TAMILNADU SKILL DEVELOPMENT CORPORATION

Naan Mudhalvan Niral Thiruvizha

BDA in Healthcare using Analytics tool

A Project Report

Submitted by

Yogeshwaran G

ABSTRACT:

Big Data Analytics (BDA) has garnered significant attention in both academia and industries, particularly in sectors such as healthcare, owing to the exponential growth of data and advancements in technology. The integration of data from diverse sources and the utilization of advanced analytical techniques has the potential to revolutionize healthcare by improving diagnostic accuracy, enabling personalized medicine, and enhancing patient outcomes. In this paper, we aim to provide a comprehensive literature review on the application of big data analytics in healthcare, focusing on its ecosystem, applications, and data sources. To achieve this, an extensive analysis of scientific studies published between 2013 and 2023 was conducted and overall, 180 scientific studies were thoroughly evaluated, establishing a strong foundation for future research and identifying collaboration opportunities in the healthcare domain. The study delves into various application areas of BDA in healthcare, highlights successful implementations, and explores their potential to enhance healthcare outcomes while reducing costs. Additionally, it outlines the challenges and limitations associated with BDA in healthcare, discusses modeling tools and techniques, showcases deployed solutions, and presents the advantages of BDA through various real-world use cases. Furthermore, this study identifies and discusses key open research challenges in the field of big data analytics in healthcare, aiming to push the boundaries and contribute to enhanced healthcare outcomes and decision-making processes.

1.1 TABLE OF CONTENTS:

2 INTRODUCTIONS:

Now let's dive into the project that we proposed. We spent quite some time finding all the references we needed just to test if our idea was possible or not. Luckily, it turns out that our project is very much possible to do in the current pace of the tech world.

2.1 OVERVIEW:

Our project "BDA in Healthcare with Prediction" is comprised of two sectors. One is the analytics sector, which is already seen many times on many platforms. The Analytics sector is used for all those big data techniques like data cleaning, data filtering, and data validation. This converts the raw and unstructured data into viable and structured data; it is the data with which we work and produce an overall visual representation for everyone's understanding. Analytics extract valuable insights from the raw data and we use those insights to push further our understanding of the data. The input normal data which is also known as raw data may consist of duplicates; it will be removed in analytics. It can also consist of wrong data which can be avoided by data validation which should be done while inputting the data.

The next sector is a special one that not many can just go on and create with a head start. That is a prediction. The prediction module uses past insights and analytics to predict the future. The prediction model has a saying "I use the past to judge the future" and it keeps its point and value up to the saying as well. When comparing both analytics and prediction, they both are very useful but they have a difficult difference between them. Prediction is more complicated than analytics in many aspects. It is because the complexity of prediction is much higher than analytics; such that, the user needs to train the model to its capacity and ability to predict future outcomes with history.

Our project has both analytics and prediction to help us in all ways possible. This project is slightly work different from the currently existing tools that are widely used in today's economy and business. Our idea was to replicate the process of the business tool's analytics process and convert/port it in the format of a file just to make sure of one thing – which is portability. Portability is only one of the aspects of this project idea, we will discuss about more aspects and information briefly in all the upcoming chapters and sections of this report.

2.2 PURPOSE:

The main purpose of this project is to show, to make people understand, and to imply the effectiveness of analytics using the tool. Raw data which is full of alphanumerical values and symbols as a dataset, will contain tons of data. It will be hard for anyone to understand anything or everything about it. It might contain duplicates of data or even wrong data. But if the same data was in a visual representation, everyone could at least get a grasp of what is going on and what are they trying to convey. This is the phase where data analytics tools come and handle the situation.

Our main goal for this project is to make people understand how important analytics is for healthcare or even any business field and also teach them the importance of analytics tools and their effectiveness. The analytics tool makes the process of an analyst so much easier; we don't need any paperwork or any unwanted/weighted records of data that store what's happening and what's going on for a prolonged period.

TOOL DETAILS:

We have to talk about the tool for sure since it's the main and core part of the project. The tool we were using is none other than "IBM Cognos Analytics" developed and produced by IBM themselves. Thanks to IBM for having a free trial version of their product which made this project possible in the first place. This free version allows users to experience all the privileged functions and operations for free for a limited period. Now back to the project, we use this tool for our all analytics techniques and processes including data cleaning, data filtering, data validation, and data visualization. IBM Cognos is one of the market's top analytics tools which is used by all those companies under IBM affiliation and some other companies as well. "Why do so many people use IBM Cognos?" the reason is so simple, it is because of its effectiveness and efficient way of processing all the commands that we, users provide.

Enhanced Decision-Making:

Data analytics empowers hospital administrators and healthcare professionals with real-time, evidence-based decision-making capabilities. Whether it's predicting patient admission rates, optimizing inventory levels, or identifying high-risk patients, analytics enables informed decision-making that drives better outcomes for both patients and the institution.

FIELD OF AIM:

The process of analytics and analytics tools are in every single successful field because every field needs the analytics results to improve themselves from their current position and to prevent themselves from any dangers that may not be foreseeable without the

analytics process. Nowadays, not a single company that aims to grow higher will ever avoid an analyst and analytics because, without the analytics, they lose so much from their hands the corporation. There is very little or near zero chance of surviving in the corporate industry without the help of analytics; that is also known as 'luck'. But luck won't last long and a company can't rely on luck to keep surviving in the corporate industry, analytics is the only way. Only through analytics, a company can improve itself, cut unwanted costs, find out and lessen wastage, and most importantly help to predict the future.

We can't stress this enough because data analytics is very essential for an organization or company's consistent growth. Here we are going to go all out to make the best use of our knowledge about the analytics field and the usage of IBM Cognos. From our project, you will understand the difference between raw data and insights, structured and unstructured data, data insights, and visualization. We get data from a healthcare organization to represent everything using the IBM Cognos tool.

3 LITERATURE SURVEY:

1] By Ashwin Belle (2015):

The rapidly expanding field of big data analytics has started to play a pivotal role in the evolution of healthcare practices and research. It has provided tools to accumulate, manage, analyze, and assimilate large volumes of disparate, structured, and unstructured data produced by current healthcare systems. Big data analytics has been recently applied to aid the process of care delivery and disease exploration. However, the adoption rate and research development in this space are still hindered by some fundamental problems inherent within the big data paradigm. In this paper, we discuss some of these major challenges with a focus on three upcoming and promising areas of medical research: image, signal, and genomics-based analytics.

Recent research that targets the utilization of large volumes of medical data while combining multimodal data from disparate sources is discussed. Potential areas of research within this field that can provide a meaningful impact on healthcare delivery are also examined.

2] By Revanth Sonnati (2017):

Data is a powerful resource which is found in many forms. Big data does not have a universal definition and it is discussed in different ways. The term big data is referred to describe the exponential growth of the data flow in various sectors which is too large to process using the available traditional database and software techniques. Often big data is presumed to be scary, yet it is an explosion in the field of information. It helps to perform various analytics, which can make an impact on economic growth, creating opportunities, and improving efficiency over other organizations. This significant heap of data is often defined as three-dimensional namely Volume, Velocity, and Variety which some even define with Veracity.

- Volume: Data volume is a contribution by various factors. It can be transactional data, which has been used through the years, or the data flow over social media. The volume of the data is the total quantity of the mass data within an organization. The volume of data generated in an organization increases daily at an unpredictable rate, which can be in petabytes and zeta bytes on the production activities and the type of the organization.
- Velocity: This refers to the data in the total data transmitted currently in an organization or motion. The speed of the data that an organization produces processes and analyzes normally keeps on accelerating. It influences the creation and delivery of the data from one point to the next. It is often time-sensitive.
- Variety: The variety, which is diverse in forms, type of data, and origin. It defines the complexity of the data and the Occurrences

of data. It is in any form like structured, semi-structured, and unstructured data. Some forms of structured data are Numerical data, traditional databases, business information, and unstructured data like Audio, Video, and Pictures.

3] By Ambigavthi and Sridharan (2018):

The pace of both digital innovation and technology disruption is refining the healthcare industry at an exponential rate. The large volume of healthcare data continues to mount every second, making it very difficult to find any form of useful information. Recently, big data has shifted the traditional way of data delivery into valuable insights using big data analytics methods. Big data analytics provides a lot of benefits in the healthcare sector to detect critical diseases at the initial stage and deliver better healthcare services to the right patient at the right time so that it improves the quality of life care. Big data analytics tools play an essential role in analyzing and integrating large volumes of structured, semi-structured, and unstructured vital data rapidly produced by various clinical, hospitals, other social web sources, and medical data lakes. However, there are several issues to be addressed in the current health data analytics platforms that offer technical mechanisms for data collection, aggregation, process, analysis, visualization, and interpretation. Due to the lack of detailed study in the previous literature, this article inspects the promising field of big data analytics in healthcare. This article examines the unique characteristics of big data, big data analytical tools, and different phases followed by the healthcare economy from data collection to the data delivery stage. Further, this article briefly summarizes the open research challenges with feasible findings, and then finally offers the conclusion.

4] By Sahil Varma (2021):

This chapter provides a better understanding and use cases of big data in healthcare. The healthcare industry generates a lot of data every day, and without proper analytical tools, it is quite difficult to extract meaningful data. It is essential to understand big data tools since traditional devices don't maintain this vast data, and big data solves the major issue in handling massive healthcare data. Health data from numerous health records are collected from various sources, and this massive data is put together to form big data. Conventional databases cannot be used for this purpose due to the diversity in data formats, so it is difficult to merge, and so it is quite impossible to process. With the use of big data, this problem is solved, and it can process highly variable data from different sources.

5] By Clinpath (2023):

The introduction of Big Data Analytics (BDA) in healthcare will allow to use of new technologies both in the treatment of patients and health management. The research is based on a critical analysis of the literature, as well as the presentation of selected results of direct research on the use of Big Data Analytics in medical facilities. The direct research was carried out based on a research questionnaire and conducted on a sample of 217 medical facilities in Poland. Literature studies have shown that the use of Big Data Analytics can bring many benefits to medical facilities, while direct research has shown that medical facilities in Poland are moving towards data-based healthcare because they use structured and unstructured data, reaching for analytics in the administrative, business, and clinical areas. The research positively confirmed that medical facilities are working on both structural data and unstructured data. The following kinds and sources of data can be distinguished: databases, transaction data, unstructured content of emails and documents, and data from devices and sensors. However, the use of data from social media is lower as in their activity they reach for analytics, not only in the administrative and business but also in the clinical area. It clearly shows that the

decisions made in medical facilities are highly data-driven. The results of the study confirm what has been analyzed in the literature that medical facilities are moving towards data-based healthcare, together with its benefits.

3.1 EXISTING SYSTEM:

Healthcare institutions are grappling with the challenge of effectively utilizing the vast amount of data generated within hospital systems. The existing problem revolves around inefficient data utilization, leading to suboptimal patient care, operational inefficiencies, and resource wastage. Hospitals are often overwhelmed with disparate data sources, ranging from electronic health records (EHRs) to medical imaging, patient demographics, billing information, and more. This data, if harnessed effectively, holds immense potential to drive insights for improving patient outcomes, operational workflows, and overall efficiency.

While preparing our Systematic Literature Survey we noticed several surveys have been undertaken in the extant literature to investigate the prospects and challenges associated with big data analytics and the healthcare domain. We also observed that current surveys remain focused on foundational basics and challenges in big data healthcare. Further authors discussed big data growth expectations for the year 2015, then statistics shown for spending by geography. Lastly, they enlighten healthcare infrastructure. Following a literature study presented by Raghupathi et al, they presented a comprehensive examination of the key attributes of big data, explored an architectural framework, and elucidated several application possibilities within the healthcare domain. Andreu-Perez et al performed a systematic literature review (SLR) spanning the years 2008 to 2015. The objective of their study was to offer a thorough examination of advancements in the field of biomedical and health informatics within big data. Luo et al conducted a comprehensive

examination of the recent progress made in the utilization of big data in several healthcare domains. The authors emphasized the substantial expansion observed within the last five years. In their study, Khanra et al did a systematic literature review (SLR) spanning the years 2013 to 2019. The authors identified and analyzed five distinct viewpoints about the application of big data analytics in the healthcare domain. Ikegwu et al conducted a systematic literature review (SLR) on the topic of big data analytics in data-driven industries. Their study aimed to explore the current state of knowledge in this area.

3.2 PROPOSED SYSTEM:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To predict the length of stay of patients and accommodation. To Evaluate the average waiting time of patients
2.	Idea / Solution description	The length of stay can be predicted using either Fuzzy logic or Tree-bagger algorithms. Along with the algorithm certain parameters like age, stage of disease, average waiting time progression, etc., are used for prediction. IBM Cognos is used for analytics purposes.
3.	Novelty / Uniqueness	It predicts the result with more accuracy using which overstays and average waiting time can be

		reduced. Proper resources and therapy can be provided.
4.	Social Impact / Customer Satisfaction	Patients can get better treatment and care than before. Length of stay and waiting time of patient prediction minimizes the overflow of patients therefore hospital resource management and utilization will be maximized. Reduces expense for treatment.
5.	Business Model (Revenue Model)	• This system can be used in all government hospitals, private hospitals, and even small clinics. • Activities – -Determine the average Satisfaction by age group. • Key Resource – Medical records. • Determine the average Wait Time by age group and Patient race
6.	Scalability of the Solution	This model will predict the length of stay and Patient visits every month for all kinds of patients. Determine the average Satisfaction by age group.

4 THEORETICAL ANALYSIS:

Let's get to all the theories and hypotheses that we collected and used for this project. Before we begin, I'd like to thank all the people who did so much research on this "BDA" because, from them and their ideology, we get a grasp of our idea and therefore begin our hypothesis and theories. Our theories are not entirely new to the analytics field or anything like that. The basic scratch was already developed, and we made some minor augmentations to the point, where our whole target is to help all of those who want to be a part of this project or even if you are someone who is reading as a journal.

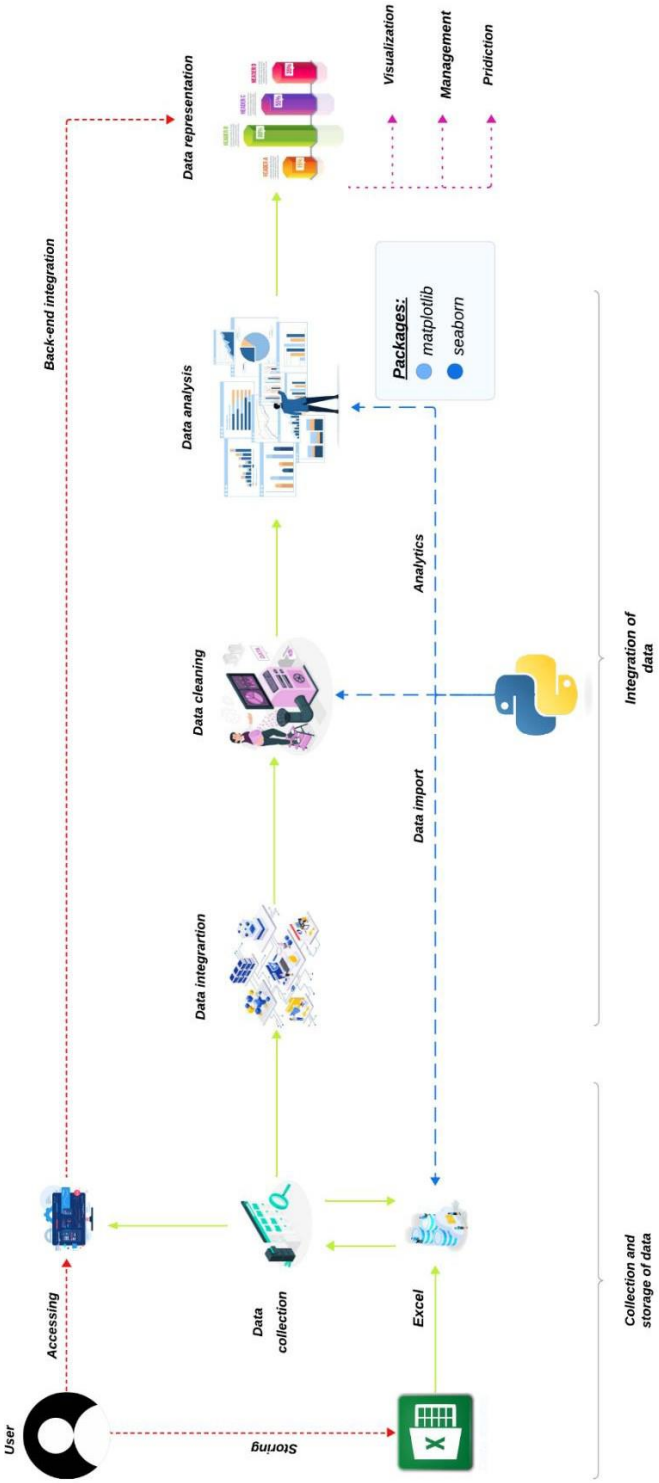
Ever wondered how companies analyze their data and get insights from them? Well, it is by using such tools as well as their analytics skillset and knowledge. Tool is an important part of analytical play role but it's not like anyone can use the tool and become a data analyst in a matter of time, no definitely not. The knowledge we possess is a type of skill set that we use to interpret things. Using tools also requires a specific skillset which helps as an integral part of analytics. The tool is not a single-function method; it consists of several sets of operations integrated into one platform and that platform is known as a tool.

Each operation has a specific set of instructions and commands that need to be followed to perform the initiated action. Some actions may support and correlate with other actions but some may be good if used alone. This statement will make more sense if you know the struggles we face in the representation section of the tool.

Now let's dive into visualizing the theory so we can get a better understanding of what is happening in this project and how it is happening in this project.

4.1 ARCHITECTURE DIAGRAM:

It represents what is happening in a diagrammatic format instead of only words. We will explain all modules one by one under this section itself, so hold tight.



The architecture diagram includes all modules and all processes that take place in the entirety of this analytics project. Let's go from start to end to get a full understanding of what's going on.

STORAGE:

Here is where all data is stored and the whole system process takes input from this storage module. But storage is not something we can say by words like 'we store this in here'. It's properly known as a database. Database all data is stored and organized but not only used for that. Also used for retrieving any data present inside the database.

The database is the starting of any field because it needs data to process anything. No one can get anywhere without data in a working field. For our project, we use MS Excel as our database. 'Why excel?' you may have this doubt. It is because Excel on of the best if not the best tools for storage and analytics platforms out there. Excel should be known to anyone who tries to learn or do analytics. Without Excel, no analyst can grow their skill set to become a senior or even superior analyst.

DATA INTEGRATION:

Now we store all of our data in our database which is Excel, what's next? Well, we have to integrate our database into our tool.

Integrating our database with our tool is to fetch our data on what we need or when we need it. Integration may be difficult on time due to the nature of overlapping or colliding data inside the database. An expert can easily avoid this but new users might not be aware of this.

To avoid this collision between data, we must clear unwanted data and duplicate data from our database. We use Excel as our database; we can easily clear all those chances of data collision in Excel itself. This is why Excel is a must-know for any analyst.

DATA CLEANING:

Data cleaning is also or should be done with our database storage i.e., excel. Removing unwanted data and clearing all duplicates from our raw dataset are the methods that come under data cleaning. We have mentioned duplicates and unwanted data for few times repeatedly. It is because it is so essential to get better results such as high-accuracy output and those data might affect the final result, especially in visualization and prediction.

Raw data without unwanted data processes faster and gives more accurate information about how we should play our data to ensure our efficiency and effectiveness.

DATA ANALYSIS:

It is the process of an analyst in the field of analytics. Now we have our data in our database where we removed all unwanted and duplicates and what we have is a pure or unique data set but still unstructured. You doubt what is structured and unstructured data. Well, unstructured data have no meaning whatsoever, and structured data will have a proper meaning to it.

Now we have what we need and now it's time to analyze all the data and extract meaningful insights from the raw data. These insights will give meaning to the data, hence turning unstructured data into structured data. The analysis is not a one-tap function that does whatever. It has layers of functions with each operation having a whole set of separate instructions. Analysts need to know what they are supposed to do and know what to do to get the results to further steps. Based on the outcomes, the analyst and com

pany decides to take steps for improvement. We need to make improvements at a low cost of expense.

DATA REPRESENTATION:

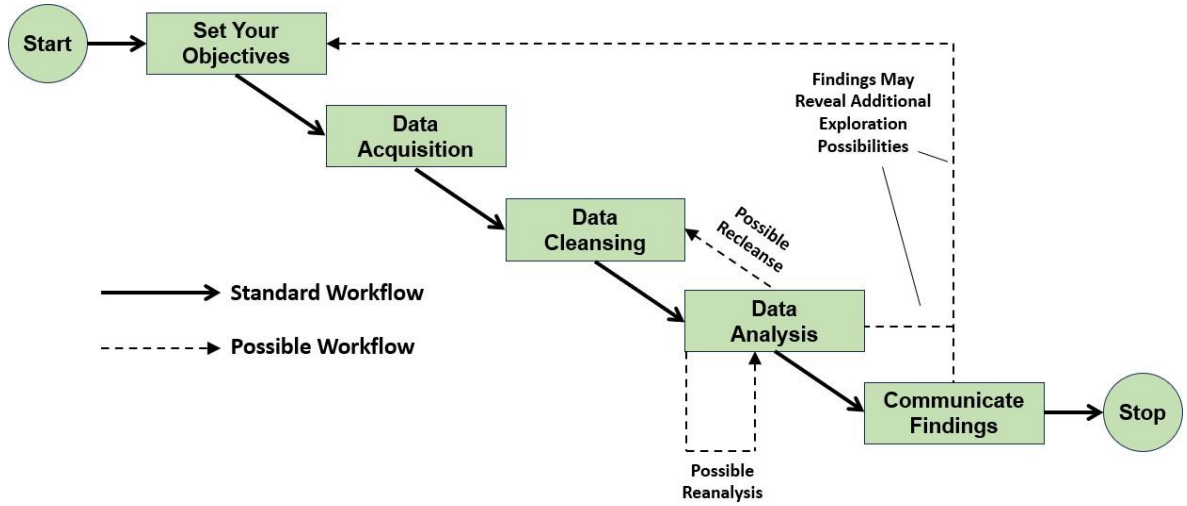
We discussed at the very beginning that visualization of data is better than usual raw data. So how or in what way do we visualize our valuable insights? Well, there isn't one particular way for data representation. There are a lot of ways such as bar charts, pie charts, graphs, histograms, line charts, etc.

OPTIONALITY:

PYTHON:

Python programming language is one of the best to support data analytics without tools. Python is not a must learn for data analytics but it is also useful in the fields of data analysis. Python means not the whole Python but only a few modules such as matplotlib, numpy, seaborn, pandas, and of course the basics of Python. These libraries provide functions and methods to calculate and display various statistical measures and tests, such as summary statistics, correlation, t-test, ANOVA, and regression. Data analysts can also use Python libraries to structure large datasets and make mathematical operations more manageable. Pandas, a Python library, offers a data

structure called a data frame to effectively work with large tables of data.



This is how the data analysis process is done in Python. If you notice, it is indeed quite similar to what we have seen in our project's architecture. Because analytics tools and analytics by Python do the same but with different functions and commands. Python requires a great amount of coding knowledge whereas in tools, all we need to do is place data and use commands instead of coding.

4.1.1 MODULES:

Module Overview:

This module provides a comprehensive overview of developing healthcare analytics solutions aimed at optimizing patient care and operational efficiency in healthcare institutions. It covers key concepts, methodologies, and best practices for leveraging healthcare data to derive actionable insights, streamline operations, and enhance overall healthcare delivery.

Module Objectives:

1. Understand the importance of healthcare analytics in improving patient outcomes and operational efficiency.

2. Learn essential data cleaning and preprocessing techniques for preparing healthcare data for analysis.

3. Explore predictive analytics methods for forecasting patient outcomes and identifying at-risk populations.

4. Discover descriptive analytics techniques for summarizing and visualizing key healthcare metrics.

5. Gain insights into prescriptive analytics approaches for providing actionable recommendations to healthcare practitioners and administrators.

6. Learn how to develop interactive dashboards and customizable reports for monitoring performance metrics and making data-driven decisions.

7. Understand compliance and security considerations in healthcare analytics, including HIPAA regulations and data protection measures.

8. Explore case studies and real-world examples of successful healthcare analytics implementations.

Module Structure:

1. Introduction to Healthcare Analytics

- Overview of healthcare analytics

- Importance of data-driven decision-making in healthcare

- Applications of analytics in optimizing patient care and operational efficiency

2. Data Cleaning and Preprocessing

- Importance of data quality in healthcare analytics
- Data profiling and exploration techniques
- Handling missing values, duplicates, outliers, and inconsistencies
- Standardization, normalization, and data transformation methods

3. Predictive Analytics in Healthcare

- Overview of predictive modeling techniques
- Forecasting patient outcomes
- Identifying at-risk populations
- Resource allocation and capacity planning

4. Descriptive Analytics

- Summarizing and visualizing healthcare metrics
- Dashboard design principles
- Key performance indicators (KPIs) in healthcare
- Trend analysis and pattern recognition

5. Prescriptive Analytics

- Overview of prescriptive analytics methods
- Providing actionable recommendations for healthcare improvement
- Optimization of care pathways and resource allocation
- Cost-effectiveness analysis

6. Data Visualization and Reporting

- Developing interactive dashboards
- Customizable reporting for healthcare stakeholders
- Visualization best practices for healthcare data

7. Compliance and Security

- HIPAA regulations and healthcare data privacy
- Ensuring compliance with regulatory requirements
- Data protection measures and security best practices

8. Case Studies and Real-world Examples

- Case studies of successful healthcare analytics implementations
- Lessons learned and best practices from real-world projects
- Opportunities and challenges in healthcare analytics

Module Conclusion:

This module provides a comprehensive understanding of healthcare analytics and its role in optimizing patient care and operational efficiency. By mastering data cleaning and preprocessing techniques, predictive and descriptive analytics methods, prescriptive analytics approaches, data visualization, and compliance considerations, participants will be equipped with the knowledge and skills to drive impactful healthcare analytics initiatives in their organizations.

4.2 HARDWARE REQUIREMENTS:

- Processor – Minimum Intel i5 (u)(h)
- RAM – Minimum 8GB
- Memory – Minimum 128 GB

Processor is for data processing power and computational capacity. The higher the power, the faster it performs. Meaning it will have more power to handle more data simultaneously and faster.

RAM is for the overall performance of the system. We have seen several modules above; all modules need a specific set of RAM from the system's RAM. If the RAM is already occupied by some other applications, then the current running will struggle to get RAM space to work properly.

Memory or ROM is essential to store our data. Our data will be stored in a database; our database will be stored in system memory unless we use cloud storage as a database such as Google's G-Drive or Microsoft's Azure or AWS. But for now, we are not going to use any virtual memory for our project, we use only physical memory for our project. Both virtual and physical memory have their ups and downs, nothing stays at the top with 100% win and 0% loss.

