

# Analytics for Hospitals Health-Care Data

Team ID	PNT2022TMID43551
Project name	Analytics for Hospital Health Data
Team Members	1.Subramanyam Shashikumar Ayyar (TL) 2.Anirudh NK 3.Niyas N 4.Vishnunathan R

## 1. INTRODUCTION

### 1.1 Project Overview

The healthcare industry is multidimensional, with multiple data sources involving healthcare systems, health insurers, clinical researchers, social media, and government, generating different types and massive amounts of data. It is impossible to handle this big data with traditional software and hardware and the existing storage methods and tools. Data analytics is the process of the analysis of data to identify trends and patterns to gain valuable insights. The data generated in the health industry are characterized by the four Vs of big data, namely volume, velocity, variety, and veracity, which play crucial roles in health data analytics. Also, evidence-based decision making has gained importance, which involves the sharing of data among various data repositories. According to Deloitte Global Healthcare Outlook, it is expected that global healthcare expenditure will continue to increase at an annual rate of 5.4% between 2017 and 2022. This is due to the increased importance of personalized medicine, the use of advanced technologies, the demand for new payment models, improvement and expansion of care delivery sites, and competition. Various research attempts, based on big data, have provided strong evidence that the efficiency of healthcare applications is dependent upon the basic architecture, techniques, and tools used. Statistical data and reports can be generated with the use of patient records, aiding in knowledge discovery,

and thereby influencing value-added services to the patients, improving healthcare quality, the making of timely decisions, and minimizing the costs incurred. Hence, there is a need to incorporate and integrate big data analytics into existing healthcare systems. Despite healthcare analytics having massive potential for value-added change, there are many technological, social, organizational, economic, and policy barriers associated with its application.

## **1.2 Purpose**

The main aim of this paper is to provide a deep analysis on the research field of healthcare data analytics, as well as highlighting some of guidelines and gaps in previous studies. This study has focused on searching relevant papers about healthcare analytics by searching in seven popular databases such as google scholar and springer using specific keywords, in order to understand the healthcare topic and conduct our literature review. The paper has listed some data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical operations, reports, decision making, and prediction and prevention system. Moreover, the systematic review has showed an interesting demographic of fields of publication, research approaches, as well as outlined some of the possible reasons and issues associated with healthcare data analytics, based on geographical distribution theme.

## 2. LITERATURE SURVEY

### 2.1 Existing Problem

The list below highlights a number of the major problems that we face today :-

- Data management, security and privacy issues. Issues such as data integrity and privacy lead to poor data management Privacy violation and discrimination. Disclosure of Personal Health Information is also a major risk.
- Technological issues, Lack of required infrastructure cannot produce safe conclusions Social inequality, as data are only open to a small elite of technical specialists who know how to interpret and use it, and to those who can employ them
- Skilled Resource set, There is a need to have a Data scientist and Data analyst to perform big data analysis. There is already a huge shortage in the required skill set for Big Data Analytics.
- Data Ownership, There is a lot of big data flowing which includes genomics, remote sensing, social media, mobile app and many other data types.
- Healthcare Models, There is a need to have sufficient business case evidence in health to measure investment return.
- Limited awareness and support, It will cause lack of funding and awareness. Dependency on private funding will support a few big players that will further lead to international economic competitiveness. Funding models have to be revisited to ensure better care.

### 2.2 References

- 1)Mohammad Alkhatib , Amir Talaei-Khoei (University of Nevada,Reno)Amir Talaei-Khoei University of Nevada, Reno | UNR · Department of Accounting and Information Systems PhD of Information Systems-Amir Ghapanchi
- 2)From:"Book of Data Analytics" Chandank Reddy(Wayne State University) Charu C.Aggarwal(Watson Research Center)

3) From: Hoyt,RE,Yoshihashi,A,Eds.(2014).Health Informatics:Practical Guide for Healthcare and formation Technology Professionals,Sixth Edition.Pensacola,FL,Lulu.com.

4)Panagiota Galetsia , Korina Katsaliakia , Sameer Kumarb,\* a School of Economics, Business Administration & Legal Studies, International Hellenic University, 14th km Thessaloniki-N. Moudania, Thessaloniki, 57001, Greece b Opus College of Business, University of St. Thomas Minneapolis Campus, 1000 LaSalle Avenue, Schulze Hall 435, Minneapolis, MN 55403, USA

5)from"n book: Innovative Data Communication Technologies and Application (pp.83-96)" P. Nagaraj-Professor (Assistant) at Kalasalingam University

