MATH244 Lab Project Members:

-Oussama Benrjab

-Taissyr staali

Exercise 2.7:

1. How many diabetic patients are there?

Instruction : > nrow(t[t$diabetique==1,])
Answer:
[1] 229 diabetic
2. Add " aged_person" column which is true if the age is over 65.

Instruction : > t$aged_person = t$age>65
3. Extract the lines from the table that correspond to diabetic patients and put them in a new table called d. The same is true for non-diabetic patients in table nd.

Instruction : > d = t[t $ diabetique == 1,]
> nd = t[t $ diabetique == 0,]
4. In tables d and nd, add a new column "satisfactory_treatment" indicating whether the treatment of hypertension is satisfactory, i.e. whether the blood pressure is 140/90mmHg or less (130/80 mmHg in diabetics).

Instruction : > d$satisfactory_treatment = d$pas2/d$pad2 <= (130/80)
> nd$satisfactory_treatment = nd$pas2/nd$pad2 <= (140/80)
5. What is the rate of diabetic patients with satisfactory treatment?

Instruction : > nrow(d[d$satisfactory_treatment == TRUE,]) / nrow(d)
Answer:
[1] 0.3144105
6. Assemble the two tables d and nd, in order to have an overall table with the new column "satisfactory_treatment":

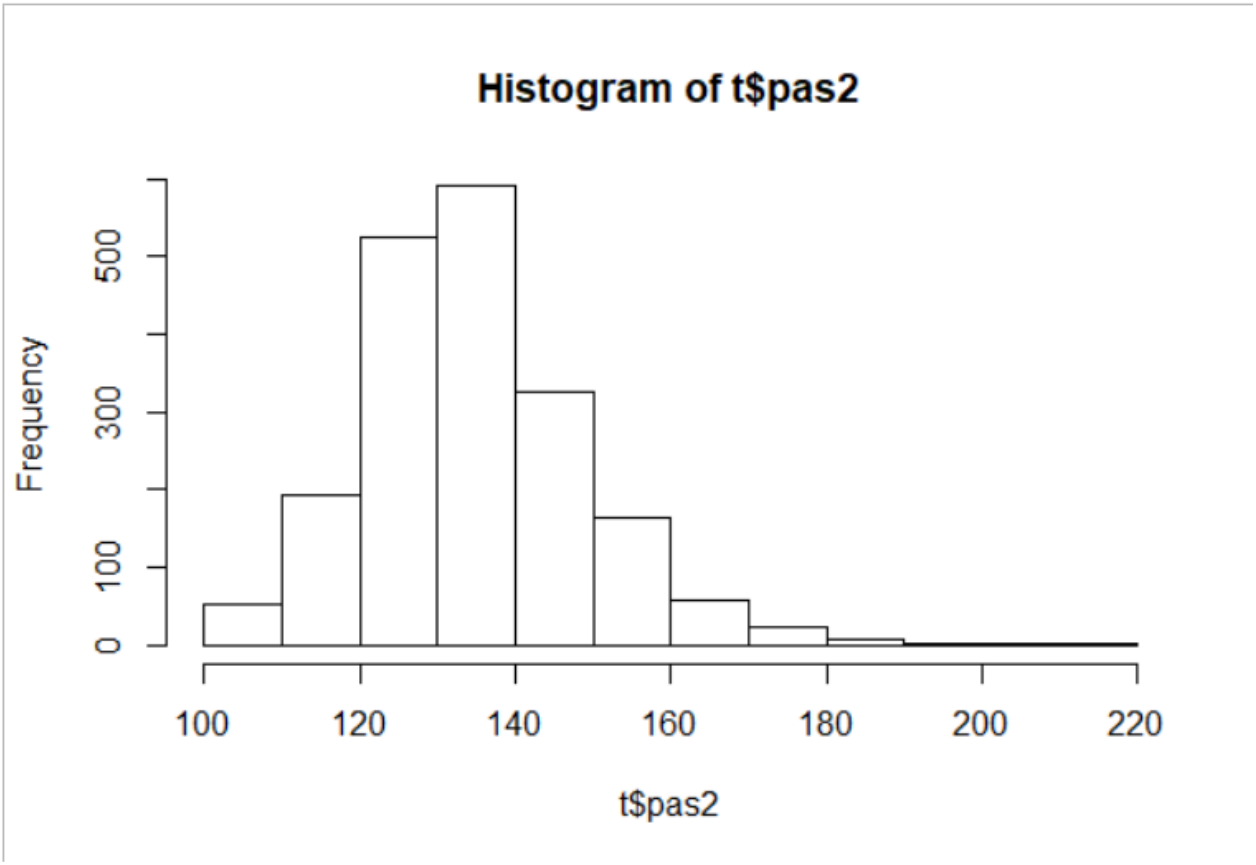
Instruction : > t = merge(d, nd, all=TRUE)
7. Extract the lines from the table that correspond to patients with unsatisfactory treatment and put them in a new table called ns.

