ABSTRACT

The aim is to develop an web application 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. In recent times due to the pandemic many people suffered to lack of sufficient facilities. The health-care facilities also suffered due to the rise in the number of cases. Due to the rise in the number of cases the patients were unable to find the hospital bed to get the required treatment. The most important thing that was lacking is that analysis for health-care facilities, it is a very tedious task to maintain the registration in a manual workbook for the availability and to check the patient health record. To optimize it and make it more understandable, digital data has to be collected and it should be made available for the respective persons and public to check the availability or services they are looking for. To carry out a real-time data analysis, we are proposing the web app with intriguing dashboard and analysis on the patient data. At the time of admission, it helps the hospitals in identifying patients who are at high LOS-risk (patients who will stay longer). Once identified, patients at high risk for LOS can have their treatment plans improved to reduce LOS and reduce the risk of infection in staff or visitors. It gives us a full-fledged analysis on the hospital data regarding the patient health conditions and case study analysis, which can result in the enormous amount of growth in the health-care industry to analyse the health-care data. IBM Cognos analytical tool has been used to make the analysis and create an interactive dashboard to make the real-time analysis on the incoming data from the database. The scope of this project is to make the intuitive dashboard, report and story to present it to the user's perspective to make them understand in better manner. Apart from the length of stay for each patient, it is used to analyse the various fields related to health using various visualization plots.

CHAPTER-1

INTRODUCTION

1.1 PROJECT OVERVIEW

We have helping 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. Our prior knowledge of LOS can aid in such as logistics room and bed allocation planning.

1.2 PURPOSE

The aim is to develop an web application accurately predict the length of stay for each patient on case by case basis so that hospitals can use this information for optimal resource allocation and better functioning. In recent times due to the pandemic many people suffered to lack of sufficient facilities. The health-care facilities also suffered due to the rise in the number of cases. Due to the rise in the number of cases the patients were unable to find the hospital bed to get the required treatment. The most important things that was lacking is that analysis for health-care facilities, it is a very tedious manual workbook for the availability and to check the patient health record. To optimize it and make it more understandable, digital data has to be collected and it should be made available for the respective persons and public to check the availability or services they looking for. To 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.

CHAPTER-2

LITERATURE SURVEY

Big Data Analytics in Healthcare at Maharaja Yeshwantrao Hospital

Author: Mimoh Ojha[1], Dr.Kirti Mathur[2]

Year:2016

This paper focuses on utilizing the big data characteristics to keep track and make use of the Hospital data of every patients and to improve the healthcare domain. It was done at the Maharaja Yeshwantrao hospital which is located at the Maharastra. It is central india one of the largest Government hospital. It is intended to digitalize the patient record and make it into the EHR [Electronic Health Record]. The hospital generates enormous amount of data in the form structures, semi-structured and unstructured. It makes use of Hadoop distributed system to process and analyse the data. By analysing the incoming data such as patient's health records, laboratory test result, electronic medical equipment, health insurance data, social media, drug research, genome research, clinical outcome, transaction and from Mahatma Gandhi Memorial medical college which is under MY hospital.

Advantages:

It is used to convert the paperwork and tedious task into paperless digital-format. Data analytics at the hospital will provide insights and benefits in terms of money saving and doctor's time as well. It can resist up to high level of data storage.

Disadvantages:

The main disadvantage of this system is, that it does not have a cloud storing facility, irrespective of the data collected and stored in it, it could not be processed from anywhere at time of need, so as of now it is just stored in the single system.

BIG DATA ANALYTICS IN HEALTHCARE

Author: Nkemakolam Chinenye Onyemachi[1,] Ogwueleka Francsiska Nonyelum[2]

Year: 2019

The amount of data being generated in the healthcare industry is growing at a very fast rate. This has generated immense interest in leveraging the availability of healthcare data to improve health outcomes and reduce costs. Big data analytics has earned a remarkable interest in the health sector as it could be used in the diagnosis and prediction of diseases. The goal is to predict the epidemic weeks in advance using the geo-map to outburst the plague or virus in the environment. The second method is using a tool called Resistance Open which is used to discover the immunities that has been present in the patients naturally. The third method is to monitor the ratio and spreaders of the disease among the surroundings. Data mining can help health care insurance organizations to detect hypocrisy and misuse, health care institutions to make decisions of customer relationship management, providers to identify effective treatments and best practices and patients now receive enhanced and more economical health care services

Advantages:

Different kinds of methods were presented and explained to do the classification, prediction and analysis on the big-data. It states the role of data analytics in the healthcare sector by make use of the Big data tool and Classification techniques.

A RESEARCH ON BIG DATA ANALYTICS IN HEALTHCARE INDUSTRY

Author: Mohaiminul Islam, Rezaul Karim, Shamim Reza, MST Asha Khatum

Year:2020

In this article, it is solely describing to show the need of the data analytics advantages in the health-care industry. Data analytics has been a very effective tool to analyse and make predictions on the patient data, that is used for further analysis and make sure to prevent the plague or epidemic outburst into the nature. Before the era of data analytics, lots of lots data was loading over the hospital paperwork but the none of the paperwork comes handy when it comes to analysis. But storing the data in the computers and perform data analysis on it by using the data analytical tool. The best finding is that can give insights and the early indications of the disease can make better and lifesaving. The need of the data analysis led to the invention of two amazing tool such as SAS (using HANA) and Dell. One of the functions of the SAS HANA tool is to analyse data from its cancer and medical admin databases to find the clinical trial for the most suited patients. It provides accurate and quick analytics regardless of the structured or unstructured data.

Advantages:

Analytical tool has its own impact on the health-care industry for finding insights from the structured or unstructured data.

A REVIEW OF ANALYTICS AND CLINICAL INFORMATICS IN HEALTH CARE

Author: Allan F Simpao,Luis M Ahumada,Jorge Galvez,Mohamed A Rehman

YEAR: 2014

Federal investment in health information technology has incentivized the adoption of electronic health record systems by physicians and health care organizations; the result has been a massive rise in the collection of patient data in electronic form (i.e. "Big Data"). Health care systems have leveraged Big Data for quality and performance improvements using analytics-the systematic use of data combined with quantitative as well as qualitative analysis to make decisions. Analytics have been utilized in various aspects of health care including predictive risk assessment, clinical decision support, home health monitoring, finance, and resource allocation. Visual analytics is one example of an analytics technique with an array of health care and research applications that are well described in the literature. The proliferation of Big Data and analytics in health care has spawned a growing demand for clinical informatics professionals who can bridge the gap between the medical and information sciences.

CHARACTERIZING MAMMOGRAPHY REPORTS FOR HEALTH ANALYTICS

Author: Carlos C Rojas, Robert Patton, Barbara G Beckerman

Year: 2011

As massive collections of digital health data are becoming available, the opportunities for large-scale automated analysis increase. In particular, the widespread collection of detailed health information is expected to help realize a vision of evidence-based public health and patient-centric health care. Within such a framework for large scale health analytics we describe the transformation of a large data set of mostly unlabeled and free-text mammography data into a searchable and accessible collection, usable for analytics. This paper also describes several methods to characterize and analyze the data, including their temporal aspects, using information retrieval, supervised learning, and classical statistical techniques. This paper presents experimental results that demonstrate the validity and usefulness of the approach, since the results are consistent with the known features of the data, provide novel insights about it, and can be used in specific applications. Additionally, based on the process of going from raw data to results from analysis, this paper presents the architecture of a generic system for health analytics from clinical notes.

Advantage:

In this paper the experimental results demonstrate the validity and the usefulness of the approach, since they both conformed to what was expected from the data and helped to get novel insights about it.

Disadvantage:

Since this paper has discussed a notion of time and space for the patient data, questions about the rate of change and the direction of the trajectory naturally arise which makes it complex.

DEVELOPMENT OF THE HEALTH INFORMATION ANALYTICS DASHBOARD USING BIG DATA ANALYTICS

Author: Anisatul Afifah, Krisostomus Nova Rahmanto

Year: 2020

This paper states about the development of digital technology that has an impact on healthcare facilities in Indonesia, one of which is the digitization of medical records. This will generate abundant clinical data from various sources including electronic medical records. Therefore, a large infrastructure is needed to store data from various sources that can facilitate the process of data aggregation to then be processed into information. Health Information Analytics Dashboard is the solution to get accurate, complete, and real-time insight from big data in healthcare. Data collection is carried out from various sources of health service facilities in Indonesia that are integrated into the system. With a user-friendly display, the analytic dashboard can be used to create monitoring reports with just one click. The method of this study uses big data analytics. The data analysis results are visualized through display charts/graphs that make it easier for users to understand the data analysis results and interpretation. This dashboard is useful to facilitate decision making so that stakeholders can find out more quickly to be able to respond appropriately and also improve the quality of health services so as to improve the degree of public health.

CURRENT PRACTICES IN CLINICAL ANALYTICS: A HOSPITAL SURVEY REPORT

Author: Dana Womack, Rosemary Kennedy, Bill Bria

Year: 2012

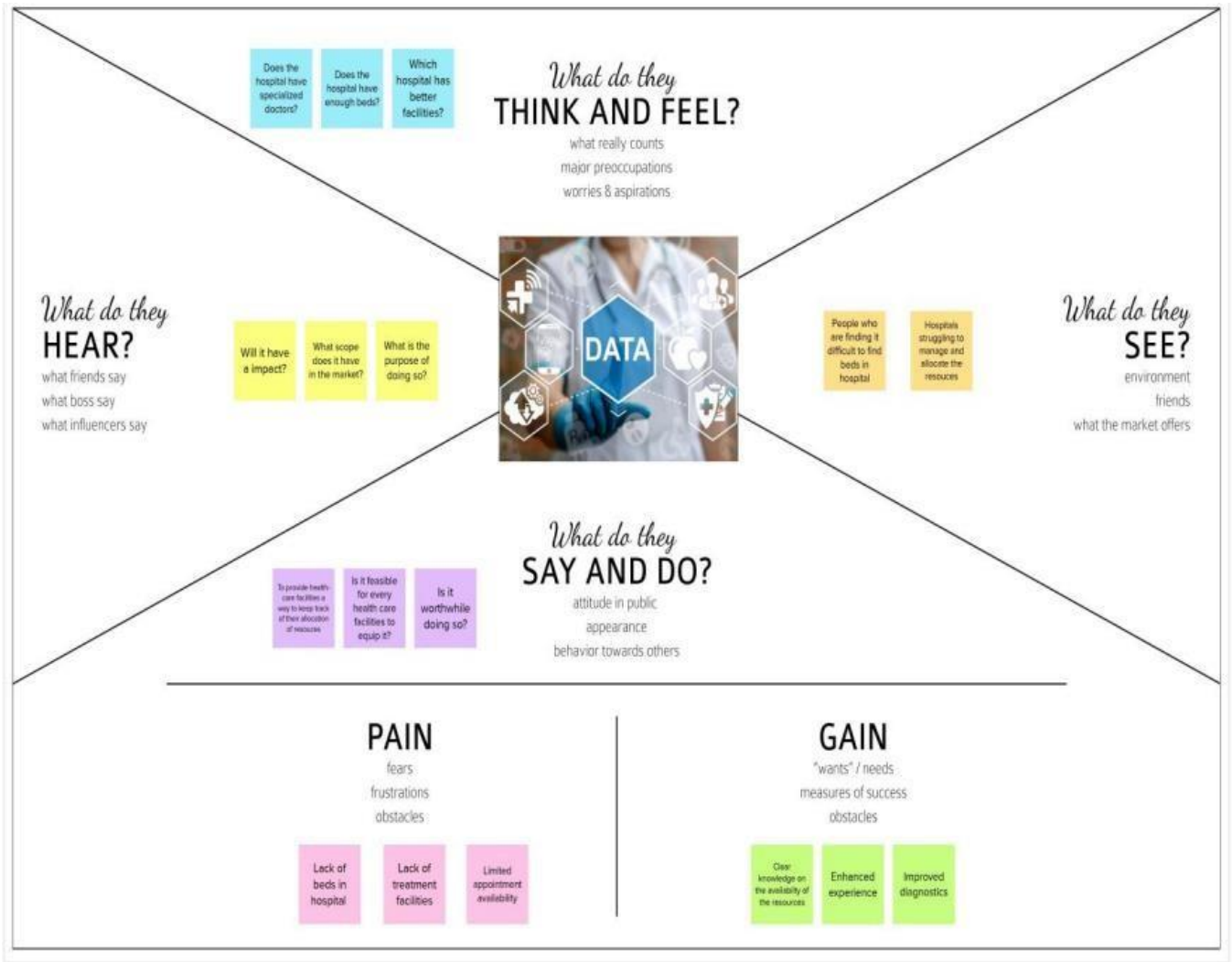
This paper is mainly based on the clinical analytics which must become a pervasive activity in healthcare settings to achieve the global vision for timely, effective, equitable, and excellent care. Global adoption of the Electronic Health Record (EHR) has increased the volume of data available for performance measurement and healthcare organizational capacity for continuous quality improvement. However, EHR adoption does not automatically result in optimal use of clinical data for performance improvement. In order to understand organizational factors related to use of data for clinical analytics, a survey was conducted of hospitals and hospital-based clinics. The survey revealed sub-optimal use of data captured as a byproduct of care delivery, the need for tools and methodologies to assist with data analytics, and the need for disciplined organizational structure and strategies. Informatics nurse professionals are well-positioned to lead analytical efforts and serve as a catalyst in their facility's transformations into a data-driven organization.

CHAPTER-3

IDEATION & PROPOSED SYSTEM

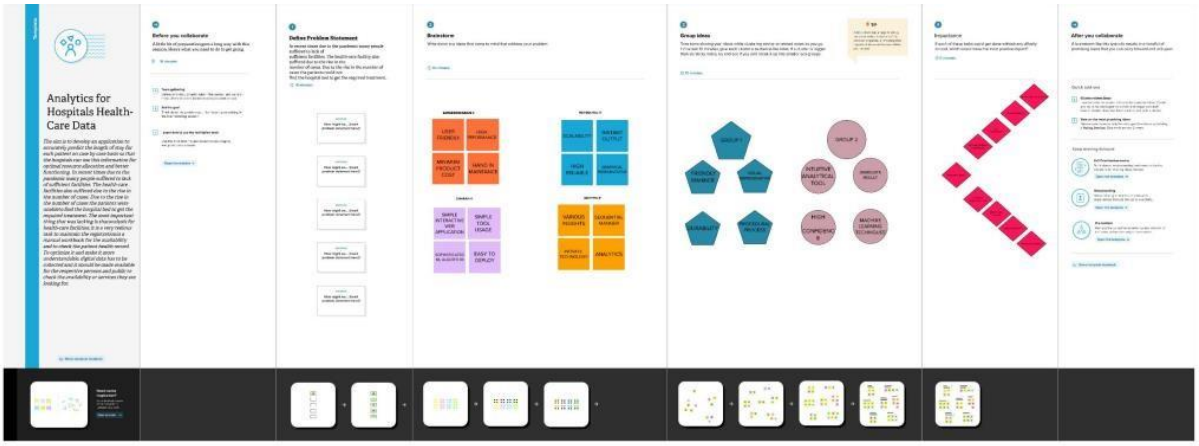
3.1 EMPATHY MAP CANVAS

Teams can utilize an empathy map as collaborative tool to learn more about their clients. An empathy map can depict a group of users, such as a Consumer segment, in a manner similar to user personas. The agile community has embraced the empathy map.



3.2 IDEATION & BRAINSTORMING

Ideation and the pratise of brainstorming, a particular method for coming up with fresh ideas, are frequently closely related. The main distinction between ideation and brainstorming is that whereas brainstorming is nearly often done in groups, ideation is typically seen as being more of a solitary endeavor.



3.3 PROPOSED SOLUTION

PROBLEM STATEMENT:

We propose a solution to analyse the healthcare data and derive the useful insights to patient and doctor to help them in a friendly manner. In hospitals, it is very necessary to keep track of patient data for future references and treatment.

IDEA/SOLUTION DESCRIPTION

We propose a solution to build a simple web application which takes input as patient-data and returns us output with the useful analysis with the help of the IBM Cognos analytical tool. The results will be displayed to the end user in the form of interactive dashboard, story and in report format.

NOVELTY / UNIQUENESS

The innovative and additional perk to make this solution stronger and the results more reliable, we use machine learning algorithms to develop a predictive analysis model which will be used to make predictions either on the patient health-status or the necessary input data. Along with the dashboard, prediction for these results will be shown in the user friendly-manner.

SOCIAL IMPACT / CUSTOMER SATISFACTION

The solution can never go unnoticed, though it is new to the society, because it is in a proactive way of analysis and prediction. It will address the concern of the key stakeholders, so it will create the impact in the customer as well as the social side. It is new to the hospital environment, there has not been any kind of project are implemented in the recent years.

BUSINESS MODEL (REVENUE MODEL)

The take-away of this project in a business scope of manner is mean to be plenty, it can be beneficial for the users (Patients and Doctors) more intriguing way. It is in need for the community of people, where it comes to handy in day-to-day life. It is a part of the live saving analysis and insights. Doctor is used to treat the patients and keep their medical history in hands for future references.

