TIME SERIES ANALYSIS

CASE STUDY

ANUPAM MAJHI

HARI NYSHADAM

LIJO THOMAS

KIRAN VENKAT

INTRODUCTION

Objective:

"Global Mart" is an online store super giant having worldwide operations. It takes orders and delivers across the globe and deals with all the major product categories - consumer, corporate & home office. As a sales/operations manager, you want to finalise the plan for the next 6 months. So, you want to forecast the sales and the demand for the next 6 months, that would help you manage the revenue and inventory accordingly.

The store caters to 7 different market segments and in 3 major categories. But not all of these 21 market buckets are important from the store's point of view. So you need to find out 2 most profitable (and consistent) segment from these 21 and forecast the sales and demand for these segments.

Methodology:

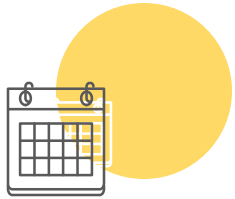
Our methodology follows the **Cross Industry Standard Process for Data Mining** (CRISP–DM) framework. The process and the steps involved can easily be visualized in the following info-graphic

DATA CLEANING

AND

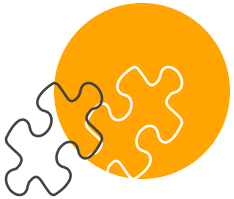
PREPARATION

DATA CLEANING



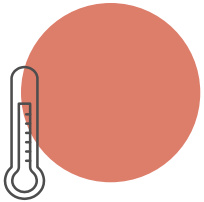
DATE - TIME

Order date was converted to a date format. Floor_date function from the *lubridate* library was used to floor all values to the first day of every month. This was essential for Monthly aggregation later.



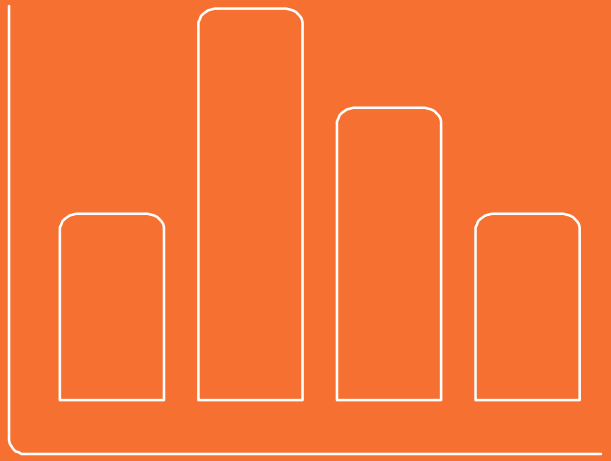
MISSING VALUES

There were no missing values in our variables of interest, 'Sales', 'Quantity' and 'Profit'.



OUTLIERS

Outliers were detected in the dataset. The outliers were capped with the appropriate values.



EXPLORATORY DATA ANALYSIS

MOST PROFITABLE AND MOST CONSISTENT SEGEMENTS

	cv	profit	avgprofit
APAC.Consumer	62.0569333942351	220645.2304	4596.77563333333
EU.Consumer	62.491325405943	186191.1375	3878.98203125
APAC.Corporate	69.7534073470841	127551.5728	2657.32443333333
LATAM.Consumer	66.1482844072086	120632.93196	2513.1860825
EU.Corporate	69.7542384017353	119052.4095	2480.25853125
US.Consumer	91.6813800539831	113159.7638	2357.49507916667
US.Corporate	89.3998263129078	82274.7001	1714.05625208333
APAC.Home Office	104.507495596202	82028.6538	1708.9302875
EU.Home Office	112.392072078045	60349.6545	1257.28446875
LATAM.Corporate	81.1121749333764	57875.42136	1205.737945
US.Home Office	105.208297182131	55571.7616	1157.74503333333
Africa.Consumer	132.099709028884	47685.099	993.4395625
LATAM.Home Office	117.569780549626	43135.13376	898.64862
EMEA.Consumer	218.827087593379	25532.574	531.928625
Africa.Home Office	178.999565401693	20412.567	425.2618125
Africa.Corporate	173.741524079629	19526.205	406.7959375
EMEA.Corporate	449.613514808496	12376.254	257.838625
Canada.Consumer	139.531215242653	9677.7	230.421428571429
EMEA.Home Office	588.07467161514	5866.263	122.2138125
Canada.Corporate	155.277513803186	5036.46	148.131176470588
Canada.Home Office	224.346068905675	3103.23	124.1292

