

Tarea 1

Daniel Camarena

18/8/2017

Tarea 1

Usando el paquete *gapminder* realiza al menos 3 gráficas y explica las relaciones que encuentres. Debes usar lo que revisamos en estas notas: al menos una de las gráficas debe ser de paneles, realiza una gráfica con datos de México, y (opcional) si lo consideras interesante, puedes crear una variable categórica utilizando la función `cut2` del paquete *Hmisc*.

Gapminder

Cargamos las librerías necesarias:

```
library(dplyr)
library(ggplot2)
library(gapminder)
```

Veamos de que trata el paquete *gapminder*:

```
?gapminder
```

Con el comando anterior podemos ver que el paquete *gapminder* contiene datos de la esperanza de vida, *GDP* per capita, y densidad de población por país.

Sería pertinente echar un vistazo a los datos con la función *glimpse*

```
glimpse(gapminder)
```

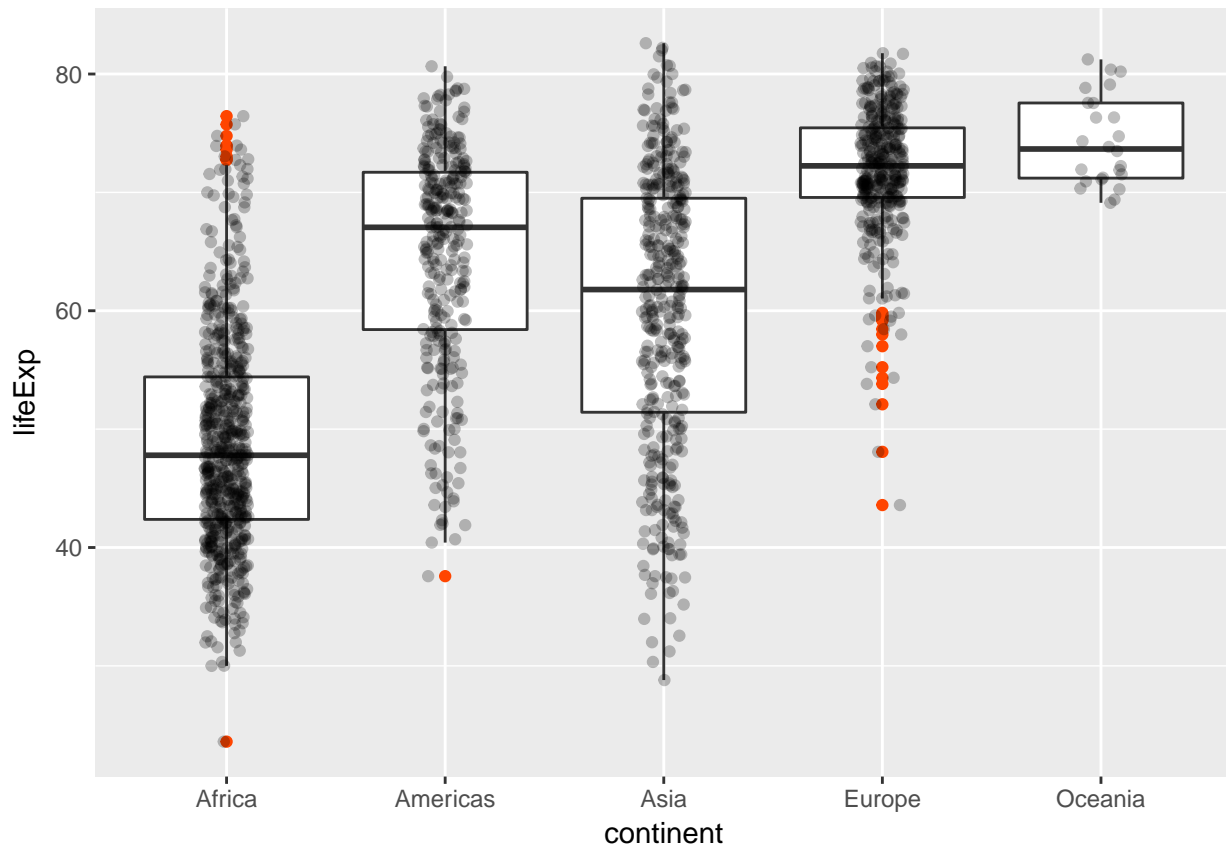
```
## Observations: 1,704
## Variables: 6
## $ country   <fctr> Afghanistan, Afghanistan, Afghanistan, Afghanistan,...
## $ continent <fctr> Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asi...
## $ year      <int> 1952, 1957, 1962, 1967, 1972, 1977, 1982, 1987, 1992...
## $ lifeExp   <dbl> 28.801, 30.332, 31.997, 34.020, 36.088, 38.438, 39.8...
## $ pop       <int> 8425333, 9240934, 10267083, 11537966, 13079460, 1488...
## $ gdpPercap <dbl> 779.4453, 820.8530, 853.1007, 836.1971, 739.9811, 78...
```

