Operations Research
Applications of Queuing theory in a fast food restaurant

Saarthak Kanade, Pratyush Bokadia, Rishika Bansal, Rushil Gulyani, Saniya Jain

Student, NMIMS Anil Surendra Modi School of Commerce, Mumbai Campus

Abstract
Bob's burrito is a fast-food chain in Bath. Like every other restaurant, it aims to serve maximum customers with efficiency and find out how to achieve the end objective; we will use queuing theory, which will help us establish a thumb rule for the same. Queuing theory is a study of waiting lines. A 95% confidence level was used in the analysis. After mathematical analysis, we can establish the prediction of waiting time and queuing line, which will help Bob's burrito understand the pattern and work to serve maximum customers. The data used in the research paper was collected through surveys. The paper concludes with the limitations of queuing theory.

Keywords
Queuing theory, Restaurant, Understaffing, Overstaffing

Introduction

The operation we have decided to investigate is 'Bob's Burrito', where the task is to serve consumers their personalised burrito through offering a range of fillings. Like any business,' Bobs Burrito' aims to serve as many customers as they can in the most efficient way possible, with queue size, queue time and serving time being important indicators of how successful they are in performing this objective.

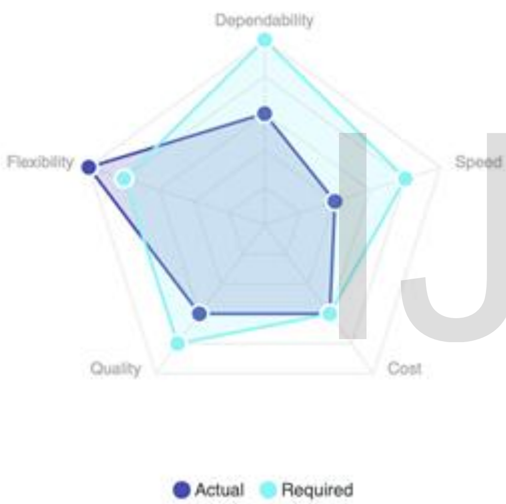

Fig. 1

From these objective performance maps, we can see that flexibility is a current priority due to the personalised aspect of each burrito; however, this is coming at a cost to the speed and efficiency of the operation, a problem we are looking to analyse. We believe these are the main challenges facing this operation and aim to find the root cause and find an appropriate solution so that 'Bobs Burrito' can fulfil its task to a more excellent capability.
According to a 1992 survey (Romsey), most customers, around 80-90 percent, would switch retailers merely to have shorter checkout lines. Another research conducted by Keurig and Brew Over Ice found that 65 percent of Americans cannot set aside time for themselves at least once a day, demonstrating the negative impact of having people wait.

Queuing theory is a branch of mathematics that studies and models the act of waiting in lines.
A simple queueing system is a service system in which "customers" arrive at a bank of "servers" and ask one of them for assistance. It is critical to remember that a "client" is anything that is waiting for service, and it does not have to be a human. (Green, p. 16) Queues, as we all know, are a regular occurrence. Because resources are limited, queues form. In reality, having lines makes financial sense. When building queueing systems, we must strike a compromise between customer service and cost concerns. All queuing systems may be split into individual subsystems, consisting of entities queuing for a specific activity.

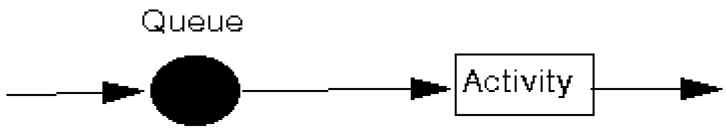

Fig. 2

This particular subsystem is typically associated with consumers queuing for service.

Literature Review

There are many metrics used to compare how good a restaurant really is; taste, cleanliness, ambience and restaurant layout are all contributing factors. However, none are more important (especially when the other goals have been achieved) than queuing time. Even with great food and service, one may avoid going to a restraint or end up with a bitter taste if the service and queuing time took too long. Using queuing theory, one can derive analysis for the expected waiting time in queue, the length of the queue, probability of the restaurant to be full or empty, the probability of a balking customer, average time in the system and expected serving time. A common sight in restaurants is waiting lines, queuing theory analyses and helps find the optimal system for queuing for customers who cannot be served immediately have to queue.

