

Global Sales Data Analytics

Team Id : PNT2022TMID22429

INTRODUCTION

The objectives of the project **GLOBAL SALES DATA ANALYTICS** is

- 1.To know fundamental concepts and can work on IBM Cognos Analytics.
- 2.To Gain a broad understanding of plotting different visualizations to provide a suitable solution.
- 3.Able to create meaningful Visualizations and Dashboard(s).

LITERATURE SURVEY

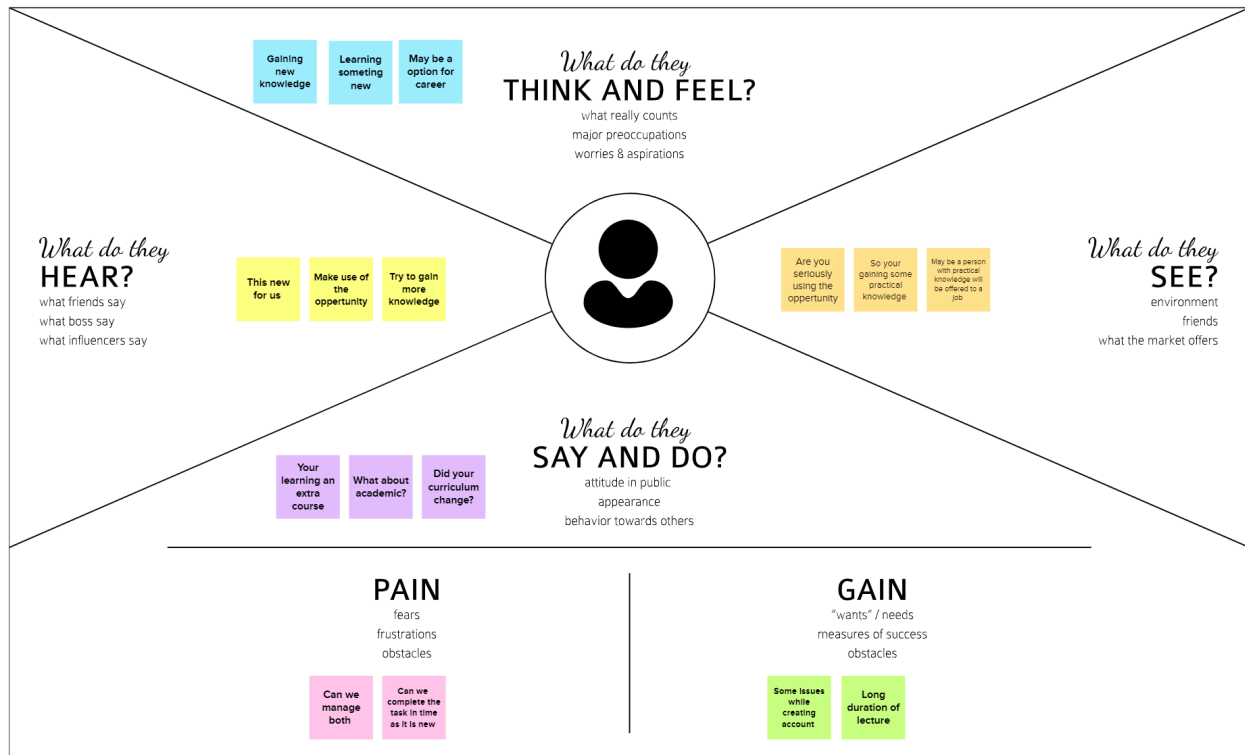
S.NO	TITLE	AUTHOR	PUBLISHED YEAR	REFERENCE
1	Drug sales data analysis for outbreak detection of infectious diseases	Mathilde Pivette	2014	https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-014-0604-2#Abs1
2	Fast fashion sales forecasting with limited data and time	Tsan-Ming Choi Na Liu	2012	https://www.sciencedirect.com/science/article/abs/pii/S0167923613002558
3	Crypto economics: Data Application for Token Sales Analysis	Jin, Seungmin Ali, Rashid	2017	https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1001&context=icis2017b
4	How Online Product Reviews Affect Retail Sales: A Meta - analysis	Kristopher Floyd Ryan Freling	2014	https://www.sciencedirect.com/science/article/abs/pii/S0022435914000293

PROBLEM STATEMENT

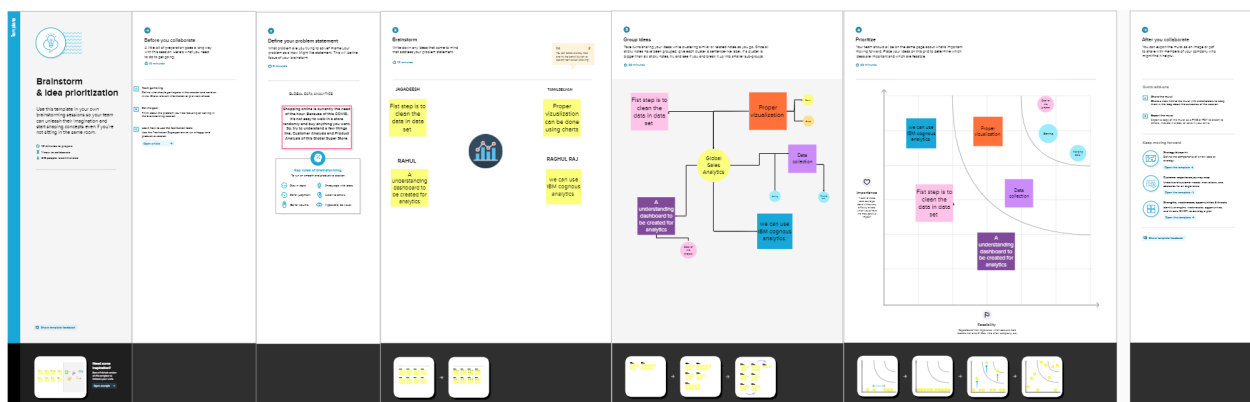
To Analyze the data and creating charts ,dashboard for the data.

IDEATION & PROPOSED SOLUTION

1. Empathy Map



2. Ideation & Brainstorming



3. Proposed Solution

S.NO	Parameter	Description
1	Problem Statement (Problem to be solved)	To create a dashboard and visualization for Analysis of this Global Super Store.
2	Idea / Solution description	Analysis can be done using IBM cognos analytics and google colab
3	Novelty / Uniqueness	Uniqueness is that the visualization for each section in the data
4	Uniqueness is that the visualization for each section in the data	Customer can easily understand the analysis on the global sales
5	Business Model (Revenue Model)	This analysis will be useful when it is implemented in business and form predecive analysis
6	Scalability of the Solution	The scalability is that the number of data in the dataset

