

Question 1 (a)

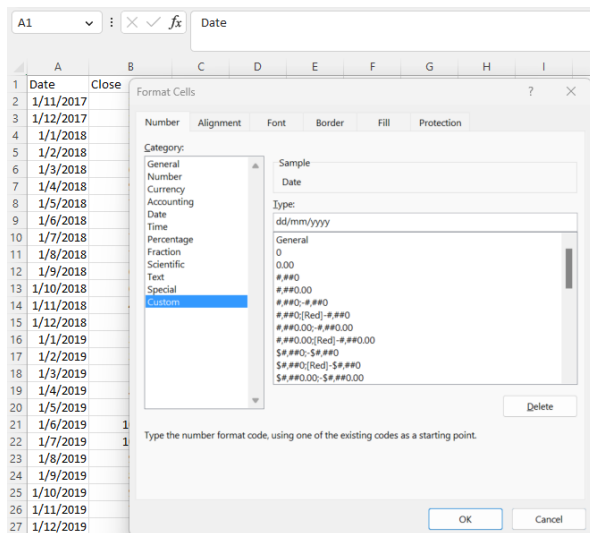
The dataset is observed to be from Nov 2017 to Oct 2022 with monthly prices and volume that were captured in chronological order. Hence, we can forecast the future values using time series method. To forecast the price in the next 12 months, we will take a subset from the original dataset and select only Date and Close variables for forecasting. Close price is the latest price that determines the value of the stock until the market resumes trading and presents the market sentiment over time. Adjusted Close price is not selected because dividend pay-outs and stock splits does not affect the monthly stock prices of company B. Figure 1 shows the dataset after removing the other variables.

Figure 1 Part of the StockPriceB.csv after variables removed.

Date	Close
1/11/2017	10233.6
1/12/2017	14156.4
1/1/2018	10221.1
1/2/2018	10397.9
1/3/2018	6973.53
1/4/2018	9240.55
1/5/2018	7494.17
1/6/2018	6404
1/7/2018	7780.44
1/8/2018	7037.58
1/9/2018	6625.56
1/10/2018	6317.61
1/11/2018	4017.27
1/12/2018	3742.7
1/1/2019	3457.79
1/2/2019	3854.79
1/3/2019	4105.4
1/4/2019	5350.73
1/5/2019	8574.5

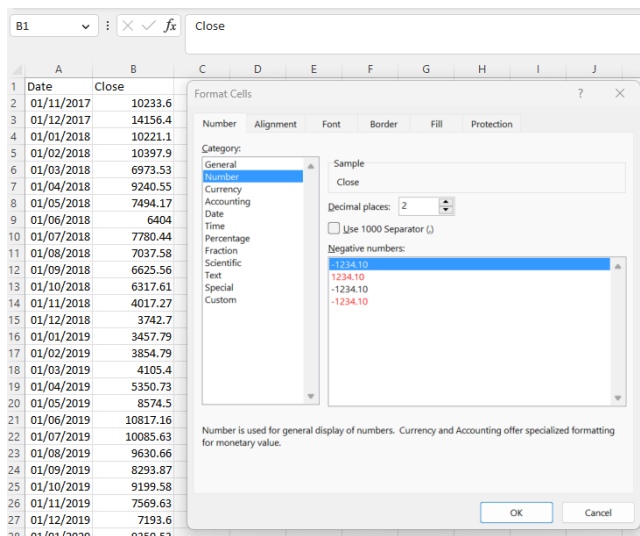
Before importing into SAS environment, we will be required to define the format of the timestamp clearly in the dataset to base on the raw data. Since Date variable resembles the format of DD/MM/YYYY, we will select the whole Column A and format the cells with the type as shown in **Figure 2**.

Figure 2 Setting the timestamp for Date variable



After the Date variable timestamp format is set, we will check through the Date variable to ensure that the time series dataset has equal interval (i.e. month) and with no missing data. There are altogether 60 records in the dataset starting from Nov-2017 to Oct-2022. For the Close variable, we will also define the format correctly as Number type with 2 decimal places as shown in **Figure 3**.

Figure 3 Setting the correct format for Close variable



To import the Time Series dataset into SAS, we must define the format with the below code so that the Date variable can be used for forecasting.

Code: DATA ANL317G.StockQ12;

INFILE "C:/Users/SASUser25/Documents/GBA/StockPriceB_RawT2.csv" DLM=","

FIRSTOBS=2;

INPUT Date DDMMYY10. Close;

FORMAT Date DDMMYY10.;

RUN;

Question 1 (b)

- i) Once the time series dataset is imported to SAS, we can derive the descriptive statistics as shown in **Figure 4**. Mean is 20163.14, Median is 10591.20, and Variance is 289519095. Standard deviation is 17015 indicating the dispersion around the mean. The average price may fluctuate between 3148.14 and 37178.14 with average discrepancy of 34,030. These statistics indicate that the dataset fluctuates with a wide range as the standard deviation is 84% of the mean. We can foresee a low confidence in our forecasted value basing on this dataset.

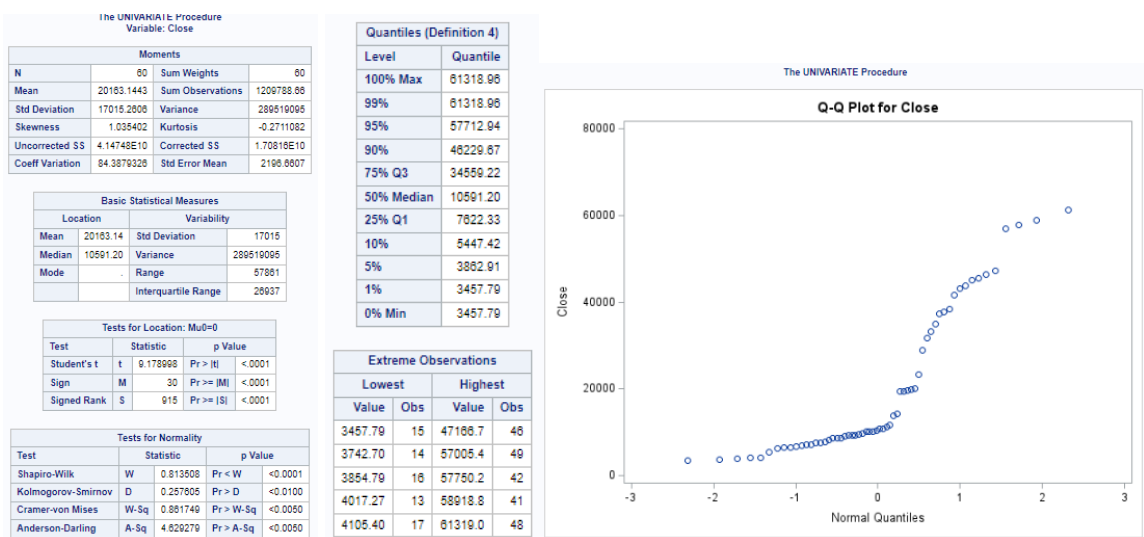
Code : PROC UNIVARIATE DATA=ANL317G.StockQ12 NORMAL PCTLDEF=4;

VAR Close;

QQPLOT Close;

RUN;

