Proyecto Fin de Curso

Karla Castro M

20/7/2021

Título del proyecto

Comparacion del crecimiento de trucha arcoiris (*Oncorhynchus mykiss*) con tres tipos de dietas diferentes.

Autor.

Karla Castro Monje

Planteamiento del problema

El beneficio economico de la acuicultura intensiva y semi-intensiva se encunetra intimamente relacionado con el suministro y el costo de alimento, debido a que los cultivos intensivos de la trucha "Arco Iris" requieren alimentos con niveles elevados de proteinas y el costo de la fuente proteica es el determina las utilidades de produccion.

Las fuentes proteicas por su alto costo son las que determinan el precio del alimento balanceado por lo cual se deben considerar fuentes alternativas de alimentacion de bajo costos reñlativos y disponibilidad.

1 Correlacion entre peso vs longitud

```
Datos_1 <- read_excel("Datos.xlsx", sheet = "Hoja 6")
summary(Datos_1)
```

```
##    Estanque             Pesof           Longf
##  Length:752         Min.   :196.0   Min.   :22.50
##  Class :character   1st Qu.:252.4   1st Qu.:26.70
##  Mode  :character   Median :285.5   Median :28.00
##                     Mean   :292.0   Mean   :28.00
##                     3rd Qu.:320.2   3rd Qu.:29.32
##                     Max.   :496.5   Max.   :35.00
```

```
Datos_1$Estanque <- as.factor(Datos_1$Estanque)
summary(Datos_1)
```

```
##     Estanque       Pesof           Longf
##  E15    : 54   Min.   :196.0   Min.   :22.50
##  E9     : 53   1st Qu.:252.4   1st Qu.:26.70
##  E10    : 52   Median :285.5   Median :28.00
```

