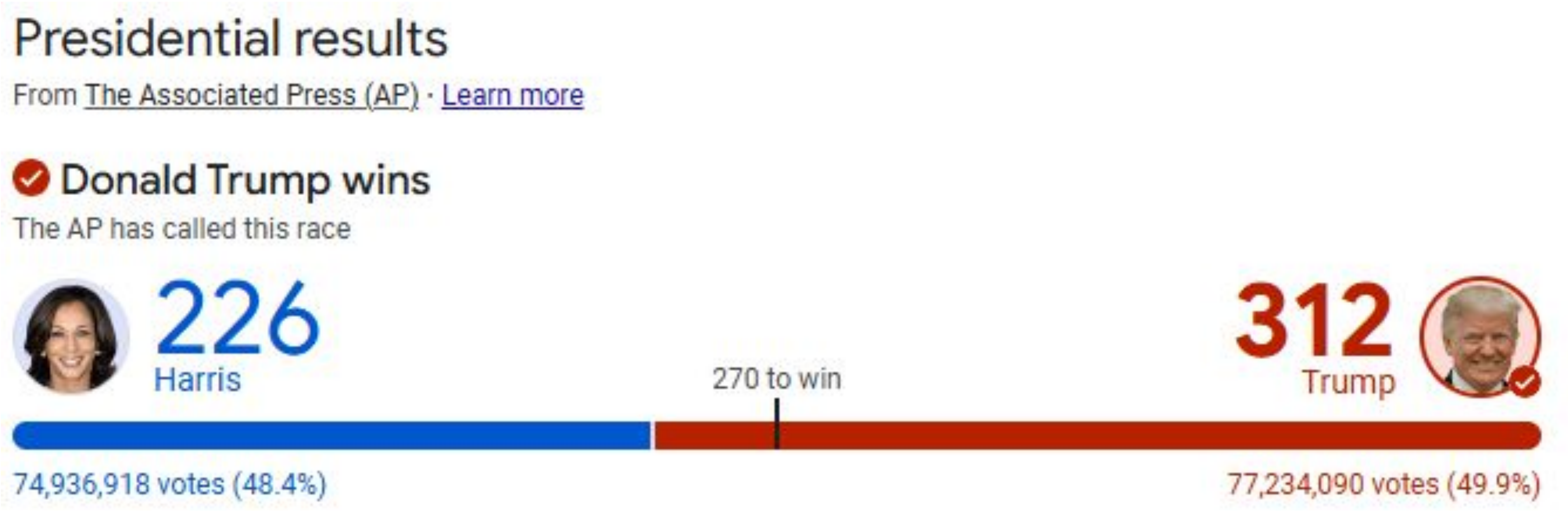
Election-Year Winds: How Politics can Shape Healthcare and Tech Stocks in 2024

PS5842 | Dennis Goldenberg, Samaa Nadkarni, Jiawen Shao

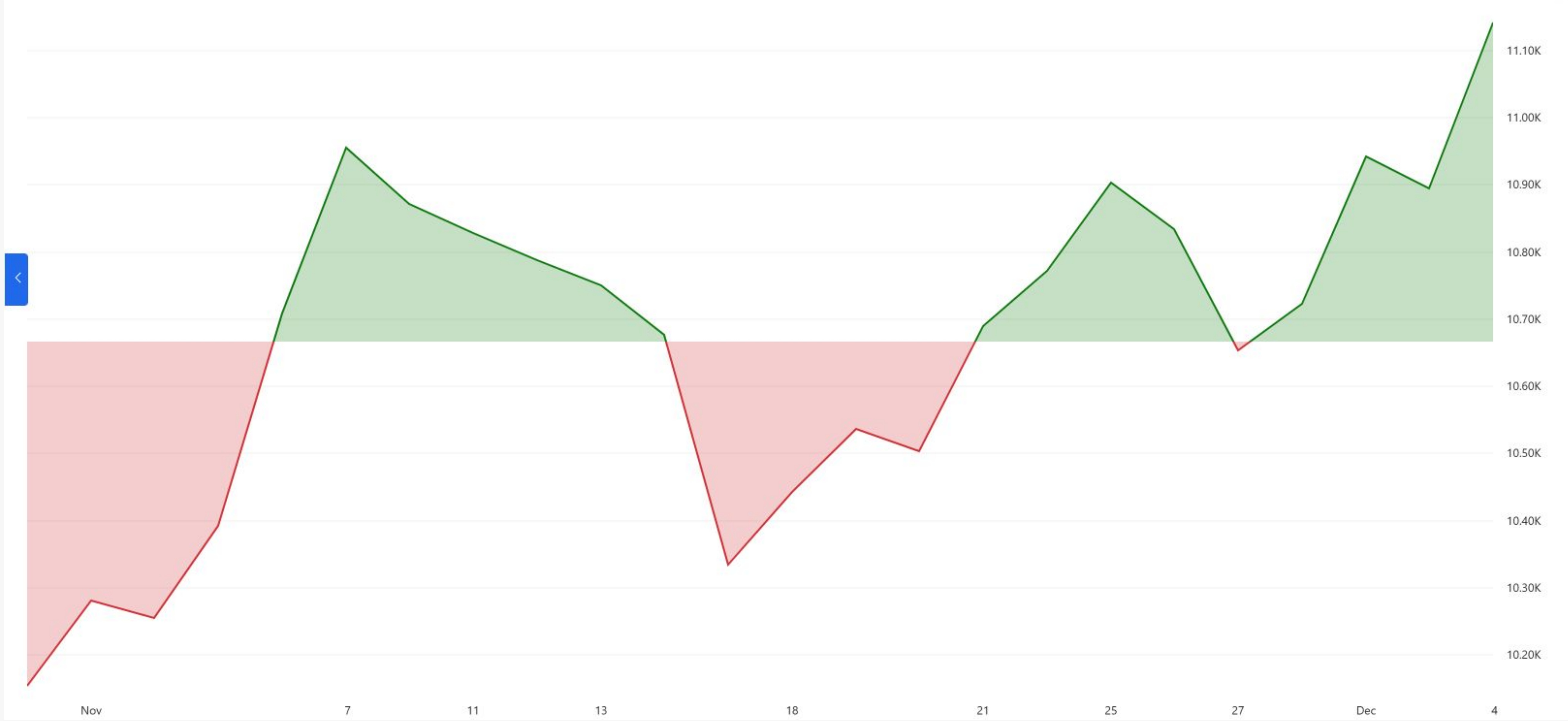
American Election and Stock Market



Election Day:
November 5th, 2024

Post Election : Tech Sector Stock

Index Shown: NASDAQ 100 tech Index (top companies by Market Cap)



Source:
https://www.msn.com/en-us/money/chart?id=a3yzqh&timeFrame=1M&chartType=baseline&projection=false

Post Election : Healthcare Sector Stock

Index Shown: S&P Healthcare Index (top companies by Market Cap)

DATE	OPEN	HIGH	LOW	CLOSE
11/18/2024	1,649.37	1,651.83	1,651.83	1,651.83
11/15/2024	1,677.58	1,651.81	1,651.81	1,651.81
11/14/2024	1,707.44	1,683.50	1,683.50	1,683.50
11/13/2024	1,715.10	1,709.79	1,709.79	1,709.79
11/12/2024	1,735.84	1,714.97	1,714.97	1,714.97
11/11/2024	1,741.74	1,738.27	1,738.27	1,738.27
11/08/2024	1,740.42	1,748.63	1,748.63	1,748.63
11/07/2024	1,731.21	1,736.46	1,736.46	1,736.45
11/06/2024	1,740.46	1,725.37	1,725.37	1,725.37

Source:
https://www.marketwatch.com/investing/index/sp500.35/download-data?startDate=11/6/2024&endDate=11/18/2024&countryCode=xx

Project Impetus

In a historic 2024 showdown, Trump vs. Harris became the defining election of the decade, setting the stage for political and economic winds that rippled through industries like healthcare and tech.

Background

- Election years are characterised by significant political and economic shifts, such as policy changes, fiscal priorities, and regulatory discussions.
- Healthcare and tech sectors are particularly sensitive to these shifts due to their reliance on government policy (e.g. healthcare reforms, tech regulations).

Use Case

- Understand how external political and economic winds affect the performance of the healthcare and tech sectors differently.
- Insights can guide policymaker strategies and help businesses mitigate risks or leverage opportunities during election cycles.
- Use sentiment indices and time-series models to forecast sector-specific trends for hedge funds, portfolio managers, and financial analysts.

Research Questions

Hypothesis: Political and economic sentiment influences these sectors, potentially in different ways.

