



## SEMINARI 9.

# Interacció i Animació en R (Respostes)

#### 1. OBJECTIUS

Conèixer i utilitzar eines d'interacció i animació en R com plotly, gganimate, shiny per a crear entorns interactius i d'animació en Visualització de Dades.

## 2. PART 1. Entorns de Navegació

En aquesta primera part crearem entorns interactius de navegació amb plotly, a partir de ggplot() o directament amb plot\_ly(). Treballarem amb el data set Poblacio\_Mundial\_1960-2022.csv de la població mundial des del 1960 al 2022 segmentat per país i continent.

Abans de fer els exercicis, dos passos previs:

1. Carregar les llibreries que utilitzarem.

```
> library(tidyverse)
> install.packages('plotly') # Càrrega paquet plotly (1a vegada)
> library(plotly)
```

#### 2. Llegir el dataset.

```
> setwd("C:/Users/enric/Documents/R")
> PWorld <- read.csv('./Poblacio_Mundial_1960-2022.csv')
> str(PWorld)
>
> setwd("C:/Users/enric/R")
> Pworld <- read.csv('./Poblacio_Mundial_1960-2022.csv')
> str(Pworld)
'data.frame': 16695 obs. of 4 variables:
$ Pais.Nom : chr "Aruba" "Aruba" "Aruba" "Aruba" ...
$ Continent: chr "America" "America" "America" "America" ...
$ Any : int 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 ...
$ Poblacio : num 54608 55811 56682 57475 58178 ...
> |
```

#### Variables:

- Pais.Nom: Nom del país observat o nom de la zona geogràfica que agrupa varis països.
- Continent: Continent del país. En alguns registres agafa el valor "Zone" referint-se a una zona geogràfica.
- Any: Any de l'observació
- Poblacio: Nombre d'habitants del país i any de l'observació





#### **EXERCICIS:**

1.- Mostra el codi per a definir una gràfica interactiva de barres i una gràfica interactiva de línies de l'evolució temporal de la població en els cinc continents, excloent les zones geogràfiques o polítiques(eliminar registres amb Continent="Zone").

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING. Filtrar (eliminar) zones geogràfiques i posteriorment agrupar per continents i sumar les poblacions:

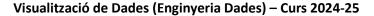
```
> PMundial <- PWorld %>% filter(Continent != "Zone") %>%
group_by(Any,Continent) %>% summarize(NHab = sum(Poblacio))
> PMundial
```

```
> PMundial <- Pworld %>% filter(Continent != "Zone") %>% group_by(Any,Continent) %>% summarize(NHab=sum(Poblacio)) `summarise()` has grouped output by 'Any'. You can override using the `.groups` argument.
# A tibble: 315 \times
# Groups:
      Any Continent
                                   NHab
   <u>1</u>960 Africa
                            276<u>157</u>583
     1960 America
                            422093874
     <u>1</u>960 Asia
                          1934<u>891</u>710
     <u>1</u>960 Europe
                            493<u>657</u>644
    <u>1</u>960 Oceania
<u>1</u>961 Africa
                             15466870
     <u>1</u>961 America
                            431722162
    1961 Asia
                          1955905117
 9 <u>1</u>961 Europe
10 <u>1</u>961 Oceania
                            15795468
# i 305 more rows
# i Use `print(n = ...)` to see more rows
```

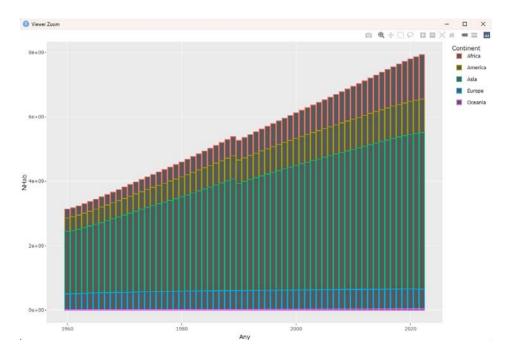
#### PAS 2: Pot haver diferents formes de visualitzar:

#### **DIAGRAMA DE BARRES INTERACTIVA:**

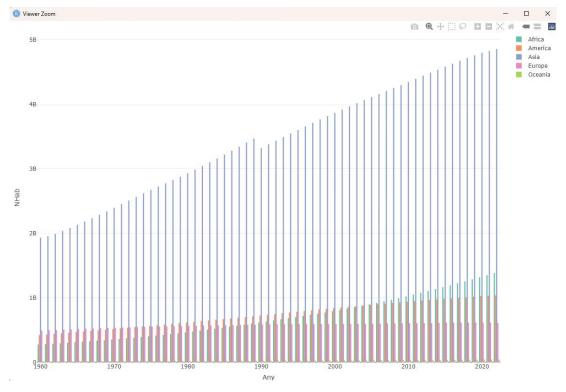
```
# DIAGRAMA AMB BARRES AMB ggplotly
> gplotPMundial<-ggplot(PMundial, aes(x=Any, y=NHab, color=Continent))
+ geom_col()
> ggplotly(gplotPMundial)
```







- # DIAGRAMA AMB BARRES AMB plotly
- > plot\_ly(PMundial,x=~Any, y=~NHab, color=~Continent) %>% add\_bars()
- > plot\_ly(PMundial,x=~Any, y=~NHab, color=~Continent, type ='bar')



#### **GRÀFICA DE LÍNIES INTERACTIVA:**

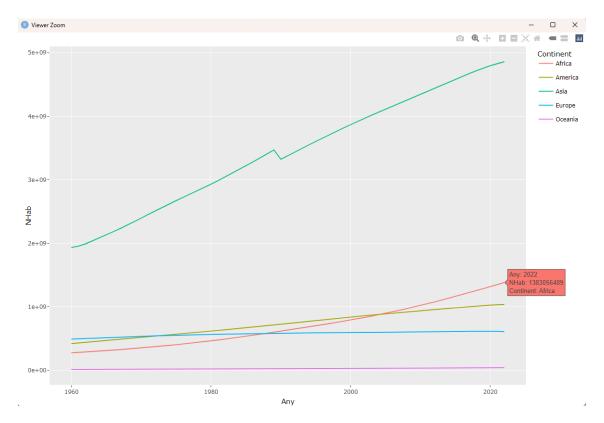
```
# GRÀFICA DE LÍNIES AMB ggplotly
```

- > ggplotPMundial <- ggplot(PMundial, aes(x=Any, y=NHab,</pre>
- color=Continent)) + geom\_line()
- > ggplotly(ggplotPMundial)





# GRAFICA DE LÍNIES AMB plotly
> plot\_ly(PMundial,x=~Any, y=~NHab, color=~Continent, type='scatter',
mode='line')



#### Respon a les següents preguntes veient la gràfica:

a) Descriu tres característiques importants que es veuen en la gràfica de línies.

#### **RESPOSTA:**

- 1. La població del continent asiàtic predomina respecte els altres continents en tots els anys (1934M 1960 a 4858M el 2022). Té una davallada l'any 1990, però després es va recuperant.
- 2. La resta de continents té un creixement més moderat o una mica de decreixement. Destaca Africa a partir del s.XXI que te un lleuger creixement (794M el 2000 a 1383M el 2022), sobrepassant la població del continent americà.
- 3. El continent europeu té un creixement molt baix al llarg dels anys (49M 1960 a 61M 2022.
- 4. Amèrica tenia menys població que Europa el 1960 (422M America per 493M Europa) i acaba el 2022 amb una diferència de 40M (1033M Amèrica a 610M Europa)35-54 anys del 2006 al 2015, 55-74 anys del 1985 al 2002.
- b) Enumera els 3 continents amb menys població l'any 2012. Continent i població en ordre creixent.

#### **RESPOSTA:**

- 1. Oceania 37475300
- 2. Europe 608142862





- 3. America 958228468
- c) En quin parell d'anys any fa una davallada a població d'Asia?. Indica la població els dos anys.

#### **RESPOSTA:**

- 1989 34070598904
- 1990 3323719552
- d) A partir de quin any la població del continent africà supera al d'Amèrica?..

RESPOSTA: I fent un zoom en la zona propera al creuament Africa-America

L'any 2005:

- Africa 899781392
- America 891694507
- 2.- Fes una gràfica interactiva de línies de l'evolució temporal de població dels països del continent americà.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING. Abans de crear la gràfica cal filtrar els registres amb valor de la variable Continent "America".

