PROJECT REPORT

Team ID	PNT2022TMID41454		
Project Name	Global Sales Data Analytics		
Team Members	P.Poovarasan - 620619104023 A.Ashokkumar - 620619104001 V.Gayathri - 620619104002 P.Gomathi - 620619104004 S.Pavithra - 620619104021		

1. INTRODUCTION

- 1.1. Project Overview
- 1.2. Purpose

2. LITERATURE SURVEY

- 2.1. Existing problem
- 2.2. References
- 2.3. Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

- 3.1. Empathy Map Canvas
- 3.2. Ideation & Brainstorming
- 3.3. Proposed Solution
- 3.4. Problem Solution Fit

4. **REQUIREMENT ANALYSIS**

- 4.1. Functional Requirements
- 4.2. Non-functional Requirements

5. **PROJECT DESIGN**

- 5.1. Data Flow Diagrams
- 5.2. Solution & Technical Architecture
- 5.3. User Stories

6. PROJECT PLANNING & SCHEDULING

- 6.1. Sprint Planning & Estimation
- 6.2. Sprint Delivery Schedule

7. CODING & SOLUTION

- 7.1. Feature 1
- 7.2. Feature 2
- 7.3. Feature 3

8. **TESTING**

- 8.1. Test Cases
- 8.2. User Acceptance Testing
- 9. **RESULTS**
- 10. ADVANTAGES & DISADVANTAGES
- 11. CONCLUSION
- 12. FUTURE SCOPE
- 13. APPENDIX
 - 13.1. Source Code
 - 13.2. Github & Project Demo Link

1. INTRODUCTION

1.1 Project Overview

Shopping online is currently the need of the hour. Because of this COVID, it is 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. Global Sales covers all activities involved in selling a product or service to a consumer or business. It is important for sales and marketing teams to review their strategies and performance in order to make improvements. Understanding performance with sales data analytics helps sales and marketing teams to review their strategies and performance in order to make improvements. Sales analytics provides valuable information like Customer Analysis and Product Analysis to improve sales methodologies. Users create multiple analytical graphs/charts/Visualizations. One way to measure performance is with sales analytics.

1.2 Purpose

Global Sales covers all activities involved in selling a product or service to a consumer or business. It is important for sales and marketing teams to review their strategies and performance in order to make improvements. One way to measure performance is with sales analytics. Sales data analytics refers to the use of technology to collect and use sales data to identify actionable insights. It is used to identify, optimize, and increase sales. An efficient sales model that generates higher revenue for the business.

2. LITERATURE SURVEY

		Technique	Findings
Title & Author(s)	Year		_
Big Data Analytics and Deep Learning Based Sentiment Analysis System for Sales Prediction - Khatiwada, Aamod and Kadariya, Pradeep and Agrahari, Sandip and Dhakal, Rabin.	2019	Big arldata apl (BDA) com ons in e- ce.	Merits: Used to understand complex datasets in a matter of time with beautiful visual representations. Demerits: Lack of security since large data processed simultaneously
COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions - Sheng, Jie and AmankwahAmoah, Joseph and Khan, Zaheer and Wang, Xiaojun	2021	Descriptive and diagnostic analytics, Predictive analytics	Merits: By comparing with machine learning models, we find that the proposed model is superior to others. Demerits: The experiment only considers the features of the product and does not consider external influences, such as the impact of regulations on sales. It uses only small dataset.
Sales Forecasting Based on	2020	CatBoost algorithm.	Merits: The search stops
CatBoost - Jingyi Ding, Ziqing Chen.			when no improvements over the current best solution have been foun in 300 iterations. Demerits: Dataset is

