Student name: Kwong Hoi Chan
Student ID: 301608563
Date: 2025/08/13

Report: Analysis of the Impact of Trump's Tweet Activity on the S&P 500

Client: Hedge Fund Research Team

Background

Donald Trump is widely known for his prolific use of Twitter (now renamed as X) as a primary channel for social interaction. During his first term as President of the United States (2017–2021), his tweets often generated immediate media coverage, influenced public sentiment, and occasionally moved financial markets

Now, with Trump having returned to the presidency, there is renewed interest in understanding whether his tweet activity could again influence market movements. For a hedge fund research team, even small, consistent predictive signals could be valuable for short-term trading strategies. If certain patterns in tweet activity are linked to predictable stock market reactions, this could be used to inform market timing decisions or hedge positioning.

This study examines Trump's tweet engagement metrics specifically favorites and retweets to determine whether they correlate with, or can predict, subsequent S&P 500 index returns.

Introduction

The purpose of this analysis was to determine whether Trump's Twitter activity has a measurable and potentially exploitable effect on the S&P 500 index. We also investigated whether tweet engagement metrics (favorites, retweets) could be used as predictive signals for future market returns over short horizons.

We refined the problem to focus on examining the relationships between tweet activity and S&P 500 returns at multiple lags (t+1, t+3, t+5, t+10 days) using correlations, t-tests, Mann–Whitney U tests, ANOVA with Tukey HSD post-hoc comparisons, Chi-square tests, and linear regression, allowing us to assess both linear and non-linear effects. The goal was to determine not just whether a correlation exists, but whether it is strong, consistent, and statistically significant enough to be considered a viable trading signal.

Problem Definition

To make the analysis concrete and actionable, we refined this idea into the following specific objectives:

1. Focus on measurable tweet engagement metrics — favorites and retweets, as indicators of public reaction and reach.
2. Analyze the S&P 500 index returns as the market response variable, given its role as a benchmark for U.S. equity performance.
3. Test multiple short-term lags (t+1, t+3, t+5, t+10 days) to capture both immediate and delayed effects of tweet activity.
4. Apply multiple statistical techniques — correlations, independent t-tests, Mann–Whitney U tests, ANOVA with Tukey HSD post-hoc comparisons, Chi-square tests, and linear regression to detect both linear and non-linear relationships.
5. Assess trading viability by determining whether any detected relationships are strong, consistent, and statistically significant enough to inform market timing or hedging strategies.

This refinement ensured that the analysis remained focused, reproducible, and aligned with the hedge fund team's goal of identifying potential predictive signals from Trump's tweet activity.

Data Collection and Cleaning

1. S&P 500 Data
   Description: Daily historical S&P 500 index price from 2017/01/20 to 2021/01/20
   Source: https://ca.investing.com/indices/us-spx-500-historical-data
   Collect Method: Extract from publicly available financial market data
   Processing Steps:
   a. Convert the date to proper datetime objects
   b. Sorted the dataset in ascending order by date
   c. Convert the close from a string to a numeric value
   d. Add a new column return that computes daily percentage returns
   $$Return_t = \frac{Close_t - Close_{t-1}}{Close_{t-1}}$$
   e. Kept only Date, Close, and Return
2. Trump Tweet Data
   Description: Public dataset of Donald Trump's tweets
   Source: https://www.thetrumparchive.com/faq
   Collect Method: Extract from publicly available archived Twitter data

Processing Steps:

a. Removed retweets (isRetweet flagged as true)
b. Converted the date column to proper datetime objects and removed timezone information
c. Sorted tweets in ascending order by date.
d. Filtered to include only tweets within Trump's presidential term from 2017/01/20 to 2021/01/20
e. Aggregated daily totals for favorites and retweets

Merging Data

Once the S&P 500 daily return data and the aggregated daily Trump tweet engagement data were prepared, the two datasets were merged to form a unified analysis table

Merging Process:

1. Left Join: Merged on the Date column using an inner join to ensure each row represented a market trading day with corresponding tweet activity. Moreover, weekends and market holidays were excluded automatically since the S&P 500 dataset contains only trading days.
2. Handling Missing Values: For trading days where no tweets were posted, favorites and retweets were set to zero. Days with no market data (e.g., weekends) were not included, as they would have no return values.
3. Filtering: Removed rows where both favorites and retweets were zero to focus on days with at least some tweet activity.

