

Informe del Proyecto

Rosalia

2025-08-08

```
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.4.3
## Warning: package 'ggplot2' was built under R version 4.4.3
## Warning: package 'tibble' was built under R version 4.4.3
## Warning: package 'tidyr' was built under R version 4.4.3
## Warning: package 'readr' was built under R version 4.4.3
## Warning: package 'purrr' was built under R version 4.4.3
## Warning: package 'dplyr' was built under R version 4.4.3
## Warning: package 'stringr' was built under R version 4.4.3
## Warning: package 'forcats' was built under R version 4.4.3
## Warning: package 'lubridate' was built under R version 4.4.3

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## vforcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr    1.3.1
## v purrr    1.0.4
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(readxl)
library(readr)
library(dplyr)
library(ggplot2)
library(ggrepel)

## Warning: package 'ggrepel' was built under R version 4.4.3
```

```
library(countrycode)

## Warning: package 'countrycode' was built under R version 4.4.3

library(scales)

## Warning: package 'scales' was built under R version 4.4.3

## 
## Adjuntando el paquete: 'scales'
##
## The following object is masked from 'package:purrr':
## 
##     discard
## 
## The following object is masked from 'package:readr':
## 
##     col_factor

library(maps)

## Warning: package 'maps' was built under R version 4.4.3

## 
## Adjuntando el paquete: 'maps'
##
## The following object is masked from 'package:purrr':
## 
##     map

library(DT)

## Warning: package 'DT' was built under R version 4.4.3

library(lorem)

## Warning: package 'lorem' was built under R version 4.4.3

library(png)
library(grid)
library(tinytex)

## Warning: package 'tinytex' was built under R version 4.4.3

library(corrplot)

## Warning: package 'corrplot' was built under R version 4.4.2
```

```
## corrplot 0.95 loaded
```

Introducción

Esta propuesta de diseñar una investigación en ciencias de datos me brinda la libertad de explorar con mayor profundidad temas sociales que me atraviesan personalmente. Al revisar las bases de datos disponibles en Gapminder, encontré una gran variedad de información relacionada con las mujeres, lo que despertó en mí una inquietud sobre la posible relación entre la longevidad femenina y la maternidad. En Costa Rica, existe una expresión popular según la cual “tener hijos te riega las bilis”, lo que en el imaginario colectivo podría asociarse con la idea de que la maternidad acorta la vida. Esta creencia me llevó a preguntarme si realmente existe una relación significativa entre la esperanza de vida de las mujeres y la cantidad promedio de hijos o hijas que tienen. En lo personal, mi familia ilustra esta interrogante de forma particular. Mi madre tuvo una sola hermana, quien no tuvo hijos y falleció a una edad relativamente baja para los estándares actuales. Por otro lado, espero que mi madre —quien sí tuvo hijos— viva muchos más años que el promedio nacional. Aunque anecdótica, esta observación familiar me motivó a profundizar en la investigación científica sobre el tema.

Entre los estudios que llamaron mi atención fue Vaupel (2001) plantea que podría existir una relación positiva entre el número de hijos y la longevidad en las mujeres, especialmente cuando el último parto ocurre a una edad avanzada. Por su parte, Zhang et al. (2023) afirman que las mujeres que tienen entre tres y cinco hijos tienden a vivir más tiempo, mientras que aquellas con uno o ningún hijo presentan una esperanza de vida más corta. Me resulta interesante cómo estas investigaciones se reflejan, de alguna forma, en la experiencia de mi propia familia.

Por otro lado, Hampton (2021) analiza la relación entre desarrollo y fecundidad, señalando que, aunque existe la idea de que a mayor desarrollo económico se podría esperar un aumento de la fecundidad, en la práctica no se observa una asociación directa. Su estudio concluye que, al menos en la última década, los niveles elevados de desarrollo humano ya no están vinculados con repuntes en la tasa de natalidad.

Ahora bien, Ziepple, Reeve y Peniston (2024) realizaron un estudio sobre la relación entre maternidad y esperanza de vida, en el cual plantean que la interacción madre-cría desempeña un papel fundamental en la extensión de la vida tanto en primates humanos como no humanos. A través de la formulación de hipótesis evolutivas, concluyen que la conexión entre la supervivencia materna y la aptitud reproductiva —entendida como el éxito de la descendencia— puede favorecer vidas más largas con menor frecuencia reproductiva. Este hallazgo sugiere que, en ciertos contextos, la longevidad podría estar más vinculada al cuidado prolongado que al número de hijos, abriendo nuevas perspectivas sobre el vínculo entre maternidad y salud a largo plazo.

Este estudio se inserta en uno de los debates sobre el crecimiento poblacional, en el cual se observa que la población sigue en aumento. Sin embargo, dependiendo del continente, el debate se matiza según la cantidad de hijos por mujer.

En este contexto, considero relevante abrir la discusión sobre lo que representa la decisión de tener hijos y su posible vínculo con la longevidad. Estos hallazgos no solo enriquecen el debate científico, sino que también invitan a repensar la maternidad desde una perspectiva más informada, contextualizada y consciente de sus implicaciones a largo plazo.

Objetivos

Ante estas inquietudes el Objetivo General es:

Analizar la relación entre la esperanza de vida femenina, la fecundidad y el nivel de desarrollo económico en países del mundo, utilizando datos históricos de Gapminder entre 1950 y 2021,

En cuanto a los objetivos específicos son:

1. Comparar la esperanza de vida femenina, la fecundidad y el producto interno bruto per cápita entre regiones geográficas y a lo largo del tiempo, para identificar asociaciones significativas y variaciones contextuales

2. Explorar combinaciones bivariadas entre esperanza de vida femenina, fecundidad y PIB per cápita para identificar patrones y tendencias que evidencien diferencias significativas entre países con distintos niveles de desarrollo económico.
3. Evaluar la relación conjunta entre PIB per cápita, fecundidad y esperanza de vida femenina, con el fin de caracterizar la correlación multivariable y su comportamiento en distintos contextos socioeconómicos.
4. Diseñar y presentar un dashboard interactivo que visualice los hallazgos derivados del análisis de las tres variables, facilitando la exploración dinámica de patrones por región, año y nivel de desarrollo económico.

Metodología de investigación

Esta investigación se basa en un enfoque exploratorio utilizando información agregada a nivel país, lo que permite analizar datos históricos sobre fecundidad, esperanza de vida y producto interno bruto per cápita. Es importante aclarar que, debido a la naturaleza agregada de los datos, no es posible establecer relaciones causales individuales. Lo que se observa son asociaciones a nivel macro, que pueden orientar futuras investigaciones más específicas.

Como se menciona en los objetivos, he decidido trabajar con tres conjuntos de datos disponibles en Gapminder para desarrollar un análisis exploratorio de datos:

- *Life_expectancy_female*: esperanza de vida al nacer de mujeres (en años).
- *Children_per_women_total_fertility*: número promedio de hijos por mujer a lo largo de su vida.
- *Gross domestic product per person*: producto interno bruto per cápita ajustado por paridad de poder adquisitivo al año 2021. Estos conjuntos de datos cubren períodos desde 1800 hasta 2021, por lo que el análisis se centrará en los datos comunes disponibles a partir de 1950.

Al revisar las bases de datos utilizadas, se identifican cinco variables principales: *year*, *country*, *esperanza_de_vida*, *nacimientos_por_mujer* y *producto_por_capita*.

Las variables principales utilizadas fueron: *year*, *country*, *esperanza_de_vida*, *nacimientos_por_mujer* y *producto_por_capita*. Aunque Gapminder actualiza sus datos periódicamente, la mayoría de las series disponibles tienen como año más reciente el 2021. Dado que se trata de indicadores globales comparables entre países, la actualización depende de fuentes oficiales como el Banco Mundial o la ONU. Por ello, se recomienda revisar las series anualmente, aunque algunas pueden actualizarse con menor frecuencia.

Datos Utilizados

```
Datos_FEGDP <- readr::read_csv("../Datos/Base_Datos_depurada/base_completa.csv", locale = locale(encoding = "UTF-8"))

## # Rows: 35636 Columns: 7
## -- Column specification -----
## Delimiter: ","
## chr (3): country, continente, gdp_grupo
## dbl (4): year, esperanza_de_vida, producto_por_capita, nacimientos_por_mujer
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

Vars <- paste0(c("esperanza_de_vida", "producto_por_capita", "nacimientos_por_mujer", "continente", "gdp_grupo"))
```

