TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 PROJECT OVERVIEW

The main objective of this project is to help the people by analyzing the data and help to make decision.Data is being generated very rapidly due to increase in information in everyday life. Huge amount of data get accumulated from various method that is difficult to analyze and exploit. Data created by considering and analyzing the profit and loss, sales percentage , low price, and quality of the product etc.., are example of huge data. Processing, analyzing and communicating this data are a challenge. Online shopping websites get flooded with voluminous amount of sales data every day. Analyzing and visualizing this data for information retrieval is a difficult task. Therefore a system is required which will effectively analyze and visualize data. Our project 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.

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, It is try to understand a few things like, Customer Analysis and Product Analysis of this Global Super Store and very helpful to people to make decision. Sales Analysis is the process of understanding how your business performs in terms of sales. It provides insights into the past, present, and future performance

of a business and can be used to help you forecast trends, identify opportunities for growth, and develop a strategic action plan for your company. The goal of sales analytics is always to simplify the information available to you. It should help you clearly understand your team's performance, sales trends, and opportunities.

CHAPTER 2

LITERATURE SURVEY

2.1 EXISTING PROBLEM

2.1.1 Dr.lakshmi,"Big Data Analytics in service industry",April 2016.

The report focuses on the growth prospects, restraints, and trends of the big data as a service market analysis. The study provides Porter's five forces analysis to understand the impact of various factors such as bargaining power of suppliers, competitive intensity of competitors, threat of new entrants, threat of substitutes, and bargaining power of buyers on the global market.

The growth of the big data as a service market is expected to continue during the Covid-19 pandemic outbreak. The operations in the IT industry are carried out normally by 'working from home' structure. Therefore the adoption of big data as a service in its application sectors such as BFSI, e-commerce, it & telecom, healthcare, and government are carried out normally.

On the other hand, the demand for big data as a service from sectors such as manufacturing and retail has been decreased due to the operational halt by government of all nations during the novel coronavirus outbreak.

2.2.2 Johannes Habel, Sascha Alavi,Nicholas Heinitz,"Predictive Sales analytics apoption",Dec 2021.

Due to the pervasive data ubiquity, sales practice is moving rapidly into an era of predictive analytics, using quantitative methods including machine learning algorithms to reveal unknown information such as customers' personality, value, or churn probabilities. However, many sales organizations face severe difficulties when implementing predictive analytics applications. This article elucidates these difficulties by developing the PSAA Model—a

conceptual framework that explains how predictive sales analytics applications support sales employees' job peformance. While the former is explained by well-established technology adoption theories, the extent to which adoption improves decision-making is determined by the value potential in the PSA application and the decision-making environment. Thereby, this paper provides a theoretical frame for future studies on predictive sales analytics.The collection of data and organizing the data make a huge problem in the visualization.

2.2.3 Manisha M.Patil,"Visual analysis of sales records",January 2021.

There is a need to gather various sales records that make it easy to measure monthly sales performance by product segment and by product category. The data gathered is stored with the help of an excel sheet which is a fast and reliable source to maintain sales records. BI analysts can use a wide variety of business intelligence BI tools such one is Tableau which gives them more flexibility in representing the results.

Tableau a visual analysis tool helps to measure sales targets from the whole dataset easily and gives the results in the form of desktop view very fast. These visual views can be used further to achieve the best sales targets in the future.The tools choosing for the analysis make an accuracy problem and also data is not in a standard format.

2.2 REFERENCES

- Morgan, B. (2019a). Descriptive Analytics, Prescriptive Analytics And Predictive Analytics For Customer Experience https://www.forbes.com/sites/blakemorgan/2019/02/21/descriptiveanalytics-prescriptive-analytics-and-predictive-analytics-for-customer experience/.
- Hyken, S. (2020). A PSA For Sales Leaders: Ditch Your Sales Scripts. https://www.forbes.com/sites/shephyken/2020/03/01/a-psa-for-sales-leaders-ditch-your-salesscripts/?sh=5b21ad8f6999.
- Gordini, N., & Veglio, V. (2017)Google Scholar. (2022). Predictive Sales Analytics. https://scholar.google.de/scholar?q=predictive+sales+analytics&hl=de&as_sdt=0%2C5&as_ ylo=&as_yhi=2022.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014)HBR. (2021). Embracing Data Analytics for More Strategic Valuehttps://hbr.org/resources/pdfs/comm/sisense2/EmbracingDataAnalytics.pdf.

2.3 PROBLEM STATEMENT DEFINITION

The main objective of this project is to help the people by analyzing the data and help to make decision.Data is being generated very rapidly due to increase in information in everyday life. Huge amount of data get accumulated from various method that is difficult to analyze and exploit. Data analyzing is one of the difficult task to finish. We have working with large amount of data which has been collected. processing and visualizing the data, No one expert

you to be a domain expert in everything, so don't let your lack of experience in an area dissuade you from trying to solve a problem in that space. That being said, if you are new to a domain area, make sure that you do your research and ideally connect with people actually solving problems.

For instance, if you want to solve a problem about homelessness in your city, but you don't know much about it and have never experienced homelessness yourself, you should look for some insider knowledge before defining your problem statement. You don't want to repeat work that's already been done or solve a "problem" that didn't actually need solving. Folks working at homeless shelters, organizing nonprofits, or otherwise trying to solve similar problems will have a wealth of knowledge and may be able to help make sure you're embarking on a useful, actionable mission.

CHAPTER 3

IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

An empathy map is a collaborative visualization used to articulate what we know about a particular type of user. It externalizes knowledge about users in order to,

1) create a shared understanding of user needs

2) Aid in decision making. An empathy map canvas serves as a foundation for outstanding user experiences, which focus on providing the experience customers want rather than forcing design teams to rely on guesswork.

Fig 3.1 Empathy Map Canvas

Empathy map canvases help identify exactly what it is that users are looking for so brands can deliver. They can be particularly beneficial for getting teams on the same page about who users are and what they want from the brand. An empathy map canvas is a more in-depth version of the original empathy map, which helps identify and describe the user’s needs and pain points. And this is valuable information for improving the user experience. Teams rely on user insights to map out what is important to their target audience, what influences them, and how they present themselves. This information is then used to create personas that help teams visualize users and empathize with them as individuals, rather than just as a vague marketing demographic or account number.

3.2 IDEATION & BRAINSTORMING

Ideation is not just a one time idea generation or a brainstorming session.In fact, we can divide ideation in these three stages: Generation, Selection and development. Brainstorming is one of the primary methods employed during the ideation stage of a typical design thinking process. Brainstorming combines a relaxed, informal approach to problem solving with lateral thinking. It encourages people to come up with thoughts and ideas that can, at first, seem a bit crazy. Some of these ideas can be crafted into original, creative solutions to a problem, while others can spark even more ideas. This helps to get people unstuck by "jolting" them out of their normal ways of thinking. Therefore, during brainstorming sessions, people should avoid criticizing or rewarding ideas. You're trying to open up possibilities and break down incorrect assumptions about the problem's limits. Judgment and analysis at this stage stunts idea generation and limit creativity. Brainstorming is a method of generating ideas and sharing knowledge to solve a particular

commercial or technical problem, in which participants are encouraged to think without interruption. Brainstorming is a group activity where each participant shares their ideas as soon as they come to mind. At the conclusion of the session, ideas are categorized and ranked for follow-on action. When planning a brainstorming session it is important to define clearly the topic to be addressed.

Fig 3.2 Ideation & Brainstorming

A topic which is too specific can constrict thinking, while an ill-defined topic will not generate enough directly applicable ideas. The composition of the brainstorming group is important too. It should include people linked directly with the subject as well as those who can contribute novel and unexpected ideas. It can comprise staff from inside or outside the organization. Ideation is not just a one time idea generation or a brainstorming session.In fact, we can

divide ideation in these three stages: Generation, Selection and development. Brainstorming is one of the primary methods employed during the ideation stage of a typical design thinking process.

