Retail Giant Sales Forecasting

Global Mart Sales and Quantity Forecasting

Sohini Chaudhuri

Business Requirement

1. Store Sales & Operations Manager wants to finalize a plan for 6 months to manage revenue & inventory
2. The store caters to 7 different Market under 3 major Segments
3. Forecast of Sales & Quantity for the next 6 months is required for the 2 most consistently profitable Market segments

Data Understanding & Preparation(2)

- The month wise aggregated Profit (AggProfit) for each subset is calculated

- Coefficient of variation (CoV) = sd(AggProfit)/mean(Profit) is tabulated below

	Africa	APAC	Canada	EMEA	EU	LATAM	US
Consumer	1.319585	0.6321323	1.395312	2.188271	0.6243052	0.6614828	1.012390
Corporate	1.776105	0.6980869	1.552775	4.467102	0.7638072	0.8111217	1.002409
Home Office	1.789996	1.0459784	2.243461	5.880747	1.1165073	1.1756978	1.096147

- The most consistently profitable segments correspond to the ones with least CoV
  1. EU Consumer
  2. APAC Consumer

- Time Series corresponding to EU Consumer Sales, EU Consumer Quantity, APAC Consumer Sales and APAC Consumer Quantity are created for Model

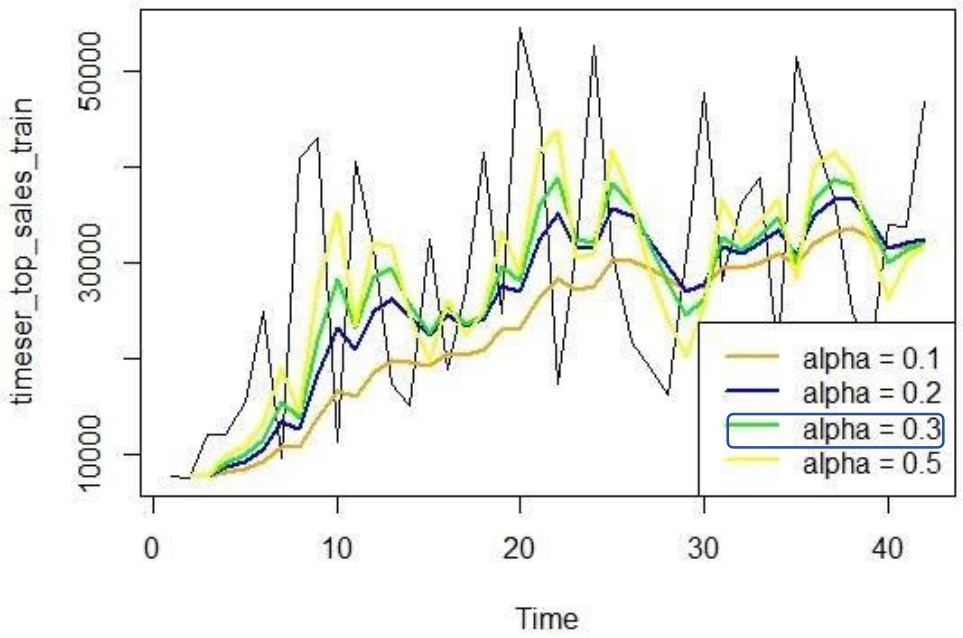
Modelling EU Consumer Sales

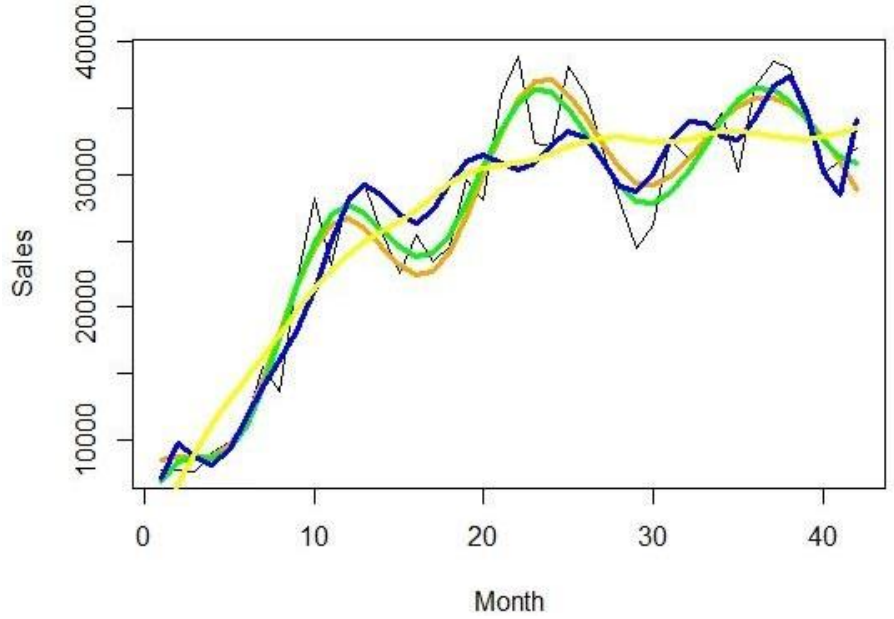
- The timeseries is separated into training and testing data. The last six months data is used to evaluate the forecast made by model
- Both classical decomposition and autoarima methods are used and their accuracies are compared using MAPE
- Holt Winter's Method is used for smoothening the data.
  - Various alpha values are tried
- The smoothened series is modelled for seasonality and trend using multiplicative, sinusoidal-polynomial expression
  - Various degrees of polynomial and coefficient in sine and cosine are tried

Modelling EU Consumer Sales – Classic Decomposition

HoltWinters smoothening . Choosing alpha=0.3(green)

MAPE= 26.82





Best model

Multiplicative Models

lmfit <- lm(Sales~ (sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2), data=smootheddf)

lmfit <- lm(Sales~ (sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3), data=smootheddf)

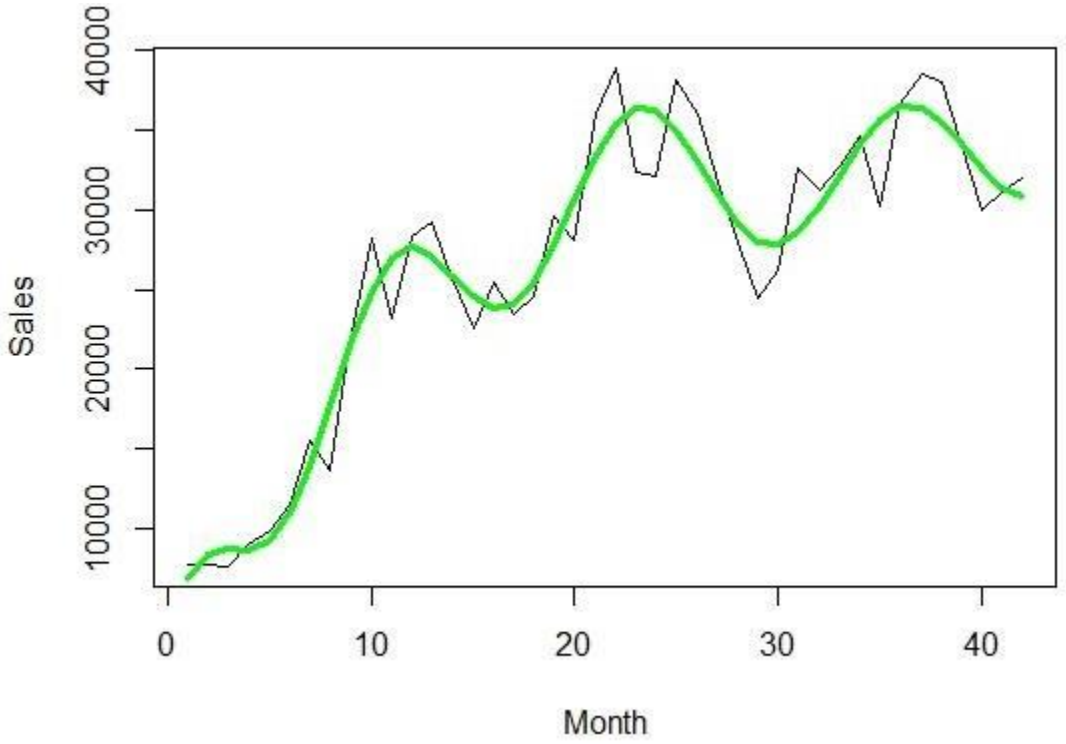
lmfit <- lm(Sales~ (sin(0.8*Month) *poly(Month,3) + cos(0.8*Month)) * poly(Month,3), data=smootheddf)

Additive Model

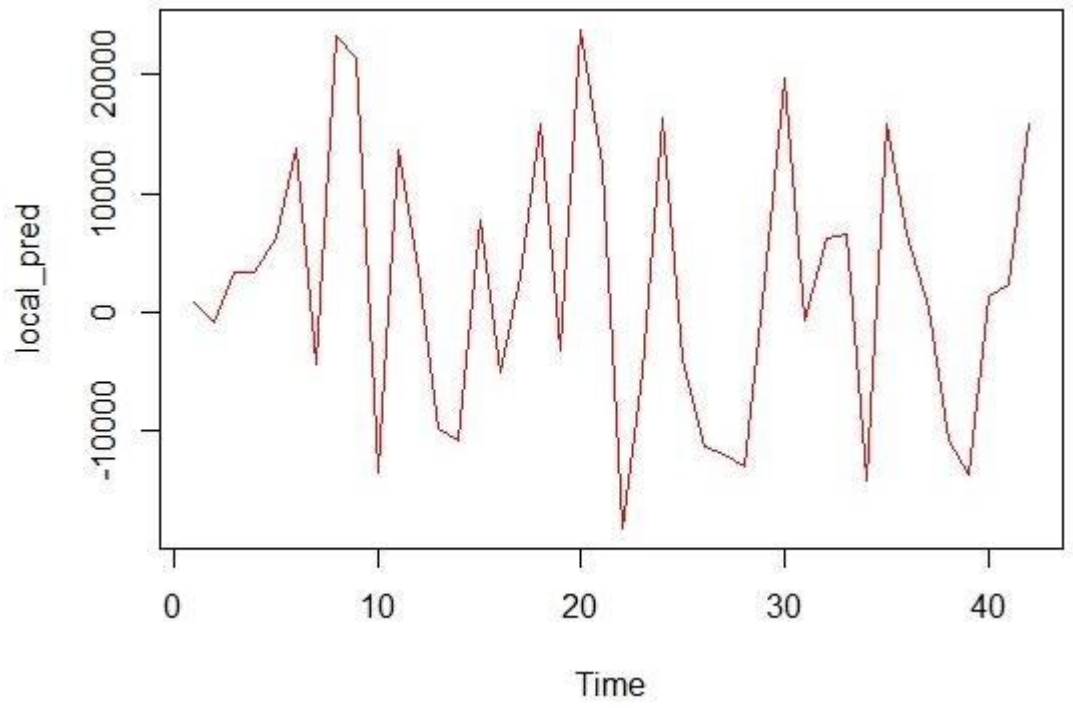
lm(Sales~ (sin(0.8*Month) + poly(Month,3) + cos(0.8*Month)) , data=smootheddf)

