# 01 Que es un DF

Téllez Gerardo Rubén

13/6/2021

## Que es

Consisten en bases de datos de una sola tabla donde notan distintos tipos de datos, siempre serán tablas de doble entrada donde se pueden ver variables, cada fila representa la observación de dichas variables para un mismo individuo.

Un DF es simialar a una matríz, pero con la capacidad de almacenar distintos tipos de datos.

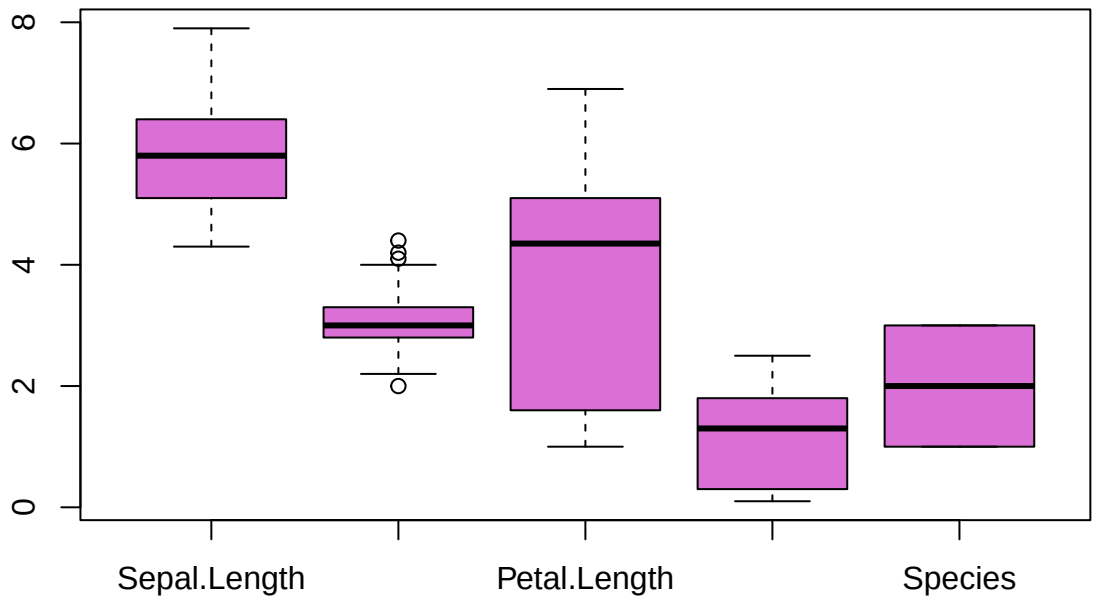
Es posible ver lo dataframes prefeinidos en R.

**data()**: Permite ver los data sets cargados de serie, o los cargados por ti mismo. Pero no se observan en el lm(). Puede recibir el nombre de los paquetes para ver los DF que vienen con los paquetes

**data(package = .packages(all.avaliable = TRUE))**: Permite ver todos los DF por defecto de los paquetes.

*iris*: Data set creado por Fisher en su trabajo *The use of multiple measurements in taxonomic problems*, conjunto multivarible de varias características de *Iris spp*

```
boxplot(iris,
        col = "orchid")
```




Cuando se usa un DF ajeno, es recomendable guardarlo en una variable, para no fastidiar el original.