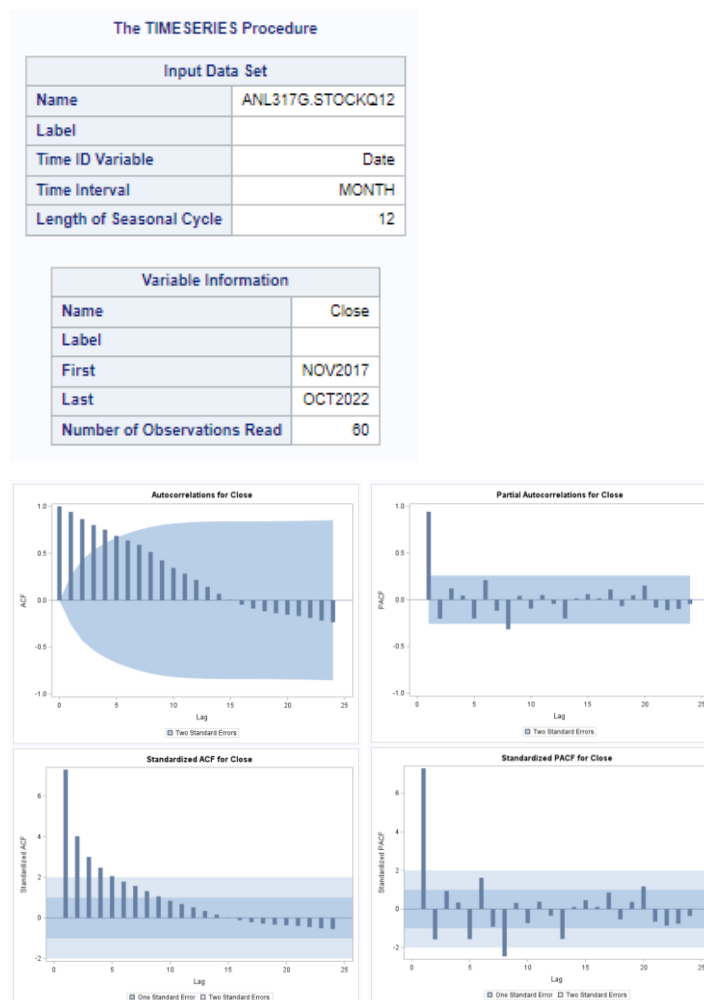
Figure 4 Descriptive Statistics of the time series.



When we apply the autocorrelation to the Close variable, we can observe in **Figure 5** that there is a uniform decay in the ACF plot and become negative correlations after lag 15. This show that the Time Series is stationary and has a trend. Although only lag 8 in the PACF plot show statistically significant, there are no correlation pattern observed between the months. Therefore, closing stock prices are random and uncorrelated with that of previous months.

```
Code : PROC TIMESERIES DATA=ANL317G.StockQ12 PLOTS=(acf pacf);
      ID Date INTERVAL=month ACCUMULATE=average;
      VAR Close;
      RUN;
```

Figure 5 Autocorrelations for Close variable.



- ii) For graphical explanation, we construct the line plot (**Figure 6**), Box Plot (**Figure 7**) and Histogram (**Figure 8**) to find the relationship of Close variable against the time series dataset.

Line Plot Code: PROC SGPLOT DATA=ANL317G.StockQ12 NOBORDER;
 SERIES X=Date Y=Close / MARKERATTRS=(SIZE=5)
 LINEATTRS=(Colour=Green);
 REG X=Date Y=Close / LINEATTRS=(THICKNESS=2)
 MARKERATTRS=(SIZE=1);
 RUN;

Box Plot Code: PROC SGPLOT DATA=ANL317G.StockQ12 NOBORDER;
 VBOX Close;
 RUN;

Histogram Code: PROC SGPLOT DATA=ANL317G.StockQ12 NOBORDER;
 HISTOGRAM Close / NBINS=20;
 DENSITY Close;
 RUN;

Figure 6 Line Plot of Closing Stock Prices

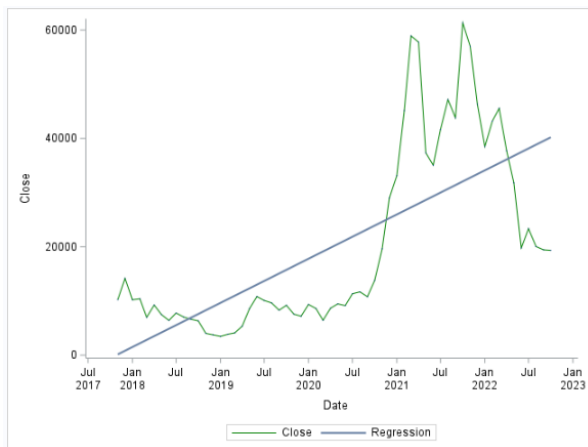


Figure 7 Box Plot of Closing Stock Prices

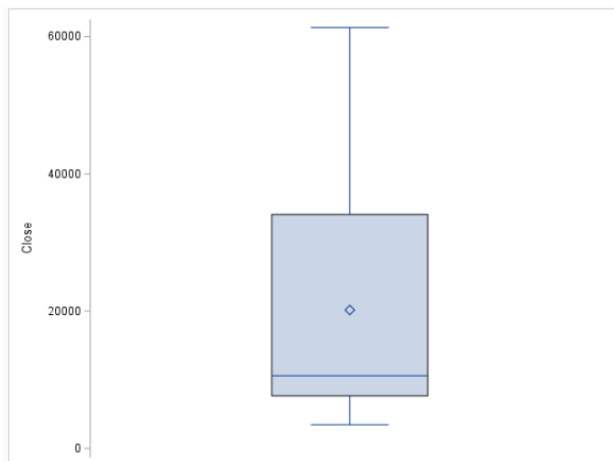


Figure 8 Histogram of Closing Stock Prices

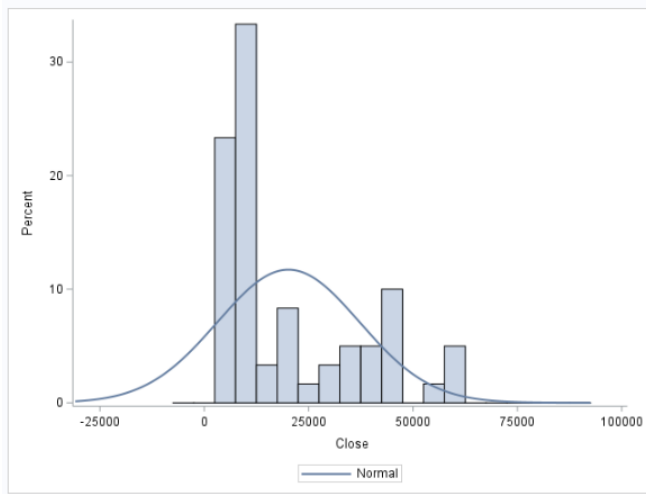


Figure 6 shows that there is a general positive upward trend in the Closing stock price with major cyclical components above the trend line between December 2020 and June 2022. Troughs were noted in Nov 2018, May 2021, and June 2022.

Figure 7 shows that Closing stock price range are from 3400 to 61000. Most observations are recorded from 8000 to 35000 (falling within 25th to 75th percentile). Mean is around 20163.14 and Median is around 10591.2. There is no outlier in the dataset.

Figure 8 shows the distribution of Closing stock price with highest frequencies of Closing stock price at around 10000 and it does not follow a normal distribution.

Figure 9 shows the Time Series decomposition plots of Closing stock price. TC plot shows the closing price fluctuated significantly over time. There is long term growth for the company as the stock price rose gradually from 2017 to 2022. The greatest price increase happened in May 2019 (60.25%) and the greatest price decrease occurred in June 2022 (-37.8%).

CC plot shows that there are 2 short term cycles between Jan 2020 to Jan 2022 which may be due to the Covid 19 circuit breaker in April 2020, easing the country controls of the pandemic from Jan 2021, Phase 2 heightened alert in July 2021 and tension in Russia grew in Jan 2022.

