GLOBAL SALES DATA ANALYTICS

(TEAM ID: PNT2022TMID36362)

PROJECT REPORT

Submitted by

GOWTHAM PRABHAKARAN R
PRABHAKARAN G
JACOB JABAKUMAR S
SURYA J
HARISH S R

in partial fulfilment for the award of degree of BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



JAYA ENGINEERING COLLEGE, CHENNAI

ANNA UNIVERSITY: CHENNAI 600 025

NOVEMBER 2022

ACKNOWLEDGEMENT

We acknowledge our sincere thanks to Head of the Department Mr. KUMARAN M for giving us valuable suggestion and help towards us throughout this Project.

We are highly grateful to thank our Project coordinator **KUMARAN M** and our Project Evaluator **Sundararaju G JAYA ENGINEERING COLLEGE** for the coordinating us throughout this Project.

We are very much indebted to thank all the faculty members of Department of Computer science and Engineering in our Institute, for their excellent moral support and suggestions to complete our Project work successfully.

Finally, our acknowledgment does our parents, sisters, and friends those who had extended their excellent support and ideas to make our Project a pledge one.

GOWTHAM PRABHAKARAN R PRABHAKARAN G JACOB JABA KUMAR S SURYA J HARISH S R

ABSTRACT

Data is being generated very rapidly due to increase in information in everyday life. Huge amount of data get accumulated from various organizations that is difficult to analyse and exploit. Data created by an expanding number of sensors in the environment such as traffic cameras and satellites, internet activities on social networking sites, healthcare database, government database, sales data etc., are example of huge data. Processing, analysing and communicating this data are a challenge. Online shopping websites get flooded with voluminous amount of sales data every day. Analysing and visualizing this data for information retrieval is a difficult task. Therefore a system is required which will effectively analyse and visualize data. This paper focuses on a system which will visualize sales data which will help users in applying intelligence in business, revenue generation, and decision making, managing business operation and tracking progress of tasks. The main Achievement of this Global Sales Data Analytics is to reduce the manufacturing cost of the raw material and improved the sales forecasting by identifying the key factors like the total sales revenue on a monthly and quarterly basis on the region and the sale amount. And the decision support system Data Warehousing Project is focused on analysing the entire business process. In order to provide critical information like daily revenue, Weekly Revenue, Monthly Revenue, total sales, goals, information on employees and vision of the company developed Business Intelligence System.

INDEX

1. INTRODUCTION

- 1. Project Overview
- 2. Purpose

2. LITERATURE SURVEY

- 1. Existing problem
- 2. References
- 3. Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

- 1. Empathy Map Canvas
- 2. Ideation & Brainstorming
- 3. Proposed Solution
- 4. Problem Solution fit

4. REQUIREMENT ANALYSIS

- 1. Functional requirement
- 2. Non-Functional requirements

5. PROJECT DESIGN

- 1. Data Flow Diagrams
- 2. Solution & Technical Architecture
- 3. User Stories

6. PROJECT PLANNING & SCHEDULING

- 1. Sprint Planning & Estimation
- 2. Sprint Delivery Schedule
- 3. Reports from JIRA

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

7.1 Feature 1

8. TESTING

8.1 User Acceptance Testing

8.2 Test Cases

9. **RESULTS**

- 1. Performance Metrics
- 10. ADVANTAGES & DISADVANTAGES
- 11. CONCLUSION
- 12. FUTURE SCOPE 13. APPENDIX

GitHub & Project Demo Link

1. INTRODUCTION

1.1 PROJECT OVERVIEW

Category : Data Analytics Team

ID: PNT2022TMID36362

SKILLS REQUIRED:

IBM COGNOS, IBM CLOUD, IBM DB2, HTML, CSS, Data Cleaning.

PROJECT DESCRIPTION:

Data analysis is the process of cleaning, changing, and processing raw data and extracting actionable, relevant information that helps businesses make informed decisions. The main objective of this project is always to simplify the information available to the users. It should help you clearly understand your team's performance, sales trends and opportunities. In this project we have used the IBM COGNOS, IBM DB2 services for loading the dataset, data preparation and vizualization of the data. The first step of the process is to download the dataset from the kaggle. The next step of data preparation IBM COGNOS provides the tool for data preparation. The third step of your data analytics project begins after obtained the data. By diving deep, we would determine what the data have and how might use it to attain their initial aim. After preparing the data, dashboard story and report are created through the IBM COGNOS and they were embedded into a web page using HTML and CSS.

1.2 PURPOSE

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.

A data dashboard is an interactive analysis tool used by businesses to track and monitor the performance of their strategies with quality KPIs. Armed with real-time data, these tools enable companies to extract actionable insights and ensure continuous growth.

6

The main objective of the project is 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. The overall purpose of a data analytics dashboard is to make it easier for data analysts, decision makers, and average users to understand their data, gain deeper insights, and make better data-driven decisions.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

The aim of sales analytics is to predict revenue more accurately and make the most of the opportunities in your reach. It provides a visual representation of your most recent performance metrics. It gives you a concise view of results-based data like sales- to-date, sales-by-region, lead conversionrate, sales growth, and so on. Dashboards are an essential tool for any business with plans to increase revenue and set ambitious growth goals. Without a sales dashboard, you're left to analyse dizzying amounts of data on your own. Trying to compile all those sales analytics metrics manually is an impossibly exhausting task with a massive risk of critical human errors. IBM Cognos tools allows users to receive insights immediately with a simple setup, no training required, and dashboards for services like sales force, Google Analytics, and Microsoft Dynamics.

2.2 REFERENCES

S NO	PAPER TITLE	AUTHOR	MONTH / YEAR	LINK
	Impact of big data analytics on sales performan ce pharmaceut ial organizatio s: The role of customer relationship management capabilities	MuhammadSh ahbaz, RimshaZahid	April/2021	https:/www.researchgate.net/public ation/351153157_Imp
2.	Exploring the Impact of Sales Technology on Salesperson Performance	Eggert A,Serdaroglu	MAY/2006	https://doi.org/10.2753/MTP1069- 6679190204
3.	Data-driven decision making via sales analytics	J.Ricky ferguson	July/2020	www.academia.edu/48422249/Data_driven_decisio

4.	A Conceptural View of the Dynamics of GlobalSales	Allison Smith, Ricky Fergurson and David Fleming	2020	C:\Users\jbs\Downloads\sef-grant-executive-summary.pdf
5.	Salesperson Ambidexterity and Customer Satisfaction: Examining the role of customer demandingness, adaptive selling, and role conflict	Raj Agnihotri	2014	C:\Users\jbs\Downloads\Agnihot ri-Executive-Summary- Salesperson-Ambidexterity-and- Customer-Satisfaction.pdf

2.3 PROBLEM STATEMENT DEFINITION

PROBLEM STAEMENT:

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. In this I am trying to understand few things like, Customer Analysis and Product Analysis of this Global Super Store.

