Assignment 1

Download Datasets 1 from the following link.

Database kept confidential. It cannot be shared.

NOTE: The data is confidential and can be used just for educational purpose in this course. You should not upload it for sharing.

NOTE2: Use all available resources to solve the problems. You can find a solution to most of the coding problems from the Internet. Google it, if you're stuck in the middle.

There are six data files in the folder. Three files are for mobile channel adopters, and the other three are for non-adopters.

- Mobile channel adopter data: MobileMember, MobileOrder, MobileBasketDiscount
- Mobile channel non-adopter data: OnlineMember, OnlineOrder, OnlineBasketDiscount

Mobile channel adopters purchase products either through the PC or Mobile channel, while non-adopters purchase only through the PC channel.

First, we want to compare demographics of adopters and non-adopters. Note that there are business users in the datasets and their information on gender and age is not available.

1) What is the average age of the adopter group ("MobileMember.dta")? What is the average age of the non-adopter group? ("OnlineMember.dta"). Is there any age difference between adopters and non-adopters?

    Hint: To generate the age, you need to change the Birth year to numeric type and subtract it from 2017. Refer to the following code to change it to the numeric type

    OnlineMember$Birth<-as.numeric(as.character(OnlineMember$Birth))

    MobileMember$Birth<-as.numeric(as.character(MobileMember$Birth))

2) Investigate the proportions of female in the adopter group ("MobileMember.dta") and non-adopter group ("OnlineMember.dta"). Is there any gender difference between adopters and non-adopters?

Load "MobileOrder.dta" and generate a dummy variable "Mobile" where 1= if Mall == "03" & (AccessRoute=="1000132495" | AccessRoute =="1000132496" | AccessRoute =="1000013091") and 0=otherwise, which indicates whether the transaction happened through the mobile channel or not.

3) Compare PC and Mobile transactions regarding OrderPrice. Which one has larger OrderPrice? Is it consistent with your conjecture? Do you have an explanation for this?

4) Compare the confirmation rate between PC and Mobile transactions. Make a dummy variable, "CRate" where 1=confirmed and 0=otherwise, which indicates whether the transaction was confirmed or not (Hint: You can compare OrderQuantity and ConfirmedQuan and generate 1 if they have the same value and 0 otherwise.) How much is the confirmation rate (=sum(CRate)/Total Number of

Transaction) for PC transactions? How much is for Mobile transactions? Is it consistent with your conjecture? Do you have an explanation for this?

5) Compare the dependence on certificates (OkSeller, QuickSeller, BigSeller) between PC and Mobile transactions. Which transaction depends more on certificates for ordering products?

Guideline for Assignment 1:

Submit the answer sheet and R-code or RapidMiner Screenshots used for the analysis to the course Blackboard. Please include your student number and name in the header of the document. Your answer sheet should be in the same format as the final report (See the syllabus). Answer sheet should not exceed one A4 page. Append all R-code or RapidMiner screenshots used for the analysis. There is no page limit for the appendix. The due is Oct 11, 23:59 PM.