SI618 Individual Project Proposal
Chongdan Pan • pandapcd • pandapcd@umich.edu

1 Overarching Research Question

Are cryptocurrencies good assets for normal people and retail investors to invest in?

2 Summary and Motivation

Recently, cryptocurrencies like Bitcoins have become the hottest topic all over the world. Cryptocurrencies are being used in multiple areas including decentralized finance, application development, art collections, and etc. The market value also increased dramatically in the past 10 years, from less than one hundred dollars to more than two trillion dollars.

To make a profit, many people and organizations are getting involved in this area. There is news that Grayscale, an investment institution, owns cryptocurrencies assets worth more than 20 billion dollars. Through mobile trading applications like PayPal, it's easy for even normal people or retail investors to buy and trade cryptocurrencies at any time. However, such a low entry barrier and mania also implies high risk. Investors can make a great fortune thanks to the 10 times increment of price in less than half of one year, but there is also news that people lose a family fortune because the price may plummet 50% in 10 hours.

This project aims at analyzing whether cryptocurrencies are good assets for retail investors based on previous trading data of some cryptocurrencies and other typical indexes like S&P 500 or 10-year treasury. Simple trading strategies will be built based on retail investors' habits and psychology. The process of investment will be simulated by using these strategies on different assets. Finally, the project will come to a conclusion based on the quantitative result of the simulation.

3 Choose and describe (at least) two different datasets

- Binance has a lot of public REST API providing cryptocurrencies' historical market data of different frequencies at different times. (https://binance-docs.github.io/apidocs/spot/en/#market-data-endpoints) The data are well-structured and have fields including High, Low, Open, Close, Volume, etc of a certain frequency. This project will mainly use daily historical data of main cryptocurrencies including Bitcoins and Etherum from Jan. 2017 to Sep. 2021.

- Yahoo Finance provides datasets of different symbols in certain periods. Unlike Binance's datasets, Yahoo's only have four fields about price High, Low, Open, Close. For example, US 10-year treasury data can be gathered by downloading the csv file. (https://query1.finance.yahoo.com/v7/finance/download/T̂NX?period1=1601929807&period2=1633465807&interval=1d&events=history&includeAdjustedClose=true) These data will mainly servers as a reference for our analysis of the cryptocurrencies, because they're the most widely used index and reference.

4 Describe how you might manipulate and join the two datasets

- It's worth mentioning that different assets have different trading hours. Cryptocurrencies can be traded 24/7 and never pause, while stock or treasury can be only traded during certain hours of weekdays. Therefore, I'll like to join these datasets based on the daily series. When focusing on the relation between cryptocurrencies and other assets, I'll analyze it week by week. For example, I'll calculate the return of all assets in each week for comparison

- For the US 10-year treasury, people typically mark its yield rate rather than its face value. Hence, I'll transform it to its inverse number, which can be regarded as the market price.

5 Describe at least three large-scale computation tasks you will perform to gain insights from the datasets

- Firstly, I will do a basic time series analysis on each asset week by week. The analysis including calculating the return, maximum drawdown, fluctuation range, and etc. Then I'll sort based on these values to find which asset in which week has the highest return or largest maximum drawdown. This information can be very helpful for us to get a general idea about these assets.

- Secondly, I may perform backtesting on some simple trading strategies and compute its return on different assets. These strategies typically conform to normal people or retail investors' psychology, such as automatic investment, trading based on last week's return, and etc. Like the first task, the return, maximum drawdown, and fluctuation range will also be calculated based on the strategies' return, but the calculation interval for these indexes is the whole range of the data rather than each week. For now, I'm not sure whether I will take the commission fee into consideration.

- Thirdly, I'll try to find the pattern or relation within or between different cryptocurrencies. Such calculation requires me to do some shifting so that I can try to predict the price in the future based on time series or cross-section analysis.

6 Describe at least one visualization you might create that highlights insights you hope to gain

- The first visualization I want to create for this project is quite common, which is a line chart showing the performance of the strategies as well as the raw price data. In addition to just showing the return and value of each strategy, I'd also put emphasis on where the return or drawdown is noticeably large. In addition, I'd probably change the axis into a log scale and do Z-score normalization so that we can make a fair comparison between these assets as well as strategies.

- Another visualization I want to create is a bar chart, which shows the number of weeks where the asset's certain index is out of a certain interval. For example, Bitcoin may have 30 weeks with a return higher than 25%, and I will count it as 30.