Little's theorem can be explained as the connection between throughput rate cycle time and work in progress. Little's theorem gives us the formula: $L=\lambda T$.

Where λ is the mean rate of customer arrival, T is the overall time for customer service, and L is the restaurant's total population (customers).

From Little's theorem, we derive three fundamental relationships between the given variables.

L increases if λ or T increases.

λ increases if L increases or T decreases.

T increases if L increases or λ decreases.

These derivations explain that to control the three variables, only two need to be controlled. Using these variables, a restraint can optimise and increase its overall customer output and decrease wait times. For example, If customer service time can be doubled while maintaining customer rate, the total number of customers in the restaurant will be doubled. (Patidar, Abhishek, Bisoniya, & Ray, 2014)

The main focus for Kendall's notation is the customer and the server. While in a queue, "human" customers may either jockey between lines or baulk from joining due to the length. In most cases, the behaviour can be characterised under six factors:

- Arrival time distribution: falls under Poisson distribution, Deterministic distribution or General distribution.
- System capacity: can range between 1 to infinite.
- Service time: distribution: can be constant, exponential, hyper-exponential, hypo-exponential or general.
- Number of servers: vary based on whether there are multiple servers (a single line awaiting multiple cashiers to be free) or a single server (individual line for each cashier)
- Queue length (optional): can be finite or infinite.
- Queuing discipline (optional): sequencing can be First in first out, last in first out, or service in random order.

The decisions taken by the customers also affect the waiting time and the functioning of a restaurant and in turn affects the queuing theory. Customers may decide to balk from joining the queue if they anticipate a long delay and want to avoid getting in one. Many times customers also jockey the queues, which means they switch between queues in the hopes of being served faster and avoid the extra waiting time. Customers also renege the queue, which means they leave the queue if they have waited for too long for the service. (LÓPEZ, 2016)

The quality of a restaurant depends on several factors like food quality, ambience, location, atmosphere, taste, cleanliness, reputation. Consumers use various variable attributes in order to decide on the type of restaurant they want to visit. Food quality, however, plays a significant role in affecting the diner's intent to return to a restaurant. Nowadays, consumers are following the health concern trend and have become very conscious about the quality of food they are being served. Quality service also plays a considerable role while choosing a restaurant. Many fast-food chains are better than their rivals solely

ISSN 2229-5518

because of the difference in the service quality offered. In addition, fast food restaurants are attempting to provide their customers with good food and good services in a favourable restaurant environment to gain competitive advantages in today's markets (Yeoh). Nowadays, customers who visit fast-food restaurants expect quality, value, and a desirable environment away from the pressure of daily life (Ribeiro Soriano, 2002).

Analysis derived from the queuing theory is waiting time, the average time in the system, expected queue length. Constant, exponential, hypo-exponential, hyper-exponential, or generic service time distributions are possible. The service time is unaffected by the interval between arrivals. A system's queue can be described as either infinite or finite in length. This includes consumers who are waiting in line.

Furthermore, regardless of peak or off-peak hours, the number of waiters or waitresses stays constant. The number of clients who can be served per minute will rise if the customer waiting time is reduced. When the service rate is greater, the utilisation rate is lower, which reduces the likelihood of consumers leaving. (Bhavani & Jayalalitha, 2021)

Restaurants are notorious for having long queues, especially during lunch and supper. As a result, queuing theory is appropriate for use in a restaurant environment. It has an accompanying queue or waiting line where clients who cannot be served quickly must queue (wait). Queuing theory has previously been used to simulate restaurant operations, minimise cycle time in a crowded fast-food restaurant, and enhance throughput and efficiency. This study helps to stabilise the system from the way of working in a restaurant. The results from this study might be used as a starting point for analysing and improving the present system. Because the restaurant can now predict how many people will queue and how many will leave each day. The restaurant may establish a daily profit objective by anticipating the many guests who will come and depart during the day. The formulae used to complete the study are suitable for future research and might also build more complicated theories. This study aims to demonstrate the utility of using queuing theory in a real-world scenario. (Bhavani & Jayalalitha, 2021)

Research objectives-

- To analyse the problem of understaffing and overstaffing
- To evaluate whether the level of customisation they currently offer is necessary
- To identify the average waiting time in a restaurant

4VS Analysis

Volume: Medium to low volume of products demanded and supplied; currently ~100 per day, with an estimated adequate capacity of at least 150 - but unrealised due to understaffing at peak times.

