PASPE 2023: R básico

Tarea 5

Rossana Torres Alvarez

July 26, 2023

## Pregunta 1. Set up

1. Carga la base de Antropometría y las librerías haven, tidyverse, survey y gtsummary

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.0 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ ggplot2 3.4.2 ✔ tibble 3.1.8  
## ✔ lubridate 1.9.2 ✔ tidyr 1.3.0  
## ✔ purrr 1.0.1   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(survey)

## Loading required package: grid  
## Loading required package: Matrix  
##   
## Attaching package: 'Matrix'  
##   
## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack  
##   
## Loading required package: survival  
##   
## Attaching package: 'survey'  
##   
## The following object is masked from 'package:graphics':  
##   
## dotchart

library(haven)  
library(readstata13)  
library(gtsummary)  
knitr::opts\_chunk$set(echo = TRUE)  
  
Antropometria <- read\_stata("~/Documents/GitHub/Rbasico/files/Antropometria.dta")

1. Remueve las etiquetas de las variables

# Remuevo las etiquetas de las variables como en lab5.5  
Antropometria<-haven::zap\_labels(Antropometria)

1. Crea la variables , , e como en el lab5.5.R

# Crear mi identificador  
Antropometria$identifier <- factor(paste0("folio\_",  
 Antropometria$folio,  
 "\_\_intp\_",  
 Antropometria$intp))  
# Crear sexo\_lab  
Antropometria <- Antropometria %>%   
 mutate(sexo\_lab = case\_when(sexo == 1~ "Hombre",  
 sexo == 2 ~ "Mujer"))  
  
  
# Crear imc e imc\_cat  
Antropometria <- Antropometria %>%   
 mutate(imc = peso/(talla/100)^2,  
 imc\_cat = case\_when(imc < 18.5 ~ "Bajo peso",  
 imc >= 18.5 & imc < 25 ~ "Normal",  
 imc >= 25 & imc < 30 ~ "Sobrepeso",  
 imc >= 30 ~ "Obesidad"))  
  
Antropometria$imc\_cat <- factor(Antropometria$imc\_cat,  
 levels = c("Bajo peso",  
 "Normal",  
 "Sobrepeso",  
 "Obesidad"))

1. Selecciona las variables: , , , , y las variables de diseño de muestra: , , , de la base Antropometría y guardalas en un nuevo dataframe llamado . Asegurate de que la variable no tenga valores

# Selecciono variables que quiero, tiro NAs de ponderador  
antro\_mini <- Antropometria %>%   
 select(sexo\_lab, imc\_cat, identifier, code\_upm, pondef, est\_var) %>%   
 drop\_na(pondef)

1. Intenta recrear la siguiente tabla:

| **Characteristic** | **Hombre**, N = 55,590,4641 | **Mujer**, N = 59,495,3601 |
| --- | --- | --- |
| Categoria de IMC |  |  |
| Bajo peso | 12,474,803.0 (22.4%) | 12,045,058.8 (20.2%) |
| Normal | 17,772,647.7 (32.0%) | 17,700,020.4 (29.8%) |
| Sobrepeso | 15,781,771.3 (28.4%) | 15,115,038.3 (25.4%) |
| Obesidad | 9,561,241.7 (17.2%) | 14,635,242.5 (24.6%) |
| 1n (%) | | |

options(survey.lonely.psu = "adjust")  
survey::svydesign(ids = ~identifier,  
 strata=~est\_var,  
 weights = ~pondef,  
 data=antro\_mini)%>%  
 tbl\_svysummary(by = "sexo\_lab",   
 include = c(imc\_cat),   
 label = imc\_cat ~ "Categoria de IMC",  
 digits = everything() ~ 1)

| **Characteristic** | **Hombre**, N = 55,590,4641 | **Mujer**, N = 59,495,3601 |
| --- | --- | --- |
| Categoria de IMC |  |  |
| Bajo peso | 12,474,803.0 (22.4%) | 12,045,058.8 (20.2%) |
| Normal | 17,772,647.7 (32.0%) | 17,700,020.4 (29.8%) |
| Sobrepeso | 15,781,771.3 (28.4%) | 15,115,038.3 (25.4%) |
| Obesidad | 9,561,241.7 (17.2%) | 14,635,242.5 (24.6%) |
| 1n (%) | | |