| | |
|---|---|
| **TEAM ID** | **PNT2022TMID19527** |
| **PROJECT TITLE** | **Analytics for Hospital Health-Care Data** |

# Analytics for Hospital Health-Care Data Report

## 1.INTRODUCTION

### 1.1 PROJECT OVERVIEW

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare Management. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital. This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. In addition, prior knowledge of LOS can aid in logistics such as room and bed allocation planning. What if I 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. 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.

### 1.2 PURPOSE

Healthcare data analytics refers to the process of working on raw datasets related to healthcare and analyzing them to find hidden patterns,trends,etc.,thus paving a way for further improvements at patient-level as well as business-level.

## 2.LITERATURE SURVEY

### 2.1 EXISTING PROBLEM

| S.NO | TITLE | AUTHOR | DESCRIPTION |
|---|---|---|---|
| 1 | Data-driven Methods for Typical Treatment Pattern Mining | Chonghui Guo and jingfeng Chen | A huge volume of digitized clinical data is generated and accumulated rapidly since the widespread adoption of Electronic Medical Records [EMRs] |

| | | | |
|---|---|---|---|
| 2 | Big Data in Supply Chain Management | Aniket Nargundkar and J.Kulkarni | A robust model is designed to help hospital administration to predict patient's Length Of Stay [LOS] to resolve the issues faced by hospital |
| 3 | A Systematic Mapping study | Anil Pandit and Sharvari Shukla | The current study performs a Systematic Literature Review (SLR) to synthesise prior research on thee applicability on Big Data Analytics (BDA) in healthcare |

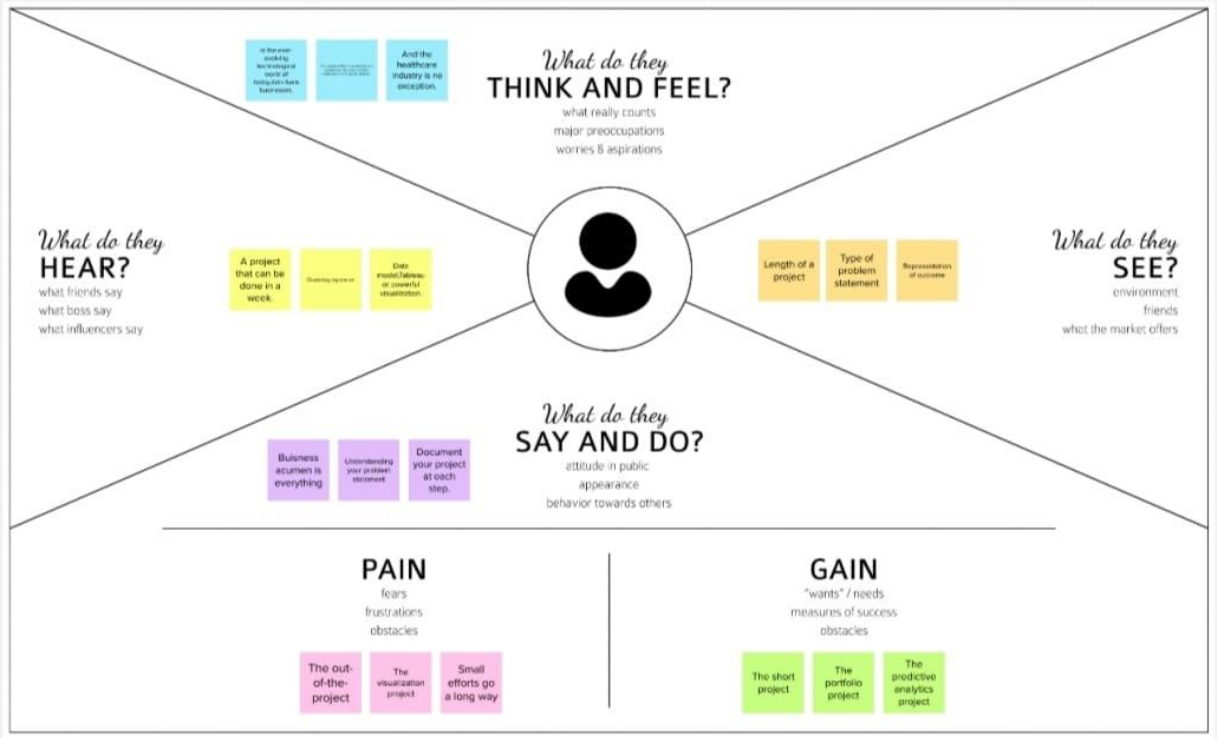
2.2 REFERENCE

2.3 PROBLEM STATEMENT

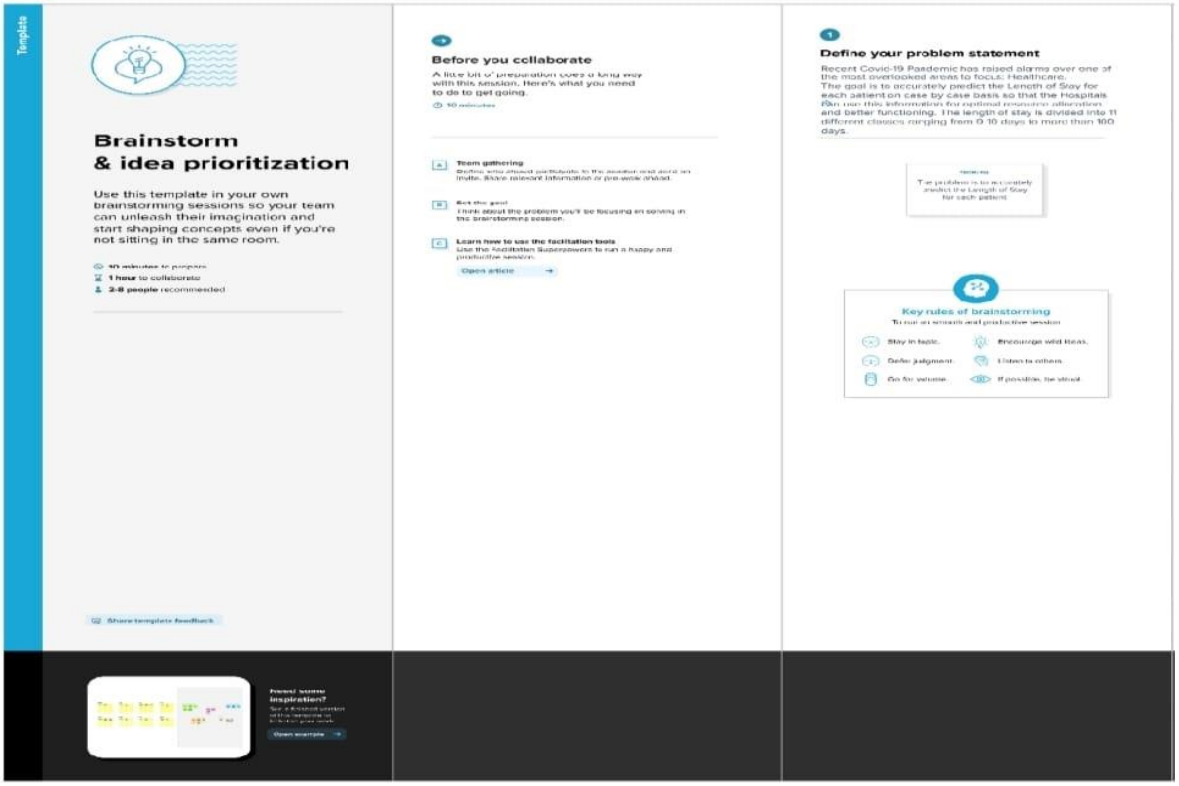
Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love. A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

