**ETL Process:**

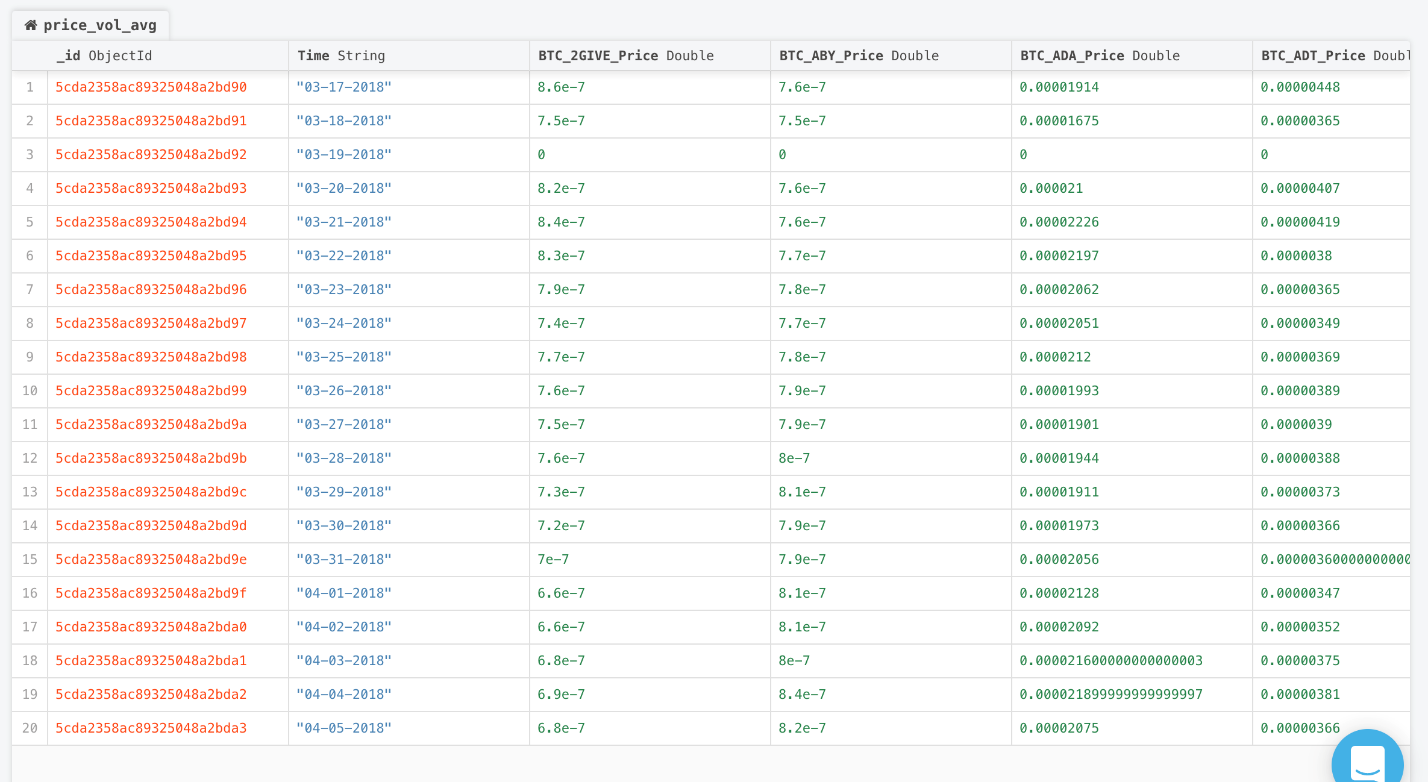
We extracted our datasets from a couple of csv files from a Kaggle repository entitled “bittrex-bitcoin-pairs”. The data can be accessed from this URL: https://www.kaggle.com/mhansinger/bittrex-bitcoin-pairs

