GLOBAL SALES DATA ANALYTICS

A PROJECT REPORT

Submitted by

Pragatheeswari V	95191902074
Navina S	95191902067
Nivethika M	95191902069
Shyamala R	95191902096

In partial fulfillment for the award of the degreeOf

BACHELOR OF ENGINEERING

IN ELECTRONICS AND COMMUNICATION ENGINEERING

PSR ENGINEERING COLLEGE

SIVAKSI -626140

(An Autonomous Institution, Affiliated to Anna University, Chennai)

Table of Content

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 into a store randomly and buy anything you want. So, this project is done to try to understand a few things like Customer
Analysis and ProductAnalysis of this Global Super Store.If you want to achieveyour salesgoals month after month, then guesswork and intuition aren't your best friends. You need to perform strategic sales analysis and get cold, hard data.

1.2 PURPOSE

By the end of this Project, you will :

- Know fundamental conceptsand 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). Regular sales data analysis provides an understanding of the products that your customers are buying and helps you dissect why they are behaving in a certain way. You can also find patterns in your lead conversions and drop offs. All of these aspectsenable you to optimize you.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEMS

- The huge blast of information and Internet gadgets has prompted a fastapproach to Big Data in the later past. The administration industry which is a noteworthy client for these Big Data applications will promp treal change to the conveyance process and new bits of knowledge into utilization examples and work processes, which thusly will help with new worldwide conveyance models incorporating new innovations anddis persion of work comprehensively. The Service Industry will utilize Big Data for different choices making an information framework and making the work processmore ideal. The idea of large-scale manufacturing lead to the Industrial Revolution, likewise, Big Data is relied upon to drive new typesof financial movementin the Service industry with connectedhuman capital, achievinga new level of monetary action, development, and development.

- In the information era, enormous amountsof data have become availableon han d to decision-makers. Big data refersto datasets that arenot only big but also high in variety and velocity, which makes them difficult to handle using traditional tools and techniques. Due to the rapid growth of such data, solutions need to be studied and provided inorder to handle and extract value and knowledge from these datasets. Furthermore, decision-makers need to be able to gain valuable insightsfrom such varied and rapidly changingdata, ranging from daily transact ions to customer interactions and social network data. Such value can be providedusing big data analytics, which is the applicationof a dvanced analytics techniques on big data. This paperaims to analyzesome of the different analytics methods and tools which can be appliedto big data, as well as the opportunities provided by the application ofbig data analytics in various decision domains.

- In the modern era of higher education, it is exceptionally challenging for teachersto counsel studentsin terms of academic matters.Teach ershave abundant data related to different aspects of students but derivingappropriate insights from them is very challenging. Predictive analytics plays an important role to cope with such challenges. However, for effective predictive analytics, numerous factors must be considered such as the selection of proper academic and social variables, the appropriate volumeof the data, and qualityof the data, and the adoptionof suitable predictive algorithms. This paper deals with an exhaustive literature survey of predictive analytics in the educational domainsp ecifically for higher education. The paper also discussed the analysisof the existingliterature survey and identified research gaps.

2.2 REFERENCES

□ A literature survey on big data analyticsin the service industry.(https://www.researchgate.net/publication/301720427_A_literature_survey_on_Big_Data_Analytics_in_Service_Industry).

□ Big Data Analytics: A literature reviewpaper. (https://www.researchgate.net/publication/264555968_Big_Data_Analytics_A_Literature_Review_Paper).

□ Iterature survey usingpredictive analytics for student counselling inhig her education.

□ (https://www.researchgate.net/publication/355481331_Literature_Survey_using_Predictive_Analytics_for_Student_Counselling_in_Higher_Education).

2.3 Problem Statement

Customer Problem Statement Template:

Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love.

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

Reference: https://miro.com/templates/customer-problem-statement/

Example:

I am	Describe customer with 3-4 key characteristics - *who are they?*	Describe the customer and their attributes here
I'm trying to	*List their outcome or "job" the care about - what are they trying to achieve?*	List the thing they are trying to achieve here
but	Describe what problems or barriers stand in the way – *what bothers them most?*	Describe the problems or barriers that get in the way here
because	Enter the "root cause" of why the problem or barrier exists – *what needs to be solved?*	Describe the reason the problems or barriers exist
which makes me feel	Describe the emotions from the customer's point of view – *how does it impact them emotionally?*	Describe the emotions the result from experiencing the problems or barriers

Problem Statement(PS)	**I am (Customer)**	**I'm tryingto**	**But**	**Because**	**Which makesme feel**
PS-1	MNC Owner	Increase myProductivity	Don't know how to increase	Future stockprediction isdifficult	Frustrated
PS-2	A person inmarketing team	Develop productmarketingplan and strategy	Difficult to find targeted customer	Time restrictionsmay apply	Stress

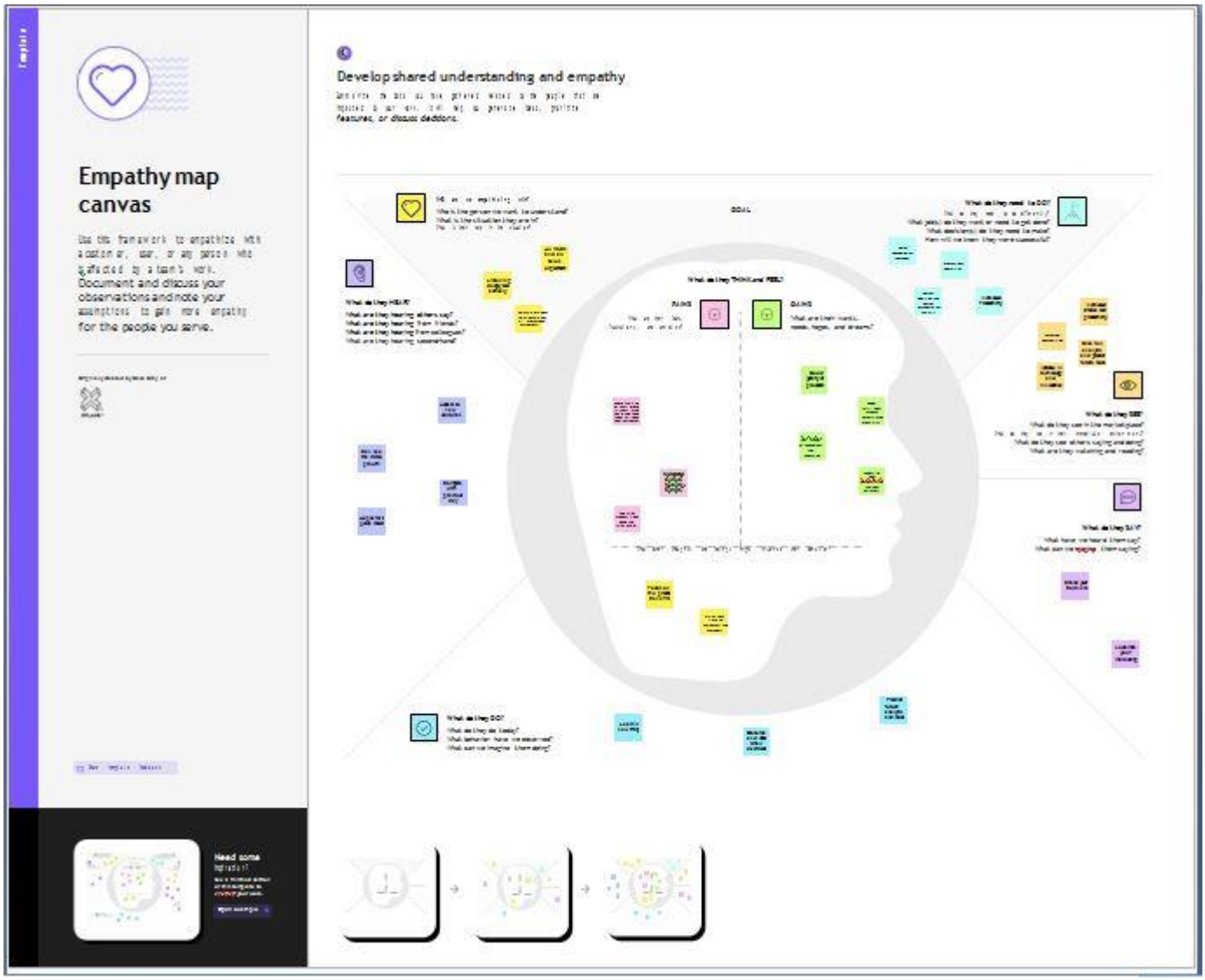
3.IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

◎ An empathy map is a simple, easy-to-digest visual that captures knowledgeabout a user's behaviors and attitudes.

◎ It is a useful tool to help teams better understand their users.

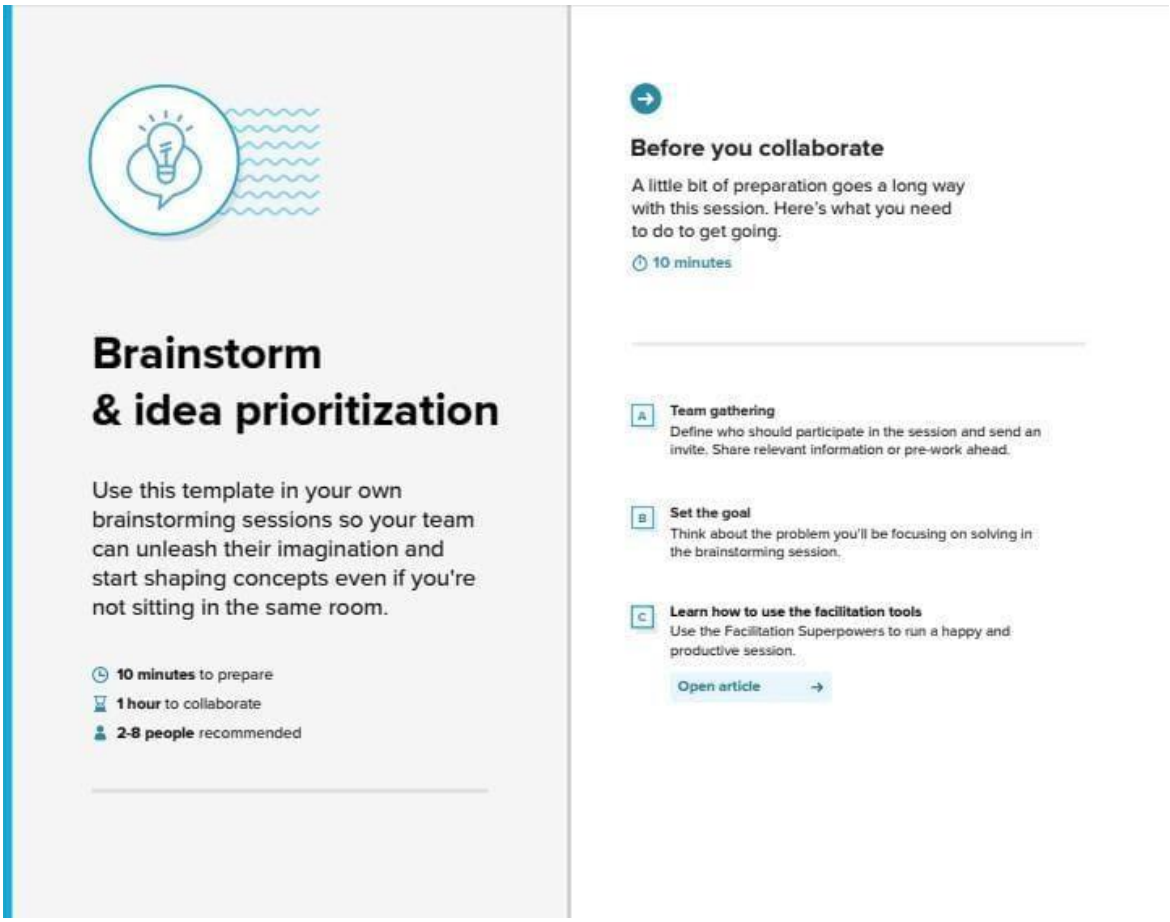
◎ Creating an effective solution requires understanding the true problemandth e person who is experiencing it. The exerciseof creating the map helpsparticipa nts consider things from the user's perspective along with his or her goals and challenges.



3.2 Ideation & Brainstorming

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Step-1: Team Gathering, Collaboration and Select the problem Statement



Step-2: Brainstorm, Idea Listing and Grouping

2

Brainstorm

Write down any ideas that come to mind that address your problem statement.

10 minutes

TIP

You can select a sticky note and hit the pencil [switch to sketch] icon to start drawing!

