Project 1: Predicting Catalog Demand

## **Step 1: Business and Data Understanding**

*Provide an explanation of the key decisions that need to be made. (500 word limit)*

### **Key Decisions:**

*Answer these questions*

**1. What decisions needs to be made?**

Should the company need to send these catalogs to the new 250 customers. If yes, how much profit can except return from the new customers.

**2. What data is needed to inform those decisions?**

Previous sales details of customers i.e.; how many products purchased and average sales amount

From the previous sales details, we have to calculate average sales per customer then multiply average sales value with score(yes) to get revenue. Then calculate profit by multiplying revenue by the gross margin which is 50% then subtracting 6.50(catalog cost).

## **Step 2: Analysis, Modeling, and Validation**

*Provide a description of how you set up your linear regression model, what variables you used and why, and the results of the model. Visualizations are encouraged. (500 word limit)*

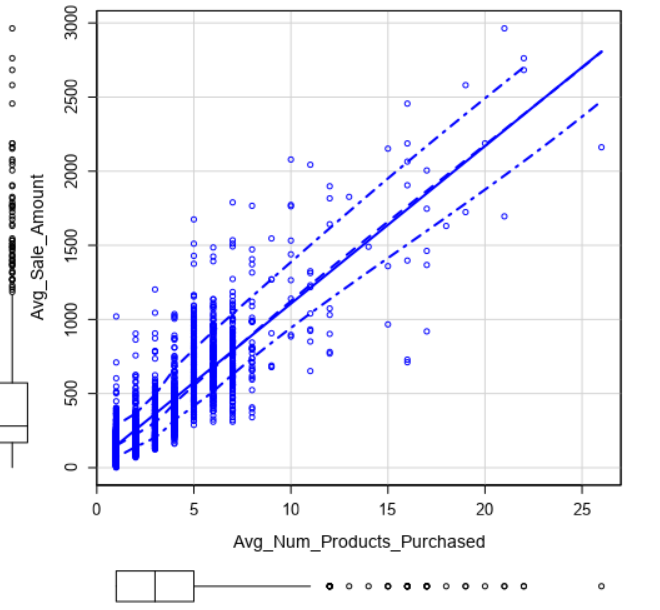
***Important:******Use the p1-customers.xlsx to train your linear model.***

*At the minimum, answer these questions:*

**1. How and why did you select the predictor variables in your model? You must explain how your continuous predictor variables you’ve chosen have a linear relationship with the target variable. Please refer back to the “Multiple Linear Regression with Excel” lesson to help you explore your data and use scatterplots to search for linear relationships. You must include scatterplots in your answer.**

