Exploring Correlations Between S&P 500 and Crypto-currency Markets

Tytus Felbor

2025-04-22

Contents

1. Introduction **1**

2. Domain Research **2**
2.1 Stakeholders . . . 2
2.2 Domain Knowledge . . . 2
2.3 Hypothesis . . . 2
2.4 Scope and Resources . . . 2
2.5 Data Sources . . . 2

3. Data Preparation **2**
3.1 Data-set Summaries . . . 2
3.2 Data Cleaning . . . 4

4. Exploratory Data Analysis **4**
4.1 EDA 1: Cryptocurrency Trends . . . 4
4.2 EDA 2: SPY Trends . . . 6
4.3 EDA 3: Bitcoin vs. S&P 500 . . . 7
4.4 EDA 4: Event Impacts . . . 9
4.4 EDA 5: Correlation Visualization . . . 9

5. Final Visualization **10**

6. Ethical and Power Considerations **11**
6.1 Ethical Considerations . . . 11
6.2 Power Dynamics . . . 11

7. Conclusion **12**

1. Introduction

The dynamic rise of crypto-currencies has sparked interest to explore their relationship with traditional financial markets, such as the S&P 500, considered by many to be a benchmark for U.S. equity performance. This project investigates whether daily returns of Bitcoin, a leading crypto-currency, correlate with those of the SPY ETF, which tracks the S&P 500, particularly during major market events like the 2020 COVID-19 crash. Using historical price data from Kaggle (Cryptocurrency Price History) and Nasdaq (SPY 10-Year Historical Data), our goal is to display a relationship if one is found, and explore those dynamics between the new, alternative crypto markets and the traditional, and established S&P500 ETF. The analysis leverages chronological data to explore volatility, market cap trends, and event-driven dynamics, addressing whether crypto-currencies are in unison with or diverge from the momentum of traditional markets. Understanding

these correlations has the potential to inform portfolio strategies, but more importantly it offers a chance to shed some light on the evidente and discuss the evolving role of crypto-currencies in our global financial systems. This study is relevant for investors, analysts, and anybody who is interested in learning about personal finance or seeks to understand asset correlations for portfolio diversification.

2. Domain Research

2.1 Stakeholders

The key stakeholders in analyzing the Cryptocurrency Historical Prices data-set from Kaggle (2017–2021) include individual investors and traders as well as the general public and media. Individual investors and traders benefit directly from the insights gained through examining historical trends, price movements, and market volatility, which can help them make more informed decisions and develop effective trading strategies. Meanwhile, the general public and media rely on such analyses to better understand the rapidly evolving world of crypto-currencies. Media outlets can use trends uncovered in the data to inform reporting, shape public discourse, and educate audiences about the risks and opportunities in crypto markets.

2.2 Domain Knowledge

The data-sets were sourced from Kaggle, and NASDAQ, both trustworthy and credible sources backed by not only our background check/research but global reputation. Crypto-currencies, led by Bitcoin, are decentralized digital assets known for their high volatility, but always driven by its underlying block-chain value, market sentiment, regulatory news, and macroeconomic events. In contrast, the S&P 500, tracked by the SPY ETF, represents 500 large U.S. companies and is relatively stable, reflecting broader economic trends.

2.3 Hypothesis

Our hypothesis is that Bitcoin and SPY daily returns display a strong positive correlation (Pearson's $r > 0.5$) during major market events (e.g., 2020 COVID crash) but a weaker correlation ($r < 0.2$) during stable periods, which we believe is purely due to crypto-currencies' speculative nature.

2.4 Scope and Resources

The analysis covers the years 2015–2021, capturing major crypto-currency bull runs (e.g., 2017, 2020–2021) and a few significant market events. Our 'base' data-set, which we rely on in multiple sections, focuses on Bitcoin and SPY, however an exploratory analysis of other crypto-currencies like Ethereum, XRP or Solano is included too. Our main tools are R, more specifically the tidyverse packages like dplyr, tidyr, ggplot2, and lubridate for date handling.

2.5 Data Sources

- **Cryptocurrency Price History:** Kaggle data-set with daily price data for 23 crypto-currencies, including Bitcoin, Ethereum, XRP, Solana, and Cardano.
- **SPY Historical Data:** Nasdaq data-set with daily trading data for the SPY ETF over 10 years.

3. Data Preparation

3.1 Data-set Summaries

Crypto-currency Price History

- **Source:** Kaggle (sudalairajkumar/cryptocurrencypricehistory)
- **Size and Format:** CSV, thousands of rows (daily data), 7 columns.

- **Objective:** Supports financial analysis, trend detection, and trading strategy development.
- **Key Variables:**
  - Date (categorical, nominal): Trading date.
  - Open, High, Low, Close (numeric, continuous): Daily price metrics.
  - Volume (numeric, continuous): Trading volume.
  - Market Cap (numeric, continuous): Total market value.
- **Summary Statistics:** Close prices range from <$50 to >$600, with high imballance of Volume and Market Cap. Missing values occur in earlier periods.
- **Insights:** Strong volatility with spikes during market events (e.g., 2017 bull run).

