****

**THE COMPLETE JOURNEY**

USER GUIDE

**Lorem ipsum dolor sit amet**

**THE COMPLETE JOURNEY**

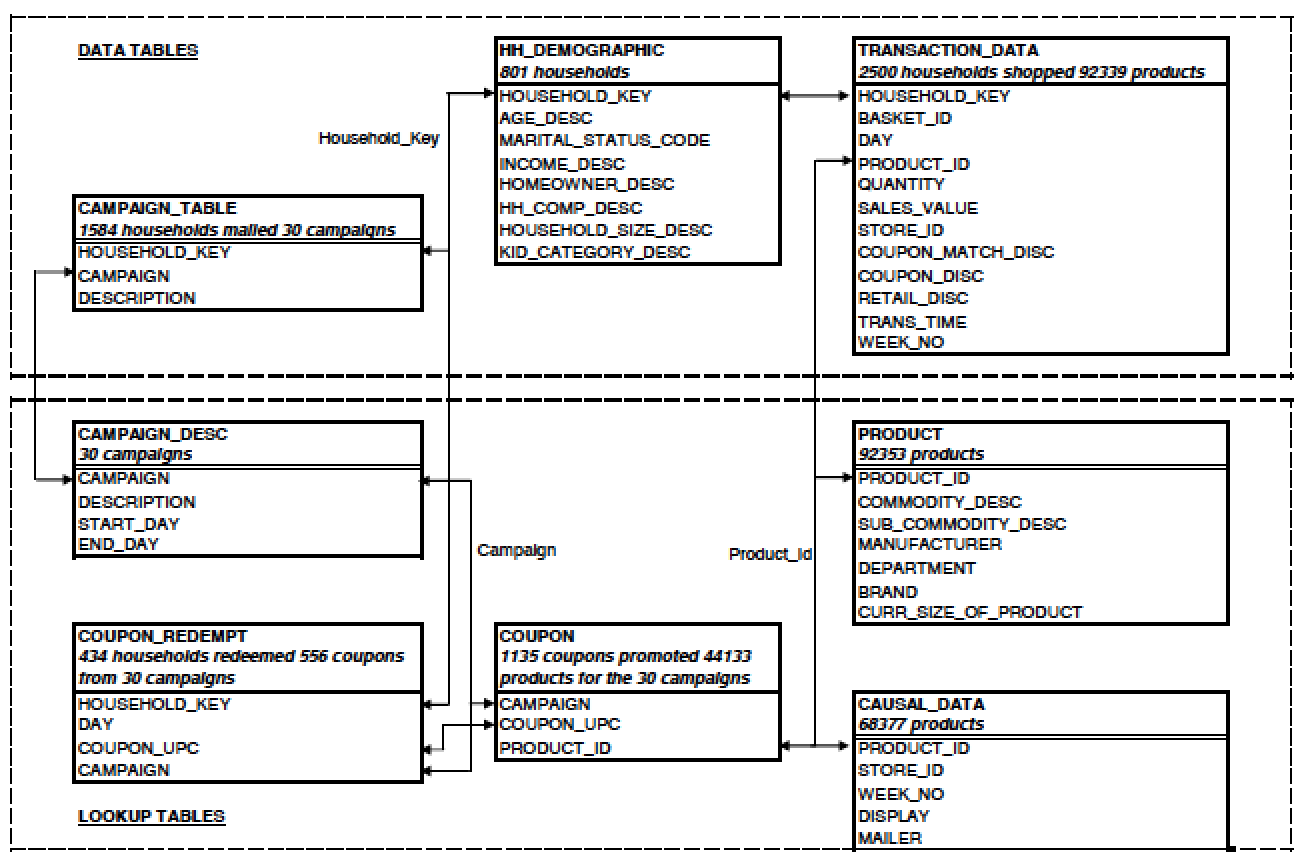
This dataset contains household level transactions over two years from a group of 2,500 households who are frequent shoppers at a retailer. It contains all of each household’s purchases, not just those from a limited number of categories. For certain households, demographic information as well as direct marketing contact history are included.

Due to the number of tables and the overall complexity of The Complete Journey, it is suggested that this database be used in more advanced classroom settings. Further, The Complete Journey would be ideal for academic research as it should enable one to study the effects of direct marketing to customers.

