Analytics for Hospitals' Health-Care Data

A Project Report Submitted by

Padmacharan D SSNCE193002069

Prasanna K SSNCE193002073

Sidhesh R Allu SSNCE193002099

Tharun Arasu SK SSNCE193002111

Under the Guidance of

Industrial Mentor: Shivam Shivare

Faculty Mentor: Dr.Ramani B

Team ID: PNT2022TMID52975

Project Report Format

1. **INTRODUCTION**

    1.1. Project Overview

    1.2. Purpose

2. **LITERATURE SURVEY**

    2.1. Existing problem

    2.2. References

    2.3. Problem Statement Definition

3. **IDEATION & PROPOSED SOLUTION**

    3.1. Empathy Map Canvas

    3.2. Ideation & Brainstorming

    3.3. Proposed Solution

    3.4. Problem Solution fit

4. **REQUIREMENT ANALYSIS**

    4.1. Functional requirement

    4.2. Non-Functional requirements

5. **PROJECT DESIGN**

    5.1. Data Flow Diagrams

    5.2. Solution & Technical Architecture

    5.3. User Stories

Chapter 1

INTRODUCTION

1.1. Project Overview

Project Title: Analytics for Hospitals' Health-Care Data

Proposed Solution: Create dashboard for ease of use and understanding and to predict the length of stay of patients from past health care records.

Problem: Patients with high Length of Stay (LOS) have higher risks of spreading infection to other patients and hospital staff.

Goals:

- Know fundamental concepts and can work on IBM Cognos Analytics
- Gain a broad understanding of plotting different visualizations to provide the suitable solution.
- Able to create meaningful Visualizations and the Dashboard(s).

Services Used: IBM Cognos Analytics.

Deliverables:

- Create a website with dashboard that can visualise different patient data parameters
- Predict the length of stay of patients from other health records.

1.2. Purpose

The main purpose of this project is to accurately predict the Length of Stay for each patient on case-by-case basis so that different Hospitals and other Health Institutes 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.

Chapter 2

LITERATURE SURVEY

2.1. Existing Problems

Hospital Management face several challenges even after many years of its implementation They include

- There is a shortage of professional healthcare faculty with in-depth knowledge of HMS and other similar technologies.
- Poor acceptance of Hospital Management System Software.
- Lack of health informatics professionals capable of establishing and implementing the techniques.

There also exists the problem of patients that stay at medical institutions for longer than reqiured and also increase the risk of spreading the infection to other patients and hospital staff.

2.2. References

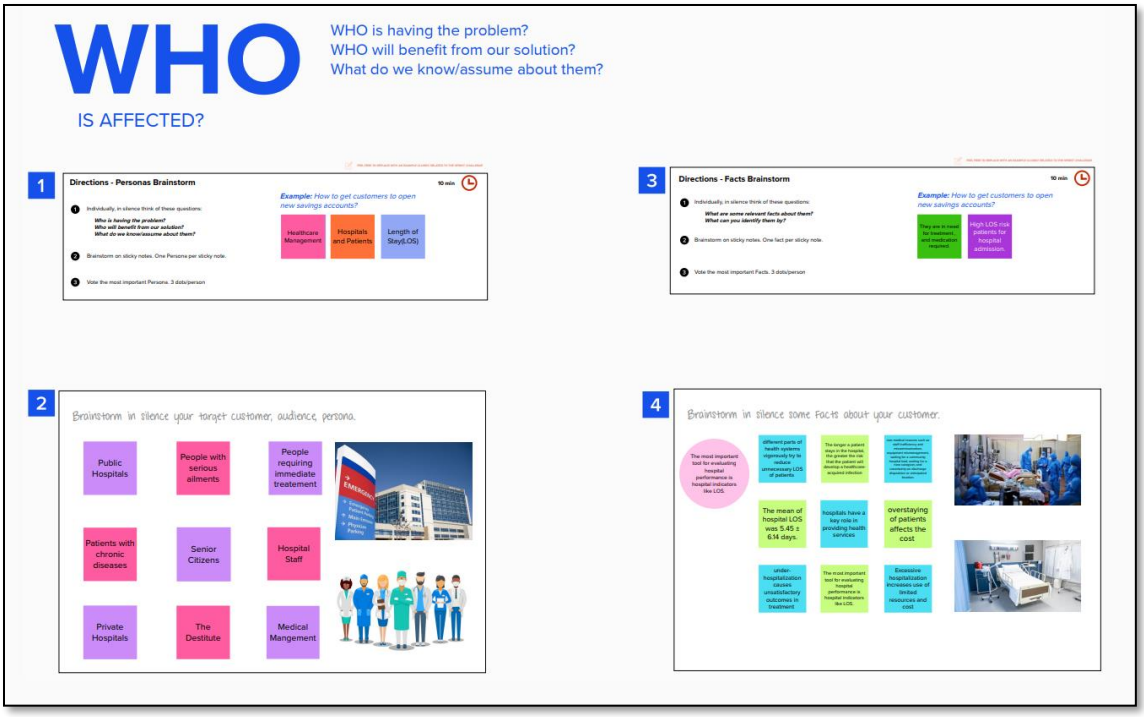
[1] A. Menon, A. M. S, A. Maria Joykutty, A. Y. Av and A. Y. Av, "Data Visualization and Predictive Analysis for Smart Healthcare: Tool for a Hospital," *2021 IEEE Region 10 Symposium (TENSYMP)*, 2021, pp. 1-8, Doi: 10.1109/TENSYMP52854.2021.9550822.

[2] Islam MS, Hasan MM, Wang X, Germack HD, Noor-E-Alam M. ASystematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining.Healthcare (Basel). 2018 May23;6(2):54. Doi: 10.3390/healthcare6020054. PMID: 29882866; PMCID: PMC6023432.

[3] S. Crossfield, O. Johnson, and T. Fleming,"Large ScaleInfrastructure for Health Data Analytics," 2016 IEEE International Conference on Healthcare Informatics (ICHI), 2016, pp. 306-306, Doi:10.1109/ICHI.2016.48.

[4] Z. Fu, X. Gu, J. Fu, M. Moattari and F. Zulkernine, "Predicting the Length of Stay of Patients in Hospitals," 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021, pp. 3150-3156, Doi: 10.1109/BIBM52615.2021.9669527.

2.3. Problem Statement Definition



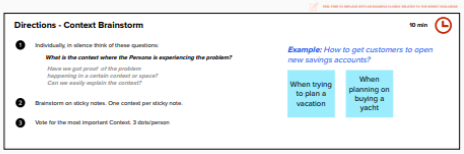
WHERE/WHEN

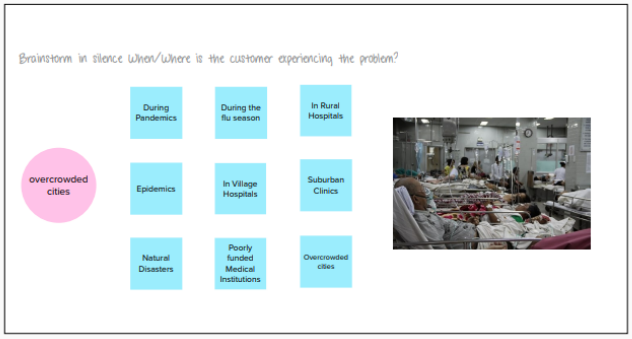
DOES IT OCCUR?

What is the context where the Persona is experiencing the problem?

Can we easily explain the context?

Have we got proof of the problem happening in a certain context or space?



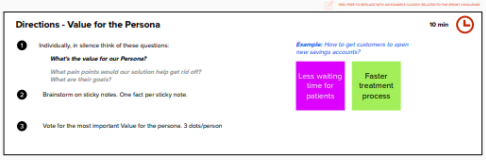


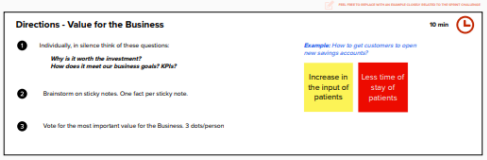
WHY

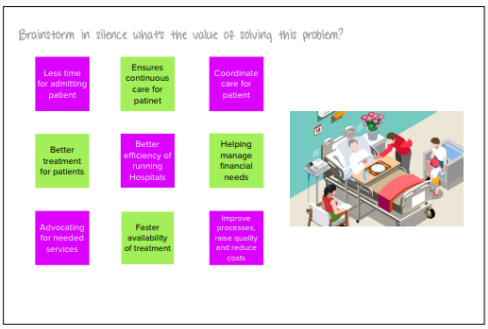
CARE ABOUT IT?

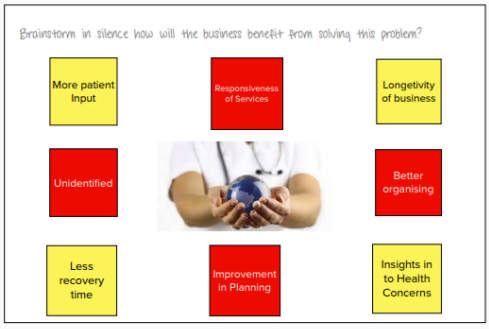
What is the most important value for the user?
What pain points would a solution help get rid of?

Why is it worth our investment?
How does it meet our business goals? KPIs?





