Retail Store Stock Inventory Analysis

Team ID: PNT2022TMID12898

Bachelor of Engineering

Electronics and Communication Engineering

PSG College of Technology

Coimbatore - 641 004

Faculty Evaluator	:	Dr.P.Visalakshi
Faculty Mentor	:	Dr.R.Venkateshwari
Industry Mentor	:	Shivani Kapoor, Rakesh Miskin

Team Members:

718019L213 - Gayathri G
718019L216 - Haritha K
718019L244 - Sanjith S
718019L245 - Sendhilnathan E

1.INTRODUCTION

1.1.PROJECT OVERVIEW:

In recent times, the employment of analytics in all kinds of business sectors, especially the retail sector has proven to increase success in their daily operations. This project aims to prove that, in addition, will identify what factors are actually contributing to this roaring success in the retail sector. Of course, the use of analytics in the business processes has its own pros and cons, but majority of the organizations feel that the introduction of analytics in their business processes has made things easier for them.Some of the drawbacks of using big data analytics in the retail sector has risen concerns among the customers as well the retailers. Privacy concerns are one of them. Customers feel that their privacy is being snatched away when retailers track their location or store their purchase information for targeting them with personalized advertisements. Although big data analytics help employees to fasten up their work, it also poses a high cost for managing such a huge amount of data. Software needed to sort and analyze this data is very expensive. On the other hand, it requires skilled people to work with them. Data quality decreases because of automation of data gathering, sorting and analyzing them.

1.2.PURPOSE:

The use of analytics decreases the use of man force as it automates all the processes but on the other hand. It helps in product development as analytics can carry out sentiment analysis of a lot of actual and potential customers through social media and find out their preferred types of products, developing their future products accordingly. The use of analytics lets the retailers predict future demands while analyzing their stocks. Micro targeting the customers can be easy when the location of customers can be easily known to the retailers by the use of analytics. Although there are many cons of adapting big data analytics in the business or retail sector, the pros are more and outweighs all the cons. This aims to prove that.

2.LITERATURE SURVEY

2.1.EXISTING PROBLEM:

In Retail store stock inventory analysis,miscommunication can cause irreversible damage to efficiency. As inventory management has numerous components, clear communication is vital for a seamless flow.Generally, insufficient access to information would lead to miscommunication issues. Warehouse management is a core component of brick-and-mortar retail inventories. Hence, ineffective warehouse management would affect the complete retail inventory process. A decentralized inventory management system would comprise the accuracy of the operations.Overselling is a result of flawed communication and process flow of your inventory management. This issue can also be a result of incorrect stock counting.

2.2.REFERENCES:

1) Inventory Management for Retail companies:

Authors: Cinthya Vanessa Muñoz Macas,Jorge Andrés Espinoza Aguirre ,Rodrigo Arcentales-Carrión
Published : IEEE 2021
Description: analyze and present an extensive literature concerning inventory management, containing multiple definitions and fundamental concepts for the retail sector. A systematic literature review was carried out to determine the main trends and indicators of inventory management in Small and Medium-sized Enterprises (SMEs). This research covers five years, between 2015 and 2019, focusing specifically on the
retail sector. The primary outcomes of this study are the leading inventory management systems and models, the Key Performance Indicators (KPIs) for their correct management, and the benefits and challenges for choosing or adopting an efficient inventory control and management system. Findings indicate that SMEs do not invest resources in sophisticated systems; instead, a simple Enterprise Resource Planning (ERP) system or even programs such as Excel or manual inventories are used.

2) Inventory Management Models and Their Effects on Uncertain Demand:

Authors: N. Nemtajela, C. Mbohwa
Published: IEEE 2016
Description: This paper focuses on the use of inventory models to control the material flow and purchased inventory items in manufacturing companies. The objectives of this paper are to assess the effects of demand uncertainty on inventory management and to evaluate the difference on uncertain demand subject to demand controls as determined and the models used. Three inventory management models are studied; the Economic Order Quantity (EOQ), the Activity-Based Costing (ABC), and Just-in-time (JIT). The paper was descriptive in nature and was conducted through the use of quantitative research methods. Survey questionnaire was compiled to gather primary data from five FMCG companies in manufacturing organizations. Survey data of 255 respondents from FMCG manufacturing
companies were used in the analysis.

3) Development of Inventory management System:

Authors: Yang Fan
Published: IEEE 2010
Description: Agent technology into domestic storage management and uses the autonomy, reactivity and sociality of Agent to realize the seamless connection among

enterprises by defining interaction and cooperation mechanisms among different Agents, thereby achieving the aim of reducing and even eliminating inventory, so it is a feasible thought and method for enterprises to realize effective storage management.
This paper mainly designs a storage management system model describing the main agent co-operation process of the system.

4) Implementation of Inventory Analysis Tool for Optimization and Policy Selection:

Authors: Siong Sheng Chin, Edmund Chan, Terence Yeo
Published: IEEE 2008
Description: This paper serves to describe the development and application of a web based, low cost, user friendly Inventory Analysis Tool for stock availability optimization and enhanced delivery performance. The inventory optimization attempts to find dynamically the best inventory policy and safety stock for Stock Keeping Units with independent demands. The analysis is based on supply and demand data, which includes forecast variability and measurements. Important supply chain parameters are modeled and estimated with graphical visualization to identify potential opportunities for improvement. The tool gathers all historical and up-to-date information to effectively track the replenishment level, safety stock level and reorder the level of finished goods within minutes. A case study from National Heart Center Singapore on the use of the tool is presented. The results should encourage more inventory managers to use the tool to lower inventory dollar level and put forecasting errors in check and control.

5) Comprehensive Analysis on Intelligent Retail Management system using classification techniques:

Authors: Phanindra Kakumanu, Saiteja Mothe, Ravi Kumar Tata, Arpita Roy
Published: IEEE 2020
Description: Retailers/Businessmen search for quick benefits with fewer speculations. This paper focuses on structuring an application to yield more profits for retailers by utilizing Machine Learning. By considering the properties such as the spot of retail, the season of retail, the impact of season on the product(s), and many more to produce a yield where the product(s) can gain benefits for the retailers/business people. By knowing the proper item to the right season and spot, it benefits the retailers to purchase the required item through the application. By utilizing Machine Learning it helps to discover the "Pace of Recommendation (exactness)." Through that precision, finding whether the item is best for that season to sell or for that spot to sell, the retailer can acquire benefits as indicated by the season and item. At last, the need to consolidate prescient qualities and prescriptive qualities by accepting the perceptive conditions as a contribution to the further calculations, and the application gives the rate of suggestion for the recommended item. This paper focuses on using models Rpart, Naïve Bayes, and ID3 Algorithm.

6) Big Data Analytics: Enterprise Resource Planning

Authors: Mr Dhananjaya Kumar

Published: IRJET 2018
Description: ERP system is the source of planning of Enterprise Resource Organization and it is an integrated application software
solution offered by a vendor to support the seamless integration of information or data flows through an organization.
It is provided as a package comprising different modules, such as product management, quick billing details,finance or accounting,human resources management, supply product chain and customer information. ERP system process implementation is mainly the lengthy process and completely more complex resulting in many cases of unsuccessful implementation which have negatively impacted on the performance of an organization's business and up to 90% of implementations did not achieve all the desired benefits. A majority of ERP problems are discovered in the last stage of the ERP life cycle known as the post-implementation phase or the after-go-live phase. Much research has been undertaken in relation to the critical success factors of ERP implementation in developed countries whereas research on problems encountered in the ERP post-implementation phase are very limited in
developed across the countries. So overcome all the problem we are find the solution which is MapReduce technology in hadoop
System.

7) Inventory management in retail industry - Application of big data analytics:

Authors: Hien Vu
Published: Research gate 2018
Description: The report articulates the core problem of inventory management is the trade-off between shortage cost and overage costs. Again, the “performance frontier” graph indicates a pragmatic solution is introducing innovation to shift the efficiency curve. In this context, that innovation is BDA. The report finds the prospects of integrating BDA in the conventional inventory management techniques and promoting the viability and appropriateness of these models in the big-data era. However, the limitations of BDA underlie data challenges, processing challenges and management challenges. Finally, the connection between BDA and Traditional operation concepts are presented with insightful lessons from the personal perspective.

The main objective of this report is to explore inventory management practices in the retail industry. In the second part, the report summarizes the literature review of inventory management models. Then the center trade-off in inventory management is presented in the third part, together with the necessity of applying BDA to understand BD. The fourth part will highlight the connection between BDA and operation concepts. The fifth part will discuss the major contribution of BDA in inventory management from the personal viewpoint. Finally, the conclusion will summarize the attempts to answer the question of BDA application in inventory management and note limitations on BDA implementation.