- > PMAmerica <- PMundial %>% filter(Continent == "America")
- > PMAmerica

> PMAmerica

11

```
> PMAmerica <- PWorld %>% filter(Continent !="Zone", Continent == "America")
```

59106

58816

```
Pais.Nom Continent Any Poblacio
1
                          America 1960
                                          54608
                  Aruba
2
                          America 1961
                                          55811
                  Aruba
3
                  Aruba
                          America 1962
                                          56682
4
                  Aruba
                          America 1963
                                          57475
5
                  Aruba
                          America 1964
                                          58178
6
                  Aruba
                          America 1965
                                          58782
7
                          America 1966
                  Aruba
                                          59291
8
                  Aruba
                          America 1967
                                          59522
9
                          America 1968
                                          59471
                  Aruba
10
                  Aruba
                          America 1969
                                          59330
```

PAS 2: Visualitzar per gràfica interactiva de línies

Aruba

Aruha

#### **GRÀFICA INTERACTIVA DE LÍNIES:**

```
# GRAFICA DE LÍNIES amb ggplotly
> gplotPMAmerica <- ggplot(PMAmerica, aes(x=Any, y=Poblacio,
color=Pais.Nom)) + geom_line()
> ggplotly(gplotPMAmerica)
```

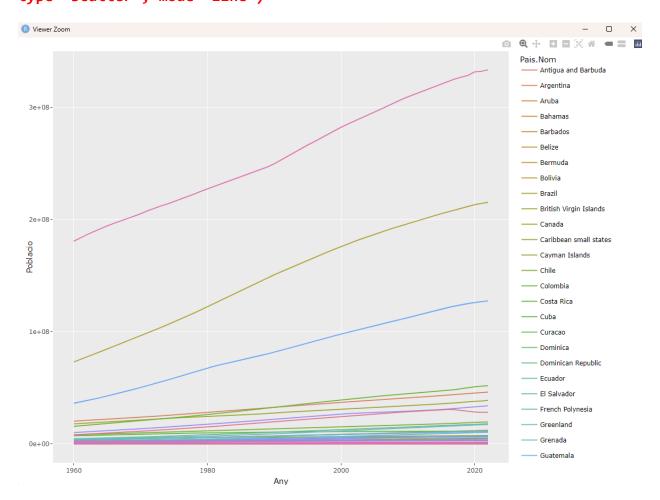
America 1970

America 1071





# GRAFICA DE LÍNIES amb plotly
> plot\_ly(PMAmerica,x=~Any, y=~Poblacio, color=~Pais.Nom,
type='scatter', mode='line')



Respon a les següents preguntes veient la gràfica:

a) Quins tres països predominen en població durant tots els anys?. Trobes a faltar alguns països que per extensió podrien tenir més població?.

#### **RESPOSTA:**

- 1) USA
- 2) Brazil,
- 3) Mexico

Es troba a faltar països com Canadà i Argentina que per extensió podrien tenir més població.

b) A partir del 2016 hi ha un país americà que té un descens de població. Digues quin país és, quina població tenia els anys 2016 i 2022.

RESPOSTA: Venezuela, i 28317M a 2022.

c) Dels tres països amb més població entre 1960 i 2022, quin és el país amb més creixement de població entre 1960 i 2022?. Dona les diferencies en milions de persones dels tres.





#### **RESPOSTA:**

- 1) USA: 333M (2022) 180M (1960) = 153M
- 2) Brazil: 215M (2022) 73M (1960) = 142M
- 3) Mexico: 127M (2022) 36M (1960) = 91M
- 3.- Defineix una gràfica interactiva de línies de població dels 10 països més poblats del món.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING: Afegim la variable rank que correspon al ranking decreixent per països del nombre de suïcidis i seleccionem els 10 primers de cada any segons la variable rank.

```
> PWorld_formatted <- PWorld %>% filter(Continent != "Zone") %>% group_by(Any)
%>% mutate(rank = rank(-Poblacio)) %>% filter(rank <= 10)</pre>
```

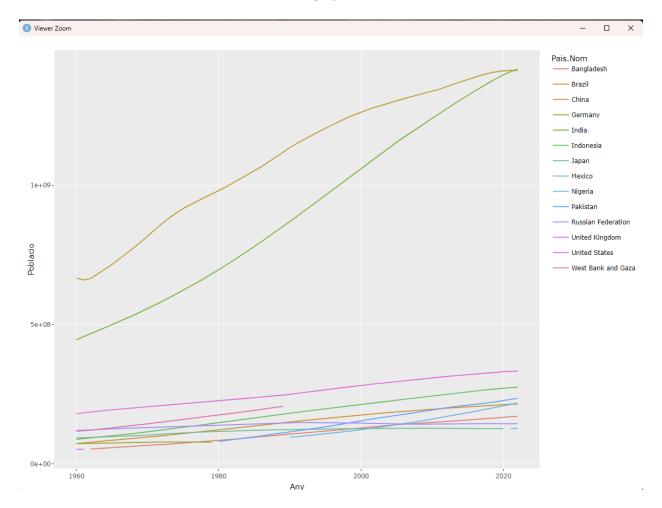
```
> PWorld_formatted <- PWorld %>% filter(Continent != "Zone") %>% group_by(Any) %>% mutate(rank = rank(-Poblacio)) %>%
                                filter(rank <=10)
> PWorld_formatted
# A tibble: 630 x 5
   Pais.Nom Continent Any Poblacio rank
                          <int>
                <chr>>
                                     <db1>
                            <u>1</u>962 53<u>461</u>661
 1 Bangladesh Asia
 2 Bangladesh Asia
                            <u>1</u>963 55<u>094</u>115
 3 Bangladesh Asia
                            1964 56774465
                                               10
 4 Bangladesh Asia
                            1965 58500159
                                               10
 5 Bangladesh Asia
                            1966 60265259
                                               10
 6 Bangladesh Asia
                            <u>1</u>967 62<u>104</u>488
                                               10
 7 Bangladesh Asia
                            1968 63<u>995</u>652
 8 Bangladesh Asia
                            <u>1</u>969 65<u>866</u>908
                                               10
 9 Bangladesh Asia
                            1970 67541860
                                               10
10 Bangladesh Asia
                            <u>1</u>971 68<u>376</u>204
                                               10
# i 620 more rows
# i Use `print(n = ...)` to see more rows
>
```

#### PAS 2: Visualitzar la gràfica interactiva de línies:

```
# GRÀFICA AMB LÍNIES AMB ggplotly
> ggplotPM_formatted <- ggplot(PWorld_formatted, aes(x=Any, y=Poblacio, color=Pais.Nom)) + geom_line()
> ggplotly(ggplotPM_formatted)
# GRÀFICA AMB LÍNIES AMB plotly
> plot_ly(PWorld_formatted,x=~Any, y=~Poblacio, color=~Pais.Nom, type='scatter', mode='line')
```







Respon a les següents preguntes veient la gràfica:

a) Quins tres països predominen en població durant tots els anys?. Mirant la gràfica, quin país ha tingut un creixement més gran entre 1960 i 2022?.

#### **RESPOSTA:**

- 1) China
- 2) Índia
- 3) USA

Índia és la que ha tingut un major creixement entre 1960 i 2022.

b) Quin és el país més poblat del món a 2022?. Digues el país i la població en milions de persones.

RESPOSTA: India amb 1417M a 2022

c) Calcula la densitat de població (Poblacio / superfície en km2) l'any 2022 pels tres països més poblats del món i mira si es manté el ranking per densitat de població. població entre 1960 i 2022, quin és el país amb més creixement de població entre 1960 i 2022?. Dona les diferencies en milions de persones dels tres.

**RESPOSTA:** 





- INDIA: Població: 1.417.173.173 (2022) Superfície 3.287.000 km2 Densitat: 431.14 hab./km2
- CHINA: Població: 1.412.175.000 (2022) Superfície: 9.597.000 km2 Densitat: 147,14 hab./km2
- USA: Població: 333.287.557 (2022) Superfície: 9.867.000 km2 Densitat: 33,77 hab./km2.
- 4.- Visualitza dues gràfiques interactives de línies per a comparar-les.
  - a) Representa una gràfica interactiva de línies amb la població dels 15 països menys poblats de l'Asia. Quins són els tres països menys poblats el 2022 amb la diferència de població entre 2022 i 1960?.

