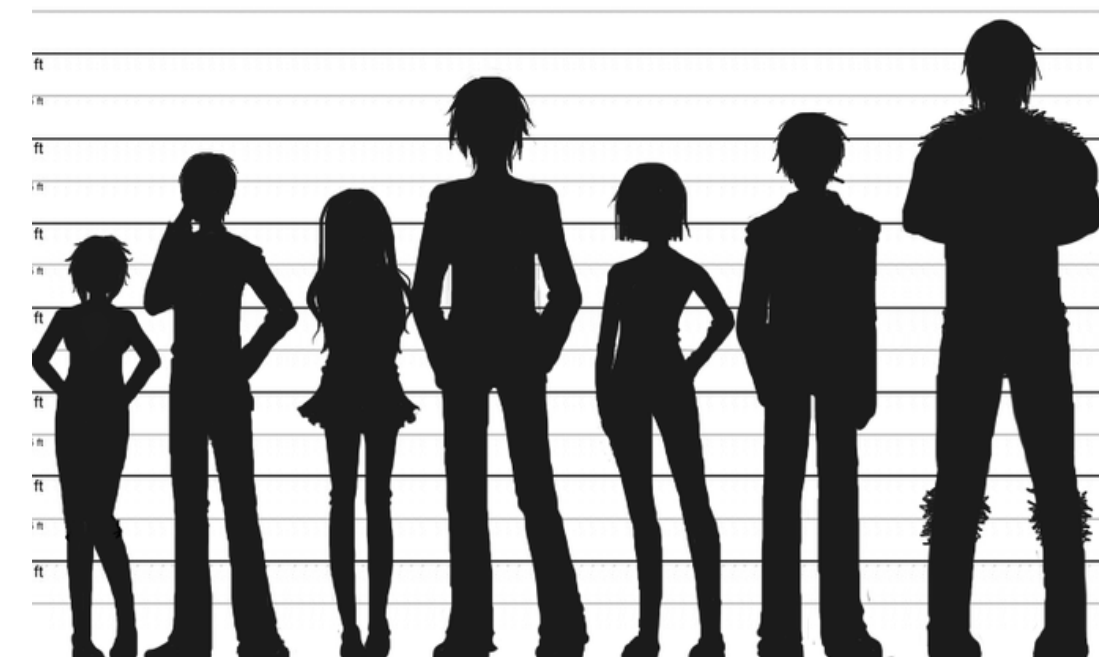


10 DE JULIO 2023

"ANÁLISIS DE ALTURA Y FACTORES ASOCIADOS"

Trabajo Final - Introducción a Ciencia de Datos

Julián Rolando - Grupo 1

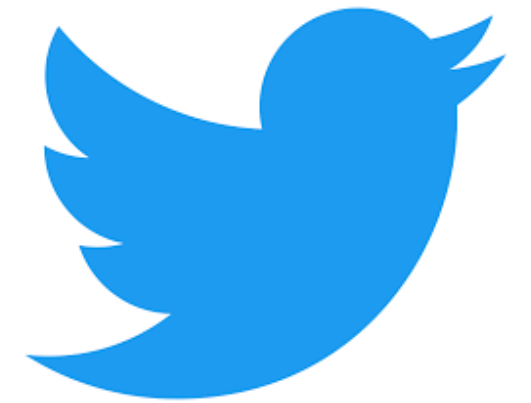


CONTEXTO


Navegando un día por Twitter, me encuentro con un usuario que subió un mapa sobre las alturas de las personas en distintos países de América y me surgió la duda de si podía analizarlo.

PREGUNTA

¿Qué factores influyen en la altura de las personas?
¿Cómo cambia a lo largo de los años?



POSIBLES FACTORES

- PBI per cápita.
- Alimentación (consumo proteína animal).
- Tasa de mortalidad de niños menores a 5 años.
- Expectativa de vida
- **Genética** 

La estatura humana es un rasgo parcialmente hereditario. Sin embargo, queremos analizar si existen factores no genéticos que influyan en la distribución de la estatura entre la población. Por ello, tenemos la hipótesis de que las variaciones de estatura en el mundo indican no sólo diferencias genéticas, sino también diferencias generales en el nivel de vida.

DATASETS

Fuentes: ourworldindata.org, databank.worldbank.org

ALTURA
País
Sexo (H o M)
Año (1914 - 2019)
Grupo de edad (5 a 19 años)
Altura promedio (CM)

PROTEÍNAS
País
Código de país
Año (1960 - 2021)
Consumo diario de calorías de proteína animal (kCal)

CONTINENTE
País
Código de país
Continente

EXPECTATIVA DE VIDA
País
Código de país
Año (1950 - 2021)
Expectativa de Vida

PBI PER CAPITA
País
Código de país
Año (1960 - 2020)
PPC (ajustado a inflación sobre US\$ 2015)

TASA MORTALIDAD NIÑOS
País
Código de país
Año (1960 - 2020)
Tasa de mortalidad

LIMPIEZA Y UNIÓN

- 1) 'Alturas' y su código de país → renombrar países como en 'Continente'.
- 2) Meterle código como columna a 'Alturas'.
- 3) Unir a 'Alturas' con todos los datasets por CÓDIGO y AÑO con left join, eliminando previamente PAÍS en cada uno de los otros.
- 4) Mantener los NA, en todo caso se filtran para graficar.
- 5) Hombres de 19 años.