Pragatheeswari V

Dynamic and Real time	Growth analysis	Observe and make small decisions
Better Decision	Qualified Technical Support	Economic

Navina S

Better Prediction	Easy deliverable process in sales	Better Customer Support
Better Customer Support	Can predict loss	Affordable Subscriptions

Nivethika M

Simple UI	Sales analytics	Predictive analytics
Better for Budget	Visualization is made easier	Sales will productive

Shyamala R

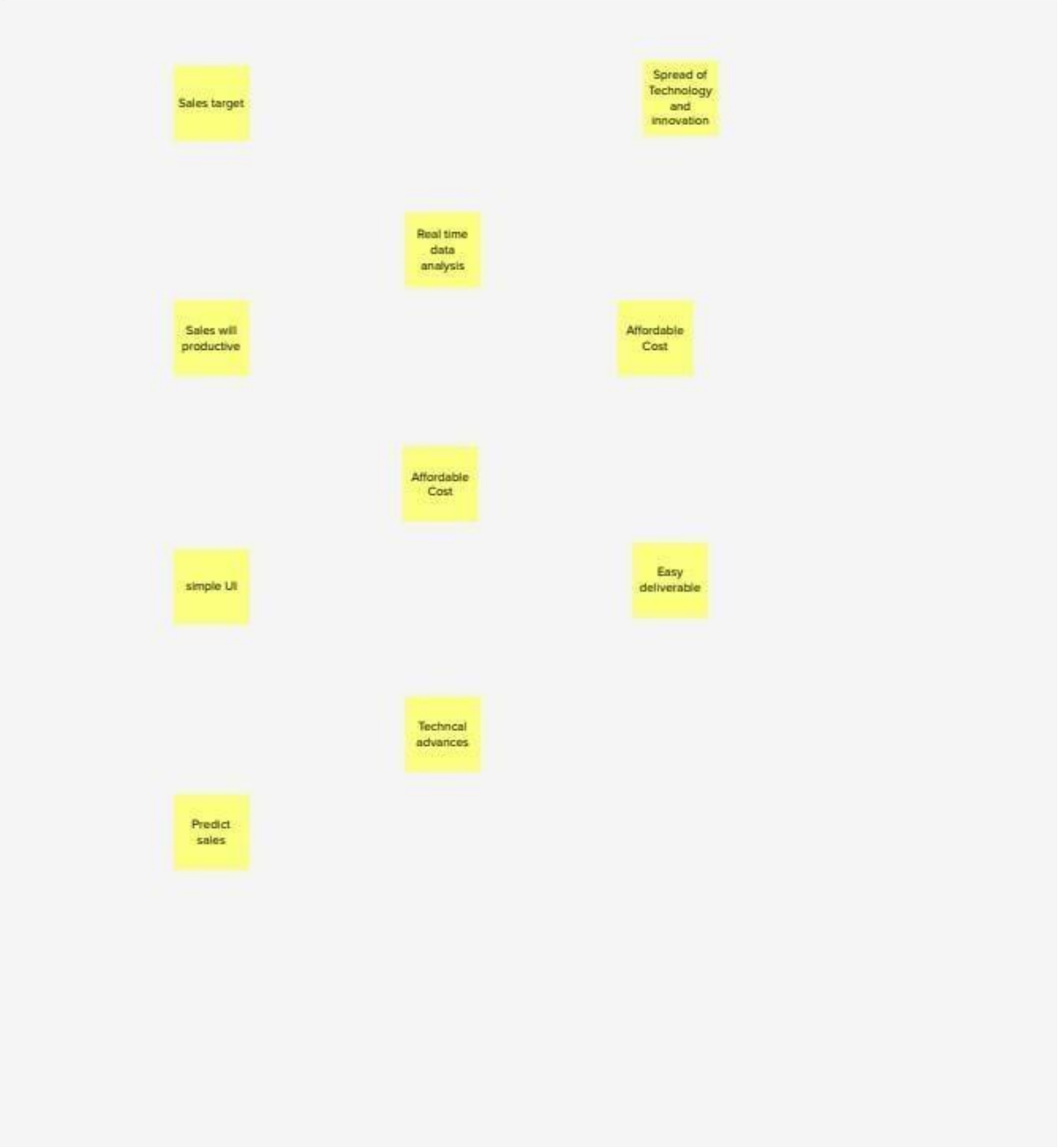
Increase Productivity	Development of global sales	Give full analysis
Satisfies target audience	Low bit rate	Predict future anlaysis

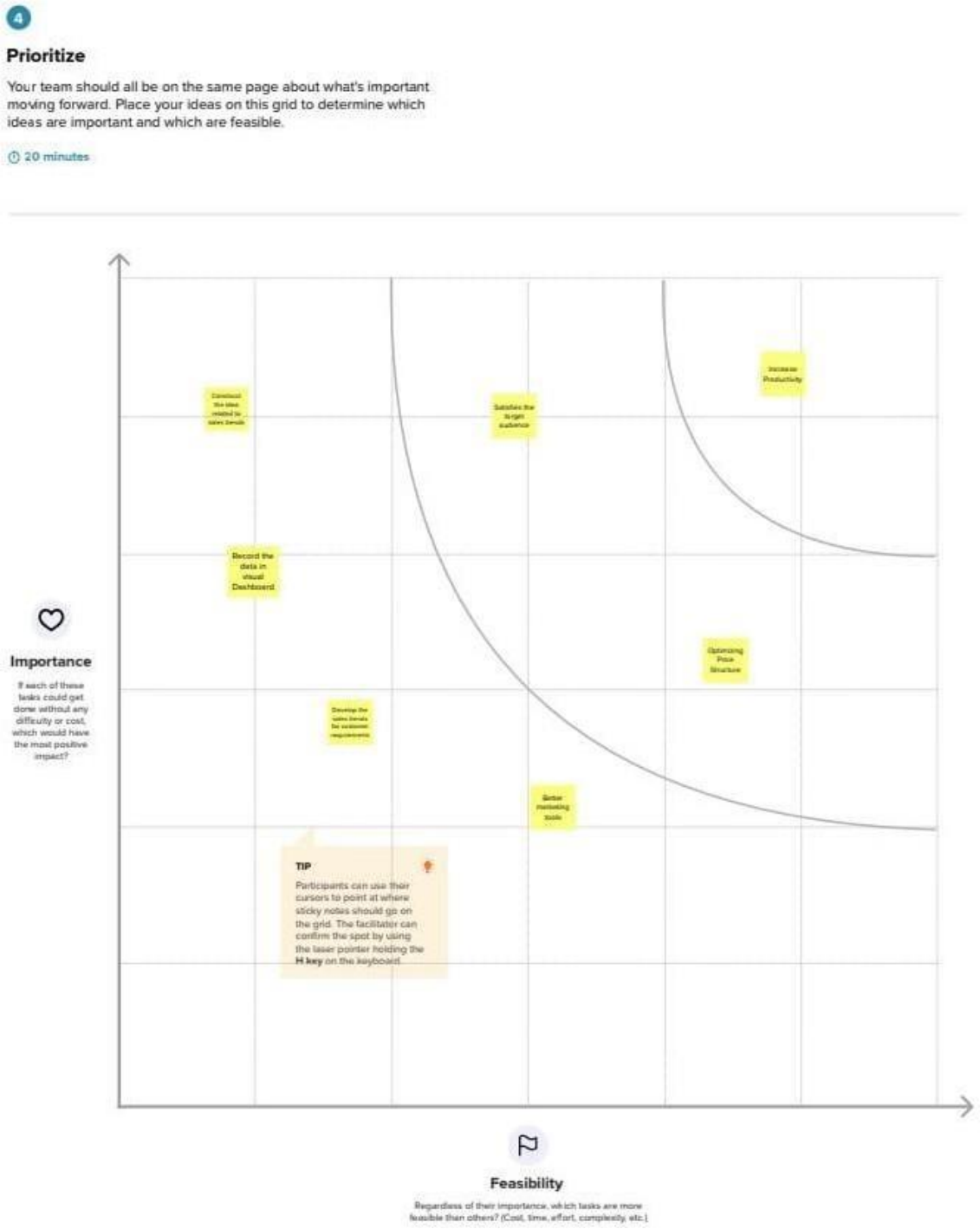
3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes



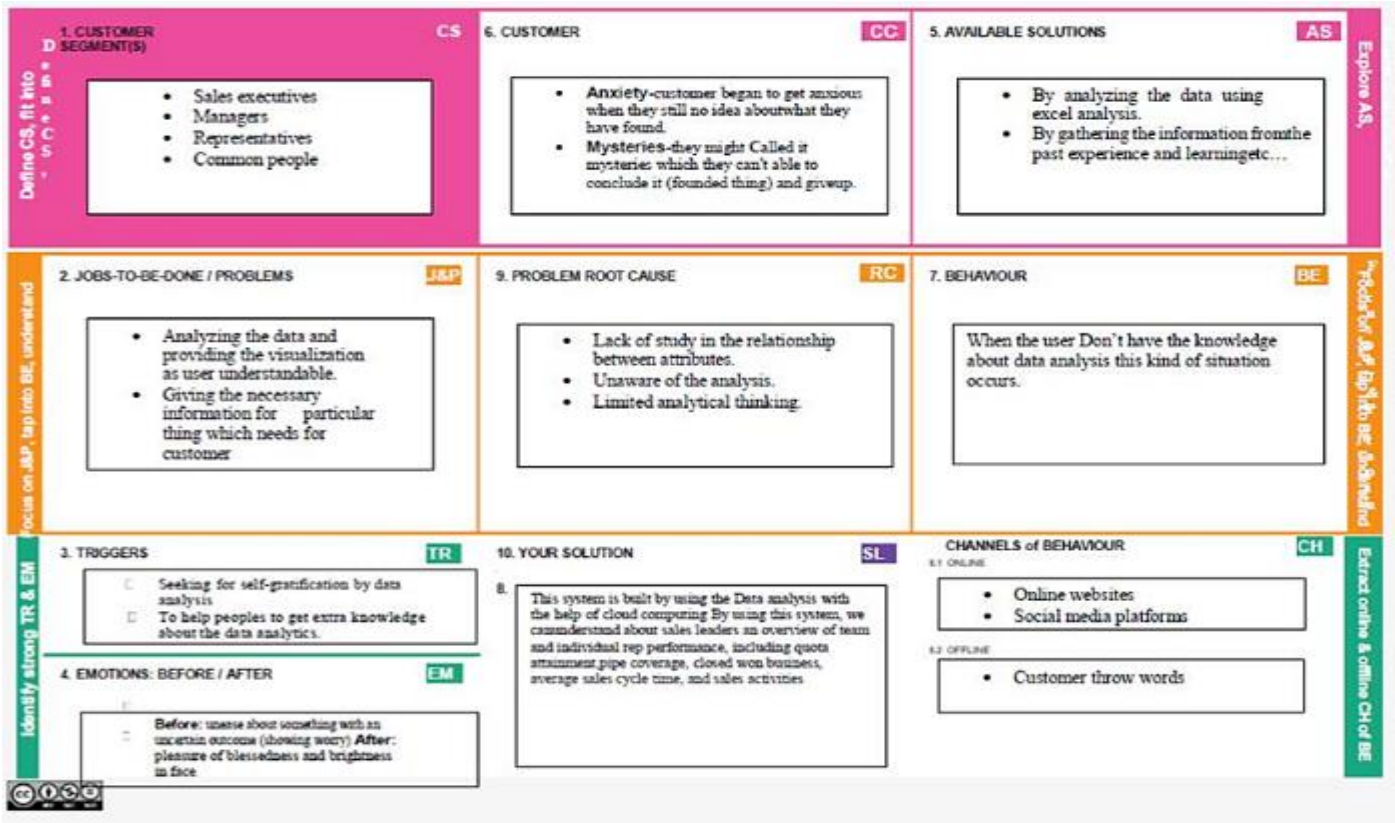


3.3 Proposed Solution

S.NO	Parameter	Description
1.	Problem Statement (Problem to be solved)	○ Online shopping websites get flooded with voluminous amount of sales data everyday. ○ Processing, analyzing and communicating ○ this data are a challenge. ○ Hence understanding these things like, customer Analysis and product analysis is essential.

