

QTL-SRS

Marcos Mancilla

16/10/2021

```
dat <- read_excel("dataset2.xls")
summary(dat)
```

```
##      Fecha                Centro          Unidad
## Min.      :2019-02-09 00:00:00 Length:12825 Length:12825
## 1st Qu.:2019-06-01 00:00:00 Class :character Class :character
## Median :2019-07-26 00:00:00 Mode  :character Mode  :character
## Mean    :2019-07-25 16:53:46
## 3rd Qu.:2019-09-20 00:00:00
## Max.    :2019-12-09 00:00:00
## Subgrupo      N_peces      Biomasa_kg      Mortalidad
## Length:12825 Min.      : 0 Min.      : 0 Min.      : 0.000
## Class :character 1st Qu.:16308 1st Qu.: 23766 1st Qu.: 0.000
## Mode  :character Median :28340 Median : 42046 Median : 1.000
## Mean    :24143 Mean    : 44360 Mean    : 9.889
## 3rd Qu.:29028 3rd Qu.: 59946 3rd Qu.: 8.000
## Max.    :29209 Max.    :114599 Max.    :1982.000
## Mortalidad_Biomasa_kg N_cosecha Biomasa_cosecha Alimento_kg
## Min.      : 0.000 Min.      : 0.0 Min.      : 0.0 Min.      : 0.0
## 1st Qu.: 0.000 1st Qu.: 0.0 1st Qu.: 0.0 1st Qu.: 285.0
## Median : 2.015 Median : 0.0 Median : 0.0 Median : 446.0
## Mean    : 27.549 Mean    : 76.9 Mean    : 270.1 Mean    : 440.6
## 3rd Qu.: 17.845 3rd Qu.: 0.0 3rd Qu.: 0.0 3rd Qu.: 604.0
## Max.    :7671.965 Max.    :22602.0 Max.    :78009.3 Max.    :1560.0
## Temperatura
## Min.      : 0.00
## 1st Qu.:10.79
## Median :11.20
## Mean    :11.15
## 3rd Qu.:11.70
## Max.    :15.00
```

```
head(dat)
```

```
## # A tibble: 6 x 12
##   Fecha                Centro Unidad Subgrupo N_peces Biomasa_kg Mortalidad
##   <dtm>                <chr> <chr> <chr>      <dbl>      <dbl>      <dbl>
## 1 2019-08-24 00:00:00 A      205 QTL2      16302      35625.      0
## 2 2019-08-25 00:00:00 A      205 QTL2      16302      35879.      0
## 3 2019-08-26 00:00:00 A      205 QTL2      16302      36261.      0
## 4 2019-08-27 00:00:00 A      205 QTL2      16302      36684.      0
## 5 2019-08-28 00:00:00 A      205 QTL2      16302      37100.      0
## 6 2019-08-29 00:00:00 A      205 QTL2      16302      37435.      0
```

