BAT 3303 | XYZ Project

FINAL PROJECT REPORT

CHLOE SCROGGINS
GARRETT MORAN
AN VU NGUYEN DIEU

Table of Contents

I. Project Overview

XYZ is a major technological company who has been leading the market for over 20 years. To maintain its position as a leader in the industry, XYZ seeks to achieve a low-cost, low-risk model by enhancing its supply chain strategically.

The COVID-19 pandemic sent the global economy into turmoil with unprecedented fluctuations in supply and demand. Companies with complex supply chains, like XYZ, are especially vulnerable to disruption due to shortages of inputs from other businesses. Inoperative ODMs and unpredictable lead times from 3PL carriers prove difficult to manage during a time when demand for desktops and laptops is constantly on the rise. Facing unprecedented risks, XYZ seeks to protect its competitive advantages by maximizing accuracy in orders delivery to meet with customers' expectations. Failure to attain accurate lead times for products can lead to higher inventory carrying costs, late fees and loss of trust from customers.

To assist XYZ in improving its service, we analyzed the relationship between different ODM sites and different Ship Modes to understand its influence on Intransit Lead Time – the amount of time to ship an order from an ODM to the Company's Warehouse. Such analysis allows us to identify the factors with the most powerful impact on Intransit Lead Time and thus, determine the direction for enhancing delivery accuracy. Our method involves developing three multivariable linear regression model in which the logistic lead time is the outcome variable whilst elements such as LOB, Origin, or Ship Mode are independent (input) variables. During this process, we encountered several problems including missing data and multicollinearity. However, we have been able to overcome all obstacles by analyzing the impact of multicollinearity on the Model Parameters and Key Statistics to improve the models step-by-step. The final result of this project is a reliable multivariable linear regression model with an adjusted R-squared of 0.8614 and low variance inflation factors.

II. Research Questions

This project seeks to address the following questions:

1. Has the logistics lead time reduced since Q2, when compared with Q4 and Q1 for AIR and GROUND ship mode?
2. How many in-transit days can be saved if we bring inventory by AIR from Site A compared to Site C?
3. Optimization Question:
    a. Can we fulfil all 100K units before the deadline with the available budget based on the current manufacturing and logistics lead times?
    b. If not, please come up with an optimized volume of product A needs to be lifted by air/Ocean/Fast Boat to meet maximum customer demand depending on the ship mode lead times and budget available?
    c. Create a proposal to request for additional budget to fulfil all 100K units by the deadline?
    d. Provide an analysis to leadership team on total logistics cost vs total revenue vs profit margin with the initial budget provided and proposed budget for them to decide what trade-offs to make?

III. Data Overview

1. Raw Data Overview and Description

The data file from XYZ consists of three excel sheets which are Raw-Data, Calendar, and Cost Details. In this section, we only discuss the two main data-frames – Raw-Data and Calendar. The third data-frame Cost Details is discussed later in the report during Optimization.

The Raw-Data sheet contains 9151 observations of six variables describing order details: LOB, Origin, Ship Mode, PO Download Date, Ship Date, and Receipt Date. These variables are described in Table 1.

Variable	Description
LOB	Line of Business or the product being manufactured and shipped. 3 types of products: product A, product B, product C.
Origin	Original Design Manufacturers or ODMs: Where the product is manufactured and shipped from. 4 sites: Site A, Site B, Site C, Site D
Ship Mode	The means used to ship an order. 4 modes: Air, FastBoat, Ground, Ocean.
PO Download Date	The date on which an order is created and downloaded at a site.
Ship Date	The date on which an order is shipped from an ODM to the company's Warehouse.
Receipt Date	The date on which the Warehouse received an order.

Table 1: Raw-Data columns descriptions

The Calendar table contains six rows which identify six consecutive time periods of orders. These match up with a Quarter and Year. Table 2 below provides the description of all relevant variables in Calendar:

Variable	Description
Start_Date	The start date of each period.
End_Date	The end date of each period.
Quarter	The corresponding quarter, based on End_Date.
Year	The corresponding year, based on Quarter.

Table 2: Calendar columns descriptions

2. Data Computation

We added a total of 4 columns to the Raw-Data dataframe, which are: Quarter, Year, Intransit Lead Time, and Manufacturing Time. Table 3 describes the new variables we added to Raw-Data:

Variable	Description
Quarter	The Quarter when an order was delivered to the company's warehouse.
Year	The Year when an order was delivered to the company's warehouse.
Manufacturing_Time	The amount of time to manufacture a product for an order. Manufacturing_Time = Ship Date - PO Download Date
Intransit_Lead_Time	The amount of time to ship an order from an ODM to a company warehouse. Intransit_Lead_Time = Receipt Date - Ship Date

Table 3: New variable descriptions

3. Cleaning Data

a. Check rows and columns of Raw-Data for NAs - Determine the number of rows to include for analysis.

To check for any missing data, we found the percentage of NA values in each row of the total 9151 rows in Raw-Data. None of the considered rows had 50% or greater of their original data populated with NAs. Therefore, we removed no rows from the Raw-Data dataframe and the total number of rows retained for further analysis is 1951.

