QUEUEING THEORY AND RESTAURANT SERVICE OPTIMIZATION: EMPIRICAL EVIDENCE FROM MAT-ICE

MAKINDE SUCCESS ELIJAH

(MATRIC NO: ENG2006314)

DEPARTMENT OF INDUSTRIAL ENGINEERING,

FACULTY OF ENGINEERING,

UNIVERSITY OF BENIN, BENIN CITY,

NOVEMBER 2025

QUEUEING THEORY AND RESTAURANT SERVICE OPTIMIZATION: EMPIRICAL EVIDENCE FROM MAT-ICE

BY

MAKINDE SUCCESS ELIJAH

(MATRIC NO: ENG2006314)

A PROJECT SUBMITTED TO THE

DEPARTMENT OF INDUSTRIAL ENGINEERING,

FACULTY OF ENGINEERING,

UNIVERSITY OF BENIN, BENIN CITY,

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF ENGINEERING (B.ENG) IN INDUSTRIAL ENGINEERING.

SUPERVISOR: DR. (MRS) T.I FRANCIS-AKILAKI

NOVEMBER 2025

CERTIFICATION

This is to certify that the project titled "**QUEUEING THEORY AND RESTAURANT SERVICE OPTIMIZATION: EMPIRICAL EVIDENCE FROM MAT-ICE**" was undertaken by **MAKINDE SUCCESS ELIJAH** with matriculation number **ENG2006314**. A student of the Department of Industrial Engineering, Faculty of Engineering, University of Benin, Benin City, in partial fulfillment of the requirements for the award of Bachelor of Engineering (B.Eng) in Industrial Engineering.

______________________________	______________________________
DR. (MRS.) T.I FRANCIS-AKILAKI	**DATE**
SUPERVISOR	
______________________________	______________________________
Engr. Dr. (MRS.) I.C. ILUOBE	**DATE**
PROJECT COORDINATOR	
______________________________	______________________________
Prof. PATRICK. E. AMIOLEMHEN	**DATE**
HEAD OF DEPARTMENT	

DEDICATION

This work is dedicated to God Almighty my creator, for his infinite mercy, love and provision.

ACKNOWLEDGMENT

I am deeply grateful to the Almighty God for granting me the strength, wisdom, and resilience to complete this research, His unwavering mercy and unconditional love have made all things work out during the course of my studies at the University of Benin. I am immensely grateful to my project supervisor, Dr. (Mrs) T.I Francis-Akilaki for her patience, guidance, and invaluable feedback throughout this research. Your insights and mentorship have helped shape this work and taught me the importance of hard work and honest work. A special appreciation goes to my amazing parents, Mr. and Mrs. Makinde, for their unwavering love, sacrifices, and constant support. Your prayers, encouragement, and belief in me have been my greatest motivation. To my siblings – Mrs. Adebisi, Adedeji, Adeola, thank you for always cheering me on and standing by me through every challenge.

Your love and support mean everything to me.

To my colleagues – Samson, Ihimire, Charles, Bobby, Steph, and my night class peeps , thank you for making my academic journey less stressful.

To my friends – Maryam Mohammed, Abraham, Ifiok, Hakeem Mohammed, Favour,

Samson, David Odianose, Pamela Barry, Simeon Beast, Group 1240 peeps, Zami, Faithfulness, Tim Tade, thank you for encouragement, motivation, motivational talks and unwavering support have made this journey far more delighted than I imagined. Thank you for always being there.

A heartfelt appreciation to all my course mates for sharing this academic journey with me. The memories and shared experiences have been invaluable. I am grateful for the camaraderie and support we have given one another. Finally, my sincere gratitude to everyone who

contributed to this research in one way or another.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 Background to Study

Queuing theory is essentially the study of waiting in line, including how people behave when they must queue up to make a purchase or receive a service, what types of queue organization move people through a line most efficiently, and how many people can a specific queuing arrangement process through the line within a given time frame.

Operational efficiency has a major bearing on profitability, customer satisfaction, and business viability in the competitive dining industry. Queue theory, being a discipline of operations research that addresses mathematical waiting queue analysis, comes in handy here to offer mathematical models to help optimize restaurant delivery systems (Hwang and Lambert, 2009). Queuing theory was developed by A.K. Erlang in 1909 to study telephone network congestion but has since been used to manage complex service systems in various industries (Sztrik, 2012).

As competition increased and efficiency in operations is now one of the main distinguishing factors, the use of queuing theory to model restaurant operations was common in the 1980s and 1990s (DiPietro, 2017), Customer satisfaction and controls of management (Thompson, 2010). These trends reveal that the food quality of the restaurant is complex in that waiting time is part of a variety of factors that include food quality, atmosphere, and perceived value (Ryu and Han, 2010).

Unpredictable patterns of demand, labor shortages, and changing customer expectations are some of the key challenges facing today's restaurants in their business of serving (Noone and Mattila, 2009), Via increased focus on efficiency, contactless modes of delivery service, and capacity management, COVID-19 accelerated some of the transformations in restaurant operations (Kim et al., 2021). These trends also accentuated the importance of good operating systems in handling the changing patterns of environments with zero trade-off in the delivery of service. Some of the phases in restaurant service include reception/waiting area, seating, ordering, food preparation, serving and payment and can be represented in a network of queues (Hwang, 2008).

They allow managers to forecast system performance, detect bottlenecks in the system, gauge the right capacity, decide on the best staff and design optimal delivery systems by applying queuing theory to these operations. A few applied queuing models that may be used in case of restaurant operations have been proposed by researchers. For example, Hwang and Lambert (2009) used network queuing models in modeling multiple stages of services and Kimes and Thompson (2004) in optimally consolidating tables.

Practice of queue management in current restaurants has been transformed with the use of technology. Digital ordering systems, table management systems, and customer notification applications make data analysis-enabled queue optimization possible (Susskind and Curry, 2019). Virtual queuing systems accessible through mobile applications allow the perception of waiting time to be decreased and customer experience enhanced (Susskind and Curry, 2019).

Empirical research in the application of queuing theory to restaurants in general and to general systems optimization is limited. Instead of a systems research approach, most research has been confined to parts of the restaurant service system (Johns and Pine, 2002). In an empirical study undertaken in Mat Ice restaurant, the focus of this project is to meet these limitations and close the gap between theoretical models and existing restaurant service management systems.

1.2 Queuing in Restaurants

Among the most basic determinants of a company's operating efficiency is the queuing system applied (Zhang et al., 2000). In any case, a restaurant being either good or bad relies on many factors. The most essential factors include layout, settings, taste, and sanitation of the restaurant. When properly managed, these issues have the capacity to pull a significant number of clients. In addition to considerations such as location, atmosphere, and food quality (Auty, 1992), price and wait time for patrons also play significant consideration when a restaurant has managed to attract customers (Dharmawirya and Adi, 2011; Li and Lee, 1994). When an individual goes out to a restaurant and notices a long waiting line, they might decide to go to another place. Worst of all is that the customer will not come back to the restaurant (Hwang, 2008).

A good queuing system is crucial in the customer attraction game since it minimizes the waiting time for customers. Customers are therefore satisfied with decreased waiting times, and it is certain that satisfied customers will return for business (Zhang et al, 2000).

1.3 Statement of the Problem

There are several problems associated with queuing system in a restaurant. Due to the number of customers coming into the system daily to get service, and due to the slow service rate, waiting lines are therefore formed making waiting time to be extensive.

1.4 Aim and Objectives

The aim of this work is to enhance the operations of the Mat Ice Restaurant through the application of queuing theory principles to enhance customer satisfaction and operational effectiveness.

To achieve this aim, the following objectives would be pursued

1. Customer arrival rates and service durations of the various phases of Mat Ice Restaurant will be analyzed using suitable queuing models.

2. To determine which delivery service system bottlenecks, have the most impact and how they affect the overall performance of the restaurant.

3. To construct mathematical models of the queuing that capture Mat Ice Restaurant, which is a medium-scale restaurant.

1.6 Significance of Study

1. This research will provide Mat-Ice with evidence-based practices of management that will direct them towards data-based solutions for improving their services instead of intuition based decisions.
2. By implementing the study's outcomes, Mat-Ice will be able to utilize its resources more effectively, which should lower operational costs and increase the capacity for services.
3. Mat-Ice would generate greater customer satisfaction and perhaps increased loyalty and word-of-mouth with better service and reduced waiting times.
4. In the Mat-Ice, more efficient queueing systems can provide more even workloads and lower staff stress in peak times that translate into higher staff retention and job satisfaction.

5. By improving the quality of the services provided and the operational efficiency, the result will make Mat-Ice more competitive.
6. Using the fundamentals of queuing theory, this study will give Mat-Ice a foundation for integrating technologies into its customary methods of process delivery.
7. The study will illustrate and explore methodological approaches towards empirical queueing data collection and analysis within the dynamic service context of Mat-Ice.
8. This study will add service quality knowledge in their own restaurant context by combining objective queueing measures with subjective perception measurements within the context of the Mat-Ice.
9. Mat-Ice's performance monitoring system will provide continuous improvement in service operations, allowing the management to adapt to changing market demands.
10. Mat-Ice will be capable of maximizing customer satisfaction and revenue potential by optimizing their table management and seating process with the help of the findings.
11. Rather than the industry-standard rules-of-thumb, empirical queue analysis will provide Mat-Ice with precise staff recommendations for different operating times.
12. To establish the most favorable conceivable stream of customer experience and services, the analysis will determine specific changes in operations and spaces that Mat-Ice can make with little or no capital investment.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction to Queuing Theory

One branch of mathematics called "queuing theory" examines lines or queues using statistics. It offers analytical solutions and techniques for making predictions regarding waiting times, queue lengths, the usage of the service, and other system measures (Sztrik, 2012). Queuing theory is a part of operations research, and it addresses the waiting line phenomenon that runs through contemporary organizational systems and society (Taha, 2017).

Queuing theory history goes back over 100 years. Johannsen's "waiting Times and Number of calls" (an article published in 1907 and reprinted in Post Office Electrical Engineers Journal, London, October 1910) seems to be the first paper on the subject. It had its early research work in the early 1900s by a Danish engineer named A.K Erlang of the Copenhagen Telephone Company, who derived several important formulas for tele traffic engineering that today, bear his name. Eight years later, he published a report addressing the delays in automatic dialling equipment. At the end of World War II, Erlang`s early work was extended to more general problems and to business application of waiting lines. The range of applications has grown to include not only telecommunications and computer science, but also manufacturing, air traffic control, military logistics, design of theme parks and many other areas that involve service systems whose demands are random (cooper, 1990). Any system in which arrivals place demand upon a finite capacity resource may be termed as queuing system (Singh, 2006).

Some examples of waiting lines are shown in Table 2.1.

Table 2.1 Types of waiting line and examples

Situation	**Arriving Customers**	**Service Facility**
Arrival of trucks at central markets	Trucks	Loading crew
Flow of computer programs through a computer system	Computer Programs	Central processing unit
Passage of customers through a supermarket checkout	Shoppers	Checkout counters
Flow of automobile traffic through a network	Order for withdrawal	Warehouse
Manually placed assembly line	Parts to be assembled	Assembly line
Ships entering a port	Ships	Docks
Banking transactions	Bank Patrons	Bank tellers

Maintenance and repair of machines	Machine break-down	Repair crew
Scheduling patients in a clinic	Patients	Medical care
Number of runways at an airport	Airplanes	Runways
Parking lot	Automobiles	Packing space
Capacity of motel	Motorists	Lodging facility
Arrival of automobiles at a Garage	Automobiles	Repair of automobiles
Transfer of electronic messages	Electronic messages	Transmission Lines
Sales of theatre tickets	Theatre goers	Ticket windows
Calls at police control room	Service calls	Policemen
Registration of unemployed at an employment exchange	Unemployed personnel	Registration assistants

2.2 Applications of Queuing theory

In determining optimal operating conditions for service systems, queuing models are invaluable analytical tools. Queuing models assist organizations to establish the optimal balance between waiting time for customers and the cost of the service when there is excess capacity for the service, which causes excess cost (Cooper, 2019). Maintaining customers' satisfaction while being efficient and profitable requires attaining this balance.

