



QF624 Machine Learning and Financial Applications Project

Group 2:

Wang Li

Aloysius Voo

Qu Wenqin

Jeriel Wong

Bryan Seah



Business Problem & Objective

- Conventional methods used to predict stock price returns rely heavily on numerical data, and cannot fully utilise all the available information about a company.
- These were noted to be highly dependent on the type of stocks whose returns were being predicted.
 - Cazalet and Roncalli (2014) found that one of the issues with the Fama-French risk factors is the value premium, which is larger for small stocks in general.
 - Cremers and Zitzewitz (2013) found that the Fama-French factors placed a disproportionate weight on small value stocks.
 - Lambert and Hübner (2014) found that this was a result of a theoretical bias in the definition of the premium.



Business Problem & Objective

- To address the limitations of conventional methods, there have been increased studies on the usage of machine learning (ML) models to predict stock returns.
 - Apply linear regression models to predict stock returns
 - Application of deep neural networks, specifically long short-term memory (LSTM) networks for stock return prediction
- Financial markets also generate large amount of unstructured data including news articles, company presentation and more.
- Thus, we seek to determine whether stock price movements can be predicted with natural language processing (NLP) techniques to analyze earning call transcripts/conference presentation to increase alpha from trading strategies.



Data

Population and source

- META, AAPL, GOOGL, NFLX, AMZN
- Period: Mar 2013 to Mar 2023
- Quarterly financial ratios and multiples from CIQ, and US macroeconomic data from Bloomberg
- Quarterly earnings transcripts from CIQ

Pre-processing of data

- Input variables with more than 4 NaN values would be dropped with remaining missing data forward filled
- All variables were normalised

Output variable

- Quarterly returns computed from price data from yfinance



Methodologies

- Linear Regression
- Neural Network
- Natural Language Processing (NLP)



Linear Regression

- Split training-test dataset (80-20)
- All of the input variables were used.
- Separate regressions were trained for each ticker using different degrees of polynomial features (utilising `sklearn.preprocessing.PolynomialFeatures`).
- Using all variables result in the issue of multicollinearity:
 - Potential overfitting
 - Increased difficulty in determining individual impact of each input variable on output variable

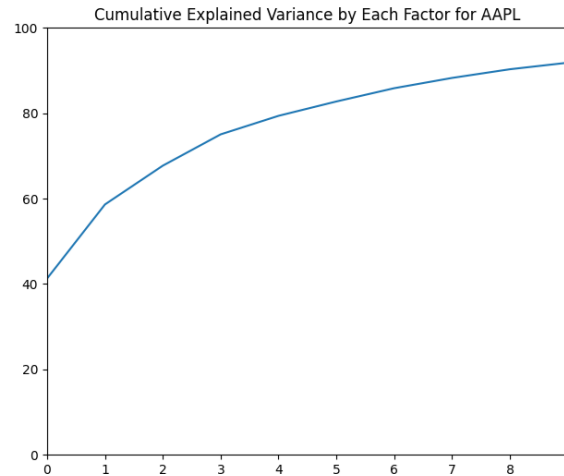
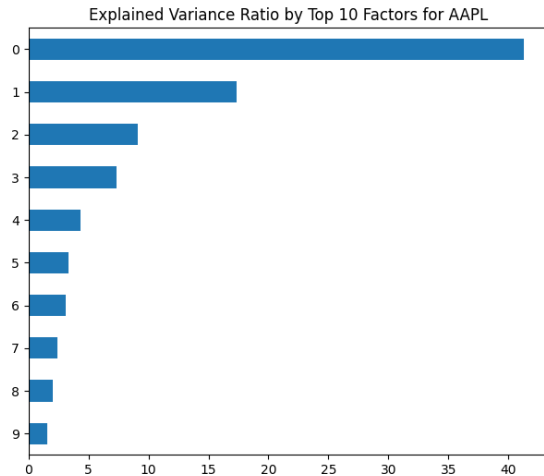


Regularisation

- Both LASSO and ridge regressions were run separately for each ticker using different degrees of polynomial features.
- Penalty would be added to the model's loss function to restrict coefficients, which would lead to zeroising / shrinkage of coefficients.
- Thus, we expect the impact of multicollinearity to be reduced.

Principal Component Regression (PCR)

- Principal component analysis: converts the existing input variables of the training set to linearly uncorrelated variables (i.e. principal components)

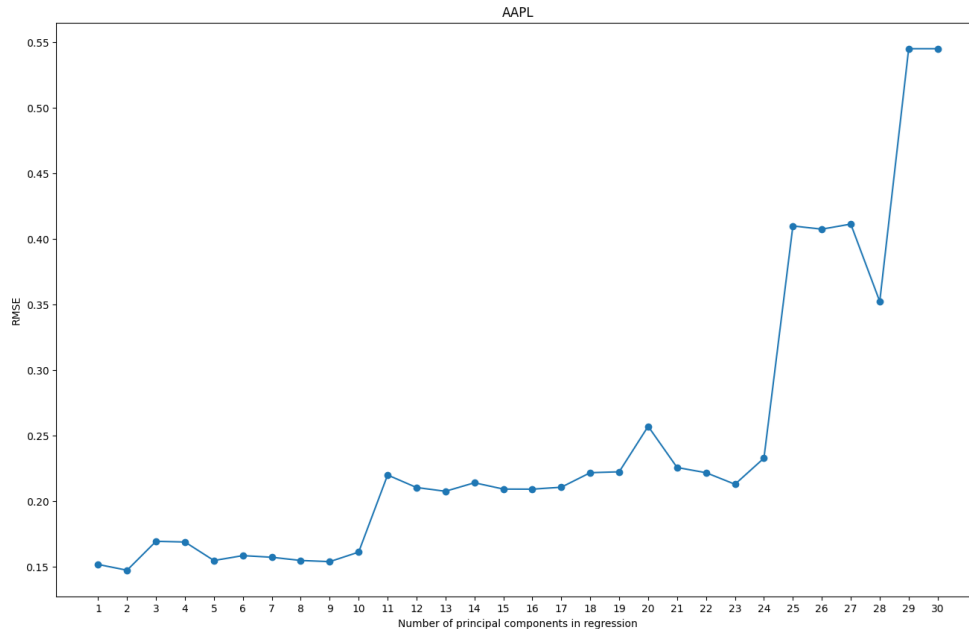




Principal Component Regression (PCR)

