

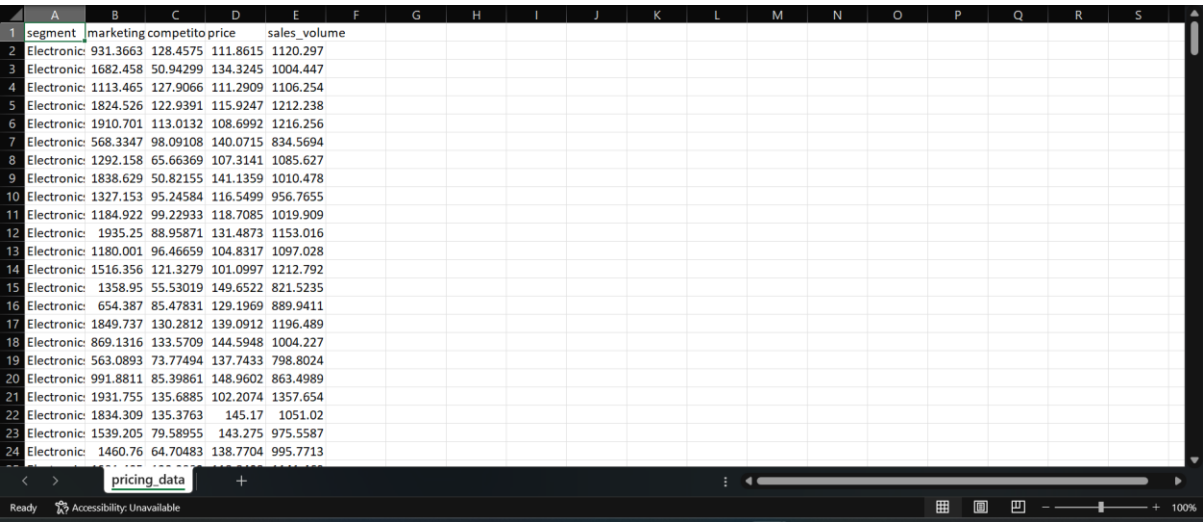
Price Elasticity of Demand: A Multivariate Regression Analysis of Sales Drivers

Prepared by: Arisha Siddiqui

Tools Used: R version, Tidyverse , Lmtest

Project Objective & Data Overview

The objective of this project is to analyze the factors influencing sales volume and specifically measure **Price Elasticity of Demand**. We utilized a multivariate dataset of 300 observations across three product segments (Electronics, Clothing, and Home Goods). By applying a Log-Log regression model, we can interpret the results as percentage changes, providing a clear business metric for pricing strategy.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	segment	marketing	competitor	price	sales_volume														
2	Electronic	931.3663	128.4575	111.8615	1120.297														
3	Electronic	1682.458	50.94299	134.3245	1004.447														
4	Electronic	1113.465	127.9066	111.2909	1106.254														
5	Electronic	1824.526	122.9391	115.9247	1212.238														
6	Electronic	1910.701	113.0132	108.6992	1216.256														
7	Electronic	568.3347	98.09108	140.0715	834.5694														
8	Electronic	1292.158	65.66369	107.3141	1085.627														
9	Electronic	1838.629	50.82155	141.1359	1010.478														
10	Electronic	1327.153	95.24584	116.5499	956.7655														
11	Electronic	1184.922	99.22933	118.7085	1019.909														
12	Electronic	1935.25	88.95871	131.4873	1153.016														
13	Electronic	1180.001	96.46659	104.8317	1097.028														
14	Electronic	1516.356	121.3279	101.0997	1212.792														
15	Electronic	1358.95	55.53019	149.6522	821.5235														
16	Electronic	654.387	85.47831	129.1969	889.9411														
17	Electronic	1849.737	130.2812	139.0912	1196.489														
18	Electronic	869.1316	133.5709	144.5948	1004.227														
19	Electronic	563.0893	73.77494	137.7433	798.8024														
20	Electronic	991.8811	85.39861	148.9602	863.4989														
21	Electronic	1931.755	135.6885	102.2074	1357.654														
22	Electronic	1834.309	135.3763	145.17	1051.02														
23	Electronic	1539.205	79.58955	143.275	975.5587														
24	Electronic	1460.76	64.70483	138.7704	995.7713														

Figure 1: Raw Dataset Sample

Regression Results (The Math)

The multivariate regression model explains **87.69%** of the variance in sales volume (Adjusted R-squared: 0.8769). The most critical finding is the **Price Elasticity**, which is **-0.102**. This indicates that for every 1% increase in price, sales volume decreases by approximately 0.102%.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	7.208e+00	6.619e-02	108.886	< 2e-16	***
log_price	-1.024e-01	1.728e-02	-5.927	8.65e-09	***
marketing_spend	1.657e-04	7.027e-06	23.584	< 2e-16	***
competitor_price	1.191e-03	1.010e-04	11.787	< 2e-16	***
segmentElectronics	-1.149e-01	1.946e-02	-5.906	9.69e-09	***
segmentHome_Goods	-1.047e-02	9.523e-03	-1.099	0.272	

