INTRODUCTION
PROJECT DESIGN AND PLANNING
IDEATION PHASE

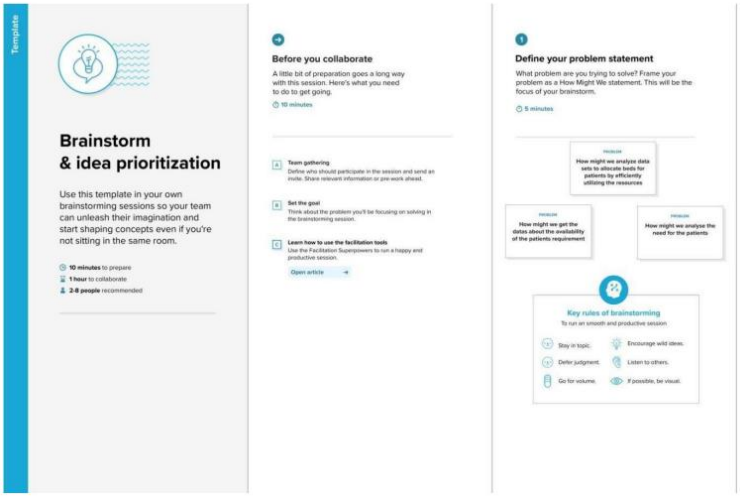
ANALYTICS FOR HOSPITALS HEALTH-CARE DATA

ABSTRACT:

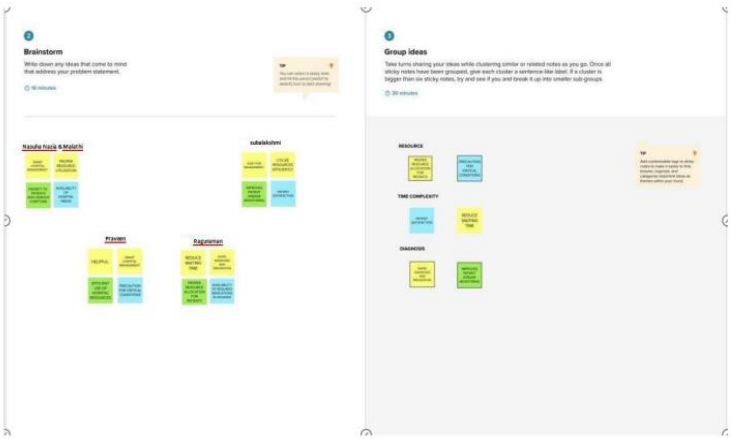
Recent Covid-19 Pandemic has raised alarms over one of the most overlooked are as to focus which is absolutely Healthcare Data Management. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiencyof 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 resourceallocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10 days to more than 100 days. The tools that we are using for data analytics is Cognos Analytics.

Brainstorm & Idea Prioritization Template

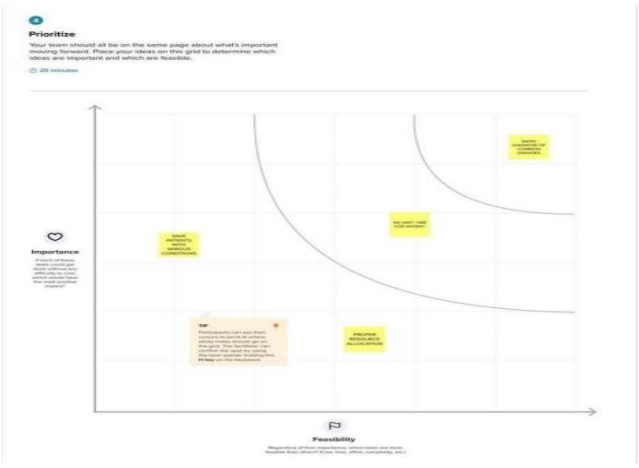
Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



LITERATURE SURVEY

Analytics For Hospitals' Health-Care Data

INTRODUCTION: The healthcare costs have been constantly rising, the quality of care provided tothe patients in the many states have not seen considerable improvements. Recently, several researchers have conducted studies which showed that by incorporating the current healthcare technologies, they are able to reduce mortality rates, healthcare costs and medical complications at various hospitals. The recent advances in information technology have led to an increasing ease inthe ability to collect various forms of healthcare data. Data analytics, in particular, forms a critical component of these computing technologies.

Title: Data analytics in healthcare: Management, Analysis and Future prospects. J BigData Volume - 6, 19th June 2019. Healthcare is a multi-dimensional system established with the sole aim for the prevention, diagnosis, and treatment of health-related issues or impairments in human beings. The combined pool of data from healthcare organizations and biomedical researchers has resulted in a better outlook, determination, and treatment of various diseases. This has also helped in building a better and healthier personalized healthcare framework. Modern healthcare fraternity has realized the potential of big data and therefore, have implemented big data analytics in healthcare and clinical practices.

Algorithm: Artificial intelligence (AI) algorithm, Novel fusion algorithm.

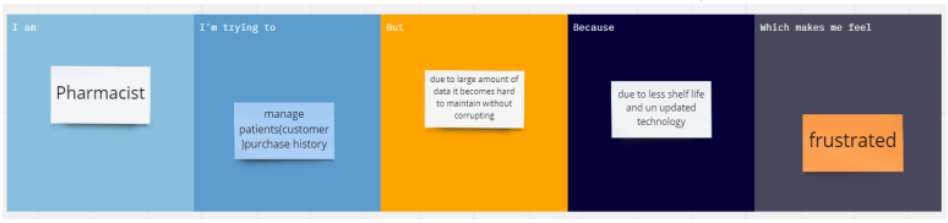
Advantages: The analyses investigate methods of improving the provision of clinical care, enhancing disease prevention, and measuring the effectiveness of various treatment options.

Disadvantages: One of the major drawbacks in the application of big data in healthcare industry is the issue of lack of privacy.

Title: Electronic Health Records in Chiropractic Practice: Common Challenges and Solutions, Journal of Chiropractic Humanities, Volume 24, Issue 1, 5th 2017, ISSN 1556-3499. The purpose of this study was to review the literature on current challenges and propose solutions for the optimal utilization of the electronic health records (EHRs) in chiropractic practice. Algorithm: Semi-Supervised Machine Learning Advantage: Medical and office staff no longer has to waste time sorting through cumbersome paper records. Users can access electronic health records quickly and efficiently with just a few strokes on a keyboard. Disadvantage: As with just about every computer network these days, EHR systems are vulnerable to hacking, which means sensitive patient data could fall into the wrong hands. Title: Big data analytics for drug discovery,IEEE International Conference on Bioinformatics and Biomedicine, September 2013. Drug discovery is related to big data analytics as the process may require the collection, processing and analysis of extremely large volume of structured and unstructured biomedical data stemming from a wide range of experiments and surveys collected by hospitals, laboratories, pharmaceutical companies or even social media. These data may include sequencing and gene expression data, drug data including molecular data, protein and drug interaction data, clinical trial and electronic patient record data, patient behaviour and self-reporting data in social media, regulatory monitoring data, and literatures where trends and drug repurposing and protein-protein interaction data may be found.
Algorithms: K- means clustering
Advantage: The use of healthcare databases in the evaluation of medical products has become very popular in recent years.
Disadvantage: A drawback of clinical trials is that they are highly controlled and highly monitored to ensure strict adherence to protocol; however, that's not how people take drugs in the real world.

