STOCK SENTIMENT ANALYSIS

Lakhyani Varun

The Stock Sentiment Analysis project aims to leverage natural language processing (NLP) and machine learning techniques to analyze and predict the sentiment of stock-related news and social media content. By extracting sentiment from various textual sources, the project seeks to provide insights into the trends and return of the company stocks

Overview:

- Project flow:
  - ☐ Collecting the data

- downloaded all the important libraries and imported them
- find out various sites to collect news headlines and data which can affect stock price of the target company
- using web scraping with the help of beautiful soup and textblob and other libraries the data of news tags are collected

  - ☐ Cleansing

- The collected is then cleansed, which includes removing unnecessary colons and slashes from the data so the data is easily readable for the model
- using basic python libraries also available in vadar the data can be cleansed and used further

  - ☐ Sentiment score

- for calculation of the sentiment score we can use the vadar sentiment

VADER sentiment analysis uses natural language processing (NLP) techniques to calculate sentiment scores by:

1. **Tokenization:** Breaking down the text into individual words or tokens.
2. **Sentiment Lexicon:** Using a pre-built lexicon that assigns sentiment scores to words based on their semantic meaning. This lexicon includes words, emoticons, and slang commonly found in social media.
3. **Contextual Analysis:** Considering the placement of words in the sentence and applying rules to handle intensifiers, negations, and other modifiers that affect sentiment.
4. **Score Aggregation:** Combining the individual sentiment scores of words in the text to generate an overall sentiment score.
5. **Normalization:** Adjusting the sentiment score based on punctuation (like exclamation marks) and capitalization (to indicate intensity).

By integrating these techniques, VADER provides a quick and effective way to analyze sentiment in text, particularly for social media content where expressions can be informal and context-dependent.

- ☐ Best returns

- the sentiment score from the model is now used in the best possible way for calculating the returns of that company stock
- multiple graphs using matplot are set for the sentiment score and returns and the graph trend is used for making trading stratergies
- sharpe ratio, win ratio , maximum drawdown and number of trades were calculated

For comparing the sentiments we used three companies:-

- ❖ NETFLIX
- ❖ DISNEY
- ❖ AMAZON

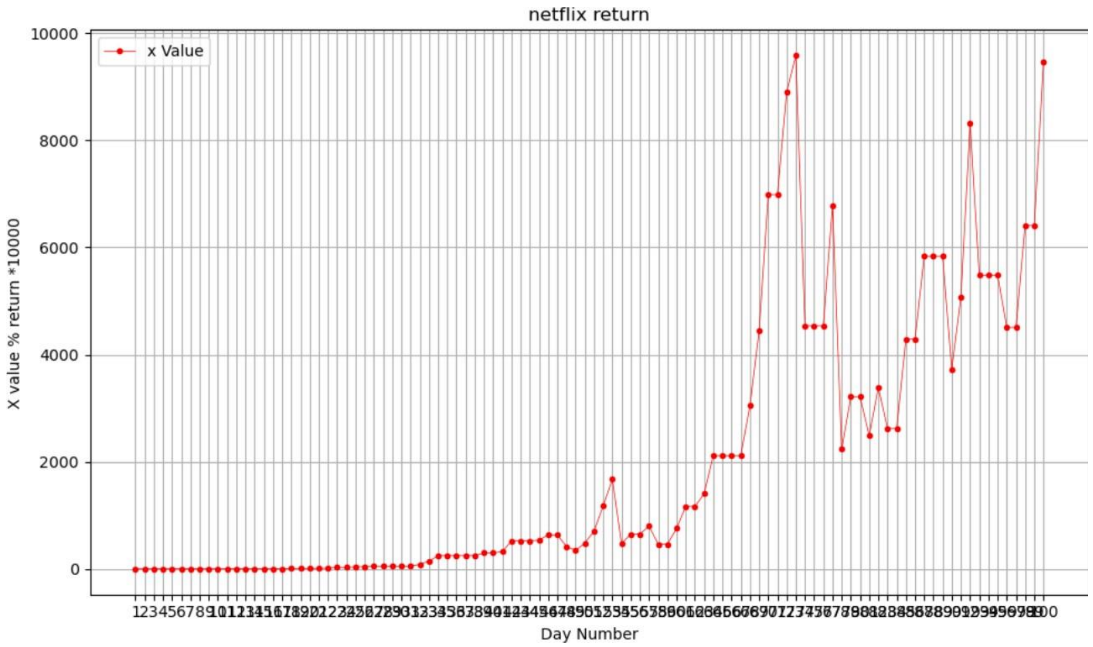
The data of the particular company was taken from sites like finviz and new york times

```
from urllib.request import urlopen, Request
from bs4 import BeautifulSoup
import nltk
from nltk.sentiment.vader import SentimentIntensityAnalyzer
import pandas as pd
import matplotlib.pyplot as plt
import yfinance as yf
import pandas as pd
import requests
from textblob import TextBlob



finviz_url = 'https://finviz.com/quote.ashx?t='
nyt = 'https://www.nytimes.com/international/'
tickers = ['NFLX']


news_tables = {}
```

