

PROJECT REPORT
GLOBAL SALES DATA ANALYSIS

TEAM ID: PNT2022TMID12427

Submitted by

HARISHDAVID J	19P111
KAMESHWARAN P	19P114
SAKTHI HARISH G	19P126
SARANKUMAR K	19P133

KARPAGAM COLLEGE OF ENGINEERING
(Autonomous)

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	i
1	INTRODUCTION	1
	1.1 PROJECT OVERVIEW	
	1.2 PURPOSE	
2	LITERATURE REVIEW	4
	2.1 EXISTING PROBLEM	
	2.2 REFERENCES	
	2.3 PROBLEM STATEMENT DEFINITION	
3.	IDEATION & PROPOSED SOLUTION	13
	3.1 EMPATHY MAP CANVAS	
	3.2 IDEATION AND BRAINSTORING	
	3.3 PROPOSED SOLUTION	
	3.4 PROBLEM SOLUTION FIT	
4.	REQUIREMENT ANALYSIS	18
	4.1 FUNCTIONAL REQUIREMENT	
	4.2 NON-FUNCTIONAL REQUIREMENT	
5	PROJECT DESIGN	20
	5.1 DATA FLOW DIAGRAMS	

	5.2 SOLUTION AND TECHNICAL ARCHITECTURE	
	5.3 USER STORIES	
6	PROJECT PLANNING AND SCHEDULING	23
	6.1 SPRINT PLANNING & ESTIMATION	
	6.2 SPRINT DELIVERY SCHEMA	
7	SOLUTIONING	25
	7.1 FEATURE 1	
	7.2 FEATURE 2	
8	TESTING	30
	8.1 TEST CASES	
	8.2 USER ACCEPTANCE TESTING	
9	RESULTS	32
	9.1 PERFORMANCE METRICS	
10	ADVANTAGES AND DISADVANTAGES	35
11	CONCLUSION	38
12	FUTURE SCOPE	39
13	APPENDIX	40
	GITHUB LINK	

ABSTRACT

In order to realize the automation of large data analysis a series of advanced algorithms, such as data cleaning and clustering, are used to deeply excavate the market sales data of sales and sideline products and it provide the necessary data basis for the decision layer of the enterprise. We designed and developed a visual analysis system of market sales data orients to products and business circles in the article. First, we clean and filter the original market sales data, and then establish the model based on clustering analysis in machine learning. Finally, the automatic display of analysis result views are achieved through excellent visualization framework technology. The experimental results show that in the actual sales data onto products, the characteristics of the business circle that the system automatically aggregates are obvious, and the combined display mode of the main and auxiliary views can show the characteristics of the sales data of a commercial circle from the macro and micro perspectives, respectively. It not only utilizes the ability of computer automatic analysis, but also fully exploits people's cognitive ability of visual information and it helps people understand the information, knowledge and wisdom behind the big data of the market sales more intuitively and efficiently. Based on web visualization technology, a case of visual analysis of market sales data onto agricultural products and business circles is presented in this paper. And it has practical reference significance of other typical marketing analysis, especially for products market.

CHAPTER 1

INTRODUCTION

1.1 PROJECT OVERVIEW

“Data analysis is a process of inspecting, cleansing, transforming and modelling data with the goal of discovering useful information, informing conclusions and supporting decision-making.”

Data analysis, in simple terms, means drawing important conclusions after collecting and organising relevant data. The main purpose of data analysis is to extract useful data and make key informed decisions. This process of data analysis involves systematic application of statistical tools and logical reasoning to gain information and evaluate the data. This is relevant for businesses as they analyse past and future activities, performance, and decisions and transform information into initiatives for improvement.

Data analysis makes it easier for users to interpret data to identify patterns and trends. There are different types of techniques used for data analysis and the method used depends upon the size and type of business. Some techniques are mentioned below:

- Text Analysis
- Statistical Analysis
- Diagnostic Analysis
- Predictive Analysis
- Prescriptive Analysis
- Descriptive Analysis

Data analysis helps the organisation to make the right choices and is used by businesses for market analysis, customer analysis, enterprise analysis, product analysis etc. Data integrity is an important component of analytics. The data used for this technology has to be reliable, complete, appropriate and accurate for absolute and relevant analysis.

The process of data analysis is based on extracting the information, preparing, integrating and finally consuming the data. It consists of the following:

- Requirement gathering

- Collection
- Cleaning
- Analysis
- Interpretation
- Visualisation

Data analytics technology is an aid for businesses to be more cohesive, efficient and informative than before. Businesses will be able to take useful actions when the quality of the collected data is checked and the analysis is thorough, precise and accurate.

In sales, many tasks are now managed through centralized cloud software, including CRMs, email marketing platforms and integration tools, making sales data readily available. Many global, industry-leading brands are now using their sales data in ingenious ways to make better business decisions, but any company can take advantage of insights and reporting tools to achieve data-driven sales success. However, the prospect of sifting through the many sales metrics available to make sense of the data can be overwhelming, while knowing what to do with that information once you've got it is another challenge.

1.2 PURPOSE

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.

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.

Here are some other KPIs you need to consider:

- Regional Sales
- Average purchase value

- Sales per rep
- Quote-to-close
- Cannibalization rate
- Sell-through rate
- Lead conversion rate
- Product performance
- Sales to date
- Sales opportunities
- Sales targets
- Sales growth

CHAPTER 2

LITERATURE REVIEW

2.1 EXISTING PROBLEM

Crafting a good sales pitch from **sales data analysis** can be difficult. Getting the right data, hitting the right client pain points, crystallizing why your services are better than the competitors, all takes hard work. Honing your sales pitch to an art takes time, and even with a perfect picture, new clients take time to acquire. One of the best ways we've found to build a good sales pitch is to use data you already have.

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

Sales Data Analysis

Finding ways to mash those together into meaningful new understanding is the key. Here's an example of an analysis we performed for a client. They wanted to know how they could meet more of their current customers' challenges. This would result in higher client satisfaction, increased revenue (from cross-sells), and could help them in future sales efforts with potential clients.

Sample Information List For Sales Data Analysis

We took their entire list of current clients and added the following data:

- Industry
- Number of employees
- NAICS/SIC code
- Site locations with latitude/longitude coordinates

- Number of times they had been serviced (grouped by dates)
- Customer satisfaction survey data

Resources To Get Info From For Sales Data Analysis

This gave us a holistic view of challenges faced by specific industries. To obtain this, we used data from our client's data warehouse as well as used LinkedIn, Yahoo Finance, Dun & Bradstreet databases, Client websites, and social media. Doing this definitely took work, but it highlighted thousands of dollars worth in future revenue and client retention. See a portion of the analysis below.

1.1 REFERENCE

PAPER 1

TITLE: Data analysis and visualization of sales data

AUTHOR: Kiran Singh,Rakhi Wajgi

YEAR OF PUBLICATION: 2016 **SUMMARY:**

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 analyze 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, 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. 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.

PAPER 2

TITLE: Drug Sales Data analysis for outbreak detection of infectious diseases.

AUTHOR: Mathilde Pivette, Judith E Mueller, Pascal crepey, Avner Bar-Hen **YEAR**

OF PUBLICATION: 2014

SUMMARY:

A search for relevant publications was conducted in Pubmed, Embase, Scopus, Cochrane Library, African Index Medicus and Lilacs databases. Retrieved studies were evaluated in terms of objectives, diseases studied, data sources, methodologies and performance for realtime surveillance. Most studies compared drug sales data to reference surveillance data using correlation measurements or indicators of outbreak detection performance (sensitivity, specificity, timeliness of the detection).

We screened 3266 articles and included 27 in the review. Most studies focused on acute respiratory and gastroenteritis infections. Nineteen studies retrospectively compared drug sales data to reference clinical data, and significant correlations were observed in 17 of them. Four studies found that over-the-counter drug sales preceded clinical data in terms of incidence increase. Five studies developed and evaluated statistical algorithms for selecting drug groups to monitor specific diseases. Another three studies developed models to predict incidence increase from drug sales.

Drug sales data analyses appear to be a useful tool for surveillance of gastrointestinal and respiratory disease, and OTC drugs have the potential for early outbreak detection. Their utility remains to be investigated for other diseases, in particular those poorly surveyed.

PAPER 3

TITLE: Modelling and Forecasting Sales Data by Time Series Analysis

AUTHOR: S.G.Kapoor, P.Madhok, S.M.Wu **YEAR**

OF PUBLICATION: 1981

SUMMARY:

Time series modeling technique is used to model a series of sales data in which seasonality causes distinct spike peaks. The analysis of actual sales data shows that the seasonality in the data can be approximated by a deterministic function and the stochastic component is a sixth order autoregressive moving average model. Use of the combined deterministic and stochastic models to derive the minimum mean squared forecast yields reliable results.

PAPER 4

TITLE: Sales Data and The Estimation of Demand

AUTHOR: S.A.Conrad

YEAR OF PUBLICATION: 2017

SUMMARY:

The differences between sales and demand are discussed. For the newsboy problem the probability distribution of demand is determined from sales data, under the assumption that customers arrive at random. The optimum stock level is shown, in general, to be different from that in the literature.