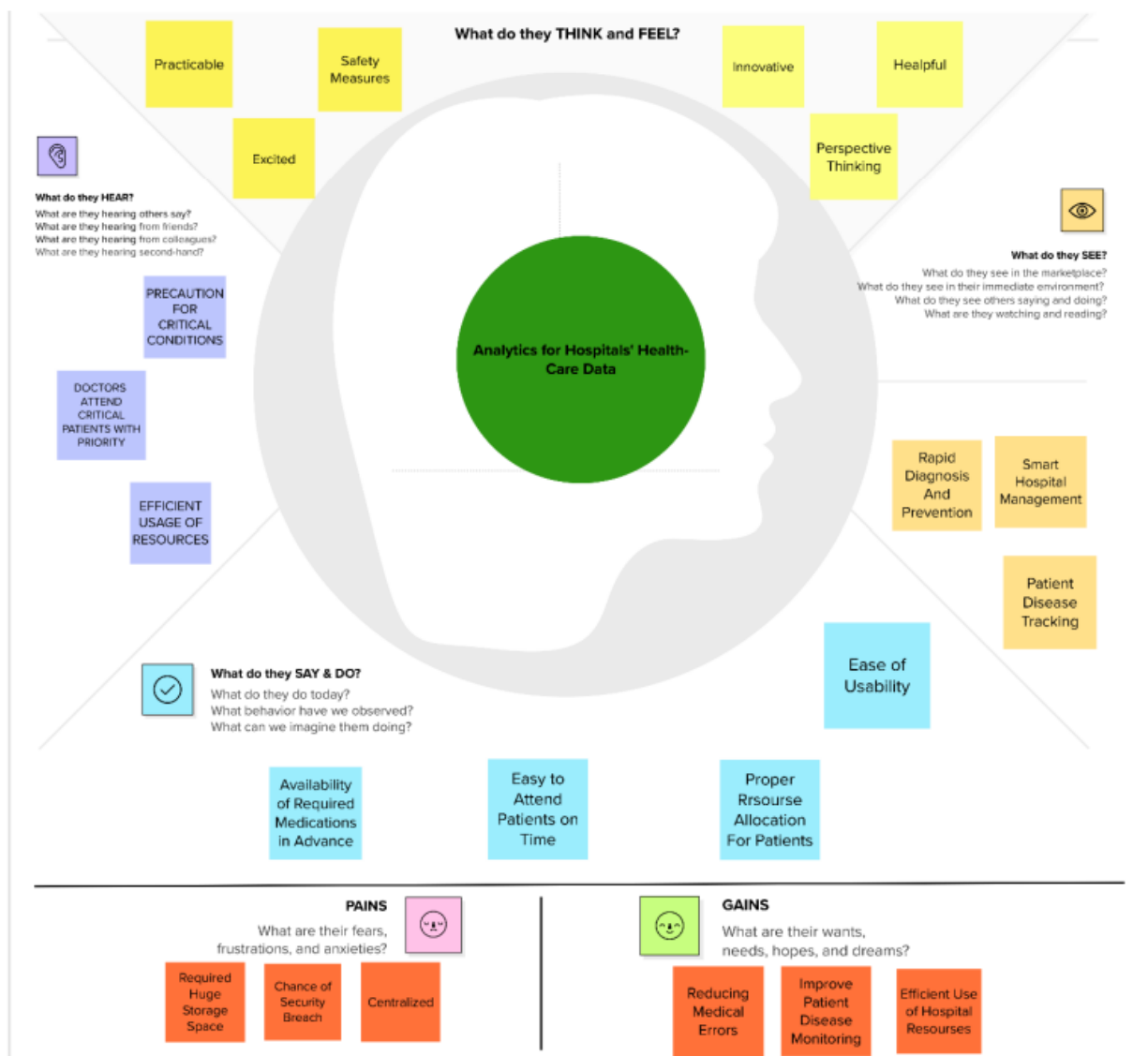
## **2.3 Problem Statement Definition**

- Collection dataset.
- Upload the dataset into cognos.
- Open the properties->data module.
- If null value is present in character field use mode method.
- If the null value is present in continuous field use average or medium.
- Display the data in respective charts.
- Create conclusion using summary.

### 3. IDEATION & PROPOSED SOLUTION

#### 3.1 Empathy Map Canvas


- An empathy map is a tool which aids in understanding another person's perspective.
- Empathy maps have up until now not been used in a medical education setting
- Objective: To assess the attitudes towards, applicability and usefulness of empathy maps as part of medical student's communication skills training.



## 3.2 Ideation and Brainstorming

To try to solve a problem or come up with new ideas by having a discussion that includes all members of a group : to discuss a problem or issue and suggest solutions and ideas.

Template



### 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

[Share template feedback](#)

➔

#### 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 a smooth and productive session

🗣️ Stay in topic.

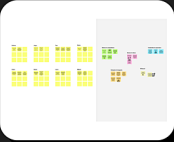
💡 Encourage wild ideas.

🕒 Defer judgment.

👂 Listen to others.

🗣️ Go for volume.

👁️ If possible, be visual.



#### Need some inspiration?

See a finished version of this template to kickstart your work.

[Open example](#) ➔

2

## Brainstorm

Write down any ideas that come to mind that address your problem statement.

