12/4/2018

**Submitted to:**

B.P.S. Murthy

ALL LIQUID DETERGENT

Market Analysis Report

**Submitted to:**

* Eeshpaul Saggu
* Pooja Arora
* Priyanshi Khatri
* Rajdeep Arora
* Sumeet Kumar

**STRATEGY**

To study the collected data and gain insights out of it, we have followed the following techniques to come out with recommendations for improvement in market share of our brand ALL and Liquid brand category.

1. **Price Elasticity Study** – Propose strategy for store level pricing and promotions

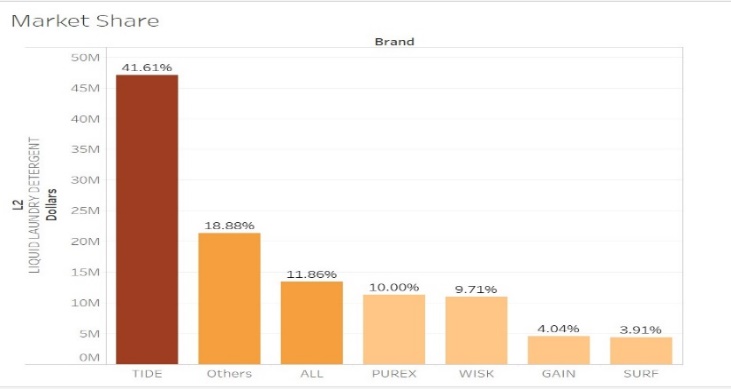
2. **Monetary Segmentation through RFML** – Use panelist purchase behavior to create strategic segments that will help us understand our present customer base and our target customer base as well. This will help us perform effective targeting marketing

3. **How Long customer stays** – Using survival analysis we found inter-arrival times between purchases for ALL and Tide brand. We have tried to understand what factors increase/decrease inter-arrival time in order to know when to best deploy marketing

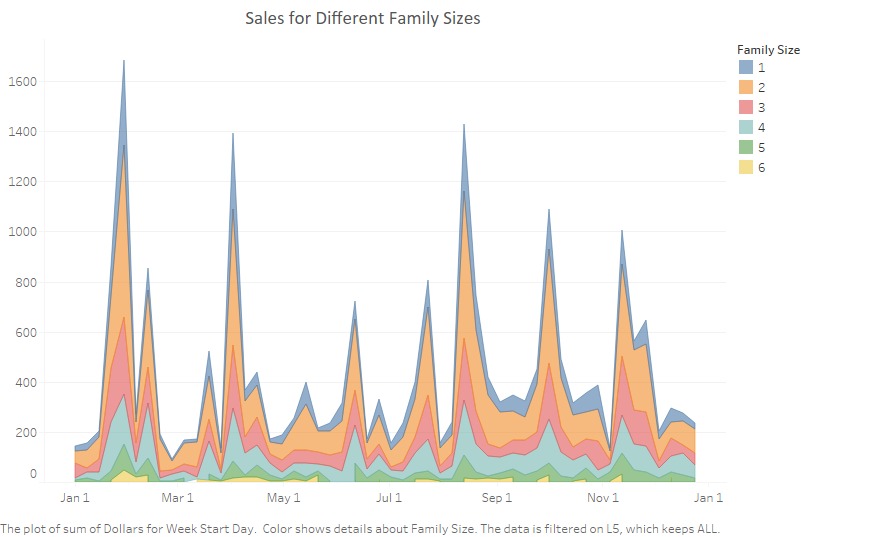
**EXPLORATORY DATA ANALYSIS**

**Brand Selection:**

We selected the detergent brand named ALL and analyzed the data for its liquid category. TIDE is the market leader in liquid category with the overwhelming market share of 41.61% and our brand holds the second largets share in the market which is 11.86%.