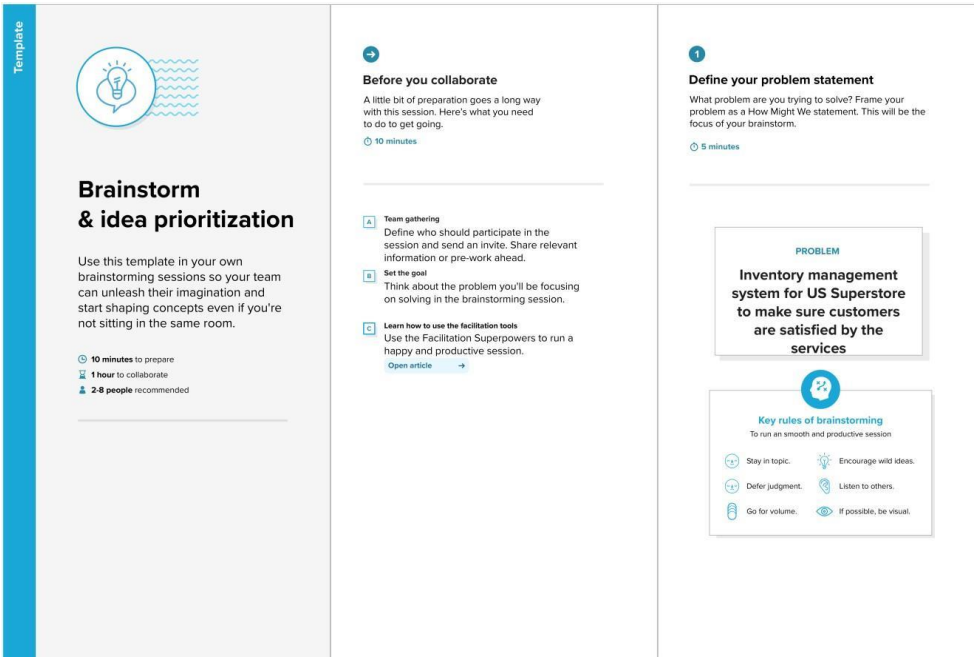
2.3.PROBLEM STATEMENT DEFINITION:

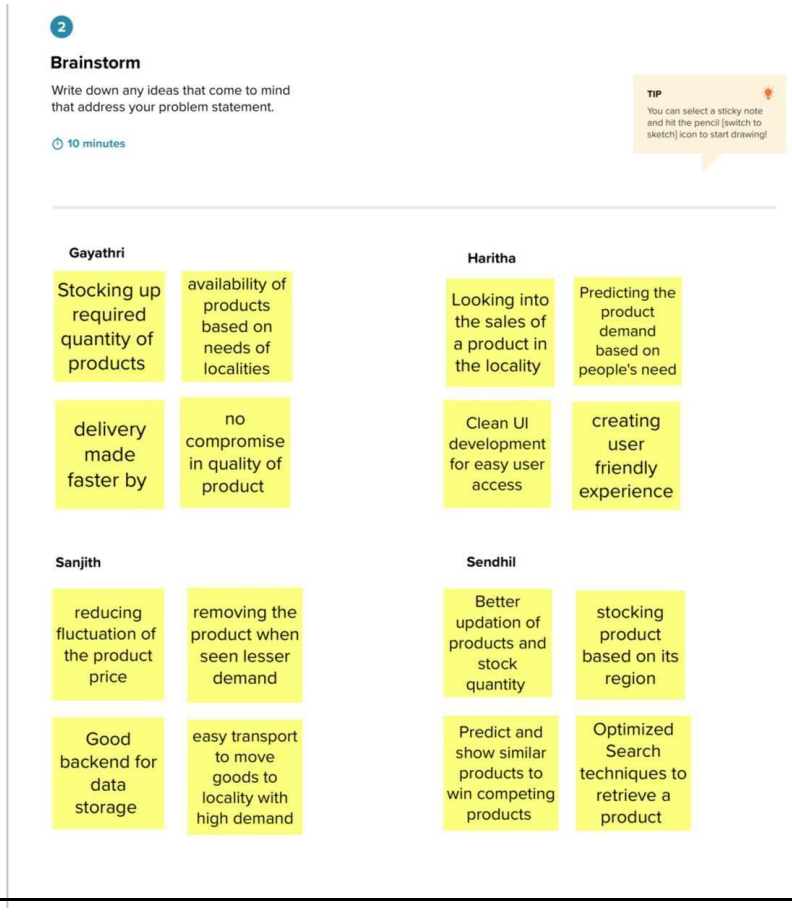
A consumer, residing in San Francisco, wants to buy accessories like a mouse, pendrive, etc. But he/she finds out that availability of accessories from the US Superstore is very less to their locality which makes him/her pay much for the shipping rate to buy the product from the farther areas.This makes him/her worried about the expenses.This happens to several customers in the locality.

3.IDEATION AND PROPOSED SOLUTION

3.1.EMPATHY MAP

3.2.IDEATION AND BRAINSTORMING





3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

TIP

Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.

prediction and analysis

predicting the demand in a locality and accordingly deciding the stock inventory to avoid losses

service

In case of area with high demand, providing exclusive services in order to increase sales in the area

feature

Depending on its demand, the products availability should be varied depending on its region and provide high features and discounts to increase cutomers

management and maintanence

maintaining the market in the locality and competing with the other products ; managing the stocks to yield good profit

4

Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes

After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

Quick add-ons

A **Share the mural**
Share a view link to the mural with stakeholders to keep them in the loop about the outcomes of the session.

B **Export the mural**
Export a copy of the mural as a PNG or PDF to attach to emails, include in slides, or save in your drive.

Keep moving forward

Strategy blueprint
Define the components of a new idea or strategy.
Open the template →

Customer experience journey map
Understand customer needs, motivations, and obstacles for an experience.
Open the template →

Strengths, weaknesses, opportunities & threats
Identify strengths, weaknesses, opportunities, and threats (SWOT) to develop a plan.
Open the template →

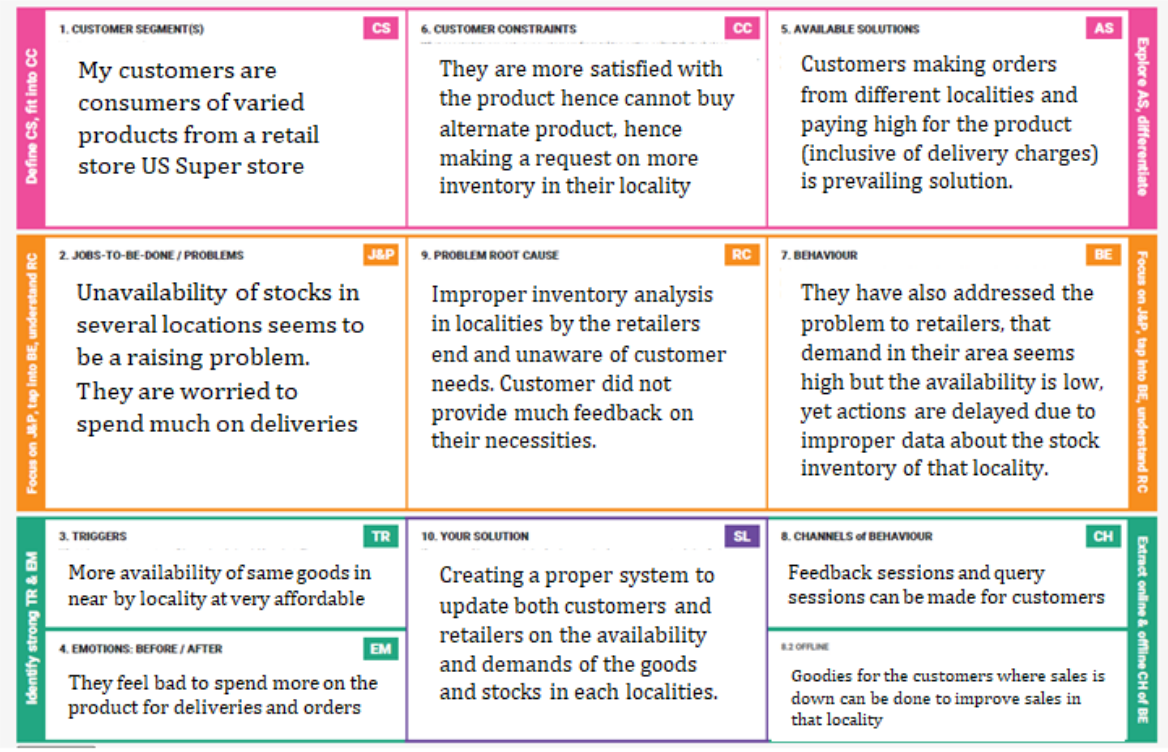
Share template feedback

3.3.PROPOSED SOLUTION

S. No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	▪ Retailers facing issues in tracking the stock stores/bought for each and every locality. Since it is necessary to make customers feel satisfied ▪ Customers are unsatisfied due to less availability or unavailability of the required stock in their area ▪ Retailers are unaware of the fact that which locality requires how much amount of the stock inventory to be done. ▪ So it is necessary for the retailer to make sure to keep a track on the stock inventory in localities and to compete with the other products and to satisfy their customers
2.	Idea / Solution description	▪ Create a system to keep track on the stocks in every locality, though maintaining of databases seem to be overhead process at retailers end ▪ Providing facilities for customers to know the availability of their every product in nearby stores. ▪ In case of emptied stock or out of stock, then retailers are to be notified by some systems. So based on demand analysis of that locality, retailers can refill stocks/goods in that area.
3.	Novelty / Uniqueness	▪ Creating an alert/notifying system to alert both customer when their required stock/product/good is available in their locality and retailers when the stock goes out of order in the stores. ▪ Creating a system to notify the retailers to give discounts to their products in the locality where the sales is less and to provide higher stocks in the locality where demand is high, and occasionally increase the cost. ▪ New arrivals can be brought in to retails of the locality where demand is high and people are more interested in their products

4.	Social Impact / Customer Satisfaction	▪ Due to unavailability of the goods/stocks in the locality , people tend to spend much on the delivery charges. So based on their queries retailers can invest at least little stock near their locality. ▪ Feedback can be got from the customers regarding the level of satisfaction on their product, so that retailers can improve their customer service.
5.	Business Model (Revenue Model)	▪ Retailers can use deep learning models to analyse the sales in every locality so that they can improvise their investments in required localities ▪ Implementing feedback strategies to avoid losses in areas ▪ Good advertising models to improve sales in the low demand areas
6.	Scalability of the Solution	▪ Day to day updating of the stock and availability to retailers and customers, via server applications make it reliable. ▪ Large data can also be easily analyzed using analytics tools and data interpretation can be well obtained.

3.4.SOLUTION FIT



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn Registration through the website(created for specific store)
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Login	Log in with User id /Username and Password

FR-4	Profile Update	Update the user credentials and contact details
FR -5	Uploading data	After collecting all user details and product details, uploading the details
FR - 6	Data Management	Based on the products availability and price updating the data frequently using deep learning to avoid miscommunication
FR - 7	Recommendation	The user will request the item. Get the item recommendations.
FR - 8	Notification	Based on the availability notifying the retailers as well as customers
FR - 9	Payment	For customers there are many types of secure billing will be prepaid as debit or credit card, postpaid as after delivery, cash on delivery
FR - 10	Review	Customers can give their ratings and review.