The exponential growth in the data collected in the form of electronic health records, registries, or wearable sensors has bought a big data revolution in the sales industry. A large amount of information is required to be extracted from this substantially available data that provides the delivery system. The other issues associated with big data include non-uniform data, large number of variables, and need for real-time data analysis. In this paper, we examine the existing literature in all the databases within the ISI Web of Science with the purpose of reviewing the current research and develop a new agenda. Reviewing past work is focused on identifying areas where big data analytics (BDA) is being applied in sales industry management. The outcome of the review is a proposed framework on BDA capability and its impact on sales organization performance.

PAPER 5:**TITLE:** A Cointegration Analysis of Advertising and Sales Data**AUTHOR:** Caroline Elliott**YEAR OF PUBLICATION:** 2001**SUMMARY:**

It is argued that the nature of the industry level relationship between advertising and sales can give some indication of the form of competition in an industry. Hence, this paper examines whether there is a long-run, stable, equilibrium relationship between advertising and sales for food and soft drinks industries. Results suggest that the variables are non-stationary, but do not contain seasonal unit roots. Cointegration is not identified between soft drinks industry advertising and sales, which, together with the results of differenced variable regressions, suggests that rivalry between firms in this industry may be intense.

PAPER 6**TITLE:** Sales and Advertisement Relationship for Selected Companies Operating in India**AUTHOR:** Supran Sharma**YEAR OF PUBLICATION:** 2021**SUMMARY:**

The study tries to examine the growth pattern and trend of sales and advertisement expenses for the selected companies over a period from 1992-93 to 2006-07. Further it seeks to evaluate the effectiveness of advertisement expenses on sales of selected companies operating in India at aggregate and disaggregate levels. It also tries to analyse the behaviour of share of advertisement expenses in total sales for the above mentioned categories. The study is based on panel or pooled secondary data collected for advertisement expenditure and sales revenue of 134 randomly selected sample companies operating in India over the period from 1992/93 to 2006/07 which are further classified on the basis of amount of sales revenue as well as on the basis of type of product produced. In this study of Panel or Pooled data, Fixed Effect approach with and without dummy variables are applied to evaluate the effectiveness of advertisement expenses on sales. Further, annual compound growth rates and summary statistics are also estimated. The contribution has found that the growth rate of sales revenue of manufacturing and companies, whose sales revenue is more than 1000 crore, is highest in spite of the negative compound growth rate of advertising expenses of these two types of companies. Further, in

case of non-manufacturing companies, it has been found that these companies are less popular among the consumers and they are also spending less on advertisements as compared manufacturing companies. While answering, how much advertisement expenses need to be incurred, the study concludes that to a large extent it depends on the nature and size of industries. Paper Classification: Research Paper.

PAPER 7

TITLE: Visual Analysis System for Market Sales Data of Agricultural products

AUTHOR: Kaiyi zhao, Ruizhi sun, Chao Deng, Li Li, Qiannan Wu, Sicong Li

YEAR OF PUBLICATION: 2018

SUMMARY:

In order to realize the automation of large data analysis and mining, a series of advanced algorithms, such as data cleaning and clustering, are used to deeply excavate the market sales data of agricultural and sideline products and it provide the necessary data basis for the decision layer of the enterprise. We designed and developed a visual analysis system of market sales data orients to products and business circles in the article. First, we clean and filter the original market sales data, and then establish the model based on clustering analysis in machine learning. Finally, the automatic display of analysis result views are achieved through excellent visualization framework technology. The experimental results show that in the actual sales data onto agricultural products, the characteristics of the business circle that the system automatically aggregates are obvious, and the combined display mode of the main and auxiliary views can show the characteristics of the sales data of a commercial circle from the macro and micro perspectives, respectively. It not only utilizes the ability of computer automatic analysis, but also fully exploits people's cognitive ability of visual information and it helps people understand the information, knowledge and wisdom behind the big data of the market sales more intuitively and efficiently. Based on web visualization technology, a case of visual analysis of market sales data onto agricultural products and business circles is presented in this paper. And it has practical reference significance of other typical marketing analysis, especially for agricultural products market.

REFERENCE LINKS:

https://careereducation.smartinternz.com/Student/guided_project_info/12917#

<https://www.sciencedirect.com/science/article/abs/pii/S0148296317300231>

<https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-014-0604-2>

<https://journals.sagepub.com/doi/abs/10.1177/002224378101800110>

<https://www.tandfonline.com/doi/abs/10.1057/jors.1976.13>

<https://www.sciencedirect.com/science/article/pii/S2405896318312242>

<https://www.sciencedirect.com/science/article/pii/S187705091730412X>

<https://www.journals.uchicago.edu/doi/abs/10.1086/511995>

<https://www.sciencedirect.com/science/article/abs/pii/S0167811609000512>

2.2 PROBLEM STATEMENT DEFINITION

Problem statements are definitions of the actual problems faced by consumers or people and reframing them in human-centric ways. The process of the defined problem statement is part of the second stage of the design thinking process. This stage is preceded by a phase called empathize, which involves the identification and observation of problems your users are facing.

- The data problem statement should be the primary focus of a leader involved in Data Analytics. That said, I will lament to ye dear reader the trials, tribulations, and “epics” of Data Problem Statements made in haste or driven by ego! In Law School, you learn the term *In Media Res* or “in the midst of”. It’s an allegory to deliver a broader message.
- First, when assessing if a problem requires the resources of data analytics to run it by more than one person. Get multiple perspectives without going to those people who will say yes to everything you say. You want scrutiny, and you know where to find it! Be clear in your intent by prefacing the request. Explaining that you’re looking for other ways the problem could be stated and that you are not asking them to brainstorm solutions possible solutions will focus the conversation. This is not a formal meeting. Do not fall prey to human instinct and try to solve the problem!
- Second, a good Data Problem Statement passes the Categorical sniff test when it is described in a more general term. Can it be categorized? If a problem statement is too specific it is likely only your problem. A crude example, is if the problem statement takes the form of a proper noun. If your problem statement is “Dr. Smith is a problem”, you’ll find it is difficult to solve with normative data.
- Third, your data problem is probably not unique. We, humans, are seldom content with the present, making it probable that your problem statement has something to do with time-series forecasting. For a data science purist the phrase “time-series forecasting” sounds like nails on a chalkboard. Technically, time-series is the tool. I add forecasting to help keep some context to what it is we are discussing. Sue Me.

Knowing what type of Data Problem you are dealing with makes articulating and getting value from the model all the likely. A quick primer of Data Analytics Tools and Techniques will strengthen your understanding of some basic outcomes data analytics can provide.

- The Fourth *piece* of advice will require the figurative sledgehammer! Focus on the problem statement one final time. Imagining the sledgehammer in your hands, find the center of the problem statement, now pull the hammer back and strike the center of the problem! Hopefully, that didn't fall too flat, and your imaginary problem statement is now broken into smaller pieces. There is nothing novel here, unless you are refereing to data. We are talking about data. For a thorough explanation on the differences of problem solving and data problem statements see my post [The Philosophy of Data](#).

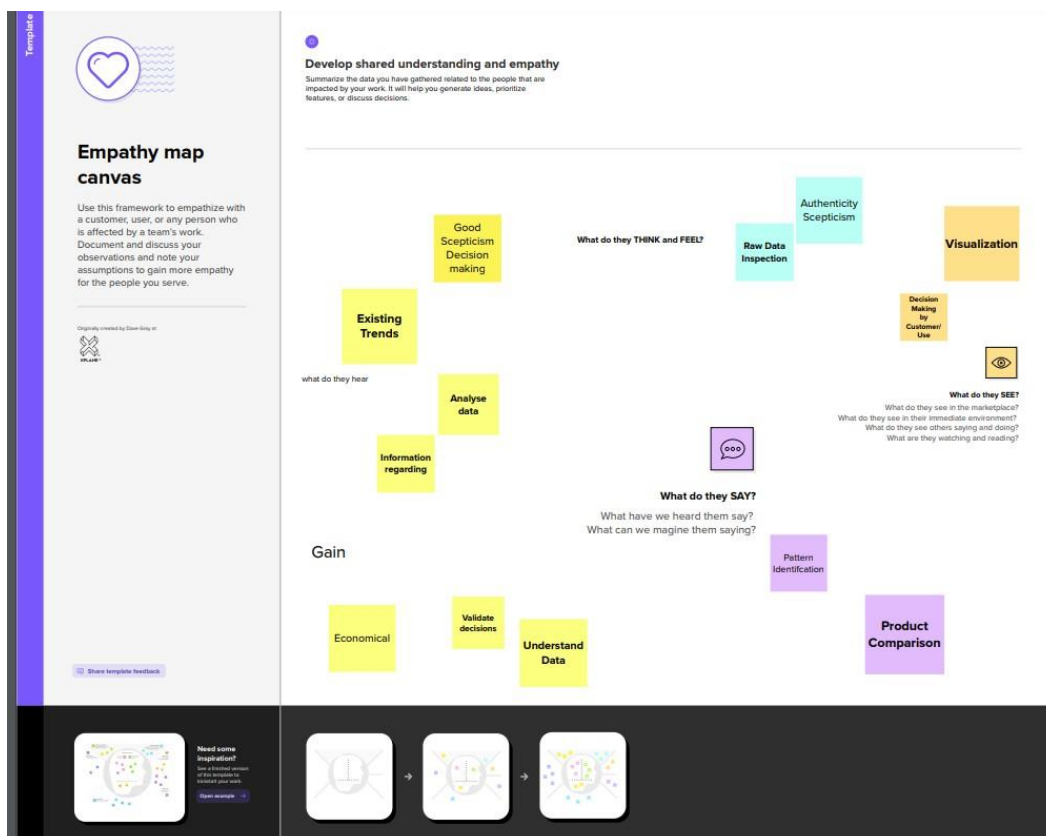
CHAPTER 3

IDEATION & PROPOSED SYSTEM

3.1 EMPATHY MAP

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes.

