HX8001 - PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

GLOBAL SALES DATA ANALYTICS

A PROJECT REPORT

SUBMITTED BY

TEAM ID: PNT2022TMID50147

KISHOR SANTHOSH.P

MANAK.K

MANOJ KUMAR.M

MATHAN.M

in partial fulfillment for the award of degree of Bachelor of Engineering in

COMPUTER SCIENCE AND ENGINEERING

JAYARAJ ANNAPACKIAM CSI COLLEGE OF ENGINEERING NAZARETH- 628617.

ANNA UNIVERSITY: CHENNAI 600 025

NOV/DEC 2022

ABSTRACT

Data analytics is applied in various fields, including business companies performance forecasting, but struggle with its implementation. Following a cross sectional field study approach, we make two contributions. First, we elaborate on the central role played by the head controller in generating trust in analytics solutions and thus, making the project successful. Second, we identify three patterns in the way companies plan, implement, and then use data analytics in the context of business performance forecasting. The two successful patterns are the ones that start with a limited but tangible objective (either in term of information precision, or rapidity of processing) that can be expended in a second time.

TABLE OF CONTENT

CHAPTER	CONTENTS	PAGE NO
1	INTRODUCTION	
	1.1PROJECT OVERVIEW	4
	1.2 PURPOSE	
2	LITERATURE SURVEY	
	2.1 EXISTING PROBLEM	5
	2.2 REFERENCES	
	2.3 PROBLEM STATEMENT DEFINITION	
3	IDEATION & PROPOSED SOLUTION	
	3.1 EMPATHY MAP CANVAS	
	3.2 IDEATION AND BRAINSTORMING	10
	3.3 PROPOSED SOLUTION	
	3.4 PROBLEM SOLUTION FIT	
4	REQUIREMENT ANALYSIS	
	4.1 FUNCTIONAL REQUIREMENTS	16
	4.2 NON-FUNCTIONAL REQUIREMENTS\	
5	PROJECT DESIGN	
	5.1 DATA FLOW DIAGRAMS	17
	5.2 SOLUTION AND TECHNICAL ARCHITECTURE	
	5.3 USER STORIES	
6	6. PROJECT PLANNING AND SCHEDULING	
	6.1 SPRINT PLANNING AND ESTIMATION	21
	6.2 SPRINT DELIVERY SCHEDULE	
	6.3 REPORTS FROM JIRA	

7	CODING & SOLUTIONING	
	7.1 FEATURE 1	26
	7.2 FEATURE 2	
	7.3 DATABASE SCHEMA	
8	TESTING	
	8.1 TEST CASES	31
	8.2 USER ACCEPTANCE TESTING	
9	RESULTS	
	9.1 PERFORMANCE METRICS	32
10	ADVANTAGES & DISADVANTAGES	33
11	CONCLUSION	34
12	FUTURE SCOPE	36
13	APPENDIX	37

INTRODUCTION

Nowadays, many companies have started to exist and a few of them have grown to top positions. Some companies have very large amounts of data while some small companies have fewer amounts of data. This paper aims to research the company's growth in terms of the sales of the company's products. The first aim of the paper is to make a web application that analyses a company's sales data of varied products. This analysis is going to be effective if we use graphs and charts. This process is named Visualization. Manual Visualization of knowledge may be a time taking process. There is already some software that performs this task, but there are many disadvantages. Software like Jupyter Notebook can't be used on mobile phones whereas websites are often accessed using mobiles. Our paper takes company data of sales of products and data of a salesperson working in a particular company and depicts graphs between fields required.

LITERATURE SURVEY

1.Strategies for data analytics projects in business performance forecasting: a field study

Source: Journal of Management Control(2022) 33:241–271 https://doi.org/10.1007/s00187-022-00338-7

Abstract: Data analytics is applied in various fields, including business forecasting, companies performance but struggle with its implementation. Following a crosssectional field study approach, we make two contributions. First, we elaborate on the central role played by the head controller in generating trust in analytics solutions and thus, making the project successful. Second, we identify three patterns in the way companies plan, implement, and then use data analytics in the context of business performance forecasting. The two successful patterns are the ones that start with a limited but tangible objective (either in term of information precision, or rapidity of processing) that can be expended in a second time.

