Project Report

Analytics for Hospitals' Health-Care Data

1.Introduction

1.1 Project overview

Healthcare organizations are under increasing pressure to improve patient care outcomes and achieve better care. While this situation represents a challenge, it also offers organizations an opportunity to dramatically improve the quality of care by leveraging more value and insights from their data. Health care analytics refers to the analysis of data using quantitative andqualitative techniques to explore trends and patterns in the acquired data. While healthcare management uses various metrics for performance, a patient's length of stay is an important one.

Being able to predict the length of stay (LOS) allows hospitals to optimize their treatment plans to reduce LOS, to reduce infection rates among patients, staff, and visitors.

1.2. Purpose

The goal of this project is to accurately predict the Length of Stay for each patient so that the hospitals can optimize resources and function better.

2.Literature survey

2.1Existing problem

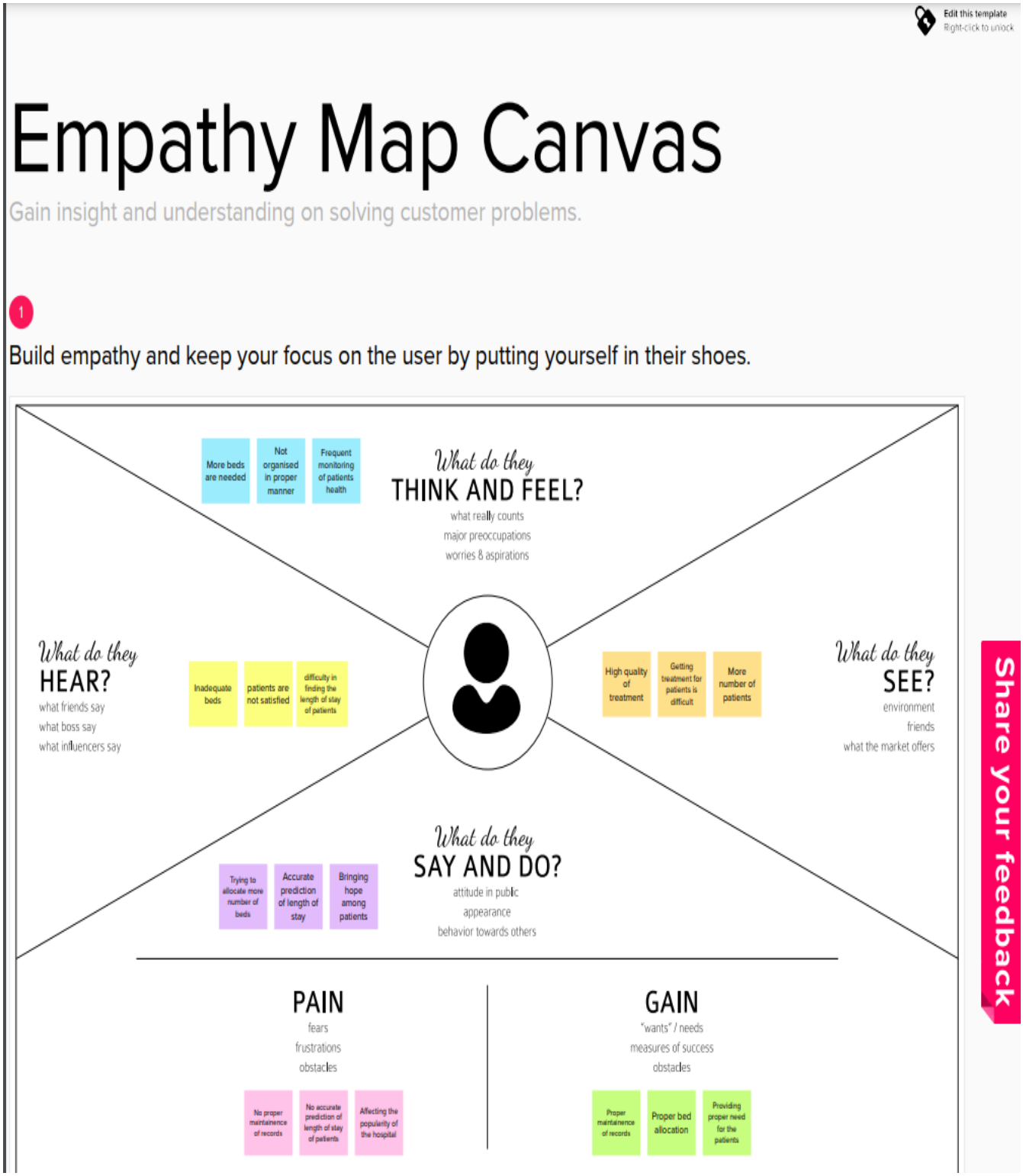
Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare Management. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of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

2.2. Problem statement

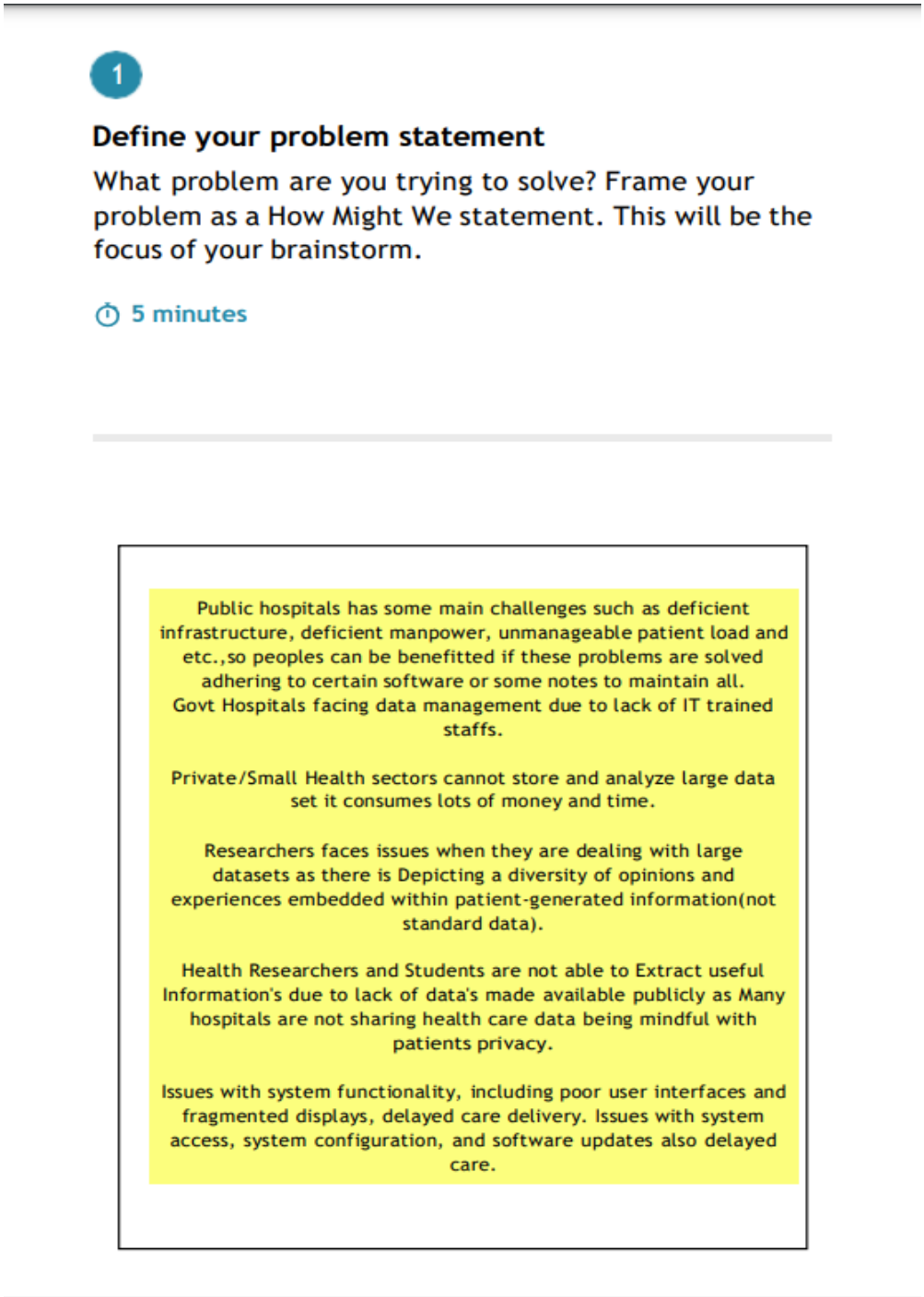
The task is to accurately predict the Length of Stay for each patient on case-by-case basisso that the Hospitals can use this information for optimal resource allocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10 days tomore than 100 days.

3.Ideation & proposed solution

3.1Empathy map Canvas



3.2 Ideation and Brainstorming



2

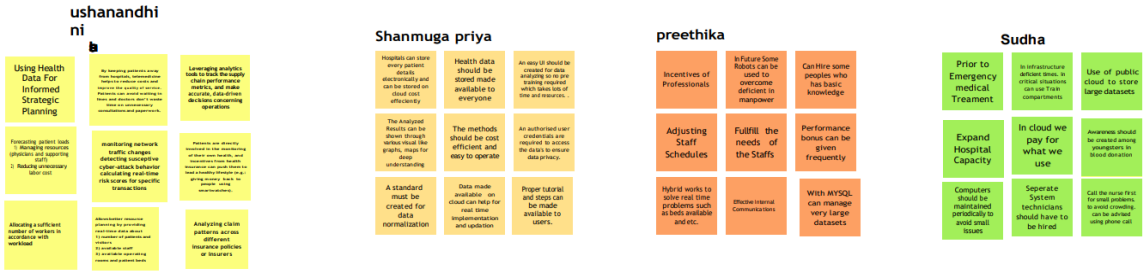
Brainstorm

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 [switch to sketch] icon to start drawing!

