



Mapas a través de Library en R

Estadística Espacial

Jose Angel Urquijo Parra

August 1, 2024

información sobre nacimientos y muertes por Síndrome de Muerte Súbita del Lactante

Los datos proporcionados contienen información sobre nacimientos y muertes por Síndrome de Muerte Súbita del Lactante (SID) en varios condados de Carolina del Norte durante los periodos 1974-78 y 1979-84. A continuación se describen las variables del conjunto de datos:

- CNTY.ID: ID del condado
- BIR74: Nacimientos en 1974-78
- SID74: Muertes por SID en 1974-78
- NWBIR74: Nacimientos no blancos en 1974-78
- BIR79: Nacimientos en 1979-84
- SID79: Muertes por SID en 1979-84
- NWBIR79: Nacimientos no blancos en 1979-84
- east: Coordenadas este del asiento del condado en millas, proyección local
- north: Coordenadas norte del asiento del condado en millas, proyección local
- lon: Longitud del asiento del condado
- lat: Latitud del asiento del condado
- L.id: Índice L de Cressie y Read (1985)
- M.id: Índice M de Cressie y Read (1985)

```
if (!require(spData)) {  
  install.packages("spData")  
}  
  
library(spData)  
library(dplyr)  
library(ggplot2)  
library(knitr)  
  
data("nc.sids")
```

```
print(summary(nc.sids))
```

CNTY.ID	BIR74	SID74	NWBIR74
Min. :1825	Min. : 248	Min. : 0.00	Min. : 1.0
1st Qu.:1902	1st Qu.: 1077	1st Qu.: 2.00	1st Qu.: 190.0
Median :1982	Median : 2180	Median : 4.00	Median : 697.5
Mean :1986	Mean : 3300	Mean : 6.67	Mean :1051.0
3rd Qu.:2067	3rd Qu.: 3936	3rd Qu.: 8.25	3rd Qu.:1168.5
Max. :2241	Max. :21588	Max. :44.00	Max. :8027.0

BIR79	SID79	NWBIR79	east
Min. : 319	Min. : 0.00	Min. : 3.0	Min. : 19.0
1st Qu.: 1336	1st Qu.: 2.00	1st Qu.: 250.5	1st Qu.:178.8
Median : 2636	Median : 5.00	Median : 874.5	Median :285.0
Mean : 4224	Mean : 8.36	Mean : 1352.8	Mean :271.3
3rd Qu.: 4889	3rd Qu.:10.25	3rd Qu.: 1406.8	3rd Qu.:361.2
Max. :30757	Max. :57.00	Max. :11631.0	Max. :482.0

north	x	y	lon
Min. : 6.0	Min. :-328.04	Min. :3757	Min. :-84.08
1st Qu.: 97.0	1st Qu.: -60.55	1st Qu.:3920	1st Qu.: -81.20
Median :125.5	Median : 114.38	Median :3963	Median : -79.26
Mean :122.1	Mean : 91.46	Mean :3953	Mean : -79.51
3rd Qu.:151.5	3rd Qu.: 240.03	3rd Qu.:4000	3rd Qu.: -77.87
Max. :182.0	Max. : 439.65	Max. :4060	Max. : -75.67

lat	L.id	M.id
Min. :33.92	Min. :1.00	Min. :1.00
1st Qu.:35.26	1st Qu.:1.00	1st Qu.:2.00
Median :35.68	Median :2.00	Median :3.00
Mean :35.62	Mean :2.12	Mean :2.67
3rd Qu.:36.05	3rd Qu.:3.00	3rd Qu.:3.25
Max. :36.52	Max. :4.00	Max. :4.00

```
# Estadísticas descriptivas
```

```
mean_rate <- mean(nc.sids$SID79, na.rm = TRUE)
```

```
median_rate <- median(nc.sids$SID79, na.rm = TRUE)
```