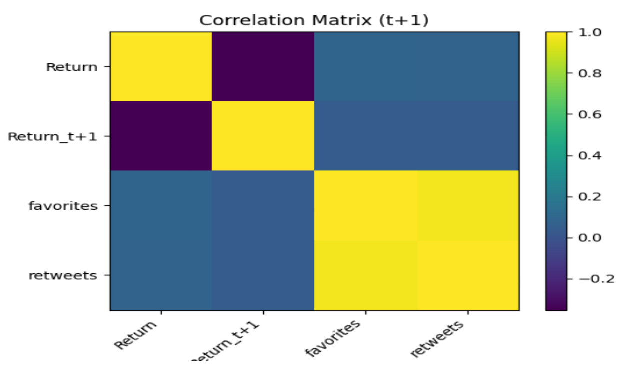
Techniques Used to Analyse the Data

To address the research questions and test for both linear and non-linear relationships, the following statistical methods were applied:

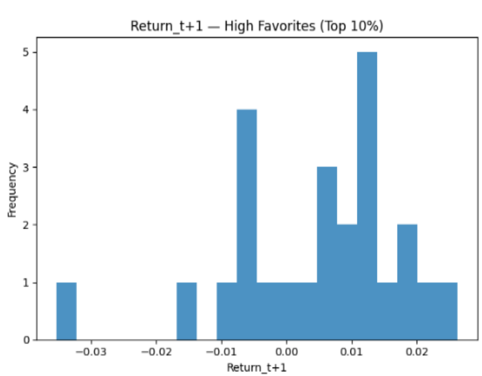
1. Correlation Analysis: Measured the strength and direction of linear relationships between tweet engagement metrics and S&P 500 returns at various lags
2. Normality Testing: Checked whether return distributions followed a normal distribution to guide test selection
3. T-test: Compared mean returns between high-activity (top 10% favorites) and low-activity tweet days under the assumption of normality
4. Mann–Whitney U test: Compared median returns between high and low activity days without assuming normal distribution
5. One-way ANOVA: Tested for differences in returns across low, medium, and high-engagement groups
6. Tukey HSD Post-hoc Test: Identified which specific engagement groups differed when ANOVA results were significant
7. Chi-square Test of Independence: Assessed whether the proportion of positive versus negative return days differed between high and low activity tweet days
8. Linear Regression Analysis: Modeled future returns as a function of tweet favorites to evaluate predictive power

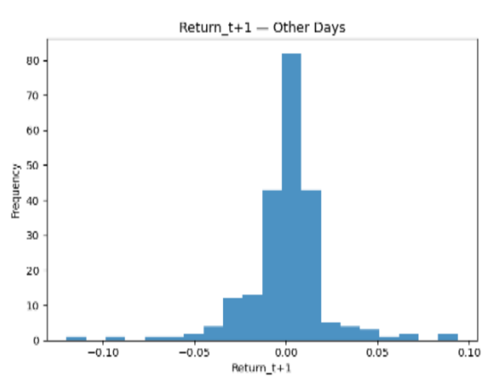
Result

t+1

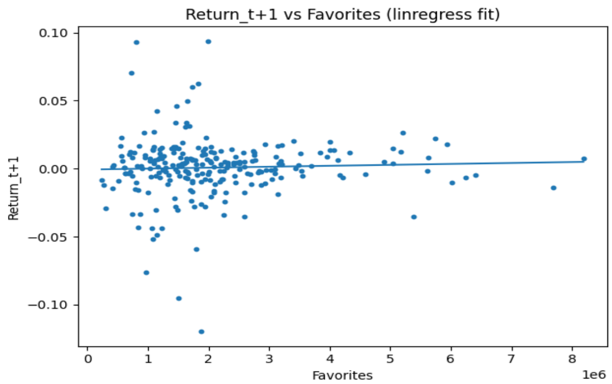


Correlations: Small correlations between favorites and next-day returns (r = 0.041). Retweets almost perfectly correlated with favorites (r = 0.976).





