APPENDIX 1

ANALYTICS FOR HOSPITALS' HEALTH CARE DATA

IBM-DOCUMENTATION

Industry Mentor(s) Name : Shivam Shivare, Hari Prabu

Faculty Mentos(s) Name   : KEERTHIKA K

TEAM ID : PNT2022TMID12449

Submitted by

Sudarasan K

Sam Alvin I

Muthukumar T

Ajith kumar B M

in partial fulfillment for the award of the degree

Of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

KARPAGAM COLLEGE OF ENGINEERING

ANNA UNIVERSITY: CHENNAI 600 025

NOVEMBER 2022

1. INTRODUCTION

1.1 PROJECT OVERVIEW

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare Management. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital.

This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning.

Suppose you have been hired as Data Scientist of Health Man – a not for profit organization dedicated to manage the functioning of Hospitals in a professional and optimal manner.

1.2 PURPOSE

The goal is to accurately predict the Length of Stay for each patient on case by case basis so that the Hospitals can use this information for optimal resource allocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10 days to more than 100 days.

a) Length of stay for each Case of Patients.

b) Stay by patient ID using Column Chart.

c) Severity of illness by Patient-ID using Tree Map.

d) Age, Department wise patient using Table.

e) Room availability by Pie Chart.

f) Dashboard Creation.

g) Department wise no. of Admissions by Waterfall Chart.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

Hospitals currently use a manual system for the management and maintenance of critical information. The current system requires numerous paper forms, with data stores spread throughout the hospital management infrastructure.

2.2 REFERENCES

1) Nagaraj P, Muneeswaran V, Veera Reddy L, Upendra P, Vishnu Vardhan Reddy M Programmed multi-classification of brain tumor images using deep neural network, 2020

2) Kornelia M. Batko, Andrzej Slezak. The use of Big Data Analytics in healthcare. J. Big Data, 9(1):3, 2022

3) Wantao Yu, Gen Zhao, Yongtao Song, Role of big data analytics capability in developing integrated hospital supply chains and operational flexibility, 2021

4) Care data analytics, Liuliu Fu , Impact of hospital size on healthcare information system effectiveness: evidence from health, 2022

5) Sachin S. Kamble, Angappa Gunasekaran, Milind Goswami & Jaswant Manda, A systematic perspective on the applications of big data analytics in healthcare management, 2019

6) Evarististis Didik Madyatmadja, Implementation of Big Data in Hospital Using Cluster Analytics, 2021

7) Sohali Imran, Tariq Mahmood, Ahsan Morshed, Timos Sellis ,Big Data Analytics in Healthcare ,2021

8) Wei Lei ,Smart Healthcare System using big Data analytics, 2021

9) A Afifhah,Development of the health information analytics Dashboard using big data analytics, 2021

10) N Mehta,Concurrence of big data analytics and healthcare, 2021

2.3 PROBLEM STATEMENT DEFINITION

1. A Patient is a Customer who needs a better treatment because they don't need to suffer.

2. A Patient is a Customer who needs doorstep treatment because it's helpful for the patients, who are aged and can't able to travel.

3. A User needs a way to do something that addressed to the respective patients so that the patient benefits directly.

4. A Patient needs a way to analyze and choose Doctor by their profile because of their experience and comfortable.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

3.2 IDEATION & BRAINSTORMING

Brainstorm & idea prioritization

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.

- 10 minutes to prepare
- 1 hour to collaborate
- 2-8 people recommended

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

🕐 10 minutes

A — Team gathering
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

B — Set the goal
Think about the problem you'll be focusing on solving in the brainstorming session.

C — Learn how to use the facilitation tools
Use the Facilitation Superpowers to run a happy and productive session.

Open article →

---

1. Define your problem statement

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

🕐 5 minutes

> PROBLEM
> How might we [your problem statement]?

Key rules of brainstorming
To run an smooth and productive session

- Stay in topic.
- Encourage wild ideas.
- Defer judgment.
- Listen to others.
- Go for volume.
- If possible, be visual.

3.3 PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to besolved)	To analyze how to reduce the length of patient stay in hospital and to identify the risk of LOS andmake it more convenient.
2.	Idea / Solution description	To use the Cognos analytics, create an interactive dashboard that will answer both the hospitality andpatient queries.
3.	Novelty / Uniqueness	One of the Efficient ways to analyze the problemis Visualization technique. The dashboard will easily provide the insight with the help of patient data.
4.	Social Impact / Customer Satisfaction	It clearly makes a huge impact on Society, Even common people can easily understand the datasetby just simply looking into it.
5.	Business Model (Revenue Model)	Business model can make revenue to the particularhospital by answering the simple queries from the patient's side on how to reduce the risk of LOS.
6.	Scalability of the Solution	The provided Solution is highly Scalable and Feasible as compared to the other common Techniques Such as Data Mining, ML,DL and so on.

3.3 PROBLEM SOLUTION FIT

	1. CUSTOMER SEGMENT(S) CS	6. CUSTOMER CONSTRAINTS	5. AVAILABLE SOLUTIONS	
Define CS, fit into CC	Who is your customer? Patients in all age categories, Hospitals	Inadequate information about the availability of required resources.	Length of stay can be predicted by the tool and patients can plan according to that	**Explore AS, differentiate**

	2. JOBS-TO-BE-DONE / PROBLEMS	9. PROBLEM ROOT CAUSE RC	7. BEHAVIOUR BE	
Focus on J&P, tap into BE, understand RC	Reduce the length of stay of customers	No proper system to track patient's data Less efficient prediction system	Customer can register themselves and analyse the data with available technologies	**Focus on J&P, tap into BE, understand RC**

	3. TRIGGERS TR	10. YOUR SOLUTION SL	8. CHANNELS of BEHAVIOUR CH	
Identify strong TR & EM	rent of hospital rooms which is so difficult to afford.	**Existing:** ratio of discharges in given period of time to no of beds in hospital during the time period **Proposing:** Using predictive analysis powered by AI	8.1 ONLINE Use of data from all the place through data exploration 8.2 OFFLINE Collect data from with the available technologies	
	4. EMOTIONS: BEFORE / AFTER EM			
	BEFORE: Feeling bad and Frustrated AFTER: Feeling better and relaxed			

1. **CUSTOMER SEGMENT**

   Who is your customer?

