University of Westminster

School of Finance and Accounting

MSc FinTech with Business Analytics

7FNCE045W.Y Project

Do Elon Musk's Tweets Move Tesla? Evidence from Daily Models and High-Frequency Event Study

Name: J Daniel Martin

Student ID: w1834001

Submission Date: 04th of September 2025

Word Count: 9904

Abstract

This dissertation investigates the impact of Elon Musk's tweets on Tesla's share price and trading activity. It employs a daily regression framework alongside a high-frequency event study to differentiate between longer-term effects and immediate intraday responses. Tweets are analysed using VADER sentiment analysis, with replies and short posts filtered out to minimise noise.

At the daily level, while associations are weak in the initial data, filtering reveals a statistically significant but economically small relationship between sentiment and returns. The coefficients range between 0.017 and 0.019, with a modest model fit ($R^2 \approx 0.02$–0.04). Tweet count does not show predictive ability. In contrast, the event study indicates notable intraday effects, with positive tweets leading to significant price increases; for instance, cumulative average returns (CAR) can rise to approximately 0.42% within 15-minute windows and about 0.89% in 60-minute windows. Negative tweets appear to have no notable impact. Volume spikes coincide with these events, particularly at market open, highlighting market reactions.

This study offers a comprehensive analysis of the effects of Musk's tweets on Tesla, revealing that while daily data may underestimate impact, intraday analysis captures significant, brief reactions following positive tweets. Limitations include potential sentiment misclassification and the influence of concurrent news. The findings suggest that close monitoring of intraday movements and tweet timing is crucial for evaluating the market impact of corporate communications.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

Can a single tweet from an individual like Elon Musk influence a multi-billion-dollar stock like Tesla? This dissertation explores this question by analysing whether Musk's tweets impact Tesla's returns and volatility, and if tweet sentiment can predict short-term market movement or trading activity.

The study navigates the complexities of sentiment, visibility, and market dynamics. It employs a two-tier approach: first, a daily analysis through regression models to assess the relationship between tweet sentiment and daily returns, controlling for market conditions; second, a high-frequency event study that examines price and volume changes immediately following each tweet.

Focusing on Musk's tweets from 2018 to 2023, the analysis minimises the noise from less significant interactions by filtering tweets into distinct datasets. Various metrics are used to evaluate sentiment, and the study aligns tweet timestamps with US market hours. The findings utilise ordinary least squares for daily models and cumulative average returns for the event study to understand the impact of Musk's social media activity on Tesla's stock performance.

Three research questions guide the analysis:

1. Do Elon Musk’s tweets influence Tesla’s stock price and volatility?
2. Does tweet sentiment predict short-term returns or trading volume?
3. Are effects asymmetric across positive and negative tweets?

First, it bridges daily and high-frequency perspectives within a single framework, showing how conclusions depend on the time scale.
Secondly, it separates content from noise by filtering replies and very short posts, aligning the sentiment measurement with tweets more likely to be informative to the market.

Thirdly, it explicitly incorporates attention and market microstructure by normalising volume around events and discussing how the treatment of out-of-hours tweets interacts with predictable opening-auction dynamics.

The results indicate that daily correlations between returns and sentiment are generally weak in unfiltered data, but they become more significant when excluding replies and very short tweets. In regression analyses, sentiment shows a slight but notable impact only in these filtered datasets, with better model fit when including volatility and volume controls, though overall explanatory power remains limited.

At an intraday level, the findings are clearer: positive tweets lead to significant price increases within the trading day. In contrast, neutral tweets have lesser effects, and negative tweets show no consistent impact. This suggests that markets respond more favourably to positive sentiment than negative or ambiguous messages. There are noticeable volume spikes around tweet times, but these can be complicated by trading patterns at the market open, highlighting the need for precise timing in analyses.

However, the study has its limitations. Sentiment classification may miss nuances like irony and somewhat subjective filtering criteria. Daily regressions might be affected by other news on the same day, and aligning tweets from outside market hours with trading sessions adds complexity. These issues are addressed through clear design choices, consistency filtering, and supplementary intraday data.

This dissertation explores a straightforward yet significant proposition: although no single individual can completely "control" a stock or the market, a prominent communicator can trigger short-lived and asymmetric price reactions with real economic impact within a single trading day. By integrating daily models with high-frequency data, the study examines the timing, methods, and magnitude of how Musk’s tweets influence Tesla’s stock price, distinguishing the genuine sentiment from attention distractions and the trading day’s structure.

2. Literature review

2.1 Social Media, Investor Sentiment, and Markets

Recent research has increasingly recognised social platforms as significant economic information channels. Early studies at a macro level indicated that overall Twitter mood could align with or even forecast major market indices. For instance, Bollen, Mao, and Zeng (2011) found that the "calm" dimension of public sentiment extracted from millions of tweets could predict future movements in the Dow Jones Industrial Average. This suggests a broader economy-wide sentiment channel rather than relying solely on firm-specific news.

Subsequent research has focused on signals at the firm level, positing that stock-specific microblogs convey information that is not fully reflected in market prices in real-time. Sprenger et al. (2014) analysed approximately 250,000 stock-tagged tweets and observed systematic relationships between tweet bullishness and stock returns, message volume and trading volume, and disagreement and volatility. They also identified how mechanisms of information diffusion, such as followers and retweets, amplify the influence of high-quality contributors.

Beyond immediate co-movement, several studies have explored the predictive capabilities of this information. Sul, Dennis, and Yuan (2017) examined 2.5 million tweets related to S&P 500 firms and discovered that sentiment from users with lower follower counts, who did not retweet, could predict returns 10 to 20 trading days in advance. This slower-diffusing opinion led to meaningful trading gains, while faster-diffusing sentiment appeared less predictive, aligning with the idea that it gets incorporated into prices more quickly. Additionally, research by Cookson et al. (2025) highlighted how Twitter heightened investor attention and withdrawal during the Silicon Valley Bank run, revealing that social amplification could contribute to price pressure and funding stress. This underscores that attention and sentiment polarity dynamics can significantly impact market outcomes.

These findings emphasise the importance of distinguishing sentiment, attention, and diffusion speed, as each may influence market prices over different time horizons.

2.2 Evidence from Events and Specific Companies

Focusing on specific events, one notable study investigates Twitter's effects during significant occurrences. Ranco et al. (2015) noted that during heightened tweet activity for publicly traded companies, negative sentiment correlated with abnormal returns and temporary mispricing. However, the effect sizes were relatively modest. Further research by Gu and Kurov (2020) reinforced that firm-specific Twitter sentiment offers insights beyond traditional analyst forecasts and has predictive value for returns.

Crucially, Sprenger et al. (2014) cautioned that daily aggregation of tweets could obscure timing. For example, aligning tweets posted after 16:00 to the next trading day can mix same-day and next-day information, potentially complicating causal interpretations. Their observations also aligned with the idea that contemporaneous relationships are often stronger than lagged ones, with bullish sentiment showing a more immediate influence on returns. In contrast, predictive elements might emerge from slower-diffusing sentiment or high-frequency analysis conducted shortly after posting tweets.

2.3 High-Frequency Impacts: Tweets as Immediate Disruptions.

Recent studies utilising high-frequency data have enhanced the identification of price responses by correlating minute-level price changes with tweets. In a study by Yang, Fernandez-Perez, and Indriawan (2024) published in The Financial Review, a specific high-frequency framework demonstrates how tweets can lead to immediate price impacts, followed by a partial reversal over the following days. This phenomenon suggests a pattern of short-term mispricing that corrects itself as the market digests the information.

This approach aligns with traditional event-study methodologies at an intraday level: it involves synchronising the “event time” with the first bar occurring after a tweet, calculating returns and accumulating these returns across a designated event window to form Cumulative Average Return (CAR) profiles.

In a related vein, machine-learning studies have examined the predictive potential of tweet attributes. For instance, Karlemstrand and Leckström (2021) found that integrating sentiment analysis with metadata, such as likes, followers, retweets, and verification statuses, enhances stock forecasting, highlighting the significance of attention and credibility factors. Similarly, using machine learning pipelines, Mokhtari et al. (2023) identified strong correlations between social media sentiment and equity trends. Milikich and Johnson (2023) introduced Taureau, a modular inference framework where company-specific Twitter sentiment is a predictive element.

Collectively, these studies convey a coherent narrative: immediate price responses can be detected with high frequency; the effectiveness of predictive signals increases when tweets are informative, attention-grabbing, or diffuse over time; and incorporating metadata assists in distinguishing between noise and actionable content in trading.

2.4 Methods: Event Study Logic and Measurement Choices

As MacKinlay (1997) outlined, the standard event-study framework guides modern tweet-event analyses. Researchers define an event window, calculate returns for each period. These returns are then accumulated to generate cumulative average returns (CARs). Statistical inference typically involves cross-sectional t-tests and, when needed, heteroskedasticity-robust errors. Daily studies often employ volatility controls, using range-based estimators like Garman–Klass (1980) for more accurate daily volatility measures.

Two main technical challenges arise. First, the timing of tweets is critical, as mapping post-close tweets to the next trading day can lead to misinterpretations due to market-open dynamics. Sprenger et al. emphasise the importance of considering intraday data for more accurate relationships. Second, extracting sentiment from tweets poses challenges. At the same time, dictionary-based tools (e.g., VADER) are standard for short texts, and machine learning approaches can better capture context but require labelled data and validation. The literature showcases both methods: broad dictionaries facilitate scalability, whereas ML approaches offer more profound insights.

2.5 What Influences Prices: Sentiment, Attention, or Both?

Distinguishing sentiment (polarity) from attention (salience) is essential. Research by Gu and Kurov (2020) shows that firm-specific sentiment offers insights beyond analysts' signals. Spikes in message volume and diffusion networks also play an essential role. Sprenger et al. found that message volume closely correlates with trading volume ($r \approx 0.44$), and that retweet structures amplify voices from higher-quality contributors, acting as an attention-weighting mechanism.

Furthermore, Sul et al. indicate that slow-diffusing sentiment from low-follower accounts can predict multi-day returns, suggesting that attention gradually reaches marginal investors. Cookson et al.'s findings from a modern bank run illustrate that attention propagation can overshadow sentiment during stressful times and quickly coordinate investor actions.

These insights apply to CEO-focused Twitter activity. For example, Elon Musk's posts blend strong sentiment with high attention and media amplification, leading to immediate market responses, often rewarding positive tweets more than penalising negative ones.

2.6 Asymmetry and Horizon: Positive Versus Negative News

Asymmetric effects have been noted in various studies. Ranco et al. found that peaks in negative tweet sentiment correlate with small abnormal returns. Sprenger et al. observed that bullish sentiment correlates more strongly with market reactions than message volume, with quick market responses to positive news and weaker reactions to negative information.

