Stocking up on Data

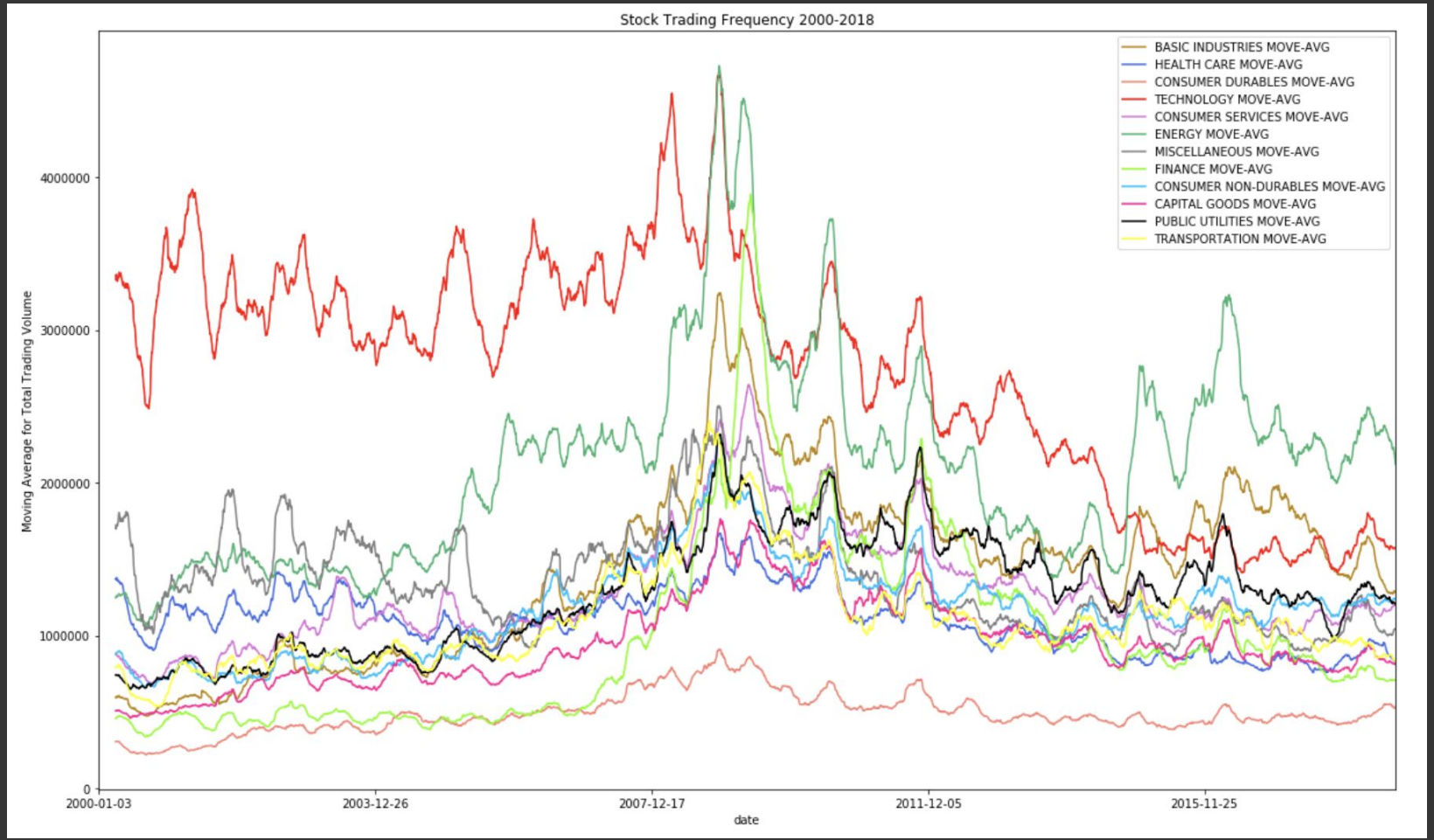
Devin Barger, Sophie Wang, Amy Wei

Trading Volume vs. Stock Price

- Both stock price and volume increase
- Stock price increases and volume decreases
- Both price and volume decrease
- Stock price decreases and volume increases

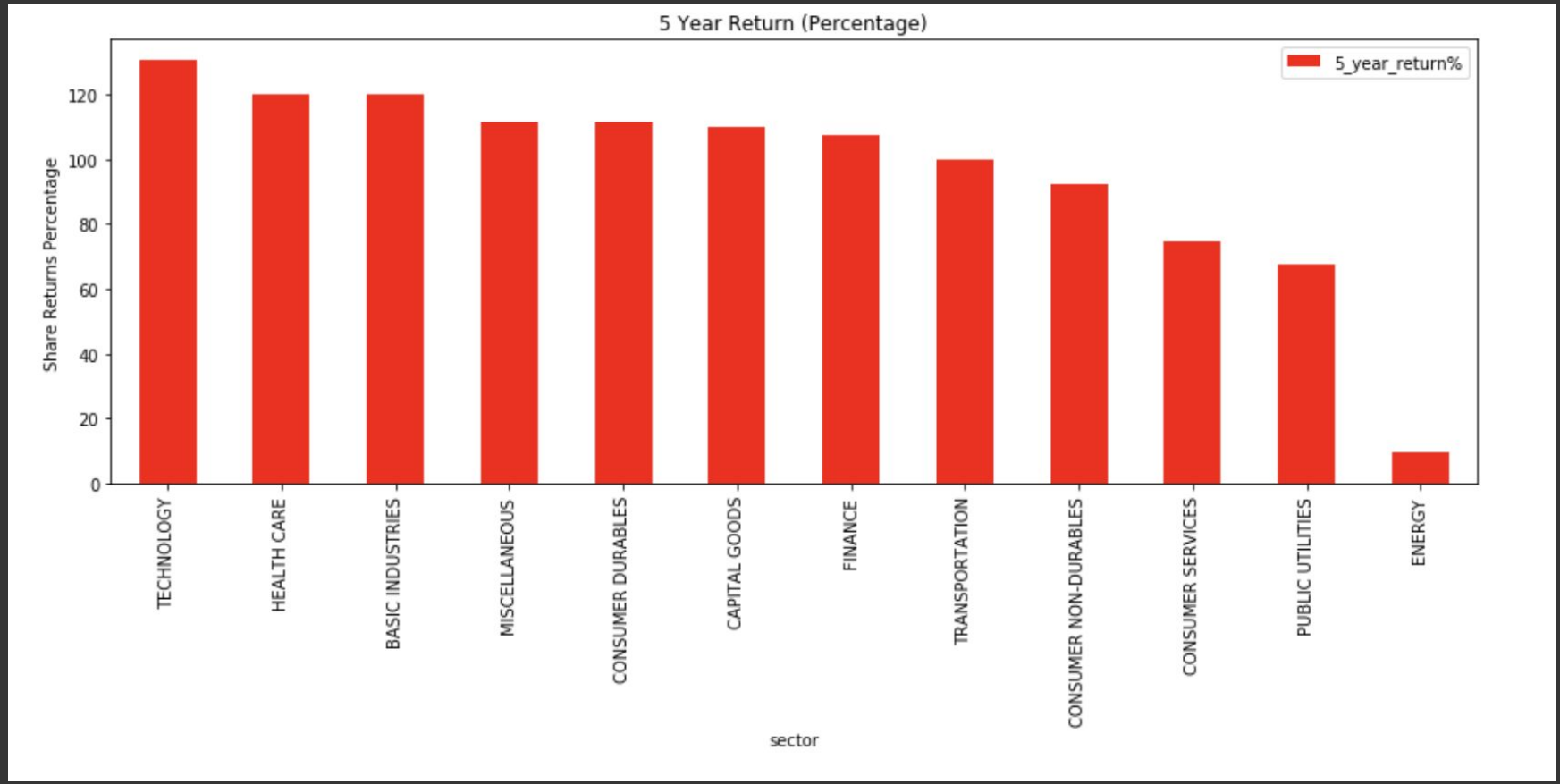
What is the overall trend of trading volume by industry?

