

Market Data Assignment

MFIT 841: Digital Capital Markets

Overview

This is going to be hard. The data are large and confusing. Your job is to find a way to distill the important information into a simple and straightforward analysis. Good luck!

Direction

Students will analyze two datasets that include quotes and trades. Your job will be to:

- 1) Describe the quotes dataset (use averages, sums, distributions, ...)
- 2) Describe the trades dataset (use averages, sums, distributions, ...)
- 3) Merge the quotes and trades datasets by symbol and datetime
- 4) Calculate bid-ask spreads, effective spreads, trading volume, minute-to-minute returns, volatility of minute-to-minute returns and any other variable that you think is important to describe market quality.

Answer the following questions:

- Are these stocks liquid or illiquid?
- Are all stocks the same?
- Why is the effective spread different than the bid-ask spread?
- Are these deep markets?
- Are spreads and depth correlated?
- Can you “explain” price movements using only this data?
- Can you predict future price movements using only this data?

You are going to have to dig deep, Google a lot and figure out how to analyse this data. I recommend using R or Python. Here are some resources that I use:

- 1) Install and initial steps to use a data science package:
 - a. R and R Studio: <https://towardsdatascience.com/how-to-install-r-and-rstudio-584eeefb1a41>
 - b. <https://www.codecademy.com/articles/install-python-data-analysis>
- 2) Import with:
 - a. R - <https://www.datacamp.com/community/tutorials/r-data-import-tutorial>
 - b. Python - <https://towardsdatascience.com/how-to-read-csv-file-using-pandas-ab1f5e7e7b58>
- 3) Lots of initial code in both Python and R by searching here: <https://stackoverflow.com>
- 4) Manual for the quotes and trades dataset are included. You do not have all of the data, I only included the most relevant fields.

The report should be **no longer** than 5 pages double spaced with 12-point font. This includes tables, figures, references, and links. Writing for financial professionals requires very compact writing, if it isn't necessary to understand it shouldn't be there. No repetition!

Describe the data <ul style="list-style-type: none"> - How many symbols - Averages, sums, distributions - Tell me what the data is telling you 	5
Analysis <ul style="list-style-type: none"> - Calculate spreads, volatility, depth, ... - Are these variables correlated - Are the markets liquid - Are all symbols comparable? - Can you predict or explain the dynamics of the data? - Tell me what you have and can learn from the data 	10
Professional Quality of Case write-up <ul style="list-style-type: none"> - Double spaced, 12 pt. font and 5 pages maximum - Free of grammar and spelling errors - Clear and concise presentation of ideas and analysis 	5
TOTAL	20