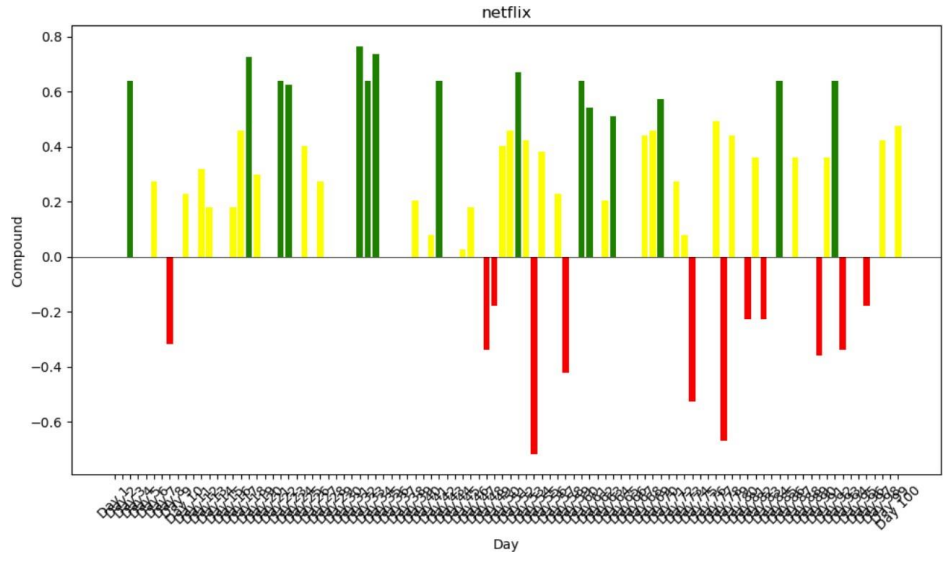
The scraped data was further cleaned and inserted in vadar sentiment model to get the sentiment score of newslines, we used the dataset for previous 100 days and calculated sentiment scores for it

	ticker	title	compound
0	NFLX	Judgement Day for Roaring Kitty? 3 Meme Stocks...	0.0000
1	NFLX	Netflix to open retail centers in Texas and Pe...	0.0000
2	NFLX	Is Netflix Inc (NASDAQ:NFLX) the Best Underval...	0.6369
3	NFLX	Investors Heavily Search Netflix, Inc. (NFLX):...	0.0000
4	NFLX	Analyst reboots Trade Desk stock price target ...	0.0000
..	...	...	...
95	NFLX	3 Up-and-Coming Media Stocks Set to Shake Up t...	-0.1779
96	NFLX	Is It Too Late to Buy Netflix Stock?	0.0000
97	NFLX	Google And Eli Lilly Have Things In Common Bes...	0.4215
98	NFLX	Netflix, Novo Nordisk And These IPOs Fuel Smor...	0.0000
99	NFLX	Analysts Say These 3 Stocks Are the Biggest Wi...	0.4767