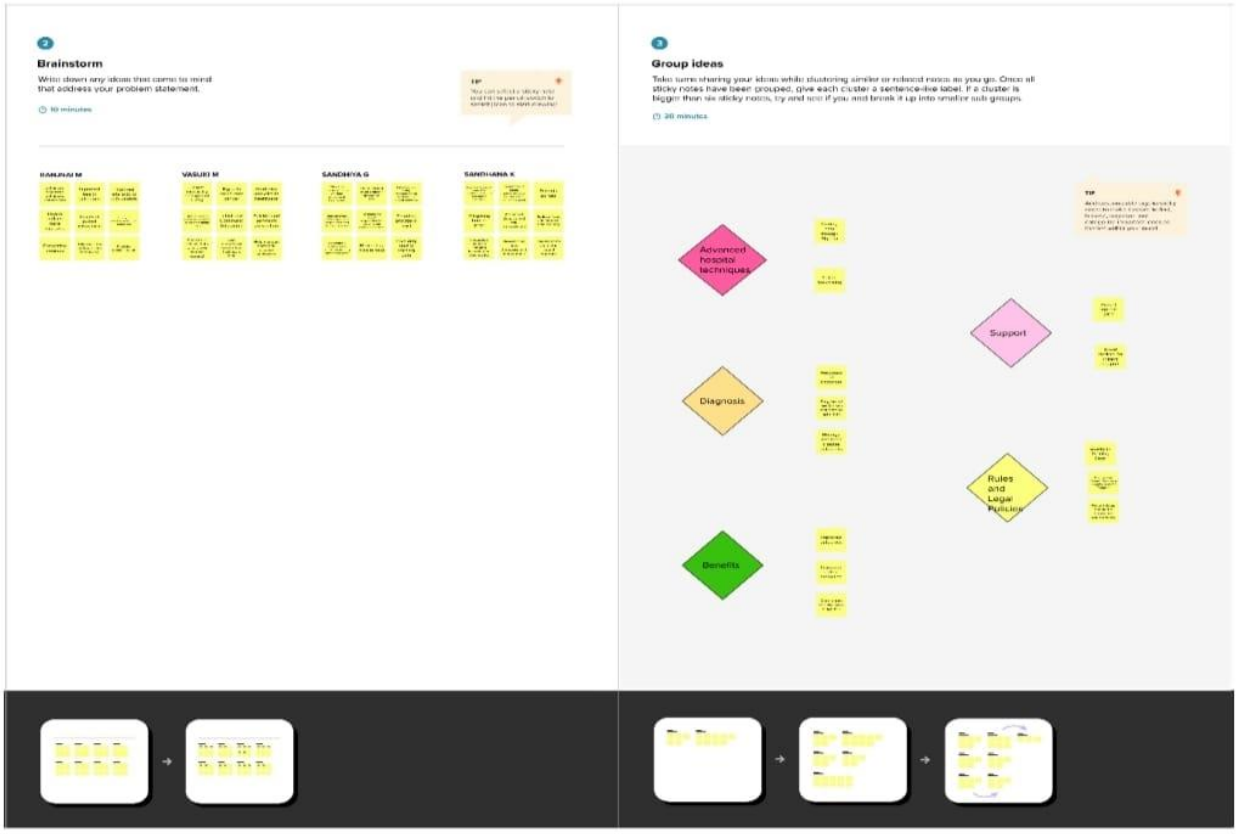
## 3. IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS




3.2 IDEATION AND 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
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
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
Define your problem statement
Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare. 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.
The problem is to accurately predict the Length of Stay for each patient
Key rules of brainstorming
To run an smooth and productive session
Stay in topic.
Encourage wild ideas.
Defer judgment.
Listen to others.
Go for volume.
If possible, be visual.
Need some inspiration?
Open example

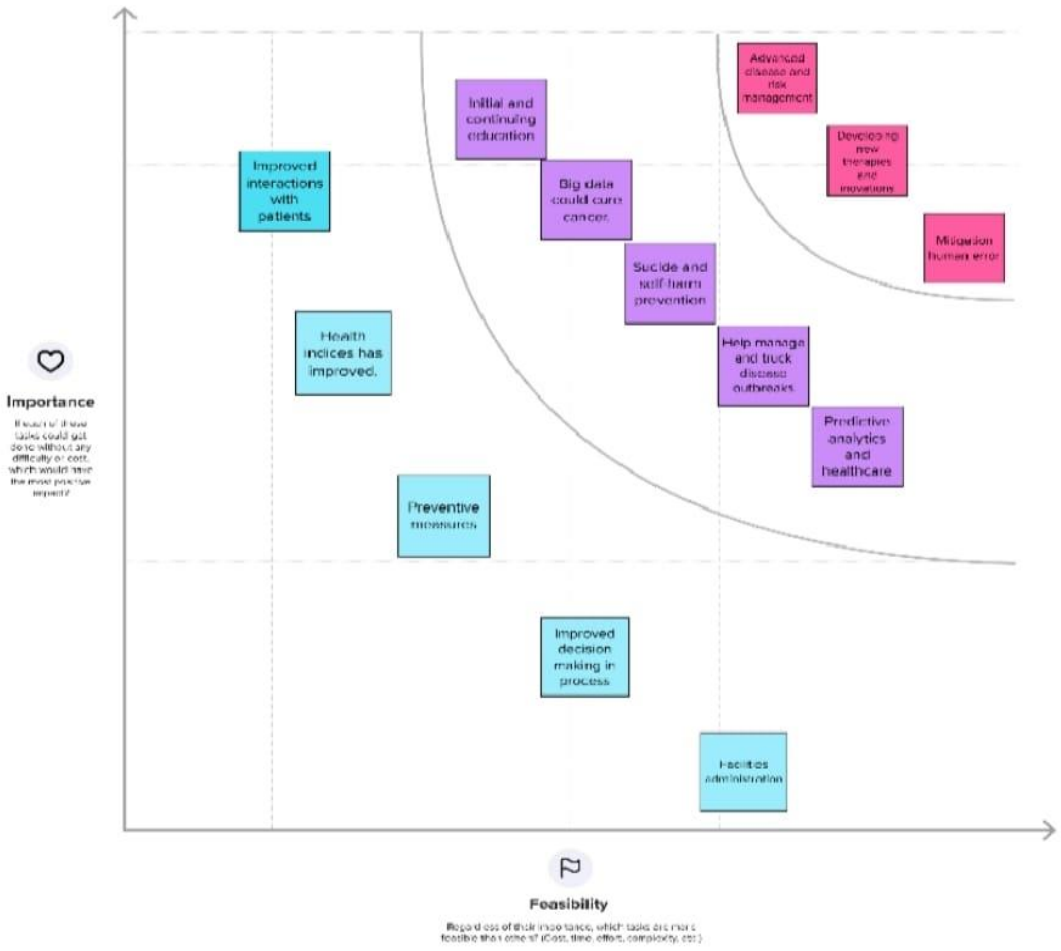

Brainstorm
Write down any ideas that come to mind that address your problem statement.
10 minutes
TIP
RANJANI M
VASUKI M
SANDHIYA G
SANDHANA K
Group ideas
Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like 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
TIP
Advanced hospital techniques
Support
Diagnosis
Rules and Legal Policies
Benefits

