# Analysis of B-Craft Beer

Mirna Arivalagan - 220142881 | Descriptive Analytics and Visualisation – MIS771
Trimester 2 2021

# Introduction:

This business report provides a thorough analysis of B-Craft brewery's sales data and pale ale production data. This report will cover an analysis and projection of sales order quantities, the likelihood of B-craft being recommended, and the report will also cover the forecasted amount of pale ale that will be produced over the next fiscal year. The analysis was done using two datasets, where the first dataset are survey responses from 200 of B-Craft's customers, where the customers were rating B-Craft on nine different attributes. The second dataset used in this report is a summary of pale ale production in litres between quarter 1 of 2011 to quarter 4 of 2020.

# 2 – Estimating Order Quantities:

## 2.1 – Scatter Plot and Correlation Analysis:

Prior to building the multiple regression model to estimate the order quantity, I have done some scatter plot and correlation analyses to understand the relationships between the Dependent Variable (DV) of Order quantity and the 14 Independent Variables (IV) that were present in the dataset provided for the purposes of this analysis. By analysing the relationships prior to building the model, this will help us decide what variables to include in the model. Below is a breakdown of what I have found in the relationships between the DV and IV:

**Positive Linear Relationship**

- Loyalty v Order Quantity – We can observe an increase in order quantity as the loyalty increases. (Correlation: 0.41)
- SM Presence v Order Quantity – As the perception of the social media presence increases, there is an increase in order quantities. (Correlation: 0.67)
- Quality v Order Quantity – There is an increase in order quantities as the perception of the beer quality increases. (Correlation: 0.43)
- Dist Channel v Order Quantity – There is a slight increase in order quantities for the direct distribution method. (Correlation: 0.40)
- Advert v Order Quantity – There is a slight increase in order quantities as the perception of B Craft's advert campaign increases. (Correlation: 0.24)
- Brand Image v Order Quantity – As the perception of B-Craft's brand image improves, we can observe an increase in Order Quantities. (Correlation: 0.34)
- Order Fulfillment vs Order Quantity – As the perception and efficacy of the order fulfillment increases, the order quantities increase as well. (Correlation: 0.31)
- Shipping Speed vs Order Quantity – The order quantities increase as the speed of shipping increases. (Correlation: 0.43)
- Shipping Cost v Order Quantity – As the perception of the affordability of the shipping costs increases to being more affordable, the order quantities increase as well. (Correlation: 0.50).

- Recommendation vs Order Quantity – There is a positive increase in order quantities for customers who have rated that they would recommend B-Craft compared to those who have rated that they will not recommend. (Correlation: 0.51).

**Negative Linear Relationship**

- Cust Type vs Oder Quantity – There is a decrease in order quantities for purchases made in pubs, bars and restaurants compared to bottle shops. (Correlation: -0.05).
- Region vs Order Quantities – There is a decrease in order quantities in Regional South Australia compared to metropolitan Adelaide. (Correlation: -0.12).
- Comp Pricing vs Order Quantity – We can observe a decrease in order quantities as the perception of competitive pricing improves. (Correlation: -0.22).

**No Linear Relationship**

- Flex Price v Order Quantity – There is no relationship between these 2 variables as an improvement in price flexibility does not correlate to a change in order quantities. (Correlation: 0).

From the correlation matrix, we have identified that there is a set of IVs that are highly correlated with each other ($R > 0.8$). Shipping Speed and Shipping Costs are highly correlated with each other ($R = 0.84$), this will cause the issue of multicollinearity if we include both these IVs in our regression model, so we needed to drop one of them. I have dropped Shipping Speed as the correlation of this IV with order quantity is 0.43 which is lower than Shipping Costs correlation of 0.50. I have also dropped Flex Price from the model, as there is virtually no relationship of this IV with the DV. (Refer to Appendix 2.1)

## 2.2 – Initial Regression Model – Multicollinearity Addressed:

For the first Linear Regression model, we ran it with the IVs of Shipping Speed and Flex price removed due to multicollinearity. We ran the linear regression model on a total of 200 observations. The following is the list of initial IVs that were included:

- Loyalty
- Cust Type
- Region
- Dist Channel
- Quality
- SM Presence
- Advert
- Brand Image
- Comp Pricing
- Order Fulfillment

- Shipping Cost
- Recommend

The R Square value for the initial model was 66%. This means that 66% of the variation in order quantities can be explained by the joint variation of the remaining 12 IVs that we have included in the model. The remaining 34% of the variation in order quantities can be explained by other IVs that were not included in this mode. 66% also indicates that this model has moderate predictive powers.

The standard error of this model sits at 0.81, which means that the average error that we may commit using this model is a total of 810 beer bottles. This is the standard deviation of values around the prediction line, meaning we may over predict or under predict around 810 bottles.

At 5% significance level, there is sufficient evidence to conclude that this model is statistically significant, and that at least one IV has statistically significant relationship that can either drive up or decrease the order quantities. We were able to come to this conclusion due the F Significance value being 0.00 < 0.05 alpha.

In our initial model, the following IVs have been identified as not being statistically significant due to the P-Value being greater than 0.05 alpha, thus we will need run the model in multiple iterations to remove these IVs one by one until our final model has all IVs P-values below 0.05.

- Cust Type (P-Value: 0.29)
- Region (P-Value: 0.14)
- Dist Channel (P-Value: 0.79)
- Advert (P-Value: 0.91)
- Brand Image (P-Value: 0.09)

The intercept of our initial model is 7.24. What this means is that average order quantities B-Craft can expect is 7240 beer bottles, where there is no quality, SM Presence, Advert, Brand Image, Comp Pricing, Order Fulfillment and Shipping Cost. And where the beers are sold in bottle shops, in Adelaide and through a distribution network, and where the customers will not recommend B-Craft beer. This intercept has no practical use, as this would mean that B-Craft is able to sell beers where there are no orders being filled. (Refer to Appendix 2.2)

## 2.3 – Final Regression Model

After running the regression model through 6 iterations and removing each IVs that are not statistically significant, I was able to produce a final model that can be used to predict order quantities.

The R Square value of this final model has decreased from 66% of the initial model to 65%. However, the adjusted R Square value has stayed the same across both models, this is a good indication that we have dealt with independent variables that are insignificant to the model,

thus addressing the limitation of R square where when additional IVs are included in the model, more often than not, the R Square score will either stay the same or increase, thus giving us the perception that the quality of the model has improved when it has not. From this final model, we can conclude that 65% of the variation in order quantities can be explained by the variation in Loyalty, Quality, Social Media Presence, Competitive Pricing, Order Fulfillment, Shipping Cost and Recommend. The remaining 35% of the variation in order quantities can be explained by other statistically significant IVs that are not included in this model. This indicates that the model has moderate predictive powers.

Again, as per the initial model, the standard error has remained at 810 beer bottles.

I was also able to establish that this model is statistically significant as the F Significance is 0.00 < 0.05, and all the remaining IVs in the model has a P-Value that is less than 0.05. Thus at 5% significance level, there is sufficient evidence to conclude that at least one IV has a statistically significant relationship with Order Quantity.

The intercept for this model is 6.84, meaning that on average B-Craft can expect to sell 6840 beer bottles where all numerical variables are 0, and the categorical variable is set to the beers being sold in Bottle Ships, in Adelaide, through a distribution network and to customers who would not recommend B-Craft. Again, there is no practical use of this intercept here.