Developing and lead-assetted	2020	Dia data	NA - vita - Ti
Developing and Implementing	2020	Big data s, R tool.	Merits: The
Big Data Analytics in Marketing -		S, K (00).	proposed method
Dina Darwish			is based
Dina Dai Wish			similarity
			measurem
			without complex tra
			so that forecast ca
			completed in a s time, and
			performs w small
			scale data Demerits:
			The
			experiment only conside
			the features of the
			product and does not
			consider external
			influences, such as
			the impact of
			regulations on sales. It uses
			only small dataset.
Social media big data analytics for	2020	Definitional ig	Merits: Captured linearity
demand forecasting:		aspects of b data	and non linearity better
development and case		analyti	than ARIMA and ARNN
implementation of an		(BDA) in	gave the best result of
innovative framework-Iftikhar,		ecommerce	565
Rehan and Khan, Mohammad			RMSE. Demerits: Hybrid
Saud			Technique can fail if
			nonlinear model fails to
			capture residue patterns

2.1 Existing problem

1. Lack of security since large data processed simultaneously

- 2. Hybrid Technique can fail if nonlinear model fails to capture residue patterns
- 3. The experiment only considers the features of the product and does not consider external influences, such as the impact of regulations on sales. It uses only small dataset.
- 4. Dataset is limited
- 5. The experiment only considers the features of the product and does not consider external influences, such as the impact of regulations on sales. It uses only small dataset.

2.2 References

- Big Data Analytics and Deep Learning Based Sentiment Analysis System for Sales Prediction
 Khatiwada, Aamod and Kadariya, Pradeep and Agrahari, Sandip and Dhakal, Rabin.
- 2. COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions Sheng, Jie and Amankwah-Amoah, Joseph and Khan, Zaheer and Wang, Xiaojun
- 3. Sales Forecasting Based on CatBoost Jingyi Ding, Ziqing Chen.
- 4. 2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP)-Wisesa, Oryza and Adriansyah, Andi and Khalaf, Osamah Ibrahim.
- 5. Developing and Implementing Big Data Analytics in Marketing Dina Darwish
- 6. Social media big data analytics for demand forecasting: development and case implementation of an innovative framework-Iftikhar, Rehan and Khan, Mohammad Saud

2.3 Problem Statement Definition

- 1. Structured data focuses on demographic data including name, age, gender, date of birth, address, and preferences, unstructured data includes clicks, likes, links, tweets, voices, etc.
- 2. The methodological innovations in studying big data analytics and. We provide insights on methods in descriptive/diagnostic, predictive and prescriptive analytics, and how they can be leveraged to study 'black swan' events such as the COVID-19-related global crisis.
- 3. It proposed a sales forecasting system based on CatBoosting. The algorithm is trained on the Walmart sales dataset, by far the largest dataset in this field. We performed effective feature engineering to boost prediction accuracy and speed.

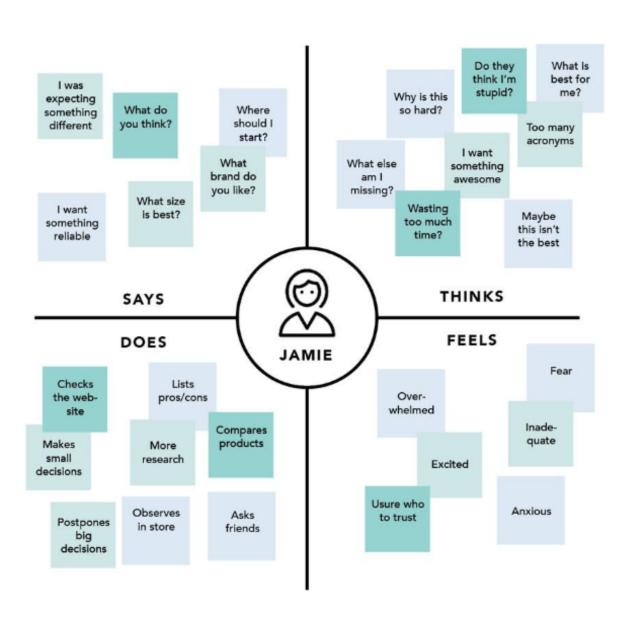
- 4. The results of this analysis are expected to generate reliable, accurate and effective forecasting data, a valuable resource for sales predictions. It shows good accuracy in forecasting.
- 5. Companies take informative business decisions in different fields, such as, health care, banking, manufacturing, media and entertainment, education and transportation and many others.
- 6. Social media big data offers insights that can be used to make predictions of products' future demand and add value to the supply chain performance