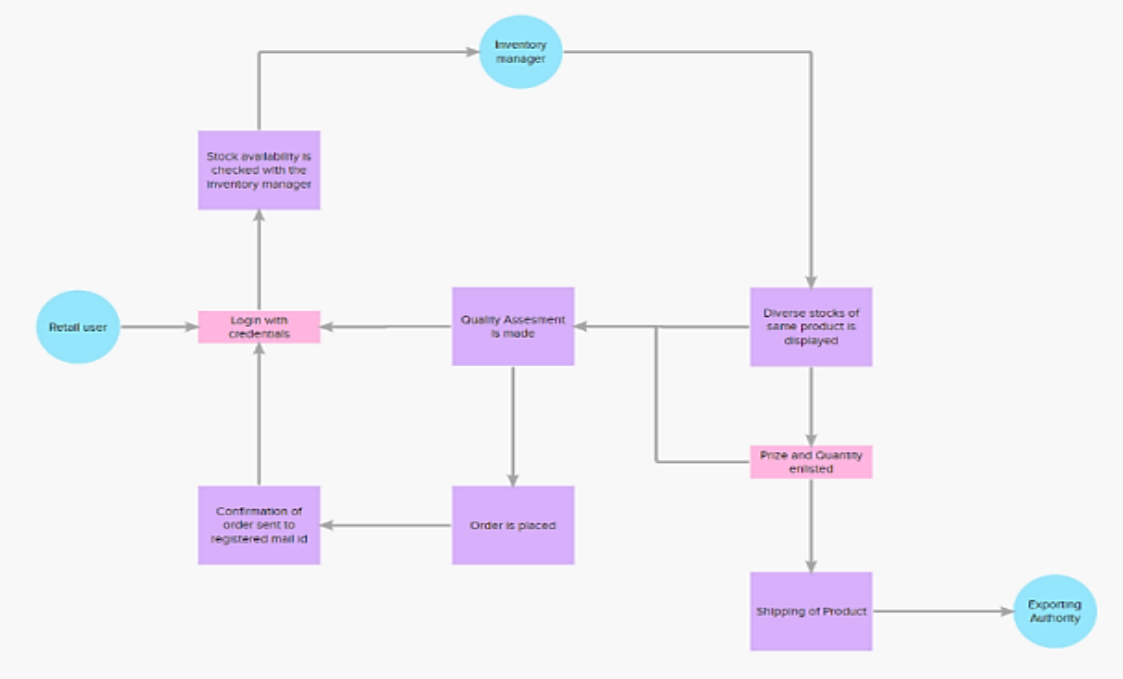
4.2 NON – FUNCTIONAL REQUIREMENTS

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	The system should have user-friendly interface. So that each user can customize for their needs. They are more likely to have the right amount of inventory for tracing.
NFR-2	Security	This can be used only by users who have their correct login credentials
NFR-3	Reliability	It avoids under or over stocking..Ensure accurate inventory valuation.
NFR-4	Performance	The system provides the analysis quickly since it uses visualization and prediction techniques. This model can predict deadstock as well as highly profitable stock

NFR-5	Availability	Users should be able to access the system across the world. To achieve this needed a reliable network and servers. This system can give retailers visibility to stock levels etc
NFR-6	Scalability	The system should support the dataset which varies according to the users. Many users can access it at the same time without any issues

5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAM



Retail users initially log in with their credentials which are registered already.

They will check the availability of stocks based on that, products will be purchased by the retailer.

Before purchasing retailer will do quality assessment after that order will be placed.

Once the order is placed confirmation will be sent to the retailer through email.

5.2 SOLUTION AND TECHNICAL ARCHITECTURE

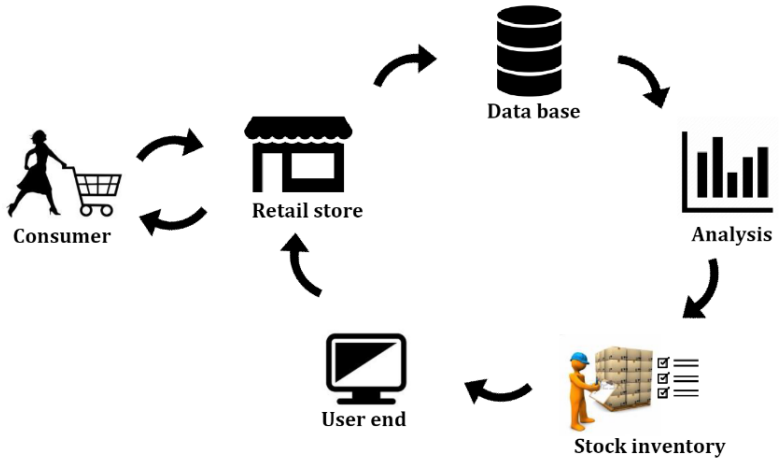


Table 1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	User may have direct interaction with retailer or through user-end applications to find out the stock availability	HTML, CSS, JavaScript, HTTP
2.	Application Logic-1	The logic for a process in theapplication	IBM Cognos analytics
3.	Database	Data Type, Configurations etc.	Excel, xls formatted dataset
4.	Cloud Database	Database Service onCloud	IBM cloud service
5.	File Storage	File storage requirements	IBM cloud storage
6.	Analysis	Analysing the previousstock supply to the particular retail store	IBM Cognos
7.	Data visualization	Understanding the needfor changes in stock supply to the store	Cognos IBM
8.	Update	Change the stock inventory to the store	User end contacts

Table 2: Application Characteristics

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	IBM Cognos analytics, cloud

2.	Security Implementations	use of firewalls etc.	Encryption algorithms
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Web server – HTTP Application server – python Database server –IBM cloud
4.	Availability	For all consumers of theretail store	IBM cloud
5.	Performance	Retail store owners get to know how to managetheir stock inventory toavoid losses and increase profit in demanded areas	ML algorithms , but usually IBM cognosfor predictions

5.3 USER STORIES

User Type	Functi-onal Requir ement (Epic)	User Stoy Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (mobile user)	Registr ation	USN-1	As a user, I can register forthe web application by entering my email, password, and confirming my password.	I can access my account/das hboard	High	Sprint-1
		USN-2	As a user, after completingthe registration I will receive a confirmation email once I have registered for the web application	I can receive a confirmation email &click confirm	High	Sprint-1
		USN-3	As a user, I can	I can register & access the	Low	Sprint-2

			register for the web application through LinkedIn	dashboard with LinkedIn Login		
		USN-4	As a user, I can register for the web application through my Google account	I can register & access the dashboard with a Gmail login	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password after installing the web application	I can access the dashboard by login into the application	High	Sprint-1
	Dashboard	USN-6	As a user, I can view the charts and graphs representation of the dataset and the information shown in the dashboard	I can analyze the stocks in my retail store.	High	Sprint-1
customer (Web user)		USN-1	As a user, I can register for the web application by entering my email, and password, and confirming my password.	I can access my account/dashboard	High	Sprint-1
Customer Care Executive		USN-2	As a user, after completing the registration I will receive a confirmation email once I have registered for the web application	I can receive confirmation email & click confirm	High	Sprint-1
Administrator		USN-3	As a user, I can register for the web application through LinkedIn	I can register & access he dashboard with LinkedIn Login	Low	Sprint-2

		USN-4	As a user, I can register for the web application through my Google account	I can register & access the dashboard with Gmail login	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the web application by entering email & password after installing the application.	I can access the dashboard by login into the application	High	Sprint-1
	Dashboard	USN-6	As a user, I can view the charts and graphs representation of the dataset and the information shown in the dashboard.	I can analyze the stocks in my retail store.	High	Sprint-1
Customer Care Executive		CCE-1	As a customer care executive, I will always be available for the interaction with the customer to clarify the queries.	An executive will analyze the customer complaints,rectify their problems	High	Sprint-2
Administrator		ADMIN-1	As an administrator, I will manage backup and recovery, data modeling and design, distributed computing, database system, and a data security	Administrator can evaluate, design, review and implementing a data,they are also responsible for updating and maintaining the data.	High	Sprint-2

