

Introduction to R for Data Analysis in the Health Sciences
BIOST 509
In-Class Exercise 1
Due by 5:00 Friday on 9/27/2019

Instructions

Begin by downloading and install R and RStudio following the instructions on the slides. Then, install and load the "tidyverse" package.

Submit your answers to the below questions in a R Script (.R), Word (.doc or .docx) or pdf file to Canvas. If you are waiting for a spot to open up in the course and do not have access to Canvas, save your responses and you can submit them when you enrol.

Include the R commands that you used to obtain your answers as well as the answers themselves.

Questions

0. Ensure you can read in the "fev.csv" dataset used in the lecture, and that you obtain the same data summaries seen in the slides. Datasets are available from Canvas in Pages/Week 1 materials (or from the github site under "datasets"). You do not need to report anything here.

1. The dataset "psa.txt" is available from Canvas in Pages/Week 1 materials (or from the github site under "datasets"). This file is a space delimited file, and not a comma separated file (not .csv). Read in the dataset using **read_table** and store it as an object, using a sensible name.

psa.txt" contains data on 50 men having hormonally treated prostate cancer. The first line of the file contains the following variable names separated by white space (spaces and/or tabs). Each successive line contains data pertinent to one of the 50 study subjects. Missing data are denoted with "NA". Data are available for the following measurements, with individual measurements on a line separated by white space.

ptid = patient identifier
nadirpsa = lowest PSA value attained post therapy (ng/ml)
pretxpsa = PSA value prior to therapy (ng/ml)
ps = performance status (0= worst, 100= best)
bss = bone scan score (1= least disease, 3= most)
grade = tumor grade (1= least aggressive, 3= most)
age = patient's age (years)
obstime=time observed in remission (months)

- a) What is the mean age of men in the study?
- b) How many men in each tumor grade?
- c) What is the mean age of men with tumor grade 3?

2. Please provide the following information on your background and interests:

- a) Which biostatistics or statistics courses have you taken? (briefly describe the level and/or content of the courses)
- b) Which biostatistics or statistics courses are you taking concurrently with this course?
- c) Have you had any prior experience with any of the following statistical packages: R, SPSS, SAS, Stata, or Other? If yes, please list them.
- d) Are there any special R topics that you would like to see covered in this course? Please list them.

Welcome to the course!
We look forward to supporting your learning and
getting to know you this quarter!