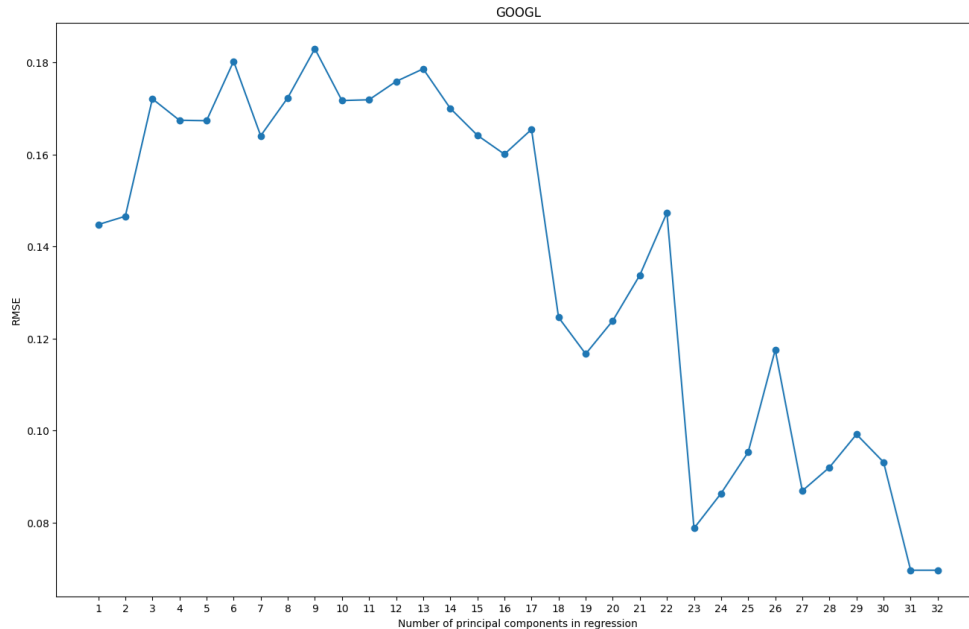
- Resulting principal components cannot be directly interpreted.
- Separate regression models were trained for each ticker by varying the number of PCs to be included in the model.
- This allows us to determine the number of PCs that would give rise to the minimum RMSE.

Principal Component Regression (PCR)



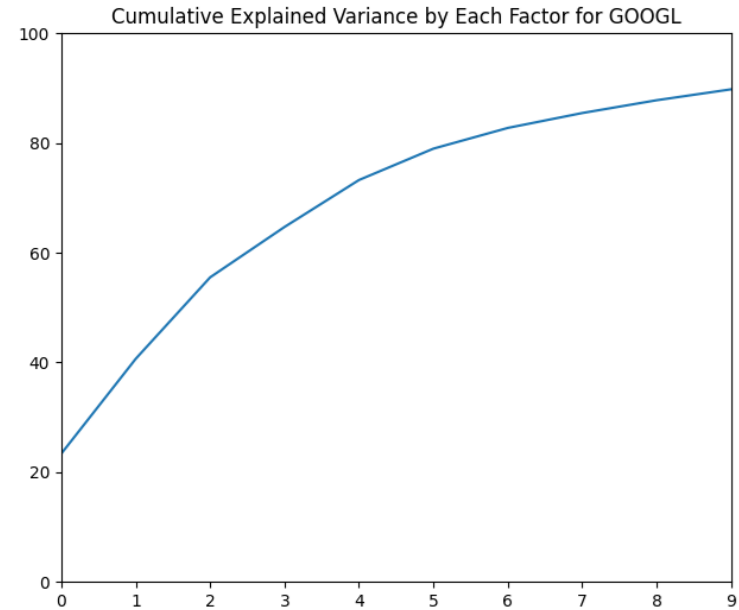
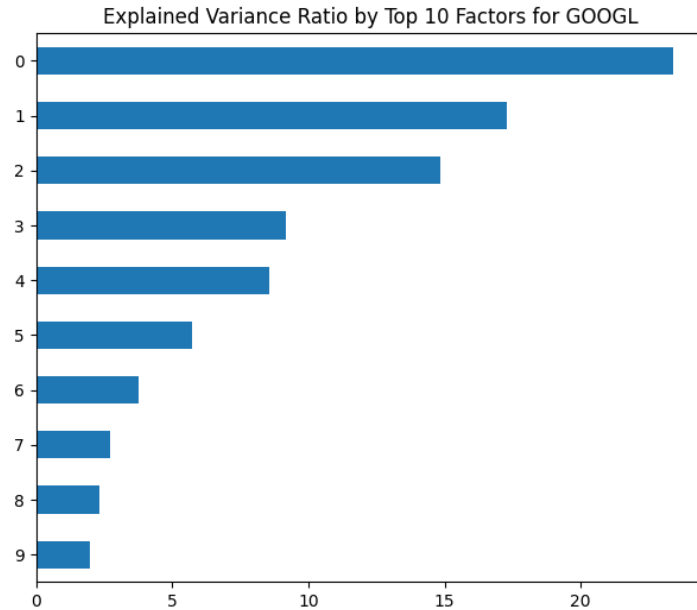
- Min RMSE per testing set for AAPL is generated with 2 component(s)
- Min RMSE per testing set for AAPL: 0.14753

Principal Component Regression (PCR)



- Min RMSE per testing set for GOOGL is generated with 31 component(s)
- Min RMSE per testing set for AAPL: 0.06964

Principal Component Regression (PCR)