Los países que se encuentran en la base de datos fueron:

```
unique(Datos_FEGDP$country)
```

```

## [1] "Aruba"
## [3] "Angola"
## [5] "Albania"
## [7] "UAE"
## [9] "Armenia"
## [11] "Antigua and Barbuda"
## [13] "Austria"
## [15] "Burundi"
## [17] "Benin"
## [19] "Bangladesh"
## [21] "Bahrain"
## [23] "Bosnia and Herzegovina"
## [25] "Belize"
## [27] "Bolivia"
## [29] "Barbados"
## [31] "Bhutan"
## [33] "Central African Republic"
## [35] "Switzerland"
## [37] "China"
## [39] "Cameroon"
## [41] "Cook Is"
## [43] "Comoros"
## [45] "Costa Rica"
## [47] "Cayman Islands"
## [49] "Czech Republic"
## [51] "Djibouti"
## [53] "Denmark"
## [55] "Algeria"
## [57] "Egypt"
## [59] "Western Sahara"
## [61] "Estonia"
## [63] "Finland"
## [65] "Falkland Is (Malvinas)"
## [67] "Faeroe Islands"
## [69] "Gabon"
## [71] "Isle of Man"
## [73] "Georgia"
## [75] "Gibraltar"
## [77] "Guadeloupe"
## [79] "Guinea-Bissau"
## [81] "Greece"
## [83] "Greenland"
## [85] "French Guiana"
## [87] "Guyana"
## [89] "Honduras"
## [91] "Croatia"
## [93] "Hungary"
## [95] "India"
## [97] "Iran"
## [99] "Iceland"
## [101] "Italy"
## [103] "Jersey"
## [105] "Japan"
## [107] "Kenya"
## [109] "Lithuania"
## [111] "Luxembourg"
## [113] "Macao"
## [115] "Macedonia"
## [117] "Madagascar"
## [119] "Maldives"
## [121] "Mali"
## [123] "Malta"
## [125] "Mauritania"
## [127] "Moldova"
## [129] "Montenegro"
## [131] "Morocco"
## [133] "Mozambique"
## [135] "Myanmar"
## [137] "Namibia"
## [139] "Nepal"
## [141] "Niger"
## [143] "Nigeria"
## [145] "Oman"
## [147] "Pakistan"
## [149] "Panama"
## [151] "Papua New Guinea"
## [153] "Paraguay"
## [155] "Peru"
## [157] "Philippines"
## [159] "Poland"
## [161] "Portugal"
## [163] "Qatar"
## [165] "Romania"
## [167] "Russia"
## [169] "Saint Lucia"
## [171] "Saint Vincent and the Grenadines"
## [173] "Samoa"
## [175] "Sao Tome and Principe"
## [177] "Senegal"
## [179] "Serbia"
## [181] "Seychelles"
## [183] "Sierra Leone"
## [185] "Sri Lanka"
## [187] "Sudan"
## [189] "Syria"
## [191] "Tajikistan"
## [193] "Tanzania"
## [195] "Thailand"
## [197] "Timor-Leste"
## [199] "Togo"
## [201] "Tunisia"
## [203] "Turkey"
## [205] "Uganda"
## [207] "Ukraine"
## [209] "United Arab Emirates"
## [211] "United Kingdom"
## [213] "United States"
## [215] "Uruguay"
## [217] "Vanuatu"
## [219] "Venezuela"
## [221] "Yemen"
## [223] "Zambia"
## [225] "Zimbabwe"

```

```

## [109] "Cambodia"
## [111] "St, Kitts and Nevis"
## [113] "Kosovo"
## [115] "Lao"
## [117] "Liberia"
## [119] "St, Lucia"
## [121] "Sri Lanka"
## [123] "Lithuania"
## [125] "Latvia"
## [127] "St, Martin (French part)"
## [129] "Monaco"
## [131] "Madagascar"
## [133] "Mexico"
## [135] "North Macedonia"
## [137] "Malta"
## [139] "Montenegro"
## [141] "Northern Mariana Islands"
## [143] "Mauritania"
## [145] "Martinique"
## [147] "Malawi"
## [149] "Mayotte"
## [151] "New Caledonia"
## [153] "Nigeria"
## [155] "Niue"
## [157] "Curaçao"
## [159] "Nepal"
## [161] "New Zealand"
## [163] "Pakistan"
## [165] "Peru"
## [167] "Palau"
## [169] "Poland"
## [171] "North Korea"
## [173] "Paraguay"
## [175] "French Polynesia"
## [177] "Reunion"
## [179] "Russia"
## [181] "Saudi Arabia"
## [183] "Senegal"
## [185] "St, Helena"
## [187] "Sierra Leone"
## [189] "San Marino"
## [191] "St,-Pierre-et-Miquelon"
## [193] "South Sudan"
## [195] "Sao Tome and Principe"
## [197] "Slovak Republic"
## [199] "Sweden"
## [201] "Sint Maarten (Dutch part)"
## [203] "Syria"
## [205] "Chad"
## [207] "Thailand"
## [209] "Tokelau"
## [211] "Timor-Leste"
## [213] "Trinidad and Tobago"
## [215] "Turkey"

## [109] "Kiribati"
## [111] "South Korea"
## [113] "Kuwait"
## [115] "Lebanon"
## [117] "Libya"
## [119] "Liechtenstein"
## [121] "Lesotho"
## [123] "Luxembourg"
## [125] "Macao"
## [127] "Morocco"
## [129] "Moldova"
## [131] "Maldives"
## [133] "Marshall Islands"
## [135] "Mali"
## [137] "Myanmar"
## [139] "Mongolia"
## [141] "Mozambique"
## [143] "Montserrat"
## [145] "Mauritius"
## [147] "Malaysia"
## [149] "Namibia"
## [151] "Niger"
## [153] "Nicaragua"
## [155] "Netherlands"
## [157] "Norway"
## [159] "Nauru"
## [161] "Oman"
## [163] "Panama"
## [165] "Philippines"
## [167] "Papua New Guinea"
## [169] "Puerto Rico"
## [171] "Portugal"
## [173] "Palestine"
## [175] "Qatar"
## [177] "Romania"
## [179] "Rwanda"
## [181] "Sudan"
## [183] "Singapore"
## [185] "Solomon Islands"
## [187] "El Salvador"
## [189] "Somalia"
## [191] "Serbia"
## [193] "St, Barthélemy"
## [195] "Suriname"
## [197] "Slovenia"
## [199] "Eswatini"
## [201] "Seychelles"
## [203] "Turks and Caicos Islands"
## [205] "Togo"
## [207] "Tajikistan"
## [209] "Turkmenistan"
## [211] "Tonga"
## [213] "Tunisia"
## [215] "Tuvalu"

```

```

## [217] "Taiwan"
## [219] "Uganda"
## [221] "Uruguay"
## [223] "Uzbekistan"
## [225] "Venezuela"
## [227] "Virgin Islands (U.S.)"
## [229] "Vanuatu"
## [231] "Samoa"
## [233] "South Africa"
## [235] "Zimbabwe"

## [217] "Tanzania"
## [219] "Ukraine"
## [221] "USA"
## [223] "St, Vincent and the Grenadines"
## [225] "British Virgin Islands"
## [227] "Vietnam"
## [229] "Wallis et Futuna"
## [231] "Yemen"
## [233] "Zambia"
## [235] "Zimbabwe"

```

Preparación de los Datos

Para la preparación y análisis de los datos se utilizó el lenguaje de programación R, dada su capacidad para realizar análisis estadísticos, visualizaciones y transformaciones de datos de forma eficiente

Durante la depuración de los datos se identificaron outliers al explorar la variable esperanza de vida femenina al nacer y la tasa de fecundidad (nacimientos por mujer). Utilizando técnicas estadísticas (por ejemplo, el método del rango intercuartílico), se detectaron aproximadamente:

- 1175 valores atípicos en la variable de esperanza de vida.
- 765 valores atípicos en la variable de fecundidad. Estos valores extremos pueden deberse a múltiples factores, como conflictos armados, crisis sanitarias, cambios abruptos en la política pública o deficiencias en el registro estadístico. Además, reflejan la alta variabilidad entre países y a lo largo del tiempo, por lo que deben ser considerados con atención en el análisis. No necesariamente deben ser eliminados, ya que pueden contener información relevante para explicar fenómenos estructurales.

Se realizaron transformaciones adicionales para corregir errores en nombres de países y se generaron nuevas variables categóricas, como continente y nivel de GDP (alto/bajo)

Para el análisis de los datos se empleó la estadística descriptiva e inferencial y se utilizaron las siguientes librerías de R - tidyverse - readxl - readr - dplyr - ggplot2 - ggrepel - countrycode - scales - maps - DT - lorem - png - grid - tinytex

Visualización

Se emplearon diversos tipos de gráficos para explorar las relaciones entre variables

1. Boxplot: a. Función: Comparar la distribución de una variable continua entre grupos categóricos. b. Aplicación: Se utilizará para mostrar diferencias en la esperanza de vida según continente, nivel de GDP (alto/bajo) o año.

2. Gráfico de dispersión:

- a. Función: Visualizar relaciones bivariadas y detectar correlaciones.
- b. Aplicación: Exploración de relaciones como:

3. Mapa de calor:

- a. Función: Visualizar patrones de densidad o intensidad en una matriz de dos variables.
- b. Aplicación: Relación entre país y año respecto a esperanza de vida o fecundidad.

4. Histograma:

- a. Función: Mostrar la distribución de una variable cuantitativa, destacando su frecuencia y variabilidad.
- b. Aplicación: Se utilizará para mostrar la distribución de la esperanza de vida femenina y de la fecundidad (hijos por mujer).

Reproducibilidad

Para garantizar la reproducibilidad del análisis se utilizaron herramientas como RStudio, RMarkdown y Git/GitHub, permitiendo documentar el proceso, versionar el código y compartir los resultados de forma transparente.

Resultados Generales

A partir de las investigaciones presentadas en la introducción, surge la pregunta central ¿Tener menos hijos realmente alarga la esperanza de vida femenina, o tener entre dos y cinco hijos genera un efecto similar?

Este interrogante guía el análisis exploratorio presentado en el dashboard. En particular, el gráfico que relaciona fecundidad y esperanza de vida muestra una tendencia general decreciente: A mayor número de hijos por mujer, menor esperanza de vida.

Este patrón sugiere que podría existir una relación inversa entre reproducción y longevidad, aunque se requiere un análisis más profundo para considerar factores contextuales y causales.

