Global Sales Data Analytics

Team ID PNT2022TMID47824

G.SARAVANAPANDIYAN

P.YUVARAJ

K.SURIYA SOUNDARAM

P.KARUPPAIAH

CHAPTER-1

INTRODUCTION:

We all are constantly thinking about the future and what is expected to happen in the coming weeks, months and even years, and to be able to do so, a look at the past is mandatory. Business needs to be able to see their progress and the factors affecting their sales [1]. In this technological era of large scale data, businesses need to rethink on the modern approaches to better understand the customers to gain a competitive edge in the market. Data is worthless if it cannot be analysed, interpreted and applied in context [2]. In this work, we have used the Walmart's sales data to create business value by understanding customer intent (sentiment analysis) and business analytics. A picture speaks a thousand words and business analytics would help paint a picture through visualization of data to give the retailers insights on their business. With these insights the businesses can make relevant changes to their strategy for the future to maximize profits and success. Most of the raw data, particularly large scale datasets do not offer value in its unprocessed state.

By applying the right set of tools [3], we can pull powerful insights from this stockpile of bits. The main focus here is to read and analyse the Walmart's available datasets to produce insights and the company's overall overview. The retail stores sell products and gain profit from it. There are a lot of subsidiaries of the stores network which are scattered on various geographical locations. As the network of stores is huge and located at different geographical locations, the company would not fully understand the customer needs and market potentials at these various locations. In this work, we used the gathered store sales datasets of Walmart to understand the factors affecting the sales for example, the unemployment rate, fuel prices, temperature and holidays in the different stores located at different geographical locations so that the resources can be managed wisely to maximize on the returns. These insights can help retailers comprehend market conditions of the various factors affecting sales for example Easter holiday would induce a spike in sales and retailers can better allocate resources (supply of goods and human resources). Thus, customer demands are observed accordingly based on the above factors

1.1 Why Need Analytics:

Improve operational efficiency through their daily activities. Assist businesses to understand their customers more precisely. Business uses data visualization to offer projections for future outcomes. These insights help in decision making and planning for the future.
The Benefits of Business Analytics.

1.2 How To Analysis Data:

The Business Analytics Lifecycle Helps Organisations Use *Data* to Identify What to Do Next. Make Confident Decisions & Drive Operational Efficiency. Read Our Client Success Stories. Fast and Reliable. Cloud Deployment. Reduce Costs. Increase Efficiency.

1.3Which software Use:

IBM® Cognos® Analytics integrates reporting, modeling, analysis, dashboards, stories, and event management so that you can understand your organization data, and make effective business decisions. After the software is installed and configured, administrators set up security and manage data sources.

CHAPTER-2

2.1 Solution Requirement:

A solution requirement is **aimed at the concerns of the people who will build and deliver the solution**. It tells those people what the functional and non-functional requirements for the solution will be and how the solution will deliver on the business and stakeholder requirements.

2.1.1 Data Analysis:

Data analysis is the process of cleaning, changing, and processing raw data, and extracting actionable, relevant information that helps businesses make informed decisions.The procedure helps reduce the risks inherent in decision-making by providing useful insights and statistics, often presented in charts, images, tables, and graphs.

2.1.2 Collect sales data:

A spreadsheet can help you collate your data, sales-related or otherwise, but a CRM platform with strong insights features is the best option With the sheer amount of information on leads, deals and communications to keep track of, you'll need a streamlined piece of software with clear access to your data.

2.1.3 prepare a data:

Data preparation is the process of preparing raw data so that it is suitable for further processing and analysis. Key steps include collecting, cleaning, and labeling raw data into a form suitable for machine learning (ML) algorithms and then exploring and visualizing the data.

2.1.4 Create Dashboard:

A dashboard for data analytics is a tool used to multi-task, organize, visualize, analyze, and track data. The overall purpose of a data analytics dashboard is to make it easier for data analysts, decision makers, and average users to understand their data, gain deeper insights, and make better data-driven decisions.

2.1.5 Skill Recruitment:

IBM Account,IBM Cognos ,Python or R.

2.1.6 Project Objectives:

By the end of this Project, you will:

- Know fundamental concepts and can work on IBM Cognos Analytics.

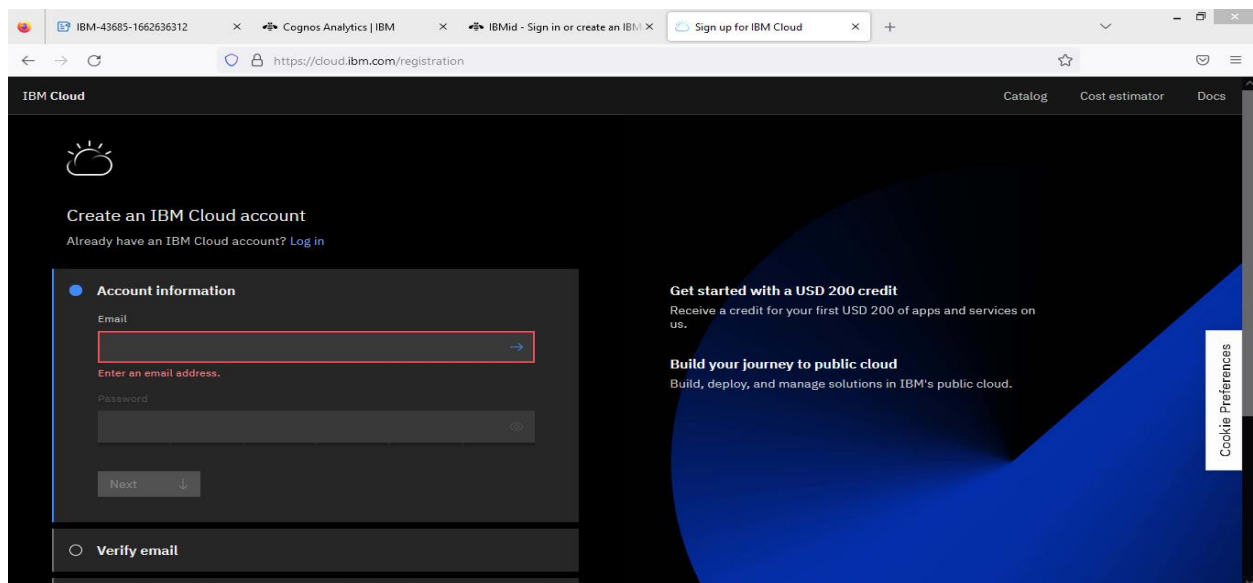- Gain a broad understanding of plotting different visualizations to provide a suitable solution.

Able to create meaningful Visualizations and Dashboard(s).