#### RESPOSTA:

**PAS 1:** DATA MASSAGING: Afegim la variable rank que correspon al ranking decreixent per països del nombre de suïcidis i seleccionem els 10 primers de cada any segons la variable rank.

```
> PWorldAsia_formatted <- PWorld %>% filter(Continent != "Zone", Continent ==
"Asia") %>% group_by(Any) %>% mutate(rank = rank(Poblacio)) %>%
filter(rank <= 15)</pre>
```

```
> PworldAsia_formatted <- Pworld %>% filter(Continent != "Zone", Continent == "Asia") %>%
group_by(Any) %>% mutate(rank = rank(Poblacio)) %>%
                              filter(rank <=15)
> PWorldAsia_formatted
# A tibble: 945 x 5
   Pais.Nom
                          Continent Any Poblacio rank
   <chr>
                                    <int>
                                              <db1> <db1>
                          <chr>
                                             <u>133</u>426
 1 United Arab Emirates Asia
                                     1960
 2 United Arab Emirates Asia
                                     1961
                                             140984
 3 United Arab Emirates Asia
                                      <u>1</u>962
                                             <u>148</u>877
                                                         8
4 United Arab Emirates Asia
                                             <u>157</u>006
                                                         8
                                      1963
 5 United Arab Emirates Asia
                                     1964
                                             165305
 6 United Arab Emirates Asia
                                     <u>1</u>965
                                              <u>173</u>797
                                                         8
                                      1966
                                                         8
 7 United Arab Emirates Asia
                                              <u>182</u>509
 8 United Arab Emirates Asia
                                      1967
                                              191404
                                                         8
 9 United Arab Emirates Asia
                                      1968
                                              <u>213</u>582
                                                         9
                                      <u>1</u>969
10 United Arab Emirates Asia
                                             <u>253</u>261
                                                        10
# i 935 more rows
# i Use `print(n = ...)` to see more rows
> |
```

#### PAS 2: Visualitzar per gràfica interactiva de línies:

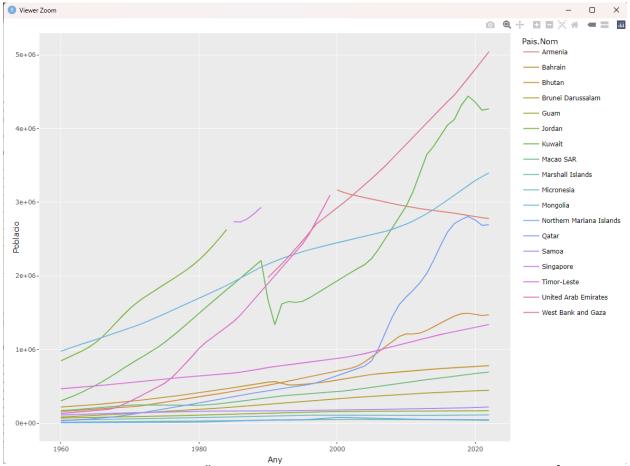
```
# GRAFICA DE LÍNIES amb ggplotly
> ggplotAsia_formatted <- ggplot(PWorldAsia_formatted, aes(x=Any,
y=Poblacio, color=Pais.Nom)) + geom_line()
> ggplotly(ggplotAsia_formatted)

# GRAFICA DE LÍNIES amb plot_ly
> plot_ly(PWorldAsia_formatted, x=~Any, y=~Poblacio, color=~Pais.Nom,
type='scatter', mode='line')
```



#### Visualització de Dades (Enginyeria Dades) - Curs 2024-25





RESPOSTA TRES PAÏSOS MENYS POBLATS D'ASIA AMB LA DIFERÊNCIA DE POBLACIÓ ENTRE 2022 i 1960:

- 1) MARSHALL ISLANDS: 41569 (2022) 8702 (1960) = 32867
- 2) MICRONESIA: 114164 (2022) 42986 (1960) = 71178
- 3) GUAM: 171774 (2022) 72374 (1960) = 99400
- b) Representa una gràfica interactiva de línies amb la població dels 15 països menys poblats d'Europa. Quins són els tres països menys poblats el 2022 amb la diferència de població entre 2022 i 1960?.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING: Afegim la variable rank que correspon al ranking decreixent per països del nombre de suïcidis i seleccionem els 10 primers de cada any segons la variable rank.

- > PWorldEuropa\_formatted <- PWorld %>% filter(Continent != "Zone", Continent
  == "Europe") %>% group\_by(Any) %>% mutate(rank = rank(Poblacio)) %>%
  filter(rank <= 15)</pre>
- > PWorldEuropa\_formatted



# GRÀFICA DE LÍNIES amb plot\_ly

type='scatter', mode='line')

#### Visualització de Dades (Enginyeria Dades) – Curs 2024-25



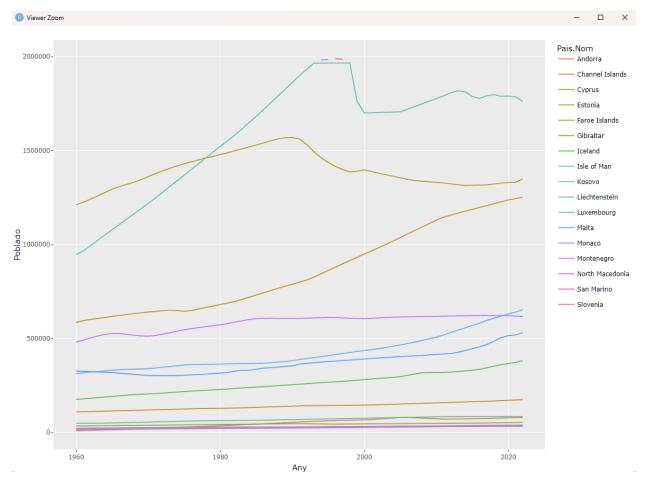
```
> PWorldEuropa_formatted <- PWorld %>% filter(Continent != "Zone", Continent == "Europe") %>%
group_by(Any) %>% mutate(rank = rank(Poblacio)) %>%
                                                                                    filter(rank <=15) %
 > PWorldEuropa_formatted
 # A tibble: 945 × 5
   Pais.Nom Continent Any Poblacio rank
                      <int> <db1> <db1>
    <chr> <chr>
                       <u>1</u>960 <u>9</u>443
<u>1</u>961 <u>10</u>216
<u>1</u>962 <u>11</u>014
 1 Andorra Europe
 2 Andorra Europe
 3 Andorra Europe
                      <u>1</u>963 <u>11</u>839
 4 Andorra Europe
                                              1
                      <u>1</u>964
<u>1</u>965
 5 Andorra Europe
6 Andorra Europe
                                  <u>12</u>690
                                   <u>13</u>563
 7 Andorra Europe
                        <u>1</u>966
                                  <u>14</u>546
                        <u>1</u>967
 8 Andorra Europe
                                  <u>15</u>745
                                              1
                       <u>1</u>968
9 Andorra Europe
10 Andorra Europe
                                   <u>17</u>079
                         <u>1</u>969
                                   <u>18</u>449
# i 935 more rows
# i Use `print(n = ...)` to see more rows
>
PAS 2: Visualitzar per gràfica interactiva de línies:
# GRÀFICA DE LÍNIES amb ggplotly
> ggplotEuropa_formatted <- ggplot(PWorldEuropa_formatted, aes(x=Any,
y=Poblacio, color=Pais.Nom)) + geom_line()
> ggplotly(ggplotEuropa_formatted)
```