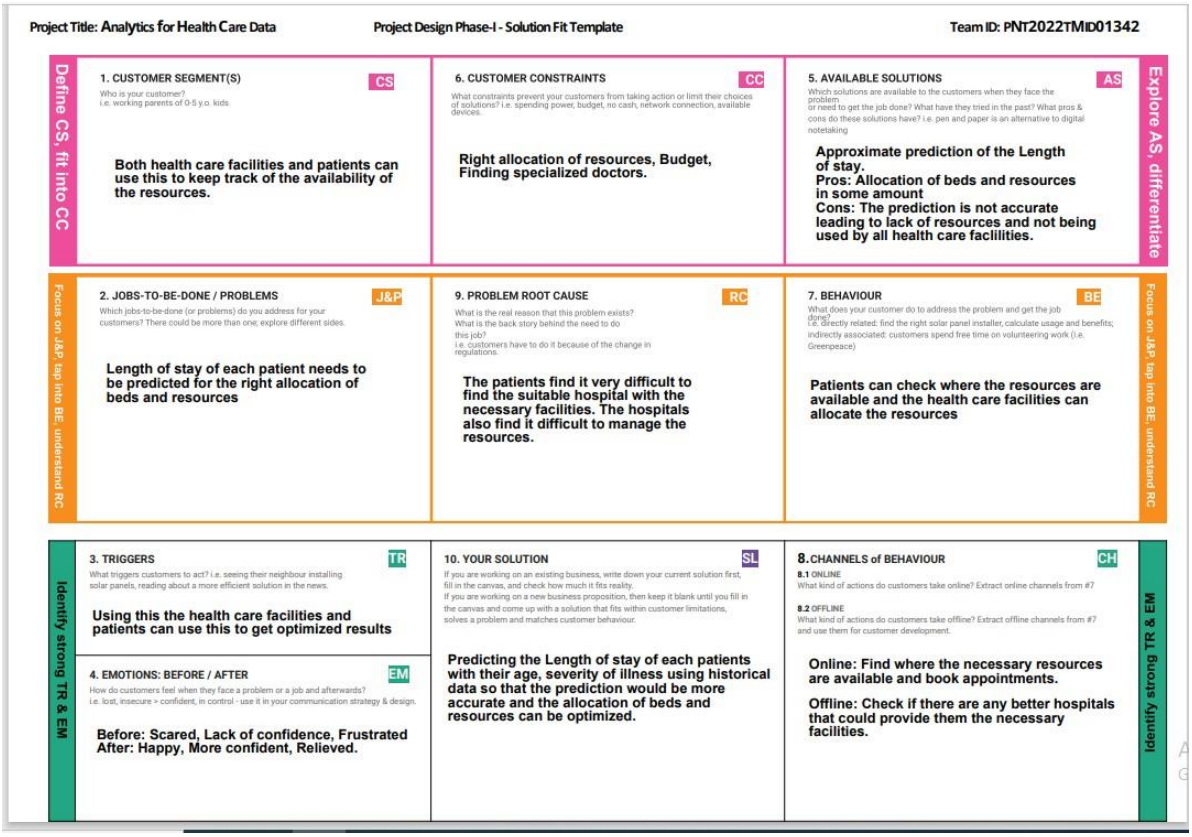
SCALABILITY OF THE SOLUTION

Scalability is the measure of the system performance against the increase or decrease in user demand. The system can handle the user request and return the results on time. It does not require much of the Graphical processor unit; it can be even run on the system with the minimum graphics or system with integrated graphics

3.4 PROBLEM SOLUTION FIT

The term "problem-solution fit" simply refers to the fact that you have identified a problem with a client and that the solution you have developed to address it truly resolves the client's issue.

1. Minimum Viable Product (MVP).
2. Early adopters who use your MVP and are pleased with it.
3. A verified issue you resolve for early-problems.



CHAPTER-4

REQUIREMENT ANALYSIS

4.1 FUNCTION REQUIREMENT

Functional requirements specify what a system should be able to do through calculations, technical details, data manipulation and processing, and other specific functionality. Use cases, which are used to represent behavioural requirements, describe all the instances in which the system makes use of the 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**	The User has his/her own ID to get registered in theportal
FR-2	**User Confirmation**	Confirmation via Email Confirmation via OTP
FR-3	**Dashboard**	The collected data are found in visualized format andthe prior data are analyzed.
FR-4	**Dataset**	The patients record and staffs record are collectedand consolidated as dataset
FR-5	**Report Generator**	The periodic reports of patients and the LoS are reported
FR-4	**Exploration**	The data exploration on available dataset

4.2 NON-FUNCTION REQUIREMNT

System qualities including security, reliability, performance, maintainability, scalability, and usability are defined by non-functional requirements (NFRs). They act as limitations or restrictions on how the system is designed for the various backlogs.

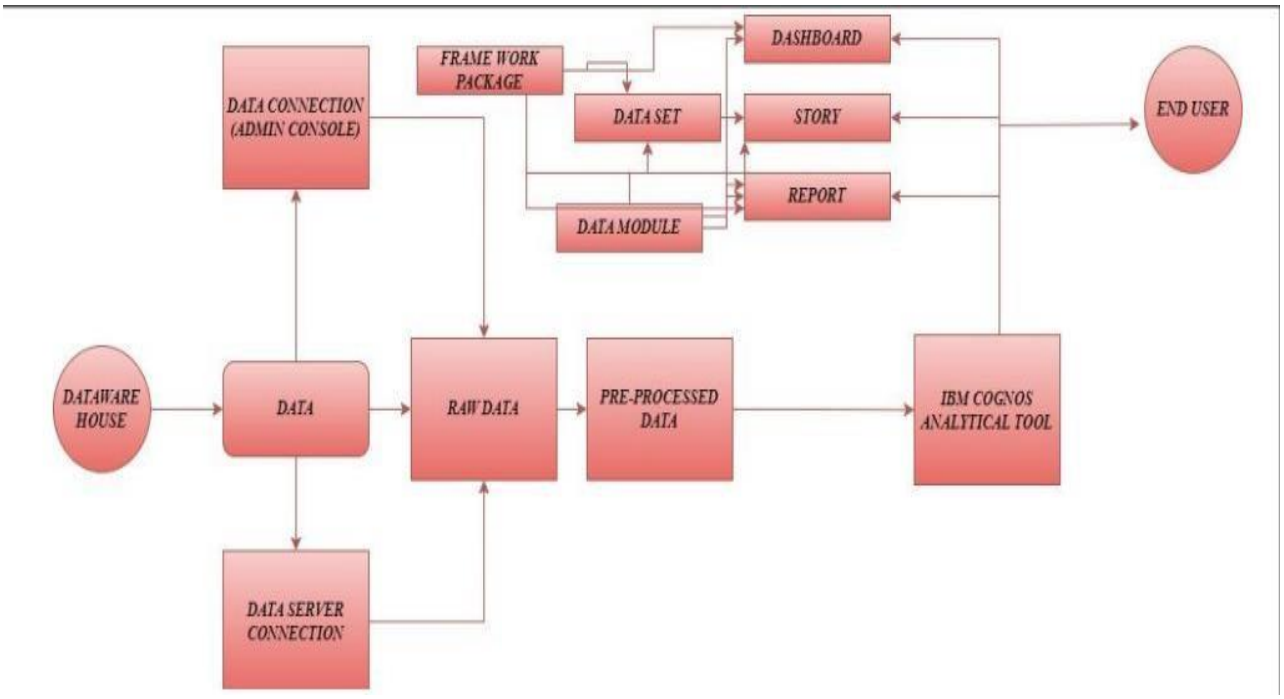
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	No prior experience required to use the dashboard. People with basic understanding can use the system.
NFR-2	**Security**	Only registered user can use this application.
NFR-3	**Reliability**	The Analytics system ensures the reliability
NFR-4	**Performance**	Gets updated regularly to improve theperformance of the application.
NFR-5	**Availability**	The availability of dataset must be constrained foraccurate data
NFR-6	**Scalability**	Any kind of data can be explored and the system is quiet expandable

CHAPTER-5

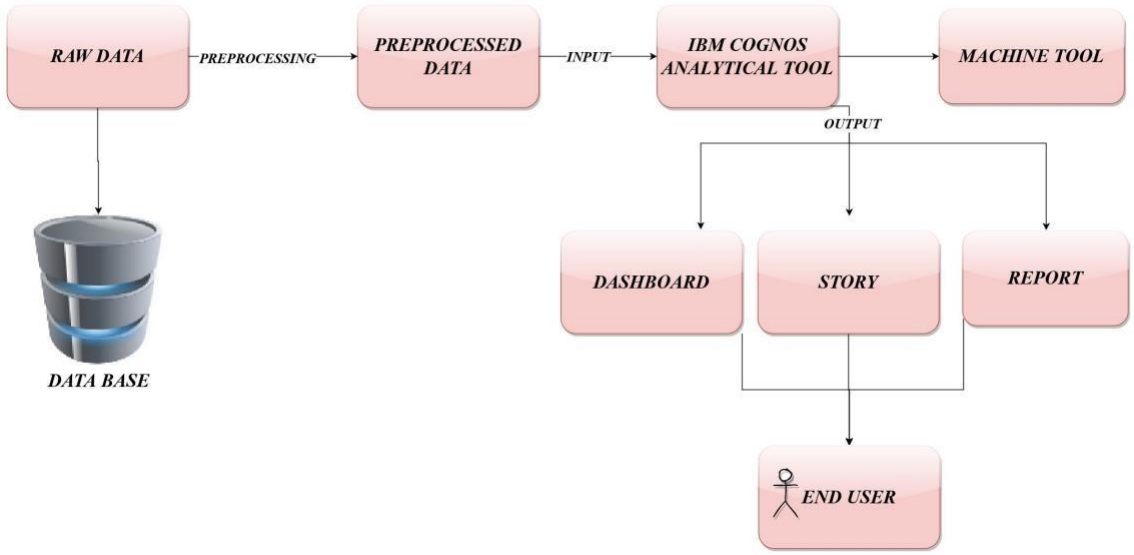
5.1 DATA FLOW DIAGRAM

Data flow diagrams are typically used by IT and engineering teams to show the flow of information, source of data inputs, and how that data is stored. These visual representations of a system can help be used to explain complex processes to key stakeholders or to build out new structures with your team. Data flow diagrams visualize relationships between external entities, processes, data stores, and data flows. You can visualize data flows with both parallel and asynchronous behaviours using our data flow diagram template.



5.2 SOLUTION & TECHNICAL ARCHITECTURE

An architectural description of a particular solution is called a solution architecture (SA). SAs combine recommendations from the enterprise solution architecture and various enterprise architecture perspectives (business, information, and technical).



User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Patient	Hospitalization	USN-1	Patients are required to hospitalize if they have any medical issues	Direct Hospitalization	High	Sprint-1
	Treatment Report	USN-2	Patients should collect them treatment report and get further doctor consult	They can receive the report from hospital	High	Sprint-1
Hospital Management	Resource Allocation	USN-3	Hospital Management should allocate the Necessary resource for treating the Patients	Should be ready for any circumstance	High	Sprint-2
	Predicting Length of Stay	USN-4	The Doctors should be awareof condition of Patients to predict the LoS	Exploring the data about the patient health condition and predicting LoS	High	Sprint-1
	Resource Availability	USN-5	The Hospital Staff should be aware of available resources in hospital	Visualizing the about the resource availability	High	Sprint-1
	Staff Welfare	USN-6	The working staff should be safe and Conscious about the COVID			

6. PROJECT PLANNING & SCHEDULING

Sprint Planning & Estimation:

The project has been divided into four sub-modules namely sprint.

Sprint1 - Data Collection and Preprocessing

Sprint2 - Data visualization

Sprint3 - Dashboards

Sprint4 - Report & Story

6.1 Sprint Planning

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection	USN-1	As a user, I need to gather the data in the CSV form	2	High	R ADITHYA
Sprint-1	Upload Dataset	USN-2	As a user , I can view the dataset	1	Low	R ADITHYA
Sprint-1	Data Preparation	USN-3	As a user, I can easily visualize the data	2	High	R ADITHYA
Sprint-2	Data Visualization	USN-4	As a user, I can visualize the charts	6	Medium	K CHARAN
Sprint-3	Dashboard	USN-5	As a user I can summarize the dashboard	4	Medium	K CHARAN
Sprint-3	Dashboard	USN-6	I must be able to gather insights from the dashboard	4	Medium	C LINGESHWARAN
Sprint - 4	Report	USN-7	As a user, I can view the full Report of analysis	5	High	V MOHANRAJ
Sprint-4	Story	USN-8	As a user, I can view the health care data in the form of story	5	High	V MOHANRAJ

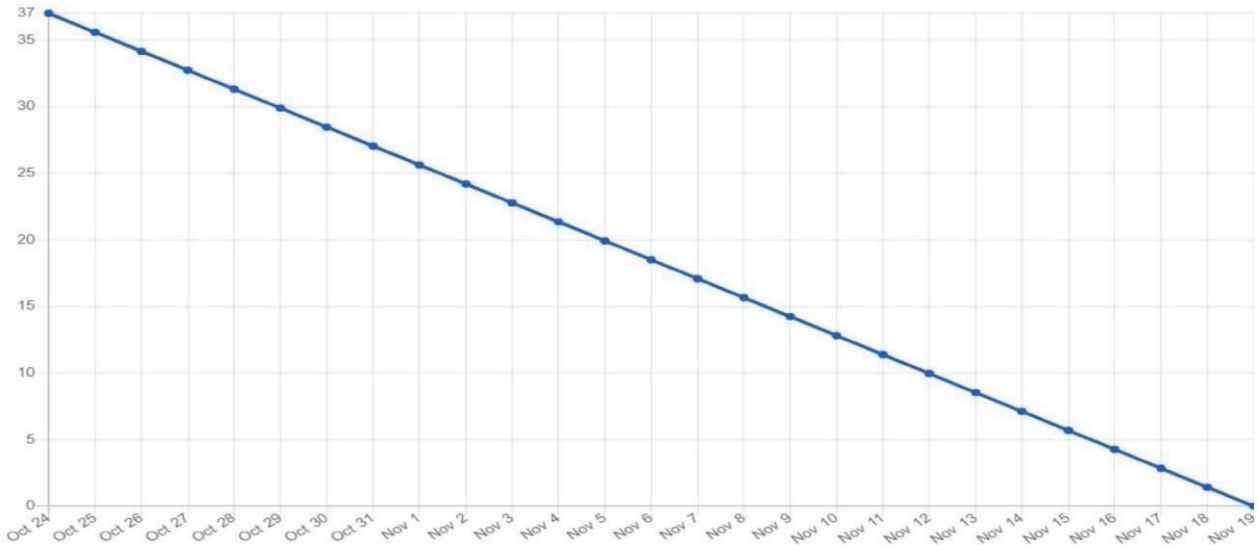
6.2 Sprint Delivery Schedule

Sprints	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	5	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	6	05 Oct 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	8	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	10	19 Nov 2022

6.3 Reports from JIRA:

BurnDownChart:

A burndown chart is used to keep of track of amount of work completed in the epic or sprint and the to-do work , that are yet to be finished.

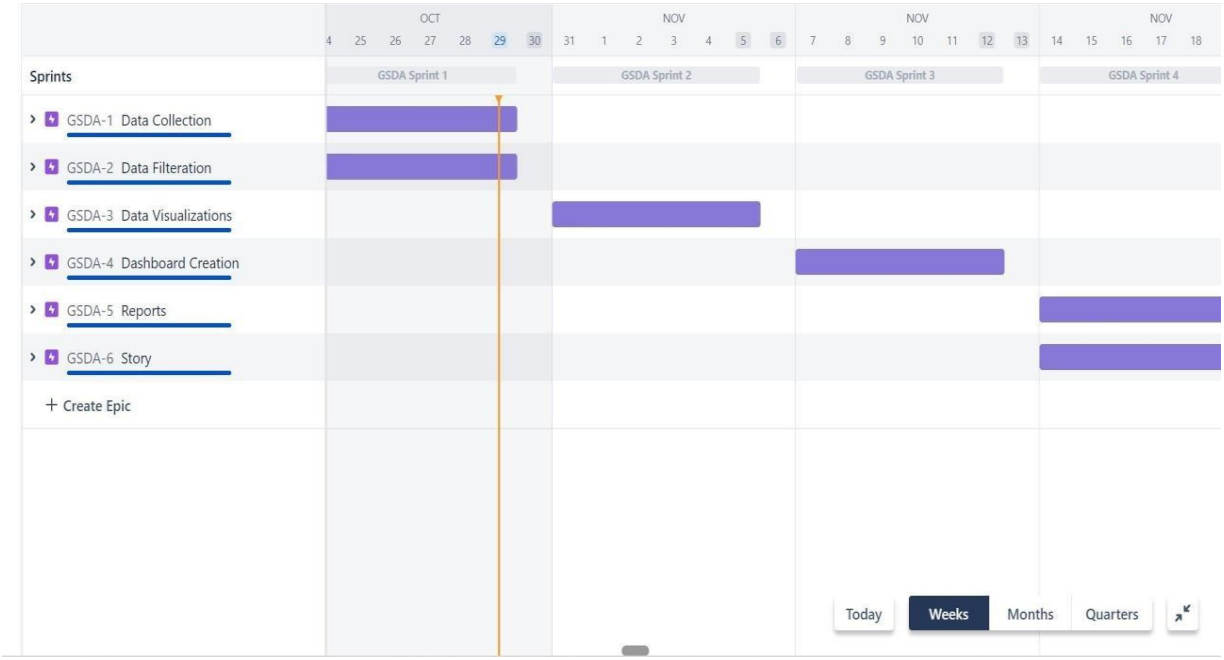


JIRA File:

Jira software is embedded into the software development cycle , it is used to design the issues for the project and to keep check in about the task and the bugs to be resolve.

