

## ▼ Nombre: Celic Gabriel Hernández Archundia

Matrícula: 2877240

```
%%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/Evidencia2_Casos_DEN_Defunciones_20230417.csv")
```

## ▼ Ejercicio 1. Promedio de defunciones ocurridas a nivel nacional en el mes de su cumpleaños durante el año 2021

```
%%R
octubre.2021 <- datos %>%
  select(nombre, ends_with("10.2021")) %>%
  filter(nombre == "Nacional") %>%
  select(-nombre)

octubre.2021 %>%
  t() %>%
  mean()

[1] 261.6129

%%R
datos %>%
  filter(nombre == "Nacional") %>%
  select(ends_with("10.2021")) %>%
```

```
mutate(promedio = rowMeans( pick(everything()) )) %>%
select(promedio)
```

```
promedio
1 261.6129
```

## ▼ Ejercicio 2. Número total de defunciones a nivel nacional en el año 2020

```
%%R
head(datos[,1:10])
```

	cve_ent	poblacion	nombre	X17.03.2020	X18.03.2020	X19.03.2020
1	1	1434635	AGUASCALIENTES	0	0	0
2	2	3634868	BAJA CALIFORNIA	0	0	0
3	3	804708	BAJA CALIFORNIA SUR	0	0	0
4	4	1000617	CAMPECHE	0	0	0
5	7	5730367	CHIAPAS	0	0	0
6	8	3801487	CHIHUAHUA	0	0	0

  

	X20.03.2020	X21.03.2020	X22.03.2020	X23.03.2020
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0

```
%%R
defunciones.2020 <- datos %>%
select(nombre, ends_with("2020")) %>%
filter(nombre == "Nacional") %>%
select(-nombre)
```

```
defunciones.2020 %>%
t() %>%
sum()
```

```
[1] 149455
```

## ▼ Ejercicio 3. Estado con el mayor número de defunciones en el mes de Agosto 2020 (sin contar la Ciudad de México ni el Estado de México)

```
%%R
datos %>%
select(nombre, ends_with("08.2020")) %>%
filter(nombre != "DISTRITO FEDERAL",
nombre != "MEXICO",
nombre != "Nacional") %>%
mutate(num.defunciones = rowSums( pick(where(is.numeric) ))) %>%
arrange(desc(num.defunciones)) %>%
select(nombre, num.defunciones) %>%
head(1)
```

```
nombre num.defunciones
1 NUEVO LEON 1247
```

## ▼ Ejercicio 4. Mes con el mayor número de defunciones a nivel nacional en el año 2022

```
%%R
resultados <- datos %>%
```

```

mutate(ene.2022 = rowSums(select(datos, ends_with("01.2022"))),
      feb.2022 = rowSums(select(datos, ends_with("02.2022"))),
      mar.2022 = rowSums(select(datos, ends_with("03.2022"))),
      abr.2022 = rowSums(select(datos, ends_with("04.2022"))),
      may.2022 = rowSums(select(datos, ends_with("05.2022"))),
      jun.2022 = rowSums(select(datos, ends_with("06.2022"))),
      jul.2022 = rowSums(select(datos, ends_with("07.2022"))),
      ago.2022 = rowSums(select(datos, ends_with("08.2022"))),
      sep.2022 = rowSums(select(datos, ends_with("09.2022"))),
      oct.2022 = rowSums(select(datos, ends_with("10.2022"))),
      nov.2022 = rowSums(select(datos, ends_with("11.2022"))),
      dic.2022 = rowSums(select(datos, ends_with("12.2022")))) %>%
filter(nombre == "Nacional") %>% # filter para filas
select(ene.2022, feb.2022, mar.2022, abr.2022, may.2022, jun.2022, jul.2022, ago.2022, sep.2022, oct.2022, nov.2022, dic

%%R
resultados %>%
select(apply(resultados,1, which.max))

      feb.2022
1      9980

```

### ▼ Ejercicio 5. Día (fecha) en la que la Ciudad de México tuvo más defunciones

```

%%R
dia.CDMX <- datos %>%
filter(nombre == "DISTRITO FEDERAL") %>%
select(-nombre, -cve_ent, -poblacion)

%%R
dia.CDMX %>%
select(apply(dia.CDMX,1, which.max))

      X26.01.2021
1      270

```