Health Spending Analysis

Hugo Jal

2024-03-22

```
knitr::opts_chunk$set(echo = FALSE)
```

In this study, I plan to analyze the health expenditure per person across countries and how it has evolved during the last years. Moreover, I will also evaluate this spending by comparing the expenditure by continent.

```
library(ggplot2)
library(tidyr)
library(dplyr)
library(tidyverse)
library(ggforce)
library(patchwork)
library(knitr)
```

```
#Load the data
health <- read.csv('total_health_spending_per_person_us.csv')
```

```
head(health)
```

```
##       country  X1995  X1996  X1997  X1998   X1999  X2000  X2001  X2002  X2003
## 1 Afghanistan     NA     NA     NA     NA      NA     NA     NA   14.8   18.3
## 2      Angola   15.6   11.3   13.5    9.1    8.82   15.8   21.4   18.1   23.9
## 3     Albania   27.9   43.0   36.1   47.1   65.00   75.2   79.9   90.3  113.0
## 4     Andorra 1390.0 1510.0 1460.0 1860.0 1430.00 1330.0 1290.0 1490.0 1890.0
## 5         UAE  738.0  731.0  751.0  839.0  847.00  885.0  815.0  917.0  968.0
## 6   Argentina  615.0  619.0  686.0  706.0  726.00  709.0  678.0  225.0  283.0
##    X2004  X2005  X2006  X2007  X2008  X2009  X2010
## 1   20.7   21.9   23.8   28.8   31.8   33.7   37.7
## 2   25.9   36.4   64.1   85.3  149.0  201.0  123.0
## 3  161.0  178.0  192.0  232.0  275.0  260.0  241.0
## 4 2190.0 2360.0 2630.0 3010.0 3390.0 3360.0 3100.0
## 5  995.0 1030.0 1110.0 1200.0 1530.0 1700.0 1450.0
## 6  332.0  400.0  464.0  562.0  699.0  734.0  742.0
```

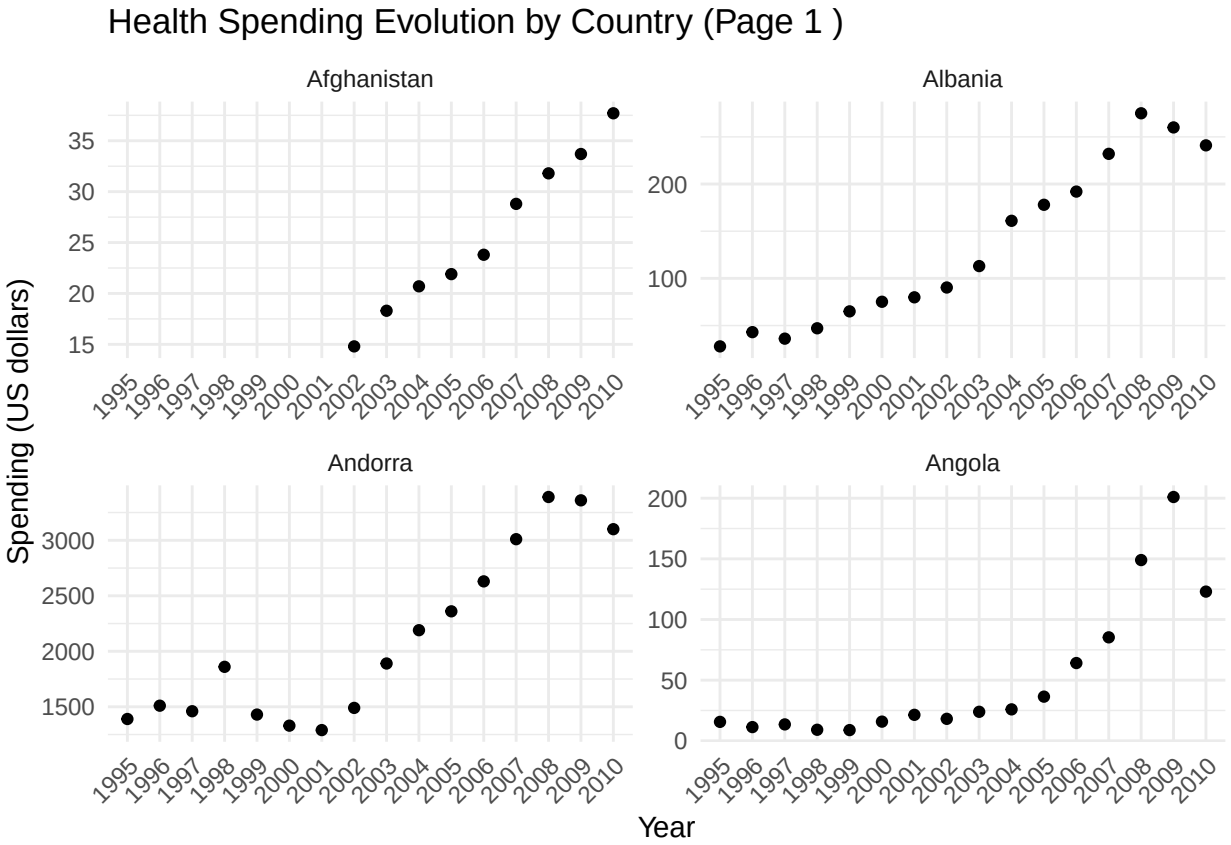
```
health_long <- health %>%
  pivot_longer(
    cols = starts_with("X"),  # Select columns that start with 'X'
    names_to = "Year",        # Name of the new 'Year' column
    names_prefix = "X",       # Remove the 'X' prefix from the column names
    values_to = "Spending"    # Name of the new 'Spending' column
  )
```

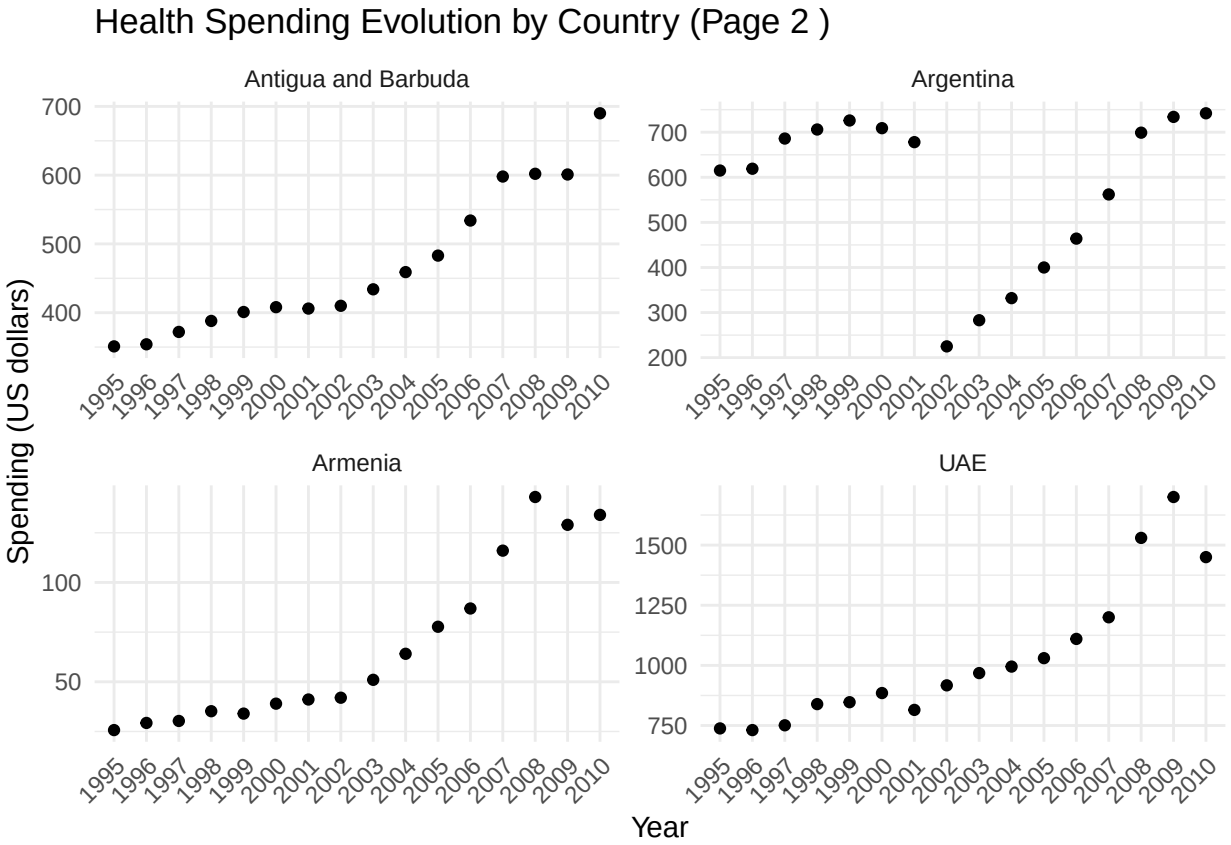
```r
head(health_long)
```

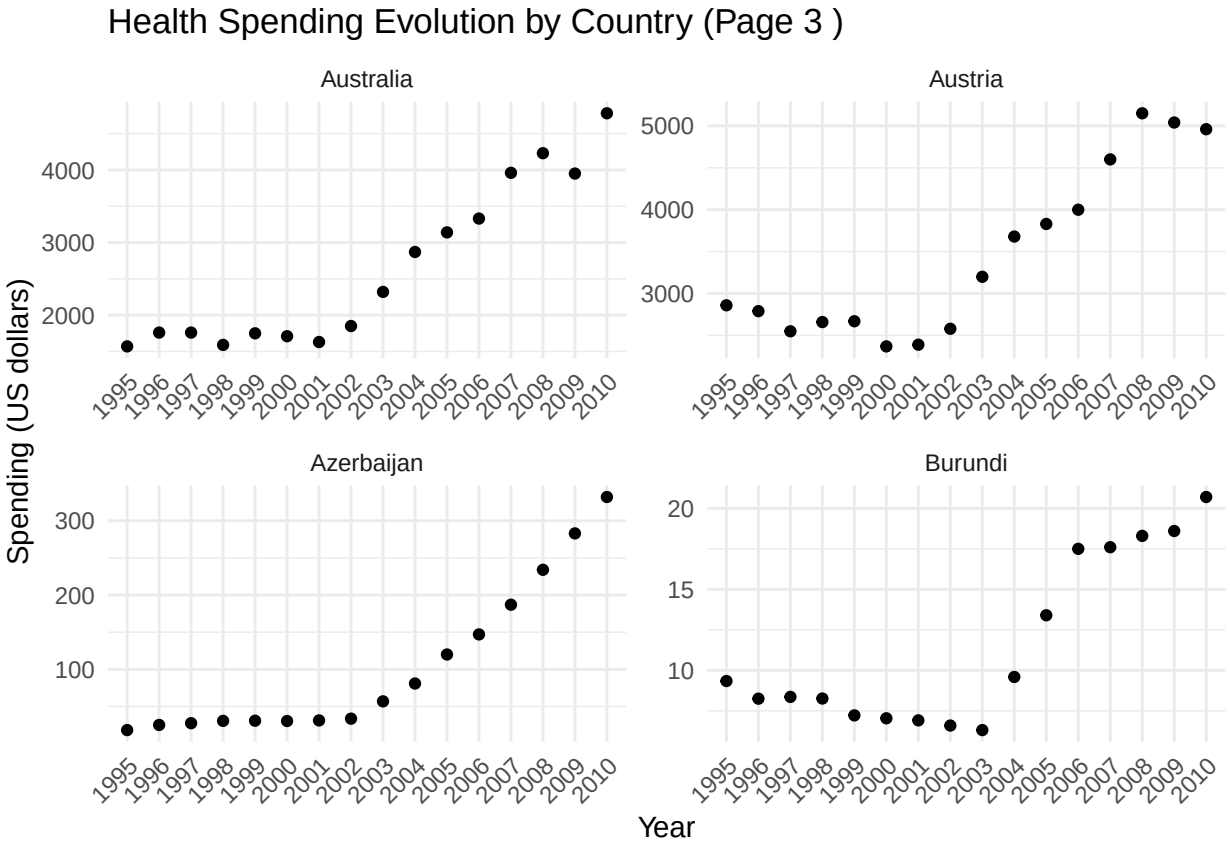
```
## # A tibble: 6 x 3
##   country     Year  Spending
##   <chr>       <chr>    <dbl>
## 1 Afghanistan 1995        NA
## 2 Afghanistan 1996        NA
## 3 Afghanistan 1997        NA
## 4 Afghanistan 1998        NA
## 5 Afghanistan 1999        NA
## 6 Afghanistan 2000        NA
```

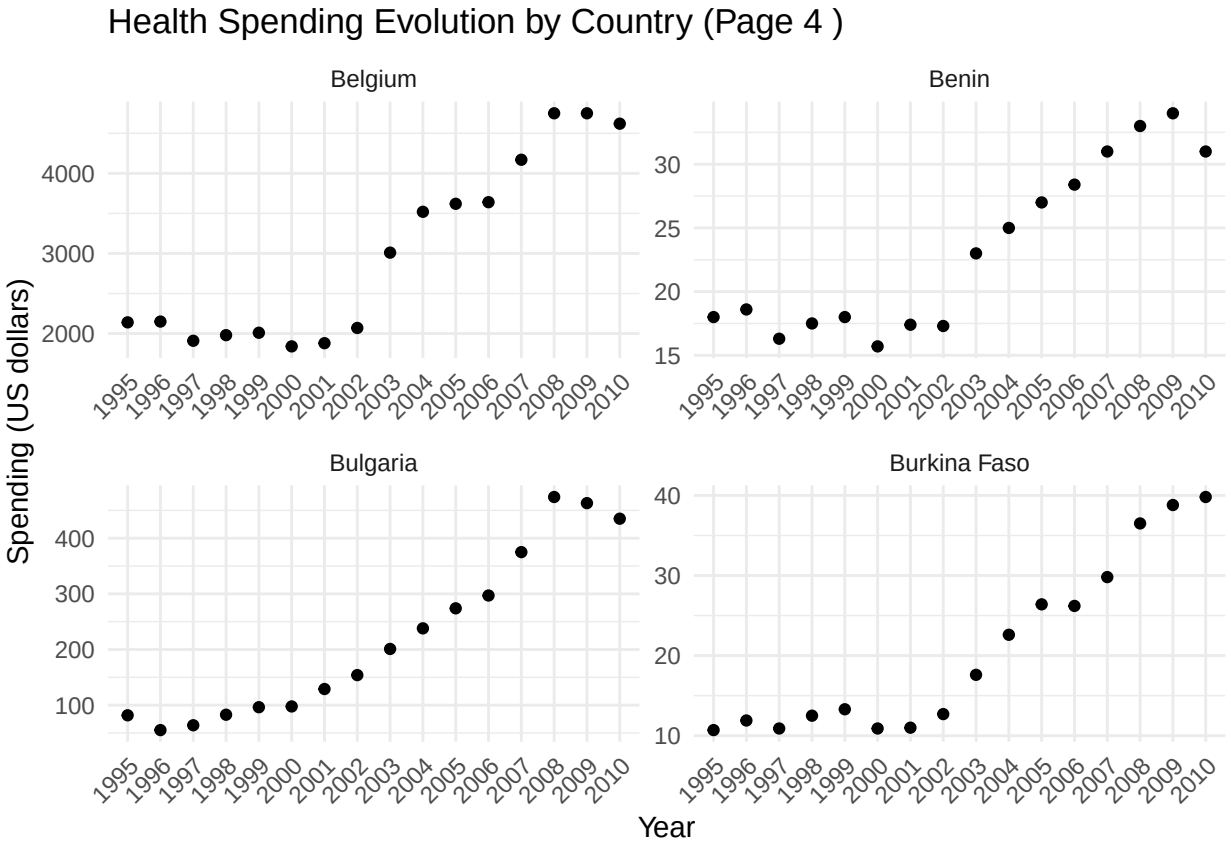
Now, I will plot the evolution of health spending by country from 1995 to 2010.

