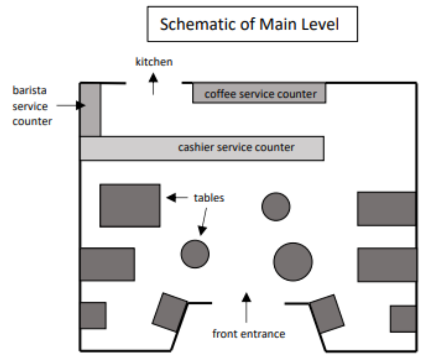
Introduction

XYZ Café specializes in breakfast, brunch, and lunch with menu items ranging from sandwiches, soups, and bakery products to coffee and specialty drinks. The café is particularly busy from late morning through lunch time. In the interest of growing and maintaining a strong customer base through high quality food and service, we are examining aspects of the operation that create queues, which in turn causes customers to wait and potential customers to balk, resulting in lost sales.

Seating is not an issue, but there is limited space on the main floor of the café, which also contains the waiting line to both the cashier and barista, in the case of customers waiting for their specialty drinks (see Figure 1). Thus, the main floor can easily get congested, giving the impression to customers that they will be waiting long for their order, motivating balking.



The focus of this report will be on the front-end operations of the café. An excel spreadsheet of data collected on business days between January 10th and 28th and a menu was provided by the owners, along with the following details:

1) Currently one barista and one cashier are employed at a rate of $16/hour.
2) The busiest time is from 10:30 am to 1:00 pm.
3) Regular coffee and hot water for tea are dispensed by the cashier without requiring any additional time compared to orders without these items.
4) The barista makes specialty drinks one at a time in the order they arrive. This does not include coffee, tea, soda, or juices.
5) Customers commonly arrive in parties and order separately. As tables at the establishment can accommodate a max of 4, the spreadsheet will have a max of 4 for party size.

The objectives are as follows:

1) Develop a model to investigate and predict queue lengths, wait times and prevalence of balking.
2) Perform a cost-benefit analysis of the possibility of adding a second cashier to process orders faster.
    a. Alternatively, have the barista act as a cashier between specialty drinks.

For objective 1, several parameters will have to be outlined, such as arrival rate, party size, cashier queue, balking, order types, order duration, pricing of orders, and barista duration. Once these parameters are properly understood, we will replicate novel data for each parameter individually, then bring them all together to simulate a single average day.

After objective 1 is finished and the model is validated, a cost-benefit analysis will be run based upon a statistical analysis of the three novel models. The statistics will include average parties served, balking, maximum queue lengths, and more. These statistics will be generated from many Monte Carlo simulations with varying random variables, which will be further outlined below.

Methods

To examine the objectives outlined, models needed to be created to examine three distinct scenarios:

1. A day with one cashier and one barista.
2. A day with two cashiers and a barista.
3. A day with one cashier and a barista who can double as a second cashier.

These models are stochastic in nature, using several random variables to generate data for projected outcomes. These random variables were based on parameters that arose from analyses of the data provided by XYZ Café; specifically, arrival rate, party size, cashier queue, balking, order types, order duration, pricing of orders, and barista duration. These parameters will be detailed in turn.

Arrival Rate

The data provided was collected for three weeks, Monday through Friday, from 10:30 am to 1 pm. The parties were recorded as arriving between the start at 10:30 am (0^{th} minute) to 1 pm (150^{th} minute), at a precision of .01 minutes. Several steps were taken to discover the arrival rate of parties over the 15-day period.

First, the interarrival period for every party (after the first) was calculated. This was done by taking the difference between a party's arrival time and the previous party's arrival time, followed by calculating the mean of each day's interarrival times. This showed how often, on average, a party arrived each of the 15 days. Then, the average of the three samples of each weekday was taken to yield a single interarrival mean for each day of the week.

The first Monday's mean interarrival period was approximately 4.61 minutes, next Monday 2.97, and the third Monday 3.69. These combined for an average Monday interarrival period of approximately 3.64 minutes. The averages for each weekday are as follows (rounded to two decimal points).

