DEM 5283/7283
Homework 1
10 points
Due February 8

Survey statistics

**NOTE Use Rstudio to create your homework, and include all discussion and analysis within that document. Submit the link to your rpubs.com site to Blackboard when you publish the document

If you need help figuring out what your PSU, weight and sample strata variable are called in YOUR data, please ask.
   a) Define a <u>binary outcome variable</u> of your choosing and define how you recode the original variable.
   b) State a research question about what factors you believe will affect your outcome variable.
   c) Define <u>at least</u> 2 predictor variables, based on your research question. For this assignment, it's best if these are categorical variables.
   d) Perform a descriptive analysis of the outcome variable by each of the variables you defined in part b. (e.g. 2 x 2 table, 2 x k table). Follow a similar approach to presenting your statistics as presented in <u>Sparks 2009</u> (in the Google drive). This can be done easily using the tableone package!
      a. Calculate descriptive statistics (mean or percentages) for each variable <u>using no weights or survey design</u>, as well <u>as with full survey design and weights</u>.
      b. Calculate percentages, or means, for each of your independent variables for each level of your outcome variable and present this in a table, with appropriate survey-corrected test statistics. (tableone package helps)
      c. Are there substantive differences in the descriptive results between the analysis using survey design and that not using survey design?