ISSS608 Visual Analytics and Applications

Cryptocurrency Price Analysis and Time Series Forecasting

Group 7

Chen Peng

Guo Yichao

26 April 2020

1. Overview

There are approximately 5,392 cryptocurrencies being traded with a total market capitalization of $201bn (as of April 22, 2020). Considerable growing rate is estimated in the following years. With the help of cryptography and a collective booking system called blockchain, cryptocurrencies build a distributed, safe and decentralized payment system, which does not need banks, intermediates, an organization or a central technical infrastructure to work.

However, mainly due to lack of regulatory support, Cryptocurrencies are highly volatile, which makes it a risky investment. Therefore, in this project, a Shiny web application is created with the cryptocurrency analysis in mind. It integrates two types of analysis: Time Series Analysis and Correlation Analysis within one application. This application is made to be informative and allows potential investors of cryptocurrency to interact with the data embedded within the application, view the visualization and generate results based on the selected filters before they make decision on cryptocurrency investment.

In this project, historical price data of top 10 cryptocurrencies from 2018 to 2019 are extracted from *https://coinmarketcap.com/* .

- **Bitcoin (BTC)** - is a consensus network that enables a new payment system and a completely digital currency, born on October 31st, 2008. Powered by its users, it is a peer to peer payment network that requires no central authority to operate.
- **Ethereum (ETH)** - is a smart contract platform that enables developers to build decentralized applications (dapps) conceptualized by Vitalik Buterin in 2013. ETH is the native currency for the Ethereum platform and also works as the transaction fees to miners on the Ethereum network.
- **XRP** - is an independent digital asset that is native to the XRP Ledger. With governance and fast transaction confirmations, XRP is said to be the most efficient settlement option for financial institutions and liquidity providers seeking global reach, accessibility, and fast settlement finality for interbank flows.
- **Tether (USDT)** - is a cryptocurrency with a value meant to mirror the value of the U.S. dollar. The idea was to create a stable cryptocurrency that can be used like digital dollars. Coins that serve this purpose of being a stable dollar substitute are called “stable coins.”
- **Litecoin** - is a peer-to-peer cryptocurrency created by Charlie Lee. It was created based on the Bitcoin protocol but differs in terms of the hashing algorithm used.
- **Binance Coin (BNB)** - is the cryptocurrency of the Binance platform. As of 2019, many businesses accept BNB as a form of payment.
- **EOS** - EOS.IO is a blockchain protocol powered by the native cryptocurrency EOS. The smart contract platform claims to eliminate transaction fees and also conduct millions of transactions per second.
- **Tezos (XTZ)** - Developed by ex-Morgan Stanley analyst Arthur Breitman. Tezos (XTZ) is a multi-purpose platform that supports dApps and smart contracts, and is powered by XTZ.
- **Stellar (XLM)** – is the intermediary currency on the Stellar network.

- **Cardano (ADA)** - is a decentralized platform that will allow complex programmable transfers of value in a secure and scalable fashion.

There are also two major cryptocurrencies: Bitcoin Cash and Bitcoin SV, with market capitalization of $4,459,706,172 and $3,633,013,064 (as of April 26, 2020). This research decides to use the historical price data from 2018 to 2019 for time series analysis and forecasting. However, a hard fork chain split of Bitcoin Cash occurred between two rival factions called Bitcoin Cash and Bitcoin SV on 15 November 2018. Hence, Bitcoin Cash and Bitcoin SV are not under consideration in this project scope in order to keep the consistency of the length of the data among multiple cryptocurrency.

2. Motivation and Objectives

Despite some of the limitations that cryptocurrencies presently face – such as computer crash or hacker interfering. What will be harder to overcome are governance and regulation. The more popular the cryptocurrencies become, the more regulation and government monitor they might attract. What's more, their relative complexity compared to conventional currencies will likely deter most people. However, there is no doubt that cryptocurrencies have shown their great potential on both investment and the possibility of becoming part of the mainstream financial system. On the other hand, cryptocurrency also influences the balance between regulation and free market, which require us to find out the possibility of making use of cryptocurrency and innovating the economic system. Therefore, Analysis on cryptocurrency is beneficial not only on economic but also academic.

We aim to achieve the following objectives in our project:

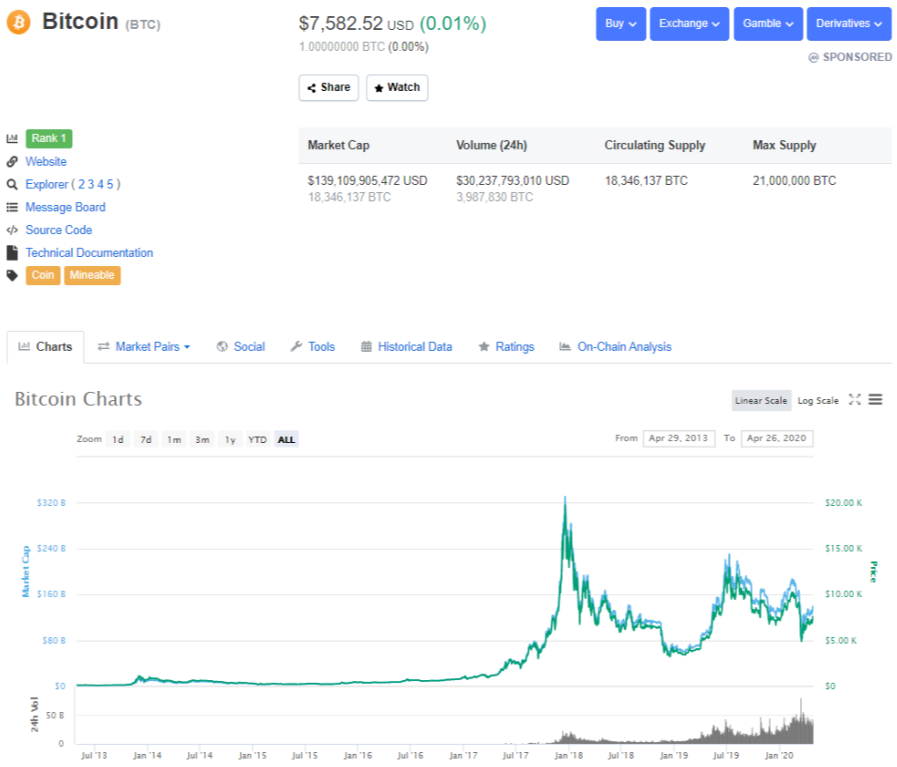
1. An overview of different cryptocurrency in time series, including open/close price, high/low price shown in a candlestick diagram and distribution of each cryptocurrency capital.

2. Correlation analyze among all cryptocurrency. Correlation analysis between Bitcoin and others cryptocurrency, which will help us reveal the similar price pattern of cryptocurrency.

3. Time Series analyze to forecast the future close price. In this project, we will use this data as a starting point for a detailed analysis into what is driving price action, and what can be done to predict future movement.

3. Review and critic on past works

The image above shows a screenshot of the cryptocurrency and graph of Bitcoin on CoinMarketCap.



Firstly the graph only shows the close price of the currency each day. Audience don't know the open price, high price and low price. Therefore, candlestick chart is a good choice for visualization.

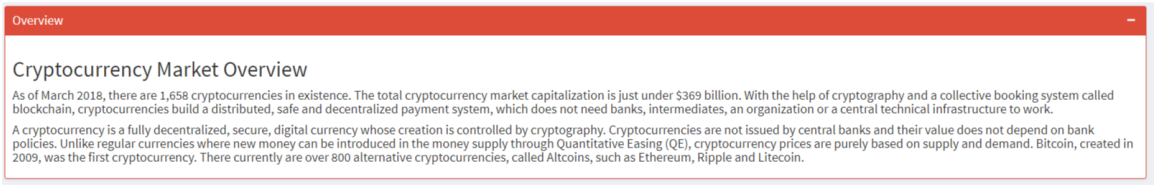
The cryptocurrency information in this website doesn't contain further analysis, such as correlation analysis and time series analysis. In this application, we will do the correlation analysis and time series analysis.

4. Design framework & Demonstration

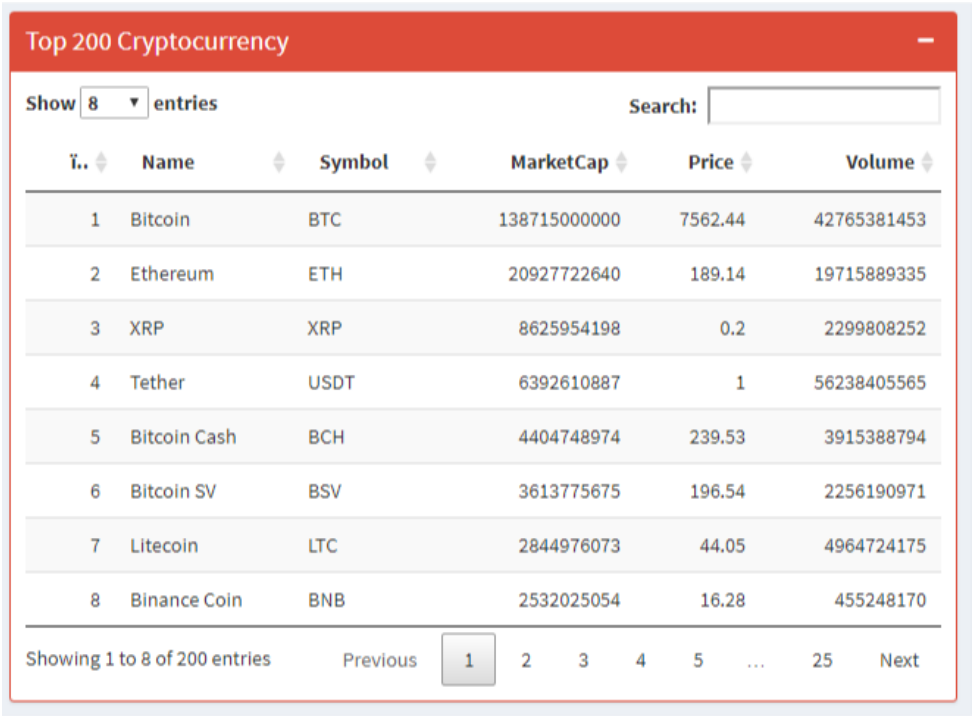
4.1 Design Principles

- User Centric: Create user guide to guide and educate on how to use the application.
- Interactivity: Use functions to give audiences the application in an interactive way.

