Temporally Dynamic Visualizations of the Cryptocurrency Market

Shawn Anderson, Ramanpreet Singh, Rishabh Singh

Abstract—In this work the authors explore dynamic visualizations of windowed periodic returns of cryptocurrencies. The first step is to collect the global price matrix (GPM) for some selection of coins, as well as a selection of a cash. The cash is what the value of all other coins is measured in. The authors use both Bitcoin and USD as the cash, and use popular cryptocurrencies such as Bitcoin Cash, Ripple, Ethereum, and Tether as the selected coins. All visualizations are done on a 30 perdiod sliding window over the GPM. Temporally interactive visualizations are implemented using the python library holoviews. Visualizations include line charts over prices and periodic returns, a heatmap of coin correlations, a correlation network, a scatterplot of summary statistics, and a kernal density plot of returns. Visualizations reveal structure in the correlation properties of cryptocurrencies.

Index Terms—Data science, cryptocurrency, bitcoin, visualization, dynamic networks

1 Introduction

1.1 Domain

The domain for this project is cryptocurrency markets. A cryptocurrency is a digital or virtual currency deisgned to work as a medium of exchange. It uses strong cryptography to secure financial transactions, control the creation of additional units, and verify the transfer of assets. The total market cap of the cryptocurrency market is over 125 Billion USD [1]. It is a relatively new market as compared to the traditional markets such as equities and commodity markets, and although it is built on the revolutionary premise of blockchain technology, but some inspiration to exploit it can be learned from the more traditional markets.

1.2 Users

The visualizations are primarily aimed towards two types of users in the financial world - traders and investors. Investors gradually build wealth over an extended period of time through the buying and holding a portfolio of financial assets, whereas traders on the other hand try to outperform the market by frequent buying and selling of assets, most often paired with high volumes. The static visualizations are geared towards the investors, whereas the streaming ones are for the traders, although either may prove to be useful for the other based on the purpose of their usage.

1.3 Data

The data in our experiments is from the Cryptocompare website [2]. We get daily open high low close (candlestick) data for our selected coins from January 01 2018 until the present day. From here, for each coin, and for each day, we extract the closing price for each coins, resulting in the Global Price Matrix (GPM). The GPM is an MxN matrix where M is the number of selected coins, and N is the number of time periods. GPM[m,n] is the price of coin m on period n. From the GPM, a returns matrix R is computed where R[:,n] = GPM[:,n] / GPM[:,n-1] for $n \in (1,N)$. From R, we select a window of 30 periods. All visualizations are performed on a 30 day window from R. Temporally interactive visualizations are achieved by applying a static visualization to each window over R, and then providing a slider for which users can slide over the final day of the window.

1.4 Task

The visualizations created in this project can be leveraged by users to answer important questions in the domain of cryptocurrencies. Since all the visualizations are interactive, this assists the users to visualize the same data at different time windows. Some of the key questions that can be answered through these visualizations are:

- How to utilize visual insights to improve/maximize portfolio returns?
- Are there any trends/patterns in coins price chart?
- Are there any outliers/market correction in coins value?
- How to diversify the portfolio and spread risk?
- Are there any correlation clusters?
- How to summarize data over a time window?

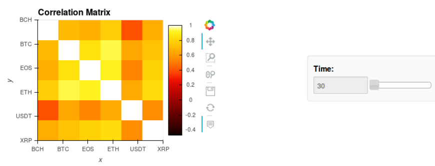


Fig. 1: Correlation Matrix

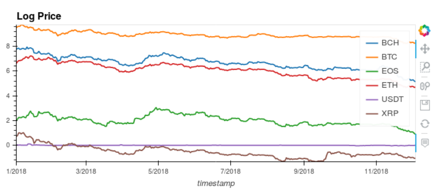
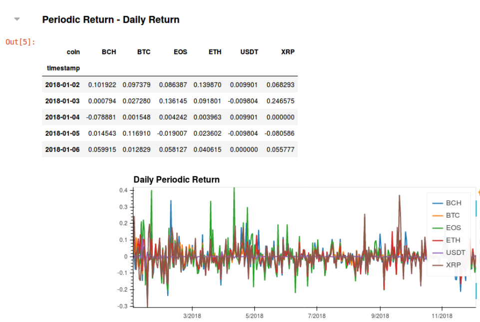