Instruction : > ns = t[t$satisfactory_treatment == FALSE,]

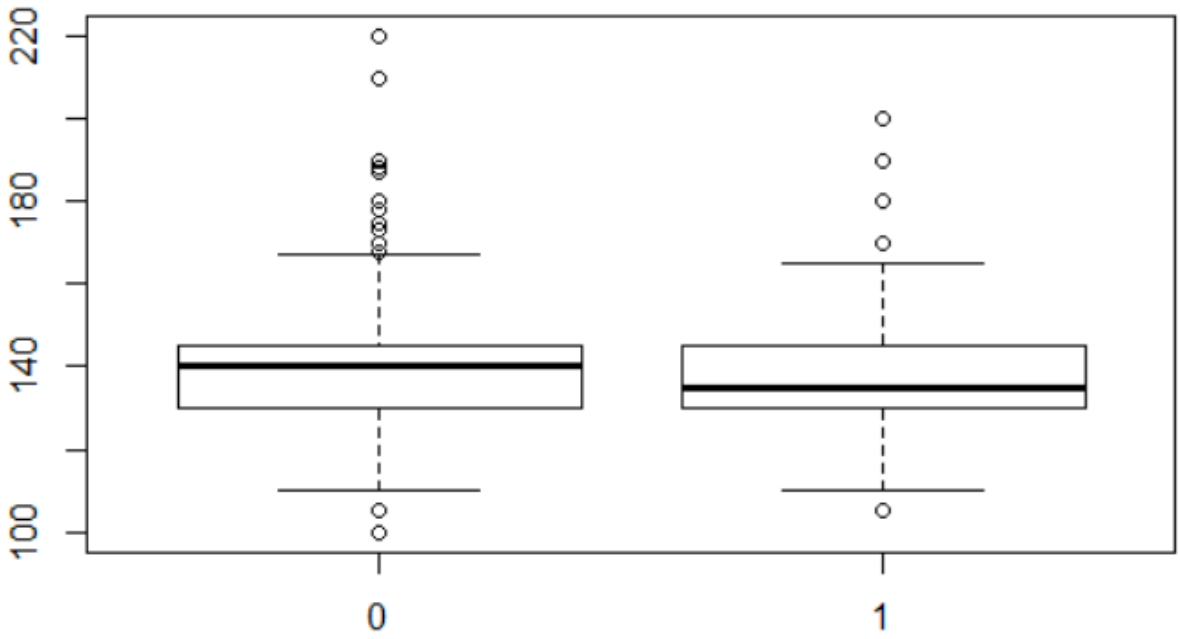
Exercise 3.6 :

1. Draw a histogram representing the systolic blood pressure.

Instruction : > hist(t$pas2)
Answer :



