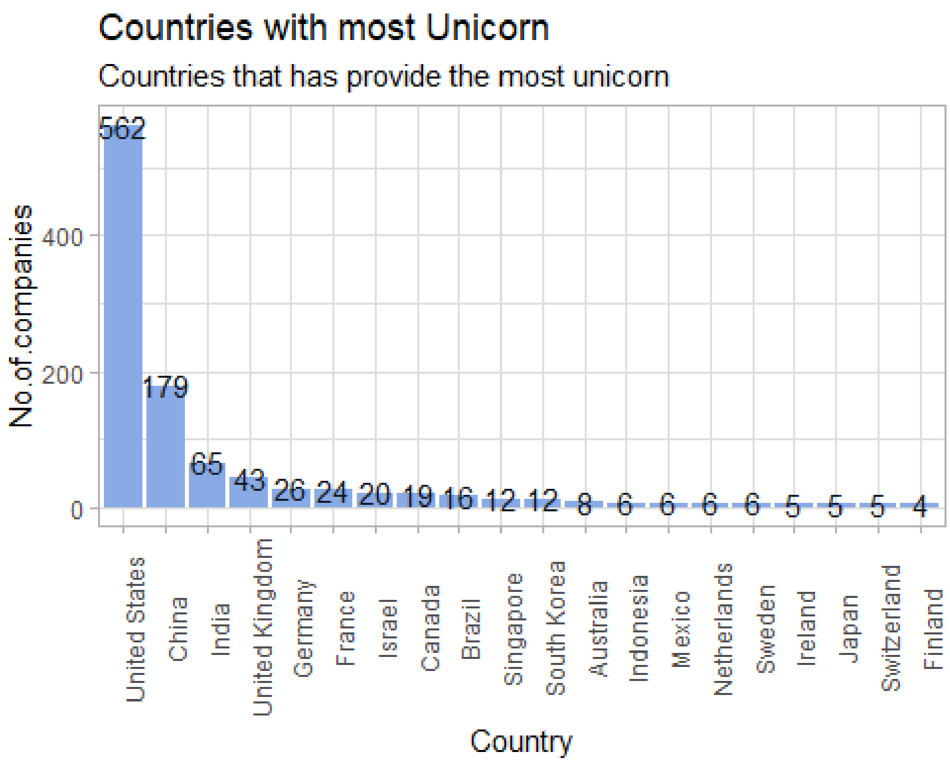
Let us now determine where our primary focus will be, our market. Wants to understand. Now we'll see how many unicorns come from each country to determine where our main focus will be placed.

```
#Change Hong kong to China in region
Unicorn_Companies<-mutate(Unicorn_Companies, Country = gsub("Hong Kong",
"China",Unicorn_Companies$Country))

# Grouping to see how many unicorn from each country
CountryCount<-Unicorn_Companies%>%
  count(Country,name = "CountofCountry", sort = TRUE)%>%
  arrange(desc(CountofCountry))
head(CountryCount)
```

```
##          Country CountofCountry
## 1  United States            562
## 2          China            179
## 3          India             65
## 4 United Kingdom             43
## 5        Germany             26
## 6         France             24
```

From 1919 to 2021, the top six countries in terms of unicorn births were the United States, China, India, the United Kingdom, Germany, and France. With more information, we can select the best market for Polarity Ventures..

```
#Top 20 countries
Top.20.countries<-CountryCount[1:20,]

# How many unicorns each Country have provided
ggplot(Top.20.countries,aes(x= reorder(Country,-CountofCountry),y=
CountofCountry))+
  geom_col(fill="#8AAAE5")+
  labs(title = "Countries with most Unicorn",
       subtitle = "Countries that has provide the most unicorn",
       x= "Country",
       y= "No.of.companies")+
  geom_text(aes(label=signif(CountofCountry)), nudge_y =
1,color="#101820FF")+
  theme_light()+
  theme(axis.text.x = element_text(angle = 90))
```