10 minutes

**TIP**  
You can select a sticky note and hit the pencil icon to start drawing!

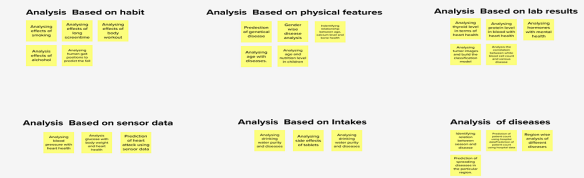


3

## 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 can break it up into smaller sub-groups.

20 minutes

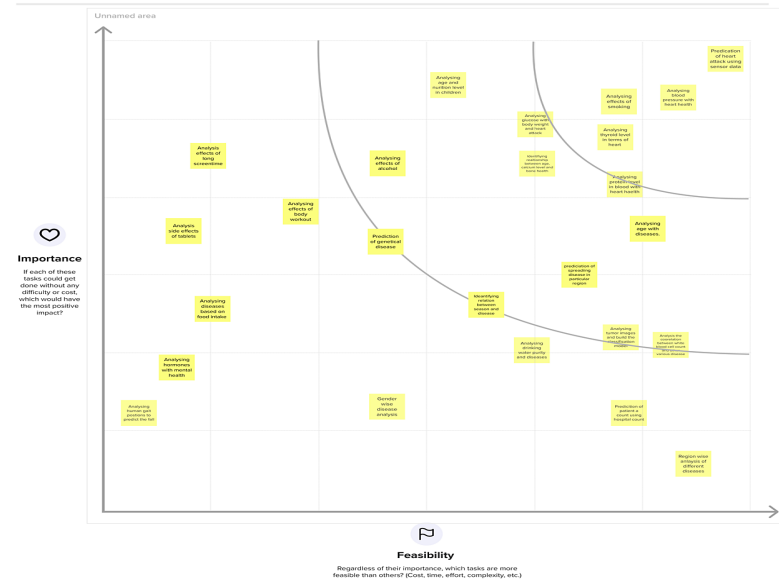


4

## 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.

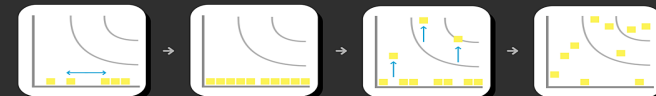
### Quick add-ons

- Share the mural**  
Share a view link to the mural with stakeholders to keep them in the loop about the outcomes of the session.
- 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)	During the covid-19 pandemic, we have faced one of the difficult times of our life. Everyone seeks to survive from the great disaster. At the time of pandemic, no one get to know about which hospital has vacant beds (free beds) to admit themselves or others infected by covid. This situation made the death rate higher.
2.	Idea/Solution description	Predictive analytics can create patient journey dashboards and disease trajectories that helps us to know about the patient's period of stay. It improves effective allocation of beds and other resources, treatment delivery, improves efficiencies, and so on.
3.	Novelty / Uniqueness	Healthcare data frequently resides in several locations. The Collected data should be stored in central system (like centralized storage). This data becomes accessible and usable when it is combined into a single, central system, such as an enterprise data warehouse (EDW). Uniqueness of our project is that we can able to use data for different things such as which medicine is more effective and for understanding behavioural pattern of particular disease.



4.	Social Impact / Customer Satisfaction	<ul style="list-style-type: none"> <li>• Effective use of resource</li> <li>• Enhanced diagnosis</li> <li>• Improved Treatment</li> <li>• Enhancing the overall quality of treatment and life of patients</li> </ul>
5.	Business Model (Revenue Model)	With the gathered data, redirecting the patients to particular hospital based on the vacancy, leading retailers used methods like market-basket analysis to discover insights about consumer purchase behaviour and used these insights to optimize the physical store experience, target relevant ads and streamline the supply chain, among other strategic initiatives
6.	Scalability of the Solution	A variety of institutions must store, evaluate, and take action on the massive amounts of data being produced by the health care sector as it expands quickly. India is a vast, culturally varied nation with a sizable population that is increasingly able to access centralised healthcare services.

### 3.4 Solution Fit

- The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem
- In an age where medical science has made noteworthy advancements, inefficiencies and healthcare errors are still persistent because of the healthcare industry's traditional technology for management.

- One specific area of concern is the exchange of patient data in case of patient transfer from one department or hospital to another. Patient record sharing, when done the traditional way, is time-consuming and inefficient and exposes patient information to a breach.
- To deliver a holistic and satisfactory patient experience, different parties involved in healthcare – doctors, scheme providers, insurance providers, doctors, and patients – should be able to exchange information among themselves securely.

Analytics For Hospitals' Health-Care DataAnalytics For Hospitals' Health-Care Data			TEAM ID-PNT2022TMD43551		
Define CS, fit into CC	<b>1. CUSTOMER SEGMENT(S)</b> <small>Who is your customer? i.e. working parents of 0-5 y.o. kids</small>  <b>CS</b>  Hospital management and patients.	<b>6. CUSTOMER CONSTRAINTS</b> <small>What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.</small>  <b>CC</b>  Not able to predict the patient LOS properly especially during the pandemic period	<b>5. AVAILABLE SOLUTIONS</b> <small>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 &amp; cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking</small>  <b>AS</b>  Effective hospital bed management using data mining technique	Explore AS, differentiate	
	<b>2. JOBS-TO-BE-DONE / PROBLEMS</b> <small>Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.</small>  <b>J&amp;P</b>  Need the proper data analysis of allocation of beds and other needs of patients	<b>9. PROBLEM ROOT CAUSE</b> <small>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 regulations.</small>  <b>RC</b>  Insufficient analysis in data ,human error and poor scheduling.	<b>7. BEHAVIOUR</b> <small>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)</small>  <b>BE</b>  Regularly monitoring the database of patients and measures to avoid error	Focus on J&P, tap into BE, understand RC	
Identify strong TR & EM	<b>3. TRIGGERS</b> <small>What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.</small>  <b>TR</b>  Prevailing emergency situations and Pandemic period situations and Pandemic period	<b>10. YOUR SOLUTION</b> <small>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.</small>  <b>SL</b>  Using predictive analysis powered by the AI which is used in analytics technique Proper Data analysis and implementation in Interactive dashboard.	<b>8. CHANNELS of BEHAVIOUR</b> <b>8.1 ONLINE</b> <small>What kind of actions do customers take online? Extract online channels from #7</small>  <b>CH</b>  Secure login ,Usage of data exploration	Extract online & offline CH of BE	
	<b>4. EMOTIONS: BEFORE / AFTER</b> <small>How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure &gt; confident, in control - use it in your communication strategy &amp; design. BEFORE: Unstable physical and psychological state during the pandemic period AFTER : Physical and psychological comfort and security to the patients. Improved critical care bed allocation decisions.</small>  <b>EM</b>		<b>8.2 OFFLINE</b> <small>What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development</small>  Preparing the data set on the patients occupancy period, predicting the LOS with doctors		

## 4. REQUIREMENT ANALYSIS

### 4.1 Functional requirement

No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
1	User Registration	Utilizing a Form for Registration, Signing up with Gmail
2	User Confirmation	Email confirmation required
3	Interoperability	A dashboard makes it possible to quickly and inter-operably transmit patient information with hospitals.
4	Accuracy	Based on LOS (Length of Stay), the dashboard accurately predicts the patient's health risks.
5	Compliance	The use of a dashboard for compliance by hospitals is quite dynamic and takes place in real time.
6	Concise	These dashboards are easy to understand, simple to customize, and interactive.