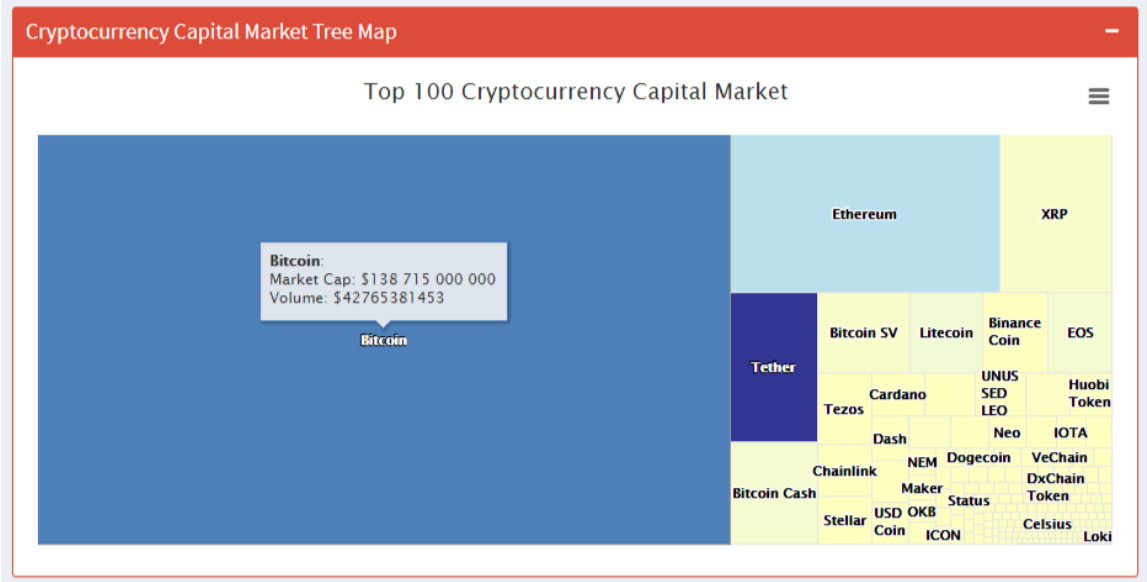
4.2 Cryptocurrency Market Overview



The textbox tells us an overview of cryptocurrency: what is cryptocurrency and what is cryptocurrency market nowadays.



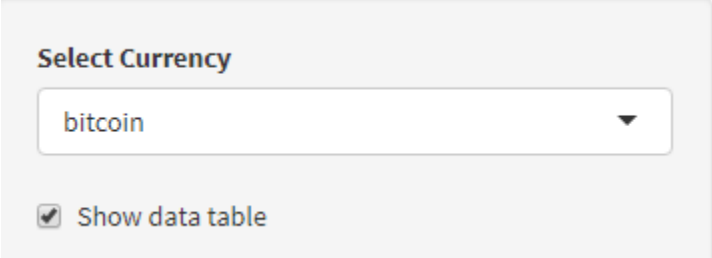
This table shows the top 200 crypto currency in the world according to their market capital. We can not only get the market capital from this table, but also other useful information, such as trading volume and close price.



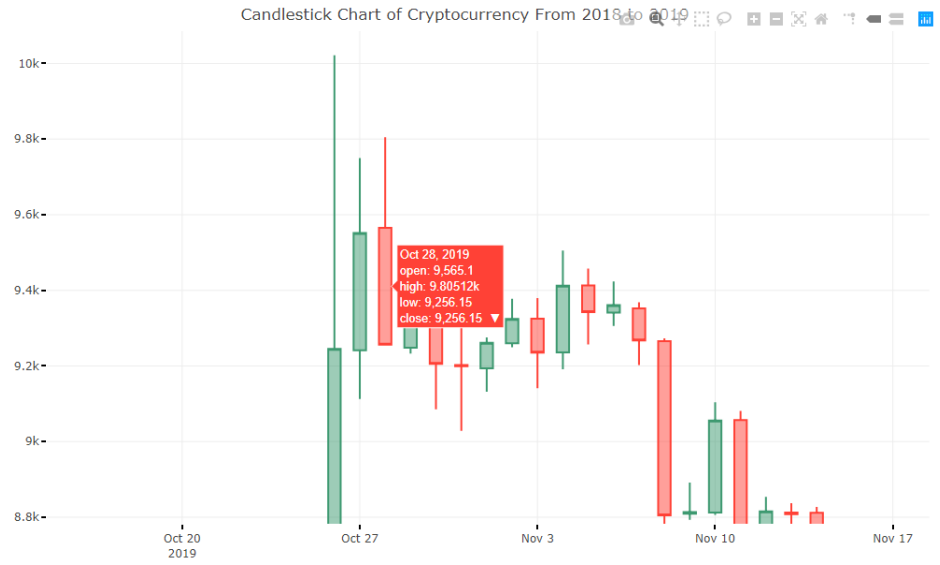
Tree map are basically rectangles placed adjacent to each other. The levels in the hierarchy of the tree map are visualized as rectangles containing other rectangles. Each set of large rectangles on the higher level in the hierarchy represents the Industry which stocks belong to. And each individual rectangle on lower level shows the top 100 crypto currency in the market. The size of each rectangle is directly proportional the Market Capitalization of each crypto currency in the market. The color in each rectangles represents the trading volume. The blue shows the higher stock price whereas the yellow shows lower. For example, as you can see from the tree map, Bitcoin has the largest capital among the others. In this receipt, we implement tree map using the “treemap” package. If user move the cursor to the whole industry, the diagram will pop up a window showing the market capital and volume.