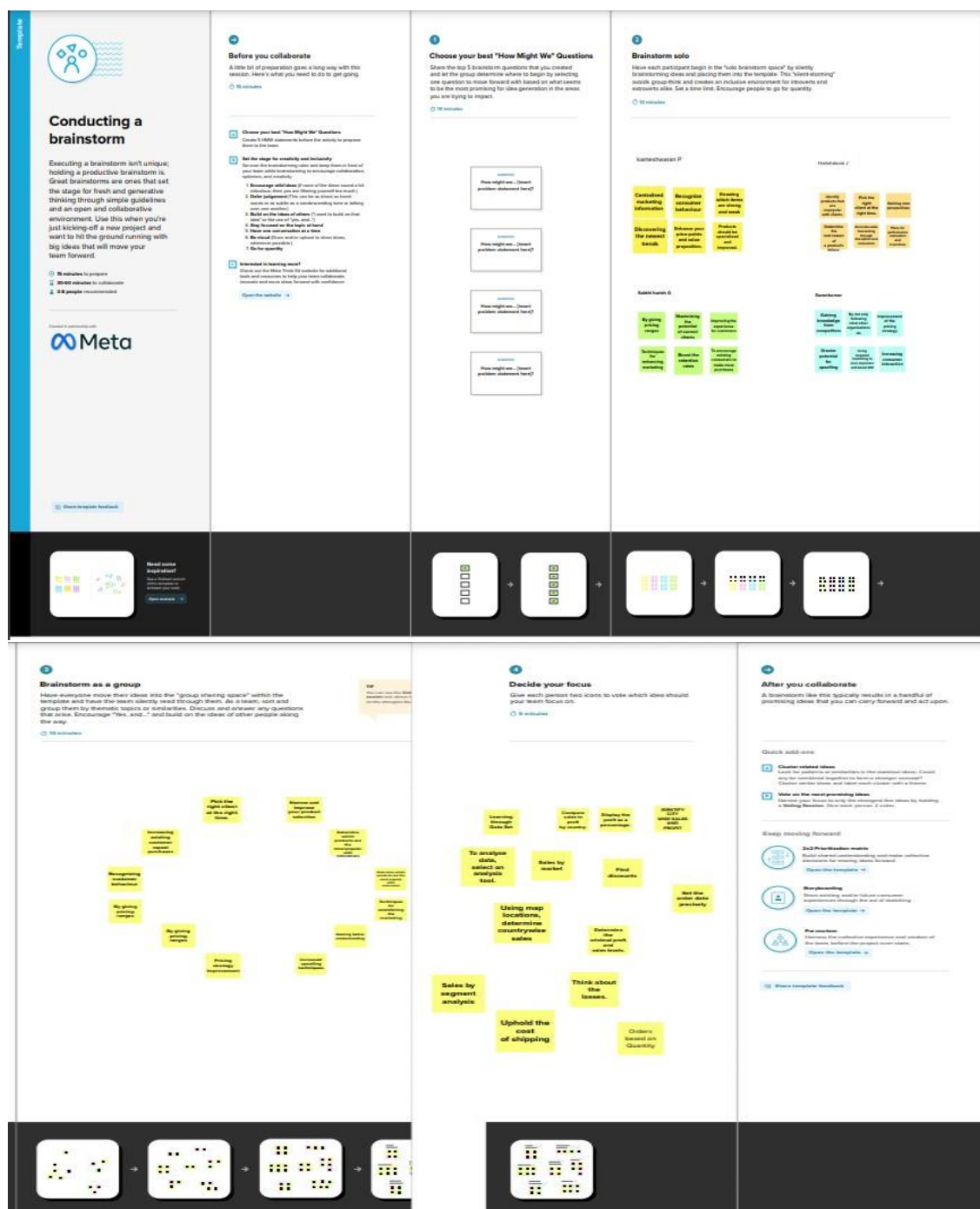
It is a useful tool to help teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



3.2 IDEATION AND BRAINSTORMING

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions. Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

Reference: <https://www.mural.co/templates/empathy-map-canvas>



3.3 PROPOSED SYSTEM

Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement(Problem to be solved)	To produce data-driven decisions by utilizing data analytics where business professionals can analyse customers, products, and new trends. A crucial component of operating a successful firm is sales analysis. We can choose which products to focus on, where to sell, and how to effectively reach customers using sales data.
2.	Idea/Solution description	Currently, analytic programmes are available, but they are not an ideal fix. To address this, we are developing a tailored analytics model that will assist companies of all sizes in increasing revenue, automating processes, making wiser decisions, and keeping you informed of changes in customer behaviour.
3.	Novelty/Uniqueness	A successful data analytics program that gives you a clear image of where you are, where you have been, and where you should go will be made possible with the help of the tailored analytics model. In line with this model. As a result, this model will be unique.
4.	Social Impact/Customer Satisfaction	The use of this strategy will benefit a variety of enterprises by enhancing efficiency, boosting revenue, and reducing loss of income.

5.	Business Model(RevenueModel)	The methodology will be successful because it helps businesses better understand their clients, assess their advertising efforts, personalise content, and develop content strategies. Given its utility, it will undoubtedly attract clients.
6.	Scalability of the Solution	According to the dataset given, the tailored model will provide a crisp visual understanding with an attractive and engaging display and understanding. The data visualisation is used to spot trends, patterns, and other things which results in making data driven decisions

3.4 PROBLEM SOLUTION FIT

1.CUSTOMER SEGMENT

People and corporations who are interested in knowing the details and a complete global sales analysis of a product and related products for effective customer making organizations which wants to know their products sales details.

2.PROBLEMS/PAINS

Sales Analysis 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.

3.TRIGGERS

TO ACT Better performance of their business competitors, fall down of their performance.

4.EMOTIONS

Satisfaction and may lead to new achievement and betterment of self and business growth.

5.AVAILABLE SOLUTIONS

Sales metrics, revenues, gross number of sales. Simply measuring revenue or the gross number of sales isn't enough. The right metrics will depend on your company, but are valuable to learn more about your company, customers, and sales process.

6.CUSTOMER LIMITATIONS

Lack of understanding over the business and the customer engagement over it.

7.BEHAVIOUR

Actions against losing customers, changes in budget, advertising and collaborations for betterment.

8.CHANNEL OF BEHAVIOUR ONLINE

Using third party services with automated insights and subscription based service to analyse data. OFFLINE Using office software to analyse complex data in un-intuitive way.

9.PROBLEM ROOT/CAUSE

Being lethargic that the business is doing fine, absence of customer involvement monitoring, presence and progress of competitors in global market.

10. YOUR SOLUTION

- Creating an interactive dashboard.
- Providing specific details about sales
- Responsive design for every screen size ➤ Manual insight for each interaction.
- One time payment

CHAPTER 4

REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

A functional requirement defines a function of a software system or its component. A function is described as a set of inputs, the behavior, and outputs (see also software). Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish.

Following are the functional requirements of the proposed solution

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Signing up with Gmail Register or Log In
FR-2	User Confirmation	Email confirmation is delivered.
FR-3	Dataset	Dataset upload to Cognos Analytics Tool.
FR-4	Visualize/Analyse	Columns can be moved around to analyse the dataset.
FR-5	Create Dashboards	Create Charts , Graphs ,Tables etc.
FR-5	Log Out	After downloading the Dashboards, log out.

4.2 NON-FUNCTIONAL REQUIREMENTS

In systems engineering and requirements engineering, a non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors.

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

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Until the Dashboard contains the appropriate Store Sales Dataset, the user can view it.
NFR-2	Security	The Dashboards/Templates are accessible to anyone with the proper Log In credentials.
NFR-3	Reliability	Templates are trustworthy since we upload and access them over the cloud.
NFR-4	Performance	The user can easily drag to any metrics they want to view, and it works as intended.
NFR-5	Availability	Anyone who is interested in learning more about Sales Data can access it for free.
NFR-6	Scalability	Dashboards and templates are quite flexible; users can change the metrics at any time.