High vs. Low Activity: No significant difference in returns (t-test p = 0.1705; Mann–Whitney p = 0.1108).

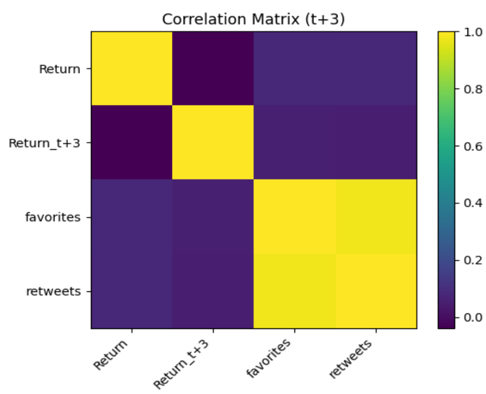


Regression: Very low R^2 (0.0017), slope near zero

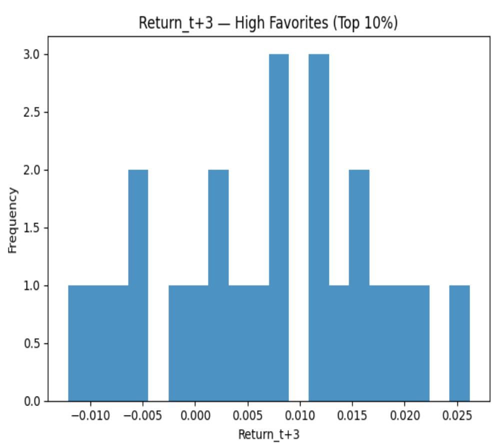
ANOVA: No significant difference across terciles ($p = 0.9598$).

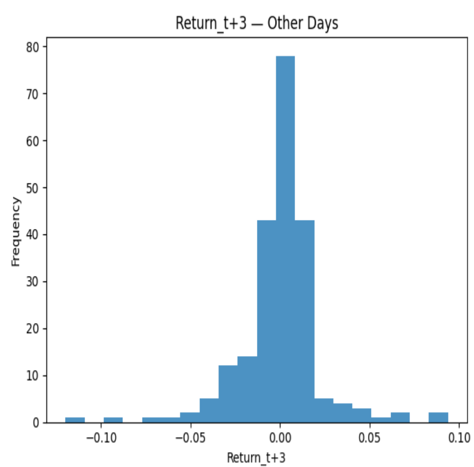
Chi-square: No association between high activity and positive/negative next-day returns ($p = 0.6045$).

t+3

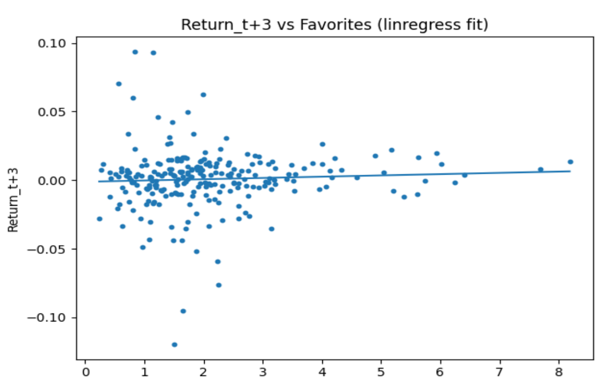


Correlations: Slightly higher correlation ($r = 0.055$) between favorites and 3-day-ahead returns.





High vs. Low Activity: Statistically significant difference (t-test $p = 0.0163$; Mann–Whitney $p = 0.0380$), indicating that high-activity days were associated with different returns 3 days later.

