GLOBAL SALES DATA ANALYTICS

NALAIYA THIRAN PROJECT BASED LEARNING

on

PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

A PROJECT REPORT

SANTHOSHKUMAR.B	111919104120
SARAVANAN.M	111919104122
VENU.G	111919104156
SAGAYAPRADEEP.A	111919104114

BACHELOR OF ENGINEERINGIN

COMPUTER SCIENCE AND ENGINEERING

S. A. ENGINEERING COLLEGE
ANNA UNIVERSITY: CHENNAI 600 025

November 2022

SA Engineering College
An Autonomous Institution Affiliated to Anna University Chennai ·
Accredited by NBA, NAAC 'A' Grade & ISO 9001:2015 Certified Institution

INTERNAL MENTOR

NATTESHAN N V S M.E.,

Assistant Professor,

Department of Computer Science and Engineering,

S.A Engineering College

Chennai-600077

INDUSTRY MENTORS

Shanawaz Anwar
Indra Prakash

TABLE OF CONTENTS

1. INTRODUCTION

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 Product Analysis of this Global Super Store. If you want to achieve your sales goals month after month, then guesswork and intuition aren't your best friends. You need to perform strategic sales analysis and get cold, hard data.

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).

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 aspects enable you to optimize you.

2. LITERATURE SURVEY

1. EXISTING PROBLEMS

- The huge blast of information and Internet gadgets has prompted a fast approach to Big Data in the later past. The administration industry which is a noteworthy client for these Big Data applications will prompt real 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 and dispersion of work comprehensively. The Service Industry will utilize Big Data for different choices making an information framework and making the work process more ideal. The idea of large-scale manufacturing lead to the Industrial Revolution, likewise, Big Data is relied upon to drive new types of financial movement in the Service

industry with connected human capital, achieving a new level of monetary action, development, and development.

- In the information era, enormous amounts of data have become available on hand to decision-makers. Big data refers to datasets that are not 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 in order to handle and extract value and knowledge from these datasets. Furthermore, decision-makers need to be able to gain valuable insights from such varied and rapidly changing data, ranging from daily transactions to customer interactions and social network data. Such value can be provided using big data analytics, which is the application of advanced analytics techniques on big data. This paper aims to analyze some of the different analytics methods and tools which can be applied to big data, as well as the opportunities provided by the application of big data analytics in various decision domains.
- In the modern era of higher education, it is exceptionally challenging for teachers to counsel students in terms of academic matters. Teachers have abundant data related to different aspects of students but deriving appropriate 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 volume of the data, and quality of the data, and the adoption of suitable predictive algorithms. This paper deals with an exhaustive literature survey of predictive analytics in the educational domain specifically for higher education. The paper also discussed the analysis of the existing literature survey and identified research gaps.

2. REFERENCES

- A literature survey on big data analytics in the service industry.(https://www.researchgate.net/publication/301720427_A_literature_survey_on_Big_Data_Analytics_in_Service_Industry).
- Big Data Analytics: A literature review paper.(https://www.researchgate.net/publication/264555968_Big_Data_Analytics_A_Literature_Review_Paper).
- Iterature survey using predictive analytics for student counselling in higher education.
- (https://www.researchgate.net/publication/355481331_Literature_Survey_using_Predictive_Analytics_for_Student_Counselling_in_Higher_Education).

3. PROBLEM STATEMENT DEFINITION

Sales and marketing teams need to review their strategies and performance to make improvements. One way to measure performance is with Sales Analytics.

Sales analytics refers to the technology and processes used to gather sales data and gauge sales performance. Sales leaders use these metrics to set goals, improve internal processes, and accurately forecast future sales and revenue. It uses different metrics and KPIs to plan an efficient sales model that generates higher revenue for the business.

The goal of sales analytics is always to simplify the information available to the sales and Marketing teams. It should help them clearly understand the team's performance, sales trends, and opportunities to gain many insights and develop strategies that are better than the previous one.

PS -1

PS -2

PS -3

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	Sales team Member	Analyze sales across multiple sources	There are inaccuracies in the sales	Sources are different from different systems	Dissatisfied
PS-2	Sales team member	Prepare data	I get errors during data entry	There is complexity in data	frustrated
PS-3	Sales team member and sales leader of the company	Understand product sales	The Sales strategy is unknown	It is difficult to understand the mindset of a customer	disappointed

3.IDEATION & PROPOSED SOLUTION

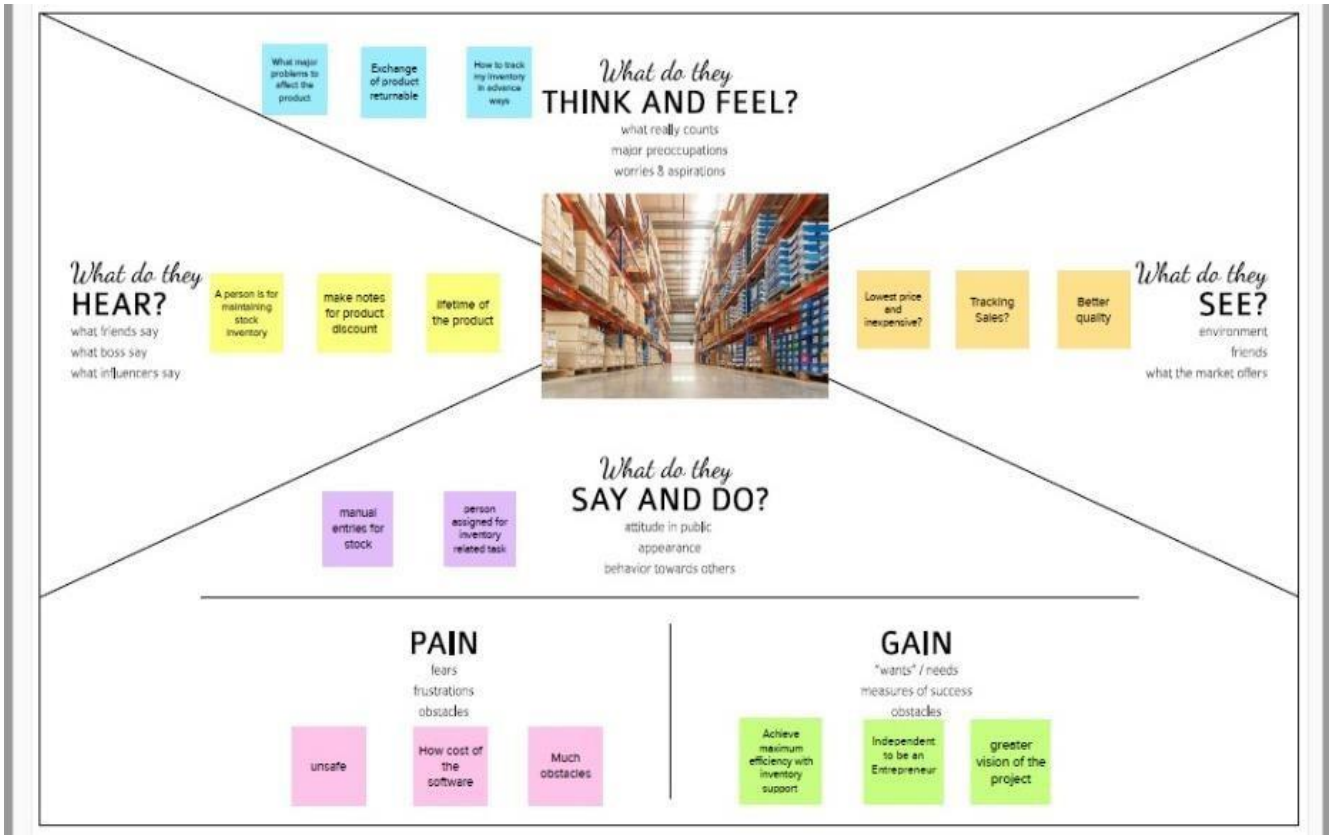
1. EMPATHY MAP CANVAS

An empathy map is a simple, easy-to-digest visual that captures knowledge about 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 problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.

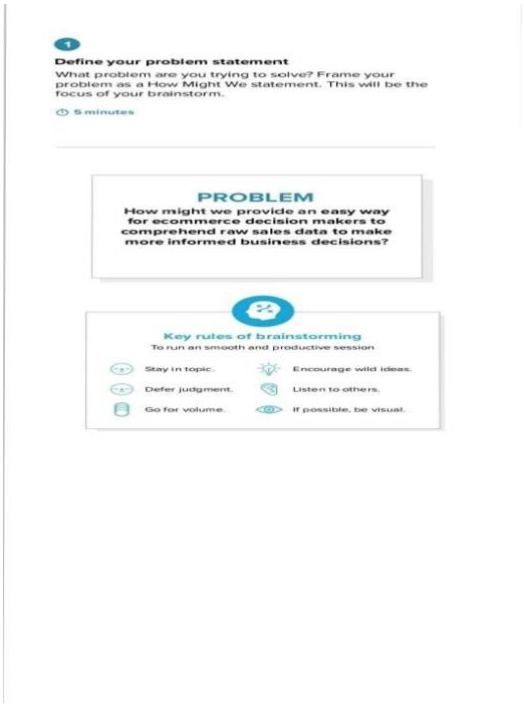
Empathy map for Global Sales Data Analytics