It focuses mainly on the Agile Project management.

User stories can be developed and monitored until the desired output for the project has been achieved.


OCT
NOV
NOV
NOV
4 25 26 27 28 29 30 31 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Sprints
GSDA Sprint 1
GSDA Sprint 2
GSDA Sprint 3
GSDA Sprint 4
GSDA-1 Data Collection
GSDA-2 Data Filteration
GSDA-3 Data Visualizations
GSDA-4 Dashboard Creation
GSDA-5 Reports
GSDA-6 Story
Create Epic
Today
Weeks
Months
Quarters

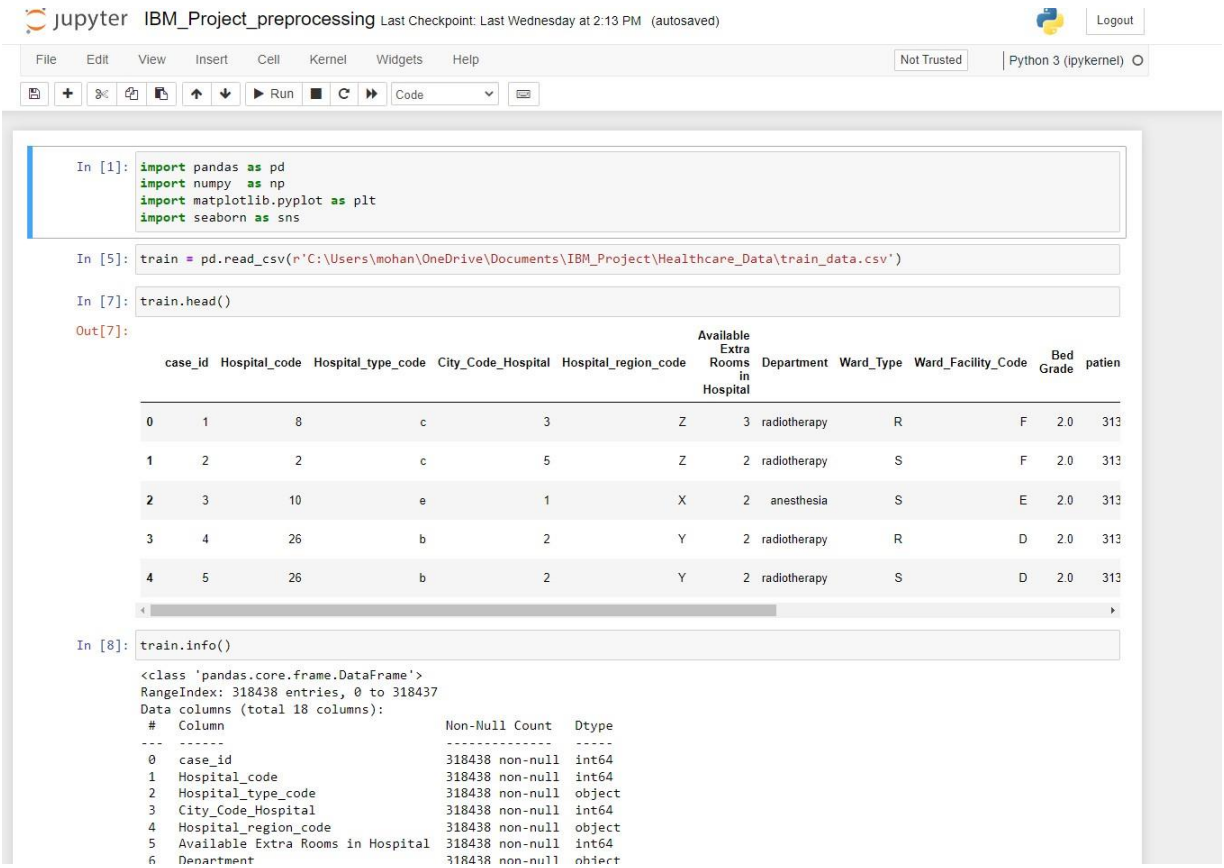
7. CODING & SOLUTIONING

7.1 Data Preparation:

The Data preparation phase starts from, collecting the input data and fetch it into the IBM Cognos analytical Tool. Before feed the input data into the Cognos tool. The data has been imported into the Jupyter Notebook to check for any missing values and needs to be drop any uncorrelated features in the dataset.

Jupyter notebook to drop the null values:

Python libraries are used to import the csv file and the null values are dropped by using the dropna() method.



```
Stay                                0
dtype: int64
```

```
In [11]: train['Bed Grade'].value_counts()
```

```
Out[11]: 2.0    123671
         3.0    110583
         4.0     57566
         1.0     26505
         Name: Bed Grade, dtype: int64
```

```
In [12]: train['Bed Grade'].unique()
```

```
Out[12]: array([ 2.,  3.,  4.,  1., nan])
```

```
In [13]: train.shape
```

```
Out[13]: (318438, 18)
```

```
In [14]: train.dropna(inplace=True)
```

```
In [15]: train.shape
```

```
Out[15]: (313793, 18)
```

```
In [16]: train.head()
```

Out[16]:

	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	patien
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0	313
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0	313
2	3	10	e	1	X	2	anesthesia	S	E	2.0	313
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0	313

```
In [17]: train.isnull().sum()
```

```
Out[17]: case_id                            0
         Hospital_code                      0
         Hospital_type_code                 0
         City_Code_Hospital                 0
         Hospital_region_code               0
         Available Extra Rooms in Hospital  0
         Department                         0
         Ward_Type                          0
         Ward_Facility_Code                 0
         Bed Grade                          0
         patientid                          0
         City_Code_Patient                  0
         Type of Admission                  0
         Severity of Illness                0
         Visitors with Patient              0
         Age                                0
         Admission_Deposit                  0
         Stay                               0
         dtype: int64
```

```
In [18]: train = train.to_csv('final_train_data.csv', index=False)
```

```
In [19]: test = pd.read_csv(r'C:\Users\mohan\OneDrive\Documents\IBM_Project\Healthcare_Data\test_data.csv')
```

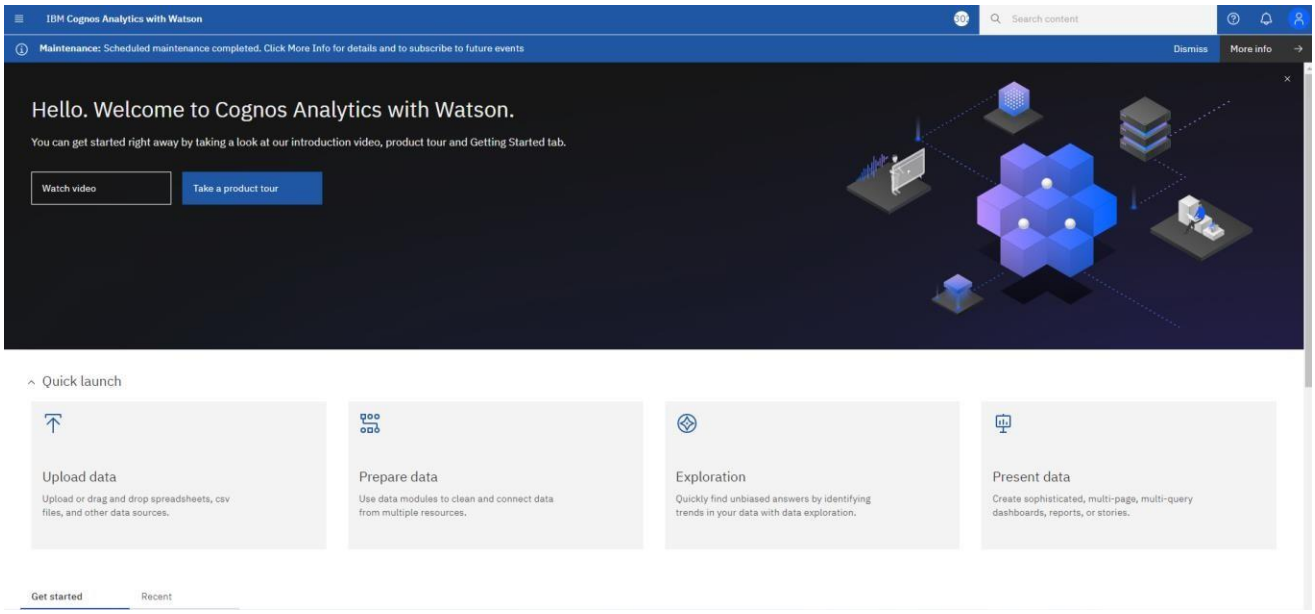
```
In [20]: test.head()
```

Out[20]:

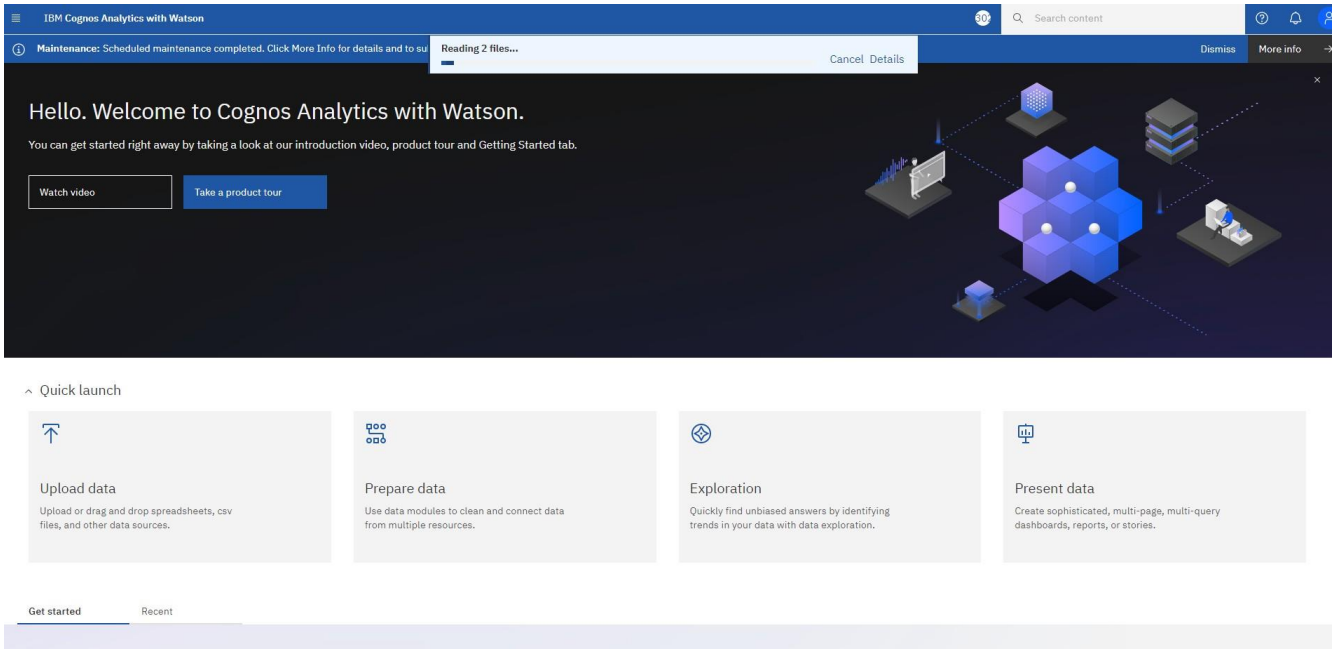
	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	patien
0	318439	21	c	3	Z	3	gynecology	S	A	2.0	170
1	318440	29	a	4	X	2	gynecology	S	F	2.0	170

IBM Cognos Tool:

It is an analytical tool , which was developed by the IBM community. In this project input data are processed and the Visualizations, Dashboards, Story and Reports are published by using the Cognos Tool.



Uploading the CSV files into the Cognos Tool:



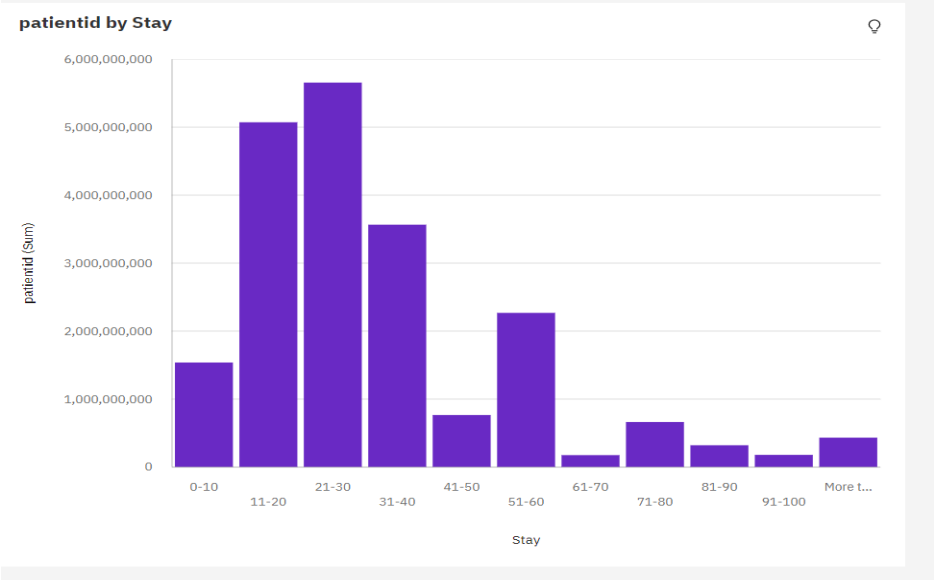
7.2 Visualizations:

The main part of the usage of this tool is to make visualizations using the input data for extracting insights out of it.

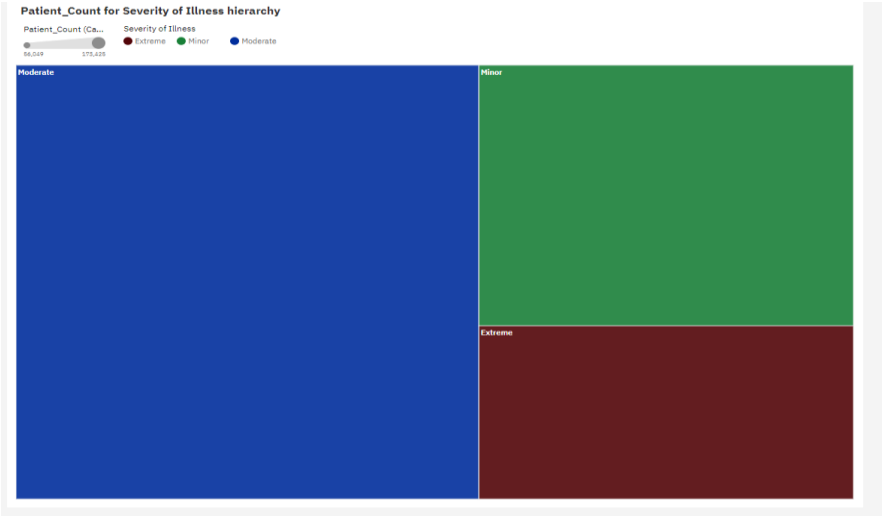
The mandatory visualizations which are mentioned in the IBM project place has been done first:

- Stay by patient ID using Column chart

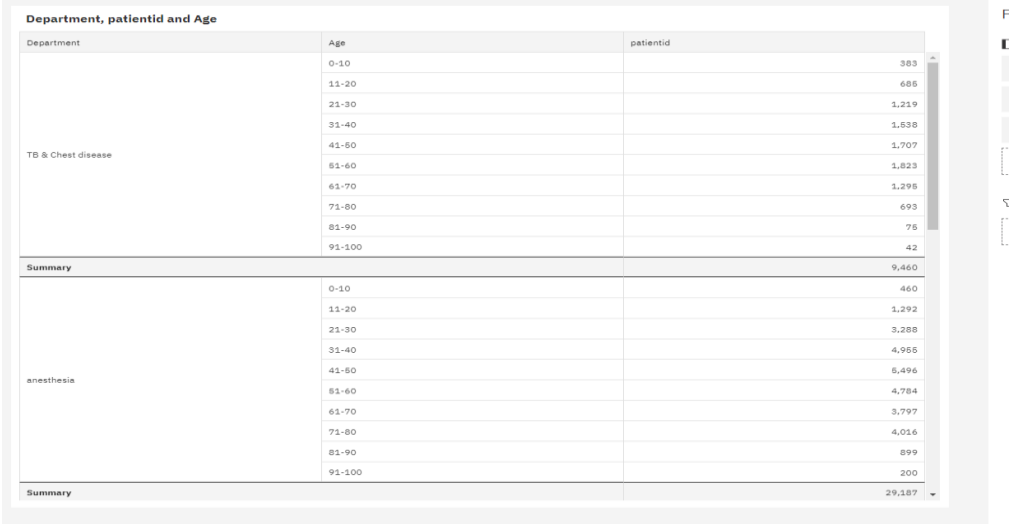
It is used to visualize the stay by patient ID in the columns chart. It describes the length of the stay



- Severity of illness by patient-Id using Tree map
  The severity of illness by count of the patient will be classified into either minor, extreme or moderate.