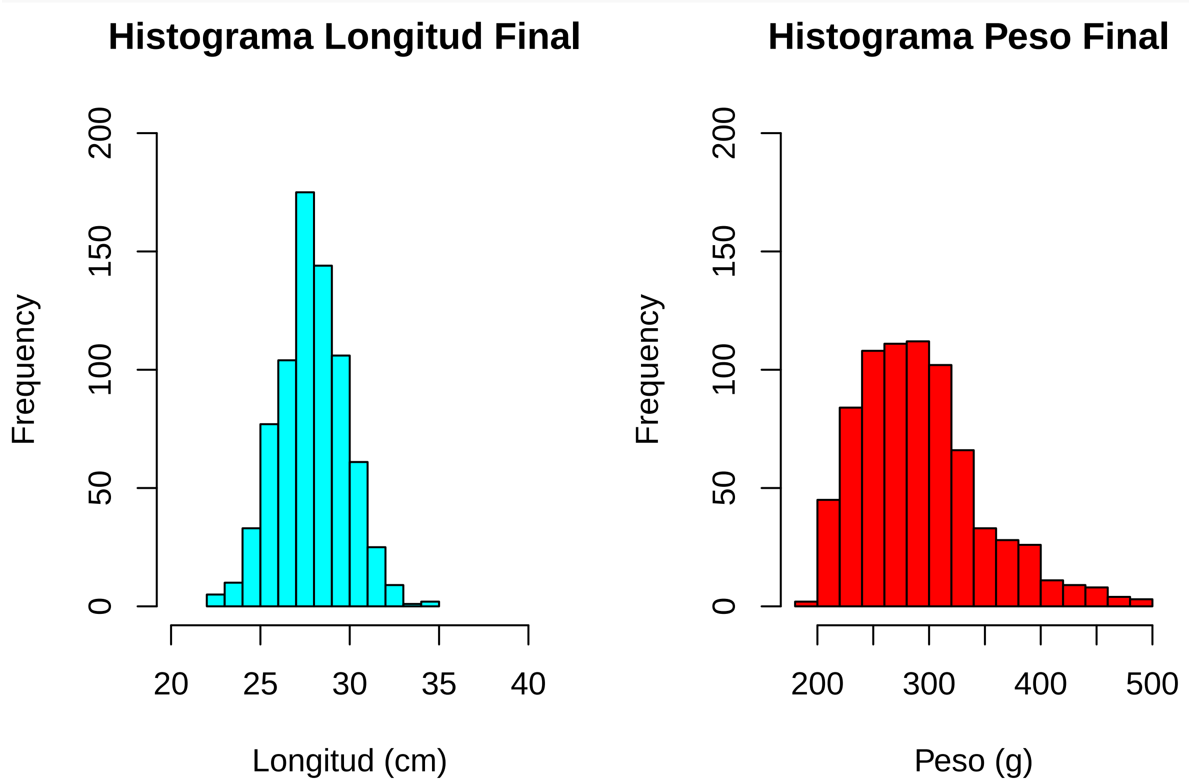
Table 1: Media y Varianza Tais

Media_Longf	Var_Longf	Media_Pesof	Var_Pesof
28.00213	3.728331	292.0395	3064.383

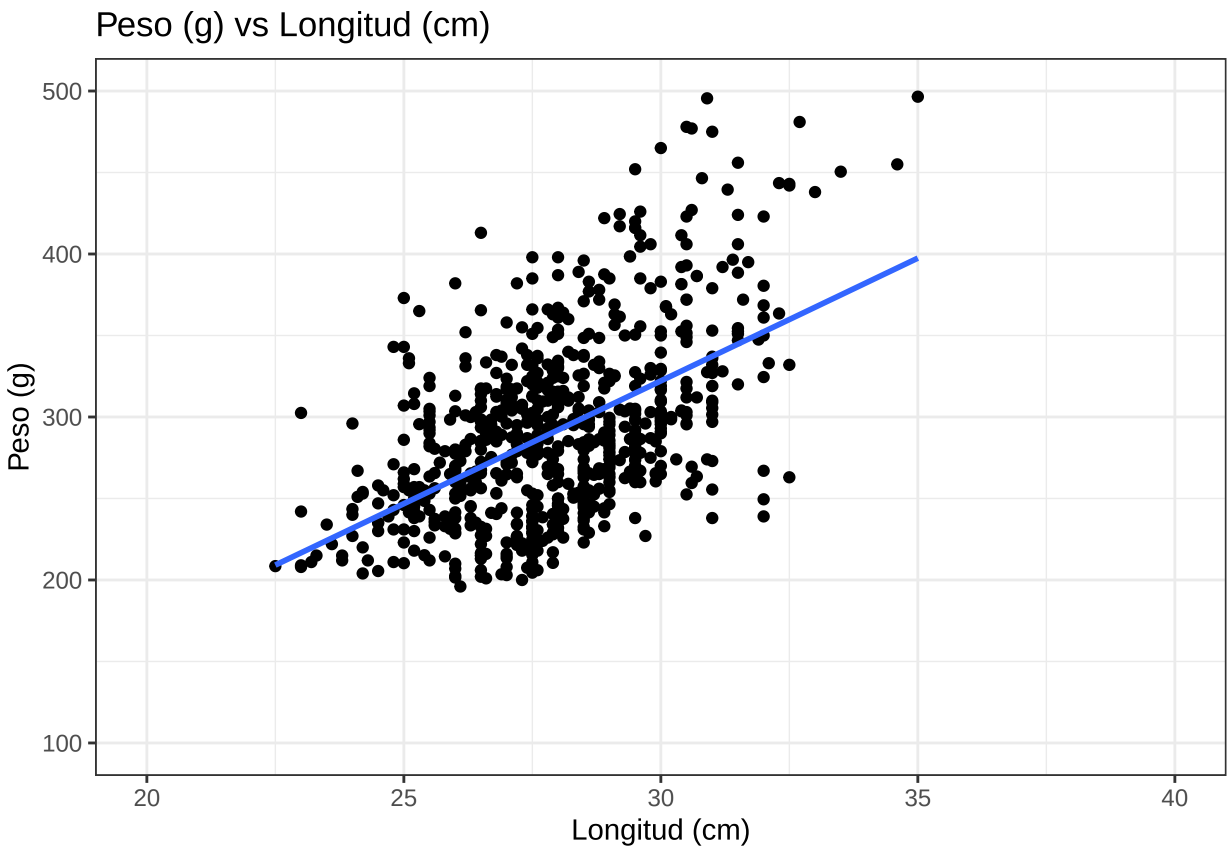
```
## E11    : 52   Mean   :292.0   Mean   :28.00
## E14    : 52   3rd Qu.:320.2   3rd Qu.:29.32
## E4     : 51   Max.   :496.5   Max.   :35.00
## (Other):438
```

```
Tabla_Lf_Pf <- Datos_1 %>%
  summarize(Media_Longf = mean(Longf),
          Var_Longf  = var(Longf),
          Media_Pesof = mean(Pesof),
          Var_Pesof = var(Pesof))
knitr::kable(Tabla_Lf_Pf, caption = "Media y Varianza Tais")
```

```
par(mfrow = c(1, 2))
hist(Datos_1$Longf, main = "Histograma Longitud Final ", xlab = "Longitud (cm)", xlim= c(20,40), ylim =
hist(Datos_1$Pesof, main = "Histograma Peso Final ", xlab = "Peso (g)", ylim =c(0,200), col = "red")
```




```
ggplot(Datos_1, aes(x = Datos_1$Longf, y = Datos_1$Pesof)) +
  geom_point() +
  ggtitle("Peso (g) vs Longitud (cm)") + geom_smooth(method = lm, se = FALSE) +
  scale_x_continuous(name = "Longitud (cm)", limits = c(20, 40), breaks = seq(20, 40, 5)) +
  scale_y_continuous(name = "Peso (g)", limits = c(100, 500), breaks = seq(100, 500, 100)) +
  theme(plot.title = element_text(hjust = 0.3), panel.background = element_blank(), axis.line = element_
  theme_bw()
```




```
cor(Datos_1$Longf, Datos_1$Pesof)
```

```
## [1] 0.5255099
```

```
cor_Lf_Pf <- cor.test(x= Datos_1$Longf, y = Datos_1$Pesof, method = "pearson", conf.level = 0.95)
cor_Lf_Pf
```

```
##
##  Pearson's product-moment correlation
##
## data:  Datos_1$Longf and Datos_1$Pesof
## t = 16.916, df = 750, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4717401 0.5753857
## sample estimates:
##       cor
## 0.5255099
```

```
IntC <- cor_Lf_Pf$conf.int
IntC
```

```
## [1] 0.4717401 0.5753857
## attr(,"conf.level")
## [1] 0.95
```

```
pander::pander(cor_Lf_Pf, caption = "Prueba de hipotesis para el coeficiente de relacion Longitud y Peso
```

Table 2: Prueba de hipotesis para el coeficiente de relacion Longitud y Peso

Test statistic	df	P value	Alternative hypothesis	cor
16.92	750	1.282e-54 * * *	two.sided	0.5255

2 Dieta, Dosis Alimento y peso final

```
Datos_2 <- read_excel("Datos.xlsx", sheet = "Hoja3")
summary(Datos_2)
```

```
##    Estanques        Dieta                 Pi              Pf
##  Min.   : 1.0   Length:15          Min.   :148.0   Min.   :244.9
##  1st Qu.: 4.5   Class :character   1st Qu.:149.8   1st Qu.:275.6
##  Median : 8.0   Mode  :character   Median :151.0   Median :288.1
##  Mean   : 8.0                      Mean   :152.5   Mean   :292.7
##  3rd Qu.:11.5                      3rd Qu.:155.9   3rd Qu.:304.5
##  Max.   :15.0                      Max.   :160.4   Max.   :358.3
##        Li              Lf              DA              Ki
##  Min.   :22.51   Min.   :26.18   Min.   :153.6   Min.   :0.01151
##  1st Qu.:22.73   1st Qu.:27.26   1st Qu.:165.1   1st Qu.:0.01207
##  Median :23.18   Median :28.37   Median :172.8   Median :0.01236
##  Mean   :23.09   Mean   :28.68   Mean   :169.5   Mean   :0.01240
##  3rd Qu.:23.38   3rd Qu.:29.16   3rd Qu.:174.3   3rd Qu.:0.01278
##  Max.   :23.50   Max.   :33.87   Max.   :181.2   Max.   :0.01374
##        Kf                FCA              CUT                 SGR
##  Min.   :0.006951   Min.   :0.8674   Min.   :0.002130   Min.   :0.4871
##  1st Qu.:0.010776   1st Qu.:1.1513   1st Qu.:0.003446   1st Qu.:0.6657
##  Median :0.011618   Median :1.2944   Median :0.003916   Median :0.6991
##  Mean   :0.012516   Mean   :1.3291   Mean   :0.004312   Mean   :0.7030
##  3rd Qu.:0.014980   3rd Qu.:1.6248   3rd Qu.:0.004734   3rd Qu.:0.7591
##  Max.   :0.015908   Max.   :1.7049   Max.   :0.008158   Max.   :0.9229
```

```
str(Datos_2)
```