Yang et al. noted rapid intraday impacts from tweets, often followed by partial reversals, indicating that mispricing is likely temporary. Sul et al. found that particular slowly diffusing sentiments can retain predictive value over several days.

The findings suggest that while positive tweets may quickly influence prices, sustained price movements are more likely when sentiment diffuses slowly or attention remains high.

2.7 Limits, Identification Challenges, and Design Implications

The literature outlines several significant challenges that can impact the effectiveness of robust design:

- Daily Aggregation and Market-Open Confounds: Assigning after-hours tweets to the next trading day can confuse their effects with trading volume spikes at market open. Aligning tweet data with the first trading bar after the tweet is more effective, though it requires complex data handling.

- Multiple Testing and Small Effects: Small coefficients can be statistically significant in extensive studies but lack economic impact. Best practice includes reporting economic magnitudes, like CARs in basis points and p-values.

- Omitted Variables and Co-Incident News: Events such as earnings announcements or product launches can coincide with tweet bursts, skewing results. Controls, as shown by Gu and Kurov for analyst forecasts, can reduce this overlap.

- Sentiment Measurement Error: Sarcasm and irony in CEO tweets can lead to inaccuracies in sentiment analysis. More sophisticated models and context features—like account credibility and engagement—are needed for better accuracy.

- Attention Vs. Sentiment: It's crucial to differentiate between sentiment and visibility; attention can influence outcomes, such as in a bank run, even if sentiment is neutral.

2.8 Positioning and Contribution

In this context, the dissertation presents three key contributions.

First, it integrates daily and high-frequency analysis by focusing on a significant speaker, Elon Musk. This is done through a thorough daily Ordinary Least Squares (OLS) analysis that incorporates controls for volatility and liquidity, in addition to an intraday tweet-event study aligned with the first observable trading activity after a tweet. This approach effectively addresses the timing concerns raised by Sprenger et al. and the mispricing dynamics identified by Yang et al.

Second, the research distinguishes between meaningful content and noise by filtering out replies and short messages. This is consistent with previous findings suggesting that higher-quality or more informative posts correlate more strongly with market returns (as noted by Sprenger et al. and Gu & Kurov).

Finally, the dissertation considers the attention channel by normalising trading volume and analysing the timing of tweets around market openings. This connection links the firm-level findings to broader insights on social amplification, as Cookson et al. (2025) discussed.

2.9 Summary of the Methodological Foundations

The study uses the event-study methodology by MacKinlay (1997) to analyse returns and define an event window for each tweet, allowing for the calculation of CAR and cross-sectional means. Daily volatility is assessed using the Garman–Klass model to enhance efficiency. A lexicon-based sentiment analysis tool, supported by finance and social media literature, is selected for its effectiveness in analysing short tweets. Additionally, metadata like likes and retweets are utilised as attention proxies, in line with the diffusion mechanisms outlined by Sprenger et al. and the predictability of slower-diffusing sentiment discussed by Sul et al.

2.10 Synthesis

The literature supports four key propositions for this research:

1. Twitter contains tradable firm-level information, but its average effects are minor and vary by message type and context (Sprenger et al.; Gu & Kurov; Ranco et al.).

2. Timing is essential: high-frequency data around tweets can reveal immediate price impacts and short-lived drifts that daily aggregates may miss (Yang et al.; Sprenger et al.).

3. The speed of attention and diffusion affects predictability: slowly diffusing sentiment can forecast multi-day returns (Sul et al.), while sudden attention can align market movements without strong sentiment polarity (Cookson et al.).

4. Methodological rigour helps reduce confounding factors: using cumulative average returns (CARs) with robust error adjustments, and careful tweet timing strengthens causal interpretations (MacKinlay; Garman–Klass; Sprenger et al.).

This dissertation applies these principles using daily regressions and intraday event studies to assess how Elon Musk's tweets influence Tesla's price and volatility, examining the symmetry of effects between positive and negative tweets.

3. Data

3.1 Data Sources

This dissertation combines social media and financial market data to study the impact of Elon Musk's Twitter activity on Tesla's stock. Three datasets are used.

- Tweets. The primary dataset is from Singh, A. (2023) *Elon Musk Tweets (Daily Updated)*, a collection hosted on Kaggle, covering all of Musk's public tweets between 2010 and 2023. The dataset contains tweet text, timestamps, tweet ID and Username. This dataset is in UTC, which is kept to ensure consistency.
- Daily stock data. Tesla daily open–high–low–close–volume (OHLCV) data were obtained from Yahoo Finance for the same period, 2010–2023. This source provides freely accessible and widely used market data for replicable academic analysis.
- Intraday stock data. Tesla high-frequency OHLCV data at 15-minute and 60-minute intervals were retrieved from the Bloomberg Terminal, covering 2018–2023. Even if Tesla is a US company, the timestamp of the dataset is in UTC.

A focus on the 2018–2023 period is adopted for the primary analysis. Before 2018, Musk's Twitter activity was relatively low and emotional, and Tesla's market capitalisation was relatively modest.

3.2 Pre-processing

When preparing the data for analysis, several key steps were taken to ensure everything was in order.

First, the timestamps were checked for consistency. The Kaggle dataset recorded tweets in Coordinated Universal Time (UTC), and it was essential to ensure all timestamps were explicitly set to UTC. Trading hours were defined from 14:30 to 21:00 UTC to align with US market hours. Tweets posted after market close were mapped to the next trading day to capture their potential impact during the upcoming session.

Next, sentiment analysis was conducted using the VADER algorithm, which is well-suited for short texts like tweets (Hutto and Gilbert, 2014). VADER provides a compound score ranging from -1 to +1. Tweets were categorised based on standard thresholds: scores over 0.05 were classified as positive, scores below -0.05 as negative, and scores between as neutral. Figure 1 shows the distribution of tweet lengths in the dataset, indicating that most messages are concise, typically containing fewer than 20 words. This supports the suitability of VADER, as traditional machine learning or deep learning models usually require longer text inputs to capture context effectively.

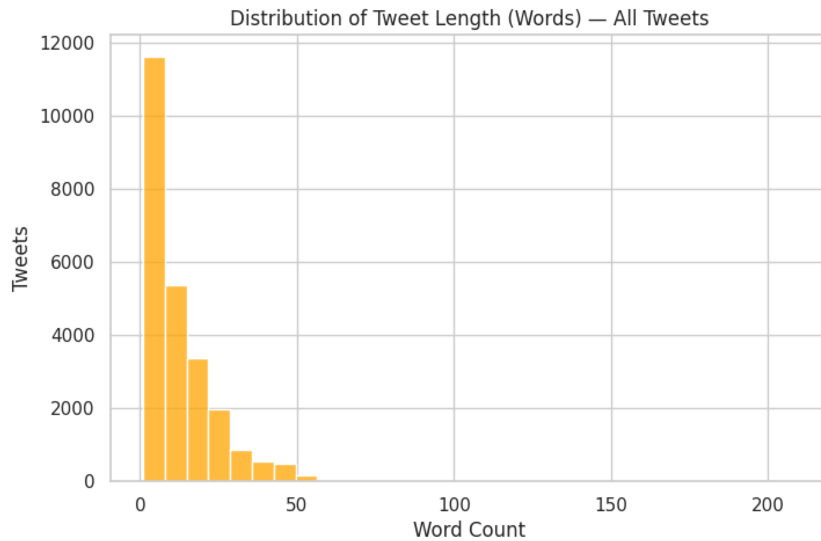

Figure 1- Distribution of Tweet Length

Three filters were applied to reduce noise in the data. Replies were excluded, as they tend to be more conversational than relevant to the market. Very short tweets—those with fewer than three words—were also filtered out. A combined dataset was created that excluded both replies and short tweets. The filtered sets will be compared with the full dataset in later analyses.

Daily returns were calculated as simple percentage changes in closing prices for the market variables, which is standard in financial econometrics. Daily volatility was estimated using the Garman–Klass formula, which combines high, low, open, and close prices to provide a solid estimate of intraday variability (Garman and Klass, 1980). Trading volume was included as a control variable to account for fluctuations in market activity.

3.3 Exploratory Data Analysis

3.3.1 Tweet Characteristics

Figure 2 shows Musk's tweet activity over time. The series reveals a sparse posting before 2018, followed by a sharp increase coinciding with Tesla's growing market relevance. This supports the decision to restrict the primary analysis to 2018–2023.

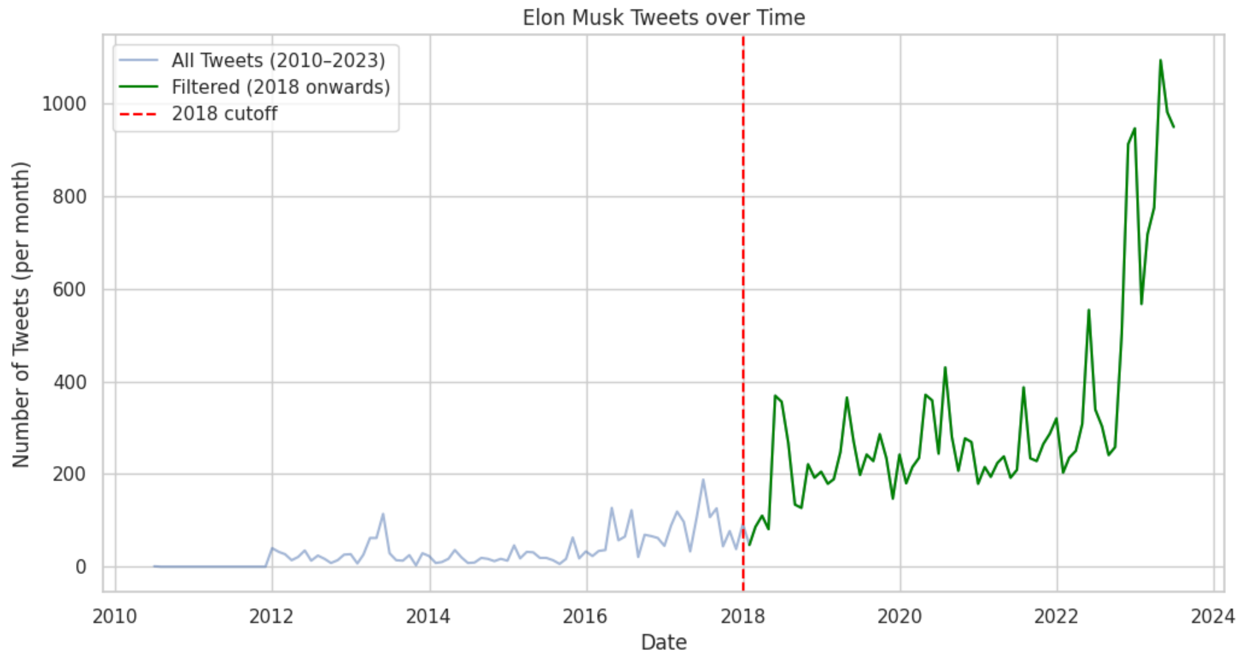

Figure 2- Elon Musk Tweets over Time

Figure 3 presents the distribution of VADER compound sentiment scores. Most tweets cluster around neutrality, but a clear skew towards positive sentiment is observable. This distribution underpins later analysis of asymmetric effects across sentiment classes.