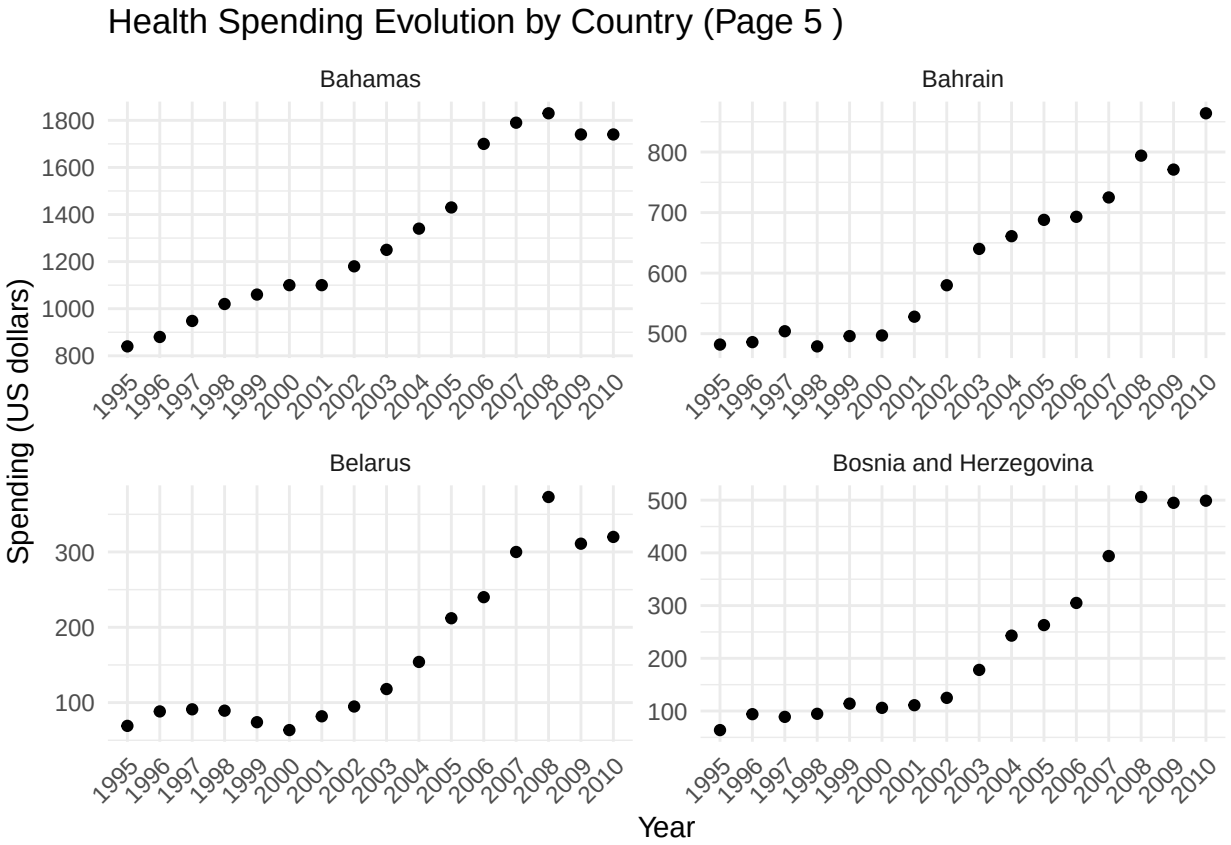
```r
# Assuming health_long is your data frame
countries <- unique(health_long$country)
pages <- ceiling(length(countries) / 4)  # ceiling(): Rounds up to the nearest whole number, ensuring t

country_plots <- list()

for (page in 1:pages) {
  start_index <- 1 + 4 * (page - 1)
  end_index <- min(4 * page, length(countries)) # start_index and end_index: Calculate the range of ind

  subset_countries <- countries[start_index:end_index]

  subset_data <- health_long %>%
    filter(country %in% subset_countries)

  plot1 <- ggplot(data = subset_data, aes(x = Year, y = Spending)) +
    geom_line() +
    geom_point() +
    facet_wrap(~ country, scales = "free") +
    labs(title = paste("Health Spending Evolution by Country (Page", page, ")"),
         x = "Year", y = "Spending (US dollars)") +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, hjust = 1))

  country_plots[[page]] <- plot1
}

# Convert plots to grobs
plot_grobs1 <- lapply(country_plots, function(plot1) ggplot2::ggplotGrob(plot1))

# Arrange plots into pages
plot_pages_arranged1 <- wrap_plots(plot_grobs1, ncol = 2)  # 2 plots per row

# Print each page
for (i in seq_along(plot_pages_arranged1)) {
  print(plot_pages_arranged1[[i]])
}
```

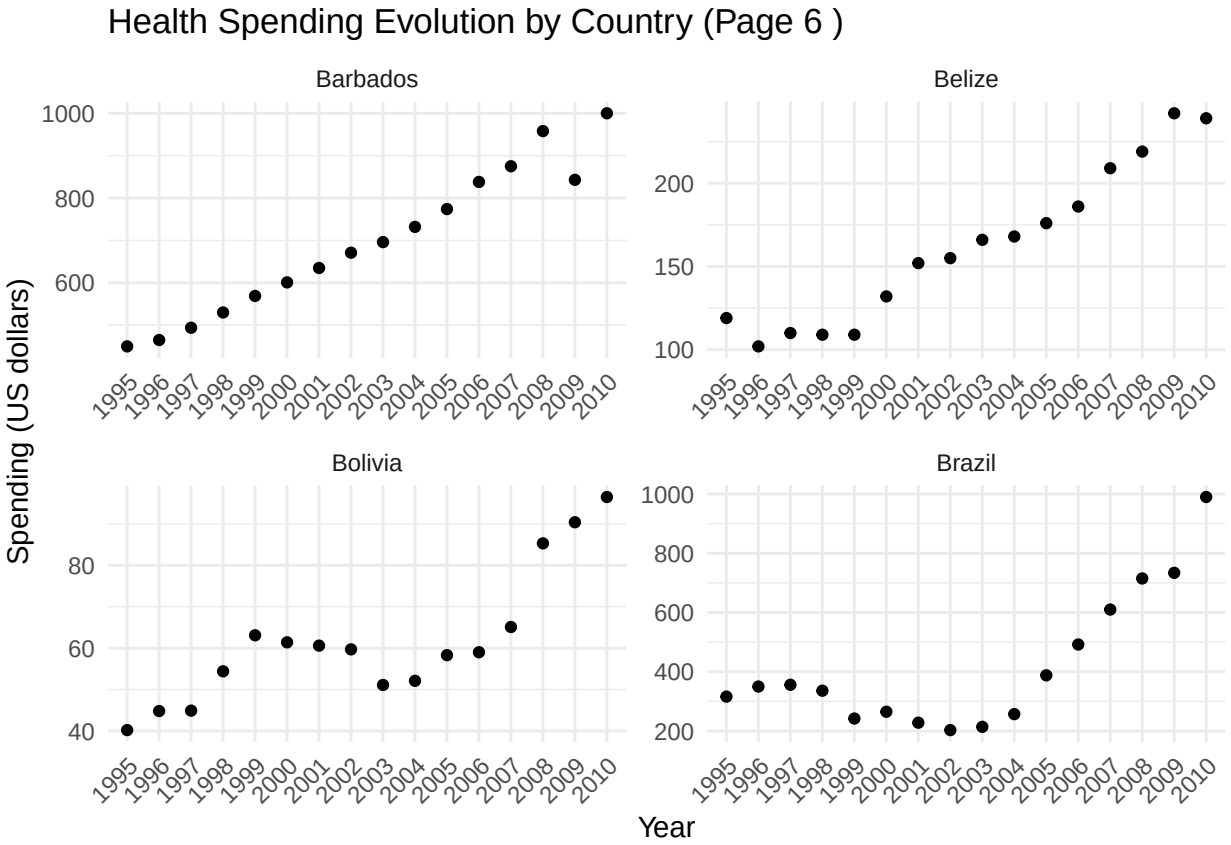

Health Spending Evolution by Country (Page 1)
Afghanistan
Albania
Andorra
Angola
Spending (US dollars)
Year

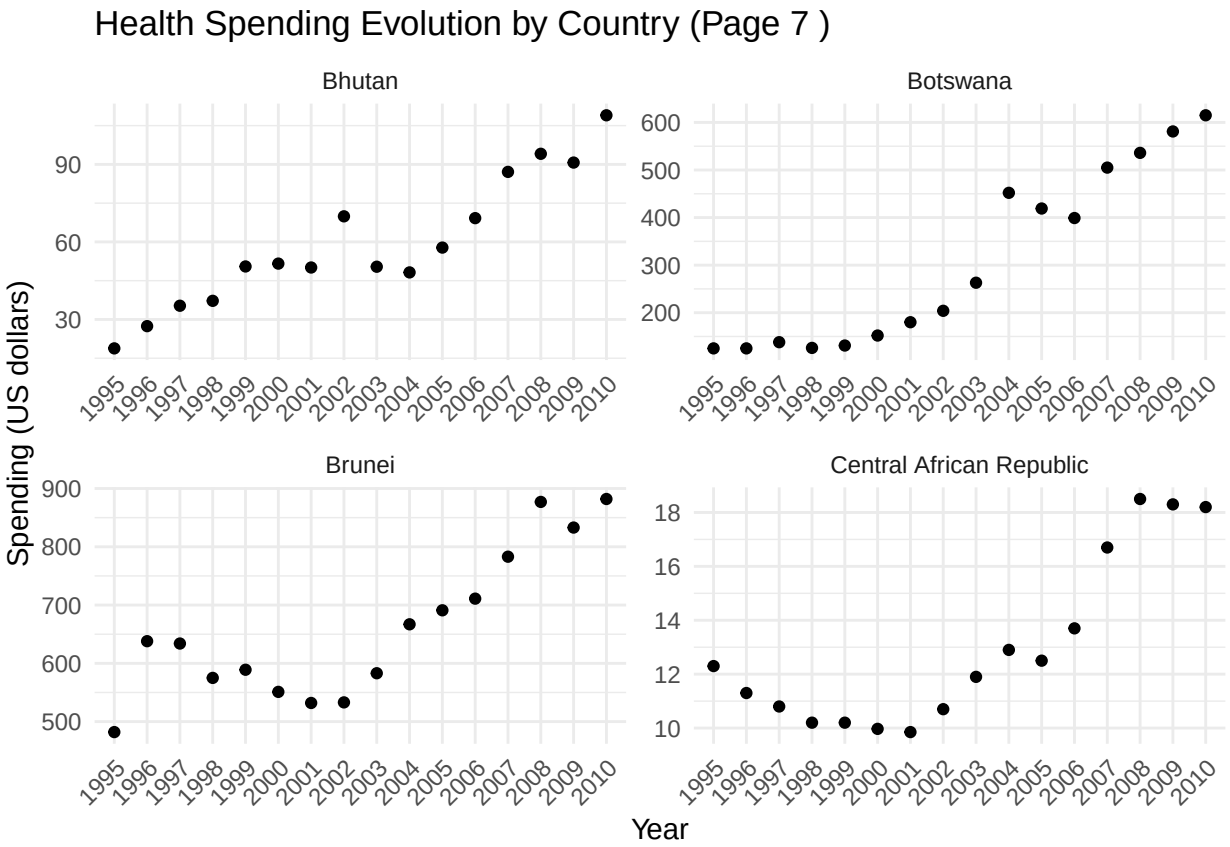

Health Spending Evolution by Country (Page 2)
Antigua and Barbuda
Argentina
Armenia
UAE
Spending (US dollars)
Year
700
600
500
400
300
200
100
50
1500
1250
1000
750
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

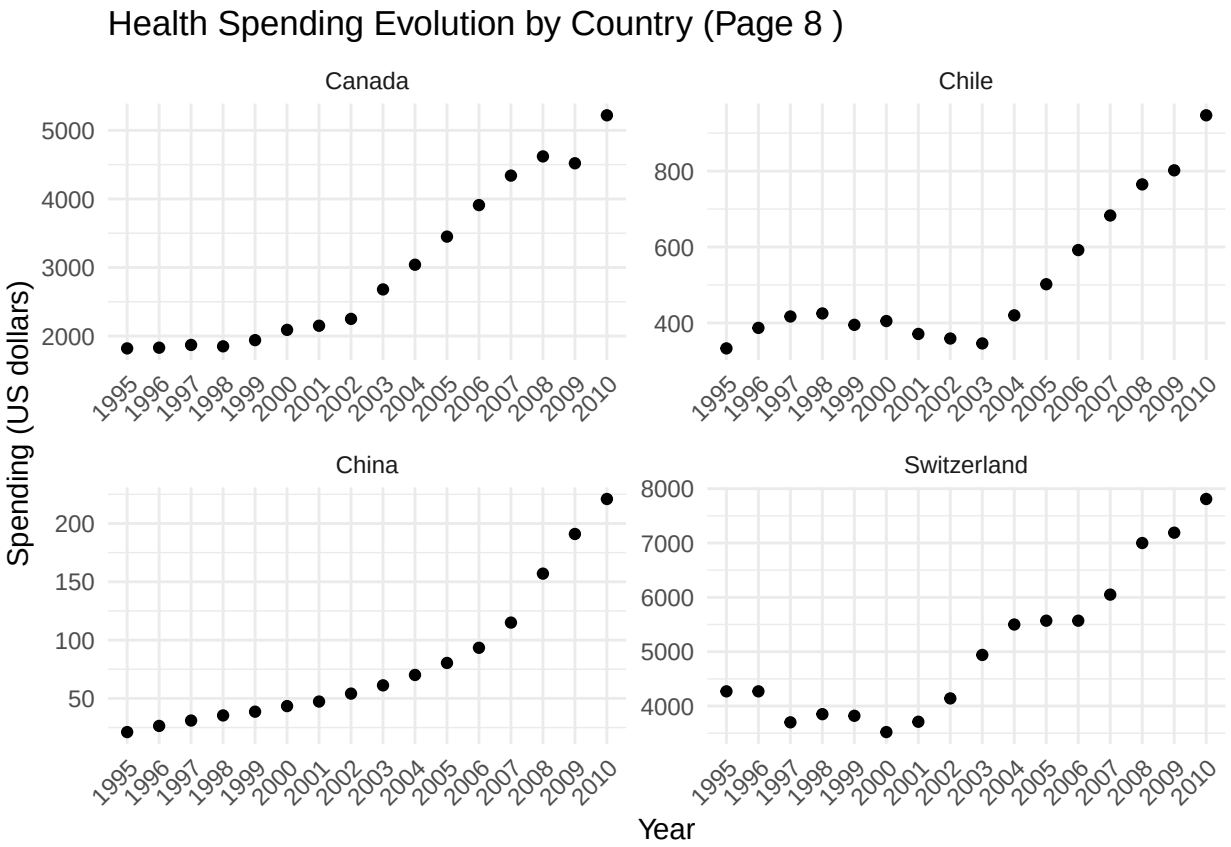

Health Spending Evolution by Country (Page 3)
Australia
Austria
Azerbaijan
Burundi
Spending (US dollars)
Year
2000
3000
4000
5000
100
200
300
10
15
20
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

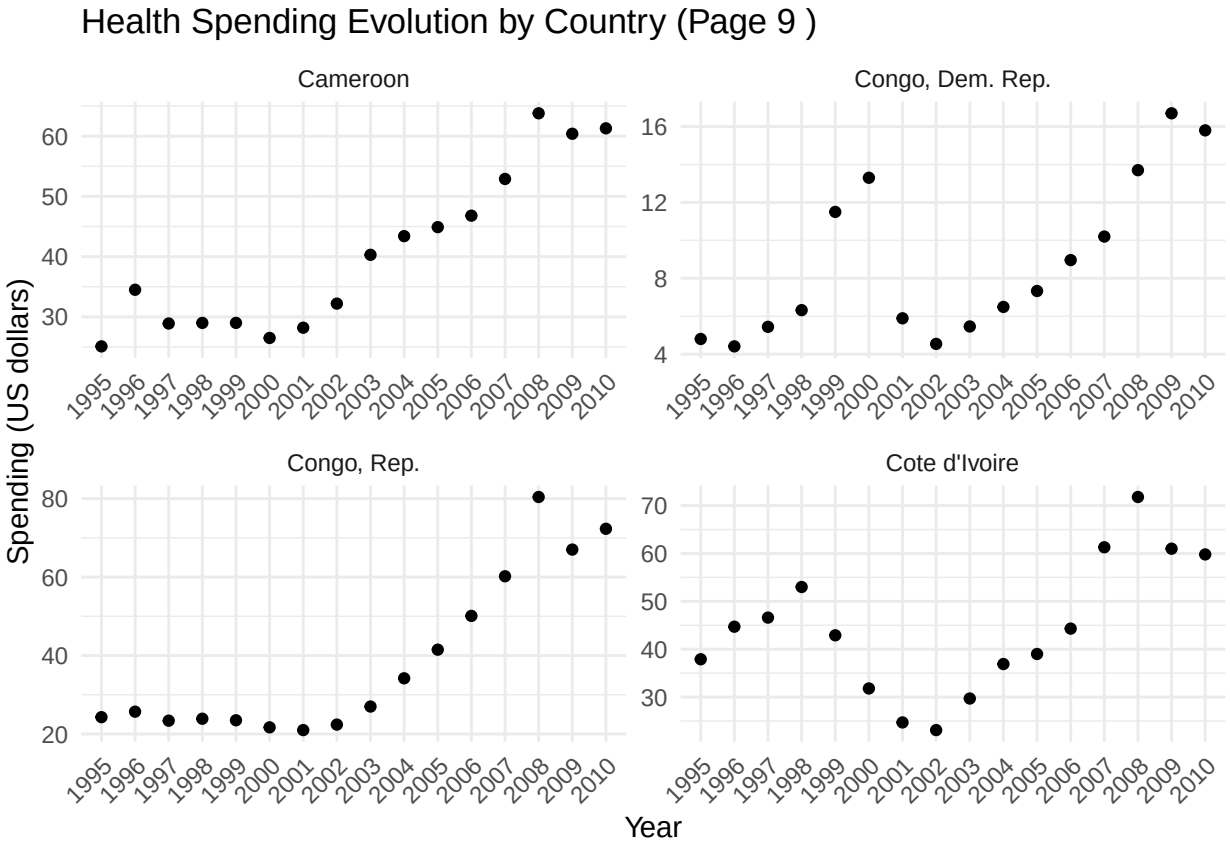

Health Spending Evolution by Country (Page 4)
Belgium
Benin
Bulgaria
Burkina Faso
Spending (US dollars)
Year
2000
3000
4000
15
20
25
30
100
200
300
400
10
20
30
40
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

