

DATOS MORFOMETRICOS

Guillermo Aramayo Tapia

2023-09-01



Sección 1. Introducción

Esta información corresponde a una pequeña parte de un tratamiento estadístico preliminar, que se encuentra en pleno proceso, el cual corresponde a un estudio de camarones en el norte de Chile, Por lo tanto, la estadística preliminar es necesaria para tener una visión general del comportamiento de distintas variables y realizar así un tratamiento robusto de los datos.

Sección 2. Lectura y estadística de datos

1. *Lectura de base de datos*

```
BD_morfo<-read.csv("BD_datos.csv",header = TRUE, dec=".",sep = ";")
```

2. *Extructura y un summary de los datos*

```
str(BD_morfo)
```

```
'data.frame':  254 obs. of  9 variables:
 $ ID          : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Localidad   : chr  "Huancarane" "Huancarane" "Taltape Bajo" "Taltape Bajo" ...
```

```

$ Fecha      : chr  "03-11-2022" "03-11-2022" "03-11-2022" "03-11-2022" ...
$ Sexo       : chr  "Macho" "Macho" "Macho" "Macho" ...
$ E_reprod   : chr  "" "" "" "" ...
$ Peso..grs.: num  123.2 114.8 149.5 79.6 55.8 ...
$ Long_cef   : num  69.1 60 64.4 53.6 47.2 ...
$ Long_Total: num  151 144 148 135 116 ...
$ Art_Pesca  : chr  "G" "G" "G" "G" ...

```

```
summary(BD_morfo)
```

ID	Localidad	Fecha	Sexo
Min. : 1.00	Length:254	Length:254	Length:254
1st Qu.: 64.25	Class :character	Class :character	Class :character
Median :127.50	Mode :character	Mode :character	Mode :character
Mean :127.50			
3rd Qu.:190.75			
Max. :254.00			
E_reprod	Peso..grs.	Long_cef	Long_Total
Length:254	Min. : 0.040	Min. : 3.340	Min. : 12.26
Class :character	1st Qu.: 0.180	1st Qu.: 6.902	1st Qu.: 19.96
Mode :character	Median : 4.045	Median :18.495	Median : 51.06
	Mean : 19.749	Mean :22.607	Mean : 59.11
	3rd Qu.: 17.692	3rd Qu.:33.175	3rd Qu.: 84.94
	Max. :203.050	Max. :69.090	Max. :160.98
Art_Pesca			
Length:254			
Class :character			
Mode :character			

Aquí necesitamos saber las fechas de los muestreos y poder asignarlos a una estación.

```
fechas_distintas <- BD_morfo %>% distinct(Fecha) # Cuantas fechas distintas hay
fechas_distintas
```

```
      Fecha
1 03-11-2022
2 07-01-2023
3 13-05-2023
4 14-05-2023
```

3. Creamos la variable Estación con las fechas detectadas.

```
BD_morfo<- BD_morfo %>%
  mutate(Estacion = case_when(
    Fecha == "03-11-2022" ~ "Primavera",
    Fecha == "07-01-2023" ~ "Verano",
    Fecha == "13-05-2023" ~ "Otoño",
    Fecha == "14-05-2023" ~ "Otoño",
    TRUE ~ ""
  ))
```

4. Revisamos los nombres de las variables de nuestra data

```
names(BD_morfo)
```

```
[1] "ID"          "Localidad"  "Fecha"      "Sexo"       "E_reprod"
[6] "Peso..grs." "Long_cef"   "Long_Total" "Art_Pesca"  "Estacion"
```

5. Dividimos la variable Long_cef en una nueva variable “grupo”, en juvenil y adulto , usando dplyr

```
BD_morfo <- BD_morfo %>%
  mutate(Categoria = ifelse(Long_cef <= 30.00, "Juvenil", "Adulto"))
```

6. Aquí le damos la clase correspondiente a las variables de tipo factor

```
BD_morfo$Localidad<-as.factor(BD_morfo$Localidad)
BD_morfo$Sexo<-as.factor(BD_morfo$Sexo)
BD_morfo$Art_Pesca<-as.factor(BD_morfo$Art_Pesca)
BD_morfo$Categoria<-as.factor(BD_morfo$Categoria)
```

7. Calculamos el valor del SMII

```
longitud <- BD_morfo$Long_cef # longitud corporal
length(longitud)
masa <- BD_morfo$Peso..grs.    # masa corporal
length(masa)

SMA <- (longitud - mean(longitud)) / sd(longitud) # Calcular el SMA
bSMA <- exp(SMA) #Calcular bSMA
BD_morfo$SMII <- masa / bSMA # Agregar la columna SMII a BD_morfo
```

8. Revisamos los niveles de la variable Localidad

```
levels(BD_morfo$Localidad)
```

```
[1] "Conanoxa"      "Cuya"          "Desembocadura" "Huancarane"
[5] "Taltape Bajo"
```

```
# [1] "Conanoxa"      "Cuya"          "Desembocadura" "Huancarane"    "Taltape Bajo"
```

