Assignment 3: Data wrangling

Due Sunday, October 27, 2019 @ 11:59pm

Public Health 280.346, Fall 2019

Your assignment for this module is to read in an unprocessed version of the National Medical Expenditures Survey (NMES) dataset using `read_csv()` and then recode all the relevant variables into factors with meaningful levels according to the given codebook.

To complete your assignment, do the following items. You can use the `Assignment3_template.Rmd` file to do your work. When completed, please submit your assignment through Github.

(1) Read in the unprocessed NMES data (`nmesUNPROC.csv`) using the `read_csv()` function. Store this data in an object called `nmesRAW`.

(2) Add an `mscd` factor variable to the dataset that is an indicator of whether the individual has a major smoking caused disease (1 = Yes, 0 = No). You can create this variable from the `lc5` and `chd5` variables as we did in class. (See class work 3-3 for hints.)

(3) Recode the categorical variables in the NMES data into factors according to the following codebook. Once all of your recoding is complete, drop the original variables and rename the dataset `nmesPROC`.

- `lc5`: indicator of lung cancer, laryngeal cancer or COPD (1 = Yes, 0 = No)
- `chd5`: indicator of CHD, stroke, and other cancers (oral, esophageal, stomach, kidney and bladder) (1 = Yes, 0 = No)
- `eversmk`: indicator of whether the individual has ever been a smoker (1 = Yes, 0 = No)
- `current`: indicator of whether the individual is currently a smoker (1 = Yes, 0 = No)
- `former`: indicator of whether the individual is formerly a smoker (1 = Yes, 0 = No, NA if `eversmk` = 0)
- `beltuse`: self-reported seat-belt use (1 = Rarely, 2 = Sometimes, 3 = Always/Almost always)
- `educate`: education level (1 = College graduate, 2 = Some college, 3 = HS graduate, 4 = Other)
- `marital`: marital status (1 = Married, 2 = Widowed, 3 = Divorced, 4 = Separated, 5 = Never married)
- `poor`: indicator of whether the individual is below the poverty line (1 = Yes, 0 = No)
- `female`: individual's sex (1 = Female, 0 = Male)

(4) Give the code need to do the following recodings:

- Recode the `eversmk` variable to be `Ever smoker` and `Never smoker` instead of `Yes` and `No`.
- Recode the `marital` variable to be `Married`, `Never married`, and `Previously married`