Project Overview

This project involves analyzing healthcare data to gain insights into patient encounters, procedures, and costs. The project utilizes Python for data preprocessing, Power BI for visualization, and Streamlit for creating an interactive web application.

Data Sources

The project uses the following CSV files:

1. patients.csv
2. encounters.csv
3. procedures.csv
4. organizations.csv
5. payers.csv

Step-by-Step Process

1. Data Collection

- Obtained the healthcare dataset containing information about patients, encounters, procedures, organizations, and payers.

- Ensured all CSV files were placed in the project directory.

2. Data Preprocessing

- Used Python and pandas for data cleaning and preprocessing.

- Steps included:

  a. Handling missing values

  b. Correcting data types (e.g., converting date strings to datetime objects)

  c. Removing duplicates

d. Standardizing column names

e. Merging relevant datasets for analysis

3. Exploratory Data Analysis (EDA)

Conducted initial data exploration to understand the structure and content of the datasets.

Used pandas to generate summary statistics and identify patterns or anomalies in the data.

4. Power BI Visualization

Imported the preprocessed data into Power BI.

Created a data model by establishing relationships between tables.

Designed various visualizations, including:

a. Dashboard overview of key metrics

b. Patient demographics analysis

c. Encounter trends over time

d. Cost analysis by procedure type and organization

e. Payer contribution analysis

5. Streamlit Web Application Development

Created a Streamlit app to provide an interactive interface for exploring patient data.

Implemented features such as:

a. Patient selection by name or ID

b. Display of basic patient information

c. List of patient encounters and procedures

d. Interactive filters and visualizations

6. Testing and Refinement

Tested both the Power BI dashboard and Streamlit app with sample queries and use cases.

Refined visualizations and interactions based on usability testing.

Ensured data accuracy and consistency across both platforms.