

ASSIGNMENT #2: Flash Crash in 2010
(Deadline: Sunday, 5 Dec 2021, 23:00)

Your boss asked you to investigate what happened during the Flash crash and think come up with recommendations and lessons to be learned for asset managers with regards to a wide range of issues from constructing profitable strategies to thinking about risks.

The data are ZIP files containing compressed CSV files of tick data for E-mini futures and ETFs. Questions are meant to be open-ended. Good answers might vary a great deal from student to student.

Download the ZIP file and clean the data.

Unzip the ZIP file, which contains four CSV files of quotes and trades for the SPDR ETF and the S&P 500 E-mini futures contract at millisecond resolution.

- Excludes bad observations with price or size of 0.
- Excludes "crossed market" and "locked market" quotes.
- Address the time zone issues. Note that the timestamps of the SPY data are New York time, but the timestamps of E-mini data are UTC time.
- Excludes observations out of trading hours, which are 9:30–16:00 in New York time.
- Examines duplicates in each data set.
- Deals with NAs in the futures data (because one row of futures data only reports the update of bid or ask).
- Summarize in the consolidated table the information on filtered out observations for each type of errors (missing, NAs, duplicate, wrong etc) for each dataset field.

List several useful observations that you can make based on the data cleaning part of the project.

Code	Description
B	NASDAQ OMX BX, Inc.
C	National Stock Exchange Inc. (NSX)
D	Financial Industry Regulatory Authority, Inc. (FINRA ADF)
I	International Securities Exchange, LLC (ISE)
M	Chicago Stock Exchange, Inc. (CHX)
P	NYSE Arca, Inc.
T	NASDAQ Stock Exchange, LLC (in Tape A, B securities)
W	Chicago Broad Options Exchange, Inc.(CBOE)
Z	Bats BZX Exchange, Inc.

Table: Exchange codes

1. Price series analysis

Normalize the futures and spot prices to 100 by dividing each price by the first price of the day, then multiplying by 100. Plot the normalized futures and SPDR prices on the same plot for the entire day for different exchanges (see codes in Table). Does any of the time series differ sharply from the rest? Can you figure out why? Exclude this series from your analysis.

2. Analysis of Volume and Liquidity.

Compare trading volume, number of trades, number of quotes, and width of bid-ask spread in the E-mini futures with SPY ETF ("cash market"). Develop a "theory" to explain why one might have more trades or quotes than the other?

3. Flash Crash.

Plot the future and SPDR prices for the time period of the flash crash. Did the arbitrage relationship between the E-mini futures contract and the SPDR ETF break down during the flash crash? If so, did the breakdown occur at the beginning, middle, or end of the flash crash?

Plot average trade size for E-mini and SPDR during the day and the flash crash. Did the average trade size change during the Flash Crash? Could you suggest any explanations?

4. Analysis of Volatility.

Plot realized volatility of the E-mini and SPY cumulated over various time periods and for various exchanges.

If typical volatility is one percent per day, how much greater than typical was volatility on the day of the flash crash? Why does realized volatility vary depending on the exchange used to construct the graphs of cumulative volatility in the sample program?

5. Do price changes on some exchanges lead price changes on others?

Examine the correlation of price changes for E-mini and SPDR for different horizons. Do you observe any patterns? How can you explain them?

6. Think about recommendations and insights for asset managers.

After you finished your analysis, think about how to present your results and main ideas to your colleagues.