2.1.7 Project Description:

Shopping online is currently the need of the hour. Because of this COVID, it's not easy to walk in a store randomly and buy anything you want. So, try to understand a few things like, Customer Analysis and Product Analysis of this Global Super Store.

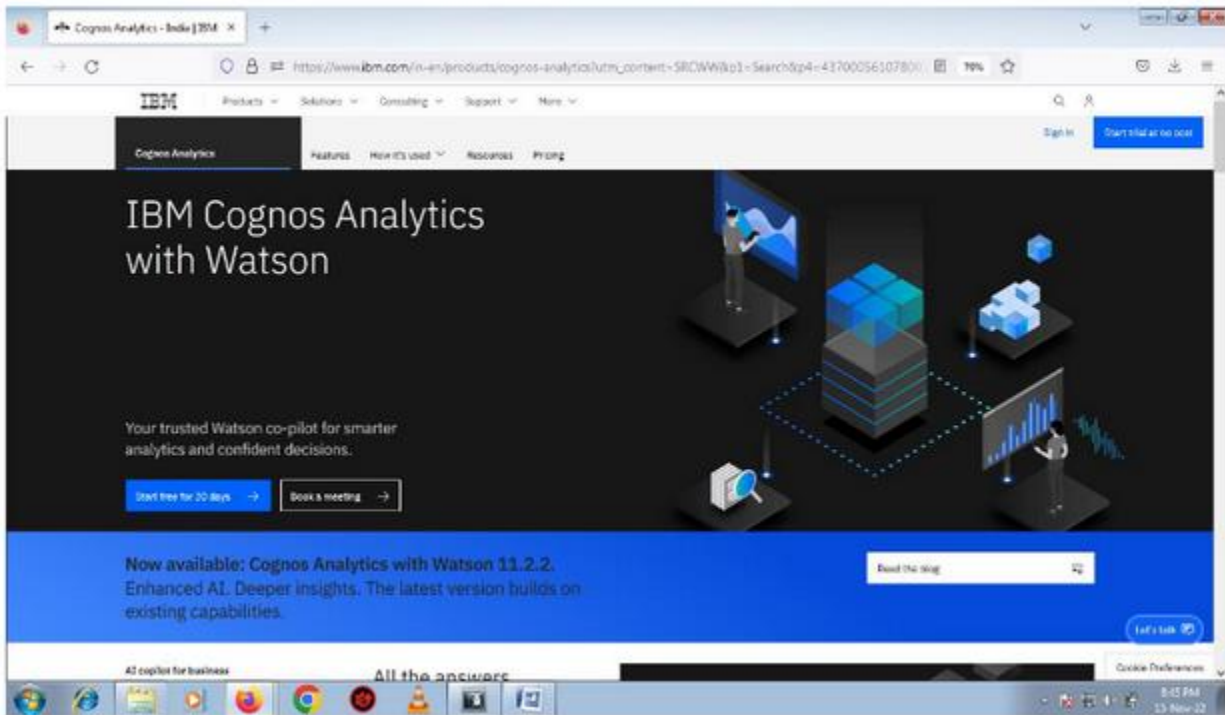## 2.2 create IBM account:

Create IBM cloud account.
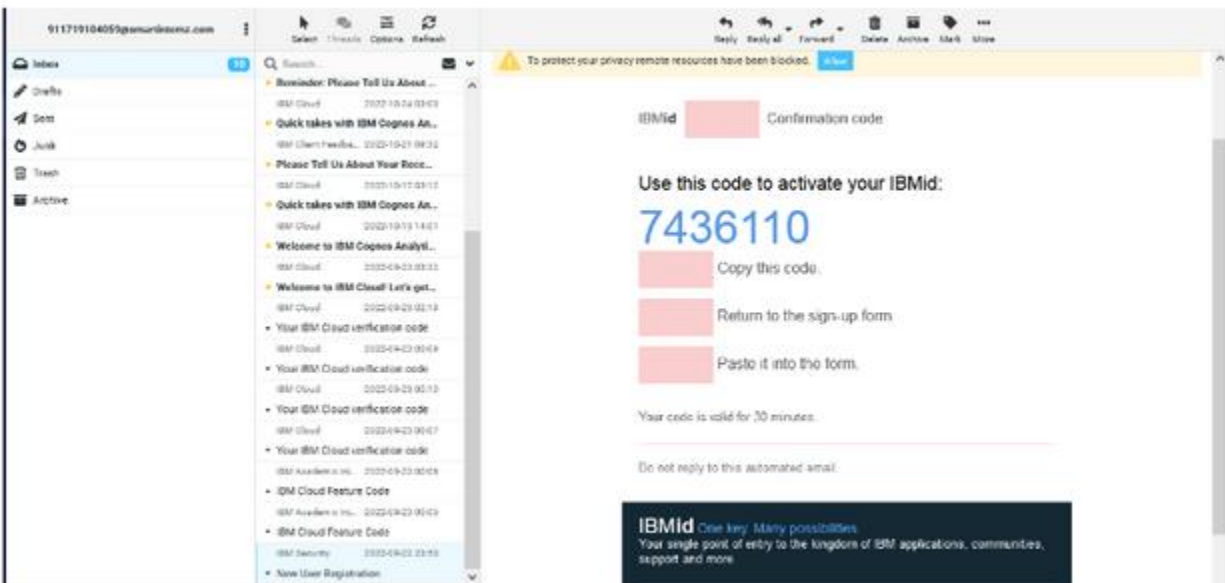
IBM CLOUD DASHBOARD

If you successfully created IBM cloud account. After create a account for ibm cognos software.