Showing 1 to 21 of 21 entries

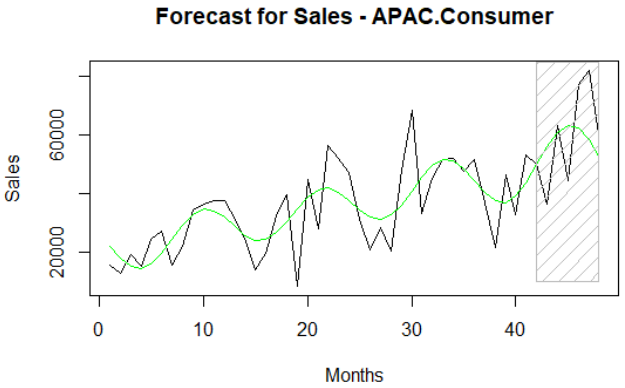
The two most profitable segments were **APAC Consumer** and **EU Consumer**. They were also the most consistent segments, according to their coefficient of variation (COV)***** values.

**COV = 100 * Std Dev(x) / Mean(x)*

MODELING THE DATA

FORECASTING FOR APAC CONSUMER - SALES

CLASSICAL DECOMPOSITION



Local Component : ARIMA(0,0,0)

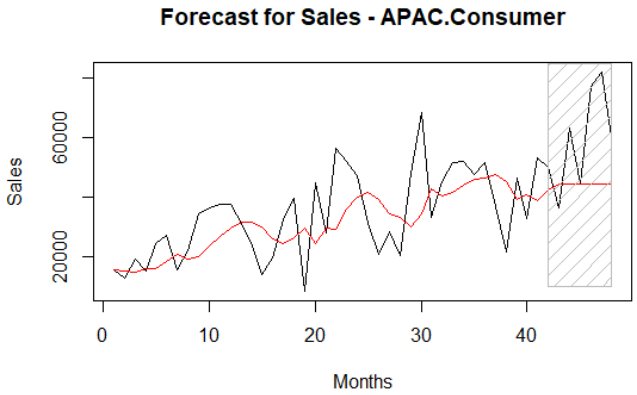
Global Component: LM fit

lmfit <- lm(Sales ~ sin(0.5*Months) * poly(Months,1) + cos(0.5*Months) * poly(Months,1), data=apacs_smoothdf)

Final Model = Local + Global

MAPE = **26.73**

AUTO ARIMA

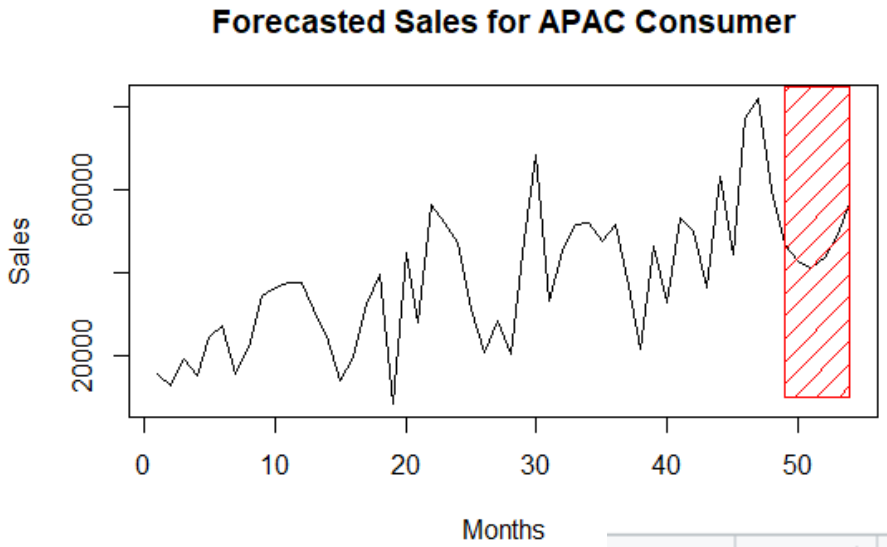


Model : ARIMA(0,1,1)