Title: Systematic review of clinical prediction models to support the diagnosis of asthma in primary care. NPJ primary care respiratory medicine vol. 29. 9 th May 2019 Diagnosing asthma is challenging. Misdiagnosis can lead to untreated symptoms, incorrect treatment and avoidable deaths. The best combination of clinical features and tests to achieve a diagnosis of asthma is unclear. As asthma is usually diagnosed in non-specialist settings, a clinical prediction model to aid the assessment of the probability of asthma in primary care may improve diagnostic accuracy. We aimed to identify and describe existing prediction models to support the diagnosis of asthma in children and adults in primary care.
Algorithm: KNN algorithm
Advantage: A validated clinical prediction model for asthma diagnosis could help healthcare professionals improve the accuracy of a diagnosis by guiding decisionmaking and reducing variability between clinicians.
Disadvantage: Variables used in the clinical prediction model were not clearly reported, or unavailable in routine clinical practice

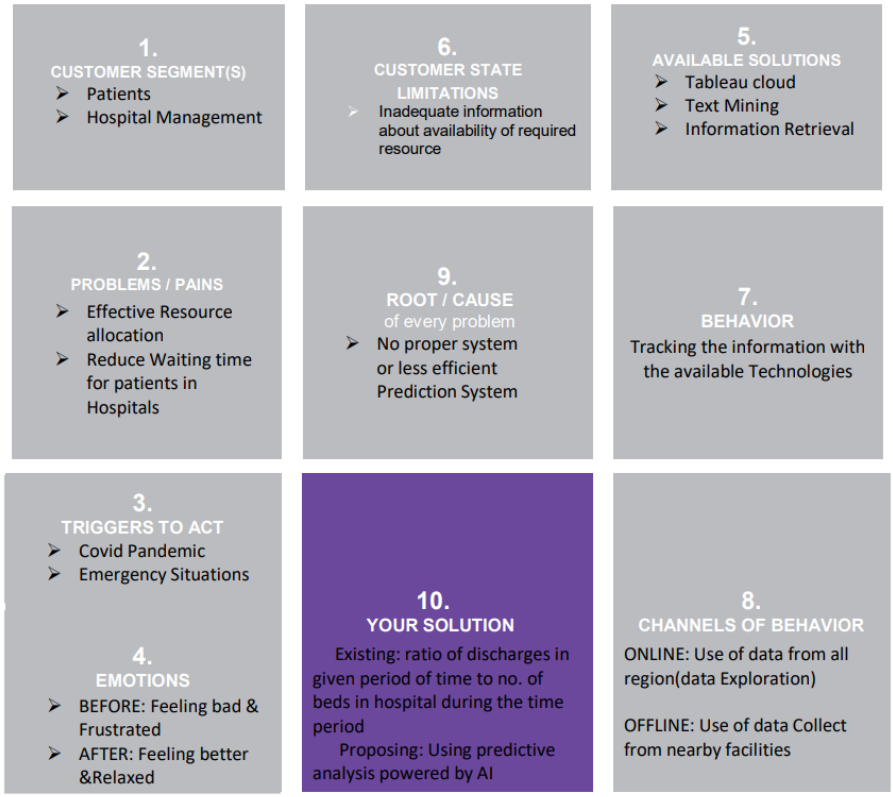
PROBLEM STATEMENT:



Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS 1	Pharmacist	Manage patients(customer) Purchase history	Due to large amount of data it becomes to maintain without corrupting	Due to less shell life and un updated technology	frustrated
PS 2	a patient suffering from covid-19	get bed in a hospital for treatment	at this pandemic time it is difficult for me to get admitted	of huge number of patients got infected with covid-19 that made hospitals run out of beds	very bad and frustrated

PROJECT DESIGN PHASE – 1

PROBLEM SOLUTION FIT



Project Design Phase-I
Proposed Solution Template

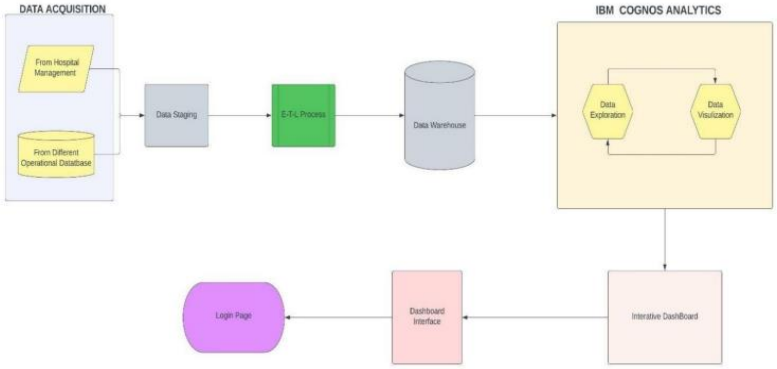
Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	During the covid-19 pandemic, we have faced one of the difficult times of our life. Everyone seeks to survive from the great disaster. At the time of pandemic, noone get to know about which hospital has vacant beds(free beds) to admit themselves or others infected by covid. This situation made the death rate higher.
2.	Idea / Solution description	Predictive analytics can create patient journey dashboards and disease trajectories that helps us to know about the patient's period of stay. It improves effective allocation of beds and other resources, treatment delivery, improves efficiencies, and so on.

