

Tarea_diplomado_final_Paula

Dra. Paula Celis-Plá

2022-10-04

Evaluación de las respuestas fisiológicas y ambientales del alga *Lessonia spicata* en la Bahía de Valparaíso en presencia de Cambio Climático

**** Descripción del Trabajo****

** La data actual corresponde a valores de variables fisiológicas y ambientales para el alga parda *Lessonia spicata*.

** Este estudio tiene por objetivo, evaluar la diferencias estadísticas de las variables fisiológicas y variables ambientales respecto a las variables categóricas estación del Año con 4 niveles; Otoño, Invierno, Primavera y Verano, y tiempo para 3 ciclos diarios con 3 niveles; día 1, día 2 y día 3 con 9 valores.

** Las variables fisiológicas analizadas en este estudio fueron Clorofila a o Cla (expresada en microgramos * gramos-1 de peso seco), Clorofila c = Clc (expresada en microgramos * gramos-1 de peso seco), Carotenos o Car (expresada en microgramos * gramos-1 de peso seco), Compuestos fenólicos o PC (expresada en miligramos * gramos-1 de peso seco), Actividad antioxidante o DPPH = actividad antioxidante (expresada en micromol de unidades de trolox * gramos-1 de peso seco), y las variables ambientales; temperatura (expresada en °C), pH (expresada en unidades de pH), Salinidad (expresada en psu) y Radiación fotosintéticamente activa o PAR (expresada en micromoles de fotones * m⁻² * s⁻¹), con un total de 108 observaciones.

Habilita Librerías

Variables categóricas Seasons, time y replicate asigna factor

Transforma variables a factores

```
Datos_Proyecto <- read_excel("Datos_Proyecto.xlsx")
summary(Datos_Proyecto)
```

```
##      Seasons              time              Chla              Chlc
## Length:108      Length:108      Min.       :0.6730      Min.       :0.01000
## Class :character Class :character 1st Qu.:0.9888      1st Qu.:0.06775
## Mode  :character Mode  :character Median :1.1660      Median :0.09500
##                                     Mean  :1.1856      Mean   :0.09365
##                                     3rd Qu.:1.3813      3rd Qu.:0.10975
##                                     Max.   :1.9690      Max.   :0.22100
##
##      Car              PC              DPPH              Temperature
## Min.       :0.4190      Min.       : 4.600      Min.       : 3.792      Min.       :11.54
## 1st Qu.:0.6240      1st Qu.: 9.953      1st Qu.: 6.539      1st Qu.:12.47
## Median :0.7520      Median :13.815      Median : 7.769      Median :13.29
## Mean   :0.7557      Mean   :15.153      Mean   : 7.444      Mean   :14.08
## 3rd Qu.:0.8780      3rd Qu.:18.990      3rd Qu.: 8.433      3rd Qu.:15.35
## Max.   :1.2080      Max.   :37.340      Max.   :10.097      Max.   :18.02
##
##      pH              Salinity              PAR
```

```
## Min. :7.620 Min. :16.44 Min. : 9.768
## 1st Qu.:7.850 1st Qu.:29.62 1st Qu.: 196.000
## Median :7.960 Median :34.10 Median : 525.793
## Mean :7.982 Mean :30.49 Mean : 569.983
## 3rd Qu.:8.150 3rd Qu.:34.63 3rd Qu.: 764.500
## Max. :8.500 Max. :34.90 Max. :1921.000
```

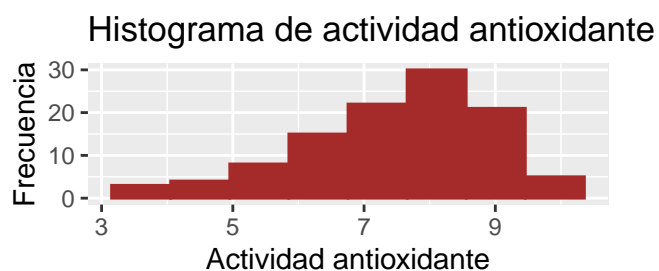
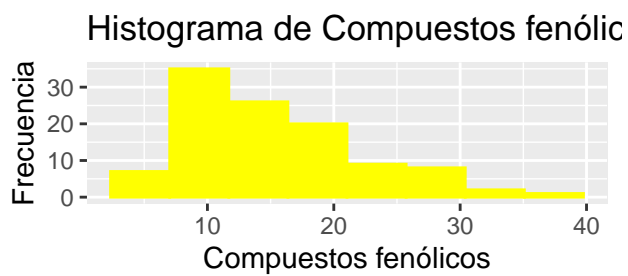
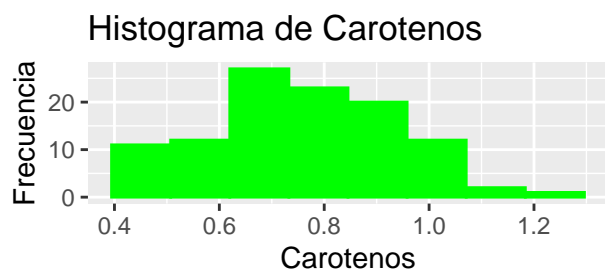
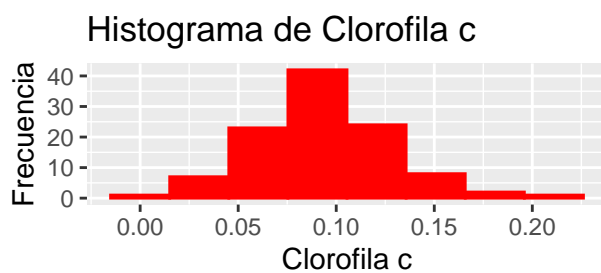
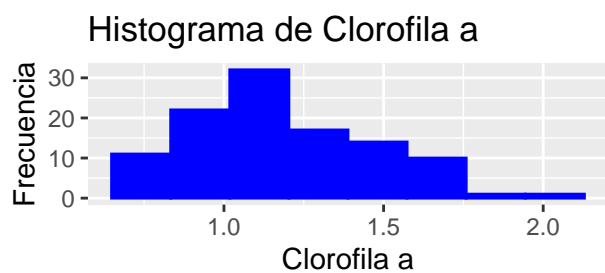
```
Datos_Proyecto$Seasons <- as.factor(Datos_Proyecto$Seasons)
Datos_Proyecto$time <- as.factor(Datos_Proyecto$time)
summary(Datos_Proyecto)
```

```
## Seasons time Chla Chlc Car
## Autumn:27 day 1:36 Min. :0.6730 Min. :0.01000 Min. :0.4190
## Spring:27 day 2:36 1st Qu.:0.9888 1st Qu.:0.06775 1st Qu.:0.6240
## Summer:27 day 3:36 Median :1.1660 Median :0.09500 Median :0.7520
## Winter:27 Mean :1.1856 Mean :0.09365 Mean :0.7557
## 3rd Qu.:1.3813 3rd Qu.:0.10975 3rd Qu.:0.8780
## Max. :1.9690 Max. :0.22100 Max. :1.2080
## PC DPPH Temperature pH
## Min. : 4.600 Min. : 3.792 Min. :11.54 Min. :7.620
## 1st Qu.: 9.953 1st Qu.: 6.539 1st Qu.:12.47 1st Qu.:7.850
## Median :13.815 Median : 7.769 Median :13.29 Median :7.960
## Mean :15.153 Mean : 7.444 Mean :14.08 Mean :7.982
## 3rd Qu.:18.990 3rd Qu.: 8.433 3rd Qu.:15.35 3rd Qu.:8.150
## Max. :37.340 Max. :10.097 Max. :18.02 Max. :8.500
## Salinity PAR
## Min. :16.44 Min. : 9.768
## 1st Qu.:29.62 1st Qu.: 196.000
## Median :34.10 Median : 525.793
## Mean :30.49 Mean : 569.983
## 3rd Qu.:34.63 3rd Qu.: 764.500
## Max. :34.90 Max. :1921.000
```

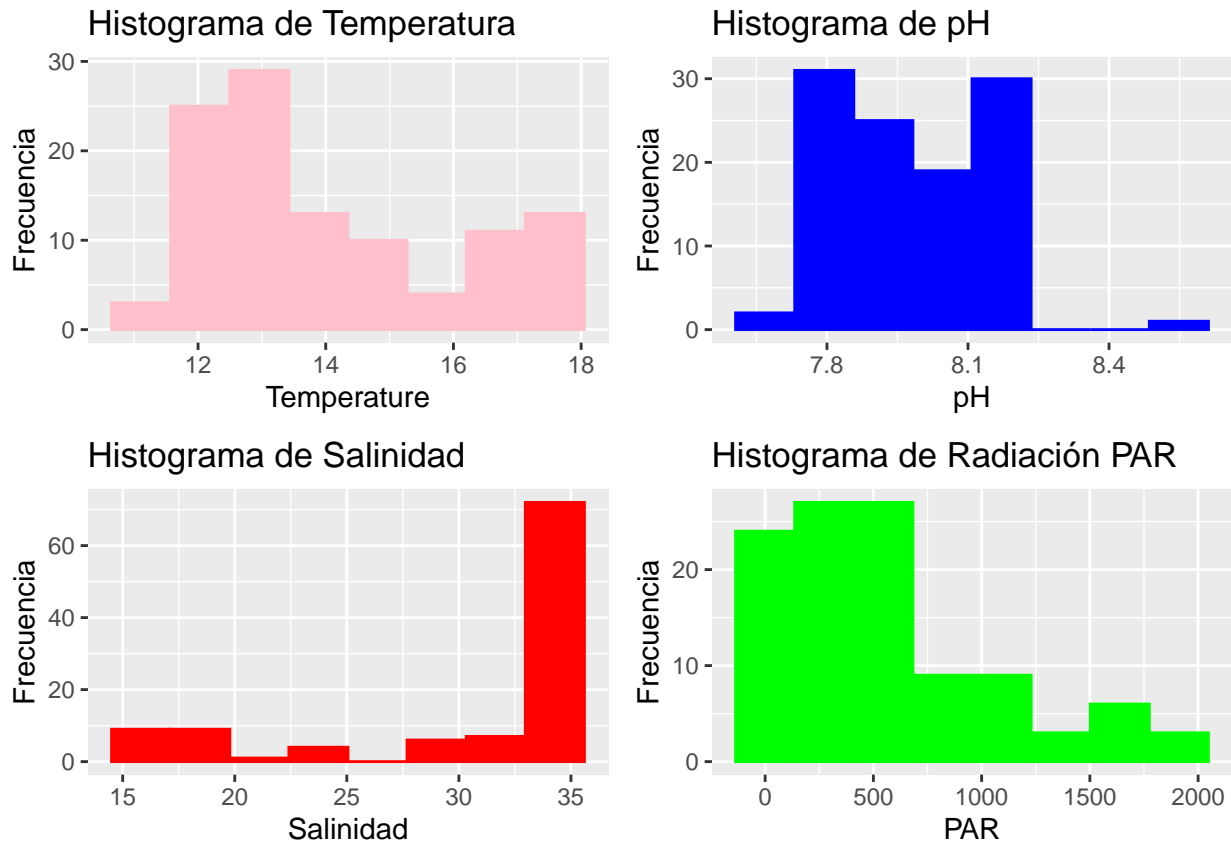
Histogramas con etiquetas y títulos

```
p1 <- ggplot(Datos_Proyecto, aes(Chla)) + geom_histogram(bins = 8, color = "blue", fill="blue")+ labs(title="Histograma de Chla")
p2 <- ggplot(Datos_Proyecto, aes(Chlc)) + geom_histogram(bins = 8, color = "red", fill="red")+ labs(title="Histograma de Chlc")
p3 <- ggplot(Datos_Proyecto, aes(Car)) + geom_histogram(bins = 8, color = "green", fill="green")+ labs(title="Histograma de Car")
p4 <- ggplot(Datos_Proyecto, aes(PC)) + geom_histogram(bins = 8, color = "yellow", fill="yellow")+ labs(title="Histograma de PC")
p5 <- ggplot(Datos_Proyecto, aes(DPPH)) + geom_histogram(bins = 8, color = "brown", fill="brown")+ labs(title="Histograma de DPPH")
p6 <- ggplot(Datos_Proyecto, aes(Temperature)) + geom_histogram(bins = 8, color = "pink", fill="pink")+ labs(title="Histograma de Temperature")
p7 <- ggplot(Datos_Proyecto, aes(pH)) + geom_histogram(bins = 8, color = "blue", fill="blue")+ labs(title="Histograma de pH")
p8 <- ggplot(Datos_Proyecto, aes(Salinity)) + geom_histogram(bins = 8, color = "red", fill="red")+ labs(title="Histograma de Salinity")
p9 <- ggplot(Datos_Proyecto, aes(PAR)) + geom_histogram(bins = 8, color = "green", fill="green")+ labs(title="Histograma de PAR")

gridExtra::grid.arrange(p1, p2, p3, p4, p5, ncol = 2)
```



```
gridExtra::grid.arrange(p6, p7, p8, p9, ncol = 2)
```