## Data frame de Iris

```
DF_iris = iris
DF_iris
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
## 1            5.1         3.5          1.4         0.2     setosa
## 2            4.9         3.0          1.4         0.2     setosa
## 3            4.7         3.2          1.3         0.2     setosa
## 4            4.6         3.1          1.5         0.2     setosa
## 5            5.0         3.6          1.4         0.2     setosa
## 6            5.4         3.9          1.7         0.4     setosa
## 7            4.6         3.4          1.4         0.3     setosa
## 8            5.0         3.4          1.5         0.2     setosa
## 9            4.4         2.9          1.4         0.2     setosa
## 10           4.9         3.1          1.5         0.1     setosa
## 11           5.4         3.7          1.5         0.2     setosa
## 12           4.8         3.4          1.6         0.2     setosa
## 13           4.8         3.0          1.4         0.1     setosa
## 14           4.3         3.0          1.1         0.1     setosa
## 15           5.8         4.0          1.2         0.2     setosa
```

```
## 16          5.7         4.4          1.5         0.4     setosa
## 17          5.4         3.9          1.3         0.4     setosa
## 18          5.1         3.5          1.4         0.3     setosa
## 19          5.7         3.8          1.7         0.3     setosa
## 20          5.1         3.8          1.5         0.3     setosa
## 21          5.4         3.4          1.7         0.2     setosa
## 22          5.1         3.7          1.5         0.4     setosa
## 23          4.6         3.6          1.0         0.2     setosa
## 24          5.1         3.3          1.7         0.5     setosa
## 25          4.8         3.4          1.9         0.2     setosa
## 26          5.0         3.0          1.6         0.2     setosa
## 27          5.0         3.4          1.6         0.4     setosa
## 28          5.2         3.5          1.5         0.2     setosa
## 29          5.2         3.4          1.4         0.2     setosa
## 30          4.7         3.2          1.6         0.2     setosa
## 31          4.8         3.1          1.6         0.2     setosa
## 32          5.4         3.4          1.5         0.4     setosa
## 33          5.2         4.1          1.5         0.1     setosa
## 34          5.5         4.2          1.4         0.2     setosa
## 35          4.9         3.1          1.5         0.2     setosa
## 36          5.0         3.2          1.2         0.2     setosa
## 37          5.5         3.5          1.3         0.2     setosa
## 38          4.9         3.6          1.4         0.1     setosa
## 39          4.4         3.0          1.3         0.2     setosa
## 40          5.1         3.4          1.5         0.2     setosa
## 41          5.0         3.5          1.3         0.3     setosa
## 42          4.5         2.3          1.3         0.3     setosa
## 43          4.4         3.2          1.3         0.2     setosa
## 44          5.0         3.5          1.6         0.6     setosa
## 45          5.1         3.8          1.9         0.4     setosa
## 46          4.8         3.0          1.4         0.3     setosa
## 47          5.1         3.8          1.6         0.2     setosa
## 48          4.6         3.2          1.4         0.2     setosa
## 49          5.3         3.7          1.5         0.2     setosa
## 50          5.0         3.3          1.4         0.2     setosa
## 51          7.0         3.2          4.7         1.4 versicolor
## 52          6.4         3.2          4.5         1.5 versicolor
## 53          6.9         3.1          4.9         1.5 versicolor
## 54          5.5         2.3          4.0         1.3 versicolor
## 55          6.5         2.8          4.6         1.5 versicolor
## 56          5.7         2.8          4.5         1.3 versicolor
## 57          6.3         3.3          4.7         1.6 versicolor
## 58          4.9         2.4          3.3         1.0 versicolor
## 59          6.6         2.9          4.6         1.3 versicolor
## 60          5.2         2.7          3.9         1.4 versicolor
## 61          5.0         2.0          3.5         1.0 versicolor
## 62          5.9         3.0          4.2         1.5 versicolor
## 63          6.0         2.2          4.0         1.0 versicolor
## 64          6.1         2.9          4.7         1.4 versicolor
## 65          5.6         2.9          3.6         1.3 versicolor
## 66          6.7         3.1          4.4         1.4 versicolor
## 67          5.6         3.0          4.5         1.5 versicolor
## 68          5.8         2.7          4.1         1.0 versicolor
## 69          6.2         2.2          4.5         1.5 versicolor
```

```
