PROJECT REPORT

TEAM ID	**PNT2022TMID26934**
PROJECT NAME	**ANAYTICS FOR HOSPITALS HEALTH - CAREDATA**

TEAM MEMBERS:

Dilli Prasath S - 310819104023
Dhakshnamoorthy B - 310819104020
Deepika K - 310819104019
Afra Jabeen S - 310819104004

CHAPTER 1

INTRODUCTION

This project deals with the analytics for hospital's health care data using dataanalytics. Data analytics (DA) is the process of examining data sets in order to findtrends and draw conclusions about the information they contain. Increasingly, data analytics is done with the aid of specialized systems and softwa re.Data analyticstechnologies and techniques are widely used in commercial industries to enable organizations to make more-informed business decisions.

a. **PROJECTOVERVIEW:**

Recent Covid-

19Pandemic has raised alarms over one of the mostoverlooked areas to f ocus:Healthcare Management.

While healthcare management has various use cases for using data

science, patient length of stay is one criticalparameter to observeand predict if one wantsto improvethe efficiency of the healthcare mana gement 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 highLOS 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 dedicatedto manage the functioning of Hospi tals in a professional and optimalmanner

b. **PURPOSE:**

Data analytics in health care is vital. It helps health care organizations to evaluateand develop practitioners, detect anomalies in scans and predict

outbreaks in illness, per the Harvard Business School. Data analytics can also lower costs for health care organizations and boost businessintelligence. Hospital dataanalytics can look over patient data and any prescribed medication to alert doctors and patients of incorrect dosages or wrong prescriptions, which lessens human error and the cost to your hospital. This in turn helps in gaining better insights and also enables healthcare practitioners to make well-informed decisions.

CHAPTER 2

LITERATURE SURVEY

The main aim of this paper is to provide a deep analysison the researc h field of healthcare data analytics.This paper is analyzing the previous studies and works in this research area,as well as highlighting some of the guidelines and gaps.This study has used seven popular databases and selected most relevant papers,in order to conduct this paper.The paper has listed some data analytics tools and techniques that have been used to improve healthcare performances in many areas such as:medical operations,reports,decision making,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[1]

This part deals with the advanced analytical methods focused on healthcare. This includes the clinical prediction models, temporal data mining methods, and visual analytics.Integrating heterogeneous data such as clinical and genomic data isessential for improving the predictive power of the data that will also be discussed.Information retrieval techniques that can enhance the quality of biomedical search will be presented. Data privacy is an extremely important concern in

healthcare. Privacy-
preserving data publishing techniques will therefore be presented.[2
].

One of the promises of the growing critical mass of clinical data
accumulatingin electronic health record (EHR) systems is secondary use (or re-use)of the data forother purposes, such as quality improvement and clinical research.1 The growth of such data has increased dramatically in recent years due to incentives for EHR adoption in the US funded by the Health Information Technology for Economic and Clinical Health (HITECH) Act.2-3 In the meantime, there has also seen substantial growth in other kinds of health-related data, most notably through efforts to sequence genomes and other biological structures and functions.4 The analysis of this data is usually called analytics (or data analytics). This chapter will define
the terminology of this field, provide an overview of its promise, describe what workhas been accomplished, and list the challenges and opportunities going forward[3].

Clinicians, healthcare providers-
suppliers, policy makers and patientsareexperiencing exciting opportunities in light of new information deriving from the analysis

of big data sets, a capability that has emerged in the last decades. Due tothe rapid increase of publications in the healthcare industry, we have conducted a structured review regarding healthcare big data analytics. With reference to the resource-based view theory we focus on how big data resources are utilized to create organization values/capabilities, and through content analysis of the selectedpublications we discuss: the classification of big data types related to healthcare, the associateanalysis techniques, the created value for stakeholders, the platforms

and tools for handling big health data and future aspects in the field. We present a number of pragmatic examples to show how the advances in healthcare were madepossible. We believe that the findings of this review are stimulating and provide valuableinformation to practitioners, policy makers and rese archers while presenting them with certainpaths for futureresearch[4].

In this modern techno-world, the term data is unavoidable and certainly, nothing is possible without its usage. The trends about how to analyze the data are the need of the hour. Data analytics is becoming a future escalating tool of all industries including medicine, robotics, etc. This article briefly explains how data analytics is used

in healthcare systems. Health care is the process of maintaining and improving the health of an individual by prevent ing, diagnosing and treatingthe diseases, illness and other physical and mental imbalances in people. Data analytics is classified into four types and they are descriptive, diagnostic, predictive and prescriptive analysis. Health care makes use of prescriptive analysis to arrive atthe best results and make better decisions. Big data plays a major role in data analytics. It helps the data analysts to collect data from the patients and store themefficiently. After the completion of this whole article, the reader will be able to getthe collective idea about health care analytics.[5]

a. **EXISTINGPROBLEM**

i. Low accuracy in prediction
ii. No feature extractiondone
iii. High complexity

b. **REFERENCES**

[1]. Mohammad Alkhatib , Amir Talaei-
Khoei (University of Nevada,Reno)Amir Talaei-Khoei University of Ne vada, Reno | UNR · Department of Accounting and Information Syst emsPhD of Information Systems-Amir Ghapanchi

[2]. From:"Book of Data Analytics" Chandank Reddy(Wayne State University) CharuC.Aggarwal(Watson Research Center)

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

[4]. Panagiota Galetsia , Korina Katsaliakia , Sameer Kumarb,* a Sc hool of Economics, BusinessAdministration & Legal Studies, Interna tional Hellenic University, 14th km Thessaloniki-N. Moudania, Thessaloniki, 57001, Greece b Opus College of Business, University of St. Thomas Minneapolis Campus, 1000 LaSalle Avenue, SchulzeHall 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

[6]. Yang J.-
J., Li J., Mulder J., Wang Y., Chen S., Wu H., Wang Q., Pan H. Eme rginginformation technologies for enhanced healthcare.Comput.ind. 2015;69:3- 11.doi:10.1016/j.compimd.2015.01.012. [CrossRef] [Go ogle Scholar]

[7]. Cortada J.W., Gordon D., Lenihan B. The Value of Analytics in Healthcare.
IBM Institute for BusinessValue; Armonk, NY, USA: 2012. Report No .: GBE03476- USEN-

00. [GoogleScholar].

[8]. Makary M.A., Daniel M. Medical error-the third leading cause of death in the US.Br. Med. J. 2016;353:i2139. doi: 10.1136/bmj.i2139. [PubMed] [CrossRef] [GoogleScholar].

[9]. Prokosch H.-U., Ganslandt T. Perspectives for medical informatics. Methods Inf.Med. 2009;48:38–44. doi: 10.3414/ME9132. [PubMed][CrossRef] [Google Scholar].

[10]. Simpao A.F., Ahumada L.M., Gálvez J.A., Rehman M.A. A review of analytics andclinical informatics in health care. J. Med. Syst. 2014;38:45. doi: 10.1007/s10916- 014-0045-x. [PubMed] [CrossRef] [GoogleScholar].

[11]. Ghassemi M., Celi L.A., Stone D.J. State of the art review: The data revolutionincritical care. Crit. Care. 2015;19:118. doi: 10.1186/s13054-015-0801-4. [PMC freearticle] [PubMed][CrossRef] [Google Scholar].

[12]. Tomar D., Agarwal S. A survey on Data Mining approaches for Healthcare. Int. J. Bio-Sci. Bio-Technol. 2013;5:241–266. doi: 10.14257/ijbsbt.2013.5.5.25. [CrossRef] [Google Scholar]

[13]. K. Jee and G. H. Kim, "Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system," Healthc. Inform. Res., vol. 19, no. 2, pp. 79–85, Jun. 2013. doi: 10.4258/hir.2013.19.2.79.

[14]. J. King, V. Patel, and M. F. Furukawa, "Physician adoption of electronic
healthrecord technologyto meet meaningful use objectives: 2009–2012," The Office of theNational Coordinator for Health Information Technology, Tech. Rep.,Dec. 2012.

[15]. V. Mayer-
Schönberger and K. Cukier,Big Data: A Revolution That Will Transfo rmHow We Live, Work, and Think. Eamon Dolan, 2014.

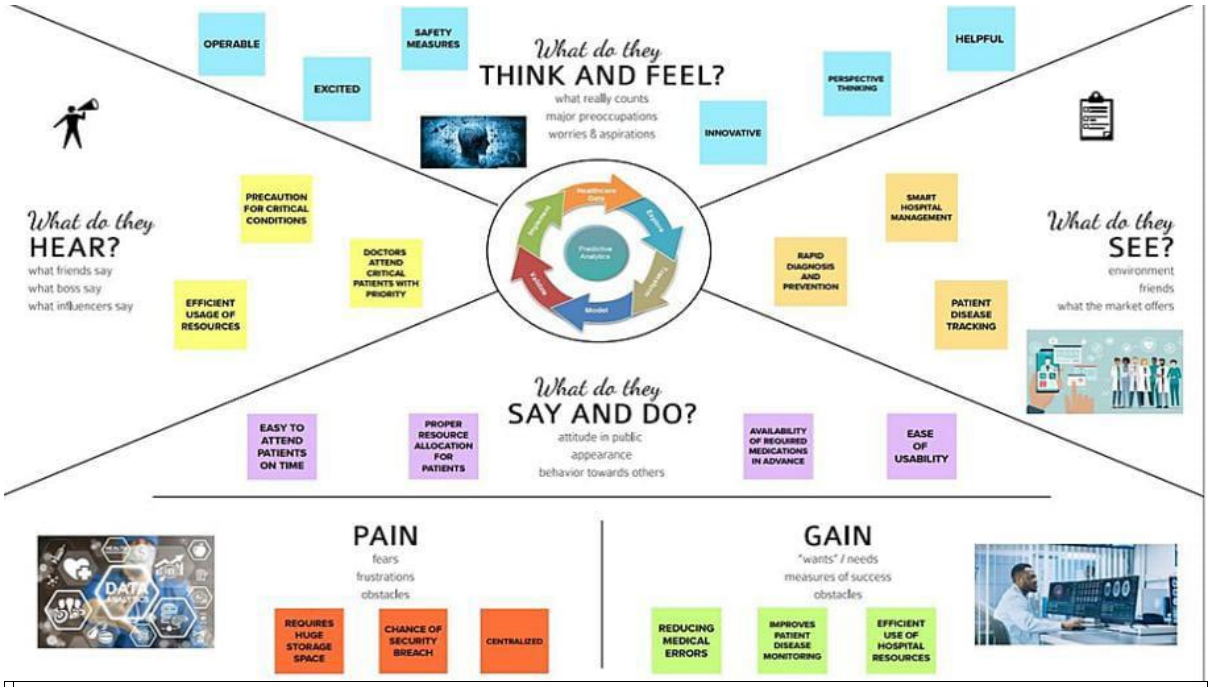
[16]. J. Rapoport,D. Teres, Y. Zhao, S. Lemeshow Length of stay d ata as a guide tohospital economicperformance for icu patients Med Care, 41 (3) (2003),pp. 386-
397

c. PROBLEMSTATEMENT AND DEFINITION

i. The aim is to accurately predictthe Length of Stay fo r each patienton case bycase basis so that the Hospitals can use this information for optimal resource allocation and better functioning.
ii. The length of stay is divided into 11 different classes ranging from 0-10 days to more than 100 days.

CHAPTER 3
IDEATION & PROPOSED SOLUTION

a. EMPATHY MAP CAMPUS



b: IDEATION & BRAINSTORMING

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

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

- **Team gathering** — Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.
- **Set the goal** — Think about the problem you'll be focusing on solving in the brainstorming session.
- **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 analyse data sets to allocate beds for patients by efficiently utilizing the resources

PROBLEM
How might we get the datas about the availability of the patients requirement

PROBLEM
How might we analyse the need for the patients

Key rules of brainstorming

To run an smooth and productive session

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

2 Brainstorm

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

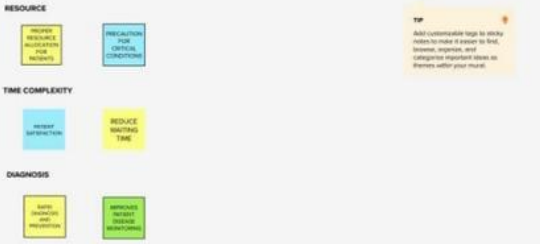
10 minutes

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. 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 and break it up into smaller sub-groups.

