

**Analysis of Mad Dog Craft Beer
Assignment-2**

MIS771 – Descriptive Analytics and Visualisation

**Shantanu Gupta
Student Id: 218200234**

Introduction:

Mad Dog Craft Beer, is an Australian micro-brewery company that sells pale ale beer in Melbourne and regional Victoria. It is sold through two market segments: a) pubs, bars and restaurants (or indirectly through a sales representative) and b) bottle shops (directly to the buyer). Despite its limited operations, the company has experienced fast growth in its production and a couple of years. Management wanted to increase its brewing capacity to 3 million litres per year of its pale ale beer, maintain a strong relationship with its potential customer and identify the critical factors that influence the quantity of beer ordered as Australians are increasingly conscious of the beer products. In addition, the internal research of the company has forecasted a shift in business climate within the next four to five years. So management, are interested in planning to put a formal procedure in place to forecast their Mad dog craft beer's Pale ale production in coming next four years. This would help them to accurately project future supply and demand and adjust production needs accordingly.

This report is especially useful for the Brewery and Wine industry, or those companies who wanted to step into Beer business in Australia. With 84% of all beer sold in Australia being made in Australia, this vibrant beer sector is a major driver of economic activity and domestic jobs, supporting vital cogs in the industry from Australian farmers upstream to brewing, packaging, distribution and freight, all the way through to retail, tourism and hospitality. By reading this report, the companies can improve their operations, marketing strategy, estimate order quantity, predict whether one customer recommends beer products to others or not etc., in this representing leading beer makers and beer lovers across Australia. From these insights of one particular company, the other brewery companies can follow similar strategies, so that they will be in an excellent position to adjust and develop plans for their own companies.

To continue successful with their operations and get solid financial turnovers in coming years, Mad Dog Craft Beer Company has approached BEAUTIFUL-DATA (a market research company) and asked to conduct a survey of their clients to better understand the characteristics of Mad Dog Craft Beer's customers, and their repurchase intention.

As a graduate intern, this report evaluates the data analysis process and displays the result of the request raised by Todd Nash. As an intern, I have

tried gather all the answers to the concerns put forward by Todd Nash and presented it with statistical output.

Q1.

1.1

- The mean number of beer bottles ordered by the customers is 7670.
- The median number of bottles ordered for the 200-sample data set is 7600 where, as I can say that the most common order quantity in a number of bottles ordered by the customer is 7200.
- The minimum number of bottles ordered by the customer is 4300, while the maximum value number of bottles ordered by the customer is 9900.
- Furthermore, the standard deviation value for a number of bottles ordered is 893, which says that there is 893 of average spread away from the mean bottles ordered value of 7670.
- The smaller the standard error, the more representative the sample will be of the overall population. The standard error of the number of bottles ordered by customer stands at 60 bottles. Deviation of 60 bottles states that the sample means deviates by 60 from the actual mean of the population.
- Order quantity is negatively skewed. As $\text{mean} < \text{median} < \text{mode}$.
- Talking about the Kurtosis value which is 0.58 indicating that the data or the distribution has positive kurtosis, stating that distribution has thinner tails than the normal distribution. Centre of the graph is a bit flatter.
- Are there any potential outliers? Yes, we can say that there are outliers by looking at the descriptive summary, as the maximum value of customer satisfaction is greater than the upper fence value. There was some customer whose order a large number of bottles than the rest of participants in the survey and thus, considered as an unusual case.
- The total number of beer bottles ordered by the customer is 1533.
- The difference between the largest and smallest number of bottles ordered by the customer is 5600.
- The lower 25% of the number of quantities ordered ranges between 4300 to 7100. The middle 50% of our data is between 7100 to 8200 bottles. The upper 25% of the quantities ordered value is 8200 or more.

1.2

Recommend is a non-numeric or categorical or qualitative variable. We encoded into a binary output which includes 0 and 1. 0 means that the customers will not recommend Mad Dog Craft beer to others while 1 means they will recommend to others. There is 101 customers that recommend Male dog beer to others and 99 people are those who don't recommend a beer to others. So it is pretty balanced dataset. There is no practical significance of mean, standard deviation and median in categorical data. Since the mod is the most common value, we can also confirm our result by seeing as the mode is equal to 1. The data is not skewed in favour of one of the class.

Q2.

2.1

First, we have analysed the relationship of all 14 independent variables with Order Quantity using scatter plot and correlation matrix.

Loyalty vs. Order Quantity

As Length of time of a particular customer has been buying from Mad Dog Craft Beer increases the order quantity also increased. Strength of this relationship between variables in the scatter plot is moderate, positive and linear. We can also see a correlation value of Loyalty with Order quantity is 0.41. We should include in the final model.

Customer type vs. Order Quantity

The relationship between them is very weak since the correlation value is -0.05. The Customer type is a weak attribute and doesn't have any influence over the Order Quantity attribute. Customer type is categorical variable, so fitting regression is not useful.

Region vs. Order Quantity

The relationship between them is very weak since the correlation value is -0.12. This is a weak attribute and doesn't have any influence over the Order Quantity attribute. As the region is a categorical variable, so fitting regression is not useful.

Distribution Channel vs. Order Quantity

The relationship between them is very weak since the correlation value is -0.40. This is a strong attribute and does have influence over the Order

Quantity attribute. As the distribution channel is a categorical variable, so fitting regression is not useful. We should include in the final model.

Quality vs. Order Quantity

As we increase the quality of mad dog craft beer, order quantity also increases. The strength of the relationship is moderate, positive and linear. We can also see the correlation value of Quality with Order quantity is 0.43. We should include in the final model.

Social Media Presence vs. Order Quantity

As we increase the social media presence, there is also an increase in order quantity. The Social Media Presence has the correlation value as 0.24 which says that the relationship is moderate depicting that the Social Media Presence has less influence over Order Quantity. The relationship is linear and positive.

Advert vs. Order Quantity

The correlation value is -0.23. The relationship is weak, positive and linear.

Brand Image vs. Order Quantity

As we can see most of the data points are clustered around the centre. The relationship is pretty good. The relationship between the Brand Image and Order Quantity is moderate because the correlation value is 0.34 and it has a moderate influence on the Order Quantity. The correlation value is -0.33. We should include in the final model.

Competitive Pricing vs. Order Quantity

Competitive pricing consists of setting the price at the same level as one's competitors. As we increase Competitive pricing have a negative relationship with Order quantity. Customers hate those companies who make monopoly and decide competitive prices to increase their profits. This method relies on the idea that competitors have already thoroughly worked on their pricing. The correlation value is -0.21.

Order Fulfilment vs. Order Quantity

The correlation value is -0.31. The relationship is weak, positive and linear.

Price Flexibility vs. Order Quantity

The perceived willingness of Mad Dog Craft Beer to negotiate product prices have a weak, negative relationship with Order quantity. As we make

prices more flexible customer will order less beer from them. If the company makes regular changes in the prices in beer bottle it causes a lot of difficulties for the customers. For example, if some customer came today and come after one week, he/she finds that there is a sudden increase in the price. So it makes them angry. The correlation value is -0.002. We should not include the final model.