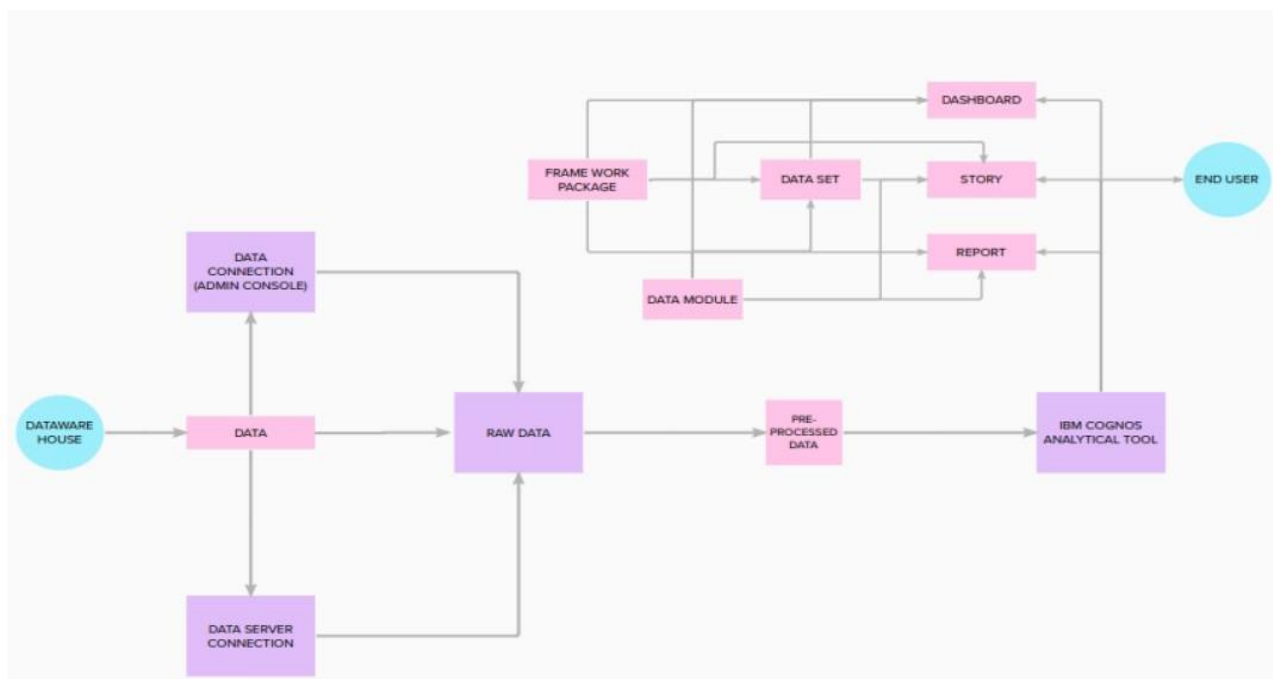
CHAPTER 5

PROJECT DESIGN

5.1 DATAFLOW DIAGRAM

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.



Data Flow Diagram For Global Sales Data Analytics:

FLOW STRUCTURES :

1. User configures credentials and starts the app.
2. User selects data files to process and load.
3. Apache tikan extracts text from the data file.
4. Extracted text is passed to Watson NLU for enrichment.

5.enriched data is visualized in the UI using D3.js library.

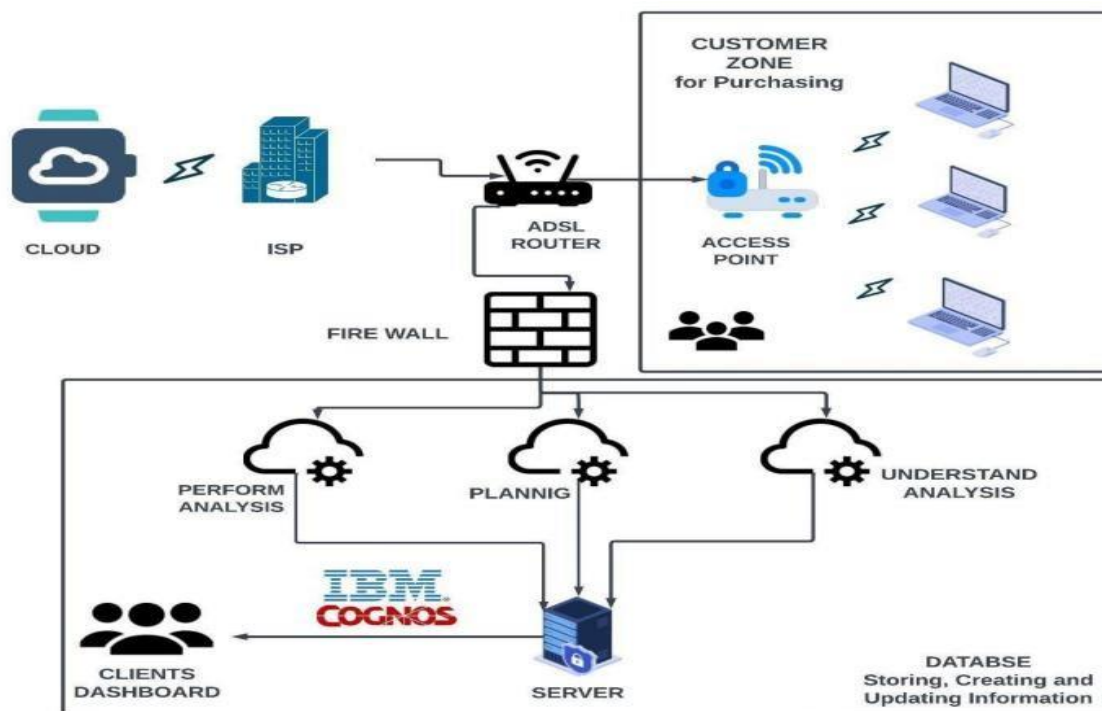
5.2 SOLUTION AND TECHNICAL ARCHITECTURE

Solution Architecture:

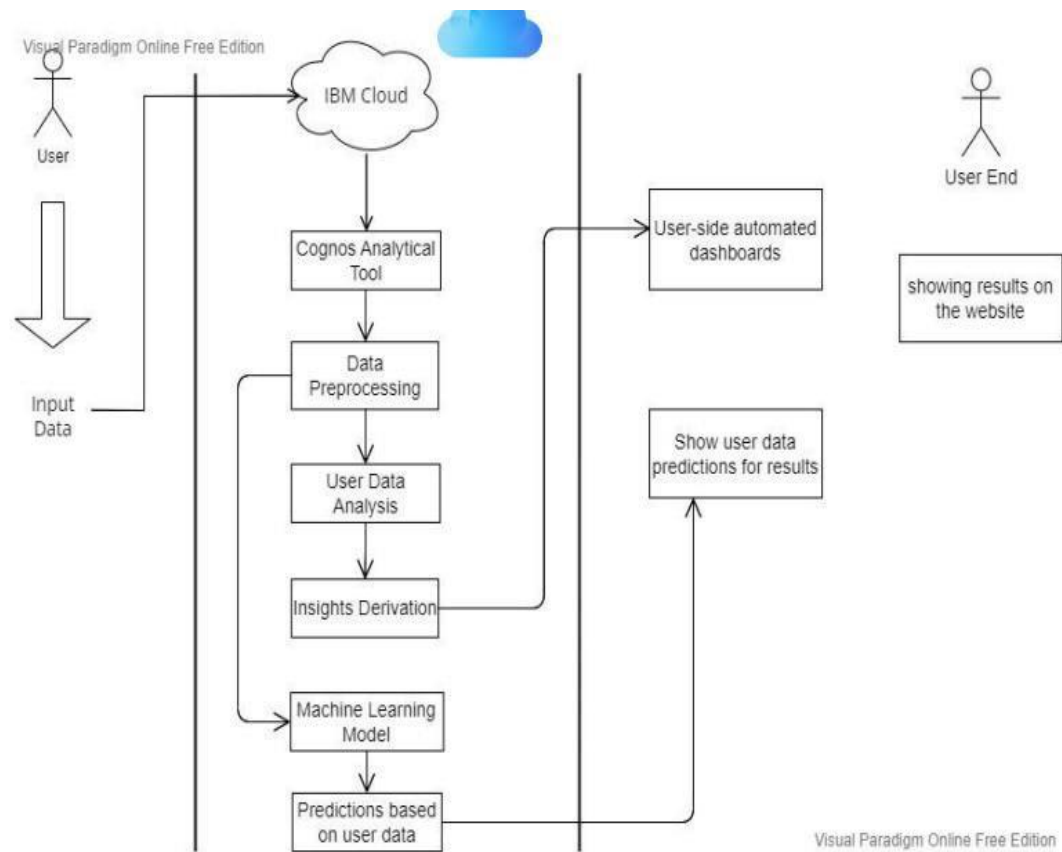
A complicated process with numerous sub-processes, solution architecture connects business issues with technological solutions. Its objectives are,

- To Solve current business issues, choose the most effective technological solution.
- To Give project stakeholders a description of the software's structure, attributes, operation, and other features.
- To Define the solution's conditions required, development stages, and functionalities.
- To Offer guidelines for how the solution is created, managed, and delivered.

Solution Architecture Diagram:



TECHNICAL ARCHITECTURE:



5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the 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 user, I can create the visualization by using the dashboard in the application		High	Sprint-3
Customer (Web user)	Login	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 and dashboard	High	Sprint-1

CHAPTER 6

PROJECT PLANNING & SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

Project Development Phase:

Sprint-1: Registration and Data upload

Sprint-2: Data preprocessing

Sprint-3: Dashboard Exploration and Creation

Sprint-4: Report and Story creating.