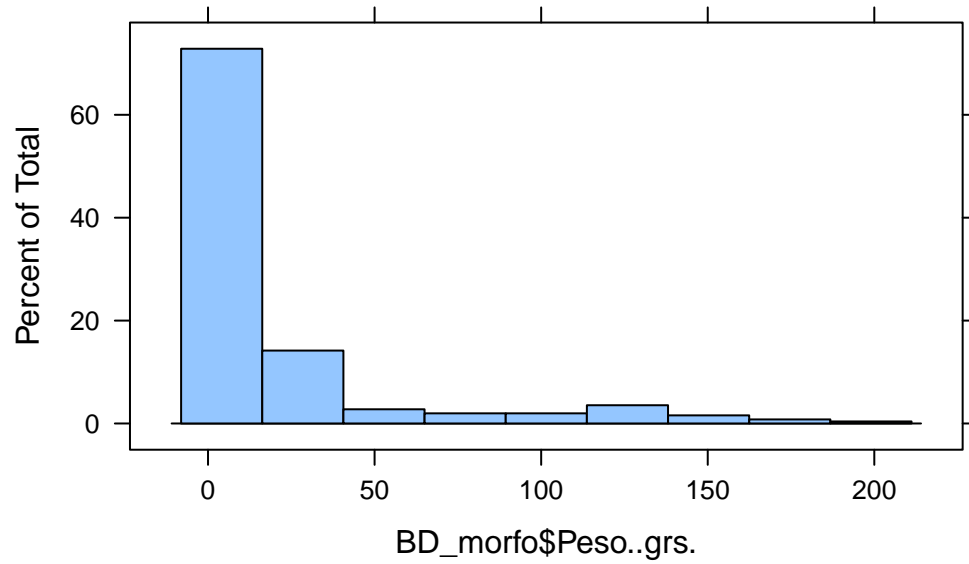
```
# Revisamos los niveles de Sexo
levels(BD_morfo$Sexo)
```

```
[1] "Hembra" "Indet"  "Macho"
```

Sección 3. Gráficos

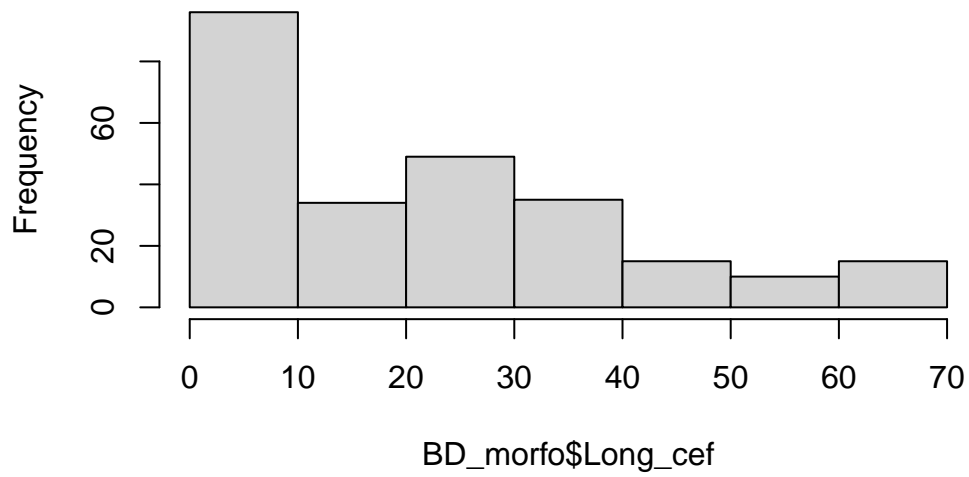
1. *Revisamos histogramas*

```
histogram(BD_morfo$Peso..grs.)
```



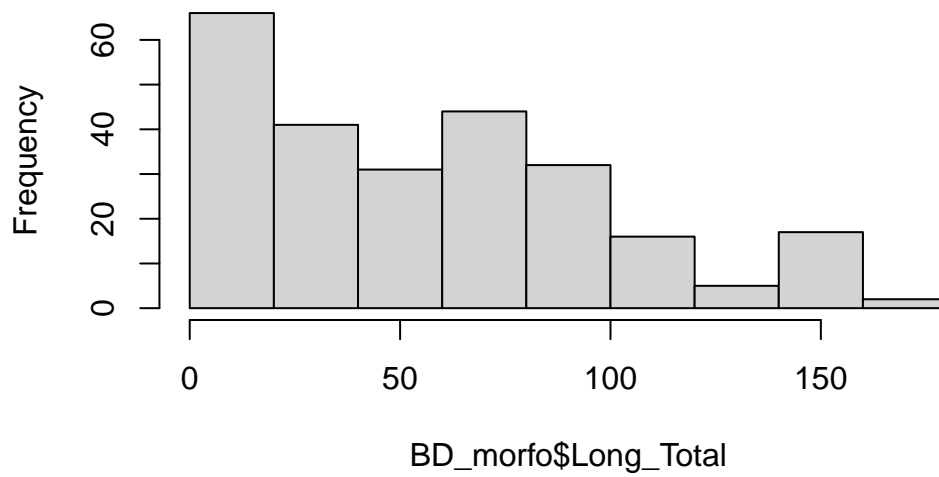
```
hist(BD_morfo$Long_cef)
```

Histogram of BD_morfo\$Long_cef



```
hist(BD_morfo$Long_Total)  
hist(BD_morfo$Long_Total)
```

Histogram of BD_morfo\$Long_Total



```
hist(BD_morfo$SMII)
```



2. Gráficos de barra

```
# Filtramos por Long_cef > 21.58 y sin categoría SEXO = indeterminados
BD_morfo_menor_inde <- subset(BD_morfo, Long_cef > "21,58" & !BD_morfo$Sexo=="Indet")
summary(BD_morfo_menor_inde)
```

