

# The R Statistical Computing Environment

## Basics and Beyond

### Getting Started With R: Exercises

**John Fox**  
(McMaster University)  
**ICPSR/Berkeley**

2016

1. If you haven't already done so, install R and RStudio on your computer (instructions are on the workshop website) and verify that it works. Install the **car** package: `install.packages("car")`.  
*Note:* This will suffice for the first day, but please visit the workshop website to insure that all R packages for the workshop are installed.

2. Write a function **MAD** to compute the median absolute deviation from the median,

$$\text{MAD} = \text{median} [|x_i - \text{median}(x_i)|]$$

Confirm that your function works by comparing its results with those from the standard R **mad** function, using, e.g., `MAD(1:100)`; then look at how **mad** is programmed and compare it to your solution. In making the comparison to your function, be sure to set the argument `constant=1` in the call to **mad** (see `?mad` for the explanation).

3. One of the data sets in the **car** package, called **States**, contains education and other data for the 50 U.S. states and Washington DC. Find out what's in the data set by looking at its help page (`?States`), and then perform a linear least-squares regression of the average SAT math score of graduating high-school students on the average teachers' salary in the states. Perform a second regression of SAT math score on both teachers' salary and percentage of students taking the SAT exam. Compare the coefficients for teachers' salary in the two regressions. How do you account for the difference? Make some graphs of the data.
3. Incorporate your analysis of the **States** data in an R Markdown document (or simply start working R Markdown).