   A people who was infected or need medical help is our customer.

2. **JOBS-TO-BE-DONE/PROBLEMS**

   Which jobs-to-be-done (or problems) do you address for your customers? Open to all, to book their appointment using website.

3. **TRIGGERS**

What triggers customers to act?

Fear of fitness.

4. EMOTIONS: BEFORE / AFTER

How do customers feel when they face a problem or a job and afterwards?

•Before : Patients feel hereafter they don't need to continue, lose faith, Feel insecure.

•After : Patients were satisfied by selecting their desired doctor and got well by quality treatment.

5. AVAILABLE SOLUTIONS

Which solutions are available to the customers when they face the problem or need to get the job done? Pre-appointed schedule for both patient and doctor, Appointing their comfortable doctor with their profile and fee structure.

6. CUSTOMER CONSTRAINTS

What constraints prevent your customers from taking action or limit their choices of solutions? Spending time in waiting, fear of bill, availability of doctors or not.

7. BEHAVIOUR

What does your customer do to address the problem and get the job done?

Updating them that, Your appointment has x-time left, because the remainder through mails or texting.

8. CHANNELS Of BEHAVIOUR

8.1 ONLINE

What kind of actions do customers take online? Extract online channels from #7

Instead of interacting with a doctor in physical mode, Patient can consult the doctor in online mode in an emergency situation.

8.2 OFFLINE

What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

In Spite of consulting the doctor in online mode, a patient can interact with the doctor would be more adequate.

9. PROBLEM ROOT CAUSE

What is the real reason that this problem exists? What is the back story behind the need to do this job? Not having friendly conversation and care with patient. Waiting time for doctor gets patient antagonize.

Reason:

Receptionist may handle up to limited number of patients in a day but a machine don't have that limitation. So patient can register their appointment through the respective website.

10. YOUR SOLUTION

A Patient needs a way to analyze and choose a Doctor by their profile because of their experience and comfort.

4. IDEATION & PROPOSED SOLUTION

4.1 FUNCTIONAL REQUIREMENTS

Following are the functional requirements of the proposed solution.

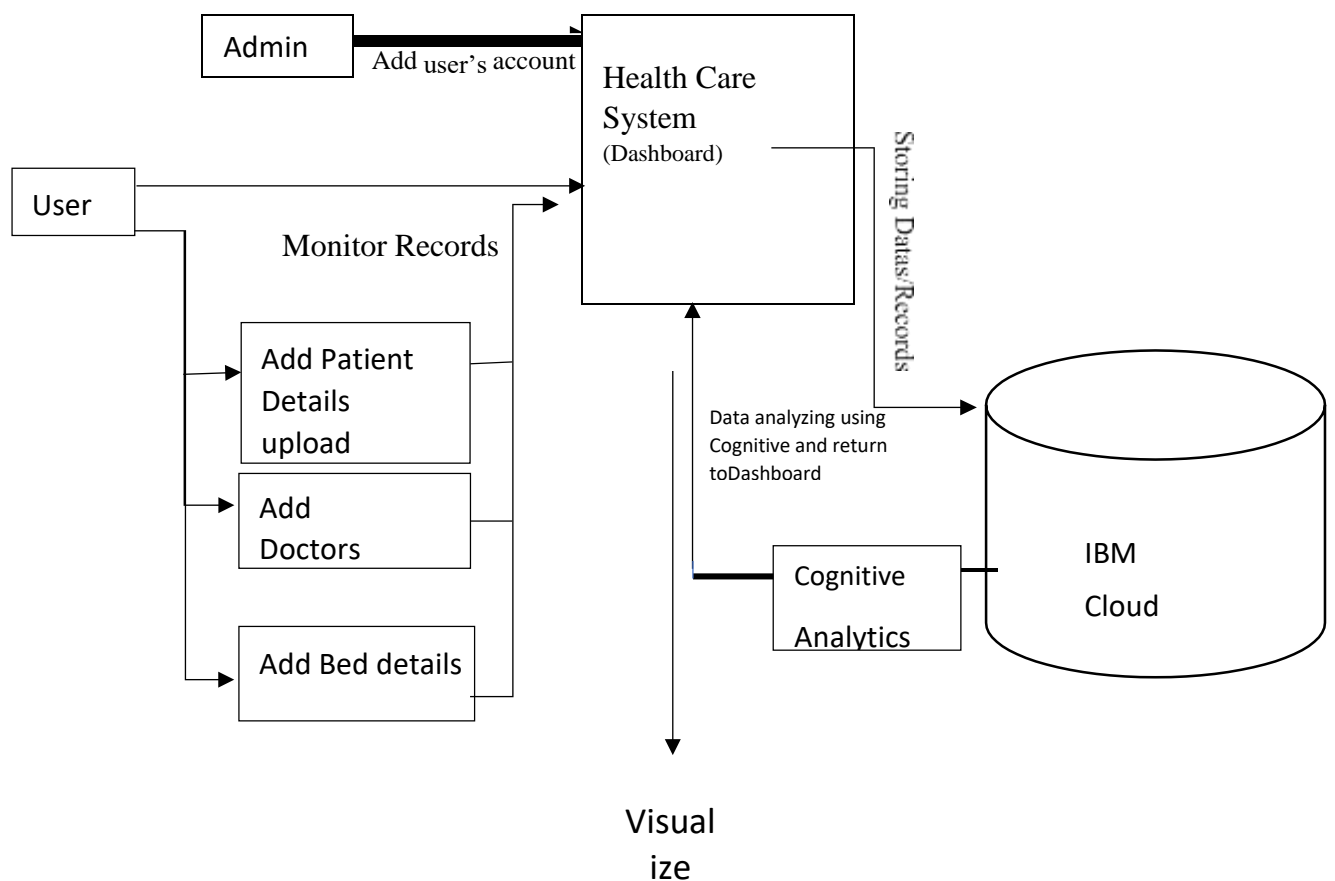| FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through website
FR-2	User Confirmation	Confirmation via Email Confirmation via Message
FR-3	Data Cleaning	We clean the data because there are many potential for data to be duplicated or incorrectly labeled when merging multiple data sources
FR-4	Reliability	Users may utilize this dashboard in an effective, efficient, and reliable manner since it is consistent and reliable for them.
FR-5	Accuracy	Dashboard accurately predicts the patient's health risks based on the length of their stay.