Customers Analysis:

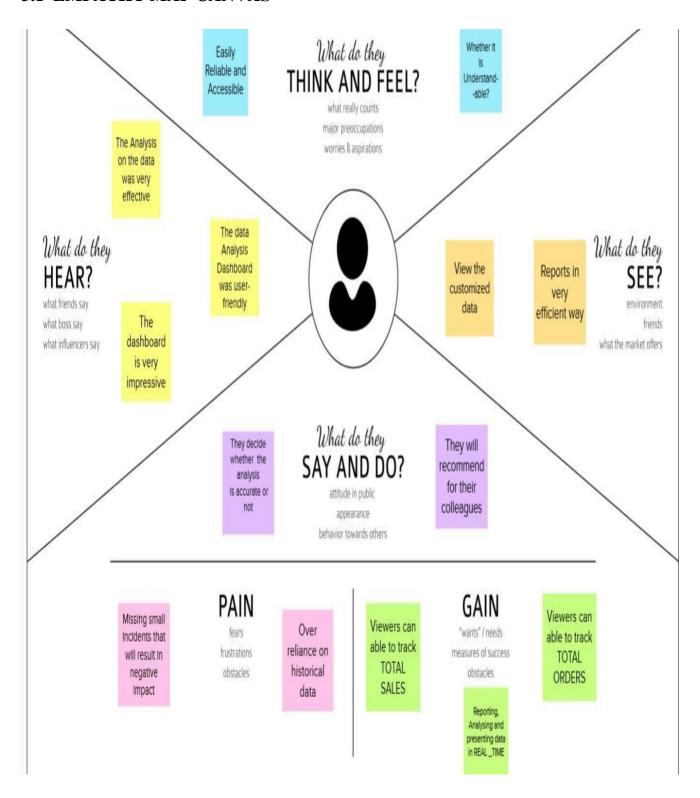
- Profile the customers based on their frequency of purchase calculate frequency of purchase for each customer.
- Do the high frequent customers are contributing more revenue.
- Are they also profitable what is the profit margin across the buckets.
- How the customer's are distributed across the countries?

Product Analysis:

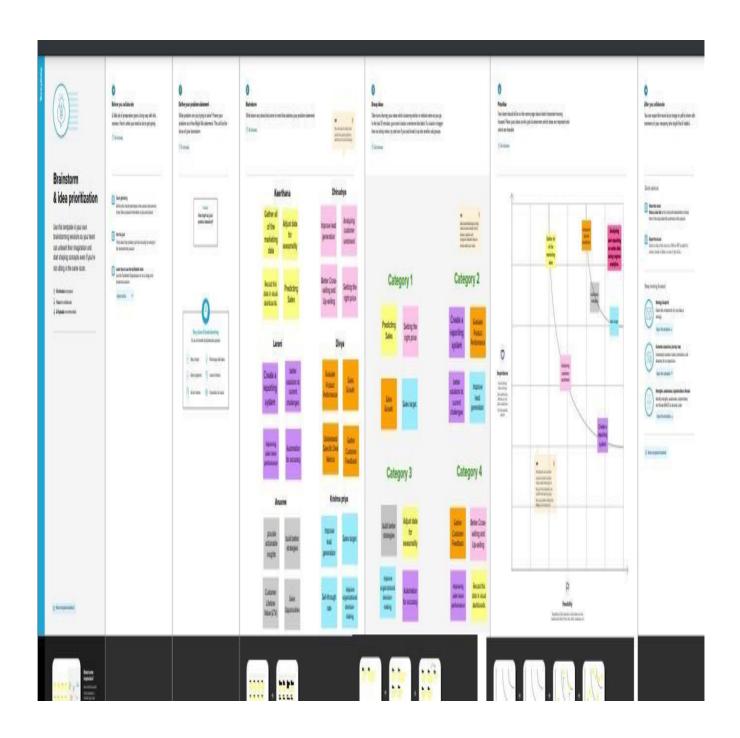
- 1. Root cause analysis
- 2. Quality of the product imported for further production.
- 3. Damage material Statement.
- 4. Collateral Damages (Including worker accidents while on work).
- 5. Defect product generation 6. Plant maintenance standard.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTROMING



3.3 PROPOSED SOLUTION

IDEA / SOLUTION DESCRIPTION:

A good sales dashboard is the solution. It organizes your most recent salesspecific data into easily understandable visual graphics saving your teams precious time and increasing understanding, motivation, and accuracy. With the right sales dashboards, you'll know exactly where you are, exactly how far away your current goals are, and where it may be necessary to tweak your sales strategy to achieve them.

UNIQUENESS / NOVALITY:

- Provides Real-Time Data.
- Can Help the Team Set Goals.
- Gives a Clear Overview of Sales Activity.
- Allows for the Identification of Sales Growth Opportunities.
- Identifies Opportunities for Improvement

SOCIAL IMPACT/CUSTOMER SATISFACTION:

- allow business decision-makers to review significant amounts valuable information at a glance.
- Mobile device accessible
- Customizable.
- Turn Data into Business Value.

