Ruida Xie

Big Data in Finance

Assignment 2

Synthetic Data and VWAP Computation

Using Q, I generated 10 days of trading data of 20 stocks. The trading frequency is adjusted in such a way that the entire data set is about 1 GB in size. The data is generated and stored as a dictionary, in which the keys are the stock tickers and the values are tables of data. However, the trading data of each stock is saved as .csv files. The prices of each stock float moderately, following a geometric Brownian motion, and likewise the volume are moderately large, following a uniform distribution capped by 5,000 shares, so that these two pieces of information are consistent with each other.

I choose to compute daily VWAP for simplicity. By reading the trading data line by line, the program updates the daily VWAP as it goes along. The program also checks if the date time is the end of a day. If it is end of a day, it stores the daily VWAP and continues to compute the next day’s VWAP.

Note: There will be more detailed explanations for the code in the code files.