Messy data is common when making statistical analyses. Your note card responses are a representative example. In groups of 3 or 4, enter the following entries into a data frame in an R script. Prior to coding, look at the responses and think about how to organize the data.

1. Reminder of the questions (do not need to enter this)
   1. What is your group name?
   2. Have you ever owned a pet?
   3. How many times did you go to the beach this summer?
   4. Who is your favorite current artist?
   5. How would you rate your R skill (A: none, B: some R, C: some other language, D: advanced other language)?
   6. It’s too early
   7. A pet fish
   8. Twice
   9. Kendrick Lamar
   10. D
   11. B Team
   12. I used to have a black cat named Dusty
   13. No clue, probably 15-20 times
   14. Ed Sheeran
   15. C
   16. It’s To Early
   17. Yes
   18. 2
   19. Vulfpeck
   20. C
   21. Q
   22. Yes I have owned a pet
   23. I went to the beach once this summer
   24. I don’t have one favorite musician
   25. I am a beginner in another language and R

How did your group organize the data? Did you follow any rules in regards to data entry for each category?  
  
  
  
Did you find it easier or harder to perform data entry in R as compared to other programs (e.g. MS Excel)? How would you improve this process?

Follow the in class instructions to download the data and sample code from Moodle. Use the sample code to import the data and answer the following questions:

1. How many people in the class own a pet? Find two ways to output this answer.
2. Find the mean, minimum, and maximum values of the number of times each person went to the beach. Find a way to display all at once and each separately.
3. Make a vector of the favorite musicians of the members of Group 2.
4. How many A-level programmers have ever owned a pet? How many D-level programmers never owned a pet? Can you find a way to display the number of people who ever owned a pet by programming level?
5. If you have time, read through data-cleaning-code.R, which shows how I took the raw data and filled in missing data. Can you understand my process? Use help() for any functions you don’t recognize. Are my methods reasonable? What other ways might you deal with missing data?