Todd Gavin

DSCI510

18 November 2022

**Ideas:**

* How weather affects global food supplies?
* How sentiment on crypto twitter influences crypto currency prices?
* Forest Fire Prediction
* How do the trades of major crypto wallets influence crypto currency prices?
  + Can the trades of major crypto wallets be used as an indicator for major crypto currency price actions?
* How do the trades of U.S. politicians influence stock market prices?
  + Can the trades of U.S. politicians be used as an indicator for major stock market price actions?
* How does crypto twitter sentiment affect the prices of major NFT collections?

**APIs:**

* Etherscan API – get transactions for a particulate wallet address
* CoinMarketCap API – get current and historical cryptocurrency price data

**DSCI 510: Principles of Programming for Data Science**

Fall 2022 Homework 2  
Plan for Final Project **Due: 11/18/2022 23:59 PM PT**

* Please submit a PDF file answering the following 3 questions.
* Please be concise and informative.
* Use of bullet points is encouraged.
* No more than 1 page.

1. What’s the name of your final project? Please describe it as a research question and provide a short description [5 points]
2. What data sources are available? Could you find multiple data sources? How are you going to collect them? How many data samples are you going to collect? [5 points]
3. What kind of analyses or visualizations do you want to do? [5 points]

Rubric:

1. The project name should:
   1. Reads like a research project. 1points.
   2. Have a short description. 2 points.
   3. Indicates the data source. 1point.
   4. Indicates the analysis method. 1 point.

2. The data sources should:

1. Specify the general nature of data, e.g. weather, hotel price, airfare, etc. 1 point.
2. Specify exact data sources, e.g. a link to the dataset, a link to an API, or a link to a website to parse. 1 point.
3. Have 2 or more data sources. 1point.
4. Includes a short description and your plan. 1 point.
5. The data source should be large and recent. It should contain at least 100 samples (recommends 1000 samples or more), and is no older than 10 years from now. 1 point.

3. The analyses and visualizations should:

1. Feasible.Itshouldbeananalysisthatcouldbedonewithyourdatasource.2 points.
2. Meaningful. It should be something that is not obvious, non-trivial, not in common sense. 2 points.
3. Impactful. It should be something having an impact and is interesting to the audience. 1 point.