Backyard Brew: Simulating My Local Coffee Shop

Aaron Tsui

MSDS 460

Professor Miller

December 8, 2024

Abstract and Introduction

This paper explores the challenges faced by my local coffee shop in minimizing lost customers due to balking and lowering overall employee turnover. The study employs simulation modeling to analyze customer flow and employee dynamics within a typical coffee shop environment. The model incorporates key variables, including customer arrival rates, service times, number of baristas, simulation hours, hours in length of the morning rush, percentage of customers who are rushed (in both the morning rush and afternoon rush), and thresholds to modify baristas, to simulate realistic scenarios.

By testing different operational strategies, such as optimizing staff scheduling and implementing queue management systems, the study identifies effective solutions to improve customer satisfaction, employee loyalty, and market competitiveness. The results recommend that splitting high traffic periods amongst multiple shifts is a solution to reducing stress and overwork amongst baristas. The results also show that dynamically adjusting baristas can significantly reduce customer balking, which contributes to a more efficient and sustainable operation. This research offers practical insights for coffee shop owners and managers seeking to balance quality customer service with employee well-being and loyalty.

Literature Review

The challenges of managing customer flow and staffing in coffee shops have been widely studied using simulation methods to optimize operations. One example was a simulation study of a Starbucks branch in Mexico City that explored ways to enhance service efficiency during peak hours. Using the Simio software, Buzali modeled customer flow and employee distribution across various scenarios. Their findings identified bottlenecks at the checkout counter and recommended increasing the number of cashiers and baristas during rush hours, resulting in a 40% reduction in wait times and improved customer throughput[1].

Another noteworthy study examined operations at a Starbucks near Bahçeşehir University. Through collecting data on beverage preparation times, customer preferences, and payment methods, the researchers created a simulation that accounted for different operational variables. Their analysis emphasized how factors like employee numbers and the introduction of efficient payment methods could streamline processes and reduce wait times[2].

Research Design and Methods

The research design of this term project leverages a discrete event simulation framework built in Python using the Simpy library.

1. Simulation Parameters

   The coffee shop environment is simulated over a 9 hour operational period with an initial allocation of three baristas. Dynamic adjustments to barista staffing are governed by queue thresholds, which dictate adding a barista when the queue exceeds eight customers and the number of current baristas is less than five. When the queue drops below three customers and the number of current baristas is greater than two, a barista will be removed.

   - Customer Arrival: Inter-arrival times follow an exponential distribution with a mean of two minutes.
   - Service Time: Service times are constrained between a minimum of five minutes and a maximum of 15 minutes, with a exponential distribution centered at 12 minutes.

- Behavioral Types: Customers are categorized as either “rushed” or “relaxed,” with behavioral probabilities adjusting based on time of day. During the morning rush, there are higher rates of “rushed” customers as compared to afternoon hours.

2. Dynamic Resource Allocation

   A dynamic baristas resource class tracks and ajustrs the capacity of the coffee shop staff in real time based on customer demand. Adjustments occur in ten-minute intervals, responding to queue sizes to balance staffing efficiency with customer satisfaction.

3. Customer Behavior
   - Balking: Customers leave immediately if the queue exceeds their tolerance threshold.
   - Reneging: Customers already in the queue may leave if their waiting time exceeds their patience threshold, determined by their behavior type; rushed customers have lower tolerance than regular or non-rushed customers.

4. Logging and Data Collection

   The simulation logs each customer’s journey through the event graph, including arrival time, queue entry, service initiation and completion, and instances of balking or reneging. Event logs are appended in real time to facilitate post-simulation analysis.

5. Performance Metrics
   - The percentage of customers served versus those who balked or reneged.
   - Average service and waiting times.
   - Frequency of barista modifications.
   - Recommendation of number of baristas on reserve.

6. Simulation Execution

   This model executes a single simulation run of nine operational hours. Random seeds ensure reproducibility, and results are analyzed to evaluate system performance under the defined parameters. Iterative adjustments to thresholds and resource allocation policies are anticipated to refine the model for better outcomes.

Results

The dynamic barista allocation system demonstrated effectiveness in addressing high traffic periods. The system performed six total adjustments across nine operational hours (the day), adding or removing baristas as queue lengths fluctuated. This adaptive strategy reduced customer wait times and improved service flow, though it required precise timing to avoid understaffing during unexpected demand surges.

Morning operations experienced one extended high traffic period and one shorter peak. Afternoon traffic was characterized by four shorter bursts, demonstrating the necessity of quick barista adjustments during these periods. Moderate queues at the end of the simulation day suggested that reduced staffing near closing time could inflate wait times. Although the simulation balanced barista capacity with expected traffic, closing time service levels with high queue length are areas to improve on in future studies.

Balking and reneging were most common during the morning rush hours, where queue thresholds for rushed customers were more frequently exceeded than they were in the afternoon hours. The dynamic adjustments helped mitigate losses in the afternoon, where shorter bursts of high traffic led to more manageable queues for the baristas in the simulation.