General Trading Frequency 2000-2018

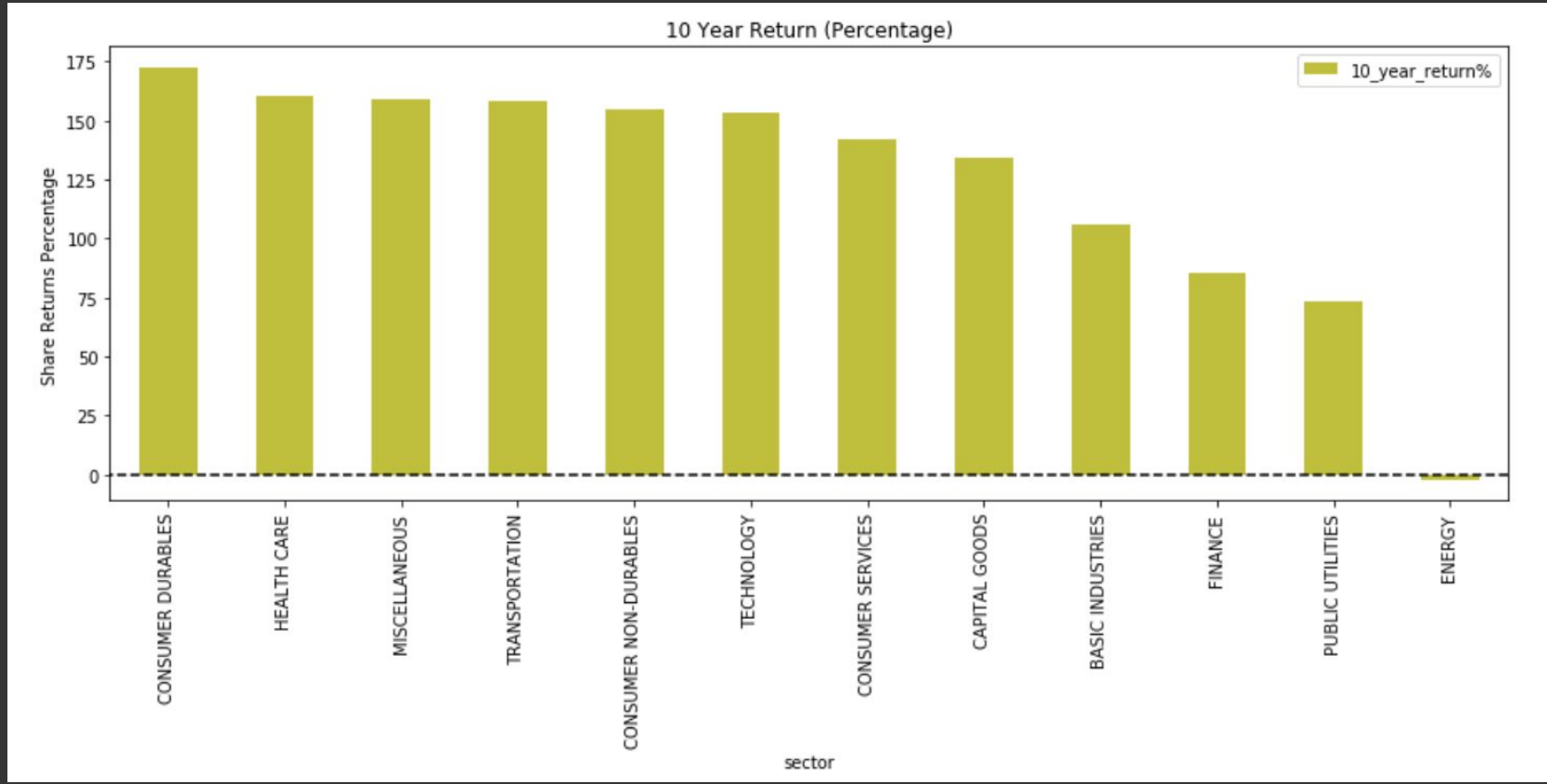


What industry sees the best return? (5-year, 10-year, 15-year)

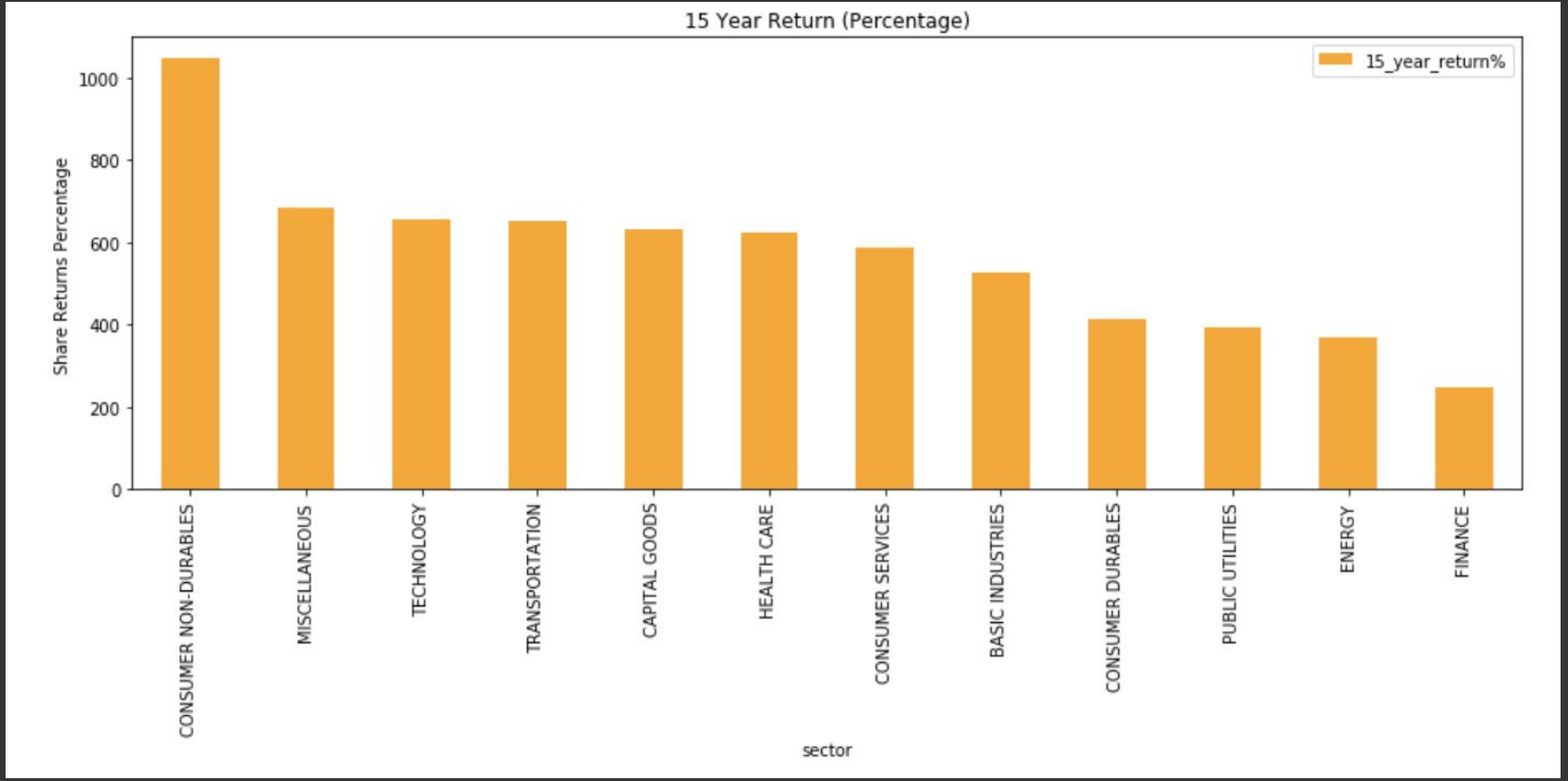
Industry Five Year Return (Percentage) 2013-2018