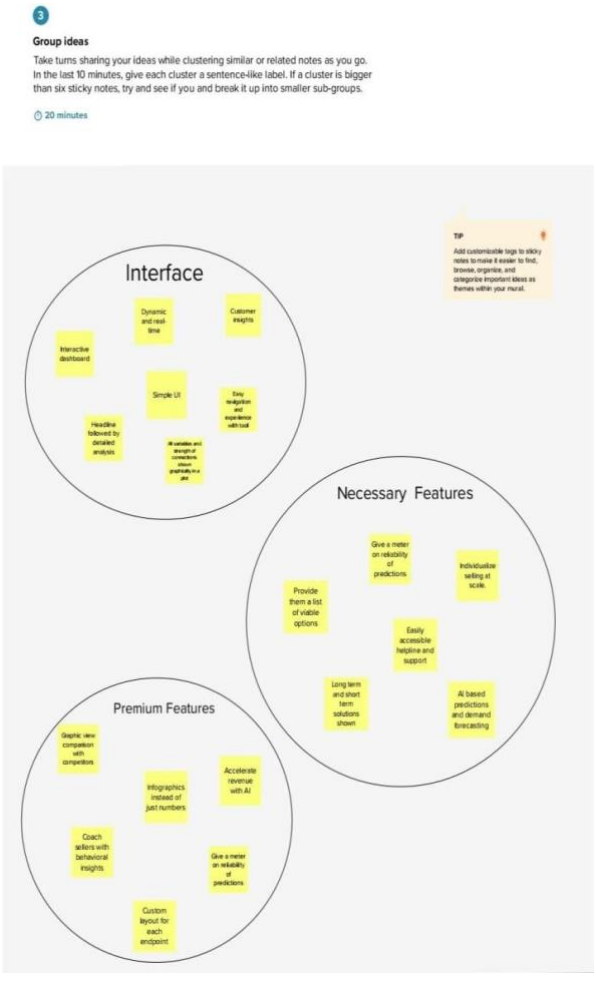
2.IDEATION & BRAINSTROMING

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.

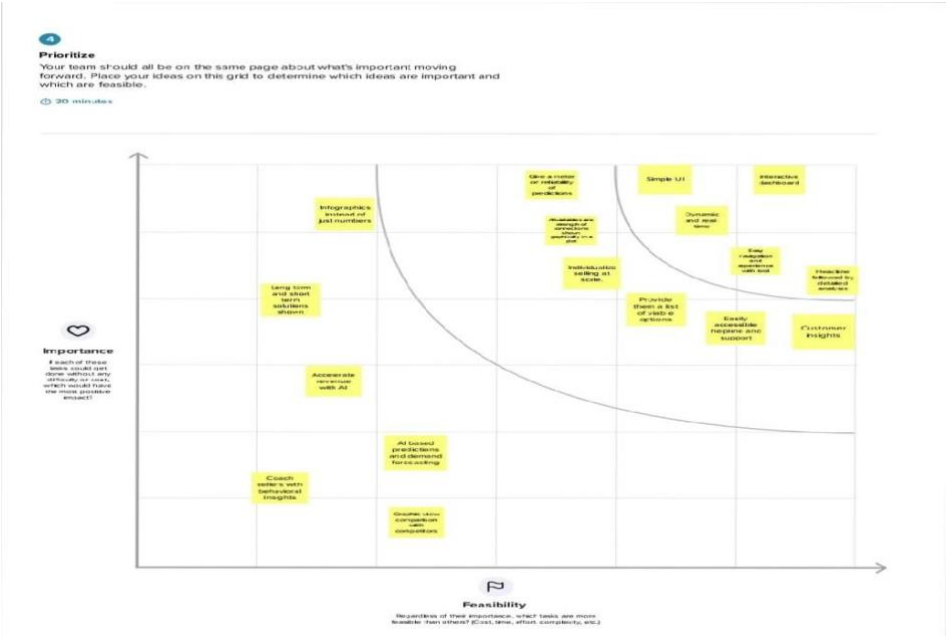
Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



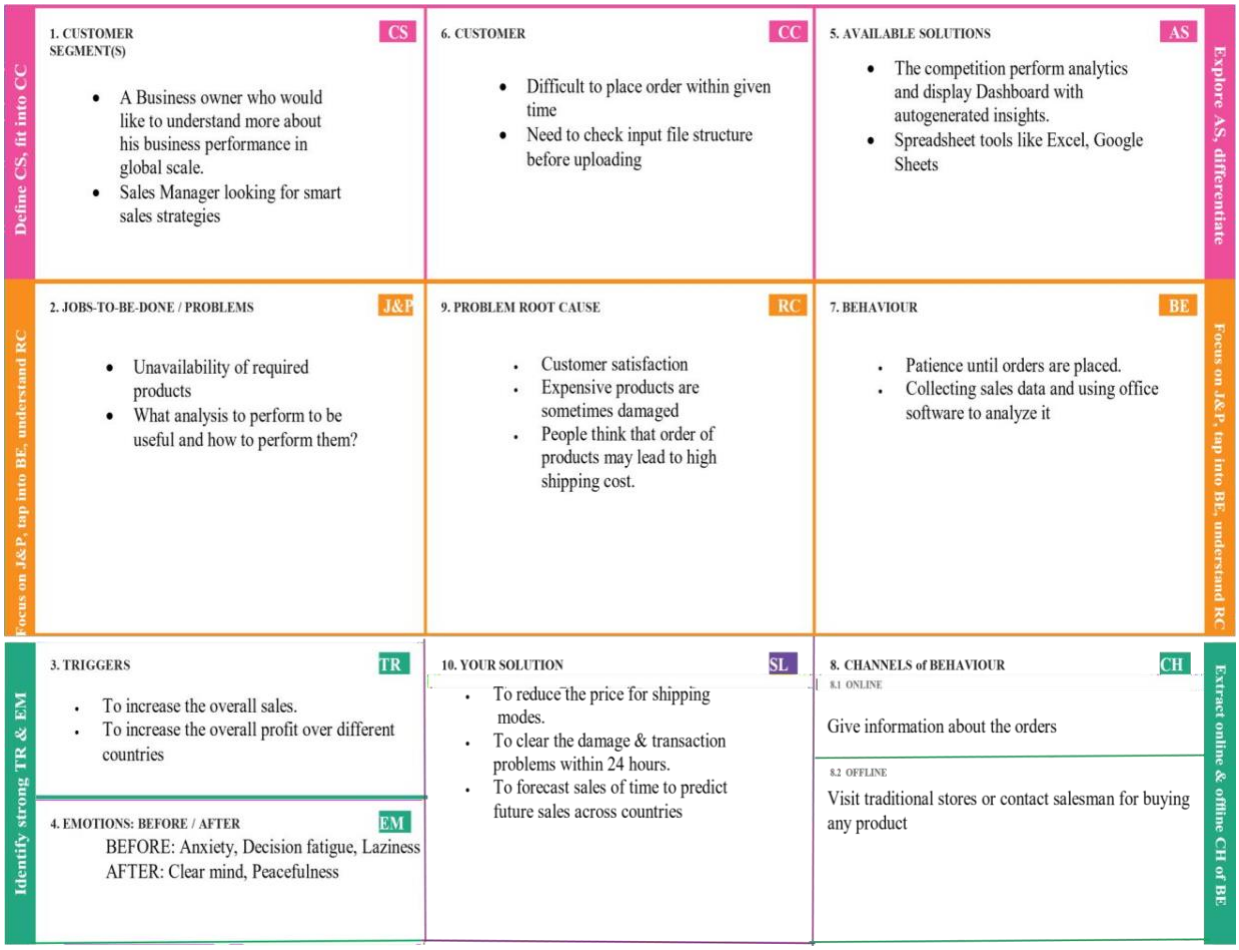
3.PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Decision makers of E-commerce companies(User)need a way to comprehend raw data, analyse and make more informed business decisions. E-commerce companies(User) need a way to understand the shift in preferences of customers and the current trend, so that they can satisfy the customers.
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 Dashboard and simple UI Dynamic and real time analytics AI based predictions and forecasting
4.	Social Impact / Customer Satisfaction	Visible profiles driven by informed decisions Optimize sales and marketing Ability to react to competitor's strategies

5.	Business Model (Revenue Model)	Three tier pricing-Basics, Standards, Enterprise 1.Basic : Limited features targeting startups and individuals 2.Standard : Limited premium features. Target customers- Medium Scale businesses. 3.Enterprise with all premium features targeted at Large corporations.
6.	Scalability of the Solution	More B2B customer services can be provided alongside Usable by all customers facing companies and startups of all scale

4. PROBLEM SOLUTION FIT

Problem-Solution Fit Design Phase-I - Solution Fit Template Team ID:PNT2022TMID16522



4.REQUIREMENT ANALYSIS

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	User Input	Data uploaded must be of proper format
FR-4	Data Verification and Validation	Data is cleaned and verified for outliers, duplications
FR-5	Data Visualization	Proper charts and graphs are chosen for a particular set of data
FR-6	Business decisions	Recommendations are made according to the data

