

# Clase Práctica

*Joshua*

*2019-06-15*

## Pregunta 0

Necesitaremos algunos paquetes para trabajar. Instale los siguientes paquetes: `babynames`, `nycflights13`, `maps` y carguelos (con `library`) junto al `tidyverse`.

## Pregunta 1

Considere, la tabla `births` del paquete `babynames`. Esto es:

```
library(tidyverse)

data("births", package = "babynames")
births
```

```
## # A tibble: 109 x 2
##   year births
##   <int>   <int>
## 1  1909 2718000
## 2  1910 2777000
## 3  1911 2809000
## 4  1912 2840000
## 5  1913 2869000
## 6  1914 2966000
## 7  1915 2965000
## 8  1916 2964000
## 9  1917 2944000
## 10 1918 2948000
## # ... with 99 more rows
```

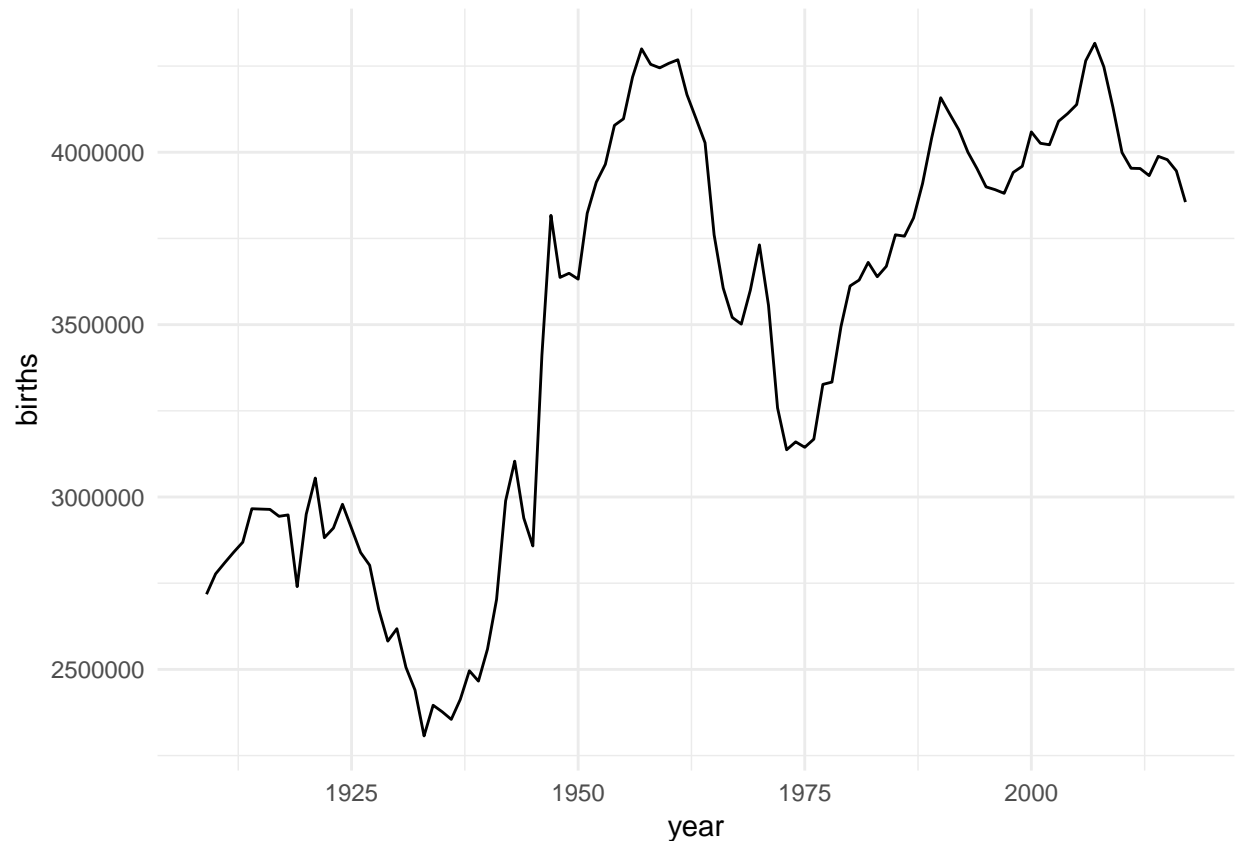
¿Cuántas columnas y filas posee la tabla *births*?

## Pregunta 2

Considere, la tabla `births` del paquete `babynames`. Estudie los nacimientos por años en EE.UU:

```
theme_set(theme_minimal())

ggplot(births) +
  geom_line(aes(year, births))
```



¿En que *decada* se obtuvo la menor cantidad de nacimientos?

### Pregunta 3

Considerando ahora la tabla `babynames` del paquete `babynames`:

```
data("babynames", package = "babynames")
glimpse(babynames)
```

```
## Observations: 1,924,665
## Variables: 5
## $ year <dbl> 1880, 1880, 1880, 1880, 1880, 1880, 1880, 1880, 1880, 188...
## $ sex <chr> "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F..."
## $ name <chr> "Mary", "Anna", "Emma", "Elizabeth", "Minnie", "Margaret"...
## $ n <int> 7065, 2604, 2003, 1939, 1746, 1578, 1472, 1414, 1320, 128...
## $ prop <dbl> 0.07238359, 0.02667896, 0.02052149, 0.01986579, 0.0178884...
```

Una tabla con cinco variables: año, sexo, nombre, `n` y `prop` (`n` dividido por el número total de solicitantes en ese año, lo que significa que las proporciones son de personas de ese género con ese nombre nacido en ese año).

