Sales and Sentiment Case Study

Sebastian Algharaballi-Yanow

Project Overview

- Preprocess, visualize, and analyze stock, earnings per share (EPS), and revenue data for Apple and Lululemon. Find unique trends and commonalities that can be correlated towards market conditions, global events, or, potentially, sentiment during earnings calls.

- Analyze sentiment in Apple and Lululemon earnings calls and compare sentiment trends across the two companies.
    - Will potential differences come from the two companies being in different sectors?

- MAIN GOAL: Investigate correlation between sentiment in earnings calls and financial metrics.

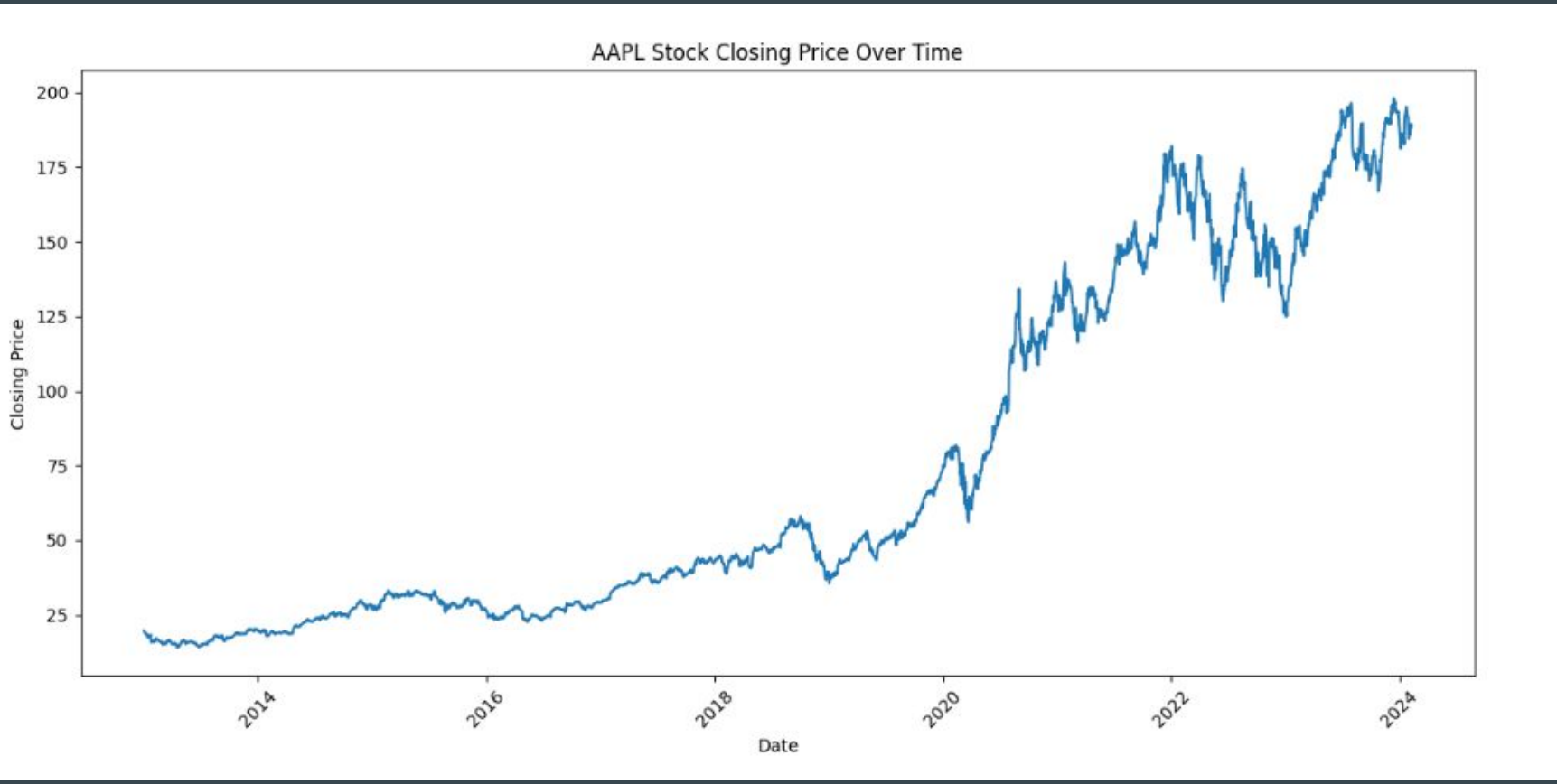
Methodology and Early Analysis

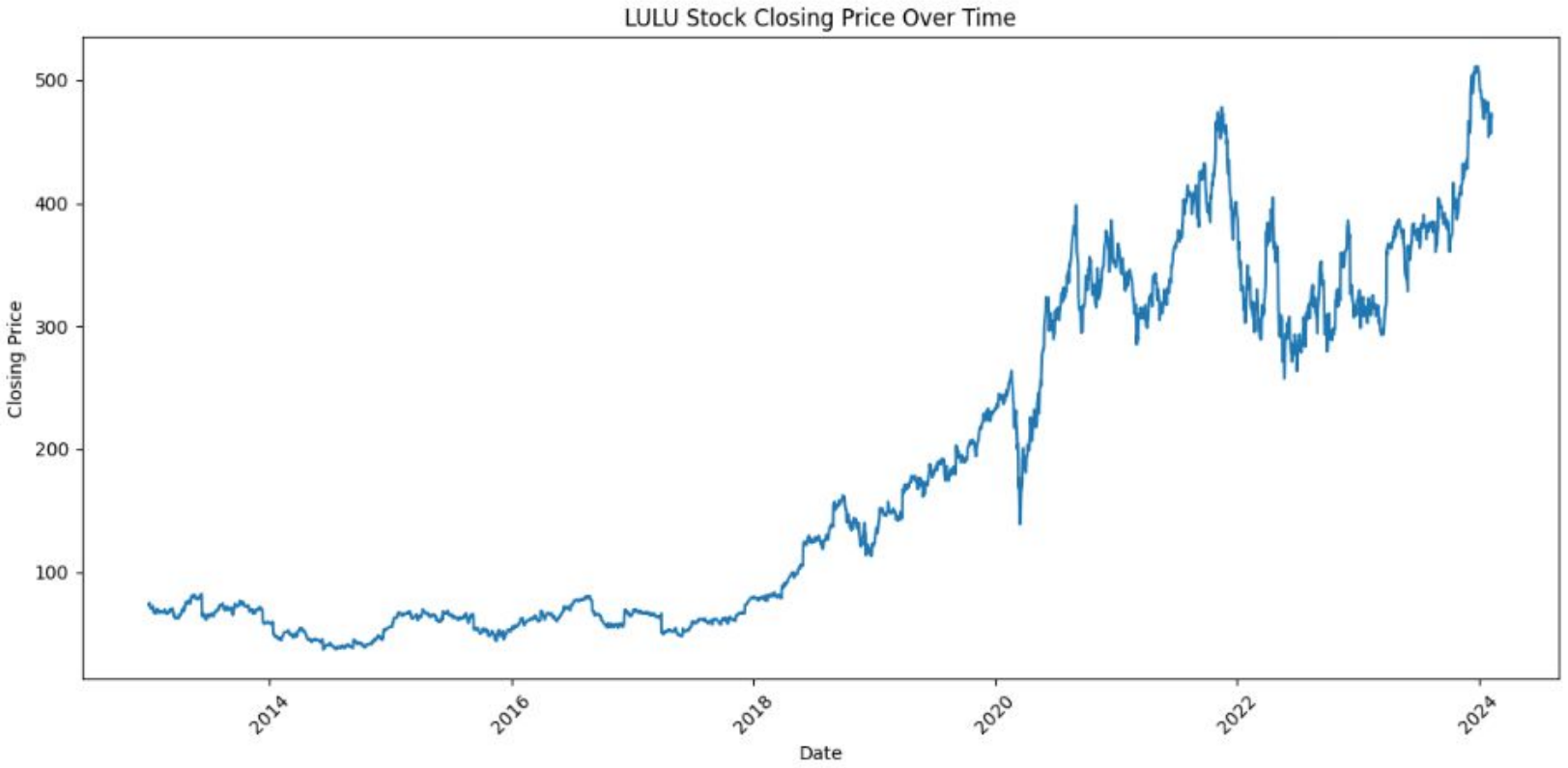
...

Data Preparation/Initial Stock Analysis

- Prepare stock data with a custom function that loads the data, converts the exchange date to datetime, and converts numeric columns to float data types for proper computation.

- Deploy function to learn more about both Apple and Lululemon stock closing price from 2014-2024.
  - Both companies show a gradual increase up to the 2020 pandemic, where significant crashes can be seen.
  - Both companies recovered to new heights in the years following the pandemic, with Apple showcasing more-steady growth and less drastic falls compared to Lululemon.

