ANALYTICS FOR HOSPITAL'S HEALTH-CARE DATA

A UG PROJECT REPORT

Submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,HYDERABAD

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

RAGAM MAHENDER **19UK1A05J4**

Under the guidance of

Mr.P. ILANNA

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
(Affiliated to JNTUH,Hyderabad)
Bollikunta,WarangaL–506005

2019– 2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
BOLLIKUNTA,WARANGAL–506005
2019– 2023

CERTIFICATE OF COMPLETION

UG PROJECT PHASE-1

This is to certify that the UG Project Phase-1 entitled **" ANALYTICS FOR HOSPITAL'S HEALTH-CARE DATA "** is being submitted by **RAGAM MAHENDER(H.NO:19UK1A05J4),** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering to Jawaharlal Nehru Technological University Hyderabad** during the academic year **2022-23**, is a record of work carried out by them underthe guidance and supervision.

Project Guide
Mr.P.ILANNA
(Assistant Professor)

Head of the Department
Dr.R.Naveen Kumar
(Professor)

External

ACKNOWLEDGEMENT

I wish to take this opportunity to express my sincere gratitude and deep sense of respect to my beloved **Dr.P.PRASAD RAO,** Principal, Vaagdevi Engineering College for making me available all the required assistance and for his support and inspiration to carry out this UG Project Phase-1 in the institute.

I extend my heartfelt thanks to **Dr.R.NAVEEN KUMAR**, Head of the Department of CSE, Vaagdevi Engineering College for providing me necessary infrastructure and thereby giving me freedom to carry out the UG Project Phase-1.

I express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their constant supervision as well as for providing necessary information regarding the UG Project Phase-1 and for their support in completing the UG Project.

I express heartfelt thanks to the guide, **Mr. P. ILANNA** Assistant professor, Department of CSE for his constant support and giving necessary guidance for completion of this UG Project Phase-1.

Finally, I express my sincere thanks and gratitude to my family members, friends for their encouragement and outpouring their knowledge and experience throughout the this.

RAGAM MAHENDER **19UK1A05J4**

ABSTRACT

The main aim of this paper is to provide a deep analysis on the research field of healthcare data analytics., as well as highlighting some of guidelines and gaps in previous studies. This study has focused on searching relevant papers about healthcare analytics by searching in seven popular databases such as google scholar and springer using specific keywords, in order to understand the healthcare topic and conduct our literature review. The paper has listed some data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical operations, reports, decision making, and prediction and prevention system. Moreover, the systematic review has showed an interesting demographic of fields of publication, research approaches, as well as outlined some of the possible reasons and issues associated with healthcare data analytics, based on geographical distribution theme.

TABLE OF CONTENTS

LIST OF FIGURES PAGE NO

1.INTRODUCTION

1.1 OVERVIEW

Today's healthcare industries are moving from volume-based business into value-based business, which requires an overwork from doctors and nurses to be more productive and efficient. This will improve healthcare practice, changing individual life style and driving them into longer life, prevent diseases, illnesses and infections.

Over the last few years, healthcare data has become more complex for the reason that large amount of data are being available lately, along with the rapid change of technologies and mobile applications and new diseases have discovered. Therefore, healthcare sectors have believed that healthcare data analytics tools are really important subject in order to manage a large amount of complex data, which can lead to improve healthcare industries and help medical practice to reach a high level of efficiency and work flow accuracy, if these data analytics tools applied correctly, but the questions are how healthcare organizations are applying these tools today, and how to think about it's future use? Also, what are the challenges they face when using such tools? And finally, what are the innovations can healthcare add to meet these challenges?

This paper aims to proof that healthcare data analytics techniques are not efficient enough and suitable anymore these days in order to manage big data issue and improve healthcare data analytics due to the rapid growth and evolution of technology. Moreover, it's also aims to promise professionals of a better quality of medical results, as well as reduce time needed to analyze healthcare data by keeping systems up to-date and sorting medical data in a logical structure along with accessing and retrieving patient's historical data fast and smoothly. Stakeholder 2 (Doctors and nurses).

1.2 PURPOSE

Healthcare has generated copious amounts of data for decades — patient records, admission rates, seasonal infections, insurance claims, prescription fulfillment, and so much more. But it's only in recent years that tools have become available to mine this data to derive meaningful insights. Healthcare analytics is being seen, and rightly so, as driving the digital transformation in healthcare. Decision-makers across the healthcare industry are leveraging analytics to optimize care, improve marketing, and raise stakeholder value. In this article, we explain what healthcare analytics is, how it's reshaping healthcare, and how it's aiding the fight against COVID-19.

Healthcare analytics has come into its own during the COVID-19 pandemic. The healthcare industry and governments are looking to analytics to curb transmission of the coronavirus, ensure care for the most vulnerable populations, predict infection rates, and now roll out vaccines. Healthcare institutions have been instrumental in developing predictive models for COVID-19.

- **Predicting the risk of transmission**: An analytics tool has been created to determine which tients are at greatest risk of coronavirus transmission. This tool is helping decide whether to discharge or
- **Optimizing COVID-19 treatment**: Developed by the Cleveland Clinic, a predictive model developed in April 2020 helps forecast patient volume, bed capacity, ventilator availability, and more.

- **Determining the spread of the virus**: Researchers at Binghamton University have designed machine-learning algorithms to predict transmission rates at different places in the U.S.

Healthcare analytics will continue to play an important role in charting the progression of the disease and how it is best mitigated. Now, analytics is being used in vaccine distribution, providing answers to such questions as how many doses are required and where the vaccine is needed most.

2.LITERATURE REVIEW

2.1 Information System & Information Technology in Healthcare Sectors

The healthcare sector is widely considered as one of the most important industries in information technology (Wager 2005). More and more, information technology has been considered as a practice that facilitates healthcare performance through using data and information efficiently within the healthcare sectors. Therefore, Wager et al (2005) said that in order to understand the relation between information technologies and healthcare, we first need to understand what are the technologies used in healthcare.

Information technology functions have developed over the last few years not only as a technology services provider, but also as a strategic provider that develops and integrates industries' infrastructures to facilitate and ensure quality of service (LeRouge et al 2007).

2.1 Healthcare Analytics & Data Mining

Data Mining is described as a process by which data is gathered, analysed and stored in order to produce useful and high-quality information and knowledge. This term also includes the way of how this data is gathered,filtering and preparation of the data for use and finally the processing of data to support data analytics and predictive modelling (Russom 2011).

2.1.1 Data collection

The first stage of data mining is the process of gathering and collecting data. However, even before gathering the data, ideas and plans should be assumed to decide which data should be gathered in order to collect specific data as desired and use it efficiently (Lamont, 2010).
Furthermore, Chordas (2001) added that a lot of projects fail and exceed estimated costs because of poor quality of gathered data which can result from poor data cleaning.

2.2 Healthcare Sectors & Big Data Analytics

2.2.1 Big data storage and management

One of the most important elements in dealing with and managing data is to know where and how this data will be stored once when it is collected. The traditional methods of storing and retrieving such data are not efficient anymore, since it was structured and stored in data warehouses and relational databases,after extracting and loading it from different outside sources. However, this data is transformed and classified before being ready to use and function (Bakshi 2012).

Furthermore, Herodotou et al (2011) agreed with Bakshi (2012) when he said that there are many numbers of data sources now and that a huge amount of data has become available, so this growth of data will absolutely require an agile database which can deal with the data logically and through data synchronization in order to adapt to the rapid data evolution.

On the other hand, Plattner and Zeier (2011) stated that databases only manage server memory data, therefore eliminating the option of managing other storage devices such as: disk and compact drivers.Accordingly, this will reduce the efficiency of database performance and real time response during the time.

2.2.2 Patients Role in Healthcare Analytics

This section is concerned about how individuals (and patients in specific) can improve healthcare analytics through understanding the small and personal data, as well as educate themselves in how to collaborate with the healthcare data analytics to reach a high level of efficiency and accuracy (Luciano 2013).

Swan (2012) was discussing the same point when he identified the term "citizen science", where nonprofessional and educated individuals are skilled enough to conduct and support healthcare analytics system. Accordingly, this will require organizations to train individuals how to follow up and track their health information, as well as self-monitoring.

Principally, to perform good data analytics, first of all we should teach individuals how to understand and realize the importance of dealing with such data, for instance how to deal with breast cancer (Hanoch 2012). However, Miron et al (2011) believed that whatever and how much our patients are educated and skilled to provide us with the data we expect, medical professionals still highly need to test and clarify this data to consider it and keep it on record. Also, he added that once when the data has been tested and clarified, we then need to find out how to change an individual's behavior, starting with parents and guardians who are responsible for raising their children.

2.2.3 Connectivity between Healthcare Analytics System and Individuals (Medical Staff and Patients)

Connectivity approaches generate thoughts and ideas from connected networks of minds and leverages prior experiences with the utilization of technologies in our everyday life. (Siemens 2004) Moreover, McHorney (2009) has added that healthcare analytics is not solely regarding technology and the knowledge however; it is also regards how much individuals are attached to and familiar with medical care systems and their personal skills such as ability to learn and adopt such systems in their life, as different people have different attitudes and reasons for not accepting such technologies, especially older people.

2.2.4 Healthcare Predictions and Decision Support System (DSS)

Healthcare prediction is another data analytics method focusing on reducing future medical costs. Predictive technique uses patient medical history to evaluate all the potential health risks and predict a future medical treatment in advance (LexisNexis 2015).

(Loginov et al 2012) stated that by retrieving and reviewing past patient details, information and diagnoses from the databases, predictive methods can take a place through forecasting, reducing time and costs.