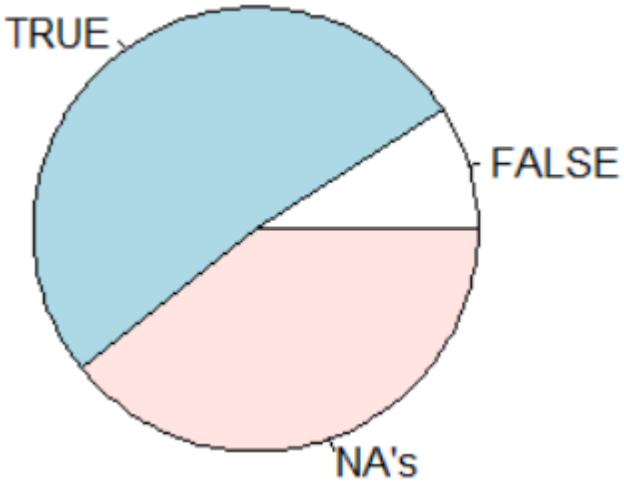
2. Plot a graph showing the systolic blood pressure as a function of the presence or absence of diabetes.
Instruction : > boxplot(t$pas2 ~ t$diabetique)
Answer :



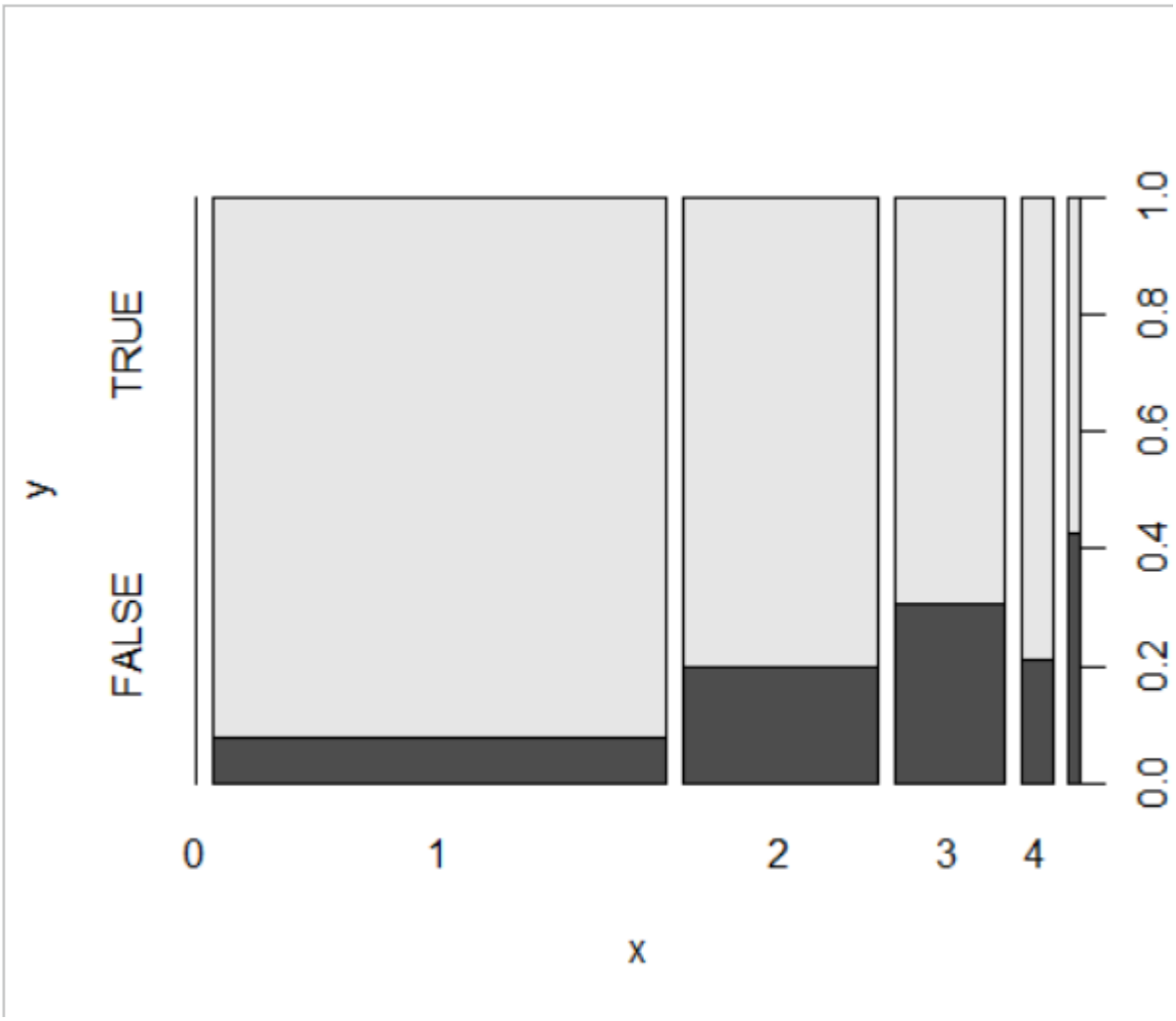
3. Draw a graph showing the change in treatment depending on the satisfactory treatment or not.
Instruction : > plot(factor(t$satisfactory_treatment), factor(t$traitement_ change))
Answer :

