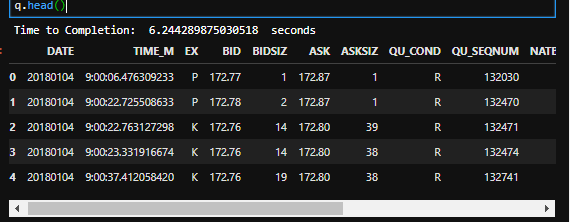
Data

Raw data TAQ is taken from WRDS – describe this part from the raw dta

Data Wrangling

One of the most challenging aspects of this project is taking the raw data from TAQ, recreating the NBBO, and then capturing instances of creates and joins. In this section we walk through how we took the raw data, and created pandas dataframes containing the information we needed.

For the purpose of our walkthrough, we first use AAPL (find day) as an example



Each line in this db represent a single quote, per exchange. Before we can start using these quotes to construct a timeseries of the NBBO, we first need Best Bid/Offer eligible tables. We do this by filtering on the “QU\_COND” for codes for BBO ellfibale quotes

One important fact about this new\ table is that each quote is considered a Level 1 quote – meaning that It is each an adjustment to the Best Bid or the Best Ask on the corresponding exchange. There are two ways to make an adjustment the Best Bid (BB) or the Best Offer (BO), the first is by adjusting the price itself, and the second is by adjusting the Volume at the BB or BO.

Using tjhis information – we are able to loop through each quote, keep track of which exchanges has the highest best bid and the lowest best offer