SC plot shows the seasonality after filtering. It is noted that the occurrence and duration occurred regularly and lasted for 12 months. It shows recurring positive spikes up to 2 and negative spikes down to 1 which repeated 5 times from Nov 2017 to Oct 2022. The magnitude is also constant.

```

Code: PROC TIMESERIES DATA=ANL317G.StockQ12 PLOTS=(decomp);
      ID Date INTERVAL=month;
      VAR Close;
RUN;

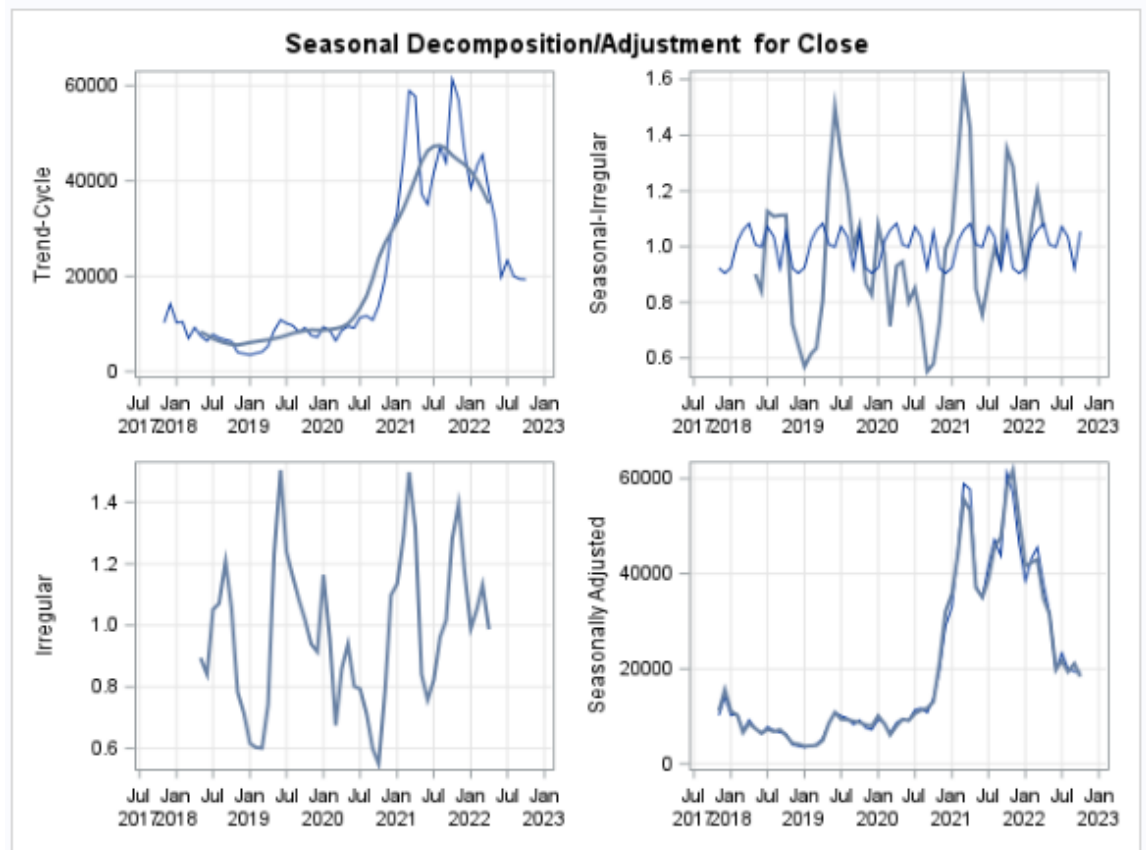
```

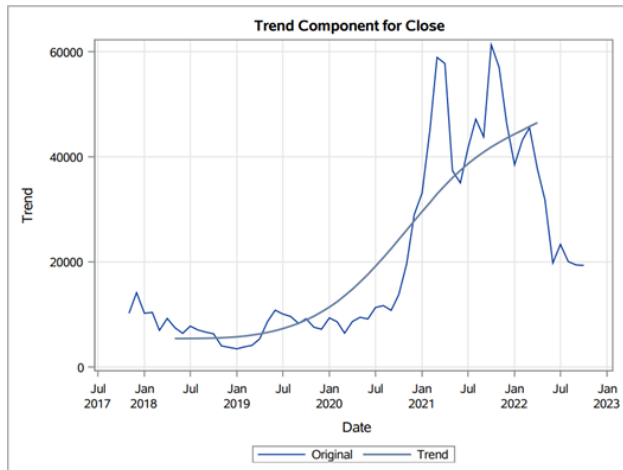
Figure 9 Time Series Decomposition Plots

The TIMESERIES Procedure

Input Data Set	
Name	ANL317G.STOCKQ12
Label	
Time ID Variable	Date
Time Interval	MONTH
Length of Seasonal Cycle	12

Variable Information	
Name	Close
Label	
First	NOV2017
Last	OCT2022
Number of Observations Read	60



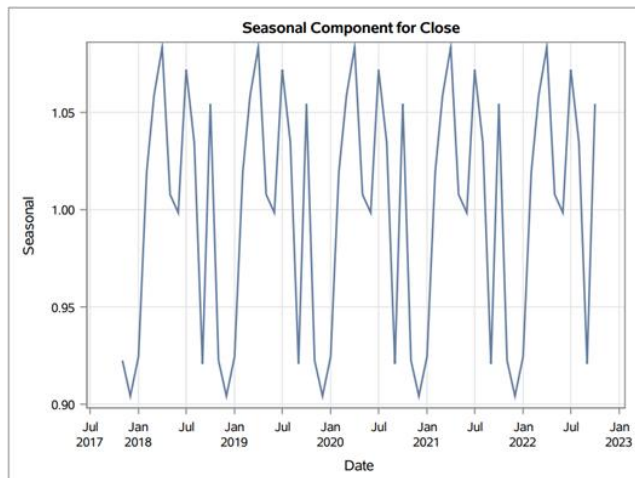


Code: PROC TIMESERIES DATA=ANL317G.StockQ12 PLOTS=(TC);

ID Date INTERVAL=month;

VAR Close;

RUN;

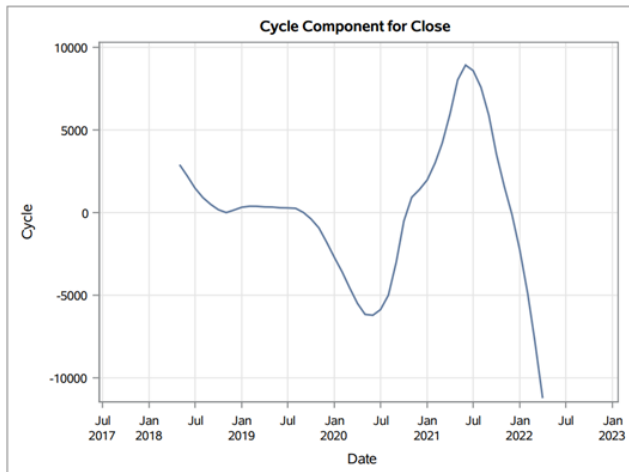


Code: PROC TIMESERIES DATA=ANL317G. StockQ12 PLOTS=(SC);

ID Date INTERVAL=month;

VAR Close;

RUN;