> plot\_ly(PWorldEuropa\_formatted,x=~Any, y=~Poblacio, color=~Pais.Nom,



#### Visualització de Dades (Enginyeria Dades) – Curs 2024-25





# RESPOSTA TRES PAÏSOS MENYS POBLATS D'EUROPA AMB LA DIFERÊNCIA DE POBLACIÓ ENTRE 2022 i 1960:

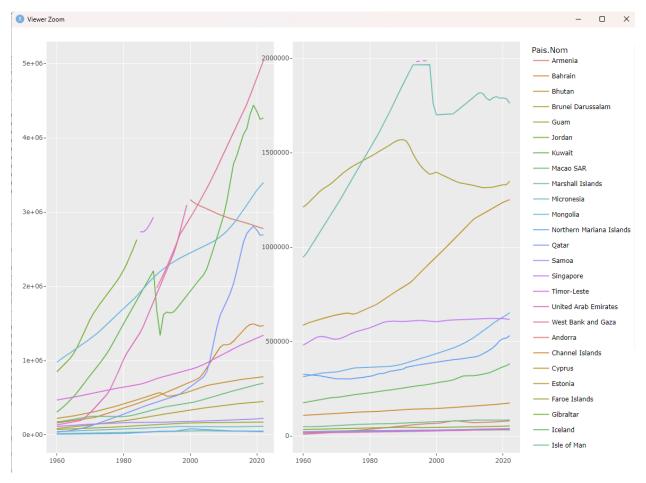
- 1) GIBRALTAR: 32649 (2022) 21822 (1960) = 10827
- 2) SANT MARINO: 33660 (2022) 15556 (1960) = 18104
- 3) MONACO: 36469 (2022) 21797 (1960) = 14672
- c) Dibuixa les dues gràfiques anteriors (Asia esquerra, Europa dreta) en una mateixa pantalla i comenta 3 idees que es poden deduir veient les dues gràfiques.

#### **RESPOSTA:**

> subplot(ggplotAsia\_formatted, ggplotEuropa\_formatted)







#### **RESPOSTA:**

- L'ordre de magnitud de la gràfica de població (eix y) en Asia es més del doble que d'Europa.
- El nivell de creixement al llarg dels anys en Asia és superior al d'Europa.
- Hi ha més canvis en el ranking en els aïsos d'Asia que en els d'Europa, on hi ha més estabilitat en les posicions.
- Kosovo pateix una davallada de població l'any 1993 producte de la guerra dels Balcans.
- Kuwait pateix una davallada de població l'any 1990 i 1991 producte de la guerra del Kuwait i del Golf.
- 5.- Guarda el darrer plot en un fitxer en format HTML en la carpeta de treball (wd) per a poder-se executar en un navegador. Digues què ocupa la versió estesa i la reduïda i prova d'executar-les en un navegador.

#### **RESPOSTA:**

```
# Funció widget_size que guarda objecte plotly a format HTML
widget_file_size <- function(p) {
    d <- getwd()
    withr::with_dir(d, htmlwidgets::saveWidget(p, "index.html"))
    f <- file.path(d, "index.html")
    mb <- round(file.info(f)$size / 1e6, 3)
    message("File is: ", mb," MB")
}</pre>
```





```
# Generar fitxer index.html
> plot <- subplot(ggplotAsia_formatted, ggplotEuropa_formatted)</pre>
> widget file size(plot)
                               # Versió completa HTML
> widget file size(partial bundle(plot)) # Versió reduida HTML
      > widget_file_size <- function(p) {
            d <- getwd()
            withr::with_dir(d, htmlwidgets::saveWidget(p, "index.html"))
f <- file.path(d, "index.html")</pre>
            mb <- round(file.info(f)$size / 1e6, 3)</pre>
            message("File is: ", mb," MB")
      + }
      > plot <- subplot(ggplotAsia_formatted, ggplotEuropa_formatted)</pre>
     > widget_file_size(plot)
      File is: 4.019 MB
      > widget_file_size(partial_bundle(plot))
      File is: 1.355 MB
      >
```

També es pot exportar amb HTML amb el botó Export del visualitzador interactiu de R.

6.- Visualitza un scatter plot 3D sobre anys, població i PIB combinant el dataset que treballem Poblacio\_Mundial\_1960-2022.csv amb el data set PIB\_Mundial\_1960\_2020.csv que conté dades del PIB de països del món des del 1960 al 2020.

Comenta mirant la gràfica algunes idees que vegis. Mira si pots trobar algun tipus de correlació entre el PIB, Continent i la població dels països.

Us podeu carregar la llibreria plyr per a fer joins entre datasets.

```
> install.packages('plyr') # Càrrega paquet plyr (1a vegada)
> library(plyr)
>

RESPOSTA:
Ara llegim el data set Poblacio_Mundial_1960-2022.csv
> setwd("C:/Users/enric/R")
> PWorld <- read.csv('./Poblacio_Mundial_1960-2022.csv')
>

Després llegim el data set PIB_Mundial_1960-2020.csv
> setwd("C:/Users/enric/R")
```

> PIBWorld <- read.csv('./PIB Mundial 1960-2020.csv')</pre>

> str(PIBWorld)

>



#### Visualització de Dades (Enginyeria Dades) - Curs 2024-25



**PAS 1:** DATA MASSAGING: Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països, no zones geogràfiques.

```
PMon <- PWorld %>% filter(Continent != "Zone")
> SWorldPIB
```

PAS 2: DATA MASSAGING: Fem un join entre els dos datasets: per país i any per a agrupar població i GDP (PIB) i eliminem registres amb PIB indefinit (PIB=="NA")..

```
> Resultat <- join(PMon,PIBWorld) %>% filter(PIB != "NA")
> str(Resultat)
```

```
> Resultat <- join(PMon,PIBWorld) %>% filter(PIB != "NA")
Joining by: Pais.Nom, Any
> str(Resultat)
'data.frame': 5124 obs. of 5 variables:
$ Pais.Nom : chr "Australia" "Australia" "Australia" "Australia" ...
$ Continent: chr "Oceania" "Oceania" "Oceania" "Oceania" ...
$ Any : int 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 ...
$ Poblacio : num 10276477 10483000 10742000 10950000 11167000 ...
$ PIB : num 1.86e+10 1.97e+10 1.99e+10 2.15e+10 2.38e+10 ...
```

#### PAS 3: SCATTER 3D PLOT:

51

1960

1970

Any 1990

Niewer Zoom

> plot <- plot\_ly(Resultat, x=~Any, y=~Poblacio, z=~PIB,
color=~Pais.Nom, type='scatter3d', mode='markers')
> plot

20T 15T



1.2B

0.8B

0.6B

0.4B

0.2B





#### **COMENTARIS:**

- 1. La majoria de països amb una baixa població mantenen un PIB baix al llarg dels anys. De tot aquest conjunt, destaquen una mica per PIB United Kingdom, France, Italy
- 2. Quatre països surten una mica de la norma: China, India, Unites States i Japan.
- 3. Dos països (China i Índia) destaquen per població, però no per PIB, en que predomina United States.
- 4. El que té una evolució més alta i predominant del PIB amb una mitjana població és US.
- 5. Japó té una evolució del PIB més modesta però amb una població força baixa. Potser és el millor país dins la majoria de països, tot i que la seva evolució del PIB ha tingut baixades els anys 1996 i 2013.
- 6. En segon lloc, la China ha tingut una alta evolució del PIB (no tant com United States) amb una població molt alta.
- 7. L'India ha tingut una evolució molt creixent de població, situant-se amb un PIB del pitjor dels 4 països predominants, i inclús més baix que algun país com UK i força igualat a France l'any 2020.

#### PAS 4: Guardar-el en html:

```
# Funció widget size que guarda objecte plotly a format HTML
widget_file_size <- function(p) {</pre>
  d <- getwd()</pre>
 withr::with dir(d, htmlwidgets::saveWidget(p, "scatterPlot3D.html"))
 f <- file.path(d, "scatterPlot3D.html")</pre>
 mb <- round(file.info(f)$size / 1e6, 3)</pre>
 message("File is: ", mb," MB")
}
# Generar fitxer scatterPlot3D.html
> widget_file_size(plot) # Versió completa HTML
        > widget_file_size(plot)
                                       # Versió completa HTML
        File is: 1.355 MB
        Warning messages:
        1: In RColorBrewer::brewer.pal(N, "Set2") :
          n too large, allowed maximum for palette Set2 is 8
        Returning the palette you asked for with that many colors
        2: In RColorBrewer::brewer.pal(N, "Set2"):
          n too large, allowed maximum for palette Set2 is 8
        Returning the palette you asked for with that many colors
        >
```





## 3. PART 2. Animate Bar Race Ranking

En aquesta segona part del seminari intentarem fer gràfiques el tipus *Animate Bar Race Ranking* sobre el data set del padró de noms (Poblacio\_Mundial\_1960-2022.csv).