- Monday: 3.64 minutes
- Tuesday: 4.07
- Wednesday: 3.55
- Thursday: 4.63
- Friday: 3.46

Table 1: Parties per Day

	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	33	32	41	29	44
Week 2	50	44	34	34	40
Week 3	38	37	54	33	48
Average:	40.3333	37.6667	43	32	44
Standard Deviation:	8.73689	6.02771	10.148892	2.64575	4
Range:	17	12	20	5	8

From this list, it appears that Friday is the busiest day, Thursday the least. Table 1 shows that the average number of parties to arrive on the three Fridays surpassed the Thursday average by 12 parties. Furthermore, the standard deviation of Thursday was only 2.64, so there is little variation among the days recorded. With the data provided, Thursday seems less busy than Friday, on average. With only three days recorded for each weekday, however, it would be tenuous to conclude such from the data provided. A much larger sample (e.g., all non-holiday weekdays of a year) would be far more reliable. Given the small sample size and relatively close interarrival average, the assumption was made that the days will have the same arrival rates.

The list of average interarrival periods for each day is helpful, but a model would require predicting how often parties tended to arrive within the given time frame. Interarrivals were presumed to follow an exponential distribution; after viewing the data compiled into a histogram (Figure 2), it was concluded as such. Knowing this, a parameter for the arrival rate, represented by lambda (λ), could be established. This was done by inverting the mean time of all the interarrival periods (Table 2).

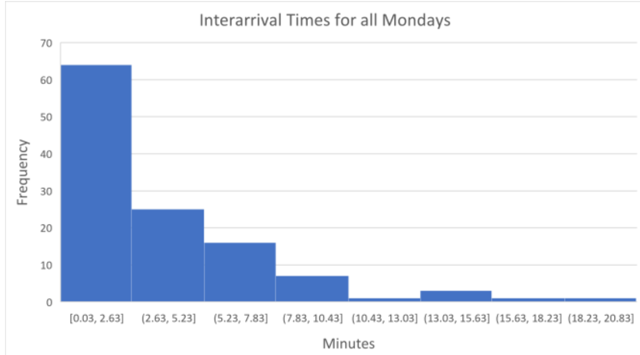


Figure 2

Party Size

The café's tables seat four customers at most, thus every party that arrives is counted in sets of four. To simulate the data, it is necessary to learn the arrival frequencies of parties sized one, two, three, and four. However, one potential problem arose with such a straightforward calculation. If a party of six arrived, for example, they would be recorded as two separate parties with the latter recorded as arriving at roughly the same time as the former. If this hypothetical party of six arrived while there was, say, 3 people already in the queue, a portion of the party would be added into the queue as one party, then the second segment of this party would see a large queue size, but realistically have no chance of balking, which had the potential of throwing off the later analysis. Trace evidence would be the only way to tell if there were instances that were indeed a party of six. Thus, the 591 were combed for parties that were possibly larger than four; every instance went as follows:

Table 2: Mean Arrival Rates		
Weekdays	Mean Interarrival Times	Arrival Rate (λ)
Monday	3.637	0.275
Tuesday	4.074	0.245
Wednesday	3.553	0.281
Thursday	4.632	0.216
Friday	3.463	0.289
Average:	3.872	0.261

On the second Monday (17 January), there was a party of four and a party of two that arrived in the same minute (the 106^{th}). Although, the second party balked, leading to conclude that they were likely not a party of six. On the second Tuesday (18 January), a party of four and a party of three arrived in the first minute of the day. The party of three balked, leading to a similar conclusion. On the first Friday (14 January), a group of four was followed by a group of two that balked. The conclusion was that these three instances were not parties of six.

On the second Tuesday (18 January), parties of four and two arrived only 0.14 minutes apart, which is roughly 8.4 seconds. Accounting for time to open and walk through the doors, this could have been a party of six. Parties of four and three showed up within the first minute of the first Wednesday (12 January), which might have been a party of six. Although, the party of three was nearly 30 seconds behind the first party, which could just as probably mean they are unassociated. Similarly, on the first Friday (14 January), parties of four and two arrived within 0.73 minutes of each other. These two groups might have been associated, but it is at least equally possible that they are unassociated. On the second Friday (21 January), parties of four and one arrived within 0.12 minutes (7.2 seconds) of each other. It is quite possible this was a party of five.

Thus, there were a total of seven occurrences in which a group of four arrived and was followed within one minute by a smaller group. Three were dismissed such that they were not associated with the larger groups of four. It is possible that four groups of six arrived, although it seems probable that two of these four were unassociated. With only two to four out of a total of 591 parties (0.34 to 0.68 percent) that arose in the three weeks of recording, it was concluded that the data was too vague to make any solid conclusions towards larger party sizes, and the few instances where the possibility arose were so few that it could be safely assumed that the parties ranged of sizes one through four. As such, the total frequencies of each party size across the data set were as follows:

- Parties of one: 30.80%
- Parties of two: 38.99%
- Parties of three: 22.60%
- Parties of four: 7.61%

The data are represented graphically by day (1 through 15) in Figure 3.