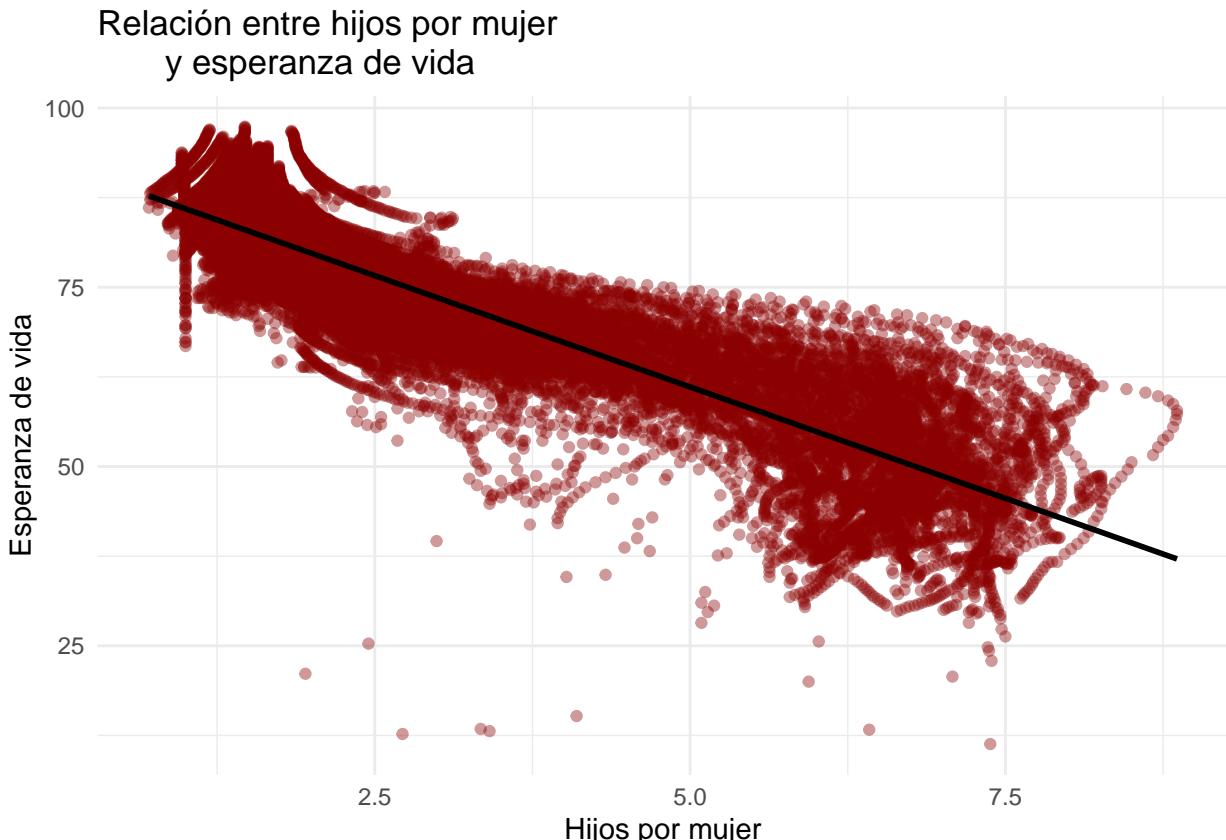
```
datos_filtrados <- Datos_FEGDP

ggplot(datos_filtrados, aes(x = nacimientos_por_mujer, y = esperanza_de_vida)) +
  geom_point(alpha = 0.4, color = "darkred") +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  labs(title = "Relación entre hijos por mujer  
y esperanza de vida", x = "Hijos por mujer", y = "Esperanza de vida") +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 5893 rows containing non-finite outside the scale range
## ('stat_smooth()').

## Warning: Removed 5893 rows containing missing values or values outside the scale range
## ('geom_point()').
```



Si bien se observa una tendencia general decreciente entre fecundidad y esperanza de vida —es decir, a mayor número de hijos por mujer, menor longevidad femenina— existen excepciones relevantes. Algunos países con tasas de fecundidad de 1 o 2 hijos mantienen niveles altos de esperanza de vida, lo que indica que no puede afirmarse categóricamente que tener menos hijos siempre incrementa la longevidad. Los datos sugieren que a partir de 2.5 hijos por mujer, la esperanza de vida comienza a descender de manera más marcada.

```
hijos <- 1.5

# Filtrar datos según 2 hijos
datos_filtrados <- Datos_FEGDP %>%
  filter(!is.na(nacimientos_por_mujer),
         !is.na(esperanza_de_vida)
  ) %>%
  filter(abs(nacimientos_por_mujer - hijos) < 0.2)

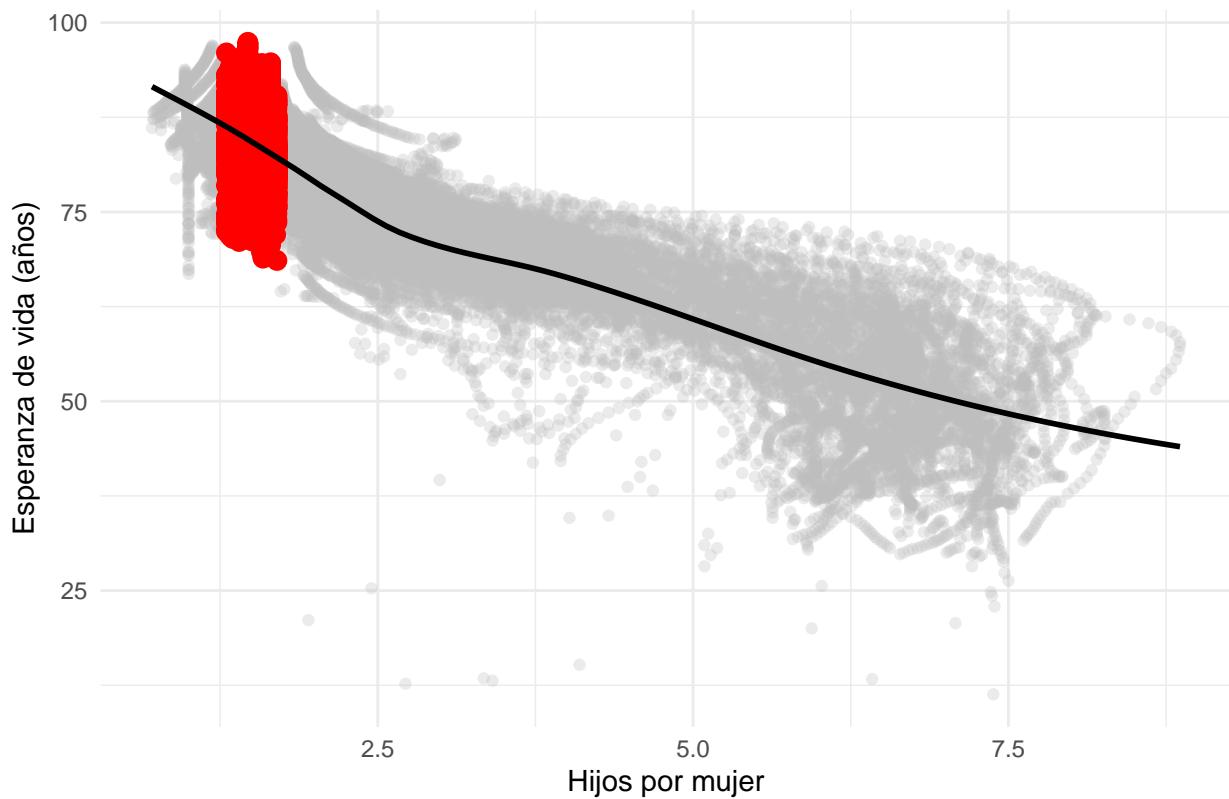
#Gráfico
ggplot(Datos_FEGDP, aes(x = nacimientos_por_mujer, y = esperanza_de_vida)) +
  geom_point(alpha = 0.3, color = "gray") +
  geom_point(data = datos_filtrados, aes(x = nacimientos_por_mujer, y = esperanza_de_vida),
             color = "red", size = 3) +
  geom_smooth(method = "loess", se = FALSE, color = "black") +
  labs(
    title = paste("Esperanza de vida según hijos por mujer",
                  if (!is.na(hijos)) paste0("~", hijos) else ""),
    x = "Hijos por mujer",
    y = "Esperanza de vida (años)"
  ) +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 5893 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 5893 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

Esperanza de vida según hijos por mujer ~1.5



Se realiza una prueba de correlación de Pearson y Spearman para identificar correlaciones positivas como negativas.

```
# Filtrar datos si deseas hacerlo por año específico (opcional)
datos_filtrados <- Datos_FEGDP %>% filter(year == 2000)

# Seleccionar solo las variables numéricas de interés
variables <- datos_filtrados %>%
  select(esperanza_de_vida, producto_por_capita, nacimientos_por_mujer)
# Correlación de Pearson
cor_pearson <- cor(variables, method = "pearson", use = "complete.obs")

# Correlación de Spearman
cor_spearman <- cor(variables, method = "spearman", use = "complete.obs")

# Mostrar resultados
print("Correlación de Pearson:")
```

```
## [1] "Correlación de Pearson:"
```

```
print(round(cor_pearson, 3))
```

	esperanza_de_vida	producto_por_capita
## esperanza_de_vida	1.000	0.60
## producto_por_capita	0.600	1.00

```

## nacimientos_por_mujer           -0.856          -0.47
##                               nacimientos_por_mujer
## esperanza_de_vida                  -0.856
## producto_por_capita                 -0.470
## nacimientos_por_mujer                   1.000

print("Correlación de Spearman:")

## [1] "Correlación de Spearman:"

print(round(cor_spearman, 3))

##             esperanza_de_vida producto_por_capita
## esperanza_de_vida           1.000            0.870
## producto_por_capita         0.870            1.000
## nacimientos_por_mujer      -0.823           -0.742
##                               nacimientos_por_mujer
## esperanza_de_vida           -0.823
## producto_por_capita         -0.742
## nacimientos_por_mujer       1.000