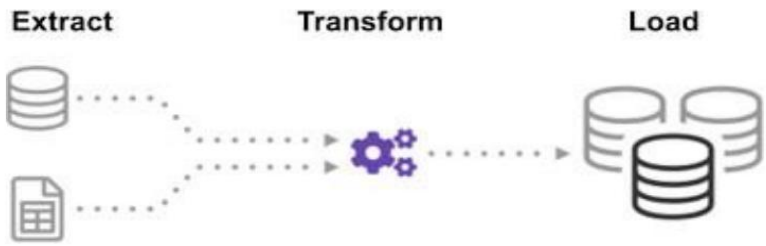
3.	Novelty / Uniqueness	Healthcare data frequently resides in several locations. The Collected data should be stored in central system(like centralized storage). This data becomes accessible and usable when it is combined into a single, central system, such as an enterprise data warehouse (EDW). Uniqueness of our project is that we can able to use data for different things such as which medicine is more effective and for understanding behavioural pattern of particular disease.
4.	Social Impact / Customer Satisfaction	effective use of resource Enhanced diagnosis Improved Treatment enhancing the overall quality of treatment and life of patients
5.	Business Model (Revenue Model)	With the gathered data, redirecting the patients to particular hospital based on the vacancy, leading retailers used methods like market-basket analysis to discover insights about consumer purchase behaviour and used these insights to optimize the physical store experience, target relevant ads and streamline the supply chain, among other strategic initiatives.
6.	Scalability of the Solution	A variety of institutions must store, evaluate, and take action on the massive amounts of data being produced by the health care sector as it expands quickly. India is a vast, culturally varied nation with a sizable population that is increasingly able to access centralised healthcare services.

PROJECT DESIGN PHASE – 1
SOLUTION ARCHITECTURE

SYSTEM ARCHITECTURE:



ETL PROCESS (DATA INTEGRATION PROCESS):



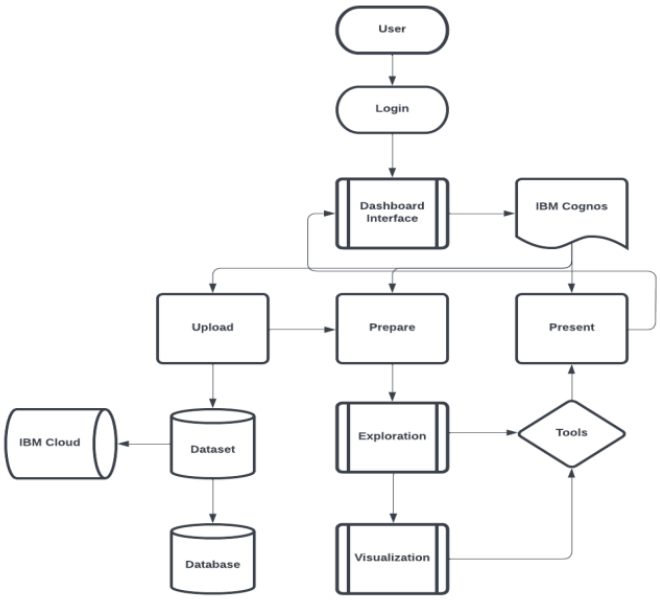
PROJECT DESIGN PHASE-II

CUSTOMER JOURNEY MAP

Journey Steps	Emergency Case	Hospitalization	Length Of Stay	Resource Allocation	Periodical Reports	Follow up consultation
Actions	Patients affected with dangerous disease will have the necessity to be admitted in hospital	Patients Should be Hospitalized after the result being positive by respective team or Management & Staffs	The length of Stay of patient should be decided by medical specialist based on the severity	The resource necessary for the treatment should be provided by the hospital Management	After complete analyze of patient's condition , periodical reports are created with particular time interval	Once recovered, Follow up consultation is necessary for being aware of the prevailing situation.
Needs and Pains	People want to Test and get the results about their physical condition	The health environment for patients and proper caring & maintenance by the respective doctors & staffs	Patient wanted to be monitored properly by the corresponding staffs and doctor	Proper resource allocation for the ongoing treatment	Patient need the report to be in a positive way	Patient need to remain healthy
Touchpoint	Test & Results	Physical mode of Admission	Analyzing the severity of the disease	Analyzing the patient's condition	Reports on Pharma Portal	Android Application or Video Conference
Customer Feeling	Nervous and Tensed	Tensed & Worried about health & recovery	Depression and Anxiety	Tensed	Positivity	Happy & Relieved
Difficulties	Difficulties in reaching the hospitals with better environment and hospitals with vacant beds	Admission process and hospitalization may take longer then expected	Depends on the case .The extreme LoS may affect the patients confidence and increase depression and anxiety	In some case, Difficulties may rise in timely allocation of resources	In some case , Report will not favor the patient and need to stay hospitalized longer	In some case, The is a chance of affecting by the same diseese again
Expectations	People wanted everything to be smooth & they expect the test reult should be negative	Patients want to get recover or cure as soon as possible with instant treatment	Patients want to return home within less period of time	Patient expects to be treated properly with the required resource	Patient expects the report to be positive	Patient expect to remain healthy with doctor's consultation

Project Design Phase-II
Data Flow Diagram & User Stories

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



User Stories

Use the below template to list all the user stories for the product.

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

Project Design Phase-II
Technology Stack (Architecture & Stack)

TECHNICAL ARCHITECTURE:

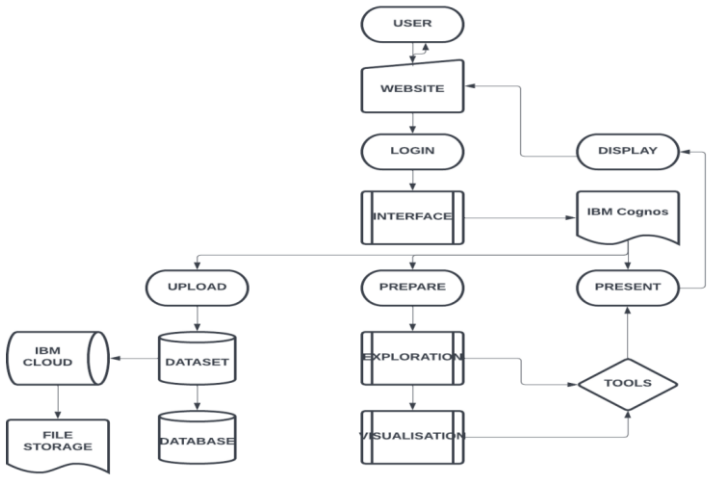


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How the user interacts with the interface e.g. Web UI, etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Dashboard Logic-1	Logic for a process in the dashboard	IBM Cognos Analytics
3.	Dashboard Logic-2	Logic for a process in the dashboard	MS Excel
4.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
5.	Cloud Database	Database Service on Cloud	IBM Cloud
6.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
7.	Uploading and Presentation	Using Exploration and Visualization	IBM Cognos Analytics

Table-2: Dashboard Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used.	IBM Cognos
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	Authentication and Authorization,Firewall,etc..
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	3-tier Architecture can be implemented so that the project can be worked by splitting up into 3 tiers namely presentation tier, application tier, data tier.
4.	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.	High availability enables your IT infrastructure to continue functioning even when some of its components fail.
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	A field of practice that uses various tools, processes, and ideas in a scientific manner to improve the desired outcomes of individuals and organizations

Project Design Phase-II
Project Milestones & Tasks

MILESTONES	TASKS
MILESTONE - 1	Data Collecting process (Datasets)
MILESTONE - 2	Required Datasets are uploaded on the IBM Cognitive Platform.
MILESTONE - 3	Data Exploration and Data Visualization
MILESTONE - 4	To Create a Interactive Dashboard.
MILESTONE - 5	Display the Insights in the Dashboard
MILESTONE - 6	Construct a Standardized Data Set and use the needed data with the Assistance of a Python Program
MILESTONE - 7	Use of different algorithm with Google Colab to achieve the desired result with more accuracy.
MILESTONE - 8	Making the output simpler and easier to understand and more efficient.
MILESTONE - 9	Deployed in the Github and waiting to review it .

