# visualización de datos de los viajes en Uber

Proyecto de visualización de datos de los viajes que se realizaron en Uber, los últimos 6 años.

```
library(readr)
library(dplyr)
```

```
##
## Adjuntando el paquete: 'dplyr'
```

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

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

```
library(lubridate)
```

```
##
## Adjuntando el paquete: 'lubridate'
```

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

```
library(ggplot2)
library(ggthemes)
```

```
df <- read_csv("uber_stock_data.csv")
```

```
## Rows: 1444 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl  (6): Adj Close, Close, High, Low, Open, Volume
## date (1): Date
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
str(df)
```

```
## spc_tbl_ [1,444 x 7] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Date     : Date[1:1444], format: "2019-05-10" "2019-05-13" ...
##  $ Adj Close: num [1:1444] 41.6 37.1 40 41.3 43 ...
##  $ Close    : num [1:1444] 41.6 37.1 40 41.3 43 ...
##  $ High     : num [1:1444] 45 39.2 40 41.9 44.1 ...
##  $ Low      : num [1:1444] 41.1 36.1 36.8 39 41.2 ...
##  $ Open     : num [1:1444] 42 38.8 38.3 39.4 41.5 ...
##  $ Volume   : num [1:1444] 1.86e+08 7.94e+07 4.67e+07 3.61e+07 3.81e+07 ...
##  - attr(*, "spec")=
##   .. cols(
```

```
##   ..   Date = col_date(format = ""),
##   ..   `Adj Close` = col_double(),
##   ..   Close = col_double(),
##   ..   High = col_double(),
##   ..   Low = col_double(),
##   ..   Open = col_double(),
##   ..   Volume = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
names(df)
```

```
## [1] "Date"      "Adj Close" "Close"     "High"      "Low"       "Open"
## [7] "Volume"
```

```
head(df)
```

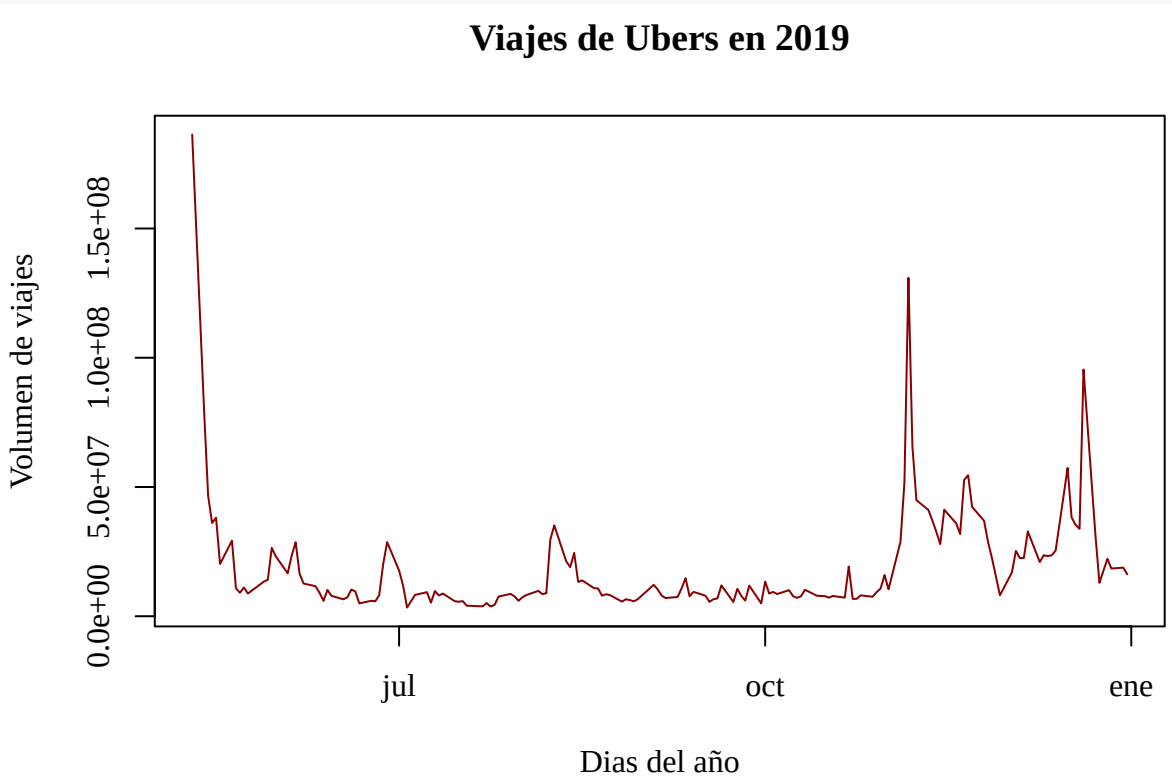
```
## # A tibble: 6 x 7
##   Date       `Adj Close` Close  High   Low  Open    Volume
##   <date>           <dbl> <dbl> <dbl> <dbl> <dbl>     <dbl>
## 1 2019-05-10        41.6  41.6  45    41.1  42   186322500
## 2 2019-05-13        37.1  37.1  39.2  36.1  38.8  79442400
## 3 2019-05-14        40.0  40.0  40.0  36.8  38.3  46661100
## 4 2019-05-15        41.3  41.3  41.9  39.0  39.4  36086100
## 5 2019-05-16        43    43    44.1  41.2  41.5  38115500
## 6 2019-05-17        41.9  41.9  43.3  41.3  42.0  20225700
```

```
# Ordenamos el dataset por fechas
df$Date <- as.Date(df$Date, format = "%Y-%m-%d")
df <- df %>%
  arrange(Date)
head(df)
```