ID	Localidad	Fecha	Sexo
Min. : 1.00	Conanoxa :32	Length:118	Hembra:64
1st Qu.: 31.25	Cuya :30	Class :character	Indet : 0
Median : 60.50	Desembocadura: 1	Mode :character	Macho :54
Mean : 77.53	Huancarane :35		
3rd Qu.: 89.75	Taltape Bajo :20		
Max. :251.00			

E_reprod	Peso..grs.	Long_cef	Long_Total	Art_Pesca
Length:118	Min. : 5.79	Min. :21.62	Min. : 40.59	D : 4
Class :character	1st Qu.: 10.50	1st Qu.:27.28	1st Qu.: 72.42	G :41
Mode :character	Median : 20.48	Median :33.97	Median : 90.95	PE-D:73
	Mean : 41.37	Mean :38.07	Mean : 95.86	

```

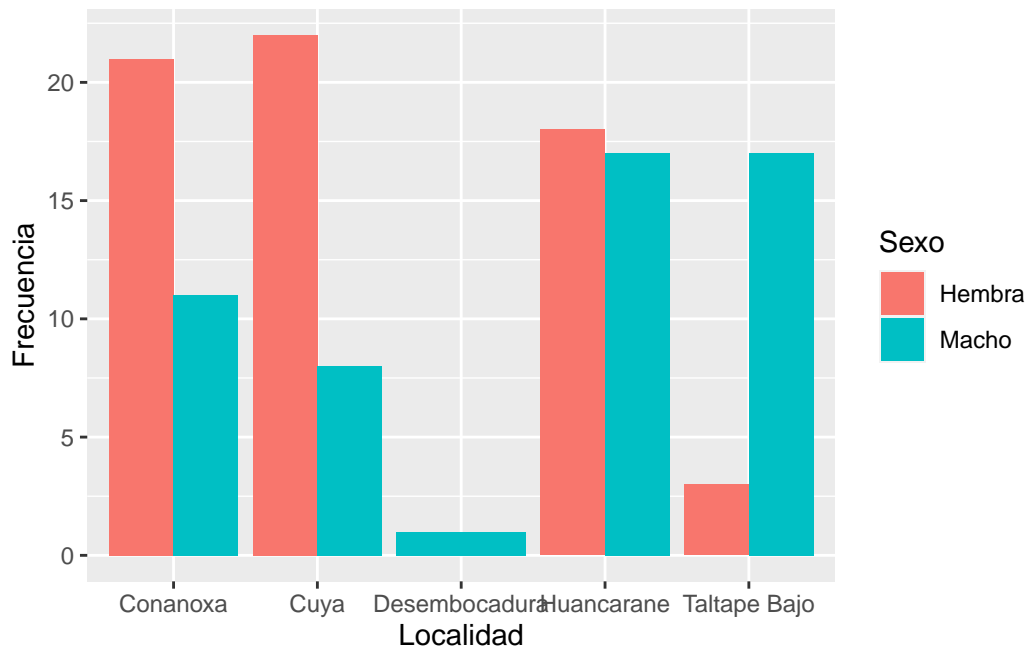
3rd Qu.: 45.87    3rd Qu.:43.52    3rd Qu.:109.56
Max.      :203.05  Max.      :69.09    Max.      :160.98
Estacion   Categoria    SMII
Length:118      Adulto :75    Min.      : 4.899
Class :character Juvenil:43    1st Qu.: 8.154
Mode  :character          Median :10.025
                                Mean  :10.468
                                3rd Qu.:13.020
                                Max.   :17.556

```

```

# Cantidad de individuos por localidad y sexo
ggplot(data = BD_morfo_menor_inde, aes(x = Localidad, fill = Sexo)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")

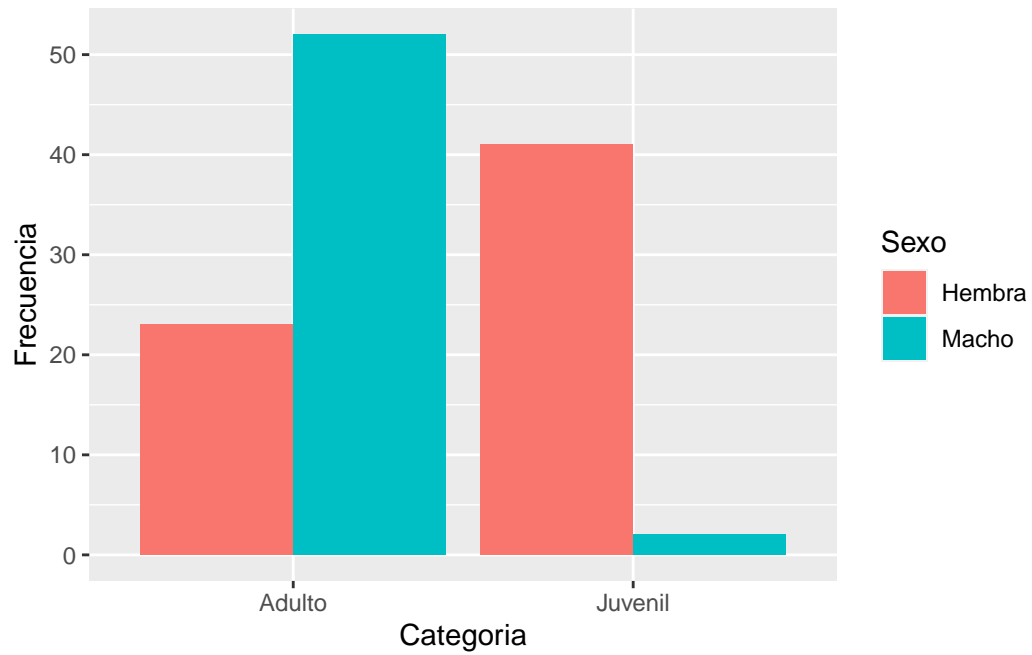
```



