Analysis Workflow

Nkasiobi

Inclusion criteria:

Excluded all participants with missing data on marital status variable.

Data preparation

1. Load libraries library(dplyr)
2. Load data
3. Select variables of interest; age group (AGEGR10), marital status (MARST) and SEX
4. Exclude participants with missing values on MARSTAT (98 don't know, 98 = Refusal)
5. Re-code sex variables, 1 = male and 2 = female
6. Re-code Age Group variables, 1 = "15 - 24" and 2 = "25 - 34", 3 = "35 - 44", 4 = "45 - 54", 5 = "55 - 64", 6 = "65 - 74" and 7 = "Age 75+
7. Check unique levels of marital status.
8. Exclude 3 = widowed and 6 = single never married
9. Recheck unique levels of marital status
10. Re-group marital status into : married/living with common law partner and divorced/separated. where 1 = married, 2 = living with common law partner,4 = separated and 5 = divorced.

Statistical analysis

Install and load packages

1. install.packages("webshot")
2. webshot::install_phantomjs()
3. library(knitr)
4. library(kableExtra)
5. library(webshot)
6. library(gt)
7. library(tidyr)
8. library(ggplot2)

Descriptive statistics

1. Group data by marital_status, sex and age group and generate count for each group
2. Pivot the data to spread it by Marital_status and SEX
3. Calculate row totals
4. Calculate column totals
5. Combine the data with totals
6. Create a table of summary statistics
7. Save the table as an HTML file

Logistic regression Analysis

1. Re-code Marital_status as binary variables in a new column "Divorced_Separated where married/living with common law partner = 0 (control group) and divorced/separated = 1(case group)
2. Confirm unique levels of Divorced_Separated
3. Run a logistic regression model with Sex as exposure and Divorced_Separated as outcome adjusting for Age-group
4. create a function to extract model result
5. Use the function to get the data from the logistic model
6. Plot model result using ggplot

7. Save the plot using ggsave