Rotman

Individual Assignment Cover Sheet

Course Code:	RSM8431 Analytics Colloquia	Date:	2026/1/16
Course Title:	Process Mining	**Student Number:**	1011886924
Instructor Name:	Arik Senderovich		
Assignment Title:	MMA_Project_2026		

Academic Integrity Compliance

In submitting this assignment, I affirm the work represents entirely my own efforts. I accept and understand the consequences of violating the University of Toronto's Academic Integrity policies as outlined in the

Code of Behaviour on Academic Matters.

I confirm that:

- I have followed the instructions for the assignment, including any specific formatting requirements.
- My work is original. Due credit is given where appropriate and I have acknowledged the ideas, research, phrases etc. of others with accurate and proper citations.
- I have kept my work to myself and did not share answers/content with others, unless otherwise directed by my instructor.
- Any proofreading by another was limited to indicating areas of concern which I then corrected myself.
- This is the first time I have submitted this assignment (in whole or in part) for credit.
- This is the final version of my assignment and not a draft.
- I am submitting this work for the correct course, via the specified platform/method (e.g. Quercus).

I agree that the statements above are true. If I have concerns, I will consult the course instructor immediately.

Optional: you may wish to follow the standard naming convention when saving files:

Complete Course Code (including Section) – STUDENT NUMBER – **Assignment Title**

Example: **RSM1234HF.2021-0108**-0123456789-**Homework1**

Rotman

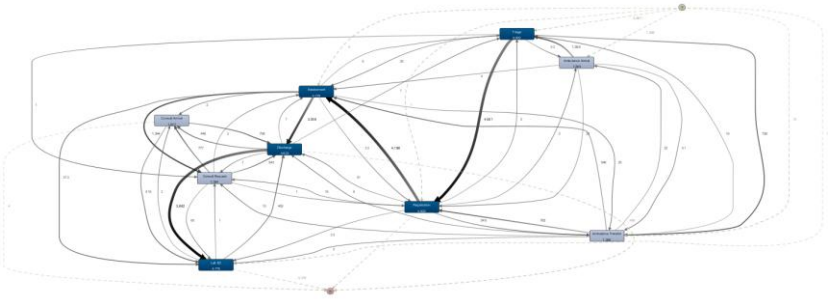
Part 1: Process Discovery and Analysis with DISCO

2.2 Tasks

2. Process Discovery

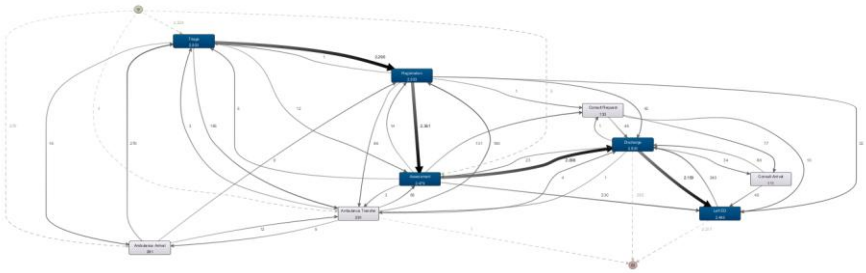
- **Identify the "Happy Path" for a CTAS Level 2 (EMERGENCY) patient.**

From the below process graph, the Happy Path is: Triage -> Registration -> Assessment -> Discharge -> Left ED. This covers 30% of all cases, and 32% of all events. The median duration of the path is: 6m + 35 m + 2.5h + 13.5m = 3h24.5m



- **Use the Attribute Filter to isolate cases that spent time in the EPZ zone. How does the process graph change?**

The process graph now has fewer paths. Around 15% of all cases, and 14% of all events involve time spent in the EPZ zone. The Happy Path is still: Triage -> Registration -> Assessment -> Discharge -> Left ED. The median duration of the path is: 5m + 25 m + 68m + 6m = 1h44m



3. Performance Analysis

- **Calculate the median time between Registration and Assessment. This is the PIA (Physician's Initial Assessment) time**

Through filtering activities under trim longest mode: Start point = Registration, End point = Assessment, the median time is calculated as 34 min.

