Production and Inventory Management for

The Ocean Breeze Soap Company

MIDN 1/C First Midshipman MIDN 1/C Second Midshipman MIDN 1/C Third Midshipman

April 2025

1 Introduction

The Ocean Breeze Soap Company is a small artisan soap shop located in downtown Simplexville. The company is struggling to manage its inventory of its three signature soap products – Atlantic Mist, Deep Sea Detox, and Sailor's Scrub – during its 13-week summer season. The shop experiences unpredictable customer traffic, with varying demand throughout the day and week. Furthermore, due to the varying availability of ingredients, the production costs of the three soap products change throughout the summer season. Small businesses like Ocean Breeze often struggle with inventory management due to limited resources, variable customer traffic, and seasonal demand fluctuations (Weissmueller, 2023). Effective inventory and production planning can mitigate these challenges, as highlighted in previous research (Atnafu & Balda, 2018).

Ocean Breeze wants to ensure that it can meet customer demand for its soap products in a timely manner while keeping its production and inventory costs at a minimum. We use three models to analyze and solve Ocean Breeze's production and inventory problem. First, we use linear regression to predict demand for the three soap products, using historical sales, weather, and special event data provided by the company. Second, we formulate and solve a linear optimization model to determine production and inventory quantities that meet the demand levels predicted by our linear regression model while minimizing cost. Finally, we develop a discrete-event simulation model to estimate the fraction of time that the products are out of stock when using the production and inventory quantities determined by our linear optimization model, incorporating historical customer arrival, cashier service time, and sales data provided by the company.

2 Data

The company provided us with five sources of data:

- weather and special event data,
- customer arrival data,
- sales data,
- cashier service time data, and
- production costs, and inventory holding costs and capacity.

2.1 Basic description of data

The weather and special event data consists of four variables, described in Table 1. The data contains a total

of 91 rows, one row for each day between 3 June 2024 and 1 September 2024. Of note, the SpecialEvent variable indicates whether the company held a sale or there was a nearby festival on that day.

Variable	Description
Date	Date of interest.
DayOfWeek	Day of the week.
Weather	Weather conditions for the given date. Possible values: sunny, rainy.
SpecialEvent	Special events for the given date. Possible values: no event, sale, or festival.

Table 1. Variables in the weather and special event data.

The customer arrival data consists of three variables, described in Table 2. The data contains a total of 819 rows, one row for each hour between 10:00 and 19:00 on each day between 3 June 2024 and 1 September 2024 (9 hours \times 91 days).

Variable	Description
Date	Date of interest.
Hour	Hour of interest.
Arrivals	Number of customer arrivals during the given hour on the given date.

Table 2. Variables in the customer arrival data.

The sales data consists of three variables, described in Table 3. The data contains a total of 273 rows, one row for each of the three soap products on each day between 3 June 2024 and 1 September 2024 (3 soap products \times 91 days).

Variable	Description
Date	Date of interest.
Soap	Soap product. Possible values: Atlantic Mist, Deep Sea Detox, Sailor's Scrub.
Sales	Number of sales for the given date and soap type.

Table 3. Variables in the sales data.

The cashier service time data consists of 168 observations of cashier service times, in seconds.

The production costs for each bar of soap over the 13-week time horizon are given in Table 4. Note that the production costs for each soap type vary over the 13-week time horizon, reflecting the availability of the necessary ingredients. The weekly inventory holding cost for each soap is \$1, and the shop can hold 500 bars of soap in inventory.

Week	1	2	3	4	5	6	7	8	9	10	11	12	13
Atlantic Mist	6	6	6	6	11	11	11	11	11	4	4	4	4
Deep Sea Detox	8	8	8	8	3	3	3	3	3	7	7	7	7
Sailor's Scrub	3	3	3	3	4	4	4	4	9	9	9	9	9

Table 4. Production costs for each soap type.

2.2 Exploratory data analysis

Tables 5 and 6 show the distribution of weather conditions and special events in the weather and special event data, respectively.

Weather	Fraction of days
Sunny	0.703
Rainy	0.297

Table 5. Distribution of weather conditions in the weather and special event data.

Special Event	Percentage
No event	0.604
Festival	0.220
Sale	0.176

Table 6. Distribution of special events in the weather and special event data.

