Scraping Coinbase for Crypto Data

Adam Gao

July 4, 2019

Summary:	Coinbase.com is a website displaying cryptocurrency prices of different cyrptoassets. Using the BeautifulSoup Python library, I scrape pricing data about each of the 52 different cryptocurrencies (Bitcoin, Ethereum, etc) from Coinbase. The data is converted from its HTML/javascript origins into a Python-friendly Pandas dataframe. The dataframes, which contain pricing information is saved as Excel files which can be re-opened by Python or Excel for easy visualization and analysis. The python files are organized as follows: • "Scrape-Coinbase.py" serves to scrape and save pricing data. • "Read-Data.py" serves an example on how to read and plot the Excel files in a useful manner.
Data Format:	"Scrape-Coinbase.py" creates a folder containing 52 Excel files and "datetime.txt" containing the date and time at the end of the extraction. Each Excel file's name corresponds to a cryptocurrency. In each file, there are 5 spreadsheets corresponding to different time durations of pricing. These durations are "hour", "day", "month", "year", and "all". Each row contains price and time. The price unit depends on the detected country.

Analysis:	"Read-Data.py" extracts each cryptocurrency's excel file into a pandas dataframe and then plots them over a chosen duration (past hour, day, month, year, or entire history)

0-Bitcoin, CAD, past year, 365 samples