Crypto Dashboard

Fintech - Project 1

TEAM


Matt Newkirk
Michael Garcia
Bryn Lloyd-Davies
Justin Bernier

Motivation & Summary

To compile data from a Cryptocurrency API and Twitter API to determine if opinions of influential figures affected price and volume of cryptocurrency traded

Goals:

- Pull historical crypto data for Bitcoin, Ethereum, Litecoin
- Pull Twitter data from interesting influencers that tweet about crypto
- Get recent sentiment on cryptocurrency

Questions

1. Does Bitcoin lead other cryptocurrencies and is there strong correlation between Bitcoin, Ethereum, and Litecoin

2. Recent sentiment analysis based on crypto mentions in twitter data

3. Can we find any relationship or correlation between influencer tweets and price/volume action in Bitcoin?

Data & Data Clean-up

Two Primary sources of data

- Coinbase Pro Crypto API for historical Crypto Data
    - fetch data from through API, 2. manually fill dates to get range back to 2017 to present. 3. create raw csv of crypto data. 4. Create three separate df's with crypto specific summary statistics. 5. create df's focused on specific statistics 6. pull in twitter data. 7. merge all dataframes. 8. filter dataframes. 9.export clean data.
- Twitter API for tweets from influencers
    - Pull the tweets from the DF as a list and then apply the polarity analyzation to the list of tweets creating a list of lists, delete the tweets from the list of lists, convert to a DF, pull only polarity as a list, and then append to a new column in the original DF. Multiplied the polarity rating by the favorite count to quantify the positive or negative rating

Data Exploration

Describe the exploration and cleanup process.

- As we were creating visuals we had to spend a lot of time thinking about the best way to represent the two data sets and which metrics would be most valuable and relevant to the other data (ie volume correlated to popularity rating)

Discuss any problems that arose after exploring the data, and how you resolved them.

- separating time and date using str.split() and creating new columns
- having to do single merge calls for each dataframe. (could have created a list and looped through them)
- using a global panda function for rounding values in dataframes <pd.options.display.float_format = '{:.2f}'.format>

Data analysis

We established baselines such as the rolling average volume traded (see charts) so we had a baseline to tell if a spike had occurred in any individual day. Overlaying the data points for tweets with a positive or negative popularity rating then showed if highly interacted with tweets corresponded to days with spikes in trading volume

Discussion

- Developed a tool to look at possible correlation
- Can see when major influencers tweeted and corresponding crypto data
- We found correlation between crypto coins
- Found a way to get social media sentiment related to cryptocurrencies from major influencers
- Bitcoin is a leader of the other cryptocurrencies
- Overall we did see some correlation, but limitations to the data made it difficult to determine causation.

Postmortem

Limitations in the free version of the Twitter API, limitations on timing of cryptocurrencies (daily values only), scaling the data in ways that were relevant to each other (ie popularity rating with volume)

With more time pulling in targeted minute level crypto data based on tweet dates to see if we can detect direct impact of price/volume action immediately after significant influencers tweet.

Questions?

THANKS!!

CONTACT

Please feel free to contact us!

Justin Bernier
justin.bernier@protonmail.com
https://www.linkedin.com/in/justin-bernier

Micheal Garcia
Garciamicheal0@outlook.com
https://www.linkedin.com/in/micheal-garcia-3345a11b6/

Matthew Newkirk
mattnewkirk555@gmail.com

Bryn Lloyd-Davies
lloyddavies.bryn@gmail.com

Fonts & colors used

This presentation has been made using the following fonts:

Roboto Squada One
(https://fonts.google.com/specimen/Squada+One?selection.family=Squada+One)

Roboto Condensed
(https://fonts.google.com/specimen/Roboto+Condensed)


slidesgo