

Project Report

1. INTRODUCTION

1.1 Project Overview

Easy to process the data manually. Data analysis and visualization programs allow for reaching even deeper understanding. The programming language Python, with its English commands and easy-to-follow syntax, offers an amazingly powerful (and free!) open-source alternative to traditional techniques and applications. Data analytics allow businesses to understand their efficiency and performance, and ultimately helps the business make more informed decisions. For example, an e-commerce company might be interested in analyzing customer attributes in order to display targeted ads for improving sales. Data analysis can be applied to almost any aspect of a business if one understands the tools available to process information. The ecommerce companies are analyzing the reviews of customer by using proper visualization method. Exploratory Data Analysis (EDA) is an approach to summarize the data by taking their main characteristics and visualize it with proper representations. EDA focuses more narrowly on checking assumptions required for model fitting and hypothesis testing, and handling missing values and making transformations of variables as needed. EDA encompasses IDA.

1.2 Purpose

It is the first stage of data analysis. Here we can know about the content of the data set and characteristic of data set. It tells about the size of the data. We can find the missing value of data. We can find the possible relationship among data. Data visualization is done by the use of tabular data and understanding the characteristics. Fundamentally, graphical exploratory data analysis is nothing but the graphical counterpart of the traditional non-graphical EDA that analyzes the data sets to help summarize their statistical characteristics focusing on the same four key aspects, like, measures of central tendency, measures of spread, the shape of the distribution and the existence of

outliers.

2. LITERATURE SURVEY

2.1 Existing problem

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.

Suppose you have been hired as Data Scientist of Health Man – a not for profit organization dedicated to manage the functioning of Hospitals in a professional and optimal manner.

2.2 References

ABSTRACT.

In the information era, enormous amounts of data have become available on hand to decision makers. Big data refers to datasets that are not only big, but also high in variety and velocity, which makes them difficult to handle using traditional tools and techniques. Due to the rapid growth of such data, solutions need to be studied and provided in order to handle and extract value and knowledge from these datasets. Furthermore, decision makers need to be able to gain valuable insights from such varied and rapidly changing data, ranging from daily transactions to customer interactions and social network data. Such value can be provided using big data analytics, which is the application of advanced analytics techniques on big data. This paper aims to analyze some of the different analytics methods and tools which can be applied to big data, as well as the opportunities provided by the application of big data analytics in various decision domains.

Introduction

Imagine a world without data storage; a place where every detail about a person or organization, every transaction performed, or every aspect which can be documented is lost directly after use. Organizations would thus lose the ability to extract valuable information and knowledge, perform detailed analyses, as well as provide new opportunities and advantages. Anything ranging from customer names and addresses, to products available, to purchases made, to employees hired, etc. has become essential for day-to-day continuity. Data is the building block upon which any organization thrives. Now think of the extent of details and the surge of data and information provided nowadays through the advancements in technologies and the internet. With the increase in storage capabilities and methods of data collection, huge amounts of data have become easily available. Every second, more and more data is being created and needs to be stored and analyzed in order to extract value. Furthermore, data has become cheaper to store, so organizations need to get as much value as possible from the huge amounts of stored data. The size, variety, and rapid change of such data require a new type of big data analytics, as well as different storage and analysis methods. Such sheer amounts of big data need to be properly analyzed, and pertaining information should be extracted.

Model Building

We use the statistical model or machine learning model to describe the variable and working of the variable. Model can be supervised or unsupervised model. We can use classification, regression model to get the output. We can visualize the result by the use of model. After that we have to evaluate the model

Univariate Graphical EDA

Univariate GEDA provides statistical summary for each field in the raw data set or the summary only on one variable.

Bivariate Graphical EDA

Bivariate GEDA is accomplished to understand the connections between each variable in the dataset and the target variable of interest or using two variables and finding connection among them. Example of these types of GEDA includes Box plot and Violin plot

Multivariate Graphical EDA

Multivariate GEDA is accomplished to understand the connections between different fields in the dataset or finding the connections between more than two variables.

2.3 Problem Statement Definition

Descriptive Statistical Analysis:

We have used descriptive statistical analysis which is used to describe the entire data sets with a single value or metric. The describe function automatically computes basic statistics for all continuous variables.

WORKING WITH THE DATA SETS

are using belongs to Amazon review data set. We are going to analyse the data with possible set of options.

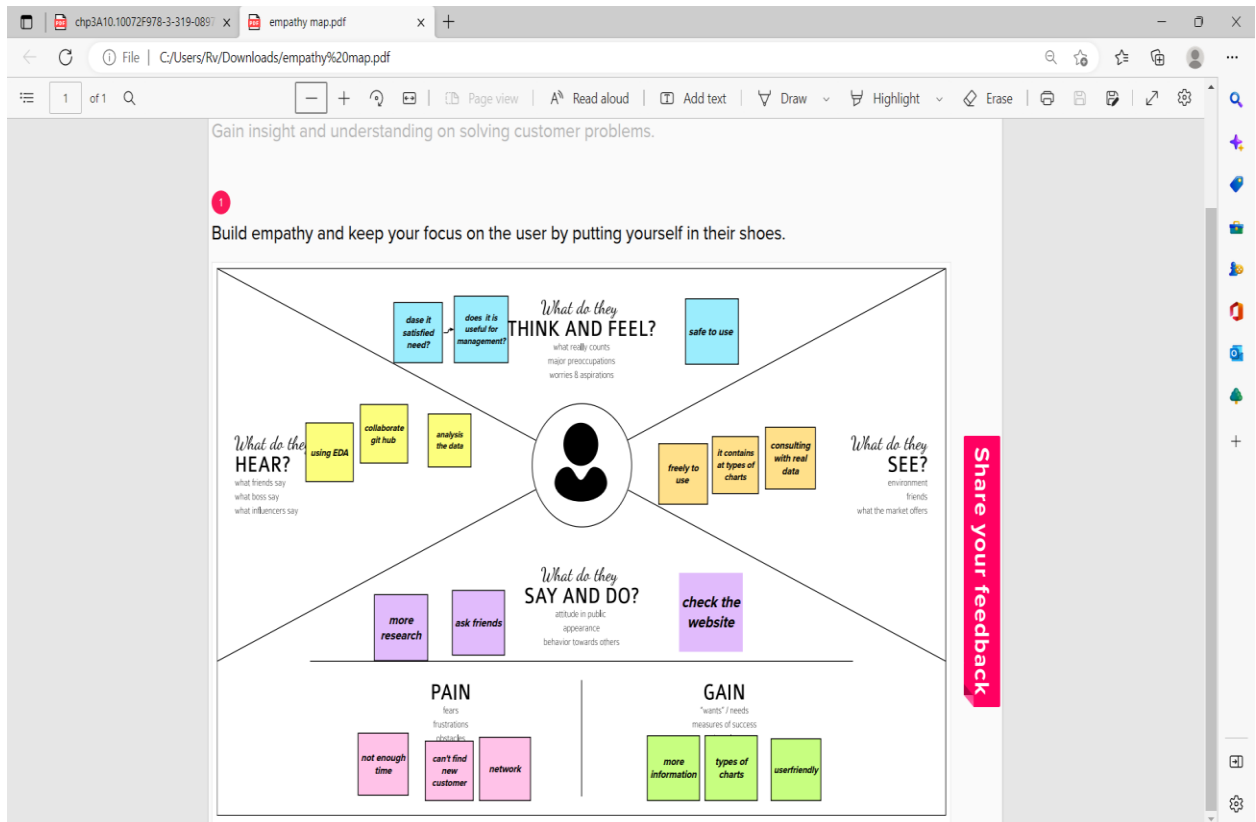
1. In the first step we have imported the Pandas libraries. numpy packages.
2. After that we have imported fairly large amazon CSV file as a data frame df. It gives the data sets in the form of rows and column.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

An empathy map is **a collaborative tool teams can use to gain a deeper insight into their customers**. Much like a user persona, an empathy map can represent a group of users, such as a customer segment. The empathy map was originally created by Dave Gray and has gained much popularity within the agile community.

An Empathy Map consists of four quadrants. The four quadrants reflect four key traits, which the user demonstrated/possessed during the observation/research stage. The four quadrants refer to what the user: Said, Did, Thought, and Felt.



3.2 Ideation & Brainstorming

Brainstorming is **an activity that will help you generate more innovative ideas**. It's one of many methods of ideation—the process of coming up with new ideas—and it's core to the design thinking process. Ideation is often closely related to the practice of brainstorming, a specific technique that is utilized to generate new ideas. A principal difference between ideation and brainstorming is that **ideation is commonly more thought of as being an individual pursuit, while brainstorming is almost always a group activity**



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

[Share template feedback](#)

4

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

Team gathering

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

Set the goal

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

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

The user needs a ways to d analysis hospital patient details. they can easily access the details quickly. analysis previous data is efficiently,so that user benefit direct.