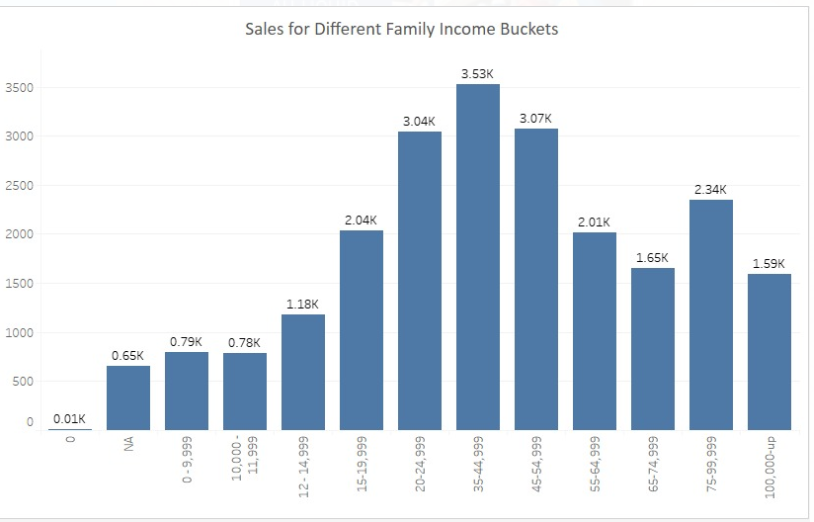
****

*Fig 1*

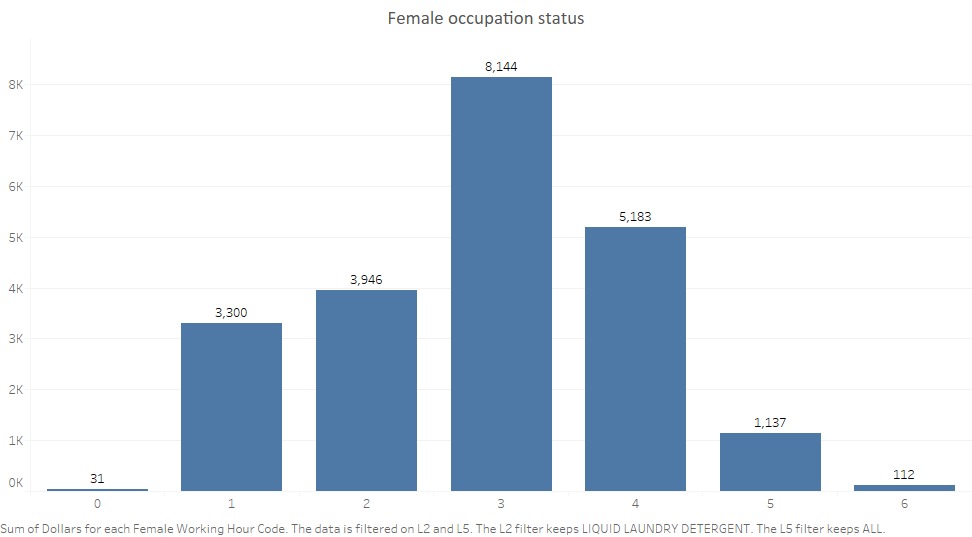


*Fig 2*

1. We observed that most of our customers have 2 people in a family



1. The sales of ALL come most from a mid-income demographic (20K-55K). Panelists in the high or low income brackets do not buy much of ALL. Tide on the other hand get high sales from High income panelists as well



**Survival Analysis**

We used survival analysis to predict how long it would take our customer (in weeks) to come back again to buy any liquid detergent product. The one-year data available are left censored as we do not know if a customer was a returning customer or it was their first purchase. We have analyzed the customers for our brand All and major competitor Tide. Using Weibul distribution, following is the output of the eight models we ran for ALL and Tide for different segments arrived through RFML analysis.

For “All” customers, following are the significant results shown. Bold represents the significance at 95% of confidence interval and otherwise not significant.

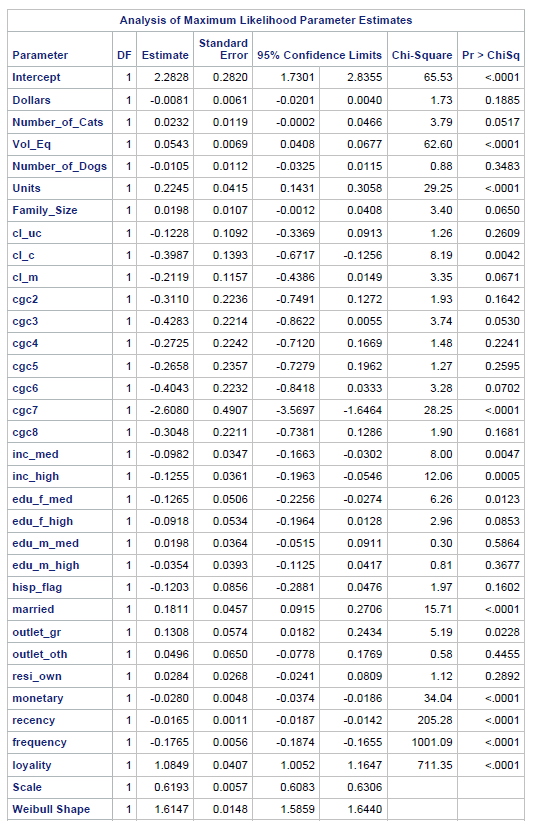
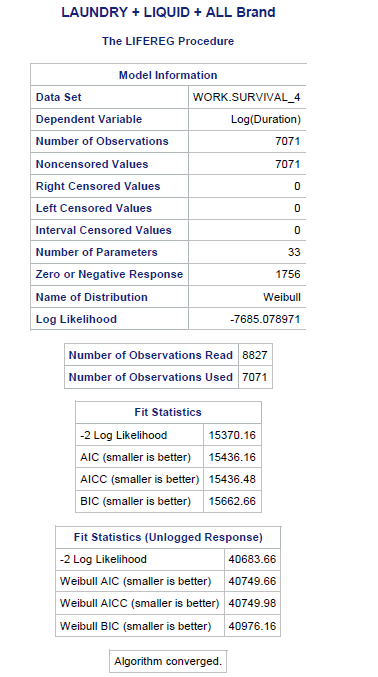
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Opportunist** | **Bulk** | **Core** | **Marginal** |
| Dollars | *-0.01* | ***0.02*** | ***0.01*** | ***0.01*** |
| Vol\_Eq | ***0.05*** | ***0.00*** | ***0.02*** | ***0.03*** |
| Units | ***0.22*** | ***0.12*** | ***0.13*** | ***0.13*** |
| Family\_Size | ***0.02*** | ***0.05*** | ***-0.01*** | ***-0.07*** |
| Resi\_own | *0.03* | ***0.14*** | *0.01* | ***-0.19*** |
| married | ***0.18*** | ***-0.38*** | ***-0.07*** | ***0.08*** |
| # cats | ***0.02*** | ***0.07*** | ***0.03*** | *0.00* |
| # dogs | *-0.01* | *-0.04* | ***0.03*** | ***-0.03*** |

For “Tide” Customers, following are the significant results shown. Bold represents the significance at 95% of confidence interval and otherwise not significant.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Opportunist** | **Bulk** | **Core** | **Marginal** |
| Dollars | ***0.01*** | ***0.02*** | ***0.02*** | ***0.01*** |
| Vol\_Eq | ***0.02*** | ***0.02*** | ***0.01*** | ***0.02*** |
| Units | ***0.12*** | ***0.13*** | ***0.13*** | ***0.19*** |
| Family\_Size | *0.00* | *-0.01* | ***-0.02*** | ***-0.04*** |
| Resi\_own | ***0.06*** | *-0.03* | ***-0.05*** | ***-0.06*** |
| married | ***-0.09*** | ***0.29*** | ***0.06*** | ***0.05*** |
| # cats | *0.02* | *0.01* | ***-0.06*** | ***-0.01*** |
| # dogs | ***-0.03*** | ***-0.12*** | ***-0.01*** | ***-0.02*** |

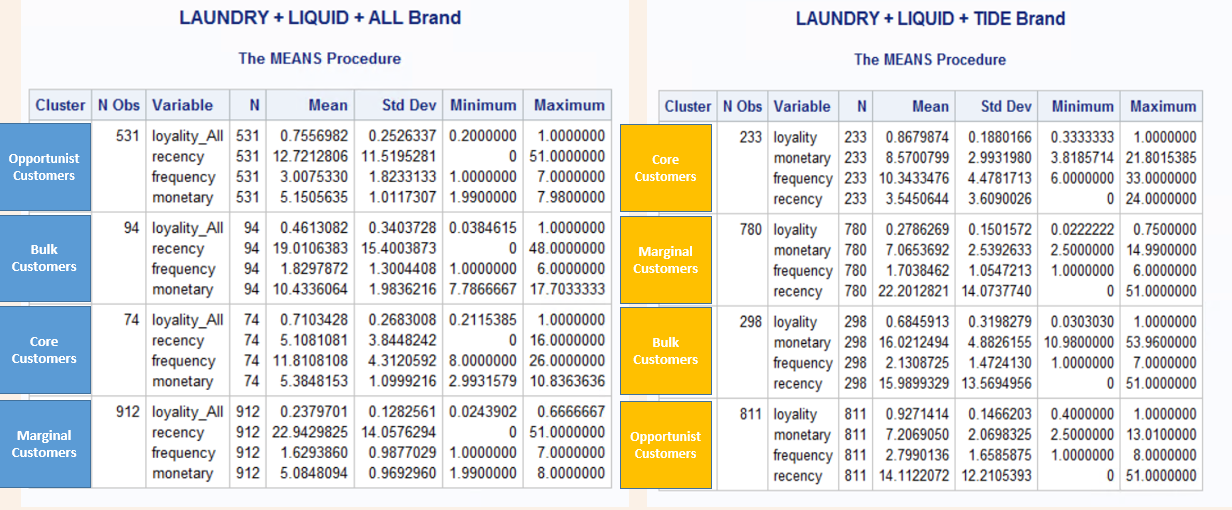
**Important Insights**:

1. Every dollar spent on the purchase of ALL liquid is associated with an increase of 2%, 1% and 1% in duration for the customer’s (Bulk, Core and Marginal) next purchase of laundry liquid detergent while controlling for other variables in the model.
2. There is a unanimous increment in inter-arrival time for increase in family size irrespective of brands explaining customers buying in bulk.
3. For loyal customers, there is a significant increment in inter-arrival time for people who owns the house as compared to marginal customers for the brand Tide
4. There is a unanimous increment in inter-arrival time for increase in volume bought irrespective of brands explaining customers buying in bulk.
5. Bulk Customer inter-arrival time increases 29% for married as compared to not married for Tide Product
6. Number of Cats and Dogs represents small family. There is an inclination for small family to buy All over Tide for all positive coefficient shown as compared to negative significant estimates shown by Tides

**Customer segmentation** (Recency, Frequency, Monetary and Loyalty)

We have looked at people who have purchased ALL and Tide at least once. On that data set of semi-loyal customers, we have created RFM segments to better understand our present customers, our target customers and their attributes and demographics.



The segments that we have identified are the following:

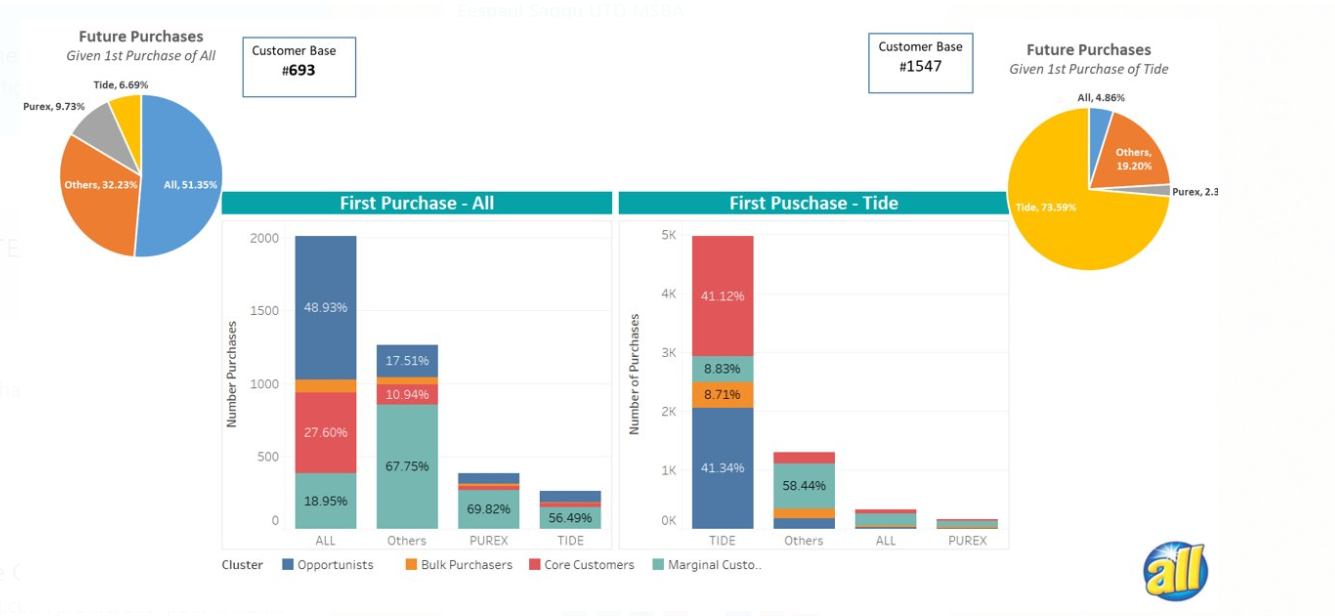
1) **Opportunistic Customers:** These customers wait for promotions and discounts to buy product. Less Recent and frequent in nature. For TIDE the have a very high recency of 14 that is couple with low frequency of 2.79.

2) **Bulk Purchasers:** These customers buy in bulk with low frequency. Average Transaction value is high for them. For TIDE the monetary component of Bulk purchasers is the highest ($22.2) and for ALL they are a close second in monetary value ($19)

3) **Core Customers:** These are loyal customers for the product. Highly loyal, more recent and more frequent in nature. However, these are less in number for both TIDE (233) and ALL (74)

4) **Marginal Customers:** These are not loyal customers and doesn’t show any inclination towards the product. Their purchase behavior is erratic, and it is difficult to assign them specific behavior. A large chunk of both TIDE (780) and ALL (912) fall into this segment

Next, we looked at people who people whose first purchase was TIDE and ALL and we compared their purchase journey across the time period that we have. (*Look at Fig below)*



**Insight:**

The majority of panelists who have first purchased TIDE or ALL at the start of our time period tend to stick with TIDE or ALL. If ALL first purchasers switch, they are least likely to go to TIDE (only 6.99%), however if TIDE first purchasers switch, they are most likely to go to ALL (More than double that of Purex). That being said, TIDE first purchasers are not very likely to switch as 73% of future purchases of TIDE first purchase continue to be TIDE.

**ANALYSIS ON PRICE ELASTICITY OF SALES (Store Level)**

To devise an effective pricing structure, it is important to study the effects of change in own price & on sales in relation with the competitors.

**Self-Price Elasticity:** It’s the measure to know the responsiveness of the sales of a product to a change in its price when nothing but the price changes

**Cross-Price Elasticity:** It’s the measure to know the responsiveness of the sales of a product to a change in the price of the competitor brand when nothing but the price changes.

**Data Preparation**

Since, the liquid laundry detergent come in different pack-sizes, we will need to first get “per lbs / per pound” price. And since there several products under one laundry-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 each store each week.

Steps for Data preparation:

Sale data is in two files: laundet\_drug\_1114\_1165 & laundet\_groc\_1114\_1165.Total sales of drug & grocery stores from separate file is joined in one single dataset.

1. Comparison is done between our brand “All” & market leader “Tide”.
2. Weighted Price for a product is calculated for a store for a week (i)
3. Dataset now has data for each IRI\_KEY for each week for both ALL & Tide.
4. Stores for which data of all 52 weeks is available in both brands are taken to make balanced Panel data
5. The data is converted from long-form to wide-form, where we will have single observation for every week when the store is open, and the price and store-promotion details are stored in columns for both brands

**SAS Results**

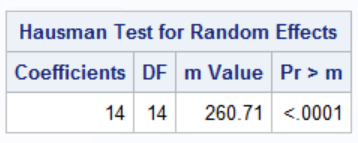
As we are now observing behavior of stores across weeks, Data is now arranged as store level weekly sales, it’s a Panel Data. We apply a simple regression (PROC PANEL) over the panel-data with “IRI\_KEY” as Cross-section ID and “week” as Time-series ID, with weighted-price and weighted-features (Feature, Display, and price-reduction) as independent variables and Total-sales of the “ALL” brand as the dependent variable.

**Hausman Test:**

To check if there is endogeneity in the model (i.e. correlation between error term & explanatory variables), we run the Hausman test with Null & Alternative Hypothesis:

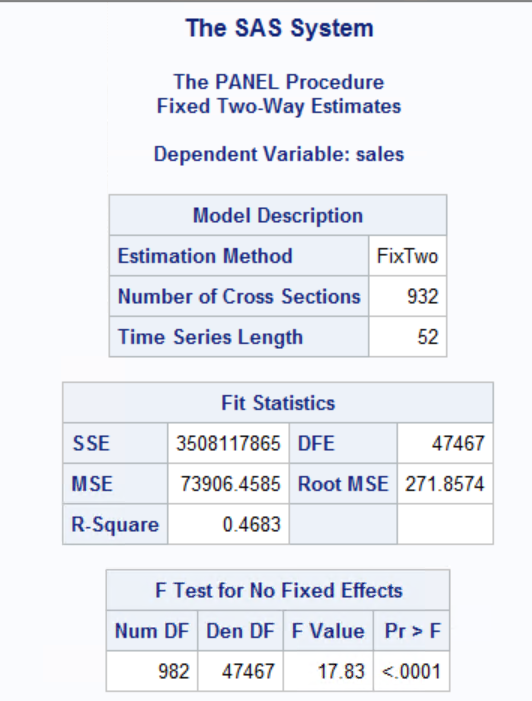
𝑯𝟎: **correlation between error term Ui and X variables**

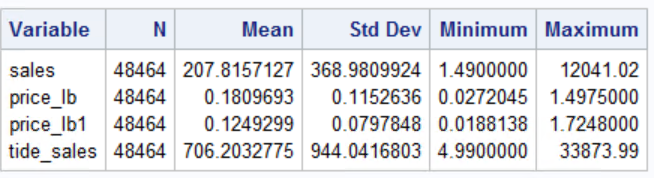
𝑯𝟏: **No correlation between error term Ui and X variables**



Since, the p-value is less than 0.05(confidence-level at 95%), so we reject null hypothesis and confirm that **Fixed-Effects model** is best model for this dataset.

1. **Model for Overall price elasticity:**





**BETA Estimates:**

**Regression equation**:

𝑆𝑎𝑙𝑒𝑠1= 𝛽0 + 𝛽1𝑃R\_wt + 𝛽2D\_wt + 𝛽3F\_wt + 𝛽4p𝑟𝑖𝑐e\_lb + 𝛽5PR\_wt1 + 𝛽6𝐷\_wt + 𝛽7F\_wt1 + 𝛽8price\_lb1 + 𝛽9price\_PR\_all + 𝛽10price\_D\_all + 𝛽11price\_F\_all + 𝛽12 price\_PR\_tide +𝛽13 price\_D\_tide + 𝛽14 price\_F\_tide

**Self-Price Elasticity:**

𝑆𝑒𝑙𝑓𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1 = (𝛽4+ 𝛽9price\_PR\_all + 𝛽10price\_D\_all + 𝛽11price\_F\_all) ∗Avg. 𝑃𝑟𝑖𝑐𝑒1 / Avg. S𝑎𝑙𝑒𝑠1

= -4.29

* From the results we observed that if All does a price reduction of 1% along with Display & Feature Ads, the sales will increase by 4.29%

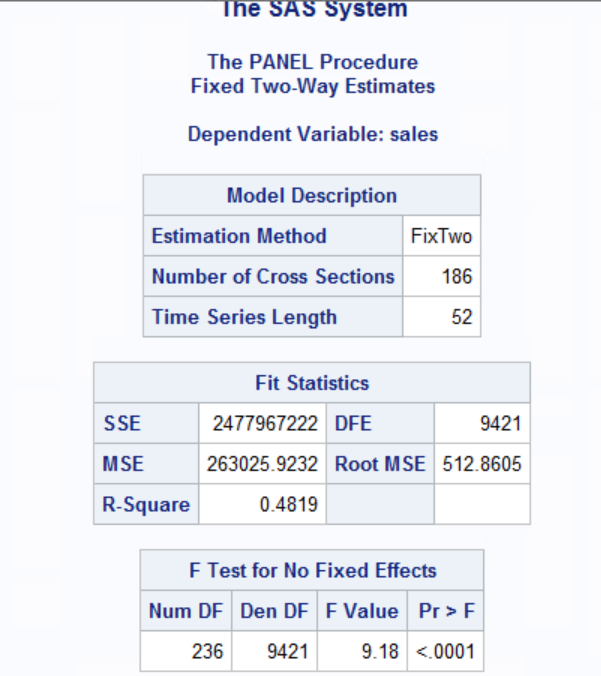
**Cross-Price Elasticity:**

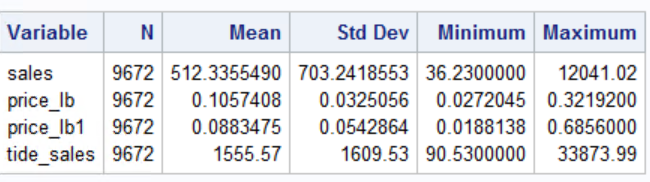
𝐶𝑟𝑜𝑠𝑠𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1= (𝛽8+ 𝛽12price\_PR\_tide + 𝛽13price\_D\_tide + 𝛽14price\_F\_tide) ∗ Avg. 𝑃𝑟𝑖𝑐𝑒2 / Avg. S𝑎𝑙𝑒𝑠1

