Although, there are many time series forecasting methods, like

- Mean (average)
- moving averages
- simple exponential smoothing
- Hot's winter method
- Damped Trend method

and others.

But some of them lag seasonality components, some lag trend and some fail to incorporate error component.

I will be using **ETS** and **ARIMA** methods for predicting the next 2 weeks orders and then selecting one based on accuracy measures.

I will be using Alteryx for my assessment as it produces far better, interactive graphs than R, python.

ETS

Step 1: Data Load and Clean

Load the data, do some cleaning and, then separate EW1 and EW2 into subsets.

Step 2: Train/Test distribution

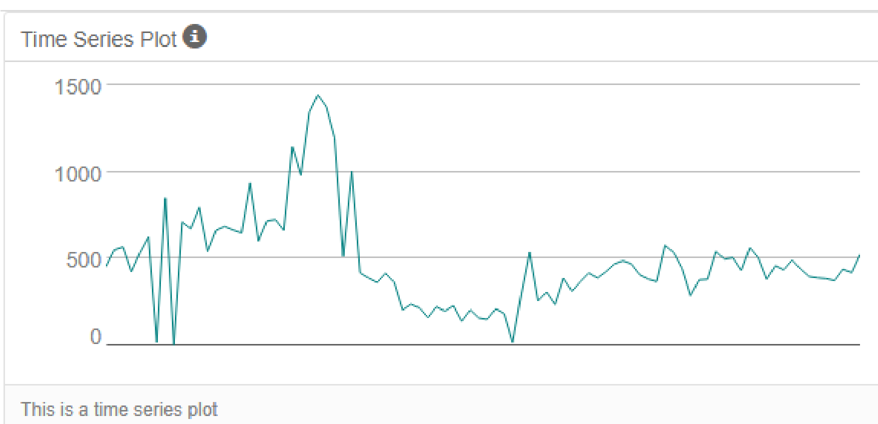
Take 80-90% of data as train dataset and remaining as test dataset. For **EW1**, I separated **76** out of 90 rows into test and remaining 14 into train dataset. For **EW2**, I separated **41** out of 55 rows into test and remaining 14 into train dataset.

Step 3: ETS configuration and implication

Check the attributes of ETS plot for each component. (Error, Trend and Seasonality)

Below are the reasons for selecting error, trend, and seasonality components.

EW1:



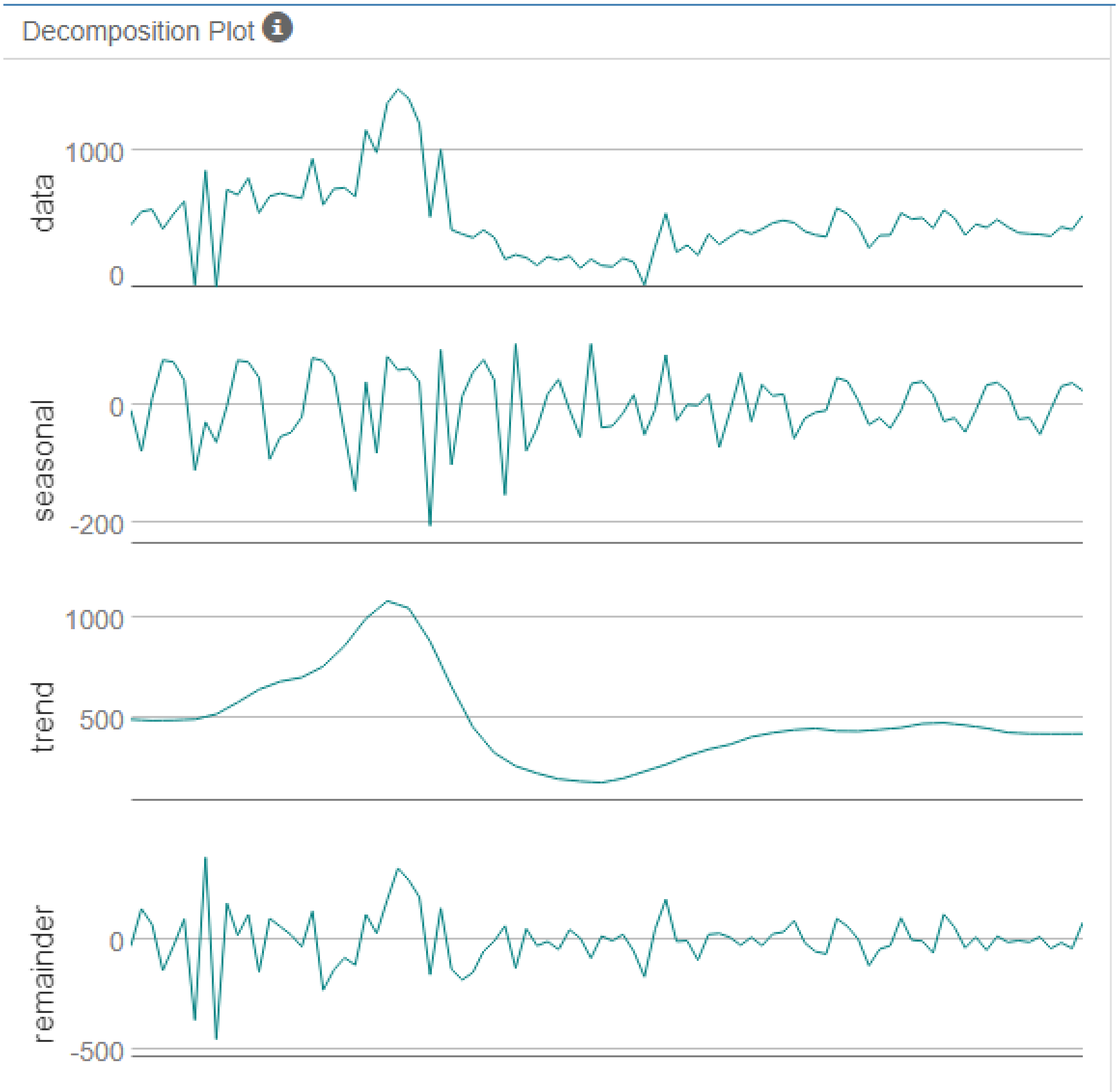
I configured my ETS tool for EW1 as **M,N,M**

Multiplicative: If decreasing or increasing exponentially.

Additive: If decreasing or increasing linearly.

Null: If there is no such pattern.

Error: Multiplicative , Trend: None, Seasonality: Multiplicative

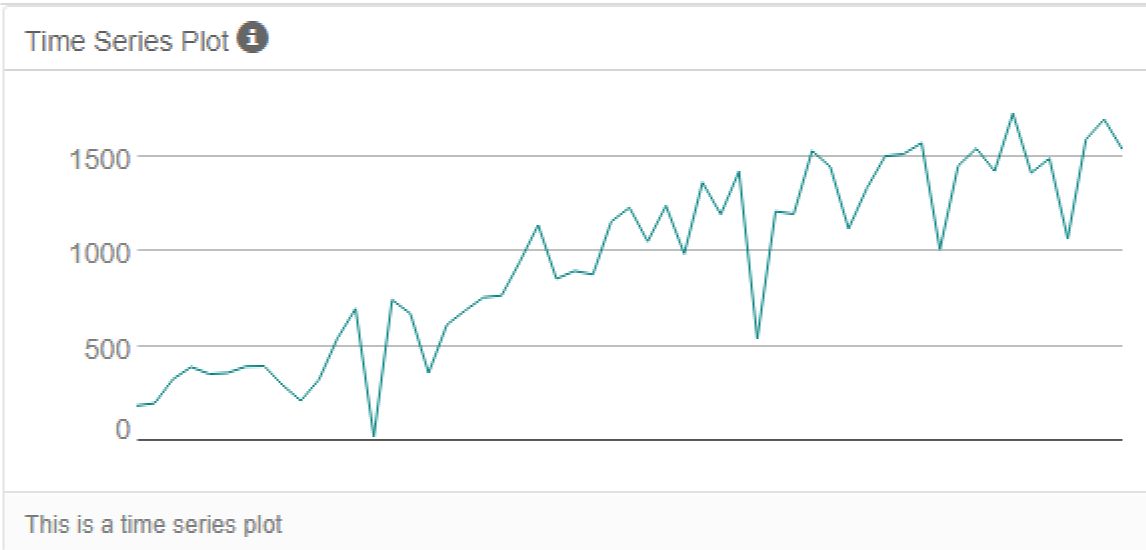


Multiplicative

None

Multiplicative

EW2:



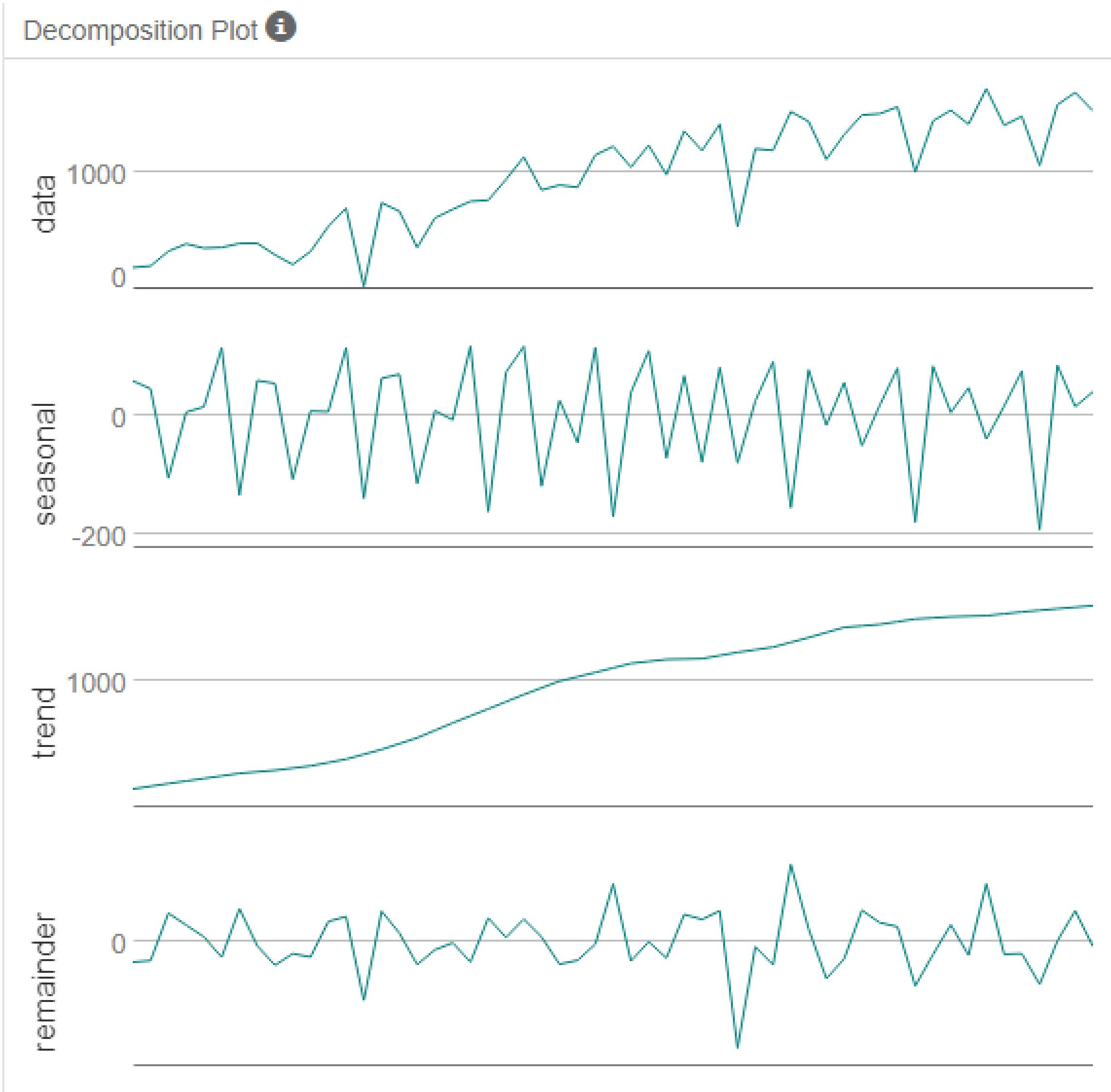
I configured my ETS tool for EW2 as **M,A,A**

Multiplicative: If decreasing or increasing exponentially.

Additive: If decreasing or increasing linearly.

Null: If there is no such pattern.

Error: Multiplicative , Trend: Additive, Seasonality: Additive



Multiplicative

Additive

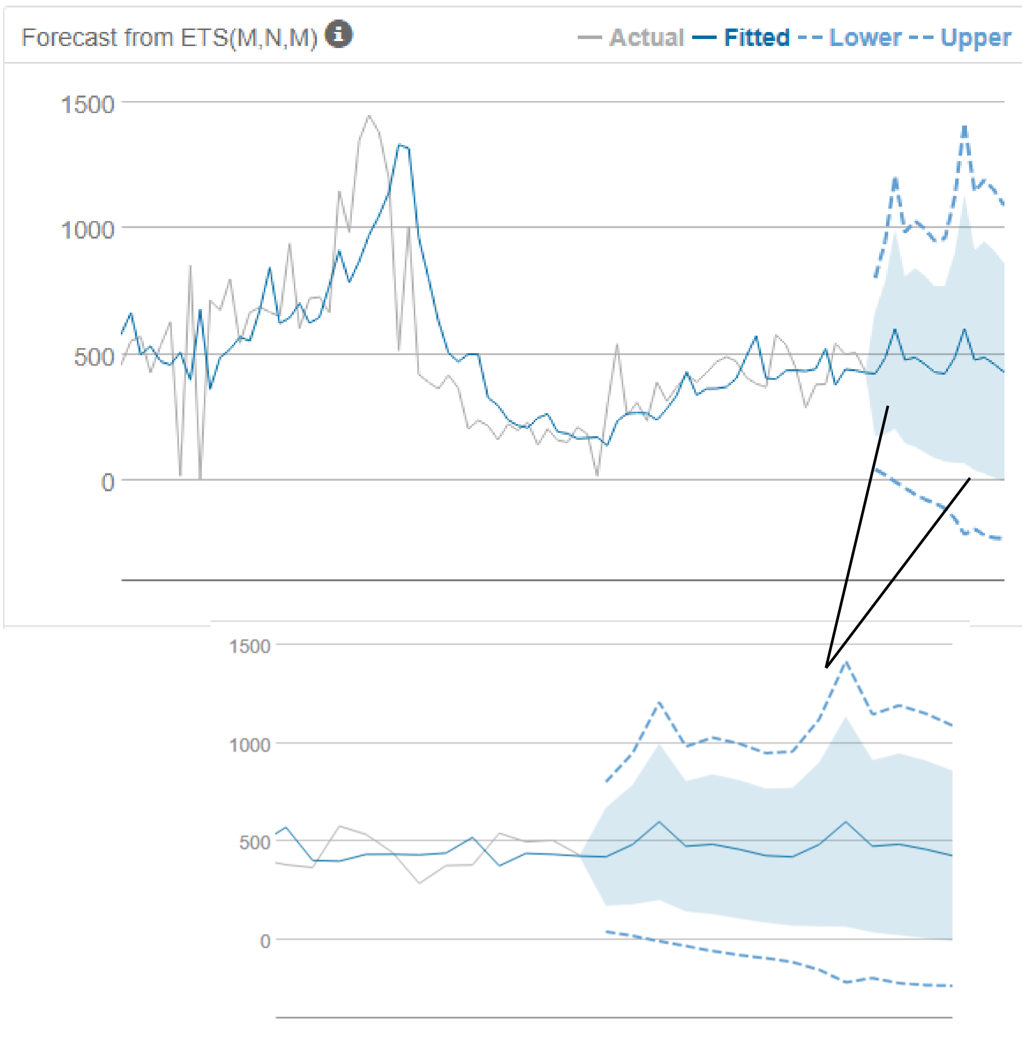
Additive

Step 4: Generating ETS forecast result and Comparing with test data

ETS tool is fitting the graph according to parameters which I have identified and then forecasting for next 2 weeks according to the fitted graph. I am then comparing the results with the test data set which I separated in step 2.

