SRI RAMACHANDRA
INSTITUTE OF HIGHER EDUCATION AND RESEARCH
(Category - I Deemed to be University) Porur, Chennai
SRI RAMACHANDRA FACULTY OF ENGINEERING AND TECHNOLOGY

AUTOMATED SALES ANALYSIS AND FORECASTING

INT 625 – INTERNSHIP 3
FINAL REVIEW

Maria Anthony Raja – E7321009 –

M.Sc. (Data Analytics)

Supervisor

Dr. P.A.Abhinand

Lecturer

Bioinformatics

SRIHER (DU)

Contents

Introduction

- The multinational retail chain Walmart operates several supermarkets, discount department stores, and hypermarkets.

- Walmart is the one of largest retailers in the world, with more than 11,000 outlets across 27 countries. The enterprise is based in the United States.

- The company employs around 2.3 million people worldwide, and it serves millions of customers every day.

Literature Survey

Project in Analytics 2022	Walmart’s Sales Analysis through Data Visualization, Aman Preet Gulati, et al.
Case study in Analytics 2023	Superstore Sales & Profit Report Using Power BI, Chaitanya Shah, et al.
Kaggle competition 2014	Walmart Sales Forecast projects
NovyPro platform 2023	Walmart's sales data analysis, K.D.Manoj Kumar, et al.

Overview

There are many projects in the internet especially in Kaggle website for the Walmart Sales Analysis. However, the proposed project deals with more automation with the help of Power BI.

Literature Survey

Comparative Review

- The available project featured in the Kaggle mostly built with the help of Python. The proposed project uses R programming for statistical analysis and forecasting.

- Same way, all available projects are using Python for data visualization whereas proposed project uses Power BI for advanced and customized data visualization into a dashboard.

Literature Survey

Gaps in Available Research

- **Limitation in data visualization**: Data analysis results can be more easily understood by users when they are visualized. Most of the projects used Python, where it is impossible to showcase the charts; instead, they can only be generated as images.

- **Missing statistical analysis**: Businesses can use statistical analysis to help them make better decisions that can drive growth, profitability, and provide them a competitive edge in their market. Available research fails to implement more on statistical analysis.

- **Start from zero**: The entire process of data preparation, forecasting, and reporting must be restarted from scratch whenever fresh data is received.

- **Not an end user product**: With the Python data visualization, the charts only can be prepared and reported. These charts are static and end users should edit the code if they want to interact with the visuals.

Research Objective

- The project's goal is to examine historical sales information for 45 Walmart locations spread across several areas, from February 2010 to November 2012. Predict sales for the future and try to derive insight from the data.

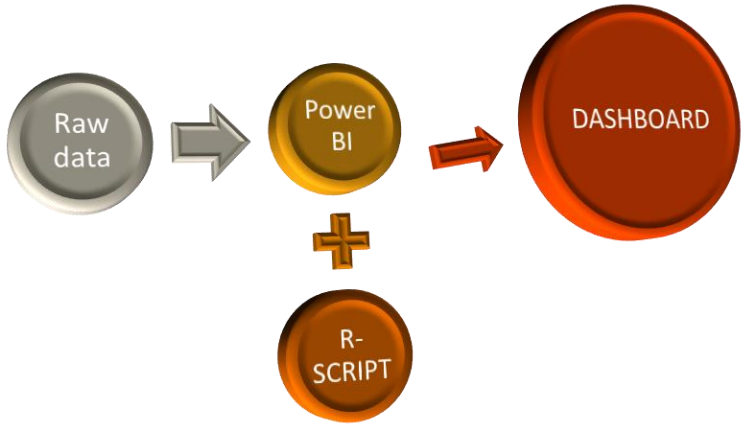
- The methodologies throughout this project comprise business intelligence tools, data analytics, and statistics.

Tools and Techniques

- PowerBI – Data loading, transformation and visualization. Also used to perform data analysis.

- R-Script – To Implement statistical analysis and advanced visuals.

Methodology



Dashboard: Using Power BI, all the charts can be consolidated to form a dashboard. This gives a great presentation experience for end users.

Automation: This part is included in proposed project which can eliminate the manual work on building the dashboard for upcoming new data.

Slicers and Filters: Using these options – Slicers and Filters in Power BI dashboard, end users can interact with the visuals easily.