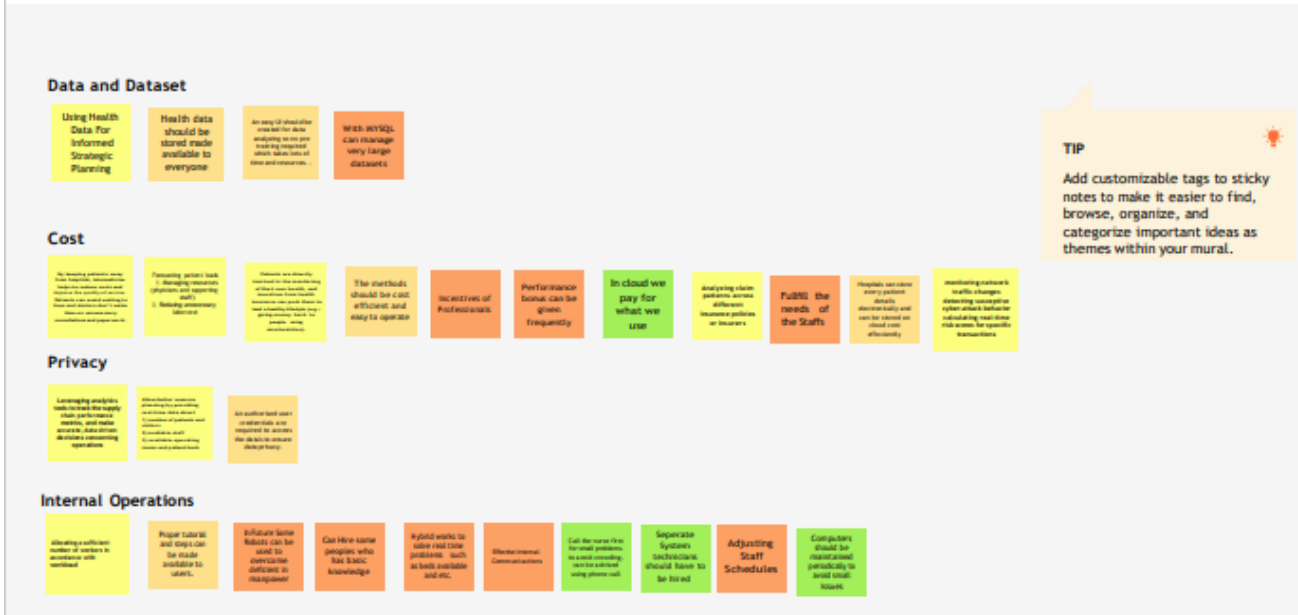


3

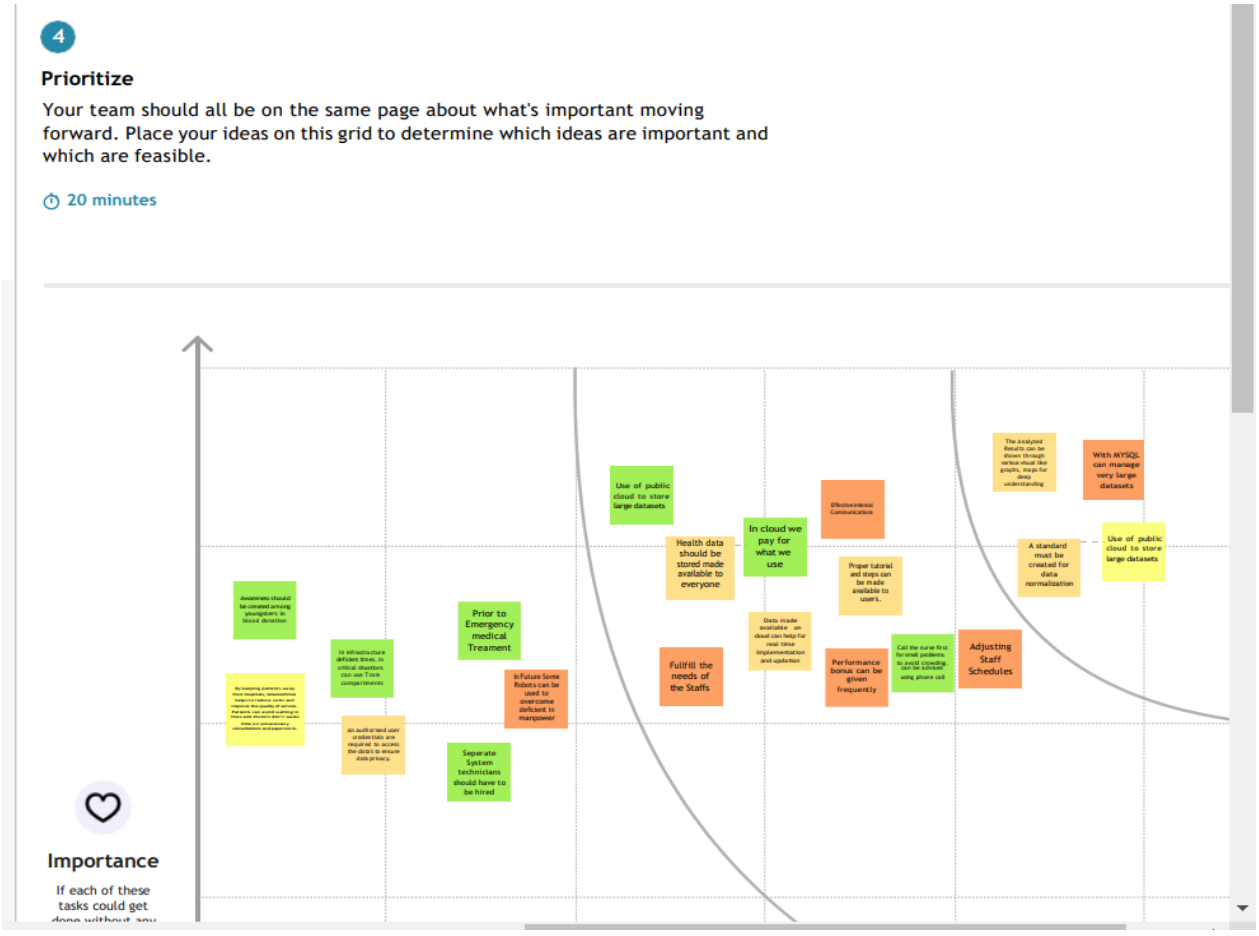
Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes



Managing Data
A standard must be created for data normalization
Cloud
Use of public cloud to store large datasets
Priority
Prior to Emergency medical Treament
Expand Hospital Capacity
Awareness


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
Use of public cloud to store large datasets
Health data should be stored made available to everyone
In cloud we pay for what we use
Proper tutorial and steps can be made available to users.
Fullfill the needs of the Staffs
Performance bonus can be given frequently
Adjusting Staff Schedules
A standard must be created for data normalization
Use of public cloud to store large datasets
With MYSQL can manage very large datasets
Prior to Emergency medical Treament
Importance

3.3Proposed solution

S.No.	Parameter	Description
1.	Problem Statement (Problem tobesolved)	• The goal is to predict the bed availability for patients at risky period and improve efficiency in the healthcare industry. • The following are the issues that are arrised because of mismanagement. ➢ Poor bed allocation planning. ➢ Difficulty in categorizing patients based on LOS ➢ Patients may get irritated and might move on to other hospitals because of the unavailability of earliest treatement in case of contagious and serious issues. • To provide proper solution for these issues ,certain procedures and techniques must be followed.

2.	Idea / Solution description	• With the help of data analytics tools and various M-L Algorithms the data patterns are monitored,analyzed to give an information about bed availability and other facilities that are currently in avail to organizations. • Creating interactive dashboard to know the bed and other facilities availabilty. • Automatic update of datas in database reflects in the dashboard and viceversa and to achieve the efficiency of management.
3.	Novelty / Uniqueness	• Responsive 24/7 Dashboard that provides the facilities that are in avail and to get the patients details which are categorised based on LOS using various factors
4.	Social Impact / CustomerSatisfaction	• This project overall manages the patients record, bed and other facilities in hospital that are in avail using a website dashboard and data analysis of bed. • Increased coordination may improve efficiency and accuracy of health care
5.	Business Model (Revenue Model)	• 24/7 support, The components provide immediate, highly focused improvements for maximum benefit.
6.	Scalability of the Solution	• For accurate prediction of available beds the best and most suits systems are in flow

3.4 Problem solution fit

Project Title: Analytics for Hospital Health-Care Data | **Project Design Phase-I - Solution Fit Template** | **Team ID:** PNT2022TMID04514

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

Who is your customer?
i.e. working parents of 0-5 y.o. kids

Hospital management and patients

6. CUSTOMER CONSTRAINTS CC

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.

Could not predict the Length Of Stay of patients properly especially during the pandemic period

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 notetaking

Hospital bed management using data mining techniques

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

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.

Proper data analysis is needed for various needs of patients

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 regulations.

Insufficient analysis in data ,human error and poor scheduling

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)

Regularly monitoring the database of patients to avoid error

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

What triggers customers to act? i.e. seeing their neighbour installingsolar panels, reading about a more efficient solution in the news

Prevailing emergency situations and Pandemic period situations

4. EMOTIONS: BEFORE / AFTER EM

BEFORE: Unstable physical and psychological state during the pandemic period

AFTER : Physical and psychological comfort and security to the patients. Improved bed allocation facilities

10. YOUR SOLUTION SL

Using **predictive analysis** powered by the AI which is used in analytics technique
Proper Data analysis and implementation in **Interactive dashboard**

8. CHANNELS of BEHAVIOUR CH

8.1 ONLINE

Usage of data exploration, Secured Login.

8.2 OFFLINE

Maintaining Data set for the patients occupancy period, predicting the LOS with doctors

Identify strong TR & EM

4.Requirements analysis

4.1 Functional requirements

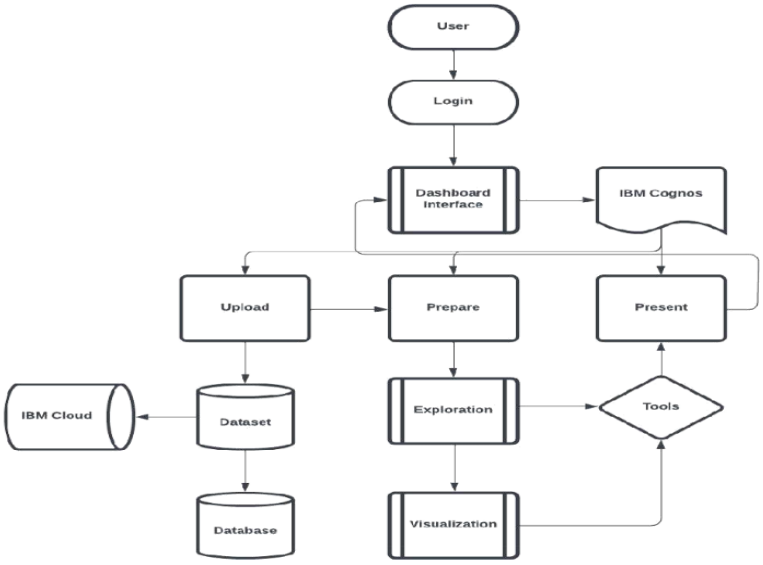
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Appointments	• Recurrent appointments and scheduling available time slots on a regular basis. • Displaying the number of appointments on a given day • After the patient is signed in asking for a mobile number and ID to avoid any issues. • Generating the appointment. • Generate the bill automatically for completed appointments • Appointment Status: 1.Confirmed 2.Pending 3.Cancelled and No Reschedule 4.Cancelled and Reschedule 5.No Show 6.Completed
FR-2	Clinical Care	• The admission for the patient must be examined thoroughly and patients who comes in a critical condition should be given treatment immediately. • Enhanced and improved reliability on reporting the data. • Access medication history from external sources and analyze the drugs(medicines) to be given. • Predicting the length of stay of inpatients.
FR-3	Patient Records	• A Proper record or documentations need to be maintained regarding the patients who are all consulted and detailed analysis of their health details. • The records must be easily accessible when required
FR-4	Bed Requirements	• Analyzing and monitoring of beds which arerequired are the most important task.

		• Using flawless systems for accurately trackingthe availability of beds.
FR-5	Providing insights of dataset	• Raw data collection and sharing of data systems are essential factors in hospital management. • Provide dataset without human error.
FR-6	Discharge Summary	• Unrestriced patients have complete discharge or deferred discharge. • Restricted patients have Conditional discharge or absolute discharge • They have to be mentioned in the discharge summary.