Variety: High variety in end products due to large number of separate customizability opportunities given to customers (7); however, it is unclear whether there is customer demand for so much variety.

Variability: Demand varies minimally between days/weeks but is highly variable throughout the day. There are several sharp peak times where large queues form and between which demand remains low and constant.

Visibility: Very high. Within this specific micro-operation, we are analysing all processes in the front of house, meaning complete visibility to customers.

Overstaffing and Understaffing

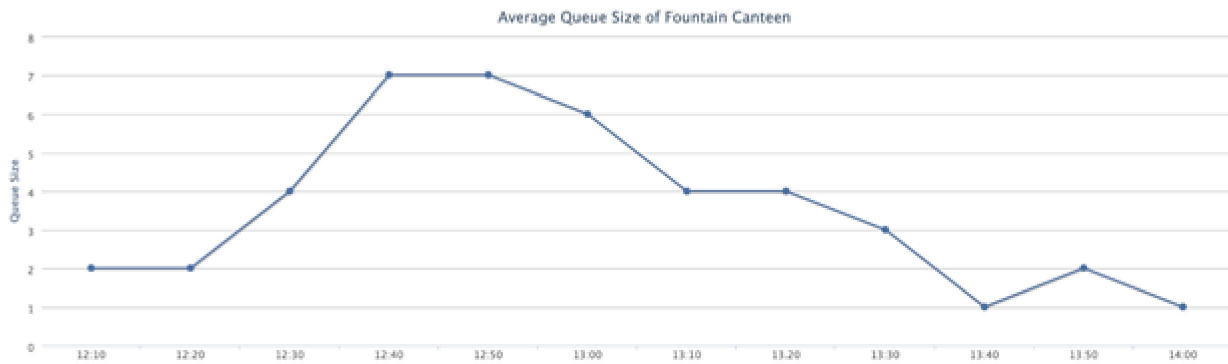

Fig. 3

Initially, to help us analyse how well this operation functioned, we gathered data on the queue size, time and serving time, and we found drastic fluctuations. The data above shows the average queue size at different 10-minute intervals that we gathered over 3 days. From this, we can see that 12:40-13:00 was peak times, with a queue size of around 7. However, this sharply drops off, and the queue size maintains under 4 people everywhere else. Despite a clear cycle of peak and off-peak times, we found the number of workers present maintained consistent despite the changes in demand.

From this Bottom-Up Perspective, we can assume that resources are not being distributed equally, with too many staff working off-peak hours and too little working in peak hours. This issue of overstaffing and understaffing can be supported further by queue times, as the graph below shows queue times rose rapidly in peak times, most likely due to understaffing. This data highlights an inefficiency problem.