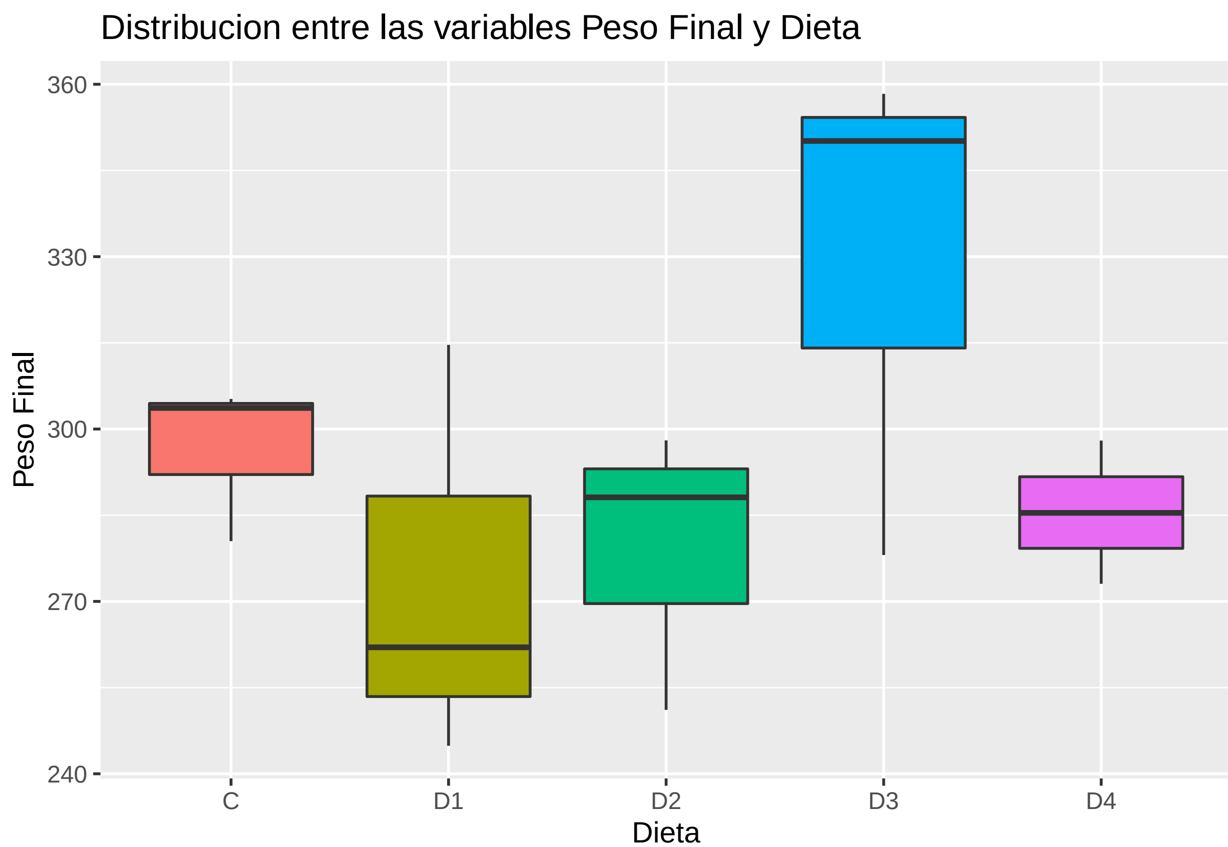
```
## tibble [15 x 12] (S3: tbl_df/tbl/data.frame)
##  $ Estanques: num [1:15] 1 2 3 4 5 6 7 8 9 10 ...
##  $ Dieta    : chr [1:15] "D1" "D4" "D3" "C" ...
##  $ Pi       : num [1:15] 148 152 150 151 156 ...
##  $ Pf       : num [1:15] 262 285 350 305 298 ...
##  $ Li       : num [1:15] 23.4 23.1 22.7 22.7 23.3 ...
##  $ Lf       : num [1:15] 26.2 26.3 28.4 27.6 26.8 ...
##  $ DA       : num [1:15] 181 177 174 173 175 ...
##  $ Ki       : num [1:15] 0.0115 0.0123 0.0129 0.0129 0.0124 ...
##  $ Kf       : num [1:15] 0.0146 0.0157 0.0147 0.0146 0.0155 ...
##  $ FCA      : num [1:15] 1.589 1.323 0.948 1.118 1.228 ...
##  $ CUT      : num [1:15] 0.00213 0.00249 0.00434 0.00376 0.00272 ...
##  $ SGR      : num [1:15] 0.621 0.686 0.923 0.767 0.704 ...
```

```
Datos_2$Dieta <- as.factor(Datos_2$Dieta)
str(Datos_2)
```

```
## tibble [15 x 12] (S3: tbl_df/tbl/data.frame)
##  $ Estanques: num [1:15] 1 2 3 4 5 6 7 8 9 10 ...
```

```
##  $ Dieta    : Factor w/ 5 levels "C","D1","D2",..: 2 5 4 1 5 3 1 4 3 2 ...
##  $ Pi       : num [1:15] 148 152 150 151 156 ...
##  $ Pf       : num [1:15] 262 285 350 305 298 ...
##  $ Li       : num [1:15] 23.4 23.1 22.7 22.7 23.3 ...
##  $ Lf       : num [1:15] 26.2 26.3 28.4 27.6 26.8 ...
##  $ DA       : num [1:15] 181 177 174 173 175 ...
##  $ Ki       : num [1:15] 0.0115 0.0123 0.0129 0.0129 0.0124 ...
##  $ Kf       : num [1:15] 0.0146 0.0157 0.0147 0.0146 0.0155 ...
##  $ FCA      : num [1:15] 1.589 1.323 0.948 1.118 1.228 ...
##  $ CUT      : num [1:15] 0.00213 0.00249 0.00434 0.00376 0.00272 ...
##  $ SGR      : num [1:15] 0.621 0.686 0.923 0.767 0.704 ...
```

```
ggplot(data = Datos_2, aes(x = Dieta, y = Pf, fill= Dieta)) +
  geom_boxplot() +
  guides(fill = FALSE, color = FALSE) +
  labs(x = "Dieta", y = "Peso Final", title = "Distribucion entre las variables Peso Final y Dieta")
```



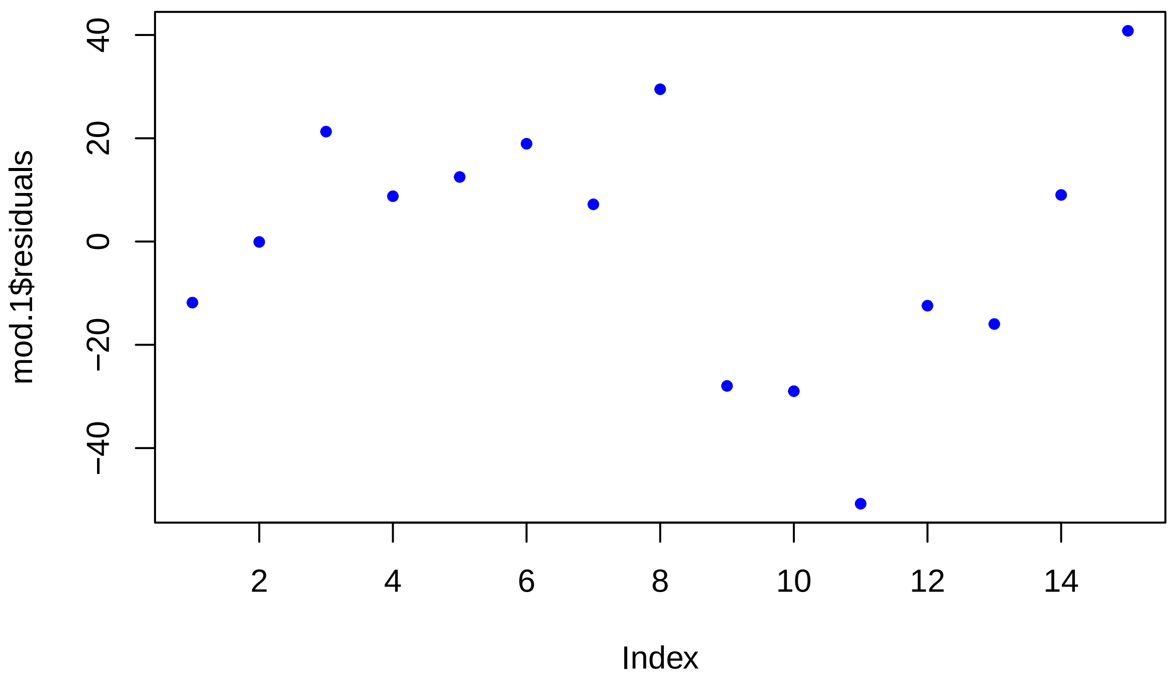

```
mod.1 <- lm(Pf ~ Dieta, data = Datos_2)
summary(mod.1)
```

```
##
## Call:
## lm(formula = Pf ~ Dieta, data = Datos_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -50.770 -14.200   7.203  15.721  40.807
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   296.47      16.79  17.652 7.25e-09 ***
## DietaD1       -22.62      23.75  -0.952    0.363    
## DietaD2       -17.38      23.75  -0.732    0.481    
## DietaD3        32.38      23.75   1.363    0.203    
## DietaD4       -10.98      23.75  -0.462    0.654    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 29.09 on 10 degrees of freedom
## Multiple R-squared:  0.4042, Adjusted R-squared:  0.1659 
## F-statistic: 1.696 on 4 and 10 DF,  p-value: 0.2268
```

