

# **ANALYTICS FOR HOSPITALS' HEALTH-CARE DATA**

**TEAM ID: PNT2022TMID25381**

**TEAM LEADER : B.ROHINI**

**TEAM MEMBER 1: V.NEENA**

**TEAM MEMBER 2: K.SIVA SANKARI**

**TEAM MEMBER 3: E.SHAMNA**

**TEAM MEMBER 4: I.SOWNDHARYA**

## **Project Report**

### **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

### **6. PROJECT PLANNING & SCHEDULING**

- 6.1 Sprint Planning & Estimation
- 6.2 Sprint Delivery Schedule
- 6.3 Reports from JIRA

### **7. CODING & SOLUTIONING (Explain the features added in the project along with code)**

- 7.1 Feature 1
- 7.2 Feature 2
- 7.3 Database Schema (if Applicable)

### **8. TESTING**

- 8.1 Test Cases

8.2 User Acceptance Testing

## **9. RESULTS**

9.1 Performance Metrics

## **10. ADVANTAGES & DISADVANTAGES**

## **11. CONCLUSION**

## **12. FUTURE SCOPE**

## **13. APPENDIX**

13.1 Source Code

13.2 GitHub & Project Demo Link

## **1.INTRODUCTION:**

### **1.1 Project Overview:**

Data analytics refers to the process and practice of analysing data to answer questions, extract insights, and identify trends. This term also includes the way of how this data is gathered, cleaned and summarized for use and finally the processing of data to support data analytics and predictive modelling.

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare Management. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital.

This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning.

The goal is to accurately predict the Length of Stay for each patient on case by case basis so that the Hospitals can use this information for optimal resource allocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10 days to more than 100 days.

### **1.2 PURPOSE:**

There are many reasons why healthcare professionals might want to predict the length of stay of their patients. For example, if a hospital is trying to forecast how many beds will be needed in the coming days or weeks, they will need to know how long their patients are likely to stay. The purpose of predicting the length of stay of the patients is to estimate the amount of time that a patient will spend in a hospital. This information improves bed turnover, allowing hospitals to match demand with capacity for elective and emergent admissions, intensive care

unit (ICU) care, and interhospital transfers. Length of Stay(LOS) is also an important metric for quality of care, so knowing how to predict it can help healthcare professionals identify areas where they can improve.

## **2. LITERATURE SURVEY:**

### **2.1 Existing Problem:**

There is an error in the existing prediction of length of stay in the hospital. The error is that the length of stay is being predicted as shorter than it actually is. This approximate prediction is causing problems for patients who are planning their hospital stay and for the hospital staff who are trying to allocate resources for the patients.

### **2.2 Reference:**

- Janice Hammond, 2015.
- LexisNexis, 2015.
- Title : Predicting Length of Stay Across Hospital Departments  
Author : Jesus Manuel Puentes Gutierrez.

### **2.3 Problem Statement Definition:**

To understand your customer's point of view create a problem statement. The Customer Problem Statement template helps you focus on what matters to create experiences people will love.

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

### **Customer Problem Statement for Hospitals' Health-Care data:**





Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	a patient	get admitted to the hospital	there is not enough availability of beds	the number of patients is more than the bed available	anxious that I might not get cured
PS-2	a nurse	Allocate resources to the patient	there is not enough available	sufficient resources have not been allocated to the hospital	scared that the patient will not get the right treatment

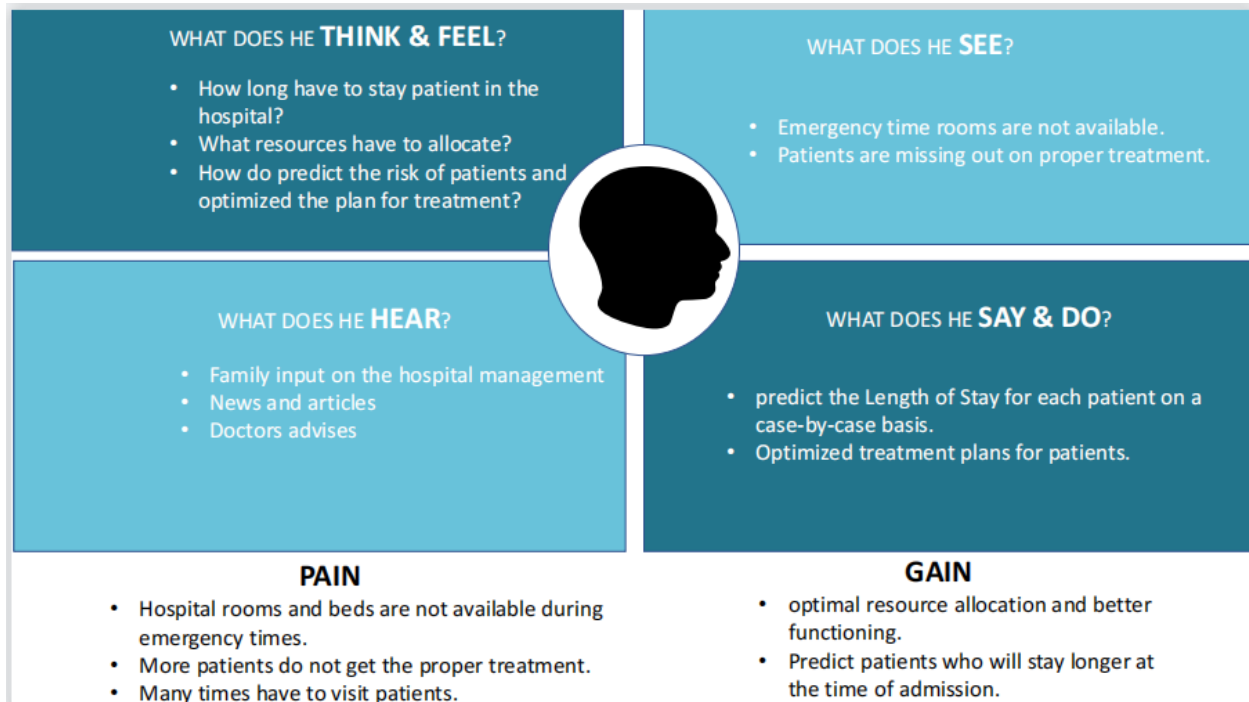
### 3. IDEATION & PROPOSED SOLUTION:

#### 3.1 Empathy Map Canvas:

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviors and attitudes. It is a useful tool to help 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.

#### Empathy map for Predicting Length of Stay (LOS):



### **3.2 Ideation & Brainstorming:**

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

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

This is used to work as a team to present the ideas of each and every person. In the Analytics for hospitals health-care data project, this brainstorming session helped in getting different possible ideas from the point-of-view of every actor involved in the hospitals health-care system.





## Brainstorm & idea prioritization

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

- 🕒 10 minutes to prepare
- 🕒 1 hour to collaborate
- 👥 2-8 people recommended



### Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

🕒 10 minutes



#### A Team gathering

Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.



#### B Set the goal

Think about the problem you'll be focusing on solving in the brainstorming session.



#### C Learn how to use the facilitation tools

Use the Facilitation Superpowers to run a happy and productive session.

[Open article](#) →



### 1 Define your problem statement

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

🕒 5 minutes

PROBLEM  
How might we predict the  
Length of the Stay (LOS)  
for each patient?



### Key rules of brainstorming

To run a smooth and productive session

- 👤 Stay in topic. 🧠 Encourage wild ideas.
- 👤 Defer judgment. 👂 Listen to others.
- 🗣️ Go for volume. 👁️ If possible, be visual.

👥 Share helpful feedback



### Need some inspiration?

See a finished version of this template to get ideas for how to use it.

[Open example](#) →

## Step-2: Brainstorm, Idea Listing and Grouping

2

### Brainstorm

Write down any ideas that come to mind that address your problem statement.

10 minutes

TP

You can select a sticky note and hit the pencil (switch to write) icon to start drawing!

**Deepika**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Sanyukthasi**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Dhanishi**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Anvashagi**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

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 can break it up into smaller sub-groups.

20 minutes

TP

Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mind.

**Based on hospital records**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Based on patient's disease**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Based on patient's health conditions**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Based on patient's previous records**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?

**Emergency situation**

- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?
- Brainstorming: How can we make the patient's life easier?



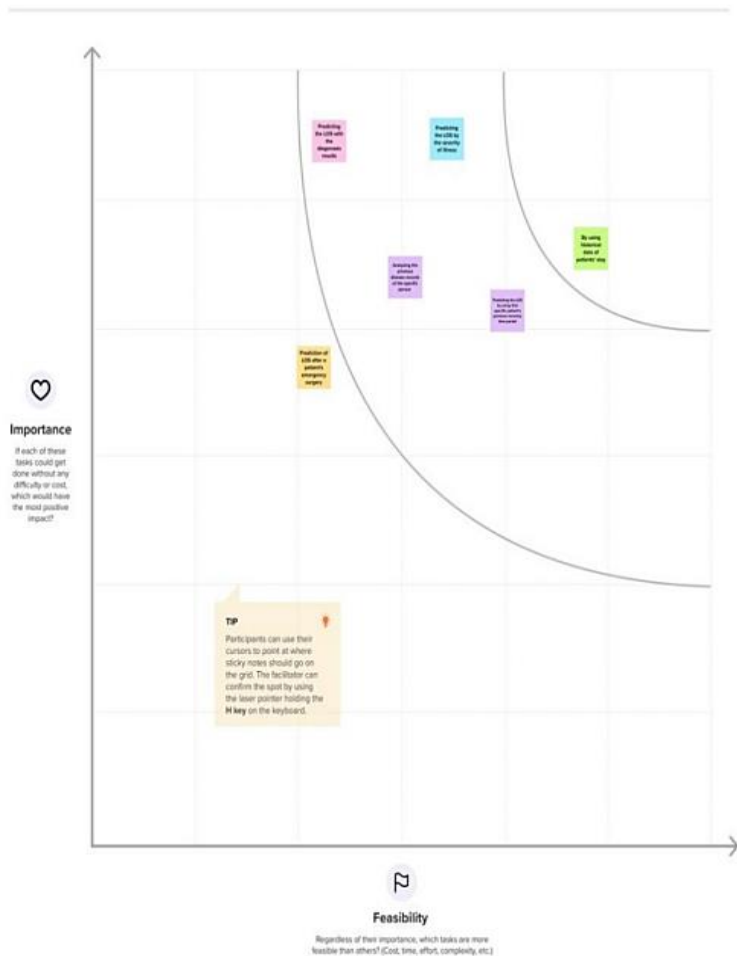
# Step-3: Idea Prioritization

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 to be solved)	To predict the Length of Stay (LOS) for each patient for the allocation of beds and resources.
2.	Idea / Solution description	Predicting the Length of stay by the patients' severity of illness, Age and allocating the beds and resources by using Data visualization tools.
3.	Novelty / Uniqueness	The traditional method's way of allocation of beds and resources by prediction of LOS are approximate and it might lead to insufficient resources. By our proposed solution with analysis of historical data we could predict a more accurate LOS.
4.	Social Impact / Customer Satisfaction	The prediction of a more accurate Length of Stay will result in the allocation of sufficient beds and resources for each patient that helps in their recovery.

5.	Business Model (Revenue Model)	The right Length of Stay prediction could help in allocation of only enough beds and resources and not more than enough. This would lessen the money both the hospital and the patient spent.
6.	Scalability of the Solution	This advanced prediction method instead of the traditional methods makes the Hospital function better by more accurate allocation of beds and resources because it uses the historical data to analyse using visualization tools.

### **3.4 Problem Solution fit:**

Define CS, fit into CC	<b>1. CUSTOMER SEGMENT(S)</b> Who is your customer? i.e. working parents of 0-5 y.o. kids <b>CS</b>  Patients in a hospital who needs to get admitted for a period of time	<b>6. CUSTOMER CONSTRAINTS</b> What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices. <b>C</b>  budget, emergency situations, no cash	<b>5. AVAILABLE SOLUTIONS</b> Which solutions are available to the customers when they face the problem? <b>AS</b> or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an  Approximate prediction of the Length of stay. Pros: Allocation of beds and resources in some amount Cons: The prediction is not accurate leading to lack of resources	Explore AS, differentiate
	<b>2. JOBS-TO-BE-DONE / PROBLEMS</b> Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides. <b>J</b>  Length of stay of each patient needs to be predicted for the right allocation of beds and resources	<b>9. PROBLEM ROOT CAUSE</b> What is the real reason that this problem exists? What is the back story behind the problem? <b>RC</b>  Predicting the Length of stay approximately leads to this	<b>7. BEHAVIOUR</b> What does your customer do to address the problem and get the job done? <b>BE</b> i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering  Check where the resources are available, ask the administration to provide beds	
Focus on J&P, tap into BE, understand RC				Focus on J&P, tap into BE, understand RC

Identify strong TR & EM	<b>3. TRIGGERS</b> What triggers customers to act? i.e. seeing their neighbor installing solar panels, reading about a more efficient solution in the news. <b>TR</b>  Knowing about other hospitals allocating resources and beds more efficiently	<b>10. YOUR SOLUTION</b> If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality. If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behavior. <b>SL</b>  Prediction of Length of stay with their age, severity of illness using historical data so that the prediction would be more accurate and the allocation of beds and resources would be better.	<b>8. CHANNELS of BEHAVIOUR</b> <b>8.1 ONLINE</b> What kind of actions do customers take online? Extract online channels from #7  <b>8.2 OFFLINE</b> What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development. <b>CH</b>  Online: find where the necessary resources are available Offline: check if there are any better hospitals that could provide them bed	Identify strong TR & EM
	<b>4. EMOTIONS: BEFORE / AFTER</b> How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure -> confident, in control - use it in your communication strategy & design. <b>EM</b>  Before: scared, no confidence, weak After: happy, more confident, relieved.			

## 4.REQUIREMENT ANALYSIS:

### 4.1 Functional Requirements:

Following are the functional requirements of the proposed solution.

<b>FR No.</b>	<b>Functional Requirement (Epic)</b>	<b>Sub Requirement (Story / Sub-Task)</b>
FR-1	User Registration	Registration through Gmail
FR-2	User Confirmation	Confirmation via gmail
FR-3	User Login	Login through Gmail
FR-4	Uploading the data	Uploading into the IBM Cognos analytics
FR-5	Preparing the data	Data module creation
FR-6	Exploring the data	Performing predictive analysis
FR-7	Visualizing the data	Creating dashboards

#### **4.2 Non-functional Requirements:**

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

<b>FR No.</b>	<b>Non-Functional Requirement</b>	<b>Description</b>
NFR-1	<b>Usability</b>	The Dashboards are created using cognos analytics to display the results of predictive analysis on a patient's LOS that could be used for easy understanding.
NFR-2	<b>Security</b>	The application helps in preventing unauthorized access.
NFR-3	<b>Reliability</b>	The results will be reliable to the users and always consistent.