Merging the weather and special event data with the customer arrival data by date of interest, we are able to identify the relationships between the number of customer arrivals with the day of the week, the weather, and special events. Figure 1 shows the average number of customer arrivals per hour for each day of the week. As Figure 1 shows, the shop is busiest on the weekends, with noticeable peaks in visitors in the early afternoon, from 12:00 to 16:00. There is also a slight uptick in customer arrivals in the evenings on Thursdays and Fridays. The average arrival rates shown in Figure 1 range from 12.46 to 39.46 customers per hour.

Figure 2 shows the average number of customer arrivals per hour under sunny and rainy weather conditions. As we see in Figure 2, the shop is about twice as busier on sunny days compared to rainy days. Figure 3 shows the average number of customer arrivals per hour under the three different special event conditions: no event, nearby festival, and sale. From this figure, we see a higher number of customer arrivals in the early afternoons on festival days, similar to the customer arrival patterns on the weekends shown in Figure 1. We also see a slight uptick in customer arrivals when the shop is holding a sale.

Day of week



Figure 1. Average number of customer arrivals per hour for each day of the week.



Figure 2. Average number of customer arrivals per hour under sunny and rainy weather conditions.



Figure 3. Average number of customer arrivals per hour under different special event conditions.

Figure 4 shows the average daily sales by soap type. The figure shows that of the three soap products, customers purchase Atlantic Mist about 50.2 percent of the time, Deep Sea Detox about 30.5 percent of the time, and Sailor's Scrub about 19.3 percent of the time.



Figure 4. Average daily sales by soap type.

Similar to the customer arrival data, merging the sales data with the weather and special events data, we can identify the relationships between sales and the day of the week, the weather, and special events. Figure 5 shows these relationships for total daily sales. As we can see, sales tend to increase from Monday to Sunday. We also see that sales are generally higher on sunny days compared to rainy days. We also see that compared to days with no special event, days with nearby festivals are associated with higher sales, while days with sales tend to be similar in sales. All of these trends mirror the customer arrival data, as seen in Figures 1, 2 and 3.



Figure 5. Relationships between total daily sales, day of week, weather, and special event.

Merging the sales data with the arrivals data, we can determine the conversion rate, or the fraction of customer arrivals that result in a sale each day. Figure 6 shows a histogram of these conversion rates. When computing these conversion rates, we assume that each customer purchases one bar of soap. We see that on most days, the conversion rate is below 40 percent. The mean conversion rate is 23.04 percent.



Figure 6. Histogram of conversion rate, or fraction of customer arrivals resulting in a sale.

To determine a suitable distribution to represent the cashier service times, we used maximum likelihood estimation to fit the service time data to the exponential, lognormal, beta, gamma, triangular, normal, and uniform distributions. We found that the exponential distribution with a mean of 282 fit the data best, with a chi-squared goodness-of-fit test statistic of 5.05 and a Kolmogorov-Smirnov test p-value of 0.917. Figure 7 shows a histogram of the cashier service time data, along with the density of the fitted exponential distribution.



Figure 7. Histogram of cashier service time data and density of fitted exponential distribution.

3 Model

In order to determine a production and inventory strategy for Ocean Breeze, we use a three-stage modeling approach:

- We used linear regression to forecast daily demand for its three soap products, based on the day of the week, weather, and special events.
- Based on the daily demand predicted by our linear regression, we used linear optimization to determine weekly production and inventory quantities that minimize cost.
- We used discrete-event simulation to estimate the fraction of time the soap products are out of stock when using the production and inventory quantities determined by our linear optimization model.

3.1 Forecasting daily demand with linear regression

To forecast demand for Ocean Breeze's three soap products, we used a linear regression with historical sales as the response variable, and the soap product, day of week, weather, and special events as the explanatory variables. After some preliminary modeling and diagnostics, we determined that using the logarithm of sales as the response variable led to models that better satisfy the linearity condition for linear

regression. We then used best subsets regression with the AIC criterion to determine which explanatory variables to include in our final model. Our fitted linear regression model is:

$$log(Sales) = \hat{\beta}_{0} + \beta_{DeepSeaDetox} DeepSeaDetox + \beta_{SailorsScrub} SailorsScrub$$
(1)
+ $\beta_{Tuesday} Tuesday + \beta_{Wednesday} Wednesday + \beta_{Thursday} Thursday$
+ $\beta_{Friday} Friday + \beta_{Saturday} Saturday + \beta_{Sunday} Sunday$
+ $\beta_{Rainv} Rainy + \varepsilon$, $\varepsilon \sim N(0, \sigma_{\varepsilon}^{2})$

where *DeepSeaDetox* and *SailorsScrub* are indicator variables for the corresponding soap products, *Tuesday*, *Wednesday*, *Thursday*, *Friday*, *Saturday*, and *Sunday* are indicator variables for the corresponding days of the week, and *Rainy* is an indicator variable for rainy weather.

3.2 Determining weekly production and inventory quantities with linear optimization

We determine weekly production and inventory quantities with the following linear optimization model:

Sets.

$$T = set of weeks$$

P = set of soap products

Parameters.

$c_{p,t}$	= production cost of soap product <i>p</i> in week <i>t</i>	for $p \in P, t \in T$
h _{p,t}	= holding cost of soap product p from week t to week $t + 1$	for $p \in P, t \in T$
$d_{p,t}$	= demand for soap product p in week t	for $p \in P, t \in T$
и	= weekly inventory capacity	
v	= safety factor	

Decision variables.

$$\begin{array}{ll} x_{p,t} & = \text{ amount of soap product } p \text{ to order in week } t & \text{ for } p \in P, t \in T \\ y_{p,t} & = \text{ amount of soap product } p \text{ to hold in inventory from week } t \text{ to week } t+1 & \text{ for } p \in P, t \in T \cup \{0\} \end{array}$$

Objective function and constraints.

minimize
$$\sum_{p \in P} \sum_{t \in T} (c_{p,t} x_{p,t} + h_{p,t} y_{p,t})$$
(2a)

subject to $y_{p,t-1} + x_{p,t} = vd_{p,t} + y_{p,t}$ for $p \in P, t \in T$ (2b)

 $y_{p,0} = 0 \qquad \qquad \text{for } p \in P \tag{2c}$

$$\sum_{p \in P} y_{p,t} \le u \qquad \qquad \text{for } t \in T \tag{2d}$$

$$\begin{aligned} x_{p,t} &\geq 0 & \text{for } p \in P, t \in T \\ y_{p,t} &\geq 0 & \text{for } p \in P, t \in T \cup \{0\} \end{aligned}$$

The objective function (2a) minimizes the total production and inventory holding costs. Constraint (2b) ensures that for each soap product, demand is met and inventory is carried over from one week to the next. The safety factor allows us to meet some multiple of the demand if desired; for example, we could set v = 1.1 to ensure we meet 110% of the demand from week to week. Constraint (2c) models the fact that we have zero inventory of all soap products at the beginning of the time horizon. Constraint (2d) ensures that we do not exceed the weekly inventory capacity. Constraints (2e) and (2f) ensure that the amounts produced and held in inventory are nonnegative.

To make this model concrete, we let

$$T = \{1, 2, ..., 13\}$$

P = {AtlanticMist, DeepSeaDetox, SailorsScrub}

The production costs $c_{p,t}$, inventory holding costs $h_{p,t}$, and the weekly inventory capacity u are given at the end of Section 2.1; in particular, see Table 4 for the production costs. For the weekly demands $d_{p,t}$, we use our fitted linear regression model (1) as follows. Let

K = set of days of the week = {Monday,Tuesday,Wednesday, ...,Saturday,Sunday}W = set of weather conditions = {Sunny, Rainy}

First, we predict the daily demand $g_{p,k,w}$ for soap product $p \in P$, day of the week $k \in K$ and weather condition $w \in W$ by substituting the appropriate values of the indicator variables in (1). For example,

$$g_{\text{DeepSeaDetox,Thursday,Rainy}} = \exp(\hat{eta}_0 + \hat{eta}_{\text{DeepSeaDetox}} + \hat{eta}_{\text{Thursday}} + \hat{eta}_{\text{Rainy}})$$

where $\hat{\beta}_i$ is the fitted estimate for coefficient β_i . We assume the weekly demand for each soap type and the fraction of sunny and rainy days is the same from week to week. Then, for each soap product $p \in P$, we can add the predicted daily demands $g_{p,k,w}$ over all the days of the week $k \in K$ and weight them by the fraction of sunny days α_{Sunny} and the fraction of rainy days $(1 - \alpha_{\text{Sunny}})$ as follows:

$$d_{p,t} = \alpha_{\text{Sunny}} \sum_{k \in K} g_{p,k,\text{Sunny}} + (1 - \alpha_{\text{Sunny}}) \sum_{k \in K} g_{p,k,\text{Rainy}} \quad \text{for } p \in P, t \in T.$$