Parkland hospital in Dallas, Texas has launched a predictive system which scans all patients' details and information to identify potentials risks and outcomes. As a result, the hospital has saved more than half a million dollars, especially in heart failure and disease predictions in terms of performing patients' monitoring and avoiding future complications (Jacob 2012).

2.2.5 Role of Predictive Analytics in Medical Healthcare

Predictive analytics supports healthcare sectors to achieve a high level of effective overall care and preventive care, as predictive systems' results allow treatments and actions to be taken when all the risks are recognized in early stages, which aids for minimizing costs. (Conley et al 2008).
Furthermore, Obenshain (2004) said that patients can also work and support medical care by following up and updating their medical status, so they can get the necessary treatment at the right time. The technology era has added significant value to the healthcare decision support system, since decision making systems in healthcare care sectors can be enhanced by focusing on patient diagnoses, behavior, and prevention in order to reach a high level of care and improve healthcare economics (Cannon & Tanner 2007).

2.2.6 Healthcare Prediction Examples

In the healthcare sectors, predictive analytics can be achieved in many ways such as; a medical care delivery success, which can be achieved by using a model that proposes algorithms in order to assist medical treatment for interacting diseases, which can reflect in capturing patient's behavior and interactions. Another method of using predictive analytics regards how to use applications and software services alongside the electronic clinical records to analyze diagnosis and confirm outcomes in order to provide the correct treatment for the right patient at the right time (Lamont 2010). Moreover, Imamura has found that the association diagnostic approach can effect efficiently in extracting desirable information from huge databases (Imamura et al 2007).

2.2.7 Financial Factors in Healthcare Predictive Analytics

The most significant and obvious result of using such technology within the healthcare sectors is its results on costs. Because of cost, information is one of the main aspects that have a big effect on the cost of healthcare predictive analytics. Medical care systems have focused on increasing healthcare analytics performance as well as minimizing the cost by simplify unstructured clinical record and reducing irregular information. Consequently, large quantities of information then will be managed and controlled smoothly and efficiently (Bertsimas et al 2008).

Predictive analytics can assist to avoid and reduce inaccurate prediction costs plus time for the reason that it makes the data sourcing cost lower by specifying the desired and necessary data only, since the data is simplified, standardized and exists in historical clinical databases (Bradley & Kaplan 2010)

2.2.8 Healthcare Analytics & Real Time

Murphy (2013) believed that real time analytics produces more accurate results and information, since it evaluates current patients' history and conditions, therefore investigating patients' diagnosis correctly and offering the best treatment.

Real time monitoring techniques guarantee to keep data up to date and increase the quality of information, as assumed so by Taylor (2010). She believed that real time matters in healthcare analytics are very significant for the reason that it generates accurate results, such as where diabetes patients can recover if their ailment is discovered and treated correctly in the earlier stages. Walker et al (2012) agreed with that

however they also highlighted some of its disadvantages, such as high cost, its high required level of training and long time to complete.

2.2.9 PROPOSED SYSTEM

A number of research articles proposed using big data analytics in varying domains especially in healthcare such as Kumar et al. 5 proposed a cognitive technology-based healthcare evaluations system using big data analytics.

3.THEORETICAL ANALYSIS

3.1 BLOCK DIAGRAM

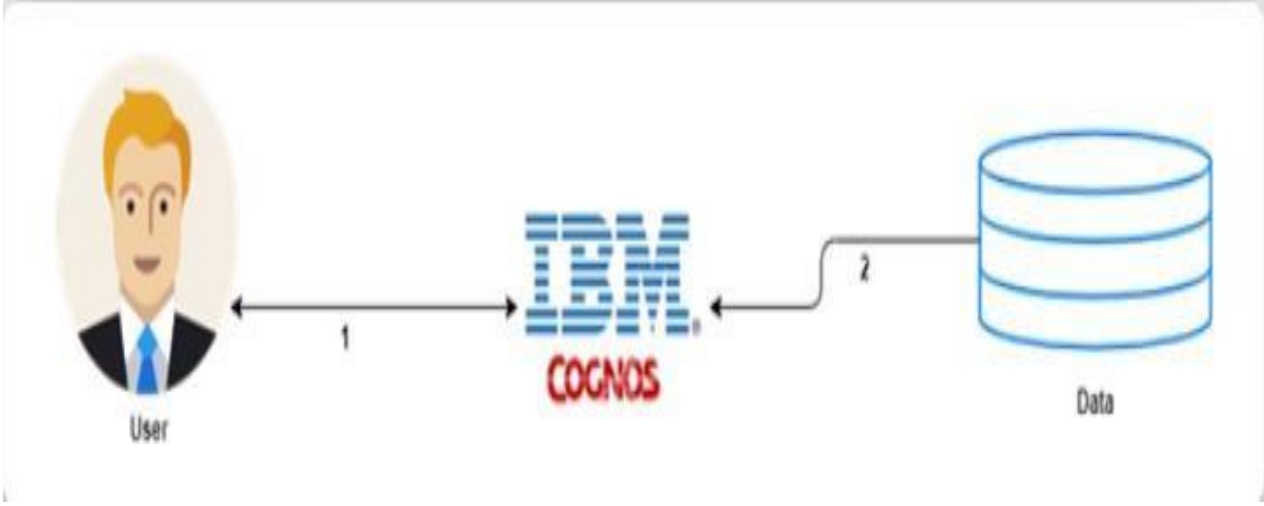

Fig 3.1.1

3.2 METHOD

The objective of this paper was to conduct a review, which encourages professionals, doctors, medical staff patients to adopt and utilize technologies in order to assist healthcare analytics and improve decision making process in our everyday life.

Our method has followed three steps: 1) searching for initial and related studies, 2) Relevance appraisal and evaluation, and finally extracting data. The next sections will explain these steps briefly.

Searching for initial and related studies: the first step in order to find the articles was to specify and identify main keywords (Dieste et al. 2009). A survey was conducted to study relevant papers published since 2010 in the information system field in general and healthcare analytics and medical decision

support system in specific. This study has found that most relevant keywords to "healthcare analytics" and "data mining" used with technology to support medical information systems.

The following searching phrases were used and structured in searching for relevant papers in many different databases – i.e the relevant and related papers should contain in its titles, keywords, abstract or full text the word "healthcare" along with any of "analytics", "metrics", "data mining", "big data" or "Decision making" see Table 1.

Group A	Group B
E-Healthcare	Analytics
Medical Practice	Metrics
Health	Decision Making
Clinical	Prediction
Hospitals	Big Data
Care systems	Data Mining
Wellness	Business
Programs	Intelligence

Table 3.2.1

Relevance and related papers step was completed by filtering relevant papers from the initial papers list and eliminating undesired ones, by filtering papers first based on titles, abstracts and finally full texts. Essentially, eliminated papers and articles have been done following exclusion criteria according to many main factors such as:

- Did not focus on utilizing technologies in order to improve healthcare analytics.
- Did not provide any applicable methods and experimental evidence.
- Were in different languages not English.
- Were not relevant enough and kind of old, which cannot be applied in these days.
- Were not available anymore.

Finally, extracting and analysis of data step comes after specifying and identifying related papers, so from he obtained papers now it can be extract 1) the year of publications, 2) methods and tools have been used professionals, 3) investigate if these tools and methods are still useful now a days or not, 4) what kind of problems that have been solved using these tools, 5) role of patients' in order to assist healthcare analytics and improve medical care decision process, 6) application areas, 7) research approaches and 8) data availability and geographical area of data gathering.

4.FLOWCHART

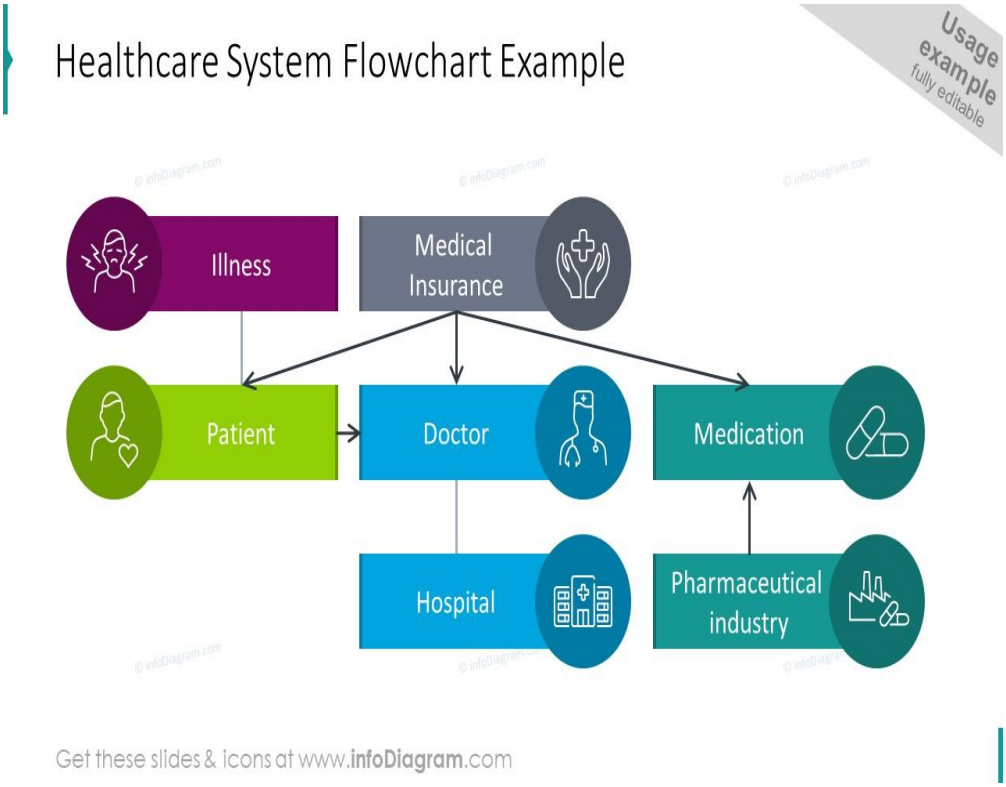

Fig 4.1

CLASS DIAGRAM:

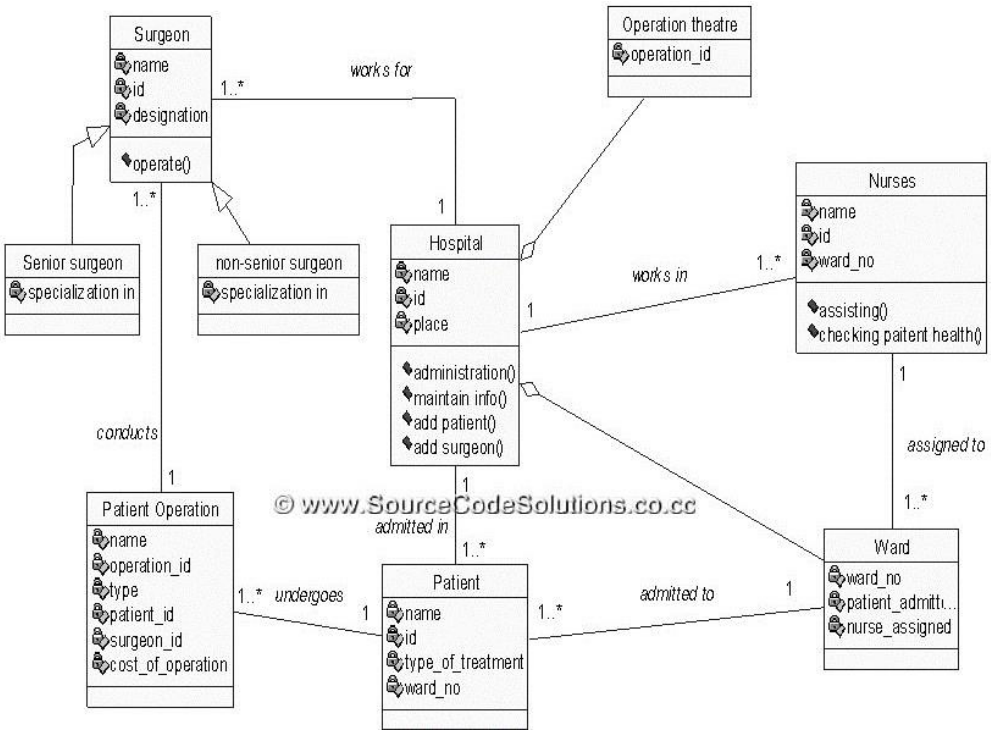

Fig 4.2

USECASE DIAGRAM:

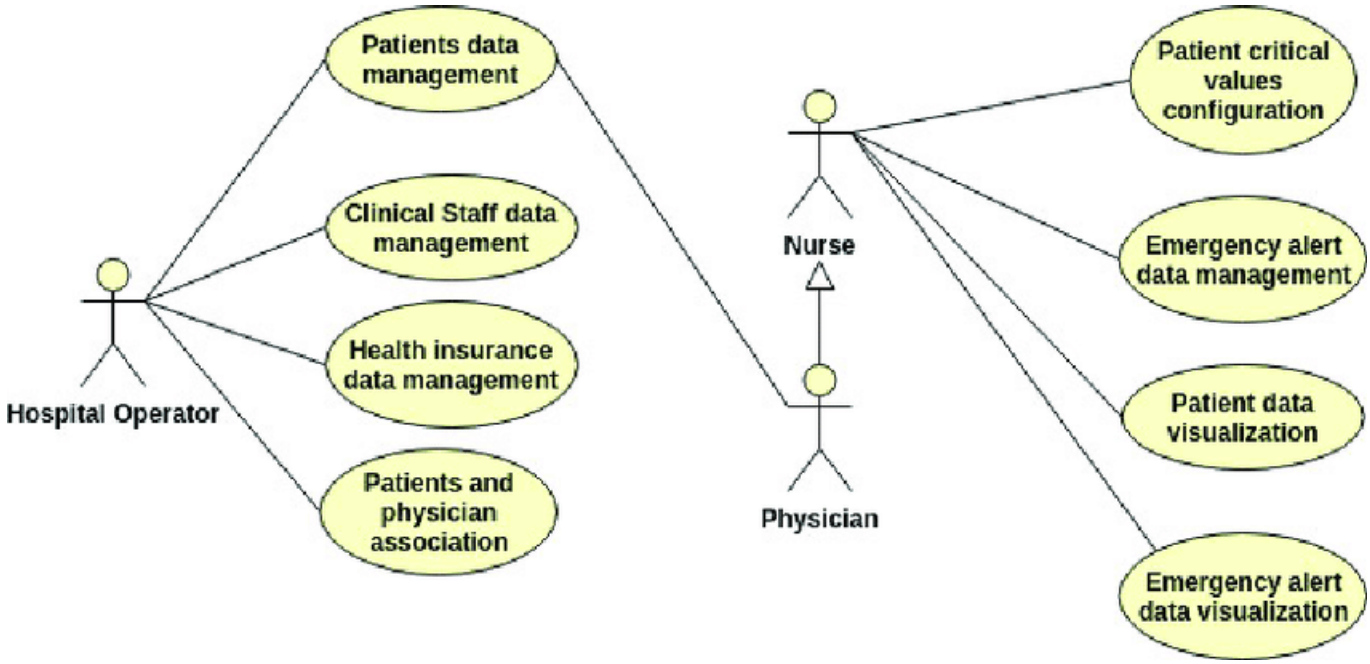

FIG 4.3

5.RESULTS

5.1 Healthcare Data Analytics Platforms and Tools:

Comparing between traditional analytics and advanced analytics, traditional analytics is focusing on business intelligence, operational research and data mining. However, advanced analytics is focusing ondescriptive, predictive and optimization (Raden 2010).

The paper has come up with different tools and techniques that would improve healthcare data analytics in order to support descriptive, predictive and prescriptive healthcare data analytics the first tool is:

• **Advanced Data Visualization (ADV**): ADV is different from other standards bars and line chart, since it can scale its visualization for millions of data points, also can handle different data types. ADV is easy to use and supports analysts to explore data widely. ADV can reduce quality problems which can occur when retrieving medical data for extra analysis. Moreover, ADV can offer rich results and fluid interactions in order to reveal clinical hidden patterns in the data. (Powell 2014; Wongsuphasawat et al. 2011)

• **Presto:** is a distributed SQL query engine used to analyze huge amount of data that collected every single day. There is nothing better for healthcare sectors to find such a product which can handle a large amount of data that will come into the system. Data can take many hours and even days to be analyzed, but with Presto data now can be analyzed in just seconds or minutes. (Wulff 2013)

Hive: is one of the programs developed in order to handle large amount of data, it's is not processing and analyzing data quickly as presto, however Hive does all excel tasks efficiently that don't need for real time performance, due to this companies can use both Presto and Hive for best performance, since presto can access data stored on Hive. (Capriolo et al. 2012)

• Vertica: program is very similar to Presto, but less expensive for the reason that Vertica eliminated costly architecture that used to associate with large amount of data. Also, Vertica has the feature of scalability which means it can cover hospital's data and analytics no matter how that data is big. Vertica can improve healthcare by reducing operational costs, accelerating medical reports and analyzing patients' patterns. (Vertica 2010; HP 2013)

• Key Performance Indicators (KPI): is a strategy evaluates in how company is executing its strategic vision. KPI can improve quality of medical healthcare for patients who are susceptible to hospital conditions when KPI used to specify significant indicators to be monitored and corrected, as well as identifying weaknesses. Also, KPI can use electronic medical record data to identify human practice and interventions. (Al-Azzawe 2014)

• **Online Analytics Processing (OLAP):** can improve healthcare system by performing statistical calculation very fast through hierarchal and multidimensional organized data, and can increase data integrity checking, quality control and reporting services. OLAP has the ability to improve healthcare decision making system by giving a better tracking of medical records and diagnoses. (Pešiet al. 2009)

• **Online Transaction Processing (OLTP):** is similar to OLAP, but it is designated to process patient care operations, such as patient registration, hospital documents and results review. (Ledbetter and Morgan 2001)

• **The Hadoop Distributed File System (HDFS):** HDFS enhances healthcare data analytics system by dividing large amount of data into smaller one and distributed it across the other systems. Eliminating data redundancy, since HDFS has such feature built into storage layer which makes professional to focus on other responsibilities. HDFS can add a value through helping medical purposes in order to personalized treatment planning, assisting diagnosis, monitoring patient's signs and fraud detections. (Shvachko et al. 2010; Datastax 2013; Nori 2014)
• **Casandra File System (CFS):** CFS is also distributed system like HDFS, however CFS is a designated system to perform analytic operation with no single point of failure. (Datastax 2013; Lakshman and Malik 2010)

• **Map Reducing System:** Map Reducing system breaks Task into subtasks and gathering its outputs, as well as it enables many of the most common of operational calculations to be performed efficiently in a large amount of data. Map Reducing system keep tracking on each server when tasks is being performed. The key strength of using Map Reducing is the high level of parallelism, since many tasks can have performed at the same time if it's not waiting for other tasks results. (Dean and Ghemawat 2008),

