GoodBelly Case Study Report

Jeff Baker Leighton Dong

October 7, 2014

Executive Summary

Armed with a relatively small budget, marketers at GoodBelly were faced with the task of justifying the expenses of promoting their new products. The team felt confident that the efforts put forth to promote their probiotic drinks were impactful, but struggled to quantitatively justify the marketing spend. This analysis was conducted to answer: "Do endcap displays or in-store demonstrations help boost sales of GoodBelly products?" The goal of this study was to address executives' hypothesis that these efforts had no effect, and to determine the impact, if any, of these marketing efforts on sales.

The data covered 126 stores over eleven weeks (1,386 total observations). We began with a preliminary examination of each of the provided variables individually, to ensure the data were clean and realistic. Next, we generated a number of linear models to determine if any model would answer the question. A series of diagnostic tests were performed on each proposed model, but none of the candidates were valid.

After a number of inconclusive findings, a search algorithm was employed to find the optimal model to explain this data. The algorithm calculated the best, valid combination of variables to fit into the model for this data. With this model, we concluded that both the presence of sales reps (in conjunction with endcaps) and in-store demonstrations had a significant, positive impact on GoodBelly unit sales, but endcaps by themselves did not. Our recommendation is that these efforts ought to continue, as long as budget and business goals permit.

Problem & Approach

The purpose of this analysis was to determine if the presence of GoodBelly demonstrations and/or endcap displays in stores justified the associated marketing costs to implement them. Did these marketing efforts impact sales? To address this question, we leveraged linear regression techniques in order to gauge the relative importance and relationship of the factors within this business problem. We began with a high-level exploration of the data, then began creating linear model candidates to help answer the question. After having a number of models to work with, we chose the one that most appropriately helped address whether the marketing efforts made an impact in sales or not.

Analysis

Data Exploration

Data was collected from 126 stores across multiple regions; each store had eleven weeks of observational data (1,386 rows of data total) for multiple variables. The variables consisted of: units sold; average retail price of the product each week; whether the store had a regional sales rep or not; the presence of an endcap; the presence of a demo for that week; whether there was a demo 1-3 weeks prior, or 4-5 weeks prior; and the number of natural retailers and fitness centers near each store. With units sold as our response variable, we began our analysis with a preliminary univariate/bivariate examination across these eight predictor variables to ensure the data were clean and realistic, and to note any interesting findings.

First, we found that retail prices were normally distributed across weeks and stores. We also found a "long tail" of units sold each week; that is, most stores sold 50-450 units/week. However, some stores sold more than 450 units over a fair number of weeks, so we flagged these observations as outliers. When we began examining paired variables, we made a similar observation: plotting units sold against retail price revealed that the majority of weekly observations lay within the \$3-\$5 price range–thus, most stores generated roughly \$150-\$2250 of unit sales each week, but there were a noticeable amount of weeks with units sold over 450, or average retail prices over \$5. Finally, we calculated correlation coefficients between units sold and the other variables, in order to determine how related these variables were to one another. We were relieved to find a weak correlation between units sold and retail price. With a value close to zero, any impact found by a winning model–good or bad–would become immediately apparent.

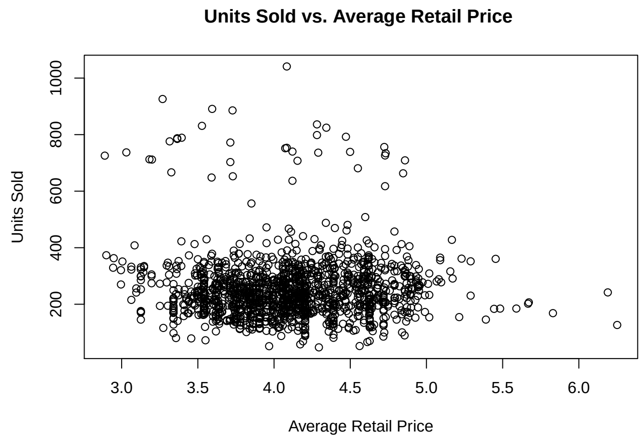


Figure 1: Most sales were between \$150-\$2,250/week

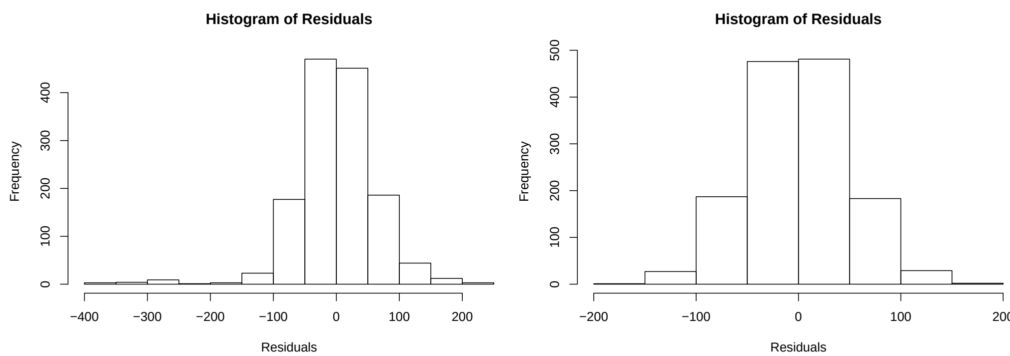
Linear Models

Once we completed the univariate and bivariate analysis, we developed a number of linear models on our data (ultimately, we generated a total of six different models). One of the goals in creating a linear regression model is to minimize the residuals, or distances of the observations from our "best fit line." The first was a straightforward full model of the eight variables, without any modifications. We found that both fitness centers and natural retailers were insignificant, while the other variables were all highly significant. The model itself explained 67% of the variation in units sold but, unfortunately, we found clear signs of non-constant variance (heteroskedasticity) in the residuals plot. In addition, we found the residuals were not normally distributed, indicating more work would need to be done to find the right model.

With the issues of non-normality and heteroskedasticity in mind, we attempted a Box-Cox transformation to correct the model, raising the units sold to the $\frac{1}{3}$ power. While not guaranteed to work, performing this modification to the response variable often "fixes" the residuals to a normal, bell-shaped distribution; however, in this case, the transformation actually did not have an impact on normality. In fact, the model explained slightly less variation (61%); additionally this new model continued to exhibit non-constant variance. As a result, we opted to continue our investigation, honing in on demos and endcaps as our primary focus.

In our third model, we narrowed our scope to simply unit price and endcap presence, testing their effect on volume. We also accounted for potential interactions between these two variables, resulting in three total variables regressed on volume. Disappointingly, the resulting model explained less of the variation in sold inventory (only 35%), and still exhibited heteroskedasticity and non-normality issues (the data actually skewed left). We were surprised to also find a clear wave pattern in the sequence plot (this is a plot of the residuals by observation); this is a clear indicator that the residuals are not independent of each other. Thus, overall, we concluded the model to be invalid. Given the same problems as the last model, we attempted to correct this version with the same transformative process as before, this time using an exponent of $\frac{1}{5}$. The transformation resulted in an even worse fourth model, without correcting any of the problematic symptoms found previously.