4. Problem Solution fit

Project Title: Global Sales Data Analytics

Project Design Phase-I - Solution Fit Template

Team ID: PNT2022TMID22429

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS Who is your customer? i.e. working parents of 0-5 y.o. kids Customer who is expecting for a sales analytics	6. CUSTOMER CONSTRAINTS CC What constraints prevent your customer from taking action or limit their choices of solution? i.e. spending power, budget, no cash, network connection, available devices. To create a dashboard for the global sales data and visualization	5. AVAILABLE SOLUTIONS AS Which solutions are available to the customers, when they face the problem or need to get the job done? what have they tried in the past? what pros & cons do these solutions have? i.e. pen and paper is an alternative to digital note-taking The project is done in IBM cognos analytics for vizualization and dashboard	Explore AS, differentiate
	2. JOBS-TO-BE-DONE / PROBLEMS J&P Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one, explore different sides. Jobs to be done is data cleaning, exporation of data vizualization and there will be no problem in the analytics	9. PROBLEM ROOT CAUSE RC What is the real reason that this problem exists? What is the back story behind the need to do this job? i.e. customers have to do it because of the change in regilts/10% The problem is to create a dashboard and visualization	7. BEHAVIOUR BE What does your customer do to address the problem and get the job done? i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace) The problem can be solved using IBM cognos analytics	
Identify strong TR & EM	3. TRIGGERS TR What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news. In developing world online sales plays vital role where customer can easily buy products from home	10. YOUR SOLUTION SL If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality. If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour. Solution for creating dashboard and visualization is done use tools in IBM cognos analytics	8. CHANNELS of BEHAVIOUR CH 8.1 ONLINE What kind of actions do customers take online? Extract online channels from #7 8.2 OFFLINE What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development. 8.1 ONLINE Customer orders their products in online through apps and websites 8.1 OFFLINE Customer buys their products on foot	Fit TR & SL factors against BE
	4. EMOTIONS: BEFORE / AFTER EM How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure > confident, in control - use it in your communication strategy & design. In this global sales analytics customer can understand how the online sales in developed with the help of dashboard and data vizualization. Before customer is not aware of online purchase where after covid the online purchase has grow and customer is aware of it			

REQUIREMENT ANALYSIS

Functional requirement

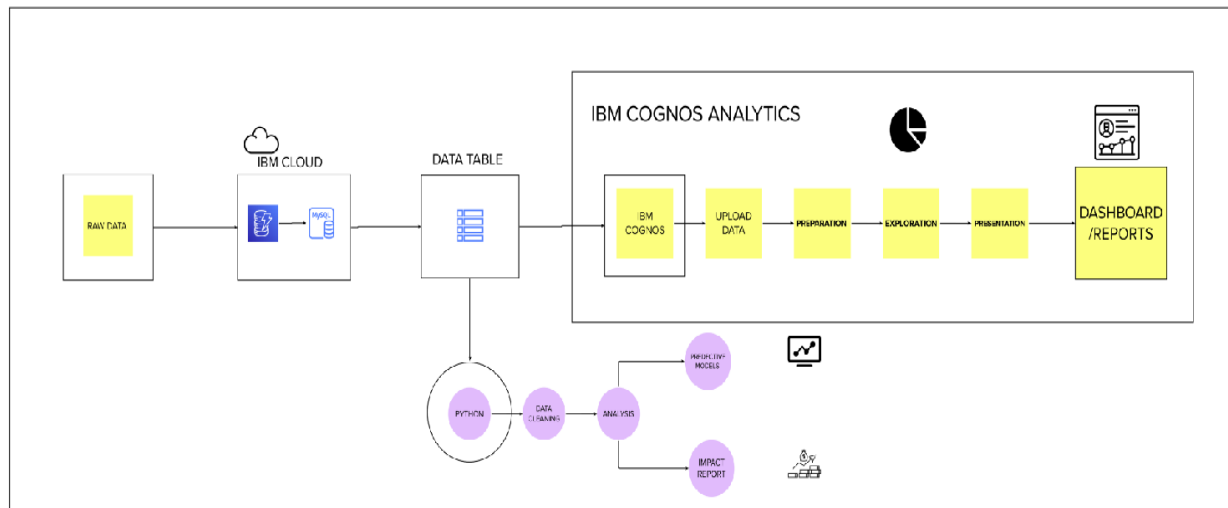
FR No	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User dashboard creation	Created by ibm cognos analytics
FR-4	Predictive analytics	Done by using Machine Learning in python programming language
FR-5	Creating visualization	Created using IBM cognos analytics

Non-functional Requirements

FR No	Non-Functional Requirement	Description
NFR-1	Usability	Used to understand data and analyse the data by vizuzlization
NFR-2	Security	Security depends on the data and the analyst
NFR-3	Reliability	The data can be updated and the task can be executed
NFR-4	Performance	Performance is high were large amont of data can be used effectively
NFR-5	Availability	Available at a time and anywhere when upload to cloud.
NFR-6	Scalability	Scalability depends on the size of the data in a task

PROJECT DESIGN

Data Flow Diagrams



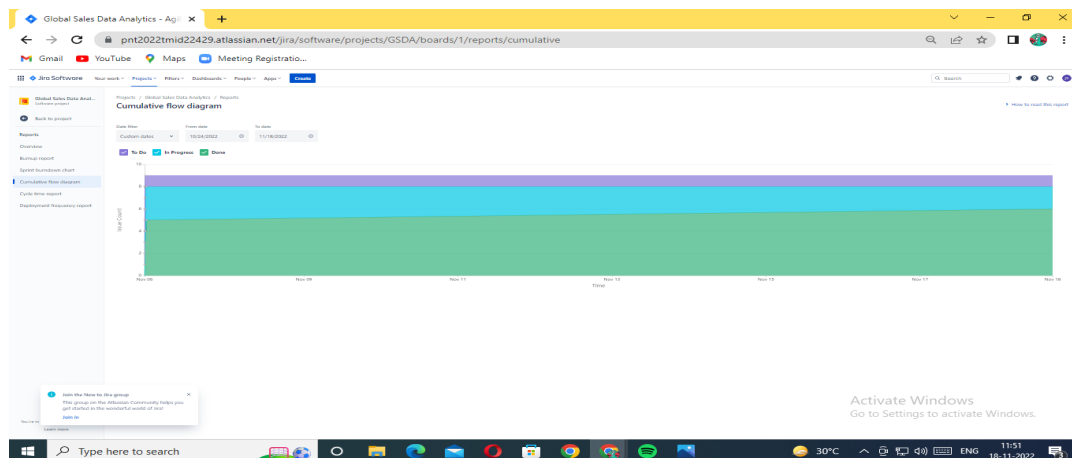
USER STORIES

User Story Number	User Story / Task	Priority
USN-1	Download the dataset from Kaggle API	High
USN-2	To understand the data in dataset	Medium
USN-3	Load the dataset in IBM cognos analytics	High
USN-4	Prepare the data with no null values	Low
USN-5	Create new calculation for perfect visualization	Medium
USN-6	Create new calculation for perfect visualization	Medium
USN-7	To track,analyze and display data.	High
USN-8	Narratives that explain how and why data changes over time,final delivery of the project.	High

PROJECT PLANNING & SCHEDULING

Sprint	Functional Requirement (Epic)	User Story / Task	Team Members
Sprint-1	Collect the dataset	Download the dataset from Kaggle API	Tamilselvan R C
Sprint-1	Understand the dataset	To understand the data in dataset	Rahul R
Sprint-2	Loading the data set	Load the dataset in IBM cognos analytics	Raghul Raj D
Sprint-2	Preparation of data	Prepare the data with no null values	Jagadeesh V
Sprint-2	Performing calculations	Create new calculation for perfect visualization	Tamilselvan R C
Sprint-3	Creating Visualization	Visualize the data for the user to understand easily.	Rahul R
Sprint-3	Creating Dashboard	To track,analyze and display data.	Raghul Raj D
Sprint-1	Report,Story and final delivery	Narratives that explain how and why data changes over time,final delivery of the project.	Jagadeesh V

REPORT FROM JIRA



TESTING

TEST CASES

1. Verify user is able to see the Login/Signup popup when user clicked on My account button.
2. Verify the UI elements in Login/Signup popup.
3. Verify whether 13 tabs available.
4. Check whether the first five tabs are interactive.
5. Dashboard responsiveness.
6. User able to use the interactive dashboard.