Modelling EU Consumer Sales – Classic Decomposition

Global Prediction

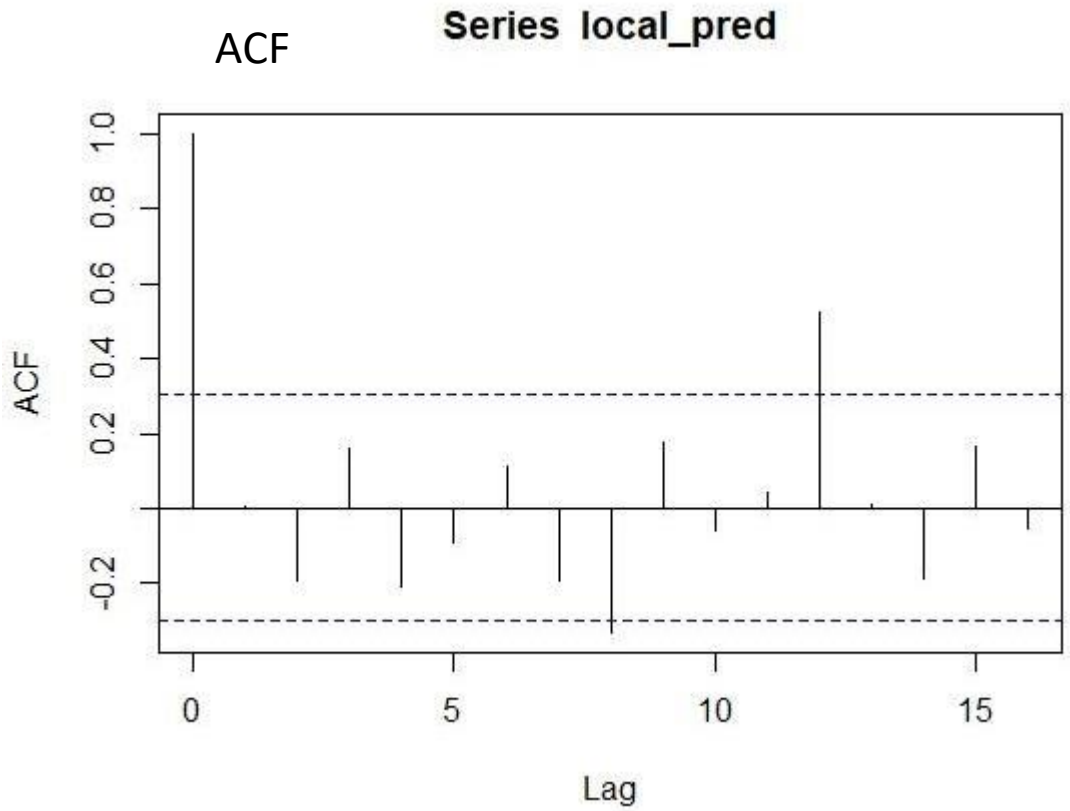


Local Prediction = Actual Time series – Global Prediction

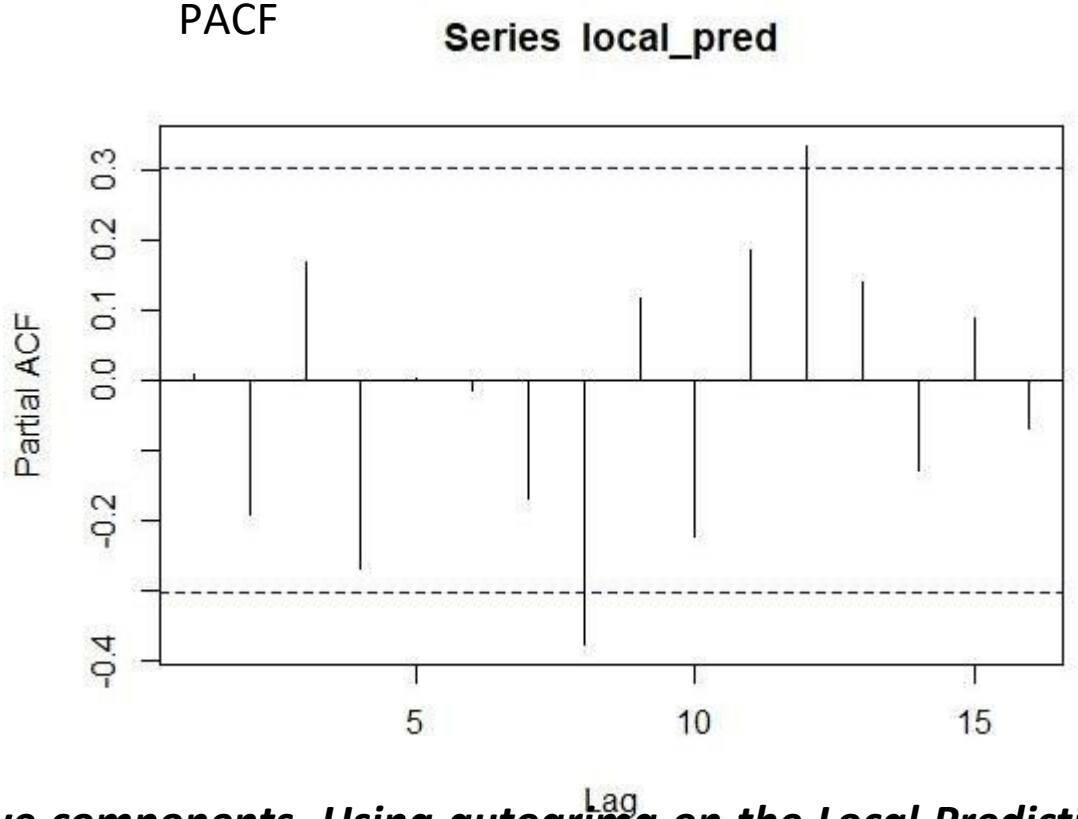


Modelling EU Consumer Sales – Classic Decomposition

ACF



PACF



ACF and PACF plots show that data has no auto regressive components. Using autoarima on the Local Prediction , we find that it is ARIMA(0,0,0). Here also, we can conclude that there is no auto regressive component .Residual series is White Noise as confirmed by KPSS and Dickey Fuller tests

Model Evaluation EU Consumer Sales – Classic Decomposition

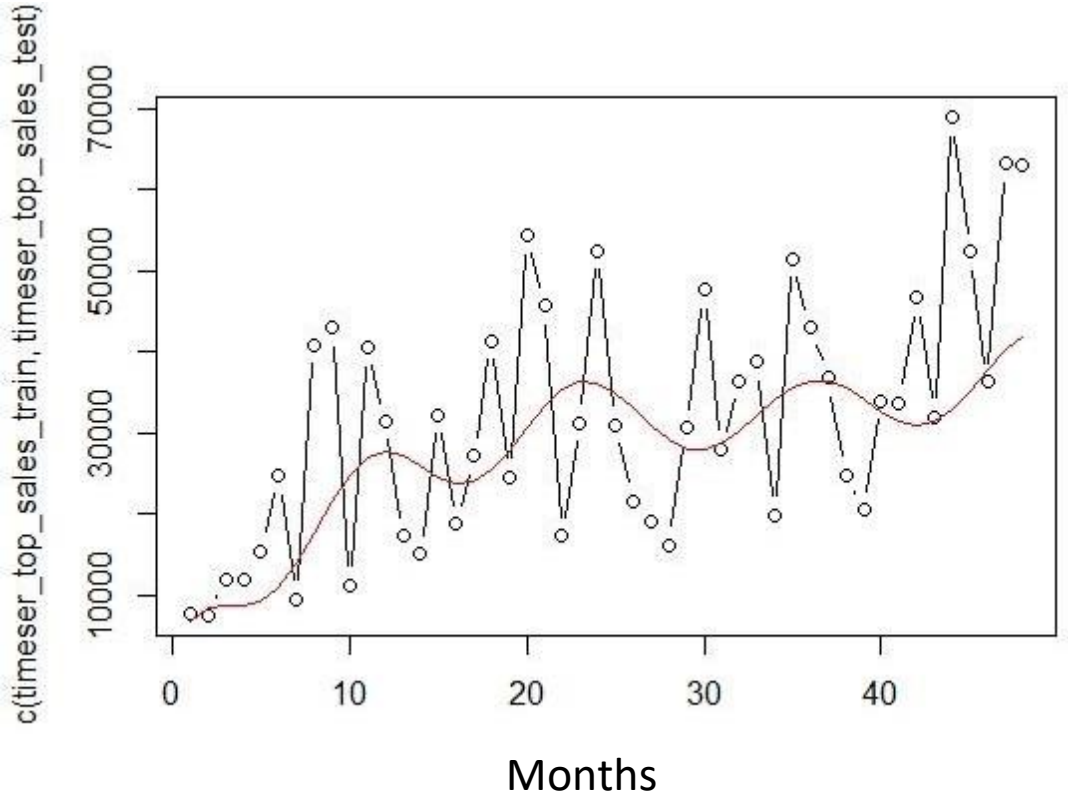


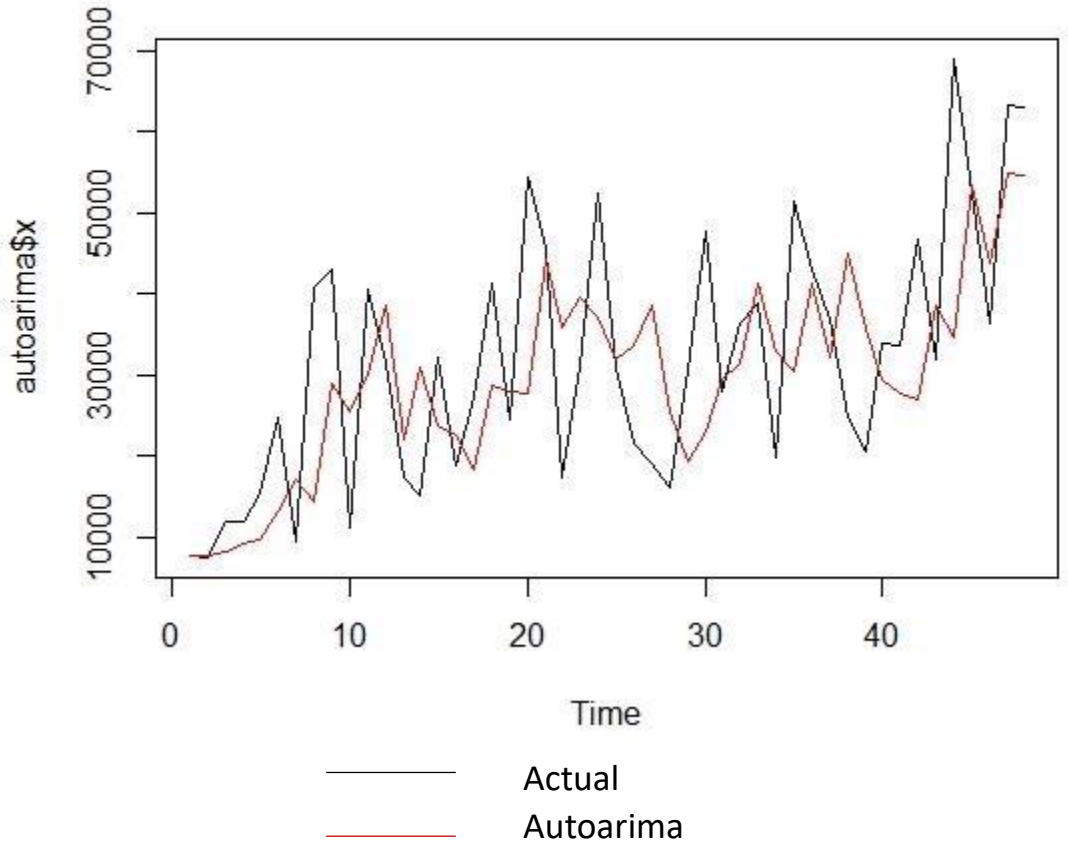
Table with actual and forecasted values

Month	Actual	Forecasted
43	31967.69	31387.85
44	68951.72	32906.64
45	52328.68	35202.61
46	36348.31	37825.45
47	63218.71	40220.44
48	63178.60	41882.25