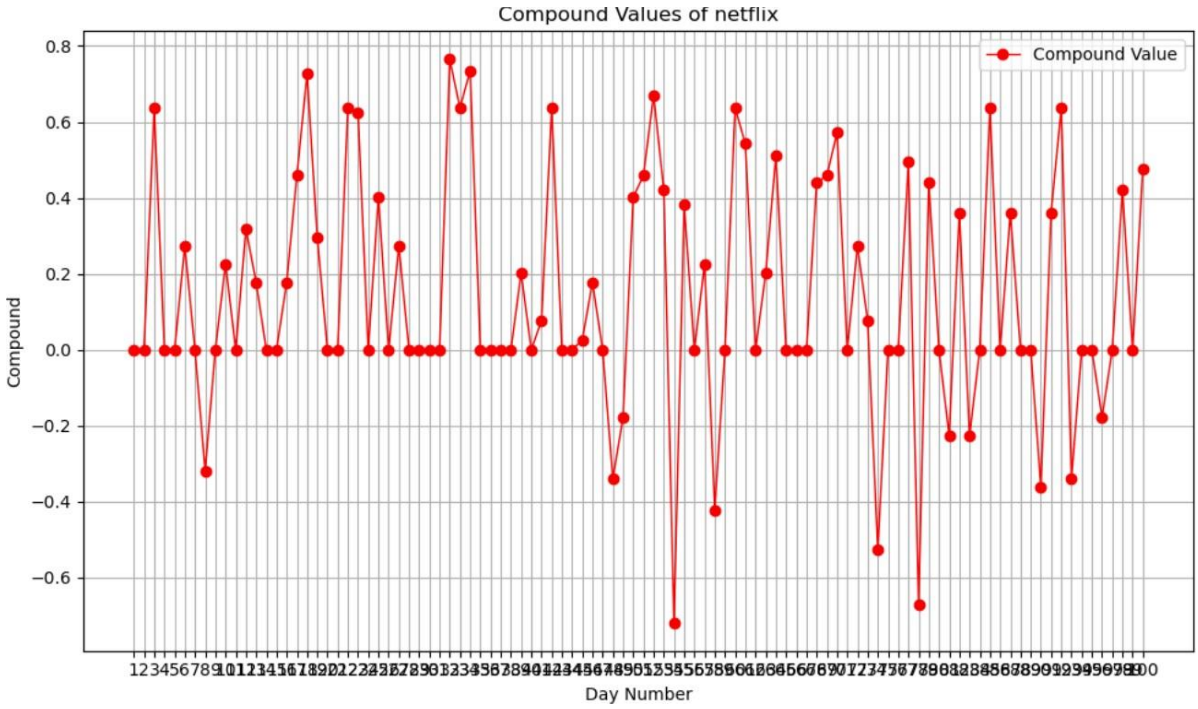
The compound score is the sentiment score of the cleaned news title



This shows the return value % of the netflix stock price for 100 consecutive days
The sentiments of the news line were used to calculate the % return values



This is the graph showing sentiment values or the compound values of netflix we got from vadar sentiment



Now we used these data to calculate sharpe ratio:
Sharpe ratio is calculated ratio of mean and std deviation of daily returns times root of 252

We calculated the sharpe ratio as 0.6828057727442582
There were total 100 trades possible throughout this time:
buy: 12
sell: 42
hold: 46
With these cases respectively

