Analytics for Hospital’s HealthCareData NALAIYA THIRAN PROJECT REPORT 2022

Submitted by Team ID:PNT2022TMID14243

D.SUPRATHIKA 111619104025
D.V.NAVYA LAKSHMI 111619104027
K.JYOTSNA 111619104052
K.G.JAISREE 111619104043

Contents

TEAM ID: PNT2022TMID14243

Analytics for Hospital's HealthCareData

1. Introduction

1.1 Project Overview

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas tofocus:Healthcare Management. While healthcare management has various use cases for using datascience, patient length of stay is one critical parameter to observe and predict if one wants to improvetheefficiency of the healthcare management in a hospital. This parameter helps hospitals toidentifypatients 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 andlower thechance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics suchas roomand bed allocation planning. Suppose you have been hired as Data Scientist of Health Man–anotfor profit organization dedicated to manage the functioning of Hospitals in a professional andoptimal manner.

1.2 Purpose

The goal is to accurately predict the Length of Stay for each patient on case by casebasisso that the Hospitals can use this information for optimal resource allocation andbetterfunctioning. The length of stay is divided into 11 different classes ranging from0-10 days tomorethan 100 days.

2. Literature Survey

2.1 Existing Problem

Data mining offers novel information regarding health care helpful for makingadministrative aswell as prediction disease, selection of treatment, health isurance policy. Thenovel corona virus pandemic outbreak is seriously threatening human health. Securityoptimization implementaion and testing on real world patients Hospitalization cost andtheinsured population all show a trend of increasing year by year. The users to helptoseeunderstand the valuable information provided by data care visual analytics huge amount ofstructured and unstructured and semi structured data have been generated by various institutions around the world.

2.2 References

1. Proposed application of big data analyticsin healthcare at Maharaja YeshwantraoHospital

This paper gives an insight of how we can store healthcare data digitally like patient'srecords as an Electronic Health Record (EHR) and how we can generate useful informationfromthese records by using analytics techniques and tools which will help in saving time andmoneyof patients as well as the doctors. This paper is fully focused towards the Maharaja YeshwantraoHospital (M.Y.) located in Indore, Madhya Pradesh, India. M.Y hospital is the central India'slargest government hospital. It generates large amount of heterogeneous data fromdifferent sources like patients health records, laboratory test result, electronic medical equipment, healthinsurance data, social media, drug research, genome research, clinical outcome, transactionandfrom Mahatma Gandhi Memorial medical college which is under MY hospital. To managethisdata, data analytics may be used to make it useful for retrieval. Hence the concept of "bigdata"can be applied. Big data is characterized as extremely large data sets that can be analysedcomputationally to find patterns, trends, and associations, visualization, querying, informationprivacy and predictive analytics on large wide spread collection of data. Big data analytics canbedoneusing Hadoop which plays an effective role in performing meaningful real-time analysisonthe large volume of this data to predict the emergency situations before it happens. This paperalso discusses about the EHR and the big data usage and its analytics at M.Y. hospital

2. A look at challenges and opportunities of Big Data analytics in healthcare

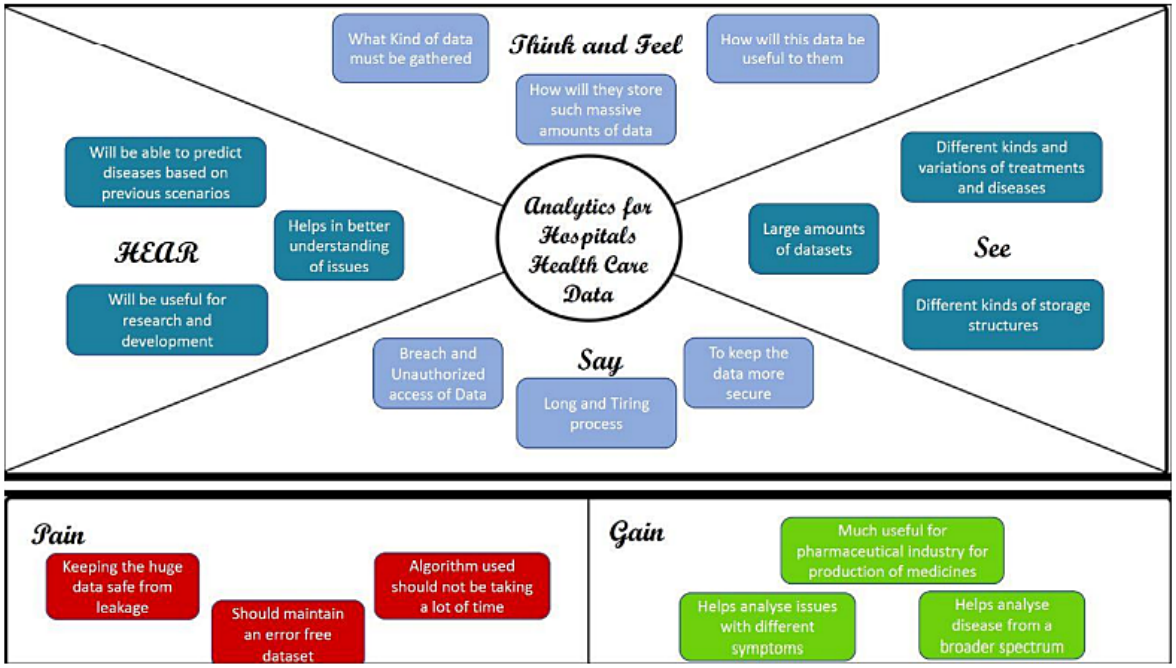
Big Data analytics can revolutionize the healthcare industry. It can improve operational efficiencies, help predict and plan responses to disease epidemics, improve the qualityofmonitoring of clinical trials, and optimize healthcare spending at all levels frompatients tohospital systems to governments. This paper provides an overview of Big Data, applicabilityof it inhealthcare, some of the work in progress and a future outlook on how Big Data analyticscanimprove overall quality in healthcare systems.