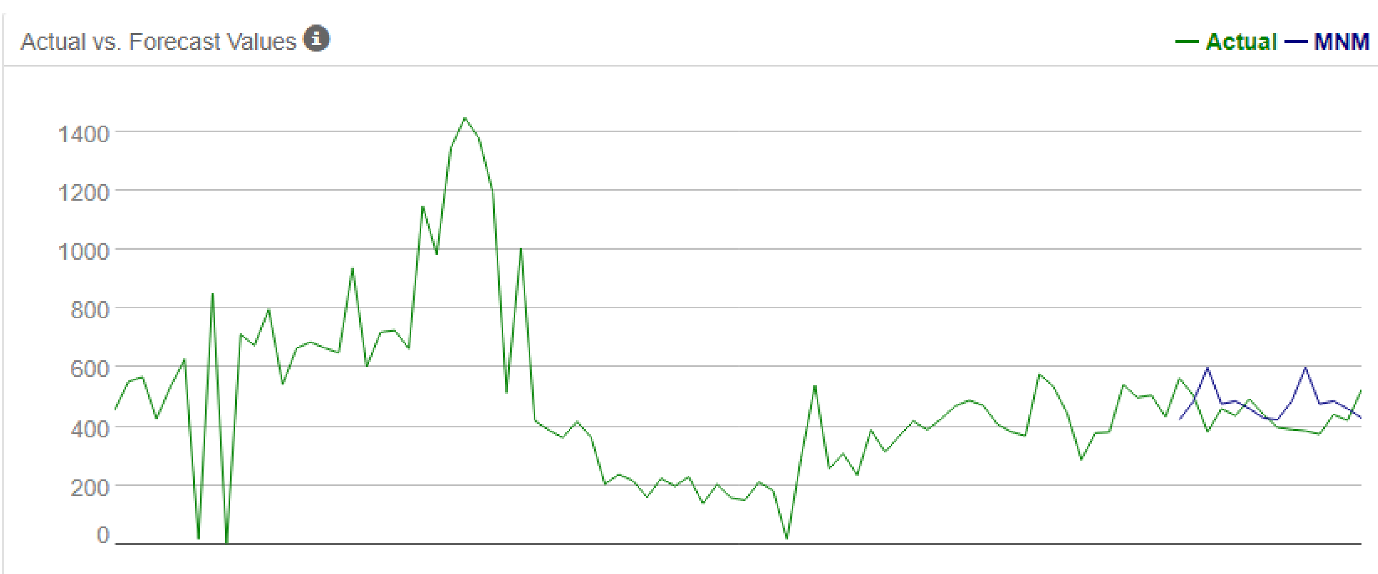
For EW1 and EW2, ETS model predicting quit well.

EW1:

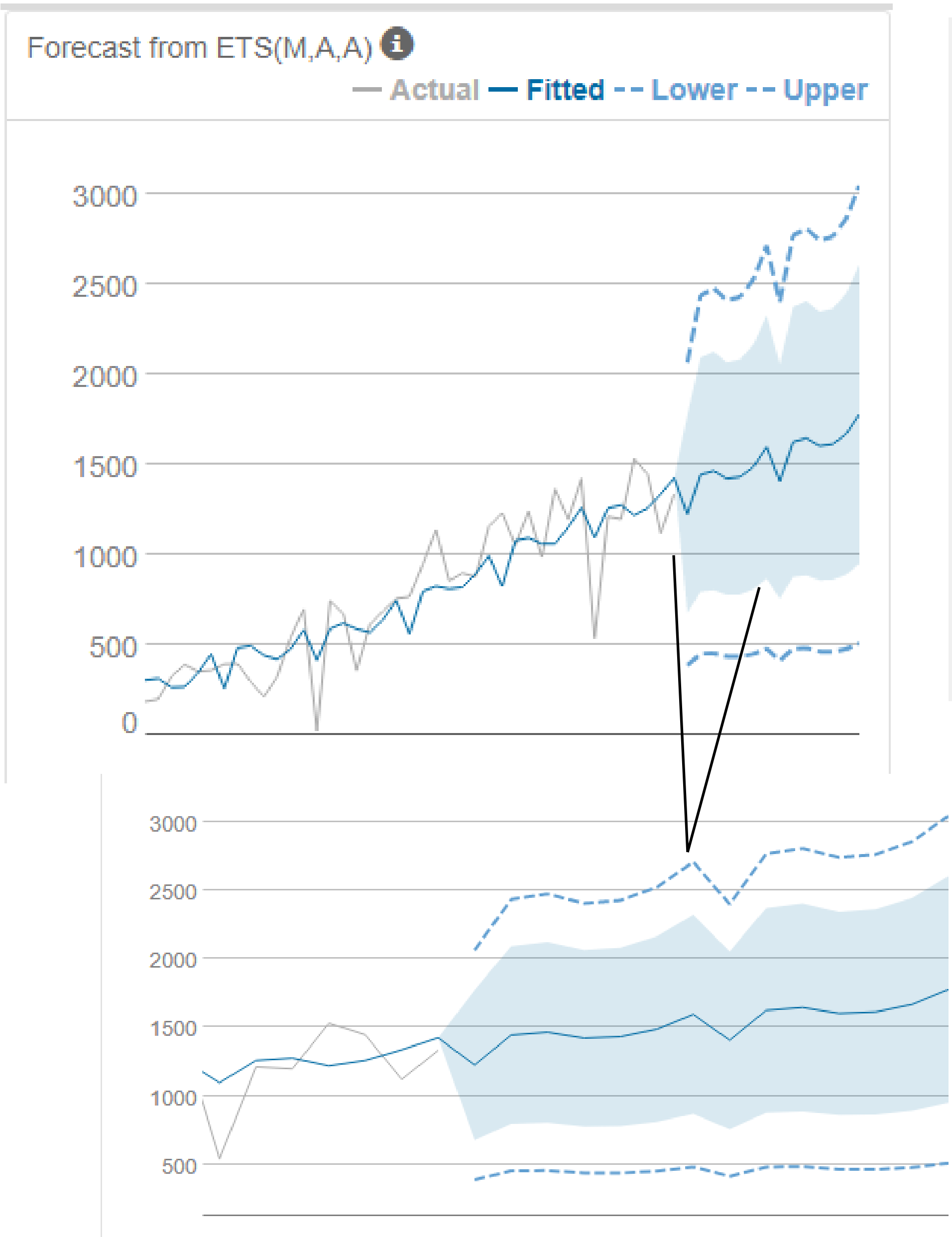


Actual and Forecast Values:

Actual	MNM
565	423.25598
506	485.51971
382	601.182
461	476.77502
437	486.66257
494	461.187
441	429.1343
398	423.25889
391	485.52305
386	601.18614
375	476.7783
441	486.66592
421	461.19017
526	429.13725