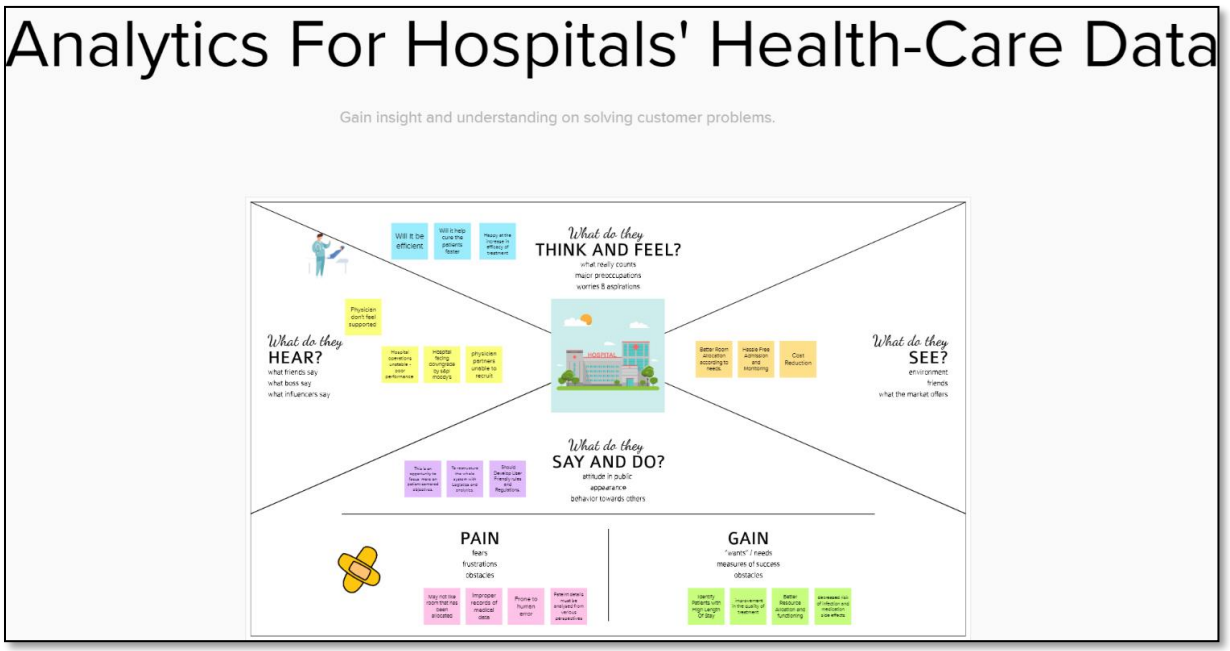




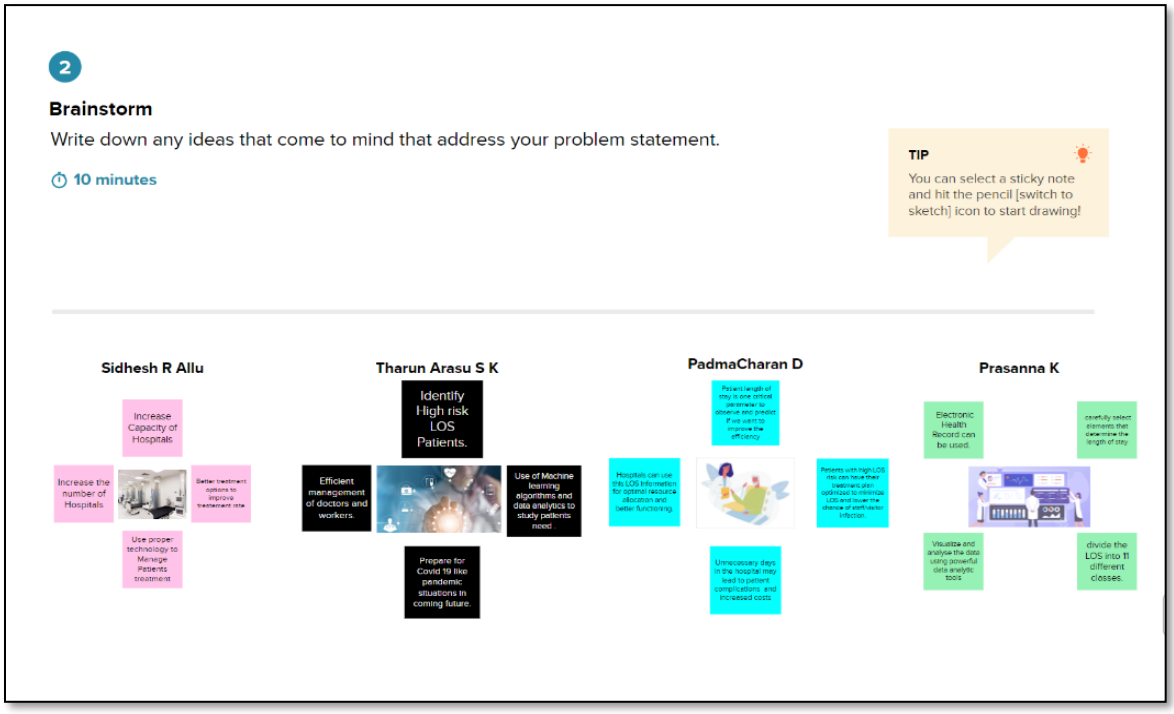
Chapter 3

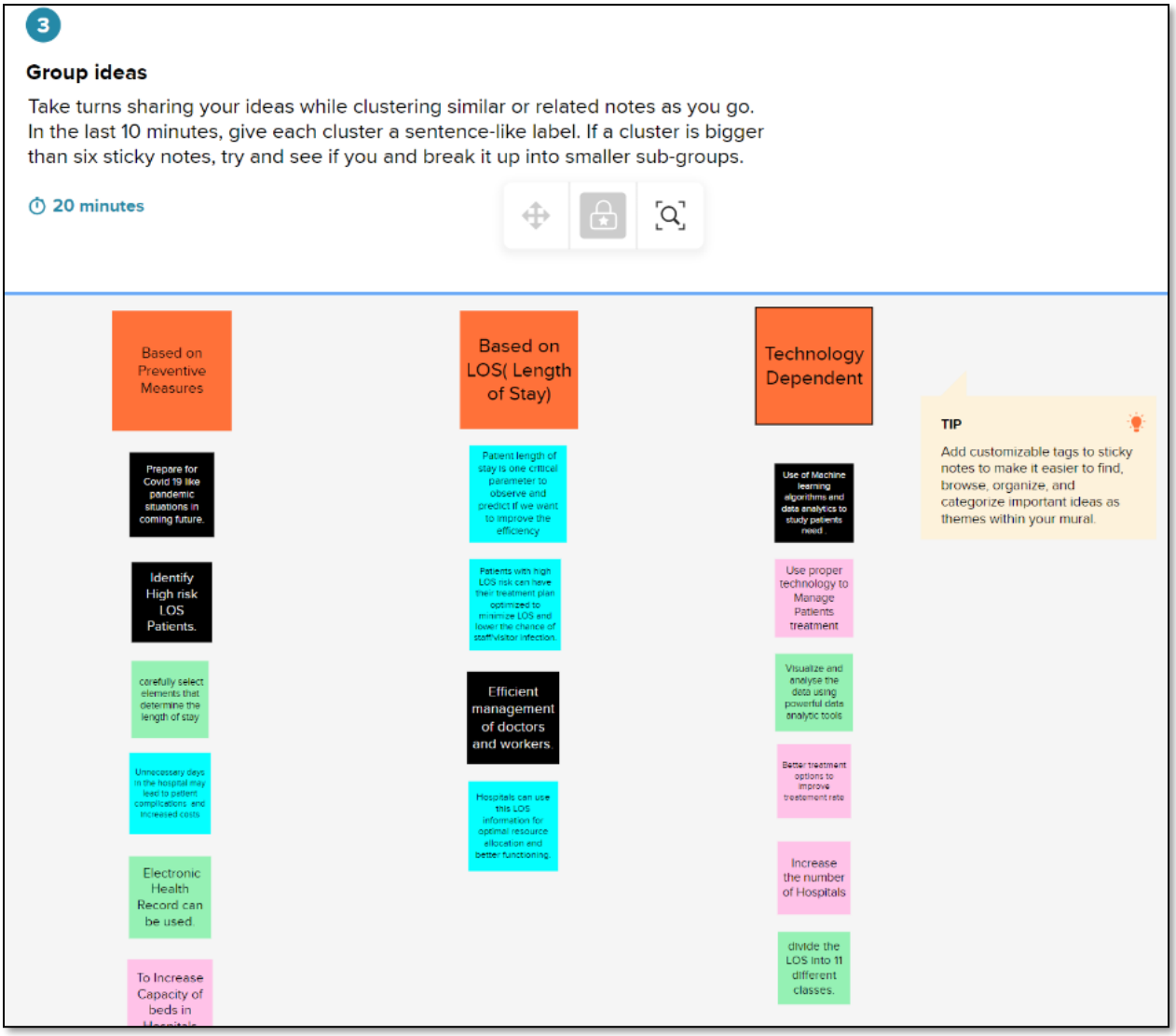
IDEATION & PROPOSED SOLUTION

3.1. Empathy Map Canvas



3.2. Ideation & Brainstorming



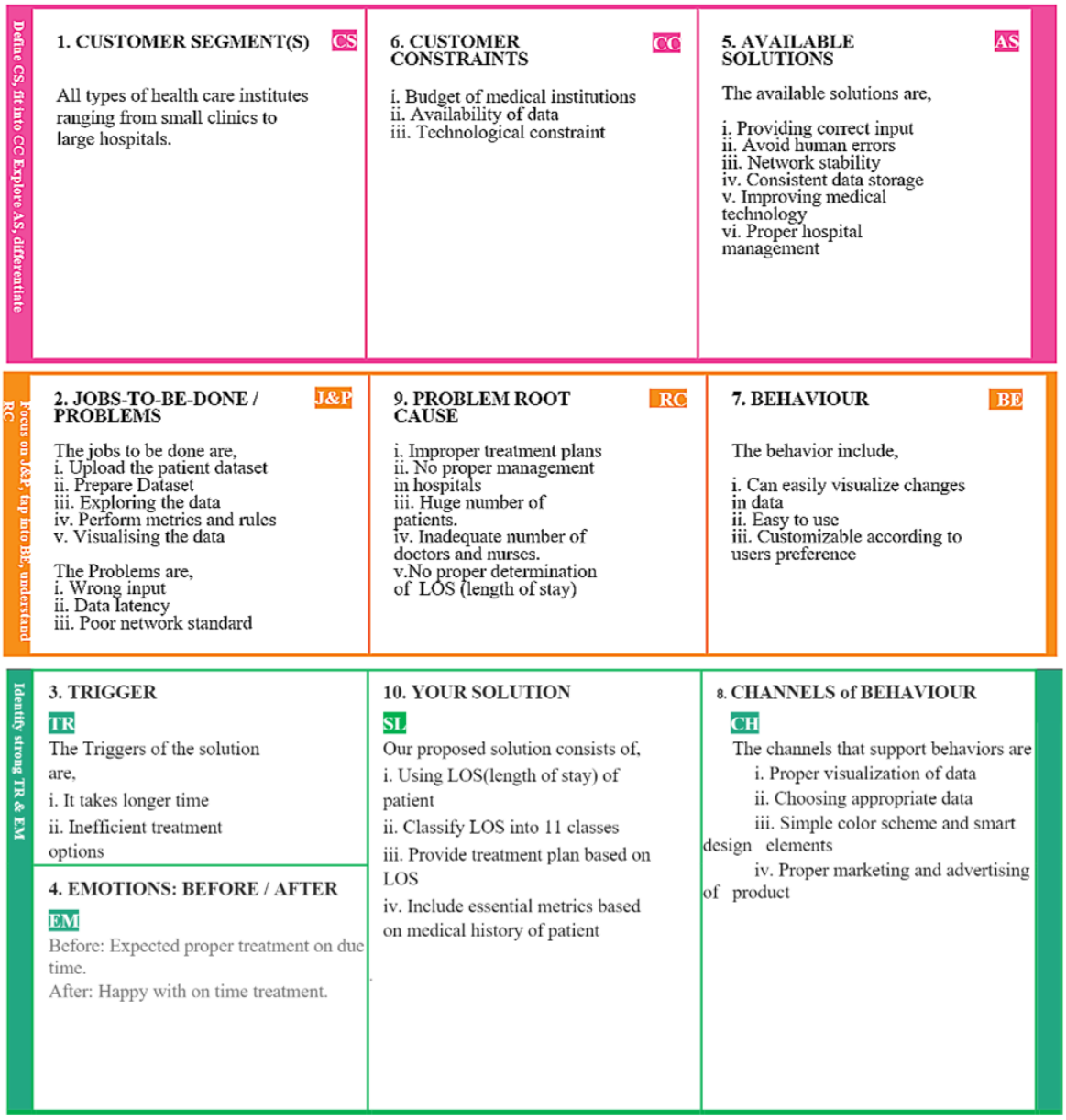

3
Group ideas
Take turns sharing your ideas while clustering similar or related notes as you go. In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.
20 minutes
Based on Preventive Measures
Prepare for Covid 19 like pandemic situations in coming future.
Identify High risk LOS Patients.
carefully select elements that determine the length of stay
Unnecessary days in the hospital may lead to patient complications and increased costs
Electronic Health Record can be used.
To Increase Capacity of beds in
Based on LOS(Length of Stay)
Patient length of stay is one critical parameter to observe and predict if we want to improve the efficiency
Patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection.
Efficient management of doctors and workers.
Hospitals can use this LOS information for optimal resource allocation and better functioning.
Technology Dependent
Use of Machine learning algorithms and data analytics to study patients need
Use proper technology to Manage Patients treatment
Visualize and analyse the data using powerful data analytic tools
Better treatment options to improve treatment rate
Increase the number of Hospitals
divide the LOS into 11 different classes.
TIP
Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.