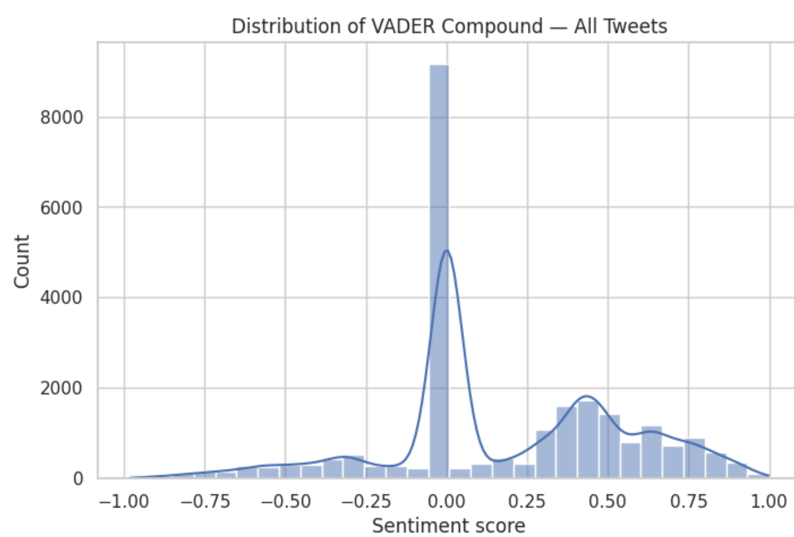

Figure 3- Distribution of VADER Compound Sentiment Score

Tweet length and type were also examined. Replies account for a substantial share of tweets and are frequently short, conversational posts. Figure 4 illustrates the split between original and reply tweets. This motivates the creation of filtered datasets, where noise from non-informative tweets is reduced.

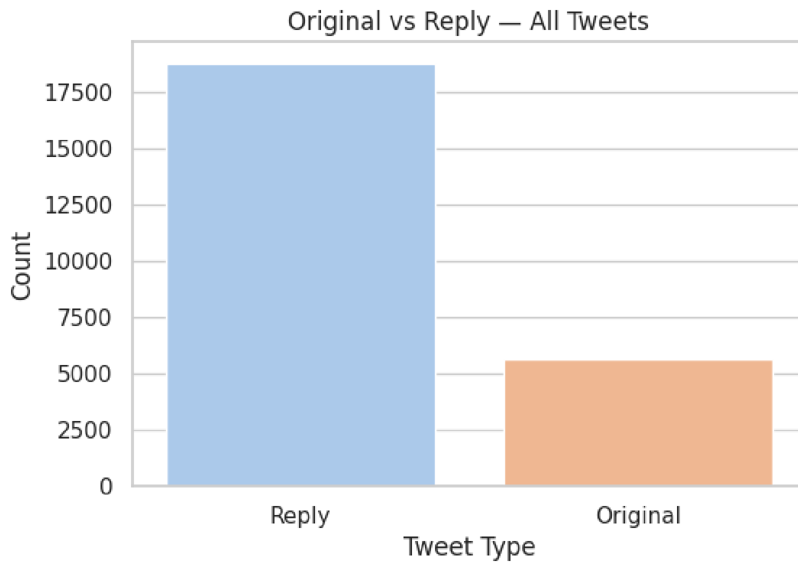

Figure 4- Original Vs Reply

3.3.2 Tesla Stock Characteristics

Tesla's stock trajectory provides essential context. Figure 5 plots daily closing prices, highlighting Tesla's significant appreciation after 2018. This rise coincided with increased public attention to the company and Musk's online persona.

Figure 5- Tesla Closing Price over Time

Volatility is a crucial control variable, as general risk conditions may confound sentiment effects. Figure 6 shows the daily volatility proxy derived from the Garman–Klass estimator. Volatility is elevated during periods of heightened market activity, which illustrates why risk adjustment is necessary in regression models.

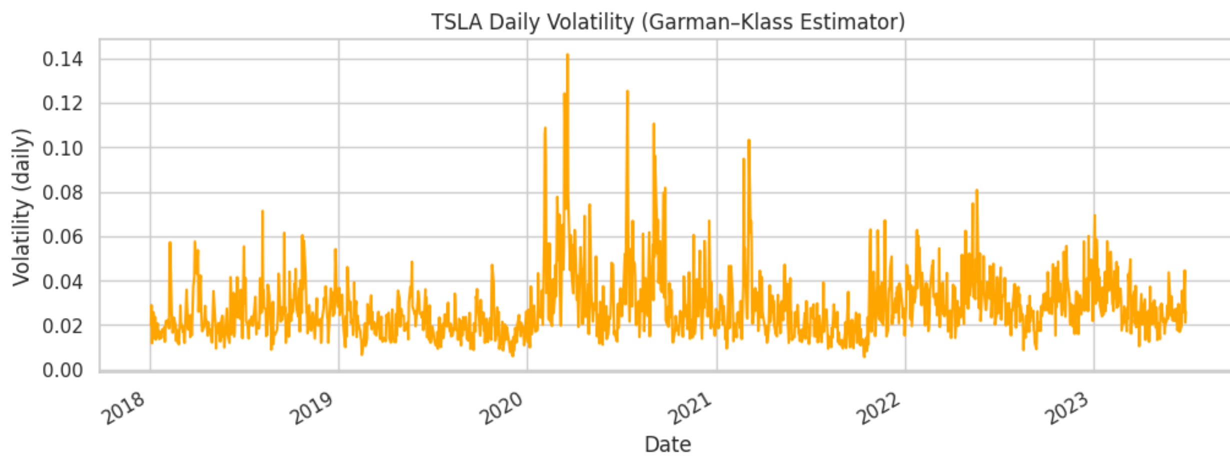

Figure 6- Tesla Daily Volatility (2018-2023)

3.3.3 Other Characteristics

In addition to analysing tweet activity and stock price movements, two other descriptive features provide valuable context.

The analysis of Musk's tweets reveals interesting trends in sentiment over time, as illustrated in Figure 11 in the Appendix. Sentiment shows a tendency to fluctuate but generally leans towards positivity. Notably, the years from 2010 to 2018 exhibit significant volatility, with sentiment scores swinging between 0.4 and 0.1. In contrast, the period starting in 2018 shows a more stable pattern, with average sentiment settling around 0.2 and fewer drastic changes. This transition towards stability supports the decision to focus the primary analysis on 2018 to 2023, as the earlier period may not reflect a consistent pattern in Musk's communication.

Second, Tesla's trading volume from 2018 to 2023 is depicted in Figure 12, Appendix. This chart highlights periods of exceptionally high trading activity, particularly between 2020 and 2021, coinciding with significant market events and the growing enthusiasm for Tesla among retail investors. The spikes in trading volume emphasise the necessity of including trading activity as a control variable in subsequent regressions.

While these factors are not directly integrated into the event study, they offer essential background information that helps to contextualise the broader environment in which Musk's tweets were made.

3.4 Summary of Analytical Datasets

The analysis focuses on the most relevant period for the market, leading to the exclusion of tweets before 2018. After removing 3,053 tweets (12.49%), the dataset was reduced to 21,397 tweets, covering January 3, 2018, to June 29, 2023. This timeframe saw significant increases in Tesla's stock activity and Elon Musk's tweeting frequency.

Following pre-processing and filtering, four daily datasets were constructed:

Table 1- Variants datasets

Dataset	Length	Percentage
"all"	21397	100.00%
"no_replies"	4102	19.17%
"no_short"	17755	82.98%
"no_replies_no_short"	3476	16.25%

The last provides the strongest signal-to-noise ratio and forms the primary dataset for the intraday event-study analysis. Table 1 summarises the size of each dataset after filtering.

4. Methodology

4.1 Research Design

The design follows a two-stage strategy. First, daily analysis explores whether sentiment extracted from Musk's tweets correlates with or predicts Tesla's returns, volatility, and trading activity. Second, the intraday event study isolates the immediate market reaction to individual tweets, focusing on short-term cumulative returns and trading volumes around tweet timestamps. This combined design allows for both broad inference and fine-grained causal identification, consistent with recent calls in the literature for multi-frequency approaches (Yang, Fernandez-Perez & Indriawan, 2025).

4.2 Sentiment Analysis

4.2.1 Choice of Algorithm

Sentiment scores were generated using the Valence Aware Dictionary and sEntiment Reasoner (VADER) tool (Hutto and Gilbert, 2014). VADER was selected for several reasons. First, it is specifically designed to perform well on short, informal texts such as tweets, which often contain slang, abbreviations, and emoticons. Second, it incorporates a lexicon for intensity modifiers and punctuation, which is common in Musk's tweeting style. Third, it has been widely applied in financial applications (Ranco et al., 2015; Gu and Kurov, 2020), providing comparability and methodological validity.

4.2.2 Classification Thresholds

VADER outputs a compound score between -1 (most negative) and +1 (most positive). Following established practice, tweets with scores greater than 0.05 were classified as positive, scores less than -0.05 as negative, and those in between as neutral (Hutto & Gilbert, 2014). These thresholds balance sensitivity with robustness and have been used extensively in prior financial sentiment studies (Bollen, Mao & Zeng, 2011; Karlemstrand and Leckström, 2021).

Before filtering, the tweets were classified using VADER thresholds into three sentiment categories: positive (compound score above 0.05), negative (below -0.05), and neutral (scores in between). Nearly half of the tweets (47.9%) were positive, 38.0% neutral, and 14.1% negative. This suggests a positive bias in Musk's public communications, which is relevant for understanding future return dynamics.

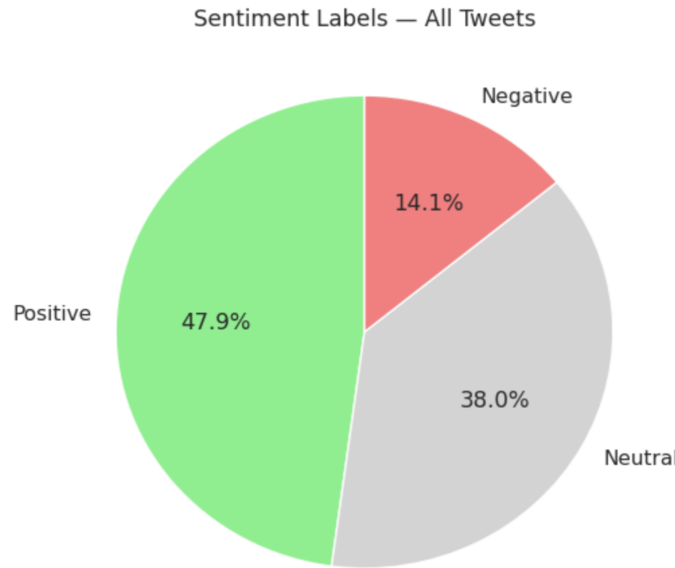

Figure 7- Sentiment Labels

4.3 Data Preparation

4.3.1 Time Zone Standardisation

US trading hours were defined as 14:30 to 21:00 UTC, corresponding to New York market hours. Tweets posted after 21:00 UTC were assigned to the following trading day, consistent with the assumption that out-of-hours tweets cannot influence same-day trading activity (Ranco et al., 2015).