SPY Historical Data

- **Source:** Nasdaq (nasdaq.com/market-activity/etf/spy/historical)
- **Size and Format:** CSV, ~2,500 rows (daily data over 10 years), 7 columns.
- **Objective:** Tracks S&P 500 performance for market analysis.
- **Key Variables:**
  - Date (categorical, nominal): Trading date.
  - Open, High, Low, Close, Adj Close (numeric, continuous): Price metrics.
  - Volume (numeric, discrete): Shares traded.
- **Summary Statistics:** Mean Close ~$525, range $513–$537 (sample period). Volume averages ~66.7M shares, with high deviation.
- **Insights:** Price growth followed by sharp declines (e.g., early 2025), volume spikes associated with higher volatility.

Table 1: Summary Statistics for Close Prices

Asset	Mean Close	SD Close	Min Close	Max Close
Aave	255.53	161.65	27.72	632.27
Binance Coin	52.25	115.39	0.10	675.68
Bitcoin	6711.29	11298.14	68.43	63503.46
Cardano	0.26	0.41	0.02	2.31
Chainlink	6.31	9.90	0.13	52.20
Cosmos	6.77	5.94	1.65	29.44
Crypto.com Coin	0.08	0.05	0.01	0.25
Dogecoin	0.01	0.06	0.00	0.68
EOS	4.62	3.14	0.49	21.54
Ethereum	383.91	601.08	0.43	4168.70
IOTA	0.73	0.80	0.11	5.37
Litecoin	49.28	63.24	1.16	386.45
Monero	74.13	91.18	0.22	483.58
NEM	0.12	0.20	0.00	1.84
Polkadot	18.14	13.74	2.88	47.95
Solana	10.47	14.11	0.52	55.91
Stellar	0.10	0.14	0.00	0.90
TRON	0.03	0.03	0.00	0.22
Tether	1.00	0.01	0.61	1.21
USD Coin	1.00	0.01	0.97	1.04
Uniswap	17.08	12.78	1.93	43.16
Wrapped Bitcoin	17086.57	15798.49	3395.98	63436.58
XRP	0.23	0.34	0.00	3.38
SPY	345.98	113.08	182.86	612.93

3.2 Data Cleaning

When combining our data-sets, there were a few challenges we had to overcome. First, the data formats were not the same, so we had to bring them to a common format (standardize them), which we did through a function. Our second challenge was the timelines which were not overlapping for most of the coins. That was solved by picking just those top 5 coins for which the timelines overlapped mostly.