From Table 5 in Section 2.1, we have that $\alpha_{\text{Sunny}} = 0.703$.

3.3 Estimating out-of-stock rates with discrete-event simulation

We use discrete-event simulation of the Ocean Breeze shop to model customer and employee behavior and validate the production and inventory quantities determined by the optimization model we described above in Section 3.2. Figure 8 shows a flow chart of our simulation model.



Figure 8. Flow chart of simulation model.

Our simulation is set to run for 13 weeks, starting on a Monday. At the beginning of each day, our simulation determines the weather for the day, according to the distribution in Table 5. Furthermore, at the beginning of week t, our simulation replenishes the inventory of soap product p according to the optimal production quantities (i.e., optimal values of $x_{p,s}$) from the optimization model described in Section 3.2.

Customers enter the shop at the *Customer Arrival* node according to a nonstationary Poisson process whose arrival rate changes hourly. For these hourly arrival rates, we use the average number of customer arrivals for the particular day of the week, hour, and weather condition from the customer arrivals data described in Section 2.

From the *Customer Arrival* node, customers either move to the *Cashier* node with probability γ , representing a customer purchasing a bar of soap, or the *Customer Exit* node with probability $1 - \gamma$,

representing a customer visiting the store without making a purchase. For the probability γ , we use a value close to the mean conversion rate of 0.23, discussed in Section 2.2.

For the customers that visit the *Cashier* node, the simulation randomly determines the soap product that customer purchases according to the distribution shown in Figure 4. Customers are served by a single cashier at this node, with a service time that follows an exponential distribution with a mean of 4.7 minutes (282 seconds). Recall this is the distribution that we found best fits the service time data provided, described in Section 3. Once these customers are served, they proceed to the *Customer Exit* node.

4 **Results**

For the regression modeling and analysis, we used the R statistical programming language. We implemented the optimization model in Python with the Pyomo optimization modeling library, and solved the model with the GLPK solver. Finally, we built and analyzed the discrete-event simulation with the Simio simulation software. All of this was done on a computer with an AMD Ryzen PRO 4750U CPU (8 cores, 1.7 GHz) and 16 GB RAM, running Windows 11.

4.1 Linear regression model

Figure 8 shows the diagnostic plots for our final linear regression model (1). As we see in Figure 8, linearity and equal variance both appear to be satisfied, since the residuals versus fitted plot shows little curvature, with the points roughly evenly distributed above and below the zero residual line and spanning a relatively constant vertical distance, moving from left to right. In addition, normality seems to be satisfied, since the normal Q-Q plot of the residuals is mostly linear. There do not appear to be any problematic observations that may distort the outcome and accuracy of the regression, since the Cook's distance of all observations is below 0.5. Finally, multicollinearity does not appear to be an issue, since the variance inflation factors of all the predictors are all less than 5.



Figure 8. Diagnostic plots for final linear regression model (1).

	Estimate	Standard error	t-statistic	p-value
Intercept	3.843	0.042	91.039	0.000
DeepSeaDetox	-0.501	0.033	-15.212	0.000
SailorsScrub	-0.876	0.033	-26.621	0.000
Tuesday	0.048	0.050	0.948	0.344
Wednesday	0.389	0.052	7.514	0.000
Thursday	0.453	0.051	8.908	0.000
Friday	0.338	0.051	6.643	0.000
Saturday	0.854	0.050	16.963	0.000
Sunday	0.777	0.051	15.285	0.000
Rainy	-0.683	0.033	-21.003	0.000
Number of c	observations	273		
	R^2	0.873		
	F-statistic	200.4	p-value	0.000

Table 7. Results for final linear regression model (1).

