GLOBAL RETAIL GIANT FORECASTING

A CASE STUDY

BY TALLURI K SAKETH RAM

PROBLEM STATEMENT

- Global Mart is an online supergiant store that has worldwide operations. This store takes orders and delivers across the globe and deals with all the major product categories — consumer, corporate and home office.
- As a sales manager for this store, you have to forecast the sales of the products for the next 6 months, so that you have a proper estimate and can plan your inventory and business processes accordingly.
- Based on the **21 unique "Market-Segments"** for which the sales forecasts can be made. That is the **dataset** needs to be **prepared** such that we get the **Order-Date**, **Sales** and **Profit** for the **21 market segments**.
- Not all of these 21 market segments are important from the store's point of view. We need to find out the most consistently profitable market-segment from the above and forecast the sales and demand for that single market-segment only

DETAILS OF DATA SETS

Dataset provided for this case had the following files:

Global+Superstore+Data.csv - The data has the details about Sales and Profits from different Markets and Segments

The columns in the above dataset as follows:

Attributes	Description
Order-Date	The date on which the order was placed
Segment	The segment to which the product belongs
Market	The market to which the customer belongs
SalesTotal	Sales value of the transaction
Profit	Profit made on the transaction

DETAILS OF DATA SETS (CONTD..)

- In the dataset, we will see that the store caters to 7 different geographical market segments and 3 major customer segments, i.e. consumer, corporate and home as can be seen in the table below

Market	Segment
Africa	Consumer
APAC (Asia Pacific)	Corporate
Canada	Home Office
EMEA(Middle East)	
EU (European Union)	
LATAM (Latin America)	
US (United States)	

PREPARING THE DATA SET

- First we read the dataset and check the details like info, shape on the data set and check for Null Values.
- Next, we find create a column for the Market Segments and populate the same by concatenating the Market Column and Segment Column.
- There are 7 Markets and 3 Segments as Shown in the Above Slide.
- After creating Market Segments, we will get 21 Market Segments in total.
- Next, we Calculate Coefficient of Variance (COV) for finding the most Consistently Profitable market and as in the Problem Statement, go ahead with Modeling on that Segment.

COEFFICIENT OF VARIANCE

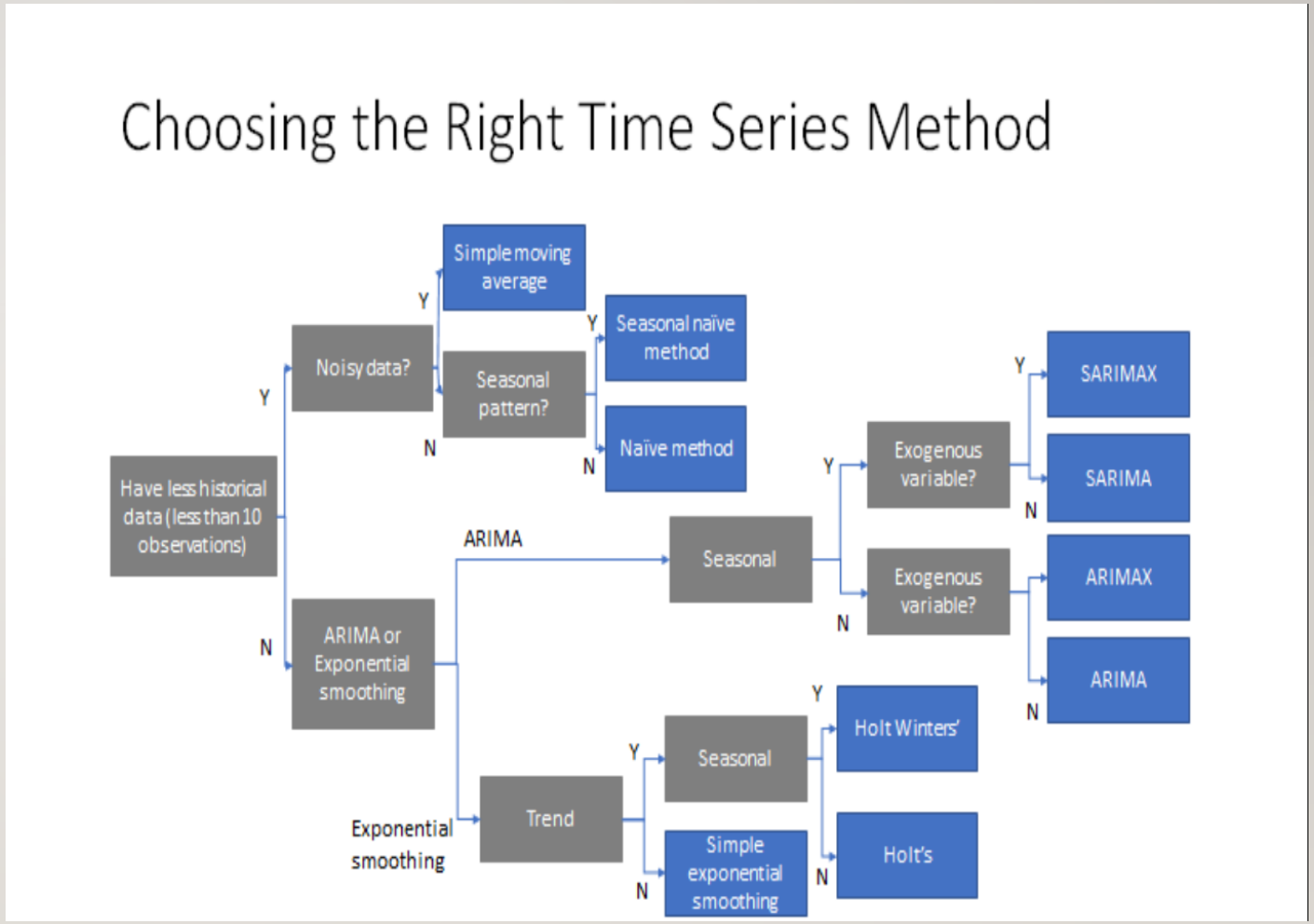
	SEGMENT	COV
0	APAC-Consumer	0.522725
1	APAC-Corporate	0.530051
12	EU-Consumer	0.595215
15	LATAM-Consumer	0.683770
13	EU-Corporate	0.722076
16	LATAM-Corporate	0.882177
14	EU-Home Office	0.938072
2	APAC-Home Office	1.008219
18	US-Consumer	1.010530
19	US-Corporate	1.071829