4. Datos balanceados y tablas de frecuencia

** Los datos estan balanceados

```
str(Datos_Proyecto)
```

```
## tibble [108 x 11] (S3: tbl_df/tbl/data.frame)
## $ Seasons      : Factor w/ 4 levels "Autumn","Spring",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ time         : Factor w/ 3 levels "day 1","day 2",...: 1 1 1 1 1 1 1 1 1 2 ...
## $ Chla         : num [1:108] 1.04 1.6 1.52 1.48 1.14 ...
## $ Chlc         : num [1:108] 0.095 0.155 0.143 0.152 0.108 0.115 0.148 0.105 0.102 0.134 ...
## $ Car          : num [1:108] 0.605 0.954 0.931 0.905 0.694 0.752 0.884 0.659 0.62 0.937 ...
## $ PC           : num [1:108] 9.12 13.87 15.83 19.89 19.81 ...
## $ DPPH         : num [1:108] 7.58 6.08 7.81 7.02 7.2 ...
## $ Temperature: num [1:108] 12.3 12.3 12.3 12.3 12.3 ...
## $ pH           : num [1:108] 7.84 7.84 7.84 7.85 7.85 7.85 7.88 7.88 7.88 7.84 ...
## $ Salinity     : num [1:108] 33.7 33.7 33.7 17.9 17.9 ...
## $ PAR          : num [1:108] 594 594 594 1563 1563 ...
```

```
knitr::kable(table(Datos_Proyecto$Seasons, Datos_Proyecto$time), caption = "Tabla de contingencia")
```

Table 1: Tabla de contingencia

	day 1	day 2	day 3
Autumn	9	9	9
Spring	9	9	9
Summer	9	9	9

	day 1	day 2	day 3
Winter	9	9	9

5. Relación entre variables cuantitativas y factores

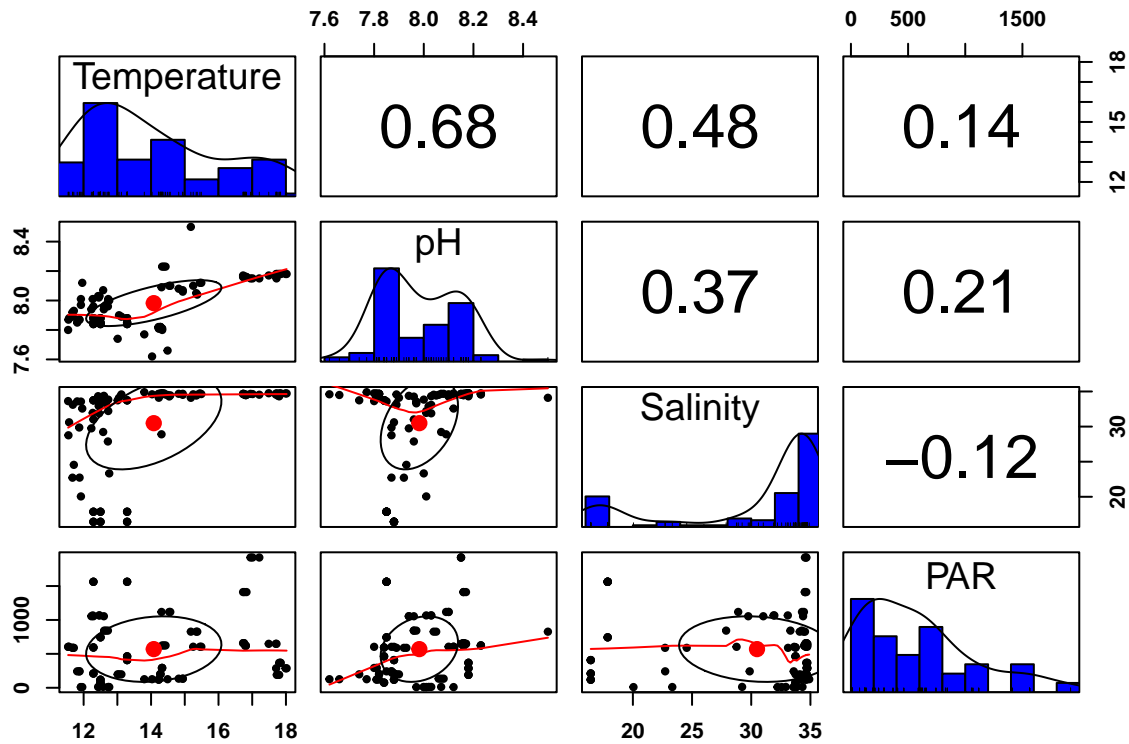
** No se incluyan las variables Seasons, time, y Replicate porque son variables categóricas.

```
summary(Datos_Proyecto)
```

```
##      Seasons      time      Chla      Chlc      Car
## Autumn:27 day 1:36 Min. :0.6730 Min. :0.01000 Min. :0.4190
## Spring:27 day 2:36 1st Qu.:0.9888 1st Qu.:0.06775 1st Qu.:0.6240
## Summer:27 day 3:36 Median :1.1660 Median :0.09500 Median :0.7520
## Winter:27      Mean :1.1856 Mean :0.09365 Mean :0.7557
##      3rd Qu.:1.3813 3rd Qu.:0.10975 3rd Qu.:0.8780
##      Max. :1.9690 Max. :0.22100 Max. :1.2080
##      PC      DPPH      Temperature      pH
## Min. : 4.600 Min. : 3.792 Min. :11.54 Min. :7.620
## 1st Qu.: 9.953 1st Qu.: 6.539 1st Qu.:12.47 1st Qu.:7.850
## Median :13.815 Median : 7.769 Median :13.29 Median :7.960
## Mean :15.153 Mean : 7.444 Mean :14.08 Mean :7.982
## 3rd Qu.:18.990 3rd Qu.: 8.433 3rd Qu.:15.35 3rd Qu.:8.150
## Max. :37.340 Max. :10.097 Max. :18.02 Max. :8.500
##      Salinity      PAR
## Min. :16.44 Min. : 9.768
## 1st Qu.:29.62 1st Qu.:196.000
## Median :34.10 Median :525.793
## Mean :30.49 Mean :569.983
## 3rd Qu.:34.63 3rd Qu.:764.500
## Max. :34.90 Max. :1921.000
```

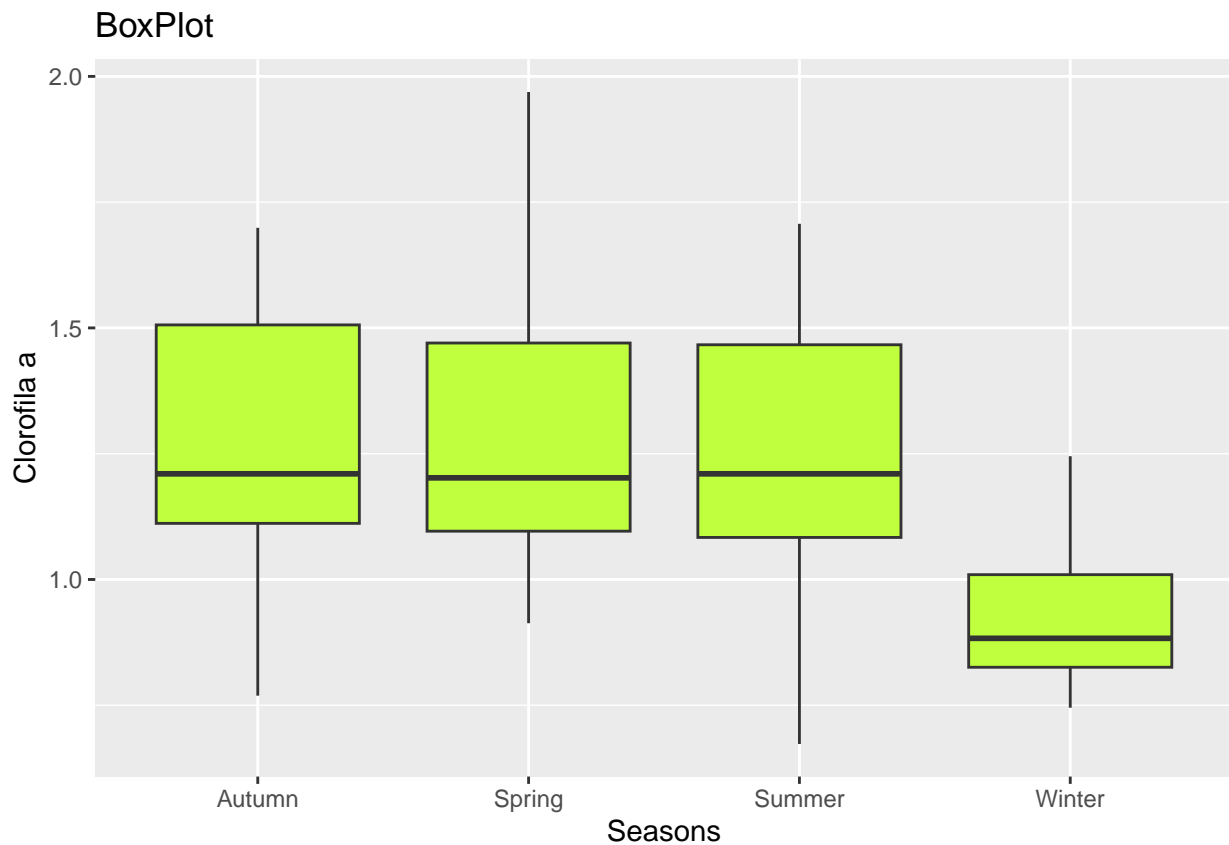
5. Gráficas de correlación de variables continuas (pearson)

```
pairs.panels(Datos_Proyecto[,3:11], method = "pearson", hist.col = "blue", density = TRUE, font=4)
```