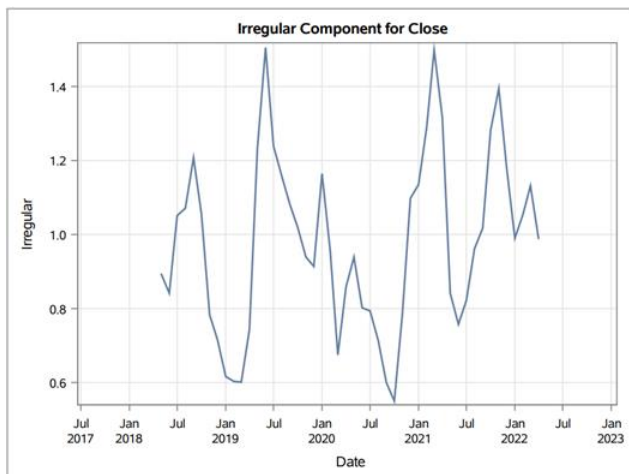


Code: PROC TIMESERIES DATA=ANL317G.StockQ12 PLOTS=(CC);

ID Date INTERVAL=month;

VAR Close;

RUN;



Code: PROC TIMESERIES DATA=ANL317G. StockQ12 PLOTS=(IC);

ID Date INTERVAL=month;

VAR Close;

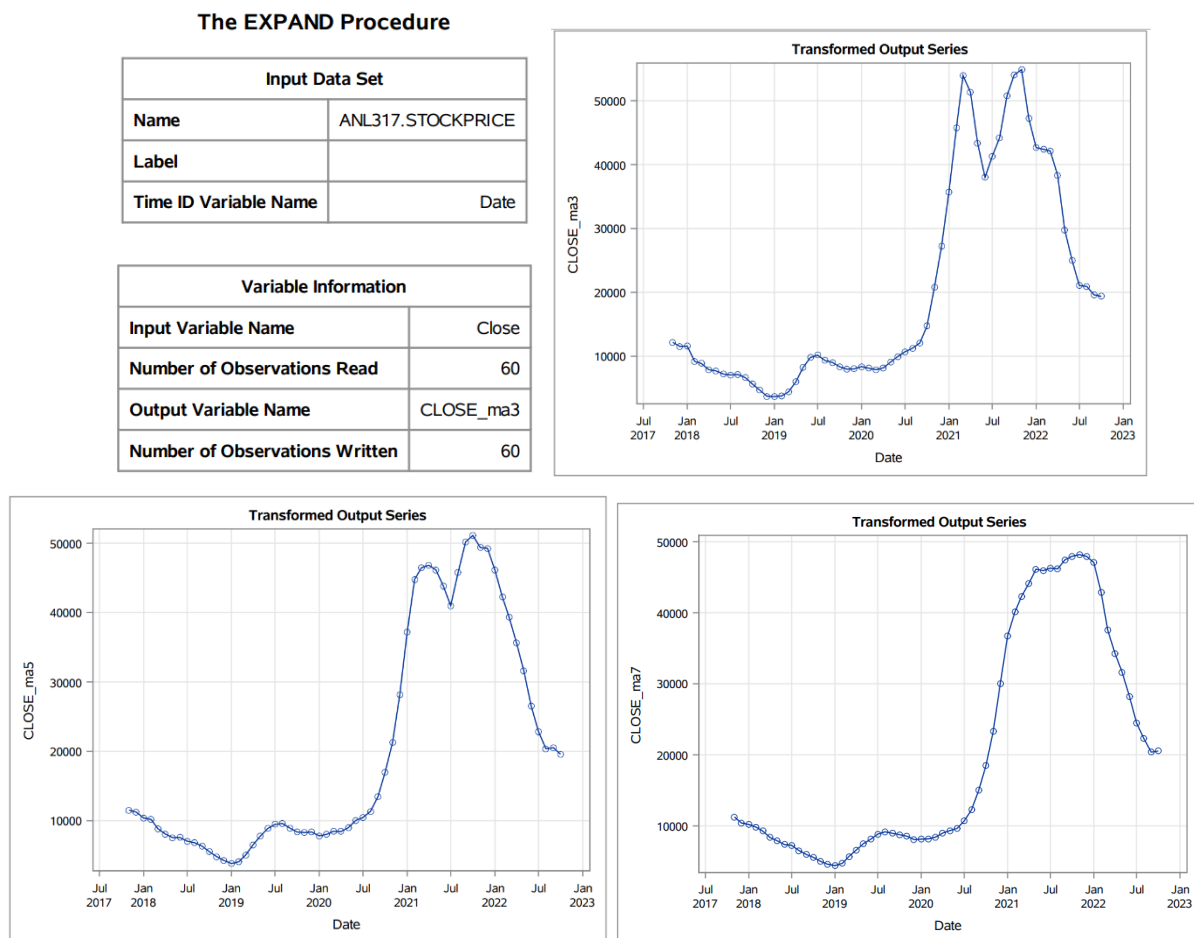
RUN;

Question 2 (a)

We use SAS code Proc Expand to calculate the moving average of the dataset for orders 3, 5 & 7 respectively. The output is plotted on **Figure 10**.

```
Code: PROC EXPAND DATA=ANL317G.StockQ12 OUT=ANL317G. StockQ12_ma  
      PLOTS=TRANSFORMOUT;  
      ID Date;  
      CONVERT CLOSE = CLOSE_ma3 / TRANSOUT=(cmovave 3);  
      CONVERT CLOSE = CLOSE_ma5 / TRANSOUT=(cmovave 5);  
      CONVERT CLOSE = CLOSE_ma7 / TRANSOUT=(cmovave 7);  
      RUN;
```

Figure 10 Output of Expand with Moving Average of 3, 5, 7.



In **Figure 10**, we could observe that moving average order 3 have smoothen out some of the minor / insignificant fluctuation while still largely resemble the trend, fluctuation, and movement of the monthly dataset. For example, in **Figure 6**, we could observe fluctuations on the downward trend during the initial periods, while in Model 3 in **Figure 10**, the overall trend during this period is still maintained with the fluctuation removed.

On the other hand, although moving average order 7 is the smoothest among the 3 models as it is the average of 7 observations, some of the significant movements / fluctuations especially the dip in between the 2 spikes between Jan 2021 and Jan 2022 were lost.

Evaluation: Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), Mean Square Deviation (MSD)

Beside comparing the 3 orders visually, we also utilize typical evaluation models such as MAPE, MAD and MSD to quantify the accuracy of the different moving average orders. We apply the backward moving average methodology to produce the forecasted values under the 3 different orders and compare it with the actual values to calculate its accuracy. The formulas for each evaluation model are as follows:

MAPE Formula: $\text{sum}((\text{actual} - \text{forecast})/\text{actual} * 100) / (\# \text{ of observations})$
MAD Formula: $\text{sum}(\text{actual} - \text{forecast}) / (\# \text{ of observations})$
MSD Formula: $\text{sum}((\text{actual} - \text{forecast})^2) / (\# \text{ of observations})$

The results of evaluation are tabulated in **Figure 11**.

Figure 11 MAPE, MAD, MSD Result for MA Order 3, 5, 7

MA Model	Observation count	MAPE	MAD	MSD
CMA_3	57	23.0587	14,817.5	282,164,993
CMA_5	55	28.1359	15,029.3	276,897,985
CMA_7	53	32.3657	15,099.6	271,821,224

Moving average order 3 has the smallest value in terms of MAPE and MAD as compared to other two orders, while moving average order 7 produces the lowest MSD. As Order 3 has the smallest number (ie. lowest deviation) in 2 out of 3 indicators, we will use Order 3 to forecast the values for the next 12 periods using Proc ARIMA.

