Table of Contents

CASE STUDY: PRICE ELASTICTY

Hi Guys,

Recently, I worked on a Capstone project as a part of my Executive Program in Business Analytics offered by SDA Bocconi partnered by Jigsaw Academy. I would like to share my work along with the insights which I was able to draw from the given data set. Hoping to get some feedback/views on my attempt so that moving ahead I can improve on my technique and approach to solve business problems.

Please Note: Due to Non-Disclosure Agreement, I am using dummy variable names.

Project Synopsis

Problem Statement:

To identify the pricing dynamics for a leading household cleaning supplies brand in US market.

- Create a robust imputation logic where price data is missing.
- Is the price elasticity increasing or decreasing over time?
- What is the impact of key competitor's pricing on the target brand?

Available Data:

- I had data for a particular retailer (say Retailer 1).
- Retailer had data for different Product Styles.Product Styles are basically different SKUs of a Brand / Product.
- My Target Product Style was "Brand T" with SKUs T1, T2, T3 and T4.
- Price Column was missing and needed to be derived using Amount Sold and Units Sold column.

Customer VS Product Styles Table:

Product Styles →	T1	T2	T3	T4	S1	S2	C1	C2	C3
Retailer 1	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ Indicates availability

- **Target Brand SKU** : T1, T2, T3 and T4
- **Competitor Brand SKU** : S1, S2, C1, C2 and C3

Scope:

I had to build a log-log linear regression model for the retailer to come up with a price elasticity equation for each SKU of Brand T.*

What is Price Elasticity?

Price Elasticity is a measure of the relationship between a change in the quantity demanded of a particular good and a change in its price.

The formula for calculating price elasticity of demand is:

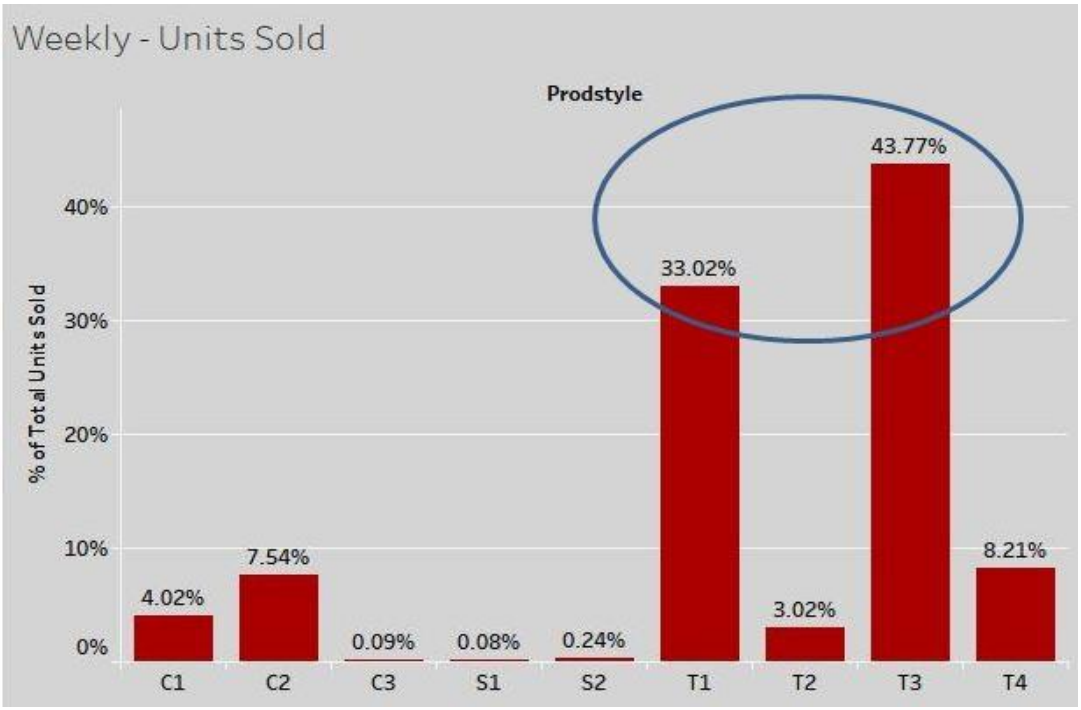
$$\text{Price Elasticity of Demand} = \frac{\%\ \text{change in Quantity Demanded}}{\%\ \text{Change in Price}}$$

For more details please check out my article on LinkedIN.