### 4.2 Non-functional Requirements

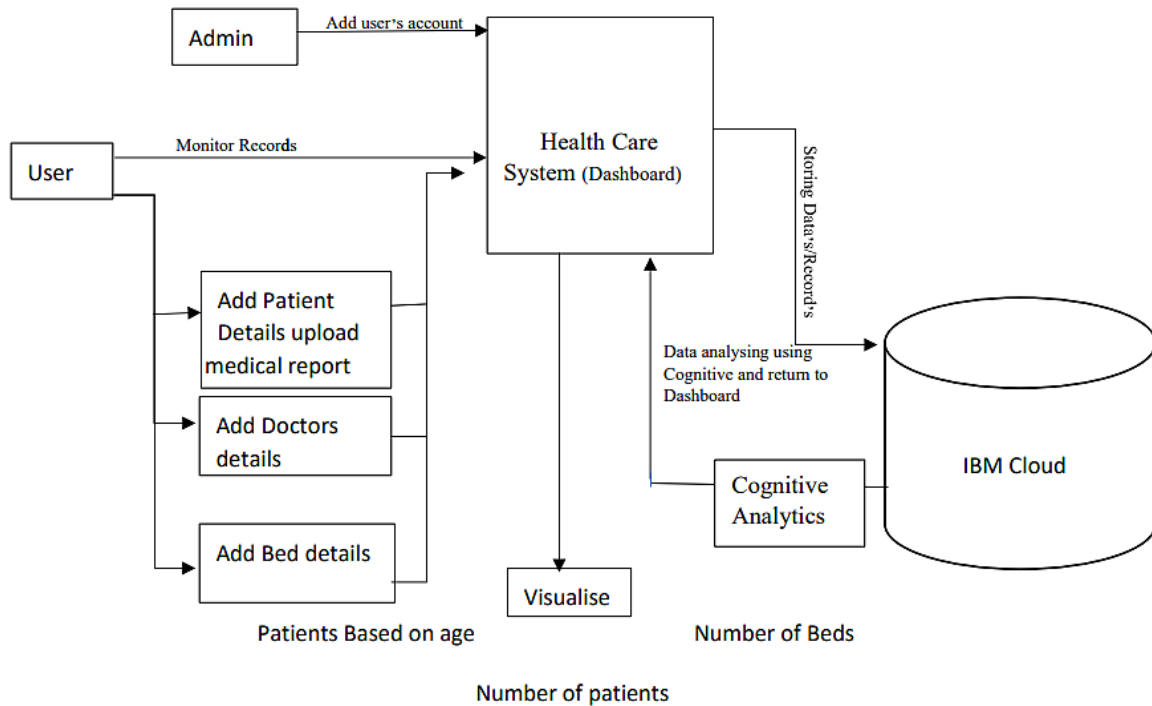
No.	Non-Functional Requirement	Description
1	Usability	In order to provide a through visual representation of the patient's LOS, this dashboard makes use of data visualization techniques including charts and graphs.
2	Security	The Dashboard aids in indicating the level of threat that currently exists for the hospitals, as well as past occurrences and incidents, authentication mistakes, scans, probes, and unwanted access.

3	Reliability	Users will find this dashboard to be consistent, dependable, and helpful in using in an effective, efficient, and dependable manner.
4	Performance	This dashboard may scan backend users, and examining how frequently they visit the dashboard might reveal relevant information about the jobs the data is beneficial for.
5	Availability	The dashboard is able to promptly satisfy user needs and aids in giving the user's dataset the relevant information.
6	Scalability	A hosted feature layer, feature layer view, or hosted tile layer are the layers that are used in the dashboard.