ARIMA Forecasting

We use SAS code Proc Arima to forecast the values of the next 12 months. The output is extracted in **Figure 12**.

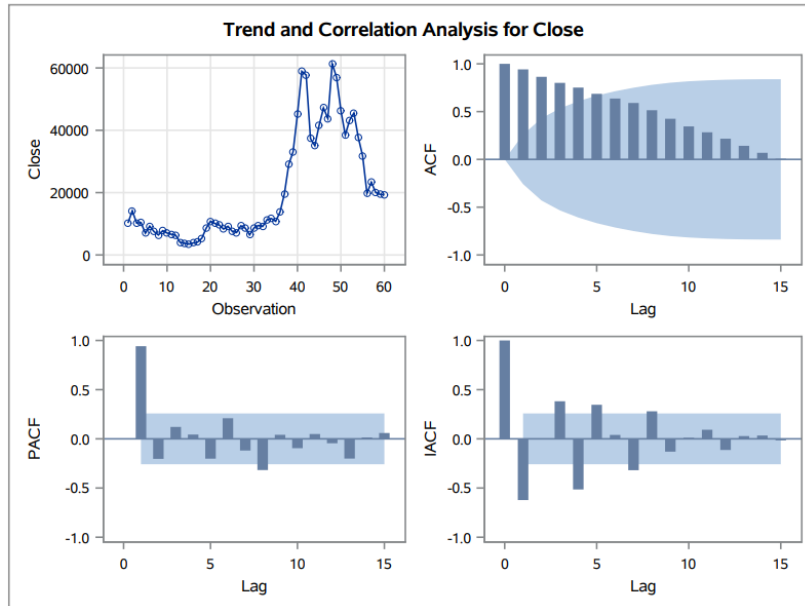
```
Code: PROC ARIMA DATA= ANL317G.StockQ12 OUT= ANL317G.StockQ12_maforecast;
      IDENTIFY var=Close;
      ESTIMATE p=(1 2 3) ar=(0.33 0.33 0.33) NOINT NOEST
      NOSTABLE METHOD=CLS;
      FORECAST LEAD=12;
RUN;
```

Figure 12 ARIMA Output for Model 3

The ARIMA Procedure

Name of Variable = Close	
Mean of Working Series	20163.14
Standard Deviation	16872.87
Number of Observations	60

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	242.90	6	<.0001	0.942	0.864	0.801	0.752	0.686	0.637
12	318.02	12	<.0001	0.590	0.514	0.425	0.345	0.283	0.216



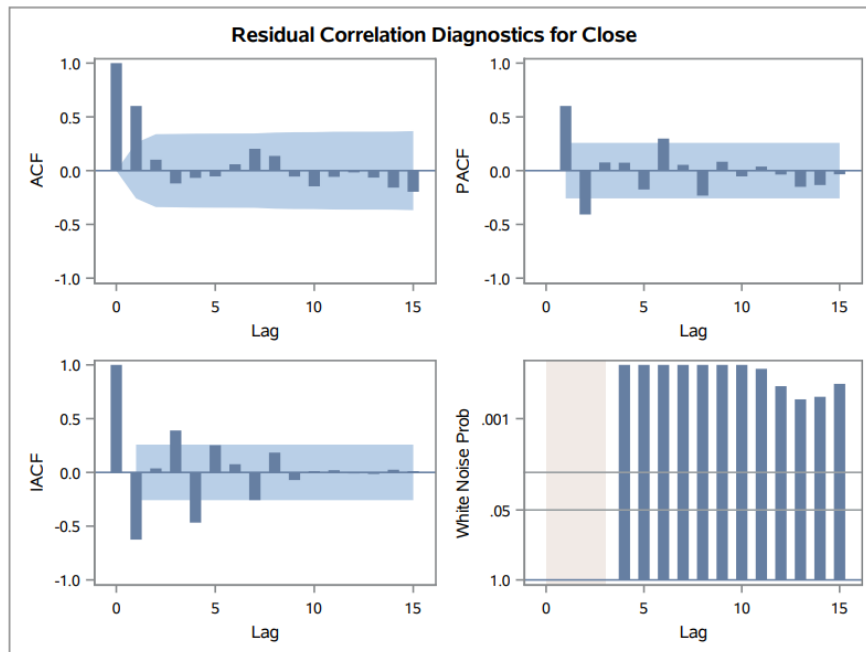
Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
AR1,1	0.33000	0	Infty	<.0001	1
AR1,2	0.33000	0	Infty	<.0001	2
AR1,3	0.33000	0	Infty	<.0001	3

The ARIMA Procedure

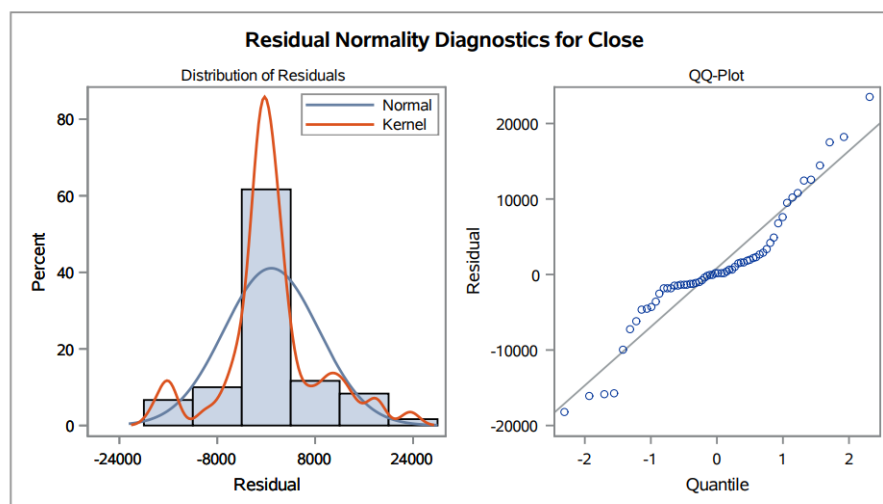
Variance Estimate	63247479
Std Error Estimate	7952.828
AIC	1250.949
SBC	1257.232
Number of Residuals	60

* AIC and SBC do not include log determinant.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	25.26	3	<.0001	0.604	0.108	-0.109	-0.056	-0.037	0.077
12	31.61	9	0.0002	0.221	0.155	-0.033	-0.119	-0.030	0.009
18	41.60	15	0.0003	-0.042	-0.134	-0.174	-0.188	-0.153	-0.101
24	43.09	21	0.0031	-0.051	0.000	0.040	0.010	-0.061	-0.083



The ARIMA Procedure

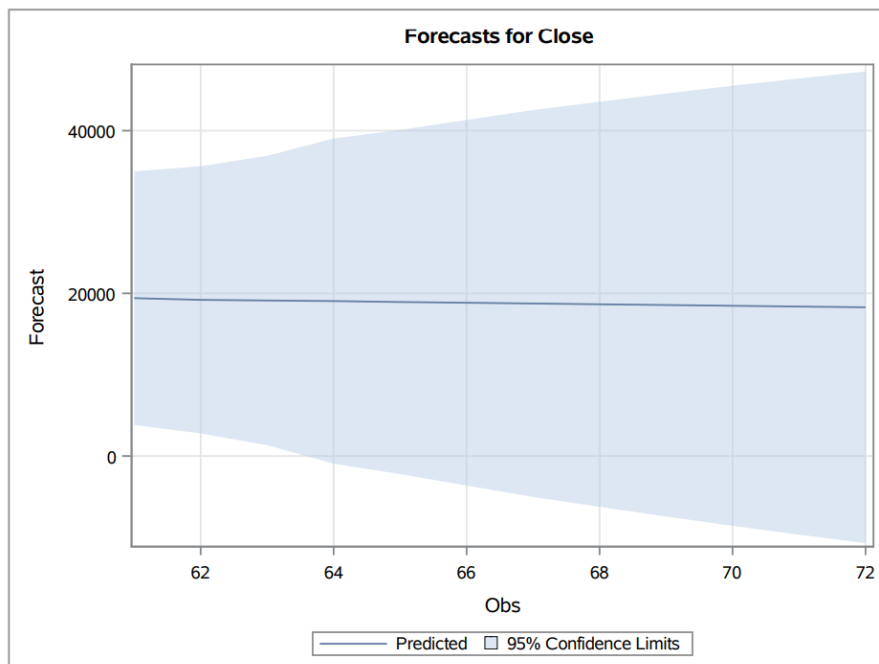


Model for variable Close

No mean term in this model.