• **Complex Event Processing (CEP):** CEP has come recently to the healthcare sectors, which means an event of changing in state, for instance suppose a patient gained more weight and moved from obesity to morbidity obesity. Now complex patient event processing will detect this new pattern and add it to the patient's events and relate it with being diabetic, which means that complex event processing is relating and linking events to the real time, as well as that will enhance EMR and HER systems. (Webster 2011)

• Text Mining: Text Mining tools can be used and add a value in healthcare in terms of analyzing clinical records from the hospital emergency departments of physician response on call, as a similar complaints called the emergency department and were treated differently depending on the person who answered the phone. Such matter can effect in the quality of healthcare, as well as costs. Therefore, text mining can offer a treatment plan which will develop some standards and protocols to understand this matter. (Raja et al. 2014)

• **Cloud Computing:** Cloud computing has increased hospital flexibility in order to respond for dynamic changes and latest medical updates, in addition to demonstrate a great healthcare value by reducing costs, increasing productivity and security and improve data analysis with minimal management effort or service provider interaction. Cloud computing reduce strain which caused by huge amount of clinical data. One of the cloud innovations is Phillips Healthsuite platform that manages healthcare data and support doctors and patients. Phillips Healthsuite platform stores a huge amount of clinical and patient data which can be used directly in the future as an actionable data, a source of diagnosis analysis and disease prediction and prevention to increase patient care. (IBM 2011; Philips 2015)
• **Mahout:** Mahouts is an apache project aims to generate applications that supports healthcare data analytics on Hadoop systems. (Hortonworks 2015)

6.CONCLUSION

In UG Project Phase-1, we have worked on problem statement, literature survey and also done the experimental analyses which are required for the project to move forward. In experimental analysis we have discussed about the machine learning concepts and models and explained the algorithms to be used in the project. We also discussed about the flowcharts, use case diagrams, decision tree and sequence diagrams which are used in the project. Based on the experimental analysis we have designed the model for the project. Entire designing part is involved in UG Project Phase-1.

7.FUTURE SCOPE

UG Project Phase-2 is the extension of UG Project Phase-1. UG Project Phase-2 involves all the coding and implementation of the design which we have retrieved from UG Project Phase-1. All the implementation is done and conclusions will be retrieved in the phase. We will also work on the applications, advantages, and disadvantages of the project in this phase. Future scope of the project will be also discussed in the UG Project Phase-2

ANALYTICS FOR HOSPITAL'S HEALTH-CARE DATA

A UG PROJECT PHASE-II REPORT

Submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,HYDERABAD

In partial fulfilment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

RAGAM MAHENDER **19UK1A05J4**

Under the guidance of

Mr.P. ILANNA

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
(Affiliated to JNTUH, Hyderabad)
Bollikunta, Warangal – 506005
2019– 2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
BOLLIKUNTA,WARANGAL–506005
2019– 2023

CERTIFICATE OF COMPLETION

UG PROJECT PHASE-II

This is to certify that the UG Project Phase-II entitled **" ANALYTICS FOR HOSPITAL'S HEALTH-CARE DATA "** is being submitted by **RAGAM MAHENDER(H.NO:19UK1A05J4),** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering to Jawaharlal Nehru Technological University Hyderabad** during the academic year **2022-23**, is a record of work carried out by them underthe guidance and supervision.

Project Guide
Mr.P.ILANNA
(Assistant Professor)

Head of the Department
Dr.R.Naveen Kumar
(Professor)

External

ACKNOWLEDGEMENT

I wish to take this opportunity to express my sincere gratitude and deep sense of respect to my beloved **Dr.P.PRASAD RAO,** Principal, Vaagdevi Engineering College for making me available all the required assistance and for his support and inspiration to carry out this UG Project Phase-II in the institute.

I extend my heartfelt thanks to **Dr.R.NAVEEN KUMAR**, Head of the Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving me freedom to carry out the UG Project Phase-II.

I express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their constant supervision as well as for providing necessary information regarding the UG Project Phase-II and for their support in completing the UG Project.

I express heartfelt thanks to the guide, **Mr. P. ILANNA** Assistant professor, Department of CSE for his constant support and giving necessary guidance for completion of this UG Project Phase-II.

Finally, I express my sincere thanks and gratitude to my family members, friends for their encouragement and outpouring their knowledge and experience throughout the this.

RAGAM MAHENDER **19UK1A05J4**

TABLE OF CONTENTS:

LIST OF FIGURES **PAGE NO**

1.INTRODUCTION

In today's digital age, hospitals generate vast amounts of data through electronic health records (EHRs), clinical systems, billing records, and various other sources. However, the true value of this data lies in its ability to be transformed into actionable insights. This is where analytics comes into play.

Analytics for hospital's health-care data refers to the process of using advanced data analysis techniques to extract meaningful insights and patterns from the massive volumes of data generated within a hospital setting. By leveraging analytics, hospitals can unlock the potential hidden within their data to drive informed decision-making, enhance patient care, optimize operations, and achieve better outcomes.

The primary goal of analytics in health care is to transform raw data into actionable intelligence that can be used to improve various aspects of hospital operations, patient care delivery, and overall organizational performance. By analyzing health-care data, hospitals can gain a deeper understanding of patient populations, identify trends, predict outcomes, and optimize resource allocation.

Analytics techniques employed in health-care settings include descriptive analytics, which provides a summary of historical data to understand past trends and patterns; predictive analytics, which uses historical data to make predictions and forecast future outcomes; and prescriptive analytics, which recommends optimal actions based on data analysis to improve decision-making.

With analytics, hospitals can improve patient care by personalizing treatments, identifying at-risk patients, and reducing medical errors. It enables hospitals to streamline operations, optimize resource utilization, and enhance efficiency, ultimately leading to improved patient satisfaction and better financial performance.

Moreover, analytics empowers hospitals to delve into research and innovation by analyzing data from clinical trials, research studies, and medical literature. This fosters evidence-based practices, drives advancements in medical knowledge, and contributes to improved patient outcomes.

However, implementing analytics in health care comes with its own set of challenges, such as ensuring data privacy and security, maintaining data quality and integrity, and developing the necessary technical

expertise. Overcoming these challenges requires robust data governance practices, secure infrastructure, and skilled professionals capable of extracting meaningful insights from complex data sets.

In summary, analytics for hospital's health-care data offers immense potential to revolutionize the way hospitals operate and deliver care. By leveraging advanced analytics techniques, hospitals can unlock valuable insights, make data-driven decisions, improve patient care, optimize operations, and contribute to research and innovation. Analytics has the power to transform health care into a more efficient, effective, and patient-centric industry, ultimately leading to better health outcomes for individuals and communities. UG Project Phase-2 involves all the coding and implementation of the design which we have retrieved from UG Project Phase-1. All the implementation is done and conclusions are retrieved in this phase. We will also work on the applications, advantages, and disadvantages of the project in this phase. Future scope of the project will be also discussed in the UG Phase-2.