USER ACCEPTANCE TEST

S.NO	TEST CASES	STATUS
1	Verify user is able to see the Login/Signup popup when user clicked on My account button.	PASS
2	Verify the UI elements in Login/Signup popup.	FAIL
3	Verify whether 13 tabs available.	PASS
4	Check whether the first five tabs are interactive.	PASS
5	Dashboard responsiveness.	PASS
6	User able to use the interactive dashboard.	FAIL

CODING & SOLUTIONING :

Feature 1 :

First download the given zip file from below link and upload it in kaggle

<https://www.kaggle.com/datasets/apoorvaappz/global-super-store-dataset>

Understanding The Dataset Let's understand the data we're working with and give a brief overview of what each feature represents or should represent.

1. Row ID
2. Order_ID
3. Order_Date
4. Ship_Date
5. Customer_ID
6. Customer_ Name
7. Segment
8. City
9. State
10. Country
11. Market
12. Region
13. Product_ID
14. Category
15. Sales

16. Profit

17. Quantity

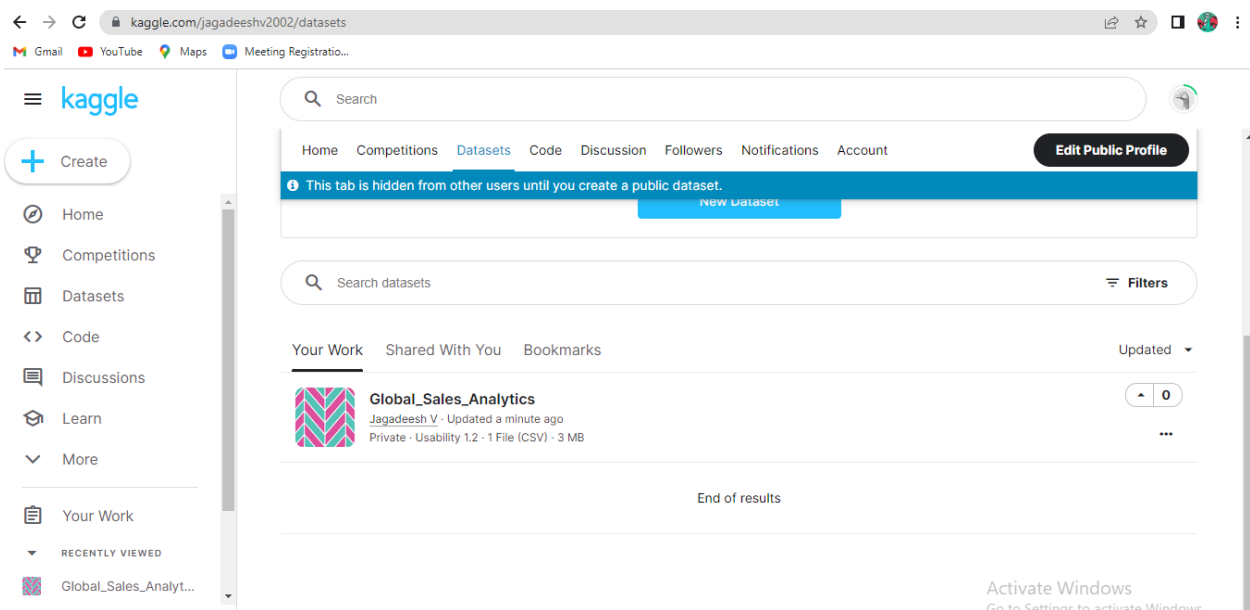
18. Discount

19. Shipping Cost

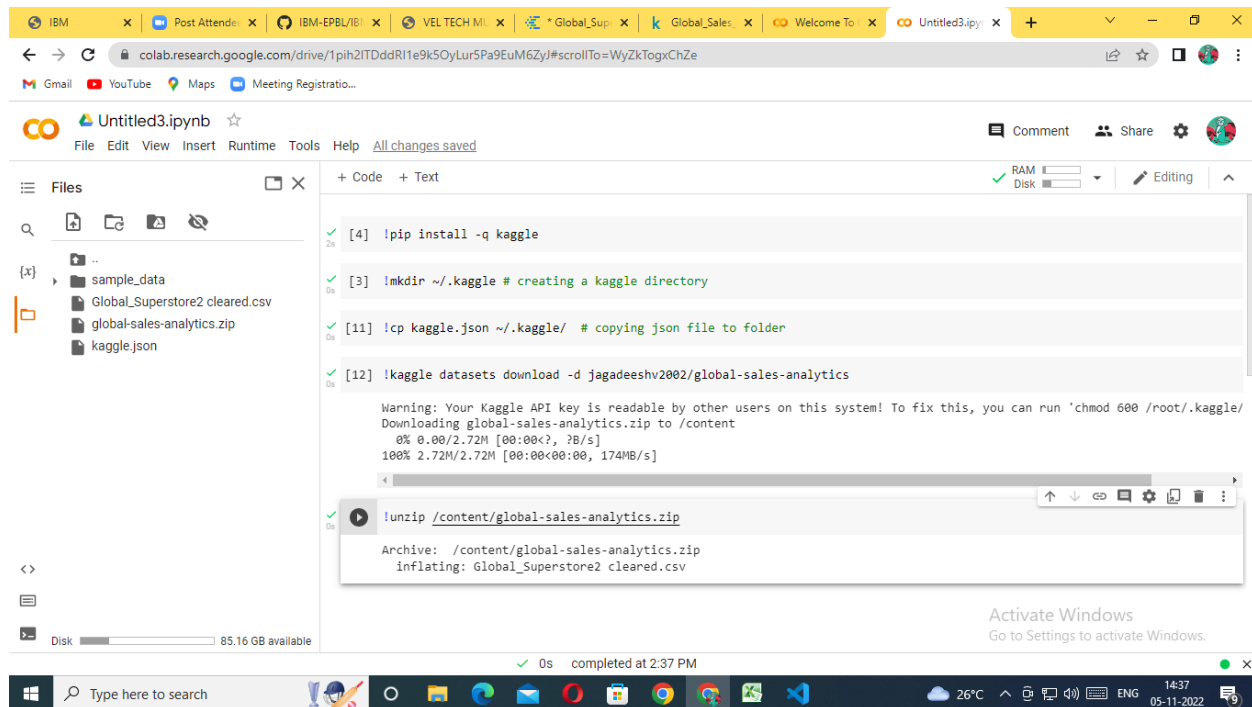
20. Order Priority

Working with dataset :

1. Downloading the data from Kaggle.



2. The downloaded content is extracted in the form of zip folder using google Colab.



Code :