Keywords: Data analytics, Performance, Forecasting, Field study From a theoretical perspective, this paper contributes to the debate on the potential radical impact of data analytics on controlling functions. Based on exploratory techniques, the observations provided an initial reading grid for further research in the field. It also validated previously developed concepts for studying performance management and data analytics forecasting regarding information systems implementation. From a practical perspective, the contributions give

managers tools to prepare for the implementation of a digitalization project. They can, at least, serve as a tool for project managers to reflect on and, at best, as guiding principles for the overall project development and implementation. Such tools could also help analyze existing but imperfect projects, or even failed ones, or those about to fail, to improve the situation and get it back on track.

2. Market Analysis and Sales Development

Source:

https://www.academia.edu/6838349/A_Project_Report_On_MARKET_ANAL

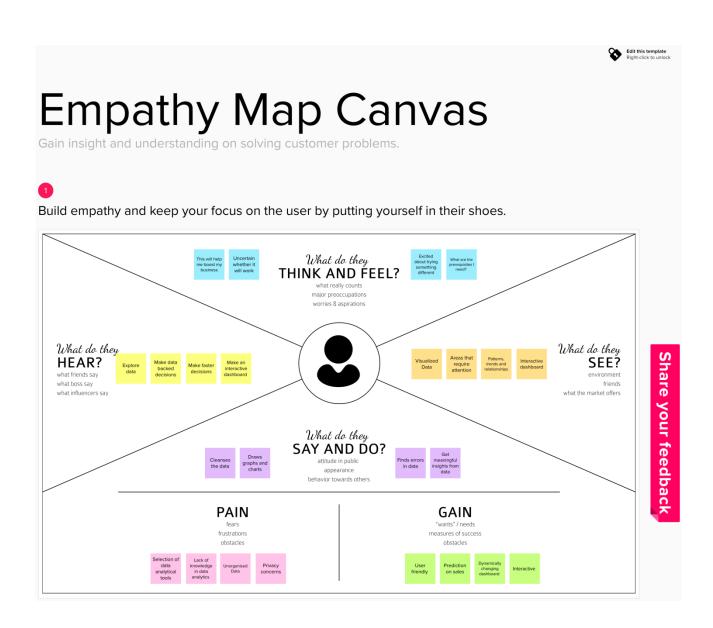
YSIS_AND_SALES_DEVELOPMENT_Submitted_By_Under_the_guidance Abstract: With the explosive growth of information sources available on the World Wide Web, it has become increasingly necessary for users to utilize automated tools to find the desired information resources, and to track and analyze their usage patterns. Association rule mining is an active data mining. research area. However, most ARM algorithms cater to a centralized environment. In contrast to previous ARM algorithms, Optimized Distributed Association Rule Mining (ODARM) is a distributed algorithm for geographically spread data sets that aimed to reduces operational/ communication costs. Recently, as the need to mine patterns across distributed databases has grown, Distributed Association Rule Mining (DARM) algorithms have been developed. These algorithms assume that the databases are either horizontally or vertically distributed. In the special case of databases populated from information extracted from textual data, existing D-ARM algorithms cannot discover rules based on higherorder associations between items in distributed textual documents

that are neither vertically nor horizontally distributed, but rather a hybrid of the two. Hence, this paper proposes a Distributed Count Association Rule Mining Algorithm (DCARM), which is experimented on real time datasets obtained from UCI machine learning repository. We are given a large database of customer transactions. Each transaction consists of items purchased by a customer in a visit. We present an efficient algorithm that generates all signicant association rules between items in the database. The algorithm incorporates buer management and novel estimation and pruning techniques. We also present results of applying this algorithm to sales data obtained from a large retailing company, which shows the effectiveness of the algorithm.

Keywords: Association rule mining, Optimized Distributed Association Rule Mining (ODARM), Distributed Count Association Rule Mining Algorithm (DCARM).