Forecasting for next 6 months and measuring accuracy gave **MAPE=26.82** which is low and hence a good model with classic decomposition

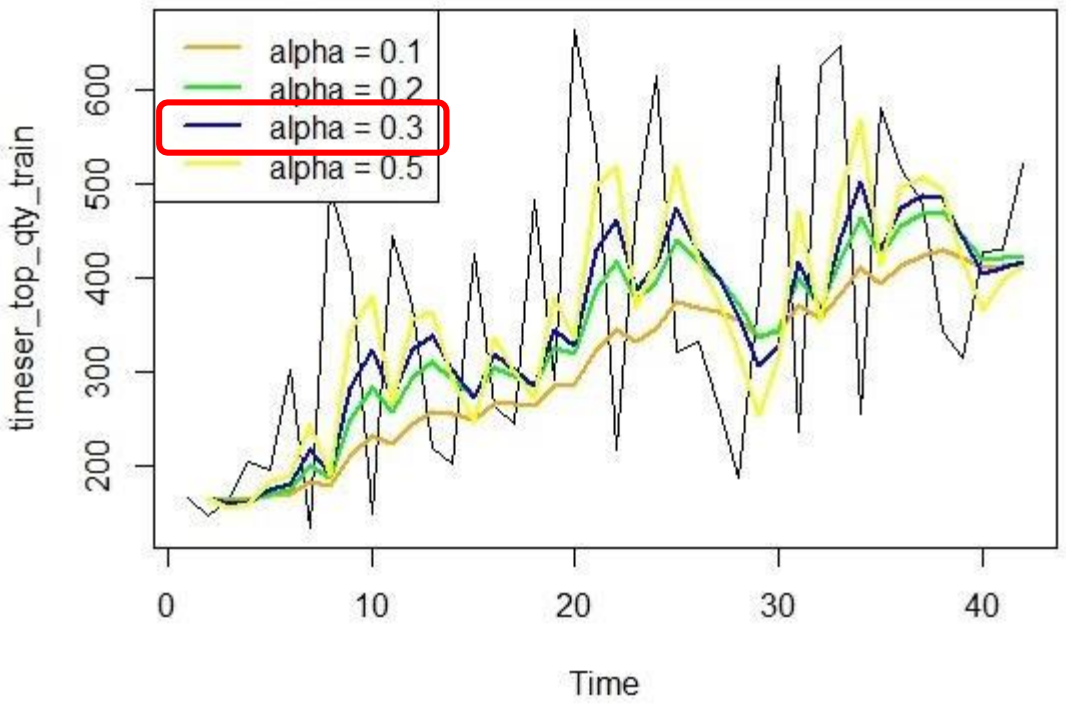
Modelling EU Consumer Sales

- Using Auto Arima Method
  - ARIMA(2,1,0)
  - log likelihood=-512.5
  - AIC=1031 AICc=1031.56 BIC=1036.55
- MAPE= 36.81

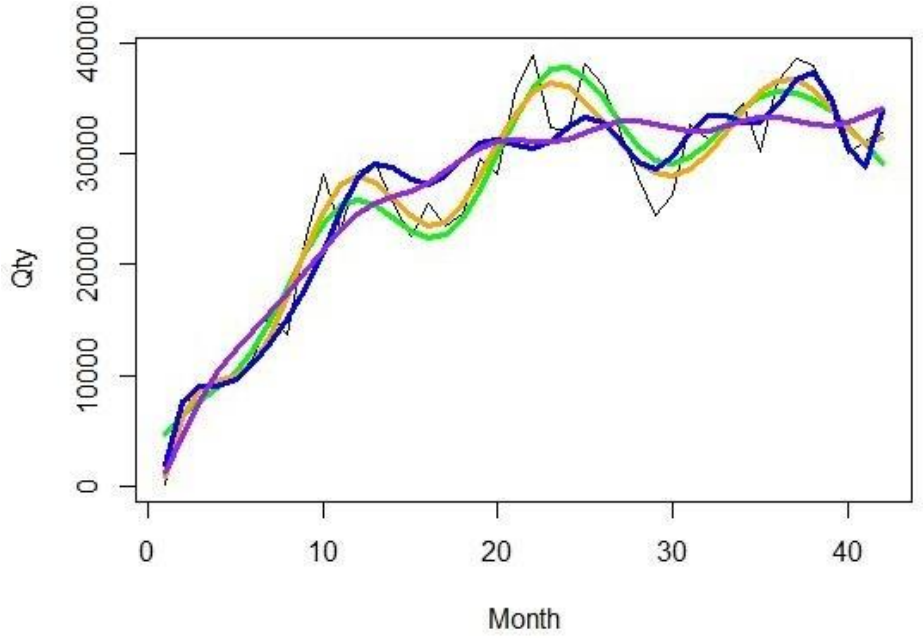


MAPE using classical decomposition method using model sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3) was 26.82 and a better model

Modelling EU Consumer Quantity – Classic Decomposition



Holts Winter smoothening, Choosing alpha = 0.3



lm(Qty~ sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2) …

lm(Qty~ (sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3)…

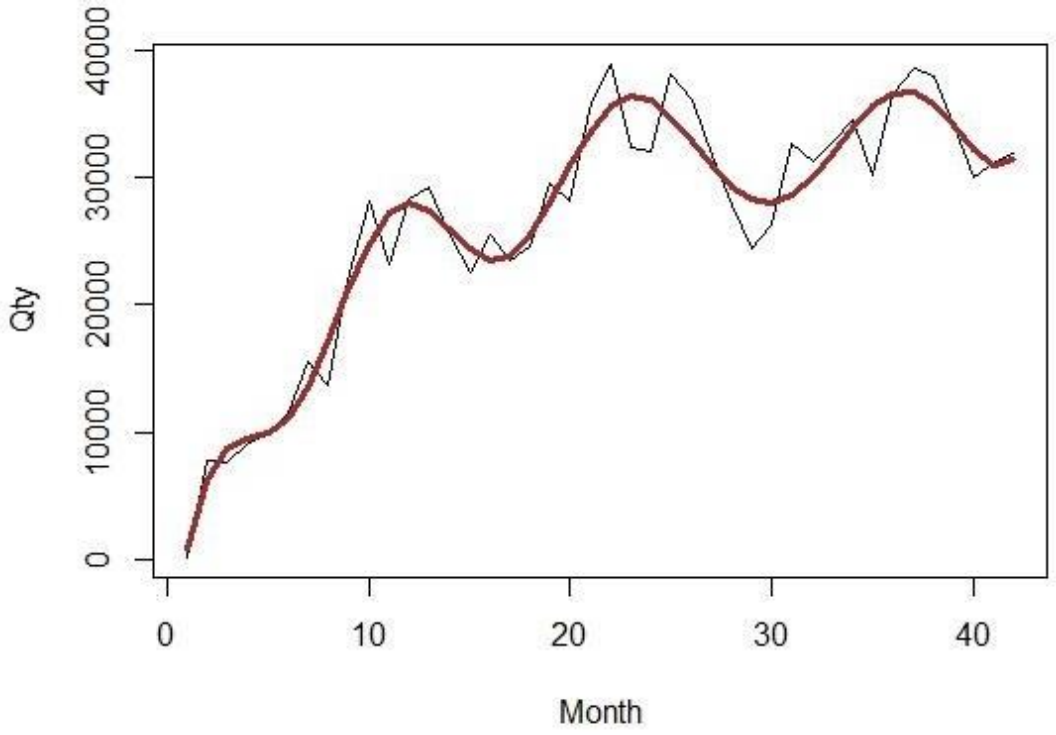
lm(Qty~ sin(0.8*Month) *poly(Month,2) + cos(0.8*Month)) * poly(Month,2) …