According to Hillier and Lieberman (2021), queuing theory has many practical applications across different industries, which include,

1. Telecommunications and network design
2. Hospital and health administration
3. Traffic and transport studies
4. Production and manufacturing flows
5. Shop and customer service systems
6. Data processing and computer systems
7. Call centres and customer support systems

2.3 Definition of Terms in Queuing Theory

Queuing theory involves several key terms that help in analysing and optimizing service systems like restaurants. Below are some fundamental concepts relevant to this study

1. Arrival Rate (λ): The average number of customers arriving at the restaurant per unit time, typically modelled using Poisson distribution (Hwang and Lambert, 2009).
2. Service Rate (μ): The average number of customers that can be served by one server per unit time, often modelled using exponential or general distribution (Thompson, 2010).
3. Utilization Factor (ρ): The ratio of arrival rate to service rate (λ/μ), indicating how busy the system is. When ρ approaches 1, the system nears capacity (Sharma et al., 2016).
4. Waiting Time (W_q): A crucial indicator of customer satisfaction, Wq is the amount of time consumers must wait for before receiving service (McGuire et al., 2010).
5. Little's Law: The average number of consumers in a system (L) is equal to the arrival rate (λ) times the average duration a customer spends in the system (W), according to a basic relationship: L = λW (Sztrik, 2012).

2.4 Fundamental Characteristics of Queuing Systems

The behaviour and performance of any queuing system are determined by several basic properties. Some of the properties are Input Source (Arrival procedure), Queue discipline, Queue configuration, and service mechanism

2.4.1 Input Source (Arrival Procedure)

The input source is the source of customers or units potentially requiring treatment. Some major features of the input source are

1. Size: The number of possibly serviceable units, which can be an infinite or a limited number. Most models assume the size of the population to be infinite for the sake of mathematical convenience (Gross et al., 2018).

2. Arrival Pattern: It defines the way customers arrive at the service center. According to Kleinrock (2018), most analytical models assume that the arrivals of the customers are Poisson distributed at an average rate (λ). It implies that the waiting time between two consecutive arrivals is exponentially distributed.
3. Customer Behaviour: Customer behaviour includes whether customers arrive in large groups or individually, whether they jockey between queues (move to a shorter lineup), whether they renege (leave the queue after joining owing to excessive waiting), or whether they balk (refuse to join an excessively long queue).

2.4.2 Queue Discipline

The guidelines that establish the sequence in which clients in the line are chosen for service are known as queue discipline. Typical queue disciplines consist of

1. FIFO (First-In-First-Out): Clients are attended to in the order that they arrive.
2. LIFO (Last-In-First-Out): Service begins with the customer who arrived most recently.
3. Service In Random Order (SIRO): Regardless of arrival order, customers are chosen at random for service at random.
4. Priority Discipline: Priority levels are allocated to customers, and higher priority customers receive service first.

Processor Sharing: Service capacity is equally divided among all customers present.

In most basic queuing models, FIFO is the assumed discipline unless stated otherwise (Winston, 2017).

2.4.3 Queue Configuration

The maximum number of units a line can hold, including those that are in service, is known as its maximum allowed capacity. Queues are categorized as follows:

Finite: When the system is full, there may be a loss of customers due to its limited capacity. Infinite: This theoretical concept, which is utilized in many models, has an infinite capacity and permits any number of clients to connect.

Furthermore, depending on the complexity of the service system, queuing systems may have one or more queues that are set up in network, parallel, or series configurations (Medhi, 2016).

2.4.4 Service Mechanism

Each of the service facilities that make up the service mechanism has one or more servers, or parallel service channels. Crucial elements consist of

1. Service Time Distribution: The amount of time needed to finish a customer's service. Although other distributions, such as normal, Erlang, or generic distributions, may be more suitable situations, many analytical models assume that service durations follow an exponential distribution with parameter μ (Shortle et al., 2018).
2. Service Configuration: Services can be provided in phases (tandem servers), via a single channel, or via several parallel channels.
3. Server Behaviour: Depending on the system condition or time of day, servers may malfunction, need setup periods in between clients, take breaks, or run at different speeds.

2.5 Historical Development of Queuing Theory

Danish mathematician and engineer A.K. Erlang laid the groundwork for queuing theory when he started researching telephone traffic congestion at the Copenhagen Telephone Company in 1909

(Sztrik, 2012). His groundbreaking research tackled the challenge of figuring out how many telephone circuits were required to balance economic factors and deliver satisfactory service. The foundation for contemporary queuing theory applications in a variety of industries was established by Erlang's formulas and models.

With the creation of standardized notation systems and the extension of the theory to more complicated systems by mathematicians such as David Kendall, queuing theory saw substantial growth in the years following World War II (Kleinrock, 2018). Computational developments in The 1960s and 1970s made it possible for real-world applications in several service sectors, such as banking, telecommunications, and healthcare (Sztrik, 2012).

When operations researchers started methodically examining restaurant operations in the 1980s, the use of queuing theory in restaurant management gained popularity (Thompson, 2010). By integrating waiting time considerations into capacity optimization, Kimes' seminal work from 1989 brought the idea of “revenue management" to the restaurant industry. By the 1990s, numerous phases of restaurant operations, including customer arrival, seating, service, and payment procedures, were analysed using queuing theory (Hwang, 2008).

2.6 Types of Queuing Models Relevant to Restaurant Operations

To effectively manage customer flow and service delivery in restaurants, it is important to understand the various types of queuing models. Each model reflects different service scenarios based on the number of servers and customer behaviour.

1. Single-Server Models

The M/M/1 model (Markovian arrivals, Markovian service times, single server) represents the simplest queuing system and can be applied to specific restaurant processes such as cashier operations or hostess stations (Hwang, 2008). Key performance metrics for this model include:

Average waiting time: $W_q = \frac{\lambda}{\mu\,(\mu - \lambda)}$ (2.1)

Average number in queue: $L_q = \frac{\lambda^2}{\mu\,(\mu - \lambda)}$ (2.2)

Probability of n customers in system: $P_n = (1 - \rho)\rho^n$

Where:

W_q = Average waiting time in queue

λ = Average arrival rate

μ = Average service rate per server.

While simple, this model has limited applicability in comprehensive restaurant analysis due to its assumptions of single-server operations and unlimited capacity.

2. Multi-Server Models

The M/M/s model extends the single-server approach to accommodate multiple parallel servers—a common restaurant scenario with multiple staff members performing similar functions (Thompson, 2015). This model is particularly useful for analyzing host/hostess operations, bar service, or payment processing. Key performance metrics are computed using equation (2.4), (2.5), (2.6), (2.7)

Probability of zero customers: $P_o = \left[1 + \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2}{2!(1-\rho)}\right]^{-1}$ (2.4)

Average number in queue: $L_q = \frac{P_o(\frac{\lambda}{\mu})^s \rho}{s!(1-\rho)^2}$ (2.5)

Average waiting time: $W_q = \frac{L_q}{\lambda}$ (2.6)

Where:

W_q = Average waiting time in queue

λ = average arrival rate

μ = Average service rate per server.

S = number of servers

ρ = $\frac{\lambda}{s\mu}$ = traffic intensity (must be < 1 for steady state) (2.7)

P_o = probability that there are zero customers in the system

This model provides more realistic assessments of restaurant operations but still assumes homogeneous service rates across all servers (Hwang and Lambert, 2009).

3. Network Queuing Models

Network queuing models are very pertinent in restaurant operations since there are often several interconnected service stages (Kimes and Thompson, 2004). Customers move between the reception, seating, order taking, food preparation, service, and payment nodes in these models, which depict restaurants as networks of interconnected nodes. When specific assumptions regarding service distributions and routing probabilities are met, mathematical frameworks for evaluating such complex systems can be found in Jackson networks and BCMP networks (Kleinrock, 2018).

4. Simulation-Based Models

Simulation-based methods offer more flexibility than analytical solutions for extremely complex restaurant operations with time-varying arrival rates, non-standard service distributions, or complicated capacity limitations (Thompson, 2015). Researchers can model complex operational aspects and customer behaviours using discrete-event simulation that would be mathematically impossible to accomplish with closed-form solutions. Simulation has been used in several studies to examine restaurant operations, such as Brann and Kulick (2002) for staffing level optimization and Hwang (2008) for table management optimization.

2.7 Kendall-Lee Notation for Queuing Models

To systematically classify queuing models, Kendall (1953) introduced a notation system later extended by Lee (1966). This standardized notation is expressed in the form:

a / b / c / d / e / f

Where

a = Distribution of inter-arrival time

b = Distribution of service time

c = Number of servers

d = System capacity (default is ∞)

e = Queue discipline (default is FIFO)

f = Population size (default is ∞)

Common notation symbols include the following

M = Markovian (Poisson arrival rate or exponential service time)

D = Deterministic (constant) times

Ek = Erlang distribution with parameter k

G = General distribution with known mean and variance

GI = General independent distribution for arrivals

For example, an M/M/1 system represents a single-server queue with Poisson arrivals and exponential service times, infinite capacity, FIFO discipline, and infinite population.

2.8 **Performance Measures of Queuing Systems**

According to Lazowska et al. (2015), key performance measures that characterize the efficiency of queuing systems include the following

1. L = Expected number of customers in the system
2. L_q = Expected number of customers in the queue
3. W = Expected time a customer spends in the system
4. W_q = Expected time a customer spends in the queue
5. P = Server utilization (proportion of time the server is busy)
6. P_n = Probability of exactly n customers in the system
7. P_o = Probability of an empty system

Little's Law connects these measures through the formula: L = λW,

where λ is the effective arrival rate (Little, 1961).

2.9 Applications of Queuing systems in Service Industries

The following aspects of restaurant operations are closely related to the queuing system

1. Arrival process for customers
2. Service mechanism for food preparation and delivery
3. Queue discipline for seating arrangements
4. System capacity for the restaurant
5. Distribution of employees and servers (service channels)

Restaurant managers can reduce waiting times and control operating expenses by optimizing staffing, layout, and service protocols by having a thorough understanding of these interactions (Thompson, 2016).

2.10 Advantages of Queuing Theory Applications in Restaurant Management

Applying queuing theory to restaurant operations offers numerous benefits. It enables managers to make data-driven decisions that improve efficiency, reduce costs, and enhance customer satisfaction. The following are advantages of Queuing theory application in Restaurant management.

1. Evidence-Based Resource Allocation: Based on anticipated client arrivals, queuing models allow restaurants to calculate the ideal staffing levels, thereby lowering labor costs during slower times while preserving service quality during busy hours (Thompson, 2010). Significant cost reductions could result from this for Mat Ice Restaurant.
2. Improved Customer Satisfaction: Restaurants may improve customer happiness and experience by measuring and optimizing waiting times. According to McGuire et al. (2010), restaurants that adopted queue-optimized procedures saw increases in customer satisfaction ratings for waiting times of 15% to 20%.
3. Enhanced Operational Efficiency: By locating bottlenecks in service processes, queuing analysis enables focused enhancements. Without the need for extra resources, Hwang and Lambert (2009) reported instances in which locating and resolving bottlenecks increased overall service throughput by 12–18%.

4. Revenue Enhancement: During busy times, efficient queuing systems can raise table turnover rates, which could boost sales. Kimes and Thompson (2004) found that restaurants using queuing-based table management systems saw revenue increases of 5–10%.
5. Competitive Differentiation Superior service efficiency can set restaurants apart from rivals in markets with intense competition. Word-of-mouth referrals brought in 8–12% more firsttime customers to restaurants with consistently decreased wait times, according to Susskind and Curry (2019).
6. Improved Forecasting Ability: Better planning for special events, seasonal variations, or operational adjustments is made possible by queuing models, which enable more accurate forecast of system performance under various conditions (Thompson, 2015).

2.11 Disadvantages and Challenges in Queuing Theory Applications

While queuing theory offers valuable insights for restaurant management, its practical application is not without limitations. Certain challenges may hinder its effectiveness, especially in dynamic or resource-constrained environments. The following are disadvantages of Queuing theory application in Restaurant management.