2.3 Problem Statement Definition

The analysis using hospital healthcare data allows us to identify between high riskcasesand low risk cases. Prior hospital data provides us with information as to howlong the patient wasadmitted to the hospital and the medication administered to them. Analysing this data canprovideus future insight as to how long the patient will stay and helps us in bed allocation andmedicineand personnel management.

3. Ideation and Proposed Solution

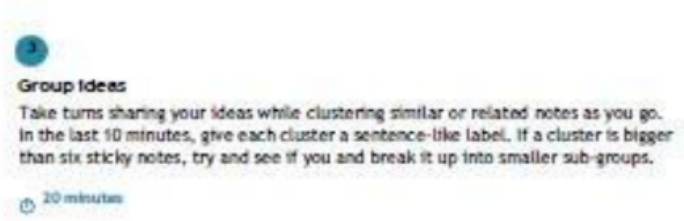
3.1 Empathy Map



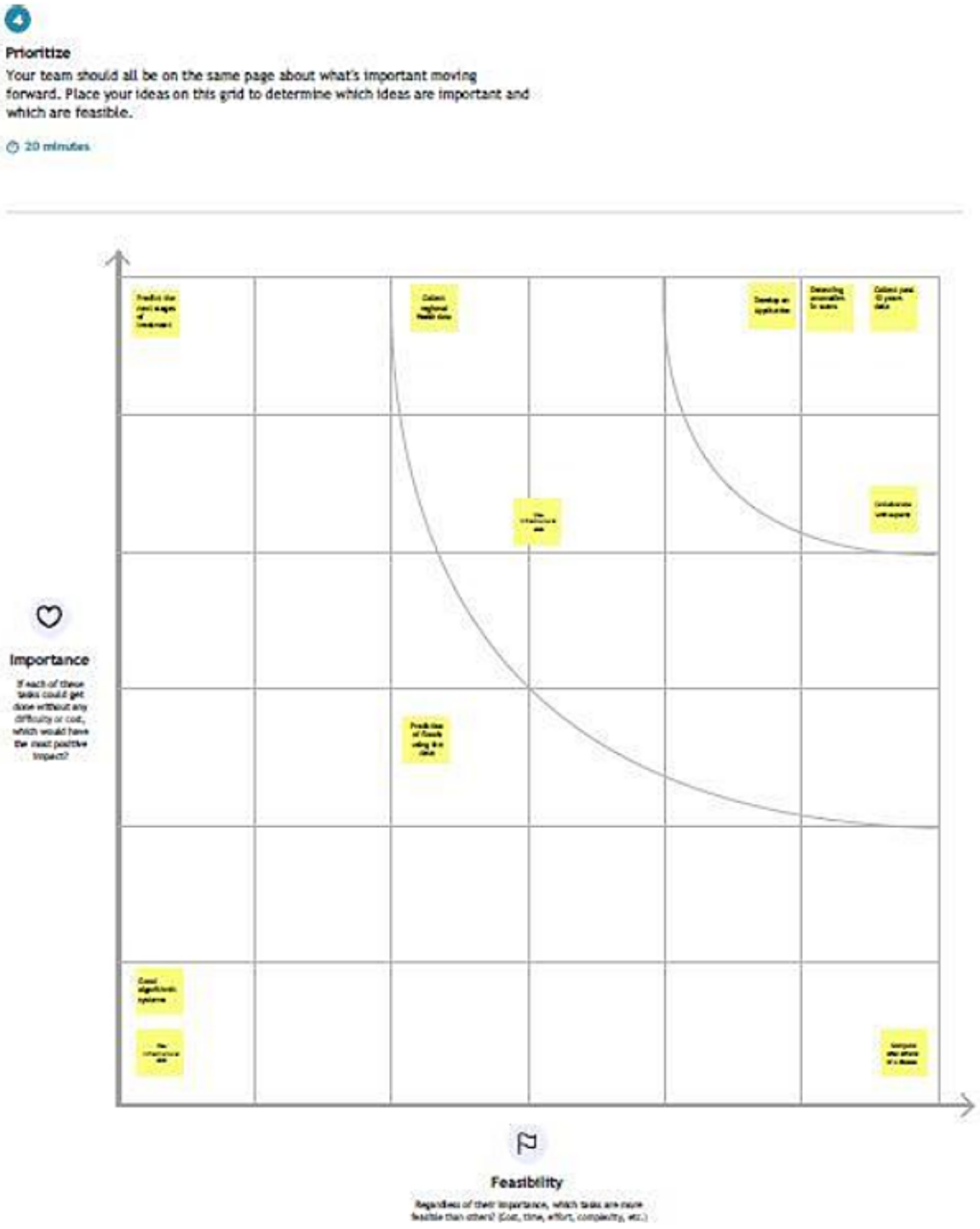
3.2 Ideation and Brainstorming

Step-1: Team Gathering, Collaboration and Select the Problem Statement

Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization

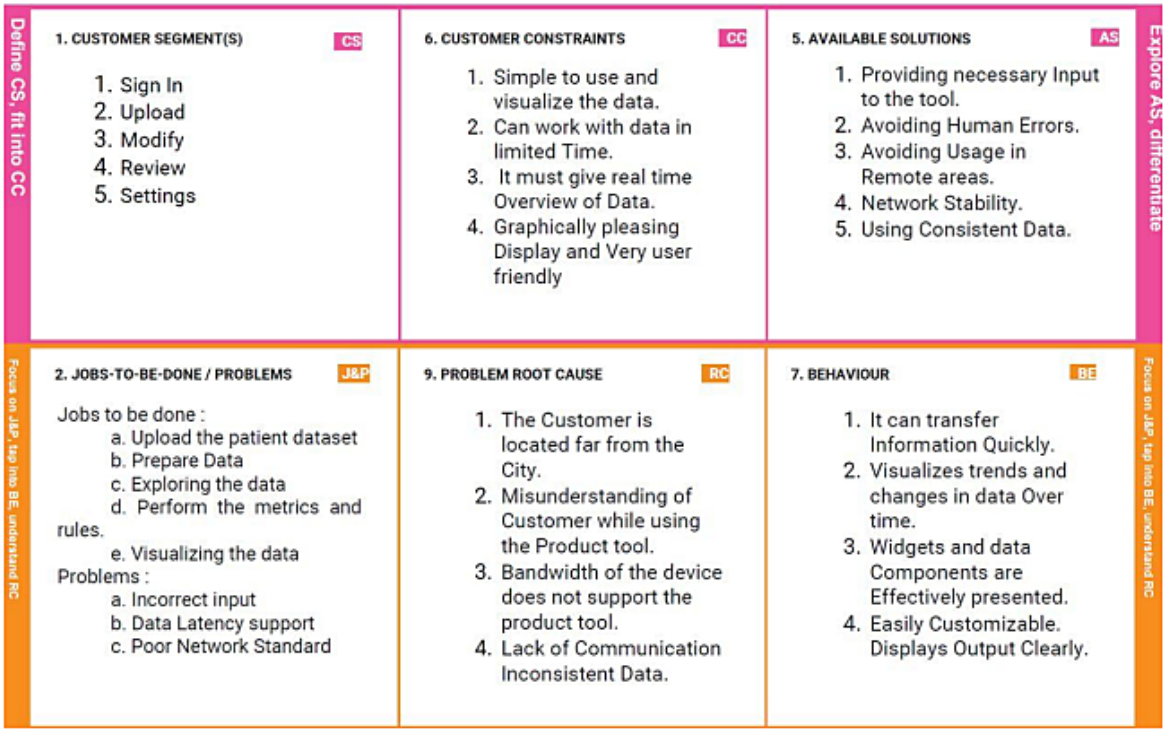


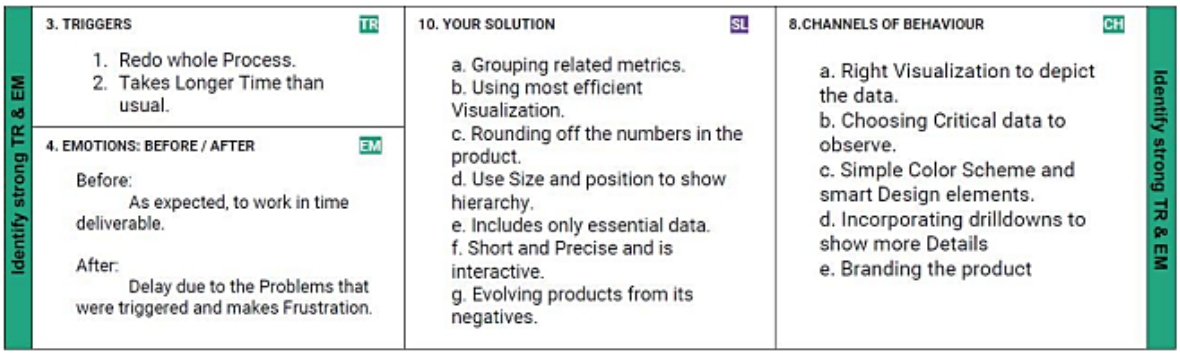
3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	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.
2.	Idea / Solution description	Use various factors which determine the LOS like: • Similarity in Patient health conditions • Similarities in previous history of illness • Predictive analysis bases on seasons • Predictive analysis based on age groups • Doctor's Advice regarding LOS • Add bed days for each discharged patient and divide the sum by the number of discharged patients • General predictions based on Oxygen Delivery Index, RBC count, CreatinineLevels etc.
3.	Novelty / Uniqueness	The uniqueness of this proposed system is that data is not gathered by a single source but from multiple departments from the hospitals thereby reducing redundant data The proposed dashboard will make it easy for anyone to analyse and accommodate based on priority and requirements
4.	Social Impact / Customer Satisfaction	The various social impacts / Customer Satisfaction are: • Easy to access by the public and prepare accordingly • People need not worry about their LOS and can simply rely on this data