(Physical memory is recommended to use for analytics processes)

The above-mentioned things are just for minimum requirements to run and perform analytics in a system. The higher the specifications than the new minimum requirements, the better.

At the lowest minimum specifications, the system will struggle and slow down while processing. We can expect the system to even crash during the process, especially in other applications running in the background out of our analytics applications such as databases and tools.

4.3 SOFTWARE REQUIREMENTS:

- Operating System – Windows 10 or above
- Tool – IBM Cognos Analytics
- Database – Microsoft Excel 2007
- Text Editor – Visual Studio (At least Version 1.64)
- [Optional] Language – Python 3.9

OS is very important as we all know so I don't explain so much in here. Just remember it responds and performs what we say or instruct it to do. Windows 10 is a better version so please at least use that rather than Windows 9 if you want to perform analytics.

IBM Cognos Analytics, we also discussed this earlier. It is the tool platform in which we perform all our analytics.

MS Excel is the most basic tool that an analyst needs to master. Without Excel, one cannot progress further in this field whatsoever. It is so much necessary as a beginner, as an intermediate, as an expert.

[Optional] It is Python and its modules

Pythons can be a great substitute for analytics in the absence of any high-end tools. It consists of all the essential modules from the source code in the form of modules and packages. The downside is, that you need to code everything and if something goes awry, you have to debug it. People are not very fond of coding and dealing with all those errors to be recovered.

To use Python as an optional substitute, you need at least version 3.9 or higher from the Python website. Otherwise, you can use Visual Studio Code to run Python on it since it is an all-in-one platform. But VS code should be version 1.64 or higher because that's where it supports the analytics processing.

5 EXPERIMENTAL ANALYSIS:

5.1 RAW DATA:

We have seen so much of words in all the past sections, especially on theoretical analysis. It must have felt like too much yapping, sorry for that. Now let's dive into some practical and hands- on analysis.

Now here, have a look at this raw dataset, this is one of the many datasets we have and use;

patient_waittime	Adult	Early Childhood	Infancy	Midddle Childhood	Teenager	Summary
African American	52,644	3,536	1,750	6,293	5,249	69,472
Asian	29,352	1,737	685	2,639	2,973	37,386
Declined to Identify	27,478	1,965	1,054	2,756	2,739	35,992
Native American/Ala...	13,694	987	328	1,404	1,362	17,775
Pacific Islander	14,929	1,084	325	1,276	1,408	19,022
Two or More Races	42,714	2,983	1,423	4,090	3,842	55,052
White	69,622	4,137	1,912	7,519	7,066	90,256
Summary	**2,50,433**	**16,429**	**7,477**	**25,977**	**24,639**	**3,24,955**

You can't tell immediately what is going on here or what insights you could get from this raw dataset. It looks like some sort of random data number that is placed in each cell of Excel. Here is where data analytics comes into play. By now, you must have known what and why we use data analytics for. If not, you should go through it once again.

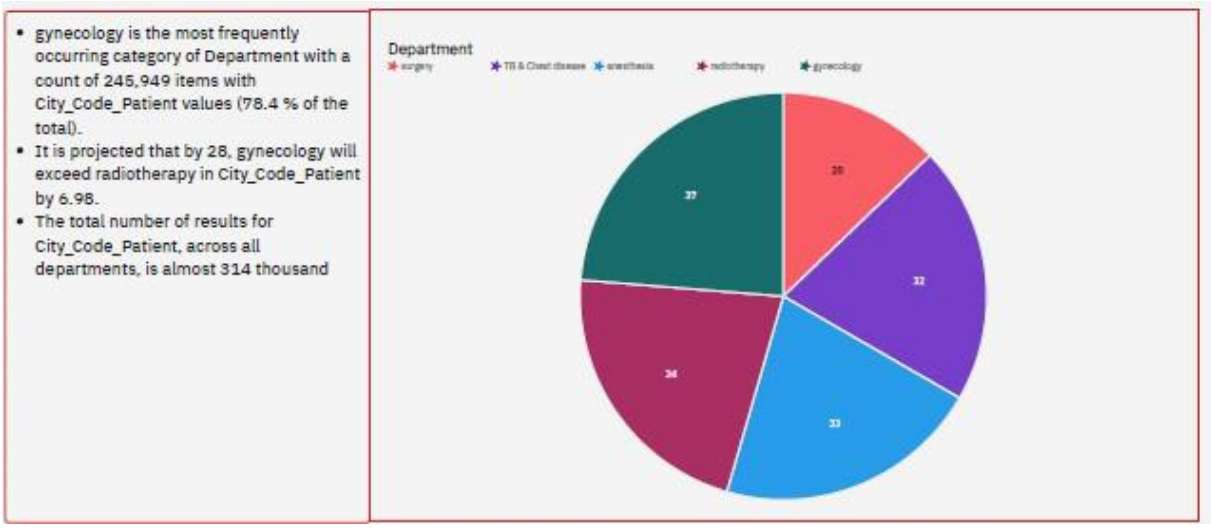
Here we have a total of seven (7) rows and eight (8) columns, each with distinct data and values. We have summary (totals) at the end of both the rows and columns which are also included in the previous counting. Now what do we analyze and how do we analyze this data? Yes by using the tool.

In the analysis phase, we create separate visual representations for easy insights we can extract. For example, we create a chart for adults, one for teenagers, and one for summary as well. By doing this,

we can extract meaningful insights and can able to tell what exactly is going on.

5.2 REPRESENTATION:

Now, it is time to convert data from the data set into visual representations with meaningful full insights;



We have created a pie chart representation from one of our datasets. Now we have seen how a dataset will look and also seen how data will be visually represented. Now we will see the in-between process of dataset and visualization.

Data Collection and Integration:

In the initial phase of our project to develop analytics for hospitals' healthcare data, we focused on collecting and integrating data from various sources within the hospital ecosystem. This process involved the following steps:

Identify Data Sources: Collaborate with relevant stakeholders to identify all potential sources of data within the hospital. This may include electronic health records (EHRs)

Extract Data: Utilize appropriate methods to extract data from each identified source

Data Cleaning and Preparation:

After the data collection and integration phase, we shifted our focus to cleaning and preparing the data for analysis. This involved the following steps:

Exploratory Data Analysis (EDA)

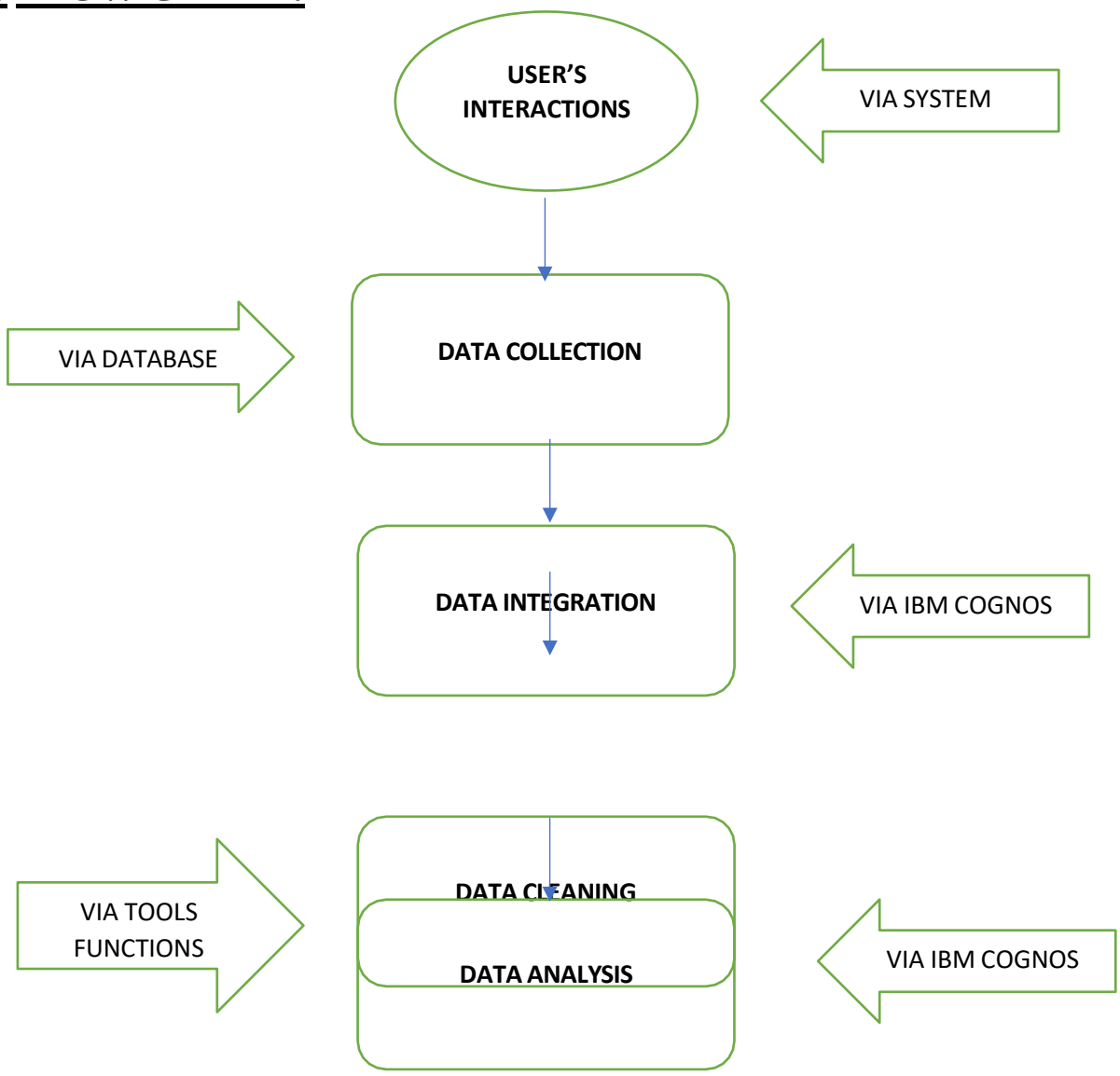
As part of the data preparation process, we conducted exploratory data analysis to gain insights into the underlying patterns and distributions within the data. This involved generating summary statistics, visualizing distributions, and identifying outliers.

Understanding the data:

Grid | Relationships | Custom tables

date	patient_id	patient_gender	patient_age	patient_sat_score	patient_last_name	patient_race	patient_waittime	department_referral
2020-06-15T11:29:36	316-34-3057	M	4	0	Methuen	Native American/Alaska Native	27	None
2020-02-04T22:34:29	358-31-9711	F	24	8	Titcombe	Native American/Alaska Native	31	General Practice
2020-09-04T17:48:27	289-26-0537	M	5	0	Gionettitti	African American	10	Orthopedics
2019-07-29T16:57:15	157-31-7520	F	47	0	Gwillim	Two or More Races	23	None
2020-03-10T22:02:15	393-38-9502	F	16	0	Corpes	White	39	Orthopedics
2019-05-04T13:16:12	662-21-6522	M	69	0	Stanlack	White	49	None
2019-09-04T16:15:52	728-31-2493	F	70	0	Chittock	Asian	50	Physiotherapy
2019-11-17T07:24:09	455-21-3671	F	11	4	Coste	Declined to Identify	30	None
2019-05-24T14:42:43	720-54-2625	F	4	0	Pavie	Native American/Alaska Native	23	None
2019-04-12T21:02:24	661-92-7059	M	42	0	Sleightholm	African American	51	None
2019-12-16T13:02:55	715-74-5338	M	22	0	Filkin	Pacific Islander	25	General Practice
2020-03-11T17:06:20	669-74-2146	F	58	0	Bilby	Declined to Identify	55	Orthopedics

6 FLOW CHART:



Let's discuss the workflow of the process in the data analytics process. Before that, I have to mention one thing; which is

that in any field, the data may vary but this process will remain the same in all fields.

Everything is done by the user's commands and the commands the system will perform. So, the motive of the user or analyst should be focused on gaining the required results.

All data will be collected and stored in a database where the data will be organized for storing and retrieving data.

Data integration must be done between the database and the tool we were about to use. This ensures that all required data is in line with fetching into the analytics process.

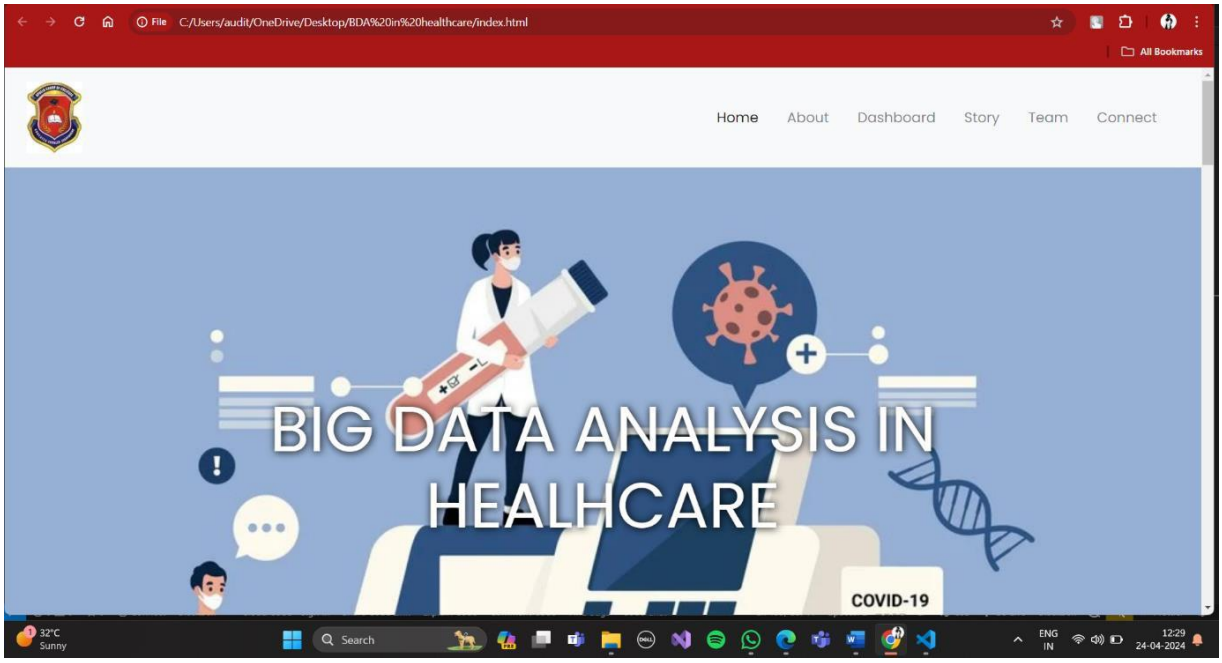
Data cleaning is done to remove unwanted data, wrong data, and remove duplicate data. This is also known as Data Filtering. By doing this, we avoid data collision, we can improve the processing speed of analytics instead of wasting in processing useless data.

Data analysis is the main core of this project as well as data analytics. It extracts meaningful insights from raw data, by converting unstructured data into structured data. By doing this, we as a user can get knowledge of what this data can be useful for and what can be improved from this set of data.

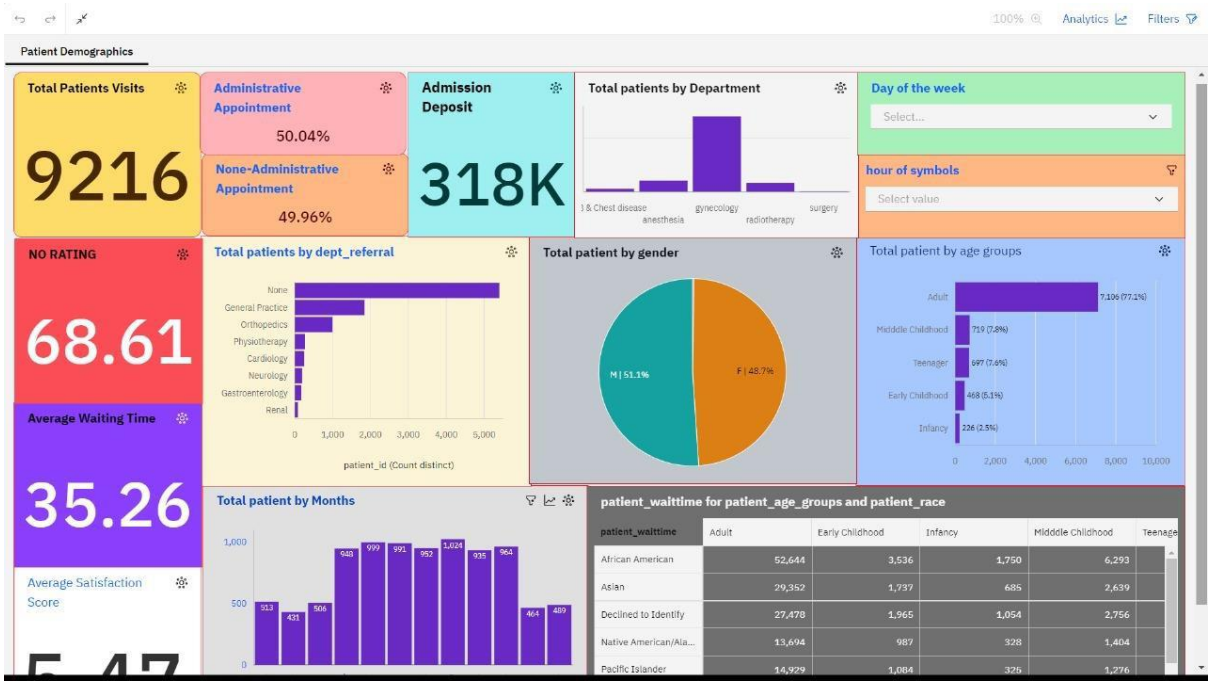
Finally, data visualization or data representation. It converts the insights of data into a visual representation in a diagrammatic format so anyone can get a grasp of what the data is trying to tell.

7 RESULTS:

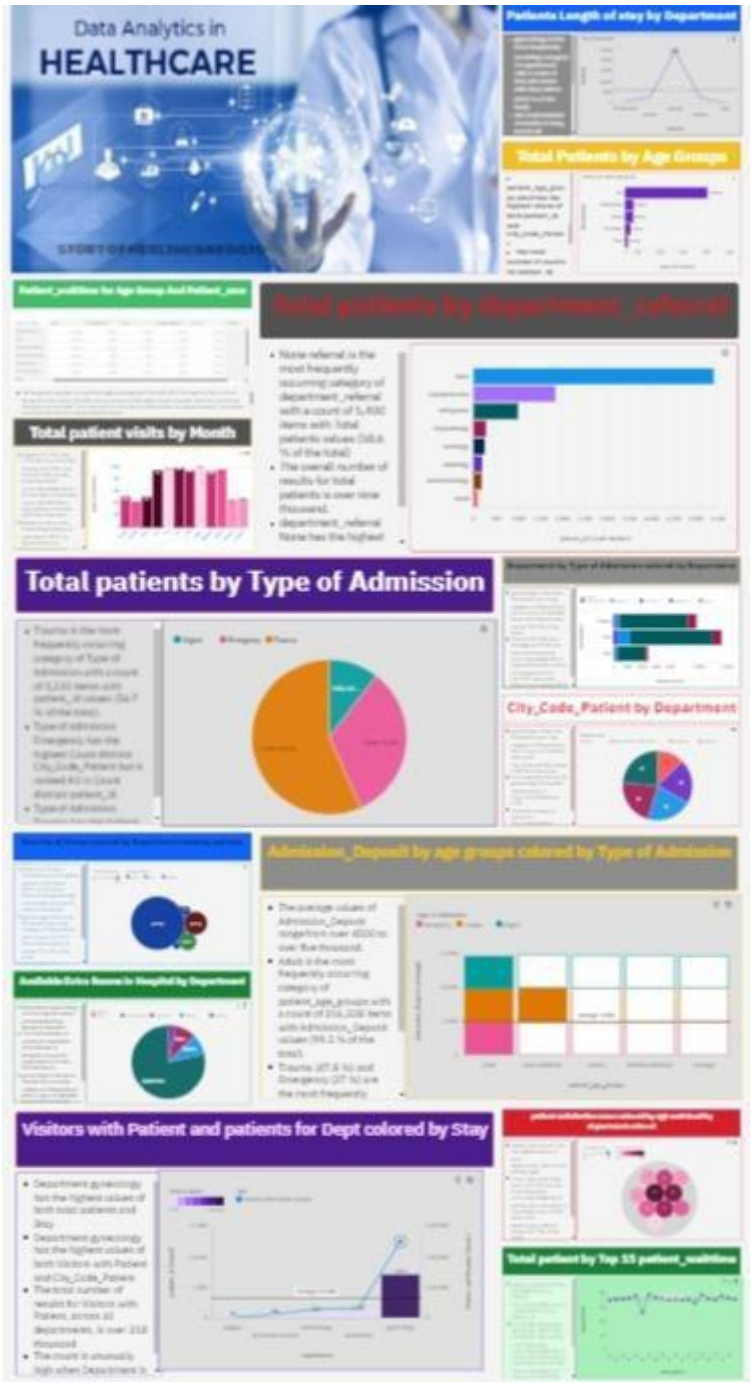
Once you get our Project, Click the index.html file to open the project. The website loads and we are presented with the home page. Then navigate to the dashboard.



After you click the dashboard on the navigation bar we will be presented with the below dashboard.



After you click the Story on the navigation bar we will be presented with the below story



8 ADVANTAGES & DISADVANTAGES:

Advantages:

1. Improved Patient Outcomes: By leveraging predictive analytics, healthcare providers can anticipate patient needs, identify risks early, and personalize treatment plans, leading to better health outcomes.

2. Enhanced Operational Efficiency: Prescriptive analytics can optimize resource allocation, streamline workflows, and reduce inefficiencies, ultimately improving the overall efficiency of healthcare institutions.

3. Cost Reduction: By optimizing resource utilization and reducing unnecessary procedures or treatments, the proposed solution can help healthcare institutions lower costs while maintaining or improving the quality of care.

4. Real-time Monitoring and Intervention: Real-time analytics enable proactive monitoring of patient data, allowing healthcare providers to intervene promptly in case of anomalies or critical events, thus preventing adverse outcomes.

5. Data-driven Decision Making: The solution provides healthcare providers with actionable insights derived from comprehensive data analysis, enabling informed decision-making at all levels of the organization.

6. Enhanced Patient Experience: By personalizing care plans and improving operational efficiency, the proposed solution can contribute to a better overall patient experience, leading to higher patient satisfaction and loyalty.

Disadvantages:

1. Complex Implementation Process: Implementing a comprehensive healthcare analytics solution requires significant investment in terms of time, resources, and infrastructure, which can pose challenges for healthcare institutions, especially smaller ones with limited budgets.

2. Data Integration Challenges: Integrating disparate data sources from various healthcare systems and ensuring data quality and consistency can be complex and time-consuming, potentially leading to delays and technical difficulties during implementation.

3. Privacy and Security Risks: Handling sensitive patient data raises concerns about privacy and security breaches. Despite robust security measures, there is always a risk of unauthorized access or data breaches, which could compromise patient confidentiality and trust.

4. Resistance to Change: Introducing a new analytics-driven approach may face resistance from healthcare professionals accustomed to traditional methods. Overcoming resistance to change and fostering a culture of data-driven decision-making may require organizational change management efforts.

5. Technical Challenges: Developing and maintaining advanced analytics algorithms and infrastructure requires specialized technical

expertise, which may be scarce or costly to acquire for some healthcare institutions.

6. Ethical Considerations: The use of predictive analytics in healthcare raises ethical concerns regarding data bias, fairness, and potential unintended consequences. Ensuring that analytics algorithms are transparent, accountable, and free from bias is essential to mitigate ethical risks.

Overall, while the proposed solution offers numerous benefits in terms of improving patient care, streamlining operations, and enhancing efficiency in healthcare institutions, addressing the associated challenges and risks is essential to ensure successful implementation and realization of its full potential.

9 APPLICATIONS:

1. Patient Care Optimization: Analytics can be applied to patient data to identify patterns, predict outcomes, and personalize treatment plans. This could include predicting readmission rates, identifying high-risk patients, and recommending preventive measures to improve patient outcomes.

2. Operational Streamlining: Hospitals can use analytics to streamline administrative processes such as scheduling, resource allocation, and inventory management. By analyzing data on patient flow, staffing levels, and equipment utilization, hospitals can optimize their operations to ensure efficient delivery of care.

3. Resource Allocation: Analytics can help hospitals allocate resources effectively by predicting patient demand, optimizing staffing levels, and identifying areas for cost reduction. This ensures that hospitals have the right resources in the right place at the right time to meet patient needs while minimizing waste.

4. Quality Improvement Initiatives: Hospitals can use analytics to monitor and improve the quality of care delivered to patients. By analyzing clinical outcomes, patient feedback, and adherence to best practices, hospitals can identify areas for improvement and implement targeted interventions to enhance the quality of care.

5. Risk Management and Compliance: Analytics can help hospitals identify and mitigate risks related to patient safety, regulatory compliance, and financial performance. By analyzing data on adverse events, compliance with quality metrics, and billing practices, hospitals can proactively address issues and ensure compliance with regulatory requirements.

6. Population Health Management: Hospitals can use analytics to manage the health of populations by identifying at-risk groups, implementing preventive interventions, and monitoring health outcomes over time. This can help hospitals improve population health outcomes and reduce healthcare costs by preventing or managing chronic conditions more effectively.

7. Research and Innovation: Analytics can support research and innovation in healthcare by providing insights into disease trends, treatment efficacy, and healthcare delivery models. By analyzing large

datasets, hospitals can generate new knowledge, drive evidence- based practice, and contribute to advancements in medical science.

8. Performance Monitoring and Benchmarking: Hospitals can use analytics to monitor performance metrics such as patient satisfaction, clinical outcomes, and financial performance. By benchmarking their performance against industry standards and peer institutions, hospitals can identify areas for improvement and drive continuous quality improvement initiatives.

Overall, the application of analytics to hospitals' healthcare data offers a wide range of opportunities to improve patient care, streamline operations, and enhance overall efficiency within healthcare institutions. By harnessing the power of data, hospitals can drive meaningful improvements in healthcare delivery and outcomes for patients.

10 CONCLUSIONS:

In conclusion, the development of analytics for hospitals' healthcare data presents a pivotal opportunity to revolutionize patient care, operational efficiency, and overall effectiveness within healthcare institutions. Through the optimization of data utilization, hospitals can harness the power of advanced analytics to drive informed decision-making, facilitate predictive and preventive measures, and ultimately improve patient outcomes. By streamlining operations and enhancing efficiency, healthcare institutions can allocate resources more effectively, reduce costs, and provide better-quality care to patients. The findings of this endeavor underscore the transformative potential of analytics in healthcare, highlighting the importance of ongoing innovation and investment in data-driven solutions to address the

complex challenges facing the industry. As we continue to advance in the realm of healthcare analytics, collaboration between stakeholders, ongoing research, and adaptation of emerging technologies will be key to unlocking the full potential of data-driven healthcare systems.