[1] !pip install -q kaggle

[2] !mkdir ~/.kaggle

[6] !cp kaggle.json ~/.kaggle/

[7] kaggle datasets download -d thayanandaramana/nycbike

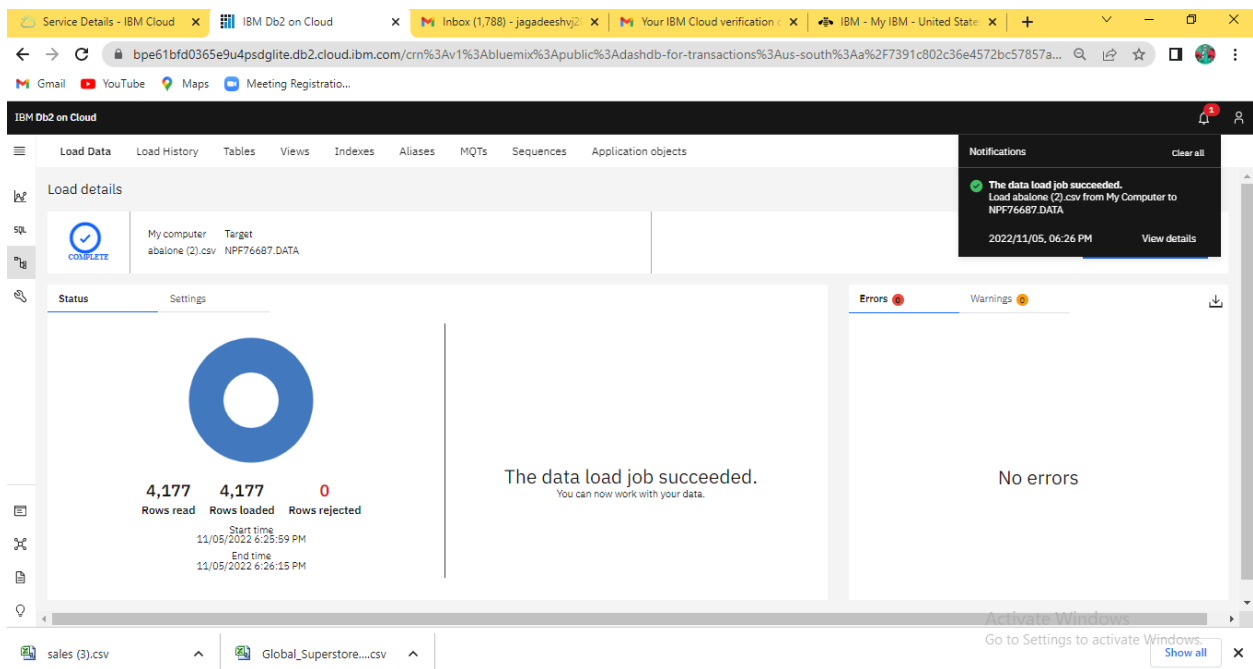
Warning: Your Kaggle API key is readable by other users on this system! To fix this, you can run 'chmod 600 /root/.kaggle/'
Downloading nycbike.zip to /content 56% 9.80M/16.1M [00:00 00:00, 27.2MB/s]
100% 16.1M/16.1M [00:00 00:00, 46.5MB/s]

!unzip /content/nycbike.zip Archive: /content/nycbike.zip inflating: 201306-citibike-tripdata.csv

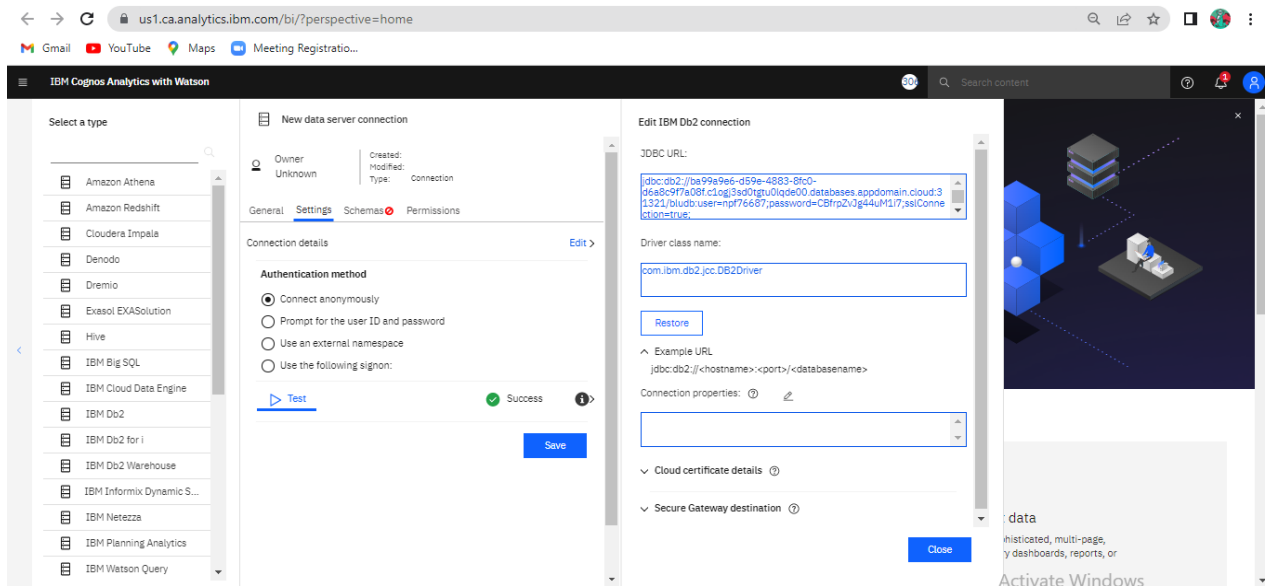
Feature 2 :

IBM Db2 service creation and Db2 connectivity with cognos

1. Create IBM cloud account and selecting IBM Db2
2. Creation of database resource and ensure it is active.
3. Load the data in IBM Db2 on cloud and ensure there is no error after the dataset is loaded (warning errors around 1000 is permitted).

