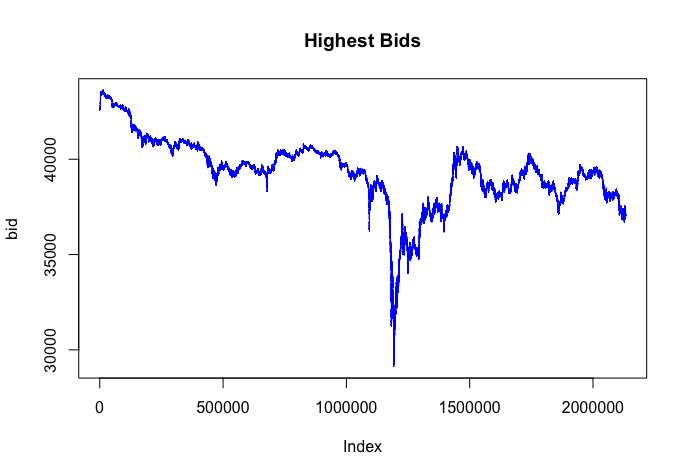
**Technical Task**

**Rishik Kumar**

**1 - Understanding the data**

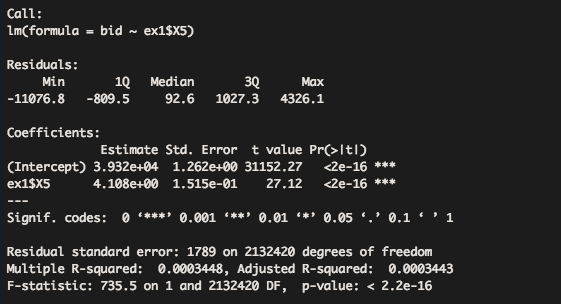
To first understand the data, I will simply plot a time series of the highest bid price to visualise the changes and patterns.



From the plot, we can see straight away that the highest bid price changes erratically which we can model using properties of Brownian motion.

Thinking of a strategy involving the order size I will first perform a regression analysis on the highest bid price using the order volume.

From R, I achieved the following output:



This output is showing that the order size variable (ex1$X5) is highly significant with a p-value (=) 0.

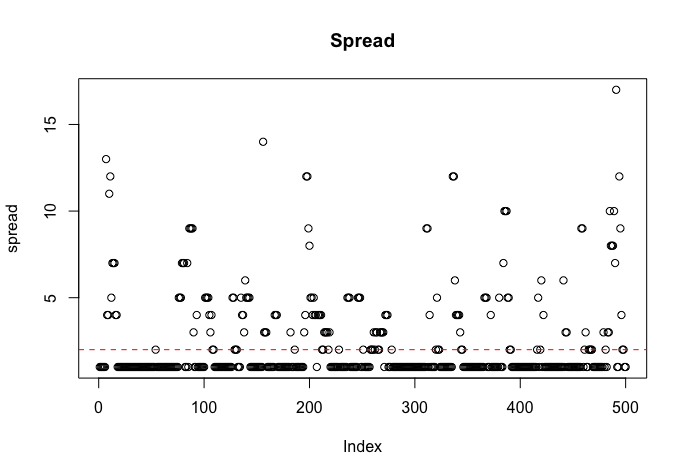
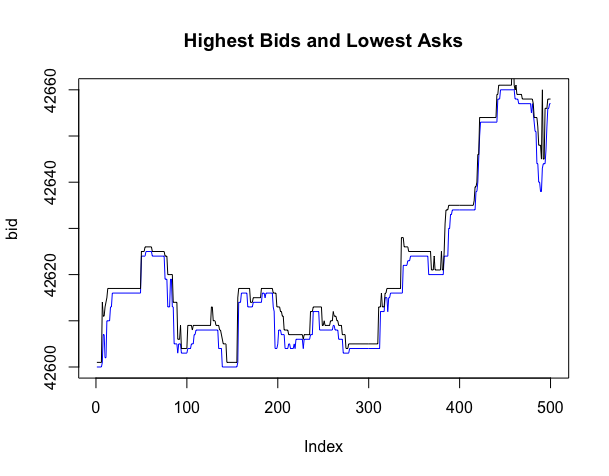
To write the full model:

Highest bid = 39320 + 4.108(order size)

The benefit of fitting this regression is that it confirms that order size (being highly significant) could potentially be used in a trading strategy. However, the most obvious limitation is that we do not know the price information at time t+1.

