Antibacterial-Consumption.R

steve

2023-04-27

# Load required packages  
library(readxl)  
  
# Read in the data from the Excel file  
data <- read\_xlsx("Annex\_1\_ESAC-Net\_report\_2020\_downloadable\_tables.xlsx",   
 sheet=8,skip=3,col\_types=c('text',rep('numeric',11)),n\_max=27)

## New names:  
## • `` -> `...1`

# Rename the columns  
colnames(w8) <- c("Country", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L")  
  
# Calculate mean, standard deviation, and median for columns B-F and K-L  
data\_stats <- data.frame(  
 Country = w8$Country,  
 Mean = apply(w8[, 2:6], 1, mean),  
 SD = apply(w8[, 11:12], 1, mean),  
 Median = apply(w8[, 2:6], 1, median)  
)  
  
  
# Find the country with the highest consumption (highest mean/median from columns B-F)  
max\_consumption <- data\_stats[which.max(data\_stats$Mean), ]  
  
# Find the country with the lowest sd (lowest mean from columns K-L)  
min\_sd <- data\_stats[which.min(data\_stats$SD), ]  
  
# Print the results  
cat("Country with highest consumption: ", max\_consumption$Country, "\n")

## Country with highest consumption: Greece

cat("Descriptive statistics for highest consumption: \n")

## Descriptive statistics for highest consumption:

cat("Mean: ", max\_consumption$Mean, "\n")

## Mean: 9.698

cat("Standard deviation: ", max\_consumption$SD, "\n")

## Standard deviation: 0

cat("Median: ", max\_consumption$Median, "\n\n")

## Median: 5.76

cat("Country with lowest sd: ", min\_sd$Country, "\n")

## Country with lowest sd: Bulgaria

cat("Descriptive statistics for lowest sd: \n")

## Descriptive statistics for lowest sd:

cat("Mean: ", min\_sd$Mean, "\n")

## Mean: 7.696

cat("Standard deviation: ", min\_sd$SD, "\n")

## Standard deviation: 0

cat("Median: ", min\_sd$Median, "\n")

## Median: 4.7

#2  
  
greece <- w8[w8$Country == "Greece", 2:6]  
bulgaria <- w8[w8$Country == "Bulgaria", 2:6]  
  
t\_test <- t.test(greece, bulgaria, var.equal = TRUE)  
  
cat("Hypotheses tested: The null hypothesis is that there is no difference between the mean antibacterial consumption of the two countries. The alternative hypothesis is that there is a difference between the mean antibacterial consumption of the two countries.\n")

## Hypotheses tested: The null hypothesis is that there is no difference between the mean antibacterial consumption of the two countries. The alternative hypothesis is that there is a difference between the mean antibacterial consumption of the two countries.

cat("Test statistic: ", t\_test$statistic, "\n")

## Test statistic: 0.371114

cat("P-value: ", t\_test$p.value, "\n")

## P-value: 0.720188

cat("95% Confidence interval: [", t\_test$conf.int[1], ", ", t\_test$conf.int[2], "]\n")

## 95% Confidence interval: [ -10.4379 , 14.4419 ]

if (t\_test$p.value < 0.05) {  
 cat("Decision: We reject the null hypothesis. There is a significant difference between the mean antibacterial consumption of the two countries.\n")  
} else {  
 cat("Decision: We fail to reject the null hypothesis. There is no significant difference between the mean antibacterial consumption of the two countries.\n")  
}

## Decision: We fail to reject the null hypothesis. There is no significant difference between the mean antibacterial consumption of the two countries.

# 3. Use a non-parametric test to compare antibacterial consumption of the two countries.  
# Subset the data for the two countries of interest  
library(dplyr)  
  
# Subset the data for the two countries of interest  
country1 <- w8 %>% filter(Country == "Greece") %>% select(B:F) %>% pull()  
country2 <- w8 %>% filter(Country == "Bulgaria") %>% select(B:F) %>% pull()  
  
# Perform the Wilcoxon rank sum test  
wilcox.test(country1, country2, alternative = "two.sided")

##   
## Wilcoxon rank sum exact test  
##   
## data: country1 and country2  
## W = 0, p-value = 1  
## alternative hypothesis: true location shift is not equal to 0