The linear regression equation for future prediction would be:

Order Quantity = 6.84 + (0.07*Loyalty) + (0.18*Quality) + (0.50*SM_Presence) + (-0.10*Comp_Pricing) + (-0.21*Order_Fulfillment) + (0.34*Shipping_Cost) + (0.33*Recommend)

Overall, where there is an increase of 1 unit in the following IVs, and assuming all other IVs are held constant, the change in order quantities will be as follows:

- **Loyalty**: Order quantities will increase by 0.07, or more precisely 70 beer bottles. I am 95% confident that on average the order quantities will increase between 40 to 100 beer bottles.
- **Quality**: Order quantities will increase by 0.18, or more precisely 180 beer bottles. I am 95% confident that on average the increase will be between 80 to 280 bottles.
- **SM Presence**: Order quantities will increase by 0.50, or 500 beer bottles. I am 95% confident that on average we will see the increase to be between 370 to 620 bottles.
- **Comp Pricing:** Order quantities will decrease by 0.10, or 100 bottles. I am 95% confident that the average reduction of orders will be between 20 to 180 bottles.
- **Order Fulfillment:** Order quantities will decrease by 0.21, or 210 bottles. I am 95% confident that on average we will see a reduction in orders between 40 to 390 bottles.
- **Shipping Cost:** Order quantities will increase by 0.34 or 340 bottles. I am 95% confident that we will see an average increase in orders between 200 to 480 bottles.
- **Recommend:** Order quantities will increase by 0.33, or 330 bottles. I am 95% confident that we will see an average increase in orders between 60 to 610 bottles.

Of the 200 observations, there was a total of 14 outliers that have been highlighted in the workbook that needs to be dealt with.

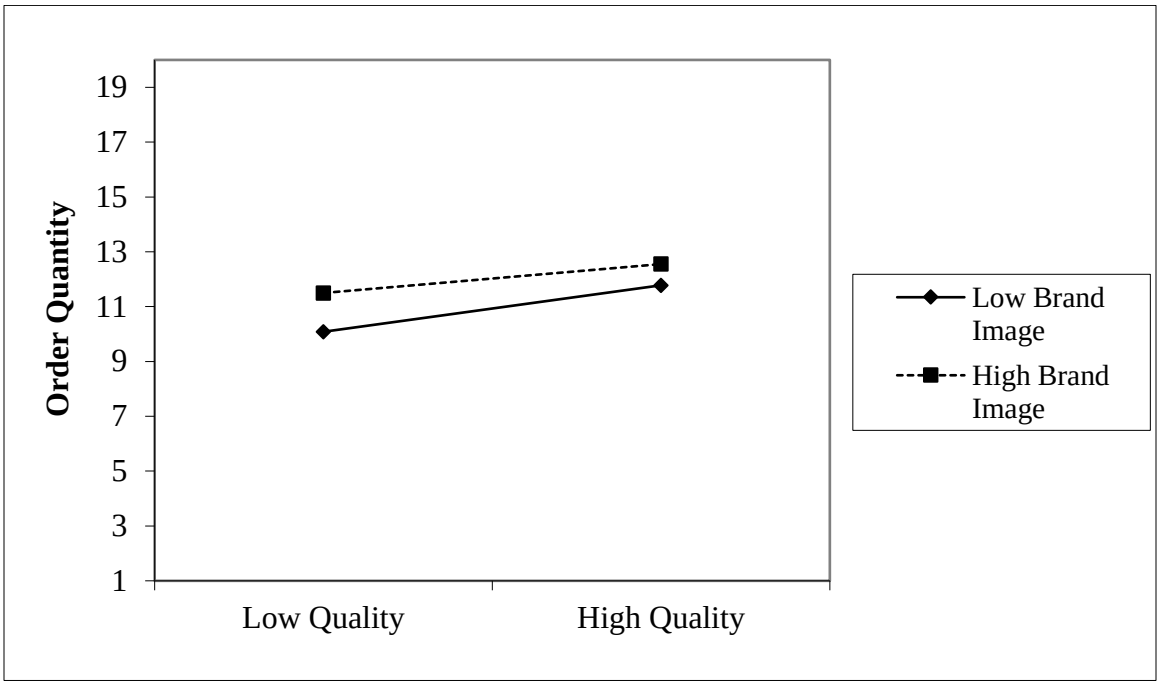
By looking at the residual plots, the dispersion of the data points looks balanced and there is no obvious pattern, which confirms that the model has not violated the independence of errors and equal variances assumptions. We can also observe from the normality plot and residuals distribution histogram that the model does not violate the assumption of normal distribution. The histogram appears to be somewhat of a bell shape, indicating normality. (Refer to Appendix 2.3)

## 3 – Interaction Effect of Brand Image and Quality on Order Quantities

To test the theory that the relationship between quality and order is stronger for customers who has higher perceptions of the brand image, a second linear regression model was built using the IV of Quality, Brand Image and Brand Image x Quality. For the interaction effect, Quality was set as the IV, Quantity is the DV and Brand Image is the moderator.

Both the R Square and Adjusted R Square values for this model was 35%, thus this means that 35% of the variation in order quantities can be explained by the joint variation in Quality and Brand Image. The remaining 65% of variation in order quantities is explained by the variation of other IV that are not included in model. This indicates that the model has low to moderate predictive powers.

As the P values for the three IVs that are in the model is below 0.05, and the F significance is 0.00, at 5% significance level, there is sufficient evidence to conclude that this model is statistically significant and that an interaction between the Quality and Brand Image exists.




We can observe in the interaction plot that Todd's theory can be justified as there is a slight interaction between Quality and Brand Image. The overall order quantity is highest for customers who have rated as high brand image and high quality. We can also observe across both instances that order quantities increase as the perception of quality increases. (Refer to Appendix 3)

## 4.1 – Predicting Likelihood of B-Craft Recommendations

To predict the likelihood of B-Craft being recommended, a Logistic Regression Model was developed using the recommended predictor variables of Dist Channel, Quality, Brand Image and Shipping Speed.

First, let's review the classification table results to assess the model's practical significance. The model's overall accuracy rate is 76%. This means that 76% of the total observations have been accurately classified as either will recommend or will not recommend. The remaining 24% misclassification rate could have been improved if additional statistically significant IVs are included in the model.

Of the 101 customers who rated that they would recommend B-Craft beers to others, 78 of them were accurately predicted as will recommend, the remaining 23 customers were misclassified as will not recommend. The overall accuracy for the segment of customers who rated will recommend is 77%.

Of the 99 customers who rated as they will not recommend B-Craft, 74 customers were accurately classified as will not recommend and the remaining 25 customers were misclassified as that they would recommend B-Craft. The overall accuracy of this segment is 75%.

The Proportional Chance Criterion (PCC) is 50%, Maximum Chance Criterion (MCC) is 51% and the Standard Rule of Thumb is 62.5% and the Cut Off is 50%. As the overall accuracy of this model is 76%, it is greater than that of PCC, MCC, Cut Off and Standard Rule of Thumb, we can conclude that this model has good practical usage. Hence it is significantly better than a random process of chance in classifying recommendation observations.

Next, we need to assess the model's statistical significance before concluding on the model's predictive powers. We can observe that all four IVs included in the model has P-values less than 0.05. There is also a reduction in LL values in this model compared to the baseline model (LL0: -138.62 & LL1: -95.38).

We can also use the Pseudo R Square values to assess the model's overall fit. Hosmer and Lemeshow R2 = 31%, Cox and Snell R2 = 35% and Nagelkerke R2 = 47%. These Pseudo R2 values means that between 31% to 47% of the variation in the likelihood of B-Craft being recommended is explained by the variation of the IVs that are included in this model.

Thus, from the evidence above, at 5% level of significance, there is sufficient evidence to conclude that this model has statistical significance and predictive powers.

As the coefficients of all four IV that are in the model have positive values, this means that the relationship of these IV with being recommended is positive in direction. From the model, were able to calculate the likelihood (odds) of being recommended. Below are the odds where there is an increase of one unit in each IV (assuming all other IVs are held constant):

- **Dist Channel:** An increase of 163.4% of the likelihood of B-Craft being recommended.
- **Quality:** An increase of 92.3% of the likelihood of B-Craft being recommended.
- **Brand Image:** An increase of 86.2% of the likelihood of B-Craft being recommended.
- **Shipping Speed:** An increase of 218.5% of the likelihood of B-Craft being recommended.