6.PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

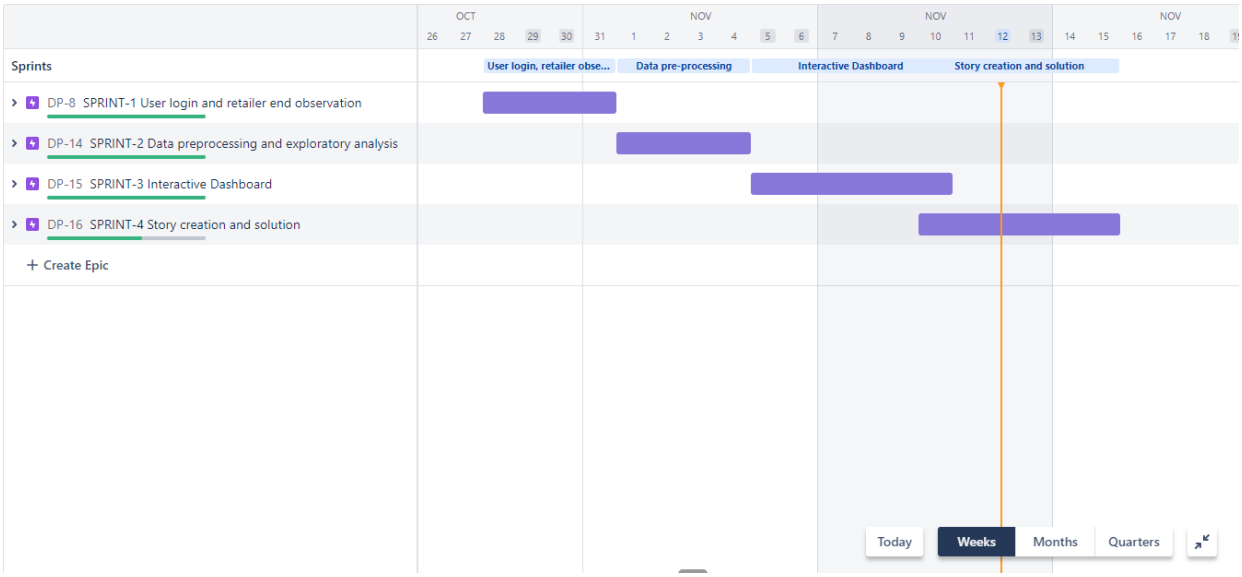
Sprint	Milestone
Sprint 1	**USER LOGIN AND RETAILER OBSERVATION:** • User can register for the application by entering name and locality • User can add his/her query to the US super store consideration • User can suggest some remedy measure • Retailer can look into the query of the customer/user and start taking action • Retailer can take data base and clean it before analysing , • He/She should fill in the missing values in case of any dataset available
Sprint 2	**DATA PREPROCESSING AND EXPLORATORY ANALYSIS** • Removing the unwanted data and add necessary columns for processing • Masking of private or sensitive data. • Create new columns in case needed to split up the dataset to work • Remove nil entry data and make sure to maintain them properly in future • Format data to a standardized pattern
Sprint 3	**INTERACTIVE DASHBOARD** • Analysing basic metrices • Learning IBM Cognos functionalities • Data visualization basics • Correlation between variables • Year-wise profit using a line graph • Year-wise quantity of utilities using line graph • Top 10 sales by year using line graph • Monthly sales using Tree Map • Monthly profit by pie chart
Sprint 4	**STORY CREATION AND SOLUTION** • Dashboard creation • Summary cards of total profit, sales, sub-categories and localities • Understanding the demand of the customer correlated with an analysed data set • Generate remedy measures for the customer’s query based on the available solution • Generate a final report for future use, for both retailers and the customers access.

6.2 SPRINT DELIVERY SCHEDULE

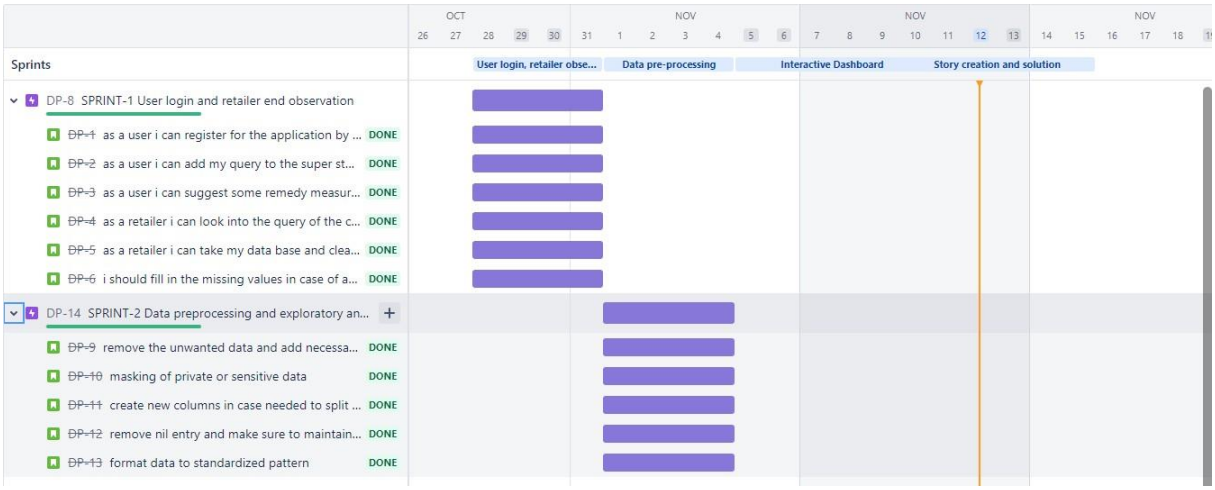
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint 1	13	4 Days	28 Oct 2022	31 Oct 2022	13	31 Oct 2022
Sprint 2	14	4 Days	01 Nov 2022	04 Nov 2022	14	04 Nov 2022
Sprint 3	21	5 Days	05 Nov 2022	10 Nov 2022	21	10 Nov 2022
Sprint 4	25	5 Days	10 Nov 2022	15 Nov 2022	25	15 Nov 2022

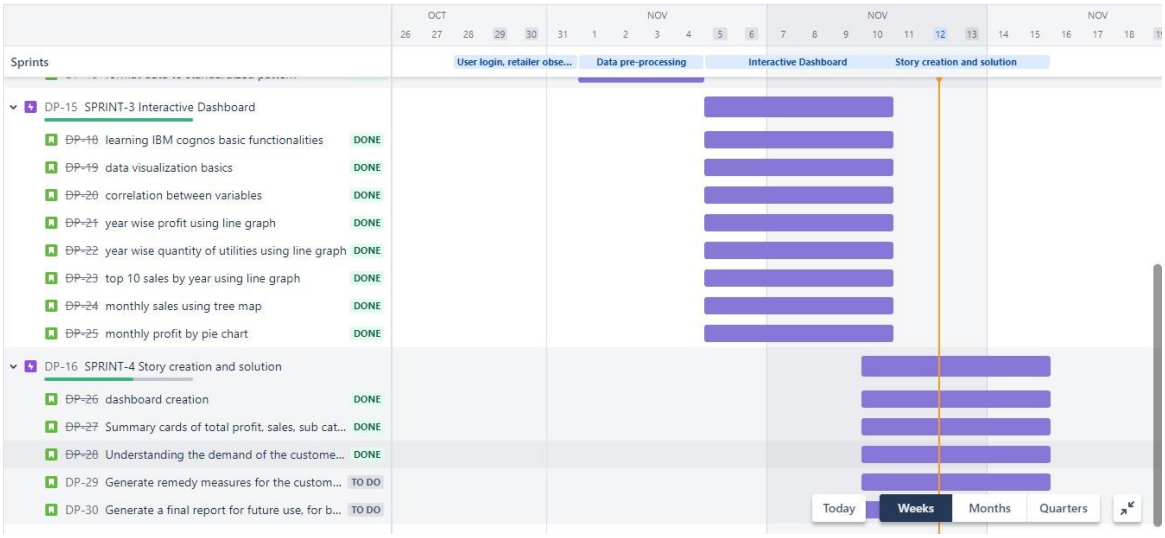
6.3 REPORTS FROM GIRA

Sprint designing

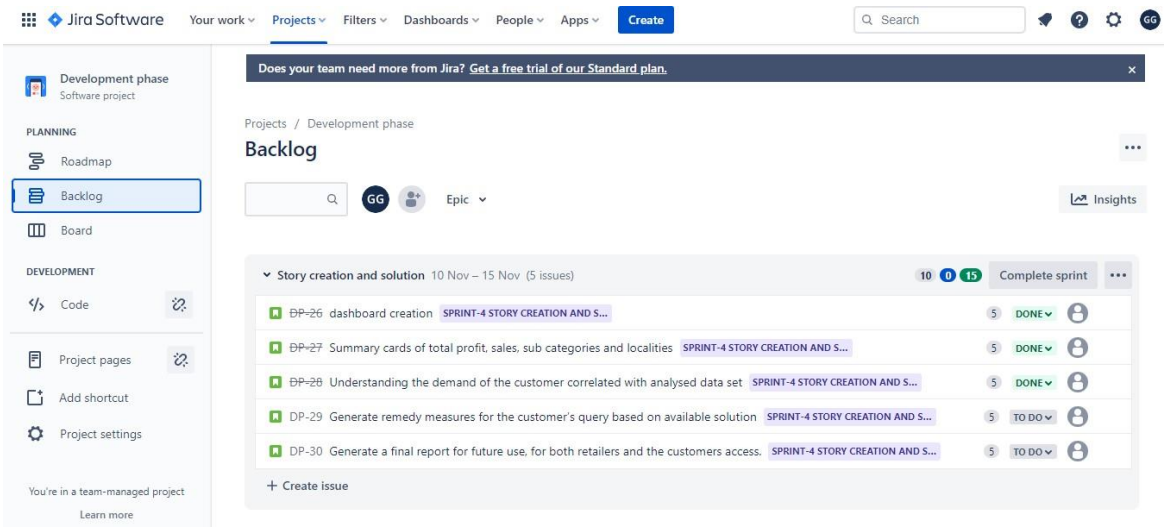


Expanded view

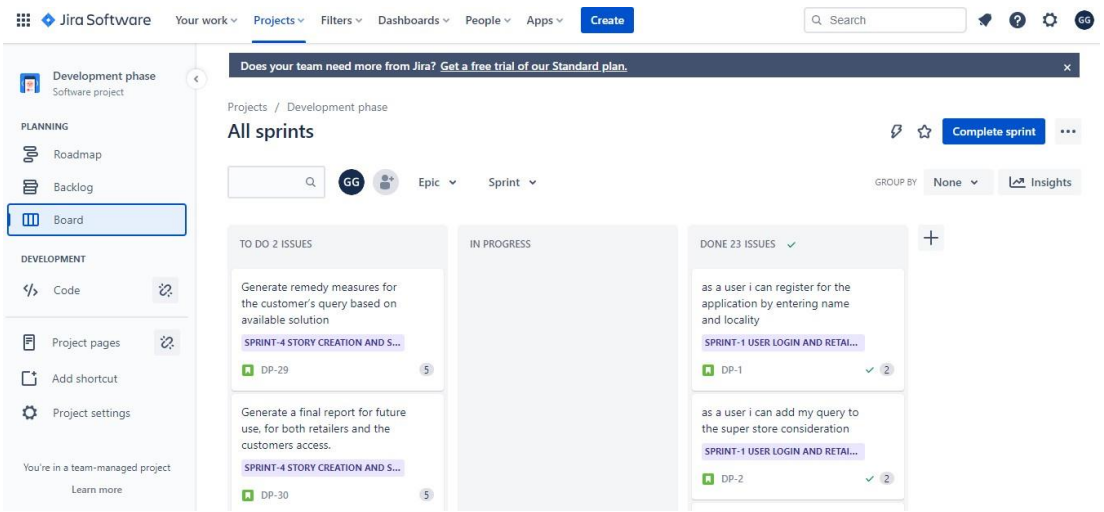




Current Backlogs (completed)