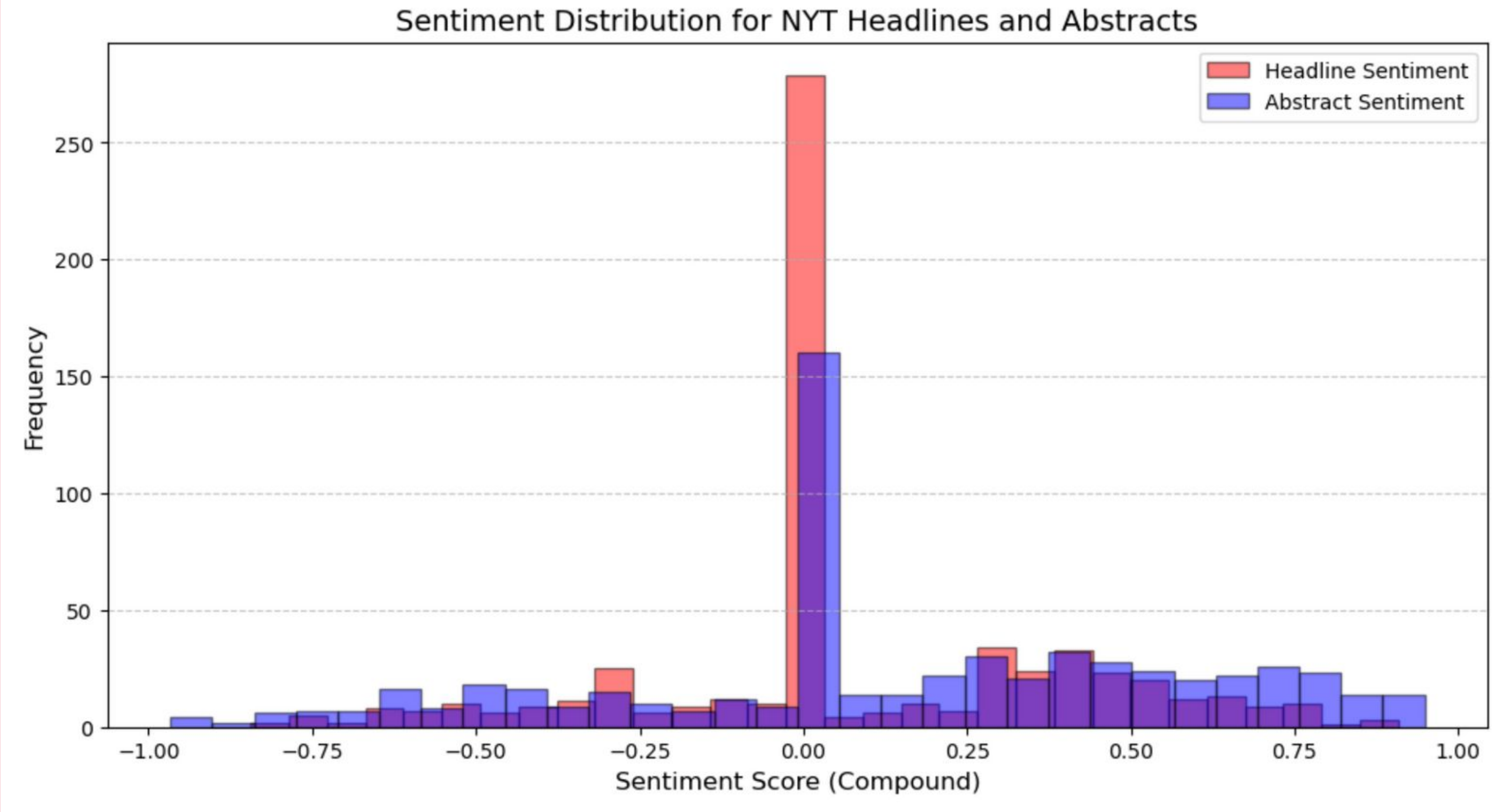
Key Questions

- Can the political and economic dynamics of an election year accurately forecast stock price trends in healthcare and tech sectors?
- Do these forces influence healthcare and tech stocks in similar ways, or do they drive them in opposing directions, **similar to post-election?**
- Is the stock movement in healthcare and tech sectors during election years a recurring pattern, or is it influenced by unique, election-specific factors?

Objectives

- Develop sentiment indices for healthcare/tech sectors and the election.
- Analyse predictive relationships using time-series analysis and modeling techniques.
- Compare sectoral reactions to political and economic sentiments.

Data Sources



NYT sentiments concentrate around 0 but skew slightly positive, reflecting neutrality and cautious optimism about how the market might perform in response to the dynamics of this election year.

Data Sources

New York Times API

- Collect recent articles and opinion pieces published about the election as well as various healthcare and tech sector stock fluctuations
- **library: requests**
  - HTTP client library for the Python programming language
- Free API allows access to headlines and article abstract.
- Created a sentiment lexicon using Vader library, assigns values between (-1,+1).

Data Sources

St. Louis FED

- Federal Reserve Economics Data (FRED) consists of thousands of economic data time series.
- Used to obtain Market Yield on T-Bills at 2/10 year constant maturity
- Nominal Treasury Yield(10 Year Treasury - "DGS10")

S&P 500 Health Care Index

- Tracks the performance of health care companies in the S&P 500.
- Considered “defensive” sector, performing relatively well during economic downturns but influenced by regulatory changes, drug approvals etc.

NASDAQ 100 Tech Index

- Represents the historical and real-time trading data of tech companies listed on NASDAQ .
- Reflects investor sentiment and market trends, especially during periods of economic or political uncertainty.

Data Sources

Polymarket

- Decentralised prediction market platform where users trade on the outcomes of real-world events, creating data reflecting collective market sentiment.
- Used to gauge public expectations and sentiment on events, aiding in decision-making for research, investments, and risk analysis.

Yahoo Finance (Treasure Data and VIX)

- Break Even Interest Rate (BEIR) as a measure of inflation
- Inflation Adjusted TIPS Yield ("DFII10")
- Data is pulled using the **R tidyquant package** and **Yahoo Finance API**, enabling streamlined and reproducible analysis.
- VIX (Volatility Index)

