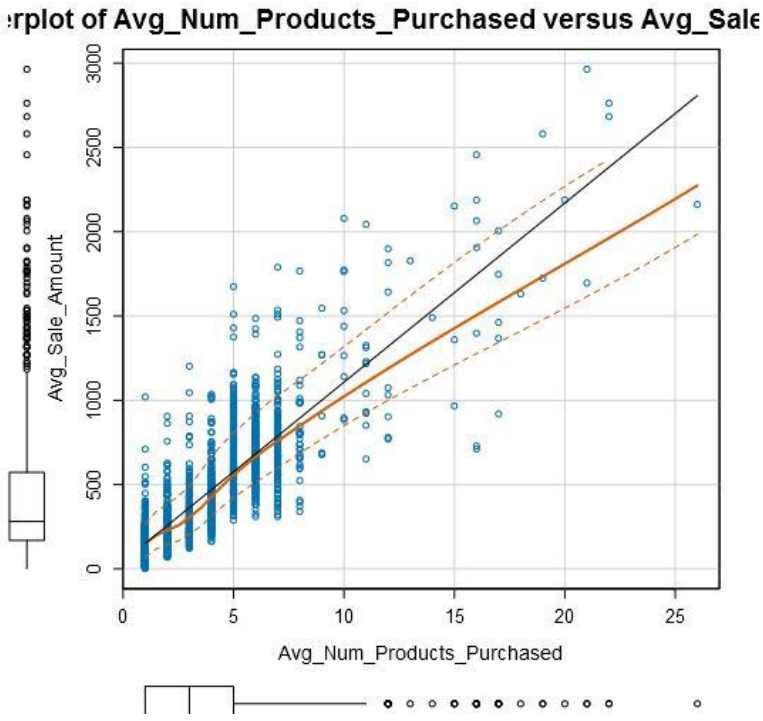
Business and Data understanding:-

1. The decision that needs to be made is whether or not sending the catalog to the 250 new consumers will be profitable and how much revenue can the company generate from sending out these catalogs. The profit margin should be greater than $10000 so as to make the decision to send the catalog to the consumers. The second decision would be whether sending the catalog is a good marketing idea or not.
2. The data required to make the decision would be the average number of products that the customers purchased in the sample data, data showing the probability whether the customer will purchase from the catalog ('Score_Yes' variable from the new customer dataset) and the customer segment data.

Analysis, Modeling and Validation:-

1. The target variable to be considered will be 'Average sale amount' as this is what we're predicting. The predictor variable used will be 'Average number of products purchased' and 'Customer Segment' as these have high adjusted r-squared values (>0.7) and low p-values (<0.05).



2. This is a good linear model as both the adjusted r-squared value and p-value are pretty good.

P-Value

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	303.46	10.576	28.69	< 2.2e-16	***
Customer_SegmentLoyalty Club Only	-149.36	8.973	-16.65	< 2.2e-16	***
Customer_SegmentLoyalty Club and Credit Card	281.84	11.910	23.66	< 2.2e-16	***
Customer_SegmentStore Mailing List	-245.42	9.768	-25.13	< 2.2e-16	***
Avg_Num_Products_Purchased	66.98	1.515	44.21	< 2.2e-16	***

As the p-value for all the predictor variables is $2.2 * 10^{-16}$, so it is way less that 0.05 and is statistically significant.

r-square value

Multiple R-squared: 0.8369, Adjusted R-Squared: 0.8366

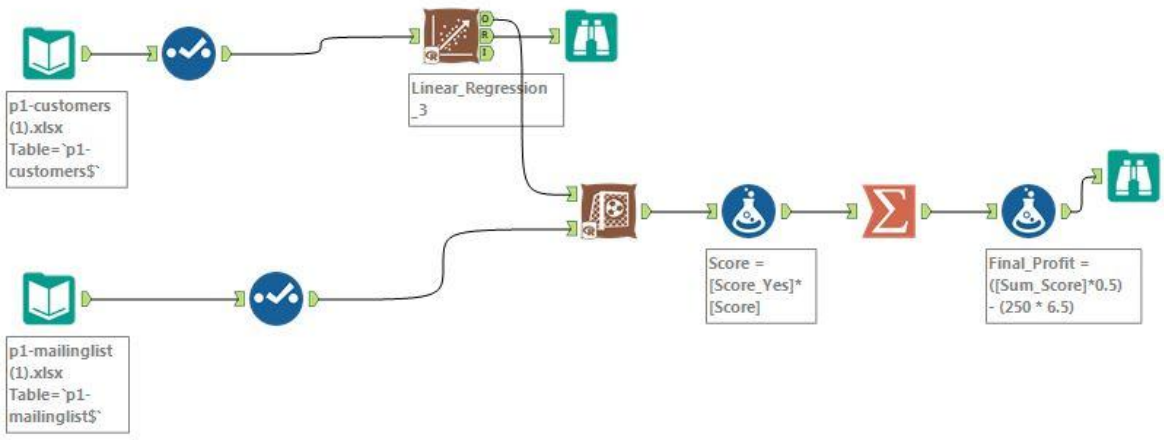
As the adjusted r-squared value is 0.8366, it is greater than 0.7 and hence is considered a good correlation.

3. Regression equation :-
   Avg Sale amount = 303.46 + 66.98 * Avg Number of Products Purchased – 149.36 * Customer_Segment Loyalty Club Only + 281.84 * Customer_Segment Loyalty Club and Credit Card – 245.42 * Customer_Segment Store Mailing List

Presentation/Visualization:-

1. My recommendation would be that the company should definitely proceed with sending the catalog.

2. First I used the linear regression model to compute the average sales of all customers on the new dataset and then multiplied it by the probability of customers buying from the catalog. After this I calculated the gross margin (price-cost) and multiplied by 0.5 (50%). Finally, I subtracted the cost of 250 catalogs from the amount to arrive at the final answer.



3. The expected profit is 21987.43.