Shipping Speed vs. Order Quantity

Shipping speed and Quantity Ordered should have Strong relationship, positive and linear. The similar trend is followed by cost. The correlation value is -0.42. We should include in the final model.

Shipping Cost vs. Order Quantity

Shipping cost and Quantity Ordered should have Strong relationship, positive and linear. The correlation value is -0.50. We should include in the final model.

Recommend vs. Order Quantity

The correlation value is -0.51. As recommend is a categorical variable, so fitting regression is not useful. We should include in the final model.

There is no variable in our model that shows a non-linear relationship i.e. quadratic, polynomial, with Order Quantity. So we did not need to do any variable transformation at all.

Now we need to check in our dataset whether there is any multi-collinearity or not? So we are considering threshold greater than 0.8 as two independent variables are highly correlated to each other. As we can see from the Table, Shipping Cost and Shipping Speed have a potential multi-collinearity problem. When we have multicollinearity we need to keep one and drop others. From a statistical point of view, it is advisable to remove Shipping Speed as it has a weaker correlation with Order Quantity. As general intuition says any item purchased is somewhat associated with shipping cost. If we completely remove shipping cost than the company would be bankrupt as prices depend on customer location. Shipping speed is not an important variable as some customers don't want fast delivery. So, this beer is not an essential item so they can wait for a few hours or 1-2 days. One more thing expedited shipping increases shipping cost, as we know customer hate to pay more. Henceforth, we consider shipping speed is an important factor for managerial decision. We will remove Shipping Speed from the model.

2.2

After analyzing the data, we came up with 13 variables that can influence the Order Quantity are Loyalty, Customer Type, Region, Distribution Channel, Quality, Social Media Presence, Advert, Brand Image, Competitive Pricing, Order Fulfillment, Price Flexibility, Shipping Cost, and Recommend.

We need to go for multiple iterations for estimating the quantity ordered using the above attributes as we have all these variables have more than $p\text{-value} > 0.05$. We started the greatest value at the time and one at the lowest.

- First Iteration: The first iteration was done, and we eliminated the Region value since its p-value (0.97) was greater than 0.05
- Second Iteration: We eliminated the customer type since its p-value is $0.91 > 0.05$
- Third iteration: We eliminated the Distribution Channel since its p-value is $0.620 > 0.05$
- Fourth Iteration: We eliminated the Advert since the p-value is $0.492 > 0.05$
- Fifth Iteration: We eliminated the Flex price since the p-value is $0.531 > 0.05$
- Sixth Iteration: We eliminated the Social media presence since the p-value is $0.251 > 0.05$
- Seventh Iteration: We eliminated the Order fulfilment since the p-value is $0.08 > 0.05$
- Eight Iteration: We eliminated the Competitive pricing since the p-value is $0.08 > 0.05$

The final regression model has the following potential independent variables: Loyalty, Quality, Brand Image, Shipping Cost, and Recommend.

Regression Equation:

Order Quantity = $3.53 + (0.05 \times \text{Loyalty}) + (0.21 \times \text{Quality}) + (0.13 \times \text{Brand Image}) + (0.22 \times \text{Shipping Cost}) + (0.34 \times \text{Recommend})$

Coefficient for Loyalty (0.05) tells us that, assuming with all other variables held constant, for a unit increase in Loyalty, Order Quantity will increase by 0.05(000), on average.

The coefficient for Quality (0.21) tells us that, assuming with all other variables held constant, for a unit increase in Quality, Order Quantity will increase by 0.21(000), on average.

The coefficient for Brand Image (0.13) tells us that, assuming with all other variables held constant, for a unit increase in Brand Image, Order Quantity will increase by 0.13(000), on average.

The coefficient for Shipping Cost (0.22) tells us that, assuming with all other variables held constant, for a unit increase in Shipping Cost, Order Quantity will increase by 0.22(000), on average.

The coefficient for Recommend (0.34) tells us that, assuming with all other variables held constant, for a unit increase in Recommend, Order Quantity will increase by 0.34(000), on average.

With all independent factors in the regression model be constant, the Order Quantity will be 3.53(000) bottles.

$R^2 = 53\%$ variation in our model is explained by our model. Approximately 54% of the variation in Productivity can be explained by the regression model (i.e., variables included in the regression model). Approximately 47% of the variation in productivity would be explained by other factors not included in the model. The value of R^2 is not high, thus this is not a strong predictive model. 61% of Standard error we can commit when using this regression model. The average error we commit by using this model.

We are 95% confident that an extra unit of increase in loyalty, on average, will increase order quantity by 30 to 80 bottles (assuming no change in the other variables).

We are 95% confident that an extra unit of increase in quality, on average, will increase order quantity by 140 to 280 bottles (assuming no change in the other variables).

We are 95% confident that an extra unit of increase in brand image, on average, will increase order quantity by 40 to 220 bottles (assuming no change in the other variables).

We are 95% confident that an extra unit of increase in shipping cost, on average, will increase order quantity by 140 to 300 bottles (assuming no change in the other variables).

There are 10 potential outliers in our data. We are meeting all 4 assumptions. No apparent problems in residual diagrams. All the model assumptions are ok.

2.3

Overall the model has some predictive power. All individual independent variables (including the interaction term) are individually significant as every independent variable p-values are $0.000 < .05$. We can see that our model is statistically significant.

Significance of the interaction term indicates that Brand Image interacts with quality in predicting Order Quantity. In other words, Brand Image moderates the relationship between quality and quantity.

As the graph is parallel there is no interaction effect. Low quality and high quality does not have any impact on order quantity. Low brand image and high brand image were performed about equally well in high quality but high brand image was performed considerably better than low brand image in the low quality. When the quality of the beer is low, the customer will order less number of bottles irrespective of their Brand image. When the quality of beer increases, those customers who believe in the better brand image of the beer more often, those customers will have more beer ordered compared to others. In other words, the effect of quality on quantity changes (increases) as the brand image becomes more improved (increases).

Regression Equation:

$$\text{Order Quantity} = 1.73 + (0.06 * \text{Loyalty}) + (0.43 * \text{Quality}) + (0.46 * \text{Brand Image}) + (-0.04 * \text{Brand Image} * \text{Quality}) + (0.21 * \text{Shipping Cost}) + (0.33 * \text{Recommend})$$

The Coefficient for Loyalty (0.06) tells us that, assuming with all other variables held constant, for a unit increase in Loyalty, Order Quantity will increase by 0.06(000), on average.

The coefficient for Quality (0.21) tells us that, assuming with all other variables held constant, for a unit increase in Quality, Order Quantity will increase by 0.21(000), on average.

The coefficient for Brand Image (0.43) tells us that, assuming with all other variables held constant, for a unit increase in Brand Image, Order Quantity will increase by 0.43(000), on average.

The coefficient for Interaction Term (0.04) tells us that, assuming with all other variables held constant, for a unit increase in Interaction Term, Order Quantity will decrease by 0.04(000), on average.

