▾ Nombre: Celic Gabriel Hernández Archundia

Matrícula: 2877240

▾ Nota 1. Considerar en todos los ejercicios que pueden existir valores nulos (NA)

Nota 2. Si no hay otra instrucción, ordenar los resultados de mayor a menor valor

```
%%capture
!pip install rpy2==3.5.1

%load_ext rpy2.ipython

from google.colab import drive
drive.mount('/content/drive')
```

```
    Mounted at /content/drive
```

```
%%R
library(readr)
library(magrittr)
library(dplyr)
```

```
    WARNING:rpy2.rinterface_lib.callbacks:R[write to console]:
    Attaching package: 'dplyr'

    WARNING:rpy2.rinterface_lib.callbacks:R[write to console]: The following objects are masked from 'package:stats':

        filter, lag

    WARNING:rpy2.rinterface_lib.callbacks:R[write to console]: The following objects are masked from 'package:base':

        intersect, setdiff, setequal, union
```

```
%%R
datos <- read.csv("drive/MyDrive/Tecmilenio/Big Data/videogames.csv")
```

▾ 1. Cantidad de videojuegos por plataforma

```
%%R
colnames(datos)
```

```
    [1] "Name"          "Platform"      "Year_of_Release" "Genre"
    [5] "Publisher"     "NA_Sales"      "EU_Sales"        "JP_Sales"
    [9] "Other_Sales"   "Global_Sales"  "Critic_Score"    "Critic_Count"
    [13] "User_Score"    "User_Count"    "Developer"       "Rating"
```

```
%%R
datos %>%
group_by(Platform) %>%
summarize(cantidad = n()) %>%
arrange(desc(cantidad))
```

```
    # A tibble: 31 × 2
      Platform cantidad
      <chr>       <int>
    1 PS2          2161
    2 DS           2152
    3 PS3          1331
    4 Wii          1320
    5 X360         1262
    6 PSP          1209
    7 PS           1197
```

```
 8 PC            974
 9 XB            824
10 GBA           822
# i 21 more rows
# i Use `print(n = ...)` to see more rows
```

▾ 2. Cantidad de videojuegos por rating

```
%%R
datos %>%
filter(Rating != "") %>%
group_by(Rating) %>%
summarize(cantidad = n()) %>%
arrange(desc(cantidad))
```

```
    # A tibble: 8 × 2
      Rating cantidad
      <chr>     <int>
    1 E          3991
    2 T          2961
    3 M          1563
    4 E10+       1420
    5 EC            8
    6 K-A           3
    7 RP            3
    8 AO            1
```

▾ 3. Cantidad de videojuegos lanzados antes del año 2000 y cantidad de videojuegos lanzados después del año 2000

```
%%R
colnames(datos)
```

```
     [1] "Name"         "Platform"        "Year_of_Release" "Genre"
     [5] "Publisher"    "NA_Sales"        "EU_Sales"        "JP_Sales"
     [9] "Other_Sales"  "Global_Sales"    "Critic_Score"    "Critic_Count"
    [13] "User_Score"   "User_Count"      "Developer"       "Rating"
```

```
%%R
datos %>%
filter(Year_of_Release < 2000) %>%
select(Name) %>%
# Cantidad de juegos lanzados antes del 2000
count() -> cantidadPre2000

datos %>%
filter(Year_of_Release > 2000) %>%
select(Name) %>%
# Cantidad de juegos lanzados después del 2000
count() -> cantidadPos2000

c(cantidadPre2000,cantidadPos2000)
```

```
    $n
    [1] 1976

    $n
    [1] 14393
```

▾ 4. Plataforma cuyos videojuegos han recibido el mayor número de críticas

```
%%R
colnames(datos)
```

```
     [1] "Name"         "Platform"        "Year_of_Release" "Genre"
     [5] "Publisher"    "NA_Sales"        "EU_Sales"        "JP_Sales"
     [9] "Other_Sales"  "Global_Sales"    "Critic_Score"    "Critic_Count"
    [13] "User_Score"   "User_Count"      "Developer"       "Rating"
```

```R
%%R
datos %>%
filter(!is.na(Critic_Count) & Platform != "") %>%
group_by(Platform) %>%
summarize(criticas = sum(Critic_Count)) %>%
arrange(desc(criticas)) %>%
print(n=1)
```

```
# A tibble: 17 × 2
  Platform criticas
  <chr>       <int>
1 X360        38583
# i 16 more rows
# i Use `print(n = ...)` to see more rows
```

▼ 5. Cantidad de videojuegos por género

```R
%%R
datos %>%
filter(Genre != "") %>%
group_by(Genre) %>%
summarize(cantidad = n()) %>%
arrange(desc(cantidad))
```

```
# A tibble: 12 × 2
   Genre        cantidad
   <chr>           <int>
 1 Action           3370
 2 Sports           2348
 3 Misc             1750
 4 Role-Playing     1500
 5 Shooter          1323
 6 Adventure        1303
 7 Racing           1249
 8 Platform          888
 9 Simulation        874
10 Fighting          849
11 Strategy          683
12 Puzzle            580
```