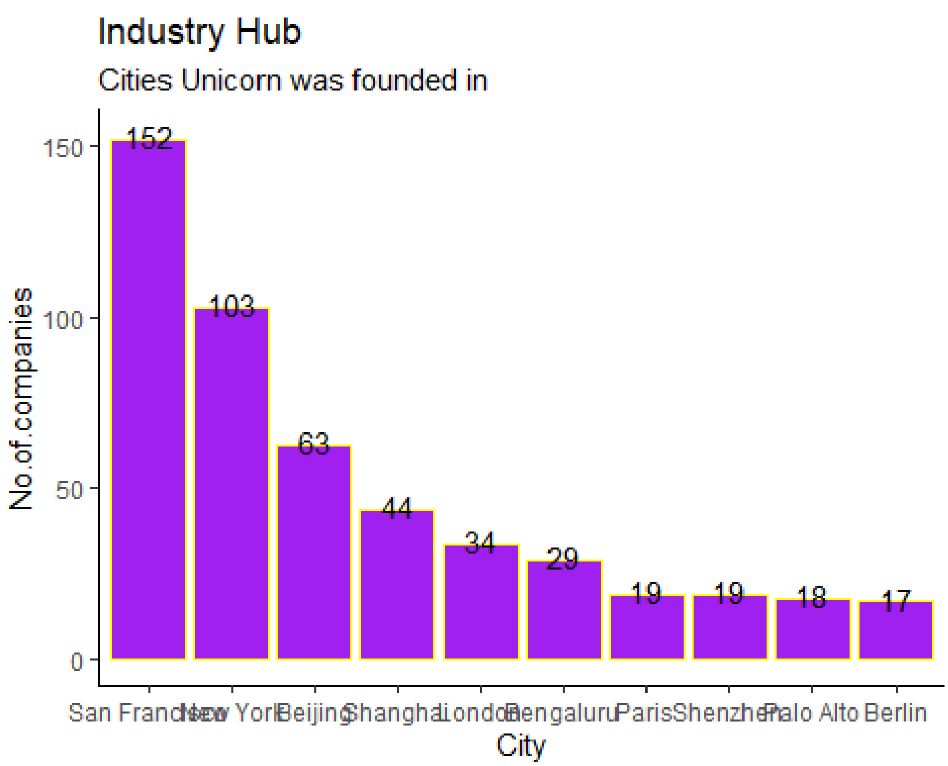
The numbers of unicorns countries have provides, with United States being the country with the most unicorns and China following suits, now we can choose and focus on a region or regions as our market.

Are there any cities that appear to be industry hubs?

Narrowing into cities to make use of the information we have concerning cities unicorn has been coming from we can investigate to see if some cities have favours compared to others and if we should be looking out for those cities.

```
#Top cities
CityCount<-Unicorn_Companies%>%
  count(City,name = "CountofCompany", sort = TRUE)%>%
  arrange(desc(CountofCompany))%>%
  mutate(CountofCompany2 = CountofCompany)%>%
  mutate(CountofCompany2 = as.character(CountofCompany2))

#Top 20 Cities with most unicorn
Top.Cities<-CityCount[1:10,]

# Presentation of the findings
ggplot(Top.Cities,aes(x=reorder(City,-CountofCompany),y=CountofCompany))+
  geom_col(color="yellow",fill= "purple")+
  labs(title = "Industry Hub",
       subtitle = "Cities Unicorn was founded in",
       x= "City",
       y= "No.of.companies")+
```

```
  geom_text(aes(label=signif(CountofCompany)),nudge_y = 1,color="black")+
  theme_classic()
```



To improve visibility, only the top ten cities were included. As a result, cities with more unicorns. If we choose a country as our target market, we can see which cities have purchased the most unicorns. The map above depicts the top cities in the world.