2.	Idea / Solution description	○ A powerful and easy-to-use sales analytics tool that automates and visualizes sales trends to optimize business outcomes.
3.	Novelty / Uniqueness	○ Interactive Dash board and simple UI ○ Dynamic and real time analytics ○ AI based predictions and forecasting
4.	Social Impact / Customer S atisfaction	○ Visible profits driven by informed decisions ○ Optimize sales and marketing ○ Ability to react to competitor's strategies
5.	Business Model (Revenue Model)	Three tier pricing- Basic, Standard, Enterprise ○ Basic: Limited feature stargeting startups andindividuals. ○ Standard: Limited premium features. Targetcustomers- Medium Scale businesses. ○ Enterprise with all premiumfeatures targetedat Lar ge corporations.
6.	Scalability of the Solution	○ More B2B customer services can be providedalongside ○ Usable by all customer facing companies andstartups of all scale

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Login	Login via Email and password
FR-4	User uploading data	To store the data set through the cloud for the further analysis
FR-5	End user benefits	Getting higher state of efficiency and also to know entire data analysis

4.2 Non-Functional requirements

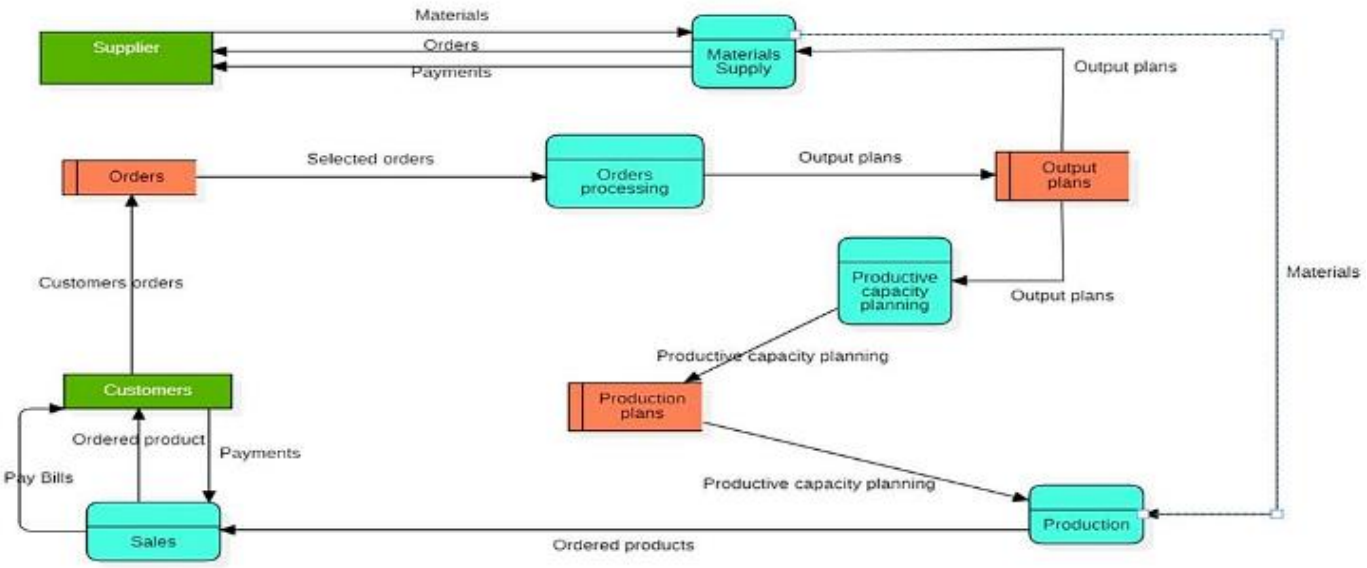
FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Optimized resources and it can be used by every person who needs to analyze their sales trend.
NFR-2	Security	The model is more securable because it has end to end encryption and the data are kept secure with high level authentication and authorization.
NFR-3	Reliability	It has high reliability based on development and analysis techniques.
NFR-4	Performance	It has high state of performance and efficiency
NFR-5	Availability	The model is available on all platforms and websites without the time constraint.
NFR-6	Scalability	The ability of a hardware and software parallel system to exploit increasing computing resources efficiency in the analysis of the (very)large datasets

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 systemrequirement graphically. It shows how data entersand leaves the system, what changes the information, and where data is stored

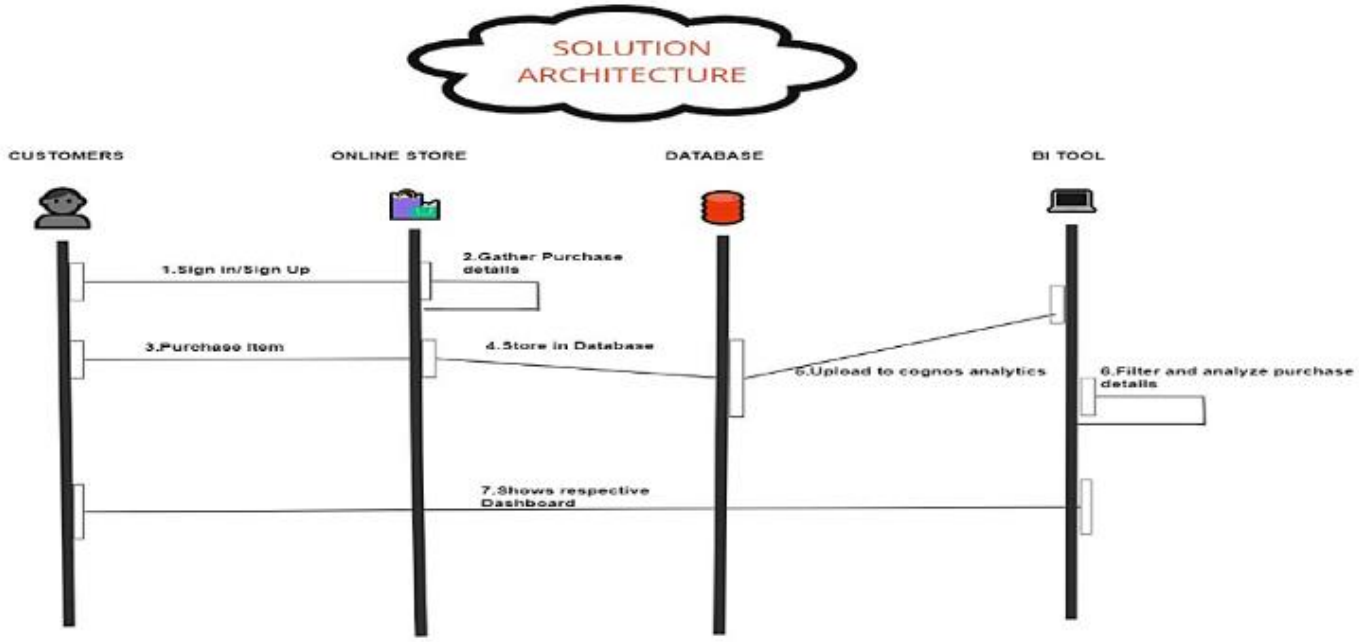
Structural Flow Diagram:



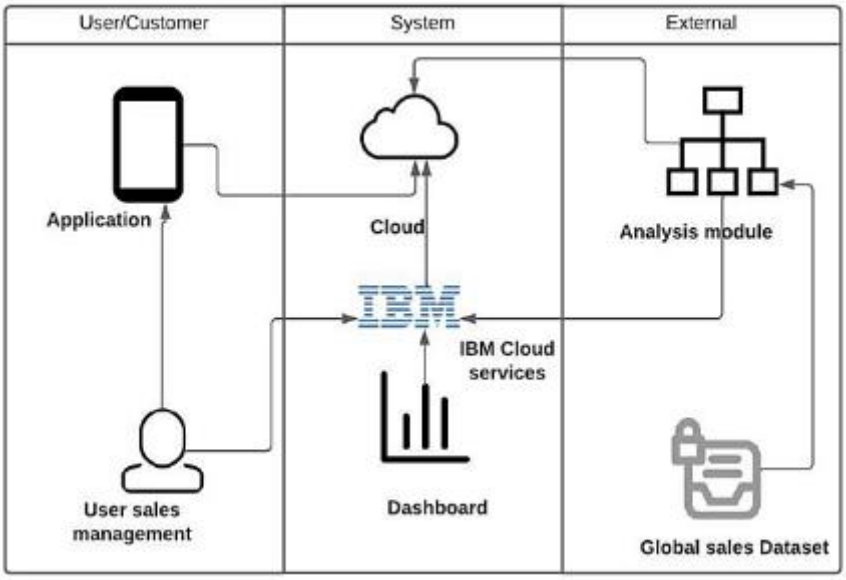
5.2 Solution & Technical Architecture

Solution architecture :

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:



Technical Architecture:



5.3 User Stories

User Type	Functional Requirement(Epic)	User StoryNumber	User Story / Task	Acceptance criteria	Priority	Release
	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
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a user, I can create the visualization by using the dashboard in the application		High	Sprint-3
Customer (Web user)	Login	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 and dashboard	High	Sprint-1
Customer care Executive	Chat box	USN-1	It can be used by easily access and responsible.	I can access by easily through application	High	Sprint-2
Administrator	Calling	USN-2	It can be used by easily access and responsible.	I can access by easily through application	High	Sprint-2
	Mail	USN-3	It can be used by easily access and responsible	I can access by easily through application	High	Sprint-1

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Product Backlog, Sprint Schedule, and Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Dataset exploration (Understanding the dataset)	USN-1	Explore the data and look for similarities, patterns and outliers and to identify the relationships between different variables. It enable to anticipate risks such as missing data, data duplicates and biases in the data. It will enable you to anticipate risks such as missing data, data duplicates and biases in the data.	2	Low	Pragatheeswari V, Navina S, Nivethika M, Shyamala R.
Sprint-2	Preparing the dataset for visualization	USN-2	In this, the dataset will be prepared for the next process by removing the unwanted values, null values, duplicate values and the missing values.	2	Medium	Pragatheeswari V, Navina S,
Sprint-3	Data visualization	USN-3	visualization used to represent information graphically, highlighting patterns and trends in data and to achieve quick insights.	3	High	Pragatheeswari V, Navina S, Nivethika M, Shyamala R.
Sprint-4	Creating dashboard, story and report	USN-4	From the visualization, we will be creating interactive dashboard will represent visual display of all the data, story and report.	3	High	Pragatheeswari V, Navina S, Nivethika M, Shyamala R.

6.2 Sprint Delivery Schedule

Project Tracker, Velocity & Burndown Chart:

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	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	30	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	30	19 Nov 2022

Velocity:

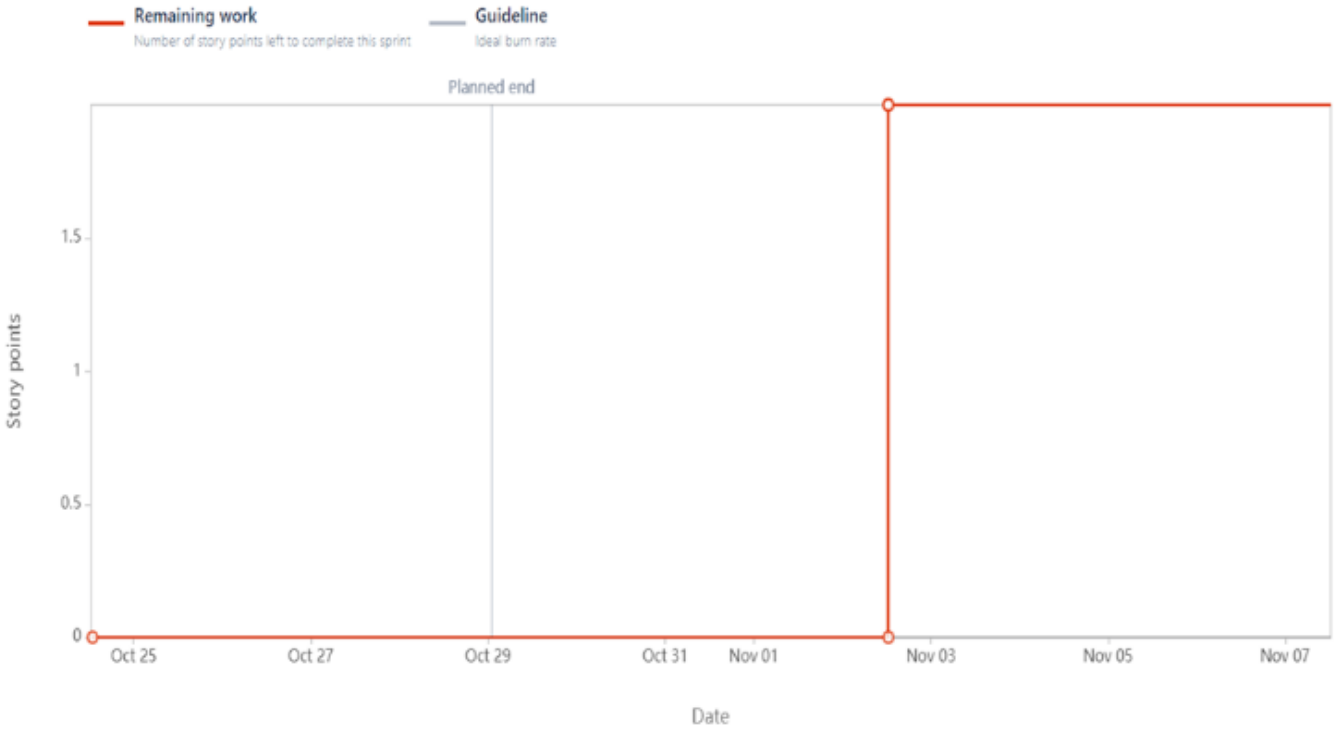
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \text{Sprint Duration}/\text{Velocity} = 20/10 = 2$$