lm(Qty~ (sin(0.8*Month) + poly(Month,3) + cos(0.8*Month))…

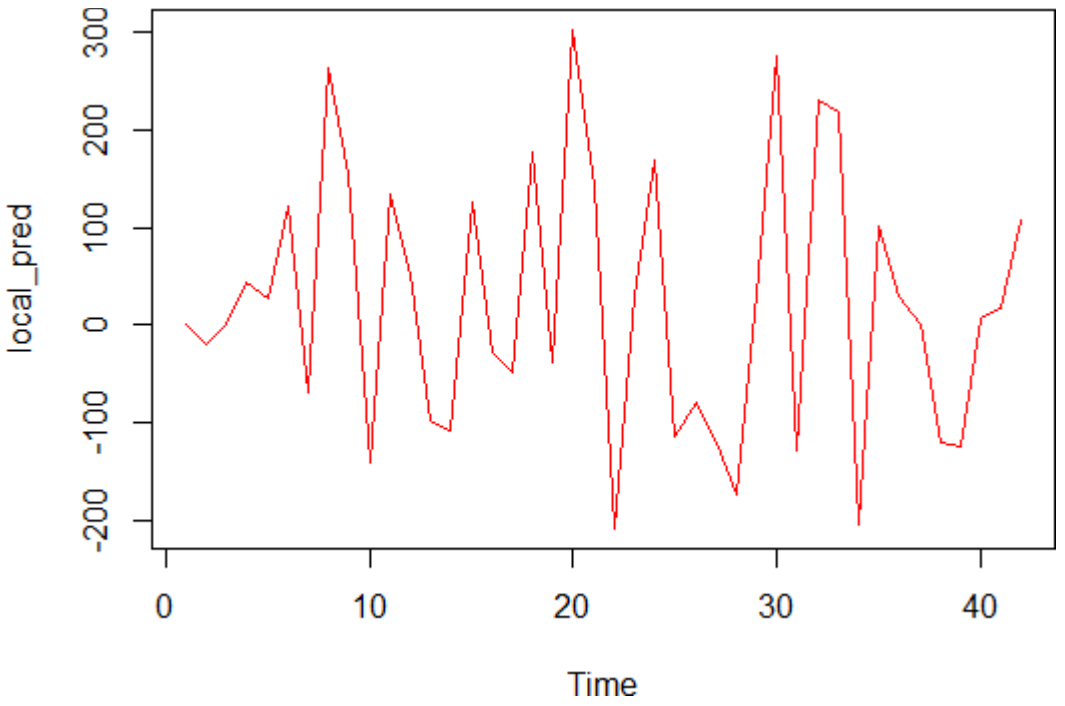
Best Model

Modelling EU Consumer Quantity – Classic Decomposition

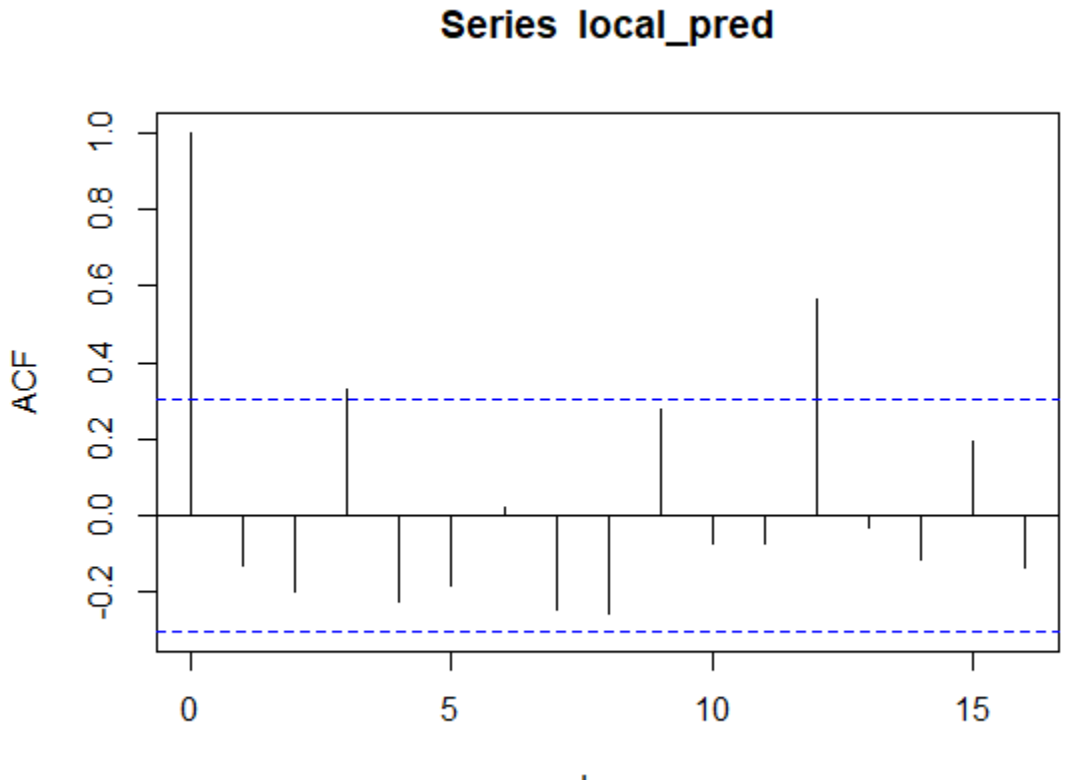
Global Prediction

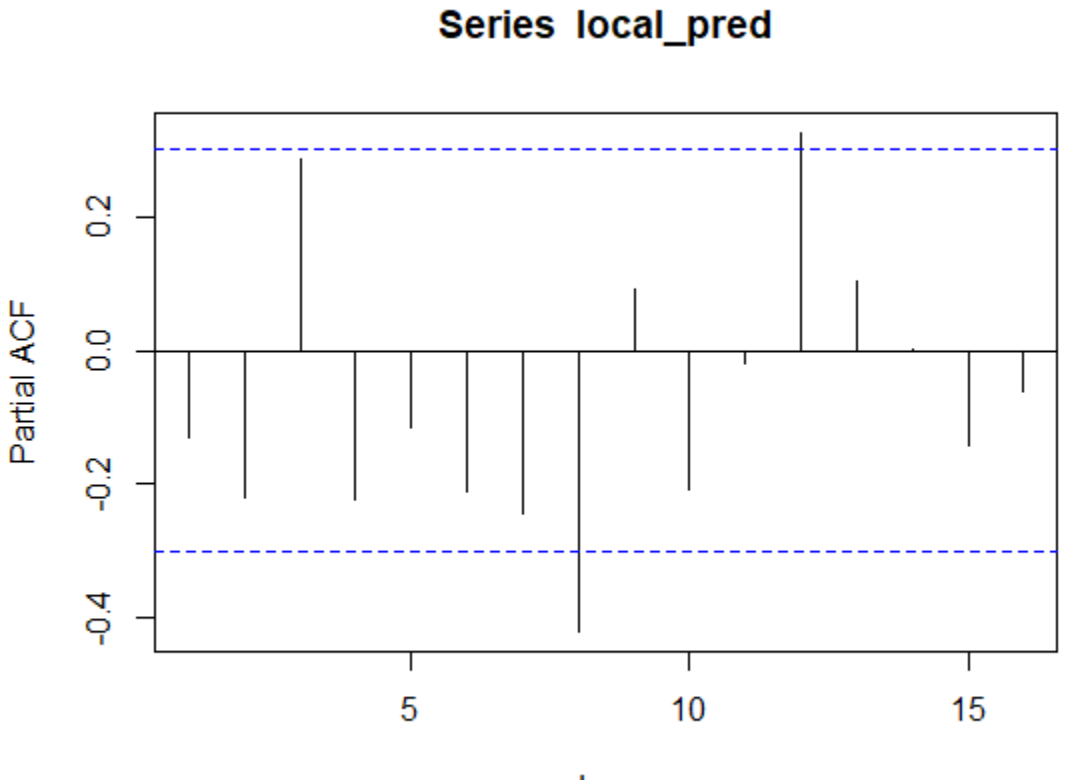


Local Prediction = Actual Time series – Global Prediction



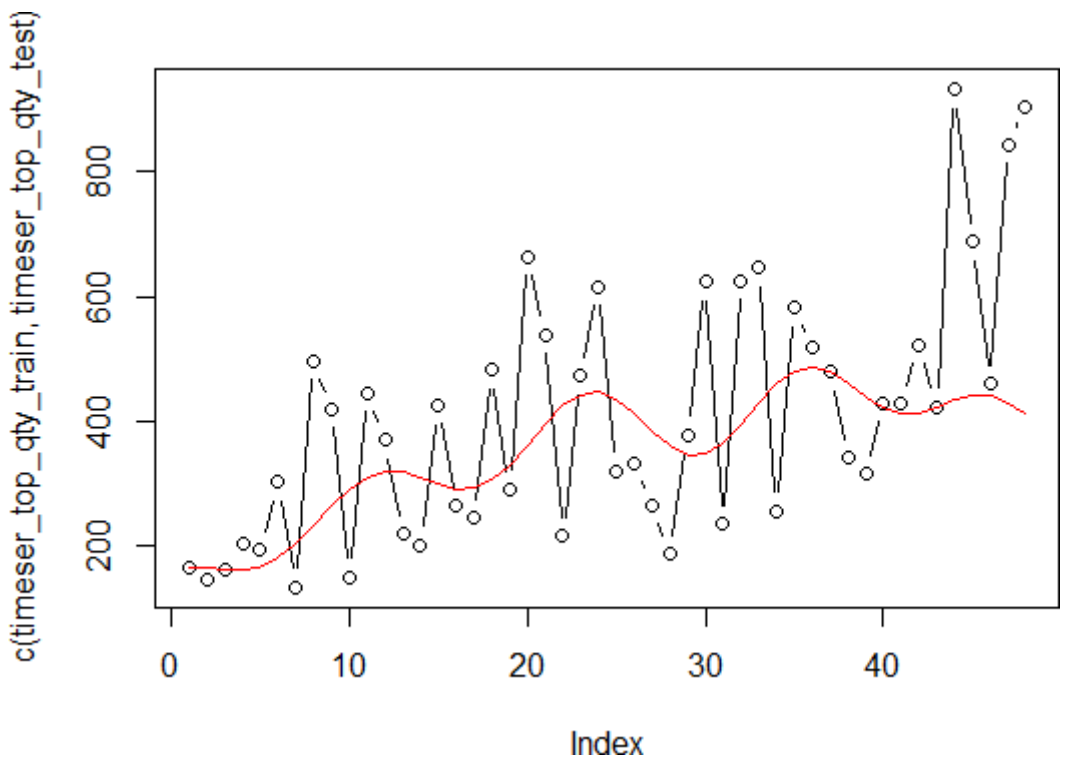
Modelling EU Consumer Quantity – Classic Decomposition





ACF and PACF plots show that data has no auto regressive components. Using auto.arima on the Local Prediction , we find that it is ARIMA(0,0,0). Here also, we can conclude that there is no auto regressive component.Residual series is White Noise as confirmed by KPSS and Dickey Fuller tests

Model Evaluation EU Consumer Quantity- Classic Decomposition



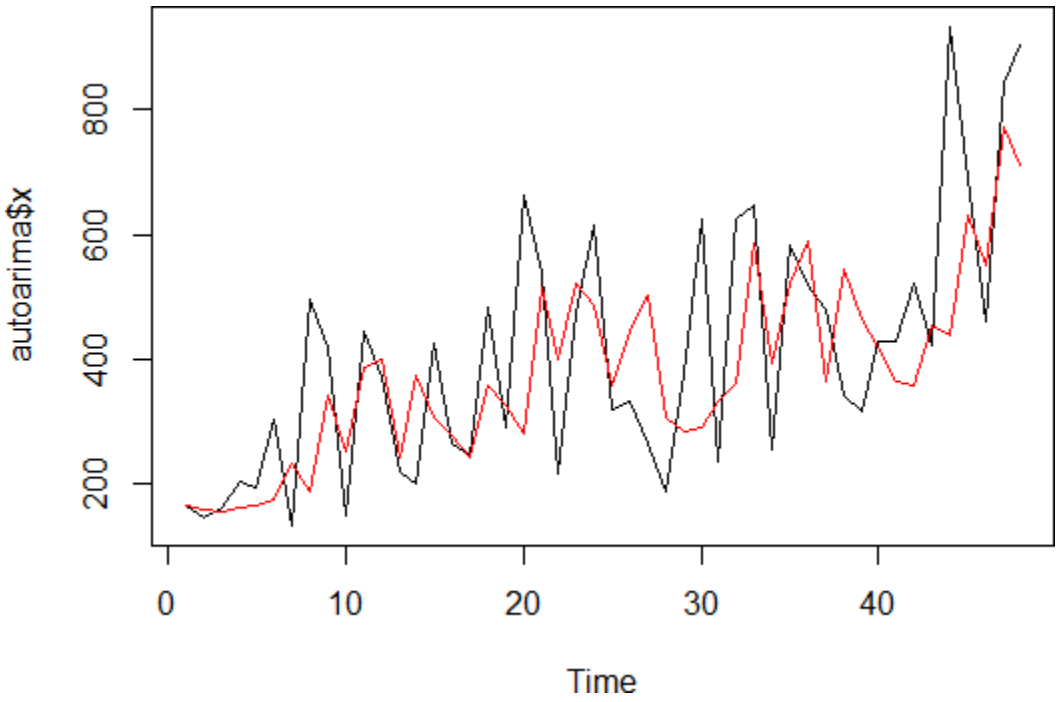
Forecasting gave MAPE of 32.79

Table with actual and forecasted Quantity values

Month	Actual	Forecasted
43	423	421.9609
44	932	434.0731
45	688	442.0687
46	459	440.9714
47	843	429.8014
48	905	412.5824