20 minutes

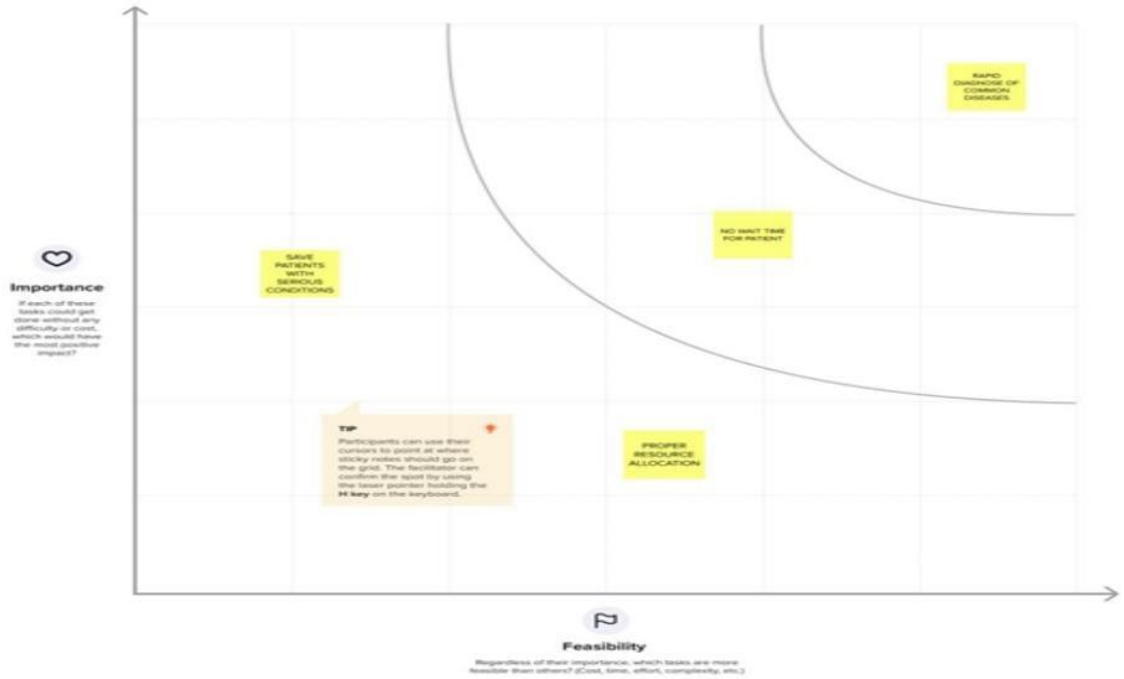


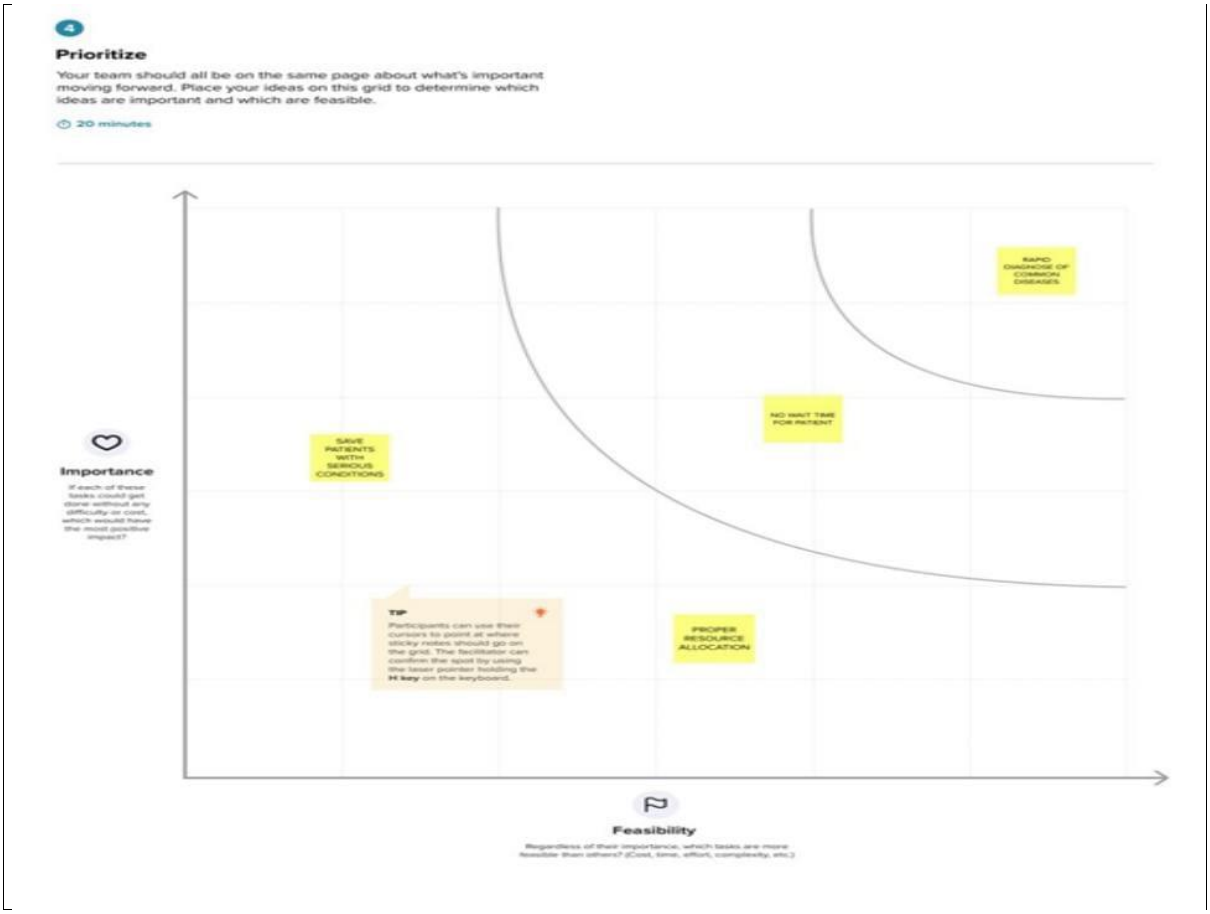
4

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





c:PROPOSED SOLUTION

Predict the lengthof stay of patients.

The length of the stay can be predicted using either Random forest or Decision Tree for more accuracy. Certain parameters like age, stage of the diseases, disease diagnosis, severity of illness, type of admission, facilities allocated, etc., are used for prediction. IBM Cognos will be used for data analytics.

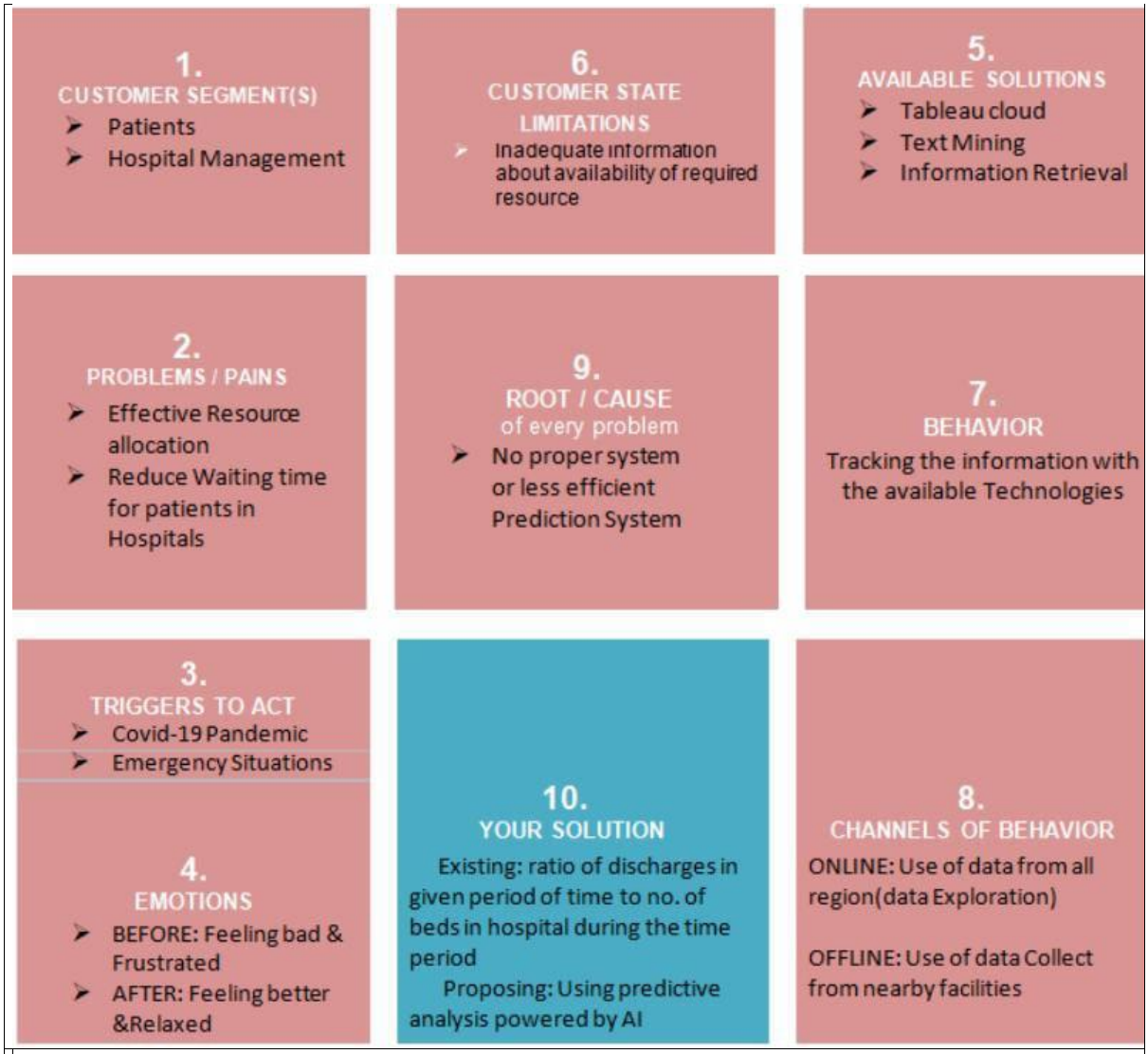
The model will be trained using collab.it predicts the length of stay(LOS)of the patients with more accuracy. As a result proper resources and therapy can be provided.Patients can get proper treatmentand better medicalc are than beforewhich helps them for their faster recovery. So the prediction minimizes the overflow of patients and helps in resource management and optimize their resource utilization. Hence this leads to faster recovery and lower the expenses fortreatment. It improvesthe trust in hospital management.

It avoids the major risk of spreading infection among the hospital staff. This leads to overall safety of hospital staff and patients. Resource consumption is optimized. This model can be used by all government hospitals, private hospitals, and this model is also trained with the real world hospital survey for better prediction smallclinics.

Length of the stay will be predicted with more accuracy. This model

predicts the length of the stay for all kinds of patients and predicts with more accuracy.

d:PROBLEM SOLUTION FIT


1.
CUSTOMER SEGMENT(S)
Patients
Hospital Management
6.
CUSTOMER STATE LIMITATIONS
Inadequate information about availability of required resource
5.
AVAILABLE SOLUTIONS
Tableau cloud
Text Mining
Information Retrieval
2.
PROBLEMS / PAINS
Effective Resource allocation
Reduce Waiting time for patients in Hospitals
9.
ROOT / CAUSE
of every problem
No proper system or less efficient Prediction System
7.
BEHAVIOR
Tracking the information with the available Technologies
3.
TRIGGERS TO ACT
Covid-19 Pandemic
Emergency Situations
4.
EMOTIONS
BEFORE: Feeling bad & Frustrated
AFTER: Feeling better &Relaxed
10.
YOUR SOLUTION
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.
CHANNELS OF BEHAVIOR
ONLINE: Use of data from all region(data Exploration)
OFFLINE: Use of data Collect from nearby facilities

CHAPTER 4
REQUIREMENT ANALYSIS

a:FUNCTIONAL REQUIREMENT

FR No.	Functional Requirement(Epic)	Sub Requirement (Story/ Sub-Task)
FR-1	**User Registration**	Registration throughForm Registration throughGmail
FR-2	**User Confirmation**	Confirmation via Email Confirmation via Message
FR-3	**Interoperability**	Dashboard helps to share the patient's informationinteroperable to the hospitals in timely manner.
FR-4	**Accuracy**	Dashboard helps predictthe patient's Healthrisksaccurately based on LOS (Length of Stay).
FR-5	**Compliance**	The compliance of a dashboard is like to use veryinteractively in realtime-by the hospitals.

FR-6	**Concise**	These dashboards are clear,intuitive, and customizableand intera ctive in manner.