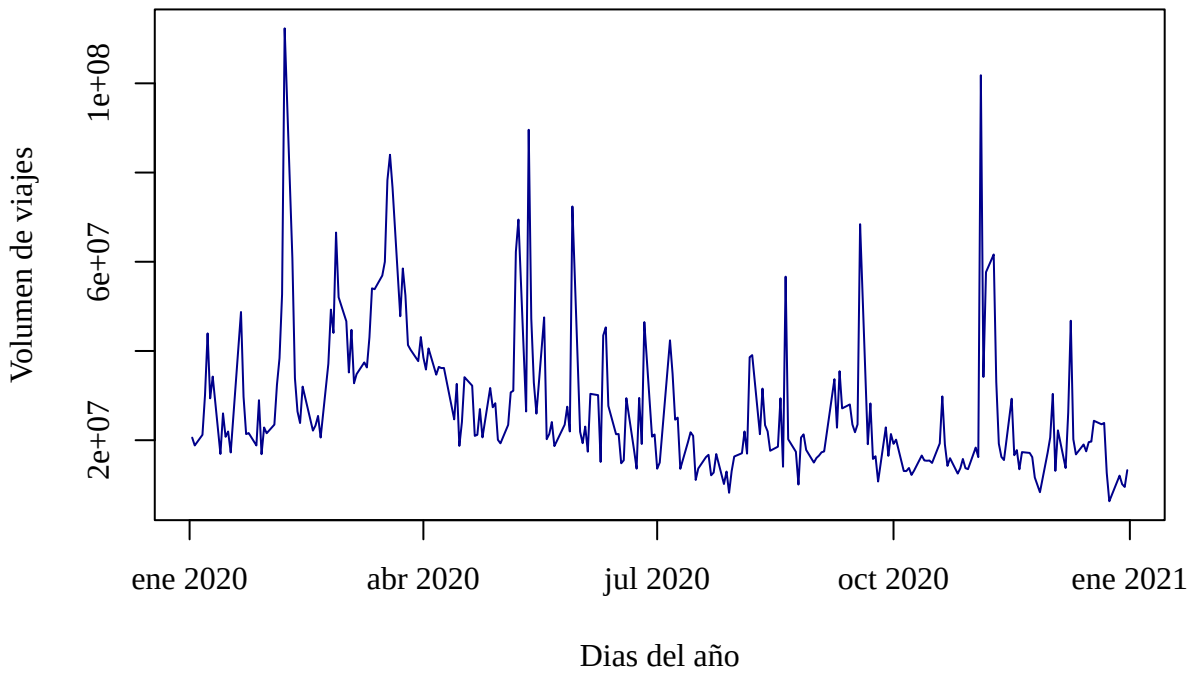
```
## # A tibble: 6 x 7
##   Date       `Adj Close` Close  High   Low  Open    Volume
##   <date>           <dbl> <dbl> <dbl> <dbl> <dbl>     <dbl>
## 1 2019-05-10        41.6  41.6  45    41.1  42   186322500
## 2 2019-05-13        37.1  37.1  39.2  36.1  38.8  79442400
## 3 2019-05-14        40.0  40.0  40.0  36.8  38.3  46661100
## 4 2019-05-15        41.3  41.3  41.9  39.0  39.4  36086100
## 5 2019-05-16        43    43    44.1  41.2  41.5  38115500
## 6 2019-05-17        41.9  41.9  43.3  41.3  42.0  20225700
```

```
# Gráfico el volumen de cada mes en cada año
df_fil19 <- df %>%
  filter(year(Date) == 2019)

df_fil20 <- df %>%
  filter(year(Date) == 2020)

df_fil21 <- df %>%
  filter(year(Date) == 2021)

df_fil22 <- df %>%
  filter(year(Date) == 2022)
```