Coefficients:	-0.7518
SE	0.1443
sigma^2	168697297:
log likelihood	=-446.43 ,
AIC	=896.86,
AICc	=897.18,
BIC	=900.29

MAPE = **27.77**

FORECASTS FOR THE NEXT SIX MONTHS



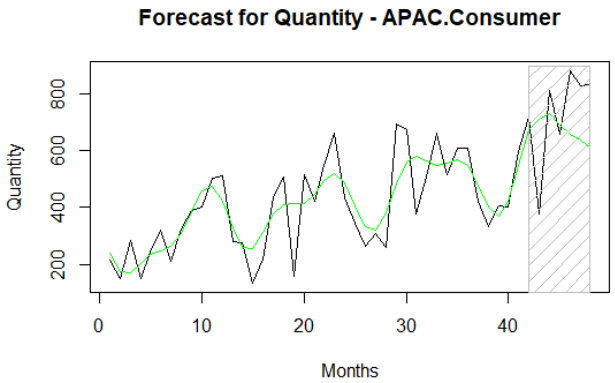
These are the Forecasted Sales for the Next Six Months for the APAC Consumer segment

	Months	Sales
1	49	47094.19
2	50	42817.90
3	51	41457.73
4	52	43687.04
5	53	49256.84
6	54	57025.24

**All Forecasts done using Classical Decomposition*

FORECASTING FOR APAC CONSUMER - QUANTITY

CLASSICAL DECOMPOSITION



Local Component : ARIMA(1,0,1)

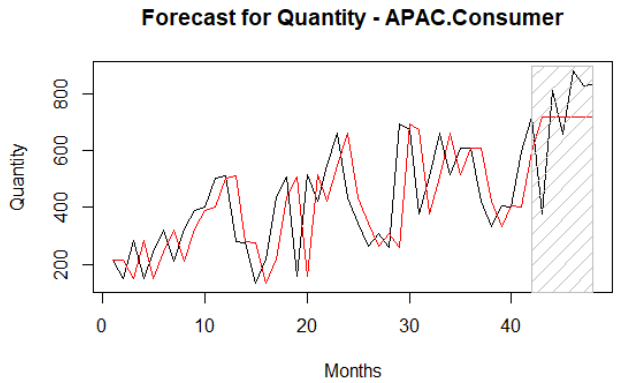
Global Component: LM fit

lmfit <- lm(Quantity ~ sin(0.5*Months) * poly(Months,1) * cos(0.5*Months), data=apacq_smoothdf)

Final Model = Local + Global

MAPE = **29.98**

AUTO ARIMA

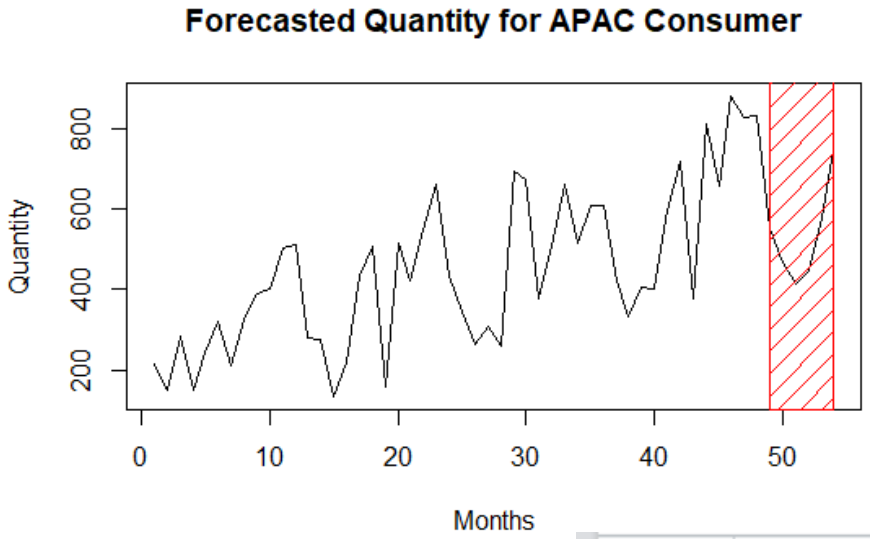


Model : ARIMA(0,1,0)