2.NON-FUNCTIONAL REQUIREMENT

Following are the non-functional requirements of the proposed solution

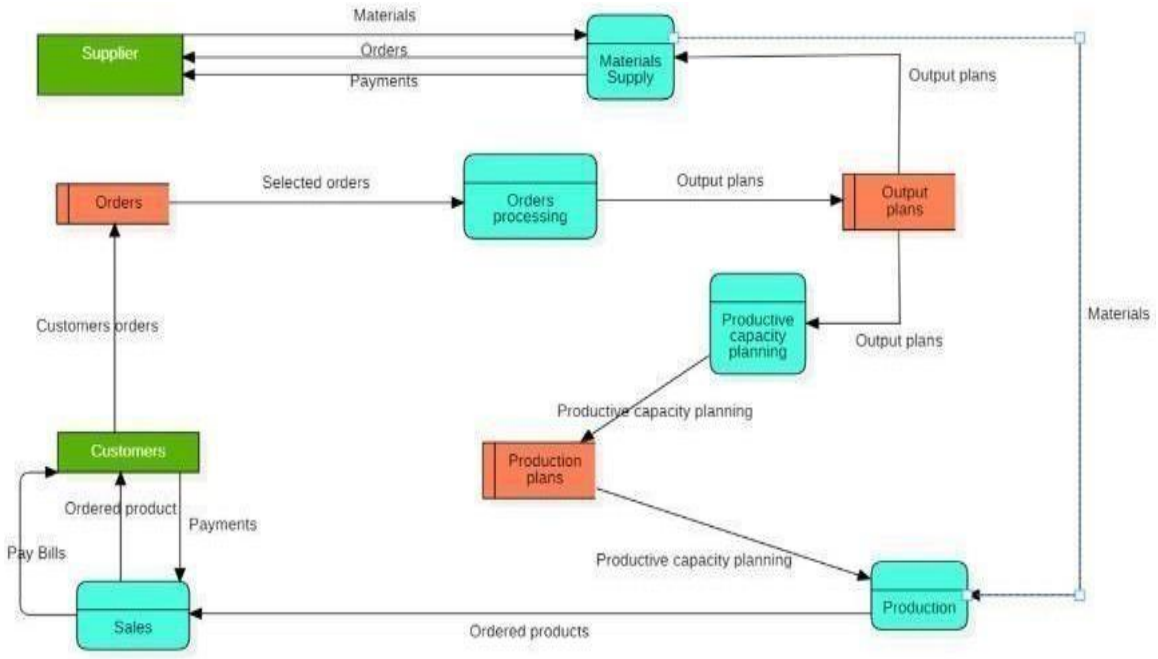
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	It is used for making critical decisions to expand their retail business and can be used by everyone
NFR-2	**Security**	It is securable because it has end to end encryption and only accessible to the user with credential details
NFR-3	**Reliability**	It has high reliability based on development and can be accessed using the cloud

NFR-4	**Performance**	It works with high accuracy and efficiency and has the high state of performance
NFR-5	**Availability**	It is available for anyone signed up for the platforms and websites
NFR-6	**Scalability**	It can be extended and elaborated with high datasets

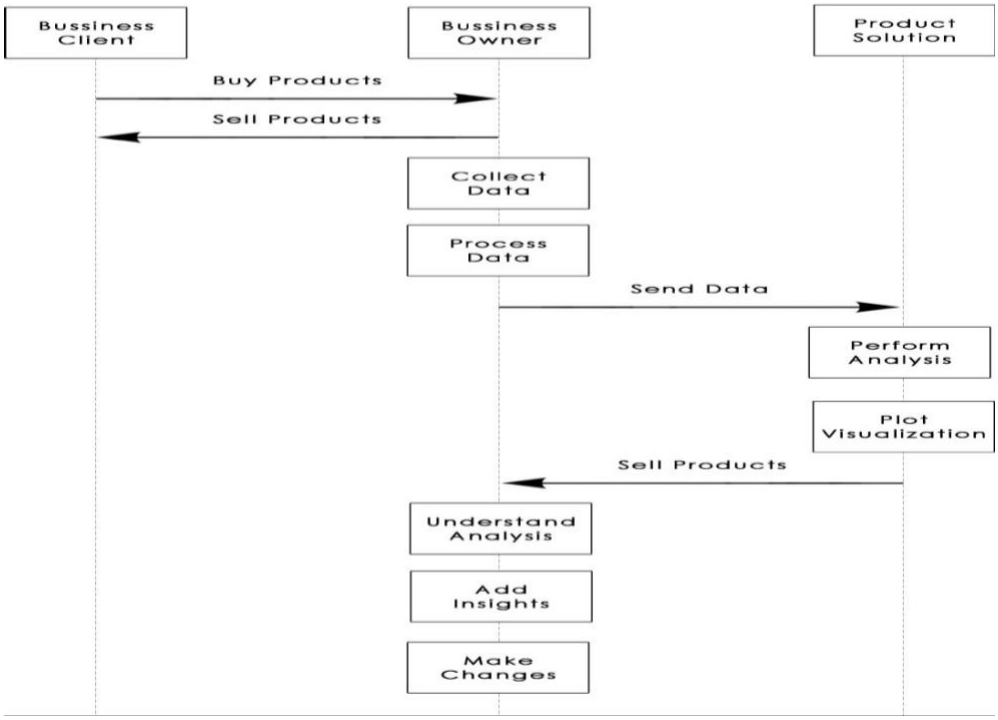
5.PROJECT DESIGN

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 **Structural Flow Diagram:**



Data Flow Diagrams:



2. SOLUTION & TECHNICAL ARCHITECTURE

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

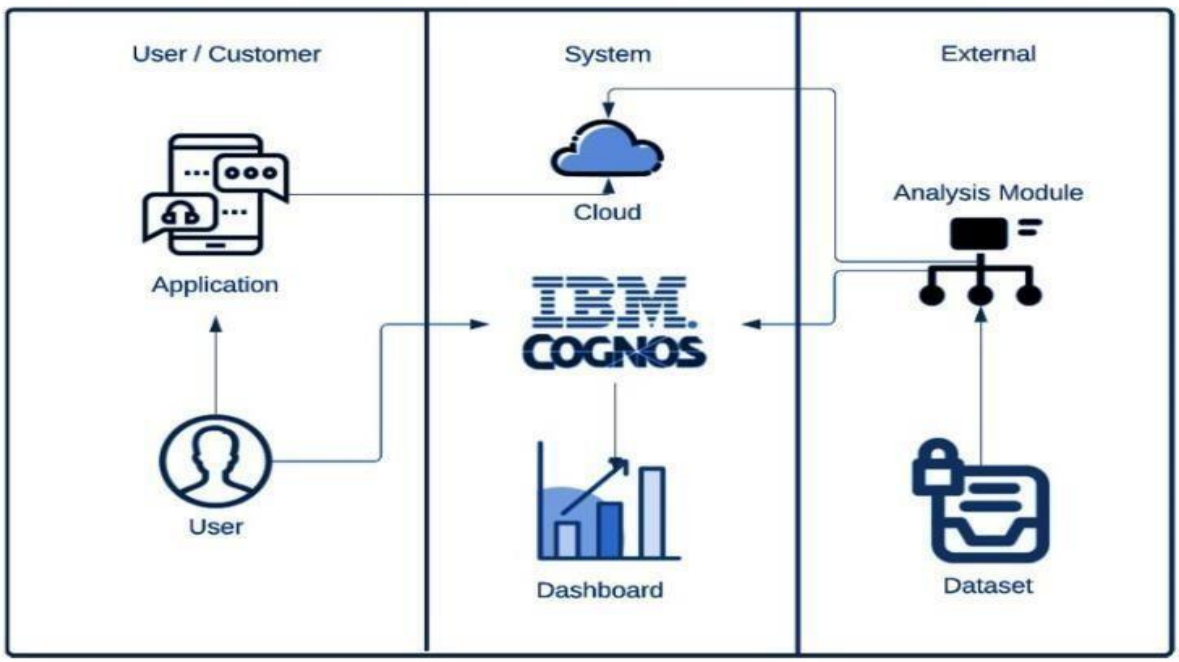


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	IBM Cognos
2.	Application Logic-1	Logic for a process in the application	IBM Cloud
3.	Application Logic-2	Logic for a process in the application	IBM Cognos Analytics
4.	Application Logic-3	Logic for a process in the application	IBM Cognos Analytics
5.	Database	Data Type, Configurations etc.	MySQL
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Cognos Analytics
9.	External API-2	Purpose of External API used in the application	Jupyter Notebook
10.	Machine Learning Model	Purpose of Machine Learning Model	Predictive Analysis Model
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Jupyter Notebook	Python
2.	Security Implementations	Unauthorised access	AES algorithm
3.	Scalable Architecture	Large data sets	IBM Cloud
4.	Availability	Multipage data visualisation charts used for everyone in the login and can be used to finding their prediction	IBM Cognos analytics
5.	Performance	Accuracy and efficiency are increased	IBM Cognos Analytics

5. PROJECT PLANNING AND SCHEDULING

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 confirmingmy password.	2	High	Santhosh kumar.B,Saravanan.M,Venu.G, Sagayapradeep.A
Sprint-1	Login	USN-2	As a user, I need valid credentials to log in to my application.	1	High	Santhosh kumar.B,Saravanan.M,Venu.G ,Sagayapradeep.A
Sprint-1	Data Collection	USN-3	As a user, I need to gather the data in the formof CSV/XLS and clean the data	2	High	Santhosh kumar.B,Saravanan.M,Venu.G,Sagayapradeep.A
Sprint-2	Upload dataset	USN-4	As a user, I can view the data of the products	1	Low	Santhosh kumar.B,Saravanan.M
Sprint-2	Data Preparation	USN-5	As a user, I need to filter it for Data visualization.	3	High	Saravanan.M,Venu.G
Sprint-2	Data visualization	USN-6	As a user, I can easily visualize the data in theform of charts.	4	Medium	Santhosh kumar.B, Sagayapradeep.A
Sprint-3	Dashboard	USN-7	As a user, I can view the summary of the product sales by the help dashboard.	2	Medium	Venu.G, Sagayapradeep.A
Sprint-3	Dashboard	USN-8	As a user, I must plan visualizations in a way that I'm able to gain insights regarding the sales based upon the category of sales and the respective region	4	High	Santhosh kumar.B,Venu.G

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-3	Dashboard	USN-9	As a user, I must be able to gain insights fromthe charts/graphs through a variety of	4	Medium	Santhosh kumar.B,Saravanan.M

			relationships established in the dashboard.			
Sprint- 4	Prediction	USN-10	As a user, I see the prediction of the specificproduct's future sales expectation.	4	Medium	Saravanan. M,Venu.G
Sprint- 4	Report	USN-11	As a user, I can view the list of categorizedproducts and their details as a report.	5	High	Santhosh kumar.B,Venu.G
Sprint-4	Story	USN-12	As a user, I can view the product and customer description and more additional information as a story.	5	High	Santhosh kumar.B,Sa gayapradee p.A