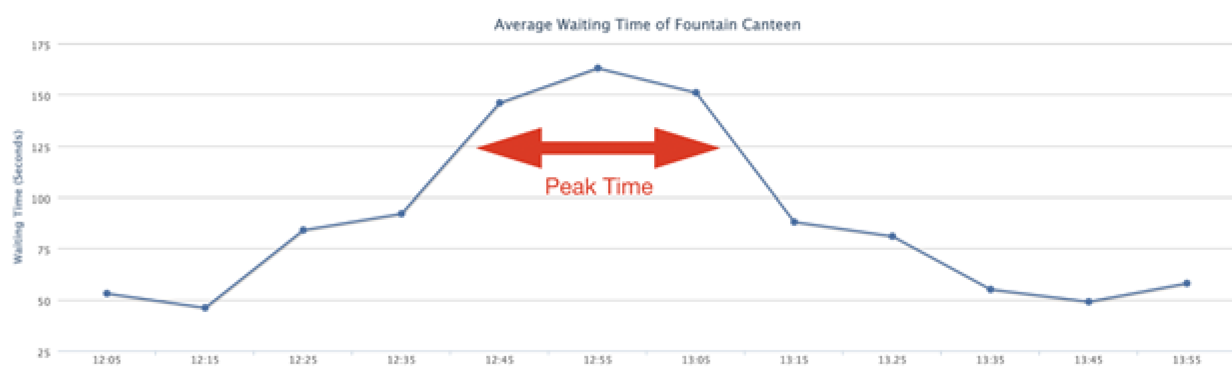

Fig. 4

Product Standardisation

In order to achieve the first objective of faster service time, we believe further standardisation of products offered to customers is necessary. Learning from the example of a leader in the fast-food industry - McDonald's – we see that this operation is suffering from setting its product specification too high, whilst its conformance is far too low.

ISSN 2229-5518

This result is long, unpredictable product assembly (to order) and, therefore, slow service times.

Instead, the overall quality should be derived from high conformance to a reasonably set bar of specificity. The core products in themselves would (by our recommendations) remain unchanged, but the level of variety available drastically reduced to a menu of 4-8 options. Through this, the products could be completely assembled and ready for the order before the shift beginning, effectively moving the capacity planning approach from lag (current) to lead. This achieves 3 of the significant contributors to McDonald's success in the fast-food industry: speed focus, standardised menus and consistent end-product quality.

The context of our operation may differ from that of McDonald's. A possible USP of the case company's product is the customizability aspect – a lesser focus in McDonald's - so we see that there are limits to applying its approach as an ideal. However, our operation is still a fast-food service, which is why we believe that the variability of the product is worth sacrificing in favour of faster service time. The need for speed across this industry is further demonstrated in an operational analysis of a struggling franchise in India. Here again, we see that in 'determining what will satisfy the customer', the primary focus of operations management should be on 'shorter time for food delivery' and 'fast order taking' (Agnes Kanyan, 2016).

Fewer options on the menu

One of the problems that arose through observation was the length of the serving time. The average serving time during the off-peak time was 1 minute 48 seconds, and the peak time was 2 minutes 31 seconds during the observation. There is a vast difference between serving time during peak time and off-peak time. Through observation, most of the customers who have ordered burritos took some time choosing the sides and optional fillings. These pauses cause bottlenecking, which delays the whole service operation massively in the current type B queue system. To solve this problem, the menu will be standardised (i.e. pre-set items, rather than fully tailored) and reduced to fewer options. The survey conducted towards the workers has shown that the most popular menu was "pork Mexican fried rice beans with fries". There is pulled chicken as a halal option and multiple vegan options such as "chimichurri veg" and "Piri Piri rice".

Along with a small selection of the most popular combinations, this would provide sufficient choice to satisfy consumer demands whilst achieving greater service time through pre-made food products and no queue bottlenecking due to customer choice delay. By having fewer options, less preparation, maintenance, and space will be needed. This would improve the cost and speed aspect of the operations performance.

Fig.5. options currently available

Type B Queueing

We were aware that the queue size and time for 'Bob's burrito' was taking too long for the business to run a time-efficient operation. The queue size averages at 3.58 persons and 6.6 during peak time which happens around 12:40-13:00. The business only has one till to get the orders. Thus a long line forms when the people are waiting.

