**Final project (5% of total grade)**

**Due date 11:59pm Thursday May 7th in Assignments on Blackboard**

**Project instructions**

First upload the NHANES SAS data into SAS studio. Use the following SAS to upload the data into the WORK folder.

/\* Change directory path in the statement below \*/

LIBNAME temp "/home/konda0/401";

DATA THYROID;

SET TEMP.TSH;

RUN;

PROC CONTENTS DATA=THYROID; RUN;

Select TSH, Thyroid stimulating hormone (uIU/mL), as the health outcome(Y) and rest of the variables as predictors except SEQN. Fit a predictive model for TSH. Generally, higher TSH values result in hypothyroidism. Predictors include several socio demographic baseline variables and concentration of few metals obtained from blood and urine. The variables in the SAS data includes

SEQN person id

TSH Thyroid stimulating hormone (uIU/mL)

FEMALE (1 female, 0 male)

AGE

BMI

PIR Poverty income ratio

ETHNICITY (AFAM African American; MXAM Mexican American, OHAM Other Hispanic/Latin American, WHAM White American)

EDUC (1= < 9th grade, 2=9 to 11grade, 3=HS, 4=SomeCollege, 5=College)

SMOKER (1 smoker, 0 non-smoker)

Arsenic

Barium

Cadmium

Lead

Tungsten

Thallium

Uranium

**Notes**

(i) Generally log transformation is recommended for TSH. Probably same log transformation on metals.

(ii) Very brief literature review on the outcome variable is optional. (iii) Your analysis should include univariate on the outcome, bivariate, multivariate analyses including an explicit equation for *good* predictive model equation. (iv) If you suspect interactions, justify them, and include them (v) Briefly discuss residual diagnostics, Partial SS, Collinearity, Outliers (v) Discussion, Limitations, conclusions must include explaining the effect of significant predictors, surprises in the analysis, what’s wrong with data/analysis/predictive model, and value of your analysis to public health. (vi) Precise report in at most 5 pages is strongly encouraged. (vii) R2 in public health data is often very small.

This is individual and independent project. If I see any similarities between your report and any of your class-mate(s) report, all of you involved in plagiarism will get zeros without any further explanation. No further discussion, negotiations, emails, meetings!!! You and your friends will be reported to SPH for further action.

**At most 5 pages report** (12 font size and single space) needed with relevant tables and graphs for the grade. The report must include brief introduction about the health outcome and the predictors. Analysis, Results, Discussion, Conclusions and References.