

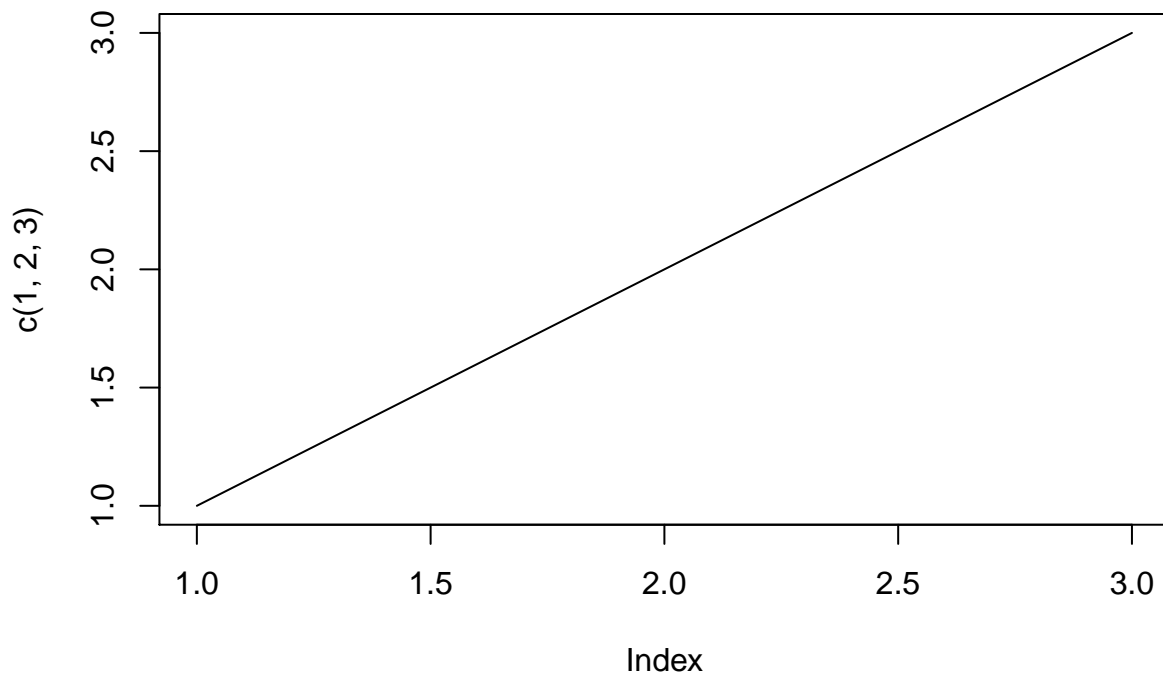
Clase de UfO's

Este es un documento de R Markdown. Markdown es una sintaxis de formato simple para la creación de documentos HTML, PDF y MS Word. Para obtener más detalles sobre el uso de R Markdown, consulta: <https://rmarkdown.rstudio.com/>

Al hacer clic en el botón **Knit** (“tejer”), se generará un documento que incluye tanto el contenido como la salida de cualquier fragmento de código R incrustado dentro del documento. Puede incrustar un fragmento de código R como este:

```
#install.packages("rmarkdown")
```

También puede incrustar gráficos, por ejemplo:



Tenga en cuenta que el parámetro **echo = FALSE** se agregó al fragmento de código para evitar la impresión del código R que generó el gráfico.

Para establecer opciones globales que se apliquen a cada fragmento de su archivo o fragmento de código **knitr::opts_chunk\$set(echo = TRUE)**

