# Performance Modelling

# EEX5362

# Mini - Project


NAME :- K.A.A.S Kodithuwakku
SID :- S22010212
REG NO :- 622515809
BATCH :- BSE 22/23
DATE :- 12/14/2025

# Contents

# 1 System Description: Commercial Bank Branch – Gelioya

## 1.1 Background and Context

The Commercial Bank branch in Gelioya is a small but important bank for the local community. It serves many people from local shop owners depositing cash to elderly residents collecting their pensions. Because it's a small branch, it only has a few resources, such as **one teller counter and one ATM machine.**

The **bank manager is often busy with other tasks**, so the teller handles most customer-facing services like deposits, withdrawals, and other inquiries. The ATM is available for basic cash withdrawals.

Performance matters a lot here. If customers have to wait for a long time, they get frustrated. This can hurt the bank's reputation in the town. Long queues can also make the bank staff feel stressed. The goal of this Mini project is to study the performance of this system. We want to measure key metrics from queuing theory.

## 1.2 System Characteristics

**Current Resource Configuration:**

- **One teller counter** - handles all in-person transactions including deposits, withdrawals, account inquiries, loan applications, and complex financial services.
- **One ATM machine** - provides basic cash withdrawal and deposit services, available for quick self-service transactions.
- **Operating hours**: 9:00 AM to 3:00 PM (6 hours daily)
- **Branch manager** - primarily engaged with administrative duties, occasionally supports during peak periods.

**Service Distribution:**

Based on observational data, approximately 60% of customers require teller services that cannot be completed through the ATM, such as business deposits, pension collection, account opening, loan inquiries, and complex transactions. The remaining 40% utilize the ATM for standard cash transactions.

**Customer Arrival Patterns:**

- The baseline average inter-arrival time is approximately 4 minutes (around 15 customers per hour) derived from Deliverable 01 observations.
- Customer arrivals exhibit stochastic variability and are modeled using a normally distributed inter-arrival process to approximate Poisson behavior.

- In the extended mini project analysis, arrival rates are treated as time-dependent, with lower arrival rates during normal hours and significantly higher arrival rates during peak periods (e.g., lunch hours), enabling realistic workload variation and stress testing of the system.

**Service Time Characteristics:**

- **Teller service**: Exponentially distributed with mean of 8 minutes (varies from simple withdrawals to complex business transactions)
- **ATM service**: Uniformly distributed between 2.5 to 5.5 minutes (includes occasional delays due to user unfamiliarity)

## 1.3 Stakeholders

The system serves multiple stakeholder groups, each with distinct needs and expectations

**1. Bank Customers:**

- Local business owners requiring daily deposits and business banking services.
- Elderly residents collecting pensions and conducting routine financial transactions.
- General community members for savings, withdrawals, and inquiries.
- Expect minimal wait times and efficient service delivery.

**2. Bank Staff:**

- Single teller managing all customer-facing transactions.
- Branch manager handling administrative functions.
- Concerned with workload balance and service quality.

**3. Bank Management:**

- Regional and corporate leadership.
- Focus on operational efficiency, customer satisfaction metrics, and cost-effectiveness.
- Must balance service quality with resource constraints.

# 2 Performance Objectives

The primary goal of this analysis is to move beyond simple observation and quantitatively evaluate the branch's performance, identify the root cause of inefficiencies, and test the impact of potential solutions.

## 2.1 Primary Performance Objectives

1. **To Identify and Quantify the Primary Bottleneck:** We will use simulation to prove which resource (teller or ATM) is constraining system performance.
2. **To Minimize Customer Wait Time (Response Time):** The key objective is to reduce the average time a customer waits in the teller queue (Wq), as this is the main source of customer dissatisfaction.
3. **To Optimize Resource Utilization (Allocation):** We aim to reduce the teller's utilization ($\rho$) from its current "saturated" state (over 90%) to a healthier, more sustainable level (e.g., 60-80%) while ensuring the ATM is not over-provisioned.
4. **To Analyze and Reduce System Queue Length:** We will measure and find solutions to reduce the Average Queue Length (Lq) and, critically, the Maximum Queue Length to prevent the physical bank lobby from overflowing.

