**Assignment 2: Quick Kits Case and Data Dictionary**

QuickKits is a meal preparation and delivery service with nation-wide franchises. From the QuickKits website, customers choose a recipe and place the order for all the ingredients, sourced and chopped. QuickKits has four main categories of recipes, namely Healthy, Vegetarian, Meaty, and Specialty. Weekly meal plans are offered, with customers choosing from plans with two, three, four, or five meal deliveries per week. A delivery can consist of one or more meals. QuickKits recently decided to test an upselling promotional discount. On March 5, 2018, (2018/03/05), 10% percent of their customers on two-meal-per-week plans were offered a promotional discount of 40% for a three-meal-per-week plan, for two weeks. As the contribution margin for this service is approximately 20%, the promotion will run at a loss. At the end of the two-week period, subscribers to the promotion would remain on the three-week plan at regular price unless they opted out.

The first challenge is to build a predictive model of subscribers to the promotion, based on the 10% of customers that received the offer. Second, utilizing your model and other information in the case to build a business recommendation plan to the managers of QuickKits on how to improve their business.

Grading:

**Part1:** 70% on documentation of the model development, using the rapid model development framework.

1. Based on the background information, build and describe your mental model for subscription to the upselling promotion among the customers on two-meal-per-week plans. (10 points)
2. Connect your mental model and the variables provided in the dataset; form a list (3~5 hypotheses) of hypotheses that can be tested using the data later; and briefly explain the rationale for each hypothesis. (10 points)
3. Document your data cleaning steps (i.e., missing information, low frequencies in the factor variables, correlation, trivially related variables, case identifiers). (15 points)
4. Document your model development steps (i.e., identify the important predictor variables, variable selection, variable transformation, and model assessment and final selection). Provide the necessary visual graphs or charts to support your decisions along the steps (15 points)
5. Interpret your final model and final model’s lift-chart on validation sample (include your final model’s lift chart in your document). (10 points)
6. Attach your R script file for verification purpose (10 points).

**Part 2:** 30% on business understanding and commutations on the recommendations on the company’s promotion plan. Here is the list of potential questions (but not limited to these questions) that might (or might not) be addressed:

* What are the characteristics of those customers who are most likely subscribe to the promotional offer?
* Are there any implications for communication content?
* What are the characteristics of those who subscribe and do not opt-out after two weeks?
* Can you identify customers who were not offered the promotion that you would now offer it to?
* Is this upselling promotion profitable? Why or why not?
* What should the next steps be for the company or your analysis? What information will be useful for your future steps?

For the business understanding, you can incorporate some point format but the rationale for your recommendation should be fully supported and well explained. The evaluation of this part will be based on the clarity of your communications, and evidence and logic of your recommendations.

(Data Dictionary on next page)

**DATA DICTIONARY:**  Variables in the Data Set **QK.csv**

Target Variable to predict probability of response to the promotion and upgrading to three-meal plan, **for the model performance competition**:

**SUBSCRIBE**: Customer signed up for the promotion or not: “Y” if signed up, “N” if not.

Customer Characteristics:

**Custid:** 7 digit customer identification number

**Disc:** Class of customer for standard discounts: “Student” or “Senior”

**Title:** “Mr” “Ms” “Mrs” or “Dr”

**LastOrder:** Date of last order: year/month/day

**Pcode:** Postal Code\*

**DA\_Income:** Mean Income of households in customer’s Census Dissemination Area, thousands of dollars

**DA\_Under20**: Number of households in customer’s Census Dissemination Area with individuals under 20 years of age

**DA\_Over60:** Number of households in customer’s Census Dissemination Area with individuals over 60 years of age

**DA\_Single:** Number of households in customer’s Census Dissemination Area with only 1 person

**NumDeliv:** Number of deliveries ordered in last 6 months

**NumMeals:** Number of meals ordered in last 6 months (each delivery can have several meals)

**MealsPerDeliv:** Average number of meals per delivery (NumMeals/NumDeliv)

**Healthy:** Number of healthy meals ordered last 6 months

**Veggie:** Number of vegetarian meals ordered last 6 months

**Meaty:** Number of meaty meals ordered last 6 months

**Special:** Number of specialty meals ordered last 6 months

**TotPurch:** Amount purchased last 6 months in dollars

**Weeks3Meals:** Number of weeks that a customer that signed up for the promotional offer stayed with the three-meal-plan, after the promotion was over.

\*Canada has 850,000 postal codes, which are combined into 54000 Dissemination Areas. A Dissemination Area typically has between 400 and 700 households.