STAT 406: HW9

- All computer code should be written using the language R. Type ALL your code into one PLAIN Text format file. Plain text format is available by default in R. Please do not use Microsoft Word .doc format or .rtf format of .pdf format. Inside your plain text file, make sure you identify each problem in a comment placed at the beginning of the problem. The file name should match your name as in 'JohnDoe.R'. Submit your R code file online (under Assignments) at or before the due date posted on ctools. And hand in a hard copy of the code with any additional writing material to your GSI, in class, **at the end of your lab. session**.

- I recommend that before submitting your homework, you also create a new directory and run your R code, to make sure that it is self-contained and runs as you planned.

1. We have seen how to process the NSF award files in class. Download the 2013 NSF awards files in XML formats from
   `https://www.nsf.gov/awardsearch/download.jsp`.
   Adapt and expand on the code written in class to find the top ten universities (or entities) which have received the most NSF funding in 2013.

2. As part of its inspection programs, the US Occupational Safety & Health Administration routinely takes samples and measures exposures to chemical hazards. The data is available online in XML format at the address
   `https://www.osha.gov/opengov/healthsamples.html`.
   Download the 2012 data and familiarize yourself with its structure. Process the data, and count for each of the 50 states, how many samples were collected in that state that tested positive for asbestos. Asbestos is recorded in the dataset as 'Asbestos (all forms)', and a sample is tested positive if the value recorded in the tag sample_result is positive.