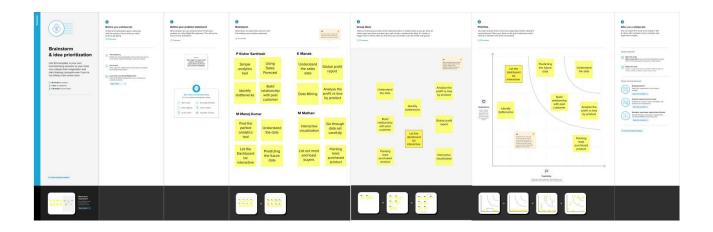
CHAPTER 3 IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



Link"https://app.mural.co/invitation/mural/mak7075/1664083900390?sender=u1e416277ea35 1baf3f797355&kev=08fa7aa5-c39a-4a2f-b6de-8f5b55e51ada"

3.2 IDEATION



Link"https://app.mural.co/invitation/mural/mak7075/16646044372
21?sender=u1e416277ea351baf3f797355&key=787c2995-856748f8-aa5f-ea3741b119f6"

3.3 PROPOSED SOLUTION

PROBLEM STATEMENT

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 conversion rate, 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 analyze 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.

IDEA/SOLUTION DESCRIPTION:

A good sales dashboard is the solution. It organizes your most recent sales-specific 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):

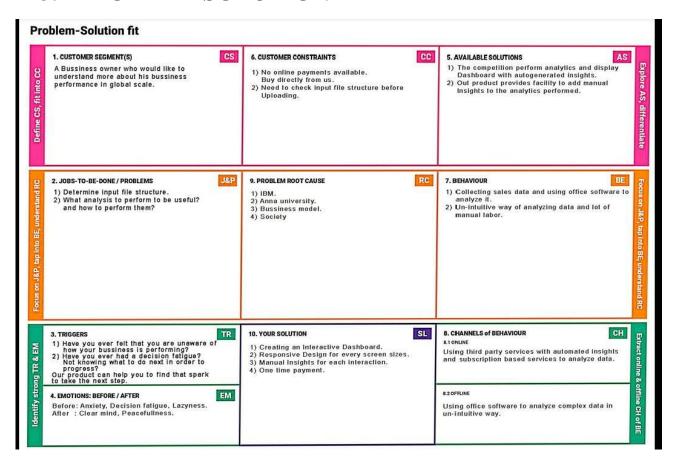
3wA Sales dashboard enables direct insight into your revenue driving force, allowing you to plan, implement and improve with data-based decisions.

SCALABILITY OF SOLUTION:

The great thing about Sales Analytics is that it gives you answers, and you 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



REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT:

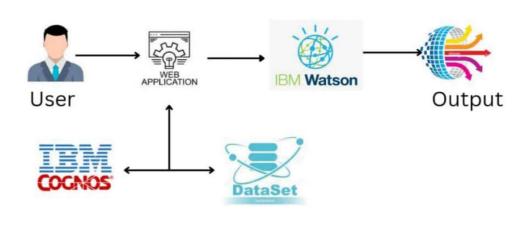
SI	Functional Requirement	Sub Requirement(Story/Sub-Task)
No.	(Epic)	
1	User Registration	Registration through Form
		Registration through Gmail
2	User Confirmation	Confirmation via Email
		Confirmation via OTP
	User Login	Login via Email and password
3		
4	User uploading	To store the dataset through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

4.2 NON FUNCTIONAL REQUIREMENTS:

SI	Non-Functional	Description
No.	Requirements	
		Optimized resources and it can be used
1	Usability	bye very one
2	Security	It has
		Securable because the send to end
		encryption
3	Reliability	It has high reliability base don-development.
		It has high state of performance and efficiency.
4	Performance	