sigma^2:	25366
log likelihood	=-266.07
AIC	=534.14
AICc	=534.24
BIC	=535.85

MAPE = **26.24**

FORECASTS FOR THE NEXT SIX MONTHS



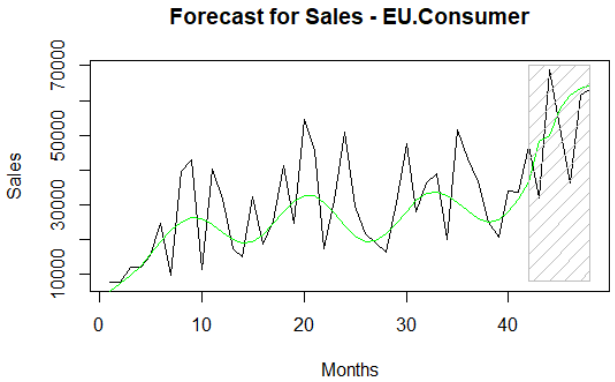
This is the Forecasted Quantitites for the Next Six Months for the APAC Consumer segment

	Months	Quantity
1	49	553.6266
2	50	469.1545
3	51	414.6781
4	52	447.2265
5	53	576.0759
6	54	747.2050

**All Forecasts done using Classical Decomposition*

FORECASTING FOR EU CONSUMER - SALES

CLASSICAL DECOMPOSITION



Local Component : ARIMA(0,0,2)

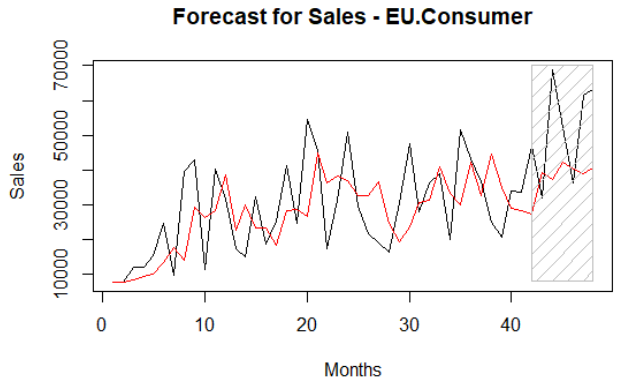
Global Component: LM fit

lmfit <- lm(Sales ~ sin(0.5*Months) * poly(Months,3) + cos(0.5*Months) * poly(Months,2) + sin(0.5*Months)*exp(0.0008*Months) + cos(0.5*Months)*exp(0.0008*Months),data= eus_smoothdf)

Final Model = Local + Global

MAPE = **27.11**

AUTO ARIMA



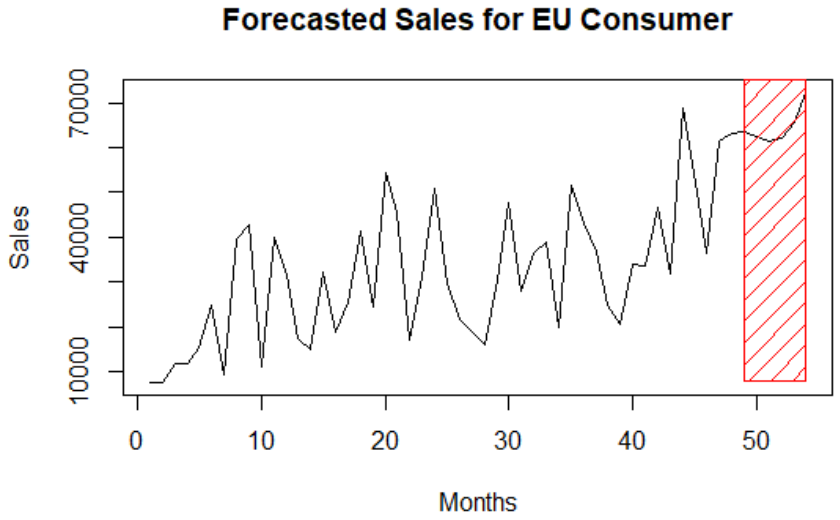
Model : ARIMA(2,1,0)