5.	Business Model (Revenue Model)	This system when sold as a business to hospitals as clients will generate huge amounts of revenue as this is a system which will help in resource allocation, budget allocations, pharmaceuticals etc. Previous medical record and records from this dashboard together will make this system reliable at all instances and help in the overall RND of the hospitals and the quality of treatment provided
6.	Scalability of the Solution	Easy to be implemented as this does not involve any special hardware. Data inputs are from the hospitals directly hence no outsourcing is required for datasets In case of sudden surges and sudden mutations of any disease the solution will require a minimum of 14 days to understand the system properly

3.3 Problem Solution Fit





4.RequirementAnalysis

4.1 Functional Requirements:

Following are the functional requirements of the proposed solution

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Operability	Share patient data and make it interoperable among the management
FR-4	Accuracy	The dashboard will be able to predict length of stay based on multiple combinations based on input sources with a n accuracy of upto 85%
FR-5	Compliance	The product is to be used within the hospital so any form of data need not be hidden
FR-6	Productivity	The dashboard is believed to improve the predictions of Length of Stay and thereby creating a scenario of providing better solution

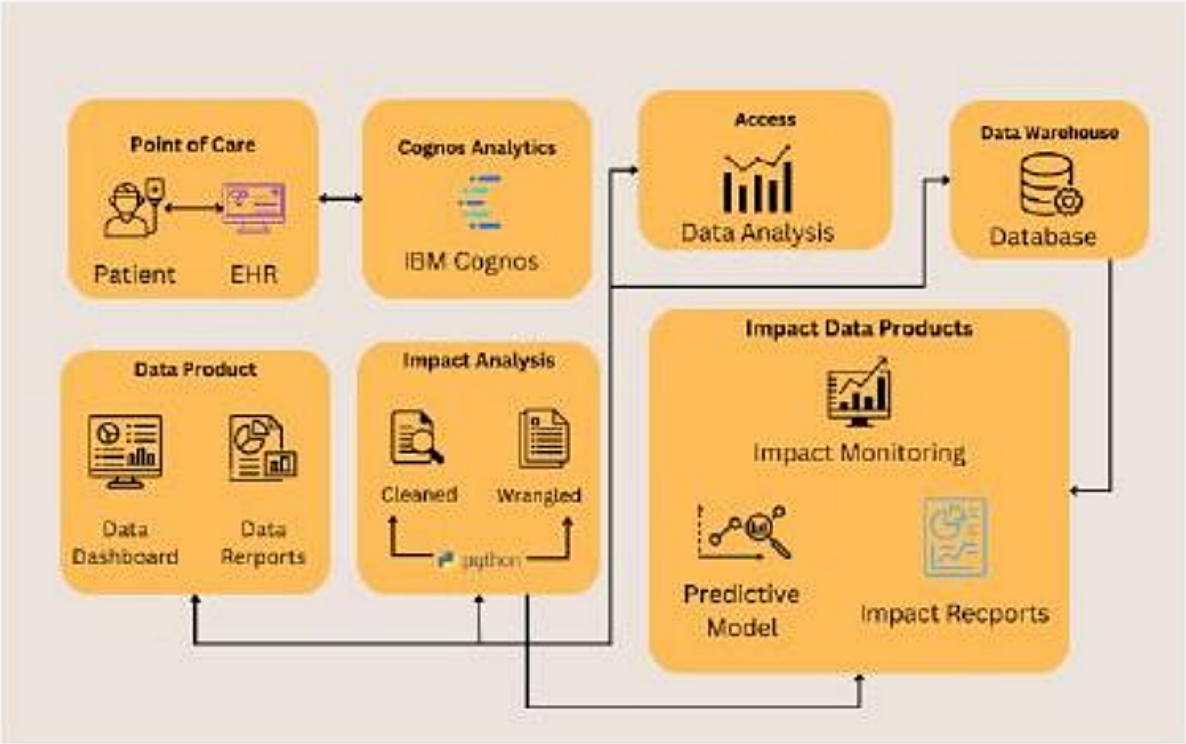
4.2 Non-functional Requirements:

Following are the non-functional requirements of the proposed solution

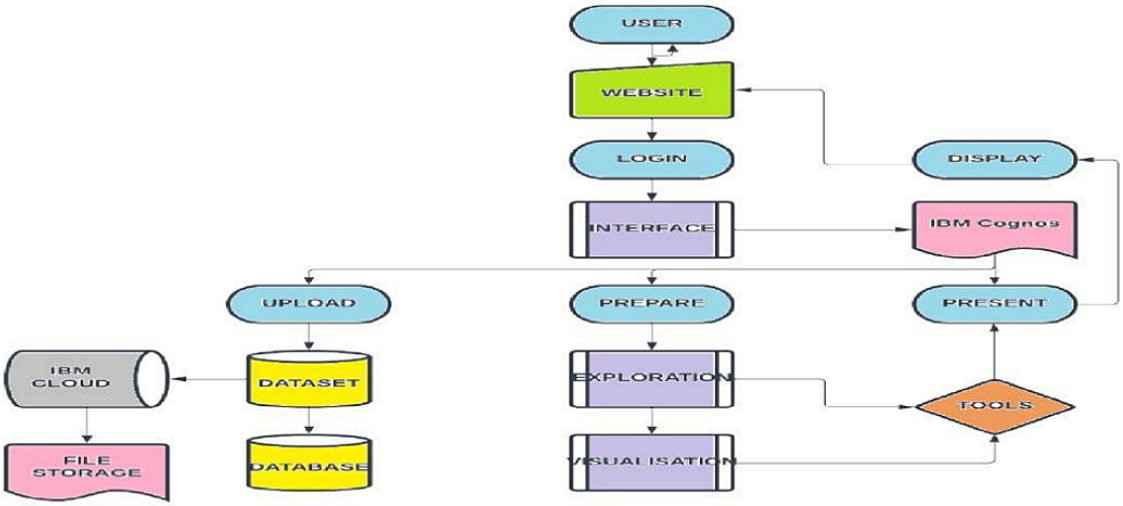
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	This Dashboards are designed to offer a comprehensive overview of patient's LOS, and do so through the use of data visualization tools like charts and graphs.
NFR-2	**Security**	General industry level security shall be provided
NFR-3	**Reliability**	This dashboard will be consistent and reliable to the users and helps the user to use in effective ,efficient and reliable manner.
NFR-4	**Performance**	The dashboard reduces the time needed for analysing data and has an automated system for that which improves the performance
NFR-5	**Availability**	The dashboard can available to meet user's demand in timely manner and it is also helps to provide necessary information to the user's dataset
NFR-6	**Scalability**	It is a multi tenant system which is capable of rimming on lower level systems as well.

5.Project Design

5.1 Data Flow Diagram



5.2 Solution and Technical Architecture



6. Project Planning andSchedulingScriptPlanning and Execution

6.1 Script Planning and Execution

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Retrieve Data	USN-1	As a user, I should get clearer clinical contextfor AIDS patient's unique case	10	High	John Nesa Kumar S
Sprint-1	Visualize the data	USN-2	As a user, I need nicely visualized dashboard ofnumber of beds occupied and number of free beds in hospital.	20	High	Mohamed Abdul Fazil
Sprint-2	Track of patientvisit of Hospital	USN-3	Tracking a patient Health care over years ofvisit and Screening of data they have in hospital.	10	Low	John Nesa Kumar S
Sprint-2	Dashboard	USN-4	As a user, I want the interactive dashboard toanalyse the data. Have the data in terms of Graph.	20	Medium	Mohamed Abdul Fazil
Sprint-3	Detailed EHR's ofpatient	USN-5	Provided greater details in the EHR's of individual patient with clear idea of what to do.	10	High	Abinesh R
Sprint-3	Story Creation	USN-6	As a user, I need the story animation of thedata set with insights	20	High	Jeno Allwyn
Sprint-4	Predict LOS	USN-7	As a user, I want the flawless system to predictthe length of stay of the patients	20	Medium	Abinesh
Sprint-4	Using ML algorithmfor Prediction	USN-8	As a user, I need prior knowledge of LOScan aid in logistics such as room and bedallocation planning.	20	High	Jeno Allwyn
Sprint	**Functional Requirement (Epic)**	**User Story Number**	**User Story / Task**	**Story Points**	**Priority**	**Team Members**
Sprint-1	Retrieve Data	USN-1	As a user, I should get clearer clinical contextfor AIDS patient's unique case	10	High	John Nesa Kumar S
Sprint-1	Visualize the data	USN-2	As a user, I need nicely visualized dashboard ofnumber of beds occupied and number of free beds in hospital.	20	High	Mohamed Abdul Fazil
Sprint-2	Track of patientvisit of Hospital	USN-3	Tracking a patient Health care over years ofvisit and Screening of data they have in hospital.	10	Low	John Nesa Kumar S