Fig. 2: Log Price Chart



Periodic Return - Daily Return

Out[5]:

coin timestamp	BCH	BTC	EOS	ETH	USDT	XRP
2018-01-02	0.101922	0.097379	0.086387	0.139870	0.009901	0.068293
2018-01-03	0.000794	0.027280	0.136145	0.091801	-0.009804	0.246575
2018-01-04	-0.078881	0.001548	0.004242	0.003963	0.009901	0.000000
2018-01-05	0.014543	0.116910	-0.019007	0.023602	-0.009804	-0.080586
2018-01-06	0.059915	0.012829	0.058127	0.040615	0.000000	0.055777

Fig. 3: Daily Periodic Return

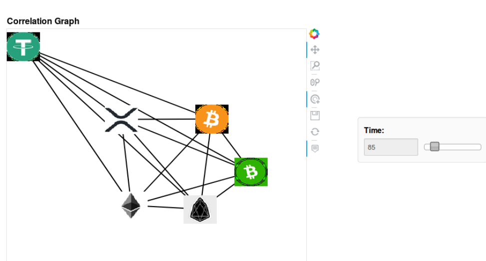


Fig. 4: Correlation Network

2 Design of Visual Idioms and Interactions

A visualization idiom is a distinct approach to create and manipulate visual representations. Visualization is incomplete and puzzling without proper choice of visual encoding and representation. Visual designs must avoid spurious visual associations so that users can visualize features of both association and distinction in the designs. It is therefore critical that we recognize tactics that are useful to ensure distinction between elements in visual representation that can also be useful to emphasize the connectedness within subgroups of those representations.

Interactive data visualization allows users the freedom to fully explore analyzed data. It enables direct actions on a plot to change elements and link between multiple plots. A major advantage of interactive visualizations is that the content can be changed by the user. The main part of such a configurable visualization becomes the template through which different structurally similar data sets are displayed, and additional controls allow the user to change what data gets displayed. When used in such a manner, an interactive visualization can make a much larger data set accessible than a comparable static graphic. Those commonly used representations also often have well-known interactions such as pan and zoom for maps. By assembling multiple standard parts and coordinating them, you can show different aspects of the data set at the same time.

All the visualizations in this project have been made by keeping in mind the significance of these two properties.For a cryptocurrency investor understanding the relationship between various coins is of utmost importance as they are always interested in diversifying risks by selecting different types of them. Therefore, to visualize which stocks behave similarly (positive correlations) or very differently (negative correlations) can reap potential economic benefits. For presenting correlation between coins, we chose two visualizations viz. correlation matrix and correlation network graph. These correlations were computed on the entire dataset for every 30-day time window using Pearson correlation coefficient which measures the linear dependence between two variables. The motive behind choosing correlation matrix is that it can be reordered/sorted according to the correlation coefficient, which is important to identify the hidden structure and pattern in the matrix. Moreover, the colormap of the correlation matrix allows the user to easily interpret and understand the hidden patterns and translate numerical data values into visible colors in a plot. A well-chosen colormap can help guide the user to notice the features of the data we want to highlight, while a poorly chosen colormap can completely obscure the data and lead to erroneous conclusions. To faithfully and intuitively represent monotonically increasing values, we should choose a colormap where the lowest values are similar in tone to the page background, and higher values become more perceptually salient compared to the page background. For our visualization, we chose fire colorbar as it effortlessly permits the user to read two opposite ends of the spectrum, where positive correlations are displayed in yellow and negative correlations in red color and color intensity is proportional to the correlation coefficients.