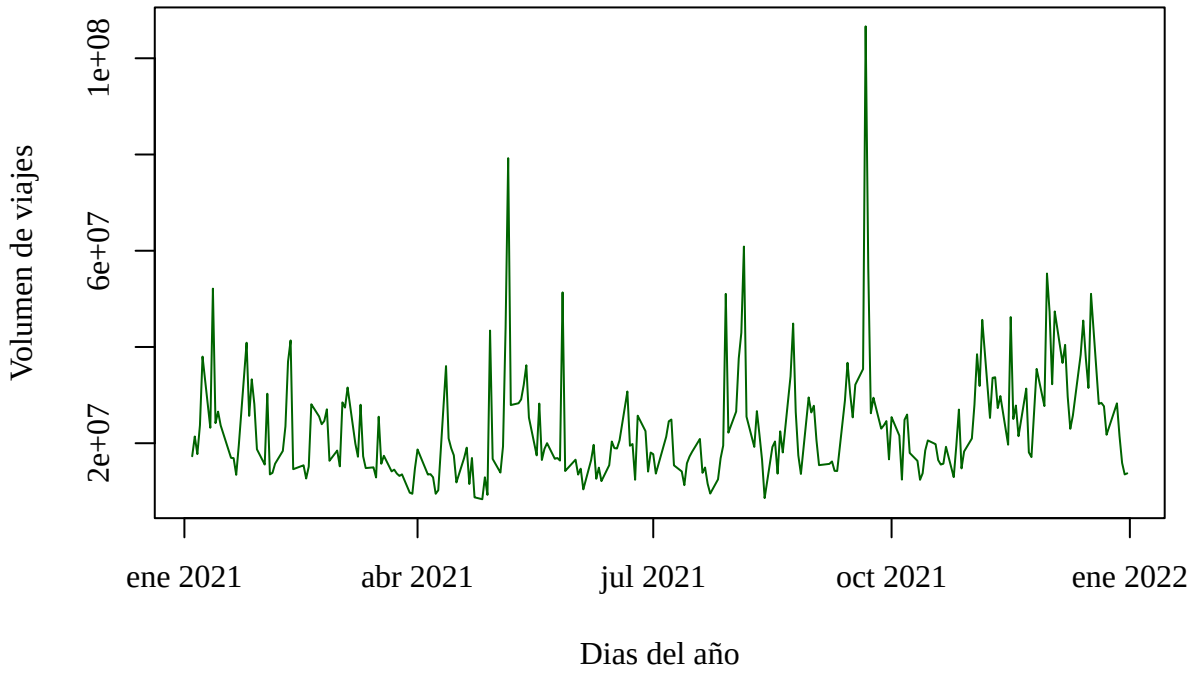
```r
# 2019
plot(df_fil19$Date, df_fil19$Volume,
     xlab = "Dias del año",
     ylab = "Volumen de viajes",
     main = "Viajes de Ubers en 2019",
     type = "l", col = "darkred",
     family = "serif")
```



