COS30018 – Option B – Task 3: Data processing 2

Jiin Wen Tan 102846565

Resources:

https://coderzcolumn.com/tutorials/data-science/candlestick-chart-in-python-mplfinance-plotly-bokeh

https://www.geeksforgeeks.org/box-plot-in-python-using-matplotlib/

https://www.geeksforgeeks.org/how-to-make-a-time-series-plot-with-rolling-average-in-python/

For this task we need to implement a few visualization techniques based on analyzing the candlestick chart representation of the stock price in project (P2) in this case we need to write two function each to display the stock market financial data using candlestick chart and boxplot chart.

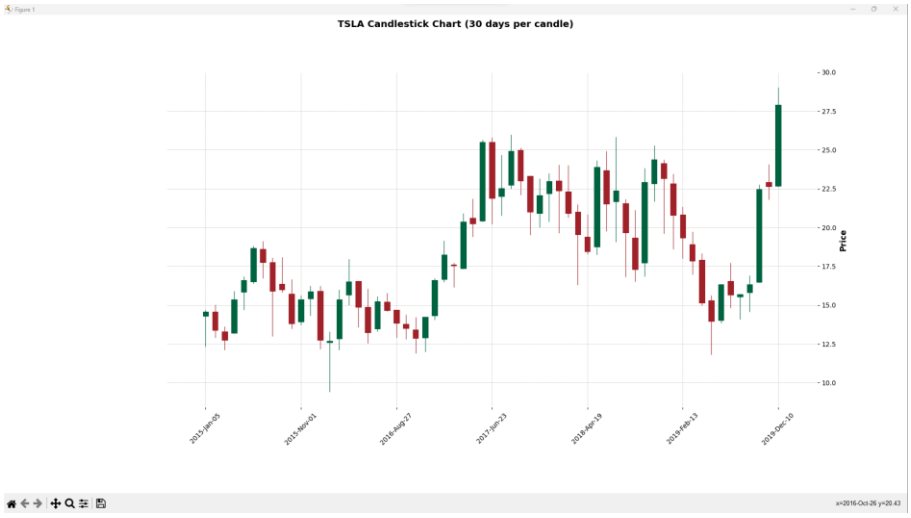
Candlestick Chart

```
def plot_candlestick(data, n=30):

    data_resampled = data.resample(f'{n}D').agg({
        'Open': 'first',
        'High': 'max',
        'Low': 'min',
        'Close': 'last'
    })

    fplt.plot(data_resampled, type='candle', style='charles', title=f'{COMPANY} Candlestick Chart ({n} days per candle)')
```

The above code is used to visualize tesla stock using candlestick charts. This function takes in two parameters data (financial data) and n (trading days for data resampling) then resamples data in this case into 30 days intervals. This process aggregates data within each interval and computes open, high, low and close. Finally, displaying the candlestick chart using mplfinance library.



Boxplot Chart

```
def plot_boxplot(data, column_name, n=30):
    chunks = [data[i:i+n] for i in range(0, len(data), n)]

    box_data = [chunk[column_name].values for chunk in chunks]

    dates = [chunk.index[0].strftime('%Y-%m-%d') for chunk in chunks]

    plt.figure(figsize=(15, 7))
    plt.boxplot(box_data)

    plt.xticks(ticks=np.arange(1, len(dates)+1), labels=dates, rotation=90)

    plt.title(f'Boxplot of {column_name} (every {n} days)')
    plt.xlabel('Date')
    plt.ylabel(column_name)
    plt.tight_layout()
    plt.show()
```

The code above is used to visualize tesla stock using boxplot charts. The function first divides the data into chunks and then extracts the value for the desired column from each chunk this is achieved through list comprehension by iterating over the data in steps of n. Then, for each chunk the values of the column are extracted to be used to plot the boxplots. For labeling purposes and visibility, I used the date of the first entry in each chunk. The ticks are set to the extracted dates and are rotated for visiblity.