We found no luck with a minimalist endcap model, so we switched to the other key variable on our fifth modeling attempt: the presence of a demonstration in the store. Combining the three time-sequenced demo variables with average retail price, we ran a regression model and checked its validity on units sold. By this point in our analysis, the invalid model was not surprising. Here, yet again, we found heteroskedasticity in the model, as well as non-normality in the residuals. With only 30% of the variation explained by this version, and a number of insignificant variables in play, we concluded that more facets of the data ought to be included to find a solution that would explain the impact of the marketing efforts to sales.



(a) (b)

Figure 2: Non-normal vs. normal distribution in residuals

Having had little success on our own in finding a clear model, we decided on our sixth attempt to employ one of many algorithmic techniques to assist us. We chose to run a backward stepwise regression search, leveraging the full spectrum of variables available to us in the data (including interaction terms). This method involves simulating multiple permutations of regression models, measuring the *Akaike Information Criterion (AIC)* metric of each version to identify the most parsimonious candidate (the most minimalistic model that still best fits the data). When the algorithm finds the model with the smallest AIC, we have our winning model.

After a handful of iterations, the algorithm ultimately deduced the most appropriate linear model for this data set, detailed below. Upon further inspection, the model made sense in justifying the relative impact of the marketing efforts. The model explains just over 80% of the variation in the residuals, which is quite good. Additionally, all issues of heteroskedasticity and non-normality were no longer present, further validating the models appropriateness for this case.

	Estimate
(Intercept)	284.6488
Average Retail Price	-22.9761
Sales Rep	**52.1411**
Endcap	2.1593
Demo (in week)	**102.1490**
Demo (1-3 weeks ago)	**71.7237**
Demo (4-5 weeks ago)	**62.6585**
Natural Retailers	-2.3592
Fitness Centers	-0.4243
Sales Rep:Endcap	**450.0663**
Sales Rep:Natural Retailers	5.3033
Demo:Demo (1-3 weeks ago)	23.4603
Demo (4-5 weeks ago):Fitness Centers	4.8149

Table 1: Final linear model, with significant variables highlighted

Having found a valid solution that passed on all initial diagnostics, we sought to interpret the model. To begin, we noted the coefficients of each variable and their respective statistical significance. While some of the coefficients were flagged to be treated as zeros, we focused on the critical variables related to demos and endcap. From this data set, we were able to conclude that, with 95% confidence, having demo activities in any of the observed weeks of this study contributed positively to unit sales. Specifically, having a demo in a corresponding week would, on average and holding all else constant, increase unit sales by approximately 102. If a store had a demo 1-3 weeks or 4-5 weeks prior, unit sales would increase, on average, by approximately 72 and 63, respectively.

With demonstrations succinctly accounted for, we turned our attention to coefficients involving endcaps within the stores. The only significant coefficient was related to the interaction term between sales representatives and endcap activities; that is to say, according to this model, any contribution by endcap activities depended on the presence of a regional sales representative. With 95% confidence, we can conclude that for every week that an endcap is displayed in a store that has a regional sales representative, unit sales should increase by 502, on average (the presence of a sales rep contributes 52 unit sales of this figure). With this in mind, the model suggests that endcap activities by themselves do not contribute to selling units, holding all else constant; the combined effects of having a sales rep along with an endcap in stores provides the greatest contribution to sales within this model.

Conclusion & Recommendations

It is clear from this model that both endcap and demo efforts make a positive impact on sales, and ought to continue as long as budget and business goals permit. Proving the effectiveness of these marketing efforts, our recommendation is that endcap activities be continued at stores that are also supported by regional sales representatives. Additionally, product demonstrations provide a noticeable impact on moving unit volume at each store, despite the small decrease in effect over time. The overall effects of these marketing plans will help boost GoodBelly sales in these Whole Foods stores across the country.

Appendix

Basic Diagnostics

Y: Units sold
X_1: Average retail price
X_2: Presence of regional sales representative
X_3: Presence of endcap activities
X_4: Presence of demo activities in that week
X_5: Presence of demo activities 1-3 weeks ago
X_6: Presence of demo activities 4-5 weeks ago
X_7: Number of natural retailers nearby
X_8: Number of fitness centers nearby

The task of justifying GoodBelly's marketing spend began with a straightforward data exploration exercise.

	mean	sd	median	min	max
Units.Sold	253.82	111.00	236.74	47.56	1041.20
Average.Retail.Price	4.11	0.46	4.10	2.89	6.25
Sales.Rep	0.55	0.50	1.00	0.00	1.00
Endcap	0.04	0.19	0.00	0.00	1.00
Demo	0.06	0.23	0.00	0.00	1.00
Demo1.3	0.16	0.36	0.00	0.00	1.00
Demo4.5	0.08	0.26	0.00	0.00	1.00
Natural	1.45	0.98	1.00	0.00	4.00
Fitness	2.48	1.59	2.00	0.00	5.00

Table 2

From the above table, we can see the relatively large distance between mean and median for natural and fitness which indicates skewness. From the following box plots, we discovered a long tail of outliers beyond approximately 450 units in units sold while the distribution is roughly normal below 450 units. The average retail price has a shorter tail of outliers, but exhibits a normal distribution overall. The plots also showed that fitness center has a roughly uniform distribution while natural retailer slightly skewed right.