4.3 Candlestick Chart

This Candlestick chart shows the historical data from 2018 to 2019 of one of the ten popular cryptocurrencies. Data contain open price, close price, highest price and lowest price of each day during this period. This chart provides an overview of the trend and historical price details in some certain periods. In order to explore the historical data of different cryptocurrencies, there is a selection panel to select certain cryptocurrency.



This chart is interactive that tools are available on the upper right corner of the chart. We can select a certain period and zoom it in order to explore the detailed information.



A check box is available below the currency selection panel to display the detailed information as a data table.

Datatable

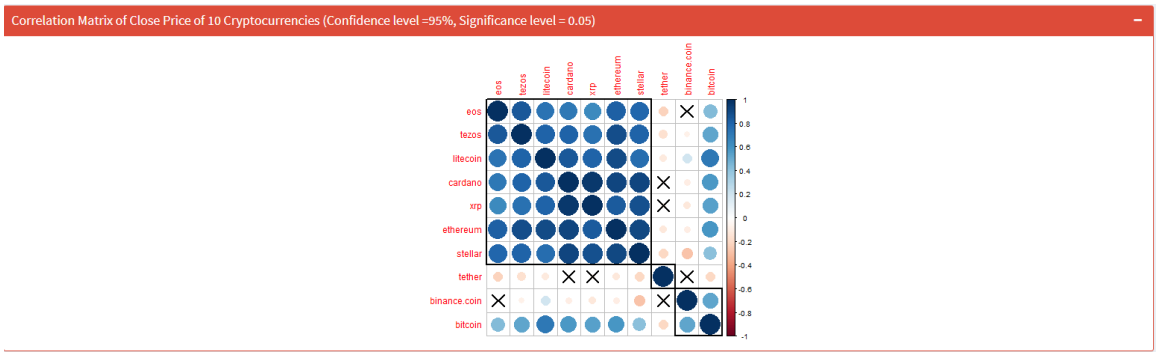
Show 10 entries Search:

Currency	Date	Open	High	Low	Close	Volume	Market.Cap
bitcoin	2019-12-31	7294.44	7335.29	7169.78	7193.6	21167946112	130000000000
bitcoin	2019-12-30	7420.27	7454.82	7276.31	7293	22874131672	132000000000
bitcoin	2019-12-29	7317.65	7513.95	7279.87	7422.65	22445257702	135000000000
bitcoin	2019-12-28	7289.03	7399.04	7286.91	7317.99	21365673026	133000000000
bitcoin	2019-12-27	7238.14	7363.53	7189.93	7290.09	22777360996	132000000000
bitcoin	2019-12-26	7274.8	7388.3	7200.39	7238.97	22787010034	131000000000
bitcoin	2019-12-25	7325.76	7357.02	7220.99	7275.16	21559505149	132000000000
bitcoin	2019-12-24	7354.39	7535.72	7269.53	7322.53	22991622105	133000000000
bitcoin	2019-12-23	7508.9	7656.18	7326.19	7355.63	27831788041	133000000000
bitcoin	2019-12-22	7191.19	7518.03	7167.18	7511.59	23134537956	136000000000

Showing 1 to 10 of 730 entries Previous 1 2 3 4 5 … 73 Next

4.4 Correlation Analysis

In this section, we use another dataset which is derived from the main dataset that its column [1] is Date, while column [2:11] are different cryptocurrencies. The numeric data are close price from 2018 to 2019.



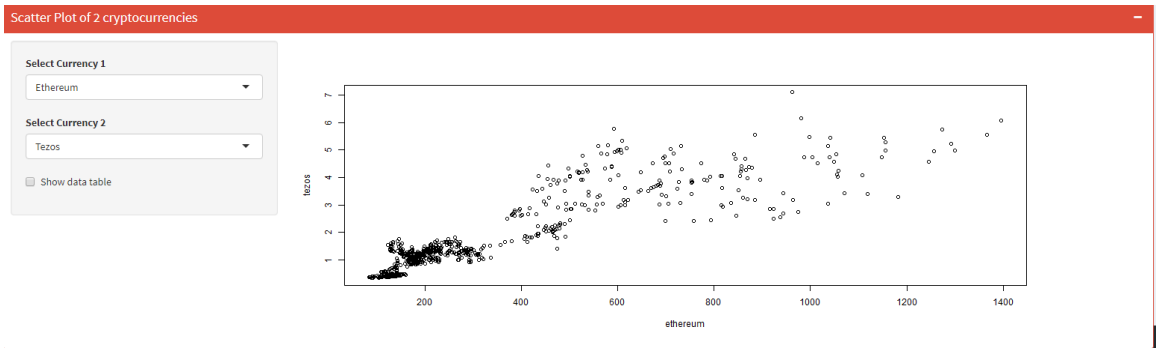
This correlation matrix is calculated using Pearson Correlation Formula:

$$r = \frac{\sum (x - m_x)(y - m_y)}{\sqrt{\sum (x - m_x)^2 \sum (y - m_y)^2}}$$

And p-value is also calculated using Pearson formula with 95% confidence level and 5% significance level. No significant coefficient will be shown as a cross on the chart.