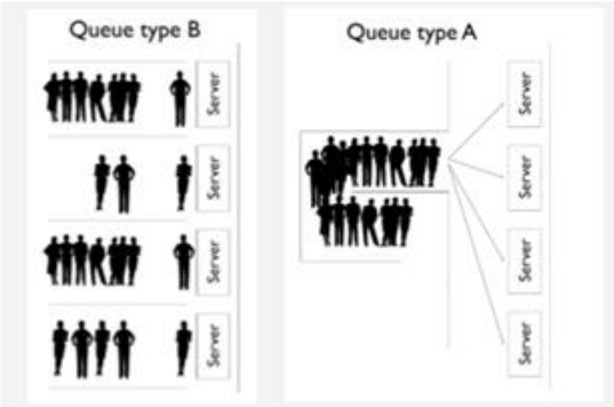

Fig. 6

Moreover, multiple servers depend on the same till to take orders which means that they have to wait for the first server to finish taking his order before doing so. This is a substantial loss of time because many servers are just waiting to serve customers but can not do that rapidly due to the setting of the place. The food outlet owns two tills, one to order and one to pay. Each server takes the order on till 1. He then makes the burrito and finally processes the payment on till 2. Many movements are going back and forth between tills; this decreases the efficiency of this operation which is making and serving burritos to customers.

A way to remediate this loss of time would be to change the queue type, which seems to be based on an A-type model: where everyone waits in the same queue and waits for the next server to be available, which means waiting for the till to be available for a server. To a type B queueing, where there is a queue to both tills (which need to process orders and payments), hence servers can simultaneously take orders from different customers, splitting the waiting time by two because there is no more dependency on the other server to finish his order first before processing server's 2 order. Ideally, using the additional till will allow an extra customer to be served instead of the idle worker waiting for the server to finish making the burrito or taking payment. While server 1 is making the burrito of the current customer in queue 1, server 2 can take the next customer's order in queue 2. Once finished, server 1 can take payment from the customer from queue 1. In contrast, server 2 makes the queue 2 customers burrito. Utilising the Queues in such a way can potentially half queue times.

Proposition

From the data collected, we believe that Bob's burritos focus on flexibility is inefficient. We identified the vast offerings, which over-complicated customers decisions and led to longer queues. Despite this, there were around four standard selections from the choices available. To achieve efficiency and speed objectives, we recommend cutting down Bob's burrito's offerings to 4 choices. Whilst reducing flexibility, they will be able to increase speed and reduce queue times. This is particularly important in the Fountain canteen as 43% of customers sampled stated that queue length had a considerable impact on their choice of where to eat. Showing speed is an order winner and provides a competitive benefit. As a result of these changes, quality would be derived from conformance while our core product offering would remain unchanged.

Clumping of demand and queueing becomes a massive problem around peak times of 12:40 to 13:00. As a result, we suggest moving to a type B queue in which both tills present are utilised. We are essentially increasing capacity by utilising the workforce more efficiently as 2 workers would be taking orders and making burritos simultaneously instead of having one idle. As a result of this change, they should deal with the increased utilisation at peak times and void queue build-ups. When in off-peak times, only one till should be open.

We would also introduce menu boards at the top of the burrito stall instead of small labels on the counter. This will allow customers to decide what they want before they are served, further reducing queue times. Moreover, in queues, occupied time feels shorter than unoccupied time.

ISSN 2229-5518

In addition to this, to solve the overstaffing and understaffing inefficiency problems, Bob's burritos should change worker shifts. We would favour a 1-hour shift from 12:00 to 1:00 in which 3 workers were present. Despite the unconventional shift pattern, this would suit university students.