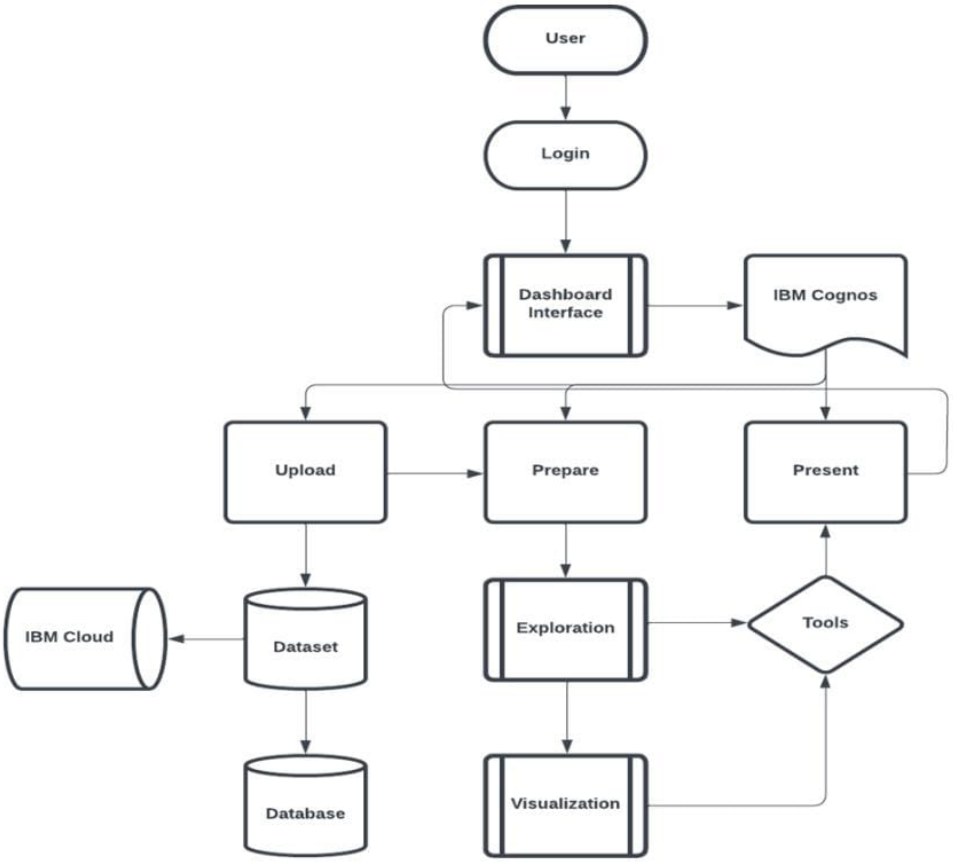
a. NON FUNCTIONAL REQUIREMENT

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	This Dashboards are designed to offer a comprehensive overview of patient's LOS,and do sothrough the use of d atavisualization tools like chartsand graphs.
NFR-2	**Security**	The Dashboard helpsto indicate the current threatlevelto the Hospitals; an indication of events and incidents that have occurred; a record of authentication errors; unauthorized access
NFR-3	**Reliability**	This dashboard will be consistent and reliable to the usersand helps th euser to use in effective, efficientand reliablemanner.
NFR-4	**Performance**	This dashboard canscan the backend users and analyzingthe frequency in which they visit the dashboard helps understand how useful and helpful the data displayed is for tasks.

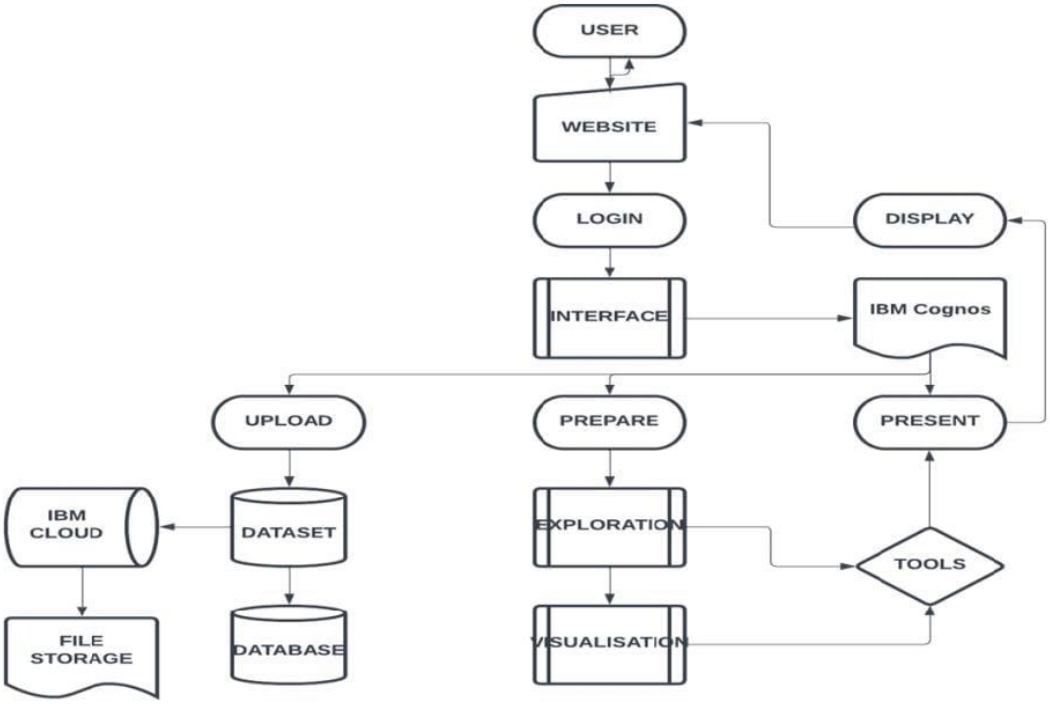
CHAPTER 5

PROJECT DESIGN

a: DATA FLOW DIAGRAMS



b:SOLUTION & TECHNICAL ARCHITECTURE



c: USER STORIES

User type	Functional requirement (epic)	User story number	User story/task	Acceptance criteria	Priority	Release
Customer	Registration	USN-1	As a user i an login to my dashboard	I can access dashboard	High	Sprint-1
	Collect data	USN-1	As a user i can provide my details	I can view my data	Medium	Sprint-1
Admin	Collect data	USN-2	As an analyst i collect the data		High	Sprint-2
	Analyze	USN-2	As an analyst i analyze the given dataset	I can analyze the dataset	High	Sprint-2
	Upload data	USN-3	As an analyst i can upload datasets	I can upload the dataset	Medium	Sprint-3
	Prediction	USN-6	As an analyst i will predict the length of stay of patient	I can predict the length of stay	High	Sprint-4
Visualization	Prepare data	USN-4	As an admin i prepare the data for visualization	I can prepare the data with visualization techniques.	High	Spint-3
	Dashboard	USN-5	As an admin i present the data that is visualized	I can present the result	High	Sprint-4

CHAPTER 6

PROJECT PLANNING &SCHEDULING

a: SPRINT PLANNING & ESTIMATION

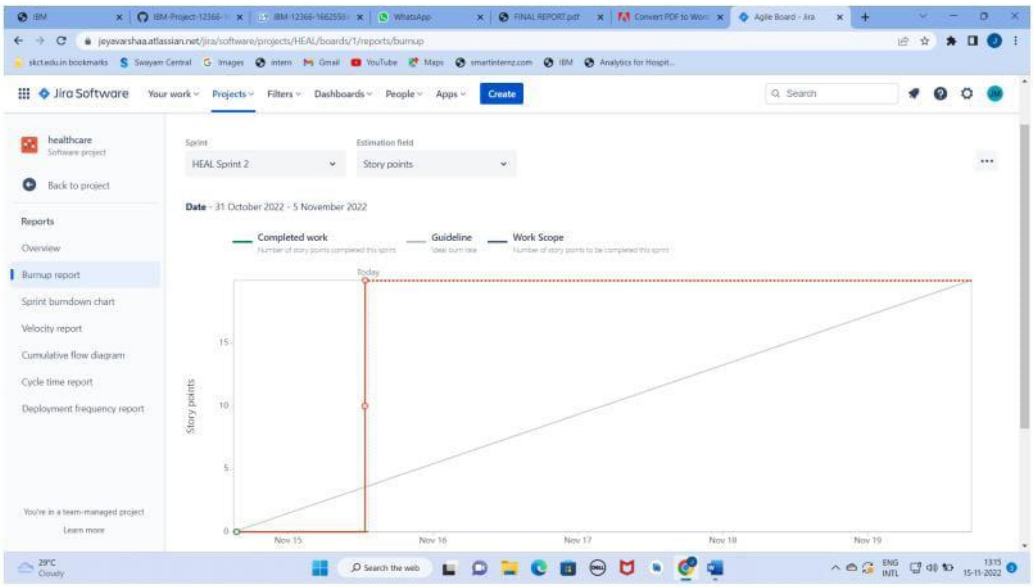
Sprint	Functional Requirement (Epic)	User Story Number	User Story/ Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a health care provider I can create account inIBM cloud and the data are collected.	20	High	Haritha Shanmathi S Jeya Varshaa M
sprint-2	Analyze	USN-2	As a healthcare provider all data are collected is cleaned and uploaded in the database or IBM cloud	20	medium	Haritha Shanmathi S Akila Umesh

b:SPRINT DELIVERY SCHEDULE

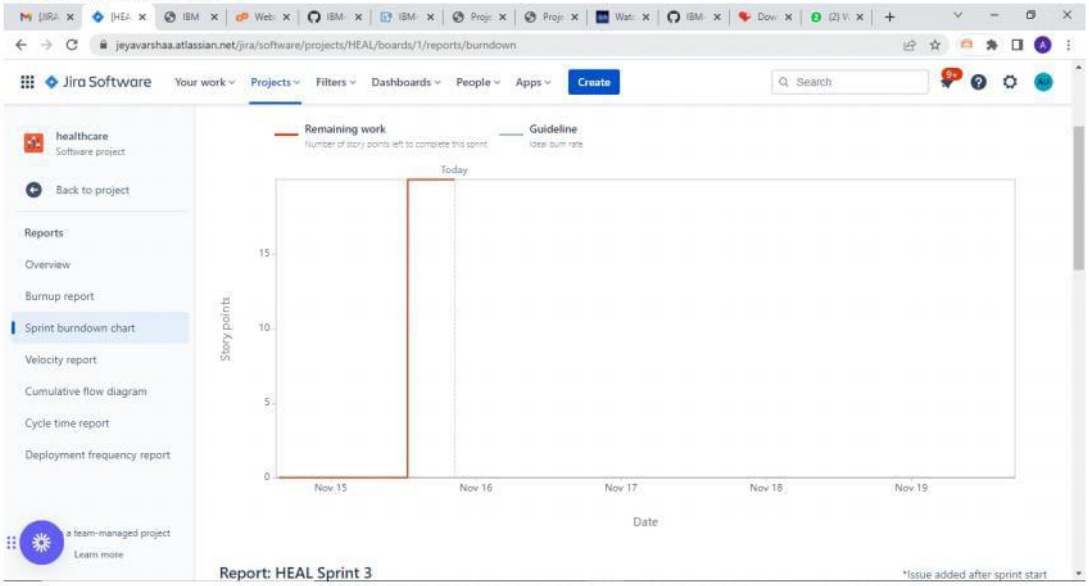
Sprint	Functional Requirement (Epic)	User Story Number	User Story/ Task	Story Points	Priority	Team Members
sprint-3	Dashbord	USN-3	As a healthcare provider I can use my account in my dashboard for uploading dataset	10	medium	Haritha Shanmathi S Gayathri M J
Sprint-3	Visualization	USN-4	As a health care provider I can prepare data forVisualization.	10	High	Haritha Shanmathi S Akila Umesh
Sprint-4	Visualization	USN-5	As a health care provider I can present data in my dashboard.	10	High	Haritha Shanmathi S Jeya Varshaa M
Sprint-4	Prediction	USN-6	As a health care provider I can predict the length of stay	10	High	Haritha Shanmathi S Gayathri M J