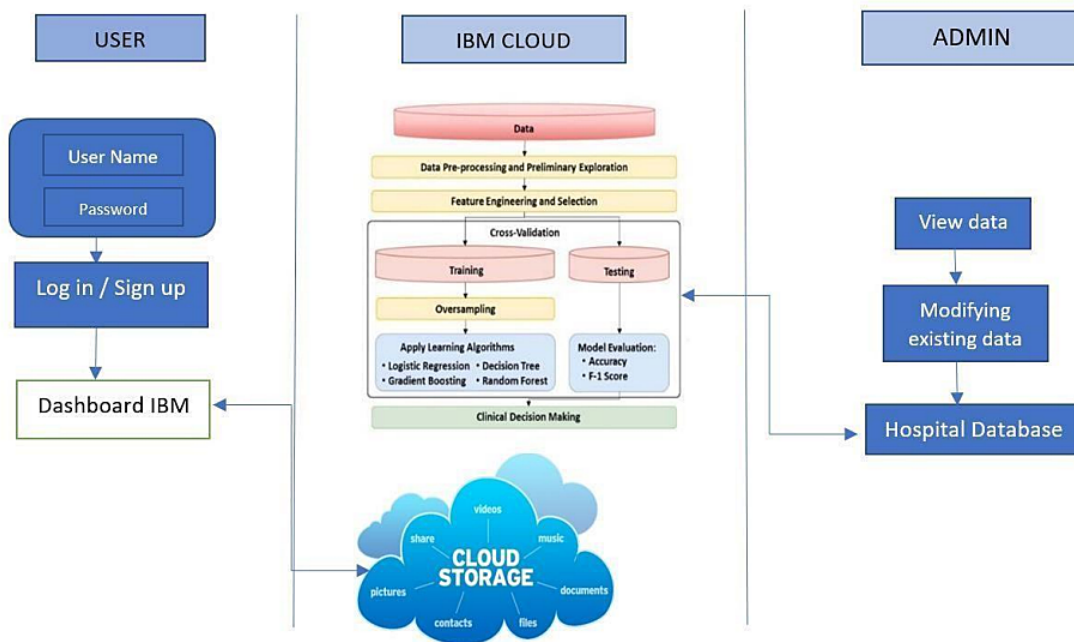
## 5. PROJECT DESIGN

### 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.



## 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	Application Logic-1	Logging in as a patient / user in the application	Python
3	Application Logic-2	Logging in as an admin in the application	IBM Watson Assistant
4	Database	All the data about patients such as disease, address and etc.	MySQL, NoSQL, etc
5	Cloud Database	IBM Watson cloud is	IBM DB2, IBM Cloudant

		used for storage, Cloud	etc.
6	External API-1	Purpose of External API used in the application	Aadhar API, etc..
7	Machine Learning Model	Purpose of Machine Learning Model	Regression Model, etc.
8	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration, Cloud Server Configuration	Local, Cloud Foundry, Kubernetes, etc.

**Table-2: Application Characteristics:**

S.No	Characteristics	Description	Technology
	Open-Source Frameworks	List the open-source frameworks used	Python
	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	Encryption.
	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Can supports higher workloads
	Availability	Justify the availability of application (e.g., use of load balancers, distributed servers etc.)	Highly available
	Performance	Design consideration for the performance of the application (number of	It performs good uses various tools and ideas in a scientific manner to

		requests per sec, use of Cache, use of CDN's) etc.	meet the desired outcomes
--	--	--	---------------------------

### 5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Administrator (Admin)	Login		As an admin, I can add users details.	I can provide username and password for the users.	High	Sprint-1
User	Login	USN-1	As a user, I can login into the website/ application using username and password	I can access Dashboard and user account.	High	Sprint-1
	Dashboard	USN-2	As a user, I can add Patient Details like Patient name, contact number, age etc.	I can view and edit patient details	Medium	Sprint-1
	Dashboard	USN-3	As a user, I can add bed details, Doctor details and other hospital detail.	I can monitor Staff and hospital performance.	Medium	Sprint-1
	Dashboard	USN-4	As a user, I can upload patient medical reports.	I can upload and Modify patient reports	High	Sprint-1



## 6. PROJECT PLANNING & SCHEDULING

### 6.1 Sprint Planning & Estimation

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	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

### 6.2 Sprint Delivery Schedule

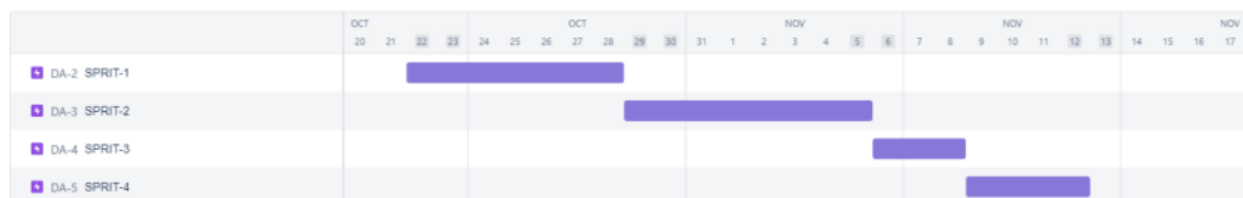
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a health care provider, I can create account in IBM cloud and the data are collected.	20	High	2 Members
Sprint-2	Analyse	USN-2	As a health care provider all the data that are collected is cleaned and uploaded in the database or IBM cloud.	20	Medium	2 Members

Sprint-3	Dashboard	USN-3	As a health care provider, I can use my account in my dashboard for uploading dataset.	10	Medium	2 Members
Sprint-3	Visualization	USN-4	As a health care provider, I can prepare data for Visualization.	10	High	2 Members
Sprint-4	Visualization	USN-5	As a health care provider, I can present data in my dashboard.	10	High	2 Members
Sprint-4	Prediction	USN-6	As a health care provider, I can predict the length of stay	10	High	2 Members

## 6.3 Reports from JIRA

### Road Map:

A roadmap is a strategic plan that defines a goal or desired outcome and includes the major steps or milestones needed to reach it. It also serves as a communication tool, a high-level document that helps articulate strategic thinking—the why—behind both the goal and the plan for getting there.



### Kanban Board:

A kanban board is an agile project management tool designed to help visualize work, limit work-in-progress, and maximize efficiency (or flow). It can help both agile and DevOps teams establish order in their daily work.

