Decoding Cryptocurrency Behavior with R: Statistical Techniques for Comprehensive Analysis

Shravan Gopalakrishnan, Vinaya Ramamorthy Venkatasubramanian, Krishnan Narayanan

SUMMARY

Cryptocurrencies are digital or virtual currencies that use cryptography for secure and decentralized transactions. They are typically based on blockchain technology, which is a distributed ledger that records all transactions across a network of computers. As with any investment, investing in cryptocurrencies comes with risks, including market volatility, regulatory uncertainty, security breaches, and potential loss of funds. It is important for investors to thoroughly research and understand the risks associated with cryptocurrencies before making investment decisions. This project proposes a statistical approach for analyzing and understanding cryptocurrency data, and to determine a reliable way to find the correlation between them. We have analyzed the market dynamics, identified the most valuable cryptocurrencies, and analyzed the distribution of each cryptocurrency using kernel density to gain insights into the spread of their values and identify which cryptocurrencies had a more stable or volatile distribution. Through methods such as correlation analysis, visualized correlations using correlation matrices, scatter plots, and heat maps, we analyzed the inter-relationships and how closely different cryptocurrencies were linked to each other. This analysis highlights the risks for investors, with matrices showing positive and negative correlation, average market capital, kernel density as well as relationships between cryptocurrencies.

METHODS

Methods section is available separately which includes Fig.1.

