

*Final Project: Analyzing grocery shopping behavior*

*File 3 of 3*

- a. How many:
  - Store shopping trips are recorded in your database?
  - Households appear in your database?
  - Stores of different retailers appear in our data base?
  - Different products are recorded?
    - i. Products per category and products per module
    - ii. Plot the distribution of products and modules per department
  - Transactions?
    - i. Total transactions and transactions realized under some kind of promotion.
- b. Aggregate the data at the household-monthly level to answer the following questions:
  - How many households do not shop at least once on a 3 month periods.
    - i. Is it reasonable?
    - ii. Why do you think this is occurring?
  - **Loyalism:** Among the households who shop at least once a month, which % of them concentrate at least 80% of their grocery expenditure (on average) on single retailer? And among 2 retailers?
    - i. Are their demographics remarkably different? Are these people richer? Poorer?
    - ii. What is the retailer that has more loyalists?
    - iii. Where do they live? Plot the distribution by state.
  - Plot with the distribution:
    - i. Average number of items purchased on a given month.
    - ii. Average number of shopping trips per month.
    - iii. Average number of days between 2 consecutive shopping trips.
- c. Answer and reason the following questions: (Make informative visualizations)
  - Is the number of shopping trips per month correlated with the average number of items purchased?
  - Is the average price paid per item correlated with the number of items purchased?
  - Private Labeled products are the products with the same brand as the supermarket. In the data set they appear labeled as 'CTL BR'
    - i. What are the product categories that have proven to be more "Private labelled"
    - ii. Is the expenditure share in Private Labeled products constant across months?
    - iii. Cluster households in three income groups, Low, Medium and High. Report the average monthly expenditure on grocery. Study the % of private label share in their monthly expenditures. Use visuals to represent the intuition you are suggesting.