Independencia

```r
plot(mod.1$residuals, pch = 20, col = "blue")
```

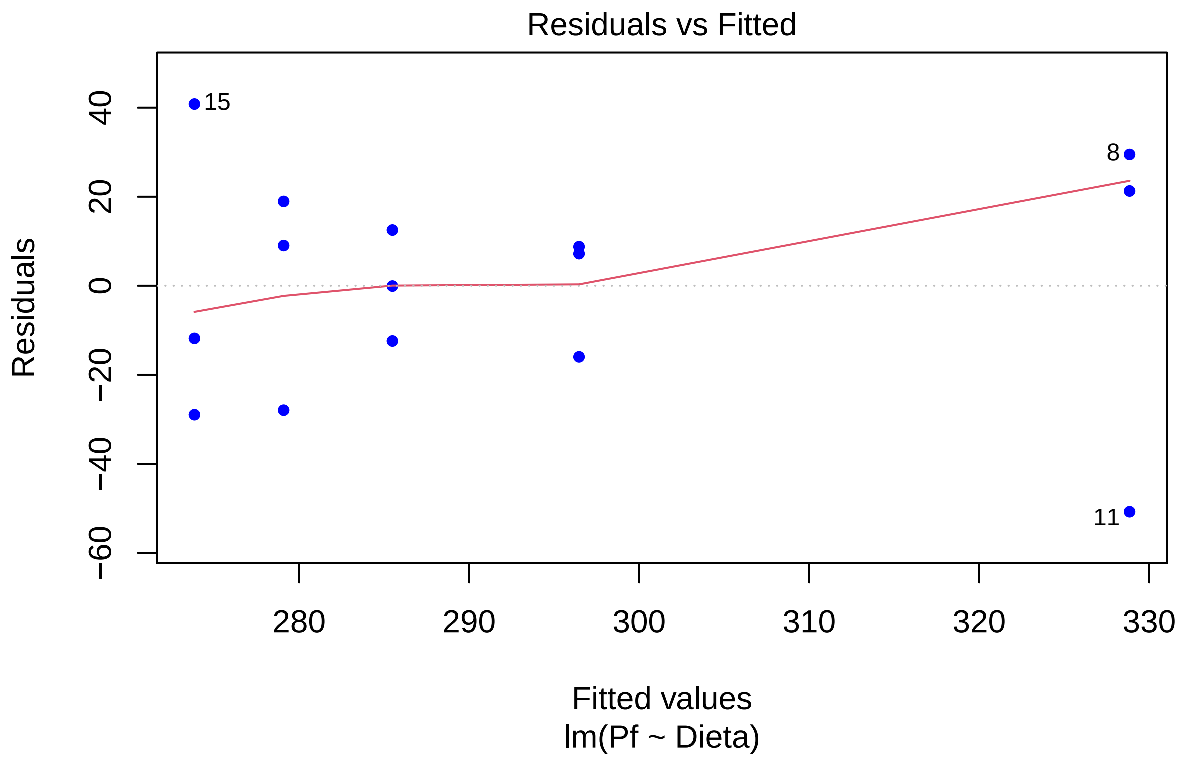



```r
dwtest(Pf ~ Dieta, data = Datos_2, alternative = c("two.sided"), iterations = 15)
```

```
## 
##  Durbin-Watson test
## 
## data:  Pf ~ Dieta
## DW = 0.98509, p-value = 0.07205
## alternative hypothesis: true autocorrelation is not 0
```

Homogeneidad de varianzas

```r
plot(mod.1, 1, pch = 20, col = "blue")
```

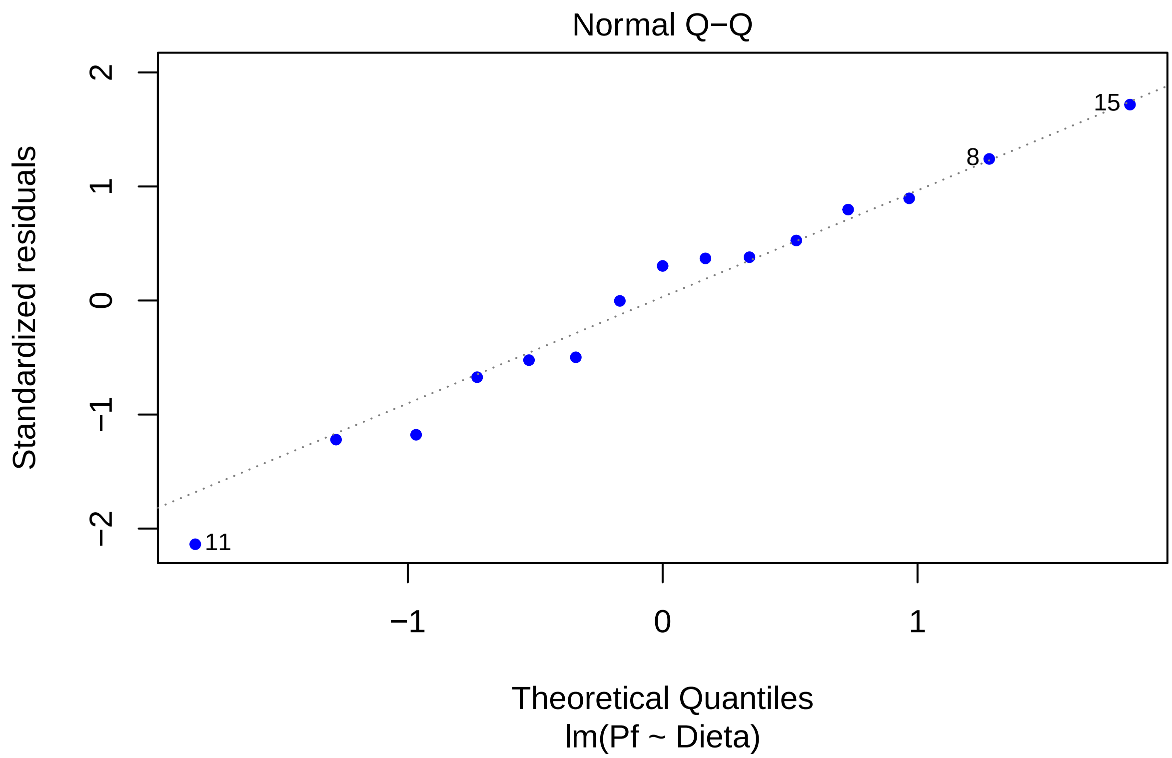



```
leveneTest(Pf ~ Dieta, data = Datos_2, center ="median")
```

```
## Levene's Test for Homogeneity of Variance (center = "median")
##       Df F value Pr(>F)
## group  4  0.3756  0.821
##       10
```