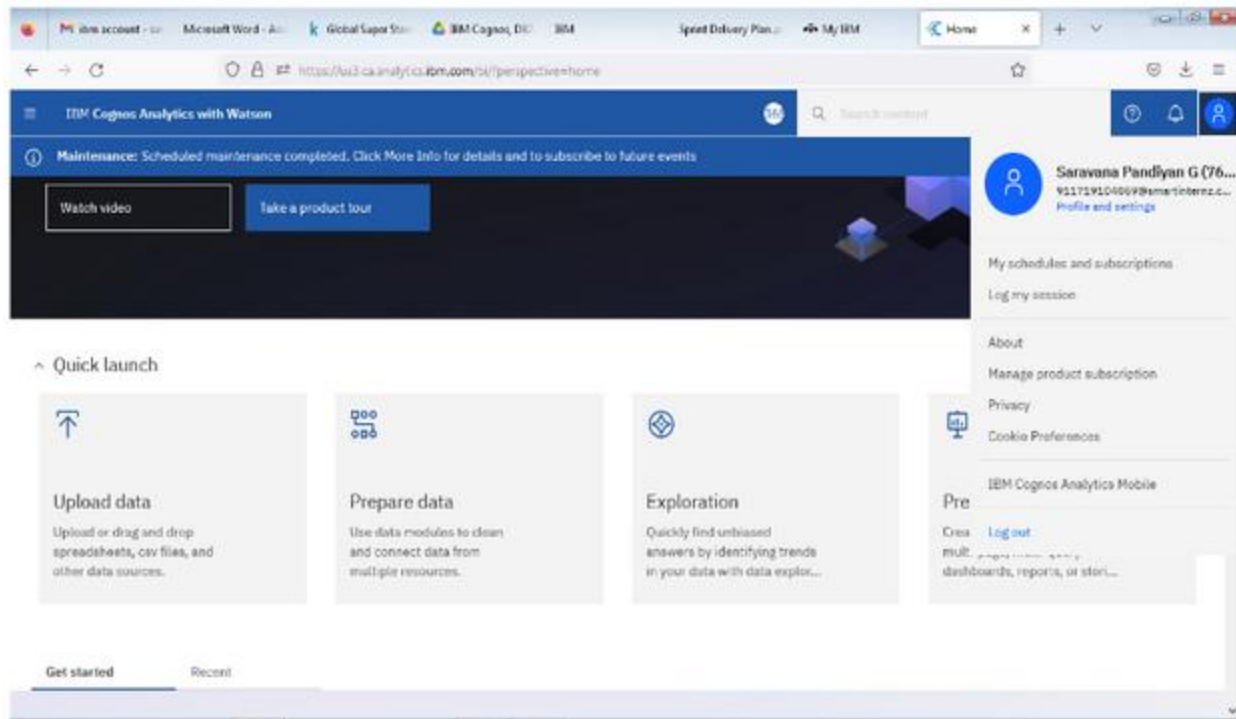## 2.2 create IBM cognos application:

Register a IBM Cognos application.

LOGIN TO WEB-MAIL AND ACCESS YOUR REGISTRATION CODE,

After registering login and conform the cognos application. AND Successfully create a IBM cognos account login

2.4 Working With Dataset:
2.4.1 Collect the dataset

Consumer and product Analytics on Global Super Store Data.Consumer and product Analytics on Global Super Store Data.

create a account.

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

2.4.2 Upload data:

select a csv file and upload

2.4.3 Prepare data:
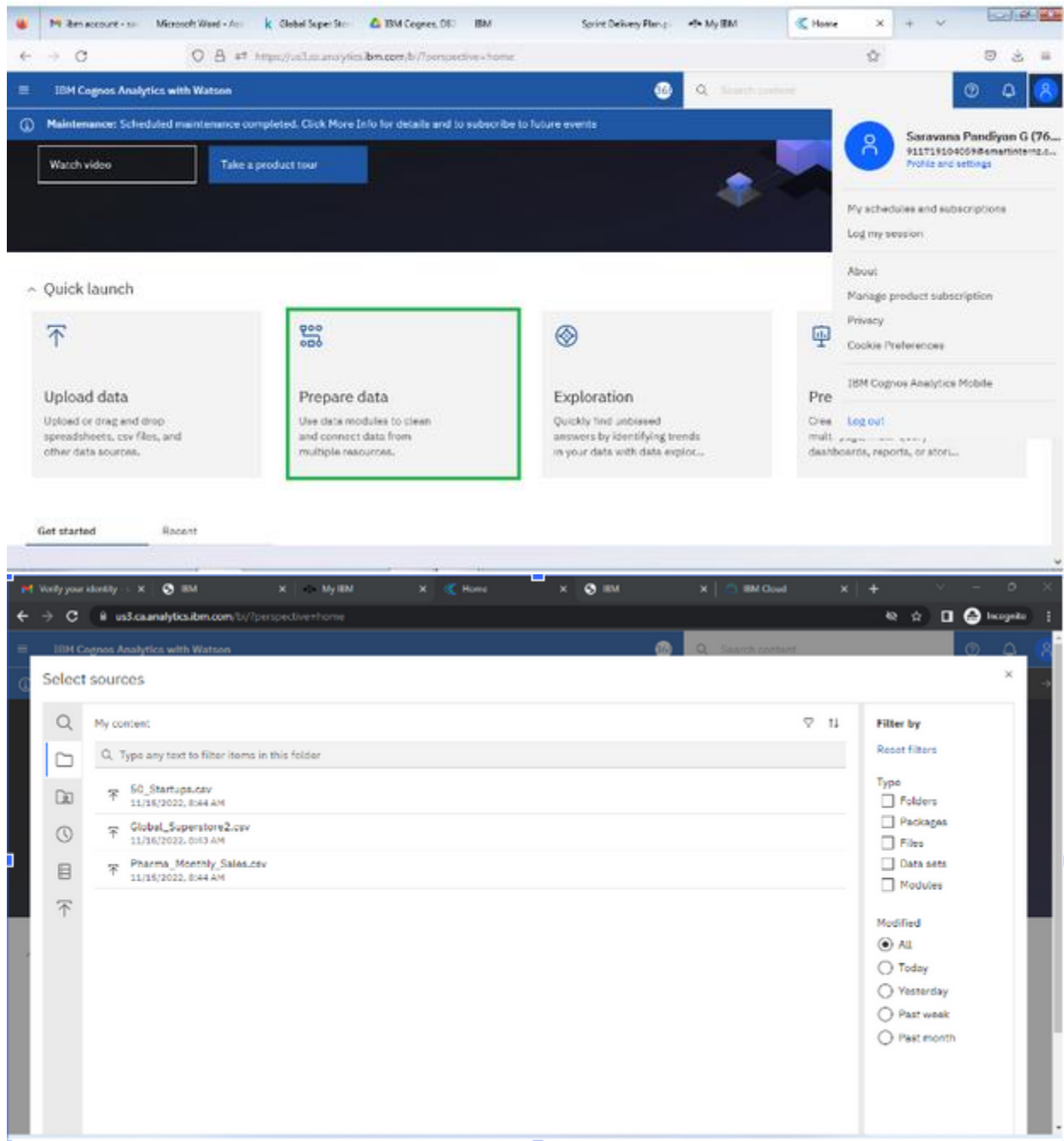
2.6 Using a different visualization type

Procedure:

1. Select the visualization that you want to work with.
2. Click the Change visualization icon in the toolbar.

   The toolbar is located at the top of the dashboard by default. Click the undock toolbar icon ⬓ if you want to move the toolbar to a selected visualization.

3. Click the visualization type that you want to use.

   Take a look at how each visualization type communicates data differently. For example, use a bar, column, or line visualization to compare a set of values. Use a line or area visualization to track relationships. Use a tree map or pie visualization to see the parts of a whole.