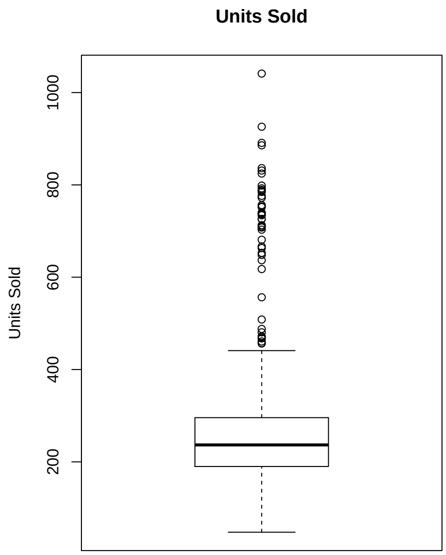


Figure 3: Box Plot

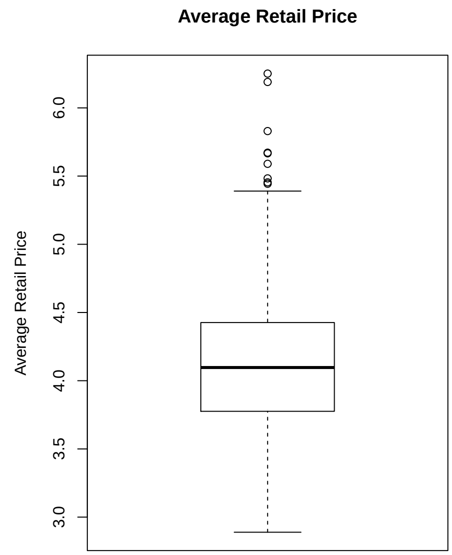


Figure 4: Box Plot

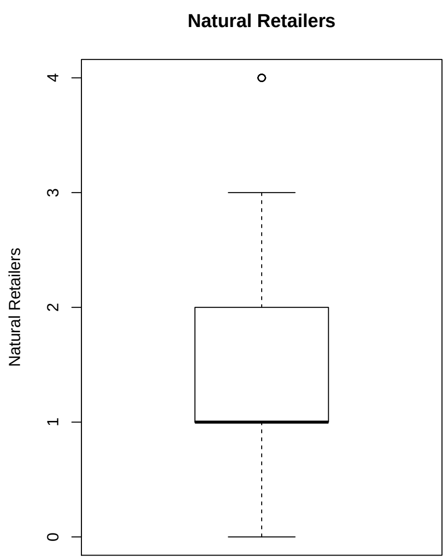


Figure 5: Box Plot

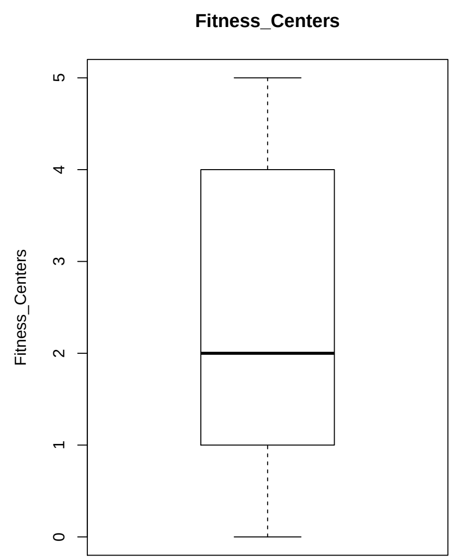


Figure 6: Box Plot

Looking at the scatter plot of units sold on the y-axis and average unit price on the x-axis, we find a random scatter with the mass formed roughly between $3-5$ and units sold ranging from 100-400. The randomness suggests a low correlation, further confirmed by a calculated correlation of 0.019. The correlation between independent variables are low in general. The relatively higher ones are 0.33 (X_1 and X_2), 0.24 (X_2 and X_5), 0.15 (X_2 and X_4) and 0.14 (X_2 and X_6).

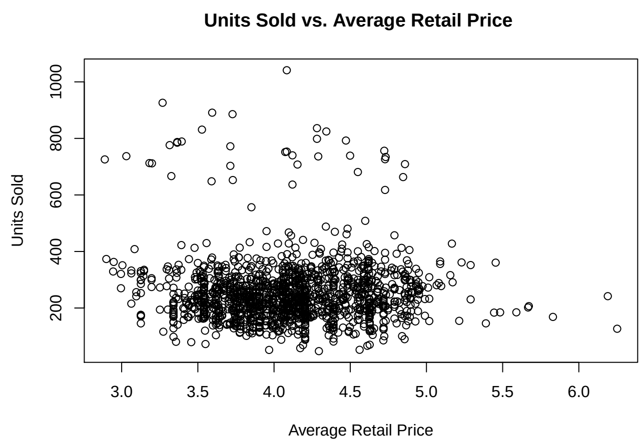


Figure 7: Scatter Plot

Linear Regression Models

After examining these variables, we then generated various linear models. We developed six in our analysis, listed below:

Model	R^2	R_a^2	F-stat	p-value	AIC
1	0.6726	0.6707	353.7	0.00	15459.36
2	0.6111	0.6088	270.5	0.00	2277.916
3	0.3555	0.3541	254.1	0.00	16388.25
4	0.2201	0.2184	130.0	0.00	3232.275
5	0.3002	0.2952	59.0	0.00	16516.24
6	0.8071	0.8054	478.6	0.00	14734.65

Table 3

1. Linear Regression Model of Data

The first model we generated represented the entire data set; we regressed units sold (Y) on X_1 through X_8. No interaction terms were included (Table 4).

$R^2 = 0.6728$, F-stat $= 353.7$, p-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	298.4881	16.1831	18.44	0.0000
Average.Retail.Price	-28.5354	3.9522	-7.22	0.0000
Sales.Rep	77.4369	3.8645	20.04	0.0000
Endcap	305.1021	9.0557	33.69	0.0000
Demo	111.1328	7.4037	15.01	0.0000
Demo1.3	73.5172	4.8954	15.02	0.0000
Demo4.5	67.5698	6.5420	10.33	0.0000
Natural	-1.5942	1.7764	-0.90	0.3697
Fitness	-1.0197	1.0840	-0.94	0.3471

Table 4

At the $\alpha = 0.05$ level of significance, the resulting model was statistically significant (F-statistic $= 353.7$; p-value $< 2.2 \times 10^{-16}$), explaining 67.28% of the variation in Y ($R^2 = 0.6728$); additionally, only b_7 and b_8 (natural retailer and fitness center) were insignificant (p-values of 0.3697 and 0.3471, respectively). We also noted the large coefficient of the end-cap variable (b_3) in this model, suggesting high positive influence of endcap activities on units sold. We also noted b_4 (111.1328), b_5 (73.5172) and b_6 (67.5698) to be statistically significant, which indicate positive effects of those variables on units sold. However, such conclusions could not be made at this point as we continued validating the model through other diagnostic measures.

Next, plotting the fitted values against residuals revealed clear heteroskedasticity in the data (Figure 8). A Breusch-Pagan test for heteroskedasticity confirmed this suspicion: $\chi^2_{BP} = 711.03$, with a p-value < our chosen 5% level of significance, we rejected the null hypothesis that the variance is constant.

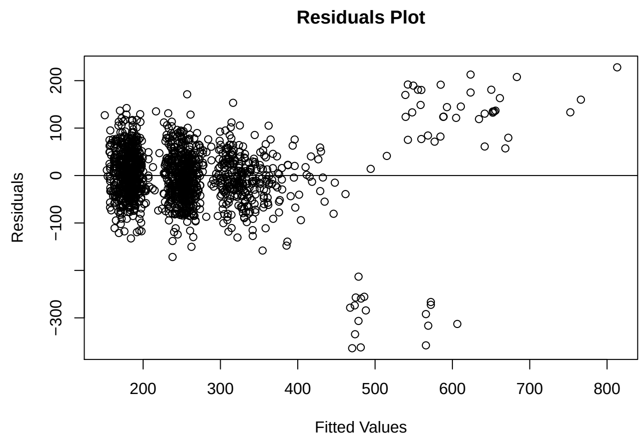


Figure 8: Fitted vs. Residuals

BP	p-value
711.03	0.00

Table 5

However, a sequence plot of the residuals showed no linear or cyclical pattern within the plot: a strong indicator that the residuals are independent for this case (Figure 9). The studentized residual plot displayed the existence of outliers (Figure 10).