4.2 NON-FUNCTIONAL REQUIREMENTS

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

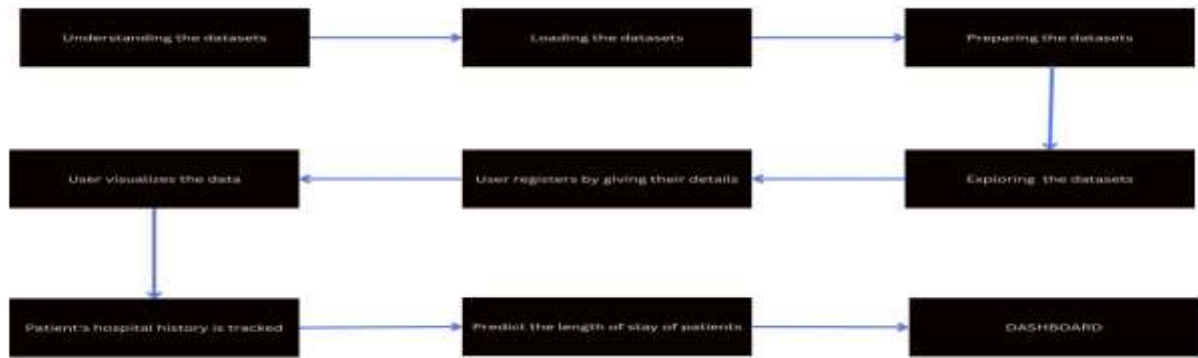| FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	In order to provide a clear understanding of the patient's Length Of Stay, this dashboard makes use of data visualization techniques including charts and graphs.
NFR-2	**Security**	Only users who have the password can access the website. High degrees of security are provided through the use of encryption techniques to secure the database.
NFR-3	**Reliability**	Users will find this dashboard to be constant and dependable, assisting them in using it effectively, efficiently, and dependably.
NFR-4	**Performance**	The project must respond quickly to the user's actions or even if the user has to wait the waiting period must be short.
NFR-5	**Availability**	The project is independent of platforms. On practically every platform, it functions flawlessly.
NFR-6	**Scalability**	The project enables concurrent usage of the data by several people. Because adding features and improving the website is simple, it is very scalable.

5. PROJECT DESIGN

5.1 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.

```
Admin ───Add user's account──→  Health Care
                                 System
                                 (Dashboard)

User ──────────────────────→                    Storing Datas/Records
        Monitor Records

      ┌─────────────┐
      │ Add Patient │                 Data analyzing using
      │ Details     │                 Cognitive and return      IBM
      │ upload      │                 toDashboard               Cloud
      ├─────────────┤
      │ Add         │            ┌──────────┐
      │ Doctors     │            │ Cognitive│
      ├─────────────┤            │ Analytics│
      │ Add Bed     │            └──────────┘
      │ details     │
      └─────────────┘

                   Visual
                    ize
```

1. Patients Based on age   3. Number of Beds
2. Number of Patients

5.2 SOLUTION & TECHNICAL ARCHITECTURE
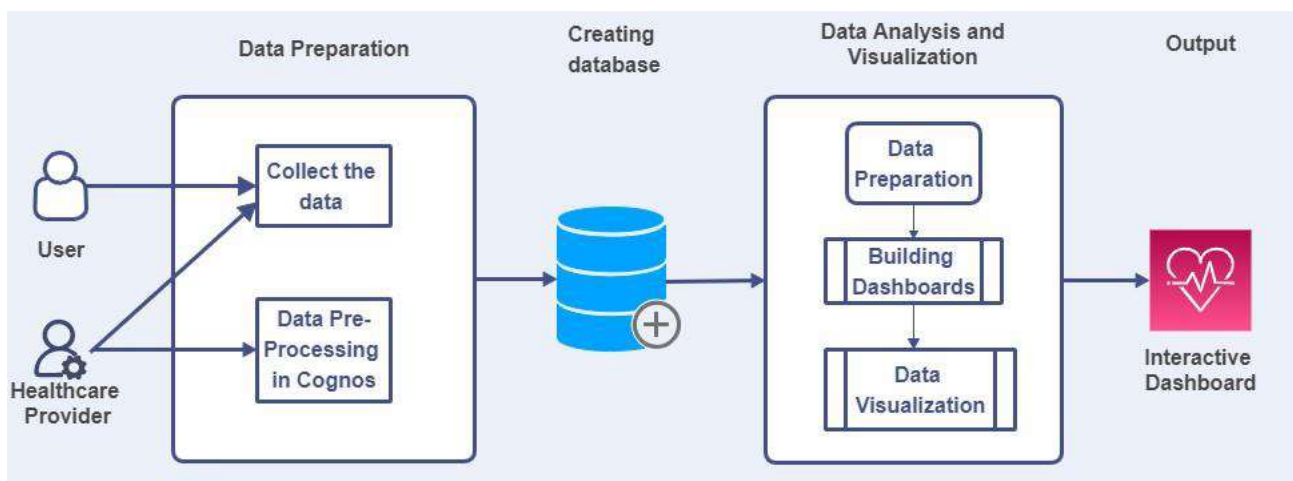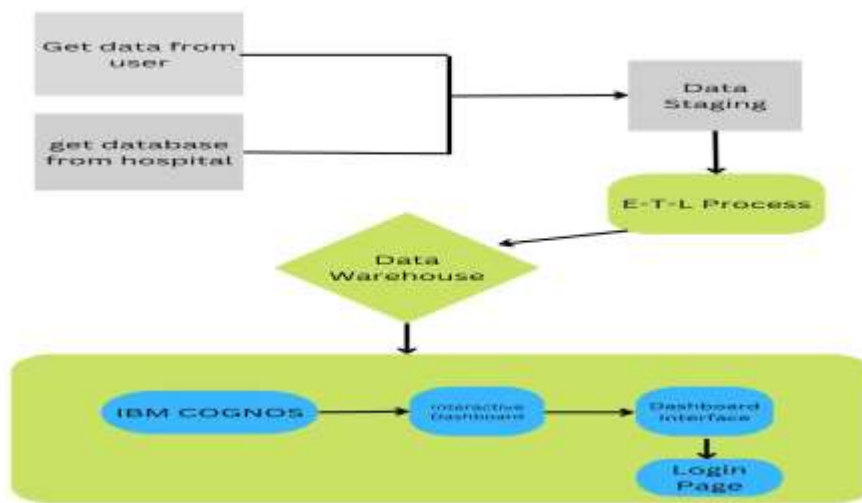
Table-1 : Components & Technologies:

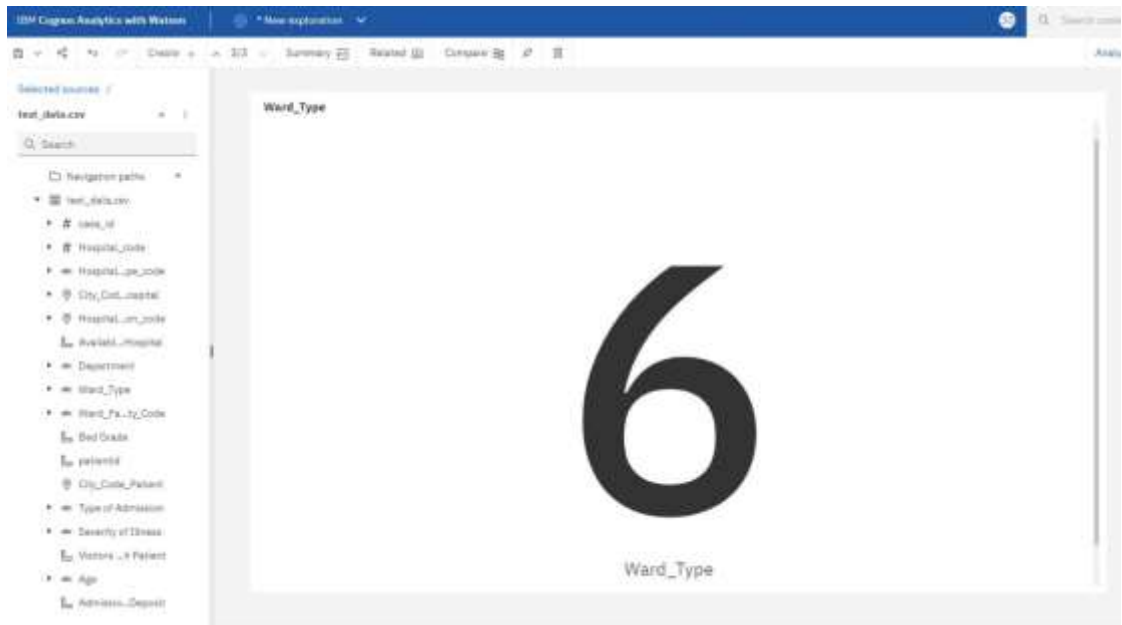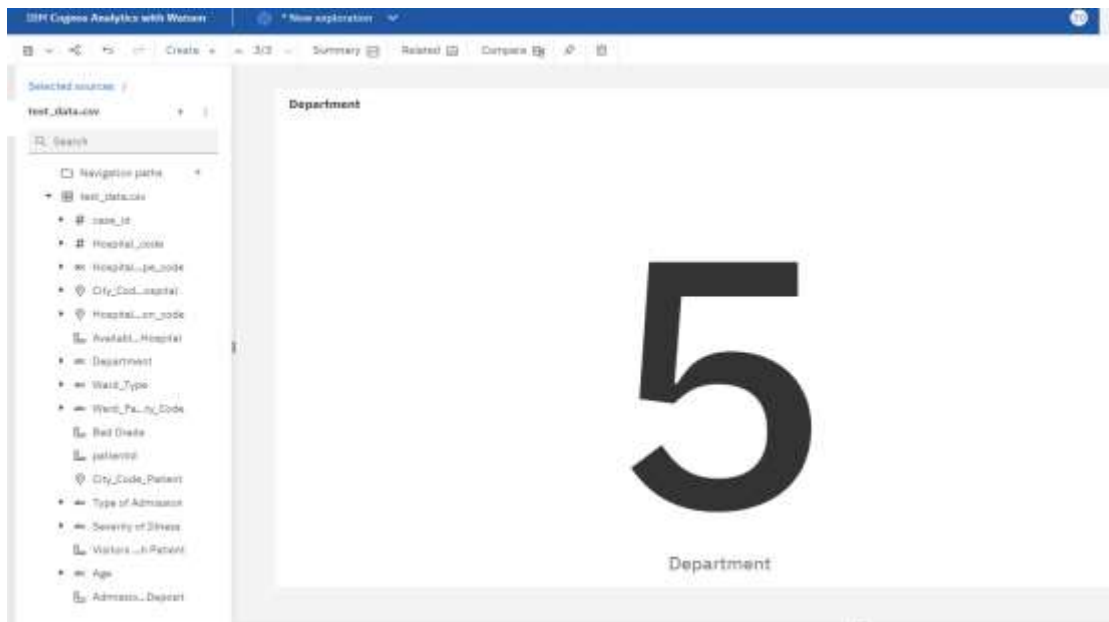| S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g.Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js /React Js etc.
2.	Data Processing	The data from the dataset is pre-processed	IBM Cognos Analytics
3.	Database	Database Service on IBM Cloud	IBM Cloud
4.	File Storage	File storage requirements	IBM Block Storage or Other StorageService or Local Filesystem
5.	Uploading Dataset	Uploading dataset to the IBM Cognos	IBM Cognos Analytics
6.	Data Visualization	The data is visualized into different forms	IBM Cognos Analytics and python
7.	Data Prediction	Prediction of the Length Of Stay (LOS)	IBM Cognos Analytics and python