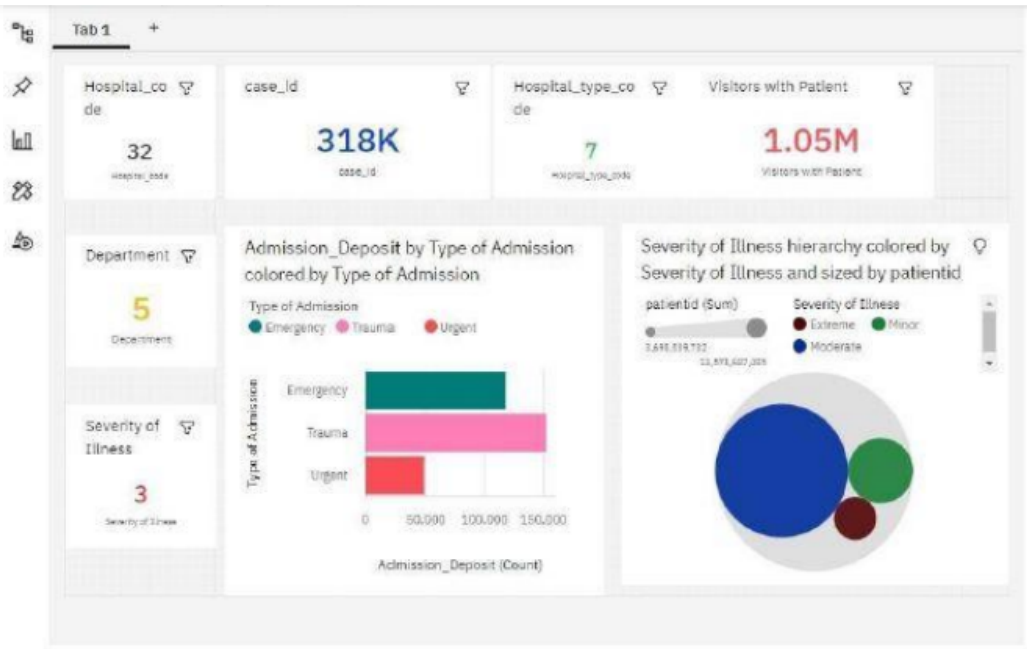
6.2 Sprint Delivery Schedule

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
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022	Sprint-1
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022	Sprint-2
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022	Sprint-3
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022	Sprint-4