- 9 Variables
- 12095 filas

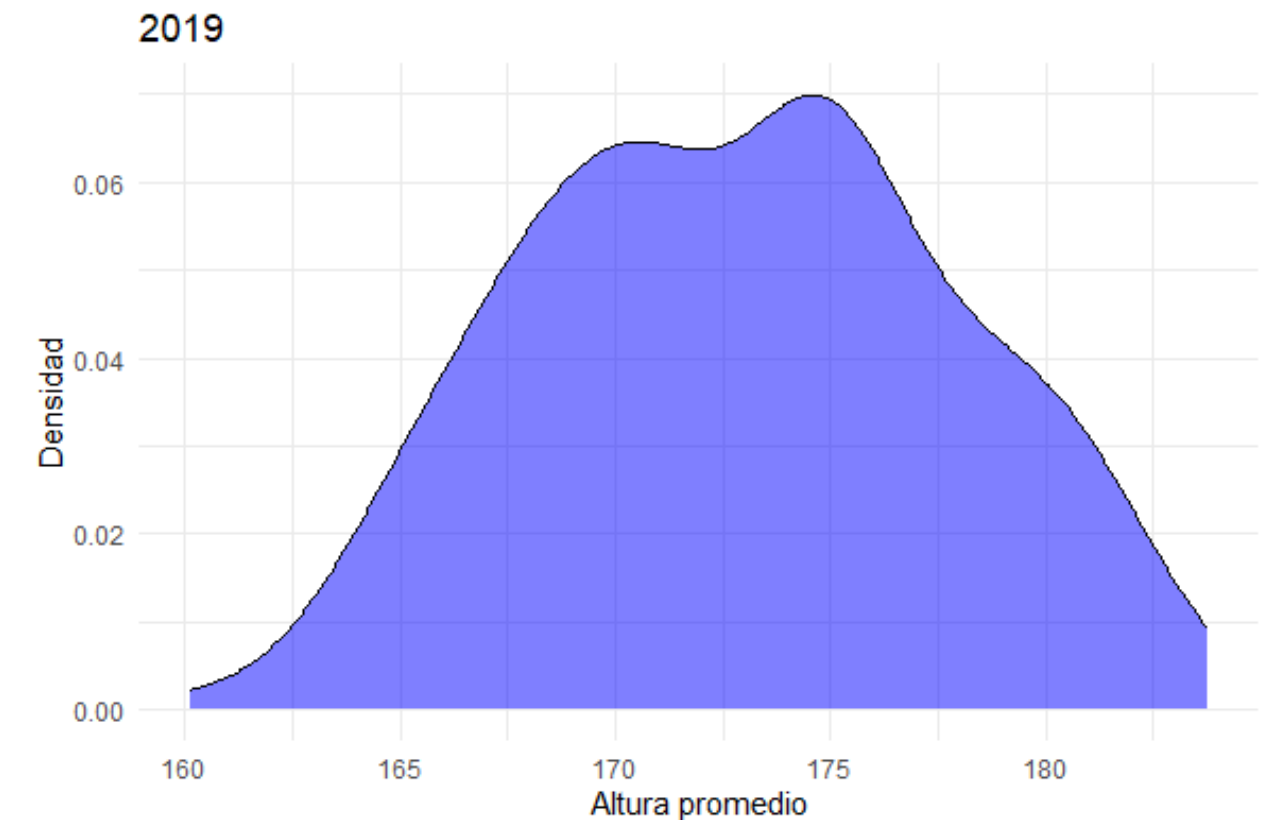
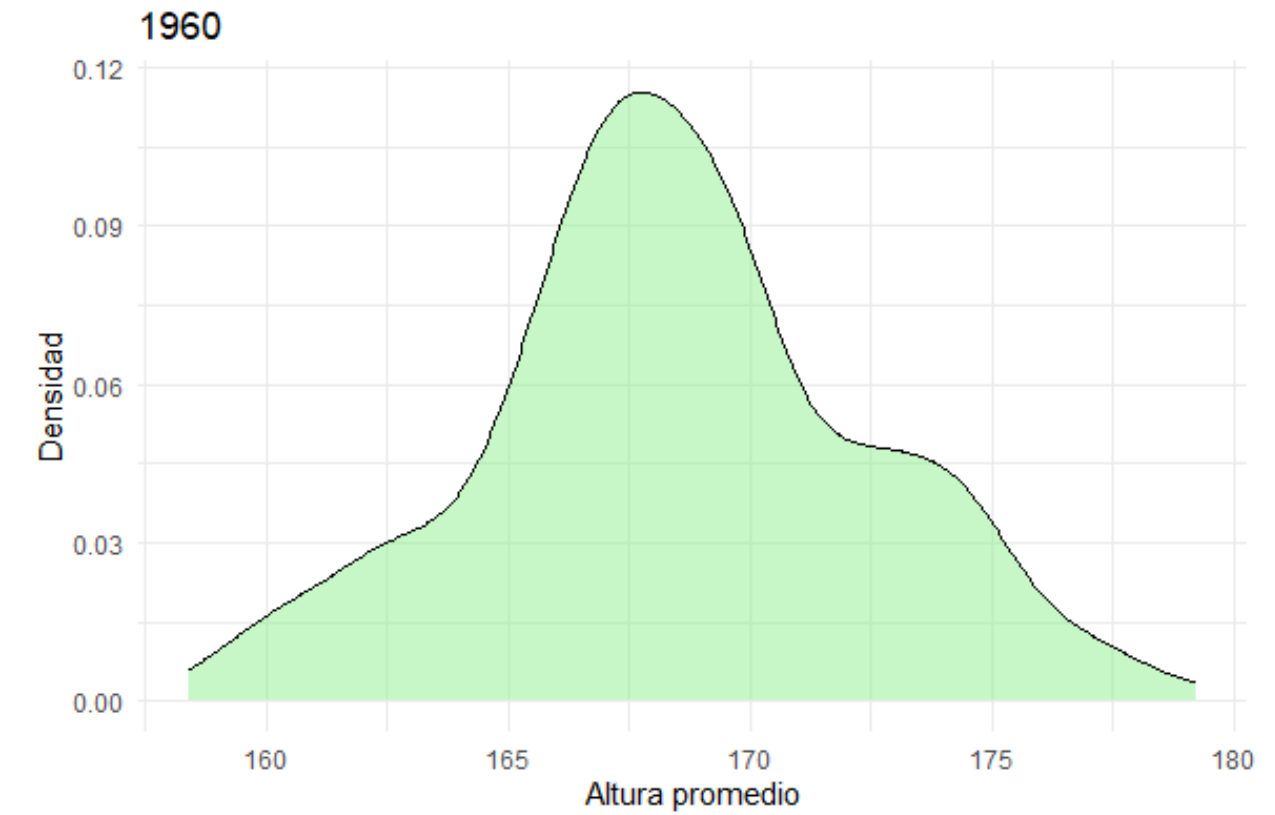
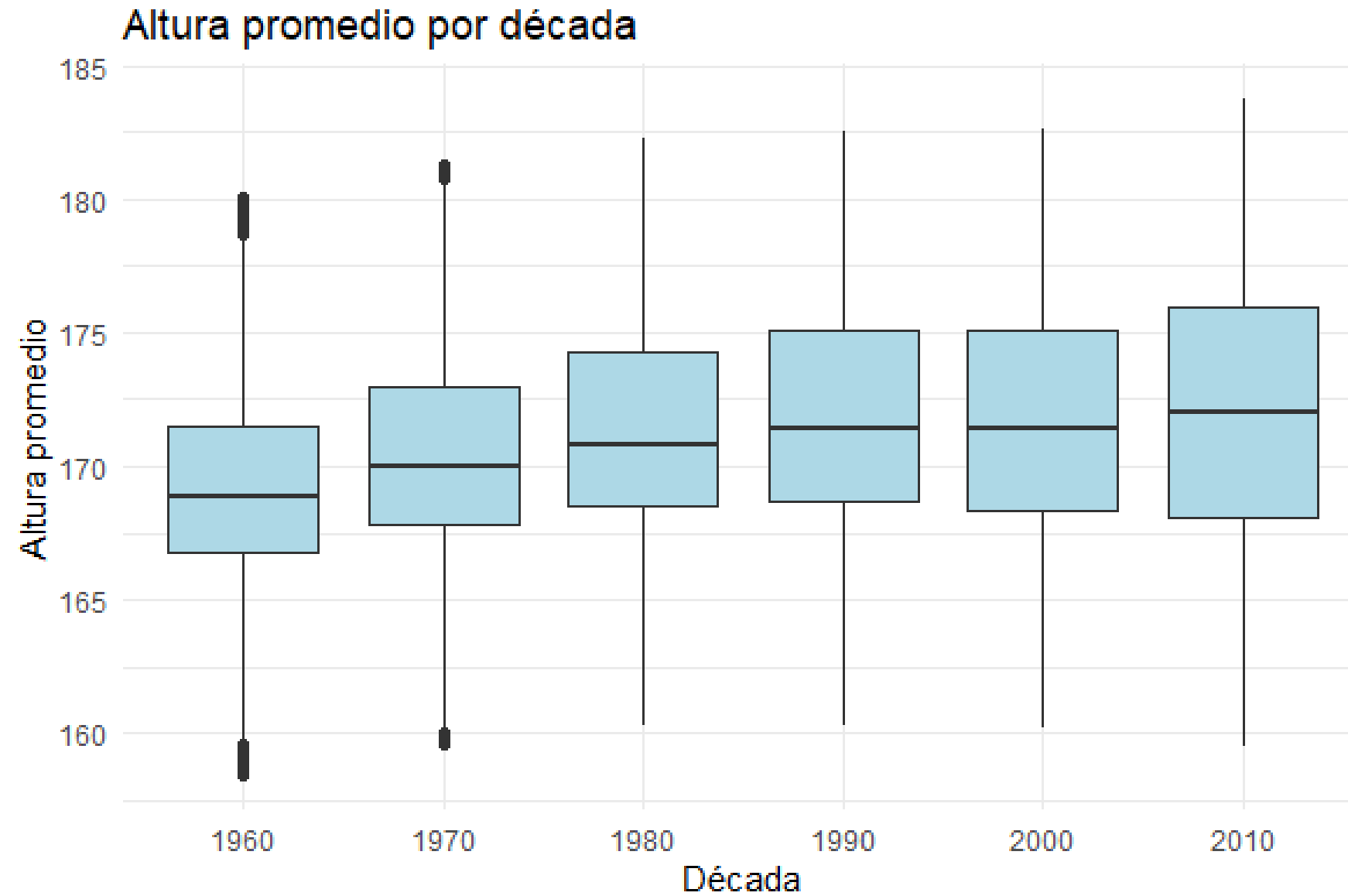
Variables categóricas

País
Continente
Código de país

Variables numéricas

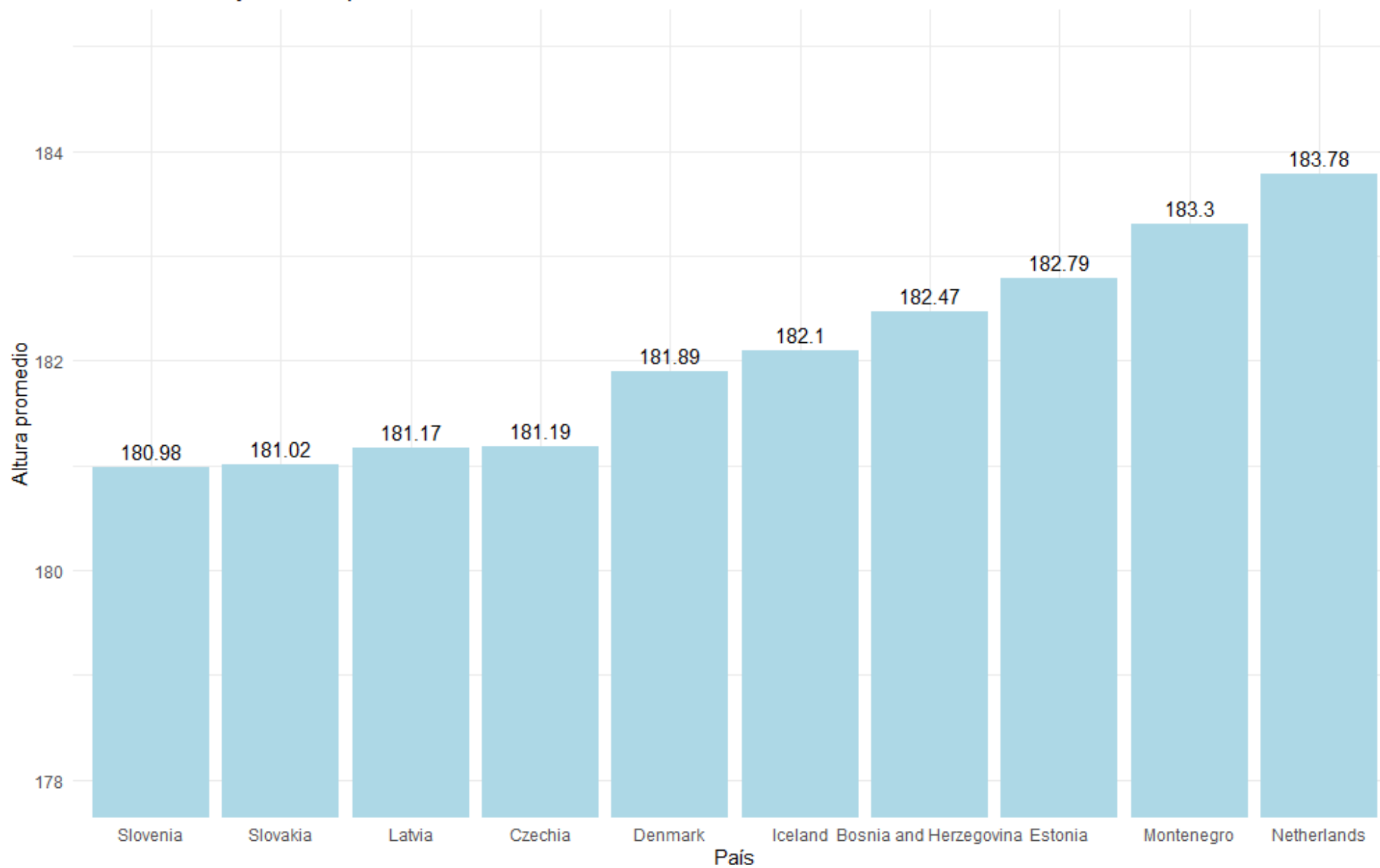
Año (1960-2019)
Altura Promedio (CM)
Calorías de proteína (10kcal - 400kcal)
Expectativa de Vida (42 años - 84 años)
PBI per cápita (144USD - 90000 USD)
Tasa de mortalidad de niños (0,42 - 43)