Methodology

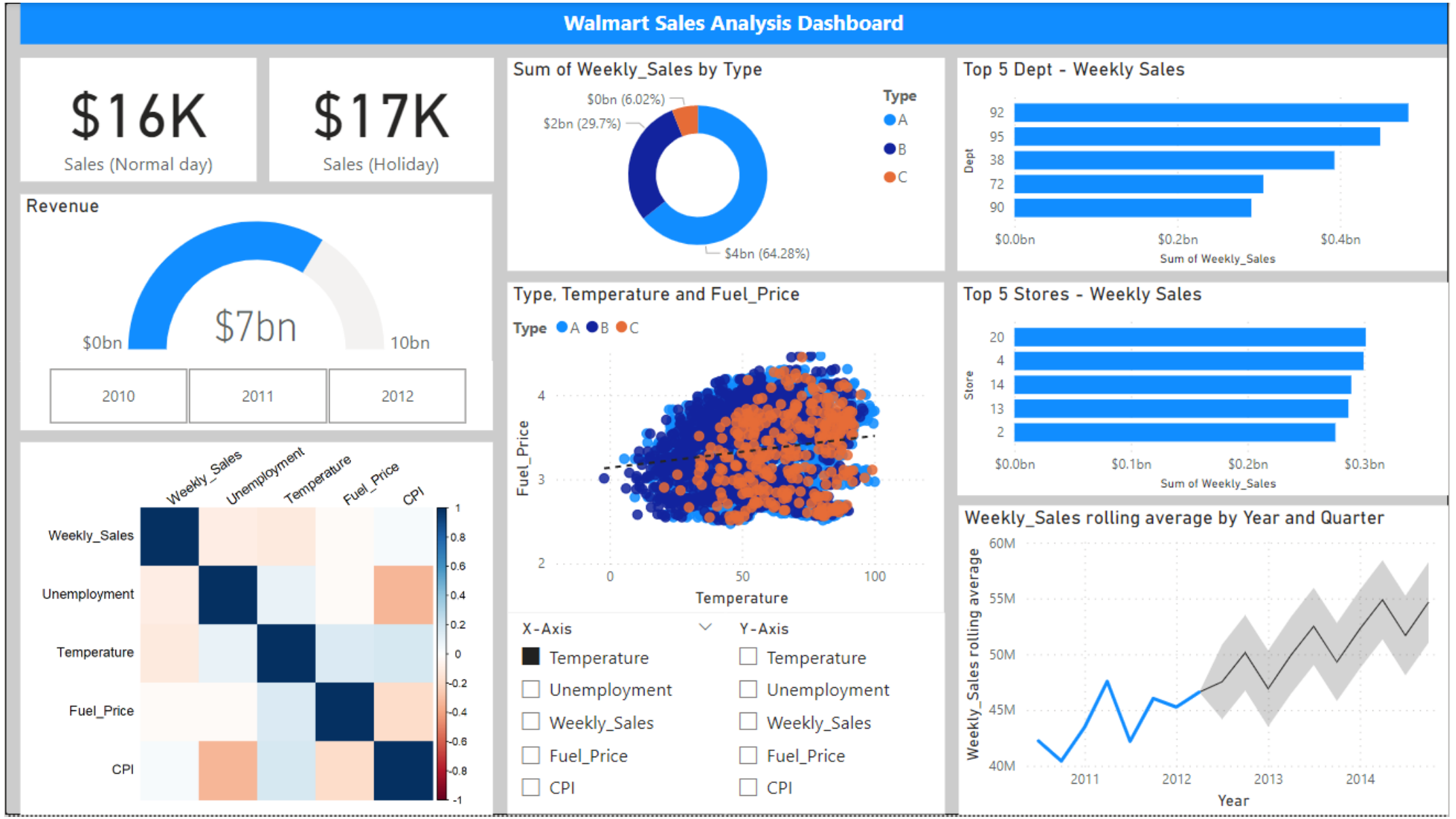
Integrating R script: Power BI can support in-built R script access which can be used to perform advanced statistical analysis. Businesses can use the statistical methods below as useful tools to understand their historical performance and make future decisions.

- **Descriptive Statistics**: The historical data on the performance of the company is summarized using fundamental statistical measures including mean, median, mode, standard deviation, and range. Descriptive statistics, such as the average revenue or profit margins for a specific time, can be utilized to gain a general understanding of the company's previous performance.