Management Recommendations

1. *Reserve Baristas on Half-Shifts*: To efficiently handle fluctuating demand, a reserve of two to three baristas should be available for backup (or on reserve) during high traffic periods. Both half-shifts are four hours forty-five minutes long and reflect how most often baristas were changed in the simulation. The half-shifts are as follows: Shift 1: 8:00 AM - 12:45 PM to cover the morning rush. Shift 2: 1:15 PM - 5:00 PM to cover afternoon demand surges. This schedule ensures adequate staffing to serve customers during both long and short high traffic periods while optimizing labor costs.
2. *Improved EOD Service*: To reduce wait times near closing, an additional barista could remain on duty until all queued customers are served. This adjustment ensures service quality is maintained throughout the day. An alternative solution to reducing wait times near closing is to advertise a "last-call" time where orders cannot be placed afterwards.
3. *Future Improvements*

- Incorporating predictive analytics based on historical data could help preempt high traffic rather than react to it; defining the morning rush period is an observational outlook on said analytics. Historical data could improve and streamline this.
  - Customer education strategies like posting real-time queue information in a public location or online (on platforms such as Yelp or a private website) would set expectations and reduce customers balking.
  - Offering mobile ordering through online methods during peak periods could further alleviate congestion and/or reduce balking or reneging.

Appendix

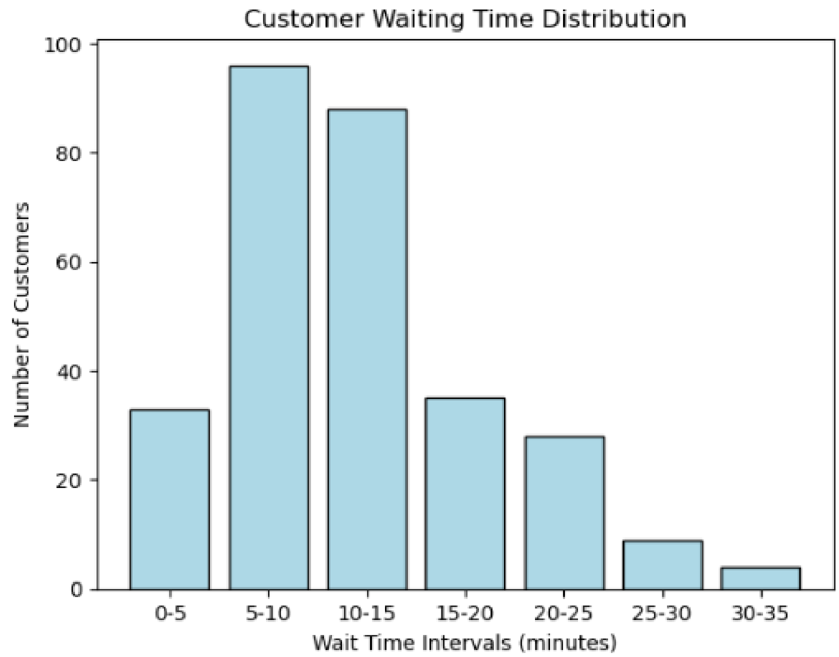
Summary Statistics

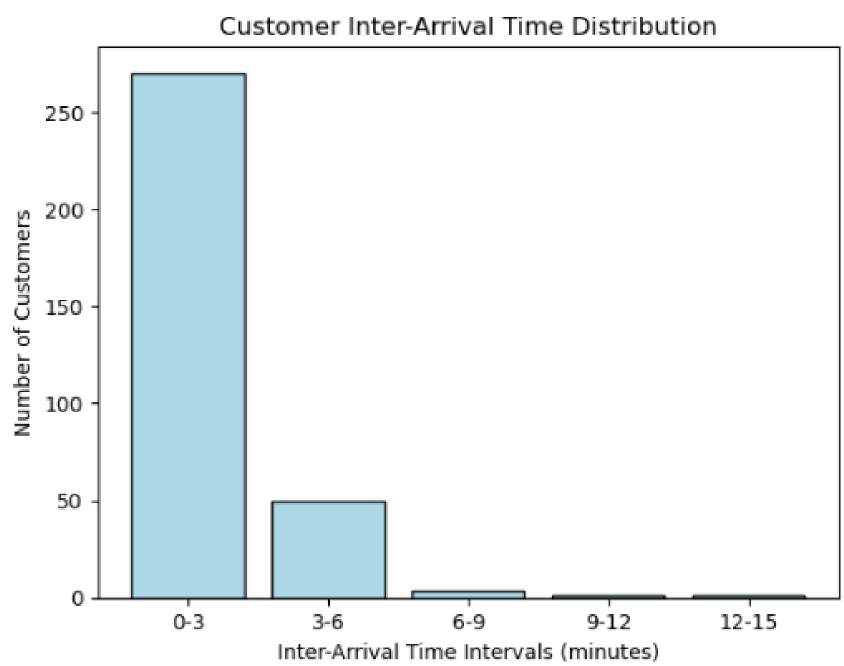
key	
Minimum inter-arrival_time	0
Average inter-arrival_time	99.33
Maximum inter-arrival_time	786
Mimimum waiting time	0
Average waiting time	714.89
Maximum waiting time	1832
Mimimum service time	300
Average service time	573.21
Maximum service time	900