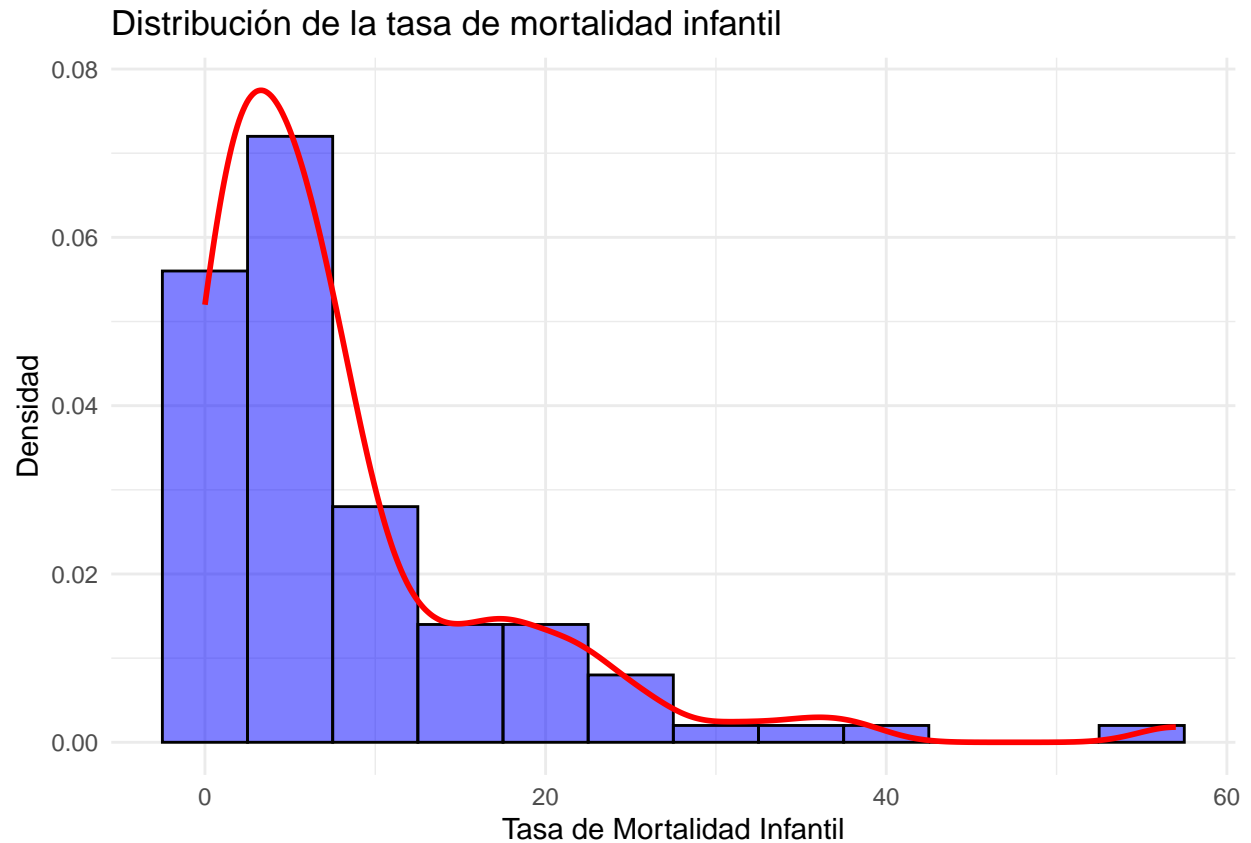
```
sd_rate <- sd(nc.sids$SID79, na.rm = TRUE)
```

```
cat("Tasa media de mortalidad infantil:",
```

```
    mean_rate, ". Mediana:", median_rate, "Desviacion estandar:", sd_rate, ".")
```