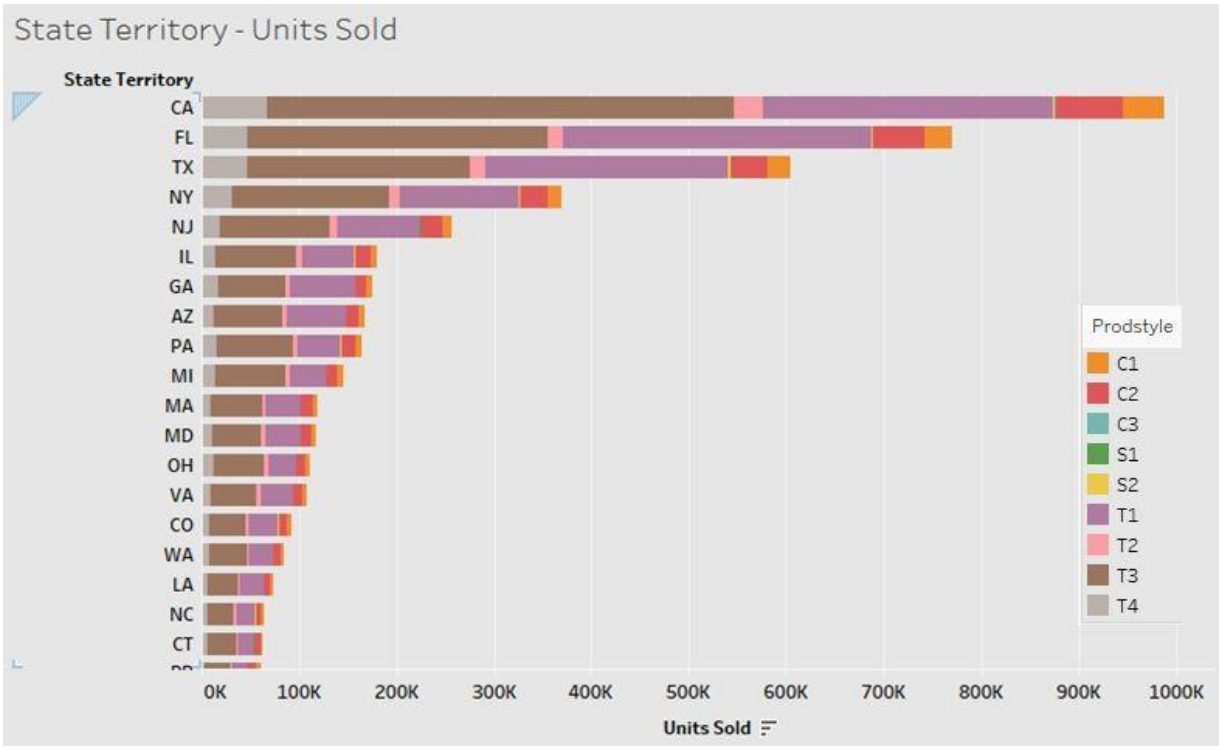
Executive Summary

Inferences:

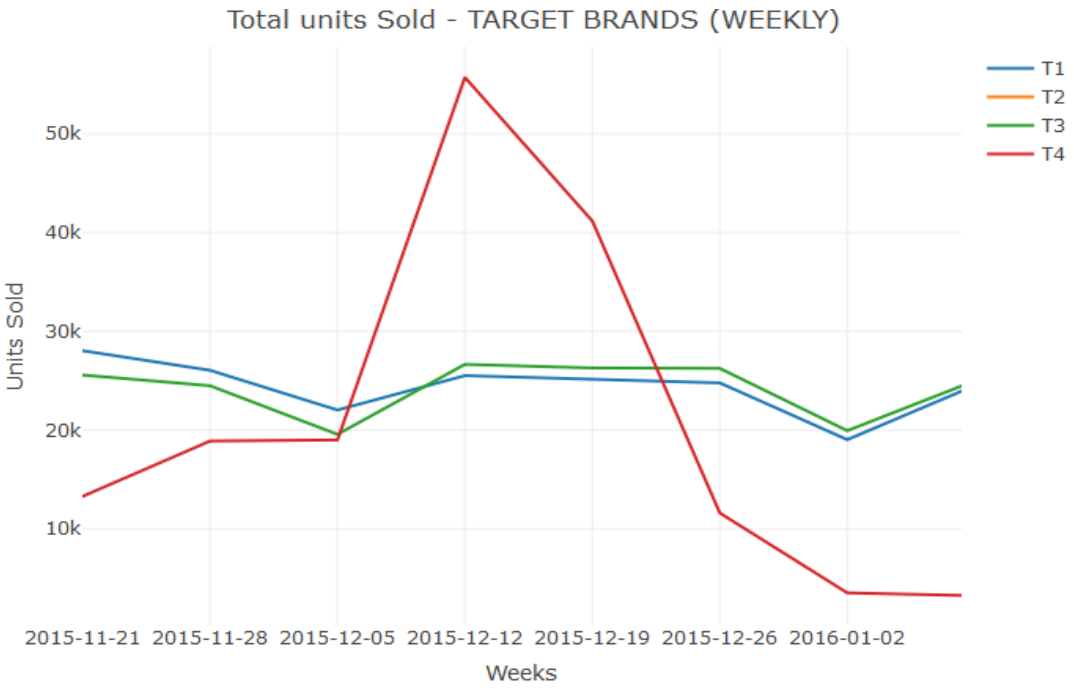
- Out of all the Product Style SKU, ***T1*** and ***T3*** contributes to ***76.79%*** of the total units sold.

