

APPM2720 Homework 2

This homework should be sent to nychka@ucar.edu before next Monday. You are welcome to work together but please send in individual assignments.

If you have not been able to install RStudio on your personal PC or Mac at this point please send me email with the issues and we can work out how to fix them.

Download all the material from [Week2](#) (Some of these files may have been corrected from lecture so download everything again.)

To change the working directory in RStudio

PC: *Tools / Change Working Dir... menu*

Mac: *Session / Set Working Directory*

(1) One strength of R is the wealth of add on packages that are contributed from a worldwide community of data scientists. For this class we will focus on the package **fields**. From R studio install **fields**. Go to the *tools* menu and select *install packages*. The default choices are OK to use and fill in **fields** in the package box. To test if fields is working try the code below (the correct result is obvious!)

```
library(fields)
data(lennon)
# this loads the lennon data set that is included in the fields package
image.plot(lennon, col=grey.colors(256), axes=FALSE)
```

(2) Load the A4 data set into your R session. This involves loading the file **AudiA4.rda** from the Week02 class folder. To make things easier to type create a data set of just the **mileages**

```
load("AudiA4.rda")
mileage <- AudiA4[,2])
```

If you divide the mileage by 1000 how do the mean and median change? How does the interquartile range and the standard deviation change? If you subtract 1400 from the mileage how does the median change? How does the interquartile range change?

(3) People often wonder if the weather for one month is related to the next months. For the Boulder temperature data (**BoulderTemperature.rda**) make a plot of April temperatures against May temperatures. If

you have a cold April how often would you expect a cold May? Make sure your plot is labeled and has a title.

(4) Use the *abline* function to add a line through your scatterplot that you feel best summarizes the relationship. (For now just pick the slope and intercept using trial and error.) Does this line seem to be useful?

(5) Average the june, july, and august temperatures together for each year. Do you see any obvious trend in these summer temperatures over time?

(6) Write a function that takes a data set and returns the 6 values: min, first quartile, median , mean , third quartile and max. Also include the option to exclude missing values (NA). Hint see the `na.rm` use in *RFunctions.R* from the scripts folder.

Handing things in:

Work on your homework as an R Script

-OPTION 1 Save your plots to a file using the export button in the plot window and then just attach them to the email with your script.

-OPTION 2 you can use the *compile notebook* button (in the editing window on the top line, it looks like a spiral notebook.) to generate an html file that has your code and figures all together.

-OPTION 3 Work on your homework as an R Markdown file. This works about the same as OPTION 2 but allows to add more comments and explanation in your report.