- **Identify the LWBS (Left Without Being Seen) rate: Analyze patients with disposition descriptions such as "Left After Triage" or "Left After Initial Assessment". What is the median time these patients spent in the ED before leaving?**

After filtering attributes:

(1) disposition_desc = "Left After Triage" or "Left After Initial Assessment"

And filtering activities under trim longest mode:

(2) Event start point = “Registration” or “Triage”

(3) Event end point = “Left ED”

The median time patients spent in the ED is 113.5 minutes.

- **Identify the ‘Ping-pong’ effect: How often do patients move back and forth between zones or repeat assessments?**

I checked all variants from the data set, but detected no cases where a certain activity was repeated over once. Therefore, there is no “Ping-pong” effect in the current data set. Because zone information is not given except for the initial zone when the patient first arrived, the analysis of patients moving back and forth between zones cannot be carried out.

- **Consultations: Filter for cases containing the Consult Request activity. What is the average delay added by waiting for a consultant to arrive?**

After filtering activities with ‘Consult Request’ under mandatory mode, the average delay added by waiting for a consultant to arrive (mean duration between consult request and consult arrival) is 2.3 hrs.

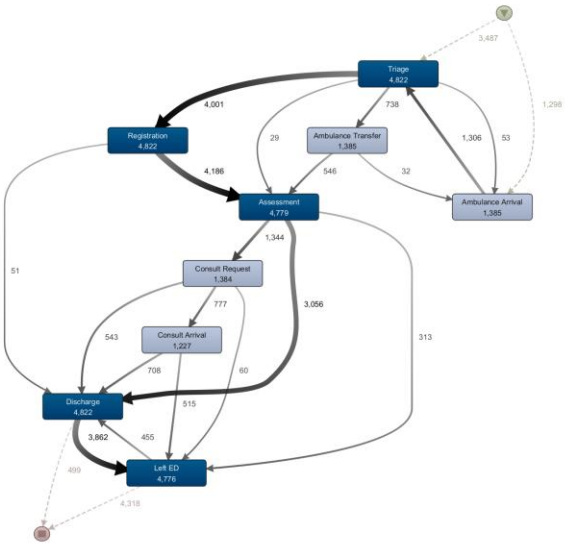
2.3 Deliverables

- **Annotated process maps for the main triage levels.**

For clarity, process maps are shown with activity slicer = 60%, path slicer = 40%.

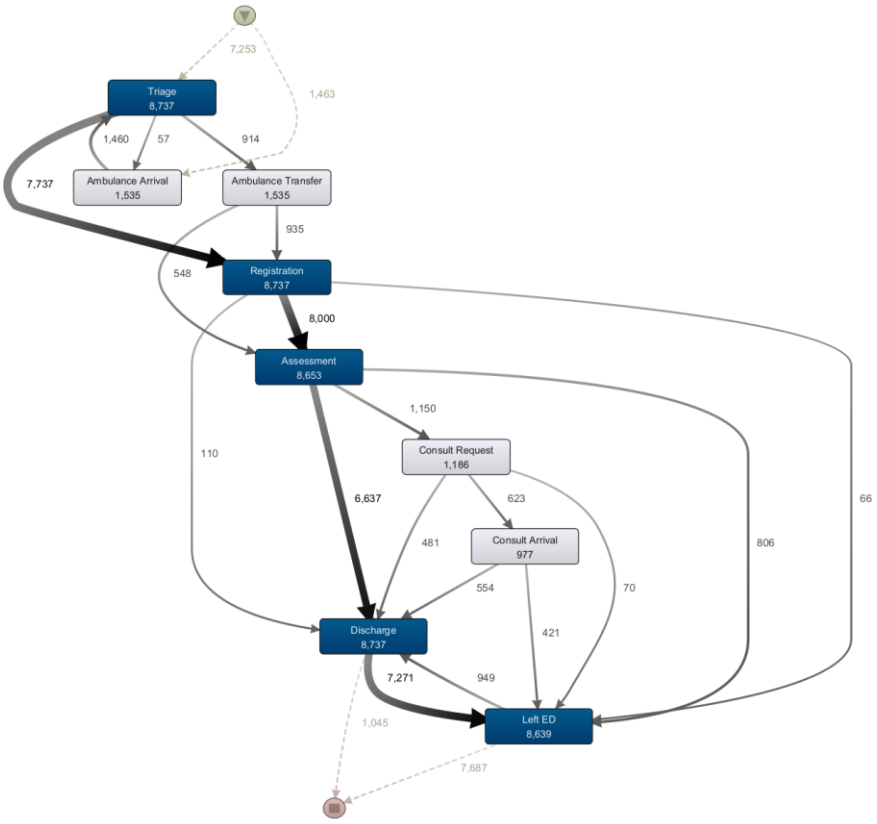
Triage Emergency (30% cases)