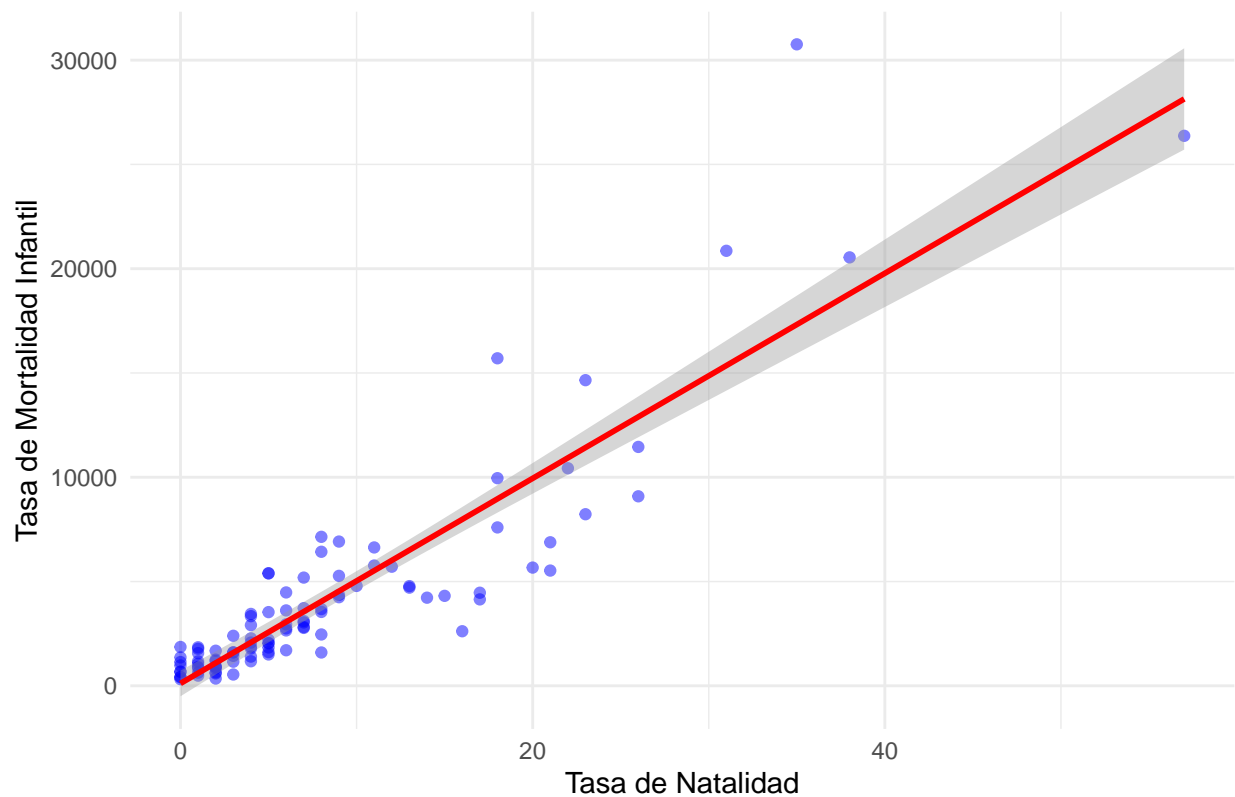
```
Tasa media de mortalidad infantil: 8.36 . Mediana: 5 Desviacion estandar: 9.431861 .
```

```
ggplot(nc.sids, aes(x = SID79)) +
  geom_histogram(aes(y = ..density..), binwidth = 5, fill = "blue",
                 color = "black", alpha = 0.5) +
  geom_density(color = "red", size = 1) +
  labs(title = "Distribución de la tasa de mortalidad infantil",
       x = "Tasa de Mortalidad Infantil",
       y = "Densidad") +
  theme_minimal()
```



```
ggplot(nc.sids, aes(x = SID79 , y = BIR79)) +  
  geom_point(alpha = 0.5, color = "blue") +  
  geom_smooth(method = "lm", color = "red", size = 1) +  
  labs(title = "Relacion entre la Tasa de Mortalidad Infantil y la Tasa de Natalidad",  
        x = "Tasa de Natalidad",  
        y = "Tasa de Mortalidad Infantil") +  
  theme_minimal()
```

Relacion entre la Tasa de Mortalidad Infantil y la Tasa de Natalidad



```
library(sf)
library(tigris)
library(ggplot2)
library(viridis)
```

```
# Cargar el mapa de los condados de Carolina del Norte
counties <- counties(state = "NC", cb = TRUE)
```

```
# Asegúrate de que el CRS de los condados sea EPSG:4326
counties <- st_transform(counties, crs = 4269)
```

```
# Convertir los datos a formato espacial utilizando las columnas 'lon' y 'lat'
nc.sids_sf <- st_as_sf(nc.sids, coords = c("lon", "lat"), crs = 4269)
```

```
# Unir los datos espaciales con el mapa de los condados
nc_map <- st_join(counties, nc.sids_sf)
```

```
# Filtrar los condados con mortalidad infantil por encima de 50
high_mortality <- nc_map %>% filter(SID79 > 30)
```