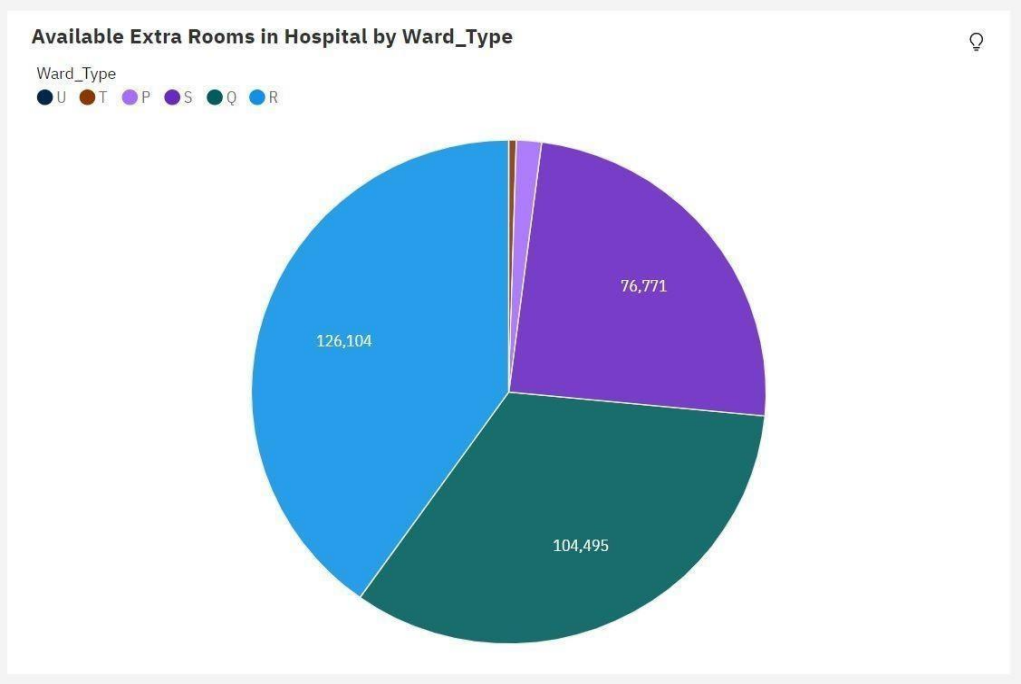


- Age, department wise patient using table

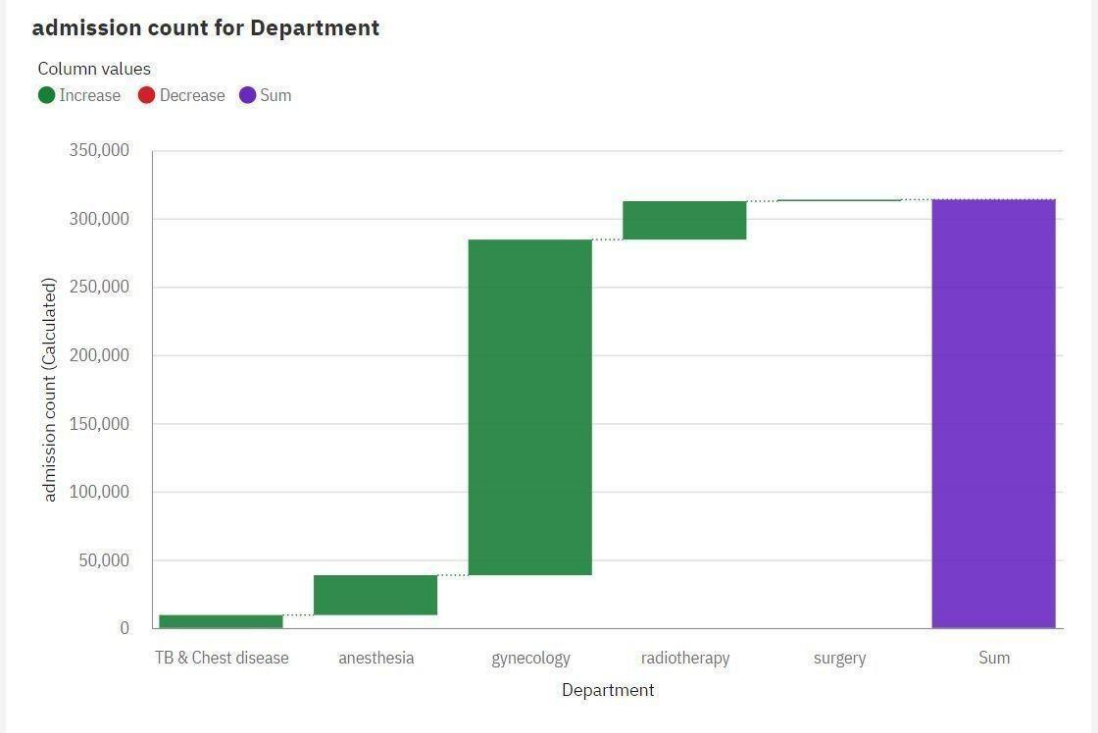


- Room availability by pie-chart

  It describes the available rooms in the specific wards

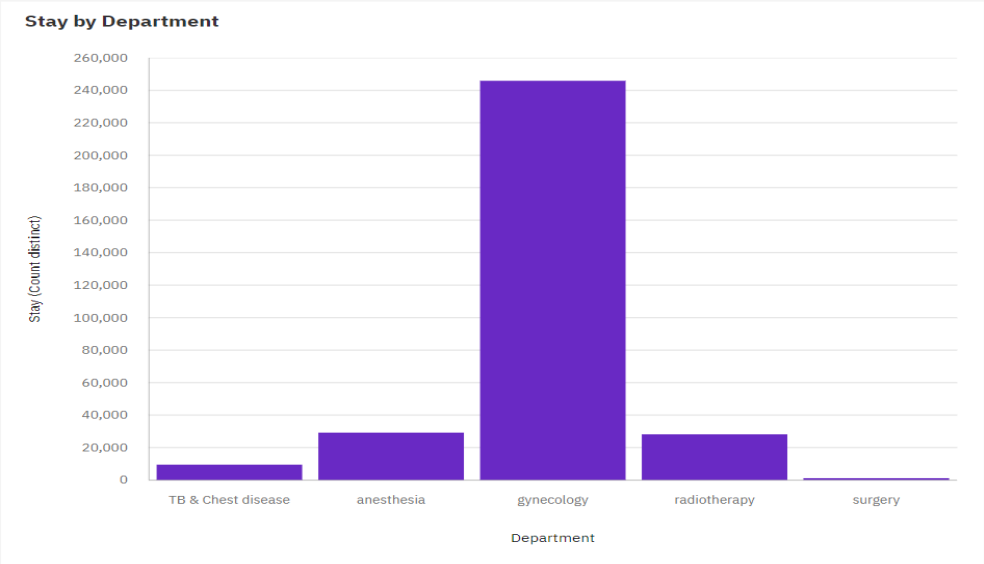


- Department wise no of admissions by waterfall chart.
  It describes the increase in manner of how much admissions are getting into the department by using waterfall chart.



- Number of patients By ward types

  A simple column chart is used to describe the count of the patients in each Department in the hospital.

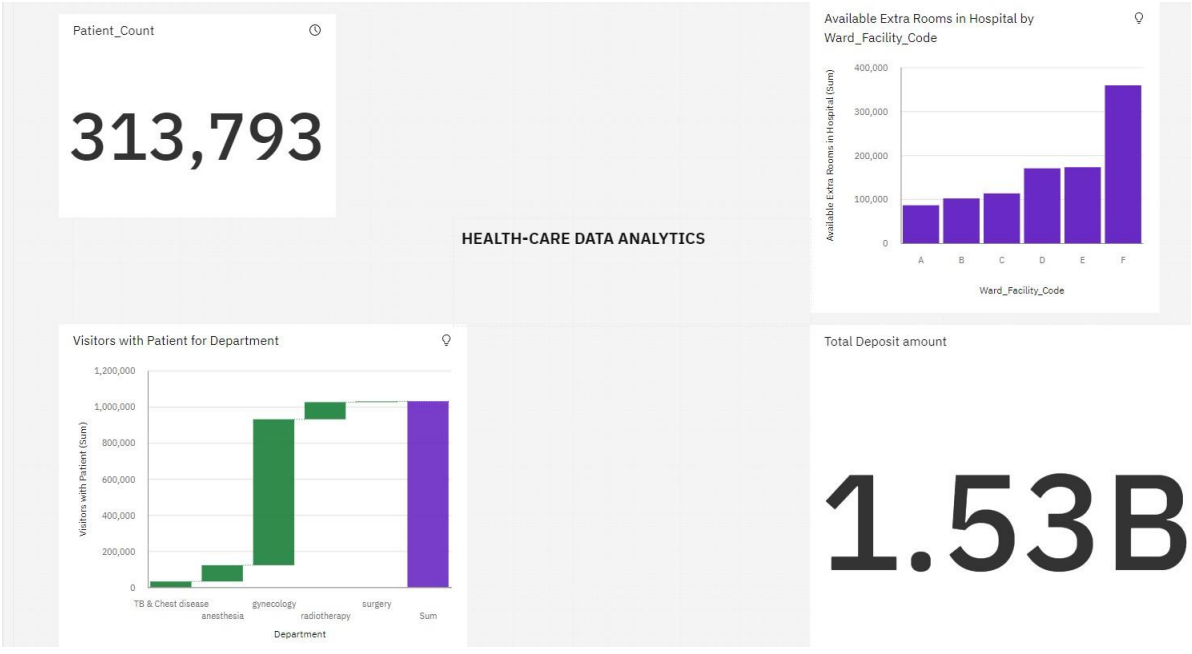


7.3 Dashboard:

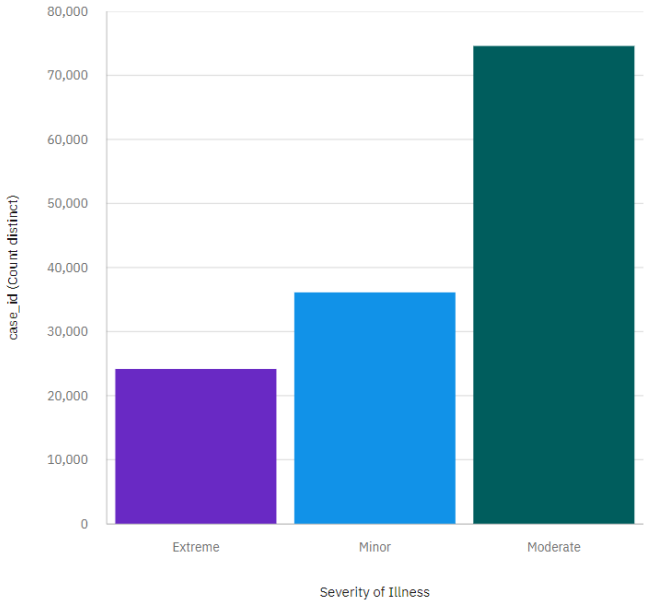
Dashboards are widely used in the business enterprise to present the analyzed data to the customer side or decision making purpose. It can be used to track, derive insights and filter out records using the filters. It is a simple way to manage the data.

It gives the overall presentation of data in a single frame.

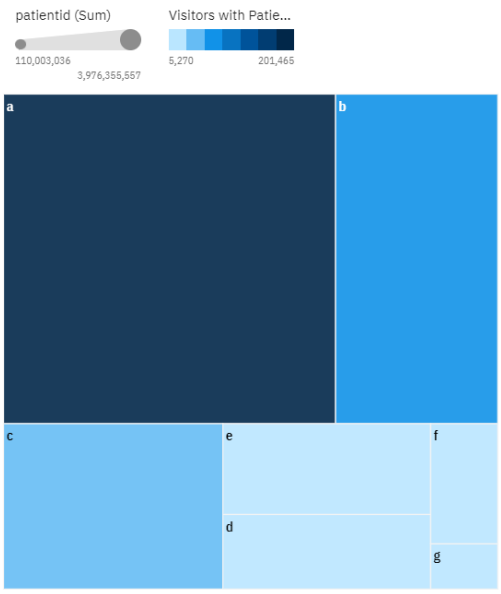
Easily understood by the customer's side.



Severity of Illness Being Reported



Patients with visitors in Different Hospitals



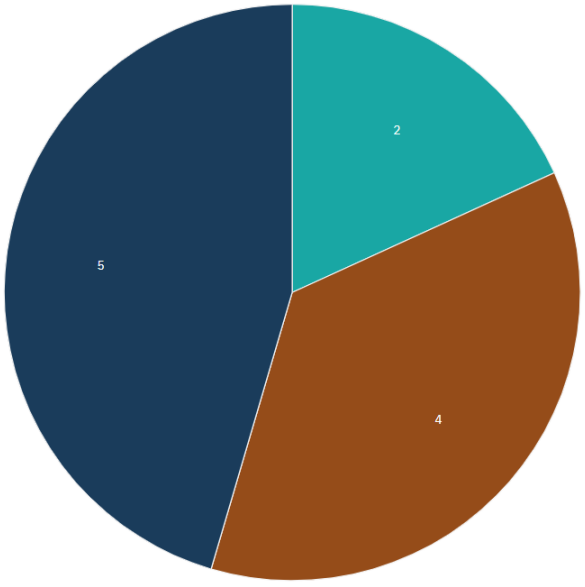
Hospital Code

Select...

Hospital type, Hospital code and Hospital region

Hospital_code	Hospital_region_code	Hospital_type_code
1	Y	d
2	Z	c
3	Z	c
4	X	a
5	X	a
6	X	a
7	X	a
8	Z	c
9	Z	d
10	X	e
11	Y	b