Autoregressive Factors	
Factor 1:	$1 - 0.33 B^{**}(1) - 0.33 B^{**}(2) - 0.33 B^{**}(3)$

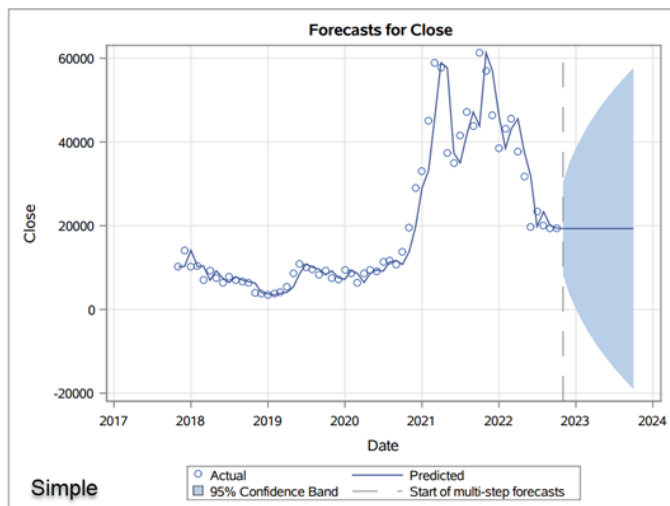
Forecasts for variable Close				
Obs	Forecast	Std Error	95% Confidence Limits	
61	19409.2701	7952.8284	3822.0129	34996.5273
62	19197.9084	8374.6719	2783.8531	35611.9638
63	19120.7275	9072.9651	1338.0428	36903.4123
64	19050.2090	10191.673	-925.1035	39025.5215
65	18931.7188	10792.221	-2220.6456	40084.0832
66	18843.8763	11460.168	-3617.6411	41305.3937
67	18752.5154	12130.196	-5022.2320	42527.2627
68	18654.2764	12698.305	-6233.9448	43542.4976
69	18562.7205	13261.486	-7429.3135	44554.7545
70	18469.9390	13798.810	-8575.2318	45515.1098
71	18376.6889	14301.788	-9654.2996	46407.6773
72	18285.0850	14787.987	-10698.8369	47269.0068



Question 2 (b)

Exponential smoothing method filter out short term volatilities from the series and it take into consideration the “importance of recency” of observation relative to older ones. Highest weightage is given to the most recent past observation and the weights will be reduced exponentially as observations get further away from the present. **Figure 13** shows the plot of smoothed values (in solid line). The smoothed values for simple, double and winters – additive are close to the long-term development of the original series. However, winter – multiplicative missed to capture volatile observations between 2021 to 2023. This could be an indication the time series is not multiplicative.

Figure 13 Result of Simple, Double, Additive Winters, Winters Exponential Smoothing

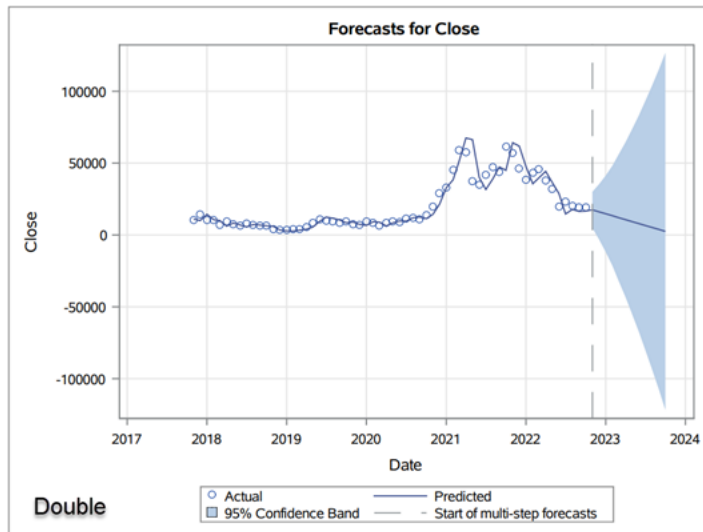


```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_ses PLOT=FORECASTS;
```

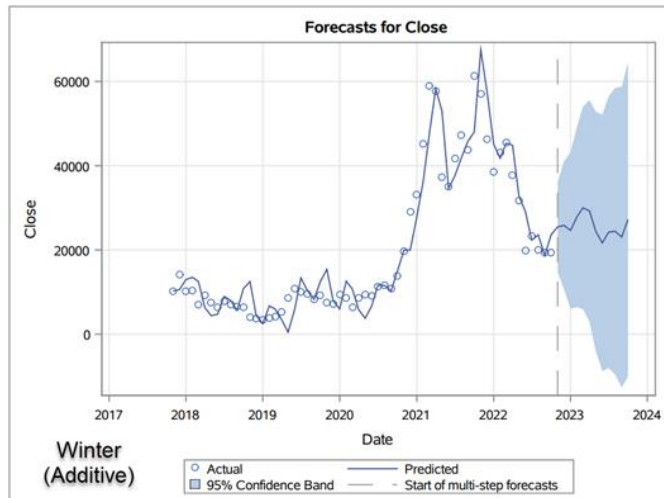
```
    ID Date INTERVAL=month;
```

```
    FORECAST Close / MODEL=SIMPLE;
```

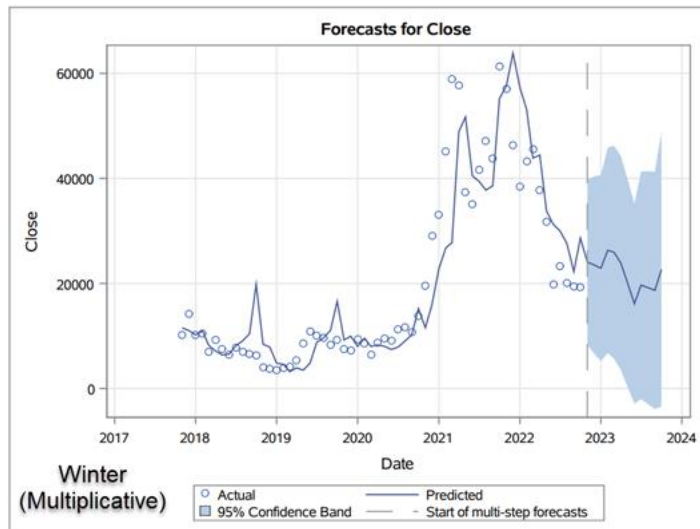
```
    RUN;
```

```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_des PLOT=FORECASTS;
      ID Date INTERVAL=month;
      FORECAST Close / MODEL=DOUBLE;
RUN;
```



```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_awin PLOT=FORECASTS;
      ID Date INTERVAL=month;
      FORECAST Close / MODEL=ADDWINTERS;
RUN;
```