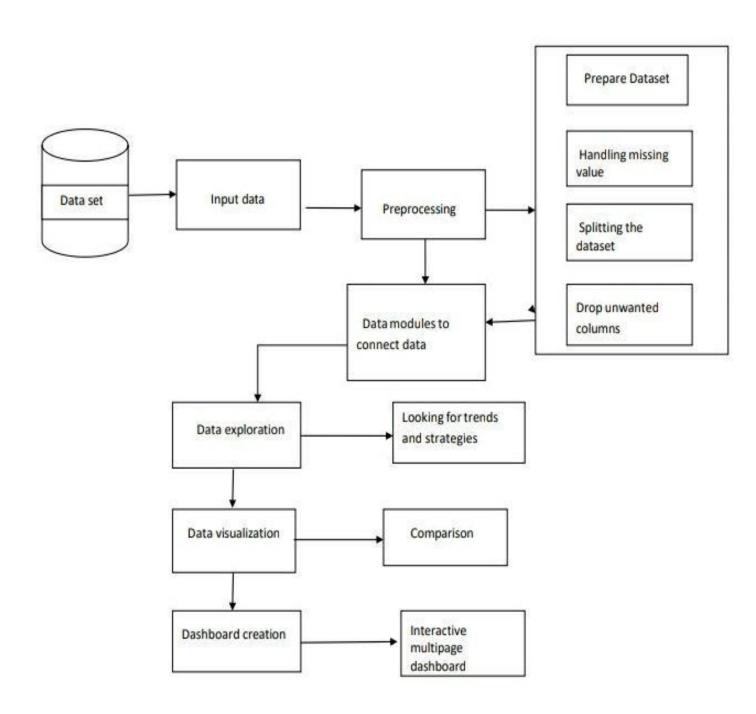
CHAPTER 5 PROJECT DESIGN

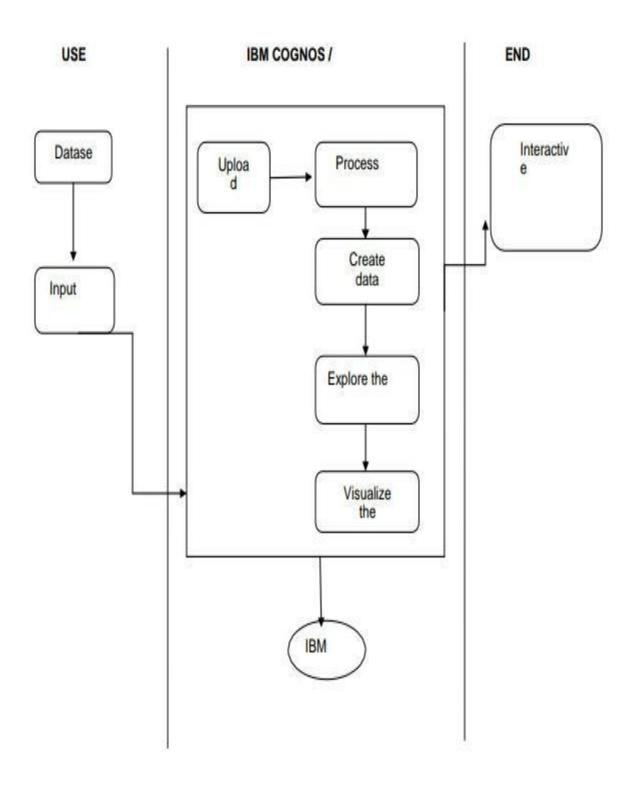
5.1 DATAFLOW DIAGRAM:



A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clearDFDcandepicttherightamountofthesystemrequirementgraphically.Itsh owshowdataentersandleavesthesystem, what changes the information, and whe redatais stored.

5.2 SOLUTION & TECHNICAL ARCHITECTURE





5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile <u>user)</u>	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a <u>user</u> , I can create the visualization by using the dashboard In the application		High	Sprint-3
Customer (Web user)	Login	USN-1	As a <u>user</u> , I can register for the application by entering my email ,password and confirming my password	I can access my account and dashboard	High	Sprint-1
Customer Care Executive	Chat box	USN-1	It can be used by easily access and responsible	I can access by easily through application	High	Sprint-2
Administrator	Mail	USN-3	It can be used by easily access and responsible	I can access by easily through application	High	Sprint-1