## 70          5.6         2.5          3.9         1.1 versicolor
## 71          5.9         3.2          4.8         1.8 versicolor
## 72          6.1         2.8          4.0         1.3 versicolor
## 73          6.3         2.5          4.9         1.5 versicolor
## 74          6.1         2.8          4.7         1.2 versicolor
## 75          6.4         2.9          4.3         1.3 versicolor
## 76          6.6         3.0          4.4         1.4 versicolor
## 77          6.8         2.8          4.8         1.4 versicolor
## 78          6.7         3.0          5.0         1.7 versicolor
## 79          6.0         2.9          4.5         1.5 versicolor
## 80          5.7         2.6          3.5         1.0 versicolor
## 81          5.5         2.4          3.8         1.1 versicolor
## 82          5.5         2.4          3.7         1.0 versicolor
## 83          5.8         2.7          3.9         1.2 versicolor
## 84          6.0         2.7          5.1         1.6 versicolor
## 85          5.4         3.0          4.5         1.5 versicolor
## 86          6.0         3.4          4.5         1.6 versicolor
## 87          6.7         3.1          4.7         1.5 versicolor
## 88          6.3         2.3          4.4         1.3 versicolor
## 89          5.6         3.0          4.1         1.3 versicolor
## 90          5.5         2.5          4.0         1.3 versicolor
## 91          5.5         2.6          4.4         1.2 versicolor
## 92          6.1         3.0          4.6         1.4 versicolor
## 93          5.8         2.6          4.0         1.2 versicolor
## 94          5.0         2.3          3.3         1.0 versicolor
## 95          5.6         2.7          4.2         1.3 versicolor
## 96          5.7         3.0          4.2         1.2 versicolor
## 97          5.7         2.9          4.2         1.3 versicolor
## 98          6.2         2.9          4.3         1.3 versicolor
## 99          5.1         2.5          3.0         1.1 versicolor
## 100         5.7         2.8          4.1         1.3 versicolor
## 101         6.3         3.3          6.0         2.5  virginica
## 102         5.8         2.7          5.1         1.9  virginica
## 103         7.1         3.0          5.9         2.1  virginica
## 104         6.3         2.9          5.6         1.8  virginica
## 105         6.5         3.0          5.8         2.2  virginica
## 106         7.6         3.0          6.6         2.1  virginica
## 107         4.9         2.5          4.5         1.7  virginica
## 108         7.3         2.9          6.3         1.8  virginica
## 109         6.7         2.5          5.8         1.8  virginica
## 110         7.2         3.6          6.1         2.5  virginica
## 111         6.5         3.2          5.1         2.0  virginica
## 112         6.4         2.7          5.3         1.9  virginica
## 113         6.8         3.0          5.5         2.1  virginica
## 114         5.7         2.5          5.0         2.0  virginica
## 115         5.8         2.8          5.1         2.4  virginica
## 116         6.4         3.2          5.3         2.3  virginica
## 117         6.5         3.0          5.5         1.8  virginica
## 118         7.7         3.8          6.7         2.2  virginica
## 119         7.7         2.6          6.9         2.3  virginica
## 120         6.0         2.2          5.0         1.5  virginica
## 121         6.9         3.2          5.7         2.3  virginica
## 122         5.6         2.8          4.9         2.0  virginica
## 123         7.7         2.8          6.7         2.0  virginica
```

```
## 124          6.3         2.7          4.9         1.8  virginica
## 125          6.7         3.3          5.7         2.1  virginica
## 126          7.2         3.2          6.0         1.8  virginica
## 127          6.2         2.8          4.8         1.8  virginica
## 128          6.1         3.0          4.9         1.8  virginica
## 129          6.4         2.8          5.6         2.1  virginica
## 130          7.2         3.0          5.8         1.6  virginica
## 131          7.4         2.8          6.1         1.9  virginica
## 132          7.9         3.8          6.4         2.0  virginica
## 133          6.4         2.8          5.6         2.2  virginica
## 134          6.3         2.8          5.1         1.5  virginica
## 135          6.1         2.6          5.6         1.4  virginica
## 136          7.7         3.0          6.1         2.3  virginica
## 137          6.3         3.4          5.6         2.4  virginica
## 138          6.4         3.1          5.5         1.8  virginica
## 139          6.0         3.0          4.8         1.8  virginica
## 140          6.9         3.1          5.4         2.1  virginica
## 141          6.7         3.1          5.6         2.4  virginica
## 142          6.9         3.1          5.1         2.3  virginica
## 143          5.8         2.7          5.1         1.9  virginica
## 144          6.8         3.2          5.9         2.3  virginica
## 145          6.7         3.3          5.7         2.5  virginica
## 146          6.7         3.0          5.2         2.3  virginica
## 147          6.3         2.5          5.0         1.9  virginica
## 148          6.5         3.0          5.2         2.0  virginica
## 149          6.2         3.4          5.4         2.3  virginica
## 150          5.9         3.0          5.1         1.8  virginica
```