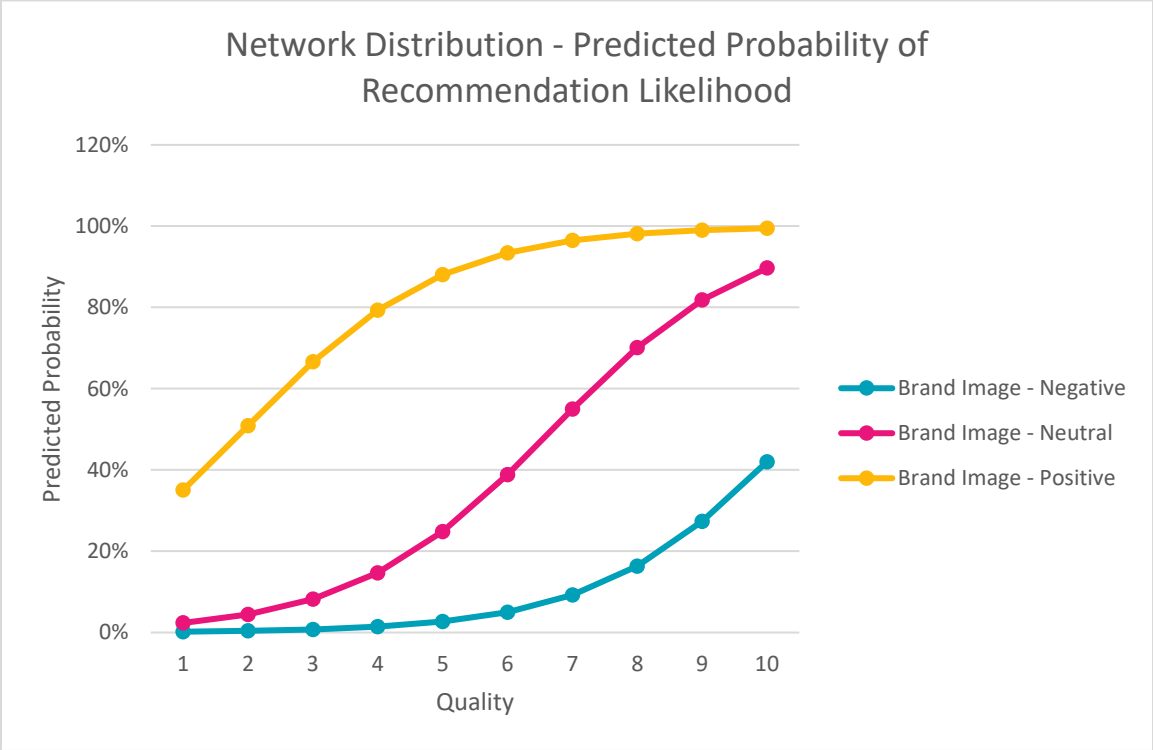
Finally, to conclude, we can analyse the ROC curve, we can observe from the plot that the area under the curve is close to 1 when it hits the accuracy ratio of 76%, this indicate that this model fits the data well. The AUC score is 0.85 which is close to 1, thus indicating that this model has moderate to strong predictive powers, and the model can discriminate between likely to recommend or not likely to recommend and that this is not due to chance. (Refer to Appendix 4.1).

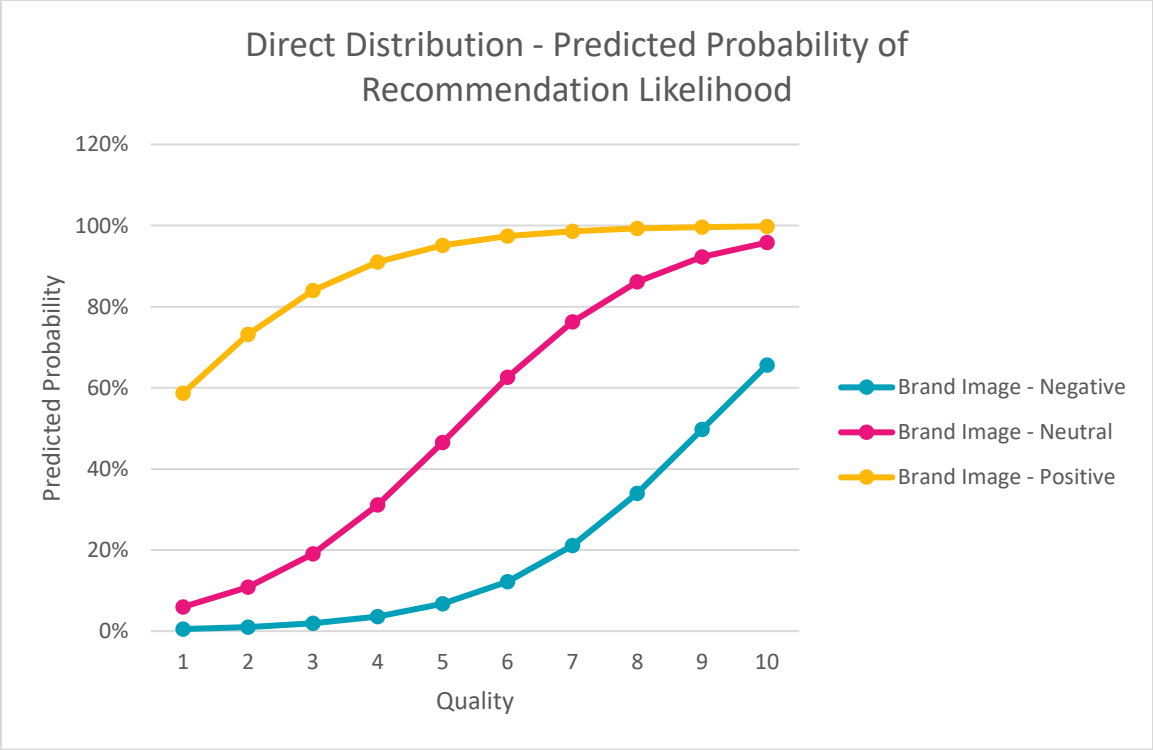
## 4.2 – Predicting Probability of Likelihood of Recommendation

To model out the probability of being recommended for a certain segment of customers, I have constructed a probability matrix for two key segments – the first being customers who are purchasing through network distribution, and the second being customers who purchase beers to direct distribution. We also included the IV of Quality and Brand Image in the matrix. (Refer to Appendix 4.2)

Across both segments, we can observe that positive brand image has the highest predicted probability of the likelihood of recommending B-Craft regardless of the perception of quality. There is a difference of about 24% between the probability for positive brand image between Network Distribution (35%) and Direct Distribution (59%) when the quality is at its lowest. The predicted probability for positive brand image across both segments evens out and is almost the same when the quality is scored at its highest value. There is also a difference of 24% between the probability for negative brand image between Network Distribution (42%) to Direct Distribution (66%) when the quality is at its highest. The only group that has similar probabilities are for Neutral brand image.

From the graph, we can also observe that the gap between Positive and Negative brand image with highest quality beer is decreased for Direct Distribution. Overall, we can conclude that the probability of being recommended is better for the segment that are purchasing through Direct Distribution, the predicted probability increases across all brand images as the quality increases, and when the quality is lowest, the predicted probability is higher in the Direct Distribution segment.


Network Distribution - Predicted Probability of Recommendation Likelihood
120%
100%
80%
60%
40%
20%
0%
Predicted Probability
1
2
3
4
5
6
7
8
9
10
Quality
Brand Image - Negative
Brand Image - Neutral
Brand Image - Positive


Direct Distribution - Predicted Probability of Recommendation Likelihood
120%
100%
80%
60%
40%
20%
0%
Predicted Probability
1
2
3
4
5
6
7
8
9
10
Quality
Brand Image - Negative
Brand Image - Neutral
Brand Image - Positive

## 5 – Timeseries Model to Forecast Pale Ale Production

To develop the pale ale production forecasting model, I have used the sales data set that has been collected between quarter 2, 2011 to quarter 4 2020. I have chosen to use the seasonal forecasting method, as when the sales across quarters were plotted, there are peaks and drops that can be observed from the historical data. Pale Ale production tends to increase every year in Q2 and there are significant drops in Q3 each year.