Podemos observar que este paquete contiene las siguientes variables:

- country: factor (142 niveles)
- continent: factor (5 niveles)
- year: entero, rangos desde 1952 a 2007 en incrementos de 5 años
- lifeExp: doble, esperanza de vida en años al momento del nacimiento
- pop: entero, población
- gdpPercap: doble, *GDP* per capita

Primero grafiquemos la esperanza de vida por continente:

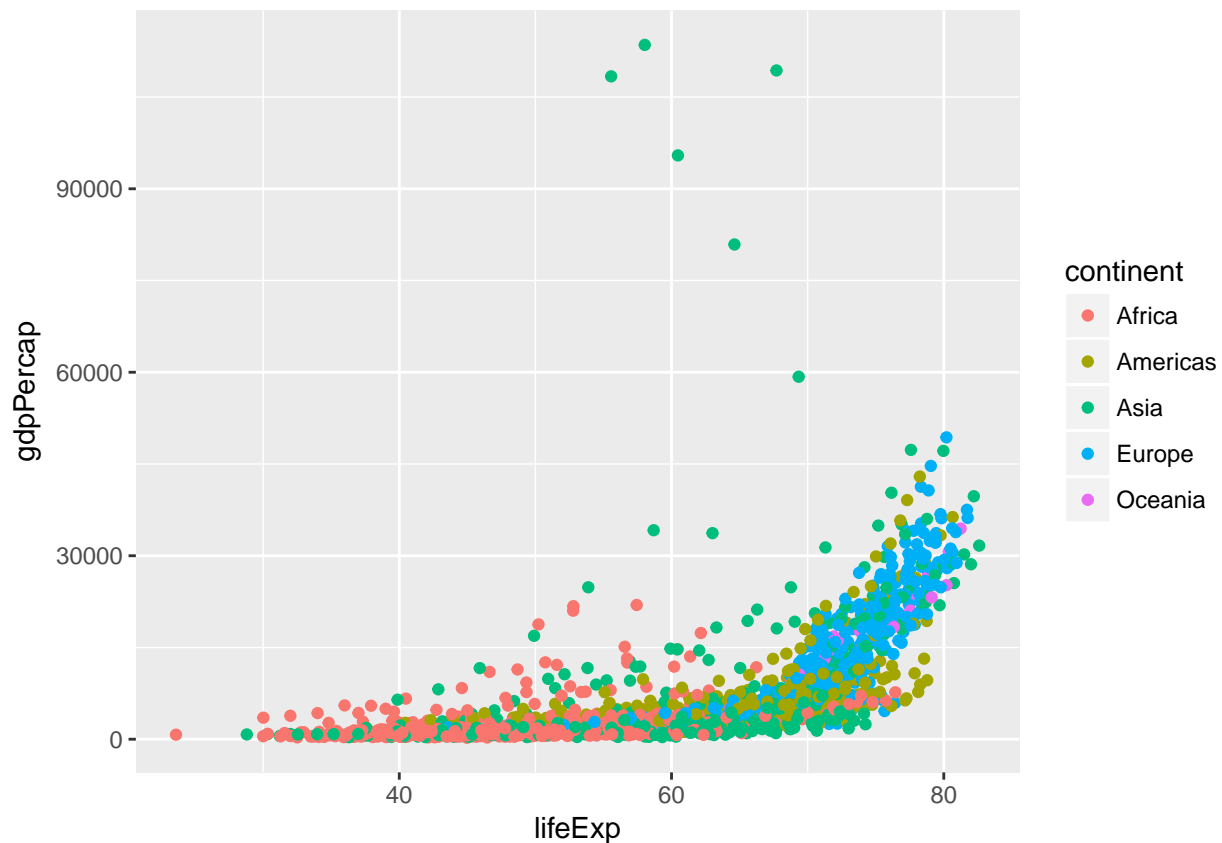
```
ggplot(gapminder, aes(x = continent, y = lifeExp)) +
  geom_boxplot(outlier.colour = "orangered") +
  geom_jitter(position = position_jitter(width = 0.1, height = 0), alpha = 1/4)
```