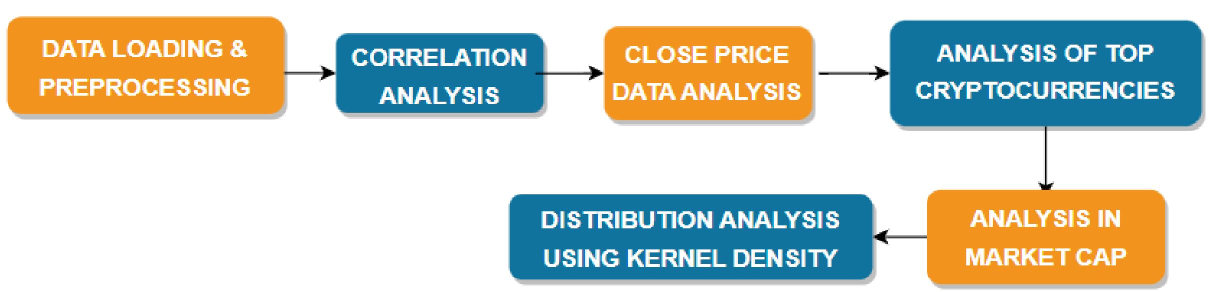

Fig.1 Proposed Research Architecture

RESULT

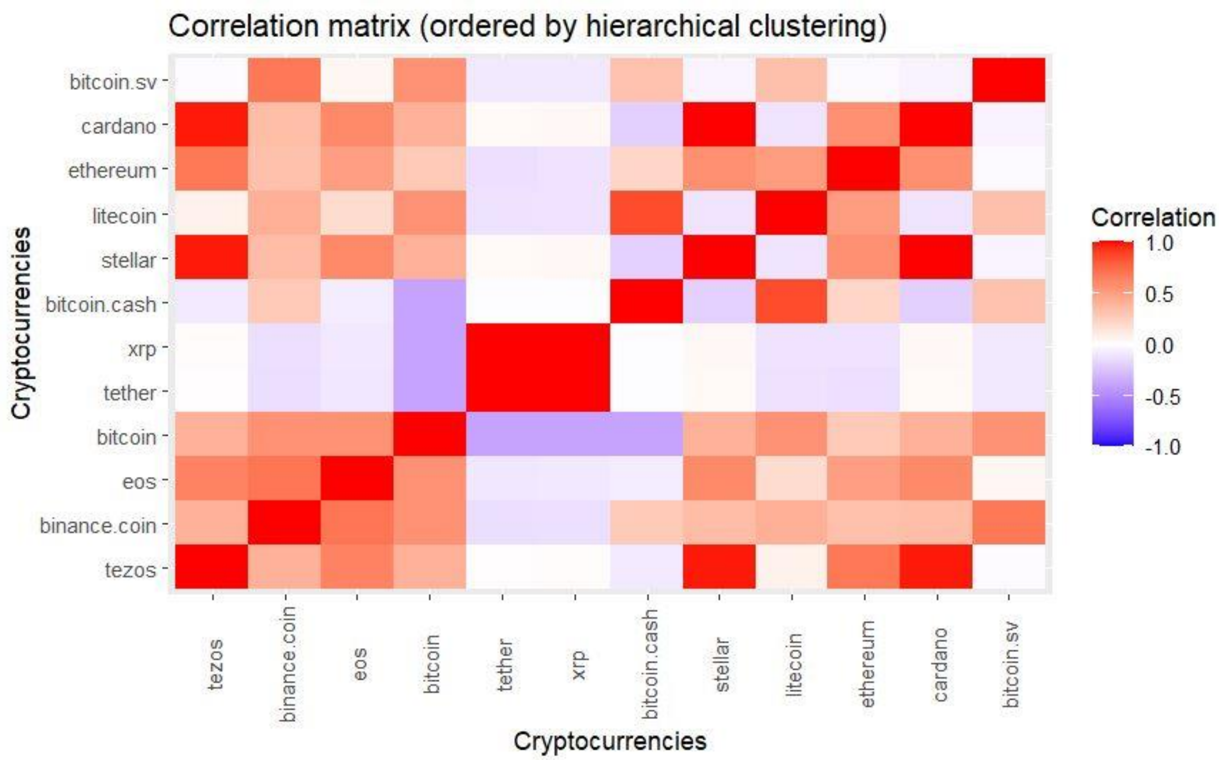

Fig.2 Correlation between Crypto Currencies

Fig.2 represents the Correlation between Crypto Currencies matrix which is ordered by hierarchical clustering generates correlation output with a range of -1 to 1, where -1 and -0.5 pertain to negative correlation, 0 being neutral or almost no correlation, and 0.5 to 1 denoting positive correlation. For instance, the heatmap data infers that the trends of Stellar and Ethereum move in the same direction together, thus have a positive correlation with each other. On the other hand, the trends of Tether and Bitcoin appear to be moving in opposite directions, confirming a negative correlation.

