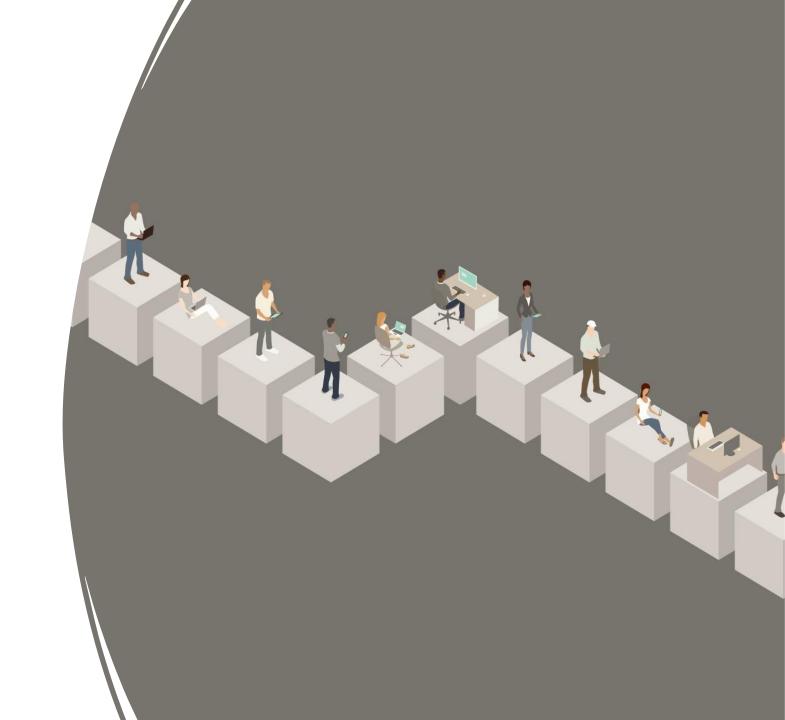
Vectores (Arreglos) en R

HERRAMIENTAS COMPUTACIONALES: EL ARTE DE LA ANALÍTICA

SEMANA TEC



Vectores en R

Estructura más sencilla en R

Colección de uno o más datos del mismo tipo

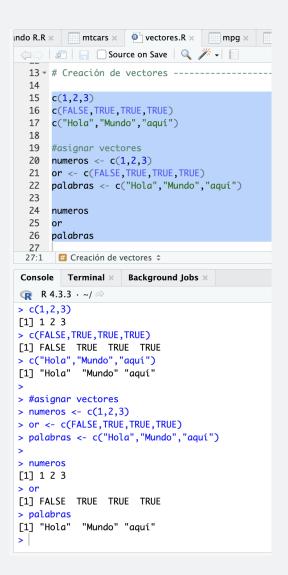
Los datos simples son considerados vectores



Creación de vectores

Se utiliza la función **c(...)** que significa combinar

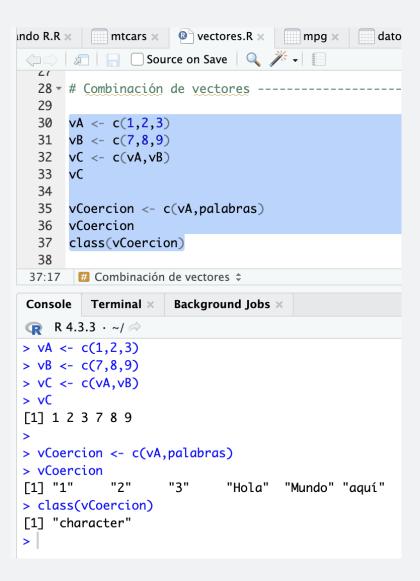
Se pueden generar vectores sin asignar a variables



Combinar Vectores

Si son del mismo tipo se pueden combinar

Si son de diferente tipo, se realiza una "coercion" de los datos para seleccionar el dato más flexible para el nuevo vector



Vectores con secuencias

Utilizando los dos puntos ":" se pueden crear secuencias numéricas para los vectores

```
ındo R.R × mtcars × vectores.R* × mpg × datos × a

    □□ Source on Save □ □ ▼ ▼ □

                                                    -→ Rur
  38
  39 - # Vectores con secuencias --
  40
      vA <- c(1:13)
  42
      VΑ
      vB <- c(100:94)
  44
      νB
      vC <- c(1.33:7.1) #Llega hasta el 7.1 buscando el 7.33
      vC
  46
  47
       # Vectores con secuencias $
 Console Terminal × Background Jobs ×
 > vA <- c(1:13)
> vA
 [1] 1 2 3 4 5 6 7 8 9 10 11 12 13
> VB <- c(100:94)
> VB
[1] 100 99 98 97 96 95 94
> vC <- c(1.33:7.1) #Llega hasta el 7.1 buscando el 7.33
> vC
[1] 1.33 2.33 3.33 4.33 5.33 6.33
```

Ejercicio

• Crea un vector para representar a las siguientes personas



Ejercicio...