When we calculated the percentage of NA values by columns, however, there were 2 emerging situations concerning NA values: the columns were either free of NAs or had 1.6% NAs. Ship Date and Receipt Date were the only two columns from the original dataset with NAs and this resulted in NAs for Quarter, Year, Manufacturing_Time, and Intransit_Lead_Time. This is due to the fact that the 4 newly added columns are imputed based on Ship Date and Receipt Date columns. Consequently, when both of these values are NAs, there will be no valid input for the new columns. We need to populate the NA cells of Ship Date and Receipt Date to ensure the imputed data is meaningful.

b. Replace the NA values and unusual values with meaningful values.

For rows with NAs, we estimated the Ship Date by adding the estimated Manufacturing Time to the PO Download Date and then added the estimated Intransit Lead Time to the Ship Date to determine the estimated Receipt Date. These estimated values were the mean of the Manufacturing Time and the Intransit Lead Time of all orders grouped by LOB, Origin, and Ship Mode.

Besides NA values, unusual data points were also replaced with the good means. These outliers included points with Intransit Lead Time less than or equal to 0, Manufacturing Time less than 0, Manufacturing Time larger than or equal to 100, and Intransit Lead Time less than or equal to 10 with OCEAN Ship Mode.

After replacing all the NAs and unusual values with meaningful values, we checked for NAs again and at this point no columns had NAs.

c. Remove irrelevant data.

Since PO Download Date, Ship Date, and Receipt Date have been used to compute Quarter, Year, Intransit_Lead_Time, and Manufacturing_Time, they are no longer relevant to the regression model and we removed them from the Raw-Data dataframe.

4. One-hot encoding

After cleaning the data, we one-hot encoded all categorical variables to ensure all levels of categorical data were represented as numerics. Such representation allows for easier calculations of correlation between variables and benefits the model-building process.

5. Descriptive Statistics

To understand the distributions and frequencies of the all variables in the Raw-Data data frame, we generated the descriptive statistics for each variable.

a. Descriptive statistics for Categorical Variables.

Tables 4-7 include the descriptive statistics for all categorical variables, including LOB, Origin, Ship Mode, and Quarter.

LOB distribution:

LOB	Count	Weight
Product A	2957	32.31%
Product B	5932	64.82%
Product C	262	2.86%

Table 4: LOB Descriptive Statistics

Origin Distribution:

Origin	Count	Weight
Site A	3477	38.00%
Site B	1572	17.18%
Site C	1659	18.13%
Site D	2443	26.70%

Table 5: Origin Descriptive Statistics

Ship Modes Distribution:

Ship Mode	Count	Weight
Air	3829	41.80%
FastBoat	532	5.81%
Ground	2443	26.70%
Ocean	2347	25.65%

Table 6: Ship Mode Descriptive Statistics

Quarter Distribution:

Quarter	Count	Weight
Q1	3874	43.04%
Q2	3638	40.42%
Q3	1446	16.06%
Q4	43	0.48%

Table 7: Quarter Descriptive Statistics

b. Descriptive statistics for Numerical Variables.

Tables 8 and 9 include the descriptive statistics for all numerical variables, including Year and Intransit Lead Time. Since Year only takes on two values despite being a numerical variable, we included a table to record the frequency of each level of Year instead of including a density plot.

Year Distribution:

Year	Count	Weight
2019	43	0.48%
2020	8958	99.52%

Table 8: Year Descriptive Statistics

Intransit Lead Time Distribution:

Min	1st Quartile	Median	Mean	3rd Quartile	Max	Std. Deviation
1	5	8	13.85	23	86	12.3384