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 rightamount 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

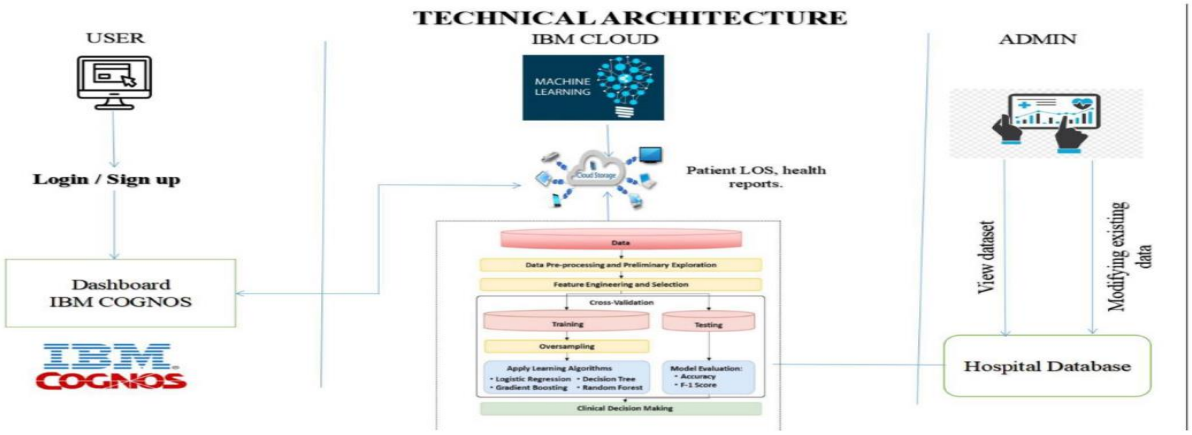


Table-1: Components & Technologies:

S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g.Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript
2.	Application Logic-1	Logic for a process in the application	Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4.	Database	Data Type, Configurations etc.	MySQL
5.	Cloud Database	Database Service on Cloud	IBM Cloud
6.	File Storage	File storage requirements	IBM Block Storage or Other StorageService or Local Filesystem
7.	External API-1	Purpose of External API used in the application	Aadhar API,etc.
8.	Machine Learning Model	Purpose of Machine Learning Model	Regression Model,etc.
9.	Infrastructure (Server / Cloud)	Application Deployment on Local System / CloudLocal Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry,etc.

Table-2: Application Characteristics:

S. No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Python
2.	Security Implementations	List all the security / access controls implemented,use of firewalls etc.	Encryption,Firewall,Antivirus
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Supports higher workloads
4.	Availability	Justify the availability of application (e.g. use ofload balancers, distributed servers etc.)	High availability enables your IT infrastructure to continue functioning even when some of its components fail
5.	Performance	Design consideration for the performance of theapplication (number of requests per sec, use of Cache, use of CDN's) etc.	A field of practice that uses various tools, processes, and ideas in a scientific manner to improve the desiredoutcomes of individuals and organizations.

5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the dashboard by entering my email, and password, and confirming my password.	I can access my accountin the dashboard	High	Sprint-1
		USN-2	As a user, I will receive a confirmation emailonce I have registered for the dashboard	I can receive a confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the dashboardthrough Social Media	I can register & access the dashboard with Social Media Login	Low	Sprint-2
		USN-4	As a user, I can register for the dashboardthrough Gmail	I can register and accessdashboard with Gmail	Medium	Sprint-2
	Login	USN-5	As a user, I can log into the application byentering email & password	I can login to the accountin my email login.	High	Sprint-2
	Dashboard	USN-6	As a user ,I can use my account in mydashboard for uploading dataset.	I can login to the accountfor uploading dataset.	Medium	Sprint-3
Customer (Webuser)	Website	USN-7	As a user ,I can use my dashboard in website	I can login into the dashboard by visiting website.	Medium	Sprint-3
Customer Care Executive		USN-8	As a user ,I can contact Customer careExecutive for my login.	I can contact customer executive for my login.	High	Sprint-4
Administrator		USN-9	As a user ,I can contact administrator for myqueries.	I can contact administratorfor solving my queries.	High	Sprint-4
Exploration	Dashboard	USN-10	As a user, I can prepare data by using Exploration Techniques.	I can prepare data byusing Exploration Techniques.	High	Sprint-3

Presentation	Dashboard	USN-11	As a user, I can Present data in my dashboard.	I can present data by using my account in dashboard.	High	Sprint-4
Visualization	Dashboard	USN-12	As a user, I can Prepare Data by using Visualization Techniques.	I can prepare data by using Visualization Techniques.	High	Sprint-3

6.Project planning & scheduling

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection	USN-1	The User needs a complete data about the patientsadmitted in the hospital and a dataset should be prepared.	2	Medium	Shanmuga Priya S, Ushanandhini R
Sprint-1	Data Exploration	USN-2	As a user, I need nicely visualized dashboard of number of beds occupied and number of free beds in hospital.	4	High	Sudharshini B,Preethika S, Shanmuga Priya S
Sprint-2	Track of patient visit of Hospital	USN-3	Tracking a patient Health care over years ofvisit and Screening of data they have in hospital.	2	Medium	Ushanandhini R,Sudharshini B

Sprint -2	Dashboard	USN - 4	As a user, I want the interactive dashboard to analyse the data. Have the data in terms of Graph.	4	High	Preethika S, Sudharshini B, Ushanandhini R
Sprint-3	Detailed EHR's of patient	USN-5	Provided greater details in the EHR's of individual patient with clear idea of what to do.	2	Medium	Shanmuga Priya S,Preethika S
Sprint- 3	Story Creation	USN-6	As a user, I need the story animation of thedata set with insights	4	High	Sudharshini B, Ushandhini R
Sprint-4	Predict LOS	USN-7	As a user, I want the flawless system to predictthe length of stay of the patients	4	High	Shanmuga Priya S, Ushandhini R,Preethika S
Sprint-4	Using ML algorithm for Prediction	USN-8	As a user, I need prior knowledge of LOS can aid in logistics such as room and bed allocationplanning.	4	High	Preethika S, Sudharshini B, Shanmuga Priya S,

Tracker, Velocity & Burndown Chart:

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

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iterationunit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

Burn Down Chart:

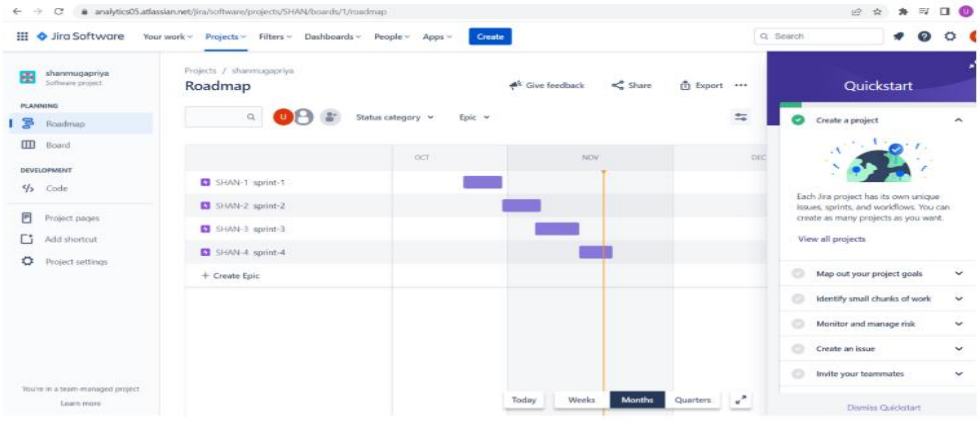
A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum.However, burn down charts can be applied to any project containing measurable progress over time.

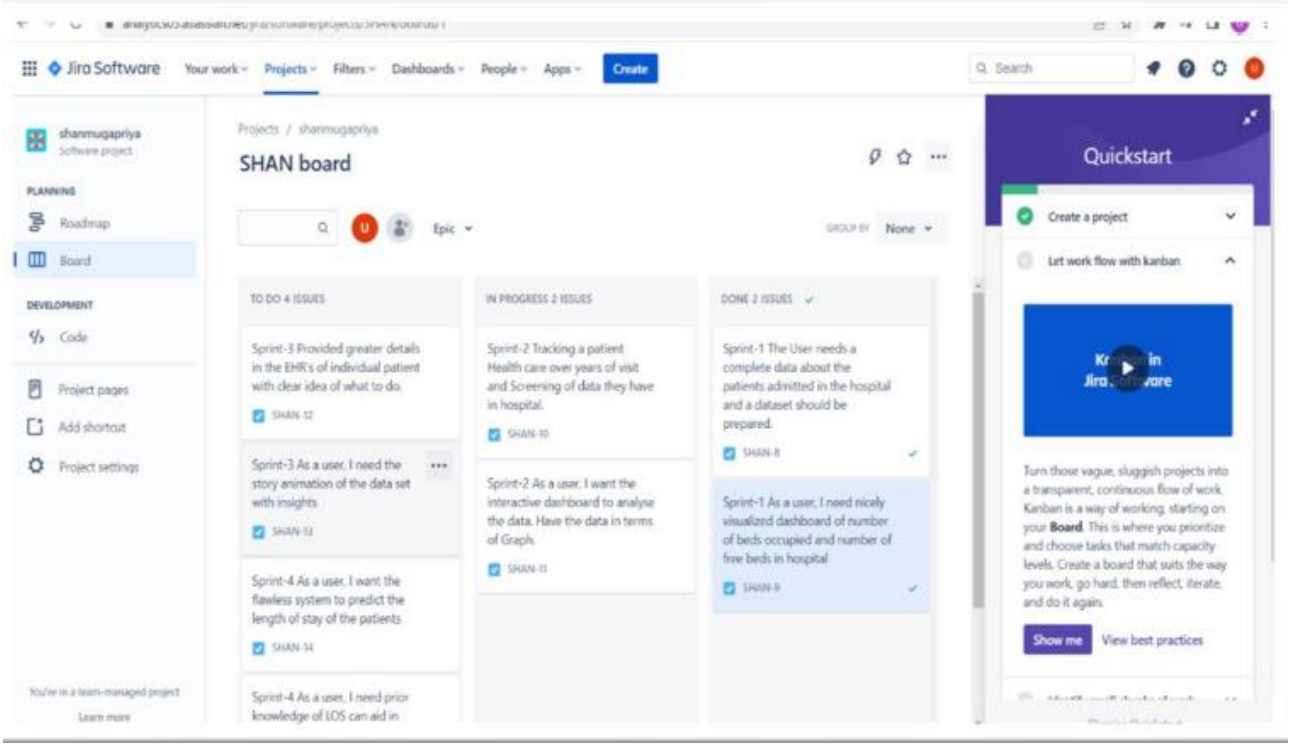
SPRINT	DATE	ESTIMATED EFFORT	ACTUAL EFFORT
SPRINT-1	24-OCT-22	20	20
	25-0CT-22	19	20
	26-OCT-22	18	19
	27-OCT-22	17	19
	28-OCT-22	17	18
	29-OCT-22	16	17
	30-OCT-22	15	15
SPRINT-2	31-OCT-22	14	13
	01-NOV-22	13	12
	02-NOV-22	12	11
	03-NOV-22	11	11
	04-NOV-22	11	11
	05-NOV-22	10	9
	06-NOV-22	9	8
	07-NOV-22	8	7