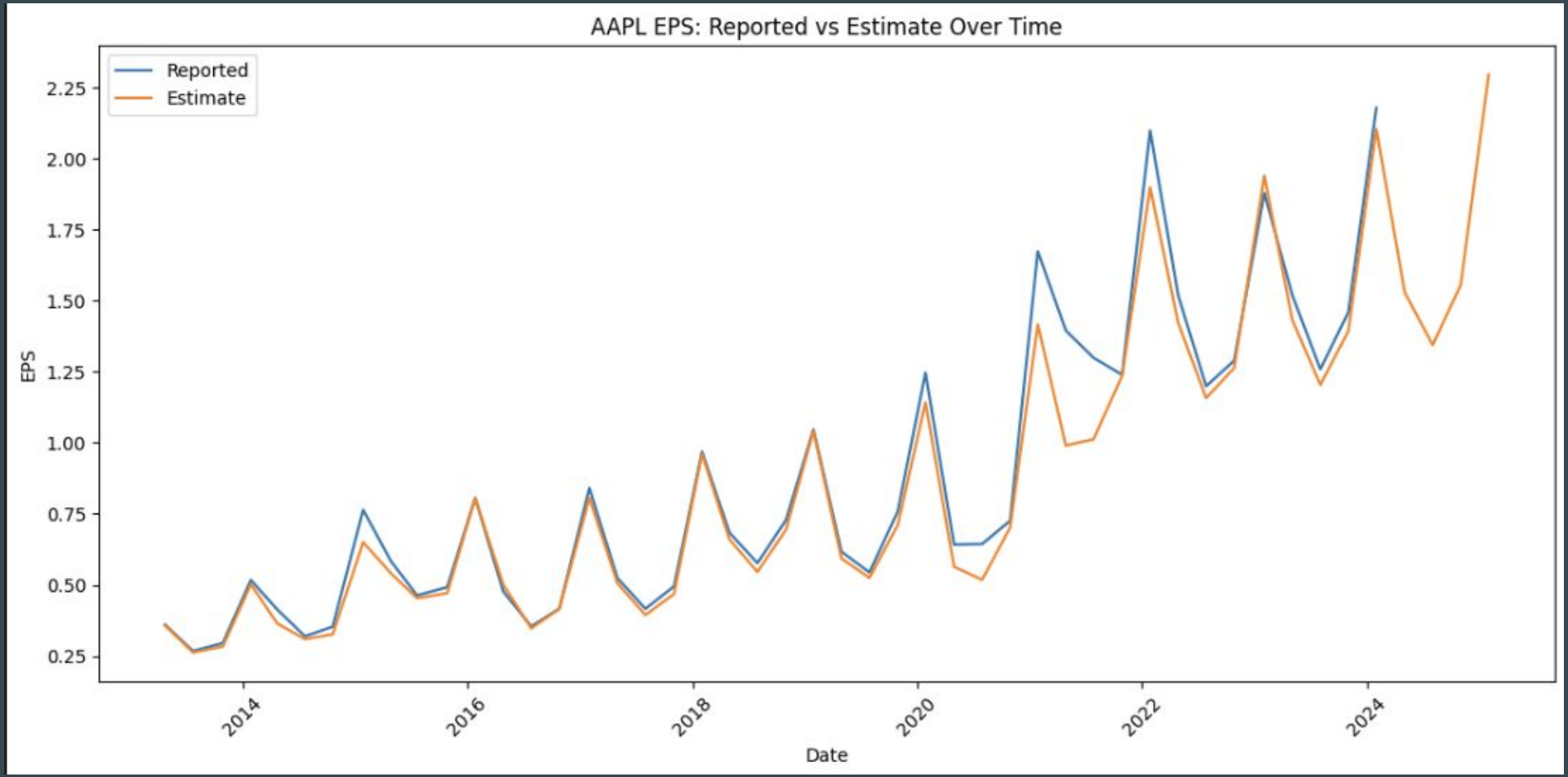

AAPL Stock Closing Price Over Time
Closing Price
200
175
150
125
100
75
50
25
2014
2016
2018
2020
2022
2024
Date

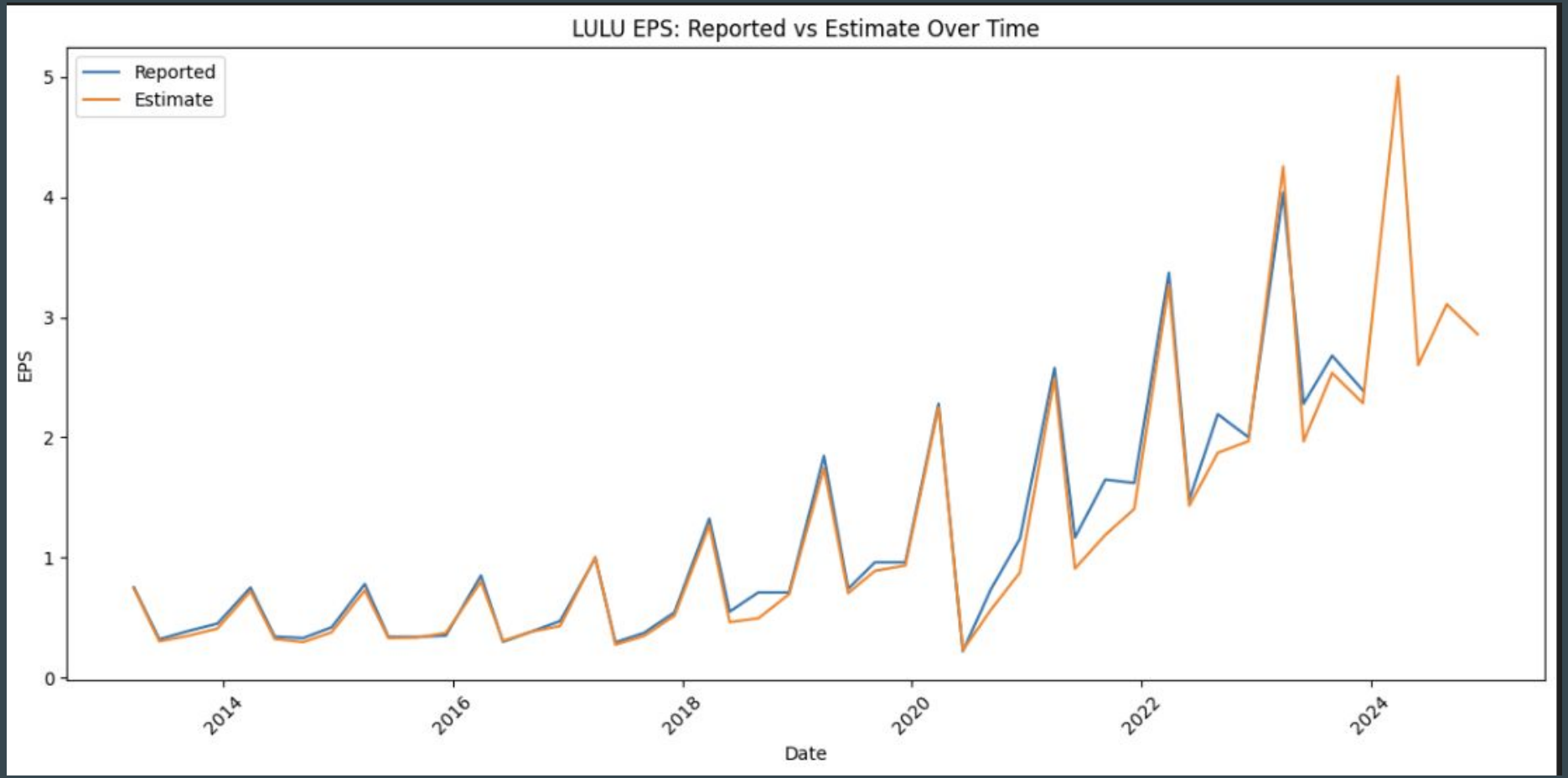

LULU Stock Closing Price Over Time
Closing Price
500
400
300
200
100
2014
2016
2018
2020
2022
2024
Date

Data Preparation/Initial EPS Analysis

- Prepare EPS data with a custom function that loads the data, converts numeric columns to floating-point numbers, and converts the “beat prediction?” variable to boolean-type.

- Deploy function to learn more about both Apple and Lululemon EPS data from 2014-2024.
  - Apple surpassed its estimated EPS in 2015, 2020, 2021, and 2024. Reported and estimated EPS was similar for all other years. Lower overall closing prices when compared to Lululemon (more people selling than buying stock).
  - Lululemon’s reported and estimated EPS was mostly steady and in line throughout the years, with slight underestimates found in 2018, 2021, 2022, and 2023.