2. CODE SNIPPETS

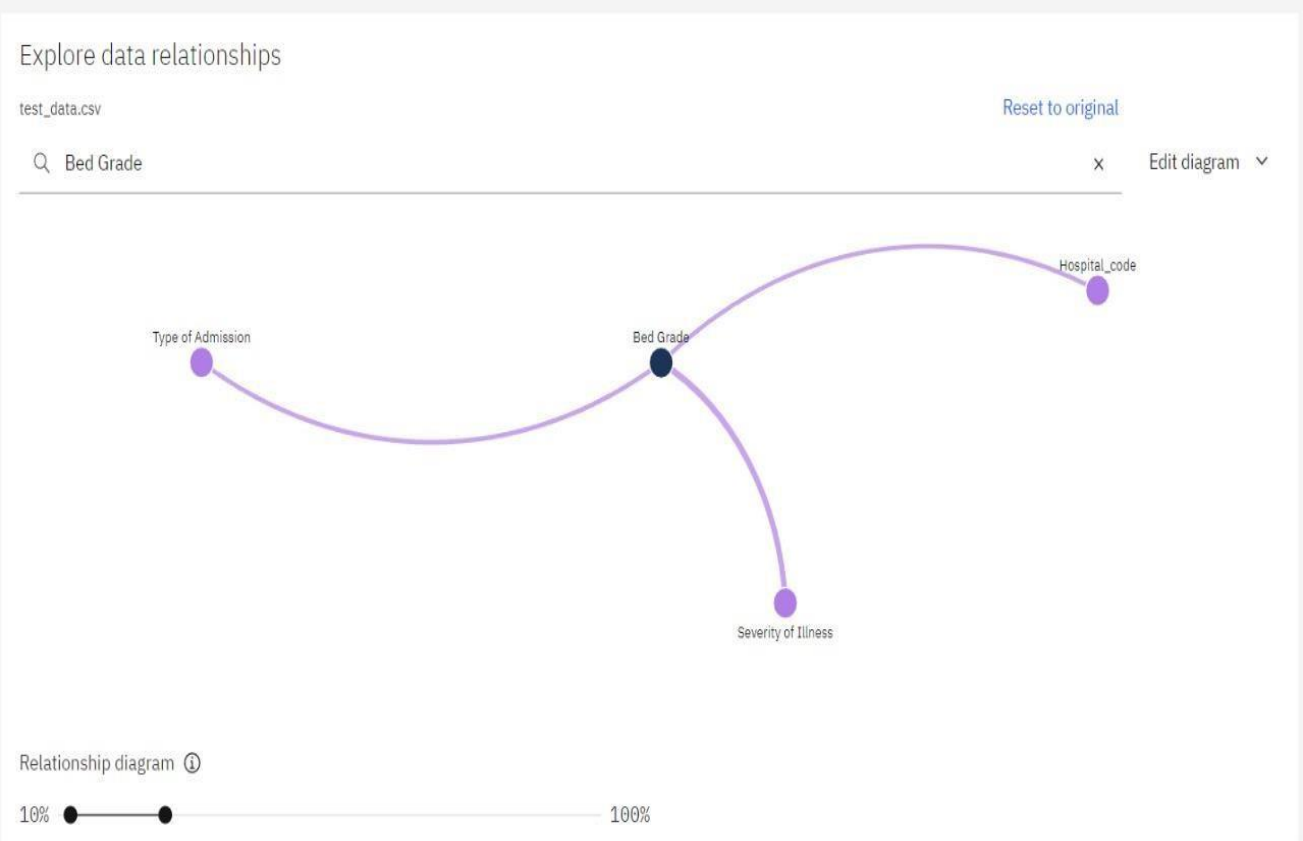

FIG 2.1

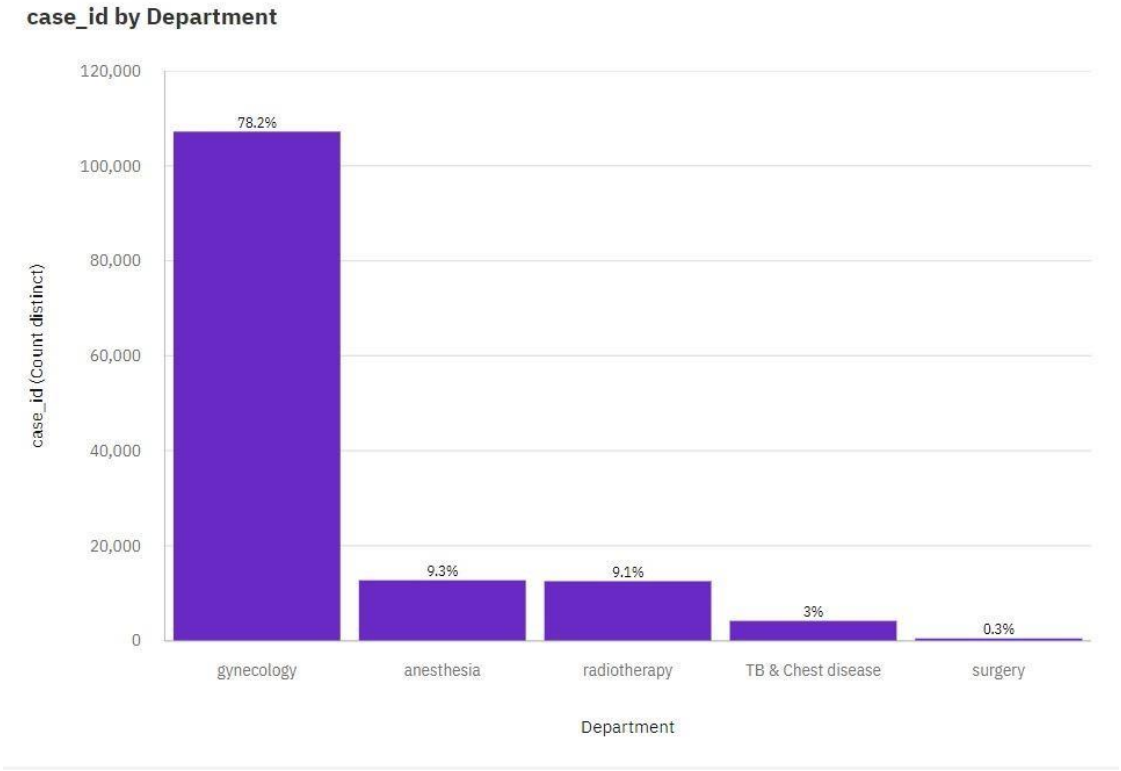
case_id by Department
120,000
100,000
80,000
60,000
40,000
20,000
0
case_id (Count distinct)
78.2%
9.3%
9.1%
3%
0.3%
gynecology
anesthesia
radiotherapy
TB & Chest disease
surgery
Department

FIG 2.2

Fig 2.3

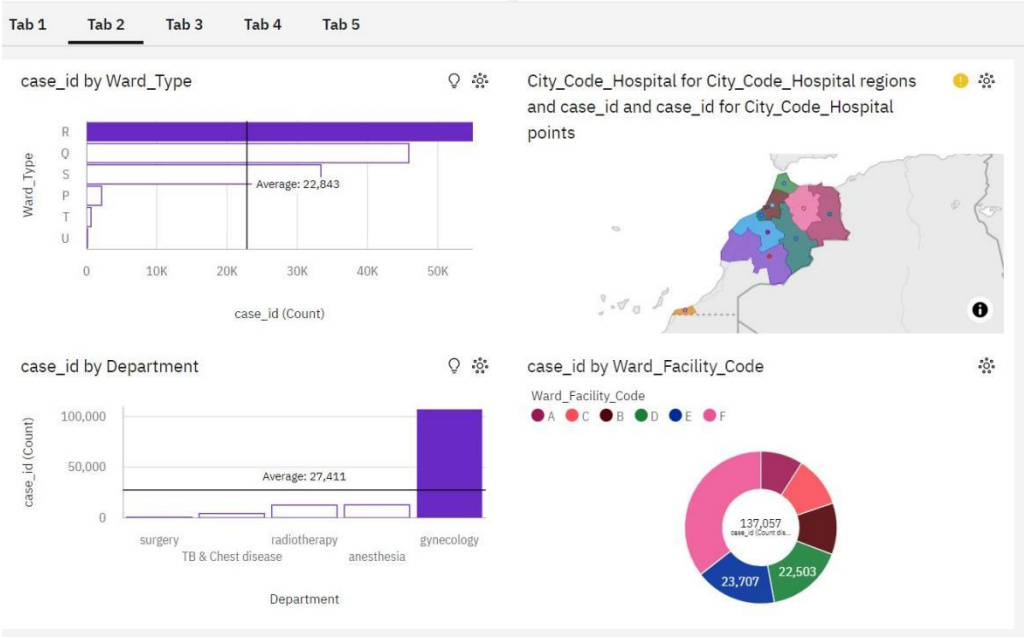
Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
case_id by Ward_Type
Ward_Type
R
Q
S
P
T
U
Average: 22,843
0
10K
20K
30K
40K
50K
case_id (Count)
City_Code_Hospital for City_Code_Hospital regions and case_id and case_id for City_Code_Hospital points
case_id by Department
case_id (Count)
100,000
50,000
0
Average: 27,411
surgery
TB & Chest disease
radiotherapy
anesthesia
gynecology
Department
case_id by Ward_Facility_Code
Ward_Facility_Code
A
C
B
D
E
F
137,057
23,707
22,503

FIG 2.4

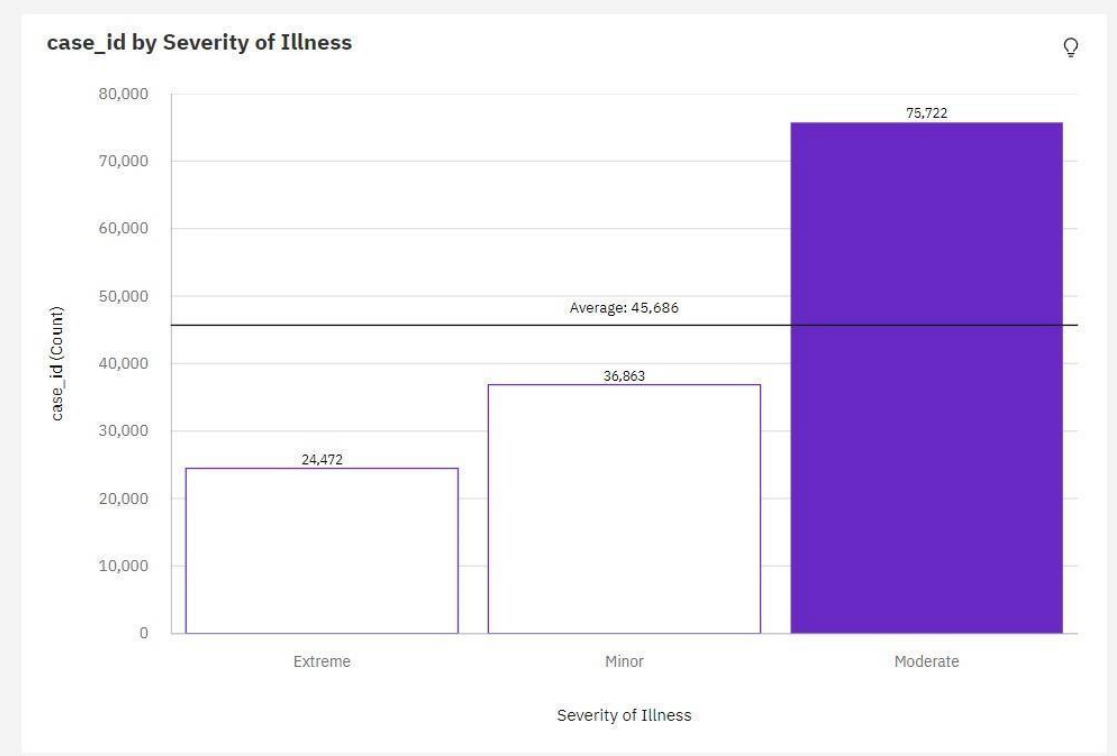
case_id by Severity of Illness
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
case_id (Count)
75,722
Average: 45,686
36,863
24,472
Extreme
Minor
Moderate
Severity of Illness