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

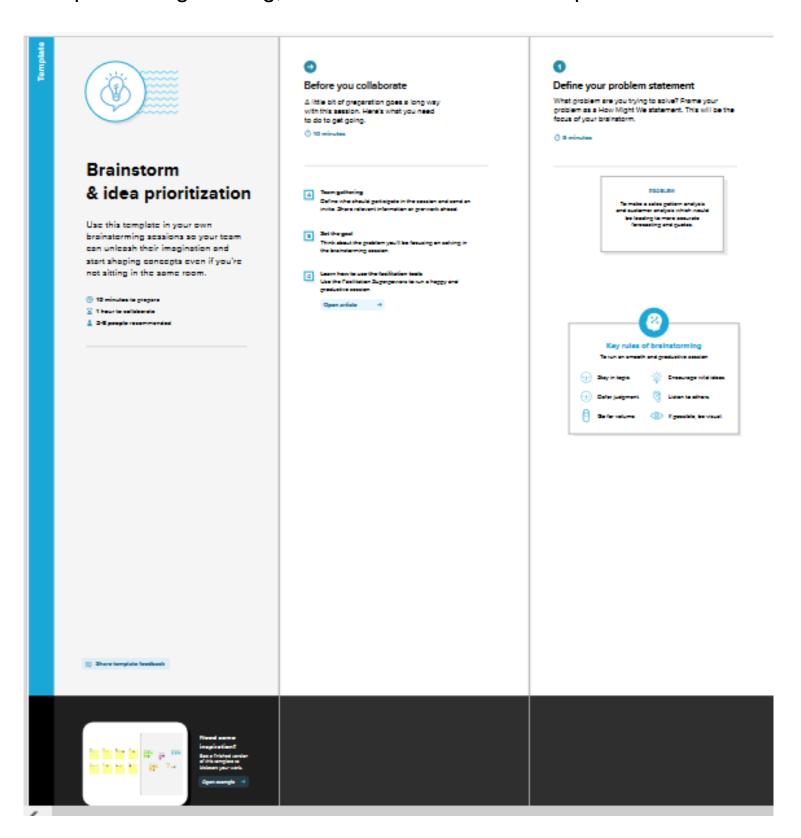
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviors and attitudes. It is a useful tool to help teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.