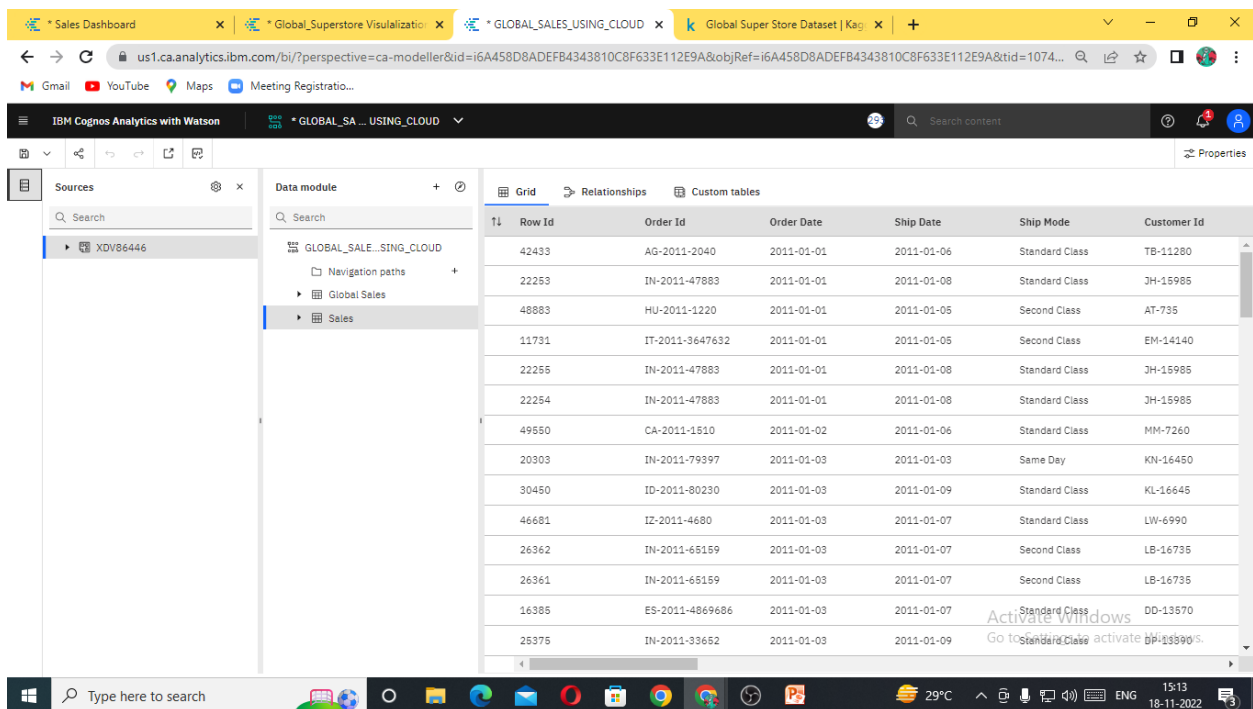


4. Connecting IBM Db2 with IBM Cognos analytics using the username and password created in database on cloud.



5. Click manage in menu then click data server connection and load the schema

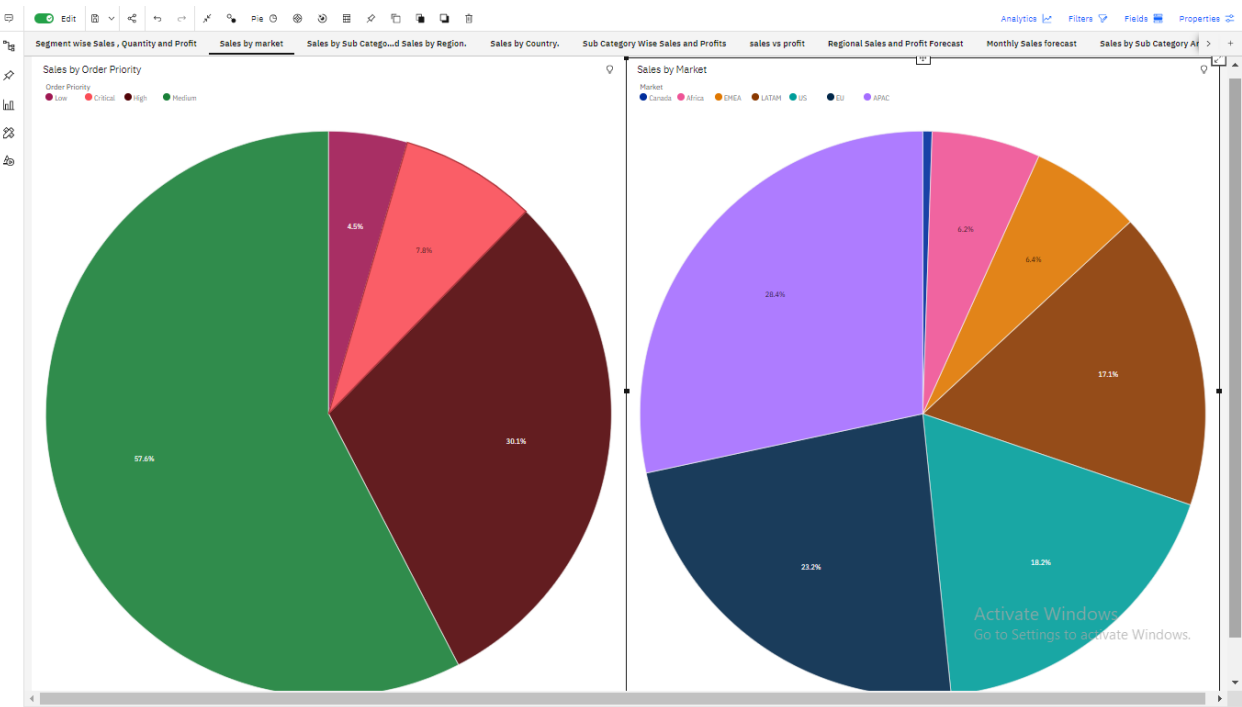
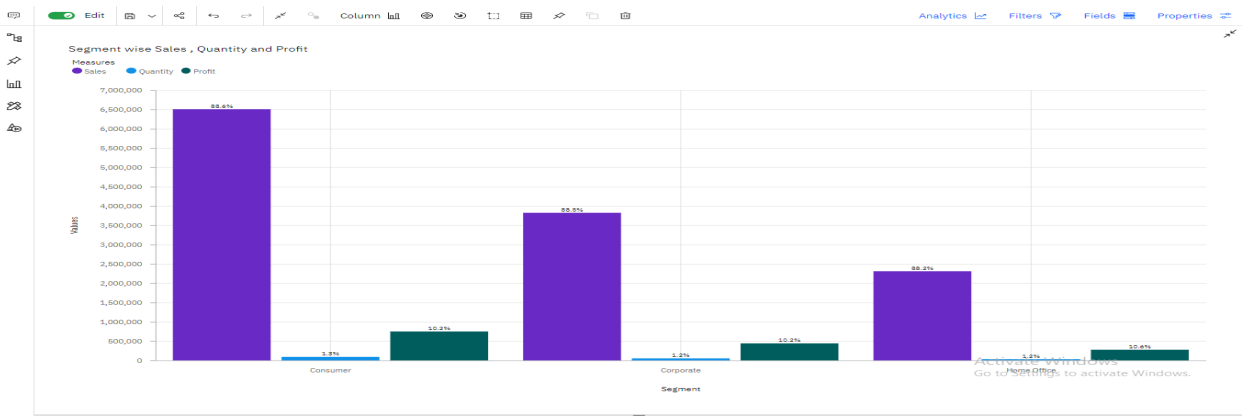
6. After loading of schema, open the loaded data in data module to perform visualization charts.

