**Interim Report**

**On**

**Actionable Insights through Analytics for a**

**Consumer Durables Retailer**

**Submitted by**

Rahul Bhardwaj

Amar Vivek

Harshvardhan Mirajkar

Badri Narayan Patnaik

Anirudh Bhatjiwale

Sandeep Pangul

***PGPBA Batch August 2016 (Pune)***

**Research Supervisor**

Mr. Snehamoy Mukherjee



**Table of Contents**

[1. Introduction 3](#_Toc482551568)

[2. Scope and Objectives 3](#_Toc482551569)

[3. Data Sources 4](#_Toc482551570)

[4. Tools 4](#_Toc482551571)

[5. Progress till Date 4](#_Toc482551572)

[6. Next Steps 5](#_Toc482551573)

[7. Challenges 5](#_Toc482551574)

[8. Exploratory Analysis 5](#_Toc482551575)

[Weekly Sales 5](#_Toc482551576)

[Branch Data 6](#_Toc482551577)

[Brand Analysis 7](#_Toc482551578)

[Product Analysis 8](#_Toc482551579)

# 1. Introduction

Today’s Consumer is markedly different from that a few years ago. They do research on pricing, delivery time, product reviews and online alternatives and make a decision to buy from a Retailer who provides maximum value. They are highly informed, enabled by new technologies and tools right from their mobile and influenced to a greater extend by Social Media before making buying decisions. Their expectations are extremely high as they expect the best of everything. This combined effect makes it challenging for Retailers to win their wallet share.

Retailers have to listen to the voice of the Consumer, differentiate their offerings to the multi-generational customer base with drastically different needs, shopping habits and interactions with the brand. Developing a loyal customer base for their brand is crucial for retailers in order to differentiate their product from the competition and build an ongoing relationship with customers in an aggressive business ecosystem. Retailers want their customers purchasing their products and recommending them to friends and family. Some of the questions that the retailers would want answers for are:-

What are they thinking? What are their needs?

Our Capstone is aimed at exploration of data from a consumer durable retail outlet based in Chennai by applying Statistical Techniques, so as to recommend actionable insights for the Retailer in delivering a compelling value proposition to the customer.

# Scope and Objectives

The scope of the project is to gain insight into the retail stores performance by studying various KPIs across different branches of the store. We will conduct exploratory analysis to investigate the data and will use unsupervised learning techniques like Clustering, RFM analysis, Market-basket analysis etc. to identify the customer segment and recommend the product basket. We will also try to build a model on sales forecast which would help the retailed in better inventory management. We will also work towards preparing a dashboard of key performance indicators for the retailer that will help in taking informed decisions. We are also considering sentiment analysis as a tool to study the impact of opinions on social media on the brand performance.

Objectives of the Research:

* To perform customer segmentation and product associations that can help the Retailer to increase the business at both customer and transaction level
* To build analytical models to forecast the sales
* To evaluate the brand performance and store performance across different stores
* To enable the retailer with appropriate dashboard based on KPIs to take decisions

The above objectives will get modified and expanded based on our exploratory analysis and further understanding of the dataset.

# 3. Data Sources

POS data from consumer durables (electronic items) from Chennai based Retailer for the year 2010-11 and 2011-12 shall be used for analysis, hypothesis testing and developing the predictive model. The 2012-13 data shall be used for validation. The dataset contains 1.2 Million records for 3 years.

The following are the variables available in the data.

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Attribute** | **Indicative Usage** |
| 1 | Bill No | Transaction Identification |
| 2 | Branch | Geographical segmentation |
| 3 | Class of Product | Product group clustering |
| 4 | Product | Exploratory analysis, market-basket |
| 5 | Brand | Exploratory analysis, wallet share, dimensions, profiling |
| 6 | QBC | Cluster Identifier |
| 7 | Item Description | Text mining for similarity, brands etc |
| 8 | Address | Customer segmentation, profiling, geospatial clustering |
| 9 | Pincode | Geospatial segmentation, Prediction |
| 10 | Qty | Detection of outliers, determine count of sales volume |
| 11 | Value | Detection of outliers, determine sales revenue |

# 4. Tools

Following tools are proposed for the project: Excel, RStudio, Tableau, KNIME

# 5. Progress till Date

1. Retail Data set has been consolidated across years and imported into R and Tableau
2. Exploratory Data Analysis has been done to understand branch, product, category and brand wise sales. Monthly and Day-wise(throughout the week) data has also been studied
3. Identifying top branches/product/brands in terms of sales. Also, studying the pareto charts to see who is contributing to the majority of sales
4. Working on coding for Market Basket Analysis based on item description and products being bought together
5. Exploratory Analysis has been listed in the section at the end