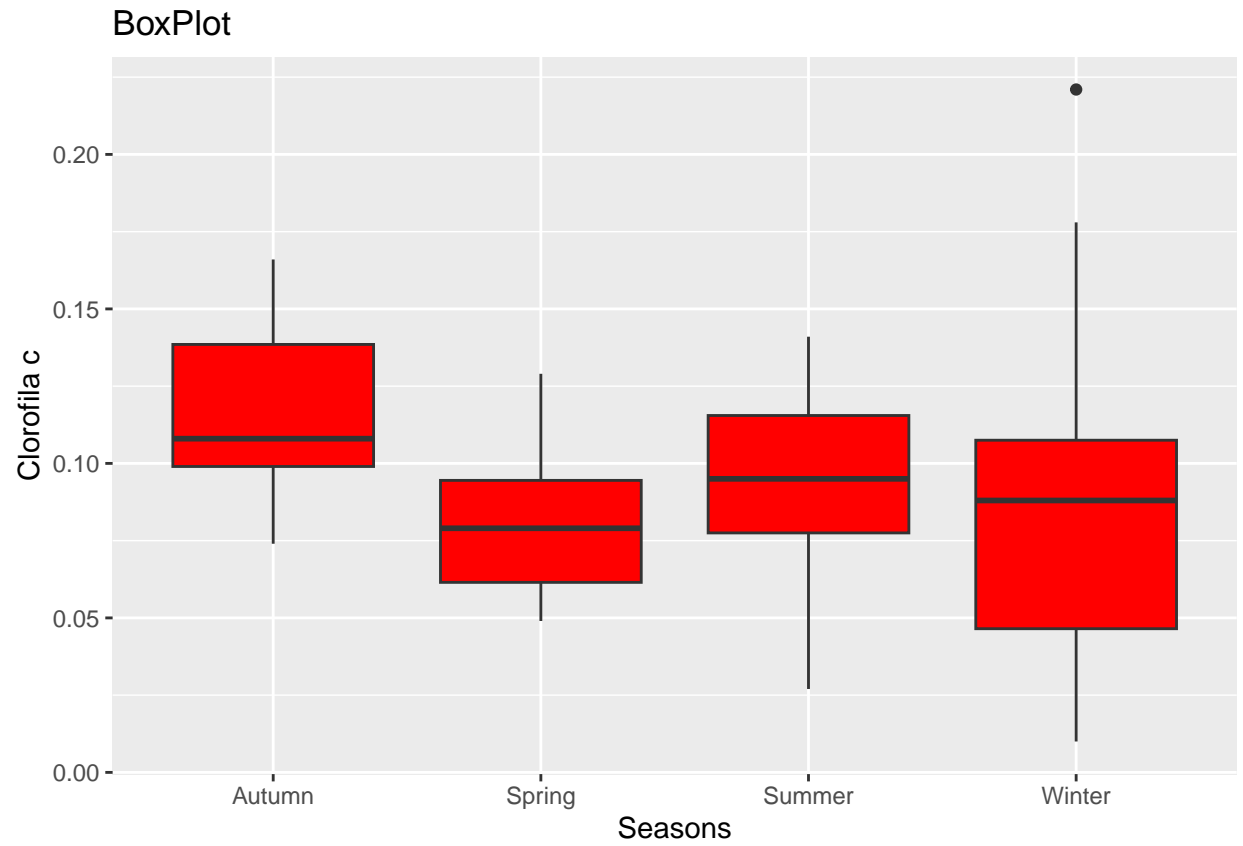



5. Relación entre variables continuas y factores (boxplot)

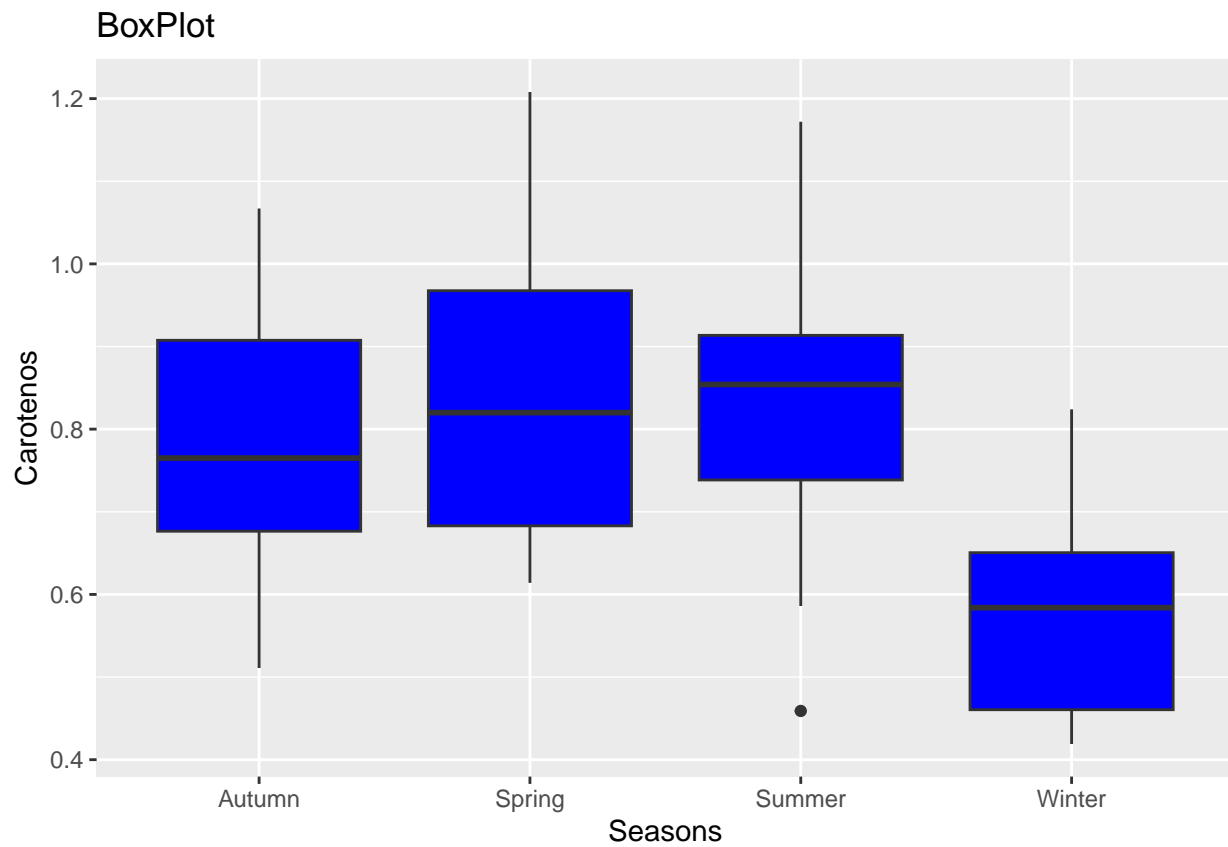
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=Chla))+geom_boxplot(fill="olivedrab1")+labs(title = "BoxPlot",
```



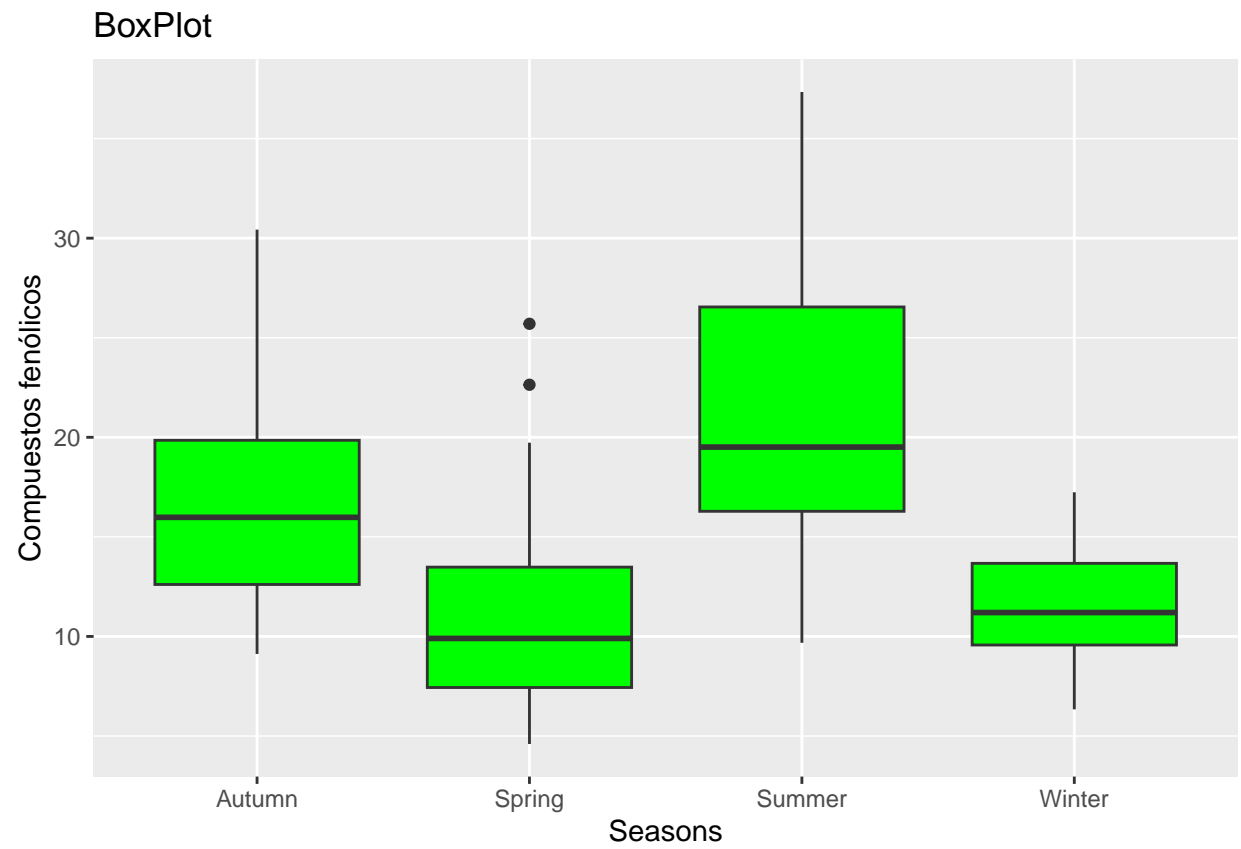
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=Chlc))+geom_boxplot(fill="red")+labs(title = "BoxPlot", x= "Se
```