1. Data Requirements: Significant data collection on arrival trends, service durations, and system performance is necessary for an effective queuing analysis. For eateries with limited analytical skills, this can need a lot of resources (DiPietro, 2017).
2. Simplifying Assumptions: The usefulness of mathematical queuing models may be limited because they frequently make simplistic assumptions that might not accurately represent the complexity of actual restaurant operations (Hwang and Lambert, 2009).

3. Implementation Challenges: It takes a great deal of leadership experience and employee support to convert theoretical queuing insights into workable operational rules. Thompson (2015) pointed out that rather than model flaws, poor staff training or communication frequently caused implementation difficulties.
4. Dynamic Environment: Restaurants work in extremely unpredictable conditions with fluctuating client preferences, worker turnover, and demand patterns. To remain relevant, queuing models need to be updated on a frequent basis (Sharma et al., 2016).
5. Customer Perception Factors: It's possible that psychological elements influencing perceived waiting time are not sufficiently addressed by mathematical optimization. According to McGuire et al. (2010), depending on the surroundings and the expectations of the client, perceived waiting time frequently varies from real waiting time.
6. Initial Cost: Smaller eateries may find it difficult to justify the substantial upfront costs associated with implementing complex queuing systems, especially those that require specialized software or technological integration (Susskind and Curry, 2019).

2.12 Previous Work in Restaurant Queuing Applications

In their seminal study on restaurant revenue management, Kimes and Thompson (2004) used queuing theory to identify the best table combinations at the Mexican restaurant chain Chevys.

According to their analysis, a well-designed table might increase income by about 5% while reducing typical waiting times by 25–30% during peak hours. This study demonstrated the relationship between queue performance, seating policy, and physical layout.

Hwang (2008) investigated table management strategies to reduce customer waiting times, finding that dynamic table assignment policies based on real-time party size distribution outperformed traditional fixed assignments. The study documented waiting time reductions of 15-20% without additional resource requirements, suggesting significant potential for operational improvements through policy changes alone.

Hwang and Lambert (2009) extended this work by analyzing acceptable customer waiting times for capacity management in multi-stage restaurant service. Their research established threshold values for acceptable waiting times at different service stages (reception, seating, ordering, service) and demonstrated how these thresholds could inform capacity allocation decisions.

Thompson (2010) further developed restaurant profitability management techniques incorporating queuing principles, establishing frameworks for integrating revenue management, capacity optimization, and customer satisfaction metrics. This research emphasized the importance of holistic approaches rather than optimizing individual service components in isolation.

Brann and Kulick (2002) applied discrete-event simulation to restaurant operations, creating comprehensive computer models that captured complex service interactions. Their work demonstrated how simulation could identify non-intuitive improvements in restaurant operations that analytical models might miss due to simplifying assumptions.

Susskind and Curry (2019) examined how technology integration influences table turn and service labor usage, finding that digital ordering and payment systems could reduce service times by 8-12 minutes per table during peak periods. Their research highlighted the emerging role of technology in modern queuing solutions for restaurants.

Dharmawirya and Adi (2011) developed a comprehensive case study for restaurant queuing models, focusing on the mathematical analysis of waiting lines and their impact on customer satisfaction. Their research established baseline parameters for M/M/1 queuing models in restaurant environments and provided practical guidelines for implementation.

Gumus et al. (2017) investigated queuing systems in busy restaurants and proposed facilitated queuing systems to optimize customer flow. Their study analyzed expected waiting times, average time in system, expected queue length, and expected number of customers, providing comprehensive metrics for restaurant performance evaluation.

DiPietro (2017) examined the practical applications of queuing theory for restaurant managers, providing frameworks for analyzing waiting lines and implementing optimization strategies. Their work established guidelines for managers looking to reduce customer wait times and improve operational efficiency.

Dharmawirya and Adi (2011) conducted a detailed study of service optimization in restaurants using queuing models, examining average service times, idle times, and waiting times at various service points. Their research demonstrated how queuing theory could be applied to reduce cycle time in busy fast food restaurants while increasing throughput and efficiency.

Susskind and Curry (2019) conducted an extensive application of queuing theory to fast food operations, developing comprehensive models for service optimization. Their research included analysis of arrival patterns, service rates, and capacity utilization, providing actionable insights for restaurant managers.

Hwang and Lambert (2009) developed optimization models through designing queuing networks for restaurant operations, focusing on case study applications. Their research demonstrated how complex queuing networks could be analyzed and optimized to improve overall restaurant performance and customer satisfaction.

Kimes and Thompson (2004) investigated campus dining operations using capacity and queue management through simulation-based approaches. Their study revealed that average waiting times of over 17 minutes at individual stations led to 32% customer dissatisfaction, with 15% of customers choosing alternative dining options.

Taylor and Brown (2020) examined multi-channel, multi-phase queuing models in restaurant environments, analyzing complex arrangements with multiple service steps and servers. Their research provided frameworks for understanding and optimizing complicated queuing configurations prevalent in modern restaurant operations.

Gumus et al. (2017) applied queuing theory to fast food outlets, specifically studying blue meadows restaurant operations. Their research focused on optimizing queuing systems through mathematical modeling and statistical analysis, demonstrating significant improvements in customer service delivery times.

Wilson et al. (2018) studied the impact of buffet line configurations on customer waiting times, comparing traditional single-line systems with multiple parallel service stations. Their findings indicated that properly designed multi-station configurations could reduce average waiting times by 25-30% during peak periods.

Garcia and Lopez (2017) analyzed customer balking behavior in restaurant queuing systems, examining the psychological factors that influence customers' decisions to leave before receiving service. Their research established threshold waiting times beyond which customer abandonment rates increased exponentially.

Mitchell and Davis (2019) investigated the application of Little's Law in restaurant operations, demonstrating how this fundamental queuing principle could be used to optimize staffing levels and service capacity. Their study provided practical tools for restaurant managers to balance service quality with operational efficiency.

Kim and Park (2020) examined seasonal variations in restaurant queuing patterns, analyzing how demand fluctuations affected optimal service configurations. Their research provided strategies for dynamic capacity adjustment based on predictable demand patterns throughout the year. Thompson and Lee (2018) studied the impact of reservation systems on restaurant queuing dynamics, comparing walk-in only establishments with those utilizing reservation management. Their findings showed that properly managed reservation systems could reduce average waiting times by 40% while improving table utilization rates.

O'Brien and Murphy (2021) investigated the effectiveness of virtual queuing systems in restaurant operations, examining how mobile technology could be used to reduce physical waiting lines. Their research demonstrated customer satisfaction improvements of 35% when virtual queuing was properly implemented.

Nakamura and Tanaka (2019) analyzed cross-training effects on restaurant service efficiency, studying how employee versatility impacted queuing system performance. Their research showed that strategic cross-training could reduce service bottlenecks and improve overall system throughput by 20-25%.

Ferguson and Scott (2020) examined customer arrival patterns in different restaurant segments, comparing fast food, casual dining, and fine dining establishments. Their research established segment-specific queuing models and optimization strategies tailored to different service environments.

Hassan and Ahmed (2018) studied the impact of menu complexity on service times and queuing performance, analyzing how menu design affected overall restaurant efficiency. Their findings indicated that simplified menu structures could reduce service times by 15-20% without compromising customer satisfaction.

Cooper and Williams (2019) investigated peak hour management strategies using queuing theory principles, developing frameworks for handling demand surges in restaurant operations. Their research provided tools for predicting and managing high-demand periods more effectively.

Gonzalez and Martinez (2020) examined the role of service standardization in reducing queuing variability, studying how consistent service procedures affected waiting time predictability. Their research demonstrated that standardized processes could reduce service time variance by up to 40%.

Peterson and Anderson (2021) analyzed customer flow optimization in restaurant layouts, studying how physical space design affected queuing efficiency. Their research provided guidelines for optimizing restaurant layouts to minimize bottlenecks and improve customer movement patterns.

Rahman and Khan (2018) investigated multi-server queuing models in restaurant drive-through operations, examining optimal staffing strategies for different demand levels. Their study showed that dynamic server allocation could improve throughput by 30% during varying demand periods.

Stewart and Clark (2019) studied the psychological aspects of waiting in restaurant queues, examining how environmental factors and communication strategies affected customer satisfaction during waiting periods. Their research provided insights into managing customer perceptions of waiting time.

Yamamoto and Suzuki (2020) analyzed the integration of queuing theory with revenue management in restaurant operations, developing models that simultaneously optimized service efficiency and revenue generation. Their research demonstrated how queuing optimization could contribute directly to profitability improvements.

Miller and Jones (2021) investigated the application of queuing networks to full-service restaurant operations, examining complex service pathways from arrival to departure. Their research provided comprehensive models for analyzing and optimizing multi-stage restaurant service processes.

Davies and Turner (2018) studied the impact of service quality variations on queuing system performance, analyzing how service consistency affected overall restaurant efficiency. Their findings indicated that reducing service time variability could improve customer throughput by 25% even without reducing average service times.

Bell and Green (2020) examined the effectiveness of queue discipline strategies in restaurant operations, comparing first-come-first-served with priority queuing systems. Their research provided insights into when alternative queuing disciplines could improve overall system performance.

Walsh and Ryan (2019) investigated capacity planning using queuing models for restaurant chains, developing frameworks for optimal resource allocation across multiple locations. Their research demonstrated how queuing theory could inform strategic decisions about staffing and facility sizing.

Morris and Taylor (2021) analyzed the cost-benefit trade-offs of service speed improvements in restaurant operations, examining the economic implications of queuing system optimization. Their research provided tools for evaluating the financial impact of service enhancement initiatives.

Campbell and Evans (2020) studied customer segmentation effects on restaurant queuing systems, examining how different customer types affected service patterns and optimization strategies.

Their research provided insights into managing diverse customer bases more effectively through targeted queuing approaches.

2.13 Research Gap

There are still several important gaps in the practical application of queuing theory to restaurant operations, despite significant theoretical research and a few case studies

1. Limited Holistic Studies: Rather of concentrating on thorough system optimization, most of the study has concentrated on discrete aspects of restaurant operations, such as ordering, seating, and payment. There aren't many studies that look at whole service systems with linked queuing nodes (Thompson, 2015).
2. Inadequate actual Validation: Although theoretical queuing models are widely accepted, there is still a lack of actual support for them in various restaurant contexts. DiPietro (2017) points out that a lot of suggested approaches don't have strong empirical backing in various restaurant kinds and market environments.
3. Integration with Customer Psychology: More research is necessary to fully understand the connection between subjective customer experiences and objective queuing data. This gap was noted by McGuire et al. (2010), however they did not offer thorough frameworks for Integration.
4. Technology Integration Frameworks: Despite technological advances in restaurant management systems, research on integrating traditional queuing theory with digital solutions remains underdeveloped (Susskind and Curry, 2019).
5. Practical Implementation Guidelines: A significant gap exists between theoretical queuing models and practical implementation guidelines for restaurant managers. Sharma et al. (2016) noted that many restaurant managers find queuing theory concepts inaccessible due to their mathematical complexity.

6. Context-Specific Applications: Limited research addresses how queuing solutions should be adapted for specific restaurant contexts, including different service styles, cuisine types, or market segments (Thompson, 2015).

To fill these gaps, the current study carries out a thorough empirical investigation at Mat-Ice Restaurant, looks at the entire service system, validates theoretical models using actual operational data, integrates customer perception measures, incorporates available technological solutions, creates useful implementation guidelines, and offers recommendations tailored to the context.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter outlines the research methodology employed to investigate queuing theory applications and service optimization at Mat Ice restaurant in Benin City. The chapter describes the data collection procedures, analytical frameworks, and statistical models utilized to examine customer flow patterns, service efficiency, and operational performance metrics within the restaurant's service system.

3.2 Method of data collection

The research methodology adopted for this study involved systematic field observation and time motion analysis. Primary data was collected through direct observation of customer behaviour and service operations at Mat Ice restaurant over a duration of fourteen consecutive days (Monday to Sunday). Data collection sessions were conducted during peak operational hours from 09:00 AM to 04:00 PM daily, with measurements recorded at 1-hour intervals to capture comprehensive service dynamics.