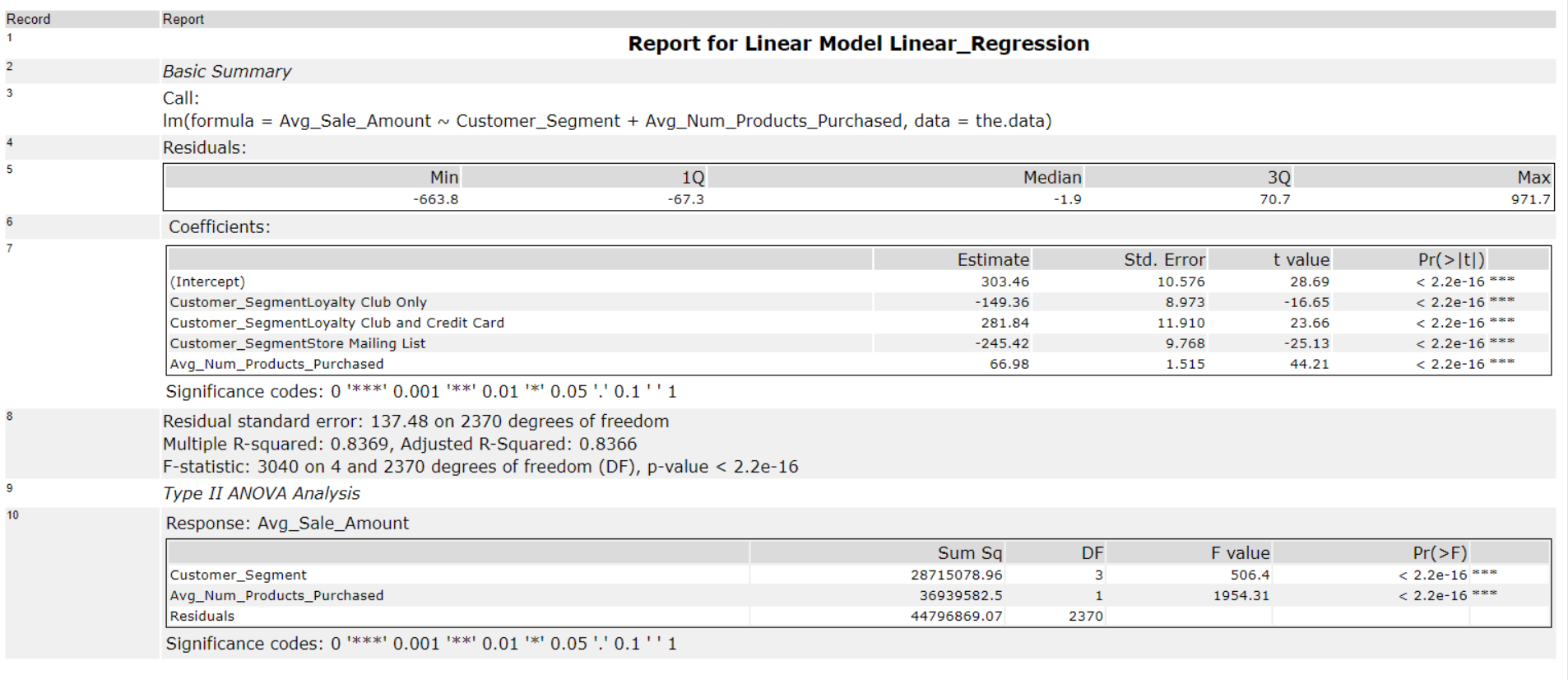
For my linear model, I have taken Customer\_Segment and Avg\_Num\_Products\_Purchased as predictor variable where as Avg\_Sale\_Amount as target variable. Below is the scatter plot between Avg\_Sale\_Amount and Avg\_Num\_Products\_Purchased.

From below can say that number of products purchased increases, sales amount increases. There is a positive correlation between these two variables.



**2. Explain why you believe your linear model is a good model. You must justify your reasoning using the statistical results that your regression model created. For each variable you selected, please justify how each variable is a good fit for your model by using the p-values and R-squared values that your model produced.**

This model is a good model since the multiple R-squared of 0.8369 and an adjusted R-Squared of 0.8366 which is a high number (>0.7) and also the P-value for both predictor variables is less than 0.05.



**3. What is the best linear regression equation based on the available data? Each coefficient should have no more than 2 digits after the decimal (ex: 1.28)**

**Important: The regression equation should be in the form:**

*Y = Intercept + b1 \* Variable\_1 + b2 \* Variable\_2 + b3 \* Variable\_3……*

Regression equation:

**Y = 303.46 + (66.98\*Average Number of Products) + (-149.36 \* Customer\_Segment Loyalty Club Only) + (281.84 \* Customer\_Segment Loyalty Club and Credit Card) + (-245.42 \* Customer\_Segment Store Mailing List) + 0 (if Credit Card)**

## **Step 3: Presentation/Visualization**

*Use your model results to provide a recommendation. (500 word limit)*

*At the minimum, answer these questions:*

**1. What is your recommendation? Should the company send the catalog to these 250 customers?**

After conducting an analysis on data, I would recommend that the company send out the catalog to these 250 customers since the profit getting from them is $21987.44which is more than $10000.

**2. How did you come up with your recommendation? (Please explain your process so reviewers can give you feedback on your process)**

I have created a model on Alteryx using Customer\_Segment and Avg\_Num\_Products\_Purchased as predictor variables and Avg\_Sale\_Amount as target variable. After running the model, got the

Intercept and coefficients values. Using these values, predicted the expected revenue (multiply revenue by the score rate) from each customer in the mailing list, and then predicted the profit (by multiplying gross-margin (0.5) and subtracting 6.50 from expected revenue).

**3. What is the expected profit from the new catalog (assuming the catalog is sent to these 250 customers)?**

The expected profit the new 250 customers will be $21987.44