Key rules of brainstorming

To run an smooth and productive session

- Stay in topic
- Encourage wild ideas
- Defer judgment
- Listen to others
- Go for volume
- If possible, be visual

2

Brainstorm

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

10 minutes

TIP
You can submit sticky notes and feedback points (points to submit) even in silent thinking!

ravichandran

anaconda platform

developer services

vasanth

analysis the patient details

use easy

losses of time

vishnu

use anywhere

usepython lib

use ibm cloud

think different idea

manikandan

development

efficient resources

open platform

chandru

easy used

reduce the field work

use only online



Need some inspiration?

See a featured session or other sessions to which you wish to add.

[Open calendar](#)



Project Design Phase-I
Proposed Solution Template

Date	24 September 2022
Team ID	PNT2022TMID44635
Project Name	ANALYTICS FOR HOSPITALS HEALTH CARE DATA
Maximum Marks	2 Marks

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)	The covid-19 pandemic has resulted in <u>uncontrollable</u> <u>may</u> local hospital were not prepared to handle this crisis. These challenges may include visualization and wide array of data integrity concerns.
2.	Idea / Solution description	Data visualization acquires the main in the health industry into focus help to identify patterns as well as correlation and makes data analysis more relevant. Example data visualization include interactive dashboard, bar charts, pie charts, heat maps, all of which have their particular uses to represent idea and data.
3.	Novelty / Uniqueness	These days data often <u>needs</u> to be visually presented in the form of interactive graphs or charts to be impactful and understand.
4.	Social Impact / Customer Satisfaction	Reducing patient waiting times Allocating and scheduling adequate staffing levels Reducing readmission rates Monitoring clinical trial data and patient

Accessibility: Good to go

		prepared to handle this crisis. These challenges may include visualization and wide array of data integrity concerns.
2.	Idea / Solution description	Data visualization acquires the main in the health industry into focus help to identify patterns as well as correlation and makes data analysis more relevant. Example data visualization include interactive dashboard, bar charts, pie charts, heat maps, all of which have their particular uses to represent idea and data.
3.	Novelty / Uniqueness	These days data often <u>needs</u> to be visually presented in the form of interactive graphs or charts to be impactful and understand.
4.	Social Impact / Customer Satisfaction	Reducing patient waiting times Allocating and scheduling adequate staffing levels Reducing readmission rates Monitoring clinical trial data and patient outcomes Less casualty
5.	Business Model (Revenue Model)	Analytics is playing a huge role in helping companies taking informed decision therapeutic areas, markets and regions t reach up to decision within the stipulated timeframe and get exposed to real world insight from competitors, payers, regulators, <u>patients</u> , etc.
6.	Scalability of the Solution	where they will be able to spend less time on grunt work and more time on actually resolving issues.

Accessibility: Good to go

3.4 Problem Solution fit

The Problem-Solution Fit simply means that **you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem.** Problem-Solution Fit - **this occurs when you have evidence that customers care about certain jobs, pains, and gains.** At this stage you've proved the existence of a problem and have designed a value proposition that addresses your customers' jobs, pains and gains

Problem-Solution fit canvas 2.0		Purpose / Vision	
Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS Who is your customer? i.e. working parents of 0-5 y.o. kids hospital management	6. CUSTOMER CONSTRAINTS CC 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. these challenges may include visualization wide array of data integrity concerns	5. AVAILABLE SOLUTIONS AS Which solutions are available to the customers when they face the problem 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 alternative to digital notetaking data visualization include interactive dashboard, bar charts, pie charts, all of which have their particular uses to represent idea and data
	Focus on J&P, tap into BE, understand RC	2. JOBS-TO-BE-DONE / PROBLEMS J&P Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides. where they will be able to spend less time on grunt work and more time on actually resolving issues.	9. PROBLEM ROOT CAUSE RC What is the real reason that this problem exists? What is the back story behind the need to do this job? i.e. customers have to do it because of the change in regulations. the covid 19 pandemic has resulted in uncontrolled may local hospital were not prepared to handle this crisis.
Identify strong TR & EM		3. TRIGGERS TR What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news. 4. EMOTIONS: BEFORE / AFTER EM 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. reducing patient waiting times reducing readmission rates	10. YOUR SOLUTION SL 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 behaviour. these days data often needs to be visually presented in the form of interactive graphs or charts to be impactful and understand



Problem-Solution fit canvas is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 license
 Created by Daria Nepriakhina / Amaltama.com



4. REQUIREMENT ANALYSIS

Requirements analysis, also called requirements engineering, is **the process of determining user expectations for a new or modified product**. These features, called requirements, must be quantifiable, relevant and detailed. In software engineering, such requirements are often called functional

TYPES OF REQUIREMENT ANALYSIS

The main types of requirements analysis include **business, customer, product, functional, and non-functional requirements**. Each one represents a stakeholder or stage of the project and communicates the project needs

4.1 Functional requirement

Functional requirements **may involve calculations, technical details, data manipulation and processing, and other specific functionality that define what a system is supposed to accomplish**. Behavioral requirements describe all the cases where the system uses the functional requirements, these are captured in use cases.

Well-written functional requirements typically have the following characteristics: Necessary. Although functional requirements may have different priority, every one of them needs to relate to a particular business goal or user requirement.

Project Design Phase-II Solution Requirements (Functional & Non-functional)

Date	15 October 2022
Team ID	PNT2022TMID44635
Project Name	Project – analytics for hospital health care data
Maximum Marks	4 Marks

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	Register portal or dashboard
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Dashboard	The collect the data are found in visualized format and the prior data as analyzed.
FR-4	Dataset	The patients record and staff record are collected and consolidated as dataset
FR-5	Report generator	The periodic reports of patients and Los are reported
FR-6	Exploration	The data exploration on available dataset

4.2 Non-Functional requirements

Nonfunctional Requirements (NFRs) **define system attributes such as security, reliability, performance, maintainability, scalability, and usability**. They serve as constraints or restrictions on the design of the system across the different backlogs.

The importance of non-functional requirements stems from the fact that **they can make or break the success of a software system or a product**. Users will not hesitate to outrightly reject a system even if it meets all expected functional requirements but fail to deliver the required quality outcomes

- Non-functional requirements or NFRs are a set of specifications that describe the system's operation capabilities and constraints and attempt to improve its functionality. ...
- Performance and scalability. ...
- Portability and compatibility. ...
- Reliability, maintainability, availability.

Non-functional Requirements:

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

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	The effectiveness, efficiency and satisfaction with which specific user can achieve aa specific set of task in a particular environment
NFR-2	Security	Educate heath care staff The data are confidential so hospital staff conscious about it. restrict access to data and application
NFR-3	Reliability	The analytics systems ensures patients Los can be identified. Data analytics in clinical settings attempts to reduced patients wait via improved scheduling and staffing ,give patients more option when scheduling appointment and receiving treatment and reduced readmission rates by using population health data to predict which patient are at greatest risk

NFR-4	Performance	Promote preventive in measures by giving patients greater insight into their health and treatment goals. The accurate results of patient Los can be identified
NFR-5	Availability	The availability od data set must be constrained for accurate data. Each year there is more demand for data analytics and scientist than there are people with the right skills to fills those roles.
NFR-6	Scalability	Any kind data can be explored and the system is quite expendable. scalability is the ability of a health intervention shown to be efficacious on a small scale and/or under controlled condition to be expanded under real world conditions to reach a greater proportion of the eligible population, while retaining effectiveness

5. PROJECT DESIGN

5.1 Data Flow Diagrams

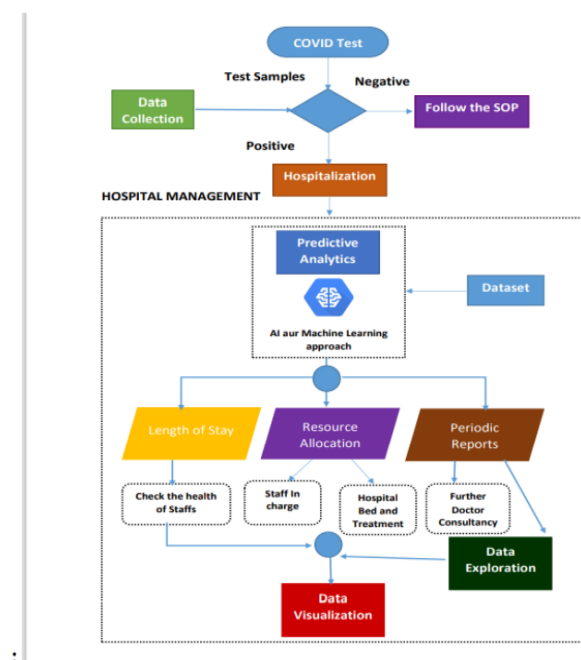
It illustrates what kinds of data will be input and output from the system, where the data will come from and go to, and where the data will be stored. A DFD is often an elaboration of a context diagram to show some of the detail of the system that was first illustrated through the context diagram.