Normalidad

```
plot(mod.1, 2, pch = 20, col = "blue")
```



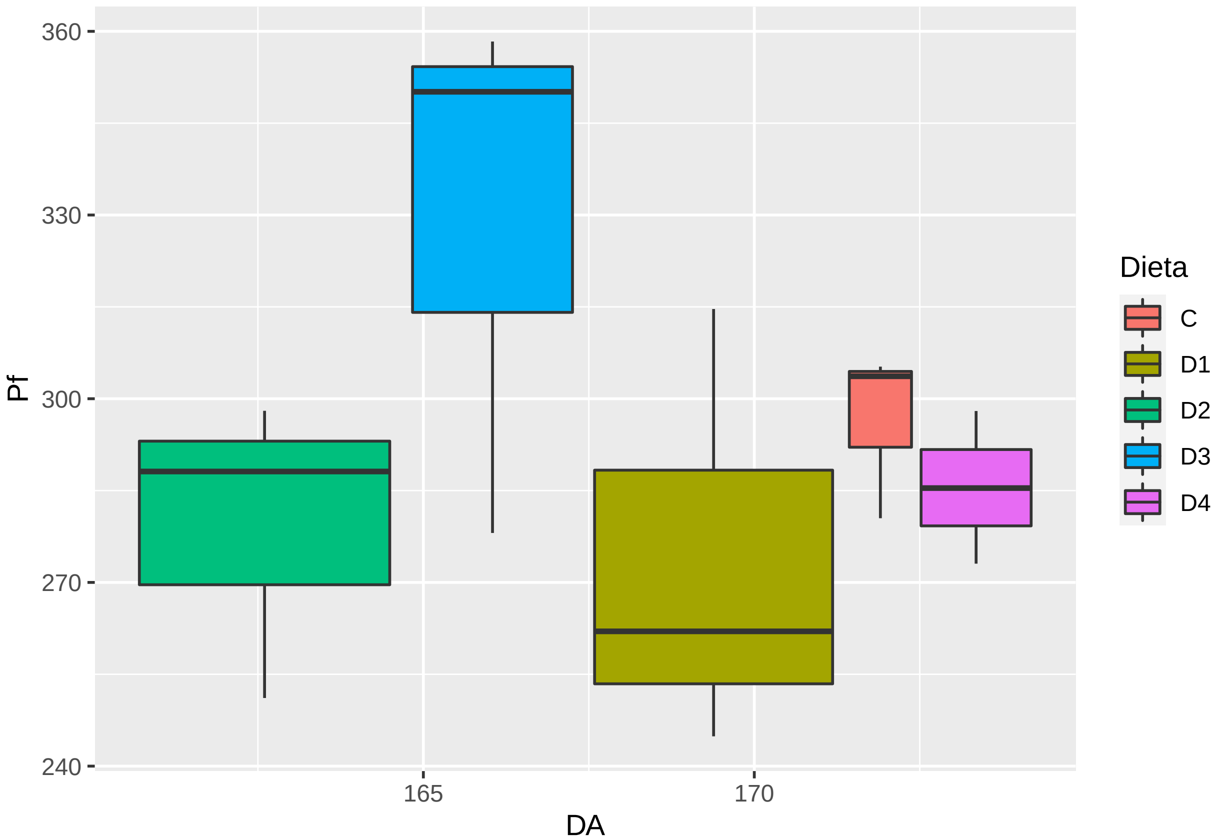

```r
ks.test(Datos_2$Dieta, Datos_2$Pf, alternative = c("two.sided"))
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  Datos_2$Dieta and Datos_2$Pf
## D = 1, p-value = 6.118e-07
## alternative hypothesis: two-sided
```

```r
Datos_2$Estanques <- as.factor(Datos_2$Estanques)
str(Datos_2)
```

```
## tibble [15 x 12] (S3: tbl_df/tbl/data.frame)
##  $ Estanques: Factor w/ 15 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ Dieta    : Factor w/ 5 levels "C","D1","D2",..: 2 5 4 1 5 3 1 4 3 2 ...
##  $ Pi       : num [1:15] 148 152 150 151 156 ...
##  $ Pf       : num [1:15] 262 285 350 305 298 ...
##  $ Li       : num [1:15] 23.4 23.1 22.7 22.7 23.3 ...
##  $ Lf       : num [1:15] 26.2 26.3 28.4 27.6 26.8 ...
##  $ DA       : num [1:15] 181 177 174 173 175 ...
##  $ Ki       : num [1:15] 0.0115 0.0123 0.0129 0.0129 0.0124 ...
##  $ Kf       : num [1:15] 0.0146 0.0157 0.0147 0.0146 0.0155 ...
##  $ FCA      : num [1:15] 1.589 1.323 0.948 1.118 1.228 ...
##  $ CUT      : num [1:15] 0.00213 0.00249 0.00434 0.00376 0.00272 ...
##  $ SGR      : num [1:15] 0.621 0.686 0.923 0.767 0.704 ...
```

```r
ggplot(data = Datos_2, aes(x = DA, y = Pf, fill = Dieta)) +
  geom_boxplot() + theme(legend.position = "right")
```




```
mod.2 <- lm(Pf ~ Dieta:DA, data = Datos_2)
summary(mod.2)
```

```
##
## Call:
## lm(formula = Pf ~ Dieta:DA, data = Datos_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.698  -7.785   0.554   8.313  42.372
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -80.5331   139.4459  -0.578   0.5777
## DietaC:DA     2.2016     0.8175   2.693   0.0247 *
## DietaD1:DA    2.0823     0.8222   2.533   0.0321 *
## DietaD2:DA    2.1826     0.8475   2.575   0.0299 *
## DietaD3:DA    2.4274     0.8275   2.933   0.0167 *
## DietaD4:DA    2.1198     0.8107   2.615   0.0281 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.62 on 9 degrees of freedom
## Multiple R-squared:  0.6757, Adjusted R-squared:  0.4955
## F-statistic:  3.75 on 5 and 9 DF,  p-value: 0.04112
```

```
mod.3 <- lm(Pf ~ Dieta + DA, data = Datos_2)
summary(mod.3)
```

```
##
## Call:
## lm(formula = Pf ~ Dieta + DA, data = Datos_2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -36.631  -7.887   0.952   9.206  41.733
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -79.0774   146.8125  -0.539   0.6032
## DietaD1     -19.5031    19.0559  -1.023   0.3328
## DietaD2      -2.9772    19.8260  -0.150   0.8839
## DietaD3      37.4362    19.1187   1.958   0.0819 .
## DietaD4     -13.9859    19.0532  -0.734   0.4816
## DA            2.1926     0.8536   2.569   0.0302 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23.29 on 9 degrees of freedom
## Multiple R-squared:  0.6562, Adjusted R-squared:  0.4653
## F-statistic: 3.436 on 5 and 9 DF,  p-value: 0.05173
```

```
anova(mod.2, mod.3)
```

```
## Analysis of Variance Table
##
## Model 1: Pf ~ Dieta:DA
## Model 2: Pf ~ Dieta + DA
##   Res.Df    RSS Df Sum of Sq F Pr(>F)
## 1      9 4606.0
## 2      9 4882.3  0   -276.27
```