The following are examples of questions that could be submitted to students or considered for academic research:

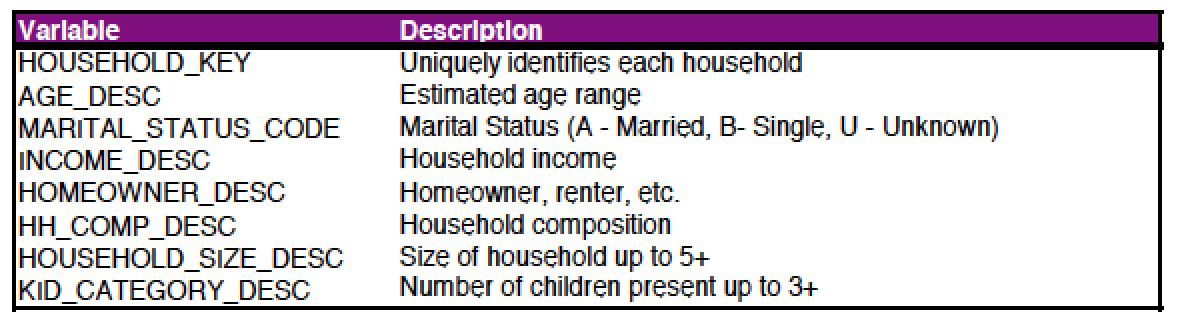
* How many customers are spending more over time? Less over time? Describe these customers.
* Of those customers who are spending more over time, which categories are growing at a faster rate?
* Of those customers who are spending less over time, with which categories are they becoming less engaged?
* Which demographic factors (e.g. household size, presence of children, income) appear to affect customer spend? Engagement with certain categories?
* Is there evidence to suggest that direct marketing improves overall engagement?

**THE COMPLETE JOURNEY: DATASET DETAILS**



hh\_demographic

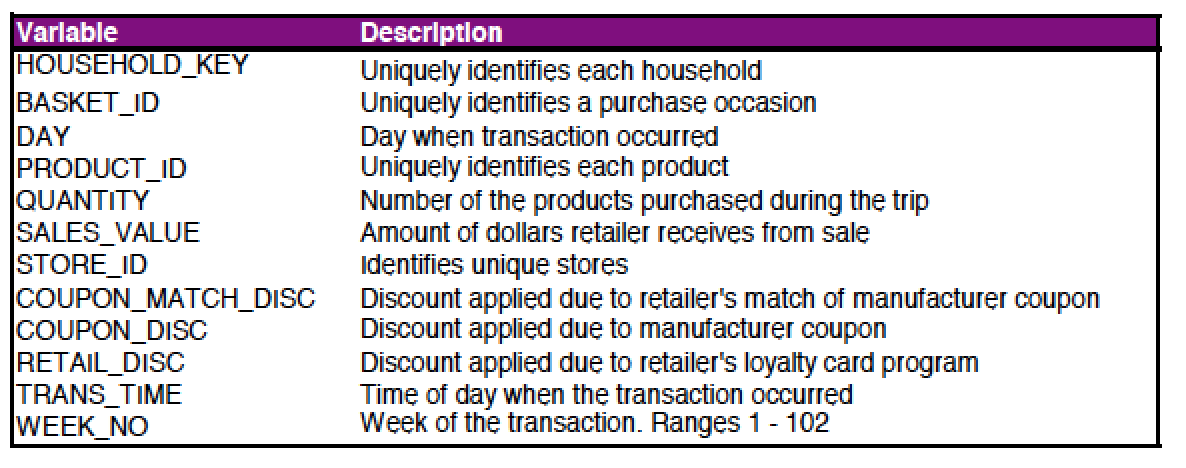
This table contains demographic information for a portion of households. Due to nature of the data, the demographic information is not available for all households.



**THE COMPLETE JOURNEY:   
DATASET DETAILS**

transaction\_data

This table contains all products purchased by households within this study. Each line found in this table is essentially the same line that would be found on a store receipt.



