

# Estadística I

## Guía de Ejercicios 2

**Profesor:** *Víctor Macías E.*

En esta guía se trabajará con el dataset **gapminder** que se encuentra disponible en el paquete **gapminder**. Se recomienda revisar el siguiente link: <https://cran.r-project.org/web/packages/gapminder/>

### Pregunta 1

- (a) ¿Cuántas observaciones y variables tiene el dataset **gapminder**? ¿Qué representa cada variable? ¿En qué unidad se miden **lifeExp**, **pop** y **gdpPercap**?
- (b) ¿Cuántas variables son cuantitativas? ¿Cuántas son categóricas?

### Pregunta 2

Muestre las primeras 8 observaciones del conjunto de datos

### Pregunta 3

¿Cuántos países por año son incluidos en la muestra? ¿Para qué años hay información disponible?

### Pregunta 4

Construya la siguiente tabla que muestra el número de países para cada uno de los años y continentes ¿Qué puede concluir?

year/continent	Africa	Americas	Asia	Europe	Oceania	Total
1952	52	25	33	30	2	142
1957	52	25	33	30	2	142
1962	52	25	33	30	2	142
1967	52	25	33	30	2	142
1972	52	25	33	30	2	142
1977	52	25	33	30	2	142
1982	52	25	33	30	2	142
1987	52	25	33	30	2	142
1992	52	25	33	30	2	142
1997	52	25	33	30	2	142
2002	52	25	33	30	2	142
2007	52	25	33	30	2	142
Total	624	300	396	360	24	1704

## Pregunta 5

Construya la siguiente tabla que muestra estadística descriptiva para las variables **pop**, **lifeExp** y **gdpPercap** para el año 2007.

Variable	Media	Desv. Estándar	Mínimo	p25	p50	p75	Máximo
pop	44021219.6	147621397.9	199579.0	4508033.5	10517531.0	31210041.8	1318683096.0
lifeExp	67.0	12.1	39.6	57.2	71.9	76.4	82.6
gdpPercap	11680.1	12859.9	277.6	1624.8	6124.4	18008.8	49357.2

## Pregunta 6

Construya la siguiente tabla que muestra estadística descriptiva para la variable **lifeExp** por continente para el año 2007.

Continente	Media	Desv. Estándar	Mínimo	p25	p50	p75	Máximo
Africa	54.8	9.6	39.6	47.8	52.9	59.4	76.4
Americas	73.6	4.4	60.9	71.8	72.9	76.4	80.7
Asia	70.7	8.0	43.8	65.5	72.4	75.6	82.6
Europe	77.6	3.0	71.8	75.0	78.6	79.8	81.8
Oceania	80.7	0.7	80.2	80.5	80.7	81.0	81.2

## Pregunta 7

Construya la siguiente tabla que identifica a los 5 países con mayor población el año 2007

País	Población
China	1318683096
India	1110396331
United States	301139947
Indonesia	223547000
Brazil	190010647

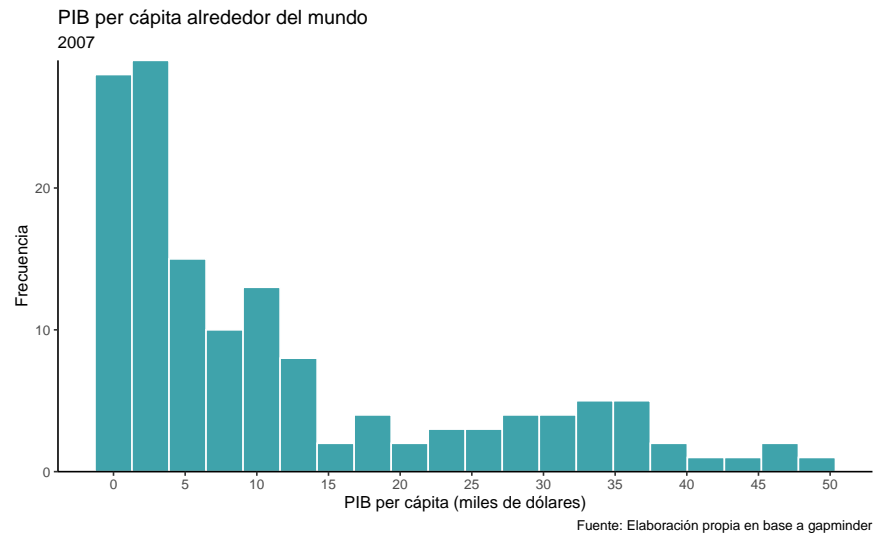
## Pregunta 8

Construya la siguiente tabla que identifica a los 7 países con menor PIB per cápita el año 2007