11 FUTURE ENHANCEMENTS:

1) Predictive Analytics for Disease Management: BDA can be used to analyze large volumes of healthcare data, including electronic health records (EHRs), genomics data, medical images, and wearable device data. This analysis can help in predicting disease patterns, identifying high-risk patients, and designing personalized treatment plans.

2) Real-Time Health Monitoring: With the advancement of IoT (Internet of Things) devices and wearable sensors, BDA can play a crucial role in real-time health monitoring. It can analyze streaming data from these devices to detect abnormalities, monitor vital signs, and provide early warnings for potential health issues.

3) Drug Discovery and Development: BDA can improve the efficiency of drug discovery by analyzing vast amounts of molecular and clinical data. It can identify potential drug candidates, predict drug responses based on genetic factors, and optimize clinical trial designs.

4) Healthcare Operations Optimization: BDA can optimize healthcare operations by analyzing data related to patient flow, resource utilization, and hospital management. It can help in reducing wait times, improving resource allocation, and enhancing overall healthcare delivery.

5) Fraud Detection and Prevention: BDA can be utilized to detect fraud and abuse in healthcare billing and claims. By analyzing patterns and anomalies in billing data, it can identify potentially fraudulent activities and prevent financial losses for healthcare providers.

12 APPENDICES:

12.1.SCREENSH

OT STEP 1

DATA COLLECTION AND INTEGRATION:

in the initial phase of our project to develop analytics for hospitals' healthcare data, we focused on collecting and integrating data from various sources within the hospital ecosystem. This process involved the following steps:

Identify Data Sources: Collaborate with relevant stakeholders to identify all potential sources of data within the hospital. This may include electronic health records (EHRs)

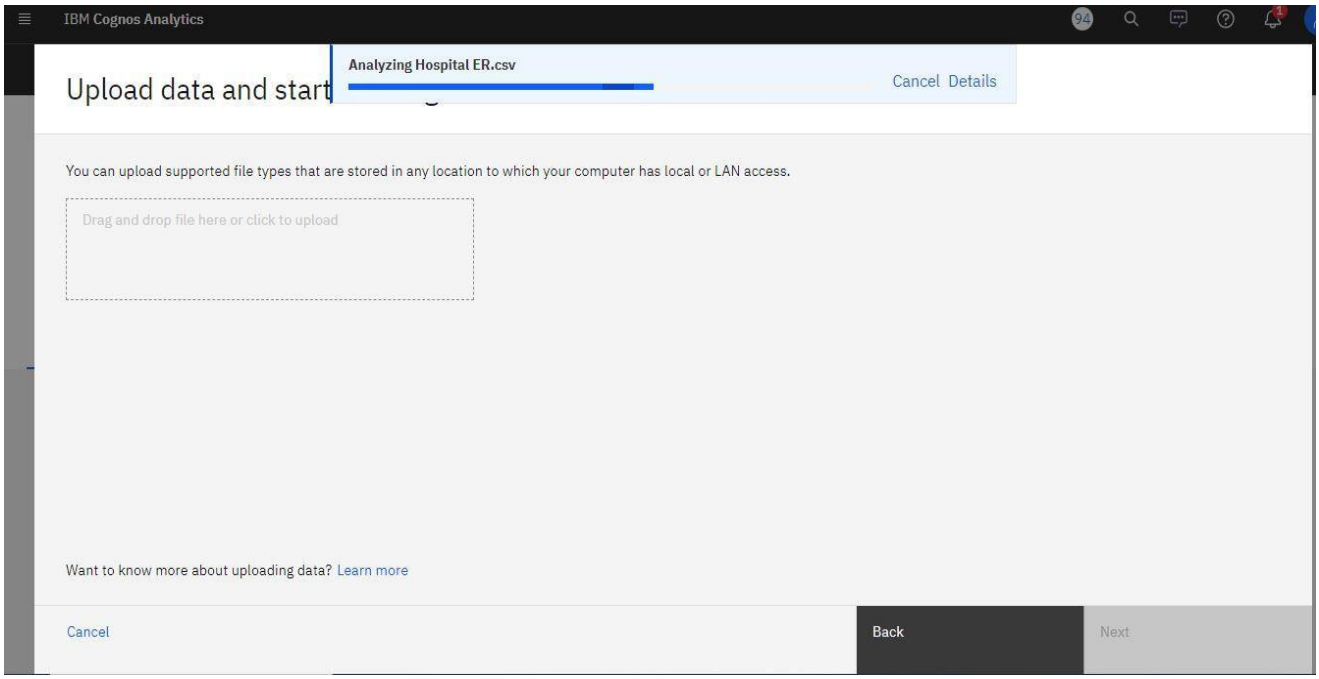
Extract Data: Utilize appropriate methods to extract data from each identified source

Download the dataset link:
https://drive.google.com/drive/folders/1p7iUIsbXUQ--XG1CQim-2plg3s5V4ueI

STEP 2

Uploading the dataset:

-Using IBM Cognos Analytics



Data Cleaning and Preparation

After the data collection and integration phase, we shifted our focus to cleaning and preparing the data for analysis. This involved the following steps:

Exploratory Data Analysis (EDA)

As part of the data preparation process, we conducted exploratory data analysis to gain insights into the underlying patterns and distributions within the data. This involved generating summary statistics, visualizing distributions, and identifying outliers.

In [1]:
```
!pip install numpy
```

```
Requirement already satisfied: numpy in c:\users\elcot\anaconda3\lib\site-packages (1.20.3)
```

In [2]:
```
#Required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

In [3]:
```
#load the data set
df=pd.read_csv('Hospital ER.csv')
df
```

Out[3]:

	date	patient_id	patient_gender	patient_age	patient_sat_score	patient_first_inital	patient_last_name	patient_race	patient_admin_flag	patient_waittime	department_referral
0	2020-03-20 08:47:01	145-39-5406	M	69	10.0	H	Glasspool	White	False	39	None
1	2020-06-15 11:29:36	316-34-3057	M	4	NaN	X	Methuen	Native American/Alaska Native	True	27	None
2	2020-06-20 09:13:13	897-46-3852	F	56	9.0	P	Schubuser	African American	True	55	General Practice
3	2020-02-04 22:34:29	358-31-9711	F	24	8.0	U	Titcombe	Native American/Alaska Native	True	31	General Practice
4	2020-09-04 17:48:27	289-26-0537	M	5	NaN	Y	Gionettitti	African American	False	10	Orthopedics
...	...	...	...	...	...	...	...	...	...	...	...
9211	2020-10-25 02:35:03	453-45-2632	M	60	NaN	A	Crich	African American	False	34	None
9212	2019-09-29 15:50:17	136-93-5822	M	63	NaN	A	Gerbl	White	False	55	None
9213	2020-08-12 14:46:37	822-41-7734	F	27	NaN	J	Dowall	Two or More Races	False	52	Orthopedics
9214	2019-10-11 16:46:15	260-14-4032	M	53	NaN	W	Rustan	White	True	51	None
9215	2019-06-01 07:48:20	703-55-3364	M	30	NaN	H	Hickeringill	Pacific Islander	True	19	None

9216 rows × 11 columns

In [4]:
```
df.head(10)
```

Out[4]:

	date	patient_id	patient_gender	patient_age	patient_sat_score	patient_first_inital	patient_last_name	patient_race	patient_admin_flag	patient_waittime	department_referral
0	2020-03-20 08:47:01	145-39-5406	M	69	10.0	H	Glasspool	White	False	39	None
1	2020-06-15 11:29:36	316-34-3057	M	4	NaN	X	Methuen	Native American/Alaska Native	True	27	None
2	2020-06-20 09:13:13	897-46-3852	F	56	9.0	P	Schubuser	African American	True	55	General Practice
3	2020-02-04 22:34:29	358-31-9711	F	24	8.0	U	Titcombe	Native American/Alaska Native	True	31	General Practice
4	2020-09-04 17:48:27	289-26-0537	M	5	NaN	Y	Gionettitti	African American	False	10	Orthopedics
5	2019-04-20 00:13:05	255-51-2877	M	58	NaN	H	Buff	Asian	False	59	None
6	2019-08-23 08:26:21	465-97-0990	F	68	NaN	F	Perrat	White	True	43	None
7	2019-07-29 16:57:15	157-31-7520	F	47	NaN	K	Gwillim	Two or More Races	True	23	None
8	2020-02-19 06:54:39	432-34-5614	F	79	1.0	E	Dewhirst	White	False	42	None
9	2020-10-11 05:25:17	609-17-8678	M	62	NaN	M	Crebo	African American	False	51	None

In [7]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9216 entries, 0 to 9215
Data columns (total 11 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   date                  9216 non-null   object 
 1   patient_id            9216 non-null   object 
 2   patient_gender        9216 non-null   object 
 3   patient_age           9216 non-null   int64  
 4   patient_sat_score     2517 non-null   float64
 5   patient_first_inital  9216 non-null   object 
 6   patient_last_name     9216 non-null   object 
 7   patient_race          9216 non-null   object 
 8   patient_admin_flag    9216 non-null   bool   
 9   patient_waittime      9216 non-null   int64  
 10  department_referral   9216 non-null   object 
dtypes: bool(1), float64(1), int64(2), object(7)
memory usage: 729.1+ KB
```

In [8]:
```
df.dtypes
```

Out[8]:
```
date                   object
patient_id             object
patient_gender         object
patient_age             int64
patient_sat_score     float64
patient_first_inital   object
patient_last_name      object
patient_race           object
patient_admin_flag       bool
patient_waittime        int64
department_referral    object
dtype: object
```

In [9]:
```
df.shape
```

Out[9]:
```
(9216, 11)
```

In [10]:
```
df.columns
```

Out[10]:
```
Index(['date', 'patient_id', 'patient_gender', 'patient_age',
       'patient_sat_score', 'patient_first_inital', 'patient_last_name',
       'patient_race', 'patient_admin_flag', 'patient_waittime',
       'department_referral'],
      dtype='object')
```

In [11]:
```
#BEFORE NULL VALUES CHECKING
df.isnull().sum().sum()
```

Out[11]:
```
6699
```

In [12]:
```
df.isnull()
```

Out[12]:

	date	patient_id	patient_gender	patient_age	patient_sat_score	patient_first_inital	patient_last_name	patient_race	patient_admin_flag	patient_waittime	department_referral
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	True	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	True	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...
9211	False	False	False	False	True	False	False	False	False	False	False
9212	False	False	False	False	True	False	False	False	False	False	False
9213	False	False	False	False	True	False	False	False	False	False	False
9214	False	False	False	False	True	False	False	False	False	False	False
9215	False	False	False	False	True	False	False	False	False	False	False

9216 rows × 11 columns

In [13]:
```
df.describe()
```

Out[13]:

	patient_age	patient_sat_score	patient_waittime
count	9216.000000	2517.000000	9216.000000
mean	39.855143	4.992054	35.259874
std	22.755125	3.138043	14.735323
min	1.000000	0.000000	10.000000
25%	20.000000	2.000000	23.000000
50%	39.000000	5.000000	35.000000
75%	60.000000	8.000000	48.000000
max	79.000000	10.000000	60.000000

In [15]:
```
df.isnull().sum()
```

Out[15]:
```
date                    0
patient_id              0
patient_gender          0
patient_age             0
patient_sat_score    6699
patient_first_inital    0
patient_last_name       0
patient_race            0
patient_admin_flag      0
patient_waittime        0
department_referral     0
dtype: int64
```

In [16]:
```
df.isnull().sum().sum()
```

Out[16]:
```
6699
```

In [18]:
```
#WORK WITH NULL VALUES
df['patient_sat_score']=df['patient_sat_score'].fillna(0)
```

In [19]:
```
df['patient_sat_score'].isnull().sum()
```

Out[19]:
```
0
```

In [20]:
```
df.isnull().sum()
```

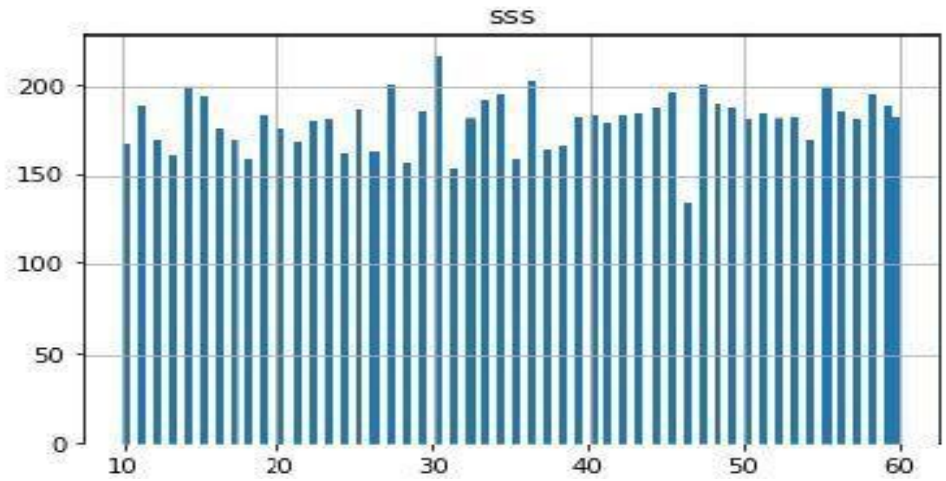
Out[20]:
```
date                 0
patient_id           0
patient_gender       0
patient_age          0
patient_sat_score    0
patient_first_inital 0
patient_last_name    0
patient_race         0
patient_admin_flag   0
patient_waittime     0
department_referral  0
dtype: int64
```

In [22]:
```
#CHECKING DUPLICATED VALUES
df.duplicated().sum()
```

Out[22]:
```
0
```

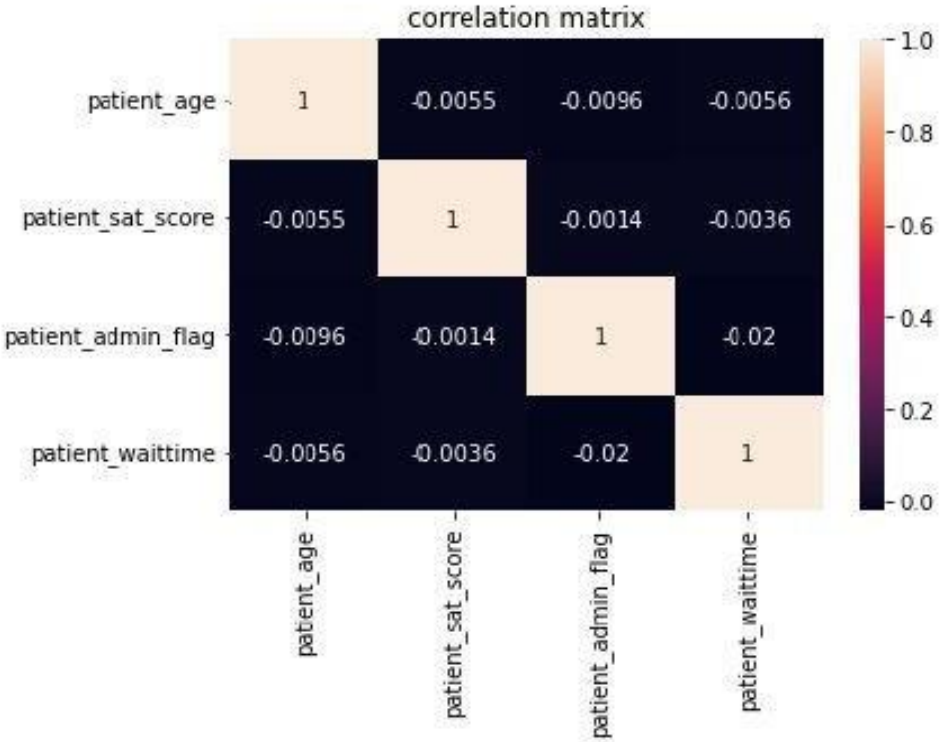
In [31]:

```
df["patient_waittime"].hist(bins=100)
plt.title("sss")
plt.show()
```

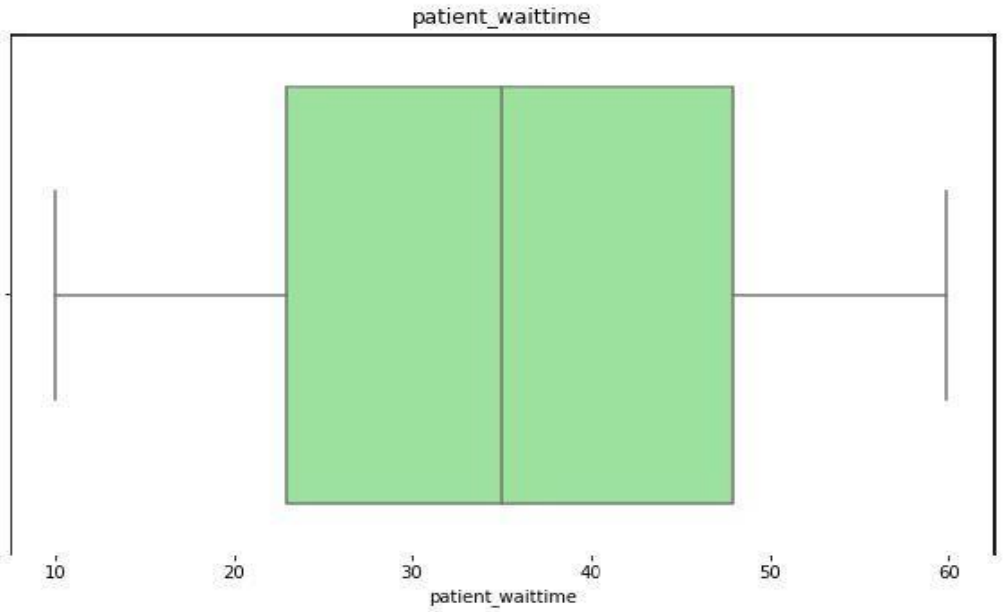


In [29]:

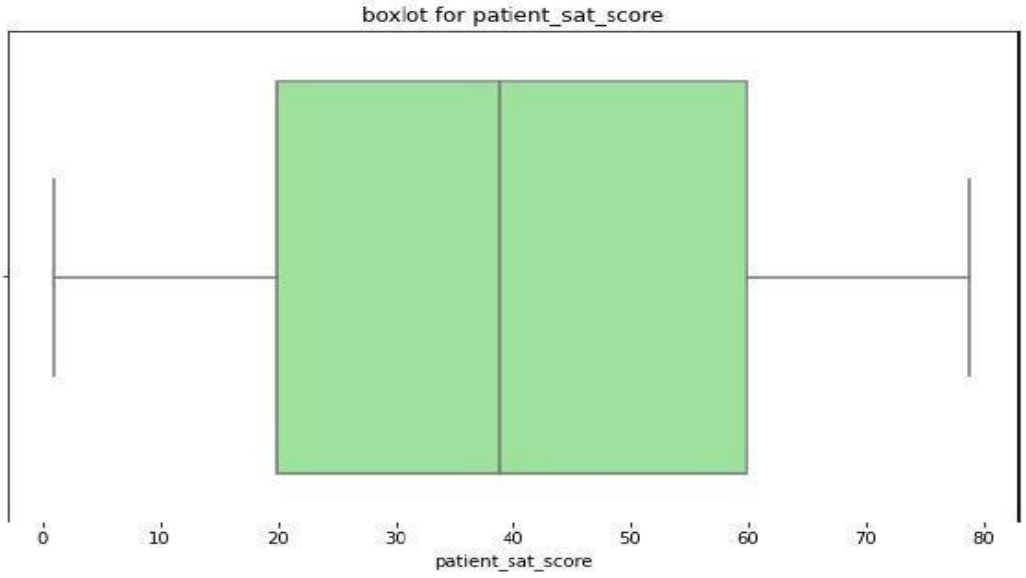
```
sns.heatmap(df.corr(),annot=True)
plt.title("correlation matrix")
plt.show()
```



In [50]:
```
#OUTLIER DETECTION
plt.figure(figsize=(10,6))
sns.boxplot(x='patient_waittime',data=df,color='lightgreen')
plt.title('patient_waittime')
plt.xlabel('patient_waittime')
plt.show()
```



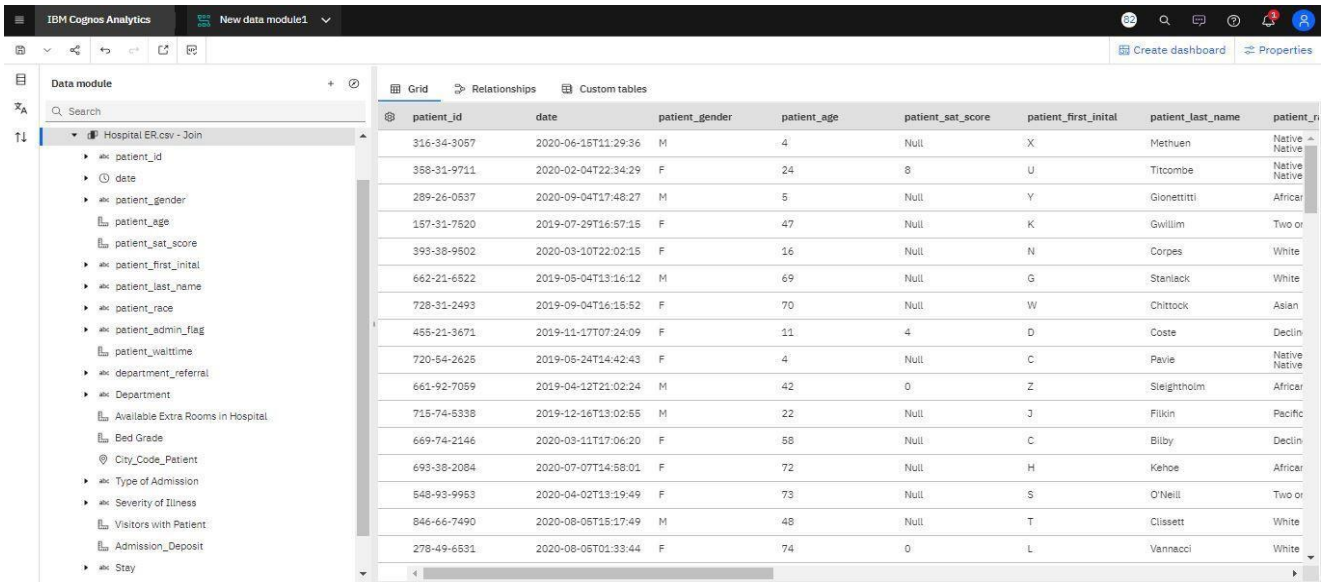
In [41]:
```
#patient_age
plt.figure(figsize=(10,6))
sns.boxplot(x='patient_age',data=df,color='lightgreen')
plt.title('boxlot for patient_sat_score')
plt.xlabel('patient_sat_score')
plt.show()
```



In [59]:
```
#TO EXPORT DATAFRAME TO CSV
df.to_csv("D:\download\Patients Healthcare dataset.csv",index=False)
```

Data warehousing:

Store the preprocessed data in a centralized data warehouse to make it easily accessible for analysis



UNDERSTANDING THE DATASET

Grid | Relationships | Custom tables

date	patient_id	patient_gender	patient_age	patient_sat_score	patient_last_name	patient_race	patient_waittime	department_referral
2020-06-15T11:29:36	316-34-3057	M	4	0	Methuen	Native American/Alaska Native	27	None
2020-02-04T22:34:29	358-31-9711	F	24	8	Titcombe	Native American/Alaska Native	31	General Practice
2020-09-04T17:48:27	289-26-0537	M	5	0	Gionettitti	African American	10	Orthopedics
2019-07-29T16:57:15	157-31-7520	F	47	0	Gwillim	Two or More Races	23	None
2020-03-10T22:02:15	393-38-9502	F	16	0	Corpes	White	39	Orthopedics
2019-05-04T13:16:12	662-21-6522	M	69	0	Stanlack	White	49	None
2019-09-04T16:15:52	728-31-2493	F	70	0	Chittock	Asian	50	Physiotherapy
2019-11-17T07:24:09	455-21-3671	F	11	4	Coste	Declined to Identify	30	None
2019-05-24T14:42:43	720-54-2625	F	4	0	Pavie	Native American/Alaska Native	23	None
2019-04-12T21:02:24	661-92-7059	M	42	0	Sleightholm	African American	51	None
2019-12-16T13:02:55	715-74-5338	M	22	0	Filkin	Pacific Islander	25	General Practice
2020-03-11T17:06:20	669-74-2146	F	58	0	Bilby	Declined to Identify	55	Orthopedics

Here's a list of descriptions for each column:

1. date: Date of the patient record entry.

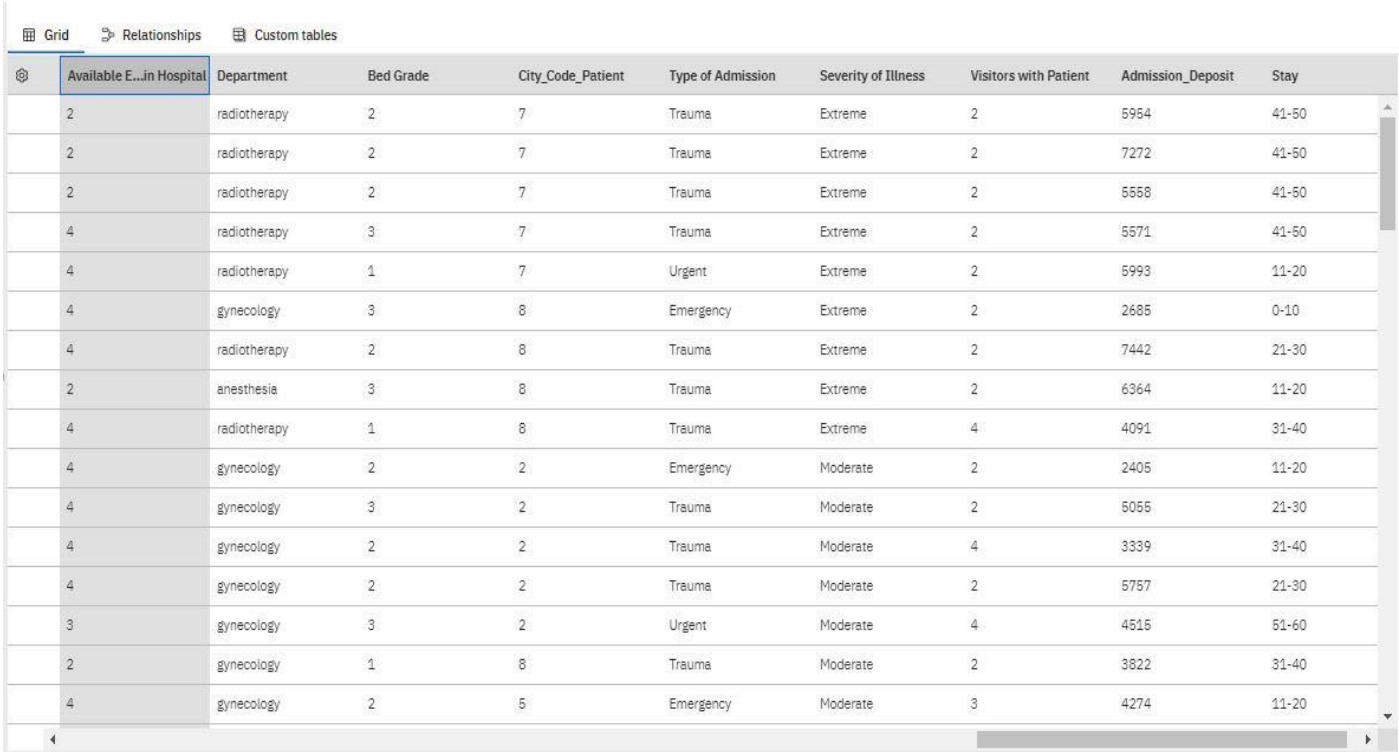
2.patient_id: Unique identifier for each patient.

