Project Assignment 3

Elikem Asudo Tsatsu Gale-Zoyiku

2023-04-01

## **Research Questions**

**What is the relationship between exercise frequency and mental health status?**

Explanatory Variable: Exercise Frequency (Quantitative) Response Variable: Mental Health Status (Categorical)

### 1. Load data set(s) and libraries

load("C:/Users/egale/OneDrive - Ashesi University/Desktop/Statistics with Probability/IPUMS\_NHIS.RData")  
library(descr)

## Warning: package 'descr' was built under R version 4.2.3

library(stats)

### 2. Create variable subset

vars=c("SEX","SEXORIEN","EDUC","MAXEDUC","WORFREQ","DEPFEELEVL","WORRX","DEPRX","DEPFREQ","POVLEV","POVERTY")  
myData = IPUMS\_NHIS[vars]  
  
str(myData)

## 'data.frame': 116291 obs. of 11 variables:  
## $ SEX : int 2 1 2 2 1 2 2 2 2 2 ...  
## $ SEXORIEN : int 2 2 2 2 0 2 2 2 2 0 ...  
## $ EDUC : int 400 201 301 303 0 400 201 400 301 0 ...  
## $ MAXEDUC : int 8 8 5 7 7 8 4 8 5 5 ...  
## $ WORFREQ : int 2 5 4 2 5 4 1 4 5 5 ...  
## $ DEPFEELEVL: int 3 0 3 0 0 1 1 1 0 0 ...  
## $ WORRX : int 2 1 1 1 0 1 2 1 1 0 ...  
## $ DEPRX : int 2 1 1 1 0 1 1 1 1 0 ...  
## $ DEPFREQ : int 4 5 4 5 5 3 4 4 5 5 ...  
## $ POVLEV : num 9.95 2.67 3.29 3.67 3.67 8.92 0.41 5.94 3.14 3.14 ...  
## $ POVERTY : int 37 32 33 34 34 37 11 37 33 33 ...

myData<-as.data.frame(myData)  
str(myData)

## 'data.frame': 116291 obs. of 11 variables:  
## $ SEX : int 2 1 2 2 1 2 2 2 2 2 ...  
## $ SEXORIEN : int 2 2 2 2 0 2 2 2 2 0 ...  
## $ EDUC : int 400 201 301 303 0 400 201 400 301 0 ...  
## $ MAXEDUC : int 8 8 5 7 7 8 4 8 5 5 ...  
## $ WORFREQ : int 2 5 4 2 5 4 1 4 5 5 ...  
## $ DEPFEELEVL: int 3 0 3 0 0 1 1 1 0 0 ...  
## $ WORRX : int 2 1 1 1 0 1 2 1 1 0 ...  
## $ DEPRX : int 2 1 1 1 0 1 1 1 1 0 ...  
## $ DEPFREQ : int 4 5 4 5 5 3 4 4 5 5 ...  
## $ POVLEV : num 9.95 2.67 3.29 3.67 3.67 8.92 0.41 5.94 3.14 3.14 ...  
## $ POVERTY : int 37 32 33 34 34 37 11 37 33 33 ...

### 3. Data management I: check for and recode errors and NAs

### 4. Data management II: further subset and create secondary variable

### 5. Descriptive statistics (sample means, standard deviations, proportions) and univariate displays

### 6. Bivariate tables and graphs

### 7. Bivariate analysis (hypothesis tests and post-hoc tests)

### 8. Moderation

### 9. Save