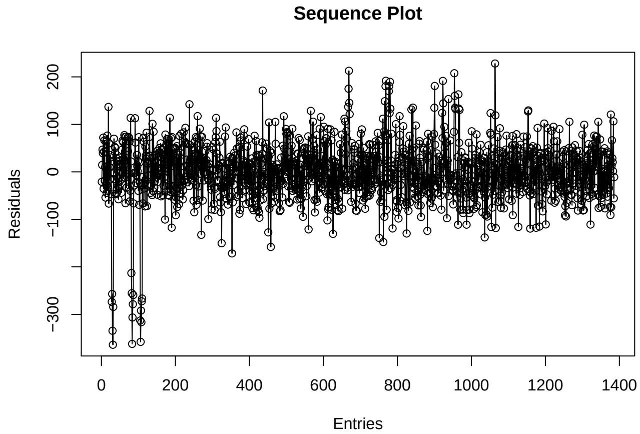


Figure 9: Residuals Sequence Plot

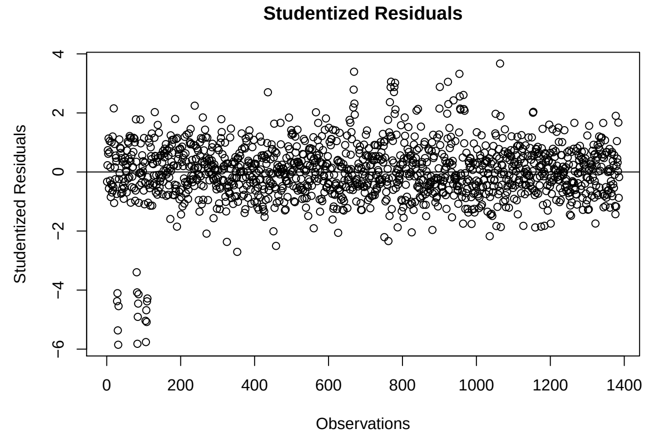


Figure 10: Studentized Residuals

Additionally, a check on normality of the data was performed by inspecting the Q-Q plot of the MLR: the heavy-tailed symmetrical pattern indicates the data are not normally distributed (Figure 11). A Shapiro-Wilk normality test on the residuals returned a p-value much smaller than our 5% significance level, so we rejected the null hypothesis that the data are normally distributed. This supported the invalidity of the model in it's current form.

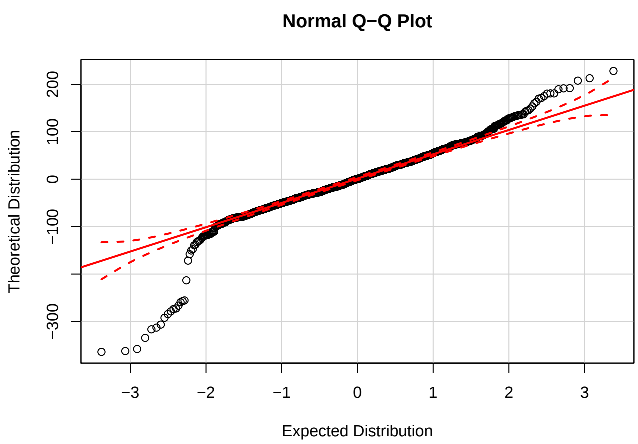


Figure 11: Q-Q Plot

W	p-value
0.93	0.00

Table 6

2. Transformed Linear Model

As a second approach to improving this model, we conducted a Box-Cox transformation on our dependent variable (units sold), using the `boxcox()` function in R to find the best lambda. In this case, we minimized SSE at $\lambda = 0.34$. We then regressed the transformed units sold variable on all eight other variables (Table 7).

$R^2 = 0.6111$, F-stat = 270.5, p-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	7.0459	0.1393	50.59	0.0000
Average.Retail.Price	-0.2646	0.0340	-7.78	0.0000
Sales.Rep	0.7214	0.0333	21.69	0.0000
Endcap	1.7988	0.0779	23.08	0.0000
Demo	0.8660	0.0637	13.59	0.0000
Demo1.3	0.6223	0.0421	14.77	0.0000
Demo4.5	0.5954	0.0563	10.58	0.0000
Natural	-0.0111	0.0153	-0.72	0.4696
Fitness	-0.0080	0.0093	-0.86	0.3912

Table 7

After the transformation, R^2 decreased by 0.0615. The model remained significant and everything else stayed the same.

A second round of diagnostics tests were performed to check the validity of this updated model: $\chi^2_{BP} = 307.77$, with a p-value$_{BP} < 2.2 \times 10^{-16}$. Since the p-value$_{BP}$ is less than the 5% significance level, we again rejected the null hypothesis that the variance is constant, despite the transformation. This conclusion was also confirmed by the residual plot, which displays non-constant variance.

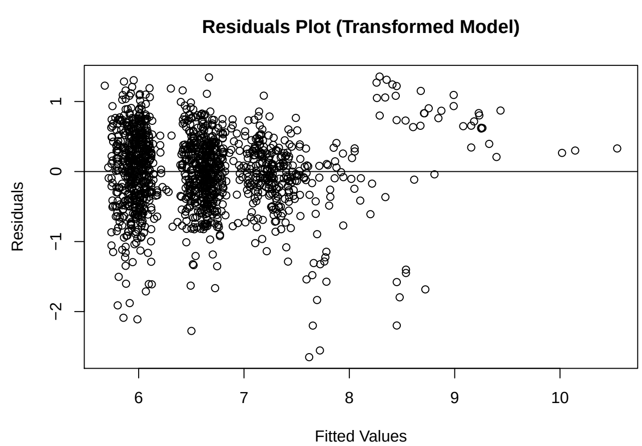


Figure 12: Fitted vs. Residuals

BP	p-value
307.77	0.00

Table 8

In addition, performing a Shapiro-Wilk normality test on the transformed data resulted in a p-value < 0.05, so we actually rejected the null hypothesis that the residuals are normal in this updated model. After looking at the Q-Q plot, our conclusion was confirmed.

W	p-value
0.98	0.00

Table 9

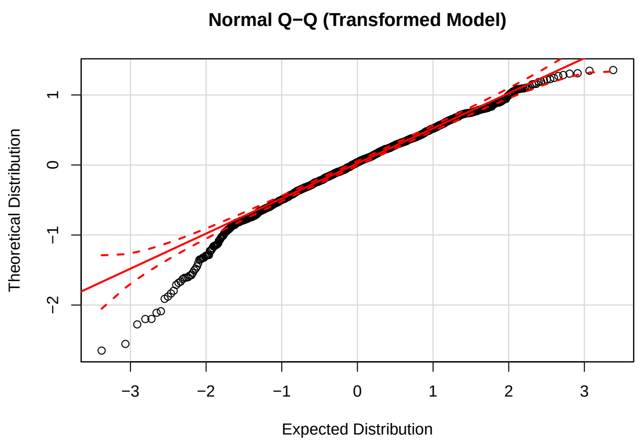


Figure 13: Q-Q Plot

However, the sequence plot did not display linear or cyclical patterns in the points, proving independence of the variables in this model (Figure 14). Moreover, the studentized residual plot still displayed outliers (Figure 15).