ANÁLISIS EXPLORATORIO

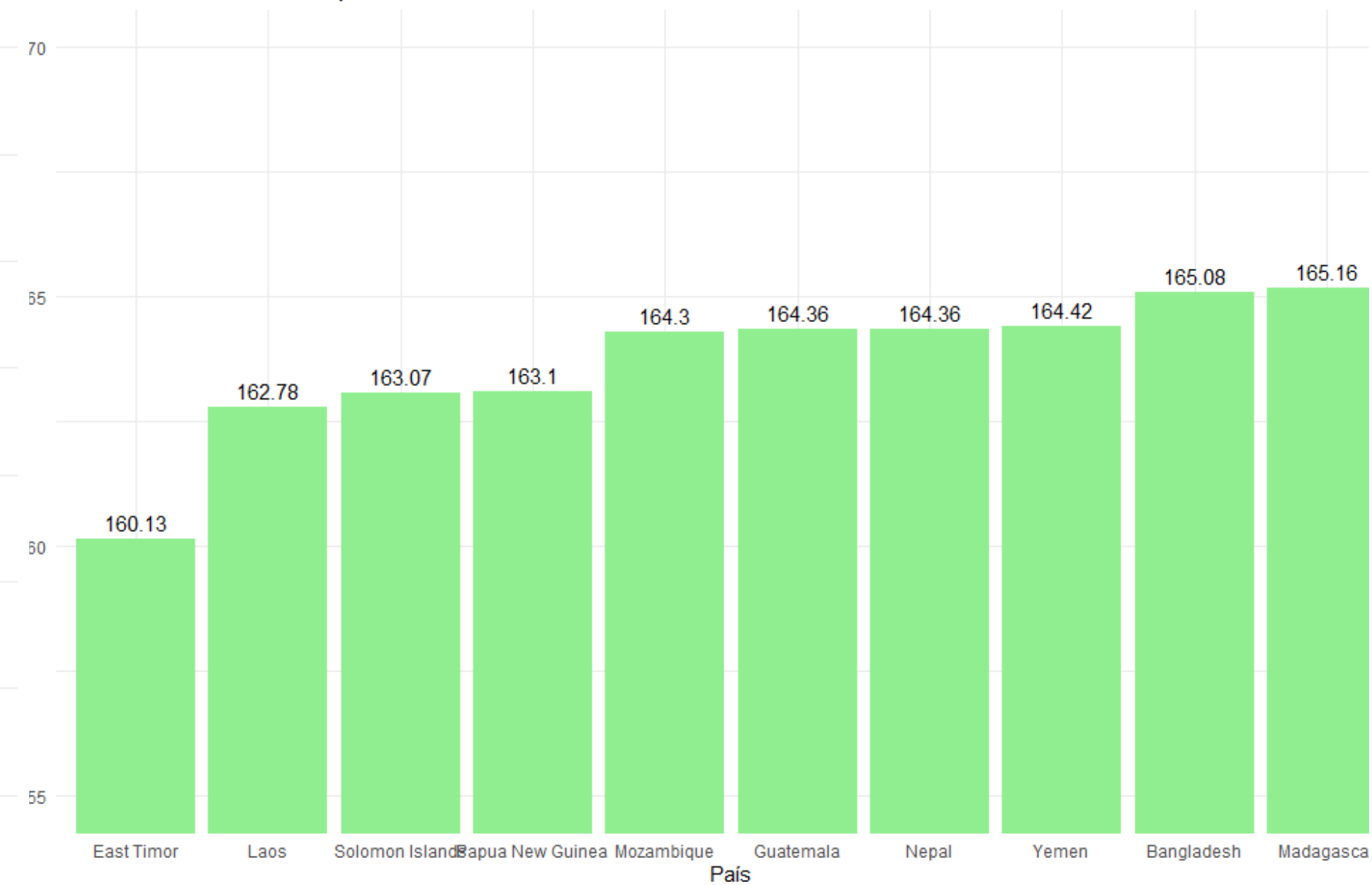


ANÁLISIS EXPLORATORIO

Países con mayor altura promedio en 2019

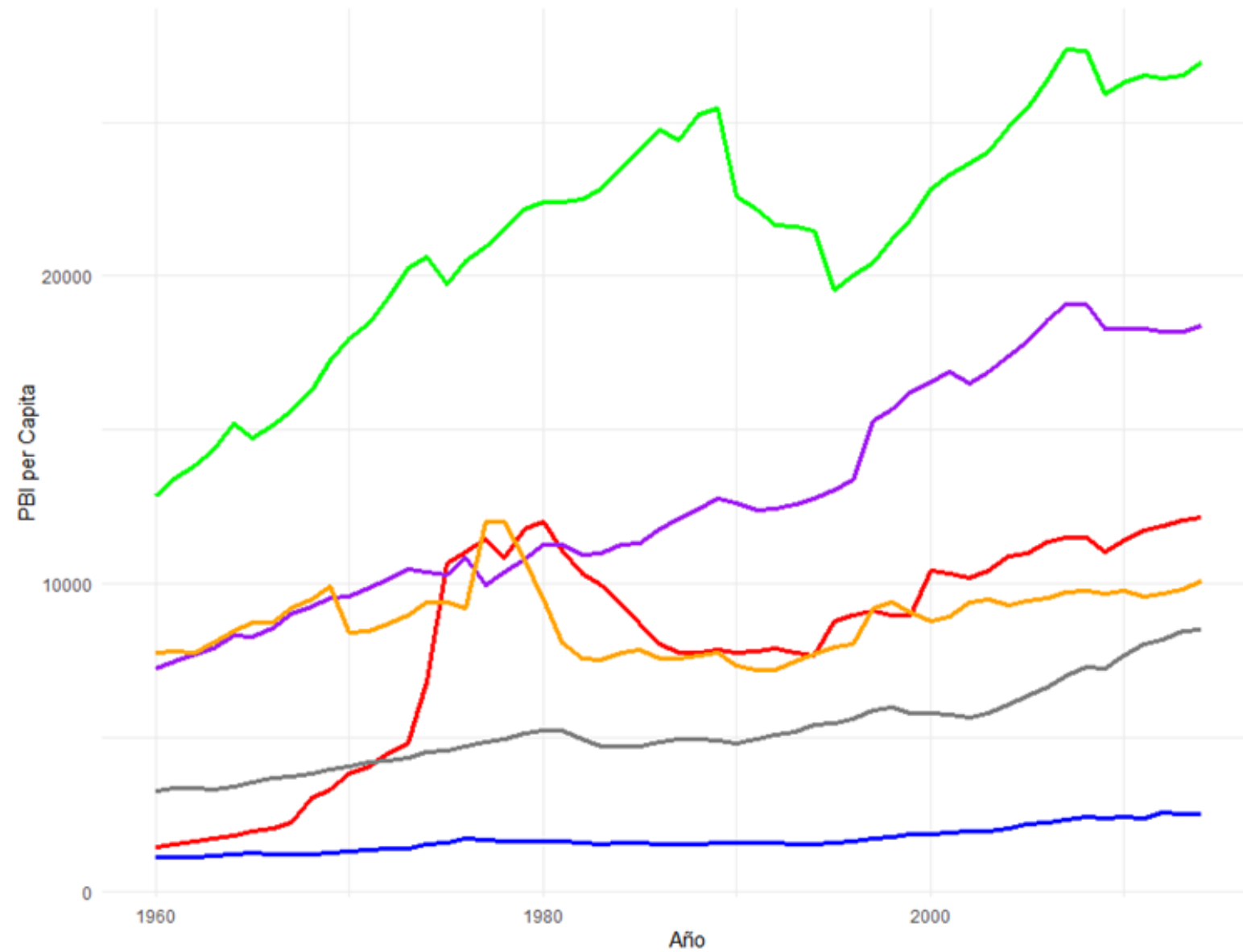


Países con menor altura promedio en 2019

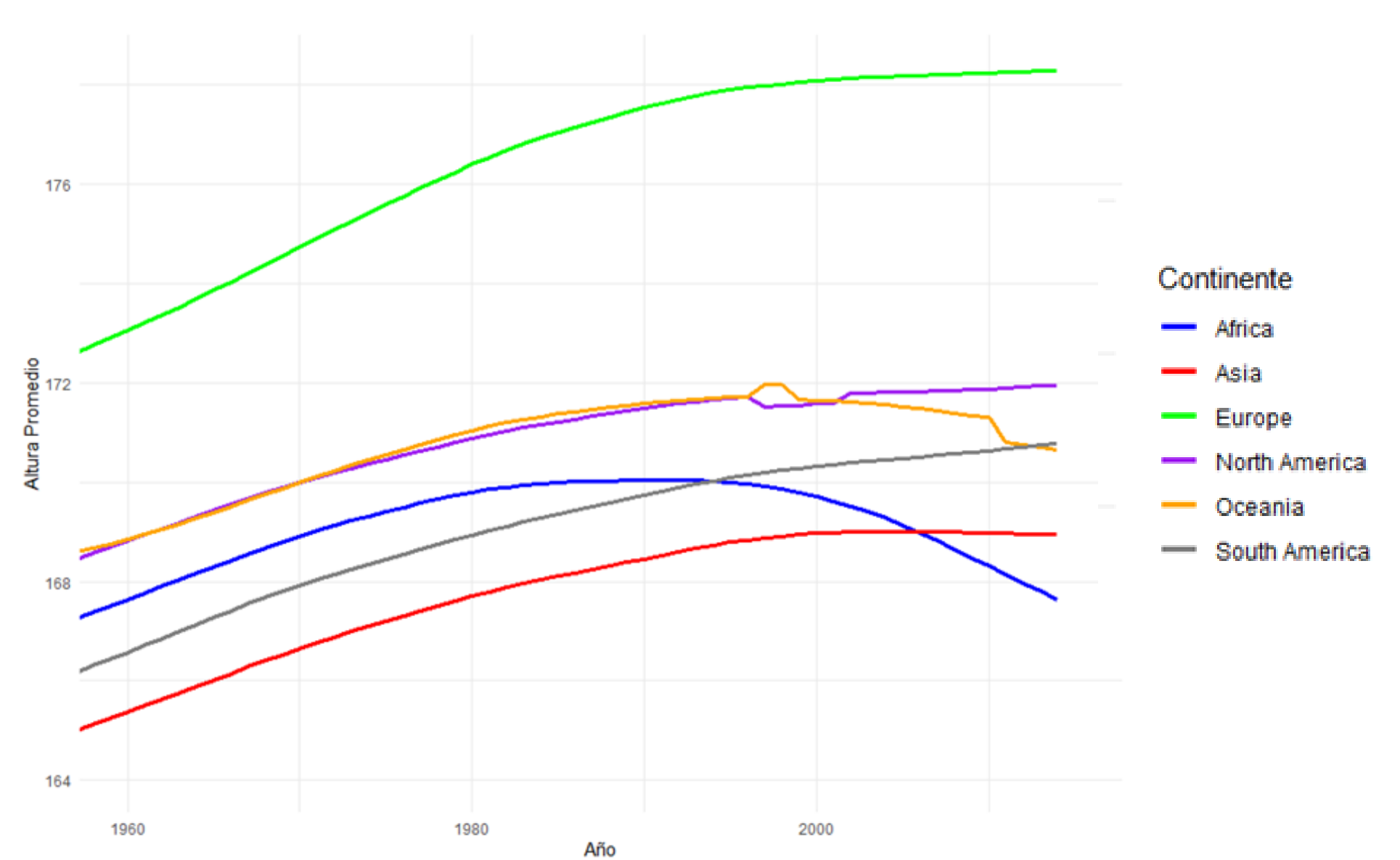