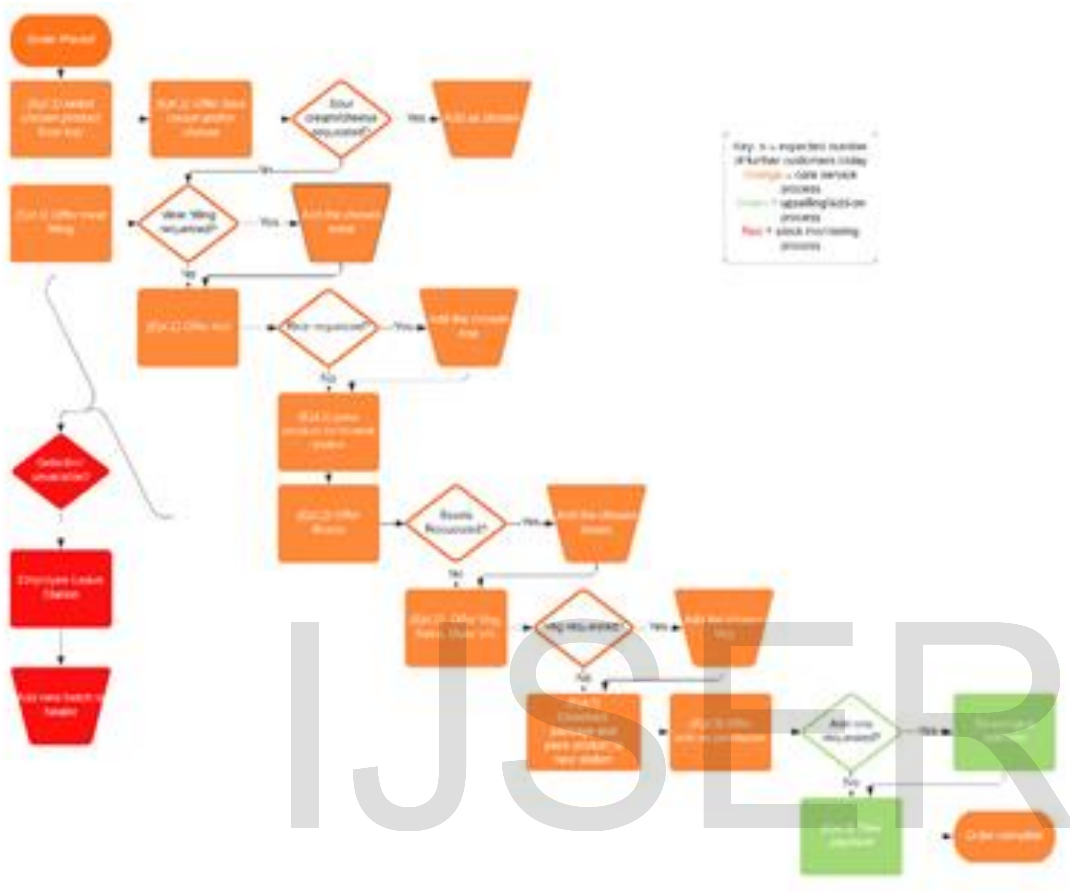

Fig. 7

Process Flow Map 1, showing current operations (limited to the specific micro-operation of food service)

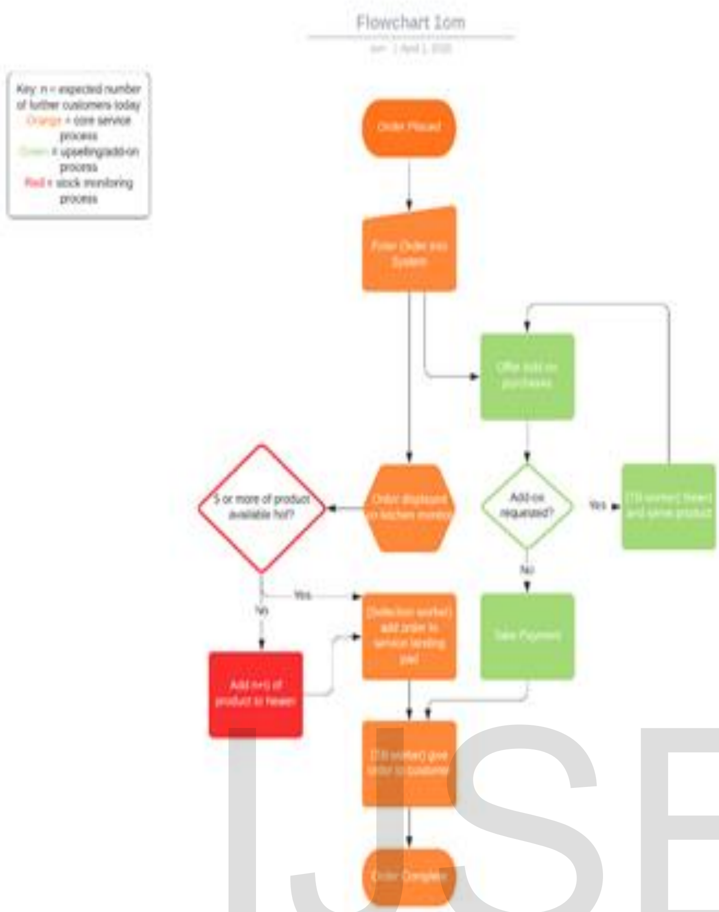

Fig. 8

Process Flow Map 2, showing our recommendation for a future operation with faster service time

The process flow map displays a simplified process that would save time, resources and, as a result, improve the overall operations process.