City Code Hospital by Hospital region code



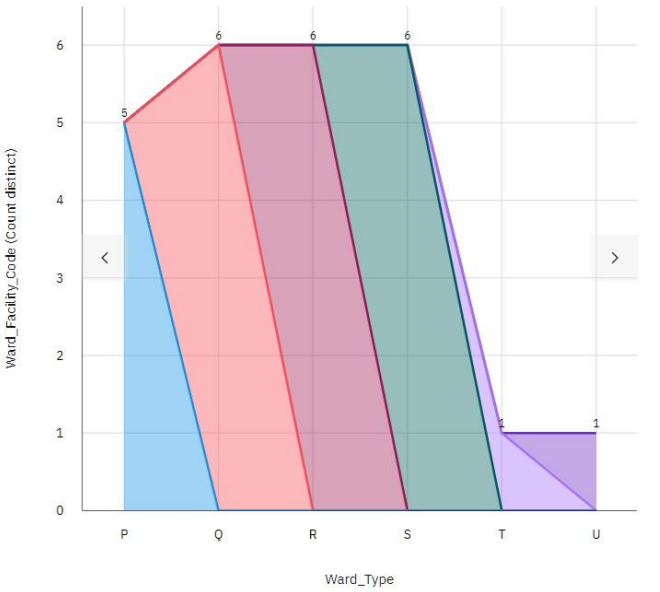
Ward Type

Ward_Type
P
Q
R
S
T
U

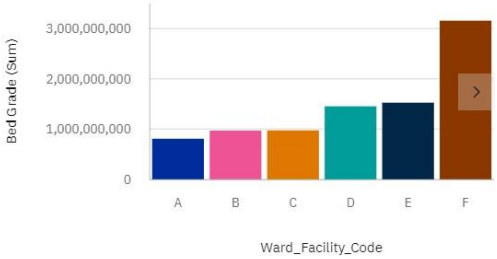
Ward Facility

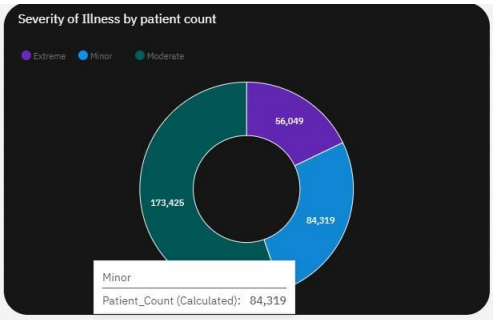
Ward_Facility_Code
A
B
C
D
E
F

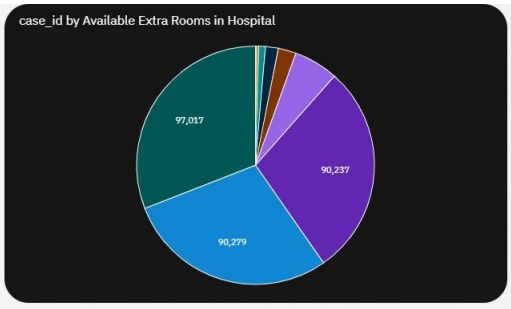
Ward Facility Code by Ward Type

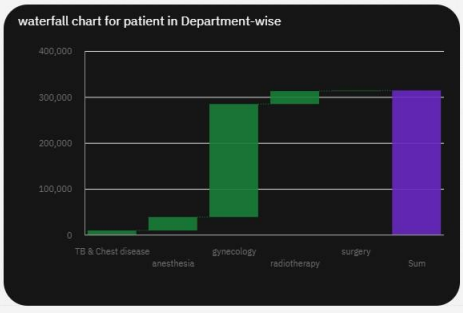


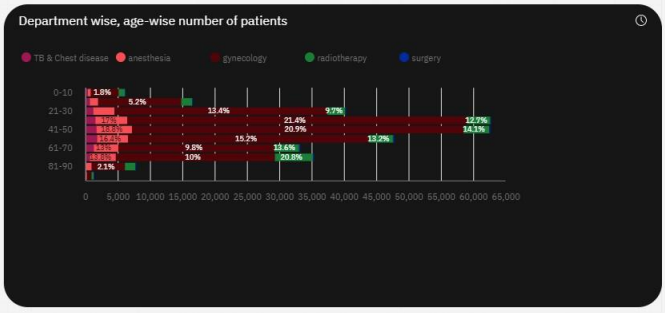
Bed Grade by Ward Facility Code

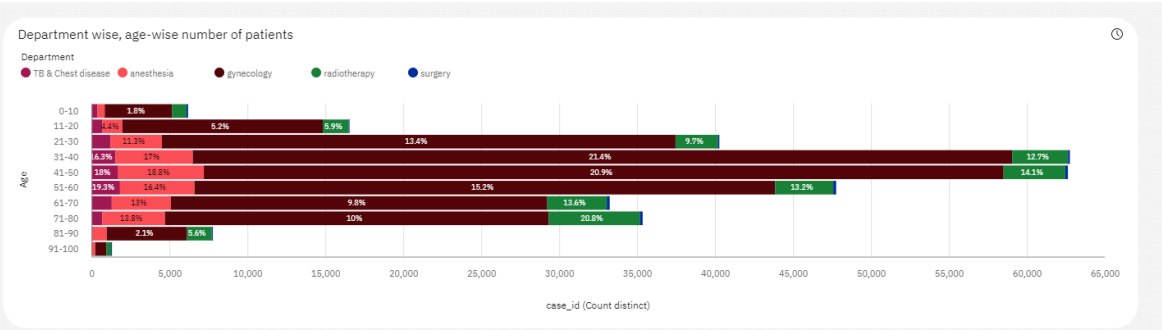


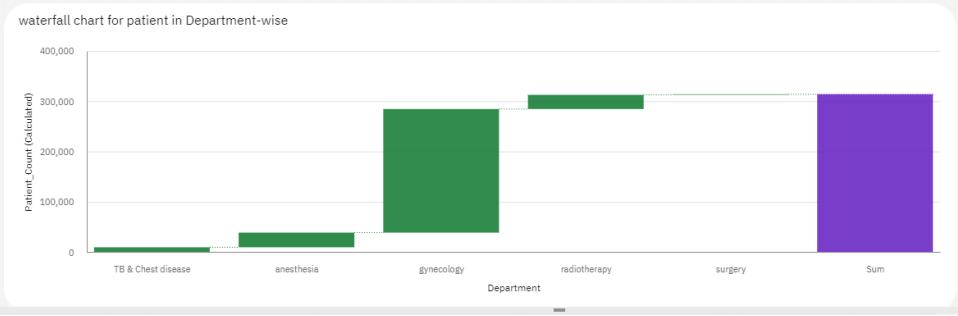


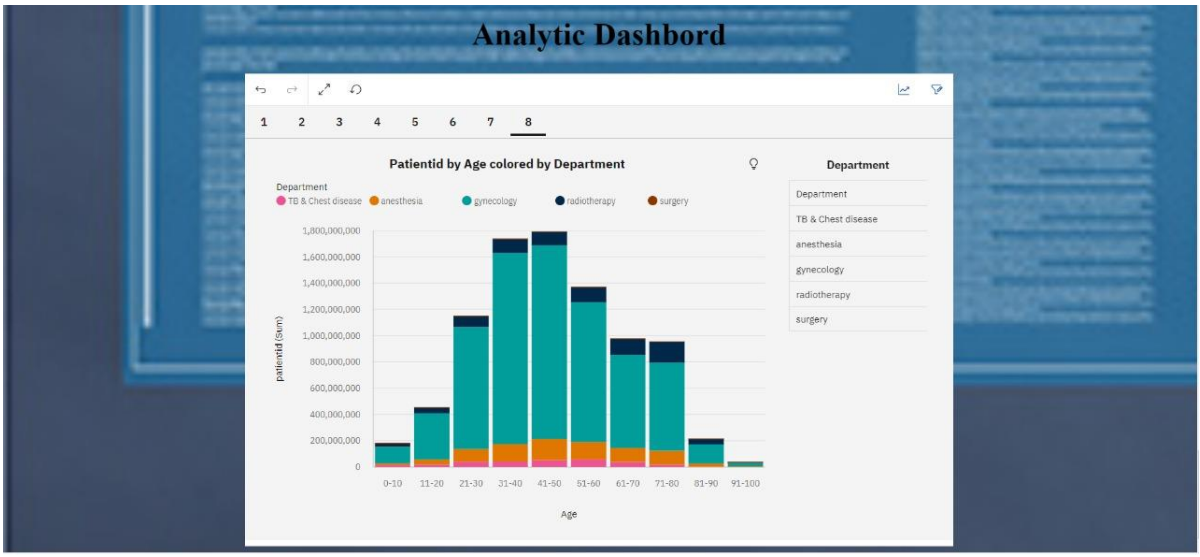




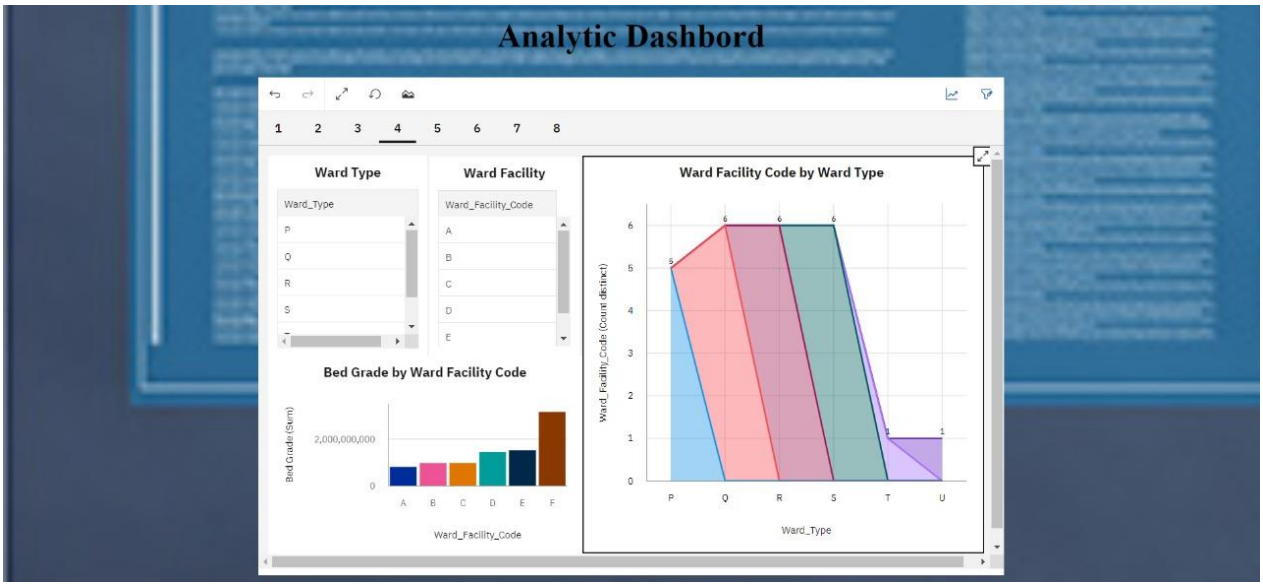


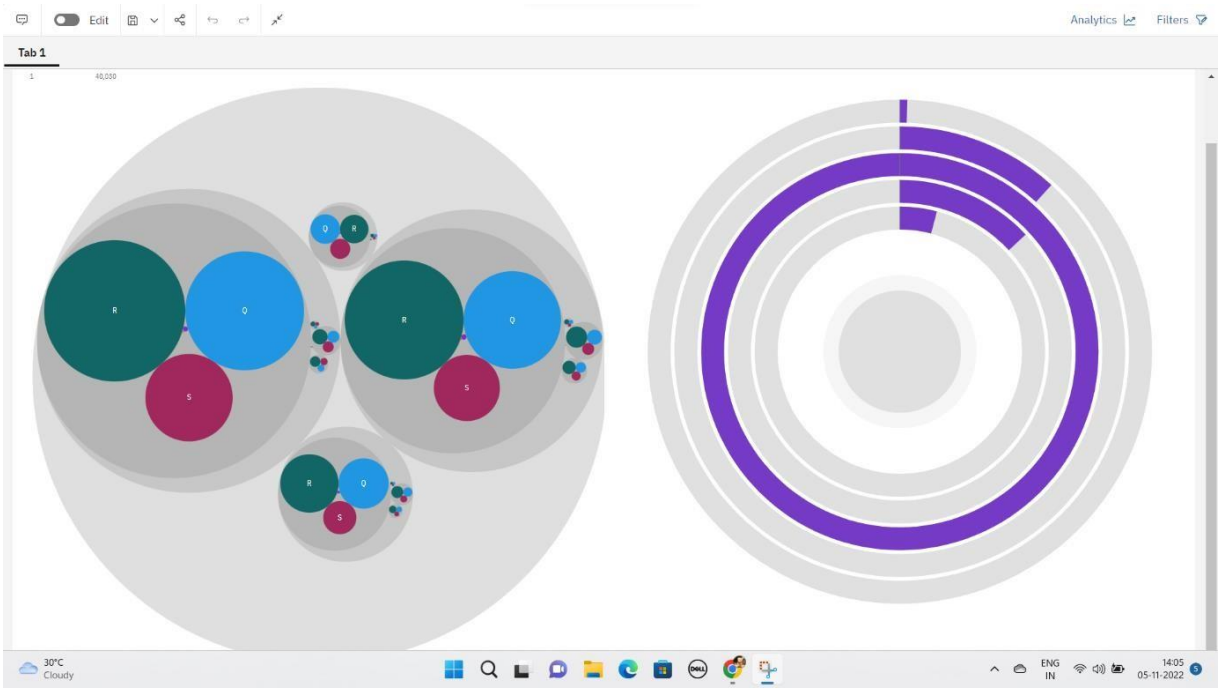

Department wise, age-wise number of patients
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
Age
case_id (Count distinct)


waterfall chart for patient in Department-wise
Patient_Count (Calculated)
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
Sum
Department


Analytic Dashbord
Patientid by Age colored by Department
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
patientid (Sum)
Age


Analytic Dashbord
Case ID With Age
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
Severity of Illness
Extreme
Minor
Moderate
case_id (Count distinct)
Age

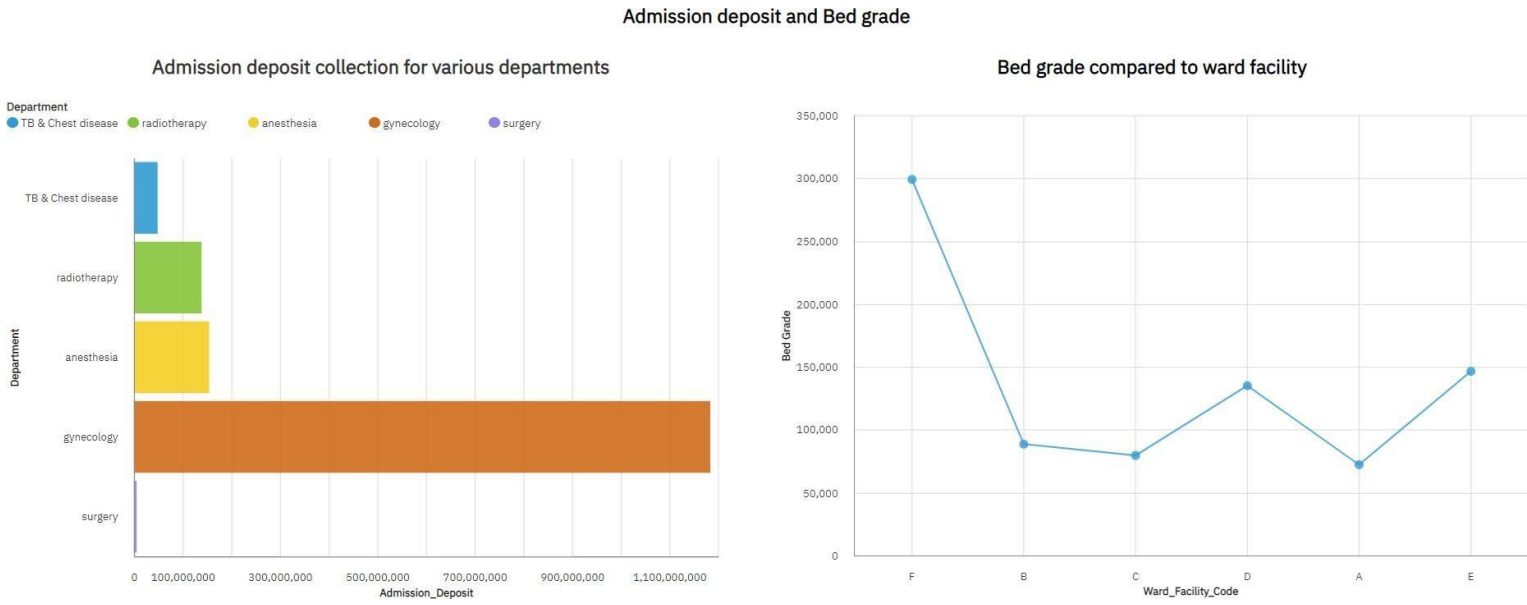

Analytic Dashbord
1 2 3 4 5 6 7 8
Ward Type
Ward_Type
P
Q
R
S
Ward Facility
Ward_Facility_Code
A
B
C
D
E
Bed Grade by Ward Facility Code
Bed Grade (Sum)
2,000,000,000
0
A B C D E F
Ward_Facility_Code
Ward Facility Code by Ward Type
Ward_Facility_Code (Count distinct)
0 1 2 3 4 5 6
P Q R S T U
Ward_Type


Edit
Analytics
Filters
Tab 1
30°C
Cloudy
ENG
IN
14:05
05-11-2022

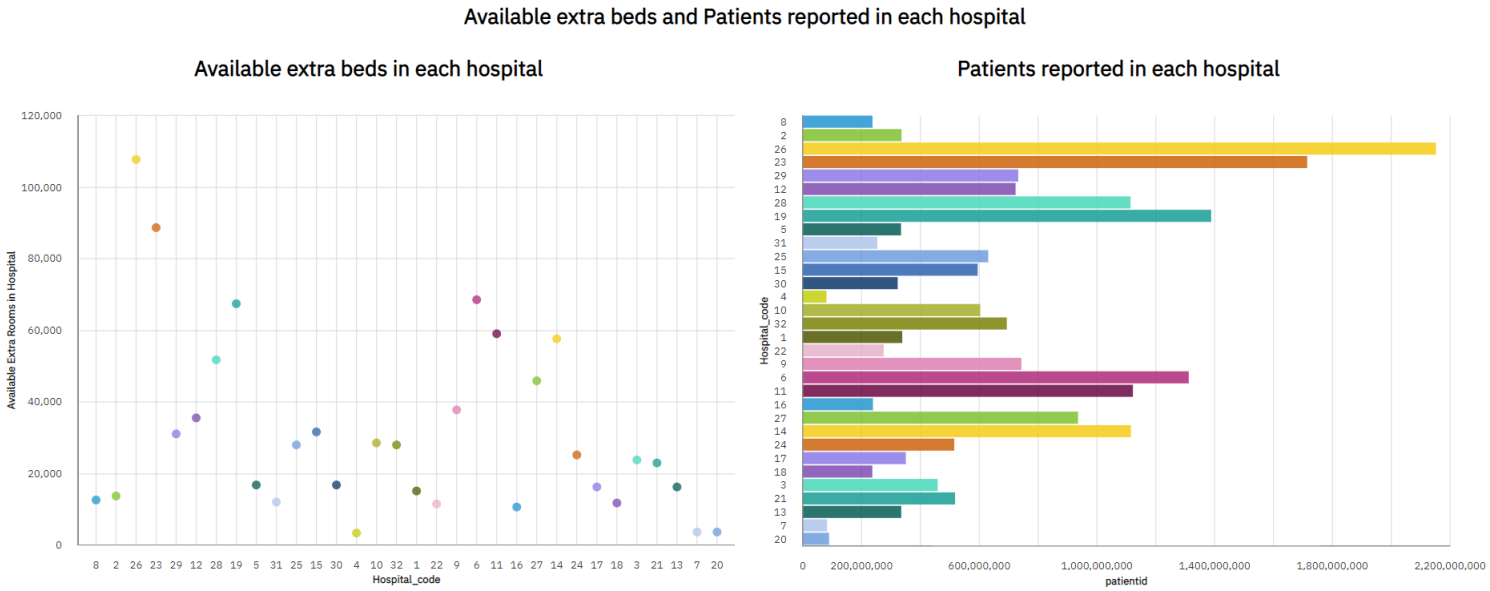
7.3 Story

It is an animation representation of the charts and graphs for better understandability.