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
Electronic Health Record can be used.
Patient length of stay is one critical parameter to observe and predict if we want to improve the efficiency
Efficient management of doctors and workers.
Use proper technology to Manage Patients treatment
Use of Machine learning algorithms and data analytics to study patients need .
Visualize and analyse the data using powerful data analytic tools
Patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection
divide the LOS into 11 different classes.
Identify High risk LOS Patients.
Increase the number of Hospitals
Better treatment options to improve treatement rate
To Increase Capacity of beds in Hospitals
carefully select elements that determine the length of stay
Unnecessary days in the hospital may lead to patient complications and increased costs
Hospitals can use this LOS information for optimal resource allocation and better functioning
Prepare for Covid 19 like pandemic situations in coming future.
Importance
If each of these tasks could get done without any difficulty or cost, which would have the most positive impact?
Feasibility

3.3. Proposed Solution

S. No.	Parameter	Description
1.	Problem Statement (Problem to besolved)	Patients with high LOS (Length of stay) need to have theirtreatment plan optimizedto reduce the chances of spreading the infection to the staffand visitors.
2.	Idea / Solution description	Allocation of hospital resources and staff based on the LOS (Length of Stay) that hasbeendetermined using DataAnalytics and Data visualisation basedon the previoushealth records of patients.
3.	Novelty / Uniqueness	The solution makes use of the patients LOS(Length of stay), which can be classified into 11 different classes. Also relies heavilyon the use of data analytics and data visualisation tools.
4.	Social Impact / CustomerSatisfaction	Efficient management of patients in hospitals leading to faster treatment rates. Hospitals startgrowing, incorporating morespecialities and treatment facilities for the patients with moreLOS (Length of stay). Judicious use of resources to treat patientsbased on LOS.
5.	Business Model (Revenue Model)	Profitable for healthcare industries due toefficient management in treatment time.
6.	Scalability of the Solution	High Scalability. Proposed solution will help improvethe workflow of the hospitalmanagement enhancing all treatment processes without hindering the dailyoperations .

3.4. Problem Solution fit



Chapter 4

REQUIREMENT ANALYSIS

4.1. Functional Requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story/ Sub-Task)
FR-1	**User Registration**	Registration through Google Forms Registration through Gmail.
FR-2	**User Confirmation**	Confirmation via Email.
FR-3	**Accuracy**	Dashboard helps predict the patient's Health risks basedon LOS (Length of Stay) .
FR-4	**Interoperability**	Dashboard helps to share the patient's information which is compatible with various other hospital management systems.
FR-5	**Compliance**	The dashboard is interactive in real time and is easy to operate.
FR-6	**Concise**	The dashboard has a clear understandable display. It is customizable and is interactive.

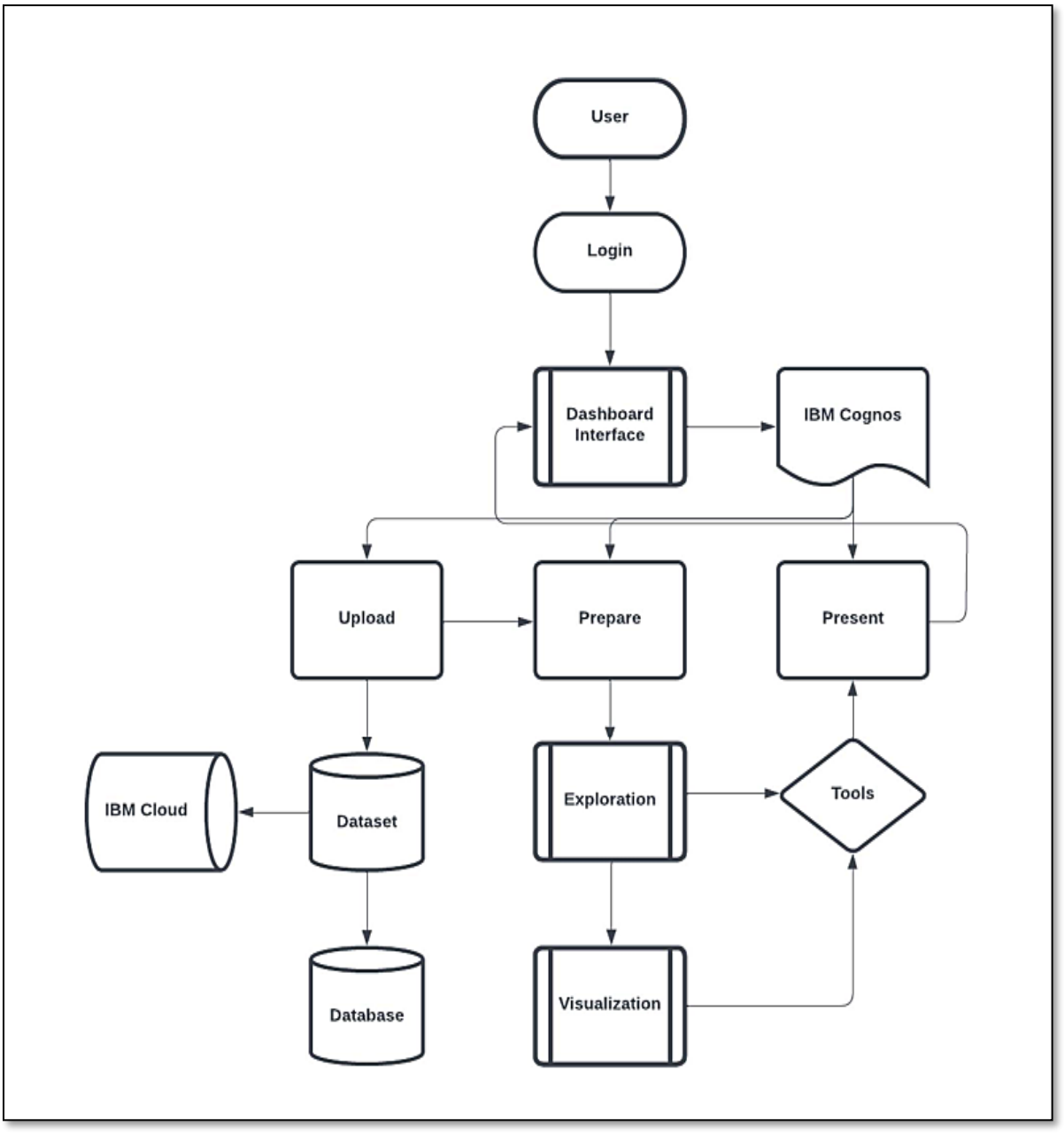
4.2. Non-Functional Requirement

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The dashboard will display the best treatment plan based on the LOS (Length of Stay) through the use of data visualization tools like pie charts and graphs.
NFR-2	**Security**	The Dashboard has inbuilt security feature to preventdata breach and loss of data. It also has the added feature of indicating authentication errors.
NFR-3	**Reliability**	The Dashboard will be consistent and reliable to the users and help the user to in an effective, efficient and reliable manner.
NFR-4	**Performance**	The Dashboard has a feedback system in which the customers can use and will be used to furtherimprove the functionality.
NFR-5	**Availability**	The Dashboard is available to meet the users demands in a timely manner and it also helps to provide necessary information to the users dataset.
NFR-6	**Scalability**	Other Hospital Management facilities can be added with the Dashboard to make a common Hospital Management System