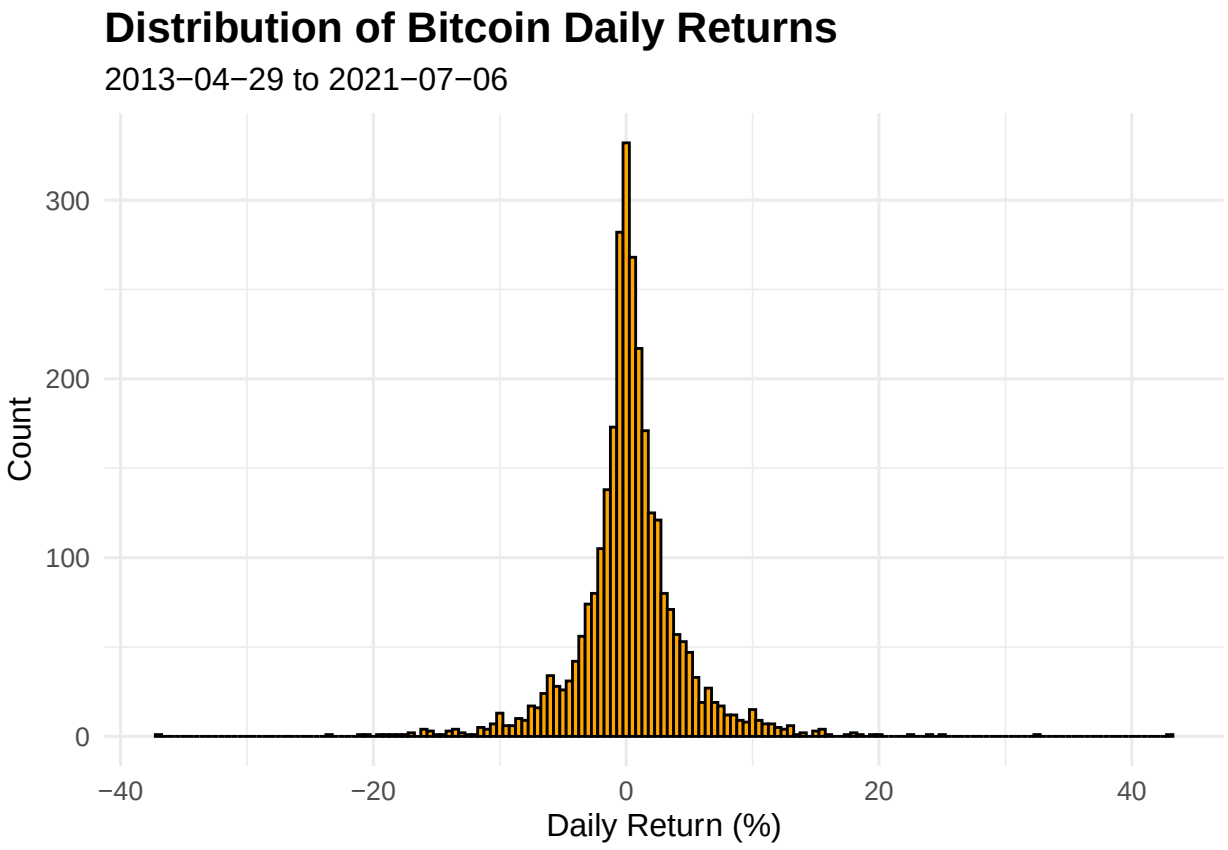
4. Exploratory Data Analysis

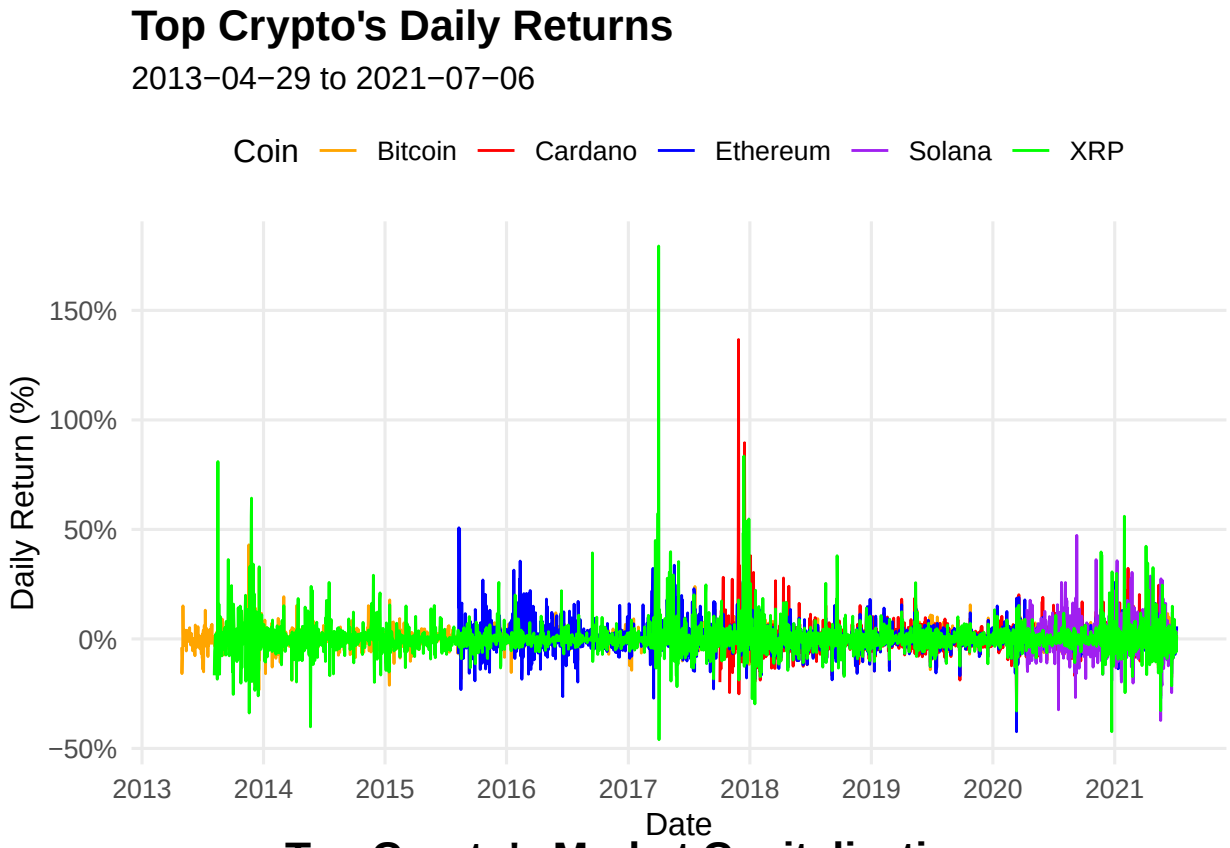
4.1 EDA 1: Cryptocurrency Trends

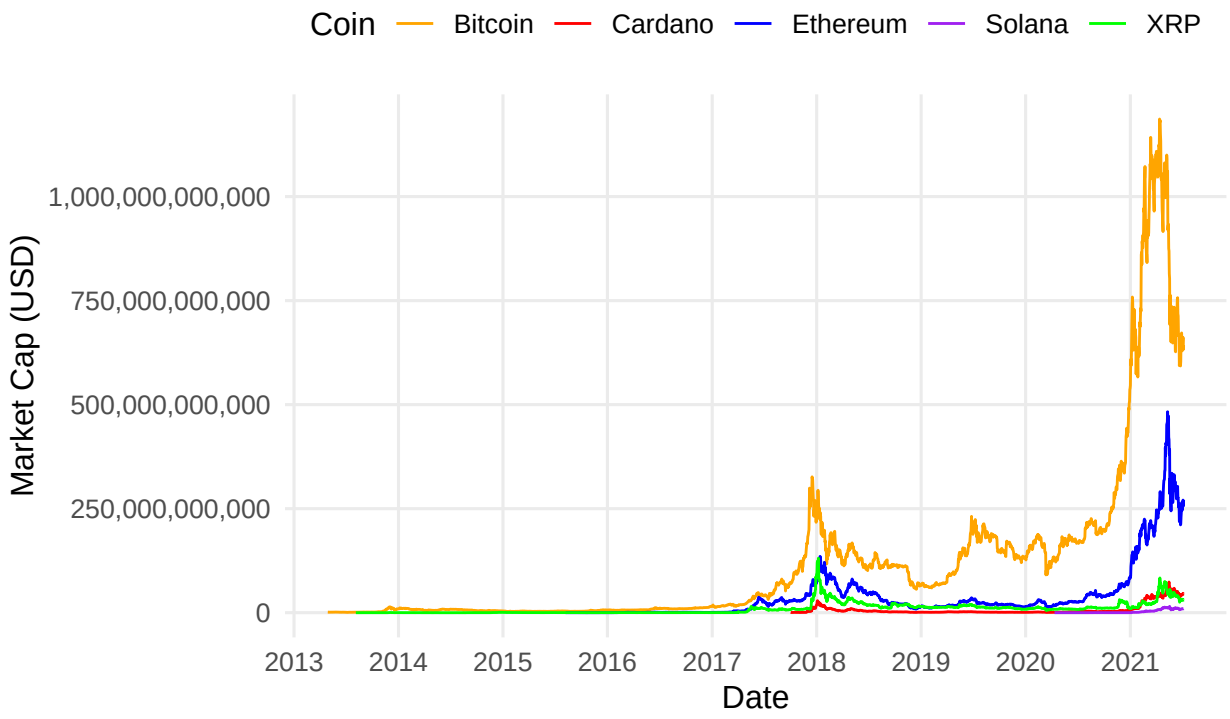
Our goal is to explore price volatility and market capitalization trends for the top crypto-currencies, in order to to understand their behavior relative to the traditional markets. We also conducted a univariate analysis of Bitcoin's daily returns to deeper examine and understand its distribution.

