UNIST
School of Business Administration

FIN552 : High Frequency Financial Data Analysis — Daejin Kim, Spring 2016

# Assignment 2
Due Date : **Thursday, April 28, 2017**

1. [**Preparing the TAQ Data**] From our class, we already learned basic properties of downloading TAQ data from the WRDS. Although the SAS provides a powerful tool for analyzing the data, it might be difficult to understand the data thoroughly. In this exercise, we'll be analyzing it offline, i.e., not on SAS platform but on Microsoft Excel for better understandings.

   (a) The excel file **sp500.xlsx** contains a list of stocks consisting of S&P 500 index on **June 6, 2013**. The excel file contains cusip information and listed stock exchange (N: NYSE, Q: NASDAQ). Choose two stocks randomly (One stock from NYSE(N) and the other from NASDAQ(Q)). What is the company's full name? What was the prior day's closing price (that is, closing price on June 5, 2013)? (You can get this from Yahoo Finance, among many other places.)

   (b) Using the information in TAQ master file (i.e., mastm_20130606), download trade file, quote file, and nbbo file for selected two stocks on the date. Note that the cusip information in the excel file and that in the master file might have different digits (Use substr() function in SAS). Filter out your data based on standard criteria from the class. Convert your data into one excel file (Use the PROC EXPORT. Same file with three different sheets).

2. [**Trade Data Analysis**] Based on your excel file, do the following analysis.

   (a) What is the largest single trade of the day? When was it executed? at what price?

   (b) How much volume(in shares) was done on each exchange? Which exchange has the largest volume? What is the market share of each reporting exchange?

   (c) Between 9:00 and 16:00, plot the trades. That is, plot price versus time.

   (d) Construct 30-minutes time buckets, i.e. 9:30-10:00, 10:01-10:30, $\cdots$. In each interval, compute the total number of trades. Again, compute the market share of each reporting exchange within time interval.

3. [**Quote Data Analysis**] Based on your excel file, do the following analysis.

   (a) During the day, how much the number of quote updates was done one each exchange? Which exchange has frequently updated the quotes? For what proportion of the quotes is the primary listing exchange at the best bid? (Note that the best bid or the best ask do not have to be alone. For example, if the primary exchange is the NYSE, the NYSE is at the best bid if the exchange of the best bis is marked as "N", "XN", or "BNZ".)

   (b) In each of the 30-minute time buckets, do the same thing for the above question.

4. [**NBBO data anaysis**]

   (a) Do the same analysis in question 3. Are there any differences from results in question 3.

   (b) In each interval, compute the average of the quoted bid-ask spread and plot it.