# A Minor Project Report

# On

# ANALYTICS FOR HOSPITALS HEALTH-CARE DATA

Submitted in partial fulfilment of requirements for the award of the

Degree of

**BACHELOR OF TECHNOLOGY**

**in**

**INFORMATION TECHNOLOGY**

Under the guidance of

## Industry Mentor

**Shivam Shivare,**

**Hari Prabu**

## Faculty Mentor

**Kalaiarasan K AP/IT**

**Team ID:** PNT2022TMID16292


Submitted by,

RAM KUMAR A (927619BIT4079)

SRIPADMESH S (927619BIT4100)

SRYAA A V (927619BIT4111)

VINOTHKUMAR L (927619BIT4125)

## TABLE OF CONTENTS

# CHAPTER-1

# INTRODUCTION

## 1.1 PROJECT OVERVIEW

Today's healthcare industries are moving from volume-based business into value based business, which requires an overwork from doctors and nurses to be more productive and efficient. This will improve healthcare practice, changing individual life style and driving them into longer life, prevent diseases, illnesses and infections. In order, it's also going to develop a healthcare data analytic technique that will promise fora better medical practice and healthcare data predictive analytics based on filling gaps of traditional healthcare data analytics techniques and overcoming its problems. Moreover, it'salso aims to promise professionals of a better quality of medical results, as well as reduce time needed to analyze healthcare data by keeping systems up to-date and sorting medical data in a logical structure along with accessing and retrieving patient's historical data fast and smoothly.

## 1.2 PURPOSE

In the area of healthcare data analytics, hospitals and clinics are looking for a new data mining techniques that will suite evolution of information technology and analyze a huge amount of complex data. The proposed technique is recommended rather than offered, since it will facilitate and enhance healthcare practice, by enabling systems to use data andanalyze it efficiently and smoothly, because it will fill the gaps of previous techniques usedin the hospitals, handle big data issue and avoid data loss, which will lead to improve care,assist diseases prediction and prevention systems and reduce cost. This technique is promising a better results and more benefits if it's applied correctly and properly.

# CHAPTER-2

# LITERATURE SURVEY

## 2.1 EXISTING PROBLEM

Implementation of Big Data in Hospital using Cluster Analytics. In the existing system technology is needed to develop rapidly, including in the health sector. Any industry and any countries, big data technology has become a significant database where the data or information to generate can be used in medicine, one of which is applied in hospitals, clinics and the private sector. Big data is evolutionary and software in a suitable environment is being developed again. Health care data drives big data, for example data on patients with heart failure and with this data, authors can use data analytic tools to detect this so that it can be anticipated so that it can reduce the death rate due to this. The problem is that many hospitals have not applied technology to detect their health, and seeing from the times the technological era is increasingly advanced, including in the field of hospitals. To get this solution, use clustering analytics with method K-means using rapid miner studio. Hopefully this analytics that can help research in medical field to do a prediction for heart failure symptoms.

## 2.2 REFERENCES


Abbott, PA & Coenen, A 2008, 'Globalization and advances in information and communication technologies: The impact on nursing and health', Nursing Outlook, vol. 56, no. 5, pp 238- 246. Al-Azzawi, H. 2014.

"Caradigm healthcare analytics." http://www.caradigm.com/media/68911/Caradigm WP-Healthcare-Analytics-Jan-2014-US-EN.PDF Retrieved 09 August, 2015.

Bakshi, K. 2014, 'Considerations for big data: architecture and approaches', Aerospace conference, IEEE, pp. 1-7.

Bertsimas, D; Bjarnadóttir, M; Kryder, J; Pandey, R; Vempala, S; Wang, G. (2008), 'Algorithmic prediction of health-care costs', Operations Research, vol. 56, no. 6, pp. 1382-2016.

Beyer, K., Ercegovac, V., Gemulla, R., Balmin, A., Eltabakh, M., Ozcan, F. and Shekita, E. 2011. "Jaql: A Scripting Language for Large Scale Semi-Structured Data Analysis." http://web.cs.wpi.edu/~meltabakh/Publications/Jaql-PVLDB2011.pdf Retrieved 09 August, 2017.

Bradley, P., & Kaplan, J. (2010), 'Turning hospital data into dollars', HFM (Healthcare Financial Management), vol. 64, no. 2, pp. 64-68.

Brownstein, CA. & Wicks, P. 2010, 'The potential research impact of patient reported outcomes on osteogenesis imperfecta', Clin orthop relat Res, vol. 468, no. 10, pp. 5-2581.

Cannon, M., & Tanner, M. (2007), Healthy competition: What's holding back healthcare and how to free it, Cato Institute, Washington, D.C.

Castro, D 2007, Improving Health Care: Why a Dose of IT May Be Just What the Doctor Ordered, The information technology and innovation foundation, U.S.

## 2.3 PROBLEM STATEMENT DEFINITION

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focusis 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 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. 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 dedicated to manage the functioning of Hospitals in a professional and optimal manner.