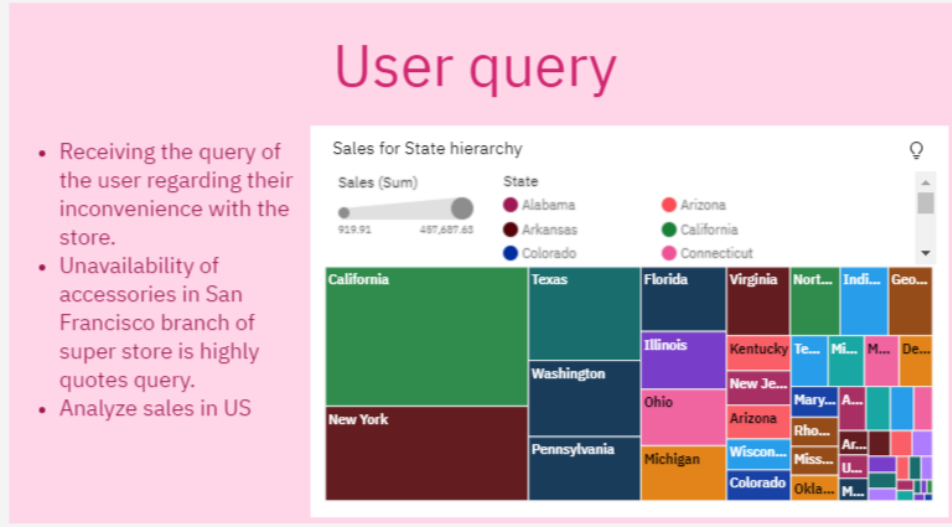
Board



7.CODING AND SOLUTIONING

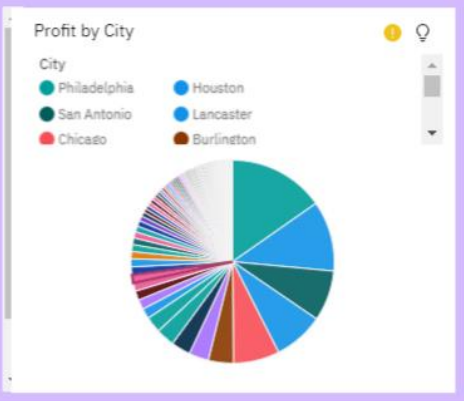
7.1 FEATURE 1 – Initial works

Story creation and solution working :



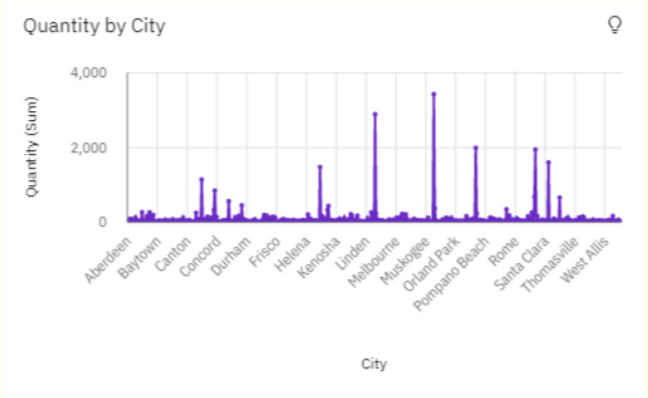
DATA PREPROCESSING

- Obtaining super store dataset and cleaning it before processing
- Creating new columns for required processing of the query



DASHBOARD

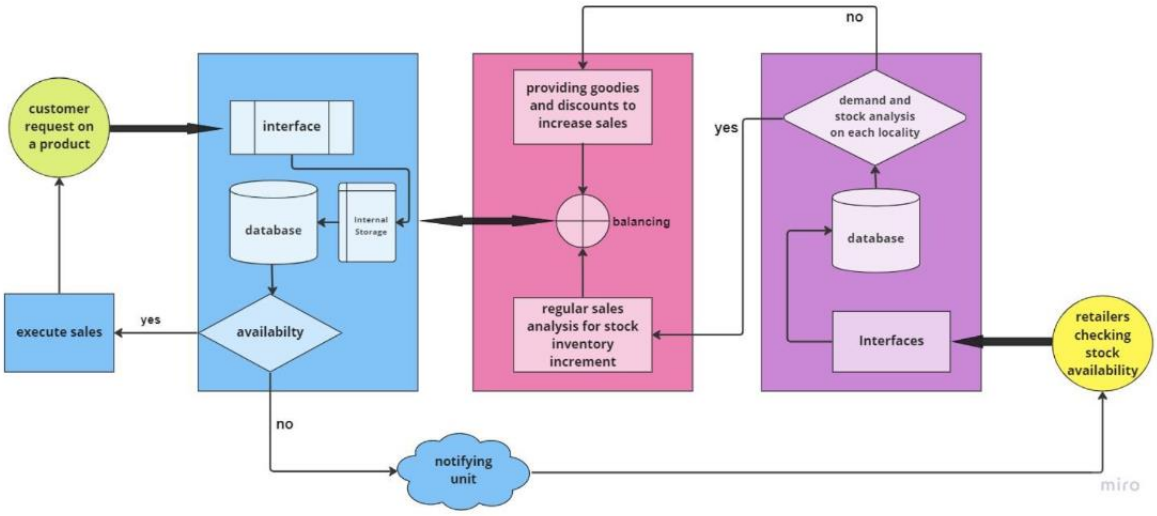
- creating a dashboard to analyze and predict the stock inventory on each locality
- Based on analysis, inventory can be managed.



Report

- The sales and profit of the locality stated by customer is analyzed and actions are taken
- The future data are also predicted using deep learning analysis.

Solution architecture



Data set cleaning : (before)

```
In [2]: import pandas as pd
        import numpy as np
        store = pd.read_csv('prob1-US-Superstore-data.csv')
        store.head()
```

Out[2]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Postal Code	Region	Product ID	Category	Sub-Category
0	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-BO-10001798	Furniture	Bookcases
1	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-CH-10000454	Furniture	Chairs
2	3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	...	90036	West	OFF-LA-10000240	Office Supplies	Labels

(after)

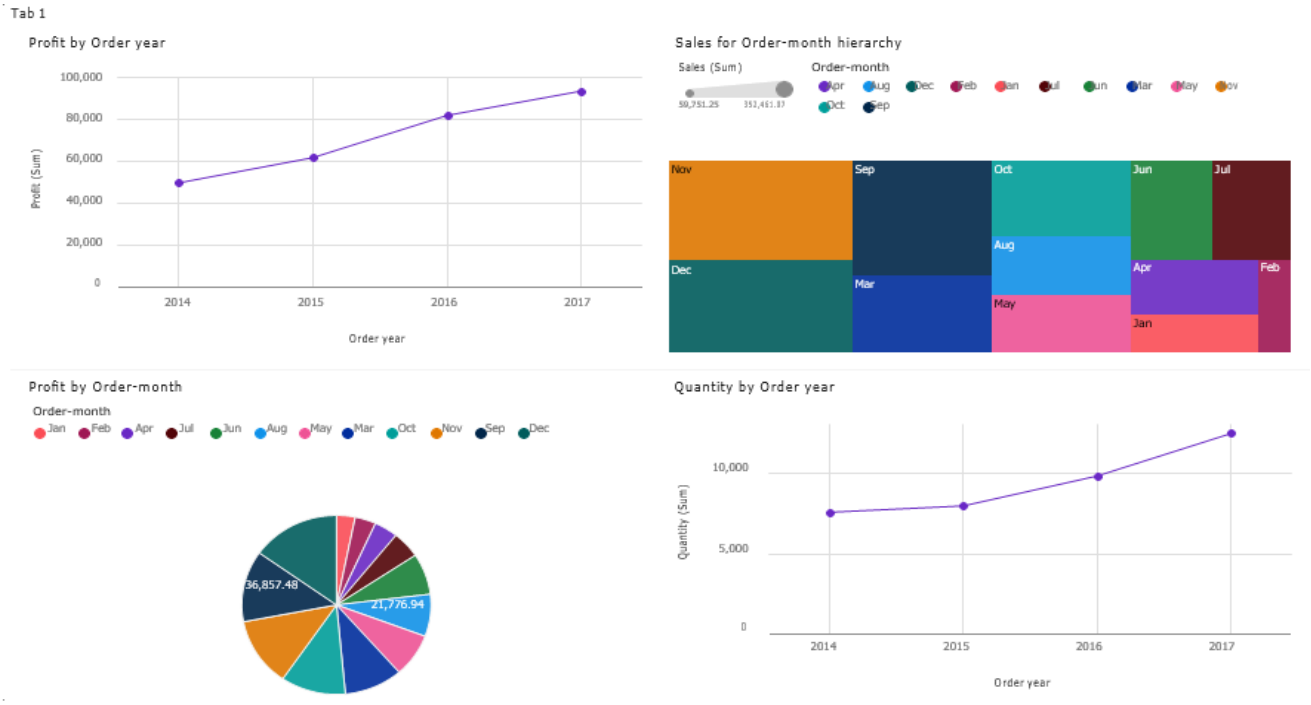
```
In [14]: new_cols = ["Row ID","Order ID","Order year","Order month","Order day","Ship year","Ship month","Ship day","Ship Mode","Customer ID","Customer Name","
         store=store.reindex(columns=new_cols)
         store.head()
```

Out[14]:

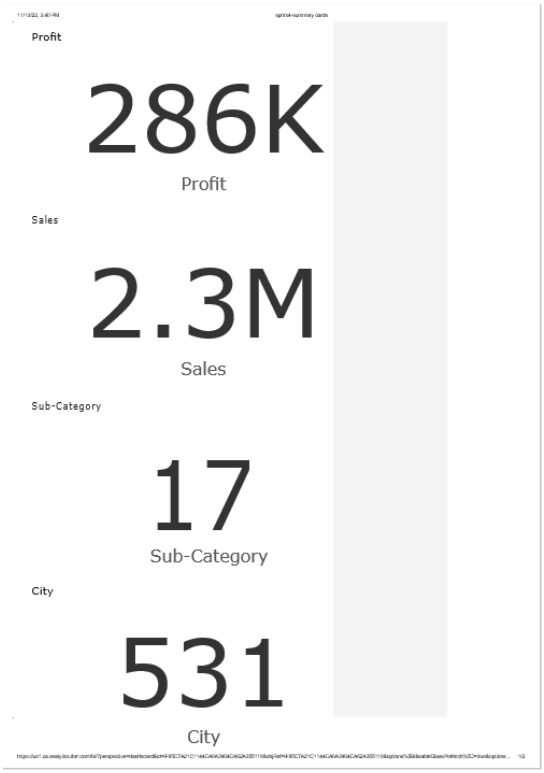
	Row ID	Order ID	Order year	Order month	Order day	Ship year	Ship month	Ship day	Ship Mode	Customer ID	...	Country	City	State	Category	Sub-Category	Product Name	Quantity	Sales	Discour
0	1	CA-2016-152156	2016	11	8	2016	11	11	Second Class	CG-12520	...	United States	Henderson	Kentucky	Furniture	Bookcases	Bush Somerset Collection Bookcase	2	261.9600	0.0
1	2	CA-2016-152156	2016	11	8	2016	11	11	Second Class	CG-12520	...	United States	Henderson	Kentucky	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	3	731.9400	0.0
2	3	CA-2016-138688	2016	6	12	2016	6	16	Second Class	DV-13045	...	United States	Los Angeles	California	Office Supplies	Labels	Self-Adhesive Address Labels for Typewriters b...	2	14.6200	0.0

7.2 FEATURE - Dashboard, Summary Cards and analysis

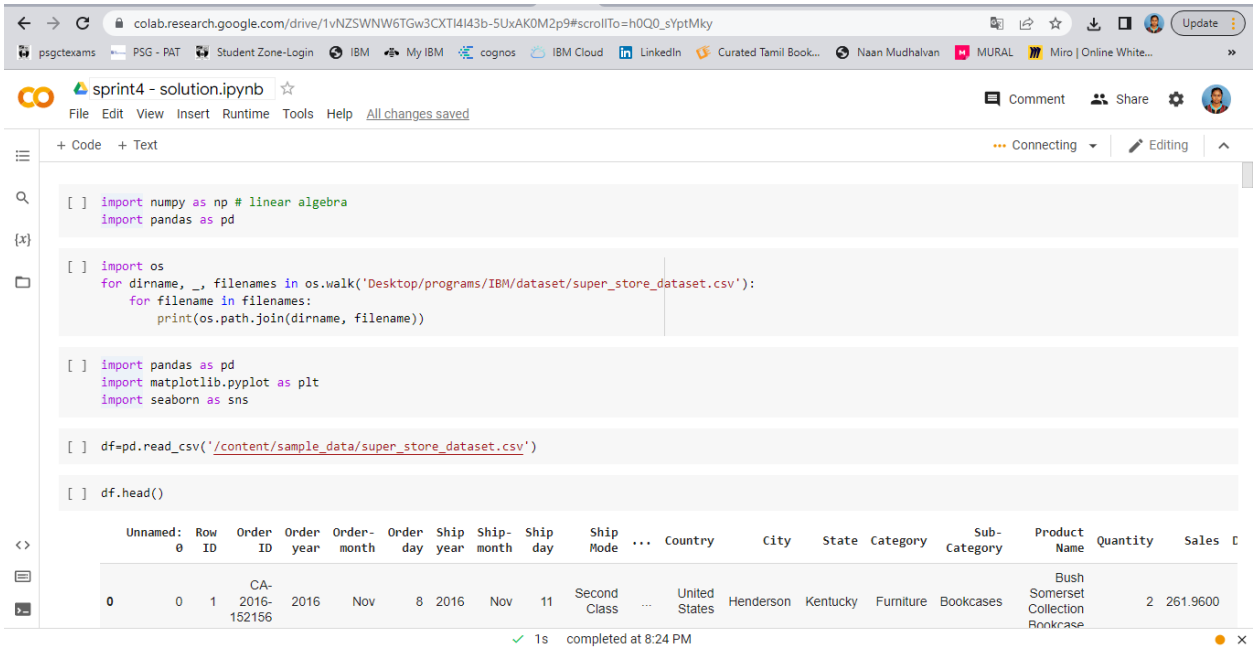
Dashboard



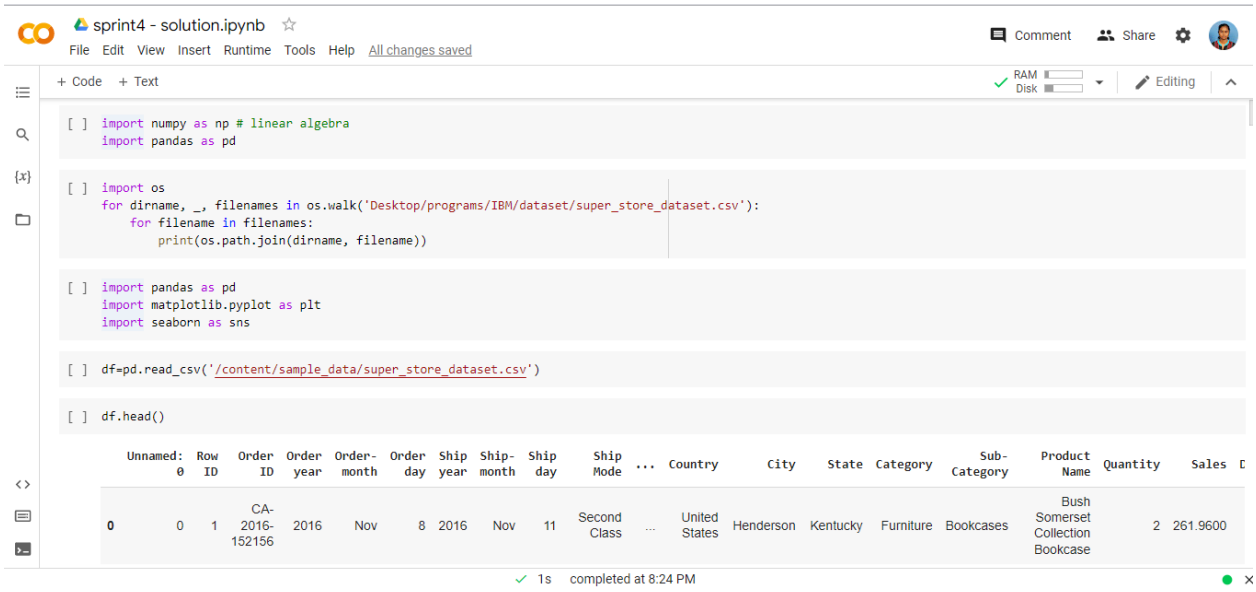
Summary cards :



Analysis using Colab interactive platform :



Analysis using cognos analytics :

