GLOBAL SALES DATA ANALYTICS

PROJECT REPORT

NALAIYA THIRAN PROJECT

SUBMITTED BY

ABISEKH M

ABDUSH SAMAD M

ABIKRISHNA M

ARAVINDAN N

OF

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

FROM

PSNA COLLEGE OF ENGINEERING AND TECHNOLOGY

CONTENT

CONTENT DESCRIPTION

PROJECT REPORT

TEAM ID	PNT2022TMID05132
PROJECT NAME	GLOBAL SALES DATA ANALYTICS
TEAM MEMBER	ABISEKH M -(921319106005) ABDUSH SAMAD M-(921319106002) ABIKRISHNA M -(921319106003) ARAVINDAN N -(921319106020)

1. INTRODUCTION

1.1 Project Overview

Shopping online is currently the need of the hour. Because of this COVID, it's not easy to walk in a store randomly and buy anything you want. So, this project is done to try to understand a few things like, Customer Analysis and Product Analysis of this Global Super Store.

1.2 Purpose

By the end of this Project, you will:

- Know fundamental concepts and can work on IBM Cognos Analytics.
- Gain a broad understanding of plotting different visualizations to provide a suitable solution.
- Able to create meaningful Visualizations and Dashboard(s).

2. LITERATURE SURVEY

2.1 Existing problem

Shopping online is currently the need of the hour. Because of this COVID, it's not easy to walk in a store randomly and buy anything you want.

2.2 References

1.Data analysis and visualization of sales data - Mar-2016 Authors: Kiran Singh,Rakhi Wajgi Data is being generated very rapidly due to increase in information in everyday life. Huge amount of data get accumulated from various organizations that is difficult to analyse and exploit. Data created by an expanding number of sensors in the environment such as traffic cameras and satellites, internet activities on social networking sites, healthcare database, government database, sales data etc., are example of huge data. Processing, analysing and communicating this data are a challenge. Online shopping websites get flooded with voluminous amount of sales data every day. Analysing and visualizing this data for information retrieval is a difficult task. Therefore, a system is required which will effectively analyse and visualize data. This paper focuses on a system which will visualize sales data which will help users in applying intelligence in business, revenue generation, and decision making, managing business operation and tracking progress of tasks. 2.Walmart's Sales Data Analysis - A Big Data Analytics Perspective - Dec2017 Authors: Manpreet singh, Bhawick Ghutla, Reuben lilo Jnr, Aesaan F S Mohammed, Mahmaad A Rashid Information technology in this 21st century is reaching the skies with large-scale of data to be processed and studied to make sense of data where the traditional approach is no more effective. Now, retailers need a 360-degree view of their consumers, without which, they can miss competitive edge of the market. Retailers must create effective promotions and offers to meet its sales and marketing goals, otherwise they will forgo the major opportunities that the current market offers. Many times, it is hard for the retailers to comprehend the market condition since their retail stores are at various geographical locations. Big Data application enables these retail organizations to use prior year's data to better forecast and predict the coming year's sales. It also enables retailers with valuable and analytical insights, especially determining customers with desired products at desired time in a particular store at different geographical locations. In this paper, we analysed the data sets of world's largest retailers, Walmart Store to determine the business drivers and predict which departments are affected by the

different scenarios (such as temperature, fuel price and holidays) and their impact on sales at stores of different locations. We have made use of Scala and Python API of the Spark framework to gain new insights into the consumer behaviours and comprehend Walmart's marketing efforts and their data-driven strategies through visual representation of the analysed data.

3.Research on Refined Sales Management, Data Analysis and Forecasting under Big Data - Oct-2020 Author: Wenhui Shan This article analyses the key points of refined sales management under big data. The main points of sales management include how to establish a sales management organization, how to improve the sales management information system, how to improve the evaluation management system, and how to strengthen internal sales control. Combining the key points of data analysis under big data, the author studies the establishment of data warehouse, data cleaning and mining, the establishment of data prediction models, and the arrangement of model analysis results. The purpose of this article is to help people give full play to the advantages of big data technology applications and promote the healthy development of the enterprise economy.

4.Advanced data analysis and exploration of large marketing databases are the subjects of many research and application projects Authors: Setia & Jyoti, 2013 The knowledge acquired can be used automatically in the processes of communication with the client including operations undertaken in order to increase the chance of making a purchase, improve customer satisfaction, reduce the risk of a customer leaving, or optimise the margin on the product.

5.Application of advanced data analysis in marketing is considered in three main perspective Authors: Jannach, Zanker, Felfernig, & Friedrich, 2010 From the managerial perspective, we use it to build a decision support system that uses large, heterogeneous data and mechanisms generating recommendations related to the sales strategy and pro-motion of the products offered. From the customer's perspective, we prepare an advisory system facilitating selection of products in accordance with one's inter-ests, needs, and preferences. From an IT perspective, advanced data analysis is related to a computing platform containing a number of exploration models, integrated with transactional systems of the online and offline store and its envi-ronment. This platform must guarantee not only access to various information resources but also scalability of applications operating on a large number of information collections

2.3

Title	Author	Journal,Year	Proposed Method	Disadvantage
Big Data Analytics:A Literature Review Paper	Nikhil Madaan, Umang Kumar, Suman Kr Jha.	18-07-2020	Big Data Storage andManagement. Customer Intelligence.Quality Management and Improvement.Risk Management and Fraud Detection	The information era we are currently living in, voluminous sorts of high velocity data.
Big Data Analytics: A Literature Review Perspective	Sarah Al-Shiakhli	2019	Supervised, unsupervised, semi-supervised, reinforcement learning.,	Thus usefully focus on big data analytics challenges with regard to security and privacy issues..