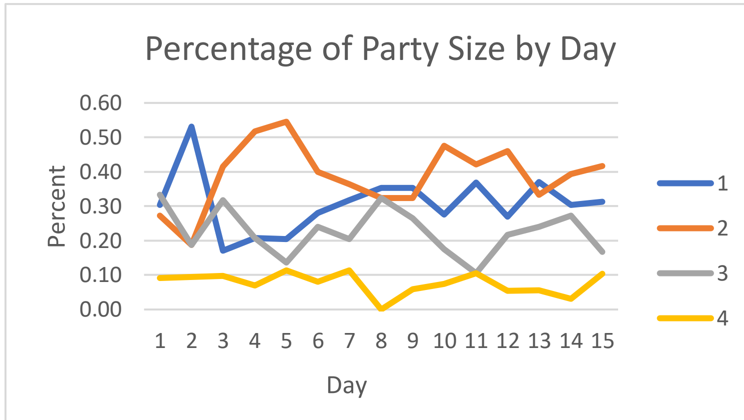


Figure 3

Order Types, Duration, and Pricing

To calculate the profitability of each model, revenues must be determined. However, the data provided did not contain any information on the frequency of any given item's purchase within an order type. As such, it was assumed that each item per order type was ordered equally often. Every item on the menu was categorized based on the five order types in the original data, and the price of those items were averaged for each order type group. The order types and their associated average price are shown in Table 3.

Table 3: Order Type Prices

Order Type	Average
Food	$ 6.03
Specialty	$ 4.46
Coffee/tea	$ 2.68
Food + Specialty	$ 10.49
Food + Coffee	$ 8.71
All	$ 4.75

In the original data, there were recordings of the time it took for each party member to be helped by the cashier. A cursory examination appears to show differences among the order types, so an examination was done to see whether the differences indeed existed. The café's data came with five order types, so the mean, minimum and maximum durations of each were calculated. Each order type was compiled and graphed into histograms to visualize the shape of the data, none of which showed a discernible pattern in its distribution (see Figure 4). Given the seemingly random distribution throughout the range of duration, it was concluded to not be weighted toward any point and thus uniform. In addition, the frequency of each order type was also calculated. These statistics can be seen in Table 4.

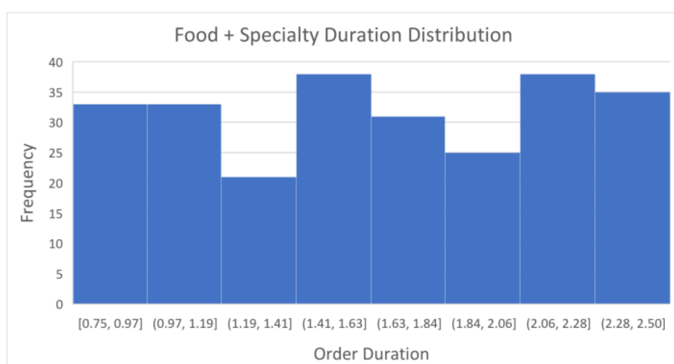


Figure 4

Barista Orders

The only order types to require the use of a barista are the ones that include menu items that fall under the category of "specialty." There are 95 of these items – 55 hot drinks and 40 cold. Once a customer is finished being helped by the cashier, they step over to the barista's queue, where they wait for their drink to be prepared by the barista. For the order types, a specialty drink is generally ordered alone or with food; Table 4 shows these two orders accounting for 54 percent of all orders (32 and 22 percent, respectively).

Table 4: Order Type Durations and Frequency

Order Type	Minimum Duration (min)	Maximum Duration (min)	Average Duration	Count	Frequency
Coffee:	0.5	0.99	0.759	184	16.17%
Food:	0.77	2.44	1.578	193	16.96%
Specialty:	0.5	1	0.747	364	31.99%
Food + Coffee:	0.75	2.49	1.614	143	12.57%
Food + Specialty:	0.75	2.5	1.646	254	22.32%

The barista duration data show that the barista took between roughly 0.75 and 1.5 minutes to make all drinks in the recorded period. This yielded a mean of approximately 1.12 minutes, with a standard deviation of 0.2 minutes (12 seconds). Like the cashier durations above, each day's barista duration time was found to follow a roughly uniform distribution. Table 5 shows the minimum, maximum, and mean of the barista duration data.

