HW2: Class One Survey Analysis

Kim Johnson & Kyle Pitzer

PART 1: INSTRUCTIONS

Please complete tasks 1-8. Use R syntax to solve all problems (i.e. do not manually change values/variables in datasets or check them manually to make sure things are correct except as a double or triple check).

1. Import class one survey data from our Github site (The dataset called 'Class 1 Survey Spring 2020.csv' is located in the class one folder), calling the R dataframe that you create *C1survey*.

2. Determine: a. the number of observations (i.e. the number of people who filled out the survey) and b. the number of variables in the dataframe.

3. Generate and display a list of column names, calling your list *varlist*.

4. a. Rename the column variables to something shorter and that is descriptive of what the variable is about (for example *like_dogs* for the 'Do you like dogs?' question variable) and b. write code to display that they are renamed.

5. Write code to determine and display the number of factor, integer, numerical, and character variables there are in the *C1survey* dataset.

6. a. Using code, check the *bday* and *bmonth* variables for any unusual or missing values. If you find any, b. describe what you will do with the unusual values in a comment before or immediately following the code that does it, and c. after you have cleaned up any unusual values, find the median bday and month.

7. a. Create a new variable called *bseason* that gives the season according to Northern Meteorological season in which respondents were born (winter=12/1 to 2/29, spring=3/1 to 5/31, summer=6/1 to 8/31, fall=9/1 to 11/30). b. Using the table function, print a table of Seasons in the columns and bmonths in the rows that allows you to check that the coding is correct c. Sum the columns using the addmargins function to answer the question of how many classmates were born in each season? Include your answer in the code you turn in.

8. Pick your favorite variable to analyze, come up with a question you want to answer with that variable, generate the code, and provide an answer to your question. Describe what you did using comments (i.e. #'s).

PART 2: INSTRUCTIONS

After completing the tasks above, go back and edit your code chunks to produce an attractive brief research report. Submit the HW2 rmd file along with the research report in html, pdf, or word format using the knit pull down menu.

1. Annotate all code for tasks 1-8. Please number tasks 1-8 and copy the instructions along with the numbering. The number plus the instructions should proceed each code chunk.

2. Edit chunks for task 1-4 to hide both the code and cleaning of the data. Instead of the code and output, provide a narrative about the data, what you did to clean and prepare it, and any relevant answers to these tasks (how many columns/rows, etc.).

3. For numbers 5 and 6, write a narrative about the tasks and output. Edit the chunks to hide the code, but show the output. Hint: I suggest breaking up these tasks into several chunks to get an attractive format.

4. For number 7, create two chunks - one for creating the new variable (a) and one for the table and addmargins function (b, c). Hide your code and output for creating and checking the new variable, but show your code and output for your table and addmargins function and provide a narrative.

5. For number 8, modify your code and output accordingly to hide code and display relevant output. Write your narrative according to the instructions of the task.

6. Add a picture of your choice to the report wherever you'd like.