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Dashboard frameworks in the form of charts,graphs and more.	IBM Cognos
2.	Scalable Architecture	Scalability 3-tier => Database, application, server	Web server – HTML, CSS Application server- Python
3.	Availability	This dashboard is available for health care providers which helps them to managing the hospital resources	IBM Cognos
4.	Performance	This dashboard can contain numerous data and provides interactive visualizing performance whichhelps in easy usage	IBM Cognos
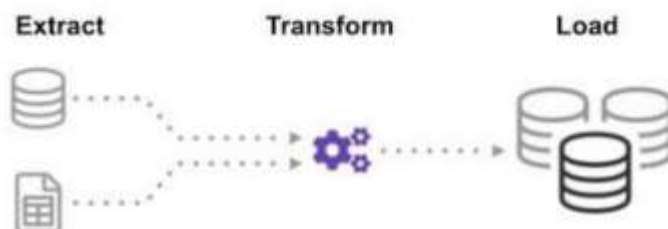
5.3 USER STORIES

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Thyroid Risk Smart Form for Data Entry	User Number-1	As a Clinician, I want to review and update a Thyroid Risk form. So that I can determine my patient's risk of Thyroid (risk category), and ensure proper remedy accordingly.	I can access patient record or data accurately.	High	Sprint-1
	BPA to Prompt Ordering Thyroid on Admission	User Number-2	As an Inpatient, I want to be prompted to Order Thyroid on admission. So that I remember to place my patient on Thyroid.	Maintain the Record for correct preference	High	Sprint-1
	Thyroid	User Number-3	As an Inpatient, I want to view to view only risk-appropriate Thyroid options in Admission Order Sets So that I can ensure my patient is getting optimal Thyroid prophylaxis	For Admission purpose of patient to get solutions among their problem	High	Sprint-2
		User Number-4	As a user, I can access the data in visualize mode		Medium	Sprint-1
	Login	User Number-5	As a user, I need to login to the application by entering email and password		High	Sprint-1
	Dashboard	User Number-6	As a user, I can access the data from Queries, graph, pie chart		High	Sprint-1
Customer (Web user)			Get the older information from the hospitals		Medium	
Customer Care Executive			Have data in graph modes		Medium	
Administrator			Access in Cloud Easily		Medium	

6. PROJECT PLANNING & SHEDULING

6.1 SPRINT PLANNING & ESTIMATION

MILESTONES	TASKS
MILESTONE – 1	Data Collecting process (Datasets)
MILESTONE – 2	Required Datasets are uploaded on the IBM Cognitive Platform.
MILESTONE – 3	Data Exploration and Data Visualization
MILESTONE - 4	To Create a Interactive Dashboard.
MILESTONE - 5	Display the Insights in the Dashboard
MILESTONE - 6	Construct a Standardized Data Set and use the needed data with the Assistance of a Python Program
MILESTONE - 7	Use of different algorithm with Google Colab to achieve the desired result with more accuracy.
MILESTONE - 8	Making the output simpler to understand and more efficient.
MILESTONE - 9	Deployed in the Github

6.2 SPRINT PLANNING & ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Register	USN-1	Data Collecting process (Datasets) Required Datasets are uploaded on the IBM Cognitive Platform	10	Medium	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan
Sprint-1	Login	USN-1	Data Exploration and Data Visualization	20	High	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan
Sprint-2	Dashboard	USN-2	To Create a Interactive Dashboard. Display the Insights in the Dashboard	10	High	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan .
Sprint-2	Dashboard	USN-3	Construct a Standardized Data Set and use the needed data with the Assistance of a Python Program	10	High	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan
Sprint-3	Dashboard	USN-4	Use of different algorithm with Google Colab to achieve the desired result with more accuracy.	20	High	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan
Sprint-4	Virtualizes	USN-5	Making the output simpler to understand and more efficient. Deployed in the Github.	20	High	Ajith Kumar B M Muthukumar Sam Alvin Sudarsanan

6.3 Sprint Delivery Schedule

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project & gathering information by referring the technical papers, research publications , journals etc.	1 SEPTEMBER 2022
Prepare Empathy Map	Prepare Empathy Map Canvas to capture the user Pains & Gains, Prepare list of problem Statements that are to be solved by this project.	7 SEPTEMBER 2022
Ideation	List the ideas by organizing a brainstorming session and prioritize the top 3 ideas based on the feasibility & importance.	14 SEPTEMBER 2022
Proposed Solution	Prepare the proposed solution document, which includes novelty, feasibility of idea, revenue model, social impact, scalability of solution, etc.	21 SEPTEMBER 2022
Problem Solution Fit	Prepare problem - solution fit document.	27 SEPTEMBER 2022
Solution Architecture	Prepare solution architecture document.	29 SEPTEMBER 2022

7.TESTING

7.1 TEST CASES

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

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	0	0	0	0
Client Application	0	0	0	0
Security	0	0	0	0
Outsource Shipping	0	0	0	0
Exception Reporting	0	0	0	0
Final Report Output	0	0	0	0
Version Control	0	0	0	0

7.2 USER ACCEPTANCE TESTING

This report shows the number of resolved or closed bugs at each severity level, and how they wereresolved.

