GAME EMAIL MARKETING CASE

March 2021

Developed by Yashar BASHIRZADEH Ph.D.

(yashar.bashirzadeh@rennes-sb.com)

Customer Data Analysis course - Rennes School of Business

A- Problem Statement:

Managers of a game company send out an email to their users inviting them to try a new feature of the game. The managers are interested in knowing the effect of this email on how much money users spend after. Namely they want to know if opening the email has any effect on money expenditure.

B- Data Description:

You have a random subsample of the whole data (n=2000) with the following variables:

ID (ID): A unique random number generated for each user.

Logged Time Before Per Day (Logged.ActivityBPD): The amount time each user spent per day in the period before the email (log transformed).

Logged Money Before Per Day (Logged.MoneyBPD): The amount of money each user spent per day in the period before the email (log transformed).

Active Days Before (ActiveDaysB): The number of days each user spent in the app in the period before the email.

Tenure (Tenure): The number of days between the very first day a user logs in to the app in the period before the email and the email date.

Recency (Recency): The number of days between the last day user logs in to the app in the period before the email and the email date.

Email (Email): Binary variable that is equal to 1 if the user received the email.

Open (Open): Binary variable that is equal to 1 if the user opened the email.

Click (Click): Binary variable that is equal to 1 if the user clicked on the link inside the email.

Unsub (Unsub): Binary variable that is equal to 1 if the user unsubscribed from the email list.

Lag (Lag): The number of days between the email date and first day that the user logs in to the app in the period after the email.

Logged Money Expenditure After (Logged.MoneyA): The total amount of money each user spent in the period after the email (log transformed).

Logged Time Expenditure After (Logged.ActivityA): The total amount of time each user spent in the period after the email (log transformed).

C- Questions:

1. what is the effect of opening the email on Logged Money Expenditure After? (Hint: Run a regression WITHOUT control variables and report the results.)

2. What control variables would you use and why? What is the effect of opening the email on Logged Money Expenditure After when you control for the relevant control variables? (Hint: Run a regression WITH control variables and report the results.)

3. What happens to the adjusted R2 after adding the control variables and why?

4. Why is the treatment effect (i.e., the effect of opening the email on Logged Money Expenditure After) different with and without the control variables? (Hint: Think about what type of users are more likely to open the email and how controlling for relevant variables might alleviate this problem.)

5. Basically, what the managers are doing is looking only at email receivers and comparing those who opened the email and those who did not. How do you evaluate this strategy? What better strategy would you recommend to the managers to investigate email effectiveness?