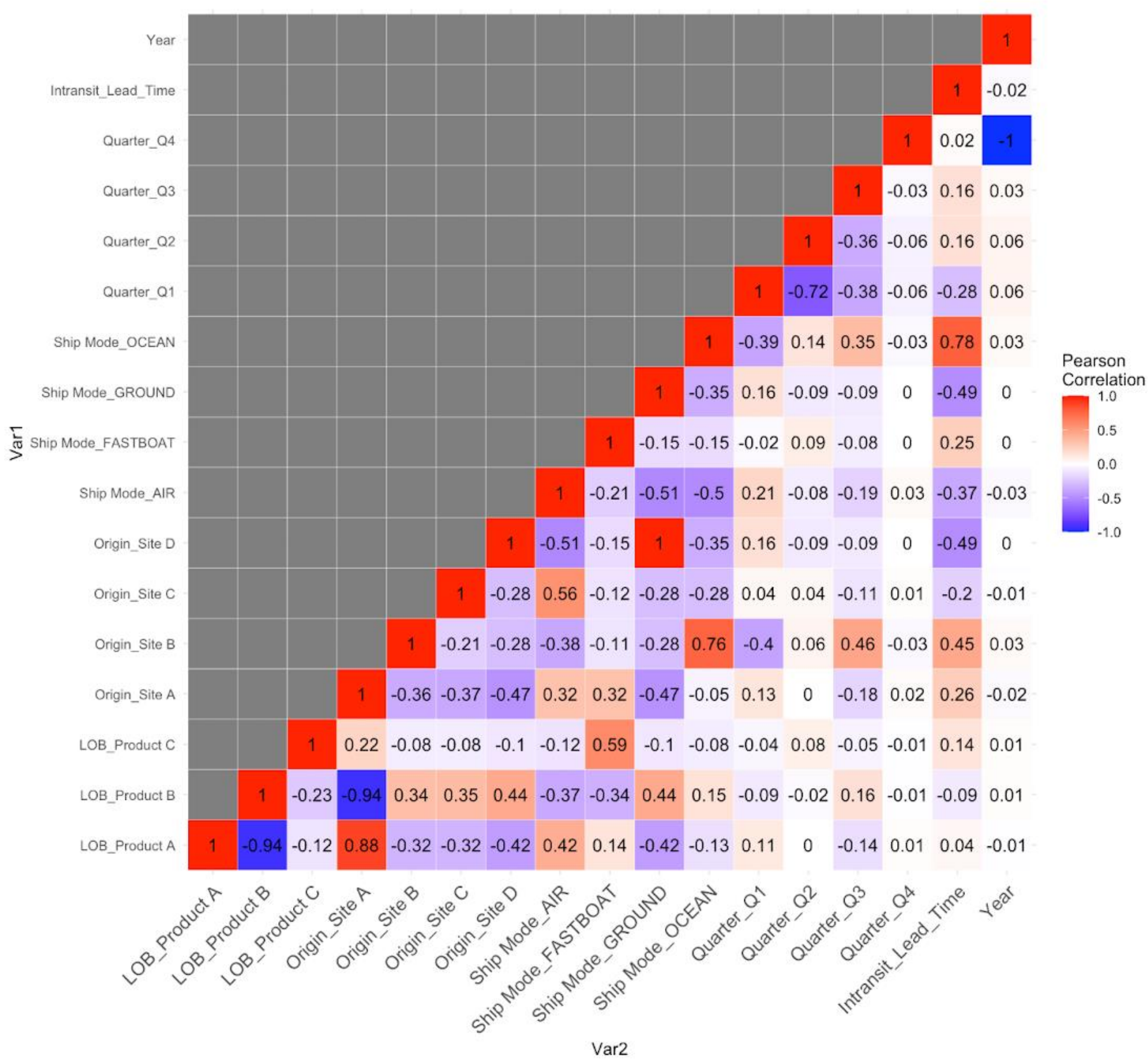
Table 9: Intransit Lead Time Descriptive Statistics

IV. Modeling Overview

1. Correlation Plots

This section consists of two visualizations: a Correlation Plot that summarizes the correlation between all variables and a Correlation Table that lists out all predictor pairs exhibiting multicollinearity.

Below is the Correlation Plot with correlation between all input variables in the regression model:



Plot 10: Correlation Plot between all variables

The following table identifies predictor pairs with high correlation or exhibiting multicollinearity:

Variable 1	Variable 2	Value
Ship Mode_GROUND	Origin_Site D	1
Origin_Site A	LOB_Product A	0.883
Intransit_Lead_Time	Ship Mode_OCEAN	0.781
Ship Mode_OCEAN	Origin_Site B	0.763
Quarter_Q2	Quarter_Q1	-0.714
LOB_Product B	LOB_Product A	-0.938
Origin_Site A	LOB_Product B	-0.0941

Table 11: Most correlated predictor pairs

2. Explanation of Model Parameters and Key Statistics

1. Intercept: The value of the outcome variable when all levels of the categorical variables are excluded from the model and all numeric variables are 0, or the base level.
2. Coefficient: The influence each independent variable has on the outcome variable, or the amount by which it will increase/decrease the outcome variable when added to the model.
3. Standard Error (std. error): The average distance that the actual values fall from the regression line.
4. p-value: The probability that an event happens entirely by accident.
5. Significant Level: The statistical significance of the relationship between a predictor variable and an outcome variable which is determined based on the p-value.
6. R-squared: The explanatory value of the model or the percentage of the variance in the outcome variable that can be accounted for by the model.
7. Adjusted R-squared: The explanatory value with a penalization for the addition of less- or non-significant predictors to the model.
8. Residual Standard Error (RSE): The average absolute size of the residuals adjusted for the degrees of freedom of the model (number of predictors included).
9. Variance Inflation Factor (VIF): A measure of the amount of multicollinearity among multiple predictor variables.
10. 'Var1 : Var2' notation: The interactive effect between two predictor variables.