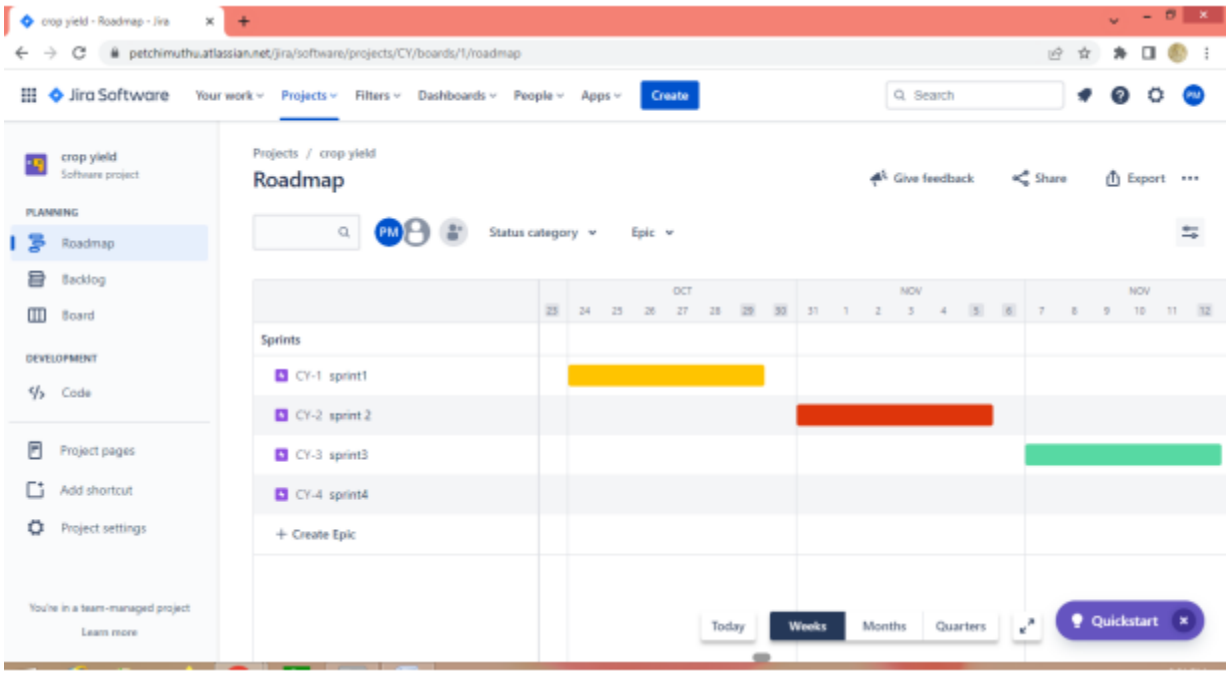
Sprint 1: AV = 20/6 = 3.333
Sprint 2: AV = 20/6 = 3.333
Sprint 3: AV = 30/6 = 5
Sprint 4: AV = 30/6 = 5

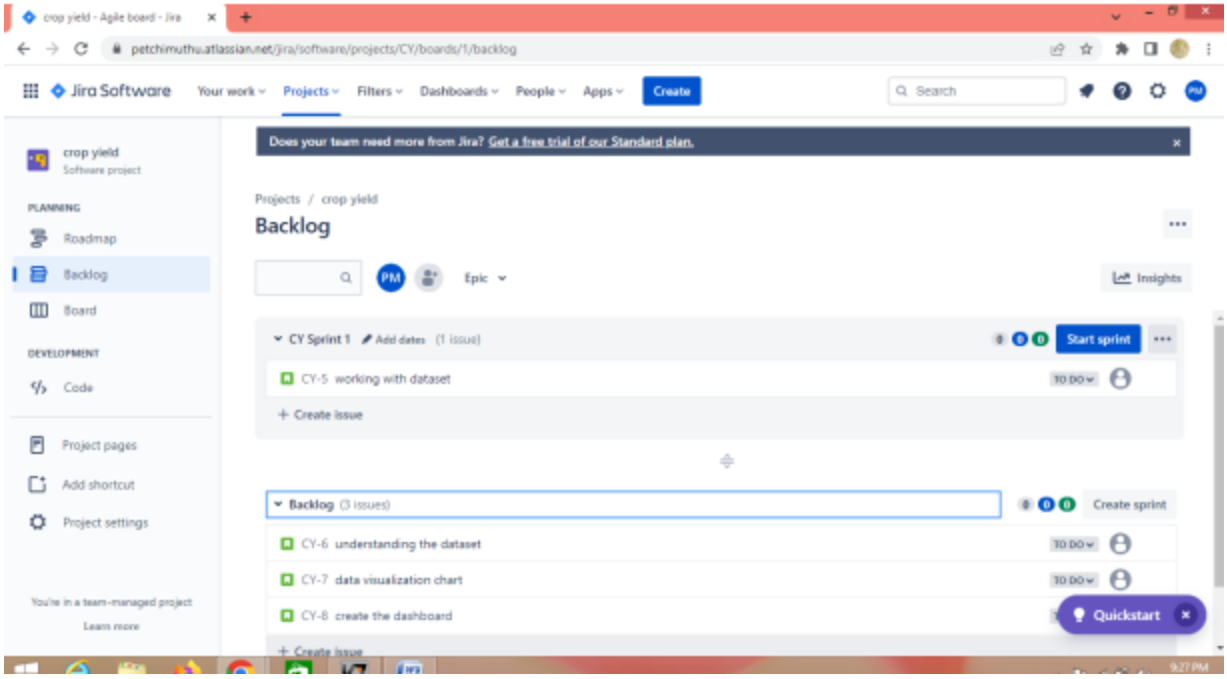
Burndown Chart:

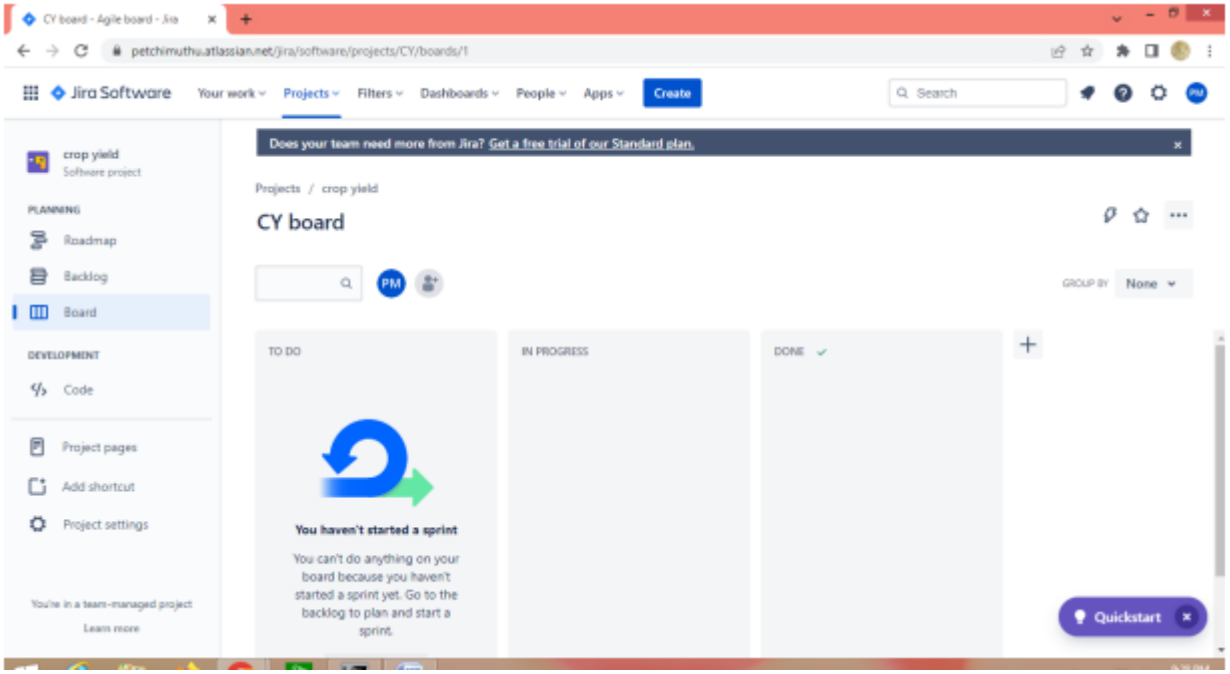
A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.



6.3 Reports from JIRA




crop yield - Agile board - Jira
petchimuthu.atlassian.net/jira/software/projects/CY/boards/1/backlog
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
crop yield
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / crop yield
Backlog
Epic
Insights
CY Sprint 1
Add dates
(1 issue)
Start sprint
CY-5 working with dataset
TO DO
Create issue
Backlog (3 issues)
Create sprint
CY-6 understanding the dataset
CY-7 data visualization chart
CY-8 create the dashboard
Quickstart
Create issue


CY board - Agile board - Jira
petchimuthu.atlassian.net/jira/software/projects/CY/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
crop yield
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / crop yield
CY board
GROUP BY
None
TO DO
IN PROGRESS
DONE
You haven't started a sprint
You can't do anything on your board because you haven't started a sprint yet. Go to the backlog to plan and start a sprint.
Quickstart