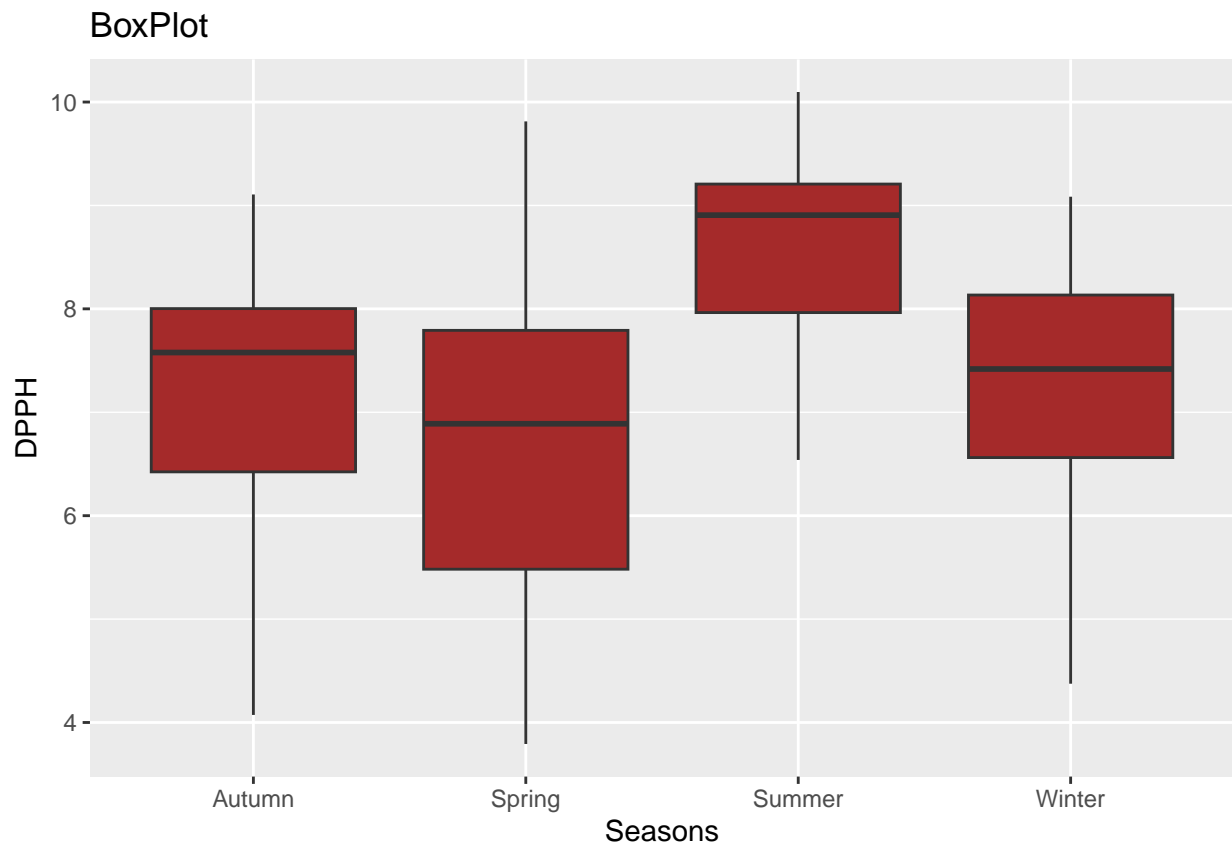
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=Car))+geom_boxplot(fill="blue")+labs(title = "BoxPlot", x= "Se
```

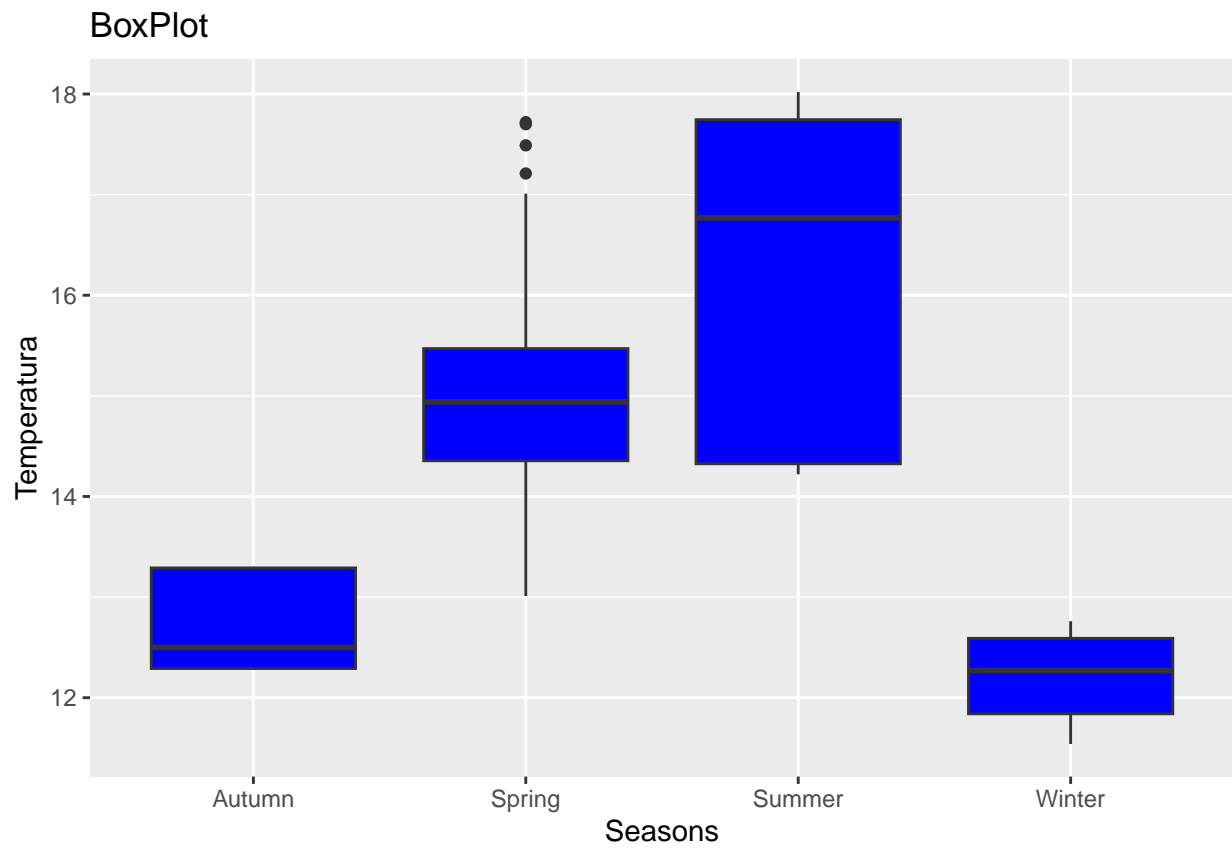
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=PC))+geom_boxplot(fill="green")+labs(title = "BoxPlot", x= "Seasons")
```



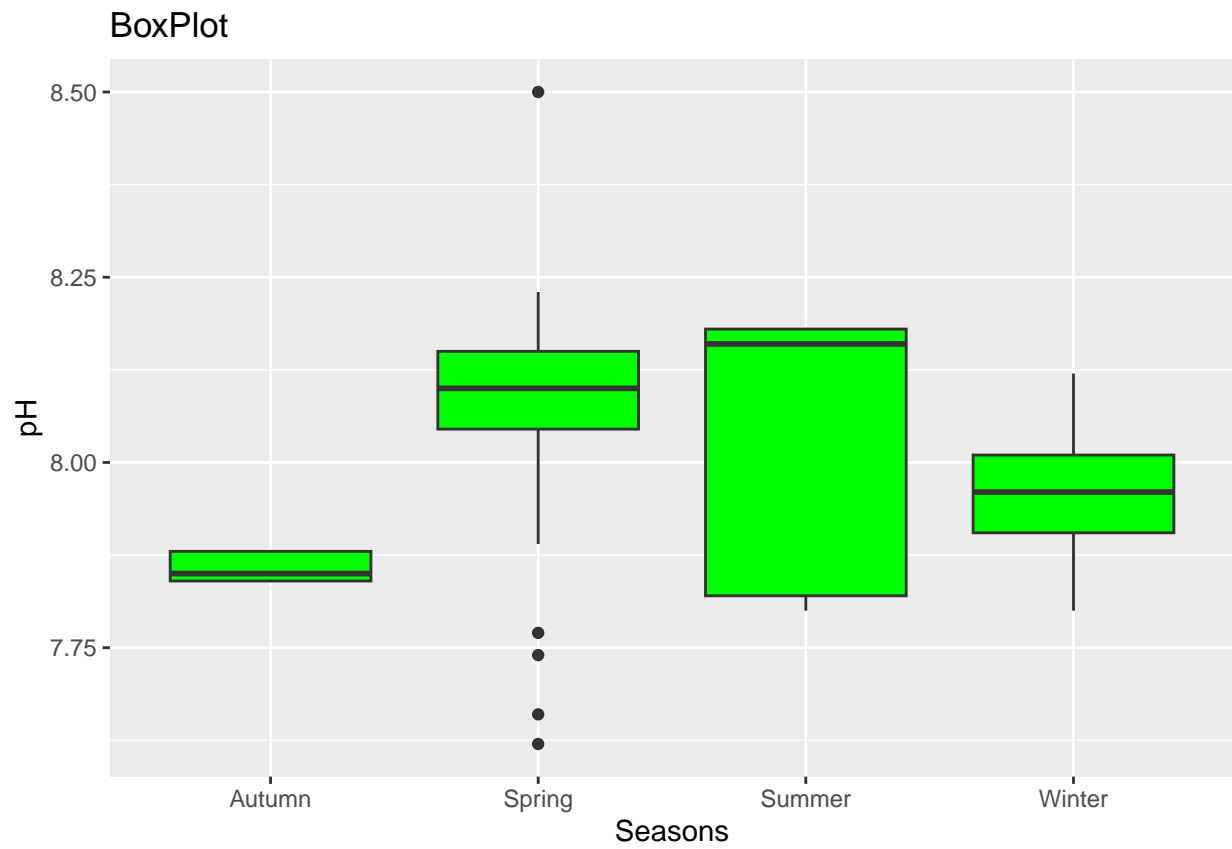
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=DPPH))+geom_boxplot(fill="brown")+labs(title = "BoxPlot", x= "Seasons", y= "Compuestos fenólicos")
```



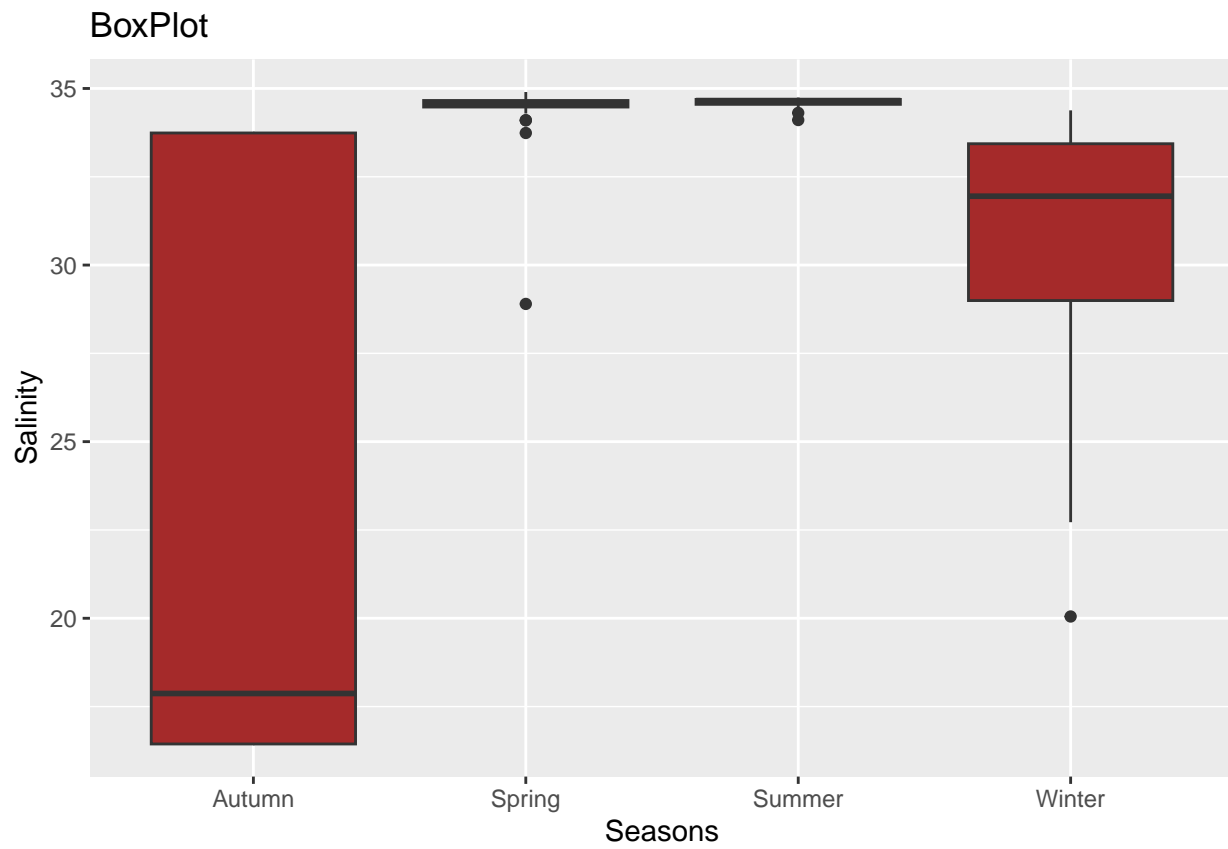
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=Temperature))+geom_boxplot(fill="blue")+labs(title = "BoxPlot")
```



