**St Joseph Engineering College**

**An Autonomous Institution**

*Affiliated to VTU-Belagavi, Recognized by AICTE, NAAC Accredited*

# Vamanjoor, Mangaluru-575028, Karnataka



**PRINCIPLES OF DATA SCIENCE**

**DATA ANALYTICS REPORT**

**ON**

# <<Healthcare Dataset>>

**Submitted in partial fulfillment of the requirements for the degree**

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

**(DATA SCIENCE )**

***Submitted by***

<< Tushar J Kotian USN >>4SO22CD061

<< Sudhanva Patil USN >> 4SO22CD055

**Vijetha U**

(Asst. Professor, ICBS Department) Course **Coordinator**

## **ST JOSEPH ENGINEERING COLLEGE**

**An Autonomous Institution**

**(*Affiliated to VTU-Belagavi, Recognized by AICTE, NAAC Accredited*)**

**DEPARTMENT OF INTELLIGENT COMPUTING & BUSINESS SYSTEMS**



# CERTIFICATE

*Certified that the Mini Project Work entitled* **“<<Healthcare>>”** *carried out by*

|  |
| --- |
| << Tushar J Kotian - 4SO22CD061 |
| << Sudhanva Patil – 4SO22CD055 |

bonafide students of V semester in partial fulfillment for the award of Bachelor of Engineering in Computer Science and Engineering (Data Science) of the Visvesvaraya Technological University, Belagavi during the year 2024-2025.

**VIJETHA U**

**Course Coordinator**

Principles of Data Science

22CDS54

# INTRODUCTION

A healthcare dataset that includes both inpatient and outpatient data provides insights into patient care across different settings. Inpatient data covers hospital stays, while outpatient data focuses on treatments or consultations without overnight admission, helping to Analyse healthcare trends and improve service delivery.

# DATASET EXPLORATION

**Detailed Description**:

* **Patient ID** (string): Unique identifier for each patient.
* **Age** (integer): Patient's age, typically ranging from 0 to 100+.
* **Gender** (categorical): Male, Female, or other.
* **Diagnosis** (string): Medical condition diagnosed during the visit.
* **Treatment Type** (categorical): Type of treatment received (e.g., surgery, medication).
* **Admission/Discharge Dates** (date): Dates of hospital admission and discharge for inpatient visits.
* **Visit Type** (categorical): Specifies if the visit was inpatient or outpatient.

**Initial Visualizations**: Include basic plots to show data distributions and correlations, with relevant observations.

# DATA SCIENCE PROCESS

Client Identification

Imagine a client or an authority to whom you would be reporting the results of your analysis.

Ex:

* For healthcare dataset, you would probably be doing analysis for some hospital
* Weather data analysis may be of use for a news channel
* Crime data analysis may be of interest to the police department

Use your imagination to pick a client for your project

Objective

Once you have selected a client try to find the objective of your analysis. The objective would be something that your client will benefit from your analysis.

Ex:

**Assess Patient Care Continuity**: Analyze the transition and coordination of care between inpatient and outpatient settings. This helps identify gaps or inefficiencies in the continuity of care, ensuring that patients receive comprehensive treatment from admission to follow-up consultations.

**Identify Healthcare Utilization Patterns**: Examine how often patients access inpatient vs. outpatient services, highlighting trends, such as frequent readmissions or the use of outpatient services for chronic disease management, helping to optimize resource allocation and patient care strategies.

**Monitor Treatment Outcomes and Effectiveness**: Compare patient outcomes in inpatient and outpatient settings, evaluating the effectiveness of treatments across both environments. This can guide clinical decision-making and the development of tailored treatment plans for different patient needs.

**Enhance Operational Efficiency**: Use the dataset to identify inefficiencies in hospital operations, such as extended inpatient stays or underutilized outpatient services. Optimizing these operations can lead to cost savings, better resource management, and improved patient experiences

**Support Predictive Analytics and Preventive Care**: Leverage the data to build predictive models for patient outcomes, readmissions, and disease progression. By understanding trends in both inpatient and outpatient care, healthcare providers can implement targeted preventive measures to reduce the need for inpatient admissions and improve overall patient health outcomes.

