**Advanced Data Analytics Final Project**

Chase Rawson

College of Professional Studies, University of New Hampshire

DAT 670: Advanced Data Analytics

Dr. Tony Sulpizio

December 11, 2023

In a digital era marked by the continuous evolution of financial markets, understanding the dynamics of cryptocurrencies is urgent. This document presents a comprehensive analysis of hourly price data spanning thirty major cryptocurrencies from January 2017 to October 2023. As the cryptocurrency landscape grows in complexity, this analysis seeks to unravel key insights, offering a strategic lens for market participants and stakeholders.

The Excel file titled *Chase Rawson Final Project Data* shows a detailed, comprehensive collection of hourly price data for thirty major cryptocurrencies, covering January 2017 to October 2023. The data includes the fields Open, High, Low, Close, Volume, and the number of trades for each cryptocurrency for each hour (as depicted on each row). (Fgjspaceman, 2023). There are 1,037,638 rows with data from cryptocurrencies Bitcoin, Ethereum, and Dogecoin, among many others. In the dataset, I was able to discover six key performance indicators (KPIs), which are Symbol by Sum of Volume, Symbol by Sum of Tradecount, Year and Quarter by Sum Tradecount, Year and Quarter by Sum of Volume, Symbol by Max of High and Min of Low, and Symbol by Standard Deviation of Open. Since January 2017, 10,266,790,509 cryptocurrency trades have been performed, equaling a total volume of $8,505,823,467,448.77. This averages around 9856 trades performed every hour, equaling a total volume of $ 8,197,301.63, a volume of $831.72 per trade.

I also performed five different regression analyses using the closing price as my dependent variable and the volume, tradecount, open price, highest price, and lowest price as my independent variables. This led me to discover the following information:

* Volume Analysis: The analysis revealed that the 'Volume' variable showed a relatively weak relationship with the 'Close' prices, as evidenced by the low R-squared value of approximately 26.4%. The coefficient was 0.0000882195, suggesting that a unit increase in 'Volume' corresponded to a minimal increase in 'Close' prices.
* Trade Count Analysis: Similarly, the 'Trade count' variable demonstrated a limited influence on the 'Close' prices, as indicated by the R-squared value of about 16.7%. The coefficient was 0.0769224621, indicating that the impact of a unit increase in 'Trade count' on 'Close' prices was moderate.
* Open Price Analysis: The analysis involving the 'Open' prices exhibited a strong positive relationship with the 'Close' prices, as denoted by the high R-squared value of approximately 99.99%. The coefficient of 0.9999786030 implied that a unit increase in 'Open' prices corresponded to nearly a one-to-one increase in 'Close' prices.
* High Price Analysis: The analysis involving the 'High' prices demonstrated a similarly robust relationship with the 'Close' prices, as indicated by the high R-squared value of around 99.99%. The coefficient of 0.994699062417 suggested that a unit increase in 'High' prices was associated with a substantial increase in 'Close' prices.
* Low Price Analysis: The analysis involving the 'Low' prices also indicated a strong positive relationship with the 'Close' prices, with an R-squared value of approximately 99.99%. The coefficient of 1.0056571544 implied that a unit increase in 'Low' prices was linked to a notable increase in 'Close' prices.

I discovered that in most cases, coefficients of the independent variables were associated with statistically significant p-values. This suggests that the relationships between the independent and dependent variables were not likely due to random chance. The R-squared values varied across the analyses, indicating the proportion of the variance in the 'Close' prices that the independent variables could explain. The R-squared values ranged from approximately 16.7% to nearly 100%, suggesting that the models could explain varying degrees of variability in the 'Close' prices. The intercepts in some cases were statistically significant, indicating that the dependent variable's value significantly differed from zero when the independent variable was zero. However, the interpretation of these intercepts might not be meaningful in the data context. These findings suggest that the 'Open,' 'High,' and 'Low' prices substantially impact the 'Close' prices, indicating a strong linear relationship. In contrast, the 'Volume' and 'Trade count' variables appeared to have relatively weaker associations with the 'Close' prices, suggesting that other factors may play a more significant role in determining cryptocurrency prices.

As we see on the Excel file, under the Dashboard tab are six visuals representing each of the six pivot tables I have created. Those visuals, in order, show Symbol by Sum of Tradecount, Symbol by Sum of Volume, Symbol by Standard Deviation of Open, Year and Quarter by Sum Tradecount, Year and Quarter by Sum of Volume, and Symbol by Max of High and Min of Low.