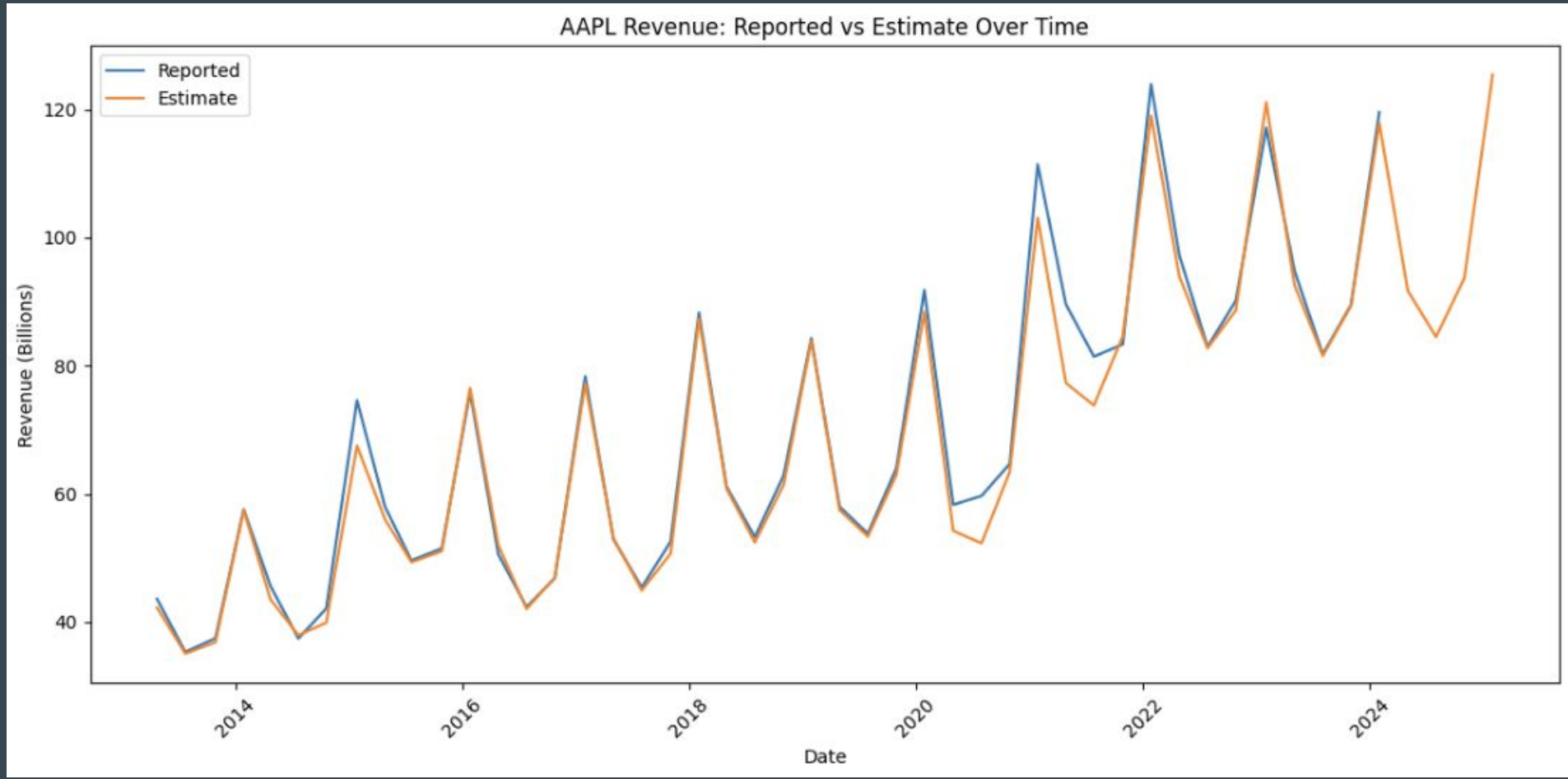

AAPL EPS: Reported vs Estimate Over Time
Reported
Estimate
EPS
2.25
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
2014
2016
2018
2020
2022
2024
Date

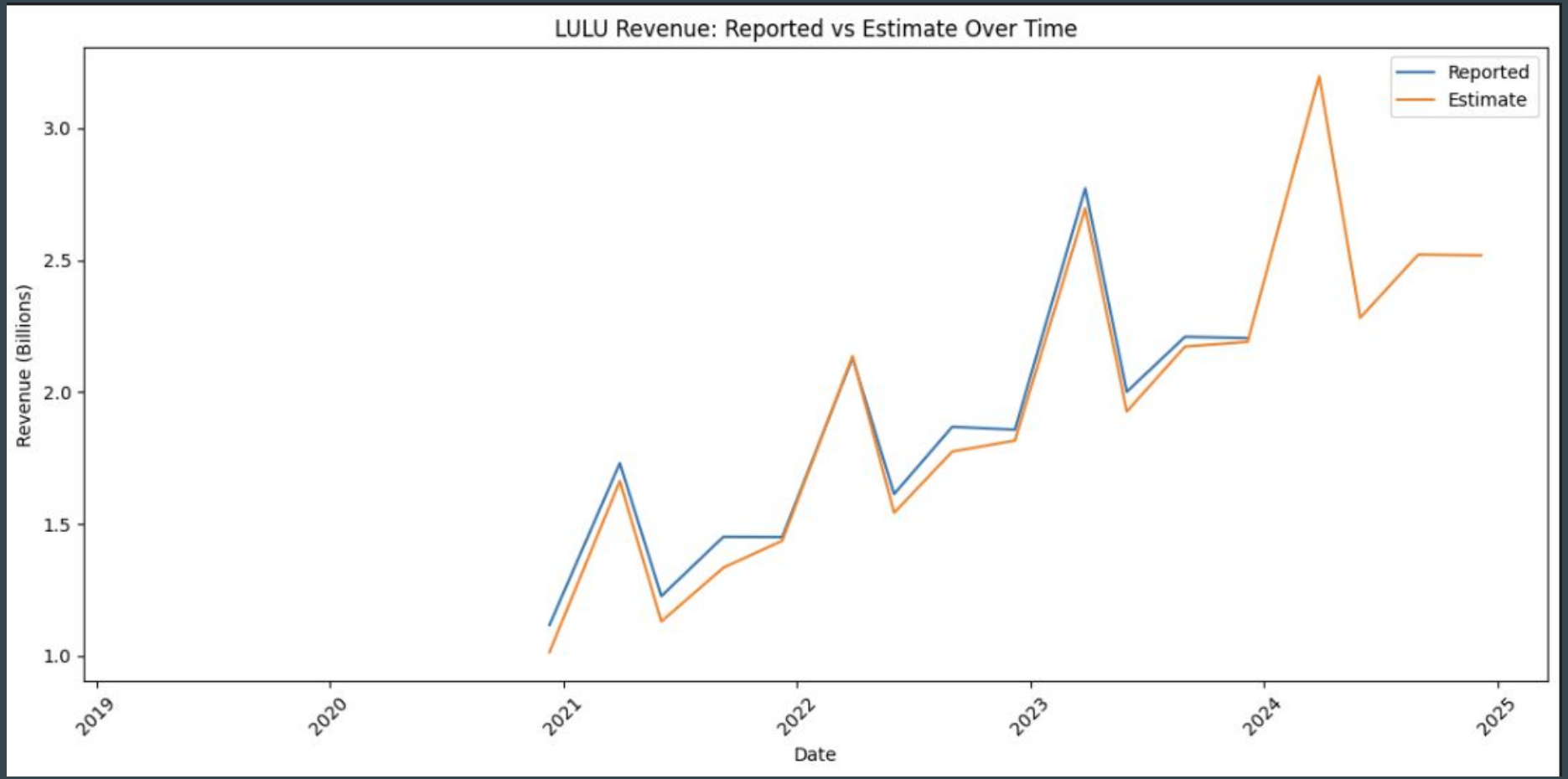

LULU EPS: Reported vs Estimate Over Time
Reported
Estimate
EPS
0
1
2
3
4
5
2014
2016
2018
2020
2022
2024
Date

Data Preparation/Initial Revenue Analysis

- Prepare revenue data with a custom function that loads the data, converts the reported revenue date to datetime, converts numeric columns to floating-point numbers, and converts the “beat prediction?” variable to boolean-type.

- Deploy function to learn more about both Apple and Lululemon revenue data from 2014-2024.
  - Apple follows very similar trends when compared to its EPS data. Surpassed estimated revenue in 2015, 2020, 2021, and 2022.
  - Lululemon surpassed its estimated revenue throughout each year we have the data for (2021, 2022, and 2023).


AAPL Revenue: Reported vs Estimate Over Time
Reported
Estimate
Revenue (Billions)
120
100
80
60
40
2014
2016
2018
2020
2022
2024
Date


LULU Revenue: Reported vs Estimate Over Time
Reported
Estimate
Revenue (Billions)
3.0
2.5
2.0
1.5
1.0
2019
2020
2021
2022
2023
2024
2025
Date

Takeaways

- Custom data preparation functions enabled precise analysis of stock, EPS, and revenue data for both companies from 2014-2024.

- Both Apple and Lululemon showed growth until the 2020 pandemic crash, followed by recovery; Apple demonstrated steadier post-pandemic growth compared to Lululemon's more volatile performance.