Prior to forecasting the production estimates, I have calculated the 4-moving average to smooth out the seasonality peaks and drops to ensure an accurate estimate for the forecast. While working out the indices, we also found that the average does not equal to 4 periods, thus this indicates that there are some seasonal irregularities, which have been handled by normalising the data. (Refer to Appendix 5).

The model has forecasted the following pale ale production amounts for 2021:

- Q1: 2159.97 litres
- Q2: 2659.97 litres
- Q3: 2001.23 litres
- Q4: 2209.71 litres

To assess the model accuracy, I have calculated the Mean Absolute Percentage Error (MAPE) and found this to be 4%. MAPE is a measure of error, where the lower the percentage, the better the model is predicting. What this means is that we can expect the actual pale ale production value for each quarter in 2021 to be off by 4% compared to the forecasted value if 2021 has similar seasonal trends as previous years.

## Conclusion

To conclude, the linear regression model that we have developed to predict the order quantity has a R2 value of 65% and adjusted R2 value of 64%, indicating that there are moderate predictive powers and that the IVs of Loyalty, Quality, SM presence, Comp Pricing, Order Fulfillment, Shipping Costs and Recommendation are variables that are driving the order quantities. Out of them, it appears that SM presence has the most weighting, where an increase in one unit of SM Presence, on average we can expect to see an increase of 500 bottles being ordered. I would recommend that B-Craft, invests in building their brand awareness on social media platforms.

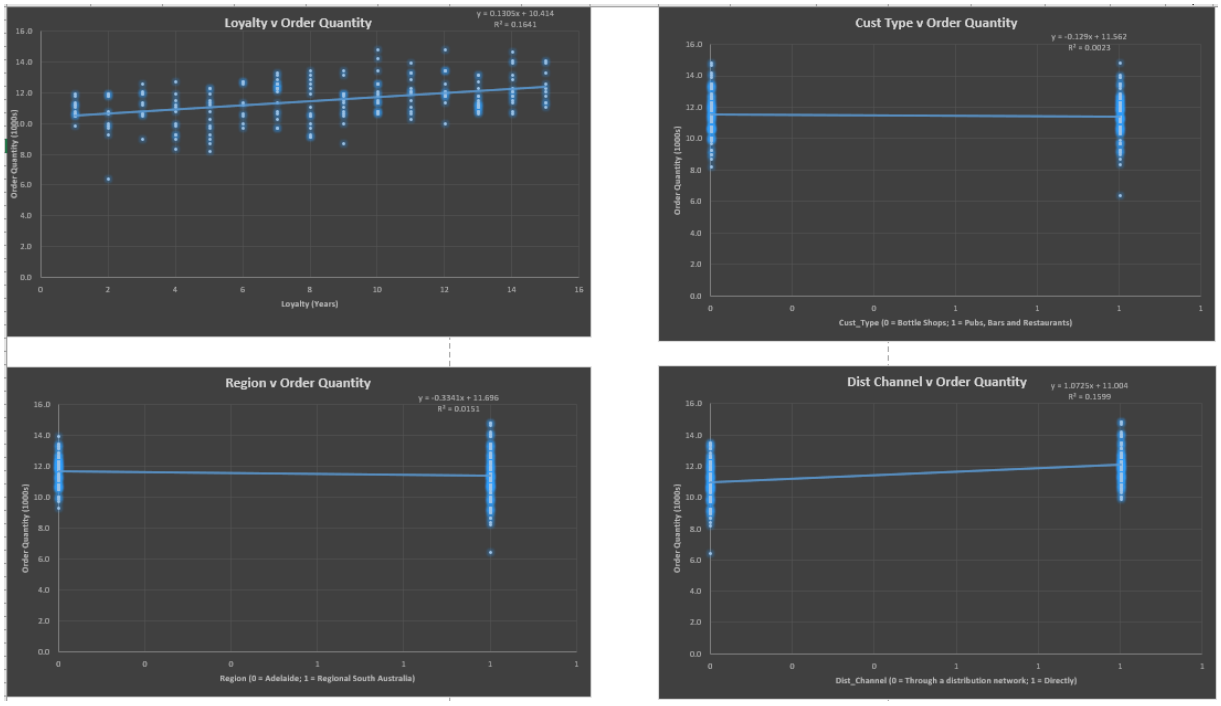
We can also conclude that Brand Image and Quality has an interaction effect that is driving up or decreasing order quantities. B-Craft should look at increasing their brand image by investing in some brand activation campaigns and investing to regularly check the quality of beers to ensure that they are consistent and of high quality.

Apart from investing further to improve brand image, we have also found that customers that are purchasing the beers through direct distribution, seem to have a higher likelihood of recommending B-Craft, and we know that once customers are happy with the service and recommending the product to their friends and family, we would expect to see an increase in order quantities. Therefore B-Craft should investigate further if there is anything that can be done to improve the experience of purchasing through the network distribution, or perhaps B-craft should look at increasing their direct distribution capacity.

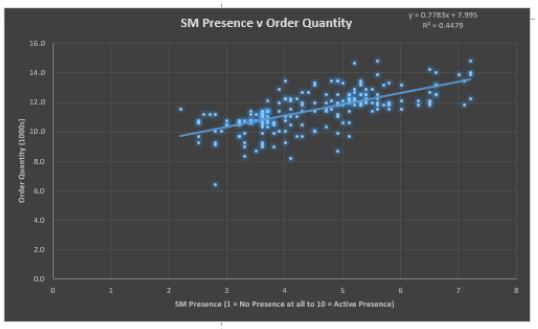
Finally, we can observe an upward trend in the pale ale production quantities, this also indicates a healthy increase in order quantities over time.

Do note that there are limitations within this experiment, one being that there are outliers present that needs to be further investigated, and the sample size of 200 is quite small, it would be beneficial to replicate the model on larger samples sets.

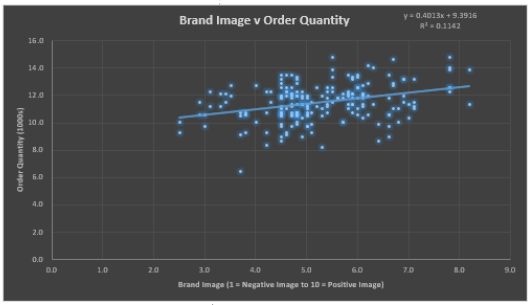
## Appendix 2.1 – Scatter Plot and Correlation Analysis:




Quality v Order Quantity
y = 0.4198x + 8.1832
R² = 0.1878
Order Quantity (1000s)
Quality (1 = Poor Quality to 10 = Excellent Quality)


SM Presence v Order Quantity
y = 0.7783x + 7.995
R² = 0.4479
Order Quantity (1000s)
SM Presence (1 = No Presence at all to 10 = Active Presence)


Advert v Order Quantity
y = 0.2769x + 10.373
R² = 0.0562
Order Quantity (1000s)
Advert (1 = Poor to 10 = Excellent)


Brand Image v Order Quantity
y = 0.4013x + 9.3916
R² = 0.1142
Order Quantity (1000s)
Brand Image (1 = Negative Image to 10 = Positive Image)


Comp Pricing v Order Quantity
y = -0.1845x + 12.783
R² = 0.0474
Order Quantity (1000s)
Comp Pricing (1 = Poor to 10 = Excellent)


Order Fulfillment v Order Quantity
y = 0.4622x + 9.5368
R² = 0.099
Order Quantity (1000s)
Order Fulfillment (1 = Poor to 10 = Excellent)