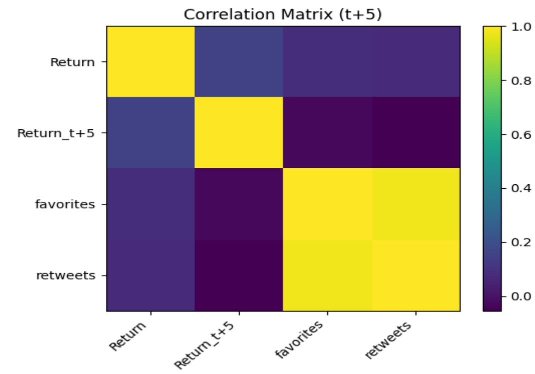


Regression: Very low R^2 (0.0030), slope near zero.

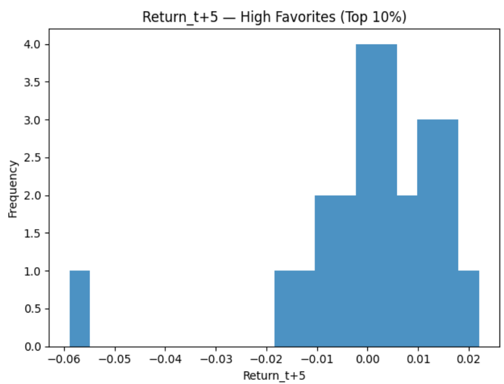
ANOVA: No significant tercile differences ($p = 0.4209$).

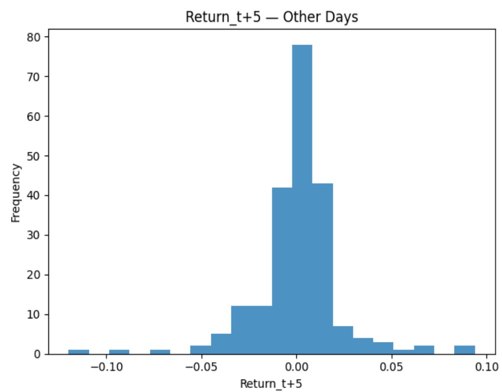
Chi-square: No significant association ($p = 0.3034$).

t+5

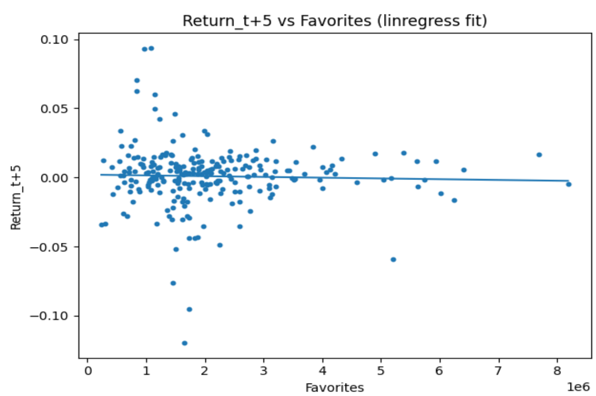


Correlations: Slight negative correlation between favorites and 5-day-ahead returns (r = -0.033).





High vs. Low Activity: No significant difference (t-test p = 0.9074; Mann–Whitney p = 0.6211).

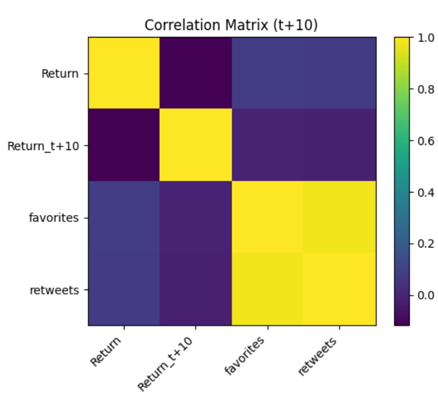


Regression: Very low R^2 (0.0011), slope near zero.

