Autoebaluazioa – Testa. EBAZPENA

Autoebaluazio test hau bi zatitan banatua dago:

- Lehenengo zatia hamabost galderaz osatua dago eta bakoitzak puntu bat balio du. Galdera hauek R programaren sintaxiari buruzkoak dira. Hortaz, R tresna informatikoaren sententzien idazkera egokiarekin erlazionatuta daude.
- Bigarren zatia bederatzi galderaz osatua dago eta bakoitzak puntu bat balio du. Galdera hauek inferentzia estatistikoari buruzkoak dira, eta beraz, hauek erantzuteko R-ko komando egokiak erabiliz lortutako emaitzak interpretatu beharko dira.

Ondorengo galderak erantzuteko **lattice** paketean dagoen **singer** datu-markoa erabili behar da.

1. ZATIA: R TRESNA INFORMATIKOAREN SINTAXIA (15 PUNTU)

lattice paketean dagoen **singer** datu-markoa kontsideratuko dugu.

1. Galdera: Instalatu **lattice** paketea

```
> install.packages("lattice")
```

2. Galdera: Kargatu **lattice** paketea

```
> library(lattice)
```

3. Galdera: Aztertu **lattice** paketearen barnean dauden datu multzoak

```
> data(package="lattice")
```

4. Galdera: Lortu **singer** datu-markoari buruzko informazioa

```
> help(singer)
```

H_singer {lattice} R Documentation

Heights of New York Choral Society singers

Description

Heights in inches of the singers in the New York Choral Society in 1979. The data are grouped according to voice part. The vocal range for each voice part increases in pitch according to the following order: Bass 2, Bass 1, Tenor 2, Tenor 1, Alto 2, Alto 1, Soprano 2, Soprano 1.

Usage

```
singer
```

Format

A data frame with 235 observations on the following 2 variables.

height

Height in inches of the singers.

voice.part

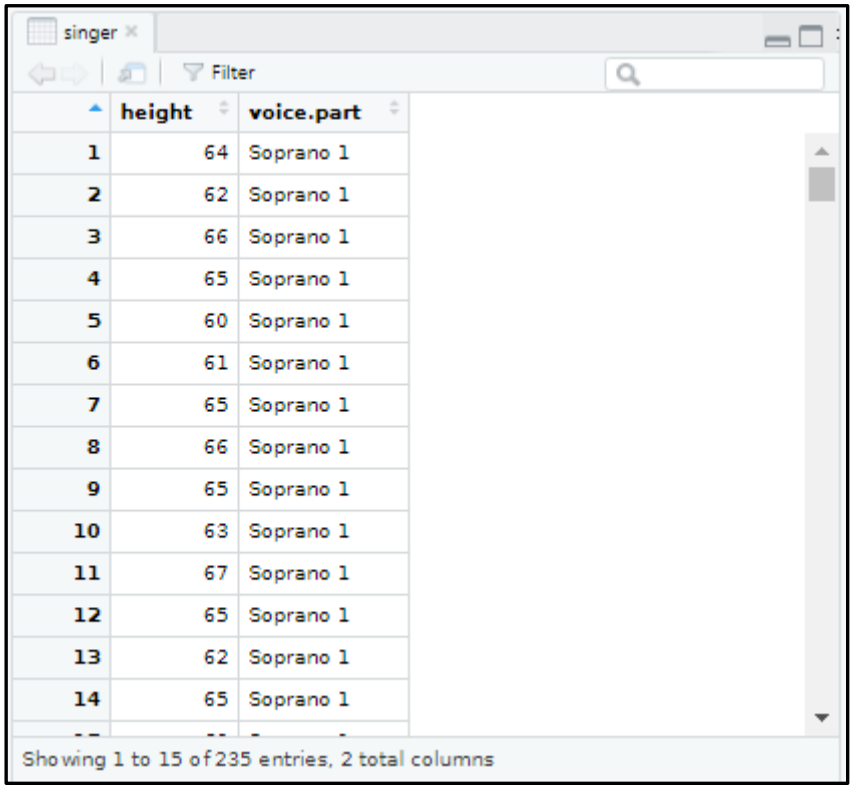
(Unordered) factor with levels "Bass 2", "Bass 1", "Tenor 2", "Tenor 1", "Alto 2", "Alto 1", "Soprano 2", "Soprano 1".