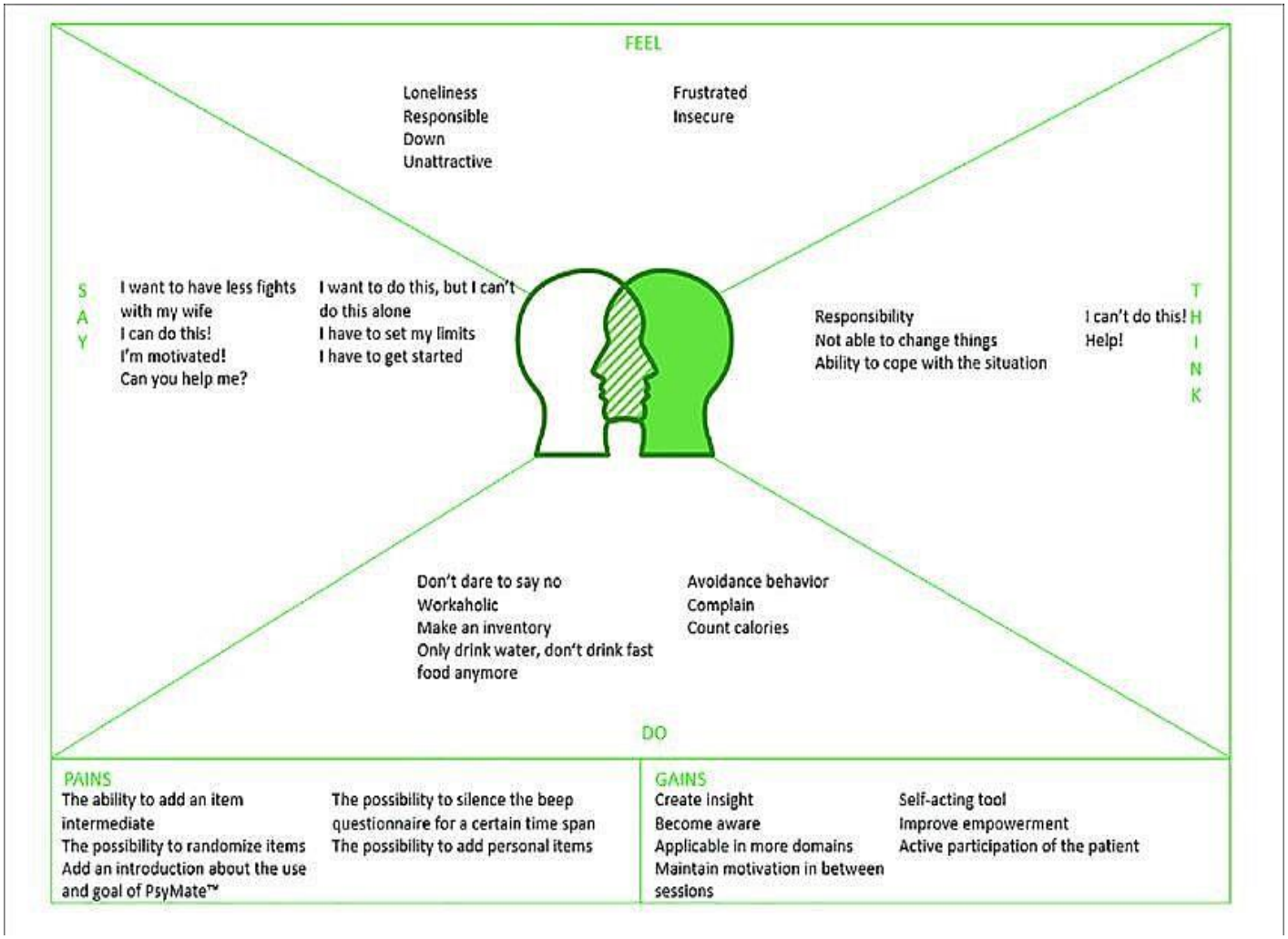
| | |
|---|---|
| I am | The person to analyze the health care data using data analytics. |
| I'm trying to | Use the recent technologies to predict the Analytics for Hospitals' Health-Care Data. |
| But | I am unaware of the existing technology that can help me a lot to analyze the health care data and I don't know to use the correct technology. |
| Because | I don't want to waste the cost and time |
| Which makes me feel | I want a best accuracy which can analyze the health care data so that the people can move with their necessary treatments. |

## 3. IDEATION & PROPOSED SOLUTION

### 3.1 EMPATHY MAP CANVAS

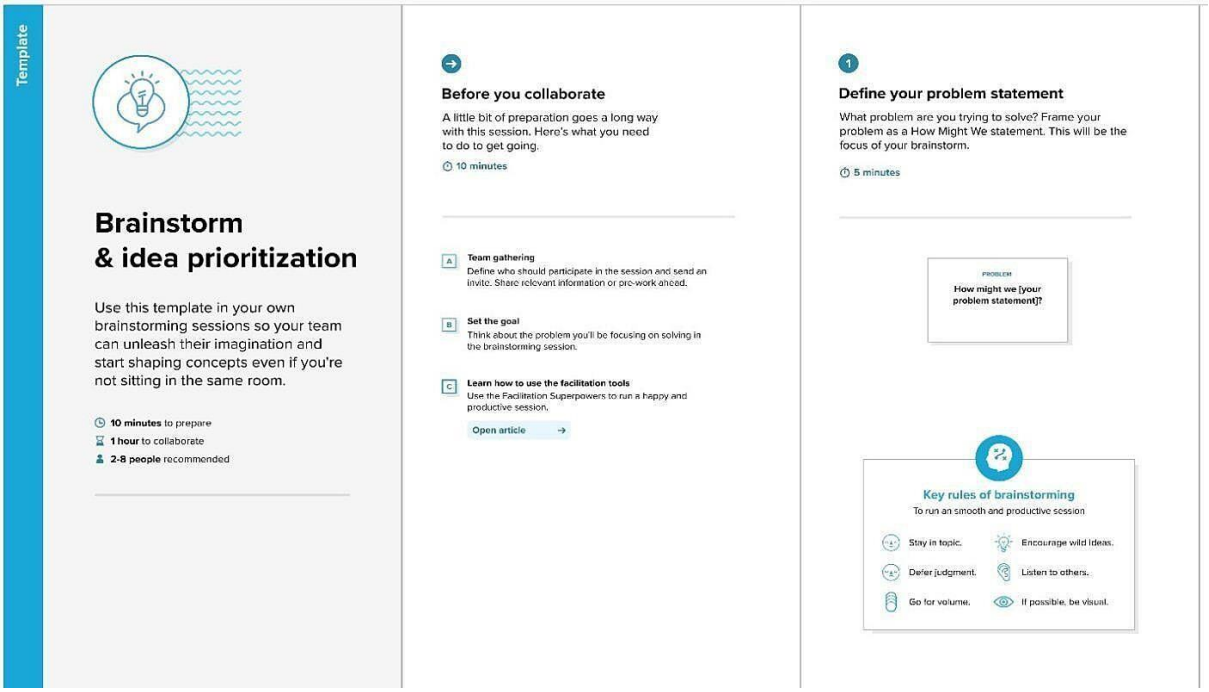
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes. It is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.