Also known as DFD, Data flow diagrams are used to graphically represent the flow of data in a business information system. DFD describes the processes that are involved in a system to transfer data from the input to the file storage and reports generation. Data flow diagrams can be divided into logical and physical

Project Design Phase-II Data Flow Diagram & User Stories

Date	15 October 2022
Team ID	PNT2022TMID44635
Project Name	Project – analytics of hospital health care data

Data Flow Diagrams

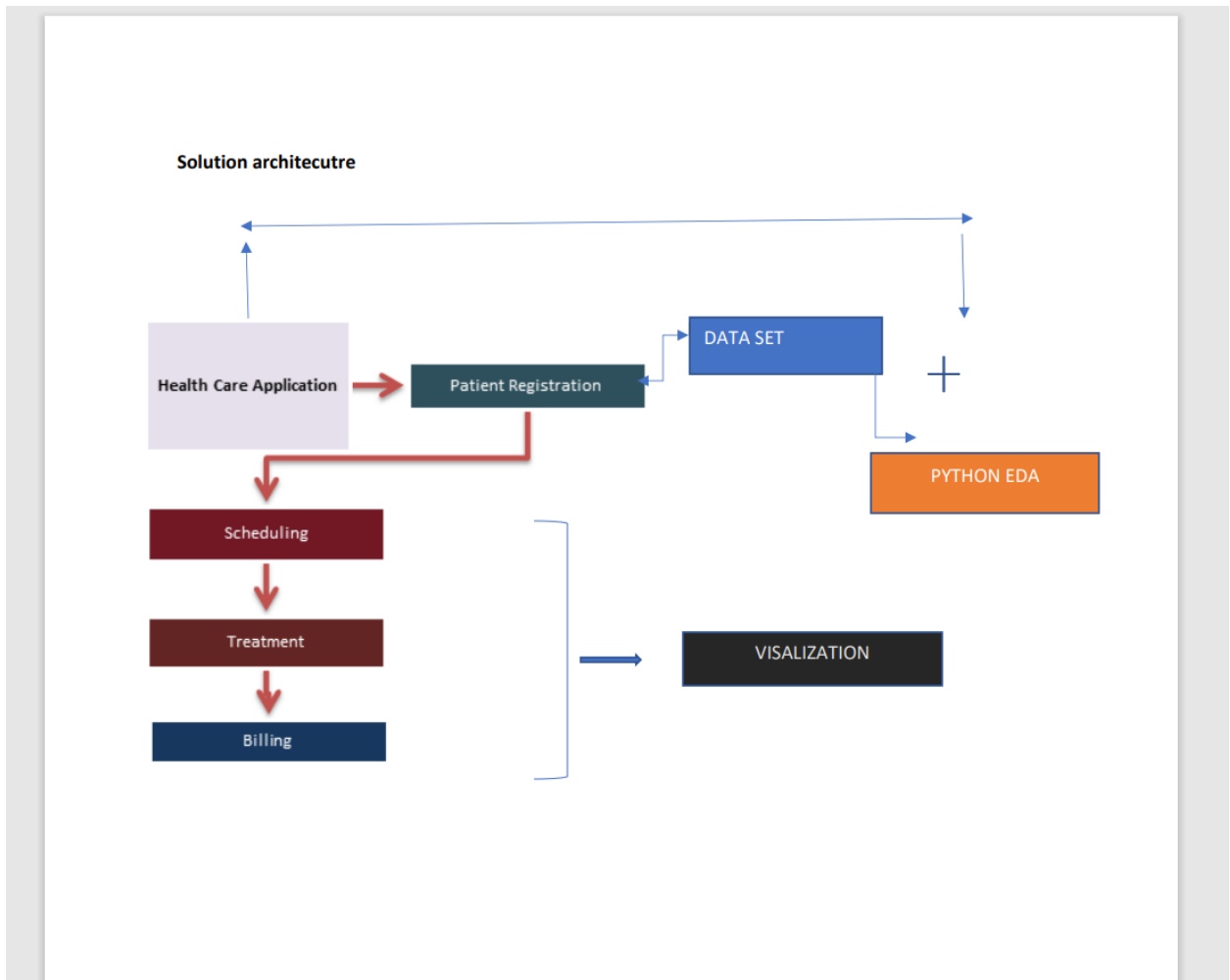


5.2 Solution & Technical Architecture

Solution Architects are most similar to project managers, ensuring that all parties, including stakeholders, are on the same page and moving in the right direction at all stages. Technical architects manage all activities leading to the successful implementation of a new application.

Technical Architecture (TA) is **a form of IT architecture that is used to design computer systems**. It involves the development of a technical blueprint with regard to the arrangement, interaction, and interdependence of all elements so that system-relevant requirements are met.

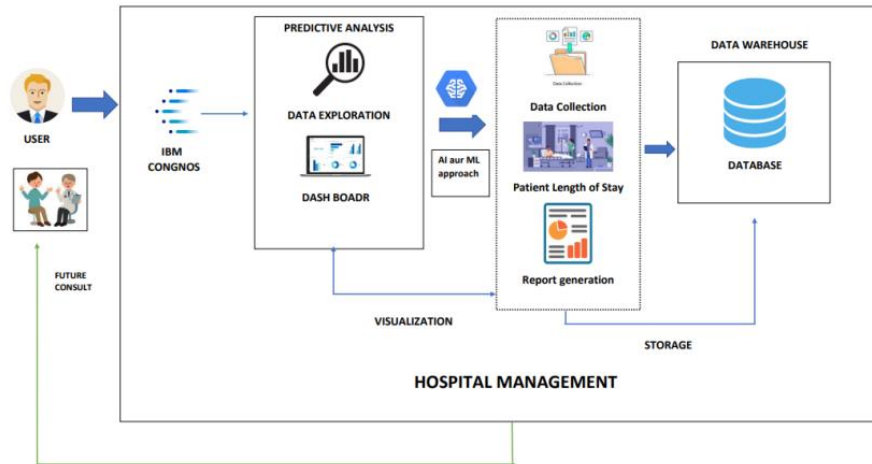
A solution architecture (SA) is **an architectural description of a specific solution**. SAs combine guidance from different enterprise architecture viewpoints (business, information and technical), as well as from the enterprise solution architecture (ESA).



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

Date	15 October 2022
Team ID	PNT2022TMID44635
Project Name	Project –Health Care Data analytics
Maximum Marks	4 Marks

Technical Architecture:



S.No	Component	Description	Technology
1.	User Interface	All the data in Ms Excel sheet	IBM Cognitive analysis
2.	Application Logic-1	Then Upload the data in IBM Watson	Ms Excel Sheet or Kaggle dataset
3.	Application Logic-2	Prepare the data	IBM Watson STT service, Kaggle data set
4.	Application Logic-3	Explore the Data	IBM Watson Assistant
5.	Database	Present data	Using numpy, pandas, seaborn lib
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	Dashboard	IBM cognos
8.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration:Tableu Cloud Server Configuration : IBM CLOUD	Eda Python, using jupyter notebook anaconda software or google colab.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Software with a source code that anyone can inspect, modify or enhance.	Visualization, charts, discriptive
2.	Security Implementations	Including administrative controls, physical security, logical controls, organizational	e.g. firewalls. authentication and authorization.

		standards	encryption. data masking
3.	Scalable Architecture	The ability of a hardware/software parallel system to exploit increasing computing resources effectively in the analysis of (very) large datasets.	a package delivery system
4.	Availability	An organization ensures that all of its business-related data is available to the organization, partners, or end-users at any time of the day, whenever and wherever required.	NoSQL Databases. Knowledge Discovery Tools. Stream Analytics. , In-memory Data Fabric. Distributed Storage.
5.	Performance	the process of quickly examining extremely large data sets to find insights	hadoop

5.3 User Stories

A user story is **an informal, general explanation of a software feature written from the perspective of the end user or customer**. The purpose of a user story is to articulate how a piece of work will deliver a particular value back to the customer.