Data Exploration and Preparation

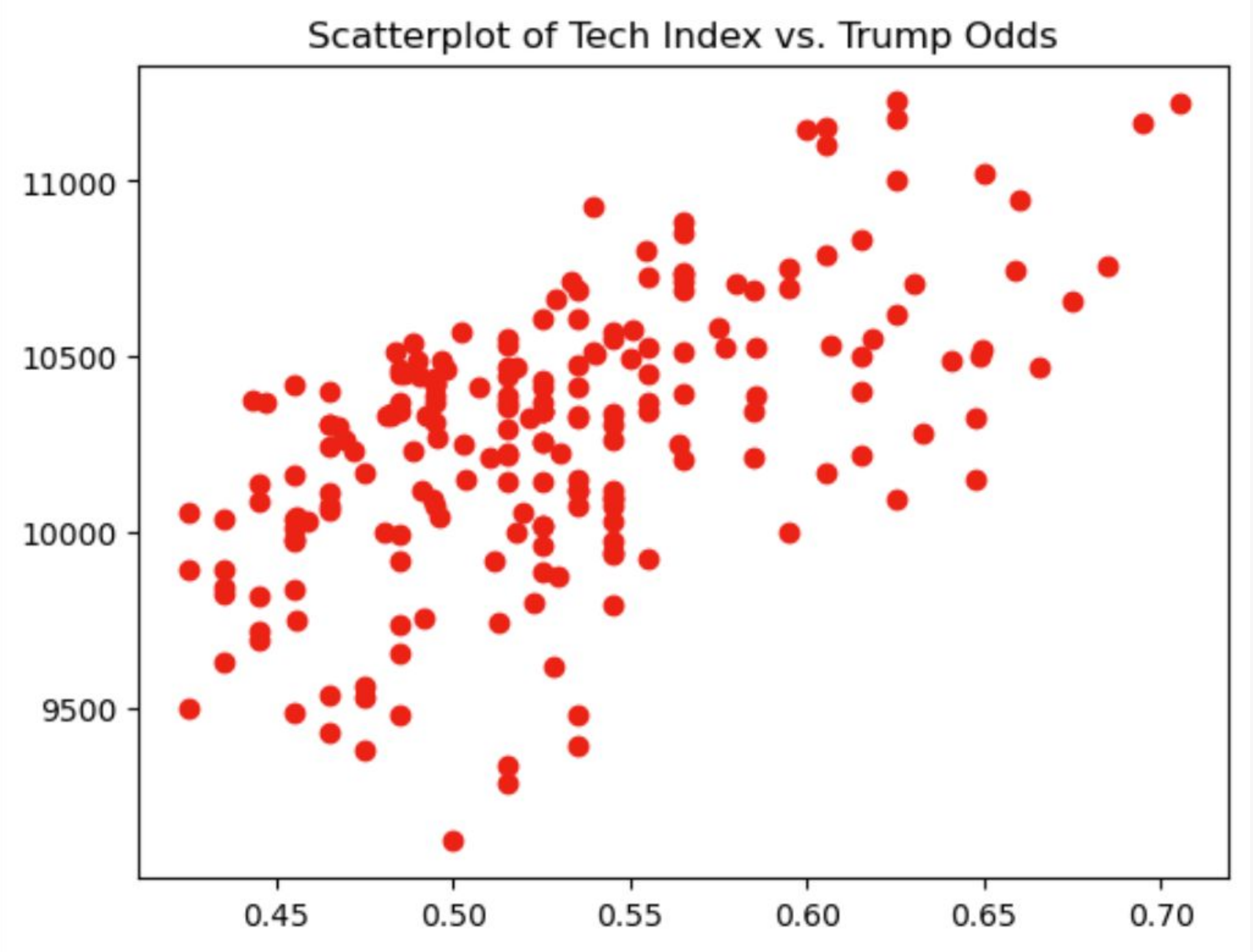
Variables Used

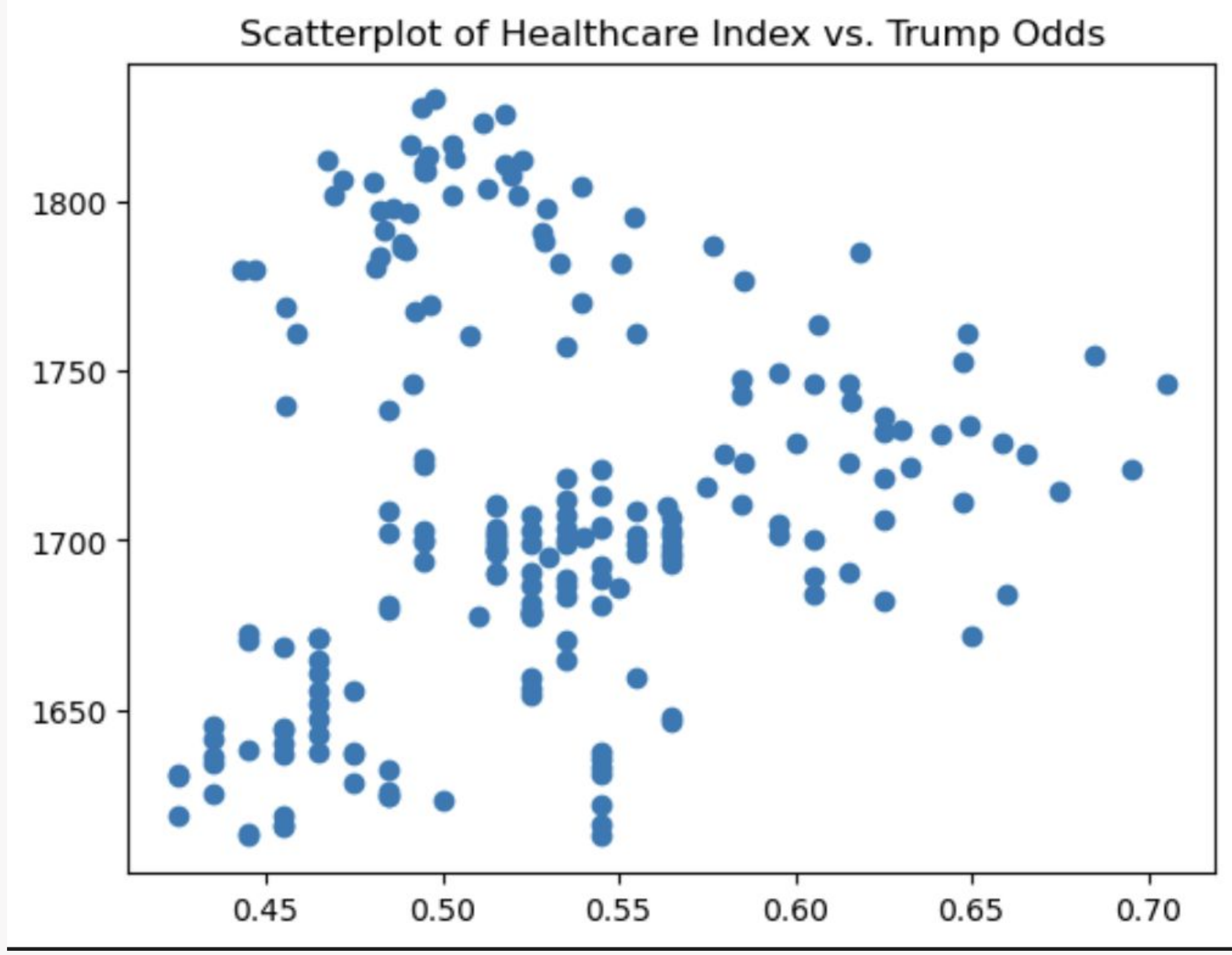
- **Trump Odds:** Political sentiment influencing tech and healthcare investor confidence.
- **BEIR** (Break-even Inflation Rate): Proxy for inflation expectations, shaping equity valuations.
- **VIX Index:** Market volatility gauge, reflecting uncertainty impact on stocks.
- **NASDAQ 100 Tech Index (Lagged):** Lagged values capture autoregressive behaviour in the tech index, past performance informs current trends
- **S&P Healthcare Index (Lagged):** Lagged values help capture autocorrelations in the series.

Gold prices, oil prices and Dollar Index(DXY), while usually very informative, were not included due to multicollinearity and the aim to increase the robustness of the model..

Data Exploration - Part 1: Correlation

	NASDAQ 100 Tech Index	S&P Healthcare Index	DGS2	DGS10	trump_odds	Dollar_Index	BEIR	VIX index
NASDAQ 100 Tech Index	1.000000	0.269505	0.113331	0.137348	0.578086	0.189297	0.136811	-0.327581
S&P Healthcare Index	0.269505	1.000000	-0.781413	-0.724043	0.182827	-0.684011	-0.686965	0.402273
DGS2	0.113331	-0.781413	1.000000	0.920299	-0.034817	0.868550	$\boxed{0.758589}$	-0.541470
DGS10	0.137348	-0.724043	0.920299	1.000000	0.032334	0.902646	$\boxed{0.877826}$	-0.359944
trump_odds	0.578086	0.182827	-0.034817	0.032334	1.000000	0.179487	0.014218	0.012486
Dollar_Index	0.189297	-0.684011	0.868550	0.902646	0.179487	1.000000	$\boxed{0.729285}$	-0.316556
BEIR	0.136811	-0.686965	0.758589	0.877826	0.014218	0.729285	1.000000	-0.332449
VIX index	-0.327581	0.402273	-0.541470	-0.359944	0.012486	-0.316556	-0.332449	1.000000

