2024-2025 Spring

DSA210 Term Project

Final Report

Bora Erbay, 30638

Investigating the Relationship Between Cryptocurrency Prices and Solar Activity

Introduction

In this project, I set out to explore whether there's any relationship between the fluctuations in cryptocurrency prices—specifically Bitcoin and Ethereum—and solar activity, as measured by daily sunspot counts. The idea was sparked by speculative claims that environmental factors such as solar storms might influence human behavior and, by extension, financial markets.

Data Sources

I used the following datasets:

- **Bitcoin price history:** daily open, high, low, close prices.
- **Ethereum price history:** similar format and source.
- **Daily sunspot observations:** covering years from 1850 to 2025.

Data Preparation

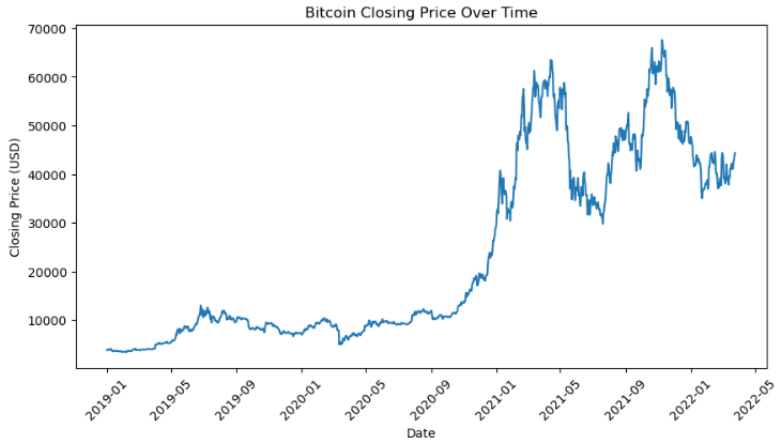
I started by loading and cleaning each dataset. This involved:

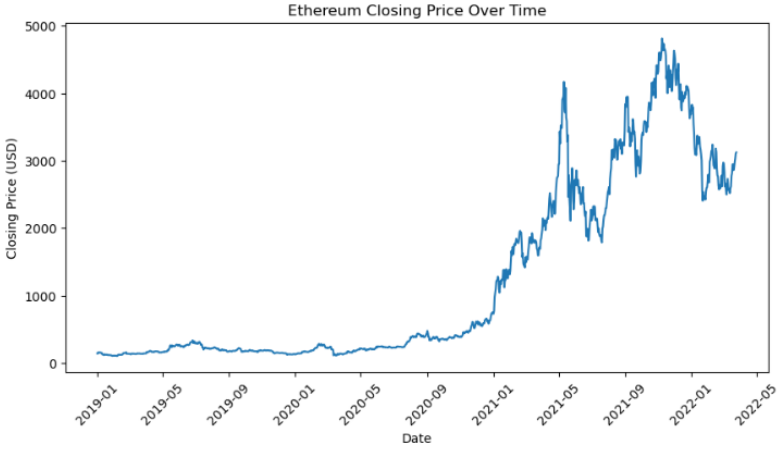
- Ensuring consistent date formats.
- Removing missing values.
- Merging datasets on the date column so that each row reflected the price of both cryptocurrencies and sunspot count on that day.