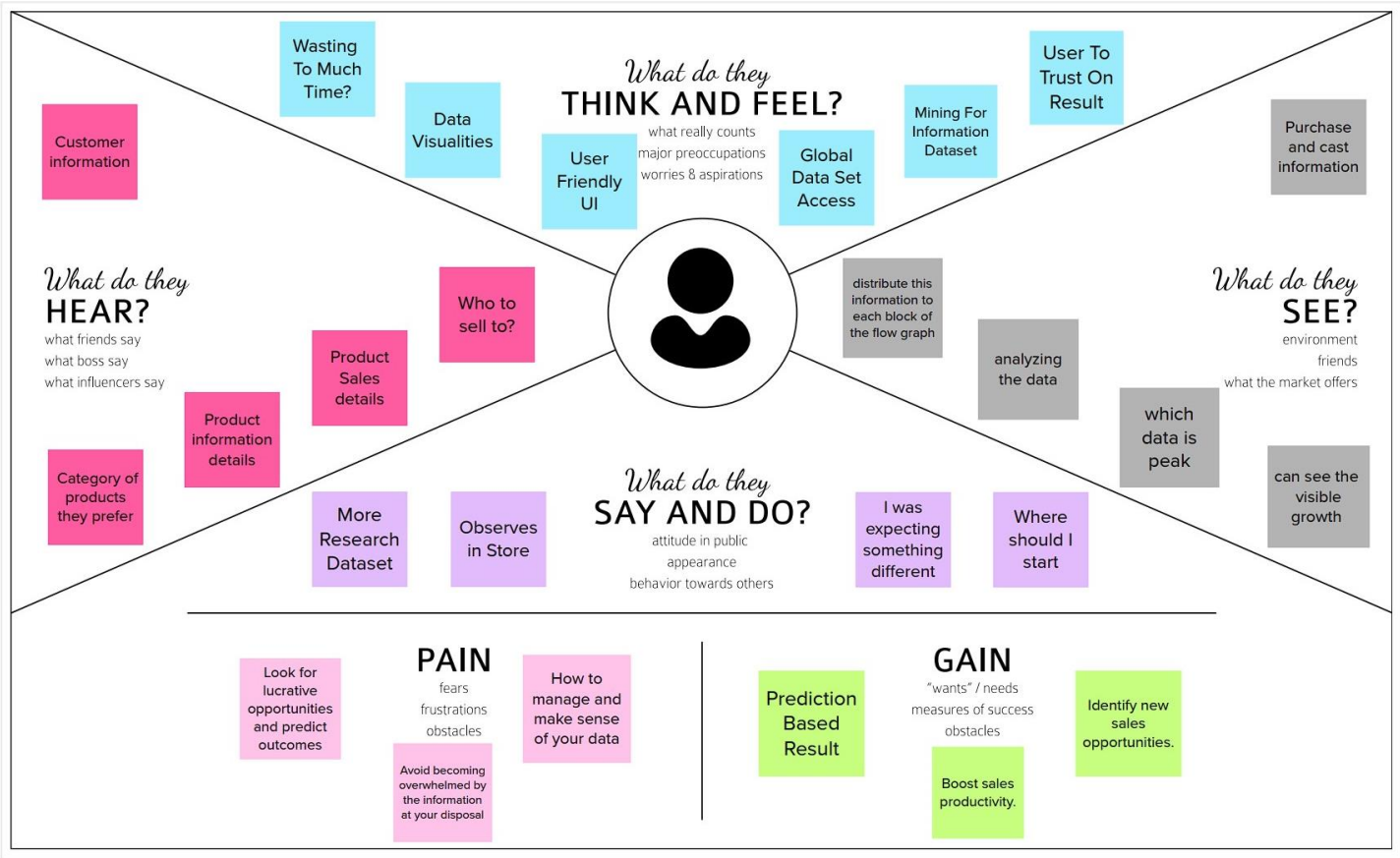
2.3Problem Statement Definition

The overall purchase power of the consumer and also sales capacity of company. Unavailability of products equally between the consumers. There is no proper distribution of products among the customer The customers are not getting the products they prefer. By hearing out to the consumers and collecting their user preference data. Data analytics and data visualization is used for this.