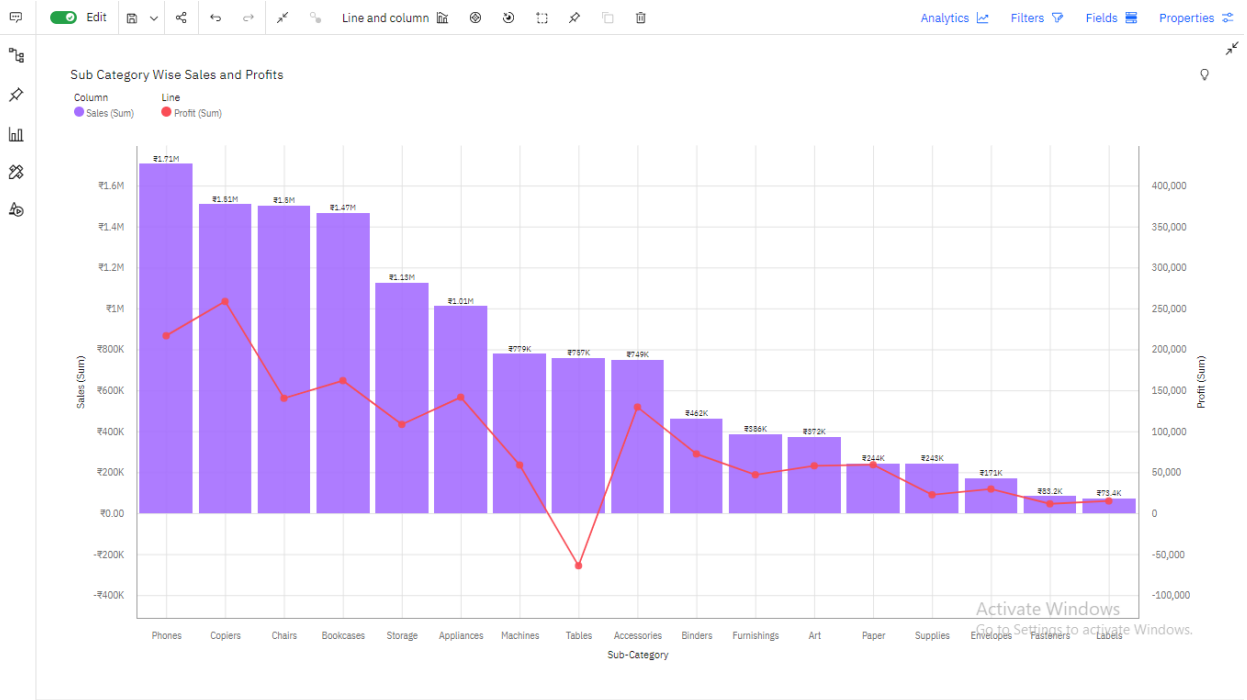
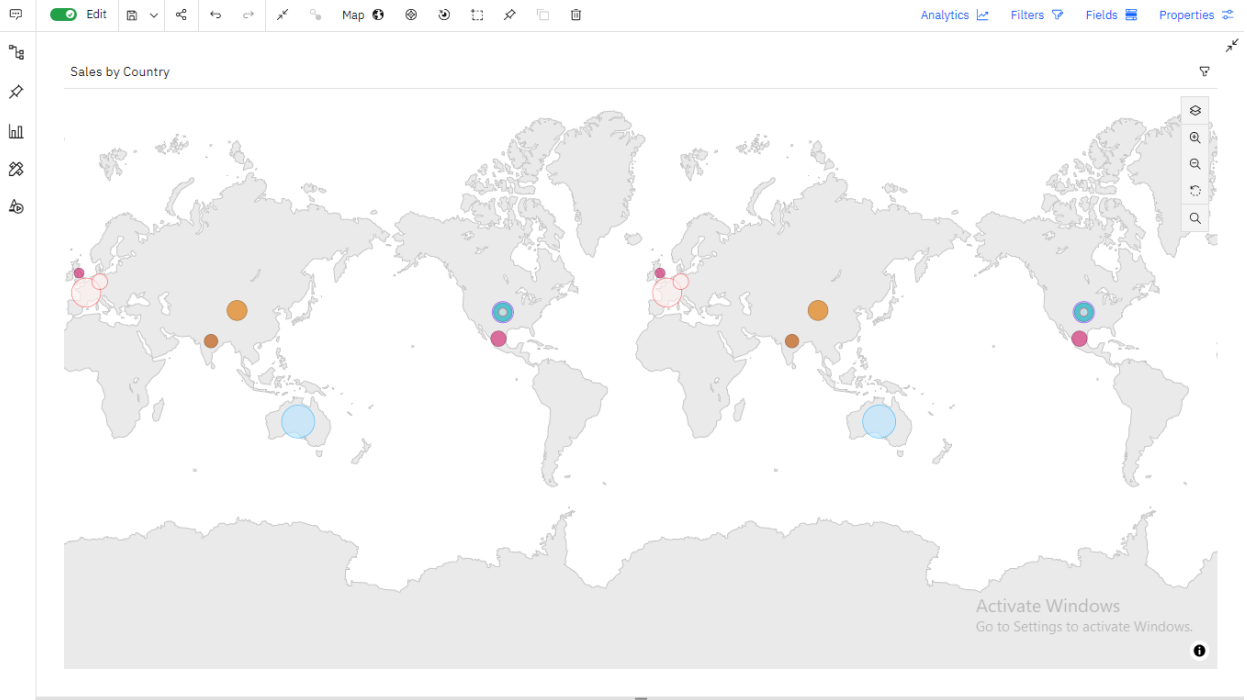


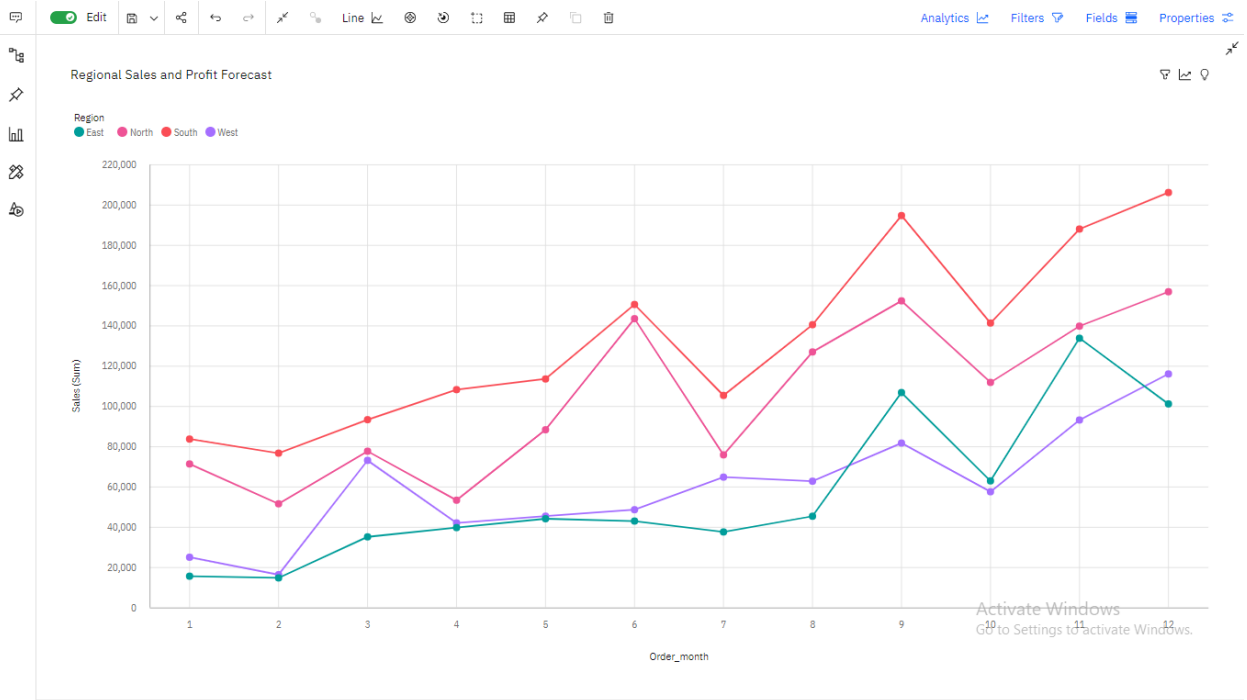
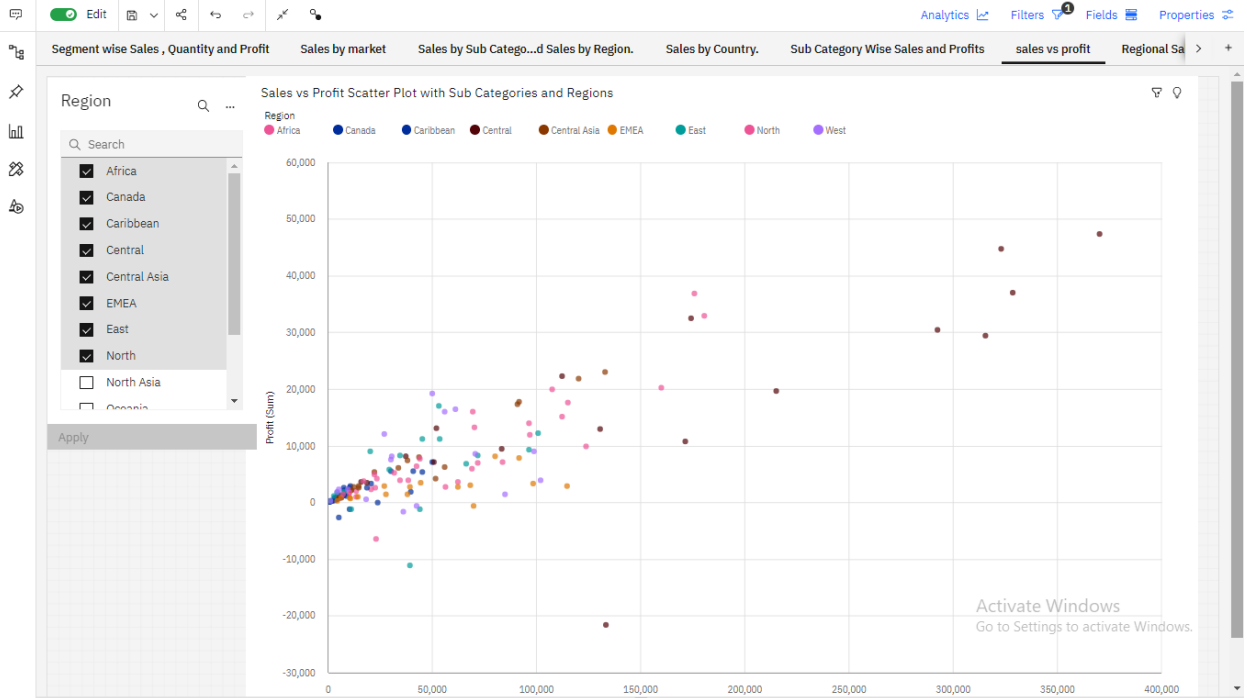
7. Ensure there is no null values, if present perform calculations accordingly.