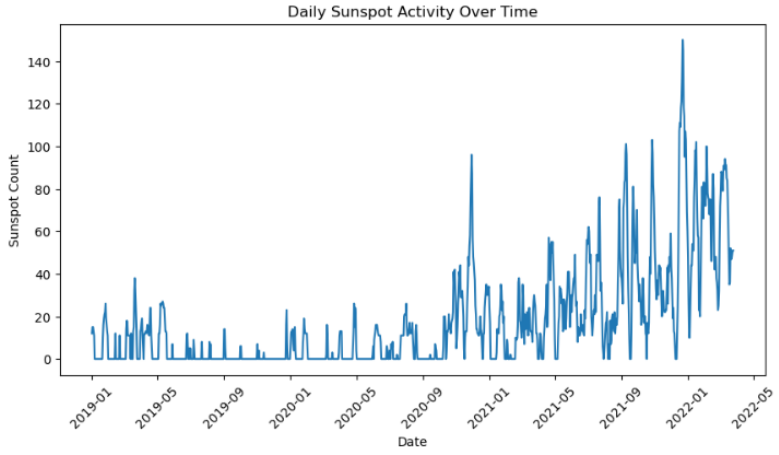
```
Merged DataFrame:
        Date     Open_btc     High_btc      Low_btc    Close_btc  \
0 2019-01-01  3746.713379  3850.913818  3707.231201  3843.520020
1 2019-01-02  3849.216309  3947.981201  3817.409424  3943.409424
2 2019-01-03  3931.048584  3935.685059  3826.222900  3836.741211
3 2019-01-04  3832.040039  3865.934570  3783.853760  3857.717529
4 2019-01-05  3851.973877  3904.903076  3836.900146  3845.194580

   Adj Close_btc  Volume_btc    Open_eth    High_eth     Low_eth  ...  \
0    3843.520020  4324200990  133.418152  141.397507  132.650711  ...
1    3943.409424  5244856836  141.519516  156.929138  140.650955  ...
2    3836.741211  4530215219  155.196045  155.863052  147.198364  ...
3    3857.717529  4847965467  148.912888  156.878983  147.907104  ...
4    3845.194580  5137609824  154.337418  160.824890  154.337418  ...

   Adj Close_eth  Volume_eth  year  month  day  date_frac  counts  std  nobs  \
0     140.819412  2258709868  2019      1    1   2019.001      12  1.0    25
1     155.047684  3328240369  2019      1    2   2019.004      15  1.5    31
2     149.135010  2676164880  2019      1    3   2019.007      15  1.2    20
3     154.581940  3126192535  2019      1    4   2019.010      13  0.8    32
4     155.638596  3338211928  2019      1    5   2019.012      11  1.4    17
```

Exploratory Data Analysis

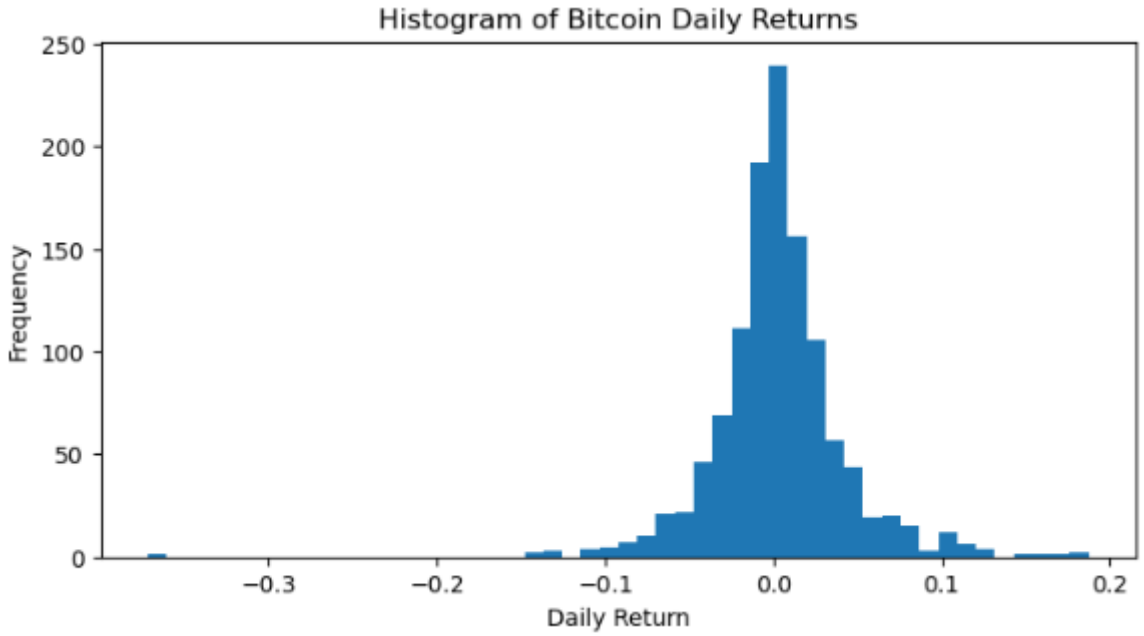
I visualized the trends in price movements and sunspot activity to look for any obvious patterns.