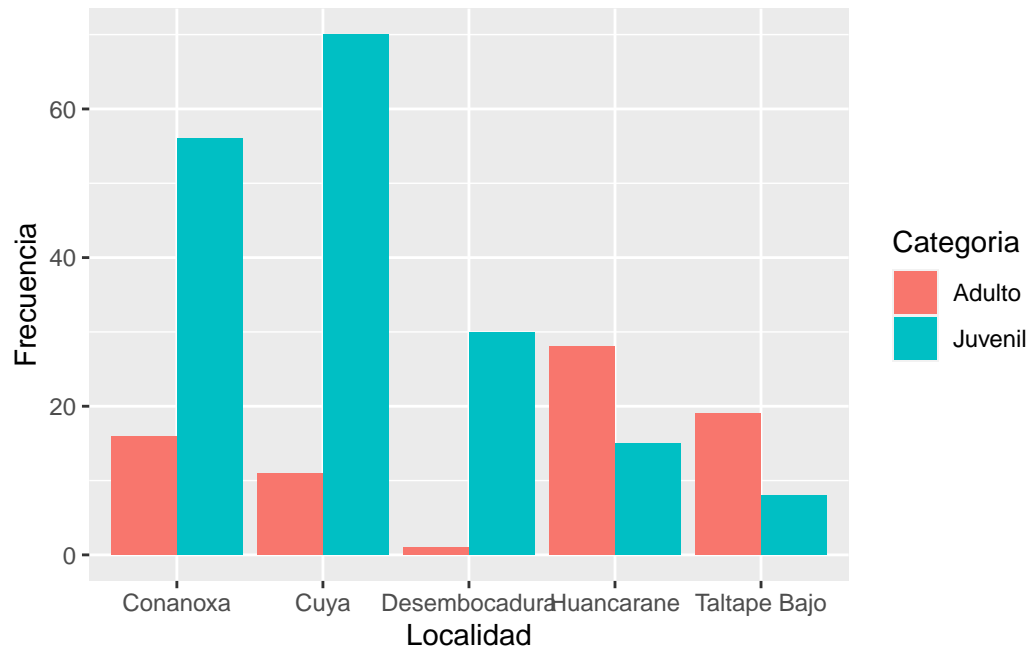
```

# Cantidad de individuos por categoría y sexo
ggplot(data = BD_morfo_menor_inde, aes(x = Categoria, fill = Sexo)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")

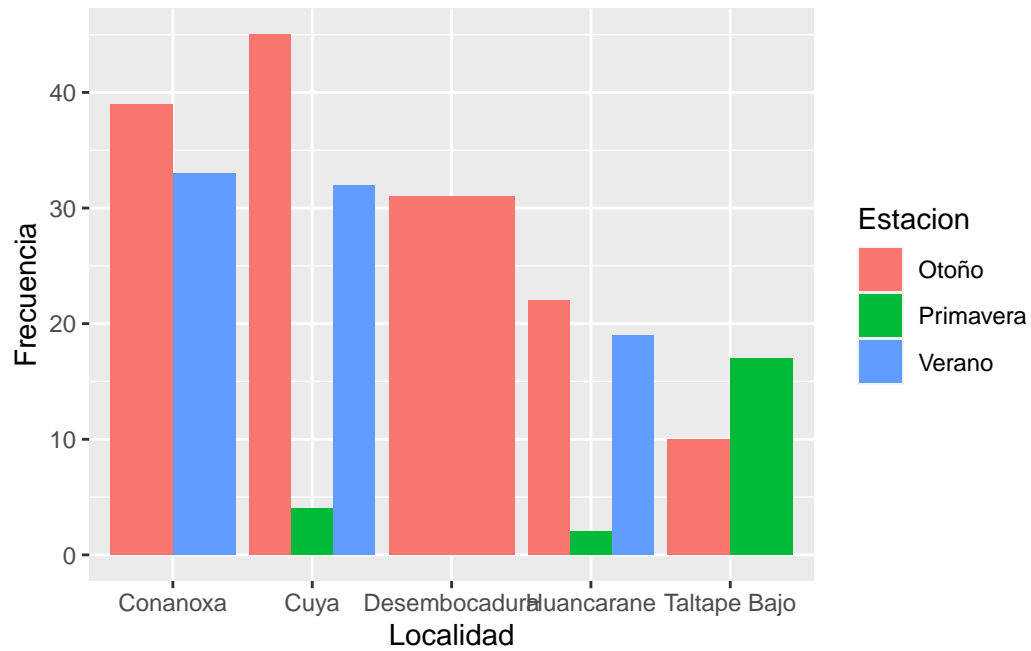
```