# 6. Next Steps

1. Unique identification of customers based on address
2. Perform RFM analysis
3. Prepare story boards and dashboards in Tableau
4. Provide recommendations based on RFM, MBA and outputs of other exploratory analysis
5. Prepare the project report and presentation
6. Identify what distinguishes top performing branches from low performing branches

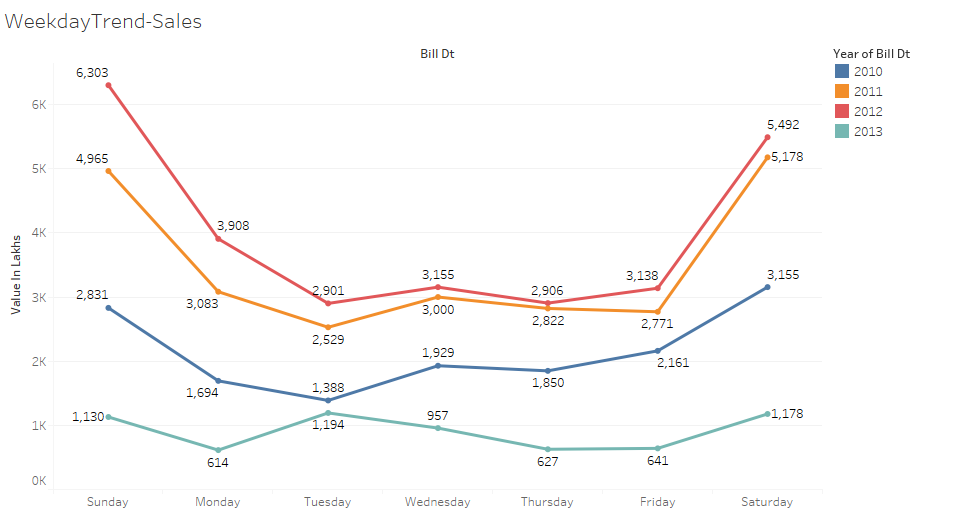
# 7. Challenges

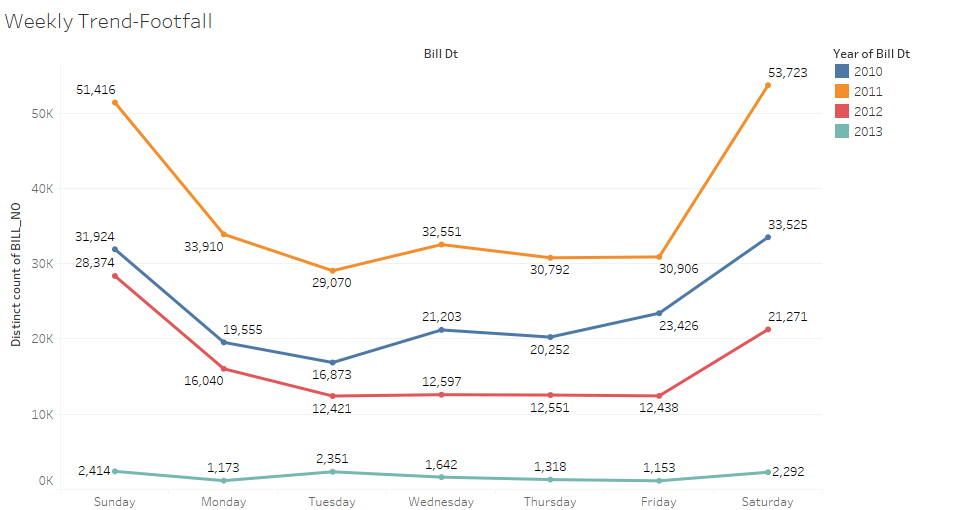
1. Unique identifier for customer is not available thus difficult to identify customers to perform RFM
2. Other supporting data like square footage of stores, neighborhood type etc. is not present. It could have helped us gain some additional insights

# 8. Exploratory Analysis

## Weekly Sales

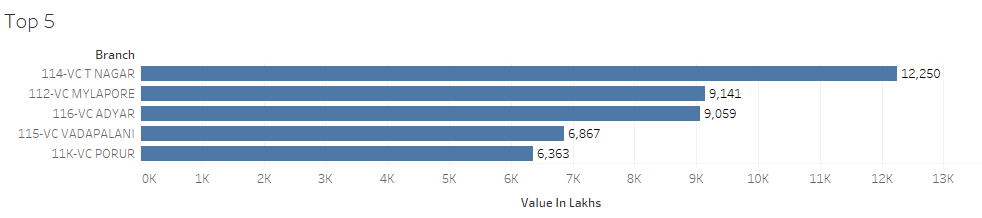
* Sum of Sales on Weekends is significantly higher than on weekdays as depicted in the graph below
* Sum of Sales has been increasing over the years. 2013 shows a dip because the data is available only for first quarter
* When comparing to the distinct count of bill generated, we see that there has been a dip from 2011 to 2012.