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprints	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	5	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	8	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	10	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	14	19 Nov 2022

2.SPRINT DELIVERY SCHEDULE

Title	Description	Date
Literature Survey and Information Gathering	Gathering Information by referring the technical papers, research Publications.	1 September 2022
Prepare Empathy Map	To capture user pain and gains Prepare List of Problem Statement	12 September 2022
Ideation	Prioritize a top 3 ideas based on feasibility and Importance	19 September 2022
Proposed Solution	Solution include novelty, feasibility, business model, social impact and scalability of solution	24 September 2022
Problem Solution Fit	Solution fit document	1 October 2022
Solution Architecture	Solution Architecture	1 October 2022
Customer Journey	To Understand User Interactions and experiences with application	9 October 2022
Functional Requirement	Prepare functional Requirement	15 October 2022
Data flow Diagrams	Data flow diagram	15 October 2022
Technology Architecture	Technology Architecture diagram	16 October 2022
Project Development-Delivery of sprint 1,2,3 &4	Develop and submit the developed code by testing it	24 October 2022 – 19 November 2022

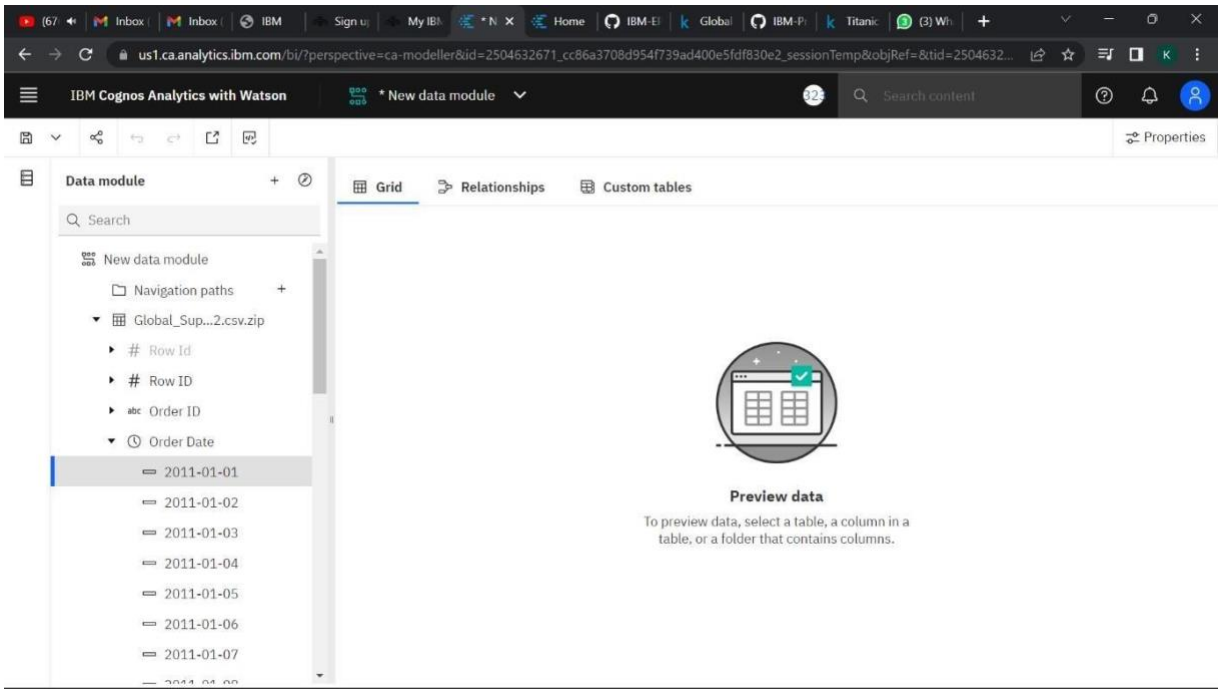
7.CODING & SOLUTIONING

FEATURE 1

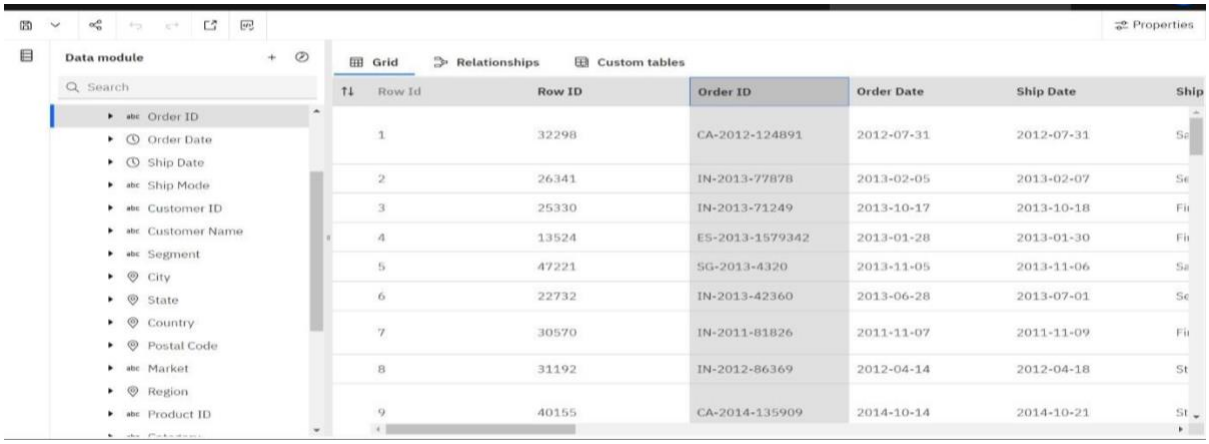
An interactive dashboard has been embedded