User stories are part of an agile approach that helps shift the focus from writing about requirements to talking about them

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
patients	hospitalization	USN-1	Patients are required hospitalized if they have covid positive	Maintain the Record for correct preference. Direct hospitalization	High

	Treatment report	USN-2	Patients should collect their treatment report and get the doctor consult	They can receive the report form hospital. For Admission purpose of patient to get solutions among their problem	High
Hospital management	Resource allocation	USN 3	Hospital management should be allocate the necessary resource for treating the covid patients	Should be ready for any circumstance	High
	Predicting	USN-4	The doctors should be aware of condition of patient the predict	Exploring the data about the patient health condition and predicting	High
	Resource availability	USN-5	As a user, I access the data in visualize	Visualize the about the resource availability	High
	Dashboard	USN-6	As a user, I can access the data from queries, graph, pie chart		High
Customer Care Executive			Have data in graph modes		
Customer (Web user)			Get the older information from the hospitals		

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

PROJECT PLANNING PHASE

Date	22.10.2022
Team ID	PNT2022TMID44635
Project Name	Analytics for Hospital Healthcare Data

Product Backlog, Sprint Schedule, and Estimation:

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	Dataset	USN-1	The user needs a complete data about the patients admitted in the hospital and a dataset should be prepared.	2	High	MANIKANDAN R
Sprint-1	Data Exploration	USN-2	Data exploration is the first step of data analysis used to explore and visualize data by charts and graphs.	2	High	CHANDRU K
Sprint-2	Data Visualization	USN-3	The patient data are graphically visualized for data verification and to know available resources	2	High	VISHNU S
Sprint-3	Dashboard	USN-4	The explored and visualized data are displayed in dashboard	2	High	VASANTH D
Sprint-4	Predictive Model	USN-5	The Predictive analysis on the data is performed by predictive model.	2	High	RAVICHANDRAN A

6.2 Sprint Delivery Schedule

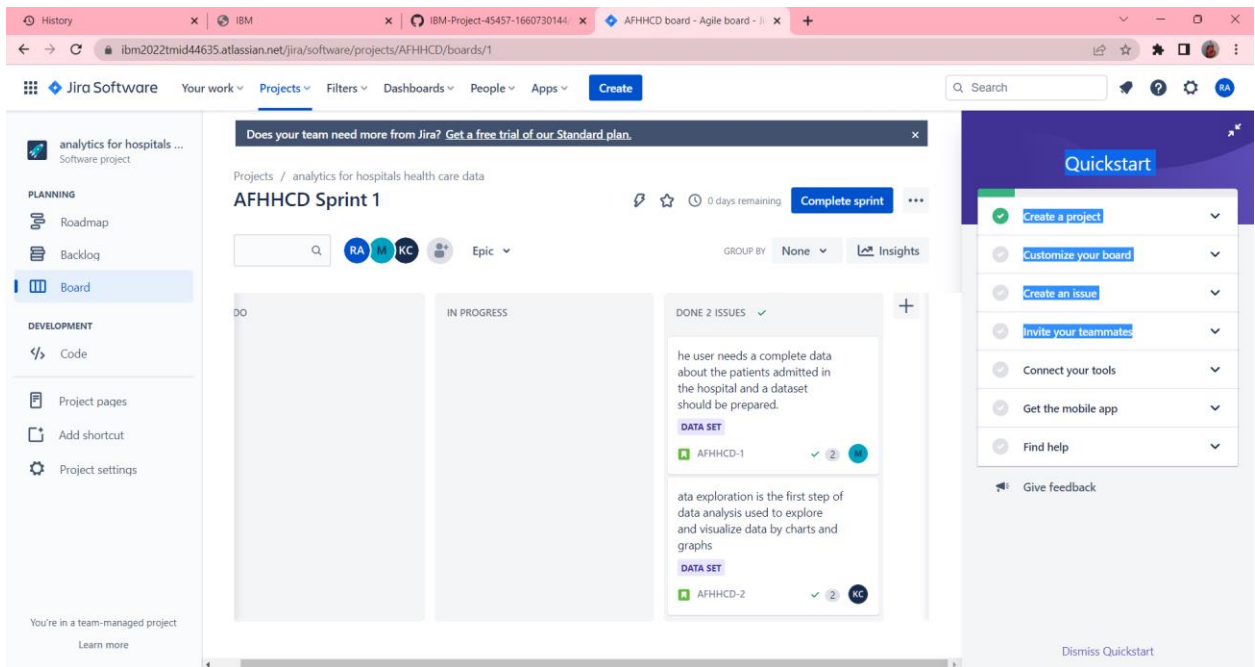
PROJECT PLANNING PHASE

Date	22.10.2022
Team ID	PNT2022TMID44635
Project Name	Analytics for Hospitals Healthcare Data

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	8 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

6.3 Reports from JIRA

	OCT
	24 25 26 27 28 29
Sprints	AFHHCD Sprint 1
> ⚡ AFHHCD-6 data set	



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

7.1 Feature 1 exploration

IBM-EPBL/IBM x IBM-Project x Colab Note x IBM x Sprint 1 explor x My IBM x Home x Your final cho x project user s x

colab.research.google.com/drive/1wd3c9xQDdCkcc8dbXO0AIRush8DwB-Te

Sprint 1 exploration.ipynb

File Edit View Insert Runtime Tools Help Last saved at November 18

+ Code + Text

Connect Editing

Sprint 1

```
[ ] from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

**import library

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings("ignore")

**load the dataset

[ ] test=pd.read_csv("/content/drive/My Drive/test_data.csv")
```

brainstrom ideatio...pdf brainstrom ideatio...pdf brainstrom ideatio...pdf Testcases Report(P...xlsx Sprint_1_explora...ipynb Removed Show all

IBM-EPBL/IBM x IBM-Project x Colab Note x IBM x Sprint 1 explor x My IBM x Home x Your final cho x project user s x

colab.research.google.com/drive/1wd3c9xQDdCkcc8dbXO0AIRush8DwB-Te#scrollTo=_Sty2UwpLf

Sprint 1 exploration.ipynb

File Edit View Insert Runtime Tools Help Last saved at November 18

+ Code + Text

Connect Editing

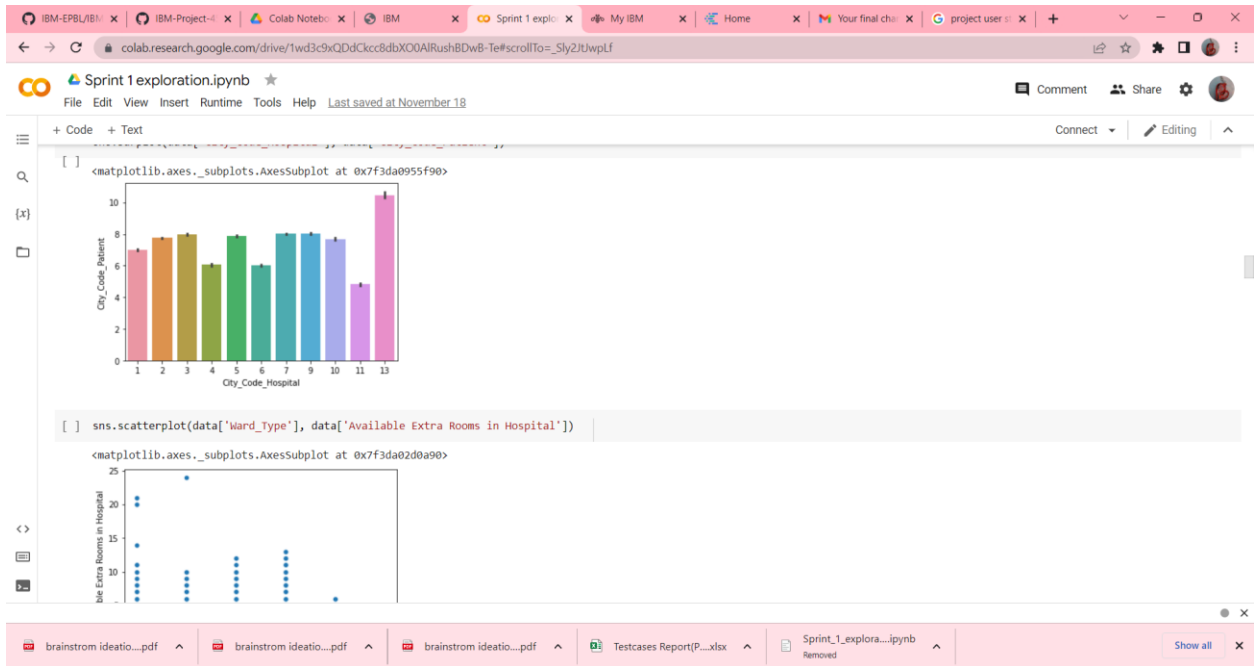
```
[ ] data.describe()
```

	case_id	Hospital_code	City_Code_Hospital	Available Extra Rooms in Hospital	Bed Grade	patientid	City_Code_Patient	Visitors with Patient	Admission_Deposit
count	318438.000000	318438.000000	318438.000000	318438.000000	318325.000000	318438.000000	313906.000000	318438.000000	318438.000000
mean	159219.500000	18.318841	4.771717	3.197627	2.625807	65747.579472	7.251859	3.284099	4880.749392
std	91925.276848	8.633755	3.102535	1.168171	0.873146	37979.936440	4.745266	1.764061	1086.776254
min	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	0.000000	1800.000000
25%	79610.250000	11.000000	2.000000	2.000000	2.000000	32847.000000	4.000000	2.000000	4186.000000
50%	159219.500000	19.000000	5.000000	3.000000	3.000000	65724.500000	8.000000	3.000000	4741.000000
75%	238828.750000	26.000000	7.000000	4.000000	3.000000	98470.000000	8.000000	4.000000	5409.000000
max	318438.000000	32.000000	13.000000	24.000000	4.000000	131624.000000	38.000000	32.000000	11008.000000

