What is a cryptocurrency? At the most basic level, cryptocurrencies are a decentralized digital medium of exchange. Unlike their centralized financial counterparts, cryptocurrencies are not maintained by oversite organizations or a central server. But rather, transactions are recorded in a public ledger built on blockchain technology. The value of the cryptocurrencies​ ​fluctuate​ ​ ​relative​ ​to​ ​other​ ​currency​ ​forms and ​thus​ ​they​ ​can​ ​be​ ​traded​ ​for​ ​profit. PROFIT! Now I have your attention. In recent months, cryptocurrencies have moved away from being an obscure tech currency to being covered in more mainstream media. The visualization provided below is a stream chart that describes volume fluctuations in trading for a few of the more mainstream cryptocurrencies.

In late spring 2017 the market experienced a significant uptick in trading volume as can be observed in the stream graph. The interest of one of our team members was peaked by this increase in trading. The outcome of this peaked interested was a researchable question: what information sources could an individual or organization using to monitor and track the cryptocurrency market.

Looking through mainstream media it would not be difficult to find topic pieces on cryptocurrencies. Articles on this topic have been covered by media outlets such as the New York Times, Fortune and The Atlantic. Despite this increase in coverage, cryptocurrency markets are not covered in the same fashion as their centralized financial counterparts. Twitter seemed like the most obvious candidate. The twitter-sphere is​ ​similar​ ​to​ ​the​ ​cryptocurrency​ ​environment in that it​ ​is​ ​also​ ​highly​ ​unregulated​ ​and​ ​attracts​ ​a segment​ ​of​ ​the​ ​population​ ​actively​ ​engaged​ ​with​ ​technology. ​

Twitter maintains a public API through which users can query the twitter data stream on topics of interest. The dataset for this analysis was developed by querying the twitter stream three times daily over the course of five days. The keywords used to query the API were types of cryptocurrencies. The names of which included Bitcoin, Etherium, Litecoin and XRP. One of the first steps in understanding an information stream is to determine the volume and velocity at which data is being generated. In this context, the unit of data is a tweet the volume is the number of tweets and the velocity is the speed at which they are produced.

The visualization on the left shows the total volume of the tweets by currency collected during the data collection period. As can be seen in this visualization, over 50% of the tweets included Bitcoin as a key word. Bitcoin was the first cryptocurrency to be invented and remains one of the most traded currencies and this result is therefore unsurprising.

The primary question of interest for this analysis was whether data contained in tweets could be used to gather information about the cryptocurrency markets. To achieve this goal, a natural language processing python library was used to analyze the sentiment of the tweets. The polarity of the tweet was defined as a number between -1 and 1 rating the degree of negativity or positivity in the language of the tweet. The results of the analysis are provided in the visualization below.

The change in sentiment was calculated for all the tweets as well as by cryptocurrency. The results for average sentiment by cryptocurrency proved unreliable. This is unsurprising given the smaller sample size of tweets associated with individual cryptocurrencies. However, the results for the sentiment values calculated using all tweets proved promising.

The Take Away

For organizations interested in the cryptocurrency markets, twitter has the potential to be a valuable information source. The average sentiment of all tweets related to the cryptocurrencies of interest experienced value fluctuations that mimicked the market. Monitoring twitter via a live data stream set to query keywords can be implemented in a straightforward manner. Therefore, setting up mechanisms to monitoring and analyzing the sentiment of tweets would likely be an asset to organizations interested in financial modeling related to cryptocurrencies.