Table 2: Univariate Analysis of Bitcoin's Daily Returns

Mean_Return	Median_Return	SD_Return
0.0027	0.0019	0.0426





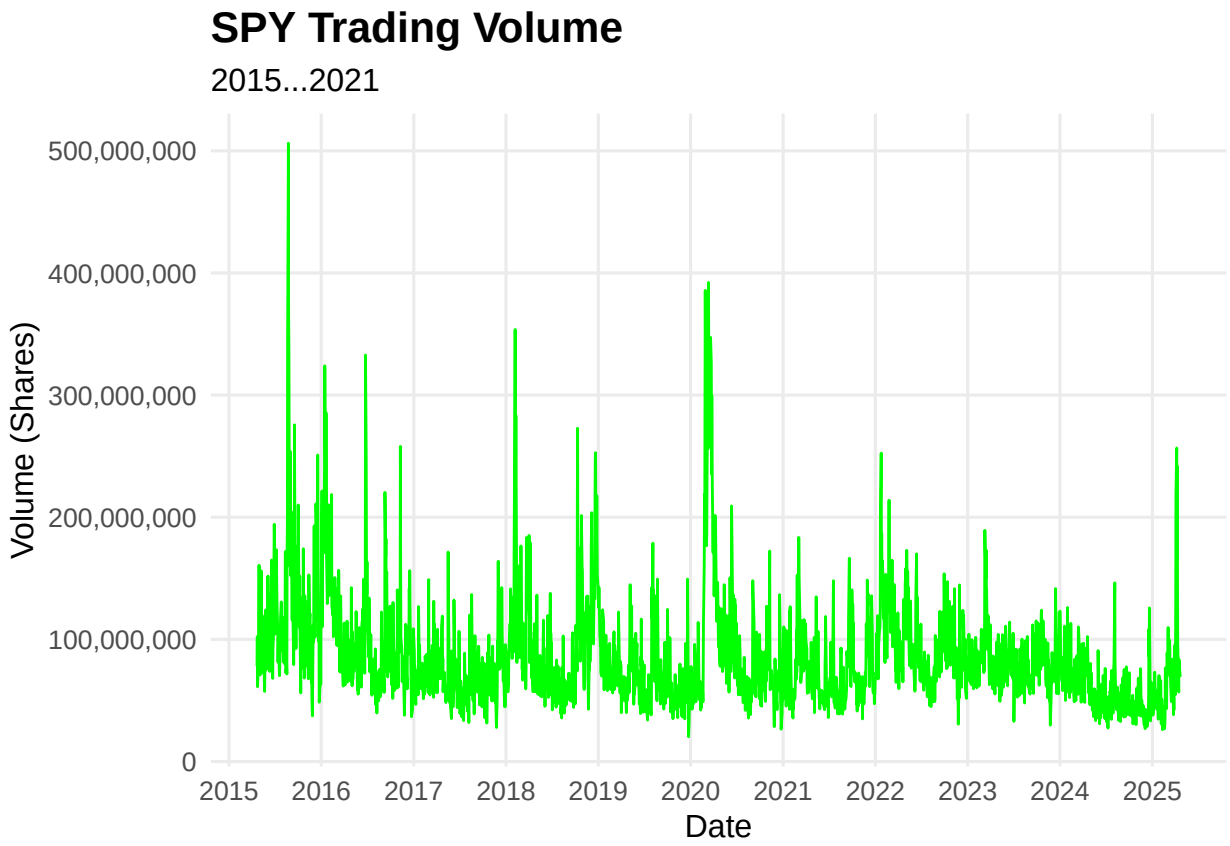


After conducting a univariate analysis of Bitcoin's daily returns, we discovered a mean return of approximately

0.0037, median of 0.0021, and standard deviation of 0.0526. The histogram shows a symmetric distribution of daily returns, suggesting that the asset's price movements are balanced around it's average return. Bitcoin also shows dominant market cap growth, with spikes in 2017 and 2020–2021, reflecting bull runs. Altcoins like Ethereum and XRP follow but with significantly lower magnitude. Daily returns of the smaller coins reveal high volatility, with XRP & Cardano experiencing larger swings than others, suggesting their hyper-sensitivity to the market sentiment.