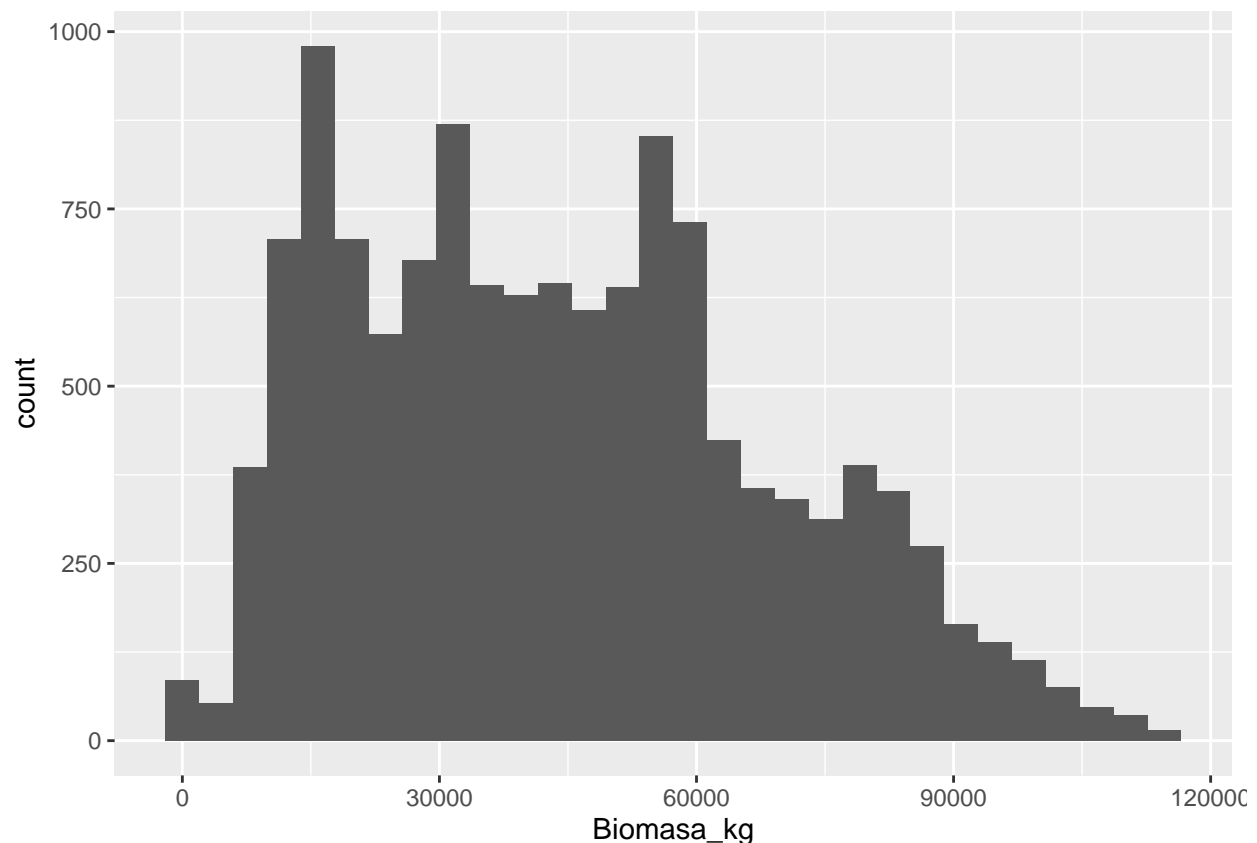
```
## # ... with 5 more variables: Mortalidad_Biomasa_kg <dbl>, N_cosecha <dbl>,
## #   Biomasa_cosecha <dbl>, Alimento_kg <dbl>, Temperatura <dbl>
```

```
str(dat)
```

```
## tibble [12,825 x 12] (S3: tbl_df/tbl/data.frame)
##  $ Fecha           : POSIXct[1:12825], format: "2019-08-24" "2019-08-25" ...
##  $ Centro          : chr [1:12825] "A" "A" "A" "A" ...
##  $ Unidad          : chr [1:12825] "205" "205" "205" "205" ...
##  $ Subgrupo        : chr [1:12825] "QTL2" "QTL2" "QTL2" "QTL2" ...
##  $ N_peces         : num [1:12825] 16302 16302 16302 16302 16302 ...
##  $ Biomasa_kg       : num [1:12825] 35625 35879 36261 36684 37100 ...
##  $ Mortalidad       : num [1:12825] 0 0 0 0 0 0 0 1 2 2 ...
##  $ Mortalidad_Biomasa_kg: num [1:12825] 0 0 0 0 0 ...
##  $ N_cosecha        : num [1:12825] 0 0 0 0 0 0 0 0 0 0 ...
##  $ Biomasa_cosecha   : num [1:12825] 0 0 0 0 0 0 0 0 0 0 ...
##  $ Alimento_kg      : num [1:12825] 455 300 450 500 490 396 472 562 435 556 ...
##  $ Temperatura      : num [1:12825] 11.3 11.1 11.1 11.3 10.3 ...
```

```
ggplot(dat, aes(x = Biomasa_kg)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

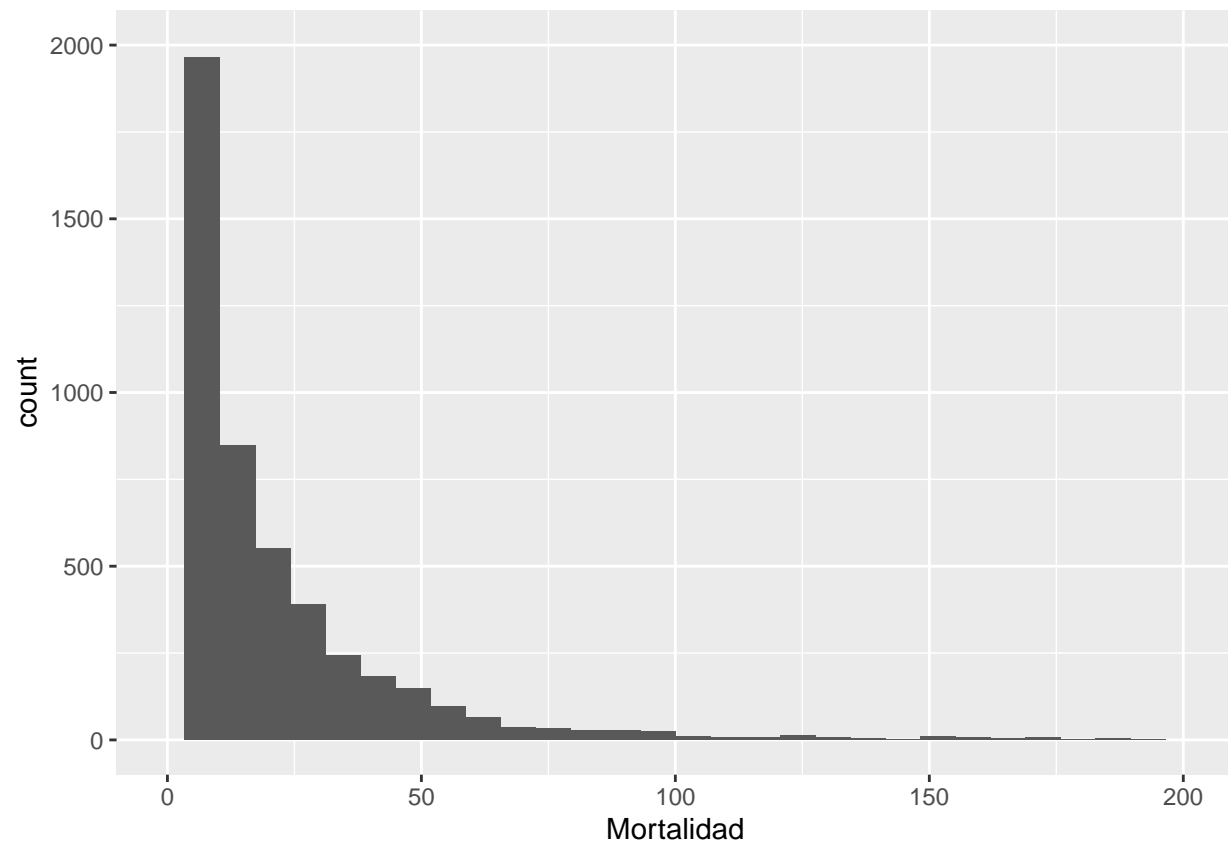