Project Planning Phase
Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a health care provider I can create account in IBM cloud and the data are collected.	20	High	Nasuha, Malathi, Praveen, Subalakshmi, Raghuraman
Sprint-2	Analyze	USN-2	As a health care provider all the data that are collected is cleaned and uploaded inthe database or IBM cloud.	20	Medium	Malathi, Praveen
Sprint-3	Dashboard	USN-3	As a health care provider I can use my account in my dashboard for uploading data set.	10	Medium	Subalakshmi, Raghuraman, Nasuha
Sprint-3	Visualization	USN-4	As a health care provider I can prepare data for Visualization.	10	High	Malathi, Praveen, Raghuraman
Sprint-4	Visualization	USN-5	As a health care provider I can present data in my dashboard.	10	High	Subalakshmi, Raghuraman, Nasuha
Sprint-4	Prediction	USN-6	As a health care provider I can predict the length of stay	10	High	Malathi, Praveen

Project Tracker, Velocity & Burndown Chart: (4 Marks)

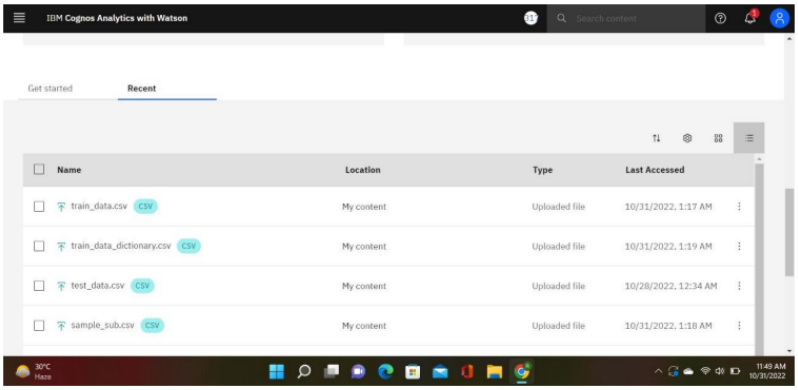
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date(Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022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

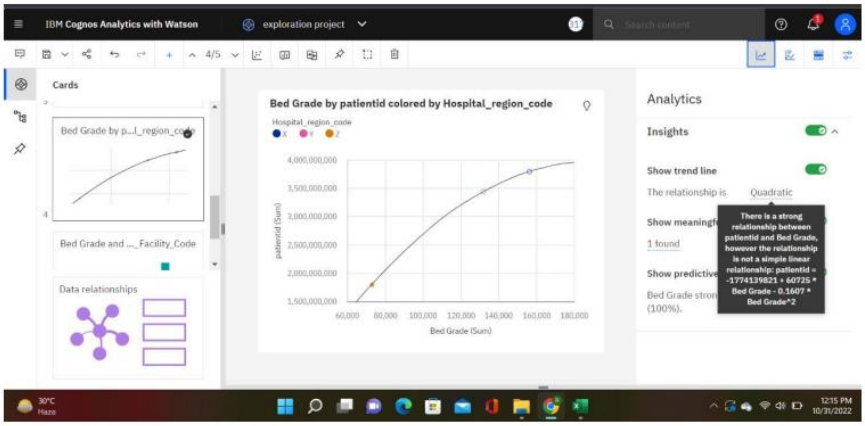
$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

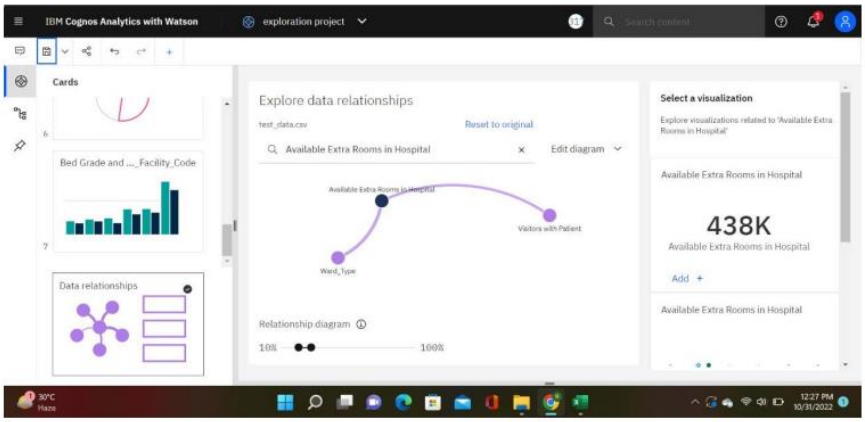
WORKING WITH THE DATA SET:

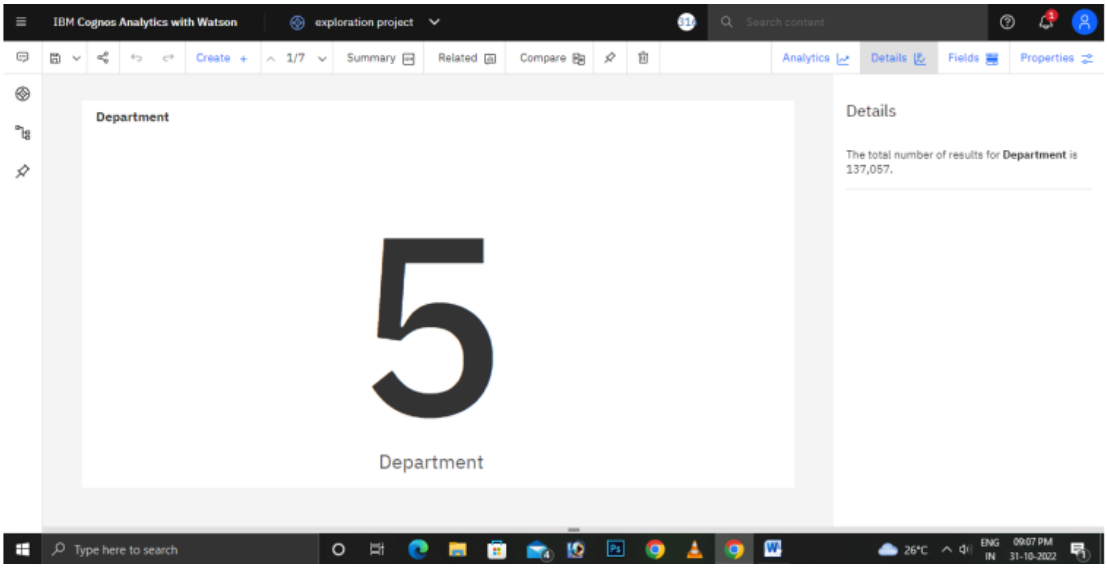
LOADING THE DATASET

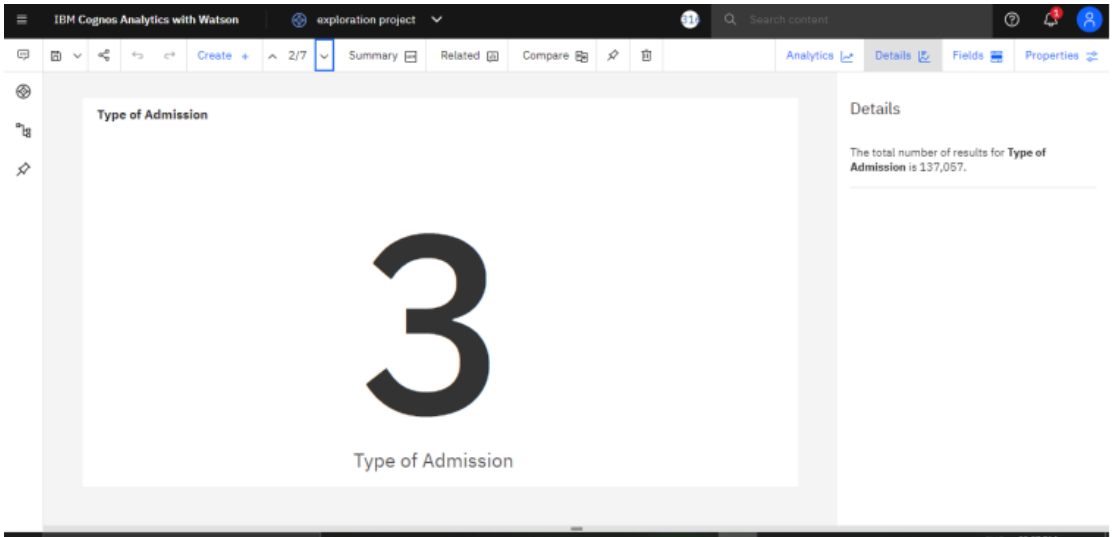


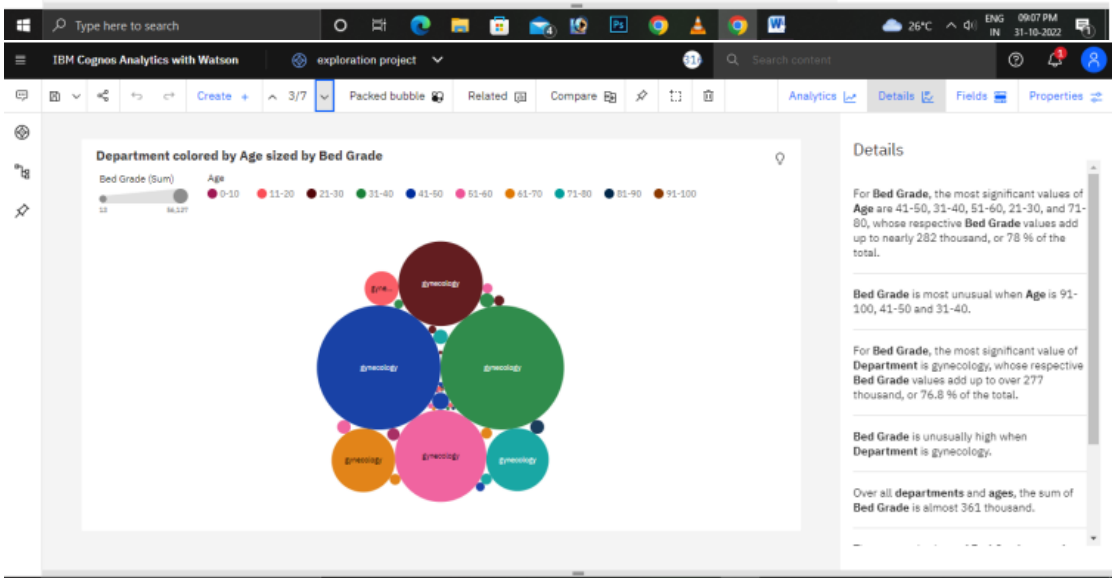
EXPLORATION OF DATA
Linkhttps://us1.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my_folders %2Fexploratio%2Bproject.