4.3.2 Filtering Rules and Variants

After selecting the range of tweets starting from 2018, four datasets were constructed to test robustness to noise:

- All tweets — unfiltered.
- No replies — excluding tweets that were direct replies (identified by leading "@"), which often represent conversations with individuals rather than market-moving statements.
- No short tweets — excluding tweets with fewer than three words, which were considered too brief to contain substantive information.
- No replies and no short tweets — combining the above filters.

This filtering strategy is consistent with prior literature emphasising the importance of focusing on market-relevant tweets (Gu and Kurov, 2020).

4.3.3 Market Variables

Tesla's daily returns were calculated as simple percentage changes in closing prices:

$$Return_t = \frac{Close_t - Close_{t-1}}{Close_{t-1}}$$

Daily volatility was estimated using the Garman–Klass estimator, which combines high, low, open, and close prices to provide an efficient measure of volatility under the assumption of zero drift (Garman and Klass, 1980):

$$\sigma^2_{GK,t} = \frac{1}{2}\left(\ln\frac{H_t}{L_t}\right)^2 - (2\ln 2 - 1)\left(\ln\frac{C_t}{O_t}\right)^2$$

Where:

H_t = High price of day *t*

L_t = Low price of day *t*

O_t = Opening price of day *t*

C_t = Closing price of day *t*

This approach has been widely adopted in volatility research for its efficiency relative to squared returns.

Trading activity was captured via daily trading volume and included as a control variable to account for fluctuations in market activity.

4.4 Daily Analysis Framework

4.4.1 Aggregation of Sentiment

To convert tweet-level data into daily metrics, the tweets were organised by trading day. For each day, several key aggregates were calculated:

The mean compound score represents the average sentiment.

The total tweet count

The count of positive tweets

The count of negative tweets.

This approach effectively captures qualitative aspects, like sentiment polarity, and quantitative factors related to Musk's Twitter activity, like the volume of tweets.

4.4.2 Correlation Analysis

Initial relationships were examined using Pearson correlation coefficients between returns and daily tweet aggregates. The coefficients were reported together with p-values to indicate statistical significance. In addition to same-day associations, lagged correlations were calculated for one, two, and three days ahead to test whether sentiment effects persisted beyond the day of the tweet.

4.4.3 Regression Analysis

Daily regressions used ordinary least squares (OLS) with heteroskedasticity-robust (HC3) standard errors. Two model classes were estimated:

Baseline model:

$$Return_t = \alpha + \beta_1 AvgSent_t + \beta_2 Count_t + \epsilon_t$$

Extended model (controls added):

$$Return_t = \alpha + \beta_1 AvgSent_t + \beta_2 Count_t + \beta_3 Vol_t + \beta_4 \log{(Volume_t + 1)} + \epsilon_t$$

Where:

$AvgSent$ is daily sentiment

$Count$ is tweet count

Vol is daily volatility

$\log{(Volume)}$ captures the trading intensity

α is the intercept term

ϵ is the error term

Ordinary Least Squares (OLS) is commonly employed for estimating linear relationships in finance, mainly due to its favourable statistical properties. It offers unbiased and efficient estimates of regression coefficients with relatively mild assumptions. Moreover, when robust standard errors are applied, they remain valid even in heteroskedasticity. In empirical asset pricing and event studies, OLS has become popular for connecting market returns with various explanatory factors such as news sentiment, liquidity, and volatility.

For this study, OLS is deemed suitable as the focus is on understanding the average marginal effect of sentiment and tweet activity on Tesla's returns, rather than emphasising forecasting accuracy or capturing non-linear dynamics. The baseline model illustrates the direct relationship between sentiment and returns, while the extended specification considers other established factors that may also impact returns, such as volatility and liquidity. By evaluating the R^2 values across different models, it is possible to assess the additional explanatory power provided by each model.

In these regressions, the intercept α indicates the baseline level of returns when all explanatory variables are set to zero. The error term ϵ accounts for any unexplained return variation due to omitted variables or random noise.

4.5 High-Frequency Event Study

4.5.1 Rationale

Event study methodology, as outlined by MacKinlay in 1997, is commonly employed to evaluate how information shocks influence asset prices. Although daily regressions can reveal broader relationships, they often miss short-lived yet statistically significant intraday effects. Given their discrete and time-stamped nature, tweets are particularly well-suited for this type of event study design, as Yang, Fernandez-Perez, and Indriawan demonstrated in 2025.

4.5.2 Alignment of Tweets to Intraday Bars

In conducting intraday analysis, tweets were matched with the first trading bar occurring at or after the timestamp of each tweet. This approach ensured that returns were calculated from the first available price following the tweet, thereby eliminating any look-ahead bias. To strike a balance between detail and sample size, both 15-minute and 60-minute intervals were examined.

4.5.3 Event Windows

Two event windows were established for analysis.

The first one, with a 15-minute frequency, spans from four bars before the event to twelve bars after, covering one hour before and three hours after.

The second window operates at a 60-minute frequency, ranging from two bars before to eight bars after the event, which translates to a two-hour pre-event period and an eight-hour post-event period.

These timeframes were selected to effectively capture immediate market responses and potential short-term drift effects, aligning with findings from previous high-frequency sentiment research (Ranco et al., 2015).

4.5.4 Cumulative Average Returns (CARs)

Cumulative Average Returns (CARs) are essential for gauging stock market reactions during tweet events. For every tweet, simple returns are compiled over a defined event window, resulting in a cumulative series that captures the average market response up to a specific point in time. The tweets are categorised by positive, neutral, or negative sentiment, and CARs are calculated distinctly for each group. To assess statistical significance, t-tests are conducted on the cross-sectional means, with 95% confidence intervals provided to ensure robust findings.

4.5.5 Volume Analysis

To examine the effects of attention, trading volume was normalised against the average volume observed before the event. This approach yielded a relative volume measure that allows for comparisons across different days.

4.6 Limitations of Methodology

The methodological approach is not without limitations.

First, the use of VADER for sentiment analysis, although efficient and widely adopted, may fail to capture subtleties such as sarcasm, irony, or context-specific meaning—features that frequently appear in Musk’s communication style.

Second, the tweet filtering rules are heuristic, and different thresholds for excluding short messages or replies could lead to alternative results.

Third, the regression analysis may be affected by omitted variable bias. Tesla’s returns are influenced by many other contemporaneous events, such as earnings announcements or macroeconomic news, which are not explicitly controlled for in the models.

Finally, the alignment procedure moves all tweets to the next available trading bar. This means that tweets posted outside of market hours are shifted to the market open, which is also when the bulk of automated transactions occur. Any increase in trading volume at the open may reflect the alignment's mechanics as much as genuine investor responses to the tweets, complicating the interpretation of volume effects.

5. Results

The findings are presented in two sections. The first part focuses on a daily-level analysis examining the connection between Elon Musk's tweets and Tesla's stock performance. The second part consists of an intraday event study to capture the quick market reactions to specific tweets.

5.1 Daily Analysis

5.1.1 Correlation Analysis

The results presented in Table 2 summarise the Pearson correlation coefficients between Tesla's daily returns and aggregated sentiment measures across the four tweet datasets.

Table 2- Correlation Analysis

Dataset	N	r(Return, Sentiment)	p_sentiment	r(Return, Count)	p_count
all	1289	0.02617	0.34780	0.00912	0.74350
no_replies	937	0.10493	0.00130	0.02156	0.50989
no_replies_no_short	898	0.11990	0.00032	0.01740	0.60256
no_short	1278	0.04298	0.12460	0.01420	0.61206

Table 2- Correlation Analysis (cont.)

Dataset	r(Return, PosCount)	p_pos	r(Return, NegCount)	p_neg
all	0.02647	0.34239	-0.02126	0.44559
no_replies	0.06904	0.03459	-0.05725	0.07984
no_replies_no_short	0.07448	0.02563	-0.07342	0.02779
no_short	0.03856	0.16834	-0.02447	0.38216

The correlations observed in the unfiltered dataset appear weak and statistically insignificant, indicating that most tweets do not have a strong immediate connection to stock returns. However, when the dataset is filtered to exclude replies and concise messages, the correlation between returns and average sentiment shows a noticeable improvement. Notably, the dataset without replies and short messages demonstrates the most substantial effect ($r = 0.12$, $p < 0.001$). While not particularly strong, the correlation observed is statistically significant enough to indicate that tweets tend to share more substantial information instead of just casual or trivial remarks. This suggests a noticeable relationship between tweet activity and Tesla's stock performance.

Table 2 also presents correlations among tweet count, positive tweet count, negative tweet count, and average sentiment. Across all datasets, total tweet count shows minimal association with daily returns, with coefficients near zero and consistently insignificant. This indicates that the volume of Musk’s Twitter activity does not systematically impact the stock price.

Positive tweets demonstrate a stronger positive correlation with returns, especially in the no_replies and no_replies_no_short datasets, where the correlation coefficients are $r = 0.07$ and $r = 0.074$, respectively. While these values are small, they are statistically significant, exhibiting p-values of less than 0.05.

All correlations are negative, as expected for the count of negative tweets, but the values are minimal and mostly lack statistical significance. The only apparent exception is in the no_replies_no_short dataset ($r = –0.073$, $p = 0.028$), reinforcing the importance of filtering out replies and short tweets. Overall, any link between negative sentiment and downward returns appears weak and less consistent than the effect of positive tweets.

5.1.2 Lagged Correlations

To test whether sentiment exerts an influence beyond the same trading day, correlations were calculated between daily sentiment and returns one, two, and three days ahead. The results are presented in Table 3.

Table 3- Lagged Correlations

Dataset	Lag	r(Sent, Return t+lag)	p_value	N
all	1	0.04819	0.08374	1289
all	2	-0.04898	0.07889	1288
all	3	-0.00738	0.79152	1287
no_replies	1	0.03608	0.26994	937
no_replies	2	-0.04355	0.18317	936
no_replies	3	0.00507	0.87694	935
no_replies_no_short	1	0.01738	0.60298	898
no_replies_no_short	2	-0.06010	0.07202	897
no_replies_no_short	3	0.00694	0.83576	896
no_short	1	0.05505	0.04913	1278
no_short	2	-0.05570	0.04660	1277
no_short	3	-0.01089	0.69752	1276

The evidence for lagged effects appears to be relatively weak across all datasets.