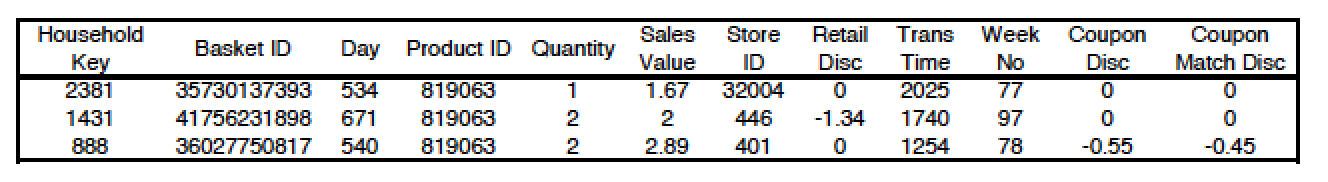
The variable sales\_value in this table is the amount of dollars received by the retailer on the sale of the specific product, taking the coupon match and loyalty card discount into account. It is not the actual price paid by the customer. If a customer uses a coupon, the actual price paid will be less than the sales\_value because the manufacturer issuing the coupon will reimburse the retailer for the amount of the coupon.

To calculate the actual product prices, use the formulas below:

* Loyalty card price = (sales\_value – (retail\_disc + coupon\_match\_disc))/quantity
* Non-loyalty card price = (sales\_value – (coupon\_match\_disc))/quantity

The example below demonstrates how to calculate the actual shelf price of the product:

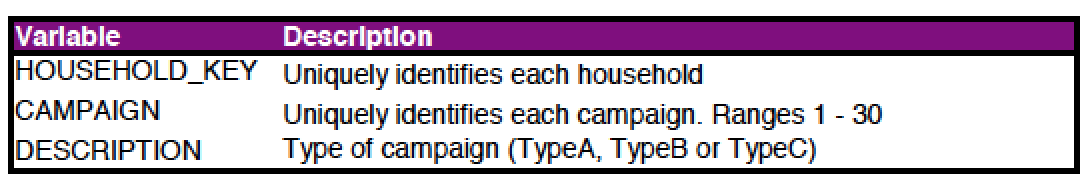
* Line 1 – When this product was purchased the retail\_disc and coupon\_disc were both zero, meaning the price of the product is the same as the amount received by the retailer.
* Line 2 – Two items of this product were purchased, and there was a retail discount applied due to a loyalty card. To determine the regular shelf price of the product (exclusive of loyalty card discount) we take the sum of the amount paid and the discount, then divide it by the quantity. ($2 + $1.34)/2 = $1.67. The shelf price of the product including loyalty card discount is $2 / 2 = $1. Also, the customer paid $2 for both of these products which is the same amount the retailer received.
* Line 3 – The actual shelf price of each product here is ($2.89 + $0.45)/2 = $1.67. Also, the customer paid $2.34 ($2.89 - $0.55) for these products, but the retailer will receive $2.89 due to the manufacturer discount.



**THE COMPLETE JOURNEY:   
DATASET DETAILS**

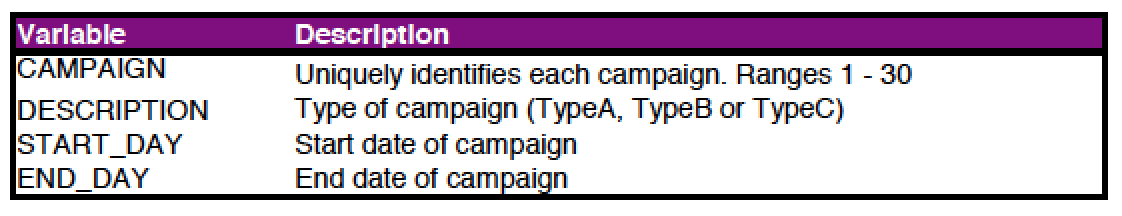
campaign\_table

This table lists the campaigns received by each household in the study. Each household received a different set of campaigns.