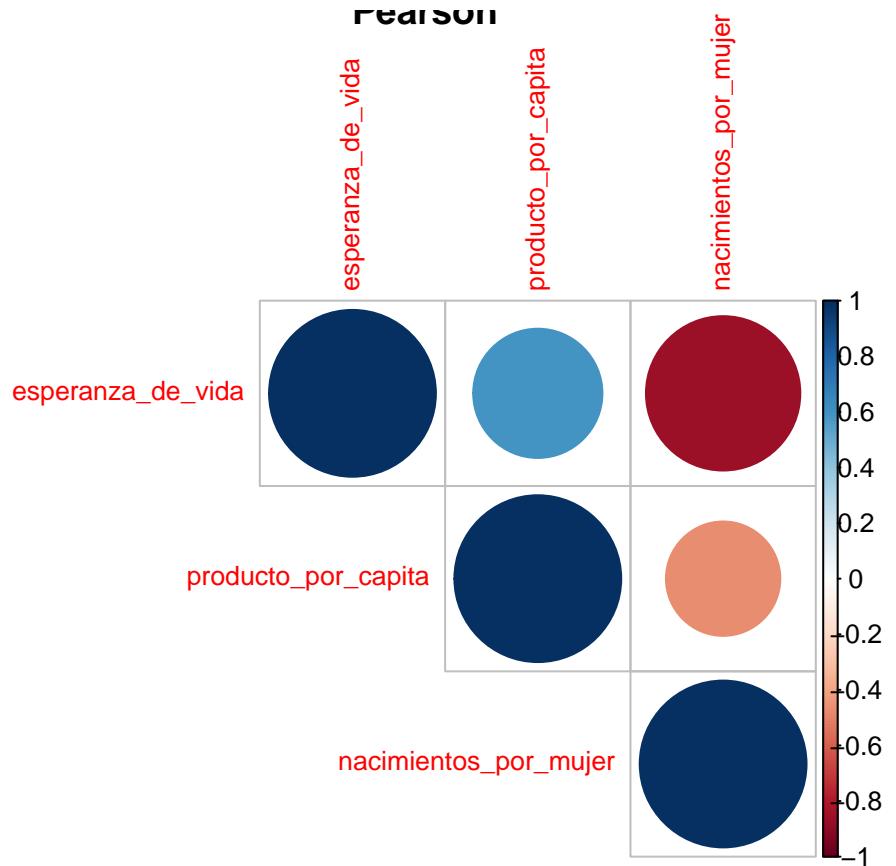
```

Como se puede ver en el gráfico, hay una correlación positiva entre esperanza de vida y PIB, mientras que hay una negativa entre esperanza de vida y fecundidad, lo mismo que se puede ver en los gráficos, a su vez, también se puede ver una correlación negativa entre el PIB per cápita y la fecundidad.

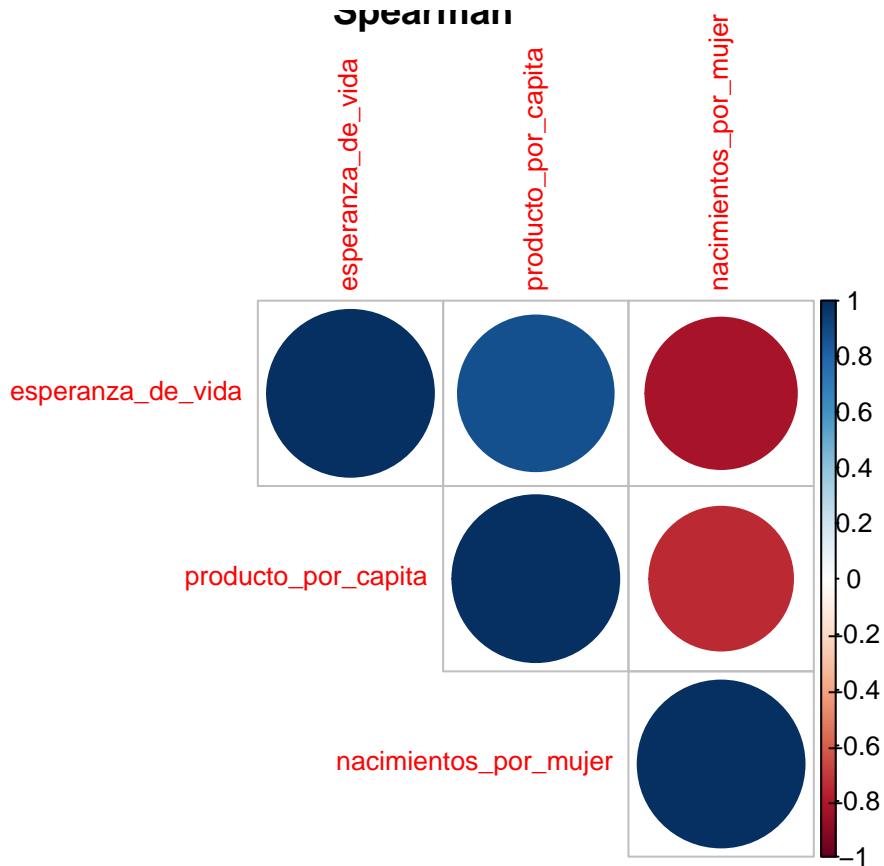
```

# Visualizar matriz de Pearson
corrplot(cor_pearson, method = "circle", type = "upper", tl.cex = 0.8, title = "Pearson")

```



```
# Visualizar matriz de Spearman
corrplot(cor_spearman, method = "circle", type = "upper", tl.cex = 0.8, title = "Spearman")
```



Los resultados respaldan los patrones visuales observados anteriormente mediante gráficos.

Análisis Regional

Europa presenta un comportamiento particular: Se observa una caída en la cantidad de hijos por mujeres y con esto también la esperanza de vida en países con 2 hijos por mujer, aunque no se trata de una reducción sistemática. De hecho, varios países con solo 1 hijo muestran menor esperanza de vida que aquellos con 2, lo cual podría deberse a factores contextuales como el envejecimiento poblacional o las políticas de salud.

```
datos_filtrados <- Datos_FEGDP %>%
  filter(continente == "Europe")

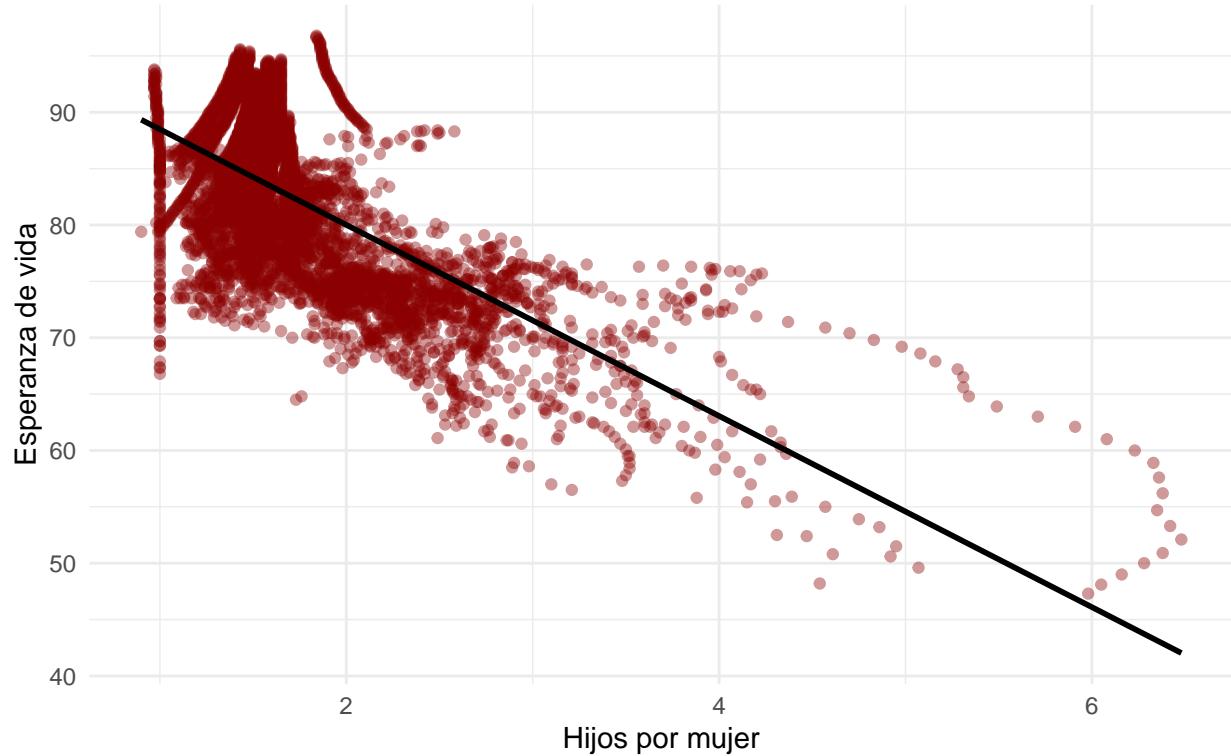
ggplot(datos_filtrados, aes(x = nacimientos_por_mujer, y = esperanza_de_vida)) +
  geom_point(alpha = 0.4, color = "darkred") +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  labs(title = "Relación entre hijos por mujer",
       y = "esperanza de vida de Europa", x = "Hijos por mujer", y = "Esperanza de vida") +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 906 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 906 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

