

**INSTITUTO TECNOLÓGICO[®]
de Pabellón de Arteaga**

TEC

DOCUMENTO

Procesamiento de Datos

Nombre del Alumno

Sara del Carmen Sanches de la Rosa

Materia

Ingeniería del Conocimiento

Docente

Flores Gallegos Eduardo

Fecha

28/05/2024

Código

```
# Instalar y cargar librerías
if (!require("readr")) install.packages("readr")
if (!require("dplyr")) install.packages("dplyr")
if (!require("ggplot2")) install.packages("ggplot2")

library(readr)
library(dplyr)
library(ggplot2)

# Leer el archivo CSV
covid_data <- read_csv("C:/Users/E1_M1/OneDrive/Imágenes/COVID19MEXICO.csv")

# Verificar nombres de columnas
print(colnames(covid_data))

print(unique(covid_data$CLASIFICACION_FINAL))

print(unique(covid_data$SEXO))

# ¿Cuál es la edad media de mujeres que enfermaron de COVID?
edad_media_mujeres <- covid_data %>%
  filter(SEXO == 2 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
  summarize(edad_media = mean(EDAD, na.rm = TRUE))

print(edad_media_mujeres)

# ¿Cuál es la edad media de hombres que enfermaron de COVID?
edad_media_hombres <- covid_data %>%
  filter(SEXO == 1 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
  summarize(edad_media = mean(EDAD, na.rm = TRUE))

print(edad_media_hombres)

# ¿Cuántos hombres enfermaron en el municipio de Pabellón de Arteaga?
hombres_pabellon <- covid_data %>%
  filter(SEXO == 1 & MUNICIPIO_RES == "Pabellón de Arteaga" & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
```

```
tally()

print(hombres_pabellon)

# ¿Cuál fue el municipio de Aguascalientes que tuvo la mayor cantidad de mujeres enfermas de COVID?
mayor_municipio_mujeres <- covid_data %>%
  filter(SEXO == 2 & ENTIDAD_RES == 1 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>% # Suponiendo
  que '1' es Aguascalientes en ENTIDAD_RES
  group_by(MUNICIPIO_RES) %>%
  tally() %>%
  arrange(desc(n)) %>%
  slice(1)

print(mayor_municipio_mujeres)

# ¿Cuántos pacientes ambulatorios hubo en el país?
pacientes_ambulatorios <- covid_data %>%
  filter(TIPO_PACIENTE == 1 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>% # Suponiendo que '1' es
  AMBULATORIO en TIPO_PACIENTE
  tally()

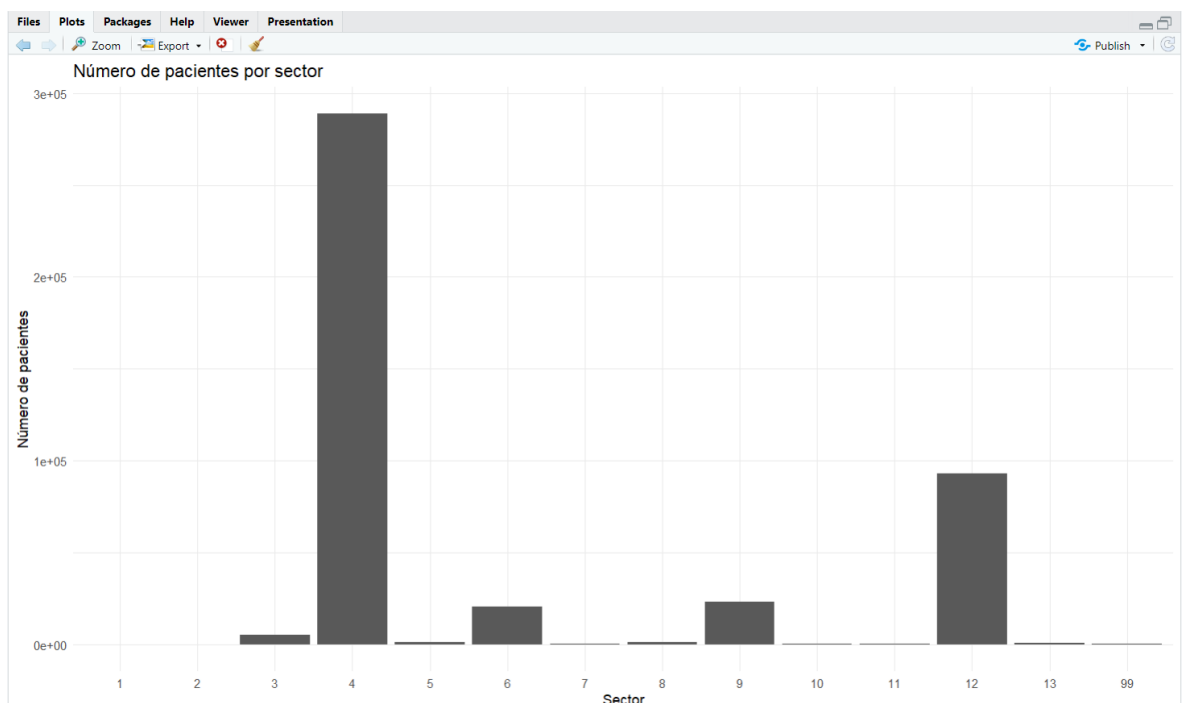
print(pacientes_ambulatorios)

# Graficar con ggplot una gráfica con los pacientes de cada sector
ggplot(covid_data %>% filter(CLASIFICACION_FINAL %in% c(1, 2, 3)), aes(x = as.factor(SECTOR))) +
  geom_bar() +
  labs(title = "Número de pacientes por sector",
        x = "Sector",
        y = "Número de pacientes") +
  theme_minimal()

# ¿Cuál es la edad máxima en la que se presenta una persona infectada?
edad_maxima <- covid_data %>%
  filter(CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
  summarize(edad_maxima = max(EDAD, na.rm = TRUE))

print(edad_maxima)
```