Correlation network graphs not only allows us to visualize correlations between coins but also helps us to identify clusters/groups, that is, finding sets of related vertices in graphs. For our use case, nodes represent coins and edge weights (length between two nodes) represents correlation between coins. In correlation network graphs the edge weights are inversely proportional to correlation, so higher the correlation shorter the edge and smaller the correlation longer the edge. Therefore, both the x-axis and y-axis of this graph were disabled to avoid clutter. These graph clusters provide macro-level view of the dataset. To make the network graph even more intuitive, the nodes (representing coins) were replaced by respective coin symbols using Holoviews Overlay (*) operator, which can aid even the non-technical user to interpret and read the graph effectively. The interactive slider was also added to the network graph to visualize the correlation/clusters at different time windows for the entire period.

For displaying trends in coins prices and returns we have plotted a date-and-cost line chart that tracks a commodity's changing price over time. It is generally used by economists to display broad market trends and predict future prices. Businesses that restock continually use this sort of graph to monitor their own expenses. It assists in tracking the fluctuation of stock prices, gives an accurate and quickly understandable assessment of the trend, acceleration, deceleration, and volatility of the cryptocurrency market. Since cryptocurrency market is relatively new, it tends to be more volatile than conventional markets. The line chart can be leveraged to show volatility indicators that can help traders to understand when they can expect to see a trend strengthen or when a new trend will be established. Furthermore, they are excellent for tracking and comparing multiple coins simultaneously over an entire time frame. As with all charts that record periodic changes, time, the independent variable is plotted on the chart's x-axis. Price/Returns, the dependent variable, goes on the chart's y-axis. All the absolute cryptocurrency prices/returns have been converted to log prices/returns because absolute price of every coin varies from thousands of dollars (Bitcoin Cash, BCH) to few cents (Ripple, XRP). Log returns are also widely preferred over raw prices for reasons such as normalization, time additivity etc. Apart from historic line chart, we also built interactive line charts showing cumulative and aggregated returns for each 30-day time window. Each time frame can be viewed by using an interactive slider that renders the line graph every time the slider value changes. These aggregated and cumulative returns can be used to decipher the coins performance over a certain time period.

Finally all these charts were composed together to form a dashboard using Holoviews '+' operator, which allows to visualize and interact multiple graphs simultaneously. The dashboard was equipped with interactive time slider, to render all the graphs together for individual time frames.

3 Implementation

The data for this project has been fetched from cryptocompare API, which is the cryptocurrency data provider, giving access to real-time, high-quality, reliable market and pricing data. The number of coins for this project has been limited to 5 viz. BCH (Bitcoin Cash), USDT (Tether), XRP (Ripple), ETH (Ethereum), and EOS, but can be scaled up to 2000 coins (available on cryptocompare). The data has been collected from May 1st, 2018 to Nov 1st, 2018. After fetching the data from API, it was converted to 3d-array using x-array with dimensions as timestamp, coins and coin features.

There were two sets of visualizations created catering to two sets of users visualizations on historic data for investors and visualizations on live streaming data for intra day traders. Mostly all the visualizations were created using Holoviews. It is a python package, when combined with the Jupyter Notebook and a plotting library, provides a rich, interactive interface for flexible and nearly code-free visualization of your results while storing a full record of the process for later reproduction. HoloViews provides a set of general-purpose data structures that allows to pair the data with a small amount of metadata. These data structures are then used by a separate plotting system to render the data interactively. HoloViews also provides powerful containers that allows to organize the data for analysis, embedding it whatever multidimensional continuous or discrete space best characterizes it. The most powerful component of Holoviews is that it makes it trivial to compose elements in the two most common ways: concatenating representations into a single figure or overlaying visual elements within the same set of axes. These components have been widely used throughout the project. The + operator has been used to compose multiple HoloViews elements together to form a dashboard, and * operator was used for overlaying coin images on top of nodes in correlation network graph.