Para controlar cuánto se imprime o se usa:

**head(df, n)**: para mostrar las primeras $n$ filas del DF. Por defecto n = 6.

**tail(df, n)**: para mostrar las $n$ últimas filas de DF. Por defecto n = 6.

**str(df)**: para conocer la estructura global del DF

**names(df)**: para producir un vector con os nombres de las columnas.

```
head(DF_iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```
tail(DF_iris)
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 145          6.7         3.3          5.7         2.5 virginica
## 146          6.7         3.0          5.2         2.3 virginica
## 147          6.3         2.5          5.0         1.9 virginica
## 148          6.5         3.0          5.2         2.0 virginica
```

```
## 149          6.2         3.4          5.4         2.3 virginica
## 150          5.9         3.0          5.1         1.8 virginica
```

```
str(DF_iris)
```

```
## '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 ...
```

```
names(DF_iris) #El último es un factor de 3 variables.
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

### Ejemplo con DF Orange

```
df_orange = Orange

str(df_orange)
```

```
## Classes 'nfnGroupedData', 'nfGroupedData', 'groupedData' and 'data.frame':   35 obs. of  3 variables
##  $ Tree         : Ord.factor w/ 5 levels "3"<"1"<"5"<"2"<..: 2 2 2 2 2 2 2 4 4 4 ...
##  $ age          : num  118 484 664 1004 1231 ...
##  $ circumference: num  30 58 87 115 120 142 145 33 69 111 ...
##  - attr(*, "formula")=Class 'formula'  language circumference ~ age | Tree
##   .. ..- attr(*, ".Environment")=<environment: R_EmptyEnv>
##  - attr(*, "labels")=List of 2
##   ..$ x: chr "Time since December 31, 1968"
##   ..$ y: chr "Trunk circumference"
##  - attr(*, "units")=List of 2
##   ..$ x: chr "(days)"
##   ..$ y: chr "(mm)"
```

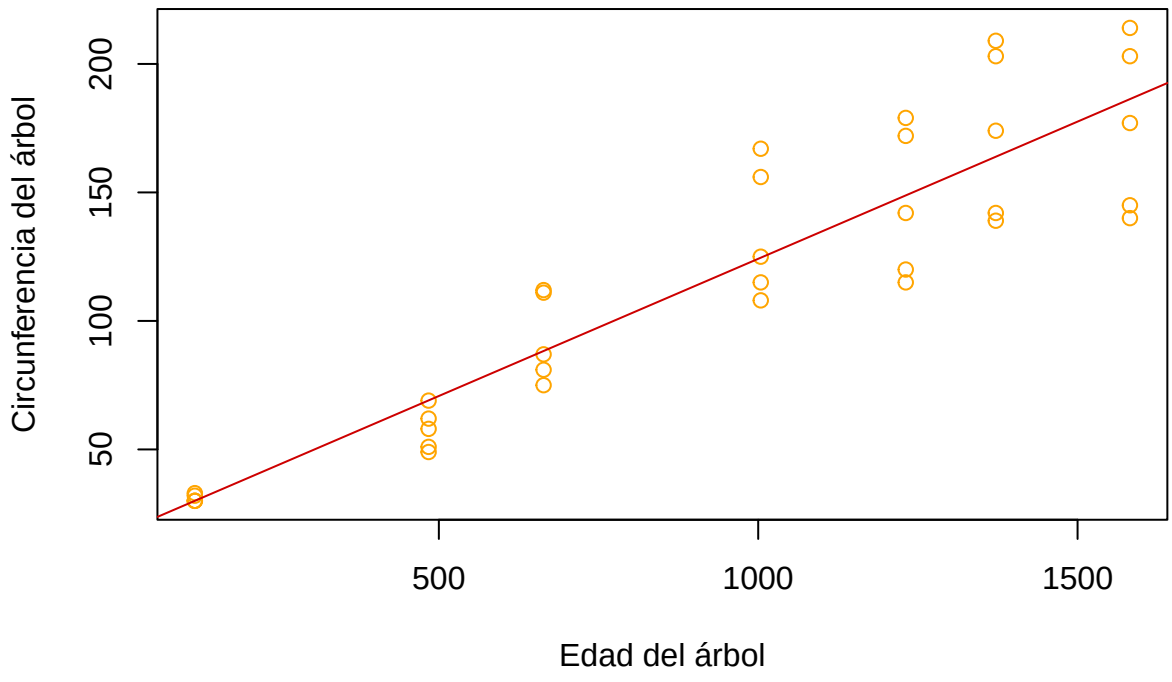
```
aov_or = aov(df_orange$circumference ~ df_orange$age)