In the unfiltered sample, the correlations are nearly zero, changing direction from day one to day two and failing to achieve conventional significance levels. A similar trend is observed in the no_replies and no_replies_no_short datasets, where the coefficients remain small and statistically indistinguishable from zero. The only results that come close to being significant are found in the no_short dataset, showing a positive correlation at a one-day lag ($r = 0.055$, $p = 0.049$) and a negative correlation at a two-day lag ($r = –0.056$, $p = 0.047$). However, by day three, these effects vanish.

These findings indicate that the connection between sentiment and returns is most pronounced on the same day and does not reliably extend into subsequent trading days.

Correlation analysis offers a preliminary understanding of the relationship between sentiment and returns, but does not account for factors like volatility or trading activity. The next step involves using regression models to estimate the effect of sentiment on returns while considering additional market variables.

5.1.3 OLS Regression Analysis

To build on the correlations, daily regressions were estimated to test whether sentiment had explanatory power for returns once other factors were considered. With HC3 robust standard errors, OLS was used to correct for potential heteroskedasticity. Two versions of the model were run. The baseline specification included only sentiment and tweet count, while the extended specification added controls for daily volatility and trading volume.

Baseline Model

Table 4 summarises the baseline OLS regressions, highlighting the coefficients related to sentiment, tweet count, and model fit (R^2). The complete regression outputs, which include constants and standard errors, can be found in Appendix Table 8 for further reference.

Table 4- OLS Baseline Model

Dataset	coef_sent	p_sent	coef_count	p_count	R2
all	0.00636	0.29551	0.00004	0.67677	0.00080
no_replies	0.01719	0.00031	0.00053	0.34664	0.01194
no_replies_no_short	0.01833	0.00006	0.00054	0.43003	0.01511
no_short	0.00985	0.10098	0.00008	0.53244	0.00214

The initial regression analysis shows that sentiment has minimal impact on the unfiltered dataset, with a coefficient of 0.0064 and an insignificant p-value of 0.30. Tweet count similarly lacks meaningful correlation.

However, when replies are filtered out, the sentiment coefficient increases to 0.0172 ($p < 0.001$), indicating that non-reply tweets aimed at a broader audience are linked to higher returns. Excluding short tweets further boosts the coefficient to 0.0183 ($p < 0.001$). Although R^2 values remain modest (around 1–1.5%), they are higher than in the unfiltered dataset, highlighting the benefit of filtering for more relevant tweets.

The number of tweets does not predict market returns, with coefficients around zero and statistically insignificant results. This indicates that Elon Musk's tweet volume does not affect market movements; instead, the sentiment in the tweets is what matters.

The findings suggest that the content and tone of Musk's original, non-reply tweets show a meaningful link with Tesla's daily returns, whereas the number of tweets he posts by itself does not appear to move the market.

Extended Model with Controls

Additional controls were introduced to capture other well-known drivers of stock returns and build on the baseline regressions. Daily volatility and trading activity, expressed as the logarithm of volume, were added to the model. Including these controls makes it possible to test whether the relationship between sentiment and returns holds once broader market dynamics are accounted for.

The results are presented in a condensed form in Table 5, while the complete set of regression outputs can be found in the Appendix Table 9.

Table 5- OLS Extended Model

Dataset	coef_sent	p_sent	coef_vol	p_vol	coef_logvol	p_logvol	R2
all	0.00578	0.33383	-0.38194	0.02688	0.01403	0.00001	0.02404
no_replies	0.01732	0.00021	-0.46461	0.00343	0.01375	0.00029	0.03638
no_replies_no_short	0.01866	0.00003	-0.46601	0.00499	0.01379	0.00041	0.03955
no_short	0.00947	0.10812	-0.36387	0.03562	0.01376	0.00002	0.02420

The extended regressions support the key trends in the baseline models while providing additional insights. In the unfiltered dataset, sentiment appears insignificant (coefficient = 0.0058, $p = 0.33$). However, when replies are excluded, it becomes positive and highly significant (coefficient = 0.0173, $p < 0.001$), and this effect strengthens even more when both replies and short tweets are removed (coefficient = 0.0187, $p < 0.001$).

The analysis reveals clear and significant relationships among the controls. Daily volatility tends to show a negative association with returns across all specifications, with coefficients ranging from about –0.36 to –0.47, and these findings are backed by statistically significant p-values. This aligns with the notion that days with higher volatility are often linked to price reversals or heightened uncertainty, which usually dampens returns. In contrast, log-transformed trading volume consistently appears to have a strong positive relationship in every dataset, indicating that increased trading activity is generally associated with higher returns.

While the R^2 values remain relatively low, there is an observed increase compared to the baseline models. The dataset without replies and short tweets shows the strongest fit, with an R^2 of 0.040, nearly three times greater than its baseline equivalent. This suggests that applying filters and including standard market controls enhances the models' ability to explain return variations.

Overall, the extended regressions indicate that when filtering for substance, Musk's tweets still have explanatory power for returns, even when accounting for volatility and trading activity. Additionally, the control variables behave as anticipated, which supports the model's validity.

R^2 - Baseline Vs Extended

R^2 measures how much of the variation in daily stock returns can be explained by the independent variables included in a model. In financial econometrics, it's common to find R^2 values to be modest because stock returns are subject to numerous unpredictable factors. However, even slight increases in R^2 can be significant, indicating that additional variables may capture essential aspects of market behaviour.

A comparison of R^2 across four datasets is illustrated in Figure 8. The baseline models, which only accounted for average sentiment and tweet counts, showed limited ability to explain return variations, with R^2 values ranging from 0.0008 to 0.015. By incorporating volatility and log-transformed trading volume as control variables, the extended models demonstrated a notable improvement; the R^2 values increased to between 0.024 and 0.040, equating to a two- to three-fold enhancement in explanatory power. A detailed table with the specific numbers is available in Appendix Table 10.

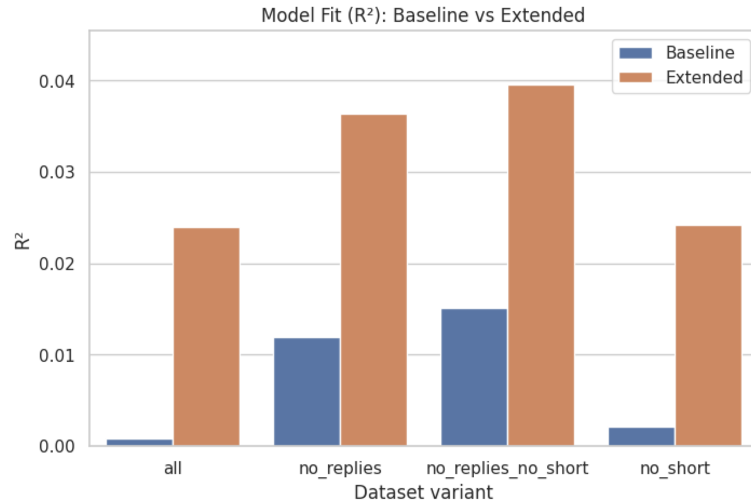

Figure 8- R2 OLS Baseline Vs Extended

The dataset labelled no_replies_no_short exhibited the strongest fit, with the extended model reaching an R^2 of 0.0395, compared to 0.0151 for the baseline. This suggests that filtering out conversational and trivial tweets can enhance the signal derived from sentiment analysis, and that including controls for volatility and trading volume leads to a more robust approach to modelling returns. While the absolute R^2 values remain relatively low—aligning with previous research highlighting the challenges of predicting stock returns—the relative increase underscores the incremental explanatory power of Elon Musk's Twitter activity when market conditions are factored in.

5.1.4 Visualising Coefficients

Figure 13 in the Appendix presents the coefficients on average sentiment and their 95% confidence intervals across baseline and extended models.

Two main observations stand out. First, in the unfiltered dataset (“all”), the coefficient is nearly zero with a confidence interval overlapping the null line, indicating a weak relationship. Excluding replies (“no_replies”) results in a larger coefficient and a narrower confidence interval, enhancing statistical significance. The effect is most potent when replies and very short tweets are filtered out (“no_replies_no_short”), yielding the most robust estimates.

Second, the similarity between baseline and extended models shows that including controls for volatility and log-volume doesn’t weaken the positive effect of sentiment; results remain stable. In contrast, the “no_short” dataset produces smaller, noisier coefficients with wide confidence intervals that overlap zero, suggesting this filter is less effective.

Overall, the visual evidence supports that refining the dataset improves the clarity and reliability of the relationship between sentiment and returns.

5.2 High-Frequency Event Study

Daily regressions can highlight general trends linking Elon Musk's tweets to Tesla's stock returns, but they might miss the brief market reactions that happen right after a tweet is posted. An event study approach was used to tackle this issue, focusing on intraday intervals of 15 and 60 minutes. The dataset utilised for this study was the no_replies_no_short set, as it demonstrated the highest performance in the previous analysis. Tweets were matched to the first trading bar after publication, allowing for an analysis of cumulative average returns (CARs) and normalised trading volumes within specified event windows.

5.2.1 15-minute Intervals Windows

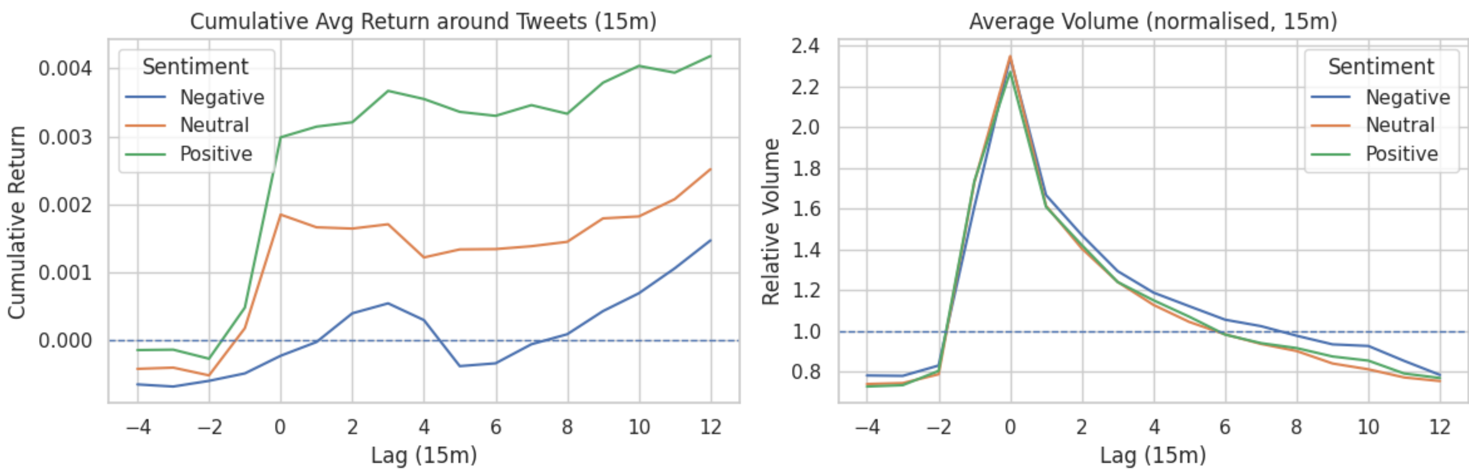

Figure 9- CAR & Volume, 15min

Figure 9 shows that an apparent asymmetry emerges across sentiment categories. A sharp and persistent upward drift follows positive tweets in returns: at Lag 0, the mean CAR is around 0.30% ($p < 0.001$), steadily increasing to 0.42% by Lag 12. Neutral tweets also display some effect, with modest but statistically significant gains at the open (0.18%, $p = 0.017$) and again at Lag 12 (0.25%, $p = 0.013$). By contrast, negative tweets show no systematic relationship with returns—their coefficients remain small and statistically insignificant across the entire window. All numerical results for CAR with a 95% confidence interval are in Table 6 below.