Podemos observar que los continentes con mayor esperanza de vida son Europa y Oceania, seguidos por America y Asia.

Podemos graficar el gdpPercapita vs la Esperanza de Vida para cada uno de los continentes:

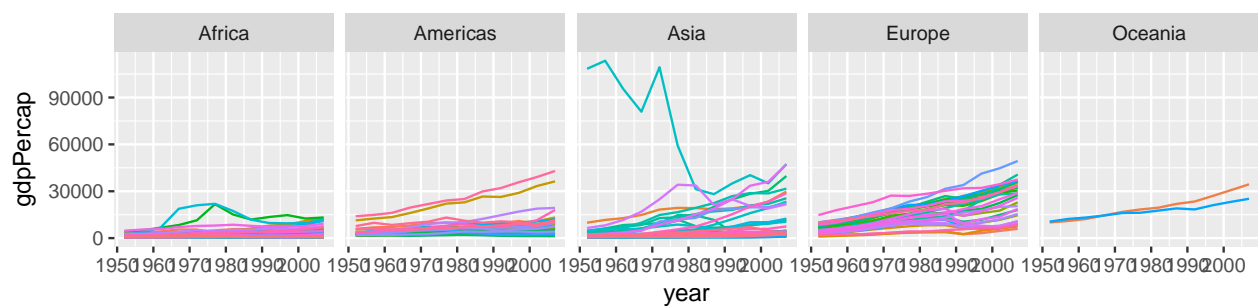
```
ggplot(data = gapminder, aes(x = lifeExp, y = gdpPercap)) +  
  geom_point(aes(color=continent))
```



Podemos notar que Africa tiene los menores niveles de ambas variable mientras que Europa tiene alta esperanza de vida y niveles medios de gdpPercapita, siendo Asia quien destaca por tener puntos mas altos en gdpPercapita.

Podemos graficar el gdpPercap de cada uno de los paises que forman un continente:

```
ggplot(data= gapminder, aes(x=year, y=gdpPercap, color=country)) +
  geom_line() + facet_grid(. ~ continent) + theme(legend.position="none")
```



