

# Title

## Hospitalist Dashboard

*A Data-Driven Tool for Analyzing Patient Flow, Outcomes, and Readmission Risk*

Reshmi Deb

BMI 8540 – Bioinformatics Informatics

Spring 2025

University of Nebraska Omaha

**Email:** rdeb@unomaha.edu

## ABSTRACT/PURPOSE

The interactive dashboard is developed with the goal of helping the clinical administrators and hospitalists to improve patient care. The dashboard presents understandable overview of patient demographics, length of stay, discharge outcomes, and readmission risks trends, created with the patient data that is similar to EHR-derived csv files. The project facilitates real-time monitoring to support data-driven decision making by combining Bash, MySQL and Python based visualization tool (streamlit). The dashboard provides trends by age group and gender, highlighting patterns in hospital stay durations and outcome distributions that can be used to improve discharge planning and reduce readmission rates.

## PROJECT OBJECTIVES / GOALS

1. Developed a pipeline that can transform, load and analyze the patient data into a structured database.
2. Built a reproducible Bash-based workflow that runs the pipeline and launches the dashboard.
3. Designed a Streamlit dashboard with python and SQL that visualizes patient statistics, filtered by age group and time range.

## BACKGROUND

The hospitals unplanned readmissions within a few weeks of discharge have been a concern among hospitals due to increasing cost and their impact on patient outcomes. Hence, hospitalists need a timely and easy access to the patient's data to improve discharge planning, resource allocation, and care coordination.

The lack of tools to analyze patient-level summary data in a user-friendly interface despite widespread utilization of EHR systems, especially at smaller clinics may hinder timely decision-making. This dashboard bridges that gap using open-source tools.

While commercial systems like Epic and Cerner offer built-in analytics, this project offers a lightweight, modular, and fully open-source alternative tailored for teaching and small research projects. Unlike general-purpose dashboards, this one focuses on key metrics important to hospitalists and supports data filtering by patient age group and date.

## PROJECT COMPONENTS

Database Implementation: `process_data.py`, `load_data.py`

Bash Shell Script: `run.sh`

Python Visualization: `app.py` (Streamlit dashboard)

## DOCUMENTATION

The README file includes setup instructions, dependencies, and usage examples.

All scripts include comments explaining key logic and functions.

Reproducibility:

1. Bash script automates the full process from raw CSV import to dashboard launch.
2. Python scripts can be run via virtual environments.
3. `requirements.txt` auto-generates required Python libraries using `pipreqs`.

License: Licensed under MIT License.

## DATA PROVENANCE

The data used in this dashboard are de-identified patient datasets that were generated using AI tools for educational and demonstration purposes only.

Simulated patient data csv files stored in input folder (person.csv, encounter.csv, outcome.csv).

These were prepared by the student based on publicly available schemas (e.g., CMS Synthetic Data, Synthea).

No real patient data is used or shared.

## USERS

Hospitalists and Clinicians can use the dashboard to monitor patient trends and outcomes.

Healthcare Analysts can extract and interpret patterns in admission/discharge data.

Biomedical Informatics Students to understand database design, data visualization, and ETL pipelines.

## IMPLEMENTATION CONSTRAINTS

Lack of access to real EHR data; used simulated/synthetic data instead.

Local MySQL setup was required due to institutional database access limitations.

Some features (e.g., real-time data updates or authentication) were excluded to maintain simplicity.

## REFERENCES

CMS Synthetic Data. <https://data.cms.gov/>.

Synthea: Synthetic Patient Generator. <https://synthea.mitre.org/>.

Streamlit Documentation. <https://docs.streamlit.io/>.

MySQL Reference Manual. <https://dev.mysql.com/doc/>.

Pandas Documentation. <https://pandas.pydata.org/docs/>.

Altair for Declarative Visualization. <https://altair-viz.github.io/>.

CDC Hospital Readmissions Facts. <https://www.cdc.gov/nchs/data/databriefs/db305.pdf>.

HealthIT.gov Interoperability Standards. <https://www.healthit.gov/topic/interoperability>.

## **PRIVACY STATEMENT**

This project does not include any real patient data. It is designed strictly for educational purposes and uses only synthetic or anonymized datasets.

## **ORIGINALITY STATEMENT**

All code, documentation, and workflows were created by the student. No copyrighted or proprietary code is included. Any AI tools used (e.g., ChatGPT) were used for editing or code refinement, not to generate full deliverables.