Retail-Giant Sales Forecasting

SUBMISSION

Submitted by:

1. Shubham Joshi
2. Ajay Gupta
3. Ashish Bhogawkar

Business Problem

1. To find out 2 most profitable (and consistent) segments from 21 segments which are formed from the combination of 3 major categories and 7 different market segments.

1. Forecast the sales and demand for these two segments.

Analysis Approach
CRISP-dm (Cross Industry Standard Process For Data Mining) Methodology
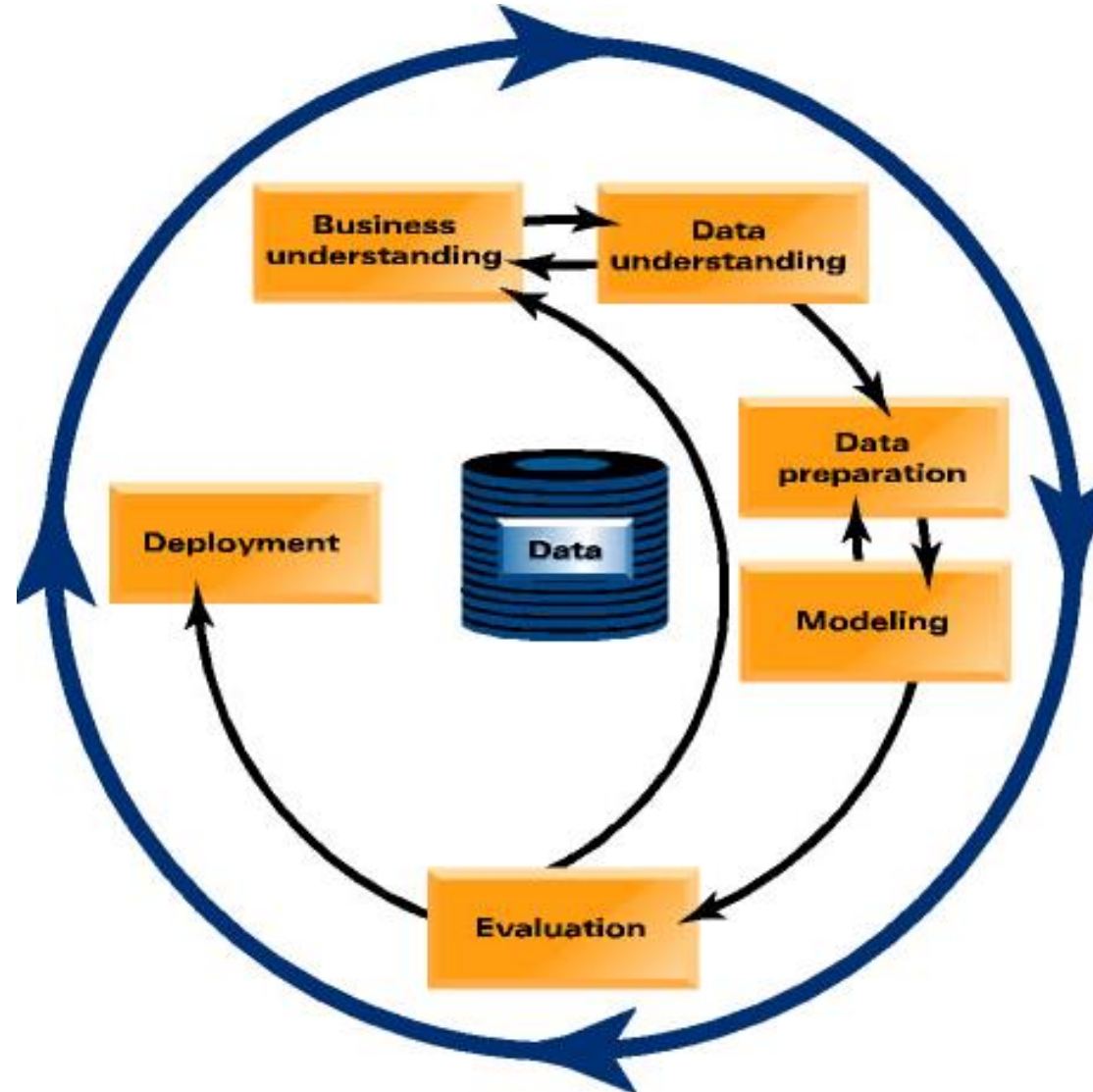
CRISP-dm (Cross Industry Standard Process For Data Mining) Methodology

1. Business Objectives

- Determine and Evaluate Business Objectives
- Determine Data mining Goals

2. Data Study

- Collect Initial Data and Map Data Needs to sources
- Identify data gaps and mismatch
- Evaluate and suggest data source
- Explore type and quantity

3. Data Preparation

Data Preparation

- Data Cleansing
- Filling up the missing values
- Data Formatting
- Labelling the data
- Converting Categorical data into numeric data
- Creating derived fields Analysis & Modeling
- Sampling

4. Modelling

DATA MODELING

- Select Modelling Techniques
- Generate Test Design
- Build Model
- Access the Model

4. Evaluation

- Select Modelling Techniques
- Build Model
- Generate Test Design
- Test Run on sample data

5. Reporting and Deployment

- Resource Allocation
- Periodic screening & scrutiny of model
- Reporting as based on Client Requirement

Global Mart is an online store super giant having worldwide operations.

