Market Data Assignment

MFIN 841: Financial Technology & Innovation

Overview

This is going to be fun. We have the number of accounts that hold a position in the symbol in the stock referenced in the file name (AAL.csv – AAL = American Airlines). We have observation multiple times per day from May 2018 up until August 13th, 2020. This of course includes the corona virus market meltdown, lockdown and recovery periods. This is a very exciting and fun dataset as it gives you a view into what a large group of retail investors are doing in quasi real-time.

Direction

Students will analyze ten datasets that include the number of Robinhood clients that hold a position in the stock mentioned. Your job will be to:

1) Retrieve information on the symbol (Name of company, listing exchange, sector, …)
2) Collect daily prices and trading volumes for each symbol from (Yahoo Finance, or any other free data service)
3) Calculate estimated returns for Robinhood clients (Do they usually buy before prices go up, or sell before they go down?)

Answer the following questions:

- Are Robinhood clients making money?
- Can you predict what Robinhood clients are going to buy or sell?
- Do you think that trading volume is affected by Robinhood client buys and sells?
- Do you think that making these information public is good for Robinhood clients?

Bonus:

I have included the full data set that includes all Robinhood client positions.

1) Do Robinhood investors hold diversified portfolios?

Tools:

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

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 - Show me the time series of buy and sells - Volumes	5
Analysis - Calculate returns - Are buys and sells correlated across stocks? - Is there a time trend in all stocks? - 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 - 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