BUSINESS MODEL (FINANCIAL BENEFIT):

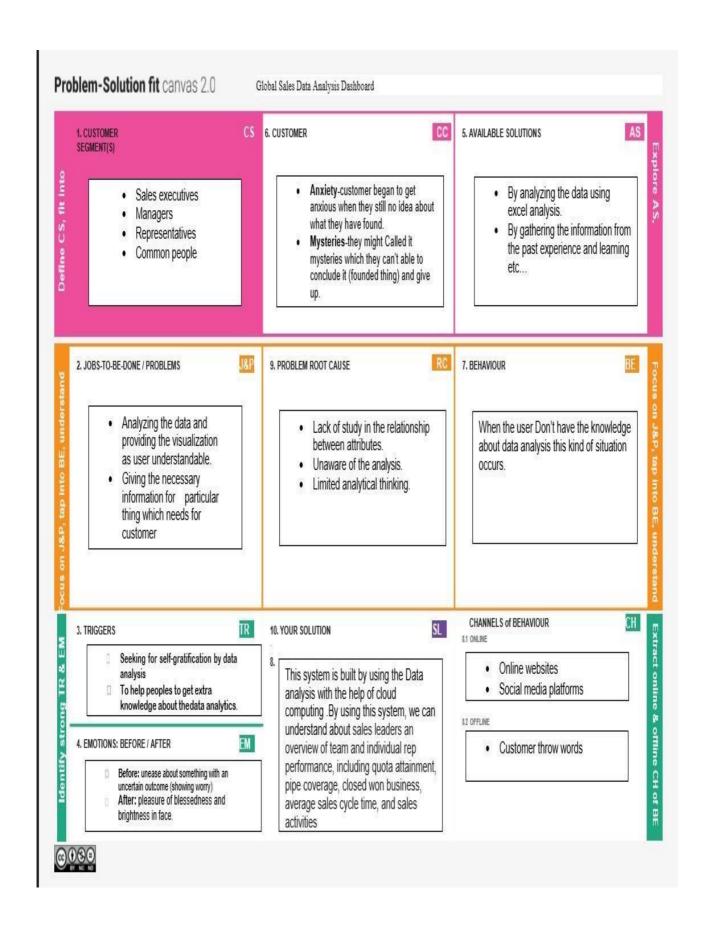
A Sales dashboard enables direct insight into your revenue driving force, allowing you to plan, implement and improve with data-based decisions.

SCALABILITY OF SOLUTION:

The great thing about Sales Analytics is that it gives you answers, and yosu need to ask the right questions. With accurate insight into current customers, a higher retention rate, as well

as increasing revenue, can be achieved. Having real-time insight into increasing and decreasing customers will allow your Sales Team to focus on the right clients at the right time, ensuring marketing opportunities are realised, and efforts are driven towards suitable clients. A Sales Dashboard helps you visualize your Sales data, which is helpful for efficient decision-making and performance analysis.

3.4 PROBLEM SOLUTION FIT



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

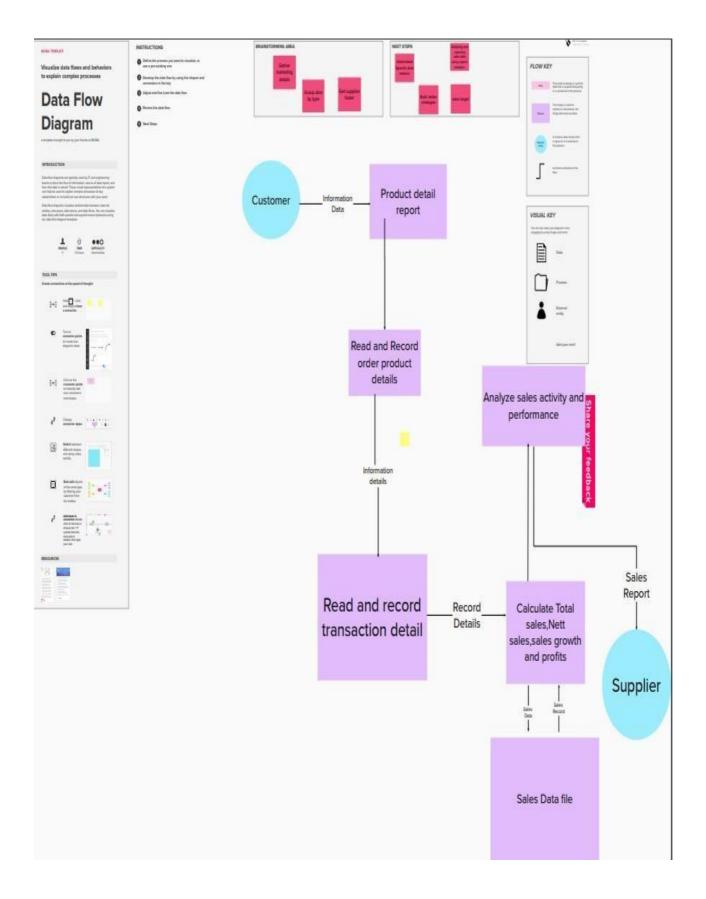
Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story/ Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail
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 userbenefits	Getting higher state of efficiency and also to know entire data analysis

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Optimized resources and it can be used by everyperson who needsto analyze theirsales trend.
NFR-2	Security	The model is more securable because it hasend to end encryption and the data are kept secure with high levelauthentication 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 websiteswithout 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 DATAFLOW DIAGRAM



5.2 SOLUTION AND TECHNICAL ARCHITECTURE

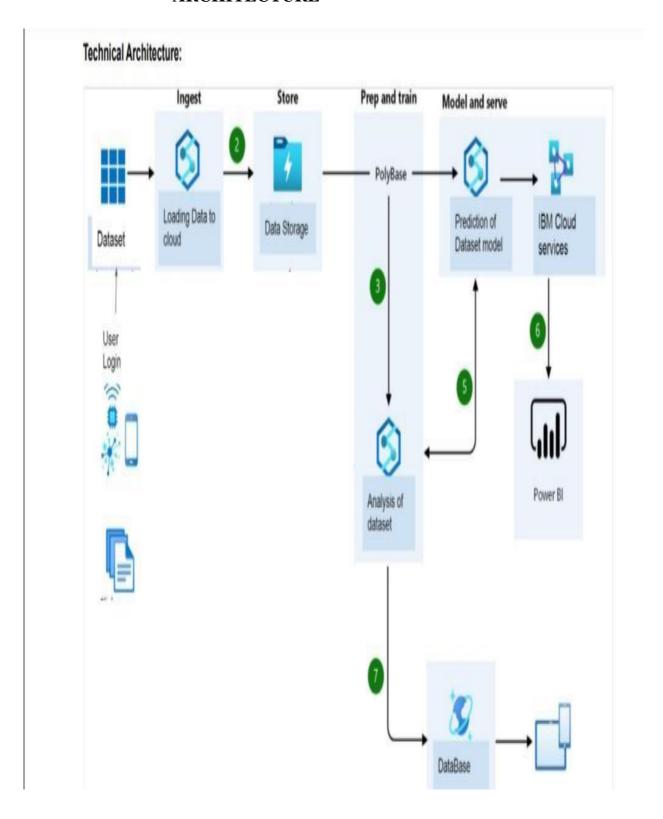


Table-1: Components & Technologies:

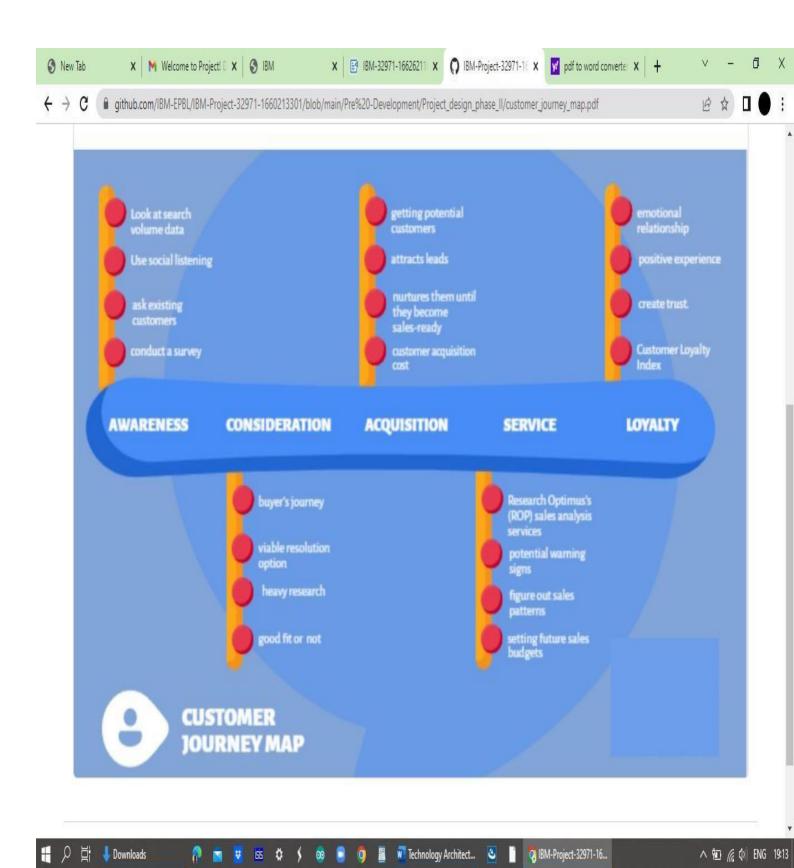
S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g.Web UI, MobileApp, Chatbot etc.	HTML, CSS, JavaScript
2.	Cloud Storage	Sales data of the customer's are uploaded to the cloud	Java / Python
3.	Working with Dataset	Uploading, Cleaning and Processing dataset	IBM Cognos Watson
4.	Data Exploration	Uploaded datais explored to identify trends	IBM Cognos Watson
5.		Multiple typesof graphs areshown according to customer data and requirements	IBM Cognos Dashboard andStory
6.	Cloud Database	Database Serviceon Cloud	IBM DB2, IBM Cloudant etc.

7.	Machine Learning Model	To do the predictive analysis on the input data	Predictive MachineLearning Model.
----	---------------------------	---	---

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology	
1.	Open-Source Frameworks	List the open-source frameworks used	IBM Cognos, IBM Cloud	
2.	Security Implementations	To securethe user information and data	256-bit AES algorithm etc.	
3.	Scalable Architecture	Provides scalability for the dataset	MS Excel Dataset, Kaggle, Google drive	
4.	Availability	Provides various typeof analysis and provides full support irrespective of platform	IBM Cognos	
5.	Performance	Withstand huge dataand process themwithout crashing.	IBM Cognos	

5.3 USER STORIES



6. PROJECT PLANNING AND SCHEDULING

6.1 Sprint Planning & Estimation

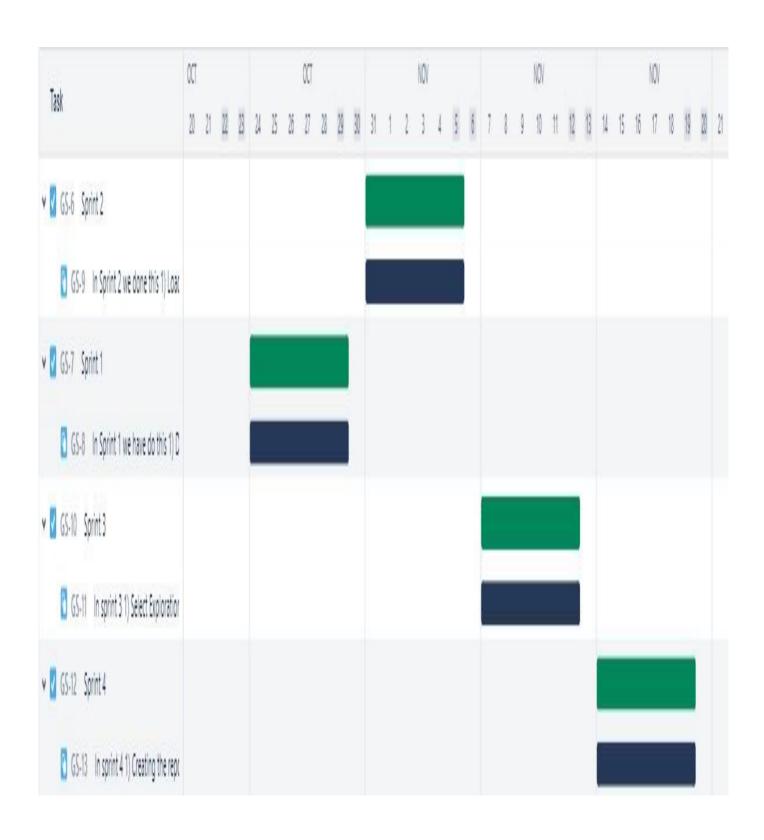
Sprint	Functional Requirement (Epic)	User Story Numb er	User Story / Task	Story Poin ts		Team Members
Sprint-1	Collect thedataset	USN-1	Download the dataset fromKaggle API	3	High	Gowthamprabhakaran.R, Prabhakaran.G, Jacobjabakumar.S, Surya.J,Harish.SR
Sprint-1	Understa nd the dataset	USN-2	To understand the data in dataset	2	Medi um	Gowthamprabhakaran.R, Prabhakaran.G, Jacobjabakumar.S, Surya.J,Harish.SR
Sprint-2	Loding the dataset	USN-3	Load the dataset in IBM cognosanalytics	5	High	Gowthamprabhakaran.R, Prabhakaran.G, Jacobjabakumar.S, Surya.J,Harish.SR
-	Preparation of data	USN-4	Prepare the datawith no nullvalues	2	Low	Gowthamprabhakaran.R, Prabhakaran.G, Jacobjabakumar.S, Surya.J,Harish.SR
Sprint-2	Perfor mIng calcula tions	USN-5	Create new calculation for perfect visualization	3	Medi um	Gowthamprabhakaran.R, Prabhakaran.G, Jacobjabakumar.S, Surya.J,Harish.SR

\mathbf{S}_{1}	\mathcal{C}	USN-6	Visualize		Gowthamprabhakaran.R,
	Visualization		thedata for	um	Prabhakaran.G,
			the user to		Jacobjabakumar.S,
			understande		Surya.J,Harish.SR
			asily.		