EMPATHY MAP:-

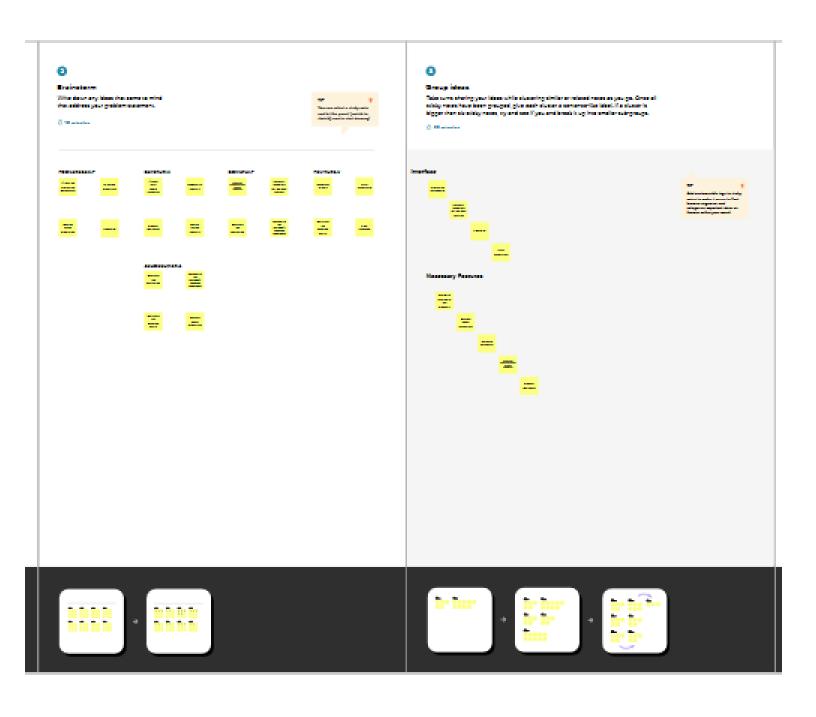


Brainstorm:-

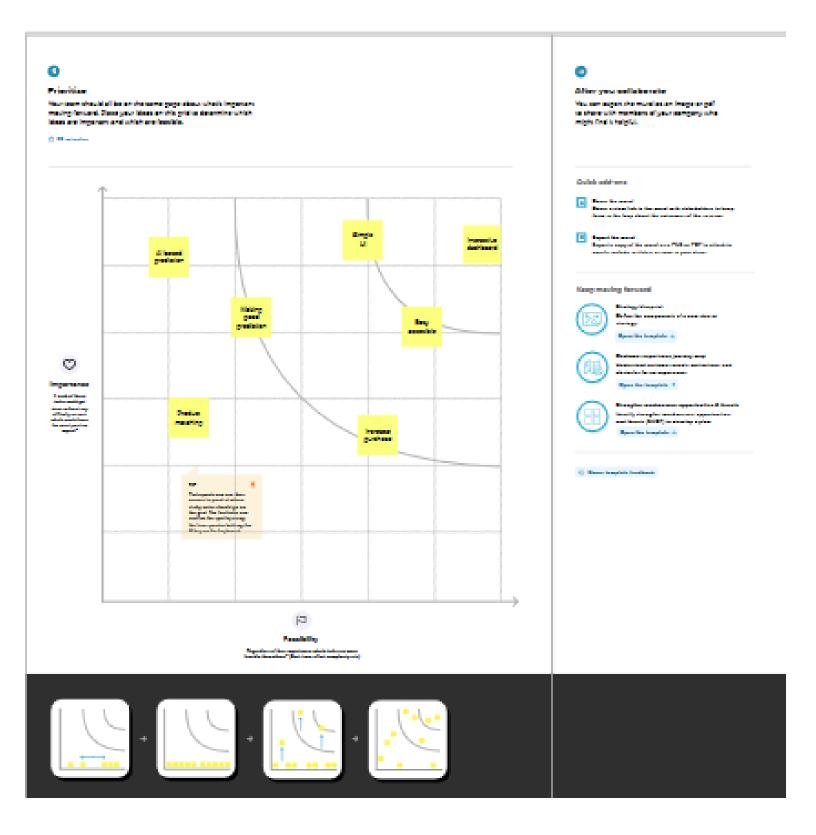
step1:Team gathering, collabration and select the problem statement



Step-2: Brainstrom, idea listing & grouping



Step-3: Idea prioritization



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 Al 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 Basic: Limited features targeting startups and individuals Standard: Limited premium features. Target customers- Medium Scale businesses. 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

3.4 Problem Solution Fit

The problem solution fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. It helps to identify behavioral patterns and recognize on sales.

Purpose:

- Solve complex problems in a way that fits the state of your customers.
- Succeed faster and increase your solution adoption by tapping into existing mediums and channels of behavior.
- Sharpen your communication and marketing strategy with the right triggers and messaging.
- Increase touch-points with your company by finding the right problem-behavior fit and building trust by solving frequent annoyances, or urgent or costly problems.
- Understand the existing situation in order to improve it for your target group.

Solution fit:

1. CUSTOMER SEGMENT(S) A Business owner who would like to understand more about his business performance in global scale. Sales Manager looking for smart sales strategies	Difficult to place order within given time Need to check input file structure before uploading	The competition perform analytics and display Dashboard with autogenerated insights. Spreadsheet tools like Excel, Google Sheets
Unavailability of required products What analysis to perform to be useful and how to perform them?	PROBLEM ROOT CAUSE Customer satisfaction Expensive products are sometimes damaged People think that order of products may lead to high shipping cost.	Patience until orders are placed. Collecting sales data and using office software to analyze it
3. TRIGGERS To increase the overall sales. To increase the overall profit over difference outries 4. EMOTIONS: BEFORE / AFTER BEFORE: Anxiety, Decision fatigue, Laz AFTER: Clear mind, Peacefulness	To reduce the price for shipping modes. To clear the damage & transaction problems within 24 hours. To forecast sales of time to predict future sales across countries	8. CHANNELS of BEHAVIOUR LI ONLINE Give information about the orders 5.2 OFFLINE Visit traditional stores or contact salesman for buying any product

