Assignment: Intro to R

Purpose: The purpose of this assignment is to reinforce the introductory skills covered in lab by applying them to specific tasks using R.

Objectives: You will review the following skills in R:

- Using RStudio
- Defining a working directory
- Install/initialize a package
- Using built-in datasets
- Reading a dataset (read.csv)
- Data types: numeric, factors, Boolean
- Vector definition and manipulation
- Using data.frames

What you need to do:

1. Create a new project in RStudio. Name it IntroR-HW-*yourname.*
2. Within your project, save the default R source file Chocolate.R.
3. Set your working directory to wherever you have saved your project. This means that you will either store your dataset in that directory, or you will read in your dataset using the full path for it. (For these homeworks, I would opt to store it in the same directory, just to make things easy. But you certainly may prefer to store all of your data files in one directory.)
4. Create a data.frame named dfChoc to capture the following scenario: There are 5 people in a survey about chocolate. They are Bob, Joe, Sam, Ted and Tom. Their ages are 23, 22, 27, 22 and 23, respectively. Their favorite chocolate snacks are Snickers, KitKat, Reeses, KitKat and Snickers. Here's how many candy bars they eat every week: 4, 3, 5, 4, 2. Create three vectors, "name", "age", "favorite" and numberBars. Construct the data.frame from these vectors. Use the "dim" command to check that your data.frame has the correct number of rows and columns.
5. What are the classes (data types) of each vector? You can use four separate commands, or you can use sapply(dfChocolate, class)
6. What are the classes of each variable of the data.frame? (Wow!)
7. Change the class of "favorite" to a factor.
8. Display the levels of "dfChocolate$favorite". (You can simply display the variable, and the levels will display, as it is a factor.)

9. Change the order of the levels of "favorite" to Snickers, KitKat, Reeses. There are several ways you can do this. Use the levels() function, which you can easily google.
10. Display the levels of the variable dfChocolate$favorite again, and it should show up differently.
11. This week is a holiday. Everyone gets an extra candy bar. Add 1 to all elements of the *original vector* numberBars.
12. Display the first two columns of the data.frame.
13. Display the second and third row of the data.frame.
14. Display the "name" column. Reference it by "name".
15. Display the third and fourth columns, rows 3 through 5.
16. Display the column names of the data.frame.
17. If you have not yet installed the package HistData, do so, using the R command. If you already installed the package, put the command in a comment line, so that I can see that you know what command to use.
18. Load the HistData package. It has a lot of historical datasets.
19. Load the DrinksWages dataset.
20. Display the first several rows of the dataset.
21. Download anesthetic.csv and the documentation file, anesthetic.docx. Store them either in your project folder, or wherever you have decided to store your data.
22. Read this file into your project in R. Assign it to anes
23. Look at the first several rows of the dataset.
24. Are there any variables that would be appropriate as factors? If so, which one(s)?

What to submit:

1. Copy your code in the source editor. Paste to a Word doc or similar.
2. Copy the contents of your console. Paste to the same Word doc or similar.
3. Submit on Sakai.
4. You may work in teams of two, one submission per team. Each team member must submit something on Sakai. If you are submitting on behalf of your team, be sure to include a comment "I am submitting on behalf of Bob" in the assignment box on Sakai. If someone else is submitting for your team, be sure to include a comment "Bob is submitting on behalf of our team".