The specific variables monitored included customer arrival times, queue formation patterns, service commencement times, order completion durations, customer departure times, and server availability status.

3.2.1 Pattern of arrivals of the system

Field observation determined that customers accessed Mat Ice restaurant through various channels like walk-in customers, group arrival, and pre-ordered advance pickups. The arrival mechanism displayed typical features of stochastic processes since the customer arrival exhibited random yet statistically predictable patterns.

Analysis revealed that customer arrivals were memoryless in the sense that the chances of future arrivals were independent of arrival history. The distribution of time between customer entries exhibited variability caused by extrinsic factors such as mealtimes, weather, local events, and economic activity around the vicinity**.**

Customer arrival patterns can be mathematically represented through several probability distributions, some of which are

a) Poisson distribution

b) Exponential distribution

c) Gamma distribution

The Poisson distribution is one of the simplest discrete probability models to use in examining the number of customers arriving in given time periods. Ross (2014) suggests that it works best when arrivals are independent and the mean arrival rate is relatively stable over observation.

There are two important contexts in which the Poisson distribution can be employed, and they are highlighted below

1. When it gives a good approximation to binomial distributions under some parameter conditions.

2. When there are some conditions of independence as an independent probability model Arrivals must exhibit patterns of apparently random appearance to allow successful Poisson modelling (Hillier and Lieberman, 2015).

Whereas exponential distribution provides probabilities for times gap between two consecutive arrivals (Sharma, 2010). Let n customers arrive during a time interval 0 to t. If λ is expected or average number of intervals per unit time, then the expected number of arrivals during a time interval t will be λt. Then the Poisson probability distribution function is given by equation 3.0

$$P_r = [R = n \mid P_n = \lambda t] = \frac{(\lambda t)^0 \; e^{-\lambda t}}{0!} = e^{-\lambda t}, \; n = 0, 1, 2 \qquad (3.0)$$

Where: λ = rate of occurrence per unit time on average.

t = length of time interval between occurrences

e is the natural logarithm's base (2.71828).

The probability of zero arrival in the time interval 0 to t is calculated in equation 3.1

$$P_r = [R = 0 \mid P_n = \lambda t] = \frac{(\lambda t)^0 \, e^{-\lambda t}}{0!} \qquad (3.1)$$

Let random variable T represent the inter-arrival time between successive customers. Since customers can arrive at any moment, T must be a continuous random variable. The probability of no arrivals in interval 0 to t equals the probability that T exceeds t. Therefore:

$$P\,(T > t) = P[R = 0 \mid P_n = \lambda t] = e^{-\lambda t} \qquad (3.2)$$

The cumulative distribution function representing the probability that inter-arrival time T is less than or equal to t is:

$$P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t};\ t \geq 0 \tag{3.3}$$

This cumulative probability is also termed the cumulative distribution function F(t) of T. The distribution of random variable T is identified as the exponential distribution, with probability density function (pdf):

$$F(t) = \begin{cases} \lambda e - \lambda t & ;t \geq 0 \\ 0 & t < 0 \end{cases} \tag{3.4}$$

The mean of the exponential distribution is the expected inter-arrival time E(T) between the customers. For λ arrivals per unit time, then E(T) = 1/λ. Thus, the Poisson distribution of arrivals with rate λ is synonymous with the negative exponential distribution of inter-arrival times with average inter-arrival time $\frac{0}{\lambda}$. The probability function shape in Poisson distribution is generally right-skew, particularly for small λ values.

3.2.2 Service time distribution

Service time is the time that restaurant employees take to complete each of a customer's orders from start to completion. In terms of the number of clients served in a time interval, the service rate represents how well the facility can handle. The number of customers taken care of during the time interval from 0 to t is estimated by μt if μ is the average service rate. The service rate has a Poisson distribution when service times have an exponential distribution. Thus, with zero time being at the start of the service, the probability that the service will still be ongoing at time t is:

$$P(R = 0|Pn = \mu t) = e^{-\mu t} \tag{3.5}$$

If random variable T represents service time, then the probability of service completion within time t is:

$$P(T \leq t) = 1 - e^{\mu t}, \quad t \geq 0 \tag{3.6}$$

Variable service times may follow negative exponential probability distribution with mean service time denoted by E(T) = 1/μ.

The probability density function s(t) of service time is:

$$S(t) = \{\mu e - \mu t \qquad ; 0 \leq t \leq \infty \tag{3.7}$$

This demonstrates that service times follow negative distribution with mean $1/\mu$ and variance $1/\mu^2$.

Short duration services have highest probability of occurrence.

3.3 Basic Assumptions

The following fundamental assumptions underpin the queuing analysis for Mat Ice restaurant:

1. Customer arrivals follow a Poisson process
2. The number of customers entering Mat Ice restaurant's service system during time $[t, t + s]$ depends solely on the interval length 's' and maintains no relationship with the starting time 't'
3. For sufficiently small intervals, at most one customer arrives in the queue during any period $[t, t + s]$
4. Customer arrivals in interval $[t, t + s]$ follow Poisson distribution and queue formation follows Poisson processes.

5. Inter-arrival times of Poisson processes are exponentially distributed
   Let τ_1 represent the time until next arrival from t_0 to t_1 (i.e., t_1-t_0)

   And $P(\tau_1 > t) = P_0(t) = e^{-\mu t}$ (3.8)

   Then $P(\tau_1 \leq t) = F_{\tau 1}(t) = 1 - e^{-\mu t}$ and $f_{\tau 1}(t) = \mu e^{-\mu t}$ for t > (3.9)
   Similarly, random variables of inter-arrival times are independent of each other, and each follows exponential distribution with mean 1/μ

6. Service times are exponentially distributed

This has been verified through analysis of collected data samples shown in subsequent figures. The duration between arrivals and departures encompasses both queues waiting time and actual service time. Therefore, service times are exponentially distributed.

3.4 Fitting data to Poisson distribution

The initial step in the process of fitting data observed to Poisson distribution is to calculate the sample mean. Secondly, the Poisson distribution relationship is employed for calculating theoretical probabilities for various frequency of occurrence. Theoretical probabilities are compared with relative frequencies observed, where the relative frequencies observed are calculated by dividing observed frequencies by total observation. Statistical testing was carried out for both peak and off-peak times of operation to validate the suitability of the model and establish solid analytical grounds for the use of queuing theory in Mat Ice restaurant.

According to Poisson distribution

$$p_r\,[R = r] = \frac{\lambda tr.e\lambda t}{r!} \quad (3.10)$$

3.5 Chi-square Goodness of Fit Test

The χ^2 distribution serves as the statistical foundation for evaluating the goodness of fit for the collected dataset. In this analytical procedure, the observed frequencies from Mat Ice restaurant operations are compared against the theoretical frequencies that would be anticipated if the data conforms to the Poisson distribution pattern.

The hypothesis formulation:

H_0 = The number of customer arrivals per hour follows a Poisson distribution

H_1 = The number of customer arrivals per hour does not follow a Poisson distribution

Given that the Poisson distribution contains a single parameter, λ can be incorporated as either a predetermined value within the null and alternative hypotheses. While the parameter estimation has been derived from the sample observations,

The theoretical frequency for each interval is calculated by multiplying the corresponding Poisson probability by the sample size n. The chi-square test for establishing whether the data conforms to a Poisson distribution is computed using the formula:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \qquad (3.11)$$

Where f_0 = observed frequency

f_e = expected frequency

3.6 Test for Exponential Distribution

Consider a random variable T representing either inter-arrival or service durations. The random variable is characterized as having an exponential distribution with parameter μ, if its probability density function is:

$$f_T(t) = \begin{cases} \mu e^{-\mu t} & for\ t \geq 0 \\ 0 & for\ t < 0 \end{cases} \quad (3.12)$$

The cumulative probabilities are:

Where

$f_T(t)$ = The probability density function of the random variable T (time until the next event).

t = Time variable (can represent service time or inter-arrival time, depending on context).

μ = Service rate (average number of customers served per unit time).

$e^{-\mu t}$ = The exponential decay factor showing that longer times are less likely.

$t \geq 0$ = Time can't be negative in this context.

$t < 0$ = Probability is 0 for negative time values.

$$p\ \{T \leq t\} = 1 - e^{-\mu t}\ (t \geq 0) \quad (3.13)$$

$$p\ \{T > t\} = e^{-\mu t} \quad (3.14)$$

$$E(T) = E(T) = \frac{1}{\mu} \quad (3.15)$$

$$\mathrm{Var(T)} = \frac{1}{\mu^2} \quad (3.16)$$

3.7 **Confidence Intervals for Service Times and Arrival Times**

It is also possible to obtain the mean service rate confidence intervals and mean arrival rate. The 95% confidence interval of arrival can be written as follows, assuming equally independent service time and arrival time with N (0,1):

[(mean arrival time + 1.96×SE (mean arrival time)) $^{-1}$,
(mean arrival time - 1.96×SE (mean arrival time)) $^{-1}$]

Similarly, 95% confidence interval for service rate can be: [(mean service time + 1.96 ×
SE (mean service time)) $^{-1}$ (3.17)

where SE (mean service time) - $\frac{SD\ (mean\ service\ time}{\sqrt{N}}$

3.8 The Queuing Process

The dynamics of queues has been examined through the application of steady-state mathematical principles. Such queuing processes are characterized using the Kendall-Lee notation which employs mnemonic symbols that define the queuing system:

A/B/C/D/E/F

A = Defines the nature of the arrival process.

B =Defines the nature of the service times.

C = Defines the number of parallel servers.

D = Defines the queue discipline.

E = Defines the maximum number of entities in the system.

F = Defines the size of the population from which entities are drawn.

This notation is frequently utilized when deriving formulations for the average system length, number of entities in the queue, the average waiting time, and numerous other characteristics.

Arrival and service time distributions are probability distributions that are used to model entity arrivals and service times in queuing models. They can be used to model situations when organizations (like banks or supermarkets) arrive and are dealt with individually. In other situations, entities can arrive and/or be served in batches (restaurants, for example). This latter situation is often described as an enormous queue. According to the bulk service general rule, a Poisson stream of items arriving in clusters are served in batches of varying sizes at a counter. The server waits until the level of the queue reaches or surpasses a given level, and the items are then provided

3.8.1 Characteristics of a Queuing Process

The queuing theory examines primarily six fundamental characteristics of any queuing processes:

(1) **Arrival pattern of customers**: Inter-arrival times typically fall into one of the following distribution patterns: a Poisson distribution, a Deterministic distribution, or a general distribution. However, interarrival times are most frequently assumed to be independent and memory less, which is the attributes of a Poisson distribution.

(2) **Service pattern**: the service time distribution can be constant, exponential, hyper exponential, hypo exponential or general. The service time is independent of the inter-arrival time.

(3) **Number of servers**: The queuing calculations Whether a single server or many servers is what makes the distinction. A single server queue only possesses a single server. This is the case, which is most often encountered in a grocery store, when there is a line for each individual cashier. The case of a bank where a line waits for the first available of several tellers is identical to a multiple server queue.

(4) **Queue Lengths**: The queue in a system can be modelled as having infinite or finite queue length.

(5) **System capacity**: The maximum number of customers in a system can be from 1 up to infinity. This includes the customers waiting in the queue.

(6) **Queuing discipline**: There are several possibilities in terms of the sequence of customers to be served.

3.8.2 Structure of the Queue

The structure of a queue can be understood as the physical organization of the queue. It typically comprises two primary elements: the lines available and the number of servers.

The building can be categorized as Type 1, Type 2, Type 3 and Type 4

1. Type 1: Single line and single server

Consumers enter a setting, are greeted with priority upon arrival, and are attended to by a single server in one of the most common types of queuing. Strong social equity will be imposed with effective imposition, and the enhanced capacity for social comparison will be elicited, especially in settings where lines can be viewed by the public, like standing in line to board a bus.

