

# MFE 230J Assignment 3 Memo

Data Science Project on Cryptocurrency Returns & Trading Venues

23 Feb 2021

The top 10 cryptocurrencies by market cap in 2019 were: ADA, BCH, BTC, EOS, ETH, LTC, XLM, XMR, XRP, and XTZ. However, the current top 10 are (descending): BTC, ETH, XRP, BNB, DOT, USDT, ADA, LINK, UNI, and XLM (cryptocompare.com). Only half remained in the top ten within a span of ~2 years. Thus, it seems that the top cryptocurrencies by market cap is not stationary across time as new coins are added to the space.

Differences in prices shown across exchanges and the exchange with the highest price is not necessarily consistent across cryptocurrencies or across time. We notice that there are some consistent spikes in price variability across cryptocurrencies though, e.g. around July 2018, coinciding with an ICO boom. The best choice of exchange seems dependent on both the cryptocurrency and the time period. We additionally consider compliance later.

Aggregating across the entire sample, Coinbase had the highest level of activity for both BTC and ETH. Bitstamp had the second highest volume for BTC, while Gemini (2017) and Kraken (2018-19) had the second highest volume for ETH. Coinbase, followed by Bitstamp, had the most volume of BTC across months; Kraken had the second highest volume of ETH, except for Jan, Aug, and Dec, when Gemini was second. Volume appears highest in summer and winter months, and lowest in spring and fall months, creating an oscillating pattern. Volume by day of week shows that from Monday, volume increases into the middle of the week, then decreases into the weekend, with the two weekend days having the least volume.

The BTC/BCH fork of 2017-08-01 saw the price of BTC fall, similar to a stock split. Log returns also appeared to drop. Unlike Bitcoin, the ETC/ETH fork of 20 July 2016 appears to have caused a rise in the price and log returns of ETH, thus resembling a stock dividend.

Results from the MA(7) strategy showed cumulative returns of ~80% for BTC and ~10% for ETH by end of sample. Both strategies were found to outperform the underlying. The random forest BTC strategy outperformed the investment in the underlying BTC for part of the test set before a trend shift, ending up underperforming against a direct investment in BTC. Results for the ETH random forest strategy were similar. We conclude that the ML trading model is not viable for a hedge fund or other type of trader as it resulted in worse returns than investing in the underlying cryptocurrencies themselves.

The largest five cryptocurrencies by market cap were: BTC, ETH, LTC, XRP, and BCH. One- and four-week samples of hourly volume (sum of volume across all exchanges) show that there appear to be daily patterns and spikes. These likely correspond to peak trading in the Asian, European, and North American markets. We also see volume spike alignment across the five exchanges, which lends credibility to the validity of the reported volumes on the exchanges.

In plotting the actual vs. the Benford count distributions, it appears that the total volume reported approximately follows the theoretical Benford distribution. The results of our chi-square goodness of fit tests, however, showed that we reject the null hypothesis for both Coinbase and Gemini. This suggests that the self-reported volumes for Coinbase and Gemini may have been faked or fraudulent.

We conclude that we should be wary of trading on Coinbase and Gemini due to potential market manipulation.