Coefficients:	-0.5892 -0.4889
s.e:.	0.1348 0.1312

log likelihood	=- 445.35
AIC	= 896.7
AICc	= 897.35
BIC	= 901.84

MAPE = **28.50**

FORECASTS FOR THE NEXT SIX MONTHS



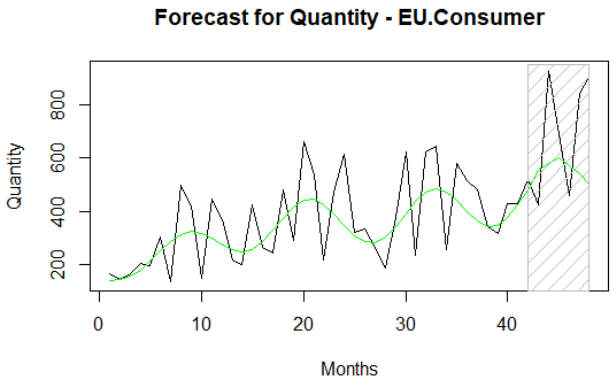
These are the Forecasted Sales for the Next Six Months for the EU Consumer segment

	Months	Sales
1	49	63666.61
2	50	62399.92
3	51	61492.49
4	52	62135.98
5	53	65498.23
6	54	72436.84

**All Forecasts done using Classical Decomposition*

FORECASTING FOR EU CONSUMER - QUANTITY

CLASSICAL DECOMPOSITION



Local Component : ARIMA(1,0,2)

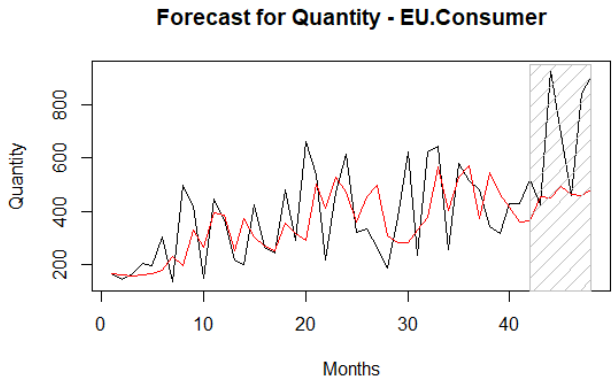
Global Component: LM fit

lmfit <- lm(Quantity ~ (sin(0.5*Months) * poly(Months,2) + cos(0.5*Months) * poly(Months,2)) *exp(0.00005*Months),data=euq_smoothdf)

Final Model = Local + Global

MAPE = **31.01**

AUTO ARIMA

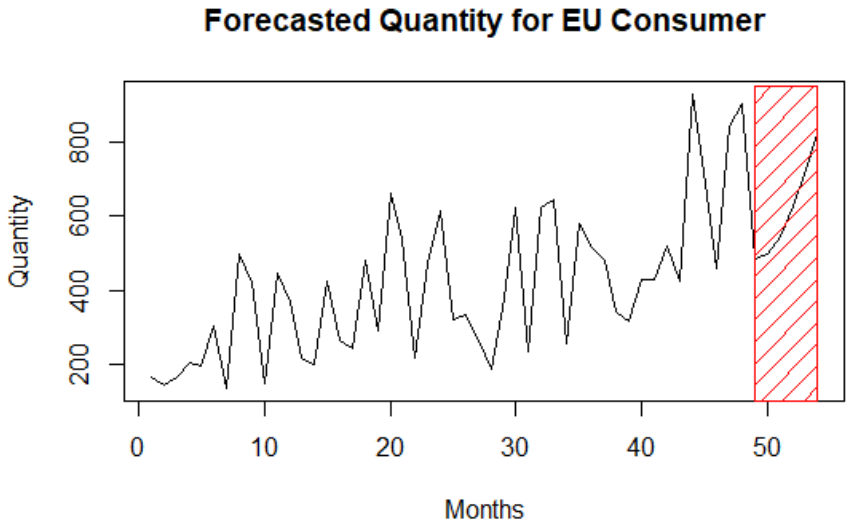


Model : ARIMA(2,1,0)