The coefficient for Shipping Cost (0.21) tells us that, assuming with all other variables held constant, for a unit increase in Shipping Cost, Order Quantity will increase by 0.21(000), on average.

The coefficient for Recommend (0.33) tells us that, assuming with all other variables held constant, for a unit increase in Recommend, Order Quantity will increase by 0.33(000), on average.

With all independent factors in the regression model be constant, the Order Quantity will be 1.73(000) bottles. Does not have a practical interpretation as the company does never have zero shipping cost, or have zero quality.

In our case R^2 (read as '*r-squared*') is 53%, which means that 53% of the variation in the order quantity can be explained by the variation in our regression model (i.e., variables included in the regression model). The remaining $1 - R^2 = 47\%$ of the variation in Order Quantity would be explained by other factors (e.g. shipping cost etc.) not included in the model. The value of R^2 is not high, thus this is not a strong predictive model. 61% of Standard error we can commit when using this regression model. The average error we commit by using this model.

We are 95% confident that an extra unit of increase in quality, on average, will increase order quantity by 111 to 757 bottles (assuming no change in the other variables).

Q3.

3.1

Compared to the baseline model, the final logistic model significantly reduced LL value, providing evidence for the statistical significance of the overall model. As we can see when we include independent variables have a p-value less than ($<$) 0.05, it implies our model is statistically significant.

Feeling toward Distribution Channel, Quality, Brand Image, and Shipping Speed have a positive relationship with the probability of recommending mad dog beers to others.

The likelihood (odds) of recommending a beer to others for a direct customer is 163.36 per cent greater than that of a through as sales-representative customer.

One unit of increase in positive feeling towards quality increases the likelihood (odds) of recommending beer to others by 92.34 per cent

One unit of increase in Brand Image increases the likelihood (odds) of recommending beer to others by 86.15 per cent

One unit of increase in Shipping Speed increases the likelihood (odds) of recommending beer to others by 218.52 per cent

The overall classification accuracy (hit ratio) was 76 %, i.e. 76 per cent of survey participants were accurately classified by the logistic regression model. Out of 200 customers, 103 were classified as successfully while 97 were misclassified. The remaining miss-classification rate could have been captured accurately if more relevant Independent variables had been included in the model. 76% of customer will recommend Mad Dog Craft beer to others

Of the 101 customers who recommend Mad dog craft beer to others (observed), 78 were accurately classified (predicted) as "recommend mad dog craft beer" and 23 were inaccurately classified (misclassified) as those who "did not recommend beer" (77% classification accuracy).

Of the 99 customers who did not recommend Mad dog craft beer to others (observed), 74 were accurately classified (predicted) as "did not recommend" and 25 inaccurately classified as those who "recommend mad dog beer" (75 % classification accuracy).

According to our dataset, the success category is recommending beers to others while the failure category does not recommend to others. In developing above predictive model I have used cut-off value is 50%. We generally used 50% when we have an equal number of observations for the success and failure category. We have a balanced dataset of recommending beers to others or not. So, we can use 50% as a criterion. When we have an unbalanced dataset and in order to check whether our model is significant or not using the statistical method we go for calculating other parameters like Maximum Chance Criteria (MCC), Proportional Chance Criteria (PCC) and Standard (Rule of Thumb). So, I have calculated all these parameters and compared with the hit ratio.

The Maximum Choice Criterion (MCC) is 51%, Proportional chance criterion (PCC) is 50% and Standard (Rule of Thumb) is 63%. While our accuracy rate (hit rate) is 76% is greater than MCC, PCC & Standard. Our logistic regression model is significantly better than a random process (chance) in classifying observations. Therefore, providing evidence for the statistical significance of the overall model.

According to Hosmer and Lemeshow R² (read as '*r-squared*'), 31 per cent of the variation in the dependent variable can be explained by the regression model. According to Cox and Snell R², 35 per cent of the variation in the dependent variable can be explained by the regression model. According to Nagelkerke R², 47 per cent of the variation in the dependent variable can be explained by the regression model. Coupled with the practical significance of the model, along with the overall model fit, these R² values are deemed as acceptable.

We do not need to check for outliers as there is no high standard error or small Wald statistics. The curve is quite distant from diagonal (.50), indicating the model's ability to discriminate between success (recommend beer) and failure (not recommend) is not due to chance. In other words, it is able to accurately classify my categories. We can also check from the Receiver operating characteristic (ROC) curve that passes through the upper left corner. The area under the curve (AUC) is very close to 1.0, pointing to a model that fits data well. Therefore the closer the ROC curve is to the upper left corner (More Area), the higher is the overall accuracy of the model.

3.2 & 3.3

Customer Segments

- 0 represents customer who made purchases directly
- 1 represents customer who made a purchase through a sales representative

Overall, the results indicated that a change in perception of quality and sense of feeling towards brand image along with two customer segments are key predictors of recommending beer to others. When the customer directly purchases the beer, the maximum predicted probability of recommending beer is 99.4%. If the customer purchases the beer through the sales representative, the maximum predicted probability is 99.8%. So beer product is highly recommended only when we have a positive brand image and very high quality, considering the speed of delivery is neutral.

It is very hard to maintain. However, it seems like if we convert those customers who have a neutral brand image and Quality score between 3 to 6 in both customer segments then we can increase our profit and production. We can increase by Launching Events That Increase Beer Sales like an Oktoberfest beer Beer Holidays, Trivia night etc. The company has limited resources, so we can't target negative brand image customer and that have low quality scores (1 to 3). Future debates should be designed to address the needs of these negative groups in order to recommend a beer to others.

Q4.

We need to forecast the Pale ale production beer for the next four quarters. As our data consist of the time series, I have used the Seasonal Forecasting method. The reason for selecting this forecasting because when we observe the data within the year which contains quarterly information. Seasonal demand has a pattern that repeats. Demand for beer has a seasonal pattern that repeats every 12 months. When I analyse the pattern I found that their seasonal patterns/demand changes quarterly. They have more sale in summer (Q2) than any other quarters. In the autumn (Q3) and spring season (Q1) season they have a very low sale in all years. In the winter season (Q4) they have a moderate sale. Time series plot exhibits the short term regular wave-like pattern. In order to get an accurate estimate of seasonal demand we need to do some kind of average for the demand in the first period of the cycle, and the second period, etc., so that the every month value is the same proportion with any other month value in each year. So, we have smoothed out the seasonality so that we can identify the linear trend. The aim of smoothing is to remove seasonality. I have used the 4-Moving average to smooth out our data because our length of pale ale production period is quarter. I have used the multiplicative model because as the data increase, so does the seasonal pattern. Seasonality seems to increase with time. In this model, the trend and seasonal components are multiplied and then added to the error/random component.

