Sentiment Analysis and Time Series of Amazon and Walmart

Roy Amador, Eric Arzoumanians, Sherleen Lee, Keyu Chen

Group 5

December 5, 2021

Leading e-commerce site in most countries

Fastest growing e-commerce site

World’s largest brick-and-mortar retailer

Introduction

- Using Sentiment Analysis, we predict the sentiments based on news articles about the two retailers.
  - VADER lexicon
  - Data: Finviz

- Using Time Series Analysis, we predict the stock price of the two retailers in four weeks.
  - Holt Winters exponential smoothing
  - Data: Yahoo Finance

Background and Research

- “Addressing corporate social responsibility in corporations: a content analysis of Amazon's and Walmart's websites”
- “Pandemic Profiteers”
- “E-Commerce and COVID-19 Pandemic in India”
- “E-commerce in the pandemic and beyond: online appendix”
- “The Impact of COVID-19 Pandemic on Amazon’s Business”

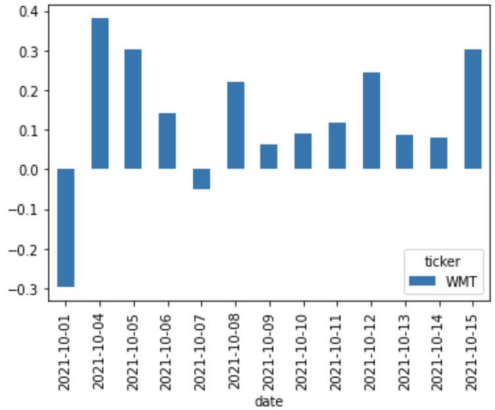
Methodology and Dataset for Sentiment Analysis

Valence Aware Dictionary and Sentiment Reasoner (VADER):

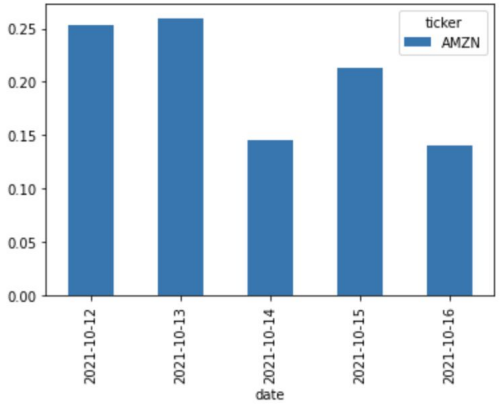
- Rule based analysis tool that is attuned to sentiments expressed in media
- The compound score is computed by summing the valence scores of each word in lexicon and then normalized between -1 and 1
- Then using the vader model applied to the latest article headlines of each company provided on the Finviz platform we retrieved the compounded sentiment score for each company.
- Process:
    - Import title articles using html link
    - Normalize text utilizing regular expression to nullify any unnecessary characters to be used in the analysis
    - Build Dataframe with article titles as a column
    - Use `from nltk.sentiment.vader import SentimentIntensityAnalyzer` library for the computation of the compound sentiment scores