4. Which investors have funded the most unicorns?

The last question is to figure out investors who have funded the most unicorn, this could be an added advantage for us at Polarity Ventures, using this information, going forward we can watch closely the which sector and companies this top investors are investing in and also making considerations of whether to go into business with them. In the unicorn data, the investors column holds the together all investors who funded a business. For us to get the top investors we will have to separate each investor distinctively.

```
# Keeping columns we only need
Investors<-Unicorn_Companies%>%
  select(Company,Select_Investors)
Investors<-data.frame(Investors)

#separate the investors into 3 different columns
Investors<-
transform(Investors,test=do.call(rbind,strsplit(Select_Investors,",", fixed =
TRUE)), stringsAsFactors = F)
```

```
## Warning in (function (..., deparse.level = 1) : number of columns of
result is
## not a multiple of vector length (arg 2)
```

```
head(Investors)
```

```
##     Company
## 1 Bytedance
## 2    SpaceX
## 3     SHEIN
## 4    Stripe
## 5    Klarna
## 6     Canva
##                                                           Select_Investors
## 1  Sequoia Capital China, SIG Asia Investments, Sina Weibo, Softbank Group
## 2             Founders Fund, Draper Fisher Jurvetson, Rothenberg Ventures
## 3 Tiger Global Management, Sequoia Capital China, Shunwei Capital Partners
## 4                                Khosla Ventures, LowercaseCapital, capitalG
## 5        Institutional Venture Partners, Sequoia Capital, General Atlantic
## 6             Sequoia Capital China, Blackbird Ventures, Matrix Partners
##                          test.1                  test.2
## 1         Sequoia Capital China     SIG Asia Investments
## 2                 Founders Fund  Draper Fisher Jurvetson
## 3       Tiger Global Management    Sequoia Capital China
## 4               Khosla Ventures         LowercaseCapital
## 5 Institutional Venture Partners          Sequoia Capital
## 6         Sequoia Capital China       Blackbird Ventures
##                     test.3                          test.4
## 1               Sina Weibo                  Softbank Group
## 2      Rothenberg Ventures                   Founders Fund
## 3  Shunwei Capital Partners        Tiger Global Management
## 4                 capitalG                 Khosla Ventures
## 5          General Atlantic Institutional Venture Partners
## 6           Matrix Partners           Sequoia Capital China
```

Now that our investors have been separated, we'll remove the original investors column and any other columns that aren't relevant to this analysis so that it doesn't affect our results. And combining the separate columns into one column so that we only have one column for investors. Most companies had three investors, but few had four, so when we split the investors columns, companies with three investors were automatically filled in the fourth column with the first investor before the comma (test.4)

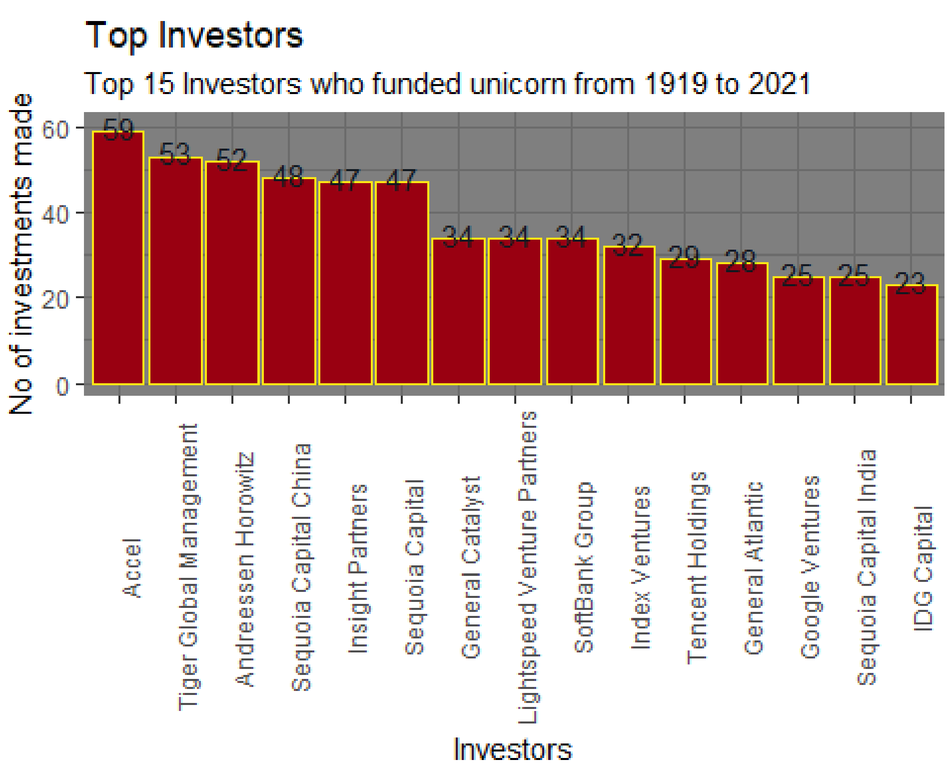
```
#remove other columns and combine remaining columns into one
Investors<-subset(Investors,select = -Select_Investors)

#Remove the fourth column
Investors<-subset(Investors,select = -test.4)

# here we'll fix in the investors of the companies with four investors
manually
```

```
Investors$test5<-ifelse(Investors$Company=="Bytedance","Softbank Group","")
Investors$test5<-ifelse(Investors$Company=="Niantic","Spark
Capital",Investors$test5)
Investors$test5<-ifelse(Investors$Company=="Rappi","Redpoint
e.ventures",Investors$test5)
Investors$test5<-ifelse(Investors$Company=="Yixia","Redpoint
ventures",Investors$test5)
Investors$test5<-ifelse(Investors$Company=="Lightricks","Goldman
Sachs",Investors$test5)

# Stacking the investors into one column
Investors<-Investors%>%
  cbind(Investors[1],stack(Investors[2:5]))

## Warning in data.frame(..., check.names = FALSE): row names were found from
a
## short variable and have been discarded

# Removing columns
drop<-c("test.1","test.2","test.3","test.4","ind","test5")
Investors<-Investors[,!(names(Investors)%in%drop)]
Investors<-subset(Investors,select = -Company)

# Renaming Columns
Investors<-rename(Investors,Investors = values, Company=Company.1)

# Trimming white spaces
Investors<-Investors%>%
  mutate_all(str_trim)

#Removing duplicates, that came about because of split and stack process.
Investors<-distinct(Investors,Company,Investors, .keep_all = TRUE)

#Top investors
Investors.count<-Investors%>%
  group_by(Investors)%>%
  count(Investors, name = "no.of.investment", sort = TRUE)%>%
  mutate(investments = as.character(no.of.investment))

#remove the row of the investor without a name
Investors.count<-Investors.count[-1,]

# Top 15 investors
Top.15.Investors<-Investors.count[1:15,]

head(Top.15.Investors)

## # A tibble: 6 × 3
## # Groups:   Investors [6]
##   Investors               no.of.investment investments
```

```
##   <chr>                                 <int> <chr>
## 1 Accel                                    59 59
## 2 Tiger Global Management                  53 53
## 3 Andreessen Horowitz                      52 52
## 4 Sequoia Capital China                    48 48
## 5 Insight Partners                         47 47
## 6 Sequoia Capital                          47 47
```