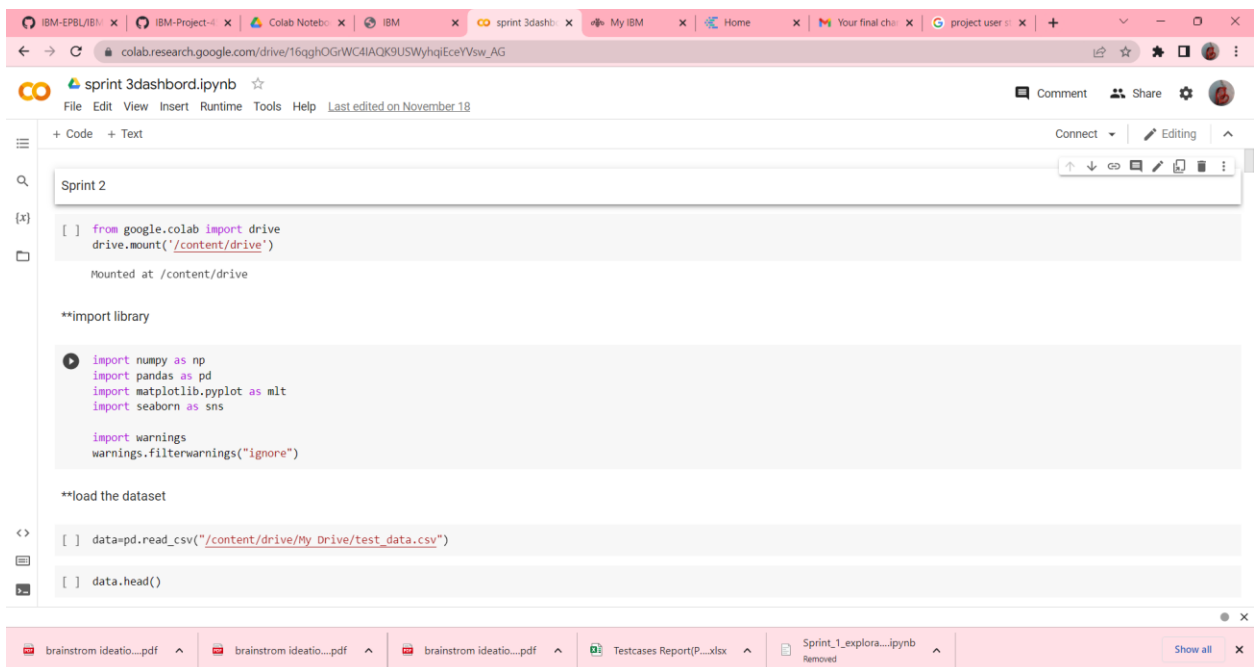
```
[ ] data.nunique()
```

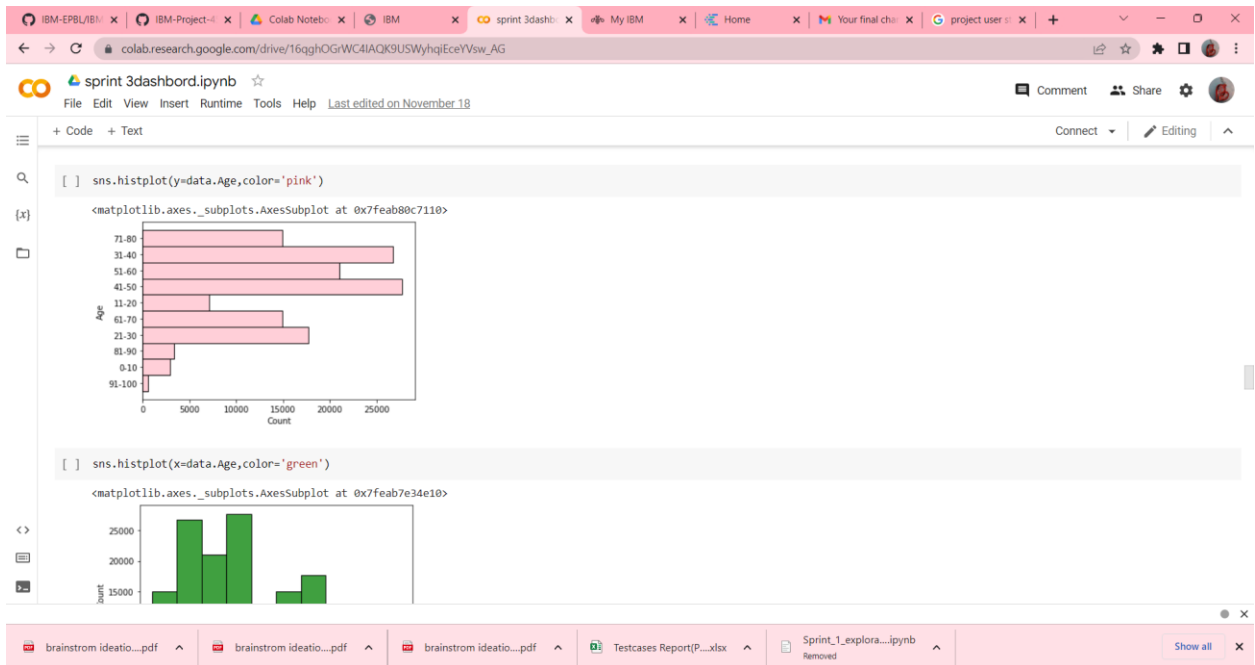
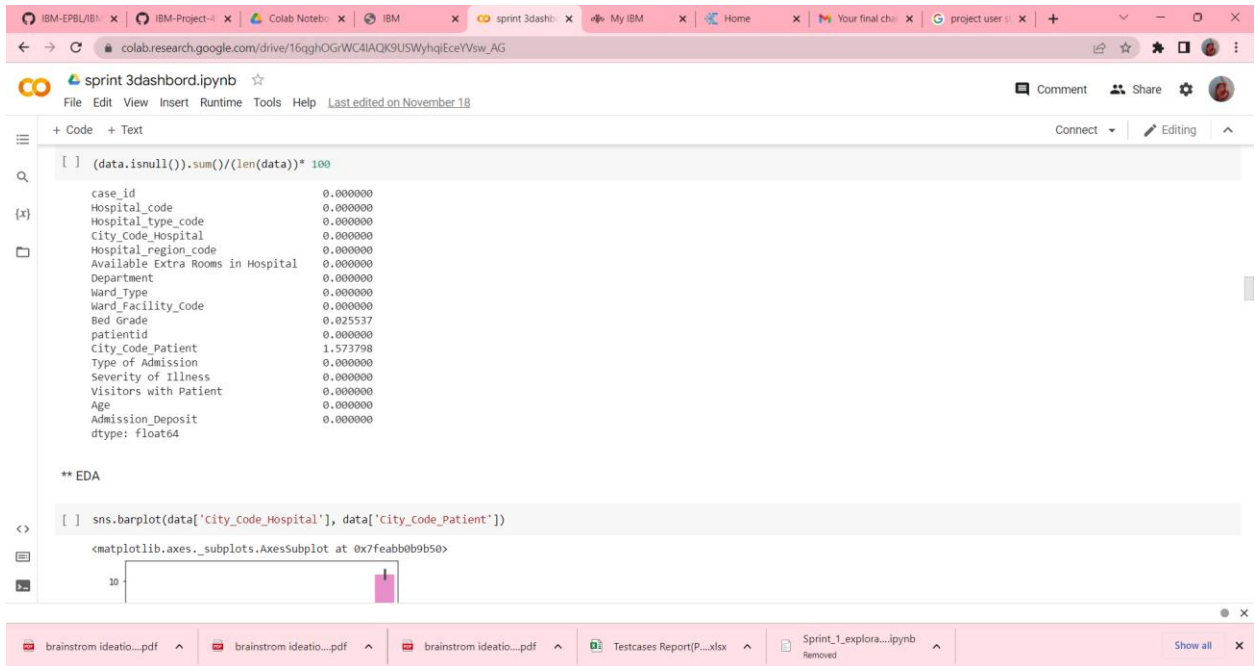
	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code
	318438	32	7	11	3	18	5	6	6