- **Time Series Analysis**: Based on historical trends, time series analysis can be used to forecast and predict future performance. Forecasting model in PowerBI is used for time series analysis.

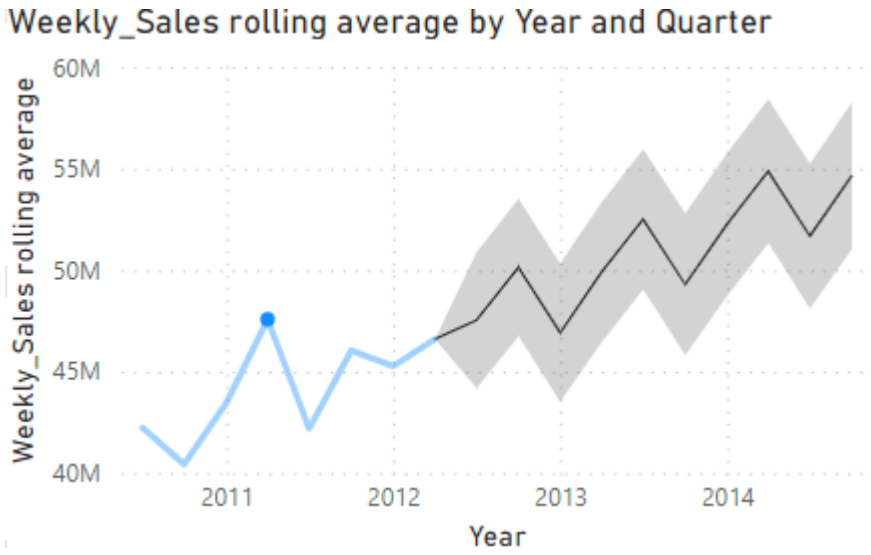
- **Regression Models**: The proposed project uses regression models to examine the relationship between independent parameters and how they affect the performance of the business. Regression analysis can be used to identify important factors that affect a company's success, including changes in consumer preferences, the state of the economy, and market rivalry.