3.3 PROPOSED SOLUTION

Your proposed solution section should offer your solution specifically, with enough detail so that your reader understands exactly what you're proposing. Indicate how your proposed solution will solve the problem and provide tangible benefits. Specifically, explain how it will meet the objectives and abide by the constrains outlined in the problem definition.

Your proposed solution should relate the current situation to a desired result and describe the benefits that will accrue when the desired result is achieved. So, begin your proposed solution by briefly describing this desired result.

Semantics of this new relationship is given by the design patterns essentials [18]. This implies that the hot- spot relationship is in fact a meta-relationship that is implemented through a design pattern that is generated taking into account the hot-spot flexibility requirements. The hot-spot cards guides this generation process,providing a systematic way for generating design patterns based on flexibility properties

• The language used to build and use the framework is a wide spectrum language, where it is not always easy

to express the user intentions;

• The complexity of framework class hierarchies and the difficulty of finding the points where the code should be written, that are the framework hot-spots or flexible points.

Project Design Phase-I

Proposed Solution

Date	23 September 2022
Team ID	PNT2022TMID42755
Project Name	Project - Global Sales Data Analytics
Maximum Marks	2 Marks

Proposed Solution:

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	We are analysing the nature of the customer, sales of different product, profit, loss, revenue and customer satisfaction.
2.	Idea / Solution description	When we are collecting previous good product with high sale, we can analyse the quality of the product. To get an increased revenue we can sell the product with less than 5 percent than actual rate
3.	Novelty / Uniqueness	Prepare a list of the all-potential differentiations of our brand and what you sell. compare your most unique angle against your audience's needs. think about Viable ways to apply it across your business.
4.	Social Impact / Customer Satisfaction	The best quality and satisfactory product can be obtained through the statistical report of the Global sales. The project will ensure cost and time effectiveness and customer satisfaction.
5.	Business Model (Revenue Model)	We can define business strategies of to ensure high profitable growth. We can gather and analyse data and provide key insights.
6.	Scalability of the Solution	We can improve scalability by monitoring internal process of sales, marketing effort, manufacturing and distribution. It maintenance the consistency of the product and quality level.

Fig 3.3 PROPOSED SOLUTION

3.4 PROBLEM SOLUTION FIT

The Problem-Solution Fit canvas is based on the principles of Lean Startup, LUM (Lazy User Model) and User Experience design. It helps entrepreneurs, marketers and corporate innovators identify behavioral patterns and recognize what would work and why. It is a template to help identify solutions with higher chances of solution adoption, reduce time spent on testing and get a better overview of the current situation. My goal was to create a tool that translates a problem into a solution, taking into account customer behavior and the context around it. None of the existing canvases or frameworks were giving me an overview and insight into the real customer situation during his/her decision-making process. With this template you will be able to take important information into consideration at an earlier stage and look at problem solving in depth. It increases your chances of finding problem-solution and product-market fit.

If you are trying to find a new solution to an existing problem, fill in this block after you get a better overview of the real situation.

When you are working on an existing solution (exploring growth strategies, problem with activation or solution adoption etc.), fill in this block first, and then see whether your solution is still relevant after all the blocks are filled in.

Try to spot patterns and repeated keywords by listing problems and behavior, related to it. For higher chances of solution adoption think of possible solutions that fit the user state limitations, take the best from alternative solutions, resemble natural triggers and tap into existing customer behavior. In short, you design a solution that is useful, understandable and accessible.

After giving it a first try you will likely realize that you were focusing on a different, less urgent / frequent problem, or you will recognize the real cause of

the situation. Extract repeated keywords (mediums, devices, situations), frequent behavior and emotions. That's your cheat sheet.

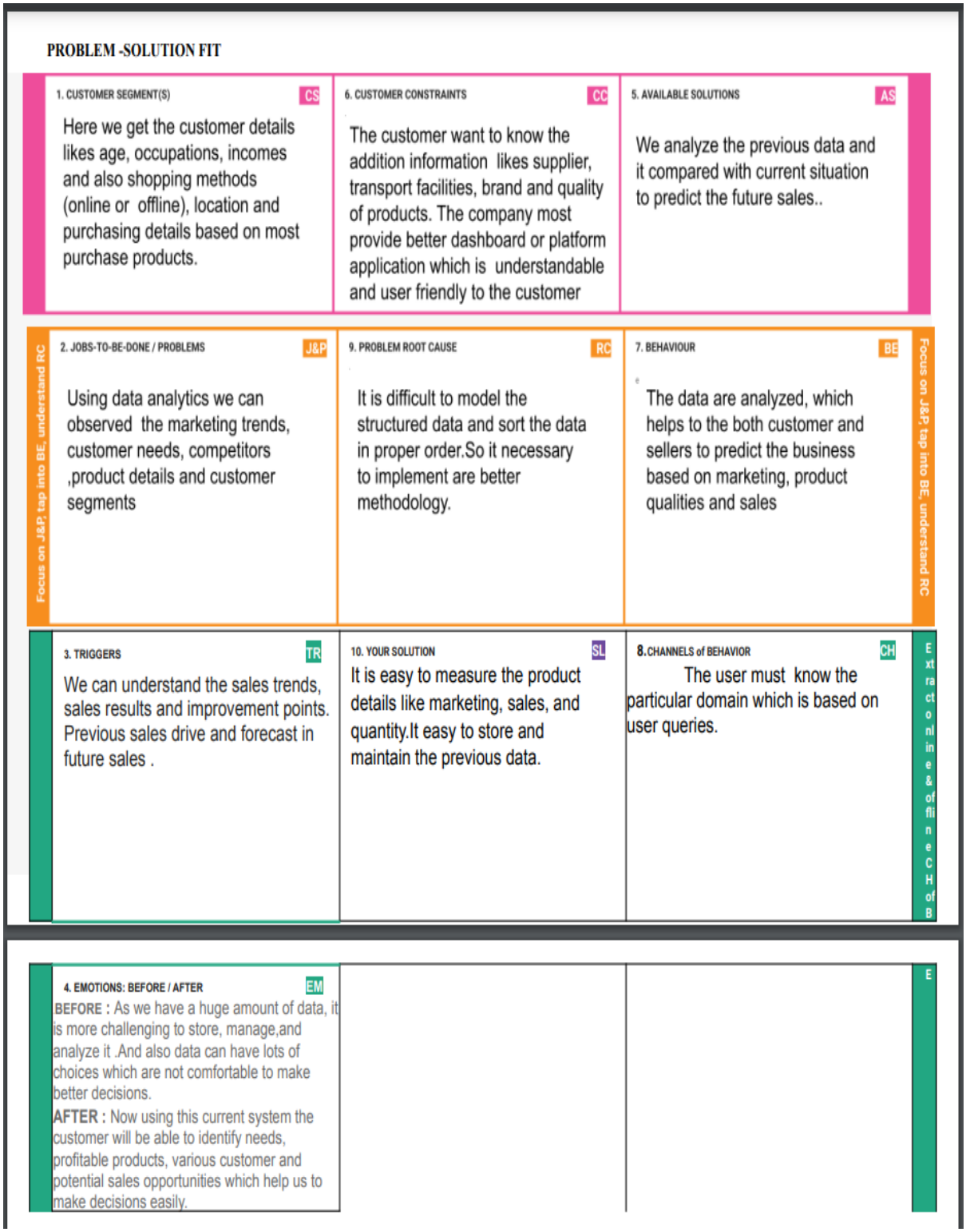


Fig 3.4 Problem Solution fit

CHAPTER 4

REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

Functional requirements are product features or functions that developers must implement to enable users to accomplish their tasks. So, it's important to make them clear both for the development team and the stakeholders. Generally, functional requirements describe system behavior under specific conditions. For example: The system sends an approval request after the user enters personal information. A search feature allows a user to hunt among various invoices if they want to credit an issued invoice. The system sends a confirmation email when a new user account is created. These are the requirements that the end user specifically demands as basic facilities that the system should offer.