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

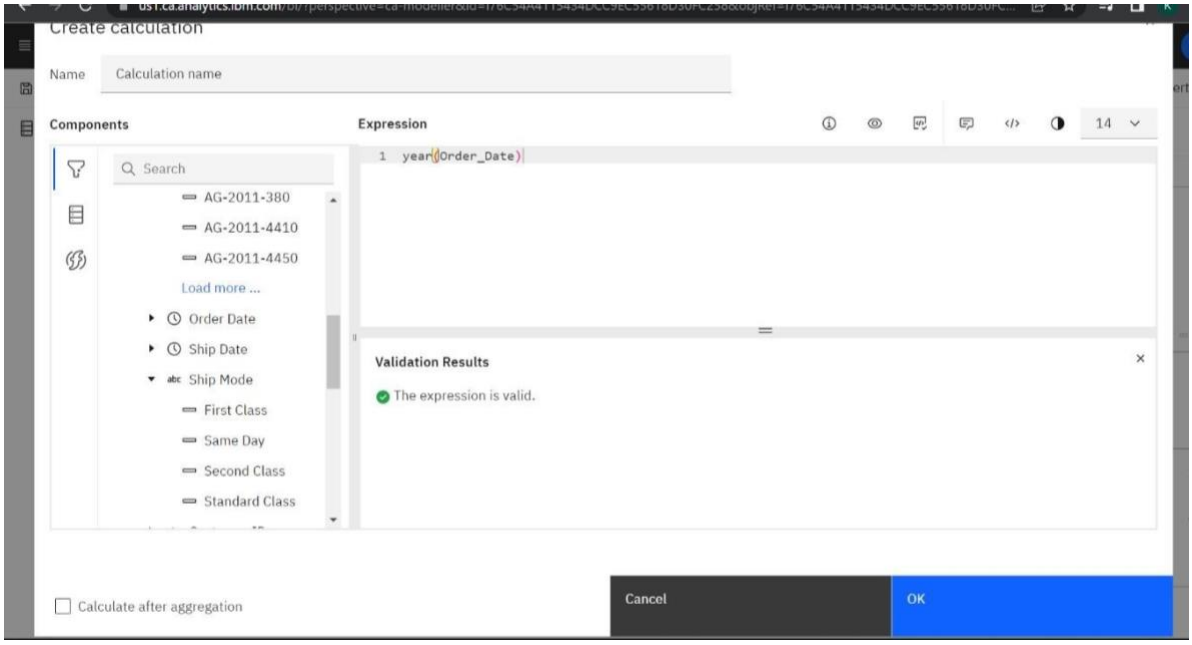
Global Superstore_Data Upload

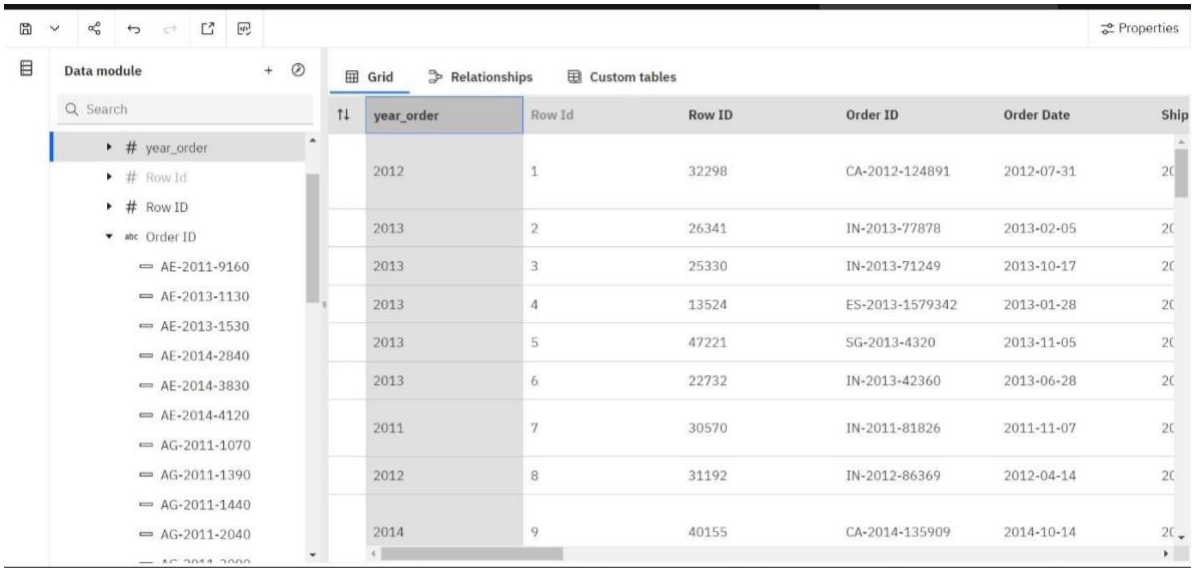


Global Superstore_DataPrep



Date Calculations and the Navigation path

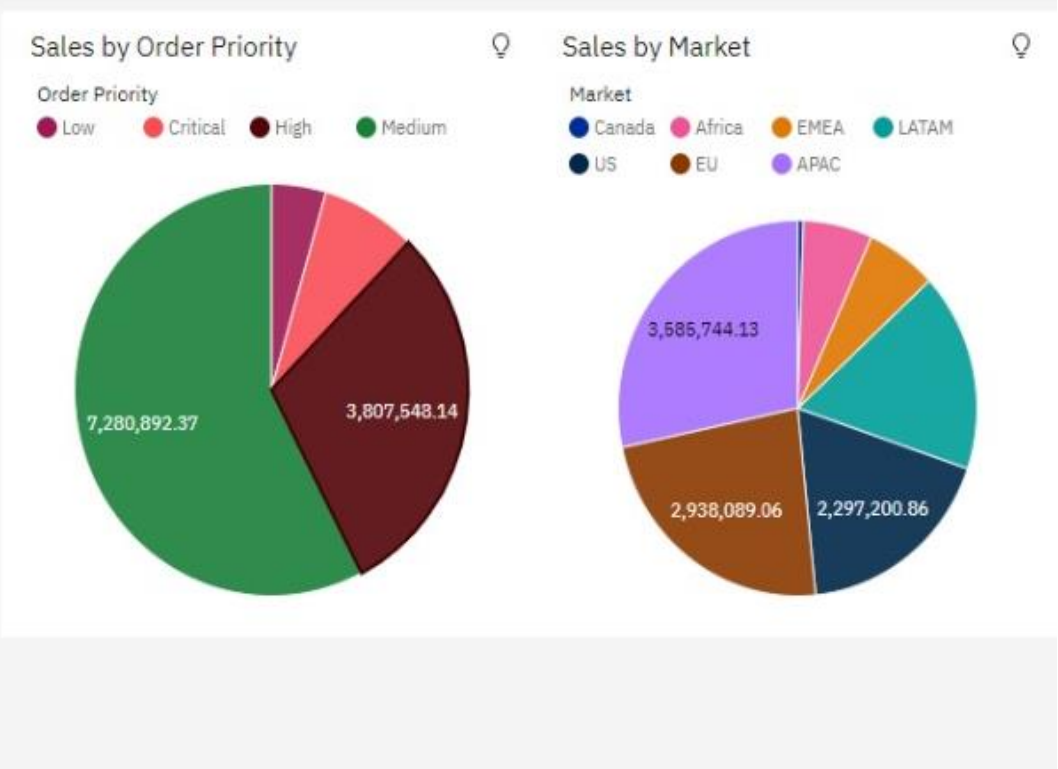




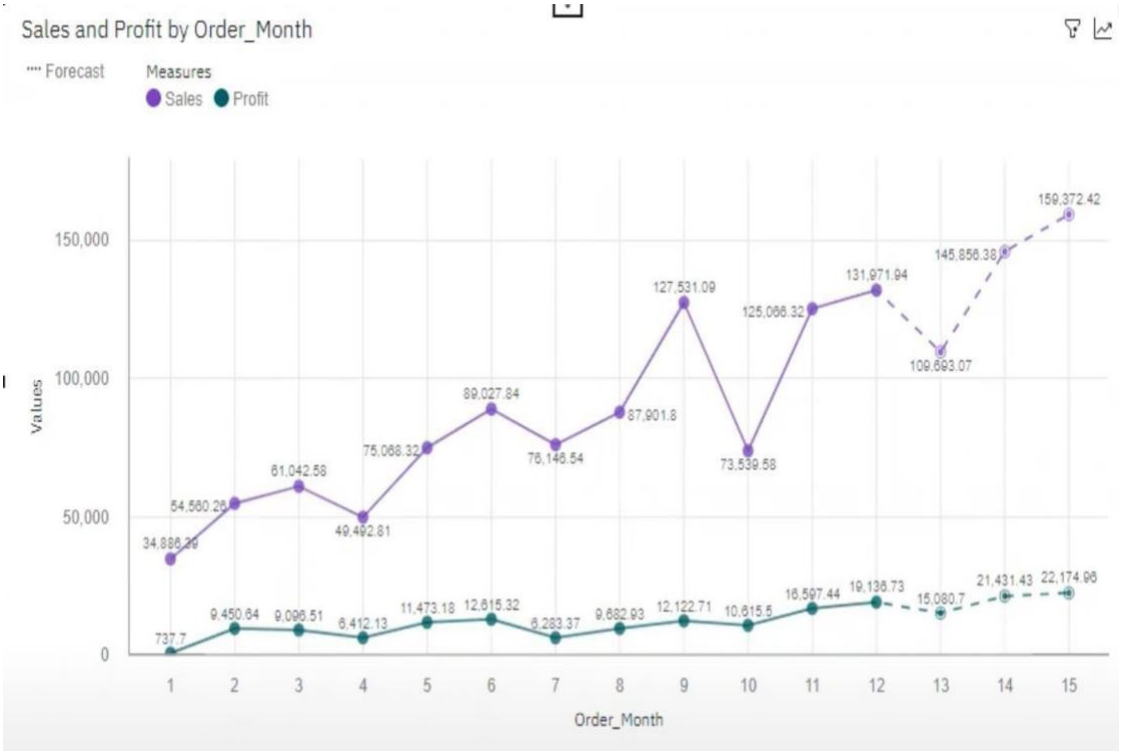
Global Superstore_DataVisualization

Segment Wise Sales, Profit, And Quantity

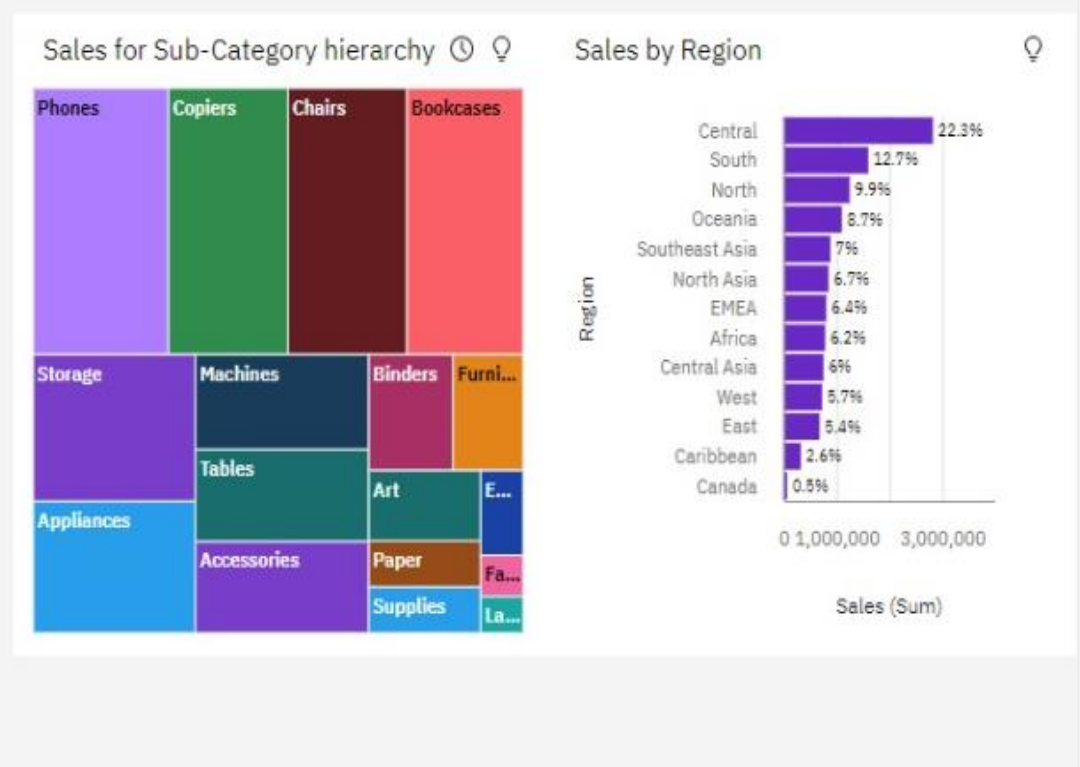