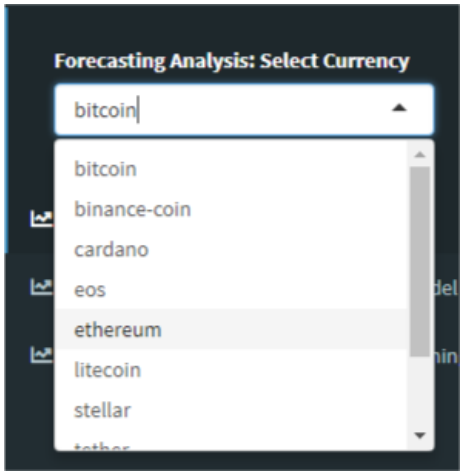
Rectangles on the chart represent the results of hierarchical clustering in order to highlight the correlated cryptocurrencies as groups.

To explore the details between two cryptocurrencies, a scatter plot is available that we can choose the pair through changing the selection box on the left. Detailed data table is also provided.

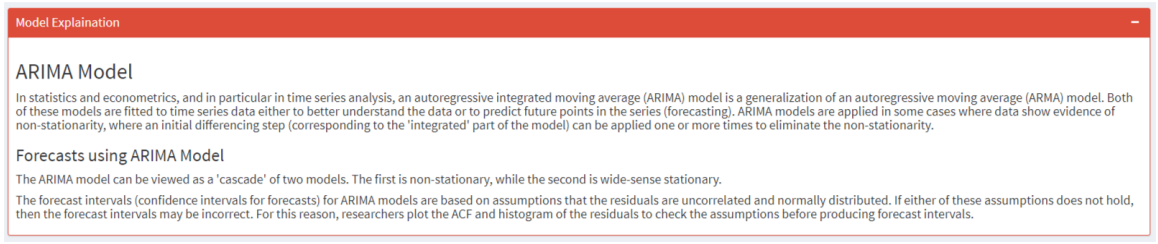


4.5 Time Series Analysis

On the technical indicator selection panel, users can choose from the following cryptocurrency for further time series forecasting.



1) ARIMA Model



Model explanation box will briefly introduce the model in this analysis. Model 1 is ARIMA model, in statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). ARIMA models are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the 'integrated' part of the model) can be applied one or more times to eliminate the non-stationarity.

In forecasting, The ARIMA model can be viewed as a "cascade" of two models. The first is non-stationary:

$$Y_t = (1-L)^d X_t$$

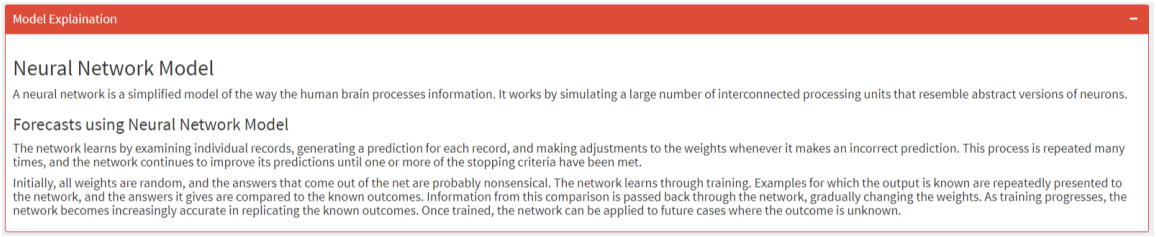
while the second is wide-sense stationary:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) Y_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t \,.$$

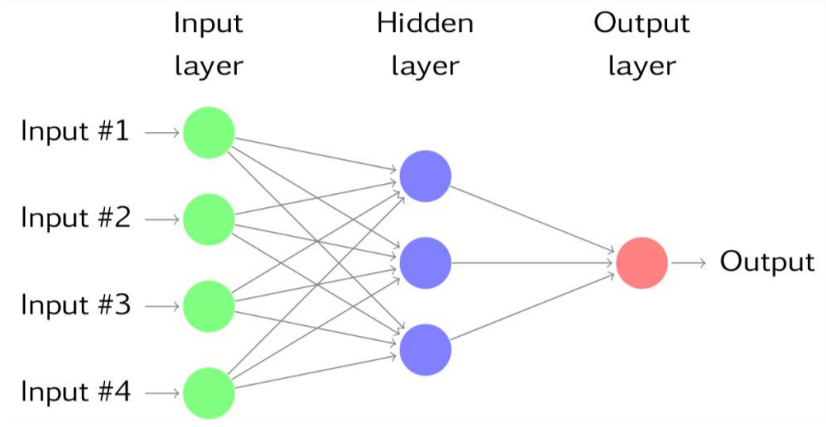
Therefore, forecasts can be made for the process Y_t, using a generalization of the method of autoregressive forecasting.

This plot not only shows the historical data for the currency that user selected. But also the predict price for the future by using ARIMA model. In addition, it also gives audience 95% confidence line and 80% confidence line for further analyze.

2) Neural Network Model



Model explanation box will briefly introduce the model in this analysis. Model 2 is Neural Network model, it is a simplified model of the way the human brain processes information. It works by simulating a large number of interconnected processing units that resemble abstract versions of neurons.



The network learns by examining individual records, generating a prediction for each record, and making adjustments to the weights whenever it makes an incorrect prediction. This process is repeated many times, and the network continues to improve its predictions until one or more of the stopping criteria have been met.