= 0.071

* From the results we observed that if Tide (Competitor brand) does a price reduction of 1% along with Display & Feature Ads, the sales of ALL will decrease by 0.071%

1. **Model for Price elasticity for top selling stores for ALL**





**Self-Price Elasticity:**

𝑆𝑒𝑙𝑓𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1 = (𝛽4+ 𝛽9price\_PR\_all + 𝛽10price\_D\_all + 𝛽11price\_F\_all) ∗Avg. 𝑃𝑟𝑖𝑐𝑒1 / Avg. S𝑎𝑙𝑒𝑠1

= -3.81

* From the results we observed that if All does a price reduction of 1% in top 20% stores in terms of Sales along with Display & Feature Ads, the sales will increase by 3.81%

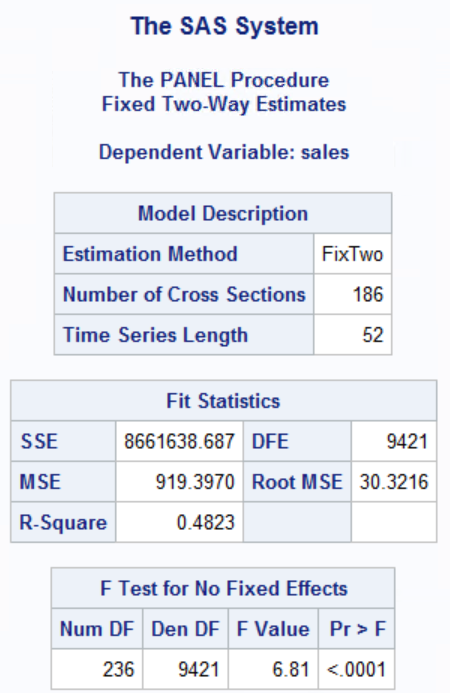
**Cross-Price Elasticity:**

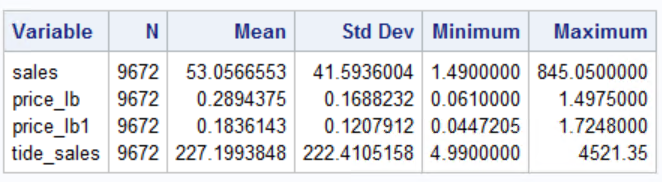
𝐶𝑟𝑜𝑠𝑠𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1= (𝛽8+ 𝛽12price\_PR\_tide + 𝛽13price\_D\_tide + 𝛽14price\_F\_tide) ∗ Avg. 𝑃𝑟𝑖𝑐𝑒2 / Avg. S𝑎𝑙𝑒𝑠1

= 0.15

* From the results we observed that if Tide (Competitor brand) does a price reduction of 1% in the top 20% stores in terms of Sales along with Display & Feature Ads, the sales of ALL will decrease by 0.15%

1. **Model for Price elasticity for bottom selling stores for ALL**





**Self-Price Elasticity:**

𝑆𝑒𝑙𝑓𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1 = (𝛽4+ 𝛽9price\_PR\_all + 𝛽10price\_D\_all + 𝛽11price\_F\_all) ∗Avg. 𝑃𝑟𝑖𝑐𝑒1 / Avg. S𝑎𝑙𝑒𝑠1

= -2.25

* From the results we observed that if All does a price reduction of 1% in bottom 20% stores in terms of Sales along with Display & Feature Ads, the sales will increase by 2.25%

**Cross-Price Elasticity:**

𝐶𝑟𝑜𝑠𝑠𝑃𝑟𝑖𝑐𝑒 𝐸𝑙𝑎𝑠𝑡𝑖𝑐𝑖𝑡𝑦 𝑜𝑓 𝐵𝑟𝑎𝑛𝑑1= (𝛽8+ 𝛽12price\_PR\_tide + 𝛽13price\_D\_tide + 𝛽14price\_F\_tide) ∗ Avg. 𝑃𝑟𝑖𝑐𝑒2 / Avg. S𝑎𝑙𝑒𝑠1

= 0.14

* From the results we observed that if Tide (Competitor brand) does a price reduction of 1% in the bottom 20% stores in terms of Sales along with Display & Feature Ads, the sales of ALL will decrease by 0.14%