## 2.2 Key Performance Metrics (Queuing Theory Parameters)

**1. Average Wait Time (Wq)**

Average time a customer spends waiting in queue before service begins (measured in minutes). Critical metric for customer satisfaction assessment.

**2. Average Queue Length (Lq)**

Average number of customers waiting in queue at any given time. Indicates system congestion level and space utilization.

**3. Resource Utilization ($\rho$)**

Percentage of time each resource (teller/ATM) is actively serving customers. Formula is $\rho = \lambda/\mu$, where $\lambda$ is arrival rate and $\mu$ is service rate. Values approaching 100% indicate resource saturation.

### 4. Maximum Queue Length

Peak number of customers waiting simultaneously during simulation period. Indicates worst-case scenarios and space requirements.

### 5. Arrival Rate (λ)

Average number of customers arriving per unit time (customers per minute). Fundamental parameter affecting all other metrics.

# 3 Modeling Approach and Assumptions

## 3.1 Modeling Technique

A **discrete-event simulation** approach was selected using the SimPy library in Python. Simulation is appropriate because it allows detailed modeling of stochastic arrivals, variable service times, and time-dependent system behavior that is difficult to capture analytically.

Conceptually, the system can be viewed as a combination of M/M/c and M/G/1 queuing structures. The M/M/c queue for teller services and the M/G/1 queue for ATM services.

## 3.2 Assumptions

Here are the main assumptions and choices I made for the model, based on observing the branch:

- **Simulation Time:** I ran the simulation for **360 minutes (6 hours)**, which is like a full banking day.
- **Customer Arrivals:** I assumed customers arrive, on average, **every 4 minutes**. I used a normal distribution for this to make it feel more realistic than customers arriving at exact 4-minute intervals.
- **Customer Choice:** Based on my observations, more people need the teller for services the ATM can't do. So, I modeled that **60% of customers go to the teller** and **40% go to the ATM**.
- **Teller Service Time:** The teller handles many different tasks, so the time varies. I used an **exponential distribution** (a standard in queuing theory) with an average service time of **8 minutes**.
- **ATM Service Time:** ATM transactions are usually faster. I modeled this using a **uniform distribution** (between 2.5 to 5.5 minutes), with a small chance of a "slow user" (like someone forgetting their PIN) to add realism.
- **Queue discipline:** First-come, first-served (FCFS) for both teller and ATM.
- No system failure

# 4 Dataset Description & Methodology

The dataset contains detailed records of individual customer transactions and each row represents a single customer interaction with the banking system. (Raw data for baseline scenario)

**Data Fields:**

- **customer_id:** Unique identifier for each customer transaction
- **arrival_time_min:** Timestamp when customer arrived at the branch (minutes from opening)
- **service_type:** Type of service utilized (Teller or ATM)
- **service_time_min:** Duration of service in minutes
- **wait_time_min:** Time spent waiting in queue before service (minutes)
- **queue_length_on_arrival:** Number of customers already waiting when this customer arrived

**Input Parameters (Model Data):**

| Parameter | Value | Description |
|---|---|---|
| Simulation Duration | 360 minutes (6 hours) | A standard banking day. |
| Mean Inter-Arrival Time | 4.0 minutes (Normal Distribution) | The average time between new customers. |
| Customer Split | 60% Teller / 40% ATM | The probability of a customer choosing a service. |
| Teller Service Time | 8.0 minutes (Exponential Distribution) | The average time the teller takes per customer. |
| ATM Service Time | 2.5 - 5.5 minutes (Uniform Distribution) | The range of time the ATM takes per customer. |

**Methodology: Two-Phase Analysis**

To ensure depth and rigor, I conducted the analysis in two phases:

- **Phase 1: Preliminary Analysis (Constant Load).** I tested three basic configurations to find the bottleneck.

- **Phase 2: Advanced Analysis (Varying Load).** I took the best solution from Phase 1 and stress tested it under a high-traffic "Lunch Rush" scenario.