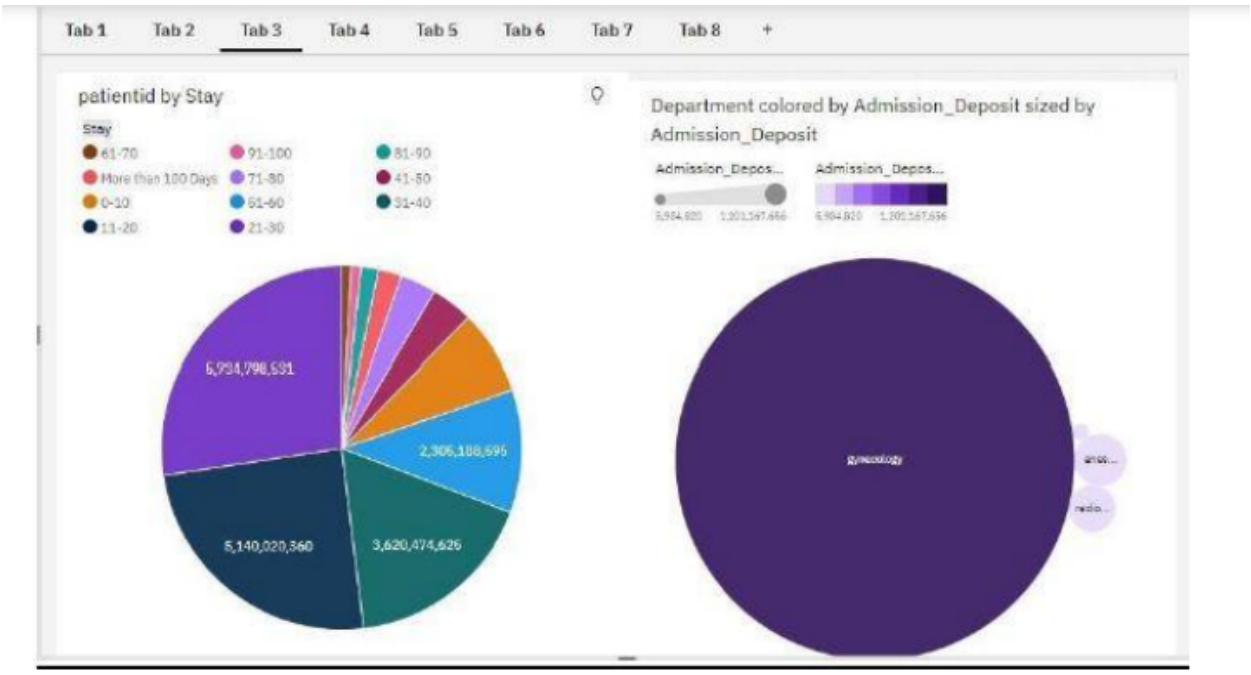
7. Coding And Solutioning

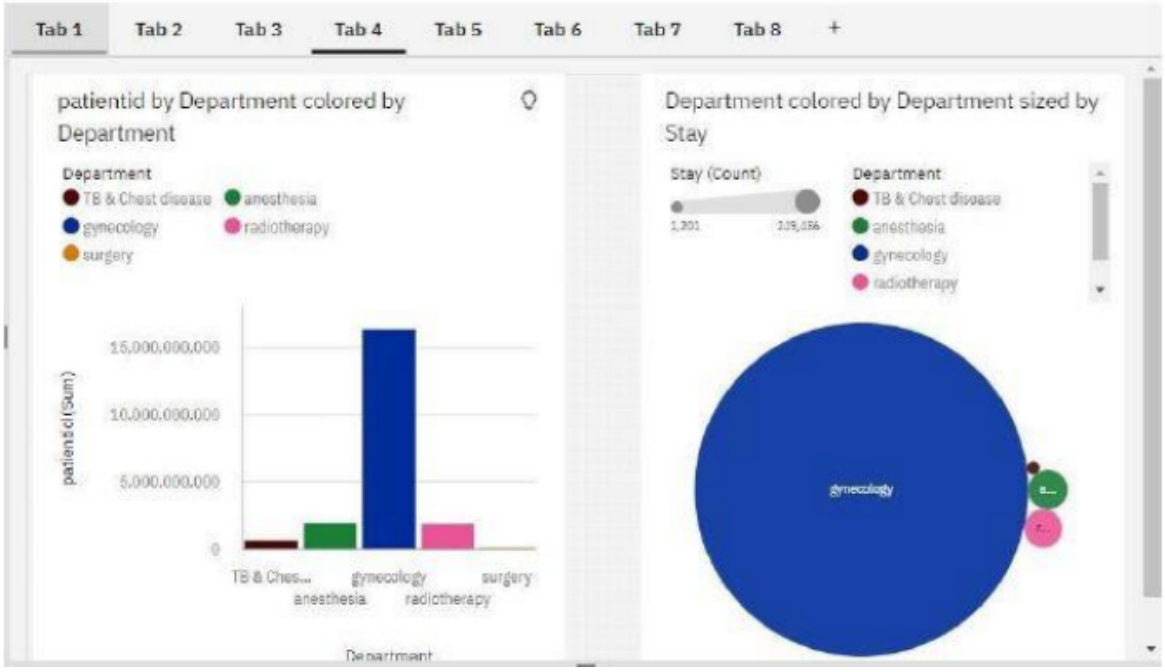
```
<!DOCTYPE html>
<html lang="en">

<head>
  <title>Data Analytics</title>
  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1">

 <linkrel="stylesheet"href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/css/bootstrap.min.css">
  <script src="https://ajax.googleapis.com/ajax/libs/jquery/3.6.0/jquery.min.js"></script>
  <script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/js/bootstrap.min.js"></script>
</head>

<body>

  <nav class="navbar navbar-inverse ">
    <div class="container-fluid">
      <div class="navbar-header">
        <a class="navbar-brand" href="#">Analytics for Hospitals' Health-Care Data</a>
      </div>
      <ul class="nav navbar-nav">
        <li><a href="index.html">Home</a></li>
        <li class="active"><a href="#">Dashboard</a></li>
        <li><a href="report.html">Report</a></li>
        <li><a href="story.html">Story</a></li>
      </ul>
    </div>
  </nav>

  <div class="container">
    <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.my_folders%2FMy%2Bfirst%
2Bdashboard&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareM
ode=embedded&amp;action=view&amp;mode=dashboard&amp;subView=model000001849013f37c_00000002
