An analysis of global life expectancy

Rebecca Barter

Table of contents

Below, we load in the gapminder dataset that we will use for analysis in this document.

```
library(tidyverse)
library(gt)
library(knitr)
gapminder <- read.csv("data/gapminder.csv")
```

Evaluating life expectancy by continent

The table below shows the average life expectancy every 5 years for each continent.

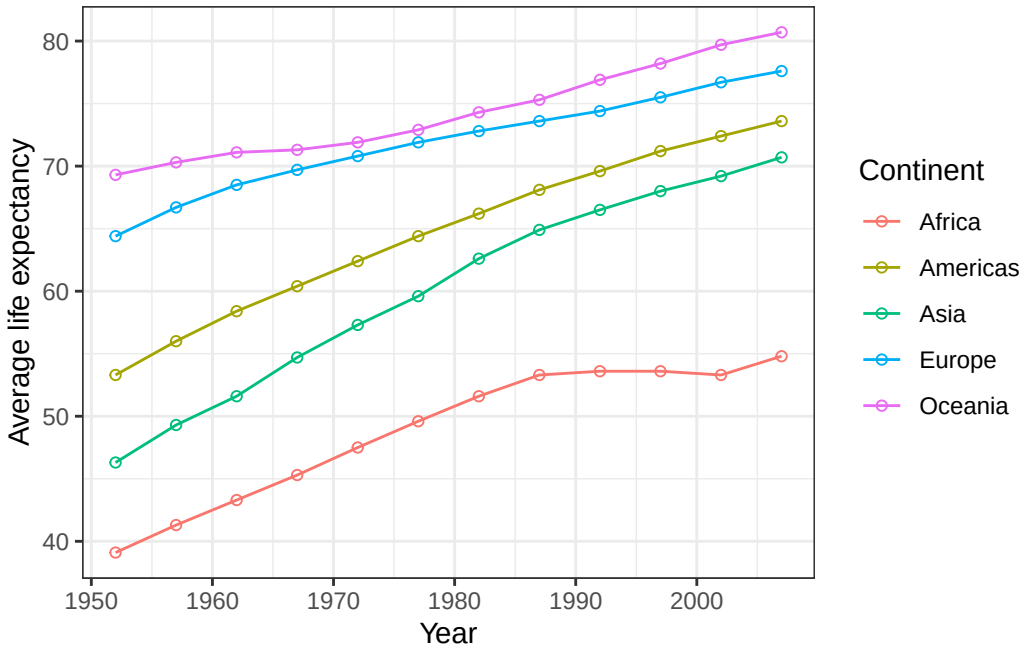
```
life_exp <- gapminder |>
  group_by(year, continent) |>
  summarise(mean_life_exp = round(mean(lifeExp), 1)) |>
  ungroup()


life_exp |>
  pivot_wider(id_cols = "year",
              values_from = "mean_life_exp",
              names_from = "continent") |>
  kable()
```

Table 1: Table describing the average life expectancy over time

year	Africa	Americas	Asia	Europe	Oceania
1952	39.1	53.3	46.3	64.4	69.3
1957	41.3	56.0	49.3	66.7	70.3
1962	43.3	58.4	51.6	68.5	71.1
1967	45.3	60.4	54.7	69.7	71.3
1972	47.5	62.4	57.3	70.8	71.9
1977	49.6	64.4	59.6	71.9	72.9
1982	51.6	66.2	62.6	72.8	74.3
1987	53.3	68.1	64.9	73.6	75.3
1992	53.6	69.6	66.5	74.4	76.9
1997	53.6	71.2	68.0	75.5	78.2
2002	53.3	72.4	69.2	76.7	79.7
2007	54.8	73.6	70.7	77.6	80.7

We can visualize these trends for each country in the line plot below:

```
life_exp |> ggplot(aes(x = year, y = mean_life_exp, color = continent)) +
  geom_line() +
  geom_point(shape = 1) +
  labs(x = "Year", y = "Average life expectancy", color = "Continent") +
  theme_bw()
```



Notice that the trends are increasing overall, with Oceania having the highest life expectancy and Africa having the lowest. Unlike the other continents, the life expectancy for African countries are stagnating around 1990, but started to increase again around 2007.