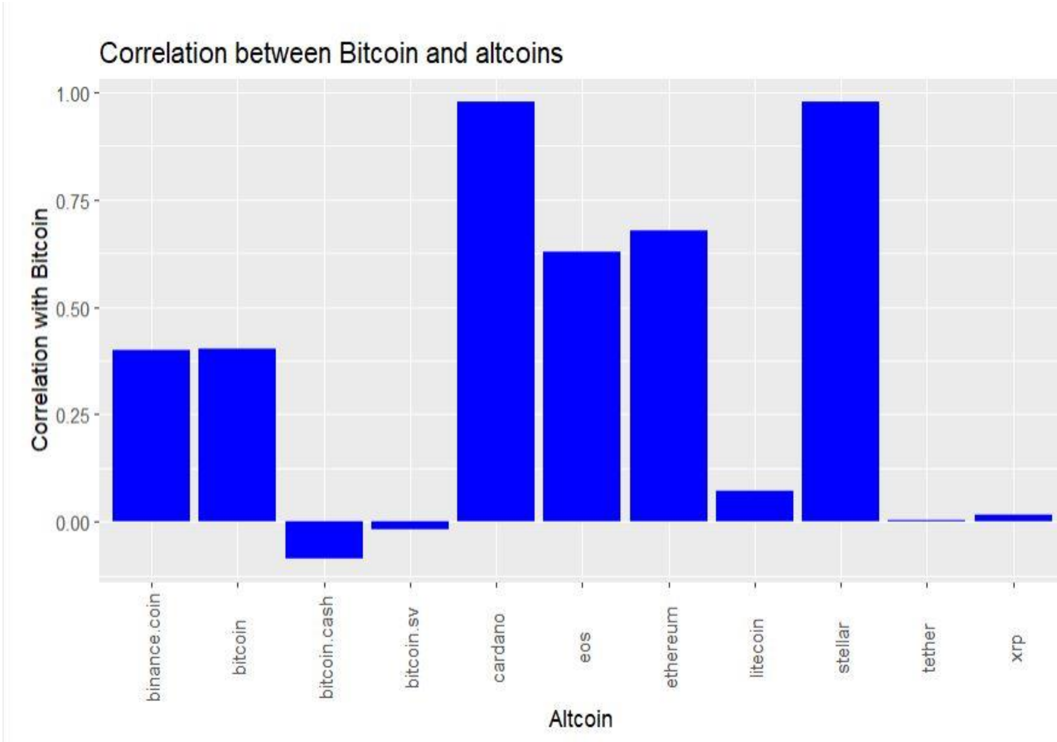

Fig.3 Correlation between Bitcoin and other coins

Fig.3 represents the Correlation between Bitcoin and altcoins. Following the same key for the bar plot that shows the correlation between Bitcoin and altcoins (alternate coins), we can clearly notice one of the prominent correlational behavior of trends being the likes of Bitcoin and Bitcoin.Cash traveling in opposite directions, thus exhibiting negative correlation. The top four positively correlated altcoins with Bitcoin are Cardano, EOS, Ethereum, and Stellar, while the top five negatively correlated altcoins are Bitcoin.Cash, Bitcoin.SV, Litecoin, Tether, and XRP.

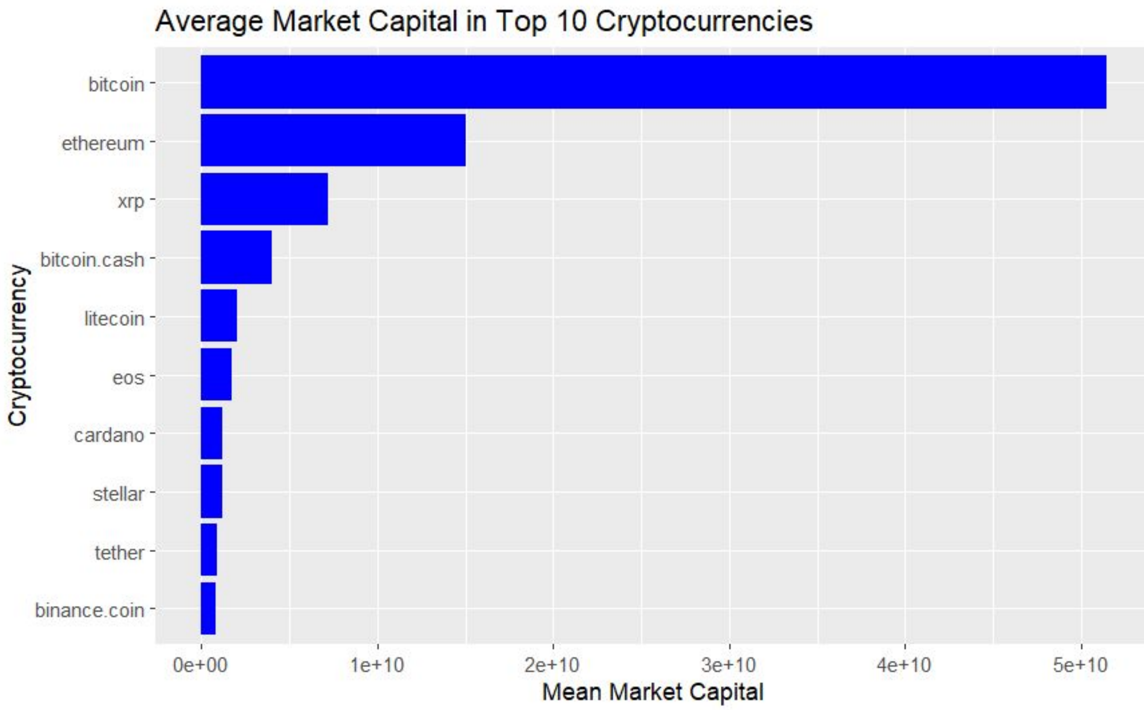

Fig.4 Cryptocurrency and their Mean Market Capital

Fig.4 visualizes bar plot that demonstrates the graph of cryptocurrencies against average / mean market capital, that orders top 10 cryptocurrencies by ascending order as Binance.coin, Tether, Stellar, Cardano, EOS, Litecoin, Bitcoin.cash, XRP, Ethereum, and Bitcoin. Bitcoin, being the most popular cryptocurrency, emerges to have the highest average of market capital.