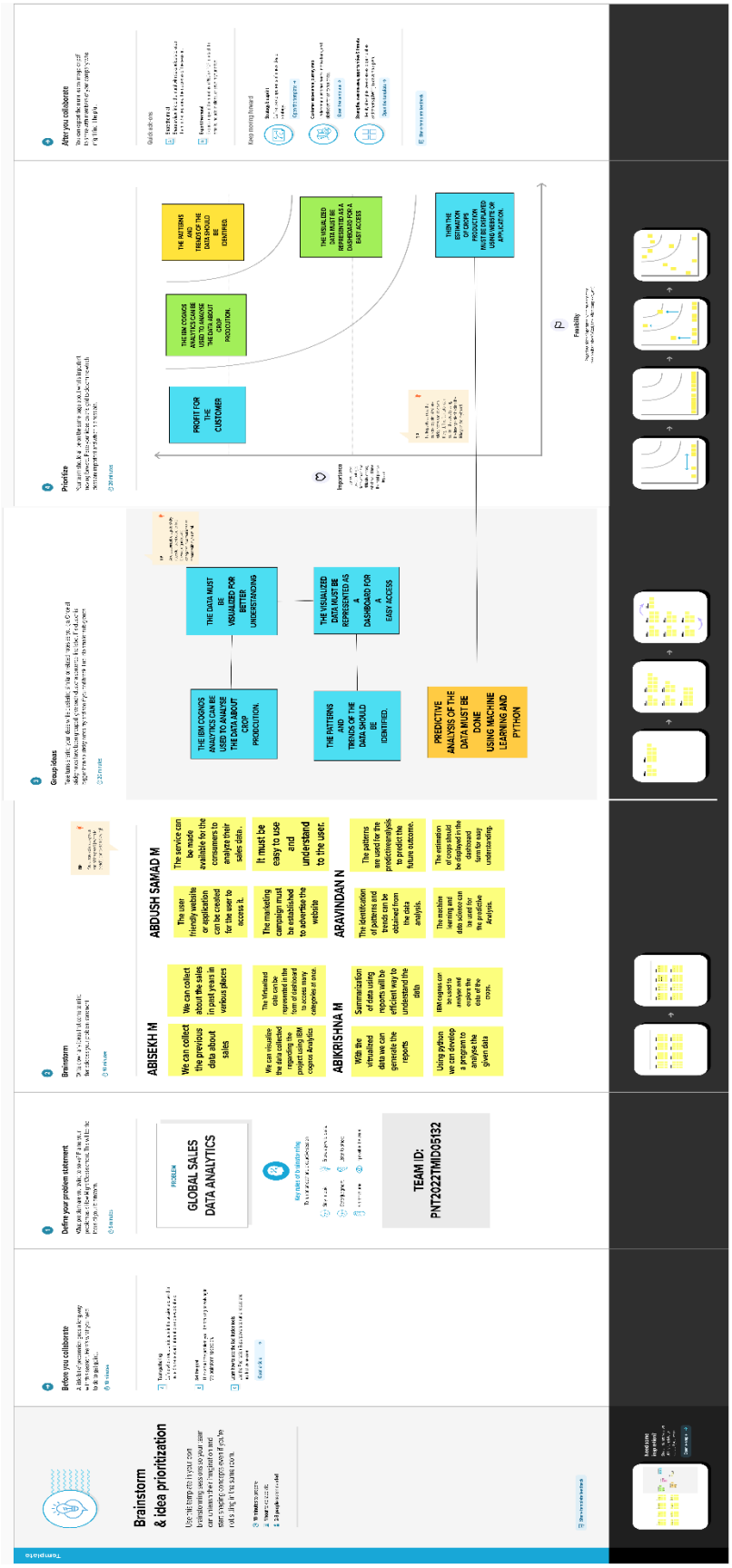
Who does the problem affect?	The overall purchase power of the consumer and also sales capacity of company.
what are the boundaries of the problem?	Unavailability of products equally between the consumers.
What is the issue?	There is no proper distribution of products among the customer The customers are not getting the products they prefer.
When does the problem occur?	If sufficient knowledge of purchase history this cause this issue.
Where does this problem occur?	It occurs to retailers who does not hear to consumers.
Why is it important that we address this problem?	So the purchasing power will increase and beneficiary for both retailers and consumers.
What solution to solve this issue?	By hearing out to the consumers and collecting their user preference data.
What methodology used to solve the issue?	Data analytics and data visualization is used for this.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



