STAT 331 Lab 9

Instructions

Submit your HTML markdown document by the beginning of class. In this lab you're going to re-do Lab 6 using loops and IF-statements for certain parts instead of what you did originally. **In particular, you should use loops and IF-statements for parts 2-6 below.** You may copy your solutions to the other parts from your previous submisson for Lab 6 (or my solution).

Exercises

A regional SAS conference will be held in a few months time. The raw data file Conference.dat contains information about the registered attendees: first name, last name, attendee ID, business phone, home phone, mobile phone, OK to contact attendee at business (Yes/No), OK to contact at home (Yes/No), OK to contact at mobile (Yes/No), registration rate, will attend Wednesday mixer, will attend Thursday lunch, whether willing to volunteer at the conference, and eating restrictions (if any).

1. Read the data set into R

2. Using the following registration rates, create a registration group variable that groups attendees in the following way. Produce a frequency table of the registration group.

- Academic Regular $350

- Student Regular $200

- Regular $450

- Academic Early $295

- Student Early $150

- Early $395

- On-Site $550

3. Create another (separate) variable that condenses the the above groups into Regular, Early, and On-Site. Produce a frequency table of the new registration group variable.

4. Create a variable that represents the participants area code using the business phone number first. If the business phone number is missing the area code, then use the mobile phone number. If the mobile phone number is missing the area code, then use the home phone number. Produce a frequency table of the area codes.

5. Revise all of the phone number variables to have the following format "XXX-XXX-XXXX". If a phone number does not have an area code, copy the area code identified in (4). Let's say it's safe to assume someone's work, home, and mobile phone have the same area code (are in the same place). Print the first name, last name, id, and three phone numbers corresponding to IDs 1058, 1060, and 1068.

6. The catering committee has already ordered food that does not contain nuts or shellfish, and all meals are kosher, but they need to figure out how many attendees need the special vegetarian or vegan meal. Create a variable that flags the attendees that require a vegan or vegetarian meal as a "Veg", that specified some other restriction as "Other", and "None" otherwise. Produce a frequency table of your variable. How many attendees require either a vegan or vegetarian meal?

7. Does there seem to be a relationship between "willing to volunteer at the conference" and registration group? Use graphs or descriptive statistics to justify your answer.

8. The catering committee has been filtering through the eating restriction information tediously, and have thought of a possible way to improve the registration website next year. The eating restrictions window should only appear for attendees who have indicated they are attending the Wednesday mixer or the Thursday lunch.

- How many people did not indicate they are attending the Wednesday mixer or the Thursday lunch, but still provided eating restriction information? Should the website be amended for next year or does it not appear to be an issue?

9. Overall, how much did SAS make from the registration fees?

- Are certain locations (area codes) contributing more to the conference than others (via registration fees)? Or is the distribution fairly equal across area codes?

- Does the distribution of registration fees by area code match the distribution of attendance by area code?