3 Correlacion entre Factores fisico quimicos en el peso final

analisis de componentes principales

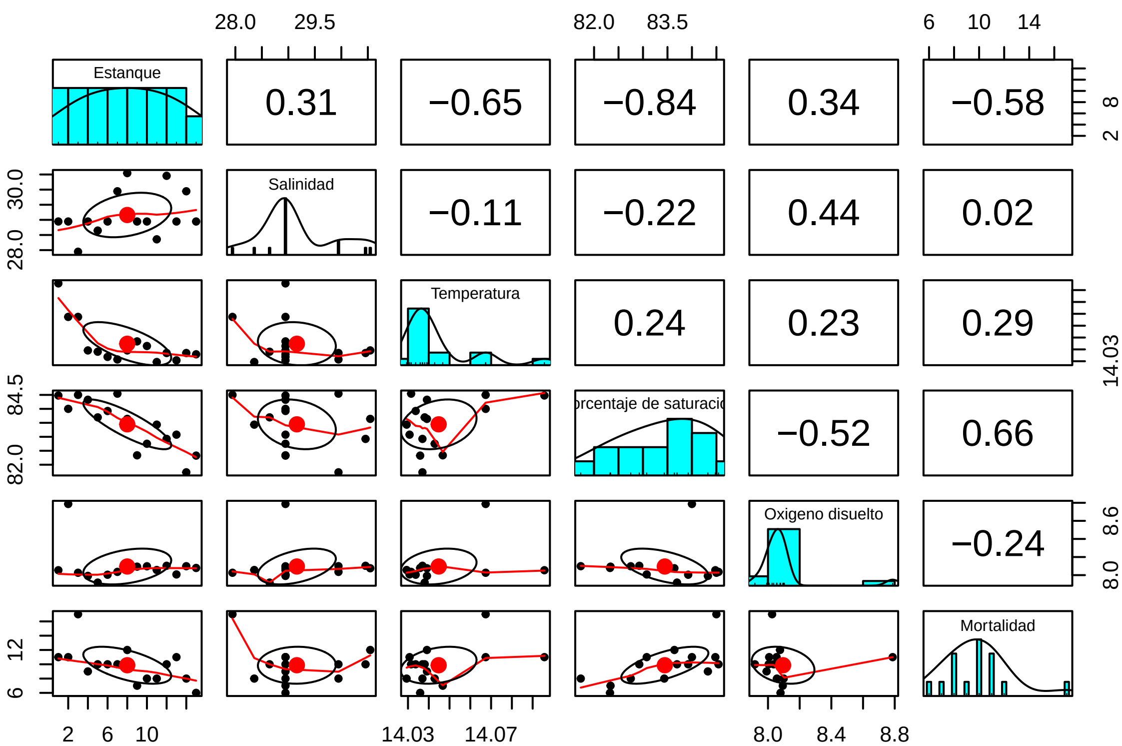
```
Datos_3 <- read_excel("Datos.xlsx", sheet = "Hoja 4")
summary(Datos_3)
```

```
##     Dieta              Estanque      Salinidad      Temperatura
##  Length:15          Min.   : 1.0   Min.   :27.95   Min.   :14.03
##  Class :character   1st Qu.: 4.5   1st Qu.:28.95   1st Qu.:14.03
##  Mode  :character   Median : 8.0   Median :28.95   Median :14.04
##                     Mean   : 8.0   Mean   :29.16   Mean   :14.04
##                     3rd Qu.:11.5   3rd Qu.:29.45   3rd Qu.:14.04
##                     Max.   :15.0   Max.   :30.55   Max.   :14.10
##  Porcentaje de saturacion Oxigeno disuelto   Mortalidad
##  Min.   :81.73            Min.   :7.917    Min.   : 6.000
##  1st Qu.:82.84            1st Qu.:8.017    1st Qu.: 8.000
##  Median :83.64            Median :8.057    Median :10.000
##  Mean   :83.45            Mean   :8.095    Mean   : 9.867
##  3rd Qu.:84.16            3rd Qu.:8.096    3rd Qu.:11.000
##  Max.   :84.54            Max.   :8.787    Max.   :17.000
```

```
str(Datos_3)
```

```
## tibble [15 x 7] (S3: tbl_df/tbl/data.frame)
##  $ Dieta                   : chr [1:15] "D1" "D4" "D3" "C" ...
##  $ Estanque                : num [1:15] 1 2 3 4 5 6 7 8 9 10 ...
##  $ Salinidad               : num [1:15] 28.9 28.9 27.9 28.9 28.6 ...
##  $ Temperatura             : num [1:15] 14.1 14.1 14.1 14 14 ...
##  $ Porcentaje de saturacion: num [1:15] 84.5 84 84.5 84.3 83.7 ...
##  $ Oxigeno disuelto        : num [1:15] 8.05 8.79 8.03 7.99 7.92 ...
##  $ Mortalidad              : num [1:15] 11 11 17 9 10 10 10 12 7 8 ...
```

```
pairs.panels(Datos_3[2:7], method = "spearman")
```




```
Datos_3_New <- as.data.frame(Datos_3[,3:7])
row.names(Datos_3_New) <- Datos_3$Estanque
head(Datos_3_New)
```

```
##   Salinidad Temperatura Porcentaje de saturacion Oxigeno disuelto Mortalidad
## 1  28.94565    14.09565                 84.47826         8.054348         11
## 2  28.94565    14.06739                 84.00000         8.786957         11
## 3  27.94565    14.06739                 84.50000         8.026087         17
## 4  28.94565    14.03913                 84.32609         7.991304          9
## 5  28.64565    14.03804                 83.69565         7.917391         10
## 6  28.94565    14.03370                 83.92391         8.004348         10
```

```
mvn(Datos_3_New, univariateTest = "SW", tol=1.51413e-60)
```

```
## $multivariateNormality
##            Test       HZ      p value MVN
## 1 Henze-Zirkler 1.142499 4.975342e-05  NO
##
## $univariateNormality
```

```
##           Test                 Variable Statistic   p value Normality
## 1 Shapiro-Wilk        Salinidad            0.8538  0.0197      NO
## 2 Shapiro-Wilk       Temperatura           0.7435   8e-04      NO
## 3 Shapiro-Wilk Porcentaje de saturacion    0.9383   0.362      YES
## 4 Shapiro-Wilk     Oxigeno disuelto        0.5296  <0.001      NO
## 5 Shapiro-Wilk        Mortalidad           0.8882   0.063      YES
##
## $Descriptives
##                           n      Mean    Std.Dev    Median       Min       Max
## Salinidad                15 29.160435 0.73619815 28.945652 27.945652 30.545652
## Temperatura              15 14.044763 0.01824863 14.038043 14.029348 14.095652
## Porcentaje de saturacion 15 83.445652 0.89122860 83.641304 81.728261 84.543478
## Oxigeno disuelto         15  8.095290 0.19803146  8.056522  7.917391  8.786957
## Mortalidad               15  9.866667 2.58751582 10.000000  6.000000 17.000000
##                               25th      75th       Skew   Kurtosis
## Salinidad                28.945652 29.445652  0.5472884 -0.7426032
## Temperatura              14.034783 14.044798  1.5639114  1.4419443
## Porcentaje de saturacion 82.836957 84.163043 -0.3558923 -1.2367449
## Oxigeno disuelto          8.016848  8.095652  2.7644607  6.9768663
## Mortalidad                8.000000 11.000000  1.0602953  1.3768227
```