```
ggplot(dat, aes(x = Mortalidad)) +
  geom_histogram()+ scale_x_continuous(limit = c(0,200))+ scale_y_continuous(limit = c(0,2000))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 34 rows containing non-finite values (stat_bin).
```

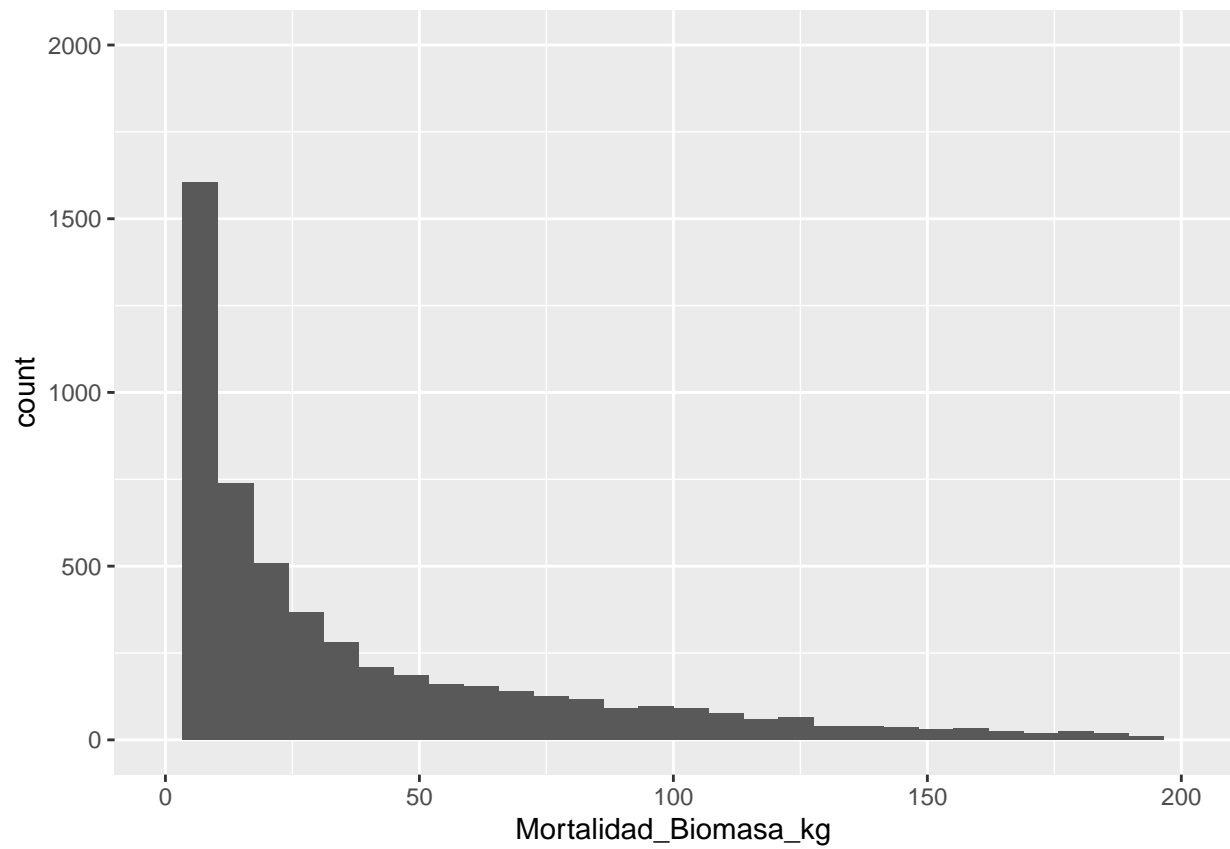
```
## Warning: Removed 2 rows containing missing values (geom_bar).
```