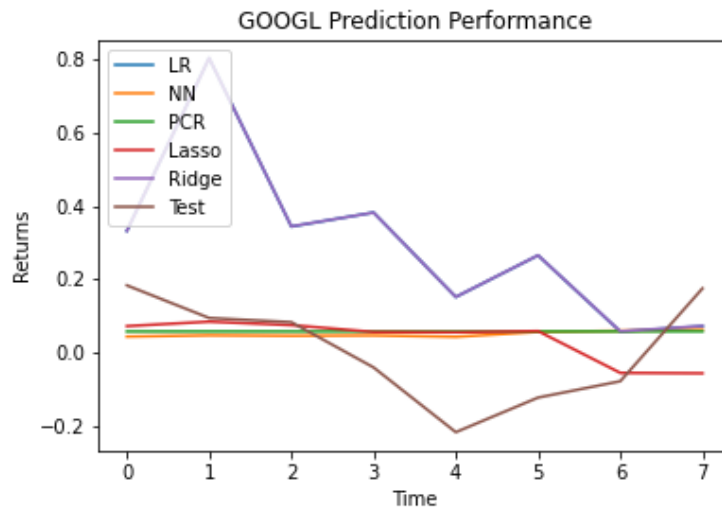
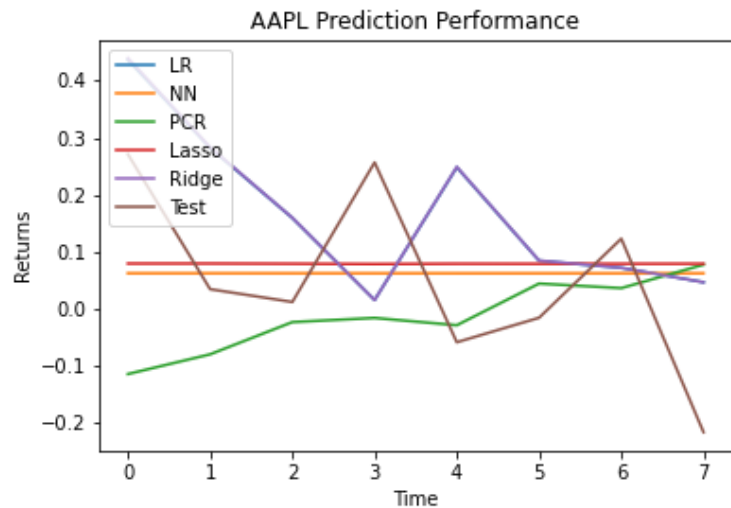




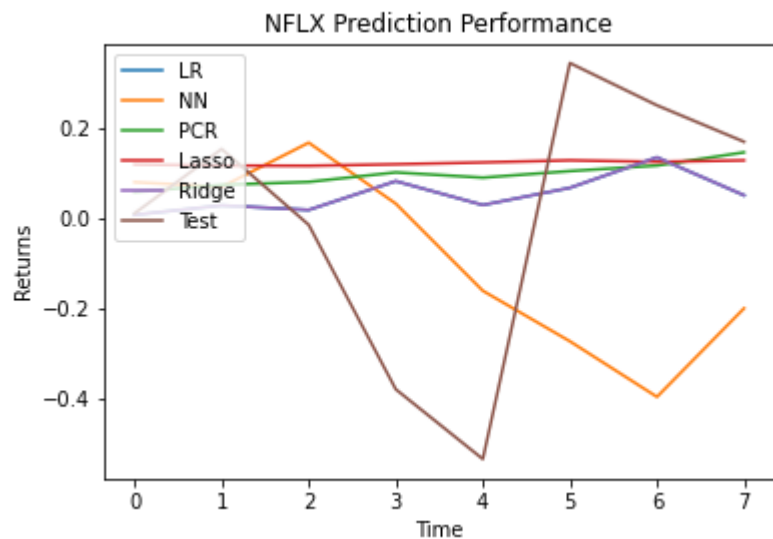
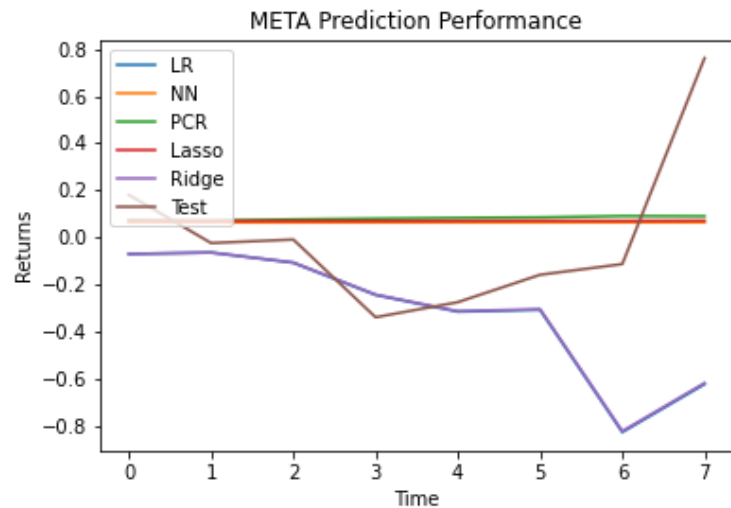
Recurrent Neural Network

- Final Model hyperparameters:
 - two LSTM layers (128 neurons, 64 neurons)
 - Optimiser = 'Adam' with a learning rate of 0.0225
 - Batch_size = 4, epochs = 50
 - Standard Scaler for input data

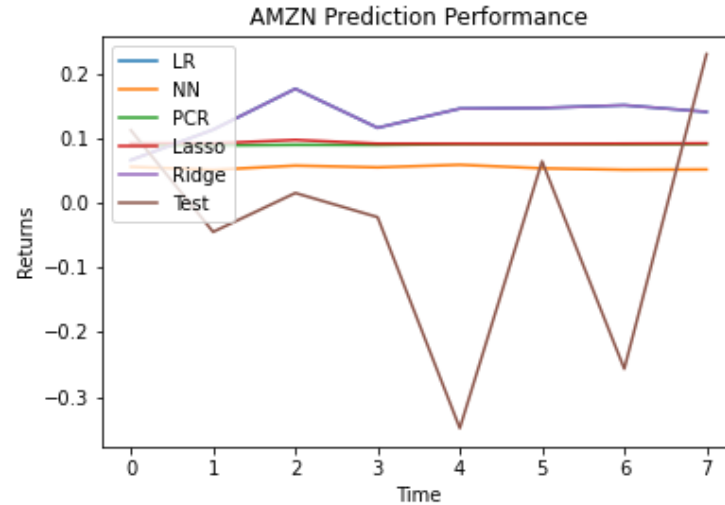
Benchmark Predictions



Benchmark Predictions



Benchmark Predictions



Evaluation

Model Performance - RMSE

	LR RMSE (Test) (1st Degree)	LR RMSE (Test) (2nd Degree)	LR RMSE (Test) (3rd Degree)	Lasso RMSE (Test) (1st Degree)	Lasso RMSE (Test) (2nd Degree)	Lasso RMSE (Test) (3rd Degree)	Ridge RMSE (Test) (1st Degree)	Ridge RMSE (Test) (2nd Degree)	Ridge RMSE (Test) (3rd Degree)	PCR RMSE (Test)	NN RMSE (Test)
META	0.93	0.30	0.56	0.33	0.33	0.33	0.39	0.30	0.56	0.27	0.34
AAPL	0.55	0.20	0.24	0.16	0.16	0.16	0.41	0.20	0.24	0.15	0.15
GOOGL	0.07	0.18	0.37	0.14	0.14	0.14	0.06	0.18	0.37	0.07	0.14
NFLX	0.33	0.38	0.29	0.31	0.31	0.31	0.31	0.38	0.29	0.21	0.30
AMZN	0.28	0.22	0.25	0.21	0.21	0.22	0.28	0.22	0.25	0.20	0.22
Average	0.43	0.26	0.34	0.23	0.23	0.23	0.29	0.26	0.34	0.18	0.23