Flex Price v Order Quantity
y = -0.0032x + 11.512
R² = 8E-06
Order Quantity (1000s)
Flex price (1 = Not Flexible at all to 10 = Very Flexible)


Shipping Speed v Order Quantity
y = 0.76x + 8.5972
R² = 0.1807
Order Quantity (1000s)
Shipping speed (1 = Very Slow to 10 = Very Fast)


Shipping Cost v Order Quantity
y = 0.5688x + 8.4337
R² = 0.2544
Order Quantity (1000s)
Shipping Cost (1 = Not Affordable at All to 10 = Definitely Affordable)


Recommendation v Order Quantity
y = 1.3662x + 10.808
R² = 0.2612
Order Quantity (1000s)
Recommend
(0 = No, would not Recommend; 1 = Yes, would Definitely Recommend)

| | Loyalty | Cust_Type | Region | Dist_Channel | Quality | SM_Presence | Advert | Brand_Image | Comp_Pricing | Order_Fulfillment | Flex_Price | Shipping_Speed | Shipping_Cost | Recommend | Order_Qty |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Loyalty | 1.00 | | | | | | | | | | | | | | |
| Cust_Type | -0.05 | 1.00 | | | | | | | | | | | | | |
| Region | 0.07 | -0.03 | 1.00 | | | | | | | | | | | | |
| Dist_Channel | 0.22 | -0.14 | -0.26 | 1.00 | | | | | | | | | | | |
| Quality | 0.08 | -0.05 | -0.54 | 0.38 | 1.00 | | | | | | | | | | |
| SM_Presence | 0.31 | 0.07 | 0.02 | 0.41 | 0.24 | 1.00 | | | | | | | | | |
| Advert | 0.26 | 0.01 | 0.22 | 0.17 | -0.05 | 0.40 | 1.00 | | | | | | | | |
| Brand_Image | 0.26 | 0.01 | 0.38 | 0.29 | -0.12 | 0.67 | 0.63 | 1.00 | | | | | | | |
| Comp_Pricing | 0.08 | 0.15 | 0.59 | -0.33 | -0.45 | 0.00 | 0.10 | 0.20 | 1.00 | | | | | | |
| Order_Fulfillment | 0.14 | -0.01 | 0.00 | 0.26 | 0.08 | 0.36 | 0.23 | 0.28 | -0.06 | 1.00 | | | | | |
| Flex_Price | 0.06 | 0.04 | 0.58 | -0.22 | -0.49 | 0.15 | 0.26 | 0.27 | 0.47 | 0.42 | 1.00 | | | | |
| Shipping_Speed | 0.20 | 0.02 | 0.01 | 0.25 | 0.07 | 0.41 | 0.32 | 0.30 | -0.06 | 0.77 | 0.51 | 1.00 | | | |
| Shipping_Cost | 0.17 | 0.02 | 0.00 | 0.24 | 0.14 | 0.42 | 0.25 | 0.30 | -0.09 | 0.70 | 0.36 | 0.84 | 1.00 | | |
| Recommend | 0.17 | -0.01 | -0.10 | 0.43 | 0.36 | 0.46 | 0.19 | 0.31 | -0.13 | 0.32 | 0.06 | 0.40 | 0.38 | 1.00 | |
| Order_Qty | 0.41 | -0.05 | -0.12 | 0.40 | 0.43 | 0.67 | 0.24 | 0.34 | -0.22 | 0.31 | 0.00 | 0.43 | **0.50** | 0.51 | 1.00 |

## Appendix 2.2 – Initial Regression Model – Multicollinearity Addressed:

Model 1 - Initial Model - Multicollinearity Addressed

SUMMARY OUTPUT

| Regression Statistics | |
|---|---|
| Multiple R | 0.81 |
| R Square | 0.66 |
| Adjusted R Square | 0.64 |
| Standard Error | 0.81 |
| Observations | 200.00 |

ANOVA

| | df | SS | MS | F | Significance F |
|---|---|---|---|---|---|
| Regression | 12.00 | 235.83 | 19.65 | 30.27 | 0.00 |
| Residual | 187.00 | 121.42 | 0.65 | | |
| Total | 199.00 | 357.24 | | | |

| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% | Lower 95.0% | Upper 95.0% |
|---|---|---|---|---|---|---|---|---|
| Intercept | 7.24 | 0.69 | 10.46 | 0.00 | 5.88 | 8.61 | 5.88 | 8.61 |
| Loyalty | 0.07 | 0.01 | 4.58 | 0.00 | 0.04 | 0.10 | 0.04 | 0.10 |
| Cust_Type | -0.13 | 0.12 | -1.06 | 0.29 | -0.36 | 0.11 | -0.36 | 0.11 |
| Region | 0.26 | 0.17 | 1.49 | 0.14 | -0.08 | 0.59 | -0.08 | 0.59 |
| Dist_Channel | -0.04 | 0.15 | -0.27 | 0.79 | -0.33 | 0.25 | -0.33 | 0.25 |
| Quality | 0.18 | 0.06 | 3.24 | 0.00 | 0.07 | 0.29 | 0.07 | 0.29 |
| SM_Presence | 0.60 | 0.08 | 7.50 | 0.00 | 0.44 | 0.76 | 0.44 | 0.76 |
| Advert | -0.01 | 0.07 | -0.12 | 0.91 | -0.14 | 0.12 | -0.14 | 0.12 |
| Brand_Image | -0.16 | 0.09 | -1.69 | 0.09 | -0.34 | 0.03 | -0.34 | 0.03 |
| Comp_Pricing | -0.12 | 0.05 | -2.47 | 0.01 | -0.21 | -0.02 | -0.21 | -0.02 |
| Order_Fulfillment | -0.20 | 0.09 | -2.26 | 0.02 | -0.38 | -0.03 | -0.38 | -0.03 |
| Shipping_Cost | 0.34 | 0.07 | 4.78 | 0.00 | 0.20 | 0.48 | 0.20 | 0.48 |
| Recommend | 0.36 | 0.14 | 2.56 | 0.01 | 0.08 | 0.64 | 0.08 | 0.64 |

## Appendix 2.3 – Final Regression Model

Model 6 - Removed Brand Image (Final Model)