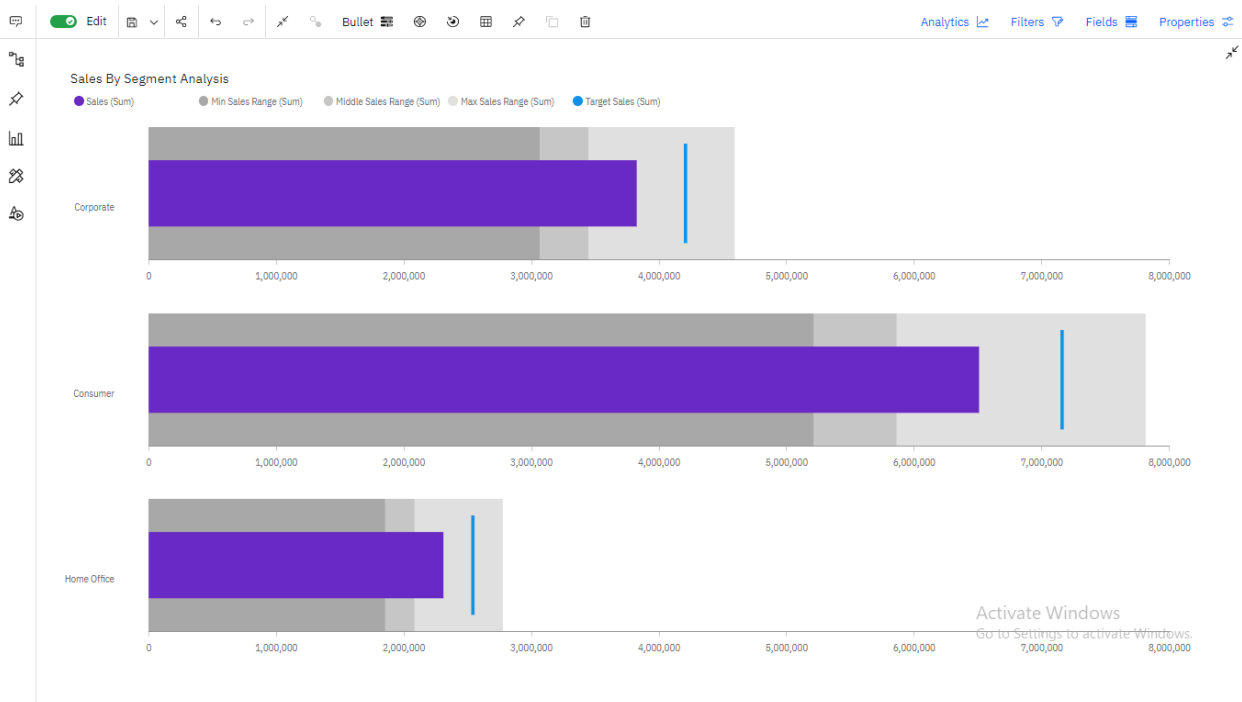
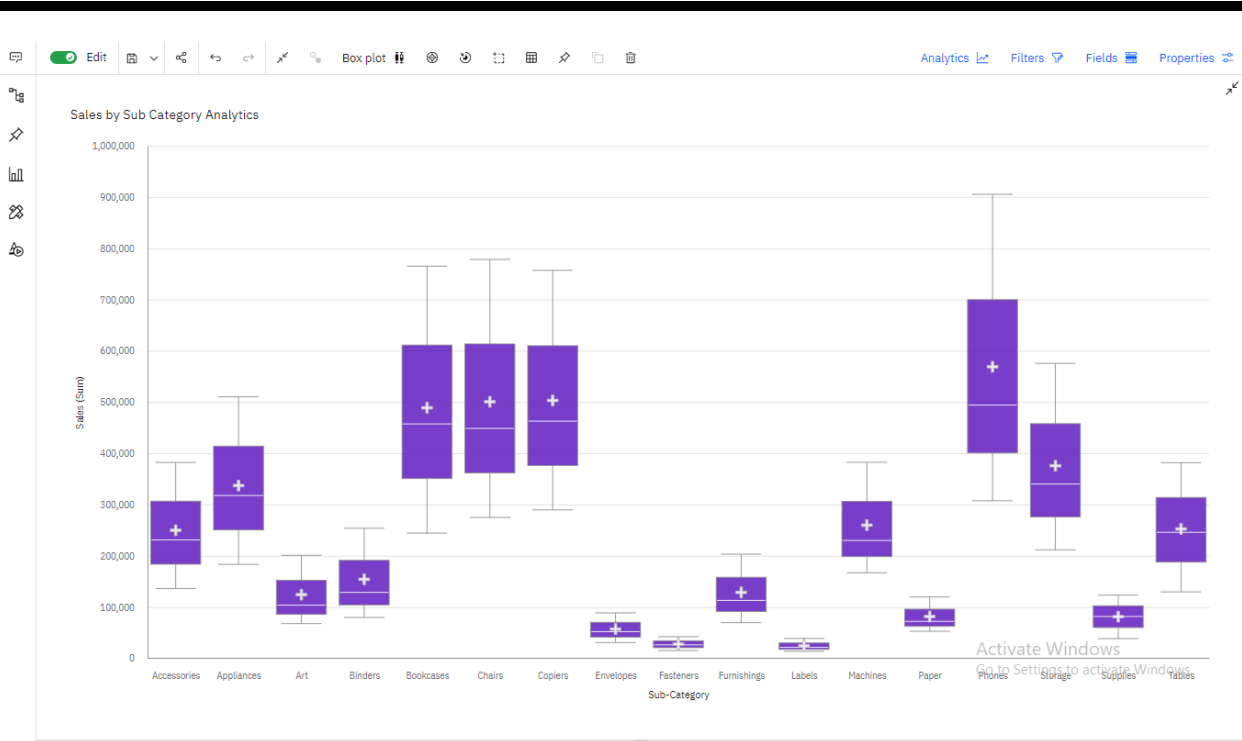
RESULTS