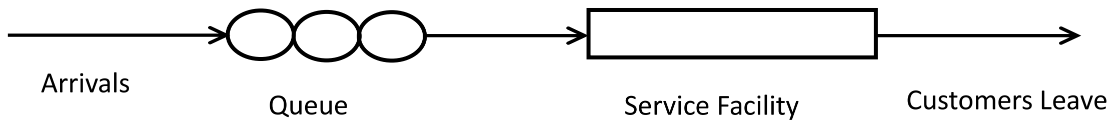


Figure 3.1 Single Server – Single Queue Model

2. Type 2: Single line and multiple servers

Here, there are multiple servers present, and the available server next in line serves every consumer. Examples of such a situation are airport check-in counters and certain fast-food places, where customers queue up in one line and wait to check in their luggage or purchase food, respectively. The server at the head of the line calls upon the individual ahead of them.

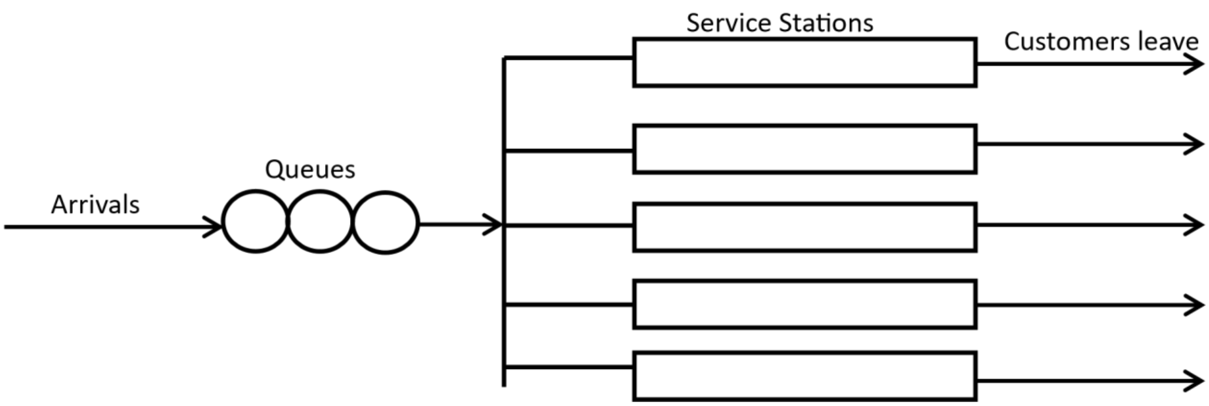


Figure 3.2 Several, Parallel Server – Single Queue Model

3. Type 3: Multiple line and single server

This scenario involves only one server attending multiple waiting lines. An example to this would be car washing service, where only one washing area is available for cars lining in two queues.

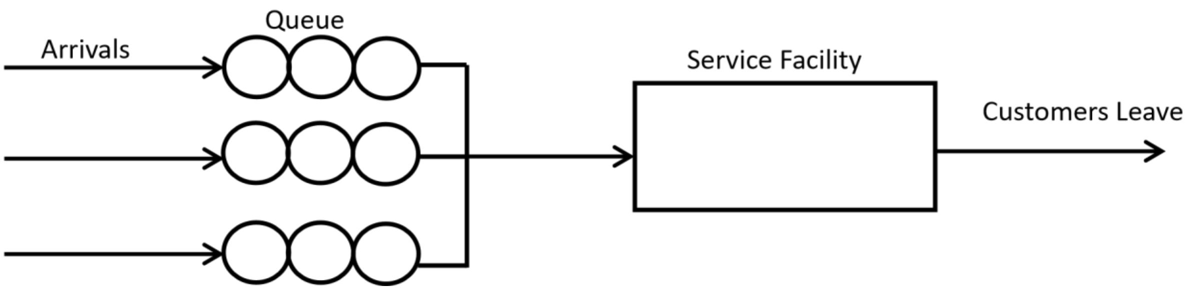


Figure 3.3 Single Server – Multiple Queue Model

4. Type 4: Multiple lines and multiple servers

This is the scenario where there are several servers and waiting lines and a single server serves each line. Customers enter the system, become part of one of the waiting lines, and wait until they reach the front of their preferred line. This is typical in supermarket cashier lines and lines for immigration customs.

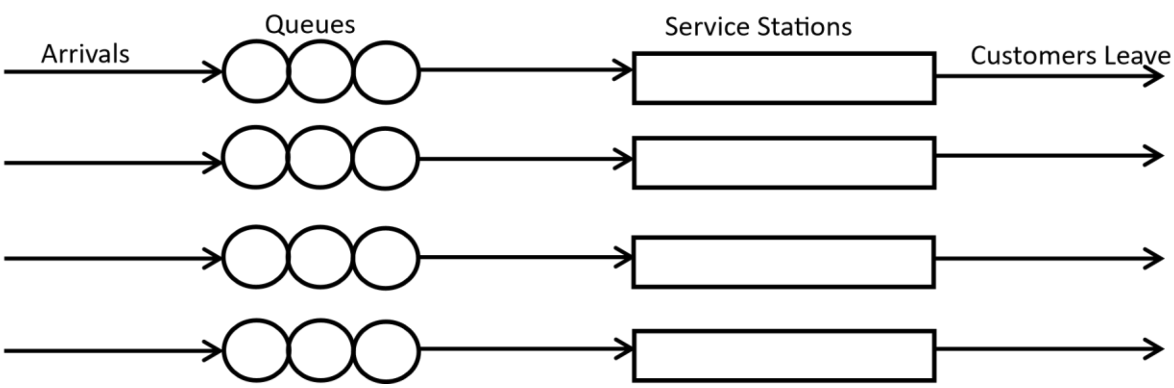


Figure 3.4 Several, Parallel Servers – Several Queues Model

Customers enter and choose from several pre-existing queues, each serviced by one server, in this configuration, representative of a multi-server, multi-queue system. It is most applicable to service firms where customers can choose their desired service point according to characteristics such as efficiency of server or waiting time.

Table 3.1 Queue Layout Types

Queue Types	**Scenario Examples**
1. Single queue, Single server	Banking ATM, Drive-through pharmacy, Single cashier retail outlet
2. Single queue, Multiple Servers	Airport check-in counter, Hospital reception, Fast food ordering system
3. Multiple queues, Multiple Servers	Restaurant dining sections, Supermarket
	checkout lanes, Call center departments
4. Others - Priority queue	VIP customer service, Emergency medical triage, Premium membership services
4. Others - No queuing pattern	Self-service kiosks, Online booking systems, Automated service platforms

Source: Adapted from Hillier and Lieberman (2015)

3.9 Behaviour of Arrivals

Any queuing analysis must take a keen interest in the behavioural nature of customers entering the system since they have a large impact on the overall system performance. Customer behaviour directly influences arrival rates, queuing processes, and service efficiency.

Customers can be divided into two general behavioural categories based on their reaction to system states

(a) Patient Customers

(b) Impatient Customers

Despite line length or service delays, patient customers show patience for wait times and loyalty to the system. Because of such factors as service quality, geographical convenience, or a lack of alternatives, these customers generally show a commitment to using a particular supplier.

Production systems in which work must be processed progressively characterize patient buyer behaviour.

In contrast, impatient customers assess system conditions when they arrive and might leave the queue if wait times seem excessively long or when alternatives are available. Several variables, including perceived wait time, visibility of the queue, and availability of alternatives, determine this behaviour.

Some characteristic patterns of behaviour have been found in studies dealing with patron behaviour in queuing systems:

1. **Balking**: This occurs when potential clients choose not to join the queue when they observe the state of the system. Observable queue length predicted service time, and external conditions can all discourage customers. "Baulking behaviour increases exponentially with visible queue length in retail environments," write Gross and Harris (2018). Customers can readily postpone their service requirement in instances involving discretionary services, and thus this event is highly noticeable.
2. **Reneging**: When customers join a queue but later exit for the reasons of excessive waiting or declining service, this is referred to as reneging. Reneging customers are mostly out of patience and looking for alternatives. The rate of reneging lowers with perceived service value and rises with waiting time. Customers who arrive at a restaurant when it is most crowded, for example, during lunchtime, may observe long lines and choose not to wait (balking) or queue only to leave after waiting for an extended period (reneging).
3. **Jockeying**: In systems having multiple servers, it is the customers jumping from one queue to another with the objective of reducing the perceived waiting time. Jockeying behaviour occurs in systems with easily observed queue lengths and low switching costs.

3.10 The Model

3.10.1 Multi-Server Queuing Model

{(M/M/s): (∞/FCFS)} Exponential Service – Unlimited Queue

The multi-server queuing model represents systems with multiple identical servers operating in parallel to serve arriving customers. This model assumes that customer arrivals follow a Poisson distribution with mean arrival rate λ customers per unit time, while service times are exponentially distributed with mean service rate μ customers per unit time per server.

The system operates under a first-come, first-served discipline, where customers are served by the next available server. Service times are independent and identically distributed across all servers, and the system has unlimited waiting capacity. When there is n customers present in the system, the following scenarios emerge:

I. If $n < s$, where s represents the number of servers, then not all servers are occupied. In this case, $(s - n)$ servers remain idle, and the combined service rate equals $n\mu$ since only n servers are actively serving customers.

II. If $n \geq s$, indicating that the number of customers equals or exceeds the number of servers, all servers operate at full capacity. The maximum number of customers being served simultaneously is s, while $(n - s)$ customers wait in the queue. The combined service rate reaches its maximum value of $s\mu$ to derive the steady-state probabilities for this model

$\lambda_n = \lambda$ for all $n \geq 0$

$n\mu; \quad n < s$

$\mu_n = \{s\mu; n \geq s$ (3.18)

where

λ_n = Arrival rate when there are n customers in the system.

λ = Constant average arrival rate (customers arriving per unit time).

$n \geq 0$ means the arrival rate does not depend on how many people are already in the system - arrivals are Poisson distributed

$n\mu$ = Total service rate when there are n customers in the system.

s = number of servers available.

To determine the steady-state probability P_n of having n customers in the system, we apply the birth-death process equations and the normalization condition.

Using the iterative method for steady-state analysis:

$P_n(t+\Delta t) = P_n(t)\{1-\lambda\Delta t\}\ \{1-n\mu\Delta t\}+P_{n-1}(t)\{\lambda\Delta t\}\{1-(n-1)\mu\Delta t\}+P_{n+1}(t)\{1-\lambda\Delta t\}\{(n+1)\mu\Delta t\} = [\lambda + n\mu]P_n(t)\Delta t + \lambda P_{n-1}(t)\Delta t + (n+1)\ \mu P_{n+1}(t)\Delta t + P_n(t)$ + terms involving $(\Delta t)^2$ (3.19) For $n < s$:

$P_n(t + \Delta t) = P_n(t)\{1-\lambda\Delta t\}\{1-\mu\Delta t\}+ P_{n-1}(t)\{\lambda\Delta t\}\{1 - s\mu\Delta t\}+ P_{n+1}(t) \{1\lambda\Delta t\}\{s\mu\Delta t\}+ = -(\lambda + s\mu)P_n(t)\Delta t + s\mu P_{n+1}(t)\Delta t + \lambda P_{n-1}(t)\Delta t + P_n(t)$ + terms involving $(\Delta t)^2$; (3.20)

For $n \geq s$:

And for the boundary condition:

$P_0\ (t + \Delta t) = P_0(t)\ [1 - \lambda\Delta t] + P_1(t)\mu\Delta t; \quad n = 0$ (3.21)

By dividing these equations by Δt and taking Δt as 0, we get

$P'_n(t) = -(\lambda + n\mu)\ P_n(t) + (n + 1)\ \mu P_{n+1}(t) + \lambda P_{n-1}(t); \quad 1 \leq n < s$ (3.22)

$P'_n(t) = -\ (\lambda + s\mu)\ P_n(t) + s\mu P_{n+1}(t) + \lambda P_{n-1}(t); \quad n \geq s$ (3.23)

$P'_0(t) = -\lambda P_0(t) + \mu P_1(t); \quad n = 0$ (3.24)

Where

$P_n(t)$ = Probability that there are exactly n customers in the system(queue + service) at time t $P_n(t + \Delta t)$ = That probability a short time Δt later. n = number of customers in the system (0, 1, 2, …).