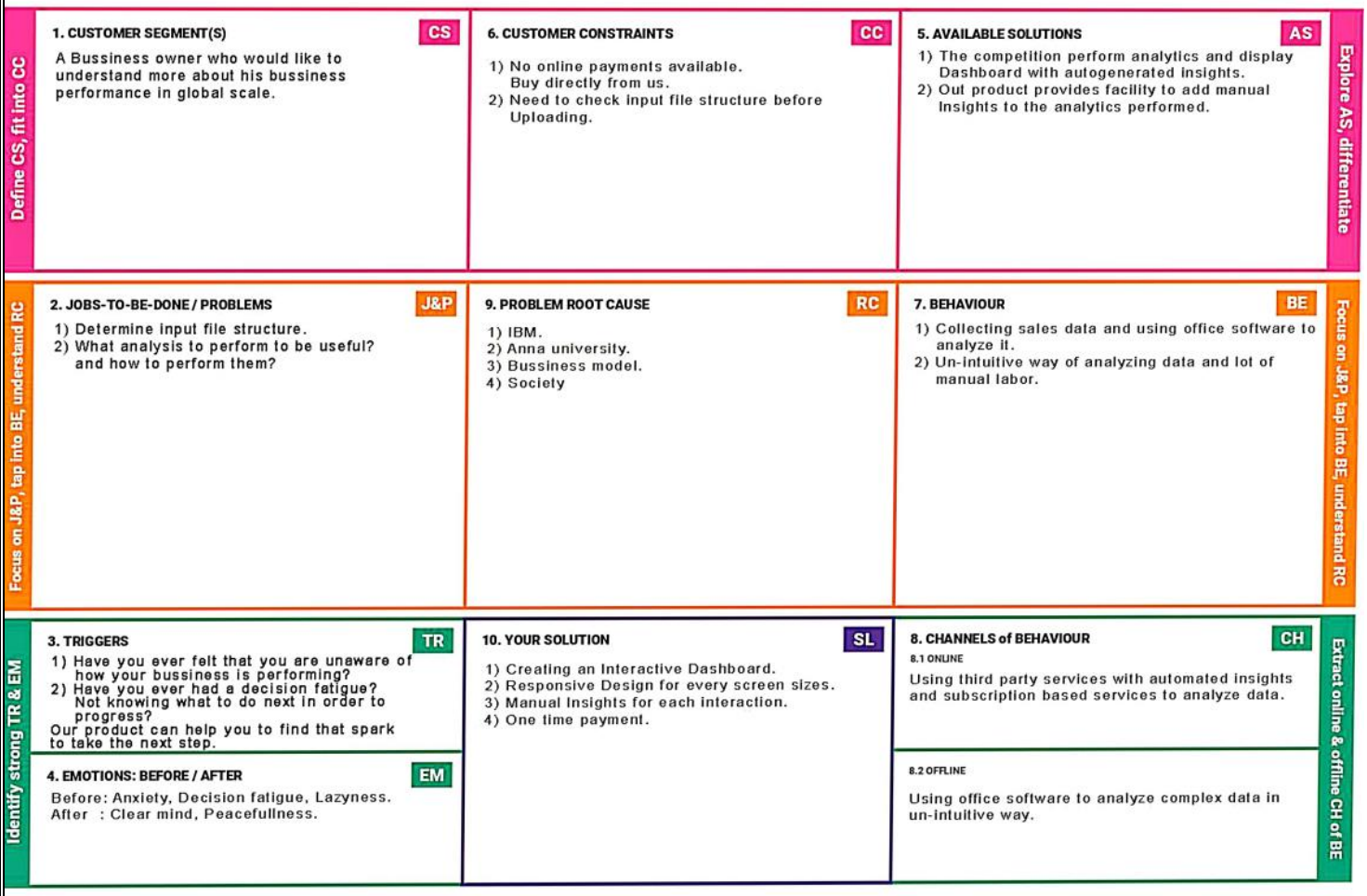
3.2 Ideation & Brainstorming



3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	The user needs a way to track and maintain overall sales data so that he can make more profit. The user needs a strategy to market the products so that it reaches all people through media. The user finds the key performance indicators so that he can boost the annual sales and reduces customer churn.
2.	Idea / Solution description	A good sales dashboard is the solution. We can monitor the sales by its geographic location and track order purchases. We can monitor the monthly sales and stocks retained on each product. Create and approve sales orders, track order purchases, improve sales tracking and optimization of goods delivery.
3.	Novelty / Uniqueness	Provides Real-Time Data, Can Help the Team Set Goals, Gives a Clear Overview of Sales Activity, Allows for the Identification sales of growth opportunities and Identifies opportunities for improvement.
4.	Social Impact / Customer Satisfaction	Ensures sustainability in global market. Creates meaningful change in business approach. Trying to attract customers of all range.
5.	Business Model (Revenue Model)	A Sales dashboard enables direct insight into your revenue driving force, allowing yo u to plan, implement and improve with data-based decisions.
6.	Scalability of the Solution	The great thing about Sales Analytics is that it gives you answers, and you need to ask the right questions. With accurate insight into current customers, a higher retention rate, as well as increasing revenue, can be achieved. Having real-time insight into increasing and decreasing customers will allow your Sales Team to focus on the right clients at the right time and efforts are driven towards suitable clients. A Sales Dashboard helps you visualize your Sales data, which is helpful for efficient decision-making and analysis performance.

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Collects Data	Providing CSV file Authentic Datasets
FR-4	Cleans the given Data	Prepares data for EDA purpose
FR-5	Visualisation of Data	Identifying trends in given data Accurate visualisation of provided numbers
FR-6	Create Dashboard	Analysation of the dataset's Key performance indicator
FR-7	Reporting	The reporting function helps users have complete control over their business. The real-time reporting collects current information and displays the data on an intuitive user interface.