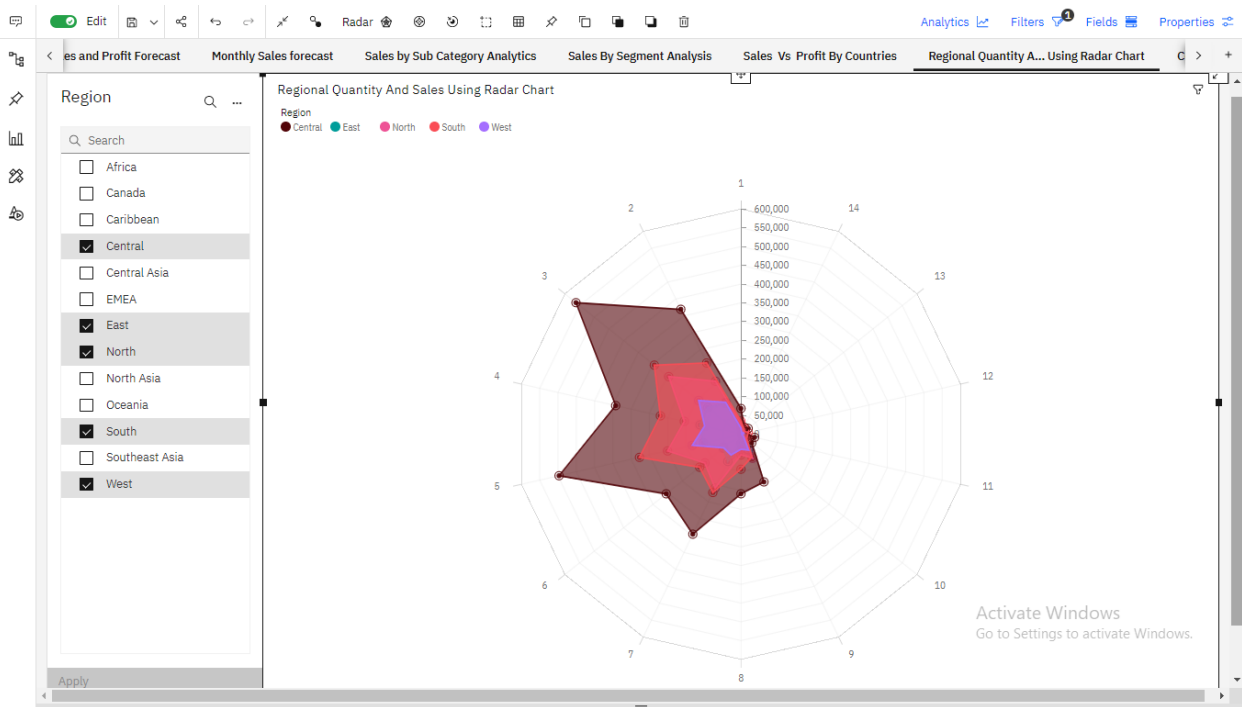
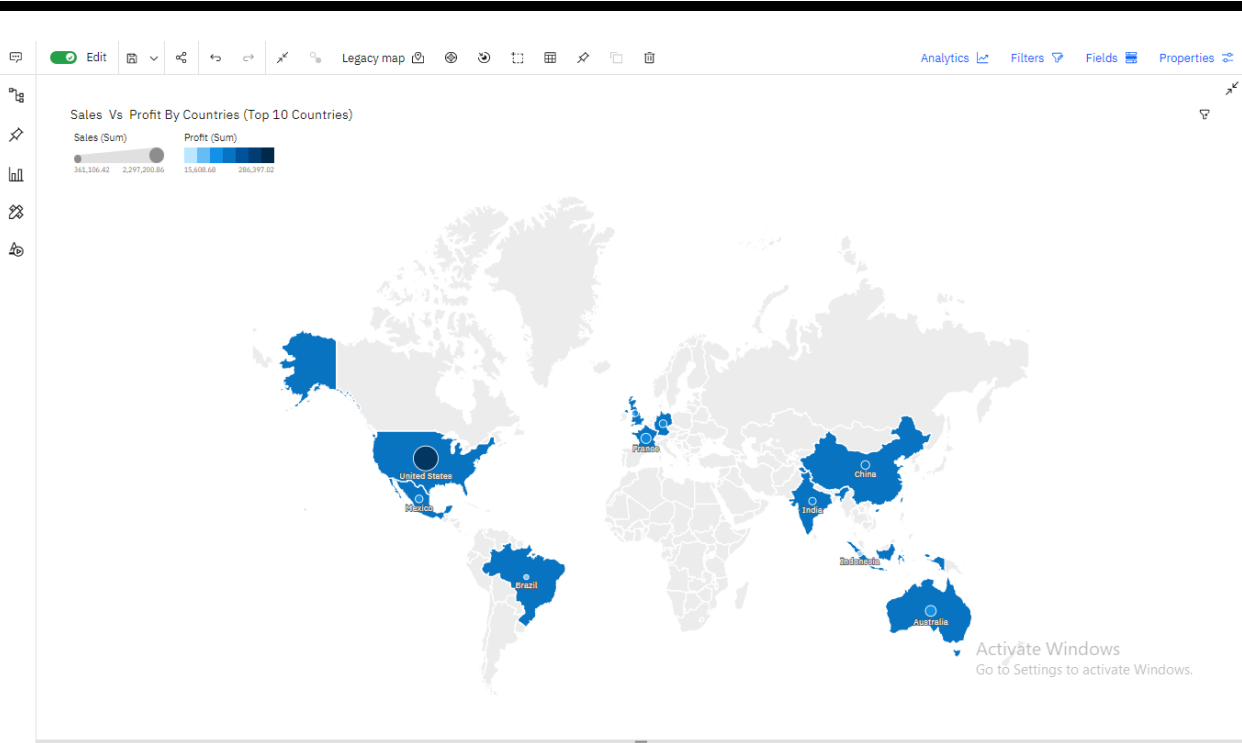
VISUALIZATION CHARTS













DASHBOARD

This is the final dashbord of the project.

We have created an interactive dashboard where by selecting any criteria can represent

the detailed information based on it by visualized charts.

Dashboard contains filters which can make user to easily understand the means.



ADVANTAGE

- 1 Main advantage is that user can easily understand.
- 2 User can easily understand the data.
- 3 With the help of ibm cognos analytics the work done effectively.

DISADVANTAGE

- 1 Lack of intelligent prioritization.

2 Small changes in data can casue a improper dashboard.

CONCLUSION

The conclusion is that the tasks are successfully done.

An intractive dashboard is created for the user.

FUTURE SCOPE

In future by useing machine learning algorithm we can do predictive analysis based on use case.

GIT-HUB AND DEMO VIDEO LINK

GITHUB LINK - <https://github.com/IBM-EPBL/IBM-Project-20598-1659756220>

DEMO VIDEO LINK - <https://youtu.be/0XRrCsPdx-0>