FIG 2.5

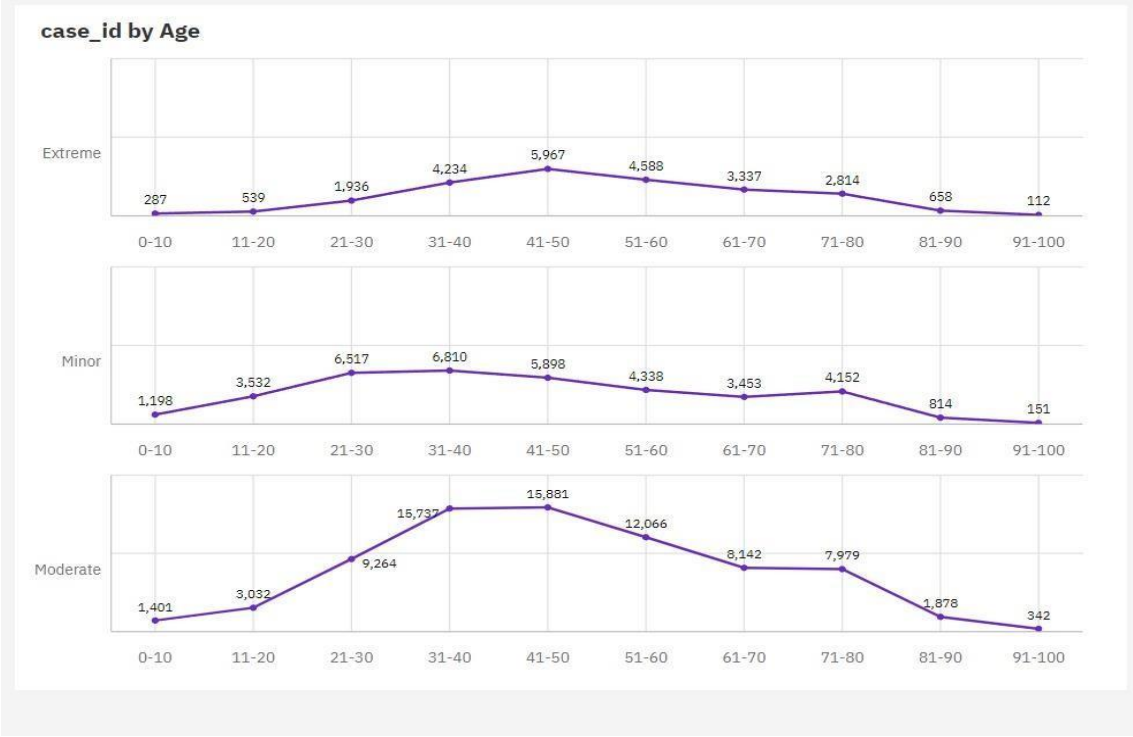
case_id by Age
Extreme
287
539
1,936
4,234
5,967
4,588
3,337
2,814
658
112
Minor
1,198
3,532
6,517
6,810
5,898
4,338
3,453
4,152
814
151
Moderate
1,401
3,032
9,264
15,737
15,881
12,066
8,142
7,979
1,878
342
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100

FIG 2.6

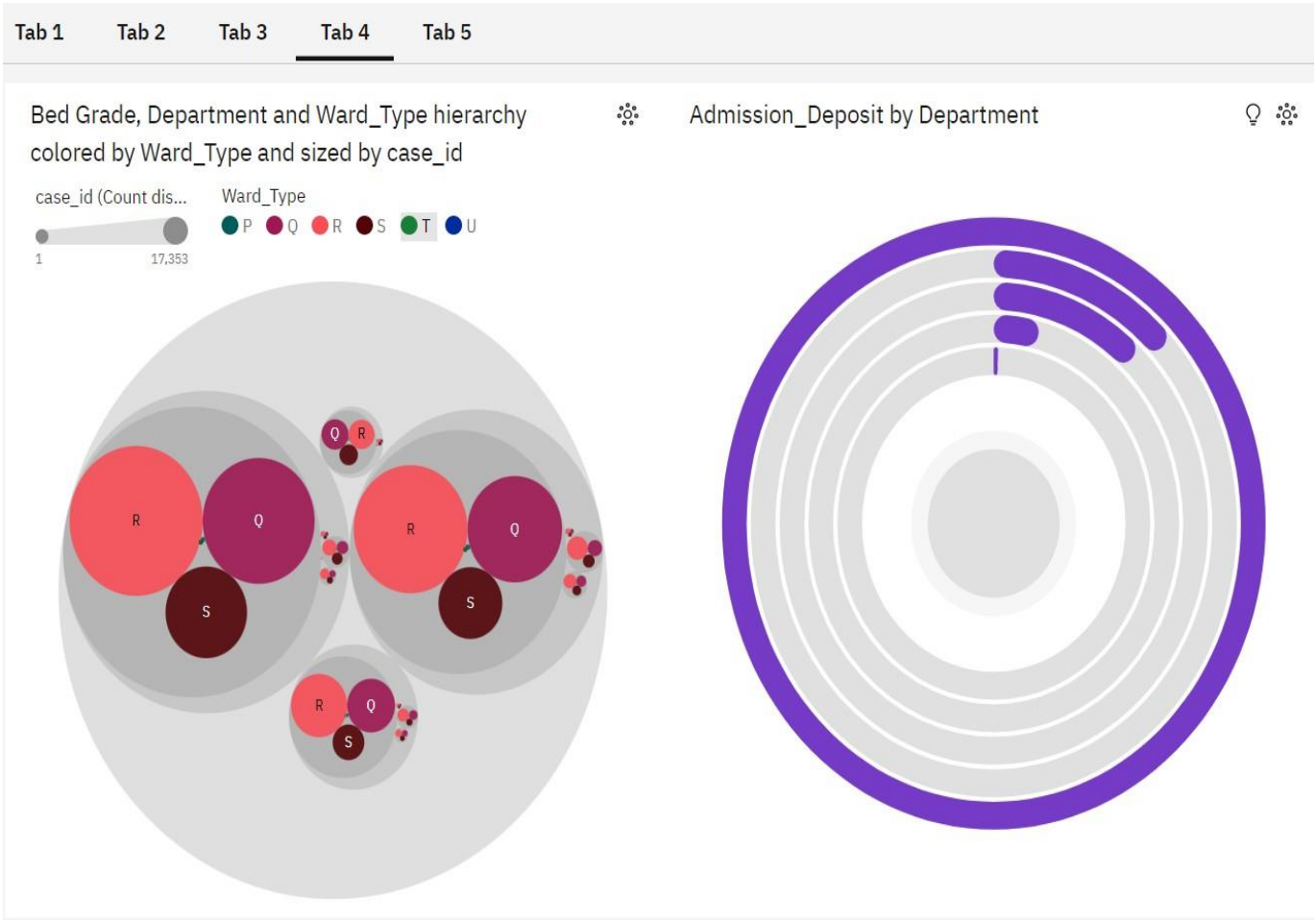
Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
Bed Grade, Department and Ward_Type hierarchy colored by Ward_Type and sized by case_id
case_id (Count dis...
1
17,353
Ward_Type
P
Q
R
S
T
U
Q
R
R
Q
S
R
Q
S
R
Q
S
Admission_Deposit by Department

