

Estudios de Informática, Multimedia y Telecomunicaciones

Visualización de datos : PEC2 - Técnicas de visualización de datos

Autor: Andres Felipe Pulgarin Rosero

Abril 2022

- 1 Introducción
- 2 Gráfico Small Multiple
- 3 Gráfico Convex hull
- 4 Gráfico Pyramid Charts

1 Introducción

Esta Prueba de Evaluación Continuada utilizaremos tres datasets(penguins,Iris y Gapminder) que vienen por defecto en el propio software de R con ellos se realizaran las primeras dos graficas.

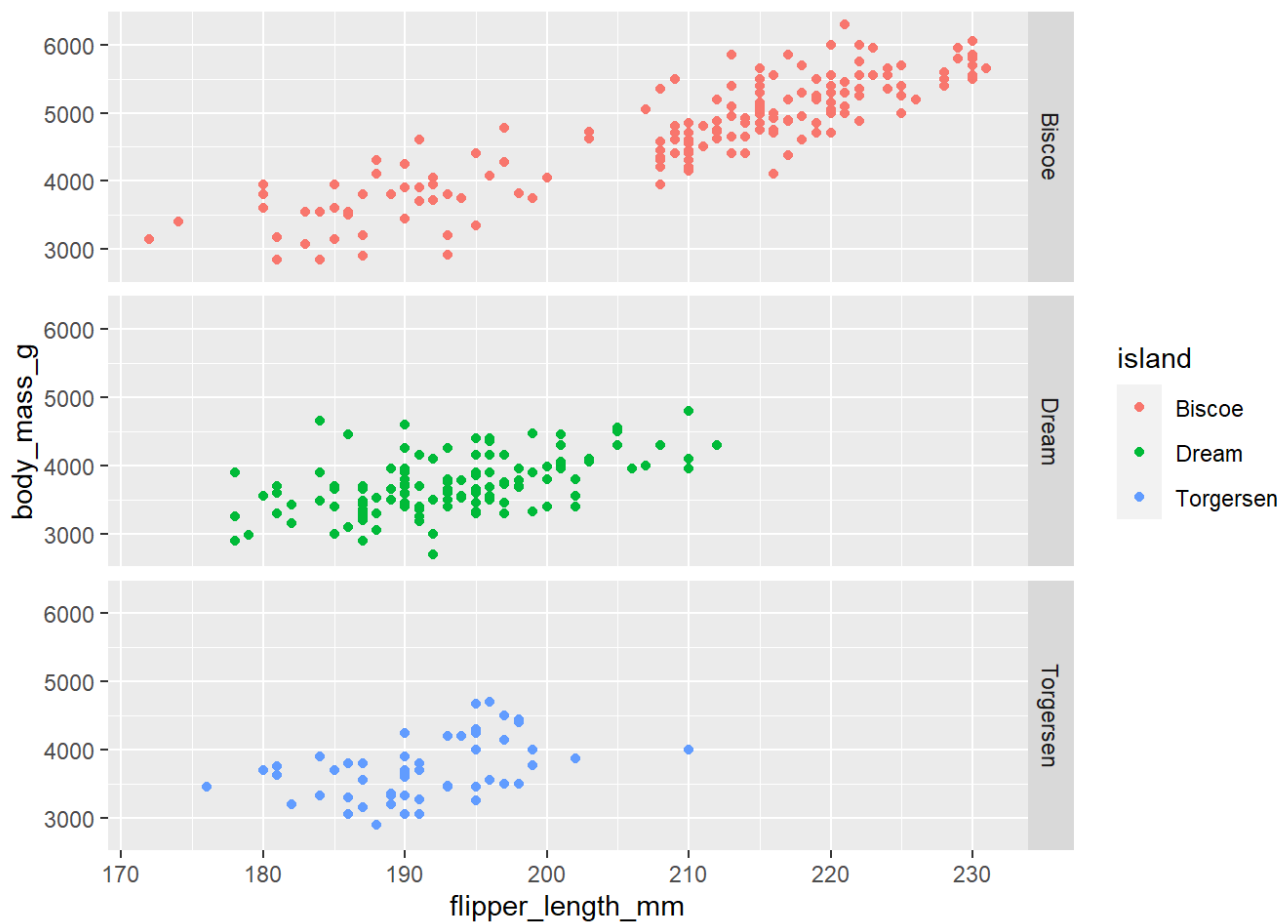
2 Gráfico Small Multiple

A continuación vamos a ver un ejemplo de cómo utilizar un gráfico multiple Small Multiple. Para ello usaremos el data set **penguins** contenido en el paquete R **palmerpenguins**. Esta base de datos se encuentra descrita en <https://cran.r-project.org/web/packages/palmerpenguins/index.html> (<https://cran.r-project.org/web/packages/palmerpenguins/index.html>) y contiene mediciones de tamaño, observaciones de puestas y proporciones de isótopos sanguíneos de tres especies de pingüinos observadas en tres islas del archipiélago Palmer, en la Antártida, durante un período de estudio de tres años.