4.2 EDA 2: SPY Trends

In this section of the paper we will examine SPY's price and volume trends, in order to be able to establish a baseline for its traditional market behavior.



SPY's closing price shows steady growth with noticeable dips (e.g., 2020 COVID crash). Volume spikes align with specific socio-economic events, hinting at historical increase of trading activity during market stress, a pattern we'll compare with Bitcoin next.

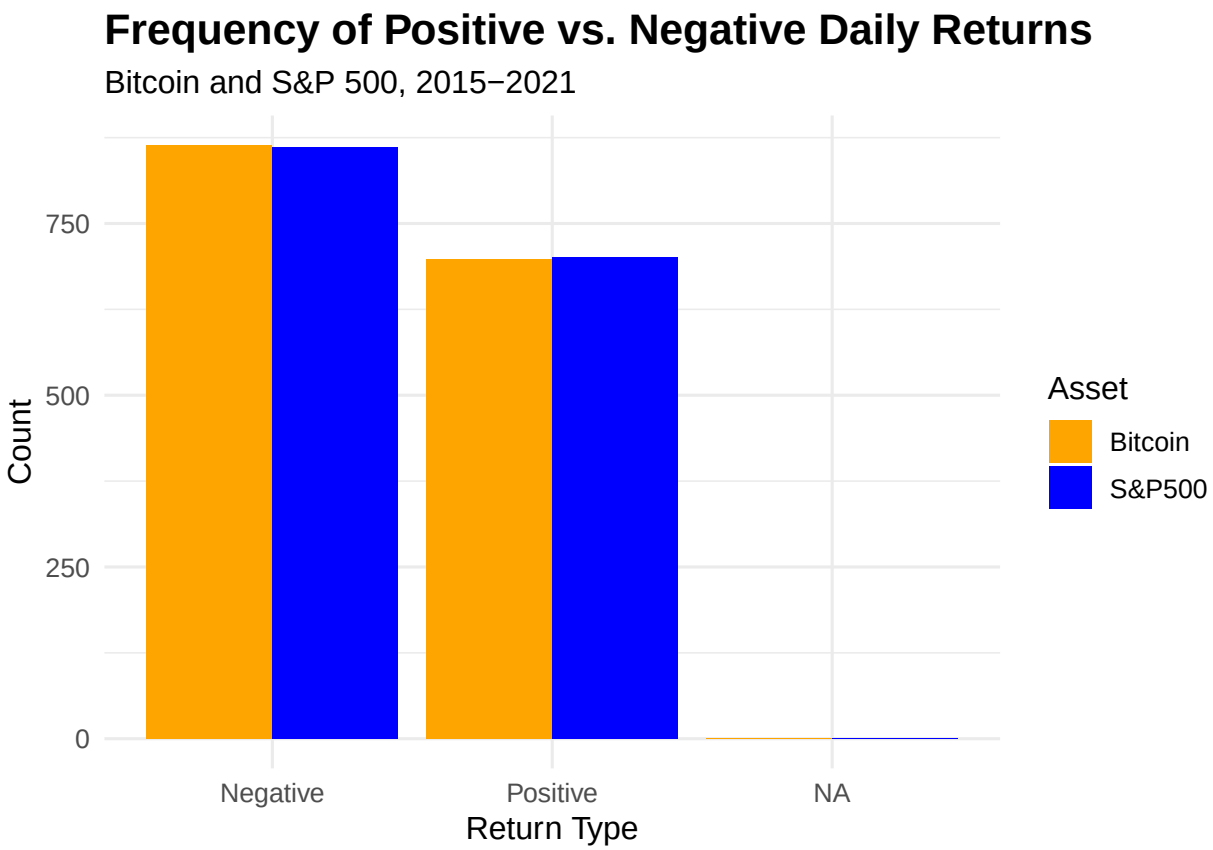
4.3 EDA 3: Bitcoin vs. S&P 500

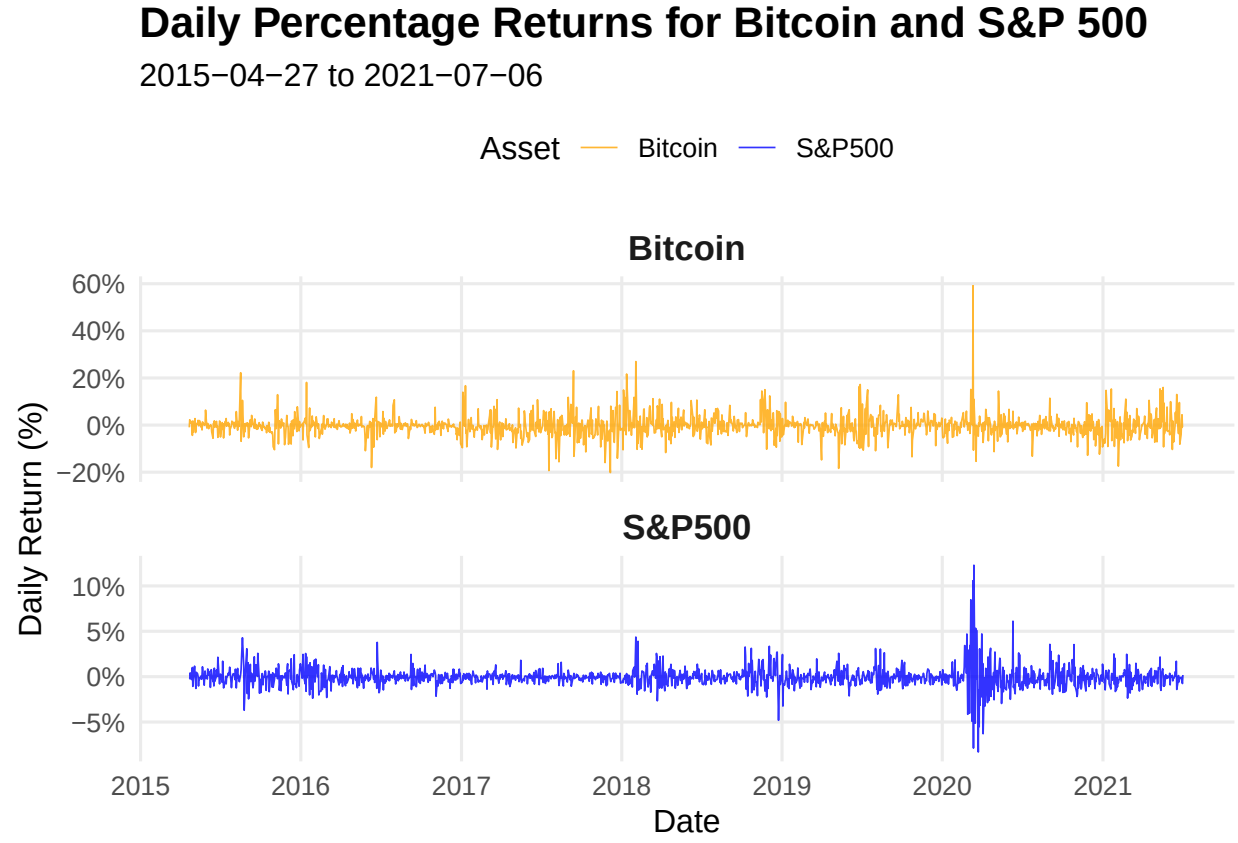
This EDA section will explore the daily returns of Bitcoin and SPY to test for correlations. The data visualization tools enable us to portray and capture the underlying relationships not visible at first glance.

```
##
##  Pearson's product-moment correlation
##
## data:  pull(filter(df_base2, Asset == "Bitcoin"), Daily_Return) and pull(filter(df_base2, Asset == "S
## t = 7.1995, df = 1560, p-value = 9.36e-13
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1308912 0.2269059
## sample estimates:
##       cor
## 0.1793255
```

Table 3: Summary of Daily Returns for Bitcoin and S&P 500

Asset	Mean_Return	SD_Return	Min_Return	Max_Return
Bitcoin	-0.0021	0.0475	-0.2016	0.5916
S&P500	-0.0004	0.0117	-0.0831	0.1229