Relación entre hijos por mujer y esperanza de vida de Europa

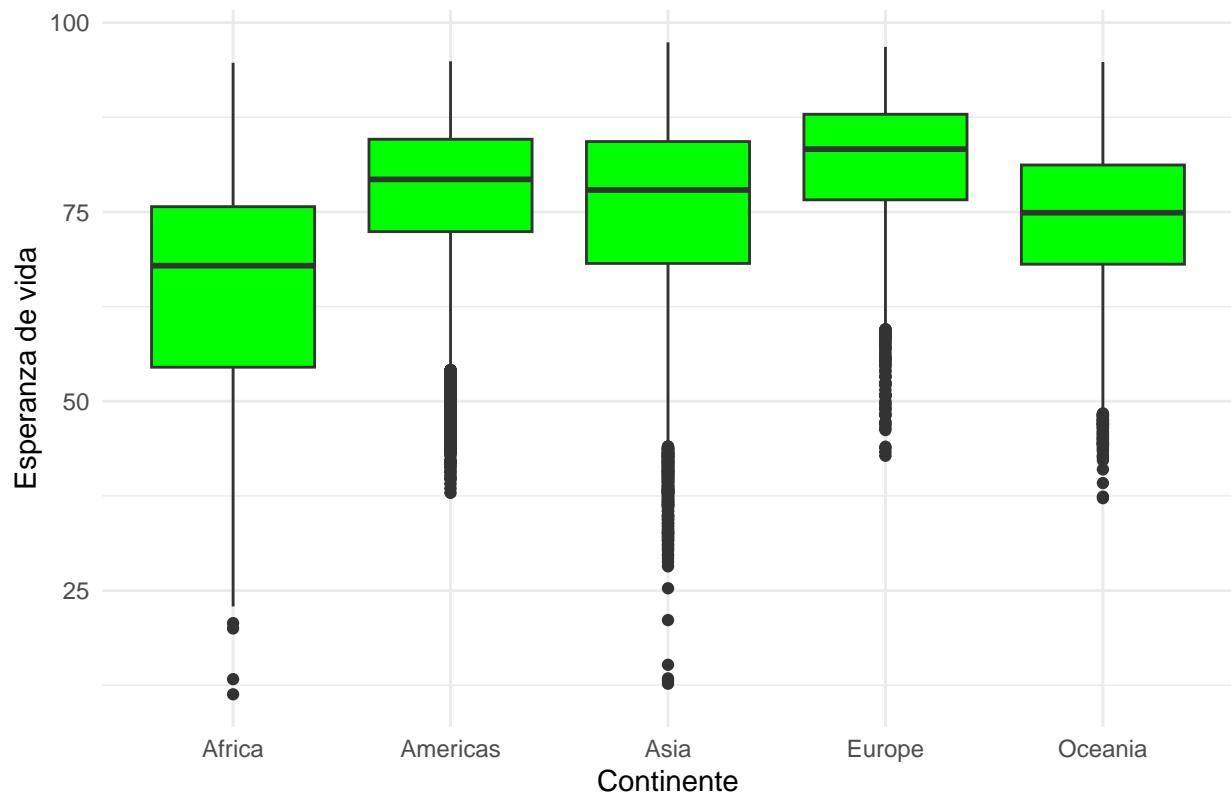


En el caso de África, se registra la esperanza de vida más baja en comparación con otros continentes, manteniéndose la tendencia decreciente a pesar de una alta fecundidad.

```
ggplot(Datos_FEGDP, aes(x = continente, y = esperanza_de_vida)) +  
  geom_boxplot(fill = "green") +  
  labs(title = "Esperanza de vida por continente", y = "Esperanza de vida", x = "Continente") +  
  theme_minimal()
```

```
## Warning: Removed 5 rows containing non-finite outside the scale range  
## ('stat_boxplot()').
```

Esperanza de vida por continente

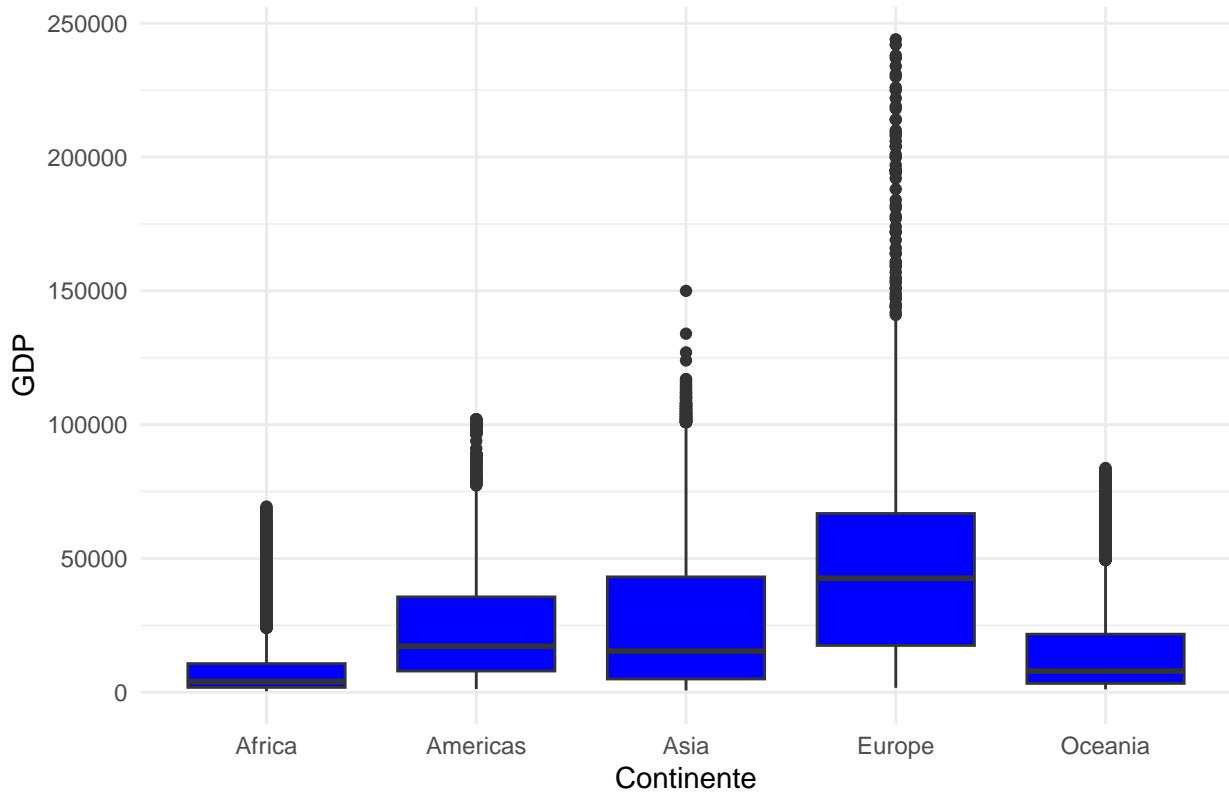


Continuando con la misma región, el bajo nivel del PIB per cápita también resalta como un elemento que podría influir en los resultados, reflejando bajos niveles de desarrollo económico.

```
ggplot(Datos_FEGDP, aes(x = continente, y = producto_por_capita)) +  
  geom_boxplot(fill = "blue") +  
  labs(title = "Producto Interno Bruto por continente", y = "GDP", x = "Continente") +  
  theme_minimal()
```

```
## Warning: Removed 7550 rows containing non-finite outside the scale range  
## ('stat_boxplot()').
```

Producto Interno Bruto por continente



En las Américas, se conserva la tendencia general: los países con mayor fecundidad tienden a registrar niveles más bajos de esperanza de vida. Este patrón refuerza la hipótesis de una correlación negativa entre estas dos variables, aunque no necesariamente causal.

```
datos_filtrados <- Datos_FEGDP %>%
  filter(continente == "Americas")

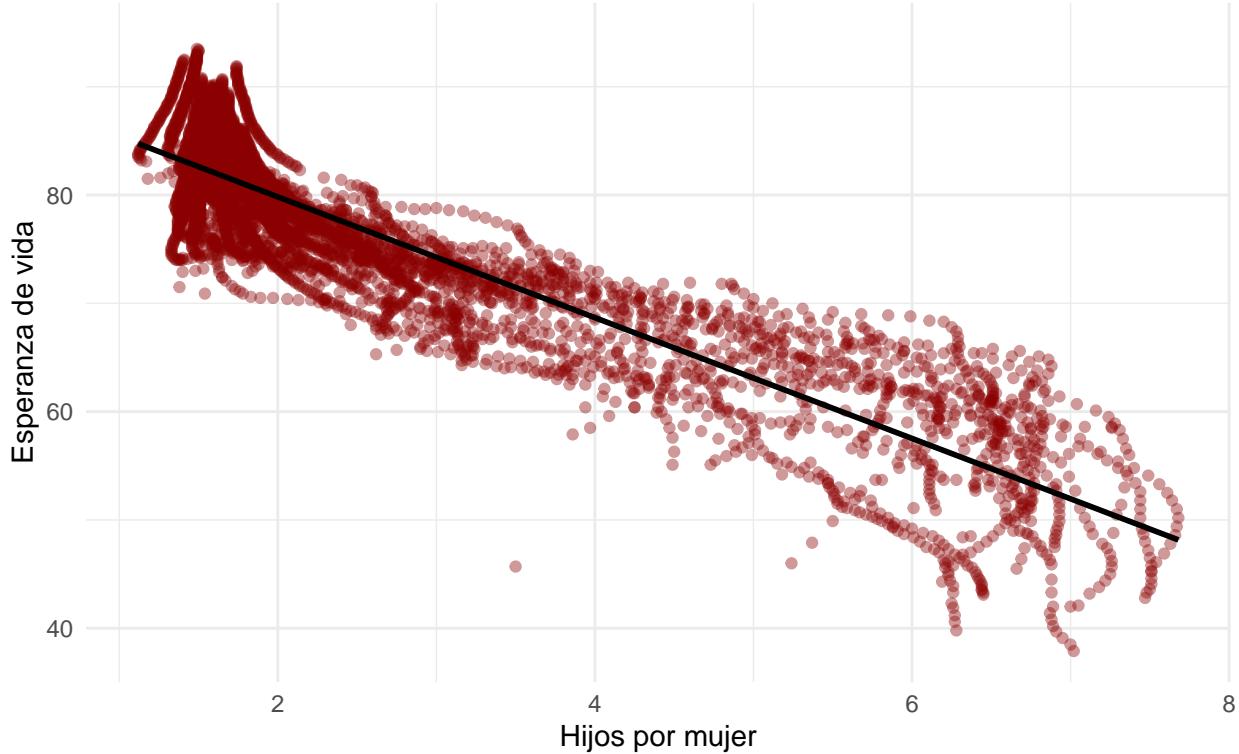
ggplot(datos_filtrados, aes(x = nacimientos_por_mujer, y = esperanza_de_vida)) +
  geom_point(alpha = 0.4, color = "darkred") +
  geom_smooth(method = "lm", se = FALSE, color = "black") +
  labs(title = "Relación entre hijos por mujer",
       y = "esperanza de vida de America", x = "Hijos por mujer", y = "Esperanza de vida") +
  theme_minimal()

## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 2869 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 2869 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

Relación entre hijos por mujer y esperanza de vida de America



En la sección de gráficos de interacción avanzada, se identifican valores altos de esperanza de vida asociados a tasas de fecundidad cercanas a 0.7 hijos por mujer, con una esperanza de vida promedio entre 86 y 89 años.