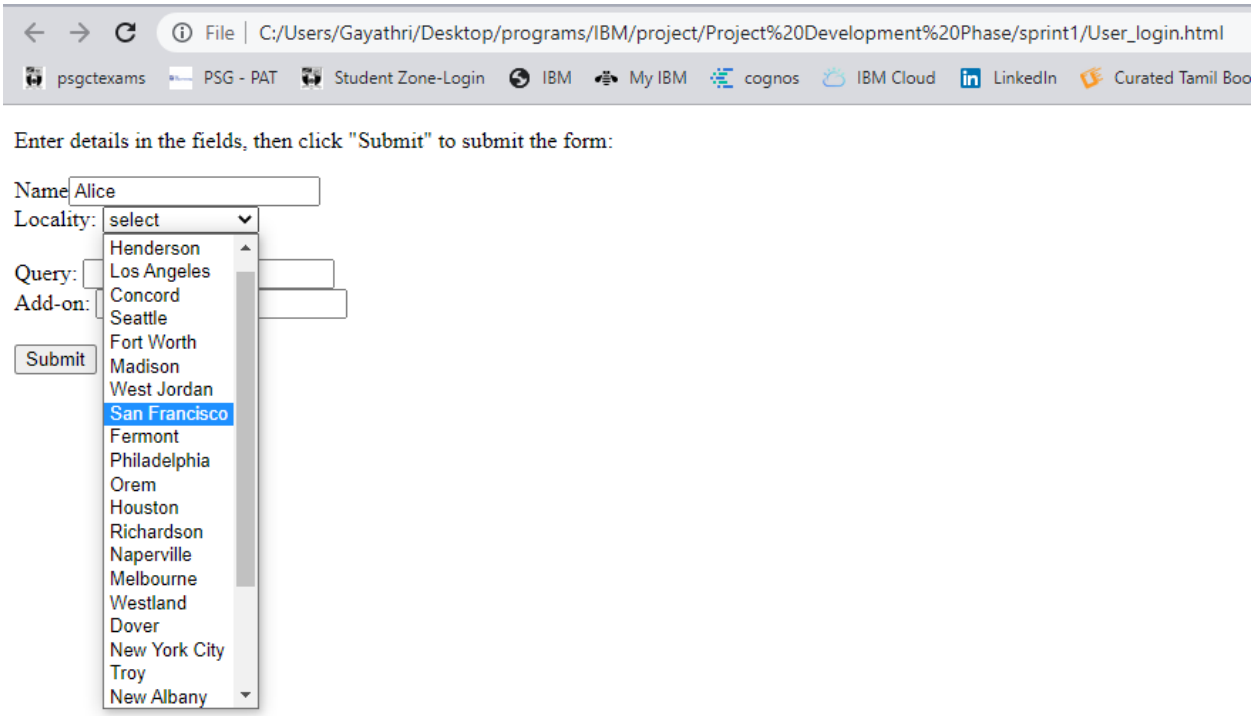


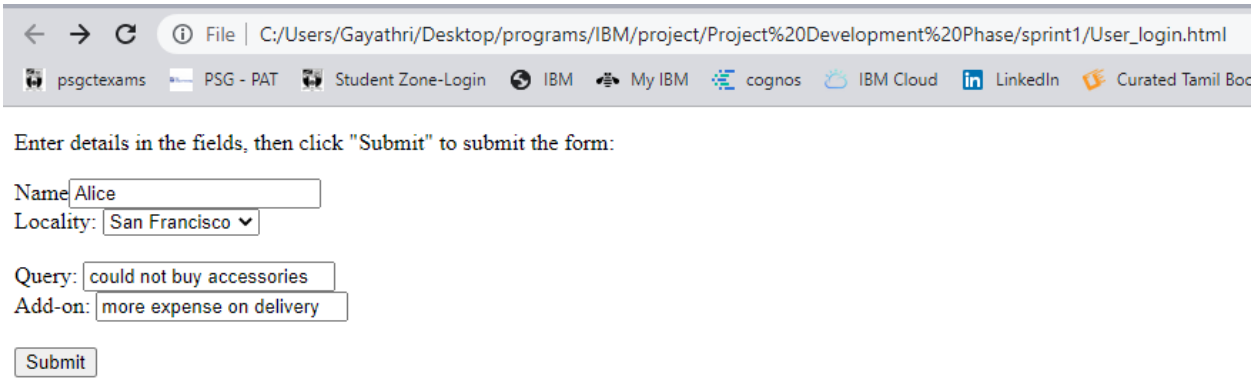
7.3 FEATURE 3 : HTML CODE (user interface)

```
<!DOCTYPE html>
<html>
<body>
<p>Enter details in the fields, then click "Submit" to submit the form:</p>
<form id="frm1" action="/action_page.php">
  Name<input type="text" name="name"><br>
  Locality:
  <select id="locality">
      <option> select </option>
      <option> Henderson </option>
      <option> Los Angeles </option>
      <option> Concord </option>
      <option> Seattle </option>
      <option> Fort Worth </option>
      <option> Madison </option>
      <option> West Jordan </option>
      <option> San Francisco </option>
      <option> Fermont </option>
      <option> Philadelphia </option>
      <option> Orem </option>
      <option> Houston </option>
      <option> Richardson </option>
      <option> Naperville </option>
      <option> Melbourne </option>
      <option> Westland </option>
      <option> Dover </option>
      <option> New York City </option>
      <option> Troy </option>
      <option> New Albany </option>
      <option> Eagan </option>
      <option> Chicago </option>
      <option> Coulumbia </option>
      <option> Rochester </option>
      <option> Portland </option>
      <option> Other </option>
</select>
<br><br>
  Query: <input type="text" name="query"><br>
  Add-on: <input type="text" name="addon"><br><br>
  <input type="button" onclick="myFunction()" value="Submit">
</form>
<script>
function myFunction() {
  document.getElementById("frm1").submit();
}
</script>
</body>
</html>
```

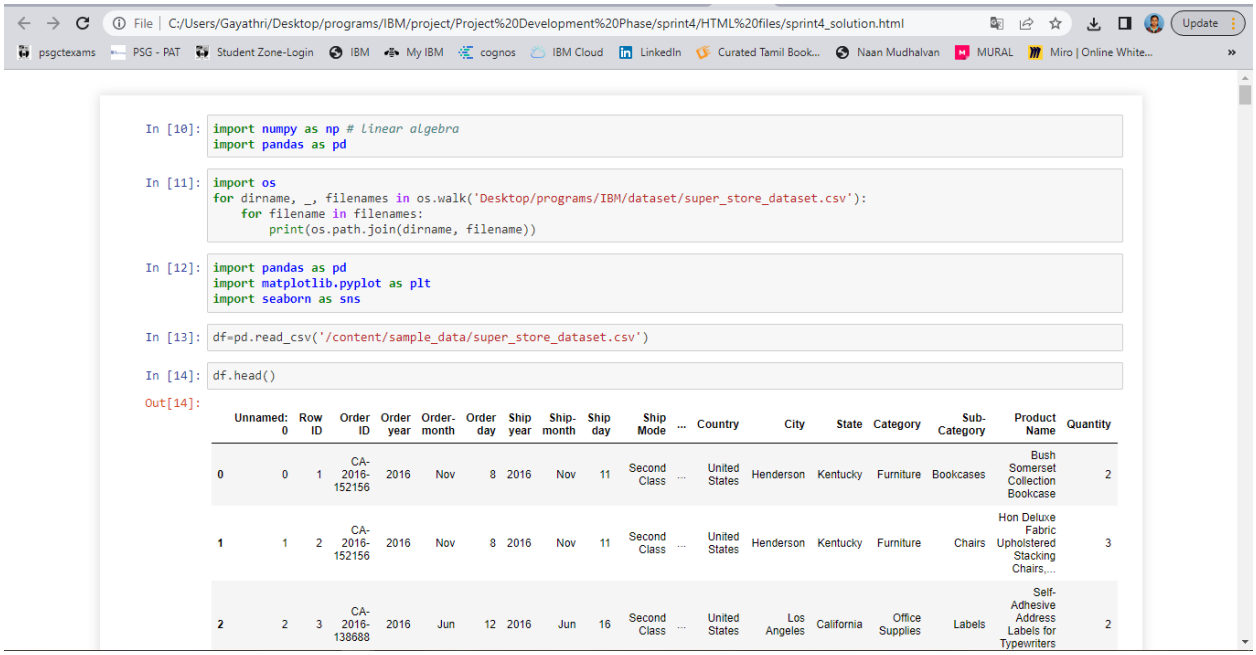
Webpage :


File | C:/Users/Gayathri/Desktop/programs/IBM/project/Project%20Development%20Phase/sprint1/User_login.html
psgctexams
PSG - PAT
Student Zone-Login
IBM
My IBM
cognos
IBM Cloud
LinkedIn
Curated Tamil Book...
Enter details in the fields, then click "Submit" to submit the form:
Name
Locality: select
Query:
Add-on:
Submit


File | C:/Users/Gayathri/Desktop/programs/IBM/project/Project%20Development%20Phase/sprint1/User_login.html
psgctexams
PSG - PAT
Student Zone-Login
IBM
My IBM
cognos
IBM Cloud
LinkedIn
Curated Tamil Boo
Enter details in the fields, then click "Submit" to submit the form:
Name Alice
Locality: select
Query:
Add-on:
Submit
Henderson
Los Angeles
Concord
Seattle
Fort Worth
Madison
West Jordan
San Francisco
Fermont
Philadelphia
Orem
Houston
Richardson
Naperville
Melbourne
Westland
Dover
New York City
Troy
New Albany


File | C:/Users/Gayathri/Desktop/programs/IBM/project/Project%20Development%20Phase/sprint1/User_login.html
psgctexams
PSG - PAT
Student Zone-Login
IBM
My IBM
cognos
IBM Cloud
LinkedIn
Curated Tamil Boo
Enter details in the fields, then click "Submit" to submit the form:
Name Alice
Locality: San Francisco
Query: could not buy accessories
Add-on: more expense on delivery
Submit

HTML report generated :



```
In [10]: import numpy as np # linear algebra
         import pandas as pd

In [11]: import os
         for dirname, _, filenames in os.walk('Desktop/programs/IBM/dataset/super_store_dataset.csv'):
             for filename in filenames:
                 print(os.path.join(dirname, filename))

In [12]: import pandas as pd
         import matplotlib.pyplot as plt
         import seaborn as sns

In [13]: df=pd.read_csv('/content/sample_data/super_store_dataset.csv')

In [14]: df.head()
```

Out[14]:

	Unnamed: 0	Row ID	Order ID	Order year	Order-month	Order day	Ship year	Ship-month	Ship day	Ship Mode	...	Country	City	State	Category	Sub-Category	Product Name	Quantity
0	0	1	CA-2016-152156	2016	Nov	8	2016	Nov	11	Second Class	...	United States	Henderson	Kentucky	Furniture	Bookcases	Bush Somerset Collection Bookcase	2
1	1	2	CA-2016-152156	2016	Nov	8	2016	Nov	11	Second Class	...	United States	Henderson	Kentucky	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	3
2	2	3	CA-2016-138688	2016	Jun	12	2016	Jun	16	Second Class	...	United States	Los Angeles	California	Office Supplies	Labels	Self-Adhesive Address Labels for Typewriters	2

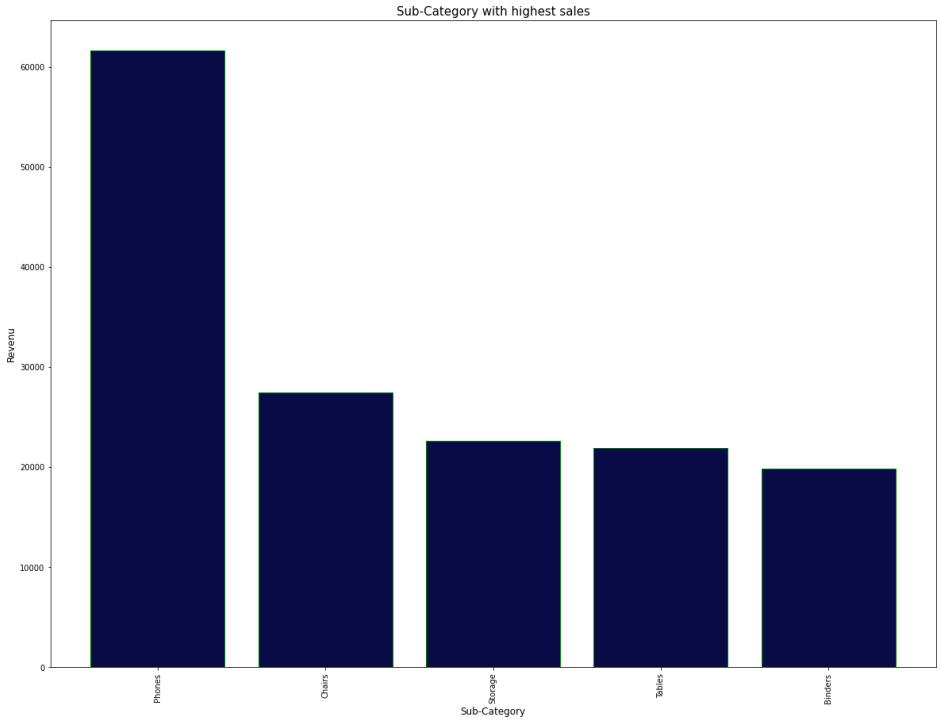
7.4 FEATURE 4 - SOLUTION AND INFERENCE

The query of the customer on sub-category :

