Math23C Spring 2018 Final Project

Patrick Watts, David Wihl

May 6, 2018

Abstract

Cryptocurrencies such as Bitcoin, Ethereum, etc. generated significant attention in 2017. Cryptocurrencies have significant volalility as there is rampant speculation. Given the high variance in prices, can data science methods explored in this class be used to model the market dynamics?

Data Source

We obtained a history cryptocurrency price data from https://www.kaggle.com/jessevent/all-crypto-currencies

This 65 mb file contains open / high / low / close (OHLC) data for 13 different currencies from 2013 through Feb 22, 2018 for a total of 702,166 observations.

(REQ: a dataframe, at least two numeric columns, at least 20 rows)

Exploratory Data Analysis

```
#df = read.csv("data/crypto.csv"); head(df)
```

Topic 1 - Volalitility Exploration

Topic 2 -

Summary of Project Requirements

Required dataset standards

- [x] A dataframe
- [ ] At least two categorical or logical columns
- [x] At least two numeric columns
- [x] At least 20 rows, preferably more, but real-world data may be limited

Required graphical displays (all graphs must be colored and nicely labeled)

- [ ] A barplot
- [ ] A histogram
- [ ] A probability density graph overlaid on a histogram
- [ ] A contingency table

Required Analysis

- [ ] A permutation test
- [ ] A p-value or other statistic based on a distribution function
- [ ] Analysis of a contingency table
- [ ] Comparison of analysis by classical methods (chi-square, CLT) and simulation methods

Required submission uploads

- [ ] A .csv with the dataset
- [ ] A long, well-commented script that loads the dataset, explores it, and does all the analysis.
- [ ] A shorter .Rmd with compiled .pdf or .html file that presents highlights in ten minutes.
- [ ] A one-page handout that explains the dataset and summarizes the analysis.

Additional points for creativity or complexity (up to 10 points)

References