ANÁLISIS EXPLORATORIO

PBI per capita promedio por continente a través de los años



Altura promedio por continente a través de los años

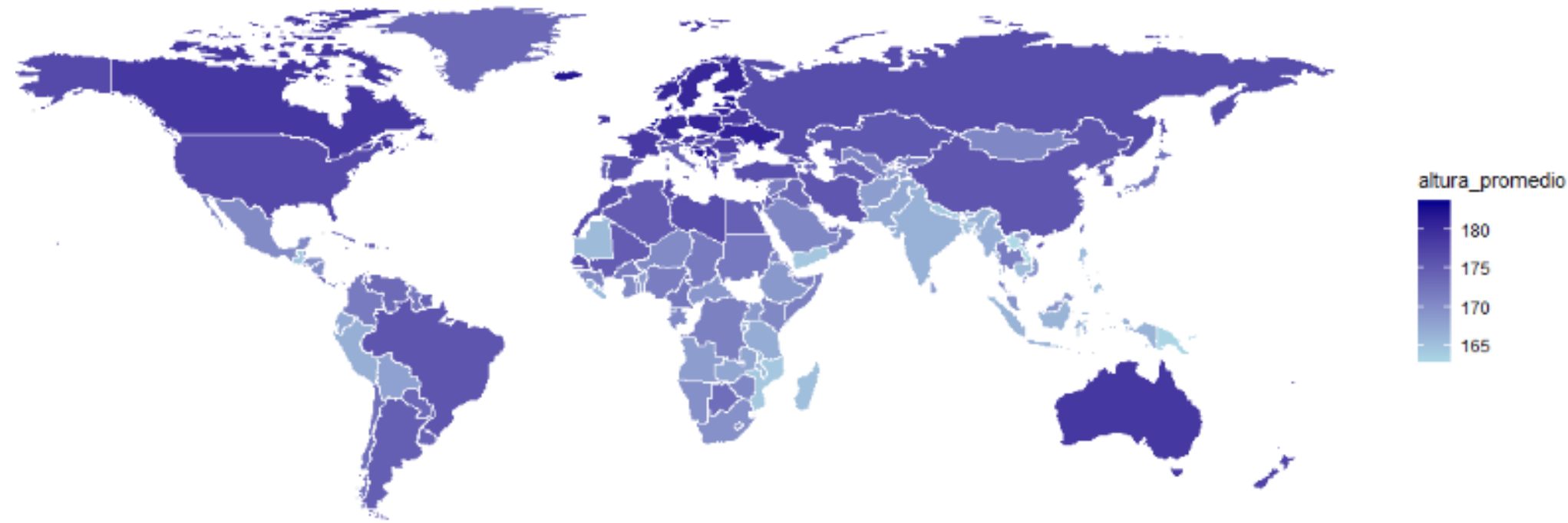


**MMM... PARECE
SER QUE EL
CONTINENTE TIENE
IMPORTANCIA**



ANÁLISIS EXPLORATORIO