3. Models Development

a. Model A

Variable	Coefficient	Std. Error	p-value	Signif. Level	VIF
Intercept	**80.8241**	**1.1002**	**$< 2 \times 10^{-16}$**	***	
Origin_Site D : Ship Mode_GROUND	-76.8091	1.1065	$< 2 \times 10^{-16}$	***	113.570382
LOB_Product C	-48.7725	1.1361	$< 2 \times 10^{-16}$	***	17.014779
LOB_Product A : Origin_Site A	-46.8548	1.0847	$< 2 \times 10^{-16}$	***	121.98276
Origin_Site C	-46.3333	1.0858	$< 2 \times 10^{-16}$	***	82.945189
LOB_Product B : Ship Mode_OCEAN	-31.2113	1.0686	$< 2 \times 10^{-16}$	***	85.947493
Ship Mode_AIR	-25.5997	0.2204	$< 2 \times 10^{-16}$	***	5.603697
Origin_Site B	-23.5022	0.3116	$< 2 \times 10^{-16}$	***	6.548398
Ship Mode_FASTBOAT	-6.7442	0.3028	$< 2 \times 10^{-16}$	***	2.37913
Quarter_Q4	6.6821	0.6755	$< 2 \times 10^{-16}$	***	1.011666
Quarter_Q3	-0.4790	0.1459	1.030×10^{-3}	**	1.367745
Quarter_Q1	-0.3783	0.1097	5.66×10^{-4}	***	1.398177

Table 12: Model A Key Statistics.

- Significant codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
- Residual standard error: 4.394 on 9139 degrees of freedom
- Multiple R-squared: 0.8733, Adjusted R-squared: 0.8732
- F-statistic: 5728 on 11 and 9139 DF, p-value: $< 2.2 \times 10^{-16}$

Model A is the first model we built after the data cleaning process. As a result, it is the only model out of the three models that has all levels of all variables. During the model building process, we used the correlation matrix to identify predictors pairs with perfect/high multicollinearity and attempted to eliminate it while retaining as many variables as possible. Instead of including the main effects of all variables, we included only the interactive effects between predictor pairs with perfect/high multicollinearity in the model. The pairs include: LOB_Product A – Origin_Site A, LOB_Product B – Ship Mode_OCEAN, Origin_Site D – Ship Mode_GROUND. Consequently, although the model has an adjusted R-squared of 0.8732 – the highest of the three models, the predictors also exhibit staggering VIFs (e.g. \`LOB_Product A\`:\`Origin_Site A\` has a VIF of 121.98, Origin_Site D : Ship Mode_GROUND has a VIF of 113.57) because the high level of multicollinearity causes the R-squared to be inflated. In this model, only Ship Mode_FASTBOAT, Quarter_Q4, Quarter_Q3, and Quarter_Q1 have reliable VIF values, all of which are below 5.

The model has an R-squared of 0.8733, an adjusted R-squared of 0.8732 and an RSE of 4.394. The adjusted R-squared indicates that the variables in the model can account for 87.32% of the variance in Intransit Lead Time. The overall p-value of the model is below 2.2×10^{-16}, signifying a statistically significant relationship between all predictors and the outcome variables.

Model A has an intercept of 80.8241, meaning that when all levels of the categorical variables are excluded from the model and all numeric variables are 0, Intransit Lead Time has an estimated value of 80.8241 days. The coefficients of model A can be divided into 2 groups – coefficients with a weight below 30 (lines in dark blue) and coefficients with a weight above 30 (lines in light blue):

Variables in Group 1 have a coefficient above 30, meaning that they increase or decrease Intransit Lead Time by more than 30 days when added to the model. The predictor with the highest coefficient is Origin_Site D:Ship Mode_GROUND and it is estimated to decrease lead time by 76.81 days. This is a significant amount given that the base level for lead time is only 80 days. These predictors also have a higher standard error compared to Group 2: all the std. error statistics for Group 1 are larger than 1, suggesting a higher margin of error. This is consequential of the fact that these predictors have higher predicting ability, thus making the Intransit Lead Time more sensitive to their changes. All variables in Group 1 have statistically significant relationships with lead time as they all have Significant Code of 3 stars or a p-value below 0.001, which is very low and indicative of high predictive value.

Variables in Group 2 have coefficients below 30, meaning that they increase or decrease Intransit Lead Time by less than 30 day when added to the model. The predictor with the most significant coefficient is Ship Mode_AIR with 25.6 days while the coefficient with the least weight is Quarter_Q1 with 0.38 days. This implies that an order placed in Quarter 1 will increase lead time by an estimate of 0.38 days or less than 10 hours. We can see that Intransit Lead Time is not as responsive to changes in values of Group 2 predictors as to those of Group 1 predictors. The lower coefficients assign less weight to these variables as they have less power in predicting lead time, resulting in lower margin of error (std. error) and less statistically significant relationships with the outcome variable. The only predictor with a p-value larger than 0.001 in this model is Quarter_Q3, which is in Group 2 and has a coefficient of -0.479. Group 1 variables have more influence over Intransit Lead Time and also more importance in the model.

b. Model B

When building model B, the goal of our group was to further lower the VIF values of model A whilst retaining as much data as possible. The variables with the highest VIFs in Model A include the 3 interactive effect variables and Origin_Site C, all of which have a VIF value of over 80, indicating extreme multicollinearity. We decided to eliminate Origin_Site C from model A to create model B because eliminating any of the other three variables significantly not only lowers the VIFs but also diminished the adjusted R-squared to below 80.