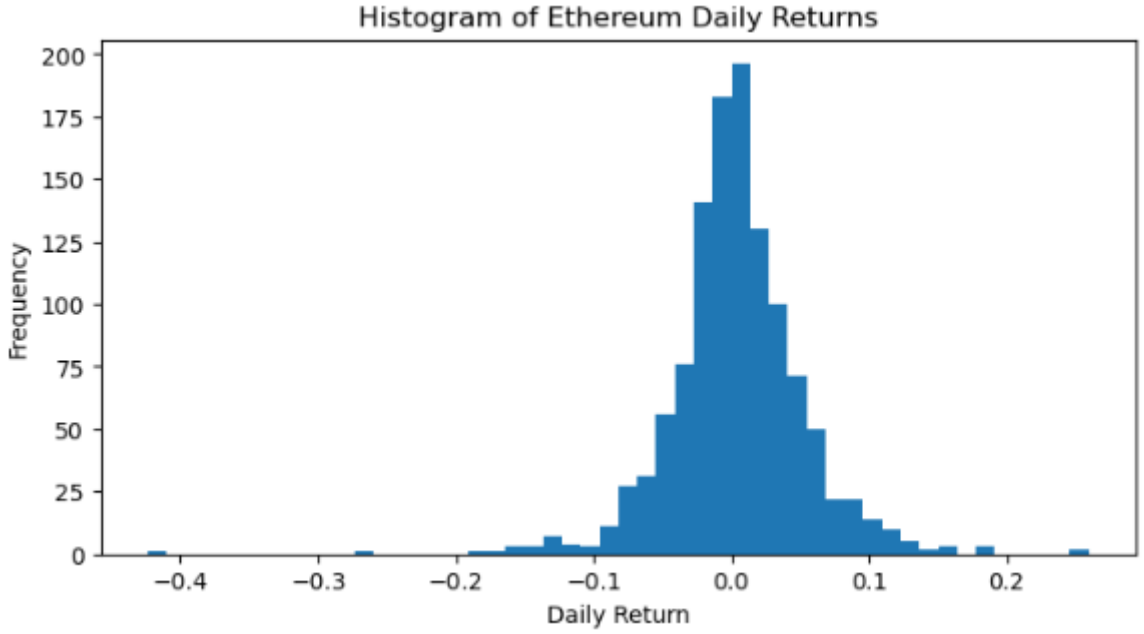






I also calculated daily returns for Bitcoin and Ethereum to analyze percentage-based movements instead of raw price changes.





Statistical Analysis

To test the hypothesis of a relationship between sunspot activity and cryptocurrency returns, I performed:

- **Pearson correlation tests** between sunspot counts and daily returns.
- **T-tests** comparing returns on days with high vs. low solar activity.

```
Bitcoin Daily Returns and Sunspot Activity:
Pearson Correlation: -0.0031, p-value: 0.9161

Ethereum Daily Returns and Sunspot Activity:
Pearson Correlation: 0.0118, p-value: 0.6856
```

```
T-test for Bitcoin Daily Returns (High vs. Low Sunspot Activity):
t-statistic: 0.2927, p-value: 0.7698

T-test for Ethereum Daily Returns (High vs. Low Sunspot Activity):
t-statistic: 0.5664, p-value: 0.5713
```

Machine Learning Models

I used multiple models to predict Bitcoin prices based on Ethereum prices and sunspot activity:

- Linear Regression:

```
Linear Regression Results:
MSE: 56423499.13638152
R² Score: 0.854092526929975
```

- Random Forest Regressor:

```
Random Forest Results:
MSE: 17384602.800481837
R² Score: 0.9550445558363317
```

- K-Nearest Neighbors (KNN):

```
KNN Regressor Results:
MSE: 19093538.963843737
R² Score: 0.9506253588519082
```

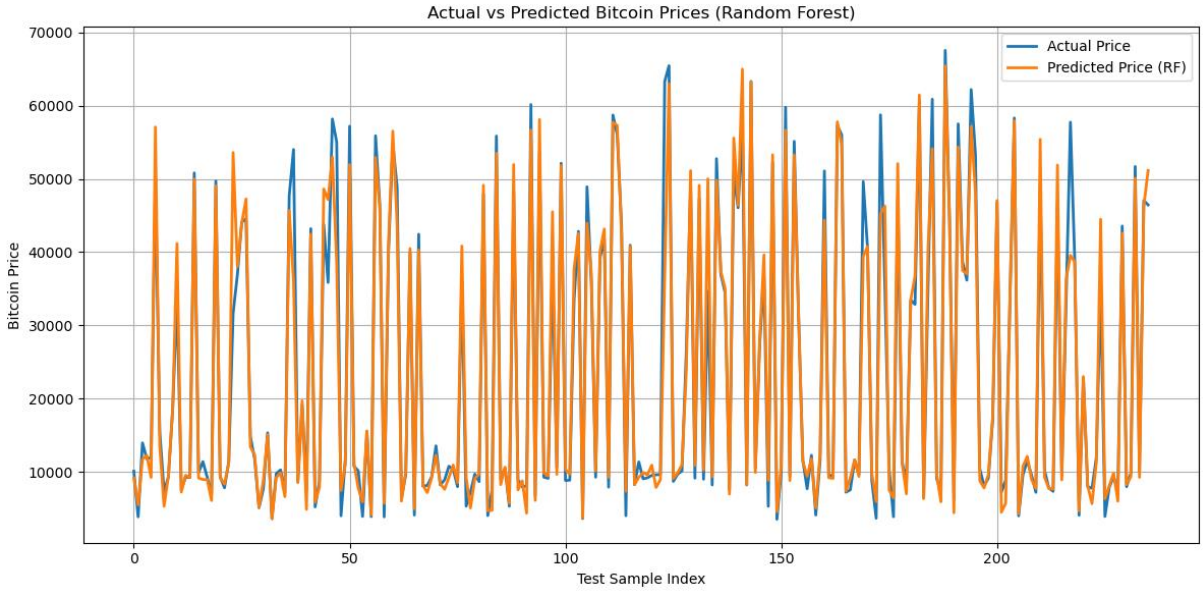
- Decision Tree Regressor:

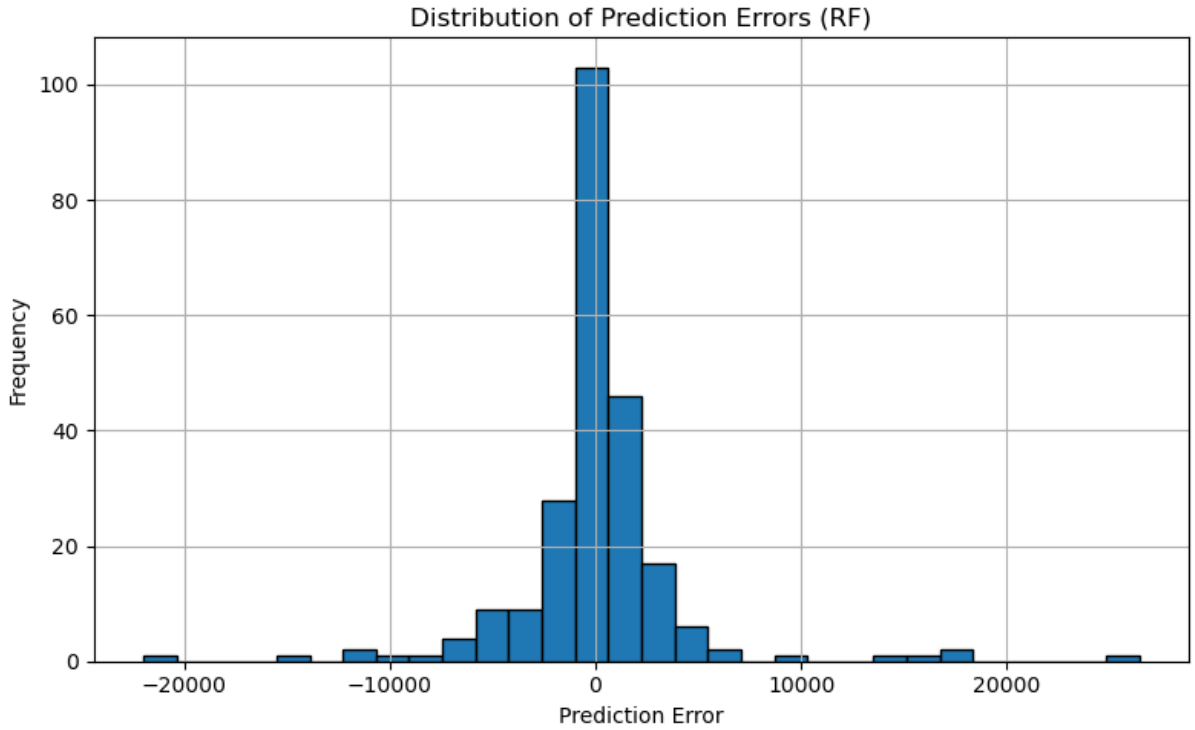
```
Decision Tree Results:
MSE: 29297819.712060064
R² Score: 0.9242377572097171
```

Each model was trained and evaluated using a split of 80% training data and 20% testing data.

Visualization of Predictions

To compare model accuracy, I plotted actual Bitcoin prices vs. predicted prices for the test set. The Random Forest model provided the closest alignment.





Conclusions

- **No statistically significant relationship** was found between sunspot activity and cryptocurrency price returns.
- **Machine learning models** predicted Bitcoin prices reasonably well based on Ethereum prices, but **adding sunspot data did not improve performance.**
- This suggests that price behavior is more likely driven by market factors than environmental phenomena like solar activity.