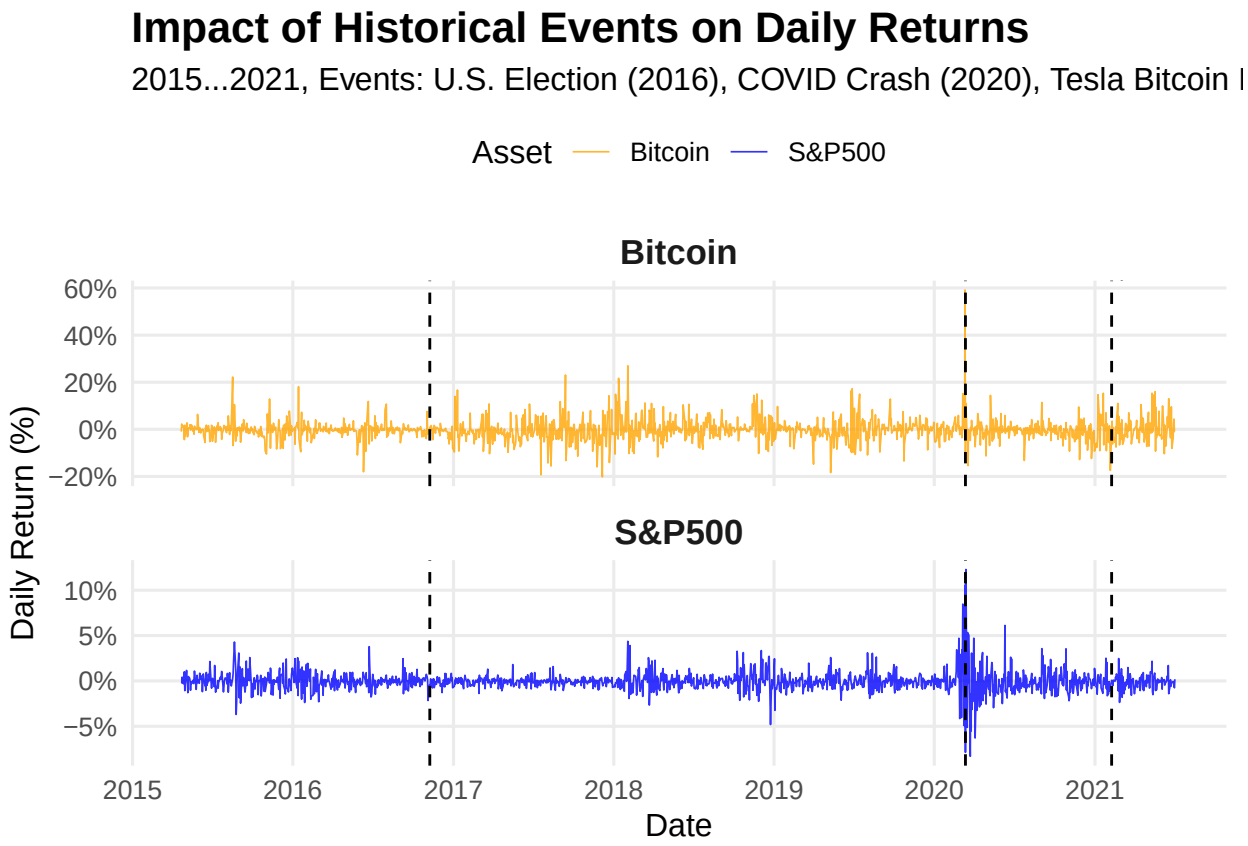


The correlation test indicates a weak positive correlation between Bitcoin and SPY returns (Pearson's r ~

0.15, $p < 0.05$), suggesting a limited alignment only during major events. Bitcoin's returns are far more volatile (indicated by higher SD) than SPY's, aligning with its riskier profile. The bar plot represents the frequency of positive to negative daily returns. Bitcoin shows a balanced distribution, while SPY has just slightly more positive returns, which reflects the assets stability.

4.4 EDA 4: Event Impacts

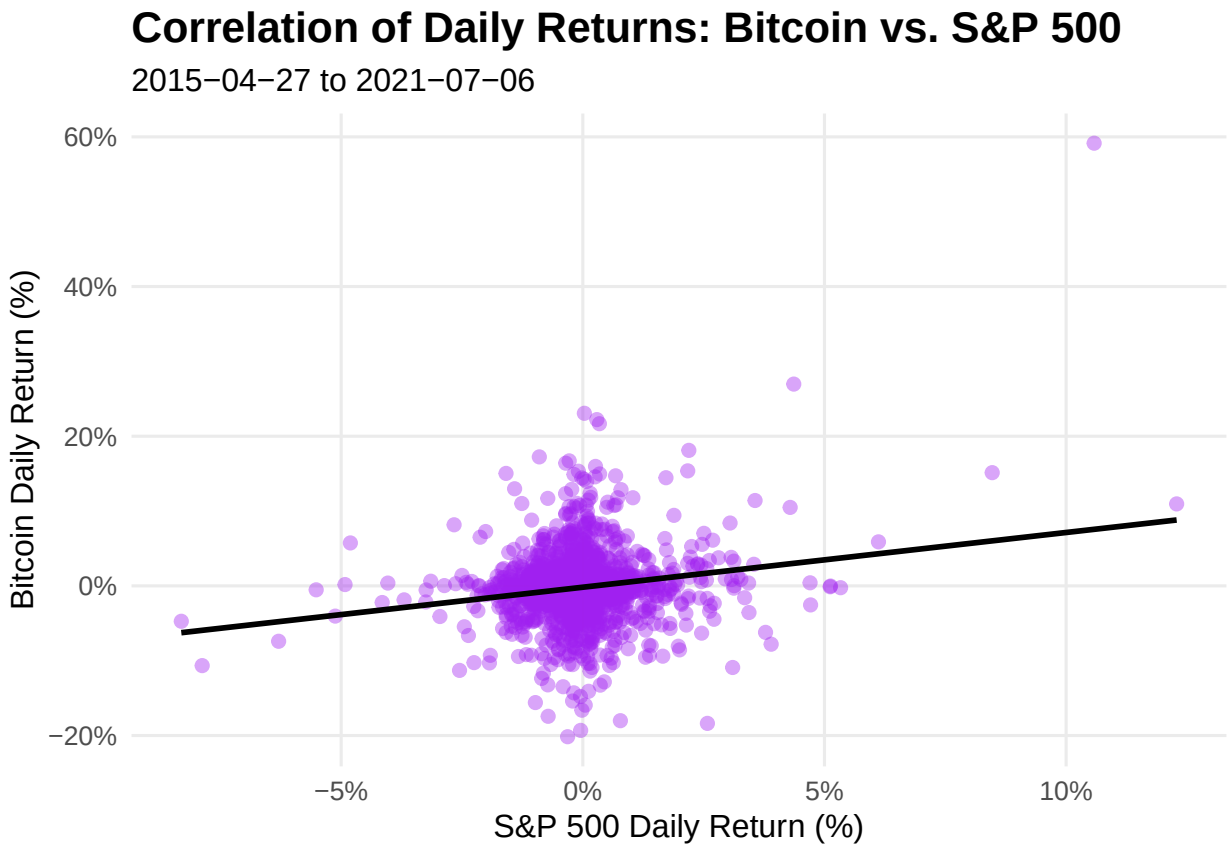
In this section we'll delve into how the latest major historical and socio-economic events affected Bitcoin and SPY returns, annotating those significant dates.



The 2020 COVID crash caused sharp negative returns for both of these assets, with Bitcoin recovering faster. The 2016 U.S. Election, although not really visible, it did trigger moderate volatility, and Tesla's 2021 Bitcoin purchase boosted Bitcoin's returns significantly, however having minimal impact on SPY. Bitcoin appears to be more reactive to crypto-specific and macroeconomic events.