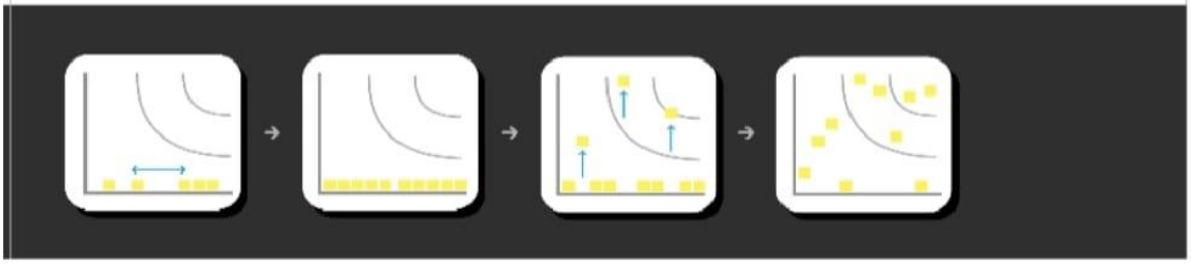
4

## 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

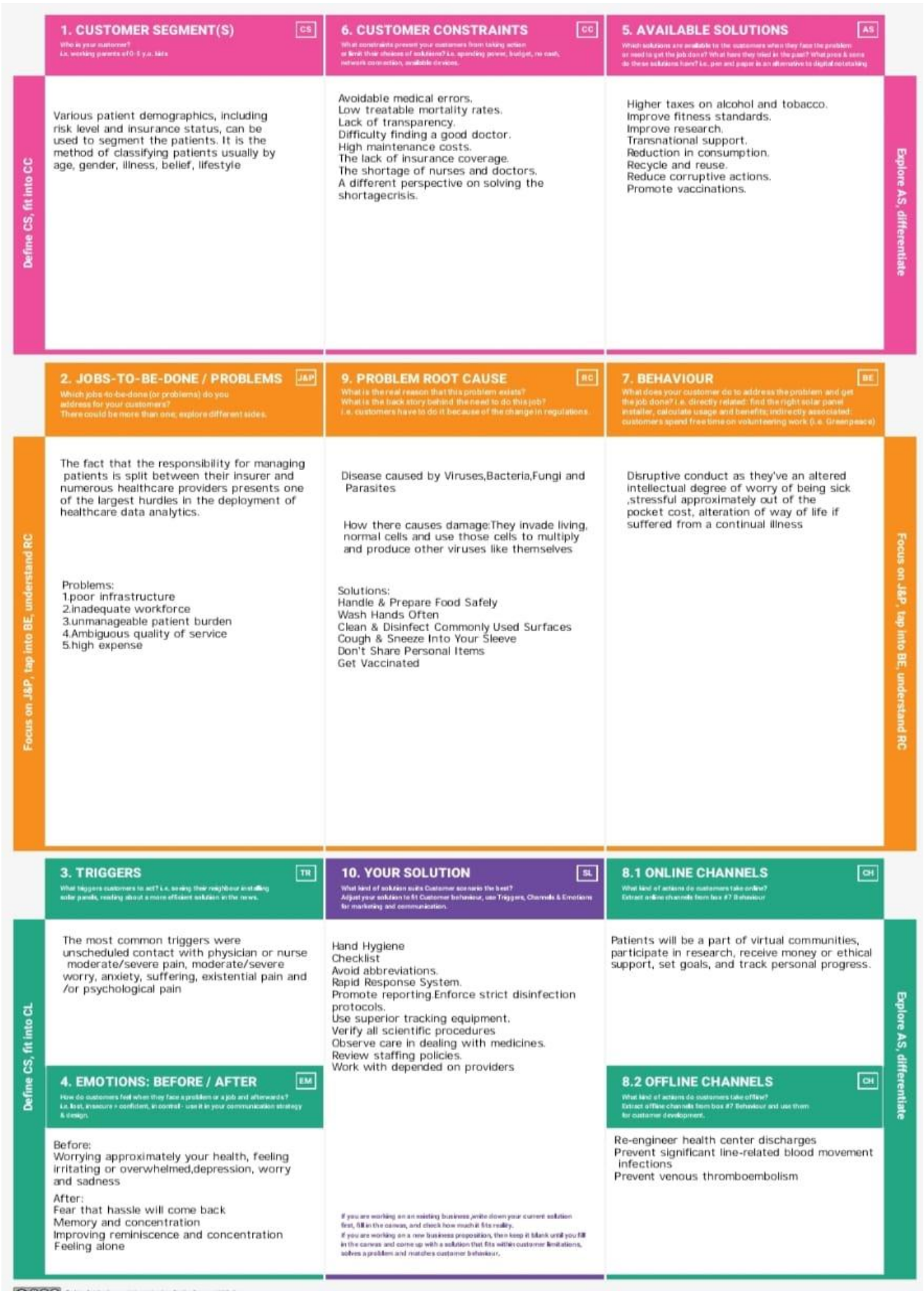


## 3.3 PROPOSED SOLUTION

| S.No. | Parameter | Description |
|---|---|---|
| 1. | Problem Statement (Problem to be solved) | EHR data matched patient-reported data in 23.5 percent of records in a study at an ophthalmology practise. Patients' EHR data did not agree in any way when they reported having three or more eye health complaints. |
| 2. | Idea / Solution description | Predictive analytics can create patient journey dashboards and disease trajectories that can lead to effective, and result-driven healthcare. It improves treatment delivery, cuts costs, improves efficiencies, and so on. |
| 3. | Novelty / Uniqueness | Healthcare data frequently resides in several locations. from various departments, such as raciology or pharmacy, to various source systems, such as EMRs or HR so²tware. The organisation as a whole contributes to the data. This data becomes accessible and usable when it is combined into a single, central system, such as an enterprise data warehouse (EDW). |
| 4. | Social Impact / Customer Satisfaction | Enhanced diagnosis<br>Improved medical treatment<br>Improved health results<br>Improved relationships with patients<br>More positive health indicators |
| 5. | Business Model (Revenue Model) | The two factors that have the biggest negative effects on hospital income are claim denials and patient incapacity to pay their part. 90% more uncollectible claim denials were written off by hospitals and healthcare systems in 2017 compared to the preceding six years. |
| 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 PROBLEM SOLUTION FIT



## 4.REQUIREMENT ANALYSIS

### 4.1 FUNCTIONAL REQUIREMENT

| FR No. | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|---|---|---|
| FR-1 | User Registration | Registration through Form |
| FR-2 | User Confirmation | Confirmation via OTP |
| FR-3 | Database | Every patient has some necessary data like phone number, their first and last name, personal health number, postal code, country, address, city, 'patient's ID number, etc |
| FR-4 | Report Generation | The Hospital Management System generates a report on every patient regarding various information like patients name, Phone number, bed number, the doctor's name whom its assigns, ward name, and more.<br>The Hospital Management system also helps in generating reports on the availability of the bed regarding information like bed numbers unoccupied or occupied, ward name, and more. |
| | Check Out | The staff in the administration section of the ward can delete the patient ID from the system when the patient checkout from the hospital.<br>The Staff in the administration section of the ward can put the bed empty in the list of beds available. |
| | Adding Patients | The Hospital Management enables the staff at the front desk to include new patients in the system. |