FIG 2.7

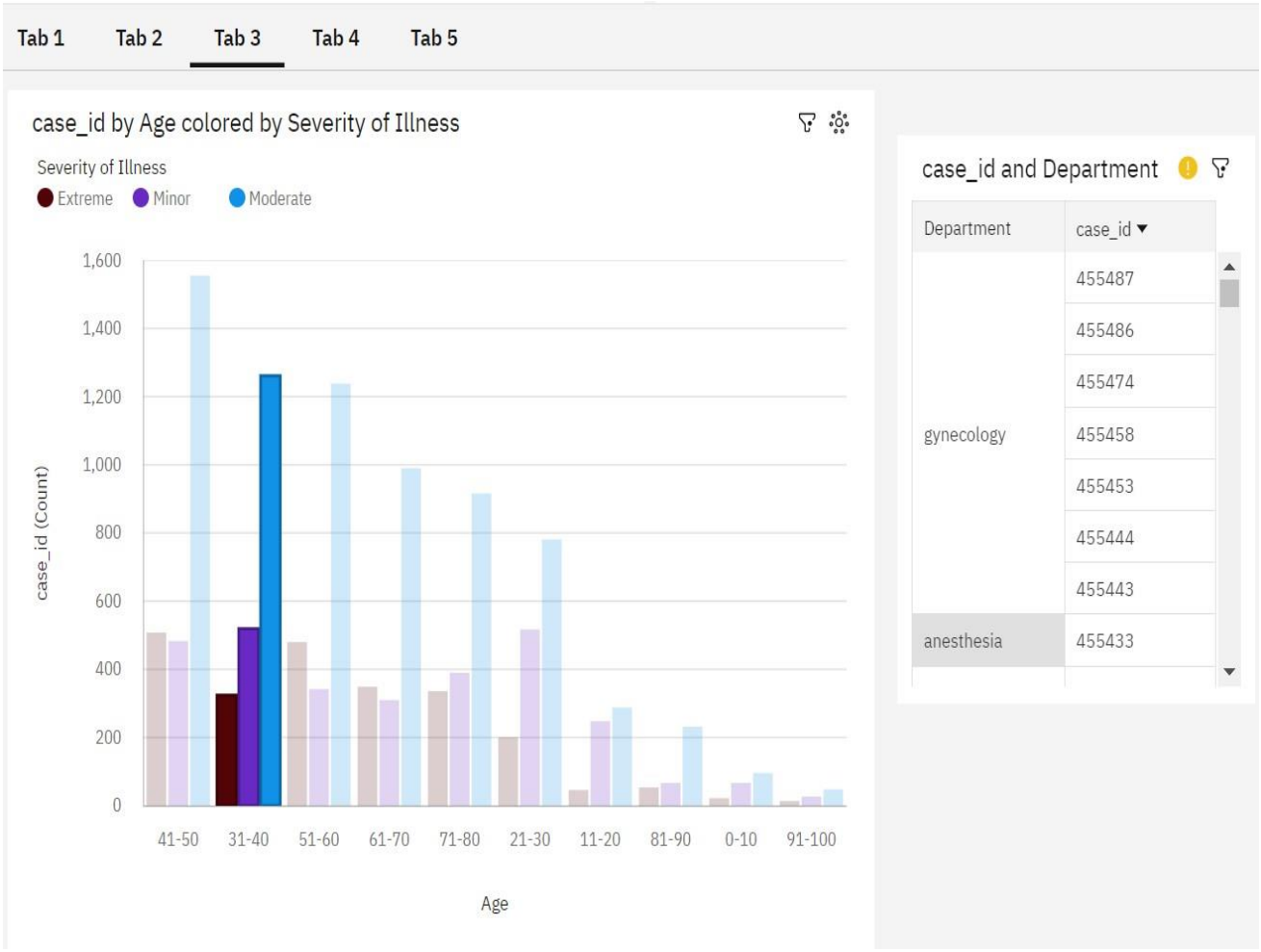
Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
case_id by Age colored by Severity of Illness
Severity of Illness
Extreme
Minor
Moderate
1,600
1,400
1,200
1,000
800
600
400
200
0
case_id (Count)
41-50
31-40
51-60
61-70
71-80
21-30
11-20
81-90
0-10
91-100
Age
case_id and Department
Department
case_id
455487
455486
455474
gynecology
455458
455453
455444
455443
anesthesia
455433

FIG 2.8

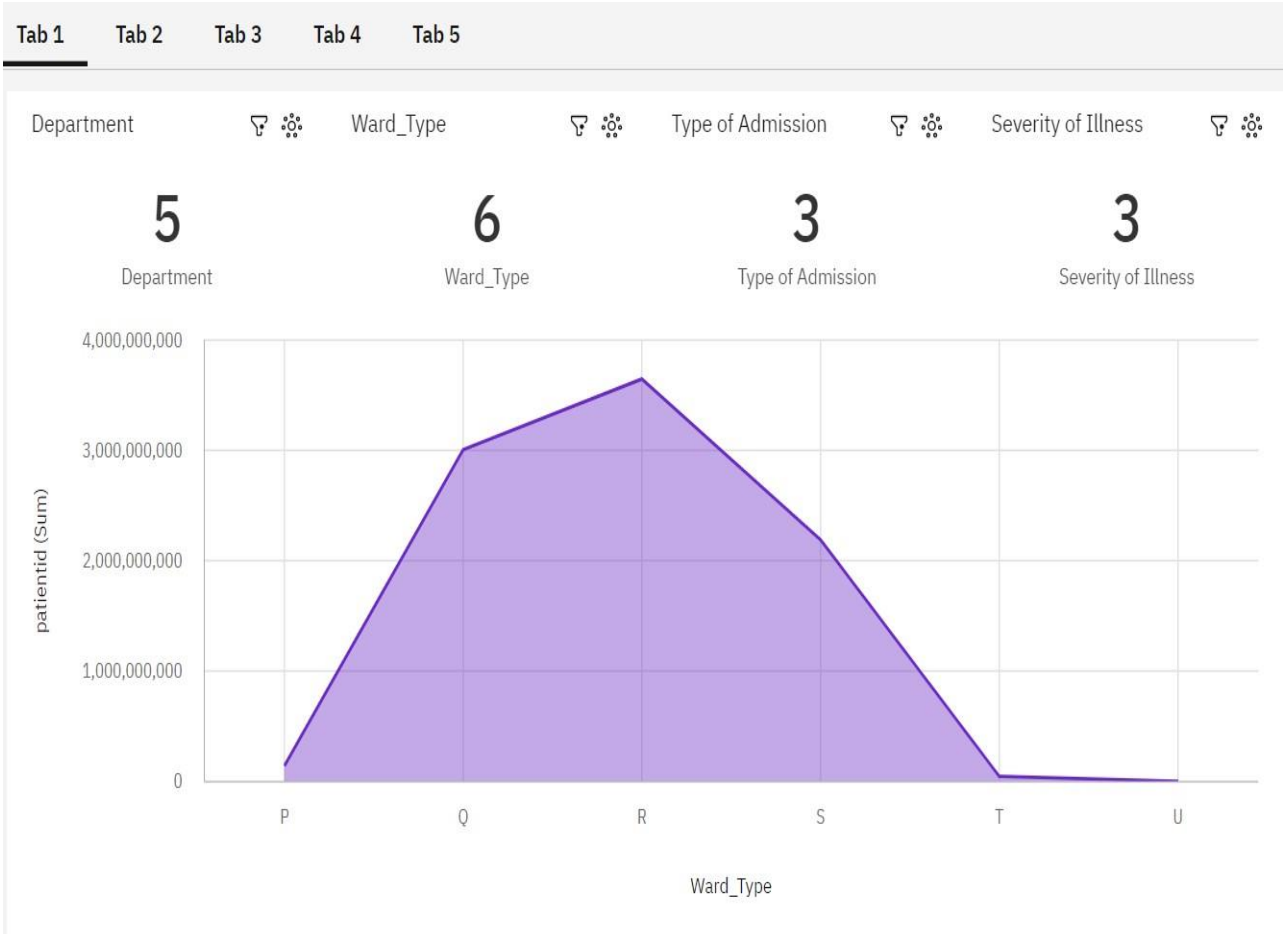
Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
Department
Ward_Type
Type of Admission
Severity of Illness
5
Department
6
Ward_Type
3
Type of Admission
3
Severity of Illness
4,000,000,000
3,000,000,000
2,000,000,000
1,000,000,000
0
patientid (Sum)
P
Q
R
S
T
U
Ward_Type

FIG 2.9

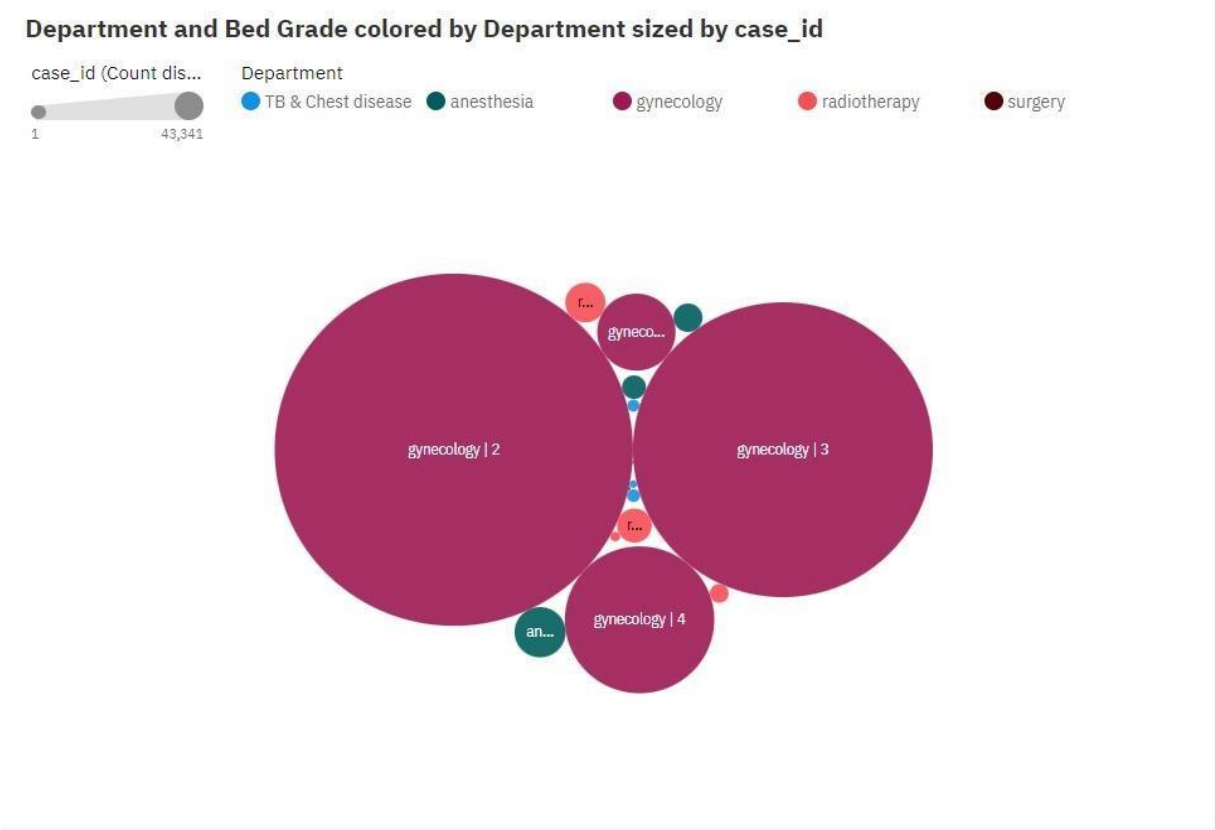
Department and Bed Grade colored by Department sized by case_id
case_id (Count dis...
1
43,341
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
r...
gyneco...
gynecology | 2
gynecology | 3
r...
an...
gynecology | 4

