Carbo Loading Dataset

Shivee Singh

Data Summary

- Four commodities - Pasta, Pasta sauce, Pancake mixes, Syrups
- 927 different products
- 387 stores
- Over 500k unique households who made over 3 million purchases.
- Over 5 million products purchased and total sales of 9 million USD over a period of 2 years.
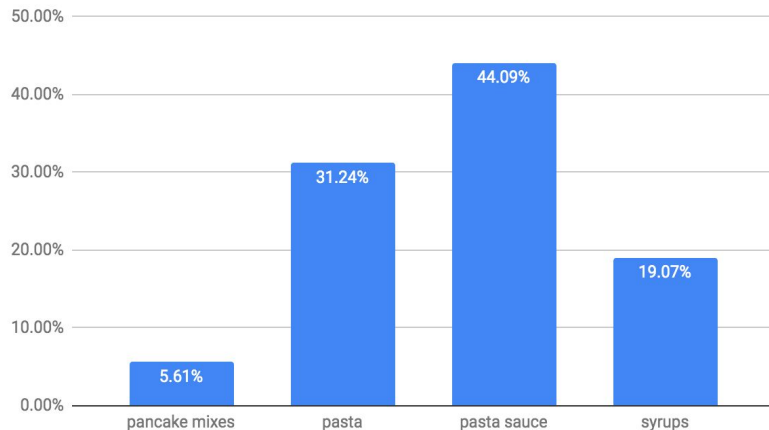
Average daily transactions: 4,100
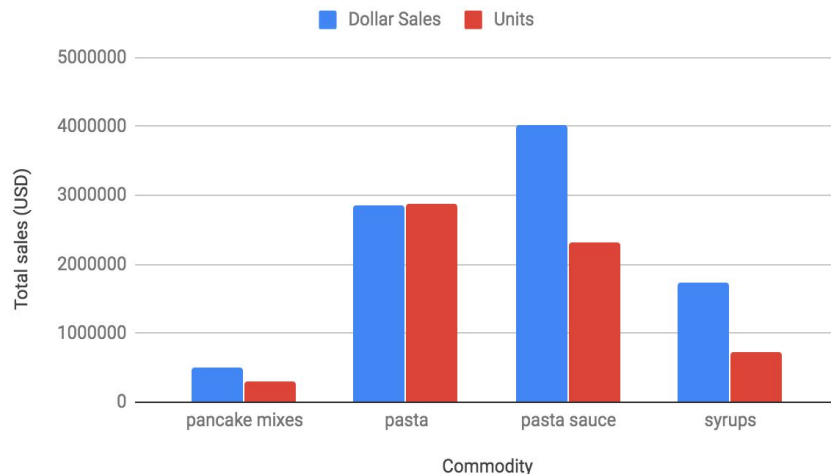
Average daily sale value: 12,500 dollars

Average purchase value per household: 18 dollars

Sales across commodities

Sales composition

Sales vs units across commodities

- 75% of the sales is composed of pasta and pasta sauce.
- Pasta sauce is highest revenue generator with 4 million in revenue.
- A single unit of pasta sauce generates more revenue than a single unit of pasta. The same is true for syrups vs pancakes.

Top 5 Pasta Products

product_description	% Pasta sales
PRIVATE LABEL THIN SPAGHETTI	2.97%
PRIVATE LABEL SPAGHETTI REGULAR	2.82%
PRIVATE LABEL ANGEL HAIR PASTA	2.32%
PRIVATE LABEL SPAGHETTI THIN	2.16%
PRIVATE LABEL SPAGHETTI REGULAR	2.12%

- All are private label.
- Top 5 products constitute 12.4% of the total pasta sales revenue.
- 257 out of 448 products in pasta category account for less than 0.1% of the revenue in this category.

Top 5 Pasta Brands

Brand	% Pasta sales
Private Label	34.90%
Barilla	15.21%
Creamette	9.52%
Mueller	9.49%
Ronzoni	7.49%

- Private labels account for ~35% of the pasta sales.
- This means customers are not very brand conscious in pasta category and are probably more price sensitive.
- There are 56 brands, top 10 brands accounted for ~95% of the pasta sales.