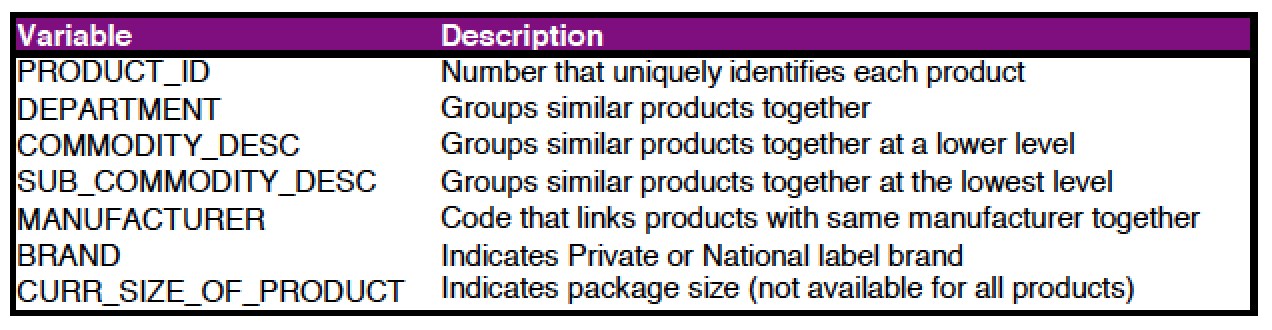
campaign\_desc

This table gives the length of time for which a campaign runs. So, any coupons received as part of a campaign are valid within the dates contained in this table.



product

This table contains information on each product sold such as type of product, national or private label and a brand identifier.

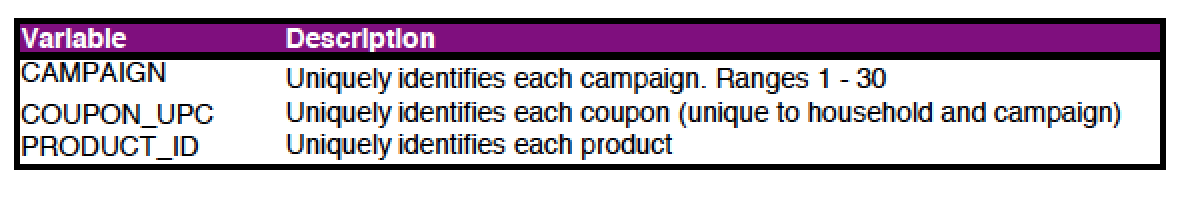


coupon

This table lists all the coupons sent to customers as part of a campaign, as well as the products for which each coupon is redeemable. Some coupons are redeemable for multiple products. One example is a coupon for any private label frozen vegetable. There are a large number of products where this coupon could be redeemed.

For campaign TypeA, this table provides the pool of possible coupons. Each customer participating in a TypeA campaign received 16 coupons out of the pool. The 16 coupons were selected based on the customer’s prior purchase behavior. Identifying the specific 16 coupons that each customer received is outside the scope of this database.

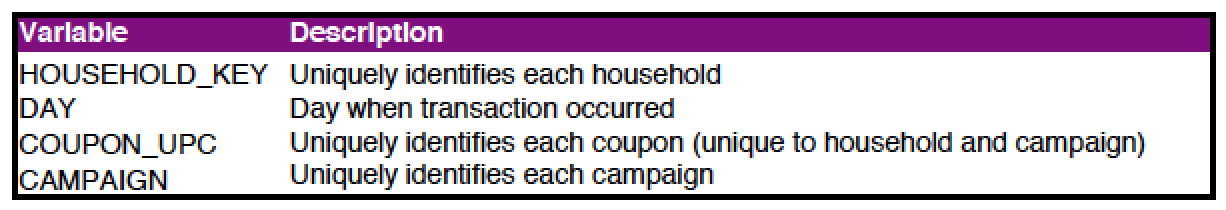
For campaign TypeB and TypeC, all customers participating in a campaign receives all coupons pertaining to that campaign.