Dependencies

```
#install.packages("tidyverse")

library(tidyverse)
```

Read and preview CSV

```
ufo <- read_csv("ufo.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
##
## -- Column specification -----
## cols(
##   X1 = col_double(),
##   datetime = col_character(),
##   city = col_character(),
##   state = col_character(),
##   country = col_character(),
##   shape = col_character(),
##   'duration (seconds)' = col_double(),
##   'duration (hours/min)' = col_character(),
##   comments = col_character(),
##   'date posted' = col_character(),
##   latitude = col_double(),
##   longitude = col_double()
## )
```

```
ufo %>% head()
```

```
## # A tibble: 6 x 12
##   X1 datetime      city state country shape 'duration (sec~' 'duration (hou~'
##   <dbl> <chr>      <chr> <chr> <chr>   <chr>         <dbl> <chr>
## 1     0 10/10/1949 ~ san ~ tx    us    cyli~         2700 45 minutes
## 2     2 10/10/1960 ~ kane~ hi    us    light         900 15 minutes
## 3     3 10/10/1961 ~ bris~ tn    us    sphe~         300 5 minutes
## 4     4 10/10/1965 ~ norw~ ct    us    disk        1200 20 minutes
## 5     5 10/10/1966 ~ pell~ al    us    disk         180 3 minutes
## 6     6 10/10/1966 ~ live~ fl    us    disk         120 several minutes
## # ... with 4 more variables: comments <chr>, 'date posted' <chr>,
## #   latitude <dbl>, longitude <dbl>
```

El operador pipeline %>% es útil para concatenar múltiples dplyr operaciones.

El número total de avistamientos de ovnis

```
ufo.count <- ufo %>% count()

paste("Se localizaron ", ufo.count, "avistamientos de ovnis")
```

```
## [1] "Se localizaron 66515 avistamientos de ovnis"
```

El número y la lista de estados, provincias y territorios

```
ufo$state %>% unique() %>% length()
```

```
## [1] 67
```

```
ufo$state %>% unique()
```

```
## [1] "tx" "hi" "tn" "ct" "al" "fl" "ca" "nc" "ny" "ky" "mi" "ma" "ks" "sc" "wa"
## [16] "co" "nh" "wi" "me" "ga" "pa" "il" "ar" "on" "mo" "oh" "in" "az" "mn" "nv"
## [31] "nf" "ne" "or" "bc" "ia" "va" "id" "nm" "nj" "mb" "wv" "ok" "ri" "nb" "vt"
## [46] "la" "pr" "ak" "ms" "ut" "md" "ab" "mt" "sk" "wy" "sd" "pq" "de" "nd" "nt"
## [61] "qc" "sa" "ns" "yk" "pe" "yt" "dc"
```

La duración promedio del avistamiento de ovnis por estado

```
ufo %>%
  group_by(state) %>%
  summarise(avg.duration = mean('duration (seconds)')) %>%
  arrange(desc(avg.duration))
```

```
## # A tibble: 67 x 2
##   state avg.duration
##   <chr>      <dbl>
## 1 ar         115893.
## 2 on          62892.
## 3 hi          26196.
## 4 wa          15265.
## 5 fl          14891.
## 6 la          12445.
## 7 va          10903.
## 8 ms           9206.
## 9 ga           7708.
## 10 wv          6778.
## # ... with 57 more rows
```

El número de avistamientos de ovnis por estado

```
ufo %>%
  group_by(state) %>%
  summarise(number.sightings = n()) %>%
  arrange(desc(number.sightings))
```

```
## # A tibble: 67 x 2
##   state number.sightings
##   <chr>      <int>
## 1 ca          8683
## 2 fl          3754
## 3 wa          3709
## 4 tx          3397
```

```
## 5 ny                2915
## 6 il                2447
## 7 az                2362
## 8 pa                2319
## 9 oh                2252
## 10 mi               1781
## # ... with 57 more rows
```

El número de avistamientos por forma de ovni percibida

```
ufo %>%
  group_by(shape) %>%
  summarise(shape.count = n()) %>%
  arrange(desc(shape.count))
```

```
## # A tibble: 28 x 2
##   shape      shape.count
##   <chr>         <int>
## 1 light        14130
## 2 triangle     6817
## 3 circle       6404
## 4 fireball     5364
## 5 unknown      4774
## 6 other        4705
## 7 sphere       4552
## 8 disk         4319
## 9 oval         3160
## 10 formation   2088
## # ... with 18 more rows
```