summary(aov_or)
```

```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## df_orange$age  1  93772   93772   166.4 1.93e-14 ***
## Residuals     33  18595     563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Naov = aov(df_orange$age ~ df_orange$Tree)
TukeyHSD(Naov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = df_orange$age ~ df_orange$Tree)
##
## $`df_orange$Tree`
##             diff       lwr      upr p adj
## 1-3 0.000000e+00 -811.8566 811.8566     1
## 5-3 0.000000e+00 -811.8566 811.8566     1
## 2-3 0.000000e+00 -811.8566 811.8566     1
## 4-3 1.136868e-13 -811.8566 811.8566     1
## 5-1 0.000000e+00 -811.8566 811.8566     1
## 2-1 0.000000e+00 -811.8566 811.8566     1
## 4-1 1.136868e-13 -811.8566 811.8566     1
## 2-5 0.000000e+00 -811.8566 811.8566     1
## 4-5 1.136868e-13 -811.8566 811.8566     1
## 4-2 1.136868e-13 -811.8566 811.8566     1
```

```
plot(df_orange$circumference ~ df_orange$age,
     xlab = "Edad del árbol", ylab = "Circunferencia del árbol", col="orange")
abline(lm(df_orange$circumference ~ df_orange$age), col="red3")
```




```
cor.test(df_orange$circumference, df_orange$age, method = "pearson")
```

```
##
```

```
##  Pearson's product-moment correlation
##
## data:  df_orange$circumference and df_orange$age
## t = 12.9, df = 33, p-value = 1.931e-14
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8342364 0.9557955
## sample estimates:
##       cor
## 0.9135189
```

## Funciones para manipular

**rownames(df)**: Para producir un vector con los identificadores de las filas.

**columnames(df)**: Para producir un vector con los identificadores de las columnas.

**dimnames(df)**: Para producir una lista formada por dos vectores, el de identificadores de filas y el de columnas.

**nrow(df)**: Consulta el número de filas de un DF.

**ncol(df)**: Consulta el numero de columnas del DF.

**dim(df)**: Produce un vector con el numero de filas y el de columnas.