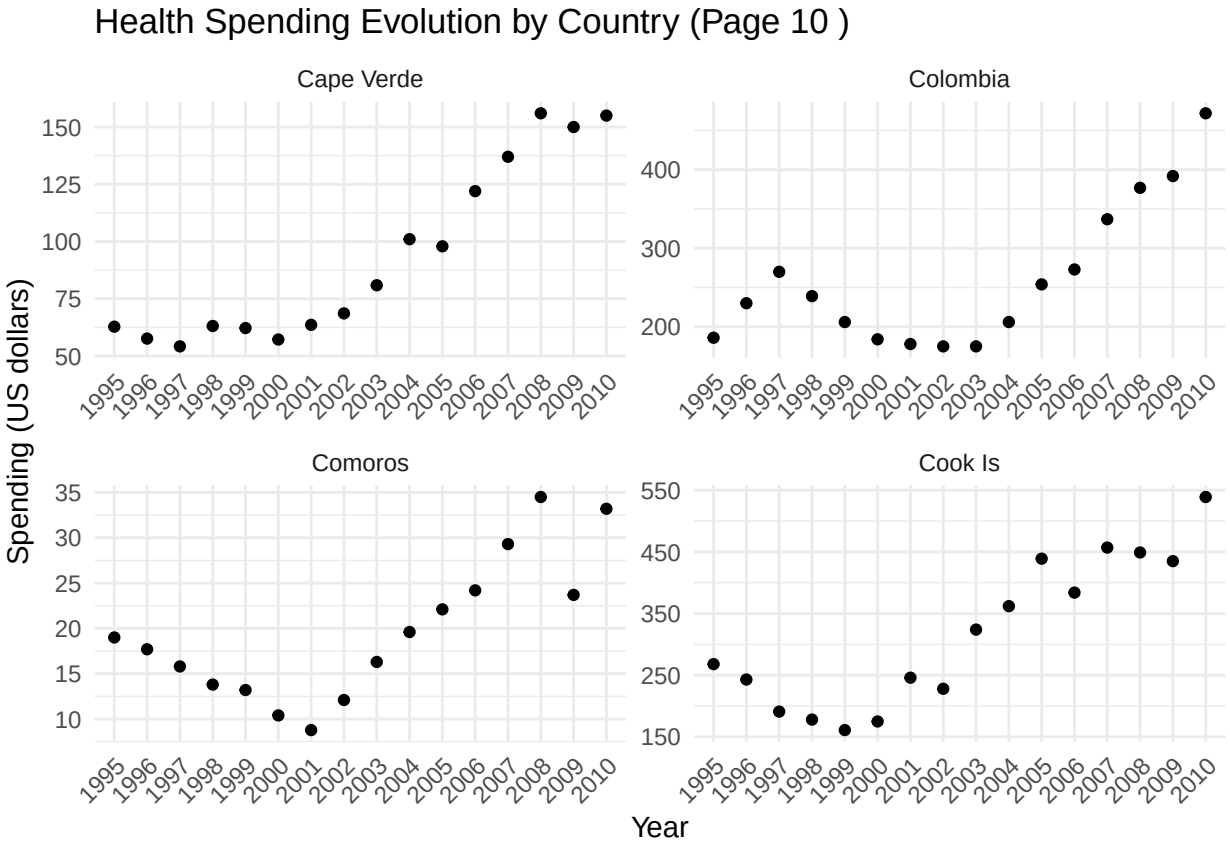

Health Spending Evolution by Country (Page 5)
Bahamas
Bahrain
Belarus
Bosnia and Herzegovina
Spending (US dollars)
Year

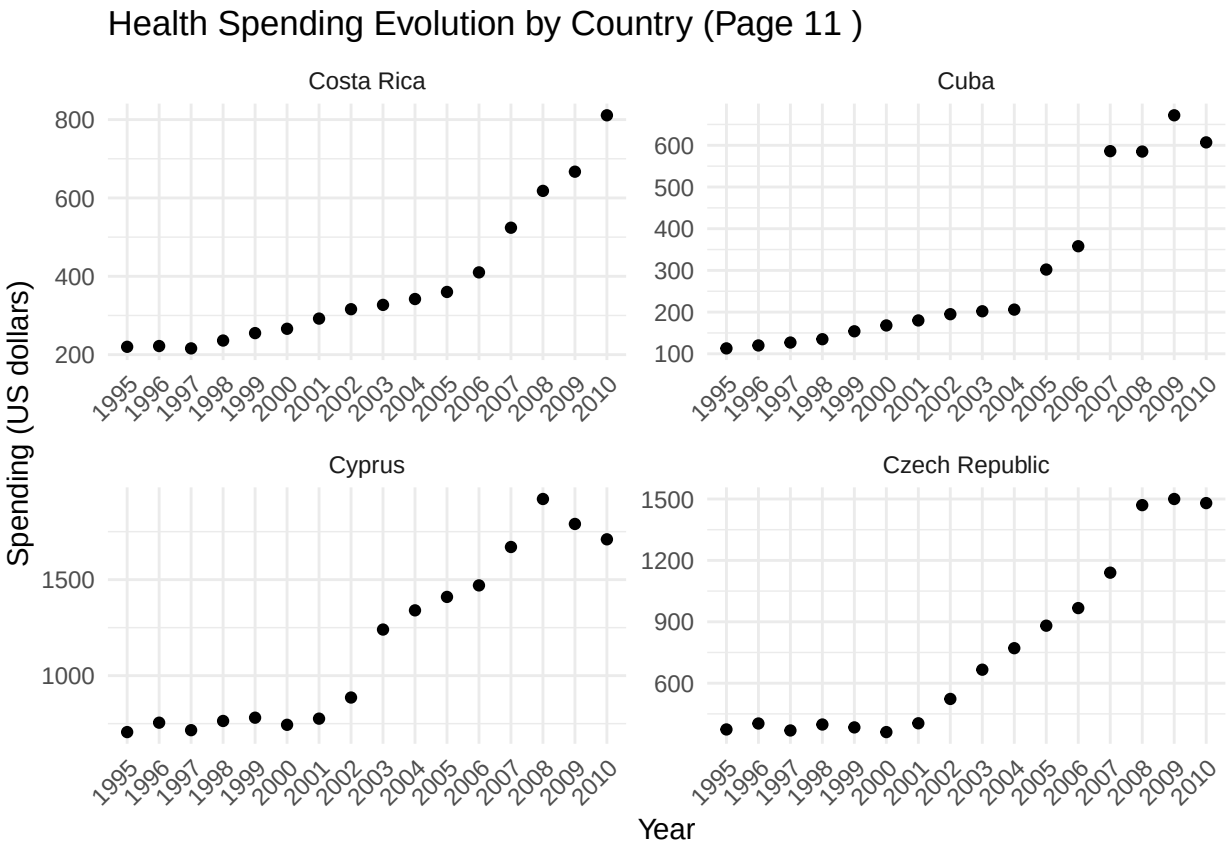

Health Spending Evolution by Country (Page 6)
Barbados
Belize
Bolivia
Brazil
Spending (US dollars)
Year

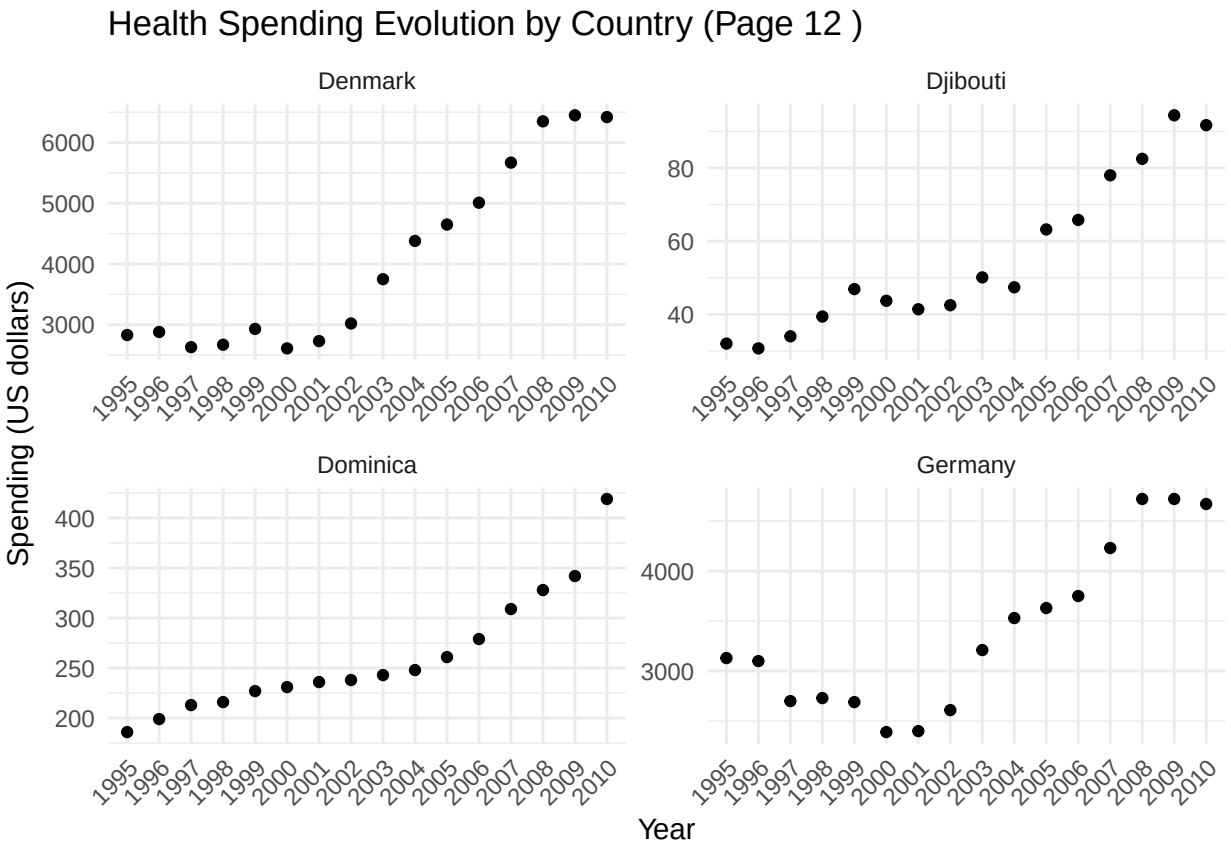

Health Spending Evolution by Country (Page 7)
Bhutan
Botswana
Brunei
Central African Republic
Spending (US dollars)
Year

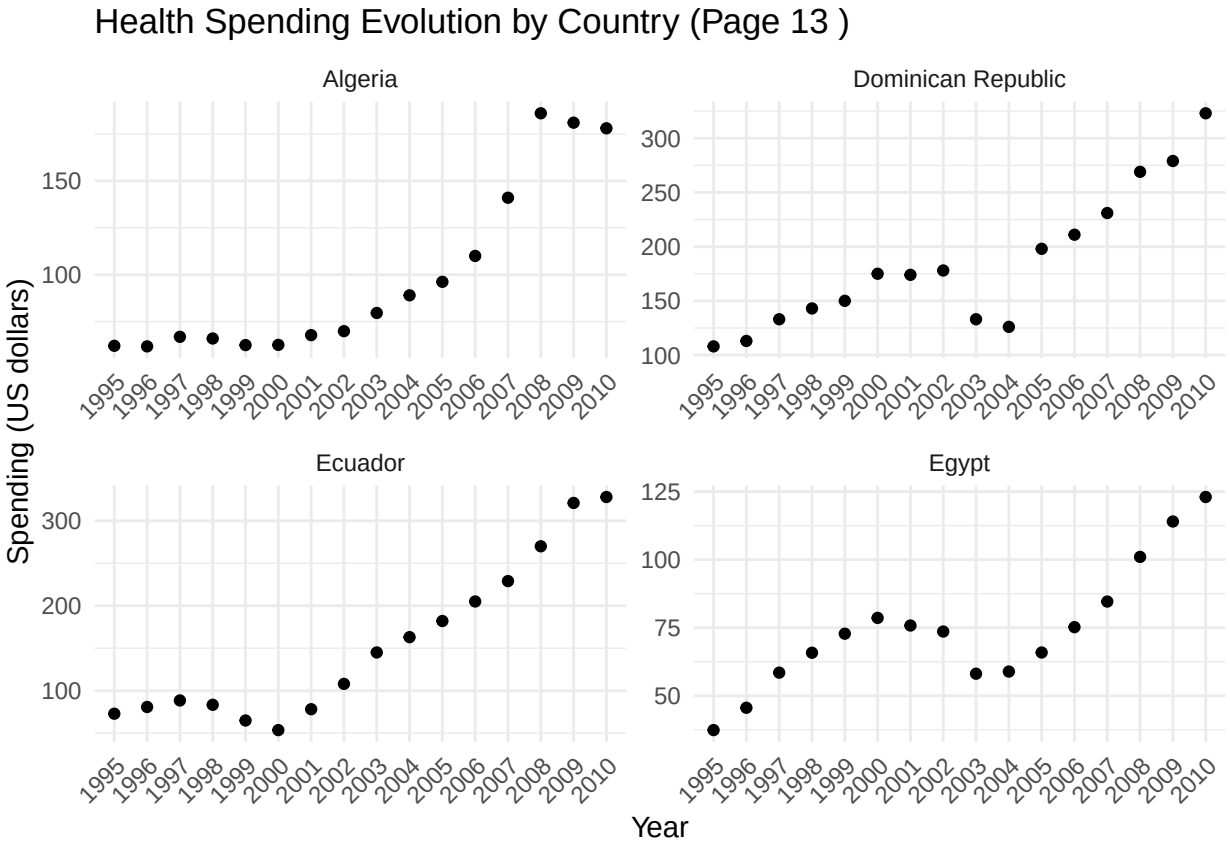

Health Spending Evolution by Country (Page 8)
Canada
Chile
China
Switzerland
Spending (US dollars)
Year

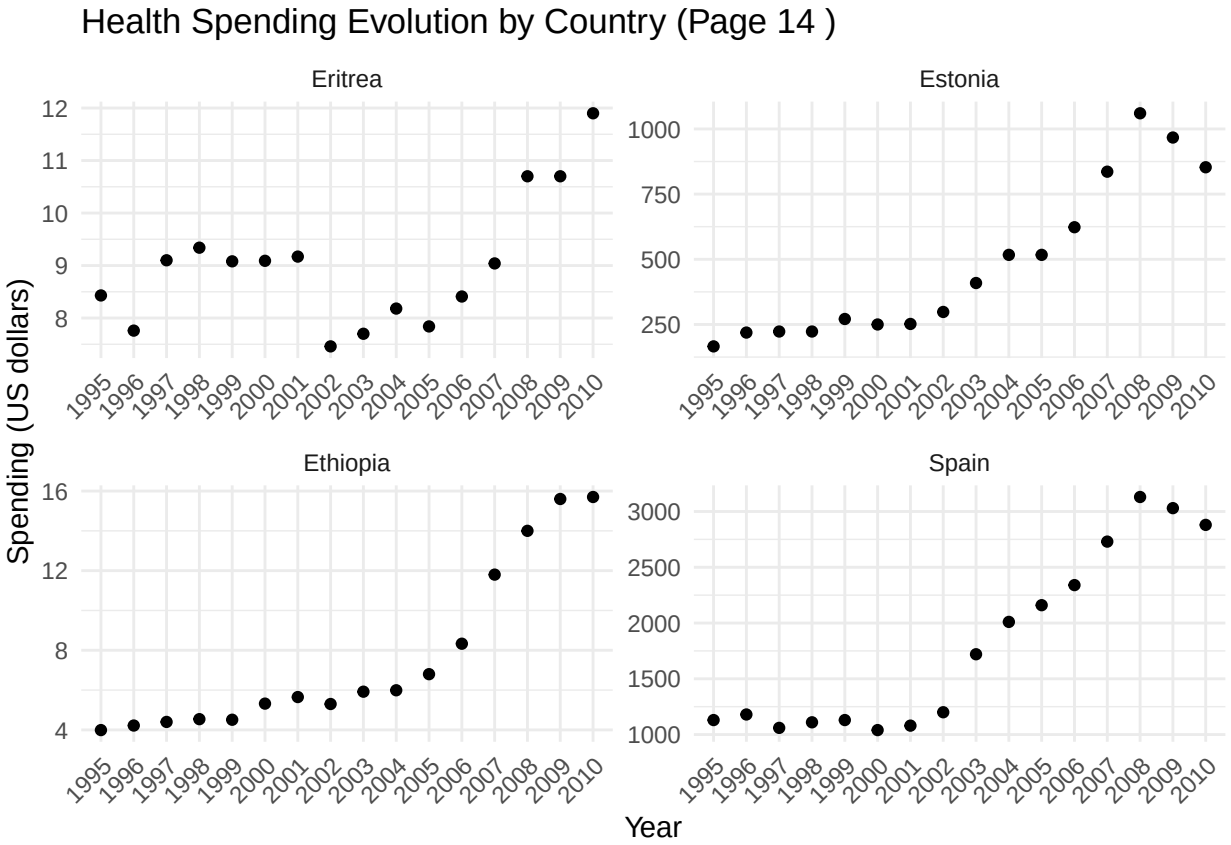

Health Spending Evolution by Country (Page 9)
Cameroon
Congo, Dem. Rep.
Congo, Rep.
Cote d'Ivoire
Spending (US dollars)
Year

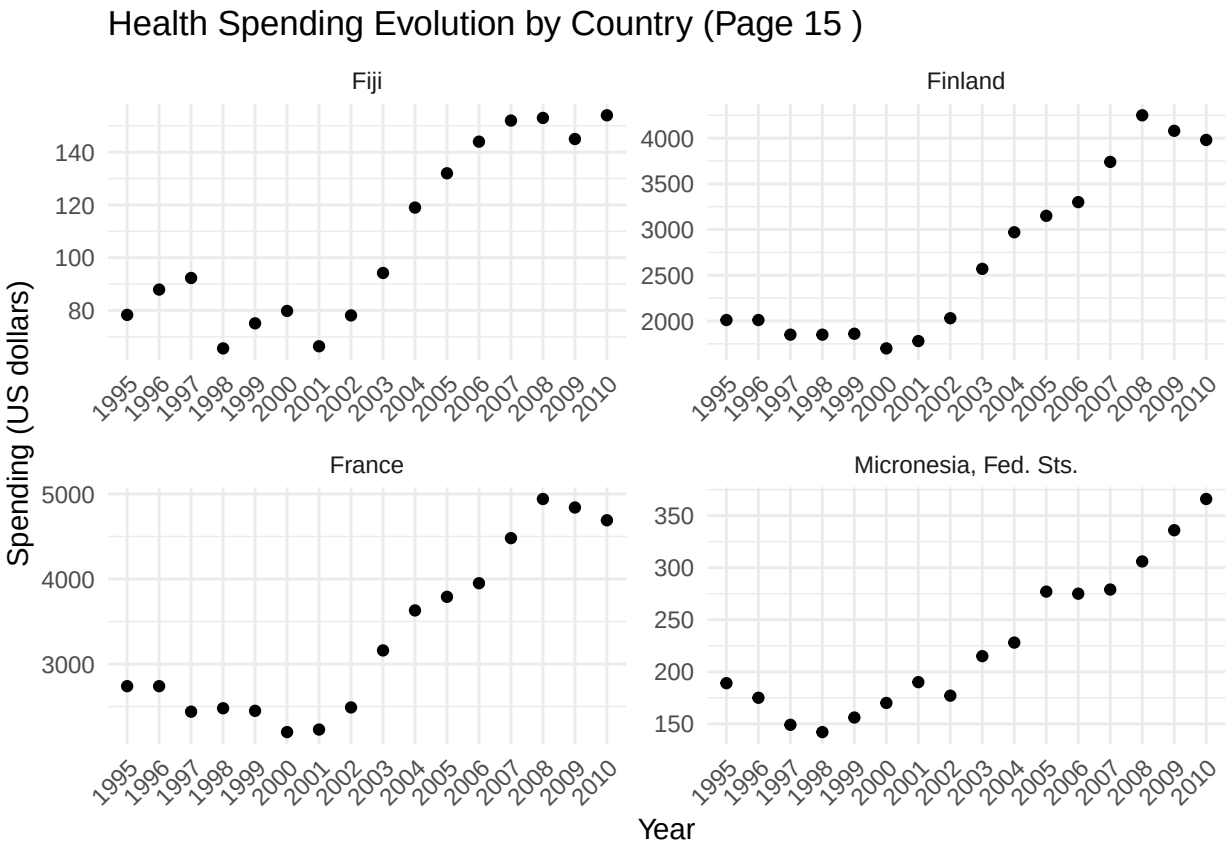

Health Spending Evolution by Country (Page 10)
Cape Verde
Colombia
Comoros
Cook Is
Spending (US dollars)
Year

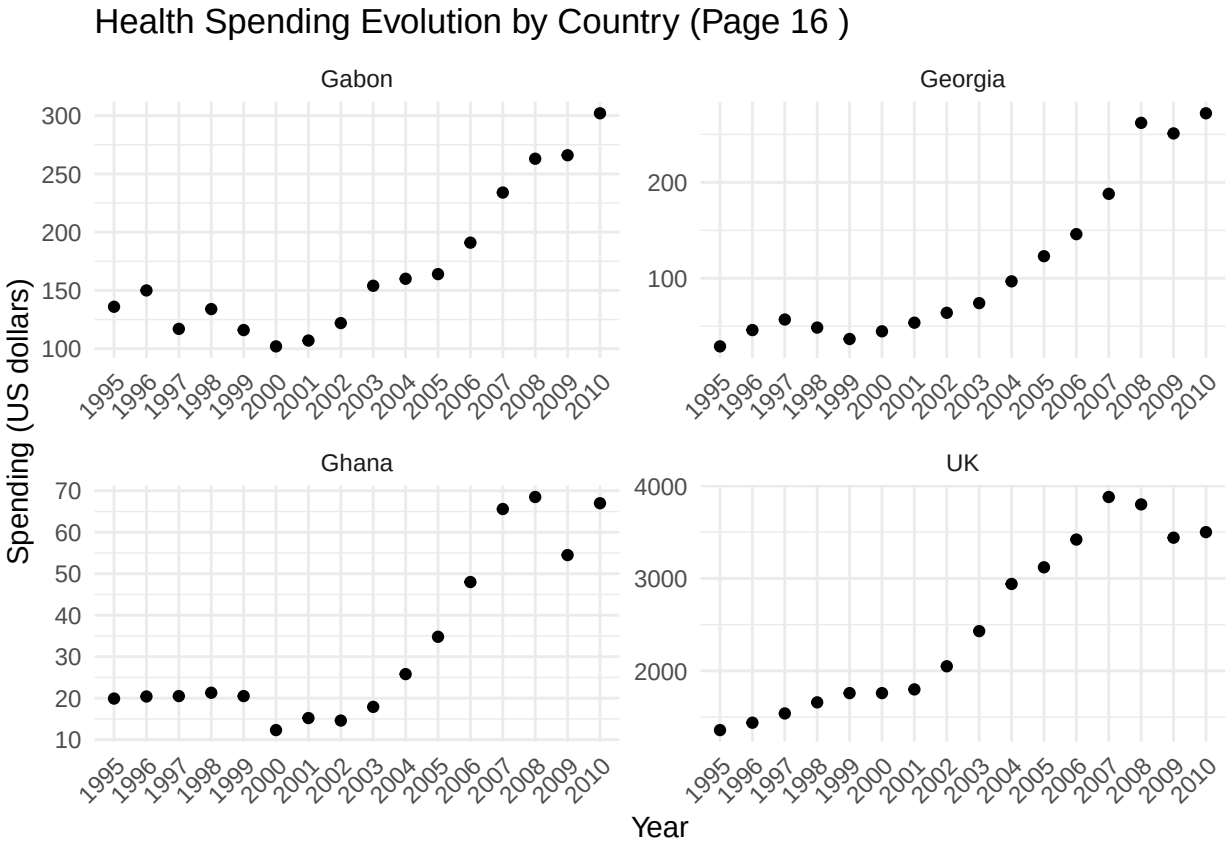

Health Spending Evolution by Country (Page 11)
Costa Rica
Cuba
Cyprus
Czech Republic
Spending (US dollars)
Year

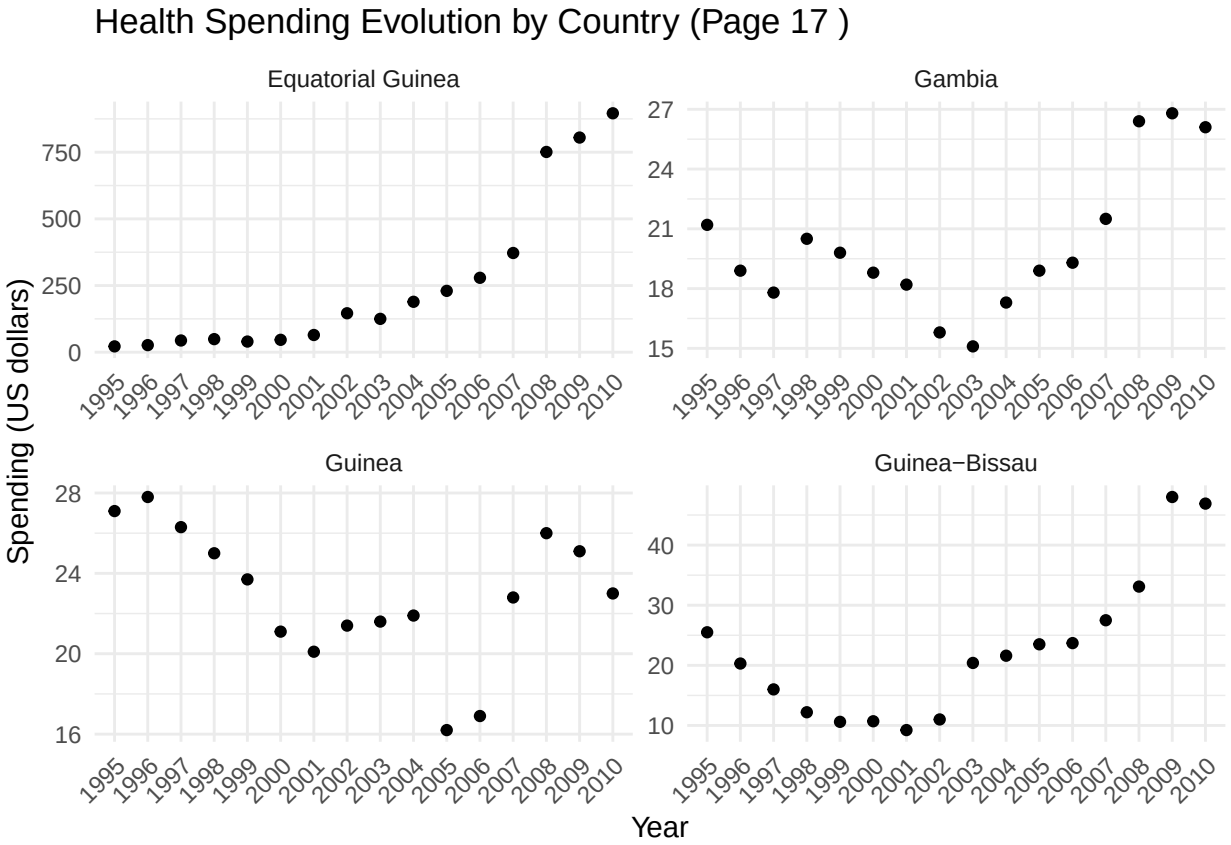

Health Spending Evolution by Country (Page 12)
Denmark
Djibouti
Dominica
Germany
Spending (US dollars)
Year
6000
5000
4000
3000
80
60
40
400
350
300
250
200
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

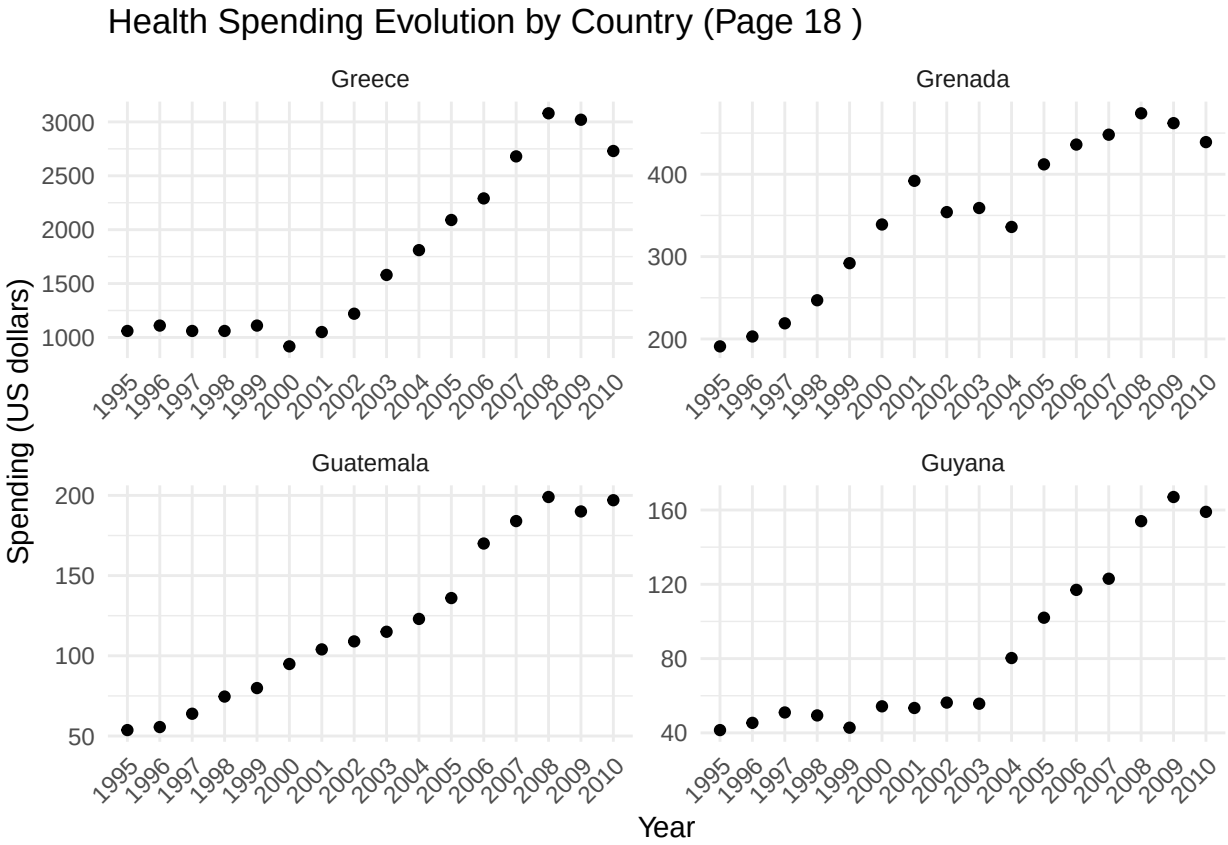

Health Spending Evolution by Country (Page 13)
Algeria
Dominican Republic
Ecuador
Egypt
Spending (US dollars)
Year

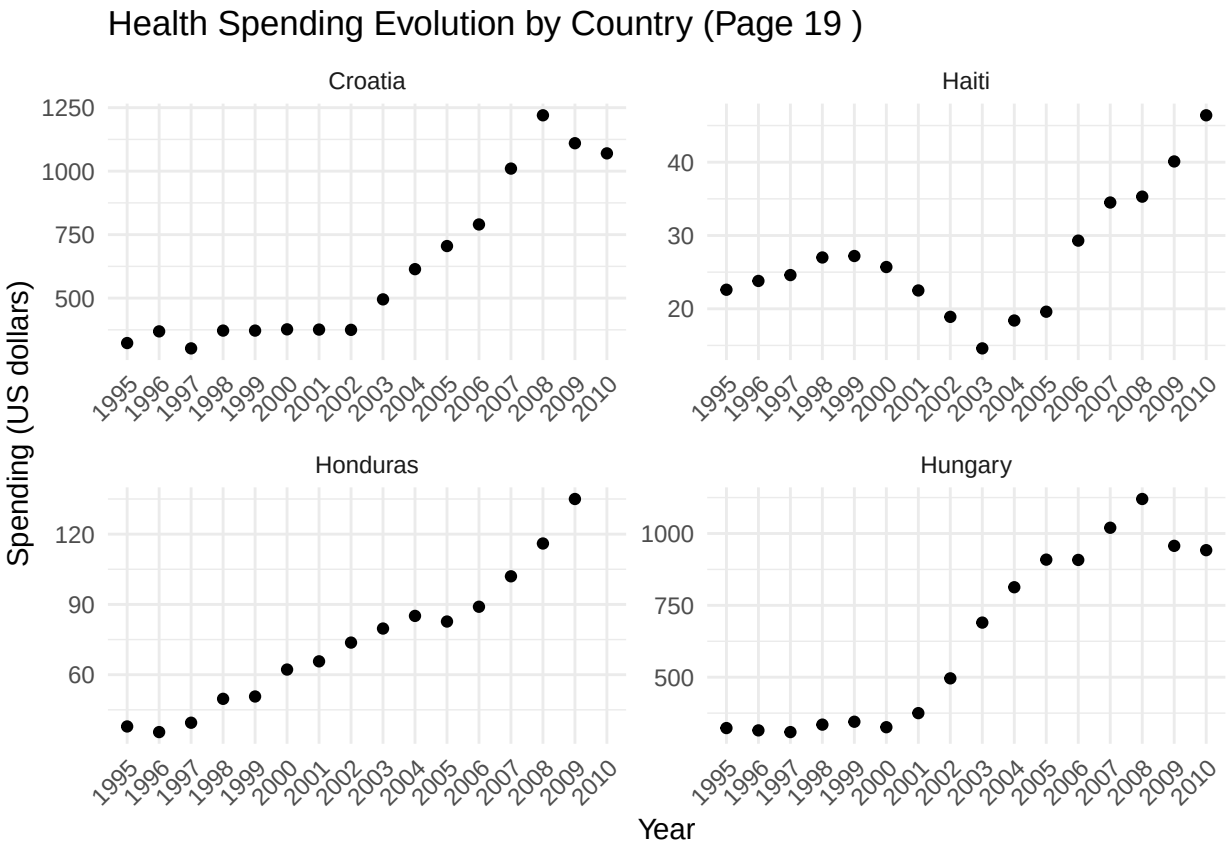

Health Spending Evolution by Country (Page 14)
Eritrea
Estonia
Ethiopia
Spain
Spending (US dollars)
Year

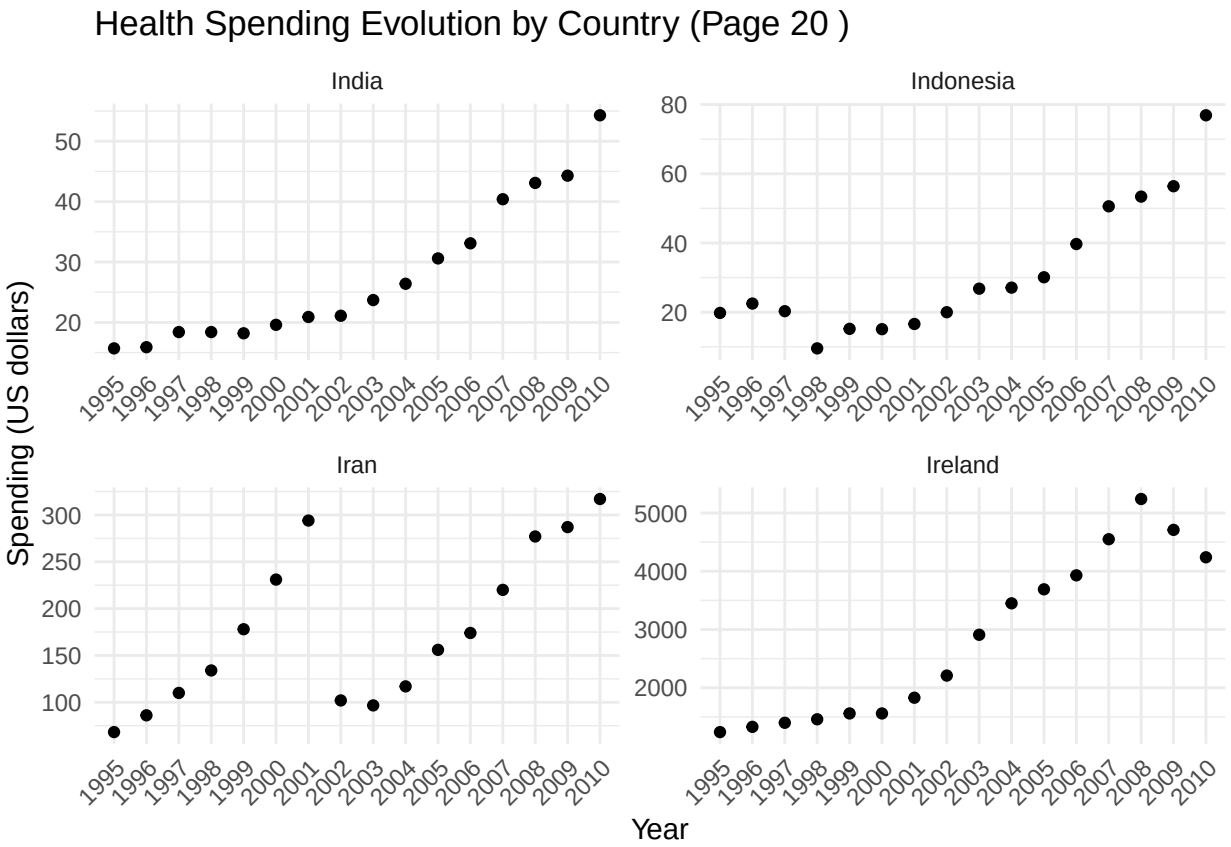

Health Spending Evolution by Country (Page 15)
Fiji
Finland
France
Micronesia, Fed. Sts.
Spending (US dollars)
Year