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	Kameshwaran P, Harishdavid j, Sakthi harish G, Sarankumar k
Sprint-1	Login	USN-2	As a user, I need valid credentials to log in to my application.	1	High	Kameshwaran P, Harishdavid j, Sakthi harish G, Sarankumar k
Sprint-1	Data Collection	USN-3	As a user, I need to gather the data in the form of CSV/XLS and clean the data	2	High	Kameshwaran P, Sakthi harish G, Sarankumar k
Sprint-2	Upload dataset	USN-4	As a user, I can view the data of the products	1	Low	Kameshwaran P, Sarankumar k
Sprint-2	Data Preparation	USN-5	As a user, I need to filter it for Data visualization.	3	High	Harishdavid j, Sakthi harish G
Sprint-2	Data visualization	USN-6	As a user, I can easily visualize the data in the form of charts.	4	Medium	Kameshwaran P, Harishdavid j

Sprint-3	Dashboard	USN-7	As a user, I can view the summary of the product sales by the help dashboard.	2	Medium	Kameshwaran P, Sakthi harish G
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	4	High	Sakthi harish G, Sarankumar k

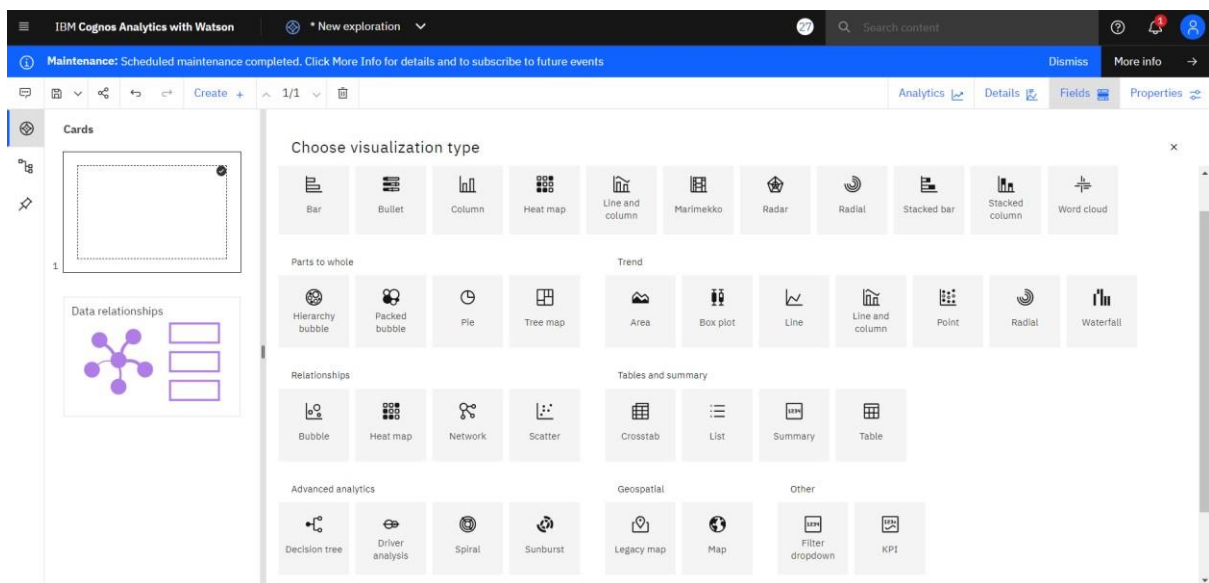
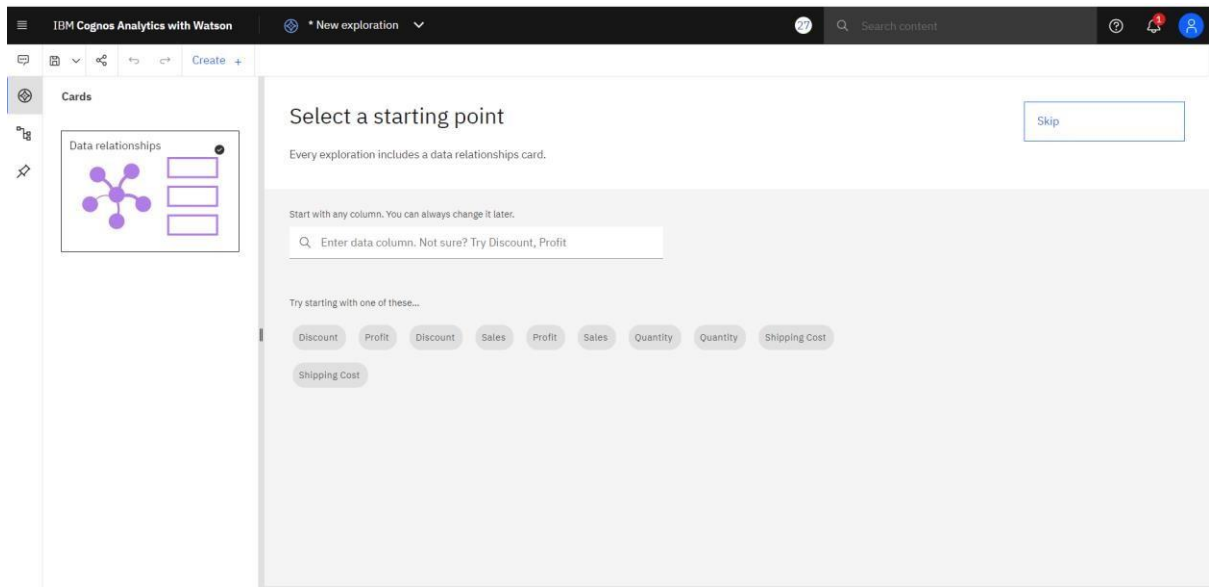
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-3	Dashboard	USN-9	As a user, I must be able to gain insights from the charts/graphs through a variety of relationships established in the dashboard.	4	Medium	Harishdavid j, Sarankumar k
Sprint- 4	Prediction	USN-10	As a user, I see the prediction of the specific product's future sales expectation.	4	Medium	Harishdavid j, Sakthi harish G
Sprint- 4	Report	USN-11	As a user, I can view the list of categorized products and their details as a report.	5	High	Kameshwaran P, Sakthi harish G
Sprint-4	Story	USN-12	As a user, I can view the product and customer description and more additional information as a story.	5	High	Harishdavid j, Sarankumar k

