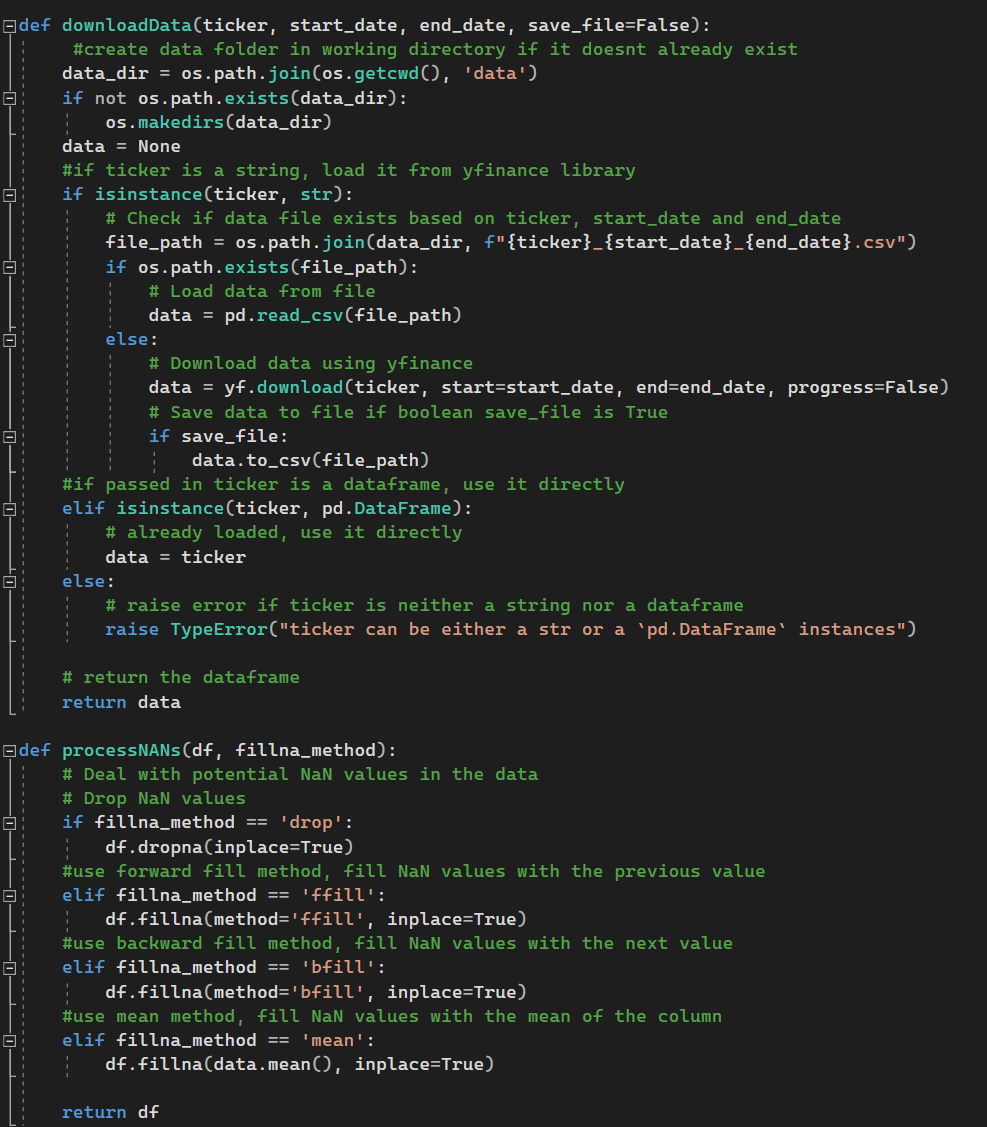
Task 3 Report

Naca Hitchman - 103072170

# Utility Functions

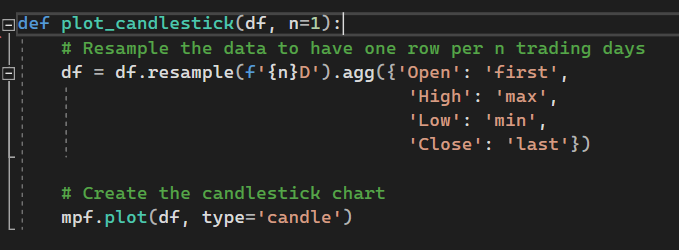
Firstly, as these functions will be plotting original downloaded finance data-frames, and not predicted data like the previous task, some code needs to be moved from the processData function from task 2 into their own separate functions. Specifically, the finance data downloading and file saving, as well as the data NANs processing.