Model Performance - MAPE

	LR MAPE (Test) (1st Degree)	LR MAPE (Test) (2nd Degree)	LR MAPE (Test) (3rd Degree)	Lasso MAPE (Test) (1st Degree)	Lasso MAPE (Test) (2nd Degree)	Lasso MAPE (Test) (3rd Degree)	Ridge MAPE (Test) (1st Degree)	Ridge MAPE (Test) (2nd Degree)	Ridge MAPE (Test) (3rd Degree)	PCR RMSE (Test)	NN MAPE (Test)
META	13.98	1.93	2.94	2.41	2.41	2.41	4.88	1.94	2.94	1.12	1.70
AAPL	11.55	3.31	2.90	2.38	2.38	2.38	8.56	3.15	2.91	1.46	1.79
GOOGL	0.45	1.22	3.65	1.12	1.12	1.12	0.43	1.21	3.65	0.45	1.12
NFLX	2.98	6.93	0.98	2.82	2.82	2.82	2.94	6.91	0.98	1.22	2.27
AMZN	3.27	1.78	3.19	2.10	2.10	2.18	3.20	1.78	3.19	1.44	2.07
Average	6.45	3.03	2.73	2.17	2.17	2.19	4.00	3.00	2.74	1.14	1.79

Natural Language Processing (NLP)

- Quarterly earning calls transcripts (Presentation and Q&A) -> perform sentiment analysis -> forecast stock price movements
- Company's financial performance, business operations, management's outlook, and address questions from analysts and investors
- Sentiment analysis - understand if the tone of the transcripts is positive, negative, neutral, or contains other emotion value
- The overall sentiment expressed in the call can then be used to help us decide if we should invest in the company's stock

AMAZON.COM, INC. FQ1 2014 EARNINGS CALL | APR 24, 2014

Presentation

Operator

Well, good day, ladies and gentlemen, and welcome to the Amazon.com Q1 2014 Financial Results Teleconference. [Operator Instructions] In addition, today's conference is being recorded. And now for opening remarks, I will turn the conference over to senior manager of Investor Relations, Mr. Dave Fildes. Please go ahead, sir.

Dave Fildes

Hello, and welcome to our Q1 2014 Financial Results Conference Call. Joining us today is Tom Szkutak, our CFO. We will be available for questions after our prepared remarks.

The following discussion and responses to your questions reflect management's views as of today, April 24, 2014, only, and will include forward-looking statements. Actual results may differ materially. Additional information about factors that could potentially impact our financial results is included in today's press release and our filings with the SEC, including our most recent annual report on Form 10-K.

As you listen to today's conference call, we encourage you to have our press releases in front of you, which includes our financial results, as well as metrics and commentary on the quarter.

During this call, we will discuss certain non-GAAP financial measures. In our press release, slides accompanying this webcast and our filings with the SEC, each of which is posted on our IR website, you will find additional disclosures regarding these non-GAAP measures, including reconciliations of these measures with comparable GAAP measures.

Finally, unless otherwise stated, all comparisons in this call will be against our results for the comparable period of 2013.

Now I'll turn the call over to Tom.

Thomas J. Szkutak

Former Chief Financial Officer and Senior Vice President

Thanks, Dave. I'll begin with comments on our first quarter financial results. Trailing 12-month operating cash flow increased 26% to \$5.35 billion. Trailing 12-month free cash flow increased to \$1.49 billion. Trailing 12-month capital expenditures were \$3.85 billion.

We continue to make additional investments in support of business growth, consisting of investments in technology infrastructure, including Amazon Web Services, and additional capacities to support our fulfillment operations. Return on invested capital is 9%, up from 1%. ROIC is 11% free cash flow divided by average total assets, minus current liabilities, excluding the current portion of long-term debt over 5-quarter ends. The combination of common stock and stock-based awards outstanding was 476 million shares compared with 471 million 1 year ago.

Worldwide revenue grew 23% to \$10.74 billion or 23% excluding the \$10 million favorable impact from year-over-year changes in foreign exchange. Media revenue increased to \$5.47 billion, up 8% or 8% excluding foreign exchange. EGM revenue increased to \$13.02 billion, up 27% or 27% excluding foreign exchange. Worldwide EGM increased to 66% of worldwide sales, up from 64%. Worldwide paid unit growth was 23%. Active customer accounts exceeded 244 million. Worldwide active seller accounts were more than 2 million. Seller operating expenses excluding stock-based compensation. Cost of sales was \$14.06 billion or 71.2% of revenue compared with 73.4%. Fulfillment, marketing, tech and content and G&A combined was \$5.18 billion or 26.2% of sales, up approximately 240 basis points year-over-year. Fulfillment was \$2.24 billion or 11.3% of revenue compared with 10.8%. Tech and content was \$1.82 billion or 9.2% of revenue compared with 7.9%. Marketing was \$843 million or 4.3% of revenue compared with 3.8%.

Now I'll discuss operating expenses excluding stock-based compensation. Cost of sales was \$14.06 billion or 71.2% of revenue compared with 73.4%. Fulfillment, marketing, tech and content and G&A combined was \$5.18 billion or 26.2% of sales, up approximately 240 basis points year-over-year. Fulfillment was \$2.24 billion or 11.3% of revenue compared with 10.8%. Tech and content was \$1.82 billion or 9.2% of revenue compared with 7.9%. Marketing was \$843 million or 4.3% of revenue compared with 3.8%.

Copyright © 2014 S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.
spglobal.com/marketintelligence

4

AMAZON.COM, INC. FQ1 2014 EARNINGS CALL | APR 24, 2014

Question and Answer

Operator

[Operator Instructions] We'll hear first from Mark Miller with William Blair.

Mark R. Miller

William Blair & Company L.L.C., Research Division

Tom, I was hoping you could maybe just lay out first what do you think the main reasons are for the acceleration in growth here in the first quarter versus the fourth quarter. And then a specific question on Prime Pantry. Is one of the motivations to get the customer on that program, it might make it easier to convert them to AmazonFresh? And if so, would you want to run this operation as a breakeven like you do with Kindle to enable a bigger business?

