Global Sales Data Analytics

PROJECT REPORT

SubmittedBy

Team ID : PNT2022TMID49365
Team Leader : GOWTHAMAN M
Team Member 1 : ARJUN T
Team Member 2 :ARUN JEEVA C
Team Member 3:NITHISH PANDI M

In Partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

GOVERNMENT COLLEGE OF ENGINEERING, BODINAYAKANUR-625582

ANNA UNIVERSITY, CHENNAI
600025

ListOfContents

1. Indroduction:

1.1 ProjectOverview:

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,try to understand a few things like, Customer Analysis and Product Analysis of this GlobalSuperStore.

1.2 Purpose:

By the end of this Project ,you will:

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

2.LITRATURESURVEY

2.1Existing Problem:

Crafting agood sales pitch from sales data analysis can be difficult.Getting the right data, hitting the right client pain points, crystallizing why your servicesare better than the competitors, all takes hard work. One of the best ways we've found to build ago sales pitch is to use data you already have.

In the digital world, there is no shortage of data, which translates into no shortage of potential competitive in sights and advantages.With databases,dataware houses,corporate intranets,best practice sharing,web analytics, voice of the customer information,and QA or Six Sigma data, you are well-poised for discovering good information.

2.2References:

1. Han Jia wei, Micheline Kamberand Jian Pei ,"**Data Mining Concepts and Techniques** "in, MK Publications,2009.Showing Context Google Scholar

2. M.Tennekes and E.deJonge,"**Top-down Data Analysis with Treemaps"** Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP'11) ,pp.236-241,March2011.Showing Context Google Scholar

3. P.Hoek,**"Parallel Arc Diagrams :Visualizing Temporal Interactions "**, Journal of Social Structure,vol.12,2011.Showing Context Google Scholar

2.3Problem Statement Definition:

Our goal is to design and create Dashboard using the Superstore Sales data (which is really close to reality) to provide answers to the following questions:

1. What are the performance indicators values for the past month ? It's necessary for stock taking and comparing it against the same period last year.

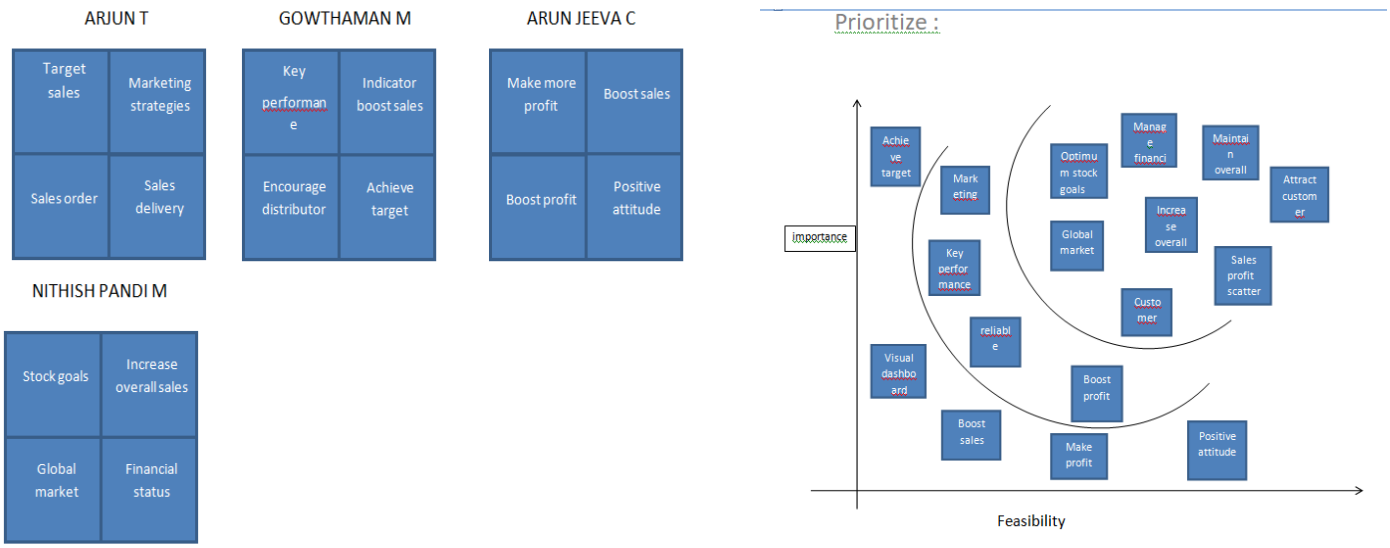
2. What key factors do affect profit growth?