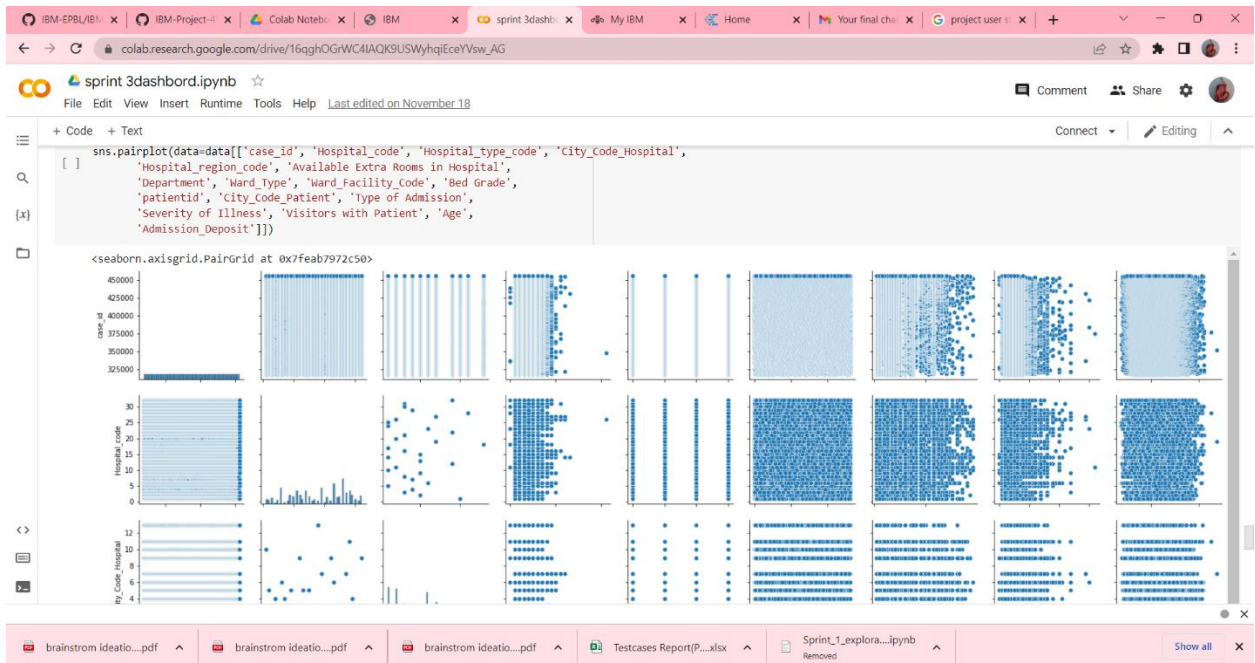
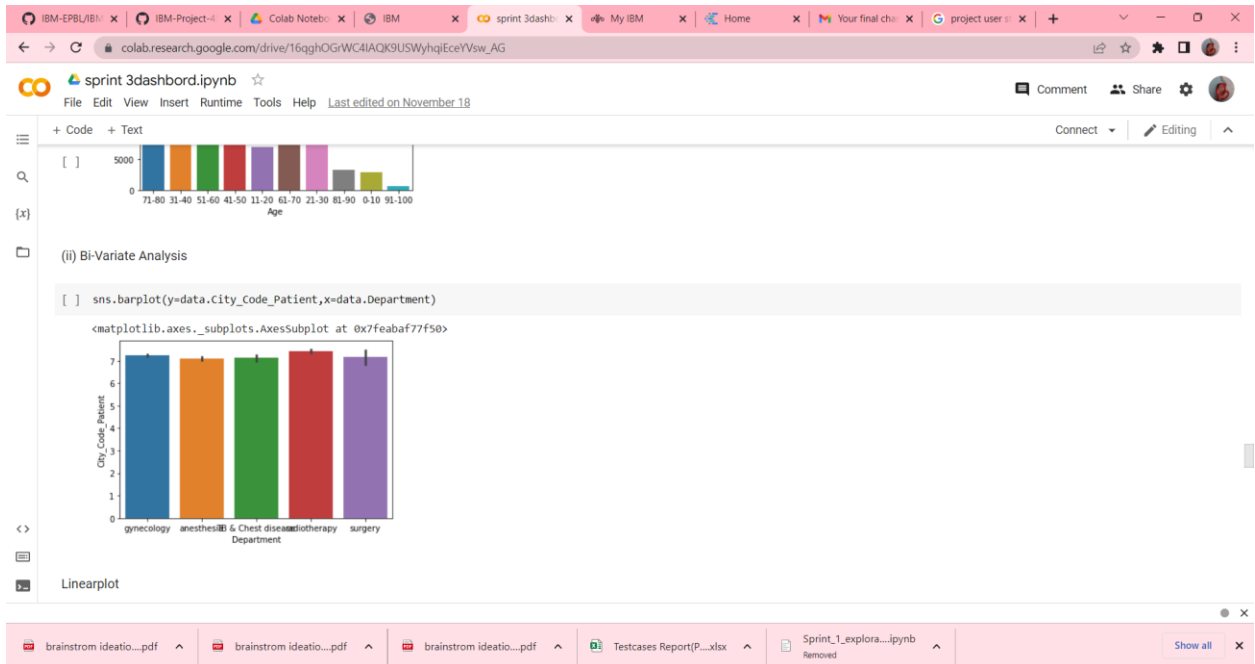
brainstrom ideatio...pdf brainstrom ideatio...pdf brainstrom ideatio...pdf Testcases Report(P...xlsx Sprint_1_explora...ipynb Removed Show all