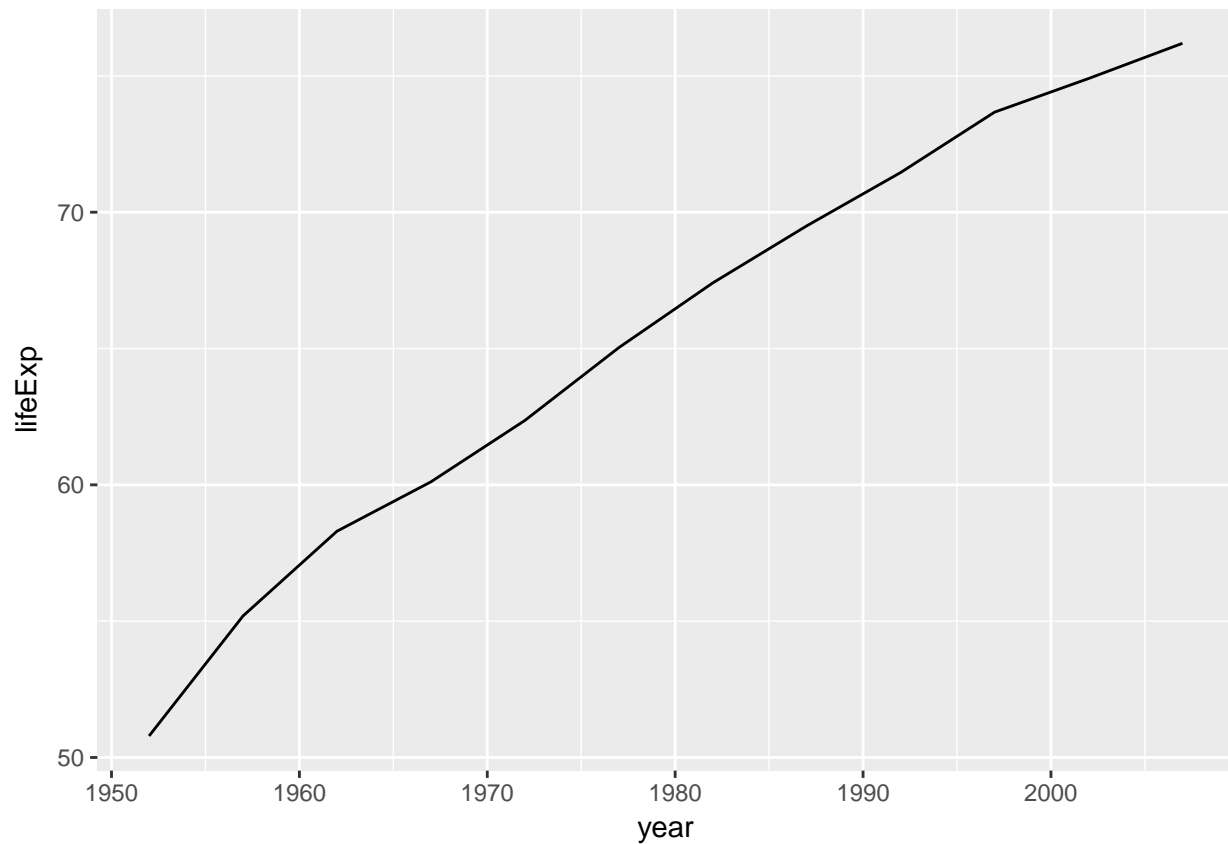
Nuevamente vemos que Africa tiene los peores niveles (excepto por dos paises que sobresalen), en Europa vemos como hay una tendencia similar para la mayoría de los paises, en Oceania ambos paises tienen un gdp similar, en America sobresalen dos paises y los demás siguen una tendencia diferente a estos dos y en Asia hay un Pais que sobresalió desde 1950 hasta 1980.

Concentremos nuestra atención en México

Podría ser interesante graficar la esperanza de vida para Mexico:

```
data <- gapminder %>% filter(country == 'Mexico')
```

```
g <- ggplot(data, aes(x=year, y=lifeExp))
g + geom_line()
```