3. What categories ,subcategories, products and clients generate more profits , and what ones that bring losses?

3. IDEATION&PROPOSEDSOLUTION:

3.1 EmpathyMapCanvas:

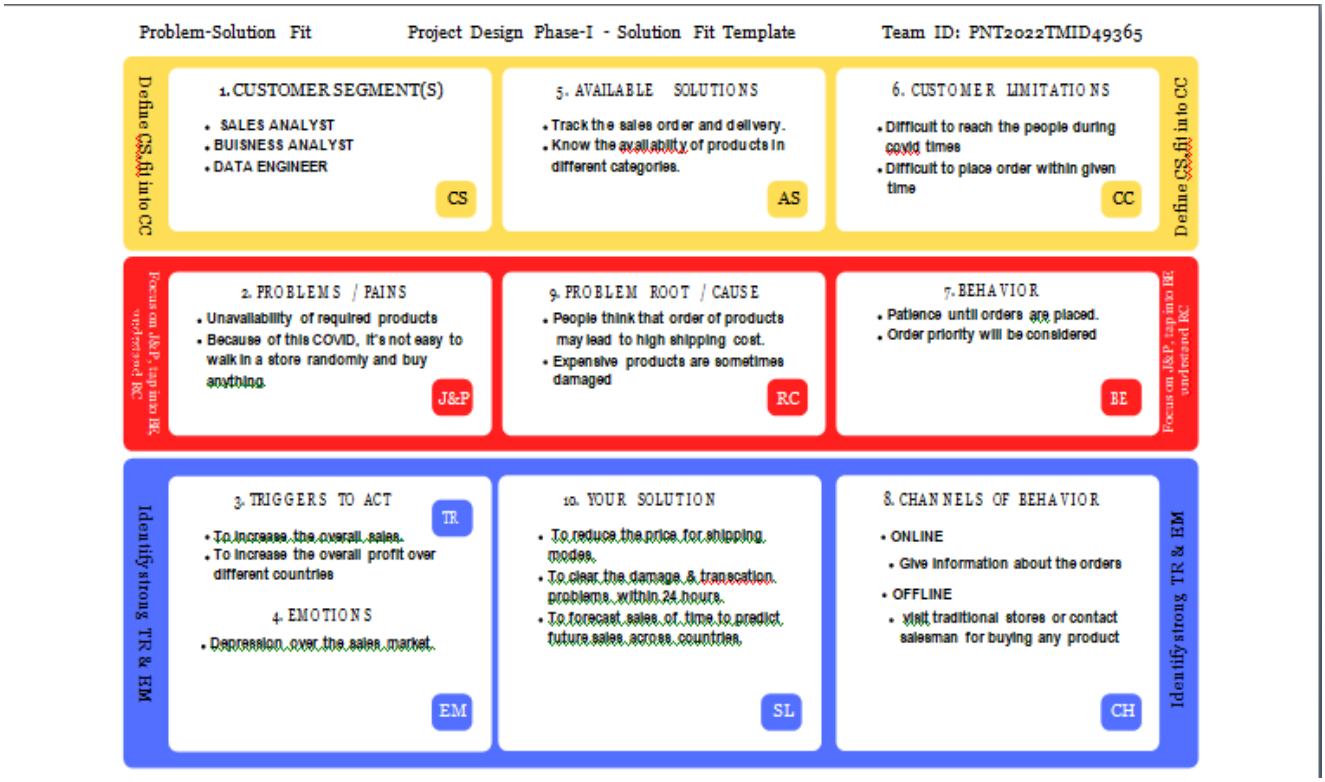
3.2 Ideation&Brainstorming:



3.3 ProposedSolution:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	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, try to understand a few things like, Customer Analysis and Product Analysis of this Global Super Store.
2.	Idea / Solution description	The described solution is by using IBM cognos we can display all the records and previous year global sales of product names, category and sub category as a graphical representation.
3.	Novelty / Uniqueness	we are going to provide discounts to the customers to increase the sales by providing free door step delivery of products to customers.
4.	Social Impact / Customer Satisfaction	Customer should know the available products and nearest location of the shops which gives the idea to customer for purchase.
5.	Business Model (Revenue Model)	This method focuses on the actual sales numbers from the customers. This helps to determine which products are top performers and multiplying the shop and increasing the product quantity.
6.	Scalability of the Solution	Using this approach, the price of products across the world are kept same so the customers will be reliable.