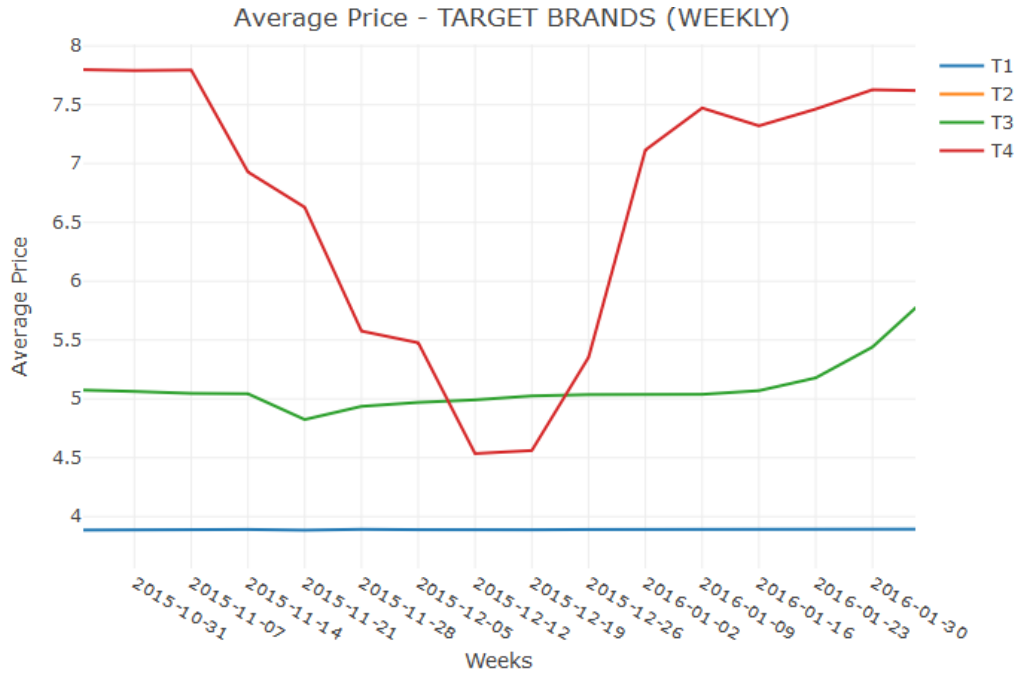


- A similar behavior is observed across all states w.r.t. Units Sold of each SKU.
- The State Territory ***TX*** seems to be a major consumer of ***T1*** and ***T3*** Product Style. Based on this behavior, it seems State Territory won't have much impact on the regression model.



- The Units Sold for Product Style SKU ***T4*** is highly elastic towards its Price.
- It is observed that the Units Sold for ***T4*** went up as the average price went down during the weeks ***05th Dec 2015 – 26th Dec 2015***. Similarly, the units sold dropped as the price of ***T4*** was increased later on.