4. If you select a visualization type that requires different data slots to be used, add a column to each empty data slot that has an asterisk (*) in it.

2.6.1 Area

Use an area visualization to emphasize the magnitude of change over time. Area charts are like line charts, but the areas below the lines are filled with colors or patterns. Stacked charts are useful for comparing proportional contributions in a category. They plot the relative value that each data series contributes to the total.Because an area visualization stacks the results for each column or item, the total of all results is easily seen.
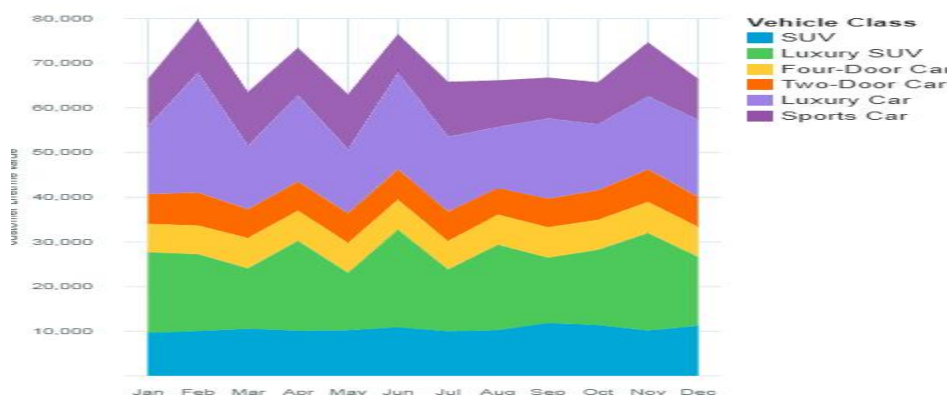
2.6.2 Bar

Use a bar visualization to compare values by one or more columns, such as sales for products or sales for products each month. Bar visualizations use horizontal data markers that are arranged in groups to compare individual values. You can use bar visualizations to compare discrete data or to show trends over time. A bar visualization can show change over a specific time period or can compare and contrast two or more columns in a time period or over time. If there are so many bars that the labels are impossible to read, filter the data to focus on a subset of the data or use a tree map.
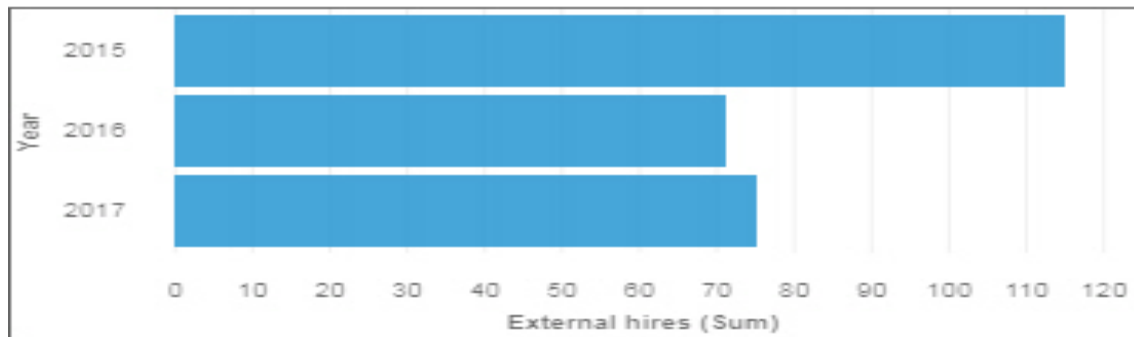
2.6.3 Bubble

Use a bubble visualization to show relationships among columns that contain numeric values, such as revenue and profit. A bubble visualization uses data points and bubbles to plot measures anywhere along a scale. One measure is plotted along each axis. The size of the bubble represents a third measure. Use bubble visualizations to represent financial data or any data where measure values are related.

The bubbles are in different sizes and colors. The x-axis represents one measure. The y-axis represents another measure, and the size of the bubbles represents the third measure. In the example shown below, color is represented by an identifier.
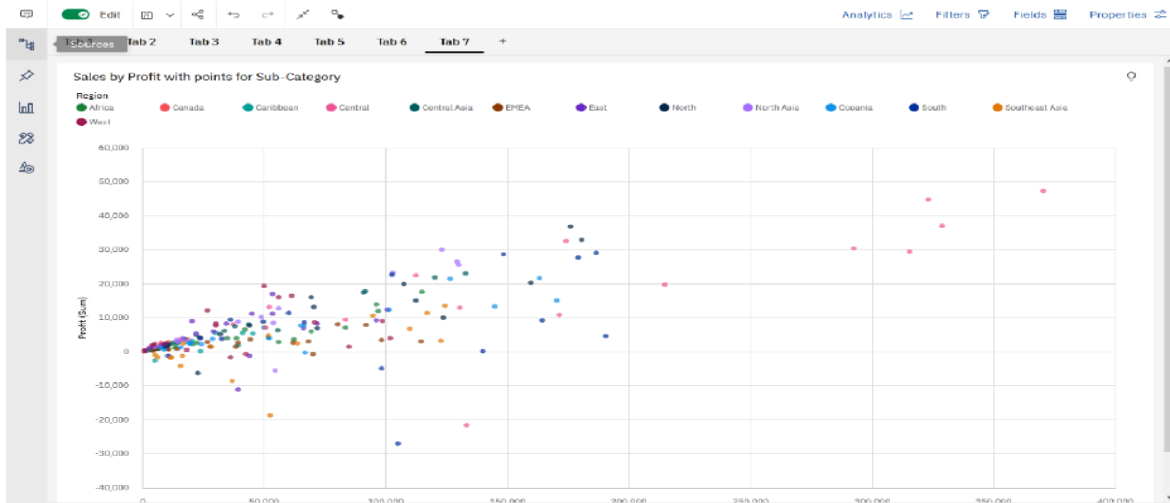
2.6.4 Bullet

Use bullet charts to show measures that need to be compared against a target value.In a call center, a bullet chart can be used to measure metrics like call volume, call answer speed, and percentage of abandoned calls. In manufacturing, a bullet chart can be used to track metrics like number of defects and orders that are shipped. In a fitness context, a bullet chart can be used to measure metrics like steps that are taken and calories that are burnt.

Bullet visualizations compare an actual measure (the bullet) to targeted measure (the target). Bullet visualizations also relate the compared measures against colored regions in the background that provide more qualitative measurements, such as good, satisfactory, and poor. Bullet visualizations can be shown at small sizes while still effectively conveying information.