All these functionalities need to be necessarily incorporated into the system as a part of the contract. These are represented or stated in the form of input to be given to the system, the operation performed and the output expected. In some cases a requirements analyst generates use cases after gathering and validating a set of functional requirements. The hierarchy of functional requirements collection and change, broadly speaking, is: user/stakeholder request → analyze → use case → incorporate. Stakeholders make a request. Functional requirements may involve calculations, technical details, data manipulation and processing, and other specific functionality that define what a system is supposed to accomplish. Behavioral requirements describe all the cases where the system uses the functional requirements, these are captured in use cases.

Project Design Phase-II

Solution Requirements (Functional & Non-functional)

Date	17 October 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global Sales Data Analysis
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	User Registration	Registration through Form Registration through Gmail
FR-2	User Confirmation	Confirmation via Email
FR-3	Account Creation	Create an account in the Profile Dashboard
FR-4	Input Credentials	Uploading your dataset Analyzing the product , sales and market rate using dashboard
FR-5	Processing Methods	Using IBM Cognos Analytics Dashboard Using Prediction algorithm to find sales and marketing data rate
FR-6	Output Credentials	Using the Dashboard and Algorithm they know about the marketing level and way to predict the future sales and marketing levels.

Fig 4.1 FUNCTIONAL REQUIREMENTS

4.2 NON-FUNCTIONAL REQUIREMENTS

These are basically the quality constraints that the system must satisfy according to the project contract. The priority or extent to which these factors are implemented varies from one project to other. They are also called non-behavioral requirements. They basically deal with issues like: Portability ,Security Maintainability, Reliability ,Scalability.

Project Design Phase-II

Solution Requirements (Functional & Non-functional)

Date	17 October 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global Sales Data Analysis
Maximum Marks	4 Marks

Non-functional Requirements:

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The user can be able to interact with the system user friendly. The system is built with simple modules and algorithms.
NFR-2	**Security**	The system administrator can only permit to access and modify the data in the system
NFR-3	**Reliability**	The database update process must roll back all related updates when any update fails. The dataset will not be modified by anyone only the users can be able to modify the dataset.
NFR-4	**Performance**	The system responds to certain user's actions under a certain workload. A user waits before the target operation happens given the overall number of users at the moment.
NFR-5	**Availability**	If the application can contain the histrological sales and marketing data and the system is accessible to a user at a time. While it can be expressed as an expected percentage of successful requests, you may also define it as a percentage of time the system is accessible for operation during some time period.
NFR-6	**Scalability**	The more users are able to use the dashboard at the same time and the model can successfully predict the future sales.

Fig 4.2 NON-FUNCTIONAL REQUIREMENT

CHAPTER 5

PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS

Data flow diagrams (DFDs) visually map your process or system, so you can uncover opportunities to improve efficiency and performance. Whether you are improving an existing process or implementing a new one, a data flow diagram will make the task easier. However, if you've never created a DFD before, getting started can be intimidating. There is a lot to take in: different levels of diagrams, symbols and notation, not to mention actually creating the diagram—navigating it all will take more than looking at a few examples. If you're new to data flow diagrams, this guide will help get you started.

Project Design Phase-II
Data Flow Diagram &User Stories

Date	16 October 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global sales Data Analytics
Maximum Marks	4 Marks

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.

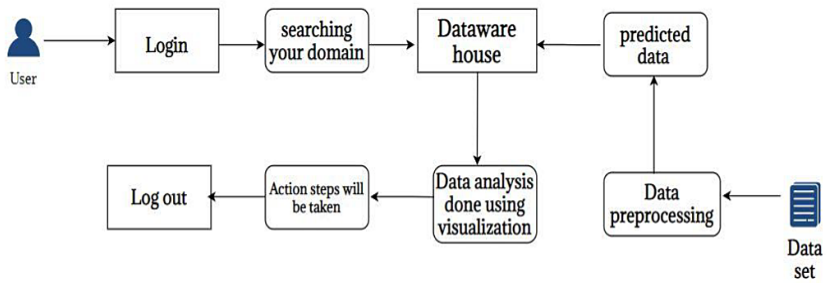


Fig 5.1 DATA FLOW DIAGRAM

5.2 SOLUTION & TECHNICAL ARCHITECTURE

Technical architecture which is also often referred to as application architecture, IT architecture, business architecture, etc.—refers to creating a structured software solution that will meet the business needs and expectations while providing a strong technical plan for the growth of the software application through its lifetime. IT architecture is equally important to the business team and the information technology team. Technical architecture includes the major components of the system, their relationships, and the contracts that define the interactions between the components. The goal of technical architects is to achieve all the business needs with an application that is optimized for both performance and security.IT architects plan for things they know are coming in the future and for things they don't yet envision or dream. Taking the time to design the architecture at the start will prevent major design changes, code refactoring, and expensive rework later in the project.

Solution architecture is the practice of designing, describing, and managing solution engineering to match it with specific business problems. For example, protecting customer data under GDPR and other privacy regulations is a business-level problem. Solution architecture defines how those requirements would translate into the way a given software operates.

A solution architect is in charge of leading the practice and introducing the overall technical vision for a particular solution.

While the practice can be managed in-house, there are companies that provide solution architecture consulting as a specific set of services.

Each of these terms includes multiple aspects that we'll discuss in the article. If you want to catch all details quickly.

5.2.1 SOLUTION ARCHITECTURE

Project Design Phase-I
Solution Architecture

Date	14 October 2022
Team ID	PNT2022TMID42755
Project Name	Global Sales Data Analytics
Maximum Marks	4 Marks

Solution 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.

Example - Solution Architecture Diagram:

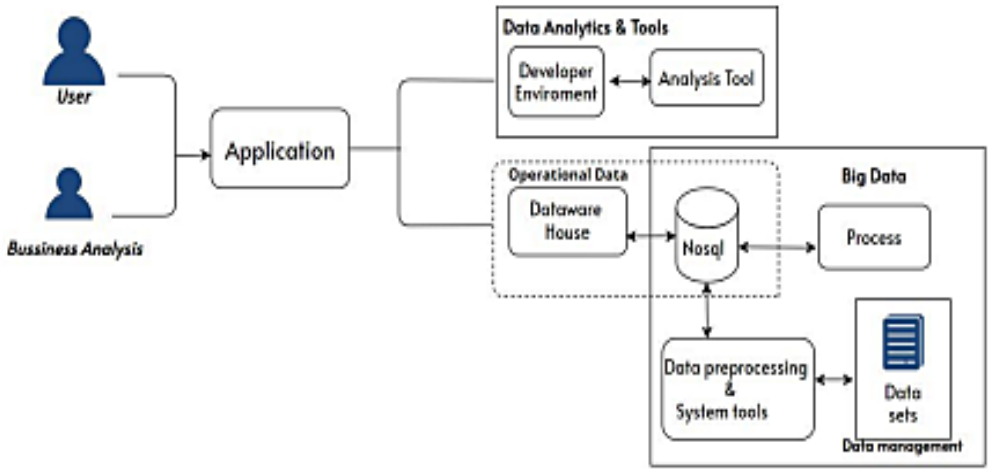


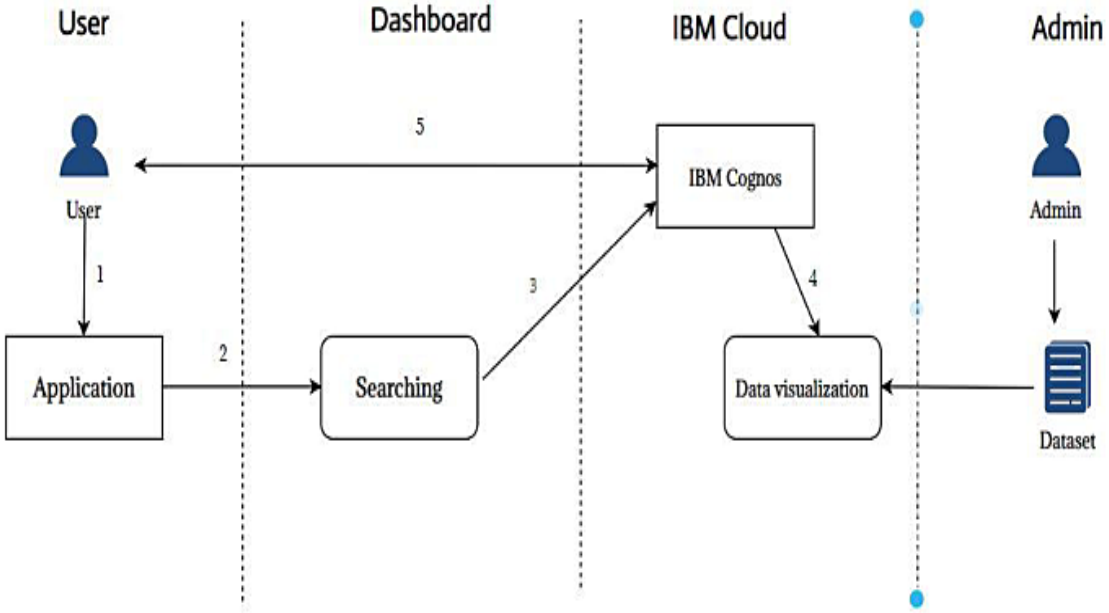
Fig 5.2 SOLUTION ARCHITECTURE

5.2.2 TECHNICAL ARCHITECTURE

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

Date	17 October 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global Sales Data Analysis
Maximum Marks	4 Marks

Technical Architecture:



Guidelines:

- Include all the processes (As an application logic / Technology Block)
- Provide infrastructural demarcation (Local / Cloud)
- Indicate external interfaces (third party API's etc.)
- Indicate Data Storage components / services
- Indicate interface to machine learning models (if applicable)

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How the user interacts with the application e.g. Web UI, Mobile App .	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4.	Application Logic-3	Logic for a process in the application	IBM Watson Assistant
5.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
6.	cognos Database	Data visualization on cognos	IBM cognos.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	Machine Learning Model	Purpose of Machine Learning Model	To analysis the sales and marketing data level
9.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	The better platform or dashboard for data visualization and predication	Python and cognos
2.	Security Implementations	The system administrator can only permit to access and modify the data in the system	Encryption and information verification
3.	Scalable Architecture	The more users are able to use the dashboard at the same time and the model can successfully predict the future sales.	Front end languages , database and backend process
4.	Availability	If the application can contain the histrological sales and marketing data and the system is accessible to a user at a time. While it can be expressed as an expected percentage of successful requests, you may also define it as a percentage of time the system is accessible for operation during some time period.	Encryption,machine learning Algorithm and the natural language process.
5.	Performance	The performance of the dashboard is flexible to every user. The front-page load time must be no more than 2 seconds for users that access the websites.	Content delivery network

Fig 5.3 TECHNICAL ARCHITECTURE

5.3 USER STORIES

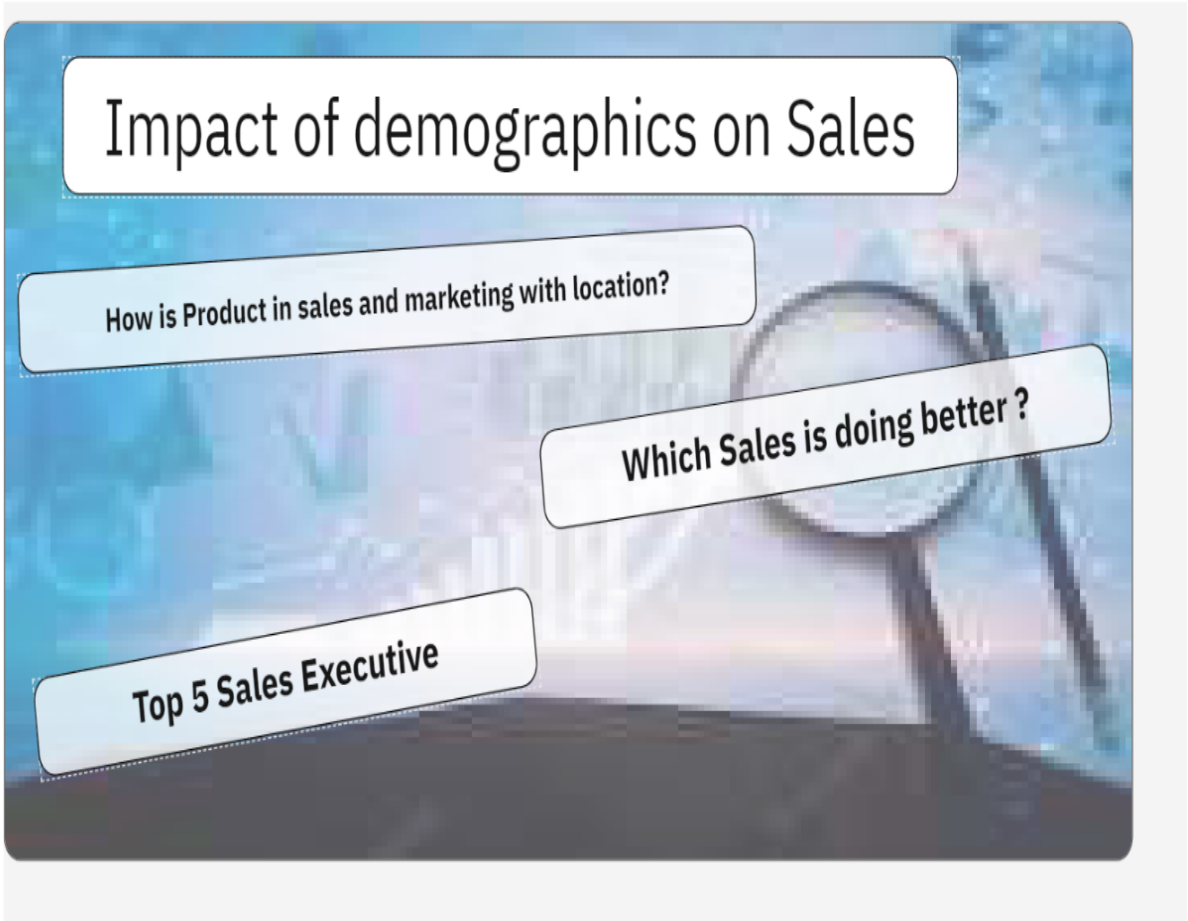
A user story is the smallest unit of work in an agile framework. It’s an end goal, not a feature, expressed from the software user’s perspective.

A user story is an informal, general explanation of a software feature written from the perspective of the end user or customer.

The purpose of a user story is to articulate how a piece of work will deliver a particular value back to the customer. Note that "customers" don't have to be external end users in the traditional sense, they can also be internal customers or colleagues within your organization who depend on your team.

User stories are a few sentences in simple language that outline the desired outcome. They don't go into detail. Requirements are added later, once agreed upon by the team.

