**Analyzing NHANES**

**Survey Data**

**Spring 2020**

**Data 606: Capstone in Data Science**

**National Health and   
Nutrition Examination Survey**

NHANES is a program designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. NHANES is a major program of the National Center for Health Statistics. NCHS is part of the Centers for Disease Control and Prevention (CDC) and has the responsibility for producing vital and health statistics for the United States.

The program was started in the 1960s. In 1999, the survey became a continuous program that has a changing focus on a variety of health and nutrition measurements.

* Records statistics for over 5000 people every year.
* Selected to represent the U.S. population of all ages. To produce reliable statistics, NHANES over-samples persons 60 and older, African Americans, and Hispanics.
* Includes demographic, socioeconomic, dietary, and health-related questions. The examination consists of medical, dental, and physiological measurements, as well as laboratory tests administered by trained medical personnel.

**Targets**

* Analyzing nutritional data and   
  Physical examination data to predict the presence or absence of diseases.
* Predict Disease and Obesity trends.
* Try to predict optimal nutritional requirements for various age groups

**Tools & Techniques**

* **Python**
  + Pandas, sklearn, keras, tensorflow, matplotlib, numpy libraries
* **Machine Learning Techniques**
  + K-means, Logistic Regression, SVM, Neural Networks, KNN   
    Random Forests, and other classification and regression methods.