PROJECT PLANNING & SCHEDULING

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

Use the below template to create product backlog and sprint schedule

				Story Points	Priority	Team Members
Sprint-1	orint-1 Registration USN-1 As a user, I can register for the application by entering my email, password, and confirming my password.		5	High	Manoj Kumar.M Mathan.M	
Sprint-1	Login	USN-2	As a user, I will receive confirmation email once I have registered for the application, and I can log into the application by entering email & password	5	High	Manoj Kumar.M Mathan.M
Sprint-1	Data Collection	USN-3	As a user, I need to gather the data in the form of CSV/XLS files and clean the data to remove the null values	10	Low	Manoj Kumar.M Mathan.M
Sprint-2	Upload dataset	USN-4	As a user, I will upload the data to IBM Cognos and view the data of the products	5	Medium	Kishor Santhosh.P Manak.K
Sprint-2	Data Preparation	USN-5	As a user, I need to filter the data for visualization in IBM Cognos	5	High	Kishor Santhosh.P Manak.K
Sprint-2	Data visualization	USN-6	As a user, I can easily visualize the data in the form of charts and graphs through IBM Cognos	10	High	Kishor Santhosh.P Manak.K

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-3	Dashboard	USN-7	As a user, I will create the dashboards based on the given data in IBM Cognos	5	High	Kishor Santhosh.P Manak.K
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	5	Medium	Kishor Santhosh.P Manak.K
Sprint-3	Dashboard	USN-9	As a user, I must be able to gain insights from the charts/graphs through a variety of relationships established in the dashboard.	10	Medium	Kishor Santhosh.P Manak.K
Sprint-4	Prediction	USN-10	As a user, I will predict the specific product's future sales expectation.	5	High	Manoj Kumar.M Mathan.M
Sprint-4	Final Analysis	USN-11	As a user, I can Analyse the list of categorized products and their details as a report.	5	High	Manoj Kumar.M Mathan.M
Sprint-4	Report	USN-12	As a user, I can prepare the product and customer description information and more additional information as a Report	10	Medium	Manoj Kumar.M Mathan.M

6.2 SPRINT DELIVERY SCHEDULE:

TITLE	DESCRIPTION	DATE
Literature Survey	Literature survey on the	28SEPTEMBER2022
&Information-	selected project & gathering	
gathering	information by referring the	
	technical papers, research	
	publications etc.	
Prepare Empathy Map	Prepare Empathy Map Canvas	24SEPTEMBER2022
	to capture the user Pains	
	&Gains, Prepare list of	
	problem statements	
Ideation	List the by organizing	25SEPTEMBER2022
	thebrainstormingsessiona	
	ndprioritize the top 3	
	ideas based on the	
	feasibility &importance.	
Proposed Solution	Prepare the proposed solution	23SEPTEMBER2022
	document, which includes the	
	novelty, feasibility of idea,	
	business model, social impact,	
	scalability of solution, etc.	
	Prepare problem - solution	30SEPTEMBER2022
ProblemSolutionFit	fitdocument.	
	Preparesolutionarchitectured	28SEPTEMBER2022
SolutionArchitecture	ocument.	
SolutionArchitecture		

Customer Journey	Prepare the customer journey	200CTOBER2022
	maps to understand the user	
	interactions & experiences	
	with the application (entry to	
	exit).	
	Prepare the function	8OCTOBER2022
Functional Requirement	AL requirement	
	document.	
Dataflow Diagrams	Draw the data flow	9OCTOBER2022
Datanow Biagrams	diagrams and	30010BEN2022
	submit the for	
	review.	
	Teview.	
	Prepare the	100CTOBER2022
Technology Architecture	technology	
	architecture diagram.	
	Prepare the milestones	10 NOVEMBER 2022
	& activity list of the	
Prepare Milestone & Activity	project.	
List	project.	
	Develop & submit the	INPROGRESS.
Project Development -	developed code by testing	
DeliveryofSprint-1,2,3&4	it.	