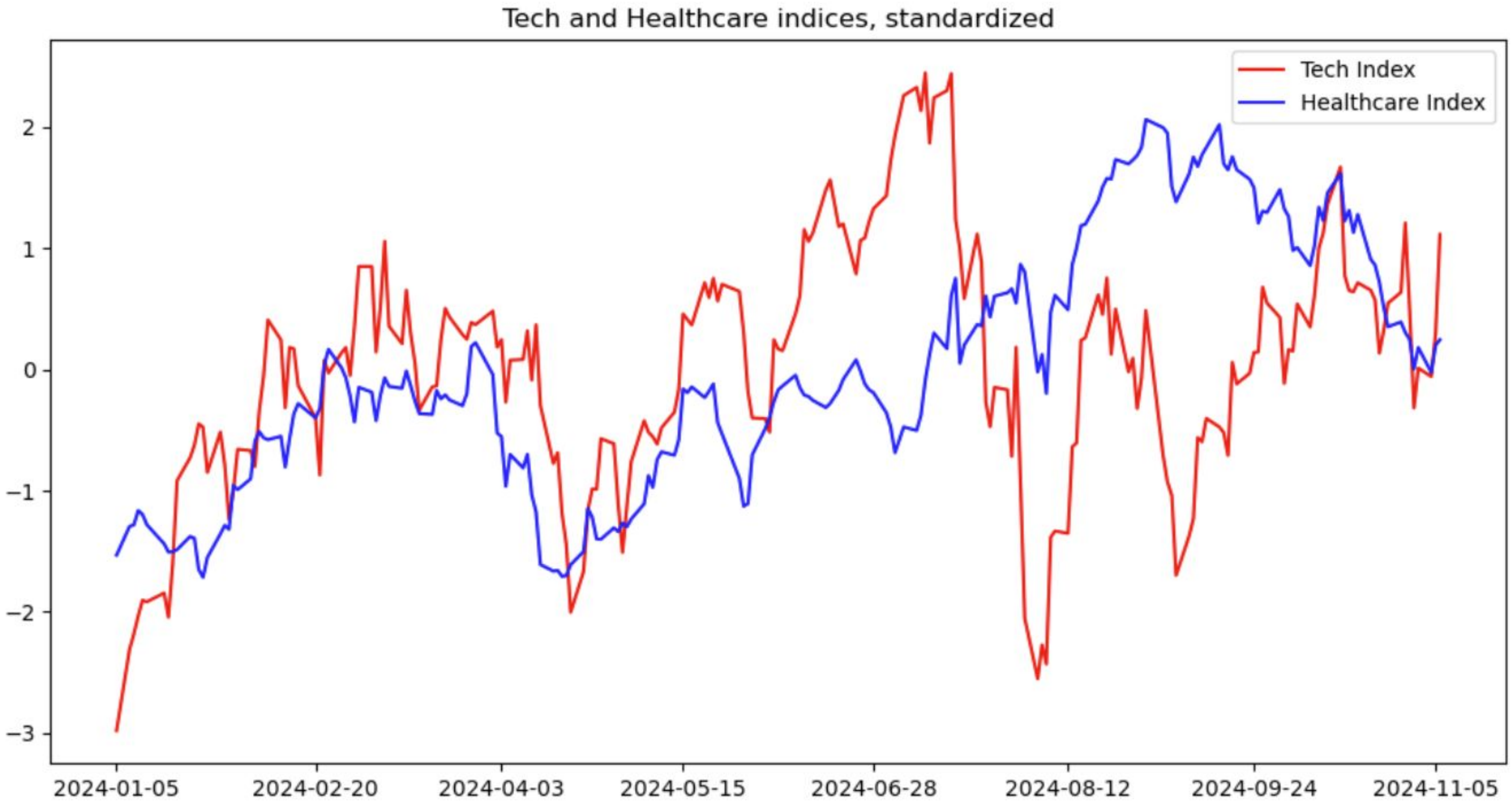




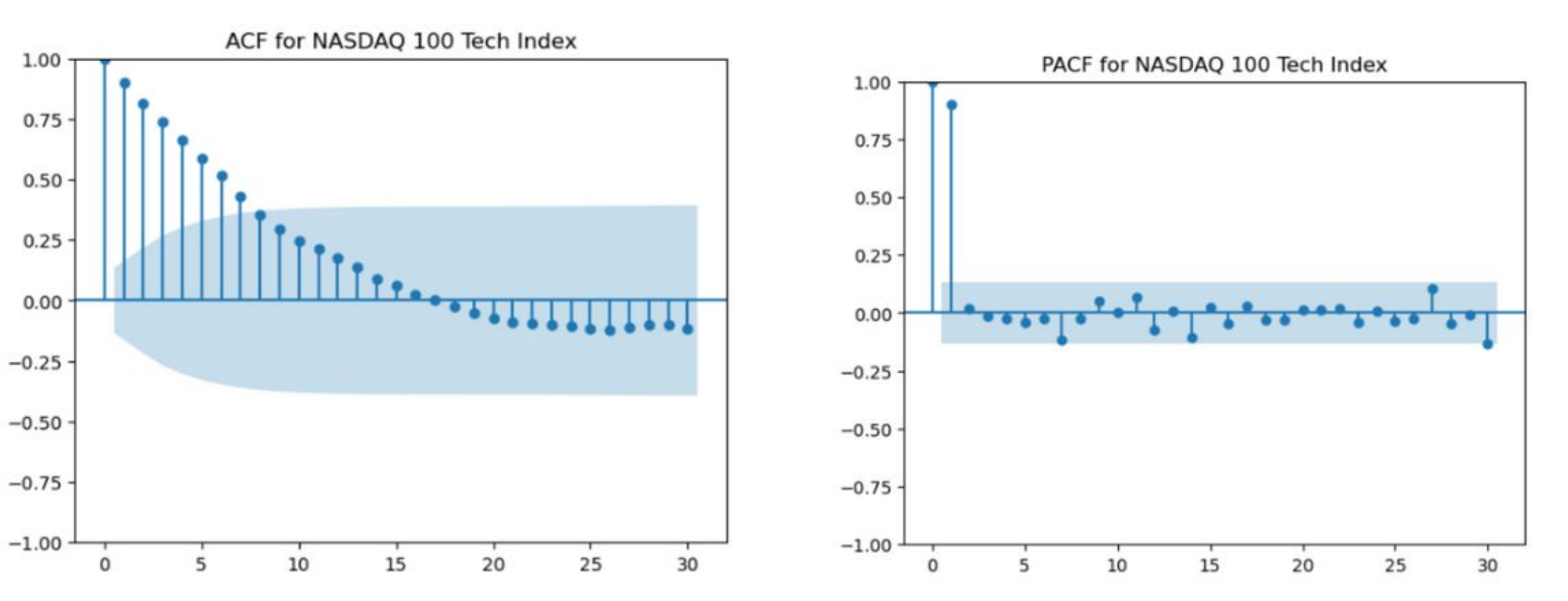
- Trump odds: strongly correlated with Tech Index
- BEIR: strongly (negatively) correlated with Healthcare Index
- DGS2, DGS10, Dollar_Index: Multicollinearity with BEIR

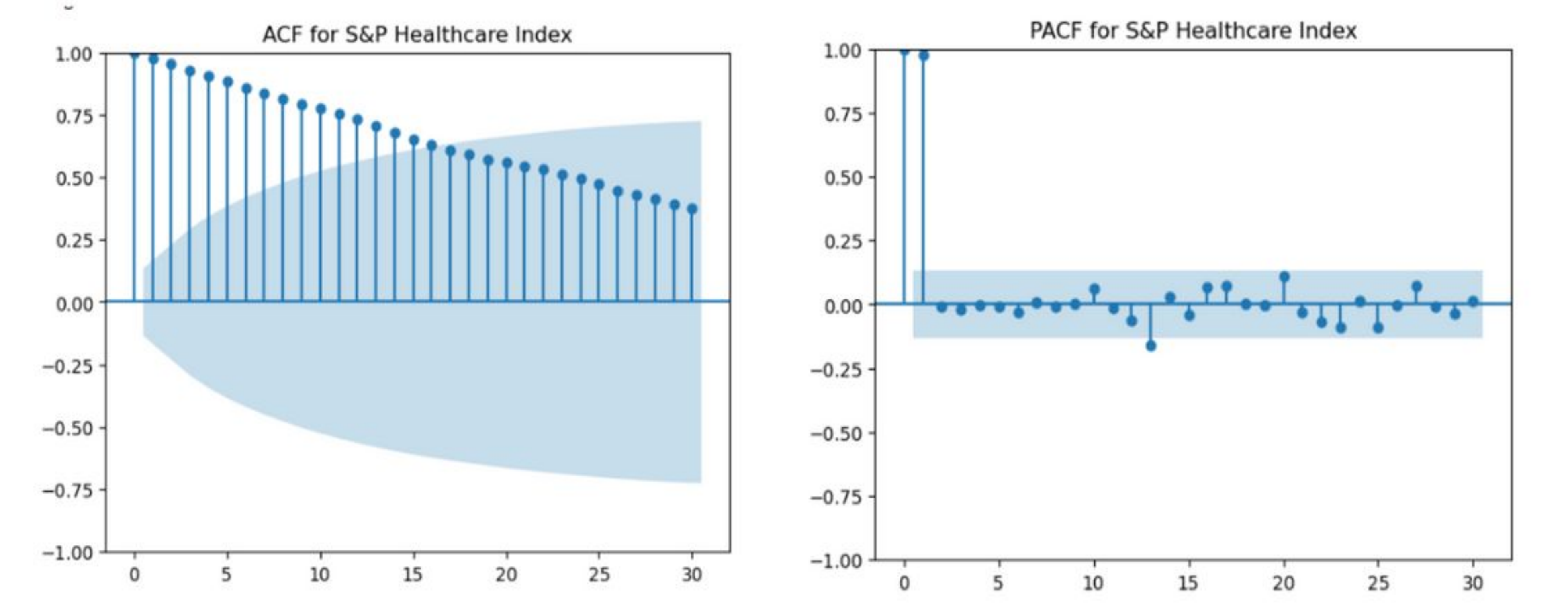
Data Exploration - Part 2: Stationarity

- **Test of Stationarity:** Ensures time series data has consistent statistical properties over time, a prerequisite for reliable modeling.



Data Exploration - Part 2: Tests for Stationarity





- **PACF Analysis**: Significant spike at lag 1 suggests an **AR(1)** process.
- **ACF Analysis**: Gradual decay after lag 1 suggests an **MA(1)** process.

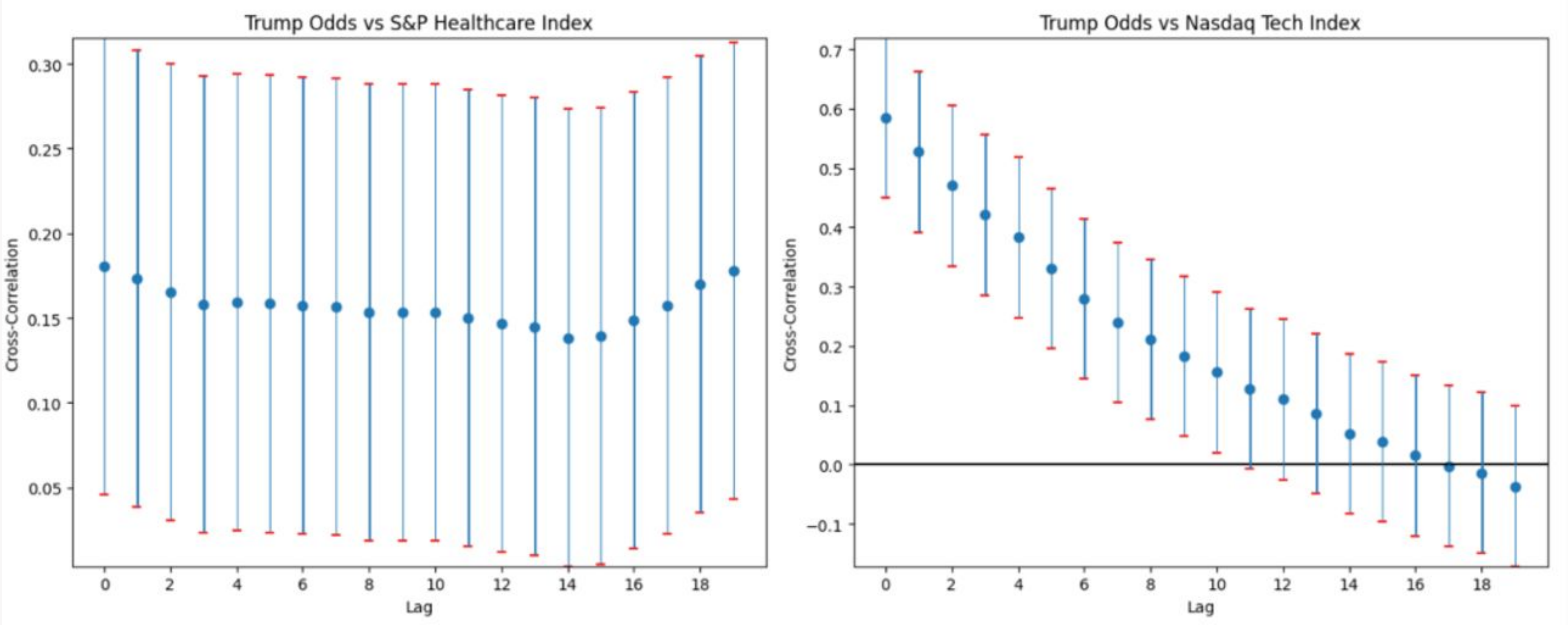
Variable	ADF Statistic	p-value	Stationary
NASDAQ 100 Tech Index	-3.6618	0.0047	Yes
S&P Healthcare Index	-1.8251	0.3681	No
Trump Odds	-3.1417	0.0236	Yes
BEIR	-1.8718	0.3454	No
VIX Index	-3.7813	0.0031	Yes