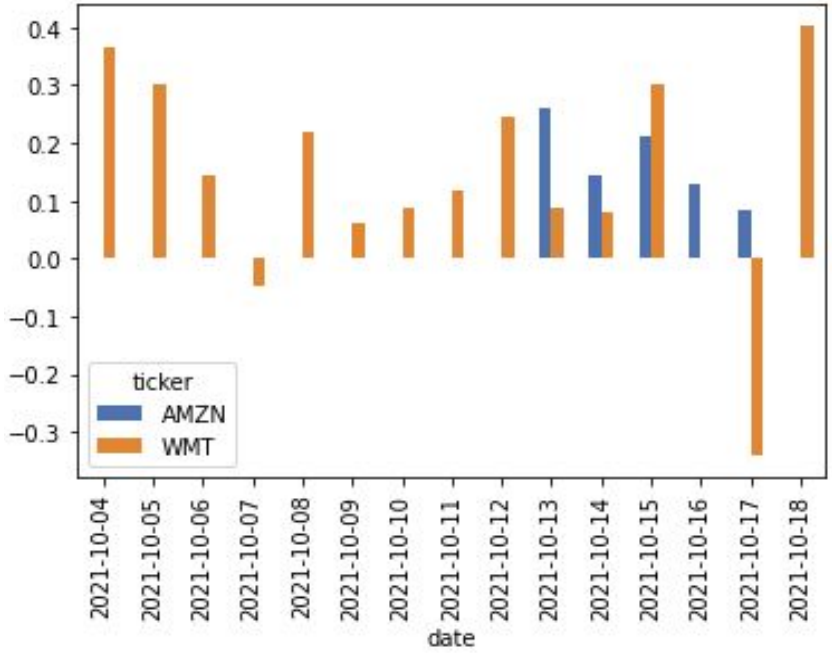
Walmart	Amazon

```
[ ] mean_df.plot(kind='bar')
    plt.show()
```




```
[ ] mean_df.plot(kind='bar')
    plt.show()
```



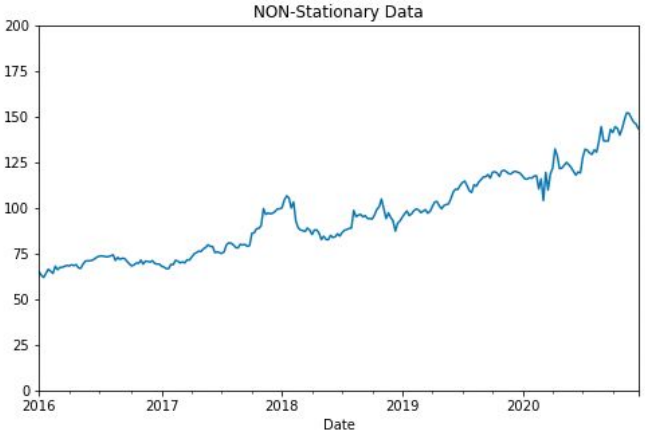

0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
ticker
AMZN
WMT
2021-10-04
2021-10-05
2021-10-06
2021-10-07
2021-10-08
2021-10-09
2021-10-10
2021-10-11
2021-10-12
2021-10-13
2021-10-14
2021-10-15
2021-10-16
2021-10-17
2021-10-18
date

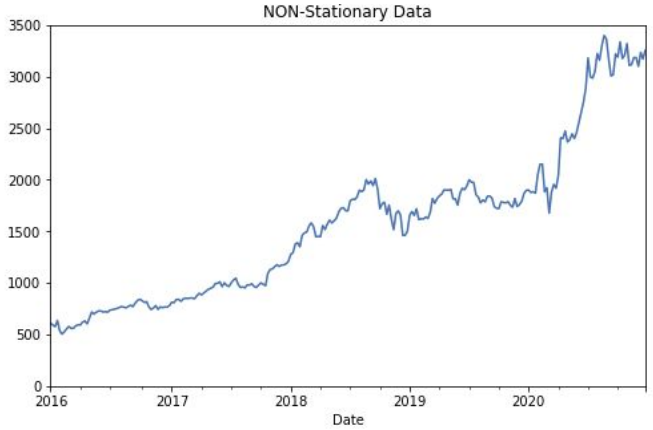
Time Series Analysis

- Data Source:
  - Yahoo Finance
- Time Duration:
  - 01/01/2016 - 01/01/2021
  - Weekly
- Method
  - ARIMA Model

Analysis

- Data Preparation
  - Missing values
  - Duplicate values
  - Stationarity



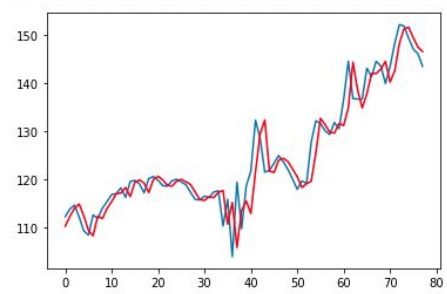


Analysis and Result

- Build up ARIMA model
  - Order: 5,1,0
  - Model for fit
- Fit data in
  - Data Split (0.7,0.3)
  - Prediction dataset
  - Compare
  - Metrics