```
colnames(DF_iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

```
rownames(DF_iris)
```

```
##   [1] "1"   "2"   "3"   "4"   "5"   "6"   "7"   "8"   "9"   "10"  "11"  "12"
##  [13] "13"  "14"  "15"  "16"  "17"  "18"  "19"  "20"  "21"  "22"  "23"  "24"
##  [25] "25"  "26"  "27"  "28"  "29"  "30"  "31"  "32"  "33"  "34"  "35"  "36"
##  [37] "37"  "38"  "39"  "40"  "41"  "42"  "43"  "44"  "45"  "46"  "47"  "48"
##  [49] "49"  "50"  "51"  "52"  "53"  "54"  "55"  "56"  "57"  "58"  "59"  "60"
##  [61] "61"  "62"  "63"  "64"  "65"  "66"  "67"  "68"  "69"  "70"  "71"  "72"
##  [73] "73"  "74"  "75"  "76"  "77"  "78"  "79"  "80"  "81"  "82"  "83"  "84"
##  [85] "85"  "86"  "87"  "88"  "89"  "90"  "91"  "92"  "93"  "94"  "95"  "96"
##  [97] "97"  "98"  "99"  "100" "101" "102" "103" "104" "105" "106" "107" "108"
## [109] "109" "110" "111" "112" "113" "114" "115" "116" "117" "118" "119" "120"
## [121] "121" "122" "123" "124" "125" "126" "127" "128" "129" "130" "131" "132"
## [133] "133" "134" "135" "136" "137" "138" "139" "140" "141" "142" "143" "144"
## [145] "145" "146" "147" "148" "149" "150"
```

```
dimnames(DF_iris)
```

```
## [[1]]
##   [1] "1"   "2"   "3"   "4"   "5"   "6"   "7"   "8"   "9"   "10"  "11"  "12"
##  [13] "13"  "14"  "15"  "16"  "17"  "18"  "19"  "20"  "21"  "22"  "23"  "24"
##  [25] "25"  "26"  "27"  "28"  "29"  "30"  "31"  "32"  "33"  "34"  "35"  "36"
##  [37] "37"  "38"  "39"  "40"  "41"  "42"  "43"  "44"  "45"  "46"  "47"  "48"
```

```
##  [49] "49"  "50"  "51"  "52"  "53"  "54"  "55"  "56"  "57"  "58"  "59"  "60"
##  [61] "61"  "62"  "63"  "64"  "65"  "66"  "67"  "68"  "69"  "70"  "71"  "72"
##  [73] "73"  "74"  "75"  "76"  "77"  "78"  "79"  "80"  "81"  "82"  "83"  "84"
##  [85] "85"  "86"  "87"  "88"  "89"  "90"  "91"  "92"  "93"  "94"  "95"  "96"
##  [97] "97"  "98"  "99"  "100" "101" "102" "103" "104" "105" "106" "107" "108"
## [109] "109" "110" "111" "112" "113" "114" "115" "116" "117" "118" "119" "120"
## [121] "121" "122" "123" "124" "125" "126" "127" "128" "129" "130" "131" "132"
## [133] "133" "134" "135" "136" "137" "138" "139" "140" "141" "142" "143" "144"
## [145] "145" "146" "147" "148" "149" "150"
##
## [[2]]
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

```
dim(DF_iris)
```

```
## [1] 150   5
```

Puede manipularse similar a las listas

```
DF_iris$Species[1:10]
```

```
##  [1] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## Levels: setosa versicolor virginica
```

### Para obtener un subdataframe

```
# Primero filas, después columnas
Iris_recortado = DF_iris[1:10, 3:5]
Iris_recortado
```

```
##    Petal.Length Petal.Width Species
## 1           1.4         0.2  setosa
## 2           1.4         0.2  setosa
## 3           1.3         0.2  setosa
## 4           1.5         0.2  setosa
## 5           1.4         0.2  setosa
## 6           1.7         0.4  setosa
## 7           1.4         0.3  setosa
## 8           1.5         0.2  setosa
## 9           1.4         0.2  setosa
## 10          1.5         0.1  setosa
```

### Filtrar por condicionales booleanas

```
DF_iris[DF_iris$Species == "setosa" | DF_iris$Species == "versicolor", 3:5]
```