Table 5: Barista Duration	Minimum Duration (min)	Maximum Duration (min)	Mean
Mondays	0.75	1.5	1.180
Tuesdays	0.75	1.5	1.145
Wednesdays	0.75	1.49	1.130
Thursdays	0.76	1.48	1.122
Fridays	0.75	1.5	1.106
Total	0.75	1.5	1.120

Cashier Queue and Balking

The cashier queue is the number of customers in line to receive service, including the one who is ordering. The length of the queue depends on the party size and the order types of each customer. The line can only be so long before the café appears overly crowded to an arriving party. If a party deems the store overly crowded, they balk at the line and do not enter, presumably taking their business elsewhere. The company wants to eliminate as much of this as possible, as then they will likely be maximizing their revenues and profits. Thus, analyzing the queue length and balking behavior is essential to a functional model.

To best predict whether a party would balk, it was clear by a cursory look of the data that balking was largely dictated by the length of the cashier's queue; however, it was deemed worthy of examination whether balking correlated with an arriving party's size or length of the barista queue as well. While queue lengths are a discrete variable, balking is a dichotomous variable; either a party balks or it does not. In the "Correlations" tab of the workbook, balking was given a value of one, not balking zero. Excel's correlation function yielded a coefficient of approximately 0.2218. Thus, as party size increases, so does balking. However, the correlation is relatively weak. The increase in party size explains less than five percent of the balking data ($r^2 = 0.0492$). Similarly, the barista queue was tested for correlation. The correlation was similar but negative ($r = -0.244$), and explained approximately six percent of the data ($r^2 = 0.0587$). This negative correlation suggested that balking tended to decrease as the barista queue got longer. It was concluded that there were no strong correlations to balking with these additional factors, and the model would be assuming that balking was purely a relation to cashier queue alone.

All balking occurred with queues of at least five customers, so all party arrivals with a queue length of at least five were isolated. There was a total of 51 of these occurrences, 38 of which resulted in a balk. The longest queue was nine, of which there was only one occurrence. Table 6 shows that parties balked nearly 69 percent of the time they saw a queue length of five. Balking increased to almost 85 percent for a length of six. However, the data available show a decrease at a queue length of seven. This is likely a result of the small sample size, as only three arrivals coincided with a queue length of seven throughout the short recording period. Thus, a trendline was used to predict a more reasonable balking percentage based on the data available from queue lengths of five and six. Additionally, for this line of best fit, it was assumed that the 100 percent balk rate at queue length eight is accurate. It may be the case that, should enough days be recorded, a party might not balk at a queue length of eight; however, not enough data exists to predict otherwise. The trendline Excel produced was

Table 6: Prevalence of Balking

Queue Length	Balks	Total	Balk Percentage
5	22	32	68.8%
6	11	13	84.6%
7	2	3	66.7%
8	2	2	100.0%
9	1	1	100.0%

$$f(x) = -0.0272x^2 + 0.4583x - 0.9231.$$

With a queue length of seven,

$$f(7) = -0.0272(7)^2 + 0.4583(7) - 0.9231 \approx 0.950.$$

It is more reasonable to assume approximately 95 percent (rather than 67 percent) of parties will balk when arriving to a queue of seven.

Models

Replication Model

This model simulates a normal lunch rush at XYZ Café with one cashier and one barista. To create this model, the parameters outlined above must replicate the provided data. The methods by which these new datapoints were produced is detailed below.

Arrival Rate

The arrival rates in Table 2 showed that approximately 0.261 parties arrived per minute, on average, over the 15-day recording period. Moreover, the arrivals of the party were exponentially distributed. Thus, to generate our model's interarrival time, the exponential cumulative distribution function,

$$F(t) = 1 - e^{-\lambda t}$$

would be used, with t being a continuous time interval greater than 0, and λ being the average arrival rate. Using algebra to rearrange the equation,

$$e^{-\lambda t} = 1 - F(t)$$

$$-\lambda t = \ln\big(1 - F(t)\big)$$

$$t = \frac{-\ln\big(1 - F(t)\big)}{\lambda}$$

the result is an equation that can produce an interarrival time t, given a random value generated between 0 and 1 input into $F(t)$. Once a party's arrival time was generated, the model would need to decide the size of that party.

Party Size, Balking, and Order Type

At each arrival, a party size was generated using another random variable. The four party size possibilities were decided by turning their respective probabilities into a cumulative distribution function. This distribution is shown in Table 7.

