Cleaning Cleaning Python Pandas-- Jupyter Notebook Reformat date for Bitcoin API/JSON distaset Innner merge on date Columns: dropped due to non-relevance Columns: renamed to keep consistentsy Once cleaned seperately, cleaned once combined New dataframe created seperately for incless variable Factor/ease of use in database loading phase than searching through string • Challenge: Remove rows that contained empty fields in the Market Cap column with Bitcoin Bitcoin dataset (API/JSON): Bitcoin CoinGecko API/JSON Create loop call 2017-2013 dates Build in try-except Datafarer created from call: date, market cap, volume, current price columns. • Challenge: Extraction was difficult because of creation of loop and try-except Bitcoin Cash

Database

PgAdmins SGL used to load data
Data relations?
Only jained on one element-date
Was balkes created
Hades primary ley
Blooch "Bloom Cash combined date-foreign ley
References primary ley
References primary ley

Bitcoin Cash generally follows

the trend of Bitcoin • Hypothesize: Pulling in any

Analysis

- other coin, they would follow a similar trend
- · Bitcoin is the "gold" standard of cryptocurrencies.

Analysis

To further normalize, used Matplotlib to generate line graph based on series of common dates





 Bitcoin Cash dataset(CSV); Kaggle
Kaggle
Daily dates & stats
Date, open, high, low, close, volume, market cap **Extract**

LOAI

BITCOIN V. BITCASH

Christos Ferderigos, Zohna Jones, Omar Kreidieh, Anu Srivastava



Purpose

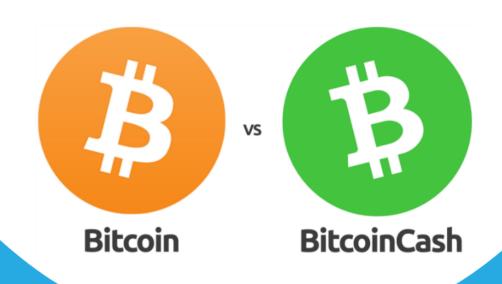
BITCOIN V. BITCASH

ristos Ferderigos, Zohna Jones, Omar Kreidieh, Anu Srivastava



Purpose

- Bitcoin and Bitcoin Cash: popular and increasingly profitable cryptocurrencies
- Purpose: Perform the ETL process on Bitcoin and Bitcoin Cash data to analyze the relationship.





Extract



Bitcoin Cash

- Bitcoin Cash dataset(CSV):
 - Kaggle
 - Daily dates & stats
 - Date, open, high, low, close, volume, market cap



Bitcoin

- Bitcoin dataset (API/JSON):
 - Bitcoin CoinGecko API/JSON
 - Create loop call 2017-2019 dates
 - Build in try-except
 - Dataframe created from call: date, market cap, volume, current price columns
- Challenge: Extraction was difficult because of creation of loop and try-except



Transform



Cleaning

- Python Pandas-- Jupyter Notebook
- Reformat date for Bitcoin API/JSON dataset
 - Innner merge on date
- Columns: dropped due to non-relevance
- Columns: renamed to keep consistentsy
- Challenge: Remove rows that contained empty fields in the Market Cap column with (-)



Cleaning

- Once cleaned seperately, cleaned once combined
- New dataframe created seperately for index variable
- Faster/ease of use in database loading phase than searching through string



LOAD



Database

- PgAdmin4 SQL used to load data
 - Data relational
 - Only joined on one element-date
- Two tables created
 - Index- primary key
 - Bitcoin/Bitcoin Cash combined dateforeign key
 - References primary key



Analysis

 To further normalize, used Matplotlib to generate line graph based on series of common dates



Analysis