Thomas J. Szkutak

Former Chief Financial Officer and Senior Vice President

In terms of growth, we did -- we saw a very solid growth, good growth in Q1, up 23% on a local currency basis. North America, we saw similar growth rates in Q1, as we saw in Q4 of last year, as well as Q1 of last year, all at 26%. Internationally, we saw the growth accelerate a little bit on an exchange-adjusted basis from 15% to 18% from Q4 to Q1. But what we've really seen is -- we're pleased with the overall fundamentals. We continue to add new customers. In stocks continue to be healthy. Third party units, as a percentage of units, are about 40%, so still very strong. FBA adoption continues to be strong around the world. We continue to add new selection. We saw strong growth in many different areas, including Web Services. And so again, a lot of different areas contributed to the growth rate that we saw in Q1. In terms of Prime Pantry, we just think it's an exciting option for Prime members. It's available only to Prime members, so they can get their everyday non-bulk items in one box, and we think that's -- it's interesting for customers. And again, it's a great way that we can add the selection and have the selection for those customers.

Operator

Moving on to Aram Robinson with Wolfe Research.

Aram Robinson

Wolfe Research, LLC

On the famous drone interview that Jeff's had with Charlie Rose, he said something that caught my attention, which was that every elasticity study that you do says that Amazon should be raising prices. I have 2 questions around that. The first is does that equation also work in reverse? Meaning that if higher prices don't dampen demand, does it also mean that lower prices don't stimulate demand like they used to?

Thomas J. Szkutak

Former Chief Financial Officer and Senior Vice President

In terms of pricing, we've been very consistent. We want to offer great value to customers. And so we work very hard to make sure that we can offer and afford to offer great prices for customers. And so that's something that we've been working on very hard over the years. That's certainly one of the reasons why you've seen the growth rates that we've experienced, along with getting closer to customers from a shipping perspective, make sure we have great in-stock and other service attributes. In terms of price increases, we certainly have increased the price on Prime. But again, that was after many years of not raising the price, even though the cost of -- transportation costs certainly had gone up, and the fact that we certainly added a lot of selection, going from a little over 1 million items in the first year to over 20 million items. And so it's still an incredibly great value for customers, and that's why we did that. So again, we're all about making sure we have great values for customers, and we'll continue to do that.

Copyright © 2014 S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.
spglobal.com/marketintelligence

7



Natural Language Processing (NLP)

- There are several open-source NLP packages available that provide powerful tools to perform various NLP tasks, including sentiment analysis. For example, TextBlob, Transformers etc.

Model	TextBlob (based on NLTK and pattern library)	FinBERT (a specialised BERT model in financial domain)
Advantages	<ul style="list-style-type: none">- Simple and easy to use, with a straightforward API- Offers a lexicon-based sentiment analysis (word/phrase-level sentiment scores based on a predefined dictionary. Then aggregate to calculate the overall sentiment score for a sentence)- Does not require extensive computational resources- Provide a good starting point for analysis	<ul style="list-style-type: none">- A State-of-the-art NLP model- Pre-trained on a large financial corpus, taking into account financial jargon and sentiment patterns- Offers domain-specific understanding and can provide better accuracy for financial sentiment analysis compared to general-purpose models
Limitations	<ul style="list-style-type: none">- Relies on predefined rules, which may not capture the nuances and complexities of language- Limited in terms of advanced NLP capabilities and domain-specific understanding- May not provide high accuracy compared to more advanced models such as transformer-based models (e.g., BERT, GPT, RoBERTa)	<ul style="list-style-type: none">- Requires computational resources and processing power- Requires labeled financial data for fine-tuning, which may be challenging to obtain



TextBlob

- TextBlob sentiment analyser can return two properties for a given input
- Polarity
 - refers to the sentiment expressed in a text, whether it is positive, negative, or neutral
 - takes a float score between [-1, +1]. -1 suggests highly negative, +1 suggests highly positive, 0 suggests neutral
- Subjectivity
 - measures how subjective or objective a text is
 - takes a float score between [0, +1]. 0 suggests maximum objectivity and +1 suggests highly subjective

```
1 TextBlob(' ').sentiment.polarity
```

```
0.0
```

```
1 TextBlob('in terms of growth, we did -- we saw a very solid growth, good growth in q1, up 23% on a local currency  
2 basis.').sentiment.polarity
```

```
0.2333333333333333
```

TextBlob

- Data pre-processing (removing duplicate characters, multiple whitespaces, and stopwords)
- Get the average polarity score of each transcript of a particular date to predict the future stock price movement
- Compare to the actual historical stock price movement for the next day and the next quarter
- Limitations: (data noise - Q&A split; model limitation)

File_Date	Ticker	File_Year	File_Quarter	Earning_Year	Earning_Quarter	Filename	Text	TextBlob_Subjectivity	TextBlob_Polarity	TextBlob_Analysis	TextBlob_Movement	Next_Day_Return	Next_Day_Movement	Next_Quarter_Trading_Date	Next_Quarter_Adj_Close	Next_Quarter_Return	Next_Quarter_Movement
2013-07-23	AAPL	2013	3	2013	2	Apple Inc., Q3 2013 Earnings Call, Jul 23, 2013...	apple inc. fq3 2013 earnings call jul 23, 2013...	0.449779	0.185650	Positive	1.0	0.051361	1.0	2013-10-23	16.384935	0.261191	1.0
2013-10-28	AAPL	2013	4	2013	3	Apple Inc., Q4 2013 Earnings Call, Oct 28, 2013...	apple inc. fq4 2013 earnings call oct 28, 2013...	0.460349	0.205227	Positive	1.0	-0.024912	-1.0	2014-01-28	15.901063	-0.038542	-1.0

Ticker	AAPL	AMZN	GOOGL	META	NFLX
Next Day Movement Prediction Accuracy	55.26%	47.37%	60.53%	65.79%	63.16%
Next Quarter Movement Prediction Accuracy	68.42%	60.53%	60.53%	63.16%	65.79%

FinBERT

NETFLIX, INC. FQ2 2018 EARNINGS CALL | JUL 16, 2018

Question and Answer

Todd Michael Juenger

Sanford C. Bernstein & Co., LLC., Research Division

All right. Thanks, Spencer. So let's start with the obvious. So for the first time in, I think, 5 quarters, net additions came in below your own forecast both in the U.S. and internationally. So whoever wants to, maybe you can help us walk through where was the source of that shortfall and what you attribute it to.

Wilnot Reed Hastings

Co-Founder, Chairman, President & CEO

David, you want to hit that?