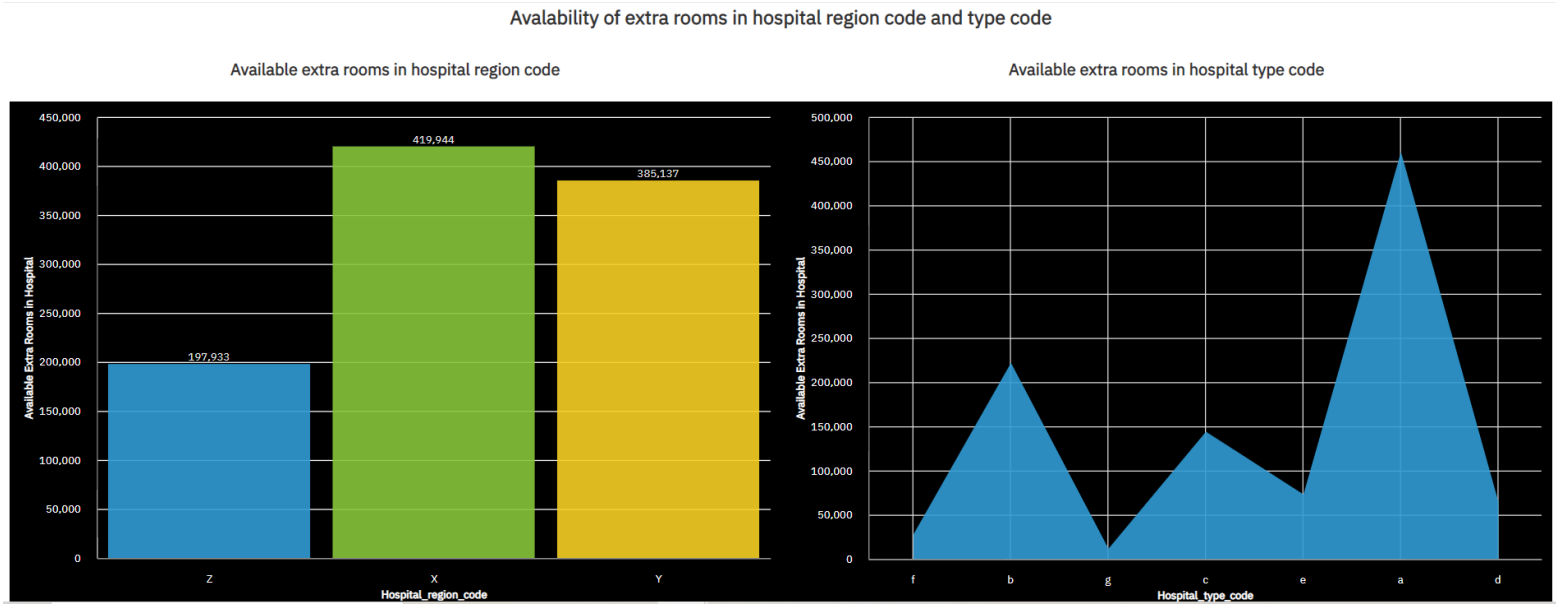
Patient's admission deposit and bed grade compared to ward facility:



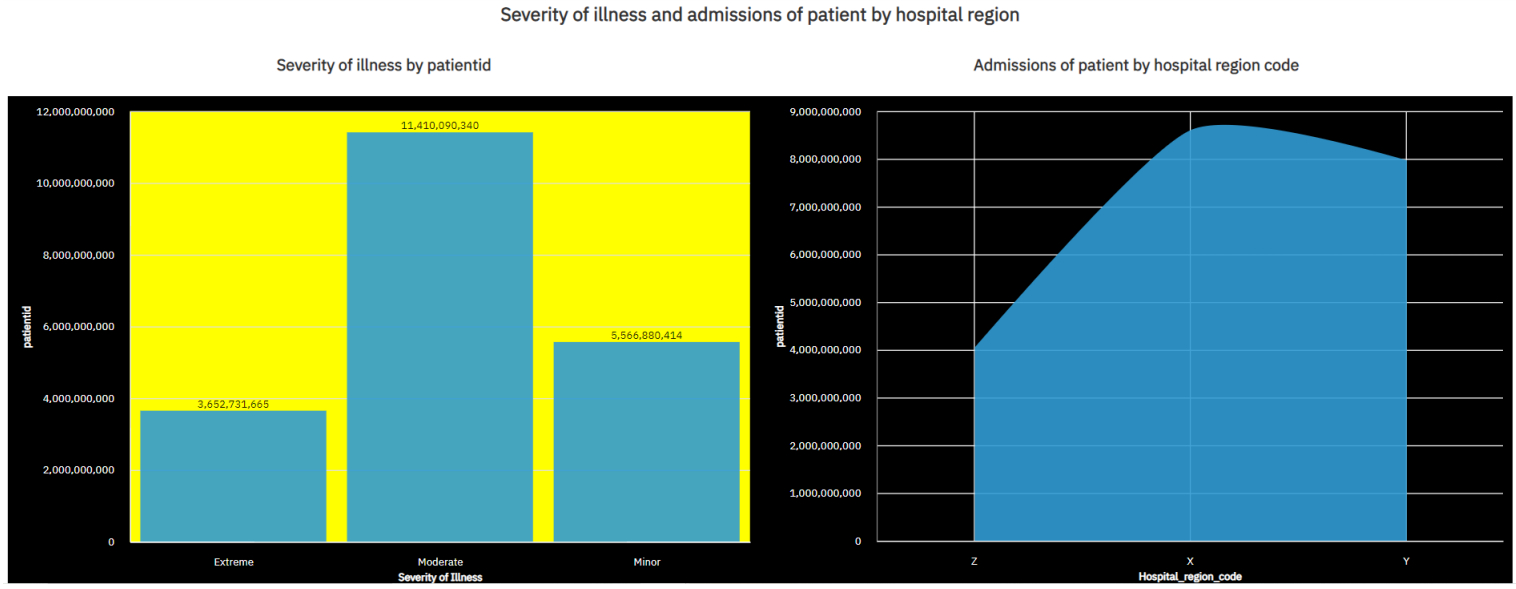
Available extra beds and Patients reported in each hospital:



Availability of extra rooms in hospital region code and type code:

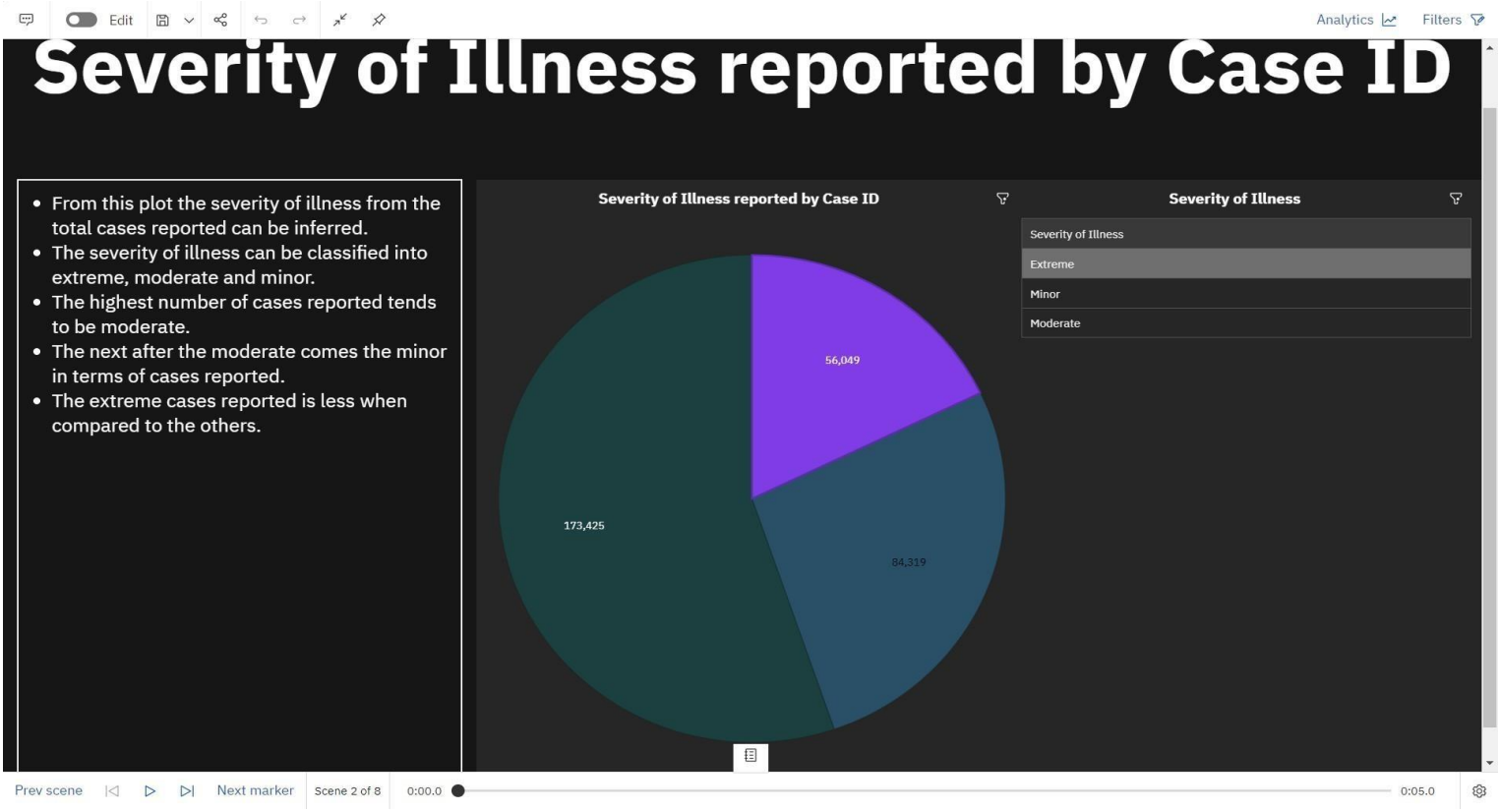


Severity of illness and admissions of patient by hospital region:

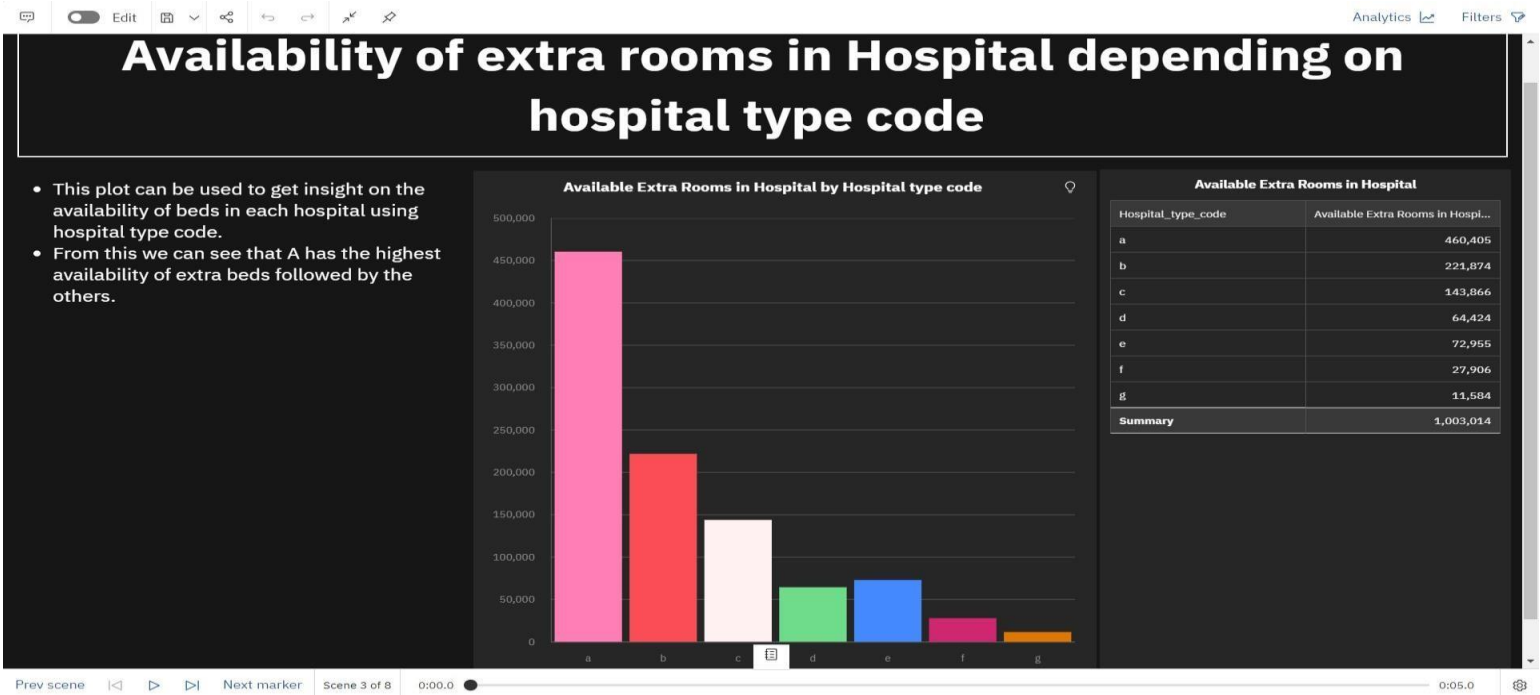


Story

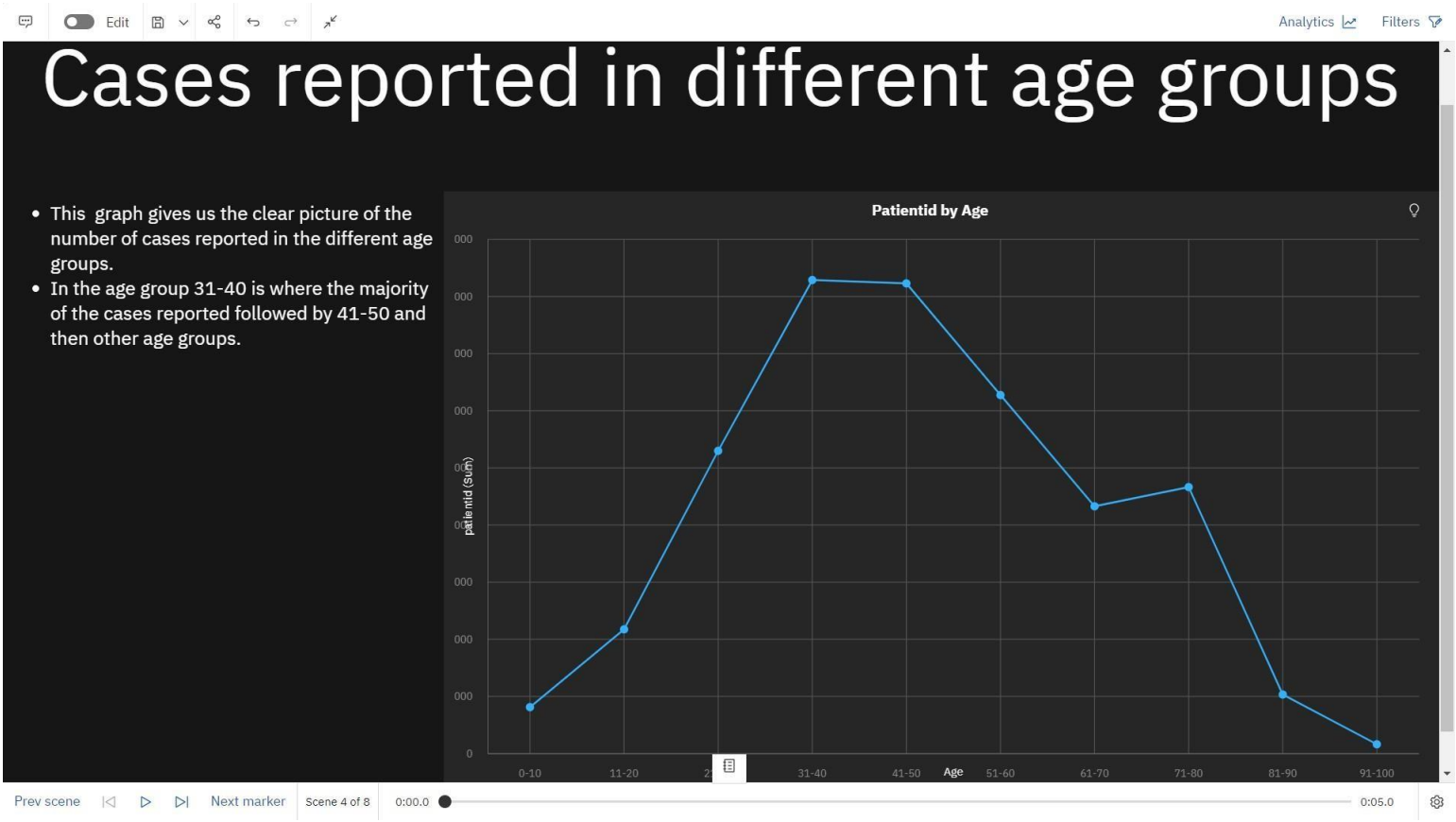
Severity of Illness reported by Case ID:



Availability of extra rooms in hospital depending on hospital type code:

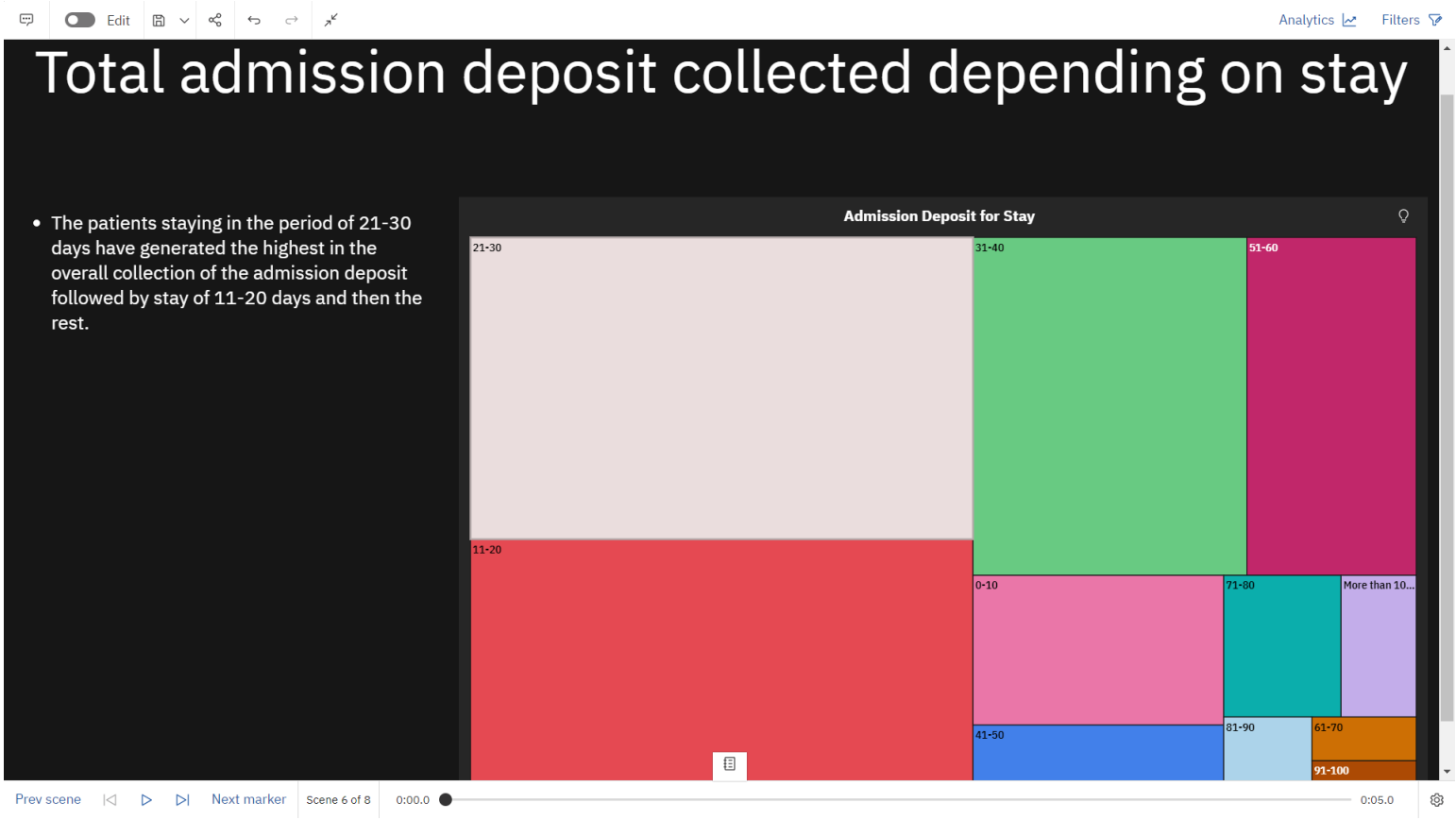


Cases reported in different age groups:

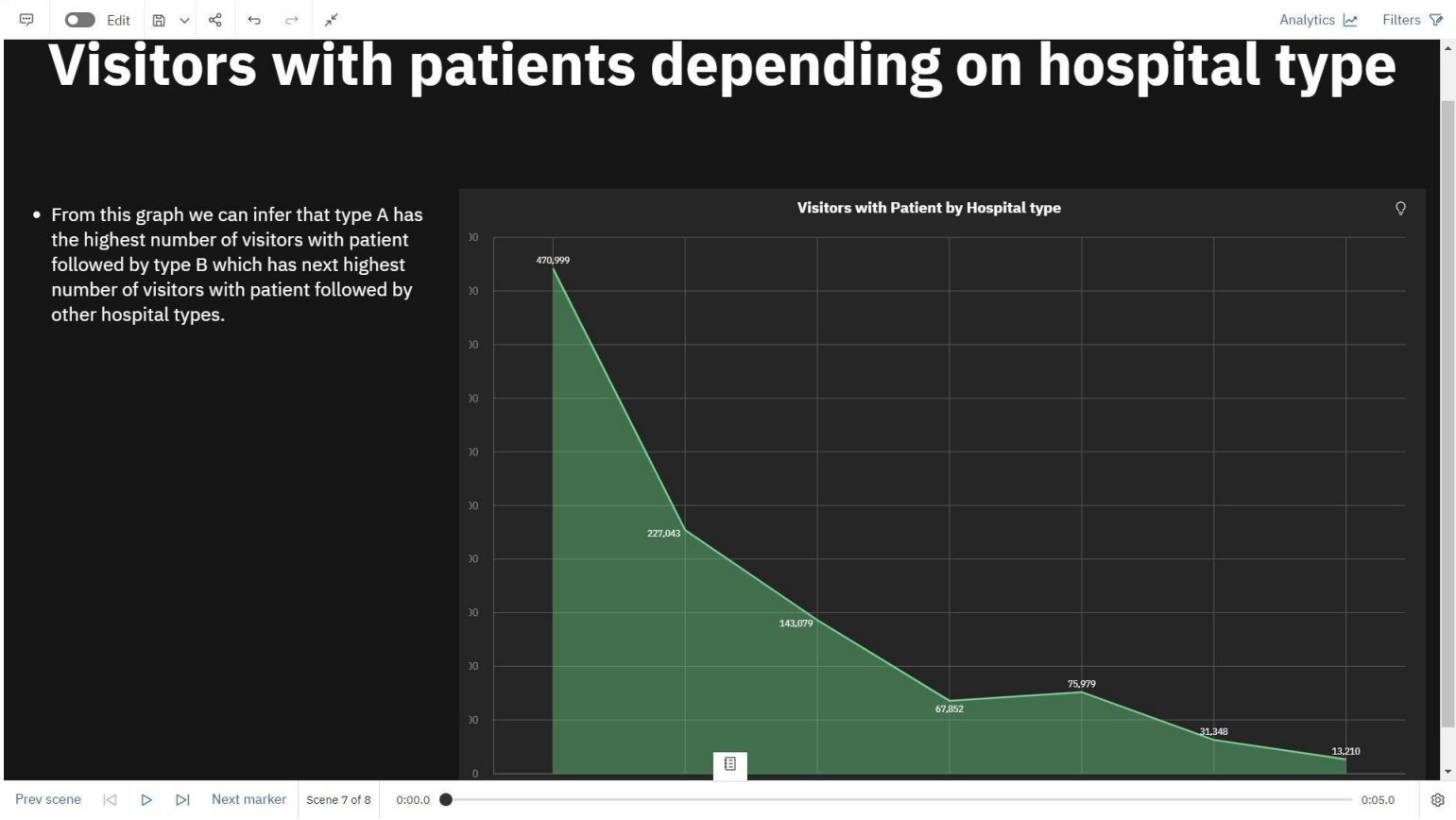


Stay by Patients:

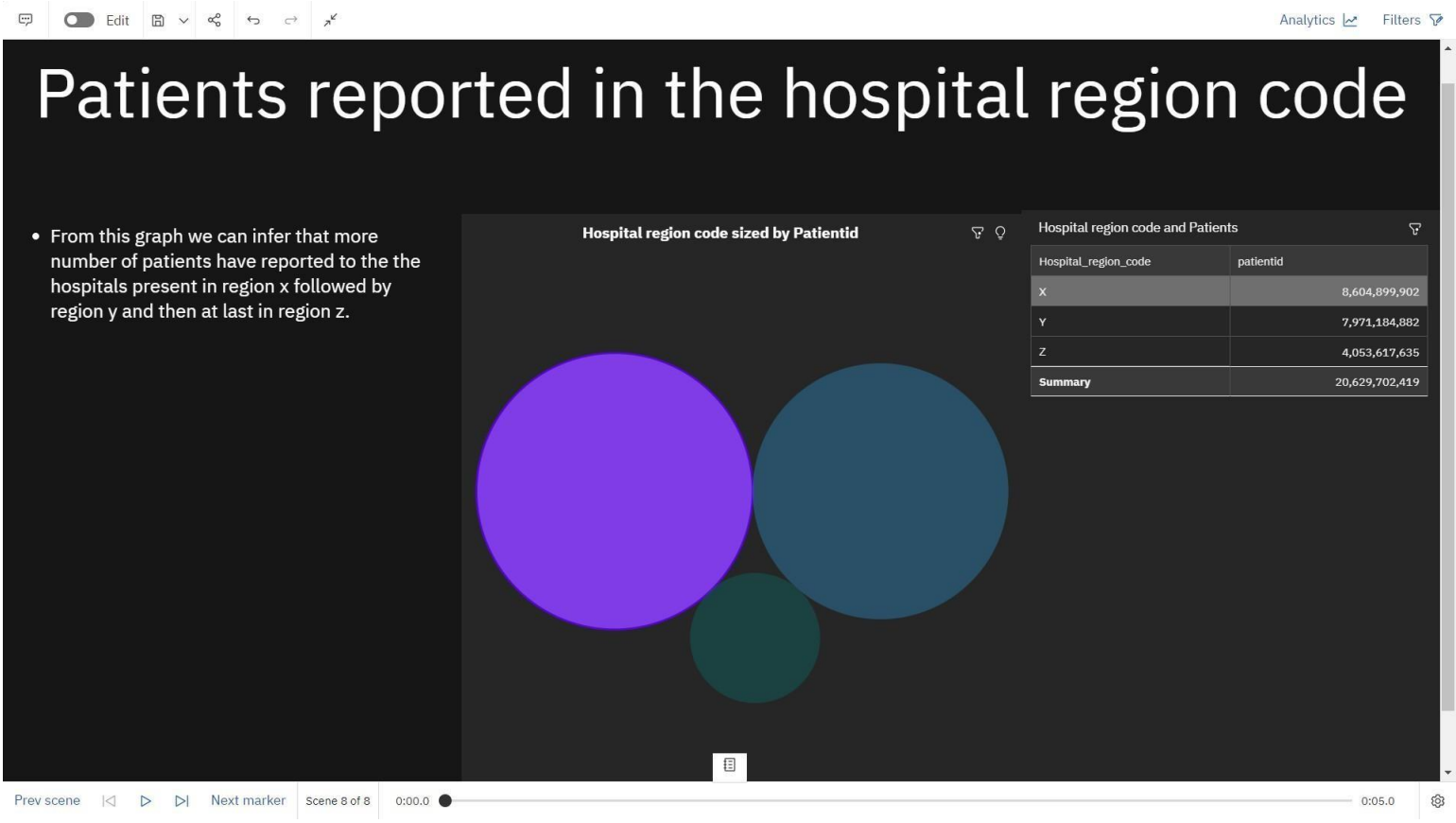
Total admission deposit collected depending on stay:



Visitors with patients depending on hospital type:



Patients reported in the hospital by region code:



Web preview:

It has been done by embedding the Cognos tool report and dashboard into the web browser using the html file.

Html code:

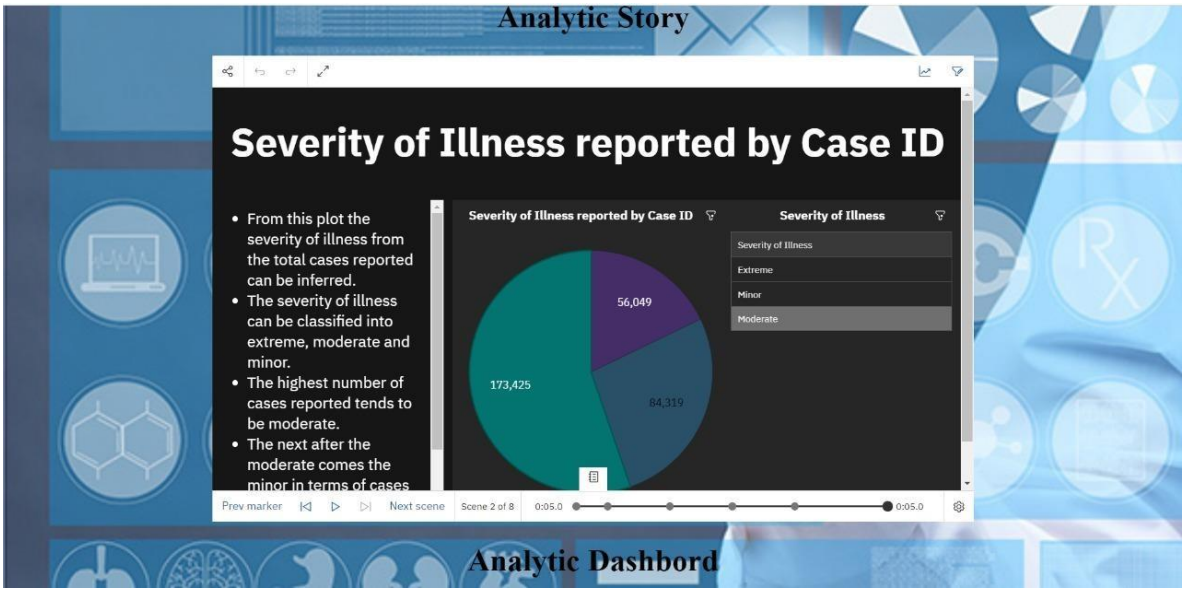
```
<html>
<head>
<title>Analytics for Hospital Health-care Data
</title>
<style>
body {
  background-image: url("ad.jpg");
  background-color: #cccccc;
  background-size: cover;
}
</style>
</head>
<body>
<center><b><h1 style="font-size:40px">WELCOME TO</h1>
</b></center>
<center><b><h1 style="font-size:55px">ANALYTICS FOR
HOSPITAL HEALTH-CARE DATA</h1></b></center>
<center><b><h1 style="font-size:40px">Analytic Story</h1>
</b></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
perspective=story&amp;pathRef=.my_folders%2Fproject%2Bsto
ry&amp;closeWindowOnLastView=true&amp;ui_appbar=false&
amp;ui_navbar=false&amp;shareMode=embedded&amp;action
=view&amp;mode=dashboard" width="980" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><b><h1 style="font-size:40px">Analytic Report</h1>
</b></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
pathRef=.my_folders%2FProject%2Breport&amp;closeWindowO
nLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&a
mp;shareMode=embedded&amp;action=run&amp;format=HT
ML&amp;prompt=false" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
pathRef=.my_folders%2Fproject%2Breport%2B1&amp;closeWin
dowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=fa
lse&amp;shareMode=embedded&amp;action=run&amp;forma
=HTML&amp;prompt=false" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
```

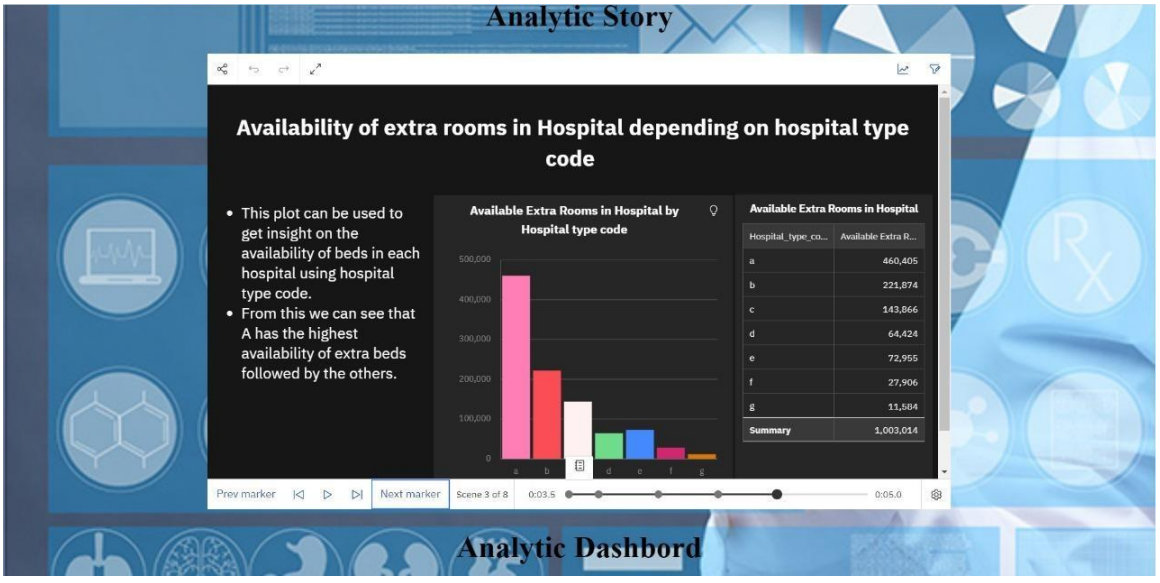
```
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
pathRef=.my_folders%2Fproject%2Breport%2B1&amp;closeWin
dowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=fa
lse&amp;shareMode=embedded&amp;action=run&amp;format
=HTML&amp;prompt=false" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
pathRef=.my_folders%2Fproject%2Breport%2B2&amp;closeWin
dowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=fa
lse&amp;shareMode=embedded&amp;action=run&amp;format
=HTML&amp;prompt=false" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
pathRef=.my_folders%2FNew%2Breport4&amp;closeWindowOn
LastView=true&amp;ui_appbar=false&amp;ui_navbar=false&am
p;shareMode=embedded&amp;action=run&amp;format=HTML
&amp;prompt=false" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>
<center><b><h1 style="font-size:40px">Analytic
Dashbord</h1></b></center>
<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?
perspective=dashboard&amp;pathRef=.my_folders%2FProject%
2BDashboard%2B1&amp;closeWindowOnLastView=true&amp;u
i_appbar=false&amp;ui_navbar=false&amp;shareMode=embed
ded&amp;action=view&amp;mode=dashboard&amp;subView=
model000001848ad04f1c_00000000" width="900" height="600"
frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></center>

</body>
</html>
```

12:11

Web preview in the browser:





8. TESTING

8.1 Unit testing:

It is the process of splitting up the entire program into sub-modules and test the sub-modules individually to check whether it is working or not.