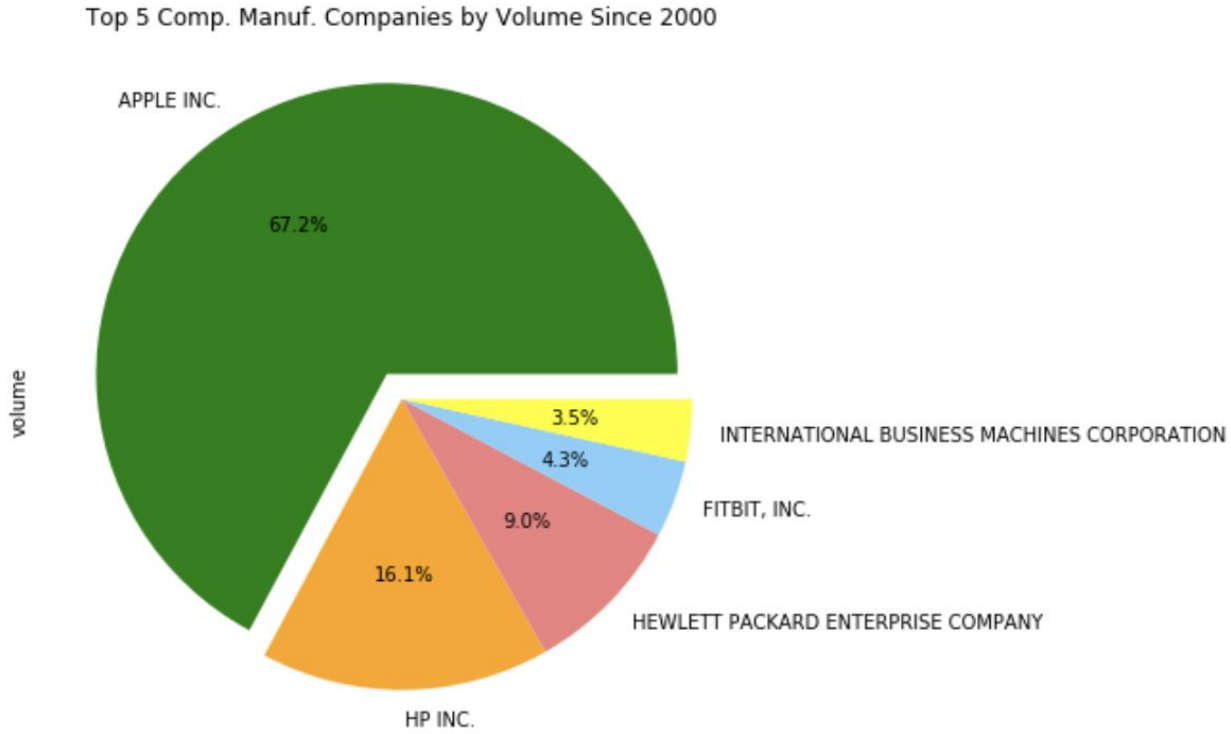
Industry Ten Year Return (Percentage) 2008-2018



Industry Fifteen Year Return (Percentage) 2003-2018



Which individual stock from our dataset is the most popular since 2000?



Apple Inc. wins!! (overwhelmingly)

Can average annual return tell everything?

What is CAGR?

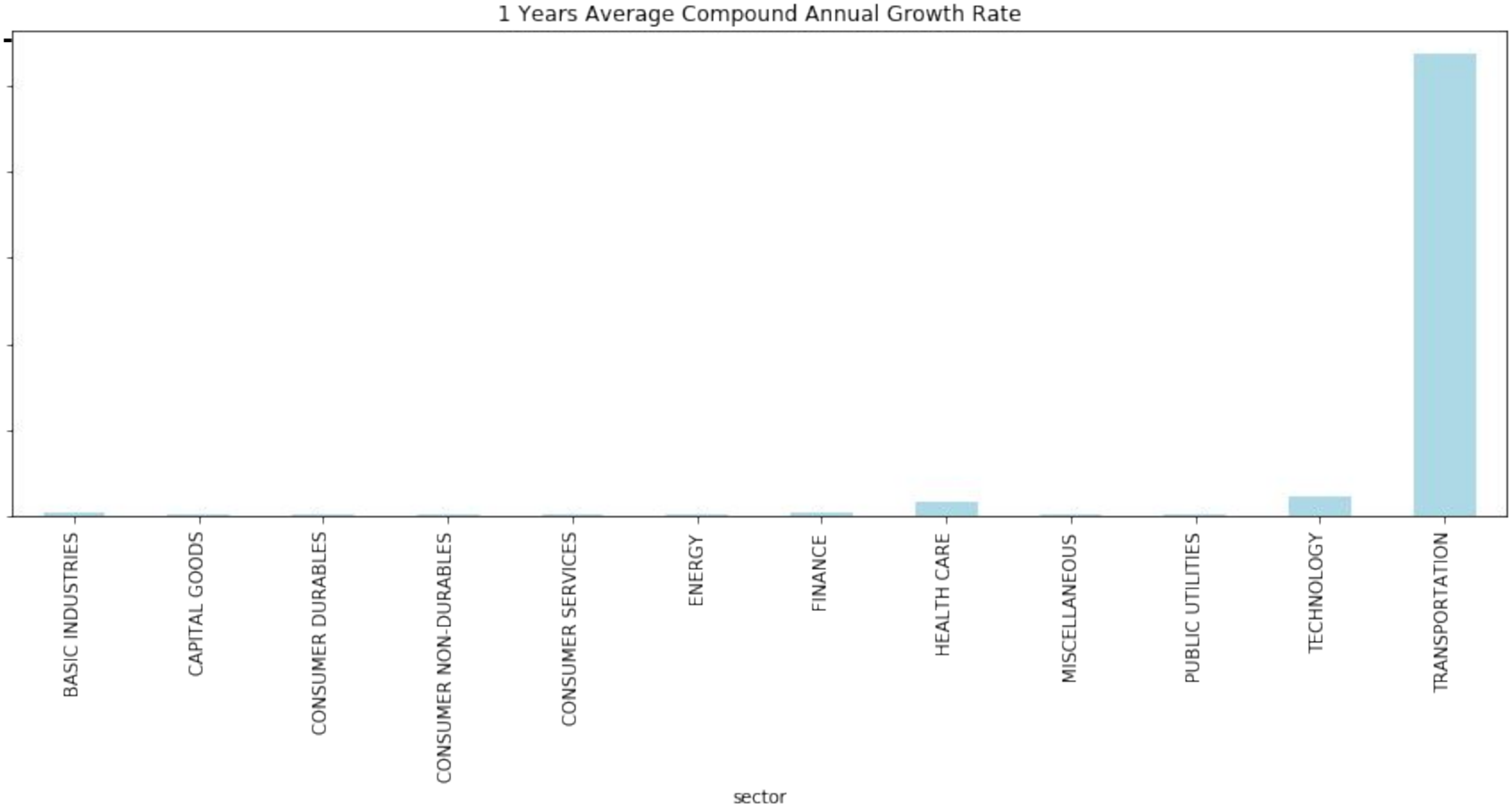
CAGR = [(EV / BV) ^ 1/n] - 1

EV=ending value of an investment;
BV=beginning value of an investment;
n = number of years

What does CAGR tells?