# 5 Detailed Analysis & Visualizations

**Phase 1 : Constant Load**

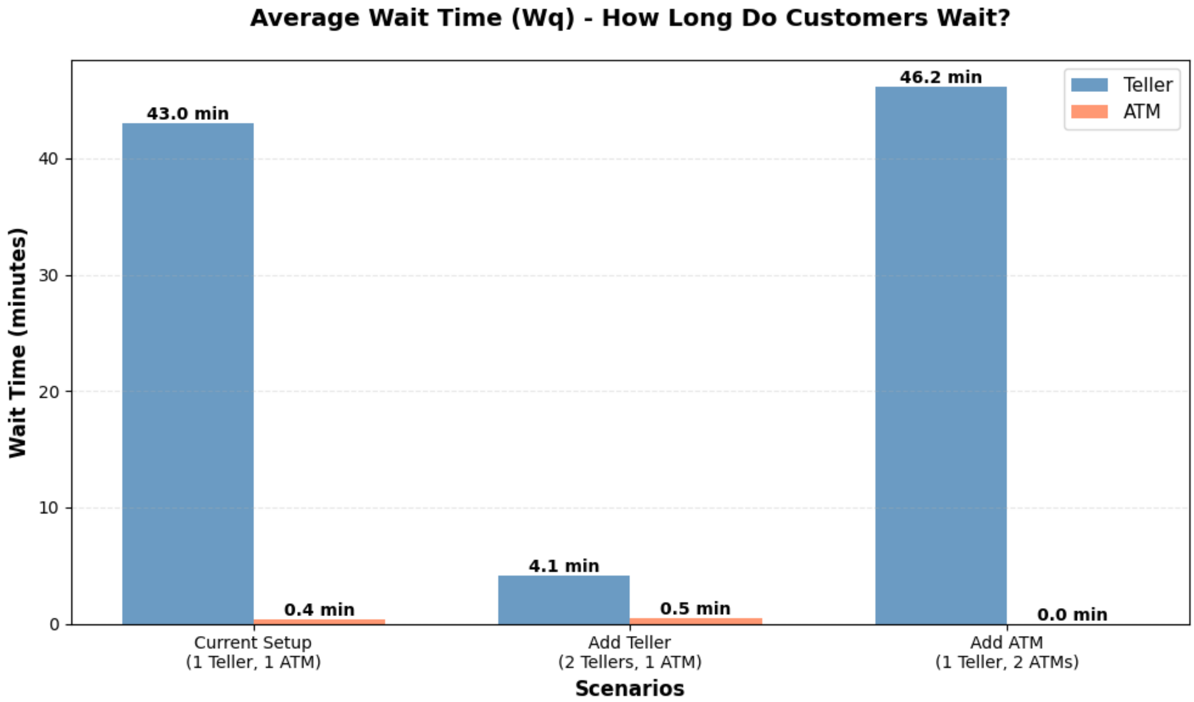
I used my simulation to test three different  scenarios to see how we could improve performance.