# DATA VISUALIZATION AND ANALYSIS

Summary of Data Visualization and Analysis Techniques for Healthcare Dataset:

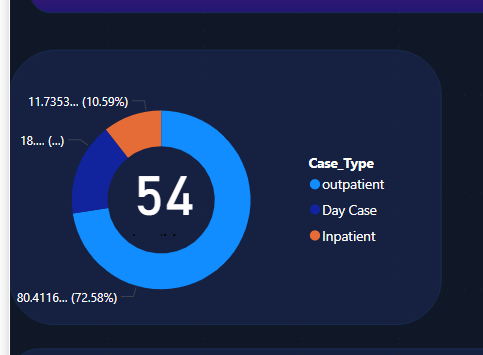
The healthcare dataset that includes both inpatient and outpatient data provides a comprehensive view of patient care across different settings, allowing for the analysis of various metrics and trends. Below is a summary of key visualization methods used to analyses this dataset and gain insights into healthcare trends:

Ex: In health care dataset…

**Q:"Which case type has the highest proportion in the donut chart, and what implications could this have on hospital resource planning? How might trends in case types affect patient care strategies?"**

**1. Donut Chart (Case Type)**

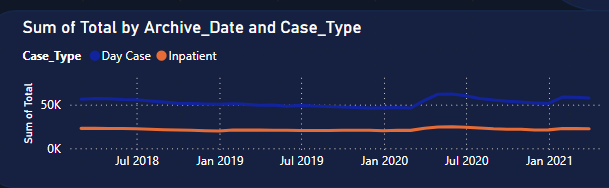
* **Purpose:** The donut chart visualizes the distribution of different case types within the dataset, which could include categories such as inpatient, outpatient, emergency, and elective cases.
* **Insight:** By analysing the proportions of each case type, this visualization helps to understand the relative share of each category within the overall dataset. It highlights trends in patient admissions and treatments across different settings.



**Q:"How do inpatient admissions and resource utilization trends change over time in the line chart? What factors might contribute to seasonal or periodic fluctuations in inpatient demand?"**

**2. Line Chart (Sum of Archive Data and Case Type - Inpatient)**

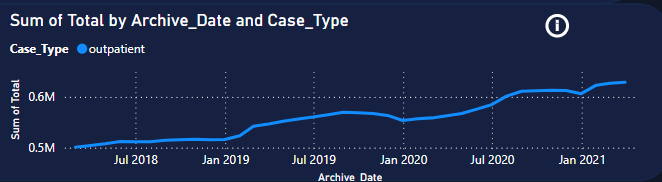
* **Purpose:** This line chart tracks the sum of archived data over time, specifically for inpatient cases.
* **Insight:** By observing this line chart, one can identify trends in inpatient admissions, lengths of stay, and resource utilization over a given time period. It can also reveal seasonal or periodic fluctuations in inpatient demand, helping healthcare providers plan for peak periods.



**Q: "How does the volume of outpatient visits change over time in the line chart? What trends in outpatient care can be observed when compared to inpatient care data?"**

**3. Line Chart** (Sum of Archive Data and Case Type - Outpatient):

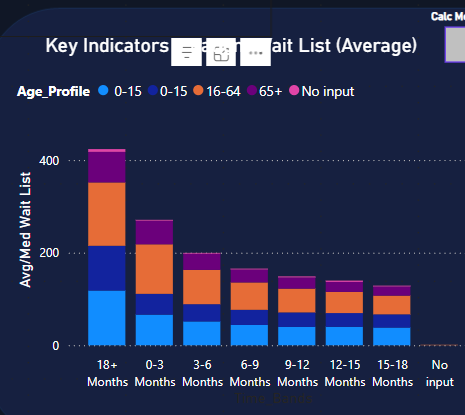
* **Purpose:** Similar to the inpatient line chart, this chart focuses on outpatient data, visualizing the sum of archived data over time for outpatient cases.
* **Insight:** This helps to track the volume of outpatient visits or procedures, highlighting trends in patient flow, outpatient treatment patterns, and the efficiency of outpatient care. Trends in outpatient care volume can be compared to inpatient care trends to see how both segments evolve.



**Q:"How do wait times vary across different case types or departments in the stacked column chart? What insights can be drawn about potential bottlenecks and areas for improvement in patient flow?"**

**4. Stacked Column Chart (Average/Median Wait List):**

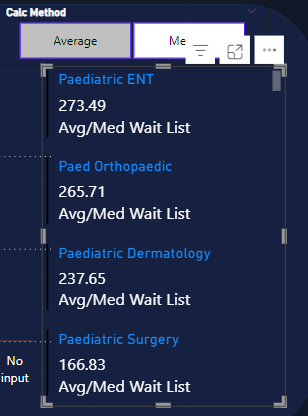
* **Purpose:** This chart shows the distribution of wait times for patients, broken down by different categories (such as case type, department, or severity).
* **Insight:** The stacked columns display average or median wait times across multiple variables, providing insights into patient wait times for both inpatient and outpatient services. Longer wait times could indicate a bottleneck in certain areas, helping administrators identify where improvements are needed.



