Test 3: Z-test for Two Population Means (Variances Known and Unequal)

# Load Required Packages:  
library(package = dplyr)  
library(package = psych)  
library(package = BSDA)  
  
# Load Required Data:  
data = read.csv(file = "data/Data\_Test\_03.csv", header = TRUE)  
  
# Prepare Data:  
set.seed(10)  
  
sample\_Sirvan = data %>% select(Sirvan) %>%   
 sample\_n(size = 55, replace = FALSE) %>% unlist()  
  
sample\_Pishgam = data %>% select(Pishgam) %>%  
 sample\_n(size = 50, replace = FALSE) %>% unlist()  
  
# Show Data:  
headTail(x = data, top = 2, bottom = 2)

|  |  |  |  |
| --- | --- | --- | --- |
| ## |  | Sirvan | Pishgam |
| ## | 1 | 3821 | 3257 |
| ## | 2 | 2880 | 2843 |
| ## | … | … | … |
| ## | 999 | 3142 | 3279 |
| ## | 1000 | 3186 | 4167 |

# Assumption Checking:  
# 1. It is necessary that the two population variances be known.  
var(x = data$Sirvan)  
var(x = data$Pishgam)

## 482324.9  
## 747119.9

# 2. The test is accurate if the populations are normally distributed.  
shapiro.test(x = data$Sirvan)  
shapiro.test(x = data$Pishgam)

## Shapiro-Wilk normality test  
## data: data$Sirvan  
## W = 0.99884, p-value = 0.7809

## Shapiro-Wilk normality test  
## data: data$Pishgam  
## W = 0.9988, p-value = 0.7541

# setting initial parameter values:  
sigma\_Sirvan = sd(x = data$Sirvan)  
sigma\_Pishgam = sd(x = data$Pishgam)  
  
# use z.test function - alternative = "two.sided":  
z.test(x = sample\_Sirvan, y = sample\_Pishgam, alternative = "two.sided", mu = 0,   
 sigma.x = sigma\_Sirvan, sigma.y = sigma\_Pishgam, conf.level = 0.95)

## Two-sample z-Test  
## data: sample\_Sirvan and sample\_Pishgam  
## z = 1.778, p-value = 0.075  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -28.027 575.590  
## sample estimates:  
## mean of x mean of y   
## 3252.982 2979.200

# use z.test function - alternative = "greater":  
z.test(x = sample\_Sirvan, y = sample\_Pishgam, alternative = "greater", mu = 0,   
 sigma.x = sigma\_Sirvan, sigma.y = sigma\_Pishgam, conf.level = 0.95)

## Two-sample z-Test  
## data: sample\_Sirvan and sample\_Pishgam  
## z = 1.778, p-value = 0.038  
## alternative hypothesis: true difference in means is greater than 0  
## 95 percent confidence interval:  
## 20.496 NA  
## sample estimates:  
## mean of x mean of y   
## 3252.982 2979.200