#### **RESPOSTA:**

**EXERCICIS:** 

Primer de tot, instal·larem les llibreries que necessitem:

```
> library(tidyverse)
> install.packages('gifski') # Càrrega paquet gifski (primera vegada)
> library(gifski) # Generació fitxer animació GIF
> install.packages('av') # Càrrega paquet av (primera vegada)
> library(av) # Generació fitxer animació AVI
> install.packages('gganimate') # Càrrega paquet gganimate (1a vegada)
> library(gganimate) # Generació de frames i fitxer animació
> Llegim el data set:
```

# > setwd("C:/Users/enric/Documents/R") > PWorld <- read.csv('./Poblacio\_Mundial\_1960-2022.csv')</pre>

# 1.- Crea un *Animate Bar Race Ranking* sobre el primer quartil (Q1) dels països europeus MENYS poblats per any.

Respecte els valors donats en les transparències, modifica els valors transition\_lenght (per exemple 4000) i state\_lenght (per exemple 400) de la comanda transition\_states() i el valor del paràmetre del nframes (per exemple a 500) de la comanda animate(objecte,nframes,...). Com canvia l'animació?

COMENTARI: Comenta l'evolució d'Andorra al llarg dels anys.

**PAS 1:** DATA MASSAGING: Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països d'Europa.

```
> PMon <- PWorld %>% filter(Continent != "Zone", Continent ==
"Europe")
```



#### Visualització de Dades (Enginyeria Dades) - Curs 2024-25

```
> PWorld <- read.csv('./Poblacio_Mundial_1960-2022.csv')</pre>
> PMon <- PWorld %>% filter(Continent != "Zone", Continent == "Europe")
> PMon
       Pais.Nom Continent Any Poblacio
        Albania Europe 1960 1608800
1
2
                   Europe 1961 1659800
        Albania
        Albania Europe 1961 1659800
Albania Europe 1962 1711319
Albania Europe 1963 1762621
Albania Europe 1964 1814135
3
4
5
6
        Albania Europe 1965 1864791
7
        Albania Europe 1966 1914573
        Albania Europe 1967 1965598
8
        Albania Europe 1968 2022272
9
```

**PAS 2:** Afegim la variable *rank* que correspon al ranking creixent per països de la població (eliminem el signe negatiu del paràmetre de la funció rank()) i seleccionem els Q1 països per cada any segons la variable rank. El guardem en l'objecte *PMon formatted*:

```
> PMon formatted <- PMon %>%
 group by(Any) %>%
 mutate(rank = rank(Poblacio)) %>%
 filter(rank <= 0.25 * max(rank)) #Q1
  > PMon_formatted <- PMon %>%
        group_by(Any) %>%
        # The * 1 makes it possible to have non-integer ranks while sliding
        mutate(rank = rank(Poblacio)) %>%
        filter(rank <= 0.25 * max(rank)) #Q1
  > PMon_formatted
  # A tibble: 756 x 5
  # Groups: Any [63]
     Pais.Nom Continent Any Poblacio rank
     <chr> <chr> <int>
                                 <db1> <db1>
   1 Andorra Europe <u>1</u>960
                                  9443
                                           1
   2 Andorra Europe
                         1961
                                 10216
                                           1
   3 Andorra Europe
                         1962
                                 11014
                                           1
   4 Andorra Europe
                         1963
                                           1
                                 11839
   5 Andorra Europe
                         1964
                                 12690
                                           1
                                           1
   6 Andorra Europe
                        1965
                                 13563
   7 Andorra Europe
                         1966
                                 14546
                                           1
   8 Andorra Europe
                         1967
                                 15745
                                           1
   9 Andorra Europe
                         <u>1</u>968
                                 <u>17</u>079
                                           1
  10 Andorra Europe
                         <u>1</u>969
                                 <u>18</u>449
  # i 746 more rows
  # i Use `print(n = ...)` to see more rows
  >
```

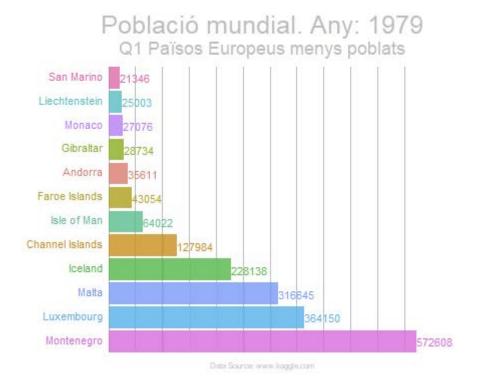
**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar els frames estàtics de les gràfiques per anys:



```
width = 0.9), alpha = 0.8, color = NA) +
     geom text(aes(y=0, label = paste(Pais.Nom," ")), vjust=0.2, hjust=1) +
     geom_text(aes(y = Poblacio, label = Poblacio, hjust = 0)) +
     coord_flip(clip = "off", expand = FALSE) +
      scale y continuous(labels = scales::comma) +
     scale x reverse() +
      guides(color = "none", fill = "none") +
     theme(axis.line=element_blank(),
     axis.text.x=element_blank(),
     axis.text.y=element_blank(),
     axis.ticks=element_blank(),
     axis.title.x=element_blank(),
      axis.title.y=element_blank(),
      legend.position="none",
      panel.background=element blank(),
     panel.border=element_blank(),
     panel.grid.major=element_blank(),
     panel.grid.minor=element_blank(),
     panel.grid.major.x = element_line( linewidth=.1, color="grey" ),
     panel.grid.minor.x = element line( linewidth=.1, color="grey" ),
     plot.title=element text(size=25, hjust=0.5, face="bold", colour="grey",
     plot.subtitle=element text(size=18,
                                                               face="italic",
                                               hjust=0.5,
     color="grey"),
     plot.caption
                      =element_text(size=8,
                                                hjust=0.5,
                                                               face="italic",
     color="grey"),
     plot.background=element blank(),
     plot.margin = margin(2,2, 2, 4, "cm")) +
     transition_states(Any, transition_length =
                                                              state_length=1,
                                                        4,
     wrap=FALSE) +
     view_follow(fixed_x = TRUE) +
     labs(title = 'Població mundial. Any: {closest_state}',
           subtitle = "Q1 Països Europeus menys poblats",
           caption = "Data Source: www.kaggle.com")
> anim
PAS 4: A partir de l'objecte anim, creem els fitxers d'animació GIF i AVI:
# Exportar frames a fitxer GIF
> animate(anim, 200, fps = 20, width = 1200, height = 1000,
        renderer = gifski_renderer("Q1Europeus.gif"), end_pause =
                                                                          15,
        start_pause = 15)
# Exportar frames a fitxer AVI
> animate(anim, 200, fps = 20, width = 1200, height = 1000,
        renderer = av renderer("Q1Europeus.avi"), end pause = 15, start pause
       = 15)
```







#### **ALTERNATIVA: Llibreria ddplot**

Primer de tot, instal·larem les llibreries que necessitem:

```
> install.packages('ddplot') # Càrrega paquet ddplot (primera vegada)
> library(ddplot) # Generació Bar Chart Race
```

PAS 1: DATA MASSAGING: Ens quedem només amb països d'Europa.