For this triage level, 54% of patients go through the happy path (Triage – Registration – Assessment – Discharge – Left ED). 22% arrive through ambulance, and 22% request consultancy.



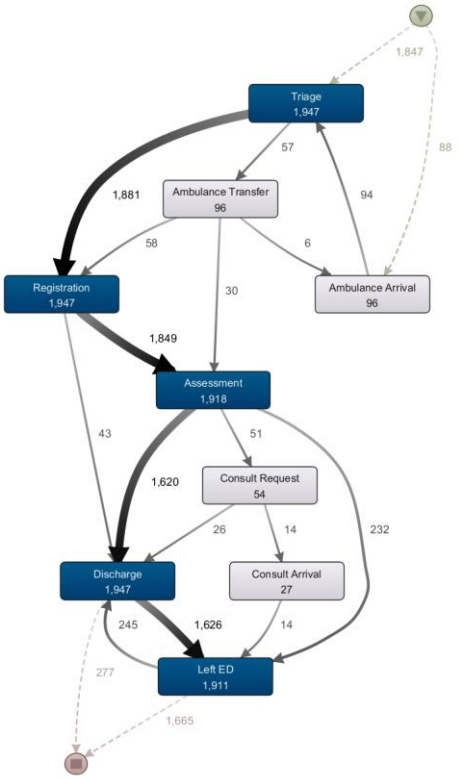
Triage Urgent (54% cases)

For this triage level, 73% of patients go through the happy path. 13% arrive through ambulance, and 8% request consultancy.



Triage Semi-urgent (12% cases)

For this triage level, around 90% of patients go through the happy path.



- **Top 3 bottlenecks found:**

Top 3 Bottleneck Stages	Median Duration
Assessment -> Consult Request	2.3 hrs
Consult Request -> Consult Arrival	117 mins
Assessment -> Discharge	92 mins

- **Strategic recommendations for the NTH COO based on your DISCO findings**

Based on the process discovery and performance analysis conducted using DISCO, several strategic opportunities emerge for improving patient flow, reducing delays, and enhancing operational efficiency in the Emergency Department (ED). The analysis indicates that while a clear and consistent “Happy Path” exists across all triage levels, significant delays arise at specific stages that materially extend patient length of stay and increase the risk of congestion and patient dissatisfaction.

The most critical area for intervention relates to consultation-related delays. The DISCO analysis identifies consultation processes as the largest bottlenecks in the ED, with a median duration of approximately 2.3 hours between assessment and consult request, followed by an additional median wait of 117 minutes before consultant arrival. These delays substantially prolong overall patient throughput and affect both emergency and urgent triage categories. To address this issue, the COO should prioritize the implementation of earlier and more structured consultation triggers, particularly for higher-acuity patients. Establishing standardized clinical criteria for early consult initiation and setting clear response-time expectations for consultants could significantly reduce waiting times and improve downstream flow without necessitating substantial increases in staffing.