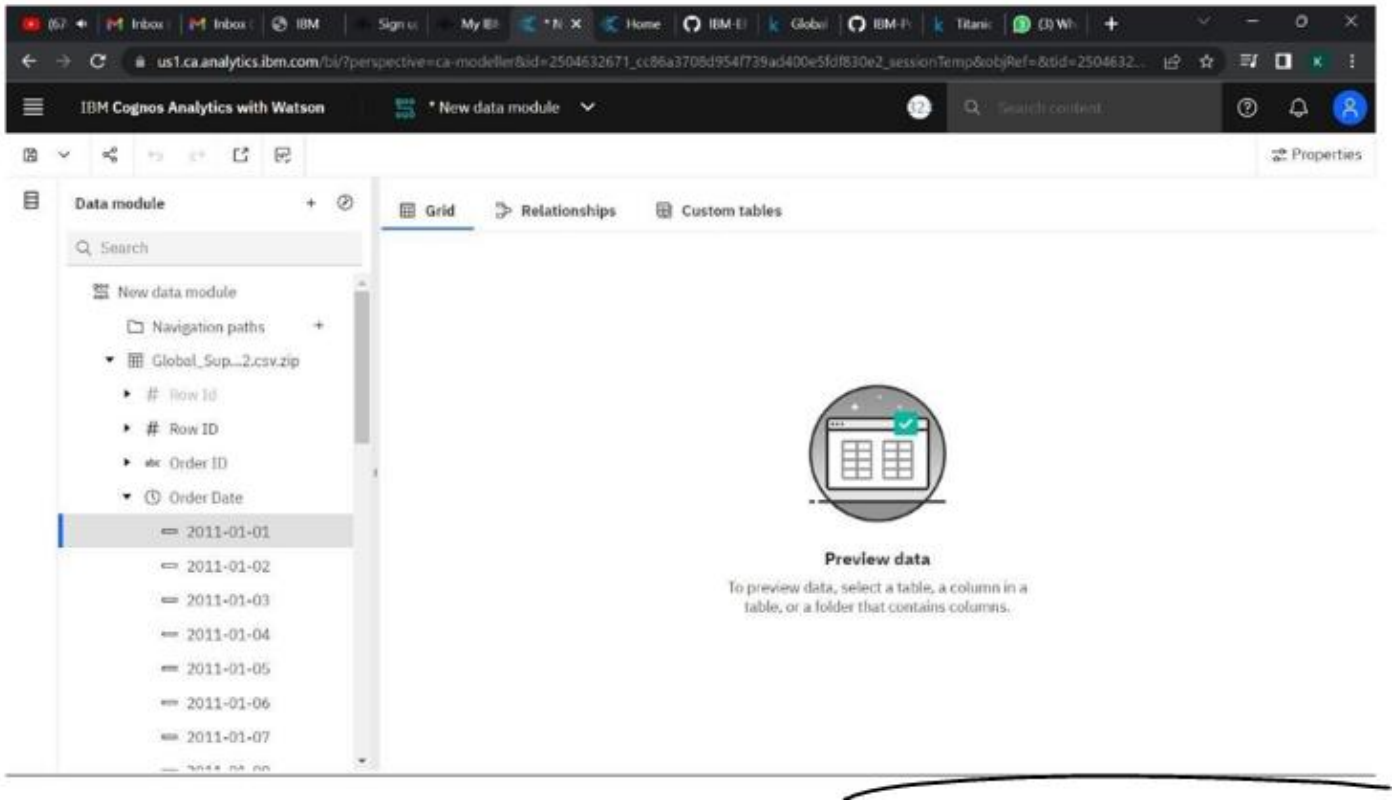
7. CODING & SOLUTIONING

7.1 FEATURE

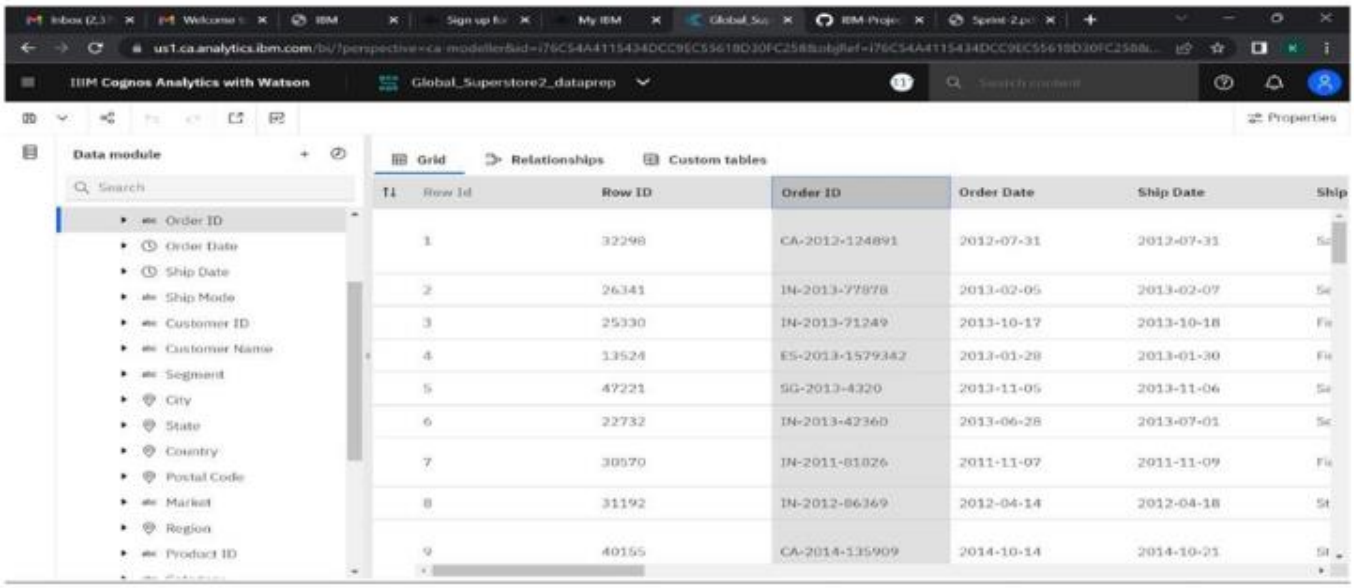
An interactive dashboard has been embedded

(https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FGlobal_Superstore2_datadashboard&action=view&mode=dashboard&subView=model0000018476edaeab_00000002).

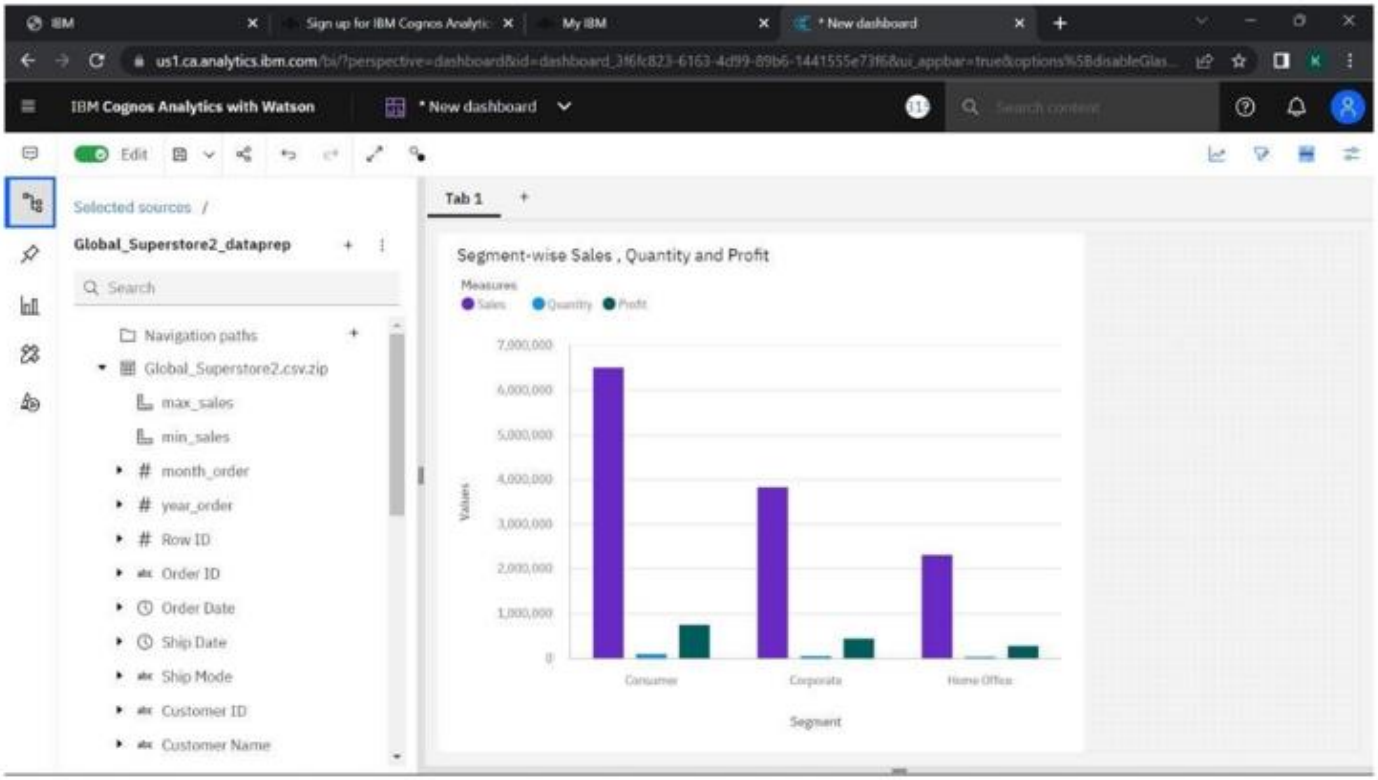
Global Superstore_Data Upload



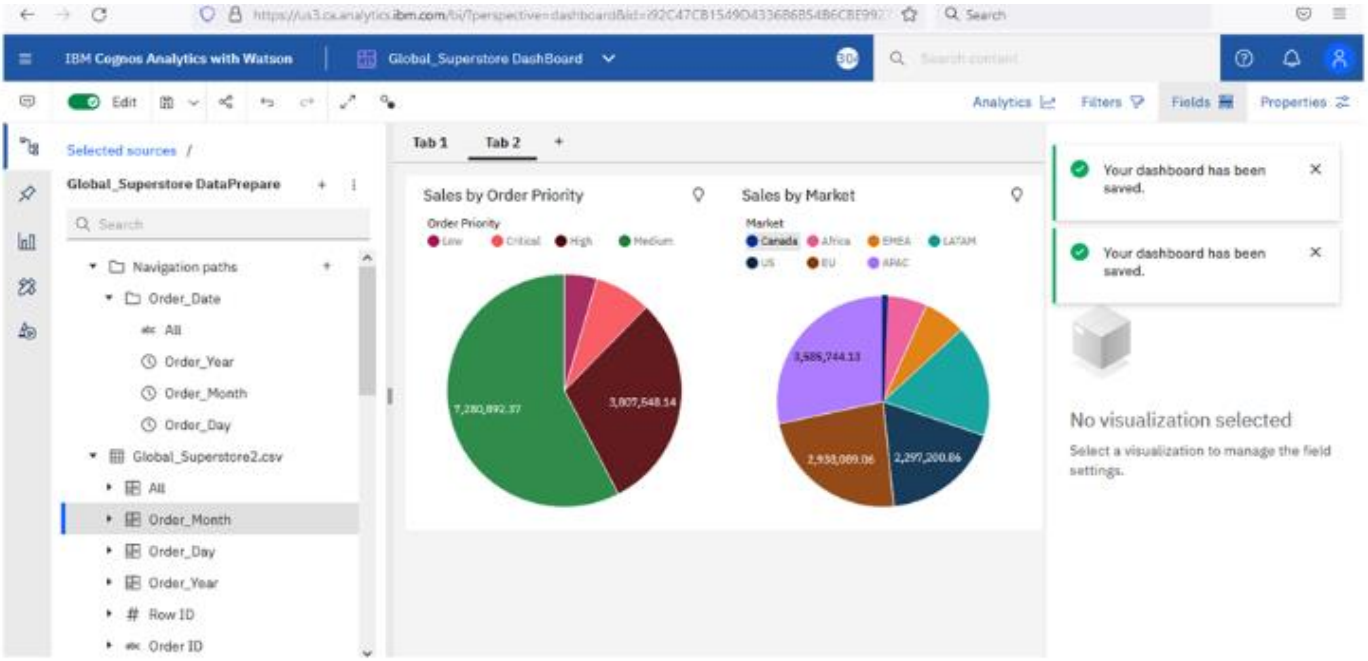
Global Super store Data Preparation:



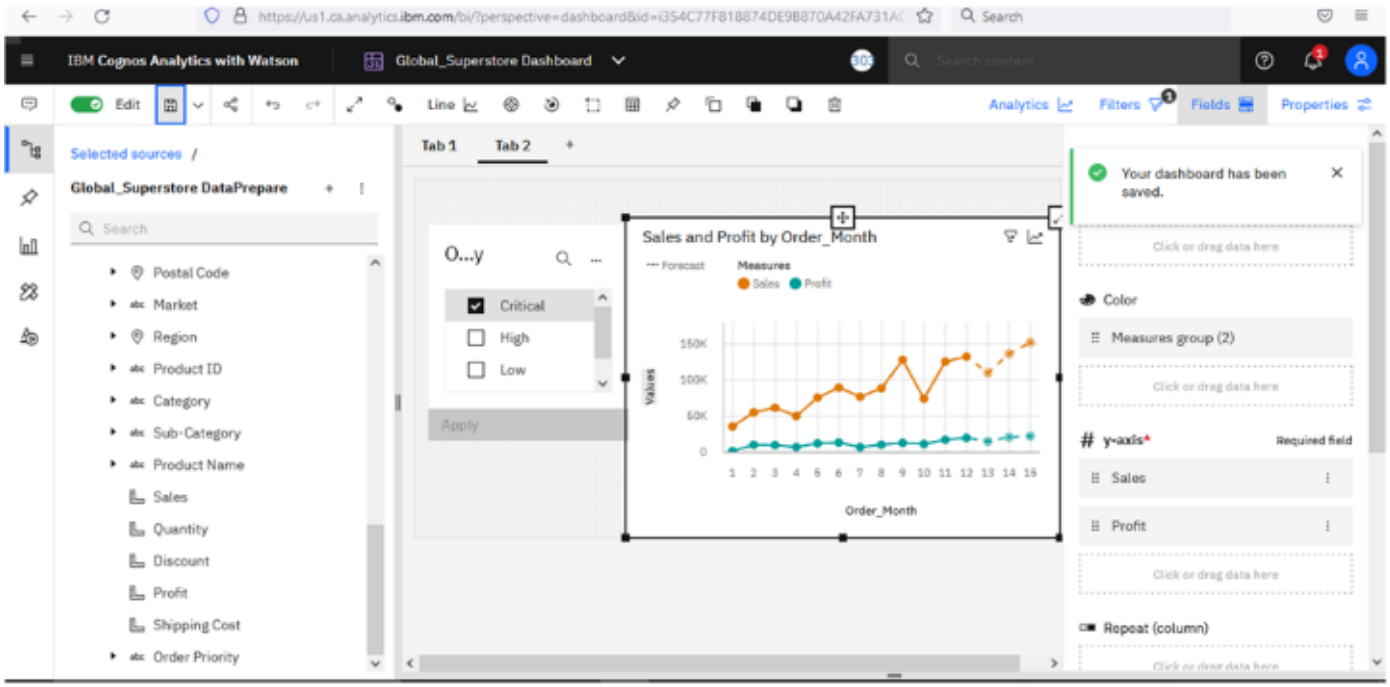
Segment wise sales, profit and quantity



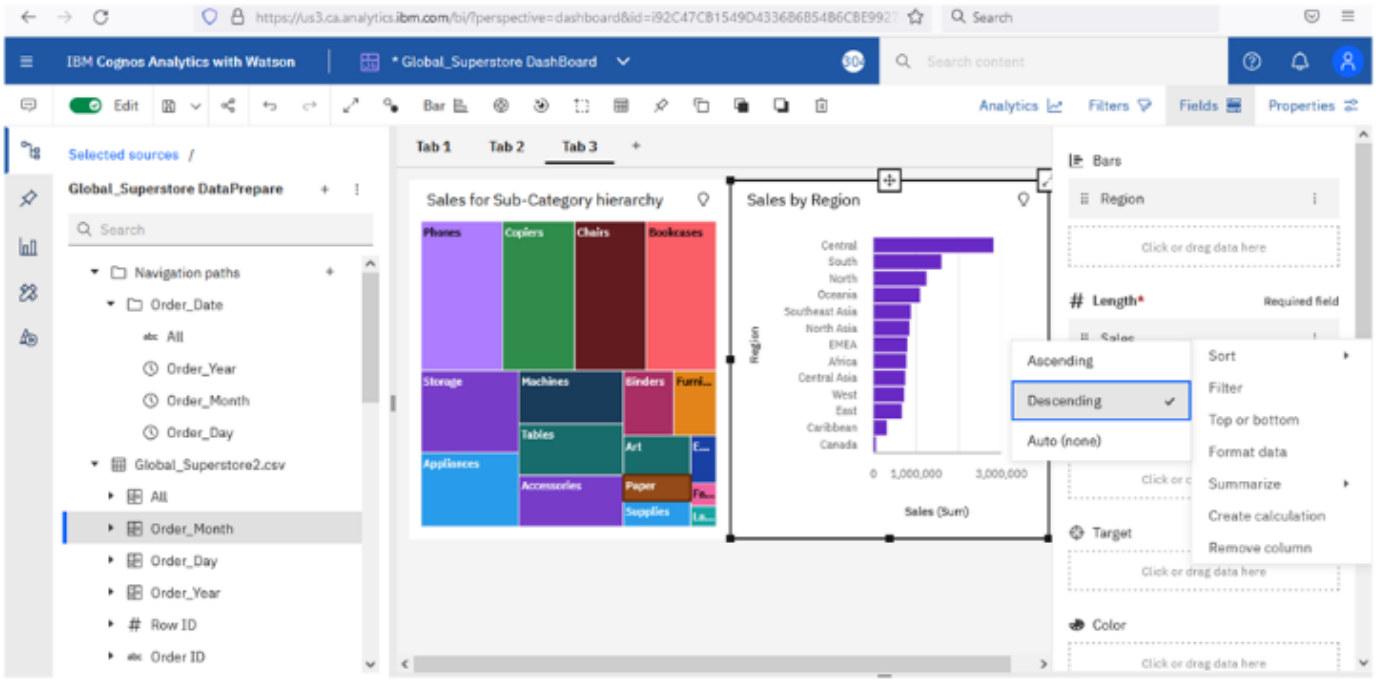
Sales by Market



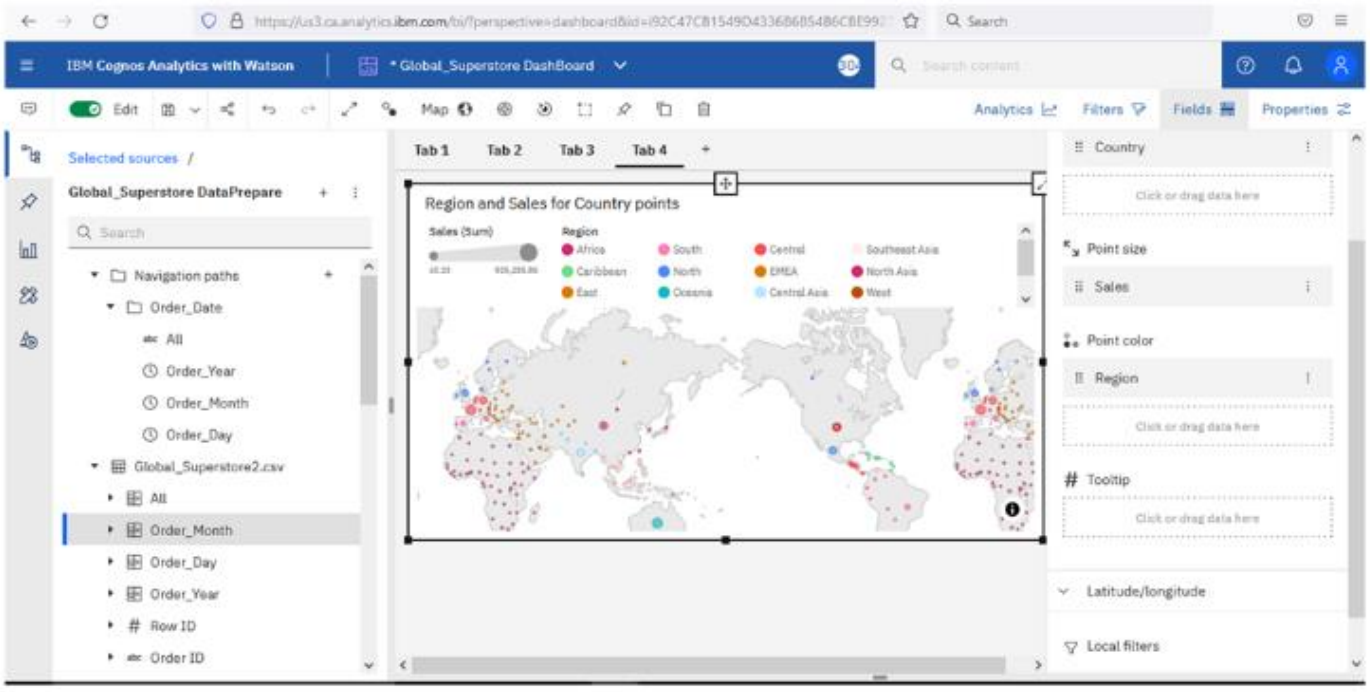
Sales forecast by order priority



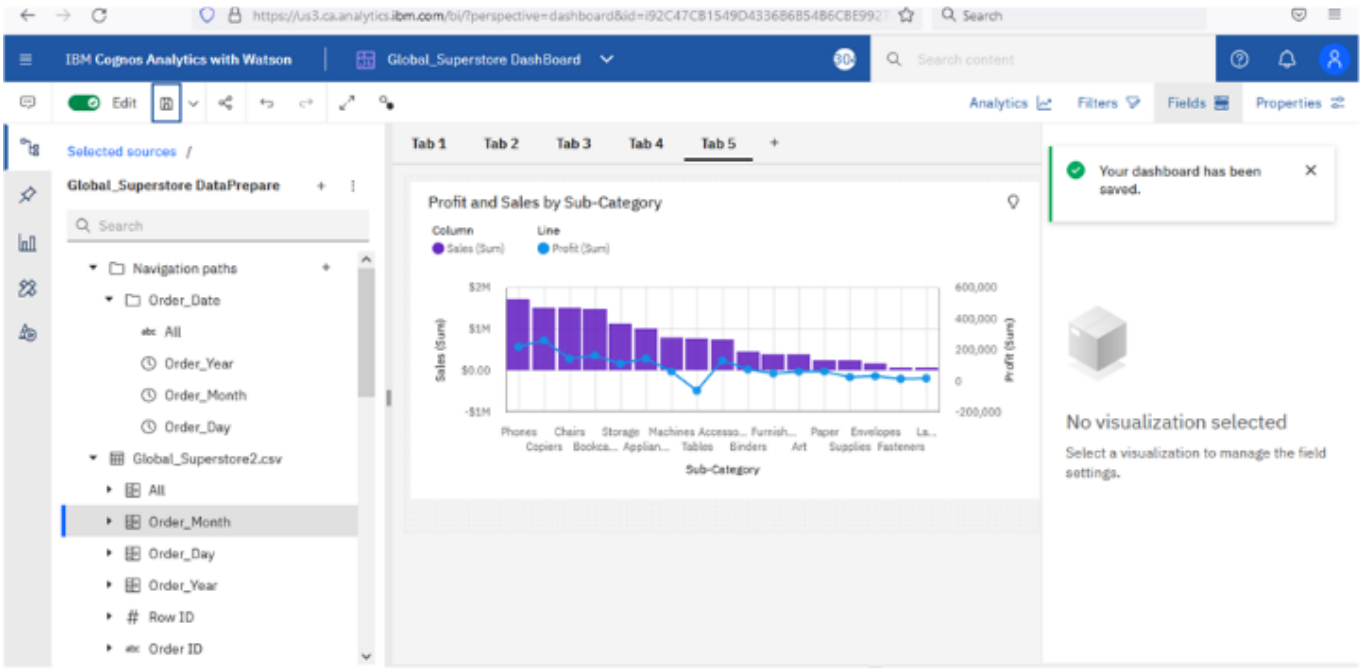
Sales By Sub Category and Sales by region



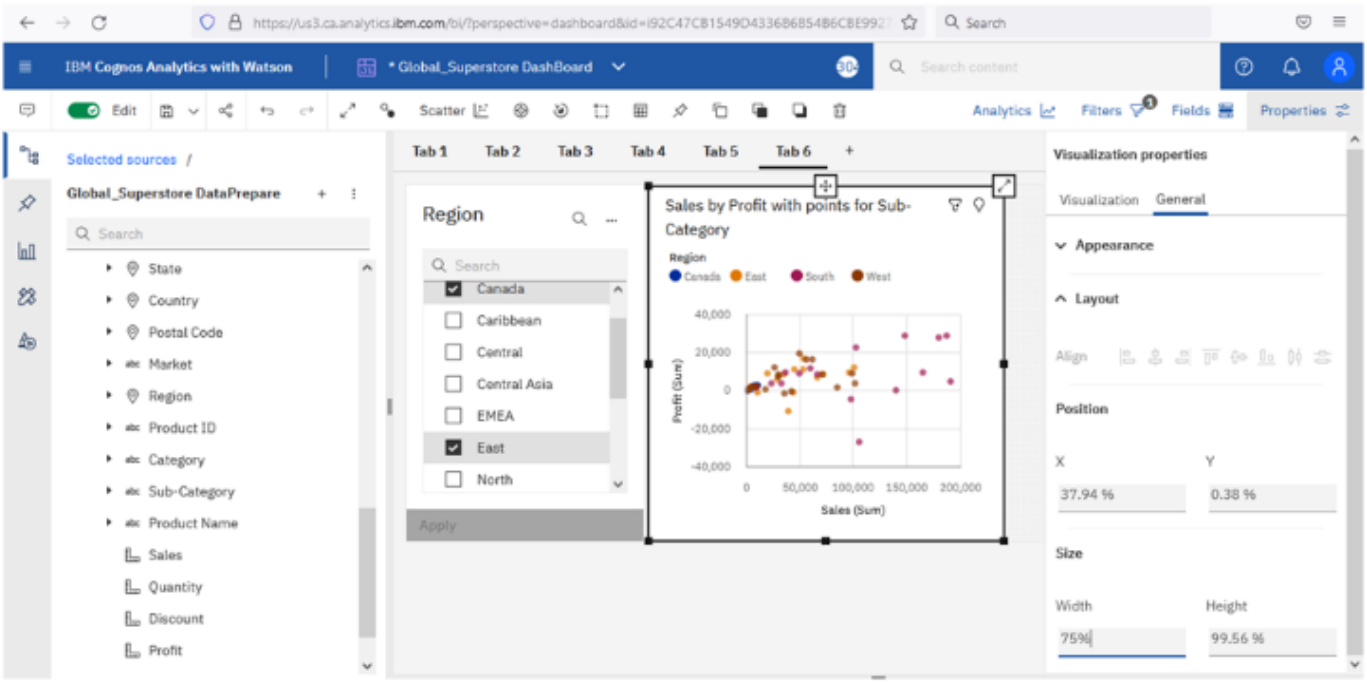
Country wise sales using Map Points



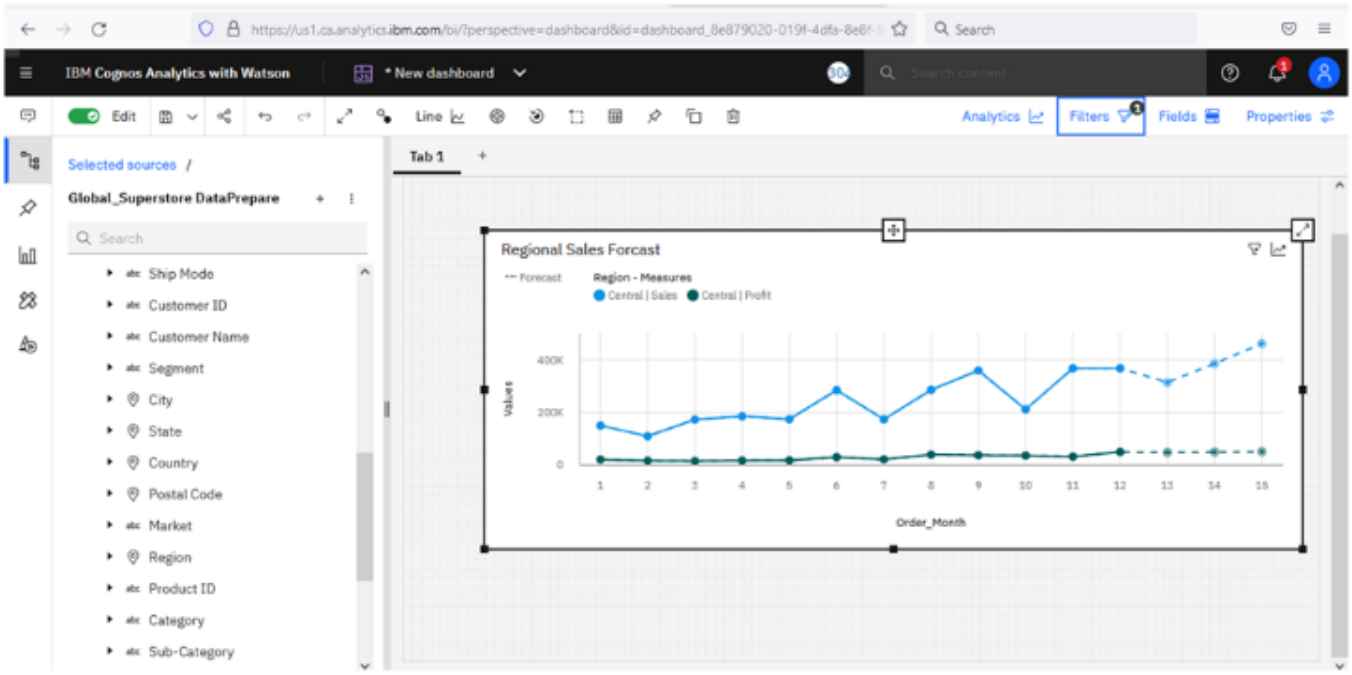
Sub category wise sales and profit using line and bar chat