In mathematical language, we are not sigma-algebra Ft+1 – measurable.

Investigating further, I decided to look at the bid-ask to spread of the data.

For ease of visualisation, I will only plot the first 1000 values.

From the plot, we can see the spread is typically 1 unit, however we notice there any times where the spread increases, this will be the basis of the trading strategy I will use.

**2 – Overview of the trading strategy**

The main idea behind my trading strategy is capitalising on the variation in the spread.

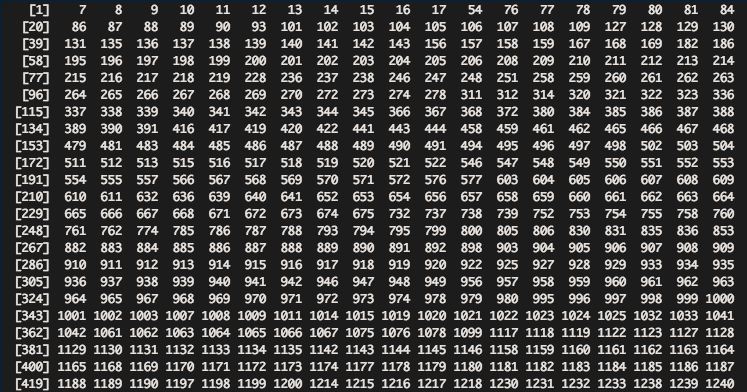
Essentially, without considering any limitations, when the spread increases we buy and then sell the same financial product almost instantaneously capitalising on the higher selling price.

Signals:

Having analysed the data, the signals I have chosen to use are when the spread increases past 1 basis point, therefore portraying a profitable opportunity when including the trading fee (mentioned in the question).

Using R, I was able to find which times have a spread greater than 0.01%.

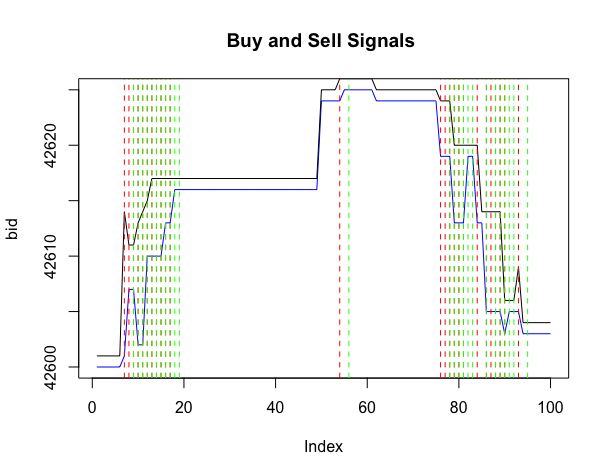
This is an example output:

The numbers in the output correspond to the elements of the time vector in which the spread is greater than 0.01% of the value.

Then using these elements, I can create the buy and sell signals.

To incorporate the trading latency of 100ms, the buy trigger will be 2 time elements after the buy signal and sell signal will be 4 time elements respectively.

Plotting below, for ease of visualisation I will only use 100 data points however the following will apply to all.

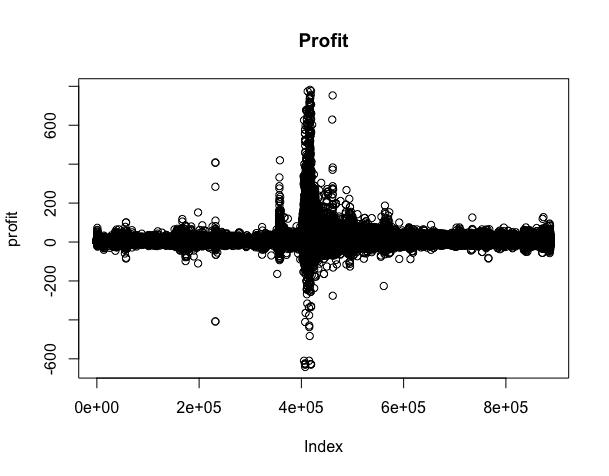


The red and green lines are the buy and sell signals respectively.