- Apple surpassed EPS and revenue estimates in select years (notably 2015, 2020-2022), while Lululemon consistently beat revenue estimates from 2021-2023 despite some EPS underestimates.

Sentiment Analysis

...

Methodology

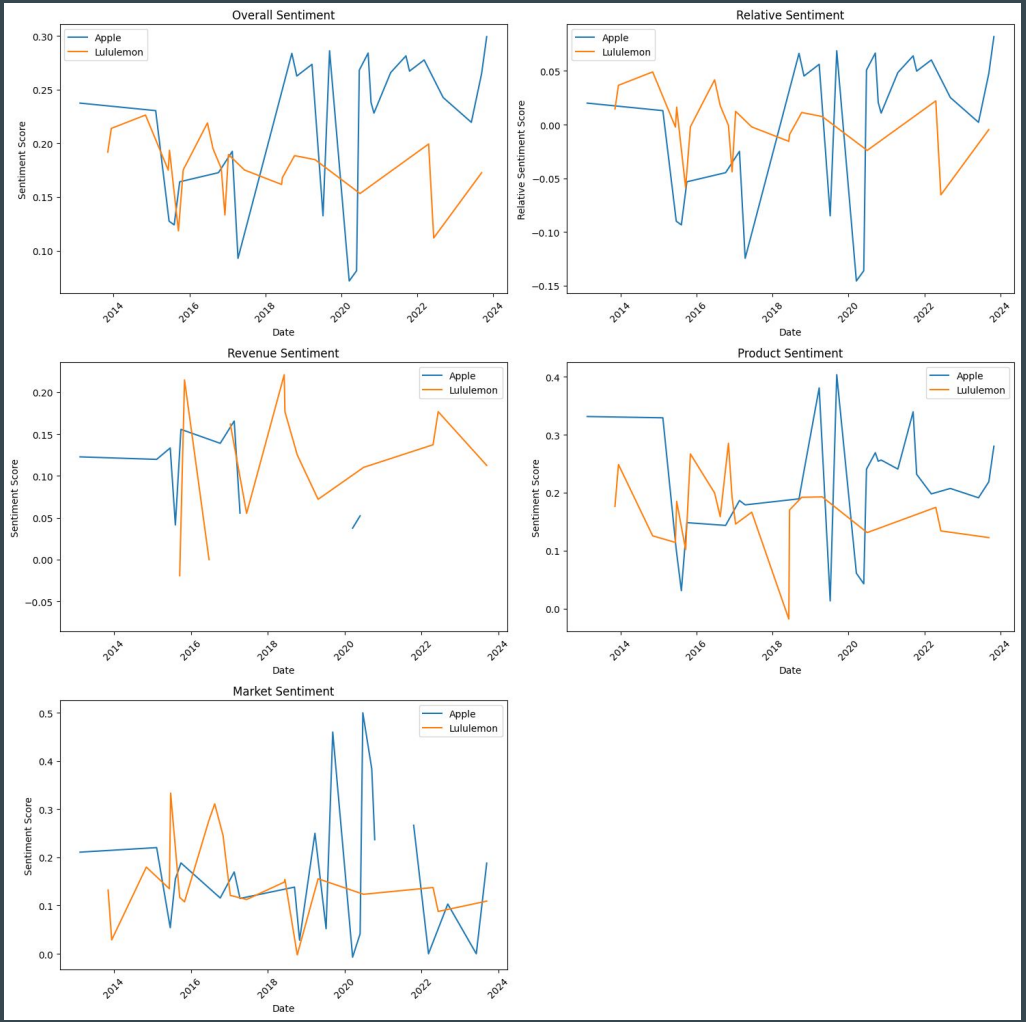
- Prepare transcript data with a custom function that loads the files, extracts dates from filenames using regex, and converts dates to datetime objects for chronological analysis. Function handles various date formats (YYYYMMDD, YYYYMM, YYYY) to ensure data inclusion.

- Deploy function to process transcripts for both Apple and Lululemon from 2013-2023.
    - Successfully loaded 27 transcripts for Apple and 21 for Lululemon.
    - Transcripts span from February 2013 to October 2023 for Apple, and November 2013 to September 2023 for Lululemon.

Methodology (cont.)

- Text chunking for consistent analysis:
  - Splits transcripts into 1000-word segments for fair comparison between short and long transcripts.
- TextBlob for sentiment scoring:
  - Provides polarity/positivity scores from -1 (negative) to 1 (positive).
  - Chosen for its pre-trained model on a large corpus of text data.
- Aspect-specific sentiment analysis:
  - Isolates sentences containing keywords: 'revenue', 'product', 'market'.
  - Calculates separate sentiment scores for each business aspect.
- Relative sentiment calculation:
  - Subtracts mean sentiment from each transcript's score.
  - Identifies transcripts with unusually positive or negative tones.
- Multi-dimensional visualization:
  - Creates 5 different time-series plots (overall, relative, revenue, product, market sentiments) to enable direct comparison of Apple and Lululemon sentiment trends over time

- Overall Sentiment:
  - Apple's sentiment generally higher and more volatile than Lululemon's.
  - Both companies show sentiment drops around 2020, likely due to pandemic.
- Relative Sentiment:
  - Apple experiences more extreme sentiment swings compared to its average.
  - Lululemon maintains more consistent sentiment relative to its mean.
- Revenue Sentiment:
  - Less data points available compared to other sentiment measures.
  - Lululemon shows higher revenue sentiment peaks than Apple.
- Product Sentiment:
  - Apple's product sentiment more variable, with higher peaks.
  - Lululemon's product sentiment more stable over time.
- Market Sentiment:
  - Apple's market sentiment highly volatile, especially post-2020.
  - Lululemon's market sentiment more consistent, with fewer extreme values.



```
Apple Most Positive Call:
date                  2023-10-31 00:00:00
overall_sentiment                0.299328
revenue_sentiment                     NaN
product_sentiment                0.280073
market_sentiment                      NaN
relative_sentiment                0.08195
Name: 22, dtype: object

Apple Most Negative Call:
date                  2020-03-17 00:00:00
overall_sentiment                0.071707
revenue_sentiment                0.037573
product_sentiment                0.061302
market_sentiment                -0.00703
relative_sentiment              -0.145671
Name: 19, dtype: object
```

- Apple most positive call : October 31, 2023
  - Holiday season anticipation
  - High product sentiment (0.280073)

- Apple most negative call: March 17, 2020
  - Early COVID-19 pandemic
  - Global market uncertainty
  - Negative market sentiment (-0.00703)

```
Lululemon Most Positive Call:
date                  2014-11-04 00:00:00
overall_sentiment                 0.226279
revenue_sentiment                -0.071019
product_sentiment                 0.125675
market_sentiment                  0.179798
relative_sentiment                0.049075
Name: 13, dtype: object

Lululemon Most Negative Call:
date                  2022-06-08 00:00:00
overall_sentiment                 0.111806
revenue_sentiment                 0.176667
product_sentiment                 0.134169
market_sentiment                  0.087828
relative_sentiment               -0.065398
Name: 14, dtype: object
```

- Lululemon most positive call: November 4, 2014
  - High market sentiment (0.179798).
  - Negative revenue sentiment (-0.071019).
  - Focus on future potential.

- Lululemon most negative call: June 8, 2022
  - Global supply chain issues.
  - Inflationary pressures.
  - High revenue sentiment (0.176667) despite overall negative tone.

Takeaways

- Apple's sentiment measures showed higher volatility across all categories (overall, revenue, product, market) compared to Lululemon's more stable sentiment trends from 2014-2024.

- Both companies experienced sentiment drops around the 2020 pandemic, with Apple's most negative call on March 17, 2020, and Lululemon's on June 8, 2022, reflecting global market uncertainties and supply chain issues.

- Apple's sentiment peaked on October 31, 2023, driven by holiday anticipation and strong product sentiment, while Lululemon's most positive call on November 4, 2014, focused on future potential despite negative revenue sentiment.

- Lululemon consistently showed higher revenue sentiment peaks than Apple, suggesting stronger market confidence in Lululemon's sales performance during the analyzed period.

Correlation Analysis with Text Embeddings

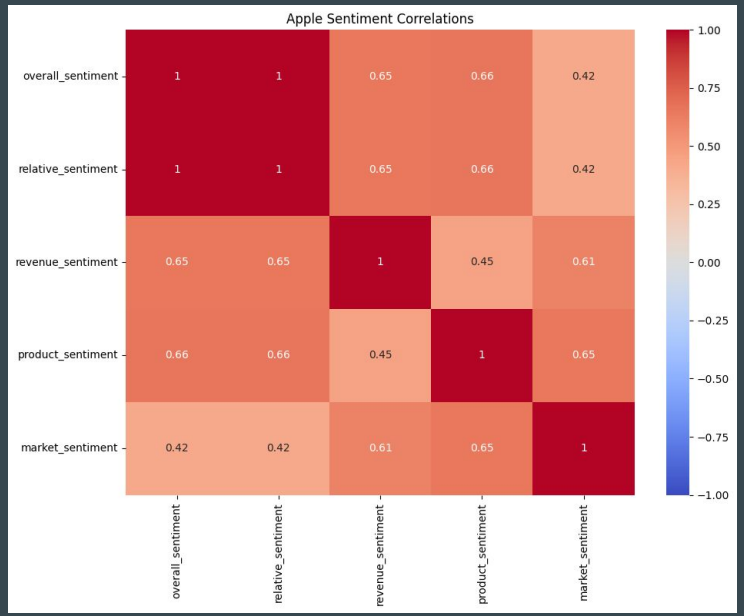
...