Inferences (Retailer 1)

The major predictors for the unit sales of T1 Brand

Change in Average Price of a Brand	Impact	Coefficient Estimate
T1	Negative	- 5.75
S2	Positive	0.11
C1	Positive	8.59
C2	Positive	19.11
C3	Negative	- 0.01

R2 = 68.51%

This model explains **68.51%** of the impact on **T1** brand's units sold with respect to change in the price of T1, S2, C1, C2 and C3 brands

The major predictors for the unit sales of T2 Brand

Change in Average Price of a Brand	Impact	Coefficient Estimate
T2	Positive	0.36
T3	Positive	9.34
T4	Negative	- 0.28

R2 = 91.13%

This model explains **91.13%** of the impact on **T2** brand's units sold with respect to change in the price of T2, T3 and T4 brands

Inferences (Retailer 1)

The major predictors for the unit sales of T3 Brand

Change in Average Price of a Brand	Impact	Coefficient Estimate
T3	Positive	2.36
T1	Positive	29.36
T4	Negative	- 0.02
C1	Positive	21.21

R2 = 71.59%

This model explains **71.59%** of the impact on **T3** brand's units sold with respect to change in the price of T3, T1, T4 and C1 brands

The major predictors for the unit sales of T4 Brand

Change in Average Price of a Brand	Impact	Coefficient Estimate
T3	Negative	- 6.63
T4	Negative	- 2.67
S1	Negative	- 4.84

R2 = 62.69%

This model explains **62.69%** of the impact on **T4** brand's units sold with respect to change in the price of T3, T4 and S1 brands.

Recommendation:

The below recommendations majorly highlights each Target brand SKU is relatively highly elastic to which competitor Brand's SKU.

Recommendation (Retailer 1)

- The units sold of T1 can be increased by 8.59% by increasing the price of C1 by 1%.
  The units sold of T1 can be increased by 19.11% by increasing the price of C2 by 1%.
- The units sold of T2 can be increased by 9.34% by increasing the price of T3 by 1%.
- The units sold of T3 can be increased by 21.21% by increasing the price of C1 by 1%.
  The units sold of T3 can be increased by 29.36% by increasing the price of T1 by 1%.
- Looking at the above inference tables, Target brand SKU - T4 has negative impact on other Target brand SKU - T2 and T3.
  T4 may be a possible substitute for T2 and T3 brand SKU.
- Overall, the regression models seems okay and the results can be considered for further actions.
- However, more weekly records would be needed to have a more accurate depiction of Price Elasticity for WD brands.

Price Elasticity Tool in Microsoft Excel

Price Elasticity Tool in MS EXCEL - Capabilities & Features

- ✓ The Tool enables the customer to estimate the % change in unit sold based on old and new average price entered for Target Brand SKU and it's respective Predictor Brand SKU (Competitor Brand SKU).
- ✓ Separate Table is maintained for each TARGET Brand SKU.
- ✓ User needs to enter the values of old price as well as new price in the cells highlighted in "Dark Grey".
- ✓ If the % change in Unit sold is greater than 0 then the value is highlighted in "Green".
- ✓ If the % change in Unit sold is less than 0 then the value is highlighted in "Red".
- ✓ If the % change in Unit sold is equal to 0 then the value is highlighted in "Yellow".
- ✓ The tool also provides an estimated value for overall % change in unit sold for each Target SKU Brand.

- *Below is the screenshot of the Elasticity Tool with sample output.*

Retailer 1

Elasticity Tool - To calculate percentage change in Units Sold

Target Brand	Predictors				
T1	T1	S2	C1	C2	C3
Old Price	3.87726	5.85537	5.98663	5.93316	0.00001
New Price	3.88418	5.95146	6.0508	5.93507	0.00001
% Change in Price	-0.17864	-1.64117	-1.07199	-0.03214	0
% Change in Unit Sold	1.02748	-0.17407	-9.21895	-0.61447	0
Overall % Change in Unit Sold	-8.980013314				

Target Brand	Predictors		
T2	T2	T3	T4
Old Price	7.81269	7.07315	0.00001
New Price	7.80902	6.97315	0.00001
% Change in Price	0.04695	1.41372	0
% Change in Unit Sold	0.01731	13.2054	0
Overall % Change in Unit Sold	13.22268494		

Target Brand	Predictors			
T3	T3	T1	T4	C1
Old Price	5.50334	3.88873	7.83446	6.35927
New Price	5.18966	3.8856	7.81315	6.36696
% Change in Price	5.69975	0.08059	0.27203	-0.12084
% Change in Unit Sold	13.4833	2.36609	-0.00589	-2.56322
Overall % Change in Unit Sold	13.28026341			

Target Brand	Predictors		
T4	T3	T4	S1
Old Price	5.50334	7.83446	27.7842
New Price	5.18966	7.81315	27.1253
% Change in Price	5.69975	0.27203	2.37148
% Change in Unit Sold	-37.8178	-0.72795	-11.4898
Overall % Change in Unit Sold	-50.03559989		

% Change in Price	(Difference between Old and New Price Value / Old Value) X 100
% Change in Unit Sold	Percentage Change X Price Elasticity Coefficient of Regression Model
Overall % Change in Unit Sold	Overall Percentage Change in Unit sold when we sum up all the "%Change in Unit Sold" values w.r.t. each predictors.