Table 7: Party Size Frequency Distribution		
Party Size	Relative Frequency	CDF
1	30.801%	30.80%
2	38.985%	69.79%
3	22.604%	92.39%
4	7.611%	100.00%

If a random value between 0 to 1 was less than or equal to the frequency corresponding to party size one, then the party generated would contain one customer. If the random value was greater than the frequency for a party size of one but less than or equal to a party size of two, then the party would contain two customers, and so on. Thus, the random value between 0 and 1 would fall into one of these four categories, generating a party size at random for the model.

Similarly, the order type was decided based on a random number and the corresponding cumulative distribution function, outlined in Table 8.

Table 8: Order Type Frequency Distribution		
Order Type	Relative Frequency	CDF
Coffee:	16.169	16.17%
Food:	16.960	33.13%
Specialty:	31.986	65.11%
Food + Coffee:	12.566	77.68%
Food + Specialty:	22.320	100.00%

A similar method was used to decide whether a party would balk. When the cashier queue was longer than four people, a random number between 0 and 1 was generated. If that random number was less than or equal to the balking probability for the given queue length greater than four, the party balked. The balking probabilities are shown in Table 9.

Table 9: Balk Chances	
Queue	Balk Probability
5	68.75%
6	84.62%
7	95.03%
8+	100.00%

Order and Barista Durations

As previously mentioned, the durations for both cashier and barista wait times were found to be uniformly distributed between the minimum and maximum duration values. These duration ranges were found to fall into three categories: coffee/specialty drink orders, orders that included food in some capacity, and the Barista orders. For the model, a random number would be generated between the minimum and maximum durations found through the original data. The potential durations are shown in Table 10.

Table 10: Order Durations

	MIN (in min)	MAX (in min)
Coffee/Specialty Drink Orders:	0.5	1
Orders w/ Food:	0.75	2.5
Barista:	0.75	1.5

With these random variables determined, a model was built in the programming language C# to simulate real world randomness in a lunch rush; in so doing, Monte Carlo simulations could be used to provide meaningful statistics for making comparisons between our different modeled situations. The pseudocode below explains the chain of logic behind the program.

The model can be broken down into three major segments; the "Trials", the "Timestamp", and the "Core", and the pseudocode of each will be examined in turn. The "Trial" segment's primary purpose is to gather input from the user on how many trials they would like performed, loop the actual program that many times, and output the statistics gathered from all trials (pseudocode in Figure 5).

The "Trial" segment:

>Output "How many trials would you like performed?"

>Input number of trials

>Check if input is valid. If so…

> >Loop "Timestamp" Segment for number of requested trials
>
> >Perform calculations on information from "Timestamp" Segment to find averages and output them.

>If the trial input wasn't valid, end program with error.

Figure 5

Next, the "Timestamp" segment is itself a loop nested within the "Trial" loop (pseudocode in Figure 6). As the name suggests, this segment acts as a timestamp for the simulation done by the "Core" segments. Because the original data was over the course of 2.5 hours (150 minutes), and the timestamp of events was recorded to the precision of 0.01 minutes, the "Timestamp" segment performs 15,000 loops of the "Core" segment.

Figure 6

The "Timestamp" segment

>Loop the "Core" segment 15000 times

>Collect the data at the end of the trial to be used by the "Trial" segment

Lastly, the "Core" segment, as its name implies, is the core of the simulation (pseudocode in Figure 7). This segment tracks and performs all events that need to be checked every 0.01 minutes, i.e., at each iteration of the Timestamp loop. It is made up of three code blocks:

- When the interarrival timer reaches 0.
- When the cashier duration timer reaches 0.
- When the barista duration timer reaches 0.

Every timestamp loop it checks for these, or otherwise decrements their respective timer variables by 1.

The interarrival block is responsible for generating a new party size, determining whether the party balks with the current cashier queue length, and if not, puts the customers into the cashier queue. It then generates a new interarrival time to wait before generating a new party. The cashier block is responsible for taking customers from the cashier queue, generating what their orders are, the duration of the order, and passing completed specialty drink orders to the barista queue. The barista block is responsible for clearing customers from the barista queue and generating a duration for the service. By performing these tasks over the course of 15,000 loops, a day of serving customers at the café is simulated, ending with a summary of the final trial and total trials.

The “Core” Segment *Figure 7*

>If the interarrival time is 0...

>Reset balk flag if last party balked

>Determine size of next party with cumulative frequency distribution

>Determines if party balks based on cashier queue size

>If party doesn't balk...