4. It is considered that there is "therapeutic inertia" when treatment is insufficient (blood pressure too high) and that the doctor did not change the treatment, and there has been no recent change. Pull the lines of the corresponding table to patients whose treatment is insufficient and place them in a new table "ns". Then add a column "inertia" indicating whether there is therapeutic inertia.
Instruction : > ns$inertia = (ns$traitement_change == 0) &
(ns$changement_recent == 0)
5. Draw a graph for therapeutic inertia.

Instruction : > pie(summary(factor(ns$inertia)))
6.



6. Plot a graph showing the therapeutic inertia depending on the number of drugs
Instruction : > plot(factor(ns$type2), factor(ns$inertia))
Answer :



7. How many patients have inertia been calculated for?
> summary(ns$inertia)
Mode FALSE TRUE NA's
logical 71 405 307
Answer :
405+71 = 476 patients have inertia

Exercise 4.4 :

1.Are diabetic patients significantly different in age from others?
Instruction : > t.test(dage, ndage)
Answer :
Welch Two Sample t-test
data: d$age and nd$age
t = 1.2522, df = 304.47, p-value = 0.2114

alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.5674364 2.5534759
sample estimates:
mean of x mean of y
64.99127 63.99825
Conclusion:
Assuming:
H0: There is no difference between the averages of the two groups.
Ha: There is a difference between the averages of the two groups.
Since p-value > 0.05, we cannot reject the null hypothesis as we do not have sufficient evidence to conclude that Ha is true.
2.Does the mean systolic blood pressure increase significantly between the last visit and the front? last?
Instruction : > t.test(t$pas1, t$pas2, paired=TRUE)
Answer :
Paired t-test
data: t$pas1 and t$pas2

t = 0.2514, df = 1939, p-value = 0.8015
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.5854451 0.7576101
sample estimates:
mean of the differences
0.08608247
Conclusion:
Assuming:
H0: There is no difference between the averages of the two groups.
Ha: There is a difference between the averages of the two groups.
Since p-value > 0.05, we cannot reject the null hypothesis as we do not have sufficient evidence to conclude that Ha is true.
3.Is systolic blood pressure significantly higher in the elderly?
Instruction : > t.test(t[t$aged_person == TRUE,]$pas2, t[t$aged_person = = FALSE,]$pas2)
Answer :
Welch Two Sample t-test
data: t[t$aged_person == TRUE,]$pas2 and t[t$aged_person == FALSE,]$pas2

t = 4.5422, df = 1805.6, p-value = 5.936e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
1.673689 4.217448
sample estimates:
mean of x mean of y
139.9446 136.9991
Conclusion:
Assuming:
H_0: There is no difference between the averages of the two groups.
H_a: There is a difference between the averages of the two groups.
Since p-value < 0.05, this is an unlikely scenario and we can reject the null hypothesis H_0. This suggests that the alternative H_a is true.
As written in the provided documentation, “na.rm = TRUE” removes NA results
Instruction : > mean(t[t$aged_person == TRUE,]$pas2, na.rm=TRUE)

Answer : [1] 139.9446
Instruction :> mean(t[t$aged_person == FALSE,]$pas2, na.rm=TRUE)
Answer : [1] 136.9991