Data: [Synthetic Patient Care](https://www.kaggle.com/datasets/prasad22/healthcare-dataset?resource=download) Contains  
  
Data Cleaning:

After downloading the CSV file, I opened it in excel and began the data cleaning process doing the following steps:

Removed Duplicates(534 Rows Removed)

Changed the column names to be fitting for SQL import removing spaces

Changed the ‘DateOfAdmission’ and ‘DischargeDate’ column to the data type of date and formatted to SQL standards for importing.

Adjust the ‘BillingAmount’ column to the currency data type

Adjusted the ‘Name’ column for PROPER casing

Importing the Data

Created a new database for this project in addition to creating a new table called ‘Patient\_Data’

Loaded the CSV by using the following query

A computer code with text

Description automatically generated with medium confidence

Tested the import to verify it loaded successfully with the following query

A close up of a computer screen

Description automatically generated

Analysis

To conduct analysis on this dataset I decided to on 5 questions that can be answered with queries.

1. What is the average age of the patients?

A close up of a text

Description automatically generated A screenshot of a computer

Description automatically generated

Using this query, I found the average age of the patients was 51 years old.

\*Continued on next page\*

1. What is the count of patients based on gender?

A white background with black text

Description automatically generated A screenshot of a computer

Description automatically generated

Based on the results from this query there were 27,496 males and 27,470 female patients.

1. What months tend to have more patients?

A white background with black text

Description automatically generated A screenshot of a computer

Description automatically generated

According to the table the top 3 months that have the most patients are August, July, and January, respectively.

1. What are the most commonly prescribed medications?

A white background with black text

Description automatically generated A screenshot of a computer

Description automatically generated

The medication called Lipitor is the most commonly prescribed for patients.

1. What is the average stay duration?



A screenshot of a computer

Description automatically generated

The average stay duration for patients is roughly 15 days.