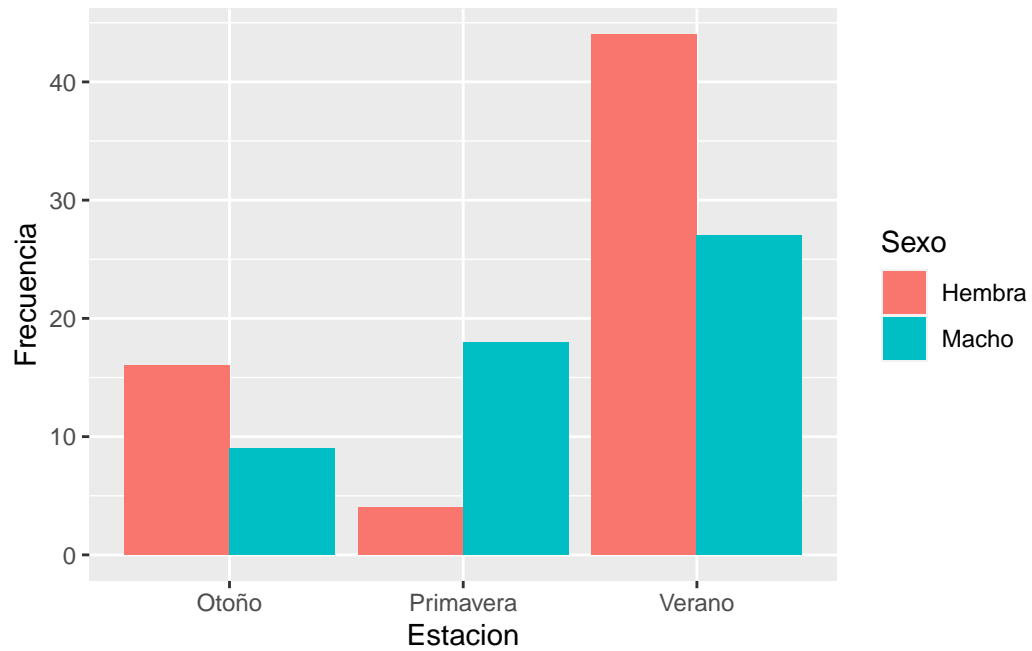
```
# Cantidad de individuos por localidad y categoría
ggplot(data = BD_morfo, aes(x = Localidad, fill = Categoria)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```



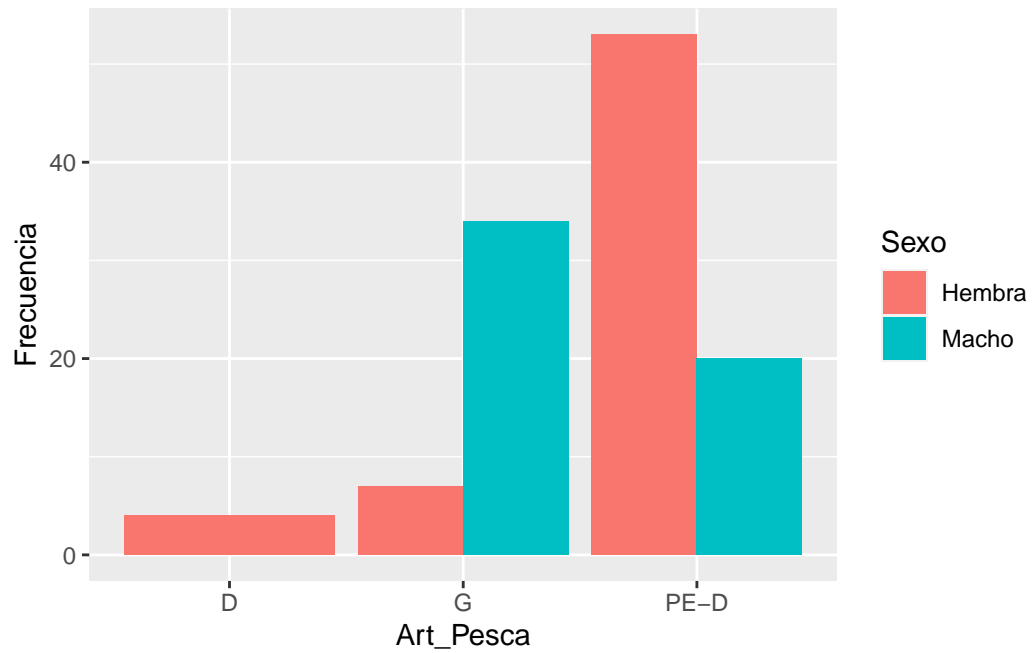
```
# Cantidad de individuos por localidad y Estación
ggplot(data = BD_morfo, aes(x = Localidad, fill = Estacion)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```