Sales By Market



Sales by order priority



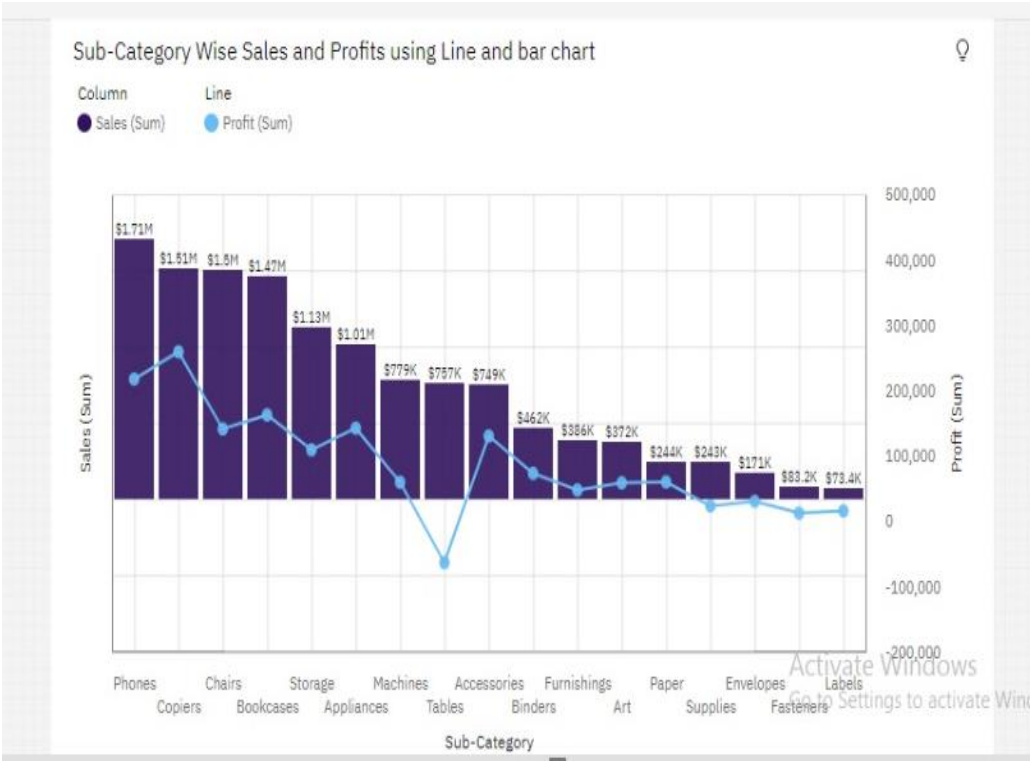
Sales By Sub Category And Sales By Region



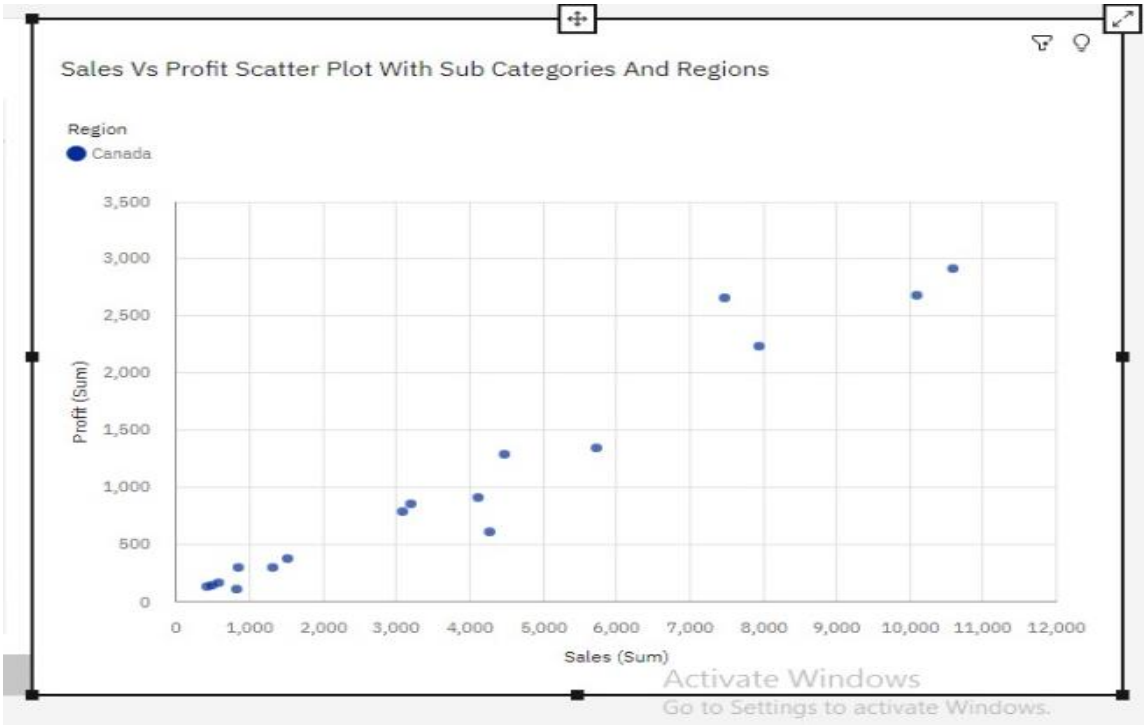
o

Country-Wise Sales Using Map Points

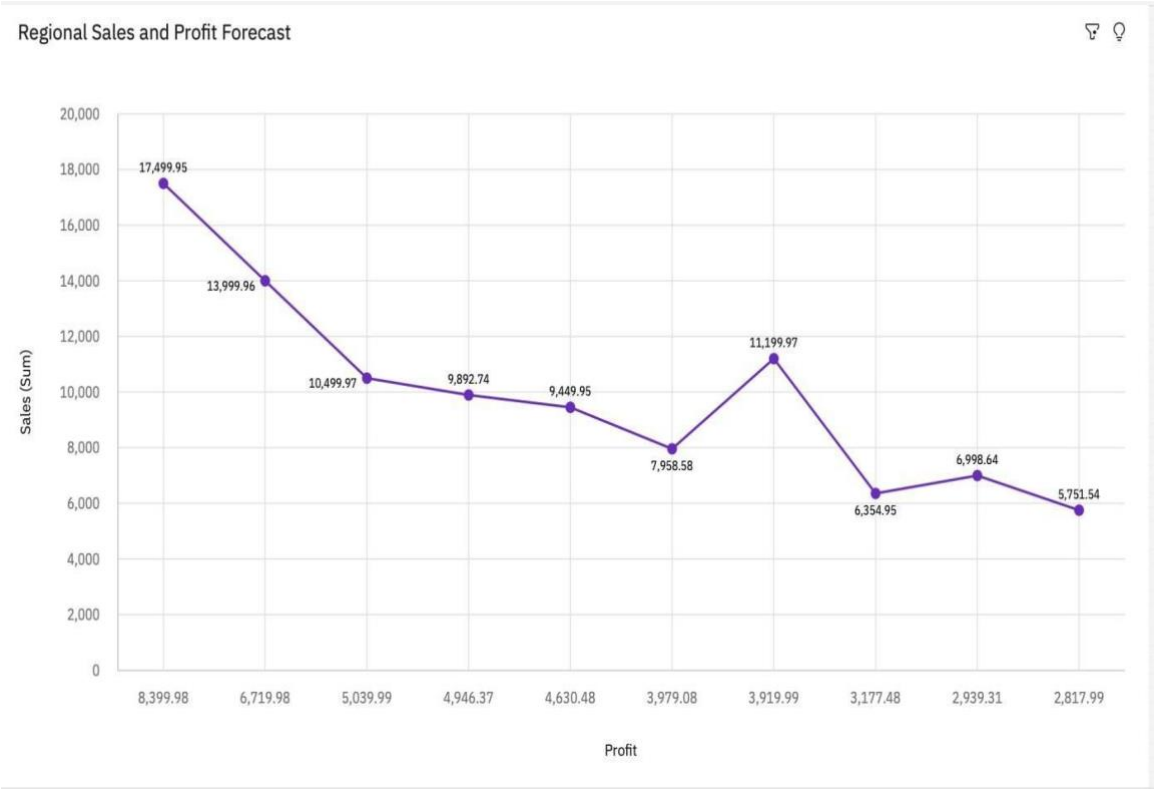
Sub Category Wise Sales And Profits Using Line And Bar Chart



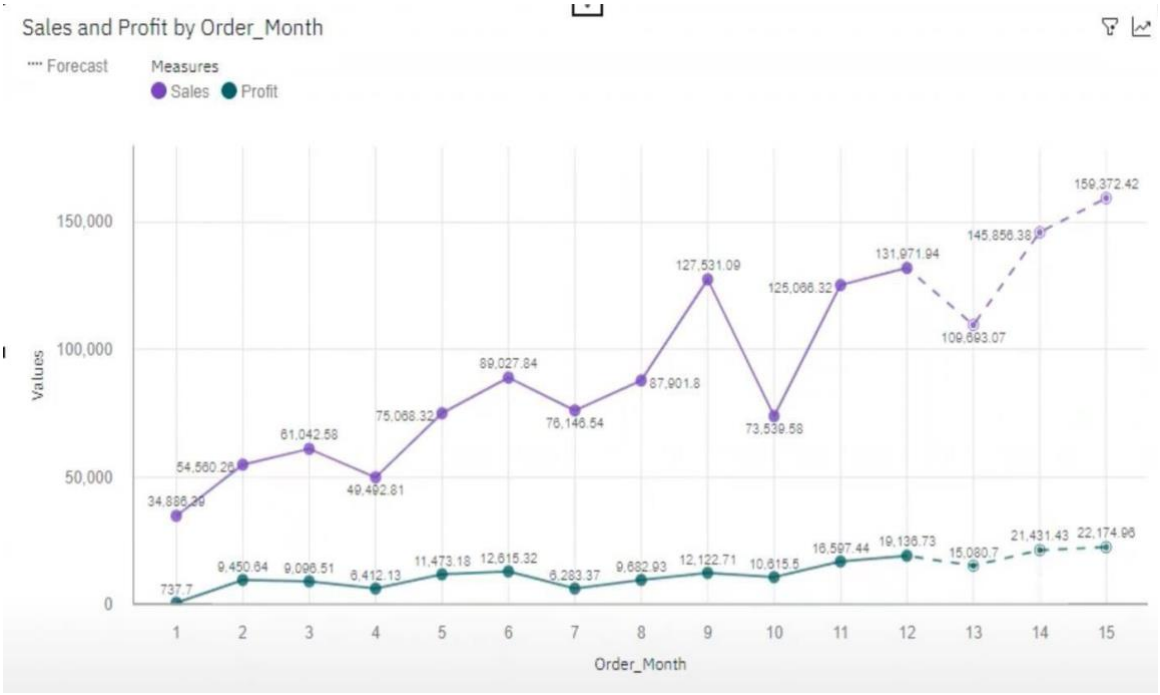
Sales Vs Profit Scatter Plot With Sub Categories And Regions