> PMon <- PWorld %>% filter(Continent == "Europe")

**PAS 2:** Afegim la variable *rank* que correspon al ranking creixent per països de la població, i seleccionem els Q1 països de cada any segons la variable rank. El guardem en l'objecte *PMon\_formatted*:

```
> PMon_formatted <- PMon %>%
  group_by(Any) %>%
  mutate(rank = rank(Poblacio)) %>%
  filter(rank <= 0.25 * max(rank))</pre>
```

**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar els frames estàtics de les gràfiques per anys:

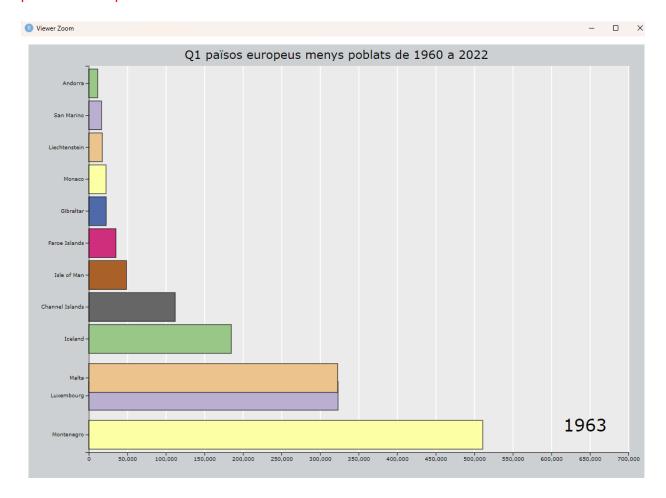
```
> barChartRace(PWorld_formatted, x="Poblacio", y="Pais.Nom", time="Any", sort
= "ascending", title = "Q1 països europeus menys poblats de 1960 a 2022")
```

**COMENTARI:** Andorra era el país menys poblat d'Andorra els anys 60. I els anys 70 va tenir un augment significatiu de població fins a passar al sisè menys poblat d'Europa.

#### Visualització de Dades (Enginyeria Dades) – Curs 2024-25



En general no hi ha han hagut grans canvis de població en la majoria dels Q1 països menys poblats d'Europa entre 1960 i 2022.



2.- Defineix un *Animate Bar Race Ranking* del segon quartil (Q2) de països MÉS poblats d'Amèrica. Comenta com evolucionen els països en general en l'animació. En quin any apareix Cuba i en quina posició acaba l'any 2022?.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING: Fem el mateix pas que l'anterior. Ens quedem només amb països d'America.

> PMon <- PWorld %>% filter(Continent == "America")

**PAS 2:** Afegim la variable *rank* que correspon al ranking decreixent per països de la població, i seleccionem els Q2 països per cada any segons la variable rank. Guardem tot en l'objecte *PMon\_formatted*:

```
> PMon_formatted <- PMon %>%
  group_by(Any) %>%
  # The * 1 makes it possible to have non-integer ranks while sliding
  mutate(rank = rank(-Poblacio)) %>%
  filter(rank >= 0.25 * max(rank) & rank <= 0.5 * max(rank)) # Q2</pre>
```





**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar en l'objecte anim els frames estàtics de les gràfiques per anys:

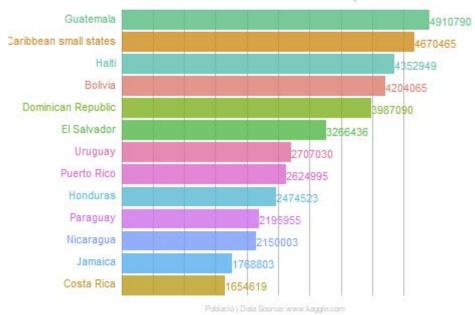
```
> anim <- ggplot(PMon formatted, aes(rank, group = Pais.Nom,</pre>
                fill = as.factor(Pais.Nom), color = as.factor(Pais.Nom))) +
 geom_tile(aes(y = Poblacio/2,
                height = Poblacio,
                width = 0.9), alpha = 0.8, color = NA) +
 geom_text(aes(y = 0, label = paste(Pais.Nom, " ")), vjust=0.2, hjust = 1) +
 geom text(aes(y = Poblacio, label = Poblacio, hjust = 0)) +
 coord_flip(clip = "off", expand = FALSE) +
 scale_y_continuous(labels = scales::comma) +
 scale x reverse() +
 guides(color = "none", fill = "none") +
 theme(axis.line=element_blank(),
        axis.text.x=element_blank(),
        axis.text.y=element blank(),
        axis.ticks=element blank(),
        axis.title.x=element blank(),
        axis.title.y=element_blank(),
        legend.position="none",
        panel.background=element_blank(),
        panel.border=element_blank(),
        panel.grid.major=element blank(),
        panel.grid.minor=element blank(),
        panel.grid.major.x = element_line( linewidth=.1, color="grey" ),
       panel.grid.minor.x = element_line( linewidth=.1, color="grey" ),
       plot.title=element text(size=25,
                                                hjust=0.5.
                                                                  face="bold",
colour="grey", vjust=-1),
       plot.subtitle=element_text(size=18,
                                                hjust=0.5,
                                                                face="italic",
color="grey"),
       plot.caption
                       =element text(size=8, hjust=0.5,
                                                                face="italic",
color="grey"),
       plot.background=element_blank(),
        plot.margin = margin(2,2, 2, 4, "cm")) +
 transition_states(Any, transition_length = 4, state_length=1, wrap=FALSE) +
 view_follow(fixed_x = TRUE) +
 labs(title = 'Població en el món. Any: {closest_state}',
       subtitle = "Q2 Països americans més poblats",
       caption = "Data Source: www.kaggle.com")
PAS 4: A partir de l'objecte anim, creem els fitxers d'animació GIF i AVI:
# Exportar frames a fitxer GIF
> animate(anim, 600, fps = 20, width = 1200, height = 1000,
        renderer = gifski renderer("Q2Americans.gif"), end pause =
                                                                           15,
       start_pause = 15)
# Exportar frames a fitxer AVI
```





> animate(anim, 600, fps = 20, width = 1200, height = 1000, renderer =
av\_renderer("Q2Americans.avi"), end\_pause = 15, start\_pause = 15)

# Població en el món. Any: 1965 Q2 Països americans mes poblats



#### **ALTERNATIVA: Llibreria ddplot**

Primer de tot, instal·larem les llibreries que necessitem:

```
> install.packages('ddplot') # Càrrega paquet ddplot (primera vegada)
> library(ddplot) # Generació Bar Chart Race
```

**PAS 1:** DATA MASSAGING: Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països d'America.

> PMon <- PWorld %>% filter(Continent == "America")

**PAS 2:** Afegim la variable *rank* que correspon al ranking decreixent per països de la població, i seleccionem els Q2 països de cada any segons la variable rank. El guardem en l'objecte PMon\_formatted:

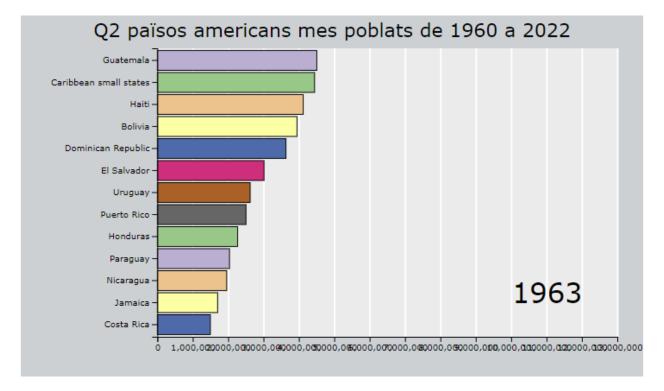
```
> PMon_formatted <- PMon %>%
  group_by(Any) %>%
  mutate(rank = rank(-Poblacio)) %>%
  filter(rank >= 0.25 * max(rank) & rank <= 0.5 * max(rank)) # Q2</pre>
```

**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar els frames estàtics de les gràfiques per anys:

> barChartRace(PMon\_formatted, x="Poblacio", y="Pais.Nom", time="Any", title =
"Q2 països americans més poblats de 1960 a 2022")







**COMENTARI:** En general hi ha canvis en la zona alta, en tres primeres posicions del ranking dels Q2 països americans entre Guatemala, Haití, República Dominicana i Cuba. I en les darreres posicions predominen Costa Rica, panama, Uruguay, Puerto Rico.

En general tots els països creixen en població al llarg dels anys.

Cuba apareix l'any 1999 en primera osició i es manté uns anys en la primera posició, acabant l'any 2022 en la quarta posició.

3.- Crea un *Animate Bar Race Ranking* sobre la primera meitat de països del Q1 MÉS poblats del món. Comenta tres aspectes significatius del que es veu en l'animació. Comenta l'evolució d'Espanya en aquesta llista.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING. Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països, no zones geogràfiques.