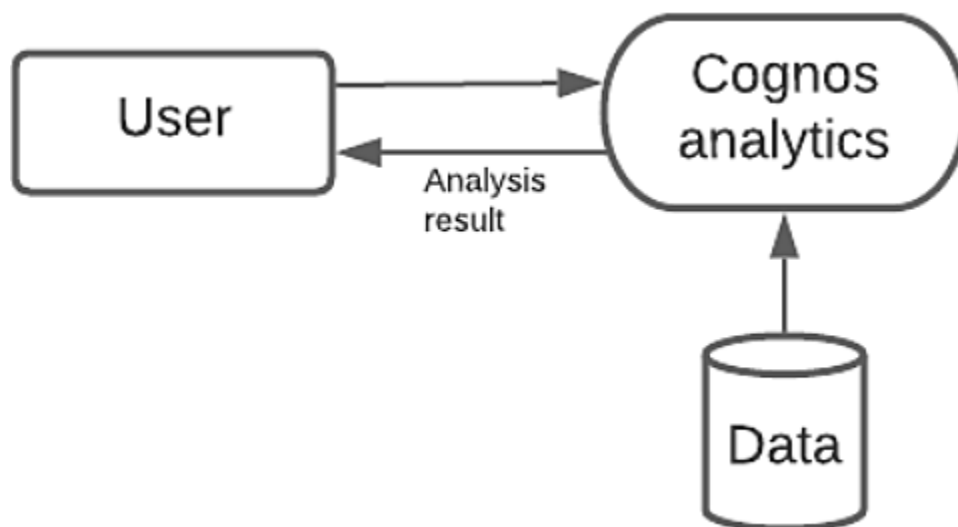
NFR-4	<b>Performance</b>	The analysis results are fast to get. Visualizations are made in less time usually seconds
NFR-5	<b>Availability</b>	The dashboards can be available always when there is a demand
NFR-6	<b>Scalability</b>	The application can perform analysis and create visualizations on any patient's data.

## 5. PROJECT DESIGN:

### 5.1 Data Flow Diagrams:

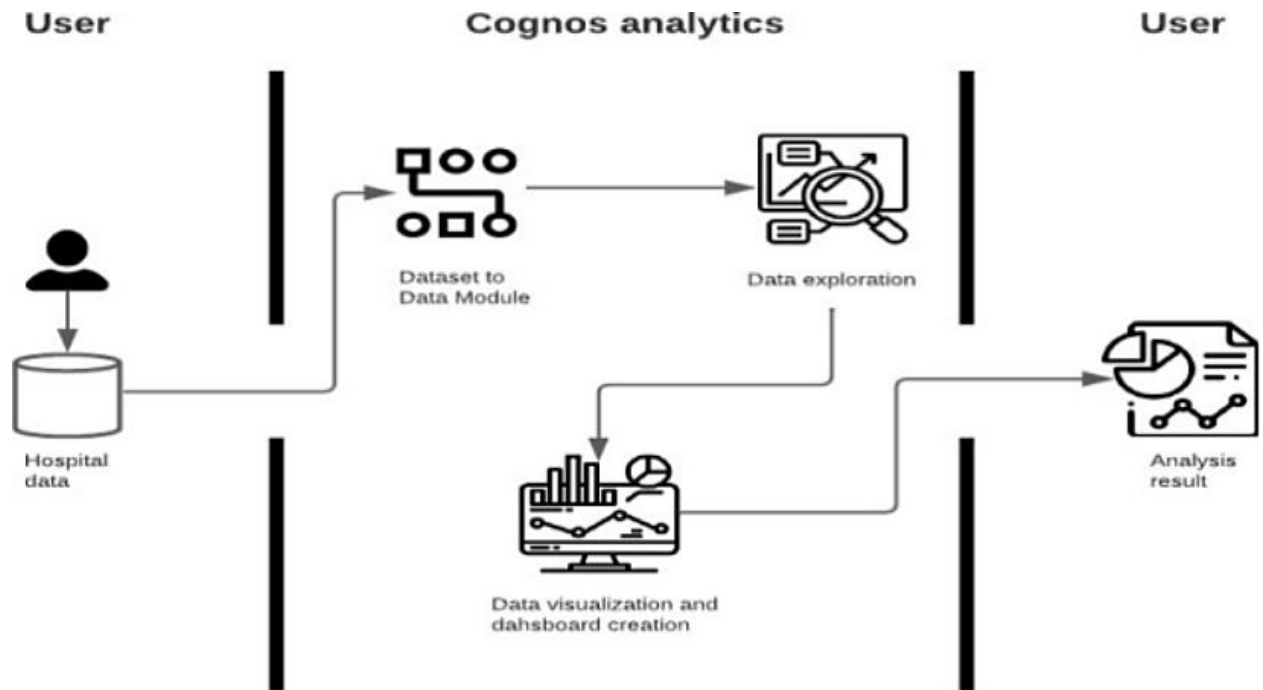
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

#### Data flow diagram level 0:



### 5.2 Solution & Technical Architecture:





### 5.3 User Stories:

The below template list all the user stories for the product.

User Type	Functional Requirement(Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
User (Data analyst)	Registration	USN-1	As a user, I can register for the Cognos analytics account by entering my gmail, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive	I can receive confirmation email &	High	Sprint-1

			confirmation gmail once I have registered for the application	click confirm		
	Login	USN-3	As a user, I can log into the application by entering gmail & password	I can receive confirmation gmail & click confirm	High	Sprint- 1
	Application dashboard	USN-4	Uploading the data through the quick launch	The data gets uploaded	High	Sprint- 1
		USN-5	Uploading the data through the hamburger menu	The data gets uploaded	High	Sprint- 1
		USN-6	Preparing the data by creating a data module	The data gets cleaned.	High	Sprint- 2
		USN-7	Exploration of data for the prediction of LOS Data visualizationw ith the creation of dashboards	The data is analysed	High	Sprint- 3
		USN-8		<i>The data is presented</i>	<i>High</i>	<i>Sprint -4</i>

## 6.PROJECT PLANNING & SCHEDULING:

### 6.1 Sprint Planning & Estimation:

The below template shows the Sprint Planning & Estimation:

<b>Sprint</b>	<b>Functional Requirement (Epic)</b>	<b>User Story Number</b>	<b>User Story / Task</b>	<b>Story Points</b>	<b>Priority</b>	<b>Team Members</b>
Sprint-1	Registration	USN-1	As a user, I can register for the cognos analytics application by entering my email, password, and confirming my password.	10	High	Rohini Neena Shamna Siva sankari
Sprint-1	Data uploading	USN-2	As a user, I will be uploading my data into the cognos analytics	10	High	Rohini Siva sankari Shamna Sowndharya
Sprint-2	Data preparation	USN-3	As a user, I will be cleaning the data for analysis and creating a data module	5	High	Shamna Neena
Sprint-2	Data Analysis	USN-4	As a user, I will be performing analysis on the data for making predictions	5	High	Rohini Siva sankari
Sprint-2	Dashboards	USN-5	As a user, I will be making visualizations and interactive dashboards from the data	10	High	Neena Sowndharya
Sprint-3	Story	USN-6	As a user, I will be making stories from the data and the dashboards	20	High	Shamna Siva sankari

Sprint-4	Report	USN-7	As a user, I will be making a report from the analysis and dashboards	20	High	Siva sankari Rohini
----------	--------	-------	---	----	------	---------------------------

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

### Velocity:

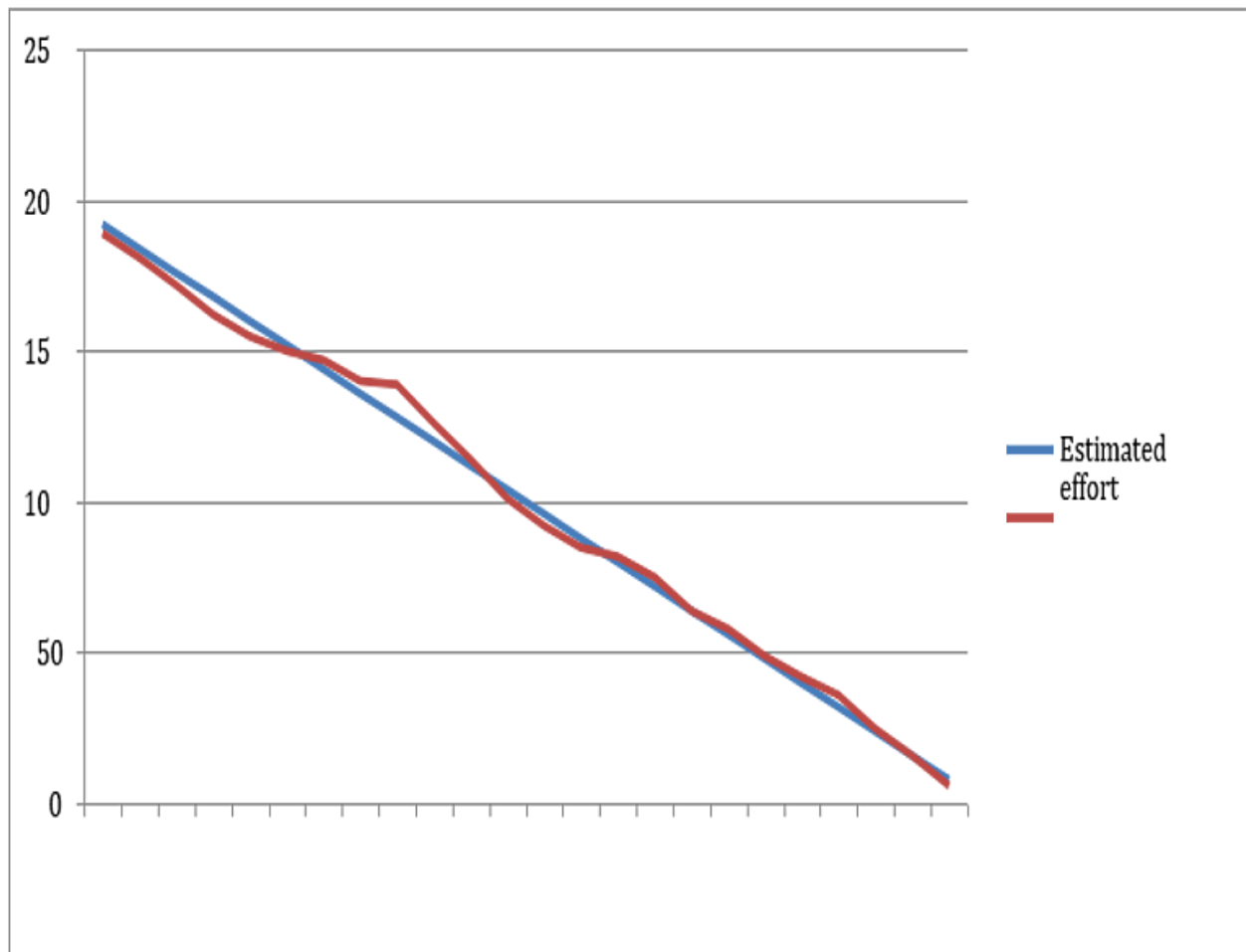
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day).

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

## 6.3 Reports from JIRA:

### Burndown Chart:

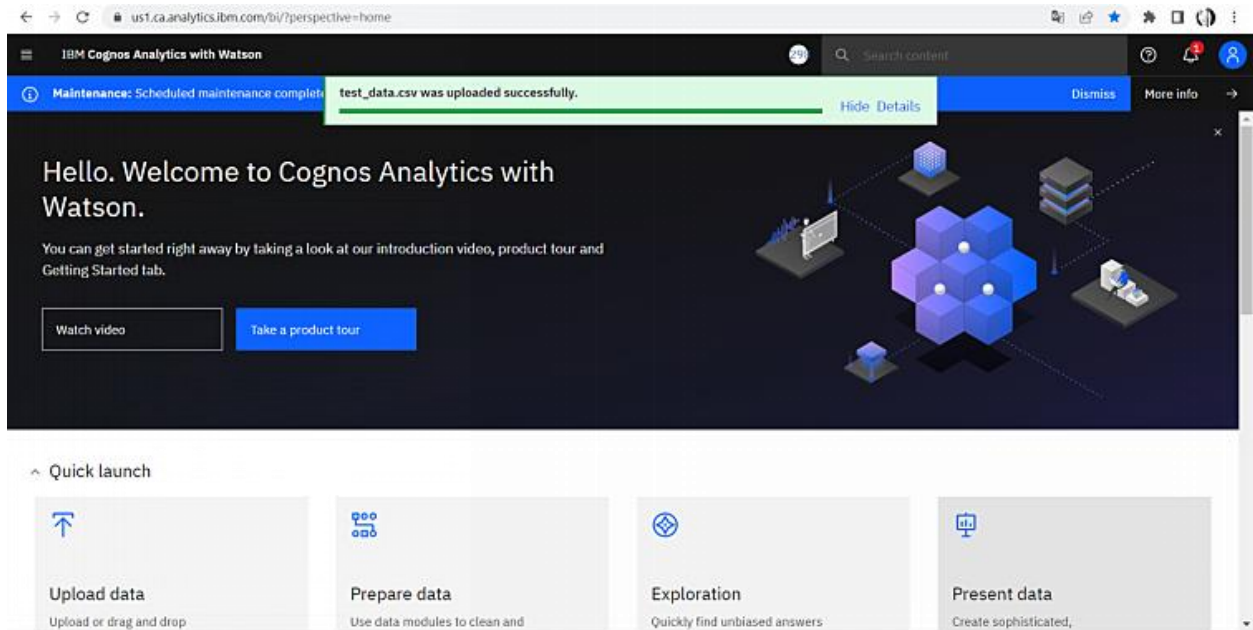
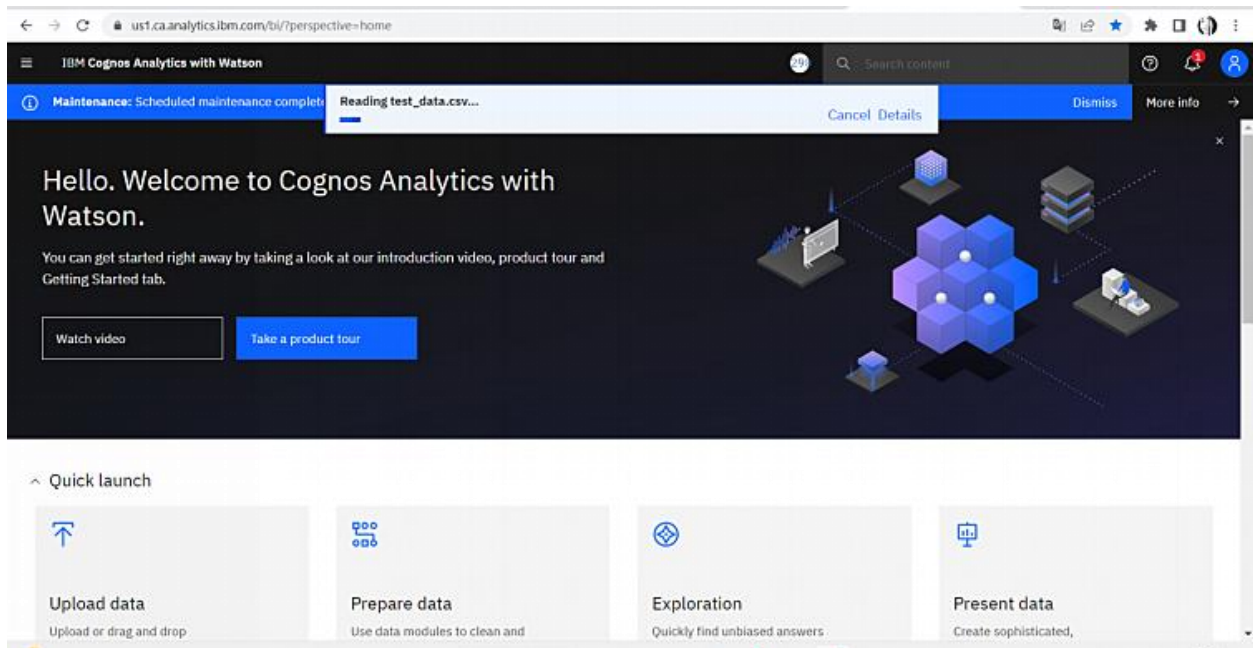
A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.