c:REPORTS FROM JIRA

Burnup chart



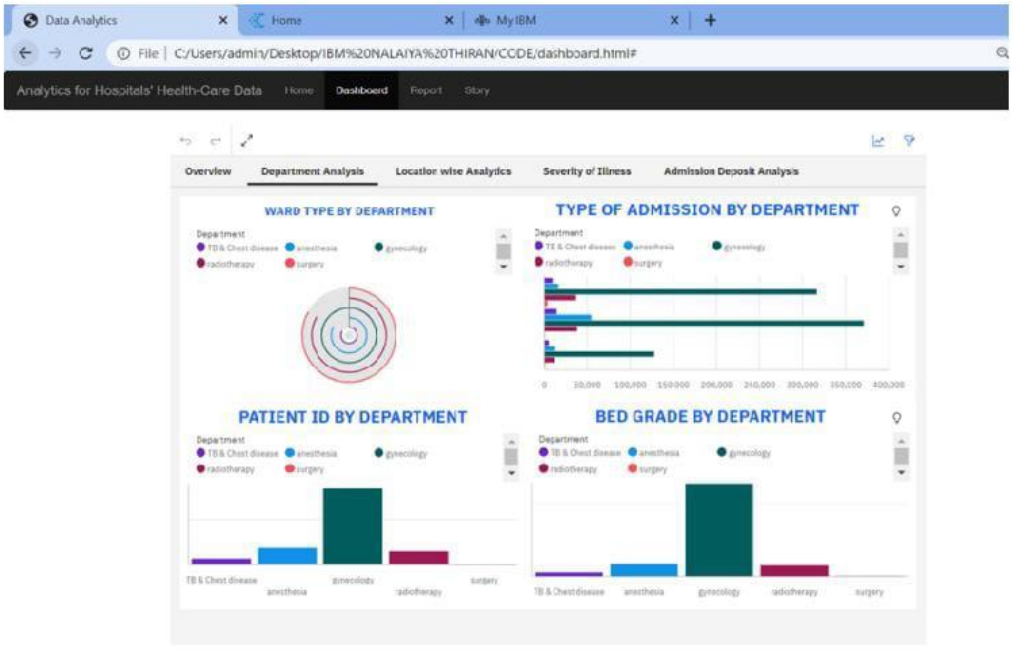
Burn down chart

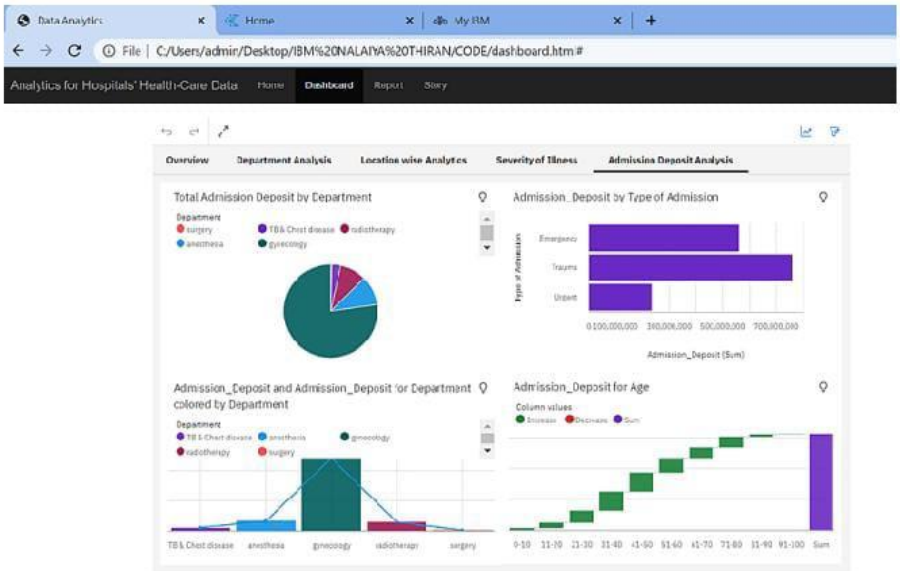


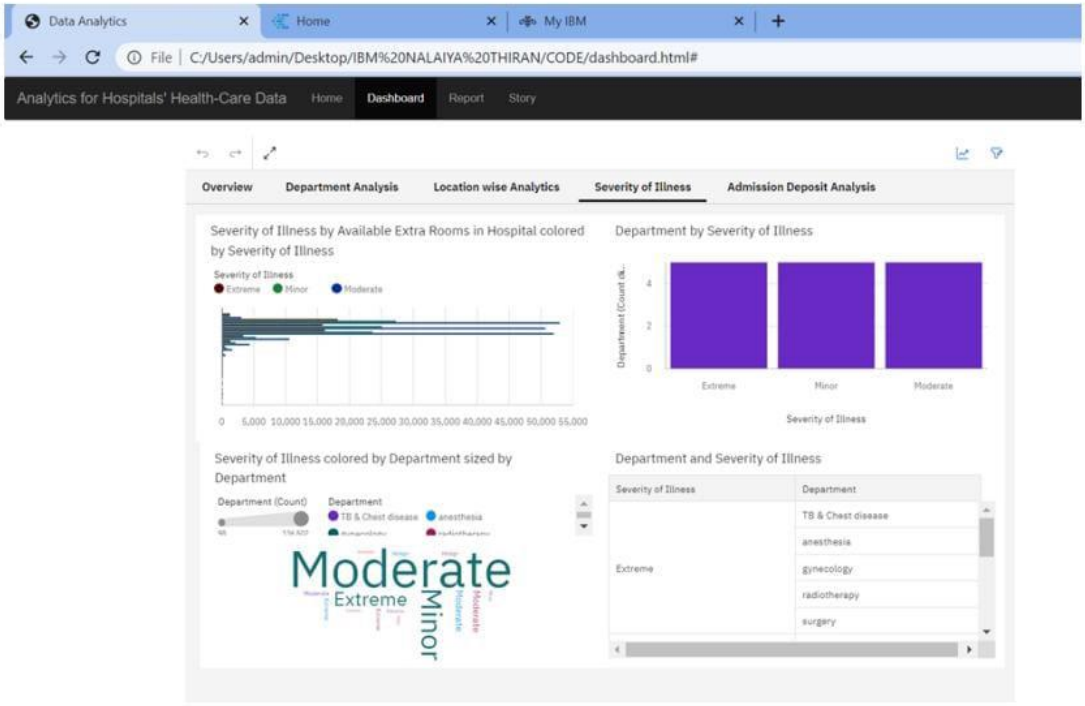
CHAPTER 7

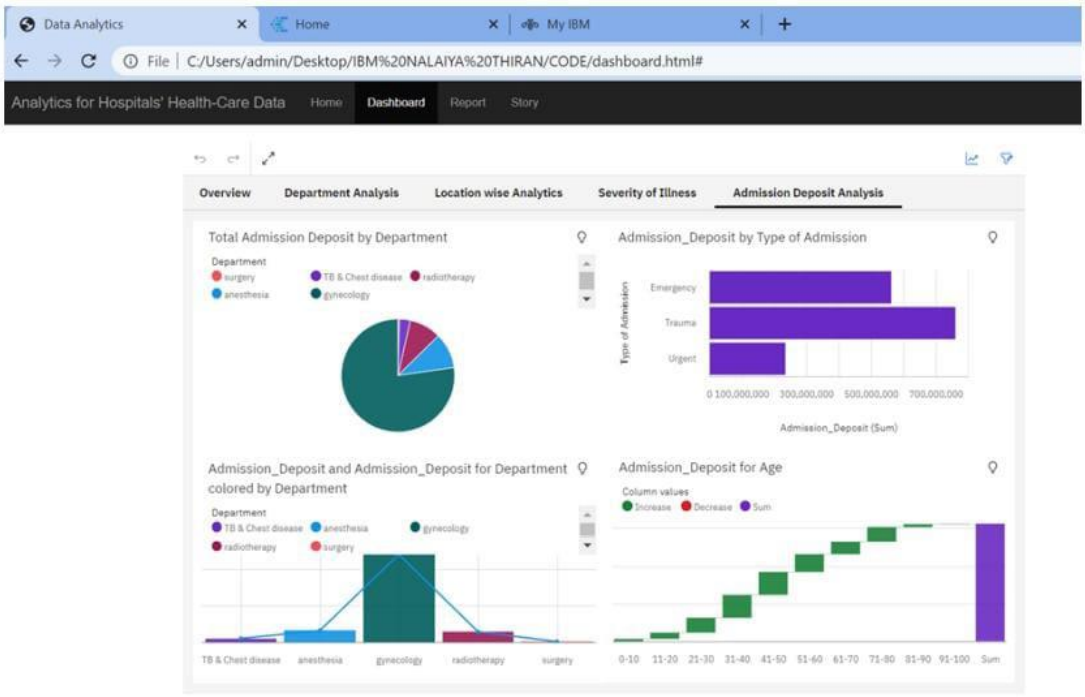
CODING & SOLUTIONING

a:FEATURE 1

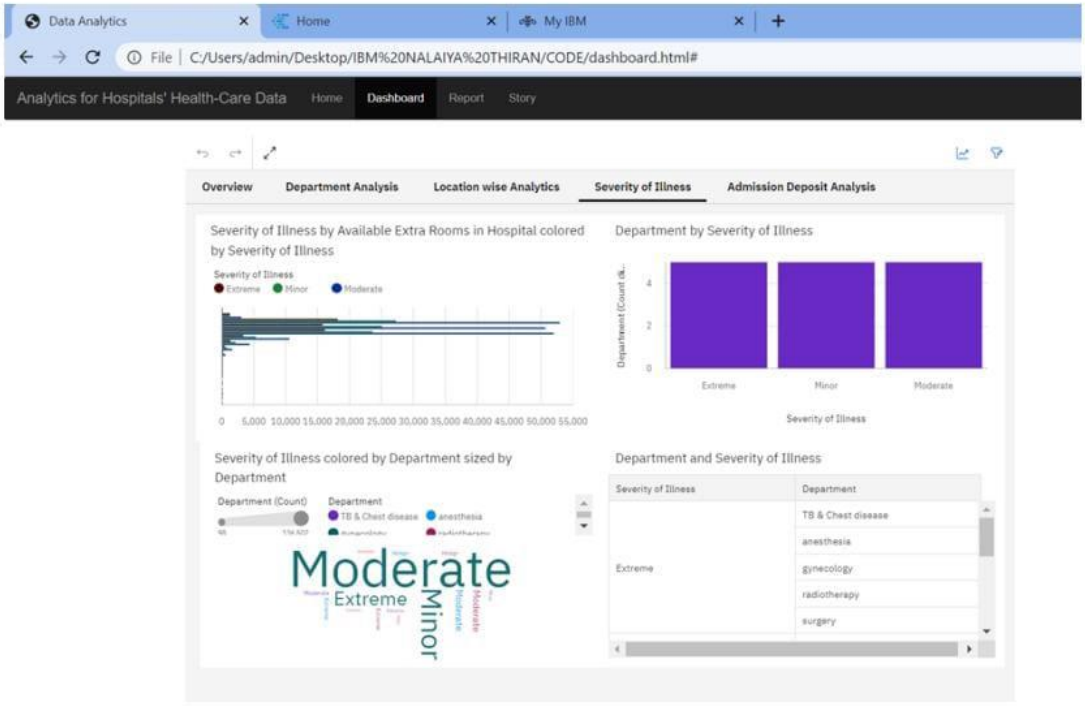


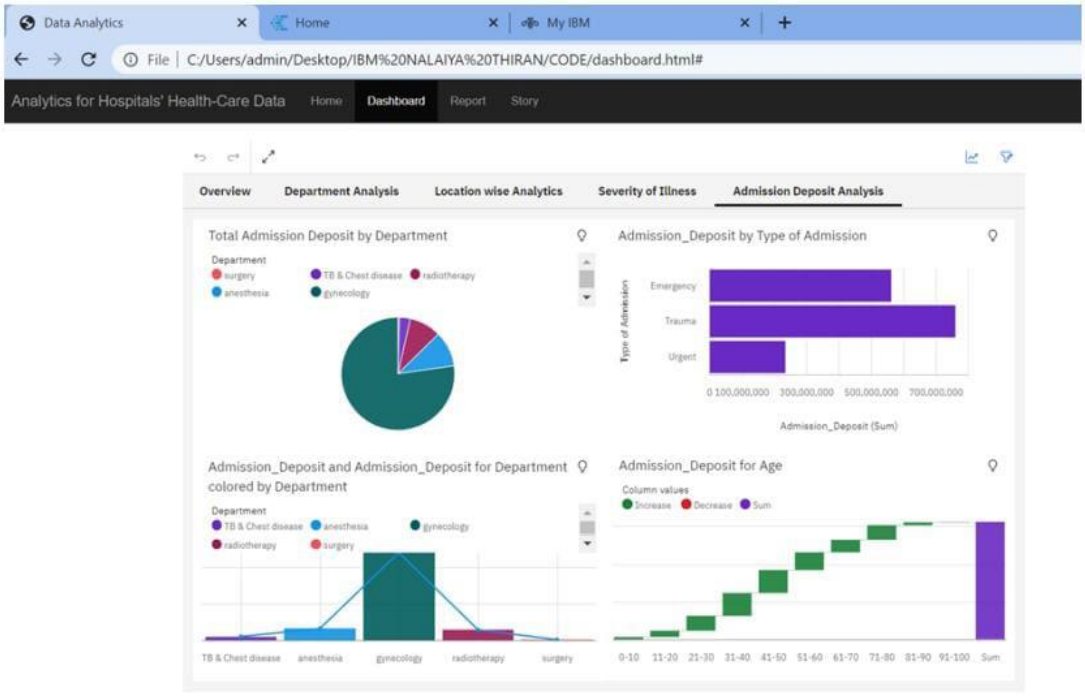


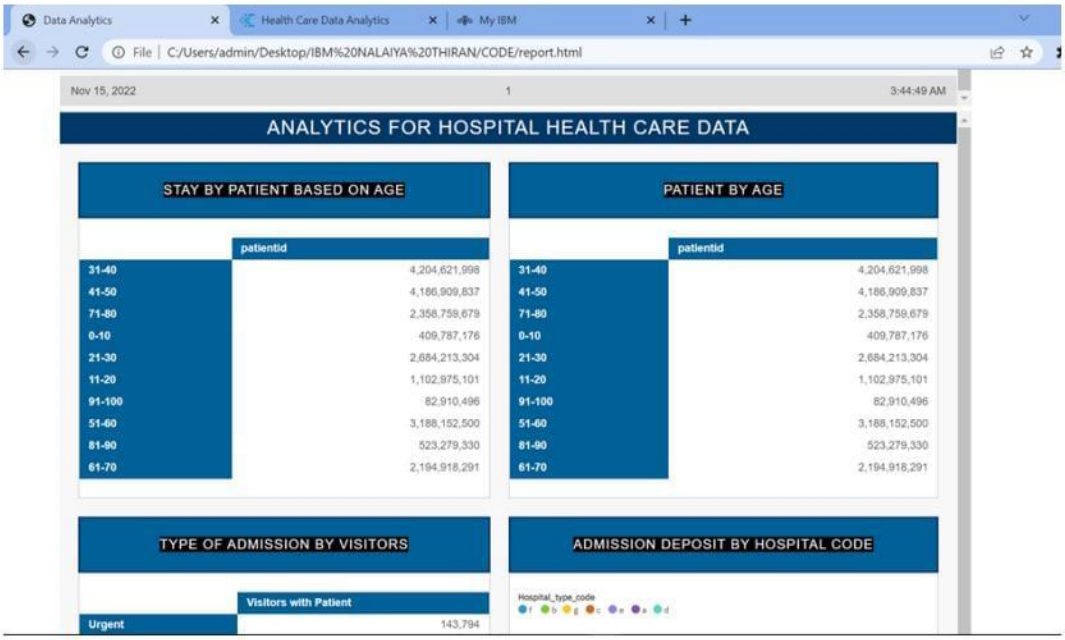

Data Analytics
Home
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/dashboard.html#
Analytics for Hospitals' Health-Care Data
Home
Dashboard
Report
Story
Overview
Department Analysis
Location wise Analytics
Severity of Illness
Admission Deposit Analysis
Severity of Illness by Available Extra Rooms in Hospital colored by Severity of Illness
Severity of Illness
Extreme
Minor
Moderate
0
5,000
10,000
15,000
20,000
25,000
30,000
35,000
40,000
45,000
50,000
55,000
Department by Severity of Illness
Department (Count di...
4
2
0
Extreme
Minor
Moderate
Severity of Illness
Severity of Illness colored by Department sized by Department
Department (Count)
Department
TB & Chest disease
anesthesia
Moderate
Extreme
Minor
Department and Severity of Illness
Severity of Illness
Department
Extreme
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery

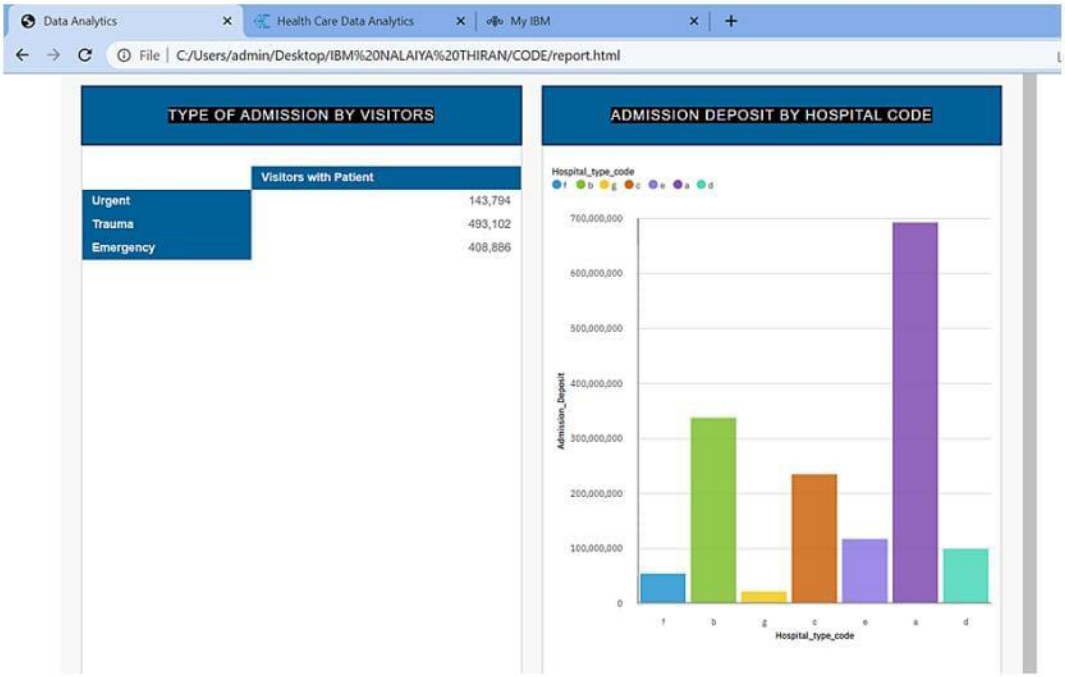

Data Analytics
Home
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/dashboard.html#
Analytics for Hospitals' Health-Care Data
Home
Dashboard
Report
Story
Overview
Department Analysis
Location wise Analytics
Severity of Illness
Admission Deposit Analysis
Total Admission Deposit by Department
Department
surgery
TB & Chest disease
radiotherapy
anesthesia
gynecology
Admission_Deposit by Type of Admission
Type of Admission
Emergency
Trauma
Urgent
0
100,000,000
300,000,000
500,000,000
700,000,000
Admission_Deposit (Sum)
Admission_Deposit and Admission_Deposit for Department colored by Department
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
Admission_Deposit for Age
Column values
Increase
Decrease
Sum
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
Sum

b:FEATURE 2


Data Analytics
Home
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/dashboard.html#
Analytics for Hospitals' Health-Care Data
Home
Dashboard
Report
Story
Overview
Department Analysis
Location wise Analytics
Severity of Illness
Admission Deposit Analysis
Severity of Illness by Available Extra Rooms in Hospital colored by Severity of Illness
Severity of Illness
Extreme
Minor
Moderate
0
5,000
10,000
15,000
20,000
25,000
30,000
35,000
40,000
45,000
50,000
55,000
Department by Severity of Illness