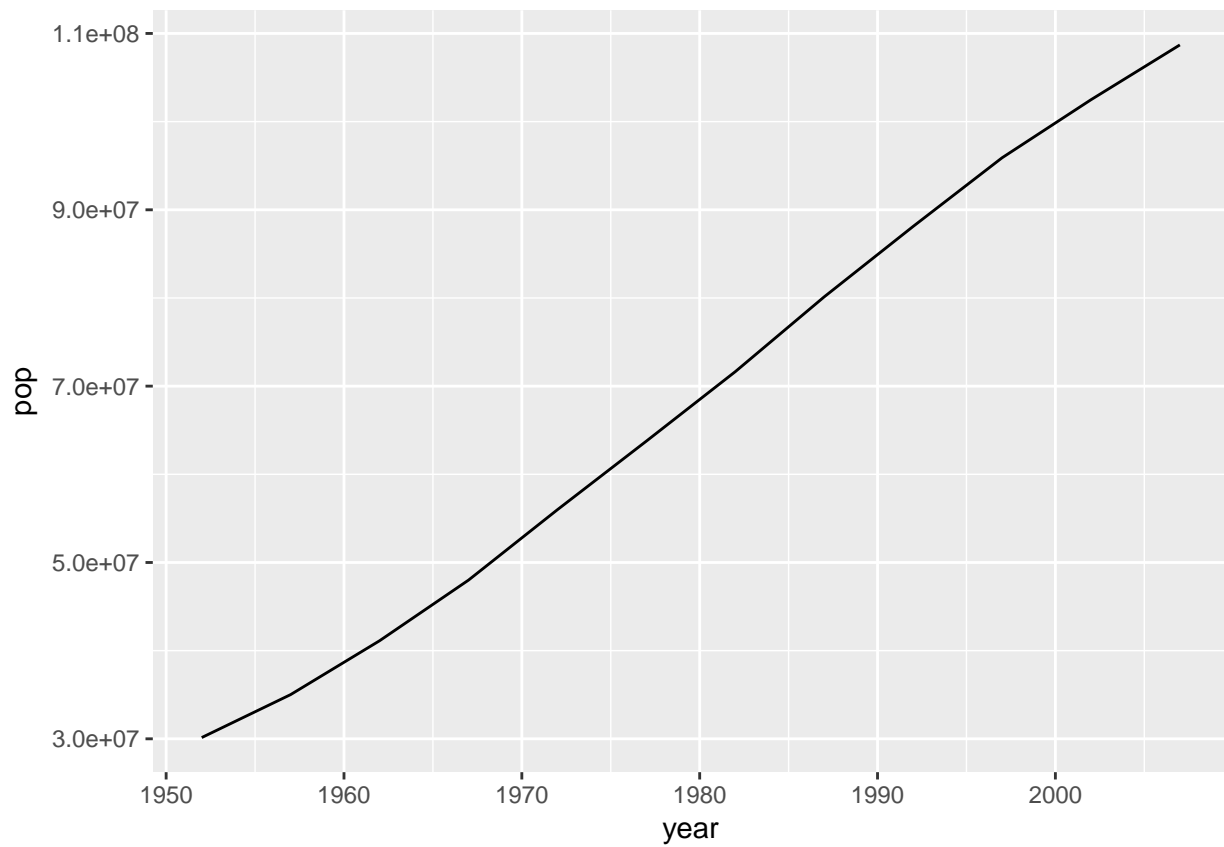


Podemos ver como es que la esperanza de vida ha seguido creciendo a lo largo de los años pasando de 51 años en 1950 a mas de 76 años en 2007

Ya estando en esto, veamos la población en México también...

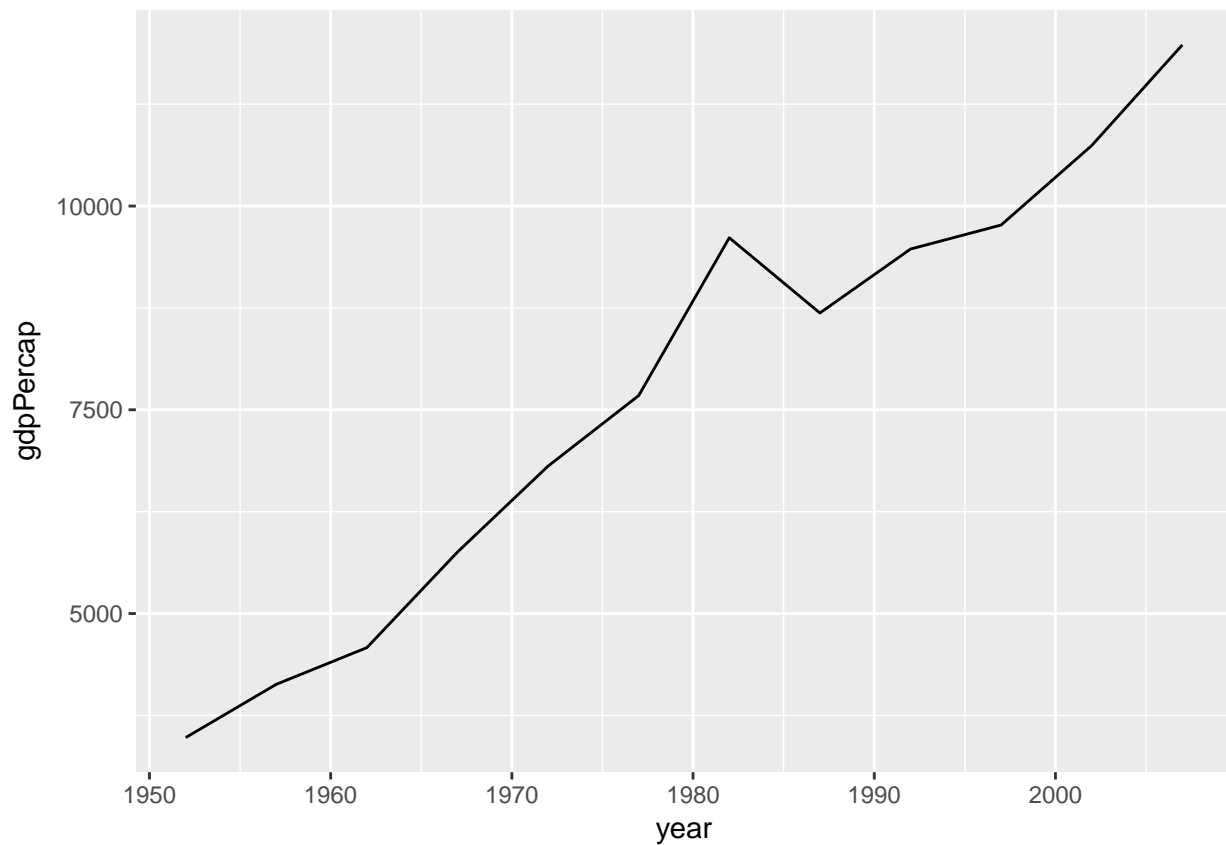
```
data <- gapminder %>% filter(country == 'Mexico')

g <- ggplot(data, aes(x=year, y=pop))
g + geom_line()
```