### 4.2 NON-FUNCTIONAL REQUIREMENTS

| FR No. | Non-Functional Requirement | Description |
|---|---|---|
| NFR-1 | **Usability** | The effectiveness, efficiency and satisfaction with which specific users can achieve a specific set of tasks in a particular environment. |
| NFR-2 | **Security** | This process of protecting data from unauthorized access and data corruption throughout its lifecycle |
| NFR-3 | **Reliability** | A highly reliable system has a lower risk of errors and process failures that can cause patients harm |
| NFR-4 | **Performance** | performance measurements include:<br>• Quality and efficiency of patient care<br>• Cost of healthcare services<br>• Disparities in performance<br>• Care outcomes |
| NFR-5 | **Availability** | inpatient, outpatient, pharmacy, and enrollment |
| NFR-6 | **Scalability** | The ability of a health intervention shown to be efficacious on a small scale and/or under controlled conditions to be expanded under real world conditions to reach a greater proportion of the eligible population, while retaining effectiveness. |

## 5. PROJECT DESIGN

### 5.1 DATA FLOW DIAGRAMS

### 5.2 SOLUTION AND TECHNICAL ARCHITECTURE

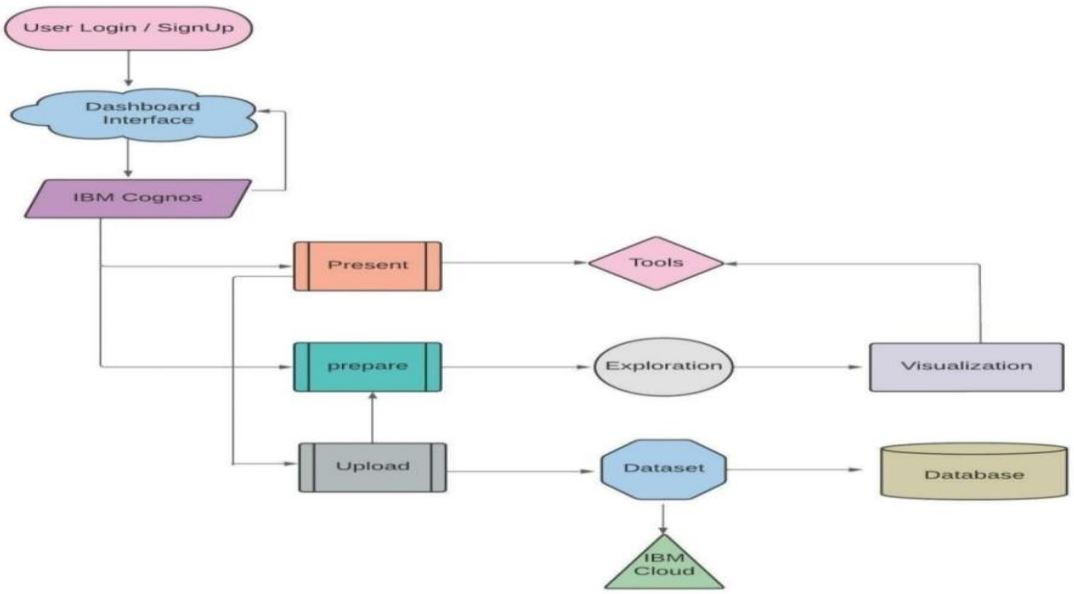


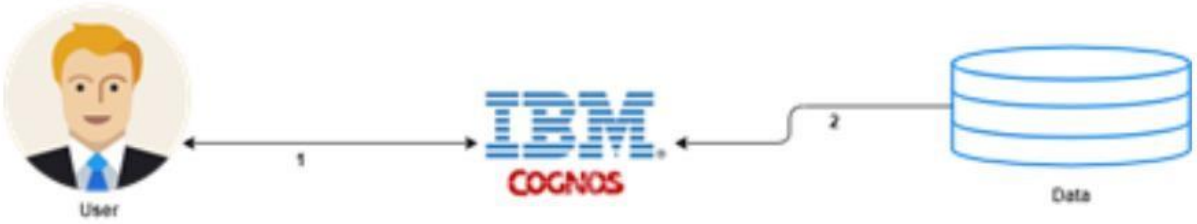

**TABLE 1:COMPONENTS & TECHNOLOGIES**

| 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. | Application Logic-3 | Logic for a process in the application | IBM Watson Assistant |
| 5. | Database | Data Type, Configurations etc. | MySQL, NoSQL, etc. |
| 6. | Cloud Database | Database Service on Cloud | IBM DB2, IBM Cloudant etc. |
| 7. | File Storage | File storage requirements | IBM Block Storage or Other Storage Service or Local Filesystem |
| 8. | External API-1 | Purpose of External API used in the application | IBM Weather API, etc. |
| 9. | External API-2 | Purpose of External API used in the application | Aadhar API, etc. |
| 10. | Infrastructure (Server / Cloud) | Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration : | Local, Cloud Foundry |

**Table-2: Application Characteristics:**

| S.No | Characteristics | Description | Technology |
|---|---|---|---|
| 1. | Open-Source Frameworks | List the open-source frameworks used | Technology of Opensource framework |
| 2. | Security Implementations | List all the security / access controls implemented, use of firewalls etc. | e.g. SHA-256, Encryptions, IAM Controls, OWASP etc. |
| 3. | Scalable Architecture | Justify the scalability of architecture (3 – tier, Micro-services) | Technology used |
| 4. | Availability | Justify the availability of application (e.g. use of load balancers, distributed servers etc.) | Technology used |
| 5. | Performance | Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc. | Technology used |

SOLUTION ARCHITECTURE:




5.3 USER STORIES

| 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 patients admitted in the hospital and a dataset should beprepared. | 20 | Medium | RANJANI M, VASUKI M |
| Sprint-1 | Data Exploration | USN- 2 | As a user, I need nicely visualized dashboard ofnumber of beds occupied and number of free beds in hospital. | 10 | High | Sandhana k, sandhiya G |
| 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. | 10 | Medium | Sandhiya G, Ranjani M |
| Sprint -2 | Dashboard | USN - 4 | As a user, I want the interactive dashboard toanalyse the data. Have the data in terms of Graph. | 20 | High | Ranjani M, vasuki M,Sandhiya G |
| 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. | 10 | Medium | vasuki M,Sandhiya G |
| Sprint- 3 | Story Creation | USN-6 | As a user, I need the story animation of thedata set with insights | 20 | High | Sandhana k, sandhiya G |
| Sprint-4 | Predict LOS | USN-7 | As a user, I want the flawless system to predictthe length of stay of the patients | 10 | High | Ranjani M, Vasuki M |
| 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. | 20 | High | Ranjani M, vasuki M,Sandhiya G, Sandhana K |