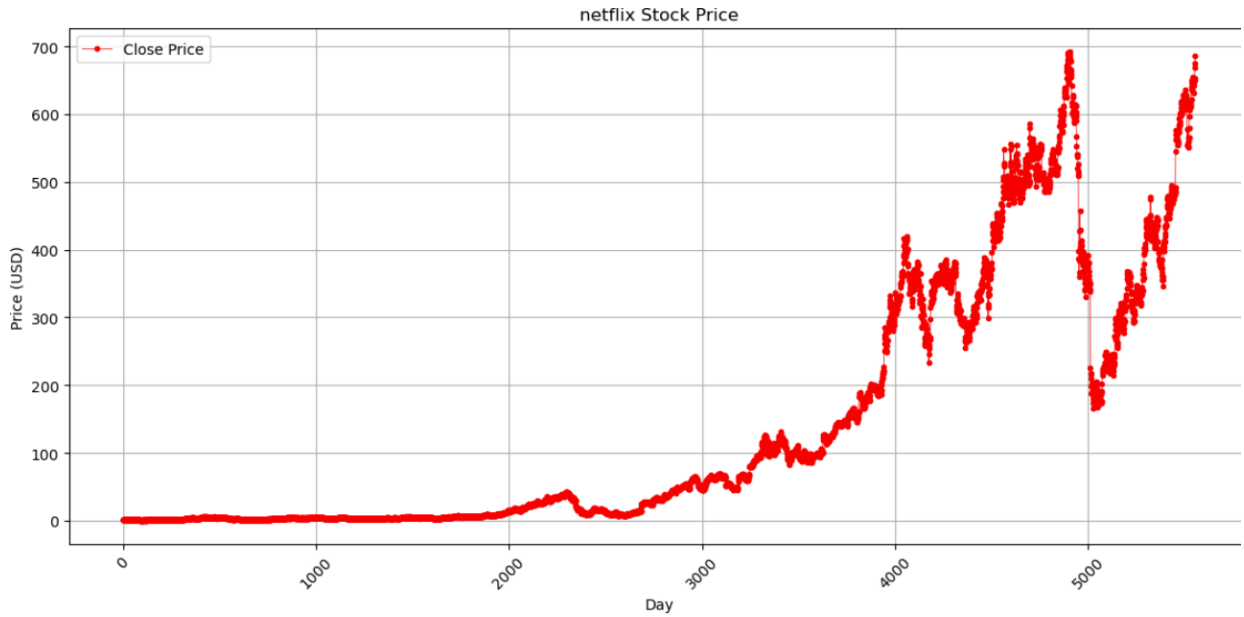
So we calculated win ration which is total number of trades with profit upon total trades
So we got it as win ratio 37.5%
return % most efficient way for past 100 days= 1.4445500000000002 %

returns 1.44% in 100 days for this company

drawdown is: min -7.18% and max 7.6%

profitable trades 45.0
Non profitable/loss trades 12.0

The below graph shows the data for change in stock return of netflix for past 10 years calculated from data collected through scraping showing usd price

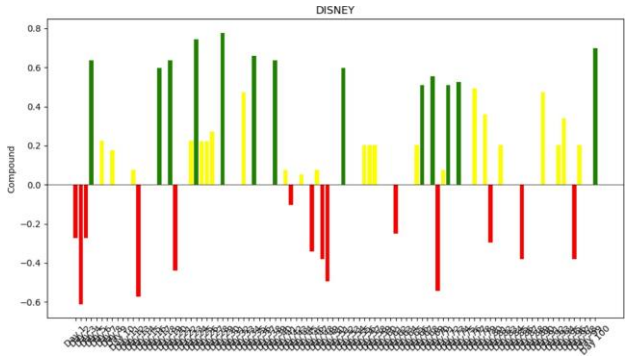


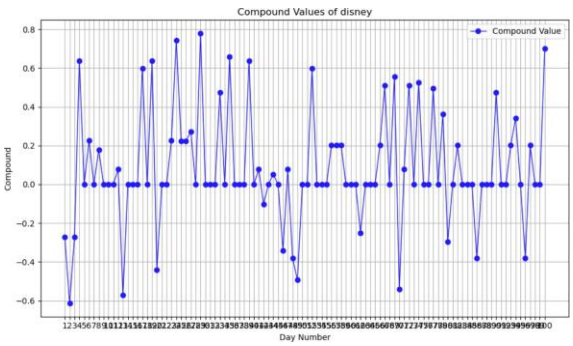
For comparing the model we similarly calculated all these graphs and values for disney and amazon also