4. REQUIREMENT ANALYSIS

4.1 Functional Requirements

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Website Registration through Gmail
FR-2	User Confirmation	Confirmation via Email
FR-3	User Login	Login via Gmail and Password
FR-4	Generating Report	User can view the product details

4.2 Non-functional Requirements

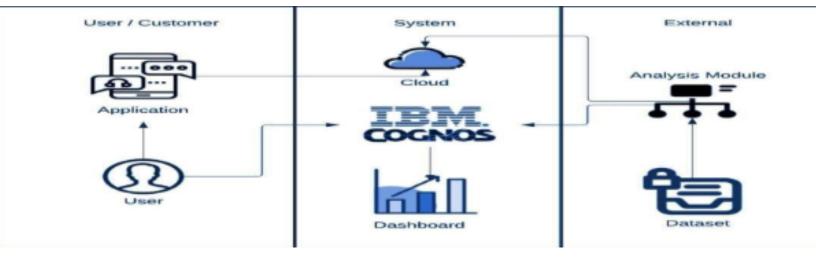
FR No.	Non-Functional Requirement	Description

NFR-1	Usability	This service will have a simple and user-fr graphical interface. Use be able to understan all the features easily.
NFR-2	Security	The main security concern is for users login information is end to end encryption should be used to avoid hacking.
NFR-3	Reliability	It has high reliability because when the system is disconnected or internet connection lost, it should save all the process of the users made.
NFR-4	Performanc e	A good internet speed while browsing the product it had high performance with efficiency.
NFR-5	Availability	It will be available 24 hours a day and seven days a week. User access anywhere at any time.
NFR-6	Scalability	A Many users can access the website simultaneously.

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 & Technical Architecture

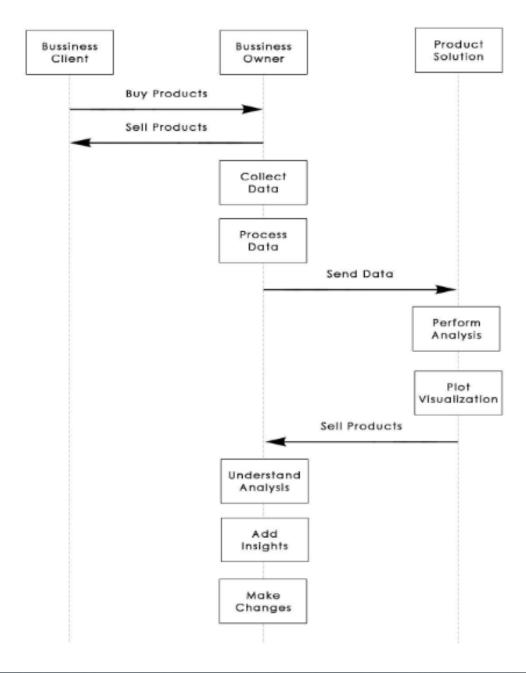
Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.

Solution Architecture Diagram:

Project Design Phase-I Solution Architecture

Date	1 October 2022
Team ID	PNT2022TMID41454
Project Name	GLOBAL SALE DATA ANALYTICS



5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the applicationby 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

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 confirmingmy password.	2	High	P. POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-1	Login	USN-2	As a user, I need valid credentials to log in tomy application.	1	High	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
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	P.POOVARASAN A. ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-2	Uplond dataset	USN-4	As a user, I can view the data of the products	1	Low	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-2	Data Preparation	USN-5	As a user, I need to filter it for Datavisualization.	3	High	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-2	Data visualization	USN-6	As a user, I can easily visualize the data in theform of charts.	4	Medium	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-3	Dashboard	USN-7	As a user, I can view the summary of theproduct sales by the help dashboard.	2	Medium	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
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 therespective region	4	High	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint-3	Dashboard	USN-9	As a user, I must be able to gain insights fromthe charts/graphs through a variety of relationships established in the dashboard.	4	Medium	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA
Sprint- 4	Prediction	USN-10	As a user, I see the prediction of the specific product's future sales expectation.	4	Medium	P.POOVARASAN A.ASHOKKUMAR V.GAYATHRI P.GOMATHI S.PAVITHRA