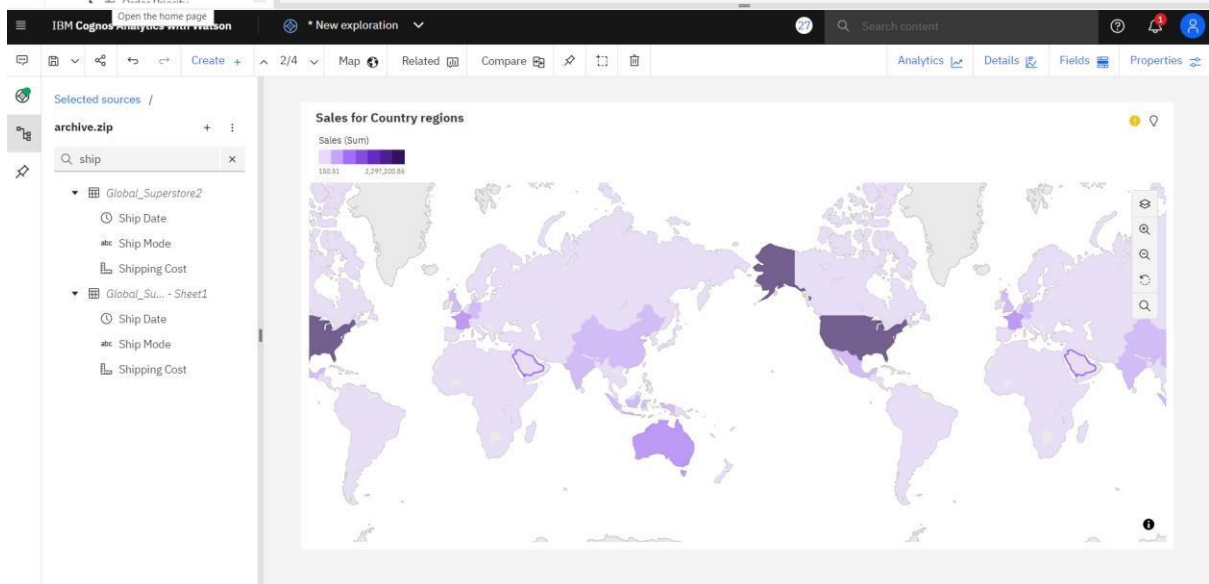
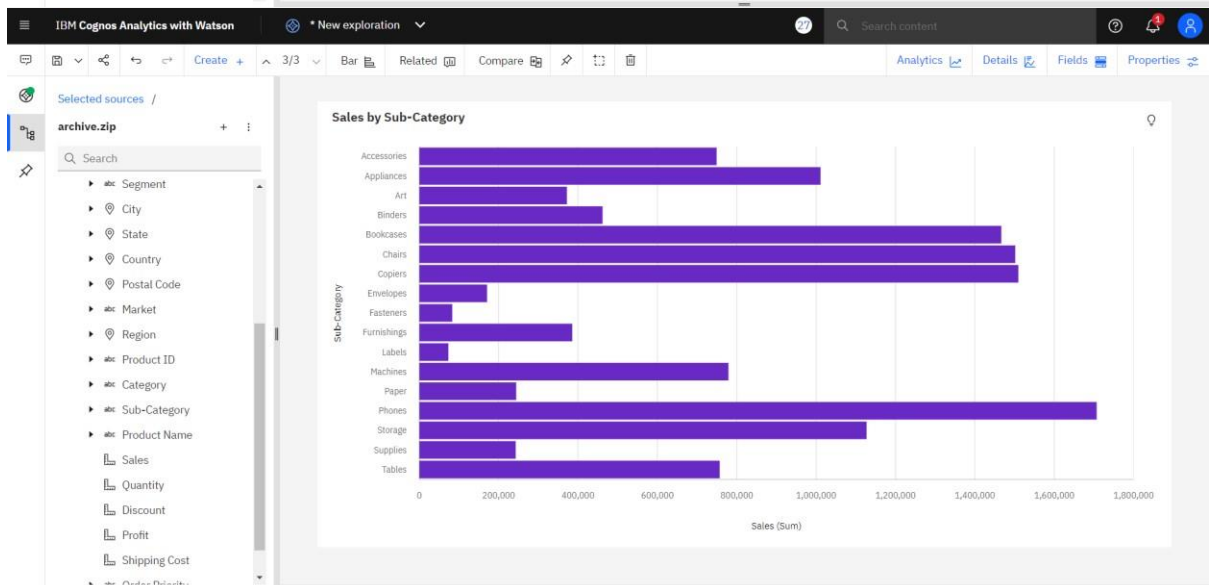
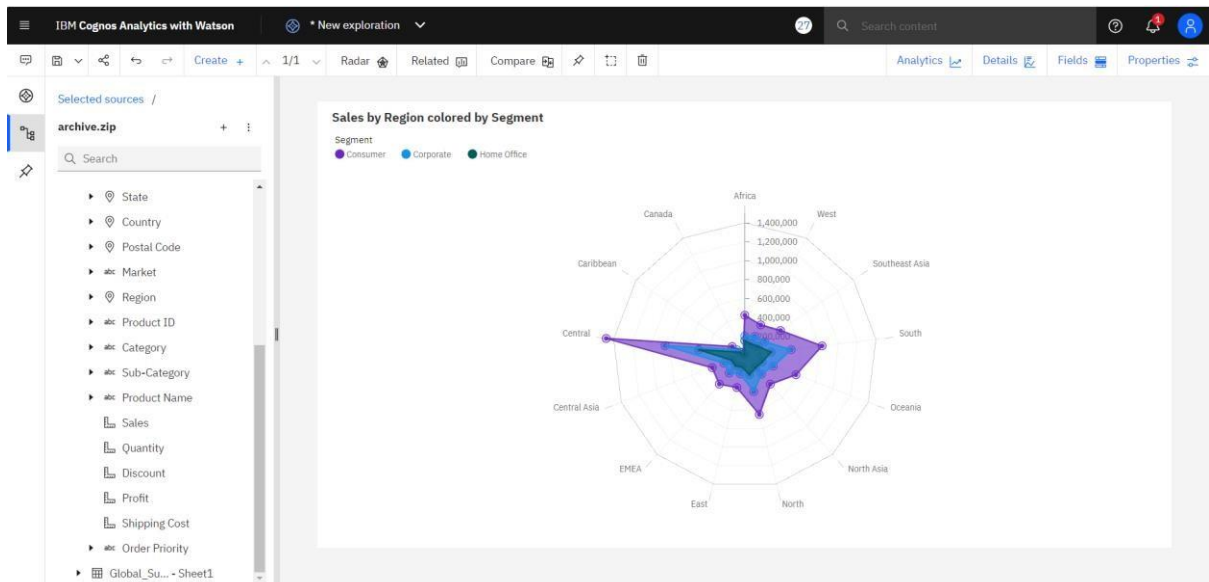
CHAPTER 7

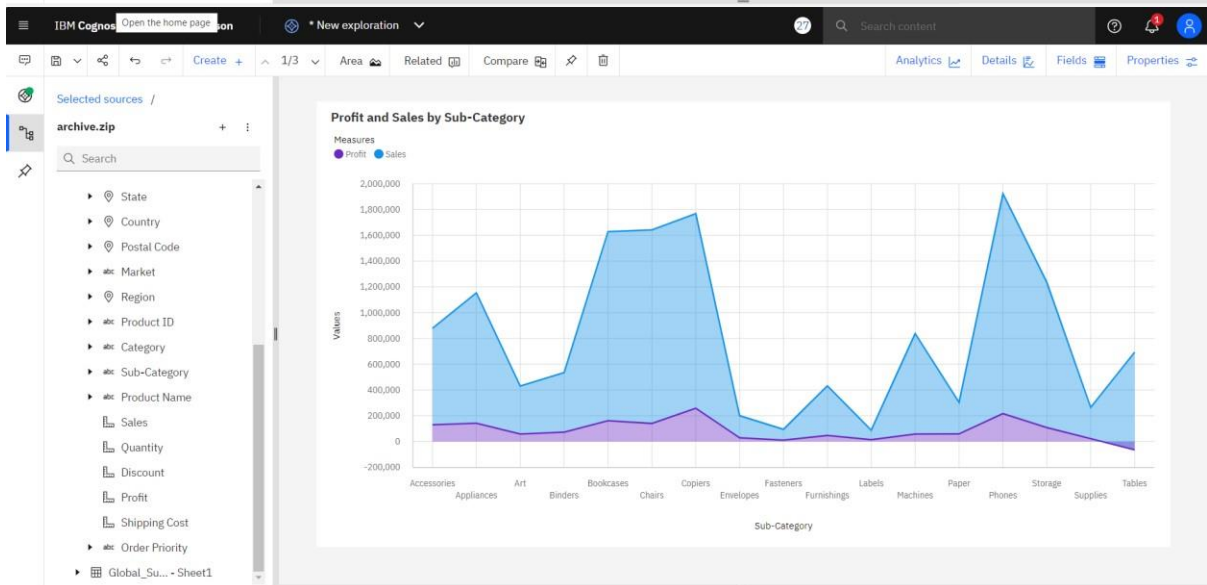
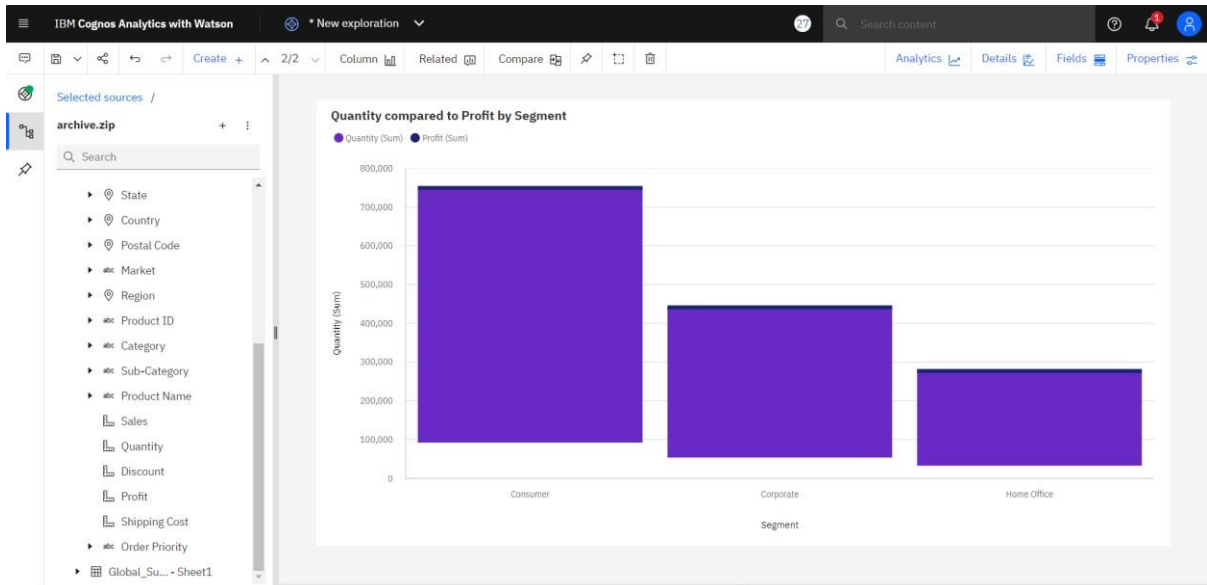
SOLUTIONING