Department (Count di...
Extreme
Minor
Moderate
Severity of Illness
Severity of Illness colored by Department sized by Department
Department (Count)
Department
TB & Chest disease
anesthesia
Moderate
Extreme
Minor
Department and Severity of Illness
Severity of Illness
Department
TB & Chest disease
anesthesia
Extreme
gynecology
radiotherapy
surgery


Data Analytics
Home
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/dashboard.html#
Analytics for Hospitals' Health-Care Data
Home
Dashboard
Report
Story
Overview
Department Analysis
Location wise Analytics
Severity of Illness
Admission Deposit Analysis
Total Admission Deposit by Department
Department
surgery
TB & Chest disease
radiotherapy
anesthesia
gynecology
Admission_Deposit by Type of Admission
Type of Admission
Emergency
Trauma
Urgent
0
100,000,000
300,000,000
500,000,000
700,000,000
Admission_Deposit (Sum)
Admission_Deposit and Admission_Deposit for Department colored by Department
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
Admission_Deposit for Age
Column values
Increase
Decrease
Sum
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
Sum


Data Analytics
Health Care Data Analytics
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/report.html
Nov 15, 2022
1
3:44:49 AM
ANALYTICS FOR HOSPITAL HEALTH CARE DATA
STAY BY PATIENT BASED ON AGE
patientid
31-40 4,204,621,998
41-50 4,186,909,837
71-80 2,358,759,679
0-10 409,787,176
21-30 2,684,213,304
11-20 1,102,975,101
91-100 82,910,496
51-60 3,188,152,500
81-90 523,279,330
61-70 2,194,918,291
PATIENT BY AGE
patientid
31-40 4,204,621,998
41-50 4,186,909,837
71-80 2,358,759,679
0-10 409,787,176
21-30 2,684,213,304
11-20 1,102,975,101
91-100 82,910,496
51-60 3,188,152,500
81-90 523,279,330
61-70 2,194,918,291
TYPE OF ADMISSION BY VISITORS
Visitors with Patient
Urgent 143,794
ADMISSION DEPOSIT BY HOSPITAL CODE
Hospital_type_code


Data Analytics
Health Care Data Analytics
My IBM
File | C:/Users/admin/Desktop/IBM%20NALAIYA%20THIRAN/CODE/report.html
TYPE OF ADMISSION BY VISITORS
Visitors with Patient
Urgent 143,794
Trauma 493,102
Emergency 408,886
ADMISSION DEPOSIT BY HOSPITAL CODE
Hospital_type_code
f b g c e a d
700,000,000
600,000,000
500,000,000
400,000,000
300,000,000
200,000,000
100,000,000
0
Admission_Deposit
f b g c e a d
Hospital_type_code

CHAPTER 8

TESTING

a: 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 story page
- verify filters are working

a: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

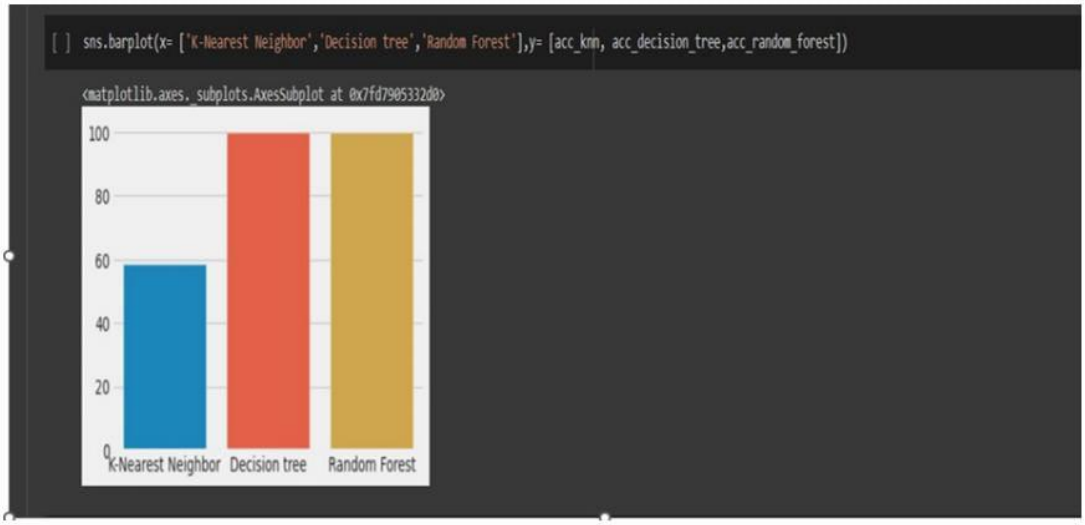
This report shows the number of test cases that have passed, failed, and untested

