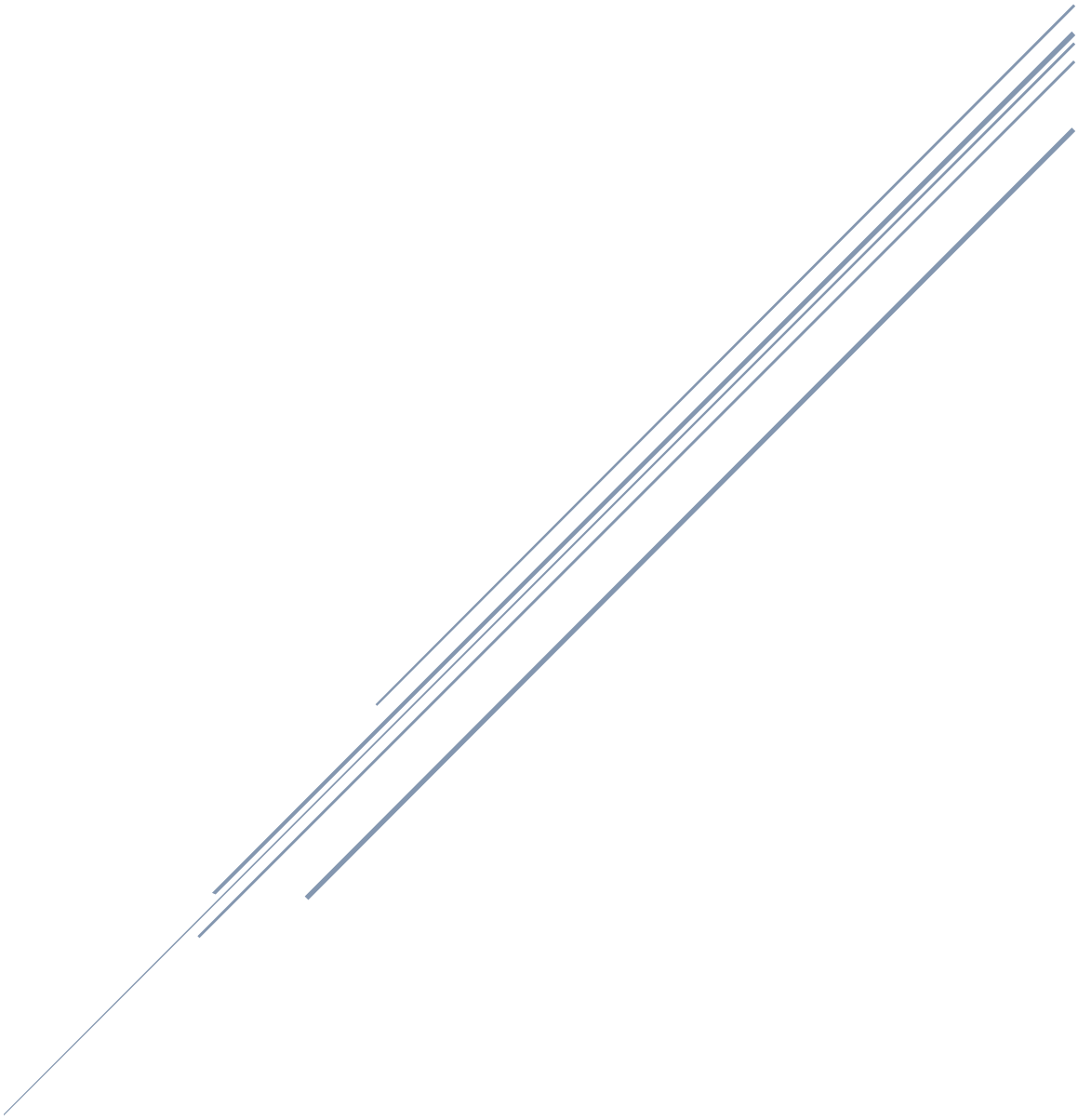
DATA ANALYSIS

Coinmarketcap vs. Coingecko



Data Analysis

Prior to Data Analysis

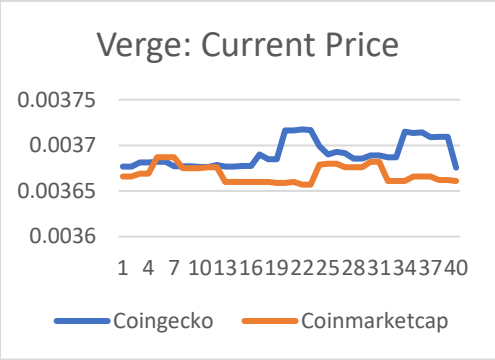
It is important, prior to any data analysis, to understand my dataset. As explained in my READme.md file, my parsed datasets are imperfect. Because I only parsed a fraction of the total files in my original folder, there was not as much of a difference between the data points, some even being almost identical. Due to this, I became increasingly interested in working with what I had.

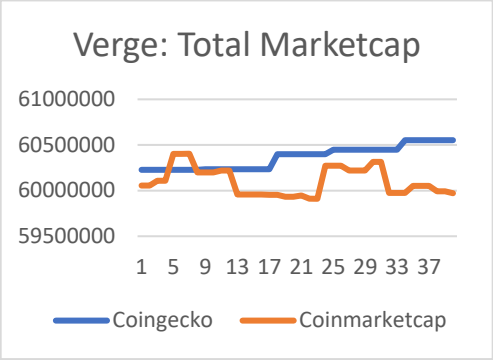
Explanation

I felt the easiest way to compare the datasets gathered from each website was through comparing individual currencies. What is important to see is the trends between individual cryptocurrencies between both websites.

I can say that both websites have many of the same cryptocurrencies (there are a few differences: i.e. Coinmarketcap has BigONE Token, while Coingecko does not), and these currencies have the same symbol. They were gathered close to the same time. I felt the differences would truly be seen in comparing individual cryptocurrency marketcap and prices between websites. Looking at my datasets, the data points were chosen on where there would potentially be a noticeable difference.

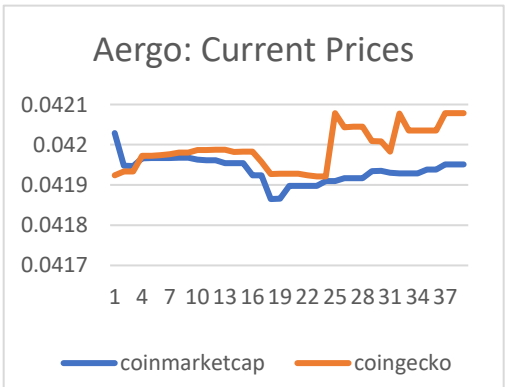
Let us first take a look at Verge (XVG), ranked 100-200, but a cheap cryptocurrency:

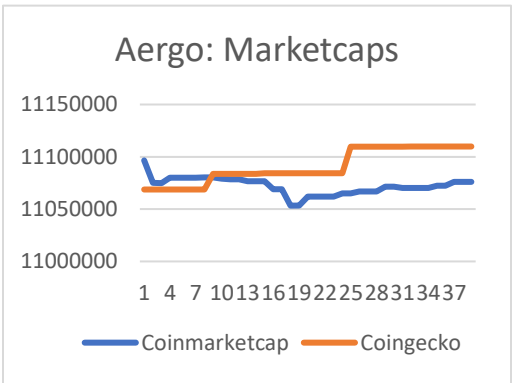




The graph on the left shows the relationship between the current price of Verge from the datasets for both Coingecko and Coinmarketcap. The y-axis shows the price, the x-axis is the number of data files were taken, meaning there were 40 different parsed files representing Verge from both Coingecko and Coinmarketcap. The graph on the right shows the total marketcap for Verge as given by both Coingecko and Coinmarketcap over 40 observations. Standardizing the data like this shows that on average, Coingecko holds a larger marketcap, and prices the cryptocurrency as higher. Keep in mind, relatively speaking, the differences are minimal from the websites, within a few cents of each other; but every hack trader who wants to make a few dollars can tell you that occasionally, a few cents makes all the difference. Regardless, I could not help but infer that whichever website has the larger marketcap, would also be charging more per cryptocurrency.