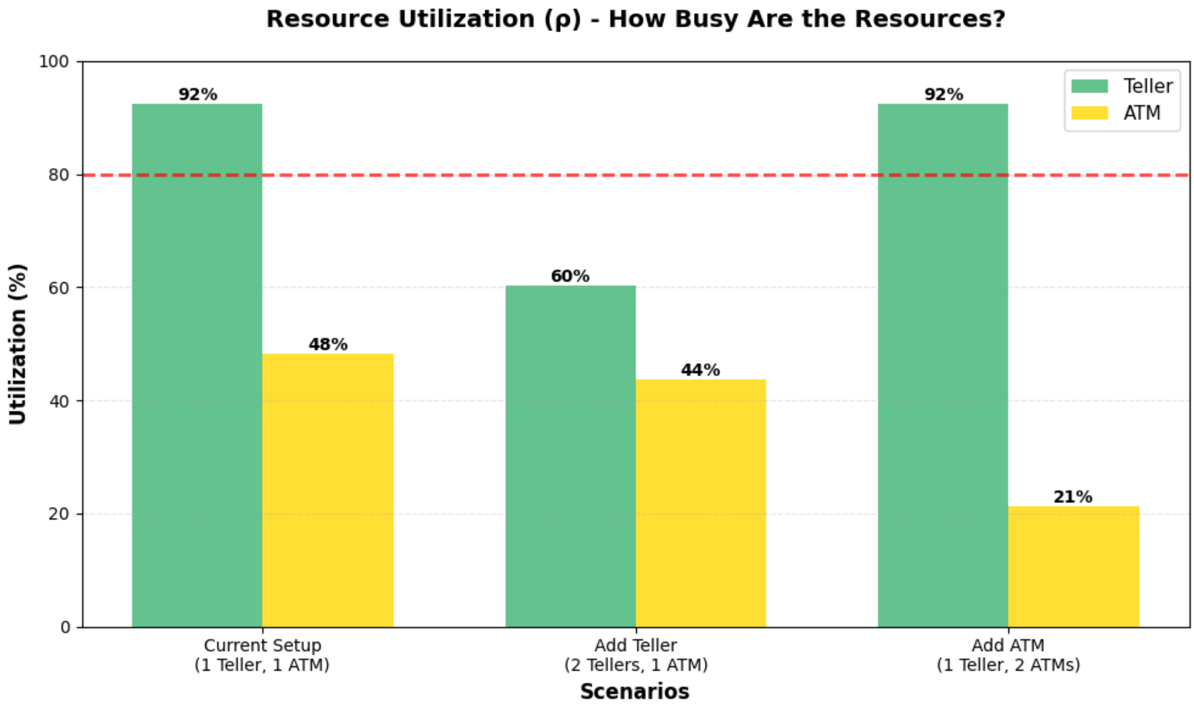
- **Scenario 1: Current Setup (Baseline)**
  - **Configuration:** 1 Teller, 1 ATM
  - **Purpose:** This is our baseline. It shows how the bank is operating right now. We use this to compare all other scenarios.
- **Scenario 2: Add a Teller**
  - **Configuration:** 2 Tellers, 1 ATM
  - **Purpose:** This tests the solution of hiring one new staff member, so two tellers are working at the same time. This is a high-cost solution (salaries) but might be needed.
- **Scenario 3: Add an ATM**
  - **Configuration:** 1 Teller, 2 ATMs
  - **Purpose:** This tests if investing in another machine is a good solution. This is a one-time cost and might be cheaper than hiring new staff.