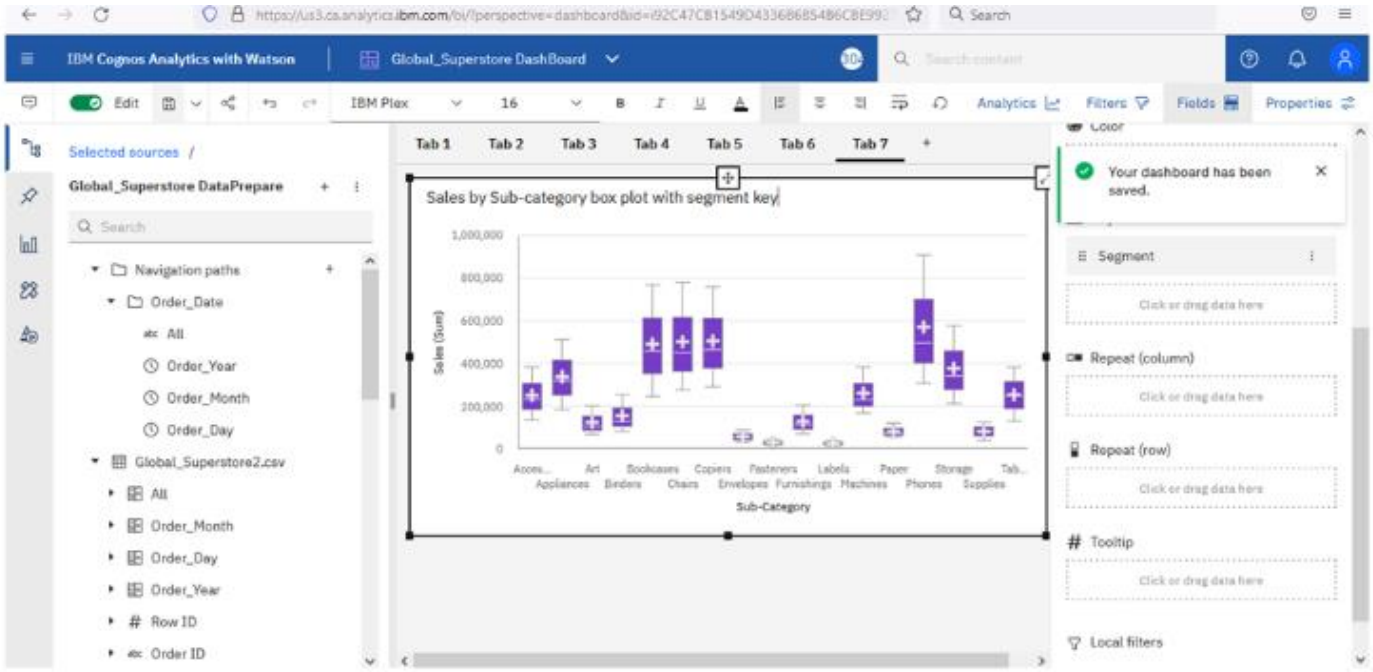
sales vs profit scatter plot with sub categories and regions