```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

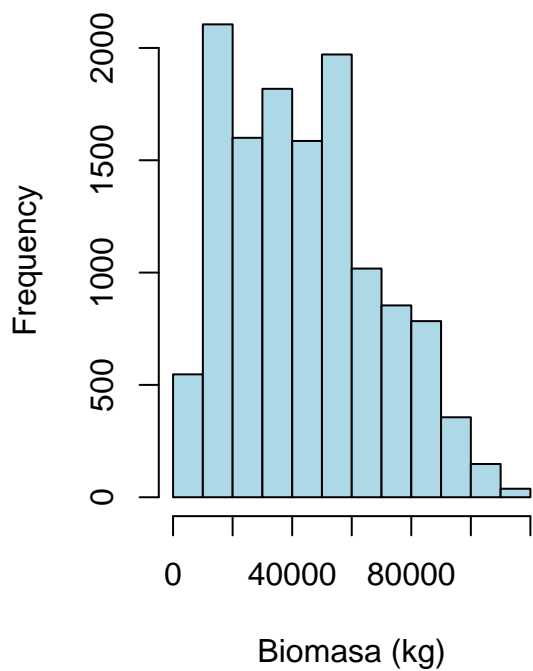
```
## Warning: Removed 254 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```

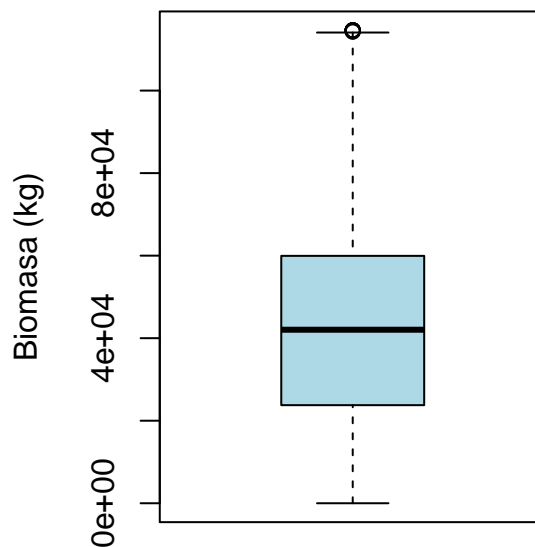