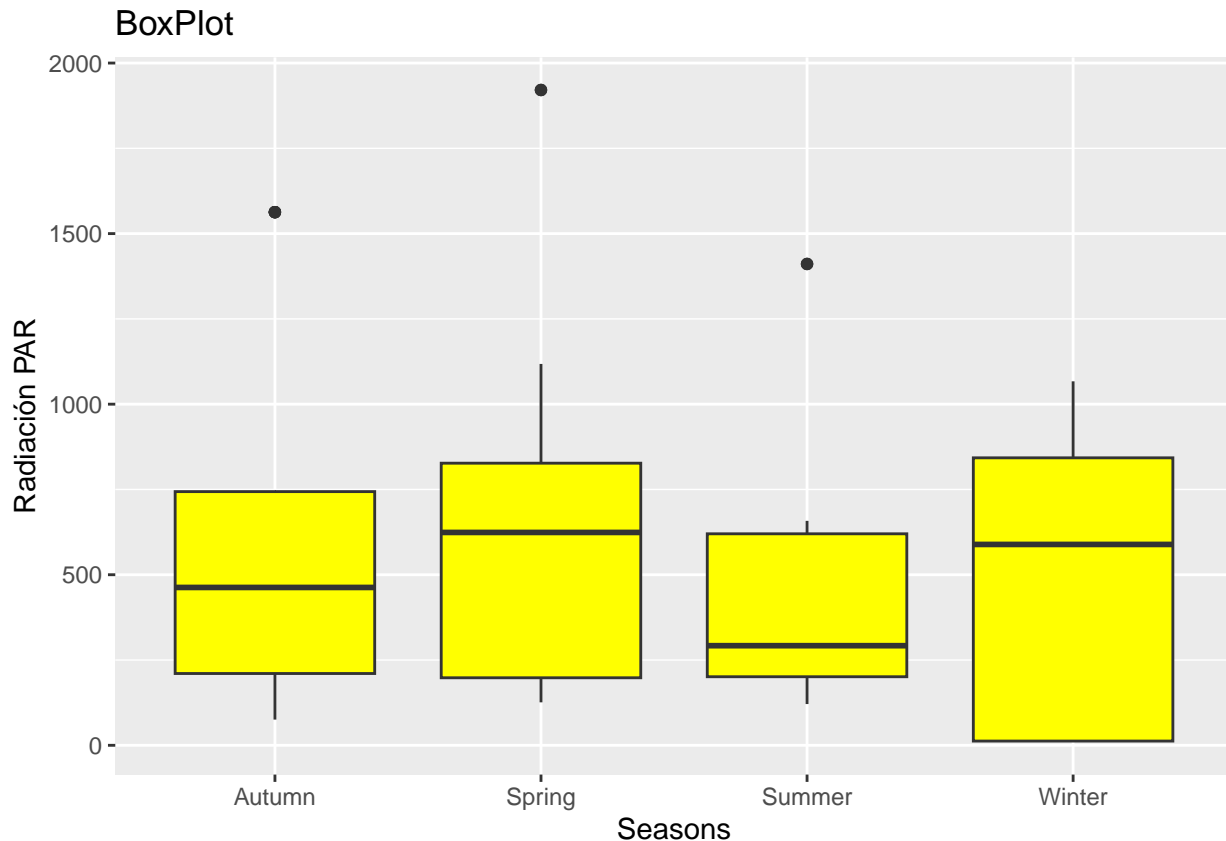
```
ggplot(Datos_Proyecto, aes(x= Seasons, y=pH))+geom_boxplot(fill="green")+labs(title = "BoxPlot", x= "Seasons")
```



```
ggplot(Datos_Proyecto, aes(x= Seasons, y=Salinity))+geom_boxplot(fill="brown")+labs(title = "BoxPlot",
```



```
ggplot(Datos_Proyecto, aes(x= Seasons, y=PAR))+geom_boxplot(fill="yellow")+labs(title = "BoxPlot", x= "Seasons", y= "PAR")
```



6. Identificación si existen errores, datos faltantes o error atípico

** En la variable Salinidad, existe poca dispersión de los datos en las estaciones de primavera y verano, se registran datos muy similares, por ellos se observó poca dispersión.

** No existen datos faltantes para cada variable.

** Los errores típicos en las variables se identifican a continuación;

** Variable Chlc : para la estación winter se registra un outlier Variable Car : para la estación summer se registra un outlier Variable PC : se registran 2 valores en Spring Variable Temperature: 3 valores en spring Variable pH : 5 valores en spring Variable Salinity: 3 valores en spring, 2 en summer y 1 en winter Variable PAR: 1 en autumn, 1 en spring y 1 en summer

7. Resumen de los datos con tablas y estadística descriptiva

```
Datos_Proyecto <- read_excel("Datos_Proyecto.xlsx" , sheet= 1)
head(Datos_Proyecto)
```

```
## # A tibble: 6 x 11
##   Seasons time   Chla Chlc   Car    PC  DPPH Temperature    pH Salinity  PAR
##   <chr>   <chr> <dbl> <dbl> <dbl> <dbl> <dbl>      <dbl> <dbl> <dbl> <dbl>
## 1 Autumn day 1  1.04 0.095 0.605  9.12  7.58      12.3  7.84    33.7  594.
## 2 Autumn day 1  1.60 0.155 0.954 13.9   6.08      12.3  7.84    33.7  594.
## 3 Autumn day 1  1.52 0.143 0.931 15.8   7.81      12.3  7.84    33.7  594.
## 4 Autumn day 1  1.48 0.152 0.905 19.9   7.02      12.3  7.85    17.9 1563.
## 5 Autumn day 1  1.14 0.108 0.694 19.8   7.20      12.3  7.85    17.9 1563.
## 6 Autumn day 1  1.17 0.115 0.752 21.6   5.12      12.3  7.85    17.9 1563.
```

```
select(Datos_Proyecto, Chla, Chlc, Car, DPPH)
```

```
## # A tibble: 108 x 4
##   Chla Chlc Car DPPH
##   <dbl> <dbl> <dbl> <dbl>
## 1  1.04 0.095 0.605  7.58
## 2  1.60 0.155 0.954  6.08
## 3  1.52 0.143 0.931  7.81
## 4  1.48 0.152 0.905  7.02
## 5  1.14 0.108 0.694  7.20
## 6  1.17 0.115 0.752  5.12
## 7  1.38 0.148 0.884  6.24
## 8  1.17 0.105 0.659  6.47
## 9  1.08 0.102 0.62   6.38
## 10 1.49 0.134 0.937  4.07
## # ... with 98 more rows
```

```
Datos_tab <- Datos_Proyecto %>% group_by(Seasons) %>% summarize(n = n(),
Promedio_Chla = mean(Chla), Maximo_Chla = max(Chla), Promedio_Chlc = mean(Chlc), Maximo_Chlc = max(Chlc),
Datos_tab
```

```
## # A tibble: 4 x 20
##   Seasons      n Promed~1 Maxim~2 Prome~3 Maxim~4 Prome~5 Maxim~6 Prome~7 Maxim~8
##   <chr>   <int>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 Autumn     27    1.28    1.70  0.117    0.166    0.790    1.07    16.5    30.4
## 2 Spring     27    1.30    1.97  0.0781   0.129    0.831    1.21    11.5    25.7
## 3 Summer     27    1.25    1.71  0.0935   0.141    0.825    1.17    21.0    37.3
## 4 Winter     27    0.920   1.25  0.0857   0.221    0.578    0.824    11.5    17.2
## # ... with 10 more variables: Promedio_DPPH <dbl>, Maximo_DPPH <dbl>,
## #   Promedio_Temperature <dbl>, Maximo_Temperature <dbl>, Promedio_pH <dbl>,
## #   Maximo_pH <dbl>, Promedio_Salinity <dbl>, Maximo_Salinity <dbl>,
## #   Promedio_PAR <dbl>, Maximo_PAR <dbl>, and abbreviated variable names
## #   1: Promedio_Chla, 2: Maximo_Chla, 3: Promedio_Chlc, 4: Maximo_Chlc,
## #   5: Promedio_Car, 6: Maximo_Car, 7: Promedio_PC, 8: Maximo_PC
```

```
knitr::kable(Datos_tab, caption = "Tabla de medidas resumen")
```

Table 2: Tabla de medidas resumen

Seasons	Promedio_Chla	Maximo_Chla	Promedio_Chlc	Maximo_Chlc	Promedio_CAR	Maximo_CAR	Promedio_PC	Maximo_PC	Promedio_DPPH	Maximo_DPPH	Promedio_Temperature	Maximo_Temperature	Promedio_pH	Maximo_pH	Promedio_Salinity	Maximo_Salinity	Promedio_PAR	Maximo_PAR
Autumn	1.2759	2.5990	0.1173	0.3336	0.7897	1.4067	0.1653	0.2643	7.1866	22.0609	6.9333	33.29	7.8566	7.88	22.6833	33.7400	67.6495	82.997
Spring	1.2997	1.9790	0.0780	0.1290	0.8307	0.3081	1.5123	2706.6831	1.8126	551451	97.72	8.0607	8.430	34.3214	8.9000	689.0000	1.000	1.000
Summer	1.2466	2.7670	0.0935	0.1811	0.8247	1.7778	221.0463	70348.5557	10.0960	829444	8.02	8.0514	8.18	34.6067	34.7426	650.8880	1.000	1.000
Winter	0.9205	1.2150	0.0856	0.2210	0.5776	0.6872	11.5140	7247.3495	2.9844	121187	82.76	7.9596	8.302	30.3392	30.3800	102.3935	17.000	17.000

```
Datos_tab2 <- Datos_Proyecto %>% group_by(time) %>% summarize(n = n(),
Promedio_Chla = mean(Chla), Maximo_Chla = max(Chla), Promedio_Chlc = mean(Chlc), Maximo_Chlc = max(Chlc),
Datos_tab2
```

```
## # A tibble: 3 x 20
##   time      n Promedio~1 Maxim~2 Prome~3 Maxim~4 Prome~5 Maxim~6 Prome~7 Maxim~8
##   <chr> <int>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 day 1    36    1.12    1.62  0.0892    0.155    0.707    1.01    14.4    34.6
```



```
## 2 day 2      36      1.20      1.70 0.0926  0.178  0.787  1.07  17.8  37.3
## 3 day 3      36      1.23      1.97 0.0992  0.221  0.774  1.21  13.3  26.8
## # ... with 10 more variables: Promedio_DPPH <dbl>, Maximo_DPPH <dbl>,
## #   Promedio_Temperature <dbl>, Maximo_Temperature <dbl>, Promedio_pH <dbl>,
## #   Maximo_pH <dbl>, Promedio_Salinity <dbl>, Maximo_Salinity <dbl>,
## #   Promedio_PAR <dbl>, Maximo_PAR <dbl>, and abbreviated variable names
## #   1: Promedio_Chla, 2: Maximo_Chla, 3: Promedio_Chlc, 4: Maximo_Chlc,
## #   5: Promedio_Car, 6: Maximo_Car, 7: Promedio_PC, 8: Maximo_PC
```

```
knitr::kable(Datos_tab2, caption = "Tabla de medidas resumen")
```

Table 3: Tabla de medidas resumen

time	Promedio_Chla	Maximo_Chla	Promedio_Chlc	Maximo_Chlc	Promedio_Car	Maximo_Car	Promedio_PC	Maximo_PC	Promedio_DPPH	Maximo_DPPH	Promedio_Temperature	Maximo_Temperature	Promedio_pH	Maximo_pH	Promedio_Salinity	Maximo_Salinity	Promedio_PAR	Maximo_PAR
day 36	1.117389	1.160089	1.089164	1.175070	1.066607	1.143684	1.098292	1.096986	0.311178	0.479667	7.976667	30.400101	7.426570	53.156299	7.1	1	1	1
day 36	1.204778	1.299009	1.092611	1.178078	1.066389	1.176973	1.347457	1.133847	1.197788	0.2	7.971385	30.371532	7.900600	60.330729	60	2	2	2
day 36	1.234750	1.069009	1.099162	1.077386	1.081332	1.227577	1.028126	1.035411	1.167780	7.998338	30.691037	7.500079	0.083021	1.000	3	3	3	3