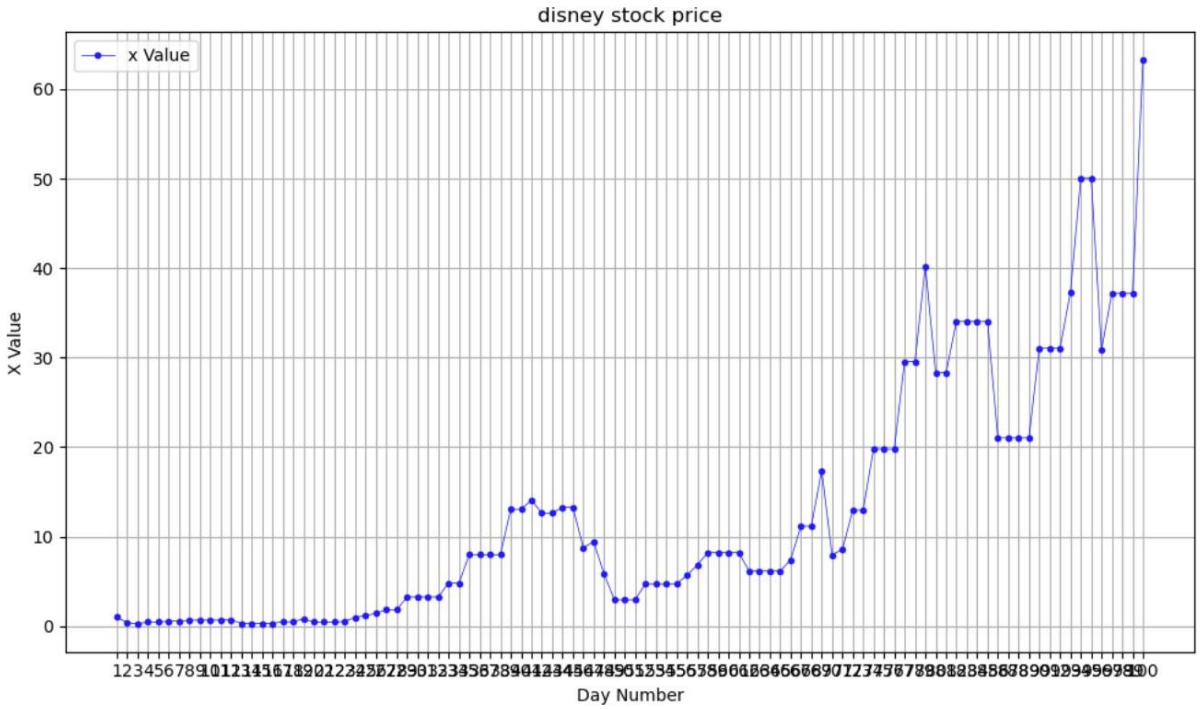
The graphs and values are mentioned below

Disney

	ticker	title	compound	\
0	DIS	These employees moved across the country for w...	-0.2732	
1	DIS	Guess How Many Parents Go Into Debt For A Disn...	-0.6124	
2	DIS	Disney told L.A. residents to move to Florida ...	-0.2732	
3	DIS	Is Netflix Inc (NASDAQ:NFLX) the Best Underval...	0.6369	
4	DIS	Is Disney Animation Back on Track?	0.0000	
..	...	...	...	
95	DIS	Dividend Duds: 3 Stocks to Dump Before They Dr...	-0.3818	
96	DIS	Here's Why Walt Disney Co (NYSE:DIS) is a Top ...	0.2023	
97	DIS	The Walt Disney Company's (DIS) Operating Inco...	0.0000	
98	DIS	Jim Cramer's 10 Stock Picks Heading Into June	0.0000	
99	DIS	3 Exceptional Growth Stocks to Buy Hand Over F...	0.7003	







The graph shows the sentiment scores as compound of the given company And x value represents the daily change or return in stock price of disney co

Sharpe Ratio: 0.427914107085862

Drawdowns:

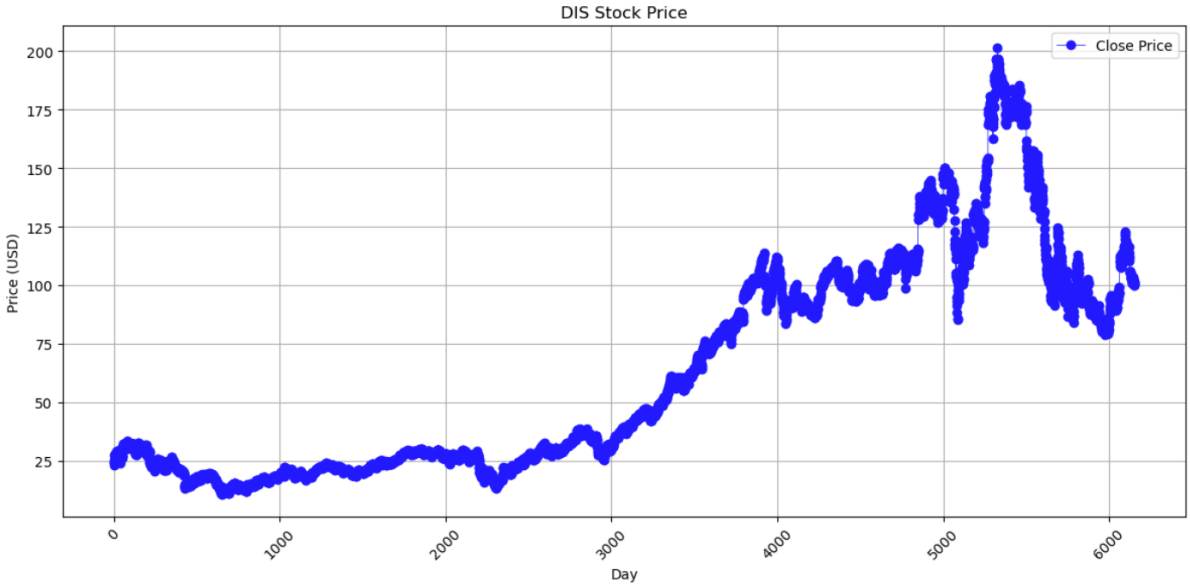
max: 0.7783 (7.7% max change)
min: -0.6124 (-6.1% max drop)

Out of 100 cases
buy: 14
sell: 31
hold: 55

positive profit trades 36.0
negative trades 13.0

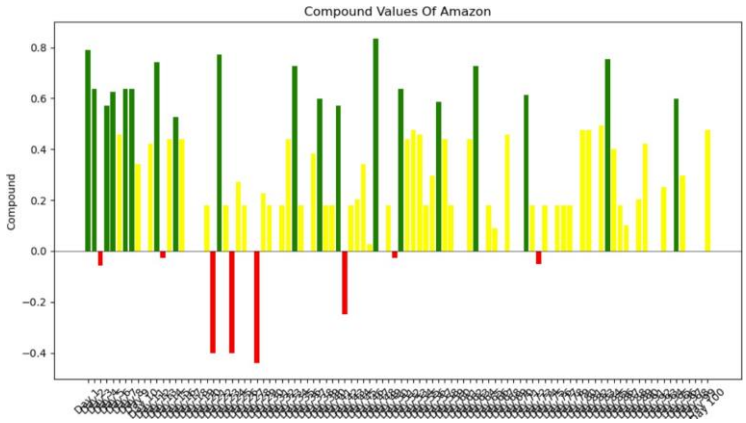
win ratio 2.769230769230769 (27%)
return % after 100 days= 0.8025600000000002 %

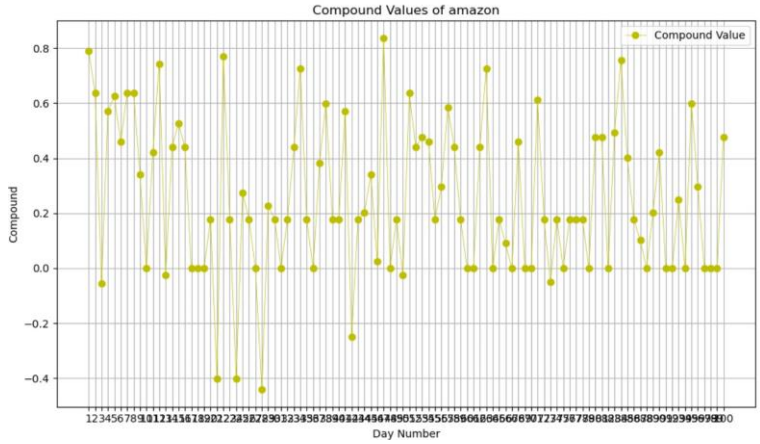
The below graph shows the stock value of disney calculated over past 10 years data for change in stock calculated for reference only