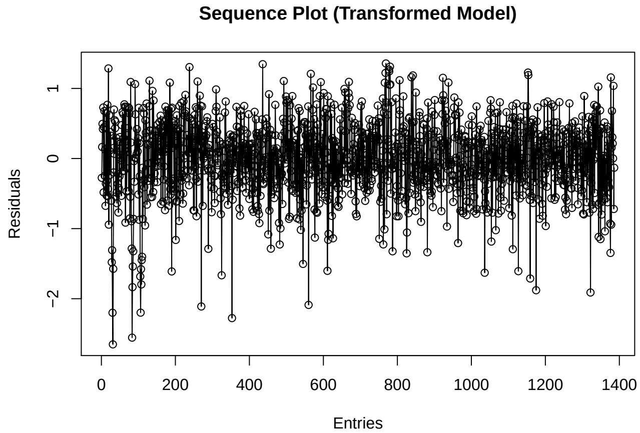


Figure 14: Residuals Sequence Plot

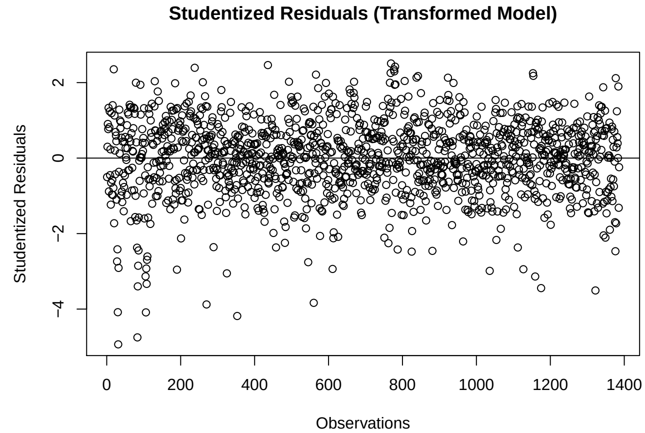


Figure 15: Studentized Residuals

3. "Endcap" Model

After developing an invalid full model, we decided to build a model focusing on one of the variables of interest – "Endcap." In this model, we regressed units sold (Y) on average retail price (X_1), Endcap (X_3) and the interaction term $X_1 : X_3$.

$R^2 = 0.3555$, F-stat = 254.1, p-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	207.1335	21.9968	9.42	0.0000
Average.Retail.Price	8.1594	5.3146	1.54	0.1249
Endcap	588.0227	94.8028	6.20	0.0000
Average.Retail.Price:Endcap	-61.6274	23.7377	-2.60	0.0095

Table 10

In this model, while the model overall was significant with a high F-statistic (254.1), it had a relatively low R^2 of 0.3555. Among the four coefficients, b_1 for average retail price was insignificant with p-value of 0.1249, which is greater than our 5% level of significance; however, all other coefficients were significant. The coefficient for Endcap (588.0227) and the coefficient for the interaction term (-61.6274) indicate that having endcap activities may change units sold by $588.0227 - 61.6274 \times$Average Retail Price, on average. However, this was deemed inconclusive until the validity of the model could be confirmed.

To check its validity, we plotted the fitted values against residuals and the plot revealed heteroskedasticity in the data (Figure 16). A Breusch-Pagan test for heteroskedasticity confirmed this suspicion: $\chi^2_{BP} = 601.84$, with a p-value < our chosen 5% level of significance, we rejected the null hypothesis that the variance is constant.

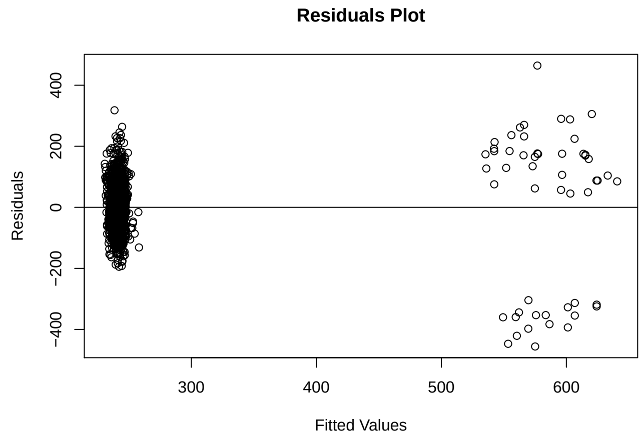


Figure 16: Fitted vs. Residuals

BP	p-value
601.84	0.00

Table 11

Furthermore, the studentized residual plot showed outliers in the data. Also, a sequence plot of the residuals showed a cyclical pattern within the plot, which indicates that the residuals are not independent for this case (Figure 17).

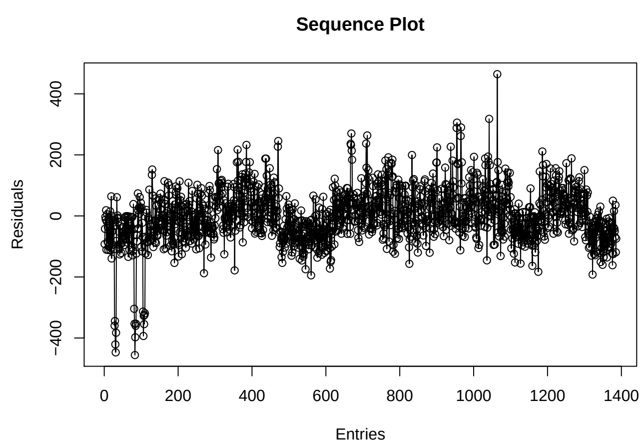


Figure 17: Residuals Sequence Plot

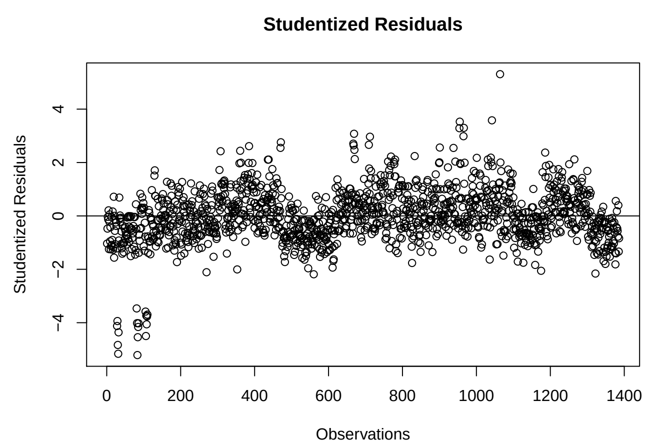


Figure 18: Studentized Residuals

Lastly, a check on normality of the data was performed by inspecting the Q-Q plot. The deviation of points from the line indicated non-normality. A Shapiro-Wilk test on the residuals returned a p-value much smaller than our 5% significance level, so we rejected the null hypothesis that the data are normally distributed. Therefore, this model was also invalid.