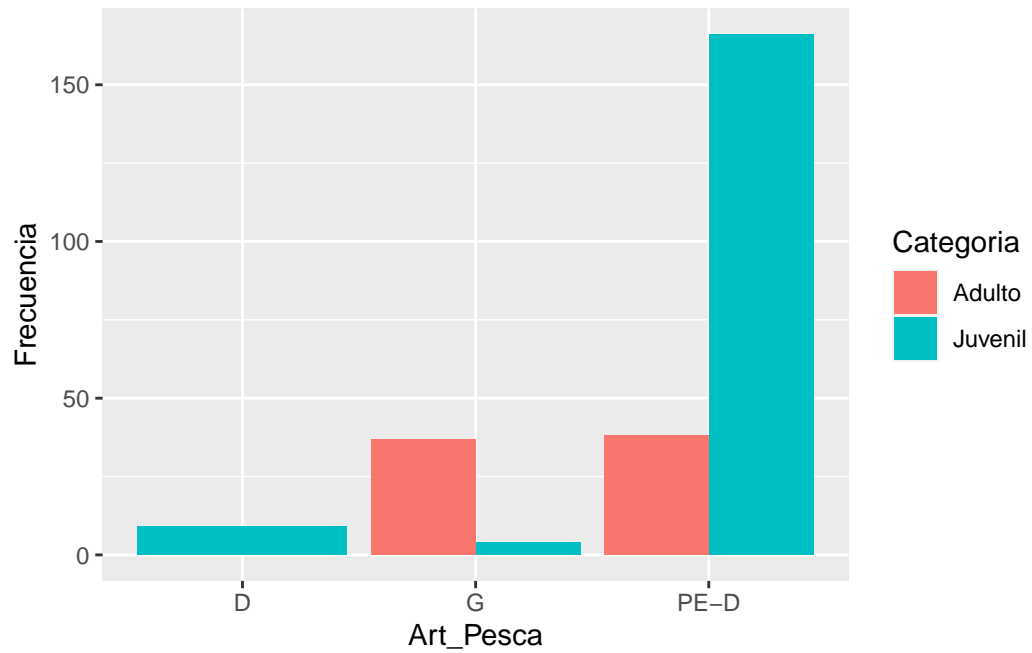
```
# Cantidad de individuos por Estación y sexo
ggplot(data = BD_morfo_menor_inde, aes(x = Estacion, fill = Sexo)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```



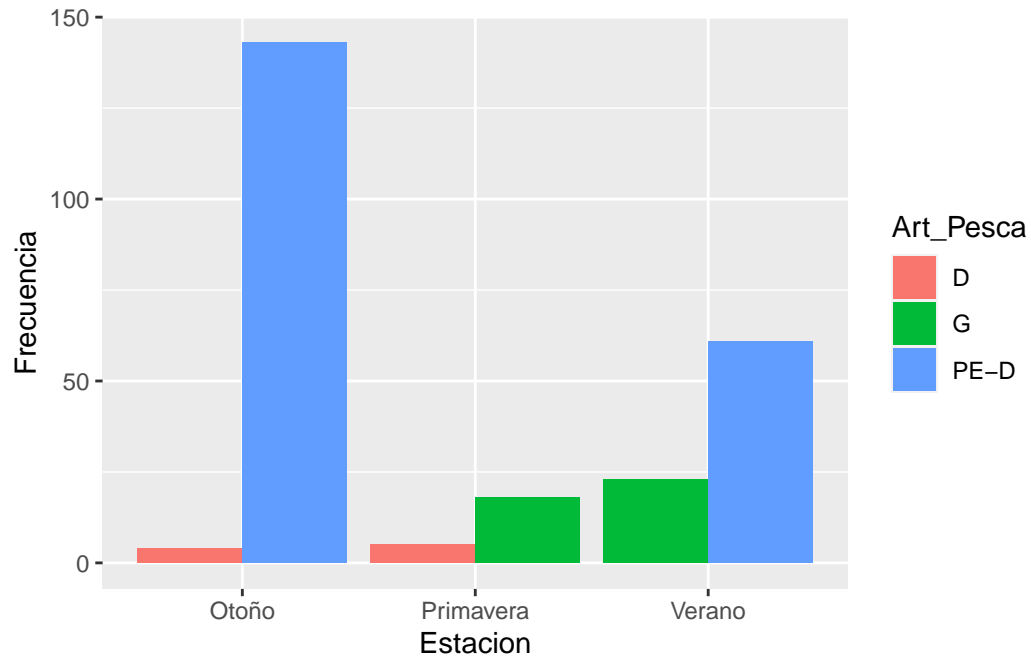
```
# Cantidad de individuos por Art_Pesca y sexo  
ggplot(data = BD_morfo_menor_inde, aes(x = Art_Pesca, fill = Sexo)) +  
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```



```
# Cantidad de individuos por Art_Pesca y Categoría
ggplot(data = BD_morfo, aes(x = Art_Pesca, fill = Categoría)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```

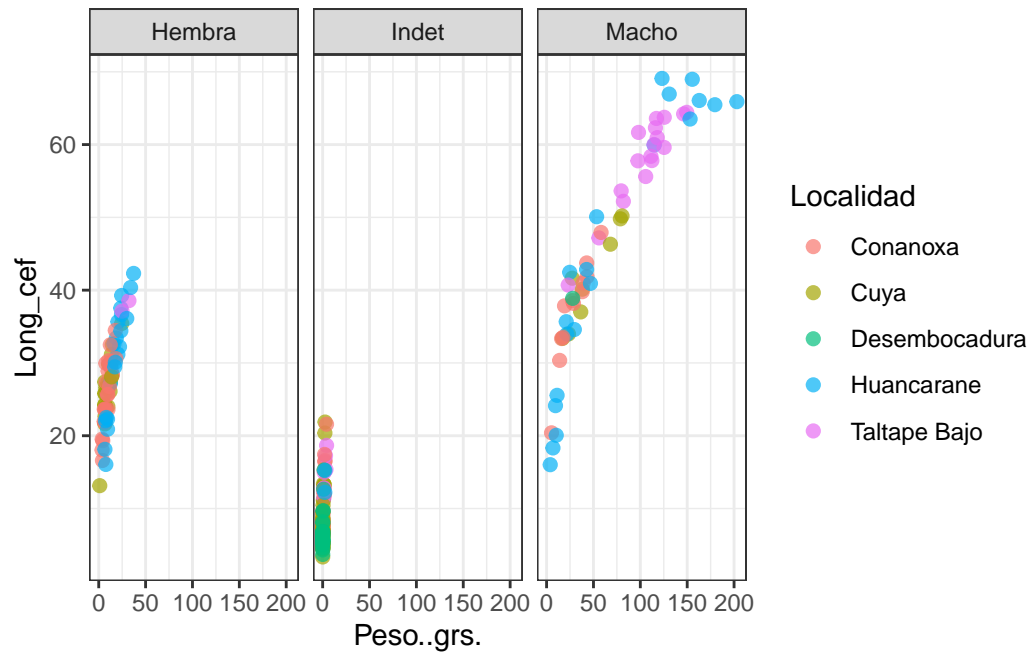