**Q:"How does the patient wait list vary across different departments and case types in the multi-row chart? What actions can be taken to optimize patient flow and reduce wait times based on this data?"**

**5. Multi-Row Chart (Patient Wait List)**

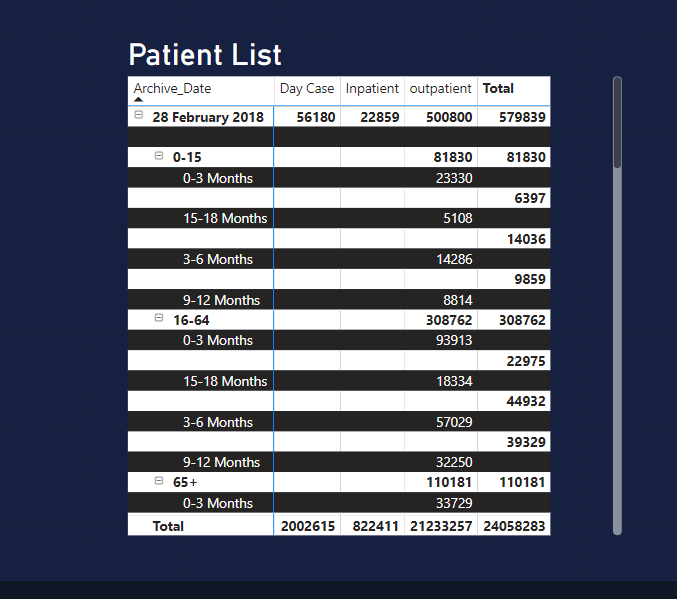
* **Purpose:** The multi-row chart displays the wait list for patients across various dimensions such as the department, case type, or wait time.
* **Insight:** This chart provides a granular view of the patient wait list, which is crucial for managing patient expectations and optimizing scheduling. It helps administrators monitor the flow of patients waiting for services and adjust staffing and resources accordingly.



**Q: "How does the patient volume fluctuate across inpatient and outpatient cases in the matrix? What strategies can be implemented to balance resource allocation based on daily admissions and patient turnover?"**

**6.Matrix (Patient List)**

* **Purpose:** This matrix presents key data about patient admissions, including:
  + Archive date
  + Daily case count (number of cases per day)
  + Inpatient data (number of patients admitted per day)
  + Outpatient data (number of patients released from the hospital per day)
  + Total number of patients in the healthcare dataset
* **Insight:** The matrix enables healthcare administrators and analysts to track patient volume, case load, and patient turnover on a daily basis. This type of matrix is especially useful for managing hospital resources and ensuring that both inpatient and outpatient departments are staffed and equipped to handle the patient load efficiently.



**Summary:**

These visualizations provide valuable insights into different aspects of healthcare operations:

* **Donut and line charts** are used to understand patient demographics, case distribution, and temporal trends (both inpatient and outpatient).
* **Stacked column charts** and **multi-row charts** help track wait times and patient backlog, crucial for improving patient experience.
* **The matrix** offers a comprehensive view of patient flow and can be used to optimize hospital capacity and staffing levels.

Together, these visualizations create a holistic view of the healthcare dataset, allowing administrators and healthcare professionals to make data-driven decisions that improve patient care and operational efficiency.

# CONCLUSION

Write a conclusion on the project – the data science process, your learning experience, useful insights identified etc.

1. **Donut Chart (Case Type):**  
The donut chart provides a clear distribution of patient case types, highlighting the balance between inpatient and outpatient cases. It helps identify which care setting is more prevalent over time.

2. **Line Chart (Sum of Archive Data and Case Type - Inpatient):**  
This line chart reveals trends in inpatient cases, helping to track fluctuations in hospital admissions. It assists in predicting peak periods and managing hospital resources effectively.

3. **Line Chart (Sum of Archive Data and Case Type - Outpatient):**  
The outpatient line chart highlights the volume of outpatient cases, allowing for comparisons with inpatient trends. It aids in understanding outpatient care demand and scheduling needs.

4. **Stacked Column Chart (Average/Median Wait List):**  
The stacked column chart visualizes patient wait times, highlighting areas where delays are most prominent. It is useful for identifying bottlenecks and optimizing patient flow across departments.

5. **Multi-Row Chart (Patient Wait List):**  
The multi-row chart provides a granular view of patient wait times by different categories, helping to manage patient expectations and optimize scheduling for reduced delays.

6. **Matrix (Patient List):**  
The matrix provides a detailed overview of patient admission and discharge patterns by day, facilitating the tracking of patient flow and helping to optimize capacity management across inpatient and outpatient settings.

We have used power bi software the data analyses of the healthcare dataset with the data (Inpatient, Outpatient)as the datasets for the data science project . we have analysed the project using various diagrams from the power bi such as Donut chart, Stacked columns chart, matrix ,line chart , multi row cart etc

Over view the dashboard with the diagram:

