

# Predictive Analytics Using SAS

## Project Report – Group 2

Akshay Pathak

Apurva Sinha

Priya Srinivas Gundevaina

Priyanka Guha

Sudharsana Rajasekaran

Varun Suresh

## Introduction:

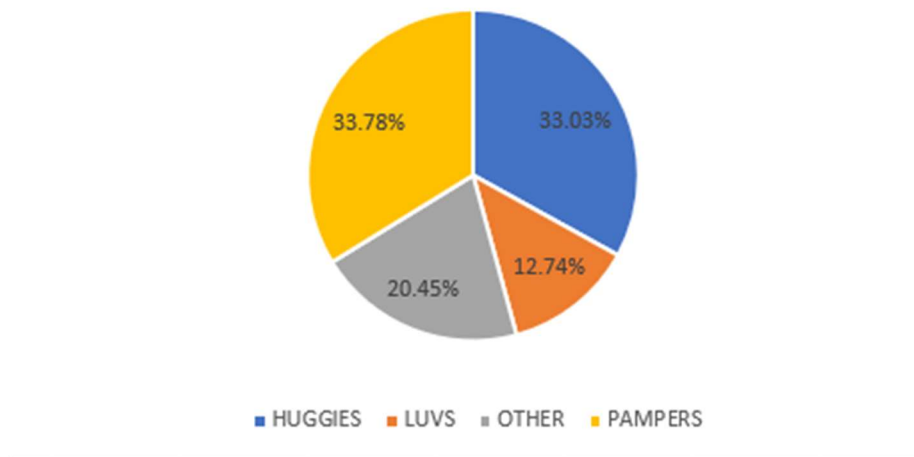
Pampers is an American household name for baby and toddler products sold by Proctor & Gamble. With data available at ProductUPC week level, we have made the following analysis to understand and answer the following questions:

1. What are the leading factors that affect the sales of diapers in stores. How do the prices of Pampers and its major competitors affect same ?
2. Understand the target segment of customer purchasing Pampers
3. Identify the reasons why customers select/or reject Pampers as a brand choice

## Market Share:

To begin with, we wanted to understand where Pampers stands with respect to other important players in the market. P&G, the parent company is the leader when it comes to capturing the baby and toddler products market, however Pampers as a brand comes second to Huggies.

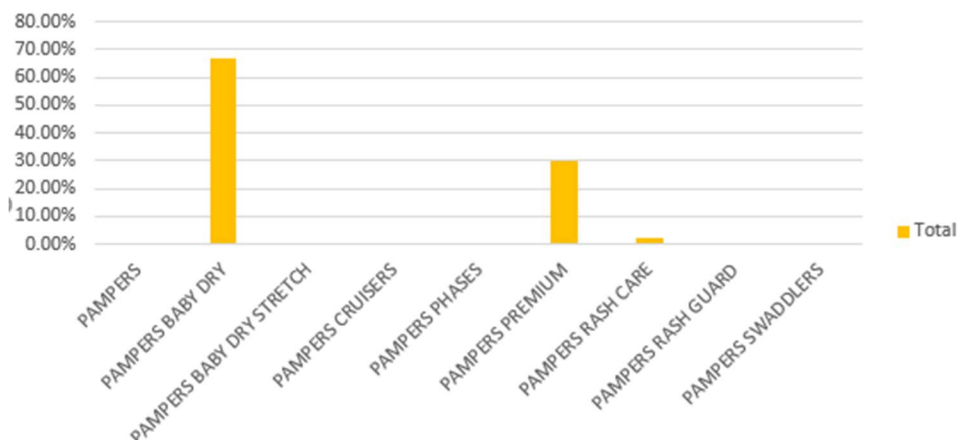
Percentage share of Brands



Both Pampers and Luvs (3<sup>rd</sup> in the market segment) contribute to P&G's growth

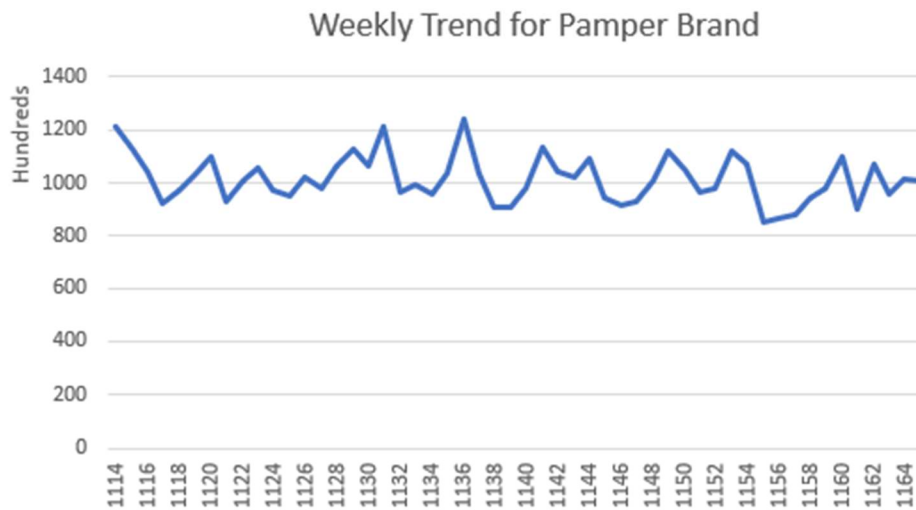
We get more clarity on Pampers by looking into the contribution of respective products/categories within Pampers that contributes to its overall sales. Pampers Baby Dry and Pampers Premium are most preferred product brands.