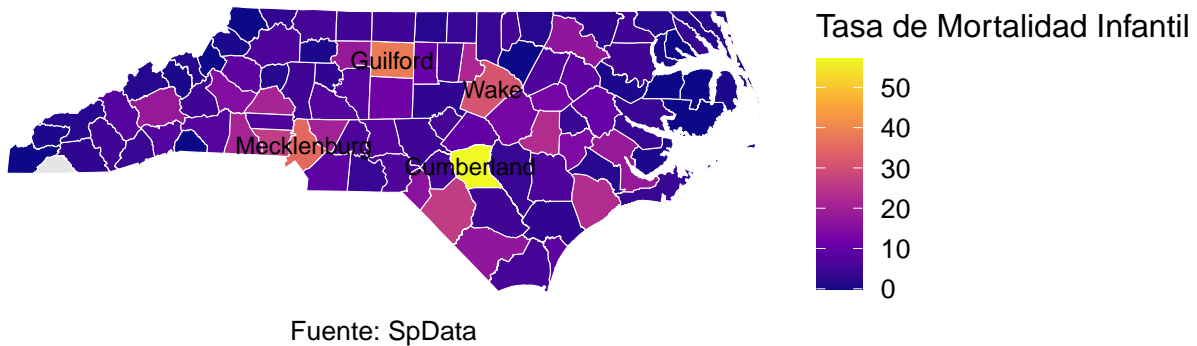
```
# Crear el mapa
ggplot(data = nc_map) +
```

```

geom_sf(aes(fill = SID79), color = "white", size = 0.2) +
scale_fill_viridis_c(option = "C", na.value = "grey90") +
labs(
  title = "Mortalidad Infantil en Carolina del Norte",
  fill = "Tasa de Mortalidad Infantil",
  caption = "Fuente: SpData"
) +
theme_minimal(base_size = 14) +
theme(
  plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
  plot.caption = element_text(size = 10, hjust = 0.5),
  legend.position = "right",
  legend.title = element_text(size = 12),
  legend.text = element_text(size = 10),
  panel.grid = element_blank(),
  axis.text = element_blank(),
  axis.title = element_blank()
) +
coord_sf(expand = FALSE) +
geom_sf_text(data = high_mortality, aes(label = NAME), size = 3, color = "black", check_overlap = TRUE)

```

Mortalidad Infantil en Carolina del Norte



Datos del mapa mundial de Natural Earth con algunas variables del Banco Mundial.

los datos contienen un `data.frame` con 177 observaciones de 11 variables:

- **iso_a2:** vector de caracteres con códigos de país ISO de 2 caracteres
- **name_long:** vector de caracteres con nombres de países
- **continent:** vector de caracteres con nombres de continentes
- **region_un:** vector de caracteres con nombres de regiones
- **subregion:** vector de caracteres con nombres de subregiones
- **type:** vector de caracteres con nombres de tipos
- **area_km2:** vector de enteros con valores de área
- **pop:** vector de enteros con población en 2014
- **lifeExp:** vector de enteros con esperanza de vida al nacer en 2014
- **gdpPercap:** vector de enteros con PIB per cápita en 2014
- **geom:** `sfc_MULTIPOLYGON`

```
library(spData)
data("world")

print(summary(world))
```

iso_a2	name_long	continent	region_un
Length:177	Length:177	Length:177	Length:177
Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character

subregion	type	area_km2	pop
Length:177	Length:177	Min. : 2417	Min. :5.630e+04
Class :character	Class :character	1st Qu.: 46185	1st Qu.:3.755e+06
Mode :character	Mode :character	Median : 185004	Median :1.040e+07
		Mean : 832558	Mean :4.282e+07
		3rd Qu.: 621860	3rd Qu.:3.075e+07
		Max. :17018507	Max. :1.364e+09
			NA's :10