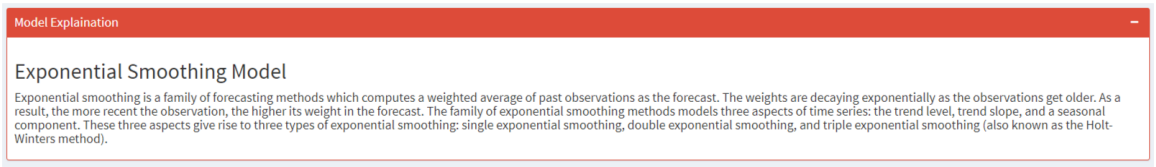
Formula:

$$z_j = b_j + \sum_{i=1}^{4} w_{i,j} x_i.$$

Initially, all weights are random, and the answers that come out of the net are probably nonsensical. The network learns through training. Examples for which the output is known are repeatedly presented to the network, and the answers it gives are compared to the known outcomes. Information from this comparison is passed back through the network, gradually changing the weights. As training progresses, the network becomes increasingly accurate in replicating the known outcomes. Once trained, the network can be applied to future cases where the outcome is unknown.

This plot not only shows the historical data for the currency that user selected. But also the predict price for the future by using Neural Network Model. This plot is an interactive plot that allow user to get the detail information when move the cursor to the line chart.

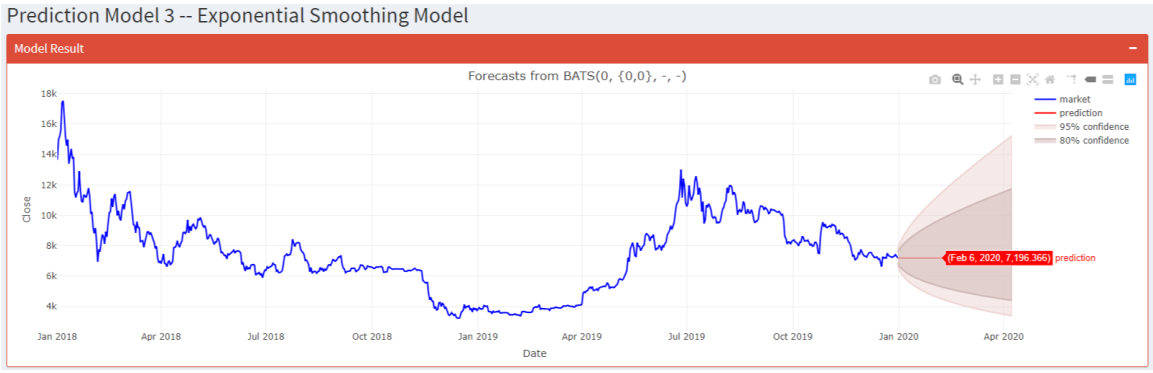
3) Exponential Smoothing Model



Model explanation box will briefly introduce the model in this analysis. Model 3 is the Exponential Smoothing Model, it is a family of forecasting methods which computes a weighted average of past observations as the forecast. The weights are decaying exponentially as the observations get older. As a result, the more recent the observation, the higher its weight in the forecast. The family of exponential smoothing methods models three aspects of time series: the trend level, trend slope, and a seasonal component. These three aspects give rise to three types of exponential smoothing: single

exponential smoothing, double exponential smoothing, and triple exponential smoothing (also known as the Holt-Winters method).

Formula: $$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t,$$



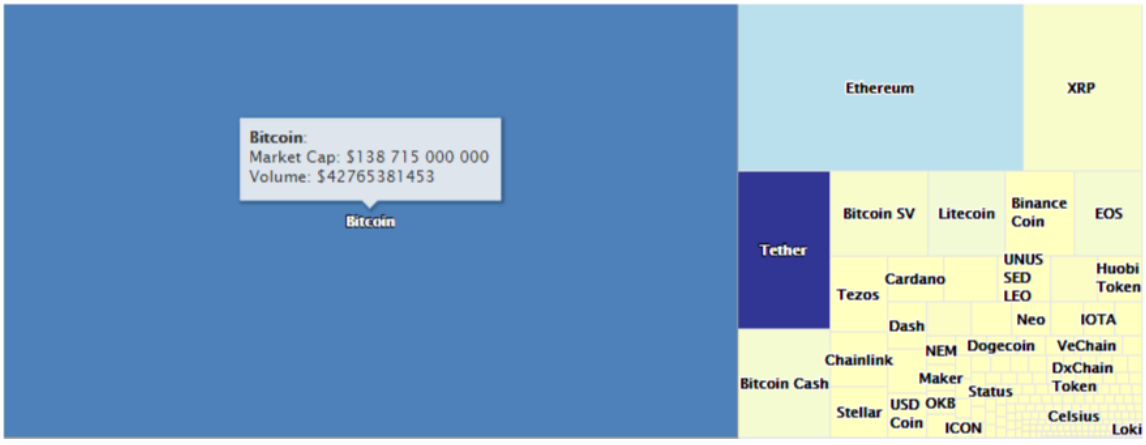
This plot not only shows the historical data for the currency that user selected. But also the predict price for the future by using Neural Network Model. This plot is an interactive plot that allow user to get the detail information when move the cursor to the line chart.

