Data science project   Chris Stokvik and Martin Bjørnerem

Our idea for a data science project is to look at the correlation between the bitcoin price and the number of google searches for 'bitcoin'. We want to analyse data from bitcoin and other cryptocurrencies. We will gather data from Google Trends and Kaggle, then merge the data for analysis. We will research whether the correlation varies between different cryptocurrencies and whether the correlation has increased or decreased in different periods.

Data on bitcoin interest and searches are available from Google Trends and we have found a dataset with daily cryptocurrency prices. Primary focus is on bitcoin and a few other well-known currencies. Also on Kaggle, we found a dataset with tweets about bitcoin containing sentiment. We are interested in how we can include public tweets in our analysis. The dataset is extremely large (1.4 million tweets), and we will need to extract a portion and analyse it before we decide exactly what we want to with it.

Once we have tidied our data and have a better understanding of it, we will merge the datasets and examine correlations between them. We will look at different time periods of the data to examine how the relationship has changed over time. We are specifically interested in the periods; before, during and after the bubble that burst at the end of 2017. We will be using a public GitHub repository to keep track of changes and progress to our project.

Current workplan

Week 1 – Access and download necessary data discussed above. Inspect the data.

Week 2 – Merge, tidy and arrange it to our needs.

Week 3 – Experiment various correlations analysis. Prices of different currencies in different time periods. This will be the foundation of our project.

Week 4 – Look closer at tweets and discuss possible ways to visualise/communicate the results. Is there a relationship between tweets and price? Do the tweets strengthen or undermine findings from previous week?

During these weeks we will also start thinking about and discussing the blueprint of our final report.

Data: https://www.kaggle.com/jessevent/all-crypto-currencies

https://www.kaggle.com/paul92s/bitcoin-tweets-14m