From the Dashboard, starting with the top left visual and moving to the right, we can determine that a great majority of Trades over the past six years came from Bitcoin, over 3.246 Billion. Ethereum is second, with 1.212 Billion trades performed. A great majority of the number of trades performed, and the amount of volume traded mostly came through these two cryptocurrencies. From Bitcoin, $3.677 Trillion and $1.583 Trillion of Ethereum were bought and sold over the past six years, as depicted in the second visual. The third visual shows the variation of the opening prices of each cryptocurrency over the past six years. I can determine that Bitcoin has the highest amount of variation, 15,829.56%, of opening prices while the cryptocurrency Shiba Inu has the lowest amount of variation, 0.0000011%.

From the bottom left visual and moving to the right, we can see the number of total cryptocurrency trades performed over each financial quarter over the past six years. The most trades performed were from 2021 Q2; over 1.5 trillion trades were performed. 2017 Q3 was when the least amount of trades were performed; just under 550,000 trades were performed. There was a slight trend of trades performed from 2017 Q3 until about the end of 2018 Q1. A massive trend in trading cryptocurrency started just after 2019 Q1, peaking at 2021 Q4 before slowing down at around 2023 Q2. I should most importantly note that for the first two visuals on Dashboard 2, there is some data from 2023 Q4 up until the middle of October 2023. However, using this data to see the view of the total amount of trades/volume over 2023 Q4 should not be trusted. The second visual is the volume traded over each financial quarter over the past six years.

Like the first visual, there is a slight trend of trades performed from 2017 Q3 until about the end of 2018 Q1. A massive trend in trading cryptocurrency started just after 2019 Q1, peaking at 2021 Q4 before slowing down at around 2023 Q2. The most volume traded was from 2021 Q2; over $1.5 Trillion was traded. 2017 Q3 was when the least amount of trades were performed. Just over $228 Million of volume was traded. Finally, visual three shows each cryptocurrency Symbol by Max of High and Min of Low. This visual can also be used to determine the amount of variation of cryptocurrency prices over the years by showing the percentage gain of the cryptocurrency, when the price was the highest and when it was the lowest, among many other statistics that can be gathered. One cryptocurrency that stood out was Chainlink; it is the maximum price was $53 while it is minimum price was $0.0001, which is a percentage gain (when dividing the low from the high) of 0.000189%, the lowest percentage gain of all of the cryptocurrencies. For reference, Bitcoin's percentage gain is 4.08%, and Ethereum's gain is 6.77%. To recap:

Key Performance Indicators (KPIs):

* Symbol by Sum of Volume and Sum of Tradecount highlight the dominance of Bitcoin and Ethereum in terms of trading activity.
* Year and Quarter-based KPIs reveal trends in trading over the past six years.

Regression Analyses:

* Volume Analysis:
  + Weak relationship with 'Close' prices (R-squared: 26.4%).
  + Coefficient: 0.0000882195, indicating minimal impact on 'Close' prices.
* Trade Count Analysis:
  + Limited influence on 'Close' prices (R-squared: 16.7%).
  + Coefficient: 0.0769224621, suggesting a moderate impact.
* Open Price Analysis:
  + Strong positive relationship with 'Close' prices (R-squared: 99.99%).
  + Coefficient: 0.9999786030, indicating a nearly one-to-one increase.
* High and Low Price Analysis:
  + Strong positive relationships with 'Close' prices (R-squared: 99.99%).
  + Coefficients: 0.994699062417 and 1.0056571544, respectively.

Located on the Power BI file titled *Chase Rawson Final Project Predictive Analysis* is six predictive analytical models using the top two cryptocurrencies by Volume, Bitcoin, and Ethereum. Six predictive models each show the Sum of Volume by Year, Quarter, Month, and Day, the Sum of Volume by Year, Quarter, Month, and Day, and the Standard Deviation of the Open Price by Year, Quarter, Month, and Day, for both Bitcoin and Ethereum. I should note that the models are based on the data from the dataset since January 1, 2020, since the monetary decisions for the general cryptocurrency and stock markets within the last three and a half years continue to factor in the financial decisions today. Also, the forecast length for each of the six models is 365 days since the last data was added to the file; the predictions last until October 2024.

Starting with the three Bitcoin models, in analyzing the predictive models for Bitcoin in Power BI, it is noticeable that the Sum of Tradecount and Sum of Volume forecasts exhibit a relatively stable trend over the next 365 days, maintaining an average level. However, a notable anomaly occurred on May 19, 2024, where both predictions spiked significantly, reaching four times the consistent average for Tradecount and displaying an even more substantial surge for Volume. This outlier on May 19 suggests a potential event or market condition significantly impacting trading activity and Volume on that specific day. Interestingly, the Standard Deviation of the Open Price forecast shows a distinct pattern. It gradually increases leading up to May 19, hinting at heightened market volatility, but then sharply declines and stabilizes near zero afterward. This unique behavior implies a sudden decrease in price variability, possibly indicating a period of market stabilization or a specific event that mitigates price fluctuations post-May 19. This event on this date could indicate a significant social or economic global event news related to Bitcoin or even regulatory changes that lead to increased volatility.