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

CHAPTER 9

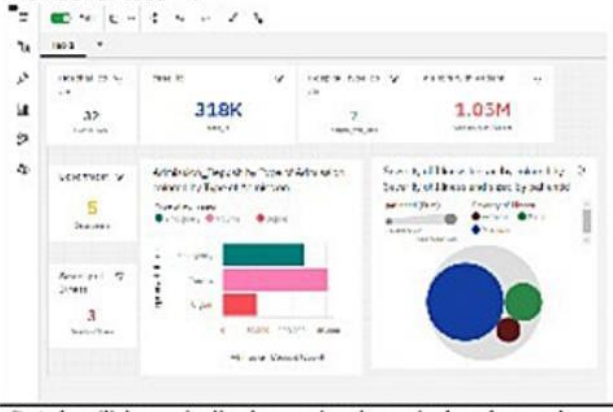
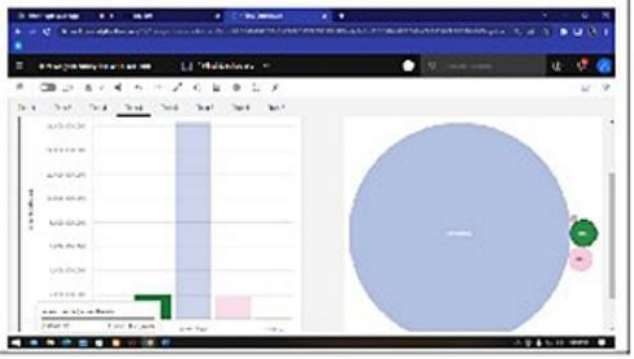
RESULTS

a: PERFORMANCE METRICS

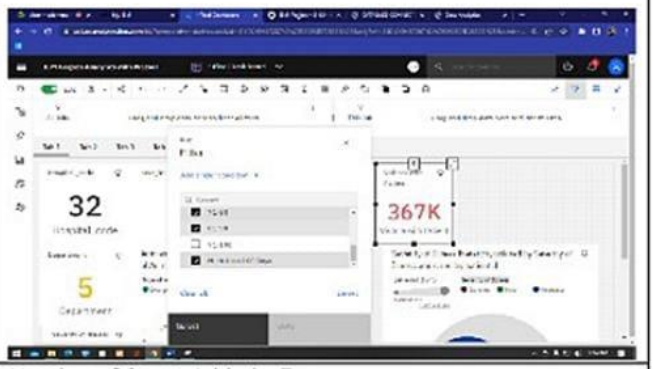
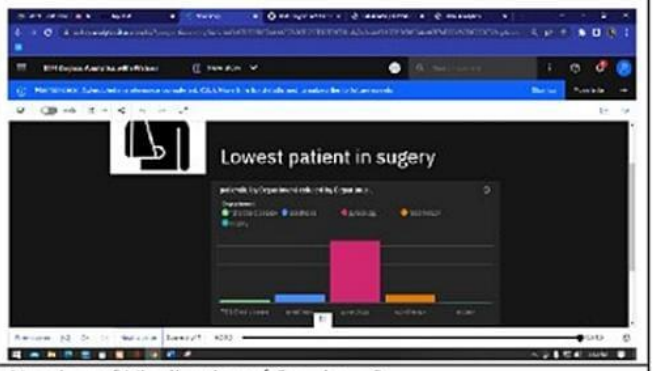


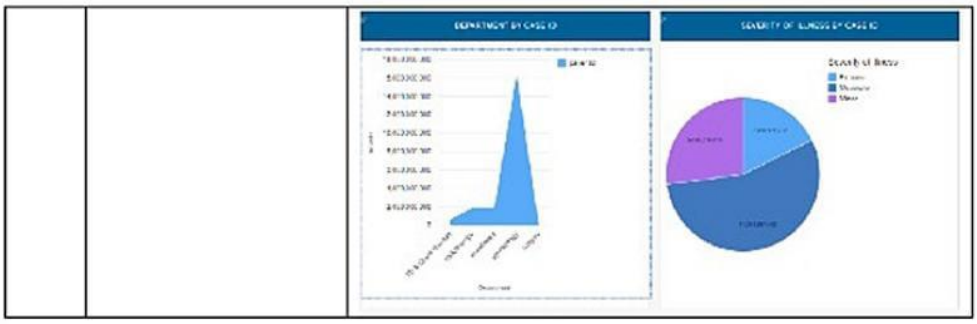
Model Performance Testing:

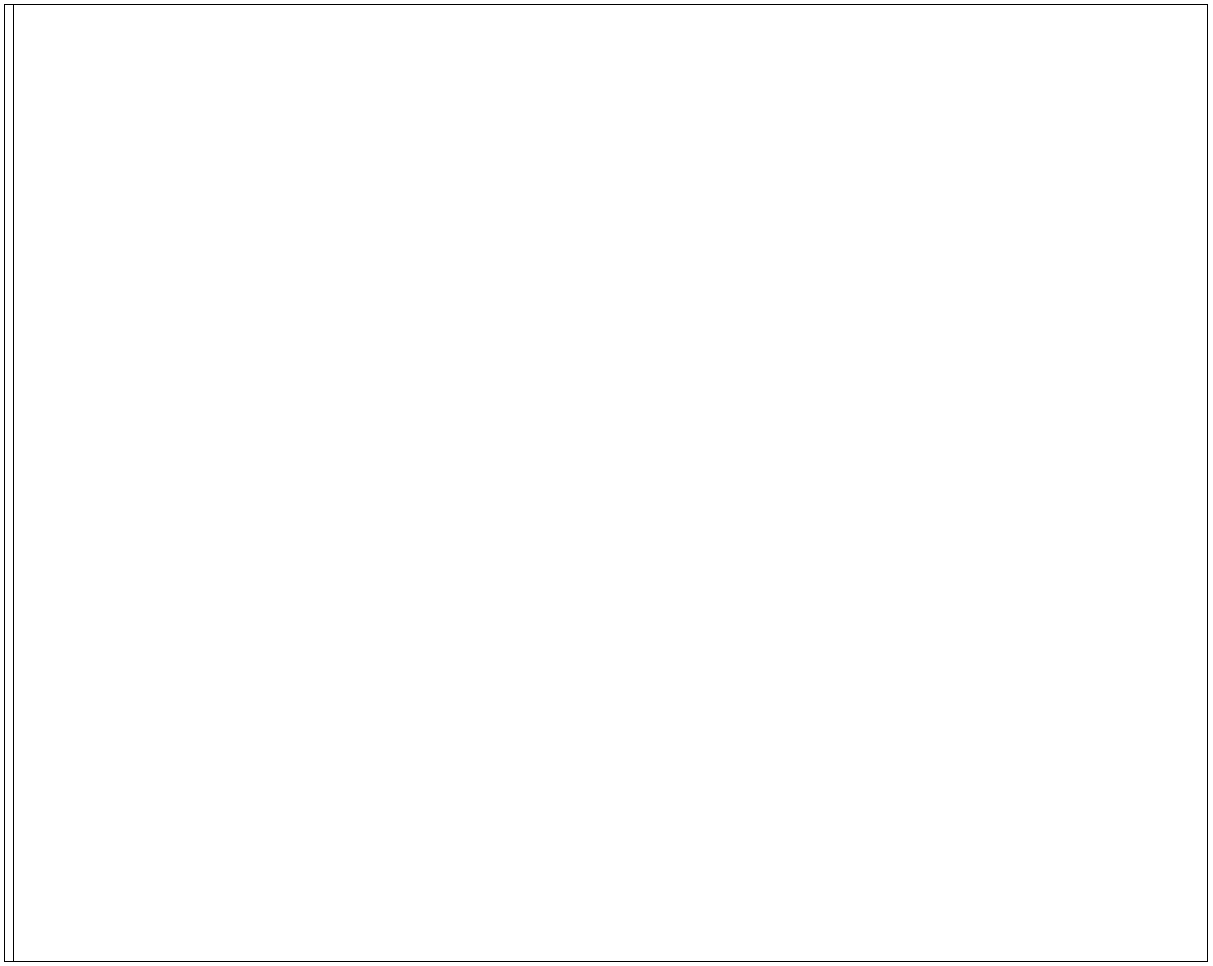
Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	Number of Visulizations / Graphs – 22 Number of tabs – 5 
2.	Data Responsiveness	Data's will dynamically changed and graph also changed. 

3.	Amount Data to Rendered (DB2 Metrics)	Number of rows read – 318438 Number of rows loaded – 318438 Number of rows rejected – 0

4.	Utilization of Data Filters	We created filters for Dashboards which is perfectly working. 
5.	Effective User Story	Number of Scene Added – 7 Animations are perfectly displayed. Images are perfectly rendered. 
6.	Descriptive Reports	Number of Visulizations / Graphs – 6





CHAPTER 10

ADVANTAGES AND DISADVANTAGES

ADVANTAGES:

i. Cost-effective use of technology

ii. Improved project management
iii. Sustaining the improvements in the result

iv. Boosting hospital capacity
v. Enhance the qualityand efficiency of healthcare

vi. benefit areas like emergency preparation, charting, administration,compliance, and financial management.
vii. Analysing clinical data to improvemedical research

viii. Using patient data to improve health outcomes
ix. Gaining operational insights from healthcareprovider data

x. Improved staffing throughhealth business management analytic s
xi. Early detection of disease.
xii. Prevention of unnecessary doctor's visit
xiii. Discovery of new drugs.
xiv. More accurate calculation of health insurancerates.

xv. More effective sharing of patient data

DISADVANTAGES:

REPLACING MEDICAL PERSONNEL:

Application of technology in every sphere of human life is improving the way things are done. These technologies are are also posing some threat to world ofworks. Robotics are replacing human labour.

DATA SAFETY:

Data security is another challenge in applying big data in healthcare. Big datastorage is usually targets of hackers. This endangers the safety of medical data.Healthcare organisations are very much concerned about the safety of patients' sensitivepersonal data. For this, all healthcare applications must meet the requirement for data security and be HIPAA compliant before they can be deployed for healthcare services.

PRIVACY:

One of the major drawbacks in the application of big data in healthcare industry is the issue of lack of privacy. Application of big data technologies involves monitoring of patient's data, tracking of medical inventoryand assets, o rganizingcollected data, and visualization of data on the dashboard and the reports. So visualization of sensitive medical data especially that of the patients creates negative impression of big data as it violets privacy

MAN POWER:

Applying big data solutions in healthcare requires special skills, and such kills are scarce. Handling of big data requires the combination of medical, technological and statistical knowledge.

CHAPTER 11

CONCLUSION

The impact of data analytics in healthcare has already ma de a substanticcal difference in the ability of healthcare providers to offer patients high-quality care inan efficient, cost-effective manner. However, the role of data analytics in improvingpatientoutcomes and healthcare processes continues to gr ow and expand as moretypes of data become available and new tools are developed that make the results of the analytics clear and easy for healthcare professionals to access.

Realizing the potential of data analytics to transform the healthcare industry begins by understanding how the technology can be applied to address healthcare providers' challenges, including staff recruitment and utilization,

operational efficiencies, and enhanced patientexperiences. Patient-centered healthcare depends on knowing what patients want and need. Data analytics holds the key to unlockingthis vital information.

CHAPTER 12

FUTURE SCOPE

Artificial Intelligence (AI) will play a significant role in data analyti cs in

healthcare for the next decade. For example, the field of AI-enabled clinical decisionsupport is just emerging. This type of support can compare patients who fit similar profileswithin a system,then it can alert doctorsto trends in d ata that may havebeen overlooked. The use of big data in healthcare will include testing for drug interactions that small studies are unlikely to catch and prevent patients from takingharmful drug combinations.

Decisions made by physicians, like what test or treatments t o give a particular patient, makeup 80-90% of all healthcare spending, so using artificial intelligence to make more educated decisions will bring down healthcare costs. It's crucial to have informed leaders at the vanguard of these innovations in healthcare.