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Durati on	Sprint Start Date	Sprint End Date (Plan ned)	Story Points Completed (as on PlannedE nd Date)	Sprint Release Date(Actua l)
Sprin t-1	5	6 Days	24 Oct 2022	29 Oct 2022	5	29 Oct 2022
Sprin t-2	10	6 Days	31 Oct 2022	05 Nov 2022	10	05 Nov 2022
Sprin t-3	15	6 Days	07 Nov 2022	12 Nov 2022	15	12 Nov 2022
Sprin t-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

6.3 REPORTS FROM JIRA



7. CODING AND SOLUTIONING

7.1 FEATURE 1

The iframes are essentially a section of your page that you "cut out." In the space that you have cut out of the page, you can then feed in an external webpage.

In essence, an iframe is another browser window set inside your web page. You see code iframes commonly used on websites that need to include external content like a Google map or a video from YouTube. Both of those popular websites use iframes in their embed code. The *iframe* element is supported by all modern desktop and mobile browsers. However, some browsers don't yet respond consistently to the three new HTML5 attributes for this element.

The *iframe* element, by itself, is not a security risk to you or your site visitors. Iframes have gotten a bad reputation because they can be used by malicious websites to include content that can infect a visitor's computer without them seeing it on the page, by incorporating links pointing to the invisible iframe, and those scripts set off malicious code.

Some computer viruses inject an invisible iframe into your web pages, effectively turning your website into a botnet.

Your site visitors are only as safe as the content of all the sites you link to. If you have reason to think a site is untrustworthy, don't link to it in any fashion.

This is a powerful capability in HTML — you can take any content from any website (with permission) and place it on your own site to enhance your content

This embedded page exists separately from the parent page in terms of its HTML, CSS, and JavaScript code.

The code snippet contains some extra attributes, width and height. These set the dimensions of the iframe region that displays the embedded file. width and height may be set as pixel values, or as percentages of the window, which scales the iframe proportionally to the size of the viewing window.

Like other HTML elements, iframes can be customized with other attributes.

7.2 SOURCE CODE

```
<!DOCTYPE html>
   <html>
   <head> <title>SALES DATA ANALYSIS</title> </head>
   <body>
   <style> body{ height:
   650px;
           background: linear-gradient(to bottom, #66ccff 0%, #cc33ff 98%);
        }
   h4{
         padding-left:300px;
         font-size: 25px;
         font-family: cursive;
         color: black;
   }
             h3{ padding-
left:800px; padding-
right:10px; font-size: 20px;
         font-family: serif; color:
         #ffe6f2;
           }
    p{
   padding-left:10px; padding-right:600px;
         font-size: 20px;
         font-family: serif; color:
         white;
```

```
}
.wrapper { margin:
  80px auto; text-
  align:
              center;
  width:
              100%;
  position: relative;
 }
 .button { padding: 15px
  100px; margin: 10px 4px;
  color: #eee; font-family:
  sans-serif; text-transform:
                  text-align:
  uppercase;
  center; position: relative;
  text-decoration:
                      none;
  display:
               inline-block;
  border: 1px solid;
 }
 .button::before {
  content:
  position:
  absolute; top: 0;
  left: 0;
  display: block;
  width: 100%;
```

height: 100%; z-

index: -1;

```
-webkit-transform: scaleY(.1);
     transform:
                        scaleY(.1);
     transition: all .4s
    }
    .button:hover { color:
    #b414ba;
    }
    .button:hover::before { opacity: 1; background-color: skyblue; -webkit-
     transform: scaleY(1); transform: scaleY(1); transition: -webkit-transform .6s
    cubic-bezier(.08, .35, .13, 1.02), opacity .4s; transition: transform .6s cubic-
    bezier(.08, .35, .13, 1.02), opacity
    }
    .glow
           { font-
    size:
              30px;
    color: #ffffff;
    -webkit-animation: glow 1s ease-in-out infinite alternate;
    -moz-animation: glow 1s ease-in-out infinite alternate; animation:
    glow 1s ease-in-out infinite alternate;
    }
    @-webkit-keyframes glow {
    from { text-shadow: 0 0 10px #eeeeee, 0 0 20px #000000, 0 0 30px #000000, 0 0 40px
#000000,
              0 0 50px #9554b3, 0 0 60px #9554b3, 0 0 70px #9554b3;
     }
          to
     {
            text-shadow: 0 0 20px #eeeeee, 0 0 30px #ff4da6, 0 0 40px #ff4da6, 0 0 50px
```

```
#ff4da6,
```