For the three Ethereum models, it is noticeable that the Sum of Tradecount forecast displays a relatively stable trend over the next 365 days, maintaining an average level with no significant changes. However, a notable exception occurred on May 19, 2024, when the forecast experienced a substantial spike, reaching five times the consistent average, suggesting a significant uptick in trading activity on that specific day. Similarly, the Sum of Volume forecast follows a consistent pattern, except for the anomaly on May 19, where it reaches 21 times the average, indicating an exceptional surge in trading volume. In comparison, the Standard Deviation of the Open Price forecast for Ethereum shows no significant increases, even on May 19; instead, it either remains constant or experiences slight decreases over the next 365 days. This consistent standard deviation implies a sustained level of market stability or a lack of significant fluctuations in Ethereum’s open price volatility during the forecast period, emphasizing a relatively calm market environment despite the isolated anomalies in trading activity and Volume on May 19.

Comparing these two cryptocurrencies, I can safely say that on or around May 19, 2024, a market-wide event or external factor will affect the entire cryptocurrency space. This could include regulatory announcements, macroeconomic developments, or global events that simultaneously impacted Bitcoin, Ethereum, and the other twenty-eight cryptocurrencies. Assuming that today's global market decisions relate to events after January 1, 2020, these cryptocurrency predictions are more likely to relate to future events shaped within the last 3.5 years, especially since January 2021. Who knows what will happen to cryptocurrency next year, but it can be safe to say that, when comparing the volatility of these cryptocurrencies from the past three years, I am 99% confident that there are not likely any significant changes to macroeconomic policy, nor will there be any significant technological innovation that will lead to levels of volatility experienced within the last three years, for the next year. Overall based on all of the analysis that has been conducted, I recommend accomplishing the following:

1. Monitor 'Open,' 'High,' and 'Low' Prices. Given their strong positive relationships with 'Close' prices, close monitoring of 'Open,' 'High,' and 'Low' prices is recommended. Any fluctuations in these variables may have a significant impact on cryptocurrency prices.
2. Consider External Factors on May 19, 2024. Anomalies in trading activity and volume for both Bitcoin and Ethereum on May 19, 2024, suggest a potential market-wide event. Stakeholders are advised to stay vigilant during this period and consider external factors that might influence the cryptocurrency market.
3. Evaluate Trading Strategies for Bitcoin and Ethereum. Given the stable trends observed in the Sum of Tradecount and Sum of Volume forecasts for Bitcoin and Ethereum, traders and investors may benefit from developing or adjusting their strategies to accommodate these consistent patterns.
4. Assess Regulatory Environment. Considering the potential impact of external events on May 19, 2024, it is crucial for stakeholders to assess the regulatory environment and stay informed about any impending changes that might affect the cryptocurrency market.
5. Explore Market Stabilization Strategies. The distinct pattern in the Standard Deviation of the Open Price forecast for Bitcoin, indicating heightened market volatility leading up to May 19, suggests the potential need for strategies to manage and stabilize portfolios during periods of increased volatility.

Lastly, analysis of cryptocurrency data spanning from January 2017 to October 2023 unravels pivotal insights into market dynamics. From the identification of key performance indicators (KPIs) to the examination of regression analyses and the unveiling of predictive models, this report serves as a comprehensive guide for understanding the intricate world of cryptocurrencies. The key findings underscore the dominance of major cryptocurrencies like Bitcoin and Ethereum, the impact of 'Open,' 'High,' and 'Low' prices on 'Close' prices, and the potential influence of external factors on specific dates, notably May 19, 2024. These insights not only provide a snapshot of the past but also pave the way for informed decision-making in the future.

As stakeholders navigate the complex cryptocurrency landscape, the report's recommendations offer strategic pathways for mitigating risks and optimizing opportunities. Whether it's monitoring price variations, evaluating trading strategies, or staying vigilant during crucial periods, these recommendations are crafted to guide stakeholders toward prudent decision-making. This report serves as a path to understand and navigate the ever-evolving cryptocurrency market. The findings presented here lay the foundation for strategic actions, and the recommendations provide practical insights for stakeholders to navigate the dynamic landscape effectively. As the cryptocurrency space continues to evolve, this report stands as a valuable resource for informed decision-making and proactive market participation.

**References**

Fgjspaceman. (2023, October 21). *Crypto Data Hourly Price since 2017 to 2023-10*. Kaggle. <https://www.kaggle.com/datasets/franoisgeorgesjulien/crypto?select=Binance_BTCUSDT_1h%2B%281%29.csv>