Methodology

- Text embedding with TF-IDF and SVD:
  - TF-IDF vectorization captures word importance in transcripts.
  - Truncated SVD reduces dimensionality to 10 components.
  - Allows for quantitative analysis of qualitative transcript data.
- Financial data preprocessing:
  - Converts string values (e.g., "5.2B") to numeric for analysis.
  - Aligns financial data with sentiment scores and embeddings.
  - Ensures consistent data types for correlation analysis.
- Correlation visualization techniques:
  - Heatmaps for different data subsets (sentiment, financial, embeddings).
  - Focused heatmap comparing sentiment/embeddings to financial metrics.
  - Thresholded heatmap to highlight strong correlations ($|r| > 0.3$).
  - Summary bar plot for average correlations with financial metrics.

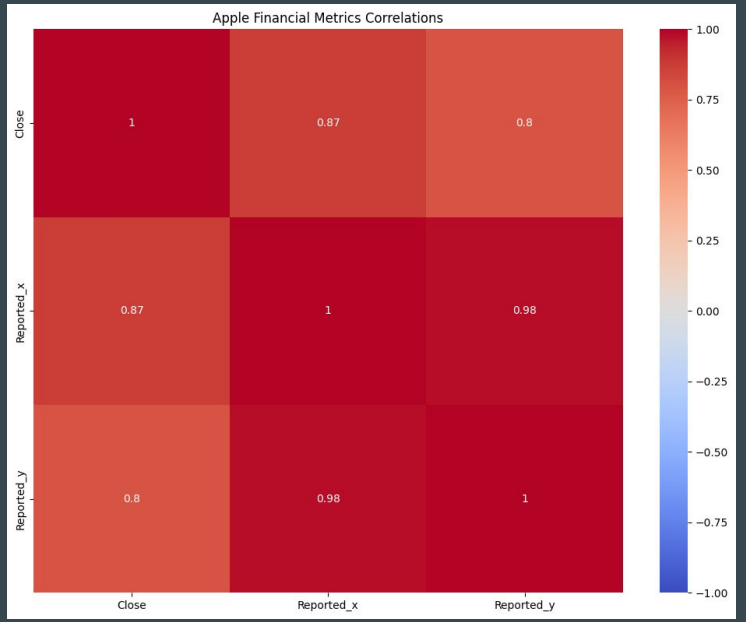
Methodology (cont.)

- Statistical tests:
  - Augmented Dickey-Fuller test checks for stationarity in time series.
  - Granger causality test examines potential causal relationships.
  - Helps understand the nature of relationships between variables.
- Time series visualization:
  - Plots stock price and overall sentiment on same timeline.
  - Enables visual inspection of potential relationships and lags.



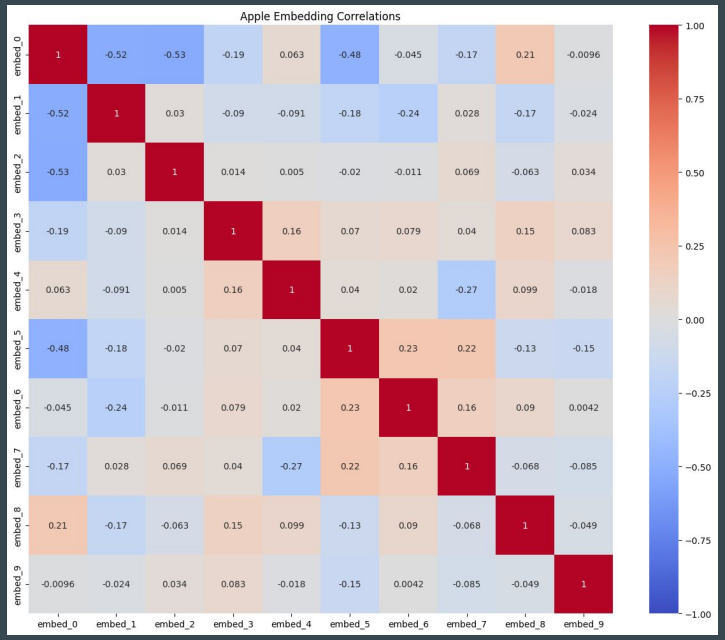
Sentiment Correlations (Apple):

- Overall and relative sentiment perfectly correlated (1.0).
- Strong correlation between overall/relative and product sentiment (0.66).
- Revenue sentiment more strongly correlated with market sentiment (0.61) than product sentiment (0.45)



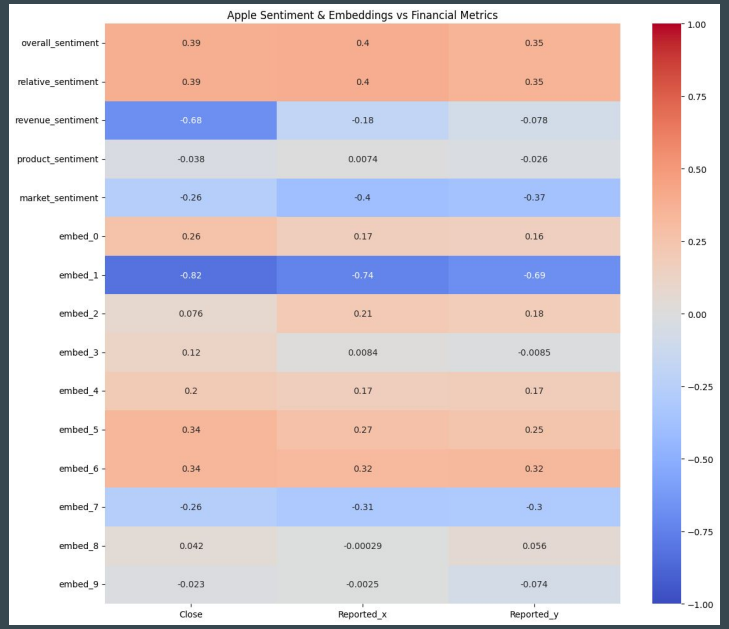
Financial Correlations (Apple):

- Very strong correlation of 0.98 between Reported_x (EPS) and Reported_y (Revenue).
- Close price strongly correlated with both Reported metrics (0.87 and 0.80).
- Suggests alignment between stock performance and financial reports.



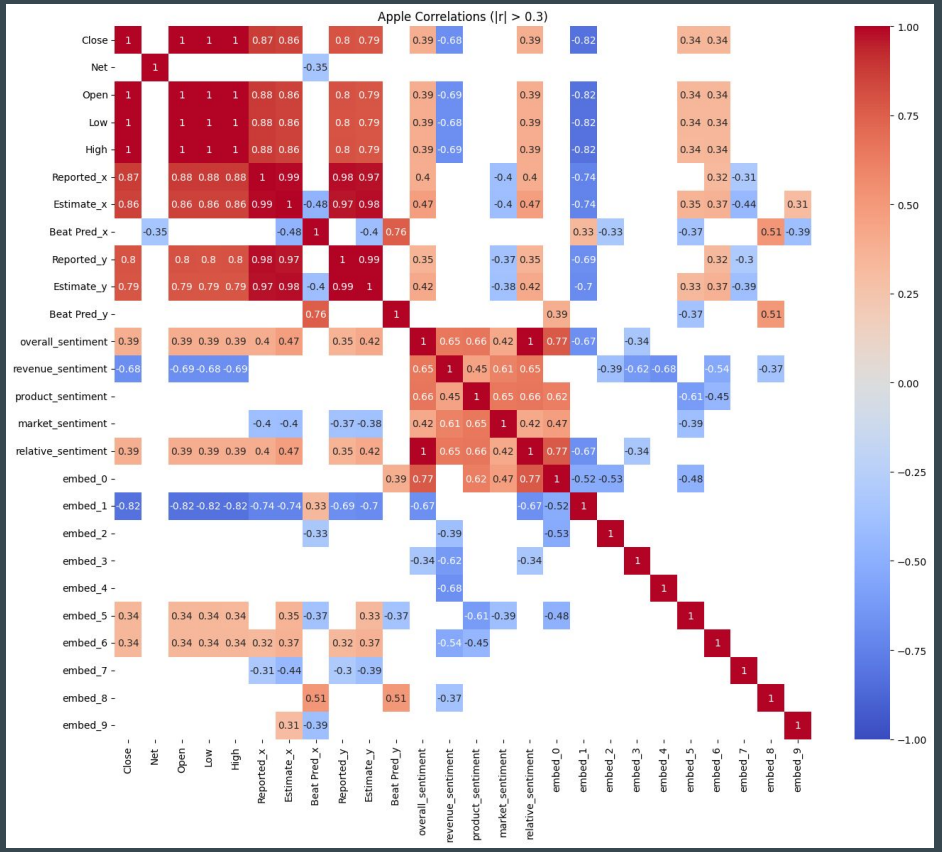
Embedding Correlations (Apple):

- Generally weak correlations between embeddings.
- Strongest negative correlation between embed_0 and embed_2 (-0.53).
- embed_5 and embed_6 show modest positive correlation (0.23).