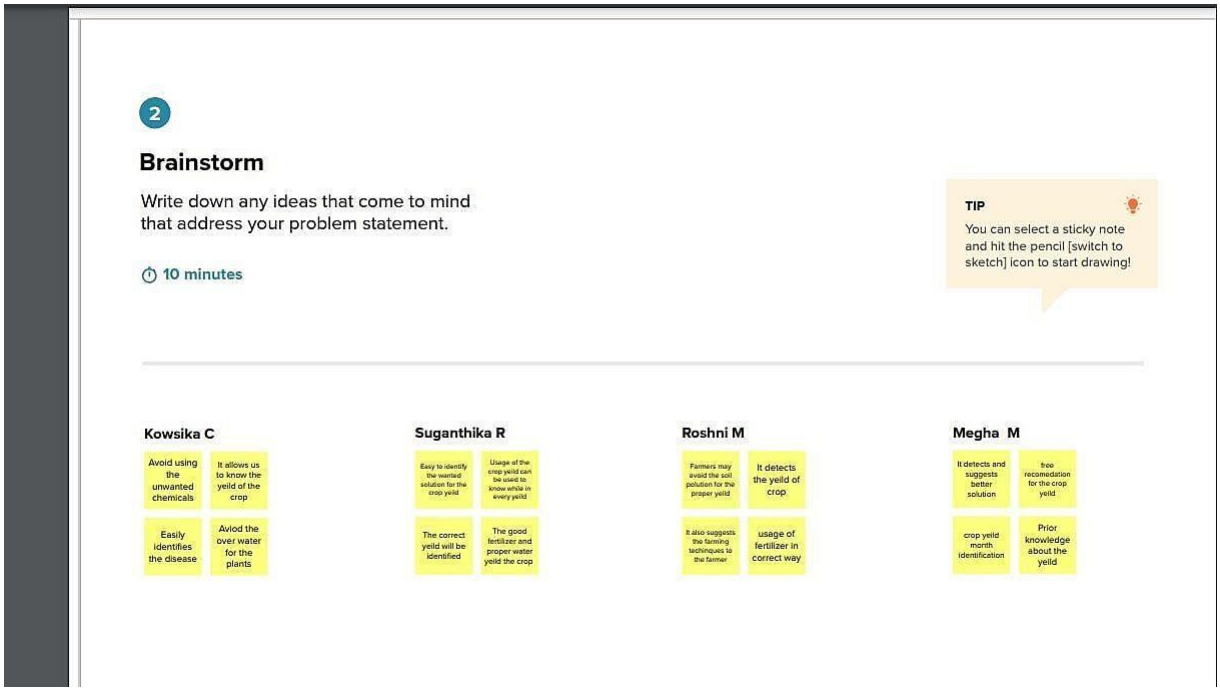
**Example:**



## 3.2 Ideation & Brainstorming

## Step 1 : Team Gathering, Collaboration and select the problem statement



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.

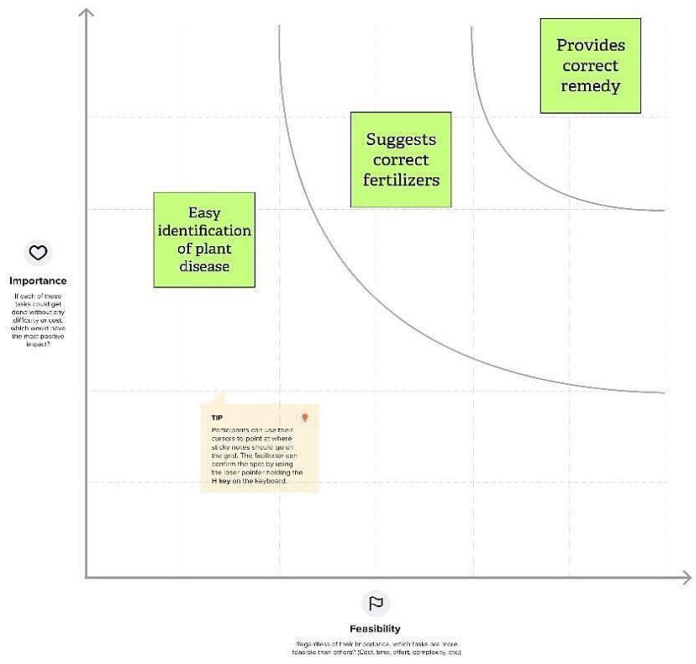
20 minutes


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




## After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

Quick add-ons

**Share the mural**
**Share a view link** to the mural with stakeholders to keep them in the loop about the outcomes of the session.

**Export the mural**
Export a copy of the mural as a PNG or PDF to attach to emails, include in slides, or save in your drive.

Keep moving forward

**Strategy blueprint**
Define the components of a new idea or strategy.
Open the template →

**Customer experience journey map**
Understand customer needs, motivations, and obstacles for an experience.
Open the template →

**Strengths, weaknesses, opportunities & threats**
Identify strengths, weaknesses, opportunities, and threats (SWOT) to develop a plan.
Open the template →

Share template feedback

### 3.3 Proposed Solution

| S.No | Parameter | Description |
|---|---|---|
| **1.** | Problem statement (problem tobe solved) | During the covid-19 pandemic, we have faced one of the difficult times of our life. Everyone seeks to survive from the great disaster. At the time of pandemic, no one get to know about which hospital has vacant beds (free beds) to admit themselves or others infected by covid. This situation made the death rate higher. |
| **2.** | Idea/solution description | Predictive analytics can create patient journey dashboards and disease trajectories that helps us to know about the patient's period of stay. It improves effective allocation of beds and other resources, treatment delivery, improves efficiencies, and so on. |
| **3.** | Novelty / uniqueness | Healthcare data frequently resides in several locations. The Collected data should be stored in central system(like centralized storage).Uniqueness of our project is that we can able to use data for different things such as which medicine is more effective and for understanding behavioural pattern of particular disease. |

| | | |
|---|---|---|
| **4.** | Social impact/customer satisfaction | effective use of resource, Enhanced diagnosis ,Improved Treatment enhancing the overall ,quality of treatment and life of patients. |
| **5.** | Business model (revenue model) | With the gathered data, redirecting the patients to particular hospital based on the vacancy, leading retailers used methods like market-basket analysis to discover insights about consumer purchase behaviour and used these insights to optimize the physical store experience, target relevant ads and streamline the supply chain, among other strategic initiatives. |
| **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 Proposed Solution Fit