4.2 **Non-Functional requirements**

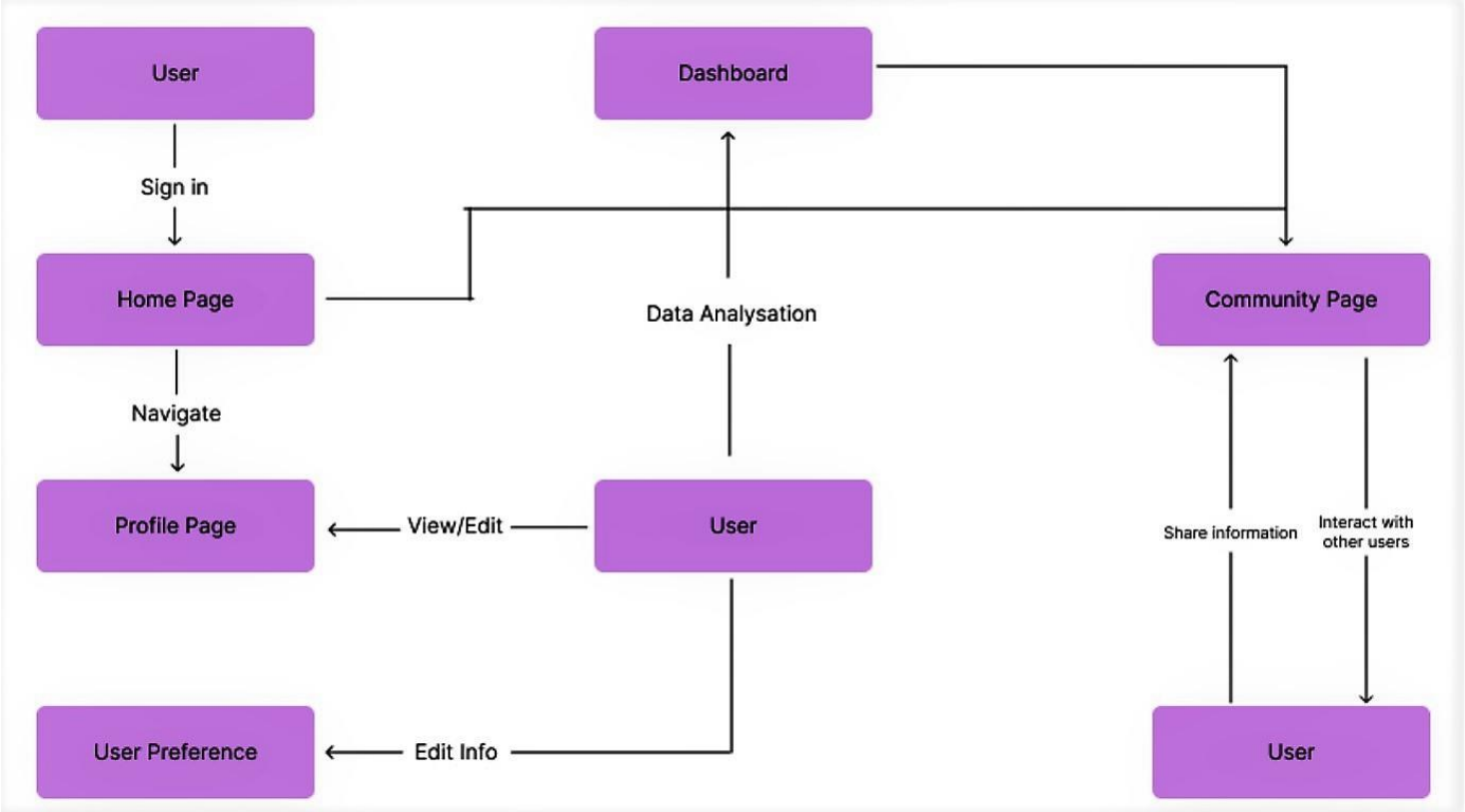
Following are the non-functional requirements of the proposed solution.

NFR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	It should be easier to understand the insights for the customers.
NFR-2	**Security**	The data is protected from unauthorized access.
NFR-3	**Reliability**	App could be run offline while server maintenance takes place. Server traffic would not be an issue.
NFR-4	**Performance**	Requires minimum system requirements, hence could be accessible in many devices with faster loading time.
NFR-5	**Availability**	Server is online 24/7 hence users could use the app at any time. App will work offline as well/
NFR-6	**Scalability**	Dashboards/Templates are very much Scalable, the user can modify the metrics whenever they want.