```
# Cantidad de individuos por estación y Art_Pesca
ggplot(data = BD_morfo, aes(x = Estacion, fill = Art_Pesca)) +
  geom_bar(position = "dodge")+labs(y= "Frecuencia")
```

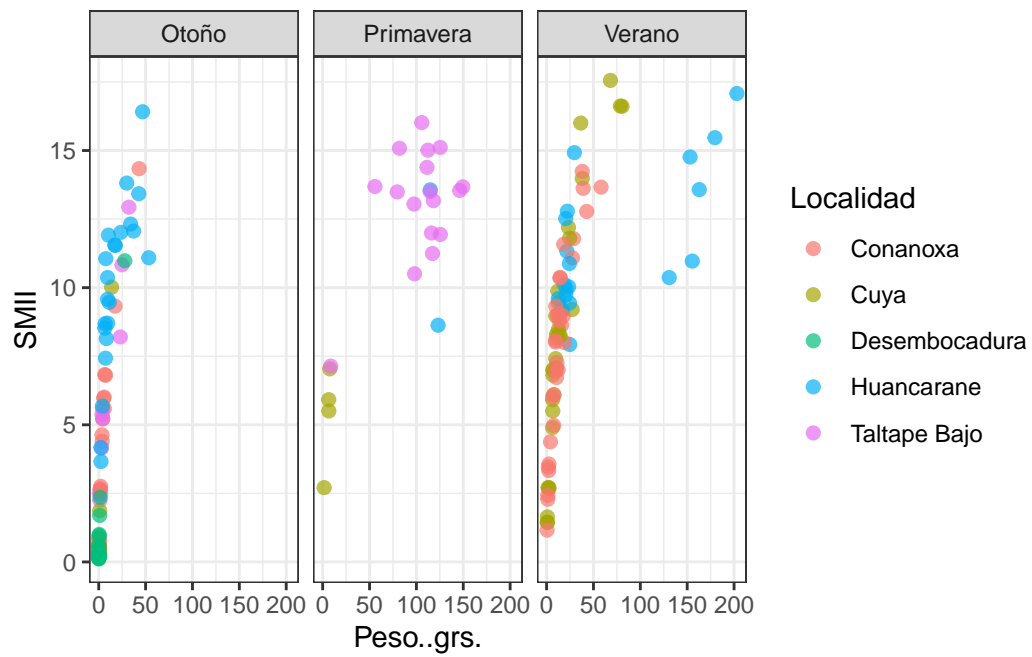


3. Gráficos triplot

```
camaron.triplots<-ggplot(BD_morfo, aes(x = Peso..grs.,
  y = Long_cef, colour = Localidad)) +
  geom_point(alpha = 0.7, size=2) +
  facet_grid(. ~ Sexo) +
  theme_bw() +
  ylab("Long_cef") +
  xlab("Peso..grs.")
camaron.triplots
```

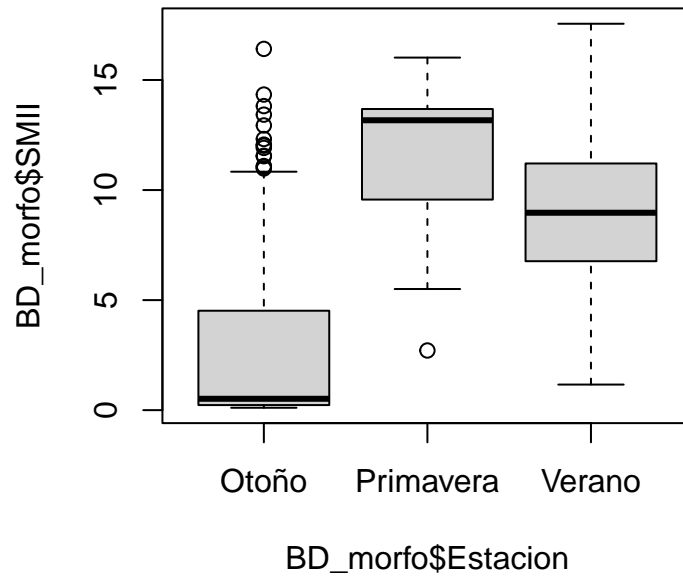


```
camaron2.triplots<-ggplot(BD_morfo, aes(x = Peso..grs.,
  y = SMII,
  colour = Localidad)) +
  geom_point(alpha = 0.7, size=2) +
  facet_grid(. ~ Estacion) +
  theme_bw() +
  ylab("SMII") +
  xlab("Peso..grs.")
camaron2.triplots
```

4. Boxplot y anovas

```
# SMII por estación  
boxplot(BD_morfo$SMII ~ BD_morfo$Estacion, data = BD_morfo)
```



```
# Hacemos un boxplot por sexo  
boxplot(BD_morfo$Peso..grs. ~ BD_morfo$Sexo, data = BD_morfo)
```



