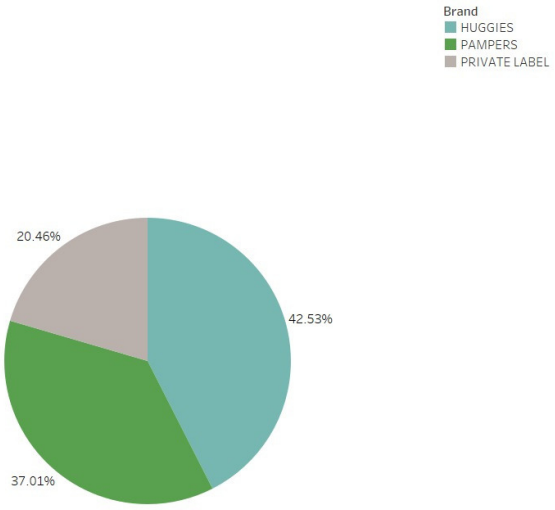
INSPIRED BY BABIES, CREATED BY PAMPERS

TEAM 4

WHAT IS THE DATA?

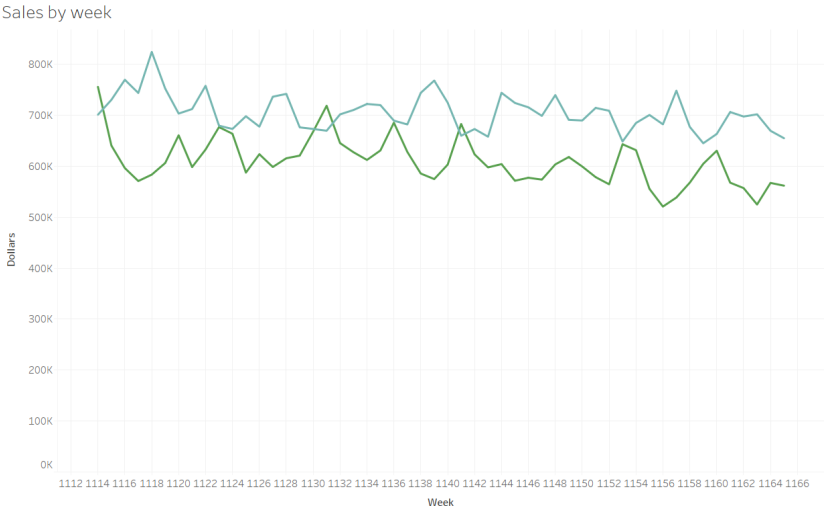
The analysis of the data in the project is broken down into two major phases, the store level behavior of data and the household level behavior of data. In order to understand how the sales behave in various contexts, exploratory visuals were plotted, for both these phases.

Market Share of Top 3 Brands



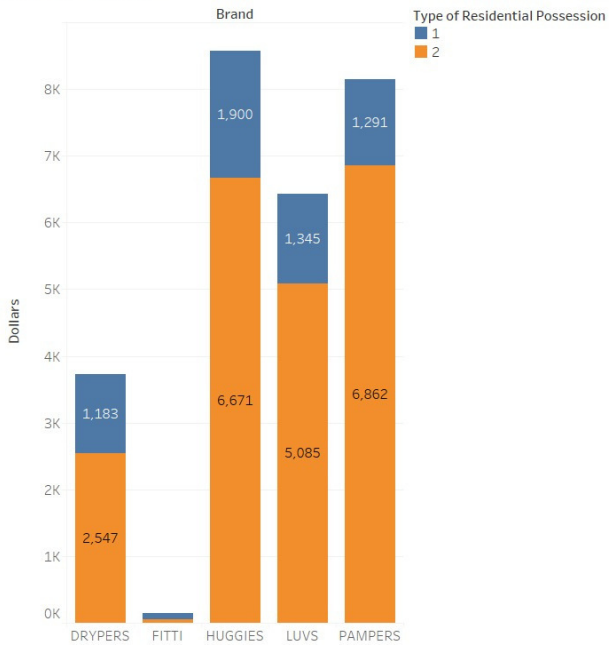
The market share of the dollar sales rolled up to the brand level makes it clear that our brand Pampers, is a competitor with Huggies in terms of Market share.

Sales by week



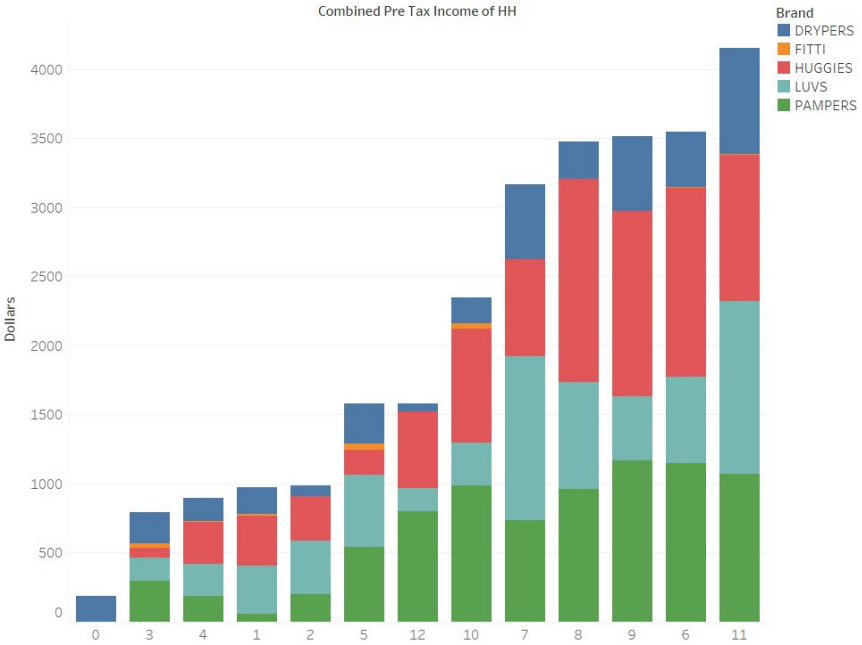
The sales of each of Pampers and Huggies varies approximately sinusoidal and out of phase.

Behaviour of buyers by housing accross barnds



People that own houses buy more diapers

Total Sales v/s Income Category across Brands



Income groups between 6 and 10 contribute our dollar sales. As expected, Huggies is our competitor.

HOW ARE PRODUCTS GETTING SOLD?

UNDERSTANDING THE IMPACT OF STORE PRESENCE AND OFFERS:

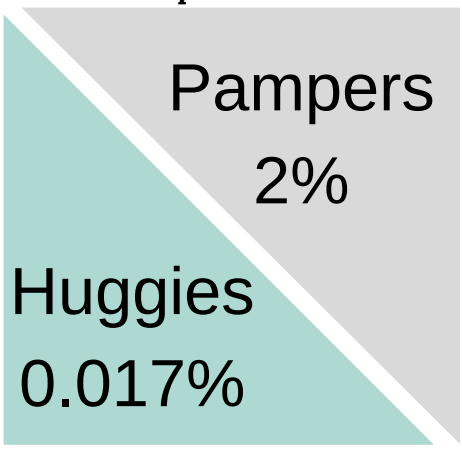
In order to recommend an optimum pricing and advertising strategy for stage 4 product, multiple linear regression was carried out using the model below:

lntotal_sales= weighted_price_per_diaper+weighted_price_per_diaper2+ weighted_price_per_diaper_HU+ weighted_Medium_Ad+weighted_Large_Add+ weighted_Cou_Reb + weighted_Min_Dis + weighted_Maj_Dis+ weighted_PR+ MAJ_DIS_MED_ADD + COU_MAJ_DIS+COU_MIN_DIS+ MIN_DIS_LAR_ADD+ PR_COU+PR_MIN_DIS+ PR_MAJ_DIS+ weighted_Maj_Dis_HU+ weighted_Large_Add_HU

Pricing Impact:

As the price of a diaper increases by 0.1 dollar, the total dollar sales show a quadratic (non-linear effect).
For every 0.1$ increase in the price per diaper then sales decreases by
-33.82+((2*68.31) *0.094)
=-20.97/10=-2.09 i.e. ~ 210%

Price Elasticity of Pampers in comparison with:

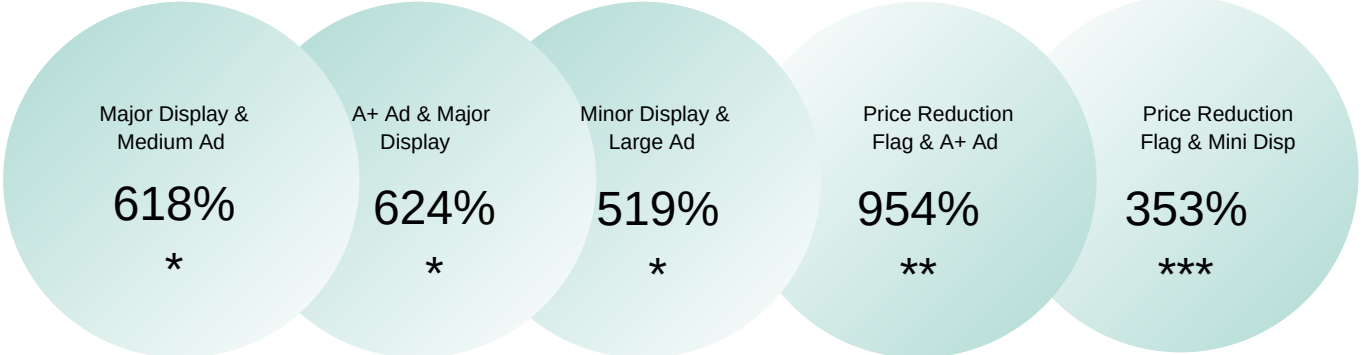


Advertising and In store Marketing Impact:

Compared to the brand with No Advertising
>The total dollar sales with medium Ad increase by 153%, and Large Ad increase by 120%, if there is no effect from interaction terms involved.
>The total dollar sales with A+ ad (retailer coupon or Rebate) increases by 88.2% if there is no effect from interaction terms involved.
Compared to the brand with No Display
>The total dollar sales with minor display increases by 71% and with major display increases by 150% if there is no effect from interaction terms involved.

- **1% decrease in our price the sales decreases by 1.97 ~ 2%**
- **1% increase in price of Huggies our sales increases by 0.017%**

Studying the interactions of various price and marketing strategies:



* more as compared to No Display and No Ad ** more as compared to No Price Reduction and No Ad *** more as compared to No Price reduction and No Display

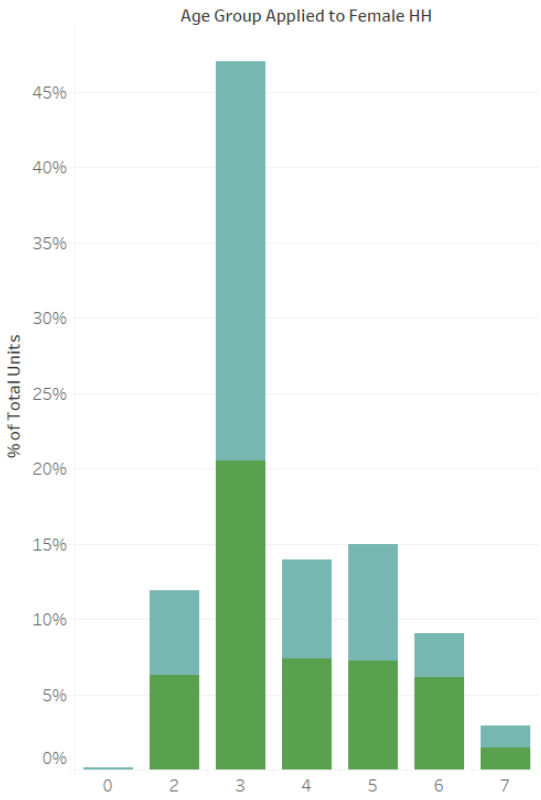
NOTE: The coefficient of PR (Price reduction) should be positive as per our understanding but it's coming as negative, but we also see that the along with price reduction if the product is on display or feature then it has an overall positive impact on sales

WHO IS BUYING?

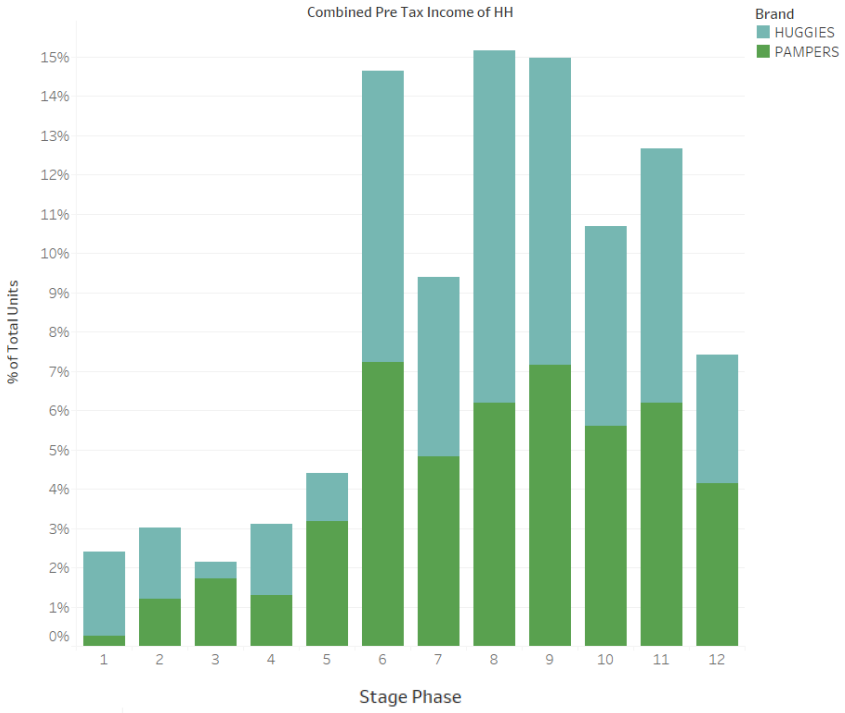
Having a good understanding of which fractions of the demographic that sales are coming from can help generate advertising strategies that pierce through the competitors.

From running the regression on some price and demographic variables, we see that 10 cents increase in unit price of a pampers diaper will decrease the number of diapers bought weekly by 12.9 units for a person. However, the distribution of sales in the given data, shows that families having age group of male and female Head of Households in the age group of 35-44 buy or products the most. The market in the lower income groups however is dominated by our competitors. Babies in stage 4 contribute to the maximum market share. There also is enough room to acquire competitor customers in this stage as it is a sizeable proportion of the economy. We seem to have imputed data for the location, however, we can see that our competitor has captured a bigger market share in the zipcode 00003.

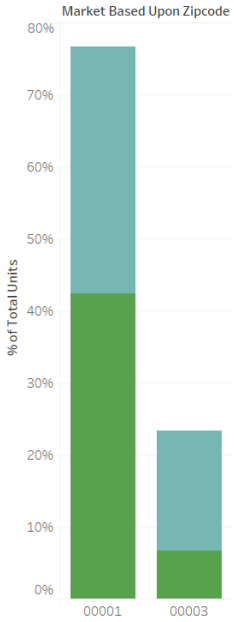
Age of Male and Female HH

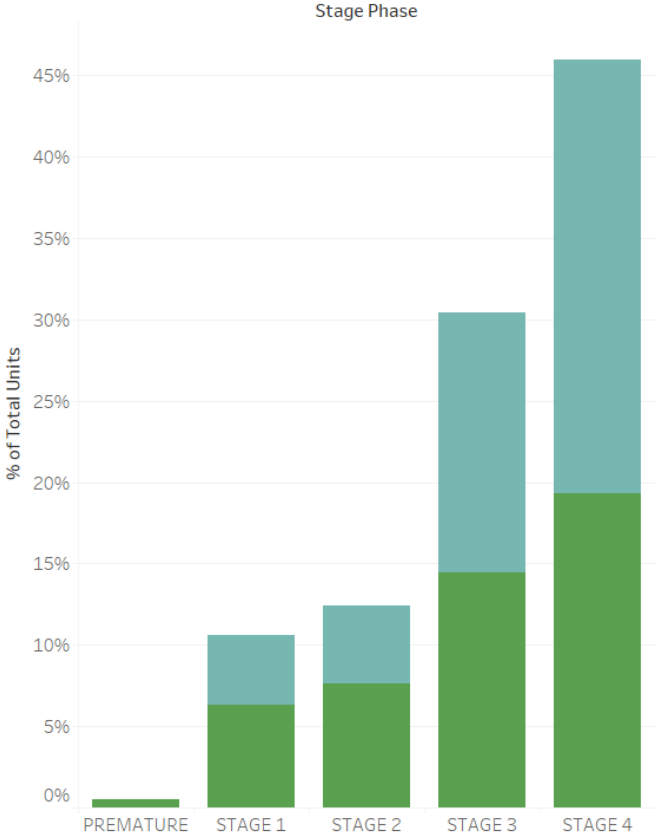


Income and Sales



Location

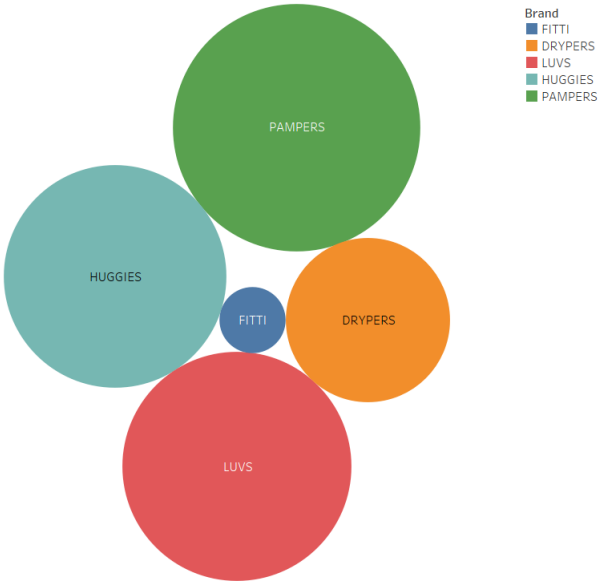




UNDERSTANDING OUR BUYERS

To understand the customer behavior, our buyers were segmented by performing RFM analysis.
From the Correlation matrix R, F & M we found that F and M are highly correlated, are considering only R and M scores to calculate total RFM scores in identifying customer loyalty
The total RFM score>=8 with M and R component >=4 each is considered as Loyal customer

Loyalty by brand



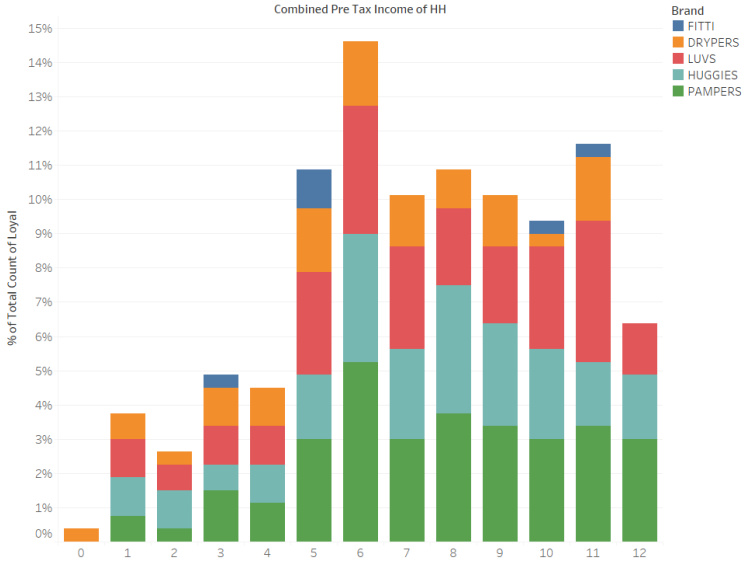
LOYALTY BY BRAND

We see that our brand Pampers has the highest loyalty amongst customers.
As identified previously, Huggies is our competitor and follows closely.