1. This online store caters to 7 different market segments and in 3 major categories.

2. Forecasting is required for the sale and the demand.

3. The forecasting is required for at granular level of market segment and category.

Data Understanding:-

The data is given in 1 different table which is as follows:-

| Sr. No	Name	Description
1	Global Superstore	This table contains transactions i.e. order specific data such as order ID, Order date, Ship date, ship mode, Sales, quantity, profits are some of the important attributes. There are total 2 attributes and total 51290 records.

Following are the Data preparation operations Performed:-

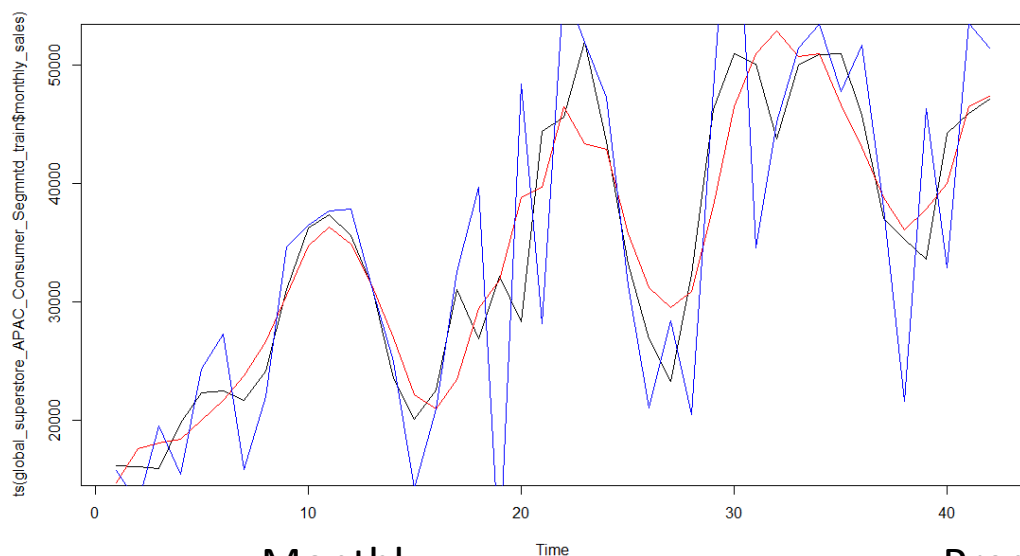| Data preparation operations	Description
Data Cleaning:	Not Required
Outlier Removal:	Not Required
Data Formatting:	Converting dates in string to Date objects.
Data Merging:	Not Required
Derived Metrics:	Not Required