```

vida_min <- 86
vida_max <- 89

# Filtrar datos según los inputs
datos_filtrados <- Datos_FEGDP %>%
  filter(!is.na(nacimientos_por_muher),
  !is.na(esperanza_de_vida))
) %>%
  filter(esperanza_de_vida >= vida_min & esperanza_de_vida <= vida_max)

#Gráfico
ggplot(Datos_FEGDP, aes(x = esperanza_de_vida, y = nacimientos_por_muher)) +
  geom_point(alpha = 0.3, color = "pink") +
  geom_point(data = datos_filtrados, aes(x = esperanza_de_vida, y = nacimientos_por_muher),
  color = "blue", size = 3) +
  geom_smooth(method = "loess", se = FALSE, color = "black") +
  labs(
    title = paste("Hijos por mujer según esperanza de vida",
    vida_min, "y", vida_max, "años"),
    x = "Esperanza de vida (años)",
    y = "Hijos por mujer")
) +
  theme_minimal()

```

```

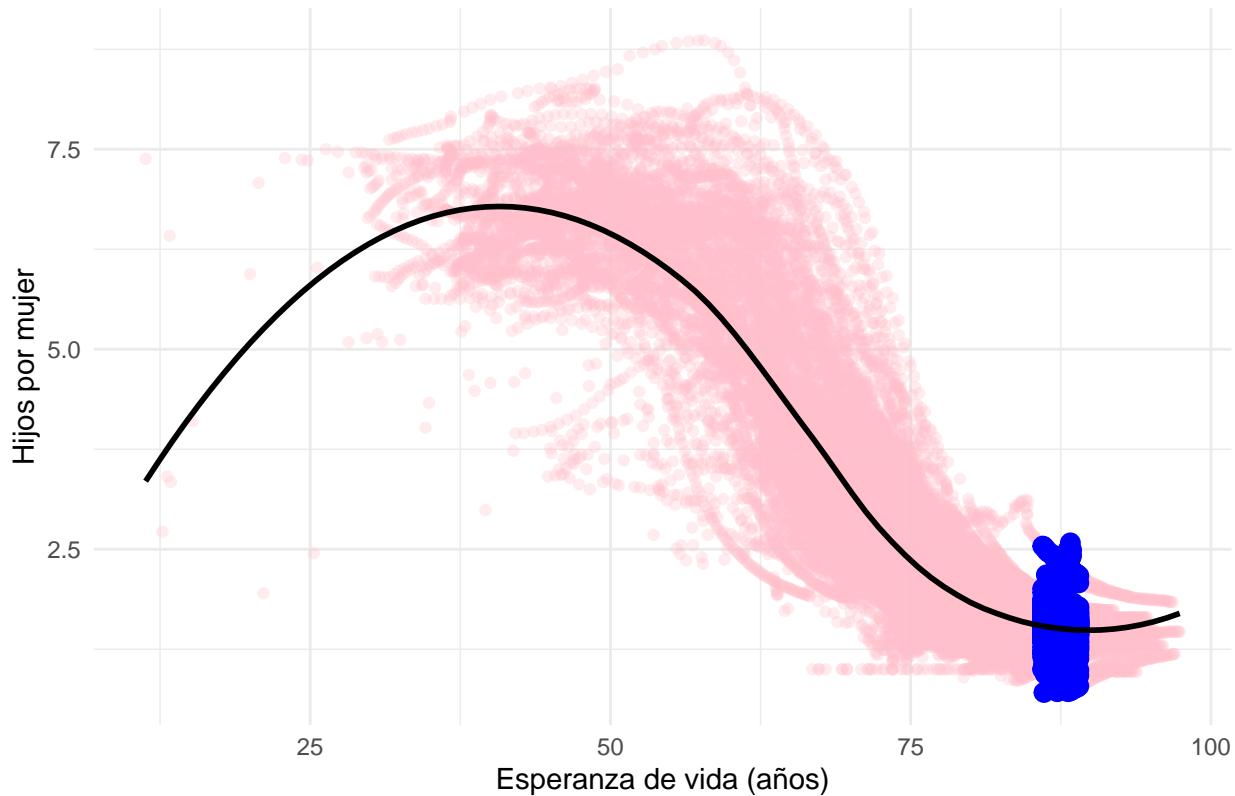
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 5893 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 5893 rows containing missing values or values outside the scale range
## (`geom_point()`).

```

Hijos por mujer según esperanza de vida 86 y 89 años



Por otro lado, los países con una fecundidad de 1 hijo muestran una esperanza de vida aún mayor, entre 93 y 96 años. Esto sugiere que los países con baja fecundidad tienden a presentar mayores niveles de longevidad. Sin embargo, no se observa un patrón lineal estricto: no puede afirmarse que a menor número de hijos por mujer siempre corresponda una mayor esperanza de vida.

```

vida_min1 <- 93
vida_max2 <- 96

# Filtrar datos según los inputs
datos_filtrados <- Datos_FEGDP %>%
  filter(!is.na(nacimientos_por_mujer),
         !is.na(esperanza_de_vida))
) %>%
  filter(esperanza_de_vida >= vida_min1 & esperanza_de_vida <= vida_max2)

#Gráfico
ggplot(Datos_FEGDP, aes(x = esperanza_de_vida, y = nacimientos_por_mujer)) +

```

```

geom_point(alpha = 0.3, color = "pink") +
geom_point(data = datos_filtrados, aes(x = esperanza_de_vida, y = nacimientos_por_mujer),
           color = "blue", size = 3) +
geom_smooth(method = "loess", se = FALSE, color = "black") +
labs(
  title = paste("Hijos por mujer según esperanza de vida",
                vida_min1, "y", vida_max2, "años"),
  x = "Esperanza de vida (años)",
  y = "Hijos por mujer"
) +
theme_minimal()

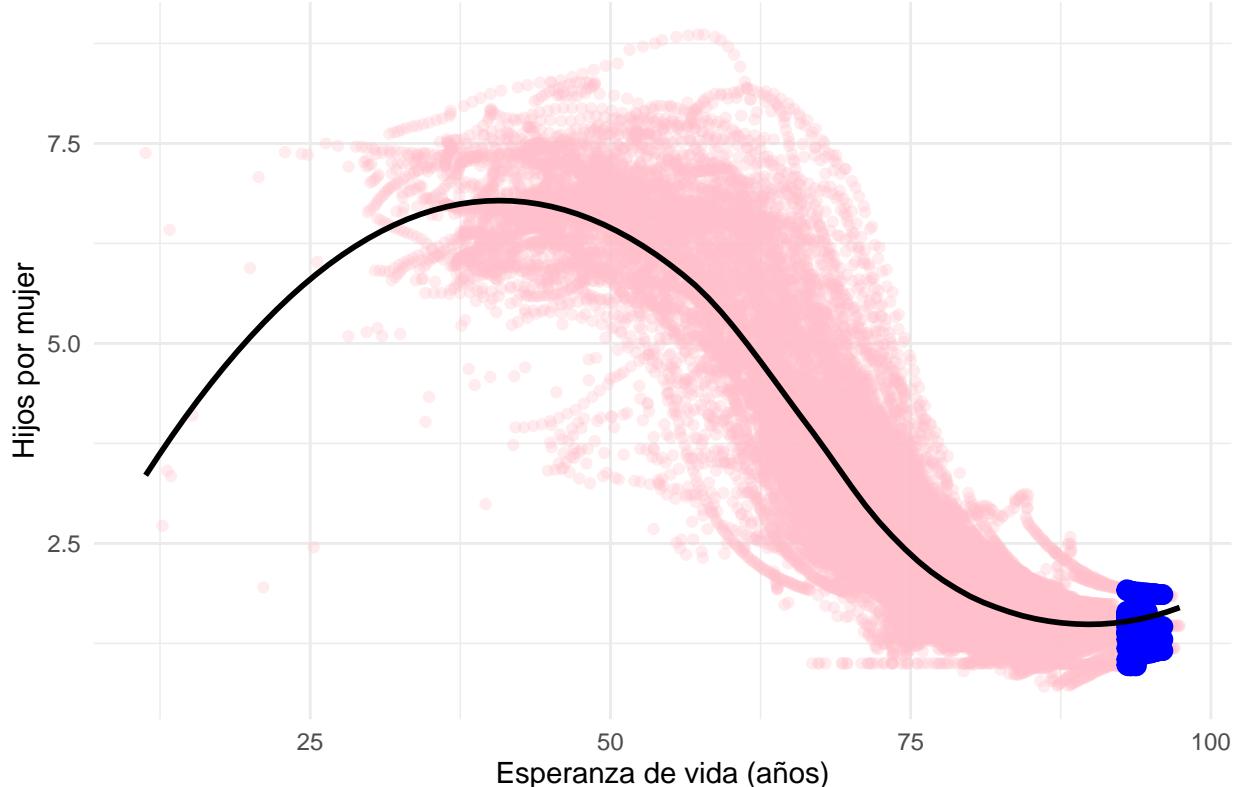
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 5893 rows containing non-finite outside the scale range
## ('stat_smooth()').