A bullet visualization features a single, primary measure. For example, current year-to-date revenue. And compares that measure to one or more other measures to enrich its meaning. For example, compared to a target. The primary measure is displayed in the context of a qualitative range of performance, such as poor, satisfactory, and good.

If you select a bullet visualization, then specify the following fields:

- The Actual bar field specifies the actual measure.
- The Target field specifies the target measure.
- The Minimum range field specifies the minimum qualitative range.
- The Middle range field specifies the middle qualitative range.
- The Maximum range field specifies the higher qualitative range.

13

2.6.5 Column

Use a column visualization to compare values by one or more columns, such as sales for products or sales for products each month. Column visualizations use vertical data markers that are arranged in groups to compare individual values. Use column visualizations to compare discrete data or show trends over time. A column visualization shows change over a specific time period or can compare and contrast two or more columns in a time period or over time. If there are so many bars that the labels are impossible to read, filter the data to focus on a subset of the data or use a tree map.
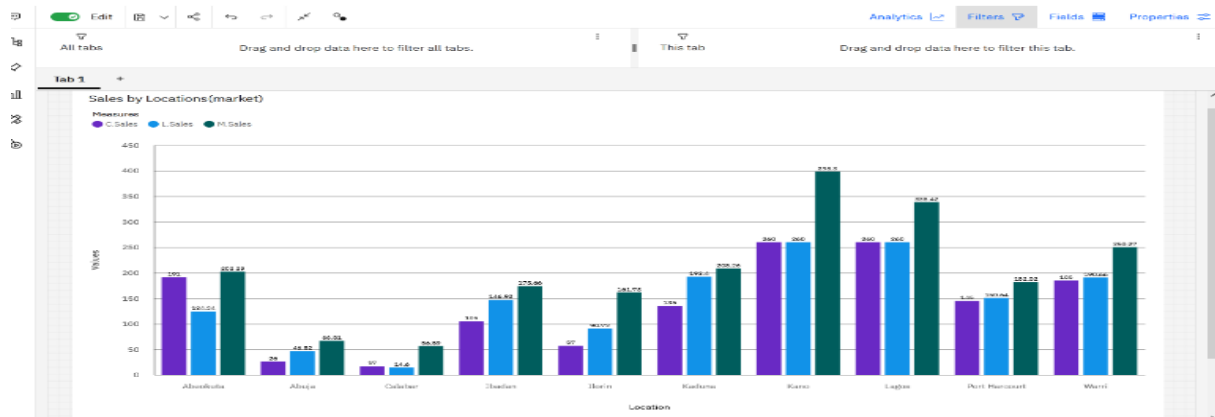
2.6.6 Line

Use a line visualization to show trends over time. A line visualization can compare trends and cycles, infer relationships between variables, or show how a single variable is performing over time. For an effective line visualization, use a time column in the x-axis, such as years, quarters, months, or days. If the x-axis shows something else, such as Canada, Netherlands, UK, and US, use a bar or column visualization.
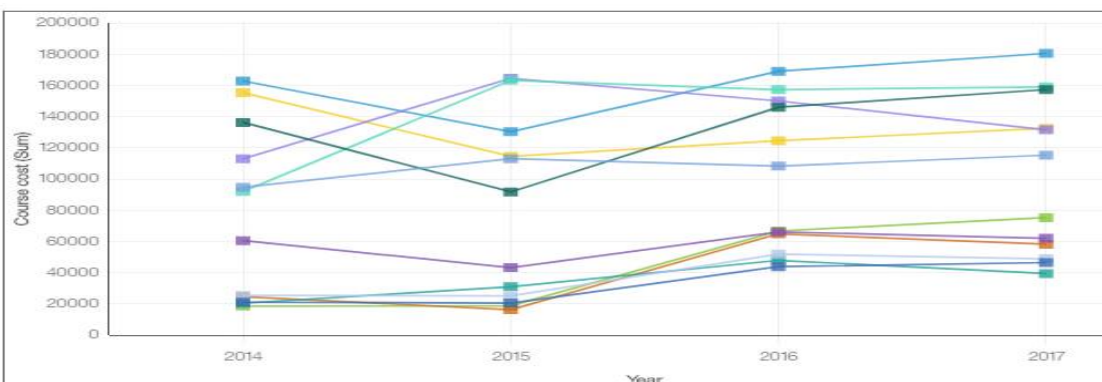
2.6.7 Line and column

Use a line and column visualization to highlight relationships between multiple data series by combining bars and lines with one visualization.

For example, this line and column visualization shows the relationship between course cost and expense totals by department.

2.6.8 Map

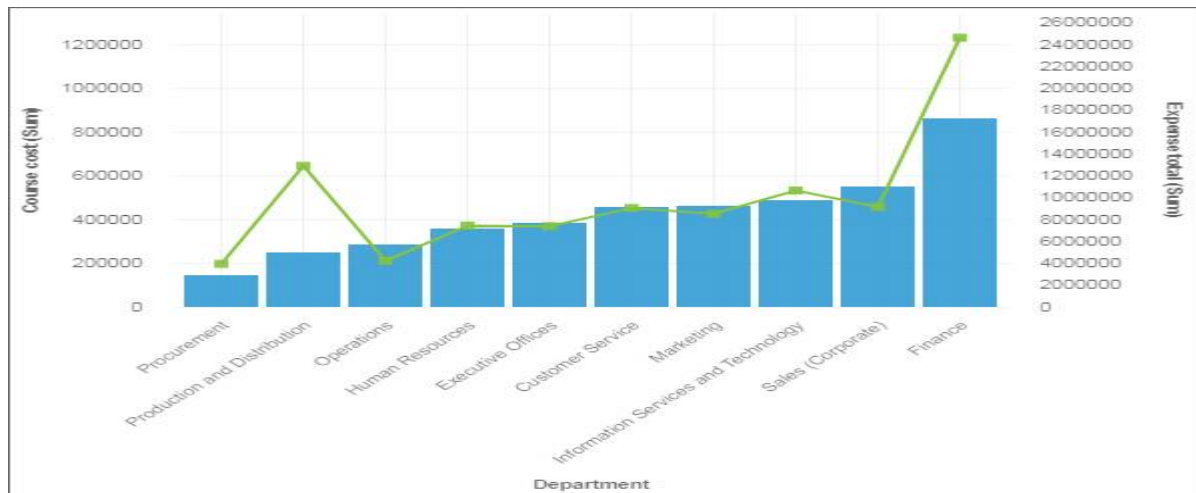Use a map when you want to see patterns in your data by geography. =Your data asset must contain geographical data, such as countries, states, provinces, or continents. To determine whether a column is mappable, Cognos Analytics analyzes a sample of 2000 values in the location column, looking for recognizable place names. If 80% or more are recognized as map values, Cognos Analytics produces a map.