```
par(mfrow=c(1,2))
hist(dat$Biomasa_kg, col="light blue", main = "Histograma Biomasa", xlab = "Biomasa (kg)")
boxplot(dat$Biomasa_kg, col="light blue", main = "Boxplot Biomasa", ylab = "Biomasa (kg)")
```

Histograma Biomasa

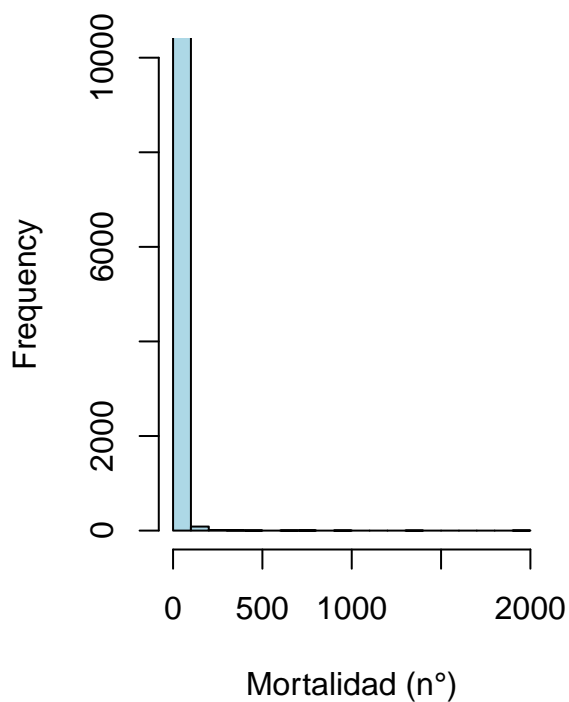


Boxplot Biomasa

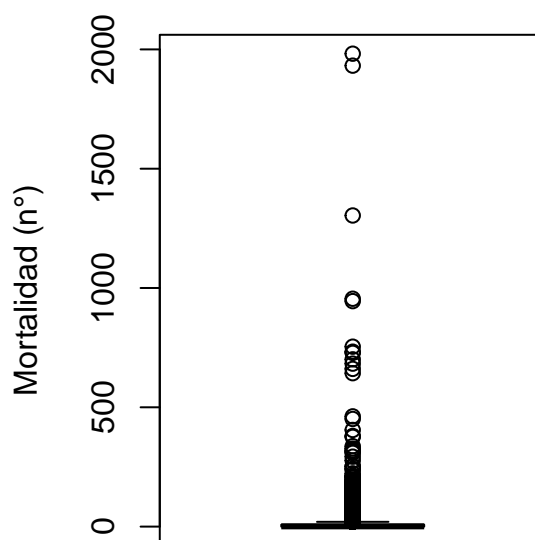


```
par(mfrow=c(1,2))
hist(dat$Mortalidad, col="light blue", main = "Histograma Mortalidad", xlab = "Mortalidad (n°)",
     xlim = c(0, 2000),
     ylim = c(0, 10000))
boxplot(dat$Mortalidad, col="light blue", main = "Boxplot Mortalidad", ylab = "Mortalidad (n°)")
```

Histograma Mortalidad



Boxplot Mortalidad