In addition to consultation delays, the transition from assessment to discharge represents another key contributor to prolonged ED stays. Even for patients following the Happy Path, the median duration between assessment and discharge remains high at 92 minutes. This suggests opportunities to improve efficiency through enhanced discharge planning, streamlined documentation, and earlier initiation of discharge readiness checks during assessment. By reducing decision-making and administrative delays at this stage, the ED can increase patient turnover and alleviate pressure on assessment areas, particularly during peak demand periods.

The analysis further highlights the effectiveness of the Emergency Processing Zone (EPZ) as a fast-track pathway. Although only approximately 15 percent of cases involve time spent in the EPZ, these patients experience substantially shorter median ED durations compared to the overall population. This suggests that the EPZ is an effective mechanism for accelerating patient flow when appropriately utilized. The COO should consider expanding EPZ eligibility criteria for suitable CTAS Level 2 and 3 patients and reinforcing triage-level decision support to ensure consistent and optimal use of this resource.

Another important finding concerns patients who leave without being seen (LWBS). These patients spend a median of 113.5 minutes in the ED before leaving, typically during the early stages of their visit following triage or initial assessment. This indicates that early-stage delays are a key driver of LWBS risk. Targeted efforts to reduce the Registration-to-Assessment (PIA) time, improve communication around expected wait times, and implement periodic patient reassessments could help mitigate premature departures and reduce associated clinical and reputational risks.

Finally, the consistency of the Happy Path across triage levels presents an opportunity to standardize and reinforce best-practice workflows. A majority of patients, particularly in urgent and semi-urgent categories, follow this optimal sequence of care. By using the Happy Path as a benchmark process, the ED can more easily identify deviations, monitor performance, and guide staff training. Standardizing care delivery around this proven pathway would support greater reliability, reduce unnecessary variation, and improve overall operational predictability.

In summary, the DISCO findings suggest that targeted improvements in consultation workflows, assessment-to-discharge efficiency, EPZ utilization, and early-stage patient flow could materially enhance ED performance. These interventions focus on optimizing existing processes rather than structural expansion, making them both practical and cost-effective strategies for improving patient outcomes and operational resilience.

Part 2: Building an Intelligent Process Mining Tool

Persona Summary – Triage Lead

The Triage Lead is responsible for managing patient inflow, prioritizing assessments, and preventing overcrowding. They are time-constrained, non-technical, and need actionable, real-time insights rather than raw data.

Application Report for ‘ED Triage Lead Decision Support Tool

1. Executive Summary

Rotman

The ED Triage Intelligence Tool is a process-mining and predictive analytics application designed to eliminate operational "blind spots." By transforming raw event logs into real-time visualizations, the tool allows Triage Leads to identify bottlenecks, monitor protocol safety, and simulate the impact of staffing changes on patient Length of Stay (LOS).

2. Technical Pre-requisites & Setup

To ensure the application runs reliably in a clinical environment, the following technical requirements must be met:

- **Software Environment:** The tool is built in Python 3.9+ using the Streamlit framework.
- **Required Libraries:** (as listed in requirements.txt)
  - pandas & numpy: For high-speed clinical data processing.
  - plotly: For interactive, touch-ready visualizations.
  - scikit-learn: For the hospital admission prediction engine.
  - networkx: For mapping complex patient flow networks.
- **Data Requirements:** The tool expects a .csv file containing the following columns: patient_id, timestamp, event, triage_code, initial_zone, and disposition_desc.

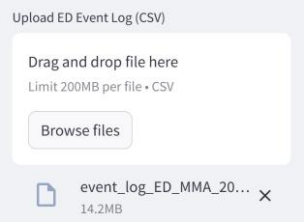
Please read the user_guide.txt to set up your virtual environment and run the application.

3. Dashboard Navigation & Interactions

The interface is designed for high-pressure environments where information must be accessible in seconds.

Global Controls (The Sidebar)

- **Data Upload:** Use the sidebar to upload the daily event log.