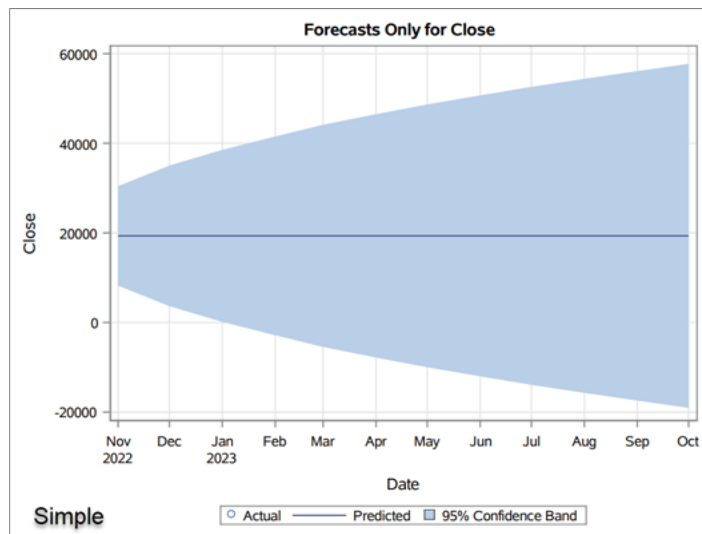


```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_mwin PLOT=FORECASTS;
      ID Date INTERVAL=month;
      FORECAST Close / MODEL=WINTERS;
RUN;
```

Figure 14 shows the predicted forecast values of the next 12-months for respective smoothing methods. Application of different exponential smoothing depends on the condition of the time series. The simple exponential smoothing method is suitable for time series without trend and seasonality. Double exponential smoothing method is suitable for time series with a trend but without seasonality. For time series with seasonality, it is suitable to use Winters method, and depending on the magnitude of seasonality, it can be additive (i.e., constant magnitude) or multiplicative (i.e., magnitude develop proportionately).

Figure 14 Forecast prediction by smoothing method.

		FORECAST - PREDICT			
		Simple	Double	Winter (Additive)	Winter (Multiplicative)
Forecast	Nov-22	19334.51799	17477.99764	25480.01565	24051.89596
	Dec-22	19334.51799	16118.47336	25842.72133	23550.22391
	Jan-23	19334.51799	14758.94908	24633.72348	22895.1369
	Feb-23	19334.51799	13399.4248	27890.71113	26317.85398
	Mar-23	19334.51799	12039.90051	29994.69967	25948.44474
	Apr-23	19334.51799	10680.37623	29288.26443	23874.56724
	May-23	19334.51799	9320.851949	24421.94232	20146.06801
	Jun-23	19334.51799	7961.327667	21673.63952	16133.78193
	Jul-23	19334.51799	6601.803385	24212.9187	19667.35089
	Aug-23	19334.51799	5242.279103	24441.0915	19204.11828
	Sep-23	19334.51799	3882.754821	23059.10401	18690.14405
	Oct-23	19334.51799	2523.230539	27239.74678	22657.58316

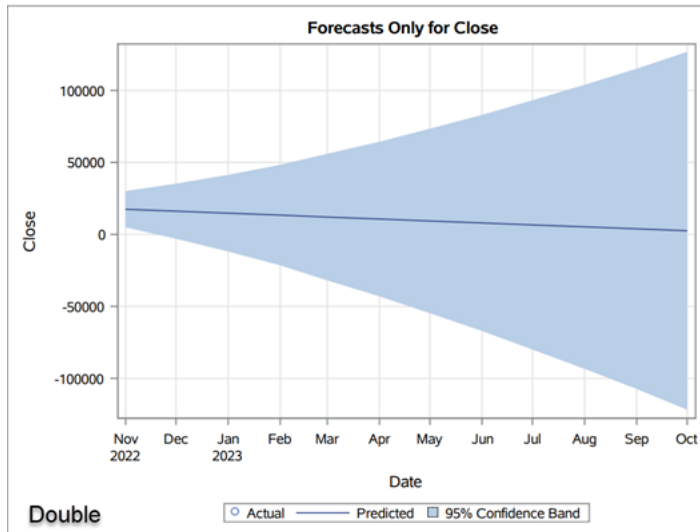


```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_sesFCstat
      PLOT=FORECASTSONLY LEAD=12 PRINT=statistics;

      ID Date INTERVAL=month;

      FORECAST Close / MODEL=SIMPLE;

      RUN;
```

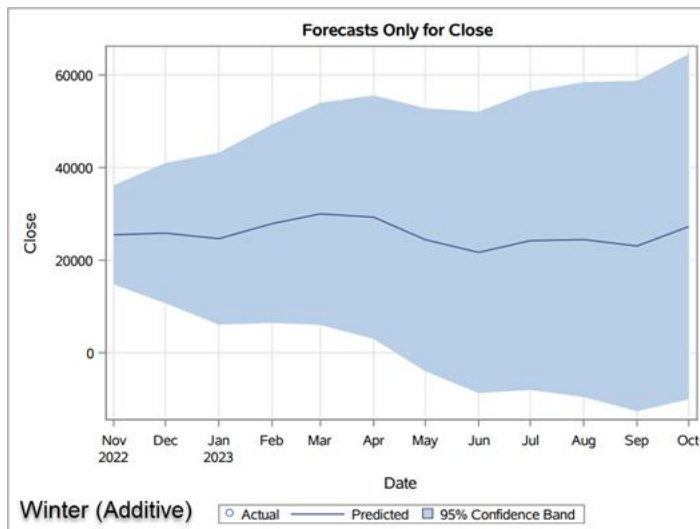


```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_desFCstat
      PLOT=FORECASTSONLY LEAD=12 PRINT=statistics;

      ID Date INTERVAL=month;

      FORECAST Close / MODEL=DOUBLE;

      RUN;
```

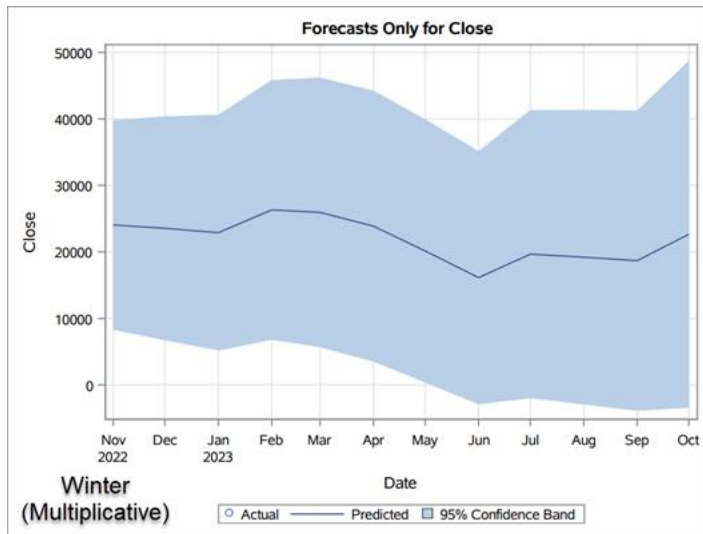


```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_awinFCstat
      PLOT=FORECASTSONLY LEAD=12 PRINT=statistics;

      ID Date INTERVAL=month;

      FORECAST Close / MODEL=ADDWINTERS;

      RUN;
```



```
Code: PROC ESM DATA=ANL317G.STOCKQ12 OUTFOR=ANL317G.stockq12_mwinFCstat
      PLOT=FORECASTSONLY LEAD=12 PRINT=statistics;

      ID Date INTERVAL=month;

      FORECAST Close / MODEL=WINTERS;

RUN;
```

Noted that the forecasted value using simple exponential smoothing method is constant for all 12-months. This may not be truly reflective of the close price time series which has seasonality and cyclical component. **Figure 15** tabulates the key evaluation criteria readings generated from SAS. Comparatively, simple exponential smoothing (which has the lowest MAPE and MAD) and winter – additive (which has the lowest MSD and AIC) are the two better models. It is worth nothing that the MAD score for Winter (additive) is not too far off from the score of Simple exponential model. Since there is a trend and seasonality cycle in the time series, Winter (additive) is the preferred model.

