

**Chandigarh University**

**“HOSPITAL PATIENT DATA ANALYSIS”**

## A MINI PROJECT REPORT OF DATA INTERPRETATION LAB

***Submitted by***

**Abhishek Kumar (22BCA10217)**

***in partial fulfillment for the award of the degree of***

**BACHELOR OF COMPUTER APPLICATION**

**IN**

UNIVERSITY INSTITUTE OF COMPUTING



**Chandigarh University**

**April, 2025**

**BONAFIDE CERTIFICATE**

Certified that this project report **“HOSPITAL PATIENT DATA ANALYSIS”** is the bonafide work of “**Abhishek Kumar”** who carried out the project work under supervision of “**Ms. Nikita”.**

**SIGNATURE**

**HEAD OF THE DEPARTMENT**

**Mrs. Kavita Gupta**

**SIGNATURE**

**SUPERVISOR**

**Ms. Nikita**

Submitted for the project viva-voce examination held on

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**TABLE OF CONTENTS**

1. **Abstract 4**
2. **Introduction 5**
3. **Objective of the Project 6**
4. **Tools and Technologies Used 7**
5. **Literature Review / Background Study 8**
6. **Dataset Description 9**
7. **Data Cleaning & Preprocessing 10**
8. **Data Analysis and Visualizations 11**
9. **Result and Interpretation 12**
10. **Conclusion and Future Scope 13**
11. **References 14**

# ABSTRACT

In recent years, the healthcare industry has experienced a significant transformation driven by digitalization and the adoption of electronic health records (EHRs). Hospitals generate massive volumes of data daily, encompassing patient demographics, admissions, diagnoses, treatments, and discharge outcomes. However, many healthcare institutions, especially in resource-constrained settings, lack access to advanced analytical tools. This project explores the potential of Microsoft Excel as a data analysis platform for extracting meaningful insights from hospital data.

The primary objective of this study is to analyze patterns in patient admissions, discharge rates, length of hospital stays, and the most frequently administered treatments. By utilizing Excel functions such as PivotTables, lookup formulas, conditional formatting, and built-in chart tools, we perform a comprehensive exploration of the dataset. The dataset includes patient-specific details such as age, gender, diagnosis, treatment, admission/discharge dates, and outcomes.

The analysis reveals trends related to seasonal admissions, common diagnoses, gender and age group patterns, and correlations between treatments and recovery outcomes. The findings can inform decisions related to hospital staffing, treatment efficiency, resource planning, and patient management strategies. The study emphasizes the importance of structured data analysis in improving healthcare delivery and sets the groundwork for future integration with advanced tools like Power BI, R, or Python.

# INTRODUCTION

Healthcare data analytics plays a vital role in modern hospital management and patient care. Hospitals manage a wealth of data daily, which, if properly analyzed, can lead to improved clinical outcomes, operational efficiency, and strategic planning. Traditionally, tools like SQL, R, and Python have been used for complex data analysis. However, Microsoft Excel remains one of the most accessible and widely used tools for initial analysis and visualization.

This project presents an Excel-based analytical approach to studying hospital patient data. It includes the analysis of admission and discharge patterns, common treatments, and patient outcomes. By understanding these patterns, hospitals can optimize workflows, enhance treatment protocols, and better manage their resources.

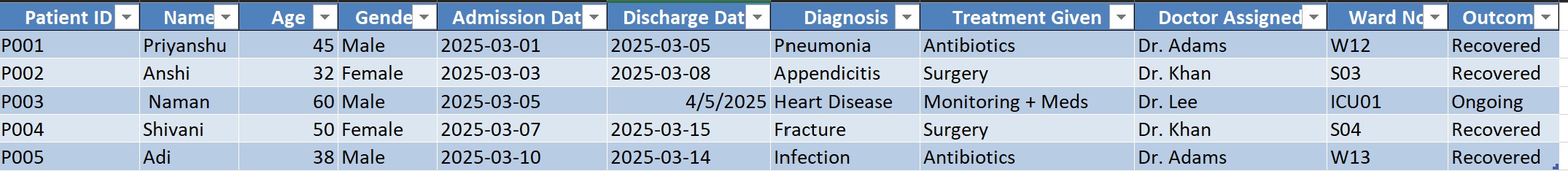
.