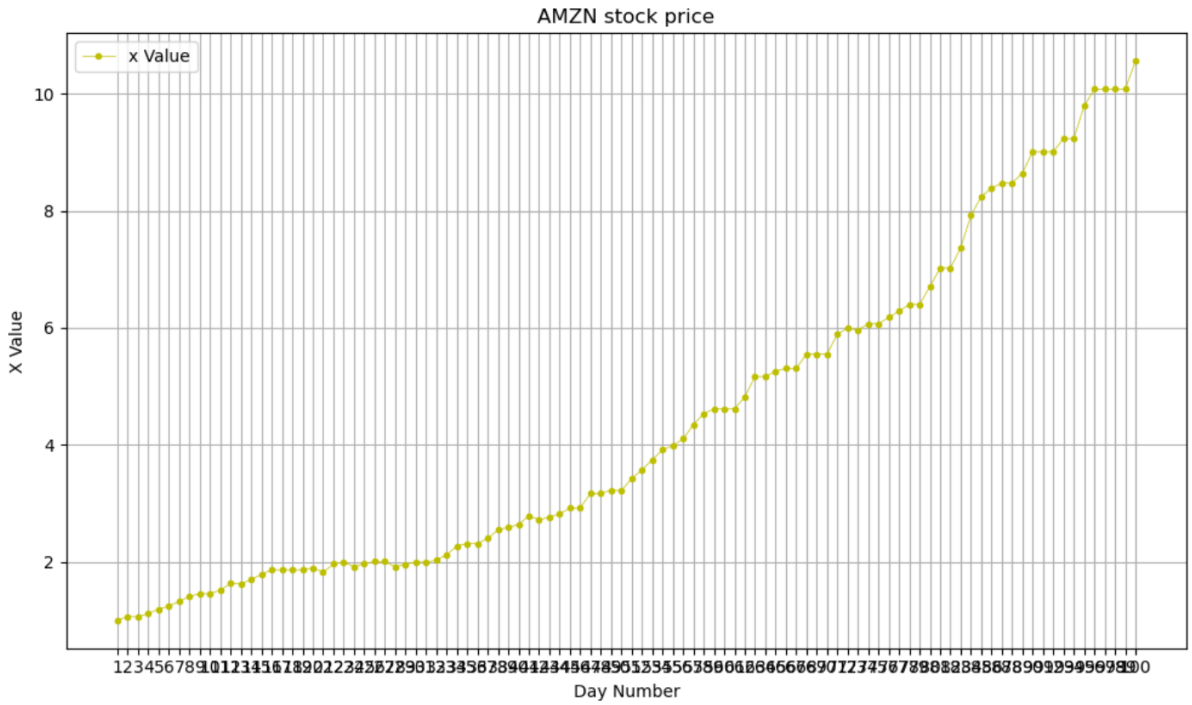


AMAZON:

	ticker	title	compound
0	AMZN	Rag to Riches: 3 Autonomous Driving Stocks Tha...	0.7906
1	AMZN	Shopifys gains with big retailers are setting ...	0.6369
2	AMZN	3 Stocks You Don't Want to Be Caught Holding W...	-0.0572
3	AMZN	Amazon (AMZN) Strengthens Blink Portfolio With...	0.5719
4	AMZN	Why Amazon (AMZN) is a Top Growth Stock for th...	0.6249
..	...	...	...
95	AMZN	Even If Stock-Split Stock Nvidia Retains Its C...	0.2960
96	AMZN	Is It Too Late to Buy Apple Stock?	0.0000
97	AMZN	2 Soaring Stocks to Hold for the Next 20 Years	0.0000
98	AMZN	Morningstar | A Weekly Summary of Stock Ideas ...	0.0000
99	AMZN	2 Artificial Intelligence (AI) Stocks That Cou...	0.4767


Compound Values Of Amazon
Compound
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4


Compound Values of amazon
Compound Value
Compound
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
Day Number