Percentage Share for Pamper Brand



Weekly trend of the tot\_units gives us a weekly outlook of fluctuations in the total units of the respective brands sold:

Tot\_units -> Y axis, Week -> X axis



### Analysis on Price Elasticity:

We perform this analysis to get a better understanding of the factors that would influence the sales of diapers with a concentration on price of Pampers and its competitors. This would help us come up with effective pricing strategies to help improve performance of Pampers.

### Analytical Dataset preparation:

Since, the diapers come in different pack-sizes, we will need to first get per Diaper Price. And since there several products under one diaper-brand, to compare the prices of one brand with other, it is important to convert the prices of each product to a weighted price and get a brand-level total of weighted price for a week and store.

The data was grouped into 4 different brands - Huggies, Luvs, Pampers and others(combining the remaining brands)

$$\text{Weighted price per diaper} = \frac{\text{Price per Diaper} * \text{Sales of Diaper for particular brand}}{\text{Total sales of diapers}} \times \% \text{ Market Share}$$

We have similarly computed weighted values for Price reduction, display and Feature. These values were then aggregated at store- week -brand level.

To capture these values for respective brands, we have converted these into columns so that for every week we now have the price, price reduction and store promotion details for every store.

Other derived variables include creation of the following interaction variables:

**PR\_F** = Weighted Price Reduction of respective Brand\*Weighted Feature of respective Brand

**Price\_F** = Weighted Price of respective Brand\*Weighted Feature of respective Brand

**Price\_PR** = Weighted Price of respective Brand\* Weighted Price Reduction of respective Brand

Alphabetic List of Variables and Attributes			
#	Variable	Type	Len
6	Feature_wt_brand1	Num	8
10	Feature_wt_brand2	Num	8
14	Feature_wt_brand3	Num	8
18	Feature_wt_brand4	Num	8
1	IR_KEY	Num	8
27	PR_F1	Num	8
28	PR_F2	Num	8
29	PR_F3	Num	8
30	PR_F4	Num	8
4	PR_wt_brand1	Num	8
8	PR_wt_brand2	Num	8
12	PR_wt_brand3	Num	8
16	PR_wt_brand4	Num	8
5	disp_wt_brand1	Num	8
9	disp_wt_brand2	Num	8
13	disp_wt_brand3	Num	8
17	disp_wt_brand4	Num	8
23	price_F1	Num	8
24	price_F2	Num	8
25	price_F3	Num	8
26	price_F4	Num	8
19	price_PR1	Num	8
20	price_PR2	Num	8
21	price_PR3	Num	8
22	price_PR4	Num	8
32	sales	Num	8
31	tot_units	Num	8
2	week	Num	8
3	wt_price_brand1	Num	8
7	wt_price_brand2	Num	8
11	wt_price_brand3	Num	8
15	wt_price_brand4	Num	8

Figure: List of variables in the data

#### Hausman Test for Random Effects:

The Hausman test is done to determine whether fixed effects model or random effects model is to be used for panel regression

**NULL Hypothesis ( $H_0$ ):** No Correlation between the error term ( $u_i$ ) and the independent variables in the model. Random effects model is to be used

**ALTERNATE Hypothesis ( $H_a$ ):** Correlation is present between the error term ( $u_i$ ) and the independent variables in the model. Fixed effects model is to be used

Hausman Test for Random Effects			
Coefficients	DF	m Value	Pr > m
19	19	181.54	<.0001

Since p-value of Hausman test is less than 0.05, there is convincing evidence to reject null hypothesis in favor of alternate hypothesis. Thus, fixed effects model needs to be used.

#### Panel Regression Model (Two Way Fixed Effects):

We observed a very strong correlation between tot\_units and sales variables in the dataset. This implies that both explain the sales of diapers in the dataset. Secondly using both the variables in the model would result in endogeneity in the model.

Since there is high correlation between the two, using tot\_units are synonymous with using sales as dependent variable.



Figure : Correlation between tot\_units and Sales

#### Dependent Variable:

Total units of diapers sold

#### Independent Variables:

WEIGHTED PRICE OF BRAND1(Huggies), WEIGHTED PRICE OF BRAND2(Luvs), WEIGHTED PRICE OF BRAND3(Pampers), WEIGHTED PRICE OF BRAND4(Others), WEIGHTED DISPLAY OF BRAND1(Huggies), WEIGHTED DISPLAY OF BRAND2(Luvs), WEIGHTED DISPLAY OF BRAND3(Pampers), WEIGHTED DISPLAY OF BRAND4(Others), WEIGHTED FEATURE OF BRAND4(Others), WEIGHTED PRICE REDUCTION OF BRAND2(Luvs), WEIGHTED PRICE REDUCTION OF BRAND4(Others), WEIGHTED PRICE\*PRICE REDUCTION OF BRAND1(Huggies), WEIGHTED PRICE\*PRICE REDUCTION OF BRAND2(Luvs), WEIGHTED PRICE\*PRICE REDUCTION OF BRAND3(Pampers), WEIGHTED PRICE\*PRICE REDUCTION OF BRAND4(Others) , WEIGHTED PRICE\*WEIGHTED FEATURE BRAND1(Huggies), , WEIGHTED PRICE\*WEIGHTED FEATURE BRAND2(Luvs), , WEIGHTED PRICE\*WEIGHTED FEATURE BRAND3(Pampers), , WEIGHTED PRICE\*WEIGHTED FEATURE BRAND4(Others)

#### Model Equation:

$$\text{Sales}_{\text{Pampers}} = B_0 + B_1 * \text{WEIGHTED PRICE OF BRAND1(Huggies)} + B_2 * \text{WEIGHTED PRICE OF BRAND2(Luvs)} + B_4 * \text{WEIGHTED PRICE OF BRAND3(Pampers)} + B_6 * \text{WEIGHTED PRICE OF BRAND4(Others)} + B_7 * \text{WEIGHTED DISPLAY OF BRAND1(Huggies)} + B_8 * \text{WEIGHTED DISPLAY OF BRAND2(Luvs)} + B_9 * \text{WEIGHTED DISPLAY OF BRAND3(Pampers)} + B_{10} * \text{WEIGHTED DISPLAY OF BRAND4(Others)} + B_{11} * \text{WEIGHTED FEATURE OF BRAND4(Others)} + B_{12} * \text{WEIGHTED PRICE REDUCTION OF BRAND2(Luvs)} + B_{13} * \text{WEIGHTED PRICE REDUCTION OF BRAND4(Others)} + B_{14} * \text{WEIGHTED PRICE*PRICE REDUCTION OF BRAND1(Huggies)} + B_{15} * \text{WEIGHTED PRICE*PRICE REDUCTION OF BRAND2(Luvs)} + B_{17} * \text{WEIGHTED PRICE*PRICE REDUCTION OF BRAND3(Pampers)} + B_{18} * \text{WEIGHTED PRICE*PRICE REDUCTION OF BRAND4(Others)} + B_{19} * \text{WEIGHTED PRICE*WEIGHTED FEATURE BRAND1(Huggies)} + B_{20} * \text{WEIGHTED PRICE*WEIGHTED FEATURE BRAND2(Luvs)} + B_{21} * \text{WEIGHTED PRICE*WEIGHTED FEATURE BRAND3(Pampers)} + B_{22} * \text{WEIGHTED PRICE*WEIGHTED FEATURE BRAND4(Others)}$$

#### Model Statistics:

From the entire dataset, the variables were treated for multicollinearity prior to building the model

The SAS System			
The PANEL Procedure			
Fixed Two-Way Estimates			
Dependent Variable: tot_units			
Model Description			
Estimation Method		FixTwo	
Number of Cross Sections		1615	
Time Series Length		52	
Fit Statistics			
S SE	11175112.90	DFE	64438
M SE	173.4243	Root M SE	13.1691
R-Square	0.8980		