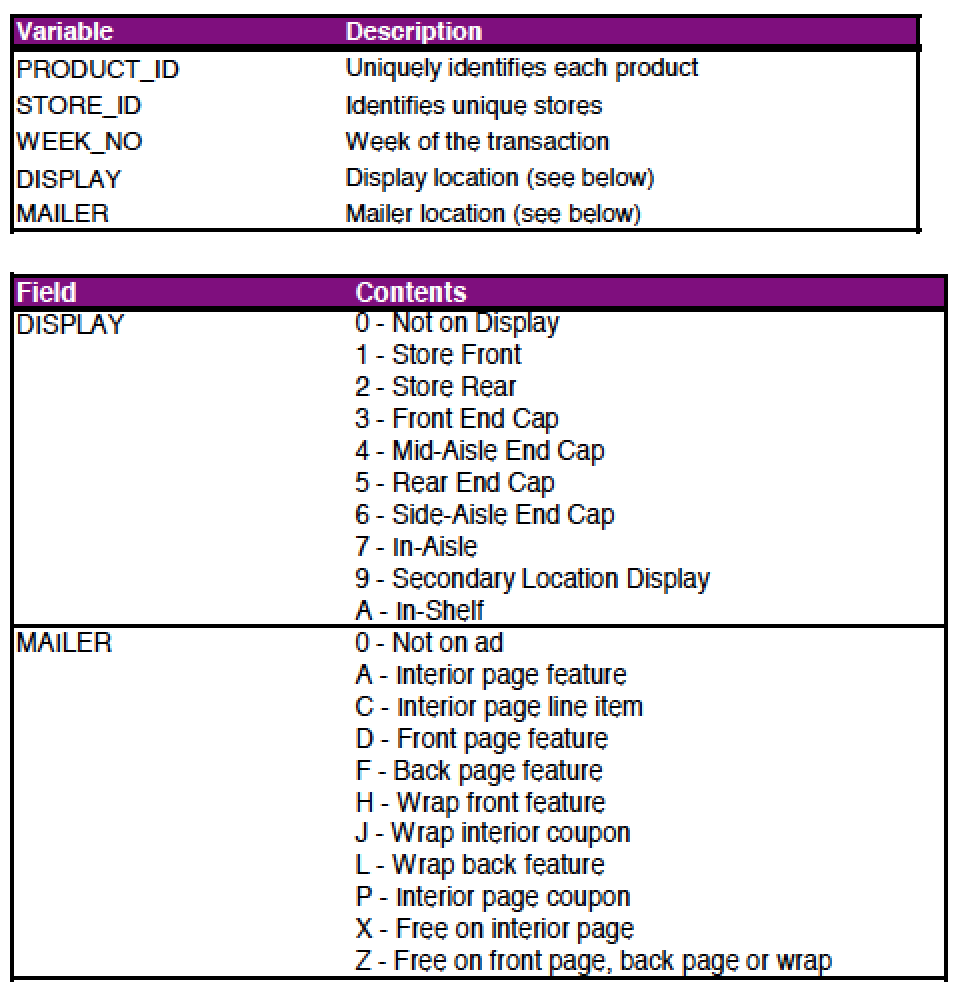
THE COMPLETE JOURNEY:   
DATASET DETAILS

coupon\_redempt

This table identifies the coupons that each household redeemed.



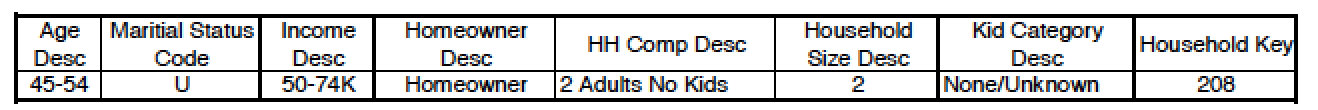
causal\_data

This table signifies whether a given product was featured in the weekly mailer or was part of an in-store display (other than regular product placement).

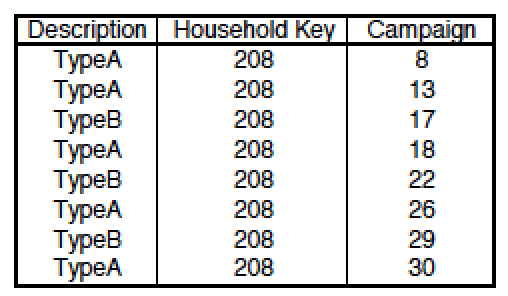
**THE COMPLETE JOURNEY:   
CASE STUDY**

John Smith is a valued customer at a national grocery retailer for which we have detailed transaction data. Throughout all the tables in the database, he is identified with a household\_key of 208.

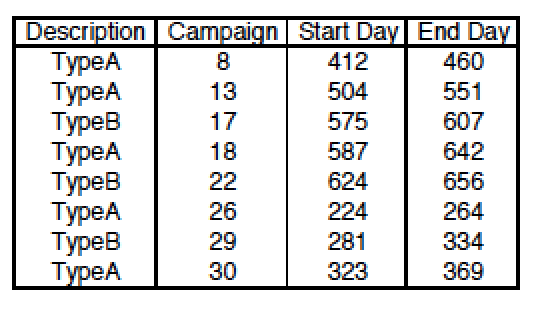
To learn a little about John, we can obtain his demographic information by looking at the record in the hh\_demographic table where household\_key = 208. The table below shows the information we receive, and tells us that he is a homeowner, who makes between $50,000 and $74,000 a year and is between 45 and 54 years old.



If we look at John’s records from campaign\_table, we can see that he received 8 different campaigns. Five of the campaigns were TypeA, and three were TypeB.