For example, you have four countries in your location column: Brazil, China, India, and Russia. The typographical error for India means that only 75% of the values are recognizable place names and you will not see a map as a starting point. But if you have five countries and one has a typographical error in it, you see a map. For example, this map visualization shows revenue by retailer country with the darker color indicating higher revenue.
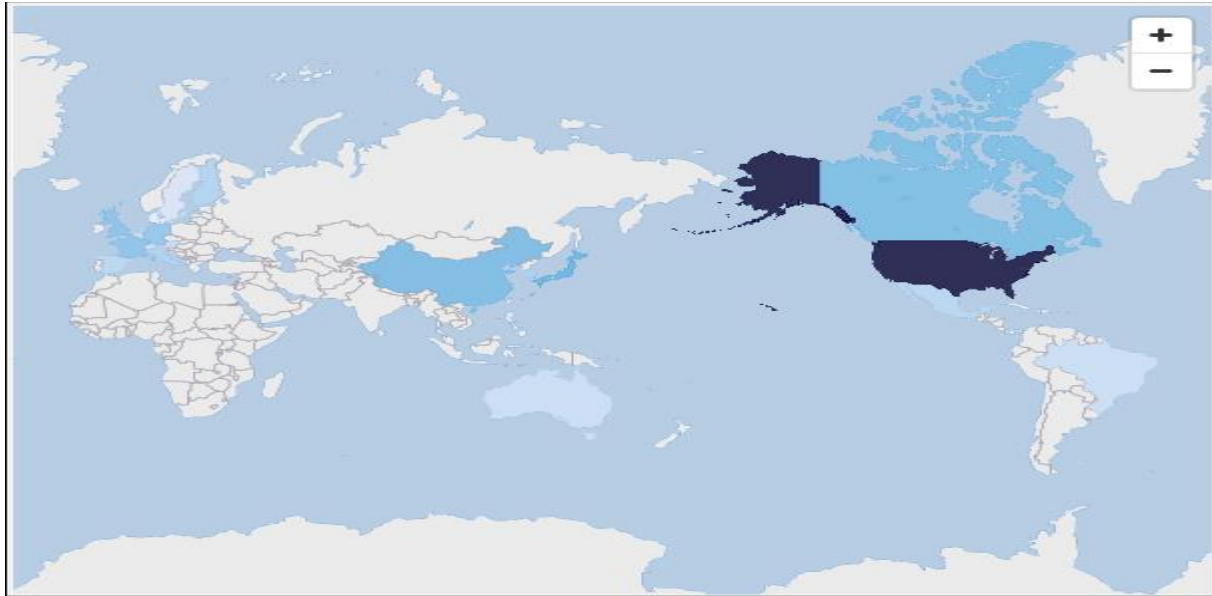
There are many types of data visualization. The most common are **scatter plots, line graphs, pie charts, heat maps, area charts, choropleth maps and histograms.**

CHAPTER 3

3.1 Ideation Phase:

Ideation is the process where you generate ideas and solutions through sessions such as Sketching, Prototyping, Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques. Ideation is also the third stage in the Design Thinking process.

3.1.1 Literature survey

The huge blast of information and Internet gadgets has prompted fast approach of Big Data in later past. Administration industry which is a noteworthy client for these Big Data applications will prompt real change to the conveyance process and new bits of knowledge into utilization example and work processes, which thusly will help with new worldwide conveyance models incorporating new innovations and dispersion of work comprehensively. The study starts with an Endeavour to comprehend contemporary choice emotionally supportive network which is a specific region of data frameworks with an emphasis on enhancing the choice making. Ackoff [2] contemplated that the objective of the administration data frameworks (MIS) was to make the data accessible to administrators for choice making purposes. Sadly, just few MIS were effective as the IT experts of the time did not comprehend the way of administrative work. Analysis of sales data from a company or retailer is a steaming topic at the moment. Numerous records may make up the sales, data, and filtering the data to discover significant insights is a common technique in sales analysis. Managers and owners can visualize aggregated data with the aid of tools like dashboards. Various tools typically display the products that are offered by various points of sale. As early as 1137, visualization was being used to understand data.

Visualization techniques have advanced greatly across all disciplines. Examining data and information can help express and visualize ideas in architecture. Visualization pertinence has increased with the advent of computer simulation. We need scientific visualization techniques, such as charts and graphs, because design data is typically conceptual or unique. Visualization of data is used to present design data with the help of drawings and diagrams. Visualization should be able to communicate effectively and present multidimensional data. It should also be synergistic. Some researchers concentrate on the data analysis tools. They place a lot of emphasis on how simple it is to use the dashboard, establish connections, store data in databases, and share information. Dashboards are frequently used to show the salesperson's revenue on a bar graph with various color-coded bars. The same application also displays sales made by salespeople and sales of products. It is easier to understand insights presented in charts or graphs than it is to rely on spreadsheets or reports because of the way the human brain processes information. Analytics provide a user with a simple, clear, and intuitive way to convey important ideas to all audiences. By making minor adjustments, it is also possible to experiment with various scenarios. According to recent studies, data analytics using visualizations can cut the length of meetings in half. Furthermore, a business intelligence strategy with visualization features boasts a return on investment (ROI) of $13.01 for every dollar invested. Therefore, data analysis and visualization are essential to a business's long-term success and to ensure that it derives the maximum data collection.

3.1.2 Empathy Map

An empathy map is **a collaborative tool teams can use to gain a deeper insight into their customers**. Much like a user persona, an empathy map can represent a group of users, such as a customer segment. The empathy map was originally created by Dave Gray and has gained much popularity within the agile community.
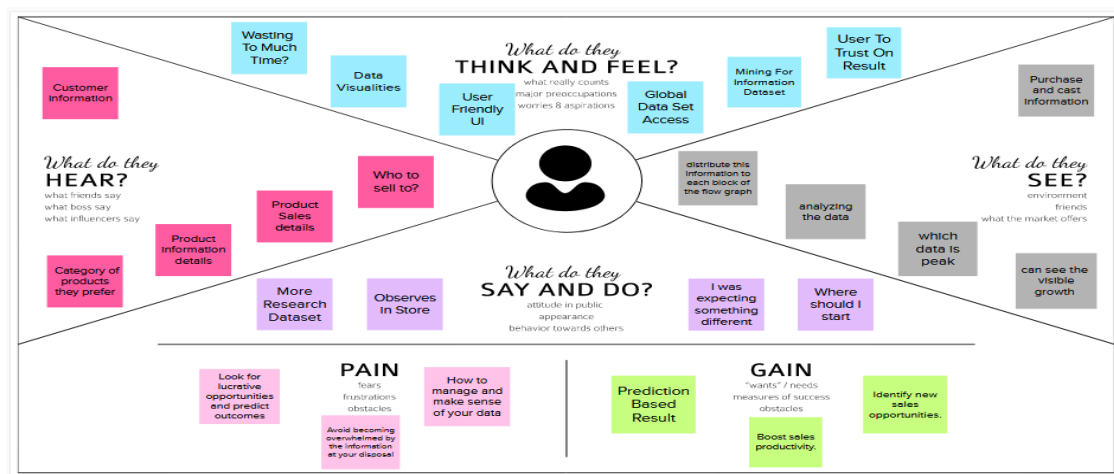
3.1.3 Ideation