Conclusion

Our analysis aims to improve the speed and efficiency of Bob's Burrito. Furthermore, we wanted to find a solution for the potential issue of understaffing for the same. Analysing and surveying the customers and staff gave us valuable insights into the operations and management issues of the chain. The major hindrance for applying our recommendations from our analysis was that, unlike other fast-food restaurants, the USP of Bob's burrito might be the variability and flexibility in options provided, hence our recommendations may not be a perfect science.

Queue theory is a significant operations research tool used to systematically organise queues of people, information or objects through a line. Its application to fast food chains is inexplicably essential, 43% of sampled customers stated waiting time had a considerable impact on where to eat. Queuing helps reduce wait time and appropriately manage queues, which answers our main question on improving efficiency and speed at the restaurant.

Per our analysis, we recommended switching to a type B queue system consisting of two tills from a 'type A' model would considerably reduce queue size and wait time. From having an idle staff member waiting for another to finish to having, two members working simultaneously and switching the capacity planning approach from a lag to a lead. We also observed that indecisiveness while ordering caused longer service times, adding boards of the options available rather than small labels at the serving station. Lastly, we also recommended having an additional employee for the 12:00 to the 13:00 shift as it is the peak time and will help solve understaffing.

After extensive analysis of the queuing model that a restaurant should follow, this study has achieved the objectives it set out to accomplish. Furthermore, other research methods can be utilised to better this using computational analysis to simulate scenarios, employing the various technological advances that have been made to this date. Although the study encompasses the queuing model, further work with detailed analysis on different retail or commercial sectors can enhance queueing theory's multiple applications and allow precise results to surface.

While the assumptions mentioned in the queueing theory model are limited, some of them are slightly irrational. Concerning human queues, some assumptions cannot hold in an everyday scenario. It is presumed that human behaviour is deterministic, based on probability, assuming specific rules that a person might follow. (Berry) One apparent constraint is the potential that there may be a limited amount of waiting for space. Another hypothesis is that the arrival rate varies by state. For instance, if there is a long line when potential consumers arrive, they are discouraged from joining it. The arrival process is not steady, which is another practical drawback of the concept. The service station will likely have peak and slack periods when the arrival rate is higher. Real-life situations are typically complicated and go beyond philosophy and mathematics, so there will always be some uncertainty, no matter how precise you are.

Works Cited

Agnes Kanyan, L. N. (2016). Improving the Service Operations of Fast-food Restaurants,. *Procedia - Social and Behavioral Sciences*, 190-198.

Berry, R. (n.d.). Queuing theory. 14.

Bhavani, L. K., & Jayalalitha, G. (2021, May 08). Applying Queuing Theory to Enhance the Service Provided by A Restaurant. *Annals of R.S.C.B., ISSN:1583-6258,, 25*(6), 4479-4484. From https://www.annalsofrscb.ro/index.php/journal/article/view/6278/4767

Green, L. (n.d.). Queuing theory and modeling. From https://www0.gsb.columbia.edu/mygsb/faculty/research/pubfiles/5474/queueing%20theory%20and%20modeling.pdf

LÓPEZ, G. O. (2016). Modeling and dimensioning of the. 56.

Patidar, L., Abhishek, A., Bisoniya, T. S., & Ray, K. P. (2014). Queuing Theory: A Case Study to Improve the Quality Services of a Restaurant.

Ribeiro Soriano, D. (2002). Customers' expectations factors in restaurants: The situation in Spain. 1055-1067.

Yeoh, C. K. (n.d.). Intention to revisit fast food restaurant in Kota Kinabalu. *Unpublished master's thesis, Universiti Malaysia Sabah.*