

R taller - Manejo de datos - Ejercicios

Anabelle Laurent

10/25/2021

```
library(dplyr)
library(tidyverse)
library(readr)
bici<-readRDS("mi_carpeta/trips.RDS")
```

Descripción de los datos “bici” bici es una base de datos que contiene información sobre cada alquiler de bicicletas:

Variable	Descripción
Duration	duración del viaje en segundos
Start.date	sello de tiempo de la fecha y hora de devolución. La fecha se recopila como año, mes y día; el tiempo se escribe como horas, minutos y segundos.
End.date	sello de tiempo de la fecha y hora de devolución. Mismo formato que “Start.date”.
Start.station.number	Identificador de estación de la estación donde se revisó la bicicleta.
Start.station	nombre de la estación de la estación donde se revisó la bicicleta.
End.station.number	Identificador de estación de la estación a la que se devolvió la bicicleta.
End.station	nombre de la estación de la estación a la que se devolvió la bicicleta.
Bike.number	identificador de bicicleta.
Member.type	tipo de socio (Regular o Casual).

```
dim(bici) # dimensión
```

```
## [1] 815370      9
```

```
str(bici) # estructura de los datos
```

```
## 'data.frame':   815370 obs. of  9 variables:
## $ Duration      : num  197 435 955 462 3357 ...
## $ Start.date    : Factor w/ 735188 levels "2017-10-01 00:00:02",...: 1 2 3 3 4 5 6 7 8 9 ...
## $ End.date      : Factor w/ 734389 levels "2017-10-01 00:03:19",...: 1 2 19 3 120 71 21 6 9 9
## $ Start.station.number: Factor w/ 488 levels "31000","31001",...: 144 105 151 112 190 190 291 215 103
## $ Start.station    : Factor w/ 488 levels "10th & E St NW",...: 59 167 67 6 106 106 84 103 12 15
## $ End.station.number: Factor w/ 488 levels "31000","31001",...: 159 290 104 103 190 219 189 132 267
## $ End.station      : Factor w/ 488 levels "10th & E St NW",...: 361 377 55 12 106 282 95 50 258 87
## $ Bike.number      : Factor w/ 4293 levels "W00005","W00007",...: 2157 381 1450 2149 3297 2227 591
## $ Member.type      : Factor w/ 2 levels "Casual","Member": 2 2 2 2 1 1 1 2 2 2 ...
```

```
head(bici) # devuelve la primera parte del conjunto de datos
```

```
##      Duration      Start.date      End.date Start.station.number
## 1  197.068 2017-10-01 00:00:02 2017-10-01 00:03:19      31214
## 2  434.934 2017-10-01 00:00:23 2017-10-01 00:07:38      31104
## 3  955.437 2017-10-01 00:00:56 2017-10-01 00:16:52      31221
## 4  461.619 2017-10-01 00:00:56 2017-10-01 00:08:37      31111
## 5 3357.184 2017-10-01 00:00:59 2017-10-01 00:56:56      31260
## 6 2235.414 2017-10-01 00:01:06 2017-10-01 00:38:21      31260
##              Start.station End.station.number
## 1      17th & Corcoran St NW      31229
## 2 Adams Mill & Columbia Rd NW      31602
## 3              18th & M St NW      31103
## 4              10th & U St NW      31102
## 5              23rd & E St NW      31260
## 6              23rd & E St NW      31289
##              End.station Bike.number Member.type
## 1      New Hampshire Ave & T St NW      W21022      Member
## 2      Park Rd & Holmead Pl NW      W00470      Member
## 3              16th & Harvard St NW      W20206      Member
## 4              11th & Kenyon St NW      W21014      Member
## 5              23rd & E St NW      W22349      Casual
## 6 Henry Bacon Dr & Lincoln Memorial Circle NW      W21107      Casual
```

Ejercicio 1: devolver los datos para los socios “Casual” qui hicieron un viaje mas largo que 1000 segundos

Ejercicio 2: devolver los datos para las estaciones “Lincoln Memorial”, “6th & K St NE” y “Kennedy Center”

Ejercicio 3: seleccionar Start.station y Duracion organizar por Duración y por orden descendente

Ejercicio 4: crear una columna llamada “Duración_min”, que es la duración en minutos

Ejercicio 5: traduce calquier columna en español

Ejercicio 6: ¿Cuál es la duración promedio de los viajes (en minutos) por tipo de socios?

```
?separate
```

Ejercicio 7: separar “Start.date” en fecha y Hora