3.patient_gender: Gender of the patient.

4.patient_age: Age of the patient in years.

5.patient_sat_score: SAT score of the patient (if applicable). **6.patient_first_inital:** First initial of the patient's first name. **7.patient_last_name:** Last name of the patient.

8. **patient_race:** Race of the patient.
9. **patient_admin_flag:** Boolean flag indicating whether the patient has administrative privileges.
10. **patient_waittime:** Wait time of the patient in minutes.
11. **department_referral:** The department to which the patient was referred.



Grid | Relationships | Custom tables

Available E...in Hospital	Department	Bed Grade	City_Code_Patient	Type of Admission	Severity of Illness	Visitors with Patient	Admission_Deposit	Stay
2	radiotherapy	2	7	Trauma	Extreme	2	5954	41-50
2	radiotherapy	2	7	Trauma	Extreme	2	7272	41-50
2	radiotherapy	2	7	Trauma	Extreme	2	5558	41-50
4	radiotherapy	3	7	Trauma	Extreme	2	5571	41-50
4	radiotherapy	1	7	Urgent	Extreme	2	5993	11-20
4	gynecology	3	8	Emergency	Extreme	2	2685	0-10
4	radiotherapy	2	8	Trauma	Extreme	2	7442	21-30
2	anesthesia	3	8	Trauma	Extreme	2	6364	11-20
4	radiotherapy	1	8	Trauma	Extreme	4	4091	31-40
4	gynecology	2	2	Emergency	Moderate	2	2405	11-20
4	gynecology	3	2	Trauma	Moderate	2	5055	21-30
4	gynecology	2	2	Trauma	Moderate	4	3339	31-40
4	gynecology	2	2	Trauma	Moderate	2	5757	21-30
3	gynecology	3	2	Urgent	Moderate	4	4515	51-60
2	gynecology	1	8	Trauma	Moderate	2	3822	31-40
4	gynecology	2	5	Emergency	Moderate	3	4274	11-20

Here's another list of descriptions for each column:
12. **Available Extra Rooms in Hospital:** Number of extra rooms available in the hospital for admission.
13. **Department**: The department to which the patient is admitted.
14. **Bed Grade:** The grade of the bed is allocated to the patient.
15. **City_Code_Patient:** Code representing the city where the patient is from.
16. **Type of Admission:** The type of admission for the patient (e.g., emergency, urgent, elective).
17. **Severity of Illness:** Severity of the patient's illness.

18. Visitors with Patient: Number of visitors accompanying the patient.

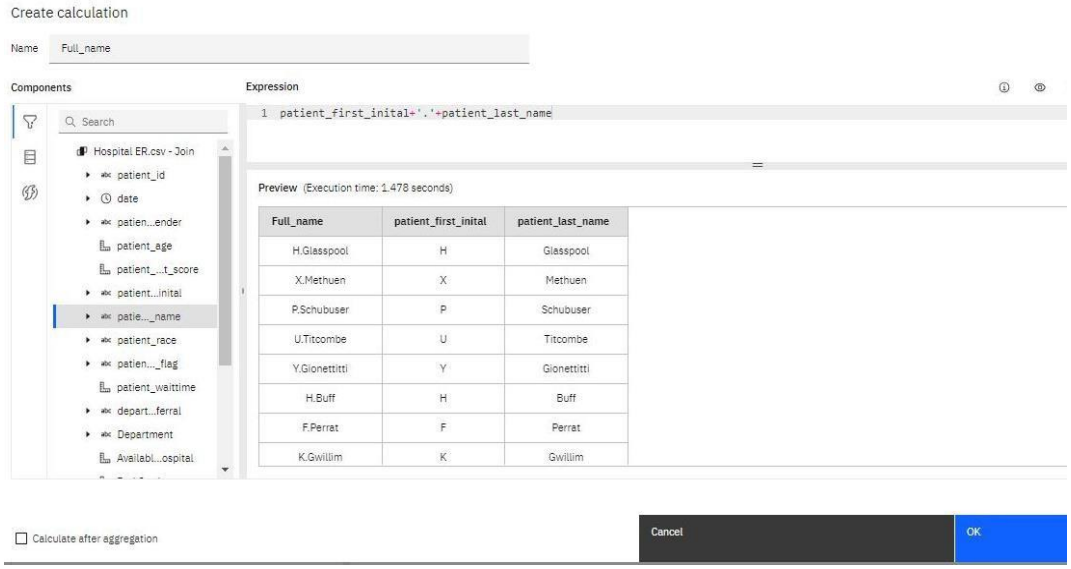
19. Admission_Deposit: Amount of deposit made during admission.

20. Stay Length of stay of the patient in the hospital.

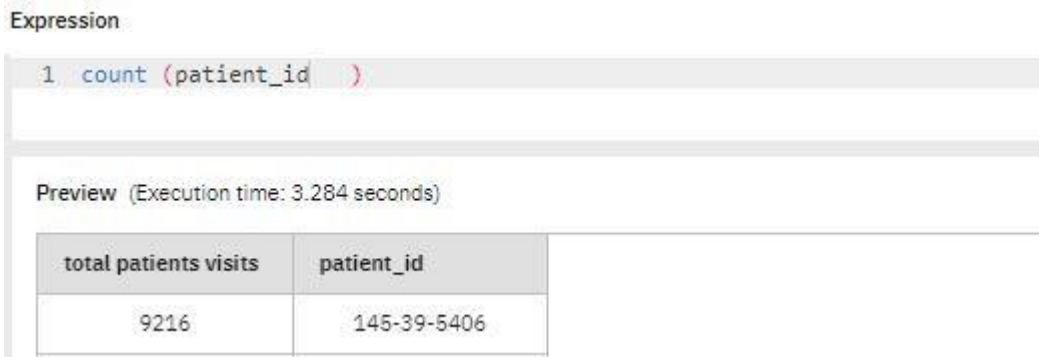
These descriptions provide a quick overview of each column's content and purpose in the dataset.

Data module in IBM Cognos analytics:

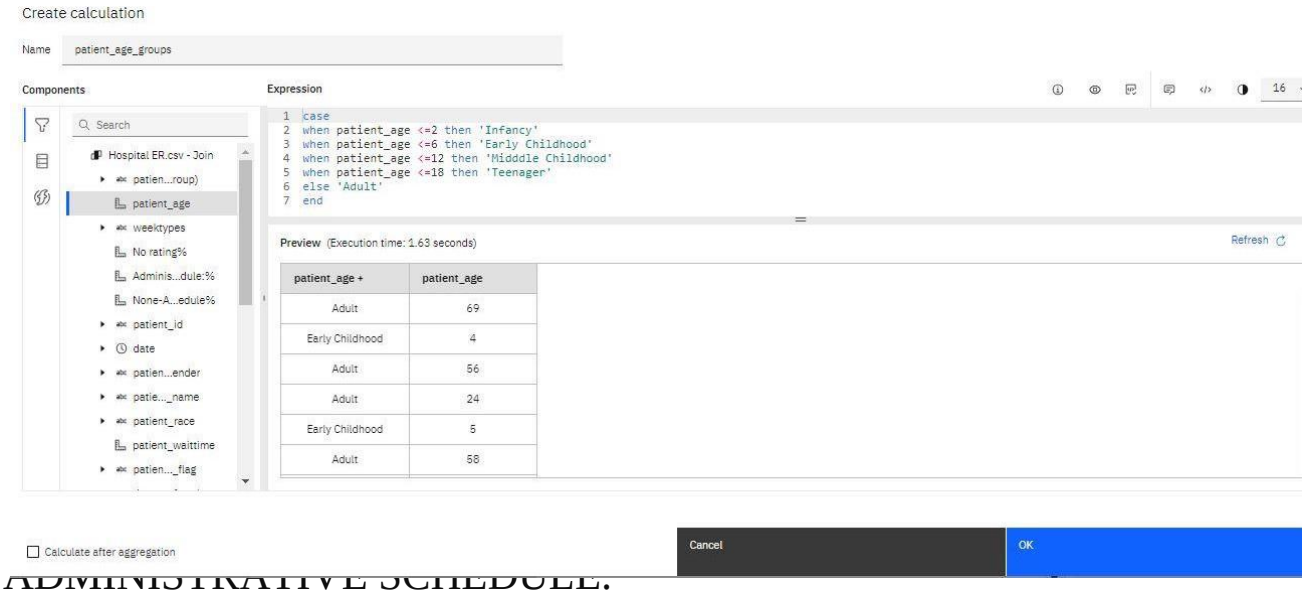
- COMBINE TWO COLUMN:



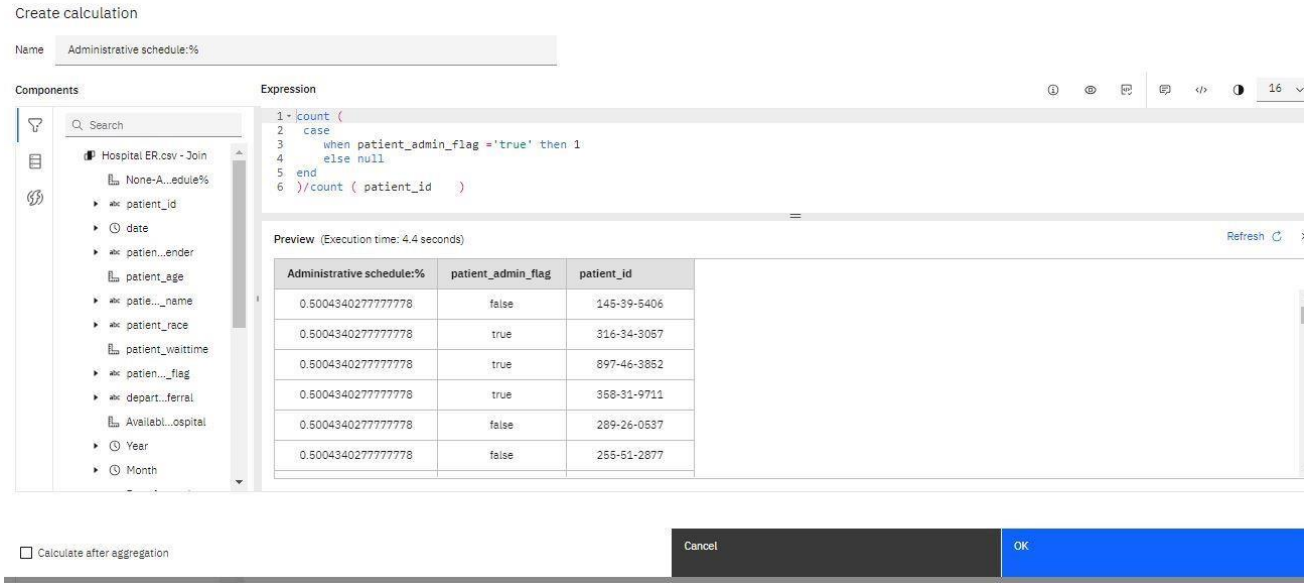
- CALCULATED TOTAL PATIENTS



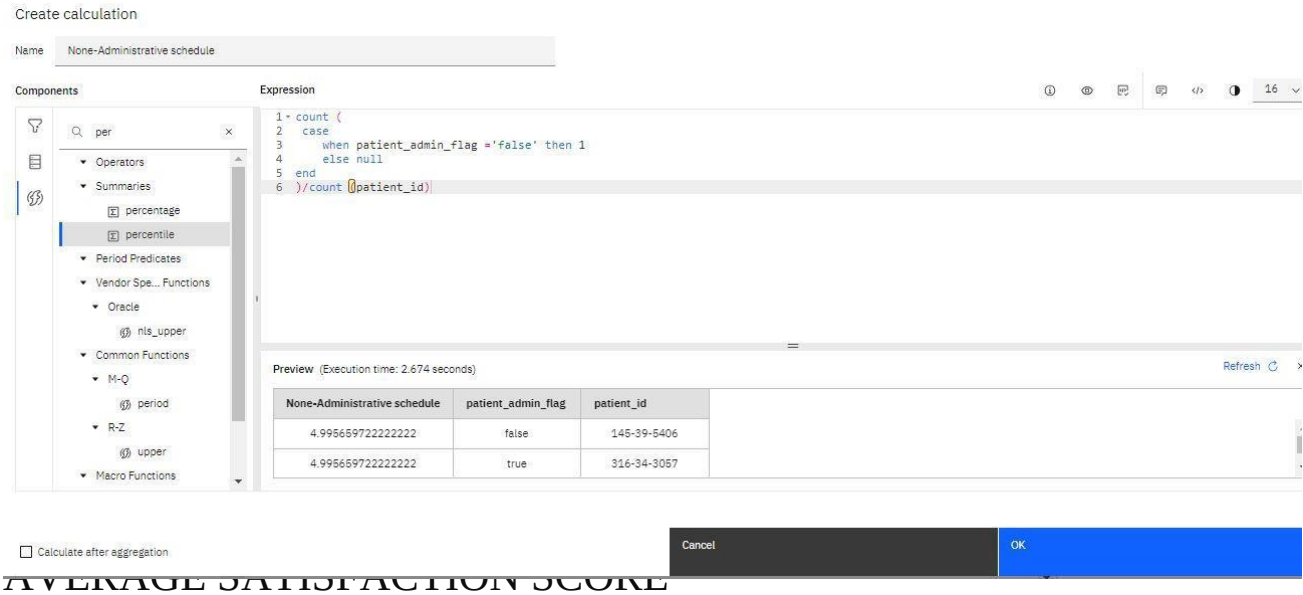
- TO CREATE PATIENTS AGE GROUPS FROM PATIENT'S AGE



- ADMINISTRATIVE SCHEDULE:

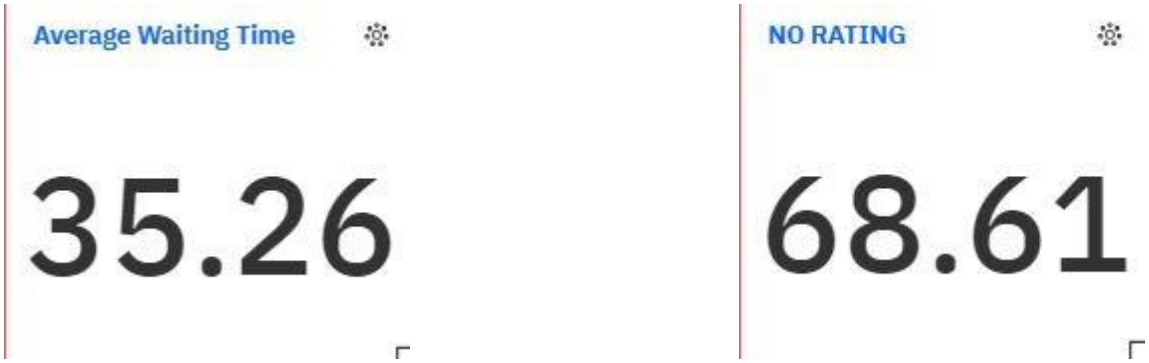


- NONE-ADMINISTRATIVE SCHEDULE:

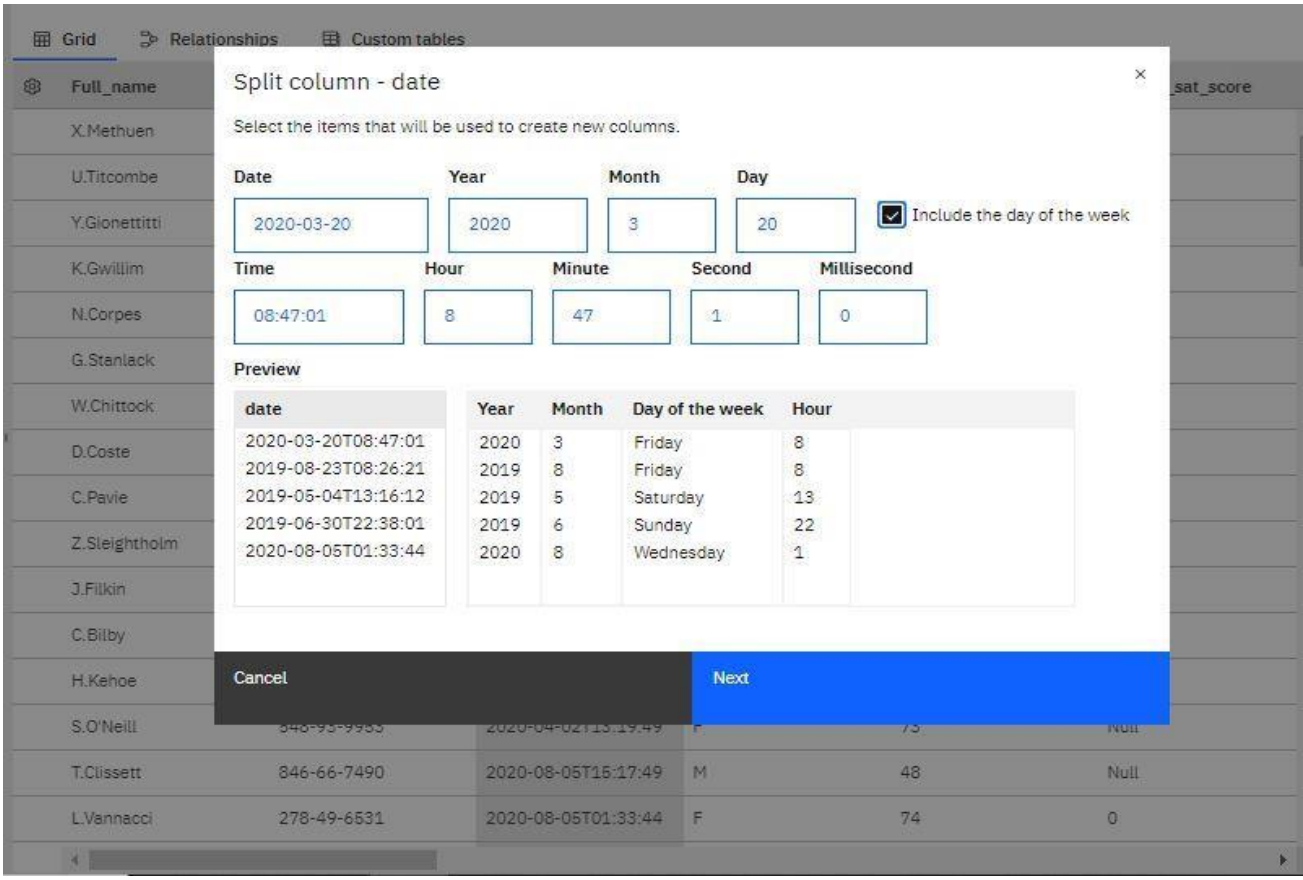


- AVERAGE SATISFACTION SCORE

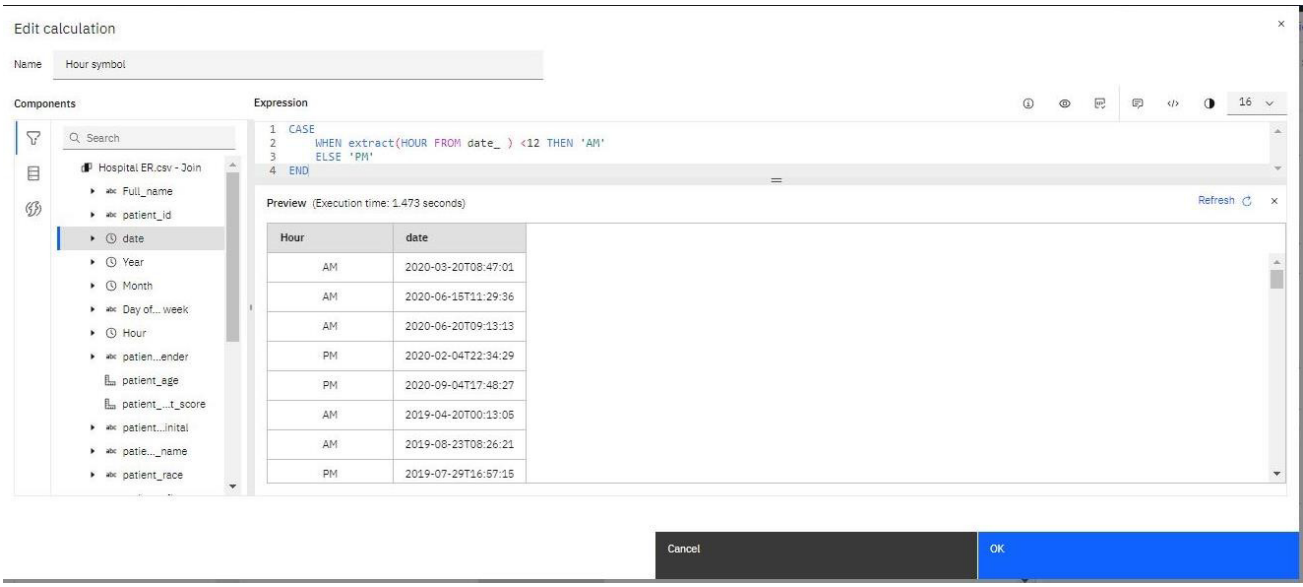
- CALCULATE AVG WAITING TIME: "CALCULATED AVG OF NO RATING



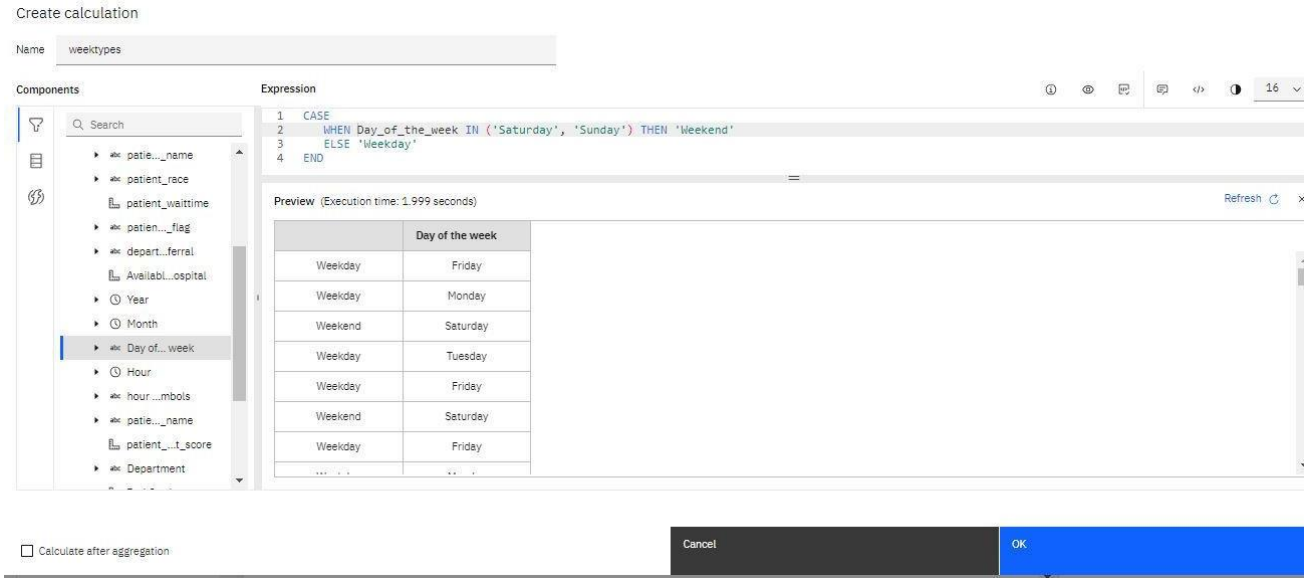
- SPLIT THE DATES: **INTO NEW COLUMNS LIKE YEAR, MONTH, WEEK OF THE DAY, HOUR**



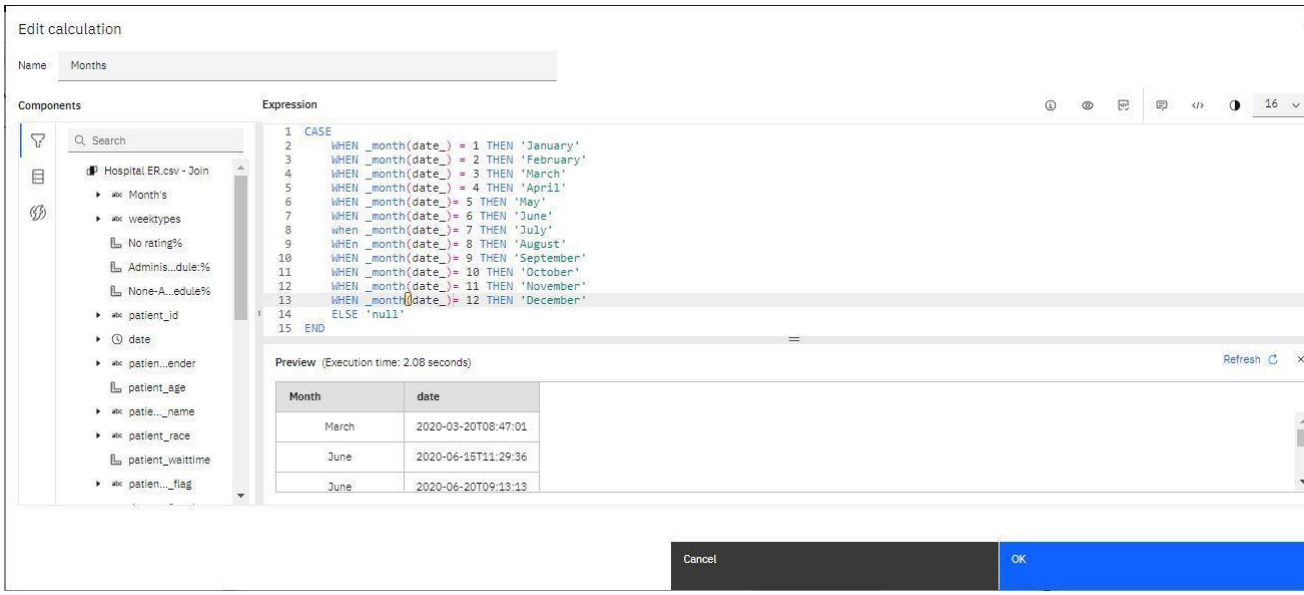
- CONVERT HOUR TIME INTO HOUR OF THE SYMBOL



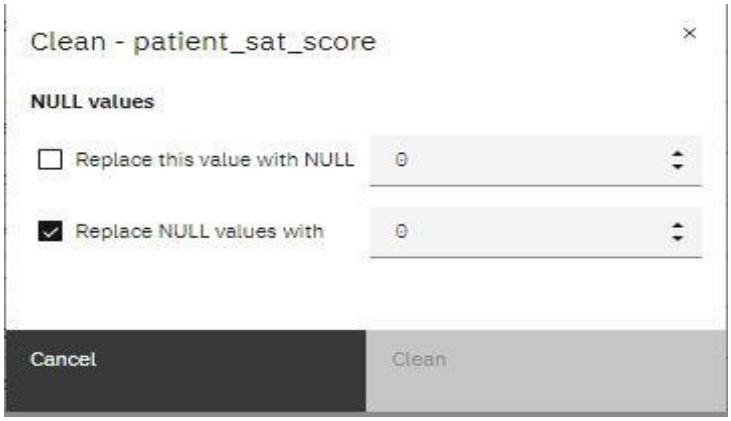
- CREATE WEEK TYPES FROM THE DAY OF THE WEEK COLUMN



- CONVENT DATE COLUMN INTO MONTHS COLUMN

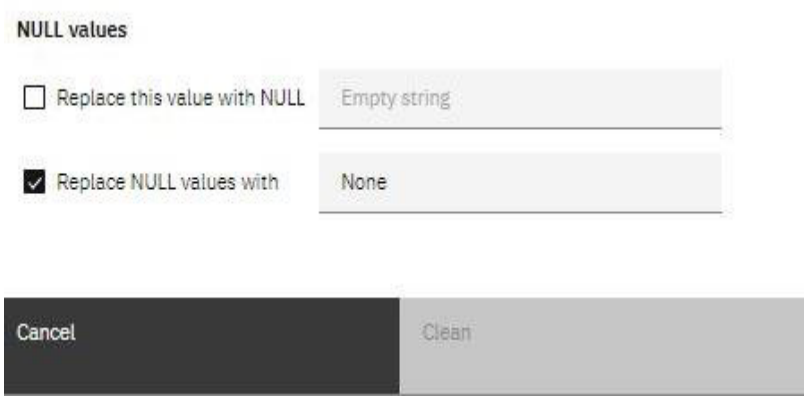


- TO CLEAN THE NULL VALUES IN SATISFACTION SCORE

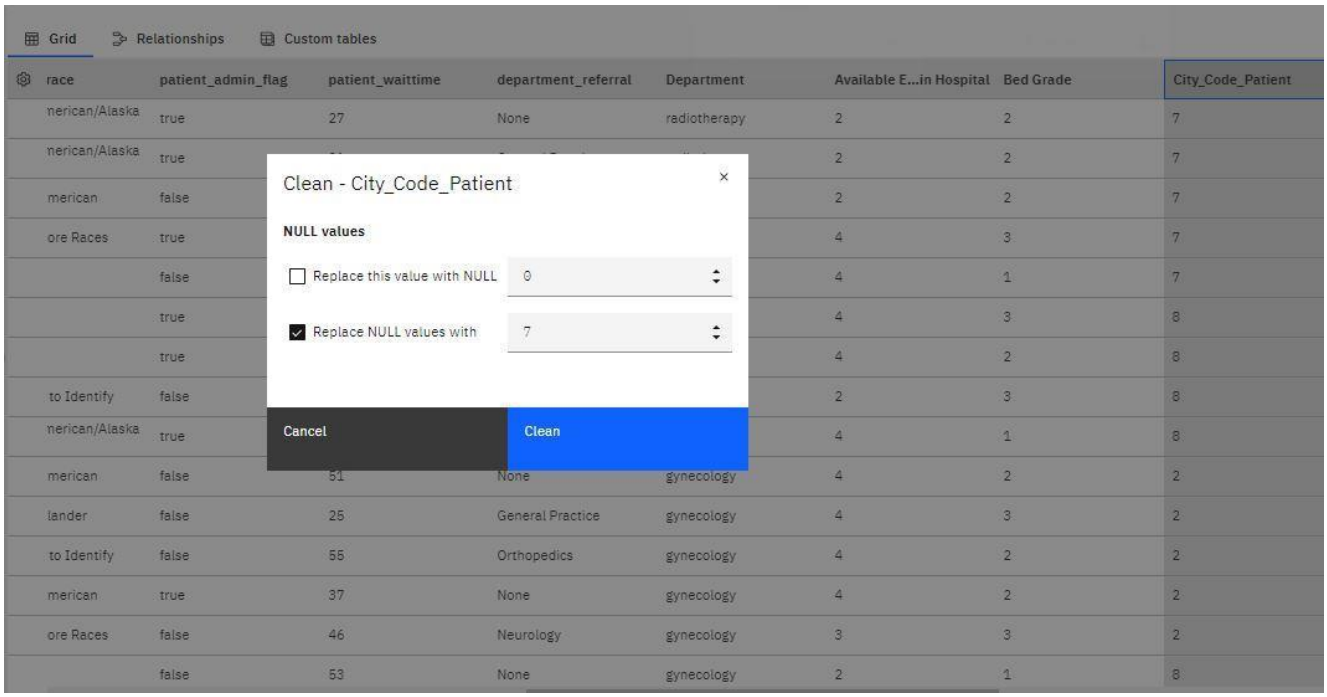


COLUMN

- TO REPLACE THE NULL VALUES WITH NONE IN THE DEPARTMENT REFERRAL



- TO REPLACE NULL VALUES WITH MEDIAN OF CITY CODE PATIENTS VALUES:



race	patient_admin_flag	patient_waittime	department_referral	Department	Available E...in Hospital	Bed Grade	City_Code_Patient
nerican/Alaska	true	27	None	radiotherapy	2	2	7
nerican/Alaska	true				2	2	7
merican	false				2	2	7
ore Races	true				4	3	7
	false				4	1	7
	true				4	3	8
	true				4	2	8
to Identify	false				2	3	8
nerican/Alaska	true				4	1	8
merican	false	51	None	gynecology	4	2	2
lander	false	25	General Practice	gynecology	4	3	2
to Identify	false	55	Orthopedics	gynecology	4	2	2
merican	true	37	None	gynecology	4	2	2
ore Races	false	46	Neurology	gynecology	3	3	2
	false	53	None	gynecology	2	1	8

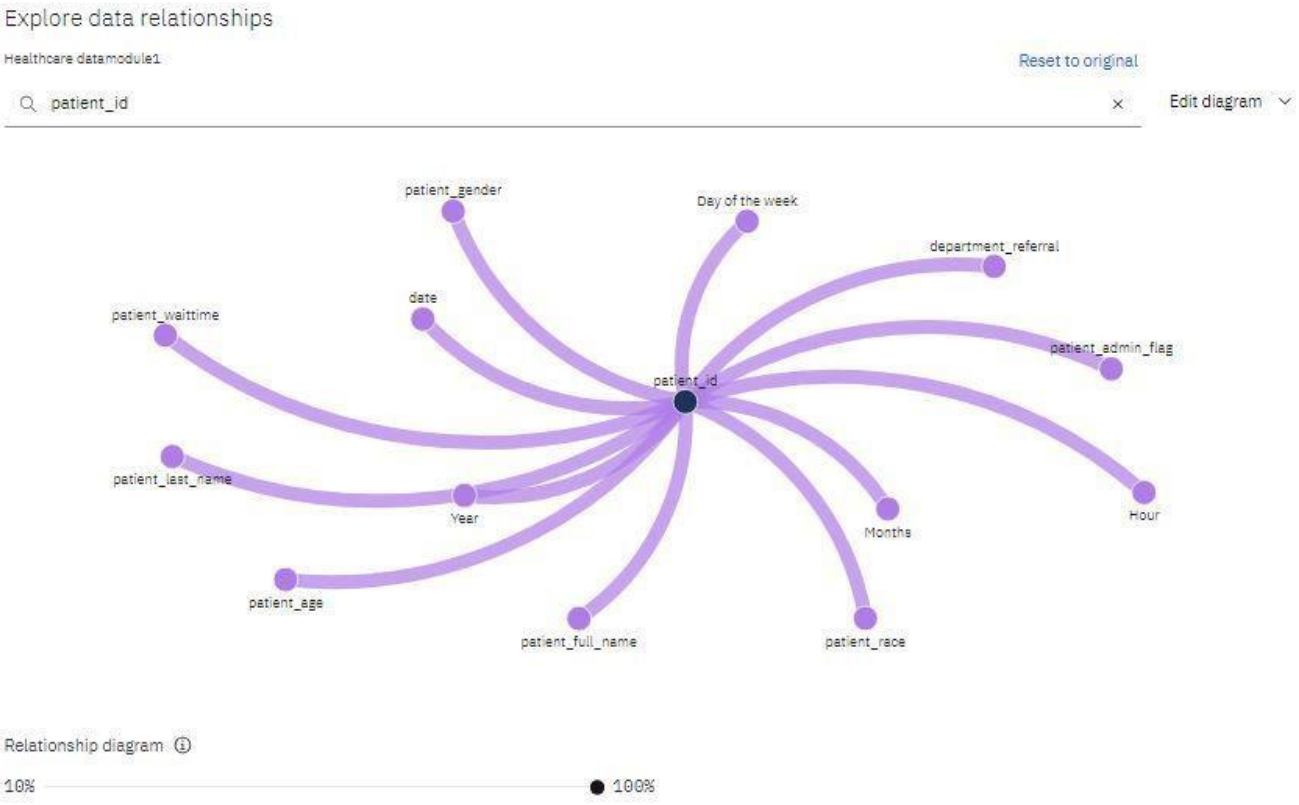
IN CONCLUSION :

In conclusion, effective data cleaning and preprocessing are essential steps in the development of analytics solutions for healthcare institutions. By following a systematic approach that includes data profiling, handling missing values, removing duplicates, standardization, outlier detection, transformation, feature engineering, data quality assurance, documentation, and version control, we can ensure that the healthcare data is cleansed, standardized, and prepared for analysis. This process enhances the reliability, accuracy, and integrity of the data, ultimately leading to more meaningful insights and informed decision-making in healthcare settings. Moving forward, continued emphasis on data quality and robust preprocessing techniques will be crucial for maximizing the value of healthcare data and driving improvements in patient care, operational efficiency, and overall healthcare outcomes.