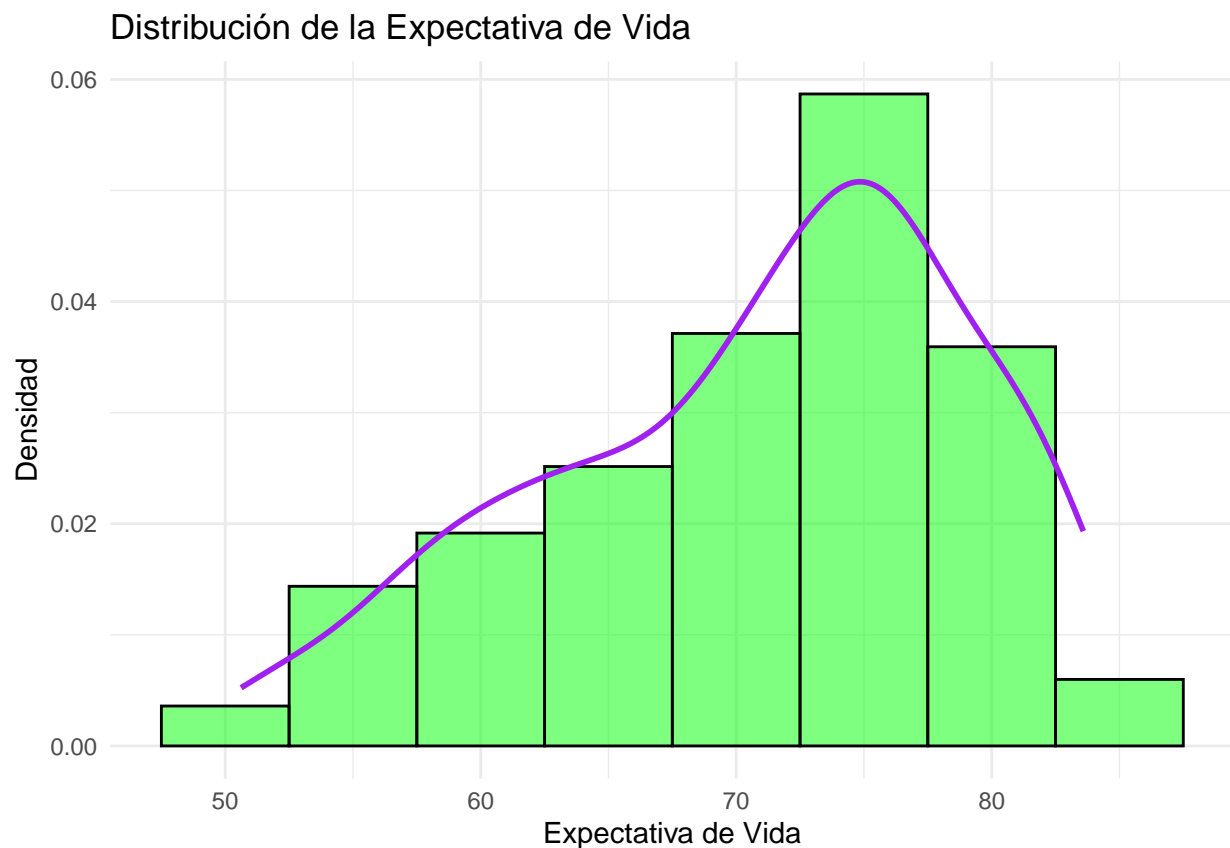
lifeExp	gdpPercap	geom
Min. :50.62	Min. : 597.1	MULTIPOLYGON :177
1st Qu.:64.96	1st Qu.: 3752.4	epsg:4326 : 0
Median :72.87	Median : 10734.1	+proj=long...: 0
Mean :70.85	Mean : 17106.0	
3rd Qu.:76.78	3rd Qu.: 24232.7	
Max. :83.59	Max. :120860.1	
NA's :10	NA's :17	

```
# Estadísticas descriptivas
mean_rate <- mean(world$gdpPercap, na.rm = TRUE)
median_rate <- median(world$gdpPercap, na.rm = TRUE)
sd_rate <- sd(world$gdpPercap, na.rm = TRUE)

cat("Tasa media del PBI:",
    mean_rate, " Mediana:", median_rate, " Desviacion estandar:", sd_rate)
```

