

Data Visualizations of MIMIC-III

Cole Stokes

AI in Healthcare

The University of Texas at Austin

[Google Colab](#) - [GitHub](#)

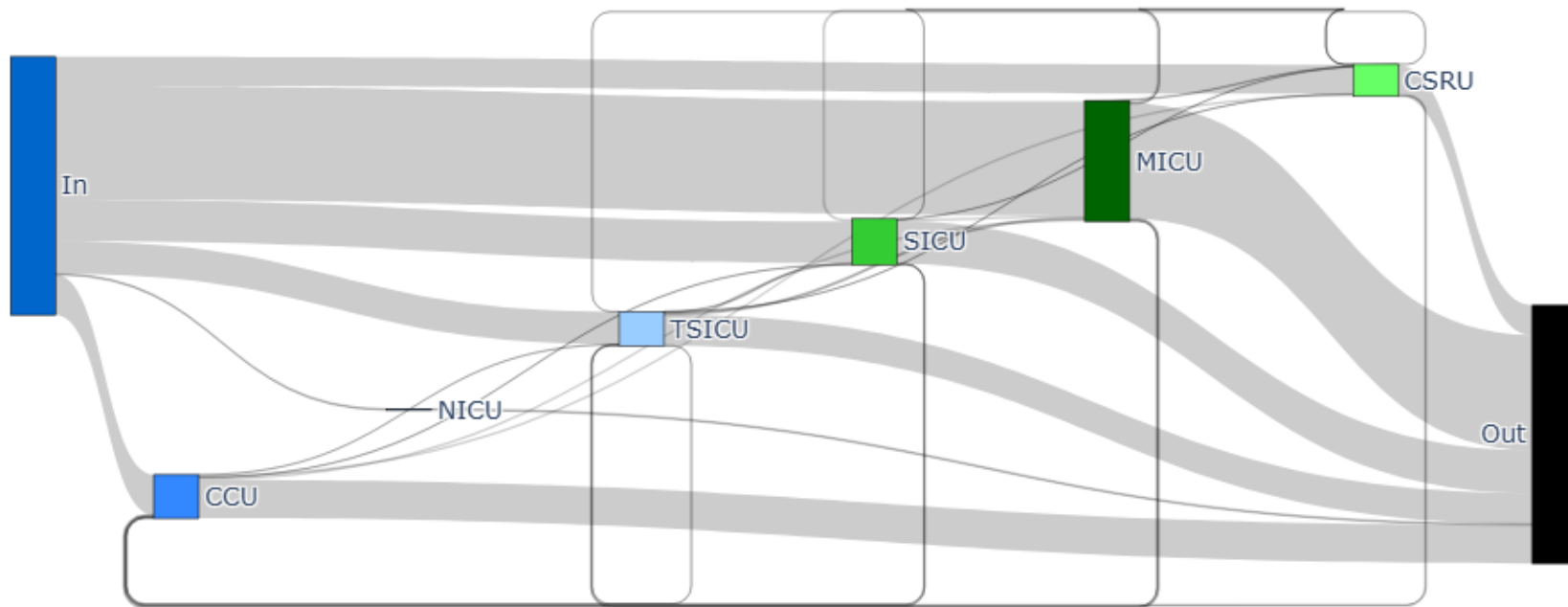
Setup

- ▶ Used a Google Colab Jupyter Notebook and Python.
- ▶ Loaded the MIMIC-III data through Google BigQuery.
- ▶ Wrote SQL commands to extract data.
- ▶ Used Pandas, NumPy, and Itertools for data processing.
- ▶ Created visualizations using Matplotlib, Seaborn, Plotly, and Holoviews.