EPBA4 - Rohan Shetty

Procedure/Methodology

The below section is more of a drill down on how data was prepared, analyzed and details on the regression model.

1. Data Preparation
2. Preliminary Data Analysis
3. Regression Model
4. Assumption Checks / Prediction

Data Preparation:

The Raw data has negative and zero values for Amount Sold and Units Sold columns. The number of records with these values is very less (3,742) compared to ***7,68,294 records*** in the data set. Hence , these rows were filtered out.

Price Column was missing hence imputed it as ***Amount Sold/Units Sold***.

Initially the data was grouped and summarized to obtain average Price and sum(Units Sold) at store week level. However, the resultant values were not significant and hence store column was dropped.

A similar behavior is observed across all states w.r.t. Units Sold of each SKU except one state territory "TX". Hence State Territory column was also dropped.

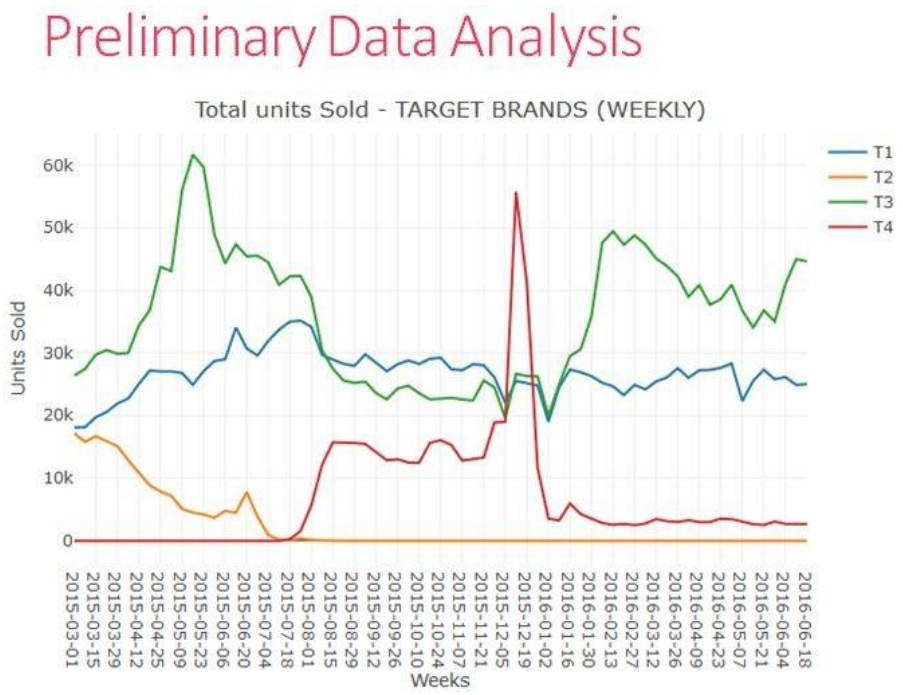
The entire data set was transposed at week level consisting of a week column and average Price , sum(Units Sold) columns for each SKU of all available Product Styles i.e. ***T**** , ***S**** and ***C****.

The newly transposed data set has ***69 records***. Each record corresponds to data for one week.

There were few records with zero values for ***T2***, ***T4*** and ***C3*** product style SKU.