```
##     Petal.Length Petal.Width    Species
## 1            1.4         0.2     setosa
## 2            1.4         0.2     setosa
## 3            1.3         0.2     setosa
## 4            1.5         0.2     setosa
## 5            1.4         0.2     setosa
## 6            1.7         0.4     setosa
## 7            1.4         0.3     setosa
## 8            1.5         0.2     setosa
## 9            1.4         0.2     setosa
## 10           1.5         0.1     setosa
## 11           1.5         0.2     setosa
## 12           1.6         0.2     setosa
## 13           1.4         0.1     setosa
## 14           1.1         0.1     setosa
## 15           1.2         0.2     setosa
## 16           1.5         0.4     setosa
## 17           1.3         0.4     setosa
## 18           1.4         0.3     setosa
## 19           1.7         0.3     setosa
## 20           1.5         0.3     setosa
## 21           1.7         0.2     setosa
## 22           1.5         0.4     setosa
## 23           1.0         0.2     setosa
## 24           1.7         0.5     setosa
## 25           1.9         0.2     setosa
## 26           1.6         0.2     setosa
## 27           1.6         0.4     setosa
## 28           1.5         0.2     setosa
## 29           1.4         0.2     setosa
## 30           1.6         0.2     setosa
## 31           1.6         0.2     setosa
## 32           1.5         0.4     setosa
## 33           1.5         0.1     setosa
## 34           1.4         0.2     setosa
## 35           1.5         0.2     setosa
## 36           1.2         0.2     setosa
## 37           1.3         0.2     setosa
## 38           1.4         0.1     setosa
## 39           1.3         0.2     setosa
## 40           1.5         0.2     setosa
## 41           1.3         0.3     setosa
## 42           1.3         0.3     setosa
## 43           1.3         0.2     setosa
## 44           1.6         0.6     setosa
## 45           1.9         0.4     setosa
## 46           1.4         0.3     setosa
## 47           1.6         0.2     setosa
## 48           1.4         0.2     setosa
## 49           1.5         0.2     setosa
## 50           1.4         0.2     setosa
## 51           4.7         1.4 versicolor
## 52           4.5         1.5 versicolor
## 53           4.9         1.5 versicolor
```

```
## 54          4.0         1.3 versicolor
## 55          4.6         1.5 versicolor
## 56          4.5         1.3 versicolor
## 57          4.7         1.6 versicolor
## 58          3.3         1.0 versicolor
## 59          4.6         1.3 versicolor
## 60          3.9         1.4 versicolor
## 61          3.5         1.0 versicolor
## 62          4.2         1.5 versicolor
## 63          4.0         1.0 versicolor
## 64          4.7         1.4 versicolor
## 65          3.6         1.3 versicolor
## 66          4.4         1.4 versicolor
## 67          4.5         1.5 versicolor
## 68          4.1         1.0 versicolor
## 69          4.5         1.5 versicolor
## 70          3.9         1.1 versicolor
## 71          4.8         1.8 versicolor
## 72          4.0         1.3 versicolor
## 73          4.9         1.5 versicolor
## 74          4.7         1.2 versicolor
## 75          4.3         1.3 versicolor
## 76          4.4         1.4 versicolor
## 77          4.8         1.4 versicolor
## 78          5.0         1.7 versicolor
## 79          4.5         1.5 versicolor
## 80          3.5         1.0 versicolor
## 81          3.8         1.1 versicolor
## 82          3.7         1.0 versicolor
## 83          3.9         1.2 versicolor
## 84          5.1         1.6 versicolor
## 85          4.5         1.5 versicolor
## 86          4.5         1.6 versicolor
## 87          4.7         1.5 versicolor
## 88          4.4         1.3 versicolor
## 89          4.1         1.3 versicolor
## 90          4.0         1.3 versicolor
## 91          4.4         1.2 versicolor
## 92          4.6         1.4 versicolor
## 93          4.0         1.2 versicolor
## 94          3.3         1.0 versicolor
## 95          4.2         1.3 versicolor
## 96          4.2         1.2 versicolor
## 97          4.2         1.3 versicolor
## 98          4.3         1.3 versicolor
## 99          3.0         1.1 versicolor
## 100         4.1         1.3 versicolor
```