INTR 4051/SOSS 7091
Lab 3: Non-numeric vectors and recoding

Create an RStudio project called `lab3`. Do this lab assignment in an R Markdown (`.Rmd`) file called `yoursurnameL03.Rmd`. When you are finished, upload the assignment to your Sakai dropbox.

More vectors

Download the file `lab03.RData` from Sakai into your project directory. Then load it in your first code chunk using

`load(file="lab03.RData")`.

(The file contains the vectors you will need for the assignment.)

The vector `numbers` contains random integers (whole numbers) between 0 and 100.

1. Apply a logical index vector to `numbers` to create a vector called `q1` that contains only the values that are between 80 and 95 (inclusive).
2. Apply a logical index index vector to `numbers` to create a vector called `q2` that contains the values that are between 10 and 30 (inclusive) or between 70 and 80 (inclusive).
3. Apply a logical index vector to `numbers` to create a vector called `q3` that contains the values that are between 10 and 30 (inclusive) and between 25 and 80 (inclusive).

The vector `country` contains the official "English short country names" of the International Standardization Organization (ISO), and the vector `code` contains the corresponding three-letter ISO country codes.

4. Create a character vector called `first.letter` containing the first letter of each country name, in lowercase
5. Apply a logical index vector to `country` to create a character vector called `m.country`, containing the names of all countries whose names start with "M."
6. Apply a logical index vector to `country` to identify the country with the longest name. (A bit tricky.)
7. (Bonus) Use `paste()` to create a character vector called `country.code` with each element consisting of the country name followed by a space and the three-letter code code in parentheses, such as:

   `"Afghanistan (AFG)"`

   (Hint: Use `sep=""`, and paste four things together: the country name, something else, the country code, and (still) something else.)
8. Use R to create a numeric vector containing all of the odd numbers from 1 to 99, and do it at least three different ways.

Coding and recoding factors

The following question (q19) is from the Pew Global Attitudes Project (GAP) survey conducted in 2009:

> How concerned, if at all, are you about the rise of Islamic extremism in our country these days? Are you very concerned, somewhat concerned, not too concerned, or not at all concerned about the rise of Islamic extremism in our country these days?

The numerical codes and response options were:

1 = Very concerned;
2 = Somewhat concerned;
3 = Not too concerned;
4 = Not at all concerned;
8 = Don't know (NOT READ);
9 = Refused (NOT READ).

The vector `egy.q19` contains numerical codes for the responses of a sample of 1000 Egyptians.

9. Create a factor called `egy.q19F` that labels each response option ("level") using the wording from the questionnaire.

10. Create a recoded factor called `egy.q19F.r` that collapses the responses into three categories ("levels") — one for those who who said they are "very" or "somewhat" concerned, one for those who said they are "not too" or "not at all" concerned, and one for those who did not express a clear opinion:

11. Use `summary(egy.q19F.b)` to get a summary of the data.
    (a) In one sentence, without using any numbers, describe what the data tell you about Egyptians' views (as of 2009) of Islamic extremism?
    (b) Off the top of your head, think of one characteristic of Egyptian individuals that might influence whether they are concerned about Islamic extremism or not.