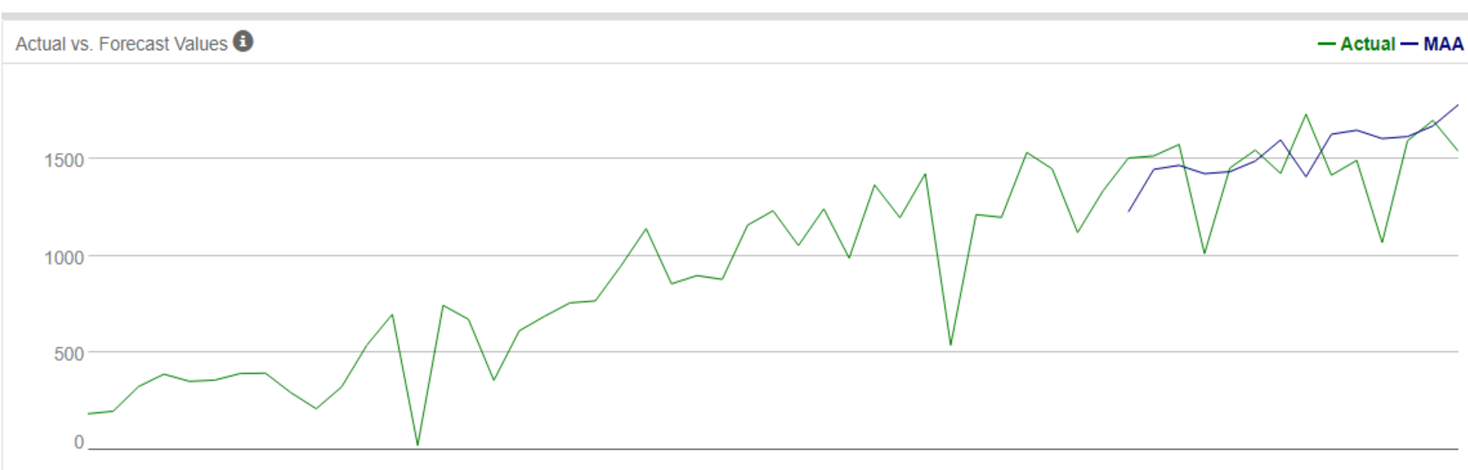


EW2

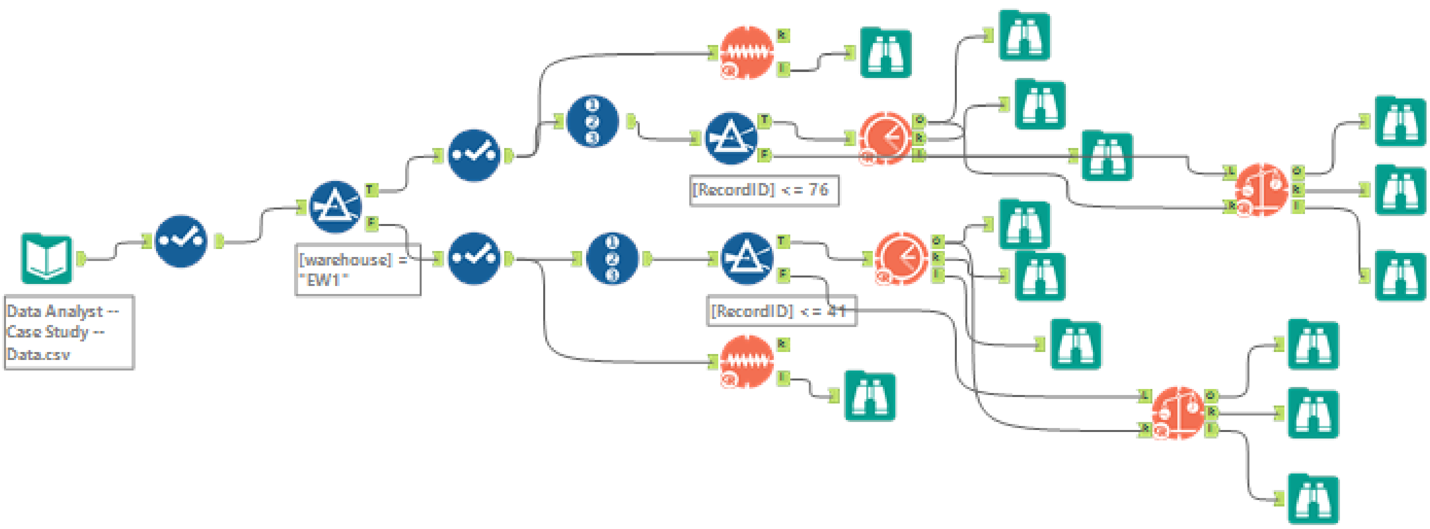


Actual and Forecast Values:

Actual	MAA
1502	1224.91587
1513	1443.81257
1572	1464.53437
1010	1421.64328
1451	1431.8749
1543	1486.5992
1423	1595.5011
1728	1405.7737
1414	1624.6704
1490	1645.3922
1067	1602.50111
1591	1612.73272
1696	1667.45703
1537	1776.35893



Alteryx Workflow for ETS model:



ARIMA

Step 1: Data Load and Clean

Load the data, do some cleaning and, then separate EW1 and EW2 into subsets.

Step 2: Train/Test distribution

Take 80-90% of data as train dataset and remaining as test dataset. For EW1, I separated **76** out of 90 rows into test and remaining 14 into train dataset. For EW2, I separated **41** out of 55 rows into test and remaining 14 into train dataset.

Step 3: ARIMA configuration and implication

As it is orders data which varies according to the days of the week, I have used seasonal ARIMA model. Regarding configuration of ARIMA, I did not only rely of ACF and PAC plots. But also tested models multiple times on test dataset to check what configuration settings on ARIMA components will produce better results.

EW1:

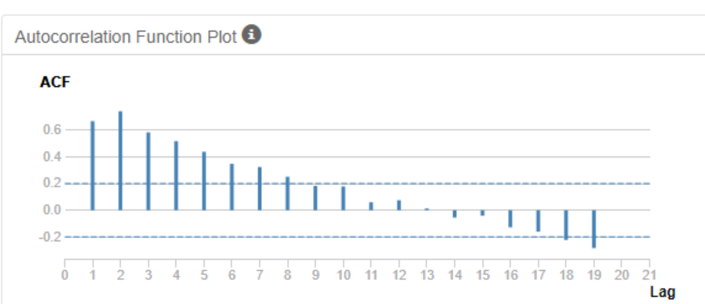
Non-seasonal differencing component: (p,d,q)

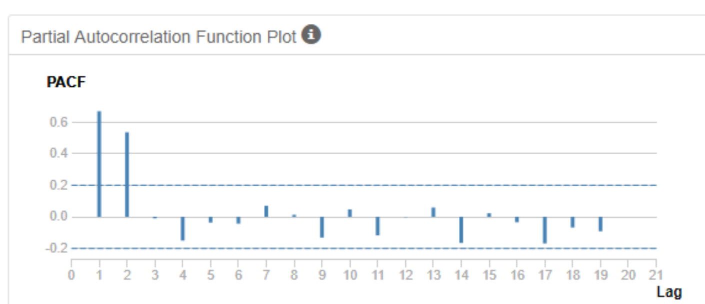
(1,1,0)

AR: Auto-regressive term is 1, because first lag is positive and there is one significant lag after first lag.

I: Differencing term is 1 because I use first differencing.

MA: Moving average term is 0, because first lag is positive.





Seasonal differencing component: (P,D,Q)

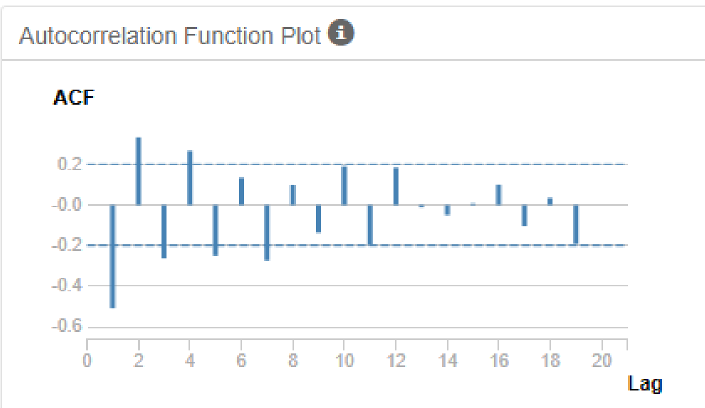
Differencing makes the ACF and PAC stationary and minimizes the correlation between the lags.

