Practica

Laura Chaparro

20/1/2022

#**Practica Penguins**

Cargamos las librerias y los datos a utilizar

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(plyr)

## ------------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.  
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:  
## library(plyr); library(dplyr)

## ------------------------------------------------------------------------------

##   
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

library(ggplot2)  
library(palmerpenguins)  
data(package = 'palmerpenguins')

Observamos los datos

View(penguins)  
View(penguins\_raw)  
str(penguins)

## tibble [344 x 8] (S3: tbl\_df/tbl/data.frame)  
## $ species : Factor w/ 3 levels "Adelie","Chinstrap",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ island : Factor w/ 3 levels "Biscoe","Dream",..: 3 3 3 3 3 3 3 3 3 3 ...  
## $ bill\_length\_mm : num [1:344] 39.1 39.5 40.3 NA 36.7 39.3 38.9 39.2 34.1 42 ...  
## $ bill\_depth\_mm : num [1:344] 18.7 17.4 18 NA 19.3 20.6 17.8 19.6 18.1 20.2 ...  
## $ flipper\_length\_mm: int [1:344] 181 186 195 NA 193 190 181 195 193 190 ...  
## $ body\_mass\_g : int [1:344] 3750 3800 3250 NA 3450 3650 3625 4675 3475 4250 ...  
## $ sex : Factor w/ 2 levels "female","male": 2 1 1 NA 1 2 1 2 NA NA ...  
## $ year : int [1:344] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 ...

summary(penguins)

## species island bill\_length\_mm bill\_depth\_mm   
## Adelie :152 Biscoe :168 Min. :32.10 Min. :13.10   
## Chinstrap: 68 Dream :124 1st Qu.:39.23 1st Qu.:15.60   
## Gentoo :124 Torgersen: 52 Median :44.45 Median :17.30   
## Mean :43.92 Mean :17.15   
## 3rd Qu.:48.50 3rd Qu.:18.70   
## Max. :59.60 Max. :21.50   
## NA's :2 NA's :2   
## flipper\_length\_mm body\_mass\_g sex year   
## Min. :172.0 Min. :2700 female:165 Min. :2007   
## 1st Qu.:190.0 1st Qu.:3550 male :168 1st Qu.:2007   
## Median :197.0 Median :4050 NA's : 11 Median :2008   
## Mean :200.9 Mean :4202 Mean :2008   
## 3rd Qu.:213.0 3rd Qu.:4750 3rd Qu.:2009   
## Max. :231.0 Max. :6300 Max. :2009   
## NA's :2 NA's :2

##**Punto 1**

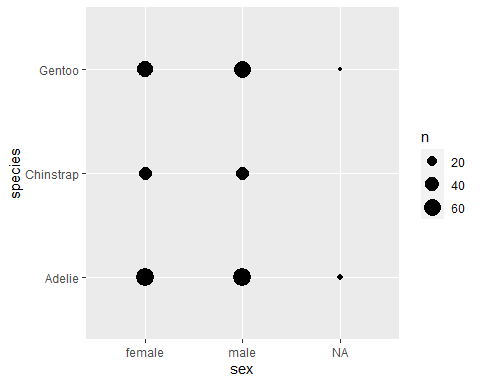
Número de individuos totales : **344**

length(penguins$species)

## [1] 344

masculinos y femeninos por especie.

ggplot(data = penguins)+  
 geom\_count(aes(x= sex, y = species))



table(penguins$sex, penguins$species)

##   
## Adelie Chinstrap Gentoo  
## female 73 34 58  
## male 73 34 61

La media, desviación estándar, valor mínimo y máximo de la longitud y profundidad del pico, la longitud de la aleta y el tamaño

*Media*

penguins <- na.omit(penguins)  
  
tapply(penguins$bill\_length\_mm, penguins$species, mean)

## Adelie Chinstrap Gentoo   
## 38.82397 48.83382 47.56807

tapply(penguins$bill\_depth\_mm, penguins$species, mean)