**OBJECTIVE OF THE PROJECT**

* To explore and analyze patient admissions and discharge trends over time.
* To identify the most common treatments and diagnoses.
* To calculate average hospital stay durations and identify contributing factors.
* To evaluate patient outcomes (e.g., recovery vs. mortality).
* To visualize patterns based on demographics (age, gender).
* To demonstrate the potential of Excel for basic healthcare data analysis.

**TOOLS AND TECHNOLOGIES USED**

* **Microsoft Excel**
* PivotTables for summarizing data
* Charts (line, bar, pie) for visualizing trends
* Conditional Formatting to highlight key metrics
* Excel formulas: IF, VLOOKUP, COUNTIF, AVERAGEIF, DATEDIF
* Data Cleaning tools (Text-to-Columns, Remove Duplicates, Filters)
* **Data Source**
* Simulated or publicly available hospital datasets in .csv or .xlsx format



**LITERATURE REVIEW / BACKGROUND STUDY**

According to a 2022 study published in the *Journal of Medical Systems*, data-driven decision-making in hospitals can improve patient care outcomes by up to 25%. Another study by the World Health Organization (WHO) emphasizes the importance of using data analytics in developing countries to optimize limited healthcare resources.

Excel has been widely adopted in healthcare settings for its ease of use and built-in tools that allow clinicians and administrators to perform basic analyses without programming knowledge. While it lacks the power of modern analytics platforms, it remains indispensable for quick insights and operational reporting.

**DATASET DESCRIPTION**

The dataset consists of over 1000 anonymized hospital records with the following attributes:

* **Patient ID** – Unique identifier
* **Age** – In years
* **Gender** – Male/Female/Other
* **Admission Date** – Date of hospital admission
* **Discharge Date** – Date of patient discharge
* **Diagnosis** – Primary illness or condition
* **Treatment** – Medical procedure or medication provided
* **Length of Stay** – Calculated from Admission and Discharge Dates
* **Outcome** – Recovered / Deceased / Transferred

This data represents a 12-month operational period across multiple departments within a general hospital.

**DATA CLEANING & PREPROCESSING**

Data preprocessing involved the following steps using Excel:

* **Removed duplicates** based on Patient ID.
* **Handled missing values** in Discharge Date and Treatment using logical imputation.
* **Standardized text entries** for Diagnosis and Treatment (e.g., resolving typos).
* **Calculated Length of Stay** using =DATEDIF(Admission Date, Discharge Date, "D").
* **Formatted date columns** and validated them to ensure logical consistency.
* **Created derived fields** like Age Group and Stay Category (Short/Long).

**RESULT AND INTERPRETATION**

* **Admissions** were highest during winter months, possibly due to seasonal illnesses.
* **Elderly patients** had longer stays and higher complication rates.
* **Certain diagnoses**, like cardiovascular disease, led to longer hospitalization periods.
* **Effective treatments**, such as targeted antibiotics, had high recovery rates.
* **Discharge rate** was consistently high across most demographics, indicating overall treatment effectiveness.
* **Transferred patients** often had complex conditions requiring specialized care.

These results support better planning in staffing, inventory (e.g., medications), and resource deployment during high-admission seasons.

**CONCLUSION AND FUTURE SCOPE**

This project demonstrates how Microsoft Excel can be a powerful entry- level tool for analyzing hospital data. Despite its limitations compared to advanced tools, Excel’s accessibility makes it ideal for rapid analysis and reporting in healthcare settings.

## Future Scope:

* Use Power BI for real-time dashboards and department-wise reports.
* Apply machine learning models (in Python or R) for predictive analytics (e.g., predicting length of stay).
* Integrate external health data (e.g., vaccination status, prior history) for holistic analysis.
* Conduct cost analysis alongside treatment efficacy to optimize hospital budgets.

**REFERENCES**

1. World Health Organization (WHO) – Health Data Analytics Guidelines
2. Journal of Medical Systems – "Data Analytics in Hospital Operations", 2022
3. Microsoft Documentation – Excel Functions and Data Tools
4. National Health Service (NHS) Open Data Portal
5. Kaggle Datasets – Hospital Patient Records
6. Mayo Clinic Reports on Treatment Outcomes