> PMon <- PWorld %>% filter(Continent != "Zone")

**PAS 2:** Afegim la variable rank correspon al ranking creixent per població i seleccionem la primera meitat del Q1 de cada any segons la variable rank. El guardem en l'objecte *PMon\_formatted*:

```
> PMon_formatted <- PWorld %>%
  group_by(Any) %>%
  mutate(rank = rank(-Poblacio)) %>%
  filter(rank <= 0.125 * max(rank))</pre>
```

**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar en l'objecte anim els frames estàtics de les gràfiques per anys:

```
> anim <- ggplot(PMon formatted, aes(rank, group = Pais.Nom,</pre>
```



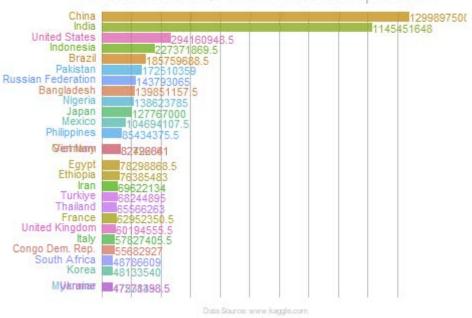


```
fill = as.factor(Pais.Nom), color = as.factor(Pais.Nom))) +
                 geom_tile(aes(y = Poblacio/2,
                 height = Poblacio,
                 width = 0.9), alpha = 0.8, color = NA) +
 geom_text(aes(y = 0, label = paste(Pais.Nom, " ")), vjust =0.2, hjust = 1) +
 geom text(aes(y = Poblacio, label = Poblacio, hjust = 0)) +
 coord_flip(clip = "off", expand = FALSE) +
 scale_y_continuous(labels = scales::comma) +
 scale_x_reverse() +
 guides(color = "none", fill = "none") +
 theme(axis.line=element blank(),
        axis.text.x=element blank(),
        axis.text.y=element_blank(),
        axis.ticks=element_blank(),
        axis.title.x=element blank(),
        axis.title.y=element blank(),
        legend.position="none",
        panel.background=element_blank(),
        panel.border=element blank(),
        panel.grid.major=element blank(),
        panel.grid.minor=element blank(),
        panel.grid.major.x = element line( linewidth=.1, color="grey" ),
       panel.grid.minor.x = element_line( linewidth=.1, color="grey" ),
        plot.title=element_text(size=25,
                                               hjust=0.5,
                                                                 face="bold",
colour="grey", vjust=-1),
        plot.subtitle=element text(size=18,
                                                hjust=0.5,
                                                               face="italic",
color="grey"),
        plot.caption
                       =element_text(size=8, hjust=0.5,
                                                               face="italic",
color="grey"),
        plot.background=element blank(),
        plot.margin = margin(2,2, 2, 4, "cm")) +
 transition_states(Any, transition_length = 4, state_length=1, wrap=FALSE) +
 view_follow(fixed_x = TRUE) +
 labs(title = 'Població en el món. Any: {closest state}',
       subtitle = "Top 15 Països Mundials",
       caption = "Població | Data Source: www.kaggle.com")
PAS 4: A partir de l'objecte anim, creem els fitxers d'animació GIF i AVI:
# Exportar frames a fitxer GIF
> animate(anim, 200, fps = 20, width = 1200, height = 1000,
        renderer = gifski_renderer("MeitatQ1Mundial.gif"), end_pause = 15,
       start pause = 15)
# Exportar frames a fitxer AVI
> animate(anim, 200, fps = 20, width = 1200, height = 1000,
        renderer = av renderer("MeitatQ1Mundial.avi"), end pause =
       start_pause = 15)
```





# Població en el món. Any: 2004 Primera meitat Q1 Països amb més població



#### **ALTERNATIVA: Llibreria ddplot**

Primer de tot, instal·larem les llibreries que necessitem:

```
> install.packages('ddplot') # Càrrega paquet ddplot (primera vegada)
> library(ddplot) # Generació Bar Chart Race
```

**PAS 1:** DATA MASSAGING. Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països, no zones geogràfiques.

> PMon <- PWorld %>% filter(Continent != "Zone")

**PAS 2:** Afegim la variable rank correspon al ranking creixent per població i seleccionem la primera meitat dels Q1 països més poblats de cada any segons la variable rank. El guardem en l'objecte *PMon\_formatted*:

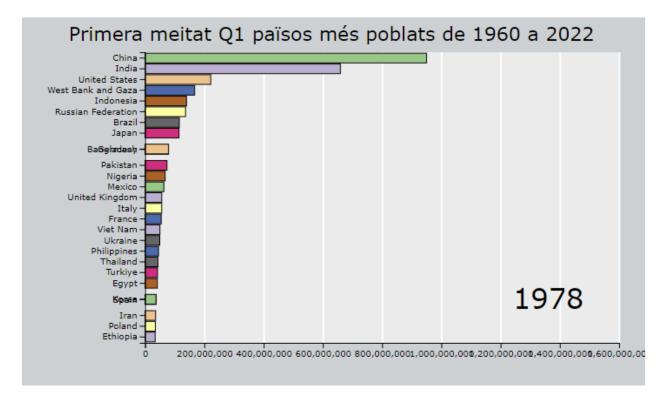
```
> PMon_formatted <- PWorld %>%
  group_by(Any) %>%
  mutate(rank = rank(-Poblacio)) %>%
  filter(rank <= 0.125 * max(rank))</pre>
```

**PAS 3:** Basant-nos en la comanda explicada en les transparències, podem generar els frames estàtics de les gràfiques per anys:

```
> barChartRace(PWorld_formatted, x="Poblacio", y="Pais.Nom", time="Any", title
= "Primera meitat Q1 països més poblats de 1960 a 2022")
```







#### **COMENTARIS:**

- 1. China s'ha mantingut al llarg dels anys com el país més poblat del món fins l'any 2022.
- 2. India ha fet un creixement fort de població el s.XXI fins a arribar a superar China l'any 2022.
- 3. Altres països asiàtics com Indonesia, Pakistan, Bangladesh han tingut augments significatius de població passant a posicions entre els 8 primers.
- 4. Pel que fa als països africans només Nigeria ha tingut un augment significatiu de població, seguit d'Ethiopia.
- 5. Russian Federation, el país amb més extensió del món ha anat baixant posicions al llarg dels anys.
- 6. Pel que fa a països americans, USA es manté en la tercera posició pel que fa a població, seguit de més lluny de Mèxic que ha anat pujant al llarg dels anys fins l'11a posició.
- 7. Els països europeus (Alemanya, UK, Italy) han anat baixant posicions en el rànking en favor dels països asiàtics.
- 8. Espanya apareix des dels anys 160 en posicions mitja-baixa de la taula i desapareix l'any 1993. I torna a aparèixer esporàdicament els anys 2009, 2010 i 2011.

# 4. PART 3. Aplicacions interactives amb Shiny

En aquesta tercera part del seminari implementarem algunes gràfiques interactives amb la llibreria Shiny

#### **RESPOSTA:**

Primer de tot, instal·larem les llibreries Plotly i Shiny.

> library (plotly)





```
> install.packages('shiny') # Càrrega paquet shiny (1a vegada)
> library (shiny)

I seguirem treballant amb el mateix data set:
> setwd("C:/Users/enric/Documents/R")
> PWorld <- read.csv('./Poblacio_Mundial_1960-2022.csv')</pre>
```

#### **EXERCICIS:**

> str(PWorld)

1.- Reprodueix l'aplicació shiny amb entrada per desplegable que dibuixi la gràfica de línies de població per països, de forma que es puguin seleccionar els països en el desplegable. Posa com a país per defecte el nom 'Spain'. Utilitza la funció plot\_ly() o ggplot(), la que vulguis.

Comenta la gràfica de la població dels països de molta extensió com Germany, France, Italy, Poland i Ukraine amb la d'Espanya. Compara aquests països amb Espanya i comenta la seva evolució.

#### **RESPOSTA:**

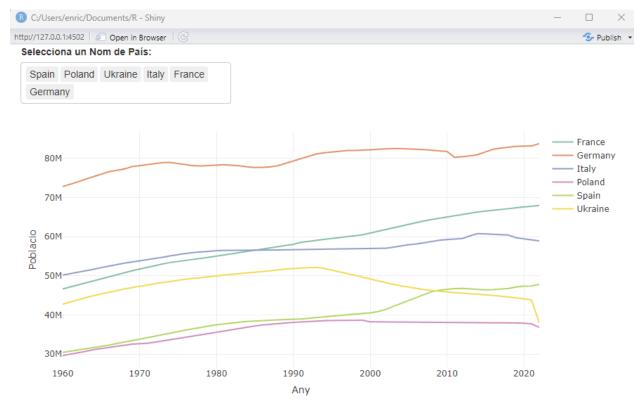
# Aplicació Shiny amb ggplot

```
PAS 1: DATA MASSAGING. Ens guedem només amb països europeus:
> PEuropa <- PWorld %>% filter(Continent == "Europe")
PAS 2: Crear l'aplicació shiny:
# Aplicació Shiny amb plot ly
ui <- fluidPage( selectizeInput(</pre>
                               inputId = "NomsPais",
                               label = "Selecciona un Nom de País:",
                               choices = unique(PEuropa$Pais.Nom),
                               selected = "Spain",
                               multiple = TRUE
                  plotlyOutput(outputId = "plot")
           )
server <- function(input, output, ...)</pre>
            output$plot <- renderPlotly (</pre>
                  plot_ly(PEuropa, x = ~Any, y=~Poblacio, color=~country) %>%
                  filter(Pais.Nom %in% input$NomsPais) %>%
                  group_by(Pais.Nom) %>%
                  add lines()
            })
shinyApp(ui, server)
```





```
ui <- fluidPage( selectizeInput(</pre>
                                inputId = "NomsPais",
                                label = "Selecciona un Nom de País:",
                                choices = unique(PEuropa$Pais.Nom),
                                selected = "Spain",
                                multiple = TRUE
                         plotlyOutput(outputId = "plot")
            )
server <- function(input, output, ...)</pre>
            output$plot <- renderPlotly(</pre>
                   plotNomP <- ggplot(filter(PEuropa,Pais.Nom %in%</pre>
                                input$NomsPais),
                                aes(x=Any,y=Poblacio,color=Pais.Nom))
                                geom path()
                   ggplotly(plotNomP)
            })
shinyApp(ui, server)
```



COMPARATIVA PAÏSOS ALEMANYA, POLAND, ITALY, FRANCE, UKRAINE AMB ESPANYA:





- Germany i France es mantenen per sobre de la població d'Espanya al llarg de tots els anys, essent la població alemanya 2,5 vegades més gran que l'espanyola i la francesa 1,5 vegades aproximadament
- 2. La població d'Ukraine ha sigut superior a la d'Espanya, però l'any 2009 la població espanyola sobrepassa la d'Ukraine.que havia baixat 7M de 1994 (52M) a 2009 (45M), quedant en 38M l'any 2022.
- 3. Poland fins l'any 2000 ha tingut una població una mica inferior a l'espanyola, però a partir del 2000 la població espanyola ha tingut un major creixement i la polaca ha disminuitr una mica o s'ha estancat, baixant 1M de persones del 2021 al 2022.
- 4. La població italiana sempre ha estat superior a l'espanyola amb un creixement similar fins l'any 2012, però a partir de l'any 2014 ha patit un cert descens d'uns 2M entre 2014 a 2022.
- 5. Destaca l'augment de la població espanyola els primers 10 anys del .XXI, quedant després l'augment més estancat.
- 6. Ukraine ha patit una davallada de població a partir del 1994 (desaparició URSS) i especialment a partir del 2021 per la guerra d'Ukraine. De ser el país més poblat dels 5 fins el 1993 amb distància, a passat a la quarta posició dels cinc.
- 7. Poland ha tingut un creixement similar a Ukraine fins el 1993, tenint una lleugera davallada de població partir d'aquell any, quedant actualment en cinquena posició.
- 2.- Reprodueix l'aplicació shiny amb entrada per desplegable que dibuixi la gràfica de línies de la població per països de tot el món, de forma que es puguin seleccionar els països en el desplegable. Posa com a país per defecte el nom 'Spain'. Utilitza la funció plot\_ly() o ggplot(), la que vulguis.

Comenta la gràfica de la població d'Argentina, Canada, Ethiopia i South Africa amb la d'Espanya. Compara aquests països amb Espanya i comenta la seva evolució.

#### **RESPOSTA:**

**PAS 1:** DATA MASSAGING. Eliminem les entrades que tinguin "Continent="Zone", ens quedem només amb països, no zones geogràfiques.





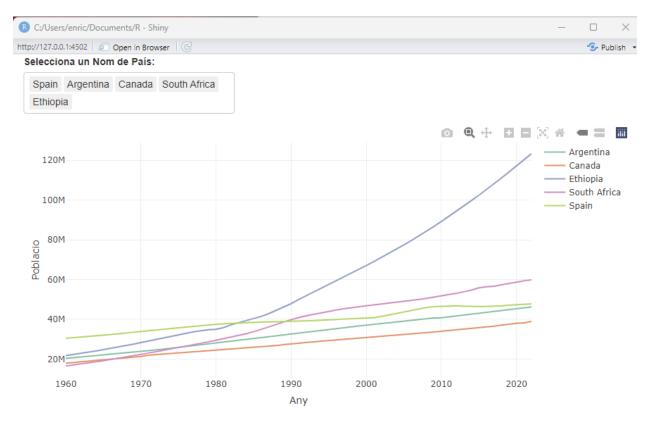
#### Visualització de Dades (Enginyeria Dades) – Curs 2024-25

```
{
                  plot ly(SWorldP, x = ~Any, y = ~Poblacio,
                               color=~Pais.Nom) %>%
                  filter(Pais.Nom %in% input$NomsPais) %>%
                  group_by(Pais.Nom) %>%
                  add_lines()
            })
shinyApp(ui, server)
# Aplicació Shiny amb ggplot
ui <- fluidPage( selectizeInput(</pre>
                               inputId = "NomsPais",
                               label = "Selecciona un Nom de País:",
                               choices = unique(PMon$Pais.Nom),
                               selected = "Spain",
                               multiple = TRUE
                               ),
                         plotlyOutput(outputId = "plot")
           )
server <- function(input, output, ...)</pre>
            output$plot <- renderPlotly(</pre>
                  plotNomP <- ggplot(filter(SWorldP,country %in%</pre>
                               input$NomsPais),
                                             y=Poblacio, color=Pais.Nom))
                               aes(x=Any,
                               geom path()
                  ggplotly(plotNomP)
            })
shinyApp(ui, server)
```



#### Visualització de Dades (Enginyeria Dades) – Curs 2024-25





#### COMPARATIVA PAÏSOS ARGENTINA, CANADA, ESPANYA, ETHIOPIA I SOUTH AFRICA:

- La població espanyola parteix en primera posició l'any 1960 amb 30M i va tenint un lleuger creixement de població, que creix una mica en el període 2001 a 2011 quedant en 47M l'any 2022.
- 2. Argentina i Canada tenen una població inferior a l'espanyola durant tots els anys, tot i tenir major superfície. El creixement d'aquests dos països ha sigut sempre superior a l'espanyol, especialment els darrers anys, de 2009 a 2022.
- 3. La població de Canadà ha sigut sempre inferior a la d'Argentina.
- 4. South Africa parteix en la darrera posició d'aquests països l'any 1960 amb 16M i va tenint al llarg dels anys un creixement molt superior a Espanya, superant-la l'any 1990 i quedant en segona posició l'any 2022 amb 59M d'habitants.
- La població d'Ethiopia parteix de la segona posició l'any 1960 (21M) tenint un moderat creixement a partir de l'any 1980 fins a arribar a 123M l'any 2022. La població etíop sobrepassa l'espanyola l'any 1984.
- La població canadenca ha tingut també un creixement notable semblant a la d'Argentina, doblant el Canadà la població en 60 anys (de 17M a 1960 a 38M a 2022), passant de la quarta posició el 1960 a la darrera el 2022.
- 7. Argentina ha tingut un comportament similar al Canadà però amb població superior a la canadenca, doblant també la població en 60 anys (de 20M a 1960 a 47M a 2022).
- 3. Executant l'aplicació, busca tres països que t'agradin i compara la seva població durant el període 1960-2022.. Què pots deduir de les gràfiques?. Inspira't amb les qüestions de les preguntes 2 i 3 de la part 1.

#### **RESPOSTA:**







Enric Martí Gòdia Bellaterra, Maig 2025