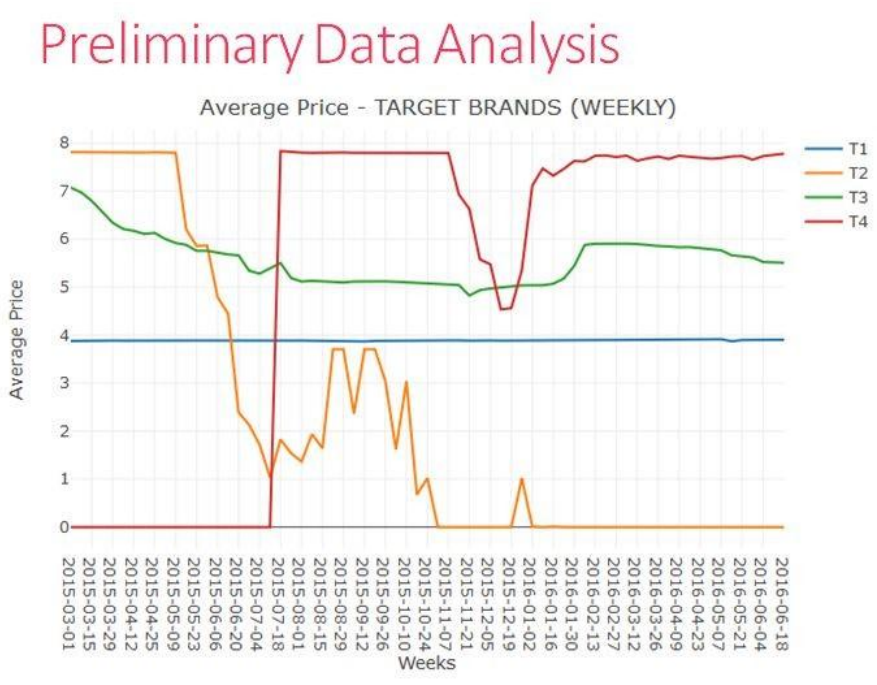
Since we were about to perform ***LOG-LOG Linear Regression model*** and LOG(0) gives “INF” error we substituted Zero value with a very insignificant value i.e. 0.00001.

Preliminary Data Analysis:



Data observations

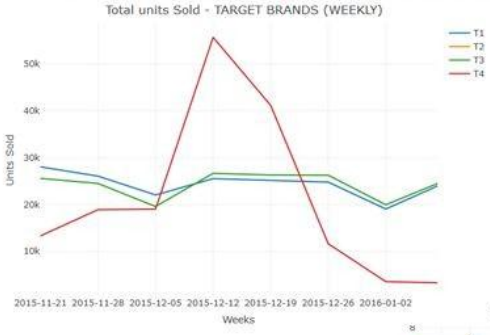
- The units sold of T2 brand seems to be zero post 23rd January 2016
- T4 seems to be newly launched or available for sale since July 2015. Also shows a huge peak in its units sold during Dec 2015 but then drops down considerably.
- T3 also shows 2 peaks during the month of May 2015 and February 2016



Data observations

- The price of t2 has been falling continuously and it seems has been discontinued since January 2016.
- There was a considerable dip in the average price of T4 during the month of Dec 2015 but since then it has been consistent.
- T1 has been consistent throughout all the weeks.
- There isn't much significant change in the price of T3

Preliminary Data Analysis



Data observations

- A closer look at the T4 brand's trend.
- It is pretty clear that as the Average Price of T4 brand went down during the week 05th Dec 2015 – 26th Dec 2015 and the Units Sold went up.
- T4 seems to be highly elastic to its price.

Regression Model and Prediction:

Regression Model

- A Price Elasticity model has to be built for each SKU of target Brand T.
- In this case, four such models has to be built. One for each SKU (**T1**, **T2**, **T3** and **T4**).
- The **LOG-LOG Linear Regression model** will directly gives us the coefficients of Price Elasticity.

Variables	Description	Variables	Description
EndDate	Start Date of a Week	**Price_mean_S1**	Average Price of S1 (weekly)
UnitsSold_sum_T1	Total Sales of T1 (weekly)	**Price_mean_S2**	Average Price of S2 (weekly)
UnitsSold_sum_T2	Total Sales of T2 (weekly)	**Price_mean_C1**	Average Price of C1 (weekly)
UnitsSold_sum_T3	Total Sales of T3 (weekly)	**Price_mean_C2**	Average Price of C2 (weekly)
UnitsSold_sum_T4	Total Sales of T4 (weekly)	**Price_mean_C3**	Average Price of C3 (weekly)
Price_mean_T1	Average Price of T1 (weekly)	**Price_mean_T3**	Average Price of T3 (weekly)
Price_mean_T2	Average Price of T2 (weekly)	**Price_mean_T4**	Average Price of T4 (weekly)

Model (Target Brand - T1)

Regression Model

Final Model (BRAND : T1)