Variable	Coefficient	Std. Error	p-value	Signif. Code	VIF
Intercept	**34.9917**	**0.2609**	**$< 2 \times 10^{-16}$**	***	
Origin_Site D : Ship Mode_GROUND	-30.9505	0.2885	$< 2 \times 10^{-16}$	***	6.438509
Ship Mode_AIR	-25.6271	0.2414	$< 2 \times 10^{-16}$	***	5.60365
Origin_Site B	-19.7931	0.3277	$< 2 \times 10^{-16}$	***	6.038895
LOB_Product B : Ship Mode_OCEAN	11.6498	0.3992	$< 2 \times 10^{-16}$	***	10.005322
Ship Mode_FASTBOAT	-6.7869	0.3315	$< 2 \times 10^{-16}$	***	2.379104
Quarter_Q4	6.7557	0.7397	$< 2 \times 10^{-16}$	***	1.011659
LOB_Product C	-2.8889	0.4016	6.83×10^{-13}	***	1.773192
LOB_Product A : Origin_Site A	-0.9791	0.158	5.96322×10^{-10}	***	2.157322
Quarter_Q3	-0.9435	0.1593	3.291351×10^{-9}	***	1.360131
Quarter_Q1	-0.3393	0.1201	4.74×10^{-3}	**	1.39808

Table 13: Model B Key Statistics.

- Significant codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
- Residual standard error: 4.811 on 9140 degrees of freedom
- Multiple R-squared: 0.8481, Adjusted R-squared: 0.8479
- F-statistic: 5103 on 10 and 9140 DF, p-value: $< 2.2 \times 10^{-16}$

Model B has an R-squared of 0.8481 and an adjusted R-squared of 0.8479, meaning that the model can account for 84.79% of the variance in Intransit Lead Time. The overall p-value of the model is below 2.2×10^{-16}, signifying a statistically significant relationship between all predictors and the outcome variables. The RSE statistic is 4.811, higher compared to model A. However, the VIF values in this model is considerably lower than those of model A – all VIF values for all variables have reduced to 10 and below. For example, LOB_Product A:Origin_Site A which used to have a VIF of 121.98 now only has a VIF of 2.16. This indicates that the extreme multicollinearity exhibited in the previous model has been eliminated and model B, whilst still subject to some level of multicollinearity, has brought the VIF statistics to a more controlled level.

The elimination of Origin_Site C did not only have a remarkable effect on the VIFs but also on coefficients in the model. Compared to model A, the weights of all coefficients dropped substantially – the top coefficients of model A lie in the 30-80 range, whereas that of model B are in the 15-35 range. The standard errors for all variables in model B are also reduced to below 1, representing a lower margin of error than the previous model. Due to a narrower range in terms of the coefficient distribution and the fewer number of predictors, we will evaluate the coefficients together as one group.

The adjustments in the model resulted in the shift in weights of the predictors – predictors with high coefficients in model A such as LOB_Product B : Ship Mode_OCEAN, LOB_Product C, and LOB_Product A:Origin_Site A now are only in the mid-range in terms of importance among all variables. Only Origin_Site D : Ship Mode_GROUND still has the highest coefficient due to the fact that Origin_Site D and Ship Mode_GROUND are perfectly correlated and both exhibit moderate correlation with Intransit Lead Time (correlation = -0.492). On the other hand, some variables with previously low coefficients are assigned more importance in this model. Despite only being somewhat influential in model A, Ship Mode_AIR and Origin_Site B now rise to become the second and third most-weighted variables with coefficients of -25.63 and -19.79 respectively. The rest of the variables in the model hold the same level of importance compared to the previous model, which include Ship Mode_FASTBOAT, Quarter_Q4, Quarter_Q3, and Quarter_Q1. The weights of their coefficients are still roughly similar with Ship Mode_FASTBOAT at approximately -6, Quarter_Q4 at 6, and Quarter_Q3 and Quarter_Q1 below 1.

c. Model C

From model B, it is evident that the VIFs are unlikely to drop to below 5 whilst maintaining the adjusted R-squared and retaining all variables. To construct model C, we experimented with all variables (including the eliminated and combined for interactive effects variables) and utilized the information we gathered from building models A and B. Our goal when building model C was to keep the adjusted R-squared at 0.8396 and above, lower the VIFs to 5 and below, and eliminate no more than 1 level from each of the LOB, Origin, Ship Mode, Quarter variables.