Table 6- CAR with 95%CI,15 min

Sentiment	Horizon	N tweets	Mean CAR	95% CI	t-stat	p-value
Negative	Lag 0 (15m)	499	-0.00013	[-0.0026, 0.0023]	-0.10352	0.91759
Negative	Lag 4 (15m)	499	0.00040	[-0.0024, 0.0032]	0.27280	0.78512
Negative	Lag 8 (15m)	499	0.00017	[-0.0030, 0.0033]	0.10437	0.91692
Negative	Lag 12 (15m)	499	0.00152	[-0.0018, 0.0049]	0.88624	0.37592
Neutral	Lag 0 (15m)	1460	0.00185	[0.0003, 0.0034]	2.39866	0.01658
Neutral	Lag 4 (15m)	1460	0.00122	[-0.0005, 0.0030]	1.37786	0.16846
Neutral	Lag 8 (15m)	1460	0.00145	[-0.0004, 0.0033]	1.50634	0.13220
Neutral	Lag 12 (15m)	1460	0.00252	[0.0005, 0.0045]	2.49393	0.01274
Positive	Lag 0 (15m)	1516	0.00299	[0.0015, 0.0044]	4.00583	0.00007
Positive	Lag 4 (15m)	1516	0.00355	[0.0018, 0.0053]	3.98110	0.00007
Positive	Lag 8 (15m)	1516	0.00334	[0.0014, 0.0052]	3.43325	0.00061
Positive	Lag 12 (15m)	1516	0.00418	[0.0021, 0.0062]	4.02305	0.00006

5.2.2 60-minute Intervals Windows

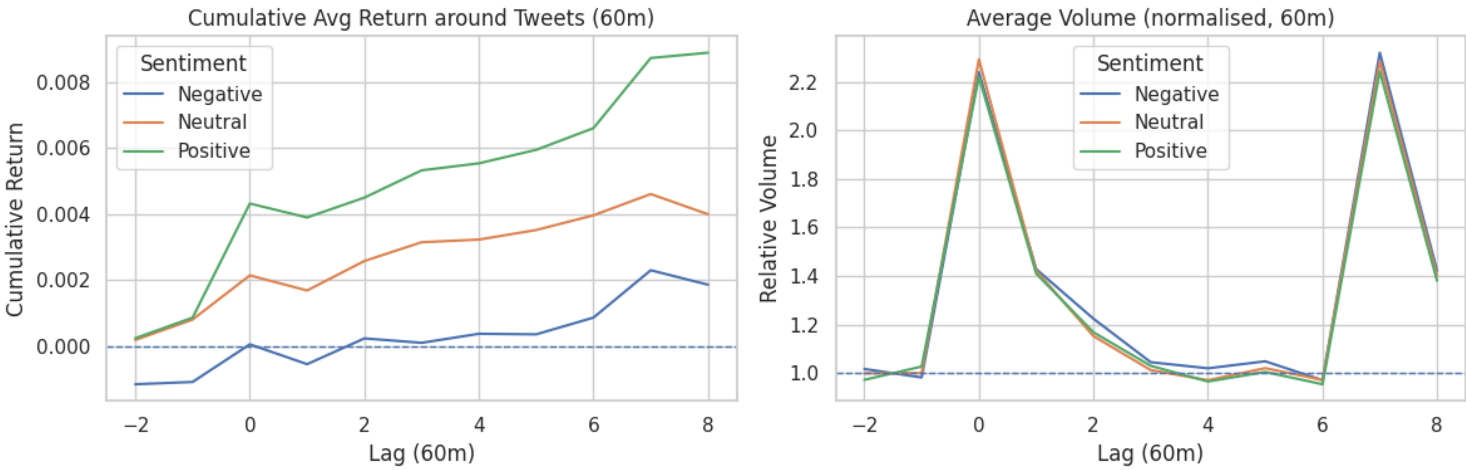

Figure 10- CAR & Volume, 60min

In Figure 10, the findings clearly illustrate how different types of tweets impact stock prices. Positive tweets lead to substantial and persistent price increases, while neutral tweets have a more moderate effect, and negative tweets tend to blend into the background noise.

The CAR for positive tweets stands out right from the beginning, showing an impressive gain of 0.43% at Lag 0, with a significance level of $p < 0.001$. This positive trend continues to escalate throughout the event window, reaching nearly 0.9% by Lag 8, which remains statistically significant ($p < 0.001$). On the other hand, neutral tweets also produce statistically significant upward movements but to a lesser extent, climbing from 0.21% at Lag 0 ($p = 0.014$) to around 0.40% by Lag 8 ($p = 0.006$). In stark contrast, negative tweets show little to no movement across the time frames analysed, with CARs hovering around zero and none achieving statistical significance. Detailed numerical results for CAR and a 95% confidence interval can be found in Table 7 below.

Table 7- CAR with 95%CI, 60 min

Sentiment	Horizon	N tweets	Mean CAR	95% CI	t-stat	p-value
Negative	Lag 0 (60m)	499	0.00016	[-0.0028, 0.0031]	0.10569	0.91587
Negative	Lag 2 (60m)	499	0.00025	[-0.0031, 0.0036]	0.14657	0.88353
Negative	Lag 4 (60m)	499	0.00045	[-0.0035, 0.0044]	0.22498	0.82208
Negative	Lag 6 (60m)	499	0.00093	[-0.0033, 0.0051]	0.43648	0.66268
Negative	Lag 8 (60m)	499	0.00185	[-0.0034, 0.0071]	0.68536	0.49343
Neutral	Lag 0 (60m)	1460	0.00214	[0.0004, 0.0038]	2.46568	0.01379
Neutral	Lag 2 (60m)	1460	0.00258	[0.0006, 0.0046]	2.54179	0.01113
Neutral	Lag 4 (60m)	1460	0.00323	[0.0010, 0.0054]	2.89360	0.00387
Neutral	Lag 6 (60m)	1460	0.00396	[0.0016, 0.0063]	3.30528	0.00097
Neutral	Lag 8 (60m)	1460	0.00399	[0.0012, 0.0068]	2.76687	0.00573
Positive	Lag 0 (60m)	1516	0.00432	[0.0026, 0.0060]	4.88901	0.00000
Positive	Lag 2 (60m)	1516	0.00450	[0.0025, 0.0065]	4.36141	0.00001
Positive	Lag 4 (60m)	1516	0.00554	[0.0033, 0.0077]	4.92146	0.00000
Positive	Lag 6 (60m)	1516	0.00660	[0.0042, 0.0090]	5.49063	0.00000
Positive	Lag 8 (60m)	1516	0.00889	[0.0058, 0.0119]	5.70820	0.00000

These 60-minute results therefore strengthen the earlier 15-minute findings: Musk's positive communications are followed by measurable and sustained price increases, whereas neutral and negative tweets have more muted or negligible return impacts.

5.2.3 Volume Effects

The volume dynamics surrounding tweet events are illustrated in the right-side pane of Figures 9 and 10.

In a 15-minute interval, trading activity depicts a significant spike at lag 0, where relative volume rises sharply above the baseline. This surge coincides with the opening of the US markets, when order books are cleared and many automatic trades are executed. Since tweets posted outside market hours are shifted to the first available bar, their influence overlaps with this naturally busy period. Consequently, the peak at lag zero may reflect market opening dynamics rather than directly responding to Musk's tweets. Following this spike, relative volume steadily decreases through lags 4, 6, and 8, aligning with the expected intraday pattern of diminishing activity. Additionally, normalisation was conducted relative to pre-event windows, which typically capture calmer trading conditions near the market close, making the market open appear even more pronounced in comparison.

When looking at a 60-minute interval, a similar pattern emerges with two distinct peaks: one at lag zero and another around lag 8. The first peak corresponds to the market open, while the second reflects the next trading day's opening. Given that six-hourly bars cover the length of a standard US trading session, lag eight approximately marks the market's reopening, explaining the recurrence of increased volume.

In summary, these patterns indicate that while there is clear evidence of systematic volume spikes, they align more with predictable features of market microstructure—particularly opening auctions—than with the tweets' sentiment. Thus, the analysis does not convincingly support that Musk's tweets drive trading volume directly. Instead, the results reveal a limitation in shifting out-of-hours tweets to the next trading session, as this methodological choice conflates tweet timing with structural patterns in market activity.

5.3 Summary of Findings

The analysis indicates that Elon Musk's tweets have varying effects on Tesla's stock performance based on data filtering and the time frame considered. Daily, weak and insignificant relationships exist between tweet sentiment and stock returns. However, excluding conversational replies and short tweets reveals a significant correlation, where positive sentiment correlates with higher returns, while negative sentiment shows minimal effects.

Further regression analysis demonstrates that while sentiment alone has limited explanatory power, incorporating market factors like volatility and trading volume enhances its significance in filtered datasets. As expected, higher volatility tends to lower returns, and increased trading volume is positively related to them, although overall explanatory power remains modest.

Positive tweets lead to immediate and lasting price increases at shorter intervals, with neutral tweets causing minor effects, whereas negative tweets generally have no significant influence. This indicates that markets react more strongly to Musk’s optimistic messages.

Additionally, trading volume spikes around tweet events, often aligning with market structures rather than sentiment, particularly during opening auctions. Overall, the findings suggest that, when relevant tweets are considered, Musk’s communications can impact the market, primarily through positive sentiment, with more noticeable effects on short-term price movements rather than trading volume.