This plot not only shows the historical data for the currency that user selected. But also the predict price for the future by using ARIMA model. In addition, it also gives audience 95% confidence line and 80% confidence line for further analyze.

5. Discussion

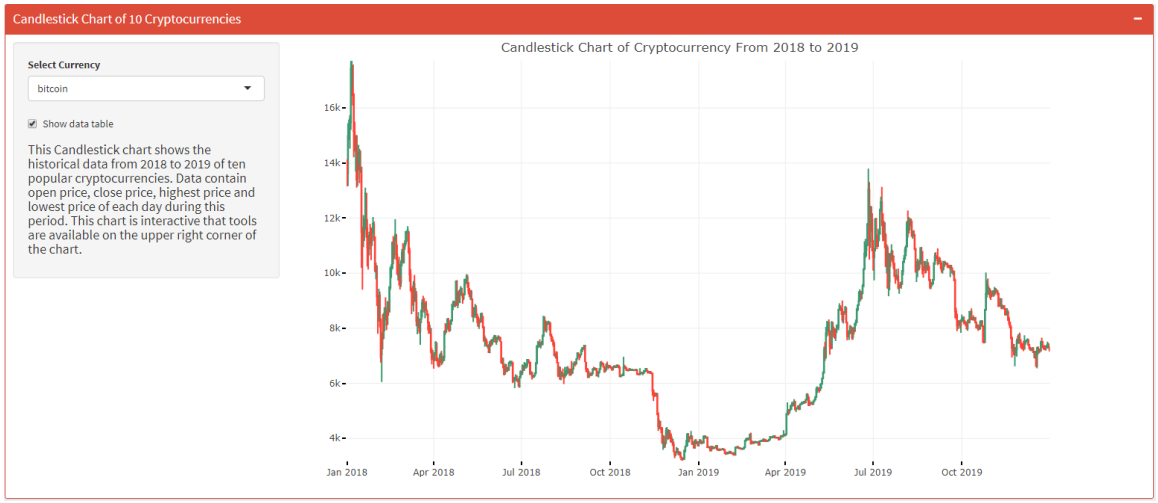
5.1 Insight

1) Market capital and Volume among crypto currency

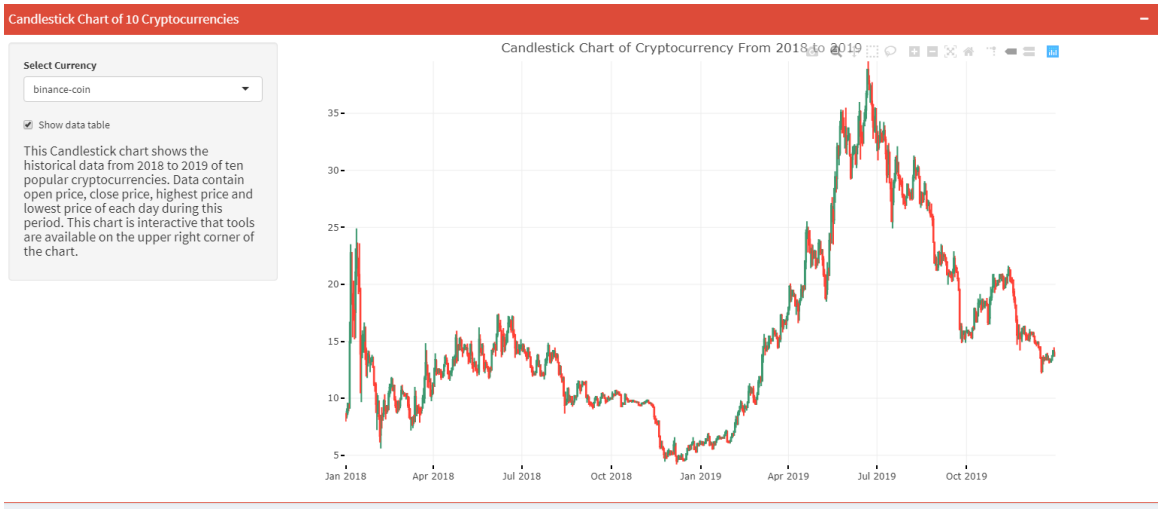


As we can see from the tree map above, Bitcoin has the largest market capital, while Tether's trading volume rank the first position. Besides Bitcoin and Tether, Ethereum, XRP, Bitcoin SV, Litecoin, Binance Coin, EOS, Tezos and Cardano are the top 10 in the crypto currency market.