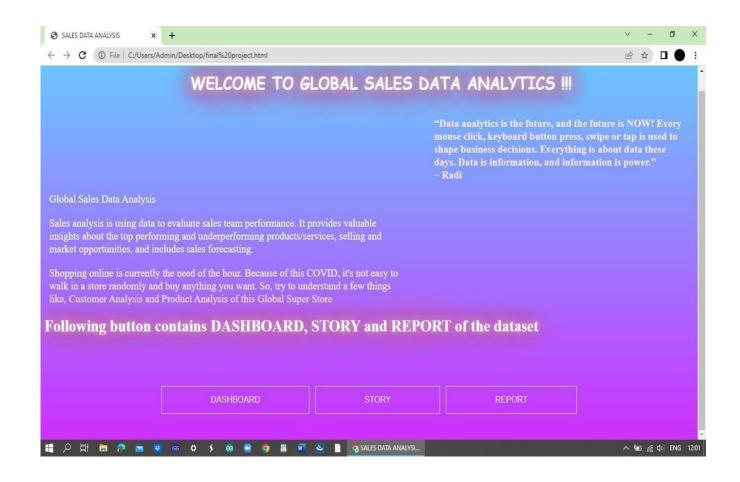
```
0 0 60px #ff4da6, 0 0 70px #ff4da6, 0 0 80px #ff4da6;
    }
   }
 </style>
<h4 class="glow"> WELCOME TO GLOBAL SALES DATA ANALYTICS !!! </h4>
<h3> "Data analytics is the future, and the future is NOW! Every mouse click, keyboard
button press, swipe or tap is used to shape business decisions. Everything is about data these
days. Data is information, and information is power."</br> ~ Radi</h3>
Global Sales Data Analysis
Sales analysis is using data to evaluate sales team performance. It provides valuable
insights about the top performing and underperforming products/services, selling and market
opportunities, and includes sales forecasting.
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
<h1 class="glow">Following button contains DASHBOARD, STORY and REPORT of the
dataset</h1>
<div class="wrapper">
                                                                     class="button"
                                                            <span
onclick="window.open('https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&path
Ref=.my_folders%2FProject%2FPROJECT%2BDASHBOARD&action=view&mode=das
hboard&subView=model000001847a4bd3cf_00000000','_blank')">DASHBOARD</span>
```

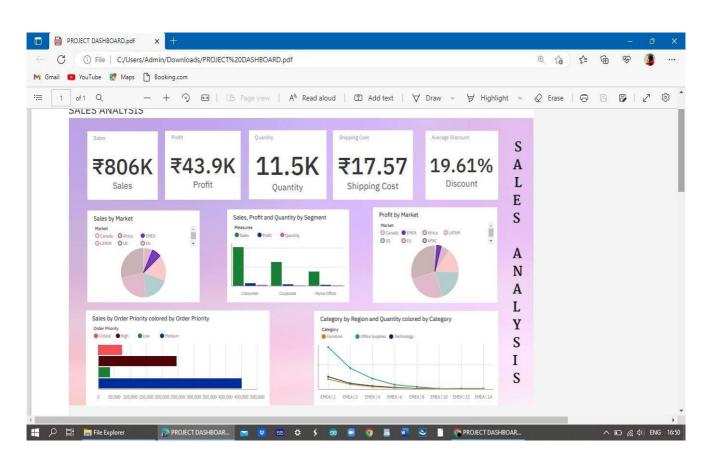
class="button"

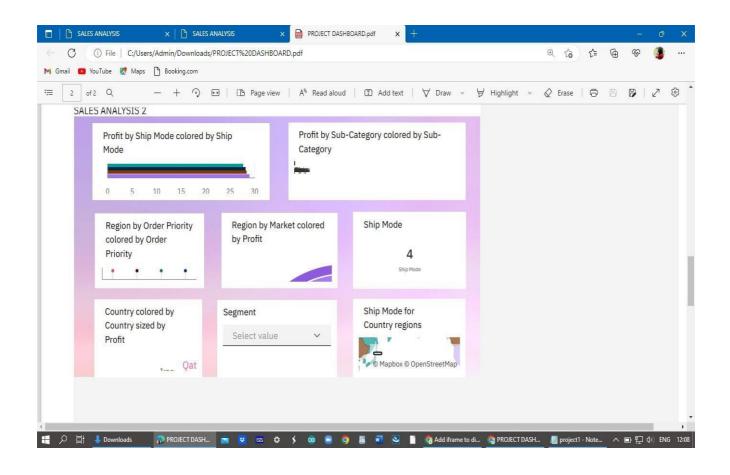
<span

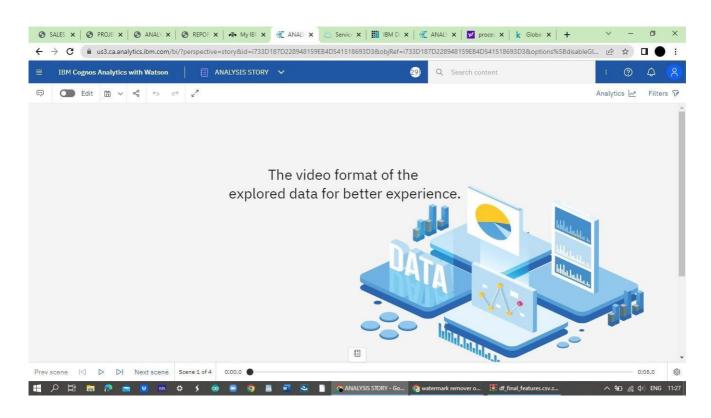
onclick="window.open('https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=. my_folders%2FProject%2FANALYSIS%2BSTORY&action=view&sceneId=model00000 1 847ab2d1f0_00000002&sceneTime=0','_blank')">STORY