David B. Wells

Former Chief Financial Officer

Yes, sure, Todd. So in general, I would say acquisition, which is up year-on-year but wasn't up as much as we thought it was going to be. So -- and it was pretty broad across multiple markets. It wasn't any one area of the world. And as you pointed out, after 4 consecutive quarters of under forecasting the business, we over forecasted the business. And we strive for accuracy. We clearly didn't pad the number. But we think based on the rolling 12 months of growth that we've had compared to the prior rolling 12 months of growth, the U.S. up slightly, internationally up significantly, that the background and underlying characteristics of the business haven't changed. Our total addressable market is intact and hasn't really changed based on those 90 days of actuals. And in general, we think that the conversion and growth to Internet-enabled entertainment is intact and people are loving it. People are adopting Netflix around the world increasingly more in our newer markets as well. And so I think we're still on track for a strong growth year this year, and maybe it's going to come in a little bit differently than we expected and others expected.

Wilnot Reed Hastings

Co-Founder, Chairman, President & CEO

And Todd, you know this probably that paid net adds are up compared to year ago and forecast to be up a year-on-year over basis in Q3. And the fundamentals have never been stronger. Our viewing is setting year-over-year records, the shows that we have coming. So we're feeling very strong about the business.

Todd Michael Juenger

Sanford C. Bernstein & Co., LLC., Research Division

Terrific. I'll stay on this for just a couple more follow-ups and then we'll move on to broader things. Just wonder, over the course of the winter, there were some well-publicized essentially global pricing increases. I wonder if you think that had any impact on either retention or gross adds relative to your forecast.

David B. Wells

Former Chief Financial Officer

Oh, we don't think so, Todd. I mean, if anything, we -- all of 2017, we sort of had rolling increases in various different parts of the world. We were able to grow continually through that and we continue to. So I don't think that is contributing to this trend.

Wilnot Reed Hastings

Co-Founder, Chairman, President & CEO

And Todd, we've seen this movie of Q2 shortfall before about 2 years ago in 2016, and we never did find the explanation of that other than there's some lumpiness in the business and continue to perform after that.

- Earnings Call Transcript Question and Answer
- Applied FinBERT on every sentence in Answers for each Question
 - If scope not defined, a lot of white spaces are deemed positive sentiment

```
1270 Sentence: In many ways, the same
1271 way I would read over the years
1272 Predicted Sentiment: positive
1273
1274
1275 Sentence: And I really do believe that this kind of shared leadership model is going
1276
1277 Copyright © 2023 S&P Global Market Intelligence, a division of S&P Global Inc
1278 Predicted Sentiment: positive
1279
1280 Sentence: All Rights reserved
1281 Predicted Sentiment: positive
1282
1283 Sentence:
1284 spglobal
1285 Predicted Sentiment: positive
1286
1287 Sentence: com/marketintelligence
1288
1289 5
1290
1291 NETFLIX, INC
1292 Predicted Sentiment: positive
1293
1294 Sentence: FQ4 2022 EARNINGS CALL | JAN 19, 2023
1295
1296 to help us to move fast to challenge each other, to challenge the company to raise to new heights
1297 Predicted Sentiment: negative
1298
1299 Sentence: And I'm
1300 just incredible what we're able to do
1301 Predicted Sentiment: positive
1302
1303 Sentence:
1304
1305 And to your point, this is the leadership team
1306 Predicted Sentiment: positive
1307
1308 Sentence: It's been pretty stable, and that's why that steady transition
1309 feels so steady
1310 Predicted Sentiment: negative
1311
```



FinBERT - Defining Question and Corresponding Answer 2

', 'Evercore ISI Institutional Equities,', 'Research DivisionMark Alan May', 'Citigroup Inc, Research Division', 'Mark R. Miller',

Call Participants

EXECUTIVES

Sean Boyle

Thomas J. Szkutak

Former Chief Financial Officer and
Senior Vice President

Mark Alan May

Citigroup Inc, Research Division

ANALYSTS

Anthony Joseph DiClemente

Barclays Bank PLC, Research
Division

Mark R. Miller

William Blair & Company L.L.C.,
Research Division

Benjamin Ari Schachter

Macquarie Research

Mark Stephen F. Mahaney

RBC Capital Markets, LLC,
Research Division

Brian J. Pitz

Jefferies LLC, Research Division

Matthew Russell Nemer

Wells Fargo Securities, LLC,
Research Division

Colin Alan Sebastian

Robert W. Baird & Co.
Incorporated, Research Division

Ronald Victor Josey

JMP Securities LLC, Research
Division

Douglas Till Anmuth

JP Morgan Chase & Co, Research
Division

Ross Adam Sandler

Deutsche Bank AG, Research
Division

Heath Patrick Terry

Goldman Sachs Group Inc.,
Research Division

Scott William Devitt

Morgan Stanley, Research Division

Jordan Elliot Rohan

Stifel, Nicolaus & Company,
Incorporated, Research Division

Stephen D. Ju

Crédit Suisse AG, Research Division

Justin Post

BofA Merrill Lynch, Research
Division

Youssef Houssaini Squali

Cantor Fitzgerald & Co., Research
Division

Kenneth Michael Sena

Evercore ISI Institutional Equities,
Research Division

FinBERT - Defining Question and Corresponding Answer 2

Question: And then quickly, is there any reason Prime Instant Video isn't available for Android devices?

Response: Thomas J. Szkutak

Former Chief Financial Officer and Senior Vice President

In terms of media growth, not a lot of callouts, except probably the obvious is digital units growing very fast relative to physical units, and we're excited to see that. And because of where we are, we're further penetrated in North America, you're seeing a bigger impact on our growth rate than you are in International, and we certainly see that, but not a lot of other callouts there. But again, we're very pleased and certainly, customers are responding to many things, including selection and great prices and everything else within those digital offerings. But also, they're responding to unique selection that we have. And we -- if you take a look at our release, you'll see the specific numbers related to some of the exclusives we have and certainly, that's having an impact. So there's many, many different things that are working for us in that space as part of our overall ecosystem for digital that we're pleased with. In terms of our device plan, we're very pleased with the devices we have and to offer customers. We think we have a great offering, both in terms of Kindle and Kindle Fire. And in terms of our future road map, we have a long-standing practice of not talking about what they'll -- what that will be prior to announcement. Operator Our next question comes from Youssef Squali from Cantor Fitzgerald. Youssef Houssaini Squali Cantor Fitzgerald & Co., Research Division 2 quick questions please. Tech and content was up to 9% of revenues. I think that's the highest it's ever been.

Find sentences that ends with ? Till the next ?