Table 7 shows the results for our final linear regression model (1). The t-tests for the coefficients indicate that each of the predictors is significantly associated with the logarithm of sales (p-values < 0.001), except for the day of the week being Tuesday. Furthermore, the F-test provides strong evidence that the overall model is effective (p-value \approx 0). Approximately 87.3% of the variability in the logarithm of sales can be explained by the predictors in our model.

Interpreting the estimated coefficients in Table 7, we see that our model indicates that on average, daily sales of Deep Sea Detox are approximately $e^{-0.501} \approx 60.6\%$ of the daily sales of Atlantic Mist, all else being equal. Similarly, our model predicts that daily sales of Sailor's Scrub are approximately 41.6% of daily sales of Atlantic Mist. In addition, our model indicates that daily sales on Saturdays are approximately 2.35 times higher than daily sales on Mondays, and daily sales on Saturdays are approximately 2.17 times higher than daily sales on Mondays, all else being equal. Finally, our model predicts that rainy weather leads to a 50% decrease in daily sales on average, holding all other variables constant.

4.2 **Optimization model**

We solved our linear optimization model described in Section 3.2 for three different safety factors, v = 1.00, 1.10, 1.20. Table 8 shows the optimal production quantities we obtained from our model, and Table 9 shows the optimal total production and inventory holding costs.

		Production quantities												
Safety factor	Week	1	2	3	4	5	6	7	8	9	10	11	12	13
1.00	Atlantic Mist	149	149	149	649	0	0	0	95	149	149	149	149	149
	Deep Sea Detox	92	92	92	92	92	92	92	92	368	0	0	0	92
	Sailor's Scrub	57	57	57	57	57	57	57	57	281	0	0	0	4
1.10	Atlantic Mist	164	164	164	664	0	0	0	156	164	164	164	164	164
	Deep Sea Detox	101	101	101	101	101	101	101	101	404	0	0	0	101
	Sailor's Scrub	63	63	63	63	63	63	63	63	260	0	0	0	55
1.20	Atlantic Mist	179	179	179	679	0	0	37	179	179	179	179	179	179
	Deep Sea Detox	110	110	110	110	110	110	110	110	406	0	0	34	110
	Sailor's Scrub	68	68	68	68	68	68	68	68	272	0	0	0	68

Table 8. Optimal production quantities from linear optimization model.

	Optimal total
Safety factor	production and
	inventory cost
1.00	\$22,419
1.10	\$24,774
1.20	\$27,104

Table 9. Optimal total production and inventory costs from linear optimization model.

Comparing the optimal production quantities with the production costs in Table 4, we see that the optimal solutions build up inventory before weeks with higher production costs. We also see that the optimal solution under the highest safety factor, v = 1.20, involves producing Atlantic Mist and Deep Sea Detox

slightly more frequently (weeks 7 and 12, respectively), likely due to the increase in demand as a result of the higher safety factor, but without a corresponding increase in the inventory capacity.

4.3 Simulation model

To test the robustness of the optimal production quantities from our linear optimization model, we ran our simulation under nine different scenarios based on different values of the conversion rate γ , the percentage of customers who make a purchase, and the safety factor v used to determine the optimal production quantities. Table 9 enumerates these scenarios.

Scenario	Conversion rate γ	Safety factor v
1	0.25	1.0
2	0.25	1.1
3	0.25	1.2
4	0.35	1.0
5	0.35	1.1
6	0.35	1.2
7	0.45	1.0
8	0.45	1.1
9	0.45	1.2

Table 9. Simulation model scenarios.

For each scenario, we ran 100 replications. In order to determine the robustness of the optimal production quantities, we focused on the out-of-stock rate for each soap product: the percentage of customers who wanted to purchase a particular product while it was out of stock. Table 10 shows the results of our simulation for these performance metrics, and Figure 9 shows these results visually as boxplots for each scenario (conversion rate and safety factor combination). As expected, as the safety factor increases, the out-of-stock rates of all three soap types decreases. In addition, also as expected, as the conversion rate increases.