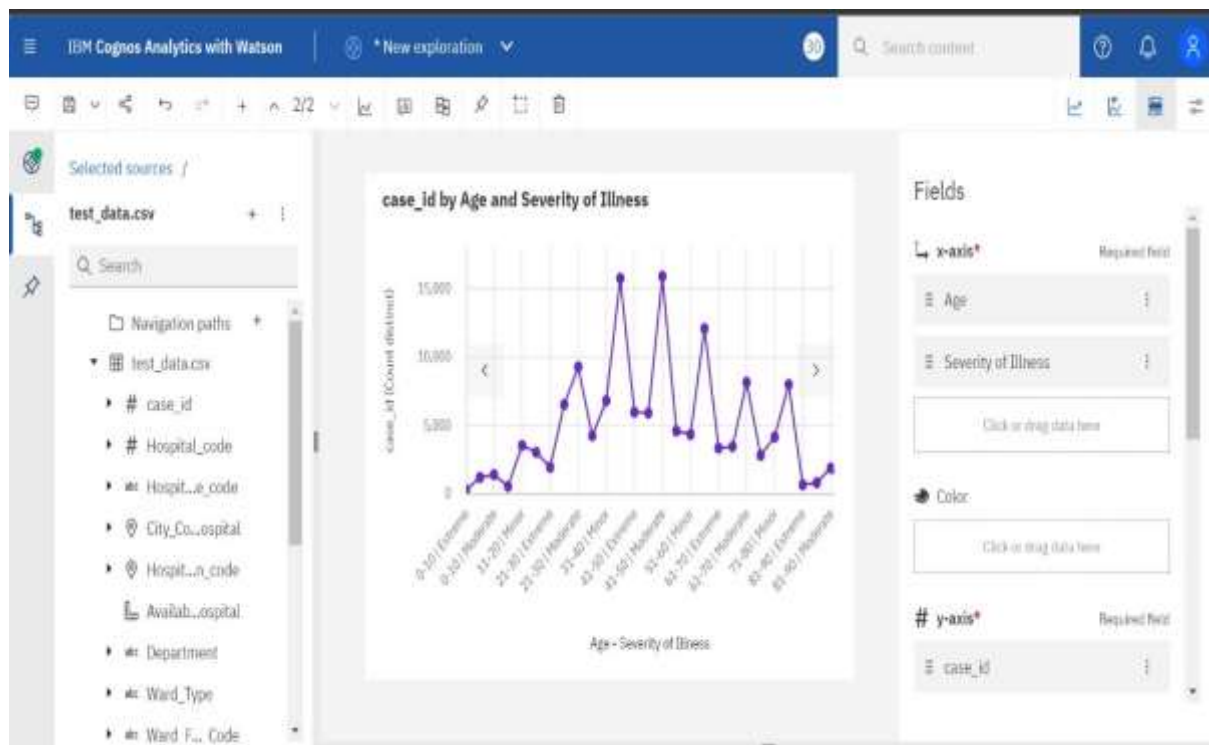| Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	5	5	3	0	13
Duplicate	0	0	0	0	0
External	7	5	1	0	13
Fixed	11	8	7	5	31
Not Reproduced	1	0	0	0	1
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	24	18	11	5	58