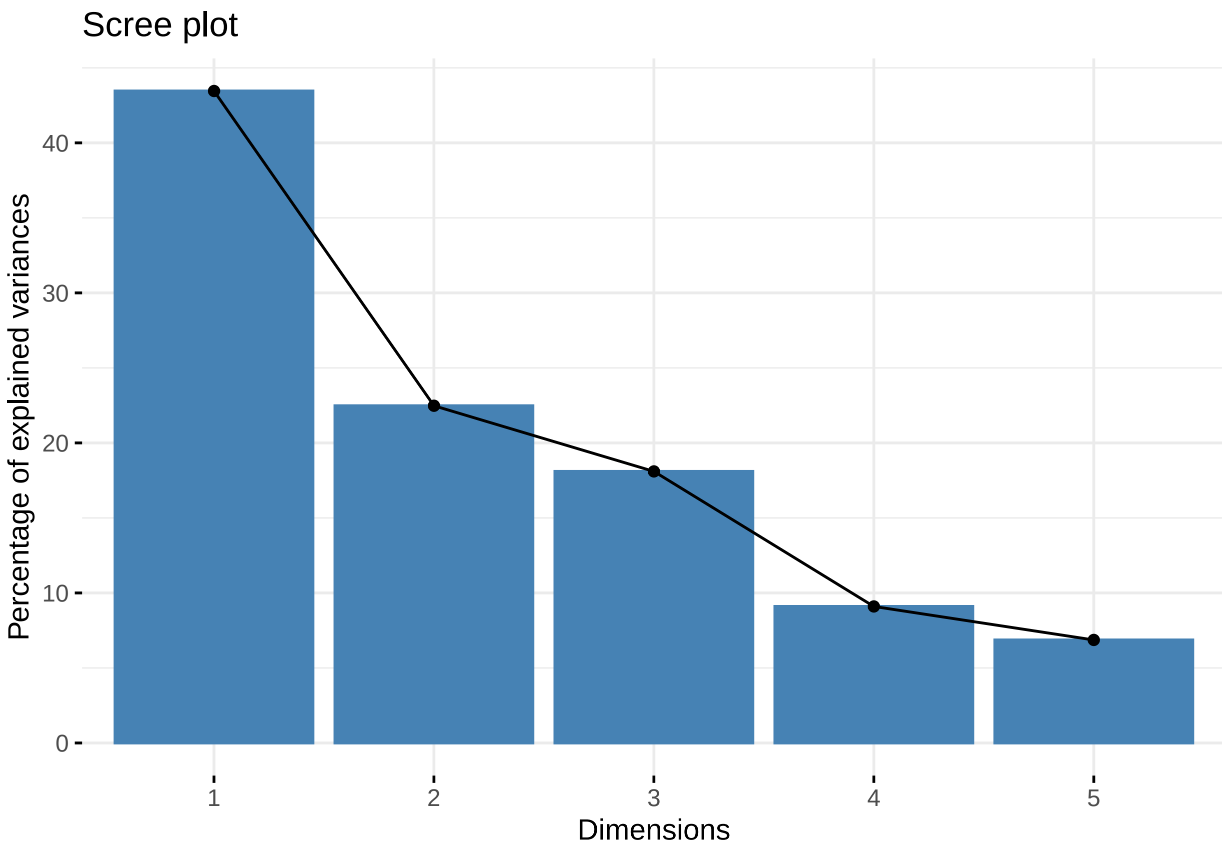
```
Datos_New <- prcomp(Datos_3_New, scale = TRUE)
Datos_New
```

```
## Standard deviations (1, .., p=5):
## [1] 1.4740353 1.0601147 0.9513236 0.6745317 0.5859753
##
## Rotation (n x k) = (5 x 5):
##                                 PC1        PC2        PC3         PC4
## Salinidad                 0.3079060  0.2416240 -0.8766286  0.23208183
## Temperatura              -0.5240637  0.3217251  0.1395325  0.74813751
## Porcentaje de saturacion -0.5494087 -0.2515154 -0.2558275 -0.40482082
## Oxigeno disuelto         -0.1753909  0.8583325  0.0327372 -0.47088253
## Mortalidad               -0.5458382 -0.1952336 -0.3814887 -0.02860146
##                                  PC5
## Salinidad                 0.15643573
## Temperatura               0.20656993
## Porcentaje de saturacion  0.63683791
## Oxigeno disuelto         -0.09849511
## Mortalidad               -0.71943931
```

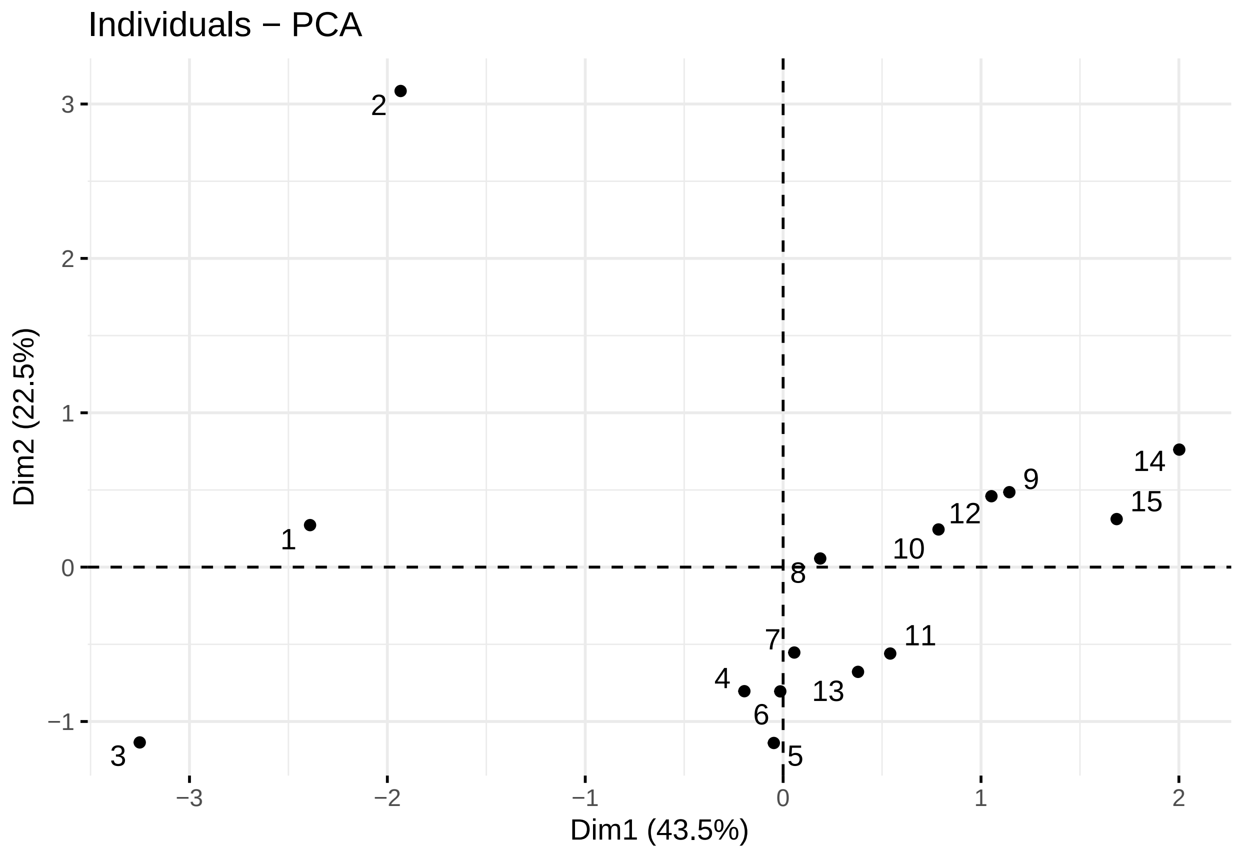
```
get_eigenvalue(Datos_New)
```

```
##       eigenvalue variance.percent cumulative.variance.percent
## Dim.1  2.1727801        43.455602                    43.45560
## Dim.2  1.1238432        22.476863                    65.93247
## Dim.3  0.9050166        18.100333                    84.03280
## Dim.4  0.4549931         9.099861                    93.13266
## Dim.5  0.3433670         6.867340                   100.00000
```