Engle-Granger Cointegration

- Data does not appear stationary, ADF tests conducted above for Stationarity
- For non-stationary series, cointegration tests suggested we can build a model without differencing with 10% level
  - Healthcare vs. BEIR Cointegration Test:
    P-value: 0.060418283035869075
  - Tech vs. BEIR Cointegration Test:
    P-value: 0.022161250196942484

Data Exploration Part 3: Cross-Correlograms

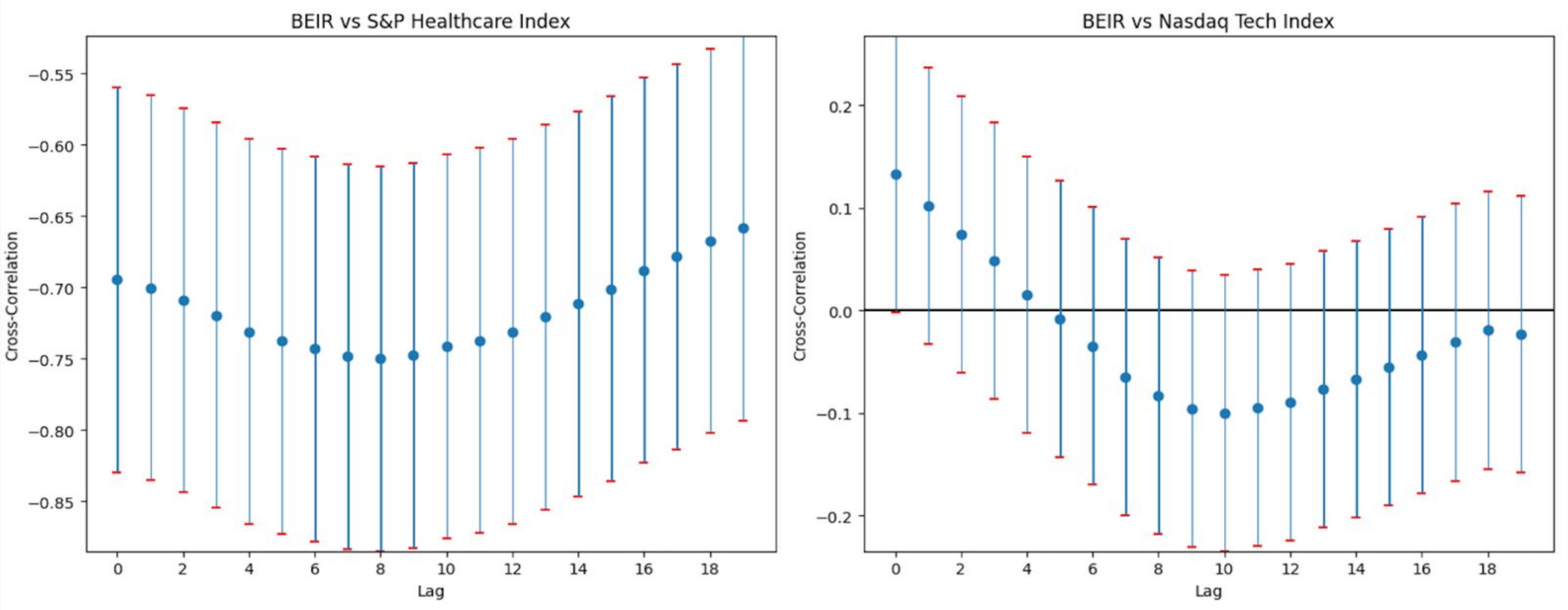
Trump Odds vs Indices



- S&P Healthcare Index: Moderate positive correlation across lags, showing steady influence.
- NASDAQ Tech Index: Strong positive correlation at initial lags, decaying gradually, suggesting immediate but diminishing impact.

Data Exploration Part 3: Cross-Correlograms

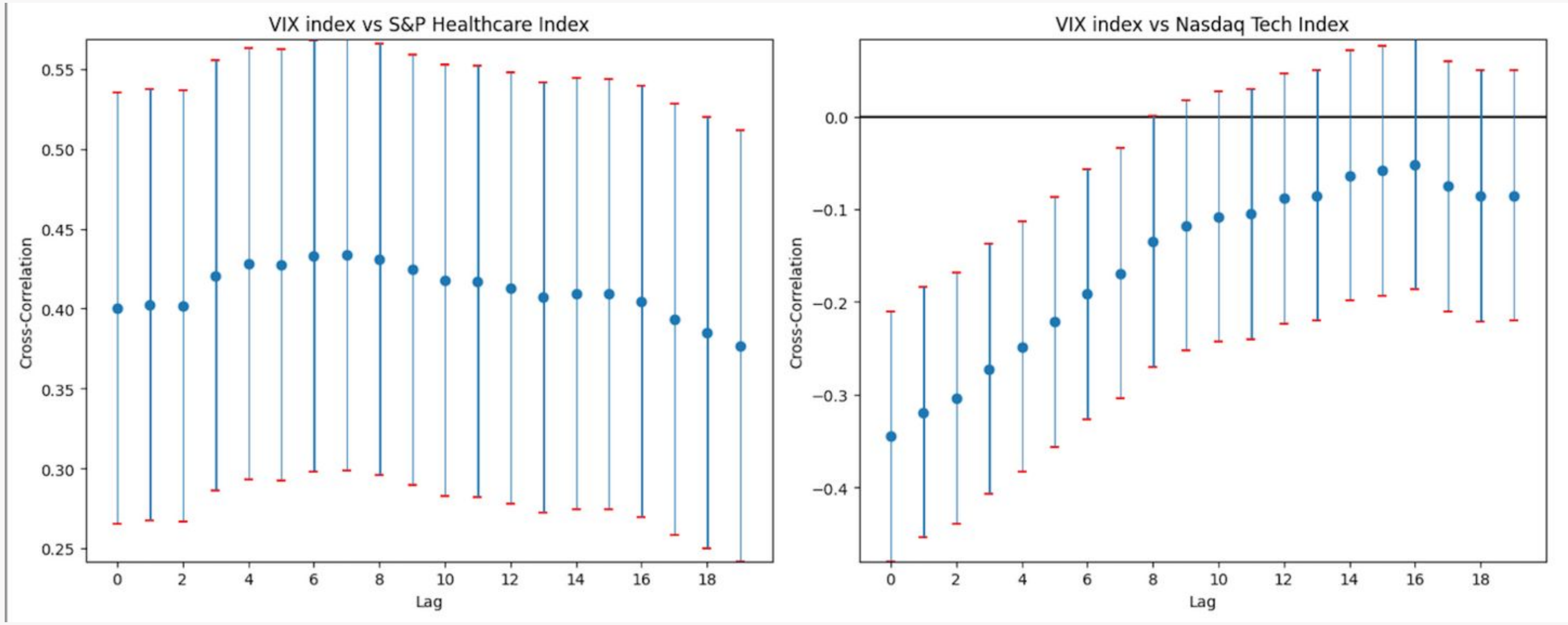
BEIR vs Indices



- S&P Healthcare Index: Strong negative correlation at all lags, with a consistent trend over time.
- NASDAQ Tech Index: Positive correlation initially, decaying gradually across lags.

Data Exploration Part 3: Cross-Correlograms

VIX Index vs Indices



- S&P Healthcare Index: Positive correlation across lags, indicating a consistent relationship with volatility.
- NASDAQ Tech Index: Negative correlation, with stronger effects at shorter lags, reflecting sensitivity to market uncertainty.

Data Preparation for Modeling

- **Train-Test Split**: 80%-20%
  - Train: Roughly 01-05-2024 - 09-05-2025 (exact dates vary due to use of lagged variables)
  - Test: Roughly 09-06-2025 - 11-05-2025
- **Standardization of variables**: For more stable results, transformation to mean 0, std. dev 1
  - Use of training data moments to standardize both (avoids *look-ahead bias*!)
- **Generation of lagged variables**: Use of pandas shift function for lagged versions of variables
  - Standardization BEFORE lagged variable generation to avoid loss of data point in generation of means and variances for standardization of lagged variables

Time Series Modeling

Baseline Model - ARIMA

Why ARIMA?