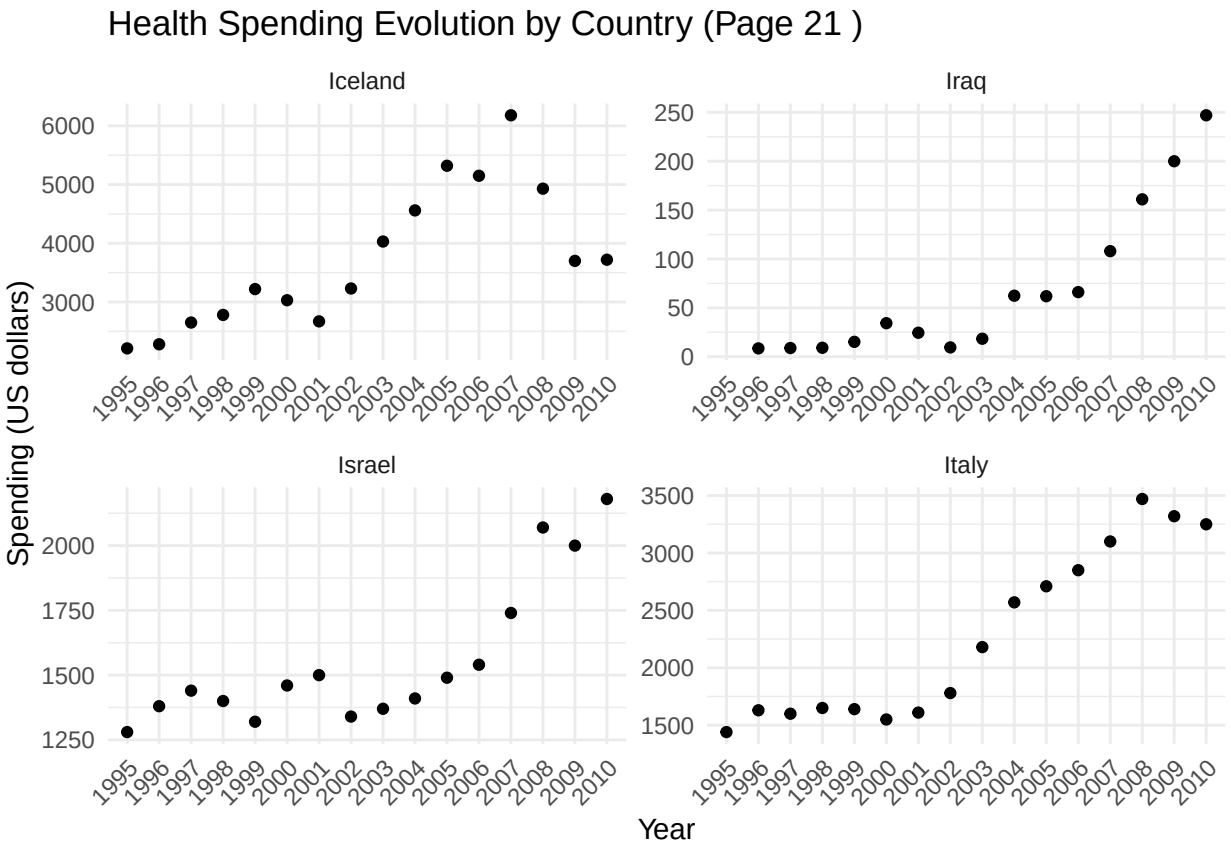

Health Spending Evolution by Country (Page 16)
Gabon
Georgia
Ghana
UK
Spending (US dollars)
Year

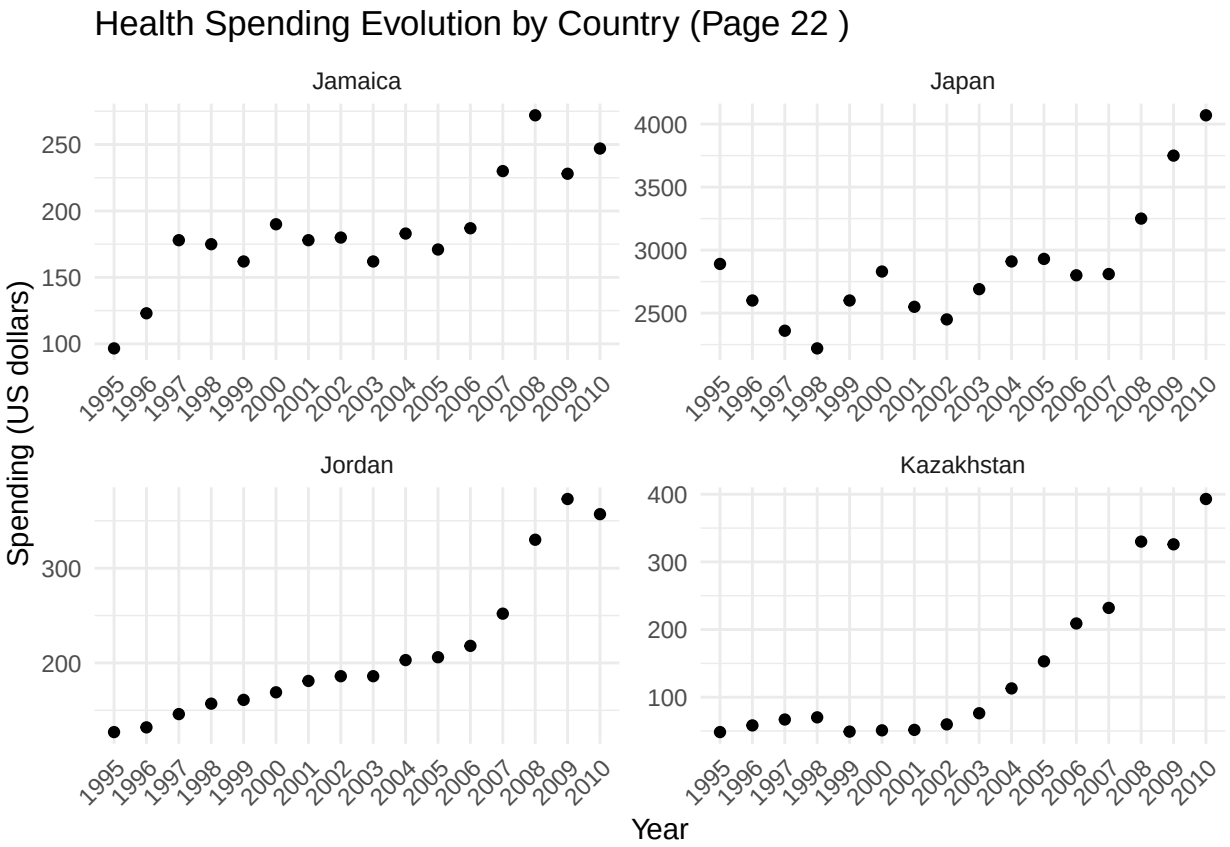

Health Spending Evolution by Country (Page 17)
Equatorial Guinea
Gambia
Guinea
Guinea−Bissau
Spending (US dollars)
Year

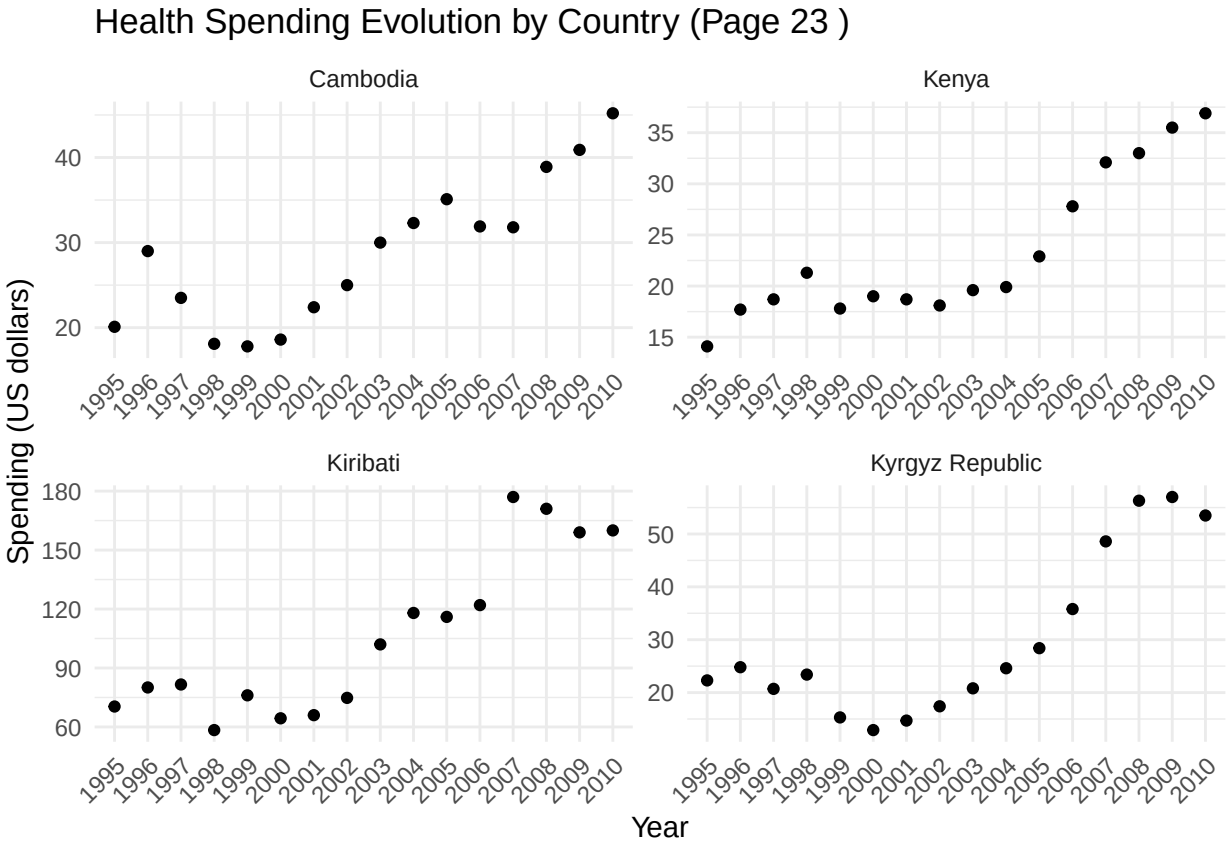

Health Spending Evolution by Country (Page 18)
Greece
Grenada
Guatemala
Guyana
Spending (US dollars)
Year

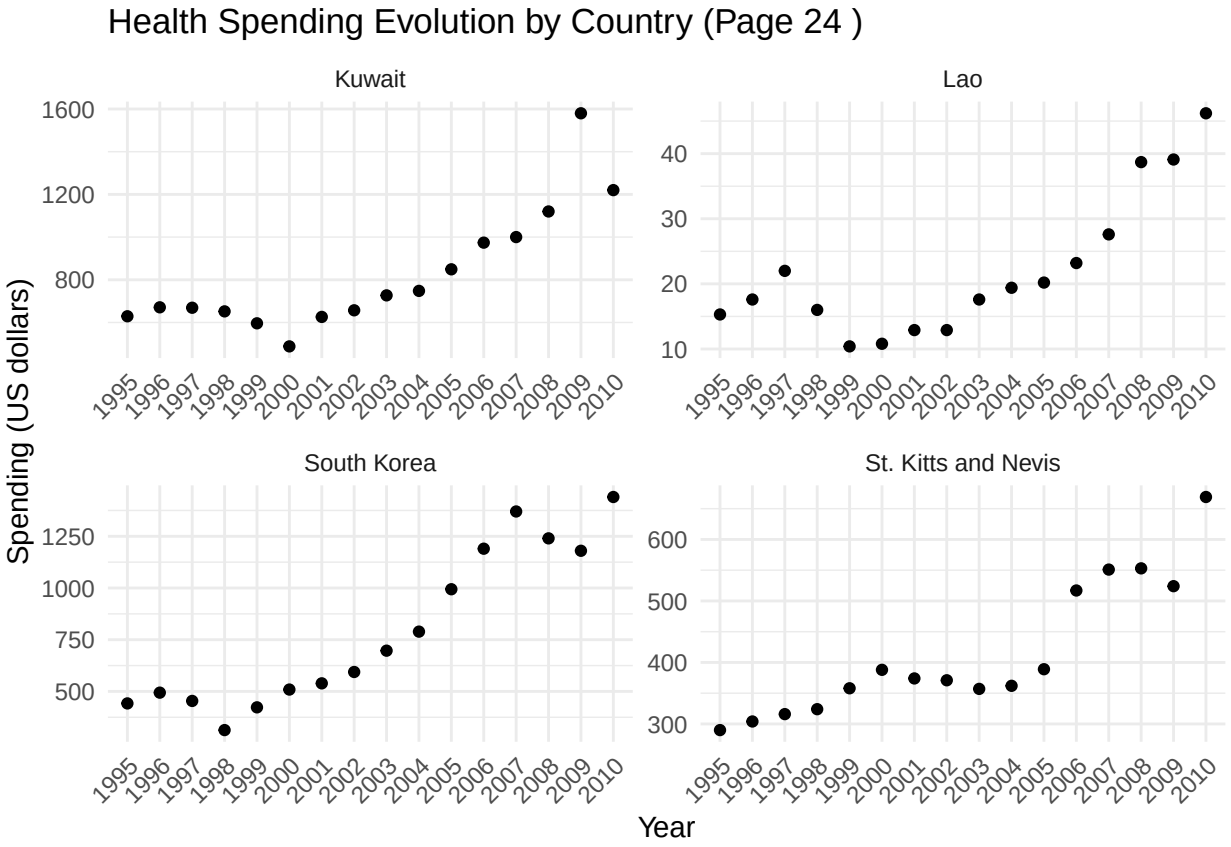

Health Spending Evolution by Country (Page 19)
Croatia
Haiti
Honduras
Hungary
Spending (US dollars)
Year
1250
1000
750
500
40
30
20
120
90
60
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

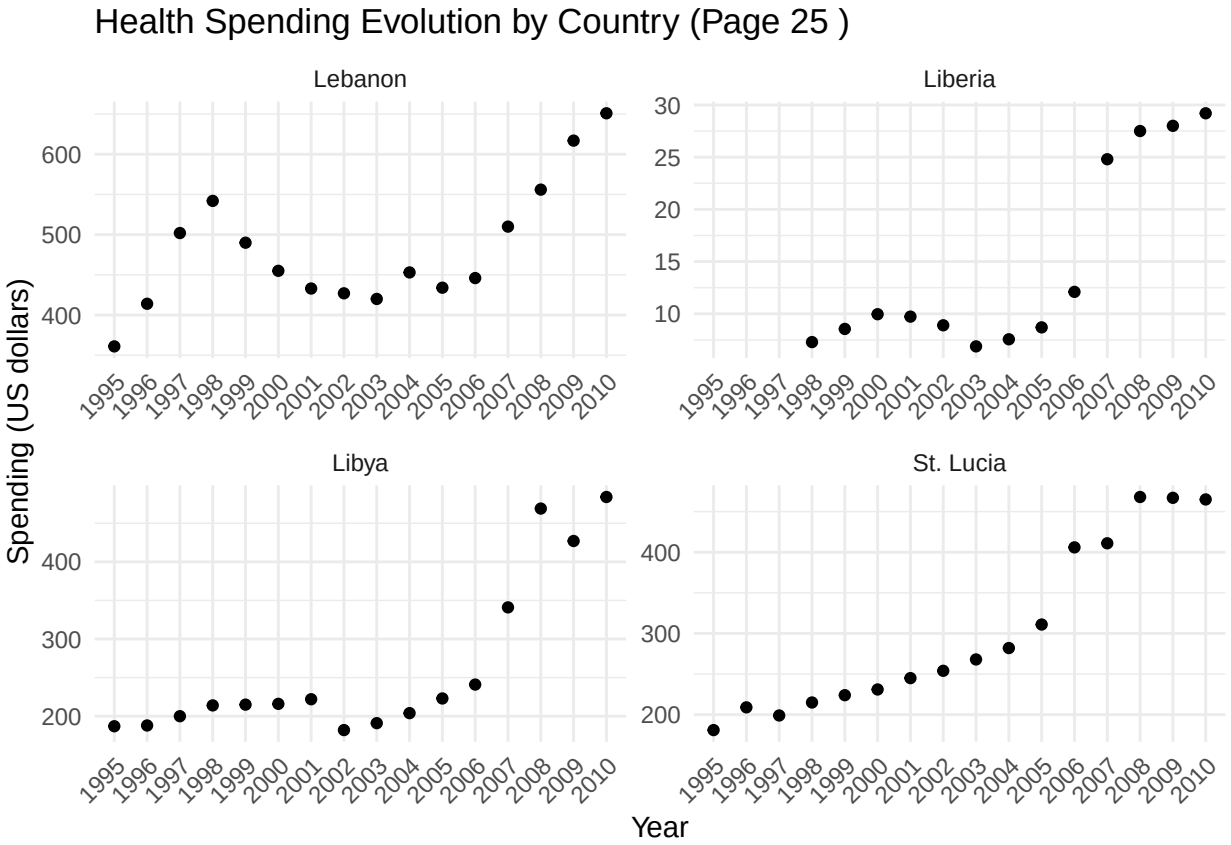

Health Spending Evolution by Country (Page 20)
India
Indonesia
Iran
Ireland
Spending (US dollars)
Year