Chapter 5

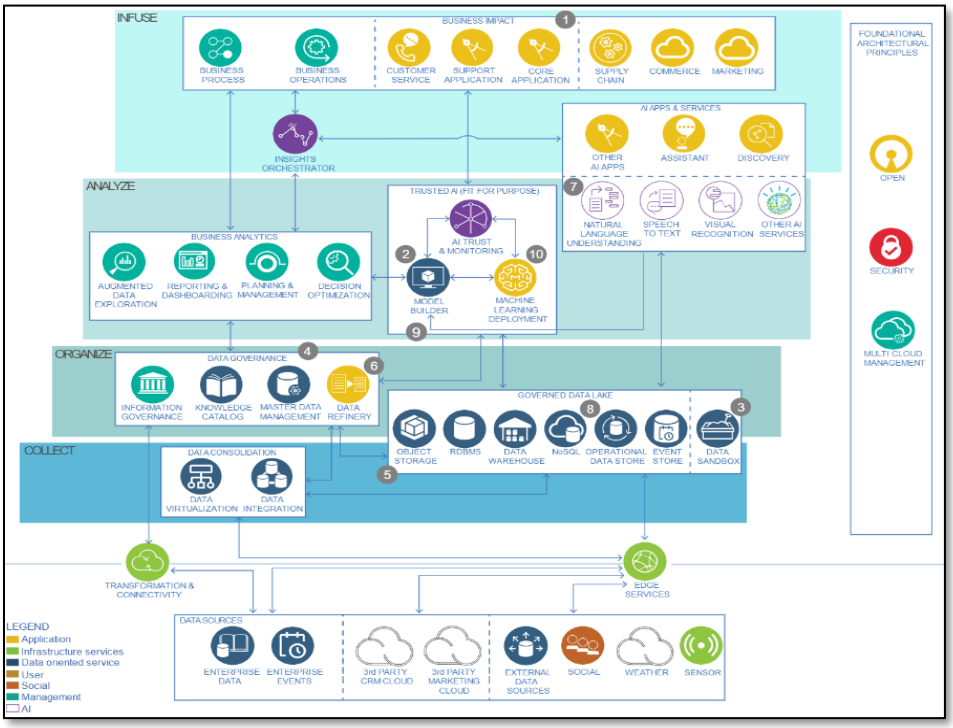
PROJECT DESIGN

5.1. Data Flow Diagram

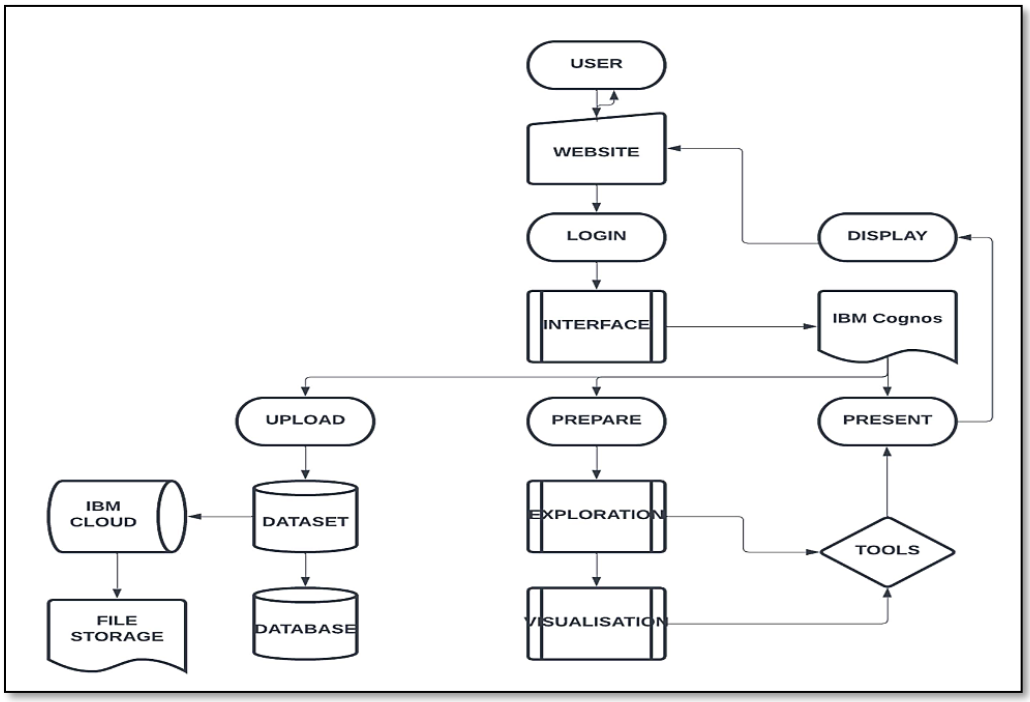


5.2. Solution & Technical Architecture

5.2.1 Solution Architecture



5.2.2. Technical Architecture



5.3. User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the dashboard by entering my email, and password, and confirmingmy password.	I can access my account in the dashboard	High	Sprint-1
		USN-2	As a user, I will receive a confirmation email once I have registered for the dashboard	I can receive a confirmationemail& click confirm	High	Sprint-1
		USN-3	As a user, I can register for the dashboardthrough social media	I can register &access the dashboard with Social Media Login	Low	Sprint-2
		USN-4	As a user, I can register for the dashboard through Gmail	I can register and access dashboard with Gmail	Medium	Sprint-2
	Login	USN-5	As a user, I can log into the application by entering email& password	I can login to the accounting my email login.	High	Sprint-2
	Dashboard	USN-6	As a user, I can use my account in my dashboard for uploading dataset.	I can login to the account for uploading dataset.	Medium	Sprint-3
Customer (Webuser)	Website	USN-7	As a user, I can use my dashboard in website	I can login into the dashboard by visiting website.	Medium	Sprint-3
Customer Care Executive		USN-8	As a user, I can contact Customer care Executive for my login.	I can contact customer executive for my login.	High	Sprint-4
Administrator		USN-9	As a user, I can contact administrator for my queries.	I can contact administrator for solving my queries.	High	Sprint-4
Exploration	Dashboard	USN-10	As a user, I can prepare data by usingExploration Techniques.	I can prepare data by using Exploration Techniques.	High	Sprint-3
Presentation	Dashboard	USN-11	As a user, I can Present data in mydashboard.	I can present data byusing my account in dashboard.	High	Sprint-4
Visualization	Dashboard	USN-12	As a user, I can Prepare Data by usingVisualization Techniques.	I can prepare data byusing Visualization Techniques.	High	Sprint-3

Chapter 6

PROJECT PLANNING & SCHEDULING

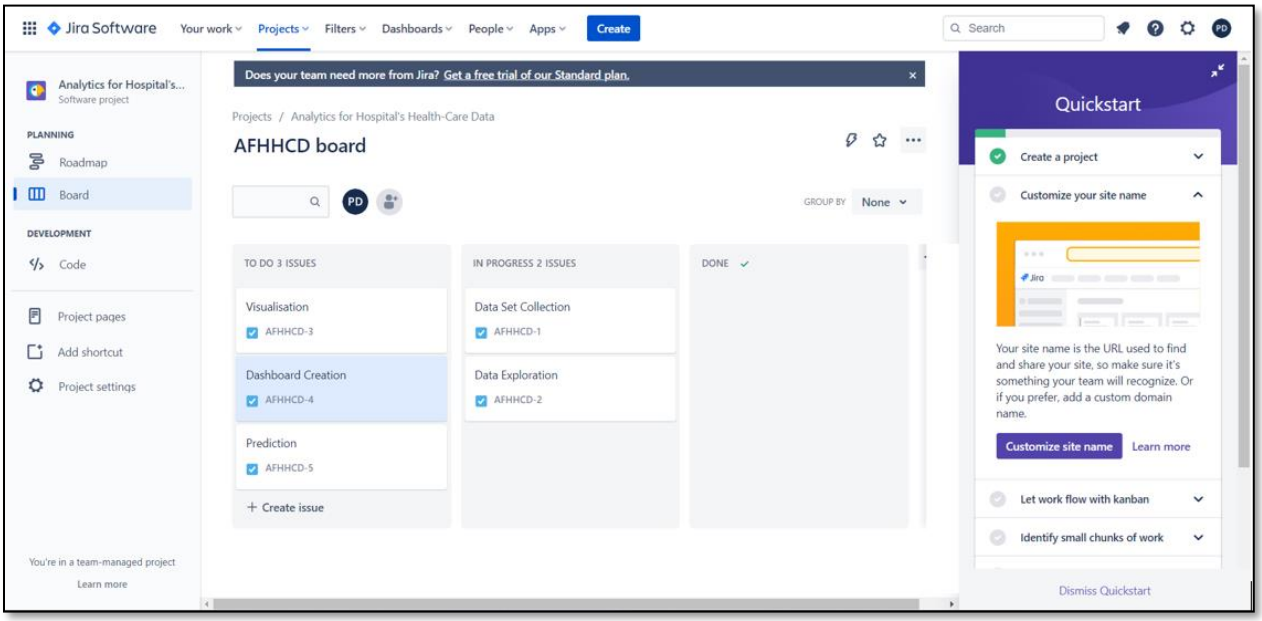
6.1. Sprint Planning & Estimation

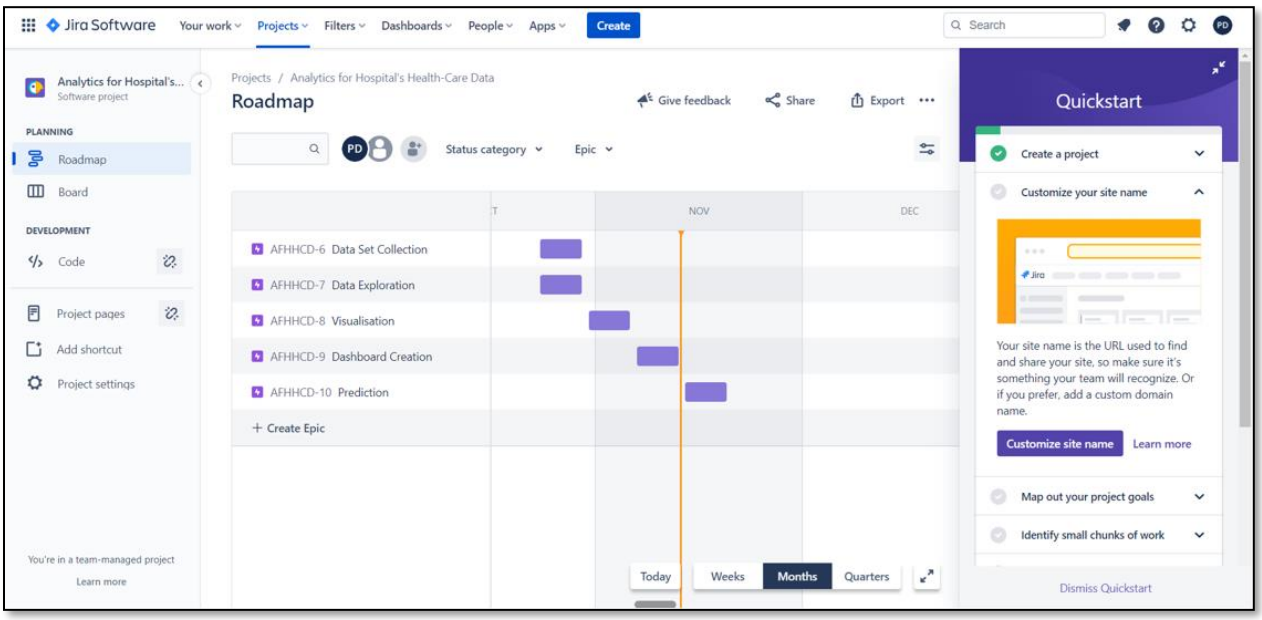
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Set Collection	USN-1	The user requires past health care data about the patient for further analysis.	5	Medium	Padmacharan D, Tharun Arasu SK
Sprint-1	Data Exploration	USN-2	Exploring the data set and cleaning the data if required.	5	High	Padmacharan D, Tharun Arasu SK
Sprint-2	Visualisation	USN-3	User can create various visualisations for better understanding of different parameters.	10	High	Prasanna K, Sidhesh R Allu
Sprint-3	Dashboard Creation	USN-4	For better user experience and viewing of visualisations dashboard can be used.	10	Medium	Padmacharan D, Tharun Arasu SK, Prasanna K, Sidhesh R Allu
Sprint-4	Prediction	USN-5	The user can predict the length of stay (LOS) of patient from past health care data.	10	High	Padmacharan D, Tharun Arasu S.K, Prasanna K, Sidhesh R Allu