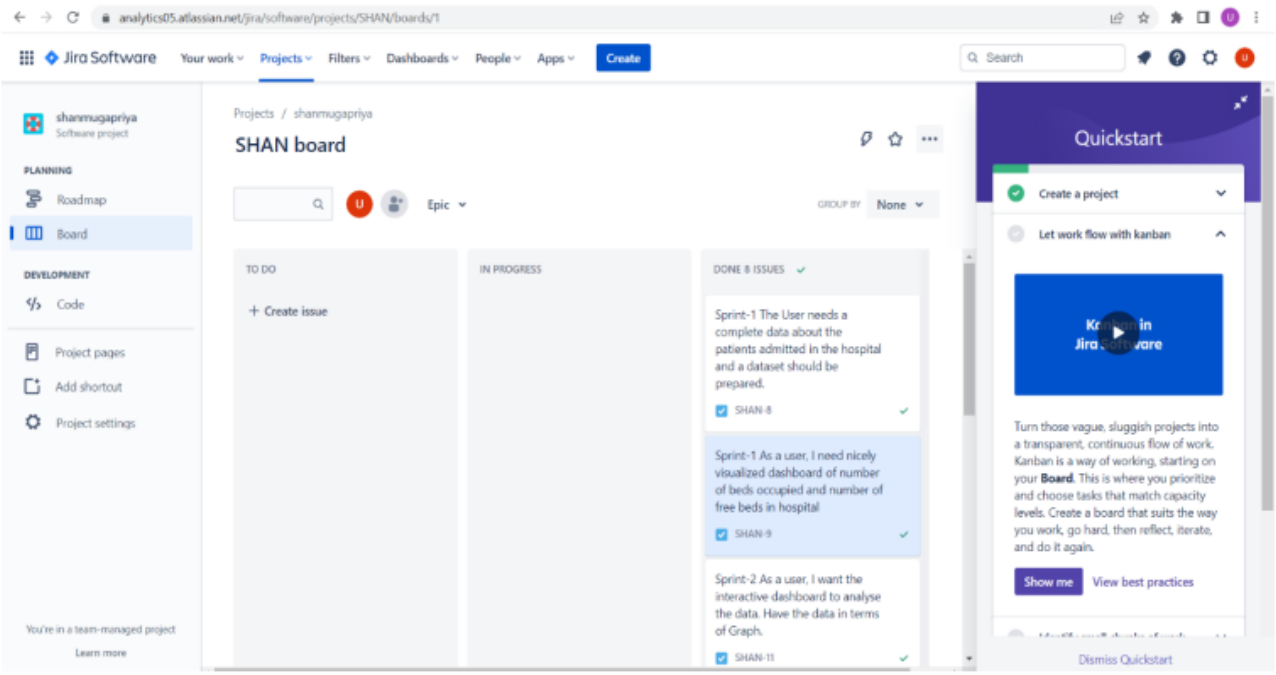
SPRINT-3	08-NOV-22	7	6
	09-NOV-22	6	6
	10-NOV-22	5	5
	11-NOV-22	5	5
	12-NOV-22	5	5
	13-NOV-22	4	4
SPRINT-4	14-NOV-22	4	4
	15-NOV-22	3	3
	16-NOV-22	2	2
	17-NOV-22	2	2
	18-NOV-22	1	1
	19-NOV-22	1	1

6.2 Reports from JIRA

Jira Sprints

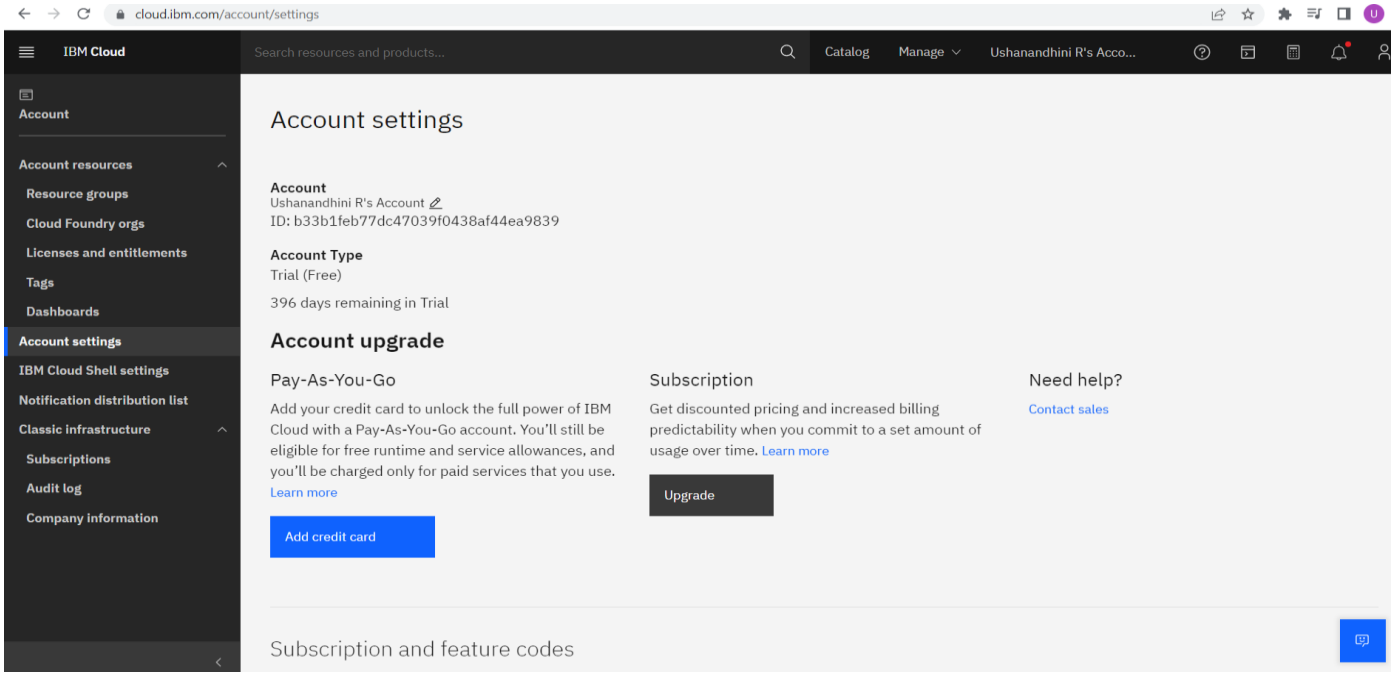


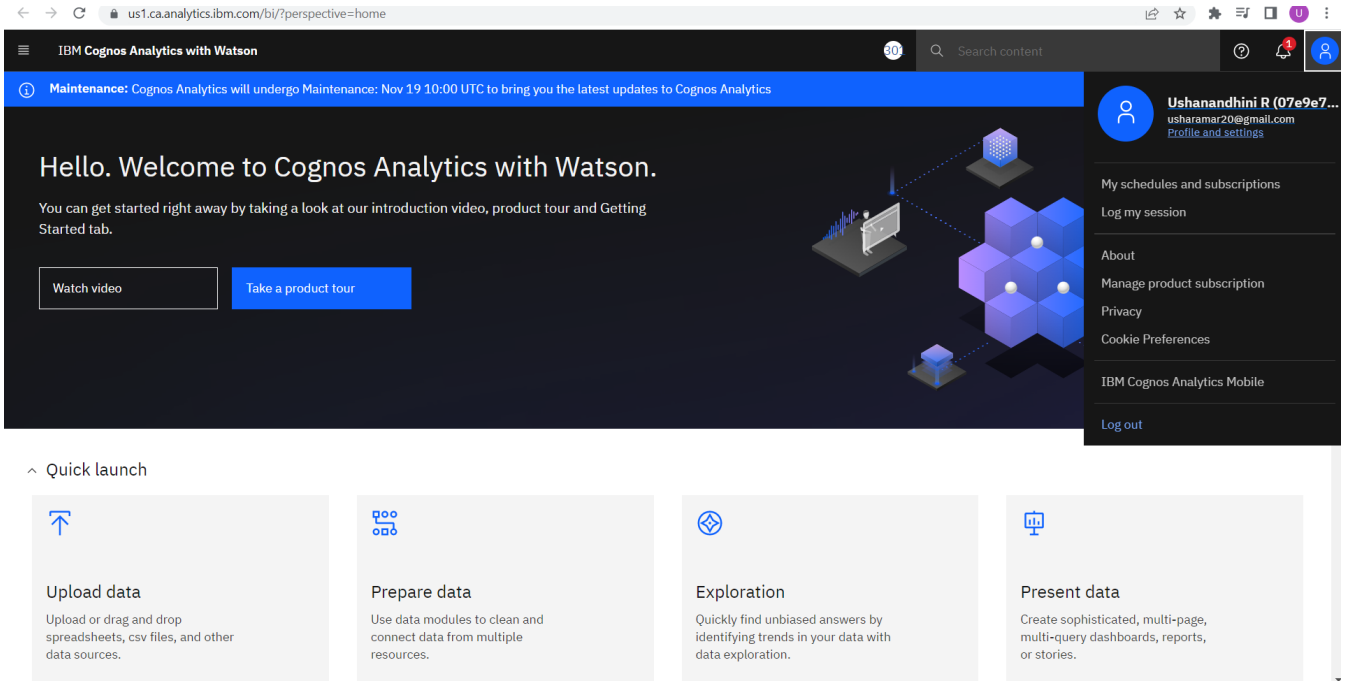

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
shanmugapriya
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / shanmugapriya
SHAN board
Epic
GROUP BY
None
TO DO 4 ISSUES
Sprint-3 Provided greater details in the EHR's of individual patient with clear idea of what to do.
SHAN-12
Sprint-3 As a user, I need the story animation of the data set with insights
SHAN-13
Sprint-4 As a user, I want the flawless system to predict the length of stay of the patients
SHAN-14
Sprint-4 As a user, I need prior knowledge of LOS can aid in
IN PROGRESS 2 ISSUES
Sprint-2 Tracking a patient Health care over years of visit and Screening of data they have in hospital.
SHAN-10
Sprint-2 As a user, I want the interactive dashboard to analyse the data. Have the data in terms of Graph.
SHAN-11
DONE 2 ISSUES
Sprint-1 The User needs a complete data about the patients admitted in the hospital and a dataset should be prepared.
SHAN-8
Sprint-1 As a user, I need nicely visualized dashboard of number of beds occupied and number of free beds in hospital
SHAN-9
Quickstart
Create a project
Let work flow with kanban
Turn those vague, sluggish projects into a transparent, continuous flow of work. Kanban is a way of working, starting on your Board. This is where you prioritize and choose tasks that match capacity levels. Create a board that suits the way you work, go hard, then reflect, iterate, and do it again.
Show me
View best practices


analytics05.atlassian.net/jira/software/projects/SHAN/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
shanmugapriya
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / shanmugapriya
SHAN board
Epic
GROUP BY
None
TO DO
+ Create issue
IN PROGRESS
DONE 8 ISSUES
Sprint-1 The User needs a complete data about the patients admitted in the hospital and a dataset should be prepared.
SHAN-8
Sprint-1 As a user, I need nicely visualized dashboard of number of beds occupied and number of free beds in hospital
SHAN-9
Sprint-2 As a user, I want the interactive dashboard to analyse the data. Have the data in terms of Graph.
SHAN-11
Quickstart
Create a project
Let work flow with kanban
Turn those vague, sluggish projects into a transparent, continuous flow of work. Kanban is a way of working, starting on your Board. This is where you prioritize and choose tasks that match capacity levels. Create a board that suits the way you work, go hard, then reflect, iterate, and do it again.
Show me
View best practices
Dismiss Quickstart