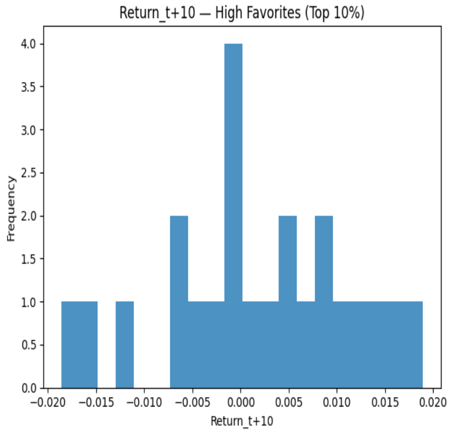
ANOVA: Significant difference across terciles (p = 0.0039), driven by Low vs. Medium group difference (p = 0.0025 in Tukey HSD).

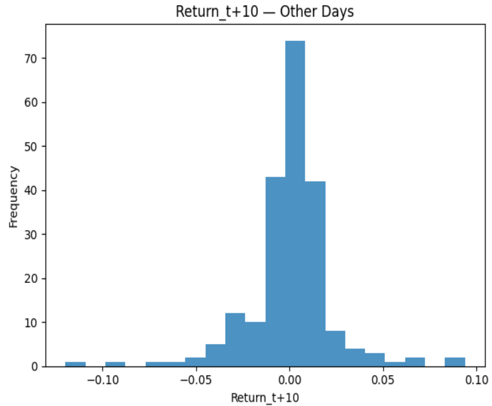
Chi-square: No significant association (p = 0.8045).

t+10

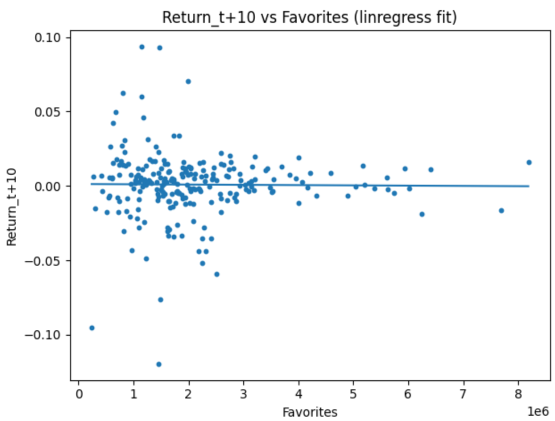


Correlations: Almost zero correlation (r = -0.011).





High vs. Low Activity: No significant difference (t-test p = 0.7329; Mann–Whitney p = 0.9781).



Regression: $R^2 \approx 0.0001$, slope near zero.

ANOVA: No significant tercile differences ($p = 0.1502$).

Chi-square: No significant association ($p = 0.5883$).

Conclusion

This analysis provides no strong or consistent evidence that Donald Trump's tweet engagement that measured through favorites and retweets where can be used to predict S&P 500 returns in the short term. While the t+3 lag produced a statistically significant difference in returns between high and low activity days, the magnitude of the effect was small, limited to a single lag, and unlikely to form the basis of a dependable trading strategy. Moreover, retweets were found to be almost perfectly correlated with favorites, offering no additional predictive power. Linear regression analyses further confirm that tweet engagement explained virtually none of the variation in future returns. Taking together, the evidence suggests that engagement volume alone without considering the content, sentiment, or context of the tweet is not a viable standalone indicator for market timing, short-term prediction in an investment strategy.

1. No consistent relationship between tweet engagement and market returns across lags.
2. Single significant result at t+3 likely due to chance, not a persistent pattern.
3. Retweets highly correlated with favorites, offering no extra predictive power.
4. Linear regression models had near-zero explanatory value.

Limitation

Several factors limit the scope and generalizability of this study. The dataset covers only Trump's first presidential term, restricting the number of high-engagement days and potentially missing patterns that could emerge over a longer period. The study also focused solely on engagement counts, without considering tweet content or sentiment, and only tested fixed-day lags, which may overlook other reaction timeframes. Finally, multiple hypothesis testing raises the risk of false positives, and external market events may confound the observed relationships. To sum up:

1. Limited sample size, especially for high-activity tweet days.
2. No analysis of tweet content, sentiment, or topic relevance.
3. Only fixed lags tested; more complex timing effects may exist.
4. Multiple comparisons increase false positive risk.
5. Potential confounding from unrelated macroeconomic or political events.