## 6. PROJECT PLANNING AND SCHEDULING

### 6.1 SPRINT PLANNING AND 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 |
| | | | | | | |

**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 iteration unit (story points per day)

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

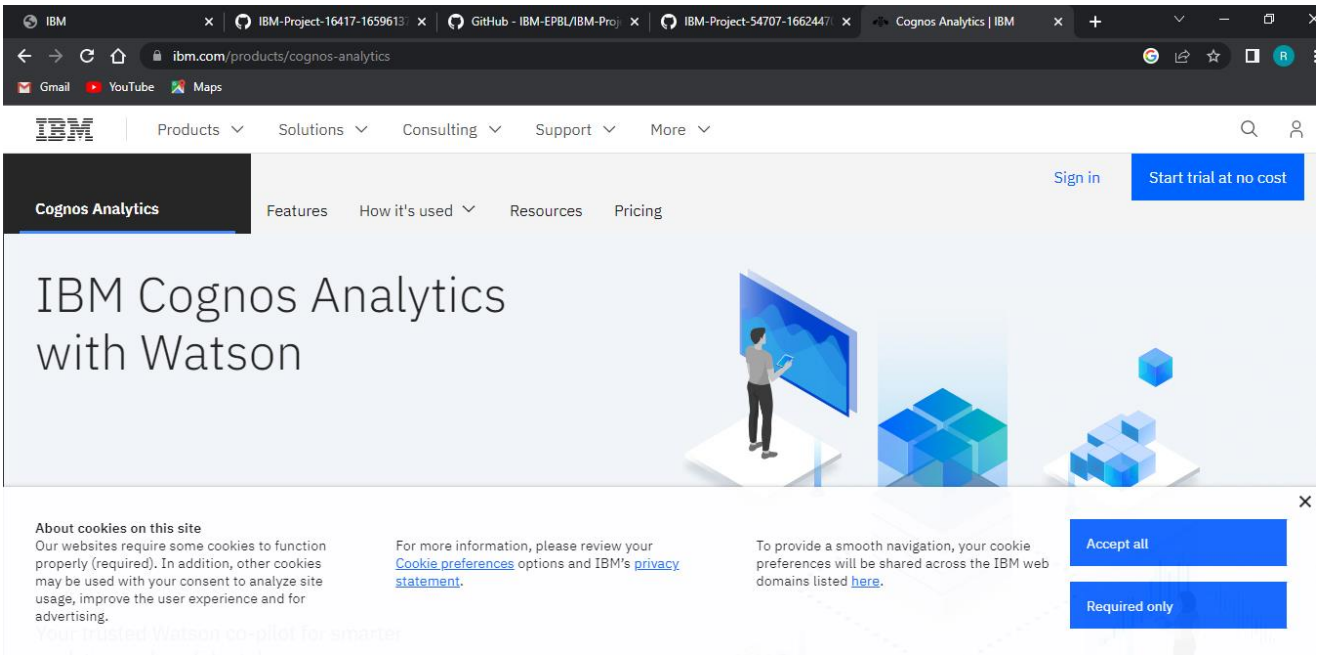
### 6.2 SPRINT DELIVERY SCHEDULE

| 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 patients admitted in the hospital and a dataset should beprepared. | 20 | Medium | RANJANI M, VASUKI M |
| Sprint-1 | Data Exploration | USN- 2 | As a user, I need nicely visualized dashboard ofnumber of beds occupied and number of free beds in hospital. | 10 | High | Sandhana k, sandhiya G |
| 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. | 10 | Medium | Sandhiya G, Ranjani M |
| Sprint -2 | Dashboard | USN - 4 | As a user, I want the interactive dashboard toanalyse the data. Have the data in terms of Graph. | 20 | High | Ranjani M, vasuki M,Sandhiya G |
| 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. | 10 | Medium | vasuki M,Sandhiya G |
| Sprint- 3 | Story Creation | USN-6 | As a user, I need the story animation of thedata set with insights | 20 | High | Sandhana k, sandhiya G |
| Sprint-4 | Predict LOS | USN-7 | As a user, I want the flawless system to predictthe length of stay of the patients | 10 | High | Ranjani M, Vasuki M |
| 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. | 20 | High | Ranjani M, vasuki M,Sandhiya G, Sandhana K |

## 7. SOLUTIONING

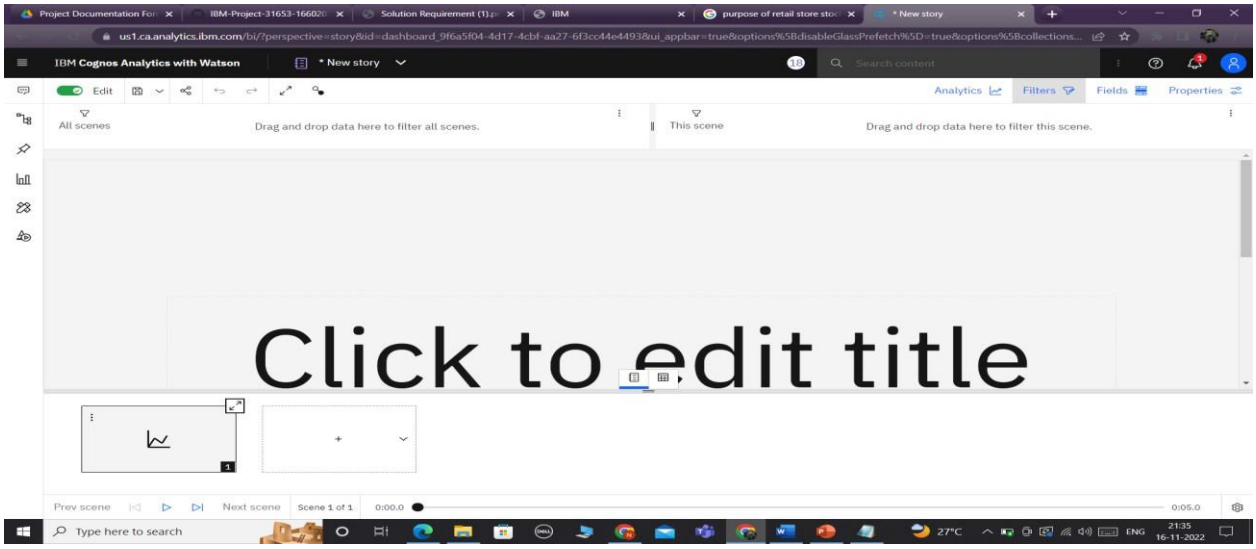
7.1 FEATURE 1

DASHBOARD AND REPORT:




7.2 FEATURE 2

STORY OF THE DATASET:

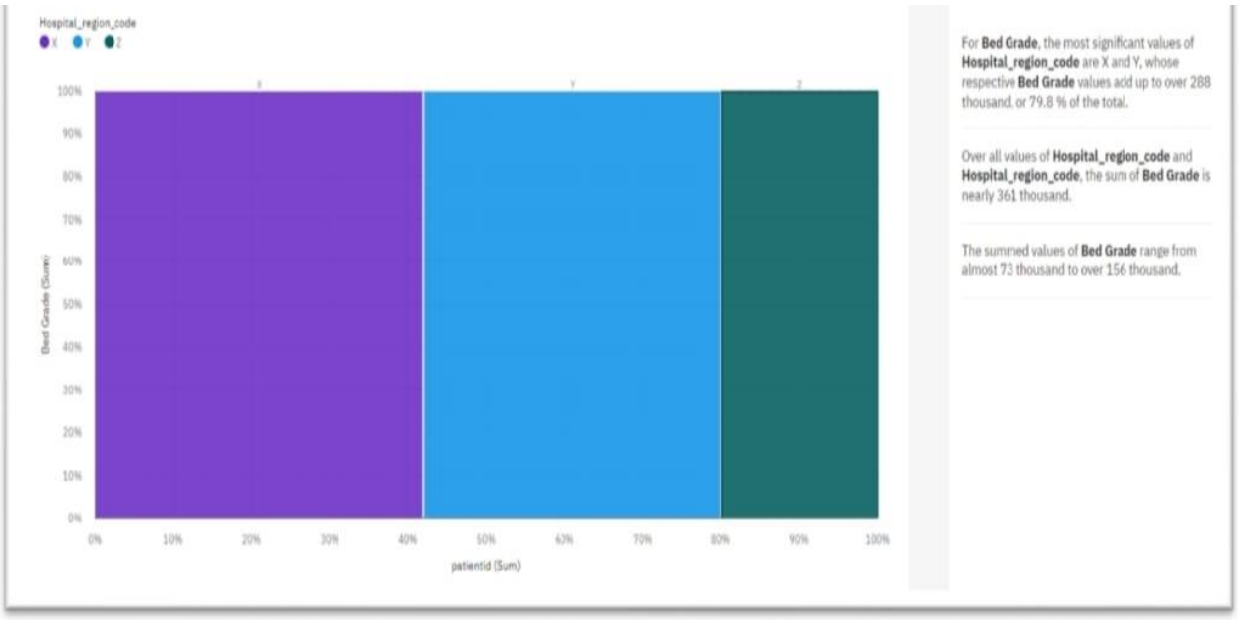



## 8.TESTING

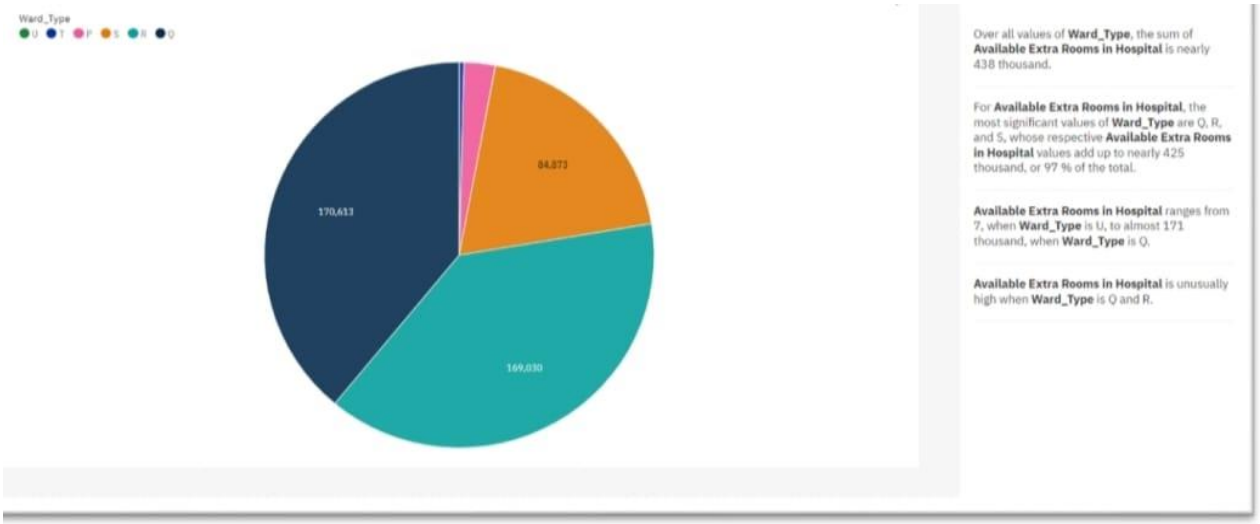
8.1 TEST CASES

### 1.EXPLORATION OF DATA

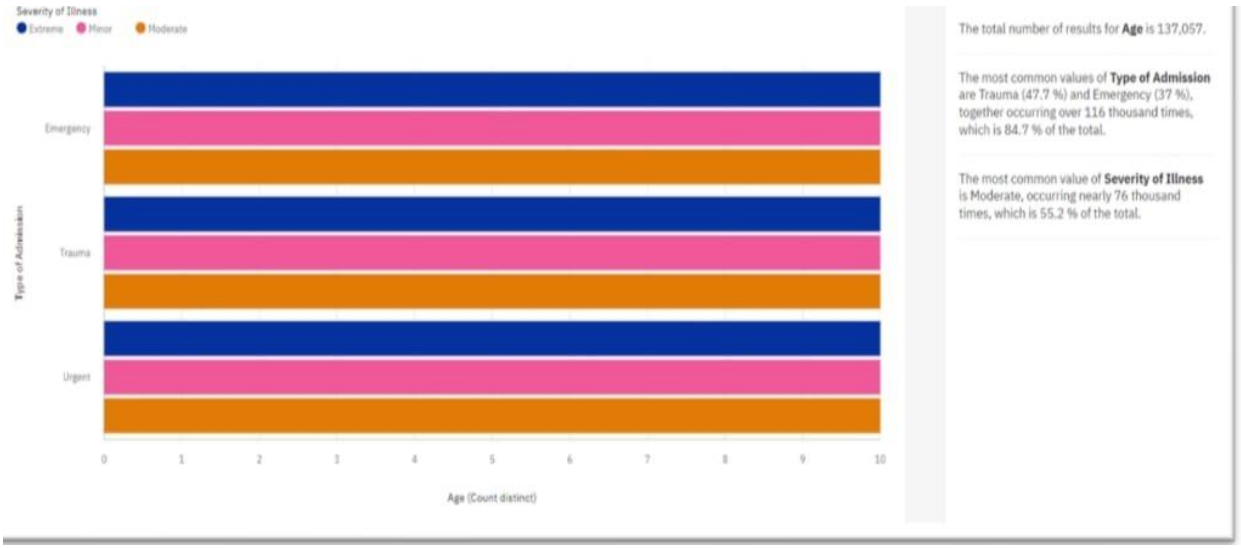

## 2.Bed Grade and Patient for Hospital_region_code coloured by Hospital_region_code



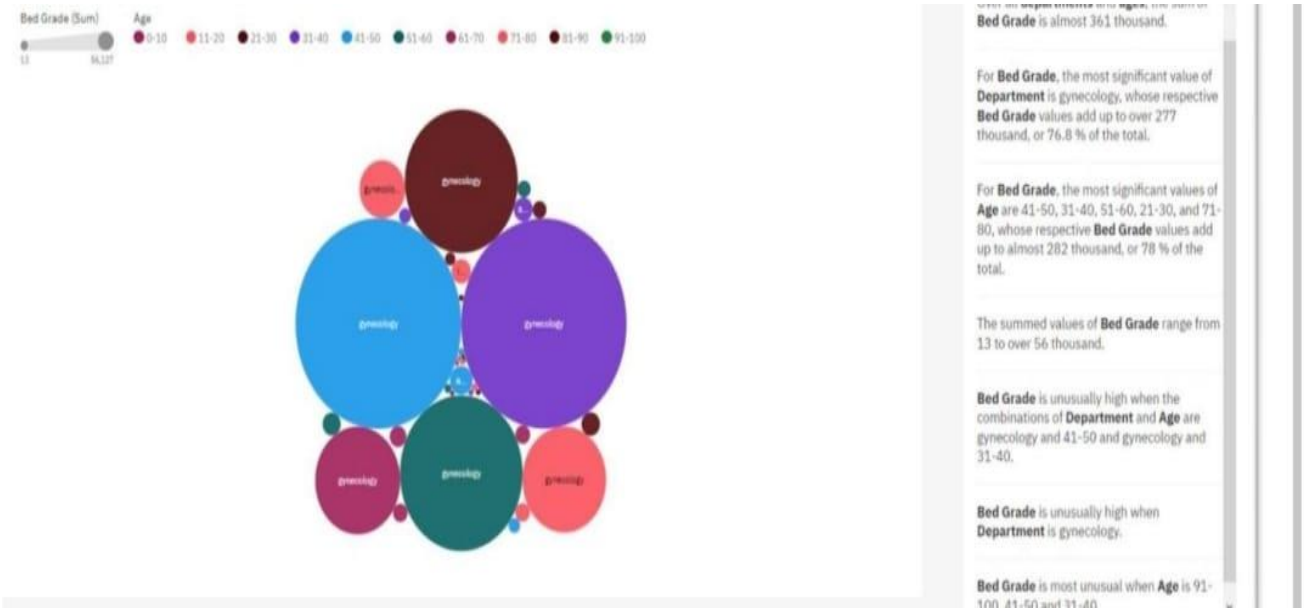

## 3.Available Extra Rooms in Hospital by Ward_Type



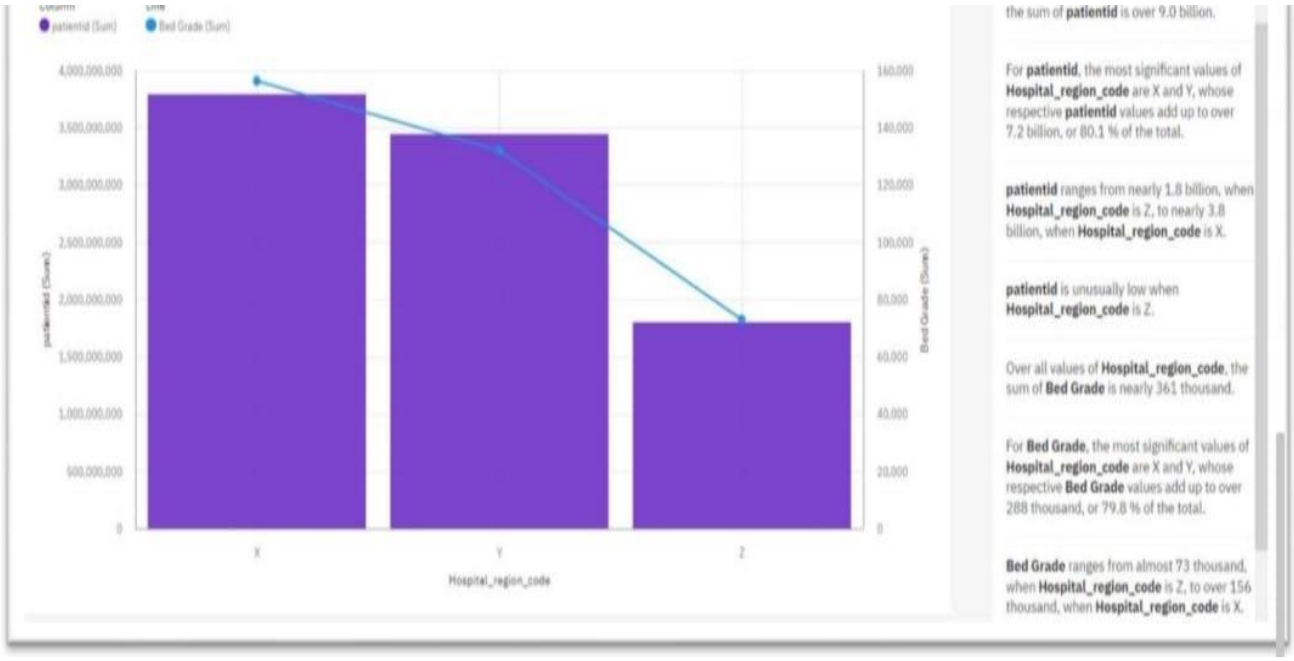

## 4.Department coloured by Age sized by Bed Grade




## 5.Age by Type Admission coloured by Severity of Illness

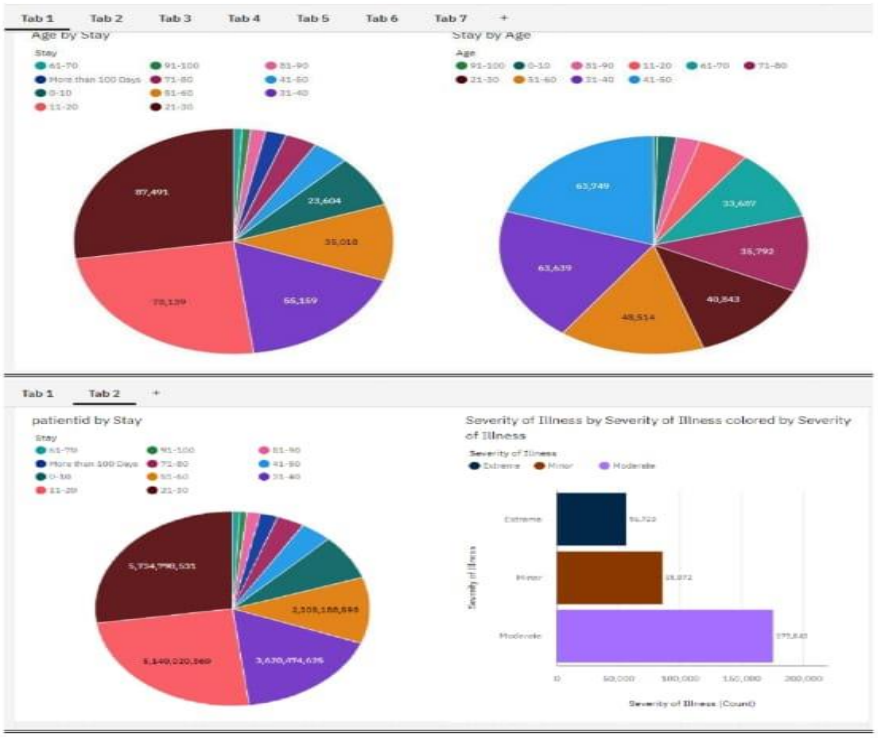


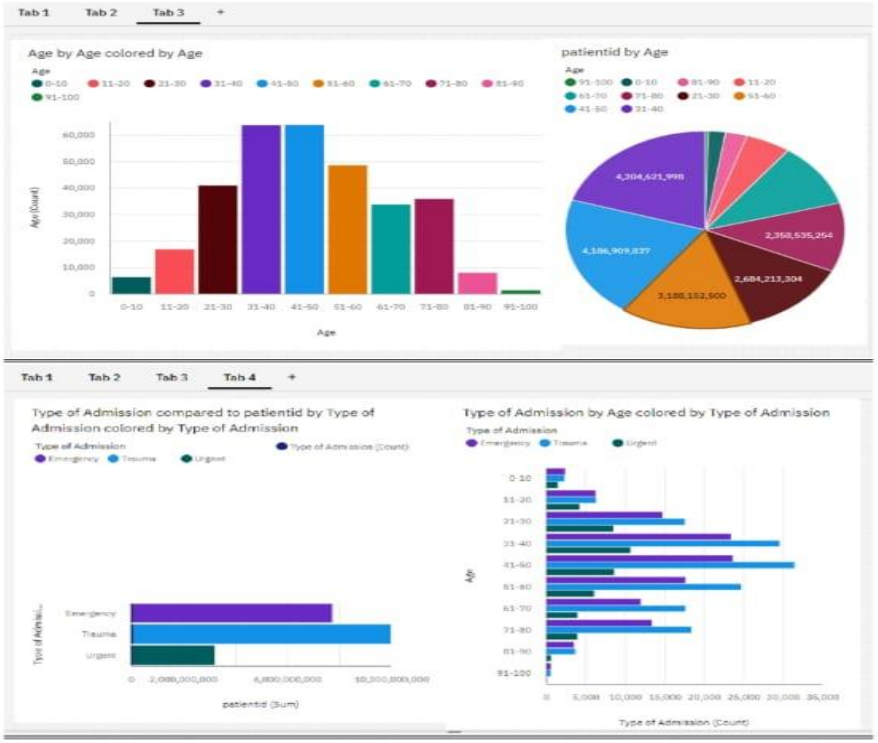

## 6.Bed Grade and Patient by Hospital_region_code



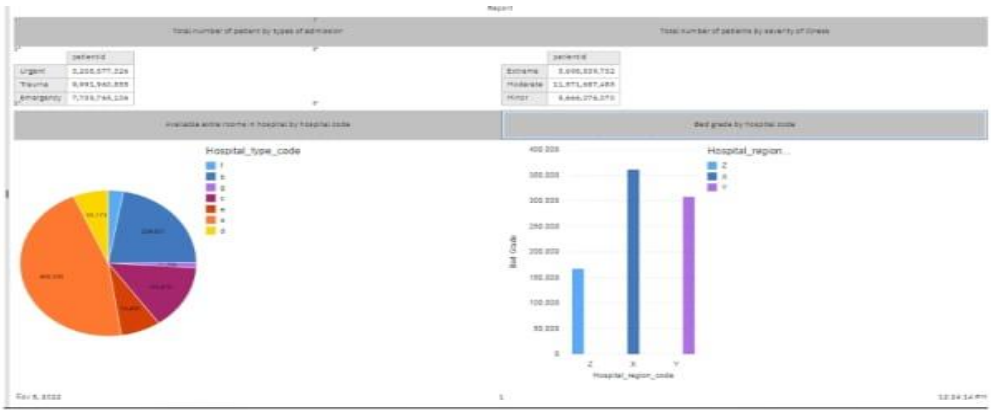

## 9.RESULTS

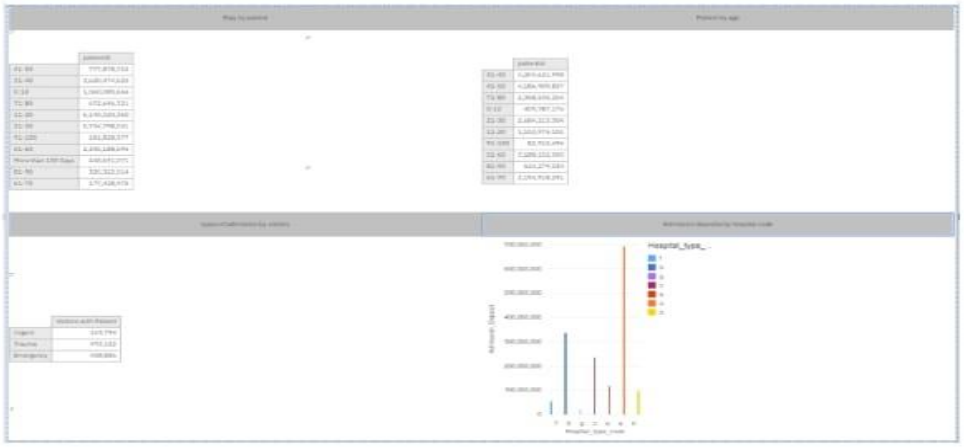
9.1 PERFORMANCE METRICS DASHBOARD:

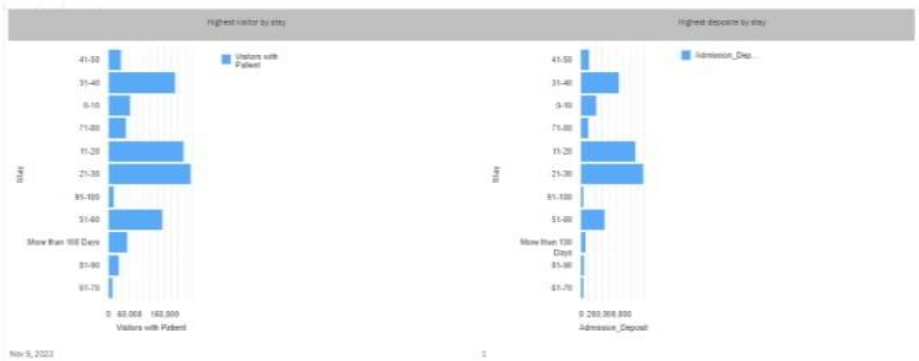

Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