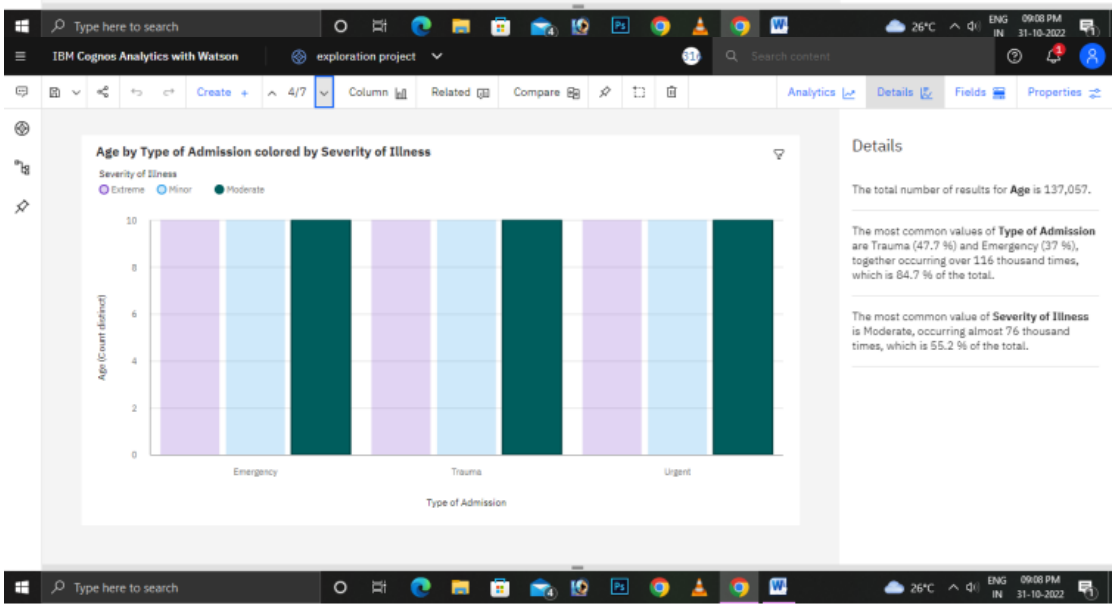

IBM Cognos Analytics with Watson
exploration project
Cards
Bed Grade by patientid colored by Hospital_region_code
Hospital_region_code
patientid (Sum)
Bed Grade (Sum)
Analytics
Insights
Show trend line
The relationship is Quadratic
There is a strong relationship between patientid and Bed Grade, however the relationship is not a simple linear relationship: patientid = -1774139821 + 60725 * Bed Grade - 0.1607 * Bed Grade^2
Show meaningful
1 found
Show predictive
Bed Grade stron
(100%).
Bed Grade and ..._Facility_Code
Data relationships

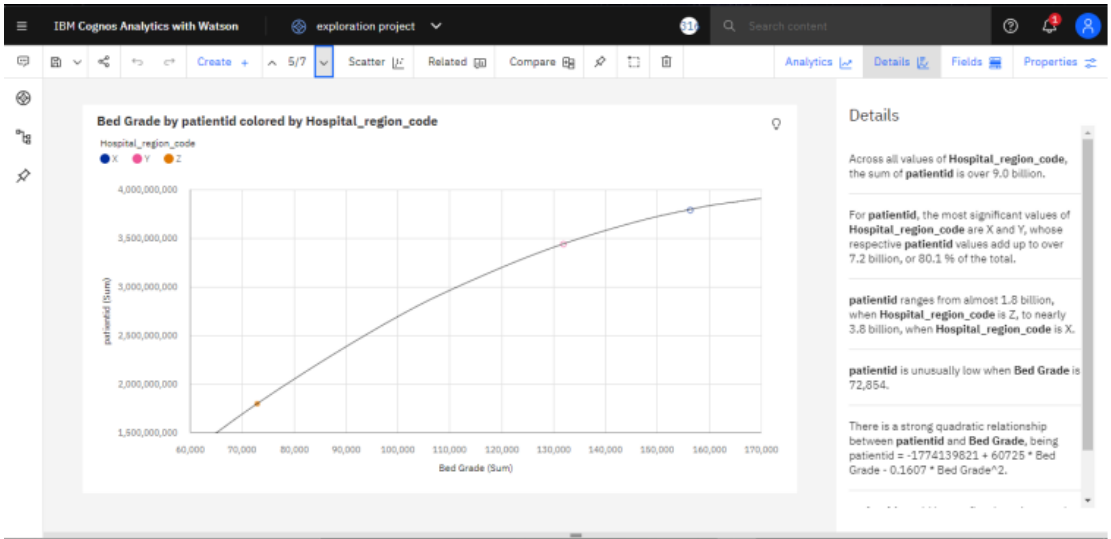

IBM Cognos Analytics with Watson
exploration project
Cards
Explore data relationships
test_data.csv
Reset to original
Available Extra Rooms in Hospital
Edit diagram
Visitors with Patient
Ward_Type
Relationship diagram
10%
100%
Select a visualization
Explore visualizations related to 'Available Extra Rooms in Hospital'
Available Extra Rooms in Hospital
438K
Available Extra Rooms in Hospital
Add
Bed Grade and ..._Facility_Code
Data relationships

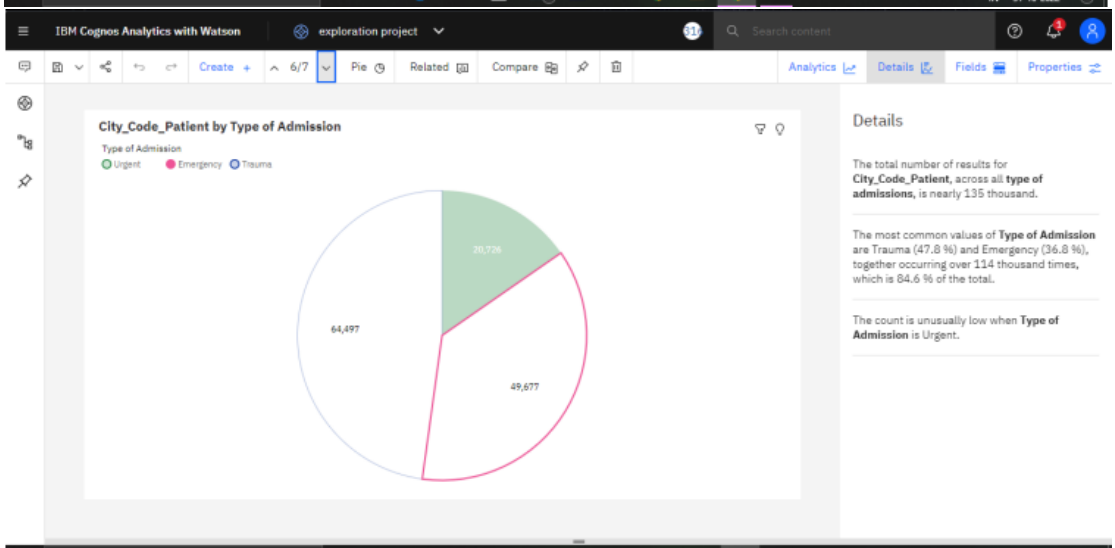

IBM Cognos Analytics with Watson
exploration project
Create
1/7
Summary
Related
Compare
Analytics
Details
Fields
Properties
Department
5
Department
Details
The total number of results for Department is 137,057.
Type here to search