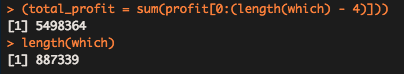
**3 – Results from day 1**

Now calculating the profit from the strategy on the first day:

As mentioned we are buying the product at two time elements after the buy signal and selling at four time elements after.



From the plot, we can see some values are negative. The strategy does make a loss when the ask price is lower than the highest bid price 4 time elements earlier, however, as we will see the strategy does make profit over the long haul by executing thousands of trades.



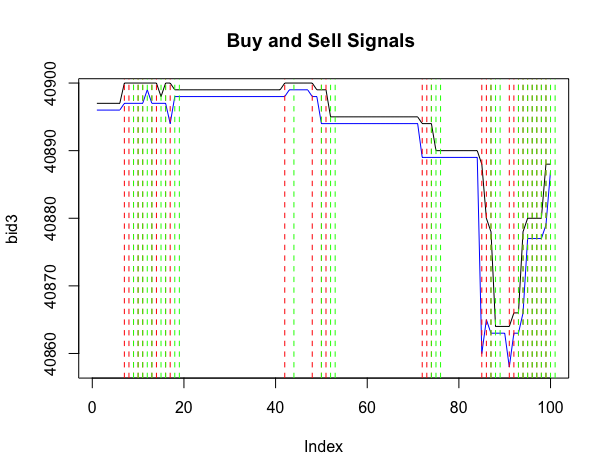
Total profit from day 1 = 5498364

Number of trades from day 1 = 887339

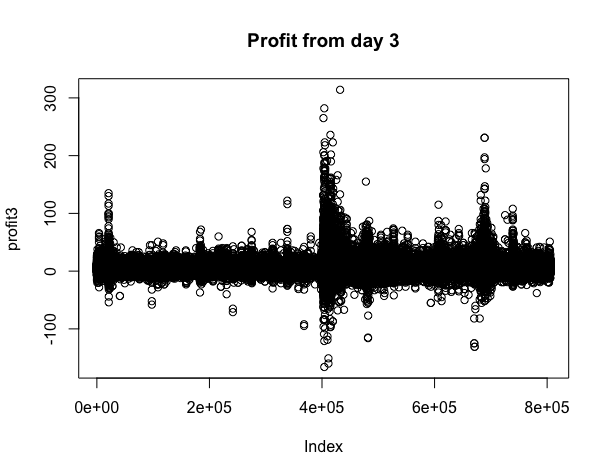
**4 – Testing on day 3**

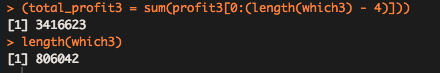
Using the same strategy on day 3 to test if it works on extrapolated data.

Plotted below are the buy and sell signals and for ease of visualisation I will only us the first 100 data points.



Then using the same function as before to calculate profit (see plot below).





Total profit from day 3 = 3416623

Number of trades from day 3 = 806042

1. **– Additional questions**

How is your approach affected by a change in position size, exchange fee, or latency?

The strategy outlined above does not take into account position size however this would affect the strategy. In order for the strategy to remain profitable the ask quantity would have to be sufficient (large enough) in order to sell position.

For my strategy, specifically the ask quantity at time t would have to be greater than bid quantity at time t – 4.

Regarding the trading fee, I have taken account of this in my code as 0.01%. However, if the trading fee gets larger, then the strategy will become less profitable. The reason for this is because the elements of which the spread is greater than 0.01% (or the new trading fee) will decrease, therefore resulting in less trades and less profit.

Regarding latency, I have factored this in by increasing the time between the buy signal and buy trigger. If latency increases further, then there is a greater risk that the strategy will make a loss.