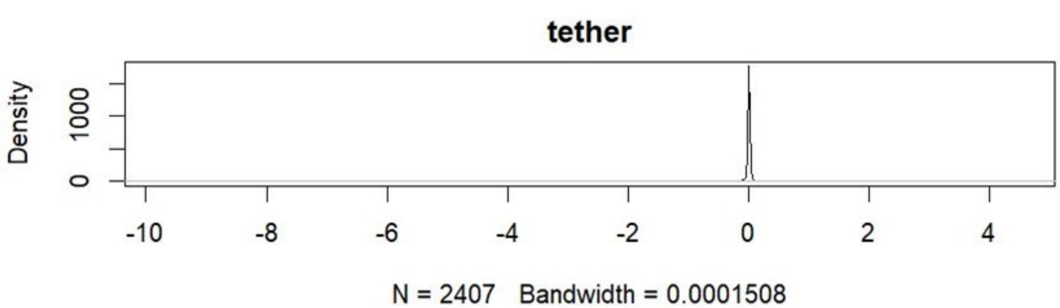

Fig.5 Kernel Density for Tether

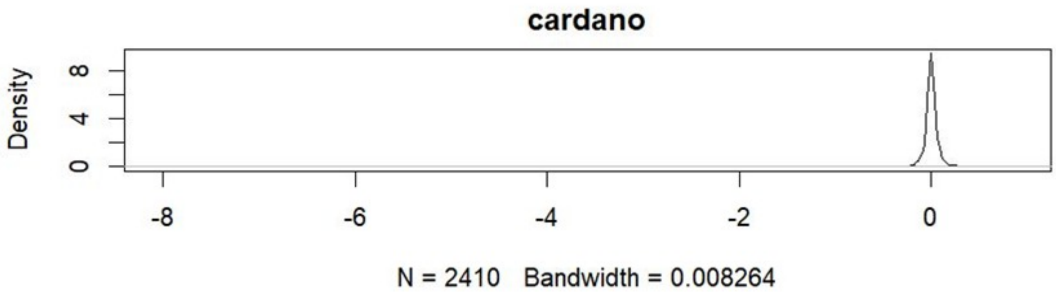

Fig.6 Kernel Density for Cardano

Fig.5 and Fig.6 represent the kernel Density of Tether and Cardano respectively. A higher kernel density for a cryptocurrency may indicate a tighter distribution of returns, which may suggest that the cryptocurrency is less volatile and less risky. On the other hand, a lower kernel density may suggest a wider distribution of returns and greater volatility, which may make the cryptocurrency riskier. The highest kernel density (top 10 points) of [0, 1.04] range was found for Tether, while the lowest was found for Cardano with a density of point range [0, 0.014]. From this, we can infer that investing in Tether is advisable, whereas investing in Cardano can be a bit risky.

Fig.7 Prices of crypto currencies Time Series Analysis

Fig.7 depicts the time series analysis of the prices of crypto currencies Bitcoin, Ethereum, XRP, Bitcoin.cash, Litecoin, EOS, Cardano, Stellar, Tether, and Binance.coin. At the beginning of the period, all cryptocurrencies had relatively low prices, with Bitcoin, Ethereum, and XRP being around $140, and other cryptocurrencies being under $5. Bitcoin and other cryptocurrencies experienced a sharp increase in price towards the end of 2017, with Bitcoin reaching a peak price of around $19,000 in mid-December 2017. However, since then, the prices of most cryptocurrencies have decreased significantly, and many of them are currently trading at prices lower than they were in early 2018.