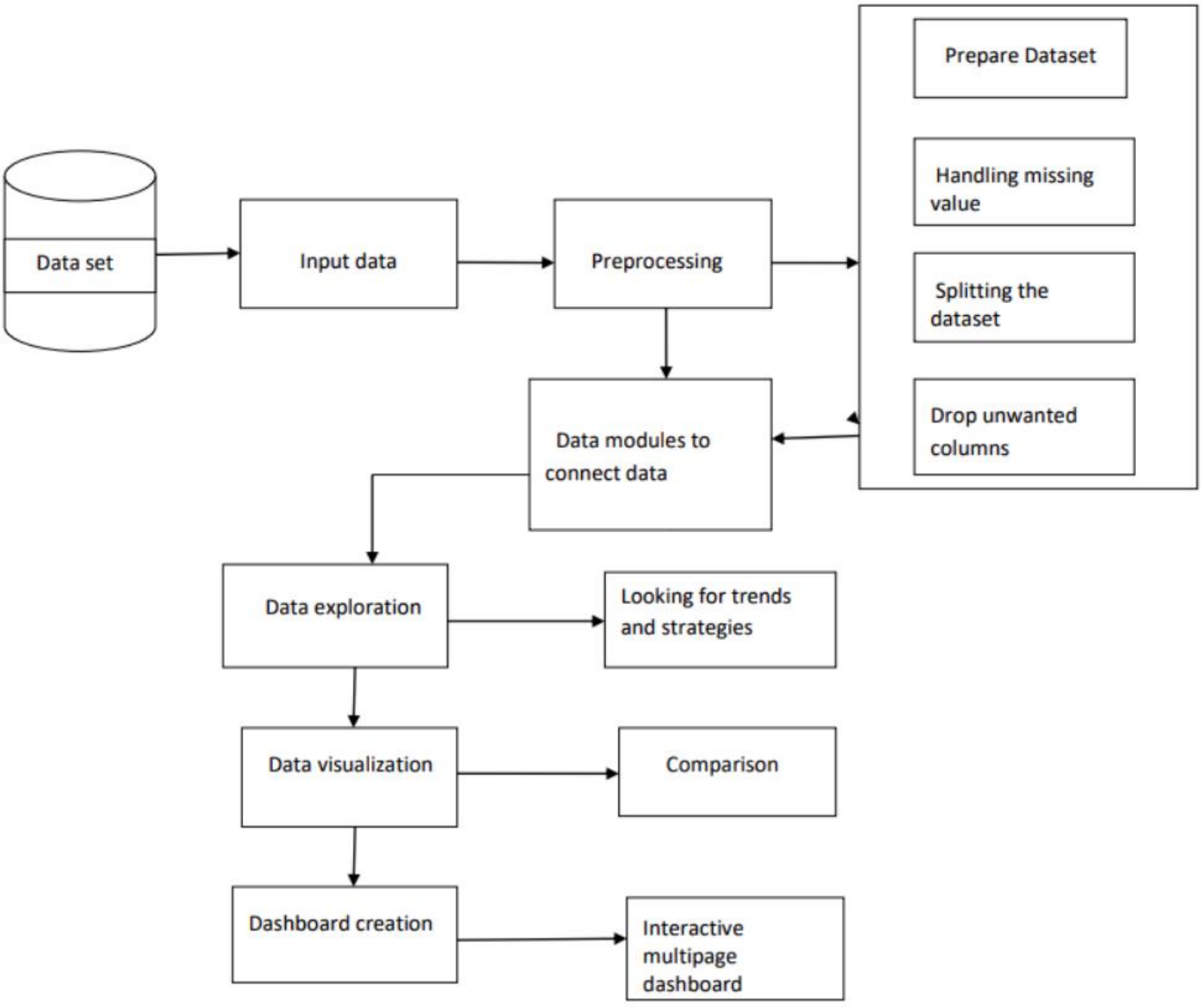
5. PROJECT DESIGN

5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution and technical architecture.



5.3 User Stories

Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
	Login	USN-3	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-4	As a user, I can enter data to conduct to conduct business analysis to make business decisions		High	Sprint-1
Customer Care Executive			As a Customer Care Executive, I can answer users' queries		High	Sprint-1
Administrator			As admin, can make changes to the interface according to the needs		High	Sprint-1

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	Abikrishna M, Abishek M
Sprint-2	Dashboard	USN-2	As a user, I can enter my sales information to clean it up and get it ready for analysis, and I can also locate my information to perform a business analysis.	3	High	Aravindan N, Abdush Samad M
Sprint-3	Customer Care	USN-3	As a user, I can enter my sales information to clean it up and get it ready for analysis, and I can also locate my information to perform a business analysis.	2	Low	Aravindan N, Abikrishna M, Abdush Samad M

Sprint-4	Administrator	USN-4	As an admin, I can modify the user interface to meet the needs of the users.	3	High	Aravindan N, Abdush Samad M, Abikrishna M, Abishek M

6.2 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	7	6	24 Oct 2022	14 Oct 2022	7	29 Oct 2022
Sprint-2	5	6	31 Oct 2022	15 Nov 2022	5	05 Nov 2022
Sprint-3	3	6	07 Nov 2022	18 Nov 2022	3	12 Nov 2022
Sprint-4	5	6	4 Nov 2022	19 Nov 2022	5	19 Nov 2022

Velocity:

The average velocity (AV) per iteration unit (story points per day):

Sprint1: AV=Sprint duration/velocity= 7/6=1.17

Sprint 2: AV=Sprint duration/ velocity = 5/6 = 0.84

Sprint3: AV=Sprint duration/velocity= 3/6=0.5

Sprint4: AV=Sprint duration/velocity= 5/6=0.84

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature

We have used many visualizations type like

- For comparison:
  - Bar
  - Bullet
  - Line and column
  - Radar
  - Word cloud
- Parts to whole:
  - Pie chart
  - Tree map
- Trend:
  - Box plot
  - Line
  - Line and column
- Relationships
  - Scatter
- Tables and summary
  - Crosstab
  - Summary
- Geospatial
  - Legacy map
  - Map