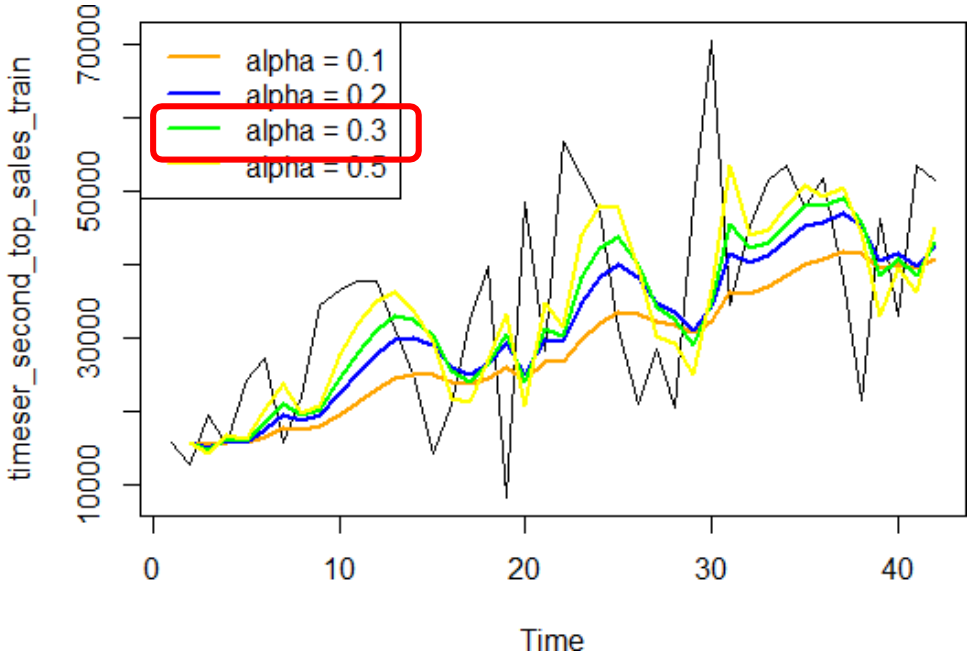
Modelling EU Consumer Quantity - ARIMA

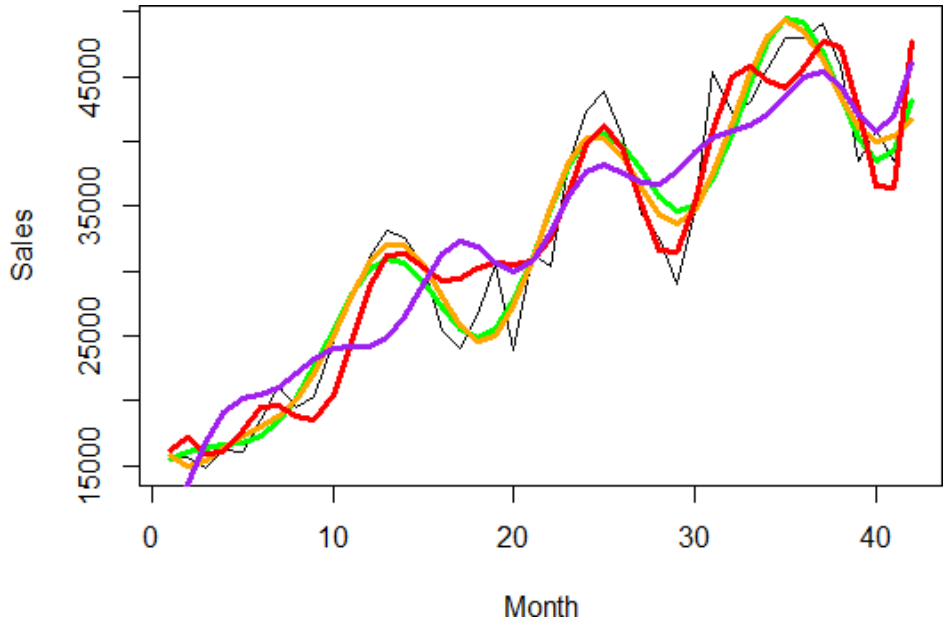
- Using AutoArima Method
  - ARIMA(2,1,0)
  - log likelihood=-261.9
  - AIC=529.8 AICc=530.44 BIC=534.95
  - MAPE=31.56



In this case MAPE is 31.56,it is comparable with the model with classical decomposition
Model used for forecast : sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3)

Modelling APAC Consumer Sales – Classic Decomposition





Holts Winter smoothening, Choosing alpha = 0.3

Best Model

lm(Sales~ (sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2)…

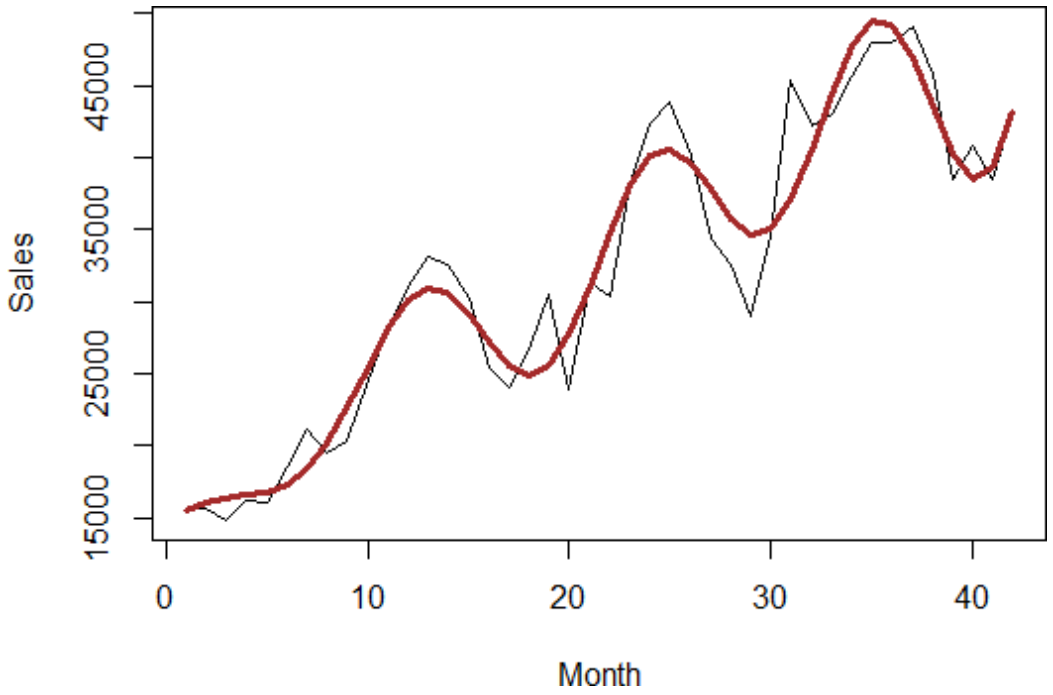
lm(Sales~ (sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3)…

lm(Sales~ (sin(0.8*Month) *poly(Month,3) + cos(0.8*Month)) * poly(Month,3)…

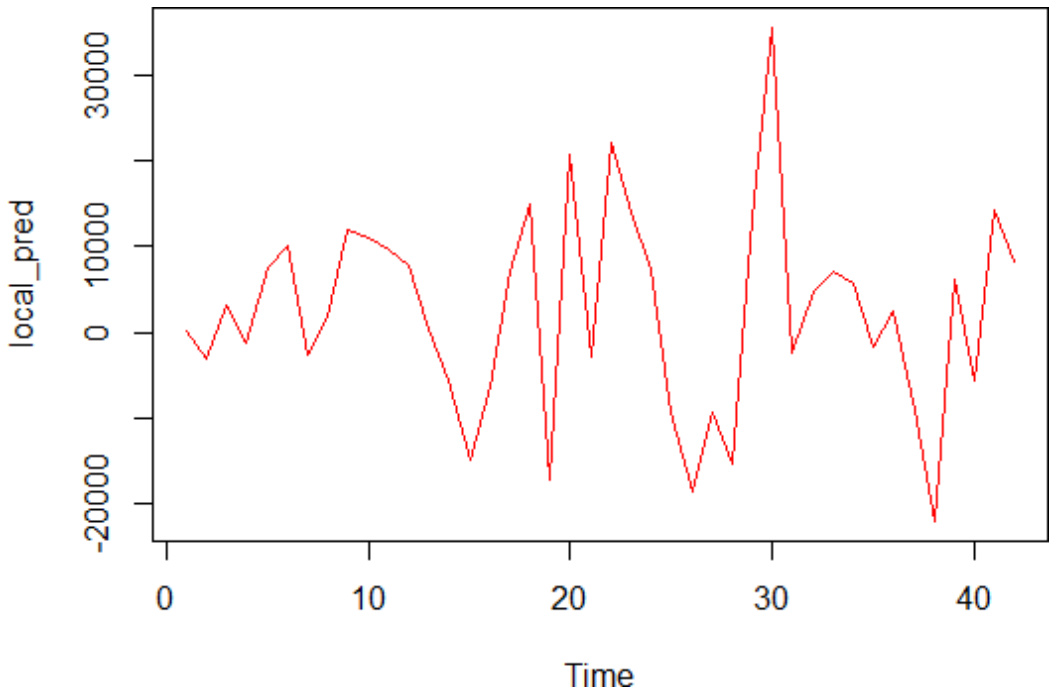
lm(Sales~ (sin(0.8*Month) *poly(Month,2) + cos(0.8*Month)) * poly(Month,2)…

