Crypto Currency Trading Data

Dataset Description:

Historical trading data of Crypto currencies:

- Source: From kaggle website, download dataset crypto-markets.csv
- URL: https://www.kaggle.com/jessevent/all-crypto-currencies
- Observations: 785,024, Variables: 13, Crypto Tokens: 1,644
- 'slug' is the unique symbol for each Crypto currency. 'symbol' and 'name' are duplicate in some cases. Type: string.
- 'open', 'high', 'low', 'close' values for each Crypto currency on trading days up to 21 May 2018, 'volume' to measure trading amount, 'market' is the total market size = units * price per units of currency. Close Ratio = (Close-Low)/(High-Low), Spread is the $USD difference between the high and low values for the day. Lifetime (in yrs) is the existing time of a currency. Type: float
- Concatenated data: 'open', 'high', 'low', 'close' for USD, Gold price and SP500. Download csv file from yahoo finance. Type: float

Interested Research Questions:

1. How is the market of crypto currency emerging? Which currency has a leading position in the market? Do they survive or die over time?
2. Explore the statistical characteristics of daily log return, how can we model it with distribution? Are different currencies show similar pattern?
3. Crypto currency are virtual assets, do they have correlation with real world financial assets, e.g. USD, Gold price, stock indices. Can we use regression models to explain the dynamics of Crypto by real world financial assets?
4. What is the best currency to buy? Use machine learning to select a currency with best performance.

Analysis Method:

1. Do visualization (seaborn barplot) of number of currencies, trading volume, market cap over time, determine if there are stages of development. Pie chart for market composition, ratio of each currency of total market size. Research the competition dynamic over time.
2. Group by each currency, calculate its daily log return. (seaborn displot) for EDA. Then model with a distribution, QQ-plot to see if it is a good match. Do it for different currencies, come up with metrics to measure their performance. (mean and std dev of returns etc)
3. Merge USD, Gold price, stock indices data with prices of Crypto currency. Lineplot for time series plot. Heatmat for correlation matrix between different asset classes. Try different regression models, maybe time series models to explain Crypto currency price movement.
4. Given the metrics like mkt cap, trading, volume, spread, survival time, use machine learning algorithms to predict the best currency. Try to add some technical indicators as well, e.g. moving average, oscillator. Try classification, which groups of currency are good candidates while others are not. Also, may use dimensional reduction, because some of 13 variables are redundant.