"
      width="1020" height="720" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>
  </div>

</body>

</html>
```

7.1 Dashboard


Tab 1
Hospital_code
32
case_id
318K
Hospital_type_code
7
Visitors with Patient
1.05M
Department
5
Severity of Illness
3
Admission_Deposit by Type of Admission colored by Type of Admission
Type of Admission
Emergency
Trauma
Urgent
Type of Admission
0
50,000
100,000
150,000
Admission_Deposit (Count)
Severity of Illness hierarchy colored by Severity of Illness and sized by patientid
patientid (Sum)
Severity of Illness
Extreme
Minor
Moderate


Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
Tab 6
Tab 7
Tab 8
Department by Age colored by Age
Age
0-10
11-20
21-30
31-40
41-50
51-60
11-20
21-30
60,000
50,000
40,000
30,000
20,000
10,000
0
Stay by Age
Age
91-100
0-10
81-90
11-20
61-70
71-80
21-30
51-60
31-40
41-50
63,749
33,687
35,792
63,639
40,843
48,514





8. Testing

8.1 User Acceptance Test

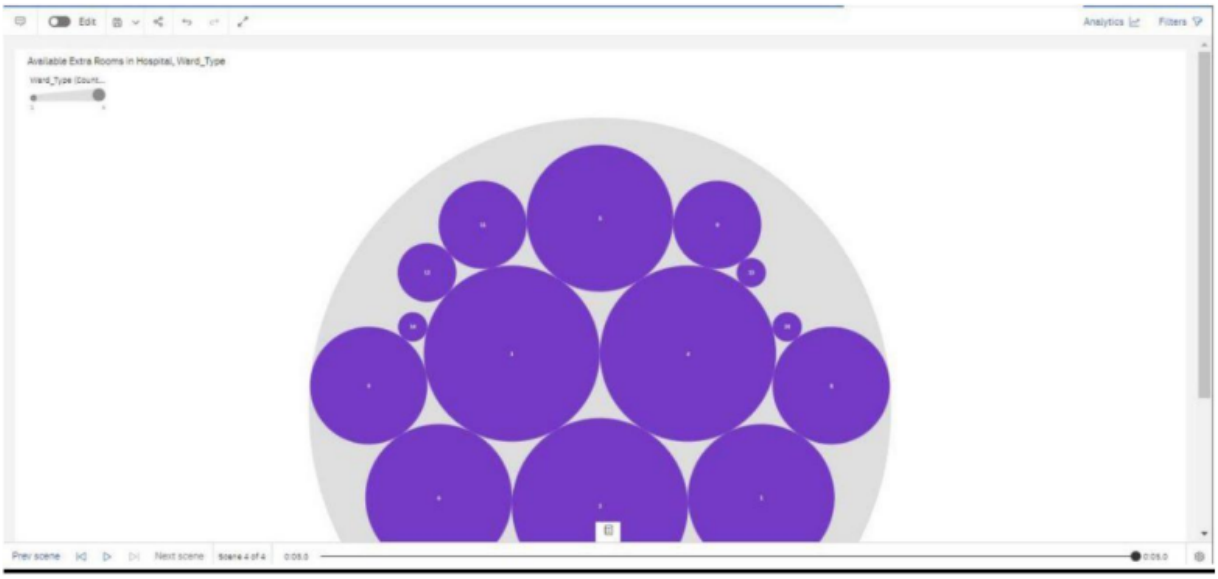
Resolution	Severity1	Severity2	Severity3	Severity4	Subtotal
By Design	5	5	3	0	13
Duplicate	0	0	0	0	0
External	7	5	1	0	13
Fixed	11	8	7	5	31
Not Reproduced	1	0	0	0	1
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	24	18	11	5	58