# CHAPTER -4

# REQUIREMENT ANALYSIS

## 4.1 Functional requirement .

| FR No | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|---|---|---|
| FR-1 | **Dashboard** | Dashboard through User can view the details |
| FR-2 | **User view** | User have different type of view that information presented in dashboard. |
| FR-3 | **Interoperability** | Dashboard helps to share the patient's information interoperable to the hospitals in timely manner |
| FR-4 | **Accuracy** | The compliance of a dashboard is like to use very interactively in real time by the hospitals. |
| FR-5 | **Compliance** | These dashboards are clear, intuitive, and customizable and interactive in manner. |

## 4.2 Non-Functional requirements

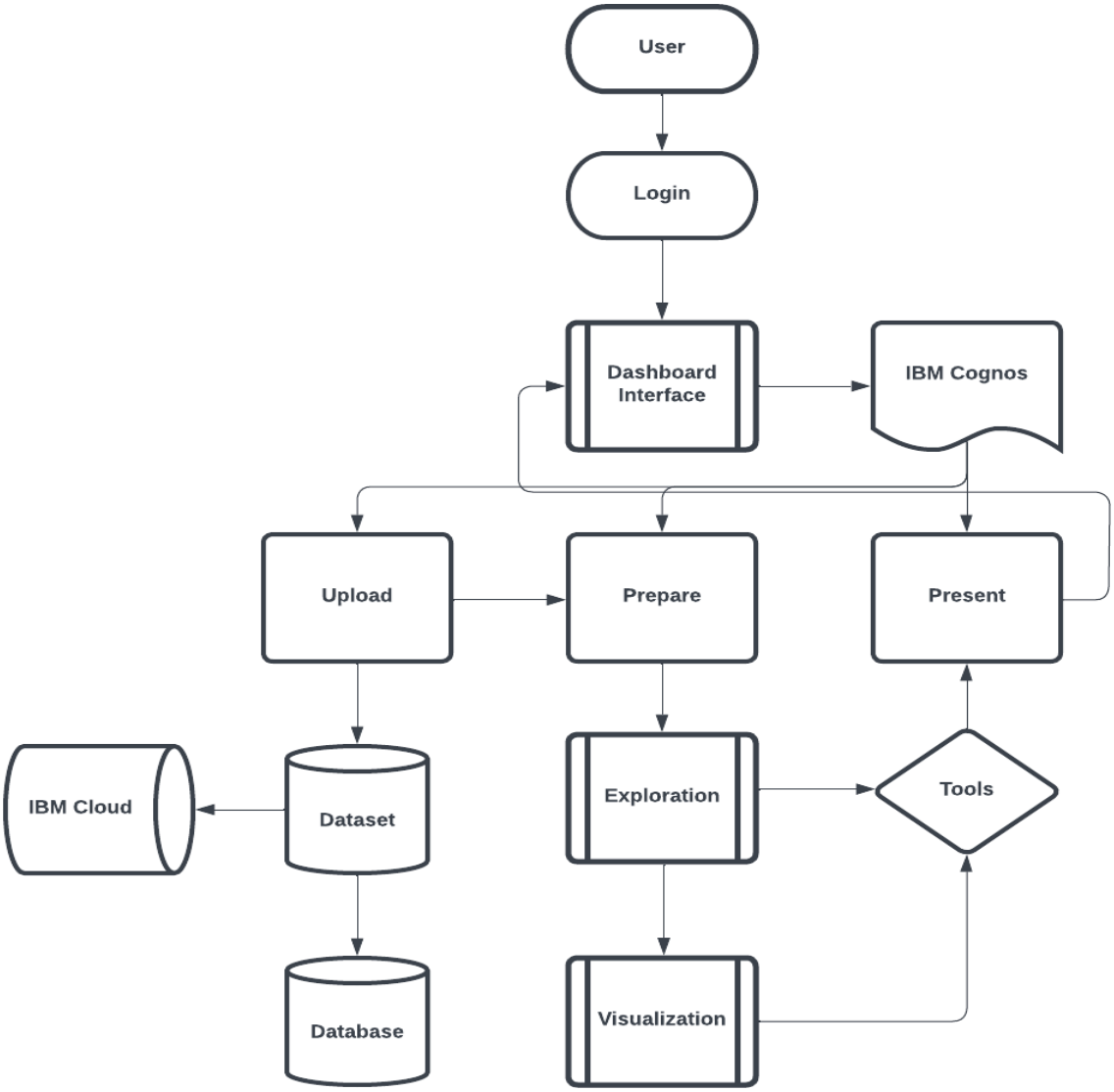
| FR No. | Non-Functional Requirement | Description |
|---|---|---|
| NFR-1 | **Usability** | This Dashboards are designed to offer a comprehensive overview of patient's LOS, and do so through the use of data visualization tools like charts and graphs |

| | | |
|---|---|---|
| **NFR-2** | **Security** | The Dashboard helps to indicate the current threatlevel to 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 users and helps the user to use in effective, efficientand reliable manner. |
| **NFR-4** | **Performance** | This dashboard can scan the backend users and analyzing the frequency in which they visit the dashboard helps understand how useful and helpful the data displayed is for tasks. |
| **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** | The layers used in the dashboard are a hosted feature layer, feature layer view, or hosted tile layer. |