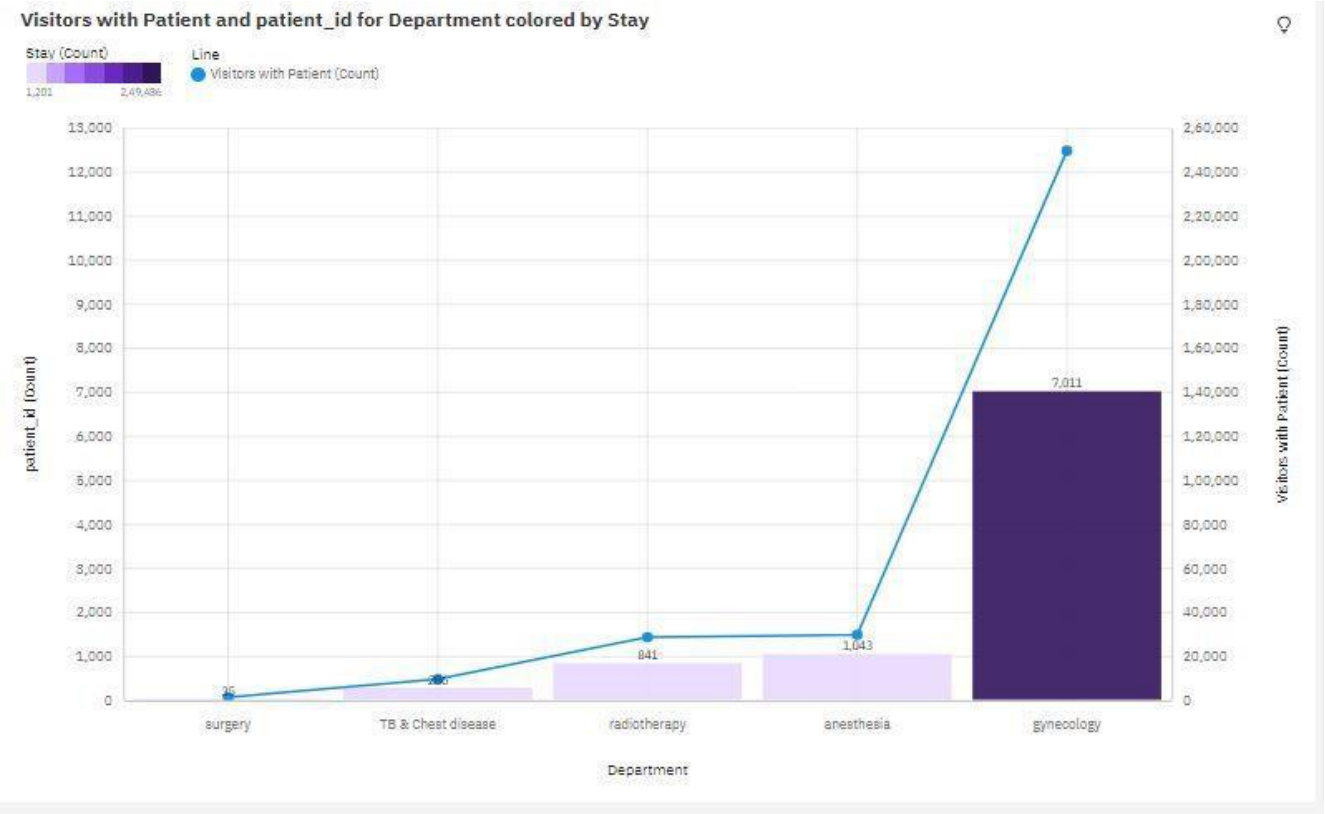
STEP 3

DATA EXPLORATION USING IBM COGNOS ANALYTICS

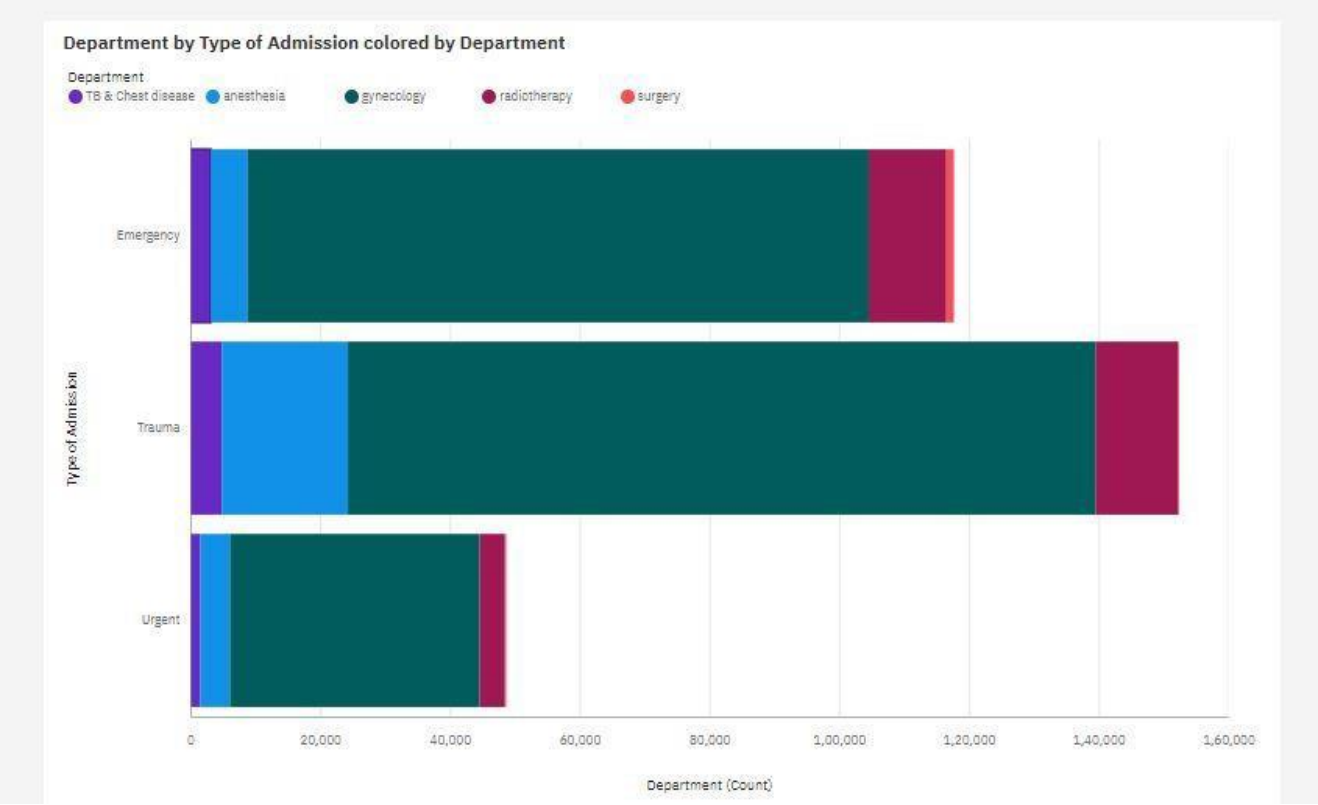
- TO EXPLORE THE DATA RELATIONSHIP FROM THE TRAINED HEALTHCARE DATASET



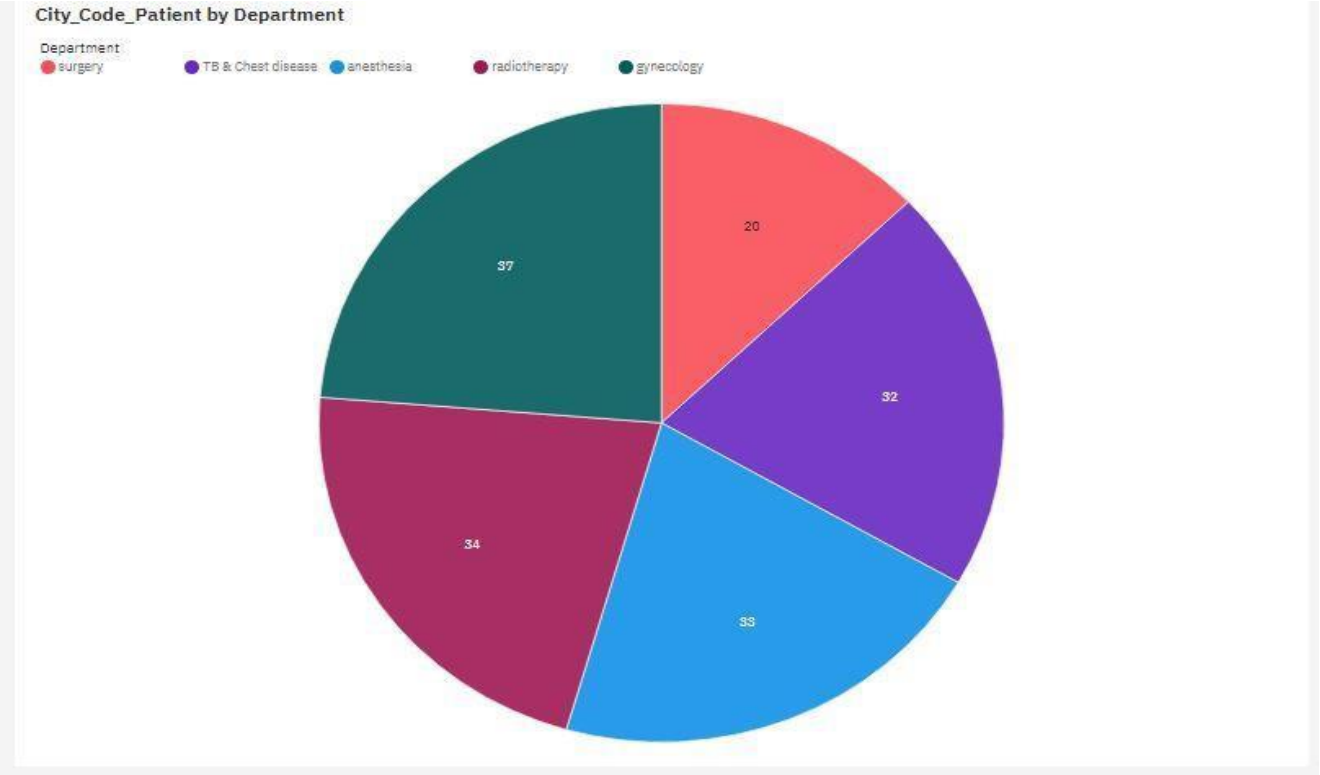
1. TO CREATE VISUALISATION



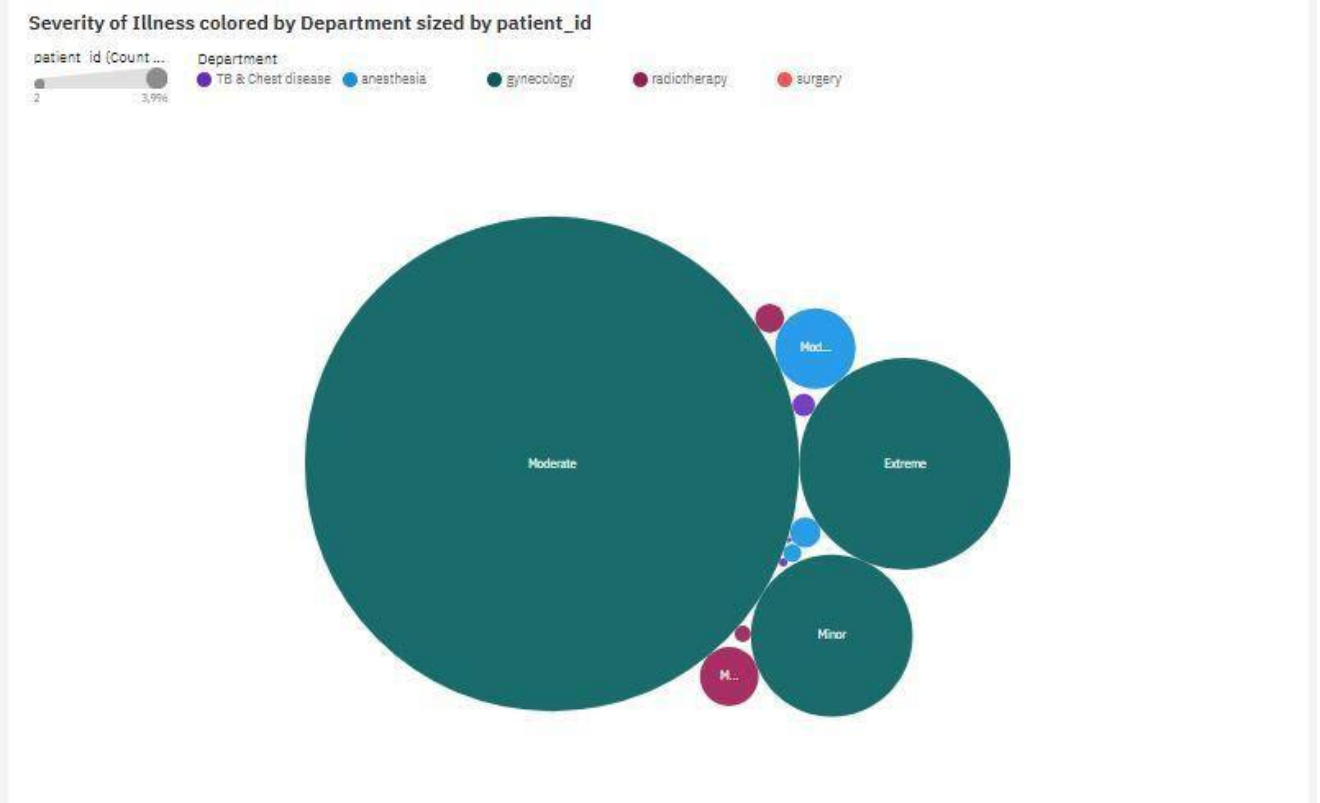
2. TO CREATE VISUALISATION



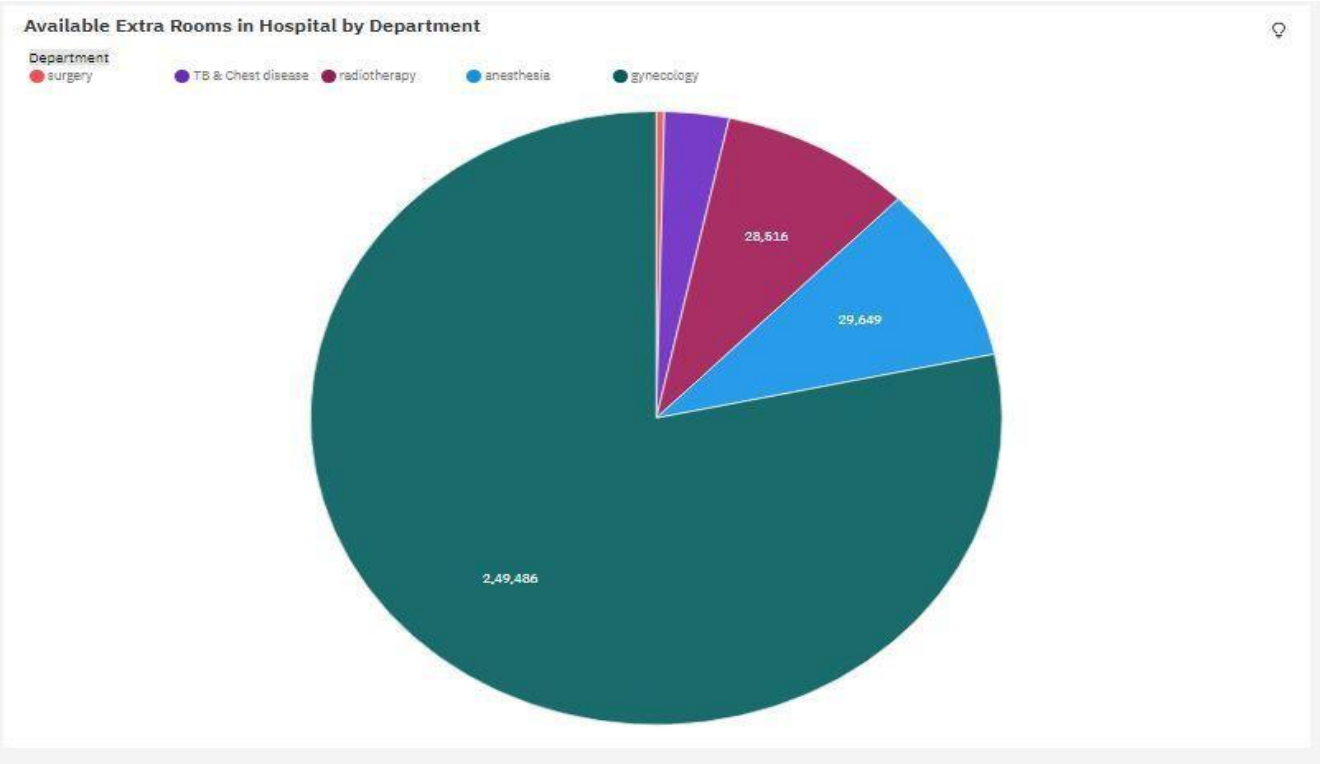
3. TO CREATE VISUALISATION



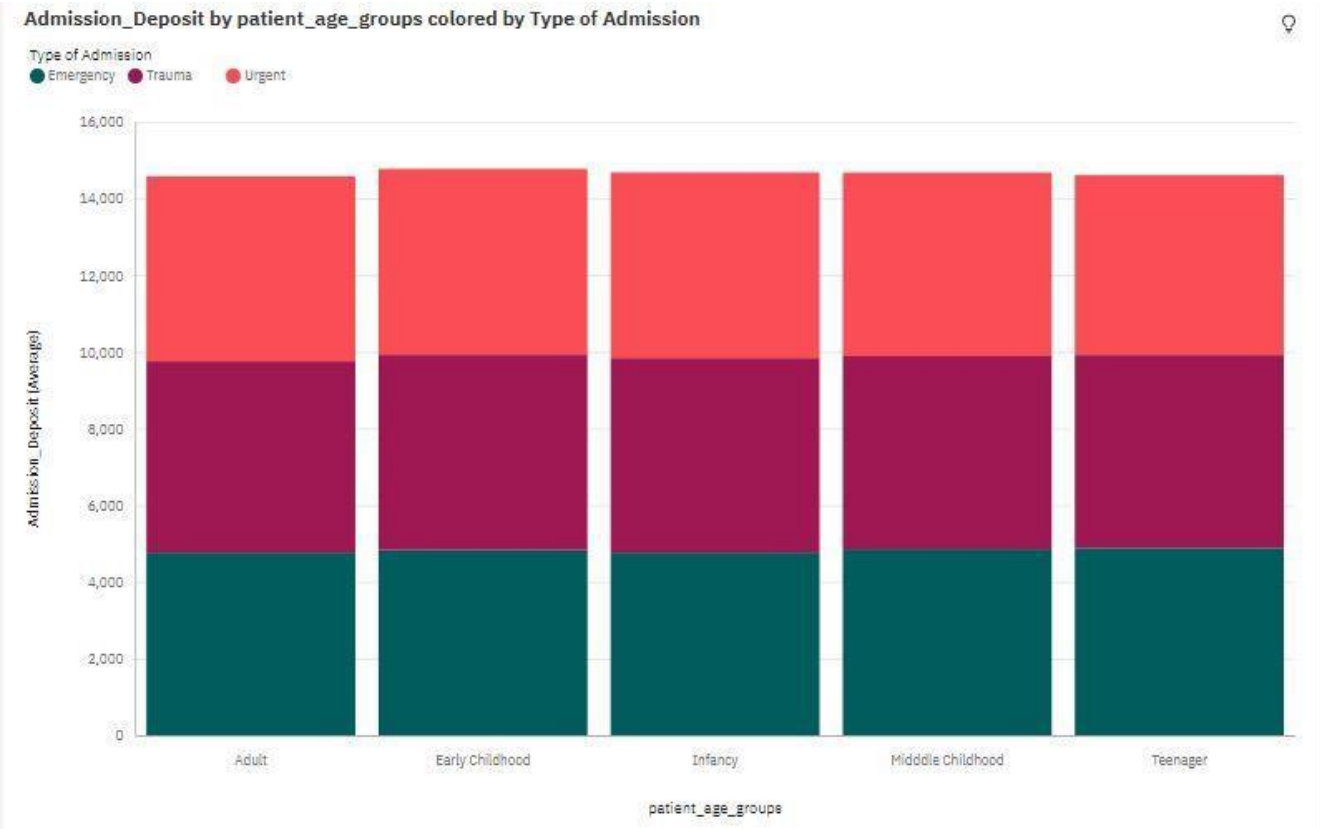
4. TO CREATE VISUALISATION



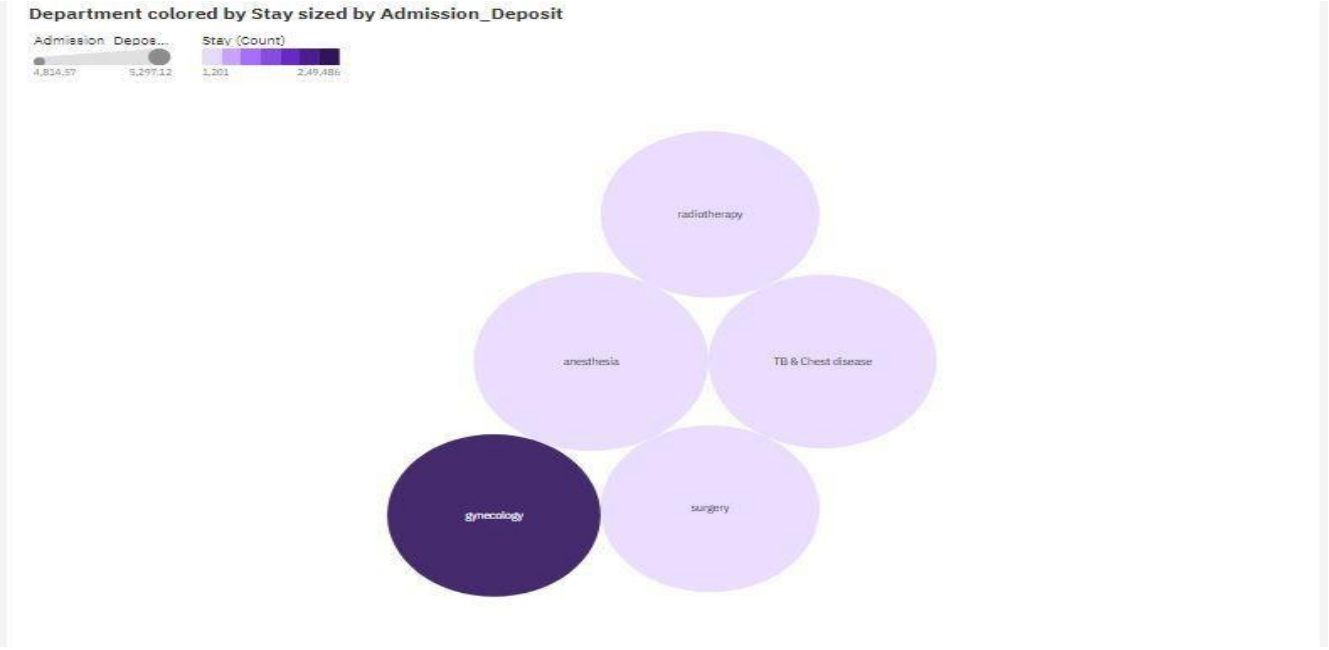
5. TO CREATE VISUALISATION



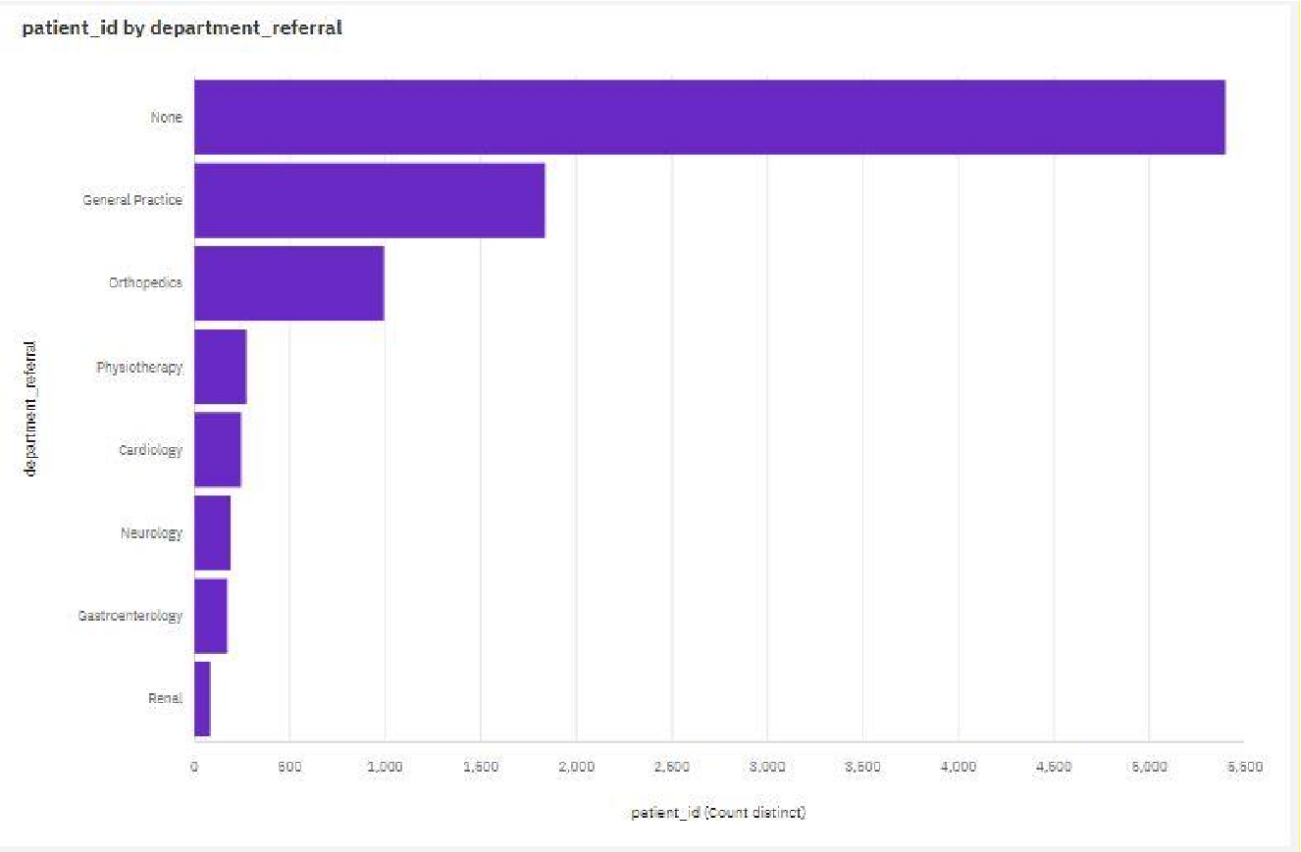
6. TO CREATE VISUALISATION



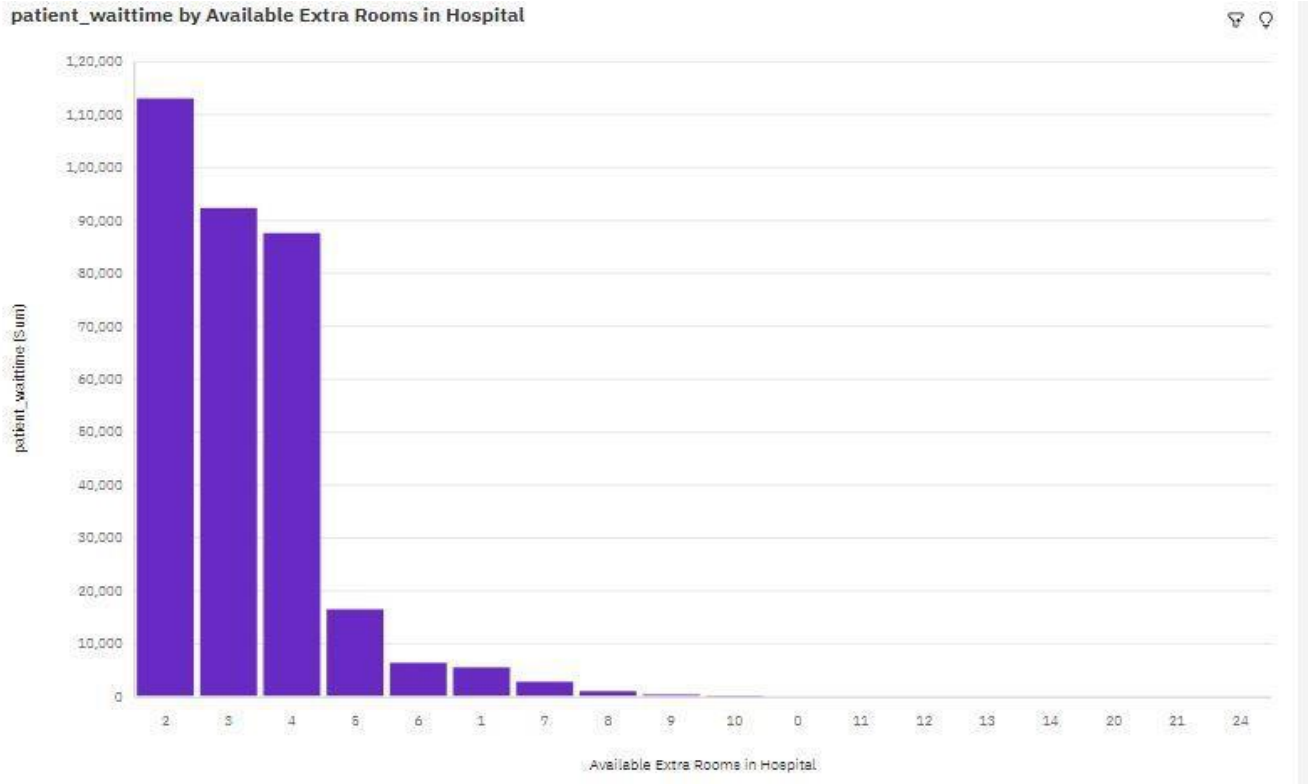
7. TO CREATE VISUALISATION



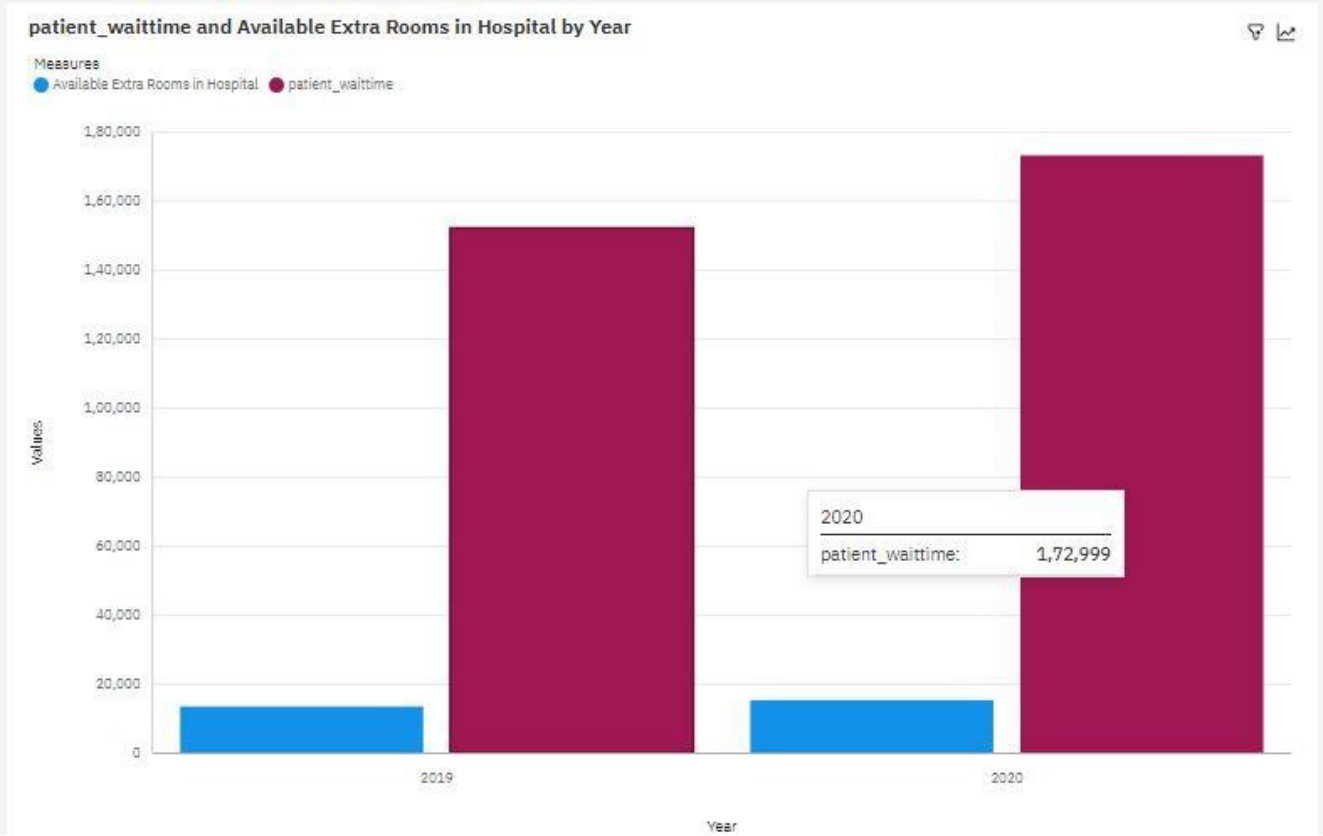
8. TO CREATE VISUALISATION



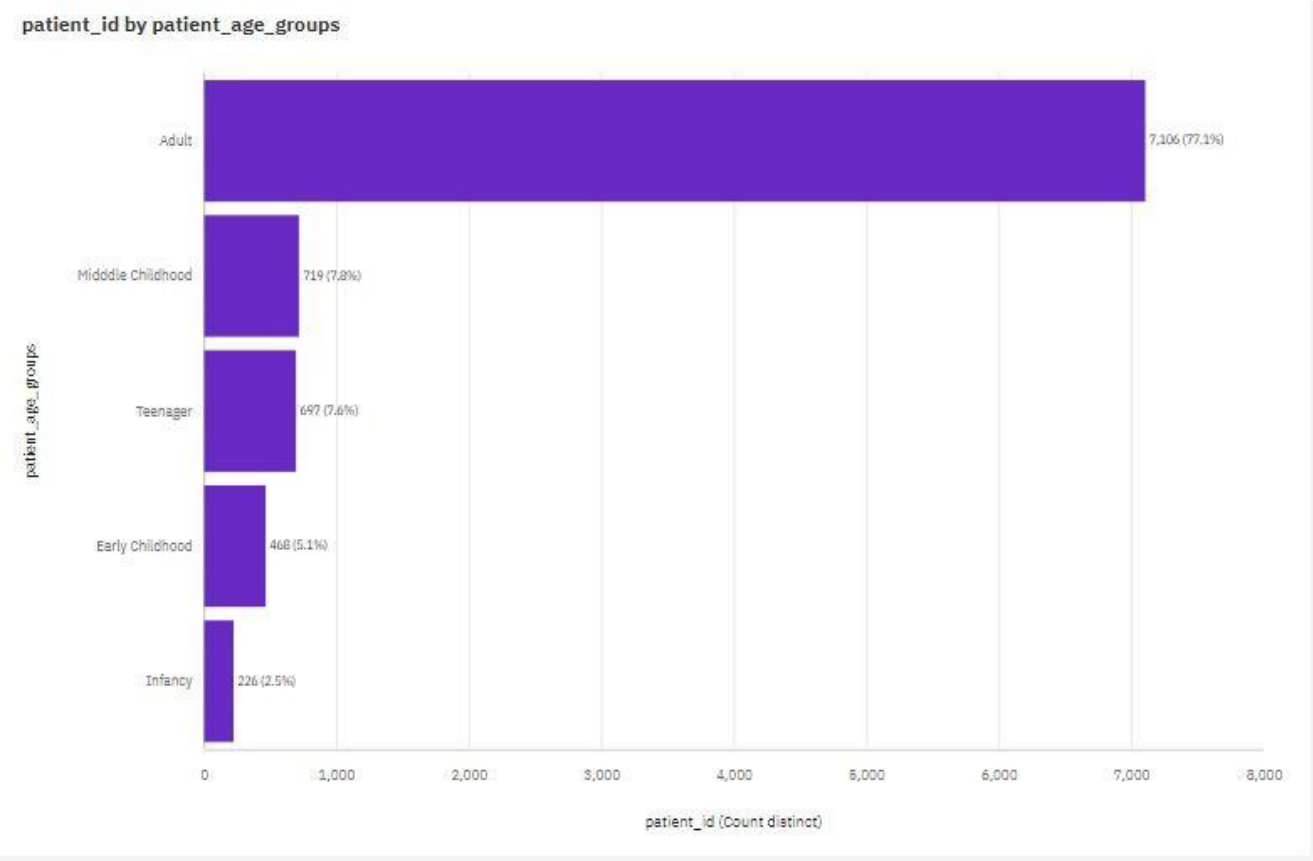
9. TO CREATE VISUALISATION



10. TO CREATE VISUALISATION



11. TO CREATE VISUALISATION



CREATE THE DASHBOARD

Title: Patient Demographics

This dashboard provides a high-level overview of patient demographics at a healthcare facility. The data is broken down into several categories, including:

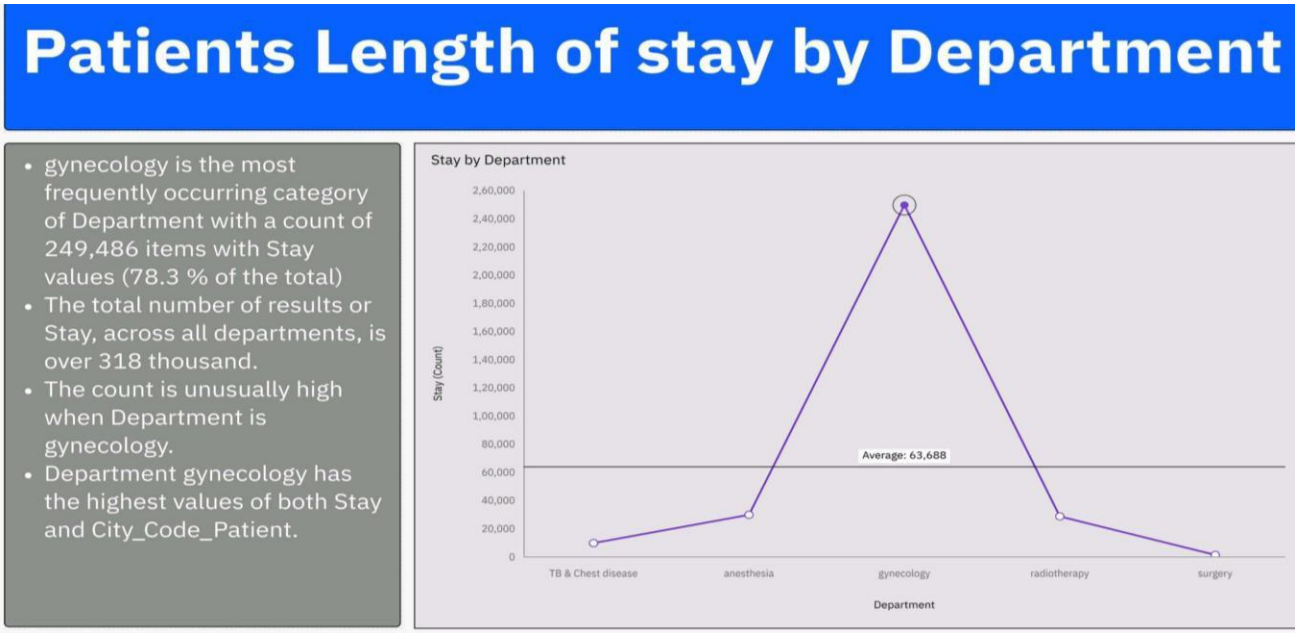
- **Total Patients by Department:** This bar chart shows the distribution of patients across different departments within the facility. Orthopedics appears to be the most popular department, followed by anesthesia.
- **Patient ID by Department:** This table likely links patient IDs to the departments they Visited.
- **Day of the Week:** This section may reveal trends in patient volume by day of the week.
- **Hour of Day:** This section may reveal trends in patient volume by hour of the day.
- **Total Patients by Months:** This bar chart shows the distribution of patients by month. It appears that the facility sees more patients in the summer months.
- **Patient ID by Patient Gender:** This table likely links patient IDs to their gender.
- **Patient ID by Patient Age Groups:** This table likely links patient IDs to their age groups.
- **Average Waiting Time:** This section may show average wait times for patients.
- **Average Satisfaction Score:** This section may show average satisfaction scores for patients.
- **Patient ID by Patient Race**: This table likely links patient IDs to their race.

- **Patient Count by Appointment Type:** This section may show a breakdown of patients by appointment type, such as administrative or non-administrative.