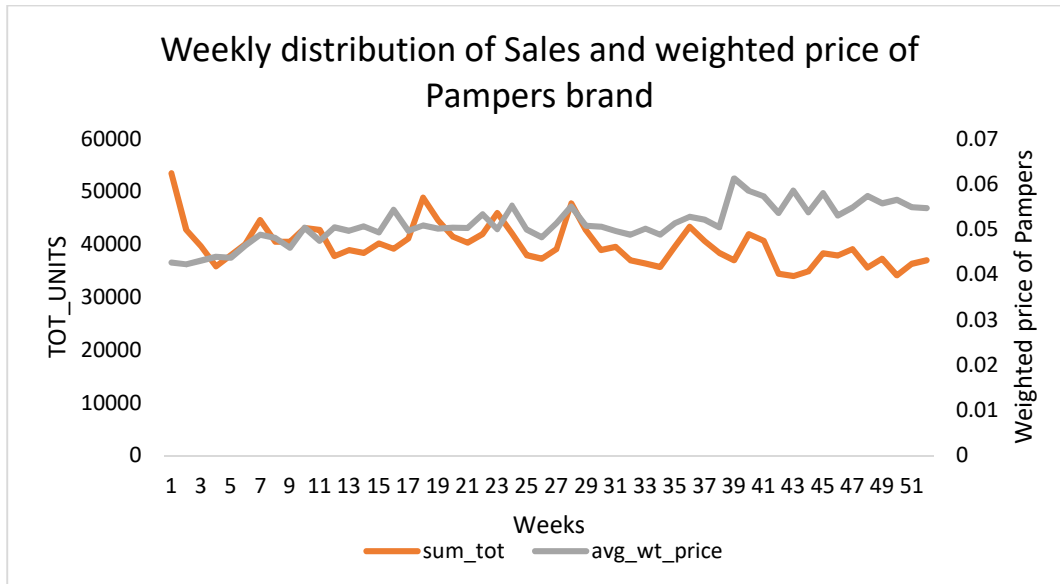
Model interpretation

From the model parameter estimates, we notice that all the variables are significant besides weighted price for Luvs, weighted display for luvs, weighted price reduction for other brands and interaction of weighted price and price reduction for other brands.

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Pr >  t
Intercept	1	55.88475	1.5827	35.30	<.0001
wt_price_brand1	1	13.17932	3.0466	4.33	<.0001
wt_price_brand2	1	8.310029	2.7452	3.03	0.0025
wt_price_brand3	1	-95.2737	2.3845	-39.96	<.0001
wt_price_brand4	1	-21.1196	3.2203	-6.56	<.0001
disp_wt_brand1	1	-3.08928	0.7991	-3.87	0.0001
disp_wt_brand2	1	-0.71788	0.6673	-1.08	0.2820
disp_wt_brand3	1	30.29519	0.6908	43.86	<.0001
disp_wt_brand4	1	-3.38559	0.6969	-5.67	<.0001
Feature_wt_brand4	1	-2.76551	0.3269	-8.46	<.0001
PR_wt_brand2	1	-1.8547	0.2286	-8.11	<.0001
PR_wt_brand4	1	0.274595	0.2912	0.94	0.3457
price_PR1	1	-18.7239	5.4513	-3.43	0.0006
price_PR2	1	-35.0113	5.3416	-6.55	<.0001
price_PR3	1	44.93703	4.7291	9.50	<.0001
price_PR4	1	-0.63762	5.6159	-0.11	0.9096
price_F1	1	-7.75616	0.9435	-8.22	<.0001
price_F2	1	-6.27937	1.3694	-4.59	<.0001
price_F3	1	-28.7524	3.9671	-7.25	<.0001
price_F4	1	108.6667	5.0263	21.62	<.0001

### Calculating Self and Cross Elasticity:

From the overall weekly trend of Sales and price of pampers , we notice that there seems to be some relation that indicates that when prices increase briefly, the sales of pampers decreases. This could be indicating that people are price sensitive when it comes to purchasing Pampers brand. This would be better explained based on the price elasticity calculation.



### Calculating Means

Variable	N	Mean	Std Dev	Minimum	Maximum
tot_units	66123	31.4197329	35.5111812	1.0000000	594.0000000
wt_price_brand1	66123	0.0513165	0.0605905	0.000074441	1.2796067
wt_price_brand2	66123	0.0896474	0.1018263	2.0834233E-6	0.6974754
wt_price_brand3	66123	0.0708028	0.0884375	9.2194716E-6	1.0649970
wt_price_brand4	66123	0.0180177	0.0246201	3.785543E-6	0.5406348
price_PR1	66123	0.0035828	0.0558816	0	4.1286777
price_PR2	66123	0.0045391	0.0272185	0	2.1357192
price_PR3	66123	0.0037354	0.0284852	0	3.1484408
price_PR4	66123	0.000875059	0.0065452	0	0.6774400
price_F1	66123	0.0016960	0.0309911	0	3.6695957
price_F2	66123	0.0021492	0.0287719	0	1.2323894
price_F3	66123	0.0026007	0.0403474	0	2.6126301
price_F4	66123	0.000564494	0.0048334	0	0.3525845

### Price Elasticity Calculation:

**Self-Price Elasticity:** A measure used in economics to show the responsiveness, or elasticity, of the quantity demanded of a good or service to a change in its price when nothing but the price changes.

$$\begin{aligned}\text{SelfPrice Elasticity of Pampers} &= \frac{\beta_4 * \text{PricePampers}}{\text{SalesPampers}} \\ &= -95.26 * 0.071 / 31.42 \\ &= -0.22\end{aligned}$$

$$\begin{aligned}\text{SelfPrice Elasticity of Pampers} &= \frac{(\beta_4 + \beta_{17} \text{PRPamper} + \beta_{21} \text{FeaturePampers}) * \text{PricePampers}}{\text{SalesPampers}} \\ (\text{considering Price reduction and Feature})\end{aligned}$$

$$= (-95.27 + 44.93 * 0.0045 - 28.75 * 0.0026) * 0.071 / 31.42$$

$$= -0.21$$

#### **Inference:**

This means that when Pampers reduces the price by 1% then there will be a 0.021% increase in sales. However, the Price reduction and advertisements for Pampers does not seem to be doing its effect. This is because the Price elasticity after considering Price reduction and Feature does not improve significantly.