País	PIB per cápita (dólares)
Congo, Dem. Rep.	278
Liberia	415
Burundi	430
Zimbabwe	470
Guinea-Bissau	579
Niger	620
Eritrea	641

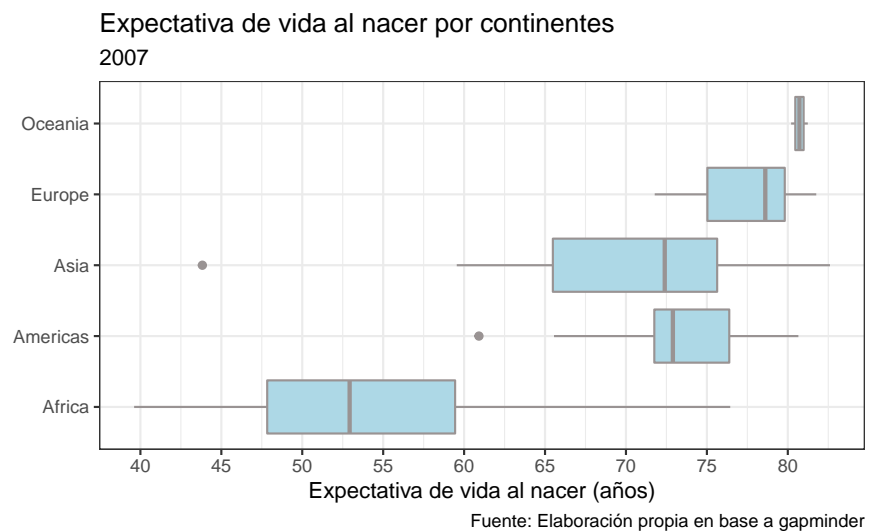
## Pregunta 9

Construya el siguiente histograma que representa la distribución de frecuencias del PIB per cápita el año 2007 ¿Qué puede concluir de la forma de esta distribución? (Ayuda: En este gráfico se usó `bins = 20` para especificar el ancho de los bins en `geom_histogram()`. Otra opción es usar `binwidth`)



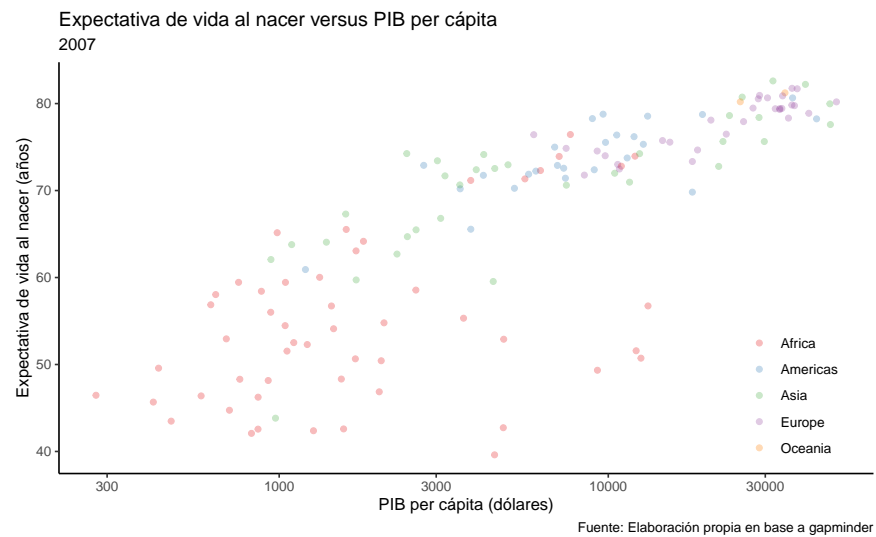
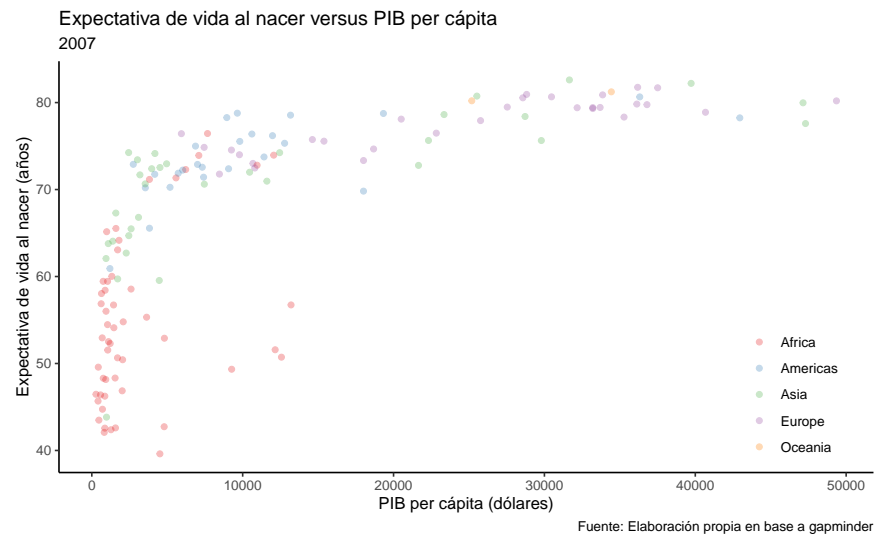
## Pregunta 10