AMZN stock price
x Value
X Value
10
8
6
4
2
Day Number

The graph shows the sentiment scores as compound of the given company
And x value represents the daily change or return in stock price of amazon

Sharpe Ratio: 1.1101107350029498

Higher sharpe ratio shows better increment

Drawdowns:

max: 0.836 (8.3%)
min: -0.4404 (-4.4%)

Buy sell conditions

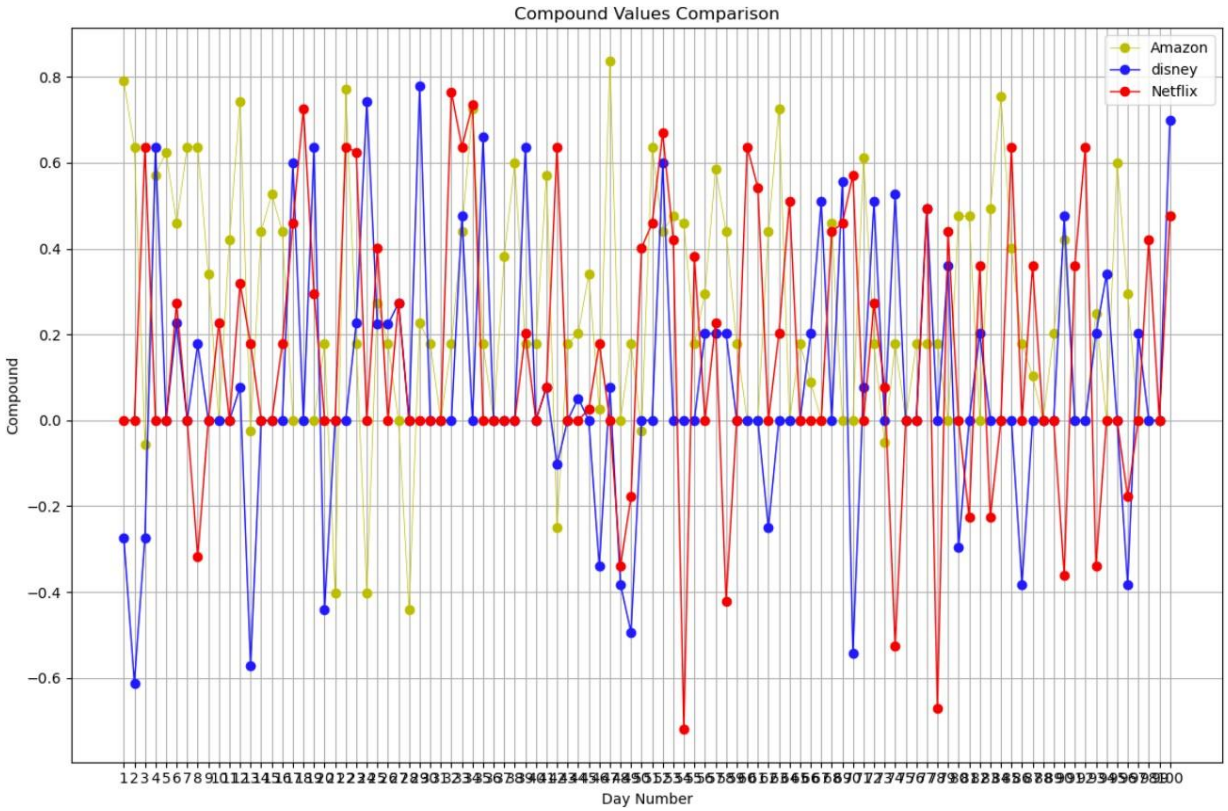
buy: 4
sell: 66
hold: 30

Positive profit trades 67.0 win ratio 8.37(83%)
Negative profit trades 8.0

return % after 100 days= 2.5002999999999997 %

The below graph shows stock price growth of amazon calculated as per data scraped from past 10 years

Now for comparing the sentiment values of all three companies together we drew a graph combined for all the three companies showing sentiment values of all three companies over past 100 days



The graph comparison and above calculations of various datas the amazon return is a better amongst the three as showing least negative sentiment