- Captures serial correlation and handles non-stationary data
- Able to includes AR, MA, white noise terms, and lagged variables (**Trump Odds**, **VIX**, **BEIR**) based on correlograms.
- Fit Separate Equations for Tech Index, Healthcare Index

Models to fit

$$TechIndex_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i TechIndex_{t-i} + w_t + \sum_{i=1}^{q} \beta_i w_{t-i} + \gamma * trumpOdds_{t-1} + \delta * VIX_{t-1}$$

$$HCIndex_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i HCIndex_{t-i} + w_t + \sum_{i=1}^{q} \beta_i w_{t-i} + \gamma * trumpOdds_{t-1} + \delta * BEIR_{t-1} + \zeta * VIX_{t-1}$$

ARIMA - Grid Search for optimal p,q

Grid Search specifics

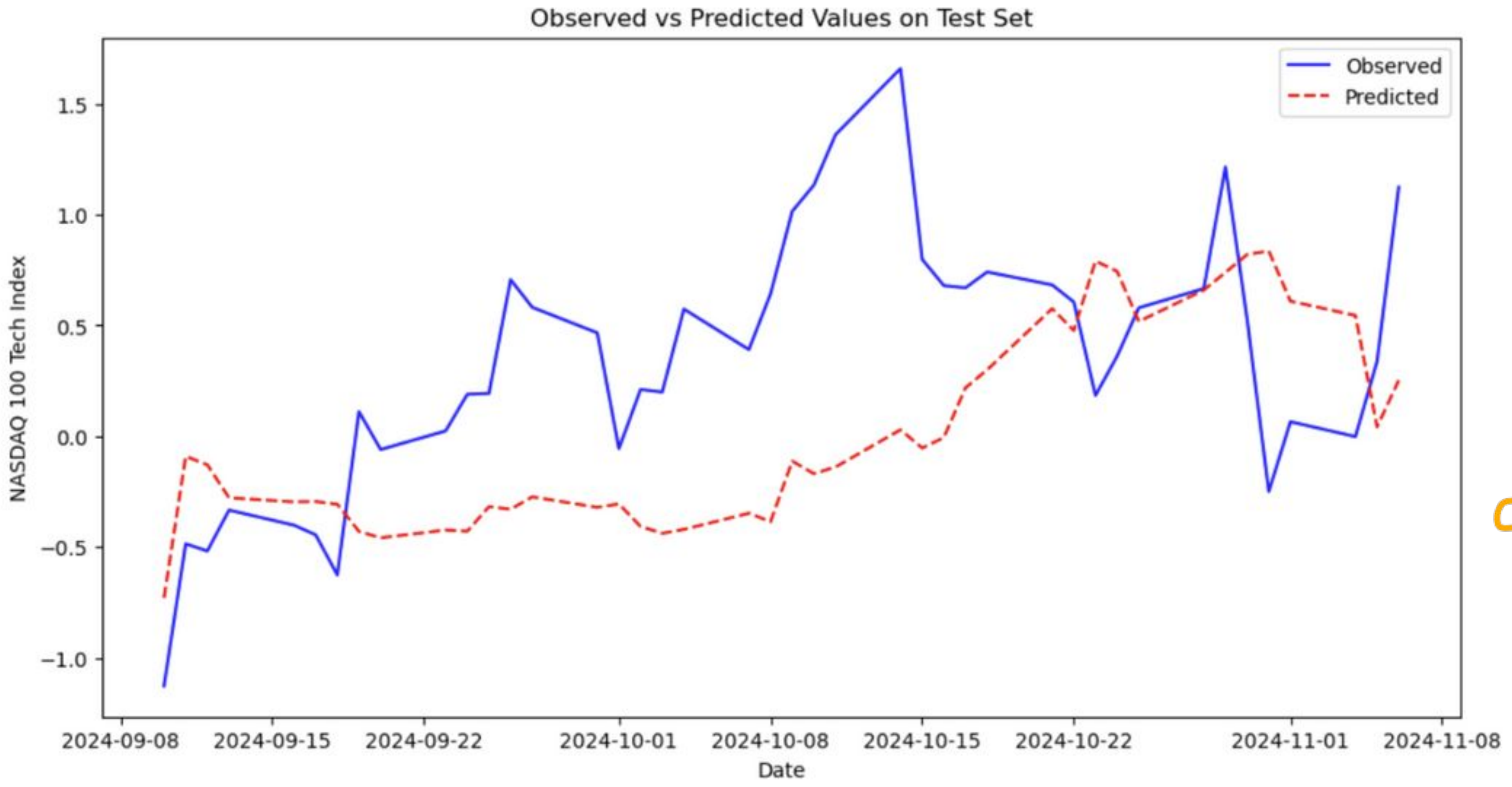
- Examined range of p’s from 0 to 2
- Examined range of q’s from 0 to 2
- Model Selection Criterion: MSE
  - “Test Set” treated more like “Validation set” in this case due to lack of data

```
max_AR = 2
max_MA = 2

model_sums = np.empty(shape = (max_AR + 1, max_MA + 1))
mse_vals = np.empty(shape = (max_AR + 1, max_MA + 1))
for i in range(0, max_AR + 1):
    for j in range(0, max_MA + 1):
        curr_model = arima_output(i,0,j, X_train, y_train, X_test, y_test)
        mse_vals[i,j] = curr_model[2]

opt_AR = np.where(mse_vals == np.min(mse_vals))[0][0]
opt_MA = np.where(mse_vals == np.min(mse_vals))[1][0]
print("Optimal p:", opt_AR)
print("Optimal q:", opt_MA)
print("MSE:", round(mse_vals[opt_AR, opt_MA],5))
```

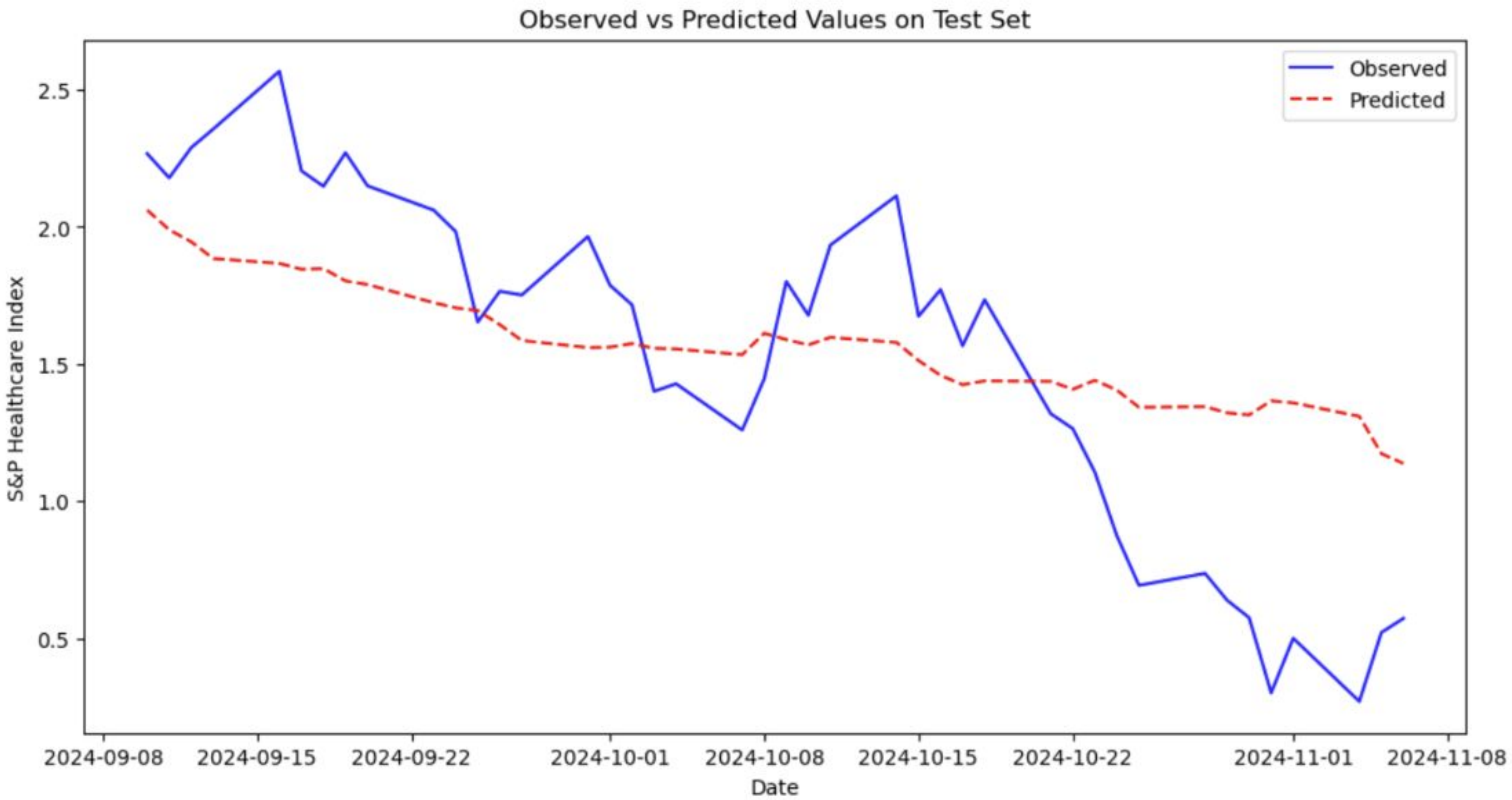
ARIMA Model Results - Tech Index




```
==============================================================================
Dep. Variable:     NASDAQ 100 Tech Index   No. Observations:                 168
Model:                    ARIMA(0, 0, 2)   Log Likelihood                -113.722
Date:                   Wed, 04 Dec 2024   AIC                            239.443
Time:                           19:37:35   BIC                            258.187
Sample:                                0   HQIC                           247.051
                                   - 168
Covariance Type:                     opg
==============================================================================
                   coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const            0.0001      0.087      0.002      0.999      -0.171      0.171
trump_odds_lag1  0.4148      0.090      4.597      0.000       0.238      0.592
VIX_lag1        -0.1125      0.066     -1.692      0.091      -0.243      0.018
ma.L1            0.8385      0.072     11.640      0.000       0.697      0.980
ma.L2            0.4628      0.072      6.412      0.000       0.321      0.604
sigma2           0.2255      0.024      9.247      0.000       0.178      0.273
==============================================================================
Ljung-Box (L1) (Q):                   3.50   Jarque-Bera (JB):               6.60
```

```
Optimal p: 0
Optimal q: 2
MSE: 0.49239
```