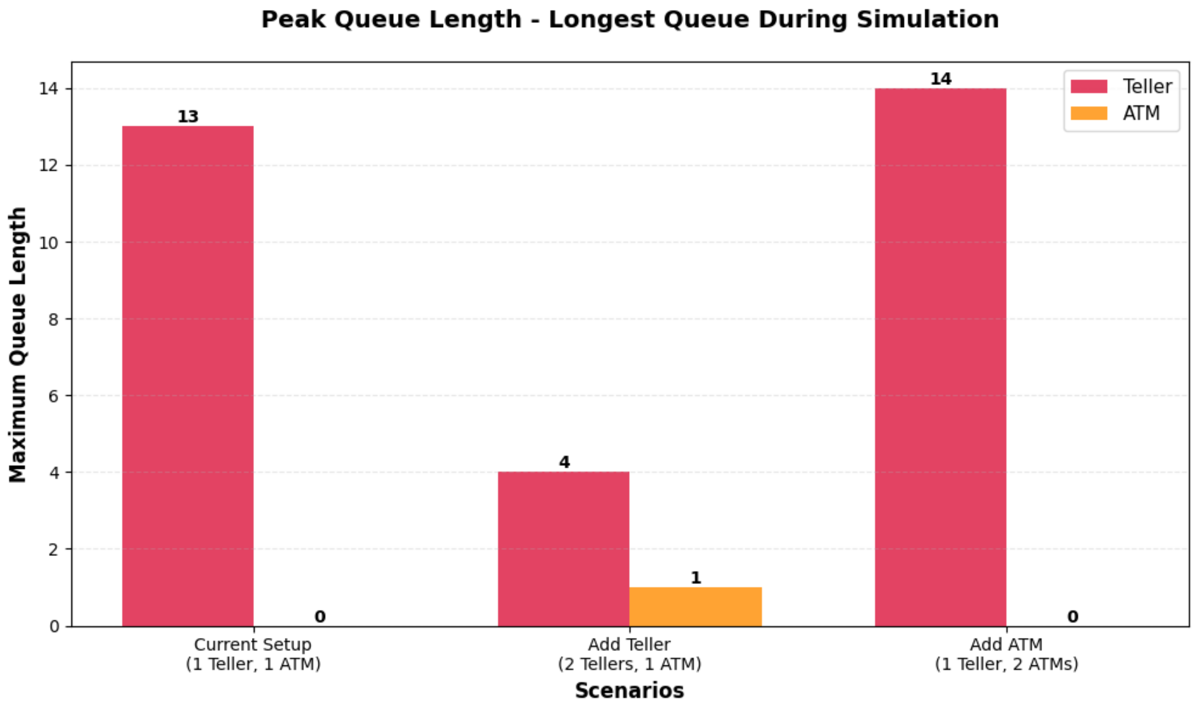
Scenarios Simulated

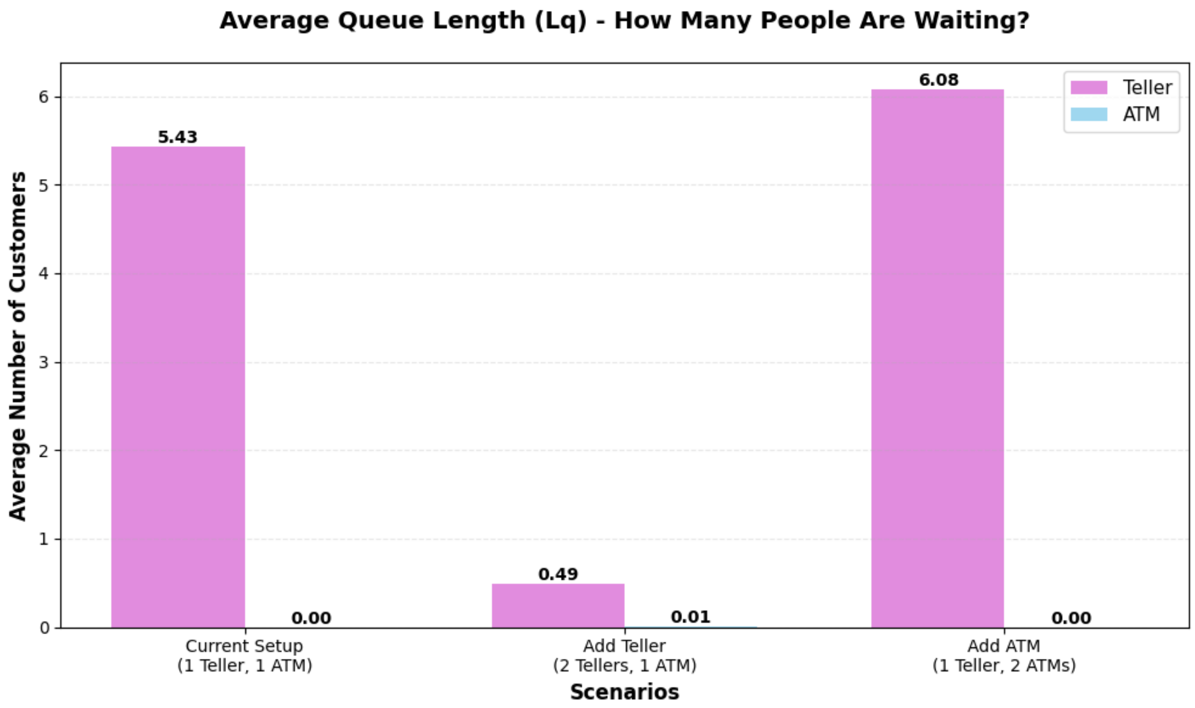
| Scenario | Teller Count | ATM Count | Mean Interarrival (min) |
|---|---|---|---|
| Current Setup | 1 | 1 | 4.0 |
| Add Teller | 2 | 1 | 4.0 |
| Add ATM | 1 | 2 | 4.0 |

I ran the simulation for all three scenarios and collected the data. The results from the simulation runs are summarized in the charts below.




Peak Queue Length - Longest Queue During Simulation
Maximum Queue Length
Scenarios
Current Setup
(1 Teller, 1 ATM)
Add Teller
(2 Tellers, 1 ATM)
Add ATM
(1 Teller, 2 ATMs)
Teller
ATM
13
0
4
1
14
0
0
2
4
6
8
10
12
14


Average Queue Length (Lq) - How Many People Are Waiting?
Average Number of Customers
Scenarios
Current Setup
(1 Teller, 1 ATM)
Add Teller
(2 Tellers, 1 ATM)
Add ATM
(1 Teller, 2 ATMs)
Teller
ATM
5.43
0.00
0.49
0.01
6.08
0.00
0
1
2
3
4
5
6

Here is a breakdown of what these results mean:

**<u>Scenario 1: Current Setup (1 Teller, 1 ATM)</u>**

- **Teller Wait Time:** 43.0 minutes
- **Teller Utilization:** 92.4%
- **Teller Max Queue:** 13 people
- **ATM Wait Time:** 0.4 minutes
- **ATM Utilization:** 48.2%

The results for our baseline are very clear. The **teller is the problem**. A 92.4% utilization is extremely high. In queuing theory, this means the resource is "saturated." The teller is working non-stop, and the queue just keeps building up. An average wait of 43 minutes is unacceptable for customers. The max queue of 13 people would fill up the entire branch.

The ATM, on the other hand, is perfectly fine. It's used less than half the time, and the wait time is almost zero.

**<u>Scenario 2: Add a Teller (2 Tellers, 1 ATM)</u>**

- **Teller Wait Time:** 4.1 minutes
- **Teller Utilization:** 60.2%
- **Teller Max Queue:** 4 people
- **ATM Wait Time:** 0.5 minutes
- **ATM Utilization:** 43.8%

This scenario shows a **massive improvement**. By adding a second teller, the average wait time for teller service dropped from 43 minutes to just **4.1 minutes**. This is a **90% reduction** in waiting. The teller utilization is now at a very healthy **60.2%**, meaning the tellers have breathing room between customers. The maximum queue length also dropped to a very manageable 4 people. This solution clearly solves the problem.

**Scenario 3: Add an ATM (1 Teller, 2 ATMs)**

- **Teller Wait Time:** 46.2 minutes
- **Teller Utilization:** 92.5%
- **Teller Max Queue:** 14 people
- **ATM Wait Time:** 0.0 minutes
- **ATM Utilization:** 21.2%

This scenario is very interesting because it shows what not to do. **Adding a second ATM did nothing to fix the main problem.** The teller wait time is still over 46 minutes, and the teller is still 92.5% busy. All this did was make the ATMs even less busy (only 21.2% utilization). This proves that the problem was never the ATM, it was always the teller.

Summary Table

| Metric | Current Setup (1T, 1A) | Add Teller (2T, 1A) | Add ATM (1T, 2A) |
|---|---|---|---|
| Teller Avg. Wait Time | **43.0 min** (Very High) | **4.1 min** (~90% Decrease) | **46.2 min** (Still High) (No Change) |
| Teller Avg. Queue | **5.43** (Long) | **0.49** (Short) (~91% Decrease) | **6.08** (Still Long) (No Change) |
| Teller Utilization | **~92%** (Overworked) | **~60%** (Healthy) | **~92%** (Overworked) |
| ATM Utilization | **~48%** (Not Busy) | **~44%** (Not Busy) | **~21%** (Very Low) |
| Total Customers | **78** (Base) | **87** (~11.5% Increase) | **74** (Slight Decrease) |

According to summary table we can clearly say that the teller is the only problem and adding an ATM is a waste of money.