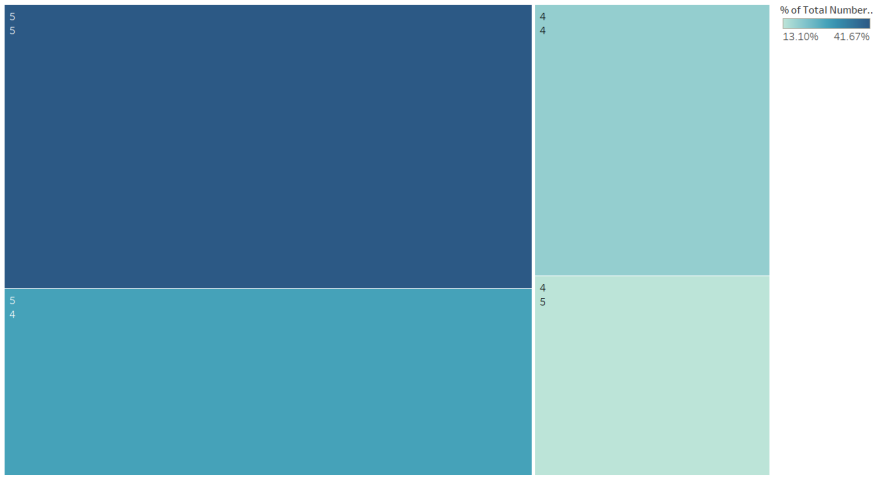
LOYALTY BY INCOME

The distribution of loyal customers by income groups shows that age group 6 and 11 are the highest contributers to loyalty. This income range between $25k and $75k is cumilatively loyal to the brands they buy. Thus aggressive marketing targeting these income groups should ensure long term income to the brand.

Loyalty by income



How recently and how much?



BREAKING IT DOWN TO RECENCY AND MONETARY ASPECTS

Highest proportion of our loyal customers show encouraging behavior. Customers doing well with both, recency and monetary aspects form the highest proportion of loyal customers. However, the customers with a high recency score (5) and low monetary score (4) may addressed by improving the store presence with featuring our brand.

WILL THEY STAY?

In order to retain the acquired customers, we study the attributes of the customers that stay and the attributes that make others leave:

The model used is:
Loyalty=b0+b1*inc1+b2*inc2+b3*resident+b4*cgc+b5*Fage1+b6* Fage2+b7* Fage3+b8* Fage4+b9* Fage5+b10* Mage1+b11* Mage2+b12* Mage3+b13* Mage4+b14* Mage5+b15*edu1 +b16*edu2

Fage2 – Compared to households with female heads with 65+ age, the female heads between the age 25-34, the odds of loyalty of a customer increase by 91.5%
Fage3- Compared to households with female heads with 65+ age, the female heads between the age 35-44, the odds of loyalty of a customer increase by 82.2%
Resident- compared to a customer renting their home, for the customers owning their house the odds of loyalty of a decrease approximately by 46%
CGC- compared to a customer with children in age range [0,5),[6,11),[12,17) the odds of loyalty of a customer decreases by approximately 44%

HIT RATIO	CONCORD-ANCE	AUC
64.8%	64.9%	0.65
percentage of accurate predictions.	all possible pairs and select loyal (1) and not loyal (0), calculate probability for both. If Prob(loyal)>Prob (Not loyal), then it is concordant.	lift of 15% in our model compared with the Naïve model (model with no explanatory variables)