Modelling APAC Consumer Sales – Classic Decomposition

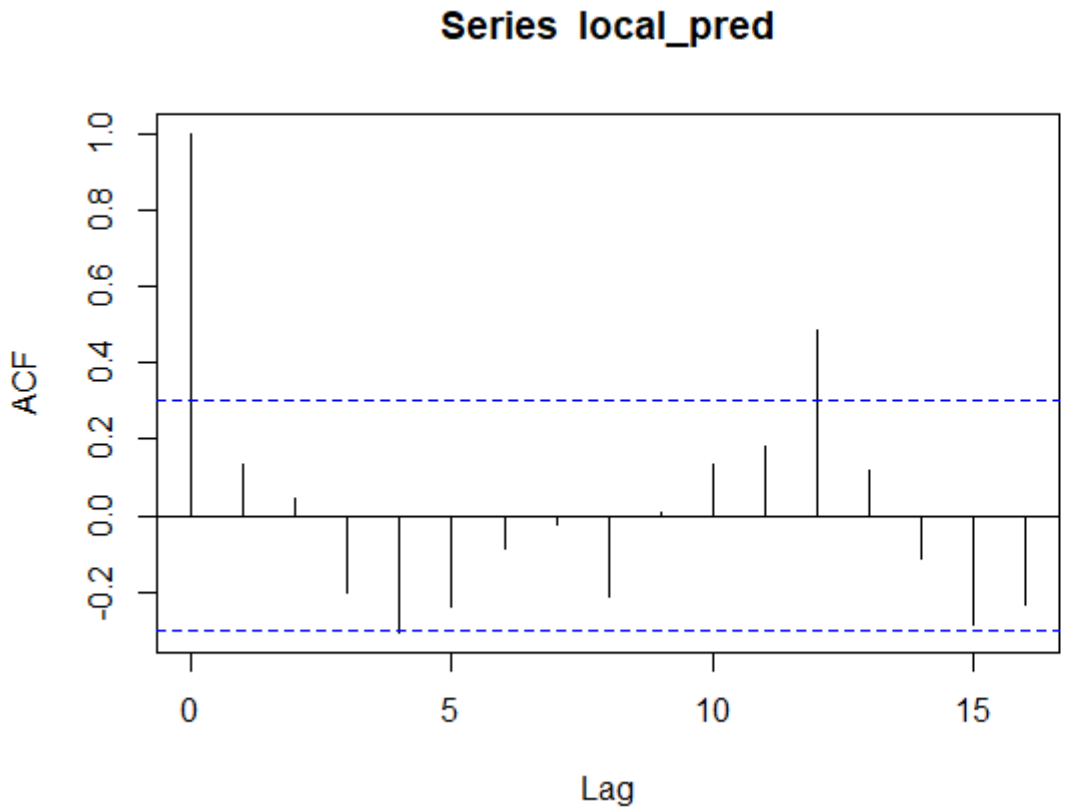
Global Prediction

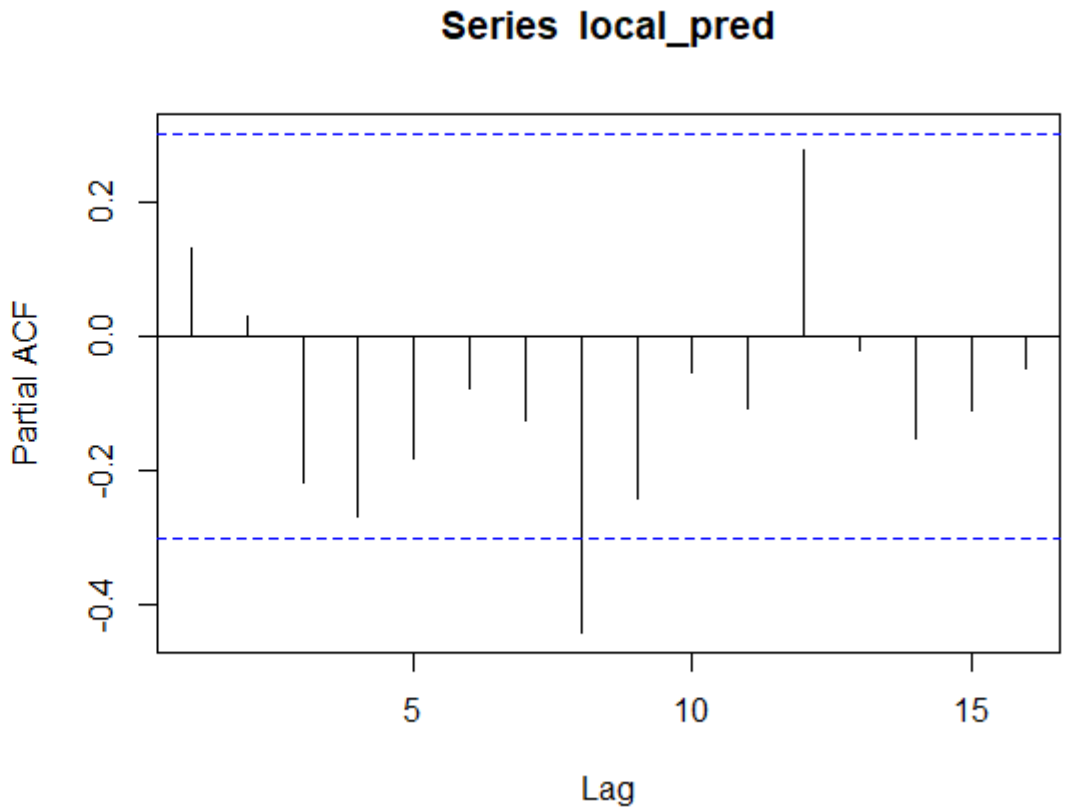


Local Prediction = Actual Time series – Global Prediction



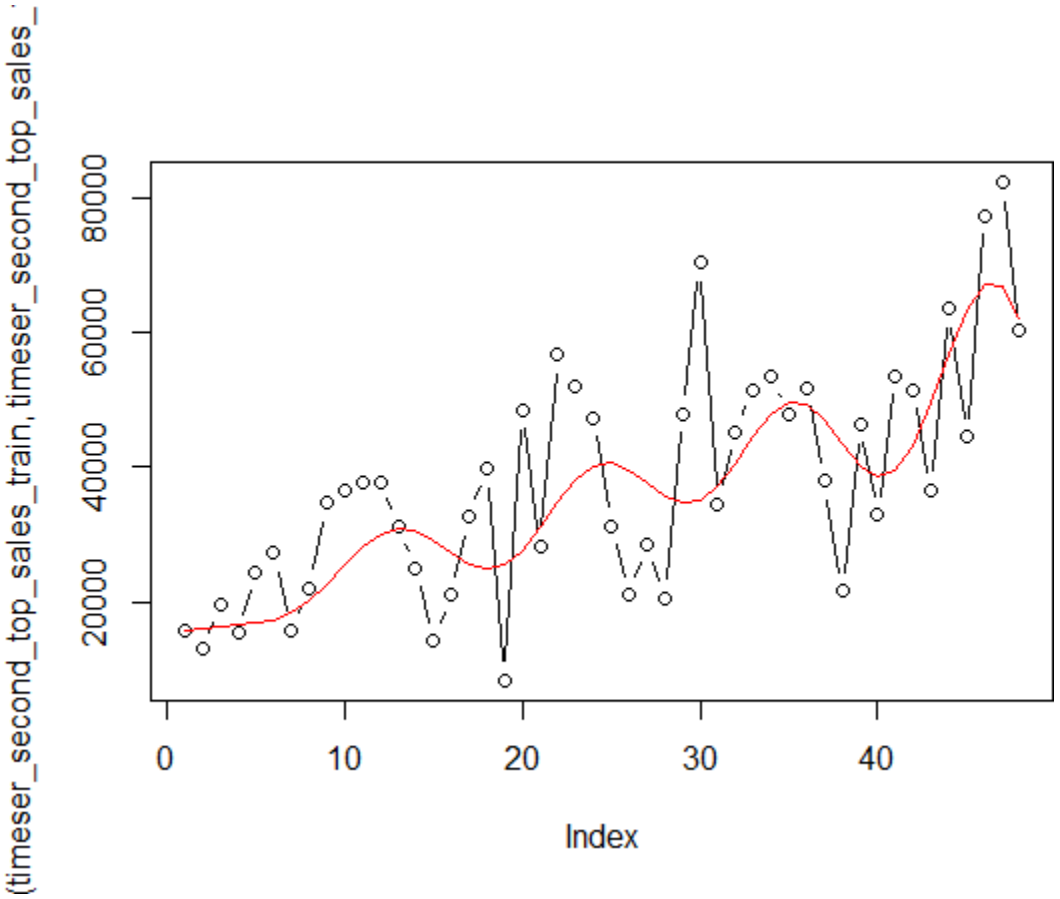
Modelling APAC Consumer Sales – Classic Decomposition





ACF and PACF plots show that data has no auto regressive components.Using auto.arima on the Local Prediction , we find that it is ARIMA(0,0,0) Here also, we can conclude that there is no auto regressive component. Residual series is White Noise as confirmed by KPSS and Dickey Fuller tests

Model Evaluation APAC Consumer Sales - Classic Decomposition



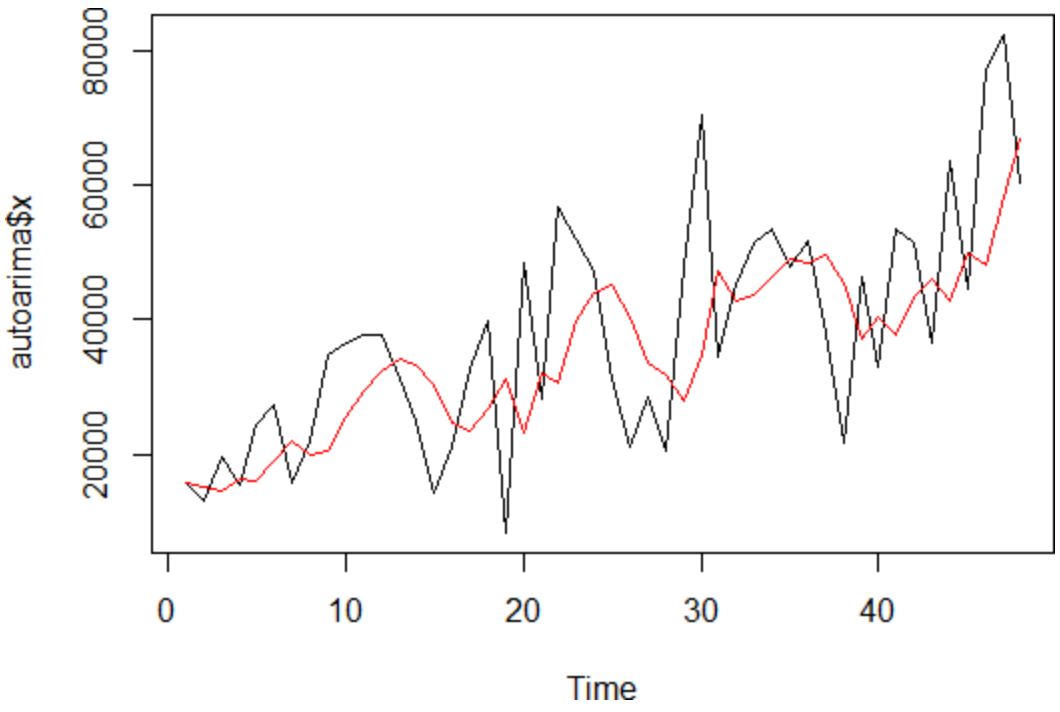
Forecasting gave a MAPE of 20.622

Table with actual and forecasted Quantity values

Month	Actual	Forecasted
43	36524.30	49502.70
44	63521.77	56841.03
45	44477.27	63338.85
46	77379.83	67101.51
47	82286.36	66816.26
48	60292.13	62218.53