Variable	Coefficient	Std. Error	p-value	Significant Code	VIF
Intercept	**8.7958**	**0.1244**	**$< 2 \times 10^{-16}$**	***	
Ship Mode_OCEAN	24.2428	0.2253	$< 2 \times 10^{-16}$	***	4.196215
Ship Mode_FASTBOAT	18.5188	0.2678	$< 2 \times 10^{-16}$	***	1.701988
Origin_Site A : LOB_Product B	16.7571	0.3543	$< 2 \times 10^{-16}$	***	1.491234
Quarter_Q4	6.4402	0.7063	$< 2 \times 10^{-16}$	***	1.011514
Origin_Site B	-6.3914	0.2336	$< 2 \times 10^{-16}$	***	3.366734
Ship Mode_GROUND	-4.6083	0.1357	$< 2 \times 10^{-16}$	***	1.563558
LOB_Product C	-1.8763	0.3566	1.465×10^{-7}	***	1.53369
Quarter_Q3	-0.8455	0.152	2.72×10^{-8}	***	1.35763
Quarter_Q1	-0.6154	0.1144	7.65×10^{-8}	***	1.390523
Origin_Site C	0.2358	0.1509	1.18×10^{-1}		1.465345

Table 14: Model C Key Statistics.

- Significant codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
- Residual standard error: 4.594 on 9140 degrees of freedom
- Multiple R-squared: 0.8615, Adjusted R-squared: 0.8614
- F-statistic: 5686 on 10 and 9140 DF, p-value: $< 2.2 \times 10^{-16}$

Model C has an R-squared of 0.8615 and an adjusted R-squared of 0.8614, meaning that the model can account for 86.14% of the variance in Intransit Lead Time. Despite having few variables compared to model A and B, model C still has a relatively high R-squared due to its inclusion of more main effects of variables and fewer interactive effects. The RSE statistic for this model is 4.594, which is higher compared to model A but lower compared to model B. The overall p-value of the model is below 2.2×10^{-16}, signifying a statistically significant relationship between all predictors and the outcome variables. For each of the input variables (LOB, Origin, Ship Mode, Quarter), only 1 level is eliminated:

- Product A for LOB.
- Site D for Origin.
- AIR for Ship Mode.
- Q2 for Quarter.

All predictors in this model have a VIF values below 5, with the highest values being VIF = 4.196 of Ship Mode_OCEAN and VIF = 3.367 of Origin_Site B. This result from the high correlation between Ship Mode_ OCEAN and Origin_Site B (correlation = 0.763). However, the remaining predictors have VIFs below 2, indicating that multicollinearity is no longer a significant problem.

With the intercept at 8.7956 and all the coefficients within the 0-25 range, model C is materially more reliable than model A and model B. The addition of predictors to the model will not increase or decrease lead time by overwhelming amounts as in model A where the intercept was over 80 and the highest coefficient had a weight of over 76. Ship Mode_OCEAN adds the most weight to the base level of lead time when added to the model because it has the highest correlation with Intransit Lead Time (correlation = 0.781). Ship Mode_FASTBOAT, however, is surprisingly assigned a higher importance than previous models. Its coefficients for model A and B stand at approximately -6, whereas the figure for this model is 18.52. Such variation can be explained by the different combinations of predictors in each model which influence the weight each predictor holds. The interactive effect between Origin_Site A and LOB_Product B ranks third in terms of importance due to its coefficient of 16.76. From the correlation table, it is evident that each pair from the Site A, Product A, Product B group exhibit an extraordinary level of correlation, or multicollinearity. To retain useful information whilst preventing the effects of multicollinearity, we excluded LOB_Product A from the model and added the interactive effect between Origin_Site A and LOB_Product B.

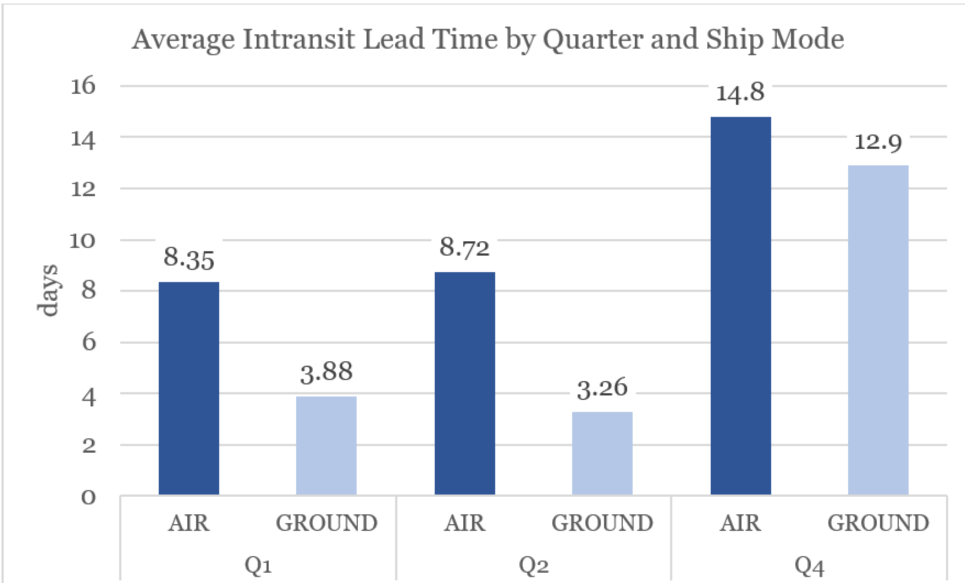
Except for Ship Mode_OCEAN, Ship Mode_FASTBOAT, and Origin_Site A : LOB_Product B, the other predictors have coefficients with weight below 7. The narrowing down in the range of the coefficients suggests lower fluctuation and more accurate determination of Intransit Lead Time. This can also be seen in the std. error statistics as most predictors have std. error below 0.4, significantly lower compared to model A and model B. All variables have a statistically significant relationships with Intransit Lead Time as all have a significant code of 3-star except for Origin_Site C – the variable with the lowest coefficient in the model.