1. **¿Cuál es la edad media de mujeres que enfermaron de covid?**
es aproximadamente 40.7 años
2. **¿Cuál es la edad media de hombres que enfermaron de covid?**
es aproximadamente 42.5 años
3. **¿Cuántos hombres enfermaron en el municipio de Pabellón de Arteaga?**
No se encontraron en la base, o bien no hubo
4. **¿Cuál fue el municipio de Aguascalientes que tuvo la mayor cantidad de mujeres enfermas de covid?**
No se encontraron en la base, o bien no hubo
5. **¿Cuántos pacientes ambulatorios hubo en el país?**
Hubo 413,461 pacientes ambulatorios en el país.
6. **Grafique con ggplot una gráfica con los pacientes de cada sector.**



7. **¿Cuál es la edad máxima en la que se presenta una persona infectada?**
La edad máxima en la que se presenta una persona infectada es de 123 años

```
> if (!require("ggplot2")) install.packages("ggplot2")
Cargando paquete requerido: ggplot2
>
> library(readr)
> library(dplyr)
> library(ggplot2)
>
> # Leer el archivo CSV
> covid_data <- read_csv("C:/Users/E1_M1/OneDrive/Imágenes/COVID19MEXICO.csv")
Rows: 1296717 Columns: 40
— Column specification —
Delimiter: ","
chr  (7): ID_REGISTRO, ENTIDAD_UM, ENTIDAD_NAC, ENTIDAD_RES, MUNICIPIO_RES, PAIS_NACIONALI...
dbl  (29): ORIGEN, SECTOR, SEXO, TIPO_PACIENTE, INTUBADO, NEUMONIA, EDAD, NACIONALIDAD, EMB...
date  (4): FECHA_ACTUALIZACION, FECHA_INGRESO, FECHA_SINTOMAS, FECHA_DEF

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.
Aviso:
One or more parsing issues, call `problems()` on your data frame for details, e.g.:
  dat <- vroom(...)
  problems(dat)
>
>
> # Verificar nombres de columnas
> print(colnames(covid_data))
[1] "FECHA_ACTUALIZACION" "ID_REGISTRO" "ORIGEN"
[4] "SECTOR" "ENTIDAD_UM" "SEXO"
[7] "ENTIDAD_NAC" "ENTIDAD_RES" "MUNICIPIO_RES"
[10] "TIPO_PACIENTE" "FECHA_INGRESO" "FECHA_SINTOMAS"
[13] "FECHA_DEF" "INTUBADO" "NEUMONIA"
[16] "EDAD" "NACIONALIDAD" "EMBARAZO"
[19] "HABLA LENGUA INDIG" "INDIGENA" "DIABETES"
[22] "EPOC" "ASMA" "INMUSUPR"
[25] "HIPERTENSION" "OTRA_COM" "CARDIOVASCULAR"
[28] "OBESIDAD" "RENAL_CRONICA" "TABAQUISMO"
[31] "OTRO_CASO" "TOMA_MUESTRA_LAB" "RESULTADO_LAB"
[34] "TOMA_MUESTRA_ANTIGENO" "RESULTADO_ANTIGENO" "CLASIFICACION_FINAL"
[37] "MIGRANTE" "PAIS_NACIONALIDAD" "PAIS_ORIGEN"
```



```