## 7. CODING & SOLUTIONING:

### 7.1 Feature 1:

- Uploading the data in the IBM cognos analytics.



- *Creating the data module.*

IBM Cognos Analytics with Watson Hospital data module 11 Search content ? 1 ?

**Maintenance:** Scheduled maintenance completed. Click More Info for details and to subscribe to future events Dismiss More info →

Grid Relationships Custom tables

**Data module** + ?

Search

Hospital data module

Navigation paths +

test\_data.csv

train\_data.csv

Type of Admission	Severity of Illness	Visitors with Patient	Age	Admission_Deposit	Stay
Emergency	Extreme	2	51-60	4911	0-10
Trauma	Extreme	2	51-60	5954	41-50
Trauma	Extreme	2	51-60	4745	31-40
Trauma	Extreme	2	51-60	7272	41-50
Trauma	Extreme	2	51-60	5558	41-50
Trauma	Extreme	2	51-60	4449	11-20
Emergency	Extreme	2	51-60	6167	0-10
Trauma	Extreme	2	51-60	5571	41-50
Trauma	Extreme	2	51-60	7223	51-60
Trauma	Extreme	2	51-60	6056	31-40
Urgent	Extreme	2	51-60	5797	21-30

IBM Cognos Analytics with Watson Hospital data module 11 Search content ? 1 ?

Grid Relationships Custom tables

**Data module** + ?

Search

Hospital data module

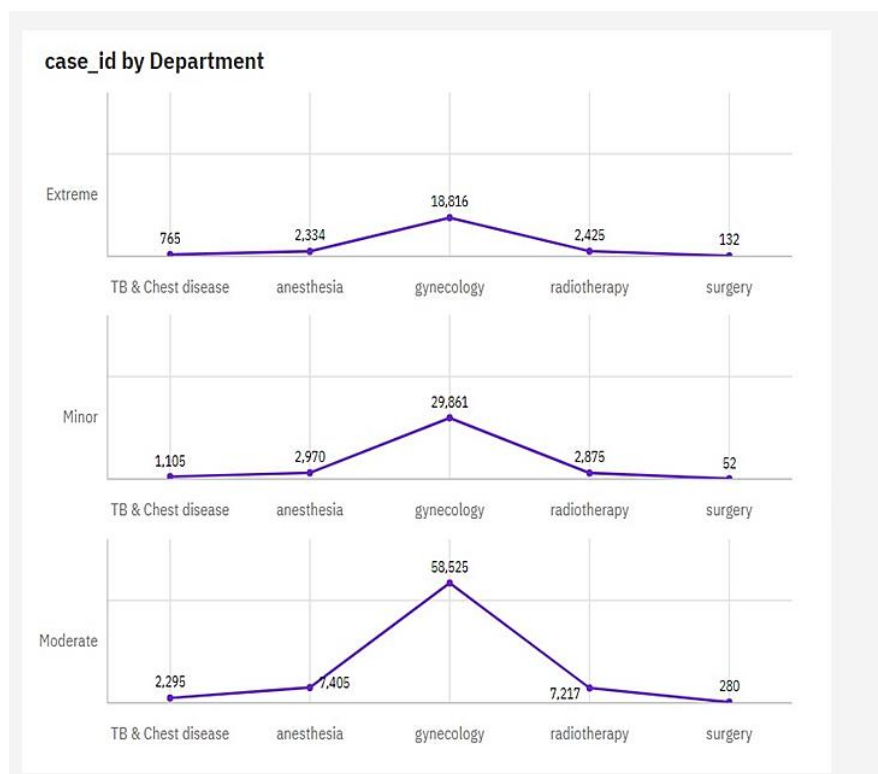
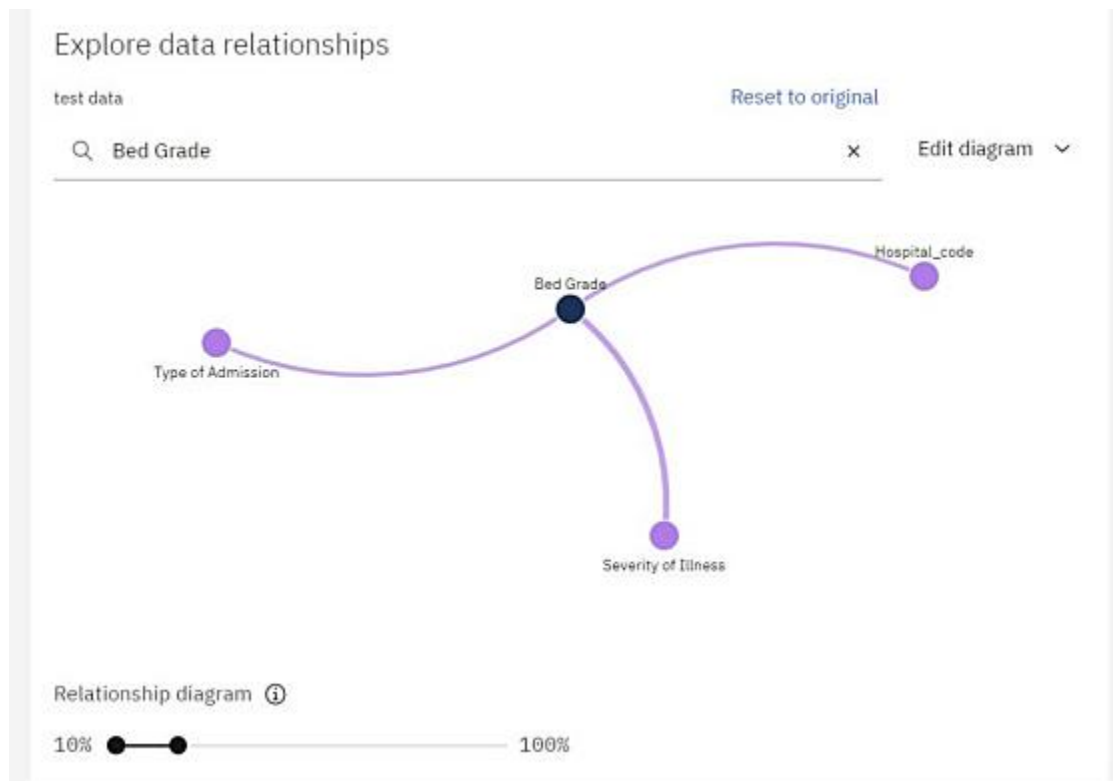
Navigation paths +

test\_data.csv

train\_data.csv

Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code
1	318439	21	c	3	Z
2	318440	29	a	4	X
3	318441	26	b	2	Y
4	318442	6	a	6	X
5	318443	28	b	11	X
6	318444	23	a	6	X
7	318445	26	b	2	Y
8	318446	25	e	1	X
9	318447	23	a	6	X
10	318448	23	a	6	X
11	318449	10	e	1	X
12	318450	4	a	4	X

- Exploring the data by using data module.



### Fields

↗ x-axis\* Required field

⋮ Department ⋮

Click or drag data here

🎨 Color

Click or drag data here

# y-axis\* Required field

⋮ case\_id ⋮

Click or drag data here

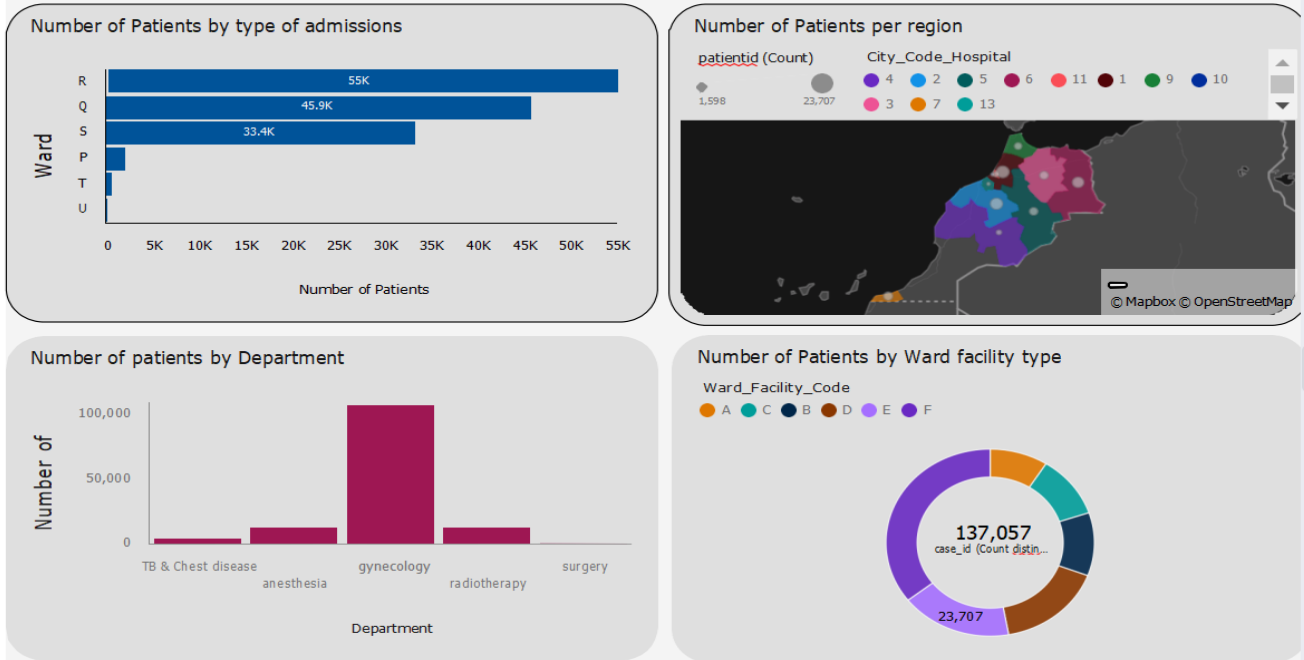
🔲 Repeat (column)

Click or drag data here

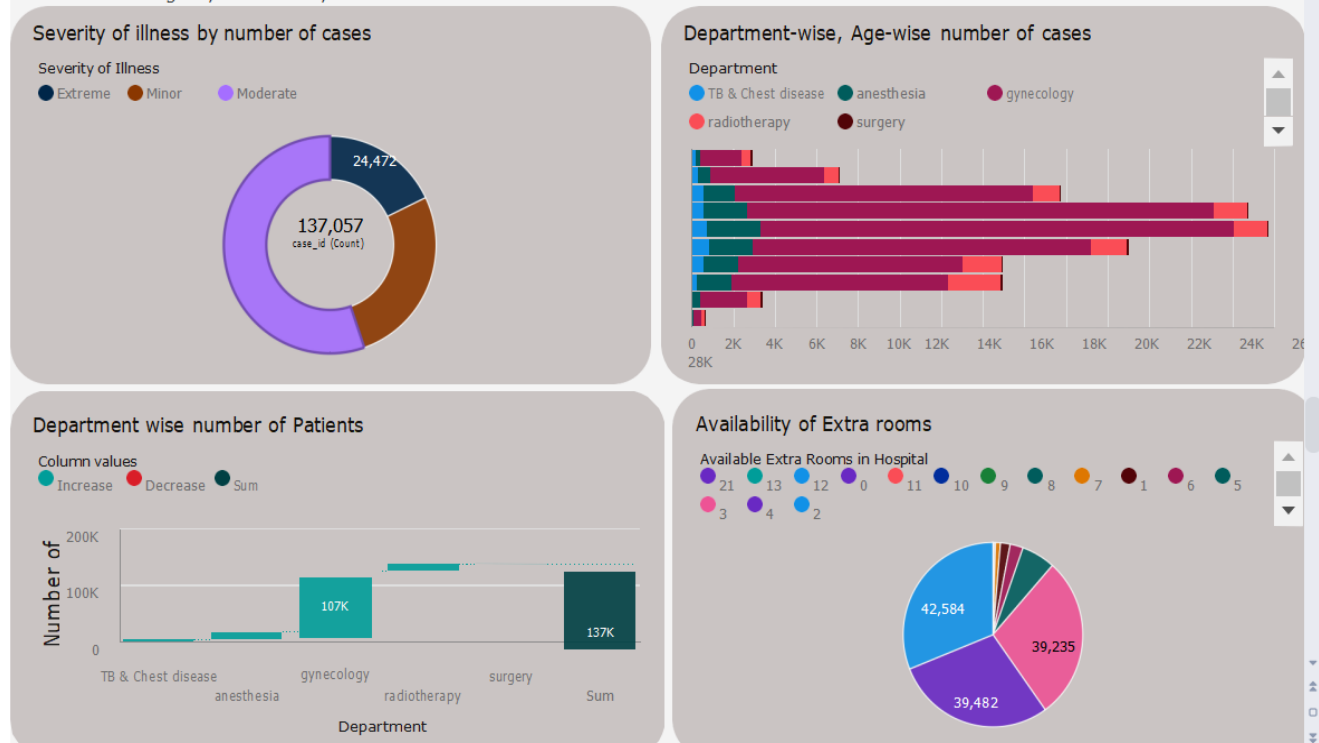


- Creation of interactive dashboards.