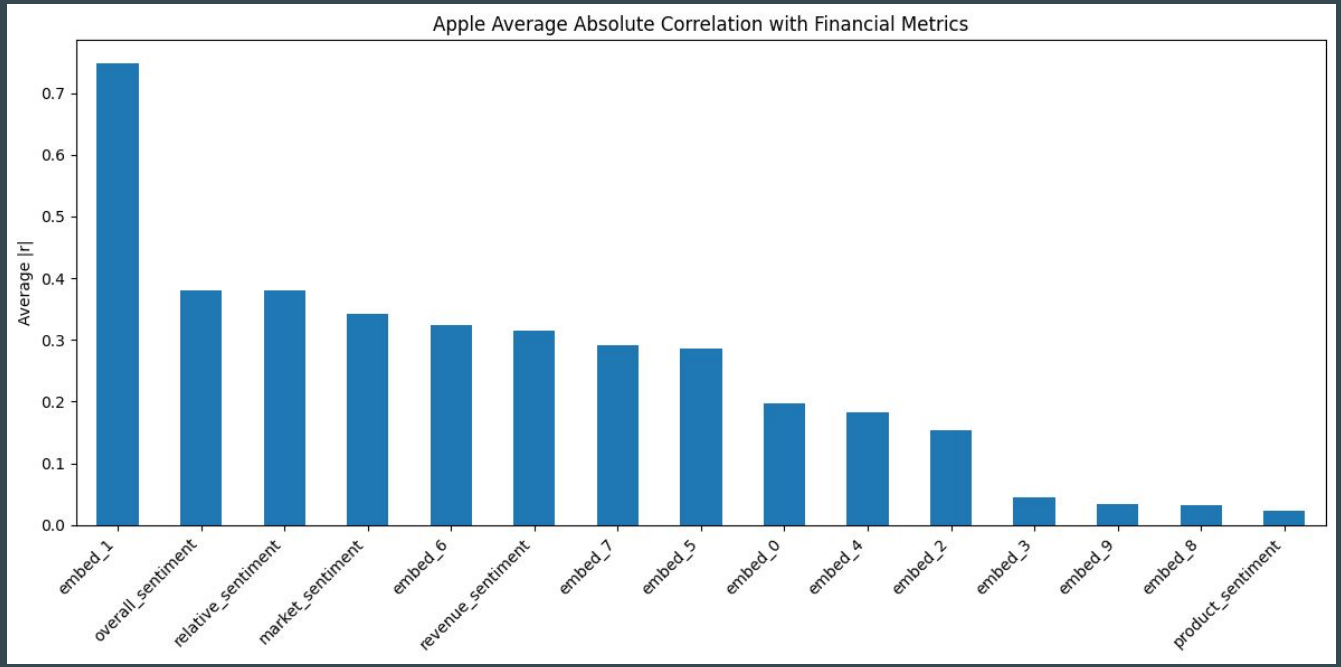
Sentiment & Embeddings vs Financial Metrics (Apple):

- Overall and relative sentiment moderately positively correlated with all financial metrics (0.35-0.40).
- Revenue sentiment shows strong negative correlation with Close price (-0.68).
- embed_1 has strong negative correlations with all financial metrics (-0.69 to -0.82).
- embed_5 and embed_6 show moderate positive correlations with financial metrics (0.25-0.34).



Embedding Correlations r > 0.3 (Apple):

- Financial Metrics Correlations:
  - Very strong correlation of 0.98 between Reported_x (EPS), Reported_y (Revenue), in addition to their estimated counterparts.
  - Close price strongly correlated with both Reported metrics (0.87 and 0.80).
  - High correlation among Close, Open, Low, and High prices (all near 1.0).
- Sentiment Correlations:
  - Overall and relative sentiment perfectly correlated (1.0)
  - Strong correlation between overall/relative and product sentiment (0.66)
  - Revenue sentiment more strongly correlated with market sentiment (0.61) than product sentiment (0.45).



Average Absolute Correlation (Apple):

- Highest correlations:
  - embed_1 shows the highest average absolute correlation (about 0.74).
  - overall_sentiment and relative_sentiment follow (both around 0.38).
  - market_sentiment ranks third (about 0.34).
- Lowest Correlations
  - embed_8, embed_9, and product_sentiment show the lowest average absolute correlations (all below 0.05).
  - embed_3 and embed_4 also show relatively low correlations (below 0.2).

```
Stationarity Test Results (p-values):
Close: 0.9786189813457074
Net: 6.996288900017537e-23
Open: 0.9774263901111406
Low: 0.9800819914164629
High: 0.98014728101566
Reported_x: 0.9989812631062924
Estimate_x: 0.9753265161062825
Beat Pred_x: 0.9865540735747521
Reported_y: 0.7828370980569936
Estimate_y: 0.8873841710111298
Beat Pred_y: 0.04343977802125083
overall_sentiment: 0.014414103501205826
revenue_sentiment: 0.24243214887419023
product_sentiment: 0.0008431890468617592
market_sentiment: 0.0004124142998247903
relative_sentiment: 0.01441410350120576
embed_0: 0.008482716147416353
embed_1: 0.42264906295173355
embed_2: 0.0009352123477497455
embed_3: 0.005924330667670108
embed_4: 0.0032358573082805864
embed_5: 0.03474718686289623
embed_6: 0.004541317212694055
embed_7: 0.008152803009816304
embed_8: 0.023219172941958615
embed_9: 0.10269596157494965
```

Stationarity Test Results (Apple):

- Non-stationary (high p-values > 0.05):
  - Most financial metrics: Close, Open, Low, High, Reported, Estimated
  - Revenue sentiment
  - embed_1, embed_9
- Stationary (low p-values < 0.05):
  - Net price (extremely low p-value: 6.996e-23)
  - Beat Pred_y
  - Overall, relative, product, and market sentiments
  - Most embeddings (0, 2, 3, 4, 5, 6, 7, 8)

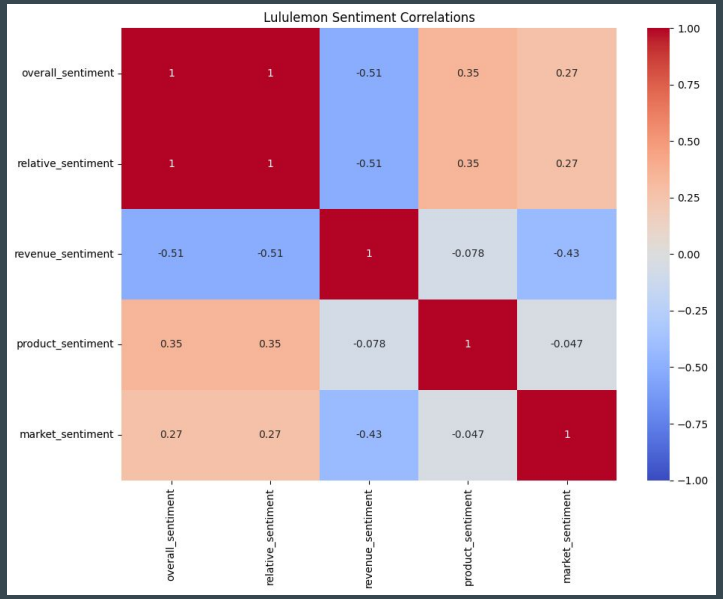
Most financial metrics are non-stationary, while sentiments and embeddings are largely stationary. This suggests that while sales data may show trends or seasonality, the earnings call text representations are more consistent over time, potentially allowing for more direct correlation analysis between calls and sales performance.

```
Granger Causality Test Results (p-values for each lag):
Lag 1: 0.06935962243134518
Lag 2: 0.1884394472706661
Lag 3: 0.33107355245297193
Lag 4: 0.4700890016837528
Lag 5: 0.6025370580551404
```

P-values increase with lag, suggesting weaker causality at higher lags. No strong evidence of Granger causality at any lag (all $p > 0.05$), but Lag 1 shows potential weak causality ($p = 0.069$). This indicates earnings call embeddings may have a slight, short-term predictive relationship with sales, but the effect quickly diminishes.