Pale Production = Seasonal effect * Trend * Residual (Irregularities)

Multiplicative time series model → which included 4 steps which are the

- Calculating seasonal indices
- Deseasonalize data

- Non-seasonal forecast
- Re-seasonalize the forecast

We calculate the ratio to moving average so we can quantify seasonality. In order to remove the seasonality component to see the underlying trend component, we need to calculate seasonal indices. The sum of the Seasonal indices should be equal to 4 as we are handling quarter data. However, because the presence of errors or irregularities in the way seasonal indices are calculated the sum will be slightly off. In order to removed irregularities, I need to normalize/adjust seasonal indices to make the sum equal to 4.

We followed the above-mentioned steps to derive the forecasting of four quarter turnover and in the end, we go the final output as: -

Production forecasted in quarter 2019-Q2 is 1728.06litres

Production forecasted in quarter 2019-Q3 is 2128.08 litres

Production forecasted in quarter 2019-Q4 is 1601.06 litres

Production forecasted in quarter 2020-Q1 is 1767.85 litres

One of the key questions in the forecasting process has to do with the measuring of the forecast accuracy. Since Mean Absolute Percentage Error (MAPE) is a measure of error, high numbers are bad and low numbers are good. So we are off around 4% by a quarter in our past amount of Pale ale beer production. The forecasted production value accuracy also differs in this range of 3-4% if similar trends follow in the coming years.

Recommendation:

- Launch Events That Increase Beer Sales like an Oktoberfest beer Beer Holidays, Trivia night etc.
- Market and Social Media to Beer Lovers. Smart Promotions to Increase Beer Sales. Connect with Food/Drink Bloggers in your region and promote the events through them.
- Use email marketing to communicate new products, specials and even wish a retailer a happy anniversary. Retailers won't buy more frequently unless they know there is a buying opportunity out there Wholesalers and Distributors should be part of a total Sales strategy.

- Do some innovation like Innovate in labelling process or made beer by Appealing to All Customers like Domestic Diehard, casual sipper, or craft guru.

Conclusion:

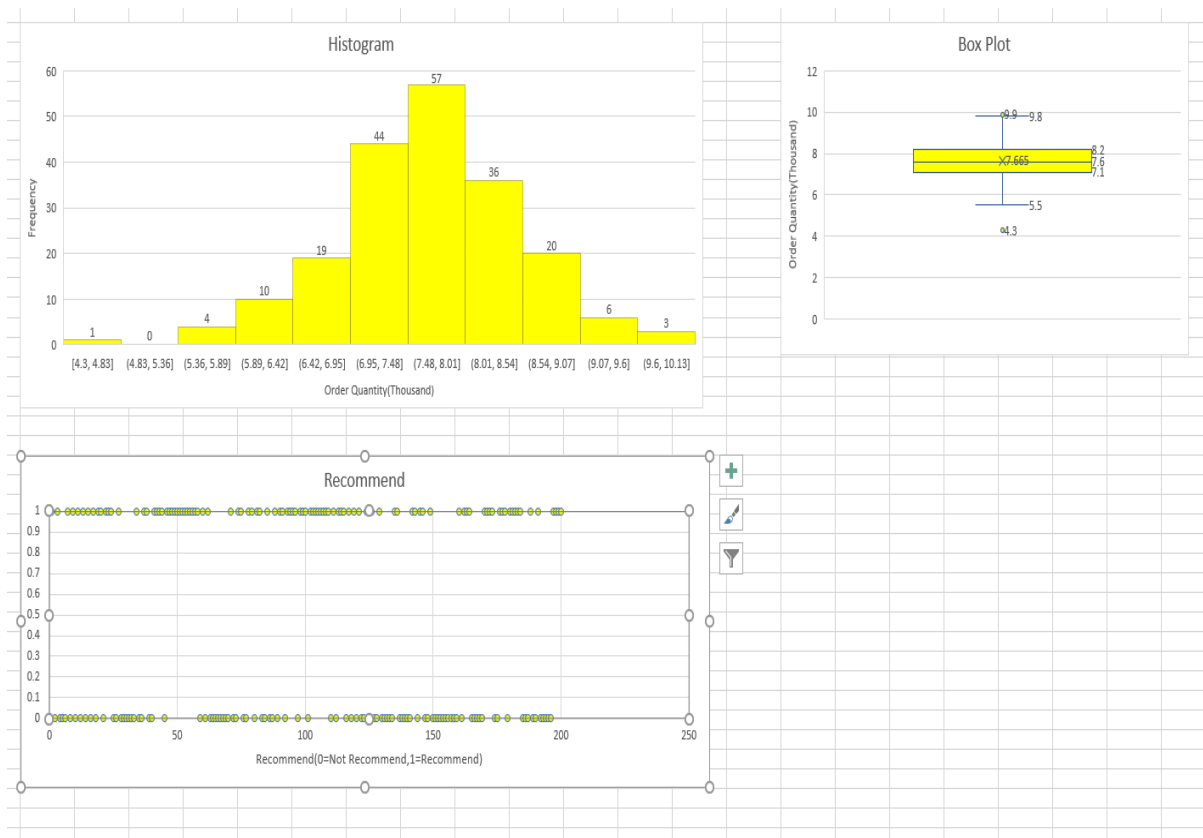
The above report has answered all the concerned questions and topics that being brought up in the minute of the meetings with respect to the Order Quantity, Quality, Brand Image and relationship between various other factors. The report provides the factors that influence the number of bottles ordered and which are the factors that are not significant in this model. We also helped Mad Dog Craft Beer to analyse its customer choices and predict its sales production of pale in the next four quarters. It must be remembered that this analysis is limited: a greater depth of understanding and evaluation can only occur if we have the bigger dataset and have more information like name of the beer, rating of the beer's taste & appearance, a text review or some kind of rating etc. The report has tried to use the best available model strategies to answer them. Improvements in every area of the sector are needed if the company or new entrant wanted to survive, grow and make a good profit. To get better insights, we can also look into other resources such as comparisons with other similar company's data, government regulation and many more factor in beer industries. Only after this process can a full appreciation of the company's current situation and possible future occurrences. The report is developed in such a way that it should be of good help, as it hopes to have satisfied every raised query from Todd Nash.

Appendices:

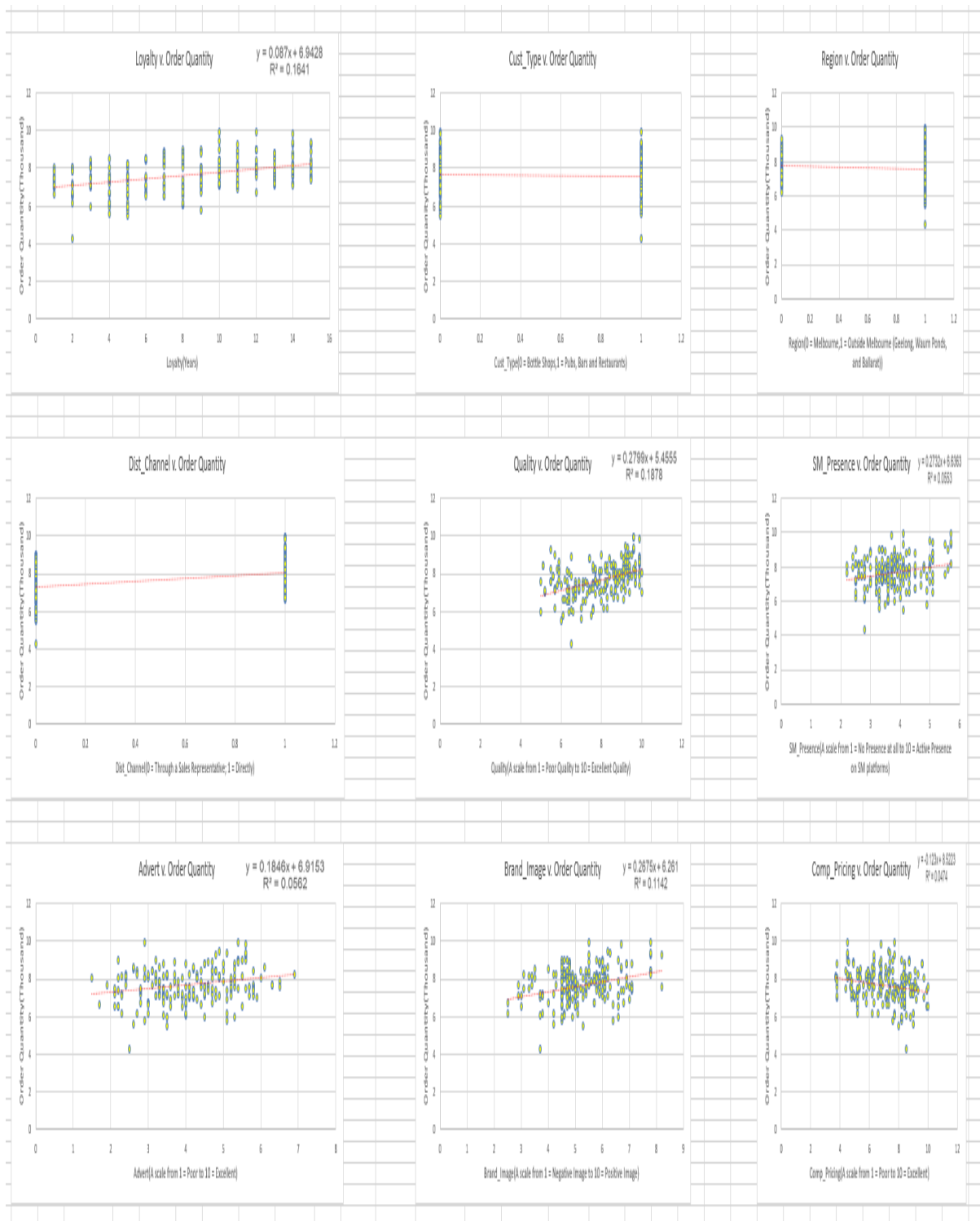
Q1.

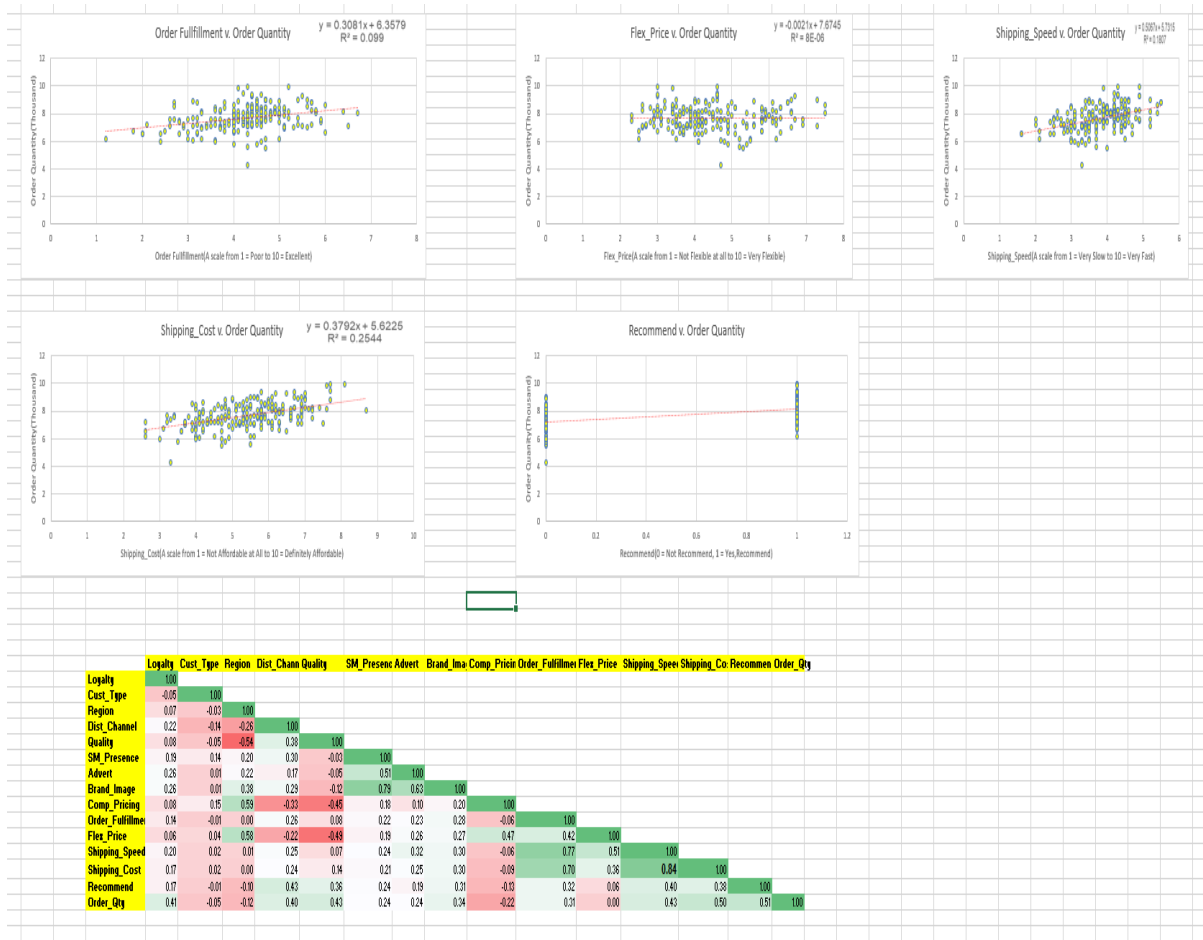
1.1

Order_Qty		Recommend	
Mean	7.67	Mean	0.51
Standard Error	0.06	Standard Error	0.04
Median	7.60	Median	1.00
Mode	7.20	Mode	1.00
Standard Deviation	0.89	Standard Deviation	0.50
Sample Variance	0.80	Sample Variance	0.25
Kurtosis	0.58	Kurtosis	-2.02
Skewness	-0.21	Skewness	-0.02
Range	5.60	Range	1.00
Minimum	4.30	Minimum	0.00
Maximum	9.90	Maximum	1.00
Sum	1533.00	Sum	101.00
Count	200.00	Count	200.00
Q1	7.1	Q1	0
Q3	8.2	Q3	1
IQR	1.1	IQR	1
LF	5.45	LF	-1.5
UF	9.85	UF	2.5
Potential Outliers	Yes	Potential Outliers	Not Possible