Tasa media del PBI: 17105.99 Mediana: 10734.07 Desviacion estandar: 18668.07

```
ggplot(world, aes(x = lifeExp)) +
  geom_histogram(aes(y = ..density..), binwidth = 5, fill = "green",
                 color = "black", alpha = 0.5) +
  geom_density(color = "purple", size = 1) +
  labs(title = "Distribución de la Expectativa de Vida",
       x = "Expectativa de Vida",
       y = "Densidad") +
  theme_minimal()
```



```
# Ajustar el modelo lineal con la transformación logarítmica del PIB per cápita
word1 <- lm(log(gdpPercap) ~ lifeExp, data = world)

# Mostrar el resumen del modelo ajustado
summary(word1)
```

Call:

```
lm(formula = log(gdpPercap) ~ lifeExp, data = world)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.38254	-0.43071	-0.02034	0.36924	2.81894

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.860804	0.490115	1.756	0.081 .
lifeExp	0.116813	0.006862	17.024	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7101 on 158 degrees of freedom

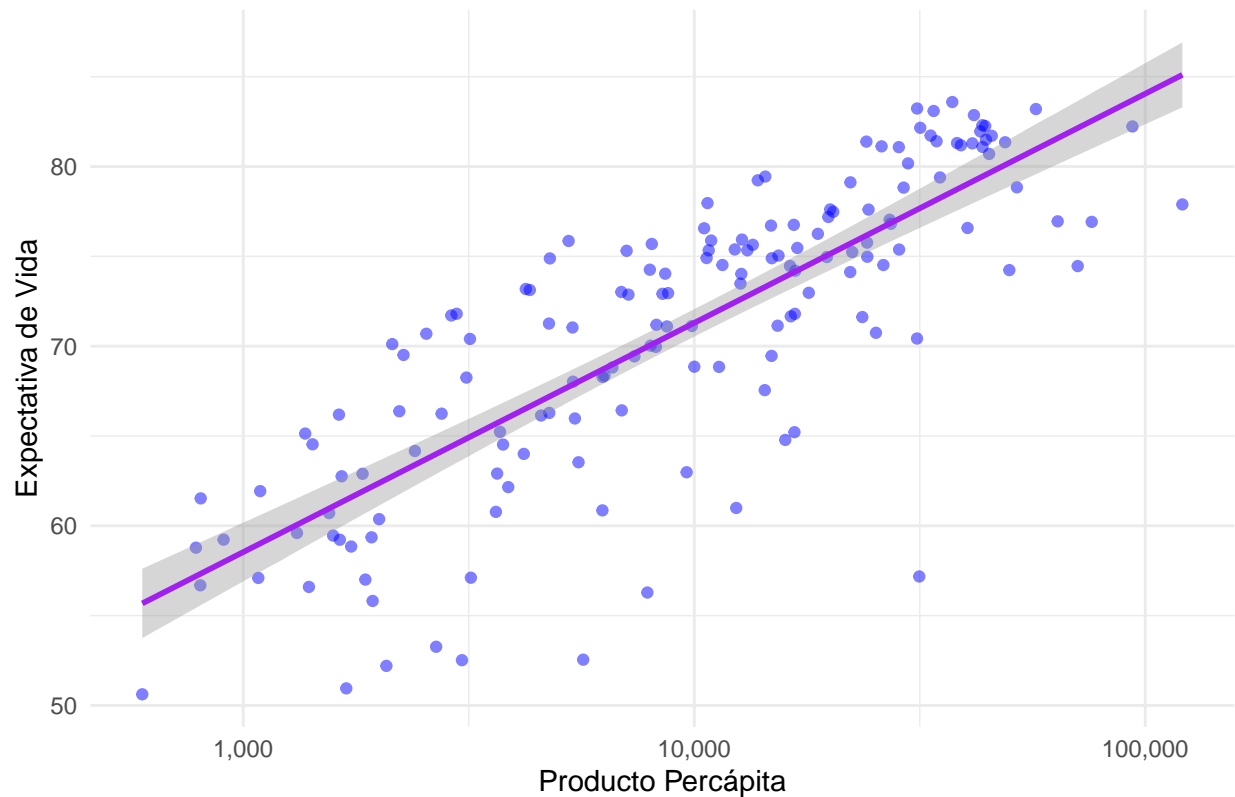
(17 observations deleted due to missingness)

Multiple R-squared: 0.6472, Adjusted R-squared: 0.6449

F-statistic: 289.8 on 1 and 158 DF, p-value: < 2.2e-16