Altura promedio en 2019



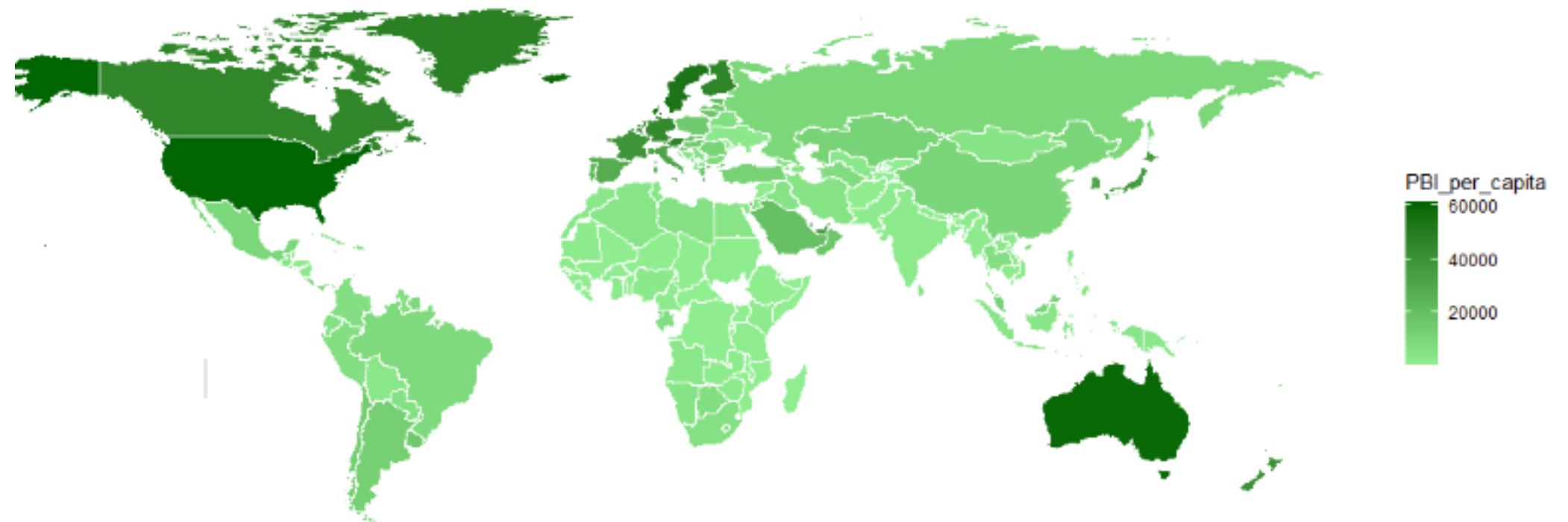
- Australia
- EE.UU
- Canadá
- Europa



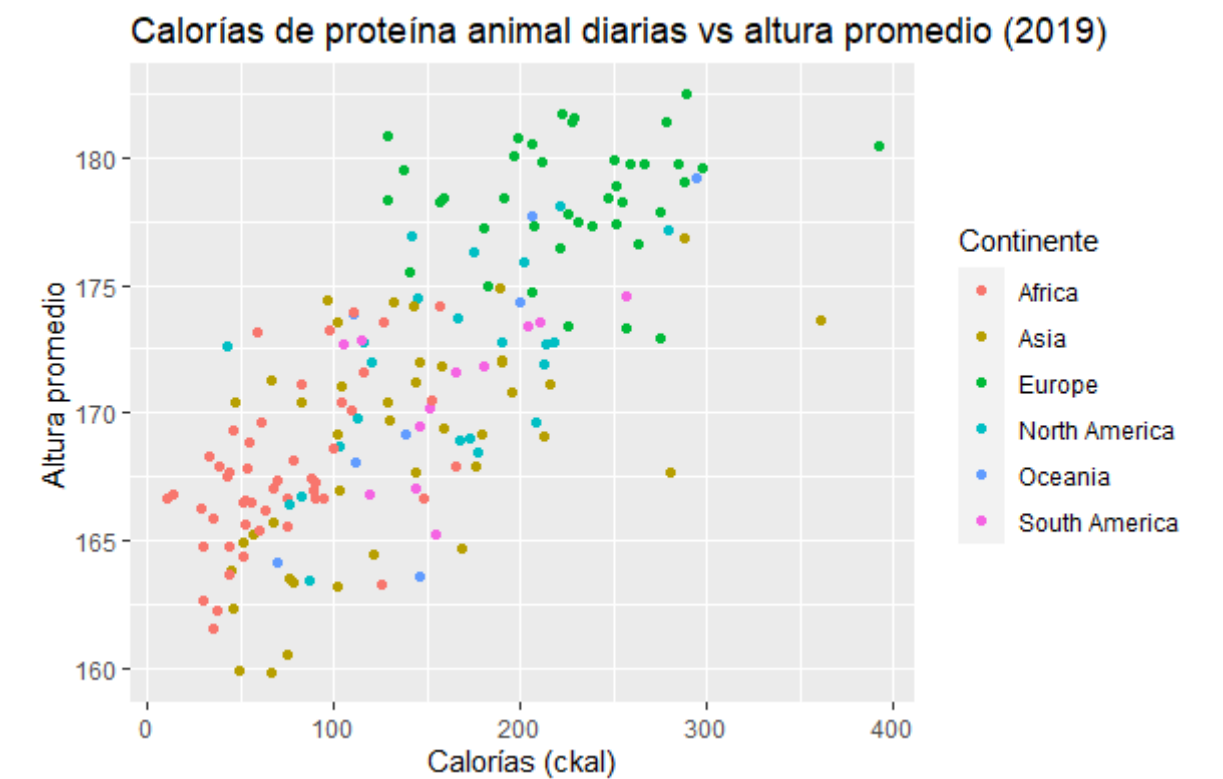
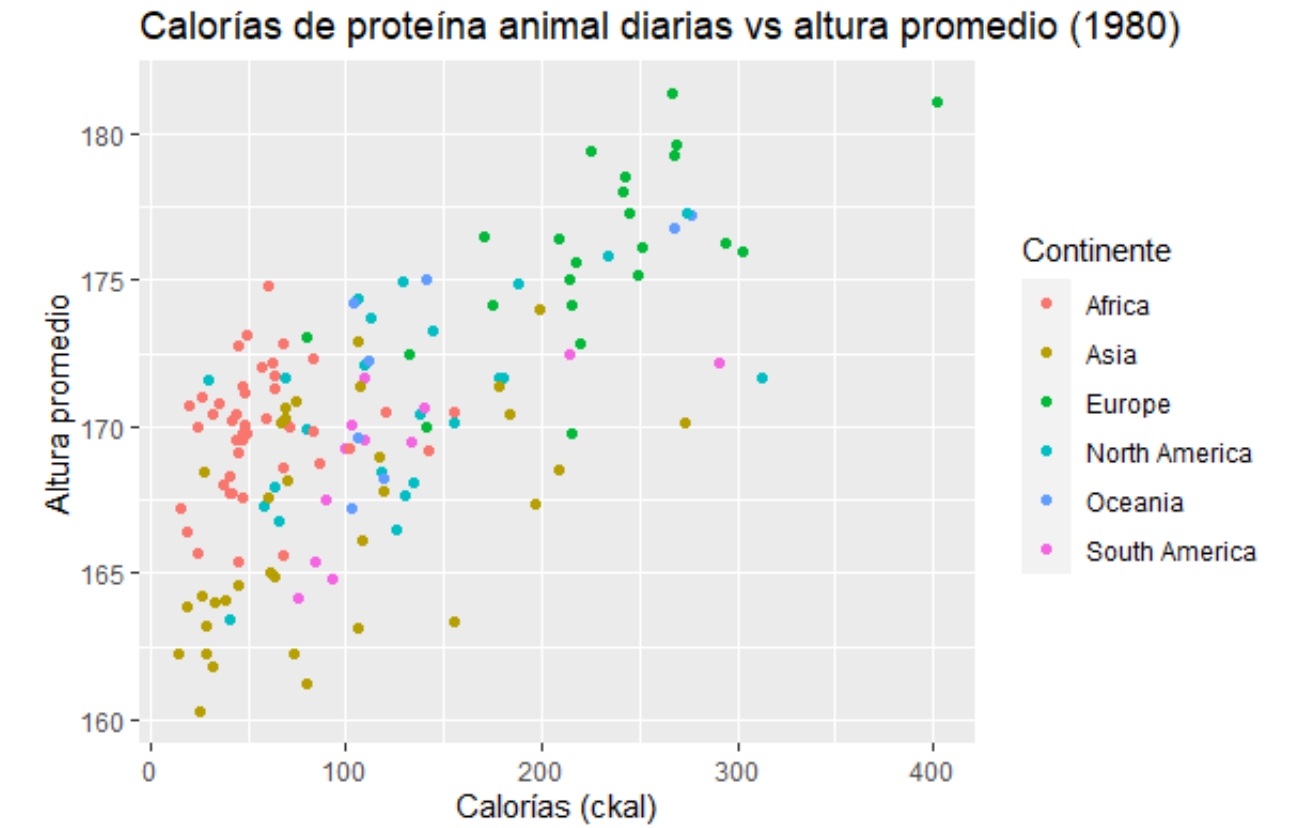
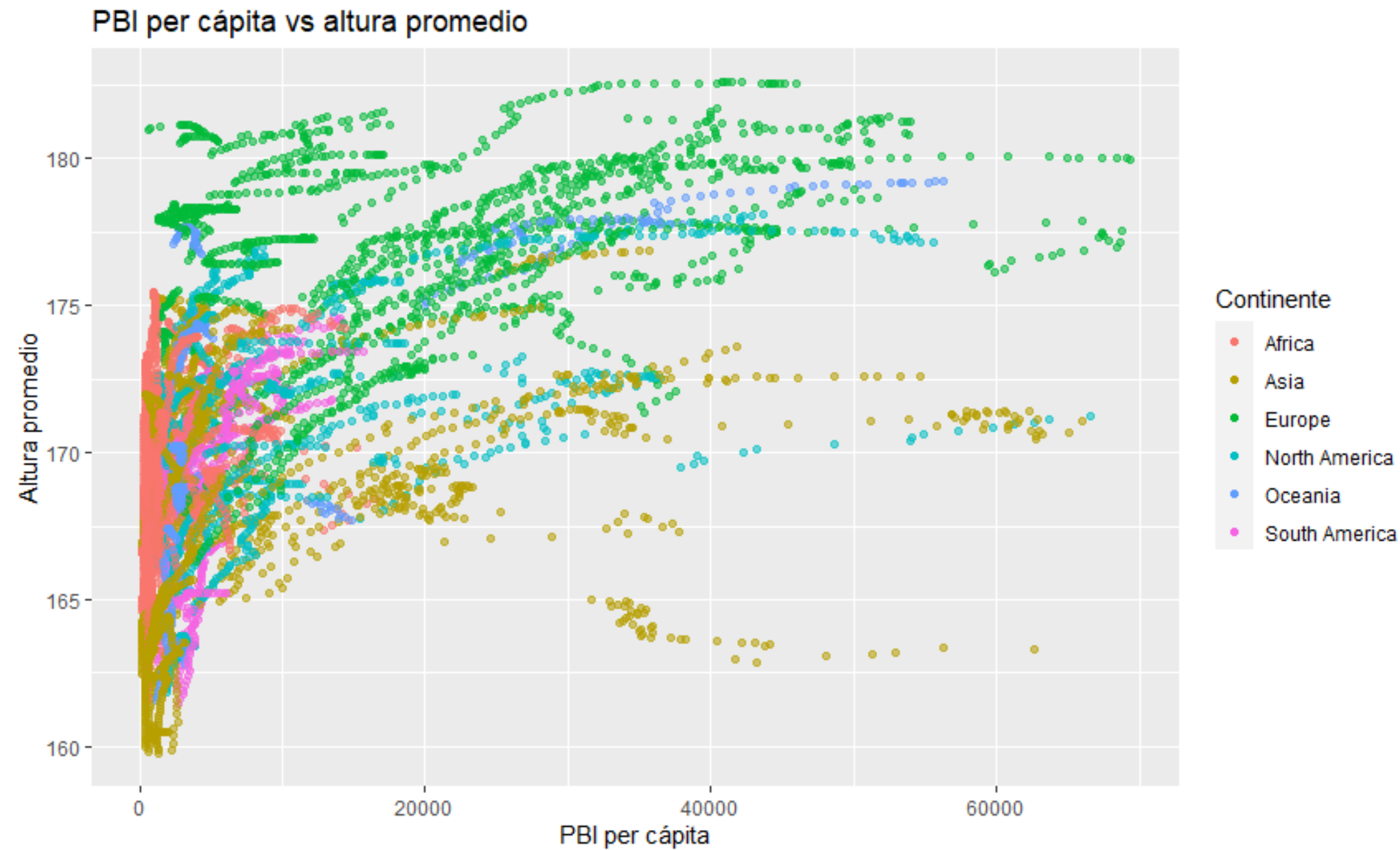
- Bolivia
- Perú
- India
- Mongolia