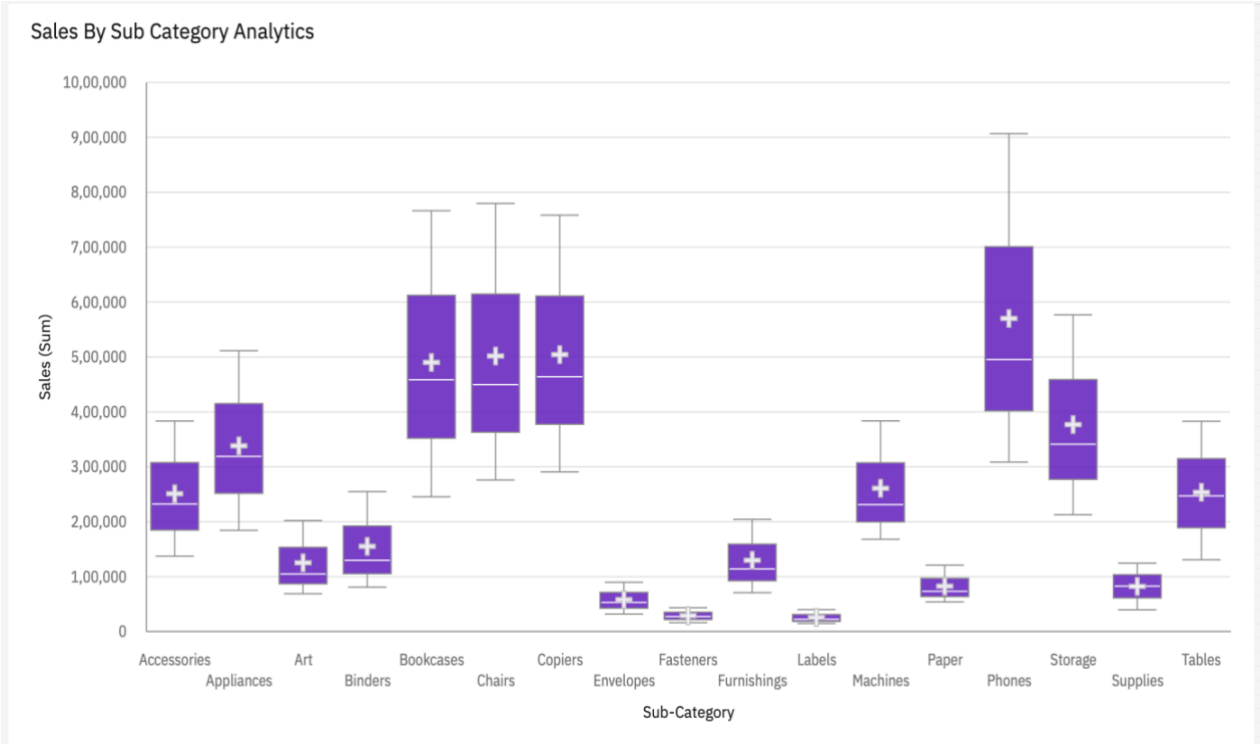
Regional Sales And Profit Forecast



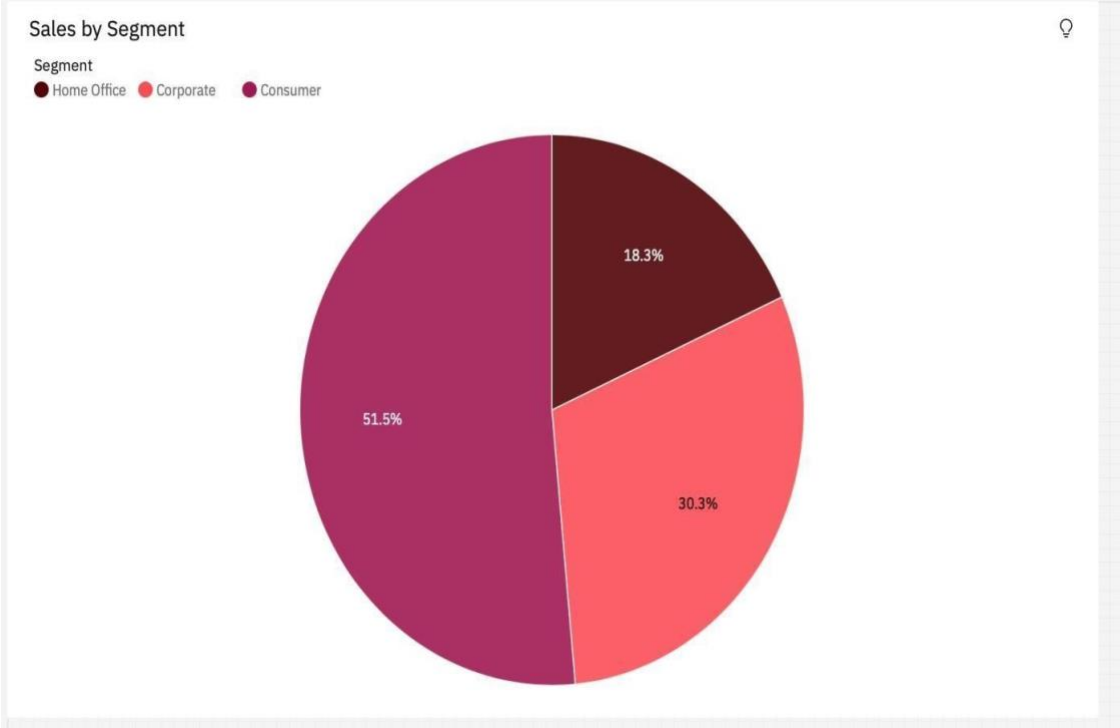
Sales Forecast By Order Priority



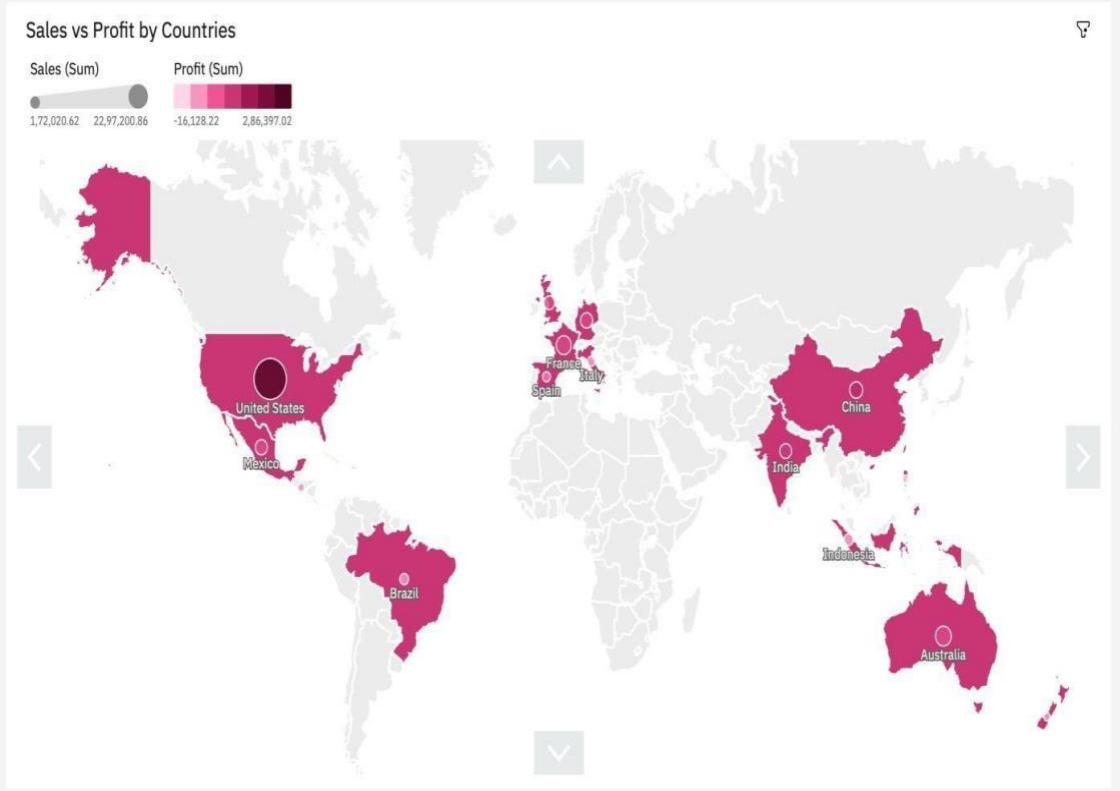
Sales By Sub-Category Analytics



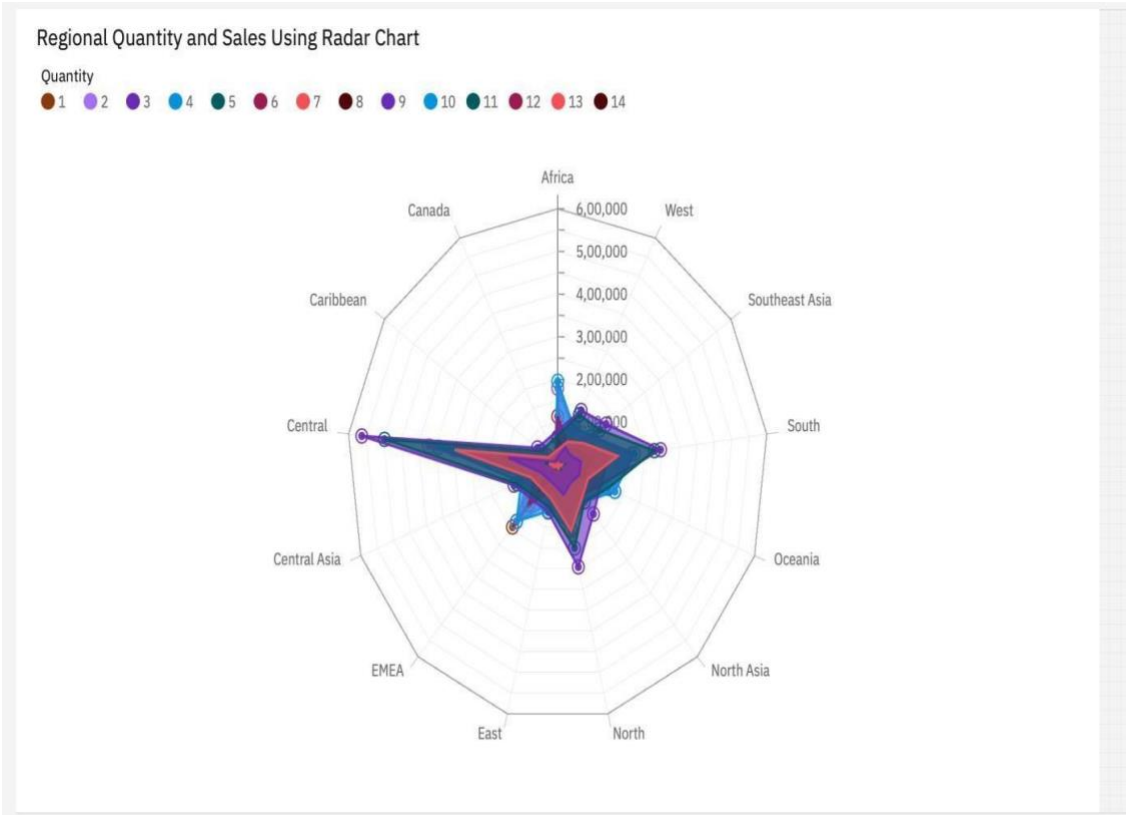
Sales By Segment Analysis



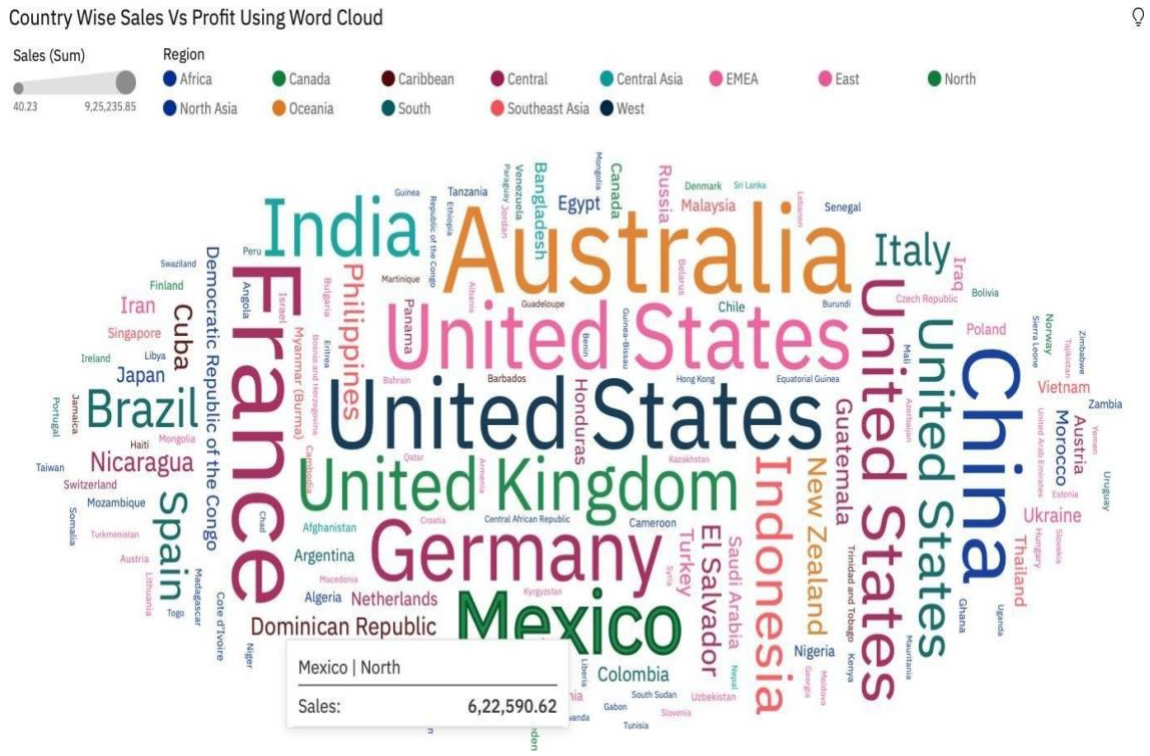
Sales Vs Profit By Countries



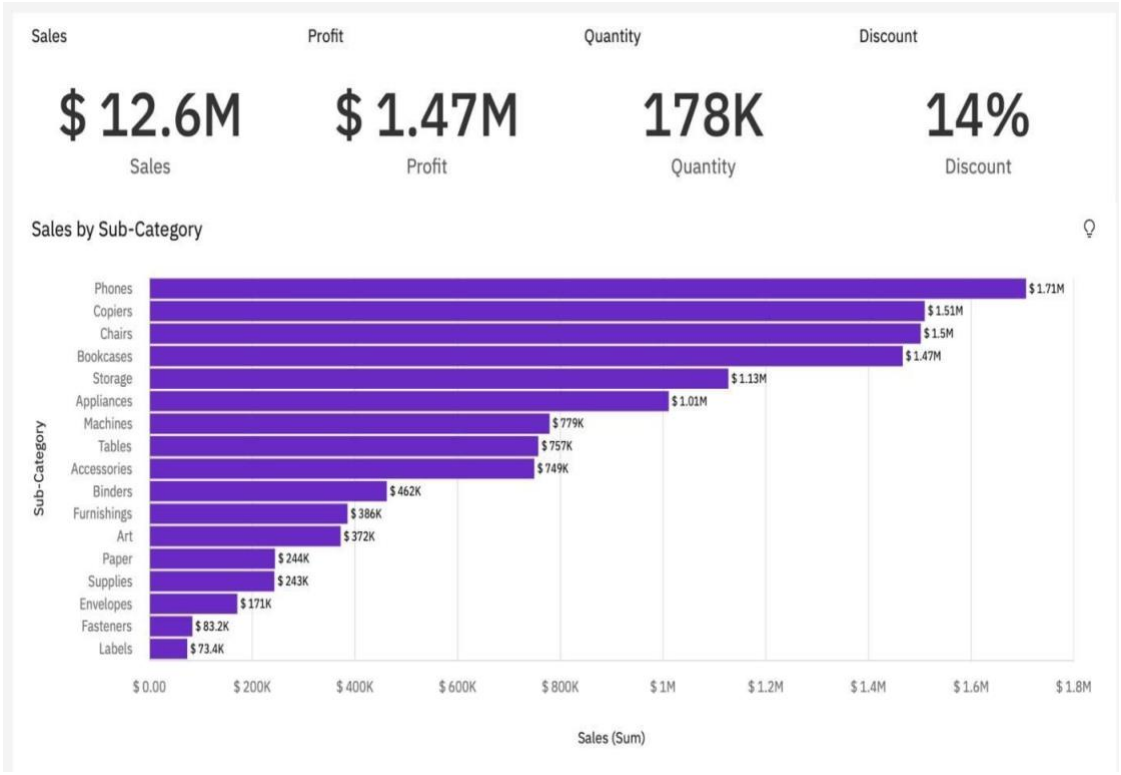
Regional Quantity And Sales Using Radar Chart



Country Wise Sales Vs Profit Using Word Cloud



Sales Dashboard



8.TESTING

1. TEST CASES

Date	03-Nov-22
Team ID	PNT2022TMID04980
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 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 or Click on the	User is able to fill in their data into to the Contact Us form	Working as	Pass	-	N	-	Kiruthika, Manujula
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_01	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_01	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,

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	Severity1	Severity2	Severity3	Severity4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won'tFix	0	0	0	1	1
Totals	24	9	11	26	71

Test Case Analysis

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

Section	TotalCases	Not Tested	Fail	Pass
PrintEngine	7	0	0	7
ClientApplication	51	0	0	51
Security	2	0	0	2
OutsourceShipping	3	0	0	3
ExceptionReporting	9	0	0	9
FinalReportOutput	4	0	0	4
VersionControl	2	0	0	2

9. RESULTS

1. PREFORMANCE 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 Visulizations / Graphs - 7-8 visualization/6-7 graphs
2.	Data Responsiveness	Users and Analyst or Developers
3.	Amount Data to Rendered (DB2 Metrics)	5 counrties
4.	Utilization of Data Filters	Sales ,profit, products, market rate and order id filtration
5.	Effective User Story	No of Scene Added - 30 user stories
6.	Descriptive Reports	No of Visulizations / Graphs - 4 visualizations/6 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

SOURCE CODE -https://github.com/IBM-EPBL/IBM-Project-27590-1660060709.

GITHUB - https://github.com/IBM-EPBL/IBM-Project-27590-1660060709.

DEMO VIDEO -https://drive.google.com/file/d/1Hr3sxsnVg32HbU1tVNrvsxVhSR-rpEx7/view?usp=drivesdk