Top 5 Pasta Sauce Products

product_description	% Pasta sauce sales
RAGU TRADITIONAL PLAIN	3.64%
PREGO REG SPAGHETTI SAUCE	2.87%
RAGU OWS SPAG SAUCE MEAT	2.22%
RAGU SPAGH SAUCE PLAIN	2.00%
RAGU\CHZ CREATION ALFRDO	1.97%

- Ragu and Prego are the most popular products in pasta sauce category.
- Top 5 products constitute 12.7% of the total pasta sauce sales revenue.
- 136 out of 294 products in pasta sauce category account for less than 0.1% of the revenue in this category.

Top 5 Pasta Sauce Brands

Brand	% Pasta sauce sales
Ragu	37.29%
Prego	19.87%
Classico	10.89%
Private Label	8.23%
Bertolli	6.41%

- Ragu and Prego are the most popular brands in pasta sauce category and account for 57% of the total pasta sauce sales.
- Customers are brand conscious in the case of pasta sauce.
- There are 43 brands, top 10 brands accounted for 99% of the pasta sales.

Top 5 Pancake Mix Products

product_description	% Pancake sales
AJ BUTTERMILK PANCAKE MIX	18.23%
AUNT JEM ORIGINAL PANCAKE MIX	10.94%
A/JEM COMPLETE PANCAKE MI	9.78%
PRIVATE LABEL COMPLETE PANCAKE MIX	8.78%
H J PANCK BTRMLK COMP MIX	8.60%

- AJ and Aunt Jem are the most popular products in pancake mix category.
- Top 5 products constitute 56.3% of the total sales revenue in the pancake mix category.
- There are 46 products in this category and top 10 account for more than 77% of the sales in this category.

Top 5 Pancake Mix Brands

Brand	% Pancake sales
Aunt Jemima	40.39%
Hungry Jack	21.77%
Private Label	14.38%
Krusteaz	5.96%
Bisquick	3.60%

- Aunt Jemima is the most popular brand in pancake mix category. It alone accounts for more than 40% of the sales in this category.
- Customers are brand conscious in the pancake mix category.
- There are 17 brands in this category and top 5 account for more than 86% of the sales in this category.

Top 5 Syrup Products

product_description	% Syrup sales
AUNT JEMIMA ORIGINL SYRUP	5.68%
PRIVATE LABEL SYRUP PLASTIC BOTTLE	5.52%
MRS BUTTERWORTH SYRUP	4.99%
PRIVATE LABEL MAPLE SYRUP	3.83%
PRIVATE LABEL BUTTERED SYRUP	3.61%

- Aunt Jemima and Private Label are the most popular products in syrups category.
- Top 5 products constitute 23.7% of the total pasta sauce sales revenue.
- 53 out of 139 products in syrups category account for less than 0.1% of the revenue in this category.

Top 5 Syrup Brands

Brand	% Syrup sales
Aunt Jemima	20.66%
Private Label	20.27%
Mrs Butterworth	9.14%
Log Cabin	7.52%
Karo	6.92%

- Aunt Jemima and Private Label are the most popular brands in the syrups category accounting for more than 40% of the syrups sales.
- Top 5 brands constitute 65% of the total sales in the syrups category.
- There are 35 brands and top 10 brands constitute 88% of the syrup sales.

Penetration level of commodities

Household penetration across commodities
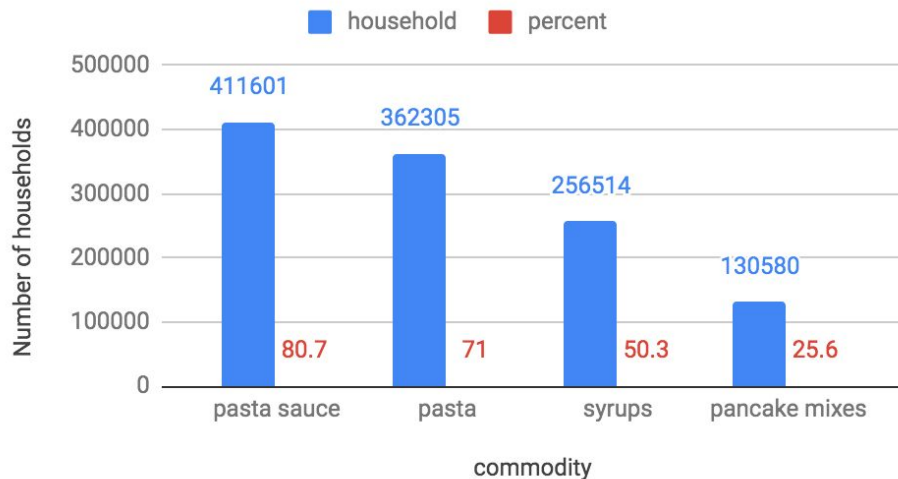
- 80% of the households have purchased pasta sauce.
- 71% of the households have purchased pasta.
- 50% of the households have purchased syrups.
- 26% of the households have purchased pancake mixes.

Commodity sales across geographies

Sales composition across geographies

Legend: syrups | pasta sauce | pasta | pancake mixes
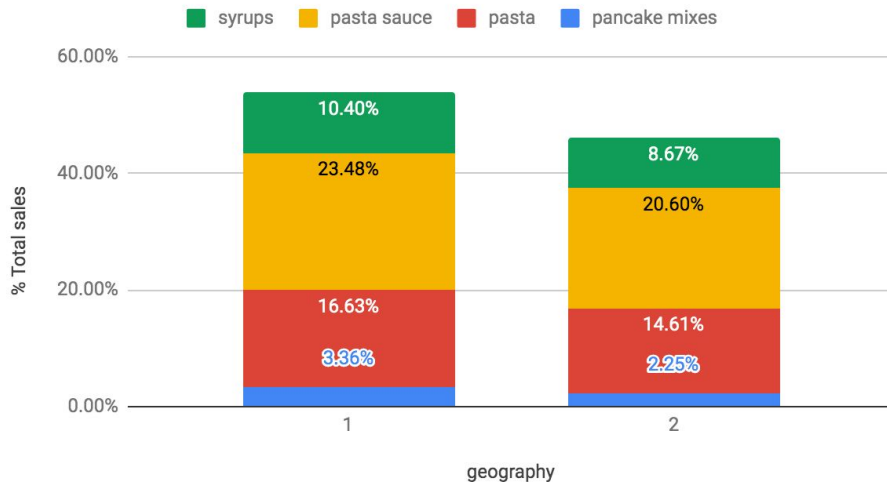
- Geography 1 has 17% higher sales as compared to geography 2.
- Sales across all commodities are higher in geography 1.

Sales distribution across households

There is no single household that contributes a major chunk to the revenue.

The top 5 households by purchase together account for less than 0.001% of sales.

Sales distribution across stores

There is no single store that contributes a major chunk to the revenue.

The top 5 stores by sales together account for less than 4% of sales.

Sales distribution across coupon and non coupon users
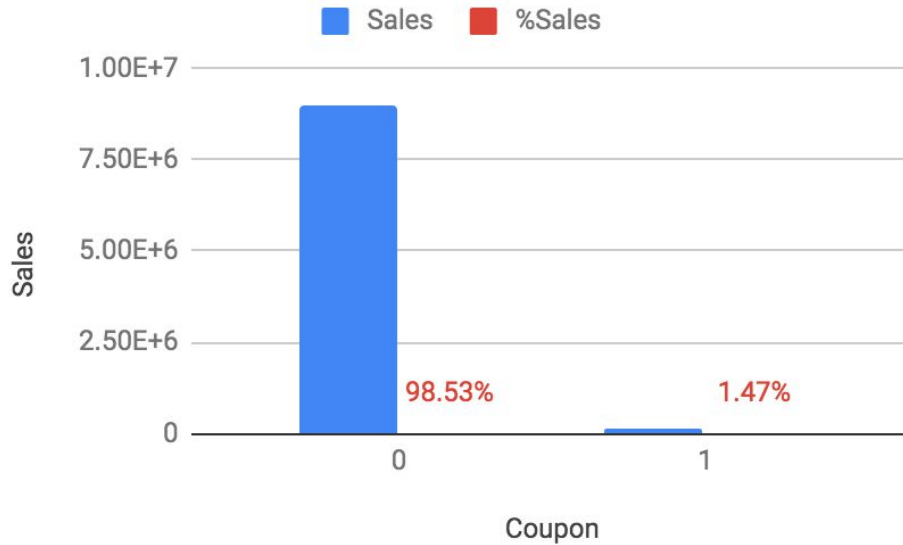
- Coupons accounted for merely 1.5% of the total sales revenue.
- 42028 households used coupons for a purchase, which is 7.6% of the total households.

Coupon usage across commodities

Commodity	% Coupon sales
pancake mixes	4.89%
pasta	12.13%
pasta sauce	55.29%
syrups	27.69%
Total coupon sales	**100.00%**

Brand	% Coupon sales
Ragu	25.05%
Bertolli	10.87%
Prego	10.61%
Aunt Jemima	10.50%
Hungry Jack	8.49%

- Pasta sauce constitutes 55% of the coupon sales.
- Syrups constitute 28% of the coupon sales.
- There can be two possibilities:
  - Customers tend to use coupons for trying out new types of pasta sauces and syrups.
  - Coupons offer heavier discounts/better deals on pasta sauces and syrups.
- 25% of the coupons sales involved purchases of Ragu brand (pasta sauce).
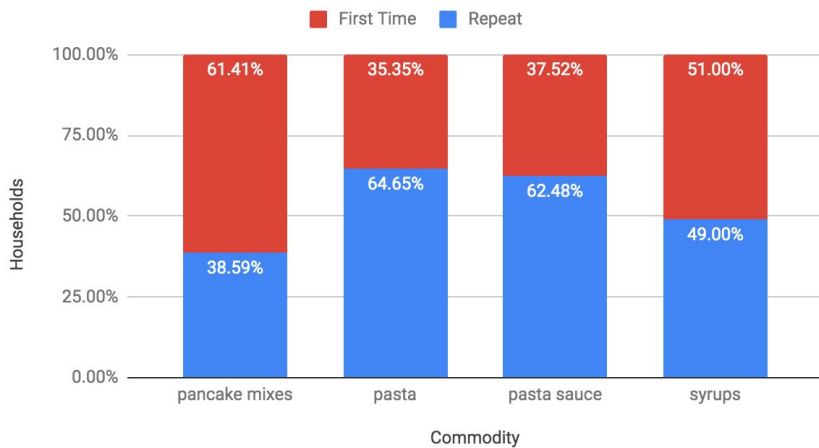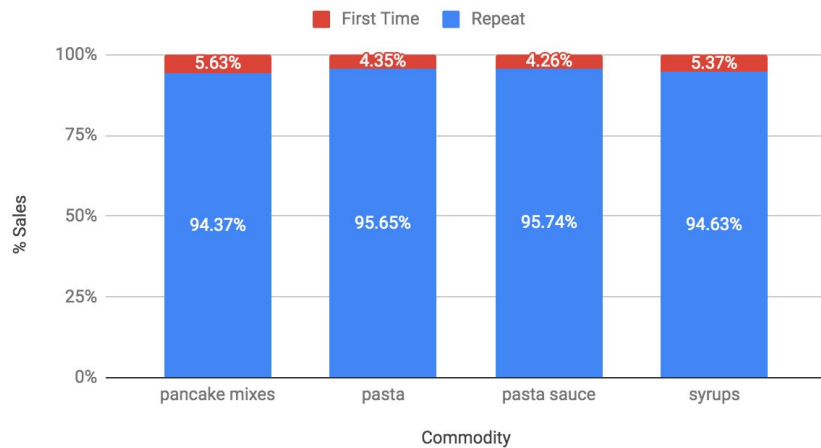
Repeat Customers
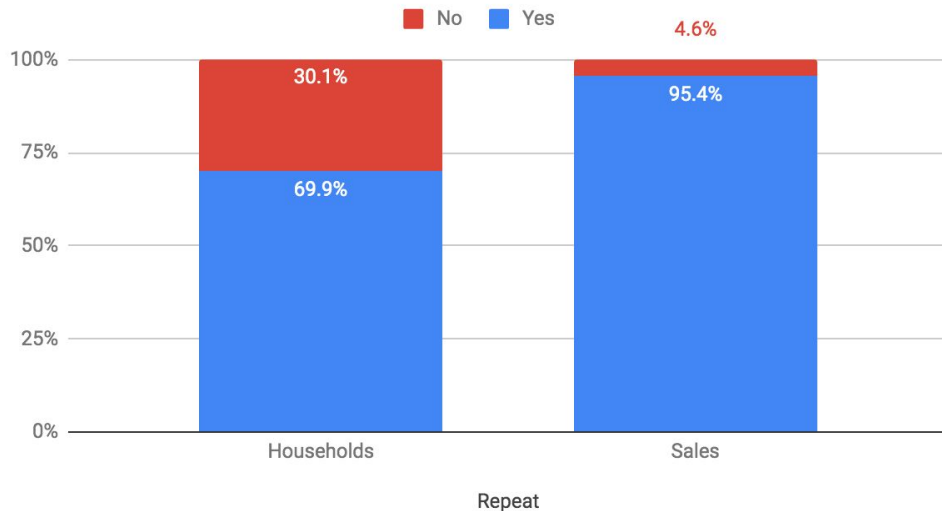
Repeat customers

- 70% of the households are repeat customers (They have purchased more than once)
- 95% of the total sales is generated by repeat customers.
- Repeat customers are very important for the company.

Value generated by repeat customers for different commodities

Composition of repeat customers across commodities

Sales generated by repeat customers across commodities

- More than 60% of the households purchasing Pasta and Pasta sauces are repeat customers.
- The repeat customers account for 96% of the total sales in Pasta and Pasta sauce category.
- This reinforces the fact that repeat purchasers are extremely important and more so for Pasta and Pasta sauce categories.

Effect of coupon on repeat purchase conversion

- 7860 customers (19% of the coupon users) used coupon for their first purchase.
- 6314 of the customers who used coupon for their first purchase became repeat customers.
- This gives a conversion ratio of 80% for first time coupon users.

Market Basket Analysis

Percentage transactions where both pasta and pasta sauce are in same basket = ~25%

Percentage transactions where both pancake and syrups are in same basket = ~3%

Proposed Machine Learning

- Develop ML model to predict whether a household shall become a repeat customer or not from their first purchase.
- Binary classification problem
- Aggregate features at the household level. Some features can be as follows-
  - Store , geography
  - Time of purchase, Day of week
  - Purchase value
  - Coupon used (1/0)
  - Units of each commodity (pasta, sauce, syrup, pancake)
- Label shall be repeat or not (1/0).
- This model would help us identify potential repeat customers.
- Since repeat customers are crucial for us, we can focus our efforts on these customers through personalized marketing or promotional offers.