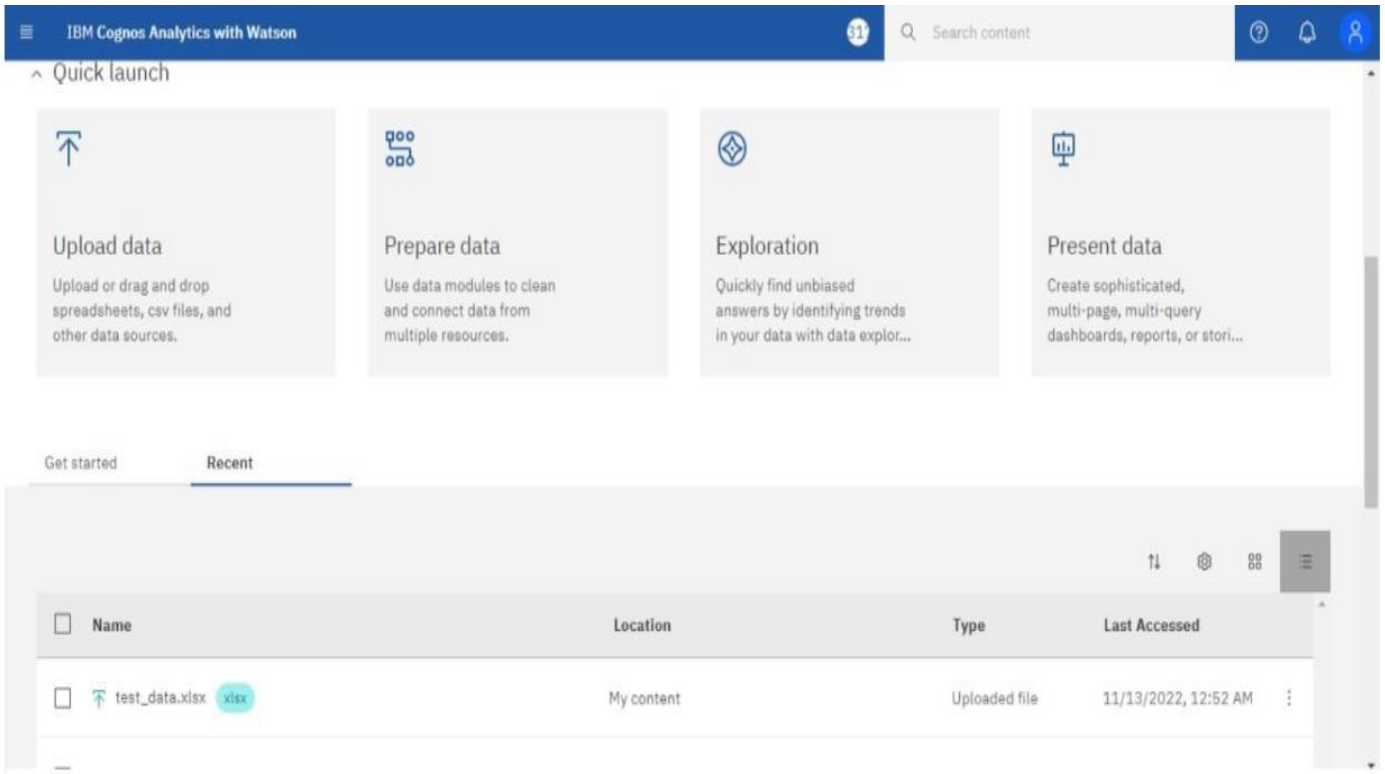
7.Coding and Solutioning

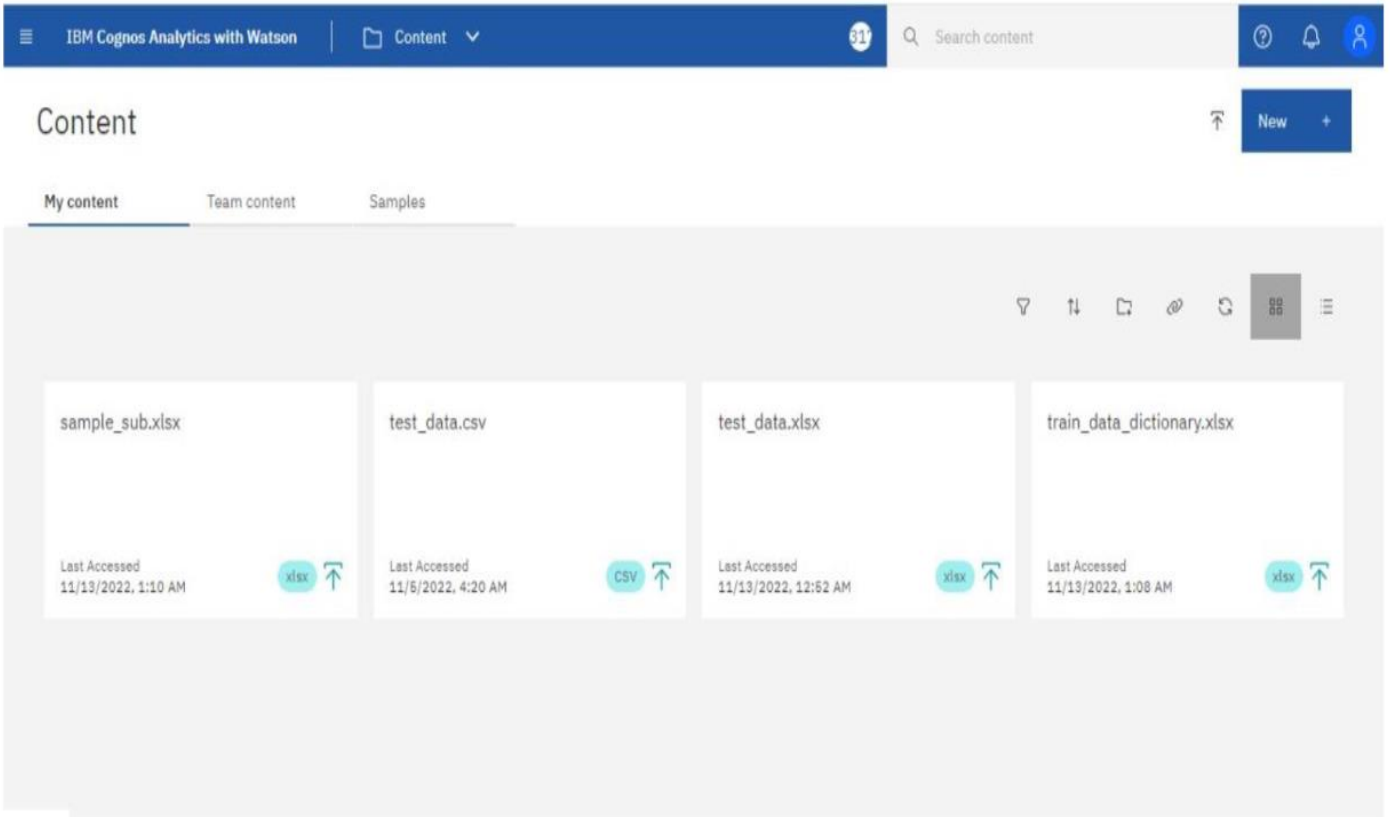
7.1. Create and configure IBM CLOUD Services