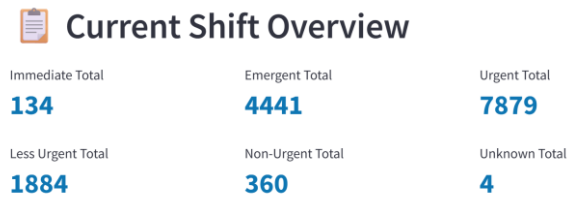


- **Acuity Filters:** Select specific Triage Levels (e.g., *Immediate* and *Emergent*) to focus only on high-risk patients. The entire dashboard updates instantly based on this selection.

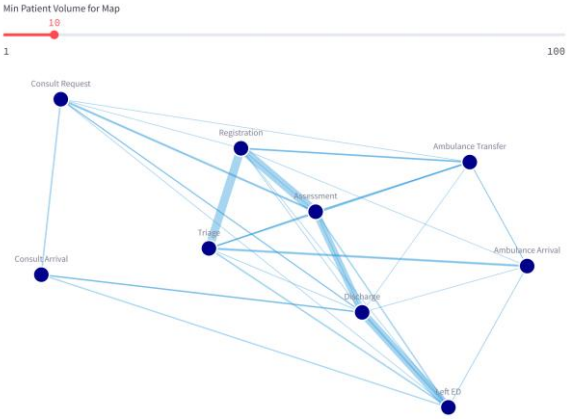
- **Urgency Guide:** A color-coded legend (Red to Green) helps maintain clinical context while navigating data.

Interactive Elements

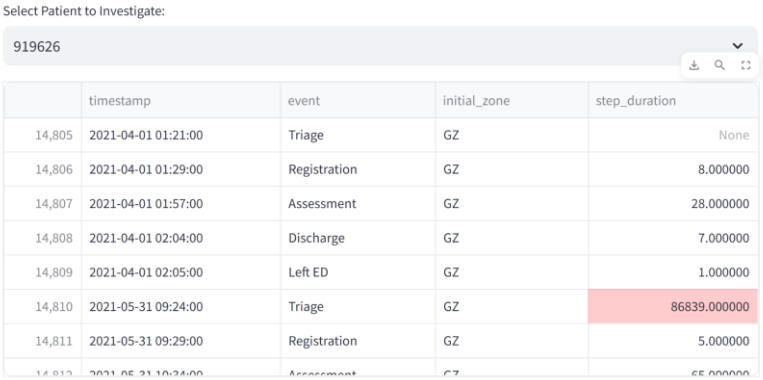
- **Metric Tiles:** Located at the top, these provide a "live pulse" of patient volumes per category.



- **The Flow Map (Section 1):** You can hover over the lines between events to see exactly how many patients followed that specific path. Use the slider to filter out "noise" and see only the most common patient routes.



- **Deep-Dive Selectors (Section 6):** Use the dropdown menu to pick a "Red Flag" patient. The table will automatically highlight the specific step where the delay occurred in red.



	timestamp	event	initial_zone	step_duration
14,805	2021-04-01 01:21:00	Triage	GZ	None
14,806	2021-04-01 01:29:00	Registration	GZ	8.000000
14,807	2021-04-01 01:57:00	Assessment	GZ	28.000000
14,808	2021-04-01 02:04:00	Discharge	GZ	7.000000
14,809	2021-04-01 02:05:00	Left ED	GZ	1.000000
14,810	2021-05-31 09:24:00	Triage	GZ	86839.000000
14,811	2021-05-31 09:29:00	Registration	GZ	5.000000

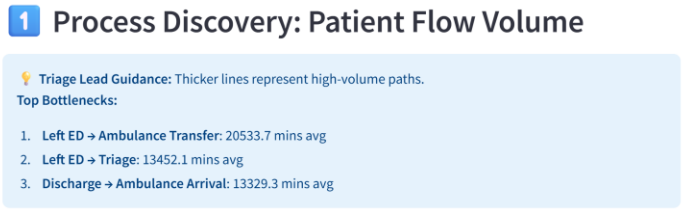
4. Understanding the Results

The results are divided into three strategic pillars: **Flow, Safety, and Prediction.**