LN(UnitsSold_sum_T1) = - 32.2012 - 5.7518*LN(Price_mean_T1) + 0.1060*LN(Price_mean_S2)
+ 8.5998*LN(Price_mean_C1) + 19.1160*LN(Price_mean_C2)
- 0.0092*LN(Price_mean_C3)

Label	Description	Coefficient Estimate
Price_mean_T1	Average Price of T1 (weekly)	- 5.75
Price_mean_S2	Average Price of S2 (weekly)	0.11
Price_mean_C1	Average Price of C1 (weekly)	8.59
Price_mean_C2	Average Price of C2 (weekly)	19.11
Price_mean_C3	Average Price of C3 (weekly)	- 0.01

R2 = 68.51%

This model explains **68.51%** of the impact on T1 brand's units sold with respect to change in the price of T1, S2, C1, C2 and C3 brands

Regression Model

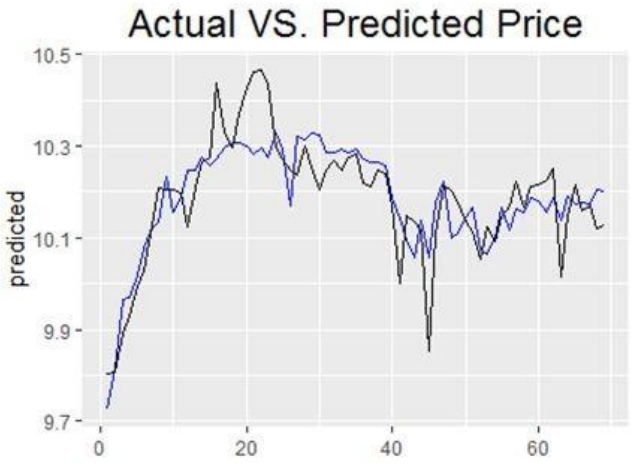
Interpretation

Price Elasticity of Demand for T1 brand would be how the price of T1, S2 , C1 , C2 and C3 impacts the units sold of T1.

Based on Model Equation:

- An increase in the price of T1 by 1% will decrease the units sold of T1 by **-5.75**%.
- An increase in the price of S2 by 1% will increase the units sold of T1 by **0.10**%.
- An increase in the price of C1 by 1% will increase the units sold of T1 by **8.59**%.
- An increase in the price of C2 by 1% will increase the units sold of T1 by **19.11**%.
- An increase in the price of C3 by 1% will decrease the units sold of T1 by **-0.01**%.

Prediction



Actual	Predicted	Error
9.80	9.72	0.07
9.80	9.81	-0.01
9.89	9.96	-0.07
9.93	9.96	-0.03
9.99	10.01	-0.02
10.03	10.07	-0.04

MAPE = 0.54%

Model (Target Brand - T2)

Regression Model

Final Model (BRAND : T2)

LN(UnitsSold_sum_T2) = - 11.9207 + 0.3686*LN(Price_mean_T2) + 9.3409*LN(Price_mean_T3)
- 0.2803*LN(Price_mean_T4)

Label	Description	Coefficient Estimate
Price_mean_T2	Average Price of T2 (weekly)	0.36
Price_mean_T3	Average Price of T3 (weekly)	9.34
Price_mean_T4	Average Price of T4 (weekly)	- 0.28

R2 = 91.13%

This model explains **91.13%** of the impact on T2 brand's units sold with respect to change in the price of T2, T3 and T4 brands

Regression Model

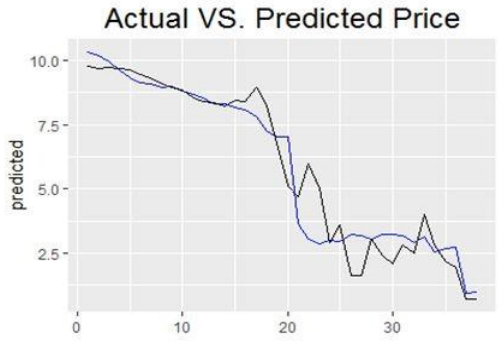
Interpretation

Price Elasticity of Demand for T2 brand would be how the price of T2, T3 and T4 impacts the units sold of T2.