Dashboard to show Number of Patients

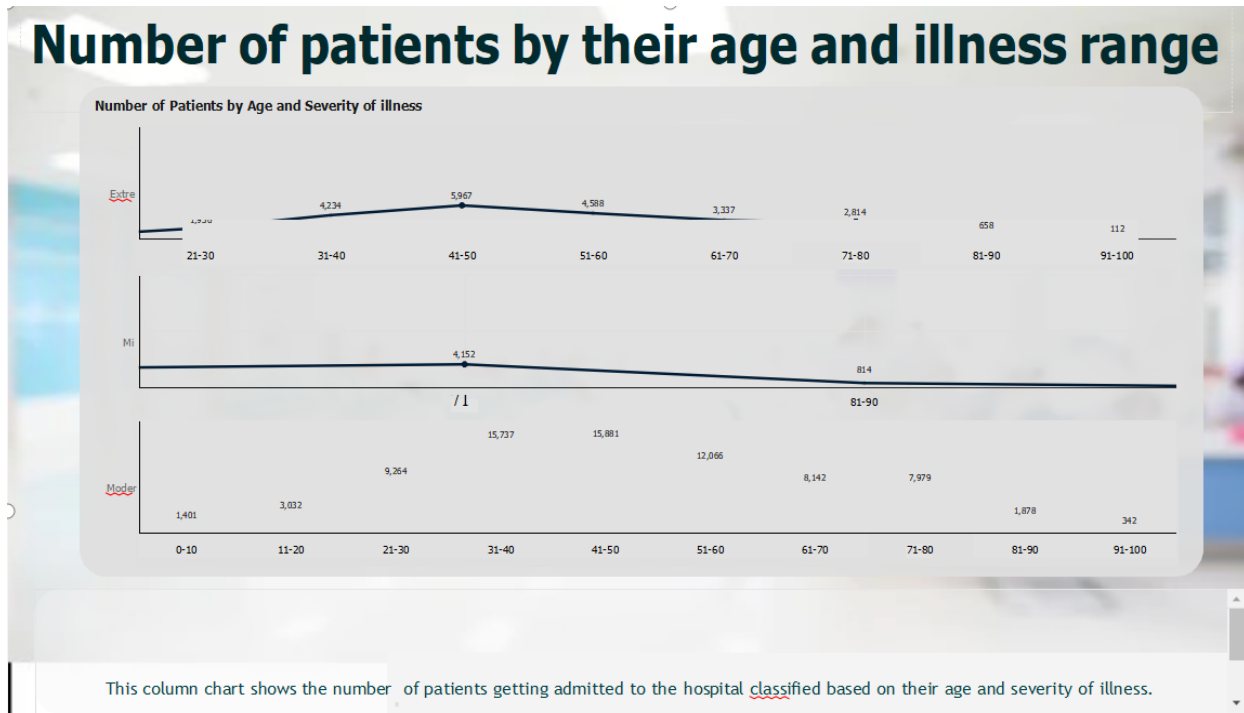


Dashboard showing Pie, Stacked bar, Waterfall and Pie charts



## 7.2 Feature 2:

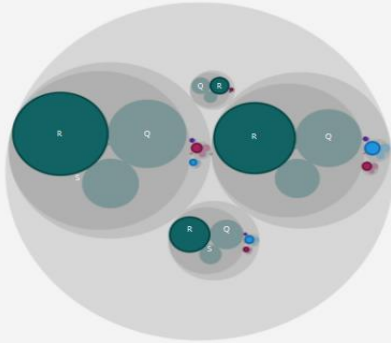
- To present the analytics and visualization results by creating a story.



# Bed grade for Patients

Bed Grade with Number of Cases by Department and Ward-wise

patients (Sum) Department  
5,610 1,150,945,531  
T3 & Chest disease anesthesia gynecology radiotherapy surgery



Ward\_Type

Ward_Type
P
Q
R
S
T
U

The hierarchy bubble shows the bed grade with number of Patients admitted in a department, ward-wise.

- To submit the results of analytics and visualization create a report.



### **7.3 Database Schema:**

- 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

## **8. TESTING:**

### **8.1 Test Cases:**

- Verify user is able to upload the data
- Verify user is able to prepare the data by creating a data module
- Verify user is able to explore the data
- Verify user is able to make dashboards
- Verify user is able to create a story
- Verify user is able to create a report

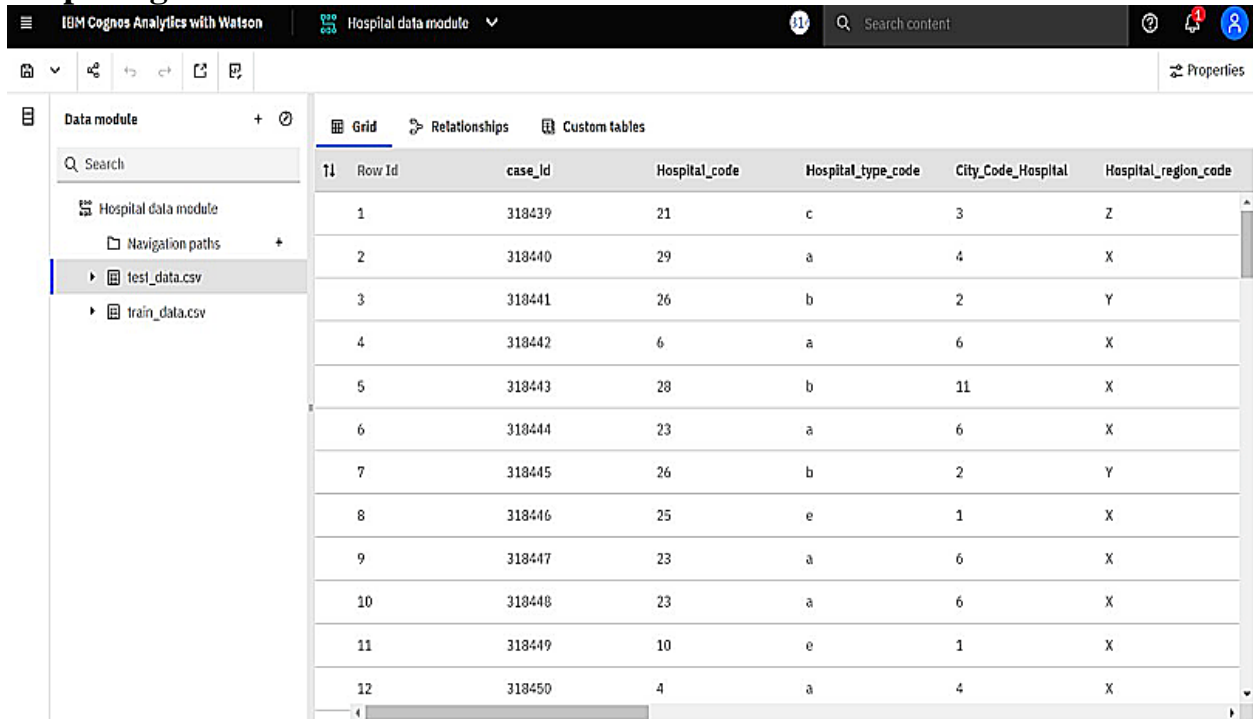
## Uploading the dataset:

The image displays two sequential screenshots of the IBM Cognos Analytics web application interface, illustrating the process of uploading a dataset.

**Top Screenshot:** The browser address bar shows the URL `us1.ca.analytics.ibm.com/bi/?perspective=home`. The top navigation bar includes the IBM Cognos Analytics logo, a search bar, and user profile icons. A blue notification banner at the top left states "Maintenance: Scheduled maintenance complete". A white progress bar in the center indicates "Reading test\_data.csv..." with "Cancel" and "Details" links. The main content area features a welcome message: "Hello. Welcome to Cognos Analytics with Watson." followed by a sub-header "You can get started right away by taking a look at our introduction video, product tour and Getting Started tab." Below this are two buttons: "Watch video" and "Take a product tour". To the right is a large graphic of interconnected blue cubes. Below the main content is a "Quick launch" section with four tiles: "Upload data" (with an upward arrow icon), "Prepare data" (with a cube icon), "Exploration" (with a magnifying glass icon), and "Present data" (with a presentation screen icon).

**Bottom Screenshot:** This screenshot shows the same interface after the upload is complete. The progress bar now displays a green status: "test\_data.csv was uploaded successfully." with "Hide" and "Details" links. The rest of the interface, including the welcome message, buttons, and quick launch tiles, remains identical to the top screenshot.

## Preparing the data:

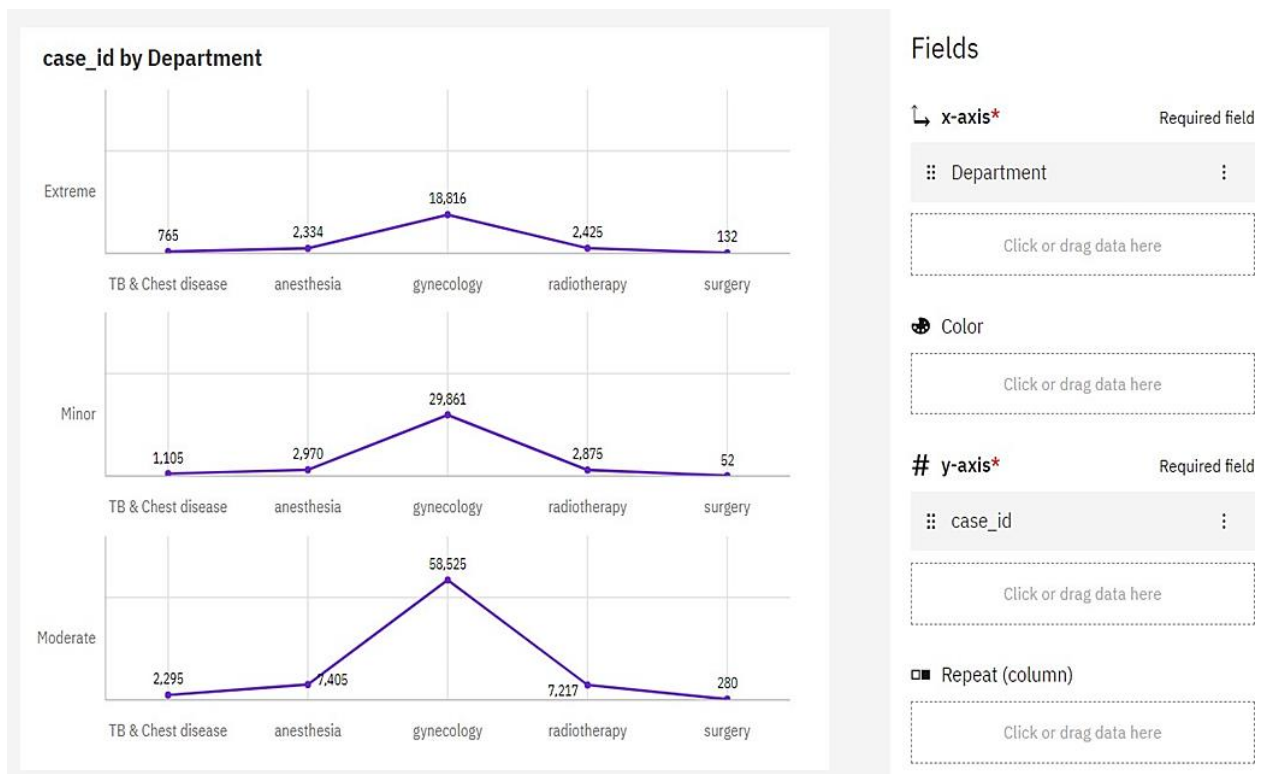


IBM Cognos Analytics with Watson | Hospital data module | Search content | Properties

Data module | Grid | Relationships | Custom tables

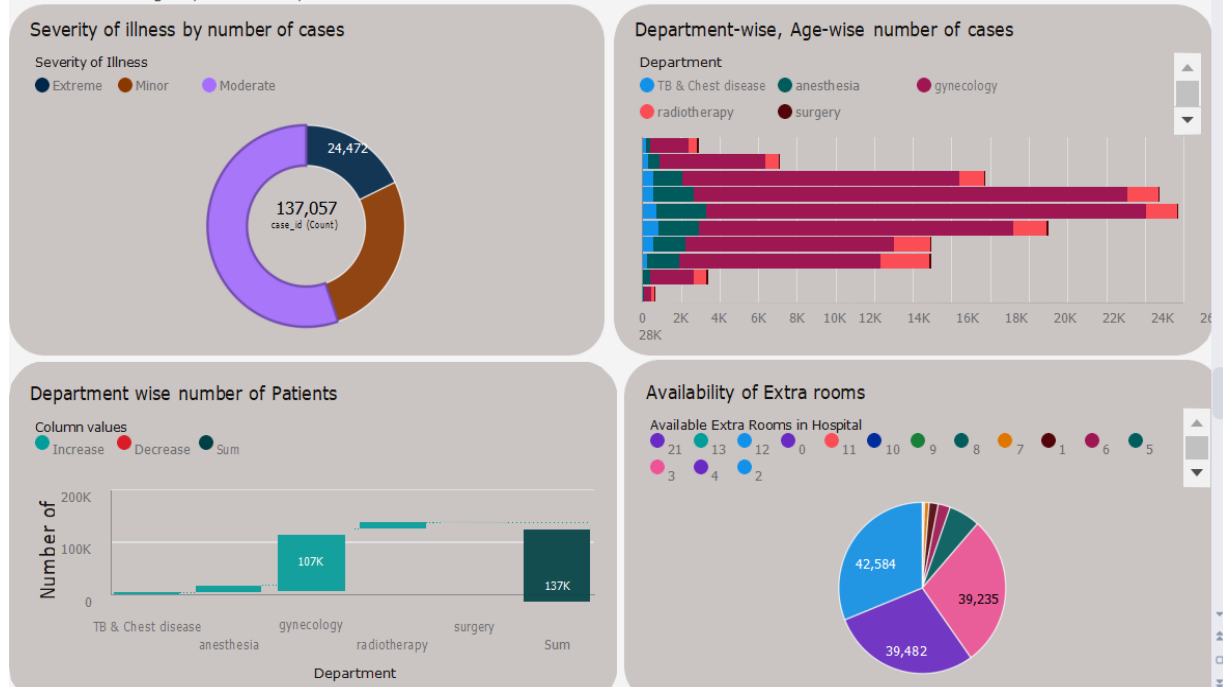
Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code
1	318439	21	c	3	Z
2	318440	29	a	4	X
3	318441	26	b	2	Y
4	318442	6	a	6	X
5	318443	28	b	11	X
6	318444	23	a	6	X
7	318445	26	b	2	Y
8	318446	25	e	1	X
9	318447	23	a	6	X
10	318448	23	a	6	X
11	318449	10	e	1	X
12	318450	4	a	4	X

## Exploration of data:



## Dashboard:

Dashboard showing Pie, Stacked bar, Waterfall and Pie charts





## Story:

The screenshot displays the IBM Cognos Analytics user interface. At the top, a dark navigation bar includes the text 'IBM Cognos Analytics with Watson', a 'Hosp story' dropdown menu, a search bar, and user profile icons. Below this, a blue banner contains a maintenance notice: 'Maintenance: Cognos Analytics will undergo Maintenance: Nov 19 10:00 UTC to bring you the latest updates to Cognos Analytics'. The main workspace features a blurred background image of a hospital corridor with the title 'Analytics for Hospitals Health-Care data' overlaid in white text. A green notification box in the top right corner states 'Your story has been saved.' The bottom of the interface shows a video player control bar with 'Prev scene', 'Next scene', 'Scene 1 of 8', and a timeline from 0:00.0 to 0:06.0.