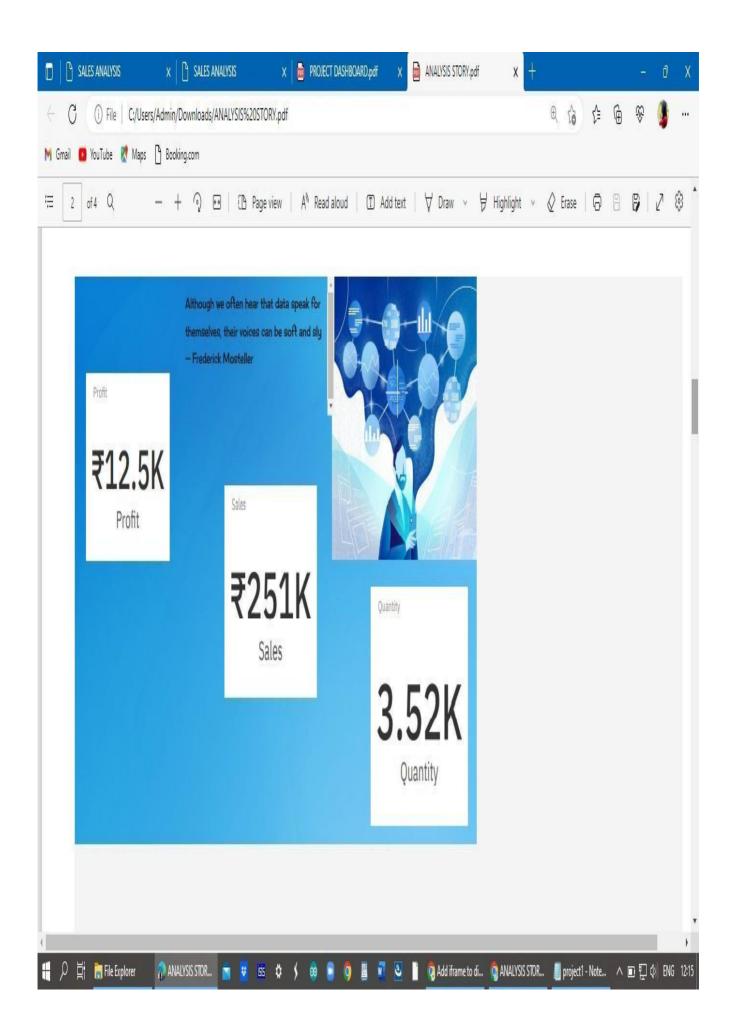
IMAGES OF THE WEBPAGE

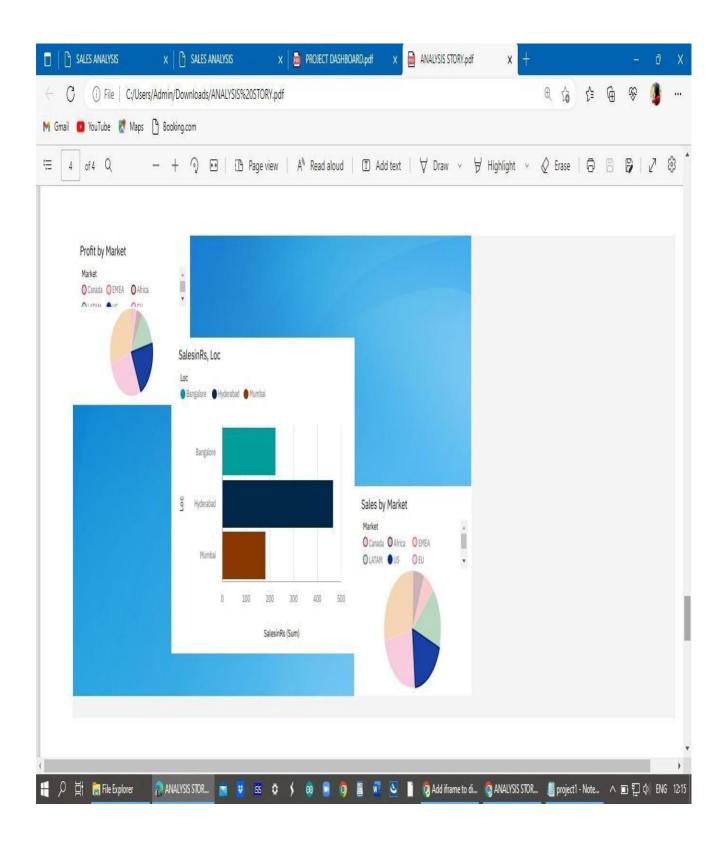


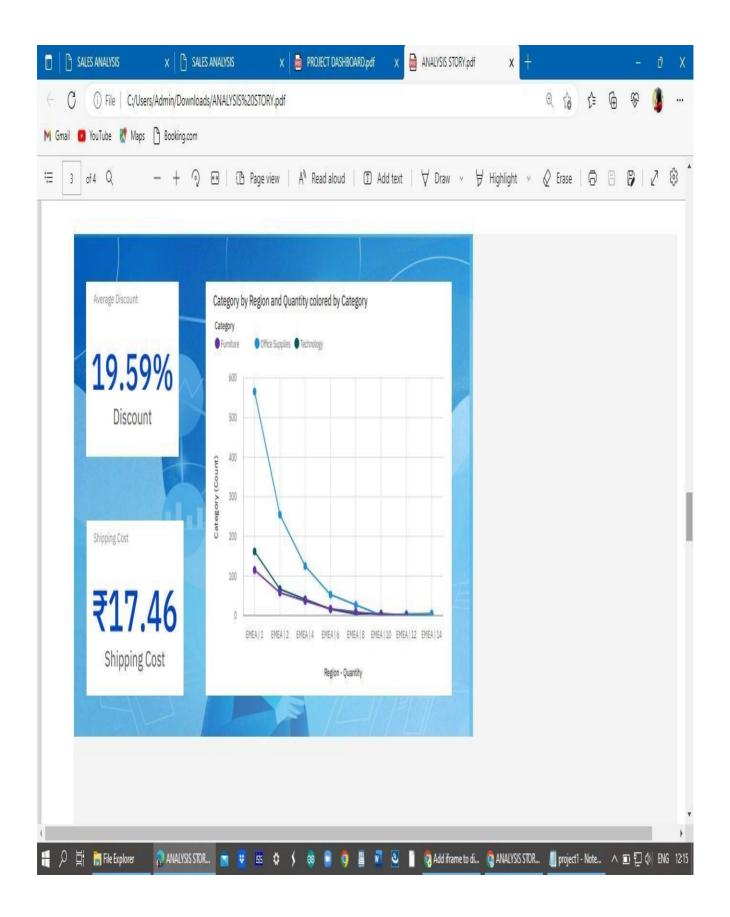


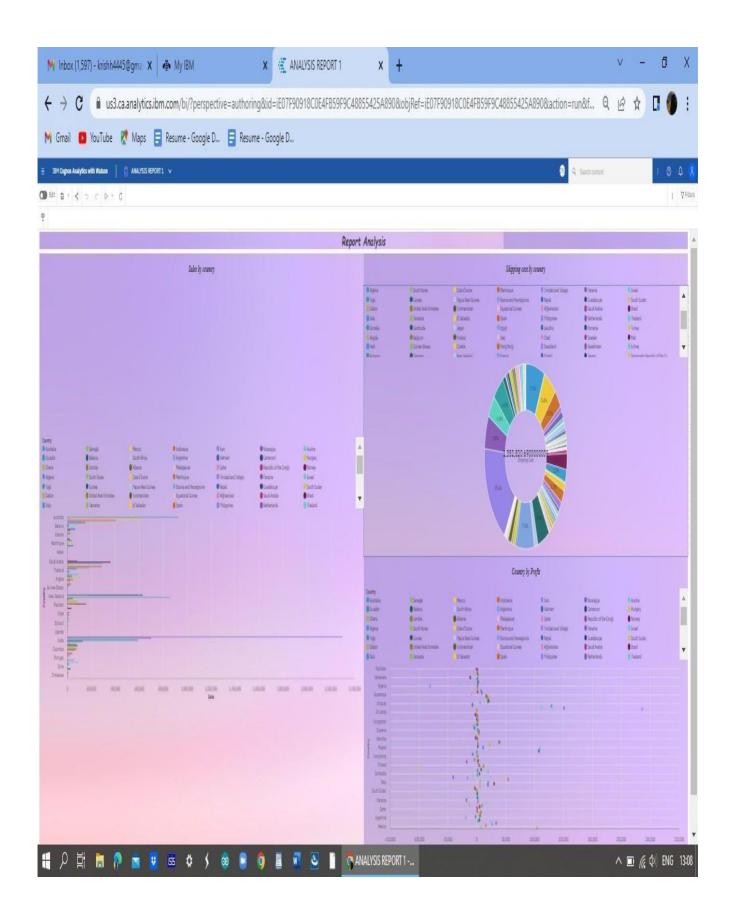


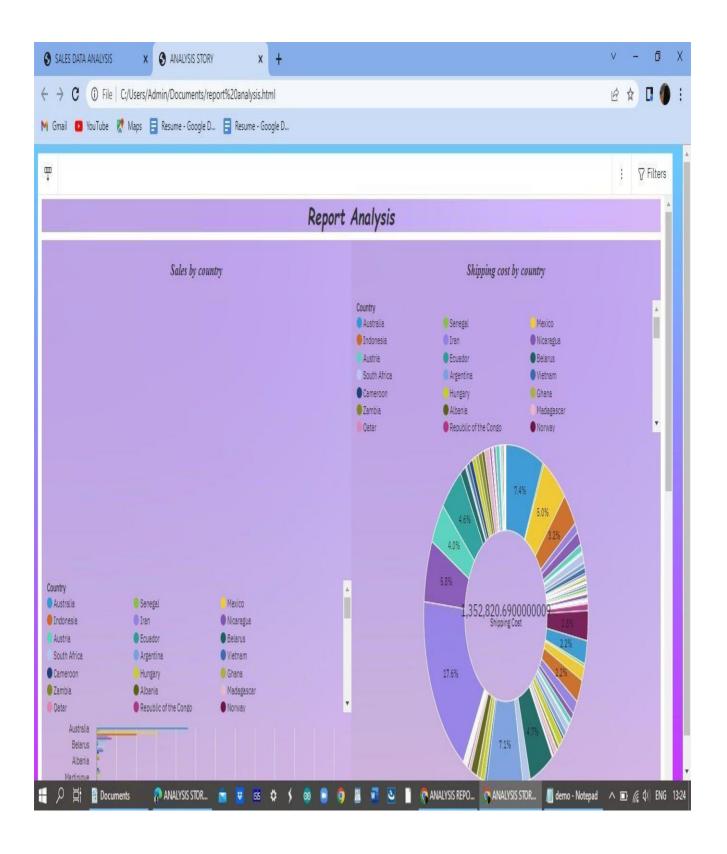












8. TESTING

8.1 USER ACCEPTANCE TESTING

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

Resolution	Severit y1	Severit y2	Severit y3	Severit y4	Subtot al
By Design	10	4	2	3	20
External	2	3	0	1	6
Fixed	11	2	4	20	37
Skipped	0	0	1	1	2
Won'tFix	0	0	0	1	1
Totals	23	9	7	26	66

8.2 TEST CASES

Section	Total Cases	Not Tested	Fail	Pass
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4

8.3 PERFOMANCE METRICES

Future sales prediction:

The prospect of future sales gives the companies that work with sales enormous relief. Those who sell have stock and must manage it intelligently. If there are too many items in stock, they risk having an insufficient room or having to sell at discounts for other items. Instead, when things are too little, the sales decline. Future sales can enable these issues to be avoided and better decisions to be taken.

Inventory Management:

The stock referred to the stocking of products and afterward used in crisis times. For enterprises to optimize resources and increase sales, inventory management is therefore vital. Retailers need to effectively manage inventories so that supply stays unimpacted, although sales suddenly rise. The supply networks and inventory chains are thoroughly analyzed to achieve this.

Merchandising:

The objective is to develop tactics to improve product sales and promotions. Customer decision-making using datasets will be influenced by merchandising. While appealing packaging and branding capture the attention of customers and increase their aesthetic look, rotary goods help to keep their products fresh and new.Data sets are gather to insights and create customer priority sets that take account of seasonality, relevance, and trends. **price prediction:**

The objective is to predict&set pricing, for both seller and buyer should satisfactory. price prdiction should be done using analysis of previous dataset.Price optimization affects the satisfaction ratings of clients directly.