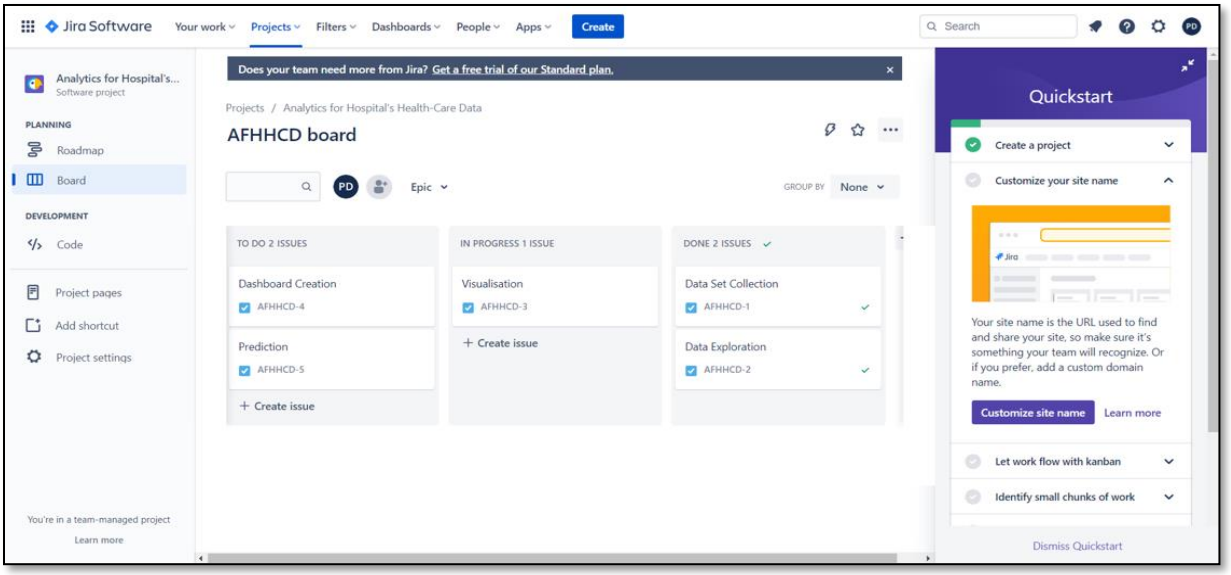
6.2. Sprint Delivery Schedule

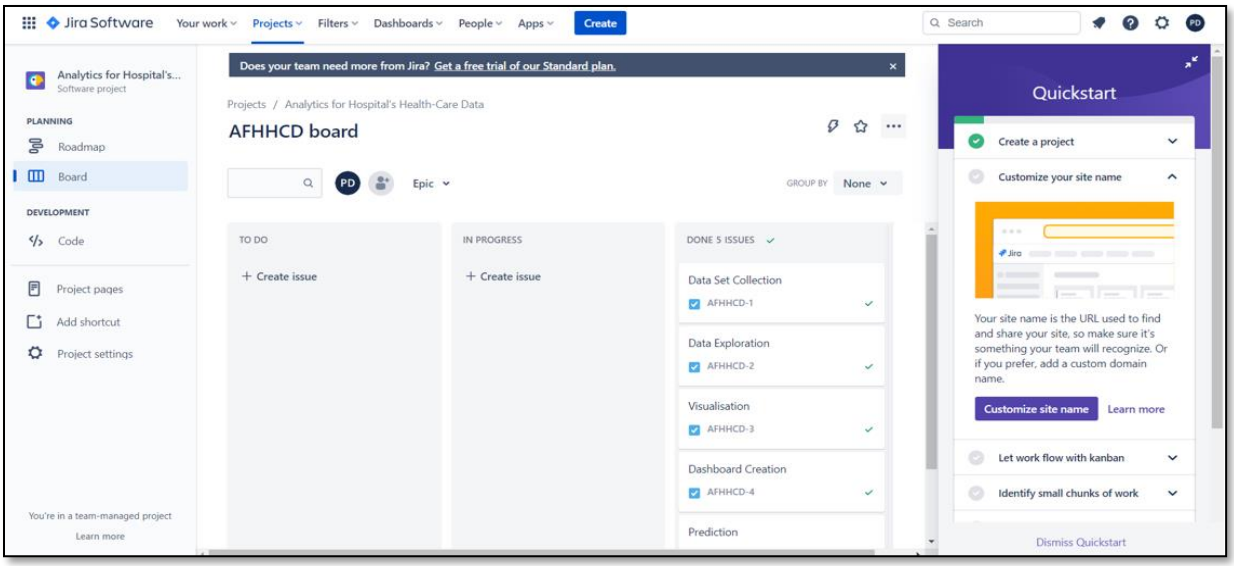
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	10	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	10	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	10	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	10	19 Nov 2022

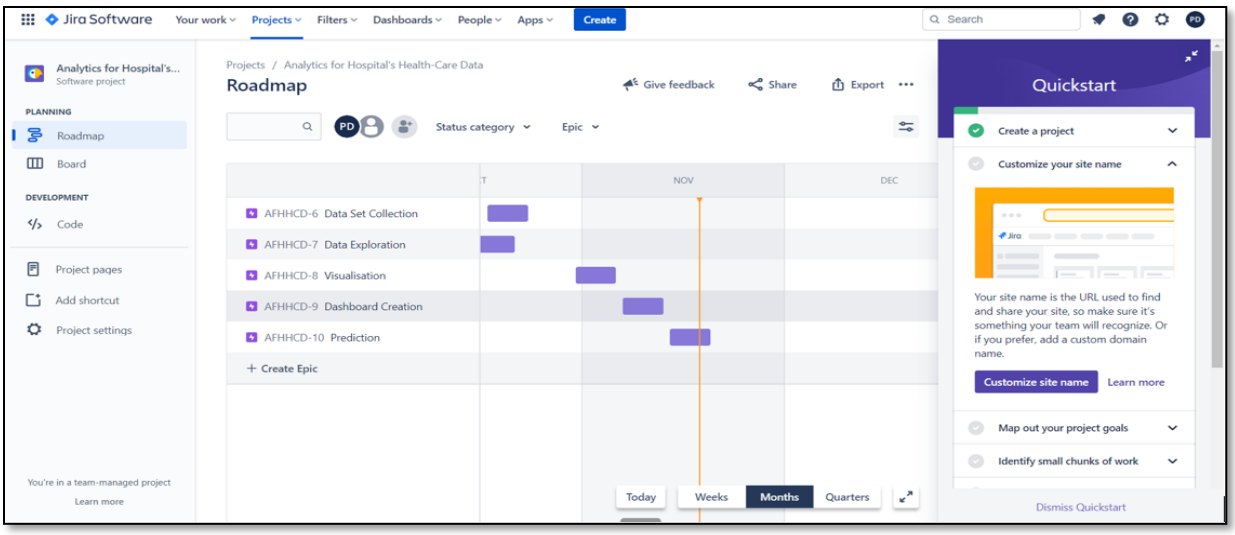
6.3. Reports from JIRA






Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Analytics for Hospital's...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Analytics for Hospital's Health-Care Data
AFHHCD board
Epic
GROUP BY None
TO DO 2 ISSUES
Dashboard Creation
AFHHCD-4
Prediction
AFHHCD-5
+ Create issue
IN PROGRESS 1 ISSUE
Visualisation
AFHHCD-3
+ Create issue
DONE 2 ISSUES
Data Set Collection
AFHHCD-1
Data Exploration
AFHHCD-2
Quickstart
Create a project
Customize your site name
Your site name is the URL used to find and share your site, so make sure it's something your team will recognize. Or if you prefer, add a custom domain name.
Customize site name
Learn more
Let work flow with kanban
Identify small chunks of work
Dismiss Quickstart


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Analytics for Hospital's...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Analytics for Hospital's Health-Care Data
AFHHCD board
Epic
GROUP BY None
TO DO
+ Create issue
IN PROGRESS
+ Create issue
DONE 5 ISSUES
Data Set Collection
AFHHCD-1
Data Exploration
AFHHCD-2
Visualisation
AFHHCD-3
Dashboard Creation
AFHHCD-4
Prediction
Quickstart
Create a project
Customize your site name
Your site name is the URL used to find and share your site, so make sure it's something your team will recognize. Or if you prefer, add a custom domain name.
Customize site name
Learn more
Let work flow with kanban
Identify small chunks of work
Dismiss Quickstart

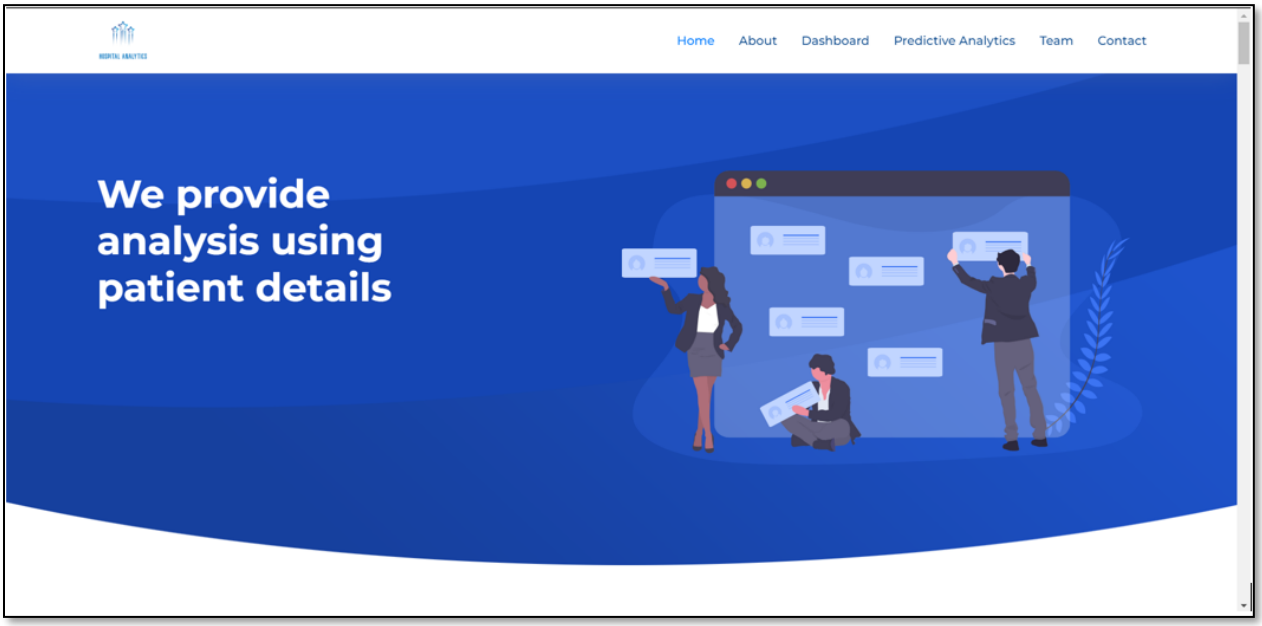

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Analytics for Hospital's...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Analytics for Hospital's Health-Care Data
Roadmap
Give feedback
Share
Export
Status category
Epic
NOV
DEC
AFHHCD-6 Data Set Collection
AFHHCD-7 Data Exploration
AFHHCD-8 Visualisation
AFHHCD-9 Dashboard Creation
AFHHCD-10 Prediction
+ Create Epic
Today
Weeks
Months
Quarters
Quickstart
Create a project
Customize your site name
Your site name is the URL used to find and share your site, so make sure it's something your team will recognize. Or if you prefer, add a custom domain name.
Customize site name
Learn more
Map out your project goals
Identify small chunks of work
Dismiss Quickstart