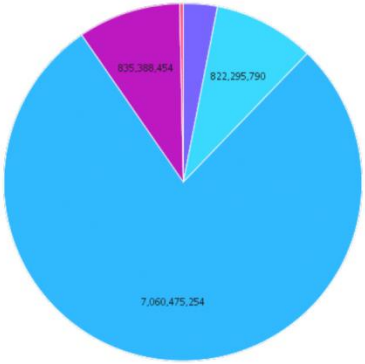
**Report (Run as PDF) :**

Hospital's Health-care analytics report

Number of Patients by departments

Department

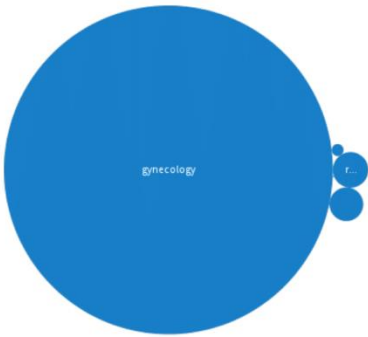
- TB & Chest disease
- radiotherapy
- gynecology
- anesthesia
- surgery



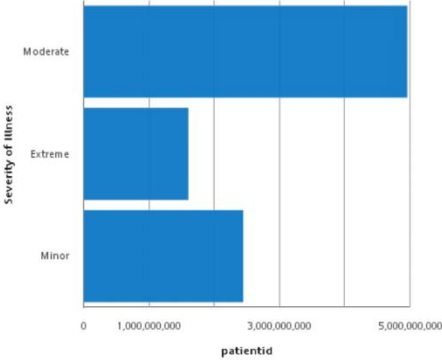
Availability of extra rooms by departments

Available Extra Rooms in Hospital

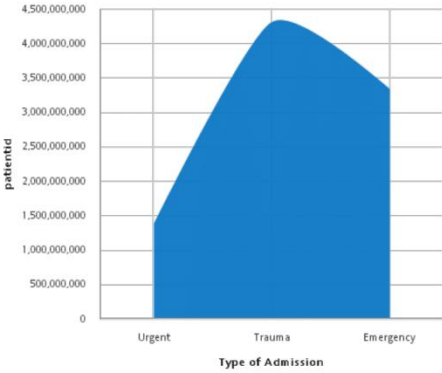
1,415 349,817



Number of patients by Severity of illness



Number of patients by type of admission



## 8.2 User Acceptance Testing:

### Purpose of Document:

The purpose of this document is to briefly explain the test coverage and open issues of the Analysis of the hospitals health-care data project at the time of the release to User Acceptance Testing (UAT).

### Defect Analysis:

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

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	1	1	1	0	3
Duplicate	0	0	0	0	0
External	1	1	0	0	2
Fixed	1	0	0	0	1
Not Reproduced	0	0	0	0	0
Skipped	1	1	1	1	4
Won't Fix	0	0	0	0	0
Totals	4	3	2	1	0

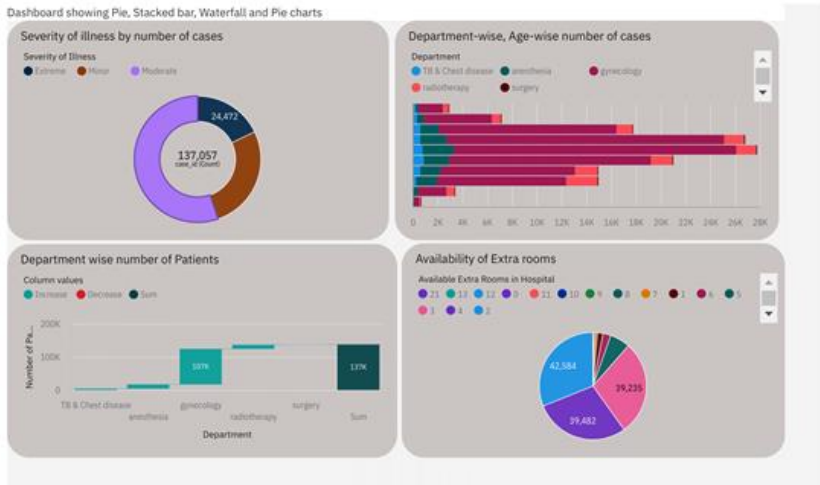
### Test Case Analysis:

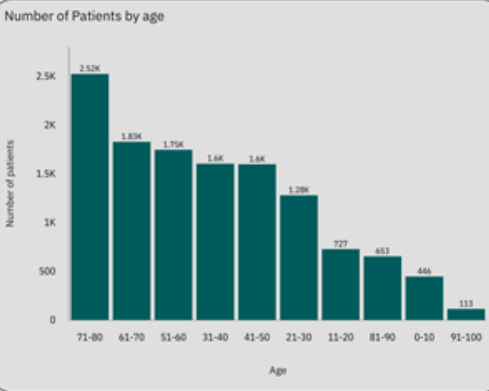
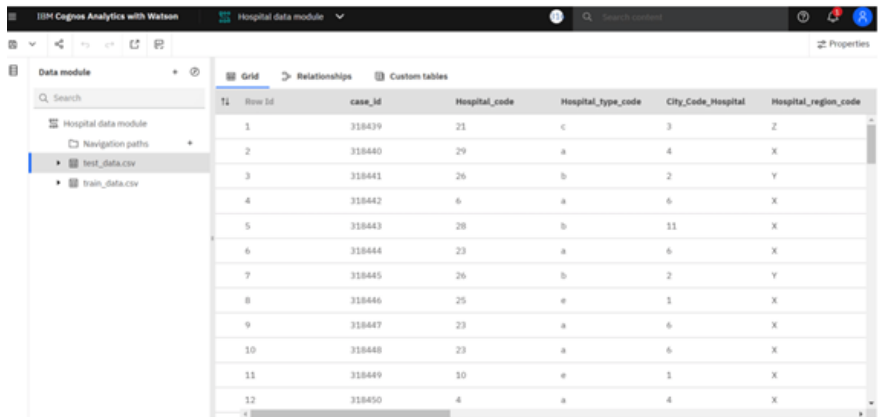
This report shows the number of test cases that have passed, failed, and untested

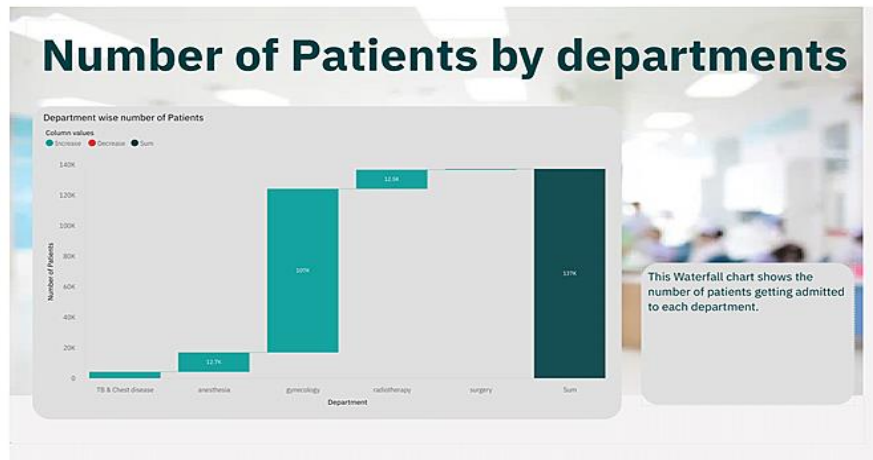
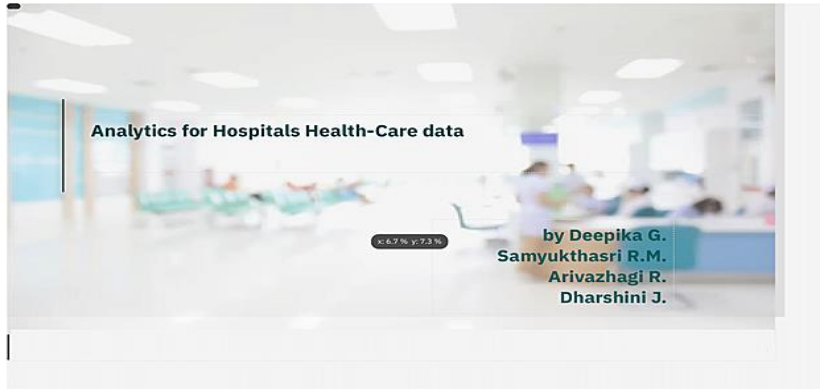
Section	Total Cases	Not Tested	Fail	Pass
Uploading the data	5	0	0	5
Preparing the data	5	0	0	5
Exploring the data	5	0	0	5
Dashboard creation	7	0	0	7
Creating a story	7	0	0	7
Creating a report	4	0	0	4

## 9.RESULTS:

### 9.1 Performance Matrics:

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	<p>No of Visualizations / Graphs – Fourteen visualizations were created using Cognos Analytics</p>  <p>The screenshot displays a dashboard with four visualizations:</p> <ul style="list-style-type: none"> <li><b>Severity of illness by number of cases:</b> A donut chart showing the distribution of illness severity. The total value is 137,057. The chart is divided into three segments: Moderate (purple), Minor (orange), and Extreme (blue).</li> <li><b>Department-wise, Age-wise number of cases:</b> A horizontal bar chart showing the number of cases across different departments and age groups. The x-axis ranges from 0 to 28K.</li> <li><b>Department wise number of Patients:</b> A vertical bar chart showing the number of patients across different departments. The y-axis ranges from 0 to 200K. The departments are TB &amp; Chest disease, anaesthesia, gynecology, radiotherapy, surgery, and Sum.</li> <li><b>Availability of Extra rooms:</b> A pie chart showing the availability of extra rooms in the hospital. The chart is divided into two main segments: 42,584 (blue) and 39,235 (pink).</li> </ul>
2.	Data Responsiveness	The visualizations can be created from the data in a fast and easy manner.

		<div><div>Age wise patients with Department and Severity</div><div><div>Age wise Patients with Department and Severity</div><div><div>Number of Patients by age</div><table><thead><tr><th>Age</th><th>Number of patients</th></tr></thead><tbody><tr><td>71-80</td><td>2.52K</td></tr><tr><td>61-70</td><td>1.83K</td></tr><tr><td>51-60</td><td>1.75K</td></tr><tr><td>41-50</td><td>1.6K</td></tr><tr><td>31-40</td><td>1.6K</td></tr><tr><td>21-30</td><td>1.28K</td></tr><tr><td>11-20</td><td>727</td></tr><tr><td>81-90</td><td>653</td></tr><tr><td>0-10</td><td>486</td></tr><tr><td>91-100</td><td>113</td></tr></tbody></table></div><div><div>Department</div><div>Department</div><div>TB &amp; Chest disease</div><div>anesthesia</div><div>gynecology</div><div>radiotherapy</div><div>surgery</div><div>Severity of Illness</div><div>Severity of Illness</div><div>Extreme</div><div>Minor</div><div>Moderate</div></div></div></div>	Age	Number of patients	71-80	2.52K	61-70	1.83K	51-60	1.75K	41-50	1.6K	31-40	1.6K	21-30	1.28K	11-20	727	81-90	653	0-10	486	91-100	113																																																								
Age	Number of patients																																																																															
71-80	2.52K																																																																															
61-70	1.83K																																																																															
51-60	1.75K																																																																															
41-50	1.6K																																																																															
31-40	1.6K																																																																															
21-30	1.28K																																																																															
11-20	727																																																																															
81-90	653																																																																															
0-10	486																																																																															
91-100	113																																																																															
3.	Amount Data to Rendered (DB2 Metrics)	<div><div>A data set with 137508 rows and 17 columns is used to analyze and create visualizations and dashboards.</div><table><thead><tr><th>Row Id</th><th>case_id</th><th>Hospital_code</th><th>Hospital_type_code</th><th>City_Code_Hospital</th><th>Hospital_region_code</th></tr></thead><tbody><tr><td>1</td><td>318439</td><td>21</td><td>c</td><td>3</td><td>Z</td></tr><tr><td>2</td><td>318440</td><td>29</td><td>a</td><td>4</td><td>X</td></tr><tr><td>3</td><td>318441</td><td>26</td><td>b</td><td>2</td><td>Y</td></tr><tr><td>4</td><td>318442</td><td>6</td><td>a</td><td>6</td><td>X</td></tr><tr><td>5</td><td>318443</td><td>28</td><td>b</td><td>11</td><td>X</td></tr><tr><td>6</td><td>318444</td><td>23</td><td>a</td><td>6</td><td>X</td></tr><tr><td>7</td><td>318445</td><td>26</td><td>b</td><td>2</td><td>Y</td></tr><tr><td>8</td><td>318446</td><td>25</td><td>e</td><td>1</td><td>X</td></tr><tr><td>9</td><td>318447</td><td>23</td><td>a</td><td>6</td><td>X</td></tr><tr><td>10</td><td>318448</td><td>23</td><td>a</td><td>6</td><td>X</td></tr><tr><td>11</td><td>318449</td><td>10</td><td>e</td><td>1</td><td>X</td></tr><tr><td>12</td><td>318450</td><td>4</td><td>a</td><td>4</td><td>X</td></tr></tbody></table></div>	Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	1	318439	21	c	3	Z	2	318440	29	a	4	X	3	318441	26	b	2	Y	4	318442	6	a	6	X	5	318443	28	b	11	X	6	318444	23	a	6	X	7	318445	26	b	2	Y	8	318446	25	e	1	X	9	318447	23	a	6	X	10	318448	23	a	6	X	11	318449	10	e	1	X	12	318450	4	a	4	X
Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code																																																																											
1	318439	21	c	3	Z																																																																											
2	318440	29	a	4	X																																																																											
3	318441	26	b	2	Y																																																																											
4	318442	6	a	6	X																																																																											
5	318443	28	b	11	X																																																																											
6	318444	23	a	6	X																																																																											
7	318445	26	b	2	Y																																																																											
8	318446	25	e	1	X																																																																											
9	318447	23	a	6	X																																																																											
10	318448	23	a	6	X																																																																											
11	318449	10	e	1	X																																																																											
12	318450	4	a	4	X																																																																											
4.	Utilization of Data Filters	<div><div>Data filtering options were used while analyzing the data and making visualizations.</div></div>																																																																														
5.	Effective User Story	<div><div>No of Scene Added – Eight scenes were added to the story to present the visualizations.</div></div>																																																																														



6. Descriptive Reports

No of Visualizations / Graphs – Five visualizations are added to the report.