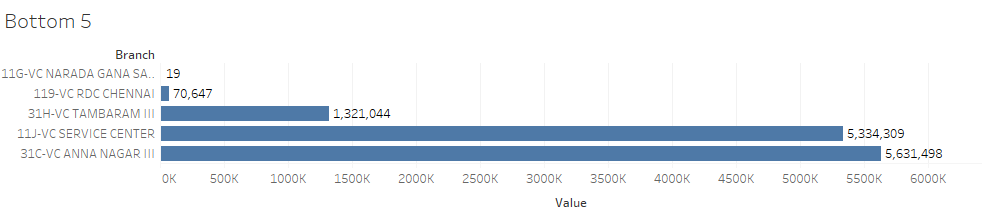


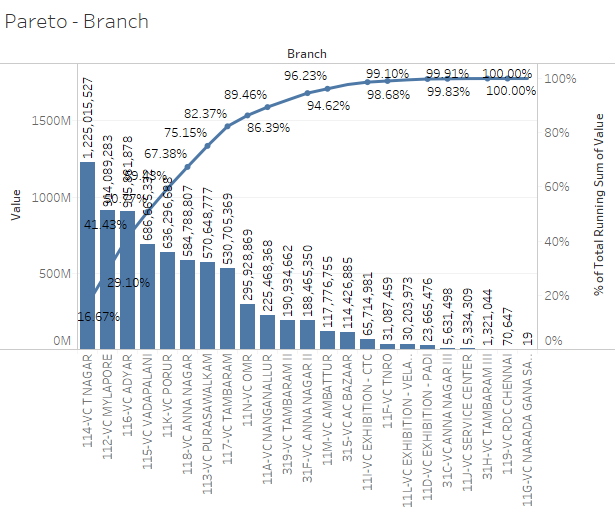


## Branch Data

* T Nagar is the branch that is doing the best in terms of sales followed by Mylapore
* On the other side Narada Gana Sabha is not doing much in sales followed by RDC Chennai branch
* Around 82% of sales are coming from 8 branches out of the 23 branches

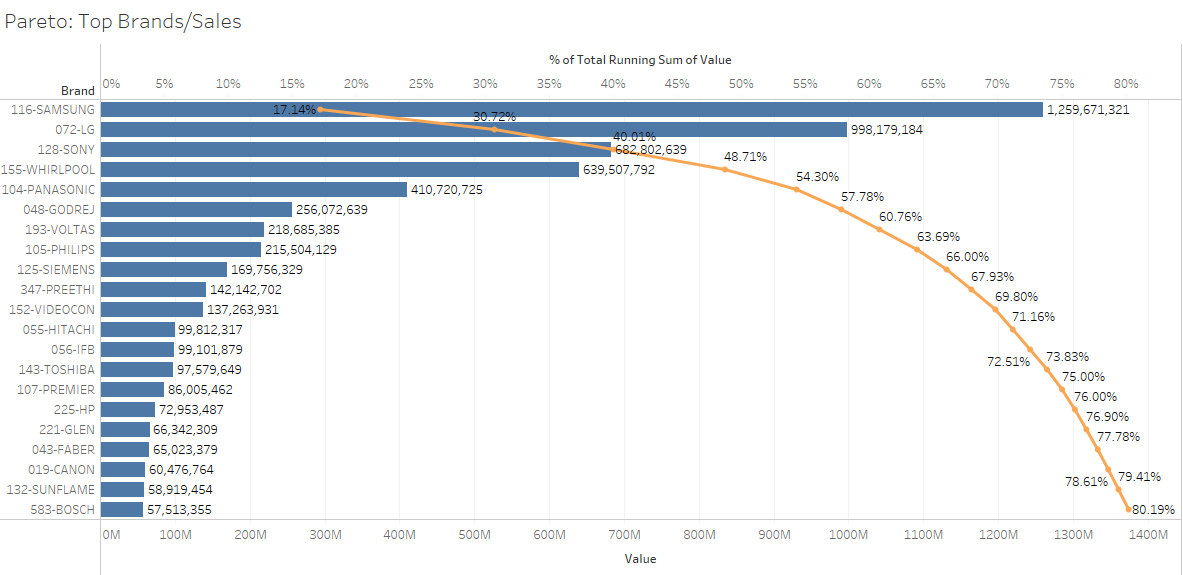






## Brand Analysis

* Samsung is the brand which is selling the most
* Around 60% of sales are coming from 7 brands



## Product Analysis

* Flat Panel TV is the most selling product in terms of sales value followed by washing machine
* Around 80% of sales are coming from 11 products