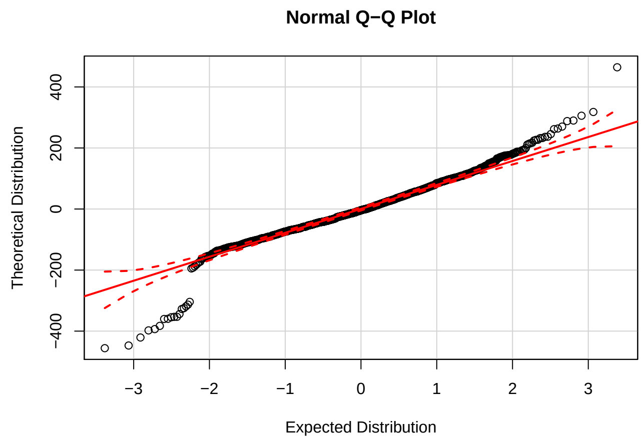


Figure 19: Q-Q Plot

W	p-value
0.93	0.00

Table 12

4. Transformed "Endcap" Model

Having an invalid model from previous section, applying transformation seemed to be an appropriate approach to take. We used the `boxcox()` function to find the best lambda (0.22 in this case) and then regressed the transformed units sold variable on average retail price (X_1), Endcap (X_3) and the interaction term $X_1 : X_3$ (Table X).

$R^2 = 0.2201$, F-stat = 130, P-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	6.2316	0.1911	32.61	0.0000
Average.Retail.Price	0.0654	0.0462	1.42	0.1568
Endcap	3.9273	0.8235	4.77	0.0000
Average.Retail.Price:Endcap	-0.4528	0.2062	-2.20	0.0282

Table 13

After the transformation, R^2 decreases by 0.1354. The model remained significant along with all coefficients except b_1.

A series of tests were the performed: $\chi^2_{BP} = 271.25$, with a p-value$_{BP} < 2.2 \times 10^{-16}$. Since the p-value$_{BP}$ is less than the 5% significance level, we reject the null hypothesis that the variance is constant. The heteroskedasticity is also confirmed by the residual plot which displays non-constant variance.

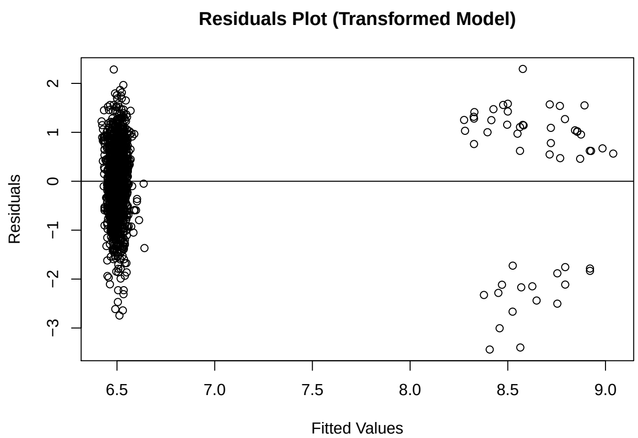


Figure 20: Fitted vs. Residuals

BP	p-value
271.25	0.00

Table 14

Additionally, after looking at the Q-Q plot, the points deviated from the line, meaning non-normality of the residuals. Performing a Shapiro-Wilk normality test on the transformed data resulted in a p-value < 0.05; we then reject the null hypothesis that the residuals are normal in this updated model, which confirmed our conclusion from the Q-Q plot.

W	p-value
0.99	0.00

Table 15

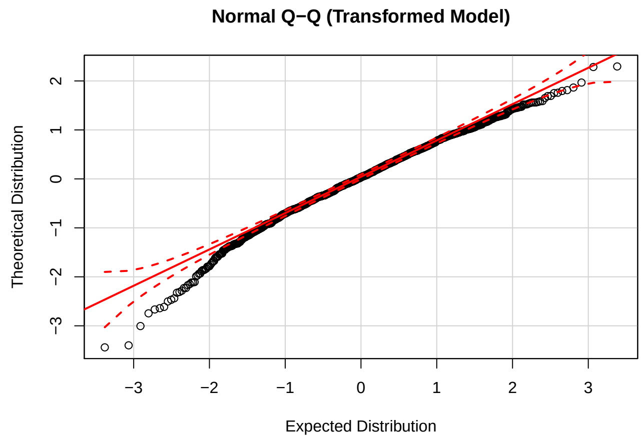


Figure 21: Q-Q Plot

However, the sequence plot did not display distinct patterns in the points, proving independence of the variables in this model (Figure 22). Moreover, the studentized residual plot still displayed outliers (Figure 23).

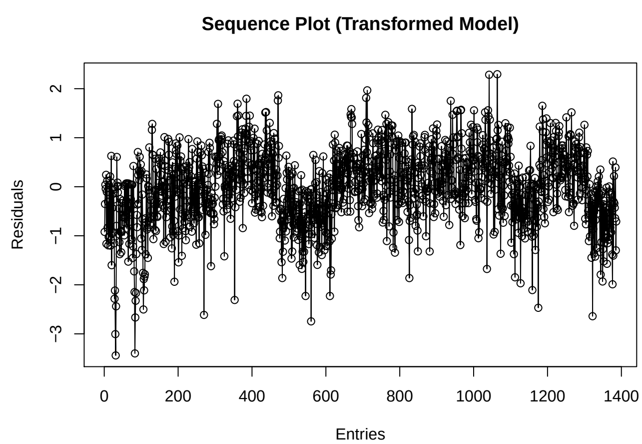


Figure 22: Residuals Sequence Plot

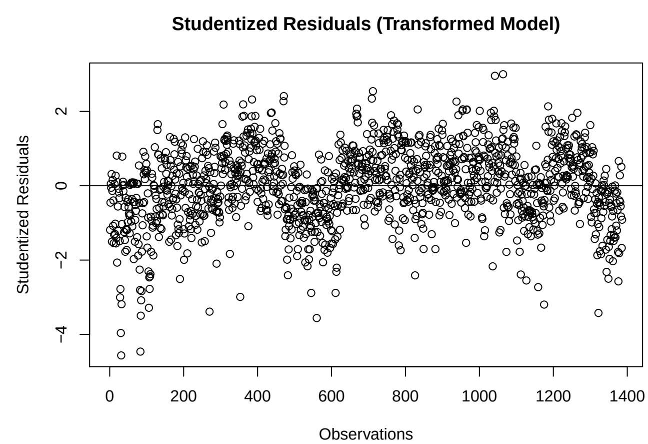


Figure 23: Studentized Residuals

5. "Demo" Model

Focusing on the "Endcap" variable did not lead to a conclusion, so we turned to the other variable of interest: "Demo." In this model, we regressed units sold (Y) on average retail price (X_1), Demo (X_4), Demo1-3 (X_5), Demo4-5 (X_6) and all the subsequent interaction terms.