## **10.ADVANTAGES & DISADVANTAGES:**

### **Advantages:**

- Allows for better resource planning.
- Can help identify patients at-risk.
- Allows for earlier interventions .
- Reduces the hospital expenditures.
- No excess usage of resources from hospitals.

### **Disadvantage:**

- The result might be less accurate sometimes.
- It may lead to over or under-treatment.

## **11.CONCLUSION:**

This project shows that reductions in hospital costs and an improvement in quality patient care are possible. We prove that estimating the length of stay is possible at the hospital level. This means that it is possible to assess which departments are a better choice to save costs. The improvement in the success rates when predicting the LOS in some hospital departments is interesting. This helps avoid unnecessary effort and provides a contextualised vision to hospitals about which departments are susceptible to cost and resource-saving options. Also, it would be interesting to analyse the length of stay after suffering a big pandemic as COVID-19 to check how it affects to this feature (LOS) and to find out how this fact affects to hospital departments.

## **12.FUTURE SCOPE:**

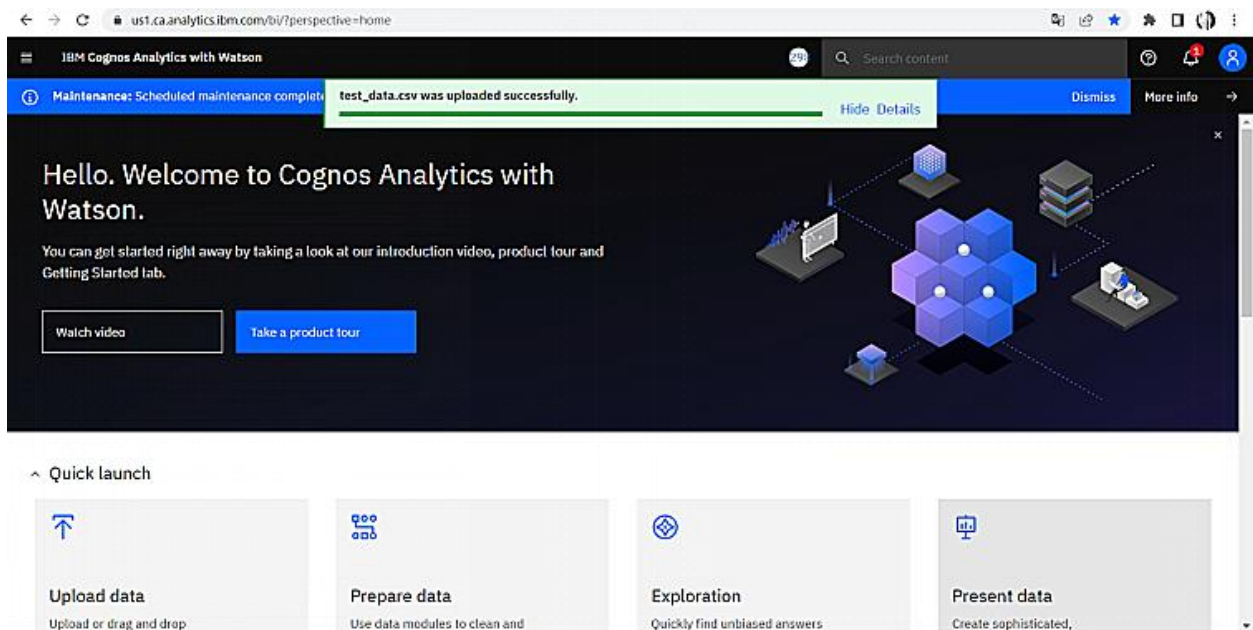
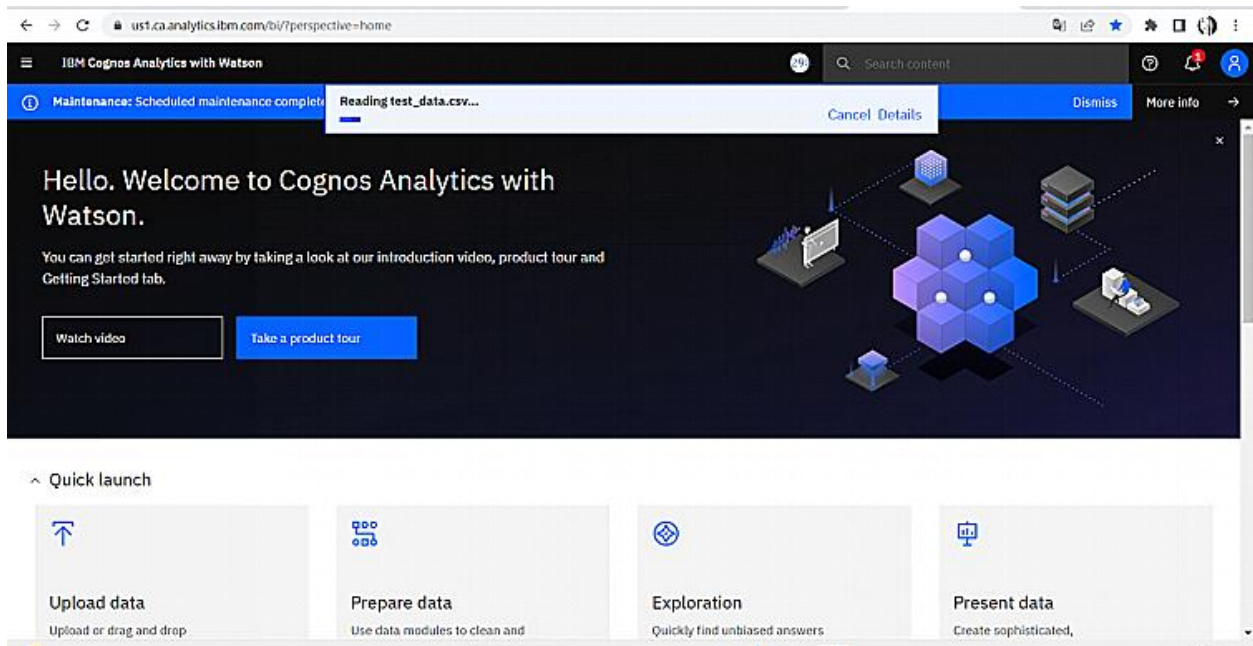
The future scope of data analytics is very promising. With the advent of new technologies and the increasing demand for data-driven decision making, organization will require more sophisticated data analytics tools and techniques. Our analysis also offers directions to future research in those hospital departments which are the most appropriate ones to predict LOS.



## 13.APPENDIX:

### 13.1 Source Code:

### Uploading dataset in Cognos:



## Preparing the Data:

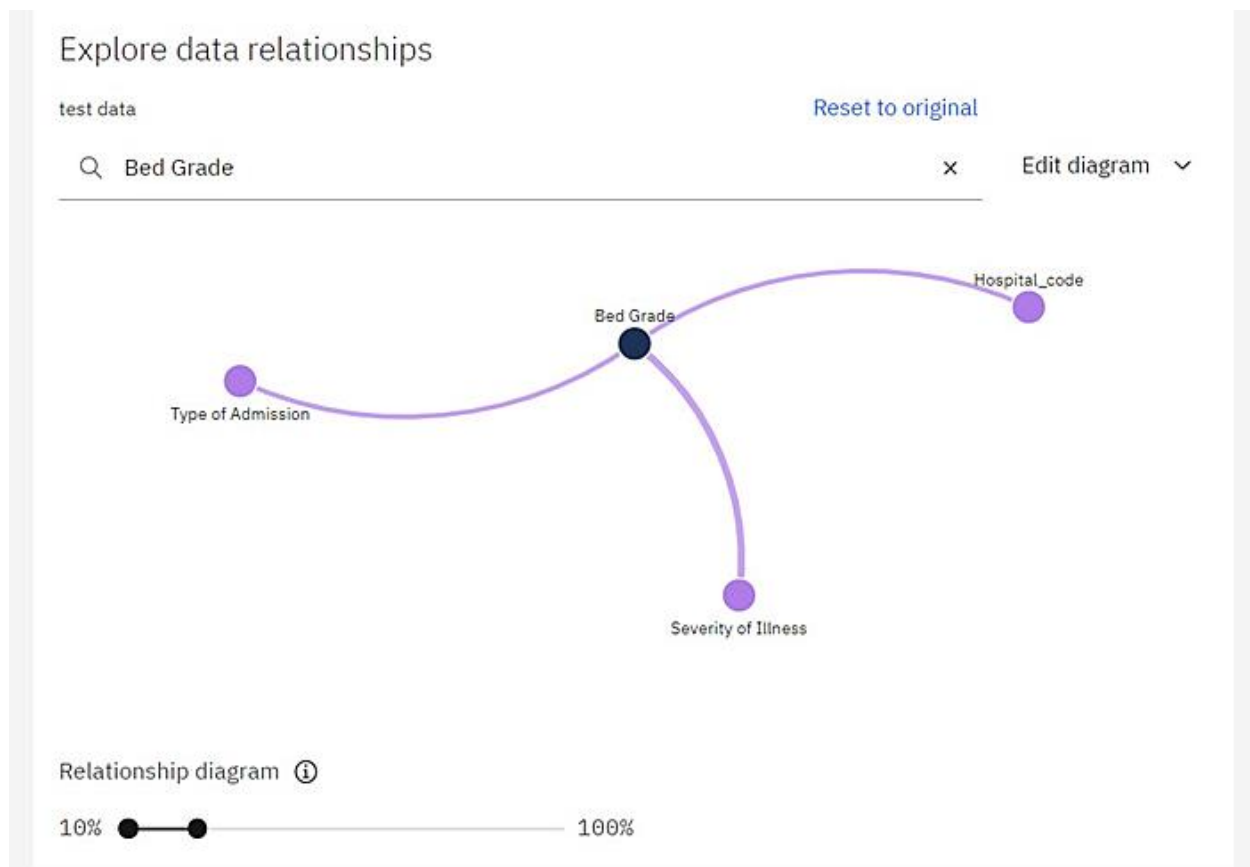
The screenshot shows the IBM Cognos Analytics interface. The top navigation bar includes the logo, 'Hospital data module', a search bar, and user icons. A blue banner at the top indicates 'Maintenance: Scheduled maintenance completed. Click More Info for details and to subscribe to future events'. Below this is a toolbar with icons for various actions. The left sidebar shows the 'Data module' structure with 'test\_data.csv' and 'train\_data.csv' listed. The main area displays a 'Grid' view of a table with the following data:

Type of Admission	Severity of Illness	Visitors with Patient	Age	Admission_Deposit	Stay
Emergency	Extreme	2	51-60	4911	0-10
Trauma	Extreme	2	51-60	5954	41-50
Trauma	Extreme	2	51-60	4745	31-40
Trauma	Extreme	2	51-60	7272	41-50
Trauma	Extreme	2	51-60	5558	41-50
Trauma	Extreme	2	51-60	4449	11-20
Emergency	Extreme	2	51-60	6167	0-10
Trauma	Extreme	2	51-60	5571	41-50
Trauma	Extreme	2	51-60	7223	51-60
Trauma	Extreme	2	51-60	6056	31-40
Urgent	Extreme	2	51-60	5797	21-30

The screenshot shows the IBM Cognos Analytics interface. The top navigation bar includes the logo, 'Hospital data module', a search bar, and user icons. A blue banner at the top indicates 'Maintenance: Scheduled maintenance completed. Click More Info for details and to subscribe to future events'. Below this is a toolbar with icons for various actions. The left sidebar shows the 'Data module' structure with 'test\_data.csv' and 'train\_data.csv' listed. The main area displays a 'Grid' view of a table with the following data:

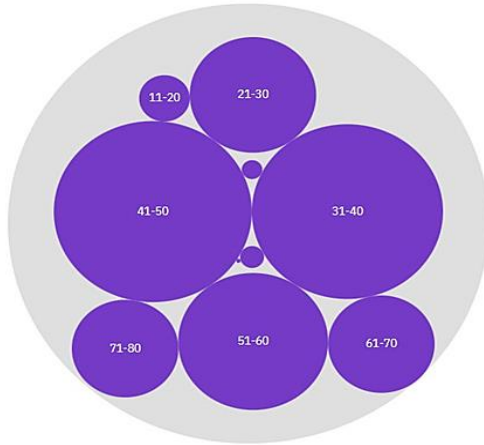
Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code
1	318439	21	c	3	Z
2	318440	29	a	4	X
3	318441	26	b	2	Y
4	318442	6	a	6	X
5	318443	28	b	11	X
6	318444	23	a	6	X
7	318445	26	b	2	Y
8	318446	25	e	1	X
9	318447	23	a	6	X
10	318448	23	a	6	X
11	318449	10	e	1	X
12	318450	4	a	4	X

