**Exercise 1**: Try to understand and run the R code contained in class slides

**Exercise 2**: Use 2013-2014 National Adult Tobacco Survey (https://www.cdc.gov/tobacco/data\_statistics/surveys/nats/). Keep variables Gender (gender), Race (RACEETHNIC\_R), General health status (GENHEALTH), Marital Status (MARITAL2), Employment status (employ2), Educational status (EDUCA2\_R), Income level (INCOME2), Smoking status (SMOK100, SMOKNOW). Perform the following:

1. Check missing frequencies for All above variables
2. Conduct proper imputation methods for all variables
3. Estimate national current smoking prevalence
4. Examine the relationship of current smoking status with other covariates
5. After imputing all variables other than income, compare following three propensity score weighting approaches for estimating frequency of income level: 1. Parametric method; 2. Weighting cell method based on covariates; 3. Weighting cell method based on predicted propensity score values. Compare the results with the results obtained with multiple imputation

**Reading assignment**:

1. Austin, P. (2013). The performance of different propensity score methods for estimating marginal hazard ratios. Statistics in Medicine, 32(16), 2837-2849
2. Franklin, J., Eddings, W, Austin, P, Stuart, E. and Schneeweiss, S. (2017). Comparing the performance of propensity score methods in healthcare database studies with rare outcomes. Statistics in Medicine, Accepted