PBI per capita en 2019



ANÁLISIS EXPLORATORIO

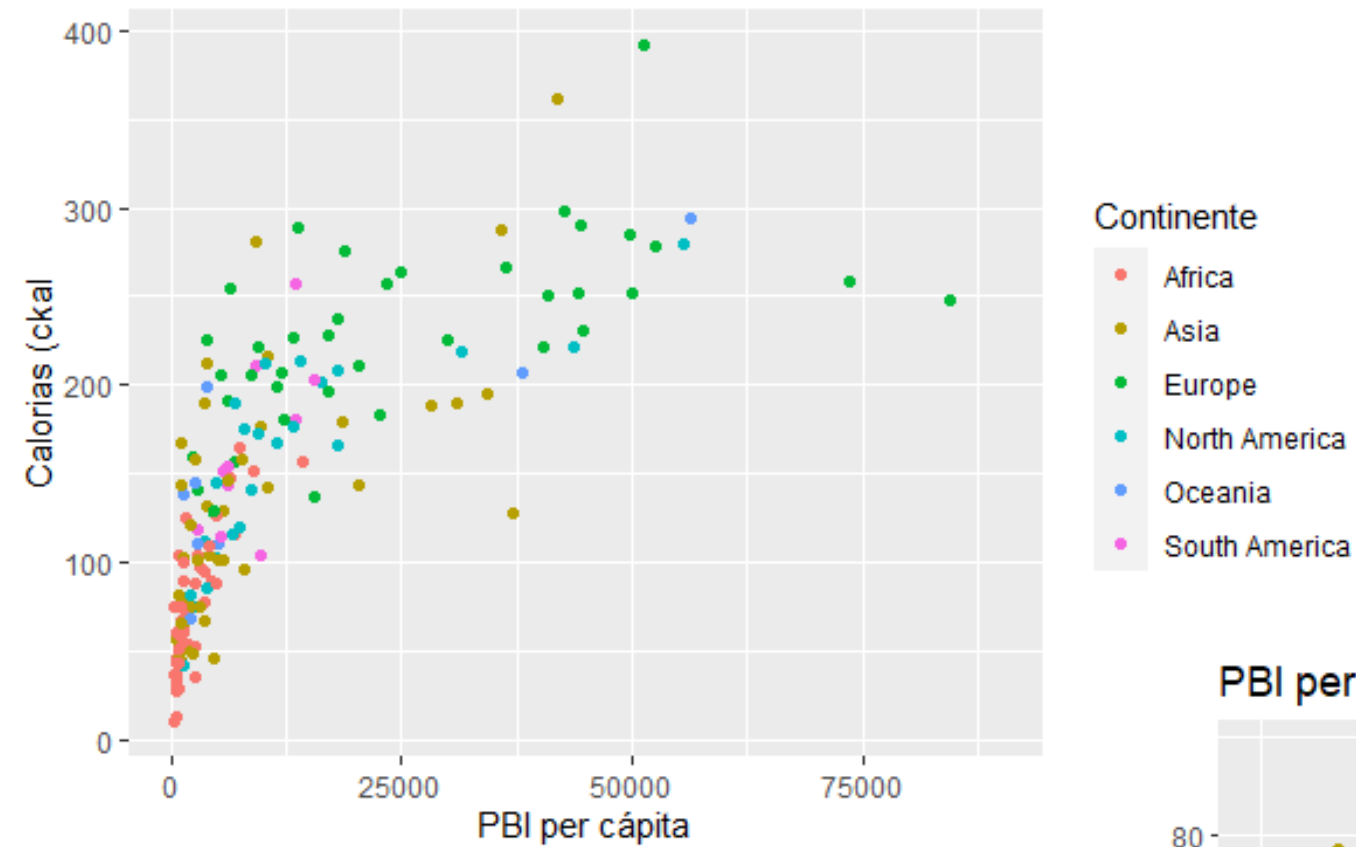


QUIERO FIJAR EL ENFOQUE EN EL PBI

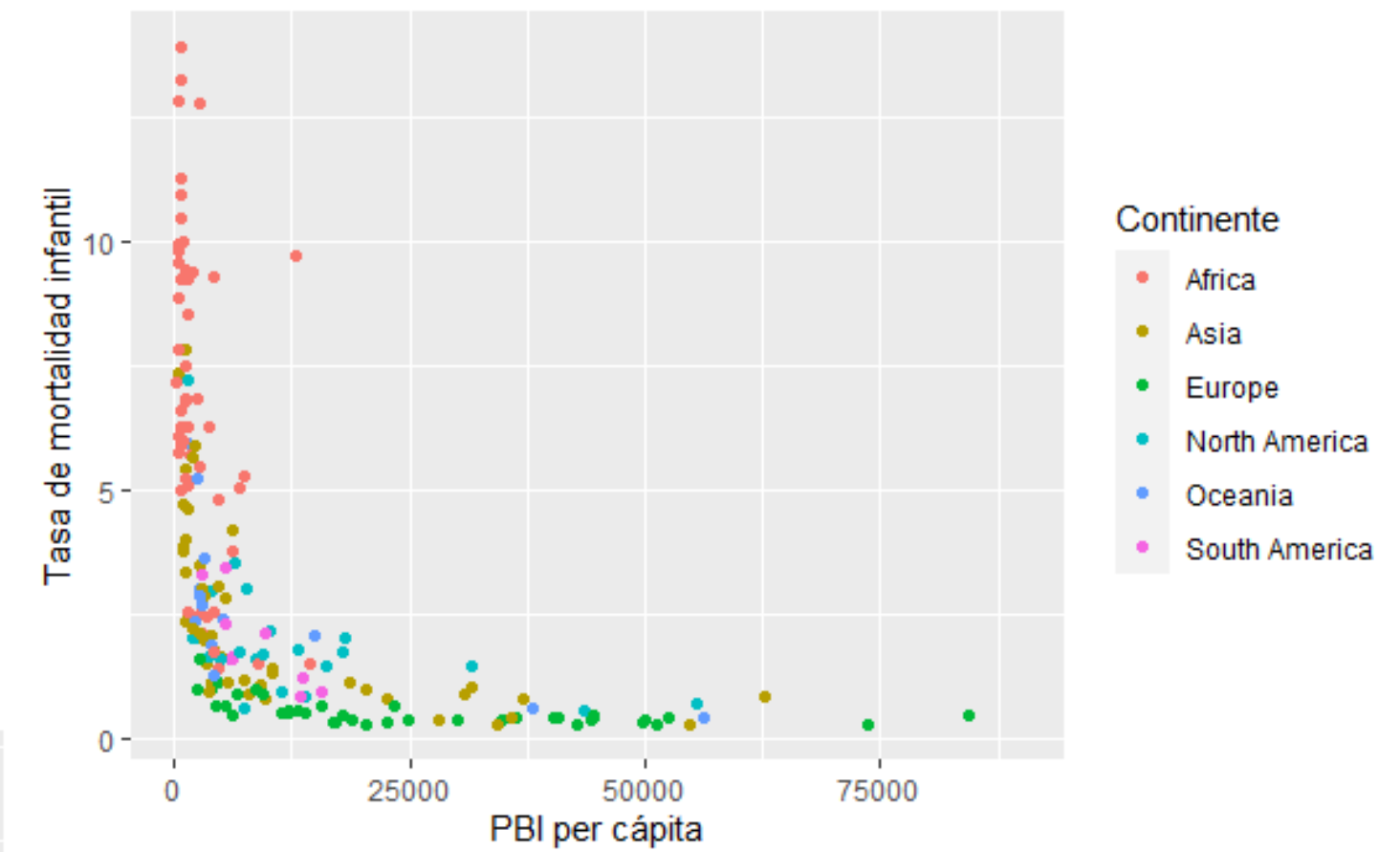


ANÁLISIS EXPLORATORIO

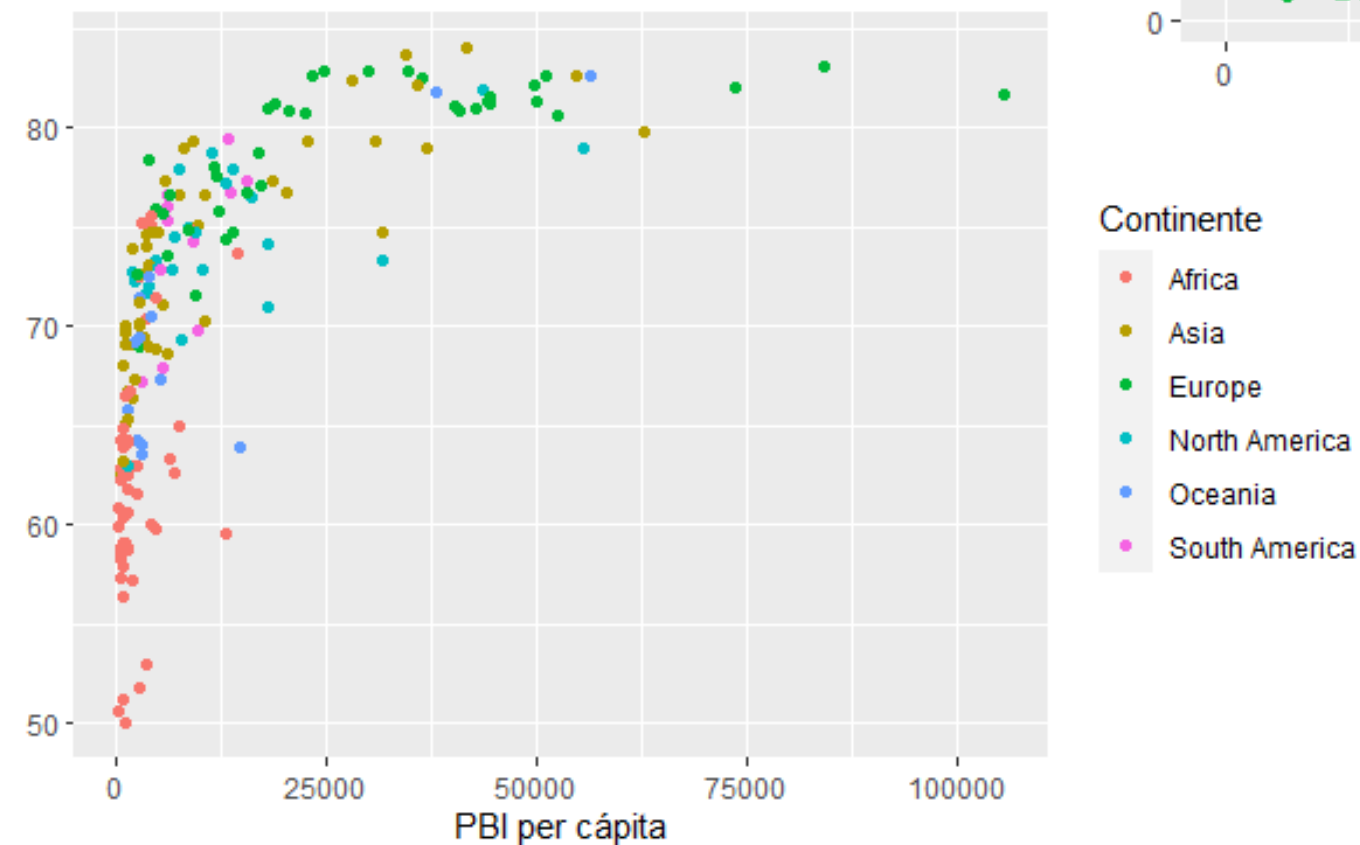
PBI per cápita vs Calorías diarias de proteína animal(2019)