7.1 FEATURE 1

Creating visualizations

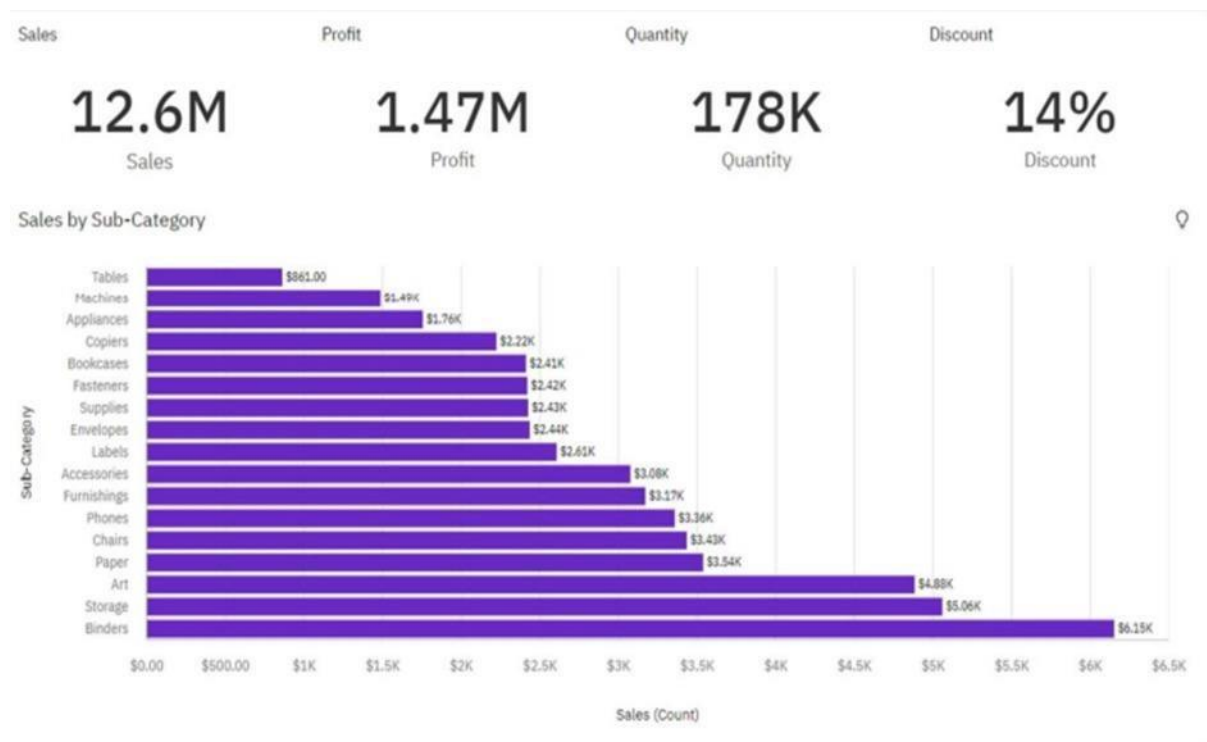
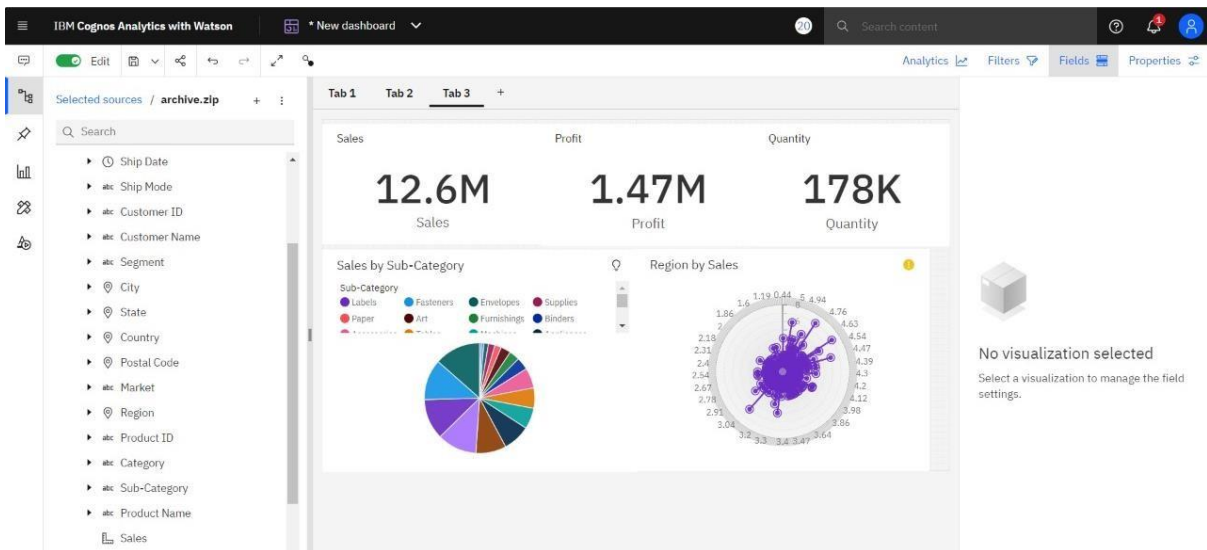






7.2 FEATURE 2

CREATING DASHBOARD



CHAPTER 8 TESTING

8.1 TEST CASES

Model Performance Testing:

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

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs - 7-8 visualization/6-7 graphs
2.	Data Responsiveness	Users and Analyst or Developers
3.	Amount Data to Rendered (DB2 Metrics)	5 countries
4.	Utilization of Data Filters	Sales ,profit, products, market rate and order id filtration
5.	Effective User Story	No of Scene Added - 30 user stories
6.	Descriptive Reports	No of Visualizations / Graphs - 4 visualizations/6 graph.

8.2 USER ACCEPTANCE TESTIN

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.

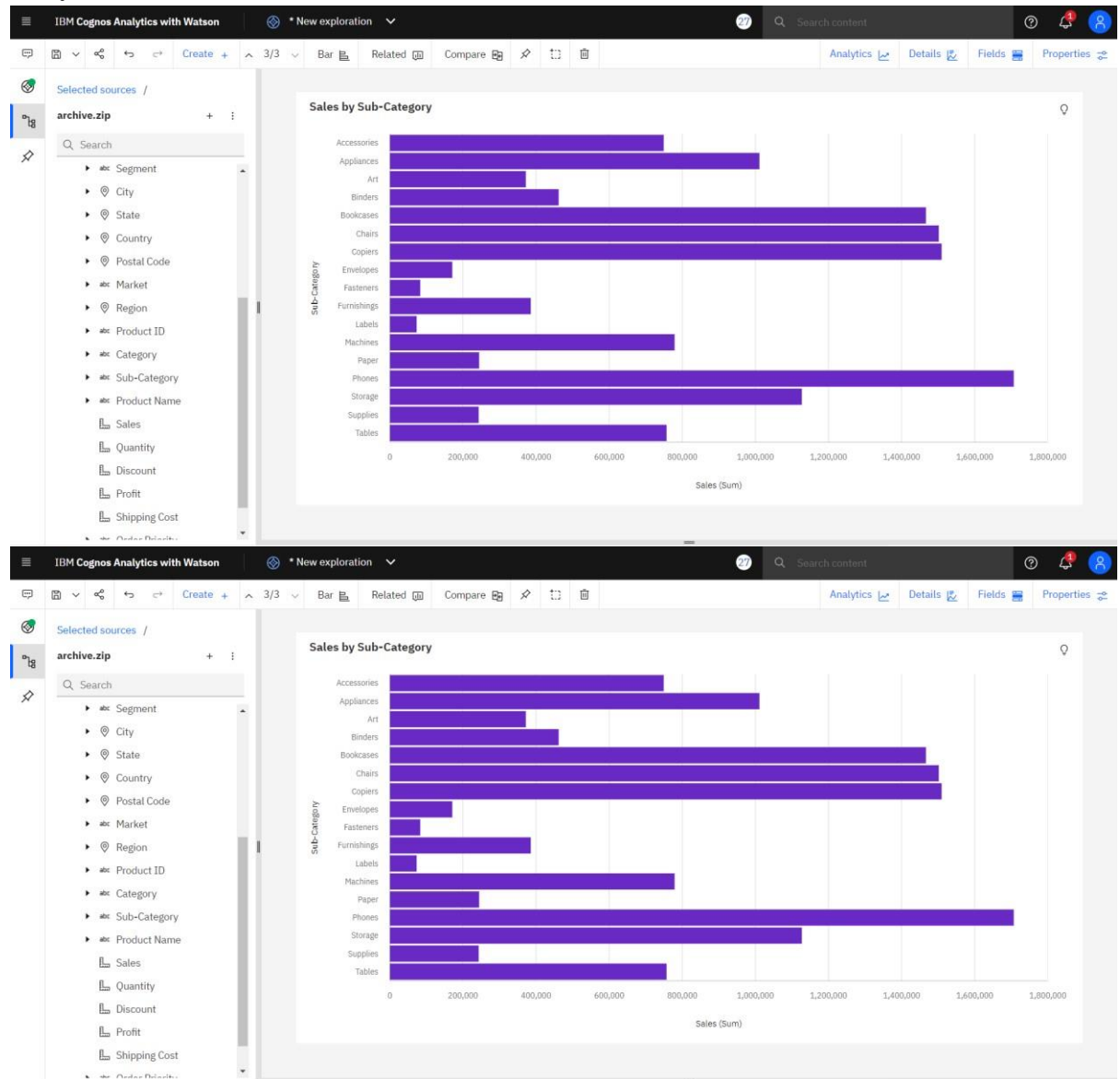
Resolution	Severity1	Severity2	Severity3	Severity4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won'tFix	0	0	0	1	1
Totals	24	9	11	26	71

