

# Análisis

Zabdiel Emilio Moreno Mendoza

2022-08-04

## Análisis

Para comenzar unifique todos los conjuntos de datos en uno solo, esto mediante full\_join()

viajes\_cyclistic1

```
## # A tibble: 50 × 10
##   ride_id      ridea...1 started_at      ended_at      start...2
##   <chr>      <chr>    <dtm>          <dtm>          <chr>
##   <chr>
##  1 940425EC360C... electr... 2022-01-08 13:29:14 2022-01-08 13:51:23 SL-012
##    TA1306...
##  2 B7EA0A01D646... classi... 2021-09-24 15:18:57 2021-09-24 15:43:48 TA1307...
##    TA1305...
##  3 8F6F196C36B4... classi... 2022-03-08 08:04:44 2022-03-08 08:12:47 TA1305...
##    623
##  4 C2FA2420FF40... classi... 2022-06-20 08:43:54 2022-06-20 08:48:45 13341
##    KA1503...
##  5 F1B3D0D7DD6B... classi... 2021-07-24 14:23:35 2021-07-24 14:55:24 TA1305...
##    TA1306...
##  6 7E3AF95BB8DB... electr... 2022-05-17 18:19:08 2022-05-17 18:23:42 <NA>
##    <NA>
##  7 651DEB6DCA2D... classi... 2021-09-08 22:47:21 2021-09-08 23:12:24 TA1307...
##    TA1307...
##  8 772F75EFD0A2... electr... 2021-09-18 19:45:41 2021-09-18 19:53:42 13191
##    <NA>
##  9 B96EDAA835F3... classi... 2021-07-03 19:33:32 2021-07-03 19:42:59 TA1307...
##    13217
## 10 E1E14726C2AE... classi... 2022-06-09 07:27:47 2022-06-09 07:37:12 TA1305...
##    13206
## # ... with 40 more rows, 4 more variables: member_casual <chr>,
## #   ride_length <time>, day_of_week <chr>, month <chr>, and abbreviated
## #   variable names 1rideable_type, 2start_station_id, 3end_station_id
## # i Use `print(n = ...)` to see more rows, and `colnames()` to see all
## #   variable names
```

Posteriormente saque promedios por datos y comparando los datos como los siguientes:

prom1

```
## # A tibble: 7 × 2
##   day_of_week Promedio
##   <chr>         <time>
## 1 domingo      21'30"
## 2 lunes        16'44"
## 3 martes       14'49"
## 4 miércoles    14'47"
## 5 jueves       15'26"
## 6 viernes      16'47"
## 7 sábado       21'11"
```

prom2

```
## # A tibble: 2 × 2
##   member_casual Promedio
##   <chr>         <time>
## 1 casual       24'25"
## 2 member       12'33"
```

prom3

```
## # A tibble: 14 × 3
##   day_of_week member_casual Promedio
##   <chr>         <chr>         <time>
## 1 domingo      casual       27'52"
## 2 domingo      member       14'20"
## 3 lunes        casual       24'44"
## 4 lunes        member       12'07"
## 5 martes       casual       21'27"
## 6 martes       member       11'42"
## 7 miércoles    casual       20'46"
## 8 miércoles    member       11'51"
## 9 jueves       casual       21'31"
## 10 jueves      member       12'02"
## 11 viernes     casual       23'01"
## 12 viernes     member       12'17"
## 13 sábado      casual       27'06"
## 14 sábado      member       14'13"
```

Filtré los resultados menores a 30 segundos y aquellos mayores a un día, pues los considere atípicos, los menores a 30 segundos que tuvieran la misma estación de comienzo y fin pues son bastante cortos como para un paseo y los mayores a un día porque los clientes casuales eran los únicos que podían superar las 25 horas.

Con esos datos también calcule la cantidad de viajes agrupado por días, tipo de clientes y meses

can1

```
## # A tibble: 7 × 2
##   day_of_week      n
##   <chr>          <dbl>
## 1 domingo        636384
## 2 lunes          569104
## 3 martes         592153
## 4 miércoles      596968
## 5 jueves         630742
## 6 viernes        611325
## 7 sábado         730451
```

can2

```
## # A tibble: 2 × 2
##   member_casual      n
##   <chr>          <dbl>
## 1 casual        1807730
## 2 member        2559397
```

can3

```
## # A tibble: 14 × 3
##   day_of_week member_casual      n
##   <chr>          <chr>          <dbl>
## 1 domingo        casual        336835
## 2 domingo        member        299549
## 3 lunes          casual        208253
## 4 lunes          member        360851
## 5 martes         casual        189619
## 6 martes         member        402534
## 7 miércoles      casual        195967
## 8 miércoles      member        401001
## 9 jueves         casual        226104
## 10 jueves        member        404638
## 11 viernes       casual        255826
## 12 viernes       member        355499
## 13 sábado        casual        395126
## 14 sábado        member        335325
```

Con esa información note algunos patrones (que señalare en el siguiente documento), entonces guarde la información necesaria en archivos .csv para poder visualizarla posteriormente.

El script con los comandos utilizados tiene el nombre de “analisis.R” y se encuentra en la misma carpeta que este informe.