```
# Realizamos un ANOVA
anova_sexo<-aov(BD_morfo$Peso..grs.~BD_morfo$Sexo,data=BD_morfo)
summary(anova_sexo)
```

```
          Df Sum Sq Mean Sq F value Pr(>F)
BD_morfo$Sexo    2 192555    96278   147.5 <2e-16 ***
Residuals      251 163858     653
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Resumen del ANOVA

Como el p-valor es menos a 0.05 se rechaza hipótesis nula H0 y se acepta la alternativa H1

Lo que implica que hay diferencias significativas en la medias.

```
datos.juvenil=BD_morfo[BD_morfo$grupo=="Juvenil",] ## Filtramos por grupo
datos.adulto=BD_morfo[BD_morfo$grupo=="Adulto",]
```

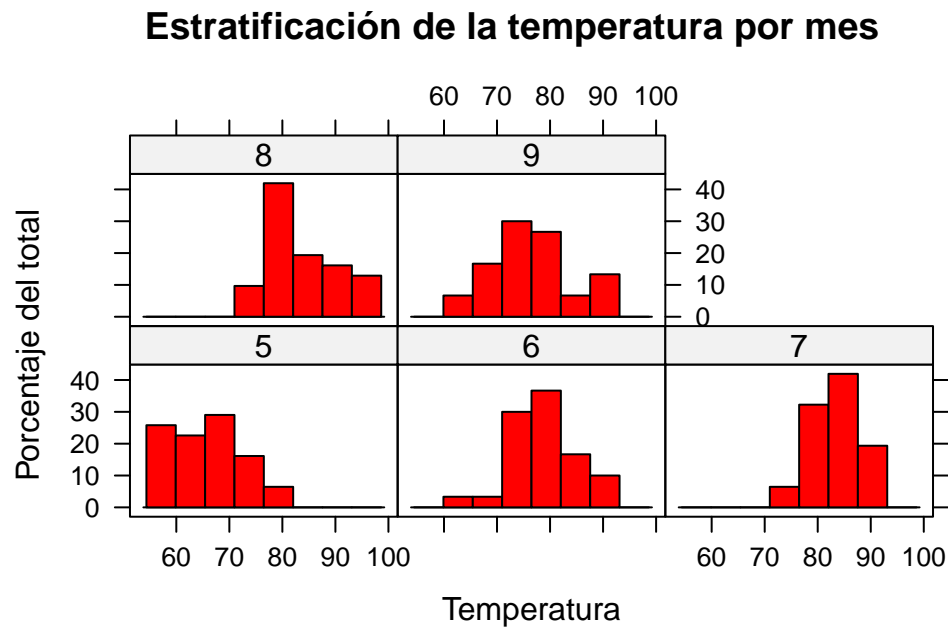
H0: Las medias son iguales entre juveniles y adultos

H1: Las media son distintas

5. Clase de talla

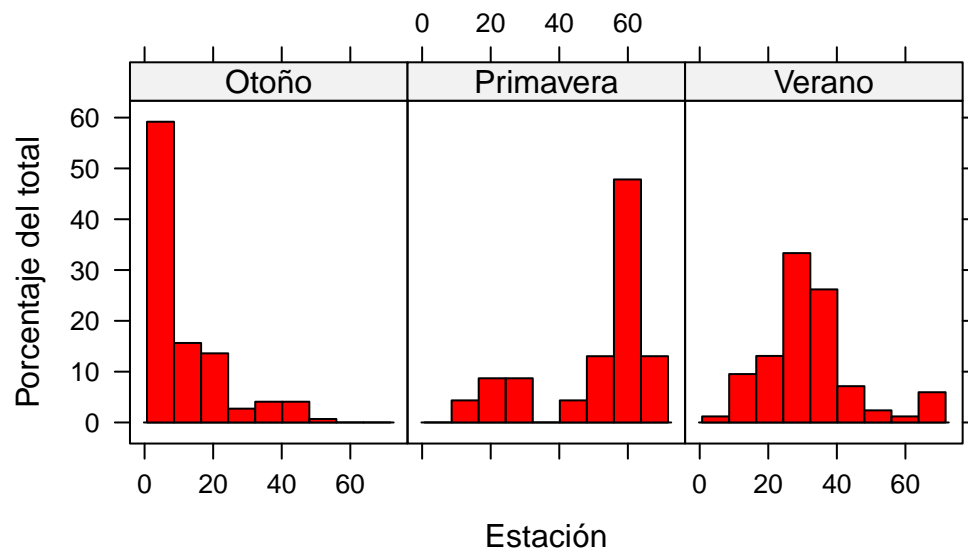
El objetivo de la clase de talla es determinar si hay diferencias en el tamaño del camarón en distintos estratos altitudinal. #

```
histogram(~Temp | as.character(Month), data=airquality,  
          main="Estratificación de la temperatura por mes",  
          xlab="Temperatura", ylab="Porcentaje del total", col="red")
```



```
histogram(~Long_cef | as.character(Estacion), data=BD_morfo,  
          main="Estratificación del Long_cef por Estación",  
          xlab="Estación", ylab="Porcentaje del total", col="red")
```

Estratificación del Long_cef por Estación



```
# Realizando un diagrama de puntos
dotplot(~Long_cef | as.character(Estacion), data=BD_morfo,
        main="Estratificación de la temperatura por mes",
        xlab="Temperatura", ylab="Porcentaje del total", col=1)
```

Estratificación de la temperatura por mes