We have now successfully obtained our Top investors from 1919 to 2021, thus the timeline of this data. Accel on top with 59 unicorn investments, a fantastic achievement, and Tiger Global management following suit with 53 investments. Putting everything together, we'll make a chart to help us understand the codes and numbers.

```
#Presenting the results in more suitable manner
ggplot(Top.15.Investors,aes(x= reorder(Investors,-no.of.investment),y=
no.of.investment))+
  geom_col(fill="#990011FF",color="#FEE715FF")+
  labs(title= "Top Investors",
       subtitle= "Top 15 Investors who funded unicorn from 1919 to 2021",
       x= "Investors",
       y= "No of investments made")+
  geom_text(aes(label=signif(no.of.investment)),nudge_y =
1,color="#101820FF")+
  theme_dark()+
  theme(axis.text.x = element_text(angle = 90))
```



Polarity Ventures can successfully navigate the venture capital business with the results of this analysis, meeting the goal of targeting small businesses or start-ups with the potential to become unicorns. However, the information contained in this data (unicorn companies) is insufficient to predict whether a company will become a unicorn; much more information will be required to reach such conclusions; the result of this analysis, to some extent, provides an idea in meeting Polarity Ventures' goal.

```
Insights
```

Finally, all of the questions posed for this analysis were duly addressed by; 1. Zappier had the highest return on investment, with a ROI of 399900%, which is an excellent return on investment.

2. How long does it take for a company to become a unicorn? On average, we found that it takes a startup 7 years, but when we take some criteria into account, such as the year founded, we find that it takes approximately 24 years for startups founded between 1919 and 2000. From 2001 to 2021, it takes an average of 7 years; we also looked into how long it will take startups in various industries.

3. Countries with the most unicorns, the United States (USA) came in first with over 500 unicorns, with China close behind, and San Franceso identified as an industry hub.

4. The final question was to identify the top investors who had funded the most unicorns. Accel has funded 59 unicorns, which is a good track record, and Tiger Global management has funded 53 unicorns.