$R^2 = 0.3002$, F-stat = 59, p-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	222.7911	26.8110	8.31	0.0000
Average.Retail.Price	-0.3727	6.5080	-0.06	0.9543
Demo	362.3445	96.1193	3.77	0.0002
Demo1.3	285.8067	56.5713	5.05	0.0000
Demo4.5	65.7833	78.4657	0.84	0.4020
Average.Retail.Price:Demo	-53.1669	22.8052	-2.33	0.0199
Average.Retail.Price:Demo1.3	-42.5808	13.7880	-3.09	0.0021
Average.Retail.Price:Demo4.5	4.8496	18.4116	0.26	0.7923
Demo:Demo1.3	48.2796	27.3053	1.77	0.0773
Demo:Demo4.5	65.3613	46.9422	1.39	0.1640
Demo1.3:Demo4.5	-31.0598	26.1722	-1.19	0.2355

Table 16

In this model, while the model was significant with a relatively high F-statistic (59), it had a relatively low R^2 of 0.3002. Among the coefficients, b_1, b_4, b_7, b_8, b_9 and b_{10} were insignificant with p-values of 0.9543, 0.4020, 0.7923, 0.0773, 0.164 and 0.2355 respectively. These were all greater than our 5% level of significance, but all other coefficients were significant. The coefficient for Demo (362.3445) and the coefficient for the interaction term (-53.1559) indicate that having demo activities in a given week would change units sold by $362.3445 - 53.1559 \times$ Average Retail Price, on average. The coefficient for Demo1-3 (285.8067) and the coefficient for the interaction term (-42.5808) indicates that having demo activities in one to three weeks ago will change units sold by $285.8067 - 42.5808 \times$ Average Retail Price. However, these conclusions could not be made until we confirmed the model was in fact valid.

A routine inspection of the model was performed. We plotted the fitted values against residuals, and the plot revealed heteroskedasticity in the data (Figure 24). A Breusch-Pagan test for heteroskedasticity confirmed this suspicion: $\chi^2_{BP} = 34.56$, with a p-value < our chosen 5% level of significance; thus, we reject the null hypothesis that the variance is constant.

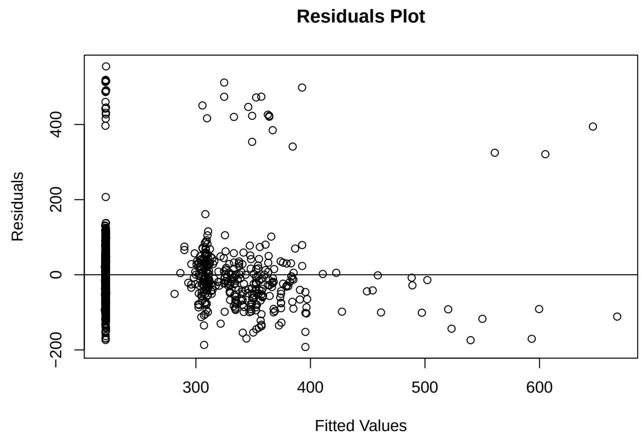


Figure 24: Fitted vs. Residuals

BP	p-value
34.56	0.00

Table 17

Furthermore, the studentized residual plot showed outliers in the data. However, a sequence plot of the residuals revealed no distinct pattern within the plot, which indicates that the residuals are independent for this case (Figure 25).

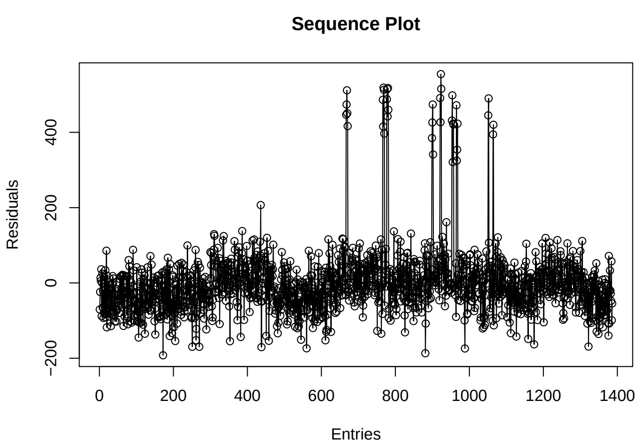


Figure 25: Residuals Sequence Plot

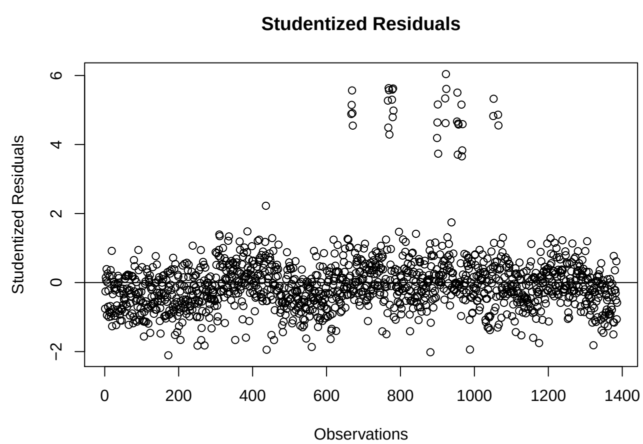


Figure 26: Studentized Residuals

Lastly, a check on normality of the data was performed by inspecting the Q-Q plot. The deviation of points from the line indicated non-normality. A Shapiro-Wilk test on the residuals returned a p-value much smaller than our 5% significance level, so we reject the null hypothesis that the data are normally distributed. Therefore, this model was also deemed invalid.

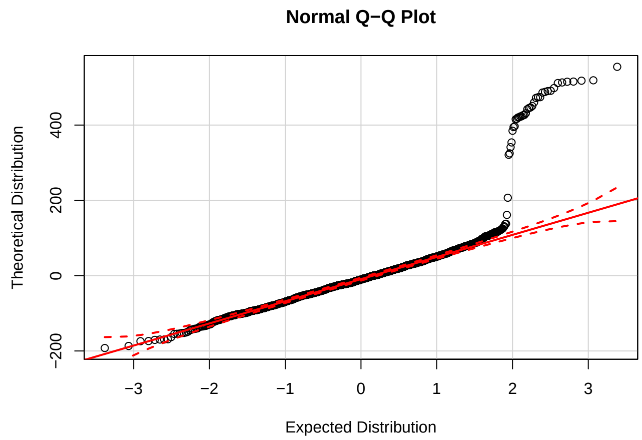


Figure 27: Q-Q Plot

W	p-value
0.75	0.00

Table 18

6. Backward Stepwise AIC Search

After all the inconclusive models produced, we turned to a search algorithm to find an optimized model. Using AIC as our metric, we used the `stepAIC` function in R, searching by adding and removing variables on the full model with all interaction terms. The best model is shown in the following table:

$R^2 = 0.8071$, F-stat = 478.6, p-value $< 2.2 \times 10^{-16}$

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	284.6488	12.9624	21.96	0.0000
Average.Retail.Price	-22.9761	3.0641	-7.50	0.0000
Sales.Rep	52.1411	4.9215	10.59	0.0000
Endcap	2.1593	12.1369	0.18	0.8588
Demo	102.1490	6.3823	16.00	0.0000
Demo1.3	71.7237	3.9095	18.35	0.0000
Demo4.5	62.6585	9.2031	6.81	0.0000
Natural	-2.3592	1.9081	-1.24	0.2165
Fitness	-0.4243	0.8808	-0.48	0.6300
Sales.Rep:Endcap	450.0663	14.8417	30.32	0.0000
Sales.Rep:Natural	5.3033	2.7917	1.90	0.0577
Demo:Demo1.3	23.4603	13.9437	1.68	0.0927
Demo4.5:Fitness	4.8149	3.2034	1.50	0.1331

Table 19

The model is significant with a relatively high F-statistic of 478.6 and a p-value less than 2.2×10^{-16}. It also has a high R^2 of 0.8071, meaning 80.71% of the variation of Y (units sold) is explained by this model. Among the coefficients, b_1, b_4, b_7, b_8, b_9 and b_{10} are insignificant with p-values of 0.9543, 0.4020, 0.7923, 0.0773, 0.164 and 0.2355, respectively, all of which are greater than our 5% level of significance. All other coefficients are significant. The coefficient for Demo (102.149) indicates that having demo activities in the current week will, on average, increase units sold by 104.149 units, holding everything else constant. The coefficient for Demo1-3 (71.7237) indicates that having demo activities within one to three weeks prior will, on average, increase units sold by 71.7237 units, holding everything else constant. The coefficient for Demo4-5 (62.6585) indicates that having demo activities within four to five weeks ago will, on average, increase units sold by 62.6585 units, holding everything else constant. The coefficient for the interaction term between Sales Rep and Endcap (450.0663) indicates that having endcap activities and a regional sales representative will on average increase units sold by 450.0663 units, holding all else consant, while having endcap activities *without* regional sales representative will have no effect. Again, such conclusions cannot be made until we confirm the model is in fact valid.

A routine inspection of the model was performed. We plotted the fitted values against residuals, and the plot revealed homoskedasticity in the data (Figure 28). A Breusch-Pagan test for heteroskedasticity confirmed this finding: $\chi^2_{BP} = 14.84$, with a p-value = 0.25 which was greater than our chosen 5% level of significance; thus, we failed to reject the null hypothesis that the variance is constant. There was enough evidence to support the null hypthesis that the model has constant variance.

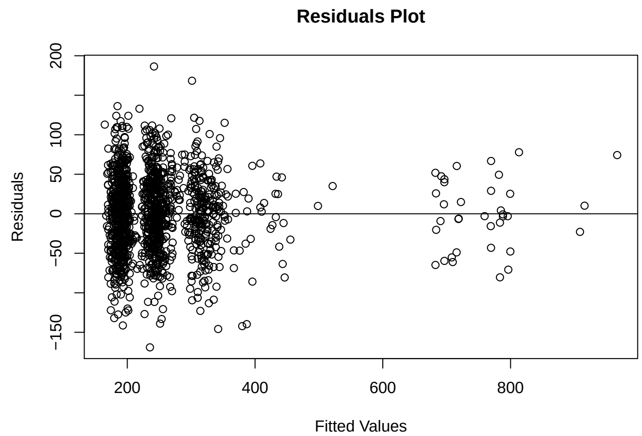


Figure 28: Fitted vs. Residuals

BP	p-value
14.84	0.25

Table 20

Furthermore, the studentized residual plot showed no outliers in the data and a sequence plot of the residuals shows no distinct pattern within the plot, which indicates that the residuals are independent for this case (Figure 29 and 30).

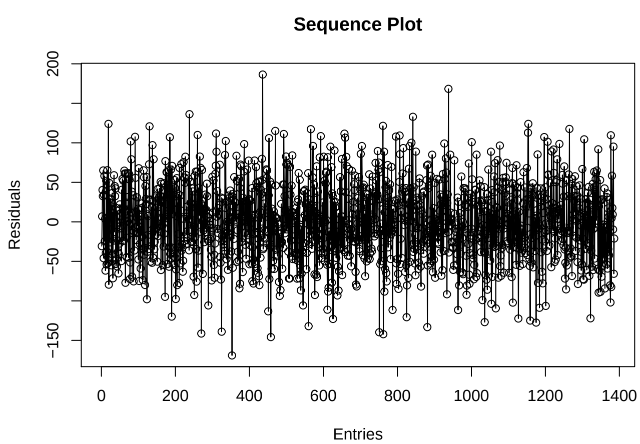


Figure 29: Residuals Sequence Plot

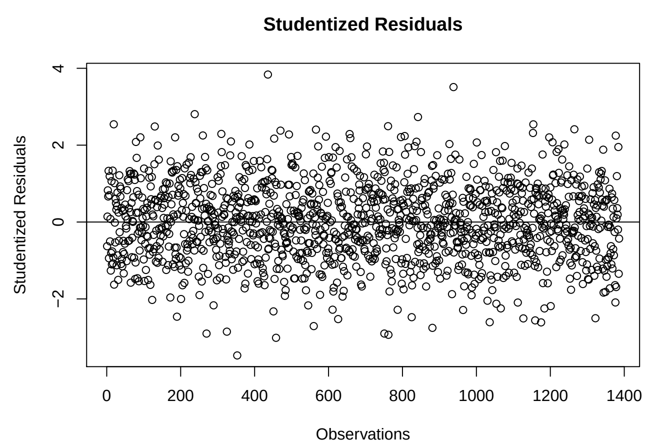


Figure 30: Studentized Residuals

Lastly, a check on normality of the data was performed by inspecting the Q-Q plot. The convergence of the points to the line indicates normality. A Shapiro-Wilk test on the residuals returned a p-value of 0.79, which was much greater than our 5% significance level, so we failed to reject the null hypothesis that the data are normally distributed. There was enough evidence to support the null hypothesis that the residuals are normally distributed. Therefore, we had a valid model and our analysis on the coefficients were in fact valid.

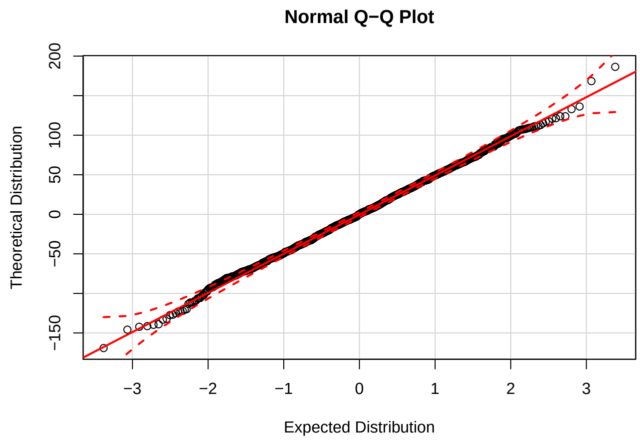


Figure 31: Q-Q Plot

W	p-value
1.00	0.79

Table 21