DISCUSSION

Meaning and Impact: The results of this analysis provide valuable insights into the correlations and relationships between cryptocurrencies, which can be useful for investors looking to diversify their portfolios. Understanding the interdependencies between different cryptocurrencies can help investors make informed decisions about which ones to invest in and how to balance their investments. For example, the positive correlation between Stellar and Ethereum suggests that investing in both cryptocurrencies may not provide as much diversification benefits as investing in cryptocurrencies with more negative correlation. Similarly, the negative correlation between Tether and Bitcoin suggests that these two cryptocurrencies may be useful in balancing each other out in a portfolio. Overall, these results highlight the importance of monitoring the cryptocurrency market and understanding the potential impact it can have on investment strategies and portfolio diversification. As the cryptocurrency market continues to evolve and mature, it will be important for investors to stay up to date with these trends and adjust their strategies accordingly.

Beneficiaries: This project would benefit investors who are interested in investing in cryptocurrencies. The analysis provides valuable insights into the correlations between different cryptocurrencies, which can help investors make more informed investment decisions. The project also highlights the inherent risks of investing in cryptocurrencies and the need for caution. Additionally, the project may benefit financial analysts, researchers, and policymakers who are interested in understanding the impact of cryptocurrencies on the broader financial landscape. The analysis of the potential impact of cryptocurrencies on the stock market and the need to monitor this impact can inform investment strategies, portfolio diversification, and overall market trends.

Informed Decisions: With the help of cryptocurrency analysis, investors can conduct various statistical and data-driven techniques to investigate the relationships and interactions between different cryptocurrencies and determine the degree of association or dependency between the prices of different cryptocurrencies. Investors can also use time series analysis techniques, to study the historical price patterns and forecast future prices of cryptocurrencies. By analyzing the historical performance of multiple cryptocurrencies using time series techniques, investors can identify potential patterns, trends, or dependencies that may exist among different cryptocurrencies. This will in turn prove to be instrumental in making investment decisions that create a significant impact on corporate and the world stock market.

Future Work: Going forward, the cryptocurrency market may have a significant impact on the stock market that could lead to changes in investment strategies, portfolio diversification, and overall market trends. As such, it is important for investors to continue monitoring the cryptocurrency market and its potential impact on the broader financial landscape. Additionally, the integration of blockchain technology into this analysis could further increase the significance of cryptocurrencies and their impact on the economy and could be a potential area for future research.

STATEMENT OF CONTRIBUTIONS

Krishnan Narayanan played a key role in the project by conducting the correlation analysis using various statistical techniques such as correlation matrices, scatter plots, and heat maps. By visualizing the correlations between different cryptocurrencies and Bitcoin, he identified which cryptocurrencies were most closely linked to Bitcoin. In addition, Krishnan analyzed the market dynamics and identified the most valuable cryptocurrencies. He also used kernel density to analyze the distribution of each cryptocurrency and provided valuable insights into the behavior of these assets.

Shravan Gopalakrishnan’s contribution to the project was crucial in converting the close price data of the top cryptocurrencies into time series and analyzing their volatility. He calculated important statistical measures such as the mean, variance, standard deviation, and coefficient of variation (CVar) to create log-return charts. This analysis was critical in identifying the potential risks and returns associated with each asset. Shravan's efforts helped to provide a comprehensive analysis of the volatility of different cryptocurrencies.

Vinaya Ramamorthy Venkatasubramanian was pivotal for extracting and averaging the market cap of each cryptocurrency to understand their market share. Her contribution to the project provided valuable insights into the market dynamics of cryptocurrencies. Vinaya also used kernel density to analyze the distribution of each cryptocurrency, which helped in identifying which assets had a more stable or volatile distribution. Her analysis was important in providing a comprehensive understanding of the spread of each asset's value.

REFERENCES

1. DeVries, Peter D. "An analysis of cryptocurrency, Bitcoin, and the future." International Journal of Business Management and Commerce 1.2: 1-9.
2. Alexander, Carol, and Michael Dakos. "A critical investigation of cryptocurrency data and analysis." Quantitative Finance 20.2 (2020): 173-188.
3. CoinMarketCap. (2023). Cryptocurrency Market Capitalizations. Retrieved March 16, 2023, from https://coinmarketcap.com
4. Moss, Jonas, and Martin Tveten. "kdensity: An R package for kernel density estimation with parametric starts and asymmetric kernels." Journal of Open Source Software 4.42 (2019): 1566.
5. Danylchuk, Hanna, et al. "Monitoring and modelling of cryptocurrency trend resistance by recurrent and R/S-analysis." S-Analysis (April 22, 2020) (2020).