## **7. CODING & SOLUTIONING**

### **7.1 Feature 1**

- Fetched the data from DB2 database.
- Creating responsive dashboard.
- Inserting filter for each chart
- Creating report
- Created reports using multiple graphs and charts

### **7.2 Feature 2**

- Creating stories and performed.
- Perform animation render image from website.
- Included graphs and charts.
- Creating web application using bootstrap.
- Embedded the cognos with web application.

### **7.3 Database Schema**

- case\_id
- Hospital\_code
- Hospital\_type\_code
- City\_Code\_Hospital
- Hospital\_region\_code
- Available Extra Rooms in Hospital
- Department
- Ward\_Type
- Ward\_Facility\_Code
- Bed Grade
- Patient id
- City\_Code\_Patient
- Type of Admission

- Severity of Illness
- Visitors with Patient
- Age
- Admission\_Deposit
- Stay

## 8. TESTING

### 8.1 Test Cases

- Verify user is able to see Home page.
- Verify user is able to see Dashboard page.
- Verify user is able to navigate to Report page.
- Verify user is able to navigate to story page.
- Verify filters are working

### 8.2 User Acceptance Testing

#### 1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the [ProductName] project at the time of the release to User Acceptance Testing (UAT).

#### 2. Defect Analysis

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

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	8	5	0	3	16
Duplicate	1	0	5	0	6
External	0	3	2	1	6
Fixed	13	4	3	16	36
Not Reproduced	0	1	0	0	1
Skipped	0	1	0	1	2
Won't Fix	1	4	2	1	8
Totals	23	18	12	22	75

### 3. Test Case Analysis

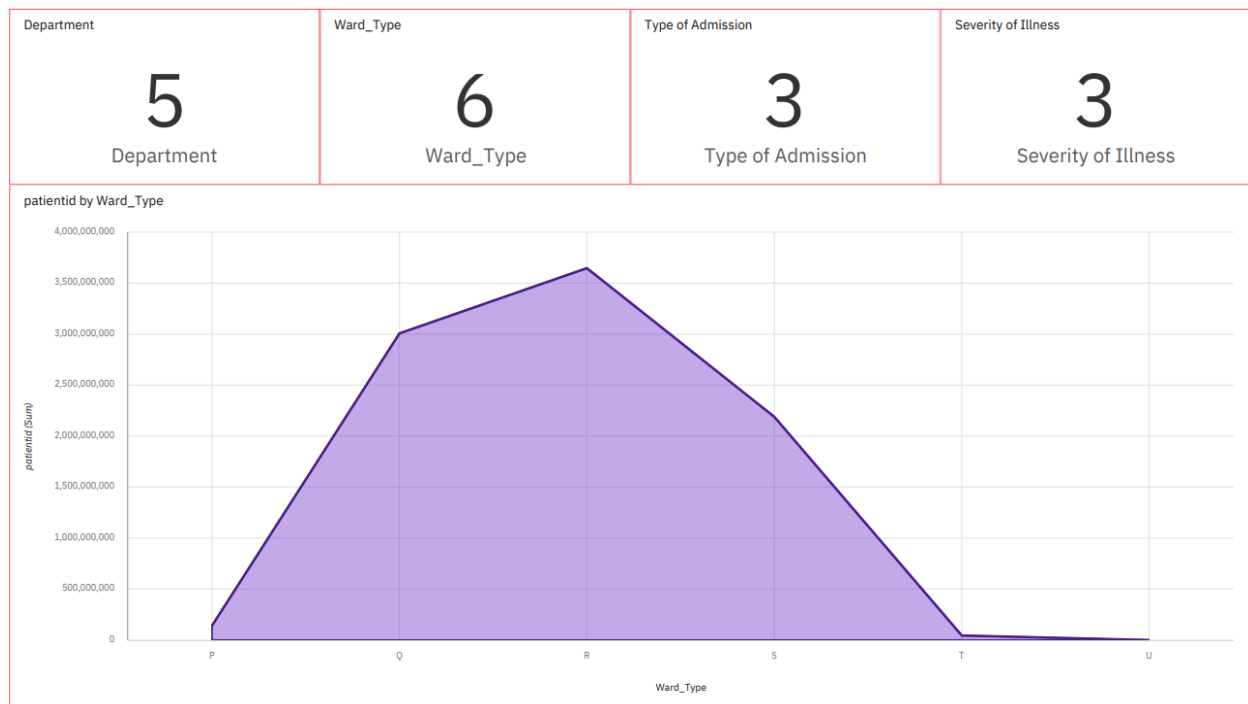
This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	9	0	0	9
Client Application	43	0	0	43
Security	1	0	0	1
Outsource Shipping	1	0	0	1

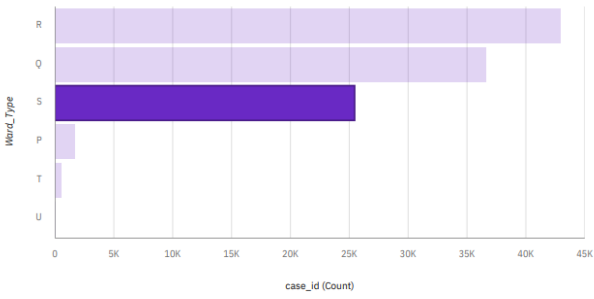
Exception Reporting	9	0	0	9
Final Report Output	10	0	0	10
Version Control	1	0	0	1