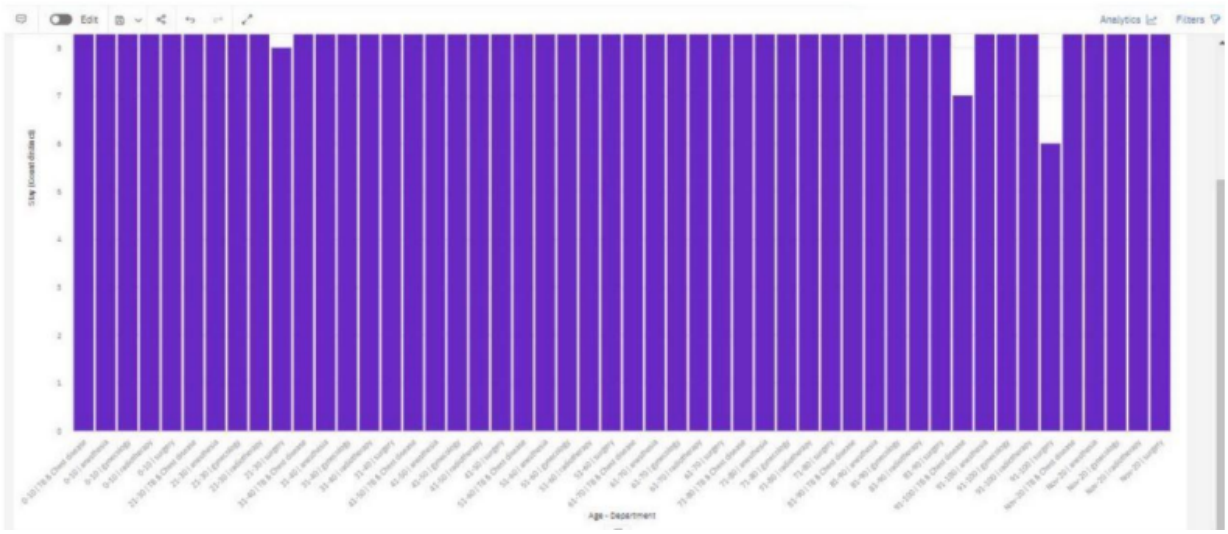
Section	Total Cases	Not Tested	Fail	Pass
Print Engine	0	0	0	0
Client Application	0	0	0	0
Security	0	0	0	0
Outsource Shipping	0	0	0	0
Exception Reporting	0	0	0	0
Final Report Output	0	0	0	0
Version Control	0	0	0	0

8.2 Performance Testing

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	17 / 5
2.	Data Responsiveness	The final output from IBM Cognos With Watson further converted into PDF or Story file, so it can be viewed by all devices.
3.	Amount Data to Rendered (DB2 Metrics)	0 KB.
4.	Utilization of Data Filters	The Utilization of data Filters like Ascending, Descending, Format and so on.
5.	Effective User Story	12
6.	Descriptive Reports	17 / 5

9. Result





10. Advantages & Disadvantages

Advantages:

- Easy to calculate a patient's length of stay.
- It is used to indicate the efficiency.
- Helpful for extension of beds based on patient's length of stay.
- User Friendly
- Easy to understand
- Secure

Disadvantages:
● Need a more dynamic User interface
● Users need to know all the fields.
● Does Not take null value as input
● Does not provide suggestions to the user

11. Conclusion

There are several complications when it comes to healthcare. We can reduce the riskofcomplications with better analysis and treatment. The solution that we provide with the analytical data will help the hospitals.

12. Future Scope

There are many possible improvements that could be explored to improve the scalabilityandaccuracy of this prediction system. As we have developed a generalized system, in future wecanuse this system for the analysis of different data sets. The performance of the health's diagnosiscan be improved significantly by handling numerous class labels in the prediction process, andit can be another positive direction of research.

13. Appendix

Github Link:
https://github.com/IBM-EPBL/IBM-Project-42107-1660649698