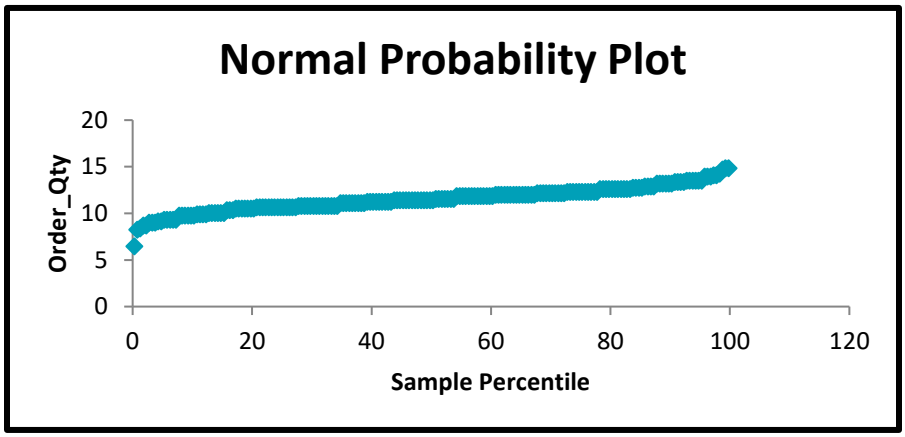
SUMMARY OUTPUT

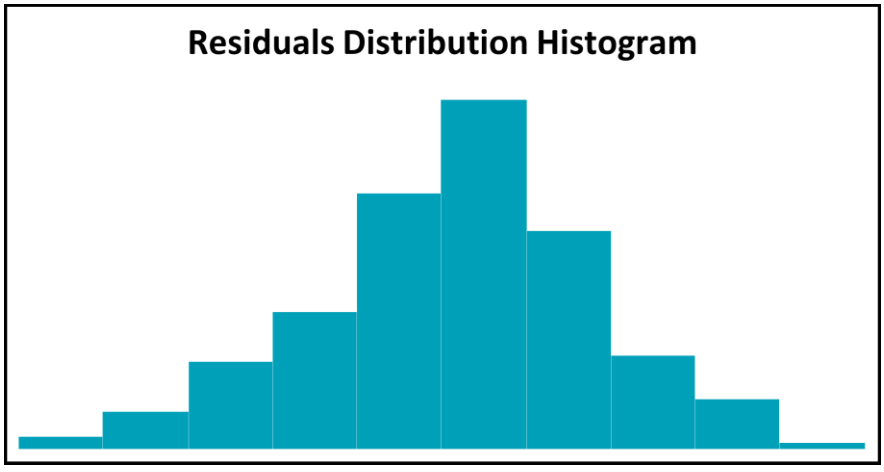
| *Regression Statistics* | |
|---|---|
| Multiple R | 0.81 |
| R Square | 0.65 |
| Adjusted R Square | 0.64 |
| Standard Error | 0.81 |
| Observations | 200.00 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 7.00 | 231.64 | 33.09 | 50.58 | 0.00 |
| Residual | 192.00 | 125.61 | 0.65 | | |
| Total | 199.00 | 357.24 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
|---|---|---|---|---|---|---|---|---|
| Intercept | 6.84 | 0.65 | 10.60 | 0.00 | 5.57 | 8.12 | 5.57 | 8.12 |
| Loyalty | 0.07 | 0.01 | 4.65 | 0.00 | 0.04 | 0.10 | 0.04 | 0.10 |
| Quality | 0.18 | 0.05 | 3.59 | 0.00 | 0.08 | 0.28 | 0.08 | 0.28 |
| SM_Presence | 0.50 | 0.06 | 8.04 | 0.00 | 0.37 | 0.62 | 0.37 | 0.62 |
| Comp_Pricing | -0.10 | 0.04 | -2.36 | 0.02 | -0.18 | -0.02 | -0.18 | -0.02 |
| Order_Fulfillment | -0.21 | 0.09 | -2.39 | 0.02 | -0.39 | -0.04 | -0.39 | -0.04 |
| Shipping_Cost | 0.34 | 0.07 | 4.88 | 0.00 | 0.20 | 0.48 | 0.20 | 0.48 |
| Recommend | 0.33 | 0.14 | 2.41 | 0.02 | 0.06 | 0.61 | 0.06 | 0.61 |



## Appendix 3 – Interaction Effect of Brand Image and Quality on Order Quantities

SUMMARY OUTPUT

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.60 |
| R Square | 0.35 |
| Adjusted R Square | 0.35 |
| Standard Error | 1.08 |
| Observations | 200.00 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 3.00 | 126.80 | 42.27 | 35.95 | 0.00 |
| Residual | 196.00 | 230.45 | 1.18 | | |
| Total | 199.00 | 357.24 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
|---|---|---|---|---|---|---|---|---|
| Intercept | 0.75 | 2.31 | 0.33 | 0.74 | -3.79 | 5.30 | -3.79 | 5.30 |
| Quality | 1.04 | 0.28 | 3.69 | 0.00 | 0.48 | 1.59 | 0.48 | 1.59 |
| Brand_Image | 1.30 | 0.40 | 3.21 | 0.00 | 0.50 | 2.09 | 0.50 | 2.09 |
| Brand Image * Quality | -0.10 | 0.05 | -2.08 | 0.04 | -0.20 | -0.01 | -0.20 | -0.01 |

## Appendix 4.1 – Predicting Likelihood of B-Craft Recommendations

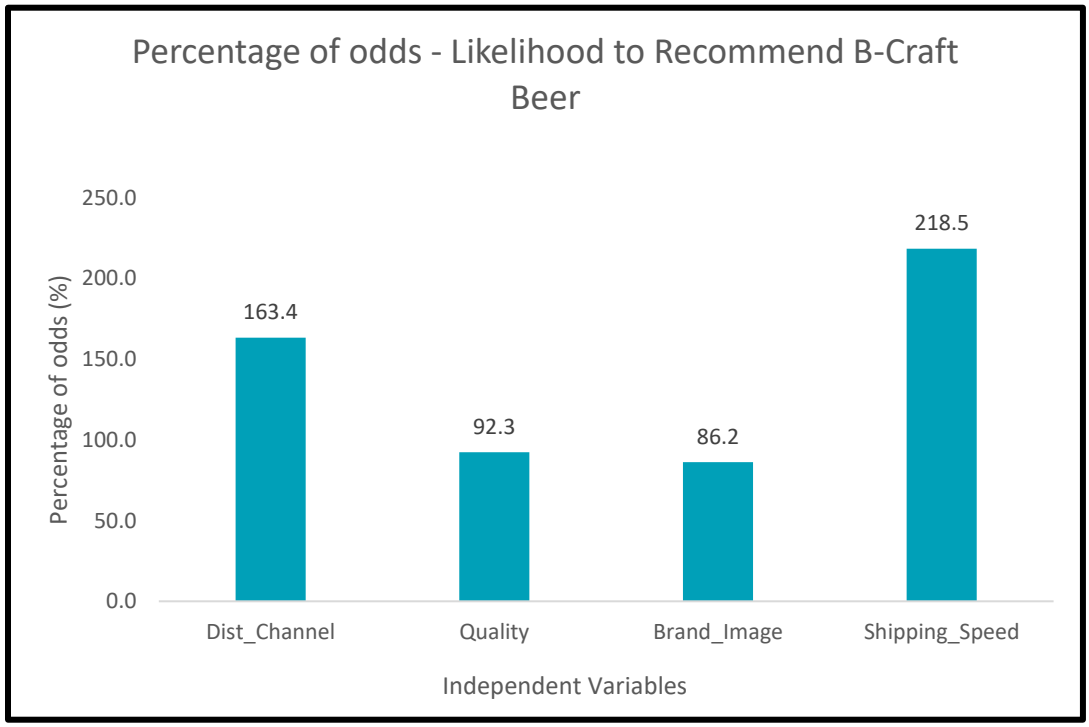
| **Classification Table** | | | |
|---|---|---|---|
| | Will Recommend (1) | Will Not Recommend (0) | ***Total*** |
| Suc-Pred | 78.00 | 25.00 | ***103.00*** |
| Fail-Pred | 23.00 | 74.00 | ***97.00*** |
| ***Total*** | ***101.00*** | ***99.00*** | ***200.00*** |
| Accuracy | 0.77 | 0.75 | 0.76 |
| Cutoff | 0.50 | | |

