

# Project Report

## Marketing Analytics for Mayonnaise Sales

### Group 8

Mayank Kumar

Rinda Sai Kiran Gunjala

Pooja Banthia

Monica Kulkarni

Vishal Taneja

Chinar Arora

## Contents

1. Exploratory Analysis
  - 1.1 Brand/Company Wise Distribution
  - 1.2 Geographical Distribution
  - 1.3 Price change and sales trend
2. Understanding effects of Promotions and Price reduction on Sales
  - 2.1 The Problem statement
  - 2.2 Approach
  - 2.3 Insights from Model Statistics
  - 2.4 Recommendation
3. Hypothesis Test
  - 3.1 Hypothesis: The sales of top brand vary across different volume equivalent categories
  - 3.2 Hypothesis: To test if larger stores have higher average price per unit as compared to smaller stores

## Introduction:

Objective: Apply statistical models to perform Marketing analytics on scanner data of mayonnaise data across grocery stores and provide recommendations for business growth.

## Tools Used:

- SAS for exploratory and statistical analysis
- Tableau and Microsoft Excel for visualization

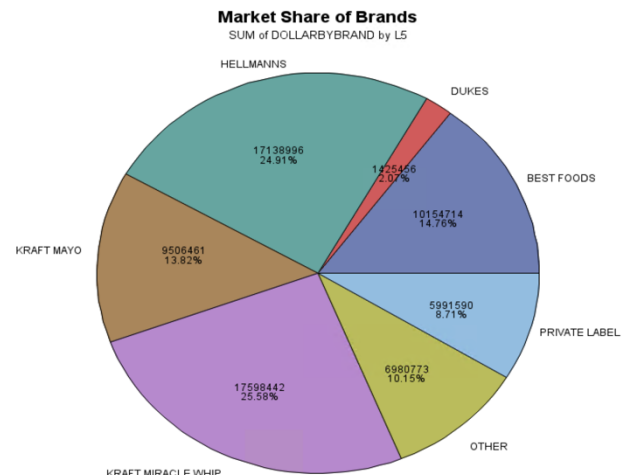
## Approach:

- Exploratory Analysis to study sales and price patterns
- Panel data analysis using two way Fixed effects to study effect of promotion for our brand(Kraft Mayo Whip) and competitors
- Hypothesis testing to test for 2 hypothesis

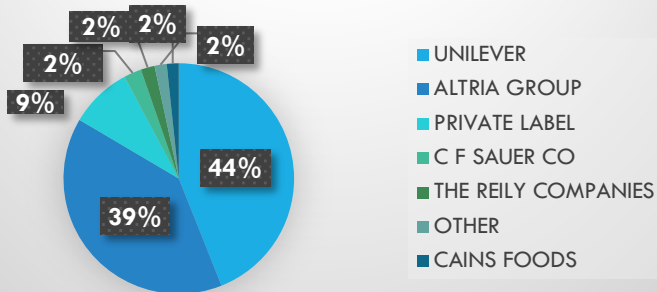
## 1. Descriptive Analysis

### 1.1 Brand/Company wise distribution

- Kraft Miracle Whip is the topmost brand with \$17.59 million in sales followed by Hellmanns which has 17.13 million sales.
- The major players in the market are Kraft Miracle Whip and Hellmanns with market share of 25.58% and 24.91% respectively.