Tab 6
Tab 7
Age by Stay
Stay by Age
patientid by Stay
Severity of Illness by Severity of Illness colored by Severity of Illness


Tab 1
Tab 2
Tab 3
Tab 4
Age by Age colored by Age
patientid by Age
Type of Admission compared to patientid by Type of Admission colored by Type of Admission
Type of Admission by Age colored by Type of Admission

REPORT:


Report
Total number of patient by types of admission
patientid
Urgent 3,205,577,226
Trauma 6,991,940,885
Emergency 7,739,744,104
Total number of patients by severity of illness
patientid
Extreme 3,695,824,752
Moderate 11,871,487,463
Minor 6,669,276,370
Available ambulance in hospital by hospital code
Hospital_type_code
Bed grade by hospital code
Hospital_region...
Bed Grade
Hospital_region_code
Nov 5, 2022
1
12:04:14 PM


Stay by severity
Patient by age
Visitors with Patient by severity
Admission deposited by hospital code
Hospital_type_...
Admission Deposit


Highest visitor by stay
Visitors with Patient
Stay
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
More than 100 Days
0 60,000 120,000
Visitors with Patient
Highest deposite by stay
Admission_Dep...
0 200,000,000
Admission_Deposit
Nov 5, 2022
1

STORY:


**10. ADVANTAGES:**

Cost Reduction  Improves Data Security  Less paperwork  Better Collaboration and Communication  Error-free administration  Improved Clinical decision -making

**DISADVANTAGES:**

- Cybersecurity Risks.
- Patient Isolation .
- Frustation with poor Implementation.

**11.CONCLUSION**

In hospital health care management , we need a healthy lifestyle to build up a healthy immune system and to avoid diseases.

## 12.FUTURE SCOPE

**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.**

## 13.APPENDIX

Github project link - https://github.com/IBM-EPBL/IBM-Project-31813-1660205338