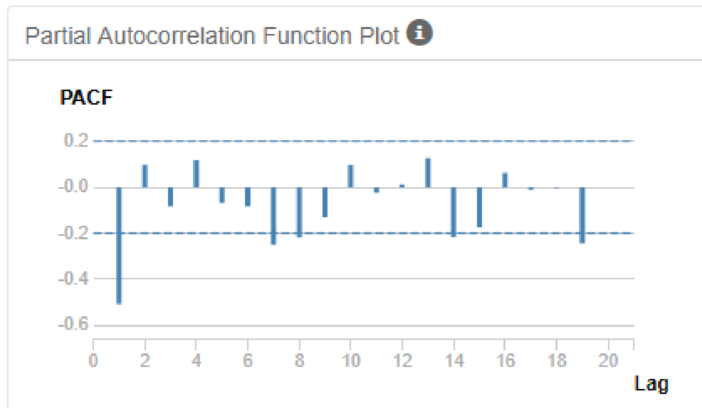
(0,1,1)

AR: Auto-regressive term is 0, because first lag is negative.

I: Differencing term is 1 because I use first differencing.

MA: Moving average term is 1, because first lag is negative and there is one significant lag after first lag.





EW2:

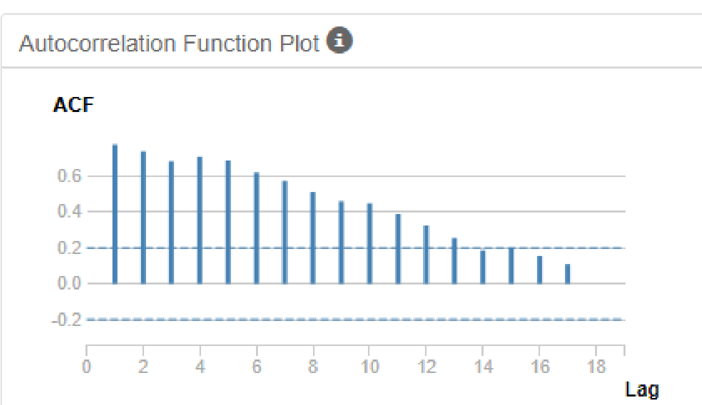
Non-seasonal differencing component: (p,d,q)

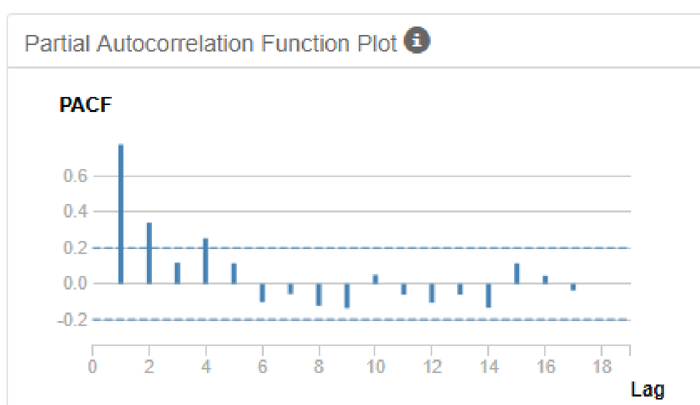
(3,1,0)

AR: Auto-regressive term is 3, because first lag is positive and there is one significant lag after first lag.

I: Differencing term is 1 because I use first differencing.

MA: Moving average term is 0, because first lag is positive.





Seasonal differencing component: (P,D,Q)

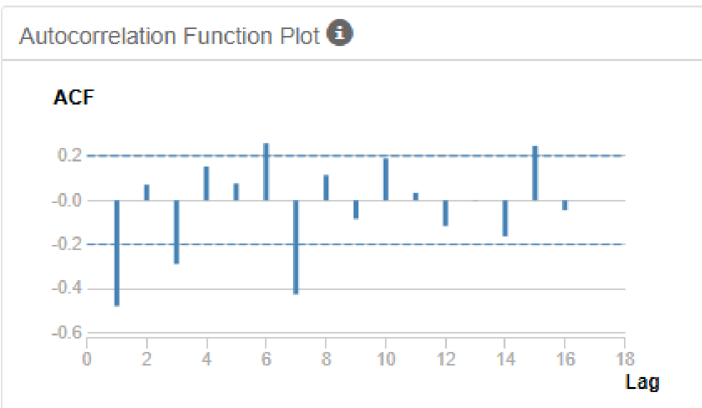
Differencing makes the ACF and PAC stationary and minimizes the correlation between the lags.

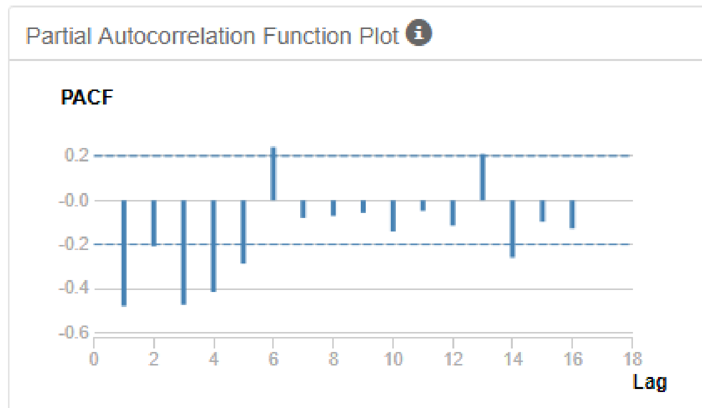
(0,1,2)

AR: Auto-regressive term is 0, because first lag is negative.

I: Differencing term is 1 because I use first differencing.

MA: Moving average term is 2, because first lag is negative and there are two significant lag after first lag.



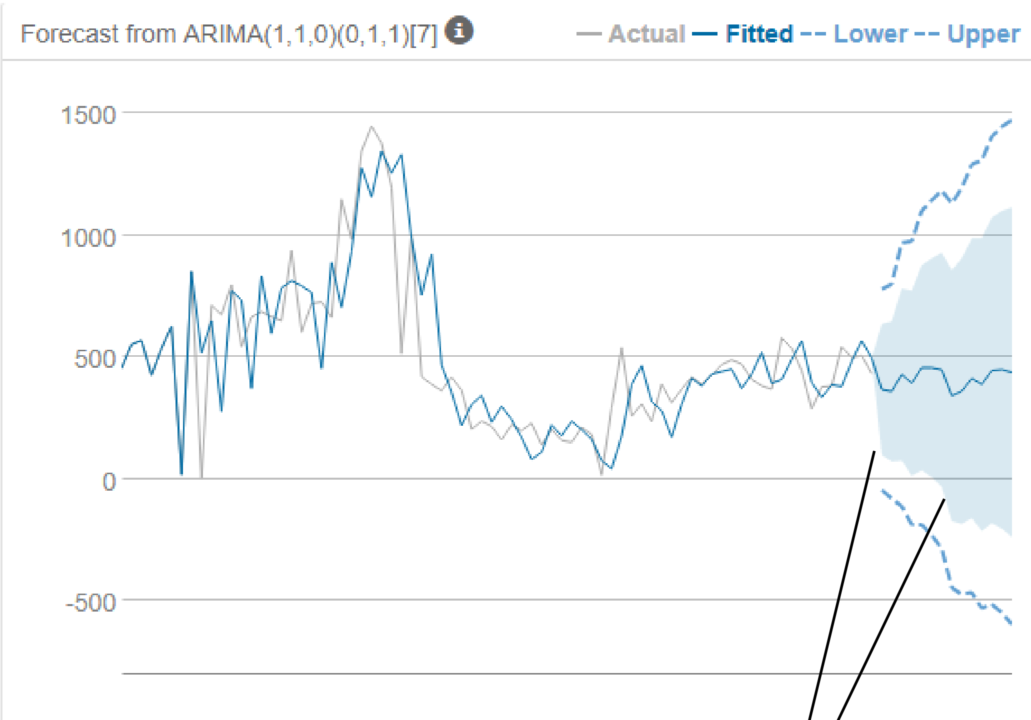


Step 4: Generating ETS forecast result and Comparing with test data

ARIMA is fitting the graph according to parameters which I have identified and then forecasting for next 2 weeks according to the fitted graph. I am then comparing the results with the test data set which I separated in step 2.