Ideation is the creative process of generating, developing, and communicating new ideas, where an idea is understood as a basic element of thought that can be either visual, concrete, or abstract. Ideation comprises all stages of a thought cycle, from innovation, to development, to actualization
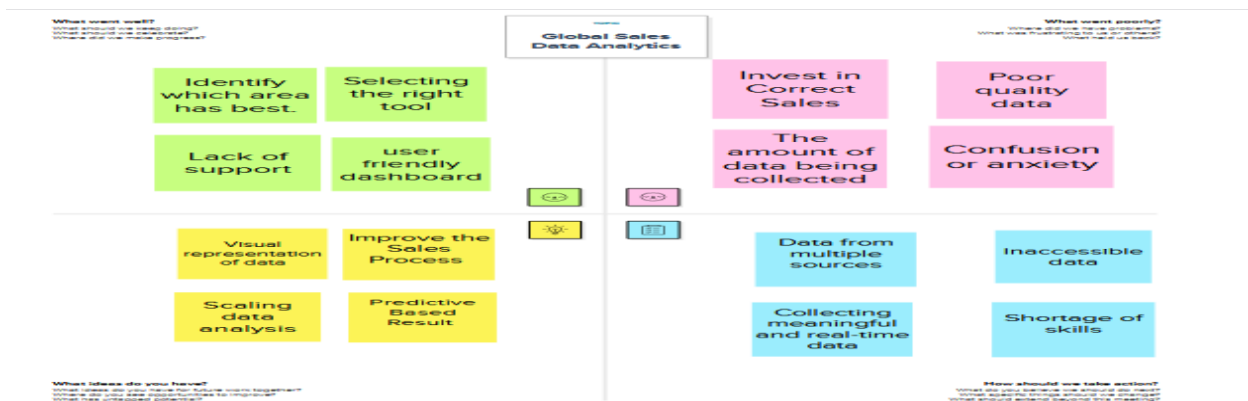
3.1.4 Problem Statement

A problem statement is a concise description of an issue to be addressed or a condition to be improved upon. It identifies the gap between the current state and desired state of a process or product.

| I am	I am trying to..	But	Because	Which makes me feel
customer	Find and analyse the save from varois sources	there are no specification or accuracy in sales	The various sources are from various system	Disappointed
Customer	Prepare all the datas needed	Get errors at the time of data entry	the data will be more complexed	Dissatisfied
Customer	Analyze the datas	Technical issues may occur	there will be poor connectivity and servers	Frustated

3.2.1 Problem-Solution Fit

Problem-solution fit is **a term used to describe the point validating that the base problem resulting in a business idea really exists and the proposed solution actually solves that problem**. Validate that the problem exists: When you validate your problem hypothesis using real-world data and feedback.

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS Direct-to-customer business and E-commerce organization owners willing to improve their sales globally.	6. CUSTOMER CONSTRAINTS CC The datasets need to be collected in an organised form. Decisions has to be made by themselves with all the coclusion produced.	5. AVAILABLE SOLUTIONS AS The dashboard is prepared to display the results of the charts and other analysis. The mannual is also created to help them use the dashboard.	Explore AS, differentiate
Focus on J&P, tap into BE, understand RC	2. JOBS-TO-BE-DONE / PROBLEMS J&P Analysing the data and identifying the trends for improving their sales.	9. PROBLEM ROOT CAUSE RC The urge to improve sales globally and earn profit in the maximum level.	7. BEHAVIOUR BE Data sets of the sales is collected and the analysis made using IBM cognos. Then the trends and patterns are identified.	Focus on J&P, tap into BE, understand RC
Identify strong TR & EM	3. TRIGGERS TR The urge to make appropriate decisions and the fear of not knowing the future of the business. 4. EMOTIONS: BEFORE / AFTER EM Before: Anxiety, Fatigue and depression After: Clear Mind and Peacefullness	10. YOUR SOLUTION SL Creating an interactive, simple and powerful dashboard to interact with the customer in the form of various subscription model.	8.CHANNELS of BEHAVIOUR CH Online: The subscription based services are made to analyse and develop insights	Identify strong TR & EM

3.2.2 Solution Architecture:

Solution architecture is a practice to provide ground for software development projects by tailoring IT solutions to specific business needs and defining their functional requirements and stages of implementation. It is comprised of many sub processes that draw guidance from various enterprise architecture viewpoints.

3.3 Project Design Phase – II

3.3.1Customer Journey:

Customer journey analytics is the science of analysing customer behaviour across touchpoints and over time to measure the impact of customer behaviour on business outcomes.