8. Utiliza Paquetes para importar datos a R como readxl o similar y paquetes tidy, dplyr, ggplot2

```
messy <- read_excel("Datos_Proyecto.xlsx")
Datos_Proyecto$Seasons <- as.factor(Datos_Proyecto$Seasons)
Datos_Proyecto$time <- as.factor(Datos_Proyecto$time)
summary(Datos_Proyecto)
```

```
##      Seasons      time      Chla      Chlc      Car
## Autumn:27   day 1:36   Min.    :0.6730   Min.    :0.01000   Min.    :0.4190
## Spring:27   day 2:36   1st Qu.:0.9888   1st Qu.:0.06775   1st Qu.:0.6240
## Summer:27   day 3:36   Median :1.1660   Median :0.09500   Median :0.7520
## Winter:27                Mean  :1.1856   Mean  :0.09365   Mean  :0.7557
##                3rd Qu.:1.3813   3rd Qu.:0.10975   3rd Qu.:0.8780
##                Max.   :1.9690   Max.   :0.22100   Max.   :1.2080
##      PC      DPPH      Temperature      pH
## Min.    : 4.600   Min.    : 3.792   Min.    :11.54   Min.    :7.620
## 1st Qu.: 9.953   1st Qu.: 6.539   1st Qu.:12.47   1st Qu.:7.850
## Median :13.815   Median : 7.769   Median :13.29   Median :7.960
## Mean    :15.153   Mean    : 7.444   Mean    :14.08   Mean    :7.982
## 3rd Qu.:18.990   3rd Qu.: 8.433   3rd Qu.:15.35   3rd Qu.:8.150
## Max.    :37.340   Max.    :10.097   Max.    :18.02   Max.    :8.500
##      Salinity      PAR
## Min.    :16.44   Min.    : 9.768
## 1st Qu.:29.62   1st Qu.:196.000
## Median :34.10   Median :525.793
## Mean    :30.49   Mean    :569.983
## 3rd Qu.:34.63   3rd Qu.:764.500
## Max.    :34.90   Max.    :1921.000
```

```
summary(messy)
```

```
##      Seasons      time      Chla      Chlc
```

```
## Length:108      Length:108      Min.    :0.6730   Min.    :0.01000
## Class :character Class :character 1st Qu.:0.9888   1st Qu.:0.06775
## Mode  :character Mode  :character Median :1.1660   Median :0.09500
##                                     Mean  :1.1856   Mean  :0.09365
##                                     3rd Qu.:1.3813   3rd Qu.:0.10975
##                                     Max.   :1.9690   Max.   :0.22100
##      Car      PC      DPPH      Temperature
## Min.    :0.4190 Min.    : 4.600 Min.    : 3.792 Min.    :11.54
## 1st Qu.:0.6240 1st Qu.: 9.953 1st Qu.: 6.539 1st Qu.:12.47
## Median :0.7520 Median :13.815 Median : 7.769 Median :13.29
## Mean   :0.7557 Mean   :15.153 Mean   : 7.444 Mean   :14.08
## 3rd Qu.:0.8780 3rd Qu.:18.990 3rd Qu.: 8.433 3rd Qu.:15.35
## Max.   :1.2080 Max.   :37.340 Max.   :10.097 Max.   :18.02
##      pH      Salinity      PAR
## Min.    :7.620 Min.    :16.44 Min.    : 9.768
## 1st Qu.:7.850 1st Qu.:29.62 1st Qu.:196.000
## Median :7.960 Median :34.10 Median :525.793
## Mean   :7.982 Mean   :30.49 Mean   :569.983
## 3rd Qu.:8.150 3rd Qu.:34.63 3rd Qu.:764.500
## Max.   :8.500 Max.   :34.90 Max.   :1921.000
```

9. Proponer hipótesis y realiza análisis estadístico de los datos, incluye evaluación de supuestos si corresponde

Hipótesis Nula: No existe una estructura de grupos separados por las variables categoricas estacionalidad y día.

Hipótesis alternativa: Existe una estructura de grupos separados por las variables categoricas estacionalidad y día.

habilita librerías

```
library(readxl)
library(ggplot2)
library(dplyr)
library(knitr)
library(pander)
library(psych) # Graficas de correlación
library(factoextra) # distancia euclidean

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3Wba
library(vegan) # Community Ecology Package: Ordination, Diversity and Dissimilarities

## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.6-4

library(dendextend) # extiende opciones de visualización

## Registered S3 method overwritten by 'dendextend':
##   method      from
##   rev.hclust  vegan
```

```
##
## -----
## Welcome to dendextend version 1.16.0
## Type citation('dendextend') for how to cite the package.
##
## Type browseVignettes(package = 'dendextend') for the package vignette.
## The github page is: https://github.com/talgalili/dendextend/
##
## Suggestions and bug-reports can be submitted at: https://github.com/talgalili/dendextend/issues
## You may ask questions at stackoverflow, use the r and dendextend tags:
##   https://stackoverflow.com/questions/tagged/dendextend
##
## To suppress this message use: suppressPackageStartupMessages(library(dendextend))
## -----

##
## Attaching package: 'dendextend'

## The following object is masked from 'package:permute':
##
##   shuffle

## The following object is masked from 'package:stats':
##
##   cutree
```

Importar datos proyecto.

```
datos_PCA <- read_excel("Datos_Proyecto.xlsx", sheet = 1)
summary(datos_PCA)
```

```
##      Seasons           time           Chla           Chlc
## Length:108      Length:108      Min.    :0.6730      Min.    :0.01000
## Class :character Class :character 1st Qu.:0.9888      1st Qu.:0.06775
## Mode  :character Mode  :character Median :1.1660      Median :0.09500
##                                     Mean  :1.1856      Mean  :0.09365
##                                     3rd Qu.:1.3813      3rd Qu.:0.10975
##                                     Max.   :1.9690      Max.   :0.22100
##
##      Car           PC           DPPH           Temperature
## Min.    :0.4190      Min.    : 4.600      Min.    : 3.792      Min.    :11.54
## 1st Qu.:0.6240      1st Qu.: 9.953      1st Qu.: 6.539      1st Qu.:12.47
## Median :0.7520      Median :13.815      Median : 7.769      Median :13.29
## Mean    :0.7557      Mean    :15.153      Mean    : 7.444      Mean    :14.08
## 3rd Qu.:0.8780      3rd Qu.:18.990      3rd Qu.: 8.433      3rd Qu.:15.35
## Max.    :1.2080      Max.    :37.340      Max.    :10.097      Max.    :18.02
##
##      pH           Salinity           PAR
## Min.    :7.620      Min.    :16.44      Min.    : 9.768
## 1st Qu.:7.850      1st Qu.:29.62      1st Qu.:196.000
## Median :7.960      Median :34.10      Median :525.793
## Mean    :7.982      Mean    :30.49      Mean    :569.983
## 3rd Qu.:8.150      3rd Qu.:34.63      3rd Qu.:764.500
## Max.    :8.500      Max.    :34.90      Max.    :1921.000
```

```
datos_PCA$Seasons <- as.factor(datos_PCA$Seasons)
datos_PCA$time <- as.factor(datos_PCA$time)
head(datos_PCA[,3:11]) %>% pander(caption = "Variables ecofisiologicas y ambientales en Lessonia spicata")
```

Table 4: Variables ecofisiológicas y ambientales en *Lessonia spicata*

Chla	Chlc	Car	PC	DPPH	Temperature	pH	Salinity	PAR
1.04	0.095	0.605	9.12	7.578	12.29	7.84	33.74	594
1.597	0.155	0.954	13.87	6.085	12.29	7.84	33.74	594
1.521	0.143	0.931	15.83	7.812	12.29	7.84	33.74	594
1.479	0.152	0.905	19.89	7.015	12.29	7.85	17.87	1563
1.139	0.108	0.694	19.81	7.201	12.29	7.85	17.87	1563
1.166	0.115	0.752	21.64	5.119	12.29	7.85	17.87	1563

```
str(datos_PCA)
```

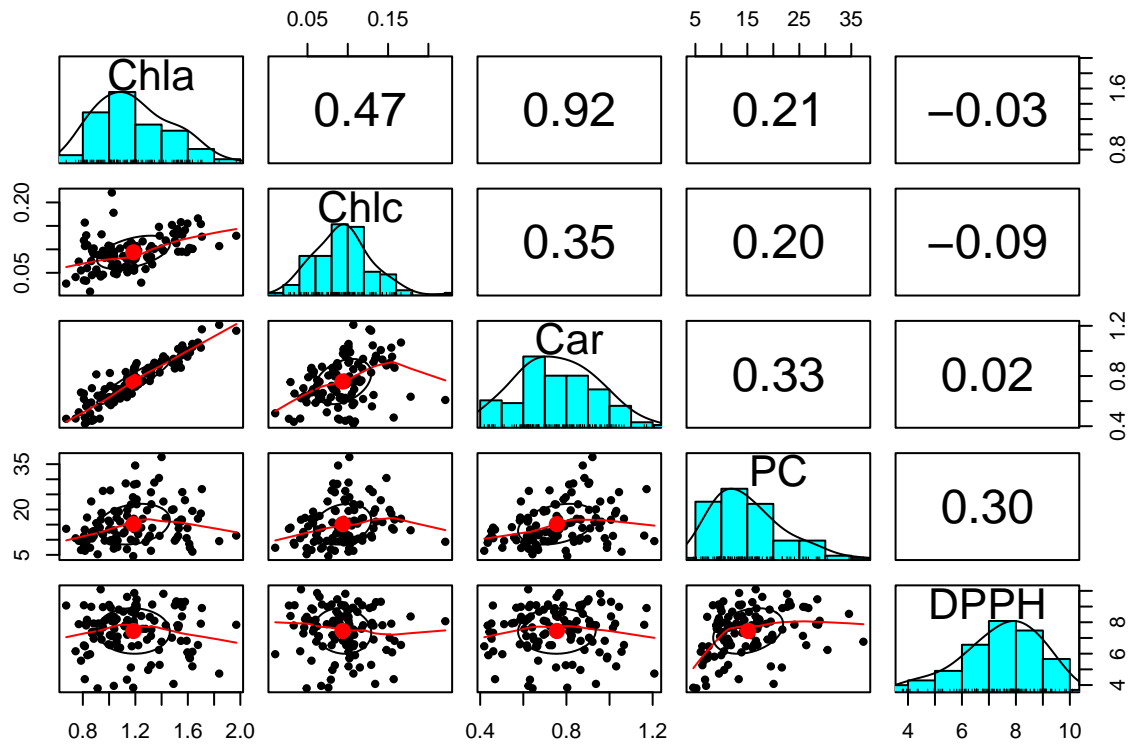
```
## tibble [108 x 11] (S3: tbl_df/tbl/data.frame)
## $ Seasons      : Factor w/ 4 levels "Autumn","Spring",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ time         : Factor w/ 3 levels "day 1","day 2",...: 1 1 1 1 1 1 1 1 1 2 ...
## $ Chla         : num [1:108] 1.04 1.6 1.52 1.48 1.14 ...
## $ Chlc         : num [1:108] 0.095 0.155 0.143 0.152 0.108 0.115 0.148 0.105 0.102 0.134 ...
## $ Car          : num [1:108] 0.605 0.954 0.931 0.905 0.694 0.752 0.884 0.659 0.62 0.937 ...
## $ PC           : num [1:108] 9.12 13.87 15.83 19.89 19.81 ...
## $ DPPH         : num [1:108] 7.58 6.08 7.81 7.02 7.2 ...
## $ Temperature: num [1:108] 12.3 12.3 12.3 12.3 12.3 ...
## $ pH           : num [1:108] 7.84 7.84 7.84 7.85 7.85 7.85 7.88 7.88 7.88 7.84 ...
## $ Salinity     : num [1:108] 33.7 33.7 33.7 17.9 17.9 ...
## $ PAR          : num [1:108] 594 594 594 1563 1563 ...
```

```
datos_PCA_mat <- as.matrix(datos_PCA[, -c(1:2)])
str(datos_PCA_mat)
```