Visualization 1

MIMIC-III Sankey Flow Diagram for Emergency-admitted Patients into ICU Care Units



How I created this:

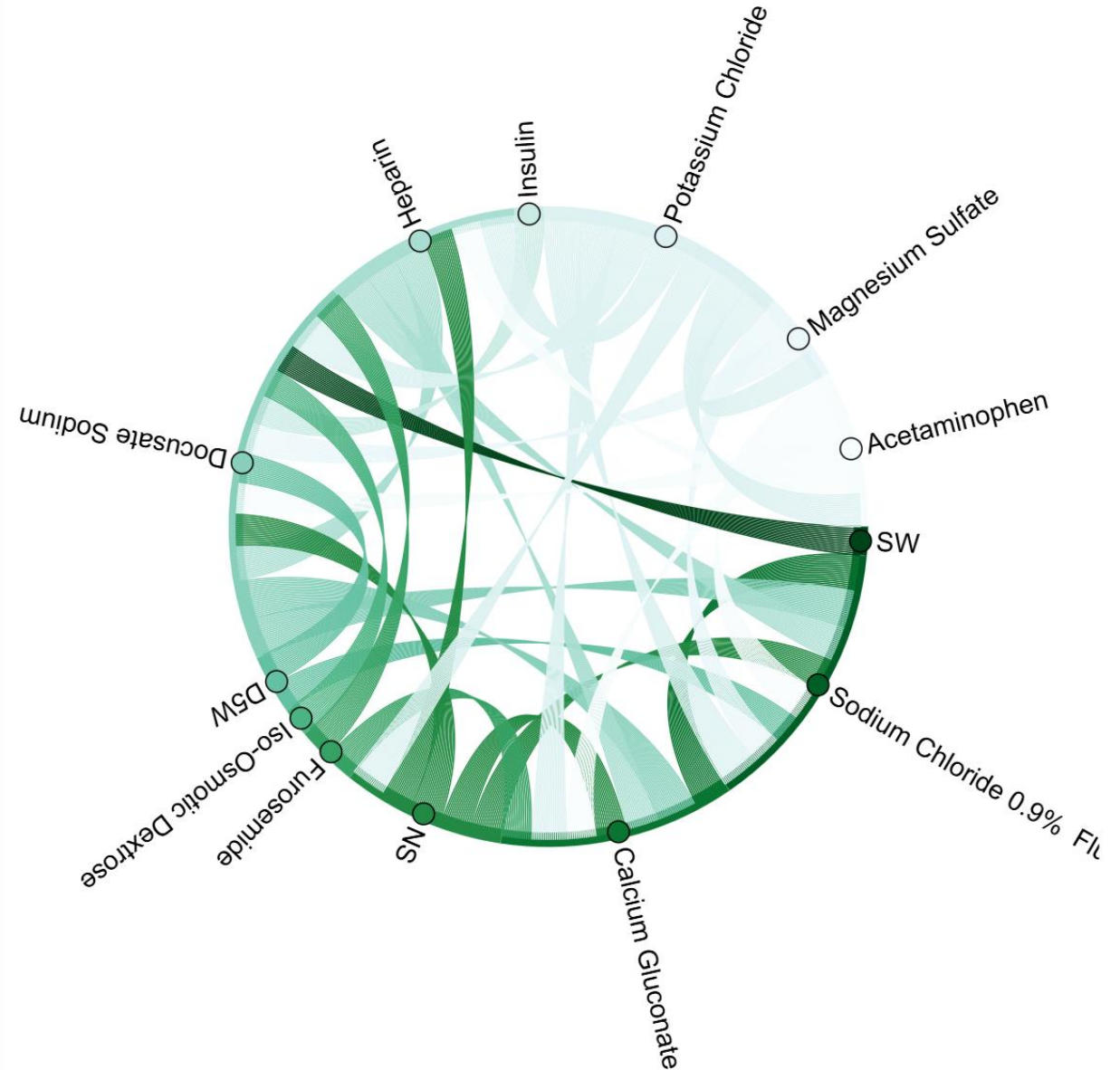
- ▶ I queried the dataset to join the admissions and icustays tables on hadm_id and filtered the data only to include emergency admissions. The IDs along with the first and last care unit information were pulled.
- ▶ Nodes for the diagram were defined as in (before hospitalization), out (when they leave), and the ICU care units. A patient flow dataframe was created showing flows between the first and last ICU unit. I then looped through the flows and added the in and outflows from the first and last ICU units. I converted each node to a numerical code and removed flows between the same unit as this indicates they stayed put.
- ▶ I used Plotly to create the Sankey flow diagram and manually adjusted colors and positioning.