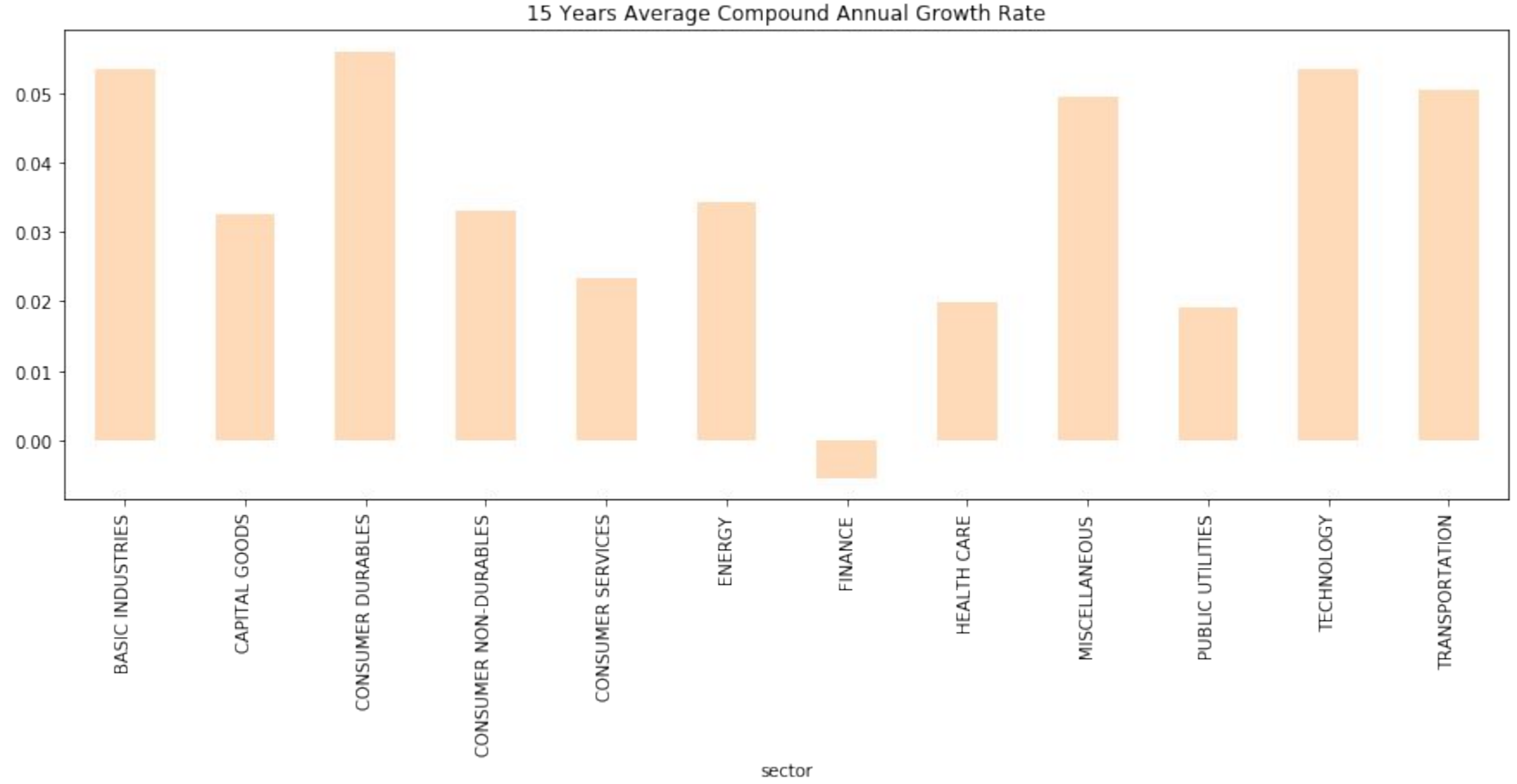
While the annual rate of return gives the amount of return every year, CAGR gives the return over the entire period of the investment. CAGR can be used to smooth returns so that they may be more easily understood when compared to alternative investments.

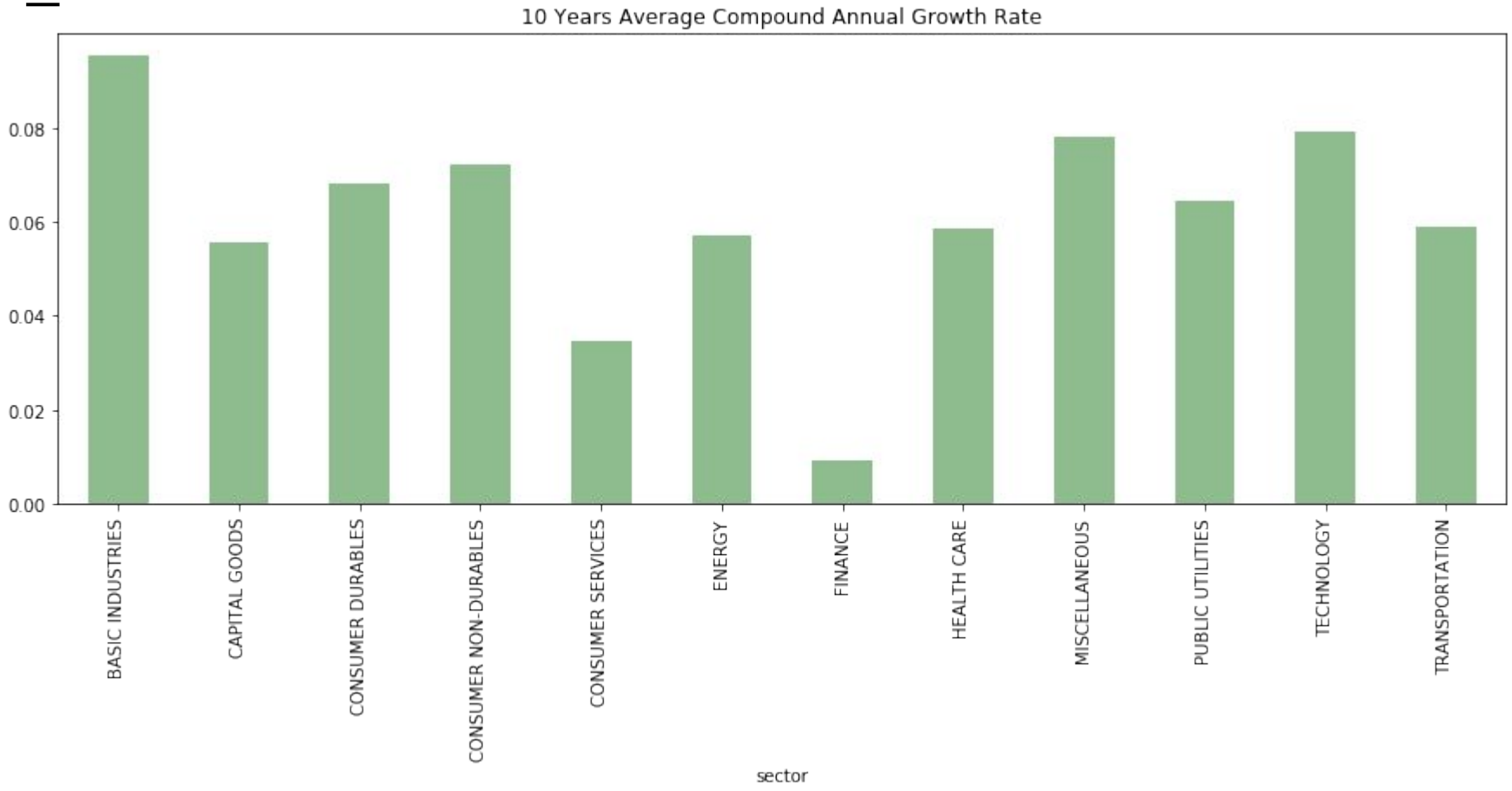
Since CAGR is is representing the return for a time period. The one year return makes no sense for CAGR.