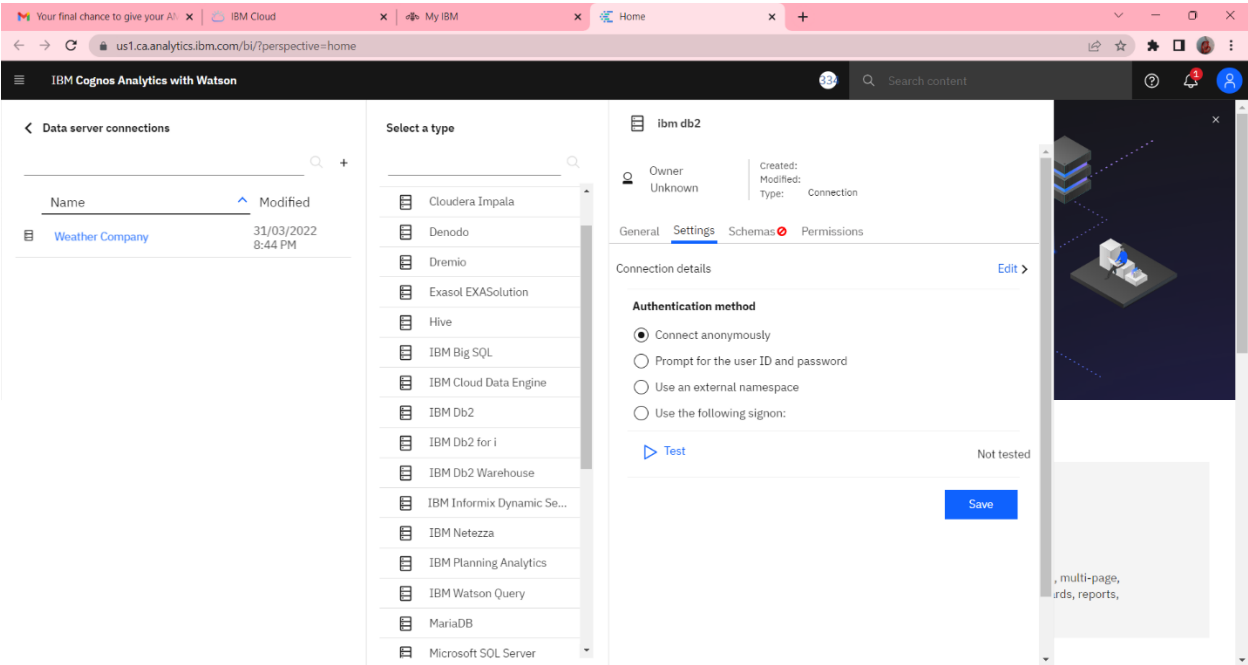
7.2 Feature 2 dashboard





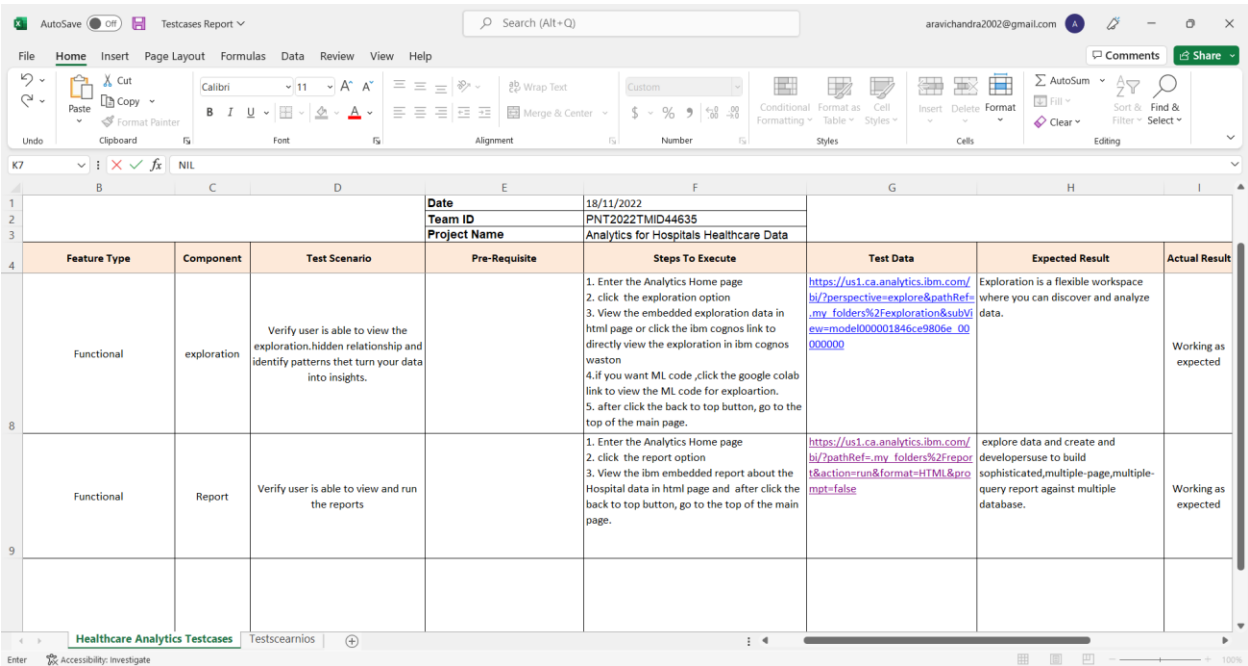


7.3 Database Schema (if Applicable)



8. TESTING

8.1 Test Cases



8.2 User Acceptance Testing

UAT(PNT2022TMID44635).pdf

File | C:/Users/Rv/Downloads/UAT(PNT2022TMID44635).pdf

1 of 2

Page view | Read aloud | Add text | Draw | Highlight | Erase

Acceptance Testing

UAT Execution & Report Submission

Date	18 November 2022
Team ID	PNT2022TMID44635
Project Name	Analytics for Hospital's Healthcare Data
Maximum Marks	4 Marks

1. Purpose of the Document:

The purpose of this document is to briefly explain the test coverage and open issues of the Analytics for Hospital's Healthcare Data project at the time of the release to User Acceptance Testing (UAT).

2. 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	9	3	0	3	15
Duplicate	1	0	3	0	4
External	3	2	0	1	6
Fixed	14	5	2	13	34
Not Reproduced	0	0	1	0	1
Skipped	0	0	2	1	3
Won't Fix	1	3	2	1	6

UAT(PNT2022TMID44635).pdf

File | C:/Users/Rv/Downloads/UAT(PNT2022TMID44635).pdf

2 of 2

Page view | Read aloud | Add text | Draw | Highlight | Erase

Totals	29	13	10	19	69
--------	----	----	----	----	----

3. Test Case Analysis:

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

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	8	0	0	8
Client Application	41	0	0	41
Security	2	0	0	2
Outsource Shipping	1	0	0	1
Exception Reporting	7	0	0	7
Final Report Output	10	0	0	10
Version Control	1	0	0	1

9. RESULTS

9.1 Performance Metrics

Performance Testing - Data Anal

File | C:/Users/Rv/Downloads/Performance%20Testing%20-%20Data%20Analytics(PNT2022TMID44635).pdf

1 of 4


Page view | Read aloud | Add text | Draw | Highlight | Erase

**Project Development Phase
Model Performance Test**

Date	18 November 2022
Team ID	PNT2022TMID44635
Project Name	Project – ANALYTICS FOR HOSPITAL HEALTH CARE DATA
Maximum Marks	10 Marks

Model Performance Testing:

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

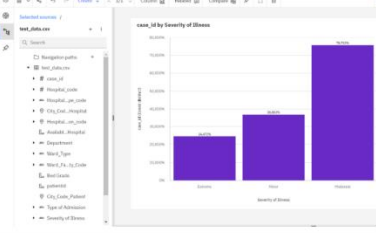
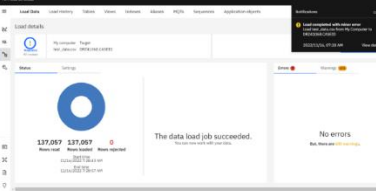
S.No	Parameter	Screenshot / Values
1.	Dashboard design	<p>No of Visualizations / Graphs -13</p> 
2.	Data	The visualization are responsive enough to view the data and the fit the

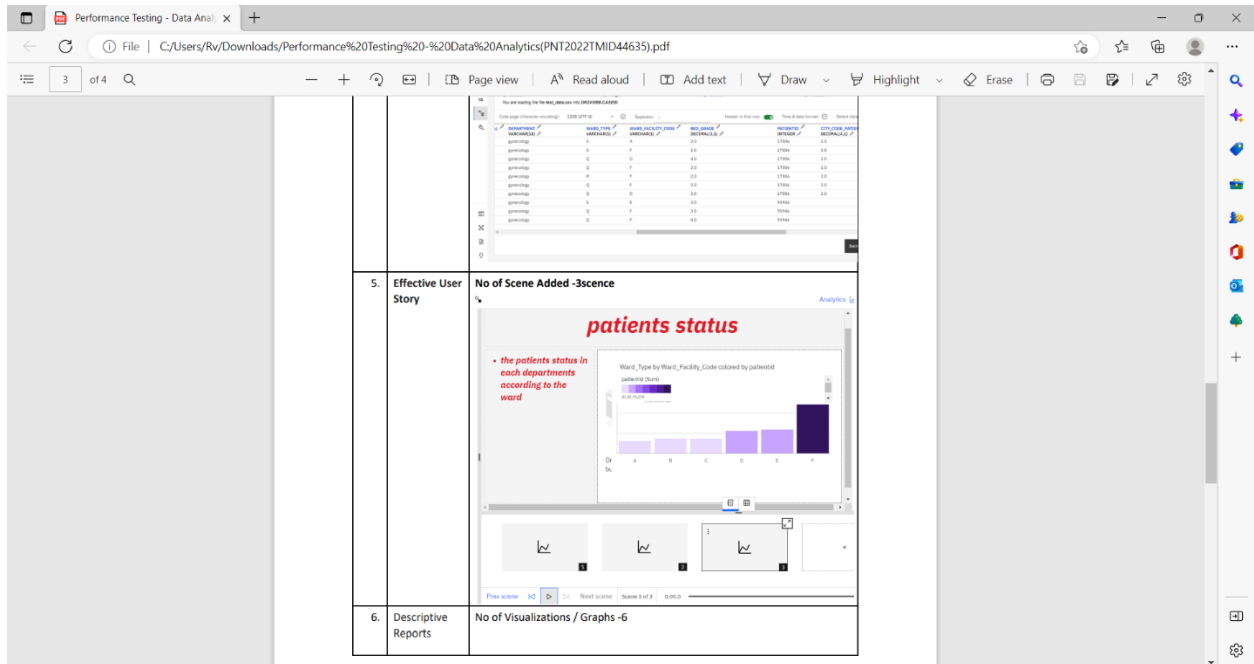
Performance Testing - Data Anal

File | C:/Users/Rv/Downloads/Performance%20Testing%20-%20Data%20Analytics(PNT2022TMID44635).pdf

2 of 4

Page view | Read aloud | Add text | Draw | Highlight | Erase

		
3.	Amount Data to Rendered (DB2 Metrics)	<p>Number of rows:137,057</p> <p>Number of loaded:137,057</p> 
4.	Utilization of Data Filters	The filters are used to see only the relevant data about the usecase



10. ADVANTAGES & DISADVANTAGES

Data analytics helps businesses get real-time insights about sales, marketing, finance, product development, and more. It allows teams within businesses to collaborate and achieve better results. It is useful for businesses to analyse past business performance and optimize future business processes.