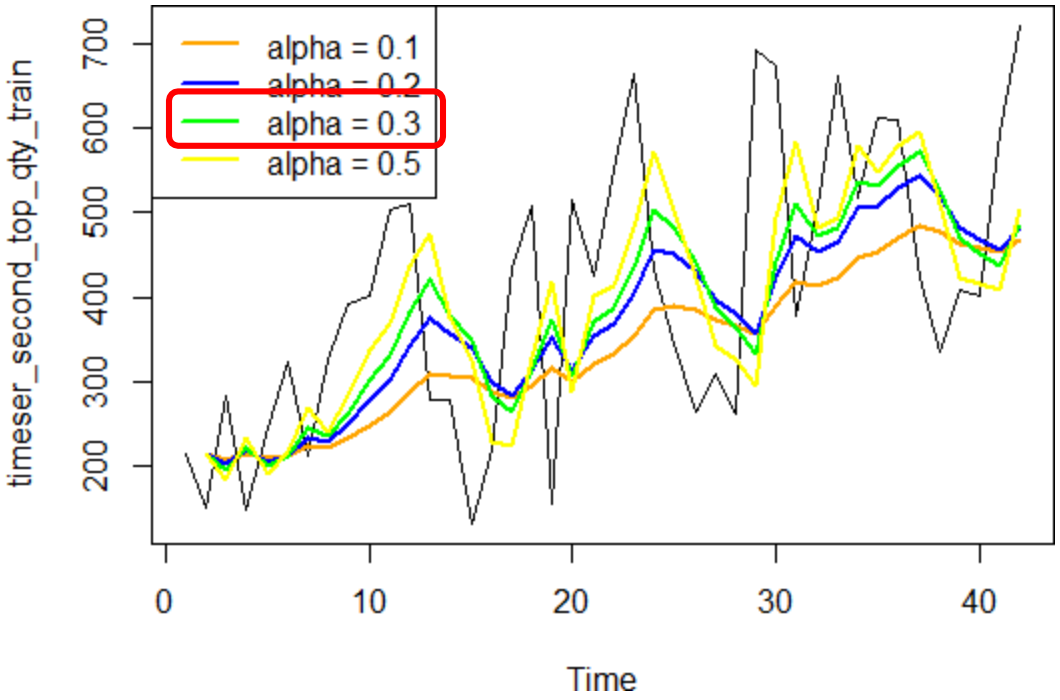
Modelling APAC Consumer Sales -ARIMA

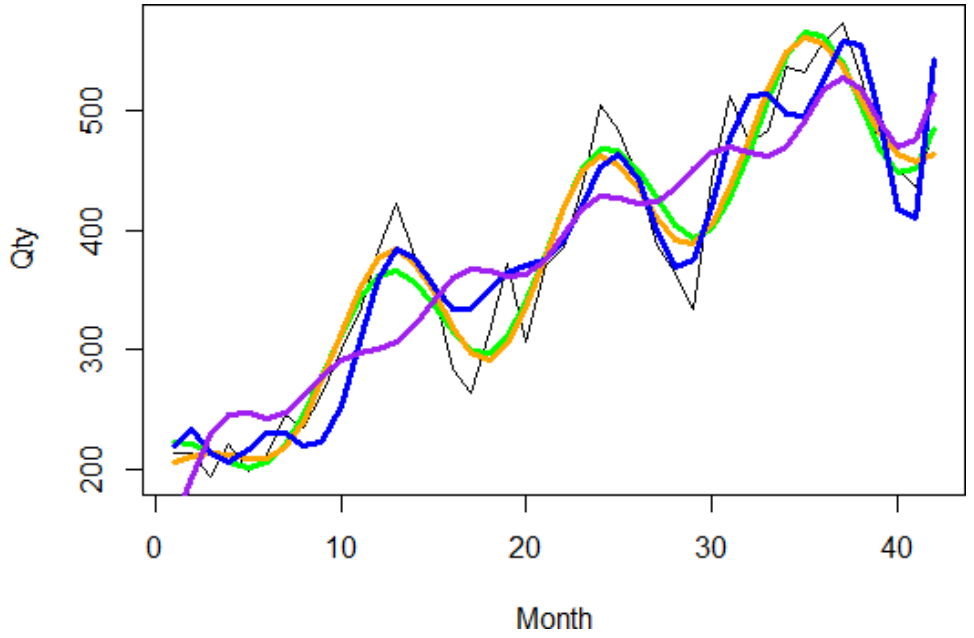
- Using the AutoArima Method
  - ARIMA(0,1,1)
  - log likelihood=-515.41
  - AIC=1034.83 AICc=1035.1
  - BIC=1038.53
  - MAPE= 45.53



In this case, classical decomposition method gives a much better MAPE of 20.622
Model type used : sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2)

Modelling APAC Consumer Quantity – Classic Decomposition





Holts Winter smoothening .Choosing Alpha=0.3

Best Model

lm(Sales~ (sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2)…

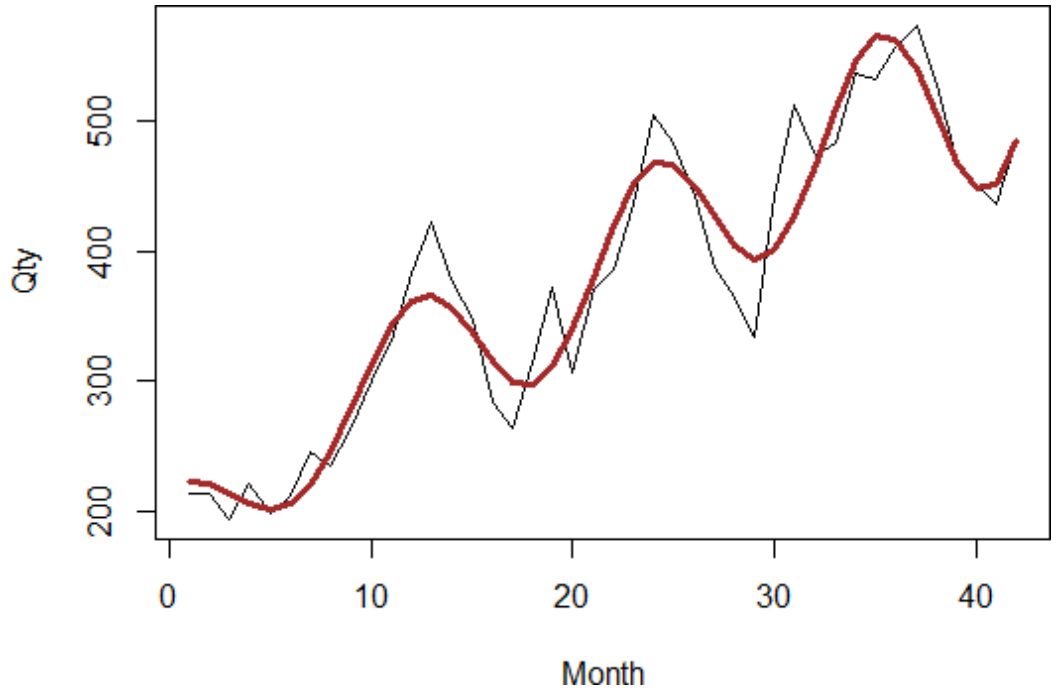
lm(Sales~ (sin(0.5*Month) *poly(Month,3) + cos(0.5*Month)) * poly(Month,3)…

lm(Sales~ (sin(0.8*Month) *poly(Month,2) + cos(0.8*Month)) * poly(Month,2)…

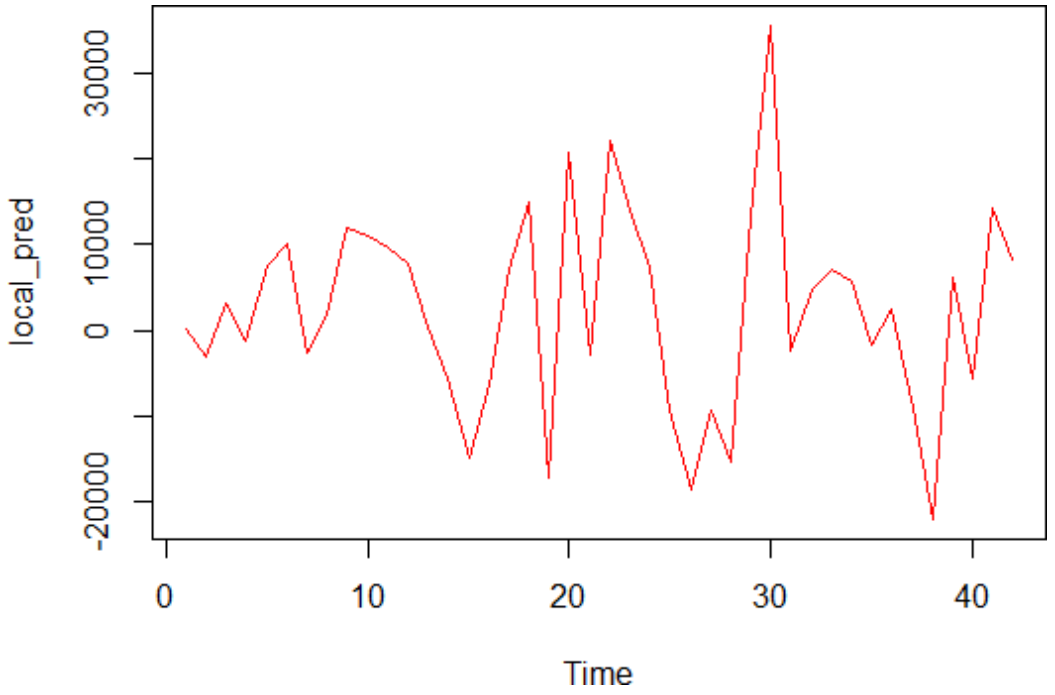
lm(Sales~ (sin(0.8*Month) *poly(Month,3) + cos(0.8*Month)) * poly(Month,3)…