			Out-of-sto	ck rate (95% confiden	ce interval)
Scenario	Conversion rate γ	Safety factor v	Atlantic Mist	Deep Sea Detox	Sailor's Scrub
1	0.25	1.0	0.0037	0.0057	0.0213
1	0.23	1.0	(0.0024, 0.0049)	(0.0040, 0.0074)	(0.0166, 0.0260)
2	0.25	11	0.0006	0.0008	0.0035
2	0.23	1.1	(0.0001, 0.0011)	(0.0002, 0.0014)	(0.0022, 0.0048)
3	0.25	1.2	0.0001	0.0001	0.0008
5	0.23	1.2	(0.0000, 0.0003)	(-0.0001, 0.0004)	(0.0003, 0.0013)
4	0.25	1.0	0.1175	0.1337	0.1699
4	0.55	1.0	(0.1137, 0.1213)	(0.1291, 0.1383)	(0.1646, 0.1752)
5	0.25	1 1	0.0325	0.0518	0.0832
5	0.55	1.1	(0.0290, 0.0360)	(0.0473, 0.0564)	(0.0776, 0.0888)
6	0.35	1.2	0.0017	0.0048	0.0251
0	0.55	1.2	(0.0010, 0.0025)	(0.0033, 0.0063)	(0.0210, 0.0292)
7	0.45	1.0	0.2419	0.2617	0.2926
/	0.45	1.0	(0.2389, 0.2450)	(0.2579, 0.2654)	(0.2880, 0.2971)
0	0.45	1 1	0.1656	0.1894	0.2181
0	0.45	1.1	(0.1623, 0.1690)	(0.1853, 0.1936)	(0.2131, 0.2231)
0	0.45	1.2	0.0931	0.1175	0.1562
9	0.45	1.2	(0.0895, 0.0967)	(0.1130, 0.1219)	(0.1508, 0.1616)

Table 10. Simulated out-of-stock rates for each scenario.



Figure 9. Boxplots of simulated out-of-stock rates for each scenario (conversion rate and safety factor).

5 Conclusion

To summarize, we used a three-stage modeling approach to analyze and solve Ocean Breeze's production and inventory problem. First, we used linear regression to forecast demand for the three soap products, using historical sales, weather, and special event data. Next, we formulated and solved a linear optimization model to determine the production and inventory quantities that minimize the cost of meeting the demand levels predicted by our linear regression model. Finally, we used a discrete-event simulation model to estimate out-of-stock rates of the three products when using the production and inventory quantities determined by our linear optimization model, incorporating historical customer arrival, cashier service time, and sales data.

Based on our results, we recommend that Ocean Breeze implement the production and inventory quantities determined by our linear optimization model with a safety factor of 1.1, shown in Table 8. As Figure 9 shows, using these quantities, the company can expect out-of-stock rates of less than 10% when the conversion rate is 0.35 or lower; historically, the conversion rate has been under 0.40 approximately 70% of the time, as shown in Figure 6. However, depending on the company's tolerance for being out of stock, they may fare better with production and inventory quantities determined with a lower or higher safety factor.

There are some limitations of our study that may affect our results and recommendation. First and foremost, we assume that historical sales patterns are indicative of future sales. Because we were only provided with aggregate daily sales data, we assumed that each customer buys at most one bar of soap when computing conversion rates. When computing weekly demand forecasts, we also assumed that product preference and weather patterns stayed constant throughout the 13-week summer season. One possible extension of this study would be to extensively test how sensitive the production and inventory quantities determined by our linear optimization model are against changes in product preference and weather patterns. It would also be interesting to extend the models and analyses in this study to accommodate percustomer sales data, instead of the aggregate daily sales data provided.

Disclosure of generative AI use

We used the generative AI tool, ChatGPT 40 mini, to

- develop the code used to produce the figures in Section 2, and
- provide feedback on the writing in Sections 1 and 4.

We have reviewed and edited the generative AI content and take full responsibility for the content of this capstone project report.

References

- Atnafu, D., & Balda, A. (2018). The impact of inventory management practice on firms' competitiveness and organizational performance: Empirical evidence from micro and small enterprises in Ethiopia. Cogent Business & Management, 5(1). https://doi.org/10.1080/23311975.2018.1503219
- Weissmueller, R. (2023, May 4). *The small business inventory dilemma and how to get out of it.* Forbes. https://www.forbes.com/councils/forbesbusinesscouncil/2023/05/04/the-small-business-inventory-dilemma-and-how-to-get-out-of-it/