3.4 ProblemSolutionfit:



4. REQUIREMENT ANALYSIS:

Project Design Phase-II
Solution Requirements (Functional & Non-functional)

Date	03 October 2022
Team ID	PNT2022TMID49365
Project Name	Project – Global sales data analytics
Maximum Marks	4 Marks

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Download the dataset	Get the data from the given resource
FR-2	Data pre processing	Fill missing values, Remove duplicate values
FR-3	Choose the tool for visualization	IBM Cognos analytics is chosen
FR-4	Data visualization	Required graph, charts are chosen for visualization and
FR-5	Prepare dashboards	Dashboards, story boards and reports are created in IBM Cognos analytics

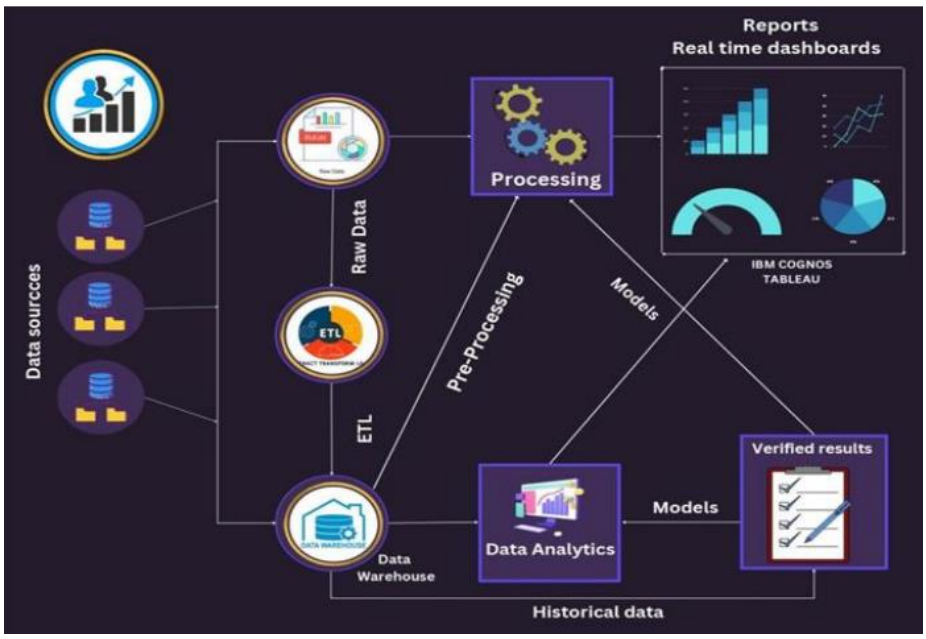
Non-functional Requirements:

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

FR 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**	Connecting the data to the software and further process.
NFR-4	**Performance**	The analysed information is recorded and updated.
NFR-5	**Availability**	The tool is only available for the authorized persons to create, update, remove and the record customer information.
NFR-6	**Scalability**	Everyday activities are monitored for the growth of work. Analytic tool should support even the size of data is increased.

5. PROJECTDESIGN:

5.1 Data Flow Diagrams:



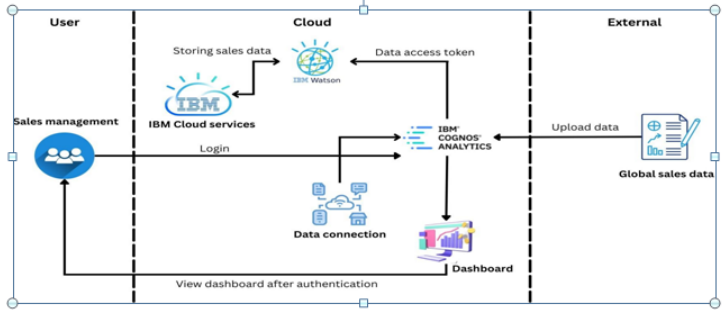
5.2 Solution&TechnicalArchitecture:

Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	03 October 2022
Team ID	PNT2022TMID49365
Project Name	Project – Global sales data analytics
Maximum Marks	4 Marks

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table2



5.3 UserStories:

User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Sales analyst/Customer	User Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can get the details of customer based on Requirements.	High	Sprint-1
			As a user, I will receive confirmation email once I have registered for the application		High	Sprint-1
	Login	USN-2	As a user, I will log in to the desired application using login credentials.	I can get the order details, sales and profit dashboards, reports, and stories used for customer analysis and product analysis.	Low	Sprint-2
	Dashboards	USN-3	As a sales analyst/customer, I can view the important sales and profit of the products and other information in real-time.		High	Sprint-1
Administrator		USN-4	As an administrator, I can access the database of the customers, sales reports, and a secure interface.	Customer relationship is managed and responsible for maintaining, updating the data.	High	Sprint-2