### Market Share of Top Companies



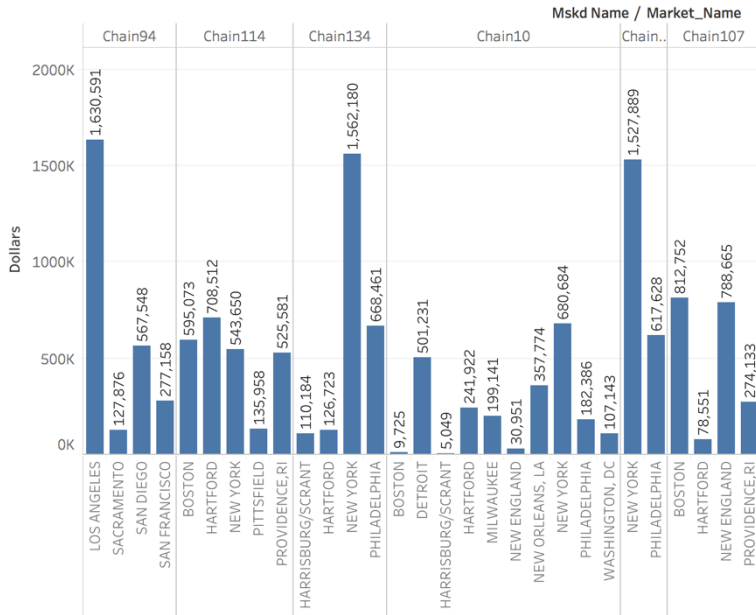
- Unilever is a major player with \$30.24 million in sales.
- Out of the top 6 brands, 2 are owned by Unilever and 2 are owned by Altria and the top brand i.e. Kraft Miracle Whip is owned by Altria Group

### 1.2 Geographical Distribution

- New York is the top region with 4.8 million sales followed by Los Angeles which has 4.7 million.
- New York and Los Angeles are top markets for Unilever and Altria Group where they compete strongly

NEW YORK	4,809,314
LOS ANGELES	4,798,609
SAN FRANCISCO	2,614,465
NEW ENGLAND	2,478,661
PHILADELPHIA	2,417,344
CHICAGO	2,227,690

Company	NEW YORK	LOS ANGELES	SAN FRANCISCO	NEW ENGLAND	PHILADELPHIA
UNILEVER	3,977,378	2,993,828	1,775,281	1,057,883	1,630,137
ALTRIA GROUP	475,328	1,396,621	509,482	645,874	539,948
PRIVATE LABEL	289,316	308,509	283,233	389,570	220,438