## Exploration of Data:

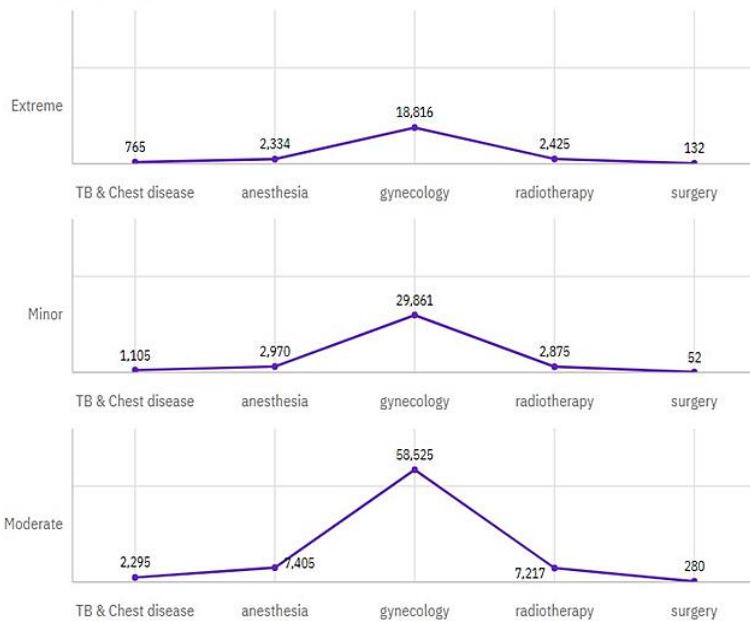


### Age, case\_id

case\_id (Count dis...  
605 27,746



### case\_id by Department



### Fields

**Bubbles\*** Required field

Age :

Click or drag data here

**Size**

case\_id :

**Color**

Click or drag data here

**Repeat (column)**

Click or drag data here

**Repeat (row)**

### Fields

**x-axis\*** Required field

Department :

Click or drag data here

**Color**

Click or drag data here

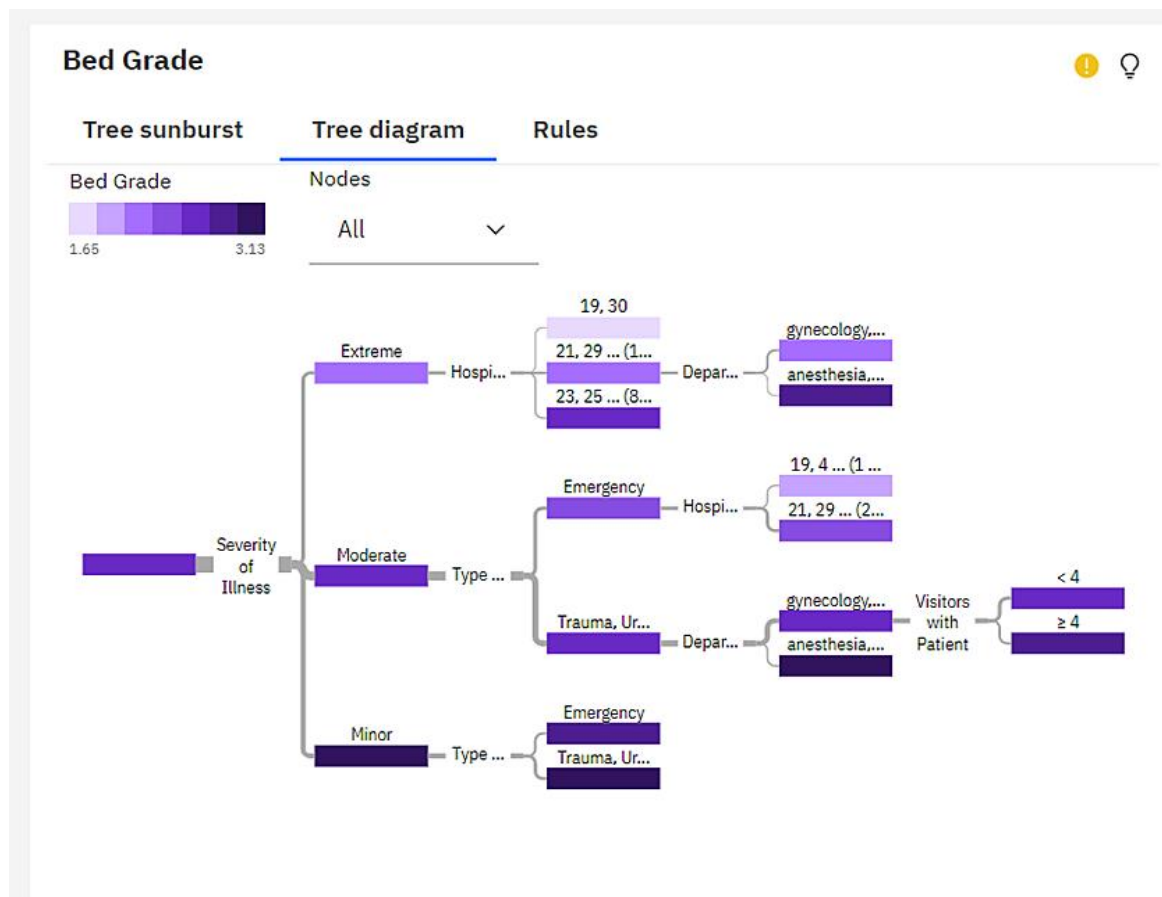
**y-axis\*** Required field

case\_id :

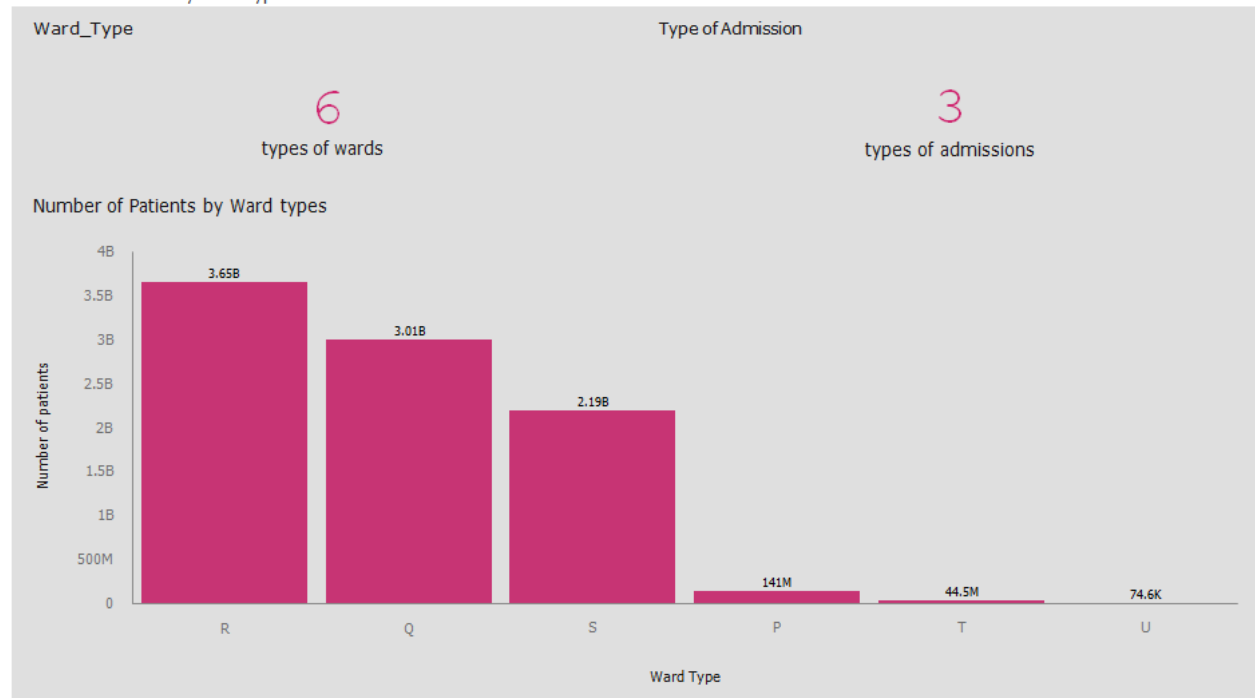
Click or drag data here

**Repeat (column)**

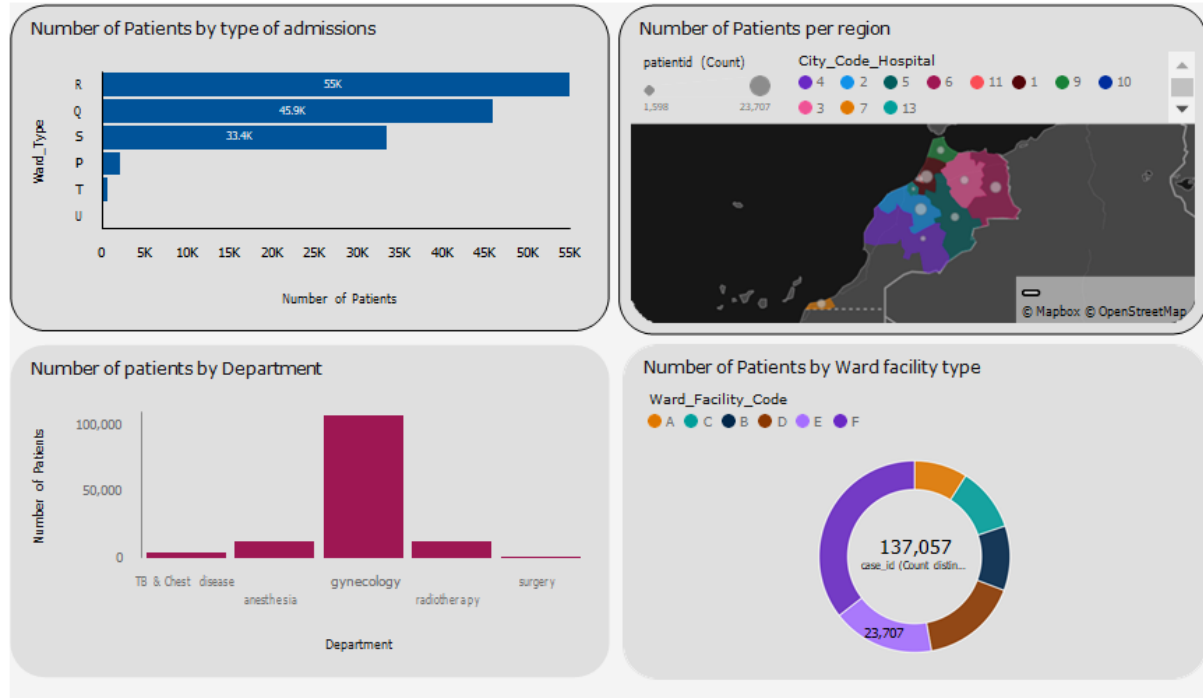
Click or drag data here



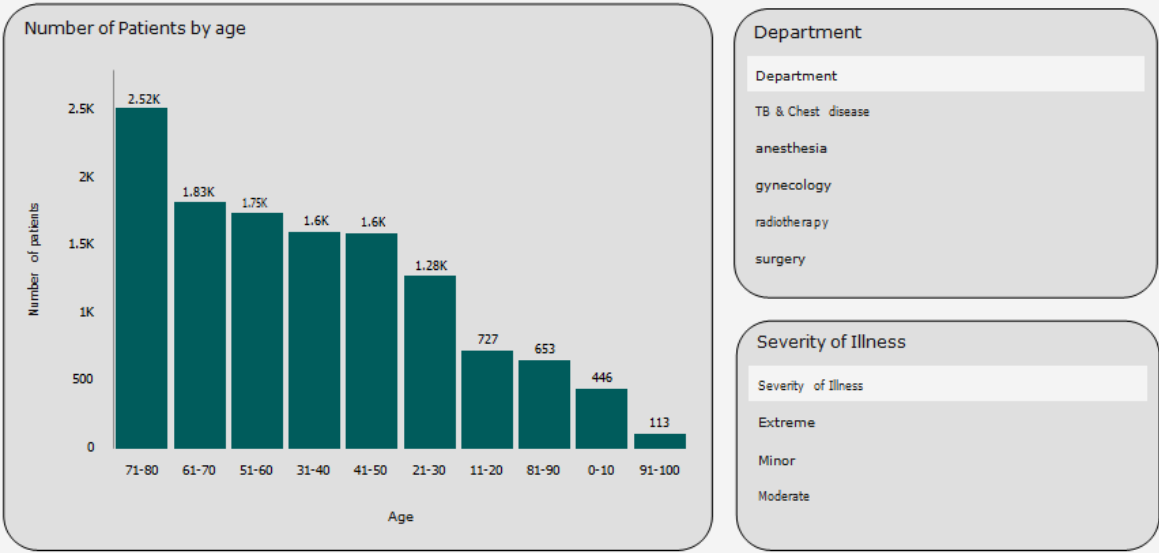
## Number of Patients by Ward types



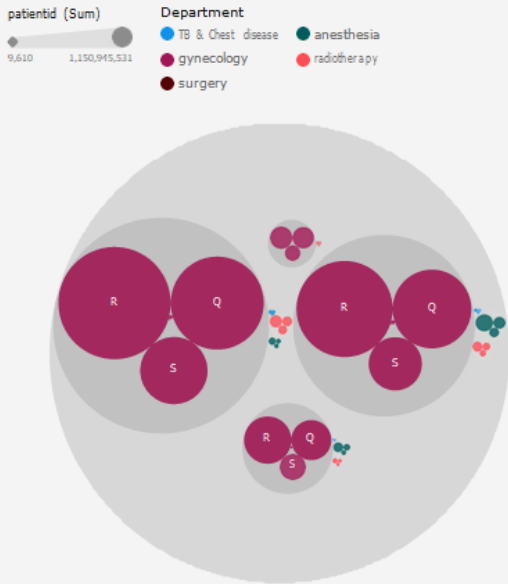
## Dashboard to show Number of Patients



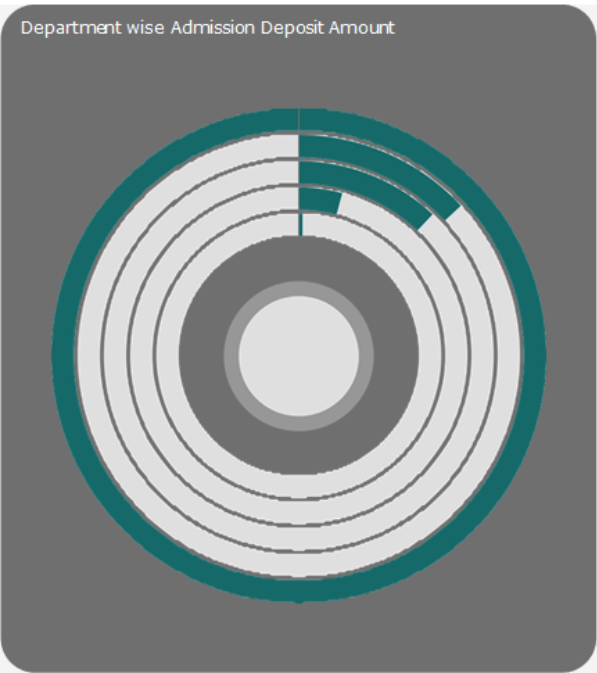
# Age wise Patients with Department and Severity