CHAPTER 13

APPENDIX

SOURCE PAGE:

HOMEPAGE:

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
   <head>
      <meta charset="utf-8">
      <title>Analytics For Hospital Data</title>
      <link rel="stylesheet" href="#">
      <meta name="viewport" content="width=device-width, initial-scale=1.0">
      <link href='https://unpkg.com/boxicons@2.1.2/css/boxicons.min.css'
rel='stylesheet'>
   </head>
   <body>
    <ul>
        <li><a class="active" href="Home.html"><strong>Home</strong></a></li>
        <li><a href="dashboard.html"><strong>Dashboard</strong></a></li>
        <li><a href="Report.html"><strong>Report</strong></a></li>
        <li><a href="Stories.html"><strong>Stories</strong></a></li>
```

```
        <li><a href="https://eu1.ca.analytics.ibm.com/bi/?perspective=home"
target="_blank"><strong>IBM Login</strong></a></li>
      </ul>

          <h1>Analytics For Hospitals' Health-Care Data</h1>

          <h2>Team ID: PNT2022TMID29634</h2>
          <h3>Team Members</h3>
          <h4>Dilli Prasath S - 310819104023<br>
              Dhakshnamoorthy B - 310819104020<br>
              Deepika K - 310819104019<br>
              Afra Jabeen S - 310819104004</h4>

          <p>Category: Data Analytics<br><br>

            <strong>Skills Required:</strong><br>
            Exploratory Data Analysis,IBM Cloud<br><br>

            <strong>Project Description:</strong><br>
            Recent Covid-19 Pandemic has raised alarms over one of the most
overlooked areas to focus: Healthcare<br>
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.<br>
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.<br>
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.<br><br>
<strong>Goal:</strong><br>
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.<br><br>
<strong>Technical Architecture:</strong><br><br>
<img src="ibm image.png" alt="IBM Cognos Analytics">
          </p>

          <style>
              @import
url('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;500;600&dis
play=swap');

body {margin: 10px;;
```

```
    padding: 40PX;
    box-sizing: border-box;
    font-family: 'Poppins', sans-serif;
    background: -webkit-linear-gradient(right, #a445b2, #9542fa);}

h2,h3,h4 {
      text-align: center;
    }

ul {
  list-style-type: none;
  margin: 5px;
  padding: 8px;
  overflow: hidden;
  background-color: #333;
  border-radius: 12px;
  position: fixed;
  top: 0;
  width: 100%;
}

li {
  float: left;
  display: flex;
    justify-content: center;
    align-items: center;
    flex-direction: column;
    text-decoration: none;
    transition: all 0.5s cubic-bezier(0.68, -0.55, 0.265, 1.55);
}

li a {
  display: block;
  color: rgb(246, 239, 239);
  text-align: center;
  padding: 14px 14px 16px 16px;
  text-decoration:none;
}

li a:hover{
  background: -webkit-linear-gradient(left, #f1dbf5, #b280ee);
  border-radius: 12px;
}

</style>
  </body>
</html>
```

DASHBOARD PAGE :

```
<!DOCTYPE html>

<html lang="en" dir="ltr">
   <head>
      <meta charset="utf-8">
      <title>Analytics For Hospital Data - Dashboard</title>
      <link rel="stylesheet" href="#">
      <meta name="viewport" content="width=device-width, initial-scale=1.0">
      <link href='https://unpkg.com/boxicons@2.1.2/css/boxicons.min.css'
rel='stylesheet'>
   </head>
   <body>
    <ul>
        <li><a class="active" href="Home.html"><strong>Home</strong></a></li>
        <li><a href="dashboard.html"><strong>Dashboard</strong></a></li>
        <li><a href="Report.html"><strong>Report</strong></a></li>
        <li><a href="Stories.html"><strong>Stories</strong></a></li>
      </ul>

      <h1>Analytics For Hospitals' Health-Care Data - Dashboard</h1>

      <iframe
src="https://eu1.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.m
y_folders%2FNew%2Bdashboard&amp;closeWindowOnLastView=true&amp;ui_appbar=false
&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=dashboard
&amp;subView=model0000018470453873_00000002" width="1300" height="550"
frameborder="10px" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>
      <p><a
href="https://eu1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_f
olders%2FNew%2Bdashboard&action=view&mode=dashboard&subView=model0000018470453
873_00000002" target="iframe_a"><strong>IBM Cognos Analytics -
Dashboard</strong></a></p>

      <style>

        @import
url('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;500;600&dis
play=swap');

body {margin: 10px;;
padding: 40PX;
box-sizing: border-box;
font-family: 'Poppins', sans-serif;
background: -webkit-linear-gradient(right, #a445b2, #9542fa);}
```

```
iframe{
    display: flex;
}
p{
    text-align: center;
}

ul {
list-style-type: none;
margin: 5px;
padding: 8px;
overflow: hidden;
background-color: #333;
border-radius: 12px;
position: fixed;
top: 0;
width: 100%;
}

li {
float: left;
display: flex;
justify-content: center;
align-items: center;
flex-direction: column;
text-decoration: none;
transition: all 0.5s cubic-bezier(0.68, -0.55, 0.265, 1.55);
}

li a {
display: block;
color: rgb(246, 239, 239);
text-align: center;
padding: 14px 16px;
text-decoration:none;
}

li a:hover{
background: -webkit-linear-gradient(left, #f1dbf5, #b280ee);
border-radius: 12px;
}

</style>
</body>
</html>
```

REPORT PAGE :

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
   <head>
      <meta charset="utf-8">
      <title>Analytics For Hospital Data - Dashboard</title>
      <link rel="stylesheet" href="#">
      <meta name="viewport" content="width=device-width, initial-scale=1.0">
      <link href='https://unpkg.com/boxicons@2.1.2/css/boxicons.min.css'
rel='stylesheet'>
   </head>
   <body>
    <ul>
        <li><a class="active" href="Home.html"><strong>Home</strong></a></li>
        <li><a href="dashboard.html"><strong>Dashboard</strong></a></li>
        <li><a href="Report.html"><strong>Report</strong></a></li>
        <li><a href="Stories.html"><strong>Stories</strong></a></li>
      </ul>

      <h1>Analytics For Hospitals' Health-Care Data - Report</h1>

      <iframe
src="https://eu1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FData%2BModule%
2FTrained%2B-
%2Breport&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=fal
se&amp;shareMode=embedded&amp;action=edit" width="1300" height="550"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>
      <p><a
href="https://eu1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FData%2BModule
%2FTrained%2B-%2Breport&action=edit" target="iframe_a"><strong>IBM Cognos
Analytics - Report</strong></a></p>

      <style>

        @import
url('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;500;600&dis
play=swap');

body {margin: 10px;;
padding: 40PX;
box-sizing: border-box;
font-family: 'Poppins', sans-serif;
background: -webkit-linear-gradient(right, #a445b2, #9542fa);}

iframe{
    display: flex;
```

```
}
p{
    text-align: center;
}

ul {
list-style-type: none;
margin: 5px;
padding: 8px;
overflow: hidden;
background-color: #333;
border-radius: 12px;
position: fixed;
top: 0;
width: 100%;
}

li {
float: left;
display: flex;
justify-content: center;
align-items: center;
flex-direction: column;
text-decoration: none;
transition: all 0.5s cubic-bezier(0.68, -0.55, 0.265, 1.55);
}

li a {
display: block;
color: rgb(246, 239, 239);
text-align: center;
padding: 14px 16px;
text-decoration:none;
}

li a:hover{
background: -webkit-linear-gradient(left, #f1dbf5, #b280ee);
border-radius: 12px;
}

</style>
</body>
</html>
```

STORY:

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
   <head>
      <meta charset="utf-8">
      <title>Analytics For Hospital Data - Stories</title>
      <link rel="stylesheet" href="#">
      <meta name="viewport" content="width=device-width, initial-scale=1.0">
      <link href='https://unpkg.com/boxicons@2.1.2/css/boxicons.min.css'
rel='stylesheet'>
   </head>
   <body>
    <ul>
        <li><a class="active" href="Home.html"><strong>Home</strong></a></li>
        <li><a href="dashboard.html"><strong>Dashboard</strong></a></li>
        <li><a href="Report.html"><strong>Report</strong></a></li>
        <li><a href="Stories.html"><strong>Stories</strong></a></li>
      </ul>

      <h1>Analytics For Hospitals' Health-Care Data - Report</h1>

      <iframe
src="https://eu1.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.my_fo
lders%2FTrained%2BData%2B-
%2BStory&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=fals
e&amp;shareMode=embedded&amp;action=view&amp;sceneId=model00000184991994ba_000
00000&amp;sceneTime=0" width="1300" height="550" frameborder="10"
gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
      <p><a
href="https://eu1.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.my_folde
rs%2FTrained%2BData%2B-
%2BStory&action=view&sceneId=model00000184991994ba_00000000&sceneTime=0"
target="iframe_a"><strong>IBM Cognos Analytics - Stories</strong></a></p>

      <style>

        @import
url('https://fonts.googleapis.com/css2?family=Poppins:wght@300;400;500;600&dis
play=swap');

body {margin: 10px;;
padding: 40PX;
box-sizing: border-box;
font-family: 'Poppins', sans-serif;
background: -webkit-linear-gradient(right, #a445b2, #9542fa);}

iframe{
```

```
    display: flex;
}
p{
    text-align: center;
}

ul {
list-style-type: none;
margin: 5px;
padding: 8px;
overflow: hidden;
background-color: #333;
border-radius: 12px;
position: fixed;
top: 0;
width: 100%;
}

li {
float: left;
display: flex;
justify-content: center;
align-items: center;
flex-direction: column;
text-decoration: none;
transition: all 0.5s cubic-bezier(0.68, -0.55, 0.265, 1.55);
}

li a {
display: block;
color: rgb(246, 239, 239);
text-align: center;
padding: 14px 16px;
text-decoration:none;
}

li a:hover{
background: -webkit-linear-gradient(left, #f1dbf5, #b280ee);
border-radius: 12px;
}

</style>
</body>
</html>
```

Importing required Packages

In [72]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
sns.set_style("darkgrid")
plt.style.use("dark_background")
```

Importing the dataset

In [73]:
```
train = pd.read_csv('/content/input/training_data.csv')
test = pd.read_csv('/content/input/testing_data.csv')
Paramters_Description = pd.read_csv('/content/input/parameter_description.csv')
sample = pd.read_csv('/content/input/testing_target.csv')
```

Viewing dataset

In [74]:
```
train.head(5)
```

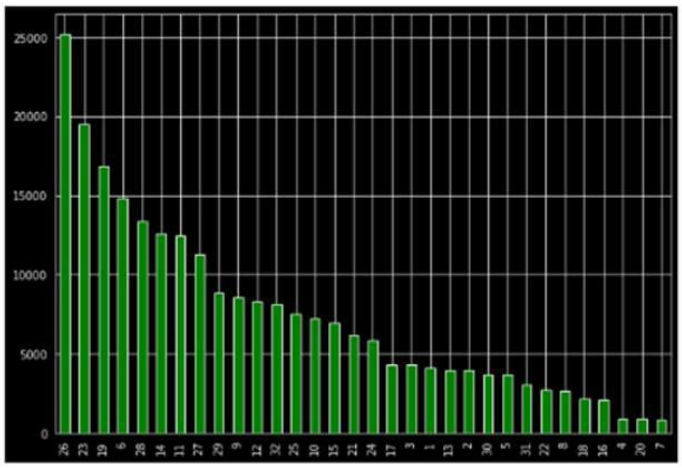
Out[74]:

	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
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0
2	3	10	e	1	X	2	anesthesia	S	E	2.0
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0
4	5	26	b	2	Y	2	radiotherapy	S	D	2.0

Dataset Column Description

```
Paramters_Description
```

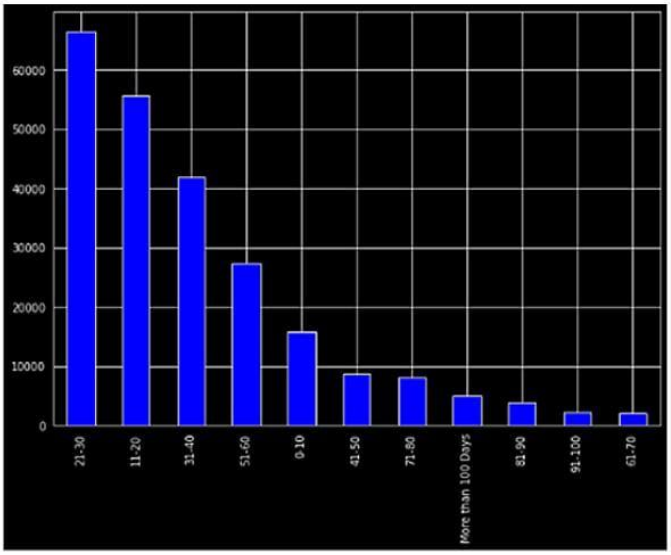
	Column	Description
0	case_id	It is identity number given by hospital admini...
1	Hospital_code	It is the code (identity number) given to the ...
2	Hospital_type_code	It is the unique code given to the type of hos...
3	City_Code_Hospital	It is the code given to the city where the hos...
4	Hospital_region_code	It is the code given to the region where the h...
5	Available_Extra_Rooms_in_Hospital	It will display the number of rooms that are s...
6	Department	The department that is overlooking the patient...
7	Ward_Type	The unique code given to the type of ward to w...
8	Ward_Facility_Code	The unique code given to the facility in the w...
9	Bed_Grade	It is the quality or condition of the bed in t...
10	patientid	It is the unique identity value given to the p...
11	City_Code_Patient	It is the unique identity code given to the ci...
12	Type_of_Admission	It is the admission type registered in the hos...
13	Severity_of_Illness	It is the severity level of the patients' illn...
14	Visitors_with_Patient	Number of the visitors with the patients to ta...
15	Age	It is the age of patients. It is given in peri...
16	Admission_Deposit	It is the deposit amount that the patient paid...
17	Stay	It is the Length Of Stay (LOS) of patients. I...