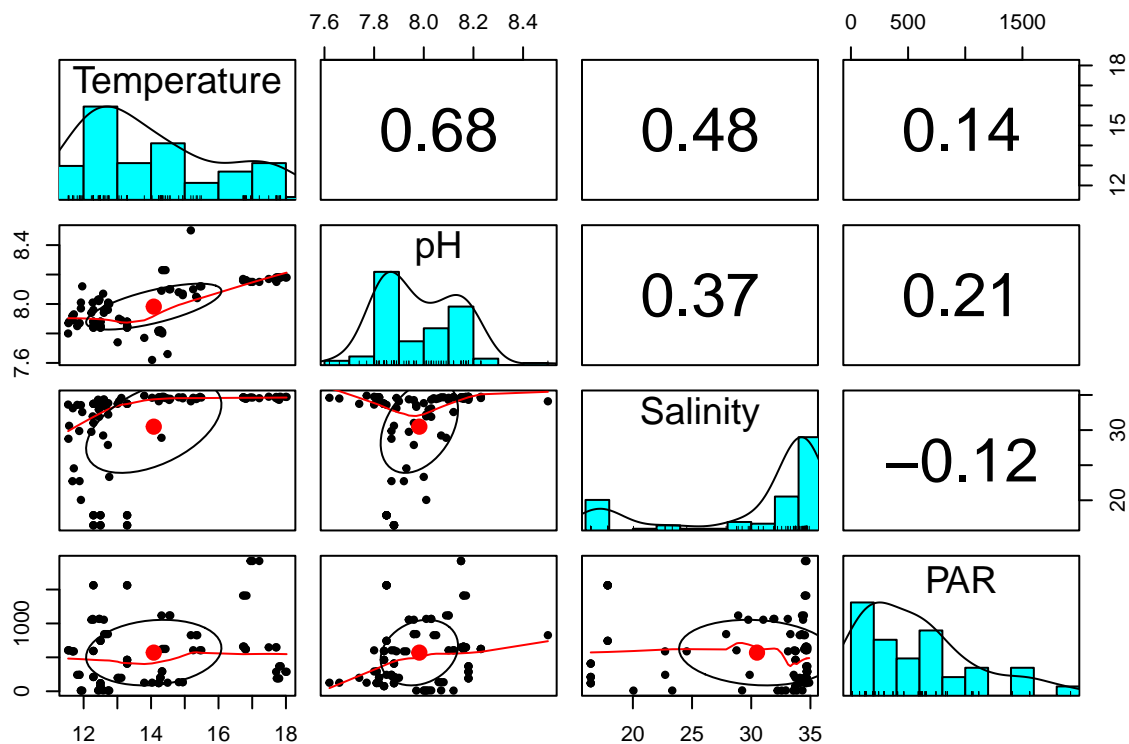
```
## num [1:108, 1:9] 1.04 1.6 1.52 1.48 1.14 ...
## - attr(*, "dimnames")=List of 2
## ..$ : NULL
## ..$ : chr [1:9] "Chla" "Chlc" "Car" "PC" ...
```

Correlación entre variables

```
pairs.panels(datos_PCA_mat[, 1:5], method = "pearson")
```



```
pairs.panels(datos_PCA_mat[,6:9], method = "pearson")
```



Realiza PCA

```
PCA_Lesso <- prcomp(datos_PCA_mat, scale = TRUE)
PCA_Lesso
```

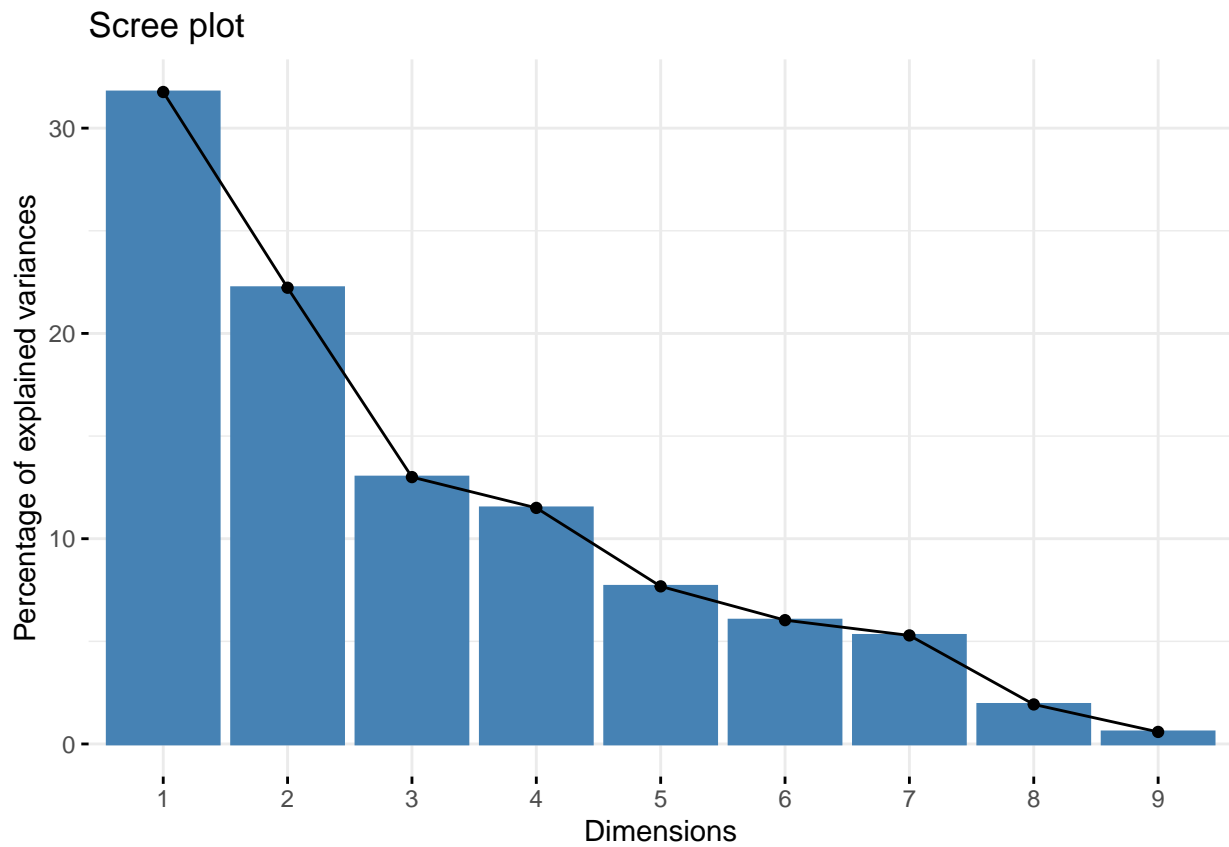
```
## Standard deviations (1, ..., p=9):
## [1] 1.6906902 1.4143173 1.0817007 1.0174003 0.8313044 0.7367822 0.6898727
## [8] 0.4164170 0.2298930
##
## Rotation (n x k) = (9 x 9):
##
##          PC1          PC2          PC3          PC4          PC5
## Chla      0.4568817 -0.34350662  0.22454424  0.12629866 -0.24706846
## Chlc      0.1751132 -0.51747885 -0.06730790  0.04046899  0.08370557
## Car       0.5000382 -0.24411934  0.17696923  0.15742553 -0.18111272
## PC        0.3194451 -0.03579808 -0.58459611  0.09918040  0.54291925
## DPPH      0.1367811  0.22149655 -0.67868393  0.10946599 -0.60014662
## Temperature 0.4805548  0.32225376  0.01808790  0.01615086  0.17659091
## pH        0.3229442  0.39571168  0.17996698 -0.28930770  0.29540096
## Salinity   0.1609386  0.48974551  0.27345642  0.27704648 -0.25120508
## PAR       0.1693368 -0.06321772 -0.06223532 -0.88039424 -0.25032258
##
##          PC6          PC7          PC8          PC9
## Chla      0.169407469 -0.11620759  0.09036154  0.7039032070
## Chlc     -0.815836745  0.05543266 -0.09303314 -0.1051605698
## Car       0.304511115 -0.01829578  0.17626415 -0.6908936552
## PC        0.177673023  0.35675969  0.27987993  0.1146503236
## DPPH     -0.132467634 -0.28410415  0.03181368 -0.0135873412
## Temperature 0.006091989 -0.01917040 -0.79563051 -0.0038316299
## pH       -0.265955836 -0.52941888  0.42981859  0.0031091494
## Salinity  -0.302870960  0.60840260  0.23442065  0.0526626327
## PAR       0.036995683  0.35254795 -0.00638049 -0.0006879714
```

Varianza explicada

```
get_eigenvalue(PCA_Lesso)
```

```
##          eigenvalue variance.percent cumulative.variance.percent
## Dim.1 2.85843341          31.760371          31.76037
## Dim.2 2.00029353          22.225484          53.98585
## Dim.3 1.17007637          13.000849          66.98670
## Dim.4 1.03510343          11.501149          78.48785
## Dim.5 0.69106702           7.678522          86.16638
## Dim.6 0.54284803           6.031645          92.19802
## Dim.7 0.47592430           5.288048          97.48607
## Dim.8 0.17340312           1.926701          99.41277
## Dim.9 0.05285079           0.587231         100.00000
```