11/15/22, 4:17 PM * Sales Analytics Story



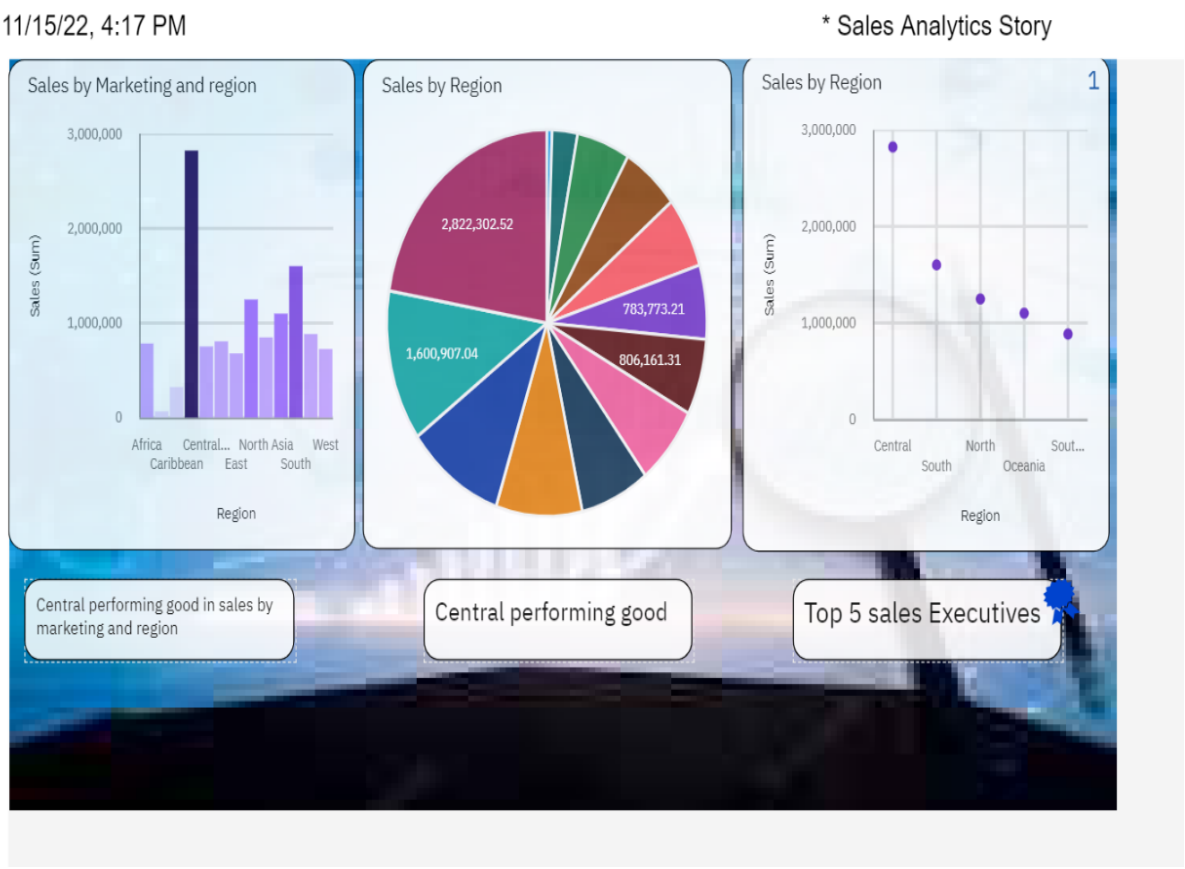


Fig 5.4 USER STORY

CHAPTER 6

PROJECT PLANNING & SCHEDULING

6.1 SPRINT PLANNING & ESTIMATION

Sprint planning is an event in scrum that kicks off the sprint. The purpose of sprint planning is to define what can be delivered in the sprint and how that work will be achieved. Sprint planning is done in collaboration with the whole scrum team. In scrum, the sprint is a set period of time where all the work is done. However, before you can leap into action you have to set up the sprint. You need to decide on how long the time box is going to be, the sprint goal, and where you're going to start.

The sprint planning session kicks off the sprint by setting the agenda and focus. If done correctly, it also creates an environment where the team is motivated, challenged, and can be successful. Bad sprint plans can derail the team by setting unrealistic expectations.As described in the Scrum Guide, Sprint Planning initiates the Sprint by laying out the work to be performed for the Sprint. This resulting plan is created by the collaborative work of the entire Scrum Team. The product owner ensures that attendees are prepared to discuss the most important product backlog items and how they map to the Product Goal. The Scrum Team may also invite other people to attend Sprint Planning to provide advice. Estimation is done by the entire team during Sprint Planning Meeting. The objective of the Estimation would be to consider the User Stories for the Sprint by Priority and by the Ability of the team to deliver during the Time Box of the Sprint.

Project Planning Phase

Milestone and Activity List

Date	29 October 2022
Team ID	PNT2022TMID42755
Project Name	Global Sales Data Analysis

Title	Description	Date
Literature Survey and Information Gathering	Gathering Information by referring the technical papers, research publications etc	1 SEPTEMBER 2022
Prepare Empathy Map	To capture user pain and gains Prepare List of Problem Statement	11 SEPTEMBER 2022
Ideat on	Prioritize a top 3 ideas based on feasibility and Importance	18 SEPTEMBER 2022
Proposed Solution	Solution include novelty,feasibility,business model,social impact and scalability of solution	24 SEPTEMBER 2022
Problem Solution Fit	Solution fit document	1 October 2022
Solution Architecture	Solution Architecture	1 October 2022
Customer Journey	To Understand User Interactions and experiences with application	9 October 2022
Functional Requirement	Prepare functional Requirement	15 October 2022
Data flow Diagrams	Data flow diagram	15 October 2022
Technology Architecture	Technology Architecture diagram	16 October 2022
Project Development-Delivery of sprint 1,2,3 &4	Develop and submit the developed code by testing it	24 October 2022 – 19 November 2022

Fig 6.1 SPRINT PLANNING

6.2 SPRINT DELIVERY SCHEDULE

Sprints also produce different deliverables for different audiences – the team, your organization at large, the public – it really depends what you want to show people to help them understand your solution.

The deliverables of a sprint aren't as predictable as they are for other projects. Sprint participants have produced sketches and drawings, writing, photographs, comic strips, videos and fully coded working prototypes. The answer is whatever's right to answer the problem.

Project Planning Phase

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

Date	29 October 2022
Team ID	PNT2022TMID42755
Project Name	Global Sales Data Analytics
Maximum Marks	8 Marks

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

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	1	Customers can register by entering the basic personal details through website	2	High	Mukesh V
	Login	2	As an authenticated user using their login credentials user can view the entire website and various options	2	High	Rajkumar A
	Working with the Dataset	3	Initially Data Preprocessing like filtering, formatting and data cleansing have to be done.	2	High	Naveen P
		4	Load the dataset in the cloud platform and analyse the data points by Visualization techniques.	10	High	Keerthivasan s

Fig 6.2 SPRINT DELIVERY

6.3 REPORTS FROM JIRA

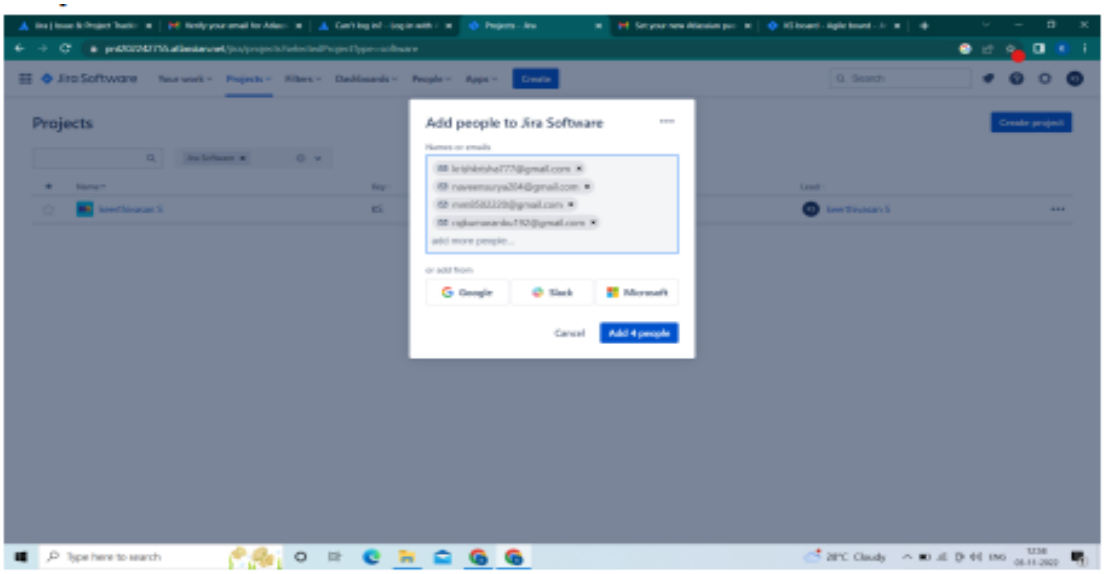


Fig 6.3 JIRA REPORT

CHAPTER 7

CODING & SOLUTIONING

7.1 FEATURE 1

7.1.1 ANACONDA NAVIGATOR

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda Distribution that allows you to launch applications and manage conda packages, environments, and channels without using command line interface (CLI) commands.

