**Project 3 – Retail Marketing**

A large retail chain store would like to conduct a “customer personality analysis” on their customers. The goal behind this analysis is to get a better understanding of the types of customers who shop at their store. This process will help them to insure that their products and their marketing messages are tailored to the specific needs, behaviors, and concerns of different types of customers.

Specifically, the retail chain is getting ready to conduct a new marketing campaign. They would like to

* identify groups (or segments) of customers that can be used to customize their marketing messaging for the next campaign
* identify customers who are likely to respond positively to the upcoming marketing campaign (based on their response to the previous campaign)

In addition to having demographic characteristics for their customers, the store has information on the amount each customer has spent on various products, the number of purchases made online vs in person, and the customer’s responsiveness to previous marketing campaigns.

The data dictionary is included below:

**Demographic information**

* ID: Customer's unique identifier
* Year\_Birth: Customer's birth year
* Education: Customer's education level
* Marital\_Status: Customer's marital status
* Income: Customer's yearly household income
* Kidhome: Number of children in customer's household
* Teenhome: Number of teenagers in customer's household
* Dt\_Customer: Date of customer's enrollment with the company
* Recency: Number of days since customer's last purchase
* Complain: 1 if customer complained in the last 2 years, 0 otherwise

**Amount Spent on Product**

* MntWines: Amount spent on wine in last 2 years
* MntFruits: Amount spent on fruits in last 2 years
* MntMeatProducts: Amount spent on meat in last 2 years
* MntFishProducts: Amount spent on fish in last 2 years
* MntSweetProducts: Amount spent on sweets in last 2 years
* MntGoldProds: Amount spent on gold in last 2 years

**Responsiveness to Previous Promotions**

* NumDealsPurchases: Number of purchases made with a discount
* AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
* AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise
* AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise
* AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise
* AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise
* Response: 1 if customer accepted the offer in the last campaign, 0 otherwise

**Purchase Channel**

* NumWebPurchases: Number of purchases made through the company’s web site
* NumCatalogPurchases: Number of purchases made using a catalogue
* NumStorePurchases: Number of purchases made directly in stores
* NumWebVisitsMonth: Number of visits to company’s web site in the last month

**Your Deliverable:**

From an analytical perspective, you will need to develop a set of clusters (and possibly a simple predictive model) that will be used as the basis for your presentation. The presentation should be directed at the marketing manager for the company. It should include:

* An overview of the business problem
* A description of each cluster including
  + A discussion of the unique characteristics of customers in that segment
  + Suggestions of how to develop a marketing message targeted at each group
  + An indication of whether the segment includes a high percentage of customers who responded positively to the previous campaign (based on the variable *response*).
* A discussion of which customers are most likely to respond positively to the next marketing campaign

You will also need to knit your R-code and results together into an HTML file which will be submitted along with your presentation materials.