>Increases cashier queue by party size and records total customers served

>Records maximum size of the Cashier Queue

>Increments the number of parties served

>Generate the interarrival time for the next party

>If cashier timer reaches 0...

>If cashier is flagged as in the process of helping a customer...

>Decrement the cashier queue

>If completed order includes Specialty, passes customer to Barista Queue

>Records maximum Barista Queue size

>Flags cashier as no longer helping a customer

>If cashier queue isn't empty...

>Flag cashier as helping customer

>Generates a random value and determines the customer’s order type with cumulative frequency distribution

>Generates a duration for Non-Food and Food orders, respectively

>Flag if order include Specialty Drink

>Otherwise, if cashier timer hasn't reached 0...

>Decrements cashier timer

>If the Barista timer has reached 0...

>If Barista is helping a customer...

>Decrement the queue by one and flag barista as no longer helping customer

>If the Barista queue has customers in it...

>Flag barista as in the process of helping a customer and generates a duration for barista timer

>Otherwise, if the Barista timer hasn't reached 0...

>Decrement the barista timer

>Decrements the Interarrival time for next party

Throughout the program, time stamped events will also output regularly during the final trial (omitted from pseudocode), as a form of debugging.

Replication Model Validation

Figure 8

```
Average Parties Served: 35.2466
Average Customers Served: 73.0966
Average Parties Balked: 3.0938
Deviation of Parties Balked: 2.5259411
Average Expected Revenue Lost from Balk: $40.47
Average Max Cashier Queue: 7.2733
Average Max Barista Queue: 2.4275
Average Revenue Earned: $455.58
Deviation of Revenue Earned: $106.19
Average Profit Earned (minus wages): $375.58

 Base Model
```

The first model (Figure 8) replicated the original data with one cashier and one barista. On average, about 35 parties were served per day. This comports well with the original data, as the average for all 15 days was 39.4 parties per day with a standard deviation of 6.3 days (Table 1). Moreover, the simulations average 3 balks per day. This fits within one standard deviation of the café's data, which averaged 2.53 balks per day (the average standard deviation per day was 0.52 balks). The average maximum of the cashier queue was 7.3 people, which is within one standard deviation of the average maximum of the original data (5.8 people with standard deviation of 1.7). Lastly, the barista queue averaged a maximum length of 2.4 people, which is also within one standard deviation of the café's data (1.9 mean with standard deviation of 0.6).

Thus, it was concluded that the replicated model closely resembled the original café data with respect to party arrivals, balks, cashier queue, and barista queue; therefore, the replication model is valid. The subsequent models utilizing a second cashier, or a floating barista will take the replication model for granted with different testing criteria.

Two-Cashier Model

This model is derived from the replication model. The structure is similar, with the only major difference being a second cashier block and additional code to prevent cashiers from helping the same customer (pseudocode in figure 9). In both this model and the barista/cashier model to follow, it is assumed that there is a singular cashier queue that both acting cashiers are pulling their next customer from. Since customers are also considered part of the queue until they are done being served, the additional code was necessary to prevent the second cashier from attempting to help the last customer when the customer was already being helped by the first cashier, which would alter the timers or the queue count.

Two Cashier Model – Cashier Block

```
>If cashier 1 timer reaches 0...
    >If cashier 1 is flagged as in the process of helping a customer...
        >Decrement cashier queue
        >If completed order is Specialty, passes customer to barista queue
            >Records maximum Barista Queue size
        >Flags cashier 1 as no longer helping a customer
    >If cashier queue isn't empty...
        >If there is only 1 customer and they are already being helped by cashier 2...
            >Do Nothing!
        >Otherwise...
            >Flag cashier 1 as helping customer
            >Generates a random value and determines the customers order type with
            cumulative frequency distribution
            >Generates a duration for Non-Food and Food orders, respectively
            >Flags if order is Specialty
>Otherwise, if cashier 1 timer hasn't reached 0...
    >Decrements cashier 1 timer
```

Figure 9

Floating Barista Model

This model is also derived from the replication model. This model has the inclusion of variables to mark the barista as busy in either line to enable logic that prevents them from working both as a cashier and barista at the same time. There are further checks to prevent the barista cashier from working on the same customer as the dedicated cashier. Likewise, the dedicated cashier has the checks from the two cashier model to ensure it does not overlap with customers being helped by the barista in the cashier queue.

Barista/Cashier Model – Barista Block

>If the Barista timer has reached 0...

>If Barista is helping a customer...

