Instructions for R functions used in paper "Statistical Analysis of a Low Cost Method for Multiple Disease Prediction" by Bayati, Bhaskar, and Montanari

1. This repository contains seven R files.
    a. functionLibrary.R contains all necessary functions that the other codes use.
    b. For each of the OLR30 and STL and each of the three methods of bootstrap, multiple imputaiton (MI), and delong there is a R code that calculates AUC confidence intervals with that method for each of the K tasks. Therefore, there are 2x3=6 codes.

2. File names are descriptive. For example, OLR30_CI_MI.R calculates confidence intervals for OLR30 using the multiple imputation approach.

3. There are a total of 4 data files. Two for each method (OLR30 and STL) that are:

    a. A file with missing values imputed via the mean imputation method and all columns centered and scaled.
    b. A data file with missing values not imputed.

4. These data files are synthetic for the codes to run. Kaggle data use agreement does not allow us to publish them on github. But anyone can go to Kaggle, accept their rules and obtain the data.