8. RESULT

● Dashboard for Number of Patients
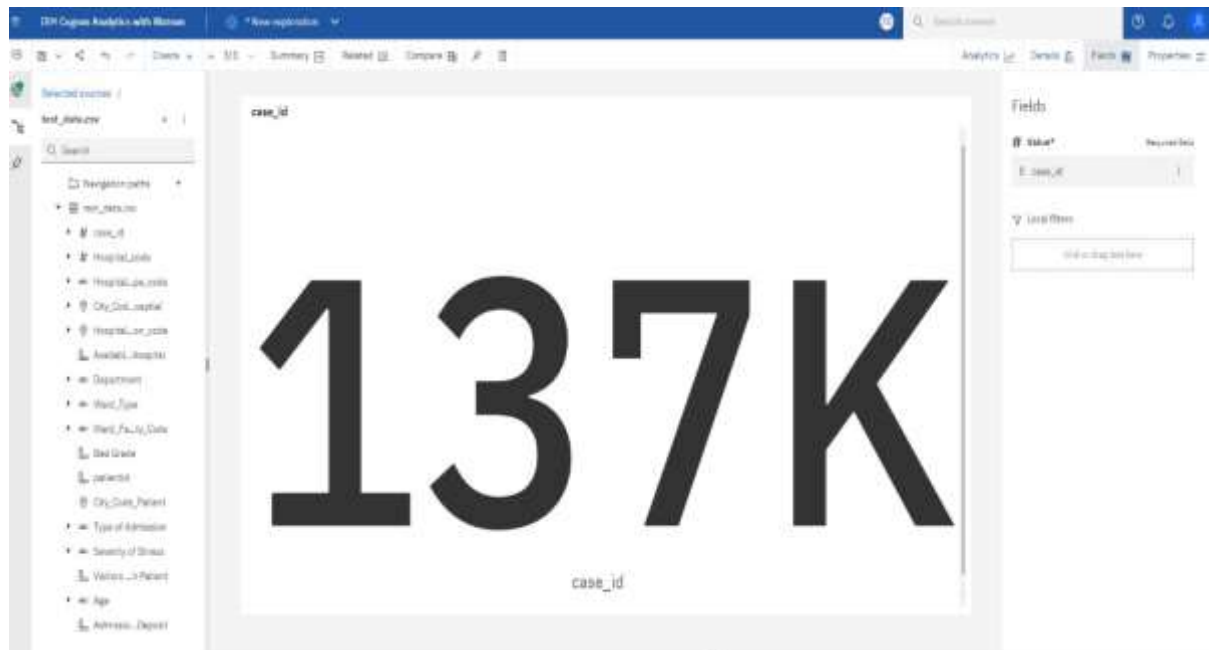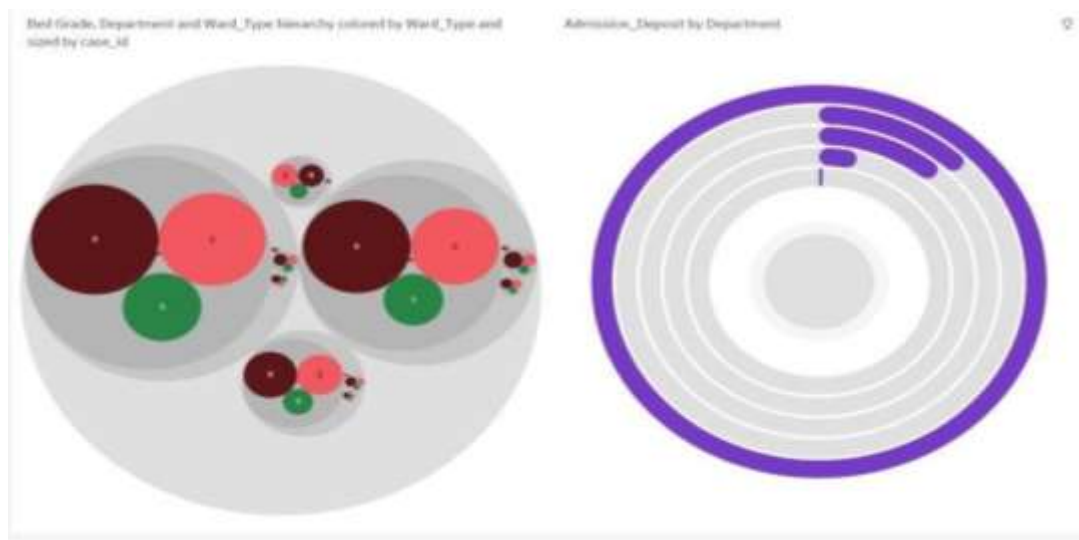

● Dashboard for Number of ward Types
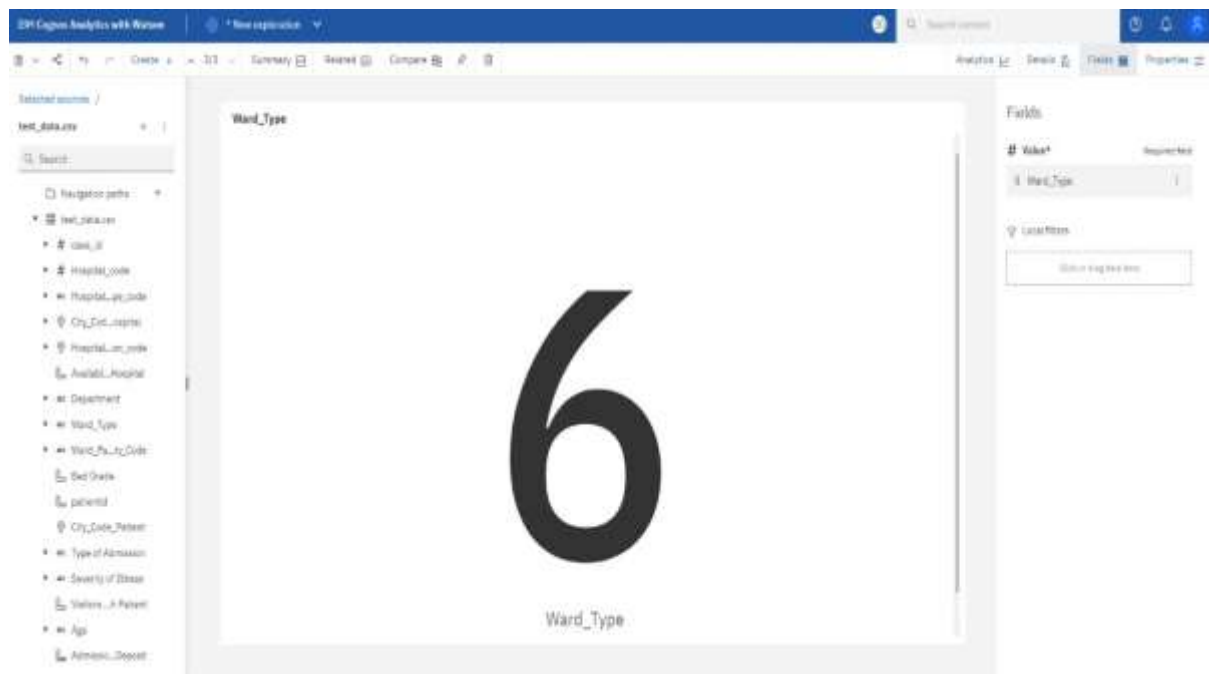
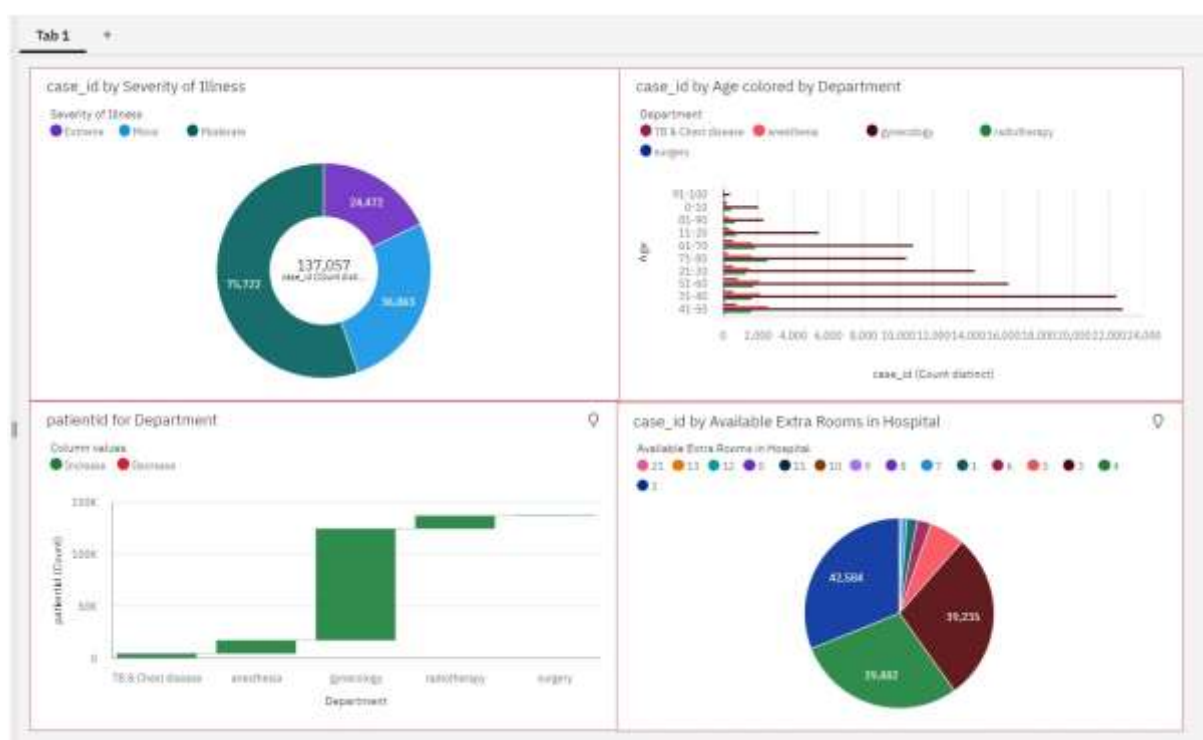● Age wise Patients with Department And Severity Filters