ARIMA Model Results – S&P Healthcare




```
                                SARIMAX Results
==============================================================================
Dep. Variable:     S&P Healthcare Index   No. Observations:                  168
Model:                   ARIMA(2, 0, 1)   Log Likelihood                  13.731
Date:                  Wed, 04 Dec 2024   AIC                            -11.463
Time:                          19:37:38   BIC                             13.529
Sample:                               0   HQIC                            -1.320
                                  - 168
Covariance Type:                    opg
=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const             0.2620      0.694      0.378      0.706      -1.097       1.621
trump_odds_lag1   0.0443      0.047      0.952      0.341      -0.047       0.135
BEIR_lag1         0.1203      0.046      2.593      0.010       0.029       0.211
VIX_lag1          0.0463      0.039      1.189      0.234      -0.030       0.123
ar.L1             1.0108      0.541      1.867      0.062      -0.050       2.072
ar.L2            -0.0341      0.537     -0.063      0.949      -1.086       1.018
ma.L1             0.1199      0.546      0.220      0.826      -0.950       1.190
sigma2            0.0488      0.004     11.434      0.000       0.040       0.057
```

```
Optimal p: 2
Optimal q: 1
MSE: 0.20724
```

Commentary on Results

- NASDAQ Tech Index:
  - Trump's implied odds of victory had a statistically significant positive effect on the Index
  - Non-zero average on the residuals
- S&P Healthcare Index:
  - AR_1 term with coefficient near one - model approximates a random walk
  - BEIR also significant, but with a positive coefficient - potential spurious relationship, as this disagrees with CCF

- Limitations of our ARIMA analysis:
  - Assumes linear relationships
  - requires separate models for Tech and Healthcare indices
  - We did not difference the response variables, and likely should have

Time Series and Deep Learning

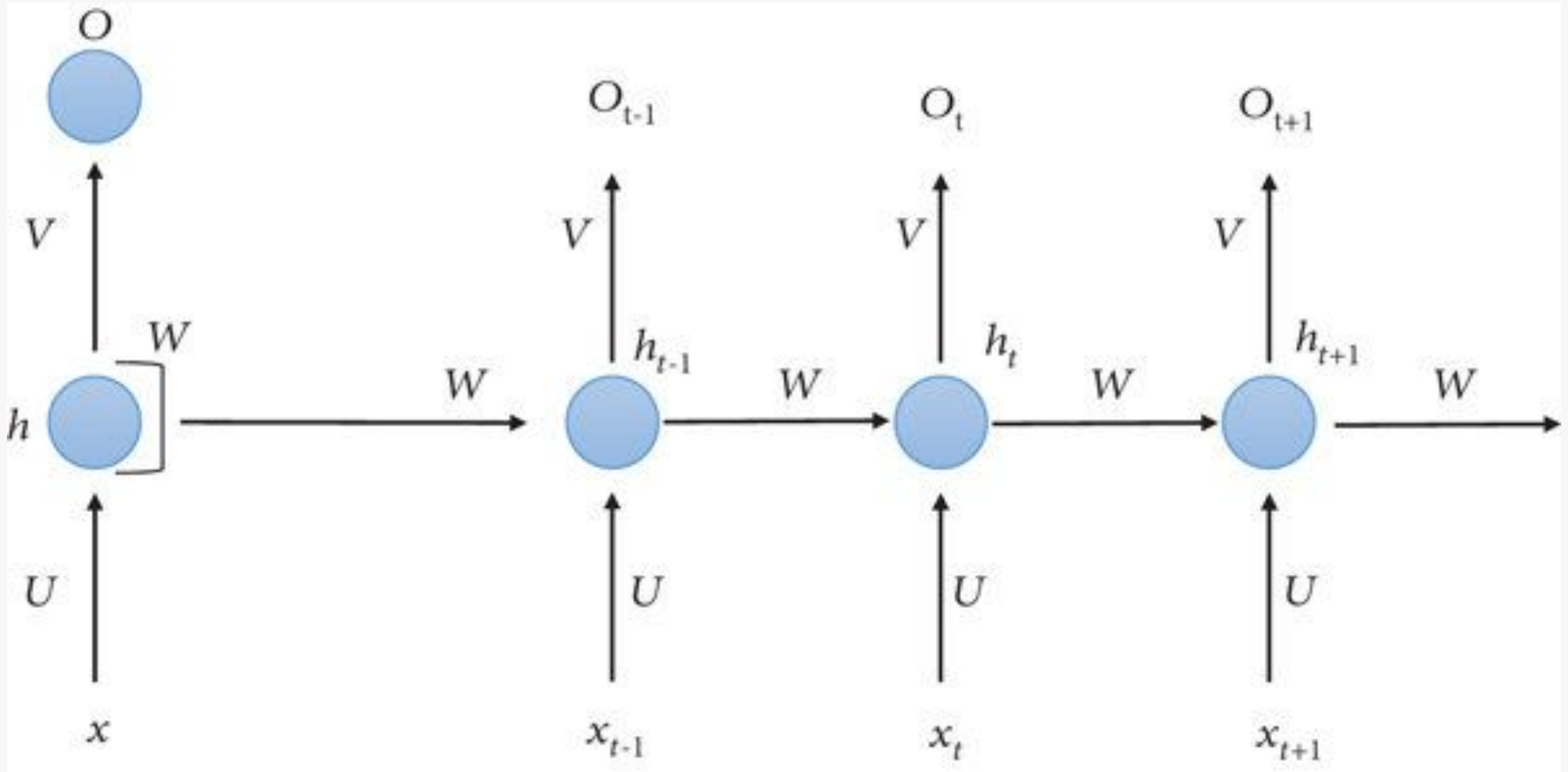
Deep Learning: RNN

Why RNN?

- Able to handle time series data
- Can handle non-stationarity
- Can model multiple output variables at once
- Can handle non-linear variable relationships

Models to fit

- Single Layer RNN
  - square matrix of inputs
  - inputs lagged versions of all variables (response variables and predictors)
  - Same lag for each variable for matrix mathematics to work



RNN Model: Hyperparameters and Specifics

Layer (type)	Output Shape	Param #
RNN_Layer (SimpleRNN)	(None, 8)	112
Output_Layer (Dense)	(None, 2)	18

Total params: 130 (520.00 B)

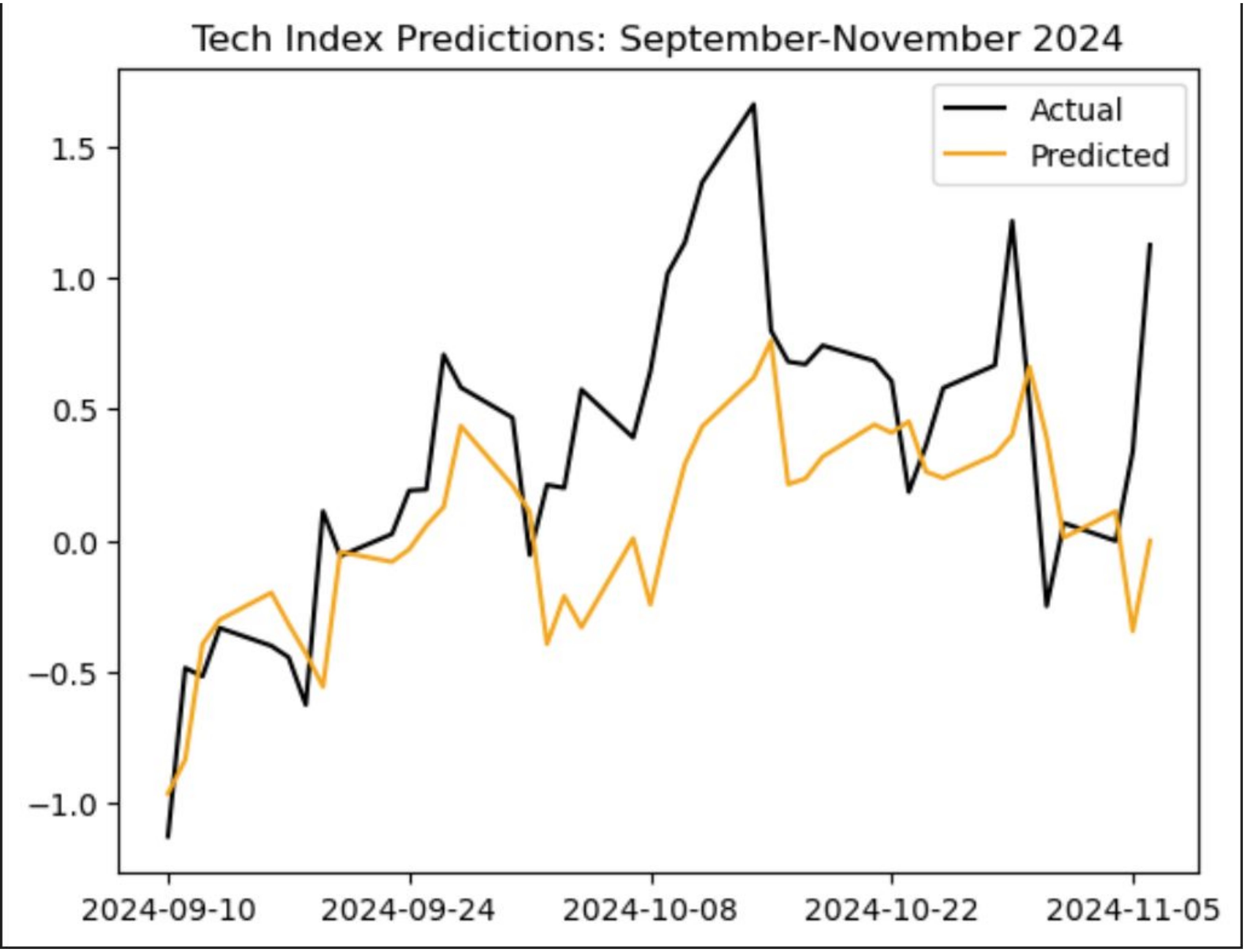
Trainable params: 130 (520.00 B)