7.2 Database Schema (if Applicable)

Build the following visualizations

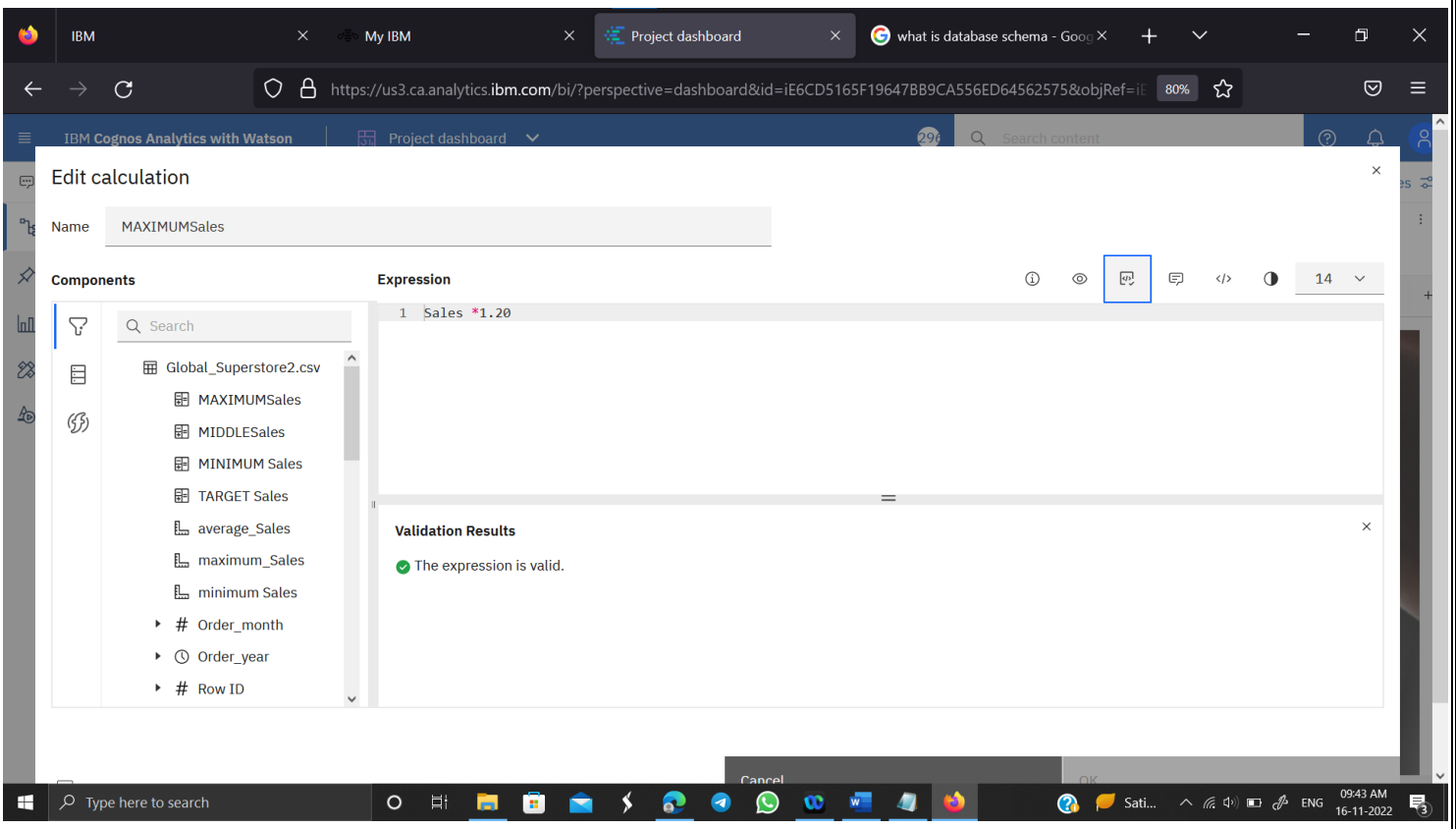
1. Global Stuperstore_Data Upload.
2. Global Superstore_DataPrep.
3. Date Calculations and Navigation path.
4. Segment wise Sales, Profit and Qty.
5. Use Pie to showcase Sales by Order Priority and Sales by Market.
6. Use a TreeMap to present Sales by Sub-Category
7. Using a Bar chart present Sales by Region by the Sales Order.
8. Present Regional Sales using Map Country points -- Showcase Top 10 countries.
9. Present Sales (Bar), Profit (line) by Sub-Category using Line and Column Chart.
10. Sales vs Profit Scatter Plot with Sub-Category points and Region in Colour.
11. Sales and Profit Forecast by Month Country as Region and Region as Filter.
12. Sales vs Profit forecast by Month by Order Priority.
13. Show the Min, Max, and Avg Sales by Sub-Category using the Box plot.
14. By setting a 10% extra Target for Sales Present Segment-wise Sales use Bullet Chart.
15. Present Sales using Hierarchy Bubbles by Market / Region.
16. Using a Legacy Map Present Sales vs Profit by Country / Region.
17. Showcase Quantity Sold by Radar Chart across various Regions.
18. Present Monthly Sales by Sub-Category using Waterfall chart.

19.Present Sales Vs Profit of Countries by Word Cloud.