The HTML file for the web app has been tested and it is working without any complications:

```
<html>

<head>

<title>Analytics for Hospital Health-care Data

</title>

<style>

body {

  background-image: url("ad.jpg");

  background-color: #cccccc;

  background-size: cover;

}

</style>

</head>

<body>

<center><b><h1 style="font-size:40px">WELCOME TO</h1></b></center>

<center><b><h1 style="font-size:55px">ANALYTICS FOR HOSPITAL HEALTH-CARE DATA</h1></b></center>

<center><b><h1 style="font-size:40px">Analytic Story</h1></b></center>

<tag background="C:\Users\Adithya Ramanathan\Desktop\story.jpg">

<center><iframe src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.my_folders%2Fproject%2Bstory&amp;closeWindowOnLastView=true&amp;ui_appbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=dashboard" width="980" height="600" frameborder="0"
```

gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>

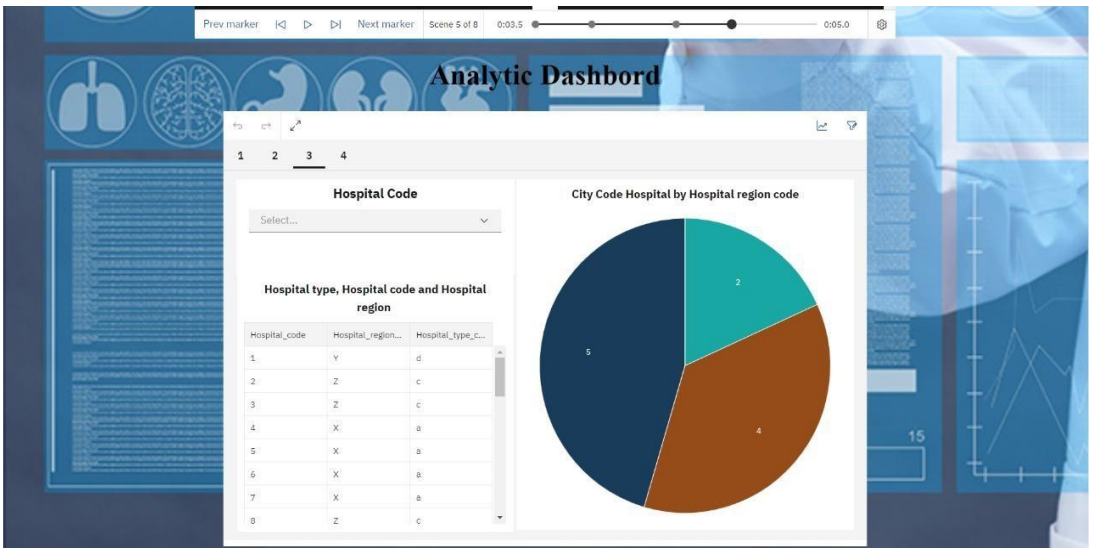
<h1 style="font-size:40px">Analytic
Dashbord</h1>

<iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef
=.my_folders%2FProject%2BDashboard%2B1&closeWindowOnLastVie
w=true&ui_appbar=false&ui_navbar=false&shareMode=embed
ded&action=view&mode=dashboard&subView=model0000018
440986489_00000003" width="900" height="600" frameborder="0"
gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>

</body>

</html>

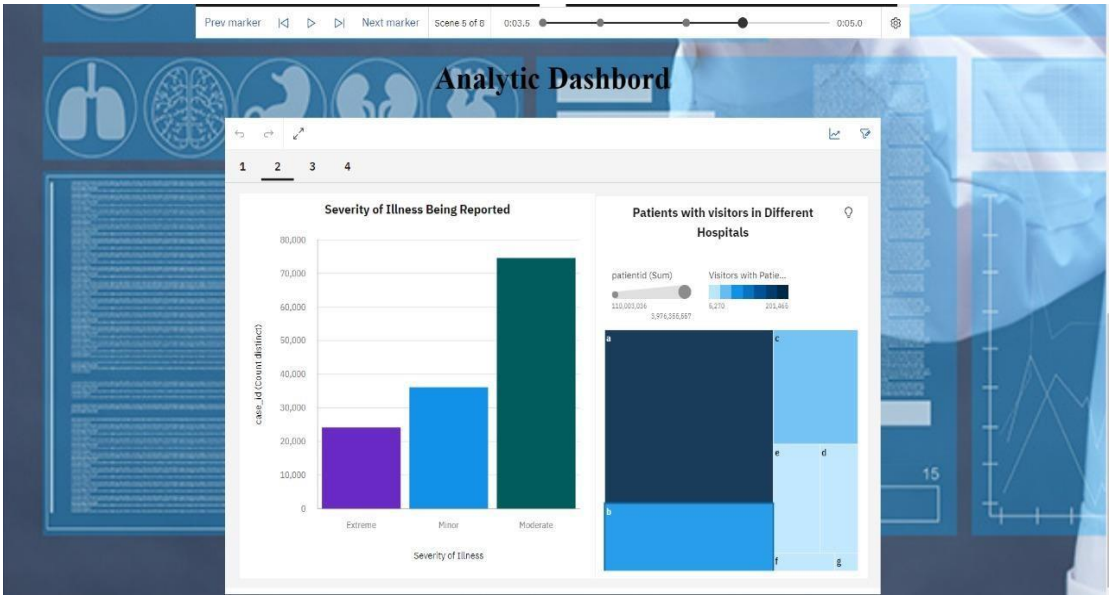
Screenshots:



Integration Testing:

It is the process in which after integrating all the modules as a whole one and make sure that the connected modules are working properly without any errors.

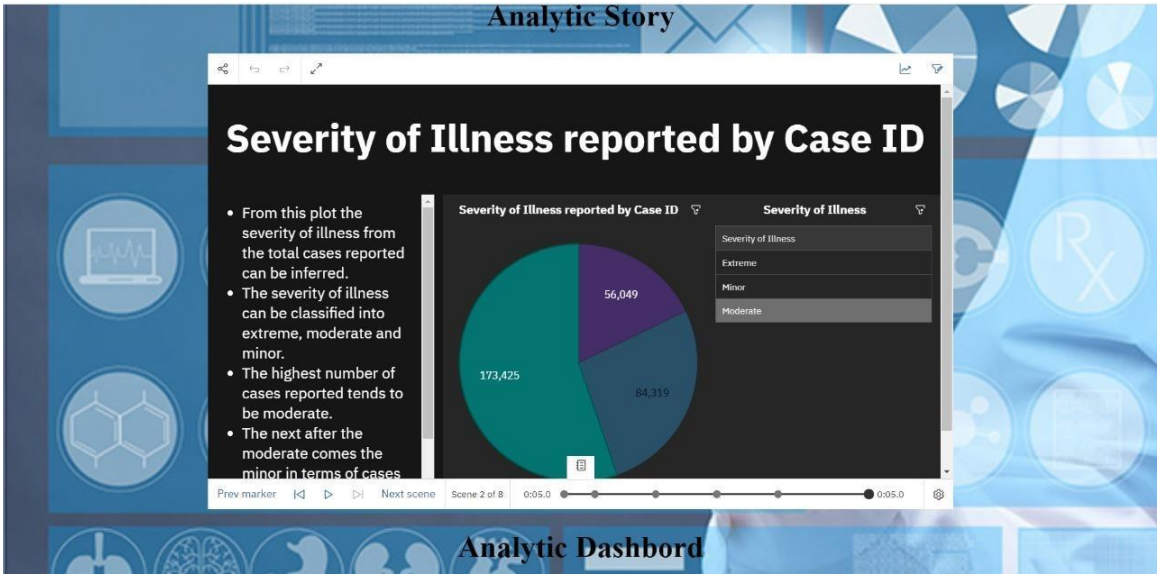
Screenshots:

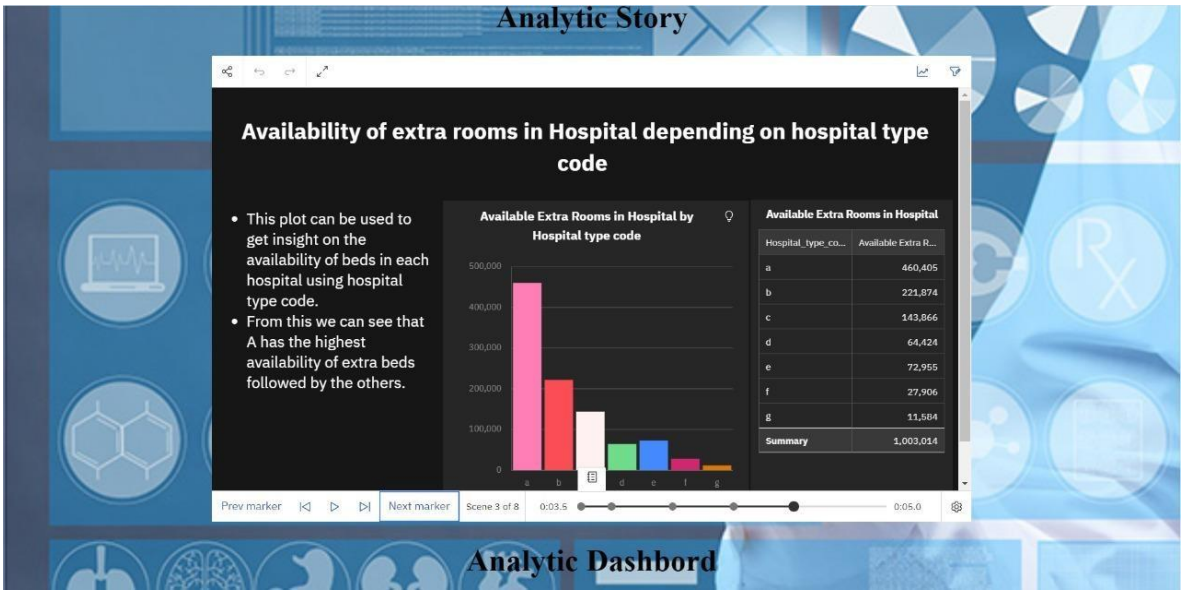


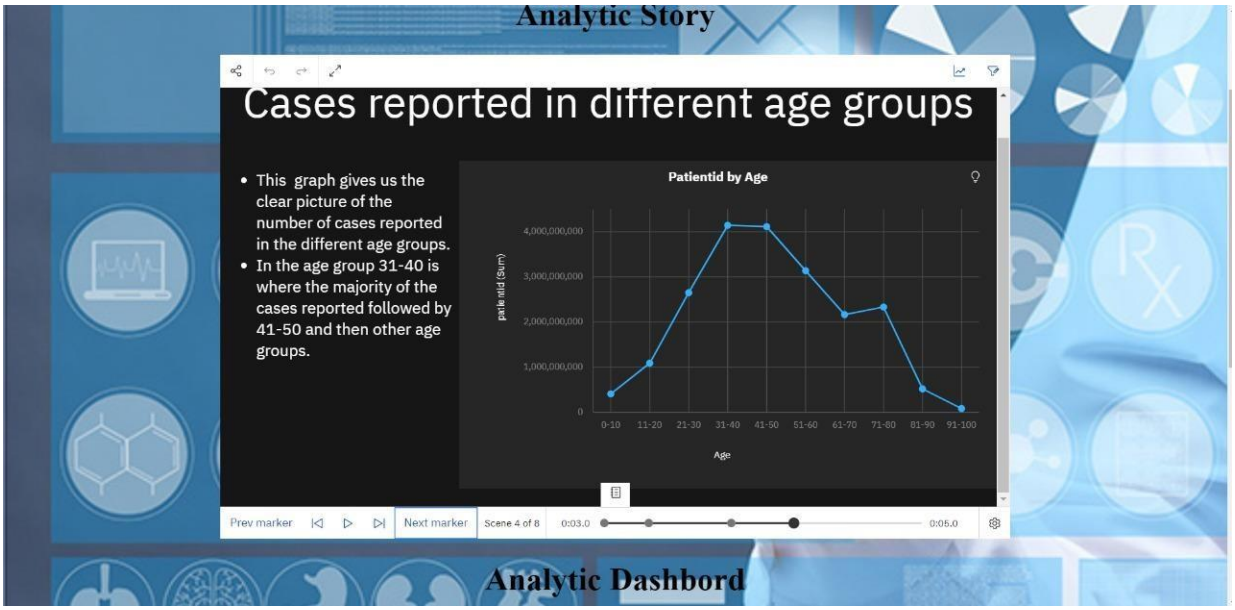
9. RESULTS

9.1 Performance Metrics

The process is likely to be more transparent and it does not take any complex pipeline to solve and compute at the runtime. It makes use of the HTML file and the Cognos analytics at the same time and fetch the data from the Cognos tool to display it in the html file. It takes time to compute, because the Cognos tool is running with the help of the backend server. After the results are fetched , it is displayed in the browser.






Analytic Story
Cases reported in different age groups
This graph gives us the clear picture of the number of cases reported in the different age groups.
In the age group 31-40 is where the majority of the cases reported followed by 41-50 and then other age groups.
Patientid by Age
patientid (Sum)
4,000,000,000
3,000,000,000
2,000,000,000
1,000,000,000
0
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
Age
Prev marker
Next marker
Scene 4 of 8
0:03.0
0:05.0
Analytic Dashbord


Analytic Story
Stay by Patients
From this graph we can infer that the most of the patients have stayed for a period of 21-30 days and the next highest is the period of 11-20 days.
The lowest number of days of occupied by the patients ranges between 91-100.
Patientid by Stay
patientid (Sum)
6,000,000,000
4,000,000,000
2,000,000,000
0
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
More than 100 Days
Stay
Prev marker
Next marker
Scene 5 of 8
0:03.5
0:05.0
Analytic Dashbord

Advantages:

- It can improved research efforts by gathering the clinical data such as EHR, Personal health records, Public health records, Electronic health records etc.
- Gathering and analyzing the available data accurately and the decision makers can make correct choices in regard to patients treatments and surgery.
- It helps improving the efficiency of the health operations, clinical procedures
- It is also improving the health outcomes means the patient outcomes in hospitals.
- Patients data analysis can help the health care providers to reduce errors, lower the fees and better identification of risk populations.
- It obtain the operational insights means that standard of care a patients and positive patients outcomes.
- Data analytics can most benefit areas like administration, emergency preparation, financial management.
- It is improves Staffing and Cost effective use of technology.
- It can help manage the labor costs in healthcare after ensuring the quality-of-care patient receives in both efficient and continuous improvement.

Disadvantages:

- Privacy is an issue in healthcare data, privacy regulations have to be in place to avoid possible data breaches.
- Data can be incomplete means the missing values in datasets, this is because doctors sometimes don’t take necessary lab measurements because of data has been lose.
- Descriptions of features can be unclear, dataset descriptions or features names are not clear, and not easy to understand.
- Values in data can be incorrect, must pay attention while putting data into Electronic Health Record(EHR).
- Communication gaps are arises that is miscommunication between the data scientists and data user.

11. CONCLUSION:

Health-care analysis plays an vital role in the hospitals. The demographic data of the each patient in the different departments with the various diseases can be classified and analysed effectively by using this Analytical tool. It is an AI improvised tool, and it makes the tool more effective and significant. It serves patient as well as doctors to look over the change in pattern over the period of time. It take the input as data, perform the analysis and present report in the form of dashboards to the end users in the web-preview. It is not expensive. It is an fully automated process, once we made the live connection with the data, it reflects whenever the new data pops in or older data pops out. It provides better user understandability of data and the description to the user. The health-care industry can avail better features by utilizing the use of the data analysis, that lead to server patient with better analysis in hand.

12. FUTURE SCOPE:

In future, along with this analysis of the input data, using Machine Learning algorithms for predictive analysis may be more effective. The predictions and classification of user data can be done by using the ML algorithms easily and a full fledged application can be created for analysis and predictions of the Patient data can be done near future. And the information can be directly send to the patient email address without delay. The end-to-end connection can be made with the patient side for providing valuable insights to the patient.

Github Link:

https://github.com/IBM-EPBL/IBM-Project-5014-1658745225

Demo Link:

https://youtu.be/O48L0tpJ4js

https://youtu.be/-VZ8xWQFazc