In order to run, many scientific packages depend on specific versions of other packages. Data scientists often use multiple versions of many packages and use multiple environments to separate these different versions.The CLI program conda is both a package manager and an environment manager. This helps data scientists ensure that each version of each package has all the dependencies it requires and works correctly.

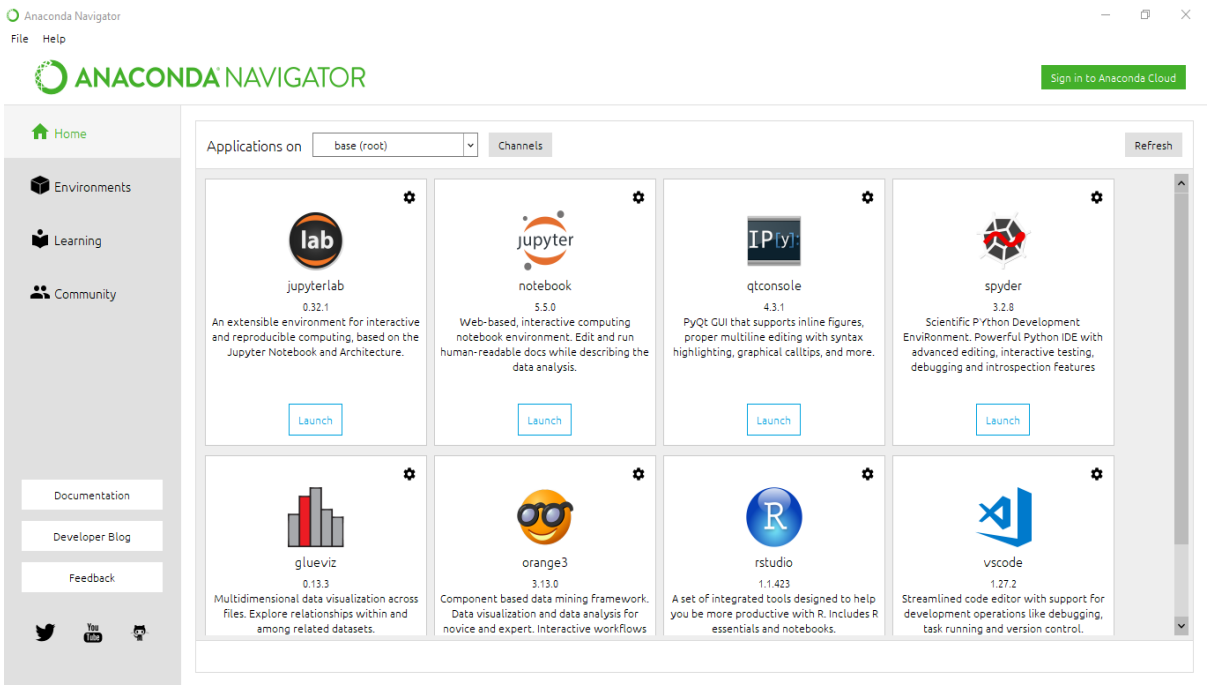


Fig 7.1 ANACONDA NAVIGATOR

7.1.2 JUPYTER NOTEBOOK

Jupyter notebooks basically provides an interactive computational environment for developing Python based Data Science applications. They are formerly known as python notebooks. The following are some of the features of Jupyter notebooks that makes it one of the best components of Python ML ecosystem −

- Jupyter notebooks can illustrate the analysis process step by step by arranging the stuff like code, images, text, output etc. in a step by step manner.
- It helps a data scientist to document the thought process while developing the analysis process.
- One can also capture the result as the part of the notebook.
- With the help of jupyter notebooks, we can share our work with a peer also.

7.2 FEATURES 2

COLLAB

Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs.

Code is executed in a virtual machine private to your account. Virtual machines are deleted when idle for a while, and have a maximum lifetime enforced by the Colab service.

Colab are prioritized for interactive use cases. We prohibit actions

associated with bulk compute, actions that negatively impact others, as well as actions associated with bypassing our policies. The following are disallowed from Colab runtime:

- file hosting, media serving, or other web service offerings not related to interactive compute with Colab
- downloading torrents or engaging in peer-to-peer file-sharing
- using a remote desktop or SSH
- connecting to remote proxies
- mining cryptocurrency
- running denial-of-service attacks
- password cracking
- using multiple accounts to work around access or resource usage restrictions
- creating deepfakes

CHAPTER 8

TESTING

8.1 TEST CASES

A test case is a set of actions performed on a system to determine if it satisfies software requirements and functions correctly. The purpose of a test case is to determine if different features within a system are performing as expected and to confirm that the system satisfies all related standards, guidelines and customer requirements. The process of writing a test case can also help reveal errors or defects within the system.

A test case document includes test steps, test data, preconditions and the post conditions that verify requirement

Project Development Phase
Model Performance Test

Date	16 November 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global Sales Data analytics
Maximum Marks	10 Marks

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	15 visualizations
2.	Data Responsiveness	Good
3.	Amount Data to Rendered (DB2 Metrics)	1
4.	Utilization of Data Filters	Yes
5.	Effective User Story	3 Stories
6.	Descriptive Reports	3 Report

Fig 8.1 TEST CASES

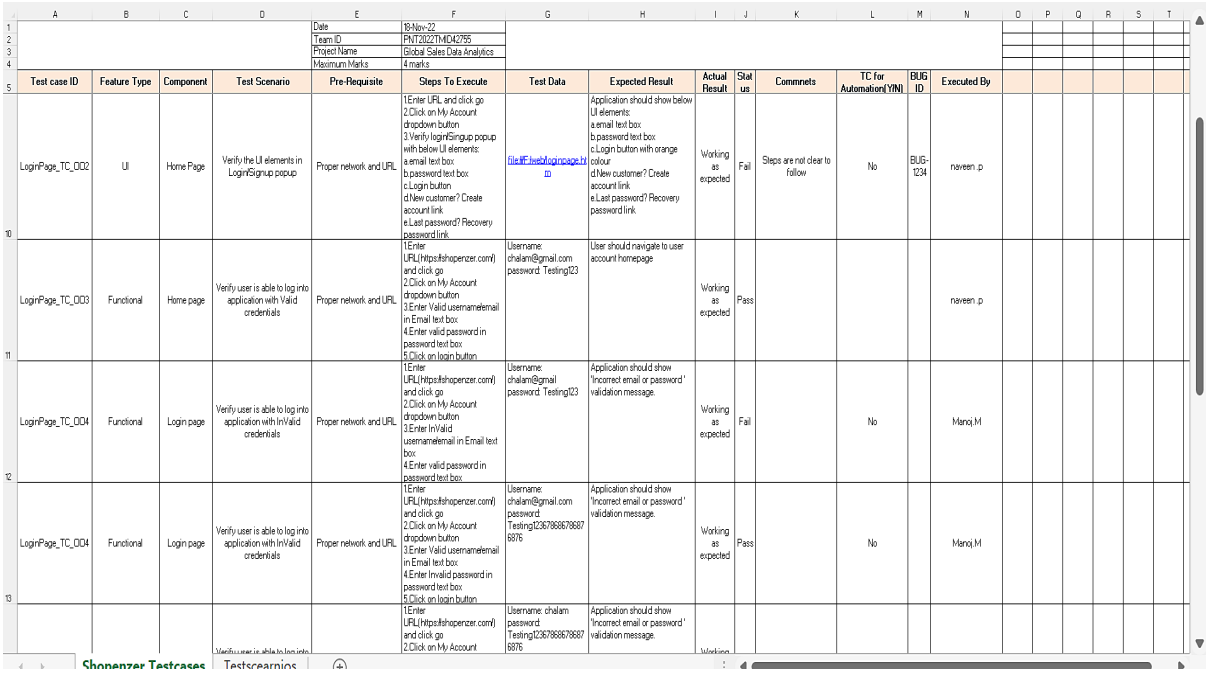
8.2 USER ACCEPTANCE TRAINING

User Acceptance Testing (UAT) is a type of testing performed by the end user or the client to verify/accept the software system before moving the software application to the production environment. UAT is done in the final phase of testing after functional, integration and system testing is done.

