

BIOST 509: Homework 7

Instructor: Amy Willis, Biostatistics, UW

Due date: 6:30pm on November 22, 2019 via Canvas

Instructions

All questions for this homework relate to two data sets from the NHANES National Youth Fitness Survey (NNYFS). The NNYFS was performed with the National Health and Nutrition Examination survey (NHANES) to obtain information about physical activity and fitness levels in children. There is a dataset with demographic information called `Y_DEMO.csv` and a dataset with body measurements called `Y_BMX.csv`. Both are available on Canvas under **Module 7 materials**.

You can review the general information about NNYHS at <https://www.cdc.gov/nchs/nyf/index.htm>, the data documentation about the demographics data at https://wwwn.cdc.gov/Nchs/Nnyfs/Y_DEMO.htm, and the data documentation about the body measures data at https://wwwn.cdc.gov/Nchs/Nnyfs/Y_BMX.htm.

Optional but encouraged: Use R Markdown to create the pdf file with your results.

Questions

1. How many unique participants (variable name is SEQN) are in the demographics data set?
2. How many unique participants (variable name is SEQN) are in the body measures data set?
3. Join the two datasets, keeping all the participants, even those who do not appear in both of the data sets. How many participants are in the merged data set?
4. In the merged data set, what is the mean age of the participants at the time of screening (RIDAGEYR)?
5. Join the two data sets, this time excluding the participants who do not appear in both data sets. How many participants are in this merged data set?
6. Based on the merged data set from Question 5, what is the mean age of the participants at the time of screening (RIDAGEYR)?
7. Using the merged dataset from Question 5, conduct a multiple linear regression analysis with body mass index (BMXBMI) as the response and with arm circumference (BMXARMC), age (RIDAGEYR), and gender (RIAGENDR) as the explanatory variables. Provide a summary of the results. Describe the association between body mass index and arm circumference after adjusting for age and gender. Is the association significant? Similarly describe the association between body mass index and gender after adjusting for arm circumference and age. Is the association significant?