APPENDIX

Relevant Code

```
## close price data

```{r message=FALSE, warning=FALSE}
library(reshape2)
prices_raw <- reshape(data[c(1, 2, 6)], timevar = "Currency", idvar = "Date", direction = "wide")
prices_raw[,"Close.Currency"] <- NULL

prices <- data.frame(sapply(prices_raw, function(z) {as.numeric(as.character(z))}))

colnames(prices) <- sub("Close.", "", colnames(prices))

library(lubridate)
dates <- parse_date_time(x = as.character(prices_raw$Date), orders = "b/d/Y", locale = "eng")
prices$Date <- dates

prices <- prices[-nrow(prices),]

```
```

```
## Correlation matrix
```{r message=F}
prices_min_180 <- prices[, colSums(!is.na(prices)) >= 180]

corr_mat <- cor(prices_min_180[,-1], use = "pairwise.complete")

library(ggplot2)
library(reshape2)

corr_long <- melt(corr_mat)

corr_long

ggplot(data = corr_long, aes(Var1, Var2, fill = value)) +
  geom_tile() +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red",
                        midpoint = 0, limit = c(-1,1), name = "Correlation") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5)) +
  labs(x = "Cryptocurrencies", y = "Cryptocurrencies",
       title = "Correlation matrix (ordered by hierarchical clustering)")
```
```

```
##Average market cap
```{r mean_markcap}
mean.cap <- data.frame(mean.cap=colMeans(markcap, na.rm = T))
mean.cap.10.name <- rownames(mean.cap[order(mean.cap$mean.cap, decreasing = T),,drop=F])[1:10]
mean.cap.10.value <- mean.cap[order(mean.cap$mean.cap, decreasing = T),,drop=F][1:10,]
mean.cap.10 <- data.frame(name=mean.cap.10.name, mean.market.cap=mean.cap.10.value)
mean.cap.10

library(ggplot2)

ggplot(mean.cap.10, aes(x = reorder(name, mean.market.cap), y = mean.market.cap)) +
  geom_col(fill = "blue") +
  coord_flip() +
  labs(title = "Average Market Capital in Top 10 Cryptocurrencies", x = "Cryptocurrency", y = "Mean Market Capital")

```
```

```
##Mean
```{r mean, collapse=TRUE}
options(digits = 3)
data.frame(mean.percent = sort(apply(ret10_clean[,1:ncol(ret10_clean)], 2,
                           function(x) mean(x, na.rm=TRUE)), decreasing = T))*100
```
```

```
##Variance
```{r variance,  collapse=TRUE}
options(digits = 3)
data.frame(variance.percent = sort(apply(ret10_clean[,1:ncol(ret10_clean)],
                                   2, function(x) sd(x, na.rm=TRUE)), decreasing = T))*100
```
```

```
##Standard deviation
```{r}
options(digits = 3)
data.frame(variance.percent = sqrt(sort(apply(ret10_clean[,1:ncol(ret10_clean)],
                                   2, function(x) sd(x, na.rm=TRUE)), decreasing = T)))*100
```
```

```
##CVar
```{r message=FALSE, warning=FALSE}
library(PerformanceAnalytics)
CVaR(ret10_clean)
```
```

```
##Log-return chart
```{r echo=FALSE}
plot.xts(ret10.xts, main="log-return", ylim = c(-3,7))


par(mfrow=c(2,1));
for(i in 1:ncol(ret10)){
  print(plot(ret10.xts[,i], main=colnames(ret10.xts)[i]))
}
```
```

```
#Distribution analysis
##Kernel Density
```{r}
par(mfrow=c(2,1));
for(i in 1:ncol(ret10_clean)){
  plot(density(ret10_clean[,i], na.rm = T), main=colnames(ret10_clean)[i])
}
```
```