## 9. RESULTS

### 9.1 Performance Metrics

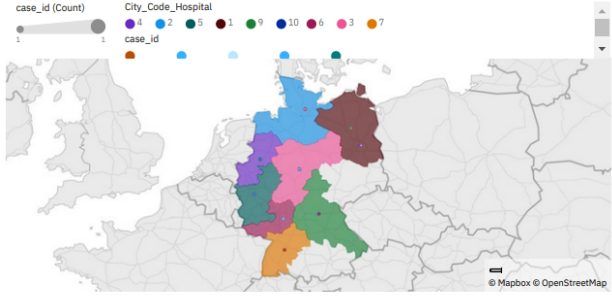


case\_id by Ward\_Type



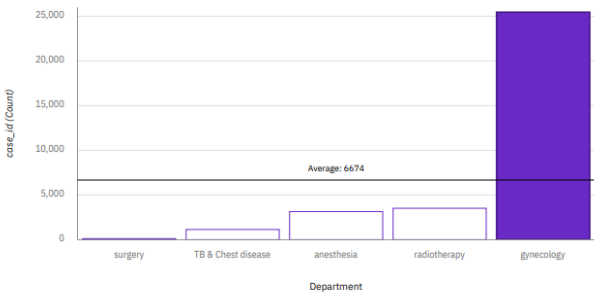
1

City\_Code\_Hospital for City\_Code\_Hospital regions and case\_id and case\_id for City\_Code\_Hospital points