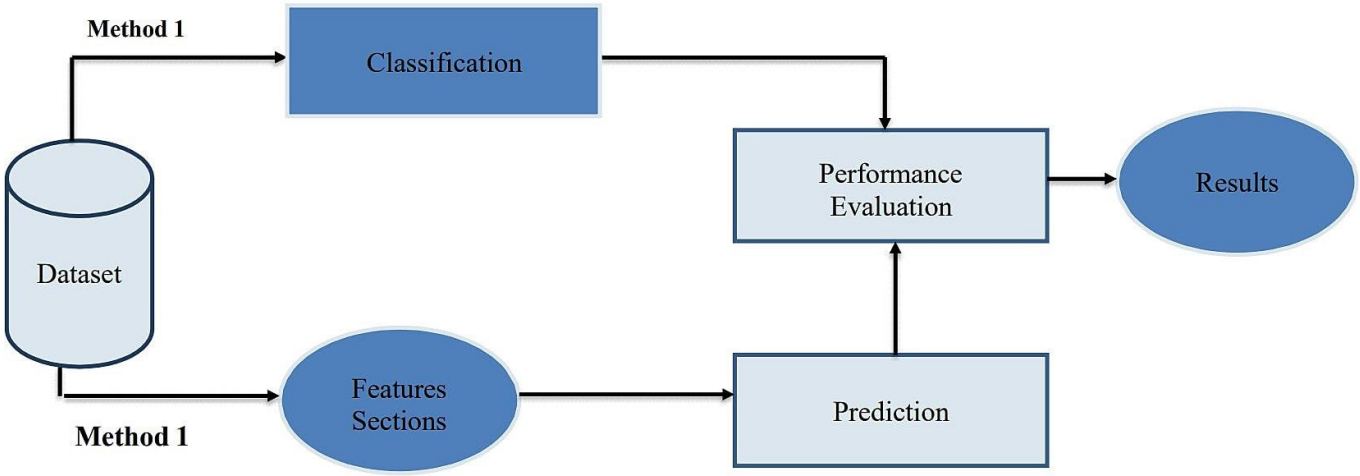
# CHAPTER -5

# PROJECT DESIGN

## 5.1 Data Flow Diagrams



## 5.2 Solution & Technical Architecture




## Table:1 Components and Technology

| S.No | Component | Description | Technology |
|---|---|---|---|
| 1. | Importing data | Data Import lets you upload data from external sources and combine it with data you collect via Analytics | Python, numpy, pandas. |
| 2. | Data Cleaning | Data cleaning is a process by which inaccurate, poorly formatted, or otherwise messy data is organized and corrected | Python, numpy, pandas |
| 3. | Data Preprocessing | Data preprocessing, a component of data preparation, describes any type of processing performed on raw data to prepare it for another data processing procedure | Python, numpy, scipy, pandas |
| 4. | Training data | Training data is the subset of original data that is used to train the machine learning model, | Numpy, scipy, pandas |
| 5. | Testing data | Test data is data which has been specifically identified for use in tests, typically of a computer program. | Numpy, scipy, pandas |
| 6. | Machine learning model | A machine learning model is a file that has been trained to recognize certain types of patterns. You train a model over a set of data, providing it an algorithm that it can use to reason over and learn from those data | Numpy, scipy, pandas, sklearn |
| 7. | Improve model performance | Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right. | sklearn |
| 8. | Checking accuracy | A data accuracy check, sometimes called a data sanity check, is a set of quality validations that takeplace before using data. | Sklearn |

## 5.3 User Stories

| User Type | Functional Requirement (Epic) | User Story Number | User Story / Task | Priority | Release |
|---|---|---|---|---|---|
| Customer (Mobile user) | Dashboard | **User Number-1** | As a user, I access the data in visualize mode | Medium | Sprint-1 |
| | | **User Number-2** | As a user, I can access the data from Queries, graph ,pie chart | High | Sprint-1 |

# CHAPTER-6

# PROJECT PLANNING & SCHEDULING

## 6.1 Sprint Planning & Estimation

Sprint planning is an event in scrum that kicks off the sprint. The purpose of sprint planning is to define what can be delivered in the sprint and how that work will be achieved. Sprint planning is done in collaboration with the whole scrum team. In scrum, the sprint is a set period of time where all the work is done. However, before you can leap into action you have to set up the sprint. You need to decide on how long the time box is going to be, the sprint goal, and where you're going to start. The sprint planning session kicks off the sprint by setting the agenda and focus. If done correctly, it also creates an environment where the team is motivated, challenged, and can be successful. Bad sprint plans can derail the team by setting unrealistic expectations. The following is the sprint planning and estimation for the project

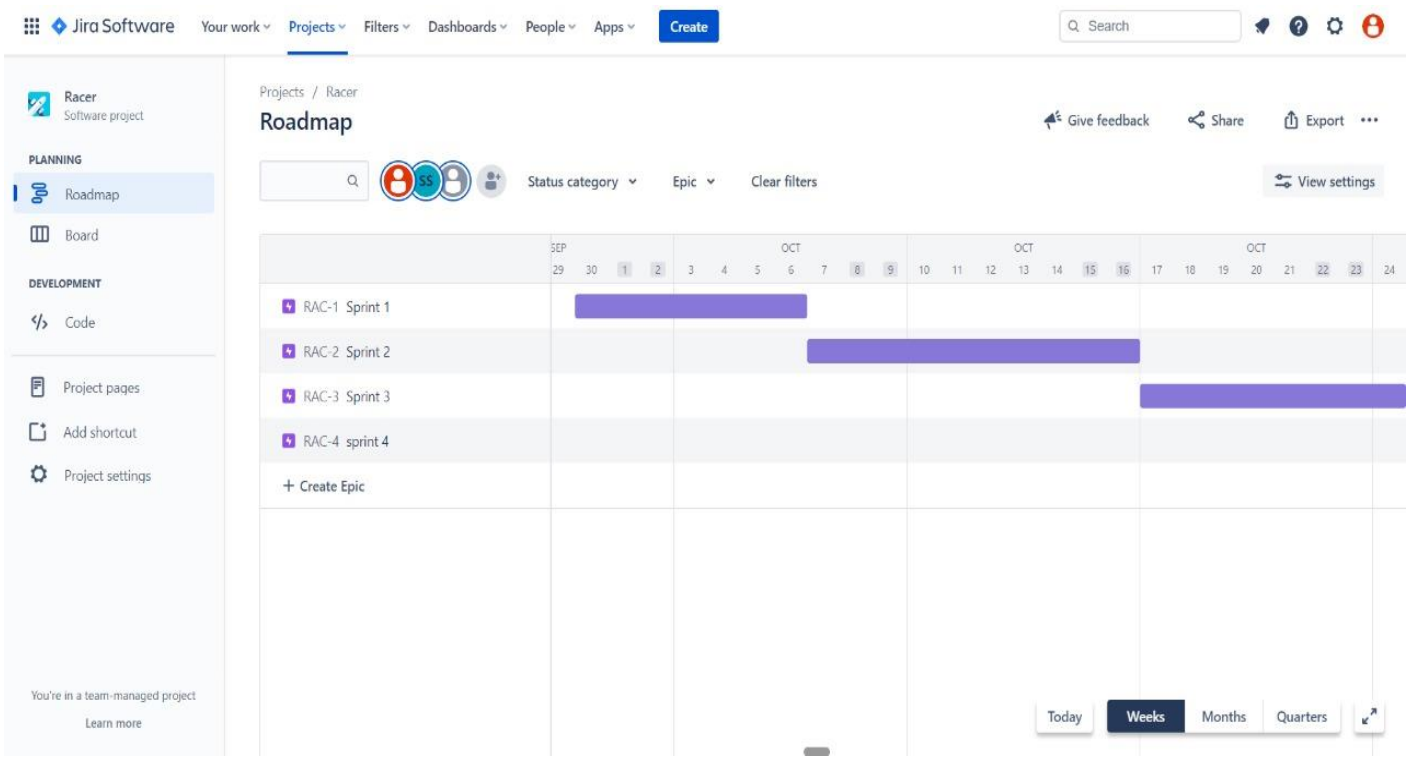
| Sprint | Functional Requirement (Epic) | User Story Number | User Story / Task | Story Points | Priority | Team Members |
|---|---|---|---|---|---|---|
| Sprint-1 | Registration | USN-1 | As a health care provider I can create account in IBM cloud and the data are collected. | 20 | High | Sweetha S<br>Swathi S |
| Sprint-2 | Analyze | USN-2 | As a health care provider all the data that are collected is cleaned and uploaded inthe database or IBM cloud. | 20 | Medium | Sruthi V G<br>Sanjika S |
| Sprint-3 | Dashboard | USN-3 | As a health care provider I can use my account in my dashboard for uploading dataset. | 10 | Medium | Sweetha S<br>Sanjika S |
| Sprint-3 | Visualization | USN-4 | As a health care provider I can prepare data for Visualization. | 10 | High | Swathi S<br>Sruthi V G |
| Sprint-4 | Visualization | USN-5 | As a health care provider I can present data in my dashboard. | 10 | High | Sweetha S<br>Sanjika S |
| Sprint-4 | Prediction | USN-6 | As a health care provider I can predict the length of stay | 10 | High | Sweetha S<br>Swathi S |