5. Galdera: **singer** datu-markoaren egitura lortu

```
> str(singer)
'data.frame':  235 obs. of  2 variables:
 $ height    : num  64 62 66 65 60 61 65 66 65 63 ...
 $ voice.part: Factor w/ 8 levels "Bass 2","Bass 1",..:
8 8 8 8 8 8 8 8 8 ...
```

6. Galdera: Bistaratu **singer** datu-markoa

```
> View(singer)
```



singer

Filter

	height	voice.part
1	64	Soprano 1
2	62	Soprano 1
3	66	Soprano 1
4	65	Soprano 1
5	60	Soprano 1
6	61	Soprano 1
7	65	Soprano 1
8	66	Soprano 1
9	65	Soprano 1
10	63	Soprano 1
11	67	Soprano 1
12	65	Soprano 1
13	62	Soprano 1
14	65	Soprano 1

Showing 1 to 15 of 235 entries, 2 total columns

7. Galdera: Bistaratu **singer** datu-markoaren lehen sei behaketak datu-markoaren egitura lortu

```
> head(singer)
```

8. Galdera: **singer** datu-markoaren aldagaiak aldagai globalak bihurtu

```
> attach(singer)
The following objects are masked from singer (pos = 3):
    height, voice.part
```

OHARRA: hemendik aurrerako galdera guztiak erantzuteko **singer** datu-markoaren aldagaiak aldagai globalak bihurtu direla suposatu da (hau da, 8. galderako erantzun zuzenaren kodea exekutatu dela suposatu da)

9. Galdera: Lortu Alto 1 ahots-sorta duten abeslarien altueren balio arraroak

```
> boxplot.stats(height[voice.part=="Alto 1"])
$stats
[1] 60.0 63.0 65.0 66.5 70.0
$n
```

```
[1] 35
$conf
[1] 64.06526 65.93474
$out
[1] 72
```

10. Galdera: Bistaratu 75 hazbete baino luzera gehiagoko abeslarien informazioa

```
> singer[height>75,]
    height voice.part
133     76    Tenor 1
155     76    Tenor 2
```

11. Galdera: Zenbat abeslarik neurtzen dute 62 hazbete baino gutxiago?

```
> length(height[height<62])
[1] 13
```

12. Galdera: Zenbat abeslarik neurtzen dute 62 hazbete baino gutxiago?

```
> length(which(height<62))
[1] 13
```

13. Galdera: Bistaratu Tenor 1 ahots-sorta duten abeslarien artetik 73 hazbete baino gehiago neurtzen duten abeslarien informazioa

```
> singer[voice.part=="Tenor 1" & height>73,]
    height voice.part
133     76    Tenor 1
134     74    Tenor 1
```

14. Galdera: Bistaratu Bass 1 ahots-sorta duten edo 74 hazbete baino gehiago neurtzen duten abeslarien informazioa

```
> singer[voice.part=="Bass 1" | height>74,]
height voice.part
133     76    Tenor 1
155     76    Tenor 2
```

```
171     72     Bass 1
172     70     Bass 1
173     72     Bass 1
174     69     Bass 1
175     73     Bass 1
176     71     Bass 1
177     72     Bass 1
178     68     Bass 1
179     68     Bass 1
180     71     Bass 1
181     66     Bass 1
182     68     Bass 1
183     71     Bass 1
184     73     Bass 1
185     73     Bass 1
186     70     Bass 1
187     68     Bass 1
188     70     Bass 1
189     75     Bass 1
190     68     Bass 1
191     71     Bass 1
192     70     Bass 1
193     74     Bass 1
194     70     Bass 1
195     75     Bass 1
196     75     Bass 1
197     69     Bass 1
198     72     Bass 1
199     71     Bass 1
200     70     Bass 1
201     71     Bass 1
202     68     Bass 1
203     70     Bass 1
204     75     Bass 1
205     72     Bass 1
206     66     Bass 1
207     72     Bass 1
208     70     Bass 1
```

```
209     69     Bass 1
211     75     Bass 2
213     75     Bass 2
224     75     Bass 2
235     75     Bass 2
```

15. Galdera: Bistaratu "Tenor 1" edo "Soprano 1" ahots-sorta duten eta 65 hazbete baino gutxiago neurtzen duten abeslarien informazioa

```
> singer[(voice.part=="Tenor 1" | voice.part=="Soprano 1")
& height<65,]
    height voice.part
1       64  Soprano 1
2       62  Soprano 1
5       60  Soprano 1
6       61  Soprano 1
10      63  Soprano 1
13      62  Soprano 1
17      63  Soprano 1
19      62  Soprano 1
22      62  Soprano 1
24      63  Soprano 1
28      62  Soprano 1
32      61  Soprano 1
36      62  Soprano 1
144     64    Tenor 1
149     64    Tenor 1
```