Time Series

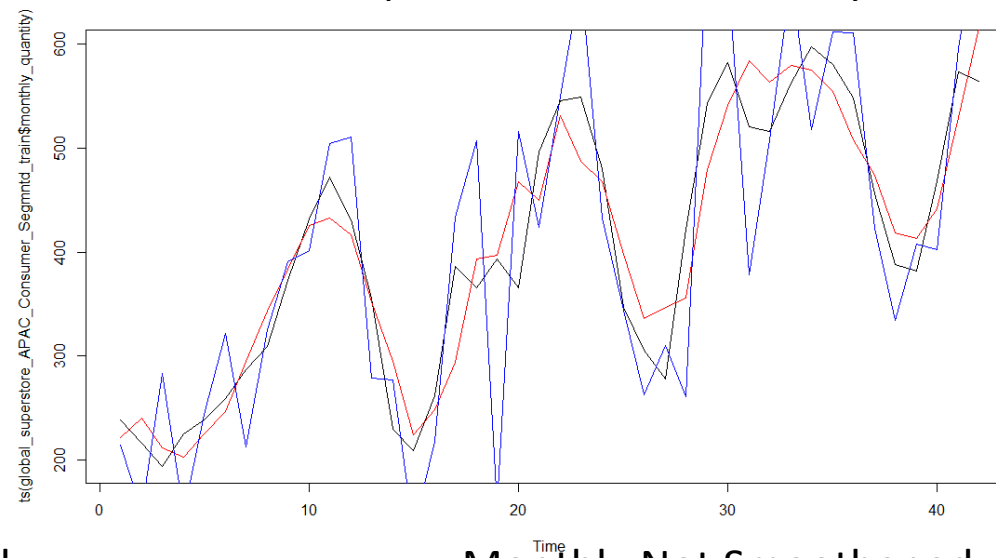## Timeseries – Europe Consumer - Sales
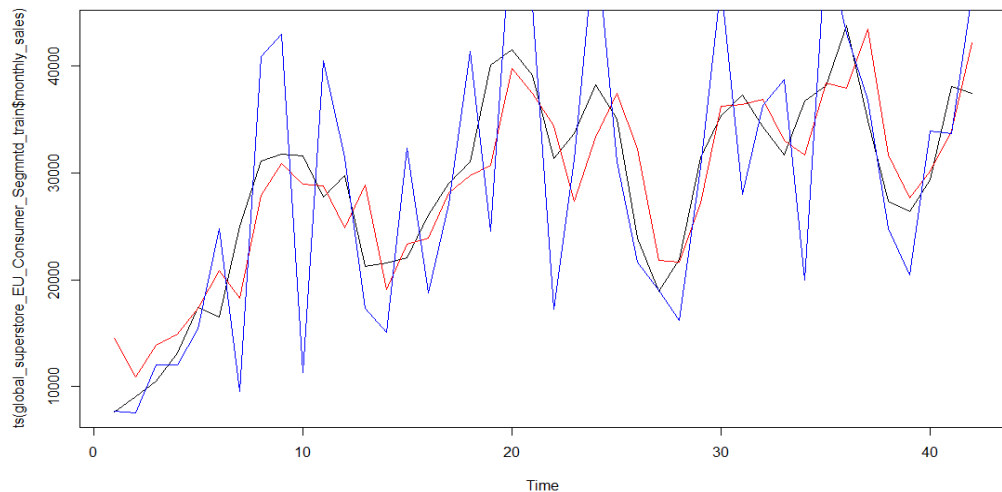
Timeseries – Europe Consumer - Quantity
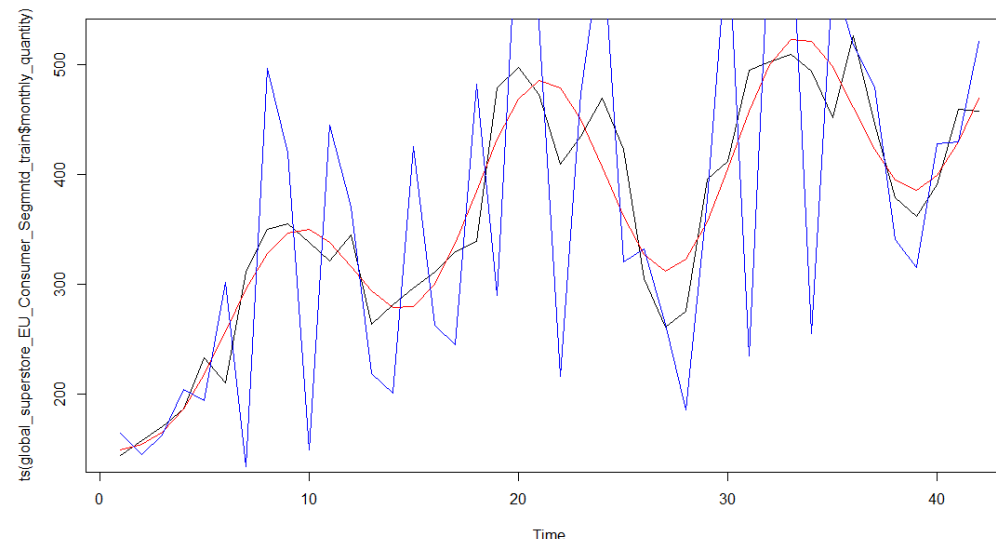
Timeseries – Apac Consumer - Sales

Timeseries – Apac Consumer - Quantity

| Monthly | Predicted | Monthly Not Smoothened |

Forecasting for Sales and Quantity – Europe and APAC

Sr. No	Month	Europe - Consumer		APAC - Consumer	
		Quantity	Sales	Quantity	Sales
01	Jan - 2015	500.75	41818.91	864.32	48248.48
02	Feb - 2015	526.51	42233.52	748.68	55131.85
03	Mar - 2015	574.81	42664.43	607.03	59345.69
04	Apr - 2015	636.83	43117.19	472.77	58345.69
05	May - 2015	697.34	43586.15	388.46	49105.84
06	Jun - 2015	739.10	44062.84	396.62	32050.28

Recommendation

Business Focus should be on following Segment and Market

- APAC – Consumer

- Europe - Consumer

Thank You