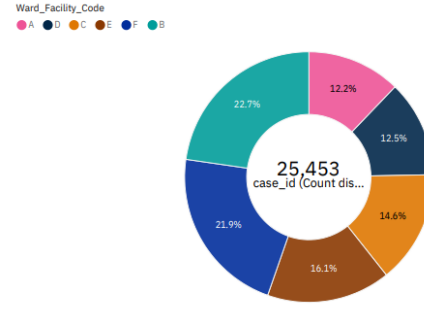
2

case\_id by Department



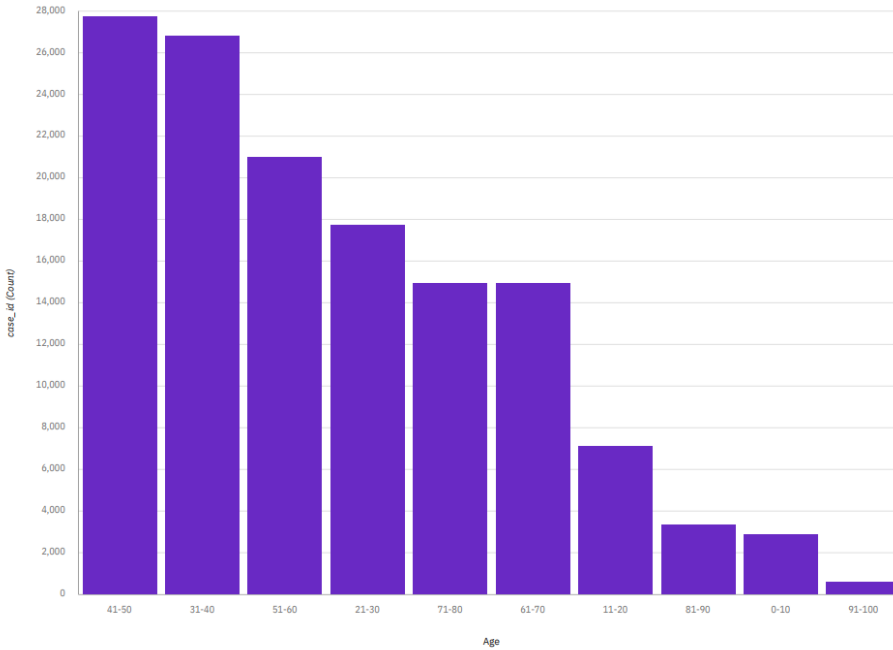
3

case\_id by Ward\_Facility\_Code



4

case\_id by Age

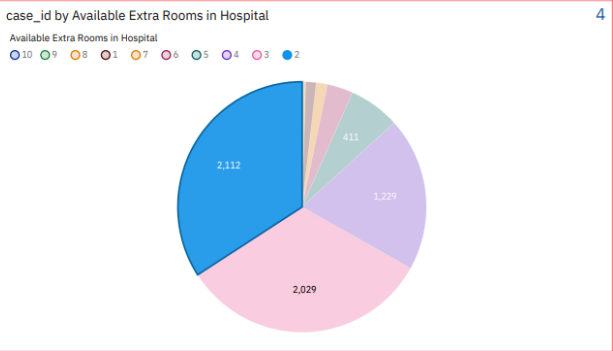
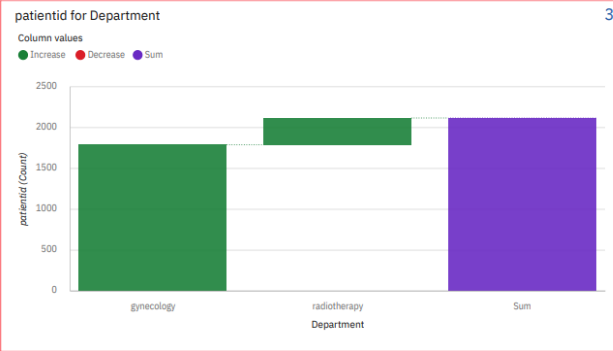
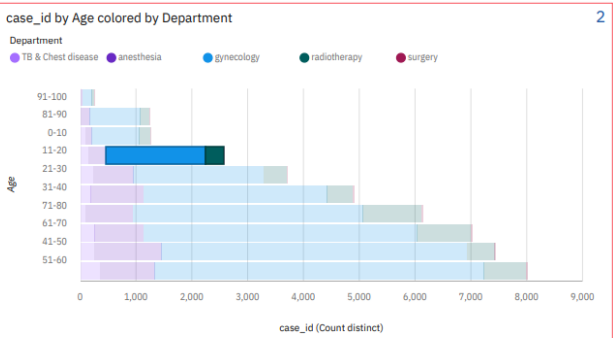
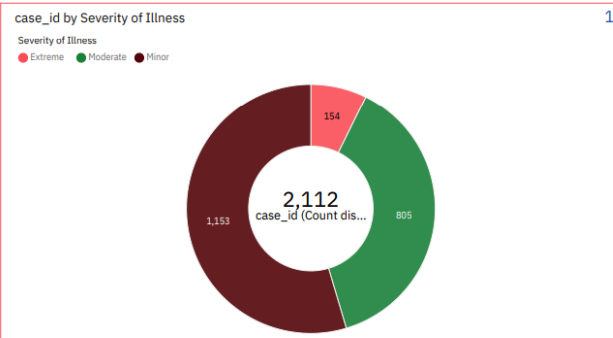
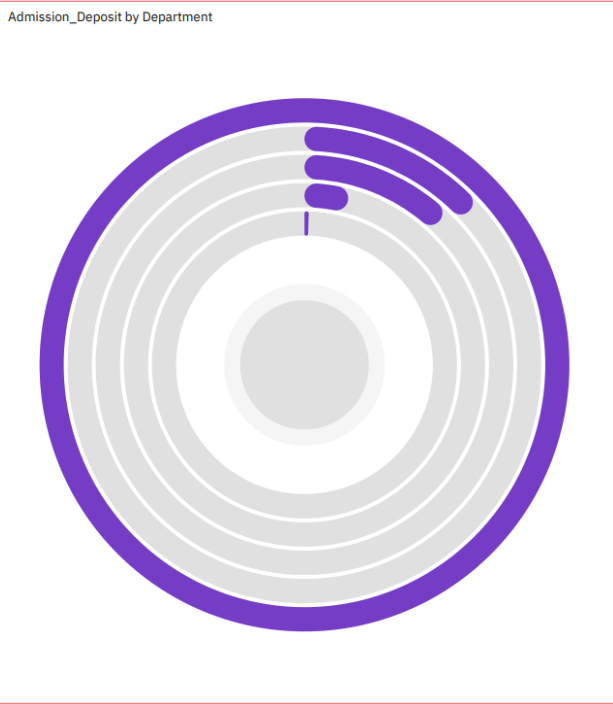
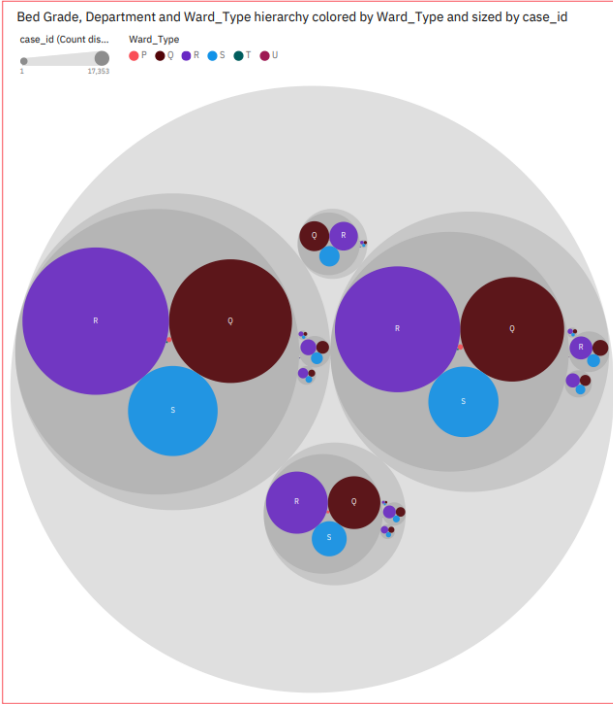


Department

Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery

Severity of Illness

Severity	of
Illness Extreme	
Minor	
Moderate	



## 10. ADVANTAGES

- Improved research efforts
- Improved health outcomes
- Obtain operational insights
- Improved staffing
- Informed strategic
- planning
- Higher-Quality Care

## DISADVANTAGES

- Privacy
- Replacing Doctors
- Frustration with poor implementation.
- Cybersecurity risks
- Healthcare Regulatory Changes.
- Healthcare Staffing Shortages

## 11. CONCLUSION

- It also means describing how health plans, health care organizations, and clinicians should be accountable to patients and society and conversely.
- How individuals can take appropriate responsibility for their own health.
- Data analytics is the science of analyzing raw datasets in order to derive a conclusion regarding the information they hold.
- It enables us to discover patterns in the raw data and draw valuable information from them.

## 12. FUTURE SCOPE

- **Improved Decision Making:** Data Analytics eliminates guesswork and manual tasks. Be it choosing the right content, planning marketing campaigns, or



developing products.

- Organizations can use the insights they gain from data analytics to make informed decisions. Thus, leading to better outcomes and customer satisfaction
- Data analytics to achieve business goals of pharmaceutical companies, payers, insurance companies, physicians, hospitals, medical equipment companies, sales reps, and other stakeholders in the healthcare business, need for this have only increased after the Affordable Act came into being.

## **13. APPENDIX**

### **GitHub link:**

<https://github.com/IBM-EPBL/IBM-Project-49175-1660816520>

### **Project Demo Link:**

<https://clipchamp.com/watch/f4XZRMQo9eT>