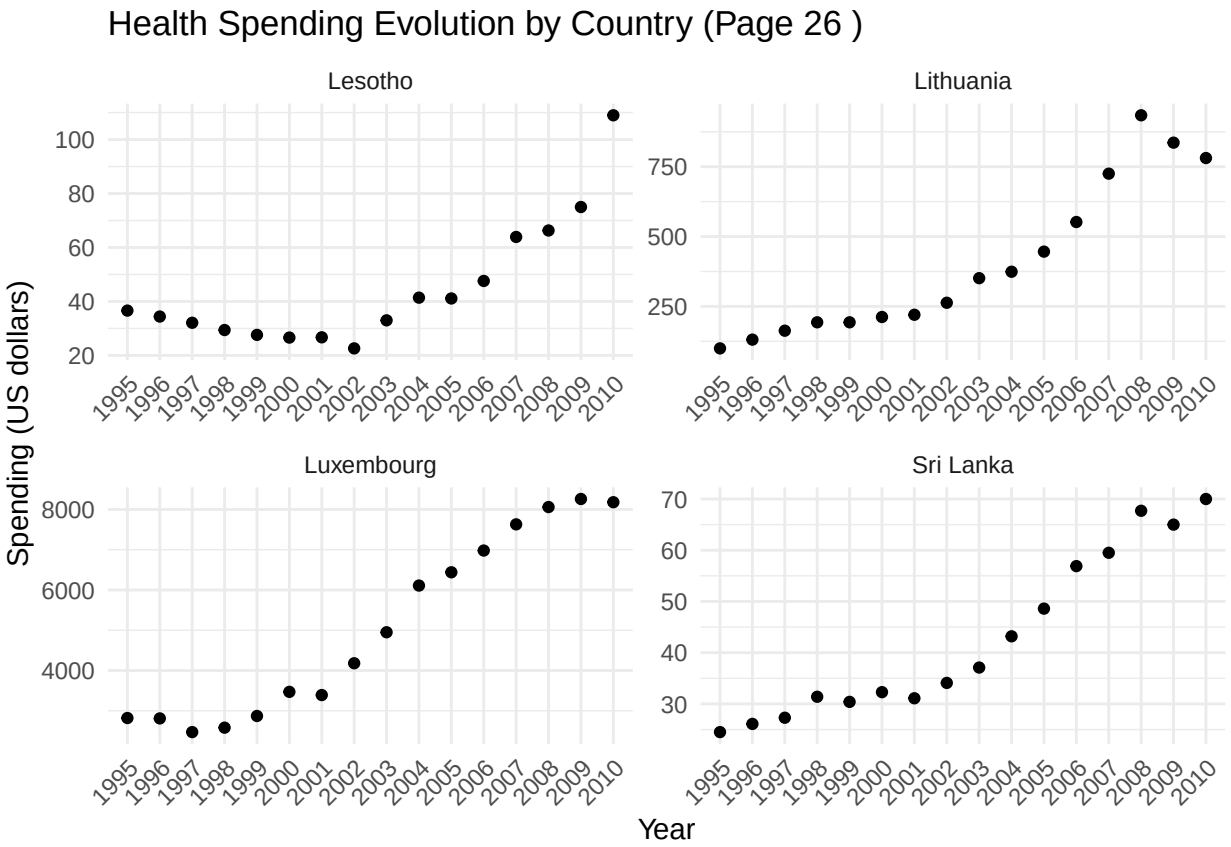

Health Spending Evolution by Country (Page 21)
Iceland
Iraq
Israel
Italy
Spending (US dollars)
Year

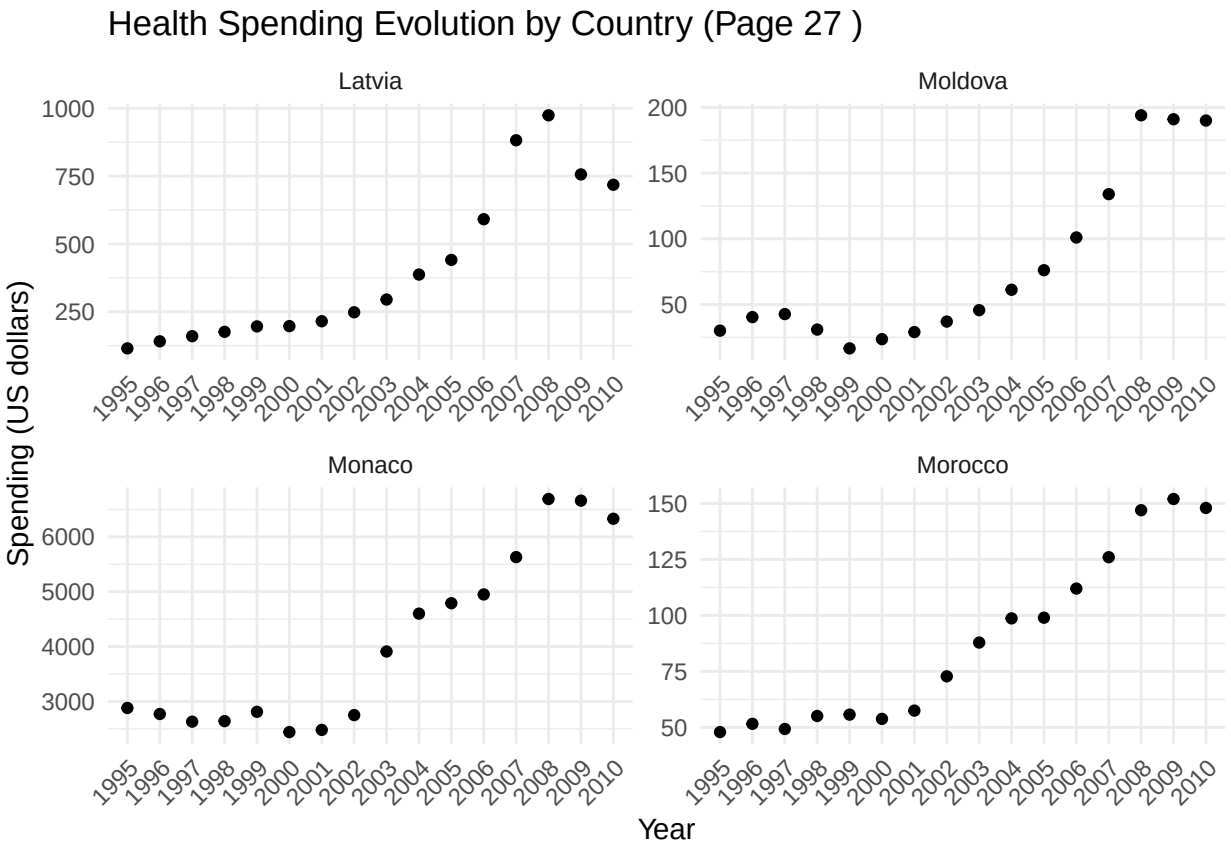

Health Spending Evolution by Country (Page 22)
Jamaica
Japan
Jordan
Kazakhstan
Spending (US dollars)
Year

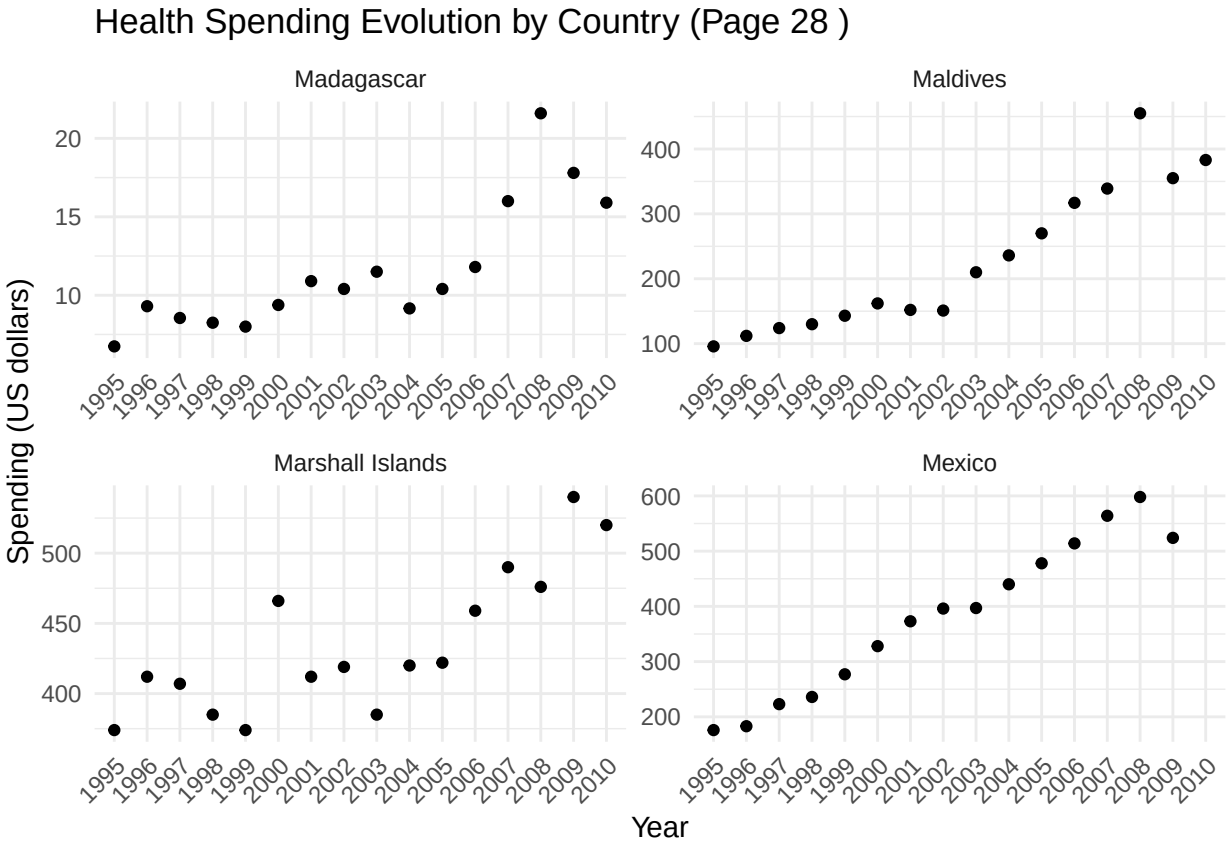

Health Spending Evolution by Country (Page 23)
Cambodia
Kenya
Kiribati
Kyrgyz Republic
Spending (US dollars)
Year

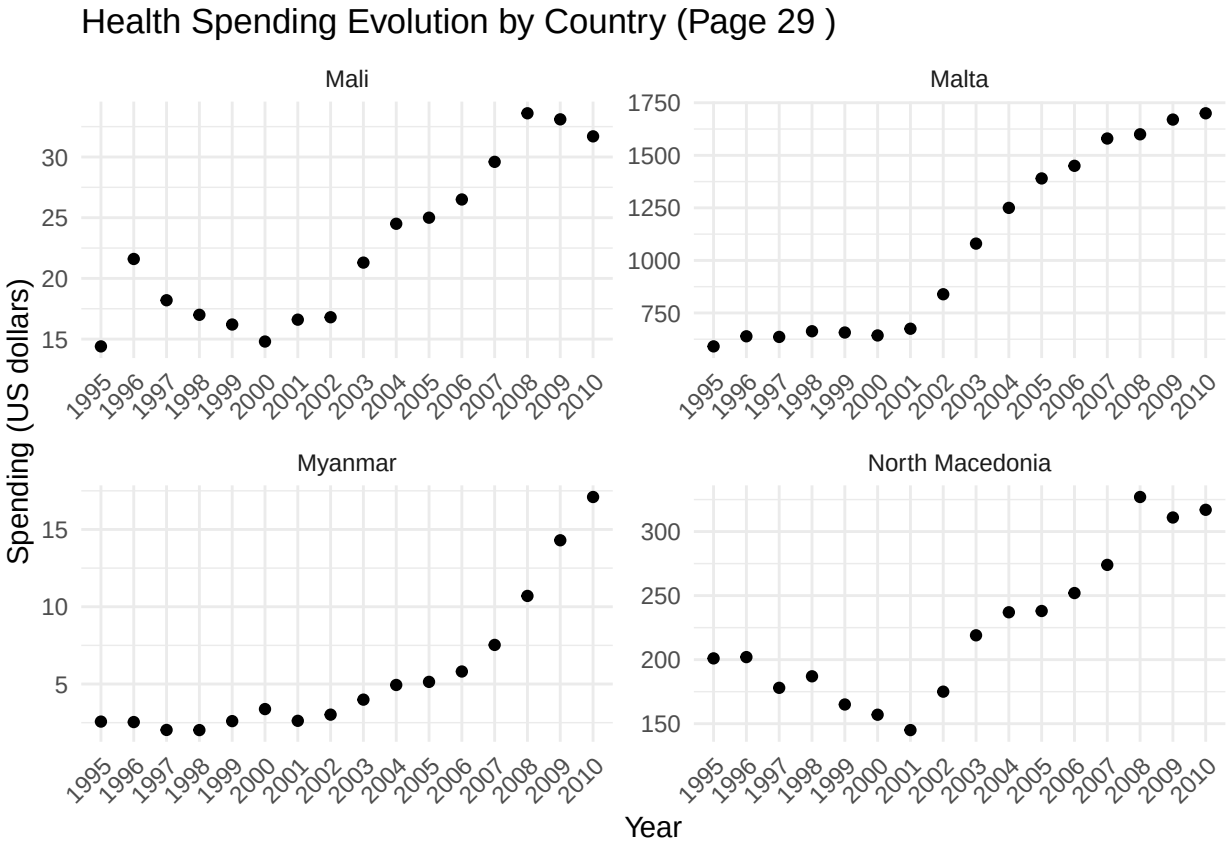

Health Spending Evolution by Country (Page 24)
Kuwait
Lao
South Korea
St. Kitts and Nevis
Spending (US dollars)
Year

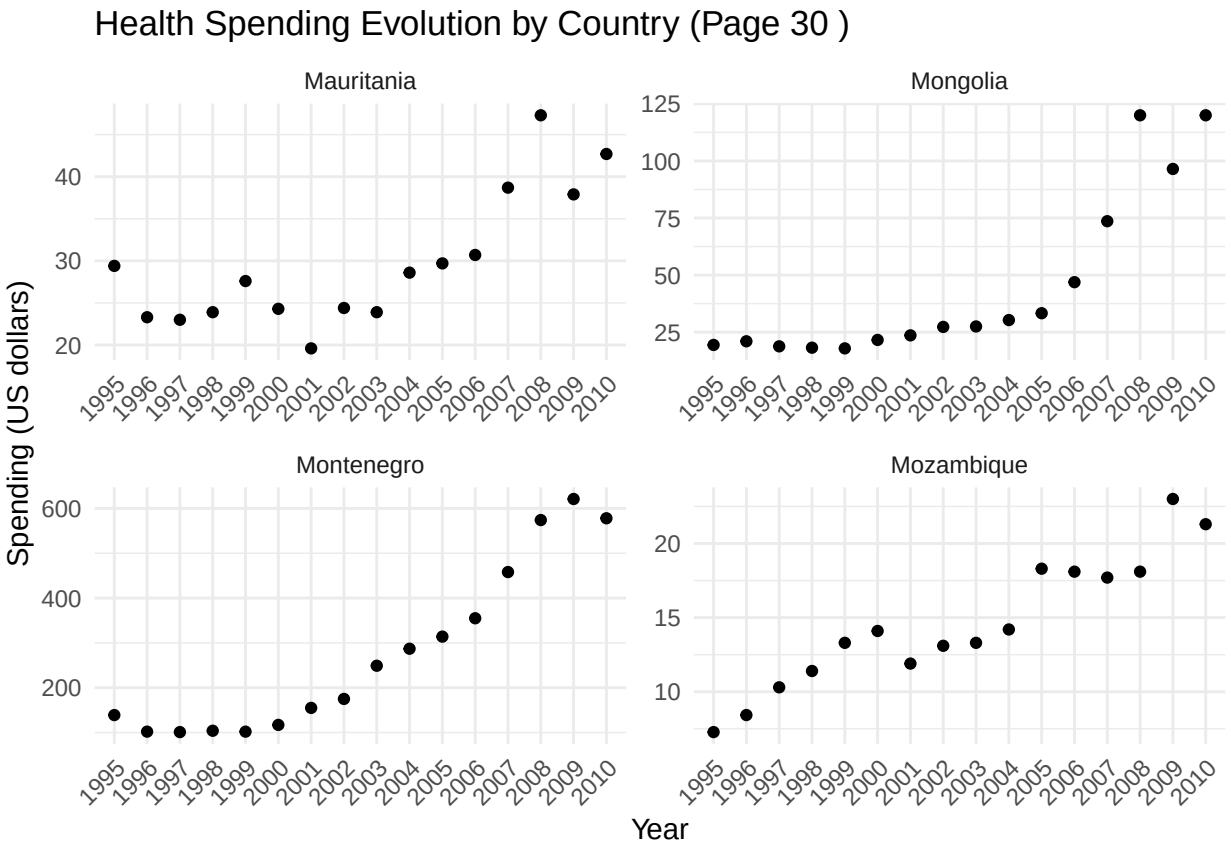

Health Spending Evolution by Country (Page 25)
Lebanon
Liberia
Libya
St. Lucia
Spending (US dollars)
Year