This indicates that company needs to devote more attention for its marketing strategies and when it decides to give price reductions.

Another way of interpreting it is that price changes in Pampers does not really affect its overall sales, that is they are price inelastic. People are probably not price sensitive, but might be more brand sensitive

$$\text{Cross Price Elasticity of Pampers} = \frac{\beta_2 * \text{PriceHuggies}}{\text{SalesHuggies}}$$

$$= 13.18 * 0.05 / 31.42$$

$$= 0.021$$

$$\text{Cross Price Elasticity of Pampers} = \frac{(\beta_2 + \beta_{15} \text{PRHuggies} + \beta_{19} \text{FeatureHuggies}) * \text{PriceHuggies}}{\text{SalesPampers}}$$

*(considering Price reduction and Feature)*

$$= (13.18 - 18.72 * 0.0035 - 7.76 * 0.0017) * 0.05 / 31.42$$

$$= 0.027$$

#### **Inference:**

Pampers is not significantly affected by the prices of Huggies which seems to be the market leader for baby products. Like the interpretation of self-price elasticity, this means that sales of Pampers are inelastic to the prices of its competitors.



## ANALYSIS -2

### Introduction:

To identify customer attributes that drive PAMPERS sales. This can be done by performing RFM analysis which helps us identify which segment of customers are more frequent, recent buyers of the product and revenue generators.

### RFM Segmentation:

Applying RFM segmentation on the overall panel dataset from Grocery, Drug and MA stores. From sales of Diaper over the observed time has the following metrics were captured that were instrumental in segmentation.

- I. Recency – Recency refers to how recent the customer has purchased. It is the difference between last purchase date and the most recent date.
- II. Frequency – Number of purchases made by a customer in a given week
- III. Monetary – Total amount spent by the customer in any given week.

Initial analysis helps us understand that a customer on an average spent around 42\$ by making 3 visits every 18 weeks. There is a high level of correlation between monetary value and frequency. So, we can consider monetary value as the prime criterion for segmentation.

3 Variables:

money freq weekcnt

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
money	346	42.36965	54.97768	14660	5.99000	378.21000
freq	346	3.01156	3.85082	1042	1.00000	28.00000
weekcnt	346	18.45087	14.65523	6384	0	51.00000

Pearson Correlation Coefficients, N = 346  
Prob > |r| under H0: Rho=0

	money	freq	weekcnt
money	1.00000	0.94827 <.0001	-0.37425 <.0001
freq	0.94827 <.0001	1.00000	-0.34704 <.0001
weekcnt	-0.37425 <.0001	-0.34704 <.0001	1.00000

We wanted to identify what demographic attributes of the customer are responsible for the customer's being high value or low value. Criterion:

- Top customers = Monetary\_score 4,5
- Bottom customers = Monetary\_score 1,2

Variable	N	Mean	Std Dev	Minimum	Maximum
PANID	132	2206066.88	1040816.33	1118786.00	3842443.00
monetary	132	85.3001515	70.0039230	27.4700000	378.2100000
freq	132	5.9242424	4.9940747	2.0000000	28.0000000
weekcnt	132	12.2045455	11.8571257	0	50.0000000
monetary_score	132	4.5227273	0.5013860	4.0000000	5.0000000
frequency_score	132	4.3181818	0.7238382	3.0000000	5.0000000
recency_score	132	3.6439394	1.2789104	1.0000000	5.0000000
child_num	132	1.0454545	0.8548348	0	3.0000000
pets_total	132	0.8560606	1.0198980	0	4.0000000
fam_size_L	132	0.5227273	0.5013860	0	1.0000000
fam_size_R	132	0.4772727	0.5013860	0	1.0000000
fam_size_O	132	0	0	0	0
fam_income_L	132	0.1439394	0.3523655	0	1.0000000
fam_income_M	132	0.4318182	0.4972164	0	1.0000000
fam_income_H	132	0.3333333	0.4732004	0	1.0000000
fam_income_VH	132	0.0909091	0.2885750	0	1.0000000
fam_income_O	132	0	0	0	0
age_mY	132	0.0075758	0.0870388	0	1.0000000
age_mM	132	0.6136364	0.4887705	0	1.0000000
age_mE	132	0.1515152	0.3599162	0	1.0000000
age_mO	132	0.2272727	0.4206667	0	1.0000000
age_fY	132	0.0454545	0.2090924	0	1.0000000
age_fM	132	0.7196970	0.4508583	0	1.0000000
age_fE	132	0.1515152	0.3599162	0	1.0000000
age_fO	132	0.0833333	0.2774383	0	1.0000000
occ_mWH	132	0.3409091	0.4758206	0	1.0000000
occ_mWL	132	0.1287879	0.3362411	0	1.0000000
occ_mB	132	0.1590909	0.3671542	0	1.0000000
occ_mNO	132	0.2196970	0.4156186	0	1.0000000
occ_fWH	132	0.4318182	0.4972164	0	1.0000000
occ_fWL	132	0.1515152	0.3599162	0	1.0000000
occ_fB	132	0.0681818	0.2530179	0	1.0000000
occ_fNO	132	0.1439394	0.3523655	0	1.0000000
one_child	132	0.3636364	0.4828783	0	1.0000000
two_child	132	0.2954545	0.4579849	0	1.0000000
three_child	132	0.0303030	0.1720729	0	1.0000000
zero_child	132	0.3106061	0.4645046	0	1.0000000