y el gdpPercapita...

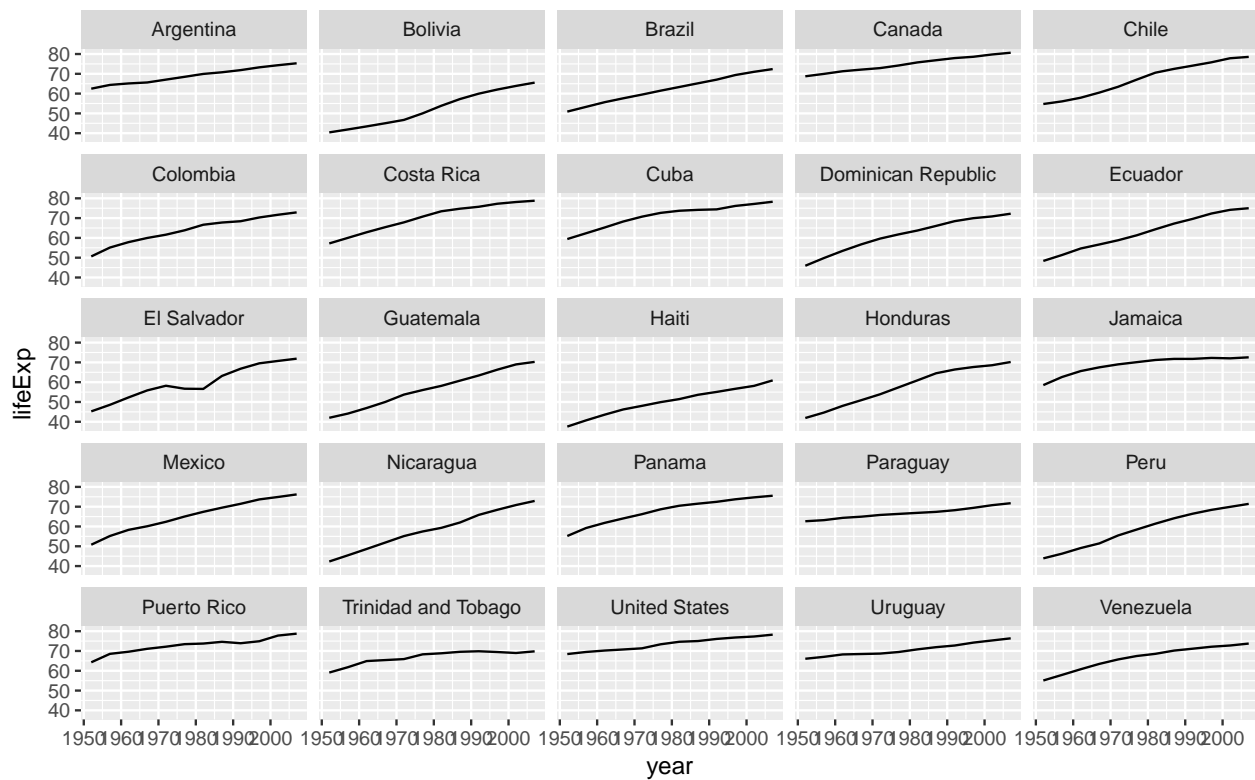
```
data <- gapminder %>% filter(country == 'Mexico')  
  
g <- ggplot(data, aes(x=year, y=gdpPercap))  
g + geom_line()
```