## Warning: Removed 5893 rows containing missing values or values outside the scale range
## ('geom_point()').

```

Hijos por mujer según esperanza de vida 93 y 96 años



Aunque no se puede concluir que tener hijos necesariamente alarga o acorta la vida, los gráficos permiten identificar tendencias consistentes, especialmente al analizar los datos por regiones o niveles de desarrollo económico.

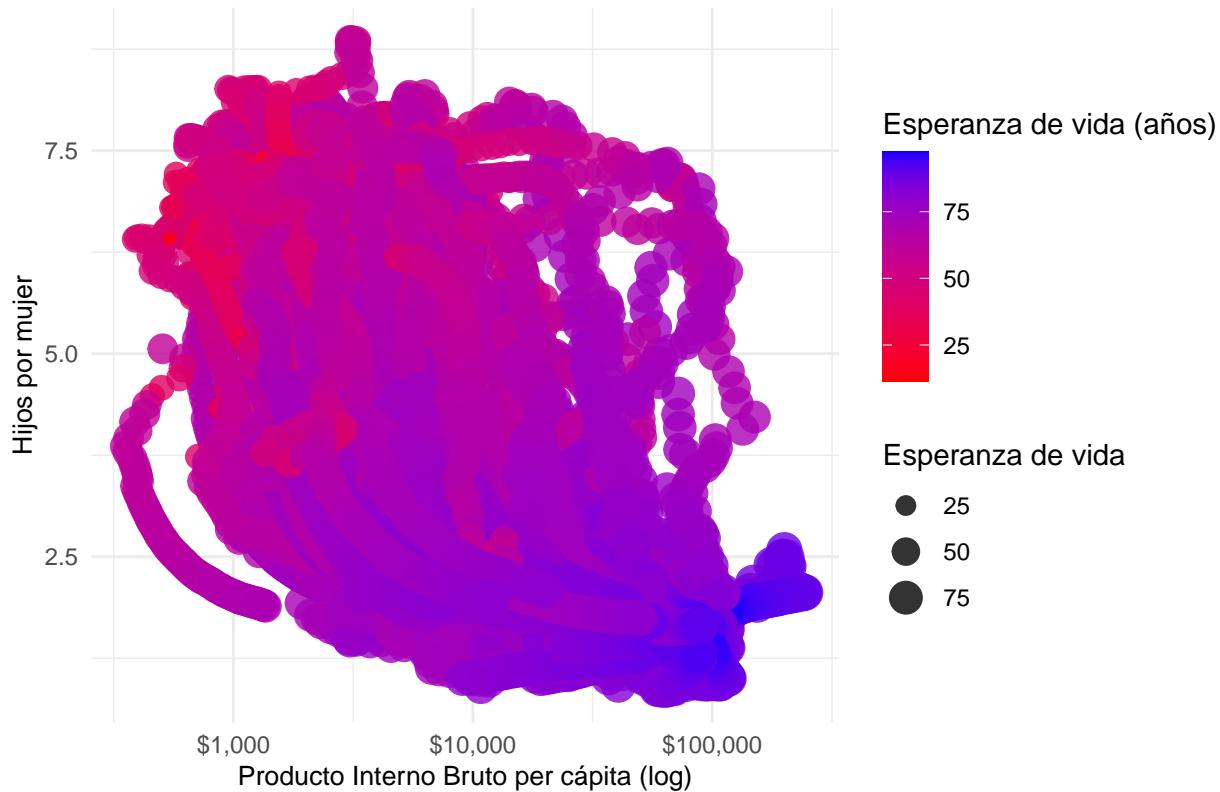
En cuanto al PIB per cápita, el gráfico de interacción avanzada muestra cómo esta variable se asocia tanto con fecundidad como con esperanza de vida. En el gráfico “Relación entre GDP, fecundidad y esperanza de

vida”, se observa que los países con PIB superiores a 100,000 dólares se ubican principalmente en la franja de 2 a 2.5 hijos por mujer, y presentan una esperanza de vida alta (representada en color azul). Esto evidencia que, si bien la baja fecundidad puede asociarse con mayor longevidad, el nivel de desarrollo económico es un factor modulador clave.

```
datos_filtrados <- Datos_FEGDP %>%
  filter(
    !is.na(producto_por_capita),
    !is.na(nacimientos_por_mujer),
    !is.na(esperanza_de_vida)
  )

# Gráfico
ggplot(datos_filtrados, aes(x = producto_por_capita,
                             y = nacimientos_por_mujer,
                             size = esperanza_de_vida,
                             color = esperanza_de_vida)) +
  geom_point(alpha = 0.8) +
  scale_x_log10(labels = dollar_format(prefix = "$")) +
  scale_color_gradient(low = "red", high = "blue") +
  labs(title = "Relación entre GDP, fecundidad y esperanza de vida",
       x = "Producto Interno Bruto per cápita (log)",
       y = "Hijos por mujer",
       color = "Esperanza de vida (años)",
       size = "Esperanza de vida") +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 15, face = "bold"),
    axis.title = element_text(size = 10)
  )
```

Relación entre GDP, fecundidad y esperanza de vida



Al comparar países según sus niveles de ingreso, se observa que en aquellos con mayor PIB per cápita, la relación entre baja fecundidad y alta esperanza de vida es más clara y estable. En contraste, los países con ingresos bajos presentan relaciones más dispersas, posiblemente influenciadas por factores externos como conflictos, pobreza estructural o debilidad institucional.

En cuanto al eje temporal, en el gráfico “Relación entre GDP, fecundidad y esperanza de vida”, se aprecia que en 1970 la esperanza de vida promedio mundial rondaba los 80 años.

```
datos_filtrados <- Datos_FEGDP %>%
  filter( year == 1970,
  !is.na(producto_por_capita),
  !is.na(nacimientos_por_muher),
  !is.na(esperanza_de_vida)
)

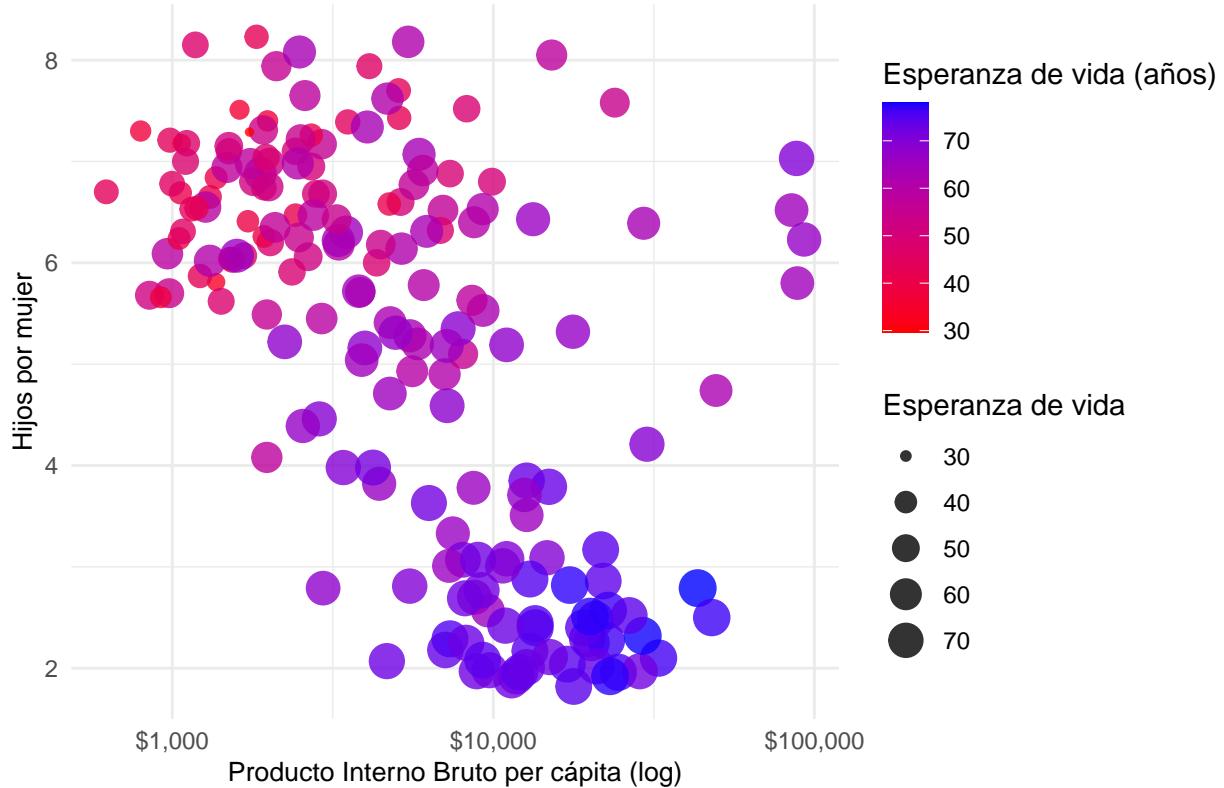
# Gráfico
ggplot(datos_filtrados, aes(x = producto_por_capita,
                               y = nacimientos_por_muher,
                               size = esperanza_de_vida,
                               color = esperanza_de_vida)) +
  geom_point(alpha = 0.8) +
  scale_x_log10(labels = dollar_format(prefix = "$")) +
  scale_color_gradient(low = "red", high = "blue") +
  labs(title = "Relación entre GDP, fecundidad y esperanza de vida en 1970",
       x = "Producto Interno Bruto per cápita (log)",
       y = "Hijos por mujer",
       color = "Esperanza de vida (años)",
```

```

size = "Esperanza de vida") +
theme_minimal() +
theme(
  plot.title = element_text(size = 15, face = "bold"),
  axis.title = element_text(size = 10)
)

```

Relación entre GDP, fecundidad y esperanza de vida en 1970



Para el año 2000, esta cifra aumentó casi en una década, acercándose a los 90 años. Este incremento evidencia mejoras globales en salud, condiciones de vida y acceso a servicios básicos.