Based on Model Equation:

- An increase in the price of T2 by 1% will increase the units sold of T2 by **0.37%.**
- An increase in the price of T3 by 1% will increase the units sold of T2 by **9.34%.**
- An increase in the price of T4 by 1% will decrease the units sold of T2 by **-0.28%.**

Prediction



Actual	Predicted	Error
9.75	10.33	-0.58
9.66	10.20	-0.53
9.72	9.96	-0.24
9.67	9.65	-0.01
9.62	9.31	0.30
9.46	9.12	0.33

MAPE= 19.28%

Model (Target Brand - T3)

Regression Model

Final Model (BRAND : T3)

LN(UnitsSold_sum_T3)= - 72.7017 + 2.3655*LN(Price_mean_T3)
+ 29.3608*LN(Price_mean_T1) - 0.0216*LN(Price_mean_T4)
+ 21.2122*LN(Price_mean_C1)

Label	Description	Coefficient Estimate
Price_mean_T3	Average Price of T3 (weekly)	2.36
Price_mean_T1	Average Price of T1 (weekly)	29.36
Price_mean_T4	Average Price of T4 (weekly)	- 0.02
Price_mean_C1	Average Price of C1 (weekly)	21.21

R2 = 71.59%

This model explains **71.59%** of the impact on T3 brand's units sold with respect to change in the price of T3, T1, T4 and C1 brands

Regression Model

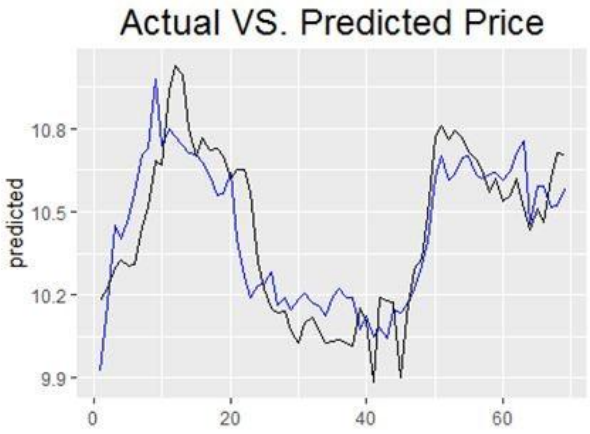
Interpretation

Price Elasticity of Demand for T3 brand would be how the price of T3, WD08, T4 and C1 impacts the units sold of T2.

Based on Model Equation:

- An increase in the price of T3 by 1% will increase the units sold of T3 by **2.36**%.
- An increase in the price of T1 by 1% will increase the units sold of T3 by **29.36**%.
- An increase in the price of T4 by 1% will decrease the units sold of T3 by **-0.02**%.
- An increase in the price of C1 by 1% will increase the units sold of T3 by **21.21**%.

Prediction



Actual	Predicted	Error
10.18	9.92	0.25
10.21	10.16	0.05
10.29	10.44	-0.14
10.32	10.40	-0.08
10.30	10.47	-0.17
10.30	10.57	-0.26

MAPE = 1.19%

Model (Target Brand - T4)

Regression Model

Final Model (BRAND : T4)

LN(UnitsSold_sum_T4) = 41.144 - 6.635*LN(Price_mean_T3) - 2.676*LN(Price_mean_T4)
- 4.845*LN(Price_mean_S1)

Label	Description	Coefficient Estimate
Price_mean_T3	Average Price of T3 (weekly)	- 6.63
Price_mean_T4	Average Price of T4 (weekly)	- 2.67
Price_mean_S1	Average Price of S1 (weekly)	- 4.84

R2 = 62.69%

This model explains **62. 69%** of the impact on T4 brand's units sold with respect to change in the price of T3, T4 and S1 brands.

Regression Model

Interpretation

Price Elasticity of Demand for T4 brand would be how the price of WD12, T4 and S1 impacts the units sold of T2

Based on Model Equation:

- An increase in the price of T3 by 1% will decrease the units sold of T4 by -**6.64**%.
- An increase in the price of T4 by 1% will decrease the units sold of T4 by -**2.68**%.
- An increase in the price of S1 by 1% will decrease the units sold of T4 by -**4.85**%.

Prediction

Actual	Predicted	Error
5.56	8.21	-2.64
7.34	8.72	-1.38
8.64	8.85	-0.20
9.40	8.84	0.56
9.66	9.91	-0.24
9.66	9.68	-0.01

MAPE = 4.85%

Hopefully you liked my attempt.

Any suggestions, comments, feedback are always welcome.

Please do like and share if you find it useful.

Thanks and Regards,
Rohan Shetty