V. Results and Discussion of Research Questions

1. Research Question 1

Has the logistics lead time reduced since Q2, when compared with Q4 and Q1 for AIR and GROUND ship mode?

The graph below shows the Average Intransit Lead Time calculated based on Quarter and Ship Mode:



Graph 15: Average Intransit Lead Time by Quarter and Ship Mode

Answer: When compared with Q4 and Q1, the logistics lead time for GROUND Ship Mode during Q2 is lower because the average for Q2-GROUND is only 3.26 days whereas the figures for Q4 and Q1 are 12.9 days and 3.88 days respectively. The same, however, is untrue for the AIR Ship Mode. From Q1 to Q2, the lead time for AIR Ship Mode increased slightly by approximately 0.4 days and jumped to 14.8 days in Q4. As a result, whilst the average logistic lead time in Q2 did not reduce since Q1, it is significantly lower compared to that of Q4.

2. Research Question 2

How many in-transit days can be saved if we bring inventory by AIR from Site A compared to Site C?

The table below shows Average Intransit Lead Time by AIR Ship Mode and Origin:

Ship Mode	Origin	Average Intransit Lead Time
AIR	Site A	8.12
	Site B	31.7
	Site C	8.72

Table 16: Average Intransit Lead Time by AIR Ship Mode and Origin

Answer: 0.6 days can be saved if inventory is brought by AIR from Site A compared to Site C.

3. Research Question 3

Optimization Question:

a. *Can we fulfil all 100K units before the deadline with the available budget based on the current manufacturing and logistics lead times?*

The total time XYZ has to manufacture and ship each 20,000-unit block of the order is calculated in the table below:

Block	Units	PO Download Date	Receipt Date	Total Logistic Time
1	20,000	9/21/2021	10/27/2021	37 days
2	20,000	9/22/2021	10/27/2021	36 days
3	20,000	9/23/2021	10/27/2021	35 days
4	20,000	9/24/2021	10/27/2021	34 days
5	20,000	9/25/2021	10/27/2021	33 days

Table 17: Total Logistic Time by unit blocks.

The table below summarizes the average amount of time to manufacture and ship Product A from Site A to the Warehouse based on Ship Mode:

LOB	Origin	Ship Mode	Average Manufacturing Time	Average Intransit Lead Time	Total Time
Product A	AIR	AIR	6.28	8.13	14.41
		FASTBOAT	8.42	27.2	35.62
		OCEAN	13.6	33.9	47.5

Table 18: Average Logistic Time for Product A from Site A by Ship Mode.

Based on the table, we can see that the average total time between PO Download Date and Receipt Date using AIR, FASTBOAT, and OCEAN Ship Modes are 14.41 days, 35.62 days, and 47.5 days respectively. However, the total time since Site A is provided with Purchase Orders (PO) until the deadline (October 27th) are only 37 days. Consequently, we will not consider the OCEAN Ship Mode for this order.

<u>Option A:</u> The order is shipped entirely using the AIR Ship Mode:

Whilst this Ship Mode works perfectly with the deadline requirements of the order – even the final 20,000 units with the last PO Download Date can arrive on time – the total costs for the 100,000 units will exceed the freight budget ($\$22 \times 100{,}000 = \$2{,}200{,}000 > \$1{,}500{,}000$). Therefore, we eliminate this option.

<u>Option B:</u> The order is shipped entirely using the FASTBOAT Ship Mode:

This option provides XYZ with a cost-effective solution for shipping as the total cost will be below the budget ($\$13 \times 100{,}000 = \$1{,}300{,}000 < \$1{,}500{,}000$). However, in terms of timing, only the first 40,000 units of the order will arrive at the Warehouse on time whereas the remaining 60,000 units will not arrive before October 27th due to their later PO Download Date. Therefore, we also eliminate this option.

<u>Option C:</u> The order is shipped using both AIR and FASTBOAT:

To meet both the requirements for time and budget, we need to distribute the 100,000 units between AIR Ship Mode and FASTBOAT Ship Mode. We should use the FASTBOAT Ship Mode for the first 40,000 units because it allows the company to meet the deadline whilst keeping cost to the minimum. As for the remaining 60,000 units, the company needs to opt for the AIR Ship Mode to ensure that the order is completed on time.

The total cost for allocating the order between two Ship Modes will be as follows:

$$\$13 \times 40{,}000 + \$22 \times 60{,}000 = \$1{,}840{,}000$$

Because the total cost calculated above exceeds the company's budget for freight ($1,500,000), this option is also eliminated.