▼ 6. Ventas totales en Norteamérica por parte de cada empresa desarrolladora

```R
%%R
colnames(datos)
```

```
 [1] "Name"          "Platform"       "Year_of_Release" "Genre"
 [5] "Publisher"     "NA_Sales"       "EU_Sales"        "JP_Sales"
 [9] "Other_Sales"   "Global_Sales"   "Critic_Score"    "Critic_Count"
[13] "User_Score"    "User_Count"     "Developer"       "Rating"
```

```R
%%R
datos %>%
group_by(Developer) %>%
summarize(ventasNA = sum(NA_Sales)) %>%
arrange(desc(ventasNA))
```

```
# A tibble: 1,697 × 2
   Developer          ventasNA
   <chr>                 <dbl>
 1 ""                    1238.
 2 "Nintendo"             231.
 3 "EA Sports"             96.8
 4 "EA Tiburon"            78.7
 5 "Ubisoft"               76.9
 6 "EA Canada"             66.2
 7 "Treyarch"              56.9
 8 "Visual Concepts"       56.0
 9 "Rockstar North"        55.2
10 "Ubisoft Montreal"      51.6
# i 1,687 more rows
# i Use `print(n = ...)` to see more rows
```

▼ 7. Ventas globales por género de videojuegos

```
%%R
colnames(datos)
```

```
     [1] "Name"          "Platform"         "Year_of_Release" "Genre"
     [5] "Publisher"     "NA_Sales"         "EU_Sales"        "JP_Sales"
     [9] "Other_Sales"   "Global_Sales"     "Critic_Score"    "Critic_Count"
    [13] "User_Score"    "User_Count"       "Developer"       "Rating"
```

```
%%R
datos %>%
group_by(Genre) %>%
summarize(ventasGLOB = sum(Global_Sales)) %>%
arrange(desc(ventasGLOB))
```

```
    # A tibble: 13 × 2
       Genre          ventasGLOB
       <chr>              <dbl>
     1 "Action"            1745.
     2 "Sports"            1332
     3 "Shooter"           1053.
     4 "Role-Playing"       934.
     5 "Platform"           828.
     6 "Misc"               803.
     7 "Racing"             729.
     8 "Fighting"           447.
     9 "Simulation"         390.
    10 "Puzzle"             243.
    11 "Adventure"          238.
    12 "Strategy"           174.
    13 ""                     2.42
```

▾ 8. Región en donde se ha obtenido la mayor cantidad de ventas de videojuegos

```
%%R
colnames(datos)
```

```
     [1] "Name"          "Platform"         "Year_of_Release" "Genre"
     [5] "Publisher"     "NA_Sales"         "EU_Sales"        "JP_Sales"
     [9] "Other_Sales"   "Global_Sales"     "Critic_Score"    "Critic_Count"
    [13] "User_Score"    "User_Count"       "Developer"       "Rating"
```

```
%%R
datos %>%
select(Global_Sales, NA_Sales, EU_Sales, JP_Sales) %>%
colSums()
```

```
    Global_Sales      NA_Sales      EU_Sales      JP_Sales
         8920.30       4402.62       2424.67       1297.43
```

▾ 9. Videojuego más vendido cada año

```
%%R
colnames(datos)
```

```
     [1] "Name"          "Platform"         "Year_of_Release" "Genre"
     [5] "Publisher"     "NA_Sales"         "EU_Sales"        "JP_Sales"
     [9] "Other_Sales"   "Global_Sales"     "Critic_Score"    "Critic_Count"
    [13] "User_Score"    "User_Count"       "Developer"       "Rating"
```

```
%%R
datos %>%
filter(Year_of_Release != "N/A") %>%
select(Name, Year_of_Release, Global_Sales) %>%
group_by(Year_of_Release) %>%
slice_max(Global_Sales) %>%
arrange(desc(Year_of_Release)) %>%
head(10)
```

```
    # A tibble: 10 × 3
    # Groups:   Year_of_Release [10]
       Name                            Year_of_Release Global_Sales
       <chr>                             <chr>                <dbl>
     1 Imagine: Makeup Artist            2020                  0.29
```

```
 2 Phantasy Star Online 2 Episode 4: Deluxe Package 2017                    0.04
 3 FIFA 17                                           2016                    7.59
 4 Call of Duty: Black Ops 3                         2015                    14.6
 5 Grand Theft Auto V                                2014                    12.6
 6 Grand Theft Auto V                                2013                    21.0
 7 Call of Duty: Black Ops II                        2012                    13.8
 8 Call of Duty: Modern Warfare 3                    2011                    14.7
 9 Kinect Adventures!                                2010                    21.8
10 Wii Sports Resort                                 2009                    32.8
```

▾ 10. Videojuego(s) más viejo(s) y videojuego(s) más nuevo(s) de acuerdo al registro

```
%%R
colnames(datos)

    [1] "Name"         "Platform"      "Year_of_Release" "Genre"
    [5] "Publisher"    "NA_Sales"      "EU_Sales"        "JP_Sales"
    [9] "Other_Sales"  "Global_Sales"  "Critic_Score"    "Critic_Count"
    [13] "User_Score"   "User_Count"    "Developer"       "Rating"
```

```
%%R
datos %>%
filter(Year_of_Release != "N/A") %>%
select(Name, Year_of_Release) %>%
arrange(desc(Year_of_Release)) %>%
head(1) -> masReciente

datos %>%
filter(Year_of_Release != "N/A") %>%
select(Name, Year_of_Release) %>%
arrange(desc(Year_of_Release)) %>%
tail(1) -> masViejo

c(masReciente, masViejo)

    $Name
    [1] "Imagine: Makeup Artist"

    $Year_of_Release
    [1] "2020"

    $Name
    [1] "Checkers"

    $Year_of_Release
    [1] "1980"
```

▾ 11. Videojuego más vendido globalmente de tu género favorito

```
%%R
datos %>%
filter(Genre == "Shooter") %>%
select(Name, Global_Sales) %>%
head(1)

          Name Global_Sales
    1 Duck Hunt       28.31
```

✓ 0 s   se ejecutó 18:50   ● ✕