6. Discussion

This section reflects on the findings, considering existing literature, theoretical implications, and methodological constraints. The results indicate a connection between Elon Musk's Twitter activity and Tesla's stock performance, revealing complexities that underscore the informational and behavioural pathways through which social media impacts financial markets.

6.1 Sentiment vs. Tweet Volume

The regression analyses consistently demonstrate that sentiment, rather than tweet frequency, serves as the primary driver of Tesla's returns. This reinforces the notion that markets react more to messages' content than their sheer quantity. The findings align with the literature on information-driven trading, suggesting that qualitative signals hold more value than mere attention-grabbing noise (Gu & Kurov, 2020). The weak explanatory power of tweet volume indicates that Musk's Twitter activity influences markets through perceived sentiment containing material or strategic information, rather than through simple attention.

6.2 Asymmetry in Market Response

A notable aspect of the high-frequency results is the asymmetric market response: positive tweets lead to statistically significant price drifts, whereas negative tweets exert little systematic impact. This may reflect an investor optimism bias, consistent with behavioural finance theories that propose good news is often overemphasised while negative information tends to be downplayed or absorbed more gradually (Ranco et al. 2015). Alternatively, it may suggest that Musk's positive communications are viewed as credible signals regarding Tesla's prospects, while negative or sarcastic tweets are dismissed as mere noise. The persistent upward CARs following positive tweets, which last several hours, align with observed mispricing and correction mechanisms documented in intraday event studies (Yang et al. 2025).

6.3 Volume Effects and Market Microstructure

The volume analysis indicates pronounced spikes at market openings but shows minimal evidence of tweet-driven surges beyond these structural patterns. This finding highlights the methodological challenge of aligning out-of-hours tweets with subsequent trading activity. Because opening auctions concentrate trading, disentangling the impact of Musk's tweets from regular market dynamics proves challenging. Thus, while volume peaks are evident, caution is warranted regarding their attribution to Twitter activity. This underscores the need to account for market microstructure in the design of event studies, as noted by Cookson et al. (2025).

6.4 Comparison with Existing Research

The findings align broadly with prior studies showing that Twitter sentiment related to firms contains information not captured by traditional sources (Gu & Kurov, 2020; Milikich & Johnson, 2023). The low explanatory power observed in daily regressions mirrors earlier research documenting the challenges of predicting stock returns overall (Ranco et al., 2015). However, the intraday analysis reveals that meaningful yet short-lived signals can be extracted from Musk's tweets, echoing high-frequency research that has identified mispricing corrected within days (Yang et al., 2025). Additionally, the asymmetry in effects adds depth to this literature by illustrating that not all sentiment categories are processed equally in the market.

6.5 Limitations

It's important to recognise several limitations that deserve attention. First, since VADER struggles with sarcasm, some of Musk's joking tweets may have been misclassified, potentially weakening the observed relationships, which often characterise Musk's communication style. Second, the filtering choices applied to the tweet datasets are heuristic; different thresholds may yield varying results.

Furthermore, the regressions may be subject to omitted variable bias, as corporate announcements, macroeconomic events, or broader market shocks could coincide with Musk's tweets, potentially confounding observed effects. Moreover, the observed volume spikes might reflect the market structure instead of just the influence of Musk's tweets. This adjustment introduces a structural bias since out-of-hours tweets were moved to the next trading session. The mapping of these tweets to the market open—an inherently high-volume and volatile period—can skew the results. This methodological compromise limits the ability to attribute volume effects to tweets cleanly.

6.6 Contribution

This study highlights an important finding: daily models tend to underestimate the impact of Elon Musk's tweets on stock prices. In contrast, more immediate, intraday event studies reveal statistically significant price changes. This suggests that the timing of the analysis is crucial when examining how social media influences financial markets. Both approaches make it clear that daily sentiment measures explain very little of Tesla's return variations. In contrast, shorter intraday event windows show notable and uneven reactions to Musk's positive tweets. This dual perspective effectively connects two areas of research—long-term predictive studies and high-frequency event analyses—demonstrating that the financial implications of Twitter interactions are better captured in the brief, intraday movements rather than through daily summaries.

7. Conclusion

The dissertation aimed to explore the potential impact of Elon Musk's Twitter activity on Tesla's stock performance through three key research questions: (1) Do Musk's tweets influence Tesla's returns and volatility? (2) Does the sentiment of tweets predict short-term returns or trading volume? (3) Are there any asymmetric effects between positive and negative tweets? A mixed methodology was employed to answer these questions, combining daily regressions with high-frequency event studies over a dataset comprising Musk's tweets and Tesla's stock data from 2018 to 2023.

In the daily analysis, the links between tweet sentiment and stock returns were weak and inconsistent when all types of tweets were included. However, eliminating replies and very brief messages revealed that sentiment coefficients became statistically significant, indicating that substantive tweets reaching a broad audience may carry more informational weight than casual interactions. Yet, the explanatory power of the daily models remained limited, with R^2 values seldom exceeding 4%. This aligns with existing literature on stock return predictability, which often highlights the dominance of noise and external shocks at daily horizons.

In contrast, the intraday event study yielded more substantial evidence of market reactions. At both 15- and 60-minute intervals, positive tweets resulted in statistically significant cumulative average returns, peaking at around 0.9% within a trading day. These effects persisted for several hours, suggesting that the market gradually absorbed signals from Musk's communications. Neutral tweets displayed modest impacts, while negative tweets did not show consistent effects. This asymmetry indicates that investors may downplay negative remarks or view them as less credible than positive communications, interpreted as informative regarding Tesla's future.

The trading volume analysis revealed challenges in aligning out-of-hours tweets with the first trading moments of the day. While spikes in trading volume were noted at market openings, these patterns appeared to reflect predictable market microstructure dynamics rather than a direct effect of Twitter activity. This underscores the need to approach timing assumptions carefully when analysing social media's influence on financial markets.

Overall, the findings suggest that Musk's tweets influence Tesla's stock price, primarily in the short term and in an asymmetric manner. The study contributes to existing literature by linking daily and intraday perspectives, demonstrating that while daily models exhibit limited predictive power, intraday analysis uncovers distinct and temporary price reactions. This suggests that Musk's tweets function more as fleeting shocks to investor sentiment than enduring informational signals.

For investors, these results suggest that tracking Musk’s tweets throughout the trading day is more beneficial than focusing solely on daily summaries. By filtering out replies and less significant posts, the signal becomes clearer, emphasising that not every tweet holds the same importance.

For researchers, the findings showcase the advantages of integrating daily and intraday methods when examining the financial impacts of social media. Refining how out-of-hours tweets are handled is critical, as current approaches may unintentionally blend genuine tweet influences with the predictable movements observed at market openings.

Several limitations of the study were acknowledged, such as reliance on heuristic sentiment classification, potential omitted variables, and biases arising from aligning out-of-hours tweets with market openings. Nevertheless, the research enhances the understanding of how social media interactions, particularly those from high-profile individuals, engage with financial markets.

Future research could expand on this study by integrating more sophisticated natural language processing models capable of understanding nuances like sarcasm and context. Additionally, refining the analysis of tweet timing would be beneficial. Investigating cross-firm or cross-market spillover effects might also yield valuable insights.

While this research offers compelling evidence of Musk's influence on Tesla, it's essential to recognise that these findings might not universally apply to other companies or CEOs. Musk's celebrity status and Tesla’s distinctive investor base play a significant role in the heightened responsiveness to his Twitter activity. Further studies could explore whether similar trends are observed in reactions to communications from other prominent executives.

References

Singh, A. (2023). *Elon Musk Tweets (Daily Updated)* [Dataset]. Kaggle. Available at: https://www.kaggle.com/datasets/aryansingh0909/elon-musk-tweets-updated-daily (Accessed: 2 September 2025).

Gu, C. and Kurov, A. (2020). *Informational role of social media: Evidence from Twitter sentiment*. Journal of Banking & Finance, 121, p.105969. Available at: https://doi.org/10.1016/j.jbankfin.2020.105969 (Accessed: 2 September 2025).

Bollen, J., Mao, H. and Zeng, X. (2011). *Twitter mood predicts the stock market*. Journal of Computational Science, 2(1), pp.1–8. Available at: https://doi.org/10.1016/j.jocs.2010.12.007 (Accessed: 2 September 2025).

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M. and Mozetič, I. (2015). *The effects of Twitter sentiment on stock price returns*. PLoS ONE, 10(9), e0138441. Available at: https://doi.org/10.1371/journal.pone.0138441 (Accessed: 2 September 2025).

Milikich, N. and Johnson, J. (2023). *Taureau: A stock market movement inference framework based on Twitter sentiment analysis*. arXiv preprint arXiv:2303.17667. Available at: https://doi.org/10.48550/arXiv.2303.17667 (Accessed: 2 September 2025).

Karlemstrand, R. and Leckström, E. (2021). *Using Twitter attribute information to predict stock prices*. Bachelor's thesis. KTH Royal Institute of Technology. Available at: https://doi.org/10.48550/arXiv.2105.01402 (Accessed: 2 September 2025).

Mokhtari, M., Seraj, A., Saeedi, N. and Karshenas, A. (2023). *The impact of Twitter sentiments on stock market trends*. Isfahan University of Technology. Available at: https://doi.org/10.48550/arXiv.2302.07244 (Accessed: 2 September 2025).

Cookson, J.A., Fox, C., Gil-Bazo, J., Imbet, J.F. and Schiller, C. (2025). *Social media as a bank run catalyst*. Working paper. Available at: https://ssrn.com/abstract=4422754 (Accessed: 2 September 2025).

Yang, N., Fernandez-Perez, A. and Indriawan, I. (2024). *The price impact of tweets: A high-frequency study*. Financial Review, 60(1), pp.147–171. Available at: https://doi.org/10.1111/fire.12406 (Accessed: 2 September 2025).

Hutto, C.J. and Gilbert, E. (2014). VADER: *A parsimonious rule-based model for sentiment analysis of social media text*. In: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media. AAAI. Available at: https://doi.org/10.1609/icwsm.v8i1.14550 (Accessed: 2 September 2025).

Garman, M.B. and Klass, M.J. (1980). *On the estimation of security price volatilities from historical data*. Journal of Business, 53(1), pp.67–78. Available at: https://www.jstor.org/stable/2352358 (Accessed: 2 September 2025).

MacKinlay, A.C. (1997). *Event studies in economics and finance*. *Journal of Economic Literature*, 35(1), pp.13–39. Available at: https://www.jstor.org/stable/2729691 (Accessed: 2 September 2025).

Sprenger, T.O., Tumasjan, A., Sandner, P.G. and Welpe, I.M. (2014). *Tweets and trades: The information content of stock microblogs*. *European Financial Management*, 20(5), pp.926–957. Available at: https://doi.org/10.1111/j.1468-036X.2013.12007.x (Accessed: 3 September 2025).