Answer: It is impossible for XYZ to fulfil all 100K units before the deadline with the available budget based on the current manufacturing and logistics lead times.

b. If not, please come up with an optimized volume of product A needs to be lifted by air/Ocean/Fast Boat to meet maximum customer demand depending on the ship mode lead times and budget available?

To determine the optimized volume of product A to be lifted based on the ship mode lead times and budget available, we first need to determine the number of units to allocate to each Ship Mode. From the analysis in Answer 1, only two out of the total three Ship Modes available for shipping Product A from Site A can allow part of or the entire order to arrive at the Warehouse on time, which are AIR and FASTBOAT. To minimize costs, the first 40,000 units of the order must be shipped using the FASTBOAT

Ship Mode. As for the remaining units, since FASTBOAT Shipping will not meet with the customer's deadline, the only option is AIR Shipping. The maximum number of units that XYZ can deliver without exceeding the budget is calculated as follows:

Maximum Volume (AIR) = ($1,500,000 –($13 ×40,000 units))/$22 = 44,545.45 units

Optimized Volume = Maximum Volume (FASTBOAT) + Maximum Volume (AIR)

= 40,000 units + 44,545 units

= 84,545 units.

Answer: The optimized volume that XYZ can fulfil whilst still meeting with the requirements for the deadline and the budget is 84,545 units, 40,000 of which is shipped by FASTBOAT and the remaining 45,545 units is shipped by AIR.

c. *Create a proposal to request for additional budget to fulfil all 100K units by the deadline?*

Answer:

Budget Proposal

XYZ has received an order placement for 100,000 units of Product A and has a responsibility to fulfil the order before October 27th, 2021 with a $1,500,000 freight budget. However, our calculations show that it would be impossible to ensure that all products arrive at the Warehouse on time whilst keeping costs below budget. The table below lays out the cost justification for each cost element of the order:

	Units	Ship Mode	CPB	Cost
Product A	40,000	FASTBOAT	$13	$520,000
	60,000	AIR	$22	$1,320,000
Total	100,000			$1,840,000
			Budget	$1,500,000
			Budget Overrun	**$340,000**

Table 19: Total Freight Cost to complete order.

The completion of this order before the customer's due date is important not only because of the order's size but also because of XYZ Company's image as a whole. Failure to deliver products on time will result in a decline in the Company's prestige and a hostile relationship with the customer. Therefore, we propose that the freight budget be raised to $1,840,000 or an additional budget of $340,000 be provided to maintain XYZ's position as a reliable supplier.

d. *Provide an analysis to leadership team on total logistics cost vs total revenue vs profit margin with the initial budget provided and proposed budget for them to decide what trade-offs to make?*

Answer:

The table below provides a comparison and cost analysis on Total Logistics Cost vs. Total Revenue vs. Profit Margin between the initial budget and the proposed budget for fulfilling the order. Production Plan 1 refers to the Company's option to maintain budget whilst fulfilling only part of the order and Production Plan 2 refers to the option to exceed the predetermined budget to fulfil the entire order.

XYZ Company	Production Plan 1	Production Plan 2
Beginning Budget Balance	$1,500,000	$1,500,000
Logistics Costs		
Freight Cost (FASTBOAT)	$520,000	$520,000
Freight Cost (AIR)	$979,990	$1,320,000
Total Logistics Costs	$1,499,990	$1,840,000
Ending Budget Balance	$10	($340,000)
Total Cost	$30,968,990	$31,309,000
Total Revenue	$28,153,485	$33,300,000
Profit (Loss)	($2,815,505)	$1,991,000
Profit (Loss) Margin	-10.00%	5.98%

Table 20: Cost Analysis and Comparison based on initial budget and proposed budget

<u>Total Cost Justification:</u>

According to the original plan, all 100,000 units need to be delivered to the Warehouse by the customer's deadline, which would allow the profit margin to be 7%. The estimated total cost was:

$$(\$333 \times 100{,}000 \text{ units}) \times 0.93 = \$30{,}969{,}000$$

For Production Plan 1, XYZ spends below budget by $10, therefore the actual total costs would be: $30,969,000 - $10 = $30,968,990

For Production Plan 2, XYZ spends above budget by $340,000, therefore the actual total costs would be: $30,969,000 + $340,000 = $31,309,000

VI. Conclusion

Out of the three models developed, Model C is the most reliable due to its high predictive value as well as low levels of multicollinearity. Whilst the weight of the independent variables in the model and their relationships with Intransit Lead Time may not give the most accurate predictions of logistic lead time, they do provide us with information about factors influencing lead time and directions for delivery improvement. In regard to XYZ Company's Case Study, we determined that it is impossible to fulfil all 100K units before the deadline with the available budget based on the current manufacturing and logistics lead times. The Company will benefit more from increasing the Freight Budget than maintaining the same budget as the former allows XYZ to recognize a positive profit margin whereas the latter prevents XYZ from fulfilling the entire order as well as maximizing profit.