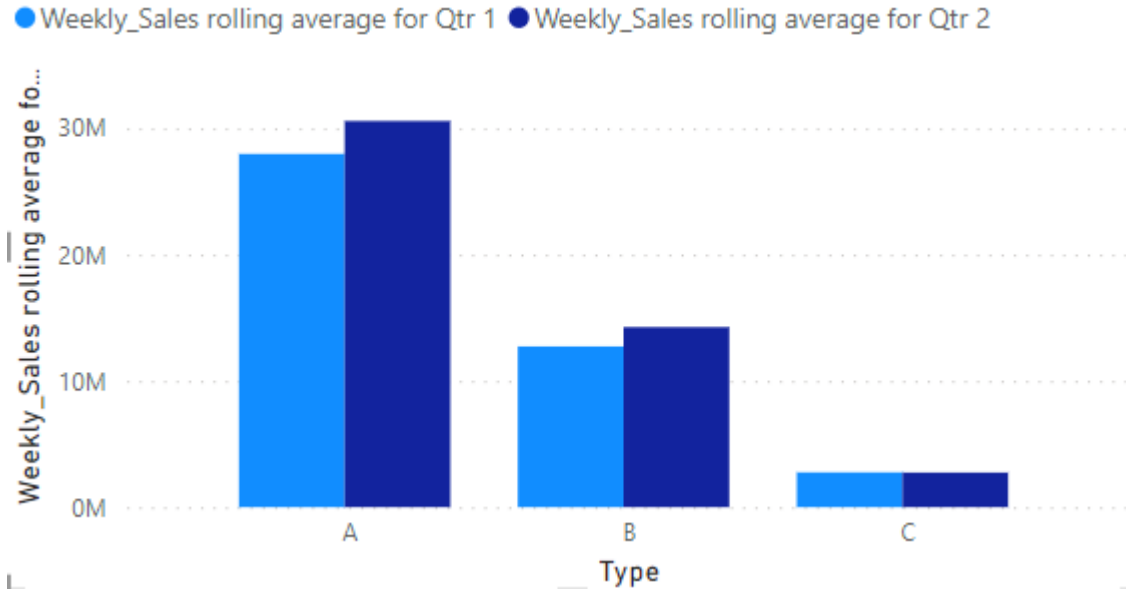
Experimental Results



Experimental Results

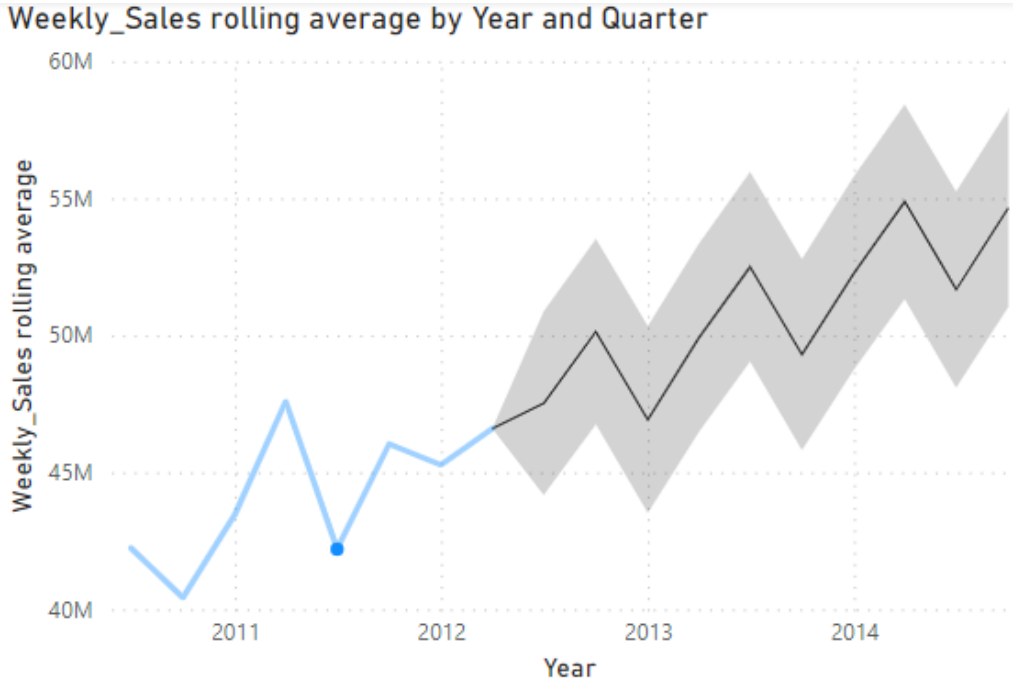
As the store type A and B received more sales in 2011 Quarter 1. Hence, there is a 9.48% increase in Overall Sales between 2011 Quarter 1 and 2011 Quarter 2.