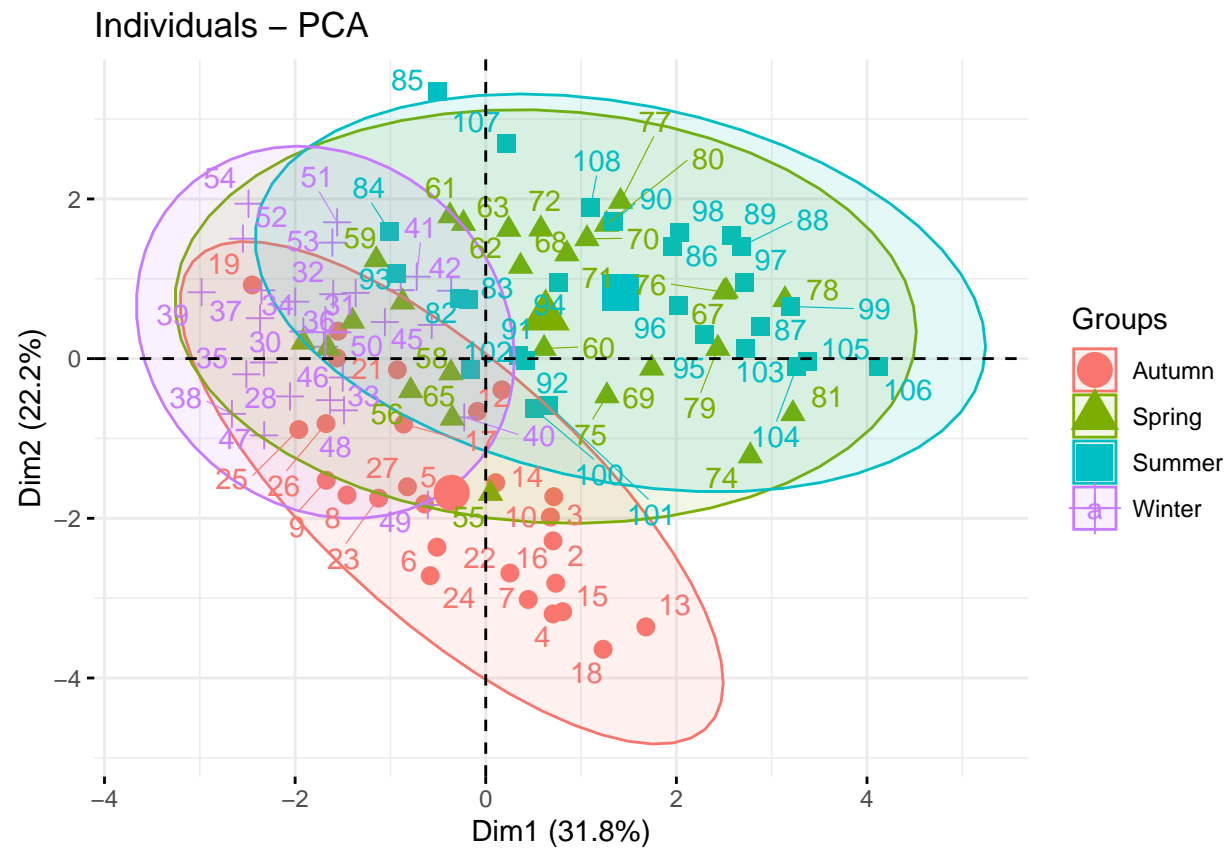
```
fviz_eig(PCA_Lesso)
```



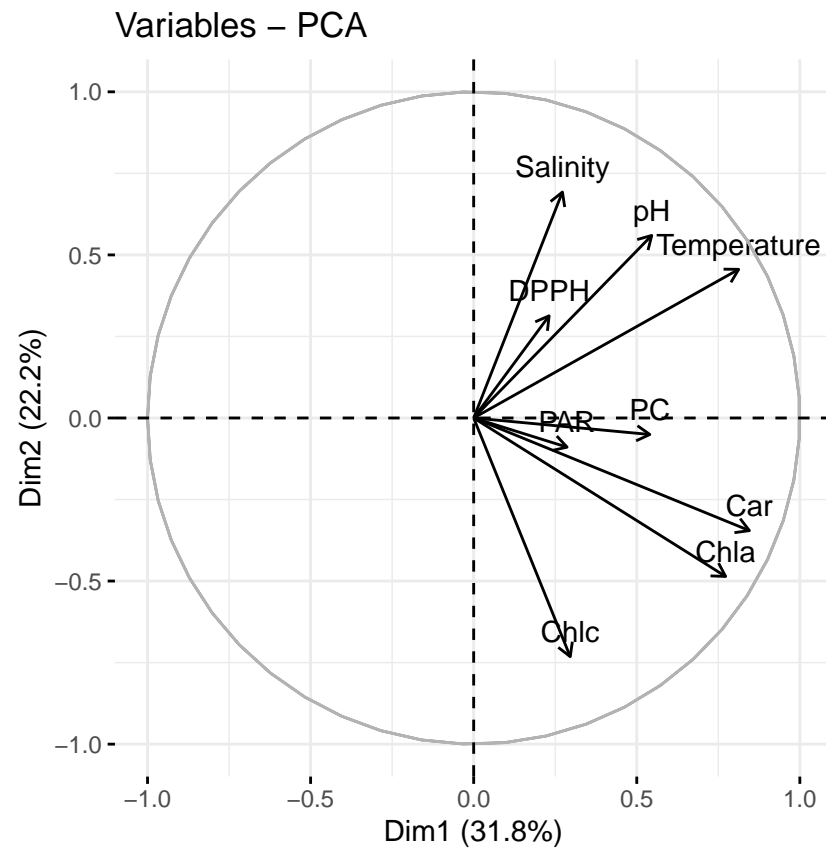
Grafica por sitio

```
fviz_pca_ind(PCA_Lesso,  
             repel = TRUE,  
             habillage = datos_PCA$Seasons,  
             addEllipses = TRUE,  
             pointsize = 3)
```

```
## Warning: ggrepel: 10 unlabeled data points (too many overlaps). Consider  
## increasing max.overlaps
```

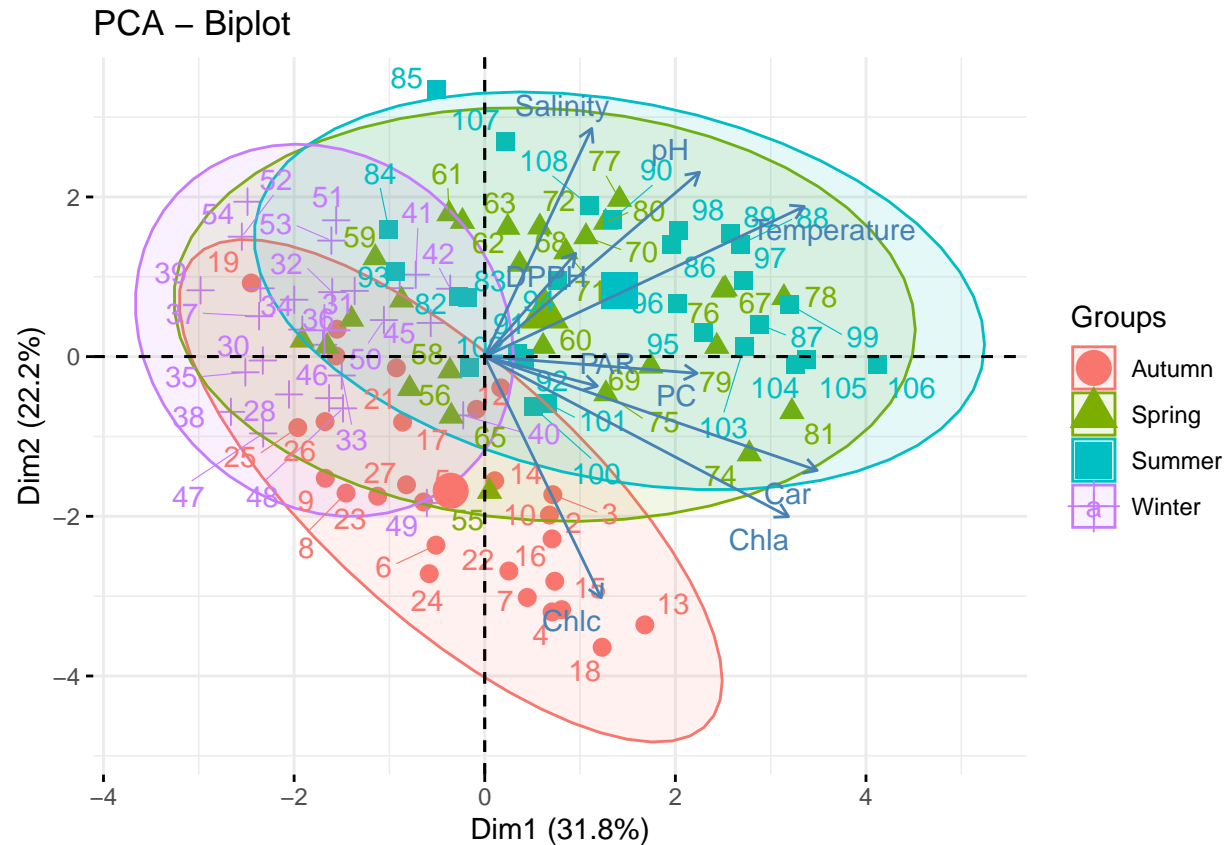


```
fviz_pca_var(PCA_Lesso)
```

```
fviz_pca_biplot(PCA_Lesso,
  repel = TRUE,
  habillage = datos_PCA$Seasons,
  addEllipses = TRUE,
  pointsize = 3)
```

```
## Warning: ggrepel: 10 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



Crea nuevas variables estandarizadas

```
val_estandarizado <- datos_PCA %>%
  select(Chla, Chlc, Car, PC, DPPH, Temperature, pH, Salinity, PAR) %>%
  mutate(Chla1 = (Chla - mean(Chla)) / sd(Chla), Chlc1 = (Chlc - mean(Chlc)) / sd(Chlc), Car1 = (Car - mean(Car)) / sd(Car),
    PC1 = (PC - mean(PC)) / sd(PC), DPPH1 = (DPPH - mean(DPPH)) / sd(DPPH), Temperature1 = (Temperature - mean(Temperature)) / sd(Temperature))
```

Calcula matriz de distancia

```
dist_euclidea <- dist(val_estandarizado[10:18]) #distancia euclidiana
```

Realiza PERMANOVA

```
permanova <- adonis2(dist_euclidea ~ Seasons:time, method = "bray", data=datos_PCA, permutations=999)
permanova %>% pander()
```

Table 5: Permutation test for adonis under reduced model

	Df	SumOfSqs	R2	F	Pr(>F)
Seasons:time	11	419	0.4351	6.722	0.001
Residual	96	544	0.5649	NA	NA
Total	107	963	1	NA	NA

```
dist_euclidea <- stats::dist(val_estandarizado[10:18], method = "euclidean")
```

10. Presenta, interpreta resultados y realiza conclusión

** Los resultados de los análisis de componentes principales, indican que la variable con mayor impacto sobre el componente principal 1 es la variable respuesta Car o Carotenos. Sin embargo para el componente principal 2 fueron las variables Chlc o Clorofila c y Salinity o Salinidad.

** Se observa que los 3 primeros componentes principales acumulan el 66.98 % de la variación total del análisis, de acuerdo a la varianza explicada.

** El gráfico de las variables PCA, se observan de manera exploratoria como se relacionan estas, con los componentes principales de manera gráfica, observándose que las flechas mas largas fueron para el componente principal 1, es decir para la variable que tiene mayor impacto o mayor inersia - peso es la variable Car o Carotenos y luego Chla o Clorofila a, pero al tener un angulo muy pequeño entre ellas, se denota tambien el alto grado de correlación que existe entre ambas variables. Así también, para el componente principal 2, la variable con mayor peso fue la varaiaible Salinity p Salidad, seguida por Chlc o Clorofila c.

** Finalmente el grafico Biplot con las variables, muestra una clara relación entre las variables respuestas para cada una de los datos obtenidos en las distintas estaciones del año y sus días de ciclos diarios respectivos, que se han compararon en los análisis. Observándose los centroides (muetsra con mayor tamaño) o muestras que representan los valores medios de ese conjunto de datos para cada estación del año, dentro de la elipse que se demarca en las 4 estaciones. Demostrando que para Winter o invierno la elipse de color morado y simbolos de cruces, se separan de los otros grupos a la izquierda del gráfico, así como Autumn o otoño, de color rosado con circulos, que se observa definido hacia abajo y el solapamiento que se observa con Spring o primavera y Summer o verano en verde y celeste con triangulos y cuadrados, respectivamnete, hacia arriba del gráfico.

** De acuerdo a Anderson et al. (2008) los datos del permanova han sido estandarizados. ** Referencias: Anderson, M. J., Gorley, R. N., and Clarke, K. R. (2008). PERMANOVA+ for PRIMER: Guide to software and statistical methods (Plymouth, UK: PRIMER-E).

** El análisis del Permanova, logra establecer que existen separación de grupos, ya que el valor del estadístico fue significativo para el anidado de Seasons o estaciones y Time o tiempo. Lo que permite concluir que se rechaza la “Hipótesis Nula” que indica que “No existe una estructura de grupos separados por las variables categoricas estacionalidad y día” y se acepta la “Hipótesis alternativa” la cual indica que “si, existe una estructura de grupos separados por las variables categoricas estacionalidad y día”.