The main Purpose of UAT is to validate end to end business flow. It does not focus on cosmetic errors, spelling mistakes or system testing. User Acceptance Testing is carried out in a separate testing environment with production-like data setup. It is kind of black box testing where two or more end-users will be involved.

UAT is performed by –

- Client
- End users



				Date	18-Nov-22								
				Team ID	PNT2022TMID42755								
				Project Name	Global Sales Data Analytics								
				Maximum Marks	4 marks								
Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Commnets	TC for Automation(Y/N)	BUG ID	Executed By
LoginPage_TC_002	UI	Home Page	Verify the UI elements in Login/Signup popup	Proper network and URL	1.Enter URL and click go 2.Click on My Account dropdown button 3.Verify login/Signup popup with below UI elements: a.email text box b.password text box c.Login button d.New customer? Create account link e.Last password? Recovery password link.	file:///F:/webblogin/page.html	Application should show below UI elements: a.email text box b.password text box c.Login button with orange colour d.New customer? Create account link e.Last password? Recovery password link	Working as expected	Fail	Steps are not clear to follow	No	BUG-1234	naveen .p
LoginPage_TC_003	Functional	Home page	Verify user is able to log into application with Valid credentials	Proper network and URL	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter Valid username/email in Email text box 4.Enter valid password in password text box 5.Click on login button	Username: chalam@gmail.com password: Testing123	User should navigate to user account homepage	Working as expected	Pass				naveen .p
LoginPage_TC_004	Functional	Login page	Verify user is able to log into application with InValid credentials	Proper network and URL	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter InValid username/email in Email text box 4.Enter valid password in password text box	Username: chalam@gmail password: Testing123	Application should show 'Incorrect email or password ' validation message.	Working as expected	Fail		No		Manoj.M
LoginPage_TC_004	Functional	Login page	Verify user is able to log into application with InValid credentials	Proper network and URL	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter Valid username/email in Email text box 4.Enter Invalid password in password text box 5.Click on login button	Username: chalam@gmail.com password: Testing12367898679887 6876	Application should show 'Incorrect email or password ' validation message.	Working as expected	Pass		No		Manoj.M
					1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account	Username: chalam password: Testing12367898679887 6876	Application should show 'Incorrect email or password ' validation message.	Working					

Acceptance Testing
UAT Execution & Report Submission

Date	17 November 2022
Team ID	PNT2022TMID42755
Project Name	Project – Global Sales Data analytics
Maximum Marks	4 Marks

1. 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).

2. 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	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	0	0	1	1
Totals	24	14	13	26	77

Fig 8.2 User Acceptance training

CHAPTER 9

RESULTS

9.1 PERFORMANCE METRICES

Performance metrics are defined as figures and data representative of an organization’s actions, abilities, and overall quality. There are many different forms of performance metrics, including sales, profit, return on investment, customer happiness, customer reviews, personal reviews, overall quality, and reputation in a marketplace. Performance metrics can vary considerably when viewed through different industries.

Performance metrics are integral to an organization's success. It's important that organizations select their chief performance metrics and focus on these areas because these metrics help guide and gauge an organization’s success.

CHAPTER 10

ADVANTAGES & DISADVANTAGES

10.1 ADVANTAGES

- **Data analytics helps an organization make better decisions**

Lot of times decisions within organizations are made more on gut feel rather than facts and data. One of the reasons for this could be lack of access to quality data that can help with better decision making. Analytics can help with transforming the data that is available into valuable information for executives so that better decisions can be made.

- **Increase the efficiency of the work**

Analytics can help analyze large amounts of data quickly and display it in a formulated manner to help achieve specific organizational goals. It encourages a culture of efficiency and teamwork by allowing the managers to share the insights from the analytics results to the employees.

- **The analytics keeps you updated of your customer behavioral changes**

In today's world, customers have a lot of choices. If organizations are not tuned to customer desires and expectations, they can soon find themselves in a downward spiral. Customers tend to change their minds as they are continuously exposed to new information in this era of digitization. With vast amount of customer data, it is practically impossible for organizations to make senses of all the changes in customer perception data without using the power of analytics

- **Personalization of products and services**

Gone are the days where a company could sell a standard set of products and services to customers. Customers crave products and services that

can meet their individual needs. Analytics can help companies keep track of what kind of service, product, or content is preferred by the customer and then show the recommendations based on their preferences.

- **Improving quality of products and services**

Data analytics can help with enhancing the user experience by detecting and correcting errors or avoiding non-value-added tasks. For example, self-learning systems can use data to understand the way customers are interacting with the tools and make appropriate changes to improve user experience.

10.2 DISADVANTAGES

- **Lack of alignment within teams**

There is a lack of alignment between different teams or departments within an organization. Data analytics may be done by a select set of team members and the analysis done may be shared with a limited set of executives. However, the insights generated by these teams are either of not much value or are having limited impact on organizational metrics.

- **Lack of commitment and patience**

Analytics solutions are not difficult to implement, however, they are costly, and the ROI is not immediate. Especially, if existing data is not available, it may take time to put processes and procedures in place to start collecting the data. By nature, the analytics models improve accuracy over time and require dedication to implement the solution.

- **Low quality of data**

One of the biggest limitations of data analytics is lack of access to quality data. It is possible that companies already have access to a lot of data, but the question is do they have the right data that they need? A top down

approach is required where the business questions that need to be answered need to be known first and what data is required to answer these questions can then be determined. In some cases, data may have been collected for historical reasons may not be suitable to answer the questions that we ask today.

- **Privacy concerns**

Sometimes, data collection might breach the privacy of the customers as their information such as purchases, online transactions, and subscriptions are available to companies whose services they are using. Some companies might exchange those datasets with other companies for mutual benefit. Certain data collected can also be used against a person, country, or community. Organizations need to be cautious of what sort of data they are collecting from customers and ensure the security and confidentiality of the data.

- **Complexity & Bias**

Some of the analytics tools developed by companies are more like a black box model. What is inside the black box is not clear or the logic the system uses to learn from data and create a model is not readily evident. For example, a neural network model that learns from various scenarios to decide who should be given a loan and who should be rejected. The usage of these tools may be easy but the logic of how decisions are made is not clear to anyone within the company.

CHAPTER 11

CONCLUSION

This report aims to increase the level of accuracy in the data analysing and prediction,and also to solve the intellectual and technical issues surrounding the analysis of massive data. This is not the first report written on massive data, and it will not be the last, but given the major attention currently being paid to massive data in science, technology

Data analytics helps companies develop new products/services that will have better responses from customers and increase their sales revenue and profits by analysing customer preferences through surveys. This will help them create new products/services that will have a better response from customers and increase their sales revenue and profits.

It provides insights into the past, present, and future performance of a business and can be used to help you forecast trends, identify opportunities for growth, and develop a strategic action plan for your company. It can be helpful for businesses to understand how their sales are doing, especially if they want to grow or make changes. It doesn’t have to be dull numbers or dry paragraphs. You can visualize it in the form of bar graphs and charts.