6.2 Sprint Delivery Schedule

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

7.CODING & SOLUTION

Feature 1 - Embedding dashboard to web app:

```
<link rel="icon" type="image/png" href="../assets/img/favicon.p</pre>
ng">
   <title>
     Global Sales Data Analytics
   </title>
   <!--
            Fonts and icons
                                 0
   <link href=https://fonts.googleapis.com/css?family=Open+Sans:30</pre>
0,400,600,700 rel="stylesheet" />
   <!-Nucleo Icons •
   <link href="../assets/css/nucleo-icons.css" rel="stylesheet" />
   <link href="../assets/css/nucleo-svg.css" rel="stylesheet" />
<!-Font Awesome Icons 7
   <script src=https://kit.fontawesome.com/42d5adcbca.js crossorig</pre>
in="anonymous"></script>
   <link href="../assets/css/nucleo-svg.css" rel="stylesheet" />
   <!-CSS Files 7
   <link id="pagestyle" href="../assets/css/soft-</pre>
uidashboard.css?v=1.0.6" rel="stylesheet" />
 </head>
 <body class="g-sidenav-show bg-gray-100">
   <aside class="sidenav navbar navbar-vertical navbar-expandxs"
border-0
            border-radius-xl
                                 my-3 fixed-start
                                                          ms-3
id="sidenavmain">
```

<div class="sidenay-header">

<I class="fas fa-times p-3 cursor-pointer textsecondary
opacity-5 position-absolute end-0 top-0 d-none d-xlnone" ariahidden="true" id="iconSidenav"></i></i>

</div>

<hr class="horizontal dark mt-0">

<div class="collapse navbar-collapse w-auto "
id="sidenavcollapse-main">

class="nav-item">

<title>shop </title>

<path class="color-background opacity-</pre>

6" d="M46.7199583,10.7414583 L40.8449583,0.949791667 C40.4909749,0 .360605034 39.8540131,0 39.1666667,0 L7.83333333,0 C7.1459869,0 6. 50902508,0.360605034 6.15504167,0.949791667 L0.280041667,10.741458 3 C0.0969176761,11.0460037 -1.23209662e-05,11.3946378 1.23209662e-05,11.75 C-0.00758042603,16.0663731 3.48367543,19.5725301 7.80004167,19.58333 33 L7.81570833,19.5833333 C9.75003686,19.5882688 11.6168794,18.872 6691 13.0522917,17.5760417 C16.0171492,20.2556967 20.5292675,20.25 56967 23.494125,17.5760417 C26.4604562,20.2616016 30.9794188,20.26 16016 33.94575,17.5760417 C36.2421905,19.6477597 39.5441143,20.170 8521 42.3684437,18.9103691 C45.1927731,17.649886 47.0084685,14.842 8276 47.0000295,11.75 C47.0000295,11.3946378 46.9030823,11.0460037 46.7199583,10.7414583 Z"></path>

51667 C40.2482243,46.9451667 41.125,45.9498589 41.125,44.7220845 L 41.125,22.2822926 C40.4887822,22.4116582 39.8442868,22.4815492 39. 198,22.4912623 Z"></path>

</g></g> </g> </g> </svg> </div> Dashboard class="nav-item"> <div class="icon icon-shape icon-sm

<title>office</title>

<g transform="translate(-1869.000000, -</pre>

293.000000)" fill="#FFFFFF" fill-rule="nonzero">

class="nav-item">

<title>credit-card</title>

<g transform="translate(-2169.000000, 745.000000)" fill="#FFFFFF" fill-rule="nonzero">

.6666667,23.2916667 L35.8333333,23.2916667 L35.8333333,26.875 Z"></path>

Story

</aside>

<main class="main-content position-relative max-height-vh100 h-100 border-radius-lg ">