## 6.2 Sprint Delivery Schedule

Since sprints take place over a fixed period of time, it's critical to avoid wasting time during planning and development. And this is precisely where sprint scheduling enters the equation.In case you're unfamiliar, a sprint schedule is a document that outlines sprint planning from end to end. It's one of the first steps in the agile sprint planning process—and something that requires adequate research, planning, and communication. Sprint Delivery Schedule for the proposed solution**.**

| 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 |

## 6.3 Reports from JIRA



# CHAPTER-7

# CODING & SOLUTIONING

## 7.1 Feature

Excel worksheets come with a standard limit of 1,048,576 rows. While performancein Excel will slow well before the said row limit, it’s a common requirement for users to analyze datasets over one million rows in size. Congo’s Analytics compresses your data so you can extract insights from large datasets. With a well- built data model, Congo’s Analytics can help you analyze datasets containing over 100 million rows.

```
import pandas as pd
import numpy as np
 import matplotlib.pyplot as plt
from matplotlib.colors import ListedColormap
 import seaborn as sns from warningsimport
filterwarnings filterwarnings('ignore')
 pd.options.display.max_columns = None
 pd.options.display.max_rows = None
 pd.options.display.float_format = '{:.6f}'.format
 from sklearn.model_selection import train_test_split
import statsmodels
import statsmodels.api as sm
 from sklearn.preprocessing import StandardScaler
from sklearn import metrics
 from sklearn.linear_model import LogisticRegression
 from sklearn.metrics import classification_report
from sklearn.metrics import cohen_kappa_score
from sklearn.metrics import confusion_matrix
from sklearn.metrics import roc_curve
```

```
from sklearn.metrics import accuracy_score
 from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
 from sklearn import tree
 from sklearn.model_selection import GridSearchCV
 from sklearn.ensemble import
AdaBoostClassifier,GradientBoostingClassifier
from catboost import CatBoostClassifier
from sklearn.feature_selection import RFE
plt.rcParams['figure.figsize'] = [15,8]
```

**Feature 2**

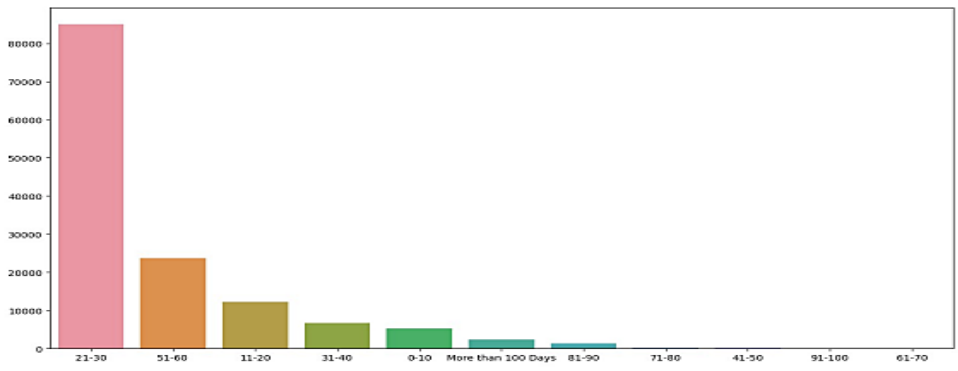
Always the user can't come to the Coding part and comes to get the results because the user don't know about the coding applications so the User builds need interactive website which user can easily catch the results in the application.

```
for i in [ 'Stay', 'Department', 'Available Extra Rooms in Hospital',
 'Ward_Type' ,'Ward_Facility_Code', 'Age', 'Type of Admission',
 'Severity of Illness', 'Bed Grade', 'Hospital_region_code',
 'Hospital_type_code' , 'City_Code_Hospital', 'Hospital_code',
 'City_Code_Patient', 'Visitors with Patient']:
 Count = train[i].value_counts()
 plt.figure(figsize=(10,5))
 sns.barplot(x=count.index.values, y=count.values,data=train)
 plt.xlabel(i)
 plt.ylabel('No of occurences')
Total = train.isnull().sum().sort_values(ascending=False)
Percent = (train.isnull().sum()*100/train.isnull().count()).sort_values(ascendin g=False)
 missing_data = pd.concat([Total, Percent], axis = 1, keys = ['Total', 'Percentage of Missing Values'])
print('X_train', X_train.shape)
print('y_train', y_train.shape)
print('X_test', X_test.shape)
 print('y_test', y_test.shape)
 X_train (219655, 16)
 y_train (219655,)
```