FIG 2.10

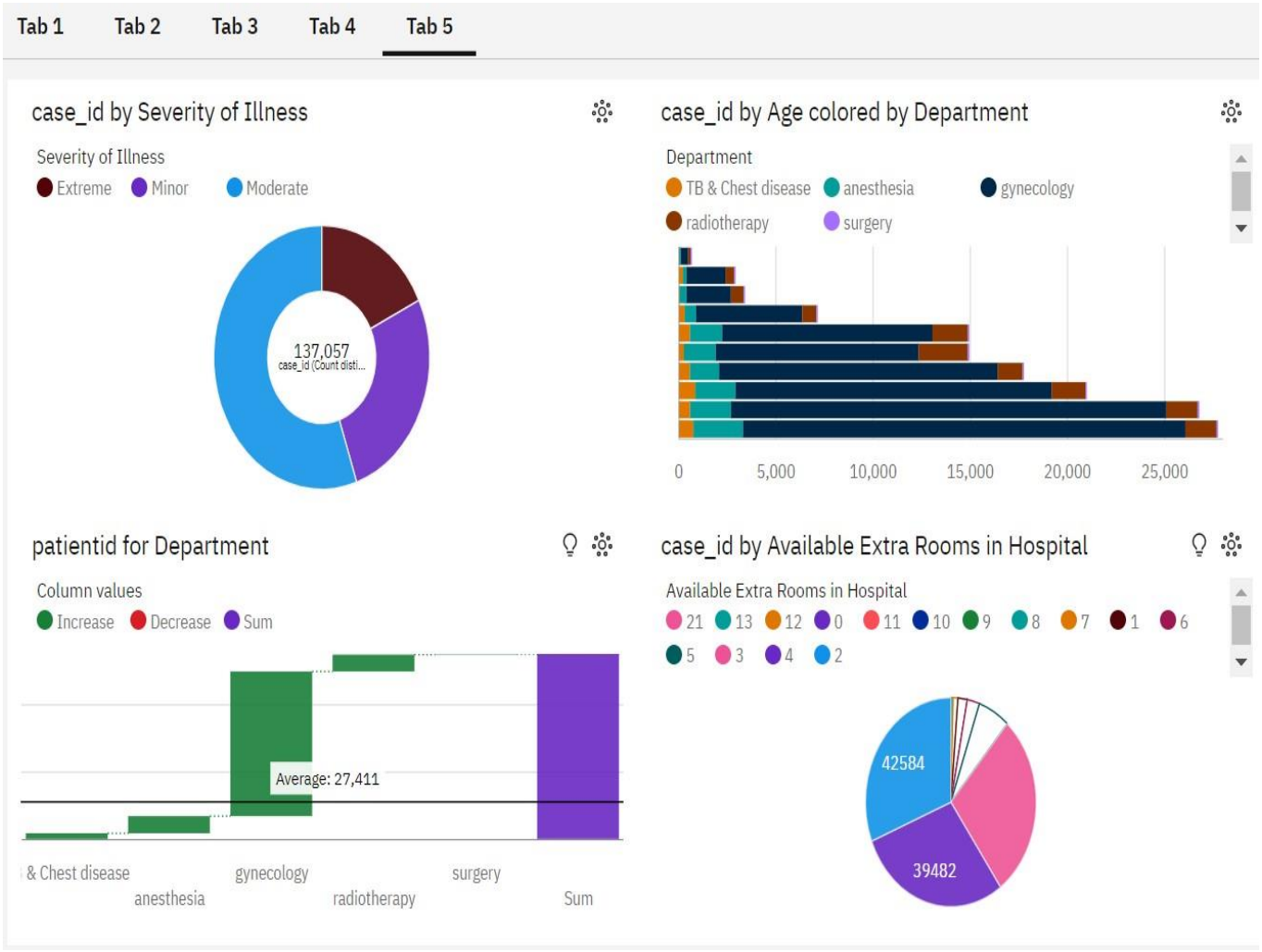
Tab 1
Tab 2
Tab 3
Tab 4
Tab 5
case_id by Severity of Illness
Severity of Illness
Extreme
Minor
Moderate
137,057
case_id (Count disti...
case_id by Age colored by Department
Department
TB & Chest disease
anesthesia
gynecology
radiotherapy
surgery
0
5,000
10,000
15,000
20,000
25,000
patientid for Department
Column values
Increase
Decrease
Sum
Average: 27,411
& Chest disease
anesthesia
gynecology
radiotherapy
surgery
Sum
case_id by Available Extra Rooms in Hospital
Available Extra Rooms in Hospital
21
13
12
0
11
10
9
8
7
1
6
5
3
4
2
42584
39482

FIG 2.11

3.CONCLUSION

The quantitative analysis of the research carried out and presented in this article made it possible to determine whether medical facilities using Big Data Analytics and if so, in which areas. Thanks to the results obtained it was possible to formulate the following conclusions. Medical facilities are working on both structured and unstructured data, which comes from databases, transactions, unstructured content of emails and documents, devices and sensors. According to analytics, they reach for analytics in the administrative and business, as well as in the clinical area. It clearly showed that the decisions made are largely data-driven. The results of the study confirm what has been analyzed in the literature. Medical facilities are moving towards data-based healthcare and its benefits.

Data Analytics has the potential for positive impact and global implications in healthcare. Future research on the use of Big Data in medical facilities will concern the definition of strategies adopted by medical facilities to promote and implement such solutions, as well as the benefits they gain from the use of Big Data analysis and how the perspectives in this area are seen.

4.APPLICATIONS

1. Evaluating and Developing Practitioners.
2. Detecting Anomalies in Scans.
3. Predicting Outbreaks.
4. Clinical Decision Support.
5. Population Health Management.
6. Health Surveillance and Disease Outbreak Management.

5.ADVANTAGES

- Analytics for Health Providers.
- High-risk Inpatient care Patient Satisfaction.
- Improved Health outcomes.
- Risk Assessment and Fraud Detection
- Obtain operational insights.
- Informed strategic planning.
- Improved Decision-Making.

6. DISADVANTAGES

1. Data Privacy and Security Risks.
2. Data Quality and Accuracy.
3. Complexity and Technical Expertise.
4. Ethical Considerations.
5. Overreliance on Technology.

7.FUTURE SCOPE

The healthcare industry is slow to act, but is in a prime position to capitalize on data and analytics insights. The future of healthcare analytics will include even larger collections of data for healthcare organizations to navigate and manage. As new technologies emerge, and consumer demand for control over personal well-being increases, it will be evermore important to understand how best to navigate the competitive landscape and how to scale your data to keep it relevant.

health-care analytics will involve integrating data from diverse sources, including EHRs, wearables, mobile health applications, social determinants of health, and genomic data. By combining and analyzing these multiple data streams, hospitals can gain a comprehensive understanding of patients' health profiles and make more informed decisions.
Natural Language Processing (NLP) and Text Analytics: NLP and text analytics techniques will enable hospitals to extract valuable insights from unstructured clinical data, such as physician notes, research articles, and patient narratives. This will enhance clinical decision support systems, improve medical coding accuracy, and facilitate knowledge discovery from vast amounts of textual information.

8.BIBILOGRAPHY

1. O'Reilly, T. (2017). Data Analytics in Healthcare. O'Reilly Media.
2. Davenport, T. H. (2013). Analytics in Healthcare and the Life Sciences: Strategies, Implementation Methods, and Best Practices. Pearson FT Press.
3. Cresswell, K. M., & Sheikh, A. (2017). Health information technology in hospitals: current issues and future trends. In International Perspectives on Primary Care Research (pp. 171-187). CRC Press.
4. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 36(4), 1165-1188.
5. Pradeep, P., & Mani, K. (2020). Healthcare analytics: An overview, challenges, and opportunities. In Handbook of Research on Machine and Deep Learning Applications for Cyber Security (pp. 245-258). IGI Global.
6. Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144.
7. Wang, F., & Ohno-Machado, L. (2014). Clinical big data analytics: emerging challenges and potential. Journal of the American Medical Informatics Association, 21(2), 173-174.
8. Garg, A. X., Adhikari, N. K., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., ... & Haynes, R. B. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. Jama, 293(10), 1223-1238.
9. Kuo, M. H., & Sahama, T. (2016). Building a framework for analytics-driven decision support systems in healthcare. Health Systems, 5(1), 1-16.
10. Shawky, M., Letaifa, B., & Youssef, M. (2016). Leveraging business intelligence and analytics for healthcare organizations: a literature review. Health Policy and Technology, 5(4), 324-335.

9. HELP FILE

PROJECT EXECUTION:

STEP-1: Go to Start, search and Goto Browser.

STEP-2: After launching Browser type IBM cognos

STEP-3: Click on the official site https://www.ibm.com/products/cognos-analytics

STEP-4: Then goto register.

STEP-5: Fill the Required Details.

STEP-6: Verify your email account.

STEP-7: Finally click on Register.

STEP-8: Goto IBM SEARCH.

STEP-9: Select analysis.

STEP-10: Select the chart or graph you want.

STEP-11: Enter the details accordingly.

STEP-12: View the analysis.