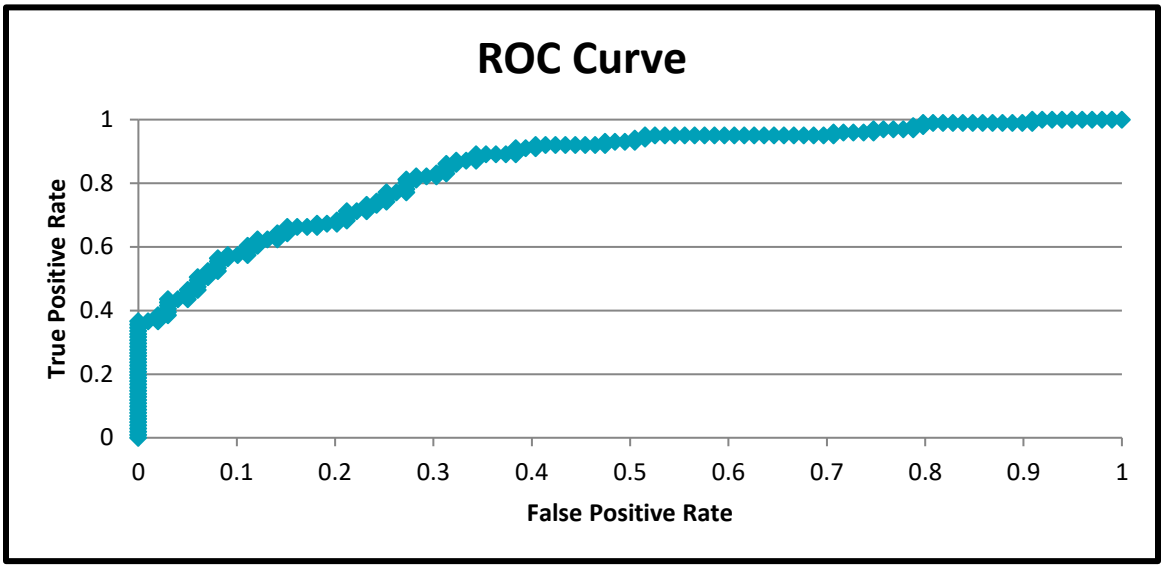
| | |
|---|---|
| Maximum Chance Criterion | **0.51** |
| PCC hit ratio = $p^2$ + $(1-p)^2$ | **0.50** |
| Standard (rule of thumb) | **0.625** |
| AUC | **0.85** |

| | |
|---|---|
| LL0 | -138.62 |
| LL1 | -95.38 |
| | |
| Chi-Sq | 86.48 |
| df | 4.00 |
| p-value | 0.00 |
| alpha | 0.05 |
| sig | yes |
| | |
| R-Sq (L) | 0.31 |
| R-Sq (CS) | 0.35 |
| R-Sq (N) | 0.47 |
| AIC | 200.75 |
| BIC | 217.25 |

| | coeff b | s.e. | Wald | p-value | exp(b) | lower | upper |
|---|---|---|---|---|---|---|---|
| Intercept | -13.28 | 2.28 | 33.90 | 0.00 | 0.00 | | |
| Dist_Channel | 0.97 | 0.38 | 6.60 | 0.01 | 2.63 | 1.26 | 5.51 |
| Quality | 0.65 | 0.15 | 18.10 | 0.00 | 1.92 | 1.42 | 2.60 |
| Brand_Image | 0.62 | 0.20 | 10.13 | 0.00 | 1.86 | 1.27 | 2.73 |
| Shipping_Speed | 1.16 | 0.29 | 16.20 | 0.00 | 3.19 | 1.81 | 5.60 |

| Percentage of odds | | | |
|---|---|---|---|
| | coeff b | exp(b) | % of odds |
| Intercept | -13.28 | 0.00 | |
| Dist_Channel | 0.97 | 2.63 | 163.4 |
| Quality | 0.65 | 1.92 | 92.3 |
| Brand_Image | 0.62 | 1.86 | 86.2 |
| Shipping_Speed | 1.16 | 3.19 | 218.5 |






## Appendix 4.2 – Predicting Probability of Likelihood of Recommendation

**Probability Matrix - Network Distribution - Netural Shipping - Various Quality & Brand Image**

| Dist_Channel | Quality | Shipping Speed | Brand Image - Negative | Predicted Probability | Brand Image - Neutral | Predicted Probability | Brand Image - Positive | Predicted Probability |
|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 5 | 1 | 0% | 5 | 2% | 10 | 35% |
| 0 | 2 | 5 | 1 | 0% | 5 | 4% | 10 | 51% |
| 0 | 3 | 5 | 1 | 1% | 5 | 8% | 10 | 67% |
| 0 | 4 | 5 | 1 | 1% | 5 | 15% | 10 | 79% |
| 0 | 5 | 5 | 1 | 3% | 5 | 25% | 10 | 88% |
| 0 | 6 | 5 | 1 | 5% | 5 | 39% | 10 | 93% |
| 0 | 7 | 5 | 1 | 9% | 5 | 55% | 10 | 96% |
| 0 | 8 | 5 | 1 | 16% | 5 | 70% | 10 | 98% |
| 0 | 9 | 5 | 1 | 27% | 5 | 82% | 10 | 99% |
| 0 | 10 | 5 | 1 | 42% | 5 | 90% | 10 | 99% |

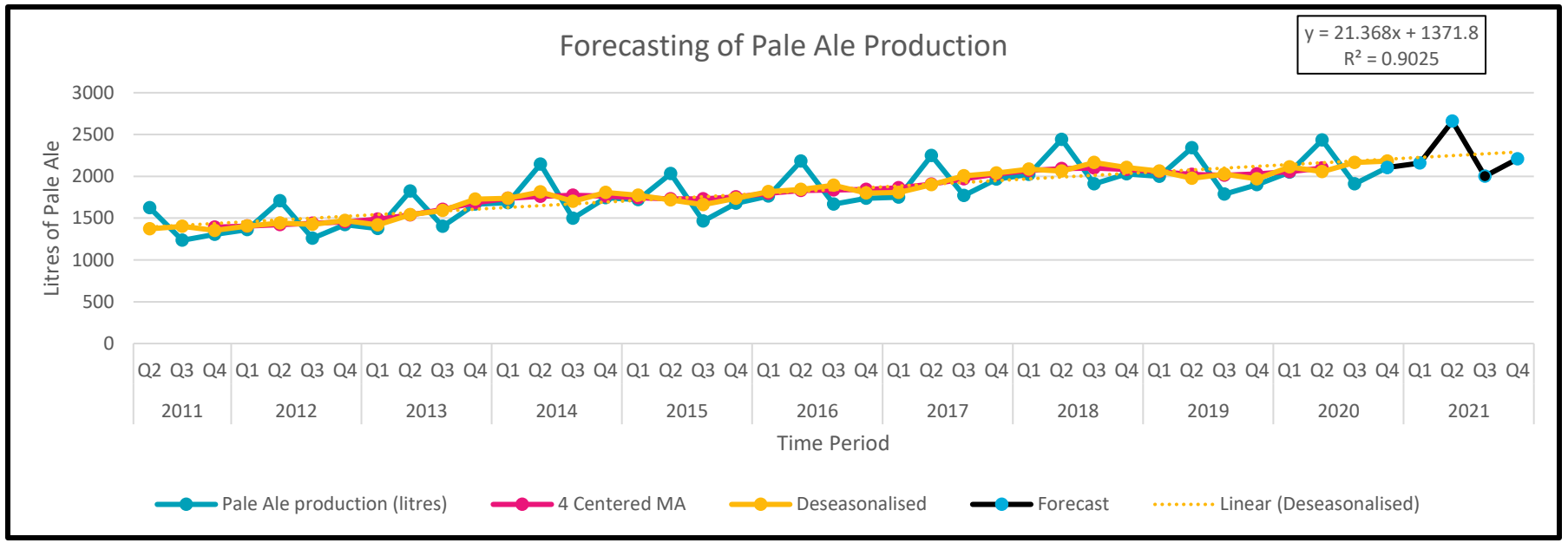
**Probability Matrix - Direct Distribution - Netural Shipping - Various Quality & Brand Image**

| Dist_Channel | Quality | Shipping Speed | Brand Image - Negative | Predicted Probability | Brand Image - Neutral | Predicted Probability | Brand Image - Positive | Predicted Probability |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 5 | 1 | 1% | 5 | 6% | 10 | 59% |
| 1 | 2 | 5 | 1 | 1% | 5 | 11% | 10 | 73% |
| 1 | 3 | 5 | 1 | 2% | 5 | 19% | 10 | 84% |
| 1 | 4 | 5 | 1 | 4% | 5 | 31% | 10 | 91% |
| 1 | 5 | 5 | 1 | 7% | 5 | 47% | 10 | 95% |
| 1 | 6 | 5 | 1 | 12% | 5 | 63% | 10 | 97% |
| 1 | 7 | 5 | 1 | 21% | 5 | 76% | 10 | 99% |
| 1 | 8 | 5 | 1 | 34% | 5 | 86% | 10 | 99% |
| 1 | 9 | 5 | 1 | 50% | 5 | 92% | 10 | 100% |
| 1 | 10 | 5 | 1 | 66% | 5 | 96% | 10 | 100% |