Any sentences in between ends with . based on the above criteria

Will be mapped as a corresponding question-response



FinBERT - Interesting Findings

```
-----
Question: How much of that is actually streaming-content related, and how do we look at it going forward?
Response:
-----
Question: Does it stay at that elevated level?
Response:
-----
Question: And then on Fresh, is that business profitable for you in the Seattle area?
Response: Thomas J. Szkutak
Former Chief Financial Officer and Senior Vice President
In terms of tech and content, we're spending in a number of different areas, but there's a few that I'd like to highlight. One is, certainly, keep in mind that the infrastructure related to our very fast-growing web services business is included in tech and content. So certainly, as we ramp up that business and it's becoming more sizable and growing very fast, you're seeing that impacting that line item. We're also investing very heavily in digital, and that's across many different parts of our digital offerings there. The Copyright © 2019 S&P Global Market Intelligence, a division of S&P Global Inc. All Rights reserved. spglobal. -- but keep in mind that Fresh was designed as a pilot. And certainly, the economics have improved over time through invention on behalf of the team there, as well as operating efficiencies. So again, that was set up as a test and which has enabled us to launch L.A. Operator
And we'll now go to Colin Sebastian from Robert Baird. Colin Alan Sebastian
Robert W. Baird & Co. Incorporated, Research Division
I guess a quick follow-up on the device strategy, and given the fairly quick pace of innovation in the tablet market overall, I wonder if you can contrast the benefits of that for Amazon, given the popularity of shopping and media apps, with the costs and complexity of maintaining your own line of hardware. And related to this, it seems as if the pace of new content acquisition and licensing has picked up a bit.
-----
```

- 2 sub-questions in a Question
- Question ends with a .
 - Inability to group such statements as a single question

117 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

118 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

119 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

120 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

121 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

122 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

123 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

124 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

125 And then also, what are the implications on Amazon's long-term margins as you go through these transitions over time?

Any help on that could help. Thomas J. Szkutak Former Chief Financial Officer and Senior Vice President Sure. In terms of the way we're looking at them, certainly, is based on the free cash flow potential, and we have a -- we're in some really interesting great businesses that have a lot of potential from a free cash flow generation standpoint, with good and high ROICs, which is exciting. From a margin standpoint, always challenging to predict where that will come out in terms of absolute numbers. But what we will do is we want to make sure that we try to maximize free cash flow. That's something that we've always said. So our strategy hasn't changed. Our outlook hasn't changed in that regard.

119 e in some really interesting great businesses that have a lot of potential from a free cash flow ge Any help on that could help.

120 e in some really interesting great businesses that have a lot of potential from a free cash flow ge Thomas J.

121 e in some really interesting great businesses that have a lot of potential from a free cash flow ge Szkutak Former Chief Financial Officer and Senior Vice President Sure.

122 e in some really interesting great businesses that have a lot of potential from a free cash flow ge In terms of the way we're looking at them, certainly, is based on the free cash flow potential, and we have a -- we're in some really interesting great businesses that have a lot of potential from a free cash flow generation standpoint, with good and high ROICs, which is exciting.

123 e in some really interesting great businesses that have a lot of potential from a free cash flow ge From a margin standpoint, always challenging to predict where that will come out in terms of absolute numbers.

124 e in some really interesting great businesses that have a lot of potential from a free cash flow ge But what we will do is we want to make sure that we try to maximize free cash flow.

125 e in some really interesting great businesses that have a lot of potential from a free cash flow ge That's something that we've always said.

126 e in some really interesting great businesses that have a lot of potential from a free cash flow ge So our strategy hasn't changed.

127 e in some really interesting great businesses that have a lot of potential from a free cash flow ge Our outlook hasn't changed in that regard.

119 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl Any help on that could help.

120 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl Thomas J.

121 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl Szkutak Former Chief Financial Officer and Senior Vice President Sure.

122 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl In terms of the way we're looking at them, certainly, is based on the free cash flow potential, and we have a -- we're in some really interesting great businesses that have a lot of potential from a free cash flow generation standpoint, with good and high ROICs, which is exciting.

123 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl From a margin standpoint, always challenging to predict where that will come out in terms of absolute numbers.

124 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl But what we will do is we want to make sure that we try to maximize free cash flow.

125 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl That's something that we've always said.

126 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl So our strategy hasn't changed.

127 we a -- we're in some really interesting great businesses that have a lot of potential from a free cash fl Our outlook hasn't changed in that regard.

[[{"label": "positive", "score": 0.3271584212779999}, {"label": "negative", "score": 0.013711443170905113}, {"label": "neutral", "score": 0.65913009643554699}], [{"label": "positive", "score": 0.06439857979944992}, {"label": "negative", "score": 0.03958840795677417}, {"label": "neutral", "score": 0.8960179618644714}], [{"label": "positive", "score": 0.036549996584653854}, {"label": "negative", "score": 0.032875820994377136}, {"label": "neutral", "score": 0.9305741786095678}], [{"label": "positive", "score": 0.55514095303293}, {"label": "negative", "score": 0.019182670712471}, {"label": "neutral", "score": 0.43391121527671813}], [{"label": "positive", "score": 0.0555229868078232}, {"label": "negative", "score": 0.06598692387342453}, {"label": "neutral", "score": 0.8784898519515991}], [{"label": "positive", "score": 0.27610865235328674}, {"label": "negative", "score": 0.009935270994901657}, {"label": "neutral", "score": 0.7139560580253601}], [{"label": "positive", "score": 0.07375810795343399}, {"label": "negative", "score": 0.01787830387353897}, {"label": "neutral", "score": 0.8884135484695435}], [{"label": "positive", "score": 0.10310310870409012}, {"label": "negative", "score": 0.01247670435979843}, {"label": "neutral", "score": 0.8644199967384318}], [{"label": "positive", "score": 0.11293850299079971}, {"label": "negative", "score": 0.036971865941319416}, {"label": "neutral", "score": 0.8500864293907366}]]

df										
32]										
..										
	company_name	QQ_YYYY	Exact_Date_Of_Releasee	Overall_Sentiment_Of_QA	Net_Sentiment_Score	class	Exact_Date_Of_Releasee2	Final_Date	ticker	
	0	Alphabet	Q1 2014	Apr 16, 2014	2	649	0	2014-04-16	2014-03-31	GOOGL
	1	Alphabet	Q1 2015	Apr 23, 2015	2	544	0	2015-04-23	2015-03-31	GOOGL
	2	Alphabet	Q1 2016	Apr 21, 2016	2	461	0	2016-04-21	2016-03-31	GOOGL
	3	Alphabet	Q1 2017	Apr 27, 2017	2	474	0	2017-04-27	2017-03-31	GOOGL
	4	Alphabet	Q1 2018	Apr 23, 2018	2	401	0	2018-04-23	2018-03-31	GOOGL