[34] "TOMA_MUESTRA_ANTIGENO" "RESULTADO_ANTIGENO" "CLASIFICACION_FINAL"
[37] "MIGRANTE" "PAIS_NACIONALIDAD" "PAIS_ORIGEN"
[40] "UCI"
>
> # Verificar valores únicos en CLASIFICACION_FINAL
> print(unique(covid_data$CLASIFICACION_FINAL))
[1] 7 3 6 1 5 2 4
>
> # Verificar valores únicos en SEXO
> print(unique(covid_data$SEXO))
[1] 2 1
>
> # ¿Cuál es la edad media de mujeres que enfermaron de COVID?
> edad_media_mujeres <- covid_data %>%
+   filter(SEXO == 2 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
+   summarize(edad_media = mean(EDAD, na.rm = TRUE))
>
> print(edad_media_mujeres)
# A tibble: 1 x 1
  edad_media
    <dbl>
1      40.7
>
> # ¿Cuál es la edad media de hombres que enfermaron de COVID?
> edad_media_hombres <- covid_data %>%
+   filter(SEXO == 1 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
+   summarize(edad_media = mean(EDAD, na.rm = TRUE))
>
> print(edad_media_hombres)
# A tibble: 1 x 1
  edad_media
    <dbl>
1      42.5
>
> # ¿Cuántos hombres enfermaron en el municipio de Pabellón de Arteaga?
> hombres_pabellon <- covid_data %>%
+   filter(SEXO == 1 & MUNICIPIO_RES == "Pabellón de Arteaga" & CLASIFICACION_FINAL %in% c(1, 2,
3)) %>%
+   tally()
>
> print(hombres_pabellon)
```

```
- group_by(MUNICIPIO_RES) %>%
- tally() %>%
- arrange(desc(n)) %>%
- slice(1)
>
> print(mayor_municipio_mujeres)
# A tibble: 0 × 2
# 2 variables: MUNICIPIO_RES <chr>, n <int>
>
> # ¿Cuántos pacientes ambulatorios hubo en el país?
> pacientes_ambulatorios <- covid_data %>%
- filter(TIPO_PACIENTE == 1 & CLASIFICACION_FINAL %in% c(1, 2, 3)) %>% # Suponiendo que '1' es AMBULATORIO e
TIPO_PACIENTE
- tally()
>
> print(pacientes_ambulatorios)
# A tibble: 1 × 1
      n
  <int>
1 413461
>
> # Graficar con ggplot una gráfica con los pacientes de cada sector
> ggplot(covid_data %>% filter(CLASIFICACION_FINAL %in% c(1, 2, 3)), aes(x = as.factor(SECTOR))) +
- geom_bar() +
- labs(title = "Número de pacientes por sector",
-       x = "Sector",
-       y = "Número de pacientes") +
- theme_minimal()
>
> # ¿Cuál es la edad máxima en la que se presenta una persona infectada?
> edad_maxima <- covid_data %>%
- filter(CLASIFICACION_FINAL %in% c(1, 2, 3)) %>%
- summarize(edad_maxima = max(EDAD, na.rm = TRUE))
>
> print(edad_maxima)
# A tibble: 1 × 1
  edad_maxima
    <dbl>
1      123
```

