Assignment 1

Complete the following assignment and upload your R script into Blackboard. We will be grading your script, not the output.

1. Enter the following dataframe into R and assign to the variable dat. Be sure that Name is recognized as a **character** and Age/BMI are recognized as **numbers**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Age | BMI | Factor A | Factor B |
| Frank | 34 | 24.2 | -1 | 1 |
| Bob | 28 | 18.3 | -1 | -1 |
| Sally | 19 | 15.4 | -1 | 1 |
| Susan | 28 | 22.7 | -1 | -1 |
| Joan | 30 | 29.2 | -1 | 1 |
| Bill | 47 | 32.4 | 1 | -1 |
| Richard | 24 | 21.0 | 1 | 1 |
| Jane | 34 | 40.4 | 1 | -1 |
| Jill | 32 | 24.8 | 1 | 1 |
| John | 64 | 34.4 | 1 | -1 |

* 1. Without manually entering the data, create a new column Factor AB that is the product of Factor A and Factor B
  2. Using the rep() command, create a new column Factor C that looks as follows

(-1,-1,1,1,-1,-1,1,1,-1,-1)

* 1. Without manually entering the data, create a new column Factor ABC that is the product of Factor A, Factor B, and Factor C
  2. Make sure that all Factor columns are recognized as **factors** in R

1. Enter this additional column into the dataframe dat, making sure it is recognized as a **factor** in R.

|  |
| --- |
| Smoking |
| Yes |
| No |
| No |
| Yes |
| Yes |
| No |
| Yes |
| Yes |
| No |
| Yes |

1. Replace the BMI of Richard with a NA
2. Without manually entering the data and using the dataframe after entering Richard’s BMI as NA , create a new column in dat with the logarithm of BMI
3. Create a new dataframe dat2 selecting only the columns from dat corresponding to logarithm of BMI, Factor A, Factor B, and Factor AB
4. Create a new dataframe dat3 selecting only the first 5 rows of dat2