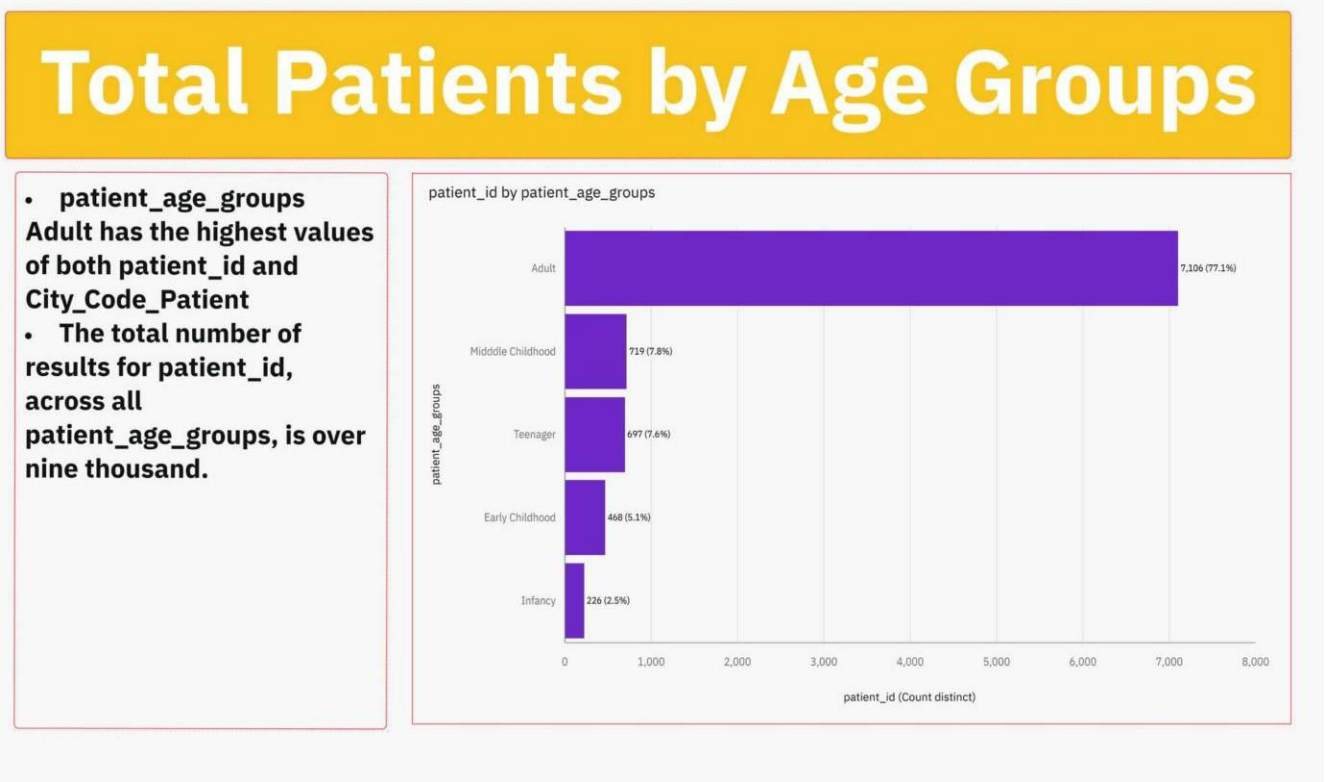
Display Dashboard in IBM Cognos

To create a

story Story 1:



Story 2:



Story 3:

Patient_waittime for Age Group And Patient_race

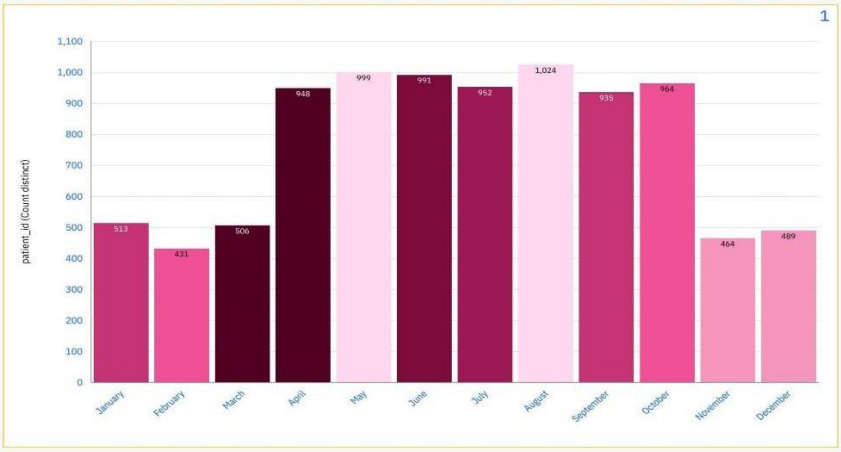
patient_waittime	Adult	Early Childhood	Infancy	Midddle Childhood	Teenager	Summary
African American	52,644	3,536	1,750	6,293	5,249	**69,472**
Asian	29,352	1,737	685	2,639	2,973	**37,386**
Declined to Identify	27,478	1,965	1,054	2,756	2,739	**35,992**
Native American/Ala...	13,694	987	328	1,404	1,362	**17,775**
Pacific Islander	14,929	1,084	325	1,276	1,408	**19,022**
Two or More Races	42,714	2,983	1,423	4,090	3,842	**55,052**
White	69,622	4,137	1,912	7,519	7,066	**90,256**
Summary	**2,50,433**	**16,429**	**7,477**	**25,977**	**24,639**	**3,24,955**

- White patients appear to have the longest average wait time (69,632), followed by Two or More Races (42,714), Asian (29,352), African American (52,644), Pacific Islander (14,929), and Native American (unrecorded). It is important to note that correlation does not equal causation, and there could be other factors influencing wait times.
- Patients age category have the longest wait time (35,992), followed by teenagers (17,775), infancy (6,293), early childhood (5,249), and middle childhood (6,947).

Story 4:

Total patient visits by Month

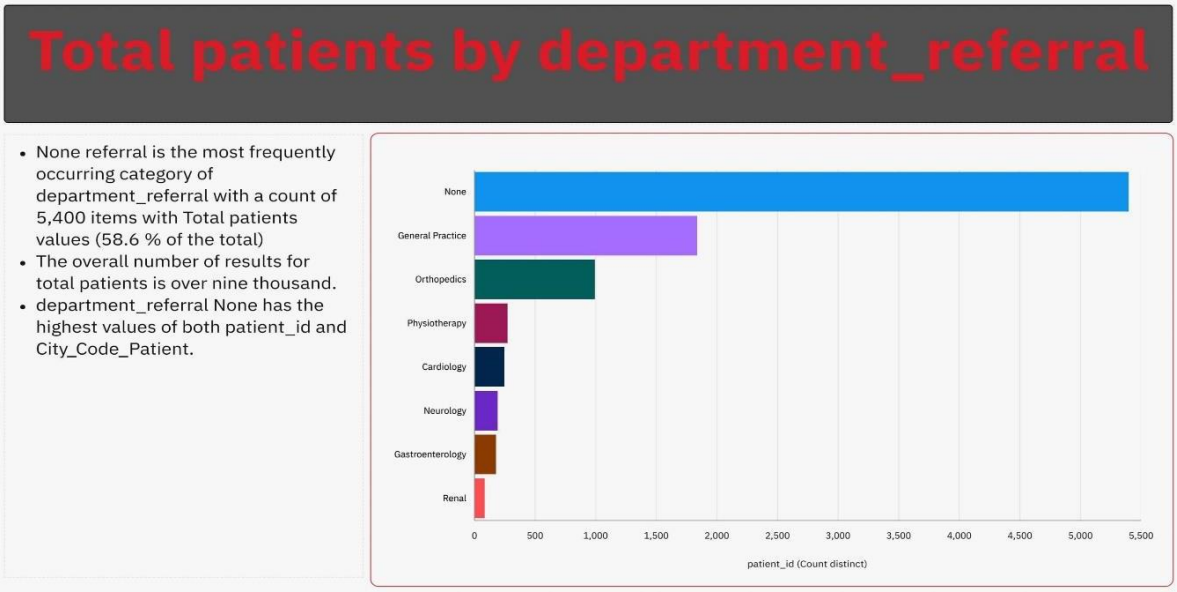
- August (11.1 %), May (10.8 %), June (10.8 %), October (10.5 %), and July (10.3 %) are the most frequently occurring categories of Months with a combined count of 4,930 items with patient_id values (53.5 % of the total)
- Based on the current forecasting, patient_id may reach -280.1 by Months March+1.
- The total number of results for patient_id, across all months, is over nine thousand.



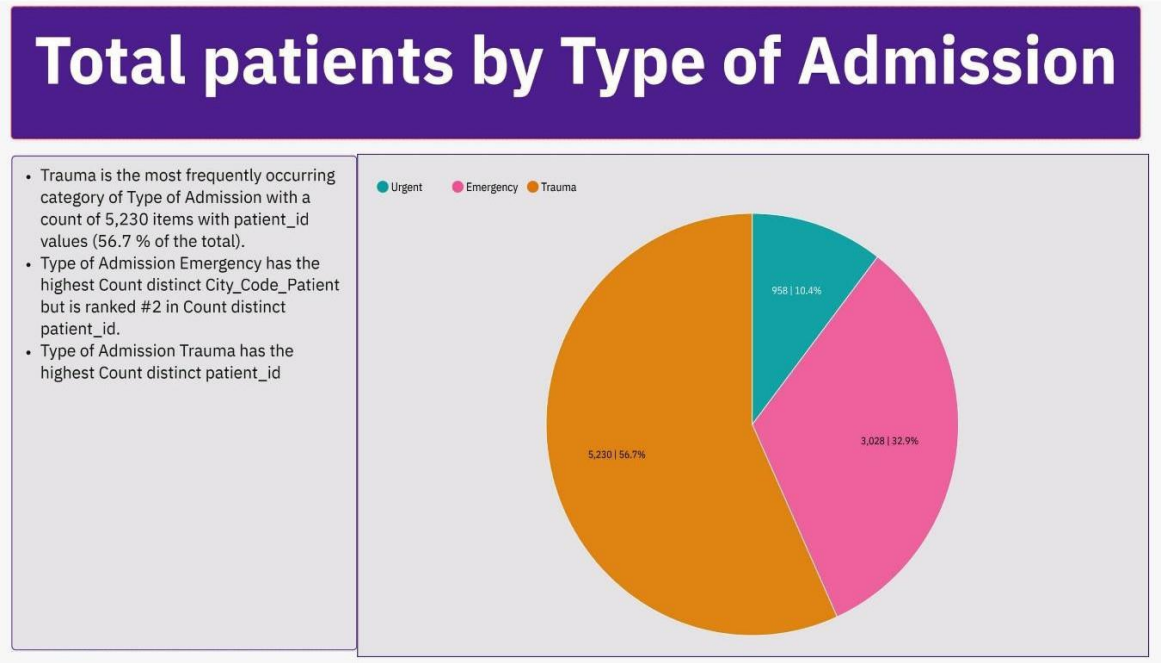
Story 5:

Filter(s) applied to the visualization(s) on the previous page:

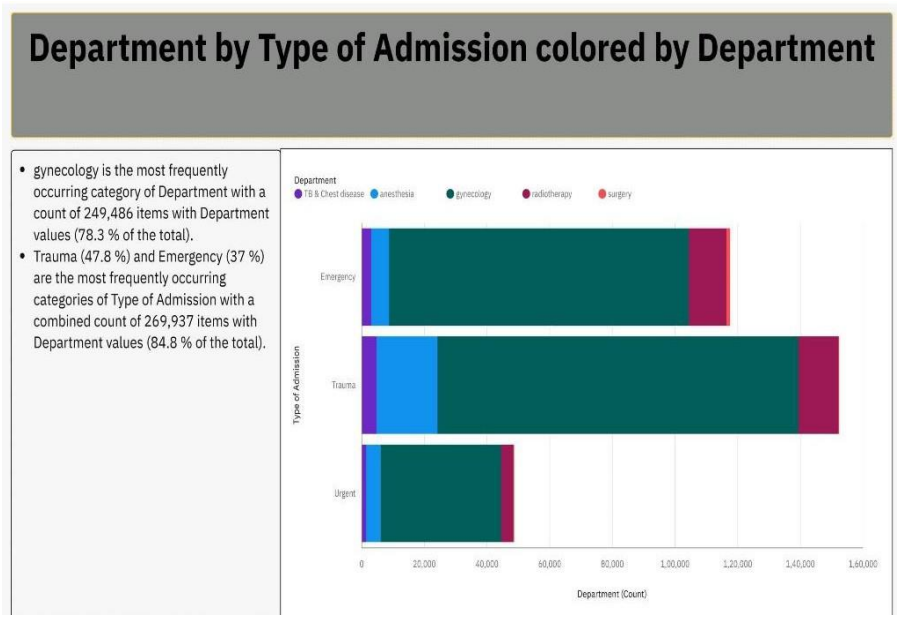
Widget 1
Months Excludes: null



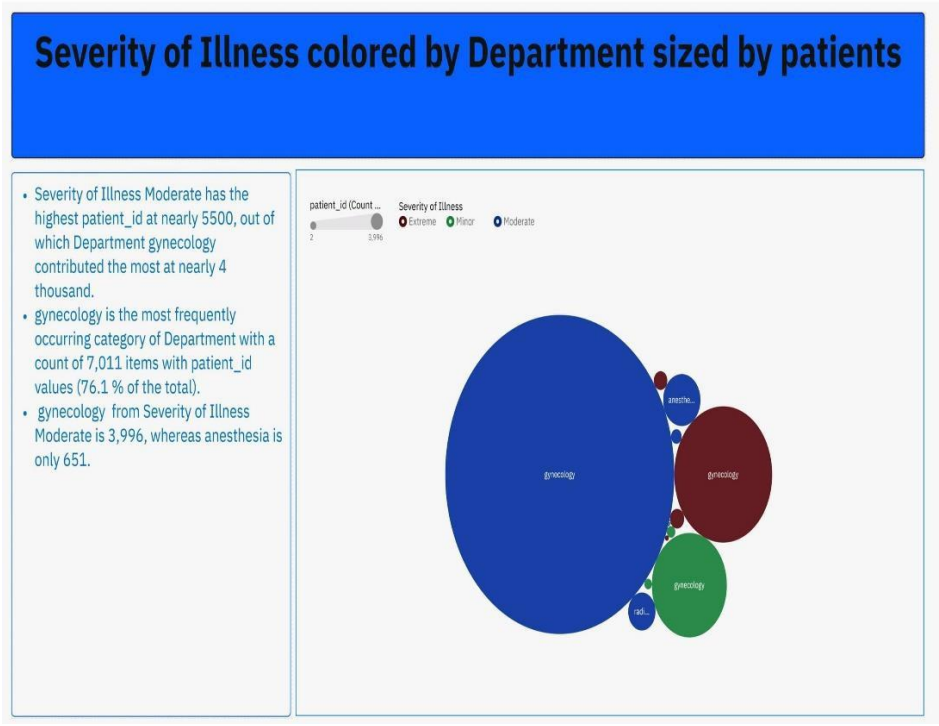
Story 6:



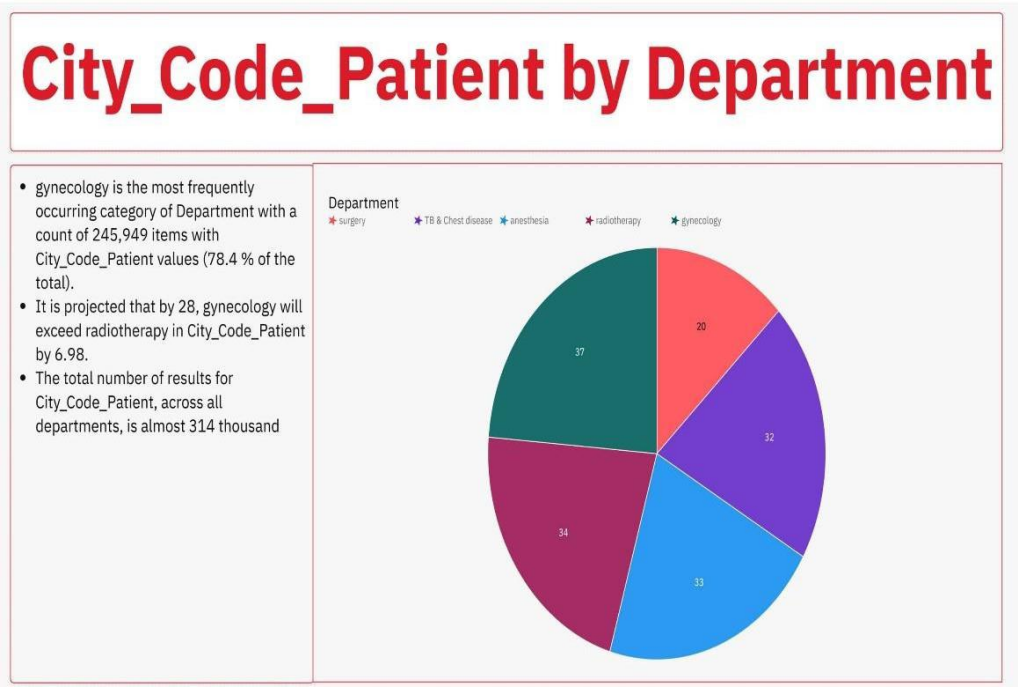
Story 7:



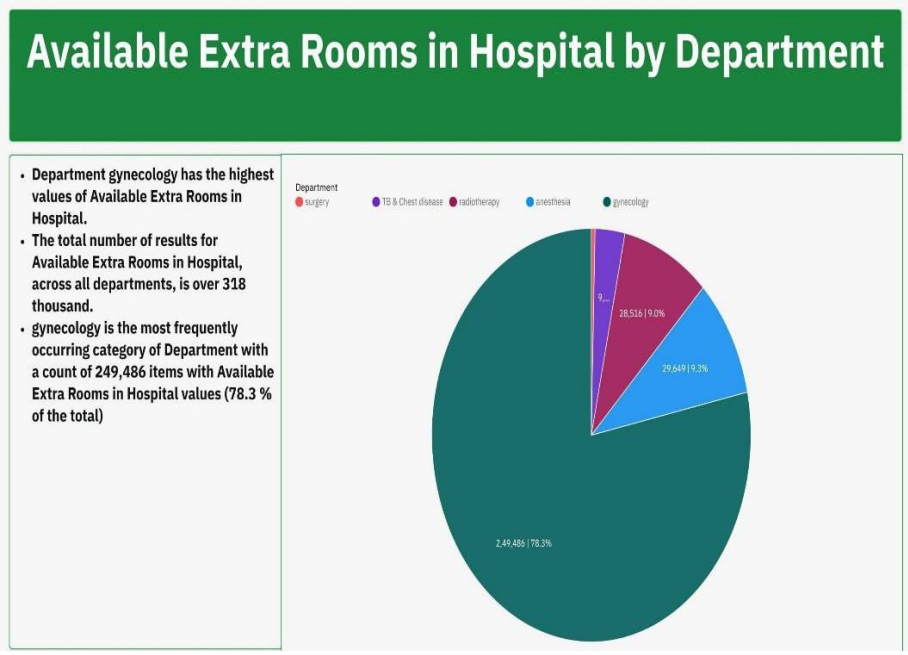
Story 8:



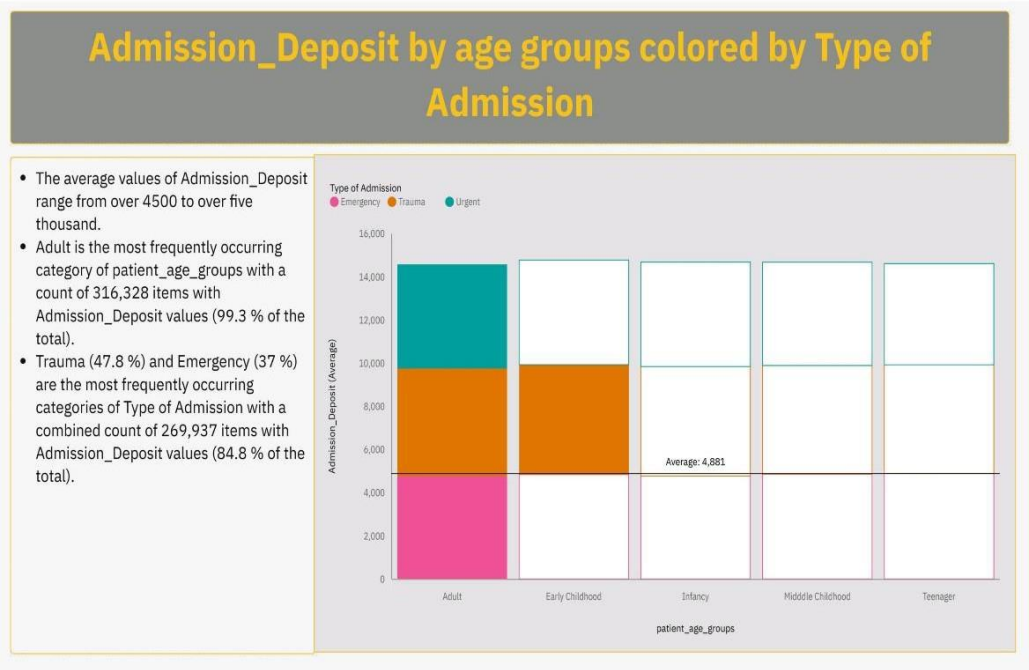
Story 9:



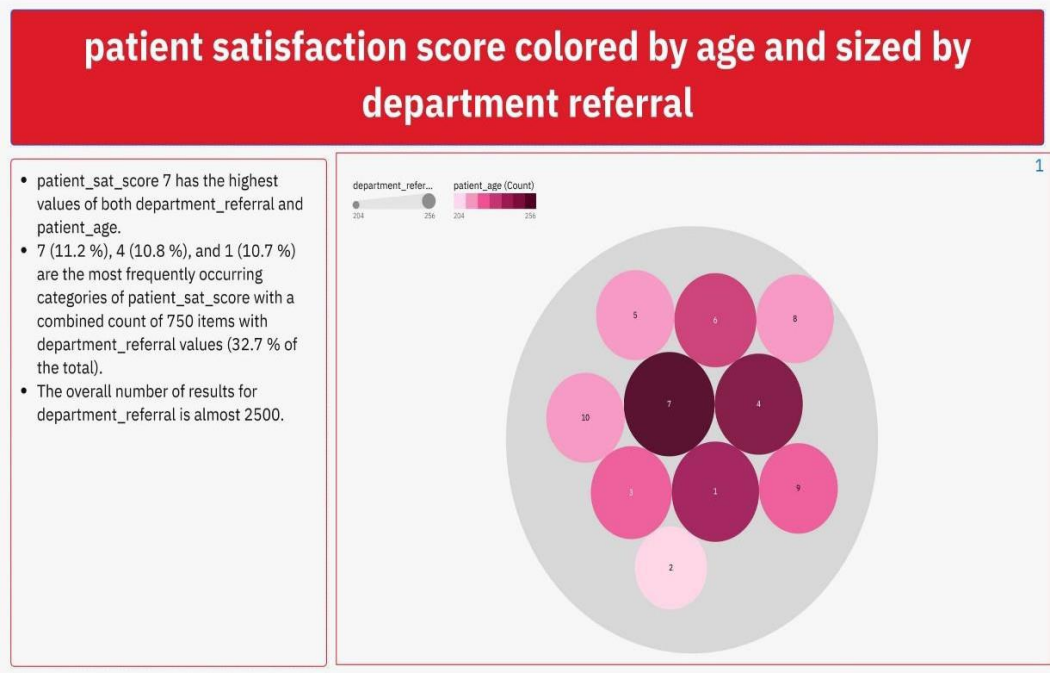
Story 10:



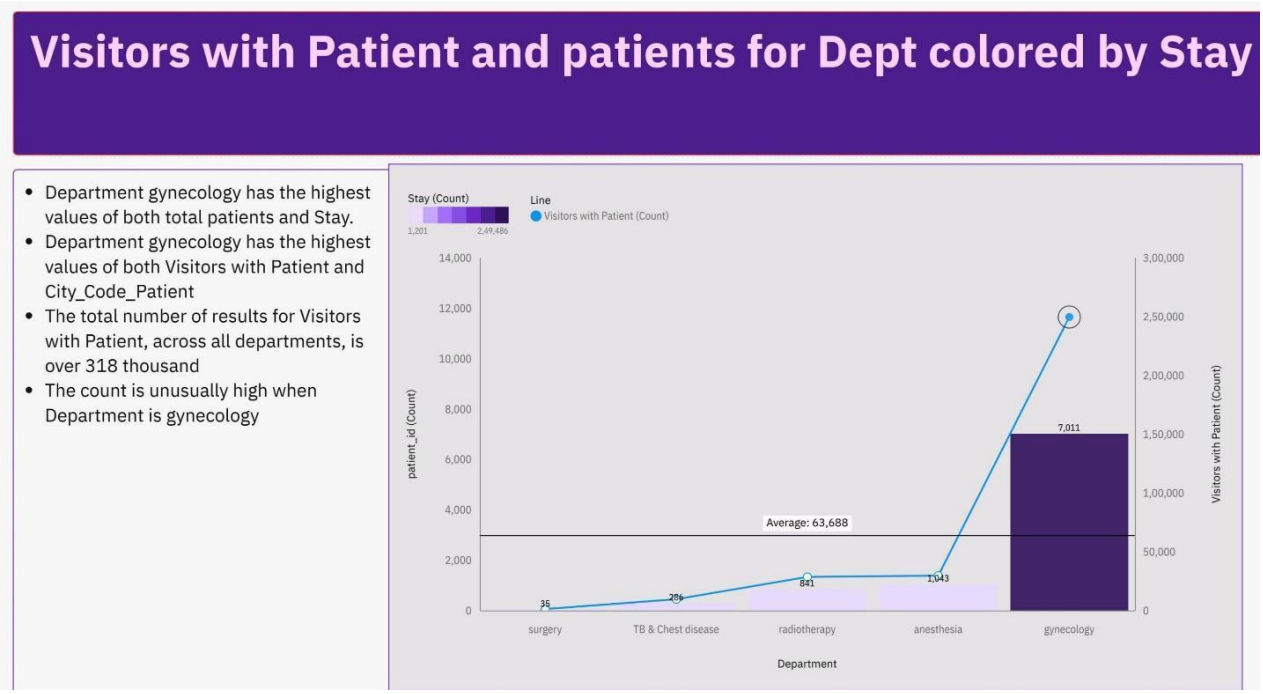
Story 11:



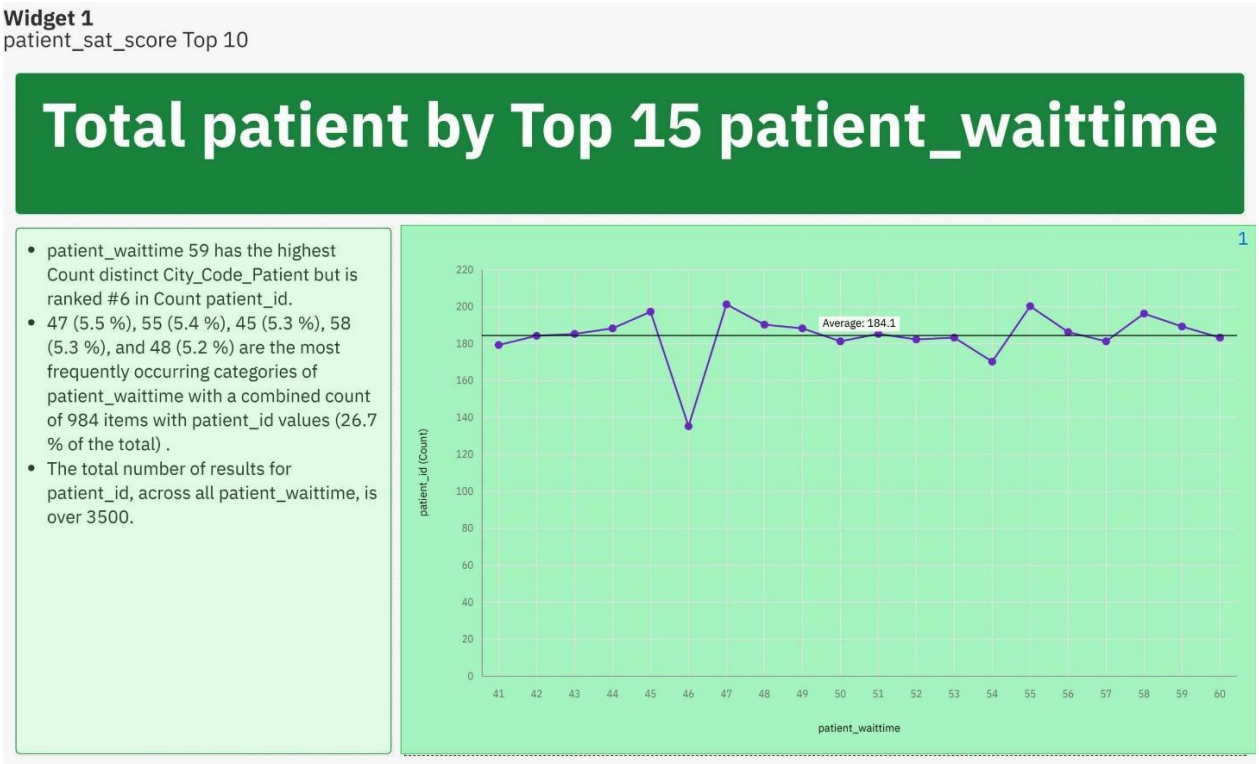
Story 12:



Story 13:



Story 14:



12.2 SOURCE CODE:

HTML:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width, initial-scale=1">
    <title>Big data Analysis in Healhcare</title>
        <link rel="stylesheet" href="https://unpkg.com/aos@next/dist/aos.css"
/>
    <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/css/bootstrap.min.css">
    <script
src="https://ajax.googleapis.com/ajax/libs/jquery/3.3.1/jquery.min.js"></scrip
t>
    <script
src="https://cdnjs.cloudflare.com/ajax/libs/popper.js/1.12.9/umd/popper.min.js
"></script>
    <script
src="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/js/bootstrap.min.js"></sc
ript>
    <script
src="https://use.fontawesome.com/releases/v5.0.8/js/all.js"></script>
    <link rel="stylesheet"
href="https://cdnjs.cloudflare.com/ajax/libs/animate.css/4.1.1/animate.min.css
">
    <link href="style.css" rel="stylesheet">
</head>
<body>

<!-- Navigation -->
<nav class="navbar navbar-expand-md navbar-light bg-light sticky-top">
    <div class="container-fluid">
        <a class="navbar-brand" href="#"><img src="img/apolloLogo.web"></a>
        <button class="navbar-toggler" type="button" data-toggle="collapse"
data-target="#navbarResponsive">
            <span class="navbar-toggler-icon"></span>
        </button>
        <div class="collapse navbar-collapse" id="navbarResponsive">
            <ul class="navbar-nav ml-auto">
                <li class="nav-item active">
                    <a class="nav-link" href="#">Home</a>
                </li>
                <li class="nav-item">
```

```
                        <a class="nav-link" href="#section1">About</a>
                    </li>
                    <li class="nav-item">
                        <a class="nav-link" href="dash.html">Dashboard</a>
                    </li>
                    <li class="nav-item">
                        <a class="nav-link" href="story.html">Story</a>
                    </li>
                    <li class="nav-item">
                        <a class="nav-link" href="#section2">Team</a>
                    </li>
                    <li class="nav-item">
                        <a class="nav-link" href="#section3">Connect</a>
                    </li>
                </ul>
            </div>
        </div>
</nav>

<!--- Image Slider -->
<div id="slides" class="carousel slide" data-ride="carousel" data-aos="fade- up"
            data-aos-offset="200" data-
            aos-delay="50"
            data-aos-duration="1000">
        <ul class="carousel-indicators">
            <li data-target="#slides" data-slide-to="0" class="active"></li>
            <li data-target="#slides" data-slide-to="1"></li>
            <li data-target="#slides" data-slide-to="2"></li>
        </ul>
        <div class="carousel-inner">
            <div class="carousel-item active">
                <img class="image" src="img/BDA HC11.webp">
                <div class="carousel-caption">
                    <h1 class="display-2">Big data Analysis in Healhcare</h1>
                    <!--h3>Big data Analysis in Healhcare</h3-->
                    <!--button type="button" class="btn btn-outline-light btn-
lg">VIEW
                                                                DEMO</button>
                                                                <button type="button" class="btn btn-primary btn-
                                                                lg">Get
        Started</button-->
                </div>
            </div>
            <div class="carousel-item">
                <img class="image" src="img/BDA HC3.webp">
            </div>
            <div class="carousel-item">
                <img class="image" src="img/BDA HC12.webp">
```

```
            </div>
        </div>
</div>

<!--- Jumbotron -->
<div class="container-fluid" data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
        <div class="row jumbotron">
            <div class="col-xs-12 col-sm-12 col-md-9 col-lg-9 col-xl-10">
                <p class="lead" id="section1">IBM Cognos Analytics is a web-based BI platform that
uses AI to automate data prep, exploration and prediction. It offers flexible deployment options, interactive
dashboard</p>
            </div>
            <div class="col-xs-12 col-sm-12 col-md-3 col-lg-3 col-xl-2">
                <a
href="https://www.bing.com/ck/a?!&&p=141eaf731adf7f08JmltdHM9MTcxMzc0NDAwMCZpZ
3VpZD0yMzVmYzYwMC1iMjNjLTYyYTUtMWU4NC1kNzFkYjM5MTYzOTcmaW5zaWQ9NTIyOA
&ptn=3&ve r=2&hsh=3&fclid=235fc600-b23c-62a5-1e84-
d71db3916397&psq=ibm+cognos+analytics&u=a1aHR0cHM6Ly93d3cuaWJtLmNvbS9wcm9kdWN0
cy9jb2dub3MtYW5hbHl0aWNz&ntb=1"><button type="button" class="btn btn-outline- secondary
                    btn-lg">Official Website</button></a>
            </div>
        </div>
</div>

<!--- Welcome Section -->
<div class="container-fluid padding"data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
        <div class="row welcome text-center">
            <div class="col-12">
                <h1 class="display-4" >Data Analytics</h1>
            </div>
            <hr>
            <div class="col-12">
                <p class="lead">Welcome to my website! This is a simple website created for
visualizing Big data analysis in healthcare .</p>
                <p class="lead">We provide you with the most relevant and effective data analysis by
integrating, aggregating and analyzing different types of data from multiple sources. This technique
saves time and gives you higher returns on your investment..</p>
                <p class="lead">This project aims to analyze healthcare data to uncover valuable
insights. We will follow a structured approach divided into four key modules</p><hr>
```

```
            </div>
        </div>
</div>

<div class="container-fluid padding"data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
          <div class="row padding">
              <div class="col">
                        <h1 > MODULE 1 : Data Collection & Understanding</h1>
            </div>
            <div class="col-md-12 col-lg-6">
                  <h2>Data Gathering:</h2>
                  <p> We will collect relevant healthcare data from various sources (specify sources).</p>
                  <h2>Understanding the Data:</h2>
                  <p> We will carefully examine the data structure, identify key variables, and assess
data quality.</p>
                        <br>

            </div>

        </div>
</div>


<!--- Three Column Section ->
<div class="container-fluid padding"data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000" >
        <div class="row text-center padding">
            <div class="col-xs-12 col-sm-6 col-md-4">
                  <i class="fas fa-code"></i>
                  <h3>HTML5</h3>
                  <p>Built with the latest version of HTML, HTML5.</p>
            </div>
            <div class="col-xs-12 col-sm-6 col-md-4">
                  <i class="fas fa-bold"></i>
                  <h3>BOOTSTRAP</h3>
                  <p>Built with the latest version of Bootstrap, Bootstrap 4.</p>
            </div>
            <div class="col-xs-12 col-md-4">
                  <i class="fab fa-css3"></i>
                  <h3>CSS3</h3>
                  <p>Built with the latest version of CSS, CSS3.</p>
            </div>
```

```
        </div>
        <hr class="my-4">
</div-->

<!--- Two Column Section -->
<div class="container-fluid padding"data-aos="fade-up" data-aos-
                offset="200"
                data-aos-delay="50"
                data-aos-duration="1000">
        <div class="row padding">
                <div class="col-md-12 col-lg-6">
                        <h2>IBM Cognos Analytics ...</h2>
                        <p>IBM Cognos Analytics is a web-based integrated business intelligence suite by
IBM. It provides a toolset for reporting, analytics, scorecarding, and monitoring of events and
metrics.</p>
                        <p>Self-Service Analytics: With just a few clicks, you can unlock the power of self-
service analytics. Create powerful visualizations, tell the story of your data, and share insights via email,
Slack, or the mobile app.</p>
                        <p>Predictive Forecasting: Cognos Analytics offers predictive forecasting capabilities.
You can analyze historical data and predict future trends, enabling better decision-making.</p>
                                <br>
                                <a
href="https://www.bing.com/ck/a?!&&p=141eaf731adf7f08JmltdHM9MTcxMzc0NDAwMCZpZ
3VpZD0yMzVmYzYwMC1iMjNjLTYyYTUtMWU4NC1kNzFkYjM5MTYzOTcmaW5zaWQ9NTIyOA
&ptn=3&ve r=2&hsh=3&fclid=235fc600-b23c-62a5-1e84-
d71db3916397&psq=ibm+cognos+analytics&u=a1aHR0cHM6Ly93d3cuaWJtLmNvbS9wcm9kdWN0
cy9jb2dub3MtYW5hbHl0aWNz&ntb=1" class="btn btn-primary">Learn More</a>
                </div>
                <div class="col-lg-6">
                        <img src="img/BDA HC8.webp" class="img-fluid">
                </div>
        </div>
</div>


<div class="container-fluid padding"data-aos="fade-up" data-aos-
                offset="200"
                data-aos-delay="50"
                data-aos-duration="1000">
               <div class="row padding">
                    <div class="col">
                                <h1 > MODULE 2: Data Preparation & Exploration</h1>
                </div>
                <div class="col-md-12 col-lg-6">
                        <h2>Loading Data into Cognos:</h2>
                        <p> The collected data will be loaded onto the IBM Cognos analytics
platform for further analysis.</p>
```

```
                    <h2>Data Cleaning & Transformation:</h2>
                    <p>We will clean the data by addressing missing values, inconsistencies, and
formatting issues. Necessary transformations will be applied to prepare the data for analysis.</p>
                    <h2>Exploratory Data Analysis (EDA):</h2>
                    <p>We will explore the data through various techniques to understand its characteristics,
identify trends, and uncover potential relationships between variables.</p>

                </div>

            </div>
        </div>
        <div class="container-fluid padding"data-aos="fade-up" data-aos-
                offset="200"
                data-aos-delay="50"
                data-aos-duration="1000">
            <div class="row padding">
                <div class="col">
                            <h1 > MODULE 3: Data Visualization & Storytelling </h1>
                </div>
                <div class="col-md-12 col-lg-6">
                    <h2>Data Visualization:</h2>
                    <p> We will create compelling visualizations (charts, graphs, etc.) to effectively
communicate insights from the data.</p>
                    <h2>Interactive Dashboard & Story Creation:</h2>
                    <p> An interactive dashboard will be built in Cognos to showcase key findings and allow
users to explore the data dynamically. Additionally, a data store will be crafted to present the insights in a
clear and engaging narrative.</p>
                    <h2>Reports Generation:</h2>
                    <p> Comprehensive reports will be generated to document the analysis process,
findings, and recommendations.</p>
                    <br>

                </div>

            </div>
        </div>
        <div class="container-fluid padding"data-aos="fade-up" data-aos-
                offset="200"
                data-aos-delay="50"
                data-aos-duration="1000">
            <div class="row padding">
                <div class="col">
                            <h1 > MODULE 4: Displaying Insights in Dashboard</h1>
                </div>
                <div class="col-md-12 col-lg-6">
```

```
                    <h2>Insights Integration:</h2>
                    <p> The key insights gleaned from the data analysis will be integrated
seamlessly into the Cognos dashboard, providing users with actionable information.</p>

                        <br>

            </div>

        </div>
</div>

<hr class="my-4">

<!--- Meet the team -->
<div class="container-fluid padding"id="section2">
        <div class="row welcome text-center">
            <div class="col-12">
                <h1 class="display-4" >Meet the Team </h1>
            </div>
        </div>
</div>

<!--- Cards -->
<div class="container-fluid padding" data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
        <div class="row padding">
                    <div class="col-md-4">
                        <div class="card">
                        <img class="card-img-top" src="">
                        <div class="card-body">
                            <h4 class="card-title">K. Manoj</h4>
                            <p class="card-text">Reg no: 210220104013.</p>
                            <a href="#" class="btn btn-outline-secondary">See
Profile</a>
                                                                    </div>
                                                                </div>
            </div>
            <div class="col-md-4">
                <div class="card">
                    <img class="card-img-top" src="">
                    <div class="card-body">
                        <h4 class="card-title">S. Sakthivel</h4>
                        <p class="card-text">Reg no: 210220104022.</p>
                        <a href="#" class="btn btn-outline-secondary">See
Profile</a>
```

```
                    </div>
                </div>
            </div>
            <div class="col-md-4">
                <div class="card">
                    <img class="card-img-top" src="">
                    <div class="card-body">
                        <h4 class="card-title">G. Yogesh waran</h4>
                        <p class="card-text">Reg no: 210220104032.</p>
                        <a href="#" class="btn btn-outline-secondary">See
Profile</a>
                                                                    </div>
                                                                </div>
            </div>
        </div>
</div>

<!--- Two Column Section -->
<div class="container-fluid padding" data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
        <div class="row padding">
            <div class="col-md-12 col-lg-6">
                <h2>How Is Data Analytics Used in Health Care?</h2>
                <p>Data analytics is the process of analyzing raw data to determine trends and
enable better decision-making. It is relevant to all types of organizations, especially healthcare
organizations.</p>
                <p>Data analytics in health care is vital. It helps healthcare organizations to evaluate and
develop practitioners, detect anomalies in scans, and predict outbreaks in illness, per the Harvard Business
School. Data analytics can also lower costs for healthcare organizations and boost business intelligence.
Most importantly, it helps healthcare companies to make better care decisions for patients.</p>

            </div>
            <div class="col-lg-6">
                <img src="img/OIP.webp" class="img-fluid">
            </div>
        </div>
        <hr class="my-4" >
</div >

<!--- Connect -->
<div class="container-fluid padding" >
        <div class="row text-center padding">
            <div class="col-12" id="section3">
                <h2 >Connect</h2>
```

```
            </div>
            <div class="col-12 social padding">
                <a href="#"><i class="fab fa-facebook"></i></a>
                <a href="#"><i class="fab fa-twitter"></i></a>
                <a href="#"><i class="fab fa-google-plus-g"></i></a>
                <a href="#"><i class="fab fa-Instagram"></i></a>
                <a href="#"><i class="fab fa-youtube"></i></a>
            </div>
        </div>
</div>

<!--- Footer -->
<footer>
        <div class="container-fluid padding" data-aos="fade-up" data-aos-
            offset="200"
            data-aos-delay="50"
            data-aos-duration="1000">
            <div class="row text-center">
                <div class="col-md-4 pt-0">
                    <div class="footer-image">
                    <img src="img/apolloLogo.webp"></div>
                    <hr class="light">
                    <p>555-555-5555</p>
                    <p>apolloengineeringcollege@gmail.com</p>
                    <p>kuthambakkam</p>
                    <p>poonamallee, chennai, Tamil Nadu , 600123</p>
                </div>
                <div class="col-md-4">
                    <hr class="light">
                    <h5>Our hours</h5>
                    <hr class="light">
                    <p>Monday: 9 am - 5 pm </p>
                    <p>Saturday: 10 am - 4 pm </p>
                    <p>Sunday: Closed</p>
                </div>
                <div class="col-md-4">
                    <hr class="light">
                    <h5>Service Area</h5>
                    <hr class="light">
                    <p>City, State, 900</p>
                    <p>City, State, 800</p>
                    <p>City, State, 700</p>
                    <p>City, State, 600</p>
                </div>
                <div class="col-12">
                    <hr class="light-100">
                    <h5>&copy; This Website created for learning purpose only</h5>
                </div>
```

```
        </div>
    </div>
</footer>

<script src="https://unpkg.com/aos@next/dist/aos.js"></script>
<script>
AOS.init();
</script>

</body>
</html>
```

Css:

```
@import url('https://fonts.googleapis.com/css?family=Poppins:400,500,700');

html,
  body{ height:
  100%;
  width: 100%;
  font-family: 'Poppins', sans-serif;
  color:#222;
}
.navbar{ padding
  : .8rem;
}
.navbar-brand
  img{width: 50px;
  height:60px;
}
.navbar-nav
  li{ padding-right:
  20px;
}
.nav-link{
  font-size: 1.1em !important;
}
.carousel-inner
  img{width: 100%;
}
.carousel-caption
  { position:
```

```
}
.carousel-caption h1 { font-
    size: 500%;
    text-transform: uppercase;
    text-shadow: 1px 1px 10px #000;
}
.carousel-caption h3 { font-
    size: 200%;
    font-weight: 500;
    text-shadow: 1px 1px 10px #000;
    padding-bottom: 1rem;
}
.btn-primary {
    background-color: #6648b1;
    border: 1px solid #564d7c;
}
.btn-primary:hover { background-
    color: #563d7c; border: 1px solid
    #563d7c;
}
.jumbotron { padding:
    1rem; border-radius: 0;
}
.padding {
    padding-bottom: 2rem;
}
.welcome { width:
    75%;
    margin: 0 auto; padding-
    top: 2rem;
}
.welcome hr {
    border-top: 2px solid #b4b4b4; width:
    95%;
    margin-top: .3rem; margin-
    bottom: 1rem;
}
.fa-code { color:
    #e54d26;
}
.fa-bold { color:
    #563d7c;
}
.fa-css3 { color:
    #2a63af;
}
.fa-code, .fa-bold, .fa-css3 {
```

```
    font-size: 4em;
    margin: 1rem;
}
.fun{
    width: 100%;
    margin-bottom: 2rem;
}
.gif {
    max-width: 100%;
}
.social a {
    font-size: 4.5em;
    padding: 3rem;
}
.fa-facebook { color:
    #3b5998;
}
.fa-twitter { color:
    #00aced;
}
.fa-google-plus-g { color:
    #dd4b39;
}
.fa-instagram { color:
    #517fa8;
}
.fa-youtube { color:
    #bb0000;
}

.fa-Facebook: hover,
.fa-Twitter:hover,
.fa-google-plus-g: hover,
.fa-Instagram:hover,
.fa-youtube:hover { color:
    #d5d5d5;
}
footer {
    background-color: #3f3f3f; color:
    #d5d5d5;
    padding-top: 2rem;
}
hr. light {
    border-top: 1px solid #d5d5d5; width:
    75%;
    margin-top: .8rem; margin-
    bottom: 1rem;
}
```

```
footer a { color:
    #d5d5d5;
}
hr.light-100 {
    border-top: 1px solid #d5d5d5; width:
    100%;
    margin-top: .8rem; margin-
    bottom: 1rem;
}
.footer-image
    img{ width: 50px;
    height: 60px;
}
.img-fluid{ width:
    auto; height: 400px;

}
.col{
    display: flex;

    align-items: center;
    padding: 100;


}

/*---Media Queries --*/ @media
(max-width: 1500px){
    .carousel-inner img{ width:
        100%; height: 550px;
    }
}

@media (max-width: 992px) {
    .social a{
        font-size: 4em;
        padding: 2rem;
    }
    .carousel-inner img{ width:
        100%; height: 400px;
    }
    }
@media (max-width: 768px) {
    .carousel-caption { top:
        45%;
```

```
    }
    .carousel-caption h1 { font-
        size: 350%;
    }
    .carousel-caption h3 { font-
        size: 140%;
        font-weight: 400; padding-
        bottom: 0.2rem;
    }
    .carousel-caption .btn { font-
        size: 95%; padding: 8px
        14px;
    }
    .display-4 {
        font-size: 200%;
    }
    .display-4 {
        font-size: 200%;
    }
    .social a{
        font-size: 2.5em;
        padding: 1.2rem;
    }
    .carousel-inner img{ width:
        100%;
        height: 100%;
    }
}
@media (max-width: 576px)
    { carousel-caption {
        top: 40%;
    }
    .carousel-caption h1 { font-
        size: 250%;
    }
    .carousel-caption h3 { font-
        size: 110%;
    }
    .carousel-caption .btn { font-
        size: 90%; padding: 4px 8px;
    }
    .carousel-indicators { display:
        none;
    }
    .display-4 {
        font-size: 160%;
    }
```

```
    .social a{
        font-size: 2em;
        padding: 1.7rem;
    }
    .carousel-inner img{ width:
        100%;
        height: 100%;
    }
}



/*---Firefox Bug Fix --*/
.carousel-item {
    transition: -webkit-transform 0.5s ease; transition:
    transform 0.5s ease;
    transition: transform 0.5s ease, -webkit-transform 0.5s ease;
    -webkit-backface-visibility: visible; backface-
    visibility: visible;
}

/*--- Bootstrap Padding Fix --*/
[class*="col-"] {
        padding: 1rem;
}




/*
Extra small (xs) devices (portrait phones, less than 576px) No media query since
this is the default in Bootstrap

Small (sm) devices (landscape phones, 576px and up) @media (min-
width: 576px) { ... }

Medium (MD) devices (tablets, 768px and up) @media
(min-width: 768px) { ... }

Large (LG) devices (desktops, 992px and up) @media
(min-width: 992px) { ... }

Extra (xl) large devices (large desktops, 1200px and up) @media (min-
width: 1200px) { ... }
*/
```

Story.html:

```
<!DOCTYPE html>
<html lang="en">
<head>
        <meta charset="UTF-8">
        <meta name="viewport" content="width=device-width, initial-scale=1.0">
        <title>Strory</title>
                <link rel="stylesheet" href="https://unpkg.com/aos@next/dist/aos.css"
/>
        <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/css/bootstrap.min.css">
        <script src="https://ajax.googleapis.com/ajax/libs/jquery/3.3.1/jquery.min.js"></scrip t>
        <script src="https://cdnjs.cloudflare.com/ajax/libs/popper.js/1.12.9/umd/popper.min.js "></script>
        <script src="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/js/bootstrap.min.js"></sc ript>
        <script src="https://use.fontawesome.com/releases/v5.0.8/js/all.js"></script>
        <link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/animate.css/4.1.1/animate.min.css
">
        <link href="style.css" rel="stylesheet">
</head>
<body>
        <div>
                <iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=story&amp;pathRef=.my_fo
lders%2FHealthcare%2Bstory1&amp;closeWindowOnLastView=true&amp;ui_appbar=false
&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;sceneId=model0
000018f04628659_00000000&amp;sceneTime=0" width="1500" height="720" frameborder="0"
gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
        </div>

        <div class="container-fluid padding"data-aos="fade-up" data-aos-
                        offset="200"
                        data-aos-delay="50"
                        data-aos-duration="1000">
                    <div class="row padding">
                        <div class="col">
                            <h1 > STORY TELLING</h1>
                            <p>I can't process images, but based on the text you provided, it looks like the
dashboard you described can be used to tell a story about
```

```
the demographics of patients at a healthcare facility. Here are some possible uses of healthcare data analytics stories on a website:</p>
                </div>
                <div class="col-md-12 col-lg-6">
                    <h2>Improve patient satisfaction:</h2>
                    <p> By understanding how patients are using the facility, you can identify areas for improvement, such as reducing wait times or making it easier to schedule appointments.</p>
                    <h2>Target marketing:</h2>
                    <p> You can use data analytics to identify which types of patients are most likely to use your services, and then target your marketing efforts accordingly. For example, you might want to run a social media campaign targeted at young parents if your data shows that you see a lot of children.</p>
                    <h2>Improve operational efficiency:</h2>
                    <p> Data analytics can help you to identify areas where you can improve efficiency, such as by streamlining scheduling or coding processes.</p>
                    <h2>Support population health management:</h2>
                    <p> Data analytics can be used to identify patients who are at high risk of certain conditions, and then target interventions to those patients. This can help to improve population health outcomes and reduce overall healthcare costs.</p>
                    <h2>Attract new patients:</h2>
                    <p> You can use data analytics to show potential patients the quality of care that you provide. For example, you might publish a blog post that shows how your facility has reduced wait times or improved patient satisfaction scores.</p>
                    <br>

                </div>

            </div>
        <div>

</body>
</html>
```

Dash.html:

```
<!DOCTYPE html>
<html lang="en">
<head>
        <meta charset="UTF-8">
        <meta name="viewport" content="width=device-width, initial-scale=1.0">
        <title>Dashboard</title>
        <link rel="stylesheet" href="https://unpkg.com/aos@next/dist/aos.css" />
        <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/css/bootstrap.min.css">
        <script src="https://ajax.googleapis.com/ajax/libs/jquery/3.3.1/jquery.min.js"></scrip t>
        <script src="https://cdnjs.cloudflare.com/ajax/libs/popper.js/1.12.9/umd/popper.min.js "></script>
        <script src="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/js/bootstrap.min.js"></sc ript>
        <script src="https://use.fontawesome.com/releases/v5.0.8/js/all.js"></script>
        <link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/animate.css/4.1.1/animate.min.css
">
        <link href="style.css" rel="stylesheet">
</head>
<body>
        <iframe src="https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.m
y_folders%2FHealthcare%2Bdashboard&amp;closeWindowOnLastView=true&amp;ui_appba
r=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mode=da
shboard&amp;subView=model0000018f00765e53_00000000" width="1500" height="700"
frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
</body>
</html>
```

12 REFERENCES:

[1] Javier Nieto León, “Big Data In Healthcare 2021.”

https://www.htechtrends.com/big-data-in-healthcare/#:~:text=Big data in healthcare refers to the collection%2C, processed by machine learning techniques and data analysts.

[2] D. Chong and H. Shi, “Big data analytics: a literature review,” J. Manag. Anal., vol. 2, no. 3, pp. 175–201, 2015, DOI: 10.1080/23270012.2015.1082449

[3] Tutorialspoint, “Big Data Analytics Tutorial.”

https://www.tutorialspoint.com/big_data_analytics/index.htm.

[4] S. Kumar and M. Singh, “Big data analytics for the healthcare industry: Impact, applications, and tools,” Big Data Min. Anal., vol. 2, no. 1, pp. 48–57, 2019, DOI:10.26599/BDMA.2018.9020031.

[5] M. Ambigavathi, “Big Data Analytics in Healthcare,” 2018 Tenth Int. Conf. Adv. Comput., pp.269–276, 2018.