```

table(dat$Subgrupo)

##
## QTL1 QTL2 SQTL
## 3256 9078 491

table(dat$Centro)

##
## A B C D
## 4130 3222 2217 3256

dat$Centro <- as.factor(dat$Centro)
dat$Subgrupo <- as.factor(dat$Subgrupo)
str(dat)

## tibble [12,825 x 12] (S3: tbl_df/tbl/data.frame)
## $ Fecha : POSIXct[1:12825], format: "2019-08-24" "2019-08-25" ...
## $ Centro : Factor w/ 4 levels "A","B","C","D": 1 1 1 1 1 1 1 1 1 1 ...
## $ Unidad : chr [1:12825] "205" "205" "205" "205" ...
## $ Subgrupo : Factor w/ 3 levels "QTL1","QTL2",...: 2 2 2 2 2 2 2 2 2 2 ...
## $ N_peces : num [1:12825] 16302 16302 16302 16302 16302 ...
## $ Biomasa_kg : num [1:12825] 35625 35879 36261 36684 37100 ...
## $ Mortalidad : num [1:12825] 0 0 0 0 0 0 0 1 2 2 ...
## $ Mortalidad_Biomasa_kg: num [1:12825] 0 0 0 0 0 ...
## $ N_cosecha : num [1:12825] 0 0 0 0 0 0 0 0 0 0 ...
## $ Biomasa_cosecha : num [1:12825] 0 0 0 0 0 0 0 0 0 0 ...
## $ Alimento_kg : num [1:12825] 455 300 450 500 490 396 472 562 435 556 ...
## $ Temperatura : num [1:12825] 11.3 11.1 11.1 11.3 10.3 ...

summary(dat)

## Fecha Centro Unidad Subgrupo
## Min. :2019-02-09 00:00:00 A:4130 Length:12825 QTL1:3256
## 1st Qu.:2019-06-01 00:00:00 B:3222 Class :character QTL2:9078
## Median :2019-07-26 00:00:00 C:2217 Mode :character SQTL: 491
## Mean :2019-07-25 16:53:46 D:3256
## 3rd Qu.:2019-09-20 00:00:00
## Max. :2019-12-09 00:00:00

## N_peces Biomasa_kg Mortalidad Mortalidad_Biomasa_kg
## Min. : 0 Min. : 0 Min. : 0.000 Min. : 0.000
## 1st Qu.:16308 1st Qu.: 23766 1st Qu.: 0.000 1st Qu.: 0.000
## Median :28340 Median : 42046 Median : 1.000 Median : 2.015
## Mean :24143 Mean : 44360 Mean : 9.889 Mean : 27.549
## 3rd Qu.:29028 3rd Qu.: 59946 3rd Qu.: 8.000 3rd Qu.: 17.845
## Max. :29209 Max. :114599 Max. :1982.000 Max. :7671.965

## N_cosecha Biomasa_cosecha Alimento_kg Temperatura
## Min. : 0.0 Min. : 0.0 Min. : 0.0 Min. : 0.00
## 1st Qu.: 0.0 1st Qu.: 0.0 1st Qu.: 285.0 1st Qu.:10.79
## Median : 0.0 Median : 0.0 Median : 446.0 Median :11.20
## Mean : 76.9 Mean : 270.1 Mean : 440.6 Mean :11.15
## 3rd Qu.: 0.0 3rd Qu.: 0.0 3rd Qu.: 604.0 3rd Qu.:11.70
## Max. :22602.0 Max. :78009.3 Max. :1560.0 Max. :15.00

table(dat$Subgrupo,dat$Centro)

##

```

```
##           A      B      C      D
##   QTL1      0      0      0 3256
##   QTL2 4130 3222 1726      0
##   SQTL      0      0 491      0

tabla1 <- with(dat, table(Subgrupo, Centro))
prop.table(tabla1)