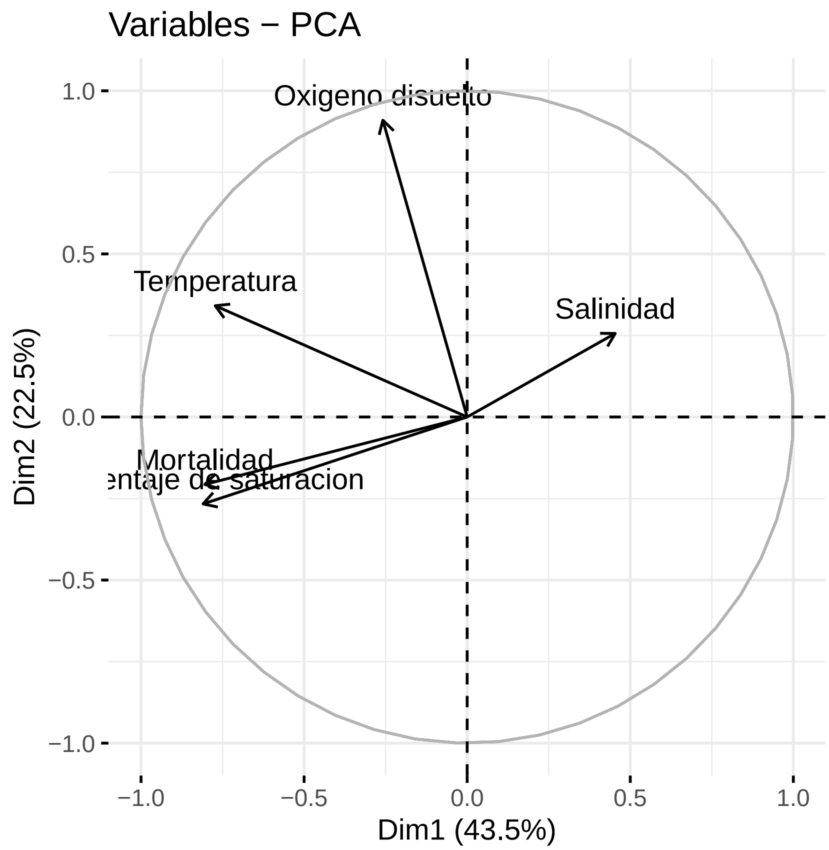
```
fviz_eig(Datos_New)
```



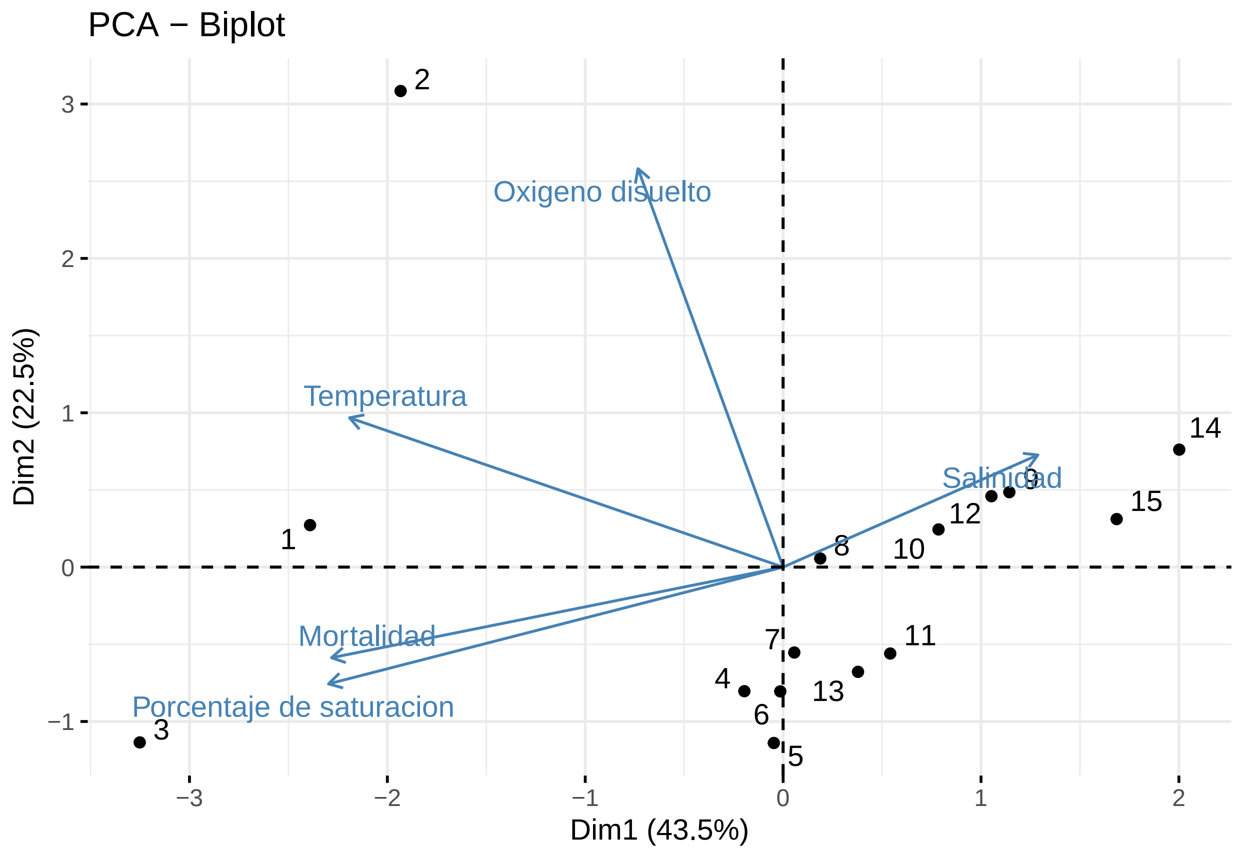

```
fviz_pca_ind(Datos_New, repel = TRUE)
```




```
fviz_pca_var(Datos_New)
```




```
fviz_pca_biplot(Datos_New, repel = TRUE)
```


PCA − Biplot
Dim2 (22.5%)
Dim1 (43.5%)
3
2
1
0
−1
−3
−2
−1
0
1
2
Oxigeno disuelto
Temperatura
Salinidad
Mortalidad
Porcentaje de saturacion
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15