● Dashboard with Hierarchy Bubble And Radial Visuals
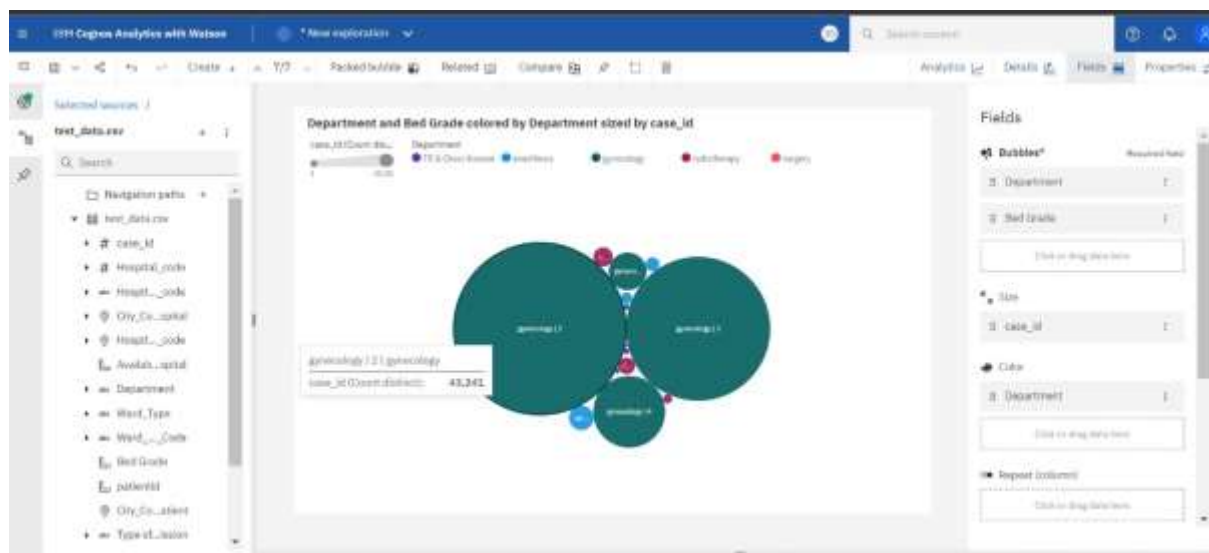
● Hierarchy Bubble and Radial Charts

● Dashboard Showing Pie, Stacked Bar, Waterfall And Pie Charts

9. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- Helps an organization to make a better decision
- Increase the efficiency of the work
- The analytics keeps you updated of your customer behavioral changes.
- Personalization of hospital details.
- Improving quality of service and health care.

DISADVANTAGES:

- Lack of alignment within teams
- Lack of commitment and patience
- Low quality of data
- Privacy Concerns
- Complexity and Bias

10.CONCLUSION

Data analytics in health care is vital. It helps health care organizations to evaluate and develop Number of patients by ward, Age wise patients with department details, Various types of visualizations to analyze the hospital's datasets and hence predict outbreaks in illness, Data analytics can also lower costs for health care organizations and boost business intelligence.

11.FUTURE SCOPE

While every fact of the industry stands to be changed by data analytics in healthcare, data has significantly improved healthcare in three areas: conducting medical studies, understanding the cost of medical tests and health insurance, and making preventative recommendations to patients.

Hospital Healthcare data analytics helps in analyzing the patient details via hospital that the availability of doctors and number of beds to the patients and hence, it reduce the man power and time of the respective Hospital.

12. FUTURE SCOPE

Github

https://github.com/IBM-EPBL/IBM-Project-14980-1659592896