##           Centro
## Subgrupo      A      B      C      D
##   QTL1 0.0000000 0.0000000 0.0000000 0.2538791
##   QTL2 0.3220273 0.2512281 0.1345809 0.0000000
##   SQTL 0.0000000 0.0000000 0.0382846 0.0000000

tabla1_prop <- with(dat, prop.table(tabla1))
knitr::kable(tabla1_prop, caption = "Proporción datos por Subgrupo de peces en diferentes Centros")
```

Table 1: Proporción datos por Subgrupo de peces en diferentes Centros

	A	B	C	D
QTL1	0.0000000	0.0000000	0.0000000	0.2538791
QTL2	0.3220273	0.2512281	0.1345809	0.0000000
SQTL	0.0000000	0.0000000	0.0382846	0.0000000

Los datos no están balanceados. Hay más observaciones para el Subgrupo QTL2 (70,7%) que para otros subgrupos. El Subgrupo en menor proporción es el SQTL que está presente solo en Centro C.

```
df1 <- dat %>%
  select(Subgrupo, Centro, Unidad, Biomasa_kg, N_peces) %>%
  group_by(Subgrupo, Centro, Unidad) %>%
  summarize(Maximo = max(N_peces), Minimo = min(N_peces))
```

```
## `summarise()` has grouped output by 'Subgrupo', 'Centro'. You can override using the `.groups` argument
knitr::kable(df1, caption = "N° de peces por unidad (jaula)")
```

Table 2: N° de peces por unidad (jaula)

Subgrupo	Centro	Unidad	Maximo	Minimo
QTL1	D	101	29132	0
QTL1	D	102	29093	0
QTL1	D	103	29209	0
QTL1	D	104	29146	0
QTL1	D	105	29209	0
QTL1	D	106	29116	0
QTL1	D	107	29125	0
QTL1	D	108	29144	0
QTL1	D	109	29179	0
QTL1	D	110	29176	0
QTL1	D	111	29209	0
QTL1	D	112	29160	0
QTL1	D	113	29209	0
QTL1	D	114	29048	0
QTL2	A	101	16259	0

Subgrupo	Centro	Unidad	Maximo	Minimo
QTL2	A	102	16393	0
QTL2	A	103	16327	0
QTL2	A	104	16430	0
QTL2	A	105	16357	0
QTL2	A	106	16430	0
QTL2	A	107	16363	0
QTL2	A	108	16430	0
QTL2	A	109	16340	0
QTL2	A	110	16379	0
QTL2	A	201	16402	0
QTL2	A	202	16416	0
QTL2	A	203	16393	0
QTL2	A	204	16413	0
QTL2	A	205	16416	0
QTL2	A	206	16394	0
QTL2	A	207	16430	0
QTL2	A	208	16418	0
QTL2	A	209	16430	0
QTL2	A	210	16409	0
QTL2	B	101	29164	0
QTL2	B	102	28956	26947
QTL2	B	103	29209	0
QTL2	B	104	29085	0
QTL2	B	105	29209	26606
QTL2	B	106	29125	27077
QTL2	B	107	29189	26952
QTL2	B	108	29118	0
QTL2	B	109	29196	0
QTL2	B	110	29209	27633
QTL2	B	111	29186	0
QTL2	B	112	29174	0
QTL2	B	113	29196	27896
QTL2	B	114	29209	27941
QTL2	C	101	29171	0
QTL2	C	102	29188	0
QTL2	C	103	29186	0
QTL2	C	104	29172	0
QTL2	C	105	4827	0
QTL2	C	106	29130	0
QTL2	C	108	29124	0
QTL2	C	109	29137	25786
QTL2	C	110	28900	0
SQTL	C	105	29184	2362
SQTL	C	107	29185	0
SQTL	C	109	16956	0

El análisis anterior indica que la mayoría de las unidades (jaulas) se cosecharon al final del periodo de observación (Mínimo = 0), quedando solo 9 jaulas con peces.

```
df2 <-df1 %>%
  select(Subgrupo, Maximo)%>%
  group_by(Subgrupo)%>%
```



```

summarize(Peces=sum(Maximo))

## Adding missing grouping variables: `Centro`
knitr::kable(df2,caption = "N° de peces por Subgrupo")

```

Table 3: N° de peces por Subgrupo

Subgrupo	Peces
QTL1	408155
QTL2	973889
SQTL	75325