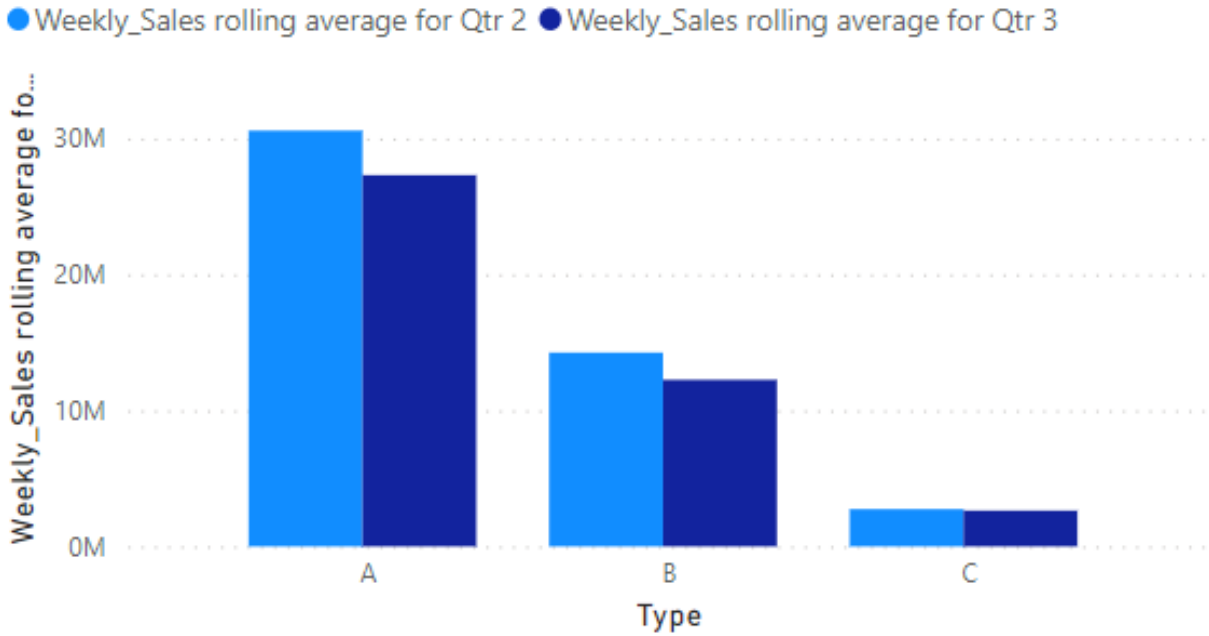




Experimental Results

As the store type A and B have lower sales in 2011 Quarter 3. Hence, there is a 11.31% decrease in Overall Sales between 2011 Quarter 2 and 2011 Quarter 3.





Conclusion

- Power BI will be the greatest choice when it comes to data visualization as the business intelligence tools feature cutting-edge in-built technologies for data visualization and analysis. Power BI's key advantage will be automation, which can eliminate manual input and the time needed to write lines of code.
- R programming is a well-liked option for statistical analysis due to its extensive libraries, powerful data handling capabilities, reproducibility, flexibility, and open-source status.
- The suggested project integrates R script and Power BI to produce the finest automated statistical results with an excellent dashboard.
- The interactive dashboard has been built which delivers the report automatically when data updated.
- The store type A produced almost 64% revenue all-time.
- Sales on holiday is higher than usual days has been proved evidently.
- The future sales have been forecasted.

References

- Walmart’s Sales Analysis through Data Visualization:
  https://www.analyticsvidhya.com/blog/2022/01/walmarts-sales-analysis-through-data visualization/

- Walmart Sales Dataset:
  https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting/data?select=features.csv.zip

- Walmart Sales Forecast projects in Kaggle:
  https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting/code

- Superstore Sales & Profit Report Using Power BI
  https://www.analyticsvidhya.com/blog/2023/01/a-case-study-superstore-sales-profit-report-using-power-bi/

- NovyPro - Walmart's sales data analysis:
  https://www.novypro.com/project/performed-kpiskey-performance-indicators-on-various-given-data-from-walmart-sales-report-for-the-time-period-between-012011-012015

Appendix

Rolling Average of Sales - DAX

```
Weekly_Sales rolling average =
IF(
    ISFILTERED('train'[Date]),
    ERROR("Time intelligence quick measures can only be grouped or filtered by the Power BI-provided date hierarchy or primary date column."),
    VAR __LAST_DATE = LASTDATE('train'[Date].[Date])
    RETURN
        AVERAGEX(
            DATESBETWEEN(
                'train'[Date].[Date],
                DATEADD(__LAST_DATE, -1, DAY),
                DATEADD(__LAST_DATE, 1, DAY)
            ),
            CALCULATE(SUM('train'[Weekly_Sales]))
        )
)
```

Correlation Matrix – R-Script

```
require("corrplot")
library(corrplot)
M<- cor(dataset)

corrplot(M, method = "color",  tl.cex=1, tl.srt = 35, tl.col = "black")
#,col = colorRampPalette(c("blue", "green"))(100)
```

Worklog

SRI RAMACHANDRA
INSTITUTE OF HIGHER EDUCATION AND RESEARCH
(Category - I Deemed to be University) Porur, Chennai
SRI RAMACHANDRA FACULTY OF ENGINEERING AND TECHNOLOGY

PROJECT WORK LOG

Unique ID: E7321009
Name of the Student: Maria Anthony Raja
Project Title: Automated Sales Analysis and Forecasting
Project Supervisor: Dr. P.A.Abhinand

Date	Duration in Hrs.	Task Planned	Task Completed	Description of the Task	Scope of the task	Skills Acquired	Review comments by supervisor	Signature of the Student	Signature of the Supervisor
25-Mar-23	0.5	Project idea discussion							
01-Apr-23	0.5	Abstract and Literature survey submission							
02-Apr-23	0.5	Guide's Suggestion in Literature survey							
04-Apr-23	0.5	Amendment in Literature survey							
02-May-23	0.5	Final review submitted							