Test 2: Z-test for Two Population Means (Variances Known and Equal)

# Load Required Packages:  
library(package = dplyr)  
library(package = psych)  
library(package = DescTools)  
  
# Load Required Data:  
data = read.csv(file = "data/Data\_Test\_02.csv", header = TRUE)  
  
# Prepare Data:  
**set.seed**(5)  
sample\_A = data %>% select(Tape\_A) %>%   
 sample\_n(size = 40, replace = FALSE) %>% unlist()  
  
sample\_B = data %>% select(Tape\_B) %>%  
 sample\_n(size = 50, replace = FALSE) %>% unlist()  
  
# Show Data:  
headTail(x = data, top = 2, bottom = 2)

|  |  |  |  |
| --- | --- | --- | --- |
| ## |  | Tape\_A | Tape\_B |
| ## | 1 | 4289 | 2936 |
| ## | 2 | 4398 | 3566 |
|  | … | … | … |
| ## | 999 | 2943 | 3033 |
| ## | 1000 | 1959 | 3666 |

# Assumption Checking:  
# 1. Both populations must have equal variances and this variance must be known.  
var(x = data$Tape\_A)  
var(x = data$Tape\_B)

## 379546.8

## 379546.8

# 2. The test is accurate if the populations are normally distributed.  
shapiro.test(x = data$Tape\_A)  
shapiro.test(x = data$Tape\_B)

## Shapiro-Wilk normality test  
## data: data$Tape\_A  
## W = 0.99854, p-value = 0.5826  
##  
## Shapiro-Wilk normality test  
## data: data$Tape\_B  
## W = 0.99922, p-value = 0.9599

# setting initial parameter values:  
sd\_pop = sd(x = data**$**Tape\_A)  
  
# use ZTest function - alternative = "two.sided":  
ZTest(x = sample\_A, y = sample\_B, alternative = "two.sided",  
 mu = 0, sd\_pop = sd\_pop, conf.level = 0.95)

## Two Sample z-test  
## data: sample\_A and sample\_B  
## z = 2.208, Std. Dev. Population = 616.07, p-value = 0.027  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 32.384 544.676  
## sample estimates:  
## mean of x mean of y   
## 3319.35 3030.82

# use ZTest function - alternative = " greater ":  
ZTest(x = sample\_A, y = sample\_B, alternative = "greater",  
 mu = 0, sd\_pop = sd\_pop, conf.level = 0.95)

## Two Sample z-test  
## data: sample\_A and sample\_B  
## z = 2.208, Std. Dev. Population = 616.07, p-value = 0.014  
## alternative hypothesis: true difference in means is greater than 0  
## 95 percent confidence interval:  
## 73.566 Inf  
## sample estimates:  
## mean of x mean of y   
## 3319.35 3030.82