## Adelie Chinstrap Gentoo   
## 18.34726 18.42059 14.99664

tapply(penguins$flipper\_length\_mm, penguins$species, mean)

## Adelie Chinstrap Gentoo   
## 190.1027 195.8235 217.2353

tapply(penguins$body\_mass\_g, penguins$species, mean)

## Adelie Chinstrap Gentoo   
## 3706.164 3733.088 5092.437

*Desviacion estandar*

tapply(penguins$bill\_length\_mm, penguins$species, sd)

## Adelie Chinstrap Gentoo   
## 2.662597 3.339256 3.106116

tapply(penguins$bill\_depth\_mm, penguins$species, sd)

## Adelie Chinstrap Gentoo   
## 1.219338 1.135395 0.985998

tapply(penguins$flipper\_length\_mm, penguins$species, sd)

## Adelie Chinstrap Gentoo   
## 6.521825 7.131894 6.585431

tapply(penguins$body\_mass\_g, penguins$species, sd)

## Adelie Chinstrap Gentoo   
## 458.6201 384.3351 501.4762

*Valor maximo*

tapply(penguins$bill\_length\_mm, penguins$species, max)

## Adelie Chinstrap Gentoo   
## 46.0 58.0 59.6

tapply(penguins$bill\_depth\_mm, penguins$species, max)

## Adelie Chinstrap Gentoo   
## 21.5 20.8 17.3

tapply(penguins$flipper\_length\_mm, penguins$species, max)

## Adelie Chinstrap Gentoo   
## 210 212 231

tapply(penguins$body\_mass\_g, penguins$species, max)

## Adelie Chinstrap Gentoo   
## 4775 4800 6300

*Valor minimo*

tapply(penguins$bill\_length\_mm, penguins$species, min)

## Adelie Chinstrap Gentoo   
## 32.1 40.9 40.9

tapply(penguins$bill\_depth\_mm, penguins$species, min)

## Adelie Chinstrap Gentoo   
## 15.5 16.4 13.1

tapply(penguins$flipper\_length\_mm, penguins$species, min)

## Adelie Chinstrap Gentoo   
## 172 178 203

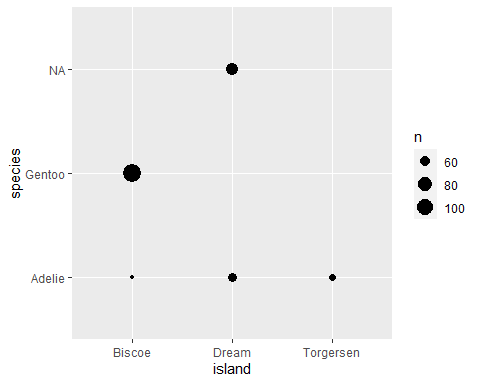
tapply(penguins$body\_mass\_g, penguins$species, min)

## Adelie Chinstrap Gentoo   
## 2850 2700 3950

##**Punto 2**

Un gráfico de barras que represente el número de individuos muestreados de cada especie en cada isla, representando las especies en diferentes colores (chinstrap – morado, gentoo – azul, adelie – naranja).

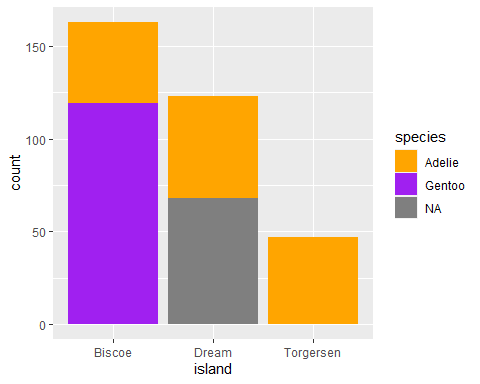
penguins <- penguins %>%  
 mutate(species = factor(species, levels = c('Adelie', 'Chinistrap', 'Gentoo')))  
  
ggplot(data = penguins) +  
 geom\_count(aes(x = island, y = species))



table(penguins$island, penguins$species )