Visualization 2

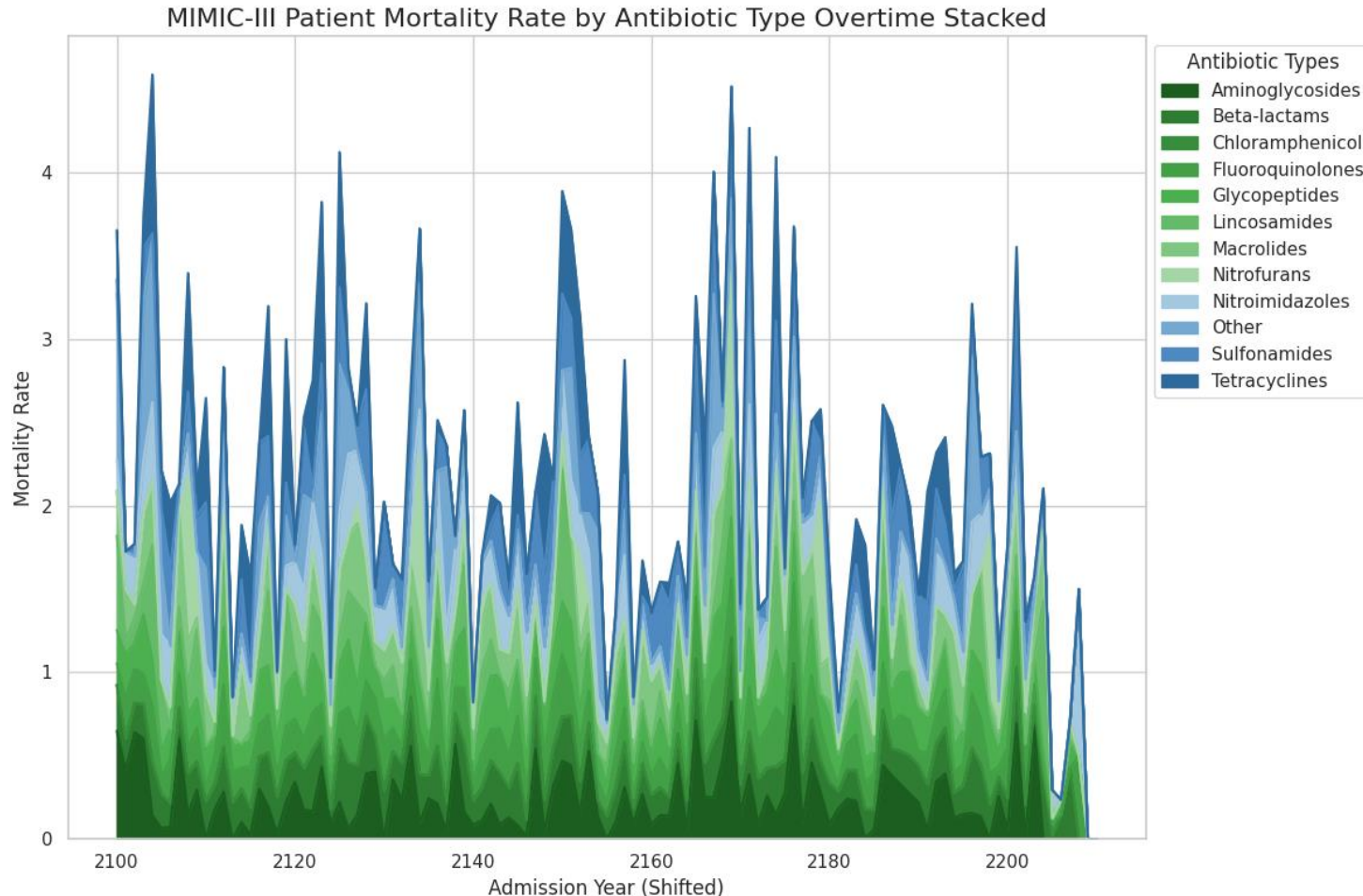
How I created this:

- ▶ In my SQL query to BigQuery, I pulled patient drug prescription data from the prescriptions table.
- ▶ Next, I created a pairs dataframe consisting of all co-occurrences of drug prescriptions to patients. A loop was then used to see what drugs had the highest co-occurrences.
- ▶ Then the drugs nodes and co-occurrence data were mapped into a Holoviews dataset.
- ▶ The Holoviews library was then used to plot a chord diagram showing which drugs are most commonly prescribed together.

Chord Diagram of the Top 30 Co-occurrences of Drug Prescriptions in MIMIC-III



Visualization 3



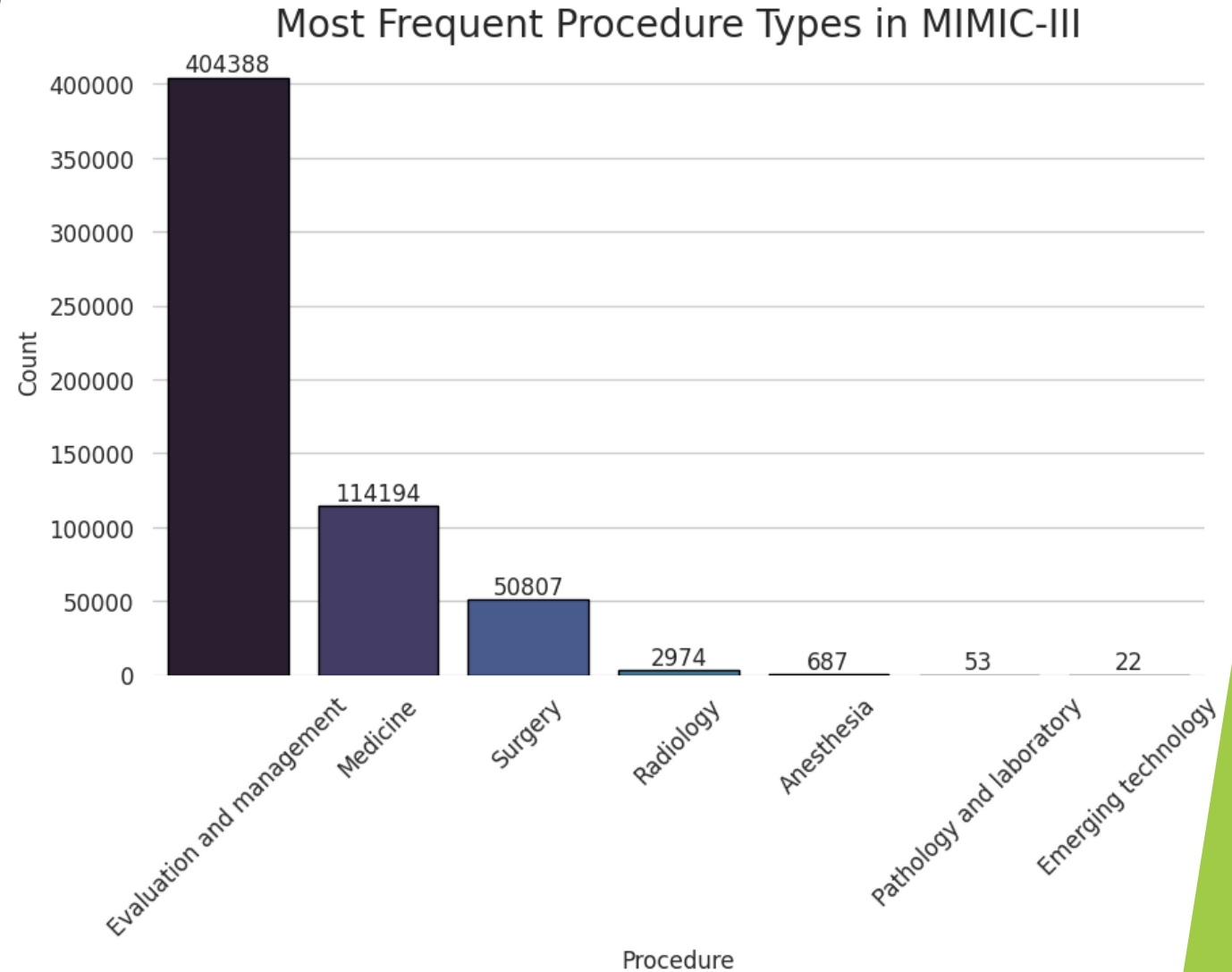
How I created this:

- ▶ From BigQuery I selected the admission year, antibiotic used, and their mortality indicator.
- ▶ I had to join the admissions and the suspicion_of_infection tables through the icustays table because icustays had both icustays_id and hadm_id.
- ▶ I then searched for groups of antibiotic types and manually created a dictionary to map each antibiotic to a group.
- ▶ Next, I aggregated the data to get the mortality rate based on year and antibiotic type.
- ▶ Matplotlib was used to create this stacked area plot.

Visualization 4

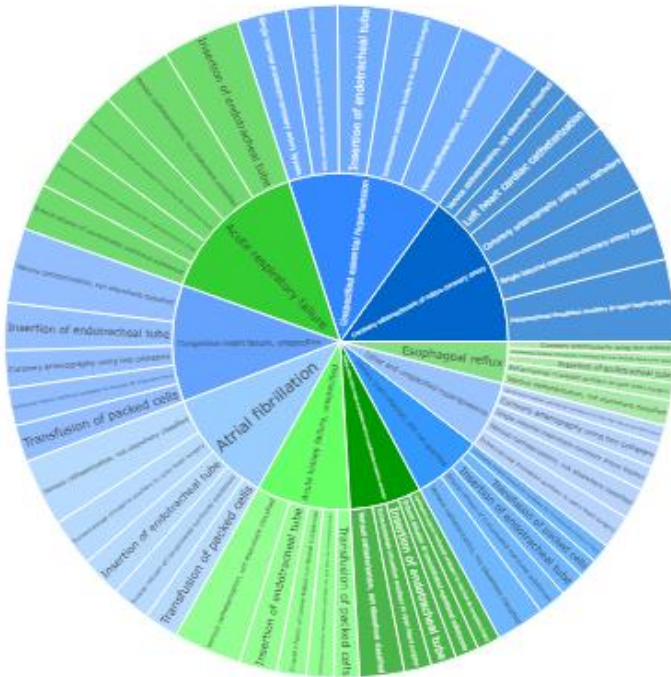
How I created this:

- I selected the section headers and their counts from the cpevents table.
- I then plotted a bar chart of these counts for each procedure type using Matplotlib and Seaborn.
- This can provide the hospital with insights on what type of procedures they perform most often.



Visualization 5

Sunburst Chart of Most Common Diagnoses and Procedures from MIMIC-III



How I created this:

- ▶ In my SQL query to BigQuery, I defined the top diagnoses by the number of occurrences in the diagnosis_icd table. I limited it to the top ten.
- ▶ I defined the top procedures by joining the diagnosis_icd and procedures_icd tables on subject_id and hadm_id and ranking the procedures based on the most common for each top diagnosis.
- ▶ I then joined the top procedures and diagnosis tables with the d_icd_diagnoses and d_icd_procedures tables on the icd9 codes to get the English names for them. I filtered to only include the top five procedures for each and sorted them by most common diagnosis and then procedure.
- ▶ Lastly, using the counts and English terminology, I used Plotly to generate a Sunburst Chart.
- ▶ This gives us information on the most common diagnoses and the most common way to treat them.