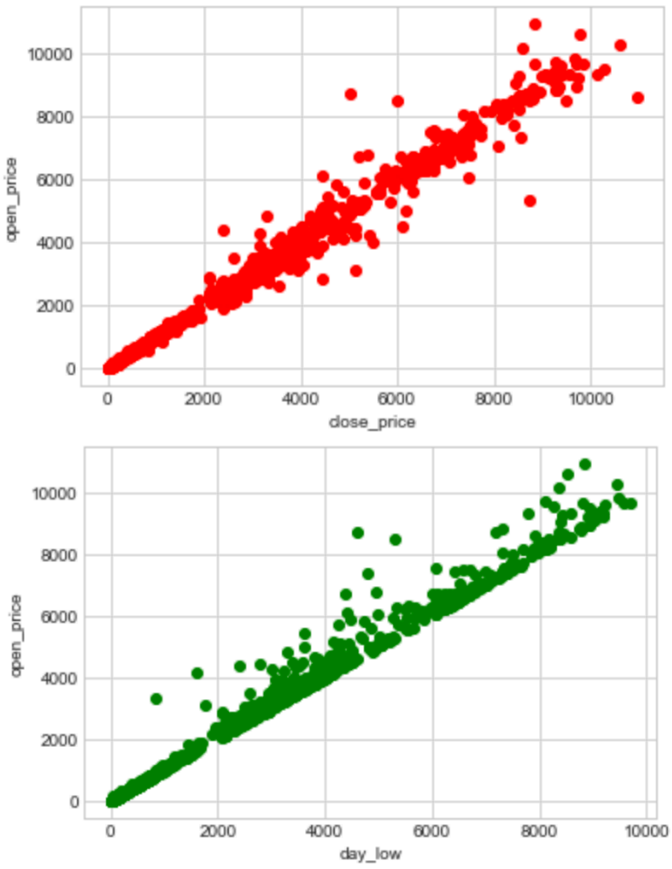
Randi Jones Due: April 23, 2019
Econ 9000- Machine Learning
Assignment 1

In this assignment, I scraped data about crypto-currencies, such as price, volume and market-cap, from coinmarketcap.com and parsed the data into a single CSV file. I then scraped the historical data for each individual crypto-currency from the past year and parsed that data into a CSV file as well. After properly parsing all of my data into workable files, I began my regression analysis.

Given the extensive range of data that is available to me with this assignment, there are a number of analytical questions that I could attempt to answer. One obvious question to consider is what factors influence a currency's opening price on a day-to-day basis? To answer this question, I first chose two variables, daily low price (day_low) and daily closing price (close_price), that I can reasonably assume will be positively and closely correlated with daily opening price (open_price). I then created two separate scatterplots of these two variables and the variable of interest, opening price. Below is the output of the scatterplots:

Randi Jones
Econ 9000- Machine Learning
Assignment 1

Due: April 23, 2019



As you can see, the two explanatory variables appear to have relationships consistent with what I would expect to see from this data. This is a good indicator that the data I am using is not corrupt or biased due to human error in the requesting and parsing procedures.

Now that I can reasonably assume that the data I am using is, in fact, good data and was not improperly coded along the way, I can continue my analysis of crypto-currency trends. First, I will check to see if there is a relationship between the closing price and the opening price. My intuition is that if a currency closes at a low price, then the opening price will be higher, as greater demand for the low-cost currency drives the price up. I ran a regression of these two variables, including a constant term. The results are posted below:

Randi Jones
Due: April 23, 2019
Econ 9000- Machine Learning
Assignment 1

Dep. Variable:	open	**R-squared:**	0.994
Model:	OLS	**Adj. R-squared:**	0.994
Method:	Least Squares	**F-statistic:**	2.518e+07
Date:	Tue, 23 Apr 2019	**Prob (F-statistic):**	0.00
Time:	20:30:47	**Log-Likelihood:**	-7.6644e+05
No. Observations:	161295	**AIC:**	1.533e+06
Df Residuals:	161293	**BIC:**	1.533e+06
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0871	0.070	1.245	0.213	-0.050	0.224
close	0.9987	0.000	5018.325	0.000	0.998	0.999

Omnibus:	261822.527	**Durbin-Watson:**	1.999
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	187145811492.046
Skew:	8.654	**Prob(JB):**	0.00
Kurtosis:	5279.949	**Cond. No.**	353.

Although the results show that the closing price has a significant effect on the opening price ($p>0.000$), the relationship is positive, indicating that a higher closing price is correlated with a higher opening price.

This is not the relationship that I was expecting, but it is possible that a low closing price does not have time to affect the opening price because demand cannot be evaluated until the opening of the day. Therefore, I restructured my hypothesis to assess whether the value of the closing price has an effect on the highest daily price. The intuition, then, is that when a crypto-currency has a high closing price, owners will want to sell immediately the next day to capitalize on gains. Conversely, if a currency has a low closing price, many people may want to purchase that currency the next day in anticipation of capitalizing on it in the future. If my intuition is correct, an exceptionally high(low) closing price should cause the daily low(high) to be extreme as well. To examine this hypothesis, I ran two separate regressions: one with day_high as the

response variable and one with day_low as the response variable. The results are posted

below:

Dep. Variable:	day_high	**R-squared:**	0.993
Model:	OLS	**Adj. R-squared:**	0.993
Method:	Least Squares	**F-statistic:**	2.459e+07
Date:	Tue, 23 Apr 2019	**Prob (F-statistic):**	0.00
Time:	20:38:40	**Log-Likelihood:**	-7.7589e+05
No. Observations:	161295	**AIC:**	1.552e+06
Df Residuals:	161293	**BIC:**	1.552e+06
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0271	0.074	0.365	0.715	-0.118	0.173
close	1.0465	0.000	4959.137	0.000	1.046	1.047

Omnibus:	567306.286	**Durbin-Watson:**	2.000
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	379312467893.026
Skew:	68.899	**Prob(JB):**	0.00
Kurtosis:	7514.396	**Cond. No.**	353.

Dep. Variable:	day_low	**R-squared:**	0.994
Model:	OLS	**Adj. R-squared:**	0.994
Method:	Least Squares	**F-statistic:**	2.661e+07
Date:	Tue, 23 Apr 2019	**Prob (F-statistic):**	0.00
Time:	20:45:57	**Log-Likelihood:**	-7.5468e+05
No. Observations:	161295	**AIC:**	1.509e+06
Df Residuals:	161293	**BIC:**	1.509e+06
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0365	0.065	0.561	0.575	-0.091	0.164
close	0.9543	0.000	5158.168	0.000	0.954	0.955

Omnibus:	548396.213	**Durbin-Watson:**	1.999
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	268755075277.032
Skew:	-62.332	**Prob(JB):**	0.00
Kurtosis:	6325.504	**Cond. No.**	353.

Randi Jones Due: April 23, 2019
Econ 9000- Machine Learning
Assignment 1

I was expecting to see that a higher closing price led to a lower daily low, but that is not what the results show. The results show that both closing price is significant with both variables, but they are also both positive.

Further analysis would be essential to uncovering the full extent of the relationship between these variables. Using machine learning, I would write a program that could scrape the crypto-currency data in real time, so that a trend analysis would be more accurate. Using a real-time scraper, I could track the changes in currency prices throughout the day to better determine what effect closing price has on currency prices the following day, and to what extent. There are a number of other consideration that I could model using machine learning, but for the purposes of this assignment, this will have to do.