6. PROJECT PLANNING& SCHEDULING:

6.1 SprintPlanning&Estimation:

TITLE	DESCRIPTION	COMPLETED DATE
Literature Survey & Information Gathering	Prepare Literature survey for the selected project & gathering information	1 NOVEMBER 2022
Prepare Empathy Map	Prepare Empathy Map Canvas to capture the user Pains & Gains, Prepare list of problem statements	1 NOVEMBER 2022
Ideation	List the by organizing the brainstorming session and prioritize the top 3 ideas based on the feasibility & importance.	1 NOVEMBER 2022
Proposed Solution	Prepare the proposed solution document, which includes the novelty, feasibility of idea, business model, social impact, scalability of solution, etc.	7 NOVEMBER 2022
Problem Solution Fit	Prepare problem - solution fit document.	7 NOVEMBER 2022
Solution Architecture	Prepare solution architecture document.	7 NOVEMBER 2022

6.2 SprintDeliverySchedule:

Project Planning Phase

Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

Date	18 October 2022
Team ID	PNT2022TMID49365
Project Name	Global sales data analysis
Maximum Marks	8 Marks

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

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	Arjun Gowthaman Arun Jeeva Nithish Pandi
		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	Low	
		USN-3	As a user, I will log in to the desired application using login credentials.	1	Medium	
Sprint-2	Pre processing	USN-4	As a user, I can do the data cleaning process.	2	High	Arjun Gowthaman Arun Jeeva Nithish Pandi
		USN-5	As a user, I can perform Extract, Transform Load (ETL) process.	2	High	
Sprint-3	Dashboard	USN-6	As a user, I can upload the data of global sales for analysis.	1	Medium	

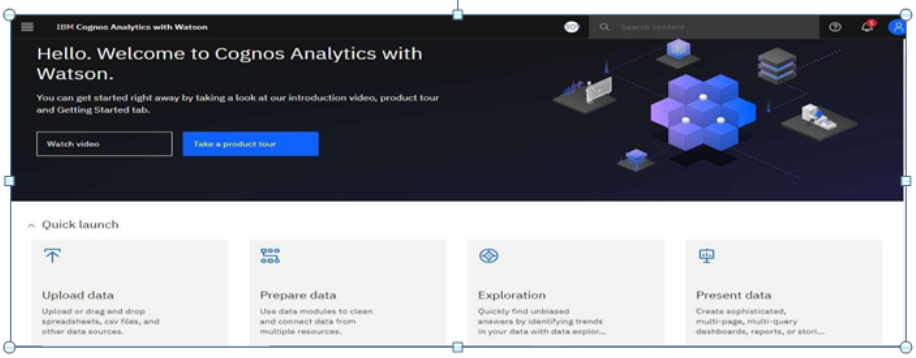
Sprint1:

DATA COLLECTION:

Sprint -1:

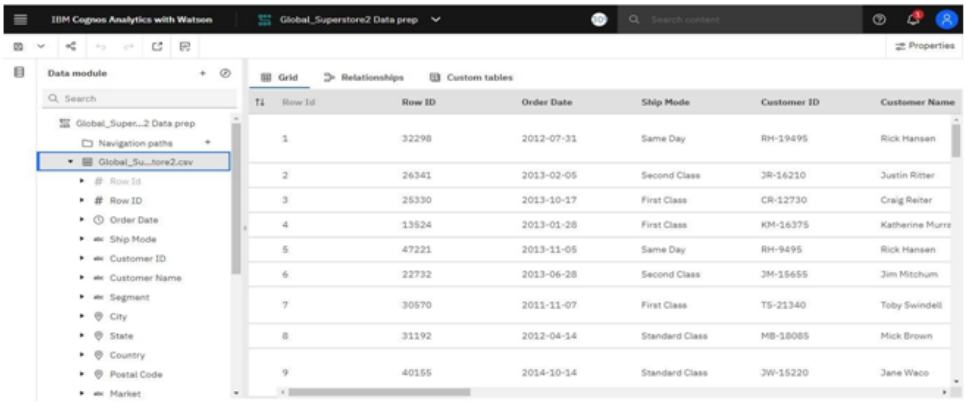
Registration and Data upload

Download the Dataset: https://www.kaggle.com/datasets/apoorvaappz/global-super-store-dataset



DATA PREPARATION

Prepare the Dataset



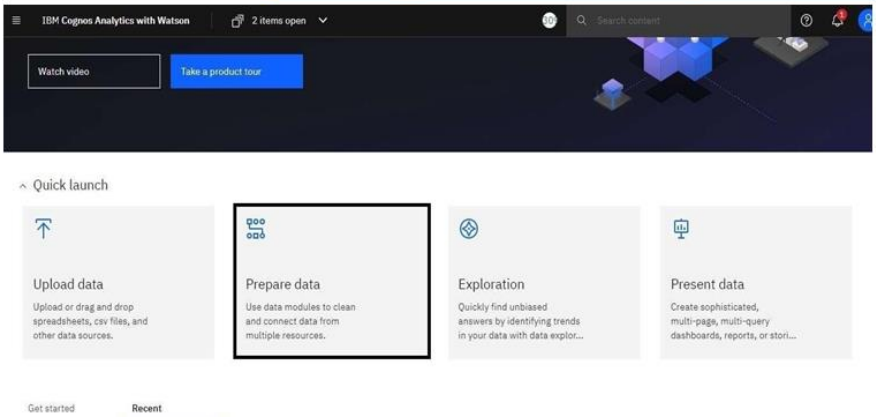
Sprint2:

DATA PREPROCESSING:

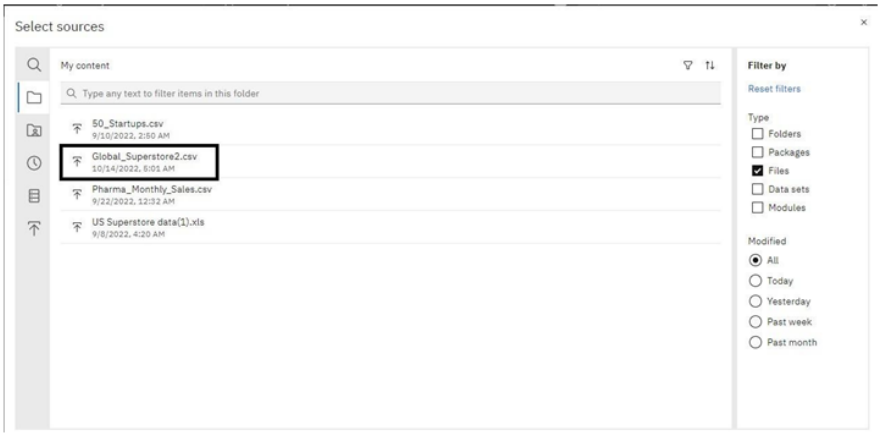
Sprint -2:

Registration and Data upload

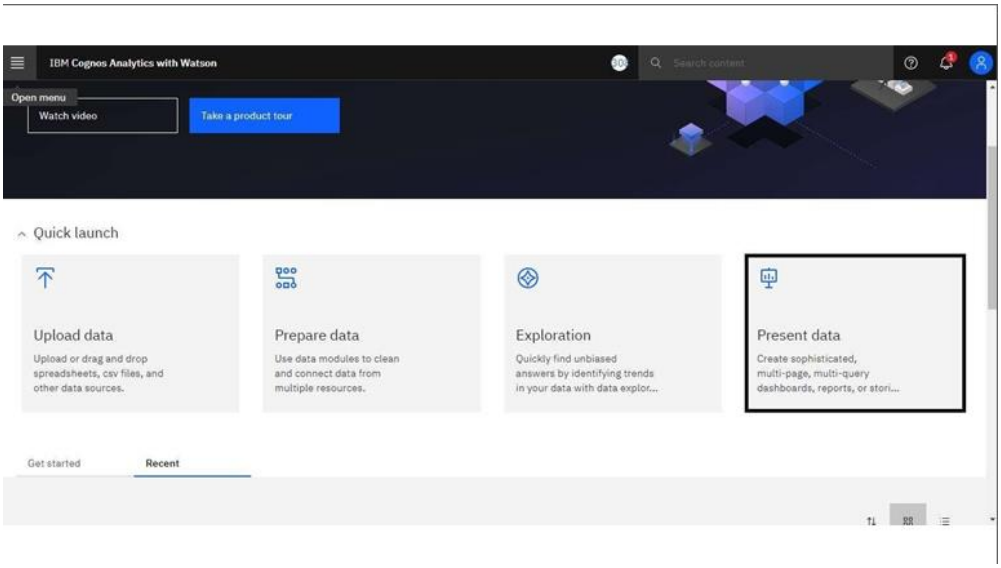
Download the Dataset: https:https://www.kaggle.com/datasets/apoorvaappz/global-super-store-dataset