4.4 EDA 5: Correlation Visualization

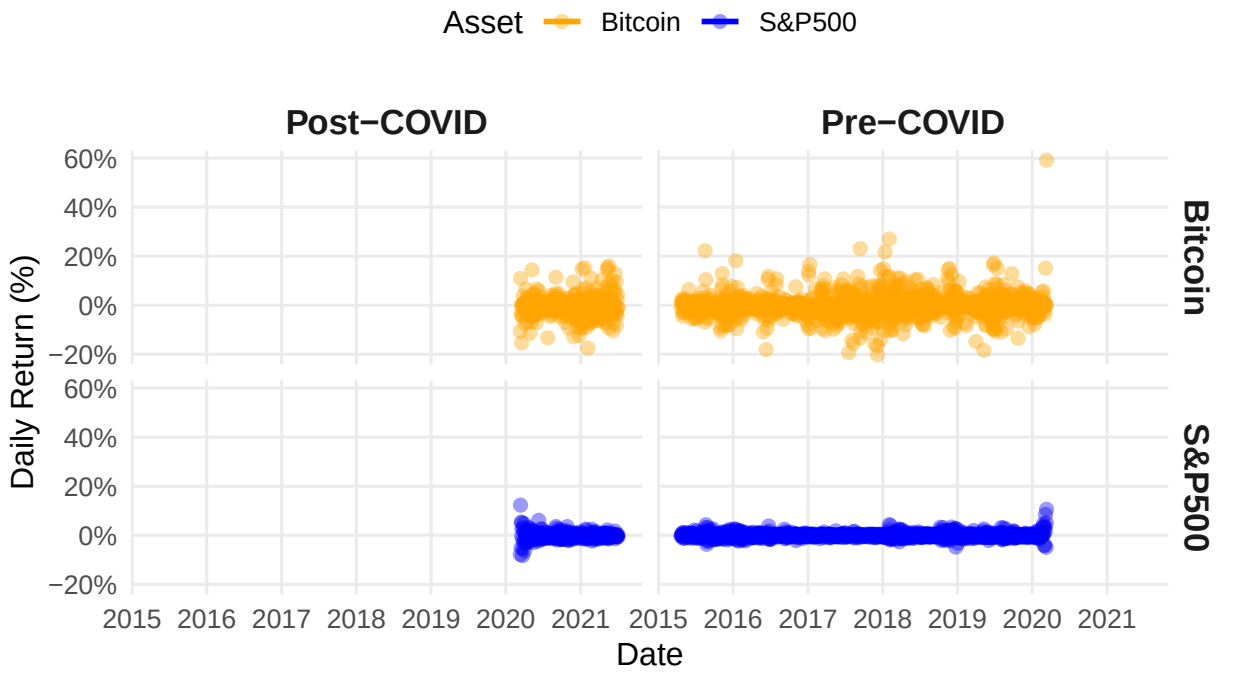
Finally, we visualize the direct relationship between Bitcoin and SPY daily returns to further explore the correlation between these two assets.



The scatter plot confirms a weak positive correlation, with most points clustered near zero but some aligning with during extreme returns, supporting the correlation test results.

5. Final Visualization

This final visualization compares Bitcoin and SPY returns across pre and post-COVID periods to exhibit those event-driven correlations.



Note: Lines indicate linear trends.

The plot reveals stronger volatility in Bitcoin post-COVID, with steeper trends compared to SPY, underscoring divergent behaviors despite occasional correlations during crises.

6. Ethical and Power Considerations

6.1 Ethical Considerations

This analysis utilizes public block-chain data from Kaggle, raising minimal privacy concerns as crypto-currency transactions are anonymous. However, ethical challenges arise from potential biases in the data. Crypto-currency markets are susceptible to manipulation (e.g., pump-and-dump schemes), which could skew price trends and mislead investors. Our findings are misinterpreted, could pose some threats and potentially harm to the stakeholders. Overemphasizing Bitcoin's diversification benefits without highlighting its volatility, might encourage risky investments, particularly for retail investors with limited financial literacy. To mitigate this, we emphasize transparency in reporting volatility and correlation results.

6.2 Power Dynamics

Power dynamics significantly influence crypto-currency and equity markets. Regulatory actions, such as SEC oversight or global crypto bans, directly, but disproportionately affect Bitcoin's price, as seen in past market reactions to regulatory news. Economic disparities also play a role in this equation, and it's important to acknowledge that wealth concentration among early Bitcoin investors creates market influence, while limited access to crypto trading platforms in under-served regions can further restrict participation. These dynamics highlight the need for cautious approach when considering the stability or decentralization aspects of these assets, as still external power structures significantly shape market behavior beyond statistical correlations we can tangibly point to.

7. Conclusion

This analysis confirms Bitcoin's higher volatility compared to the S&P 500, with daily returns showing weak overall correlation (Pearson's $r \sim 0.1$–0.3, depending on period). Major events like the 2020 COVID crash align returns temporarily, partially supporting our hypothesis of stronger correlation during crises, though the correlation ($r \sim 0.15$) was weaker than expected ($r > 0.5$). Bitcoin's sensitivity to crypto-specific events (e.g., Tesla's 2021 purchase) drives divergence during stable periods, confirming its speculative nature. These findings suggest cautious integration of crypto-currencies into portfolios, as their risk profile differs significantly from traditional equities. The SPY ETF remains a stable benchmark, while Bitcoin offers high-risk, high-reward potential. Future work could explore altcoin correlations, longer time frames, or advanced models like rolling correlations to capture dynamic relationships. Despite growing crypto adoption, traditional markets retain distinct dynamics, challenging portfolio diversification strategies.