Sul, H.K., Dennis, A.R. and Yuan, L. (2017). *Trading on Twitter: Using social media sentiment to predict stock returns*. Decision Sciences, 48(3), pp.454–488. Available at: https://doi.org/10.1111/deci.12229 (Accessed: 3 September 2025).

Appendix

Code

https://github.com/KorvenDalas/MSc_FYP/blob/main/_FYP%20Code.ipynb

AI Tools

ChatGPT was a code assistant, helping troubleshoot Python scripts and optimise workflows during data processing and analysis.

Grammarly was employed as a writing assistant to check grammar, spelling, and consistency.

Tables

Table 8- OLS Baseline Model, Full Output

Dataset	N	coef_const	se_const	p_const	coef_sent	se_sent	p_sent
all	1289	0.00093	0.00212	0.66094	0.00636	0.00608	0.29551
no_replies	937	-0.00199	0.00244	0.41455	0.01719	0.00477	0.00031
no_replies_no_short	898	-0.00232	0.00258	0.36800	0.01833	0.00457	0.00006
no_short	1278	-0.00004	0.00218	0.98682	0.00985	0.00601	0.10098

Table 8- OLS Baseline Model, Full Output (Cont.)

Dataset	coef_count	se_count	p_count	R2
all	0.00004	0.00009	0.67677	0.00080
no_replies	0.00053	0.00057	0.34664	0.01194
no_replies_no_short	0.00054	0.00068	0.43003	0.01511
no_short	0.00008	0.00013	0.53244	0.00214

Table 9- OLS Extended Model, Full Output

Dataset	N	coef_const	se_const	p_const	coef_sent	se_sent	p_sent
all	1289	-0.24827	0.05711	0.00001	0.00578	0.00598	0.33383
no_replies	937	-0.24449	0.06856	0.00036	0.01732	0.00467	0.00021
no_replies_no_short	898	-0.24575	0.07067	0.00051	0.01866	0.00451	0.00003
no_short	1278	-0.24473	0.05750	0.00002	0.00947	0.00589	0.10812

Table 9- OLS Extended Model, Full Output (Cont.)

Dataset	coef_count	se_count	p_count	coef_vol	se_vol	p_vol
all	0.00001	0.00010	0.89581	-0.38194	0.17256	0.02688
no_replies	0.00065	0.00057	0.24859	-0.46461	0.15878	0.00343
no_replies_no_short	0.00067	0.00069	0.33044	-0.46601	0.16597	0.00499
no_short	0.00004	0.00013	0.74807	-0.36387	0.17317	0.03562

Table 9- OLS Extended Model, Full Output (Cont.)

Dataset	coef_logvol	se_logvol	p_logvol	R2	Adj_R2
all	0.01403	0.00320	0.00001	0.02404	0.02100
no_replies	0.01375	0.00379	0.00029	0.03638	0.03224
no_replies_no_short	0.01379	0.00391	0.00041	0.03955	0.03524
no_short	0.01376	0.00323	0.00002	0.02420	0.02113

Table 10- R2 Comparison

Dataset	R2_extended	R2_baseline
all	0.02404	0.00080
no_replies	0.03638	0.01194
no_replies_no_short	0.03955	0.01511
no_short	0.02420	0.00214

Figures

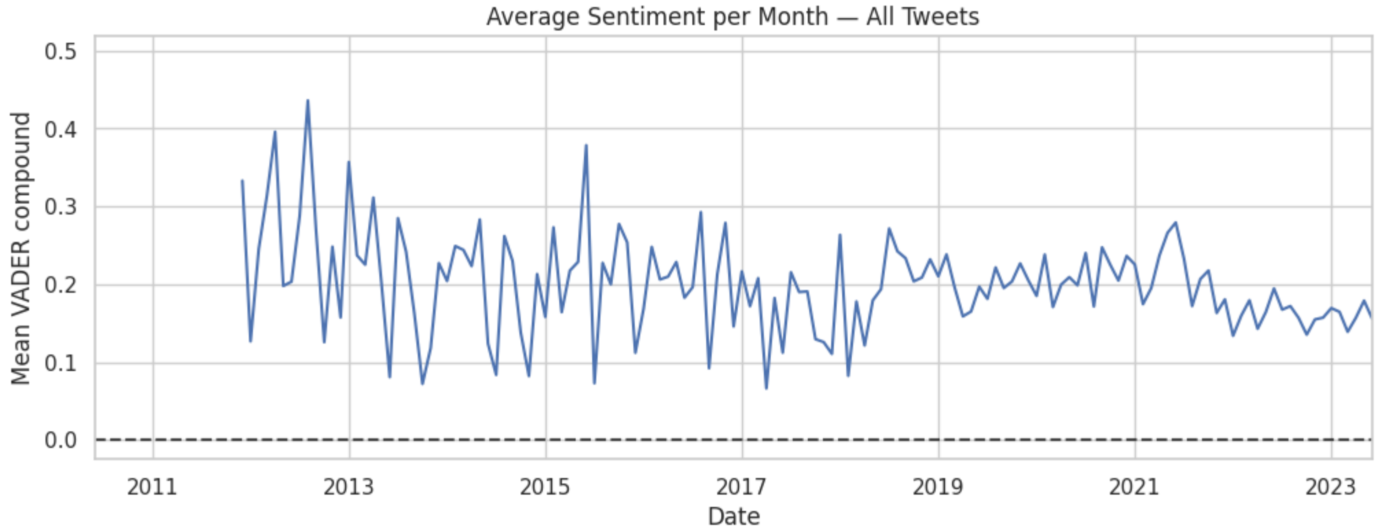

Figure 11- Average Sentiment per Month (2010 – 2023)

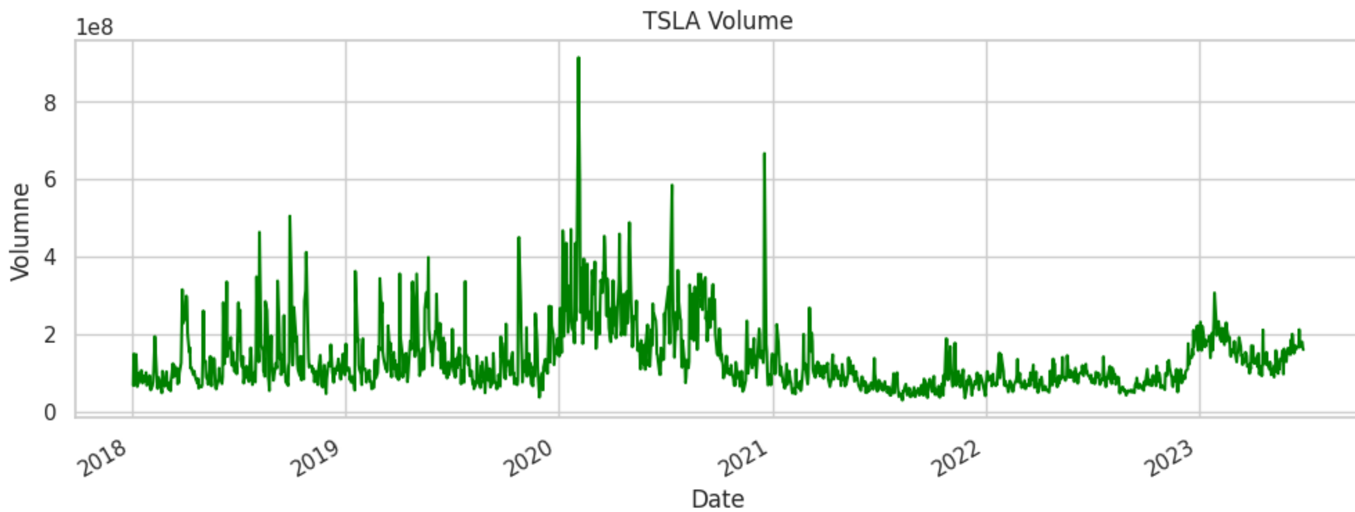

Figure 12- TSLA Volume (2018-2023)

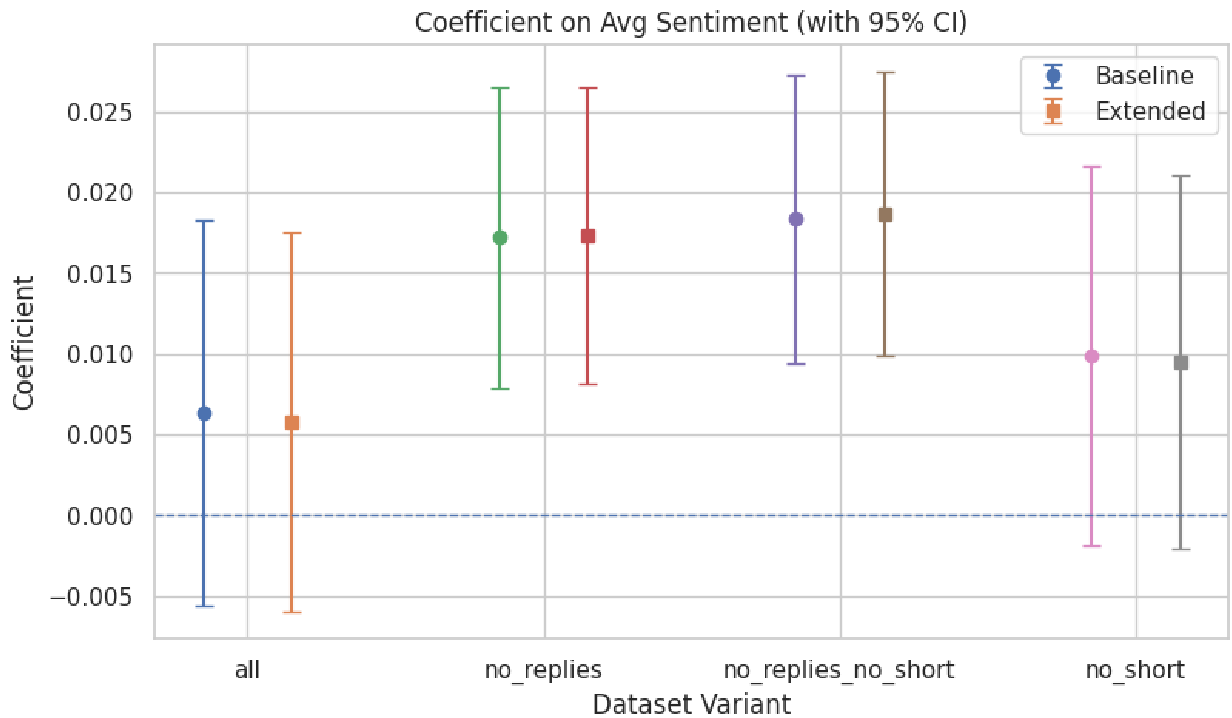

Figure 13- Coefficient on Avg Sentiment with 95% CI