	194	Apple Inc.	Q4 2018	Nov 01, 2018	2	337	-1	2018-11-01	2018-12-31	AAPL
	195	Apple Inc.	Q4 2019	Oct 30, 2019	2	389	0	2019-10-30	2019-12-31	AAPL
	196	Apple Inc.	Q4 2020	Oct 29, 2020	2	436	0	2020-10-29	2020-12-31	AAPL
	197	Apple Inc.	Q4 2021	Oct 28, 2021	2	540	0	2021-10-28	2021-12-31	AAPL
	198	Apple Inc.	Q4 2022	Oct 27, 2022	2	440	0	2022-10-27	2022-12-31	AAPL
199 rows × 9 columns										

Evaluation

Model Performance - RMSE

	LR RMSE (Test) (1st Degree)	LR RMSE (Test) (2nd Degree)	LR RMSE (Test) (3rd Degree)	Lasso RMSE (Test) (1st Degree)	Lasso RMSE (Test) (2nd Degree)	Lasso RMSE (Test) (3rd Degree)	Ridge RMSE (Test) (1st Degree)	Ridge RMSE (Test) (2nd Degree)	Ridge RMSE (Test) (3rd Degree)	PCR RMSE (Test)	NN RMSE (Test)
META	0.93	0.30	0.56	0.33	0.33	0.33	0.39	0.30	0.56	0.27	0.34
AAPL	0.55	0.20	0.24	0.16	0.16	0.16	0.41	0.20	0.24	0.15	0.15
GOOGL	0.07	0.18	0.37	0.14	0.14	0.14	0.06	0.18	0.37	0.07	0.14
NFLX	0.33	0.38	0.29	0.31	0.31	0.31	0.31	0.38	0.29	0.21	0.30
AMZN	0.28	0.22	0.25	0.21	0.21	0.22	0.28	0.22	0.25	0.20	0.22
Average	0.43	0.26	0.34	0.23	0.23	0.23	0.29	0.26	0.34	0.18	0.23

Model Performance w/ Sentiment Scores - RMSE

	LR RMSE (Test) (1st Degree)	LR RMSE (Test) (2nd Degree)	LR RMSE (Test) (3rd Degree)	Lasso RMSE (Test) (1st Degree)	Lasso RMSE (Test) (2nd Degree)	Lasso RMSE (Test) (3rd Degree)	Ridge RMSE (Test) (1st Degree)	Ridge RMSE (Test) (2nd Degree)	Ridge RMSE (Test) (3rd Degree)	NN w/ Sentiment RMSE (Test)
META	0.78	0.31	0.50	0.33	0.33	0.33	0.34	0.31	0.50	0.33
AAPL	0.53	0.20	0.22	0.16	0.16	0.16	0.40	0.20	0.22	0.15
GOOGL	0.11	0.17	0.30	0.14	0.14	0.14	0.07	0.17	0.30	0.16
NFLX	0.37	0.39	0.29	0.31	0.31	0.31	0.36	0.39	0.29	0.32
AMZN	0.28	0.22	0.25	0.21	0.21	0.22	0.28	0.22	0.25	0.21
Average	0.41	0.26	0.31	0.23	0.23	0.23	0.29	0.26	0.31	0.23

Evaluation

Model Performance - MAPE

	LR MAPE (Test) (1st Degree)	LR MAPE (Test) (2nd Degree)	LR MAPE (Test) (3rd Degree)	Lasso MAPE (Test) (1st Degree)	Lasso MAPE (Test) (2nd Degree)	Lasso MAPE (Test) (3rd Degree)	Ridge MAPE (Test) (1st Degree)	Ridge MAPE (Test) (2nd Degree)	Ridge MAPE (Test) (3rd Degree)	PCR RMSE (Test)	NN MAPE (Test)
META	13.98	1.93	2.94	2.41	2.41	2.41	4.88	1.94	2.94	1.12	1.70
AAPL	11.55	3.31	2.90	2.38	2.38	2.38	8.56	3.15	2.91	1.46	1.79
GOOGL	0.45	1.22	3.65	1.12	1.12	1.12	0.43	1.21	3.65	0.45	1.12
NFLX	2.98	6.93	0.98	2.82	2.82	2.82	2.94	6.91	0.98	1.22	2.27
AMZN	3.27	1.78	3.19	2.10	2.10	2.18	3.20	1.78	3.19	1.44	2.07
Average	6.45	3.03	2.73	2.17	2.17	2.19	4.00	3.00	2.74	1.14	1.79

Model Performance w/ Sentiment Scores - MAPE

	LR MAPE (Test) (1st Degree)	LR MAPE (Test) (2nd Degree)	LR MAPE (Test) (3rd Degree)	Lasso MAPE (Test) (1st Degree)	Lasso MAPE (Test) (2nd Degree)	Lasso MAPE (Test) (3rd Degree)	Ridge MAPE (Test) (1st Degree)	Ridge MAPE (Test) (2nd Degree)	Ridge MAPE (Test) (3rd Degree)	NN w/ Sentiment MAPE (Test)
META	11.73	1.92	2.47	2.41	2.41	2.41	3.94	1.91	2.46	1.58
AAPL	11.17	3.13	2.56	2.38	2.38	2.38	8.39	3.14	2.55	2.02
GOOGL	1.02	1.04	2.95	1.12	1.12	1.12	0.57	1.03	2.95	1.12
NFLX	3.10	7.24	1.49	2.70	2.70	2.70	2.82	7.23	1.49	1.85
AMZN	3.19	1.87	3.13	2.10	2.10	2.18	3.14	1.86	3.13	1.57
Average	6.04	3.04	2.52	2.14	2.14	2.16	3.77	3.04	2.52	1.63

Comparison of model predictions





Conclusions / Observations

- Sentiment analysis on its own had decent prediction accuracy (next day / quarter)
- However, adding sentiment analysis into our benchmark ML models did not improve performance.
- Top performing predictions were not particularly meaningful - constant predictions/small variance.

Areas for further exploration

- Extending sentiment analysis to non-company published text (i.e. news articles, twitter, analyst reports) would allow modelling on shorter time slices (monthly/weekly/daily).
- Increasing sample size beyond 10 years using the same methodology.



Q&A

Thank you :)



Appendix