Stay

```
train.Stay.value_counts()
```

```
21-30                 66497
11-20                 55691
31-40                 41951
51-60                 27458
0-10                  15866
41-50                  8665
71-80                  8061
More than 100 Days     5029
81-90                  3821
91-100                 2179
61-70                  2090
Name: Stay, dtype: int64
```

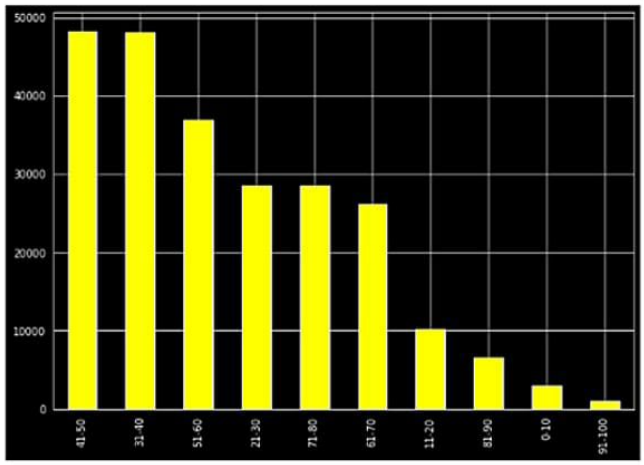
Age

```
train.Age.value_counts()
```

```
41-50    48272
31-40    48106
51-60    36969
21-30    28555
71-80    28552
61-70    26139
11-20    10141
```

```
81-90      6578
0-10       3030
91-100      966
Name: Age, dtype: int64
```

```
#Age distribution
plt.figure(figsize=(10,7))
train.Age.value_counts().plot(kind="bar", color = ['Yellow'])
```



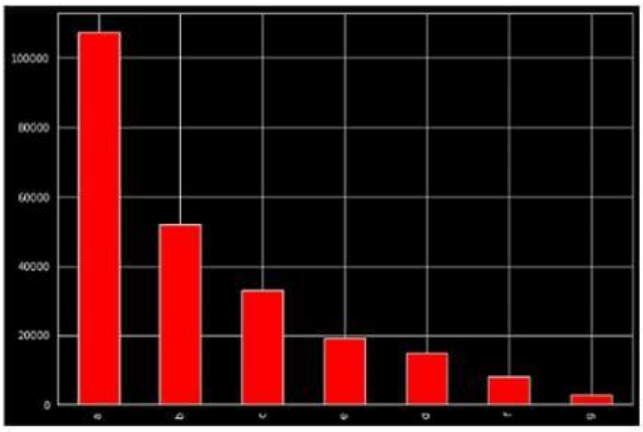
Hospital_type_code

```
train.Hospital_type_code.value_counts()
```

```
a    107545
b     51925
```

```
c    32995
e    19105
d    14833
f     8166
g     2740
Name: Hospital_type_code, dtype: int64
```

```
#Hospital_type_code distribution
plt.figure(figsize=(10,7))
train.Hospital_type_code.value_counts().plot(kind="bar", color = ['Red'])
```

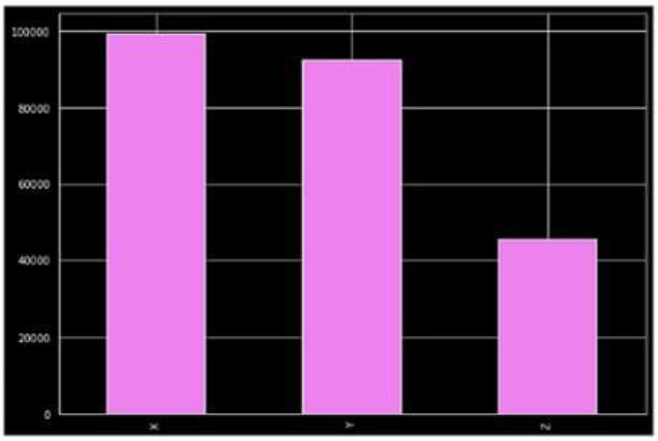


Hospital_region_code

```
train.Hospital_region_code.value_counts()
```

```
X    99568
Y    92214
Z    45527
Name: Hospital_region_code, dtype: int64
```

```
#Hospital_region_code distribution
plt.figure(figsize=(10,7))
train.Hospital_region_code.value_counts().plot(kind="bar", color = ['Violet'])
```



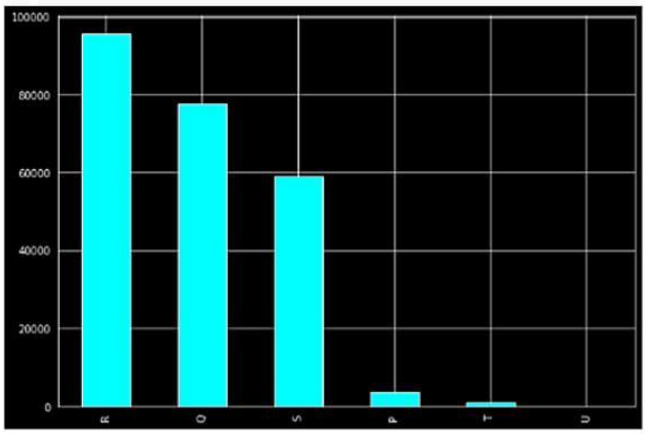
Available_Extra_Rooms_in_Hospital

```
train.Available_Extra_Rooms_in_Hospital.value_counts()
```

```
2     74877
3     68517
4     67756
5     13879
6      5344
1      4208
7      1876
8       622
9       144
10       46
```

```
R    95788
Q    77707
S    59022
P     3691
T     1092
U        9
Name: Ward_Type, dtype: int64
```

```
#Ward_Type distribution
plt.figure(figsize=(10,7))
train.Ward_Type.value_counts().plot(kind="bar", color = ['cyan'])
```



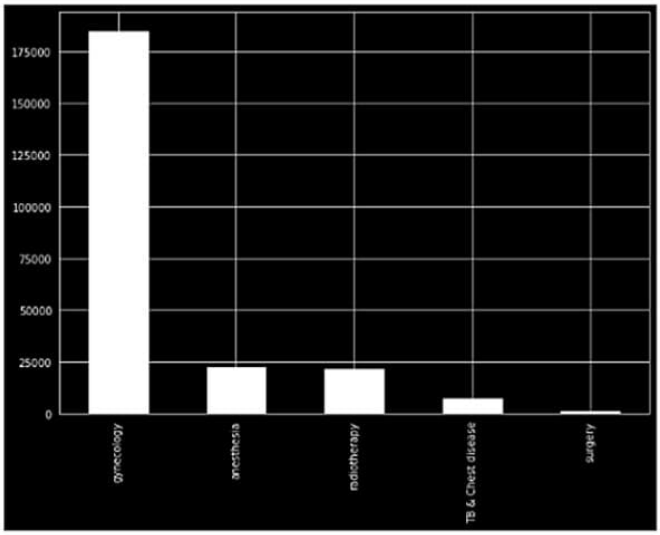
Ward_Facility_Code

```
train.Ward_Facility_Code.value_counts()
```

```
F    84438
E    41246
```

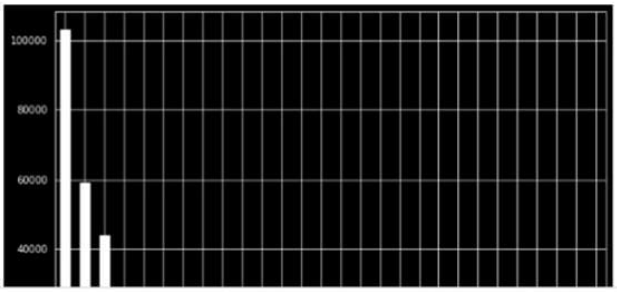
```
anesthesia              22557
radiotherapy            21725
TB & Chest disease       7017
surgery                   948
Name: Department, dtype: int64
```

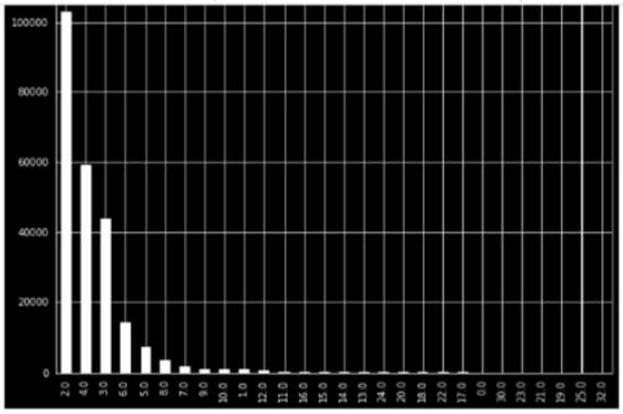
```
#Department distribution
plt.figure(figsize=(10,7))
train.Department.value_counts().plot(kind="bar", color = ['white'])
```



Ward_Type

```
train.Ward_Type.value_counts()
```





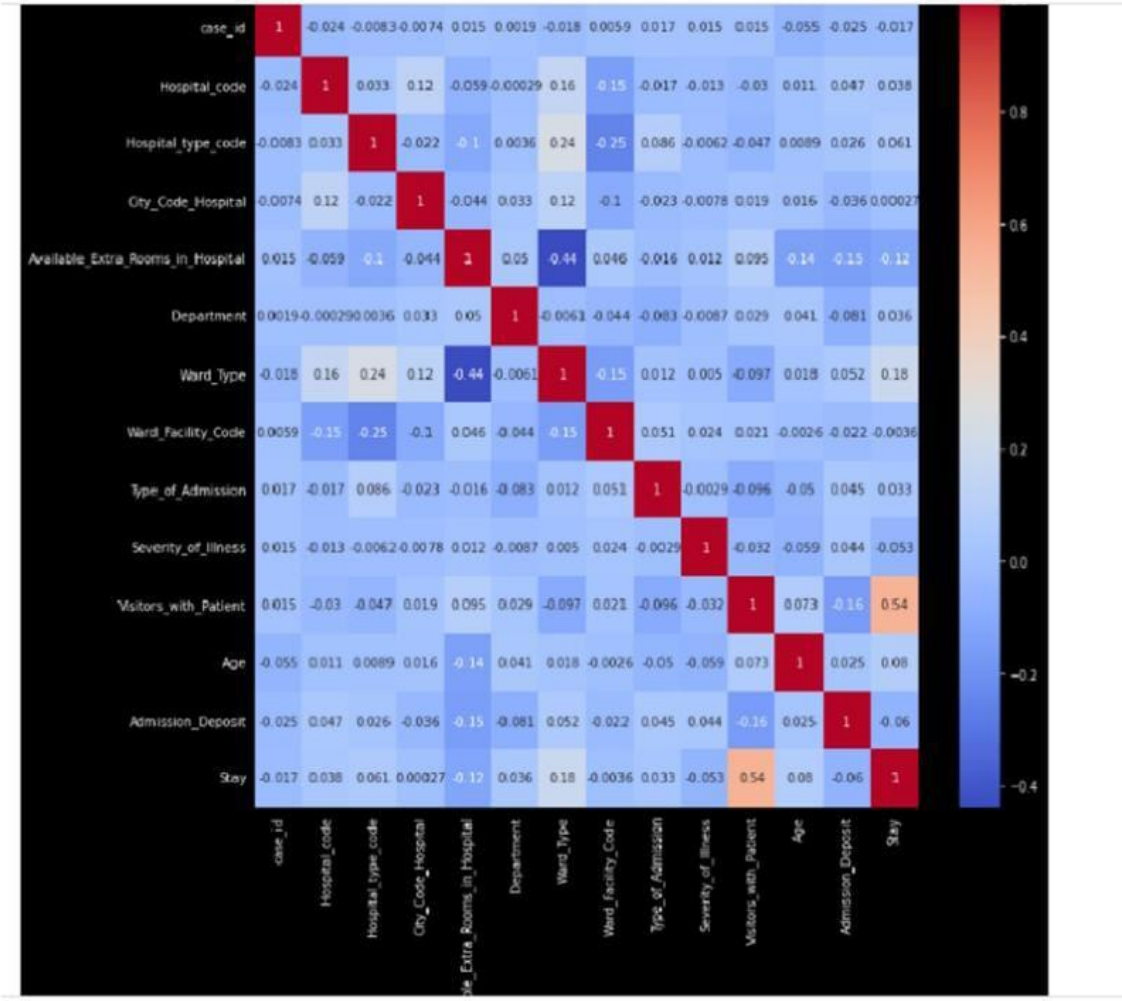
Severity of Illness

]:
```
train.Severity_of_Illness.value_counts()
```

]:
```
Moderate    134324
Minor        55665
Extreme      47319
Min              1
Name: Severity_of_Illness, dtype: int64
```

]:
```
#Severity_of_Illness distribution
plt.figure(figsize=(10,7))
train.Severity_of_Illness.value_counts().plot(kind="bar", color = ['red'])
```

]:



SSGitHub & Project Demo Links:

GithHub link: https://github.com/IBM-EPBL/IBM-Project-1054-1658337094

Project demo

link: https:/drive.google.com/file/d/1YIteqP82M5z45KuksjTocswXXMJ

OtQ45/vie w?usp=sharing