Measuring life expectancy against GDP

The scatterplot below shows the relationship between GDP per capita and life expectancy in 2007, colored by continent.

```
gg_scatter <- gapminder |>
  filter(year == 2007) |>
  ggplot() +
  geom_point(aes(x = gdpPercap, y = lifeExp,
                 color = continent, size = pop),
             alpha = 0.5) +
  scale_size_continuous("Population",
                        labels = c("1 million", "250 million", "1 billion"),
                        breaks = c(1, 250, 1000) * 1e6) +
  labs(x = "GDP per capita", y = "Life expectancy", color = "Continent") +
  theme_bw()
gg_scatter
```

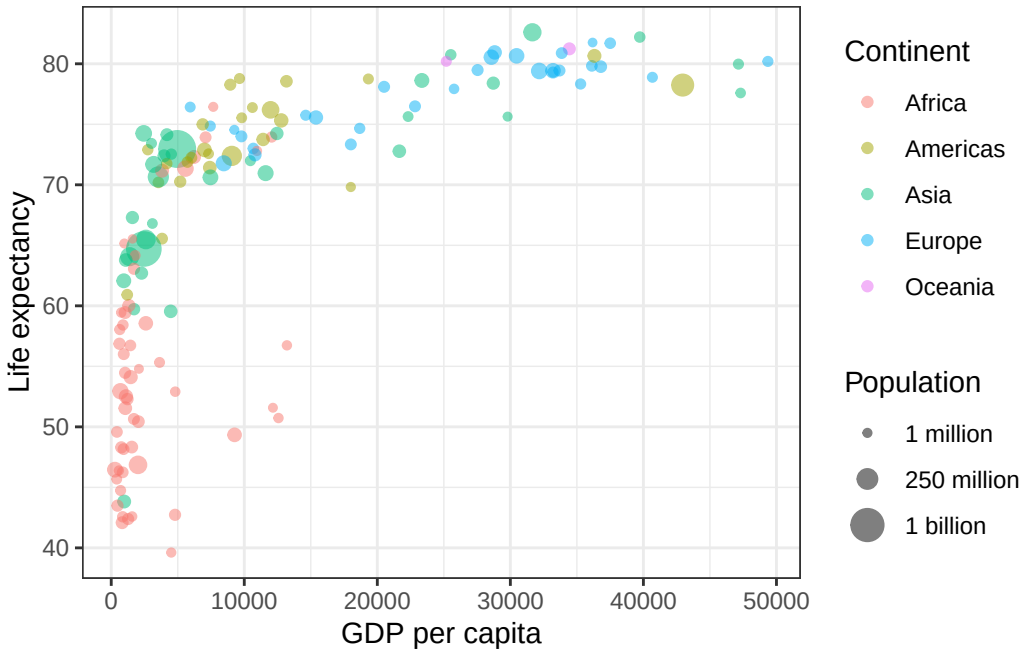


Figure 1: A scatterplot of GDP per capita against life expectancy

It seems that overall, the higher the GDP per capita, the higher the life expectancy.

This look a like an exponential relationship. In Figure Figure 2, we convert the x-axis to a log-scale.

```
gg_scatter + scale_x_log10()
```

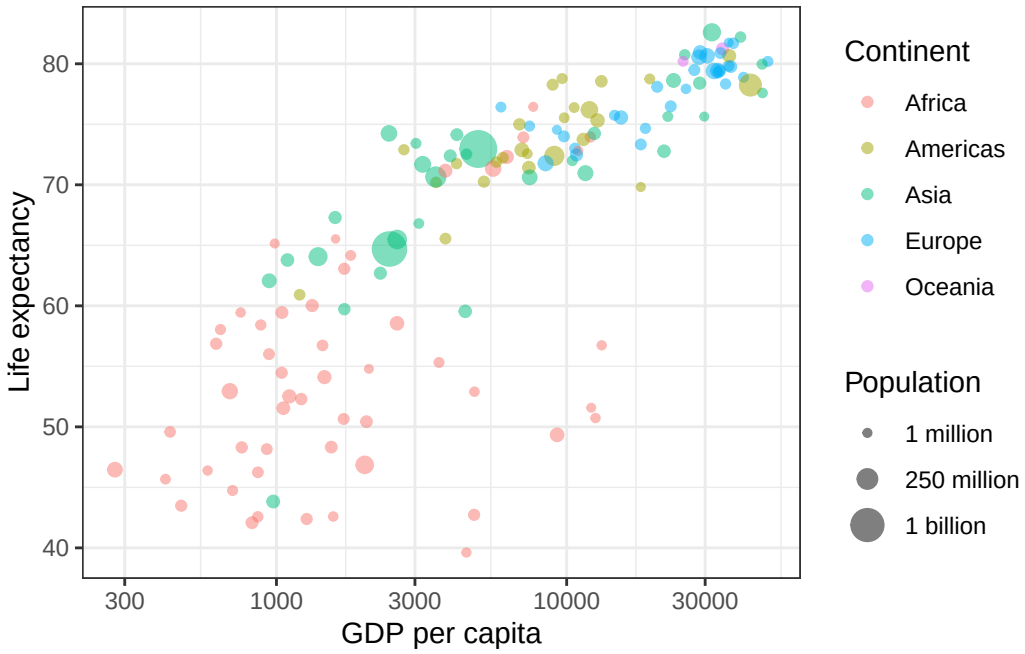


Figure 2: A scatterplot of GDP per capita against life expectancy, with the x-axis presented on a log-scale.