Modelling APAC Consumer Quantity – Classic Decomposition

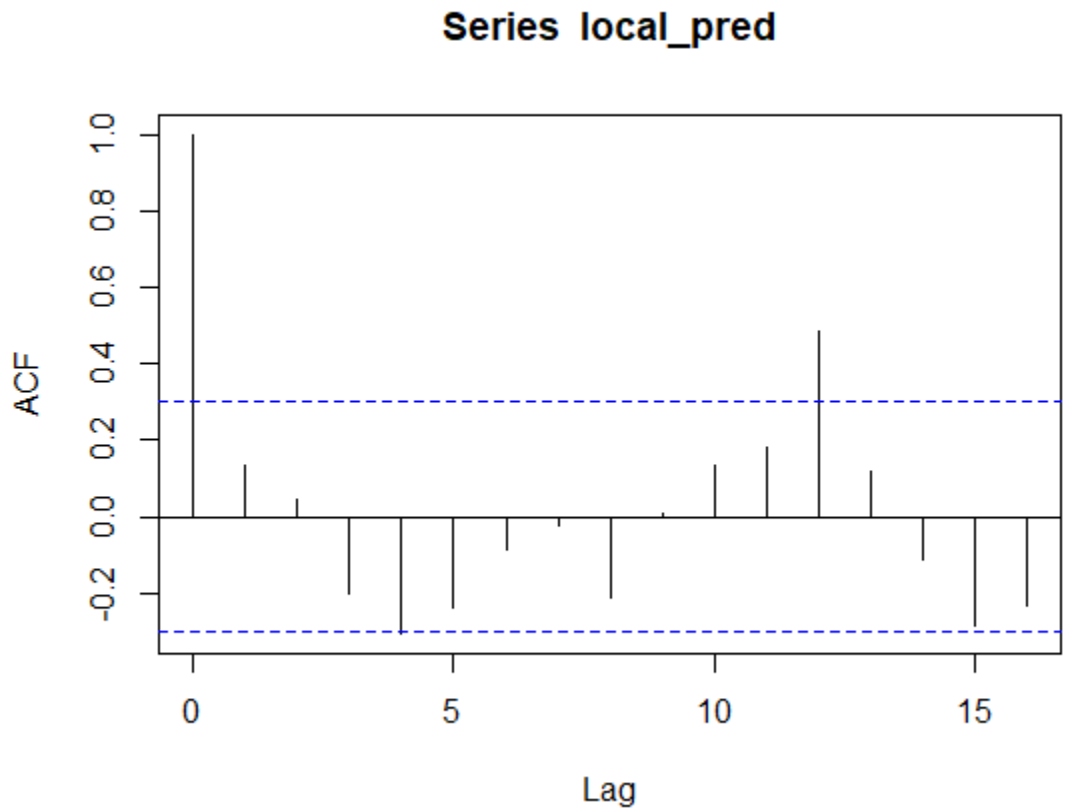
Global Prediction

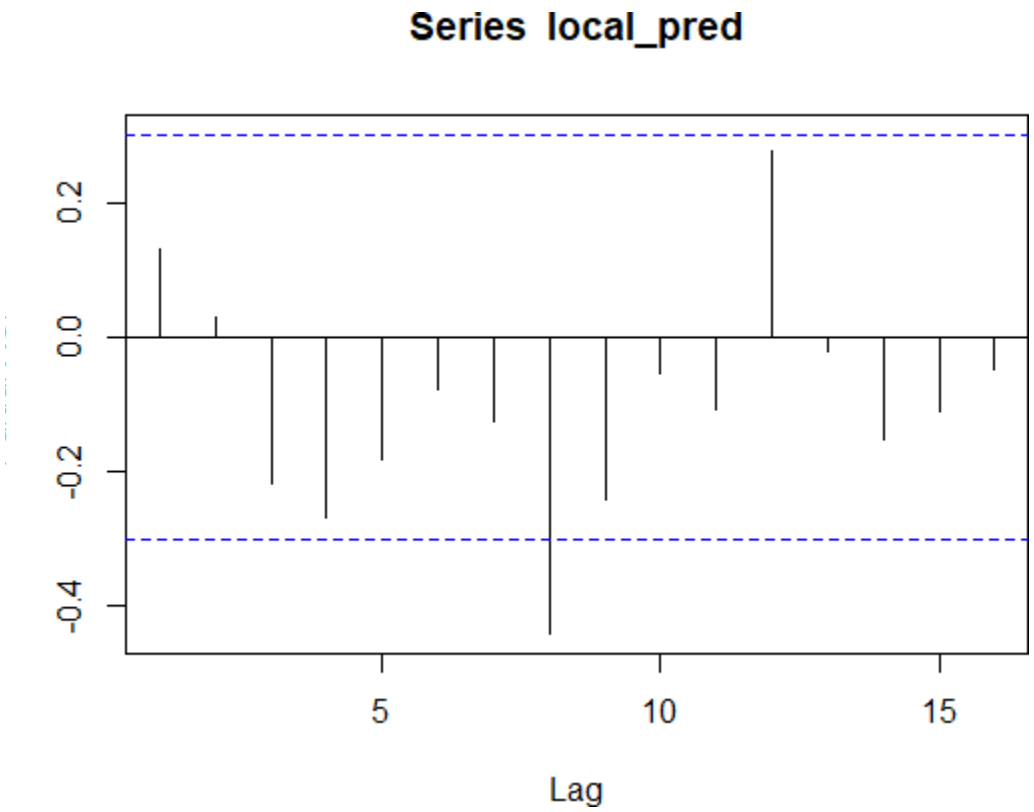


Local Prediction = Actual Time series – Global Prediction



Modelling APAC Consumer Quantity – Classic Decomposition





ACF and PACF plots show that data has no auto regressive components. Using auto.arima on the Local Prediction , we find that it is ARIMA(0,0,0). Here also, we can conclude that there is no auto regressive component.Residual series is White Noise as confirmed by KPSS and Dickey Fuller tests

Model Evaluation – Classic Decomposition

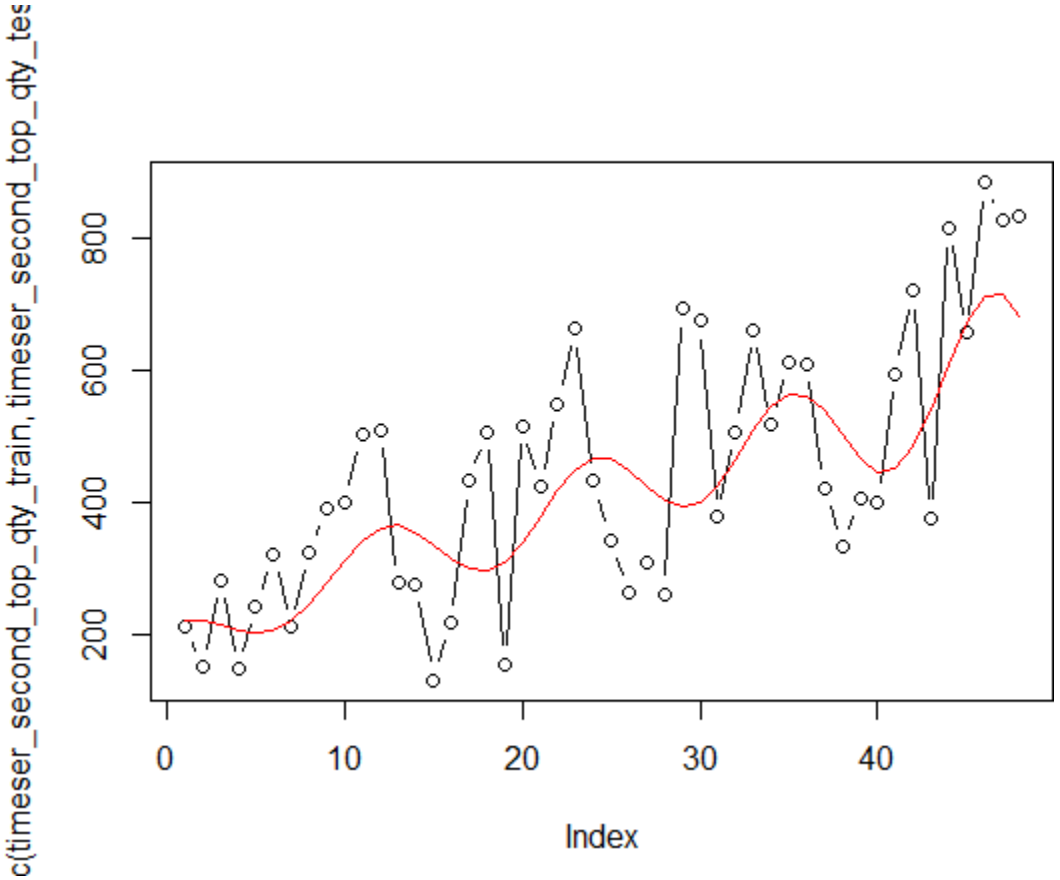


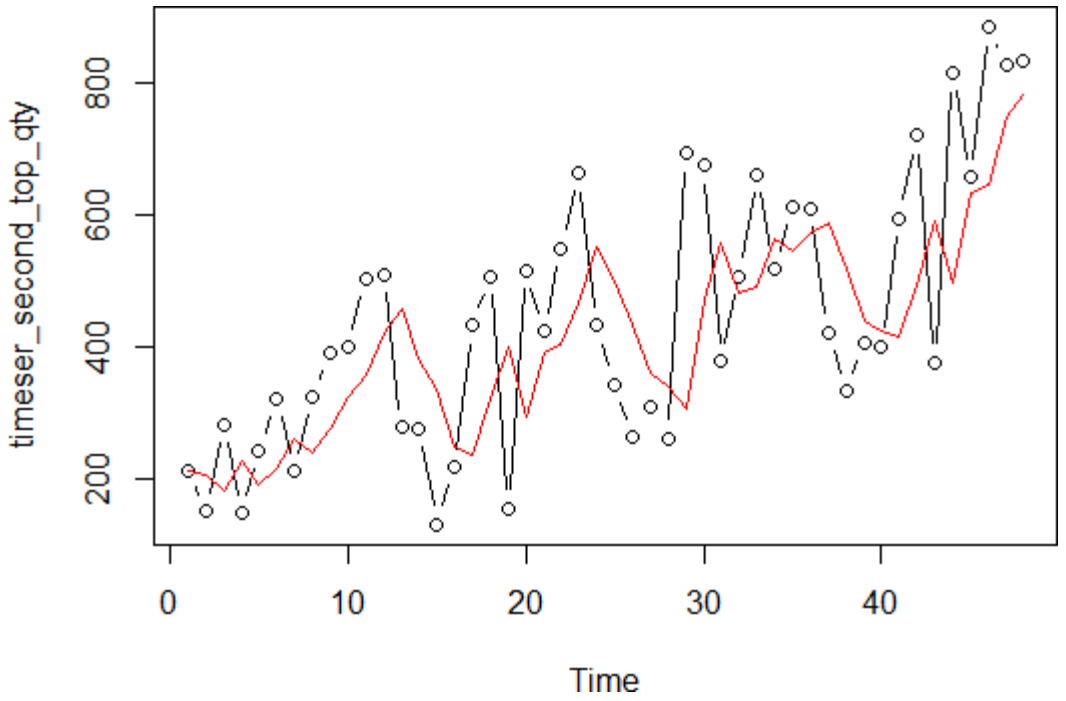
Table with actual and forecasted Quantity values

Month	Actual	Forecasted
43	377	541.0292
44	816	609.7819
45	658	672.7776
46	885	712.4482
47	829	716.3389
48	833	681.3039

Forecasting gave a MAPE of 20.38 which is low and hence accurate forecast

Modelling APAC Consumer Quantity -ARIMA

- Modelling Using Auto Arima
  - ARIMA(0,1,1)
  - log likelihood=-303.76
  - AIC=611.53 AICc=611.8 BIC=615.23
  - MAPE= 33.918



In this case classical decomposition gives a better forecast as its MAPE = 20.38
Model is of type sin(0.5*Month) *poly(Month,2) + cos(0.5*Month)) * poly(Month,2)

SUMMARY - Sales

Top 2 Profitable(consistent) Market & Segment

1. **EU - Consumer**
2. **APAC - Consumer**

Sales Forecast for next 6 months

Month	EU -Consumer Sales		Month	APAC - Consumer Sales	
	Forecast[MAPE=26.83]	Actual		Forecast[MAPE=20.62]	Actual
43	31387.85	31967.69	43	49502.7	36524.3
44	32906.64	68951.72	44	56841.03	63521.77
45	35202.61	52328.68	45	63338.85	44477.27
46	37825.45	36348.31	46	67101.51	77379.83
47	40220.44	63218.71	47	66816.26	82286.36
48	41882.25	63178.6	48	62218.53	60292.13
49	42514.79	Not Available	49	54296.07	Not Available
50	42150.4	Not Available	50	45143.74	Not Available
51	41186.26	Not Available	51	37476.69	Not Available
52	40313.85	Not Available	52	33910.15	Not Available
53	40346.2	Not Available	53	36191.84	Not Available
54	41979.81	Not Available	54	44604.11	Not Available

SUMMARY - Quantity

Quantity Forecast for next 6 months

Month	EU -Consumer Quantity					Month	APAC -Consumer Quantity				
				Total						Total	
	Technology	Office Supplies	Furniture	Forecast	Actual		Technology	Office Supplies	Furniture	Forecast	Actual
43	86	286	51	422	423	43	88	202	87	541	377
44	180	645	107	434	932	44	135	544	137	610	816
45	122	473	93	442	688	45	136	366	156	673	658
46	115	296	48	441	459	46	177	491	217	712	885
47	130	573	140	430	843	47	187	453	189	716	829
48	138	602	165	413	905	48	175	491	167	681	833
49	119	239	40	398	NA	49	123	369	123	615	NA
50	118	237	39	394	NA	50	107	322	107	536	NA
51	123	247	41	411	NA	51	94	281	94	468	NA
52	135	270	45	450	NA	52	87	260	87	433	NA
53	151	303	50	505	NA	53	89	268	89	447	NA
54	169	337	56	562	NA	54	103	308	103	513	NA

Consumer products Quantity forecast is obtained from the model for EU & APAC Markets

Forecast under 3 product categories is also mentioned based on the below approximate % percentage observed

EU

- Technology – 30% of Forecast
- Office Supplies – 60% of Forecast
- Furniture – 10% of Forecast

APAC

- Technology – 20% of Forecast
- Office Supplies – 60% of Forecast
- Furniture – 20% of Forecast