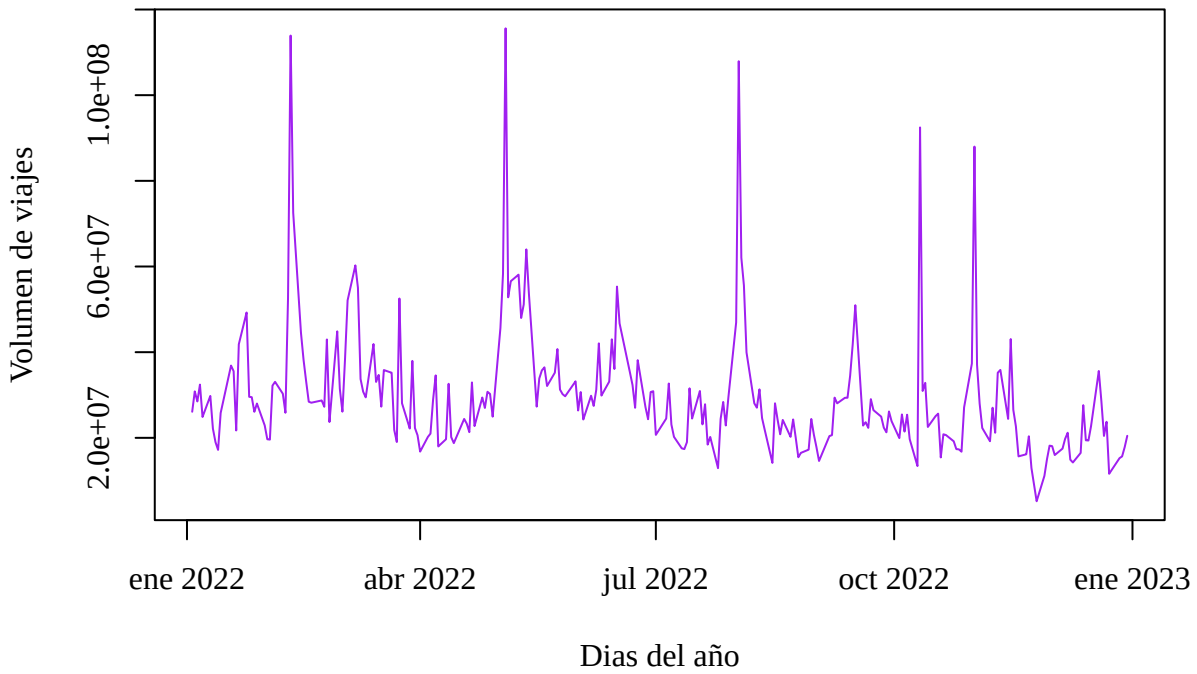

```r
# 2020
plot(df_fil20$Date, df_fil20$Volume,
     xlab = "Dias del año",
     ylab = "Volumen de viajes",
     main = "Viajes de Ubers en 2020",
     type = "l", col = "darkblue",
     family = "serif")
```




```
# 2021
plot(df_fil21$Date, df_fil21$Volume,
     xlab = "Dias del año",
     ylab = "Volumen de viajes",
     main = "Viajes de Ubers en 2021",
     type = "l", col = "darkgreen",
     family = "serif")
```

