

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTE OF COMPUTING

SCIENTIFIC METHODOLOGY FOR COMPUTING
MO430

EXERCISE 2

STUDENT: Gian Franco Joel Condori Luna
RA: 234826

UNPAIRED DATA

1.- Teste t:

```
#Read file
data <- read.csv("/exercise_2/ex2.csv", stringsAsFactors = FALSE)

set.seed(10)

#Separate data by people who have diabetes and people who do not have
diabetes
si_diabetes <- data$bp[data$type=="Yes"]
no_diabetes <- data$bp[data$type=="No"]

test <- t.test(si_diabetes, no_diabetes) # test t

print(test)
```

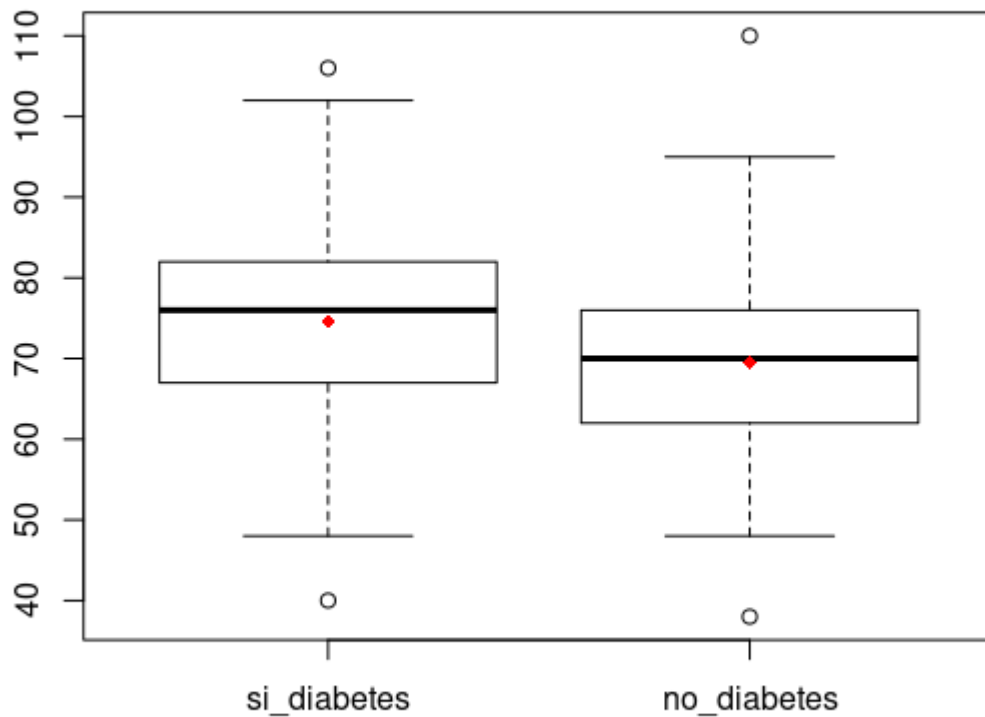
- Result:

```
data: si_diabetes and no_diabetes
t = 2.9592, df = 130.28, p-value = 0.003665
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.671482 8.414080
sample estimates:
 mean of x mean of y
74.58824 69.54545
```

- Diagrams:

```
boxplot(si_diabetes,no_diabetes,names=c("si_diabetes","no_diabetes"))#Show the diagrams
```

```
medias <- c(mean(si_diabetes),mean(no_diabetes))#Show the mean by a
point
points(medias,pch=18,col="red")#Highlight the mean of a color
```



2.- Wilcoxon rank sum

```
wilcox.test(si_diabetes, no_diabetes, alternative = "two.sided")
```

- Result:

Wilcoxon rank sum test with continuity correction

data: si_diabetes and no_diabetes

W = 5669, p-value = 0.002294

alternative hypothesis: true location shift is not equal to 0