For visualizing live streaming data, many components were used and linked together to build live visualizations. Firstly, a celery worker was built in Django (a python framework), to fetch API calls from cryptocompare after every second. Celery worker creates an asynchronous task that are executed concurrently on a single or more worker servers using multiprocessing. It is mainly focussed on real time operation but supports scheduling as well. The data fetched by the celery workers was stored in Redis database. It is used to cache the data fetch from the API calls, therefore it is faster than traditional database that writes to disk. Due to these features, Redis is often used to store state temporarily. This is common with microservice architectures, session stores and data that doesn't need long term persistence. Only 30 rows (timestamps) are cached in Redis at a time, after which it deletes those rows and add new 30 data points. Once the data is collected in Redis, Holoviews streamz and Pipe object were created, which allows data to be pushed into a Holoviews DynamicMap callback to change a visualization. Then using a sliding window of 30, these 30 rows are concatenated to pandas dataframe, which is then taken by a DynamicMap element to display graphs. We can interact with these live visualizations in a same way as we did for historic plots.

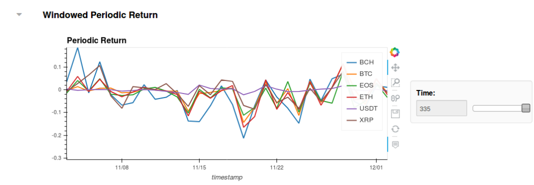


Fig. 5: Windowed Periodic Return

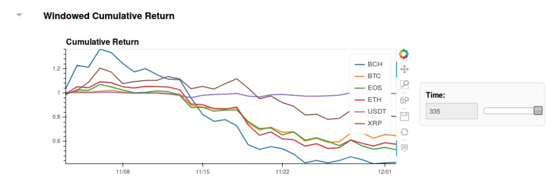


Fig. 6: Windowed Cumulative Return

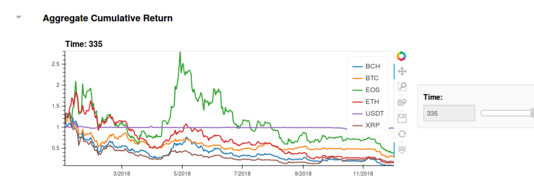


Fig. 7: Aggregate Cumulative Return

4 Discussion

From the visualizations that were created, the questions initially posed were answered. The following are some of the insights and trends that were observed through the visualizations:

- From the correlation matrix and the graph, we find that Tether(USDT) is negatively correlated to all the other coins we tested i.e. Ripple(XRP), Eos(EOS), Etherium(ETH) and Bitcoin Cash(BCH).
- Ripple(XRP), Eos(EOS), Etherium(ETH) and Bitcoin Cash(BCH) are all highly correlated to each other.
- Tether (USDT) is the least volatile currency.
- The cumulative return is highest for Tether.

  Thus, based solely on correlation, two clusters were most noticeable - one comprising of just USDT and the other comprising of XRP,BCH,EOS and ETH. One good diversification strategy can be to invest some portion of the capital in each of these two clusters, thereby minimizing the risk. In practice, the proportion of portfolio allocated to each cluster will also depend on other factors such as the performance of the individual cluster, the fundamentals of the coins involved, the risk tolerance etc, which must be taken into account when calculating the precise values.

5 Conclusion

In this work the authors have exhibited a novel toolkit for cryptocurrency market visualization. By applying common econometric analysis to a sliding window over the global price matrix, various interactive visualizations can be produced, with time being a common dimension of interactivity. Interactive visualizations include a heatmap correlation matrix which changes to colour over time, a correlation network that changes shape over time, and a scatter plot of return distribution parameters with moving data points over time. Temporally dynamic visualizations prove useful in identifying meaningful insights into the nature of cryptocurrency markets, such as contraction of correlations at certain periods, or clustering of currencies based on correlation clusters. Further work would include user testing of visualizations for usefulness to professional investors and traders in the cryptocurrency industry.

References

[1] https://coinmarketcap.com/
[2] https://min-api.cryptocompare.com/data/histoday
[3] https://www.investopedia.com/
[4] http://holoviews.org
[5] ttps://www.interaction-design.org
[6] ttps://zapier.com
[7] ttp://www.sfu.ca/siatclass/IAT814/