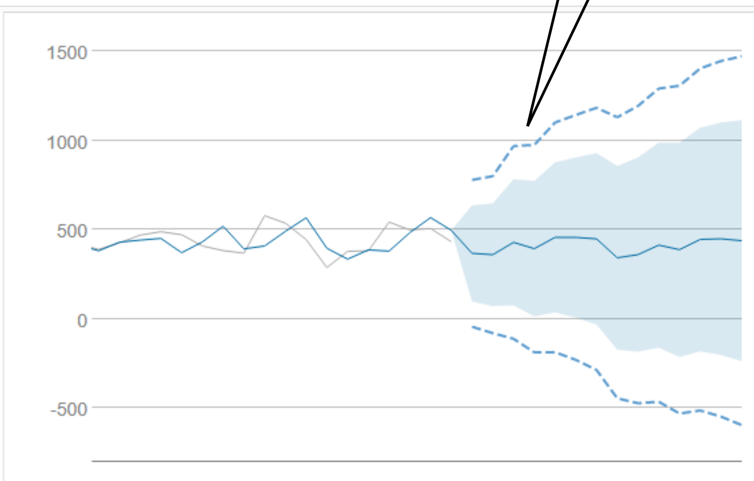
For both EW1 and EW2, *ARIMA model is predicting quite better than ETS model based on what I can observe from plots and data points.*

EW1:

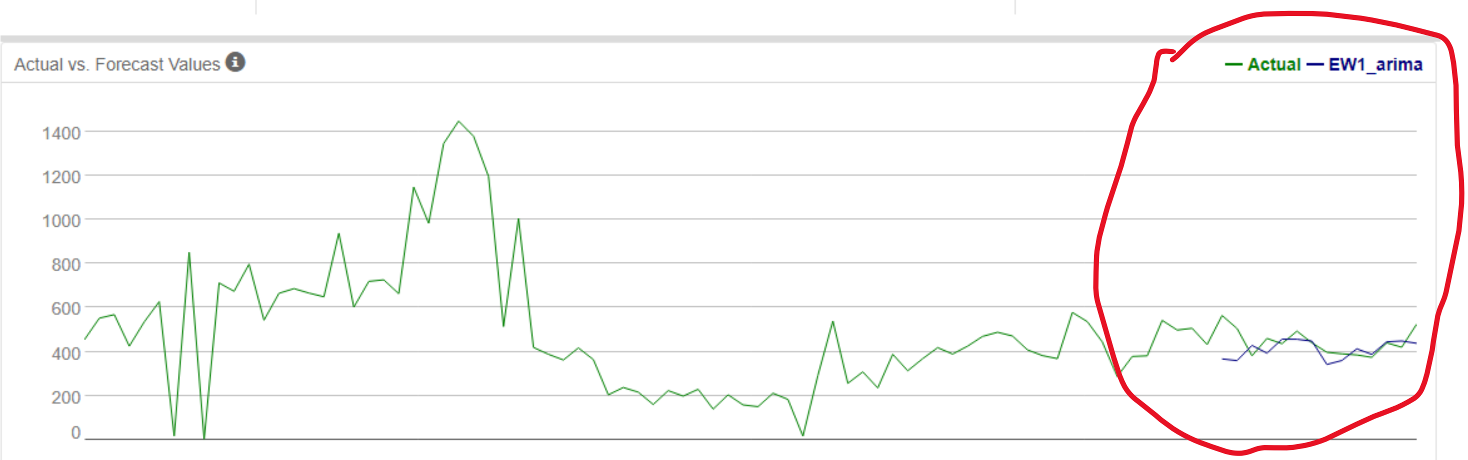


Actual and Forecast Values:

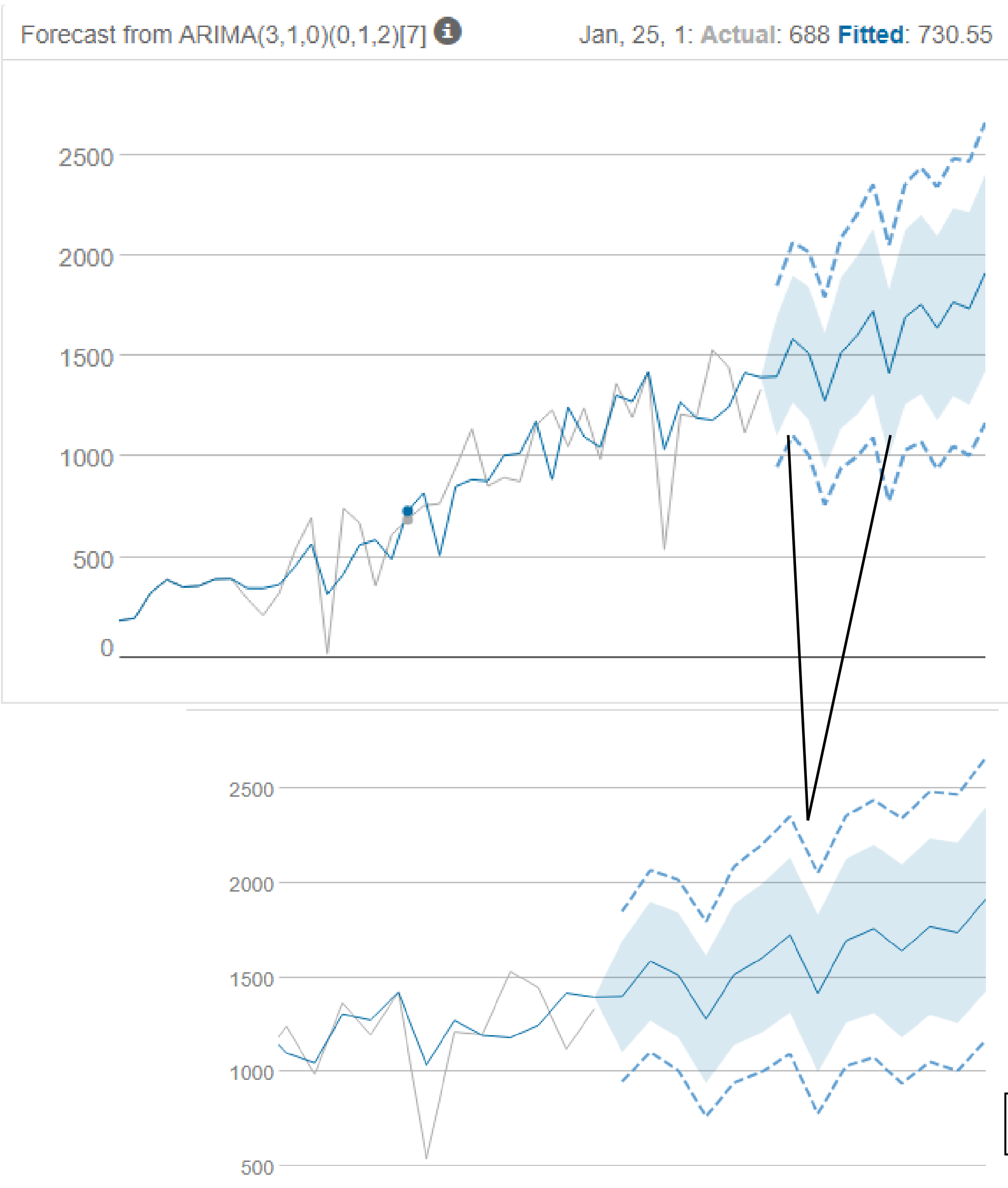
Actual	EW1_arima
565	367.82317
506	360.20503
382	429.14229
461	394.12563
437	457.36866
494	457.04863
441	448.95911
398	342.81346
391	360.66083
386	413.76917
375	388.59149
441	445.71878
421	449.20019
526	438.74777