Sales trend over the given period in pasta and pasta sauce category

Daily pasta sales
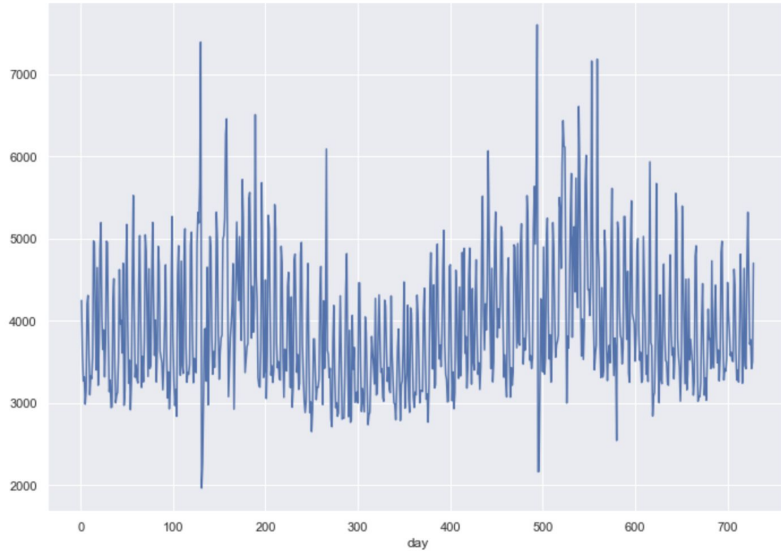
Daily pasta sauce sales
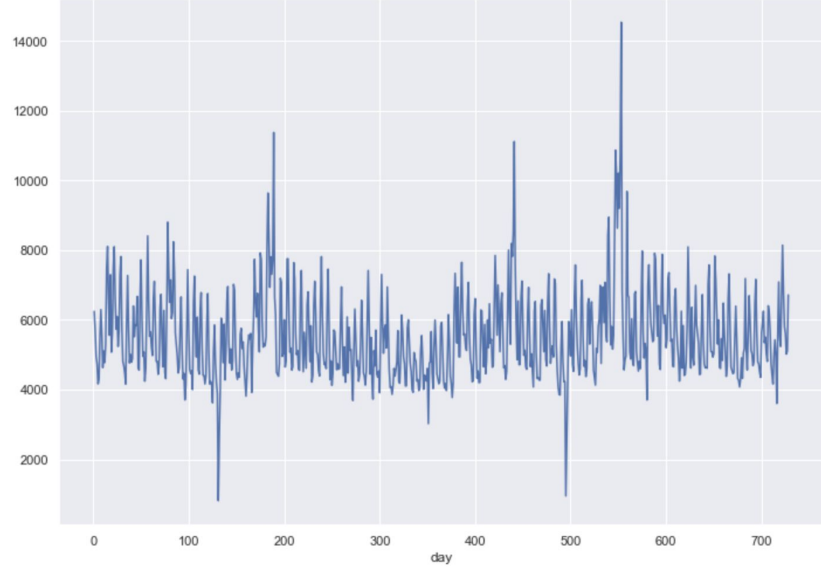
- Pasta sales are more volatile and have a cyclical pattern.
- During few times of the year there are large spikes in the demand for both pasta and sauce.
- Year on year sales for pasta and pasta sauce grew by 5.6%
- Year on year sales for pasta grew by 6.5%

Weekly sales over the given period for all categories

- Pasta and sauce sales are almost in line.
- Interesting thing to note is that when pasta and sauce demand declines it is accompanied by increase in syrups and pancake mixes demand.

Causal Analysis

Display Description	% Pasta sales	% Pasta sauce sales
Not on Display	56.85%	59.81%
Rear End Cap	10.35%	14.56%
In-Shelf	13.12%	9.24%
Front End Cap	4.30%	5.64%
Secondary Location Display	3.38%	2.14%
In-Aisle	4.68%	2.22%
Promo/Seasonal Aisle	1.77%	2.32%
Mid-Aisle End Cap	2.11%	2.32%
Store Rear	2.18%	0.90%
Store Front	1.09%	0.73%
Side-Aisle End Cap	0.17%	0.13%
Grand Total	**100.00%**	**100.00%**

- I do not see a direct effect of display on sales of pasta and pasta sauce.

Causal Analysis

Feature description	% Pasta sales	% Pasta sauce sales
Interior Page Feature	43.34%	44.24%
Not on Feature	32.68%	21.88%
Front Page Feature	8.61%	23.53%
Wrap Interior Feature	5.72%	3.81%
Back Page Feature	3.03%	2.59%
Wrap Front Feature	3.18%	1.07%
Wrap Back Feature	3.46%	0.85%
Interior Page Line Item		2.03%
Grand Total	**100.00%**	**100.00%**

- I do not see a direct effect of position on weekly mailer on sales of pasta and pasta sauce.

Conclusion

- Customers are not brand conscious in pasta category.
- We can investigate whether they are brand loyal or not.
- We should focus on repeat customers.
- We should focus on bundled deals with pasta sauce.
- If we can promote coupons on bundled deals then we can increase pasta sales.
- Pasta sales are high during weeks 25-27 every year, we could think of offering deals to first time purchasers during that time.