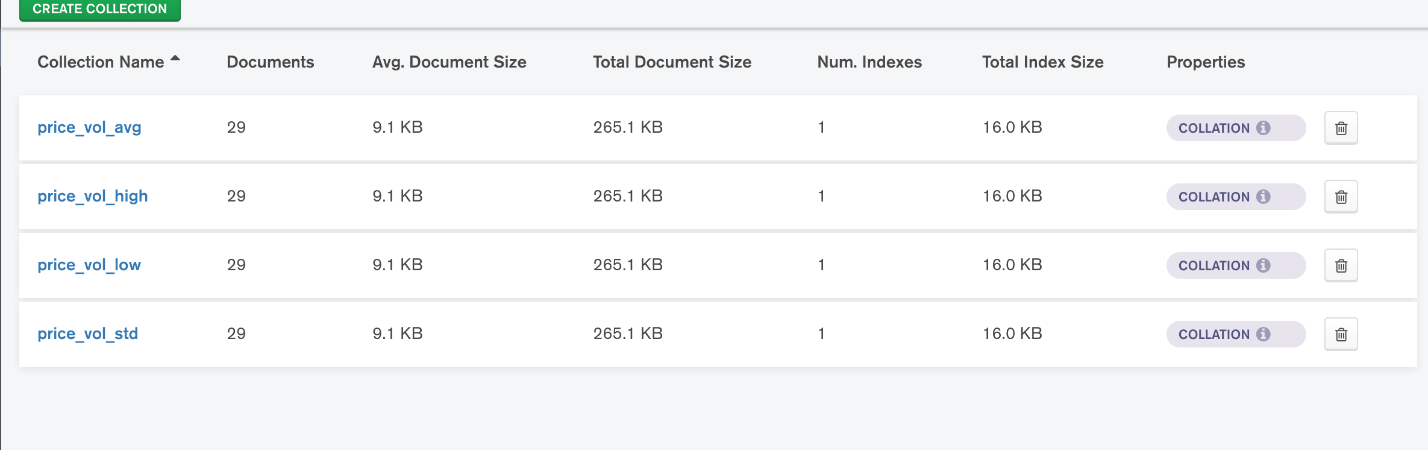
We took four csv files, and loaded them into Pandas. These datasets contained Ask, Bid, Price, and Volume information for each of 196 cryptocurrencies. The prices were listed in value relative to bitcoin. Each row of data represented one minute.

Our data had UNIX time as its time variable, so we had to convert that into a date-time string. We grouped by day, and output dataframes of the min(), max(), average, and standard deviation of the group-by object for both price and volume (for a total of 8 dataframes). We then joined the price and volume dataframes on those parameters for a total of four dataframes.

We then uploaded the joined dataframes to a local MongoDB database using this code:

mongoimport -d Cushing\_Griggs -c price\_vol\_std --type csv --file price\_vol\_std.csv –headerline





**Further Analysis:**

We wanted to see whether there was a correlation between the different cryptocurrencies’ volatilities.

We found that there was a single row which had 0 for all 196 columns. We dropped that row for the foregoing analysis.

We took the top five cryptocurrencies (decided on according to an arbitrary internet article), and wanted to analyze minute-by-minute volatility. To measure that, we took the percent difference from the previous value. Thus, we lost one row, but still had well over 40,000 to work with.

We provided descriptive statistics for each of the minute-by-minute volatilities below, and plotted them.

**Analytic result:**

Here is presented the volatility scores for each of the cryptocurrencies in question, first as descriptive statistics, and then as a plot (with the x-axis representing time).

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

We also created a correlation matrix of these values between the cryptocurrencies in R.

A screenshot of a computer

Description automatically generated

As you can see, there are no correlations. As an example, a scatterplot between the values for two cryptocurrencies is provided.

A close up of a logo

Description automatically generated

We conclude that there is no relationship between the volatilities of the value of different cryptocurrencies relative to bitcoin.