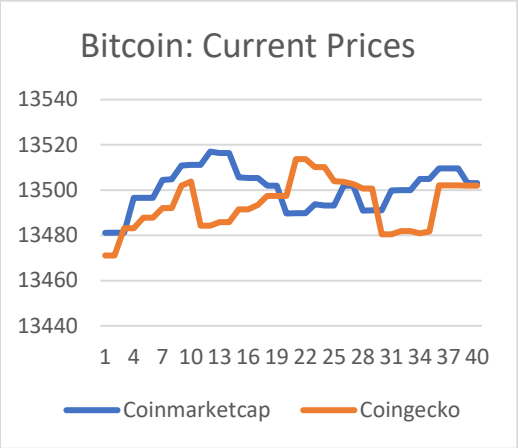
In order to test this theory, I decided to look at a few other currencies. First, I looked at a cryptocurrency ranked 300-400, Aergo (AERGO).

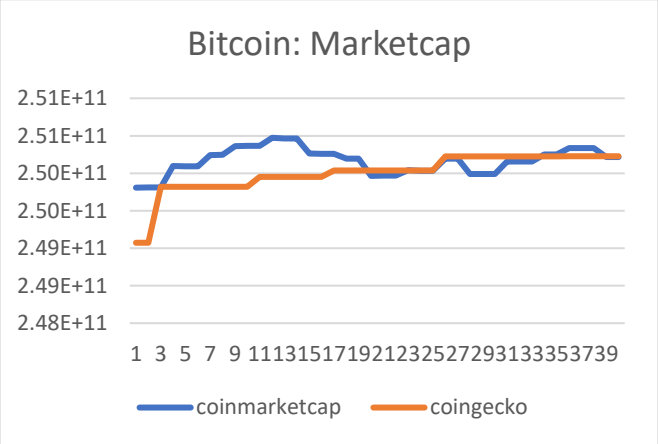




Even in lower ranked cryptocurrencies, this seems to still be true, with Coingecko leading the way in higher marketcap and higher prices. Even at the beginning of both graphs, you can see the correlation, that as Coinmarketcap had a higher marketcap, it also charged higher prices.

To test the theory once more, I decided to try it with the leading cryptocurrency, Bitcoin (BTC):





Surprisingly, even Bitcoin follows this same trend. Whichever website has a higher marketcap will also be charging more per cryptocurrency. As I said before, for these currencies, the differences are marginal. Even for a currency as large as Bitcoin, the differences here are only around $40 (compared to a currency that costs $1,500, this is not much). Because of the differences in these marketcap and prices, however, it is not out of the ordinary to assume that the mean of the data would present a significant difference between the websites (with each difference adding up, eventually the difference becomes large). Does this make one website more accurate than the other? On a micro scale, no. Each website presents marginal differences from each other that account for market share, global economy, company health, and a thousand other factors. On a macro scale, it is possible to have one website more accurate than the other, even purely based on marketcap and current prices alone. There would need to be more data gathered for a much longer period of time, as well as consumer buying habits, to truly

know the answer to this question. Luckily for each of these websites, the relatively new emergence of cryptocurrency as a legitimate market share has allowed room for these discrepancies.

Noteworthy

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.469827025
R Square	0.220737434
Adjusted R Square	0.220658545
Standard Error	1389.93803
Observations	19759

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10811403283	5405701641	2798.086889	0
Residual	19756	38167164164	1931927.726		
Total	19758	48978567446			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	152.1775687	9.928666159	15.3270909	1.01381E-52	132.7165483	171.6385891	132.7165483	171.6385891
price_change_24h	9.740981382	0.226414501	43.02278059	0	9.297189926	10.18477284	9.297189926	10.18477284
market_cap	4.94721E-08	8.454E-10	58.51920362	0	4.78151E-08	5.11292E-08	4.78151E-08	5.11292E-08

Along with my theory presented above, I had an additional theory that price_change_24h and marketcap were affecting the current price. The only difference between this theory and the one presented earlier in this paper, is this theory considers the differences in price in the past. My assumption was that the price of a cryptocurrency yesterday, and the total marketcap today, would affect the price of the cryptocurrency today.

To test this theory, I ran a regression test through Excel comparing the marketcap, price_change_24h, and current_prices for my entire Coingecko dataset. I found that while the significance of F and the P-values showed it was a trustworthy data, the value of R^2 only showed a 22% accuracy to this theory. While this may hurt my initial theory, I believe it actually shows that marketcap and the current price work in tandem, when one moves, so does the other, and this would happen regardless of what the price or marketcap were yesterday.