Laboratorio 1 Luis Tujab

Luis Tujab 1103920

2023-10-05

R Markdown

Este es el documento entregable

```
gatos <- data.frame(color= c("blanco","Negro","gris"),
       peso = c(1,2,3),
       propietario = c(1,0,1))

gatos
```

```
##    color peso propietario
## 1 blanco    1           1
## 2  Negro    2           0
## 3   gris    3           1
```

```
gatos$peso <- gatos$peso*2
gatos
```

```
##    color peso propietario
## 1 blanco    2           1
## 2  Negro    4           0
## 3   gris    6           1
```

```
paste ("el gato es color: ", gatos$color)
```

```
## [1] "el gato es color:  blanco" "el gato es color:  Negro"
## [3] "el gato es color:  gris"
```

```
class(gatos)
```

```
## [1] "data.frame"
```

```
class(gatos$peso)
```

```
## [1] "numeric"
```

```
class(gatos$color)
```

```
## [1] "character"
```

```
str(gatos)
```

```
## 'data.frame':    3 obs. of  3 variables:
##  $ color      : chr  "blanco" "Negro" "gris"
```

```
##  $ peso       : num  2 4 6
##  $ propietario: num  1 0 1
```

```
mi_vector <- c(2,6,"3")
class(mi_vector)
```

```
## [1] "character"
```

```
char_to_number <- as.numeric(mi_vector)
class(char_to_number)
```

```
## [1] "numeric"
```

```
gatos$propietario <- as.logical(gatos$propietario)
gatos
```

```
##    color peso propietario
## 1 blanco    2        TRUE
## 2  Negro    4       FALSE
## 3   gris    6        TRUE
```

Agregar librerias

```
library(nycflights13)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
flights
```

```
## # A tibble: 336,776 × 19
##     year month   day dep_time sched_dep_time dep_delay arr_time
sched_arr_time
##    <int> <int> <int>    <int>          <int>     <dbl>    <int>
<int>
##  1  2013     1     1      517            515         2      830
819
##  2  2013     1     1      533            529         4      850
830
##  3  2013     1     1      542            540         2      923
850
##  4  2013     1     1      544            545        -1     1004
1022
##  5  2013     1     1      554            600        -6      812
```

```
837
##  6  2013     1     1      554            558        -4      740
728
##  7  2013     1     1      555            600        -5      913
854
##  8  2013     1     1      557            600        -3      709
723
##  9  2013     1     1      557            600        -3      838
846
## 10  2013     1     1      558            600        -2      753
745
## # ℹ 336,766 more rows
## # ℹ 11 more variables: arr_delay <dbl>, carrier <chr>, flight <int>,
## #   tailnum <chr>, origin <chr>, dest <chr>, air_time <dbl>, distance
<dbl>,
## #   hour <dbl>, minute <dbl>, time_hour <dttm>
```

```
glimpse(flights)
```

```
## Rows: 336,776
## Columns: 19
## $ year           <int> 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013,
2013, 2…
## $ month          <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1…
## $ day            <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1…
## $ dep_time       <int> 517, 533, 542, 544, 554, 554, 555, 557, 557,
558, 558, …
## $ sched_dep_time <int> 515, 529, 540, 545, 600, 558, 600, 600, 600,
600, 600, …
## $ dep_delay      <dbl> 2, 4, 2, -1, -6, -4, -5, -3, -3, -2, -2, -2, -
2, -2, -1…
## $ arr_time       <int> 830, 850, 923, 1004, 812, 740, 913, 709, 838,
753, 849,…
## $ sched_arr_time <int> 819, 830, 850, 1022, 837, 728, 854, 723, 846,
745, 851,…
## $ arr_delay      <dbl> 11, 20, 33, -18, -25, 12, 19, -14, -8, 8, -2, -
3, 7, -1…
## $ carrier        <chr> "UA", "UA", "AA", "B6", "DL", "UA", "B6", "EV",
"B6", "…
## $ flight         <int> 1545, 1714, 1141, 725, 461, 1696, 507, 5708,
79, 301, 4…
## $ tailnum        <chr> "N14228", "N24211", "N619AA", "N804JB",
"N668DN", "N394…
## $ origin         <chr> "EWR", "LGA", "JFK", "JFK", "LGA", "EWR",
"EWR", "LGA",…
## $ dest           <chr> "IAH", "IAH", "MIA", "BQN", "ATL", "ORD",
"FLL", "IAD",…
## $ air_time       <dbl> 227, 227, 160, 183, 116, 150, 158, 53, 140,
```