```

datos_filtrados <- Datos_FEGDP %>%
  filter( year == 2000,
  !is.na(producto_por_capita),
  !is.na(nacimientos_por_mujer),
  !is.na(esperanza_de_vida)
)

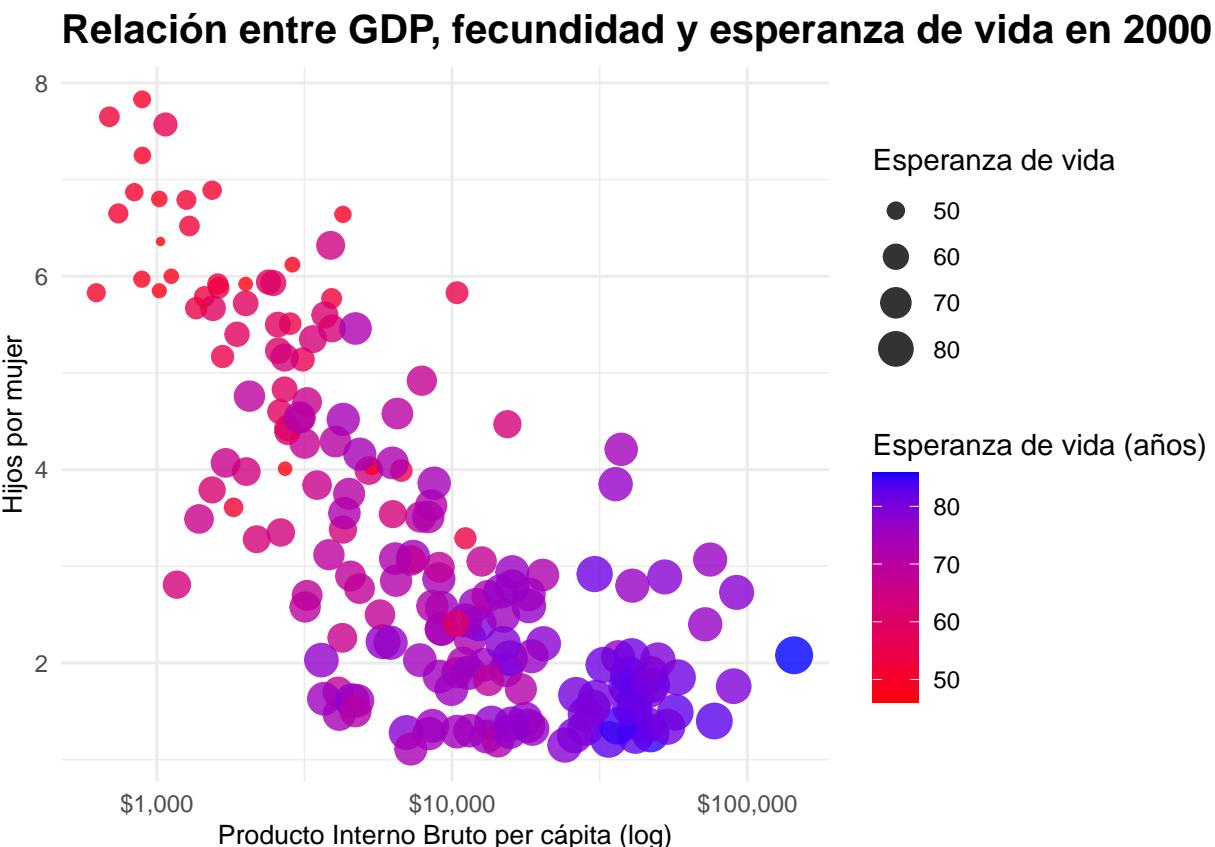
# Gráfico
ggplot(datos_filtrados, aes(x = producto_por_capita,
                               y = nacimientos_por_mujer,
                               size = esperanza_de_vida,
                               color = esperanza_de_vida)) +
  geom_point(alpha = 0.8) +
  scale_x_log10(labels = dollar_format(prefix = "$")) +
  scale_color_gradient(low = "red", high = "blue") +
  labs(title = "Relación entre GDP, fecundidad y esperanza de vida en 2000",
       subtitle = "Año 2000")

```

```

x = "Producto Interno Bruto per cápita (log)",
y = "Hijos por mujer",
color = "Esperanza de vida (años)",
size = "Esperanza de vida") +
theme_minimal() +
theme(
  plot.title = element_text(size = 15, face = "bold"),
  axis.title = element_text(size = 10)
)

```



Aunque la fecundidad no constituye un determinante único en la decisión de tener hijos ni en la esperanza de vida, sí representa un componente clave en el análisis con perspectiva de género. Las decisiones reproductivas están profundamente influenciadas por factores sociales, económicos y culturales, que varían según el contexto y el nivel de desarrollo.

En este estudio, la regresión múltiple revela que el PIB per cápita y la fecundidad explican conjuntamente más del 84% de la variabilidad en la esperanza de vida femenina. Este resultado evidencia una correlación multivariable sólida entre desarrollo económico, salud y dinámica demográfica, y refuerza la necesidad de abordar los indicadores poblacionales desde una mirada integrada y contextualizada.

Además, estos hallazgos se alinean con estudios previos que sugieren un vínculo entre decisiones reproductivas y longevidad femenina. Si bien no se puede establecer una relación causal directa, los datos apuntan a que niveles moderados de fecundidad —como tener un solo hijo— podrían asociarse con una mayor esperanza de vida, especialmente en contextos con mejores condiciones socioeconómicas.

```

# Modelo de regresión múltiple
modelo <- lm(variables, data = Datos_FEGDP)

# Resumen del modelo
summary(modelo)

## 
## Call:
## lm(formula = variables, data = Datos_FEGDP)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -58.855  -2.576   0.313   3.099  19.957
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)             8.533e+01  8.738e-02  976.5 <2e-16 ***
## producto_por_capita    1.442e-04  1.353e-06   106.5 <2e-16 ***
## nacimientos_por_mujer -5.139e+00  1.945e-02  -264.2 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.146 on 28082 degrees of freedom
## (7551 observations deleted due to missingness)
## Multiple R-squared:  0.8469, Adjusted R-squared:  0.8469 
## F-statistic: 7.766e+04 on 2 and 28082 DF,  p-value: < 2.2e-16

```

Construcción del Dashboard

La creación de un dashboard permite visualizar información que, por otros medios, no sería posible explorar con la misma claridad. El dashboard titulado “Relación entre fecundidad, esperanza de vida de las mujeres y el Producto Interno Bruto (PIB)” fue desarrollado con el propósito de cumplir los objetivos planteados previamente. Para su diseño, se tomó como referencia el boceto propuesto en el Reto 1.

Aunque no se logró replicar visualmente el diseño del boceto original debido a las limitaciones propias en la herramienta en R, sí fue posible implementar una funcionalidad que permite consultar la información de forma efectiva y dinámica.

En la primera sección del dashboard, el objetivo era ofrecer una visión general de las tres variables principales. En lugar de utilizar gráficos, se optó por presentar los datos mediante tablas dinámicas que permiten al usuario consultar la información según el año seleccionado. Esta decisión respondió al interés de mantener la lógica de visualización del boceto original, que incluye botones para retroceder en intervalos de 10 años. Sin embargo, al implementar un input en R para la selección del año, se consideró que no era necesario replicar dichos botones, ya que el usuario puede seleccionar libremente el año deseado desde un menú desplegable.

Asimismo, para mejorar la visualización, se incorporó el nombre del país, el año seleccionado, la variable correspondiente y la bandera del país, aportando un componente visual atractivo y fácil de interpretar.

La segunda sección, dedicada a los gráficos comparativos de variables, presenta tres gráficos de dispersión con distintos colores, lo que facilita su diferenciación. Además, se añadió la opción de seleccionar un país o continente, lo que permite analizar cómo varía la correlación entre las tres variables seleccionadas según el contexto geográfico.

En la tercera sección, correspondiente a los gráficos de interacción avanzada, se representa la relación entre las tres variables mediante un gráfico de dispersión: el eje x representa el PIB (GDP), el eje y la tasa

de fecundidad (hijos por mujer), mientras que el tamaño de los puntos indica la esperanza de vida. Esta visualización permite captar de forma integrada la interacción entre las variables clave del estudio.

Al lado derecho de la tercera sección, por curiosidad y siguiendo el diseño del boceto original, se incluyeron dos gráficos de dispersión adicionales que permiten explorar la relación inversa entre las variables. El primero permite al usuario seleccionar una cantidad de hijos deseada y visualizar la esperanza de vida asociada a ese nivel de fecundidad. El segundo gráfico permite seleccionar una edad específica (esperanza de vida) para observar cuántos hijos, en promedio, tienen las mujeres que alcanzan esa edad. Estas visualizaciones brindan un enfoque interactivo y didáctico para comprender los patrones entre fecundidad y longevidad femenina.

Para finalizar, se incluye una tabla dinámica que permite a las personas usuarias explorar libremente los datos, realizar diferentes combinaciones y responder a nuevas inquietudes o curiosidades que puedan surgir durante la interacción.

Bibliografía

1. Bobbi S. Low, Ashley Hazel, Nicholas Parker and Kathleen B. Welch (2008) Influences on Women's Reproductive Lives : Unexpected Ecological, Cross-Cultural Research 42: 201. <http://ccr.sagepub.com/content/42/3/201>
2. Gapminder. (s.f.). Gapminder data. Gapminder Foundation. <https://www.gapminder.org/data/>
3. Hampton, J. (2021). Does very high human development lead to fertility rebound? An analysis of the Human Life Indicator. Demographic Research, 44, 121–156. <https://doi.org/10.4054/DemRes.2021.44.5>
4. OpenAI. (2025). ChatGPT (versión GPT-4o) [modelo de lenguaje]. <https://chat.openai.com/>
5. Rees, P. (2001). Children and long life: A tale of two hypotheses. People and Place, 9(2), 14–17. <https://doi.org/10.3316/ielapa.200110709>
6. Vaupel, J. W., Carey, J. R., & Christensen, K. (2001). Longevity and aging: Controversial issues in demography. International Journal of Epidemiology, 30(4), 628–631. <https://doi.org/10.1093/ije/30.4.628>
7. Zhang, C., Cao, J., Yang, Q., & Wu, Y. (2023). Association between number of children and female longevity: Evidence from the Chinese Longitudinal ealthy Longevity Survey. Authorea. <https://doi.org/10.22541/au.169765027.7634832>
8. Zipple, M. N., Reeve, H. K., & Peniston, O. J. (2024). Maternal care leads to the evolution of long, slow lives. Proceedings of the National Academy of Sciences, 121(25), e2403491121. <https://doi.org/10.1073/pnas.2403491121>