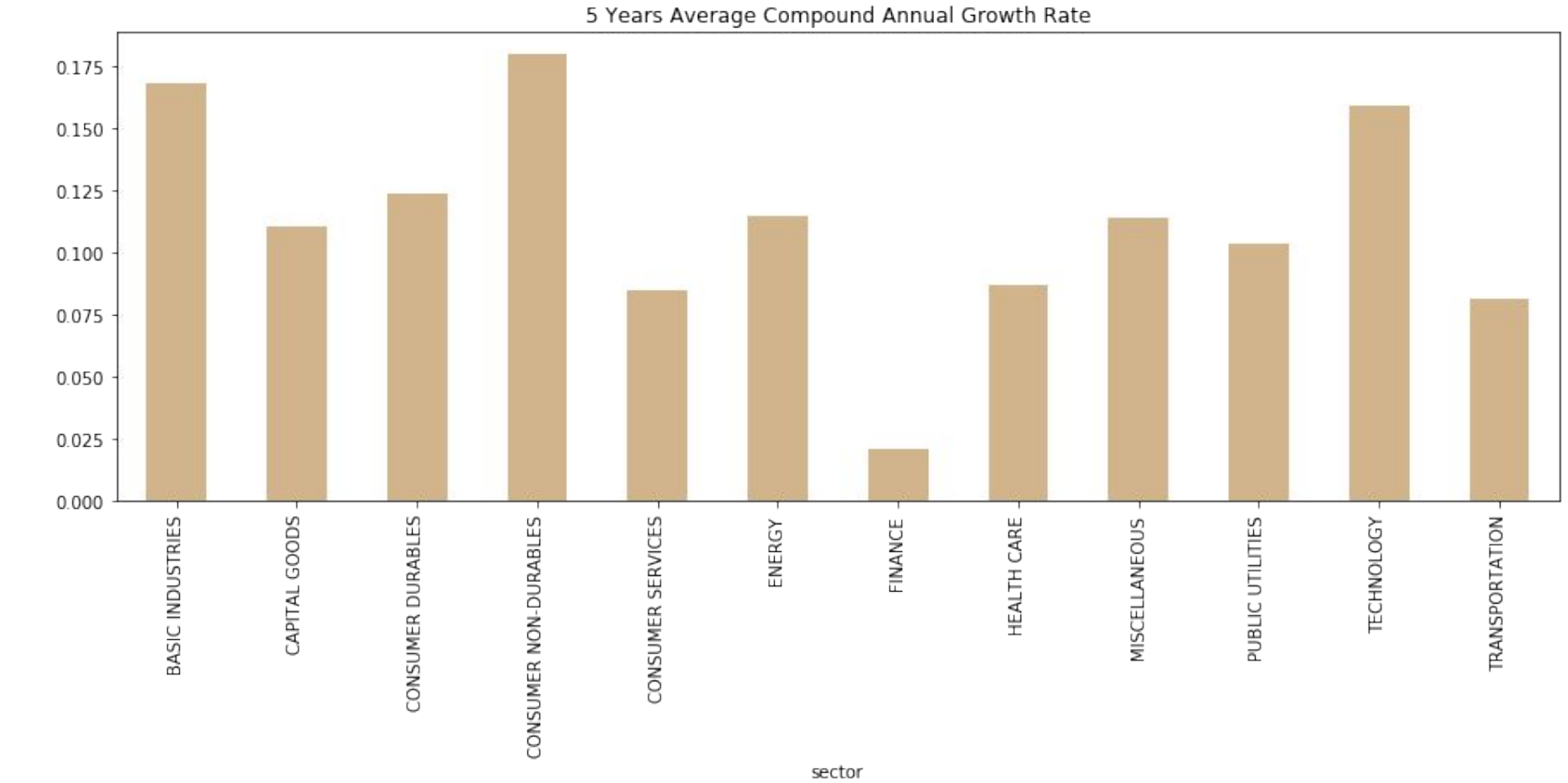

1 Years Average Compound Annual Growth Rate
500
400
300
200
100
0
BASIC INDUSTRIES
CAPITAL GOODS
CONSUMER DURABLES
CONSUMER NON-DURABLES
CONSUMER SERVICES
ENERGY
FINANCE
HEALTH CARE
MISCELLANEOUS
PUBLIC UTILITIES
TECHNOLOGY
TRANSPORTATION
sector

What is a good CAGR:

There is no definition for a good CAGR. However, anything between 18%-25% over 5 years is normally considered as a good CAGR. Anything below 12%, it's probably best to choose an alternative investment.


15 Years Average Compound Annual Growth Rate
0.05
0.04
0.03
0.02
0.01
0.00
BASIC INDUSTRIES
CAPITAL GOODS
CONSUMER DURABLES
CONSUMER NON-DURABLES
CONSUMER SERVICES
ENERGY
FINANCE
HEALTH CARE
MISCELLANEOUS
PUBLIC UTILITIES
TECHNOLOGY
TRANSPORTATION
sector