Este dataset está previamente trabajado para que los datos estén limpios y sin errores. De no ser así antes de nada deberíamos buscar errores, valores nulos u *outliers*.

```
## 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 4
2 ...
## $ 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 4
250 ...
## $ 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 ...
```

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



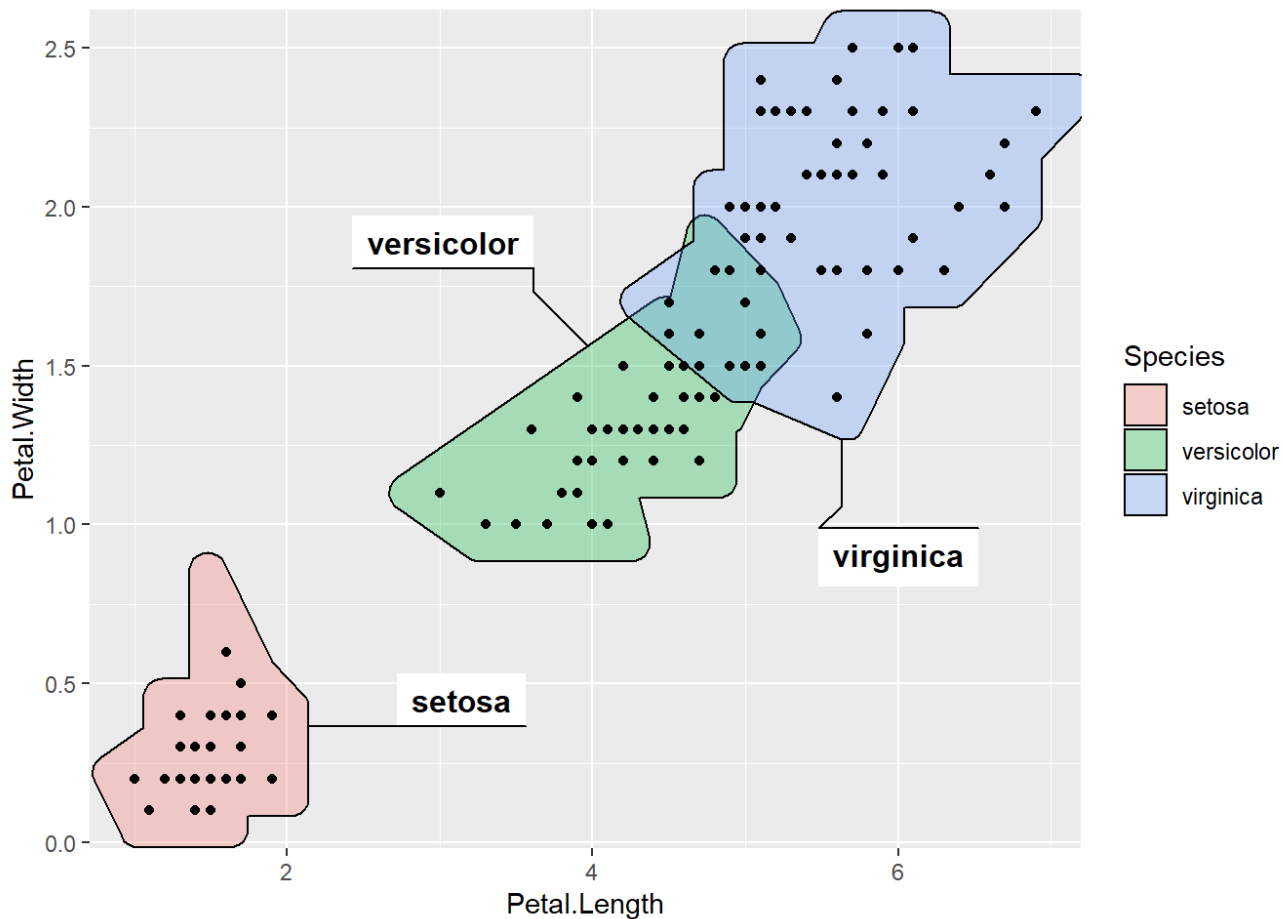
3 Gráfico Convex hull

A continuación vamos a ver un ejemplo de cómo utilizar un gráfico Convex hull. Para ello usaremos el data set

iris. Esta base de datos ha sido presentado por Ronald Fisher en su artículo de 1936 El uso de medidas múltiples en problemas taxonómicos, contiene tres especies de plantas (setosa, virginica, versicolor) y cuatro características medidas para cada muestra. Estos cuantifican la variación morfológica de la flor de iris en sus tres especies, todas las medidas dadas en centímetros.

Este dataset está previamente trabajado para que los datos estén limpios y sin errores. De no ser así antes de nada deberíamos buscar errores, valores nulos u *outliers*.

```
## 'data.frame': 150 obs. of 5 variables:  
## $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...  
## $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...  
## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...  
## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...  
## $ Species : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1  
1 1 ...
```



4 Gráfico Pyramid Charts

A continuación vamos a ver un ejemplo de cómo utilizar un gráfico Pyramid Chart, para ello utilizaremos el dataset de gapminder

```
## tibble [1,704 x 6] (S3: tbl_df/tbl/data.frame)
## $ country : Factor w/ 142 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ continent: Factor w/ 5 levels "Africa","Americas",...: 3 3 3 3 3 3 3 3 3 3 3
## $ year      : int [1:1704] 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997
## $ lifeExp   : num [1:1704] 28.8 30.3 32 34 36.1 ...
## $ pop       : int [1:1704] 8425333 9240934 10267083 11537966 13079460 14880372
## $ gdpPercap: num [1:1704] 779 821 853 836 740 ...
```