X_test (94138, 16)

y_test (94138,)

```
from sklearn.model_selection import KFold,cross_val_score
kfold=KFold(n_splits=10, shuffle=True, random_state=10)
 LR = LogisticRegression()
LR.fit(X_train,y_train
) y_pred_LR=LR.predict(X_test)
accuracy_score(y_test,y_pred_LR)*100 37.94217000573626
rf_classification = RandomForestClassifier(random_state = 10)
 rf_model = rf_classification.fit(X_train, y_train)
 y_pred_RF = rf_model.predict(X_test)
 accuracy_score(y_test,y_pred_RF)*100 38.315026875438186
print(classification_report(y_test,y_pred_RF))
rf_classification_tuned = RandomForestClassifier(criterion = 'gini',
 n_estimators = 47, random_state = 10)
 rf_model_tuned = rf_classification_tuned.fit(X_train, y_train)
y_pred_RF_tuned = rf_model_tuned.predict(X_test)
accuracy_score(y_test,y_pred_RF_tuned)*100 37.94323227602031
print(classification_report(y_test,y_pred_RF_tuned)
ls = df_full_test.columns.tolist()
in_data = df_full_test[ls]
 out_data = cb.predict (in_data)
test = pd.read_csv("D:/HealthCare/test_data.csv") submit = pd.DataFrame()
submit ['case_id'] = test['case_id'] submit ['Stay'] = out_data stay_decode = { 1 : '0-10', 2 : '11-20', 3 : '21-30', 4 :
'31-40', 5 : '41-50', 6 : '51-60', 7 : '61-70',8 : '71-80', 9 : '81- 90', 10 : '91-100', 11 : 'More than 100 Days' }
 submit ['Stay'] = submit ['Stay'].map(stay_decode)
submit.head(15)
```



# 8.TESTING

## 8.1 TEST CASE

| Test case ID | Feature Type | Test case ID | Feature Type | Test case ID | Expected Result | Actual Result | Status | Commnets | TC for Automation(Y/N) |
|---|---|---|---|---|---|---|---|---|---|
| Dashboard | Functional | LoginPage_TC_OO4 | Functional | https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.public_folders%2FHealth_Care%2FDepartment%2BPatient&closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false&shareMode=embedded&acti | Displaying the responsive Dashboard.A dashboard helps you to monitor events or activites at a glance by providing key insight and analysis about your data on one or more page or screens | Working as expected | Pass | NIL | Y |
| Story | Verify whether the story is functioned on the analytics dashboard | IBM COGONOS WITH WATSON and IBM cloud | 1. Enter the Analytics Home page<br>2. click the story option<br>3. View the ibm embedded story about the Hospital data in html page and after click the back to top button, go to the top of the main page. | "https://us3.ca.analytics.ibm.com/bi/?pathRef=.public_folders%2FHealth_Care%2FReport%2B1&closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false&shareMode=embe | A story is a type of view that contains a set of scenes that are displayed in sequence over time. | Working as expected | Pass | NIL | Y |
| Report | Verify user is able to view and run the reports | IBM COGONOS WITH WATSON and IBM cloud | 1. Enter the Analytics Home page<br>2. click the report option<br>3. View the ibm embedded report about the Hospital data in html page and after click the back to top button, go to the top of the main page. | https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.public_folders%2FHealth_Care%2Fstory&closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false& | explore data and create and developersuse to build sophisticated,multiple-page,multiple-query report against multiple database. | Working as expected | Pass | NIL | Y |

## 8.2 USER ACCEPTANCE TESTING

The purpose of this document is to briefly explain the test coverage and open issues of the project at the time of the release to User Acceptance Testing(UAT)

**Defect Analysis**

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

| 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'tFix | 0 | 0 | 0 | 0 | 0 |
| Totals | 24 | 18 | 11 | 5 | 58 |

**Model Performance Testing**

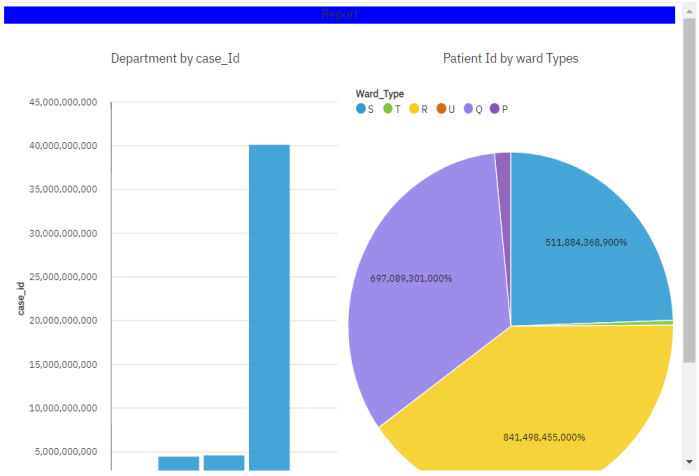
Project team shall fill the following information in model performance testing template.

| 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.RESULTS

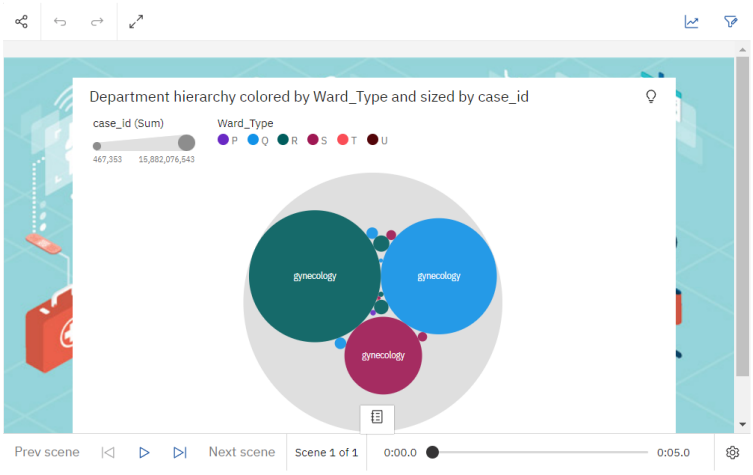
## 9.1 PERFROMANCE METRICS


## Report




## Story

## Story



# CHAPTER-10

## ADVANTAGES & DISADVANTAGES

### ADVANTAGES

By digitizing, combining and effectively using big data, healthcare organizations ranging from single-physician offices and multi-provider groups to large hospital networks and accountable care organizations stand to realize significant benefits.