2.1





2.2

Initial Regression Model - After removing variable from Multicollinearity								
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.75							
R Square	0.56							
Adjusted R Square	0.53							
Standard Error	0.61							
Observations	200.00							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	13.00	89.23	6.86	18.36	0.00			
Residual	186.00	69.54	0.37					
Total	199.00	158.78						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 99.0%	Upper 99.0%
Intercept	4.18	0.54	7.74	0.00	3.12	5.25	2.78	5.59
Loyalty	0.06	0.01	4.98	0.00	0.03	0.08	0.03	0.08
Cust_Type	0.01	0.09	0.11	0.91	-0.17	0.19	-0.23	0.25
Region	0.01	0.14	0.04	0.97	-0.27	0.28	-0.36	0.37
Dist_Channel	0.06	0.11	0.50	0.61	-0.17	0.28	-0.24	0.36
Quality	0.20	0.04	4.48	0.00	0.11	0.29	0.08	0.31
SM_Presence	-0.11	0.10	-1.14	0.26	-0.30	0.08	-0.37	0.14
Advert	-0.03	0.05	-0.66	0.51	-0.13	0.07	-0.16	0.10
Brand_Image	0.22	0.08	2.80	0.01	0.07	0.38	0.02	0.43
Comp_Pricing	-0.06	0.04	-1.61	0.11	-0.13	0.01	-0.16	0.04
Order_Fulfillment	-0.14	0.07	-1.87	0.06	-0.28	0.01	-0.32	0.05
Flex_Price	0.05	0.06	0.75	0.46	-0.07	0.17	-0.11	0.20
Shipping_Cost	0.26	0.05	4.89	0.00	0.16	0.37	0.12	0.41
Recommend	0.32	0.11	2.90	0.00	0.10	0.53	0.03	0.60

Final Regression Model - After 8 iterations (Deleting Region, Customer Type, Distribution Channel, Advert Flex Price, Social Media Presence, Order Fulfillment, Competitive Pricing and Shipping Speed; one variable at a time)

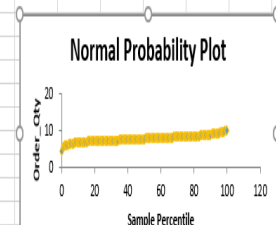
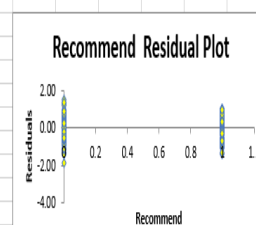
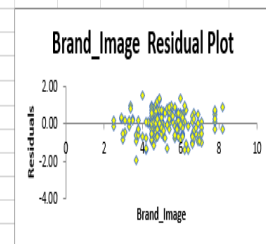
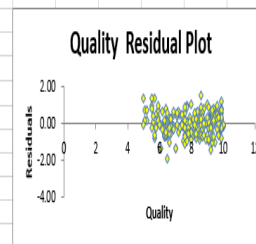
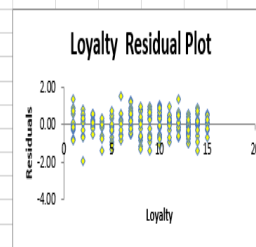
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.74
R Square	0.54
Adjusted R Square	0.53
Standard Error	0.61
Observations	200.00

ANOVA

	df	SS	MS	F	Significance F
Regression	5.00	86.10	17.22	45.96	0.00
Residual	194.00	72.68	0.37		
Total	199.00	158.78			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 99.0%	Upper 99.0%
Intercept	3.53	0.40	8.87	0.00	2.75	4.32	2.50	4.57
Loyalty	0.05	0.01	4.94	0.00	0.03	0.08	0.03	0.08
Quality	0.21	0.03	5.93	0.00	0.14	0.28	0.12	0.30
Brand_Image	0.13	0.04	2.97	0.00	0.04	0.22	0.02	0.24
Shipping_Cost	0.22	0.04	5.48	0.00	0.14	0.30	0.12	0.33
Recommend	0.34	0.10	3.24	0.00	0.13	0.54	0.07	0.61



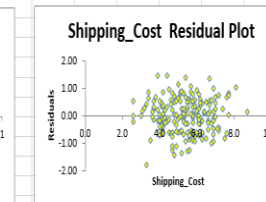
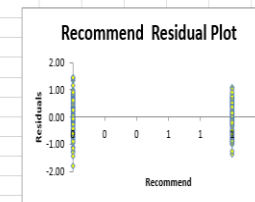
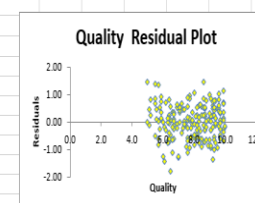
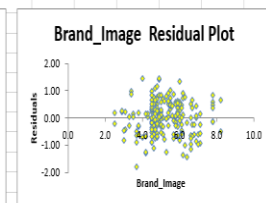
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.74
R Square	0.55
Adjusted R Square	0.53
Standard Error	0.61
Observations	200.00

ANOVA

	df	SS	MS	F	Significance F
Regression	6.00	86.84	14.47	38.84	0.00
Residual	193.00	71.93	0.37		
Total	199.00	158.78			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 99.0%	Upper 99.0%
Intercept	1.73	1.34	1.29	0.20	-0.91	4.36	-1.75	5.20
Loyalty	0.06	0.01	5.06	0.00	0.03	0.08	0.03	0.08
Brand_Image	0.46	0.24	1.94	0.05	-0.01	0.92	-0.16	1.07
Quality	0.43	0.16	2.65	0.01	0.11	0.76	0.01	0.86
Brand Image*Qualit	-0.04	0.03	-1.42	0.16	-0.10	0.02	-0.11	0.03
Shipping_Cost	0.21	0.04	5.21	0.00	0.13	0.29	0.11	0.32
Recommend	0.33	0.10	3.13	0.00	0.12	0.53	0.06	0.60

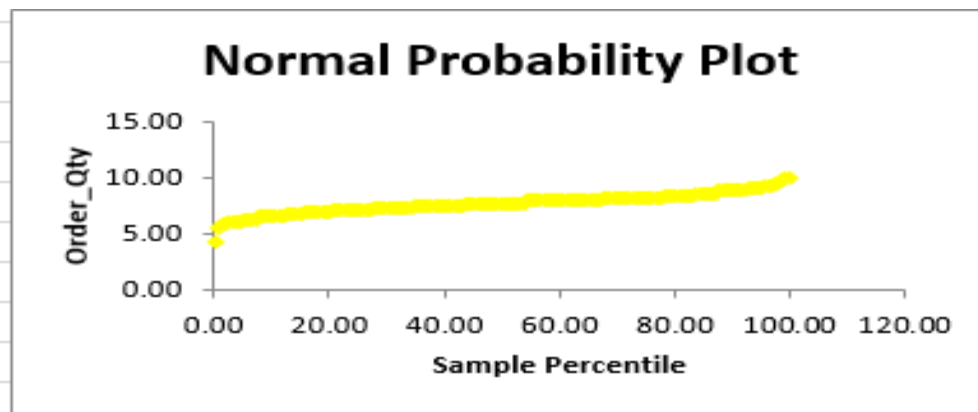


RESIDUAL OUTPUT

Observation	Predicted Order Qty	Residuals	Standard Residuals
1.00	7.86	0.54	0.90
2.00	7.93	-0.43	-0.72
3.00	8.41	0.59	0.98
4.00	6.97	0.23	0.39
5.00	8.03	0.97	1.62
6.00	6.59	-0.49	-0.81
7.00	6.68	0.52	0.86