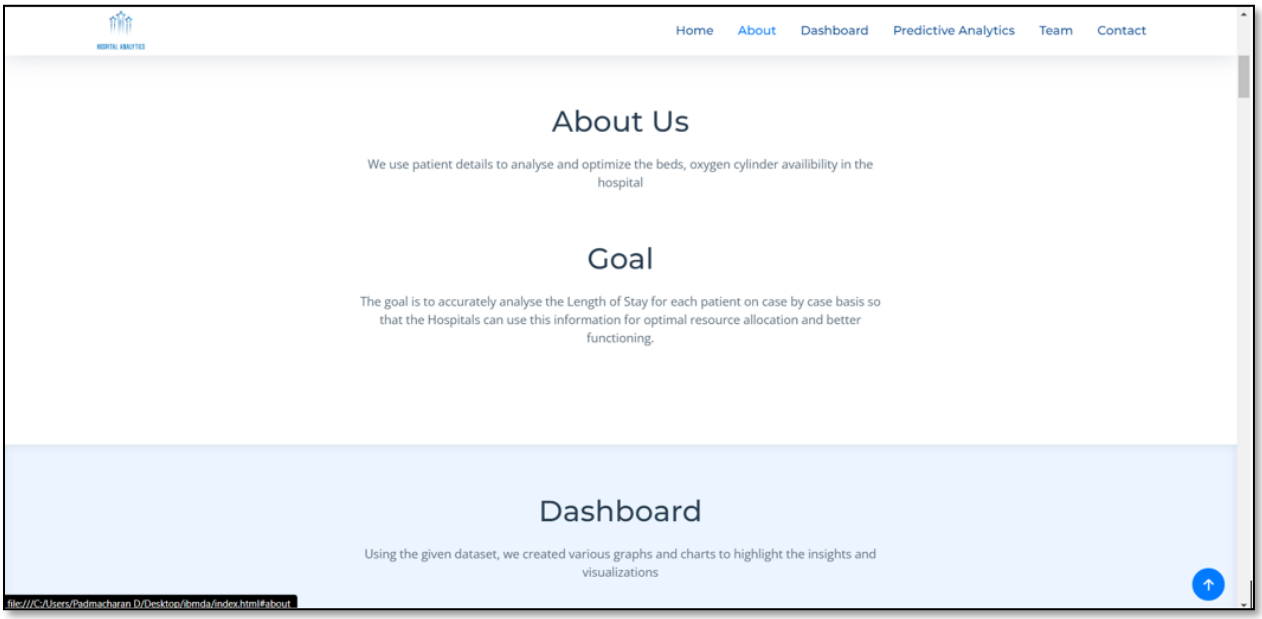
Chapter 7

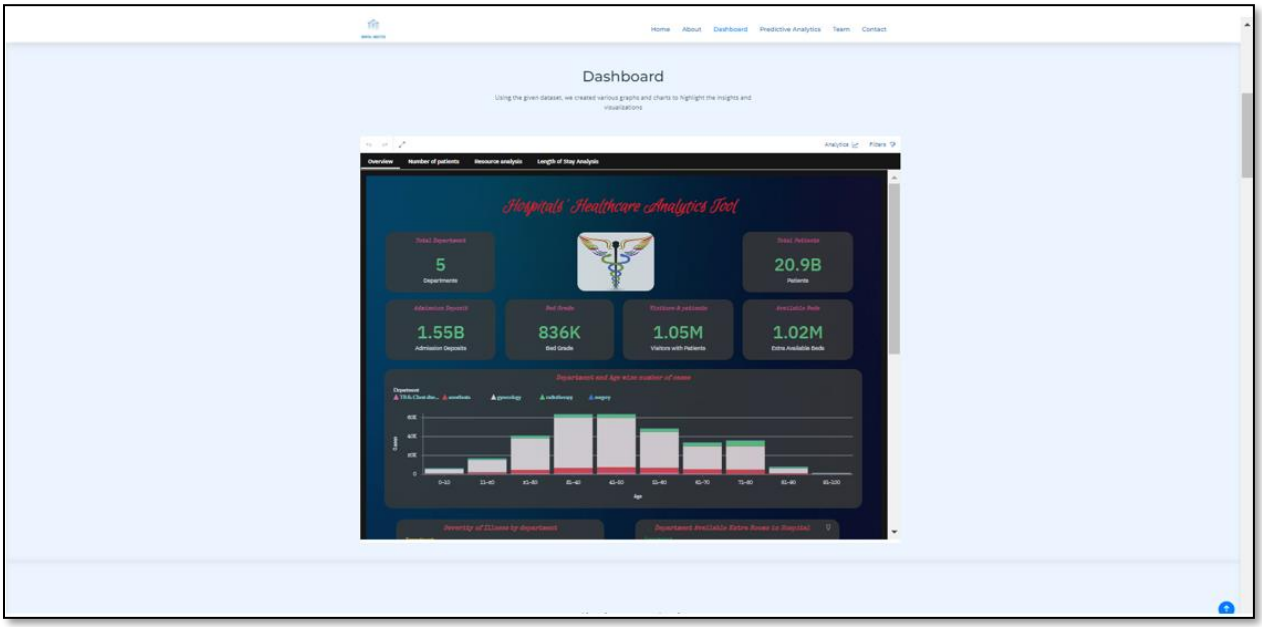
CODING & SOLUTIONING

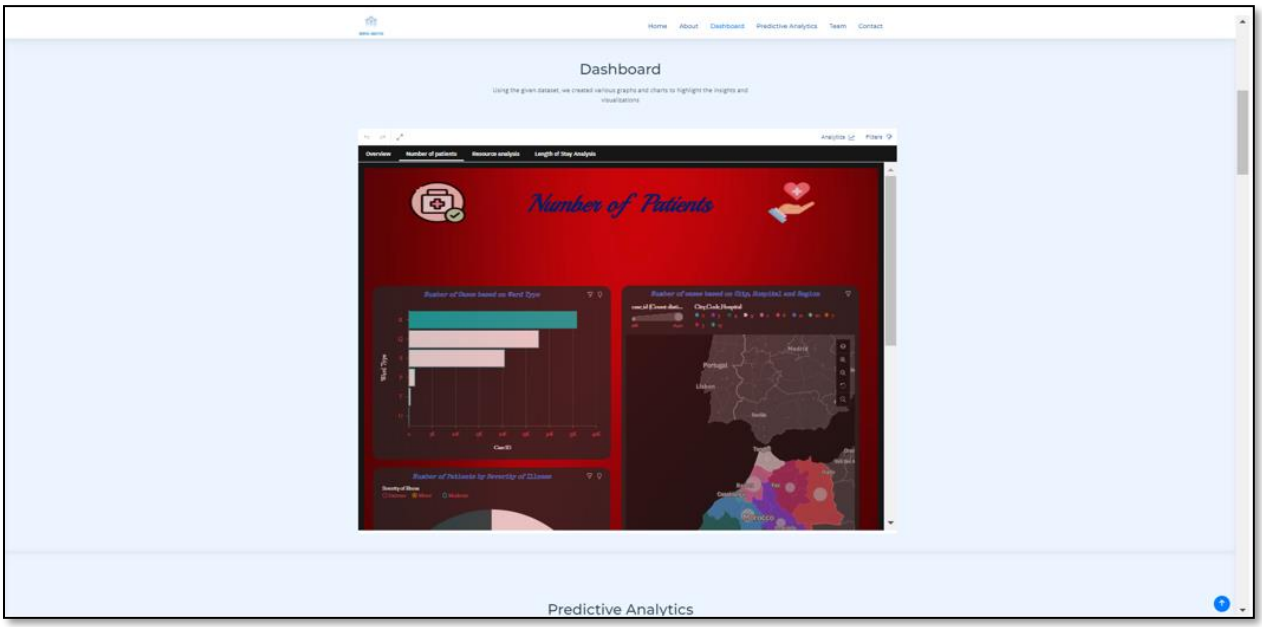
7.1. Website with dashboard

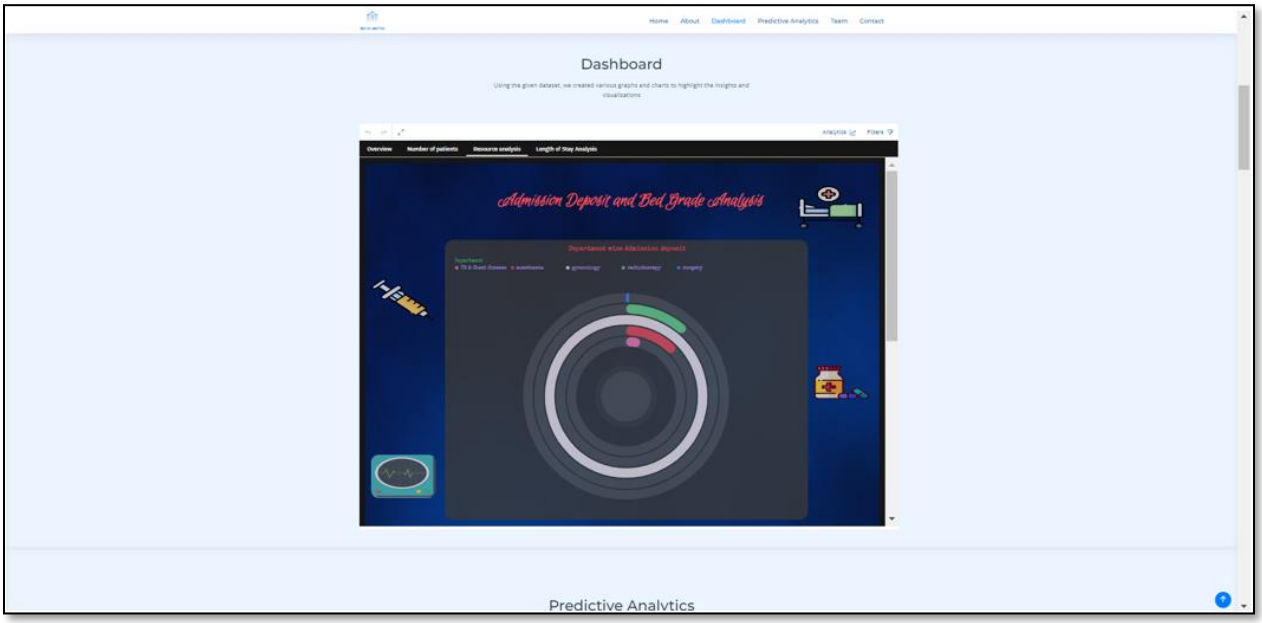
A website was created to display different visualisations between different patient data collected from previous health records.

