

Trabajando con Directorios y Archivos

Primero vemos como movernos por nuestro disco duro

Directorio de trabajo

```
getwd() setwd(dir)
```

Ver archivos del directorio actual

```
list.files()
```

Preguntar si un archivo está en el directori actual

```
file.exists("patient-info.txt")
```

Leer datos de un archivo texto

```
pacientes <- read.delim("datosPacientes.txt")
```

Leer datos de un archivo csv

```
tmp <- read.csv("datosPacientes.csv")
```

Visualizando Dimensiones de los Datos

Para ver las 10 primeras líneas de los datos

```
head(tmp)
```

O ejecute

```
pacientes[1:10,]
```

Para ver las 10 últimas líneas de los datos

```
tail(tmp)
```

Para saber el tipo de la estructura pacientes ejecute

```
class(pacientes)
```

Para mostrar cuantas columnas o filas tiene el dataframe

```
ncol(pacientes) nrow(pacientes)
```

Para ver las dimensiones de un dataframe ejecute el siguiente comando

```
dim(pacientes)
```

Para ver el nombre de las columnas de los datos ejecute lo siguiente

```
colnames(pacientes)
```

Valores perdidos. A veces los datos no son suministrados correctamente y R puede lidiar con valores como NA (diferente de NULL)

NULL es un objeto vacío NA es un valor perdido

Ejemplo

```
length(pacientes$Height) mean(pacientes$Height)
```

R puede lidiar con estos valores perdidos

```
mean(pacientes$Height, na.rm = TRUE)
```

```
mean(na.omit(pacientes$Height))
```

Podemos realizar un análisis de índice de masa corporal de la siguiente manera

```
BMI <- (pacientes$Weight)/((pacientes$Height/100)^2) upper.limit <- mean(BMI,na.rm = TRUE) +  
2*sd(BMI,na.rm = TRUE) upper.limit
```

Graficando los datos de forma básica

Podemos graficar estos resultados

```
plot(BMI) abline(h=upper.limit)
```

Redondear los valores

```
round(BMI,1)
```

Guardar el índice como una columna más del data frame

```
pacientes$BMI <- round(BMI,1) head(pacientes)
```

Seleccionar los pacientes arriba del límite

```
BMI > upper.limit
```

```
pacientes[BMI > upper.limit,]
```

Ó,

```
candidatos <- BMI > upper.limit
```

```
pacientes[candidatos,]
```

La función which toma un vector lógico y retorna los índices

```
which(BMI > upper.limit)
```

```
candidatos <- which(BMI > upper.limit)
```

Guardar Resultados

Escribir los resultados en un archivo CSV

```
write.csv(pacientes[candidatos,], file="candidatos.csv")
```