##   
## Adelie Chinistrap Gentoo  
## Biscoe 44 0 119  
## Dream 55 0 0  
## Torgersen 47 0 0

ggplot(data = penguins) +  
 geom\_bar(aes(x = island, fill = species))+  
 scale\_fill\_manual(values = c("orange", "purple", "blue"))



Contesta a las siguientes preguntas:

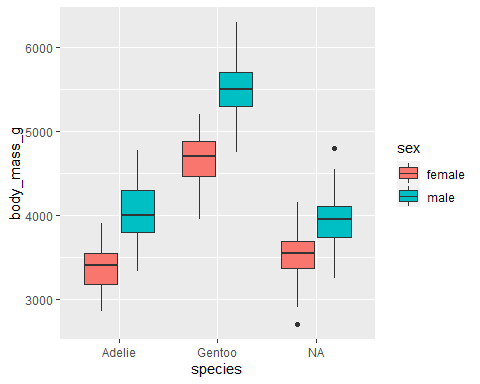
**1.**¿qué especie se ha muestreado en las tres islas?: **Adeline**

**2.**¿cuántos individuos se han muestreado de la isla Dream? **55**

##**Punto 3**

Un gráfico multipanel de cajas y bigotes del tamaño de los pingüinos según su sexo, donde aparezca un panel para cada especie.

ggplot(data = penguins)+  
 geom\_boxplot(aes(x = species, y = body\_mass\_g, fill = sex ))



Contesta a las siguientes preguntas:

**1.**¿qué especie tiene mayor tamaño? **Gentoo**

**2.**¿en qué especie las hembras y los machos tienen un tamaño más similar? **Chinistrap o (NA)**

citation("ggplot2")

##   
## To cite ggplot2 in publications, please use:  
##   
## H. Wickham. ggplot2: Elegant Graphics for Data Analysis.  
## Springer-Verlag New York, 2016.  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Book{,  
## author = {Hadley Wickham},  
## title = {ggplot2: Elegant Graphics for Data Analysis},  
## publisher = {Springer-Verlag New York},  
## year = {2016},  
## isbn = {978-3-319-24277-4},  
## url = {https://ggplot2.tidyverse.org},  
## }

citation("dplyr")

##   
## To cite package 'dplyr' in publications use:  
##   
## Hadley Wickham, Romain François, Lionel Henry and Kirill Müller  
## (2021). dplyr: A Grammar of Data Manipulation. R package version  
## 1.0.7. https://CRAN.R-project.org/package=dplyr  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Manual{,  
## title = {dplyr: A Grammar of Data Manipulation},  
## author = {Hadley Wickham and Romain François and Lionel Henry and Kirill Müller},  
## year = {2021},  
## note = {R package version 1.0.7},  
## url = {https://CRAN.R-project.org/package=dplyr},  
## }

citation("plyr")

##   
## To cite plyr in publications use:  
##   
## Hadley Wickham (2011). The Split-Apply-Combine Strategy for Data  
## Analysis. Journal of Statistical Software, 40(1), 1-29. URL  
## http://www.jstatsoft.org/v40/i01/.  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Article{,  
## title = {The Split-Apply-Combine Strategy for Data Analysis},  
## author = {Hadley Wickham},  
## journal = {Journal of Statistical Software},  
## year = {2011},  
## volume = {40},  
## number = {1},  
## pages = {1--29},  
## url = {http://www.jstatsoft.org/v40/i01/},  
## }

citation("palmerpenguins")

##   
## To cite palmerpenguins in publications use:  
##   
## Horst AM, Hill AP, Gorman KB (2020). palmerpenguins: Palmer  
## Archipelago (Antarctica) penguin data. R package version 0.1.0.  
## https://allisonhorst.github.io/palmerpenguins/  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Manual{,  
## title = {palmerpenguins: Palmer Archipelago (Antarctica) penguin data},  
## author = {Allison Marie Horst and Alison Presmanes Hill and Kristen B Gorman},  
## year = {2020},  
## note = {R package version 0.1.0},  
## url = {https://allisonhorst.github.io/palmerpenguins/},  
## }