Podríamos comparar el comportamiento de México contra su continente:

Esperanza de vida:

```
data <- gapminder %>% filter(continent == 'Americas')  
  
g <- ggplot(data, aes(x=year, y=lifeExp))  
g + geom_line() + facet_wrap(~ country)
```

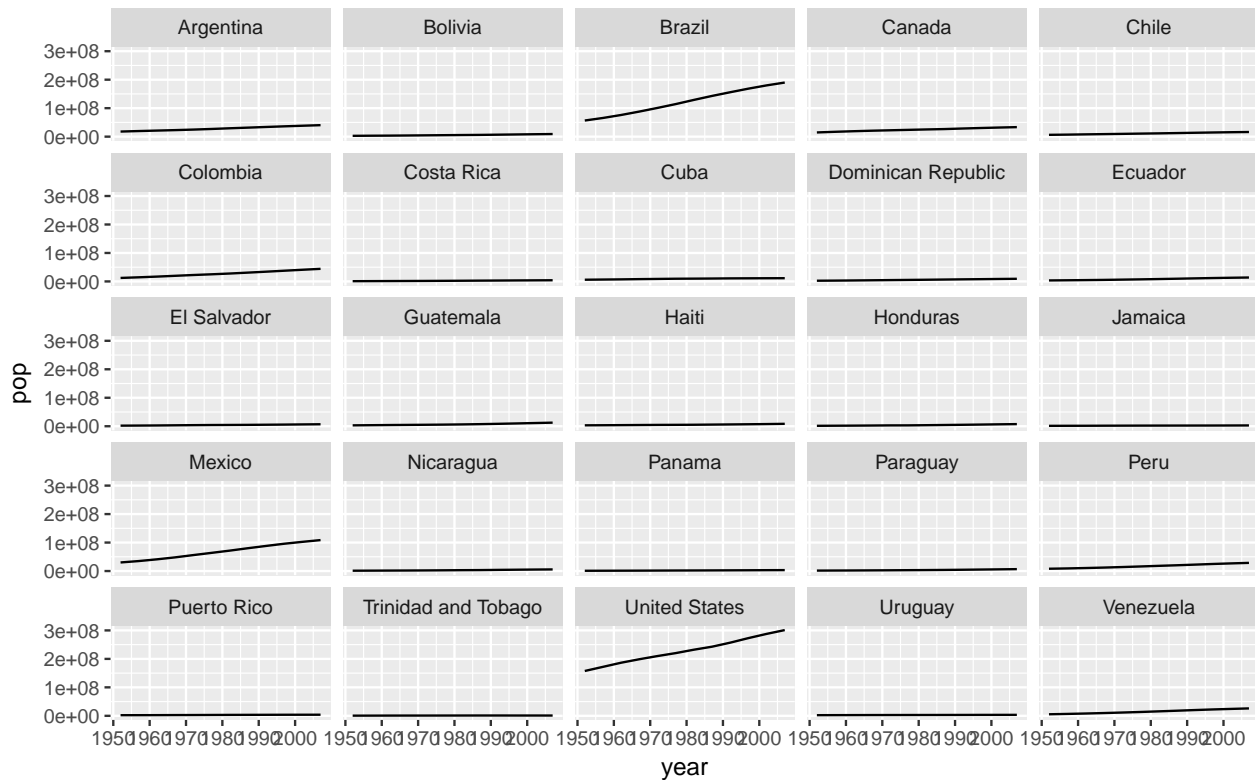


Es fácil notar que **Canadá** y **Estados Unidos** tienen los mayores niveles de *esperanza de vida*