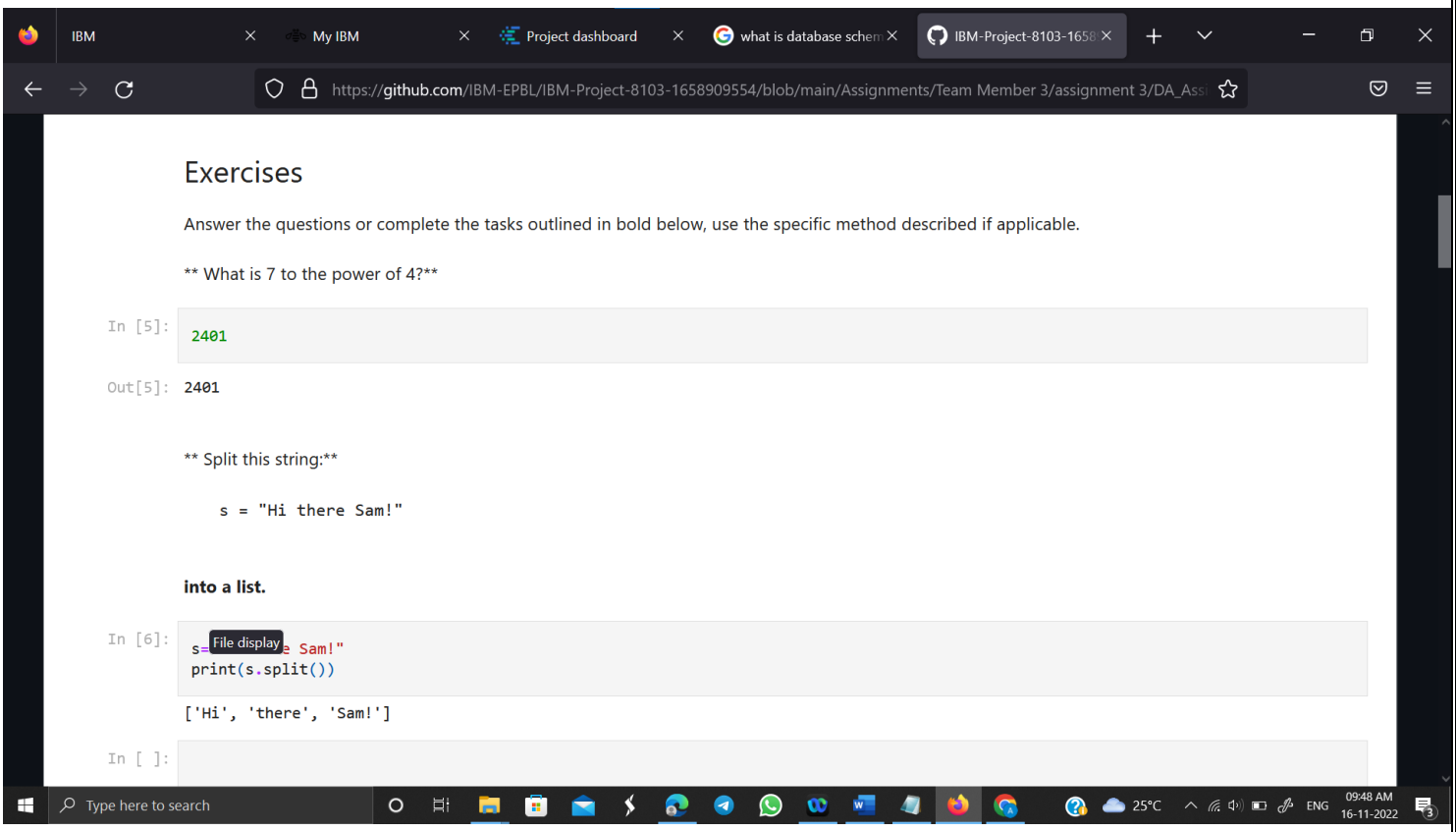
20.Sales dashboard with Summary Cards.

8. TESTING

8.1 Test Cases



Every expression is validated before calculation.



9. RESULTS

9.1 Performance Metrics

This dashboard is created to understand a few things like, Customer Analysis and Product Analysis of the Global Super Store. This can be achieved by hearing out to the consumers and collecting their user preference data So that purchasing power will increase and beneficiary for both retailers and consumers.

10. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- purchasing power will increase and beneficiary for both retailers and consumers.
- IBM Cognos analytics helps in building the dashboard and creating the exploration.

DISADVANTAGES:

- A little bit confusing to choose the type of exploration.

11. CONCLUSION

Creating this dashboard will help understand customers will So that purchasing power will increase and will be beneficiary for both retailers and consumers.

12. FUTURE SCOPE

- This dashboard is a responsive dashboard, so as we update the csv file uploaded in the IBM Cognos dashboard updates automatically so that this dashboard can be utilized in future also.
- This dashboard is also having forecast exploration which enables to predict future sales.

13. APPENDIX

GitHub: **IBM-Project-8103-1658909554**

DEMO-VIDEO-LINK:

https://drive.google.com/file/d/1Q11t0MP7sUos8AH0FN5pGfUZApy2ODkl/view?usp=share_link