Potential benefits include detecting diseases at earlier stages when they can be treated more easily and effectively; managing specific individual and population health and detecting health care fraud more quickly and efficiently.

Numerous questions can be addressed with big data analytics. Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data, such as length of stay (LOS)

McKinsey estimates that big data analytics can enable more than $300 billion in savings per year in U.S. healthcare, two thirds of that through reductions of approximately 8% in national healthcare expenditures.

### DISADVANTAGE

This technology needs to be extensively trained with curated data sets in order to perform as expected. However, the lag between data collection and processing has to be addressed

Since the healthcare industry is crucial for patient care, the medical community need proof that the important managerial issues of ownership, governance and standards have to be considered and woven through these issues are those of continuous data acquisition and data cleansing

Invasion of privacy is one of the major concerns of this technology. With companies delving into customer data until its very depths, maintaining privacy is next to impossible for the general public.

# CHAPTER-11

# CONCLUSION

In this project, different variables were analyzed that correlate with Length of Stay by using patient-level and hospital-level data. By predicting a patient's length of stay at thetime of admission helps hospitals to locate resources more efficiently and manage their patients more effectively. Identifying factors that associate with LOS to predict and manage the number of days patients stay, could help hospitals in managing resources and in the development of new treatment plans. Effective use of hospital resources and reducing the length of stay can reduce overall national medical expenses.

Big data analytics in healthcare is evolving into a promising field for providing insight from very large data sets and improving outcomes while reducing costs. Its potential is great; however there remain challenges to overcome. Big data analytics has the potential to transform the way healthcare providers use sophisticated technologies to gain insight from their clinical and other data repositories and make informed decisions. In the future we'll see the rapid, widespread implementation and use of big data analytics across the healthcare organization and the healthcare industry. To that end, the several challenges highlighted above, must be addressed. As big data analytics becomes more mainstream, issues such as guaranteeing privacy, safeguarding security, establishing standards and governance, and continually improving the tools and technologies will garner attention.

# CHAPTER-12

# FUTURE SCOPE

**Smart Staffing & Personnel Management:** having a large volume of quality data helps health care professionals in allocating resources efficiently. Healthcare professionals can analyze the outcomes of checkups among individuals in various demographic groups and determine what factors prevent individuals from seeking treatment.

**Advanced Risk & Disease Management:** Healthcare institutions can offer accurate, preventive care. Effectively decreasing hospital admissions by digging into insights such as drug type, conditions, and the duration of patient visits, among many others.

**Real-time Alerting:** Clinical Decision Support (CDS): applications in hospitals analyzes patient evidence on the spot, delivering recommendations to health professionals when they make prescriptive choices. However, to prevent unnecessary in-house procedures, physicians prefer people to stay away from hospitals

**Enhancing Patient Engagement:** Every step they take, heart rates, sleeping habits, can be tracked for potential patients (who use smart wearables). All this information can be correlated with other trackable data to identify potential health risks.

## CHAPTER -13
## APPENDIX

### 13.1 SOURCE CODE

```
<!DOCTYPE html>
<html lang="en">
<head>
  <title>Bootstrap Example</title>
  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <link rel="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>

<div class="jumbotron text-center">
  <h1>Analytics for health care-Data</h1>
</div>

<div class="container">
  <div class="row">
    <div class="col-sm-4">
      <h3>Patients</h3>
      <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.public_
folders%2FHealth_Care%2Fpatients&amp;closeWindowOnLastView=true&amp;ui_ap
pbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp
;mode=dashboard&amp;subView=model000001848c07e3e4_00000002" width="250"
height="250" frameborder="0" gesture="media" allow="encrypted-media"
```

```
allowfullscreen=""></iframe>    </div>
   <div class="col-sm-4">
    <h3>wards</h3>
    <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.public_
folders%2FHealth_Care%2Fwards&amp;closeWindowOnLastView=true&amp;ui_appb
ar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;m
ode=dashboard&amp;subView=model000001848c040509_00000002" width="250"
height="250" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>    </div>
   <div class="col-sm-4">
    <h3>Cases</h3>
    <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.public_
folders%2FHealth_Care%2Fcases&amp;closeWindowOnLastView=true&amp;ui_appba
r=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mo
de=dashboard&amp;subView=model000001848c019a6e_00000000" width="250"
height="250" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>
   </div>
  </div>
  <h3>Dashboard</h3>
  <br>
  <p>
   <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.public_
folders%2FHealth_Care%2FDepartment%2BPatient&amp;closeWindowOnLastView=t
rue&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;ac
tion=view&amp;mode=dashboard&amp;subView=model000001848b4421a7_00000002"
width="800" height="500" frameborder="0" gesture="media" allow="encrypted-
media" allowfullscreen=""></iframe>
  </p>
  <br>
  <h3>Report</h3>
```

```
<br>
<p>
<iframe
src="https://us3.ca.analytics.ibm.com/bi/?pathRef=.public_folders%2FHealth_Care%2FReport%2B1&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=run&amp;format=HTML&amp;prompt=false" width="800" height="500" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe> </p>
<br>
<h3>Story</h3>
<br>
<p>
<iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.public_folders%2FHealth_Care%2Fstory&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;sceneId=model000001848d38ab03_00000001&amp;sceneTime=0" width="800" height="500" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
</p>
</div>

</body>
</html>
```

## 13.2 GITHUB & PROJECT DEMO LINK

**GITHUB ID:https://github.com/IBM-EPBL/IBM-Project-18860-1659690812**

**PROJECT DEMO LINK:https://drive.google.com/file/d/1uNt6N5Y17g755rPc6UZuiGho5_4jTUb-/view?usp=sharing**