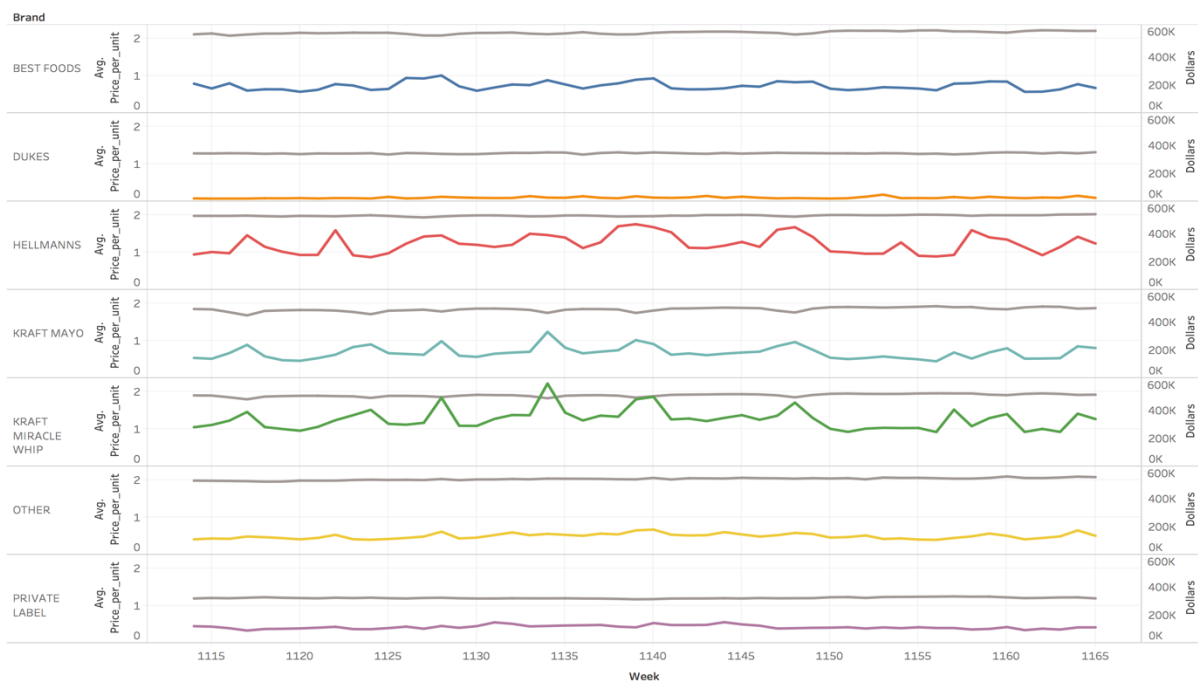


Chain 94 has the highest sales(2.6 million)in Los Angeles,Sacramento,San Diego and San Francisco followed by chain 114(2.5million) which has highest sales in Hartford,Boston, New York.

Chain94	2,603,174
Chain114	2,508,774
Chain134	2,467,549
Chain10	2,316,006
Chain89	2,145,516
Chain107	1,954,102
Chain124	1,914,186
Chain117	1,871,520
Chain75	1,840,004
Chain35	1,810,375

### 1.3 Price Change and Sales Trend

Price Change and Sales Trend



Measure Names	
Avg. Price_per_unit	
Brand	
BEST FOODS	
DUKES	
HELLMANN'S	
KRAFT MAYO	
KRAFT MIRACLE WHIP	
OTHER	
PRIVATE LABEL	

Brand	Avg. Price_per_unit	Dollars
KRAFT MIRACLE WHIP	2	17,598,442
HELLMANN'S	2	17,138,996
BEST FOODS	2	10,154,714
KRAFT MAYO	2	9,506,461
OTHER	2	6,980,773
PRIVATE LABEL	1	5,991,590
DUKES	1	1,425,456

- The average price per unit clearly divides the brands into premium(Average price=2) and marginal(average price=1) categories.
- Brands with higher average unit prices show an approximate cyclicity in their variations and seem to phase out their price variations.

## 2. Understanding effects of Promotions and Advertisements on Sales

We have done analysis keeping **Kraft Mayo Whip (KMW)** as our brand. Our competitors are Hellmanns (HM), Kraft Mayo (KM), Best Foods (BF), Private Label (PL), Dukes (DU) and Other (OTHER). In order to understand how marketing efforts (advertising and promotions) across the industry affect sales (or in this case of units sold), it is important to understand customer buying behavior patterns as different stores at different locations have different types of customers. These customers can be

- Smart Shoppers – Coupons – Who are always looking for the best deal.
- Lazy Customers – Who are not sensitive to promotions.

### a. Problem Statement-

To understand the unobserved Heterogeneity, different price sensitivities, different kinds of shoppers across time and brands, following questions need to be answered:

- What is the effect of price change on unit sales?
- How does display and feature impact unit sales?
- How does competitor pricing strategies/ promotions impact Kraft Mayo Whip's sales?

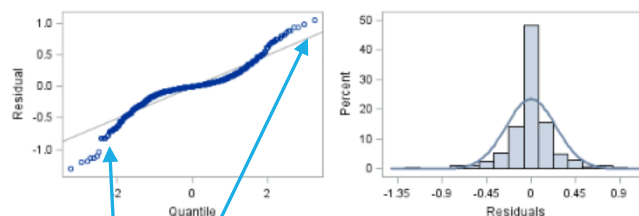
### b. Approach-

In order to account for these unobserved factors, Panel regression across both stores and weeks, is used to analyze the effect of cross-sectional variance of store and time on units sold for our brand Kraft Mayo Whip. The intent of this model is to measure the price sensitiveness of the customers and their behavior with different promotions and advertisements.

We have considered the log of dollar sales for our brand Kraft Mayo Whip as the dependent variable. Also, we have considered price per unit (PPU), price reduction flag (PR), display (Dnew) and interaction of average price reduction and display (avg\_dis\_up\_week) as independent variables. We found that the variable feature was highly insignificant, so we did not include it in the final model.