By moving both into their own separate functions, we can re-use them for this task and still also call them in the original processData function. We can see below the data download code was moved into a new downloadData function, where we pass the ticker, dates, and whether or save file Boolean. Further, the NAN processing code was moved into a processNANs function, still retaining the ability to specify if NANs are to be dropped or filled. Both of these functions return dataframes so they can be used in other areas.



# Candlestick Chart Function

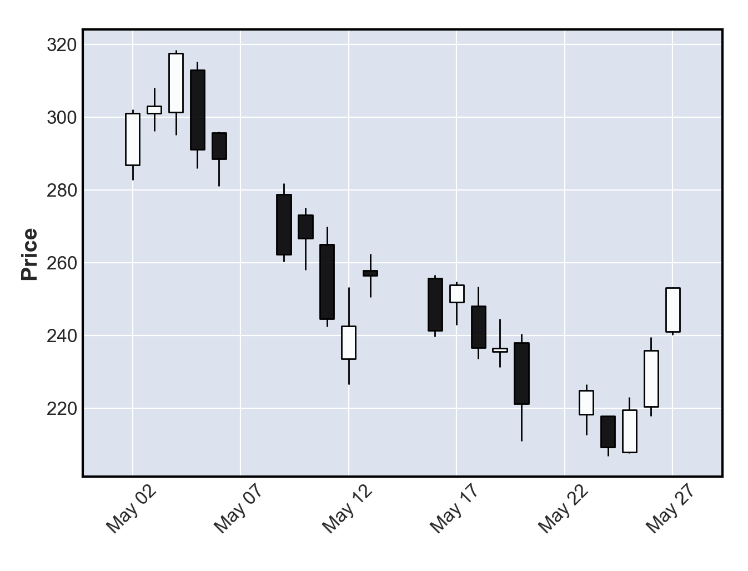
Creating the candlestick chart was extremely straightforward. The task specifies we need to be able to pass n as the number of trading days represented by each stick, so for the parameters we pass both a data frame, and an n integer with a default value of 1 if the user chooses not to pass a value.



The resample method is then run on the dataframe to adjust it to fit the n trading days parameter. “f’{n}D’” specifies to interpres n as days, as agg helps us ensure the various columns are aggegated correctly according to their candlechart usage.vThen, the mpf plot method can be used, passing the ‘candle’ type to show a candlestick chart using the data frame.

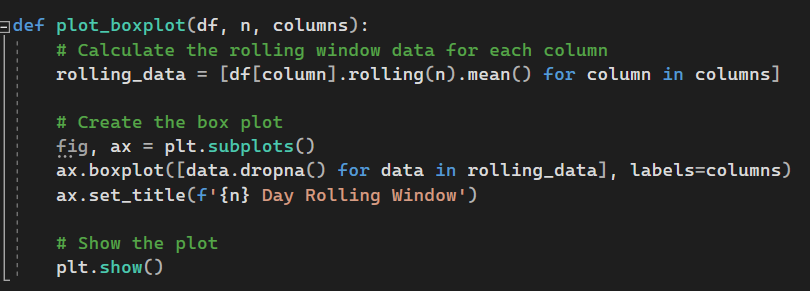
A graph with black and white rectangles

Description automatically generatedn = 1 : n = 5 :



# Boxplot Chart Function

Creating the boxplot was also relatively straight forward. The specifications asked for being able to pass n as the size of a rolling window average, so our parameters include a dataframe, n for the window size, and the list of columns to be displayed.



Firstly, the rolling data is calculated for the passed data-frame. Each column of the data is loop through, and the built in rolling() method is used, passing n, to get the data we need. The pyplot subplots() method is used to give us an ax object, which then allows us to generate a boxplot, passing a modified rolling\_data list with the nans dropped. The title of the plot is then set, before slowing the plot using pyplot again.

n = 20 : n = 40 :

A graph of a window

Description automatically generated with medium confidenceA graph of a window

Description automatically generated with medium confidence

I found that since the volume column values are significantly larger than the other columns, they can’t be displayed without the other columns appearing invisible due to how much smaller they are. This is why I had to include being able to pass which columns to display, so that the other columns could be visible by not passing volume.