2. ZATIA: INFERENTZIA ESTATISTIKOA R ERABILIZ (9 PUNTU)

16. Galdera: 0.1eko adierazgarritasun mailaz zein ahots-sortako abeslarien altuerak jarraitzen du banaketa normala?

```
> shapiro.test(height[voice.part=="Tenor 2"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Tenor 2"]
W = 0.88995, p-value = 0.02245
> shapiro.test(height[voice.part=="Tenor 1"])
     Shapiro-Wilk normality test
```

```
data:  height[voice.part == "Tenor 1"]
W = 0.96571, p-value = 0.6375
> shapiro.test(height[voice.part=="Alto 2"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Alto 2"]
W = 0.87129, p-value = 0.003132
> shapiro.test(height[voice.part=="Alto 1"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Alto 1"]
W = 0.96839, p-value = 0.4005
> shapiro.test(height[voice.part=="Bass 2"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Bass 2"]
W = 0.92896, p-value = 0.07321
> shapiro.test(height[voice.part=="Bass 1"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Bass 1"]
W = 0.95899, p-value = 0.165
> shapiro.test(height[voice.part=="Soprano 2"])
     Shapiro-Wilk normality test
data:  height[voice.part == "Soprano 2"]
W = 0.96712, p-value = 0.4637
```

17. Galdera: Onar al daiteke 0.05eko adierazgarritasun mailaz "Soprano 2" ahots-sorta duten abeslarien batezbesteko altuera 64.5 hazbete baino txikiagoa delako hipotesia?

```
> Soprano2_altuera<-height[voice.part=="Soprano 2"]
> Soprano2_altuera
[1] 63 67 60 67 66 62 65 62 61 62 66 60 65 65 61 64 68
[18] 64 63 62 64 62 64 65 60 65 70 63 67 66
> t.test(Soprano2_altuera,mu=64.5, alternative="less",
conf.level=0.95)
     One Sample t-test
data:  Soprano2_altuera
t = -1.1567, df = 29, p-value = 0.1284
alternative hypothesis: true mean is less than 64.5
95 percent confidence interval:
     -Inf 64.75012
```

```
sample estimates:
mean of x
 63.96667
```

18. Galdera: Lortu 0.95eko konfiantza maila batekin "Tenor 1" ahots-sorta duten abeslarien altueraren bariantzarako konfiantza tartea

```
> Tenor1_altuera<-height[voice.part=="Tenor 1"]
> Tenor1_altuera
 [1] 69 72 71 66 76 74 71 66 68 67 70 65 72 70 68 64 73
[18] 66 68 67 64
> n_Tenor1_altuera<-length(Tenor1_altuera)
> KTsigma95<-c((n_Tenor1_altuera-1)*
var(Tenor1_altuera)/qchisq(0.975,(n_Tenor1_altuera-1)),
(n_Tenor1_altuera-1)*
var(Tenor1_altuera)/qchisq(0.025,(n_Tenor1_altuera-1)))
> KTsigma95
[1]  6.491427 23.127377
```

19. Galdera: Onar al daiteke "Alto 1" eta "Soprano 2" ahots-sorta duten abeslarien altueraren bariantza berdina dela 0.1eko adierazgarritasun mailaz?

```
> Alto1_altuera<-height[voice.part=="Alto 1"]
> Alto1_altuera
 [1] 65 62 68 67 67 63 67 66 63 72 62 61 66 64 60 61 66
[18] 66 66 62 70 65 64 63 65 69 61 66 65 61 63 64 67 66
[35] 68
> Soprano2_altuera<-height[voice.part=="Soprano 2"]
> Soprano2_altuera
 [1] 63 67 60 67 66 62 65 62 61 62 66 60 65 65 61 64 68
[18] 64 63 62 64 62 64 65 60 65 70 63 67 66
> var.test(Alto1_altuera,Soprano2_altuera,
conf.level = 0.9)
     F test to compare two variances
data:  Alto1_altuera and Soprano2_altuera
F = 1.2245, num df = 34, denom df = 29, p-value = 0.5819
alternative hypothesis: true ratio of variances is not
equal to 1
90 percent confidence interval:
```

```
 0.6685122 2.2065605
sample estimates:
ratio of variances
          1.224504
```