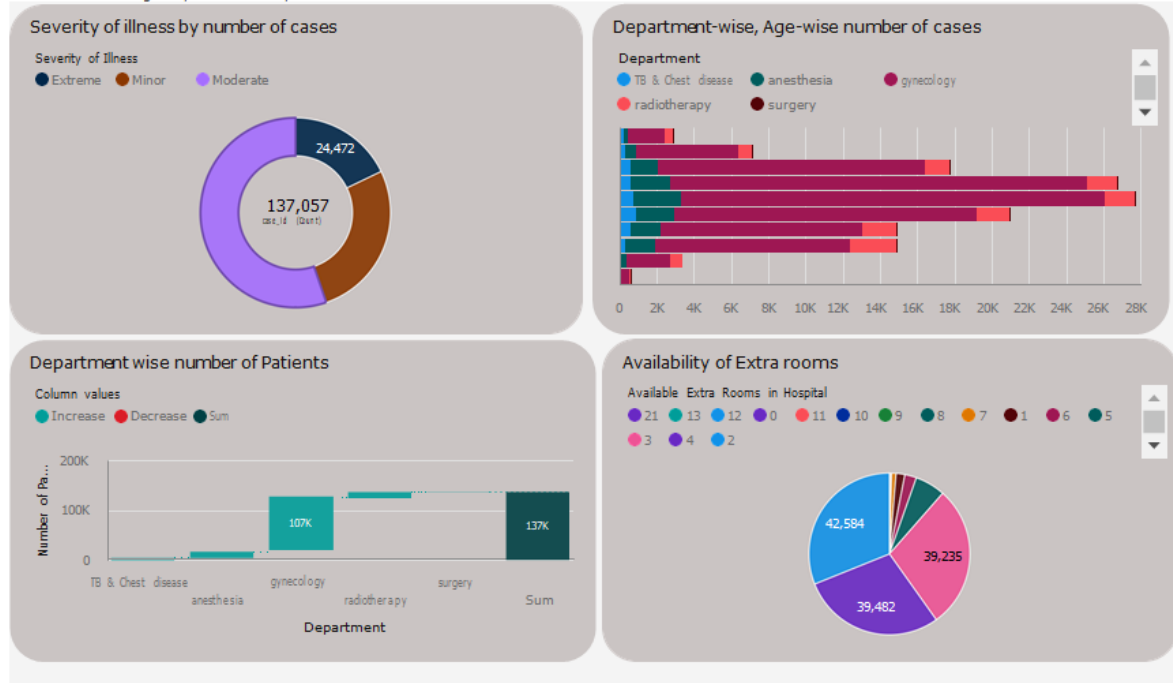
## Bed Grade with Number of Cases by Department and Ward-wise



## Department wise Admission Deposit Amount



Dashboard showing Pie, Stacked bar, Waterfall and Pie charts



Story:

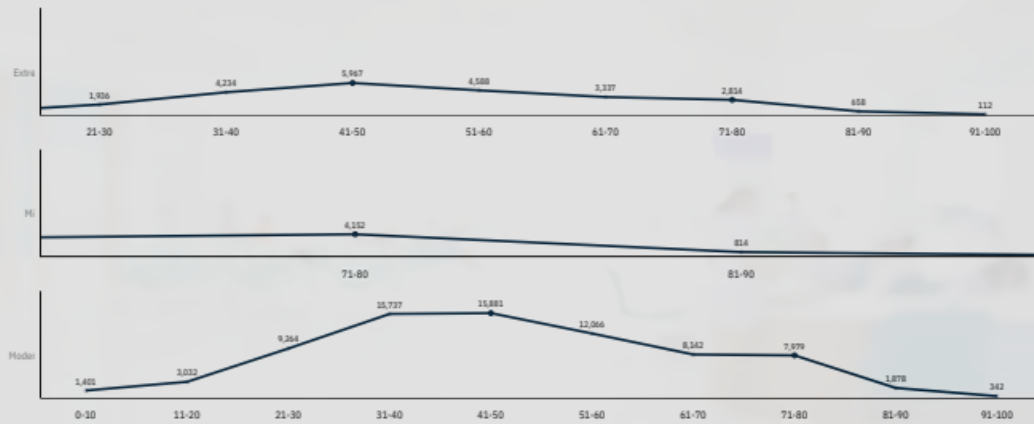
## Predicting the Length of the stay of the Patients

- The goal is to accurately predict the Length of Stay for each patient on case by case basis.
- 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.



## Number of patients by their age and illness range

Number of Patients by Age and Severity of illness

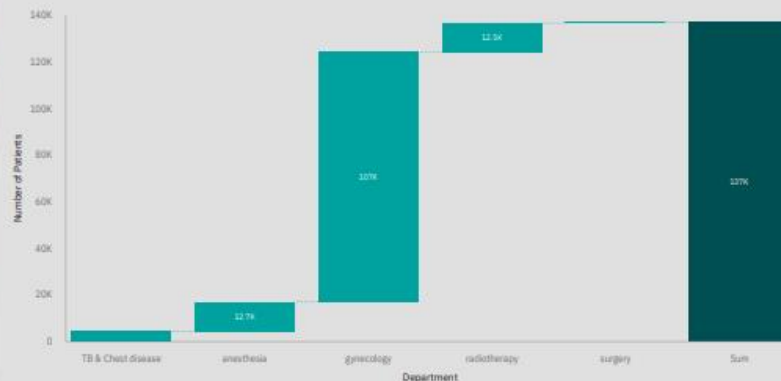


This column chart shows the number of patients getting admitted to the hospital classified based on their age and severity of illness.

## Number of Patients by departments

Department wise number of Patients

Column values  
● Increase ● Decrease ● Sum

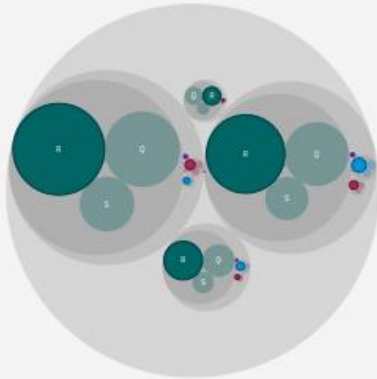


This Waterfall chart shows the number of patients getting admitted to each department.

# Bed grade for Patients

Bed Grade with Number of Cases by Department and Ward-wise

patientid (Sum) Department  
● TB & Chest disease ● anesthesia ● gynecology ● radiotherapy  
● surgery



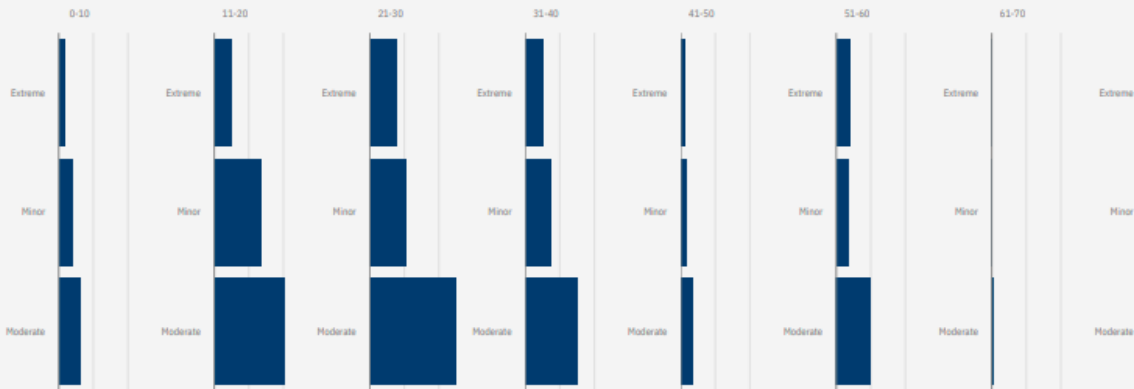
Ward\_Type

Ward_Type
P
Q
R
S
T
U

The hierarchy bubble shows the bed grade with number of Patients admitted in a department, ward-wise.

## Length of stay of Patients by their illness range

Number of patients by their Severity of illness and length of stay



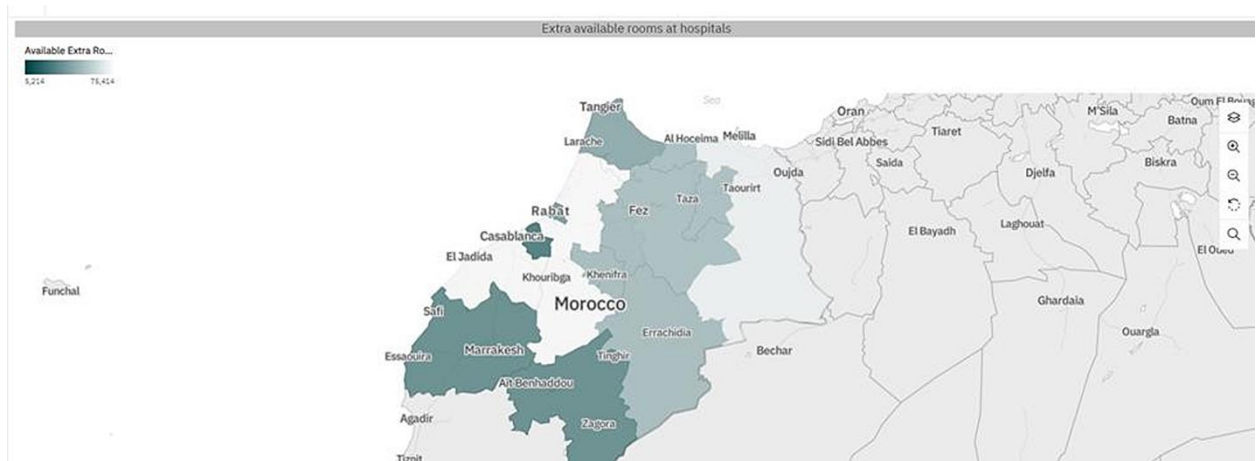
This bar charts depicts the number of patients by their Severity of illness and their Length of stay.

# Conclusion

- Based on these, the Length of Stay of Patients is identified. Once identified, beds and resources could be allocated to the departments and wards.
- Patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection.

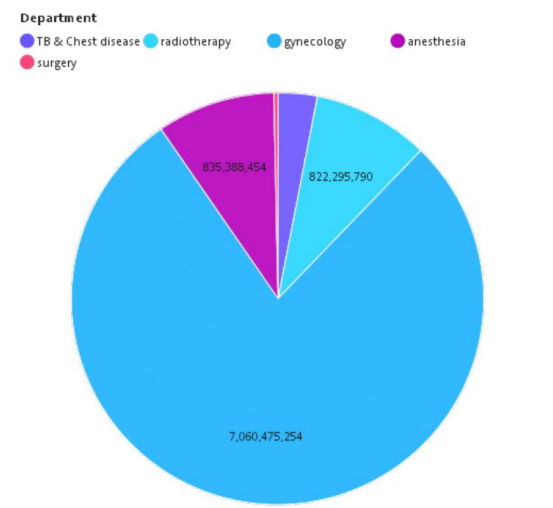
**THANK YOU !**

**Report (Run as pdf):**

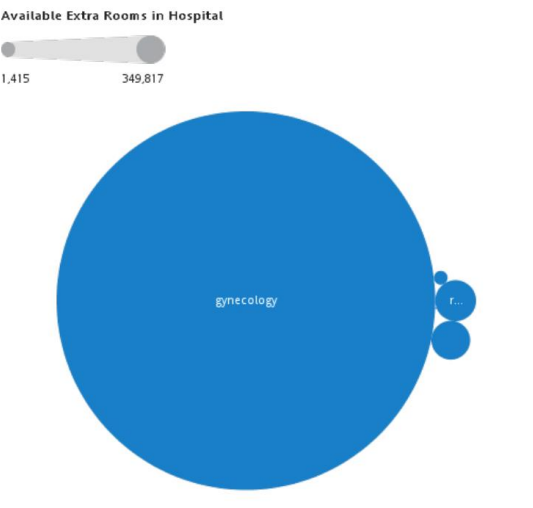


Hospitals Health-care analytics report

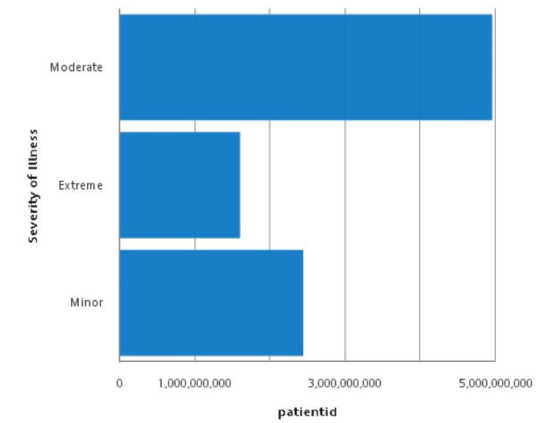
Number of Patients by departments



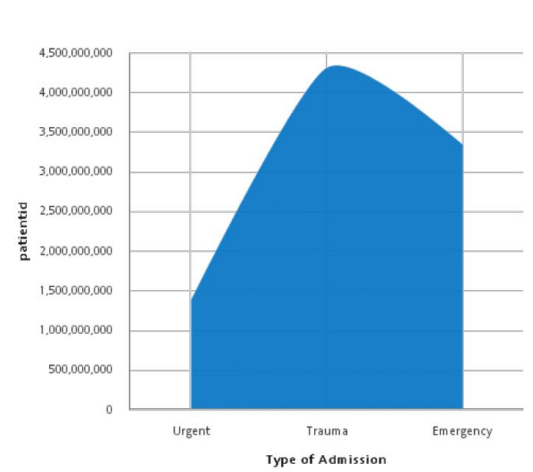
Availability of extra rooms by departments



Number of patients by Severity of illness



Number of patients by type of admission



## **Link:13.2 GitHub & Project Demo**

**GitHub link:**

**<https://github.com/IBM-EPBL/IBM-Project-47980-1660803720>**

**Project Demo Link:**

**[https://drive.google.com/file/d/1Ih3oLmJqvU24rzu1R5XfW6k-QwX\\_-](https://drive.google.com/file/d/1Ih3oLmJqvU24rzu1R5XfW6k-QwX_-IPa/view?usp=drivesdk)**

**[IPa/view?usp=drivesdk](https://drive.google.com/file/d/1Ih3oLmJqvU24rzu1R5XfW6k-QwX_-IPa/view?usp=drivesdk)**