>If Barista is acting as barista...

>Decrement barista queue

>Otherwise, if Barista is acting as cashier...

>Decrement cashier queue

>If order is flagged as speciality, increment the barista queue

>Records maximum Barista Queue Size

>If the Barista queue has customers in it...

>Flag Barista as acting as barista and helping customer and set barista timer

>Otherwise, if cashier queue has customers in it...

>If there's only 1 customer left and they're being helped by cashier 1...

>Do Nothing!

>Otherwise...

>Set Barista as acting as Cashier and helping customer

>Generates a random value and determines the customers order type with cumulative frequency distribution

>Generates a duration for Non-Food and Food orders, respectively

>Flags if order is Specialty

>Otherwise, if the Barista timer hasn't reached 0...

>Decrement the barista timer

Figure 10

Results

Three models were created to test whether XYZ Café should keep their current system of employment, add a second cashier, or use the barista as floating second cashier. Each program prompts the user for the number of simulations the user wishes to run, and ten statistics are output:

- Average number of parties and customers served
- Average and standard deviation of balking parties
- Average maximum cashier and barista queue lengths
- Average revenue earned and standard deviation
- Average expected lost revenue from balking
- Average profit earned (revenue less wages)

The calculation of most of the averages was straight-forward; cumulate the total values at the end of each trial, then divide it by the number of trials performed at the end. In the case of expected revenue lost from balking, the weighted average revenue per customer was found using the average revenue for each of the five order types (Table 3) and their frequency (Table 4) to weight each revenue. This was then multiplied by the weighted average for party size, using one to four for the party sizes and their frequency distribution to weigh them. With a weighted average revenue per customer of

$$(2.68 * 0.1617) + (4.46 * 0.3199) + (6.03 * 0.1696) + (8.71 * 0.1257) + (10.49 * 0.2232) = \$6.32$$

and weighted average party size per party of

$$(1 * 0.308) + (2 * 0.3899) + (3 * 0.226) + (4 * 0.0761) = 2.0702,$$

it's expected that each balking party averages $13.08 in lost revenue.

For this report, 15,000 simulation trials were used. The three models yielded unique values for the ten statistics. Figures 11 and 12 display the results of the two-cashier and floating barista models, respectively, with the base model's results (Figure 8) for comparison. The two-cashier and the floating barista models that build upon the replication model are as follows.

Two-Cashier Model

The two-cashier model serves about 2.3 more parties per day than the replication model. Moreover, the number of balks reduces by more than two, averaging less than one balk per day. Additionally, the maximum cashier queue length decreases by about 0.8 people, on average. Thus, with the addition of a second cashier, the café can expect to increase the number of customers served by about five per day, reduce the number of balks significantly, and keep the cashier queue length slightly shorter, on average.

However, the average maximum barista queue increases by almost 1.2 people. This is a result of the increased rate of customers served since more than half of all customers order a specialty drink, requiring the use of the barista's service.

```
Average Parties Served: 35.2466
Average Customers Served: 73.0966
Average Parties Balked: 3.0938
Deviation of Parties Balked: 2.5259411
Average Expected Revenue Lost from Balk: $40.47
Average Max Cashier Queue: 7.2733
Average Max Barista Queue: 2.4275
Average Revenue Earned: $455.58
Deviation of Revenue Earned: $106.19
Average Profit Earned (minus wages): $375.58
```

Base Model

The average revenue lost decreases from the replication model by more than $28, from $39.21 to $10.88. Over a year, this would constitute nearly $10,000 in additional revenue that would have been lost to balking with only one cashier. However, the second cashier will require an additional salary of $16 per hour. With the total wages of the two cashiers and the barista, the average profit earned per day is $370.94. That is about $2 less than average profit of the replication model. Thus, the additional $10,000 in revenue will not be realized.

Figure 8

```
Average Parties Served: 37.5013
Average Customers Served: 77.6887
Average Parties Balked: 0.8552
Deviation of Parties Balked: 1.1100557
Average Expected Revenue Lost from Balk: $11.19
Average Max Cashier Queue: 6.5358
Average Max Barista Queue: 3.5936
Average Revenue Earned: $489.25
Deviation of Revenue Earned: $121.87
Average Profit Earned (minus wages): $369.25

Two Cashier Model
```