At finally we have analized the massive amount of data to get the details about the profit ,loss and also the sales prediction to get a better result. It will be easy to understand and very knowledgable with the data and whave increased the accuracy of the prediction.

CHAPTER 12
FUTURE SCOPE

Data analytics has a bright future ahead as it has more potential, which everyone can explore. There is no shortage of opportunities for those who want to explore this field and move forward with their career in this competitive market world.

Today, data analytics is being used in many fields such as healthcare, retail, transportation, manufacturing, and many others. However, there are certain areas where it can be used more effectively.

We have planned to develop the model by collecting more and more data to increase the prediction accuracy.

Data analytics used to analyze the information on customer preferences and trends to create customized products.

Data analytics helps companies monitor social media activities to know how their customers are reacting to their products or services before launching a new product or service that has a good response from customers. This will help them.

CHAPTER 13

APPENDIX

13.1 SOURCE CODE

13.1.1 Source Code (Web page)

```
<!DOCTYPE
html>
                <html>
                <head>
                <meta name="viewport" content="width=device-width, initial-scale=1">
                <style>
                body {
                  background-image: url('https://s3.amazonaws.com/utep-uploads/wp-
                content/uploads/sparkle-box/2019/06/28113346/desk-graphs.jpg');
                  background-repeat: no-repeat;
                  background-attachment: fixed;
                  background-size: cover;
                }
                .con{
                  max-width: 400px;
                  margin: auto;
                  background: whitesmoke;
                  padding: 30px;
                }
                .a{
                  background-color: darkgrey;
                }
                </style>
                </head>
                <body>
                <div class="container-fluid " style="background-color: grey;">
                <a href="file:///F:/web/registration.html">
```

```
<img src ="https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQC1ueC3iCmaW5la1MI4pF2AY2Yp20Utb6bwQ&usqp=CAU"  style="width:38px; height:38px;" title="Goto Registration "></a>
</a>
</div>
<br>
<br>
<div class="con">

<h1>Login page</h1>
<form>
<label for="fname">User id:</label><br>
<input type="text" id="fname" name="Name"><br>
<label for="Password">Password:</label><br>
<input type="password" id="Password" name="Password"><br><br>
<a href="file:///F:/web/Mainpage.html" target="_blank">sign in</a><br><br>
<a href="file:///F:/web/forgetpassword%20.html" text-algin:center>forget pasword..</a>

</form>
</div>

</body>
</html>
```

13.1.2 SOURCE CODE (ML CODE WITH COLAB)

3498 lines (3498 sloc) | 246 KB Raw Blame

Team ID:PNT2022TMID42755

Project Name:Global Sales Data Analytics

In [ ]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
```

In [22]:
```
data=pd.read_csv('/content/drive/MyDrive/Global_Data/Global_Superstore2 - Global_Superstore2.csv')
```

In [23]:
```
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

In [24]:
```
data.shape
```

Out[24]: (51290, 24)

In [25]:
```
data.describe()
```

Out[25]:

	Row ID	Postal Code	Sales	Quantity	Discount	Profit	Shipping Cost
count	51290.00000	9994.000000	51290.000000	51290.000000	51290.000000	51290.000000	51290.000000
mean	25645.50000	55190.379428	246.490581	3.476545	0.142908	28.610982	26.375915

Fig 1

In [28]:
```
data['Order Date'] = pd.to_datetime(data['Order Date'])
```

In [29]:
```
data.info()
```

```
RangeIndex: 51290 entries, 0 to 51289
Data columns (total 24 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Row ID          51290 non-null  int64
 1   Order ID        51290 non-null  object
 2   Order Date      51290 non-null  datetime64[ns]
 3   Ship Date       51290 non-null  object
 4   Ship Mode       51290 non-null  object
 5   Customer ID     51290 non-null  object
 6   Customer Name   51290 non-null  object
 7   Segment         51290 non-null  object
 8   City            51290 non-null  object
 9   State           51290 non-null  object
 10  Country         51290 non-null  object
 11  Postal Code     9994 non-null   float64
 12  Market          51290 non-null  object
 13  Region          51290 non-null  object
 14  Product ID      51290 non-null  object
 15  Category        51290 non-null  object
 16  Sub-Category    51290 non-null  object
 17  Product Name    51290 non-null  object
 18  Sales           51290 non-null  float64
 19  Quantity        51290 non-null  int64
 20  Discount        51290 non-null  float64
 21  Profit          51290 non-null  float64
 22  Shipping Cost   51290 non-null  float64
 23  Order Priority  51290 non-null  object
dtypes: datetime64[ns](1), float64(5), int64(2), object(16)
memory usage: 9.4+ MB
```

Fig 2

In [30]:
```
a = data.groupby(['Order Date', 'Profit'])
a.first()
```

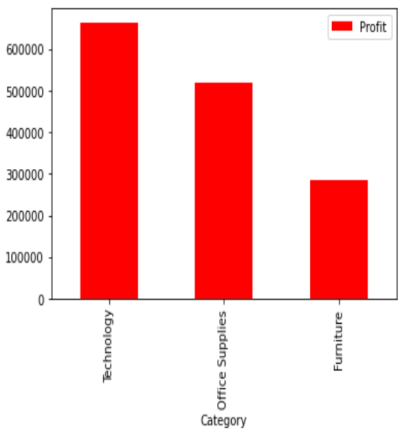
Out[30]:

Order Date	Profit	Row ID	Order ID	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	Country	...	Region	Product ID	Category	Sub-Category	Product Name
	-26.055	11731	IT-2011-3647632	05-01-2011	Second Class	EM-14140	Eugene Moren	Home Office	Stockholm	Stockholm	Sweden	...	North	OFF-PA-10001492	Office Supplies	Paper	Enermax Note Cards, Premium
	15.342	22254	IN-2011-47883	08-01-2011	Standard Class	JH-15985	Joseph Holt	Consumer	Wagga Wagga	New South Wales	Australia	...	Oceania	OFF-PA-10001968	Office Supplies	Paper	Eaton Computer Printout Paper, 8.5 x 11
2011-01-01	**29.640**	48883	HU-2011-1220	05-01-2011	Second Class	AT-735	Annie Thurman	Consumer	Budapest	Budapest	Hungary	...	EMEA	OFF-TEN-10001585	Office Supplies	Storage	Tenex Box, Single Width
	36.036	22253	IN-2011-47883	08-01-2011	Standard Class	JH-15985	Joseph Holt	Consumer	Wagga Wagga	New South Wales	Australia	...	Oceania	OFF-SU-10000618	Office Supplies	Supplies	Acme Trimmer, High Speed
	37.770	22255	IN-2011-47883	08-01-2011	Standard Class	JH-15985	Joseph Holt	Consumer	Wagga Wagga	New South Wales	Australia	...	Oceania	FUR-FU-10003447	Furniture	Furnishings	Eldon Light Bulb, Duo Pack
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Fig 3

Out[75]: (51290, 23)

In [76]:
```
data.groupby(['Category']).sum()[['Profit']].sort_values(by="Profit",ascending=False).nlargest(n=5, columns=['Profit']).plot.bar(color="Red")
plt.show()
```



In [77]:
```
data.groupby(['Category']).count()[['Order ID']].sort_values(by="Order ID",ascending=False).nlargest(n=5, columns=['Order ID']).plot.pie(subplots=Tr
plt.show()
```

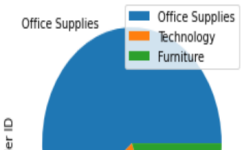


Fig 4

13.2.1 GITHUB LINK:

https://github.com/IBM-EPBL/IBM-Project-14903-1659591699

13.2.2 DEMO VIDEO LINK:

https://drive.google.com/file/d/1pFYeuR5YnjpgTCgD9qHUu4H3S0z1Y13s/view?usp=drivesdk