	SEGMENT	COV
20	US-Home Office	1.124030
17	LATAM-Home Office	1.169693
6	Canada-Consumer	1.250315
3	Africa-Consumer	1.310351
7	Canada-Corporate	1.786025
4	Africa-Corporate	1.891744
5	Africa-Home Office	2.012937
8	Canada-Home Office	2.369695
9	EMEA-Consumer	2.652495
10	EMEA-Corporate	6.355024
11	EMEA-Home Office	7.732073

We can observe from the above table that, APAC-Consumer Segment has the least COV and we can conclude that, it has most Consistent Profit.

MODELING AND FORECASTING

- Next, we head to the Modeling and Forecasting the data.
- Here the data we have is for 48 Months which is higher that 10 historical observations.
- When we have more than 10 historical observations, we can choose either Smoothing Techniques or Auto Regression Techniques as shown in the flow chart.



MODELING AND FORECASTING (CONTD.)

- We can plot the data and say that, the data we have is a Seasonal Data.
- So, when we go for Smoothing Techniques, we go for the Trend and Seasonal and go for the Holt Winters' Method and Calculate the **Root Mean Square Error (RMSE)** and **Mean Absolute Percentage Error (MAPE)** for all the methods.
- In the same way, when we go for Auto Regression Techniques, we go for the Seasonal and also calculate, the RMSE and MAPE using Auto regressive integrated moving average (ARIMA) and Seasonal auto regressive integrated moving average (SARIMA) techniques also.

FORECASTING- SMOOTHING TECHNIQUES

- So, when we forecast using different Smoothing Techniques, we get different RMSE and MAPE values as shown in the beside table.
- From the table, we can observe that, 'Holt Winters' additive method' has the lower RMSE and MAPE values.
- We can conclude that 'Holt Winters' additive method' method is best method for forecasting this problem using Smoothing Techniques.

Method	RMSE	MAPE
Naive method	12355.97	17.47
Simple average method	24146.06	34.34
Simple moving average forecast	14756.73	15.82
Simple exponential smoothing forecast	15011.49	15.99
Holt's exponential smoothing method	18976.37	34.57
Holt Winters' additive method	9555.63	9.33
Holt Winters' multiplicative method	9423.23	11.43

FORECASTING- AUTO REGRESSION TECHNIQUES

- So, when we forecast using different Auto Regression Techniques, we get different RMSE and MAPE values as shown in the beside table.
- From the table, we can observe that, 'Seasonal autoregressive integrated moving average (SARIMA)' has the lower RMSE and MAPE values.
- We can conclude that 'Seasonal autoregressive integrated moving average (SARIMA)' method is best method for forecasting this problem using Auto Regression Techniques.

Method	RMSE	MAPE
Autoregressive (AR)	10985.28	13.56
Moving Average (MA)	23360.02	33.93
Autoregressive moving average (ARMA)	22654.32	32.40
Autoregressive integrated moving average (ARIMA)	22654.32	32.40
Seasonal autoregressive integrated moving average(SARIMA)	9616.66	12.87

CONCLUSION

- For the given business problem the following methods are best for forecasting the Sales of the Retail Giant Company
- SMOOTHING TECHNIQUES - Holt Winters' additive method
- AUTO REGRESSION TECHNIQUES - Seasonal autoregressive integrated moving average (SARIMA)

END