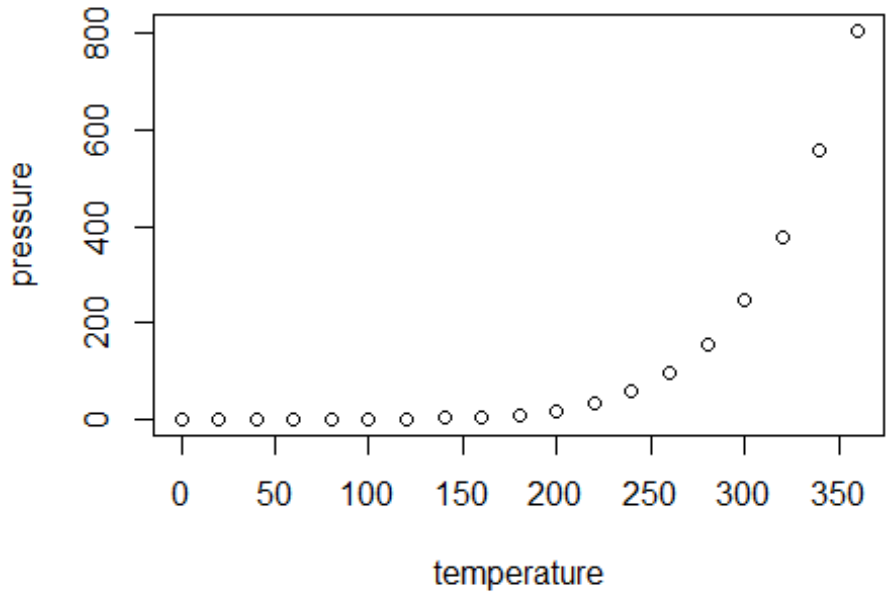
```
138, 149, 1…
## $ distance       <dbl> 1400, 1416, 1089, 1576, 762, 719, 1065, 229,
944, 733, …
## $ hour           <dbl> 5, 5, 5, 5, 6, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5,
6, 6, 6…
## $ minute         <dbl> 15, 29, 40, 45, 0, 58, 0, 0, 0, 0, 0, 0, 0, 0,
0, 59, 0…
## $ time_hour      <dttm> 2013-01-01 05:00:00, 2013-01-01 05:00:00,
2013-01-01 0…

flights$carrier <- as.factor(flights$carrier)
glimpse(flights)

## Rows: 336,776
## Columns: 19
## $ year           <int> 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013,
2013, 2…
## $ month          <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1…
## $ day            <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1…
## $ dep_time       <int> 517, 533, 542, 544, 554, 554, 555, 557, 557,
558, 558, …
## $ sched_dep_time <int> 515, 529, 540, 545, 600, 558, 600, 600, 600,
600, 600, …
## $ dep_delay      <dbl> 2, 4, 2, -1, -6, -4, -5, -3, -3, -2, -2, -2, -
2, -2, -1…
## $ arr_time       <int> 830, 850, 923, 1004, 812, 740, 913, 709, 838,
753, 849,…
## $ sched_arr_time <int> 819, 830, 850, 1022, 837, 728, 854, 723, 846,
745, 851,…
## $ arr_delay      <dbl> 11, 20, 33, -18, -25, 12, 19, -14, -8, 8, -2, -
3, 7, -1…
## $ carrier        <fct> UA, UA, AA, B6, DL, UA, B6, EV, B6, AA, B6, B6,
UA, UA,…
## $ flight         <int> 1545, 1714, 1141, 725, 461, 1696, 507, 5708,
79, 301, 4…
## $ tailnum        <chr> "N14228", "N24211", "N619AA", "N804JB",
"N668DN", "N394…
## $ origin         <chr> "EWR", "LGA", "JFK", "JFK", "LGA", "EWR",
"EWR", "LGA",…
## $ dest           <chr> "IAH", "IAH", "MIA", "BQN", "ATL", "ORD",
"FLL", "IAD",…
## $ air_time       <dbl> 227, 227, 160, 183, 116, 150, 158, 53, 140,
138, 149, 1…
## $ distance       <dbl> 1400, 1416, 1089, 1576, 762, 719, 1065, 229,
944, 733, …
## $ hour           <dbl> 5, 5, 5, 5, 6, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5,
6, 6, 6…
## $ minute         <dbl> 15, 29, 40, 45, 0, 58, 0, 0, 0, 0, 0, 0, 0, 0,
```

```
0, 59, 0…
## $ time_hour      <dttm> 2013-01-01 05:00:00, 2013-01-01 05:00:00,
2013-01-01 0…

data <- dplyr::select(flights,-year)

flights %>%
  select(contains("time"))

## # A tibble: 336,776 × 6
##    dep_time sched_dep_time arr_time sched_arr_time air_time time_hour
##       <int>          <int>    <int>          <int>    <dbl> <dttm>
##  1      517            515      830            819      227 2013-01-01
05:00:00
##  2      533            529      850            830      227 2013-01-01
05:00:00
##  3      542            540      923            850      160 2013-01-01
05:00:00
##  4      544            545     1004           1022      183 2013-01-01
05:00:00
##  5      554            600      812            837      116 2013-01-01
06:00:00
##  6      554            558      740            728      150 2013-01-01
05:00:00
##  7      555            600      913            854      158 2013-01-01
06:00:00
##  8      557            600      709            723       53 2013-01-01
06:00:00
##  9      557            600      838            846      140 2013-01-01
06:00:00
## 10      558            600      753            745      138 2013-01-01
06:00:00
## # ℹ 336,766 more rows

table(flights$origin) #Cantidades a partir de una tabla

##
##    EWR    JFK    LGA
## 120835 111279 104662

prop.table(table(flights$origin)) #Porcentajes a partir de una talba

##
##       EWR       JFK       LGA
## 0.3587993 0.3304244 0.3107763

summary(cars)

##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
```

```
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.