<!-Navbar **7**

<div class="container-fluid py-1 px-3">

```
<nav aria-label="breadcrumb">
       0 me-sm-6 me-5">
       <h6 class="font-weight-bolder mb-
0">Sales Dashboard</h6>
      </nav>
      <div class="collapse navbar-collapse mt-sm-0 mt-2 me-md-</pre>
0 me-sm-4" id="navbar">
        <div
             class="ms-md-auto pe-md-3
                                   d-flex
                                           align-
itemscenter">
       </div>
```

</div>

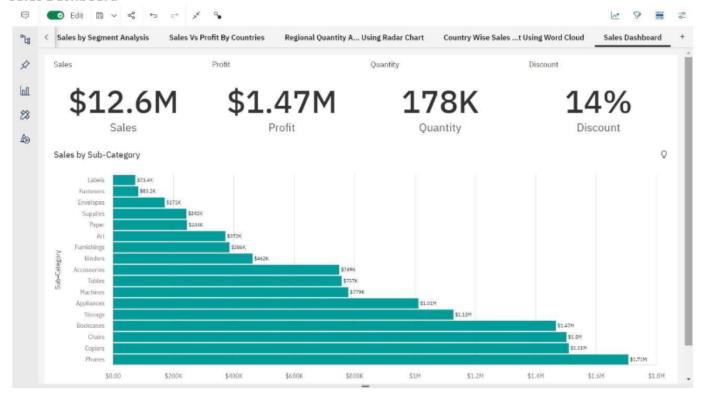
</div>
</nav>
<!-End Navbar •

<!-End Toggle Button •

</body>

</html>

Sales Dashboard



Feature 2 - Embedding report to web app:

<iframe

src=https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FDa
ta%2BModule%2FSales%2BReport&closeWindowOnLastView=true&u
i appbar=false&ui_navbar=false&shareMode=embedded&act
ion=edit width="1200" height="600" frameborder="0" gesture="media"
allow="encrypted-media" allowfullscreen=""></iframe>

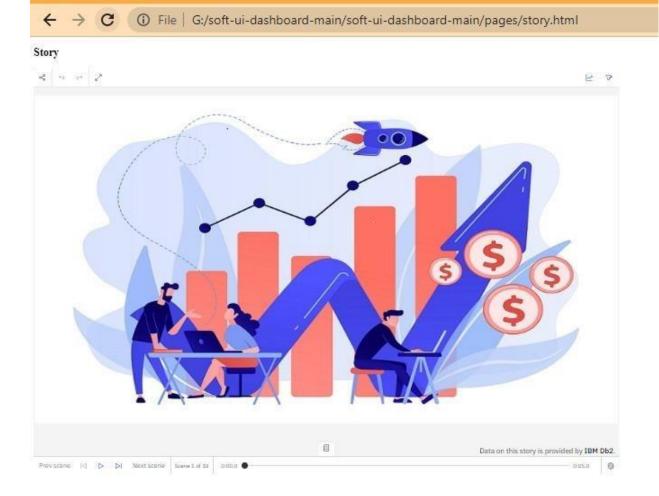
© Edit © v c so v v so v segment wise Sales, Profit and Quantity Sales by Market +

Sales by Order Priority
Ord

Feature 3 - Embedding story to web app:

<iframe

src=https://us3.ca.analytics.ibm.com/bi/?perspective=story&pa
thRef=.my folders%2FData%2BModule%2FSales%2BStory&closeWindow
OnLastView=true&ui_appbar=false&ui_navbar=false&share
Mode=embedded&action=view&sceneId=model0000018485276975_0
0000000&sceneTime=0 width="1300" height="900" frameborder="0"
gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>



8. TESTING

8.1 Test Cases

Test Scenarios

- Verify user able to see login page
- Verify user able to login to application or not?
- Verify user able to navigate to create your account page?