Población:

```
data <- gapminder %>% filter(continent == 'Americas')

g <- ggplot(data, aes(x=year, y=pop))
g + geom_line() + facet_wrap(~ country)
```

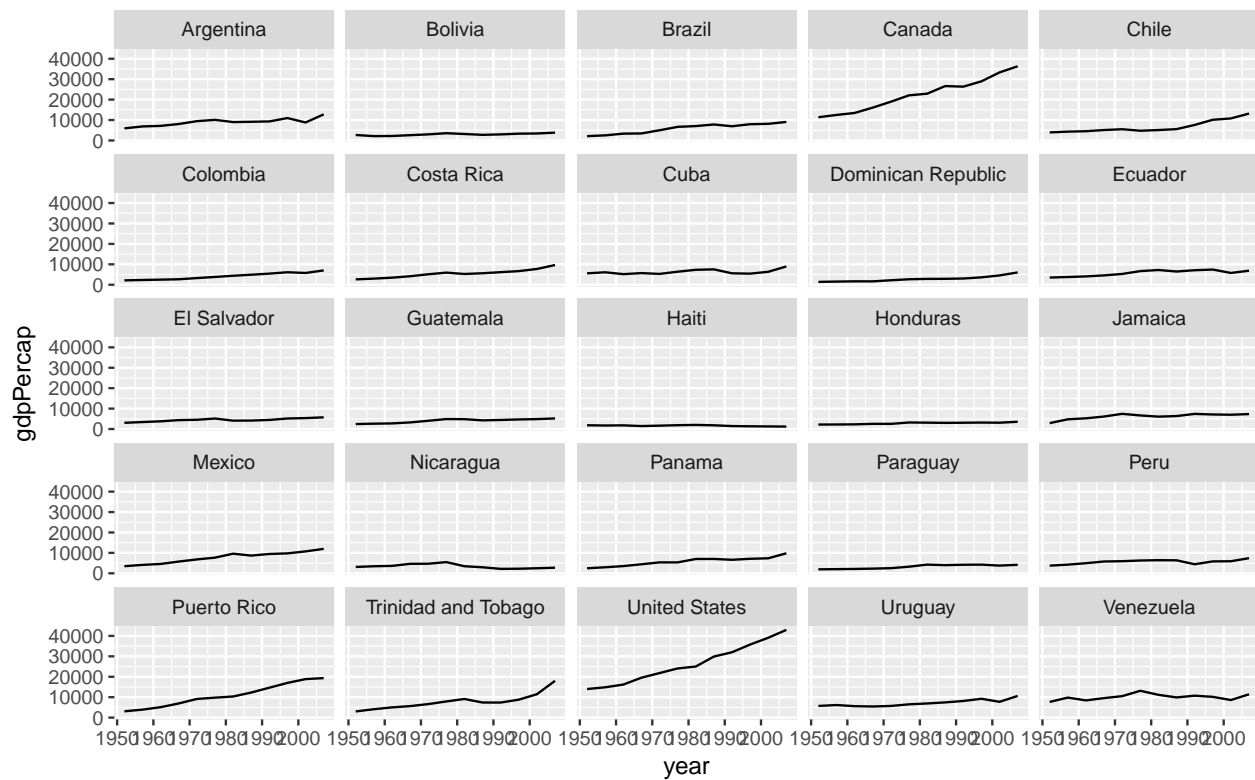


Los países que sobresalen por su crecimiento poblacional a lo largo de los años son: **Estados Unidos, México y Brasil**

GDP:

```
data <- gapminder %>% filter(continent == 'Americas')

g <- ggplot(data, aes(x=year, y=gdpPercap))
g + geom_line() + facet_wrap(~ country)
```

Aqui vemos que los dos países que sobresalen en GDP en el continente son **EEUU** y **Canadá**