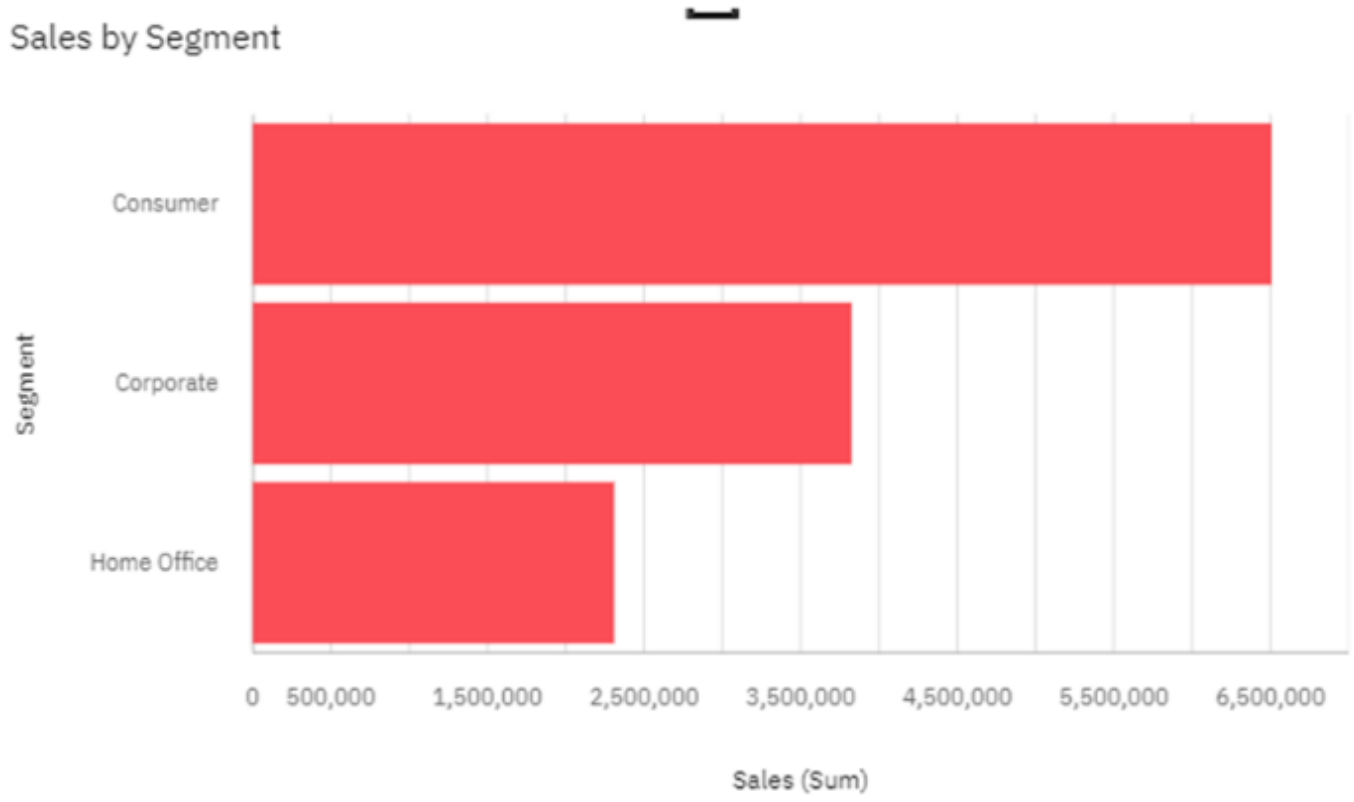
Regional Sales and Profit Forecast



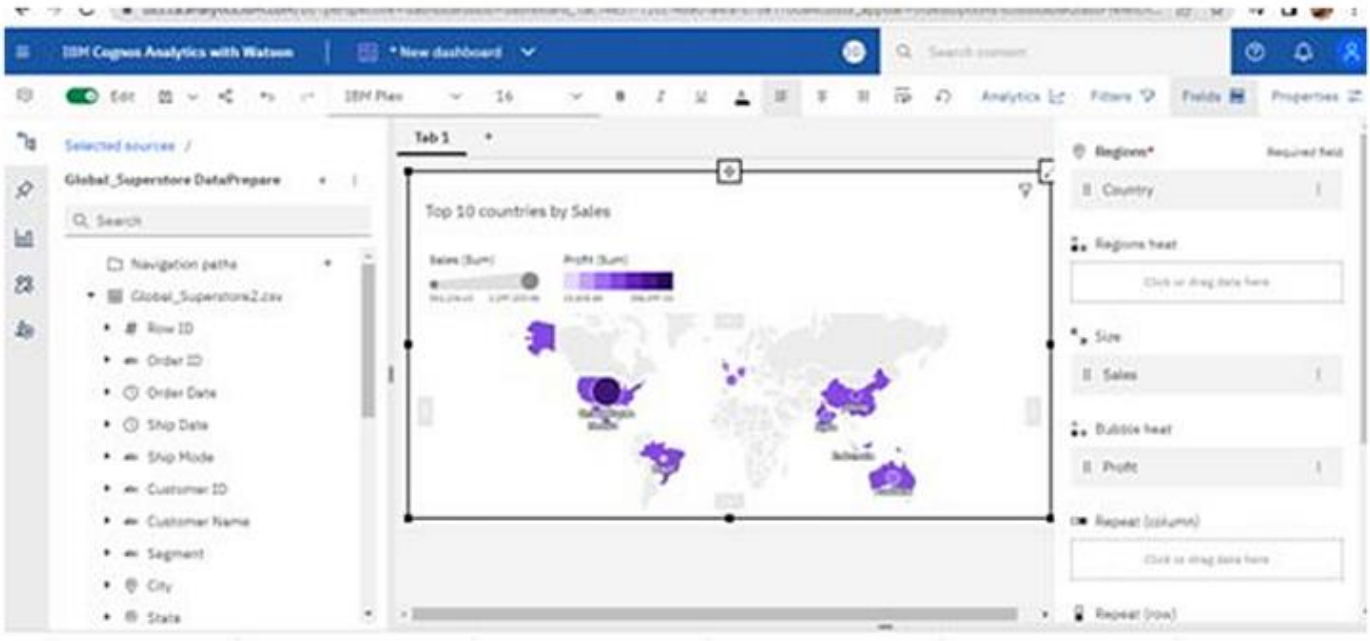
Sales by sub category analysis



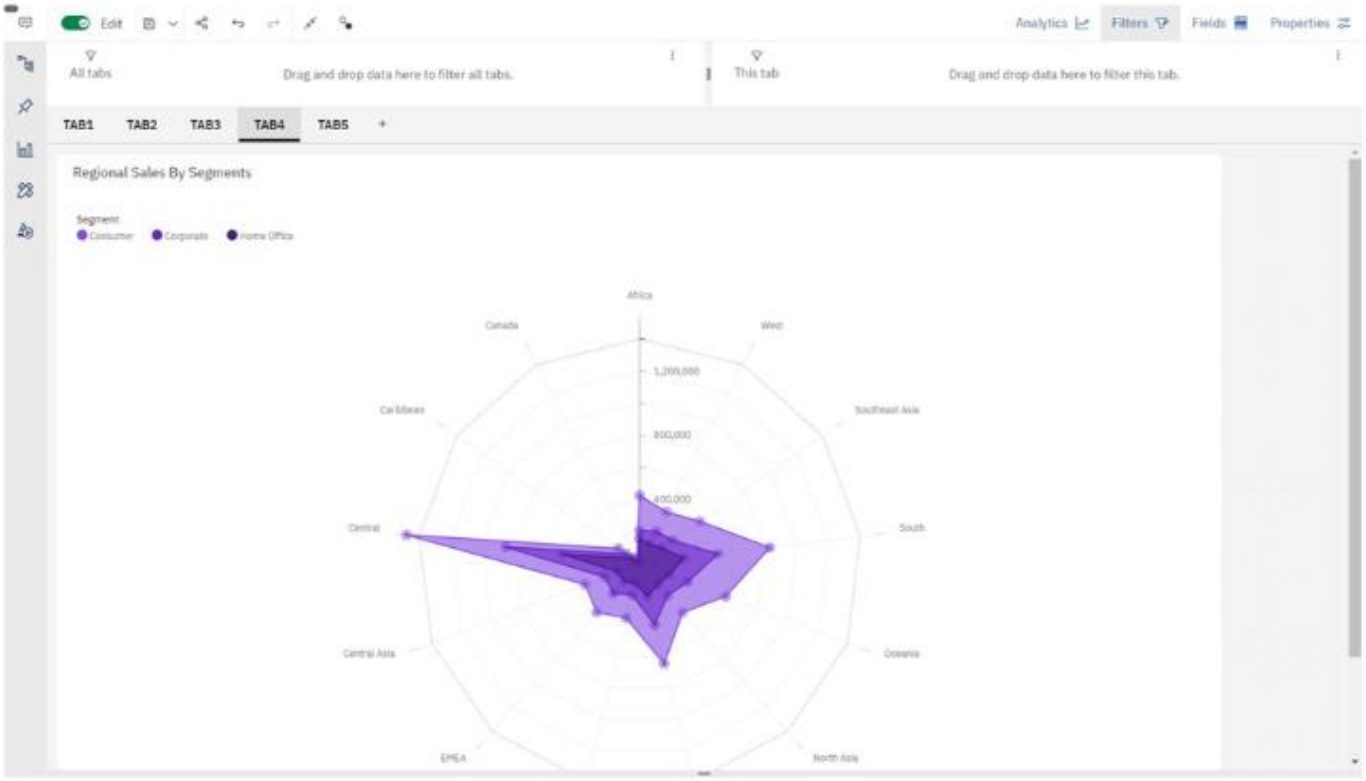
Sales By Segment Analytics



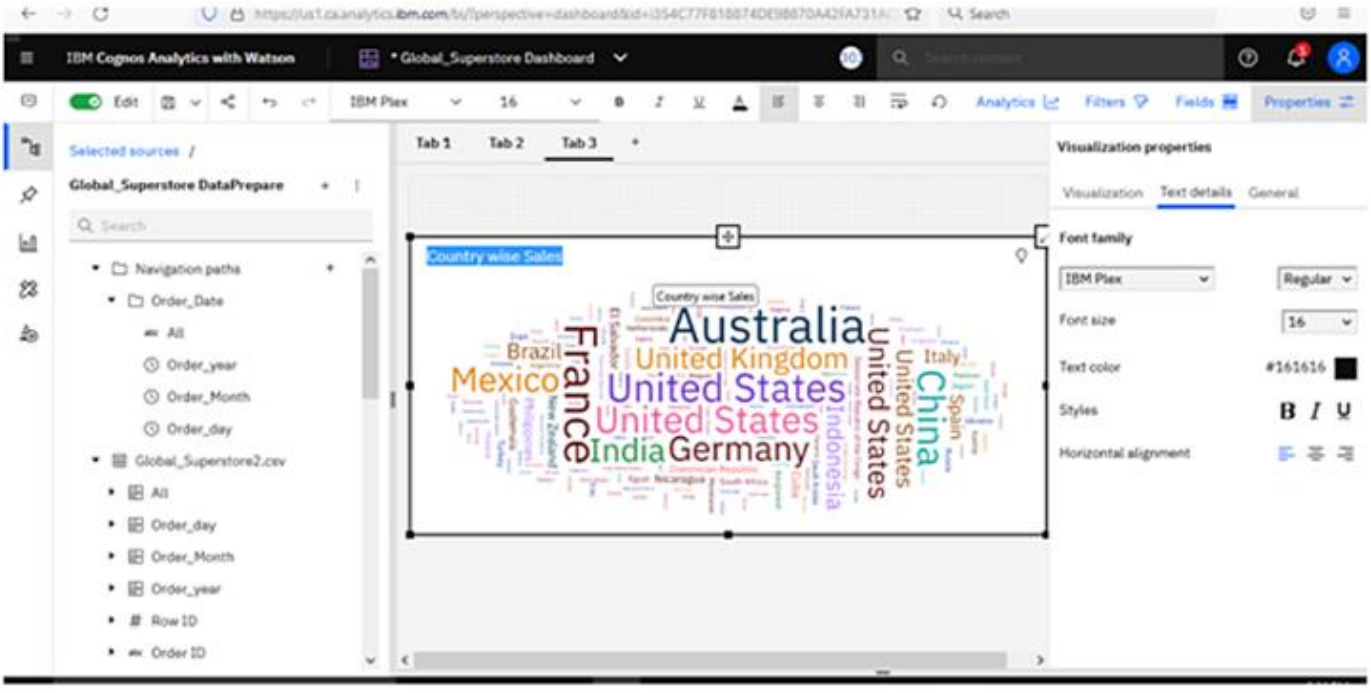
Sales Vs Profit by Countries



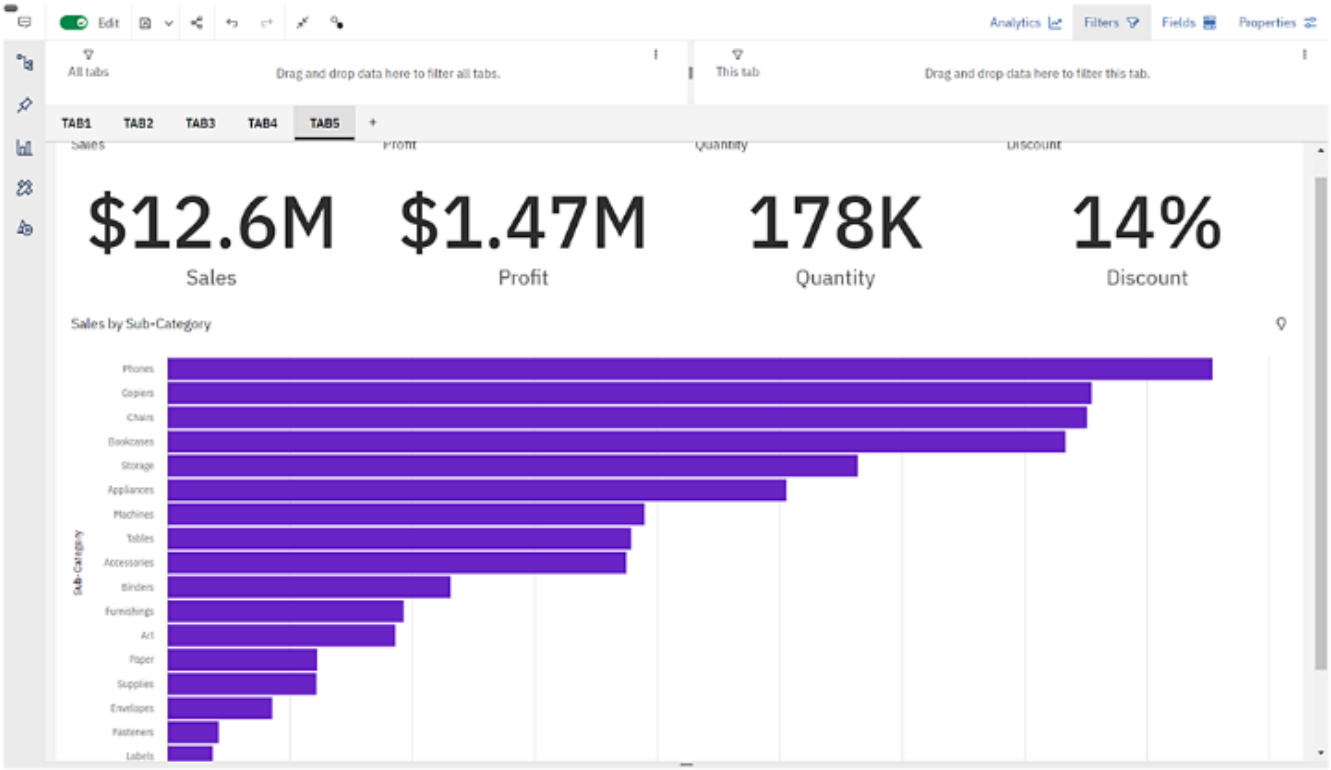
Regional Quantity And Sales Using Radar Chart



Country wise Sales Vs profit using word cloud



Sales dashboard



8. TESTING

8.1 Test Cases

Date	03-Nov-22
Team ID	PNT2022TMID04958
Project Name	Project - Global Sales Data Analytics
Maximum Marks	4 marks

Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Expected Result	Actual Result	Status	Comments	TC for Automation(Y/N)	BUG ID	Executed By
Dashboard_TC_00	UI	Home Page	Verify user is able to see	User should have good	Enter the web page	UI elements of the home page are visible, such as:	Working as	Pass	-	N	-	kiruthika, deepa,
Dashboard_TC_00	Functional	Home Page	Verify the user is able to	User should have good	User should be able to click the	User could see the UI and interact with various button and elements in the web page	Working as	Pass	-	N	-	kiruthika, deepa
Dashboard_TC_00	Functional	Home page	Verify the user is able to	User should have good	Step 1: Scroll down to the Services	User is able to interact with the drop down elements on the home page	Working as	Pass	-	N	-	manjula, dharshini
Dashboard_TC_00	Functional	Home page	Verify the user is able to enter	User should have good	Step 1: Scroll down to Click on the	User is able to fill in their data into the Contact Us form	Working as	Pass	-	N	-	Kiruthika, Manjula
Dashboard_TC_00	UI	Home page	Verify the user is able to view	User should have good	Step 1: Scroll down to the Portfolio	User is able to view the UI of the dashboard	Working as	Pass	-	N	-	Deepa,dharshini,
Dashboard_TC_00	Functional	Home page	Verify the user is able to	User should have good	Interact with various visualizations	User can interact and direct to the cognos dashboard where they can view various visualizations and	Working as	Pass	Various visual charts	N	-	kiruthika,deepa
Dashboard_TC_00	Functional	Home page	Verify the user is able to click	User should have good	Click on the hyperlink to go to the	User is able to click on the Portfolio link which leads them to a second page	Working as	Pass	-	N	-	dharshini,kiruthika
Dashboard_TC_00	UI	Portfolio page	Verify the user is able to view	User should have good	Enter the Portfolio page	User is able to view the UI elements on the Portfolio page	Working as	Pass	-	N	-	Manjula
Dashboard_TC_00	UX	Portfolio page	Verify the user is able to	User should have good	User should be able to click the	User could see the UI and interact with various button and elements in the web page	Working as	Pass	-	N	-	kiruthika, deepa
Dashboard_TC_00	UI/UX	Portfolio page	Verify the user is able to view	User should have good	Step 1: Scroll down to the Portfolio	User is able to view the UI of the dashboard	Working as	Pass	Various visual charts	N	-	dharshini,kiruthika
Dashboard_TC_00	Functional	Portfolio page	Verify the user is able to	User should have good	Interact with various visualizations	User can interact and direct to the cognos dashboard where they can view various visualizations and	Working as	Pass	-	N	-	kiruthika, manjula,

8.2 User Acceptance Testing

Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Global Sales Data Analytics project at the time of the release to User Acceptance Testing (UAT).

Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved.

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	0	0	0	0	0
Duplicate	0	0	0	0	0
External	0	0	0	0	0
Fixed	0	0	0	0	0
Not Reproduced	0	0	0	0	0
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	0	0	0	0	0

Test Case Analysis

This report shows the number of test cases that have passed, failed and untested

9. RESULT

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.

Model Performance Testing:

S.NO	Parameter	Screenshot/Values
1.	Dashboard design	No of Visualizations / Graphs - 7-8 visualization/6-7 graphs
2.	Data Responsiveness	Users and Analysts or Developers
3.	Amount Data to Rendered (DB2 Metrics)	11.5 MB (GlobalSuperstore2.csv)
4.	Utilization of Data Filters	Sales , profit, products, market rate, and order id filtration
5.	Effective User Story	No of Scene Added - 7
6.	Descriptive Reports	No of Visualizations / Graphs - 4 visualizations/10 graph

10. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- Data visualizations
- Ease of use
- Integration capabilities

DISADVANTAGES

- No prediction features available as of yet
- Need to improve security aspect of the product.

11. CONCLUSION

By implementing this analytics solution, the company brought their competitive and sales data reporting in-house, cut costs and increased the accuracy of their reporting and analysis. As the company moves forward with this new solution, their sales reporting costs will most likely be reduced by 50 to 70%. They are now able to analyze raw data themselves, respond more quickly to changes in market trends and perform root cause analysis to determine those shifts in the market. By securing quicker access to their data with the new solution, the company was also able to reduce the risk associated with delayed responses to changes in their markets.

With the new solution, the company can now process sales reports faster than the outsourced solution, reducing turnaround time between 50% to 60%. The reporting needs of the company have been streamlined, consolidating over 10 reports into the centralized dashboard solution. The company's competitive analysis group is also able to more quickly respond to internal data requests given they have the ability to pull the information themselves. With this quicker response, the company is better able to react to changes in the market and predict opportunities for its sales force. The business also experienced an increase in the overall understanding of their sales data throughout the organization. The company now has great flexibility in the presentation of their sales and competitive data, while also being able to integrate sales data with other key data points for the organization.

12. FUTURE SCOPE

Sales analytics refers to the use of technology to collect and use sales data to derive actionable insights. It is used to identify, optimize, and forecast sales. It uses different metrics and KPIs to plan an efficient sales model that generates higher revenue for the business. 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 sale.

13.APPENDIX

GITHUB - https://github.com/IBM-EPBL/IBM-Project-4773-1658739942