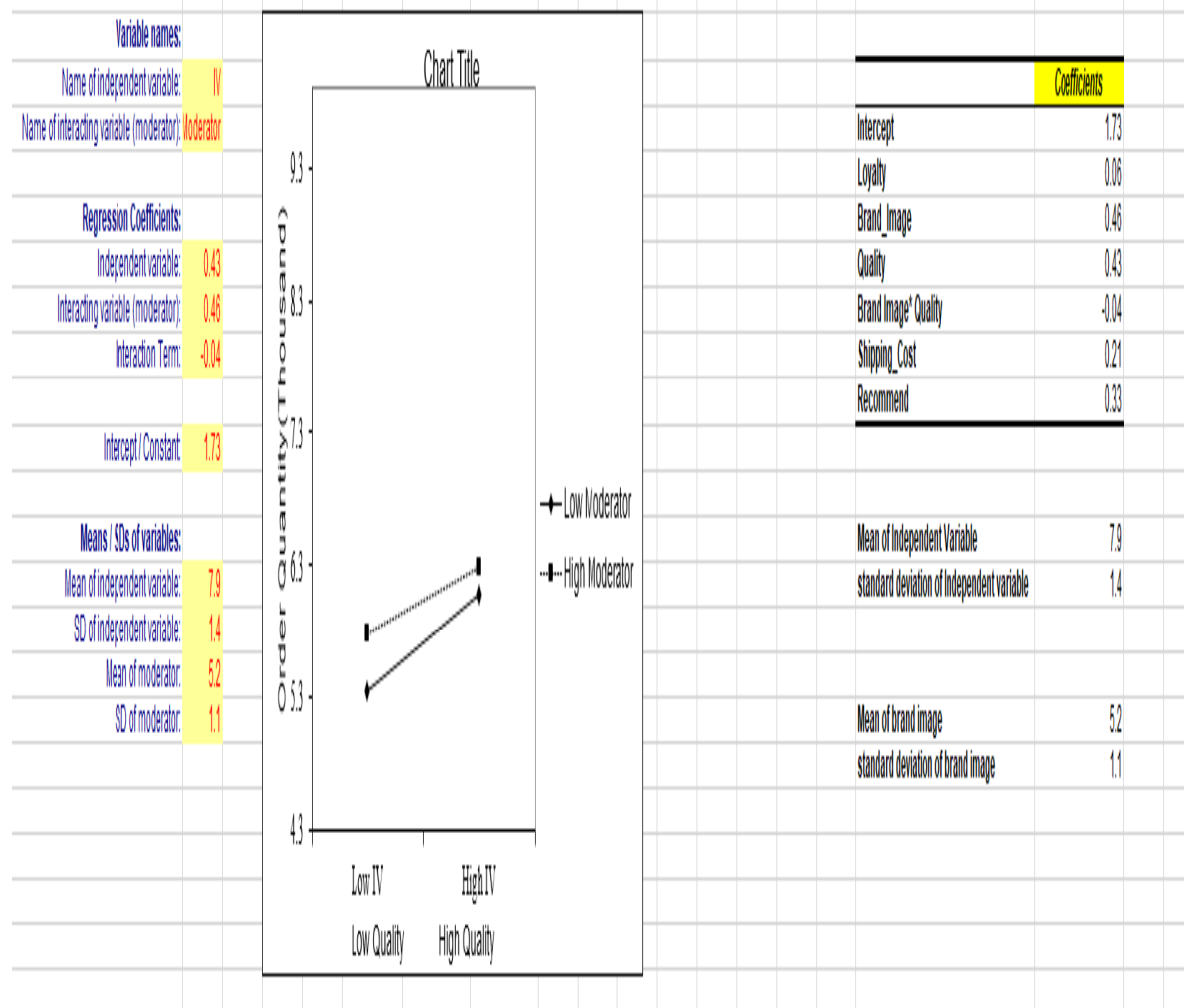
PROBABILITY OUTPUT

Percentile	Order Qty
0.25	4.30
0.75	5.50
1.25	5.60
1.75	5.80
2.25	5.80
2.75	6.00
3.25	6.00



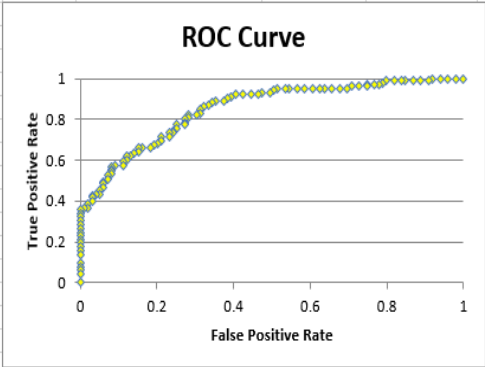
2.3

Enter information from your regression in the shaded cells



3.1

	coeff b	s.e.	Wald	p-value	exp(b)	lower	upper	%Change in Odds
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	163.36
Quality	0.65	0.15	18.10	0.00	1.92	1.42	2.60	92.34
Brand_Image	0.62	0.20	10.13	0.00	1.86	1.27	2.73	86.15
Shipping_Speed	1.16	0.29	16.20	0.00	3.19	1.81	5.60	218.52