Figure 1: Top-Customers (M = 4,5)

Variable	N	Mean	Std Dev	Minimum	Maximum
PANID	141	2174533.54	1047184.78	1104935.00	3842559.00
monetary	141	12.0736879	2.2461225	5.9900000	16.9900000
freq	141	1.0000000	0	1.0000000	1.0000000
weekcnt	141	24.4113475	15.0736624	0	51.0000000
monetary_score	141	1.0000000	0	1.0000000	1.0000000
frequency_score	141	1.0000000	0	1.0000000	1.0000000
recency_score	141	2.5035461	1.3396381	1.0000000	5.0000000
child_num	141	0.6595745	0.8002659	0	3.0000000
pets_total	141	1.0000000	1.3680017	0	6.0000000
fam_size_L	141	0.3120567	0.4649847	0	1.0000000
fam_size_R	141	0.6879433	0.4649847	0	1.0000000
fam_size_O	141	0	0	0	0
fam_income_L	141	0.1489362	0.3572948	0	1.0000000
fam_income_M	141	0.5460993	0.4996453	0	1.0000000
fam_income_H	141	0.2624113	0.4415135	0	1.0000000
fam_income_VH	141	0.0425532	0.2025671	0	1.0000000
fam_income_O	141	0	0	0	0
age_mY	141	0.0141844	0.1186722	0	1.0000000
age_mM	141	0.4184397	0.4950617	0	1.0000000
age_mE	141	0.3617021	0.4822062	0	1.0000000
age_mO	141	0.2056738	0.4056341	0	1.0000000
age_fY	141	0.0638298	0.2453210	0	1.0000000
age_fM	141	0.4893617	0.5016689	0	1.0000000
age_fE	141	0.3900709	0.4895048	0	1.0000000
age_fO	141	0.0567376	0.2321653	0	1.0000000
occ_mWH	141	0.2836879	0.4523943	0	1.0000000
occ_mWL	141	0.1276596	0.3349001	0	1.0000000
occ_mB	141	0.1347518	0.3426756	0	1.0000000
occ_mNO	141	0.3617021	0.4822062	0	1.0000000
occ_fWH	141	0.2695035	0.4452837	0	1.0000000
occ_fWL	141	0.1843972	0.3891903	0	1.0000000
occ_fB	141	0.0851064	0.2800347	0	1.0000000
occ_fNO	141	0.2624113	0.4415135	0	1.0000000
one_child	141	0.3120567	0.4649847	0	1.0000000
two_child	141	0.1418440	0.3501339	0	1.0000000
three_child	141	0.0212766	0.1448194	0	1.0000000
zero_child	141	0.5248227	0.5011638	0	1.0000000

Figure 2: Bottom-Customers (M = 1,2)

From the above output, for customers choosing 'PAMPERS' as their preferred brand, we can see that:

- 132 Top-customers have contributed up to 85.3% of the overall revenue. These are customers who visit the shop most frequently and are most recent compared to the other segment.
- 141 are the low revenue customers contribute to only 12% of the overall revenue.
- 52% of the Large family size households belong to high value customers group when compared to regular family size households which have only 47%.
- Middle-aged male and female groups in a household prefer Pampers more when compared to other age groups.
- 'PAMPERS' is preferred more by household having females occupying positions in White collar jobs whereas, females with no occupation are less likely to favor our products. This could mean that the pricing of our products is a little higher as compared to other brands.
- Households with one or two kids will more preferably buy our product.

#### Preference: high value customers:

Large family household, Middle aged male and female, female with white collar job and household with 1 or 2 children.

#### Preference: low value customers

Regular family household, Medium family income, female with no jobs and household expecting a child.

1. Households expecting a child do not have specific choice of brand as they are unaware about the benefits of the features of one brand over another. Hence, it is possible that their choice will be driven by price of the diaper with their income playing as a major role in it.
  2. Large family size prefers to buy PAMPERS due to prior experience with the brand which had led to a strong brand affinity.
  3. Middle aged Male and Female are aware of the features of the product and its benefits. Hence, they prefer PAMPERS
- Brand Choice Preference

### ANALYSIS -3

To study the impact of feature, display and price reduction on Brand selection.