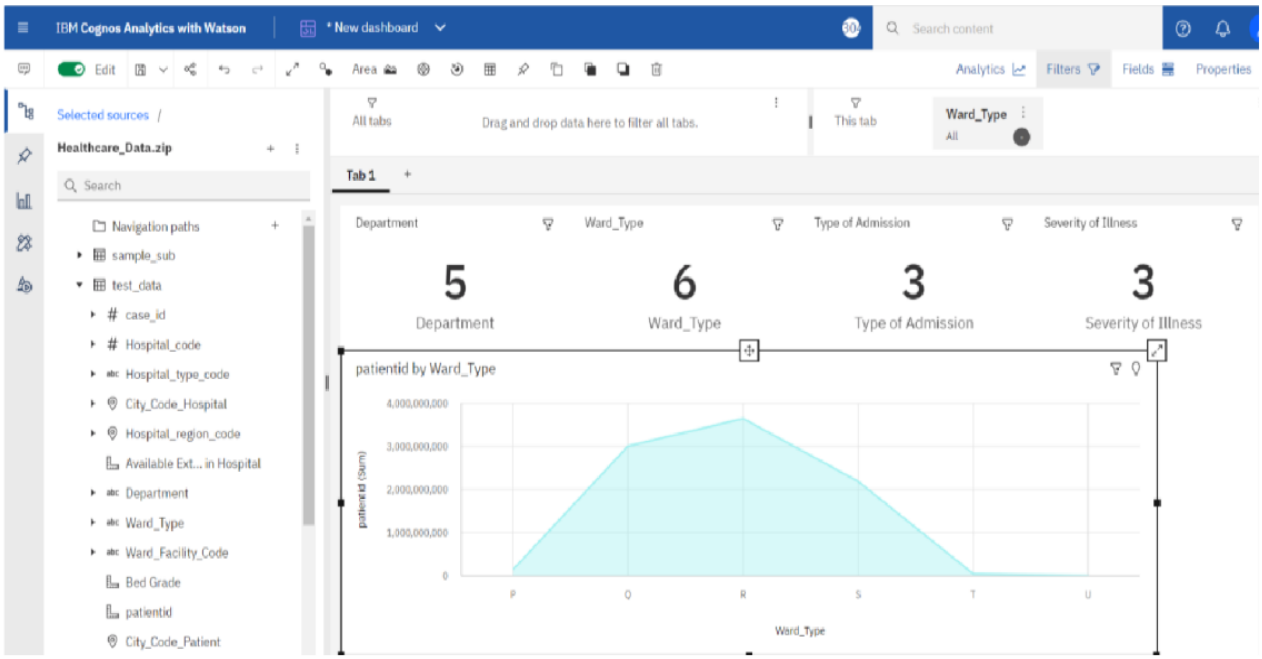
Loading Datasets



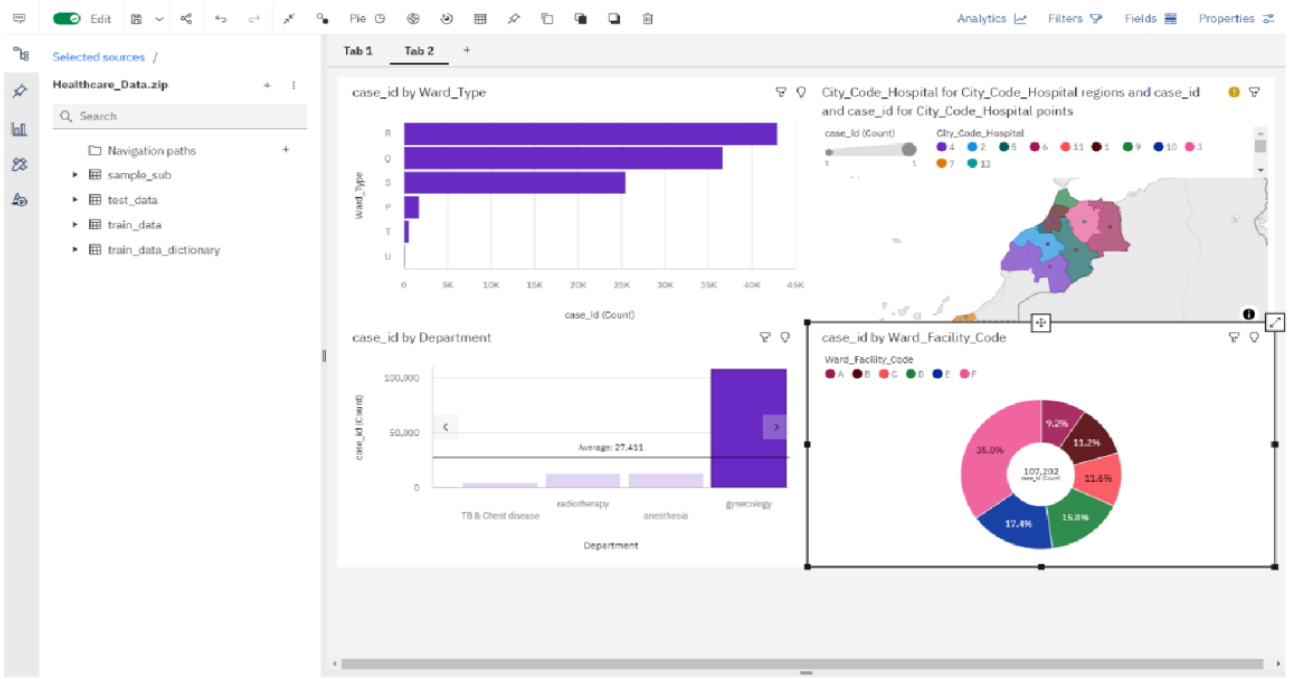


7.2.Data Visualizations

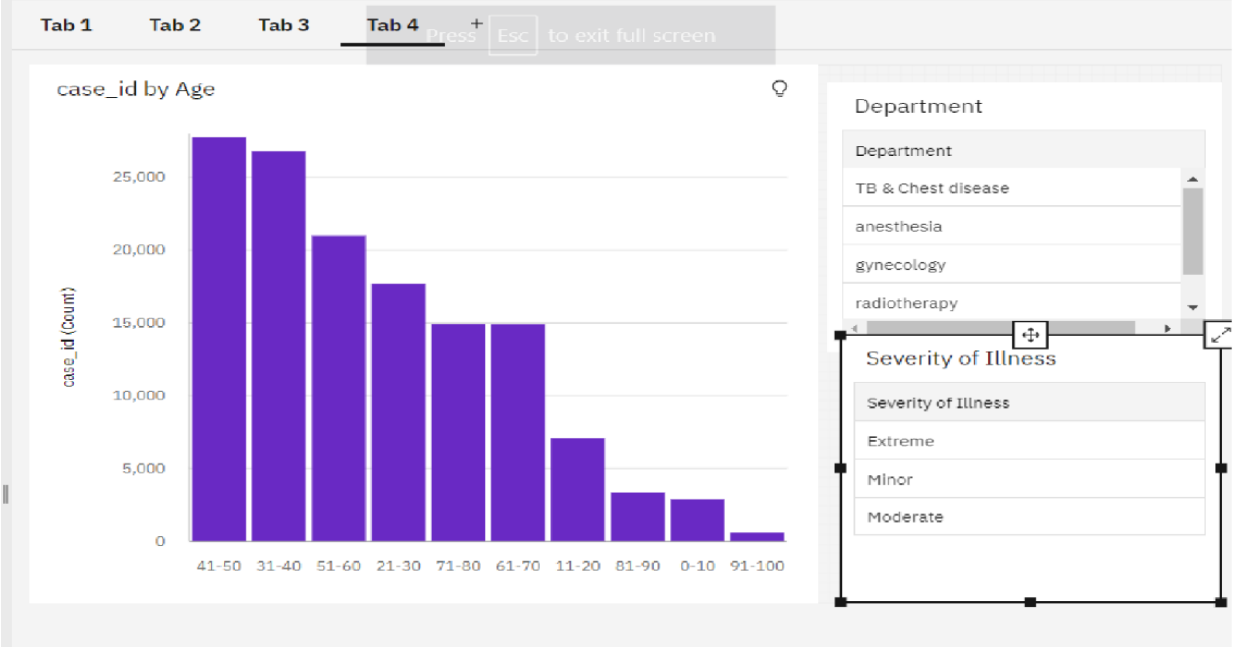
1.Number of patients by ward types



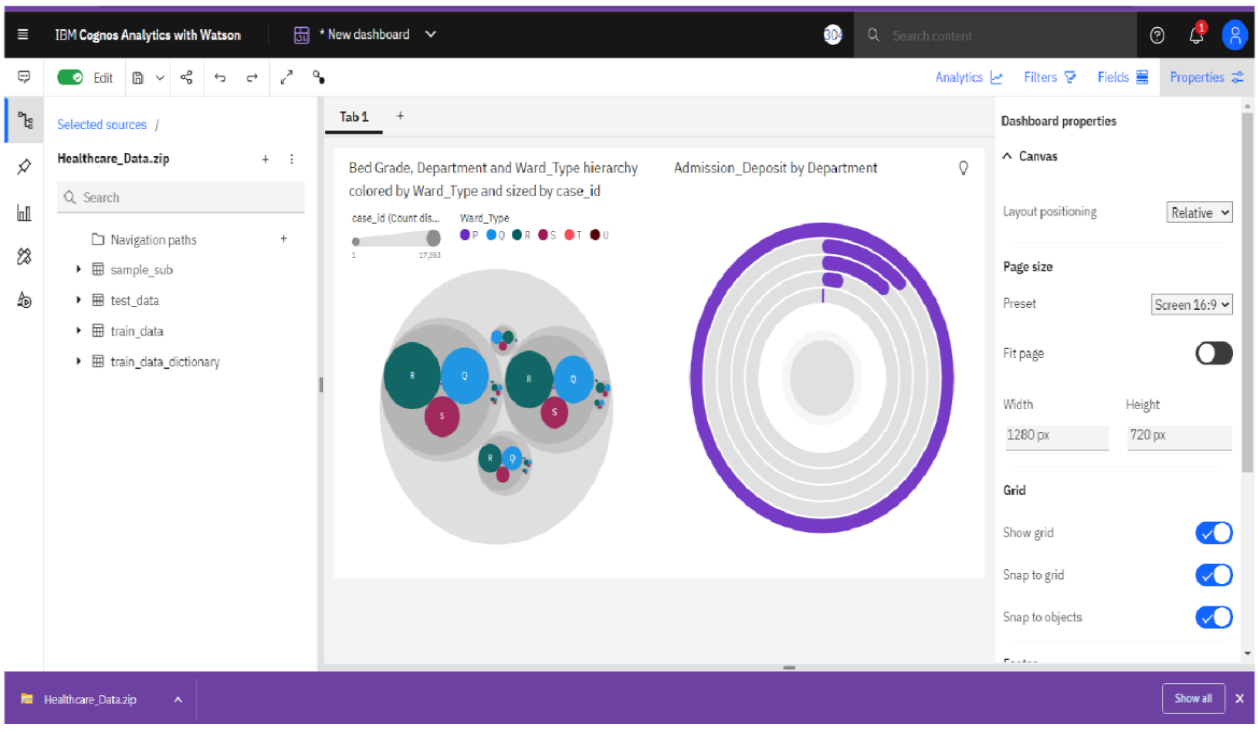
2.Dashboard to show number of patients



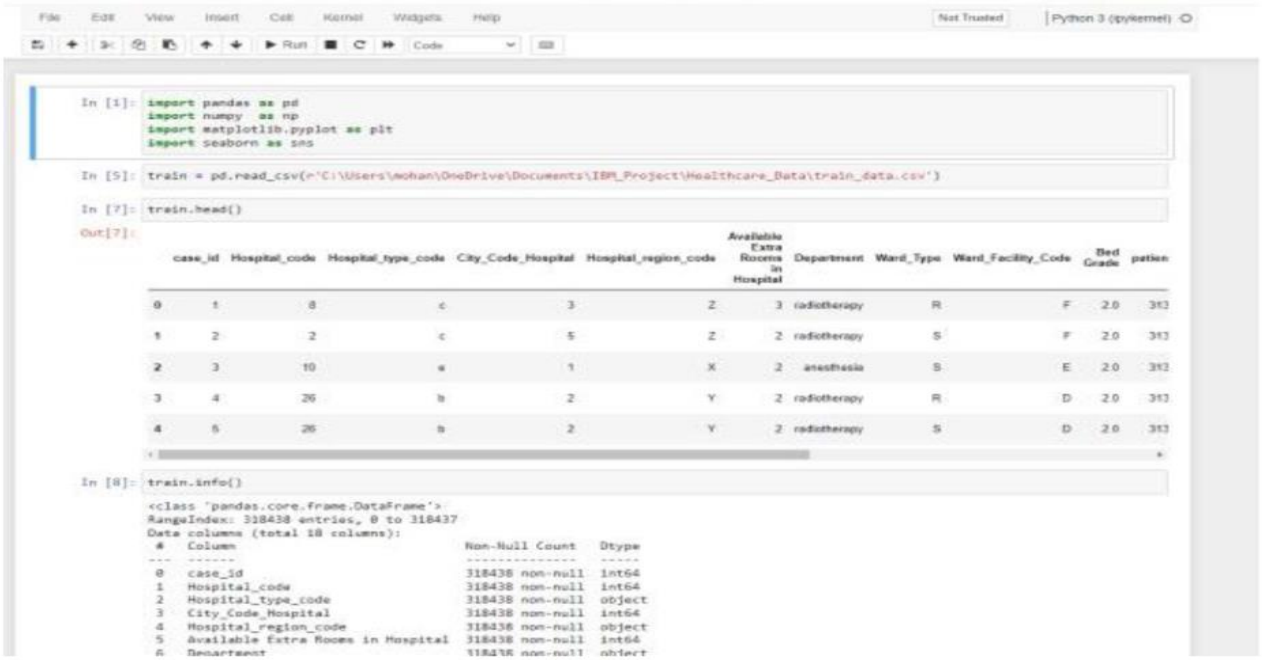
3.Age wise patients with Department and severity filters