Figure 11

While more customers are served and balking is reduced, the profitability of XYZ Café decreases after adding a second cashier. Moreover, with the estimated cost of $1,000 for the POS system, the addition of a second cashier would be a significantly costly decision in the long term.

```
Average Parties Served: 36.9033
Average Customers Served: 76.5195
Average Parties Balked: 1.3717
Deviation of Parties Balked: 1.5446148
Average Expected Revenue Lost from Balk: $17.94
Average Max Cashier Queue: 6.7459
Average Max Barista Queue: 3.2698
Average Revenue Earned: $480.36
Deviation of Revenue Earned: $117.86
Average Profit Earned (minus wages): $400.36

Barista Cashier Model
```

Figure 12

Floating Barista Model

The floating barista model averages about 1.8 more parties served per day than replication, and reduces balking by about 1.6 parties, on average. The average maximum cashier length reduces replication by about 0.6 people per day, to an average of 6.8 people. Thus, the barista's ability to float between creating specialty drinks and acting as a second cashier will increase the number of customers served per day, significantly decrease the number of balks, and reduce the length of the cashier queue.

However, the maximum barista queue length increased by about 0.8 people per day, on average. This is from an increase in parties and customers served, but also that the floating barista can add to the barista queue length while they float as a second cashier.

The average revenue lost because of balking decreases from the replication model by more than $21, to only $17.87 per day. Moreover, the average revenue per day increases by more than $27. Since the floating barista model does not require an additional hourly wage, the daily profits are increased from an average $372.86 to $400.05. That is more than $27 per day, which would total $7069.40 over the 260 fiscal working days of 2022.

With the estimated $1,000 cost in the addition of a new POS system, the decision to have the barista float between a second cashier and creating drinks would be a significantly profitable move.

Discussion

There are two objectives for this report:

1) Develop a model to investigate and predict queue lengths, wait times and prevalence of balking.
2) Perform a cost-benefit analysis of the possibility of adding a second cashier to process orders faster.

a. Alternatively, have the barista act as a cashier between specialty drinks.

The models described here are suitable for these objectives, but they do have some limitations to take into consideration. The models are built from a small sample of days, so each simulation is essentially creating an average lunch rush between January 10th to January 28th, 2022. There is no consideration for presence of holidays, differences across seasons, or growth of the business over time.

Advantages of these models would be how it comprehensively simulates a day. Each aspect of the models is directly translated to data types of the sample data, and functions off the logic of the café workers workflow.

In a study with a wider scope using a larger sample of data over a few years, these models can adapt and overcome the prior mentioned limitations. Working with a larger dataset would certainly be a more intensive task, but would enable projecting the impact of holidays, trends over seasons, and the growth of the business. If this additional data were supplied, both model parameters and the model would need to be updated. The parameters would more likely reflect the true nature of the system, and need to include ways to model holidays, differences across seasons, and expected growth of the business over time. The model itself would need to accept these new parameters, as well as simulate a year instead and create the desired summary statistics across each year.

The model could also branch off to tackle separate issues. Perhaps adding large, highly readable menus to the store will prime customers for the counter, resulting in shorter order times. The models can take hypotheses on the parameters and see the effects on queue length, wait times, and prevalence of balking, i.e., sensitivity analysis.

Recommendations

The two-cashier model improves replication on several measures.

- Balks decreased by 72.3%
- Average maximum cashier queue decreased by 11.4%
- Number of parties and customers served increased by 7%

Despite these positive statistics, adding a second cashier is ultimately unprofitable. The average maximum barista queue length increases by 47.9 percent. More importantly, the addition of a third employee increases the daily wages by 50 percent, and this drastically undercuts the increase in revenue. The average profit from the replication model is $372.86 while the two-cashier model yields only $370.94. Moreover, the new POS system needed for another cashier would cost about $1,000. Thus, adding a second cashier full-time would hurt the profitability of XYZ Café. Therefore, the two-cashier model is a net-negative intervention and is not recommended.

The floating barista model similarly affects replication on several measures.

- Balks decreased by 54.4%
- Average maximum cashier queue decreased by 8.1%
- Number of parties and customers served increased by 5%
- Maximum barista queue length increased by 34.7%

The floating barista model yields an average profit close to the two-cashier model, but there is no additional employee. Thus, the average profit of the floating barista model is $400.05. That is a 7.3 percent increase, which would yield more than $7,000 in additional profits throughout the 2022 fiscal year. Again, the new POS system will cost approximately $1,000, but over time the move to a floating barista would be significantly more profitable.

It is recommended that XYZ Café invest in a second POS system and train their baristas to double as second cashiers when not servicing specialty orders. In time, this will significantly improve the profitability of the company.