PBI per cápita vs Tasa de mortalidad infantil (2019)



PBI per cápita vs esperanza de vida(2019)



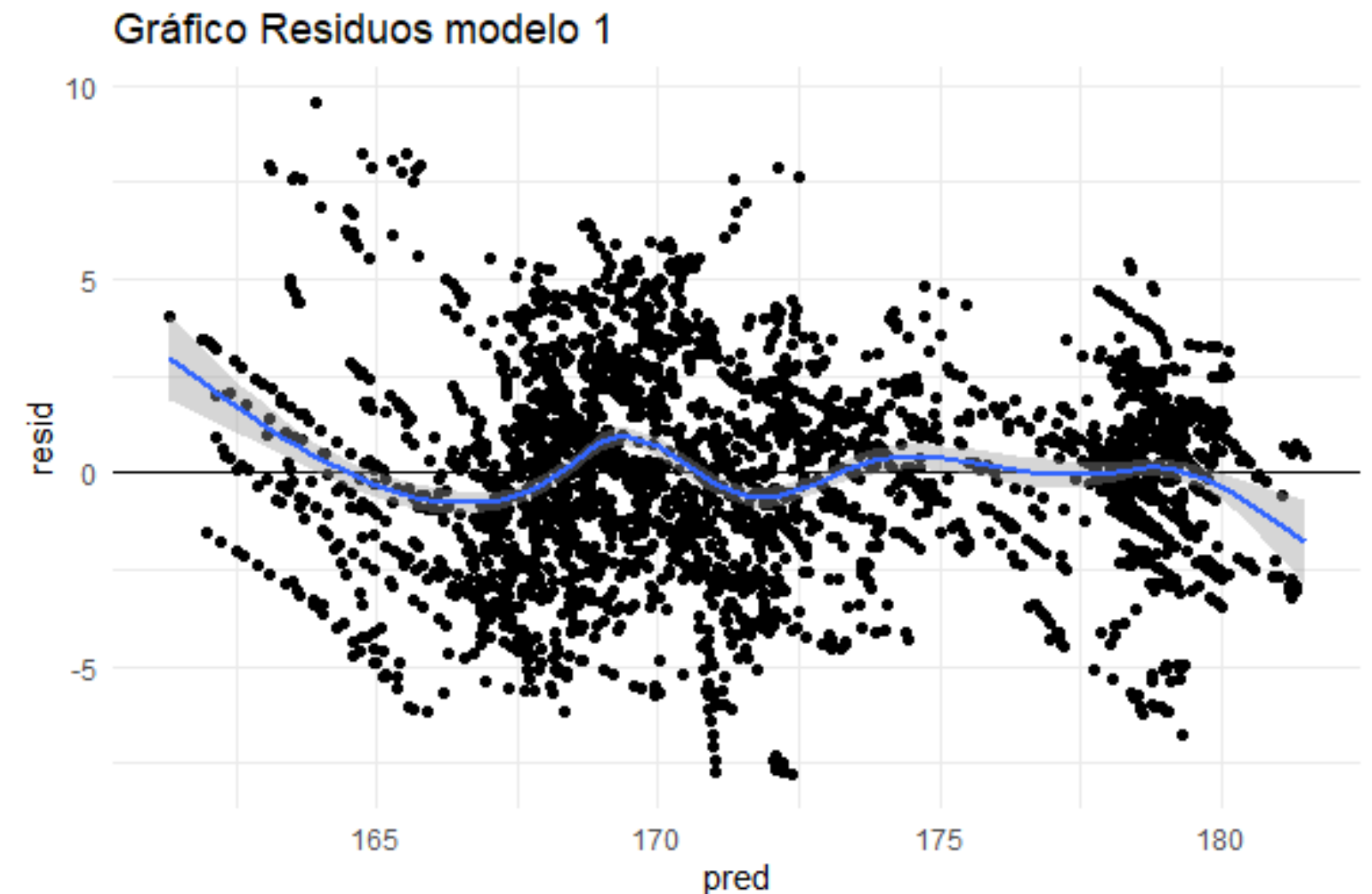
MODELO 1: MÁS SIMPLE

```
MOD1 <- LM(ALTURA_PROMEDIO ~ CONTINENTE * LOG(PBI_PER_CAPITA))
```

Variables	Coeficiente
Intercept	158,76 ***
Asia	-6,82 ***
Europa	15,03 ***
Norteamérica	-10,4 ***
Oceanía	-15,12 ***
Sudamérica	-24,53 ***
log(PBI per cápita)	1,36 ***
Asia: log(PBI per cápita)	0,64 ***
Europa: log(PBI per cápita)	-0,89 ***
Norteamérica: log(PBI per cápita)	1,26 ***
Oceanía: log(PBI per cápita)	-1,95 ***
Sudamérica: log(PBI per cápita)	2,79 ***