6.3 REPORTS FROM JIRA

BURNDOWN CHART:

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

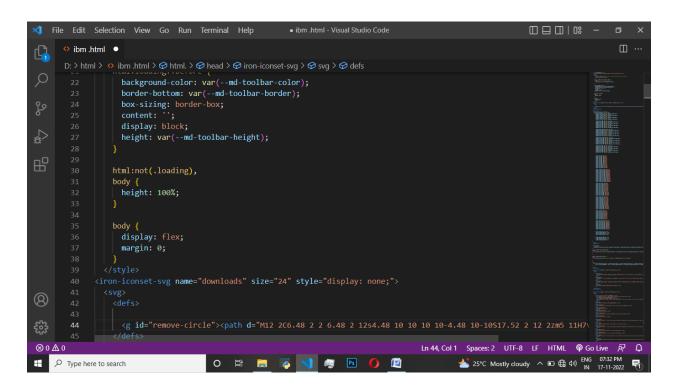
Sprint	Total	Duration	Sprint Start	Sprint	Story Points	Sprint
	Story		Date	End	Completed	Release
	Points			Date(Plan	(as on	Date(Actual)
				ned)	Planned End	
					Date)	
Sprint-1	20	6Days	24Oct2022	29Oct2022	20	29Oct2022
Sprint-2	20	6Days	31Oct2022	05Nov2022	20	05Nov2022
Sprint-3	20	6Days	07Nov2022	12Nov2022	20	12Nov2022
Sprint-4	20	6Days	14Nov2022	19Nov2022	20	19Nov 2022



CODING AND SOLUTION

7.1 FEATURE 1:

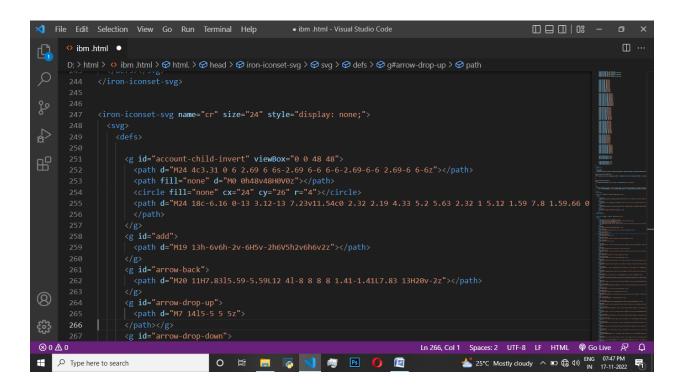
```
| File | Edit | Selection | View | Go | Run | Terminal | Help | Non-himminal visual Studio Code | Co
```

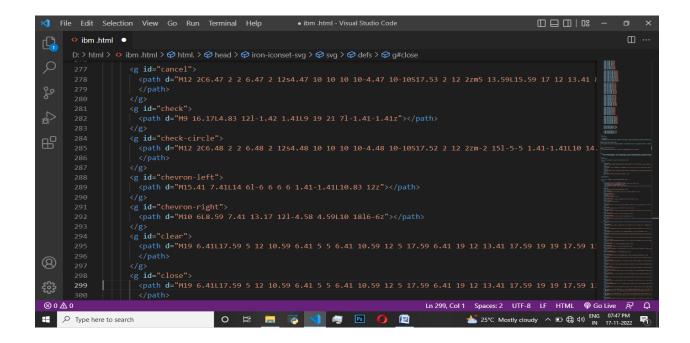


```
• ibm .html - Visual Studio Code
                                                                                                                D: > html > ♦ ibm .html > ♦ html. > ♦ head > ♦ custom-style > ♦ style > ♦ html
                 -google-red-100-rgb: 244, 199, 195;
                    --google-red-100: rgb(var(--google-red-100-rgb));
                    --google-red-300-rgb: 230, 124, 115;
                    --google-red-300: rgb(var(--google-red-300-rgb));
                    --google-red-500-rgb: 219, 68, 55;
                    --google-red-500: rgb(var(--google-red-500-rgb));
                    --google-red-700-rgb: 197, 57, 41;
                    --google-red-700: rgb(var(--google-red-700-rgb));
                    --google-blue-100-rgb: 198, 218, 252;
                    --google-blue-100: rgb(var(--google-blue-100-rgb));
                    --google-blue-300-rgb: 123, 170, 247;
                    --google-blue-300: rgb(var(--google-blue-300-rgb));
                    --google-blue-500-rgb: 66, 133, 244;
                    --google-blue-500: rgb(var(--google-blue-500-rgb));
--google-blue-700-rgb: 51, 103, 214;
                    --google-blue-700: rgb(var(--google-blue-700-rgb));
                    --google-green-100-rgb: 183, 225, 205;
                    --google-green-100: rgb(var(--google-green-100-rgb));
(A)
                    --google-green-300-rgb: 87, 187, 138;
                    --google-green-300: rgb(var(--google-green-300-rgb));
錢
                    --google-green-500-rgb: 15, 157, 88;
                      -google-green-500: rgb(var(--google-green-500-rgb));
                                                                                        Ln 73, Col 1 Spaces: 2 UTF-8 LF HTML @ Go Live 🔊 🚨
                                                                                                💒 25°C Mostly cloudy \land 🗈 🤀 예) ENG 07:32 PM
                                       O 🛱 🔚 🛜 刘 🙉 🖪
```