Realice el proceso necesario para obtener la misma estructura en cuanto a *columnas* que la tabla `births`.

## Pregunta 4

Genere un data frame partiendo de la tabla babyanes `babynames` y conteniendo los nacimientos de personas de género femenino con el nombre *Nala*, *Ariel* y *Elsa* desde los años 1980 en adelante.

## Pregunta 5

Con el data frame obtenido en la pregunta anterior genere un gráfico que contenga la información de los nacimientos por año de cada uno de los nombres mencionados y  **mencione**  una hipótesis/suposición al respecto de lo observado.

hint: Use `facet_wrap(~ name, scales = "free_y")`.

## Pregunta 6

Utilizando la tablas `airports` (aeropuertos) y `flights` (vuelos) del paquete `nycflights13` obtenga una tabla que contenga conteos de vuelos según su destino `dest` además de la longitud y latitud del aeropuerto (de destino).

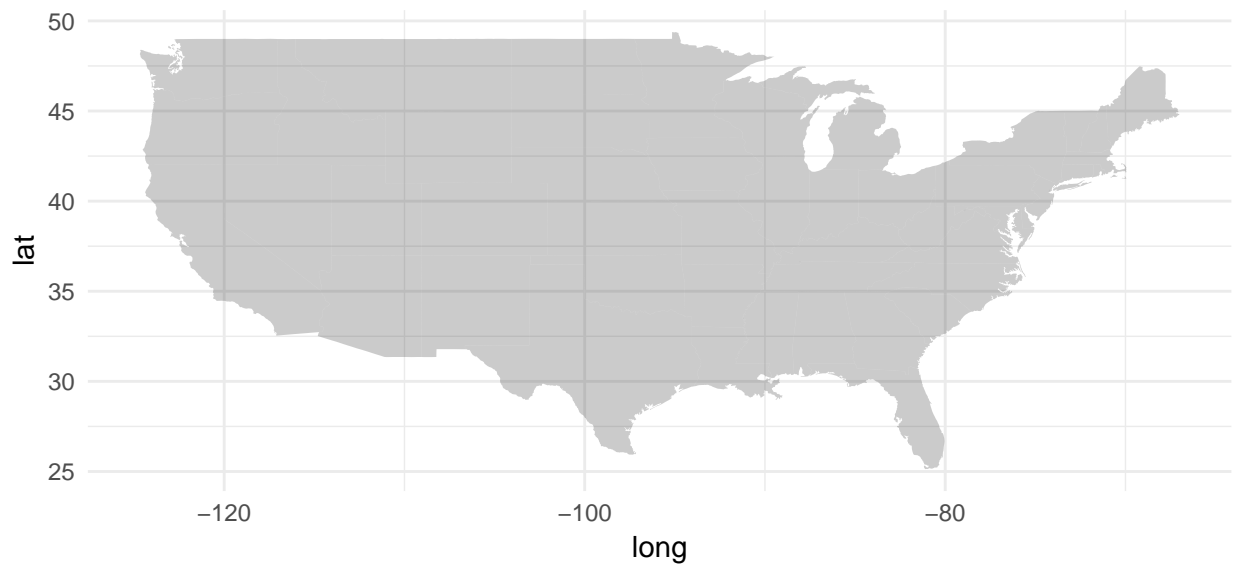
## Pregunta 7

Apoyándose del siguiente gráfico

```
us <- map_data("state")  
  
glimpse(us)
```

```
## Observations: 15,537  
## Variables: 6  
## $ long      <dbl> -87.46201, -87.48493, -87.52503, -87.53076, -87.5708...  
## $ lat       <dbl> 30.38968, 30.37249, 30.37249, 30.33239, 30.32665, 30...  
## $ group     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1...  
## $ order     <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 1...  
## $ region    <chr> "alabama", "alabama", "alabama", "alabama", "alabama...  
## $ subregion <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

```
ggmap <- ggplot() +  
  geom_polygon(data = us, aes(long, lat, group = group),  
              alpha = 0.25) +  
  coord_fixed() # esto es para mantener la razón 1:1  
  
ggmap
```



Agregue una capa de de puntos ubicando los aeropuertos obtenidos de la pregunta anterior y usando además: `aes(size = la_cantidad_de_vuelos_a_dicho_aeropuerto)`.

### Pregunta 8

A la izquierda del gráfico anterior se observan 2 puntos. Genere el/los pasos necesarios para seleccionarlos usando la tabla resultante de la pregunta 5 para identificar los nombres de dichos areopuertos y responda ¿Dónde están? ¿Que gracias tienen?

### Pregunta 9

Volviendo con `babynames`. Encuentre/invente/estudie alguna forma de detectar series de nombres que presenten anomalías o comportamientos *llamativos*.