👍 Much better forecast than ETS



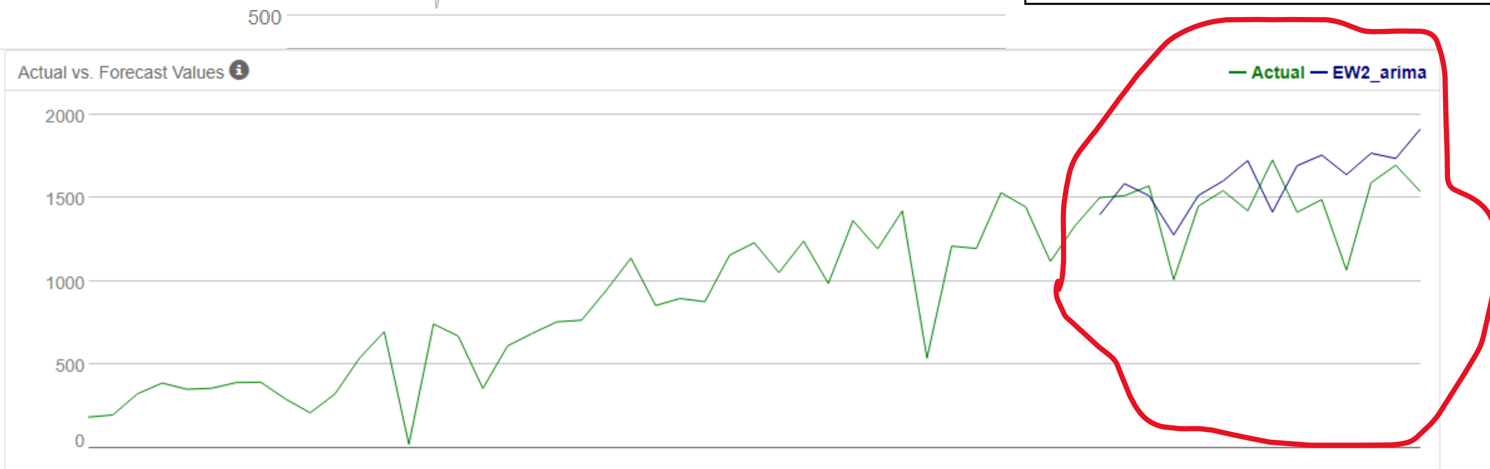
EW2:



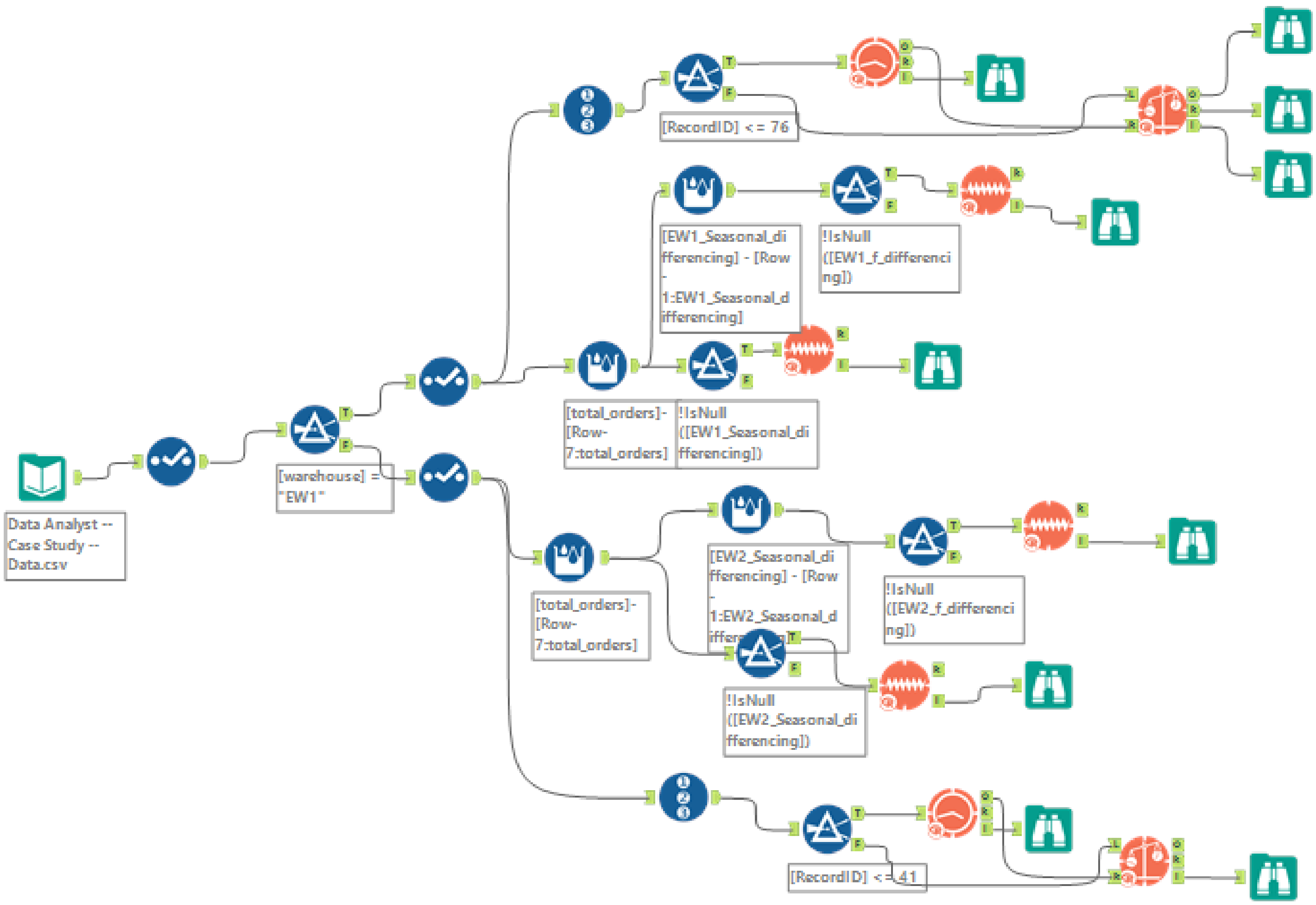
Actual and Forecast Values:

Actual	EW2_arima
1502	1398.71177
1513	1585.30785
1572	1513.29302
1010	1278.47199
1451	1514.85579
1543	1601.51168
1423	1723.62625
1728	1415.07361
1414	1692.99616
1490	1756.57955
1067	1639.34167
1591	1767.70603
1696	1736.41857
1537	1913.71659

👍 Much better forecast than ETS



Alteryx Workflow for ARIMA seasonal model:



ACCURACY &
MODEL SELECTION

PREDICTING NEXT 2 WEEKS DATA

Accuracy:

I am using Root Mean Square Error (RMSE) and AIC.

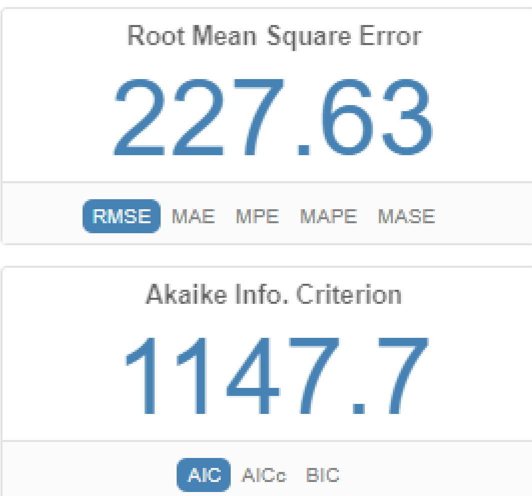
- RMSE: average distance between forecasted values and original values.
- AIC: to determine how well our model is fitting to dataset.