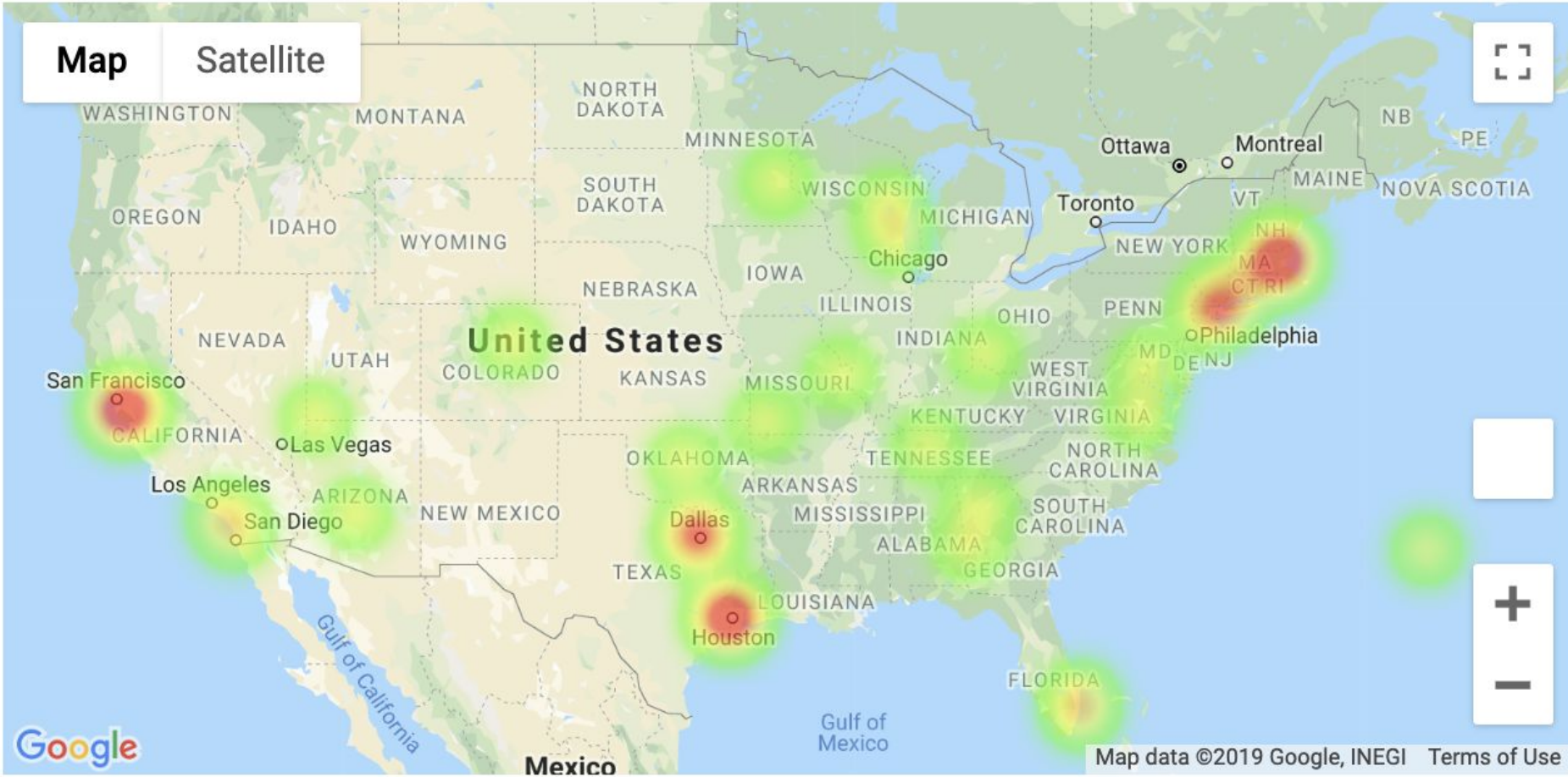

10 Years Average Compound Annual Growth Rate
0.08
0.06
0.04
0.02
0.00
BASIC INDUSTRIES
CAPITAL GOODS
CONSUMER DURABLES
CONSUMER NON-DURABLES
CONSUMER SERVICES
ENERGY
FINANCE
HEALTH CARE
MISCELLANEOUS
PUBLIC UTILITIES
TECHNOLOGY
TRANSPORTATION
sector

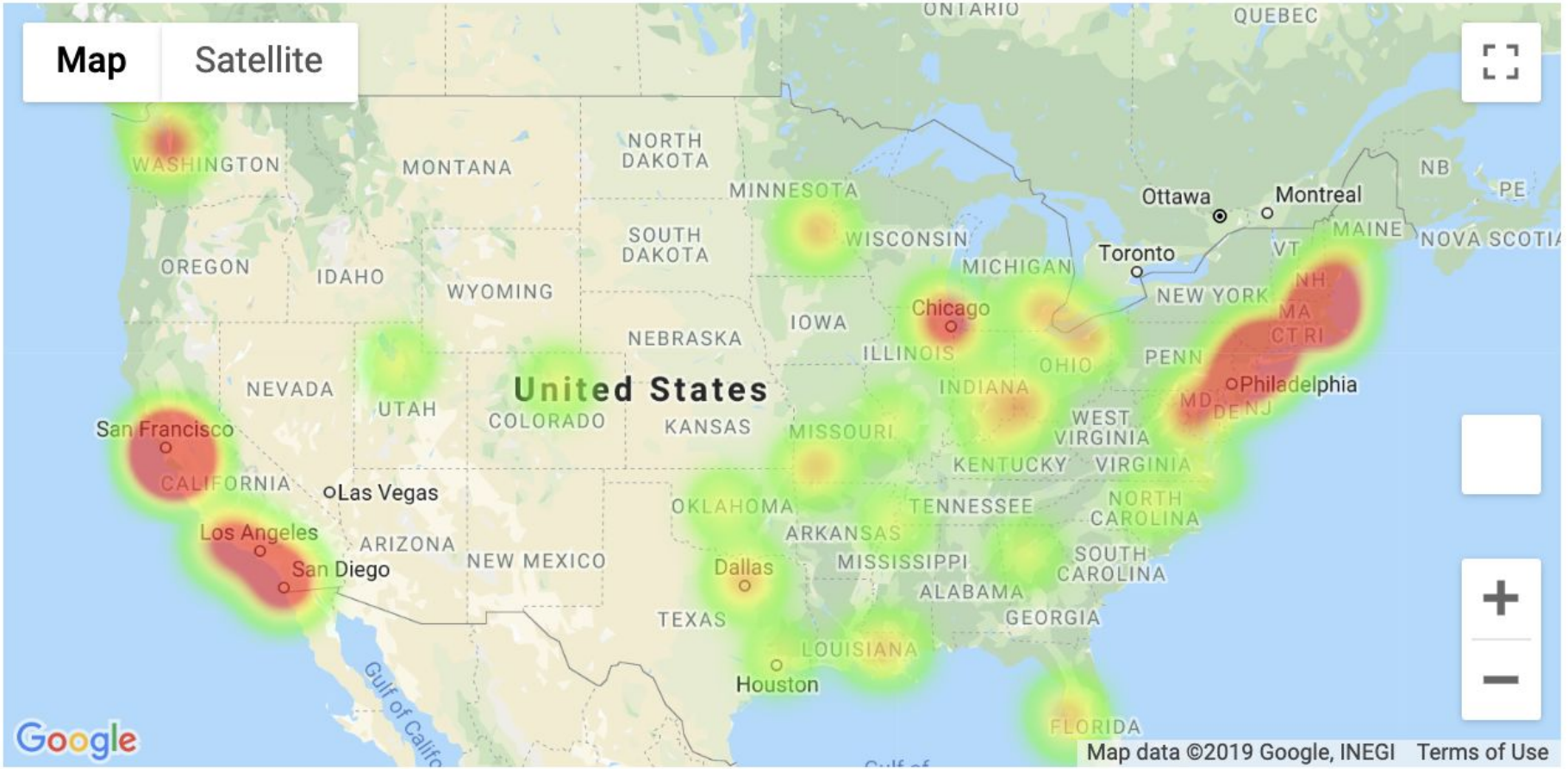

5 Years Average Compound Annual Growth Rate
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
BASIC INDUSTRIES
CAPITAL GOODS
CONSUMER DURABLES
CONSUMER NON-DURABLES
CONSUMER SERVICES
ENERGY
FINANCE
HEALTH CARE
MISCELLANEOUS
PUBLIC UTILITIES
TECHNOLOGY
TRANSPORTATION
sector

Limitation of CAGR:

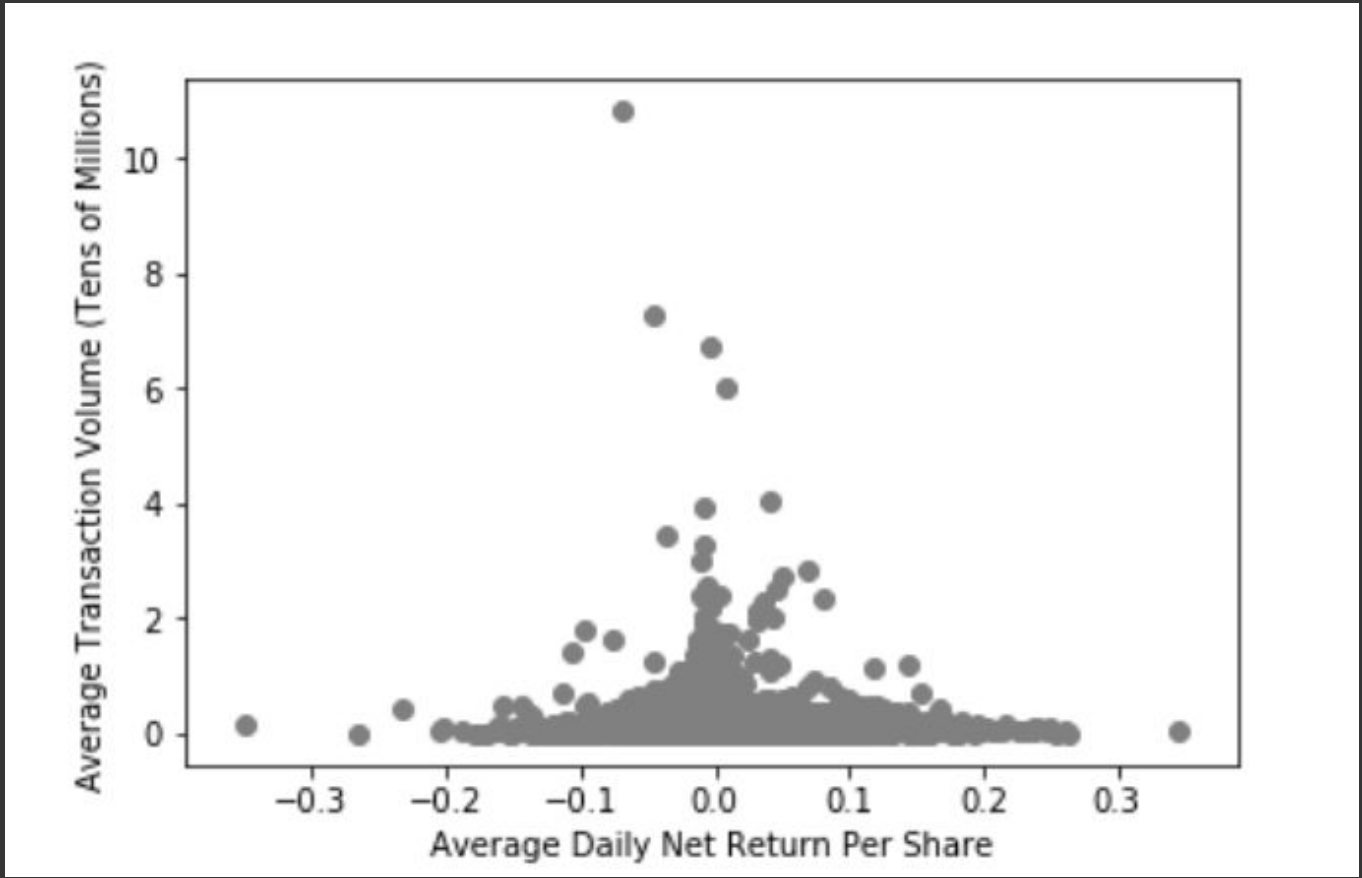
Investment returns are volatile, meaning they can vary significantly from one year to another. However, CAGR does not reflect volatility. CAGR is a pro forma number that provides a "smoothed" annual yield, so it can give the illusion that there is a steady growth rate even when the value of the underlying investment can vary significantly.

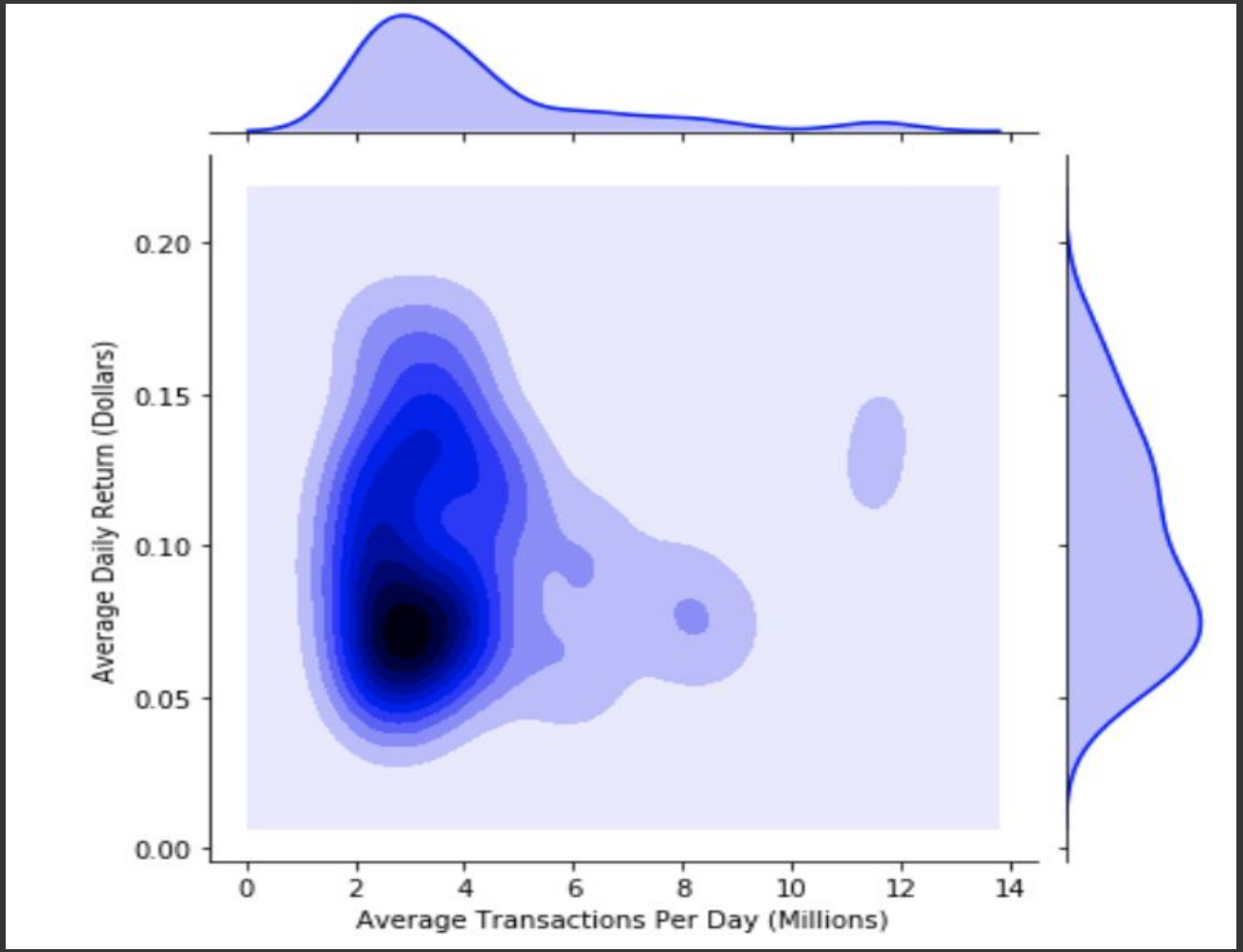
Are specific areas of the US more likely to house growing companies? Has it changed over time?