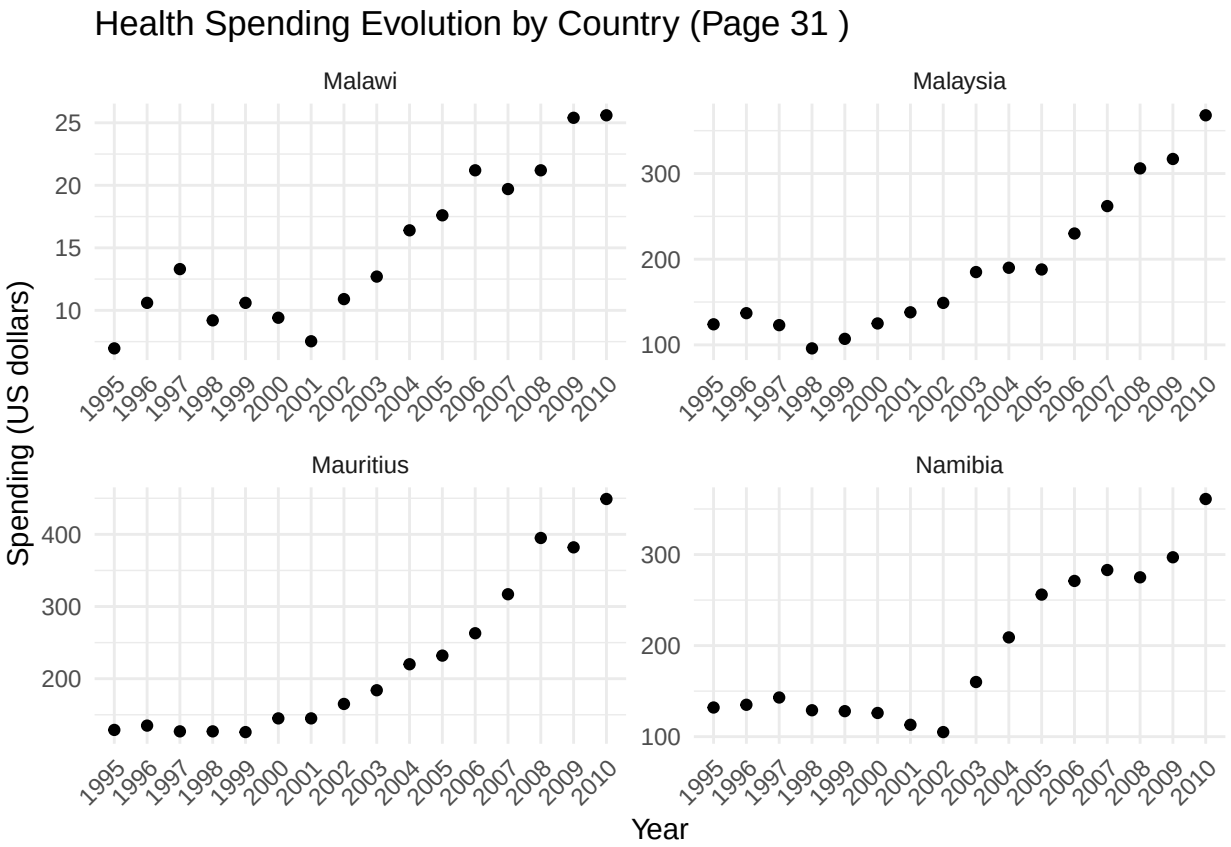

Health Spending Evolution by Country (Page 26)
Lesotho
Lithuania
Luxembourg
Sri Lanka
Spending (US dollars)
Year

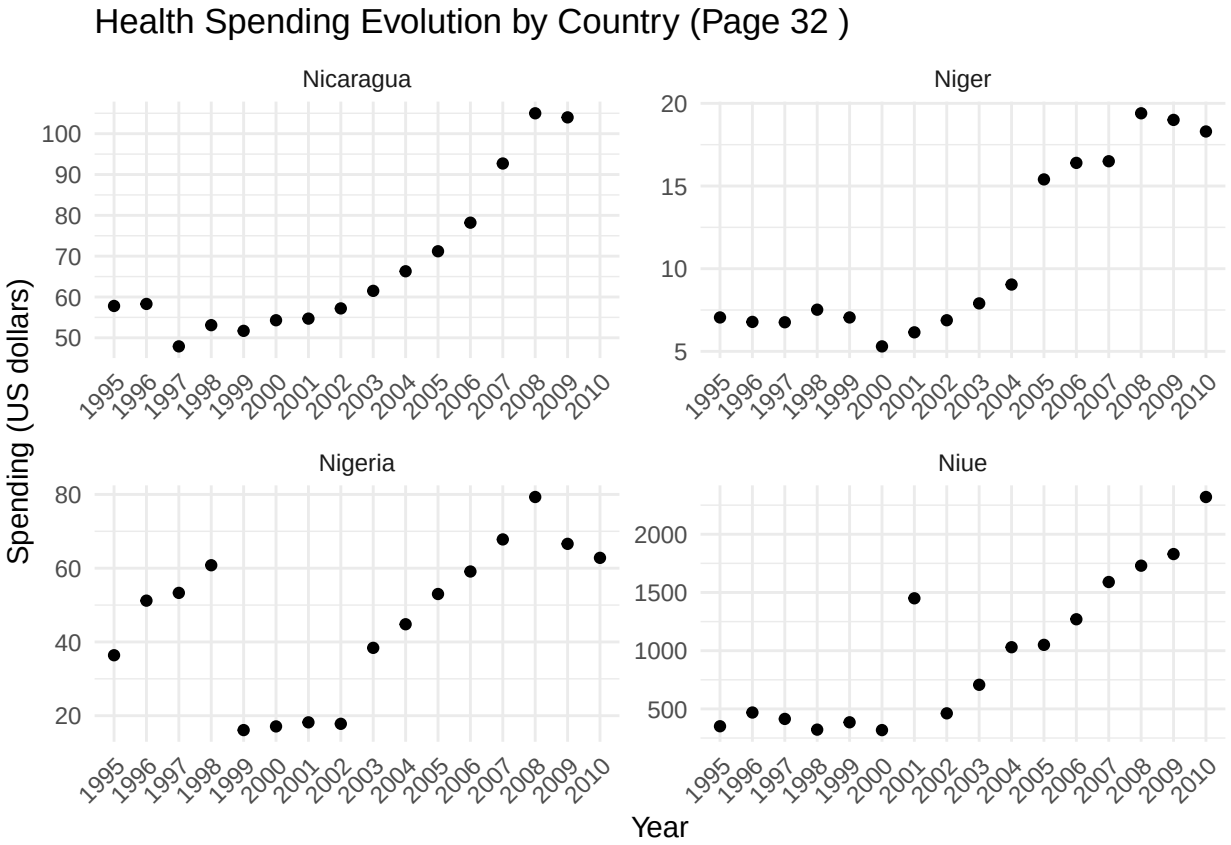

Health Spending Evolution by Country (Page 27)
Latvia
Moldova
Monaco
Morocco
Spending (US dollars)
Year
1000
750
500
250
200
150
100
50
6000
5000
4000
3000
125
75
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

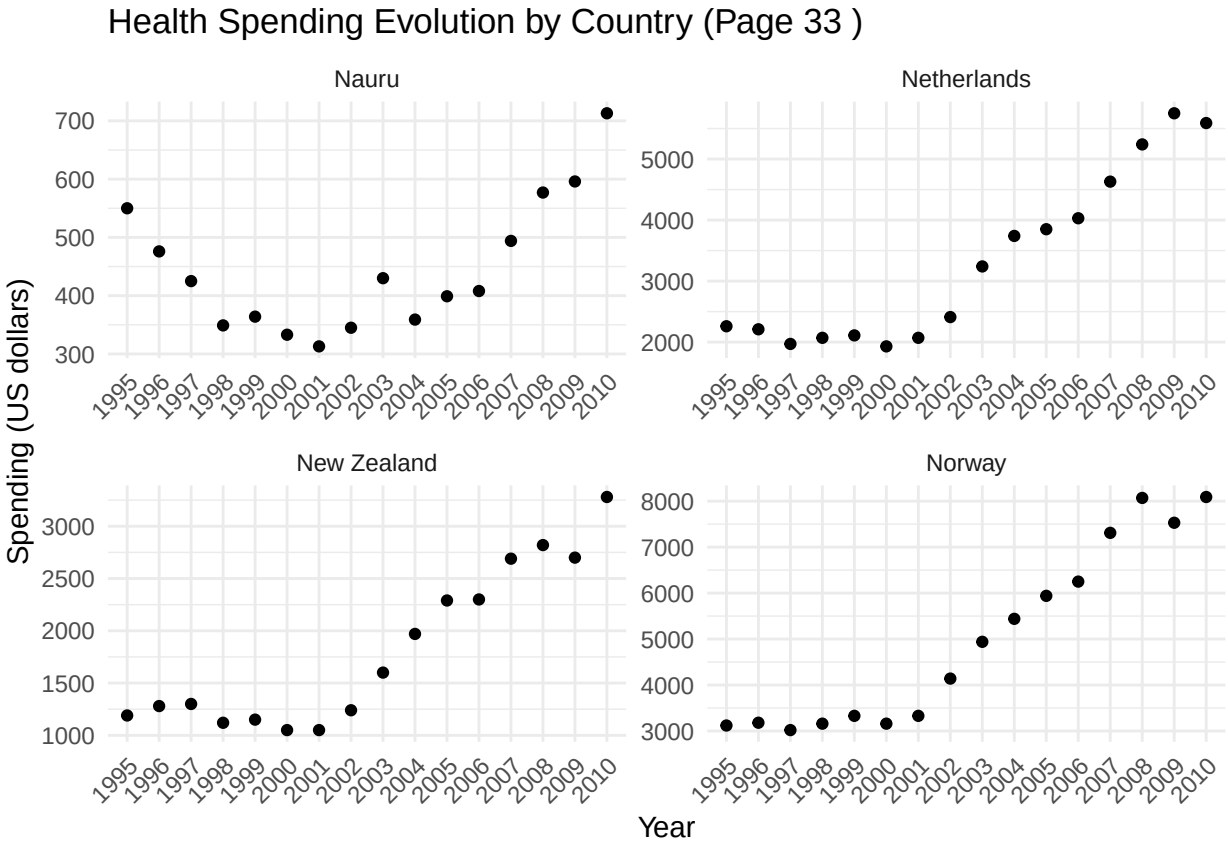

Health Spending Evolution by Country (Page 28)
Madagascar
Maldives
Marshall Islands
Mexico
Spending (US dollars)
Year

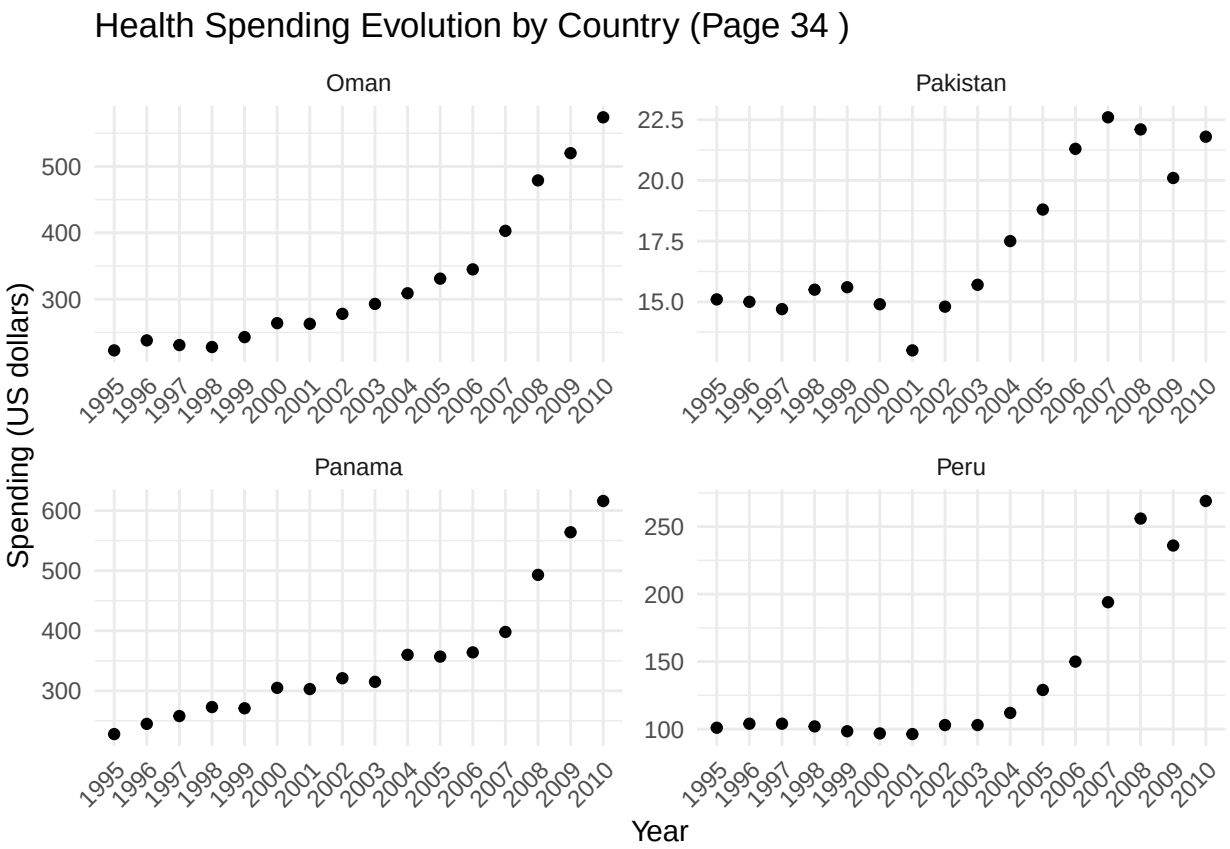

Health Spending Evolution by Country (Page 29)
Mali
Malta
Myanmar
North Macedonia
Spending (US dollars)
Year

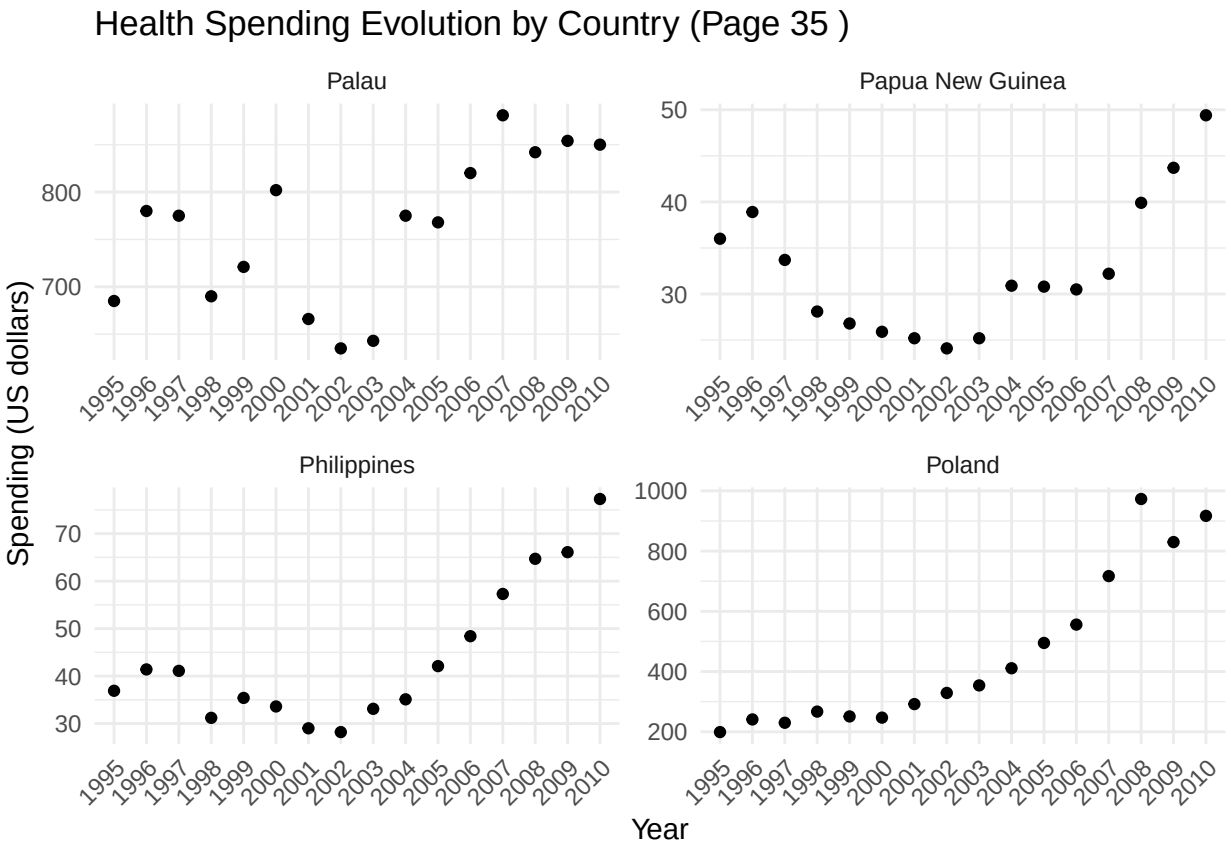

Health Spending Evolution by Country (Page 30)
Mauritania
Mongolia
Montenegro
Mozambique
Spending (US dollars)
Year

