Week 4 Quiz

The answers to the following questions should be placed in a single R script. Place your R script in a public repository on github and submitting a link to the script here. Label your answers using comments so that they can be clearly and quickly found within the script.

Week 4 quiz is due end of day on Friday September19th. *Solutions to all quiz exercises will be posted on Saturday September 20th. Your grade will be based on two randomly selected exercises.*

The first two questions are graded credit/no credit. Any attempt at a solution will receive full credit.

1. Read about the OSEMN data science process workflow in "A Taxonomy of Data Science", http://www.dataists.com/2010/09/a-taxonomy-of-data-science/. Then, *without consulting your reading or notes*, write a one sentence description for each of the five process steps.
2. Which do you think will yield a better program: Well designed function interfaces and data structures, decently implemented, or decently designed function interfaces and data structures, Well implemented? Briefly explain your reasoning.
3. Display the day of the year for the current date. For example, on February 1st, the value returned should be 32. If you want, use a third party package such as chron or lubridate.
4. Give an example of when you would choose the POSIXlt date class, versus when you would choose the POSIXct date class.
5. The hero of James Gleick's book, *The Information*, is pioneer data scientist Claude Shannon. Write R code that answers the question, "How long ago was Claude Shannon born?" (There is clearly some room for interpretation in your answer here).
6. Suppose you have a data frame where one column gives the month (in numeric format), the next gives the day, and the third column gives the year. Use R to create such a data frame (by hand is fine) and then add a fourth column with the date in date format. Use str() on the data frame to check your work.
7. Illustrate the code necessary to take a string of MM-DD-YYYY format and convert it to a date.
8. Illustrate the code necessary to take a date and extract the month of the date.
9. Create a sequence of all of the dates from January 1, 2005, to December 31, 2014.
10. When you're creating a data frame from a set of vectors, what argument can you give the data.frame() function to instruct it to *not* automatically convert character vectors into factors?