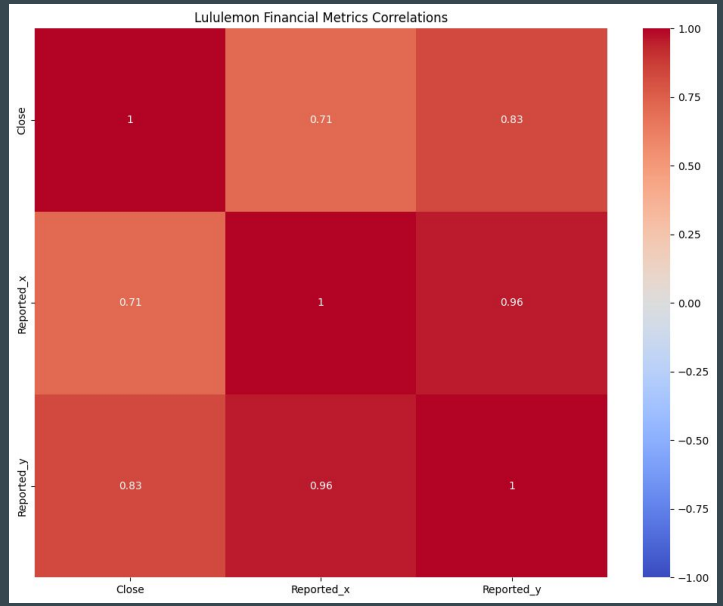
- Stock price (blue) shows clear upward trend from 2014 to 2024.
- Overall sentiment (orange) fluctuates between 0.10 and 0.30.
- No strong visual correlation between sentiment and short-term price movements.
- Sentiment appears more volatile in recent years (2020-2024).
- Notable sentiment drops: mid-2016, early 2020, mid-2022.
- Stock price growth accelerated post-2020 despite sentiment volatility.

While sentiment fluctuates, it doesn't consistently predict short-term stock price changes. Long-term stock growth seems independent of sentiment variations.



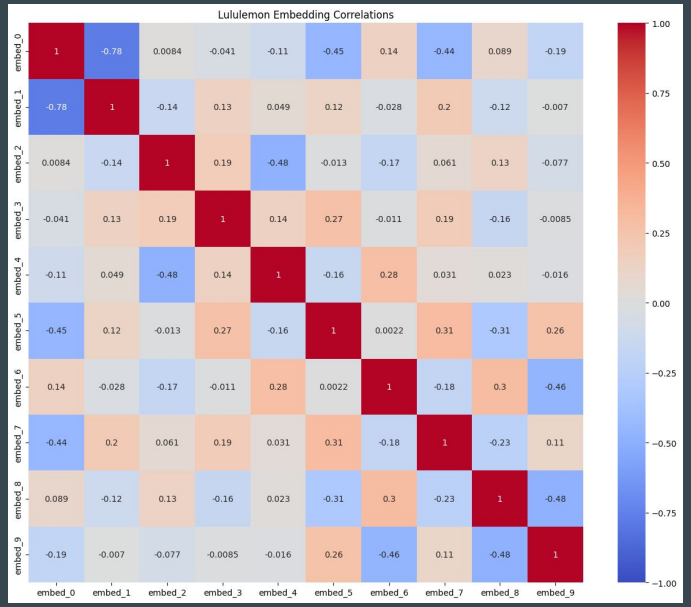
Sentiment Correlations (Lululemon):

- Overall and relative sentiment perfectly correlated (1.0).
- Strong negative correlation (-0.51) between overall/relative and revenue sentiment.
- Moderate positive correlation (0.35) between overall/relative and product sentiment.
- Weak positive correlation (0.27) between overall/relative and market sentiment.
- Revenue sentiment negatively correlated with all other sentiments.
- Overall sentiment aligns closely with product and market perceptions, but diverges from revenue expectations.



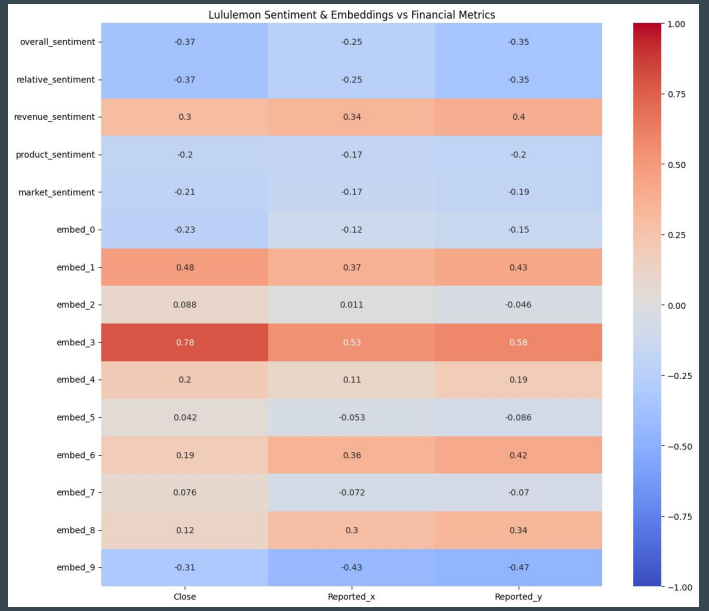
Financial Correlations (Lululemon):

- Very strong correlation (0.96) between Reported_x (EPS) and Reported_y (Revenue).
- Strong correlation (0.83) between Close price and Reported_y.
- Moderate correlation (0.71) between Close price and Reported_x.
- Reported financial metrics show high consistency and align well with stock price performance.



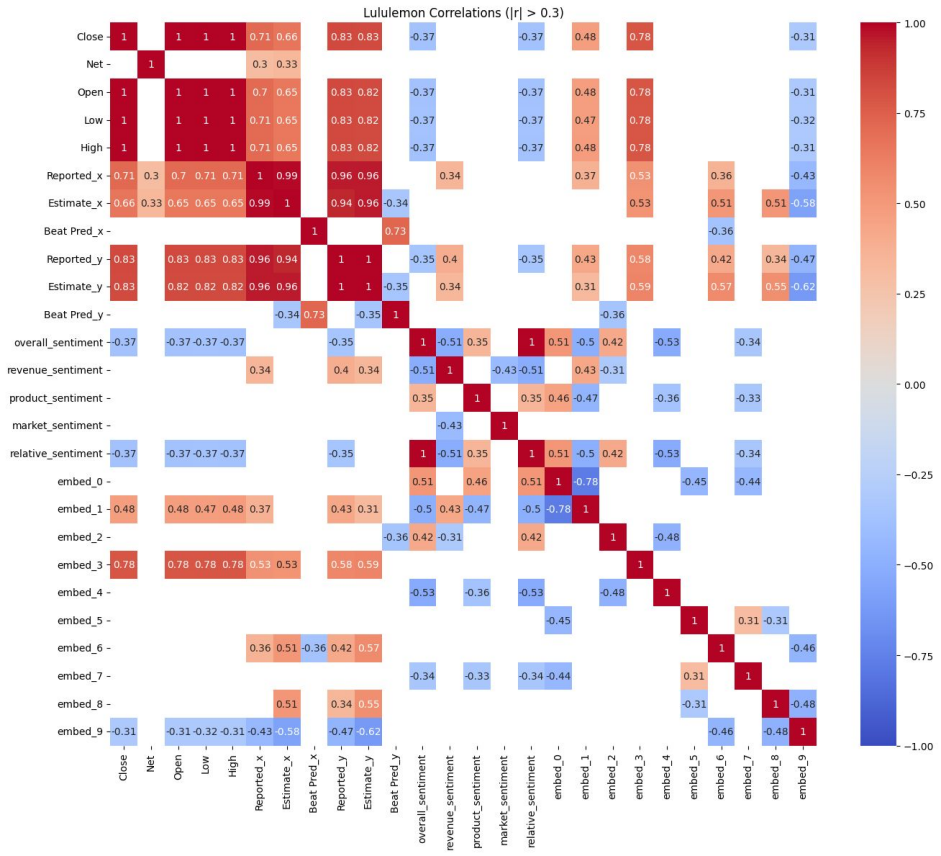
Sentiment Correlations (Lululemon):

- Strong negative correlation between embed_0 and embed_1 (-0.78).
- Moderate negative correlations between embed_0 and embed_5 (-0.45), and embed_0 and embed_7 (-0.44).
- Weak to moderate positive correlations between several pairs, e.g., embed_5 and embed_7 (0.31), embed_6 and embed_8 (0.30).
- Many embeddings show weak or near-zero correlations with each other.
- embed_3 shows relatively weak correlations with all other embeddings, strongest being 0.27 with embed_5.