Usando el vector de las personas realiza lo siguiente:

Suma las edades de las personas

Indica la longitud del vector personas

Muestra el vector completo

Muestra el primer dato

Muestra todo el vector menos el cuarto dato

Muestra las posiciones 1, 3 y 5

Muestra las posiciones 3 a 5

```
49 * # Ejercicio personas -----
  50
      personas \leftarrow c(1.35, 1.79, 1.71, 1.85, 1.70, 2.11, 2.03, 1.81, 1.70, 1.74, 1.53, 1.70, 1.88)
  52
      sum(personas)
  53
      length(personas)
      personas
      personas[1]
     personas[-4]
     personas [c(1,3,5)]
     personas[3:5]
  60
 50:1
       # Ejercicio personas $
                                                                        R Script $
Console Terminal ×
                   Background Jobs ×
                                                                          > personas <- c(1.35,1.79,1.71,1.85,1.70,2.11,2.03,1.81,1.70,1.74,1.53,1.70,1.88)
> sum(personas)
Γ17 22.9
> length(personas)
[1] 13
> personas
 [1] 1.35 1.79 1.71 1.85 1.70 2.11 2.03 1.81 1.70 1.74 1.53 1.70 1.88
> personas[1]
[1] 1.35
> personas[-4]
[1] 1.35 1.79 1.71 1.70 2.11 2.03 1.81 1.70 1.74 1.53 1.70 1.88
> personas[c(1,3,5)]
[1] 1.35 1.71 1.70
> personas[3:5]
[1] 1.71 1.85 1.70
>
```

Estadísticos básicos en R

• Valor mínimo : min(dato)

• Valor máximo : max(dato)

Promedio: mean(dato)

• Mediana, representa la tendencia central: *median(dato)*

- Varianza, la varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media: var(dato)
- Desviación estándar, indica qué tan dispersos están los datos con respecto a la media: sd(dato)
- Moda o Valor más frecuente (debes de instalar la librería modeest): mfv(dato)

```
61 * # Estadisticos básicos -----
  62
      min(personas)
      max(personas)
      mean(personas)
      median(personas)
  67
      var(personas)
      sd(personas)
  70
      install.packages("modeest")
      library(modeest)
      mfv(personas)
  74
       # Estadisticos básicos $
Console Terminal ×
                    Background Jobs ×
> min(personas)
Γ17 1.35
> max(personas)
Γ17 2.11
> mean(personas)
[1] 1.761538
> median(personas)
[1] 1.74
> var(personas)
[1] 0.03799744
> sd(personas)
[1] 0.1949293
> install.packages("modeest")
Error in install.packages : Updating loaded packages
> library(modeest)
> mfv(personas)
[1] 1.7
>
```