20. Galdera: Onar al daiteke "Alto 1" ahots-sorta duten abeslarien altuera eta "Soprano 2" ahots-sorta dutenena berdina dela 0.15eko adierazgarritasun mailaz?

```
> Alto1_altuera<-height[voice.part=="Alto 1"]
> Alto1_altuera
 [1] 65 62 68 67 67 63 67 66 63 72 62 61 66 64 60 61 66
[18] 66 66 62 70 65 64 63 65 69 61 66 65 61 63 64 67 66
[35] 68
> Soprano2_altuera<-height[voice.part=="Soprano 2"]
> Soprano2_altuera
 [1] 63 67 60 67 66 62 65 62 61 62 66 60 65 65 61 64 68
[18] 64 63 62 64 62 64 65 60 65 70 63 67 66
> t.test(Alto1_altuera,Soprano2_altuera,
alternative="two.sided", var.equal=T, conf.level=0.85)
     Two Sample t-test
data:  Alto1_altuera and Soprano2_altuera
t = 1.3813, df = 63, p-value = 0.1721
alternative hypothesis: true difference in means is not
equal to 0
85 percent confidence interval:
 -0.05055336  1.88864860
sample estimates:
mean of x mean of y
 64.88571  63.96667
```

21. Galdera: Onar al daiteke abeslarien %25ak "Soprano" ahots-sorta duela %20ko adierazgarritasun mailaz?

```
> n_S=length(voice.part[voice.part=="Soprano 1" |
voice.part=="Soprano 2"]);n_S
[1] 66
> n_A=length(voice.part[voice.part=="Alto 1" |
voice.part=="Alto 2"]);n_A
[1] 62
```

```
> n_T=length(voice.part[voice.part=="Tenor 1" |
voice.part=="Tenor 2"]);n_T
[1] 42
> n_B=length(voice.part[voice.part=="Bass 1" |
voice.part=="Bass 2"]);n_B
[1] 65
> prop.test(n_S,(n_S+n_A+n_B+n_T),0.25, conf.level=0.8)
     1-sample proportions test with continuity
     correction
data:  n_S out of (n_S + n_A + n_B + n_T), null probability
0.25
X-squared = 1.034, df = 1, p-value = 0.3092
alternative hypothesis: true p is not equal to 0.25
80 percent confidence interval:
 0.2428744 0.3220412
sample estimates:
        p
0.2808511
```

22. Galdera: Onar al daiteke abeslarien %25ak ”Tenor” ahots-sorta duela %1eko adierazgarritasun mailaz?

```
prop.test(n_T,(n_S+n_A+n_B+n_T),0.25,conf.level=0.99)
     1-sample proportions test with continuity
     correction
data:  n_T out of (n_S + n_A + n_B + n_T), null probability
0.25
X-squared = 5.9929, df = 1, p-value = 0.01436
alternative hypothesis: true p is not equal to 0.25
99 percent confidence interval:
 0.1216671 0.2539928
sample estimates:
        p
0.1787234
```

23. Galdera: Onar al daiteke abeslarien %50ak 68 hazbete baino gehiago neurtzen duten hipotesia 0.20ko adierazgarritasun mailaz?

```
> tamaina68<-length(height[height>=68]);tamaina68
[1] 107
```

```
> prop.test(tamaina68,(n_S+n_A+n_B+n_T),0.5,
conf.level=0.8)

     1-sample proportions test with continuity
     correction

data:  tamaina68 out of (n_S + n_A + n_B + n_T), null
probability 0.5
X-squared = 1.7021, df = 1, p-value = 0.192
alternative hypothesis: true p is not equal to 0.5
80 percent confidence interval:
 0.4120437 0.4992464
sample estimates:
        p
0.4553191
```

24. Galdera: Abeslarien altueraren mediana erabiliz, hauek altua edo baxua bezala sailkatu eta aldagai hori eta abeslari-mota aldagaia erabiliz kontingentzia-taula eraiki ondoren, hurrengo erantzunetatik egokia dena aukeratu:

```
> mediana<-median(height)
> height.kat<-as.factor(ifelse(height<mediana,"Baxua",
"Altua"))
> kontig_taula<-table(voice.part,height.kat)
> kontig_taula
height.kat
voice.part  Altua Baxua
  Bass 2       25     1
  Bass 1       37     2
  Tenor 2      20     1
  Tenor 1      15     6
  Alto 2        8    19
  Alto 1        9    26
  Soprano 2     5    25
  Soprano 1     2    34
```