It can be noted that as per the request of several customers, the dataset in analysed, But it becomes clear that people of the locality do not request more onaccessories in sub categories. It is seen that other sub-categories are more preferred in purchase and yield a greater sales and revenue.

```
top_cate=df.groupby(['Sub-Category']).sum().sort_values('Sales',ascending=False).head(5)
top_cate=top_cate[['Sales']].round(2)
top_cate.reset_index(inplace=True)
top_cate
```

	Sub-Category	Sales
0	Phones	330007.05
1	Chairs	328449.10
2	Storage	223843.61
3	Tables	206965.53
4	Binders	203412.73



So we can convey to the customer that, currently the sales and revenue for accessories in the areas of California, particularly in San Francisco is very low. Which makes it much harder to take decisions on investing more for the added inventory to that store. In case of an increase in demand, future steps will be taken to increase sales as requested.

8. TESTING

8.1 Test Cases

					Date	3-Nov-22								
					Team ID	PNT2022TMID12898								
					Project Name	Retail store stock inventory analysis								
					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_OO1	Functional	Home Page	Verify user is able to enter name	none	type in the details	Alice	accepted data print out	Working as expected	Pass	none	yes	-	sanjith , sendhilnathan	
LoginPage_TC_OO1	funotional	Home Page	Verify user is able to enter locality	none	choose the appropriate location from the drop down box and select it	California	selected value displayed from dropdown	Working as expected	pass	none	yes	-	sanjith , sendhilnathan	
LoginPage_TC_OO1	Functional	Home page	Verify user is able enter query	none	type the query you want to convey	unable to make an online purchase due to server error	text data visible in blank box	working as expected	pass	none	yes	-	gayathri , haritha	
LoginPage_TC_OO1	Functional	Login page	Verify user is able to enter add on suggestions	none	type in the add on details	could you please update server and state the reason for error	entered data being visible	working as expected	pass	none	yes	-	gayathri , haritha	
LoginPage_TC_OO2	funotional	output page	the analysed dataset visible for user	data analysis and modeling in colab platform	outputs are clearly visible	results for the query and solution available	query can be solved or no	working as expected	pass	none	yes	-	gayathri , haritha , sanjith , sendhilnathan	

8.2 User acceptance testing : Test case analysis

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

Section	Total Cases	Not Tested	Fail	Pass
Verify user is able to enter name	10	0	0	10
Verify user is able to enter locality	10	0	0	10
Verify user is able enter query	10	0	0	10
Verify user is able to enter add on suggestions	10	0	0	10
the analysed dataset visible for user	10	0	0	10

9. RESULTS :

9.1 PERFORMANCE METRICS :

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visulizations / Graphs : Colab source – 9 Cognos cource – 15
2.	Data Responsiveness	Great and high
3.	Amount Data to Rendered (DB2 Metrics)	9994 rows × 23 columns 9994 entries of data being used
4.	Utilization of Data Filters	• Null values removed • Discretization of certain fields • Splitting the data into multiple columns for easier access and analysis.
5.	Effective User Story	No of Scene Added – 4
6.	Descriptive Reports	No of Visulizations / Graphs – 24

10. ADVANTAGE AND DISADVANTAGE

The proposed model for the problem statement holds good in terms of the analysis point of view. The user can clearly understand the scenario at the retailer end. The analysis being done is to identify the solution to solve out the user's query. The dataset of the store is being analysed well and the results are obtained to its optimum and best. The main advantage lying here is the usage of analytics platform like IBM cognos analytics and Google colab.

The webpage interacts with the user and receives their input, but does not provide any feedback. The results will be observable to the user through the generated report in webpage. The solution is either the query will be rectified or no. But alternate solutions are currently not suggested from retailer end. This can be included in futuristic works.

11. CONCLUSION :

The proposed solution architecture and the designed work holds good to provide apt solution to the user's query. A full detailed report is generated to the user's observation and inferences are obtained to conclude the query. The analysis is done accurately to yield more trustable results. This work helps retailers as well as users to negotiate much in case of continuous queries being raised by consumers. The stock inventory can be very well managed by the retailers by analyzing the dataset obtained from the database. This greatly helps in increasing the profit and reduce losses in business. Also users can be relieved off from their query either by an approved solution from retailer or via an alternate solution.

12. FUTURE SCOPE :

In this current work, it ends with consumer readable report to understand how well their query can be solved. In future work, it could be extended to work on few lined report generation with interactive platform for the consumer's easier access. The alert messages on the updation of inventory stock and product can be done to keep the consumer in track with their query on how much it is under progress or processing.

13. APPENDIX :

Source code :

```
import numpy as np # linear algebra
import pandas as pd
import os
for dirname, _, filenames in os.walk('Desktop/programs/IBM/dataset/super_s
tore_dataset.csv'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv('/content/sample_data/super_store_dataset.csv')
```

```
df.head()
df.tail()
df.sample(10)
df.info()
df.shape
df[df.duplicated()]
df.drop_duplicates()
df.isnull().sum()
df['Order day'] = pd.to_datetime(df['Order day'],format='%d')
df['Ship day'] = pd.to_datetime(df['Ship day'],format='%d')
df.sort_values(by=['Order day'],inplace=True)
df['Ship day'].value_counts()
plt.figure(figsize=(20,15))
df['Ship Mode'].value_counts().plot.pie()
df['Country'].value_counts()
df['City'].value_counts()
plt.figure(figsize=(20,15))
df['City'].value_counts().plot.pie()
plt.figure(figsize=(20,15))
plt.xticks(rotation='vertical')
sns.countplot('State',data=df)
top_states=df.groupby(['State']).sum().sort_values('State',ascending=False
).head(10)
top_states=top_states[['Sales']].round(2)
top_states.reset_index(inplace=True)
top_states
plt.figure(figsize=(20,15))
plt.bar(top_states['State'],top_states['Sales'],color='#6a0dad',edgecolor=
'orange')
plt.xticks(rotation='vertical')
plt.title('States with highest revenu',fontsize=15)
plt.xlabel('States',fontsize=12)
plt.ylabel('Revenu',fontsize=12)
top_cities=df.groupby(['City']).sum().sort_values('Sales',ascending=False)
.head(10)
top_cities=top_cities[['Sales']].round(2)
top_cities.reset_index(inplace=True)
top_cities
plt.figure(figsize=(20,15))
plt.bar(top_cities['City'],top_cities['Sales'],color='#95dee3',edgecolor='
purple')
plt.xticks(rotation='vertical')
plt.title('Cities with highest revenu',fontsize=15)
plt.xlabel('Cities',fontsize=12)
plt.ylabel('Revenu',fontsize=12)
```

```
top_customers=df.groupby(['Customer Name']).sum().sort_values('Sales',asce
nding=False).head(10)
top_customers=top_customers[['Sales']].round(2)
top_customers.reset_index(inplace=True)
top_customers

plt.figure(figsize=(20,15))
plt.bar(top_customers['Customer Name'],top_customers['Sales'],color='#8b00
00',edgecolor='yellow')
plt.xticks(rotation='vertical')
plt.title('Customers with highest revenu',fontsize=15)
plt.xlabel('Customers',fontsize=12)
plt.ylabel('Revenu',fontsize=12)

top_products=df.groupby(['Product Name']).sum().sort_values('Sales',ascend
ing=False).head(5)
top_products=top_products[['Sales']].round(2)
top_products.reset_index(inplace=True)
top_products

plt.figure(figsize=(20,15))
plt.bar(top_products['Product Name'],top_products['Sales'],color='#080b45'
,edgecolor='green')
plt.xticks(rotation='vertical')
plt.title('Products with highest sales',fontsize=15)
plt.xlabel('Products',fontsize=12)
plt.ylabel('Revenu',fontsize=12)

df['day'] = df['Order day']

plt.figure(figsize=(20,15))
df['Ship year'].value_counts().plot.pie()

year_sales = df.groupby(['Ship year']).sum().sort_values('Sales',ascending
=False).head(5)
year_sales.reset_index(inplace=True)
year_sales

sales_2018 = df[df['Ship year']==2018]
sales_2018.head()

sales = df.groupby(['Order-
month']).sum().sort_values('Sales',ascending=False).head(5)
sales.reset_index(inplace=True)
```

```
sales

top_cate=df.groupby(['Sub-
Category']).sum().sort_values('Sales',ascending=False).head(5)
top_cate=top_cate[['Sales']].round(2)
top_cate.reset_index(inplace=True)
top_cate

plt.figure(figsize=(20,15))
plt.bar(top_cate['Sub-
Category'],top_products['Sales'],color='#080b45',edgecolor='green')
plt.xticks(rotation='vertical')
plt.title('Sub-Category with highest sales',fontsize=15)
plt.xlabel('Sub-Category',fontsize=12)
plt.ylabel('Revenu',fontsize=12)

!jupyter nbconvert --to html ///content/sample_data/sprint4_solution.ipynb
```

GitHub link – https://github.com/IBM-EPBL/IBM-Project-37320-1660304232
Demo link - https://youtu.be/2gudnQnCqd8