Coefficients:	-0.7359 -0.5879
s.e.:	0.1224 0.1185
sigma^2:	21185:
log likelihood	= -261.9
AIC	= 529.8
AICc	= 530.44
BIC	= 534.94

MAPE = **30.13**

FORECASTS FOR THE NEXT SIX MONTHS



These are the Forecasted Quantities for the Next Six Months for the EU Consumer segment

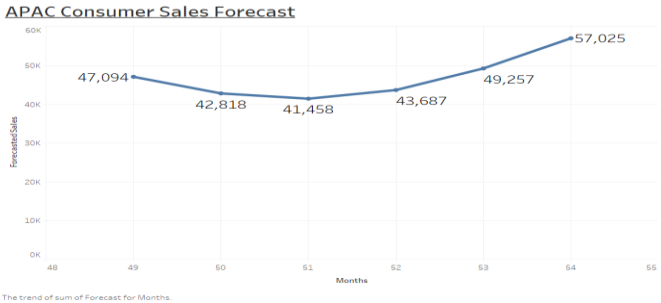
	Months	Quantity
1	49	486.8419
2	50	497.7507
3	51	546.6297
4	52	627.2299
5	53	726.2971
6	54	821.8325

**All Forecasts done using Classical Decomposition*

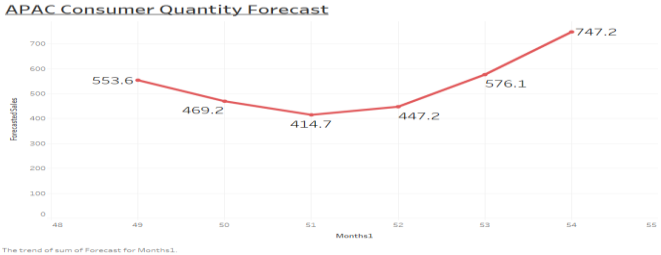
CONCLUSIONS

All Forecasts done using Classical Decomposition

APACS CONSUMER

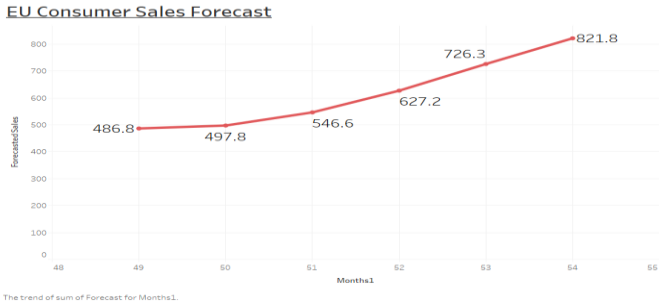


1 SALES : We expect the Sales to gradually Increase over the Next 6 Months. The Highest Sales is forecasted to be in Month 54

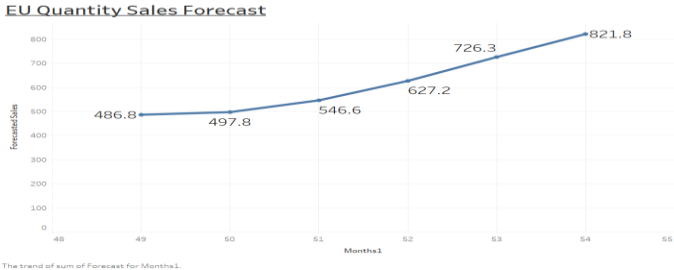


2. QUANTITY : We expect the Quantity to average around the 450 mark over the Next 6 Months. However, high quantities (>750)will be required from Month 54 onwards.

EU CONSUMER



1 SALES : We expect the Sales to gradually Increase over the Next 6 Months. The Highest Sales is forecasted to be in Month 54



2. QUANTITY : We expect the Quantity to gradually increase over the next 6 Months. Higher Quantities will be required every Month , to keep up with Demand