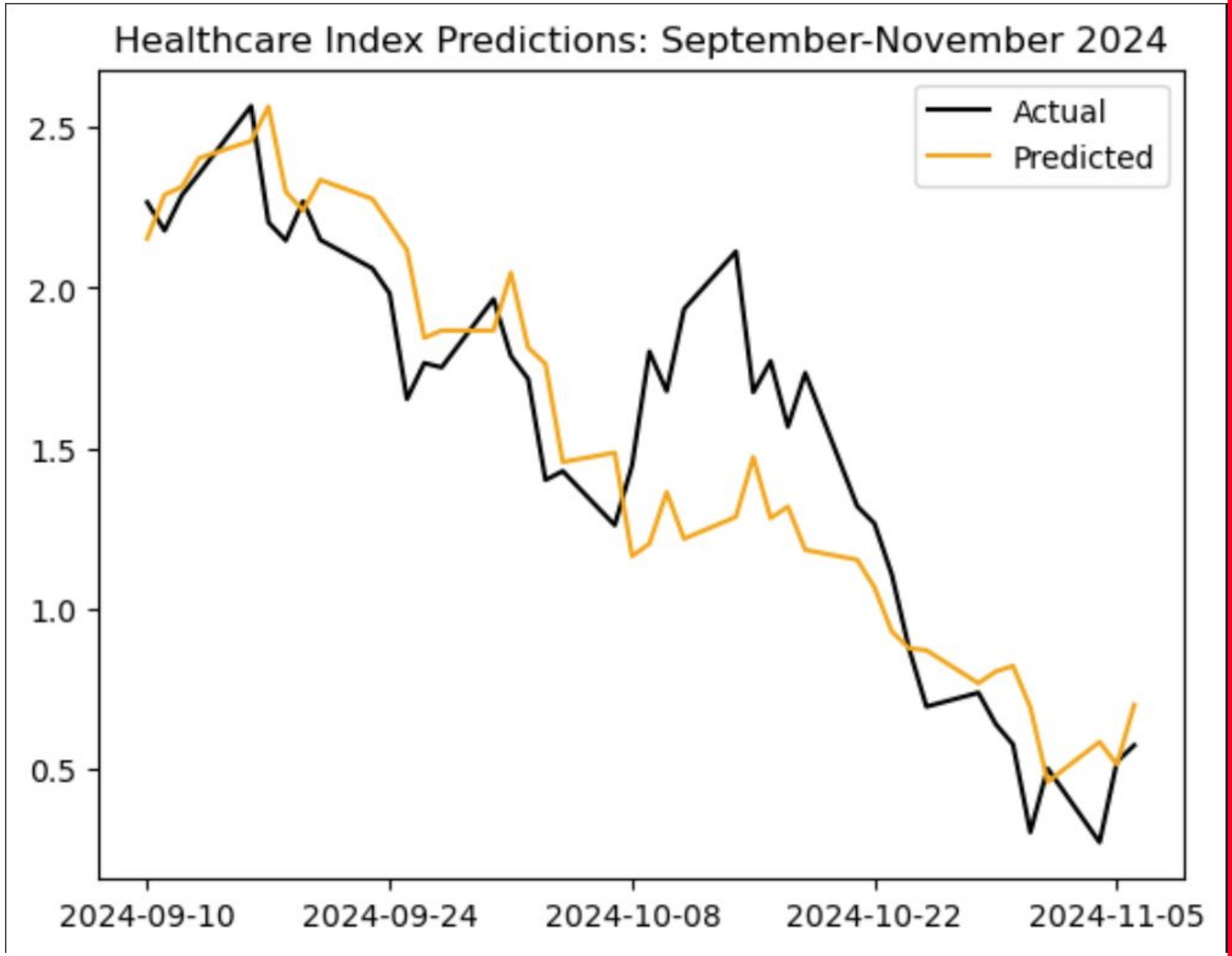
Non-trainable params: 0 (0.00 B)

- Inputs:
  - Tech Index (lag 1), Healthcare Index (lag 1), Trump Odds (lag 1), BEIR (lag 1), VIX Index (Lag 1)
- Outputs:
  - Tech Index and healthcare index (time t)

Hyperparameter	Value
Learning Rate	0.001
Number of Neurons (k)	8
Input Shape (time steps, features)	(1, 5)
Output Shape	2
Epochs	500
Early Stopping	Stop when 'loss' stops improving (save_best_only=True)

RNN Model Predictions





R-squared results:

	Tech Index	Healthcare Index
Time Series	-0.49809	0.477038
RNN	0.63553	0.876826

Commentary on Results

- NASDAQ Tech Index:
  - Network generalizes to test set effectively
  - Network not reactive enough to Trump surge in win probability in October
- S&P Healthcare Index:
  - Network generalizes to test set effectively
  - Relationship captured well
  - Could be leveraging relationship with BEIR (Healthcare is generally recession-proof)

- Limitations of our RNN analysis:
  - No time to hyperparameter tune for RNN
  - No examination of inclusion of further lags
  - A small dataset for the number of parameters used

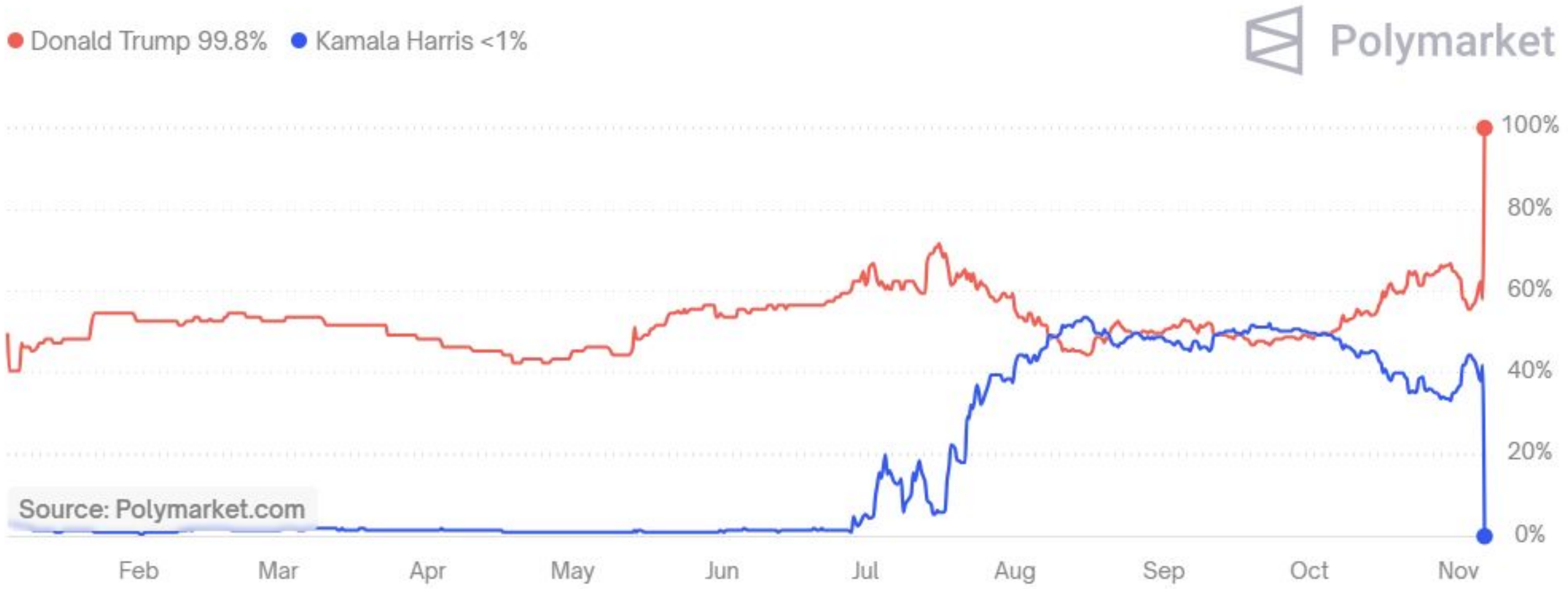
Conclusion and Next Steps

Conclusion

- *Do the Political Dynamics of the Election year effect the Healthcare and Tech sectors?*
  - Statistically significant coefficient for trump odds in time series modeling of Tech sector, reaction of RNN to bumps in Trump's implied odds
- *Do said dynamics impact the Healthcare and Tech sectors in different directions, as after the election?*
  - Trump's implied odds of victory do not seem as significant for the healthcare index, and if anything, slightly positive
- *Are these factors specific to this election, or part of a general election year trend?*
  - This is the only election with implied electoral victory odds from betting data

Limitations and Further Considerations

- *Limitations from Our Analysis*:
  - Only one year of implied Betting Odds data (stock price analysis typically over longer time frame)
  - Only election with publicly available Betting odds data (any trump/harris specific trends?)
- *Further considerations for Analysis*:
  - Hyperparameter tuning to improve both model's performances
  - Quantification of feature importance in RNN through interpretable machine learning



Thank You!

Election-Year Winds: How Politics can Shape Healthcare and Tech Stocks in 2024

PS5842 | Dennis Goldenberg, Samaa Nadkarni, Jiawen Shao