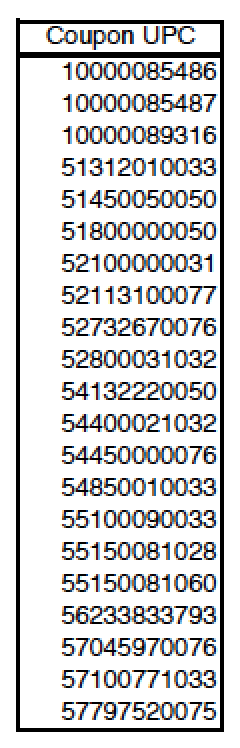


These campaigns were spread out over the 2 year period of the study. To understand the time periods of these campaigns, look at the records in the campaign\_desc table for the campaigns listed above for John.

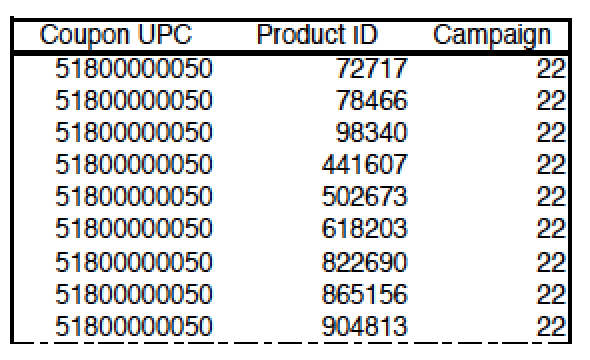


**THE COMPLETE JOURNEY:   
CASE STUDY**

Let us take a closer look at campaign 22. When we look at all the distinct coupon\_upc’s from the coupon table where campaign = 22, we see that there were 21 distinct coupons sent out as part of that campaign.



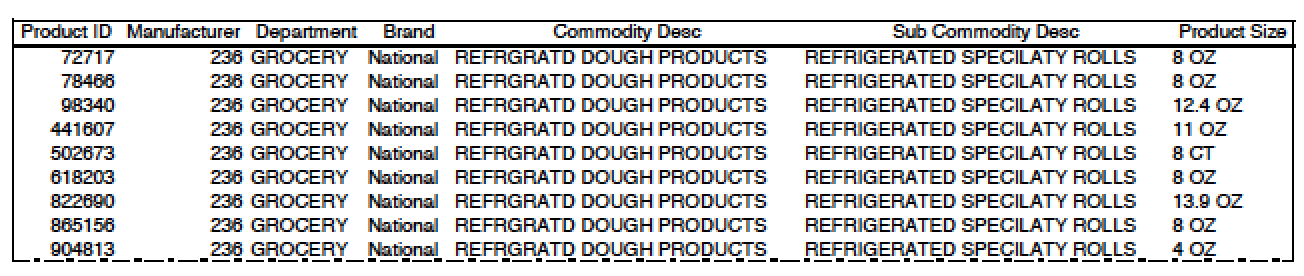
Let us take an even deeper look at one of the specific coupons offered as part of the campaign. If we print out all records from the coupon table where campaign = 22 and coupon\_upc = 51800000050, we see that this coupon could actually be redeemed for a number of products.



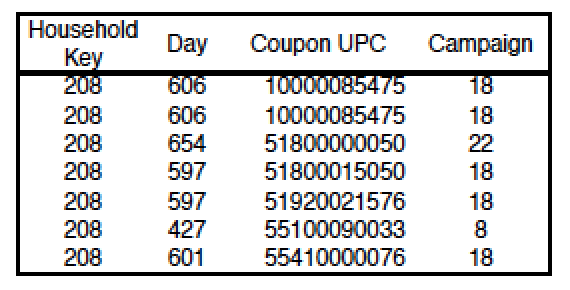
Although all the products are not displayed above, we find that this coupon is actually valid on 38 distinct products.

If we go to the product table and print out all records for the product\_id’s above (72717, 78466, etc.), we see that this coupon is valid for refrigerated specialty rolls from a national brand.