Map
Satellite
WASHINGTON
MONTANA
NORTH DAKOTA
MINNESOTA
SOUTH DAKOTA
WISCONSIN
MICHIGAN
OREGON
IDAHO
WYOMING
NEBRASKA
IOWA
ILLINOIS
NEVADA
UTAH
COLORADO
KANSAS
MISSOURI
INDIANA
OHIO
PENN
NEW YORK
VT
NH
MA
CT RI
MAINE
NB
PE
NOVA SCOTIA
Ottawa
Montreal
Toronto
Chicago
Philadelphia
MD
DE
NJ
WEST VIRGINIA
KENTUCKY
VIRGINIA
TENNESSEE
NORTH CAROLINA
SOUTH CAROLINA
GEORGIA
ALABAMA
MISSISSIPPI
ARKANSAS
OKLAHOMA
TEXAS
LOUISIANA
FLORIDA
United States
San Francisco
CALIFORNIA
Las Vegas
Los Angeles
ARIZONA
San Diego
NEW MEXICO
Dallas
Houston
Gulf of California
Gulf of Mexico
Mexico
Google
Map data ©2019 Google, INEGI
Terms of Use


Map
Satellite
ONTARIO
QUEBEC
WASHINGTON
MONTANA
NORTH DAKOTA
MINNESOTA
NB
PE
Ottawa
Montreal
SOUTH DAKOTA
WISCONSIN
MAINE
NOVA SCOTIA
OREGON
IDAHO
WYOMING
MICHIGAN
Toronto
VT
NH
NEW YORK
MA
CT RI
IOWA
Chicago
NEBRASKA
ILLINOIS
OHIO
PENN
NEVADA
UTAH
United States
COLORADO
KANSAS
INDIANA
Philadelphia
MD
DE NJ
San Francisco
MISSOURI
WEST VIRGINIA
KENTUCKY
VIRGINIA
CALIFORNIA
Las Vegas
OKLAHOMA
TENNESSEE
NORTH CAROLINA
ARKANSAS
Los Angeles
ARIZONA
San Diego
NEW MEXICO
Dallas
MISSISSIPPI
SOUTH CAROLINA
ALABAMA
TEXAS
GEORGIA
LOUISIANA
Houston
Gulf of Calif
FLORIDA
Google
Map data ©2019 Google, INEGI
Terms of Use

Is there a correlation between stock volume and return?

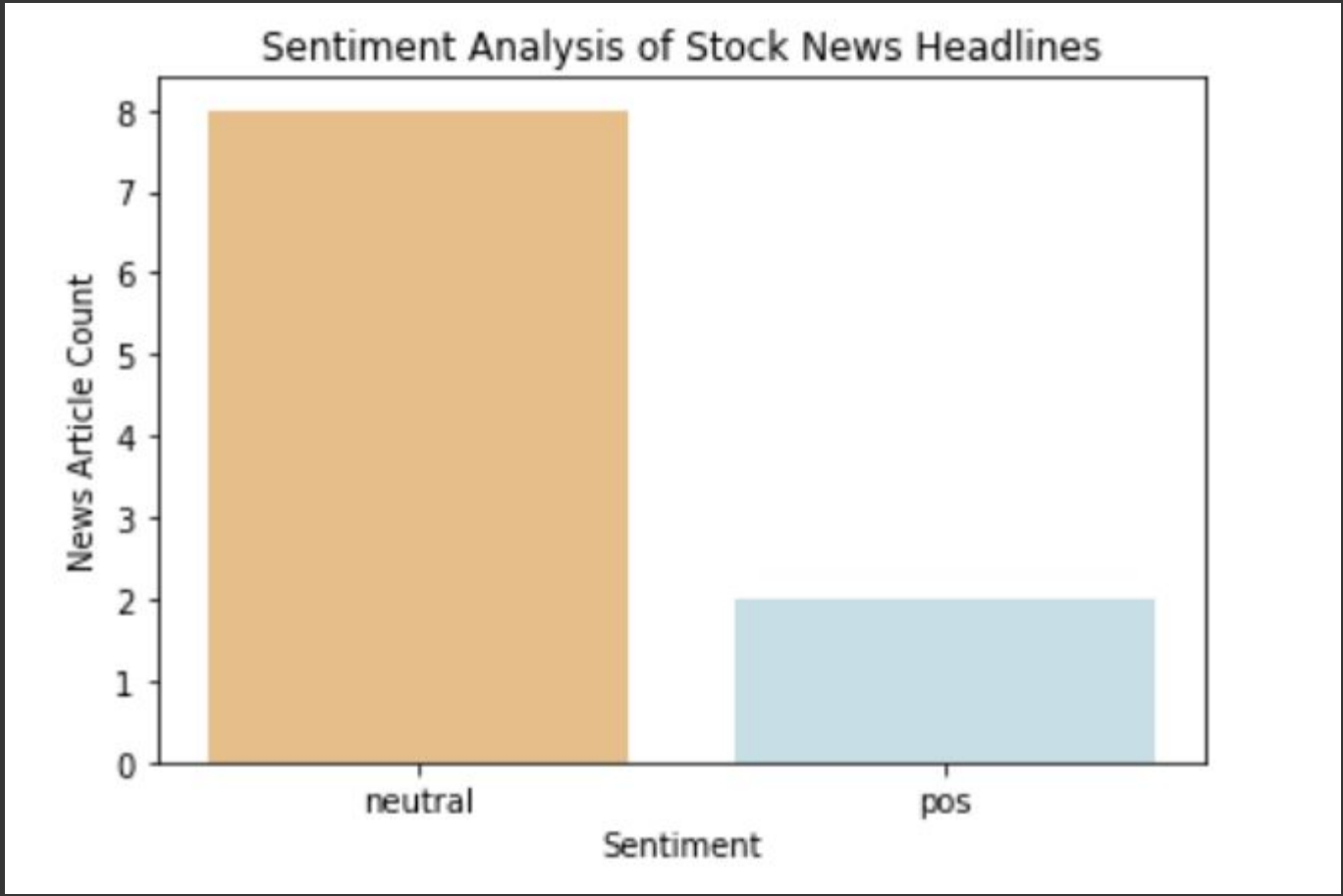

Average Transaction Volume (Tens of Millions)
10
8
6
4
2
0
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
Average Daily Net Return Per Share


0.20
0.15
0.10
0.05
0.00
Average Daily Return (Dollars)
0
2
4
6
8
10
12
14
Average Transactions Per Day (Millions)

API's: The new financial advisor

```
news_url = "https://apidojo-yahoo-finance-v1.p.rapidapi.com/stock/get-news"
user_stock = input('Please enter a US stock ticker: ')
news_querystring = {"region":"US","category":user_stock}
news_headers = {
    'x-rapidapi-host': "apidojo-yahoo-finance-v1.p.rapidapi.com",
    'x-rapidapi-key': rapid_api_key
    }
response = requests.get(news_url, headers=news_headers, params=news_querystring).json()

text_url = 'http://text-processing.com/api/sentiment/'

news_response = response['items']['result']

stock_classifications = []

for news in news_response:
    title = news['title']
    sentiment_data = {'text':title}
    sentiment_response = requests.post(url=text_url, data=sentiment_data).json()
    stock_classifications.append(sentiment_response['label'])

sns.countplot(x=stock_classifications, palette='RdYlBu')
plt.title('Sentiment Analysis of Stock News Headlines')
plt.xlabel('Sentiment')
plt.ylabel('News Article Count')
plt.savefig('Sentiment Bar Chart for Given Stock')
```

```
Please enter a US stock ticker: AMZN
```


Sentiment Analysis of Stock News Headlines
News Article Count
0
1
2
3
4
5
6
7
8
neutral
pos
Sentiment