**Phase 2: Dynamic Arrival Rate**

According to Phase 1 analysis, I identified the best configuration (2 Tellers, 1 ATM) using a constant workload. However, real banks don't have a steady stream of customers all the day, they have busy hours and quiet hours. In Phase 2, I introduced a "Dynamic Arrival" process (meaning customer arrival rates change over time) to see how the system handles stress and if it can recover after a rush.

**Simulation Scenario: The Lunch Rush**

To make the simulation realistic, we divided the operation time into three periods to mimic a typical day with a lunch peak,

1. **Morning (0–120 mins):** Customers arrive every 5 minutes on average (Light load).
2. **Lunch Peak (120–240 mins):** Customers arrive every 3 minutes on average (Heavy load).
3. **Afternoon (240–360 mins):** Back to every 4 minutes (Recovery period).

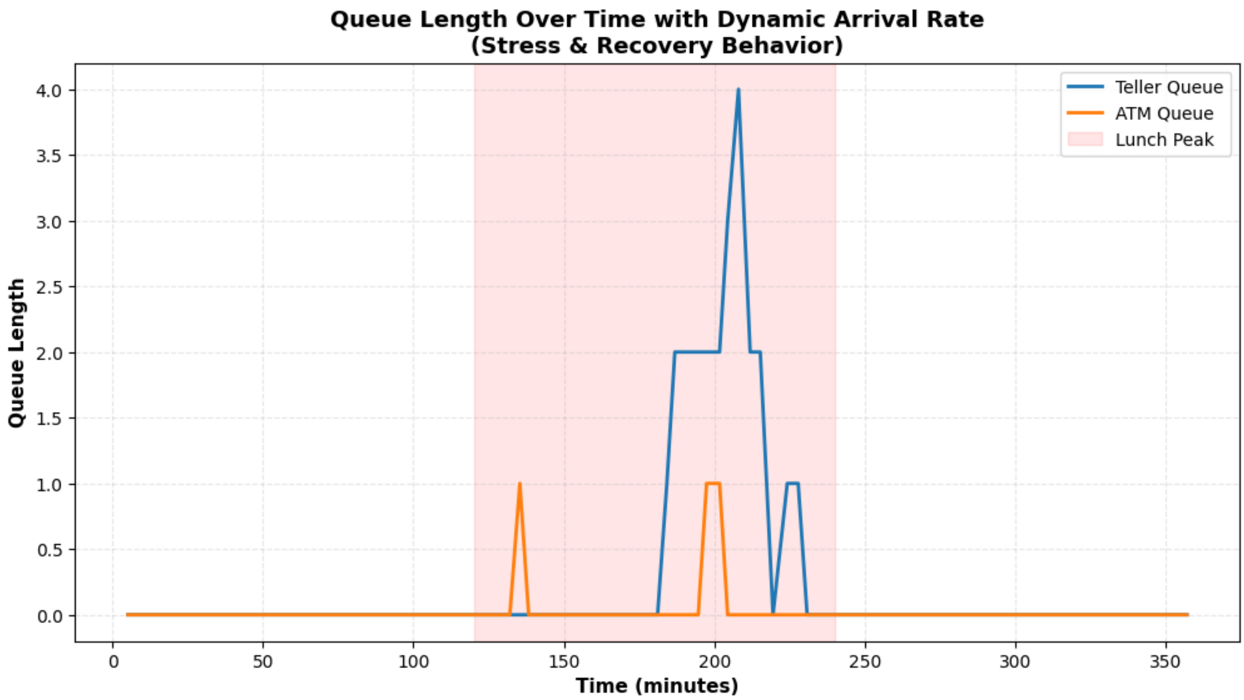
This setup creates a temporary stress test to see if the queue gets out of control during the peak and how fast it clears up afterward.

