

Ejercicios_Analisis_Explo

RLT

3/15/2021

Contents

SE ASUME QUE LEO Y HIZO LOS EJERCICIO DEL CAPITULO 7, DEL LIBRO	1
Recuerda buscar el data set nuevo al dia de hoy.	2
.	2
1. EJERCICIO (5 puntos)	2
2. EJERCICIO (5 puntos)	2
3. EJERCICIO (5 puntos)	3
4. Ejercicio (5 puntos)	3
5. EJERCICIO (5 puntos)	3
BONO (5 puntos)	4
Fecha de la ultima revisión	

El tema proviene de los siguientes sitios.

English: <https://r4ds.had.co.nz/factors.html>

Español: <https://es.r4ds.hadley.nz/factores.html>

SE ASUME QUE LEO Y HIZO LOS EJERCICIO DEL CAPITULO 7, DEL LIBRO

Los ejercicios aqui se basan en parte de la información en el capitulo y lo aprendido en la clase. En los siguientes ejercicio usar el data set de COVID_PR que utilizamos en los trabajos anteriores (tiene que poner su data set al dia). Tiene que enseñar los scripts. Enviar el trabajo por Edmodo en formato .html.

Cada estudiante tiene que subir su trabajo a Edmodo

```
library(ggplot2)
library(tidyverse)
```

Recuerda buscar el data set nuevo al día de hoy.

```
library(readr)
url_COVID_PR <- read_csv("Datos/url_COVID_PR.csv")

head(url_COVID_PR)

tail(url_COVID_PR)
```

1. EJERCICIO (5 puntos)

Haga un histograma de la cantidad de casos de Covid-19 por día, la columna “IncCasosSaludNuevo”. - añadir una línea blanca alrededor de las columnas - cambiar el color de las barras - Añadir leyendas en y y x con sean más apropiadas

```
ggplot(url_COVID_PR, aes(IncCasosSaludNuevo)) +
  geom_histogram(colour = "white", fill="orange")
```

2. EJERCICIO (5 puntos)

Añadir una nueva columna al data frame del día de la semana y enseñar los valores de esta de esta columna. Puede usar la librería “lubridate” con la función wday, recuerda que tiene que convertir la columna anteriormente a fecha.

```
library(lubridate)

url_COVID_PR$mdy=mdy(url_COVID_PR$Fecha)
url_COVID_PR$weekday=wday(url_COVID_PR$mdy)

url_COVID_PR$weekday
```

3. EJERCICIO (5 puntos)

Haz un gráfico del promedio de nuevos casos (IncCasosSaludNuevo) por el día de la semana.

- añadir una linea blanca alrededor de las columnas
- cambiar el color de las barras
- Añadir leyendas en y y x con sean más apropiadas

```
df=url_COVID_PR %>%
  select(IncCasosSaludNuevo,weekday) %>%
  group_by(weekday) %>%
  drop_na() %>%
  summarize(meanCV=mean(IncCasosSaludNuevo))

df

ggplot(df, aes(weekday, meanCV)) +
  geom_col(colour = "white", fill="orange")
```

4. Ejercicio (5 puntos)

Hacer una gráfica de linea (geom_line) del número de persona que han recibido su primera vacuna en Puerto Rico, "N1MoreDoses".

- Nota que debería estar el eje de x, la fecha
- remover del análisis las fechas que no tienen vacunas (enseñar el script que uso para lograrlo)

```
names(url_COVID_PR)

url_COVID_PR %>%
  drop_na() %>%
  ggplot(aes(mdy,N1MoreDoses)) +
  geom_line( )
```

5. EJERCICIO (5 puntos)

Usando la misma información anterior haga un analisis del % de personas vacunadas en Puerto Rico. - Asume que la población total de Puerto Rico es 2,810,607 - Deberia enseñar como cambia el % de personas vacunas en tiempo

```
url_COVID_PR %>%
  drop_na(VacAdm) %>%
  ggplot(aes(mdy,VacAdm/2810607)) +
  geom_line( )
```

BONO (5 puntos)

En la misma gráfica anterior, añade el porcentaje de personas que han recibido su segunda vacuna, “N2Doses”.

- Deberían tener dos líneas, una para la primera vacuna y la para las dos vacunas.
- Cambie el nombre de los ejes para identificar correctamente la información.

```
url_COVID_PR %>%  
  drop_na(N1MoreDoses) %>%  
ggplot(aes(mdy, N1MoreDoses/2810607)) +  
  geom_line() +  
  geom_line(aes(mdy, N2Doses/2810607))
```

““