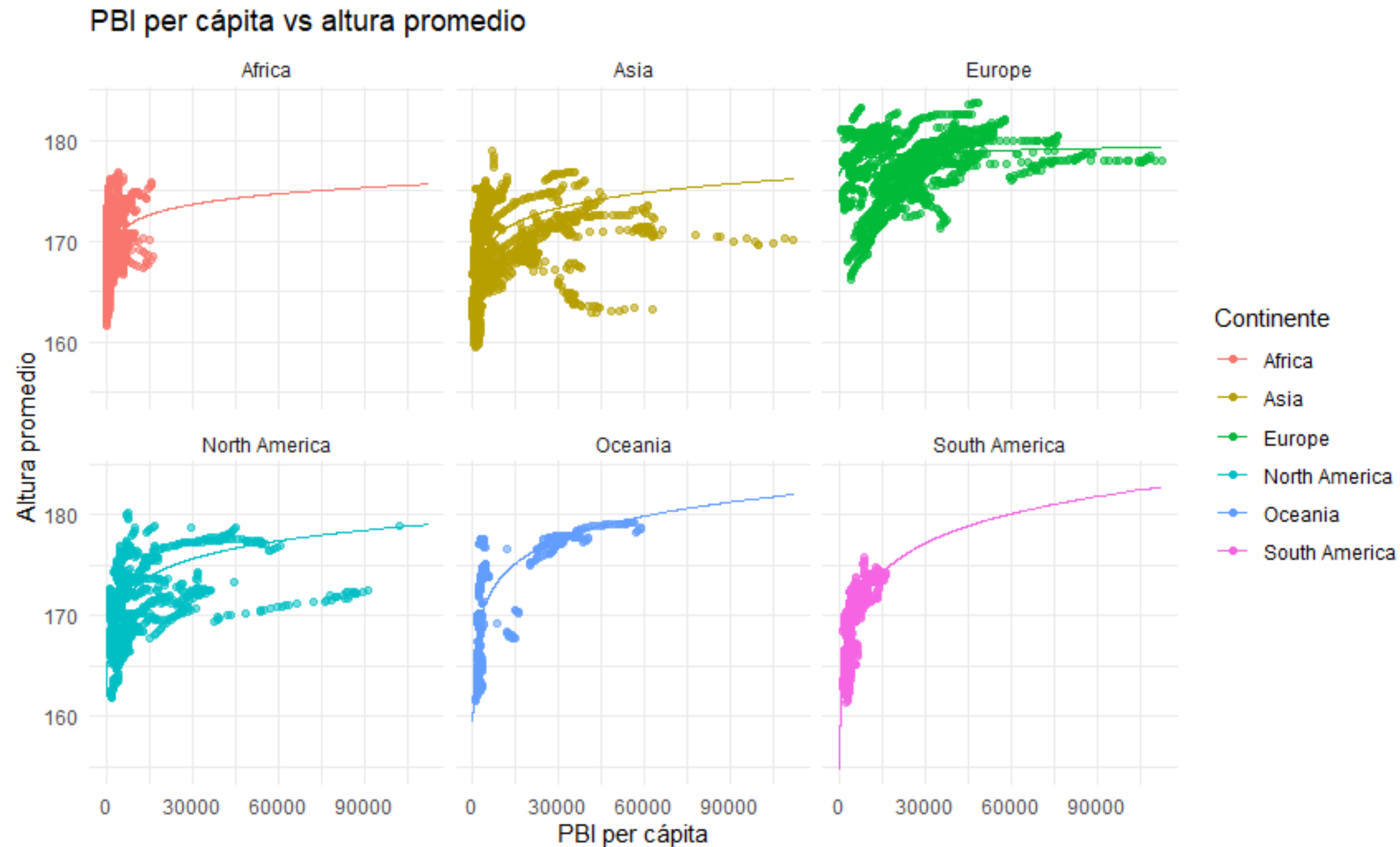
R-cuadrado: 0,725

Error residual: 2,83



MODELO 1: MÁS SIMPLE

```
MOD1 <- LM(ALTURA_PROMEDIO ~ CONTINENTE * LOG(PBI_PER_CAPITA))
```



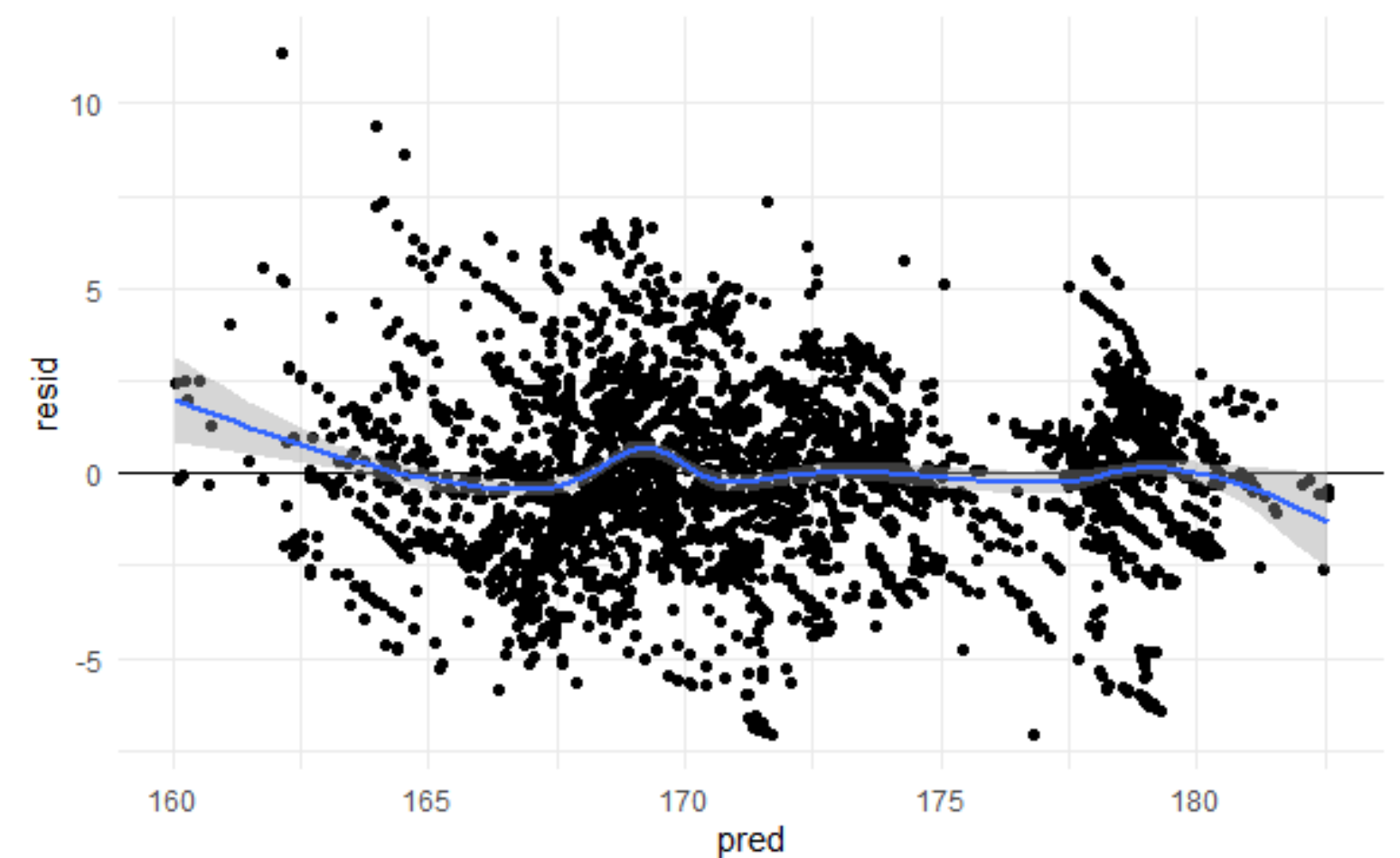
MODELO 2: MENOS SIMPLE

```
MOD2 <- LM(ALTURA_PROMEDIO ~ CONTINENTE *(LOG(PBI_PER_CAPITA)+  
CALORIAS_PROTEINA+EXPECTATIVA_VIDA+TASA_MORTALIDAD_NIÑOS))
```

Coefficients:

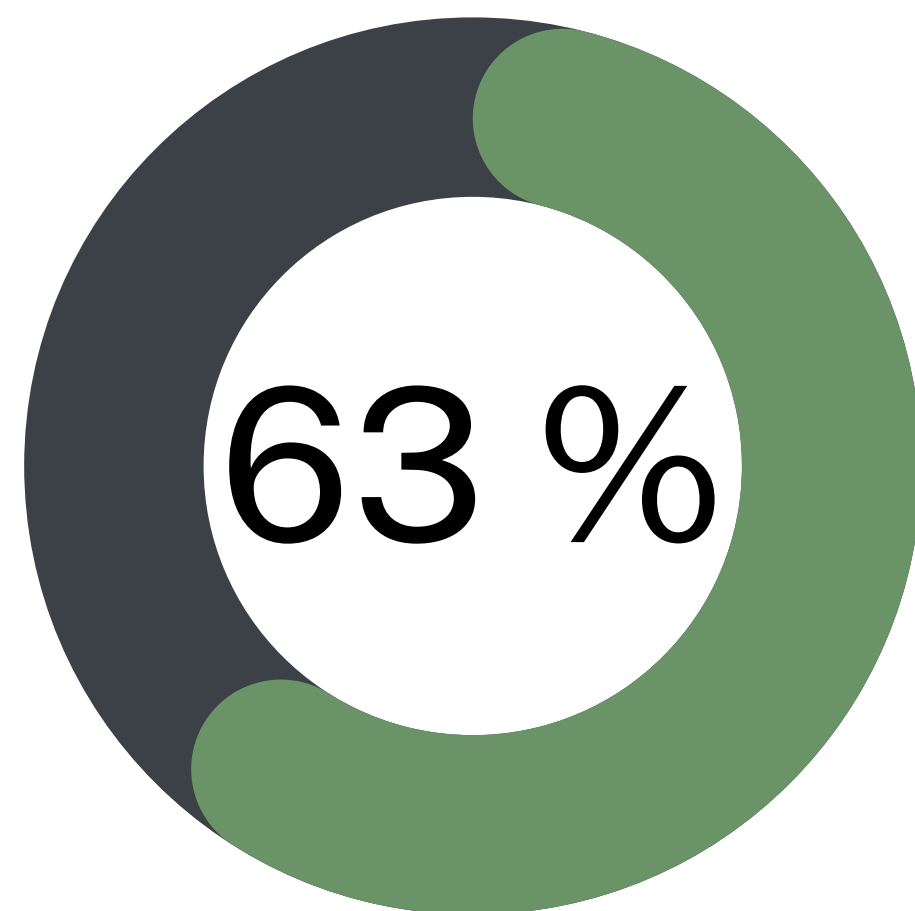
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	150.314933	2.580015	58.261	< 2e-16	***
ContinenteAsia	8.241828	4.373856	1.884	0.059661	.
ContinenteEurope	47.085336	4.402154	10.696	< 2e-16	***
ContinenteNorth America	28.715189	6.844757	4.195	2.84e-05	***
ContinenteOceania	34.012741	9.331200	3.645	0.000274	***
ContinenteSouth America	17.179954	16.909172	1.016	0.309743	
log(PBI_per_capita)	1.114017	0.216911	5.136	3.08e-07	***
calorias_proteina	0.013417	0.005389	2.490	0.012868	*
expectativa_vida	0.124559	0.026308	4.735	2.35e-06	***
Tasa_mortalidad_niños	0.225185	0.050673	4.444	9.31e-06	***
ContinenteAsia:log(PBI_per_capita)	-1.086348	0.286927	-3.786	0.000157	***
ContinenteEurope:log(PBI_per_capita)	-0.407099	0.317112	-1.284	0.199367	
ContinenteNorth America:log(PBI_per_capita)	0.893465	0.516420	1.730	0.083761	.
ContinenteOceania:log(PBI_per_capita)	0.900713	0.771697	1.167	0.243272	
ContinenteSouth America:log(PBI_per_capita)	3.260711	1.002014	3.254	0.001156	**
ContinenteAsia:calorias_proteina	0.015546	0.006112	2.543	0.011053	*
ContinenteEurope:calorias_proteina	-0.020197	0.006171	-3.273	0.001082	**
ContinenteNorth America:calorias_proteina	0.023810	0.008237	2.891	0.003885	**
ContinenteOceania:calorias_proteina	0.034023	0.010590	3.213	0.001335	**
ContinenteSouth America:calorias_proteina	0.010324	0.009315	1.108	0.267858	
ContinenteAsia:expectativa_vida	-0.018329	0.058670	-0.312	0.754763	
ContinenteEurope:expectativa_vida	-0.413828	0.055858	-7.409	1.86e-13	***
ContinenteNorth America:expectativa_vida	-0.543478	0.104015	-5.225	1.92e-07	***
ContinenteOceania:expectativa_vida	-0.603250	0.151065	-3.993	6.75e-05	***
ContinenteSouth America:expectativa_vida	-0.638950	0.147251	-4.339	1.50e-05	***
ContinenteAsia:Tasa_mortalidad_niños	-0.582019	0.110774	-5.254	1.64e-07	***
ContinenteEurope:Tasa_mortalidad_niños	-3.251627	0.477535	-6.809	1.29e-11	***
ContinenteNorth America:Tasa_mortalidad_niños	-0.229333	0.195612	-1.172	0.241179	
ContinenteOceania:Tasa_mortalidad_niños	-1.268179	0.298889	-4.243	2.30e-05	***
ContinenteSouth America:Tasa_mortalidad_niños	-0.870204	0.620933	-1.401	0.161232	

Gráfico Residuos modelo 2



R-cuadrado: 0,774

Error residual: 2,58



TÍTULO AQUÍ

TU INCREÍBLE TÍTULO AQUÍ

Purus in mollis nunc sed id
semper risus in hendrerit
gravida rutrum quisque non
tellus orci ac auctor augue
mauris augue neque gravida in
fermentum et sollicitudin ac
orci phasellus.

CONCLUSIONES

- Los países más ricos (los que tienen un PBI per cápita más alto) suelen ser más altos. Aunque esta tendencia se aplana en el extremo - en otras palabras, ser rico sólo te lleva hasta cierto punto.
- Los factores socioeconómicos en la infancia podrían influir en la estatura a través de su efecto sobre las condiciones de vida, como las malas condiciones de vivienda, la malnutrición o el acceso inadecuado a la atención sanitaria.

TRABAJO A FUTURO

- Comportamiento en mujeres.
- Expandir el análisis incluyendo más variables como la atención sanitaria, educación, medioambiente.
- Analizar a la tribu Dinka, de Sudán del Sur.

¡GRACIAS!