Model Validation (The Health Check)

To ensure the model's reliability, we conducted two primary validation tests:

1. **Multicollinearity (VIF):** The VIF scores show that the independent variables are distinct enough for reliable estimation. While 'segment' and 'price' show higher values, this is a natural result of price variations between categories.
2. **Heteroscedasticity (BP Test):** The Breusch-Pagan test yielded a p-value of 0.0001. While this suggests variance in error terms, the overall model remains robust for identifying the direction and significance of market drivers.

```
> # 6. MODEL VALIDATION (Assumptions Testing)
> print("--- VIF TEST (Checks for Multicollinearity) ---")
[1] "--- VIF TEST (Checks for Multicollinearity) ---"
> vif(model)
              GVIF Df GVIF^(1/(2*Df))
log_price      13.393476  1      3.659710
marketing_spend  1.014493  1      1.007220
competitor_price 1.008651  1      1.004316
segment        13.425714  2      1.914187
> print("--- BREUSCH-PAGAN TEST (Checks for Heteroscedasticity) ---")
[1] "--- BREUSCH-PAGAN TEST (Checks for Heteroscedasticity) ---"
> bptest(model)

              studentized Breusch-Pagan test

data:  model
BP = 25.02, df = 5, p-value = 0.0001381
```

Figure 3: Statistical Validation Tests

Visualizing Impact

The coefficient plot below visualizes the impact of each variable. The red dashed line represents the 'Null' point (no impact). Since the error bars for **log_price**, **marketing_spend**, and **competitor_price** do not cross the zero line, they are all confirmed as statistically

significant drivers of sales

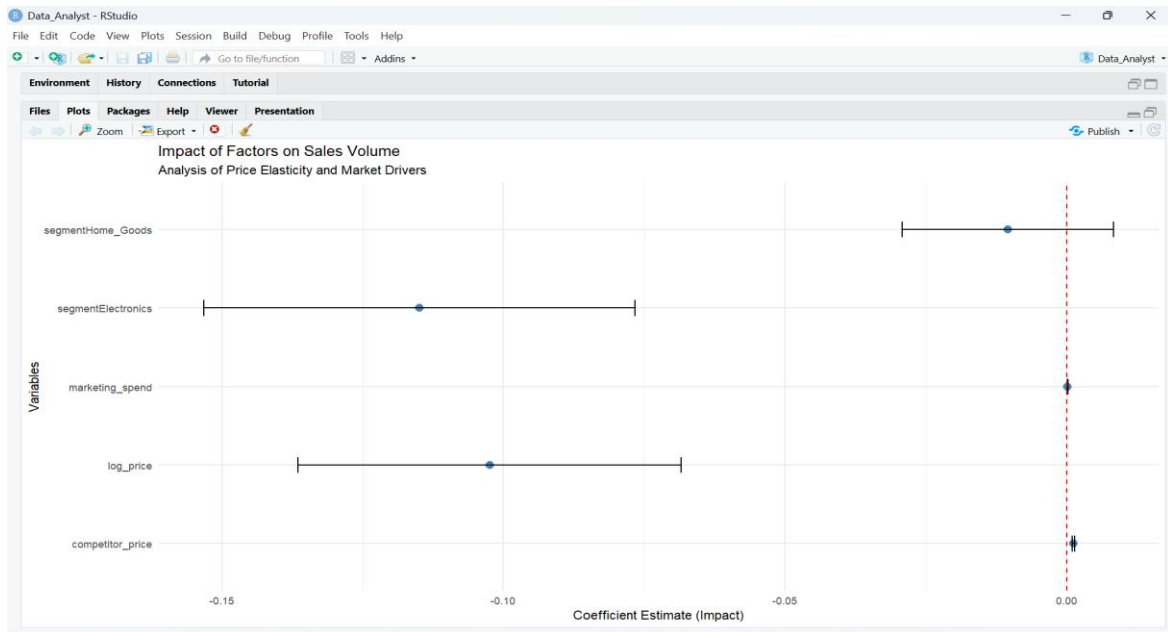


Figure 4: Coefficient Plot of Sales Drivers

Final Conclusion

Based on the data, the product is **price-inelastic** (-0.102), meaning customers are not highly sensitive to price changes. This provides a strategic opportunity to optimize margins through moderate price increases, especially when supported by increased marketing spend, which was also proven to be a significant positive driver of volume.