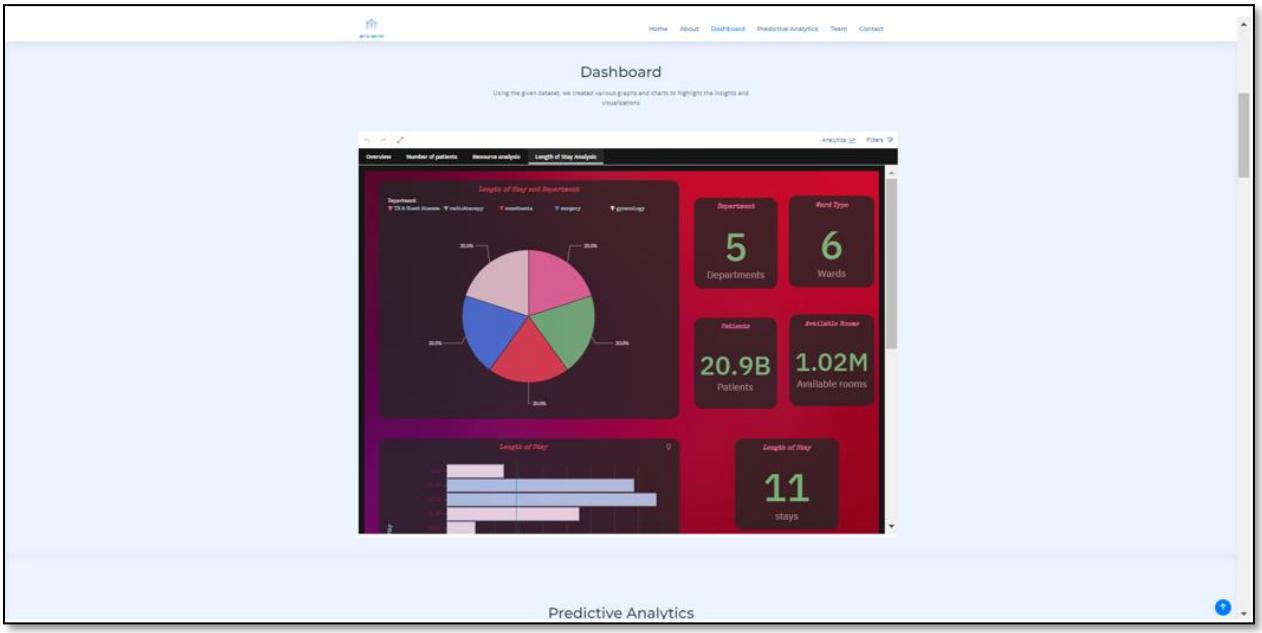





Home
About
Dashboard
Predictive Analytics
Team
Contact
Dashboard
Overview
Number of patients
Resource analysis
Length of Stay Analysis
Hospitals' Healthcare Analytics Tool
5
Departments
20.9B
Patients
1.55B
Admission Deposits
836K
Bed Grade
1.05M
Visitors with Patients
1.02M
Extra Available Beds


Home
About
Dashboard
Predictive Analytics
Team
Contact
Dashboard
Overview
Number of patients
Resource analysis
Length of Stay Analysis
Number of Patients
Predictive Analytics


Home
About
Dashboard
Predictive Analytics
Team
Contact
Dashboard
Overview
Number of patients
Resource analysis
Length of Stay Analysis
Admission Deposit and Bed Grade Analysis
Predictive Analytics


Home
About
Dashboard
Predictive Analytics
Team
Contact
Dashboard
Overview
Number of patients
Resource analysis
Length of Stay Analysis
Department
5
Departments
Ward Type
6
Wards
20.9B
Patients
1.02M
Available rooms
11
stays
Predictive Analytics

7.2. Length of Stay (LOS) Prediction

Dataset Description:

	Column	Description
0	case_id	Case_ID registered in Hospital
1	Hospital_code	Unique code for the Hospital
2	Hospital_type_code	Unique code for the type of Hospital
3	City_Code_Hospital	City Code of the Hospital
4	Hospital_region_code	Region Code of the Hospital
5	Available Extra Rooms in Hospital	Number of Extra rooms available in the Hospital
6	Department	Department overlooking the case
7	Ward_Type	Code for the Ward type
8	Ward_Facility_Code	Code for the Ward Facility
9	Bed Grade	Condition of Bed in the Ward
10	patientid	Unique Patient Id
11	City_Code_Patient	City Code for the patient
12	Type of Admission	Admission Type registered by the Hospital
13	Severity of Illness	Severity of the illness recorded at the time o...
14	Visitors with Patient	Number of Visitors with the patient
15	Age	Age of the patient
16	Admission_Deposit	Deposit at the Admission Time
17	Stay	Stay Days by the patient

```
train.head()
```

	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	patientid	City_Code_Patient	Type of Admission	Severity of Illness	Visitors with Patient	Age	Admission_Deposit	Stay
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0	31397	7.0	Emergency	Extreme	2	51-60	4911.0	0-10
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0	31397	7.0	Trauma	Extreme	2	51-60	5954.0	41-50
2	3	10	e	1	X	2	anesthesia	S	E	2.0	31397	7.0	Trauma	Extreme	2	51-60	4745.0	31-40
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0	31397	7.0	Trauma	Extreme	2	51-60	7272.0	41-50
4	5	26	b	2	Y	2	radiotherapy	S	D	2.0	31397	7.0	Trauma	Extreme	2	51-60	5558.0	41-50

```
train.Stay.value_counts()

21-30                  87491
11-20                  78139
31-40                  55159
51-60                  35018
0-10                   23604
41-50                  11743
71-80                  10254
More than 100 Days      6683
81-90                   4838
91-100                  2765
61-70                   2744
Name: Stay, dtype: int64
```

Building Model:

The ML model makes use of Logistic Regression, Random Forest Classifier, Decision Tree Classifier, KNN Classifier and XGB Classifier.

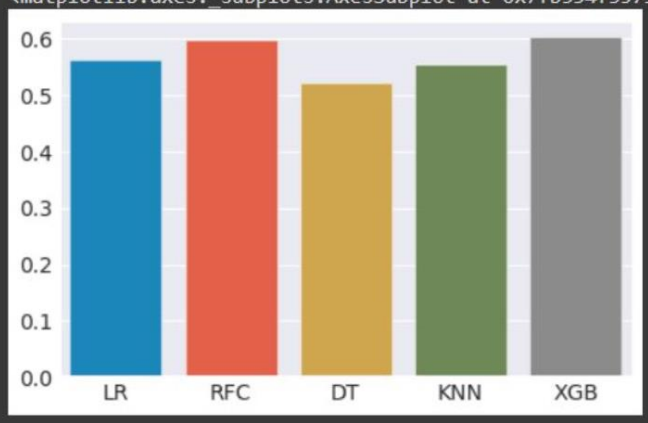
Building Model

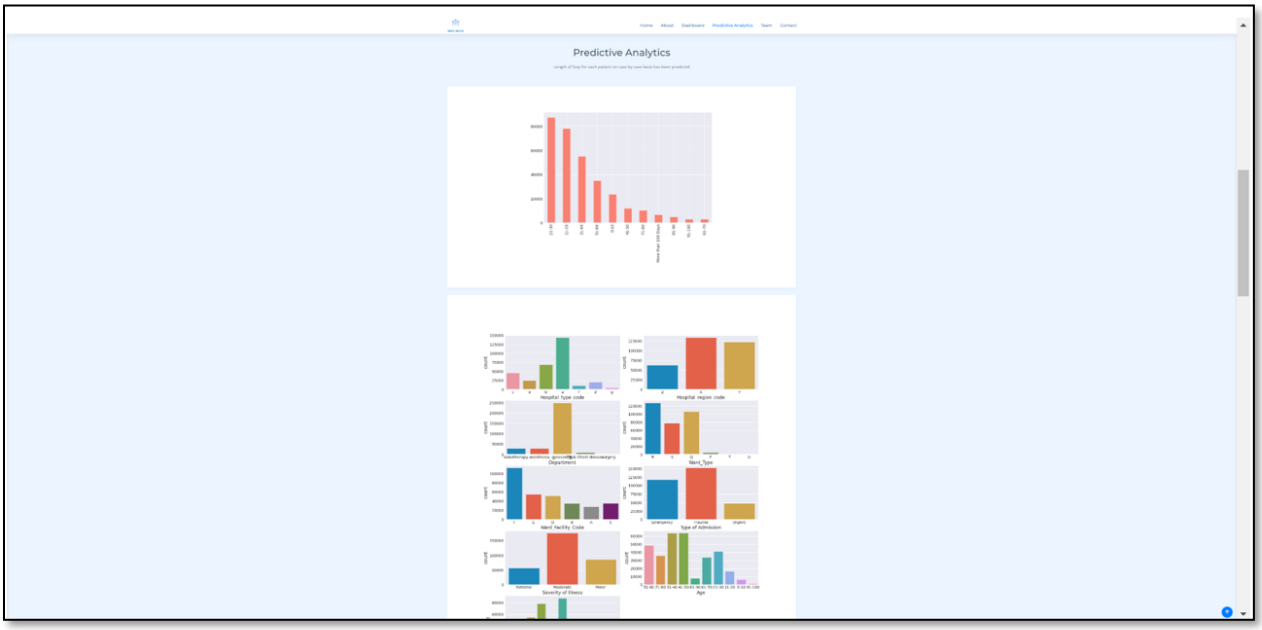
```
[ ]  from sklearn.linear_model import LogisticRegression
     from sklearn.ensemble import RandomForestClassifier
     from sklearn.tree import DecisionTreeClassifier
     from sklearn.naive_bayes import MultinomialNB
     from sklearn.neighbors import KNeighborsClassifier

     from xgboost import XGBClassifier
     from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
     from sklearn.model_selection import RandomizedSearchCV
```

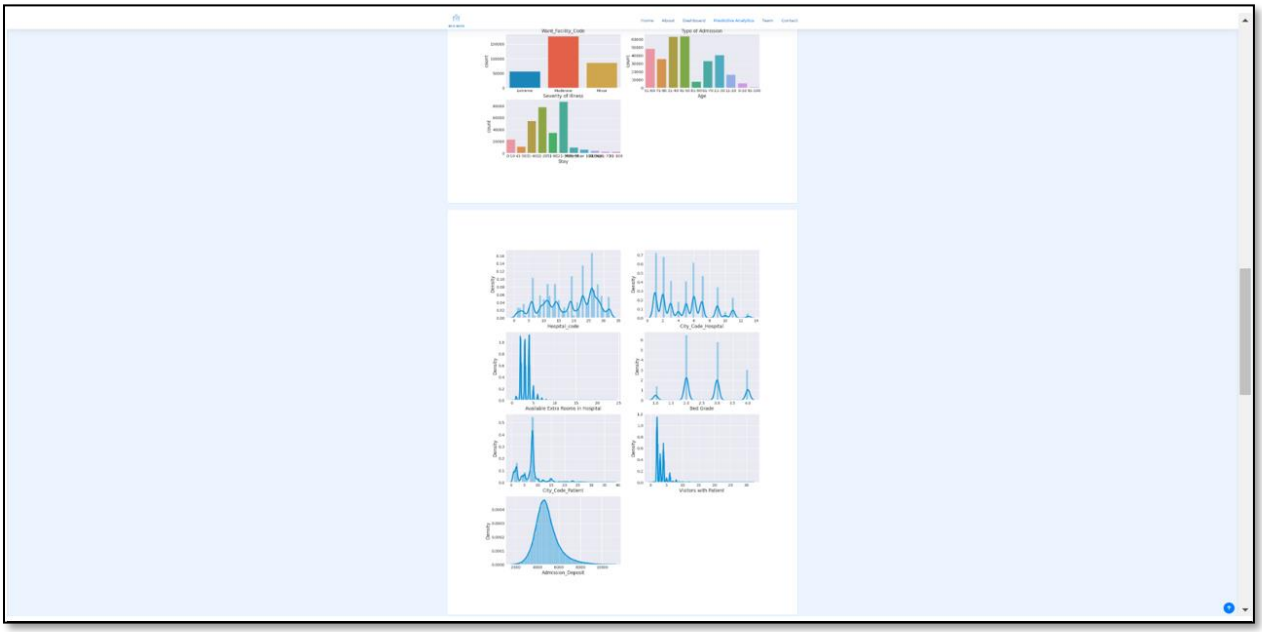
```
[ ]  value= [LogisticRegression(), RandomForestClassifier(), DecisionTreeClassifier(),  KNeighborsClassifier(), XGBClassifier()]

     key= ['LogisticRegression', 'RandomForsetClassifier', 'DecisionTreeClassifier',  'KNeighborsClassifier',  'XGBClassifier']

     models= dict(zip(key,value))
     accuracy_scores=[]
     for key,value in models.items():
         value.fit(X_train,y_train)
         y_pred= value.predict(X_test)
         accuracy= accuracy_score(y_test, y_pred)
         accuracy_scores.append(accuracy)
         print(key)
         print(accuracy)
```