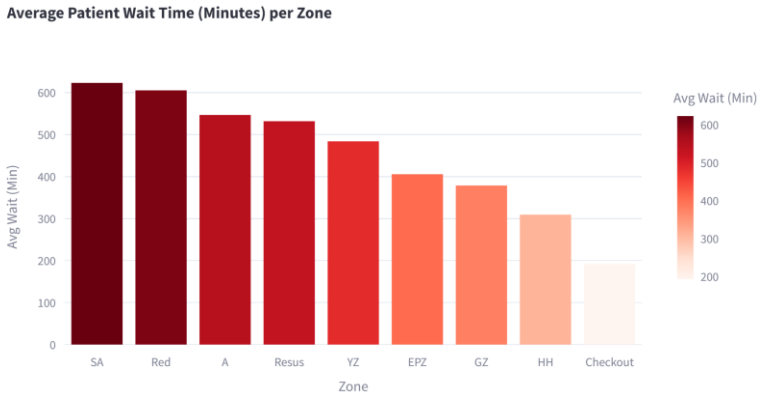
I. Flow Discovery & Load Balancing (Sections 1 & 3)

- **Top 3 Bottlenecks:** The tool automatically calculates which transitions (e.g., *Registration → Assessment*) are causing the longest delays.

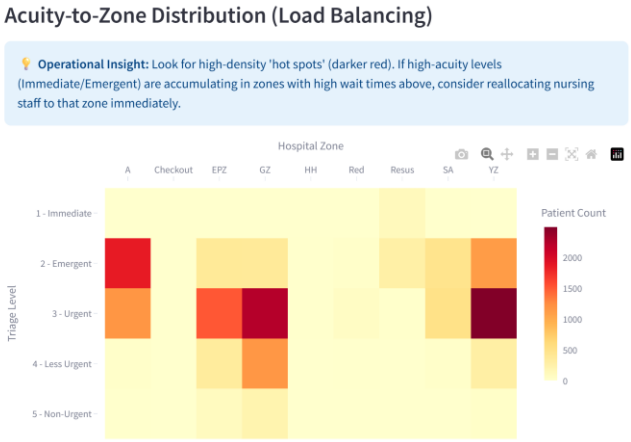
Rotman



- **Wait Time Visual:** The tool visualizes the wait time per zone, with zones of longer wait time highlighted in dark red.



- **Acuity-Zone Heatmap:** This is your Tactical Map. It shows if high-acuity patients are accumulating in zones with high wait times. If you see a dark red "hot spot" in a low-resource zone, it is a signal to reallocate nurses immediately.



II. Safety & Protocol Conformance (Section 2)

- **Compliance Rate:** This tracks how many patients followed the "Gold Standard" path (*Triage → Registration → Assessment → Discharge*).
- **Sequence Violations:** If a patient skips a step or receives an assessment before triage, the tool flags this as a potential safety risk or documentation error.

2 Safety Protocol Conformance

Clinical Gold Standard Protocol: Triage → Registration → Assessment → Discharge → Left ED