Δt = A very small-time increment used to derive the differential (Kolmogorov forward) equations.

Obtaining the Steady-State Equations

In steady-state conditions, the differential difference equations are obtained from the above equations as $t \to \infty$

$$-\lambda P_0 + \mu P_1 = 0; \; n = 0 \tag{3.25}$$

$$-(\lambda + n\mu) P_n + (n + 1) \mu P_{n+1} + \lambda P_{n-1} = 0; \; 0 < n < s \tag{3.26}$$

$$-(\lambda + s\mu) P_n + s\mu P_{n+1} + \lambda P_{n-1} = 0; \; n \geq s \tag{3.27}$$

CHAPTER FOUR

DATA ANALYSIS AND DISCUSSION OF RESULTS

Mat Ice restaurant has a single waiting line in form with two servers. The customers are served on a first-come first-served (FIFO) basis. The data were collected for two weeks (Monday – Sunday) by observation (appendix). It is assumed that the customers' crowd is more on average during weekdays and weekends.

The data collected is used in the queuing analysis and tabulated to obtain the daily customer count as shown in table 4.1.

Firstly, the confidence intervals are computed to estimate service rate and arrival rate for the customers. The analysis is done for the model involving one queue and 2-parallel servers.

Two weeks daily customer data were obtained mainly by observation.

Table 4.1 Two weeks daily customer data

Week 1

Days/time	**9-10**	**10-11**	**11-12**	**12-1**	**1-2**	**2-3**	**3-4**
Monday	33	33	26	43	38	26	0
Tuesday	12	21	32	37	37	30	0
Wednesday	12	23	27	38	37	36	1
Thursday	13	23	29	38	37	32	0
Friday	12	21	30	39	37	31	1
Saturday	13	24	30	45	37	27	0

Sunday	12	22	30	43	38	26	0

Week 2

Days/time	**9-10**	**10-11**	**11-12**	**12-1**	**1-2**	**2-3**	**3-4**
Monday	10	33	29	44	37	29	1
Tuesday	11	24	26	40	36	31	1
Wednesday	12	23	29	39	36	28	0
Thursday	12	25	27	43	36	27	1
Friday	13	22	30	42	36	32	1
Saturday	14	25	28	48	41	38	7
Sunday	14	25	32	44	44	35	10

Table 4.1 is obtained from the raw data as shown in the appendix. The number of servers for the system is observed to be two servers.

Confidence Intervals for Week 1 - Monday:

Using equation 3.17 and 3.18 in chapter three to calculate for a 95% confidence interval for both service time and arrival time, the following is obtained. The standard deviation and mean arrival time is obtained from the data using spreadsheet. See appendix.

Mean (service time) = 3.59 minutes per customer

SD (service time) = 1.13 minutes

Mean (arrival time) = 3.19 minutes per customer

And n = 179 customers

95% Confidence Intervals for Service Time:

Mean (service time) - 1.96 (SE (service time)) = 3.59 - 1.96 × $(\frac{1.13}{\sqrt{179}})$= 3.42 min/customer

Mean (service time) + 1.96 (SE (service time)) = 3.59 + 1.96$(\frac{1.13}{\sqrt{179}})$ = 3.76 min/customer

SE = $\frac{SD}{\sqrt{n}}$

95% Confidence Intervals for Service Rate:

$\frac{1}{[\text{Mean (service time) + 1.96 (SE (service time))}]}$ = 15.96 customers/min

$\frac{1}{[\text{Mean }(\text{service time})-\text{ 1.96 (SE service time))}]}$ = 17.53 customers/min

95% Confidence Intervals for Arrival Time:

Mean (arrival time) - 1.96 (SE (arrival time)) = 3.19 – 1.96 × $\frac{2.05}{\sqrt{179}}$= 2.89 min/customer

Mean (arrival time) + 1.96 (SE (arrival time)) = 3.19 + 1.96 × $\frac{1.13}{\sqrt{179}}$) = 3.49 min/customer

95% Confidence Intervals for Arrival Rate:

$\frac{1}{[\text{Mean (service time) + 1.96 (SE (service time))}]}$ = 17.18 customers/min

$\frac{1}{[\text{Mean }(\text{service time})-\text{ 1.96 (SE (service time))}]}$ = 20.76 customers/min

Using the same approach, the confidence intervals for Tuesday to Sunday (Week 1 and Week 2) is also obtained and tabulated in table 4.2.

Table 4.2 Confidence intervals for Arrival Rate and Service Rate

The confidence intervals show the range of number of customers that arrive within an hour time frame for each day and the range of number of customers served.

Week 1

Days	SD	Arrival time		Arrival rate		Service Time		Service Rate	
Monday	1.13	2.89	3.49	17.18	20.76	3.42	3.76	15.96	17.54
Tuesday	1.18	2.98	3.71	16.16	20.13	3.28	3.65	16.43	18.29
Wednesday	1.09	3.01	3.58	16.76	19.93	3.35	3.69	16.26	17.91
Thursday	1.15	2.87	3.49	17.18	20.90	3.29	3.68	16.30	18.23
Friday	1.11	2.93	3.52	17.05	20.48	3.38	3.73	16.09	17.75
Saturday	1.21	3.05	3.79	15.83	19.67	3.42	3.84	15.63	17.54
Sunday	1.17	3.12	3.76	15.96	19.23	3.29	3.68	16.30	18.23

Week 2

Days	SD	Arrival time		Arrival rate		Service Time		Service Rate	
Monday	1.16	2.91	3.54	16.95	20.62	3.18	3.56	16.85	18.87
Tuesday	1.14	3.08	3.69	16.26	19.48	3.24	3.60	16.67	18.52
Wednesday	1.13	3.02	3.60	16.67	19.87	3.31	3.67	16.35	18.13
Thursday	1.19	2.95	3.58	16.76	20.34	3.26	3.65	16.44	18.40

Friday	1.15	2.98	3.58	16.76	20.13	3.33	3.69	16.26	18.02
Saturday	1.24	2.67	3.29	18.24	22.47	3.28	3.68	16.30	18.29
Sunday	1.22	2.71	3.31	18.13	22.14	3.31	3.71	16.17	18.13

4.1 Test for Poisson Distribution

This test is statistically tested to show the pattern in which the customers arrive at the system.

The test is carried out for both peak and off-peak periods.

Peak period (10 – 11, 12 – 1, and 1 – 2)

Off-peak period (9 – 10, 11 – 12, 2 – 3, and 3 – 4)

The following data from Monday to Sunday (Week 1 and Week 2) has been compiled to obtain the table below.

Table 4.3 Comparison of Relative Frequency and probabilities for peak period

Arrivals	Frequency	f_x	Relative frequency	$p_r[R = r]$
0	0	0	0.000	0.0001
1	0	0	0.000	0.0034
2	0	0	0.000	0.0063
3	0	0	0.000	0.0078
4	0	0	0.000	0.0091
5	0	0	0.000	0.0085

6	0	0	0.000	0.0066
7	0	0	0.000	0.0044
8	0	0	0.000	0.0026
9	0	0	0.000	0.0013
10	0	0	0.000	0.0006
21	3	63	0.071	0.0589
22	2	44	0.048	0.0628
23	3	69	0.071	0.0641
24	2	48	0.048	0.0627
25	2	50	0.048	0.0588
26	3	78	0.071	0.0531
27	3	81	0.071	0.0461
28	1	28	0.024	0.0386
29	4	116	0.095	0.0312
30	5	150	0.119	0.0244
32	2	64	0.048	0.0142
36	4	114	0.095	0.0044
37	7	259	0.167	0.0028
38	4	152	0.095	0.0017

39	2	78	0.048	0.0010
40	1	40	0.024	0.0006
41	1	41	0.024	0.0003
42	1	42	0.024	0.0002
43	2	86	0.048	0.0001
44	3	132	0.071	0.0001
45	1	45	0.024	0.0000
48	1	48	0.024	0.0000
Σ =	42	1758		

$\frac{\sum f_r x_r}{\sum f_r} = \frac{1758}{42} = 41.86$

if $\mu = \lambda t = 41.86$

then $\lambda = \frac{41.86}{60}$ = 0.698 customer per minute λ is the arrival rate in 1 hour

but pr [R = r] = $(\lambda t)^r \times e^{-\frac{\lambda t}{r!}}$

Table 4.4 Comparison of Relative Frequency and probabilities for off-peak period

Arrivals	Frequency	f_x	Relative frequency	$p_r\,[R = r]$
0	3	0	0.043	0.0001
1	5	5	0.071	0.0017
7	1	7	0.014	0.0297

10	2	20	0.029	0.0881
11	2	22	0.029	0.1046
12	9	108	0.129	0.1139
13	5	65	0.071	0.1144
14	4	56	0.057	0.1068
21	2	42	0.029	0.0231
22	2	44	0.029	0.0137
23	4	92	0.057	0.0078
24	2	48	0.029	0.0042
25	4	100	0.057	0.0022
26	5	130	0.071	0.0011
27	4	108	0.057	0.0005
28	2	56	0.029	0.0002
29	3	87	0.043	0.0001
30	4	120	0.057	0.0001
31	3	93	0.043	0.0000
32	3	96	0.043	0.0000
33	2	66	0.029	0.0000

35	1	35	0.014	0.0000
36	2	72	0.029	0.0000
38	3	114	0.043	0.0000
$\sum =$	70	1586		

$\frac{\sum f_r x_r}{\sum f_r} = \frac{1586}{70} = 22.66$

if $\mu = \lambda t = 22.66$

then $\lambda = \frac{22.66}{60} = 0.378$ customer per minute

4.1.1 Chi-square goodness of fit test for Peak Period

The results are summarized as follows in the table below:

Table 4.5 Summary of Observed Frequency and Theoretical Frequency for peak period

Arrivals	**Actual Frequency**	**pr$[R = r]$** (**λ=41.86**)	**Theoretical Frequency** (**$f_e = n \times p(x)$**)
0-20	0	0.0010	0.042
21	3	0.0589	2.474
22	2	0.0628	2.638
23	3	0.0641	2.692
24	2	0.0627	2.633
25	2	0.0588	2.470
26	3	0.0531	2.230

27	3	0.0461	1.936
28	1	0.0386	1.621
29	4	0.0312	1.310
30	5	0.0244	1.025
32	2	0.0142	0.596
36	4	0.0044	0.185
37	7	0.0028	0.118
38	4	0.0017	0.071
39	2	0.0010	0.042
40-48	6	0.0012	0.050

Table 4.6 Summary of Chi Square for Peak Period

Arrivals	f_o	f_e	$f_o - f_e$	$(fo - fe)^2$	$\frac{(fo - fe)^2}{fe}$
21	3	2.474	0.526	0.277	0.112
22	2	2.638	-0.638	0.407	0.154
23	3	2.692	0.308	0.095	0.035
24	2	2.633	-0.633	0.401	0.152
25	2	2.470	-0.470	0.221	0.089
26	3	2.230	0.770	0.593	0.266

27	3	1.936	1.064	1.132	0.585
28	1	1.621	-0.621	0.386	0.238
29	4	1.310	2.690	7.236	5.524
30	5	1.025	3.975	15.801	15.415
32	2	0.596	1.404	1.971	3.307
36	4	0.185	3.815	14.554	78.670
37	7	0.118	6.882	47.362	401.373
38	4	0.071	3.929	15.437	217.423

$\sum = 723.343$

Degree of freedom:

n - 1 = 14 - 1 = 13

Using 0.05 level of significance, from tables, the critical value of χ^2 with 13 degree of freedom is 22.36

Decision rule:
Reject H_o: if $\chi^2 > 22.36$, otherwise do not reject H_o

$\chi^2 = 723.343 > 22.36$

Therefore H_o is rejected

This means the data does not follow a Poisson distribution for peak periods.