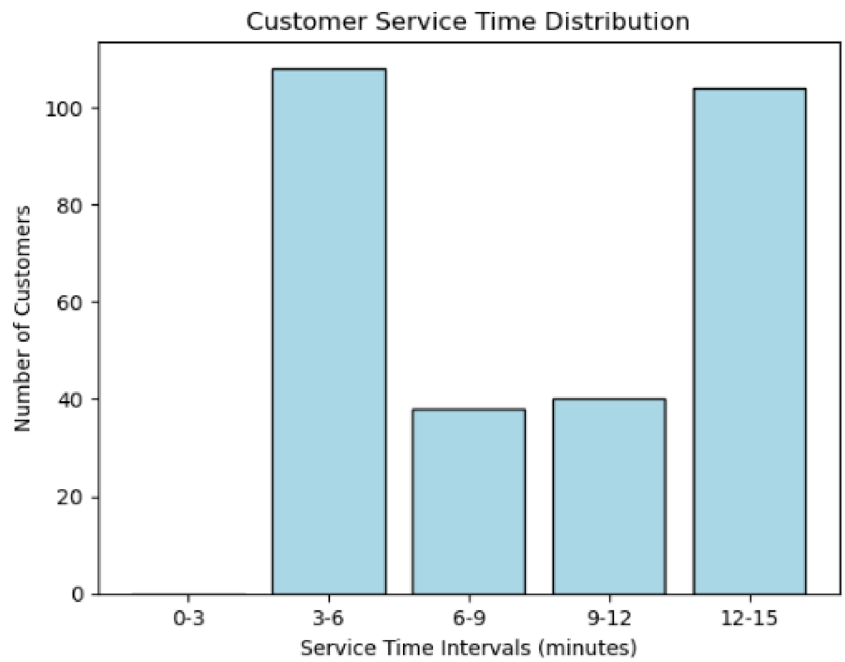
Simulation Statistics

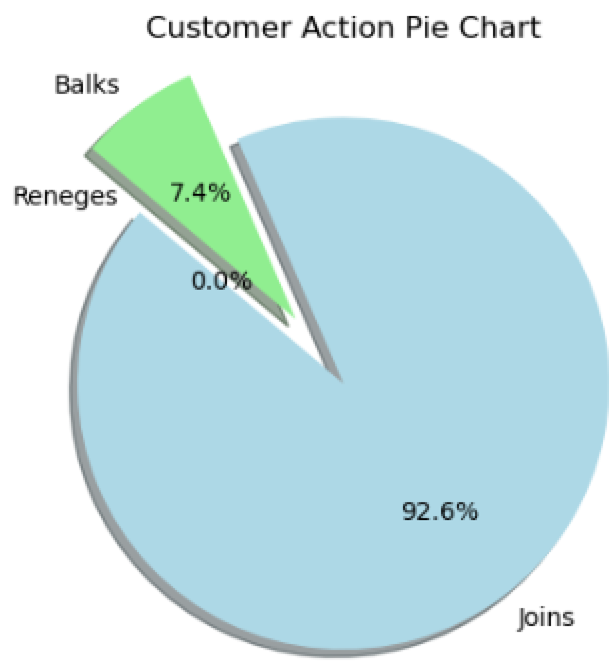
```
The simulation is set to run for 9 hours (540 minutes)

Results after 32382 seconds (539.7 minutes, 8.99 hours):
326 unique customers arrived
302 customers joined the queue for service
24 customers balked (lost business)
293 customers began service
290 customers ended service
9 customers were still in line at the end of the simulation
```

Customer Time Distribution Graphs






Customer Service Time Distribution
Number of Customers
0
20
40
60
80
100
0-3
3-6
6-9
9-12
12-15
Service Time Intervals (minutes)


Customer Action Pie Chart
Balks
7.4%
Reneges
0.0%
92.6%
Joins

References

1. Buzali, Dylan, Santiago Muñiz, Santiago Elizondo, and Maria Oroselfia Sanchez Sanchez. "Simulation-Optimization of a Coffee Shop in Business District: A Case Study of Starbucks in Mexico City." 5th South American Industrial Engineering and Operations Management Conference, Bogota, Colombia, May 7–9, 2024. https://doi.org/10.46254/SA05.20240042.
2. "An Operational Evaluation of Starbucks Coffee Shop Near Bahçeşehir University Using Simulation Techniques." IEOM Society International Conference Proceedings. Accessed December 8, 2024. https://ieomsociety.org/proceedings.