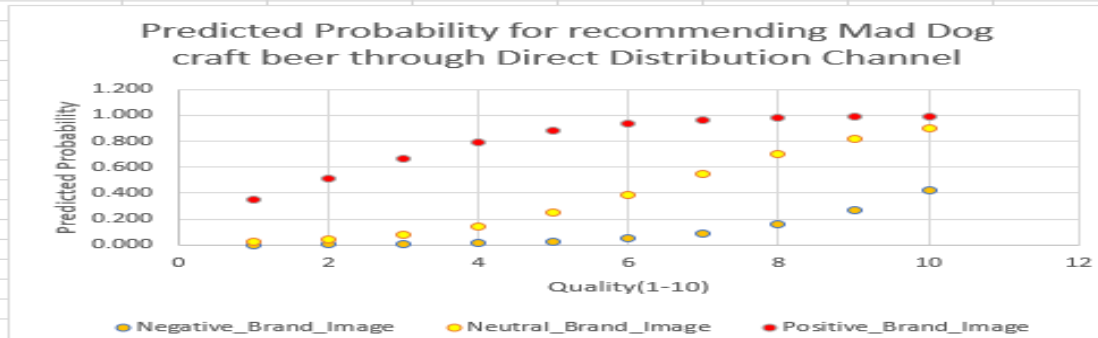
Classification Table									
Coeff	LL0	-138.62	Covariance Matrix					Converge	
	LL1	-95.38	5.20	0.13	-0.28	-0.29	-0.40	0.00	Suc-Obs
									Fail-Obs
-13.28			0.13	0.14	-0.02	-0.02	0.00	0.00	Suc-Pred
0.97	Chi-Sq	86.48	-0.28	-0.02	0.02	0.01	0.01	0.00	Fail-Pred
0.65	df	4.00	-0.29	-0.02	0.01	0.04	0.00	0.00	
0.62	p-value	0.00	-0.40	0.00	0.01	0.00	0.08	0.00	Accuracy
1.16	alpha	0.05							
	sig	yes							Cutoff
	R-Sq (L)	0.31						Maximum Chance Criterion	MCC
	R-Sq (CS)	0.35						Proportional Chance Criterion	PCC
	R-Sq (N)	0.47						Rule of Thumb	Standard
	Hosmer	103.35							
	df	106.00							
	p-value	0.55							
	alpha	0.05							
	sig	no							

3.2

0_Dist_Channel	Quality	Speed of Delivery	Negative_Brand_Image	Predicted_Probability	Neutral_Brand_Image	Predicted_Probability	Positive_Brand_Image	Predicted_Probability		Coeff
0	1	5	1	0.002	5	0.023	10	0.350		
0	2	5	1	0.003	5	0.044	10	0.509		-13.278
0	3	5	1	0.007	5	0.081	10	0.666		0.968
0	4	5	1	0.014	5	0.146	10	0.793		0.654
0	5	5	1	0.026	5	0.248	10	0.880		0.621
0	6	5	1	0.050	5	0.388	10	0.934		1.159
0	7	5	1	0.092	5	0.549	10	0.964		
0	8	5	1	0.163	5	0.701	10	0.981		
0	9	5	1	0.273	5	0.818	10	0.990		
0	10	5	1	0.419	5	0.896	10	0.994		

1	Dist Channel	Quality	Speed of Delivery	Negative Brand Image	Predicted Probability	Neutral Brand Image	Predicted Probability	Positive Brand Image	Predicted Probability
1	1	1	5	1	0.005	5	0.059	10	0.586
1	2	5	5	1	0.010	5	0.108	10	0.732
1	3	5	5	1	0.019	5	0.190	10	0.840
1	4	5	5	1	0.036	5	0.311	10	0.909
1	5	5	5	1	0.067	5	0.465	10	0.951
1	6	5	5	1	0.122	5	0.625	10	0.973
1	7	5	5	1	0.211	5	0.762	10	0.986
1	8	5	5	1	0.340	5	0.860	10	0.992
1	9	5	5	1	0.497	5	0.922	10	0.996
1	10	5	5	1	0.655	5	0.958	10	0.998

3.3



4.

Time Period	Year	Quarter	Y _t	4 MA	Centered MA	Ratio	Baseline S _t , I _t	S _t	Y _t /S _t	T _t	S _t × T _t	APE	Quarter	1	2	3	4
1	2009	Q3	1239.60						1.0	1398.27	1194.60	0.01				0.94	0.97
2		Q4	989.07						0.88	1020.40	1130.69	0.01				1.20	0.87
3	2010	Q1	1044.50					0.94	0.96	1082.74	1143.79	0.06				1.19	0.87
4		Q2	1030.57					0.97	0.97	1024.17	1055.00	0.04				1.22	0.85
5		Q3	1365.43					1.20	1.18	1053.90	1082.90	0.03				1.17	0.85
6		Q4	1006.00					0.87	0.88	1140.73	1200.07	0.05				1.19	0.91
7	2011	Q1	1135.43					0.98	0.96	1078.04	1207.07	0.03				1.18	0.90
8		Q2	1101.40					0.93	0.97	1135.33	1234.35	0.09				1.17	0.91
9		Q3	1457.37					1.19	1.18	1232.10	1297.06	0.02				1.16	0.89
10		Q4	1122.17					0.87	0.88	1272.46	1338.45	0.00				1.16	0.89
11	2012	Q1	1332.87					0.99	0.96	1381.67	1395.95	0.07					SUM
12		Q2	1348.37					0.97	0.97	1389.91	1382.64	0.06					
13		Q3	1716.77					1.22	1.18	1450.00	1381.74	0.09					
14		Q4	1193.47					0.85	0.88	1380.11	1338.83	0.02					
15	2013	Q1	1394.40					0.99	0.96	1445.46	1353.33	0.06					
16		Q2	1376.17					0.99	0.97	1418.96	1371.02	0.03					
17		Q3	1626.30					1.17	1.18	1374.86	1388.12	0.01					
18		Q4	1770.93					0.85	0.88	1327.75	1405.21	0.06					
19	2014	Q1	1340.70					0.96	0.96	1389.79	1422.01	0.02					
20		Q2	1410.90					0.98	0.97	1454.06	1439.40	0.01					
21		Q3	1745.97					1.19	1.18	1475.23	1495.90	0.01					
22		Q4	1334.83					0.91	0.88	1503.61	1473.69	0.03					
23	2015	Q1	1388.50					0.94	0.96	1438.34	1490.69	0.04					
24		Q2	1401.30					0.94	0.97	1444.47	1507.70	0.04					
25		Q3	1738.37					1.18	1.18	1518.76	1524.80	0.00					
26		Q4	1416.07					0.90	0.88	1616.72	1541.97	0.04					
27	2016	Q1	1573.17					0.97	0.96	1630.77	1599.07	0.04					
28		Q2	1618.60					0.98	0.97	1668.47	1678.16	0.06					
29		Q3	1953.87					1.17	1.18	1691.07	1693.26	0.04					
30		Q4	1528.40					0.91	0.88	1733.10	1610.05	0.07					
31	2017	Q1	1624.50					0.97	0.96	1833.80	1627.45	0.03					
32		Q2	1599.47					0.97	0.97	1840.74	1644.94	0.00					
33		Q3	1871.73					1.16	1.18	1880.76	1660.64	0.05					
34		Q4	1428.30					0.89	0.88	1818.99	1670.73	0.04					
35	2018	Q1	1519.60					0.94	0.96	1875.24	1695.83	0.08					
36		Q2	1638.10					1.00	0.97	1898.97	1712.92	0.01					
37		Q3	1947.43					1.16	1.18	1945.74	1730.02	0.05					
38		Q4	1527.30							1732.94	1747.11	0.01					
39	2019	Q1	1684.16					0.96		1745.83	1764.21	0.01	MAPE	4%			
40		Q2						0.97		1791.00	1720.06						
41		Q3						1.18		1791.40	1720.06						
42		Q4						0.88		1815.43	1670.06						
43	2020	Q1						0.96		1832.99	1767.06						