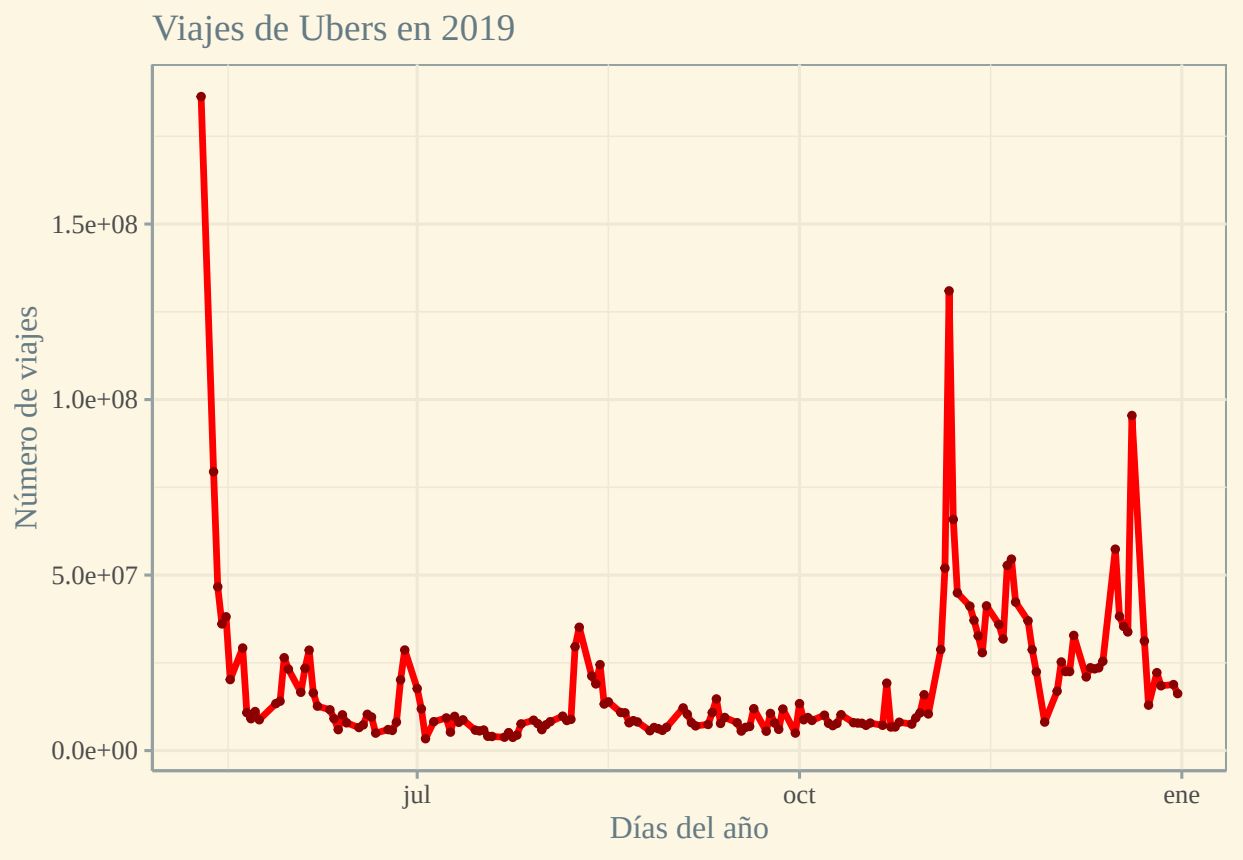



```
# 2022
plot(df_fil22$Date, df_fil22$Volume,
     xlab = "Dias del año",
     ylab = "Volumen de viajes",
     main = "Viajes de Ubers en 2022",
     type = "l", col = "purple",
     family = "serif")
```




```
# Con ggplot2
ggplot(df_fil19, aes(df_fil19$Date, df_fil19$Volume)) +
  geom_line(color = "red", size = 1.2, linetype = "solid") +
  geom_point(color = "darkred", size = 1)+
  theme_solarized(base_family = "serif") +
  labs(x = "Días del año",y = "Número de viajes",
  title = "Viajes de Ubers en 2019")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```


Viajes de Ubers en 2019
Número de viajes
1.5e+08
1.0e+08
5.0e+07
0.0e+00
jul
oct
ene
Días del año