Data Codebook- GROUP 3

**Source credentials / data owner:**

1. S&P 500 data – <https://finance.yahoo.com/quote/SPY?p=SPY>

We choose a dataset of S&P 500 index that call “SPDR S&P 500 ETF” (SPY).

SPY is the first, and most popular, ETF in the U.S. It tracks one of the most popular indexes in the world, the S&P 500 Index.

We choose this dataset of ETF and not the Index itself because we want data that is tradeable.

2. VIX data - <https://finance.yahoo.com/quote/TVIX?p=TVIX>

We choose the dataset of VIX index that call “VelocityShares Daily 2x VIX ST ETN” (TVIX).

The VIX index which reflects implied volatility of the S&P 500 Index at various points along the volatility forward curve. The ETNs are linked to a multiple (2x) of the daily return of the index and do not represent an investment in the VIX.

We choose this dataset of ETN and not for Index because that we want data about something that is tradeable.

3. Bitcoin data - <https://www.kaggle.com/johnnyjai/bitcoin/data>

**Description of the data and its purpose**

**1. S&P 500 dataset includes a data frame with 6 variables:**

**Date** - date variable that describes the trading date. (format- 25/11/2015)

**Open** - numeric variable that describes the price of the ETF at market open. (USD)

**High** - numeric variable that describes the highest price reached in this trading date. (USD)

**Low** - numeric variable that describes the lowest price reached in this trading date. (USD)

**Close** - numeric variable that describes the close price at this trading date. (USD)

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for the ETF.

**2. VIX dataset includes a data frame with 12 variables:**

**Date** - date variable that describes the trading date. (format- 25/11/2015)

**Open** - numeric variable that describes the price of the ETF at market open. (USD)

**High** - numeric variable that describes the highest price reached in this trading date. (USD)

**Low** - numeric variable that describes the lowest price reached in this trading date. (USD)

**Close** - numeric variable that describes the close price at this trading date. (USD)

**Change** - numeric variable that describe the percent of change in the price of the ETN relative to the previous day price.

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for the NTF.

**3. Bitcoin dataset includes a data frame with 7 variables:**

**Date** - date variable that describes the Trading date. format- Oct 19, 2017.

**Open** - numeric variable that describes the opening price at this date in a number of $.

**High** - numeric variable that describes the high price at this date in a number of $.

**Low** - numeric variable that describes the low price at this date in a number of $.

**Close** - numeric variable that describes the close price at this date in a number of $.

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for Bitcoin.

**Market Cup** - Total equity market value of all Bitcoin.**Description of the data and its purpose**

**1. S&P 500 dataset includes a data frame with 6 variables:**

**Date** - date variable that describes the trading date. (format- 25/11/2015)

**Open** - numeric variable that describes the price of the ETF at market open. (USD)

**High** - numeric variable that describes the highest price reached in this trading date. (USD)

**Low** - numeric variable that describes the lowest price reached in this trading date. (USD)

**Close** - numeric variable that describes the close price at this trading date. (USD)

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for the ETF.

**2. VIX dataset include a data frame with 12 variables:**

**Date** - date variable that describes the trading date. (format- 25/11/2015)

**Open** - numeric variable that describes the price of the ETF at market open. (USD)

**High** - numeric variable that describes the highest price reached in this trading date. (USD)

**Low** - numeric variable that describes the lowest price reached in this trading date. (USD)

**Close** - numeric variable that describes the close price at this trading date. (USD)

**Change** - numeric variable that describe the percent of change in the price of the ETN relative to the previous day price.

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for the NTF.

**3. Bitcoin dataset include a data frame with 7 variables:**

**Date** - date variable that describes the Trading date. format- Oct 19, 2017.

**Open** - numeric variable that describes the opening price at this date in a number of $.

**High** - numeric variable that describes the high price at this date in a number of $.

**Low** - numeric variable that describes the low price at this date in a number of $.

**Close** - numeric variable that describes the close price at this date in a number of $.

**Volume** - numeric variable that describes the number of total shares that are traded on a given day for Bitcoin.

**Market Cup** - Total equity market value of all Bitcoin.