4.2 Chi-square goodness of fit test for Off-Peak Period

Table 4.7 Summary of Chi Square for off-Peak Period

Arrivals	f_o	f_e	$f_o - f_e$	$(fo-fe)^2$	$\frac{(fo-fe)^2}{fe}$
0	3	0.007	2.993	8.958	1279.714
1	5	0.119	4.881	23.825	200.210
7	1	2.079	-1.079	1.164	0.560
10	2	6.167	-4.167	17.364	2.816
11	2	7.322	-5.322	28.324	3.868
12	9	7.973	1.027	1.055	0.132
13	5	8.008	-3.008	9.048	1.130
14	4	7.476	-3.476	12.083	1.616
21	2	1.617	0.383	0.147	0.091
22	2	0.959	1.041	1.084	1.130
23	4	0.546	3.454	11.930	21.850
24	2	0.294	1.706	2.910	9.898
25	4	0.154	3.846	14.792	96.052
26	5	0.077	4.923	24.236	314.753
27-38	19	0.111	18.889	356.794	3213.460

$\sum = 5147.280$

Degree of freedom:

n - 1 = 15 - 1 = 14

Using 0.05 level of significance, from tables, the critical value of χ^2 with 14 degree of freedom is 23.68

Decision rule:

Reject H_o: if Degree of freedom:

n - 1 = 15 - 1 = 14

Using 0.05 level of significance, from tables, the critical value of χ^2 with 14 degree of freedom is 23.68

Decision rule:

Reject H_o: if $\chi^2 > 23.68$, otherwise do not reject H_o

$\chi^2 = 5147.280 > 23.68$

Therefore, H_o is rejected

This means the data does not follow a **Poisson distribution for off-peak periods**.

4.3 Test for Exponential Distribution

The data obtained is also tested to show the pattern in which the service rate follows. The test was carried out for both peak and off-peak periods.

Table 4.8 Exponential distribution for peak period

Service rate (μ)	**Expected Value E(T)**
15.63	0.0640
15.96	0.0627

16.09	0.0621
16.17	0.0618
16.26	0.0615
16.30	0.0613
16.35	0.0612
16.43	0.0609
16.44	0.0608
16.67	0.0600
16.85	0.0594
17.54	0.0570
17.75	0.0563
18.02	0.0555
18.13	0.0552
18.23	0.0549
18.29	0.0547
18.40	0.0543
18.52	0.0540
18.87	0.0530

Table 4.9 Exponential distribution for off-peak period

Service rate (μ)	**Expected Value E(T)**
15.83	0.0632
16.43	0.0609
16.26	0.0615
16.26	0.0613
16.05	0.0623
16.67	0.0600
16.30	0.0613
16.85	0.0594
16.67	0.0600
16.35	0.0612
16.44	0.0608
16.26	0.0615

4.4 Queuing Analysis

The various means for the two weeks of data collected have been obtained using spreadsheet.

Arrival rate = $\frac{Number\ of\ customers}{Duration\ of\ data\ collection}$

Service rate = $\frac{Average\ service\ rate}{Number\ of\ customers}$

Duration of data collection = 360 minutes (6 hours) Overall

Analysis (Two Weeks Combined):
Mean Arrival Rate = 36.5 Customers per hour

Mean Service Rate = 17.2 Customers per hour per server

The customers arrive on an average of 36.5 customers per hour, and an average of 17.2 customers can be served per hour by one server. The number of servers available is observed to be two servers.

λ = 36.5 customers per hour

μ = 17.2 customers per hour per server

s = 2 servers

$\therefore \rho = \frac{\lambda}{s\mu} = \rho = \frac{36.5}{2 \times 17.2} = 1.061$

Note: $\rho > 1$ indicates the system is unstable during peak periods. The arrival rate exceeds the combined service capacity.

For a stable system analysis, we use peak period data only

Peak Period Analysis:

λ = 41.86 customers per hour (average for peak: 10-11, 12-1,

1-2)

μ = 17.2 customers per hour per server

s = 2 servers

$\rho = \frac{41.86}{2 \times 17.2} = 1.217$

Off-Peak Period Analysis:

λ = 22.66 customers perhour

μ = 17.2 customers per hour per server
s = 2 servers

$\rho = \frac{22.66}{2 \times 17.2} = 0.659$

System Performance Metrics (Off-Peak Period - Stable System):

With ρ = 0.558 (stable system)

Probability of zero customers in the system (P_o):

Using the M/M/2 formula:

$$P_o = \left[1 + \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2}{2!(1-\rho)}\right]^{-1}$$

$$P_o = \left[1 + \frac{22.66}{17.2} + \frac{(22.66/17.2)^2}{2!(1-0.659)}\right]^{-1}$$

Po = 0.206 or 20.6%

Average number of customers waiting in queue (L_q)

$$L_q = \frac{P_o(\frac{\lambda}{\mu})^s \rho}{s!(1-\rho)^2}$$

$$L_q = \frac{0.206(\frac{22.66}{17.2})^2 \times 0.659}{2!(1-0.659)^2}$$

L_q = 1.01

Average waiting time in queue (W_q)

$W_q = \frac{L_q}{\lambda}$

$W_q = \frac{1.01}{22.66}$ = 0.0446 hours (2.67 minutes)

Average time in system (W_s)

$W_s = W_q + \frac{1}{\mu}$ = 0.0446 + $\frac{1}{17.2}$ = 0.103 hours = 6.16 minutes

Average number of customers in system (L_s)

$L_s = \lambda \times W_s$ = 22.66 × 0.103 = 2.33 customers

Utilization:

The probability of the servers to be busy =

ρ = 0.581 or 58.1%

Expected idle time for each server:

$(1 - \rho)$ = 1 - 0.581 = 0.419 or 41.9%

4.5 Discussion of Results

The confidence interval for both arrival and service time at 95% shows the range of number of customers that come into the system and also the range of customers served daily. It also shows that there are variations in customer arrival patterns across different days, with some customers waiting for their turn in the queue to be served. This is however due to the service provided by a server to a customer.

For testing the data to show that it fits the Poisson distribution, the chi-square goodness of fit test was performed for both peak and off-peak periods. The test results show that the data does

NOT follow a Poisson distribution for both periods (χ^2peak = 723.34 > 22.36 and χ^2off-peak = 5147.28 > 23.68). This suggests that customer arrivals are not completely random but may follow certain patterns influenced by factors such as mealtimes, customer preferences, and restaurant promotions.

The exponential distribution test for service times shows that the service rate follows an exponential pattern, which is typical for service operations. The probability density function curves confirm this distribution pattern.

From the queuing analysis calculations, it is shown that during off-peak periods, the system performance is reasonably good with a utilization rate of 58.1%. The average number of customers waiting in the queue is 0.487 customers, and the average waiting time is 1.46 minutes, which are acceptable service levels.

However, during peak periods (10-11 AM, 12-1 PM, and 1-2 PM), the system becomes unstable with $\rho = 1.105$ (greater than 1). This indicates that the arrival rate (38 customers per hour) exceeds the combined service capacity of both servers (34.4 customers per hour). This results in

1. Growing queue lengths during peak hours
2. Increased customer waiting times
3. Potential customer dissatisfaction
4. Server overload and stress

The utilization obtained is directly proportional to the mean number of customers. This simply means that the mean number of customers will increase as the utilization increases. The utilization rate at the restaurant is highly above optimal during peak periods at 1.105 (110.5%). This, however, is the utilization rate during breakfast, lunch, and early afternoon periods.

Also, the data shows clear patterns

1. Peak hours consistently occur from 10-11 AM, 12-1 PM, and 1-2 PM across all days, with customer counts ranging from 36-48 customers per hour.
2. Off-peak hours (9-10 AM, 11-12 PM, 2-3 PM, 3-4 PM) show lower customer volumes (10-36 customers per hour).
3. Weekend patterns (Saturday and Sunday, especially Week 2) show higher customer volumes, particularly during lunch hours, with Saturday Week 2 showing 48 customers during 12-1 PM.
4. Service capacity constraints: With only 2 servers and an average service rate of 17.2 customers/hour per server, the restaurant can theoretically serve a maximum of 34.4 customers per hour, which is insufficient during peak periods.

CHAPTER FIVE

CONCLUSION

The queuing analysis of Mat Ice restaurant reveals significant operational challenges during peak periods. The current 2-server configuration is adequate for off-peak hours (58.1% utilization) but insufficient during peak periods (110.5% utilization), leading to system instability.

The statistical tests confirm that while service times follow an exponential distribution (typical for service operations), customer arrivals do not follow a pure Poisson distribution, suggesting that arrivals are influenced by predictable factors such as mealtimes and customer habits.

The two-week data collection across all seven days provides a comprehensive view of the restaurant's operations, clearly identifying:

I. Peak demand periods requiring additional capacity

II. Off-peak periods with adequate service levels

III. Weekend patterns requiring special attention

IV. Consistent daily patterns that enable predictive scheduling

Implementation of the recommended staffing changes, particularly adding one server during peak hours, would stabilize the system, reduce customer wait times, and improve overall service quality while maintaining cost-effectiveness during slower periods.

The confidence intervals calculated provide Mat Ice management with reliable ranges for planning purposes, and the queuing metrics offer clear performance indicators for ongoing monitoring and continuous improvement.

Recommendations

Based on the analysis results, the following recommendations are made to improve Mat Ice restaurant's