In order to test which model - Fixed effects or Random effects would be applicable, **Hausman Test** is conducted, which tests for correlation between intercept error term and model error term. Given that the p value close to 0 (significant), **Null Hypothesis is rejected** means, there is correlation observed between the error terms. So, in this case, **Fixed effects model is the applicable model to be used**. Fixed Two-way effects model is considered here to include effects of different brands and time.

Root MSE	1.04518	R-Square	0.0793
Dependent Mean	2.28465	Adj R-Sq	0.0792
Coeff Var	48.15190		



We obtain an R-square of 0.0793 which can be improved.

We see that from the residuals plot, there are some deviations towards the tails which can be improved.

We observe that the residuals are normally distributed.

### c. Insights from Model Statistics-

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	3.05919	0.01618	189.06	<.0001
KMW_PPU	1	-0.16264	0.00334	-48.63	<.0001
KMW_PR	1	0.04153	0.00558	7.45	<.0001
KMW_Dnew	1	0.26672	0.00837	31.87	<.0001
avg_up_dis_week	1	0.09176	0.00470	19.52	<.0001
HM_PPU	1	-0.16315	0.00349	-46.73	<.0001
HM_PR	1	-0.22451	0.00814	-27.57	<.0001
HM_Dnew	1	-0.09124	0.01374	-6.64	<.0001
avg_up_dis_week_HM	1	0.08367	0.00558	15.05	<.0001
KM_PPU	1	0.00912	0.00379	2.40	0.0163
KM_PR	1	0.03717	0.00496	7.49	<.0001
KM_Dnew	1	-0.03297	0.00948	-3.48	0.0005
avg_up_dis_week_KM	1	-0.24848	0.00534	-46.53	<.0001
BF_PPU	1	0.03392	0.00366	9.26	<.0001
BF_PR	1	0.01393	0.00877	1.59	0.1121
BF_Dnew	1	0.03109	0.01830	1.70	0.0894
avg_up_dis_week_BF	1	0.03403	0.00579	5.87	<.0001
DU_PPU	1	0.03937	0.00329	11.96	<.0001
DU_PR	1	0.07060	0.00746	9.47	<.0001
DU_Dnew	1	0.05728	0.02640	2.17	0.0300
avg_up_dis_week_DU	1	0.04401	0.00670	6.56	<.0001
PL_PPU	1	0.00181	0.00430	0.42	0.6746
PL_PR	1	0.01515	0.00399	3.80	0.0001
PL_Dnew	1	-0.05813	0.00941	-6.18	<.0001
avg_up_dis_week_PL	1	0.04564	0.00698	6.54	<.0001
OTHER_PPU	1	-0.22001	0.00165	-133.27	<.0001
OTHER_PR	1	0.03917	0.00588	6.66	<.0001
OTHER_Dnew	1	1.39924	0.01316	106.29	<.0001
avg_up_dis_week_OTHER	1	0.10953	0.00586	18.68	<.0001

Kraft Mayo Whip

Competitors

- Keeping all other variables constant, the following effects are significant at 5% significance level.

- With every one dollar increase in price per unit, sales of KMW decreases by 16%.
- Price reduction by one dollar increases sales of KMW by 4%.
- If there is display, then the sales for KMW increase by 26%.
- Price Reduction and display together leads to boost in sales of KMW by 9%.

Looking at the summary statistics of the competitors, it can be observed that-

- If there is a unit increase in price for the competitors Best Foods, Dukes or Private Label there is an increase in the sales for Kraft Mayo Whip by 0.1% to 3%.
- If there is a price reduction observed for the competitor Hellmanns, then the sales of Kraft Mayo Whip decrease by 22%.
- If there is a display at Hellmanns then the sales of Kraft Mayo Whip decrease by 9%. If there is a display at Kraft Mayo then sales of Kraft Mayo Whip decrease by 3%. However, both the brands belong to the same company. If there is a display at Private Label, then the sales of KMW decrease by 5%. Display for any other competitor (Best Foods or Dukes) does not affect the sales of Kraft Mayo Whip.