**Results** Using the **2 Teller, 1 ATM** configuration, we observed the following:

- **Total Customers Served:** 85
- **Teller Utilization:** 47.6%
- **ATM Utilization:** 45.7%
- **Max Teller Queue:** 4 customers
- **Avg. Teller Wait Time:** ~2.08 minutes

**Observations**

- **The Lunch Spike:** As expected, the teller queue grew during the busy lunch period (120–240 mins). This confirms that Tellers are the busiest resource when demand is high.
- **System Recovery:** Unlike the failed configurations in Phase 1, the queue did not explode. It peaked at 4 customers and then gradually went back down to near zero once the afternoon started. This shows the system has "good recovery" it clears the backlog quickly.
- **ATM Performance:** The ATM queue stayed very low the whole time, proving that one ATM is plenty even during rush hour.




Phase 2 confirms that our chosen solution (2 Tellers, 1 ATM) is robust. It can handle sudden surges in customer traffic without causing long delays. The system proved it is stable under stress and capable of recovering quickly after a busy period.

# 6 Conclusion and Recommendations

## 1. Key Insight: The Teller is the Bottleneck

The study of the Commercial Bank branch in Gelioya reveals that the teller counter is the main performance **bottleneck**.

- **Baseline Issue:** With only one teller, utilization hit **92%**, causing long queues and wait times.
- **ATM Status:** The ATM showed low utilization and almost no queues, meaning the current ATM capacity is already sufficient.

## 2. Theoretical Explanation (Queuing Theory)

This behavior can be explained using fundamental queuing theory principles. The arrival rate of customers requiring teller services exceeds the effective service rate of a single teller. As system utilization ($\rho$) approaches unity, the expected waiting time and queue length increase rapidly in a non-linear manner. This theoretical behavior is directly observed in the baseline simulation results, where small increases in demand lead to disproportionately large increases in waiting time.

## 3. Validation Under Stress (Dynamic Load)

Phase 2 (the lunch peak simulation) proved that our solution works in the real world.

- When demand spiked, the teller queue grew temporarily, confirming it is the stress point.
- However, with **2 Tellers**, the system handled the rush and recovered quickly once the peak ended. This proves the solution is robust enough for daily fluctuations.

## Recommendations

Based on the simulation data, we propose the following,

**Priority**: Add a Second Teller

This is the most critical fix. It significantly reduces wait times and keeps the queue length manageable. While this increases staffing costs, it is necessary to bring utilization down to a sustainable level and improve customer satisfaction.

**Avoid**: Adding More ATMs

Do not invest in a second ATM. The simulation shows the existing ATM is underutilized even during peak hours. Adding another one would be a waste of resources without improving system performance.

# 7 Limitations and Future Work

## Limitations

While this simulation provides useful data on the Gelioya branch's performance, there are some simplifications and limitations to keep in mind,

- Fixed Customer Decisions: The model assumes customers choose the Teller or ATM based on a fixed probability. In real life, a customer might see a long teller line and decide to use the ATM instead, which creates dynamic behavior we didn't model.
- No System Failures: We assumed "perfect" conditions no ATM breakdowns, network errors, or staff calling in sick. These real-world interruptions would likely degrade performance further.
- Standard Service Times: Service times were generated using standard statistical distributions (like Exponential). In reality, service times can vary wildly based on transaction complexity or the experience level of the specific staff member on duty.
- Limited Scope: The simulation covers only one branch during a single operational day. It does not account for seasonal trends (like the New Year season) or month-end salary days where demand is much higher.

## Future Work

- **Failure Scenarios:** Introducing random events like ATM outages or network delays to test the system's reliability and robustness under failure conditions.
- **Priority Queuing:** Implementing a priority system to see how handling special-needs or VIP customers affects the wait time for regular customers.
- **Long-Term Simulation:** Running the simulation over a longer period (e.g. a full month) to capture weekly patterns and seasonal variations.

# 8 Code and Outputs

To Check code :-

Open Google Colab Code Share link – Click here

GIT Repository :-

GIT Repository link - Click here