7.2 FEATURE 2:

```
• ibm .html - Visual Studio Code
                                                                                                                o ibm .html
       D: \land html \gt \diamondsuit ibm .html \gt \diamondsuit html. \gt \diamondsuit head \gt \diamondsuit iron-iconset-svg \gt \diamondsuit svg \gt \diamondsuit defs
                   <g id="block">
                     <path fill-rule="evenodd" clip-rule="evenodd" d="M10 0C4.48 0 0 4.48 0 10C0 15.52 4.48 20 10 20C15.52</pre>
                   <g id="domain">
                     <path d="M2,3 L2,17 L11.8267655,17 L13.7904799,17 L18,17 L18,7 L12,7 L12,3 L2,3 Z M8,13 L10,13 L10,15 |</pre>
                    <path fill-rule="evenodd" clip-rule="evenodd" d="M4.6327 8.00094L10.3199 2L16 8.00094L10.1848 16.8673C1</pre>
                   <g id="menu">
(2)
                     <path d="M2 4h16v2H2zM2 9h16v2H2zM2 14h16v2H2z"></path>
錢
⊗ 0 ▲ 0
                                                                                      Ln 242, Col 1 Spaces: 2 UTF-8 LF HTML @ Go Live № Q
                                                                                                O # 🔚 😼 刘 🧛
Ps ()
```





TESTING

Software testing is the process of evaluating and verifying that a software product or application does what it is supposed to do. The benefits of testing include preventing bugs, reducing development costs and improving performance. Test management plan. Types of software testing.

8.1 TESTCASE:

A test case includes information such as **test steps**, **expected results and data** while a test scenario only includes the functionality to be tested.

8.2 USER ACCEPTANCE TESTING:

User Acceptance Testing (UAT), which is performed on most UIT projects, sometimes called beta testing or end-user testing, is a phase of software development in which the software is tested in the "real world" by the intended audience or business representative.

CHAPTER 9 RESULT

Performance Metrics

This dashboard is created to understand a few things like, Customer Analysis and Product Analysis of the Global Super Store. This can be achieved by hearing out to the consumers and collecting their user preference data So that purchasing power will increase and beneficiary for both retailers and consumers.

ADVANTAGES:

- Lower marketing costs
- Consistency in brand image
- More revenue and more customer
- Optimum utilization of resources
- Growth and expansion opportunities

DISADVANTAGS:

- Differences in consumer needs, wants, usage patterns
- Lack of sales and marketing channel adaptation
- Chances of non-acceptance of product or services
- Non-specification of target markets
- Government restriction

CHAPTER 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

CHAPTER 12 FUTURE SCOPE

- Contract manufacturing/ outsourcing
- Integration of economics
- Joint venture and collaboration
- Technical and managerial knowledge
- Foreign agent and distributor

CHAPTER 13 APPENDIX

GitHub Link: "https://github.com/IBM-EPBL/IBM-Project-41459-1660642306"

Project Demo Link: "https://youtu.be/YSQuNdbw2mI"

Source code: "https://github.com/IBM-EPBL/IBM-Project-41459-

1660642306/tree/main/Final%20Deliverables/Source%20code "