# Appendix 5 – Timeseries Model to Forecast Pale Ale Production

$$Y_t = T_t \times S_t \times I_t$$

| | | $Y_t$ | $t$ | | Baseline | $S_t, I_t$ | $S_t$ | $Y_t / S_t$ | $T_t$ | $S_t \times T_t$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Quarter | Pale Ale production (litres) | Time Period | 4 MA | 4 Centered MA | Ratio (Obs/MA) | Index | Deseasonalised | Trend | Forecast | APE |
| 2011 | Q2 | 1624.5 | 1 | | | | 1.18 | 1372.83 | 1393.17 | | 0.01 |
| | Q3 | 1235.08 | 2 | | | | 0.88 | 1400.50 | 1414.54 | | 0.01 |
| | Q4 | 1305.63 | 3 | 1382.10 | 1392.39 | 0.94 | 0.96 | 1353.43 | 1435.90 | | 0.06 |
| 2012 | Q1 | 1363.21 | 4 | 1402.68 | 1405.48 | 0.97 | 0.97 | 1405.21 | 1457.27 | | 0.04 |
| | Q2 | 1706.79 | 5 | 1408.28 | 1422.65 | 1.20 | 1.18 | 1442.38 | 1478.64 | | 0.03 |
| | Q3 | 1257.50 | 6 | 1437.01 | 1438.70 | 0.87 | 0.88 | 1425.92 | 1500.01 | | 0.05 |
| | Q4 | 1420.54 | 7 | 1440.40 | 1454.85 | 0.98 | 0.96 | 1472.56 | 1521.38 | | 0.03 |
| 2013 | Q1 | 1376.75 | 8 | 1469.31 | 1487.46 | 0.93 | 0.97 | 1419.17 | 1542.74 | | 0.09 |
| | Q2 | 1822.46 | 9 | 1505.61 | 1536.31 | 1.19 | 1.18 | 1540.12 | 1564.11 | | 0.02 |
| | Q3 | 1402.71 | 10 | 1567.00 | 1605.59 | 0.87 | 0.88 | 1590.57 | 1585.48 | | 0.00 |
| | Q4 | 1666.08 | 11 | 1644.18 | 1684.61 | 0.99 | 0.96 | 1727.09 | 1606.85 | | 0.07 |
| 2014 | Q1 | 1685.46 | 12 | 1725.05 | 1737.13 | 0.97 | 0.97 | 1737.39 | 1628.22 | | 0.06 |
| | Q2 | 2145.96 | 13 | 1749.21 | 1758.82 | 1.22 | 1.18 | 1813.51 | 1649.58 | | 0.09 |
| | Q3 | 1499.33 | 14 | 1768.44 | 1772.78 | 0.85 | 0.88 | 1700.14 | 1670.95 | | 0.02 |
| | Q4 | 1743.00 | 15 | 1777.13 | 1763.08 | 0.99 | 0.96 | 1806.82 | 1692.32 | | 0.06 |
| 2015 | Q1 | 1720.21 | 16 | 1749.04 | 1744.58 | 0.99 | 0.97 | 1773.21 | 1713.69 | | 0.03 |
| | Q2 | 2033.63 | 17 | 1740.13 | 1731.73 | 1.17 | 1.18 | 1718.58 | 1735.06 | | 0.01 |
| | Q3 | 1463.67 | 18 | 1723.34 | 1728.72 | 0.85 | 0.88 | 1659.69 | 1756.42 | | 0.06 |
| | Q4 | 1675.88 | 19 | 1734.10 | 1752.66 | 0.96 | 0.96 | 1737.24 | 1777.79 | | 0.02 |
| 2016 | Q1 | 1763.25 | 20 | 1771.22 | 1796.83 | 0.98 | 0.97 | 1817.57 | 1799.16 | | 0.01 |
| | Q2 | 2182.08 | 21 | 1822.44 | 1829.91 | 1.19 | 1.18 | 1844.03 | 1820.53 | | 0.01 |
| | Q3 | 1668.54 | 22 | 1837.38 | 1835.92 | 0.91 | 0.88 | 1892.01 | 1841.90 | | 0.03 |
| | Q4 | 1735.63 | 23 | 1834.47 | 1842.70 | 0.94 | 0.96 | 1799.18 | 1863.26 | | 0.04 |
| 2017 | Q1 | 1751.63 | 24 | 1850.94 | 1863.63 | 0.94 | 0.97 | 1805.59 | 1884.63 | | 0.04 |
| | Q2 | 2247.96 | 25 | 1876.32 | 1905.18 | 1.18 | 1.18 | 1899.70 | 1906.00 | | 0.00 |
| | Q3 | 1770.08 | 26 | 1934.03 | 1967.98 | 0.90 | 0.88 | 2007.15 | 1927.37 | | 0.04 |
| | Q4 | 1966.46 | 27 | 2001.94 | 2026.23 | 0.97 | 0.96 | 2038.46 | 1948.74 | | 0.04 |
| 2018 | Q1 | 2023.25 | 28 | 2050.53 | 2068.08 | 0.98 | 0.97 | 2085.58 | 1970.10 | | 0.06 |
| | Q2 | 2442.33 | 29 | 2085.64 | 2093.66 | 1.17 | 1.18 | 2063.97 | 1991.47 | | 0.04 |
| | Q3 | 1910.50 | 30 | 2101.68 | 2098.69 | 0.91 | 0.88 | 2166.37 | 2012.84 | | 0.07 |
| | Q4 | 2030.63 | 31 | 2095.70 | 2082.86 | 0.97 | 0.96 | 2104.98 | 2034.21 | | 0.03 |
| 2019 | Q1 | 1999.33 | 32 | 2070.03 | 2054.39 | 0.97 | 0.97 | 2060.93 | 2055.58 | | 0.00 |
| | Q2 | 2339.67 | 33 | 2038.75 | 2022.36 | 1.16 | 1.18 | 1977.21 | 2076.94 | | 0.05 |
| | Q3 | 1785.38 | 34 | 2005.97 | 2012.01 | 0.89 | 0.88 | 2024.49 | 2098.31 | | 0.04 |
| | Q4 | 1899.50 | 35 | 2018.04 | 2029.87 | 0.94 | 0.96 | 1969.05 | 2119.68 | | 0.08 |
| 2020 | Q1 | 2047.63 | 36 | 2041.70 | 2057.26 | 1.00 | 0.97 | 2110.71 | 2141.05 | | 0.01 |
| | Q2 | 2434.29 | 37 | 2072.82 | 2098.54 | 1.16 | 1.18 | 2057.17 | 2162.42 | | 0.05 |
| | Q3 | 1909.88 | 38 | 2124.25 | | | 0.88 | 2165.67 | 2183.78 | | 0.01 |
| | Q4 | 2105.20 | 39 | | | | 0.96 | 2182.28 | 2205.15 | 2105.20 | 0.01 |
| 2021 | Q1 | | 40 | | | | 0.97 | | 2226.52 | 2159.97 | |
| | Q2 | | 41 | | | | 1.18 | | 2247.89 | 2659.97 | |
| | Q3 | | 42 | | | | 0.88 | | 2269.26 | 2001.23 | |
| | Q4 | | 43 | | | | 0.96 | | 2290.62 | 2209.71 | |


Forecasting of Pale Ale Production
y = 21.368x + 1371.8
R² = 0.9025
Litres of Pale Ale
3000
2500
2000
1500
1000
500
0
Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
Time Period
Pale Ale production (litres)
4 Centered MA
Deseasonalised
Forecast
Linear (Deseasonalised)