Data Analyst: Pros: Writing queries, free food, doing analysis, forecasting, creating graphs, creating presentations, flexibility to schedule. Cons: Communication, menial tasks unrelated to job, not feeling valued. Data Driven Decision Making Through Data Analysis.

11. CONCLUSION

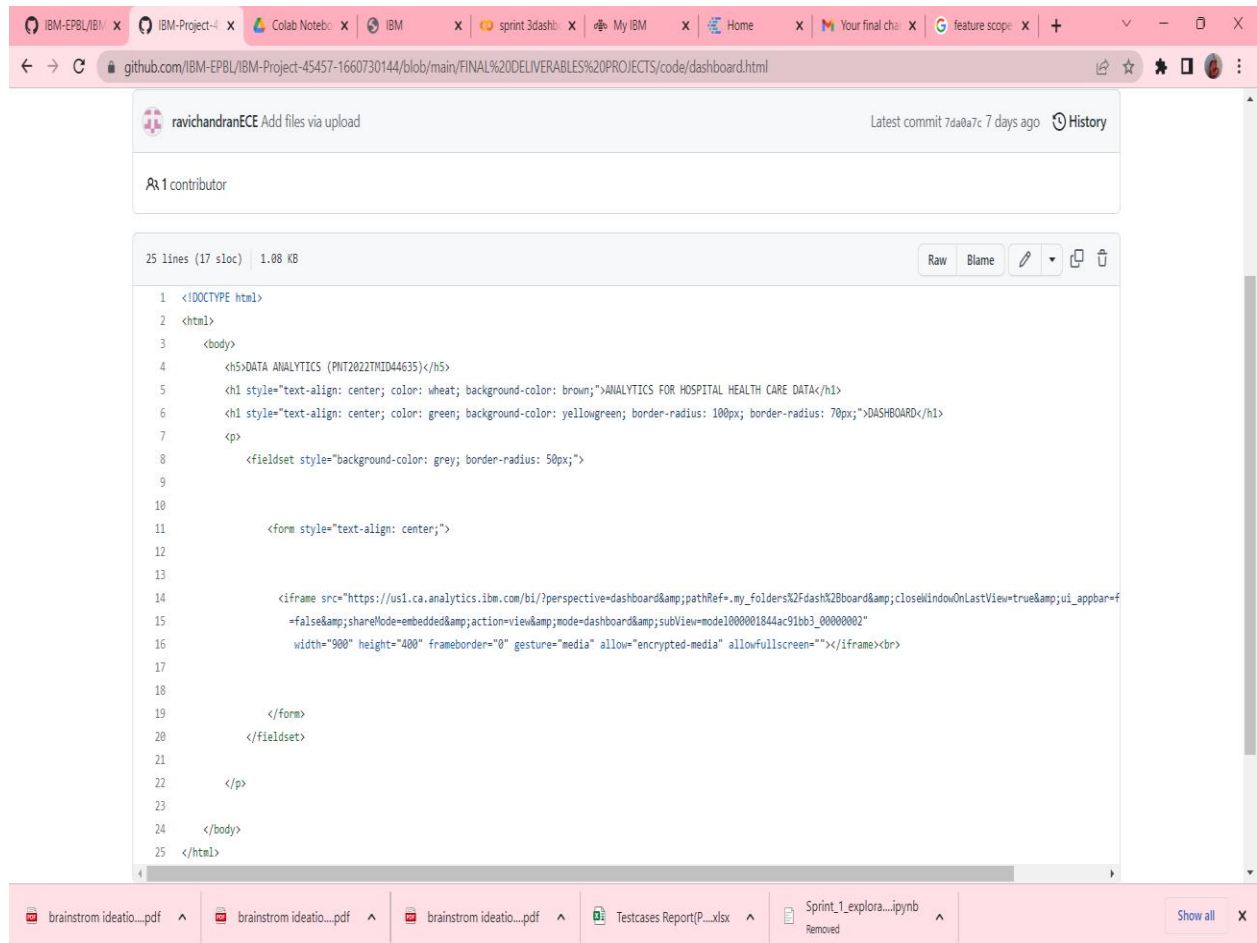
In this article we have explained the detail about explorative data analysis. We have used the language python programming language for implementation. We have used jupyter note book for detail analysis. We have implemented different library packages of python.

12. FUTURE SCOPE

Data Analytics is used to make sense of large amounts of data to derive insights and trends to improve business growth.

13. APPENDIX

Source Code



```
1 <!DOCTYPE html>
2 <html>
3 <body>
4 <h5>DATA ANALYTICS (PNT2022TMD44635)</h5>
5 <h1 style="text-align: center; color: wheat; background-color: brown;">ANALYTICS FOR HOSPITAL HEALTH CARE DATA</h1>
6 <h1 style="text-align: center; color: green; background-color: yellowgreen; border-radius: 100px; border-radius: 70px;">DASHBOARD</h1>
7 <p>
8 <fieldset style="background-color: grey; border-radius: 50px;">
9
10 <form style="text-align: center;">
11
12
13 <iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2Fdash%2Bboard&closeWindowOnLastView=true&ui_appbar=f
14 =false&shareNode=embedded&action=view&mode=dashboard&subView=mode1000001844ac91bb3_00000002"
15 width="900" height="400" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe><br>
16
17 </form>
18 </fieldset>
19
20 </p>
21 </body>
22 </html>
```

IBM-EPBL/IBM x IBM-Project-4 x Colab Noteb x IBM x sprint 3dashb x My IBM x Home x Your final cho x feature scope x + - x

github.com/IBM-EPBL/IBM-Project-45457-1660730144/blob/main/FINAL%20DELIVERABLES%20PROJECTS/code/exploration.html

28 lines (17 sloc) 1.05 KB

Raw Blame

```
1 <!DOCTYPE html>
2 <html>
3   <body>
4     <h5>DATA ANALYTICS (PHIT20227MID44635)</h5>
5     <h1 style="text-align: center; color: wheat; background-color: brown;">ANALYTICS FOR HOSPITAL HEALTH CARE DATA</h1>
6     <h1 style="text-align: center; color: green; background-color: palevioletred; border-radius: 100px; border-radius: 70px;">EXPLORATION</h1>
7     <p>
8       <fieldset style="background-color: gray; border-radius: 50px;"
9
10
11         <form style="text-align: center;"
12
13
14
15         <iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my_folders%2Fexploration&
16         closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false&shareMode=embedded&subView=mode1000001846ce9806e_000000000"
17         width="900" height="400" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe><br>
18
19
20
21
22
23       </form>
24     </fieldset>
25
26   </p>
27 </body>
28 </html>
```

brainstrom ideatio...pdf brainstrom ideatio...pdf brainstrom ideatio...pdf Testcases Report(P...xlsx Sprint_1_explora...ipynb Removed Show all x

IBM-EPBL/IBM x IBM-Project-4 x Colab Noteb x IBM x sprint 3dashb x My IBM x Home x Your final cho x feature scope x + - x

github.com/IBM-EPBL/IBM-Project-45457-1660730144/blob/main/FINAL%20DELIVERABLES%20PROJECTS/code/story.html

28 lines (17 sloc) 1.03 KB

Raw Blame

```
1 <!DOCTYPE html>
2 <html>
3   <body>
4     <h5>DATA ANALYTICS (PHIT20227MID44635)</h5>
5     <h1 style="text-align: center; color: wheat; background-color: brown;">ANALYTICS FOR HOSPITAL HEALTH CARE DATA</h1>
6     <h1 style="text-align: center; color: green; background-color: palevioletred; border-radius: 100px; border-radius: 70px;">STORY</h1>
7     <p>
8       <fieldset style="background-color: gray; border-radius: 50px;"
9
10
11         <form style="text-align: center;"
12
13
14
15         <iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.my_folders%2Fstories&closeWindowOnLastView=true&
16         ui_appbar=false&ui_navbar=false&shareMode=embedded&action=view&mode=dashboard"
17         width="900" height="400" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe><br>
18
19
20
21
22
23       </form>
24     </fieldset>
25
26   </p>
27 </body>
28 </html>
```

brainstrom ideatio...pdf brainstrom ideatio...pdf brainstrom ideatio...pdf Testcases Report(P...xlsx Sprint_1_explora...ipynb Removed Show all x

GITHUB LINK-<https://github.com/IBM-EPBL/IBM-Project-45457-1660730144>

DEMOLINK--[https://drive.google.com/file/d/1WGMsuBN-](https://drive.google.com/file/d/1WGMsuBN-F2LTssDmN6exx5ZwlLb8XrLx/view?usp=share_link)

[F2LTssDmN6exx5ZwlLb8XrLx/view?usp=share_link](https://drive.google.com/file/d/1WGMsuBN-F2LTssDmN6exx5ZwlLb8XrLx/view?usp=share_link)