Construya el siguiente diagrama de caja para la variable `lifeExp` en el año 2007 (Ayuda: En este gráfico se especificó `ggplot(aes(x = continent, y = lifeExp))`, de modo que en el eje X se encuentre la variable `continent` y en el eje Y, la variable `lifeExp`. Posteriormente, se agregó `coord_flip()` para rotar los ejes)



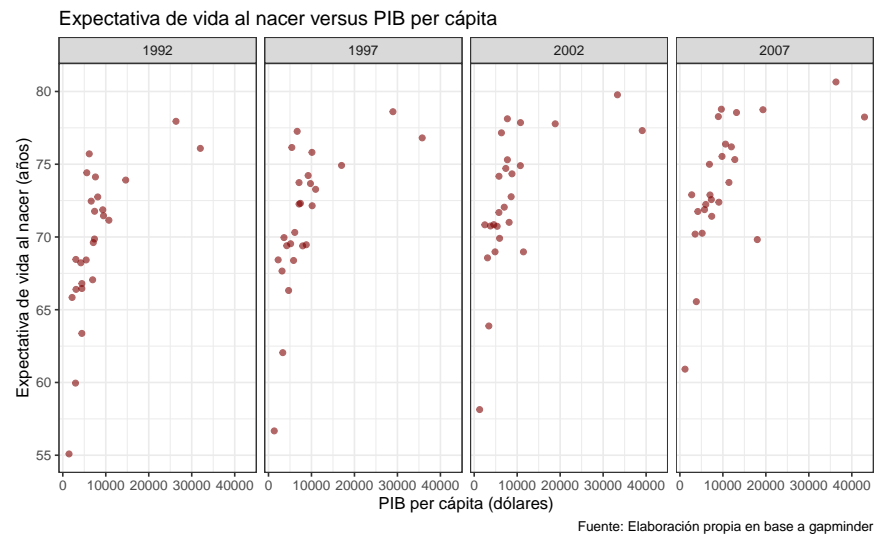
## Pregunta 11

Construya el siguiente gráfico de dispersión que muestra la relación entre de la expectativa de vida al nacer y PIB per cápita para el año 2007. Presente también un gráfico en que el eje X use una escala logarítmica en base 10 ¿Qué diferencias observa entre ambos gráficos? (Ayuda: En el segundo gráfico se usó `scale_x_log10()` para incorporar una escala logarítmica en el eje X)



## Pregunta 12

Construya el siguiente gráfico de dispersión que muestra la relación entre de la expectativa de vida al nacer y PIB per cápita para los años 1992, 1997, 2002 y 2007, incluyendo sólo a los países del continente americano. Use la misma escala para representar los valores del eje Y (Ayuda: Para que los cuatro gráficos sean adyacentes se usó `facet_grid(cols = vars(year))`)



## Pregunta 13

Construya el siguiente gráfico de línea que muestra la evolución de la expectativa de vida al nacer para Chile, El Salvador, Japón y Estados Unidos en el período 1952-2007.

