

Part 06 - Reporting with Quarto

Introduction to R for Health Data Science: Hands-on training

Carlos Matos

1 Exploring the gapminder dataset

This document is part of the Introduction to R for Health Data Science: Hands-on training.

We will take a closer look at Portugal.

In Section [1.1.1](#) we will look at how life expectancy changed in Portugal over time.

In Section [1.1.2](#) we will make a regression of lifeExp over gdpPercap in Portugal and report the results.

First, let's load the necessary packages:¹⁻³

```
```{r}
#| message: false
#| warning: false
library(tidyverse)
library(gapminder)
library(report)
```
```

Then, let's take a look at the dataset.

```
gapminder %>% glimpse()
```

```
Rows: 1,704
```

```
Columns: 6
```

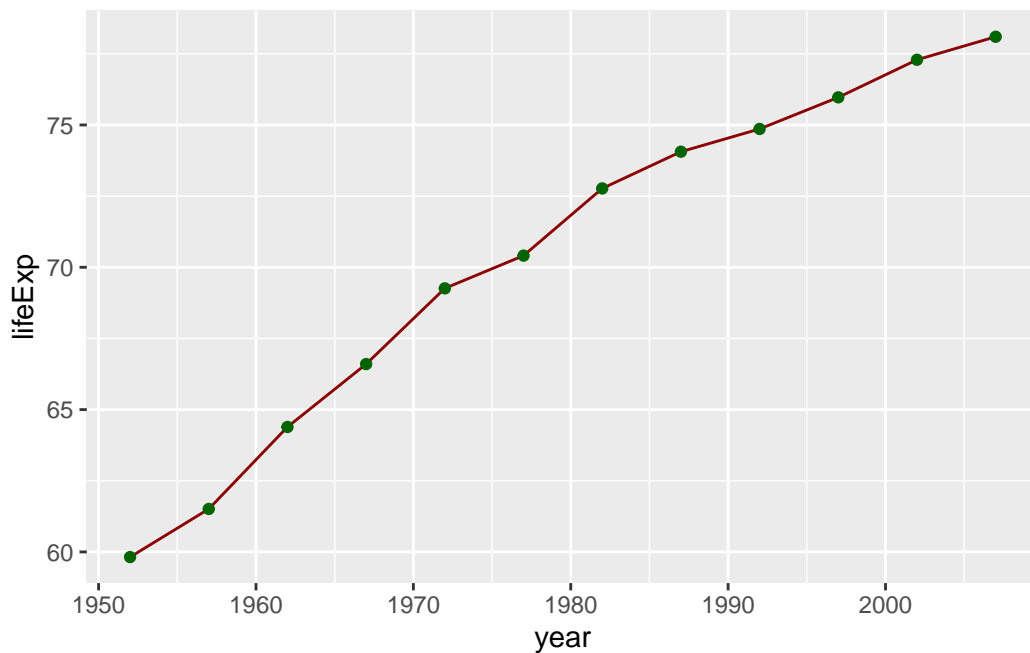
```
$ country   <fct> "Afghanistan", "Afghanistan", "Afghanistan", "Afghanistan", ~
$ continent <fct> Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, ~
$ year      <int> 1952, 1957, 1962, 1967, 1972, 1977, 1982, 1987, 1992, 1997, ~
$ lifeExp   <dbl> 28.801, 30.332, 31.997, 34.020, 36.088, 38.438, 39.854, 40.8~
$ pop       <int> 8425333, 9240934, 10267083, 11537966, 13079460, 14880372, 12~
$ gdpPercap <dbl> 779.4453, 820.8530, 853.1007, 836.1971, 739.9811, 786.1134, ~
```

We see a tibble with 1704 rows and 6 columns.

1.1 Portugal analysis

1.1.1 Over time

```
gapminder %>%  
  filter(country == "Portugal") %>%  
  ggplot(aes(x = year, y = lifeExp)) +  
  geom_line(color = "darkred") +  
  geom_point(color = "darkgreen")
```



We see that that life expectancy **increased continuously**.

1.1.2 Regression modelling

```
gapminder_pt <- gapminder %>%  
  filter(country == "Portugal")
```

```
lm(lifeExp ~ gdpPercap, data = gapminder_pt) %>%  
  report()
```

We fitted a linear model (estimated using OLS) to predict lifeExp with gdpPercap (formula: `lifeExp ~ gdpPercap`). The model explains a statistically significant and substantial proportion of variance ($R^2 = 0.94$, $F(1, 10) = 160.60$, $p < .001$, adj. $R^2 = 0.94$). The model's intercept, corresponding to $\text{gdpPercap} = 0$, is at 59.49 (95% CI [57.32, 61.66], $t(10) = 61.06$, $p < .001$). Within this model:

- The effect of gdpPercap is statistically significant and positive ($\beta = 9.63\text{e-}04$, 95% CI [$7.94\text{e-}04$, $1.13\text{e-}03$], $t(10) = 12.67$, $p < .001$; Std. $\beta = 0.97$, 95% CI [0.80, 1.14])

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald t-distribution approximation.

2 References

1. Wickham H, Averick M, Bryan J, et al. Welcome to the {tidyverse}. 2019;4:1686. doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686)
2. Bryan J. Gapminder: Data from gapminder. Published online 2023. <https://CRAN.R-project.org/package=gapminder>
3. Makowski D, Lüdtke D, Patil I, Thériault R, Ben-Shachar MS, Wiernik BM. Automated results reporting as a practical tool to improve reproducibility and methodological best practices adoption. Published online 2023. <https://easystats.github.io/report/>