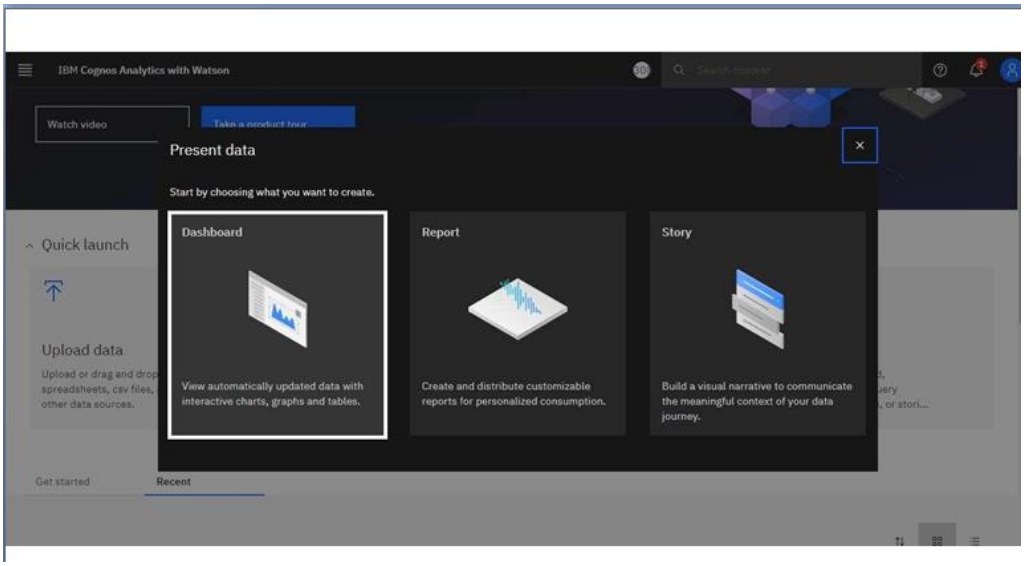


SELECTING THE SOURCE



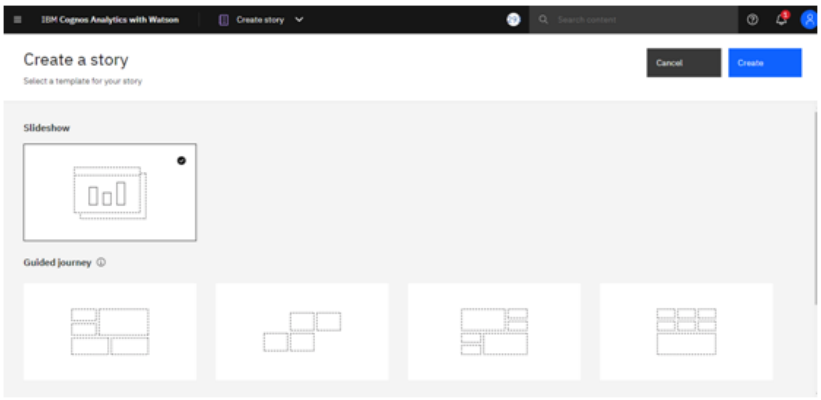
Sprint3:


IBM Cognos Analytics with Watson
Open menu
Watch video
Take a product tour
Quick launch
Upload data
Upload or drag and drop spreadsheets, csv files, and other data sources.
Prepare data
Use data modules to clean and connect data from multiple resources.
Exploration
Quickly find unbiased answers by identifying trends in your data with data explor...
Present data
Create sophisticated, multi-page, multi-query dashboards, reports, or stori...
Get started
Recent

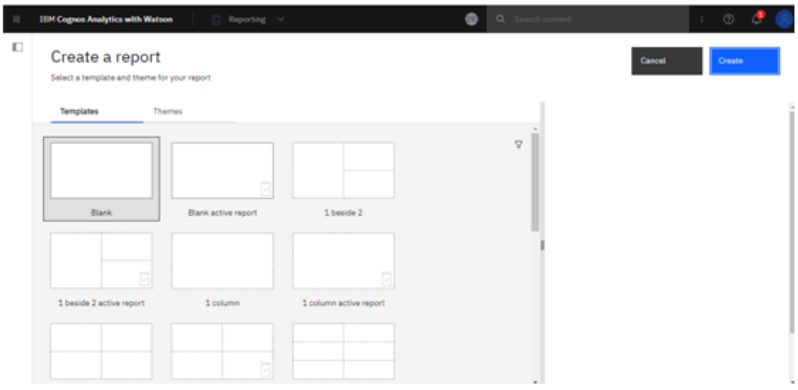

IBM Cognos Analytics with Watson
Watch video
Present data
Start by choosing what you want to create.
Dashboard
View automatically updated data with interactive charts, graphs and tables.
Report
Create and distribute customizable reports for personalized consumption.
Story
Build a visual narrative to communicate the meaningful context of your data journey.
Quick launch
Upload data
Get started
Recent

Sprint4:

STORY



REPORT

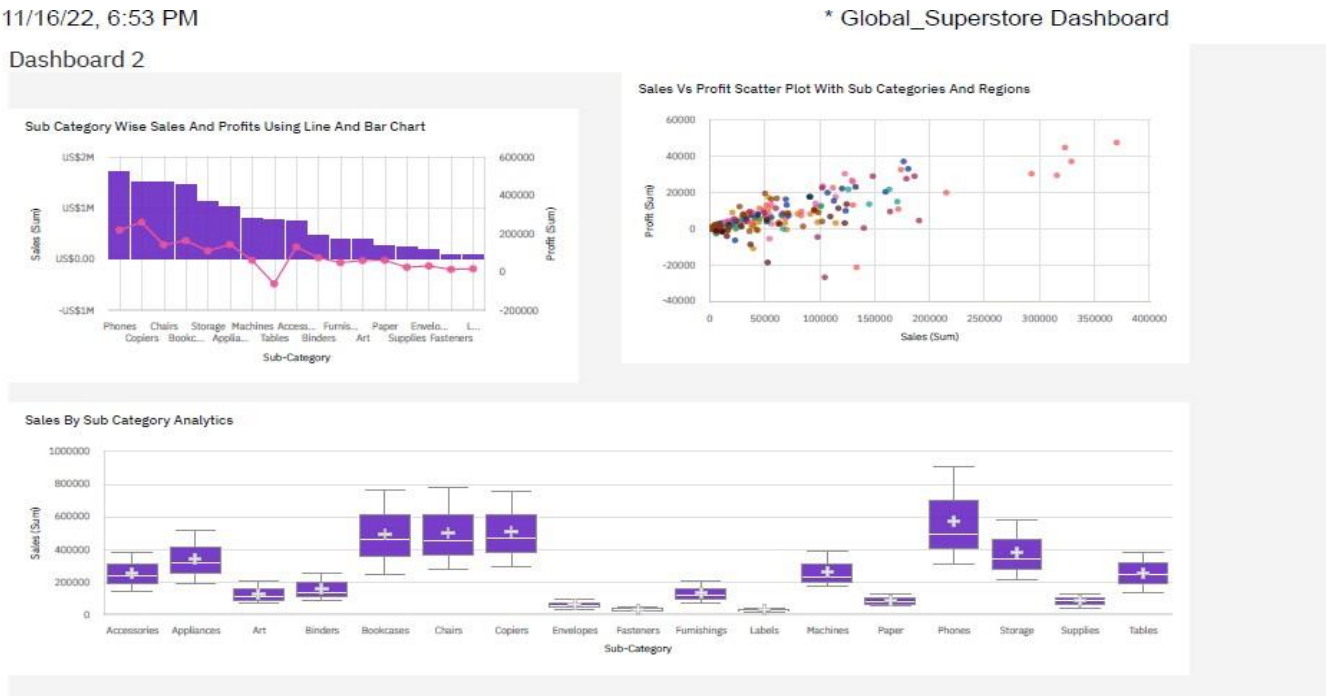


7. RESULTS:

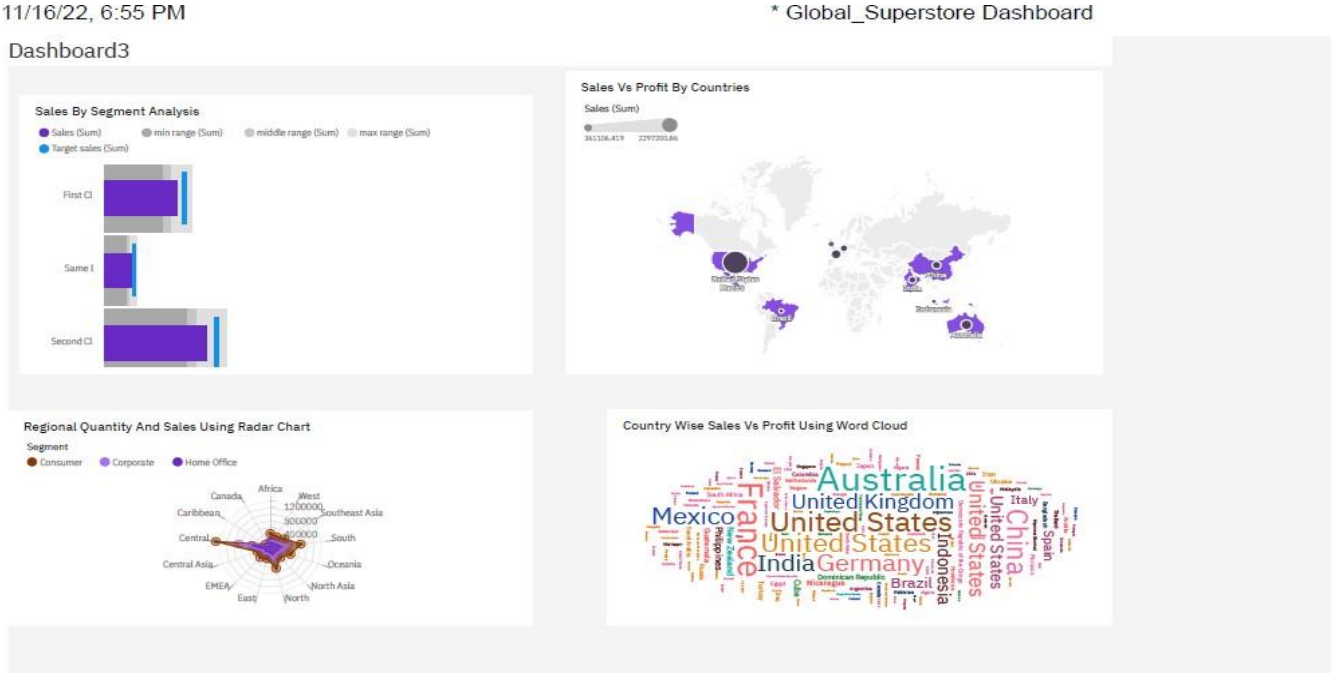
7.1 PerformanceMetrics:

Dashboard1

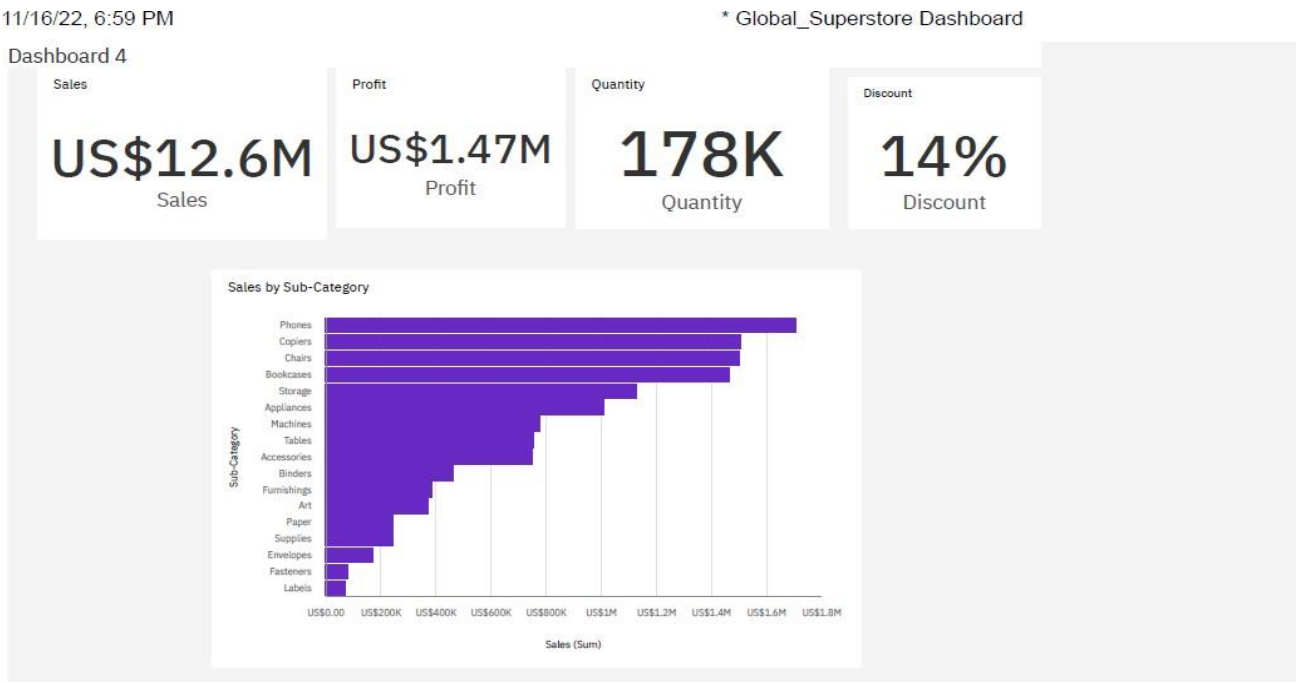
Dashboard2



Dashboard3



Dashboard4



8. ADVANTAGES & DISADVANTAGES:

ADVANTAGES:

i) It was the cost efficiency project.

ii) Receive full-scale services Maximize presentation

iii) It was the timing saving project for peoples.

DISADVANTAGES:

i) The lack of data security is the big disadvantages in this project.

ii) Risk of choosing the wrong provider.

9. CONCLUSION:

By implementing this analytics solution, the company brought their competitive and sales data reporting in-house, cut cost sand in creased 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 50to70%.

They are now able to analyse raw data themselves, respond more quickly tochanges 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 out sourced solution, reducing turn around 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 salesforce.

The business also experienced an increase in the overall understanding of theirsales data throughout the organization. The company now has great flexibility in thepresentationoftheirsalesandcompetitivedata,whilealsobeingabletointegratesalesdatawithotherkeydatapointsfortheorganization.

10. FUTURESCOPE:

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

GitHub & Project :

GithubLink: https://github.com/IBM-EPBL/IBM-Project-5164-1658749867