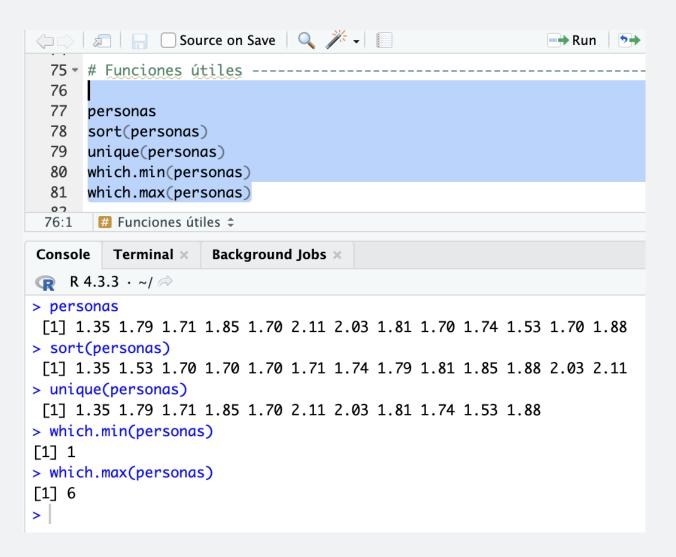
Otras funciones útiles

Ordenar los elementos

Vector sin elementos repetidos

Índice del elemento mínimo

Índice del elemento máximo

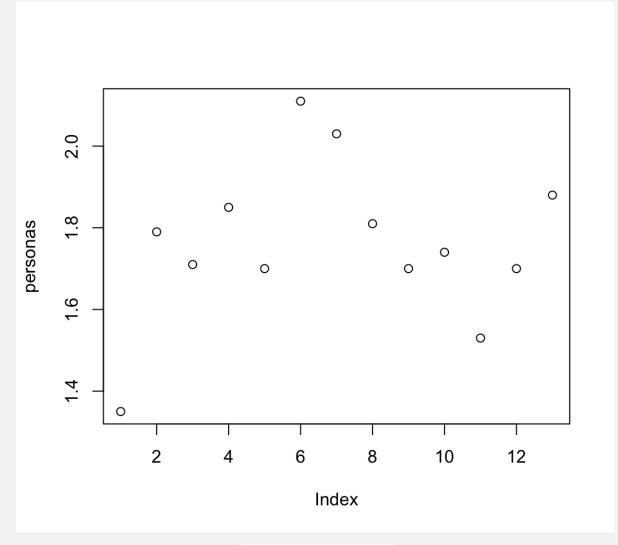


Gráficas usando vectores

El comando *plot(...)* crea una gráfica con los elementos del vector

En el eje X se encuentra el número de posición en el vector

En el eje Y se encuentran los valores en el vector

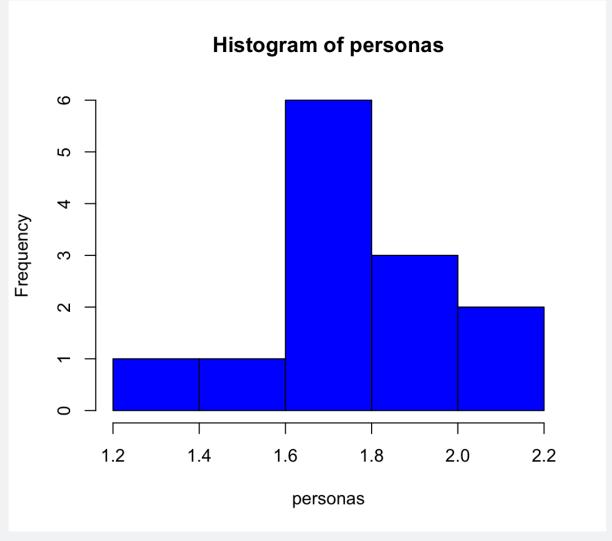


plot(personas)

Histograma

Gráfico que utiliza barras para simbolizar la manera en que se distribuye un conjunto de datos

Se utiliza el comando *hist(...)*



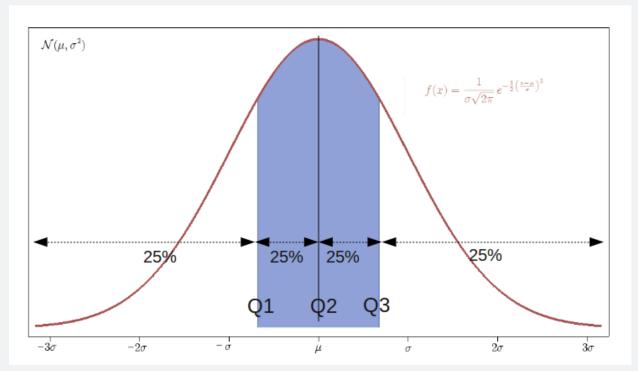
hist(personas,col="blue")

Cuartiles

Los datos deben de estar ordenados de menor a mayor

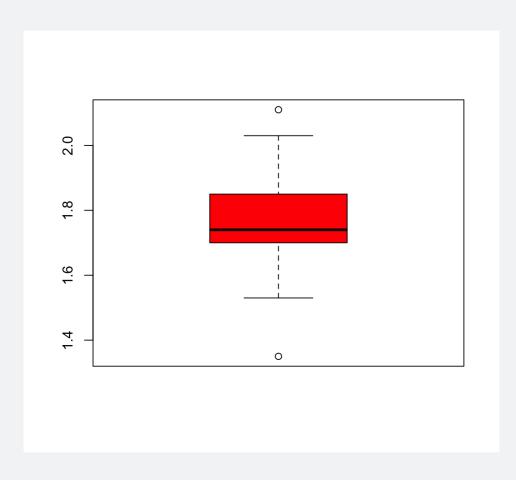
Dividen el número de datos de un conjunto en 4 partes iguales (aproximadamente):

- Q1: 25% de los datos menores por debajo de este punto
- Q2: se ubica en la media
- Q3: 25% de los datos mayores por encima de este punto

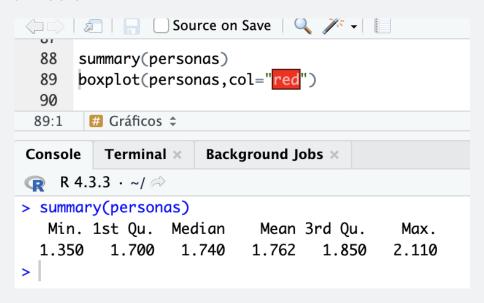


https://es.wikipedia.org/wiki/Cuantil#/media/Archivo:lqr_with_quantile.png

Boxplot y summary



- Permiten visualizer los cuartiles en un conjunto de datos
- Outlier: más de 3 desviaciones estándar de distancia de la media



Ejercicio integrador

- Equipos de 2 personas
- Genera un vector que contenga la cantidad de seguidores en Instagram que tienen tus compañeros de clase
- Realiza un análisis en R de dicha información
 - Utilza gráficos y estadísticos
- Entrega
 - Código Fuente
 - Documento PDF con lo siguiente:
 - Discusión de los resultados arrojados en el análisis en R
 - Conclusión
 - ¿Qué te dice el análisis sobre la manera en que tus compañeros de clase se relacionan mediante Instagram?



https://www.freepik.es/vector-gratis/popularidad-redes-sociales_6976389.htm