- Verify user able to recovery password
- Verify login page elements

Access visualizations

- User able to see dashboard
- User able to see report
- User able to see stories

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	9	6	3	1	19	
Duplicate	1	o	3	0	4	
External	6	5	0	2	13	
Fixed	13	3	4	18	38 0	
Not Reproduced	0	0	0	0		
Skipped	0	0	0	0		
Won't Fix	0	o	0	0	0	
Totals	29	14	10	21	74	

Test Case Analysis

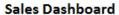
This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	5	0	0	5
Client Application	38	0	0	38
Security	3	0	0	2
Outsource Shipping	9	0	0	9
Exception Reporting	5	0	0	5
Final Report Output	4	0	0	4
Version Control	3	0	0	3

9. RESULTS

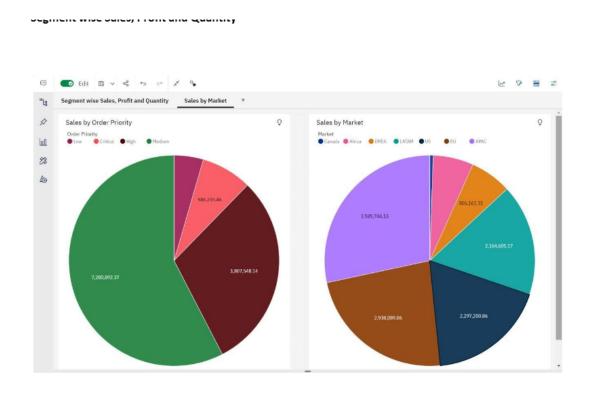
Screenshots of web application:

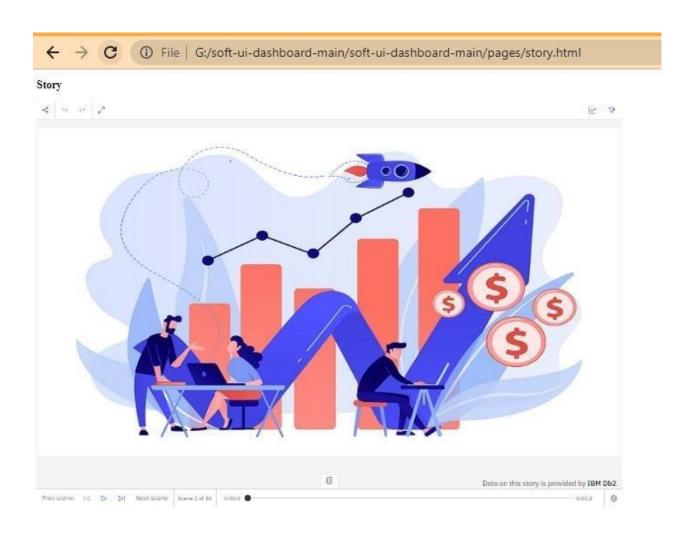
DASHBOARD:





REPORT:





10. ADVANTAGES & DISADVANTAGES

Advantages:

- It is used to identify, optimize, and forecast sales.
- Sales data will help a company to take a future decision in terms of inventory management, marketing activities, schemes or offers to be rolled and changes in manufacturing processes if applicable.
- An efficient sales model that generates higher revenue for the business.

• Better prediction, Profit function performance.

• Helps to review their strategies and performance in order to make improvements.

Disadvantages:

Sales pattern can be changed

insufficient data may lead to wrong path.

 data may have been collected for historical reasons may not be suitable to answer the questions that we ask today.

• business users do not see results immediately

11. CONCLUSION

Analyzing sales helps businesses in understanding their most profitable products and the ones that are not moving, most profitable customers, and potential sales opportunities thereby providing sales which match customer needs and meets their satisfaction. An efficient sales model that generates higher revenue for the business. It helps in the perception of profit about particular product and perception of sales in different locations and times.

12. FUTURE SCOPE

Use the technology to collect and use sales data to derive actionable insights. It is used to identify, optimize, and forecast sales. An efficient sales model that generates higher revenue for the business. An efficient sales model that generates higher revenue for the business.

13. APPENDIX 13.1 Source

Code

kaggle api to gather the dataset



Githup Link : Link : http://githup.com/IBM-EPBL/INM-project-36585-1660296381