Filter	
edad_maxima	
1	123



Untitled1* x covid_data x edad_maxima x edad_media_hombres x	
Filter	
edad_media	
1	42.509

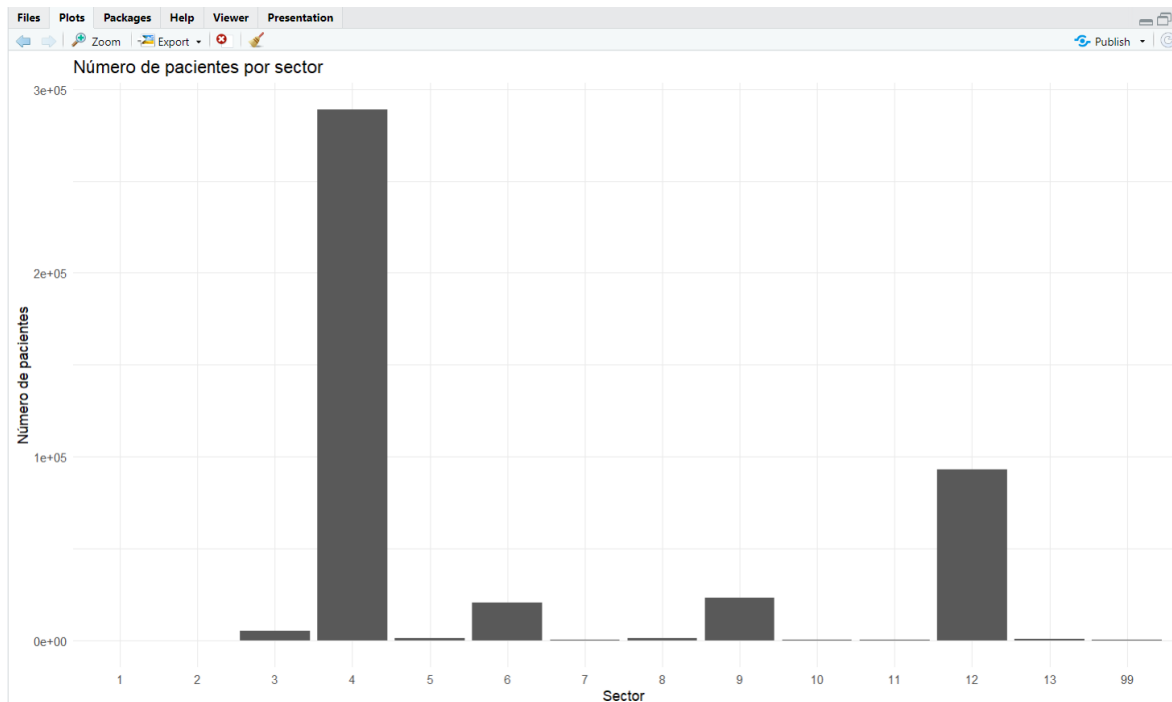
Untitled1* x covid_data x edad_maxima x edad_media_hombres x edad_media_mujeres x	
Filter	
edad_media	
1	40.72147

Untitled1* x covid_data x edad_maxima x edad_media_hombres x edad_media_mujeres x hombres_pabellon x	
Filter	
n	
1	0

covid_data x edad_maxima x edad_media_hombres x edad_media_mujeres x hombres_pabellon x mayor_municipio_mujeres x	
Filter	
MUNICIPIO_RES	n
No data available in table	

axima x edad_media_hombres x edad_media_mujeres x hombres_pabellon x mayor_municipio_mujeres x pacientes_ambulatorios x	
Filter	
n	
1	413461

Carretera a la Estación de Rincón Km. 1 C.P. 20670 Pabellón de Arteaga, Aguascalientes
Tel. 465 958-2482 Ext. 104 e-mail: plan_parteaga@tecnm.mx tecnm.mx | pabellon.tecnm.mx



Carretera a la Estación de Rincón Km. 1 C.P. 20670 Pabellón de Arteaga, Aguascalientes
Tel. 465 958-2482 Ext. 104 e-mail: plan_parteaga@tecnm.mx tecnm.mx | pabellon.tecnm.mx

