QVI

We need to present a strategic recommendation to Julia that is supported by data which she can then use for the upcoming category review. However, to do so, we need to analyse the data to understand the current purchasing trends and behaviours. The client is particularly interested in customer segments and their chip purchasing behaviour. Consider what metrics would help describe the customers' purchasing behaviour.

We have chosen to complete this task in R, however you will also find Python to be a useful tool in this piece of analytics. If you aren't familiar with R or Python we would recommend searching a few online courses to help get you started. We have also provided an R solution template if you want some assistance in getting through this Task. Whilst its possible to complete the task in Excel you may find the size of the data and the nature of the tasks is such that it is more difficult to complete in Excel.

To get started, download the resource csv data files below and begin performing high-level data checks such as:

- Creating and interpreting high-level summaries of the data
- Finding outliers and removing these (if applicable)
- Checking data formats and correcting (if applicable)

You will also want to derive extra features such as pack size and brand name from the data and define metrics of interest to enable you to draw insights on who spends on chips and what drives spends for each customer segment. Remember, our end goal is to form a strategy based on the findings to provide a clear recommendation to the Category Manager so make sure your insights can have a commercial application.

LIFESTAGE: Customer attribute that identifies whether a customer has a family or not and what point in life they are at e.g. are their children in pre-school/primary/secondary school.

PREMIUM_CUSTOMER: Customer segmentation used to differentiate shoppers by the price point of products they buy and the types of products they buy. It is used to identify whether customers may spend more for quality or brand or whether they will purchase the cheapest options.

Pro analytics Tip: While the data set would not normally be considered large some operations may still take some time to run.

ToDo

- EDA
- Data Cleansing & Preparation
- Data Integration
- Segmentation
- Segmentation Analytics
- Statistical Analysis
- Expose customer behavior
- Stores level revenue / month
- Final Report with Findings and recommendation , How your work will imact our business

Experimentation new Chips layout

Control store & Trial stores

- Store 77, 86, 88 are trial stores

- Trial period was from 2019/02/01 To 2019/04/30

- The main dataset we can choose any control store that have similar behavior as the trial store .(monthly number of customers, monthly revenue etc ….)

- Choose a contrtol store for each trial store

- Then compare each and see if the trial stores has significantly different results over specific period

Recommedations To Study

- Clustering
- Descriptive | Statistical Analytics
- Visualization