### d. Recommendations-

- Introduction of display leads to an increase in sales for Kraft Mayo Whip. The same effect is not observed for feature. So, the management should invest heavily on display.
- Price reduction also leads to a boost in the sales of KMW but display has a greater effect in terms of magnitude of increase in sales.
- Company can invest in both price reduction and display the frequency of which should be such that display is made more frequently.

### 3. Hypothesis Testing

**Hypothesis Test 1:** The sales of top brand vary across different volume equivalent categories

#### The ANOVA Procedure

Class Level Information		
Class	Levels	Values
new_vol	3	High_Volum Low_Volume Med_Volume

Number of Observations Read	545520
Number of Observations Used	545520

#### The SAS System

##### The ANOVA Procedure

Dependent Variable: DOLLARS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	120349151	60174575	9848.87	<.0001
Error	545517	3332997189	6110		
Corrected Total	545519	3453346340			

R-Square	Coeff Var	Root MSE	DOLLARS Mean
0.034850	242.2980	78.16518	32.25994

Source	DF	Anova SS	Mean Square	F Value	Pr > F
new_vol	2	120349150.9	60174575.4	9848.87	<.0001

We conducted an ANOVA test to find whether sales of top brand vary across different volume equivalent categories. We found that  $p\text{-value} < 0.001$ . We can say with 95% confidence that average dollar sales across different volume is different, which implies dollar sales varies with volume.

### The MEANS Procedure

Analysis Variable : DOLLARS		
new_vol	N Obs	Sum
High_Volum	57991	1306889.88
Low_Volume	242960	4387877.90
Med_Volume	244569	11903673.99

### The MEANS Procedure

Analysis Variable : PRICE_PER_UNIT			
new_vol	N Obs	Sum	Mean
High_Volum	57991	88380.86	1.5240445
Low_Volume	242960	546187.25	2.2480542
Med_Volume	244569	403943.37	1.6516540

**Hypothesis Test 2:** To test if larger stores have higher average price per unit as compared to smaller stores

### The TTEST Procedure

Variable: PRICE\_PER\_UNIT

STORERANK	N	Mean	Std Dev	Std Err	Minimum	Maximum
BOTT	5639	1.7473	0.7040	0.00937	0.1956	6.1843
TOP3	5494	1.7113	0.6367	0.00859	0.4950	4.7275
Diff (1-2)		0.0360	0.6716	0.0127		

STORERANK	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
BOTT		1.7473	1.7289 1.7656	0.7040	0.6912 0.7172
TOP3		1.7113	1.6945 1.7281	0.6367	0.6250 0.6488
Diff (1-2)	Pooled	0.0360	0.0150 Infy	0.6716	0.6629 0.6806
Diff (1-2)	Satterthwaite	0.0360	0.0151 Infy		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	11131	2.83	0.0024
Satterthwaite	Unequal	11070	2.83	0.0023

### Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	5638	5493	1.22	<.0001

From the results of the F-test for equality of variances, we have the 2 hypothesis:

H<sub>0</sub>: Variances of the 2 population are equal

H<sub>1</sub>: Variances are unequal

The p-value is <0.0001 which is lesser than the significance level of 0.05. So we reject the null hypothesis and conclude that the **variances of the 2 populations are unequal**.

The t-test for unequal variance has the following hypothesis:

H<sub>0</sub>: The average price per unit value for larger stores is greater than or equal to smaller stores

H<sub>1</sub>: The average price per unit value for larger stores is less than smaller stores



This is a upper ttest (right tail test). In the code this is done by setting `SIDES=U`.  
The p-value of 0.0024 is lesser than the significance level of 0.05. We reject the null hypothesis and are 95% confident that the average price per unit for larger stores is less than to the average price per unit for smaller stores.

Thank you!