4.Dashboard with Hierarchy Bubble and Radial Visuals



5.Null values cleaning process



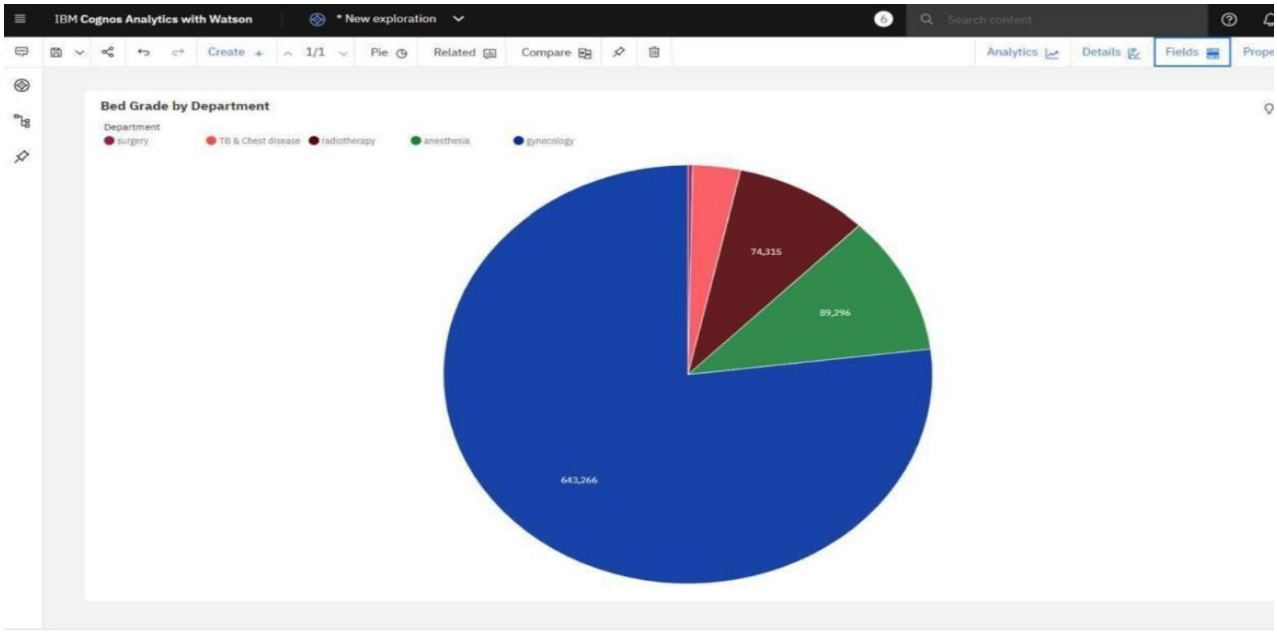
```
In [1]: import pandas as pd
        import numpy  as np
        import matplotlib.pyplot as plt
        import seaborn as sns

In [5]: train = pd.read_csv(r'C:\Users\mohan\OneDrive\Documents\IBM_Project\Healthcare_Data\train_data.csv')

In [7]: train.head()
```

Out[7]:

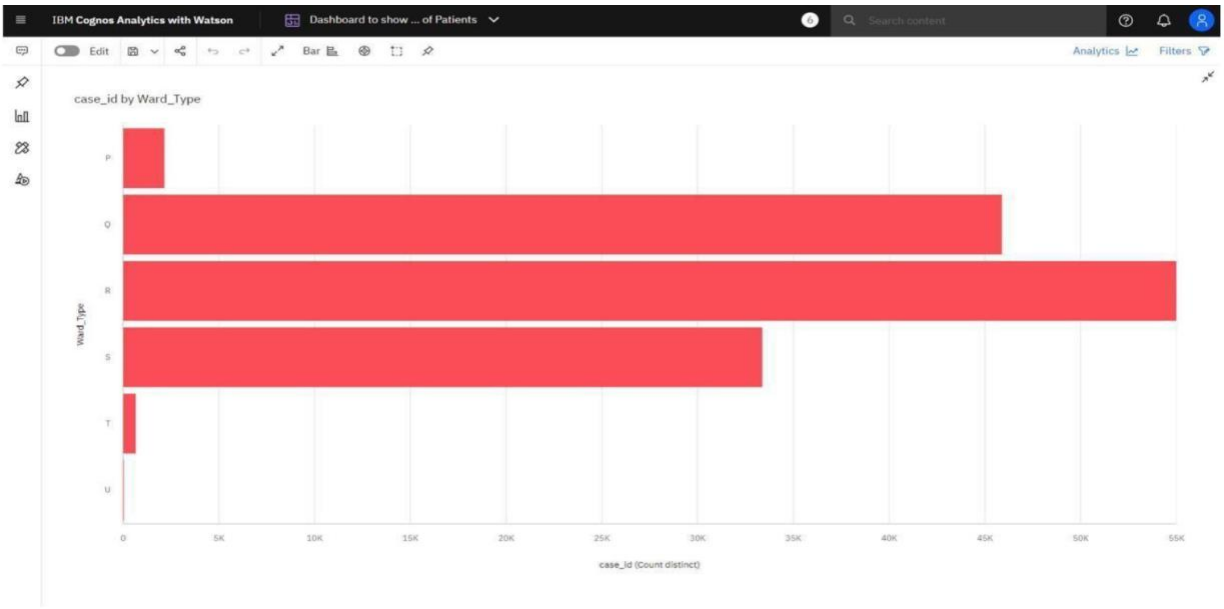
	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	patien
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0	313
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0	313
2	3	10	e	1	X	2	anesthesia	S	E	2.0	313
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0	313
4	5	26	b	2	Y	2	radiotherapy	S	D	2.0	313

```
In [8]: train.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 318438 entries, 0 to 318437
Data columns (total 18 columns):
 #   Column                             Non-Null Count   Dtype
---  ------                             --------------   -----
 0   case_id                            318438 non-null  int64
 1   Hospital_code                      318438 non-null  int64
 2   Hospital_type_code                 318438 non-null  object
 3   City_Code_Hospital                 318438 non-null  int64
 4   Hospital_region_code               318438 non-null  object
 5   Available Extra Rooms in Hospital  318438 non-null  int64
 6   Department                         318438 non-null  object
```

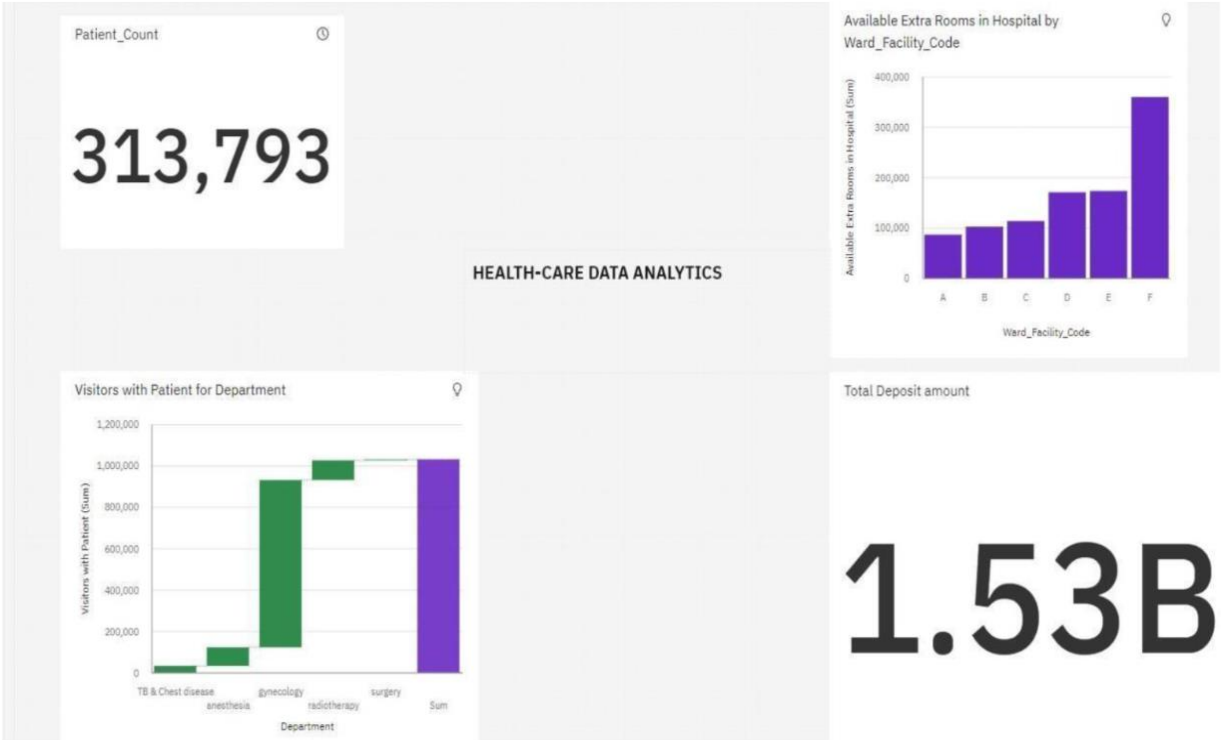
6. BedGrade by Department



7. Case Id by Ward type



7.3.Dashboard creation with datasets



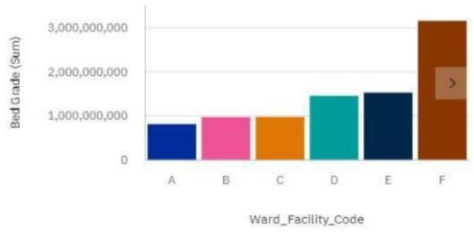
WARD TYPE AND WARD FACILITIES ANALYSIS:

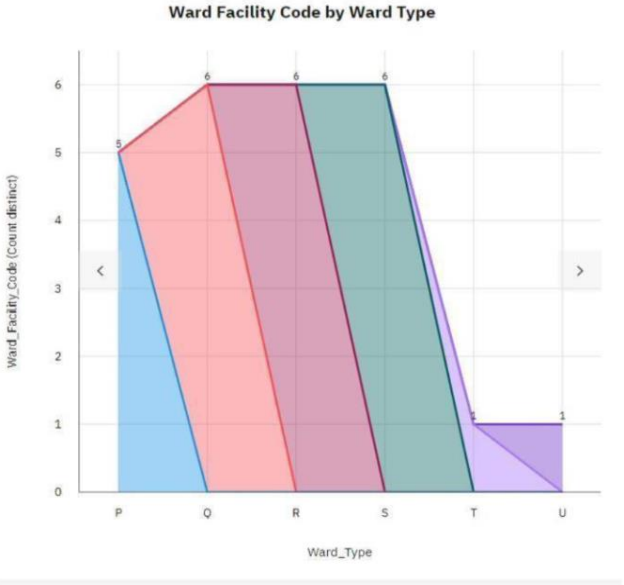
Ward Type

Ward_Type
P
Q
R
S
T
U

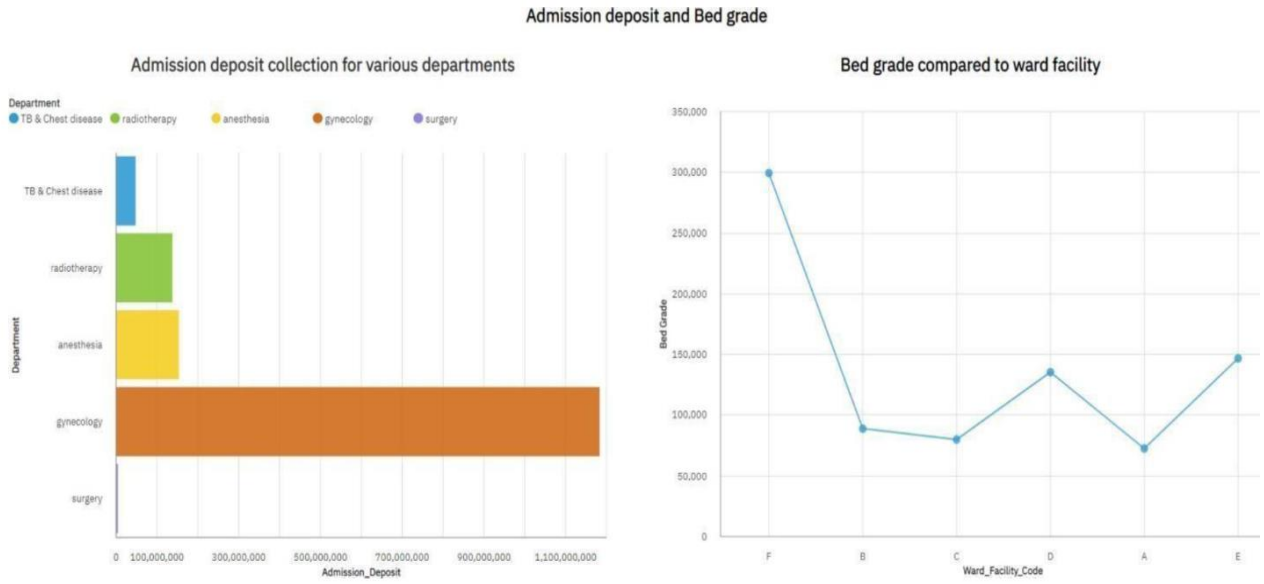
Ward Facility

Ward_Facility_Code
A
B
C
D
E
F

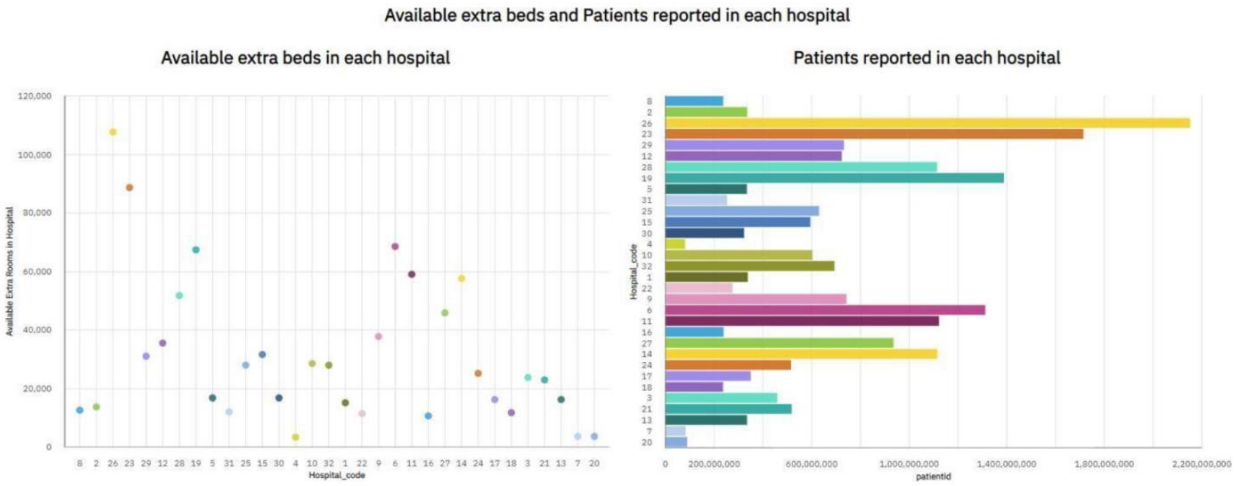




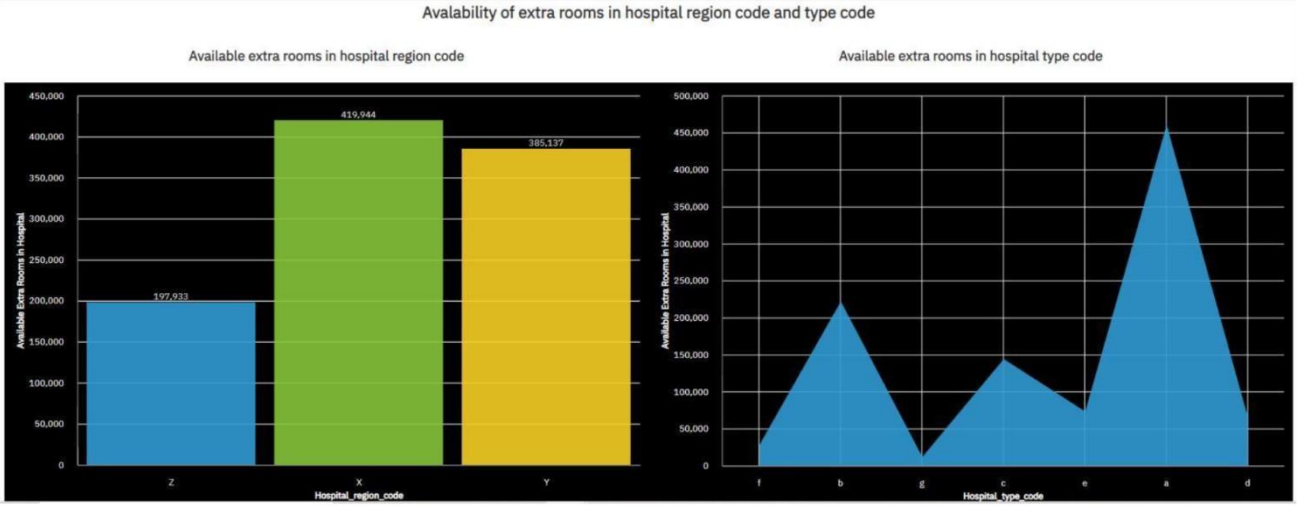
8.Results

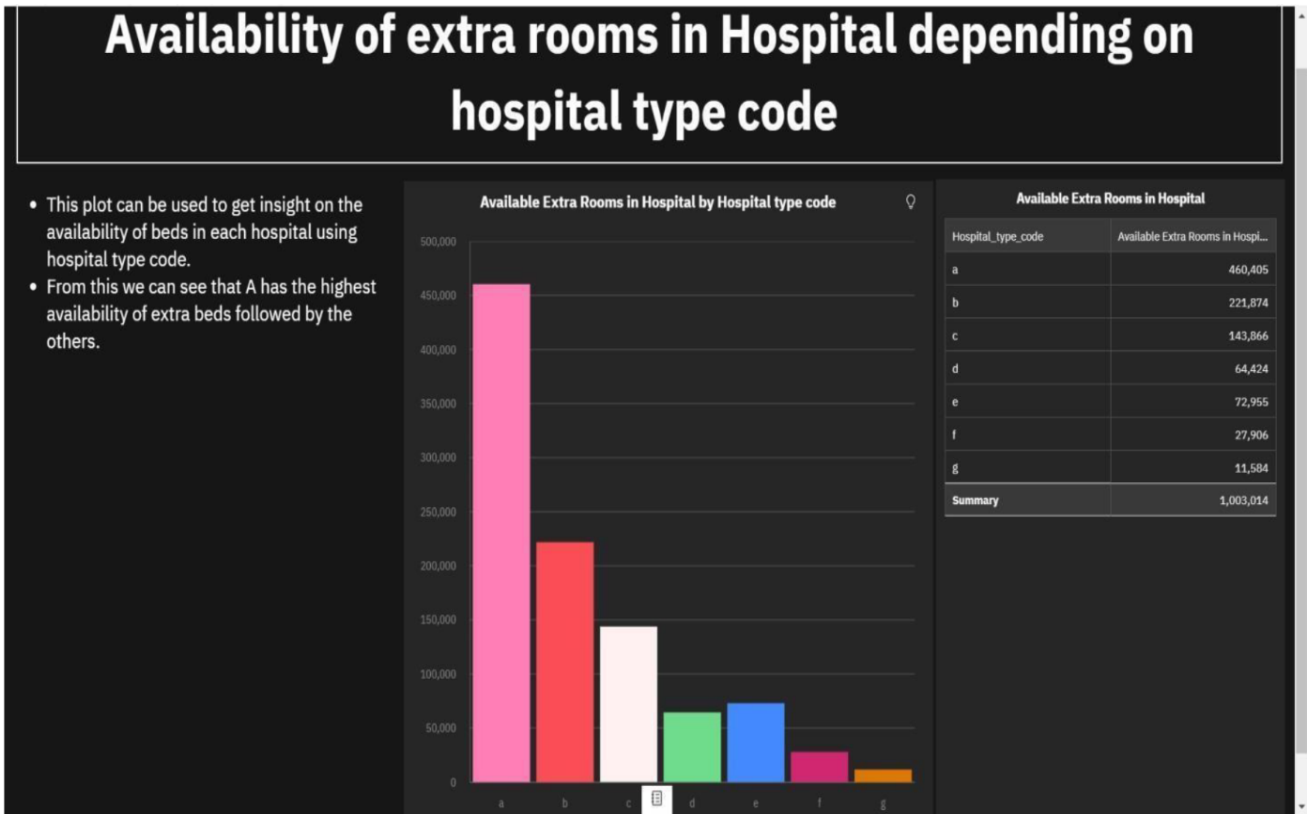


Available extra beds and Patients reported in each hospital:



Availability of extra rooms in hospital region code and type code:





9. Advantages and Disadvantages

9.1Advantages

→As the internet reaches the far ends of our world, so does digital health. With a simple internet connection, anyone can access **patient health records** online without visiting the medical center.Gone are those days of securely storing all the handwritten prescriptions and test reports. With **electronic health record** apps, you can store all the relevant health data in one place without worrying about losing one.

→ Personal digital health tools like fitness bands let you know your health-related data on a real-time basis.

→Those technological marvels constantly track your vitals and auto-dial emergency numbers in case anything wrong happens to you.

9.2 DISADVANTAGES

→Adapting to new technologies has always been a challenge for senior citizens.

→ Adapting to new technologies has always been a challenge for senior citizens. Some of them prefer the old-school treatment methods insteadof getting used to digital health facilitators.

→ If not done properly, these apps can often crash, resulting in an inconvenience.

10.CONCLUSION

This research demonstrates Analytics for hospital and health care data with data visualization and anlaytics.This information is gathered from a series of health information systems (HIS) and other technological tools utilized by health care professionals, insurance companies and government organizations. Consider the impact this has had on the COVID-19 pandemic. The data being collected is analyzed in real time to understand the effects of the virus better and predict future trends so we may slow the spread and prevent future outbreaks. Health care data management has the potential to lead to better care if used properly.

Healthcare analytics can be understood as the gathering and analysis of healthcare sector data with the purpose of deriving insights and prompting decision-making. Ranging from main areas like medical expenses, clinical data, patient behavior, or pharmaceuticals, healthcare data analytics can be employed at both the macro and micro level to sufficiently boost operations, enhance patient care, and curtail overall expenses.

Nevertheless this data, while being highly advantageous is also pretty complicated. Be it the data from electronic health records (EHR) or the data gained by assessing real-time vital signs, the data is not only derived from a number of varying sources, but it is also required to follow government regulations, making it a complex and precarious process.

11.FUTURE SCOPE

The data analytics market in the healthcare space has only increased over the last few years. Considering the rising costs of medical treatments globally, a proper body of knowledge was needed to reduce the costs at the business-level as well as the professional-level. McKinsey, in one of its reports, states that healthcare expenses constitute 17.6 percent of the GDP in the USA, which is approximately US$600 billion, more than what is the set benchmark for the ideal size of population in the country. This is a serious indicator of bigger trouble. Hence, the usage of healthcare data analytics is being promoted these days.

To some, the domain of healthcare data analytics may look new, but it has a lot of potential, especially if you wish to engage in challenging job roles and build a strong data analytics profile in the upcoming years. In this blog, we have covered some of the major topics such as what is healthcare data analytics, its applications, scope, and benefits, etc. We hope it helps you in your decision-making as a healthcare data analytics professional.