Both RMSE and AIC for ARIMA are lower than ETS for EW1 and EW2 both, suggesting better accuracy for ARIMA.

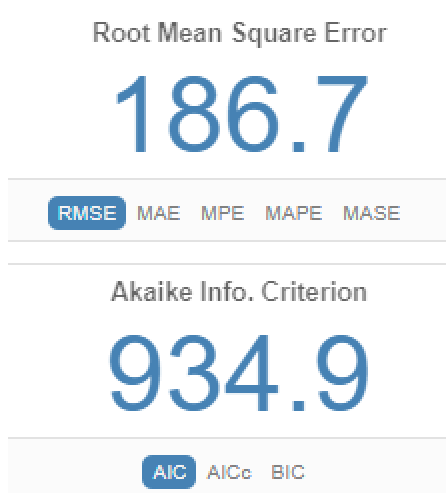
So, I am selecting ARIMA model for predicting next 2 weeks data of EW1 and EW2.

EW1:

ETS

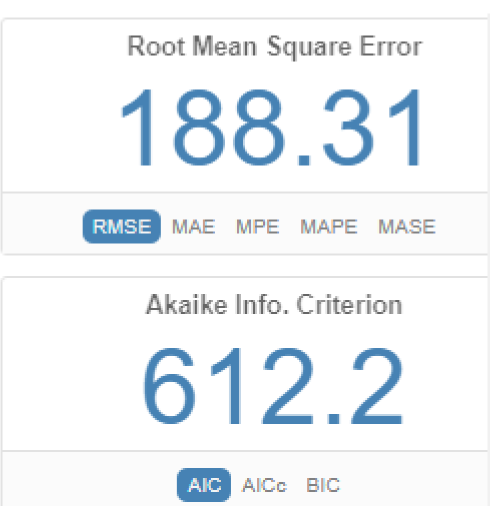


ARIMA

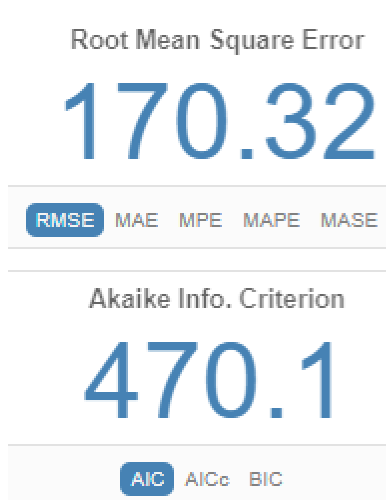


EW2:

ETS



ARIMA



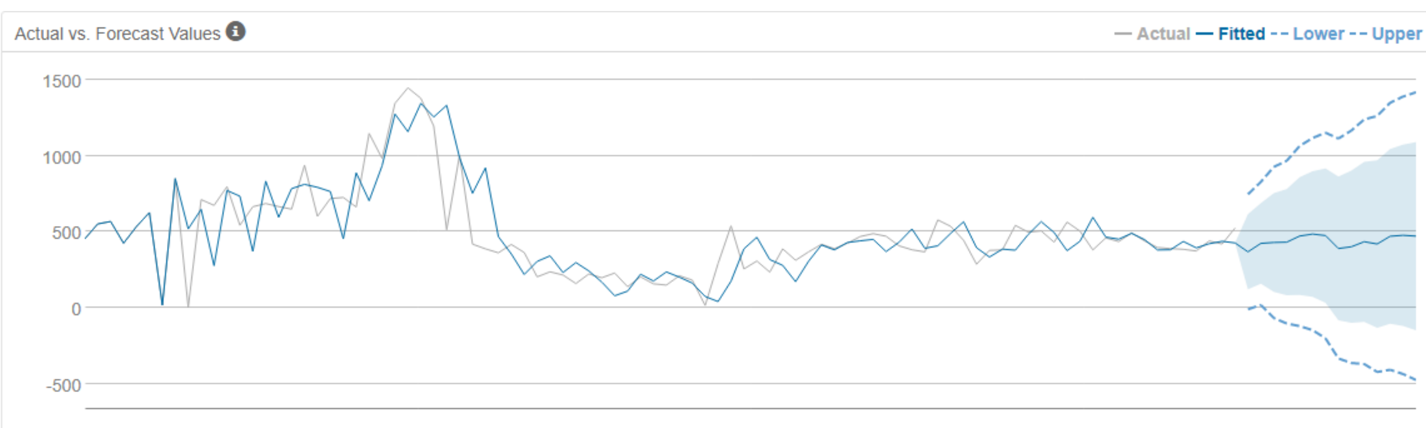
Predicting next 2 weeks data:

I have selected ARIMA to predict next 2 weeks data as it was giving better accuraacy.

Here is the forecast for the next two weeks for both EW1 and EW2 warehouses. Upper and Lower bounds are presented with both 95% and 80% confidence levels.

EW1:

Date	Period	Sub_Period	Forecasted value	95% confidence		80% confidence	
				Forecasted high	Forecasted low	Forecasted high	Forecasted low
09-11-21	13	7	369.529064	747.6067814	-8.548653424	616.7407957	122.3173323
10-11-21	14	1	425.0126402	829.4449929	20.58028751	689.4567417	160.5685388
11-11-21	14	2	430.9946819	927.6718909	-65.68252714	755.7544529	106.2349109
12-11-21	14	3	432.9740342	967.7335894	-101.7855209	782.6345129	83.31355553
13-11-21	14	4	473.2641078	1065.377632	-118.8494166	860.4263312	86.1018843
14-11-21	14	5	486.1902086	1117.599609	-145.2191921	899.0466244	73.33379269
15-11-21	14	6	476.6933538	1152.634069	-199.2473613	918.6672311	34.71947645
16-11-21	14	7	391.201003	1114.063067	-331.6610605	863.855101	-81.45309498
17-11-21	15	1	402.6837237	1166.817582	-361.450135	902.3239981	-96.95655062
18-11-21	15	2	435.9426585	1240.232706	-368.3473888	961.8396532	-89.95433622
19-11-21	15	3	421.0125919	1262.497662	-420.4724781	971.2301048	-129.2049209
20-11-21	15	4	471.7851074	1349.578706	-406.0084909	1045.743491	-102.173276
21-11-21	15	5	478.2129607	1390.353186	-433.9272644	1074.629396	-118.2034749
22-11-21	15	6	472.7444822	1418.451218	-472.962254	1091.10888	-145.6199153



EW2:

Date	Period	Sub_Period	Forecasted value	95% confidence		80% confidence	
				Forecasted high	Forecasted low	Forecasted high	Forecasted low
09-11-21	8	7	1306.213395	1767.447601	844.9791897	1607.798232	1004.628559
10-11-21	9	1	1652.983356	2133.968893	1171.997818	1967.482894	1338.483817
11-11-21	9	2	1743.895599	2244.96442	1242.826778	2071.526891	1416.264308
12-11-21	9	3	1521.790827	2037.224868	1006.356786	1858.815031	1184.766623
13-11-21	9	4	1673.072697	2247.608921	1098.536474	2048.741742	1297.403653
14-11-21	9	5	1661.658381	2260.584856	1062.731907	2053.275353	1270.041409
15-11-21	9	6	1831.523581	2454.432157	1208.615005	2238.821607	1424.225556
16-11-21	9	7	1647.589247	2286.004541	1009.173953	2065.02657	1230.151924
17-11-21	10	1	1806.132108	2473.310442	1138.953774	2242.376573	1369.887644
18-11-21	10	2	1854.22869	2542.072938	1166.384442	2303.985869	1404.471511
19-11-21	10	3	1679.519764	2387.931146	971.1083822	2142.725069	1216.314459
20-11-21	10	4	1884.670658	2610.670761	1158.670555	2359.37661	1409.964707
21-11-21	10	5	1870.831666	2616.885196	1124.778136	2358.649849	1383.013483
22-11-21	10	6	1972.239351	2736.623776	1207.854926	2472.043462	1472.435241