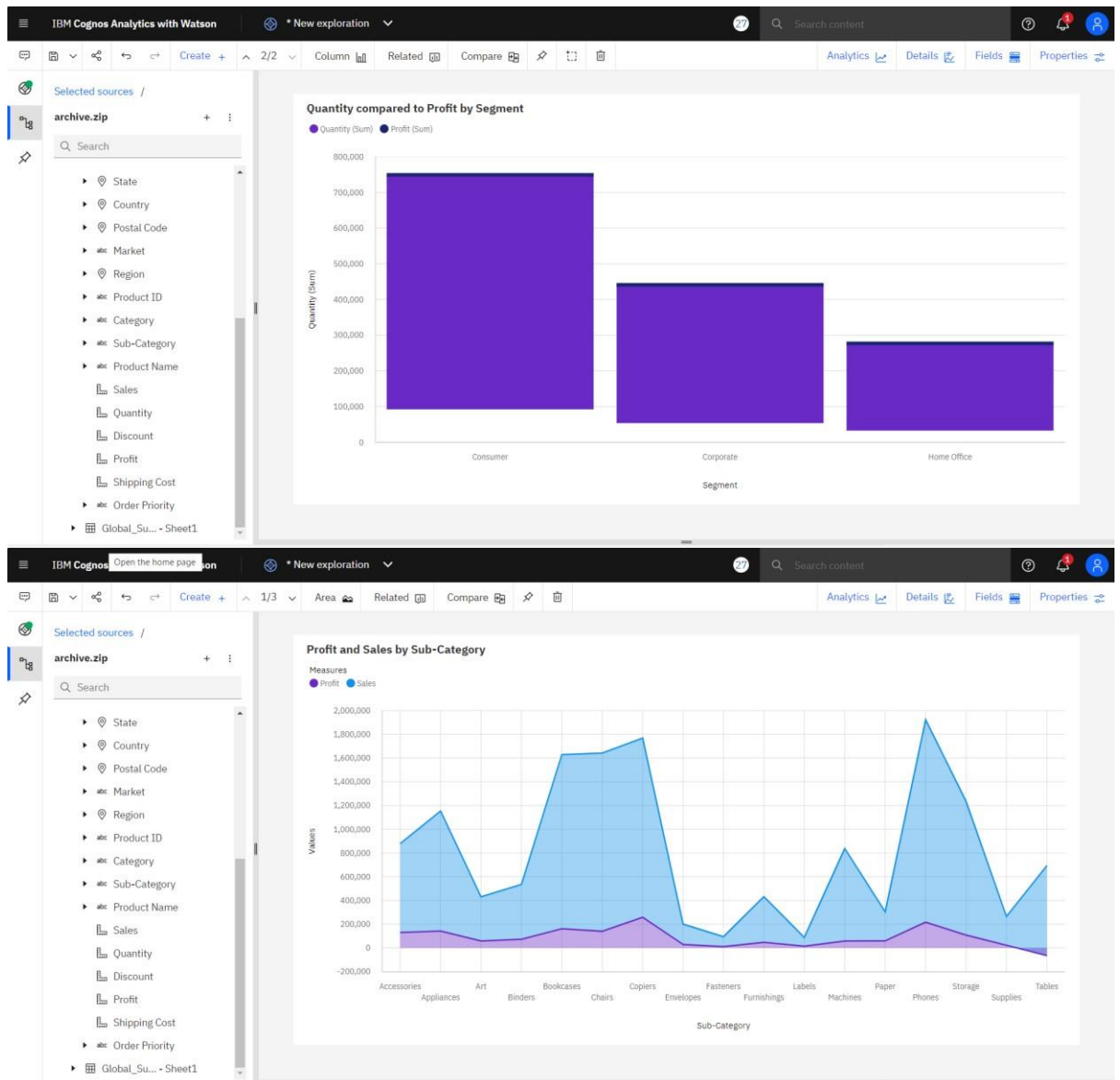
CHAPTER 9

RESULTS

9.1 PERFORMANCE METRICS

we're taking a closer look at the top sales performance metrics for kick-starting your sales analytics.





CHAPTER 10

ADVANTAGES AND DISADVANTAGES

Advantages

Cost efficiency

For many small and medium-sized companies, hiring full-time IT experts is costprohibitive. Data analytics outsourcing presents a cheaper alternative to having an in-house data analytics team.

Getting a trusted outsourcing partner lets you save on the costs associated with onboarding new employees, technology and licensing fees for equipment, paying full-time salaries, and more. The data outsourcing team you choose to go with will have its own infrastructure and cover all associated costs. You'll only have to pay your solutions provider an agreed-upon rate.

Receive full-scale services

Data analytics outsourcing doesn't just offer analytics reporting solutions. Your outsourcing partner will also use various tools and skilled data scientists to give you a comprehensive analysis of your data and actionable insights. Moreover, you won't have to worry about having the necessary infrastructure and technology for analytics since your analytics solutions provider will already have them in place from their end.

You'll also receive your analytics reports in a digestible format that will enable you to make informed business decisions. Many providers also offer ongoing support.

Maximize presentation

Data analytics outsourcing partners don't just give you the data you need to make informed business decisions. They also present it in a professional, organized, and easilycomprehensible way.

Apart from data analytics services, analytics providers also present the data in your preferred visualization method. Most packages include a dashboard that features graphs, a big-picture view, and much more.

If you ever need to create presentations for employees or stakeholders, the visual information can help inform your presentations. Essentially, all you need to do is pull the data from the reports and transfer it to your preferred data presentation application.

Save time

Regardless of the nature or size of your organization, there are high chances that you're generating more data than you can keep track of. A typical entity collects data about its customers, sales, social media visitors, website visitors, and much more. Sorting through all this data and turning it into actionable insights takes lots of time, more than what most businesses have to spare.

Moreover, you may want to go beyond simple analytics to cleanse and organize organizational data to gain insights into how your business is performing and what areas you can improve – which takes even more time.

Getting a trusted outsourcing partner takes these tedious and time-consuming tasks out of your hands, giving you more time to focus on other concerns. Additionally, since you're working with an expert, your analytics job will be done more thoroughly and efficient.

Disadvantage

Risk of choosing the wrong provider

Like any other outsourcing engagement, choosing the right outsourcing provider is always a challenge. You'll want an outsourcing partner that's affordable and has the right cultural fit and alignment with your organization. And with most providers advertising themselves as the best, it's not always easy to tell if they're really worth working with, especially when they don't have any referrals from reputable organizations.

Lack of on-site support

Some businesses are pretty reluctant when it comes to data analytics outsourcing as they feel more comfortable having an on-site analytics team. The added convenience of having an on-site analytics team affords is very beneficial for companies that require constant data analytics to streamline operations. However, having an on-site team is not always costeffective.

Less control

Some organizations don't like the idea of putting their data analytics needs in the hands of an outsourcing partner. Data analytics outsourcing essentially takes control out of your hands. Once you sign a contract and surrender your data to an analytics solutions provider, you have limited control over what they do with the data. This can prove problematic for businesses dealing with sensitive customer data. However, this can be easily managed with constant monitoring, improved communication, and collaborative project management.

Data security

When dealing with analytics outsourcing partners, organizations have to be prepared for the risk of exposing sensitive organizational data. That said, there are numerous ways to ensure that an analytics provider is trustworthy and able to keep all company data confidential.

Therefore, before working with an outsourcing partner, do your research to pick a provider that can ensure the safety of your data using clear, specific strategies. The security measures the provider plans to take should also be part of your agreement with the company.

CHAPTER 11

Conclusion

Data analysis is the process of analyzing data to discover useful information that is consistent with the objective of the research. Data analysis includes the inspection, modification, modeling, and transforming of data as per the need of the research topic. The conclusion is the final inference drawn from the data analysis, review of literature, and findings.

Data analysis is an important process of research or simply discovering information related to any work. Data derived from the observation, experiment, and other primary and secondary data collection methods is large and cannot be taken as it is. Not all data is relevant, neither can it directly signify any trends, relations, facts, and associations within the data. To find out those required trends and relations, the data needs to be reconstructed in the relevant form and modified. This process is called data analysis. Data analysis and conclusion take forward the research.

CHAPTER 12

FUTURE SCOPE

Data analytics is expected to radically change the way we live and do business in the future. Already today we use the analytics in our technology devices, for many decisions in our lives. Not only how to drive from A to B and avoid traffic-jams, but also to identify waste in business processes with the help of Lean six sigma optimization projects.

Although organizations are taking steps to turn data into insights, our global survey showed that organizations are still struggling with data quality and the problem to find the right resources to turn these insights into true value and become more data-driven.

Expectations are that data analytics will make the impossible possible, but we are still in the early stages of the data era. Basically, every company is currently investing in data analytics capabilities to keep up with known or unknown developments and competition.

where are we heading for in the next 30 years?

The known data analytics development cycle is described in stages: from descriptive (what happened) to diagnostic (why did it happen), to discovery (what can we learn from it), to predictive (what is likely to happen), and, finally, to prescriptive analytics (what action is the best to take). In general, organizations currently find themselves in the diagnostic and discovery stages.

Another way of looking at this is that data analytics initially “supported” the decision-making process, but is now enabling “better” decisions than we can make on our own. What comes to mind here are cases where analytics is applied to combine multiple data sources, resulting in new and better insights, for example to combine sales, location and weather data to understand sales increase for certain stores and improve the replenishment process.

If it turns out in the future that a decision-making process based on data analytics will produce better results, the step to “automated” decision-making will be small (e.g., artificial intelligence). Examples are the autopilot update in the Tesla model S cars, or the Google car, which has been driving around for more than 1.2 million miles without getting a single ticket.

CHAPTER 13

APPENDIX

13.1 GITHUB LINK

<https://github.com/IBM-EPBL/IBM-Project-12914-1659500233>