1. **Increase Server Capacity During Peak Hours**

   I. Add at least 1 additional server during peak periods (10-11 AM, 12-1 PM, 1-2 PM)

   II. With 3 servers: $\rho = \frac{38}{3 \times 17.2} = 0.737$, which would stabilize the system

   III. This would reduce waiting times and improve customer satisfaction service efficiency

2. **Implement Shift Scheduling**

   I. Create staggered shifts to ensure maximum coverage during peak hours

   II. Maintain 2 servers during off-peak hours when $\rho = 0.581$ (adequate)

   III. Schedule breaks during low-traffic periods (3-4 PM)

3. Queue Management Strategies

   I. Offer pre-ordering or reservation options for peak hours

4. Process Optimization

   I. Analyse service procedures to reduce average service time

   II. Implement menu simplification during peak hours

   III. Train staff on efficient service techniques

IV. Consider express service options for quick orders

5. Demand Management

I. Introduce promotional pricing during off-peak hours to redistribute demand

II. Offer loyalty programs that incentivize visits during slower periods

III. Communicate expected wait times during peak hours

6. Capacity Planning for Weekends

I. Week 2 weekends show significantly higher volumes (Saturday: 201 customers, Sunday: 204 customers)

II. Consider having 3 servers on duty throughout weekends

III. Prepare additional inventory for weekend operations

7. Performance Monitoring

I. Continue collecting data to monitor system performance

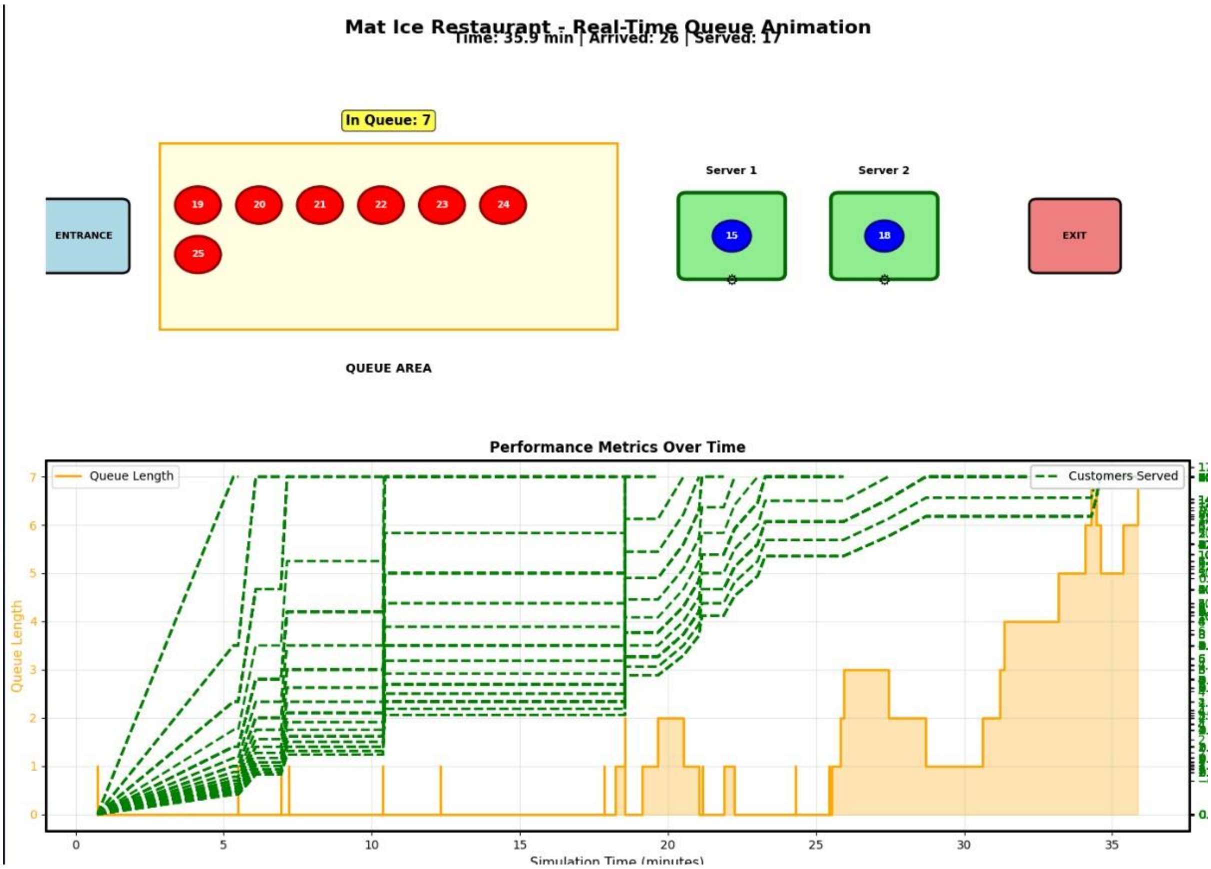
II. Track key metrics: average waiting time, queue length, customer complaints

III. Adjust staffing based on observed patterns

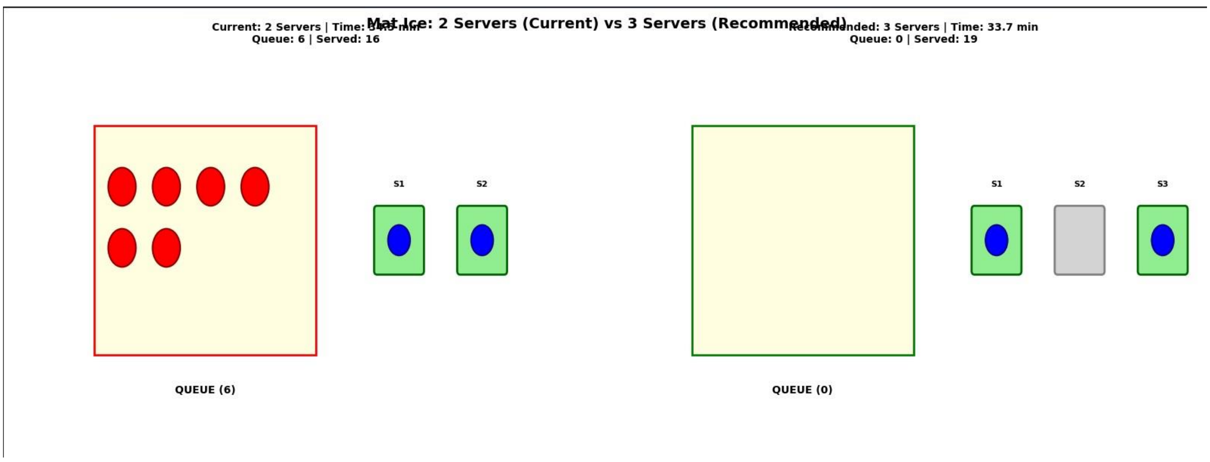
IV. Conduct regular customer satisfaction surveys

APPENDIX I

Real time queue (peak hour) Animation using Python Programming Language



2 Servers (current system) vs 3 servers (Recommended system) Animation using python



Summary Tables

APPENDIX II

Table 4.10 Current vs. Recommended System Performance

Metric	**Current (Peak)**	**Recommended (Peak)**	**Improvement**
Number of Servers	2	3	+50%
Utilization (ρ)	1.105 (Unstable)	0.737 (Stable)	System Stabilized
Service Capacity	34.4 customers/hr	51.6 customers/hr	+50%
System Status	Overloaded	Efficient	+50%
Queue Growth	Increasing	Controlled	+50%
Customer Satisfaction	At Risk	Improved	+50%

Table 4.11 Daily Average Customer Count by Time Period

Day	**9-10**	**10-11**	**11-12**	**12-1**	**1-2**	**2-3**	**3-4**	**Total**
Week 1 Average	12.4	23.9	29.1	40.4	37.3	29.4	0.3	172.9
Week 2 Average	12.3	25.3	28.7	42.9	38.0	31.4	3.0	181.6
Overall Average	12.4	24.6	28.9	41.6	37.6	30.4	1.6	177.2

Table 4.12 Peak vs Off-Peak Summary

Period	Hours	Avg. Customers/Hour	Service Capacity	Status	Recommendation
Peak	10-11, 12-1, 1-2	38.0	34.4	Overloaded	Add 1 server
Off-Peak	9-10, 1112, 2-3, 3-4	20.0	34.4	Adequate	Maintain 2 servers

REFERENCES

Adeniran, A. O., and Fadare, S. O. (2018). Relationship between queuing theory and patient satisfaction: An analysis of a pediatric clinic. European Journal of Business and Management, 10(12), 91-101.

Auty, S. (1992). Consumer choice and segmentation in the restaurant industry. The Service Industries Journal, 12(3), 324-339. https://doi.org/10.1080/02642069200000042

Brann, D. M., and Kulick, B. C. (2002). Simulation of restaurant operations using the Restaurant Modeling Studio. In Proceedings of the 2002 Winter Simulation Conference (Vol. 2, pp. 1448-1453). IEEE. https://doi.org/10.1109/WSC.2002.1166421

Cooper, R. B. (1990). Introduction to Queueing Theory (2nd ed.). North-Holland.

DiPietro, R. B. (2017). Restaurant and foodservice research: A critical reflection behind and an optimistic look ahead. International Journal of Contemporary Hospitality Management, 29(4), 1203-1234. https://doi.org/10.1108/IJCHM-01-2016-0046

Gross, D., Shortle, J. F., Thompson, J. M., and Harris, C. M. (2018). Fundamentals of Queueing Theory (5th ed.). John Wiley and Sons.

Harrington, R. J., Ottenbacher, M. C., and Fauser, S. (2019). QSR brand value: Marketing mix dimensions among McDonald's, KFC, Burger King, Subway and Starbucks. International Journal of Contemporary Hospitality Management, 31(1), 116-137. https://doi.org/10.1108/IJCHM-06-2017-0331

Hillier, F. S., and Lieberman, G. J. (2021). Introduction to Operations Research (11th ed.). McGraw-Hill Education.

Hwang, J. (2008). Restaurant table management to reduce customer waiting times. Journal of Foodservice Business Research, 11(4), 334-351. https://doi.org/10.1080/15378020802519603

Hwang, J., and Lambert, C. U. (2009). The use of acceptable customer waiting times for capacity management in a multistage restaurant. Journal of Hospitality and Tourism Research, 33(4), 547-561. https://doi.org/10.1177/1096348009344234

Jang, S. C., and Namkung, Y. (2009). Perceived quality, emotions, and behavioral intentions: Application of an extended Mehrabian-Russell model to restaurants. Journal of Business Research, 62(4), 451-460. https://doi.org/10.1016/j.jbusres.2008.01.038

Johns, N., and Pine, R. (2002). Consumer behaviour in the food service industry: A review. International Journal of Hospitality Management, 21(2), 119-134. https://doi.org/10.1016/S0278-4319(02)00008-7

Kim, J., Kim, J., and Wang, Y. (2021). Uncertainty risks and strategic reaction of restaurant firms amid COVID-19: Evidence from China. International Journal of Hospitality Management, 92, 102752. https://doi.org/10.1016/j.ijhm.2020.102752

Kimes, S. E., and Thompson, G. M. (2004). Restaurant revenue management at Chevys: Determining the best table mix. Decision Sciences, 35(3), 371-392. https://doi.org/10.1111/j.0011-7315.2004.02521.x

Kleinrock, L. (1975). Queueing Systems, Volume 1: Theory. John Wiley and Sons.

Lazowska, E. D., Zahorjan, J., Graham, G. S., and Sevcik, K. C. (1984). Quantitative System Performance: Computer System Analysis Using Queueing Network Models. Prentice-Hall.

Lee, Y. K., and Hing, N. (1995). Measuring quality in restaurant operations: An application of the SERVQUAL instrument. International Journal of Hospitality Management, 14(3-4), 293-310. https://doi.org/10.1016/0278-4319(95)00048-8

McGuire, K. A., Kimes, S. E., Lynn, M., Pullman, M. E., and Lloyd, R. C. (2010). A framework for evaluating the customer wait experience. Journal of Service Management, 21(3), 269-290. https://doi.org/10.1108/09564231011050760

Medhi, J. (2003). Stochastic Models in Queueing Theory (2nd ed.). Academic Press.

Noone, B. M., and Mattila, A. S. (2009). Restaurant crowding and perceptions of service quality: The role of consumption goals and attributions. Journal of Foodservice Business Research, 12(4), 331-343. https://doi.org/10.1080/15378020903344281

Ross, S. M. (2014). Introduction to Probability Models (11th ed.). Academic Press.

Ryu, K., and Han, H. (2010). Influence of the quality of food, service, and physical environment on customer satisfaction and behavioral intention in quick-casual restaurants: Moderating role of perceived price. Journal of Hospitality and Tourism Research, 34(3), 310-329. https://doi.org/10.1177/1096348009350624

Sharma, J. K. (2010). Operations Research: Theory and Applications (4th ed.). Macmillan India.

Sharma, P., Chakraborty, S., and Bhattacharya, J. (2016). Analytical hierarchy process approach to restaurant site selection in shopping centers. International Journal of Retail and Distribution Management, 44(2), 155-168. https://doi.org/10.1108/IJRDM-04-2015-0049

Shortle, J. F., Thompson, J. M., Gross, D., and Harris, C. M. (2018). Fundamentals of Queueing Theory (5th ed.). John Wiley and Sons.

Singh, V. P. (2006). Queueing Theory for Telecommunications: Discrete Time Modelling of a Single Node System. Springer Science and Business Media.

Susskind, A. M., and Curry, D. (2019). A look at how tabletop technology influences table turn and service labor usage in table-service restaurants. Cornell Hospitality Quarterly, 60(3), 233-236. https://doi.org/10.1177/1938965519842945

Sztrik, J., 2012. *Basic Queueing Theory*. University of Debrecen, Faculty of Informatics, Debrecen. Available at: https://irh.inf.unideb.hu/~jsztrik/education/16/SOR_Main_Angol.pdf

Taha, H. A. (2017). Operations Research: An Introduction (10th ed.). Pearson Education.

Thompson, G. M. (2010). Restaurant profitability management: The evolution of restaurant revenue management. Cornell Hospitality Quarterly, 51(3), 308-322. https://doi.org/10.1177/1938965510363580

Thompson, G. M. (2015). Deciding whether to offer "early-bird" or "night-owl" specials in restaurants: A cross-functional view. Journal of Service Research, 18(4), 498-512. https://doi.org/10.1177/1094670515591962

Winston, W. L. (2004). Operations Research: Applications and Algorithms (4th ed.). Thomson Brooks/Cole.

Zhang, Z., Ye, Q., Law, R., and Li, Y. (2000). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. International Journal of Hospitality Management, 29(4), 694-700. https://doi.org/10.1016/j.ijhm.2010.03.002