3.3.2 Data flow diagrams:

A data-flow diagram is a way of representing a flow of data through aprocess or a system. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no controlflow — there are no decision rules and no loops.
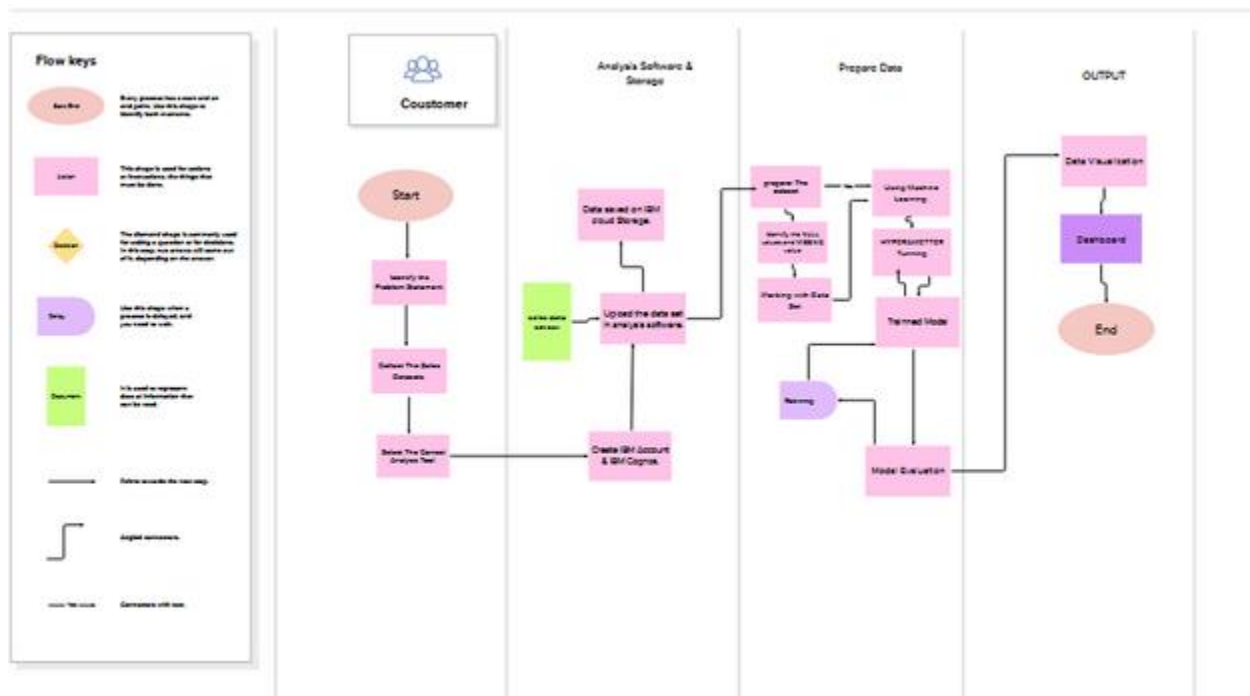
Example: Structural Flow Diagram:
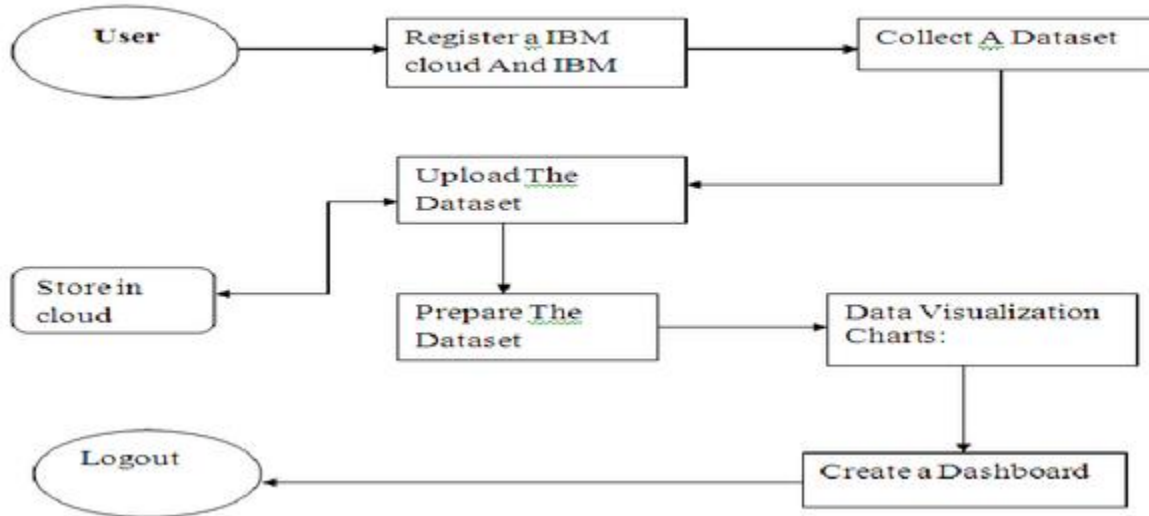
DATA FLOW DIAGRAM FOR GLOBAL

USER

Dashboard

DATASET

Dataset
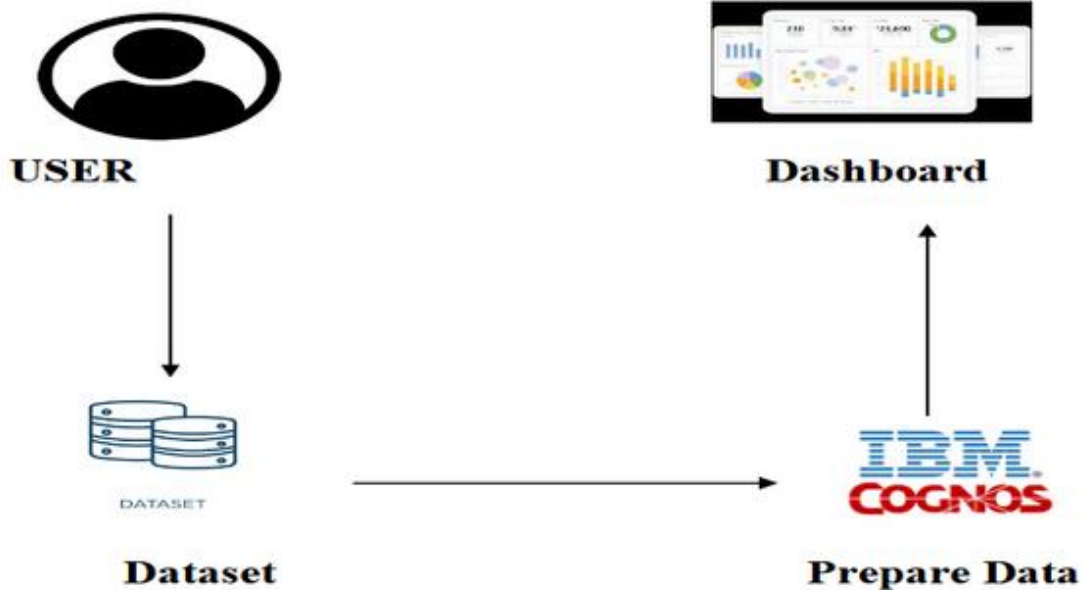
Prepare Data

- User login a IBM cloud and IBM cognos.
- Collect a sales dataset.
- Upload the data set IBM cognos.
- Prepare dataset.
- Visualized data set graphically (graph, chart).
- Crate a template.
- Create a Dashboard

3.3.3.Technical Architecture

Analytics architecture refers to **the systems, protocols, and technology used to collect, store, and analyze data**. The concept is an umbrella term for a variety of technical layers that allow organizations to more effectively collect, organize, and parse the multiple data streams they utilize.

4.1 Dashboard

A dashboard for data analytics is **a tool used to multi-task, organize, visualize, analyze, and track data**. The overall purpose of a data analytics dashboard is to make it easier for data analysts, decision makers, and average users to understand their data, gain deeper insights, and make better data-driven decisions.