2) Candlestick chart of cryptocurrency historical data



Bitcoin

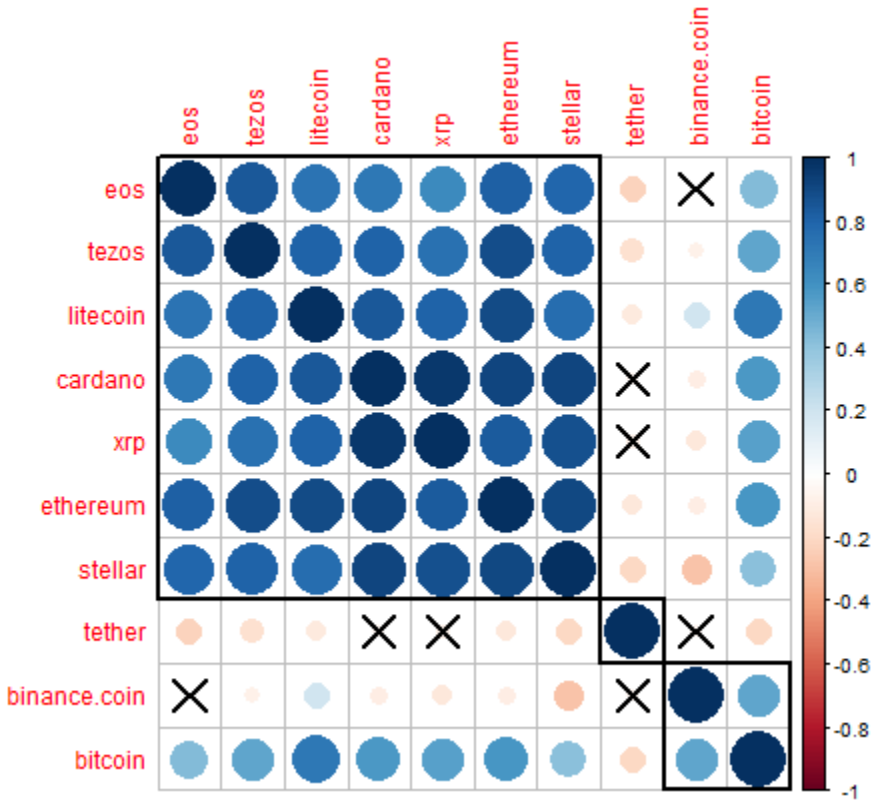


Binance-coin

From the two candlestick charts of Bitcoin and Binance-coin from 2018 to 2019, we can see that they both show the pattern of undulation, while other 8 cryptocurrency show an obvious decline trend. Cryptocurrency price dropped down because of the highly leveraged nature of the cryptocurrency market and the unwillingness of buyers to step in

amid extreme volatility and uncertainty. However, Industry executives remain positive after the big drop.

4) Correlation Analysis



From the correlation matrix we can find out that the cryptocurrencies have been divided into three clusters. EOS, Tezos, Litecoin, Cardano, Xrp, Ethereum and Stellar show relatively strong positive correlations. Along with the conclusion in former section, this situation might be because the whole industry has been influenced by same reason. Bitcoin and Binance-coin might be more sensitive facing industrial impact so that they show the pattern of undulation. This may draw a conclusion that cryptocurrency market and price might be strongly influence by common facts such as industrial changes, technology and regulation.

5) Time series forecasting

When we focus on the model 2, we can find that the model gives us the result that the price of Bitcoin will increase in the near future. Furthermore, due to the volatility of the bitcoin, it is difficult to make an accuracy prediction by using a simple predict model.

5.2 Lesson Learnt

Time Series Data: For the creation of Time Series Analysis, it was important that time IDs were present and in the right format so that the further visualization and analysis can be deployed, hence “lubridate” and “XTS” package had to be used to overcome the formatting of the time IDs.

Data preparation is very significant that a tiny data format error might lead to plotting failure.

6. Future Work

1) Real Time Interactivity

Currently, our application only visualize historical cryptocurrency market price data. Moving forwards, we will link our application to the cryptocurrency website to get real time interactivity. The number of cryptocurrencies is increasing rapidly, while this project mainly focus on top 10 cryptocurrencies. Enlarging the dataset and adding new cryptocurrency can be done in the future work.

2) Time Series Forecasting

Perform model test result for a large number of data with different forecasting models to choose the best model.

7. Package

List of R Packages/Libraries Used shiny, shinydashboard, lubridate, plotly, TTR, vcd, tseries, forecast, treemap, highcharter, tidyverse, corrplot.

8. Reference

Data Source: https://coinmarketcap.com/

Barone, A. (2019) The Future Of Cryptocurrency in 2019 and Beyond. *Investopedia*. https://www.investopedia.com/articles/forex/091013/future-cryptocurrency.asp

Bagshaw, R. (2020) Top 10 cryptocurrencies by market capitalisation. *Coinrivet*. https://coinrivet.com/top-10-cryptocurrencies-by-market-capitalisation/

Derbentsev, V., Datsenko, N., Stepanenko, O., & Bezkorovainyi, V. (2019). Forecasting cryptocurrency prices time series using machine learning approach. *SHS Web of Conferences*, *65*. https://doi.org/10.1051/shsconf/20196502001

Ehlers, S., & Gauer, K. (2019). Beyond Bitcoin: A Statistical Comparison of Leading Cryptocurrencies and Fiat Currencies and Their Impact on Portfolio Diversification. *The Journal of Alternative Investments*, 22(1), 114–125. https://doi.org/10.3905/jai.2019.1.072

Radovanov, B., Marcikic, A., & Gvozdenović, N. (2018). A Time Series Analysis of Four Major Cryptocurrencies. *FACTA UNIVERSITATIS - Economics and Organization*, *15*(3), 271–278.

Sajter, D. (2019). TIME-SERIES ANALYSIS OF THE MOST COMMON CRYPTOCURRENCIES *. *Ekonomska Misao i Praksa*, *28*(1), 267–282. Retrieved from http://search.proquest.com/docview/2249689505/

Young, J. (2020) Crypto Traders Explain What Caused the Bitcoin Price Plunge to the $3K Range. *Cointelegraph*. https://cointelegraph.com/news/crypto-traders-explain-what-caused-the-bitcoin-price-plunge-to-3-000