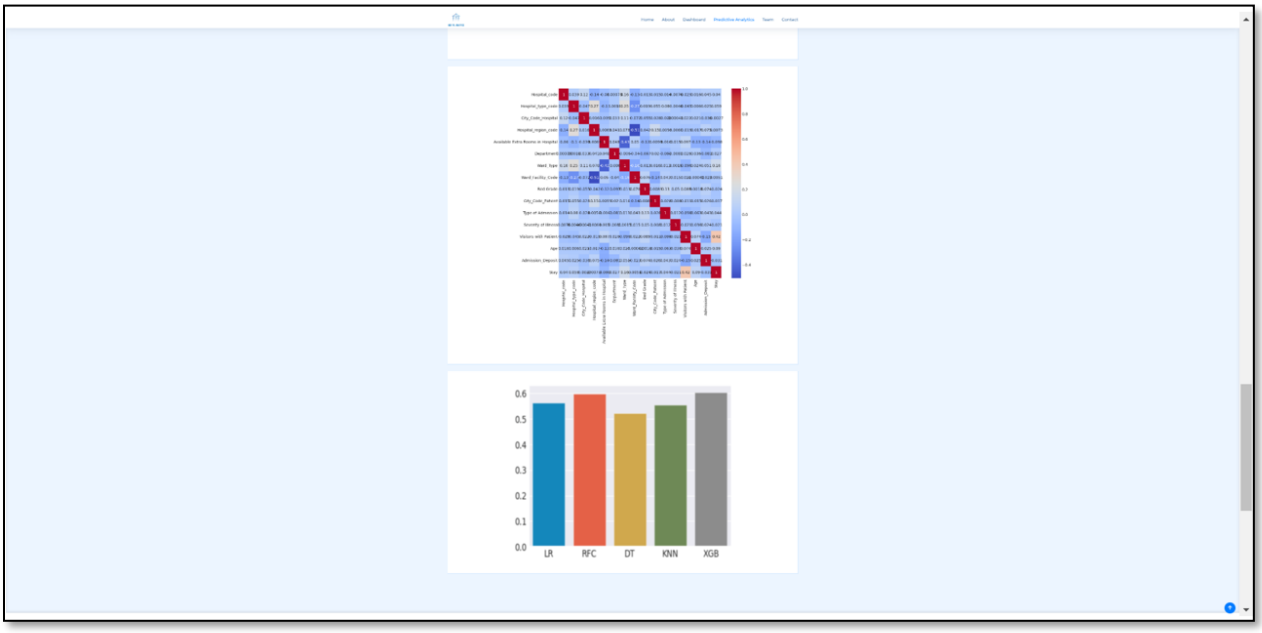
From the results we can infer that XGB Classifier gives the best accuracy for the given dataset. Thus, XGB Classifier is used to predict the Length of Stay of patients.

```
[ ] LogisticRegression
    0.5627590754930285
    RandomForsetClassifier
    0.598762718251476
    DecisionTreeClassifier
    0.5214012058786585
    KNeighborsClassifier
    0.5556462755935184
    XGBClassifier
    0.6032219570405728

[ ] sns.barplot(x= ['LR','RFC','DT','KNN','XGB'],y=accuracy_scores)

    <matplotlib.axes._subplots.AxesSubplot at 0x7fb354f35710>
```








0.6
0.5
0.4
0.3
0.2
0.1
0.0
LR
RFC
DT
KNN
XGB

Chapter 8

TESTING

8.1. Test Cases

The ML model was trained using the given train data set. The XGB Classifier's hyperparameters are tuned according to test data set and the predictions have been done using sample sub dataset. The Length of stay of patients is predicted using this model.

```
[ ]  submission['Stay']= predictions

[ ]  submission['Stay']= submission['Stay'].replace({0:'Less than 20 days', 1:'21-50 days', 2:'51-100+ days'})

[ ]  submission
```

	case_id	Stay
0	318439	Less than 20 days
1	318440	51-100+ days
2	318441	21-50 days
3	318442	21-50 days
4	318443	51-100+ days
...	...	...
137052	455491	21-50 days
137053	455492	Less than 20 days
137054	455493	21-50 days
137055	455494	Less than 20 days
137056	455495	51-100+ days

137057 rows × 2 columns

8.2. User Acceptances Testing

8.2.1. Defect Analysis

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

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	6	2	2	0	10
Duplicate	1	0	0	1	2
External	1	4	1	2	8
Fixed	5	0	6	6	17
Not Reproduced	1	1	0	1	3
Skipped	1	1	0	0	2
Won't Fix	0	1	2	1	4
Totals	15	9	11	11	46

8.2.2. Test Case Analysis

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

Section	Total Cases	Not Tested	Fail	Pass
Design	8	2	1	5
Dashboard	15	4	2	9
Responsiveness	10	2	0	8
Exception Reporting	17	5	2	10
Final Report Output	13	3	1	9

Chapter 9

RESULT

9.1. Performance Metrics

An ML model was trained to predict the Length of Stay of patients. It makes use of XGB Classifier Model. The model predicts the Length of Stay with an accuracy of 64%. The Length of Stay of 1,37,057 patients has been predicted based on 11 different parameters.

```
y_pred= grid.best_estimator_.predict(X_test)
print(accuracy_score(y_test,y_pred))
print(confusion_matrix(y_test,y_pred))
print(classification_report(y_test,y_pred))
```

```
0.6439203617635976
[[10095  9637   617]
 [ 5507 22391  2981]
 [  636  3300  8524]]
              precision    recall  f1-score   support

           0       0.62      0.50      0.55     20349
           1       0.63      0.73      0.68     30879
           2       0.70      0.68      0.69     12460

    accuracy                           0.64     63688
   macro avg       0.65      0.64      0.64     63688
weighted avg       0.64      0.64      0.64     63688
```

The final website showcases the data visualisations between various health data that has been provided in the data set. The ML model has been used to predict the Length of Stay of patients which makes it easy for the Hospital Management to give the best treatment option to patients which further makes hospitals very efficient.

The users found the deployed website very easy to use and were comfortable using the dashboard which is embedded on the website.

9.2. Model Performance Testing

S. No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations – 13 Visualizations
2.	Data Responsiveness	Yes, the dashboard visualizations change according to selected parameter.
3.	Amount Data to Rendered (DB2 Metrics)	0KB
4.	Utilization of Data Filters	Data Filters were used which include ascending, descending and format.
5.	Effective User Story	No of Scene Added – 8
6.	Descriptive Reports	No of Visualizations / Graphs - 11 Summary Values and 13 Visualizations.

Chapter 10

ADVANTAGES AND DISADVANTAGES

Advantages:

- The length of stay can be determined based on which the patient can be given a personalised treatment plan.
- The patient is given a better treatment plan to reduce length of stay there by reducing the chances of spreading infection to other patients and staff.
- The visualisations on the dashboard can help patients understand the connection of different health data.
- The dashboard is customizable and can be personalised for each patient.
- Visualising the data helps increase efficiency of hospitals form the health care management viewpoint as they can allocate proper resources for patients based on the length of stay.

Disadvantages:

- Huge amount of patient data is required for accurate results. Data Collection is a tedious and time-consuming process.
- Proper encryption should be provided to secure patient's health records.

Chapter 11

CONCLUSION

The visualisations on the dashboard help us understand the relation between different health parameters. An ML model was trained to predict the length of stay of a patient based on past medical data.
The following can be inferred from visualising and analysing past health care records

- The most common values of Age are 41-50 (20 %) and 31-40 (20 %), together occurring over 127 thousand times, which is 40 % of the total.
- The most common value of Department is gynaecology, occurring over 249 thousand times, which is 78.3 % of the total.
- 40% of patients are people within the age limit 31-50.
- Gynaecology Department is the department where patients count is very high.
- In the department of Gynaecology, available extra rooms are very high.
- Even in gynaecology department the ward types R, Q, and S have very high number of patients
- Overall, in a hospital number of cases in R type ward is very high.
- Most of the patients are moderately ill.
- In gynaecology department most of the admissions are emergency.
- In Morocco, the hospitals with city code 1 and 2 are visited by more number of patients.
- The admission deposits are high in the department of surgery.
- In the gynaecology department, the bed grade 3 is highly used.
- Ward_Type_CAT40451 is the more commonly used ward type in the hospital.
- The number of stays is high in gynaecology department.
- More number of patients are stay in the hospital for 11 – 30 days

Chapter 12

FUTURE SCOPE

The project can be further extended to help coordinate resources among all Hospitals in a particular location. Patient Health Records can be used to further predict and analyse the best type of treatment plan for a particular ailment. Data regarding different medical procedure machines can also be mentioned and their usage can be displayed on the dashboard so that hospitals which require the machine can be allocated with the appropriate resources as required.

Further resource allocation will help reduce the cost of various treatment plans and will help hospital management by making the health industry very efficient.

Chapter 13

APPENDIX

Source Code:

Website Code:
https://github.com/IBM-EPBL/IBM-Project-19798-1659706601/blob/main/Final%20Deliverables/Webpage%20Code.html

ML Code:
https://colab.research.google.com/drive/12TYMppgGEx6kD221q4mGAo96Cjtmj1A0?usp=sharing

GitHub Link:
https://github.com/IBM-EPBL/IBM-Project-19798-1659706601

Project Demo Link:
https://drive.google.com/file/d/1BwtzA2InJPyFPNfM1cBXcHMQXjbUpID0/view?usp=share_link