Figure 15 Evaluation criteria generated from SAS.

	MODEL EVALUATION			
	Simple	Double	Winter (Additive)	Winter (Multiplicative)
MAPE	17.7734	20.0811	29.3717	30.7096
MAD	3,586.3022	4,112.6116	3,899.9581	5,206.3524
MSD	31,506,684.3	40,656,314.2	28,415,554.2	61,576,319.5
AIC	1,037.94262	1,053.23988	1,035.74683	1,082.14728

Question 2 (c) Comparing the results between Moving Average Method vs Smoothing Method

In **Figure 16**, we tabulate the MAPE, MAD, MSD and AIC of the two methods to compare their forecasting accuracy. Winter (Additive) model is more superior in 3 out of 4 evaluation criteria. It has lower score in MAD, MSD and AIC, although the moving average model produces lower MAPE of 23.0587. As such, Winter (Additive) is preferred as the forecasting method since it is more accurate in 3 out of 4 evaluation criteria.

Figure 16 Model Comparison of Moving Average 3 Model and Additive Winter Model

	Model Comparison	
	Moving Average order of 3	Winter (Additive)
MAPE	23.0587	29.3716901
MAD	14,817.5150	3,899.9581
MSD	282,164,992.86	28,415,554.2
AIC	1250.949	1035.74683

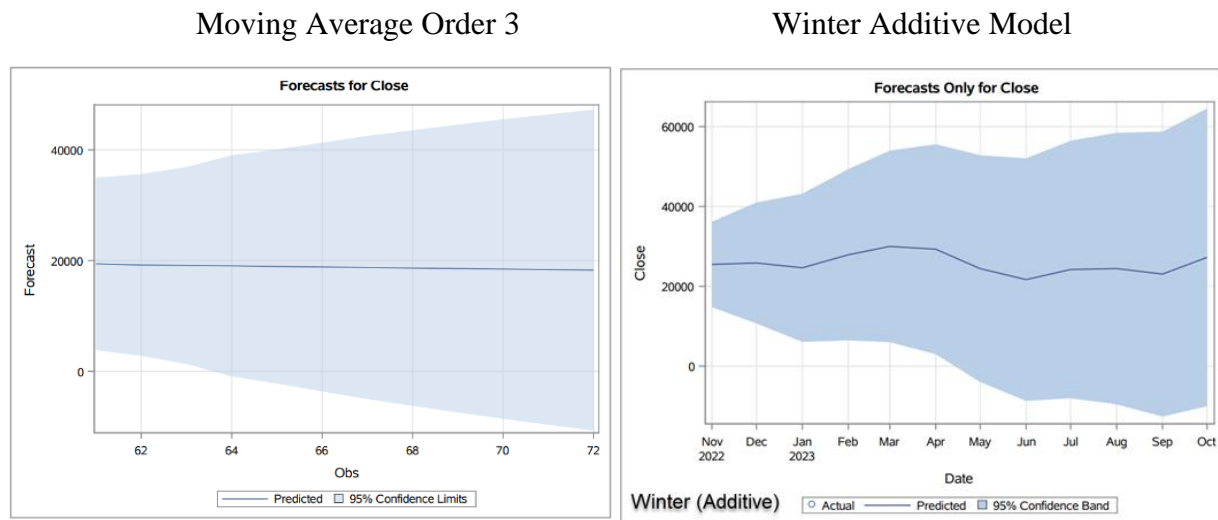
The forecasted values from the two models as well as their respective standard errors and 95% confidence range are tabulated in Figure 17 for easy reference. Moving average model produces relatively constant forecasted values (hovering around 18000 to 19000 range) in the next 12 months compared to Winter (Additive) model which fluctuates within a wider range of 21000 to 29000. The forecasted values produced by Winter (additive) model is more realistic and better resemble the real-life movement of stock price. Based on the historical trend over the last 60 months, it is unlikely that the future prices will fluctuate within such small range as forecasted by the moving average model.

We also observe that at the earlier forecasting period, the moving average model has higher standard deviation and wider 95% confidence range than the Winter (Additive) model, indicating the latter has higher confidence in its forecasted values, although the trend reverses towards the later part of the forecasting period.

Stock price is a real time reflection of a company past performance as well as future potential earnings based on recent performance, Therefore, the more recent observations should take a bigger weightage in predicting the future. Backward moving average used for forecasting allows equal weightage assigned to past observations regardless of how far the time lag it is away from y_t . Therefore, the advantage of exponential smoothing method which assigns much

higher weightage to recent observations will outweigh moving average. Overall, smoothing method (i.e. Winter Additive model specifically) is preferred over the moving average model.

Figure 17 Forecasted Values, Standard Deviation and 95% confidence range between the two models



Forecasts for variable Close				
Obs	Forecast	Std Error	95% Confidence Limits	
61	19409.2701	7952.8284	3822.0129	34996.5273
62	19197.9084	8374.6719	2783.8531	35611.9638
63	19120.7275	9072.9651	1338.0428	36903.4123
64	19050.2090	10191.673	-925.1035	39025.5215
65	18931.7188	10792.221	-2220.6456	40084.0832
66	18843.8763	11460.168	-3617.6411	41305.3937
67	18752.5154	12130.196	-5022.2320	42527.2627
68	18654.2764	12698.305	-6233.9448	43542.4976
69	18562.7205	13261.486	-7429.3135	44554.7545
70	18469.9390	13798.810	-8575.2318	45515.1098
71	18376.6889	14301.788	-9654.2996	46407.6773
72	18285.0850	14787.987	-10698.8369	47269.0068

Winter - additive							
	Date	Actual	Predict	Lower	Upper	Error	STD
61	Close Nov-22	.	25480.01565	14760.7668	36199.26451	.	5469.105015
62	Close Dec-22	.	25842.72133	10683.4218	41002.02086	.	7734.478619
63	Close Jan-23	.	24633.72348	6061.267585	43206.17936	.	9475.916923
64	Close Feb-23	.	27890.71113	6436.159357	49345.26291	.	10946.40103
65	Close Mar-23	.	29994.69967	5996.999497	53992.39985	.	12243.94957
66	Close Apr-23	.	29288.26443	2987.849733	55588.67912	.	13418.8255
67	Close May-23	.	24421.94232	-3999.257844	52843.14249	.	14500.87879
68	Close Jun-23	.	21673.63952	-8724.522607	52071.80165	.	15509.55138
69	Close Jul-23	.	24212.9187	-8044.839562	56470.67695	.	16458.34235
70	Close Aug-23	.	24441.0915	-9578.152604	58460.33561	.	17357.07614
71	Close Sep-23	.	23059.10401	-12638.04567	58756.25369	.	18213.16614
72	Close Oct-23	.	27239.74678	-10063.01149	64542.50504	.	19032.36925