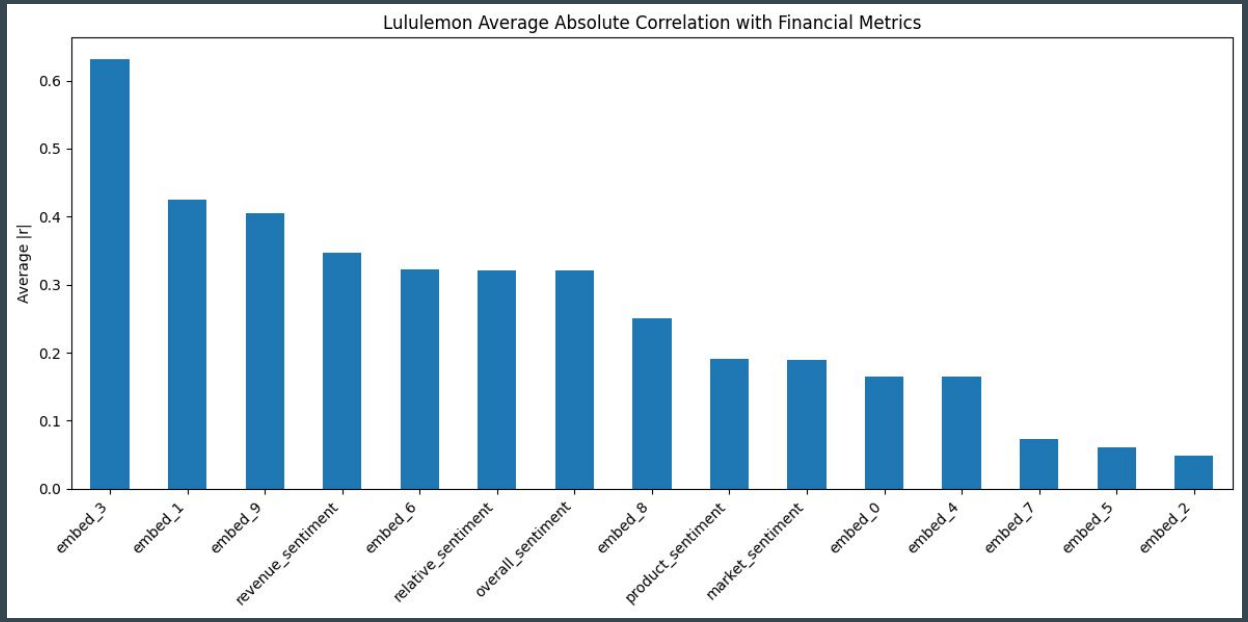
Sentiment & Embeddings vs Financial Metrics (Lululemon):

- Overall and relative sentiment moderately negatively correlated with all financial metrics (-0.25 to -0.37).
- Revenue sentiment shows moderate positive correlation with all financial metrics (0.30 to 0.40).
- embed_3 has strong positive correlations with all financial metrics (0.53 to 0.78).
- embed_1 and embed_6 show moderate positive correlations with financial metrics (0.37 to 0.48 for embed_1, 0.19 to 0.42 for embed_6).



Embedding Correlations r > 0.3 (Lululemon):

- Financial Metrics Correlations:
  - Very strong correlation of 0.96-0.99 between Reported_x, Reported_y, and their estimated counterparts.
  - Close price strongly correlated with both Reported metrics (0.83 for both).
  - High correlation among Close, Open, Low, and High prices (all 1.0 or near 1.0).
- Sentiment Correlations:
  - Overall and relative sentiment perfectly correlated (1.0).
  - Moderate negative correlation between overall/relative sentiment and revenue sentiment (-0.51).
  - Product sentiment moderately correlated with market sentiment (0.46).
- Embedding Correlations:
  - Strong negative correlation between embed_0 and embed_1 (-0.78).
  - Strong positive correlation between embed_3 and financial metrics (0.53 to 0.78).
  - Moderate negative correlations between embed_9 and financial metrics (-0.31 to -0.62).



Average Absolute Correlation (Lululemon):

- Highest correlations:
  - embed_3 shows the highest average absolute correlation (about 0.63).
  - embed_1 follows (about 0.43).
  - embed_9 ranks third (about 0.40).
- Lowest Correlations
  - embed_2 shows the lowest average absolute correlation (about 0.05).
  - embed_5 and embed_7 also show relatively low correlations (both below 0.10).
  - product_sentiment and market_sentiment have relatively low average absolute correlations (both around 0.19).

```
Stationarity Test Results (p-values):
Close: 0.9539865647820228
Net: 3.7102004397037433e-23
Open: 0.9442540417458952
Low: 0.9607124526964884
High: 0.9456175278096474
Reported_x: 1.0
Estimate_x: 0.9985755955575
Beat Pred_x: 0.15921374243748304
Reported_y: 1.0
Estimate_y: 0.9973043852269156
Beat Pred_y: 1.3669675921427389e-05
overall_sentiment: 0.01191492715494231
revenue_sentiment: 0.017965397671531786
product_sentiment: 0.003928439688391532
market_sentiment: 0.002042936396295338
relative_sentiment: 0.011914927154942617
embed_0: 0.004691453405494977
embed_1: 0.026665183015860867
embed_2: 0.0027388075280846908
embed_3: 0.43284286871493544
embed_4: 0.021728575973411466
embed_5: 0.005311862168173781
embed_6: 0.03119336465735544
embed_7: 0.0039047380246543736
embed_8: 0.05531891540457718
embed_9: 0.08196626612043872
```

Stationarity Test Results (Lululemon):

- Non-stationary (high p-values > 0.05):
  - Most financial metrics: Close, Open, Low, High, Reported_x, Reported_y, Estimate_x, Estimate_y.
  - Beat Pred_x.
  - embed_3, embed_8, embed_9.
- Stationary (low p-values < 0.05):
  - Net price (extremely low p-value: 3.71e-23).
  - Beat Pred_y.
  - All sentiment metrics: overall, revenue, product, market, and relative.
  - Most embeddings (0, 1, 2, 4, 5, 6, 7).

Most financial metrics are non-stationary, while sentiments and most embeddings are stationary. This suggests that while financial data may show trends or seasonality, the earnings call text representations and sentiment analyses are more consistent over time. This could potentially allow for more direct correlation analysis between calls and financial performance.

```
Granger Causality Test Results (p-values for each lag):
Lag 1: 0.2314833743013594
Lag 2: 0.2006430325316049
Lag 3: 0.3566039501987819 4
Lag 4: 0.3428993545693092
Lag 5: 0.3733318369515254 4
```

P-values for Granger Causality tests increase with lag, suggesting weaker causality at higher lags. There is no strong evidence of Granger causality at any lag (all $p > 0.05$). The lowest p-value is at Lag 2 ($p = 0.2006$), but it's still not statistically significant. This indicates that earnings call embeddings may not have a clear predictive relationship with financial metrics, at least not in the short term (up to 5 lags).

- Stock price (blue) shows clear upward trend from 2014 to 2024, with significant acceleration post-2018.
- Overall sentiment (orange) fluctuates between approximately 0.11 and 0.23.
- No strong visual correlation between sentiment and short-term price movements.
- Sentiment appears more stable pre-2016, with increased volatility afterwards.
- Notable sentiment spikes: late 2015, mid-2016, early 2017.
- Significant sentiment drops: mid-2016, late 2022 to early 2023.
- Stock price experienced a sharp drop in early 2020 (likely due to COVID-19) but quickly recovered and continued its upward trajectory.

While sentiment shows considerable fluctuations, it doesn't appear to consistently predict stock price movements. The long-term stock growth trend seems largely independent of sentiment variations, with the stock price showing steady increase despite sentiment volatility.

Takeaways

- Both Apple and Lululemon show no strong evidence of Granger causality between earnings call embeddings and financial metrics, suggesting limited short-term predictive power of call sentiments on stock performance.

- Lululemon's overall sentiment is negatively correlated with financial metrics (-0.25 to -0.37), while Apple's shows a moderate positive correlation (0.35-0.40). This divergence suggests industry-specific differences in how sentiment relates to financial performance.

- Revenue sentiment behaves differently for each company: Lululemon shows positive correlation with financial metrics (0.30 to 0.40), while Apple exhibits a strong negative correlation with stock price (-0.68). This indicates distinct market interpretations of revenue discussions for each brand.

Takeaways (cont.)

- Both companies display strong correlations between reported financial metrics (EPS and Revenue), but Lululemon (0.96) shows slightly higher consistency than Apple (0.98).

- Embedding correlations with financial metrics vary: Lululemon's embed_3 shows strong positive correlations (0.53 to 0.78), while Apple's embed_1 has strong negative correlations (-0.69 to -0.82). This suggests different semantic aspects of calls may be relevant for each company.

- Stationarity tests reveal that for both companies, most financial metrics are non-stationary, while sentiment measures are largely stationary. This consistency in sentiment measures over time could potentially allow for more reliable long-term analysis.

- Visual analysis of stock prices vs. overall sentiment shows that for both companies, long-term stock growth appears independent of sentiment fluctuations, with no consistent short-term predictive power observed.

In conclusion, while there are some correlations between earnings call sentiments and financial metrics, the relationship is complex and varies between Apple and Lululemon. The lack of clear causality and inconsistent correlations suggest that earnings call sentiment alone may not be a reliable predictor of short-term financial performance for either company.

Thank you!

...