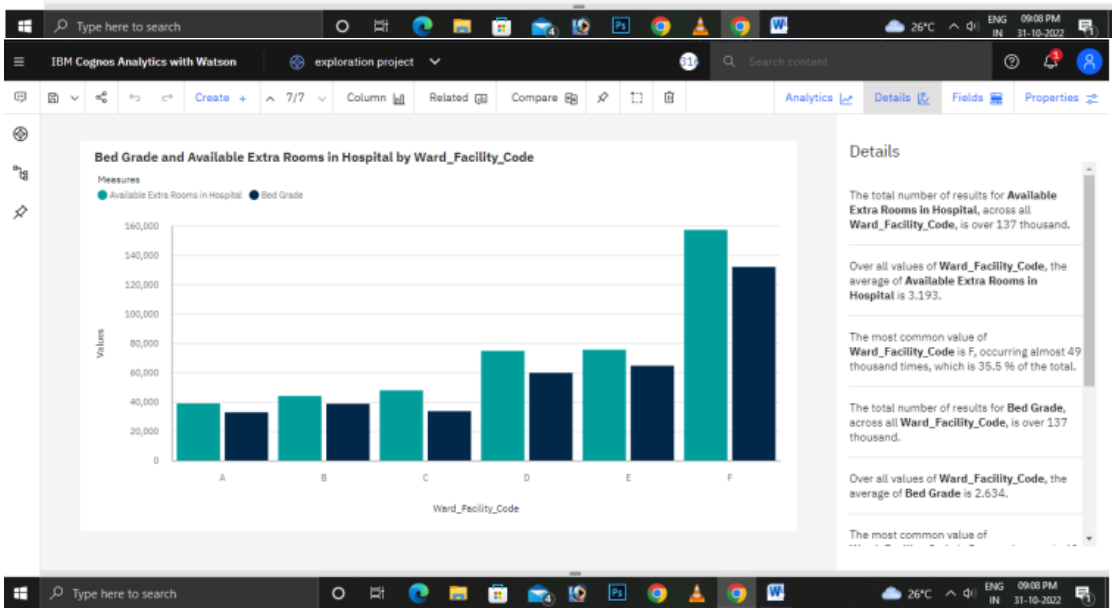

IBM Cognos Analytics with Watson
exploration project
Search content
Create
2/7
Summary
Related
Compare
Analytics
Details
Fields
Properties
Type of Admission
3
Type of Admission
Details
The total number of results for Type of Admission is 137,057.


Type here to search
26°C
ENG
IN
09:07 PM
31-10-2022
IBM Cognos Analytics with Watson
exploration project
Search content
Create
3/7
Packed bubble
Related
Compare
Analytics
Details
Fields
Properties
Department colored by Age sized by Bed Grade
Bed Grade (Sum)
Age
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
gynecology
Details
For Bed Grade, the most significant values of Age are 41-50, 31-40, 51-60, 21-30, and 71-80, whose respective Bed Grade values add up to nearly 282 thousand, or 78 % of the total.
Bed Grade is most unusual when Age is 91-100, 41-50 and 31-40.
For Bed Grade, the most significant value of Department is gynecology, whose respective Bed Grade values add up to over 277 thousand, or 76.8 % of the total.
Bed Grade is unusually high when Department is gynecology.
Over all departments and ages, the sum of Bed Grade is almost 361 thousand.


Type here to search
26°C
ENG
IN
09:08 PM
31-10-2022
IBM Cognos Analytics with Watson
exploration project
Search content
Create
4/7
Column
Related
Compare
Analytics
Details
Fields
Properties
Age by Type of Admission colored by Severity of Illness
Severity of Illness
Extreme
Minor
Moderate
Age (Count distinct)
10
8
6
4
2
0
Emergency
Trauma
Urgent
Type of Admission
Details
The total number of results for Age is 137,057.
The most common values of Type of Admission are Trauma (47.7 %) and Emergency (37 %), together occurring over 116 thousand times, which is 84.7 % of the total.
The most common value of Severity of Illness is Moderate, occurring almost 76 thousand times, which is 55.2 % of the total.





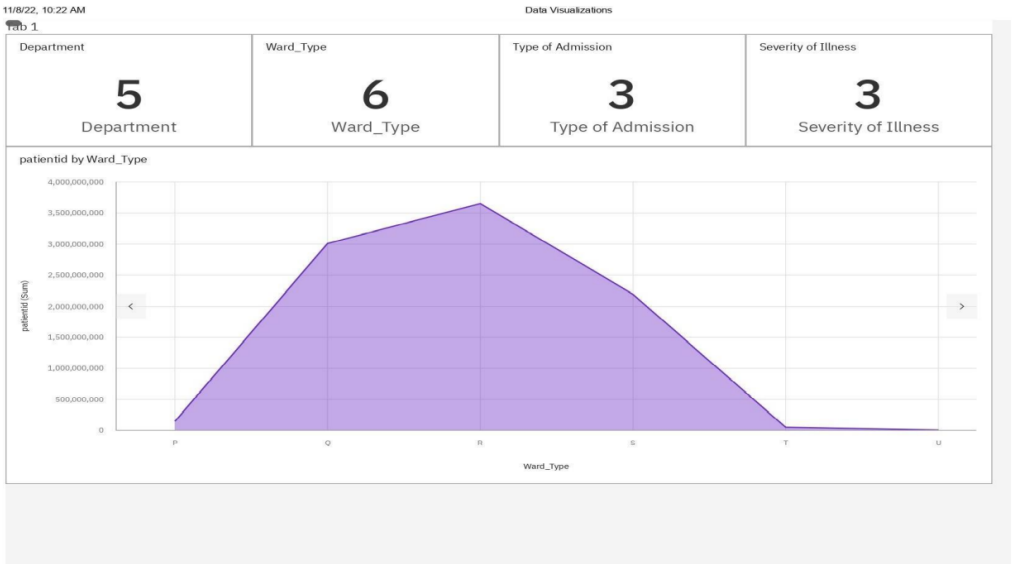


DATA VISUALIZATIONS:

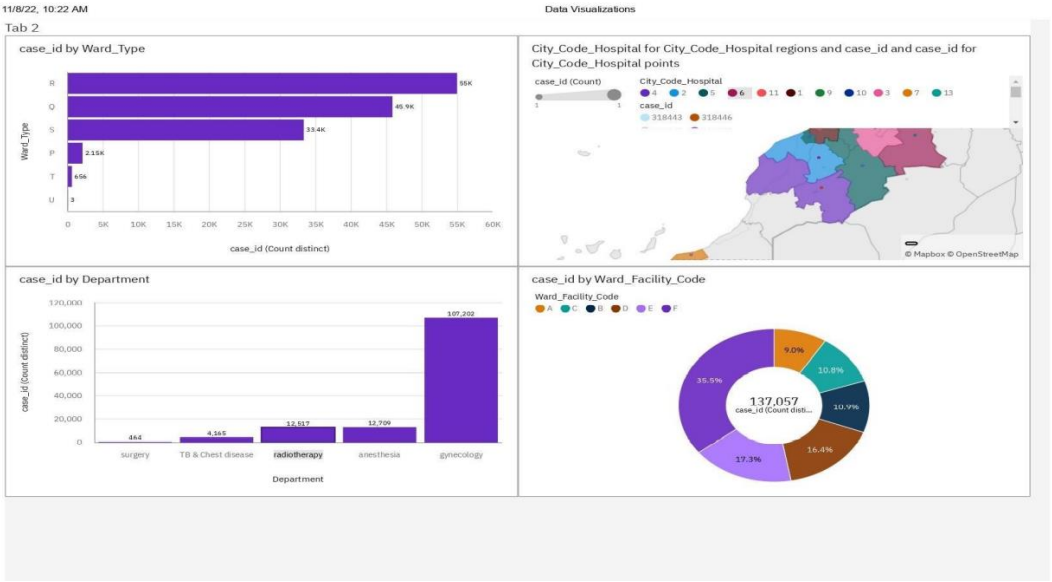
TOPICS

1) Number Of Patients By Ward Types
2) Dashboard To Show Number Of Patients
3) Age Wise Patients With Department And Severity Filters
4) Dashboard With Hierarchy Bubble And Radial Visuals
5) Dashboard Showing Pie, Stacked Bar, Waterfall And Pie Charts

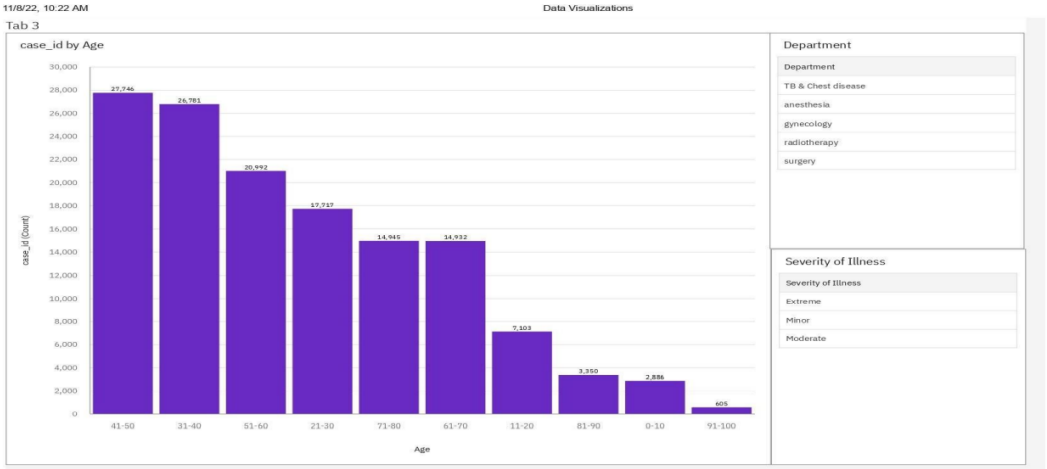
1. Number Of Patients By Ward Types



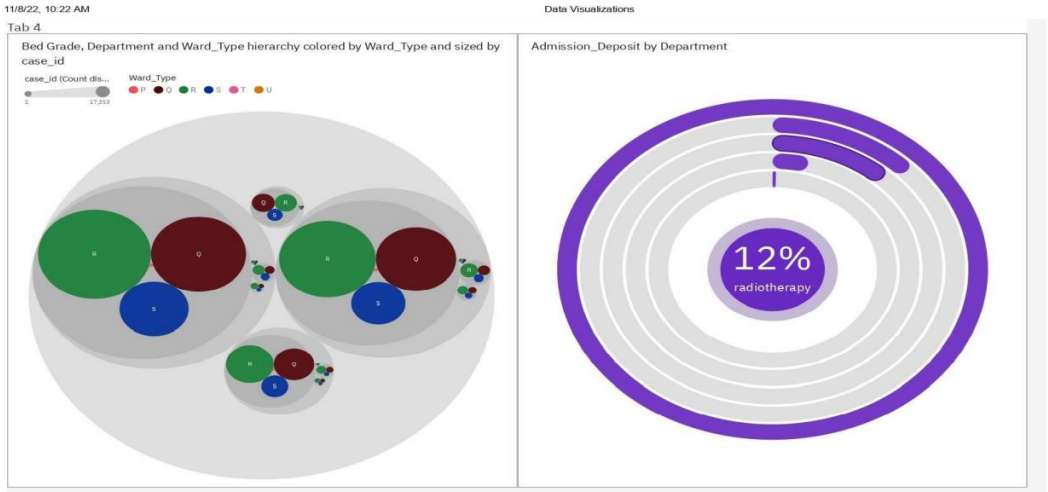
2. Dashboard To Show Number Of Patients:



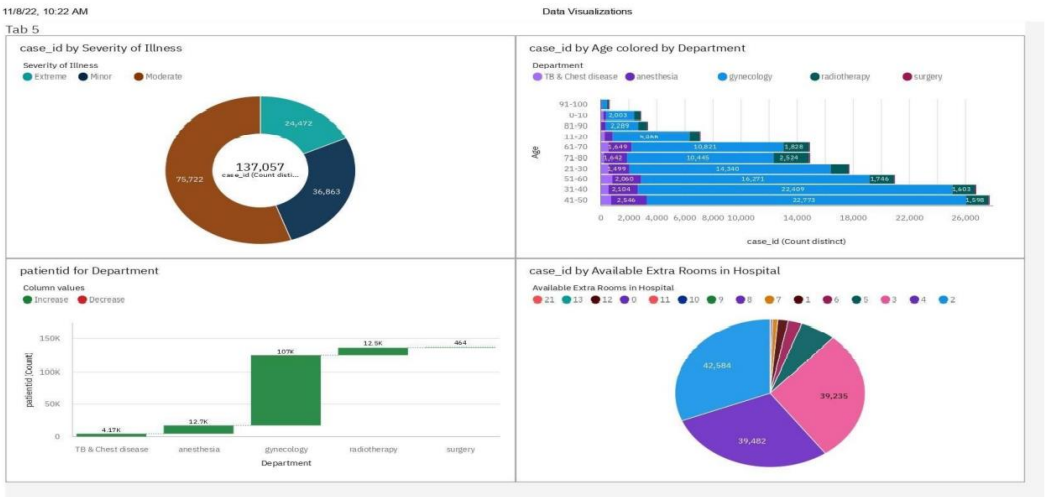
3. Age Wise Patients With Department And Severity Filters



4. Dashboard With Hierarchy Bubble And Radial Visuals



5. Dashboard Showing Pie, Stacked Bar, Waterfall And Pie Charts



Data visualization Link:
https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FData%2BVisualizations&action=view&mode=dashboard&subView=model0000018455805cb7_00000002