Churn prevention:

Now when sales players have the skills to anticipate when a client will make the next purchase, it is necessary to predict when a consumer will quit buying. Customer churn refers to the percentage of customers who have stopped buying and using the product for a specific period. data sets are used to identify trends and features in the behavior, communication, and ordering of customers who have ceased shopping through customer relationship management information.

10.ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- > Monitor the progress of sales.
- > Contribute to decision-making.
- > Personalization of products and services.
- > Increase the efficiency of the work.

DISADVANTAGES:

- > Lack of alignment within different teams and department.
- > Lack of access to quality data.
- > Increased time and cost when existing data is not avilable.
- > Privacy concerns.

11.CONCLUSION:

From this project, we are able to predict and manage the daily basis sales and it can be access through our web application by both seller and buyer. While making this project, we gained a lot of experience of working as a team. We discovered various predicted and unpredicted problems and we enjoyed a lot solving them as a team. We adopted things like video tutorials, text tutorials, internet and learning materials to make our project complete.

12.FUTURE:

The project assists well to predict and manage the sales in general.

However, this project has some limitations:

- The application is unable to maintain the backup of data once it is uninstalled.
- This application does not provide higher decision capability. So making the machine learning prediction model

To further enhance the capability of this application, we recommend the following features to be incorporated into the system:

- Mobile apps advantage.
- Provide backup and recovery of data.
- Provide better user interface for user.

13.APPENDIX:

Source Code Github Link:

https://github.com/IBM-EPBL/IBM-Project-9086-1658975183

Project Demo Link:

https://drive.google.com/file/d/1hd3EW3Rme5WGswL33y6I5ZipmN92_U_g/view?usp=share_link