#### Data Preparation:

- Combining store level data for a drug and grocery store
- Created 4 primary brands – Huggies, Pampers, Luvs and Others.
- Data filtered for customers having continuous weeks of data
- Weighted Price weighted display, feature and price reduction score are calculated for a store(i) :

$$\text{Weighted Price}_i = \Sigma \text{Price}_i * \frac{\text{Sales of Diapers}}{\text{Total Sales of all Diapers}}$$

$$\text{Weighted PriceReduction}_i = \Sigma \text{Price reduction}_i * \frac{\text{Sales of Diapers}}{\text{Total Sales of all Diapers}}$$

$$\text{Weighted Feature}_i = \Sigma \text{Feature}_i * \frac{\text{Sales of Diapers}}{\text{Total Sales of all Diapers}}$$

$$\text{Weighted Display} = \Sigma \text{Display}_i * \frac{\text{Sales of Diapers}}{\text{Total Sales of all Diapers}}$$

#### Model:

To identify brand preference based on Features, Display and Price reduction at a store level.

Looks like default printed output

##### The LOGISTIC Procedure

Model Information	
Data Set	WORK.BRAND_CHOICE_STORE_FEATURES
Response Variable	brand
Number of Response Levels	4
Model	generalized logit
Optimization Technique	Newton-Raphson

Number of Observations Read	3674
Number of Observations Used	3674

Response Profile		
Ordered Value	brand	Total Frequency
1	HUGGIES	825
2	LUVS	737
3	OTHER	1351
4	PAMPERS	761

Logits modeled use brand='PAMPERS' as the reference category.

#### Model Stats:

Model Convergence Status	
Convergence criterion (GCONV=1E-8) satisfied.	

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9937.851	9261.461
SC	9956.478	9503.614
-2 Log L	9931.851	9183.461

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	748.3900	36	<.0001
Score	846.9979	36	<.0001
Wald	693.6805	36	<.0001

AIC and BIC are very close this means that the model is getting penalized for using a higher number of variables.

$$\text{McFadden's Rsqr} = 1 - \frac{\text{LogLikelihood of the model}}{\text{LogNulllikelihood of the model}} = 7.5\%$$

Analysis of Maximum Likelihood Estimates							
Parameter	brand	DF	Estimate	Standard	Wald	Pr > Chi Sq	Exp(Est)
				Error	Chi-Square		
Intercept	HUGGIES	1	0.1473	0.0925	2.5366	0.1112	1.159
Intercept	LUVS	1	0.0316	0.0952	0.1098	0.7403	1.032
Intercept	OTHER	1	1.0648	0.0808	173.495	<.0001	2.9
disp_wt_brand1	HUGGIES	1	-0.445	0.3047	2.1321	0.1442	0.641
disp_wt_brand1	LUVS	1	-0.9218	0.3052	9.1241	0.0025	0.398
disp_wt_brand1	OTHER	1	-2.0033	0.2584	60.1138	<.0001	0.135
disp_wt_brand2	HUGGIES	1	-0.3318	0.3329	0.9938	0.3188	0.718
disp_wt_brand2	LUVS	1	1.417	0.2805	25.5123	<.0001	4.125
disp_wt_brand2	OTHER	1	-0.3691	0.2919	1.5982	0.2062	0.691
disp_wt_brand3	HUGGIES	1	0.9094	0.2362	14.8176	0.0001	2.483
disp_wt_brand3	LUVS	1	0.0913	0.2513	0.132	0.7164	1.096
disp_wt_brand3	OTHER	1	-0.4691	0.2265	4.2885	0.0384	0.626
disp_wt_brand4	HUGGIES	1	-1.24	0.6439	3.708	0.0542	0.289
disp_wt_brand4	LUVS	1	0.3814	0.6812	0.3134	0.5756	1.464
disp_wt_brand4	OTHER	1	0.3405	0.5893	0.3339	0.5634	1.406
Feature_wt_brand1	HUGGIES	1	-0.8336	0.3001	7.7178	0.0055	0.434
Feature_wt_brand1	LUVS	1	-0.2057	0.2906	0.501	0.4791	0.814
Feature_wt_brand1	OTHER	1	0.3065	0.2426	1.5957	0.2065	1.359
Feature_wt_brand2	HUGGIES	1	0.1842	0.2092	0.7754	0.3785	1.202
Feature_wt_brand2	LUVS	1	0.5456	0.2096	6.7738	0.0093	1.726
Feature_wt_brand2	OTHER	1	-0.3125	0.1995	2.4517	0.1174	0.732
Feature_wt_brand3	HUGGIES	1	-0.1791	0.2246	0.6362	0.4251	0.836
Feature_wt_brand3	LUVS	1	0.0235	0.2307	0.0104	0.9187	1.024
Feature_wt_brand3	OTHER	1	0.0287	0.2048	0.0197	0.8884	1.029
Feature_wt_brand4	HUGGIES	1	0.2692	0.4832	0.3103	0.5775	1.309
Feature_wt_brand4	LUVS	1	-1.165	0.5403	4.6487	0.0311	0.312
Feature_wt_brand4	OTHER	1	-1.7091	0.4746	12.9687	0.0003	0.181
PR_wt_brand1	HUGGIES	1	0.5782	0.431	1.7999	0.1797	1.783
PR_wt_brand1	LUVS	1	-0.8935	0.5442	2.6959	0.1006	0.409
PR_wt_brand1	OTHER	1	-0.975	0.4801	4.1231	0.0423	0.377
PR_wt_brand2	HUGGIES	1	-0.0451	0.5922	0.0058	0.9394	0.956
PR_wt_brand2	LUVS	1	-1.5084	0.7965	3.5865	0.0582	0.221
PR_wt_brand2	OTHER	1	-0.9472	0.7082	1.7892	0.181	0.388
PR_wt_brand3	HUGGIES	1	13.534	7.5047	3.252	0.0713	754289
PR_wt_brand3	LUVS	1	-822	887.1	0.8587	0.3541	0
PR_wt_brand3	OTHER	1	10.19	7.7974	1.7077	0.1913	26623.4
PR_wt_brand4	HUGGIES	1	4.4302	2.5561	3.004	0.0831	83.948
PR_wt_brand4	LUVS	1	3.1269	2.7468	1.2959	0.255	22.803
PR_wt_brand4	OTHER	1	12.891	2.2732	32.159	<.0001	396800