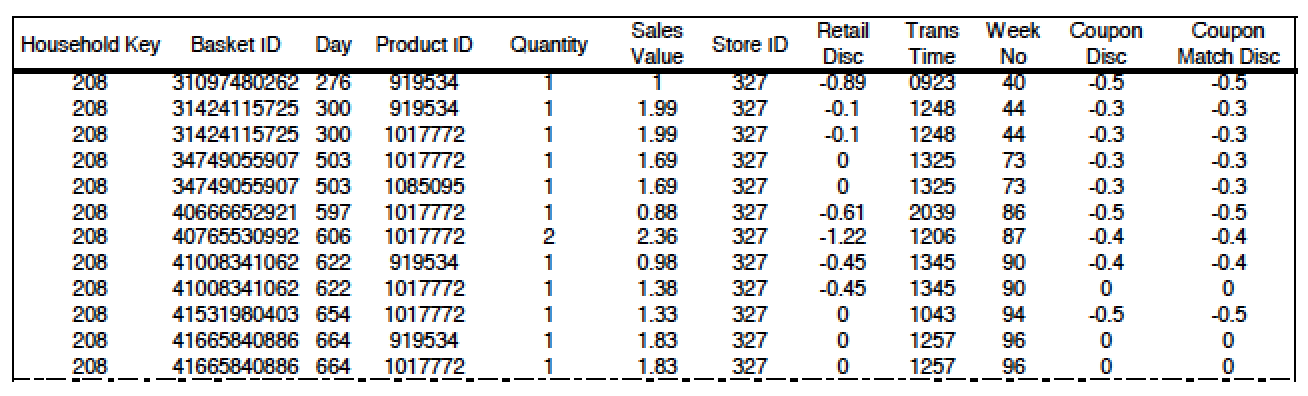
**THE COMPLETE JOURNEY:   
CASE STUDY**



As we’ve seen, John received a number of campaigns over the past two years that contained many coupons. Chances are, he did not redeem every coupon he received. So, let us take a look to see what coupons he did redeem. To do this, we need to view all records from the coupon\_redempt table where household\_key is 208. This shows us that he redeemed 7 coupons from 3 of the campaigns.



John’s coupon redemptions are only part of the overall picture of his purchasing behavior. If we look at the records from the transaction\_data table where household\_key equals 208, we can view everything that John purchased.



This gets a bit complicated, but we can combine the transaction data with the other tables to understand John’s behavior when he was redeeming a coupon (and when he wasn’t redeeming a coupon).

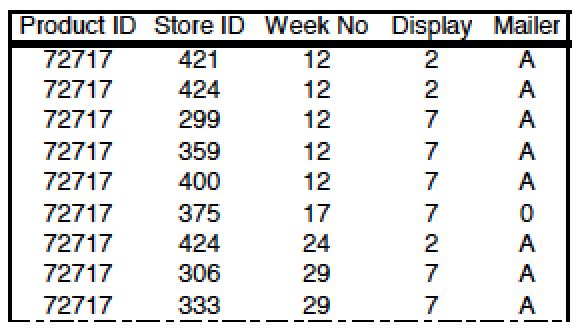
* John received offers as part of campaign 22, which occurred between days 624 and 656
* We know he redeemed coupon 51800000050 on day 654
* Through the coupon table, we know that the coupon is actually valid for a number of products,
* including product 1017772
* From the table above, we can see (3rd line from the bottom) where John purchased this item and received a discount from using a coupon

Knowing when John redeemed a coupon can help us learn a lot about him, and how the receipt of certain campaigns affected his behavior. Does the receipt of campaigns cause him to purchase more items than he did previously? Is John more likely to redeem coupons for products he already purchases, or does it entice him to try products he has never purchased before?

There is one bit of information we have not talked about yet – what is happening in the rest of the store? Is it possible that John purchased the item above because of other events occurring in the store in addition to his coupon?

We obviously do not know a customer’s reason for purchasing an item, but we do know whether an item was featured during the time of the purchase. To do this, let us look at product 72717. If we view all records from the causal\_data table where product\_id equals 72717, we see the weeks and stores where this product was featured in the weekly mailer and where it was featured as part of an in-store display. If we look at the first line, we can tell that in store 421 and week 12, the product was featured on a display in the rear of the

store and was featured on an interior page of the mailer.



We hope that this quick look at John Smith’s behavior provides clarity around The Complete Journey database, and inspires your own investigation into the purchasing behavior of these customers.

**CONTACT**

**INFORMATION**

For general questions about 84.51° or the Area 51 Files, please contact:

**REBA PERRY**

P: 513-632-0533

E: [reba.perry@8451.com](mailto:reba.perry@8451.com)