Compliance Rate

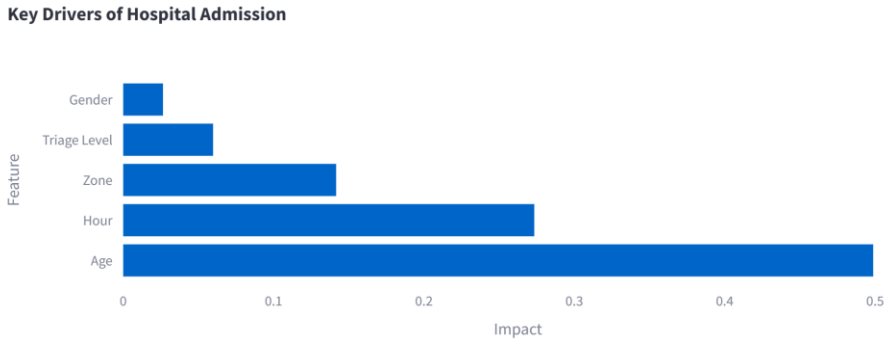
87.3%

View Conformance Exceptions

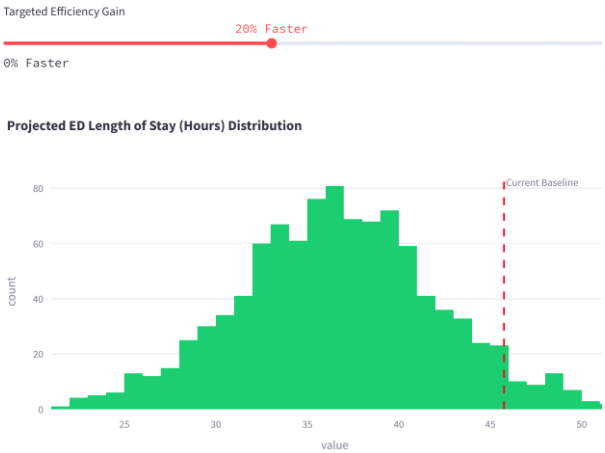
ID	Issue	Trace
77,286	Sequence Violation	Triage → Registration → Assessment → Left ED → Discharge
148,226	Sequence Violation	Triage → Registration → Assessment → Consult Request → L
166,696	Sequence Violation	Triage → Registration → Assessment → Consult Request → L
189,801	Sequence Violation	Triage → Registration → Assessment → Left ED → Discharge
206,527	Missing Step	Triage → Registration → Left ED → Discharge
208,992	Missing Step	Ambulance Arrival → Triage → Registration → Assessment →
230,327	Sequence Violation	Triage → Registration → Assessment → Left ED → Discharge
233,073	Sequence Violation	Triage → Ambulance Arrival → Registration → Ambulance Tr
236,465	Sequence Violation	Triage → Registration → Assessment → Left ED → Discharge
243,913	Missing Step	Triage → Registration → Assessment → Discharge

III. Predictive Intelligence & Simulation (Sections 4 & 5)

- **Admission Drivers:** The Machine Learning model identifies which factors (Age, Triage Level, or Arrival Hour) are most likely to lead to an admission. This helps the Lead plan for bed availability hours in advance.



- **The "What-If" Simulator:** By moving the slider, you can see how much the total ED stay would decrease if you reduced current delays by 10% or 20%.



This tool utilizes the Monte Carlo simulation as a statistical "stress test" for the Emergency Department's operational goals. Instead of providing a single, static average, the logic takes your current real-world Length of Stay (LOS) data and runs thousands of randomized trials to project a future distribution.

When you move the slider to reduce delays by 10% or 20%, the simulation shifts the mean of your current performance downward while maintaining the natural "noise" or variability found in a hospital environment. It uses a probability distribution (typically a Normal or Gamma distribution based on your data) to generate 1,000 "hypothetical shifts." This allows the triage lead to see not just a new average, but the probability of success—specifically, how a reduction in bottlenecks shrinks the "long tail" of extreme wait times, leading to a more predictable and safer patient flow.

5. Practical Use Case for the Triage Lead

Scenario: It is 4:00 PM. The dashboard shows that "Emergent" volumes are spiking (Section 3 Heatmap) and the transition from "Assessment to Discharge" is taking 45 minutes longer than average (Section 1 Bottlenecks).

Actionable Insight:

1. **Reallocate:** Move a float nurse to the zone highlighted in the heatmap.
2. **Expedite:** Identify the specific patients in Section 6 (Red Flags) who have been waiting the longest and prioritize their discharge processing.
3. **Plan:** Check the Admission Predictor (Section 4) to alert the inpatient wards about the expected number of admissions in the next 2 hours.

Rotman

Final Page

Grade: _____