```
# Crear el gráfico
ggplot(world, aes(x = gdpPercap, y = lifeExp)) +
  geom_point(alpha = 0.5, color = "blue") + # Puntos de datos
  geom_smooth(method = "lm", formula = y ~ x, color = "purple", size = 1) + # Línea de ajuste
  scale_x_log10(labels = scales::comma) + # Aplicar escala logarítmica al eje x
  labs(title = "Relación entre el Producto Percápita y la Expectativa de Vida",
       x = "Producto Percápita",
       y = "Expectativa de Vida") +
  theme_minimal()
```


Relación entre el Producto Percápita y la Expectativa de Vida



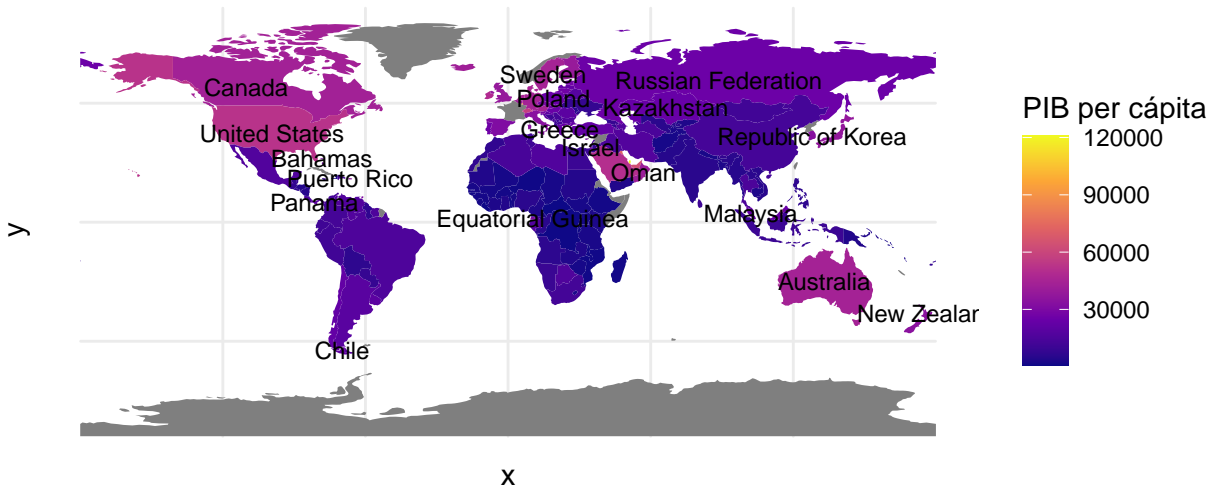
```
library(sf)
library(ggplot2)
library(viridis)
library(dplyr)

# Filtrar los países con un PIB per cápita mayor a 20,000
world <- world %>% mutate(top_gdpPercap = if_else(gdpPercap > 20000, name_long, NA_character_))

# Crear el mapa
ggplot(data = world) +
  geom_sf(aes(fill = gdpPercap), color = NA) +
  scale_fill_viridis_c(option = "C", name = "PIB per cápita") +
  geom_sf_text(aes(label = top_gdpPercap),
              size = 3, color = "black", check_overlap = TRUE) +
  theme_minimal() +
  labs(
    title = "Mapa Mundial del PIB per cápita en 2014",
    subtitle = "Datos del Banco Mundial",
    caption = "Fuente: Natural Earth y Banco Mundial"
  )
)
```

Mapa Mundial del PIB per cápita en 2014

Datos del Banco Mundial



Fuente: Natural Earth y Banco Mundial