#### Statistically significant parameters are:

- disp\_wt\_brand1<sub>Luvs</sub>, disp\_wt\_brand1<sub>Other</sub>, disp\_wt\_brand2<sub>Luvs</sub>, disp\_wt\_brand3<sub>Huggies</sub>, disp\_wt\_brand3<sub>Other</sub>
- Feature\_wt\_brand1<sub>Huggies</sub>, Feature\_wt\_brand2<sub>Luvs</sub>, Feature\_wt\_brand4<sub>Pampers</sub>, Feature\_wt\_brand4<sub>Other</sub>
- PR\_wt\_brand1<sub>Other</sub>, PR\_wt\_brand4<sub>Other</sub>

#### Interpretation:

When compared with Pampers, keeping all the other variables are constant, the likelihood of choice selection between brands:

- **disp\_wt\_brand1<sub>Luvs</sub>**: The relative odds of selecting Luvs products decreases by 61%, if there is a display for Pampers products
- **disp\_wt\_brand1<sub>Other</sub>**: The relative odds of selecting Other products decreases by 87%, if there is a display for Pampers products
- **disp\_wt\_brand2<sub>Luvs</sub>**: The relative odds of selecting Luvs products increases by 312%, if there is a display for Luvs products

- **disp\_wt\_brand3<sub>Huggies</sub>**: The relative odds of selecting Huggies products increases by 140%, if there is a display for Huggies products
- **disp\_wt\_brand3<sub>Other</sub>**: The relative odds of selecting Other products decreases by 40%, if there is a display for Huggies products
- **Feature\_wt\_brand1<sub>Huggies</sub>**: The relative odds of selecting Huggies products decreases by 57%, if there is a store level promotion featuring Pampers products
- **Feature\_wt\_brand2<sub>Luvs</sub>**: The relative odds of selecting Luvs products Increases by 70%, if there is a store level promotion featuring Luvs products
- **Feature\_wt\_brand4<sub>Luvs</sub>**: The relative odds of selecting Luvs products decreases by 69%, if there is a store level promotion featuring Other products
- **Feature\_wt\_brand4<sub>Other</sub>**: The relative odds of selecting Other products decreases by 82%, if there is a store level promotion featuring Other products
- **PR\_wt\_brand1<sub>Other</sub>**: The relative odds of selecting Other products decreases by 63%, if there is a store level store level price reduction for Pampers products
- **PR\_wt\_brand4<sub>Other</sub>**: The relative odds of selecting Other products becomes very high if there is store level price reduction for Other products

**Insights:**

- If there Display in a store for Pampers, then people prefer it over Luvs and Other brands
- If there Pampers is Featured in a store, then it is preferred over Huggies
- If there is a Price reduction for Pampers, then it is preferred over Other brands

**Recommendation:**

- Price Sensitivity analysis:
  - Sales of Pampers is price insensitive. People are probably not price sensitive but might be more brand sensitive. So, company needs to devote more attention for its marketing strategies and when it decides to give price reductions.
- RMF Analysis:
  - Large family household, Middle aged male and female, female with white collar job and household with 1 or 2 children prefer Pampers over other brands. The marketing team can reallocate some of the promotion budget to other group of customers with higher marketing potential.
  - Regular family household, Medium family income, female with no jobs and household expecting a child prefer other brands over Pampers. Special promotional strategies must be designed to target this segment.
- Band Choice model:
  - Feature and Display in stores for Pampers improves the odds of getting purchased by the customers over other brands. Hence, Store Managers should try to predict the periods where the product sales and put the product on display and feature ads to promote sales.