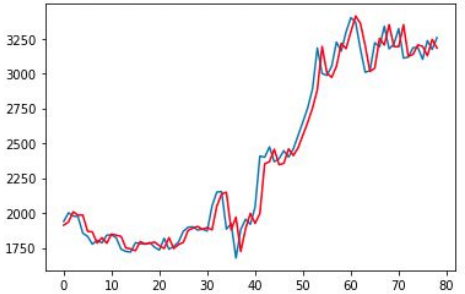
Walmart

```
RMSE:  4.047705662500223
R2:  0.883045736696225
```



Amazon

```
RMSE:  114.2978988070329
R2:  0.9644554229562872
```



Analysis and Result

- Forecast
  - Stock Price in Next 4 weeks

Walmart	
2020-12-25	146.54
2021-01-01	146.68
2021-01-08	146.74
2021-01-15	146.59

Amazon	
2020-12-25	3182.51
2021-01-01	3190.35
2021-01-08	3211.21
2021-01-15	3211.24

Discussion

- News headlines and stock prices are correlated
- Limitations
  - Limited dataset - news only from established sources
  - VADER is attuned to sentiments expressed in social media
  - Limited amount of articles displayed
  - Forecasted for public holidays
- Future Scope
  - Use data from multiple sources - e.g., Twitter
  - Public holidays as exogenous variables

Conclusion

- Most sentiments are positive towards Amazon and Walmart
- Changes during time period created new opportunities while also creating new problems
- Time series models were able to predict future stock price
- Sentiment analysis measured general sentiment towards companies
- Limitation of dataset

References

Weise, K., & Corkery, M. (2021, August 17). People Now Spend More at Amazon Than at Walmart. *The New York Times*. Retrieved October 12, 2021 from https://www.nytimes.com/2021/08/17/technology/amazon-walmart.html

Alfonso, V., Boar , C., Frost , J., Gambacorta, L., & Liu, J. (2021, January 12). *Online appendix - e-commerce in the pandemic and beyond*. BIS. Retrieved October 18, 2021, from https://www.bis.org/publ/bisbull36_appendix.pdf.

Ertem-Eray, T. (2020, October 16). *Addressing corporate social responsibility in corporations: a content analysis of Amazon's and Walmart's websites*. Emerald Insight . Retrieved October 17, 2021, from https://www-emerald-com.proxy.library.cpp.edu/insight/content/doi/10.1108/CCIJ-03-2020-0060/full/html.

Işık, S., İbiş, H., & Gulseven, O. (2021, January 16). *The impact of the COVID-19 pandemic on Amazon's business*. SSRN. Retrieved October 18, 2021, from https://poseidon01.ssrn.com/delivery.php?ID=056112118000093005127098118001091122118082063037061028081121126073113006086004094066101022100000122017012003111108079121093112061072021040019083119124029077067072098060085038118101001014071069071084092096124087084114005083122015076089002122122001016009&EXT=pdf&INDEX=TRUE.

Memon, S., & Kathar, G. N. (n.d.). *E-commerce and covid-19 pandemic in ... - excelpublication.com*. Excel Publication. Retrieved October 17, 2021, from http://excelpublication.com/wp-content/uploads/2021/08/Dr.-Ganesh-C.A-Sohel-Memon.pdf.

PANDEMIC PROFITEERS. Americans for Tax Fairness. (2020, August). Retrieved October 12, 2021, from https://americansfortaxfairness.org/wp-content/uploads/For-the-Many-ATF-CPD-NJ-Pandemic-Profiteers-Brief-August-2020-FINAL.pdf.