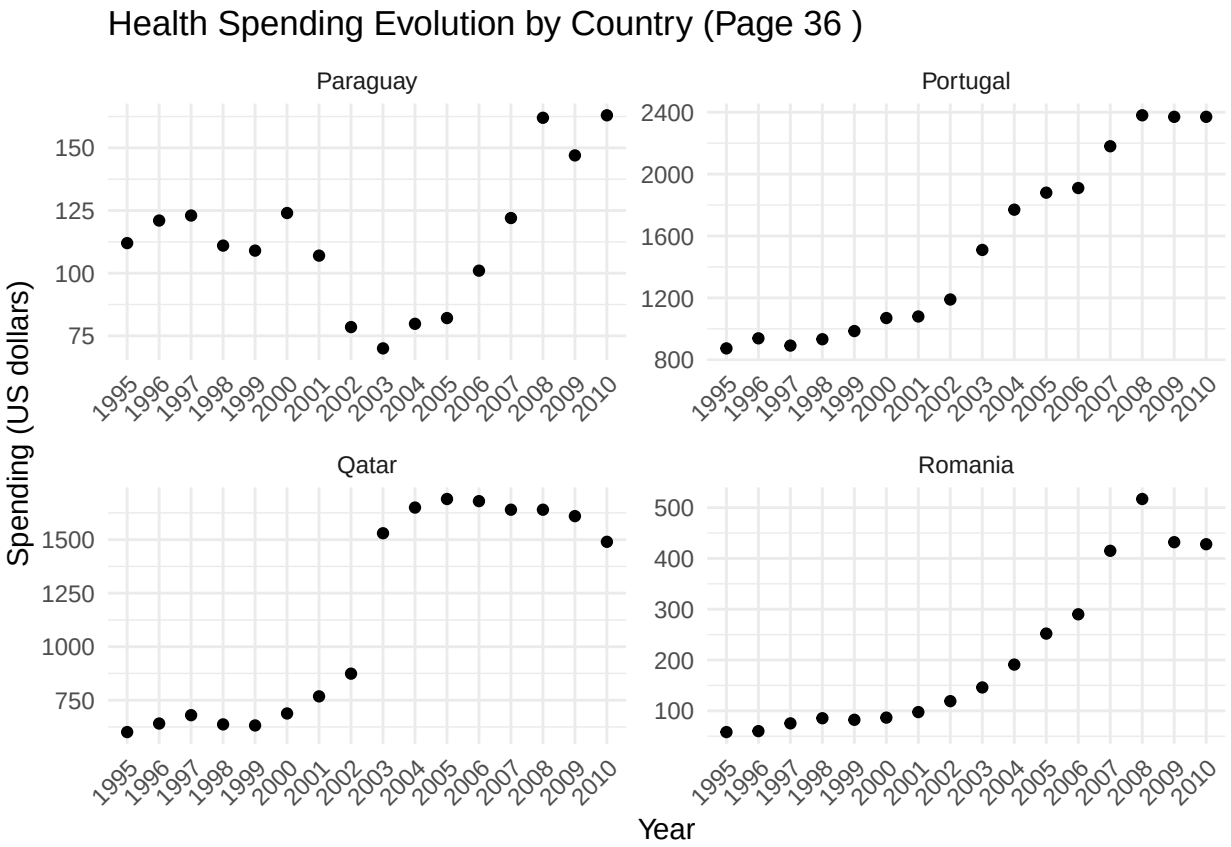

Health Spending Evolution by Country (Page 31)
Malawi
Malaysia
Mauritius
Namibia
Spending (US dollars)
Year

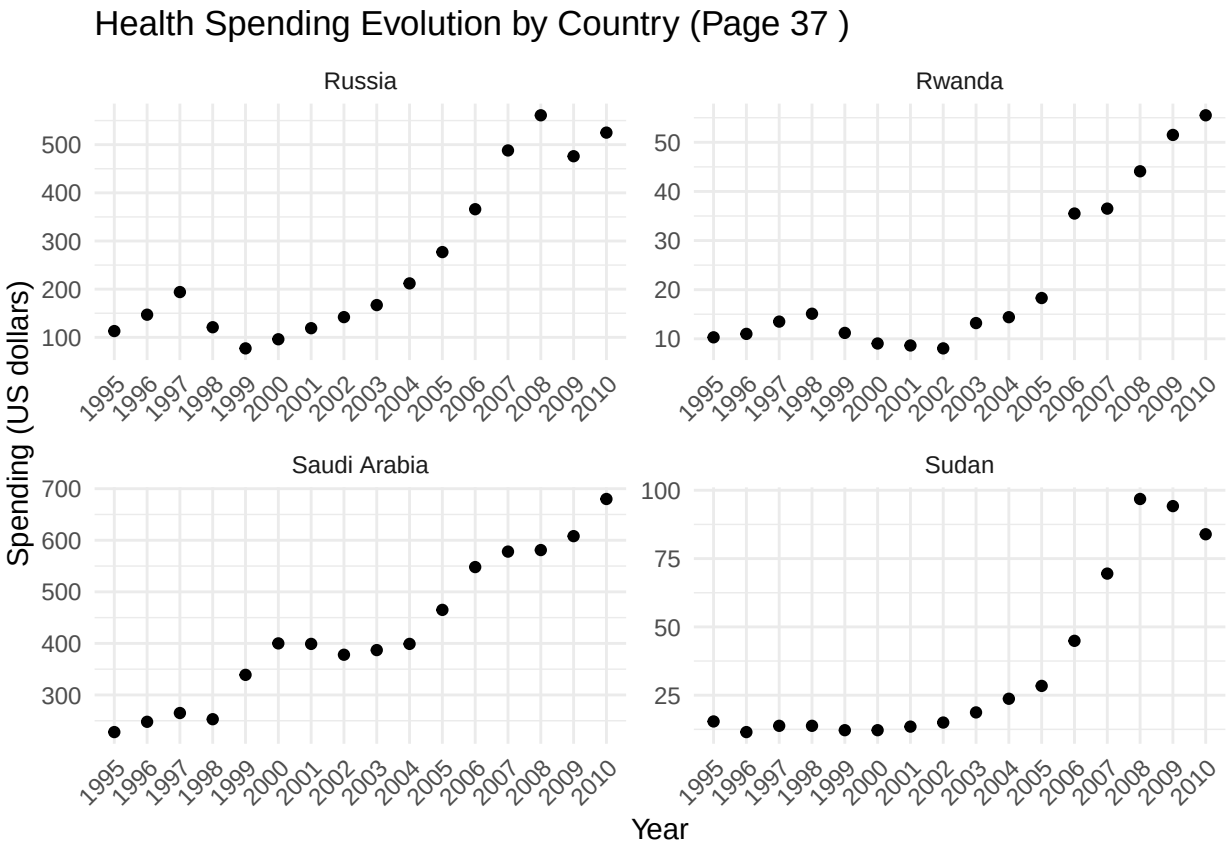

Health Spending Evolution by Country (Page 32)
Nicaragua
Niger
Nigeria
Niue
Spending (US dollars)
Year

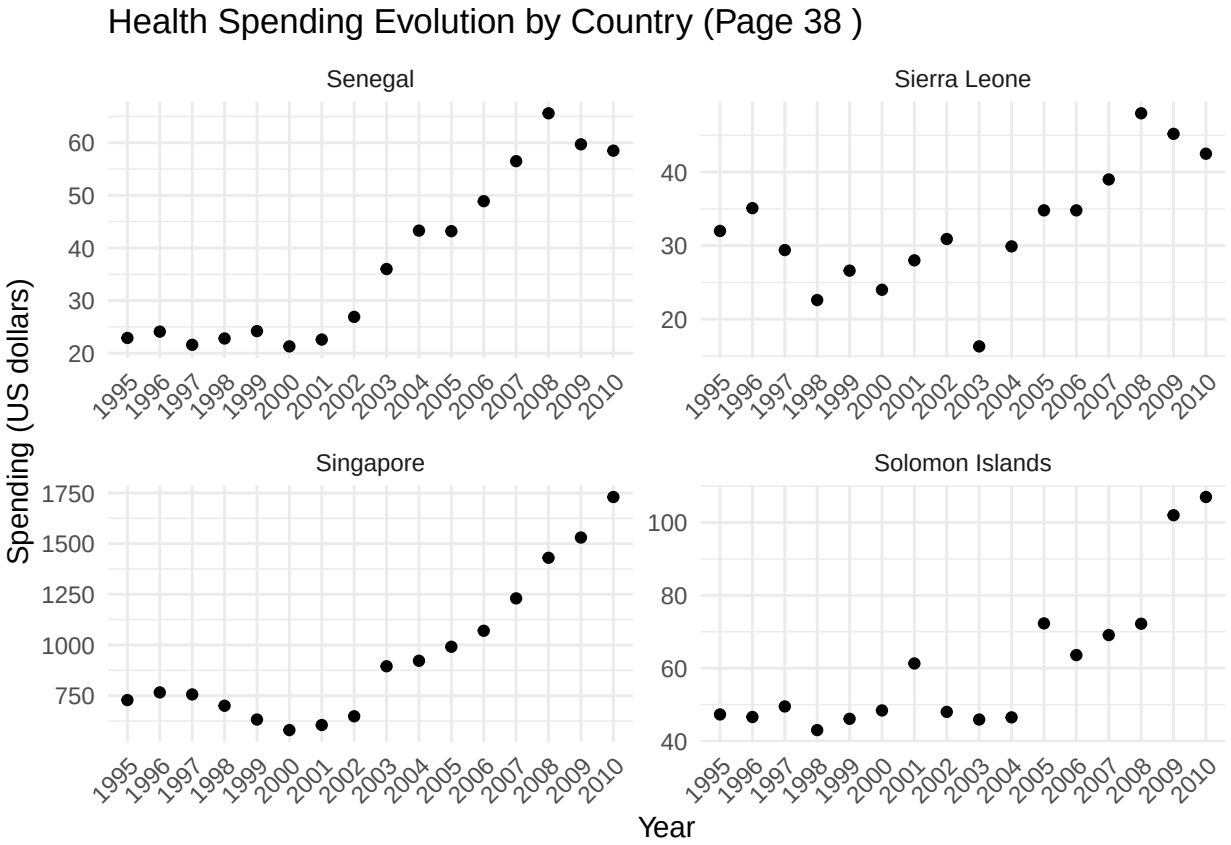

Health Spending Evolution by Country (Page 33)
Nauru
Netherlands
New Zealand
Norway
Spending (US dollars)
Year

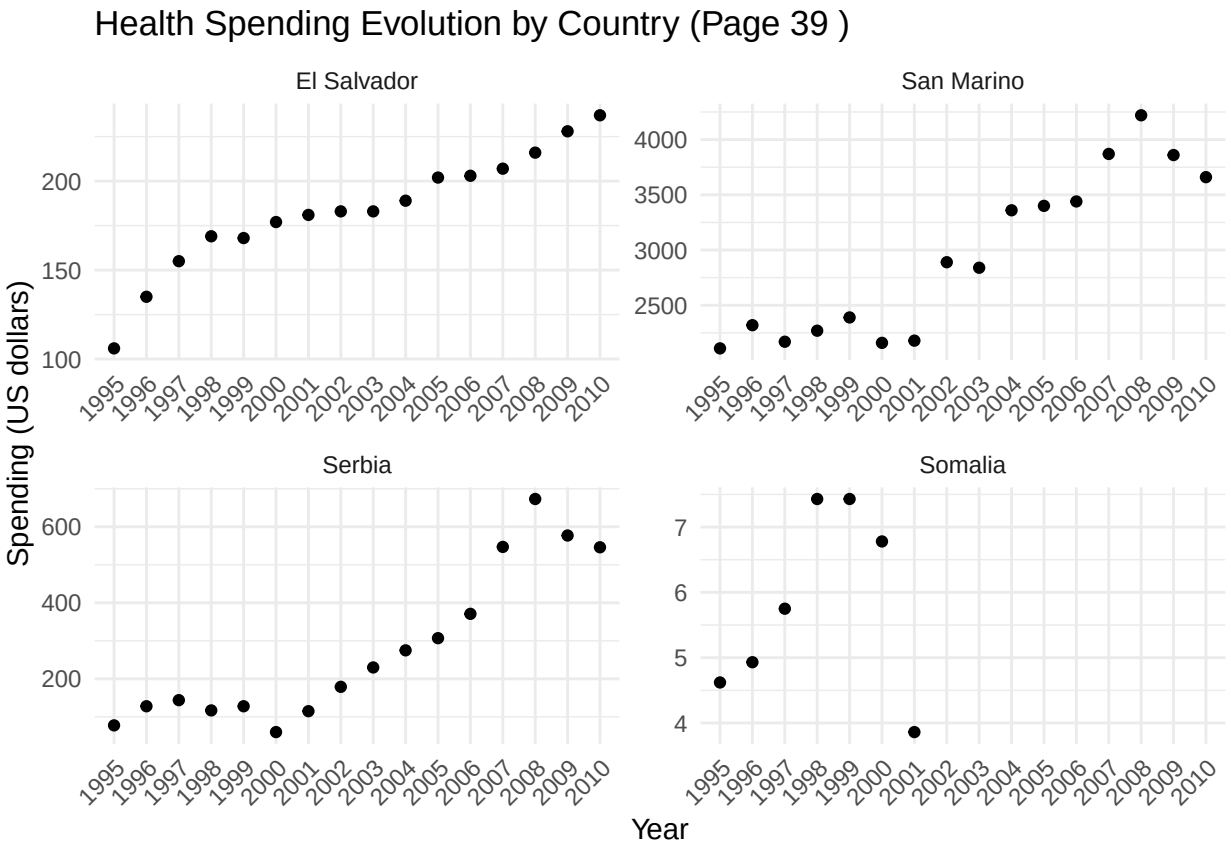

Health Spending Evolution by Country (Page 34)
Oman
Pakistan
Panama
Peru
Spending (US dollars)
Year
500
400
300
22.5
20.0
17.5
15.0
600
250
200
150
100
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

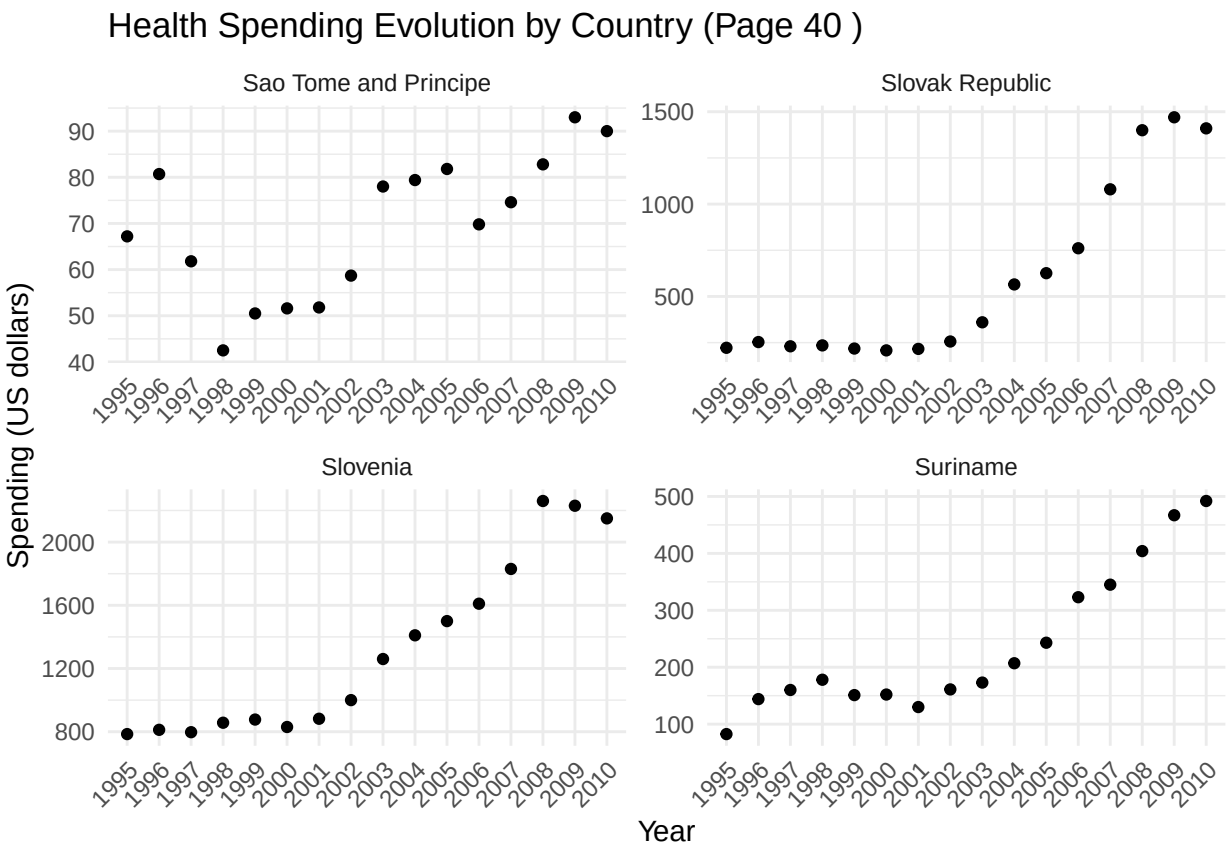

Health Spending Evolution by Country (Page 35)
Palau
Papua New Guinea
Philippines
Poland
Spending (US dollars)
Year

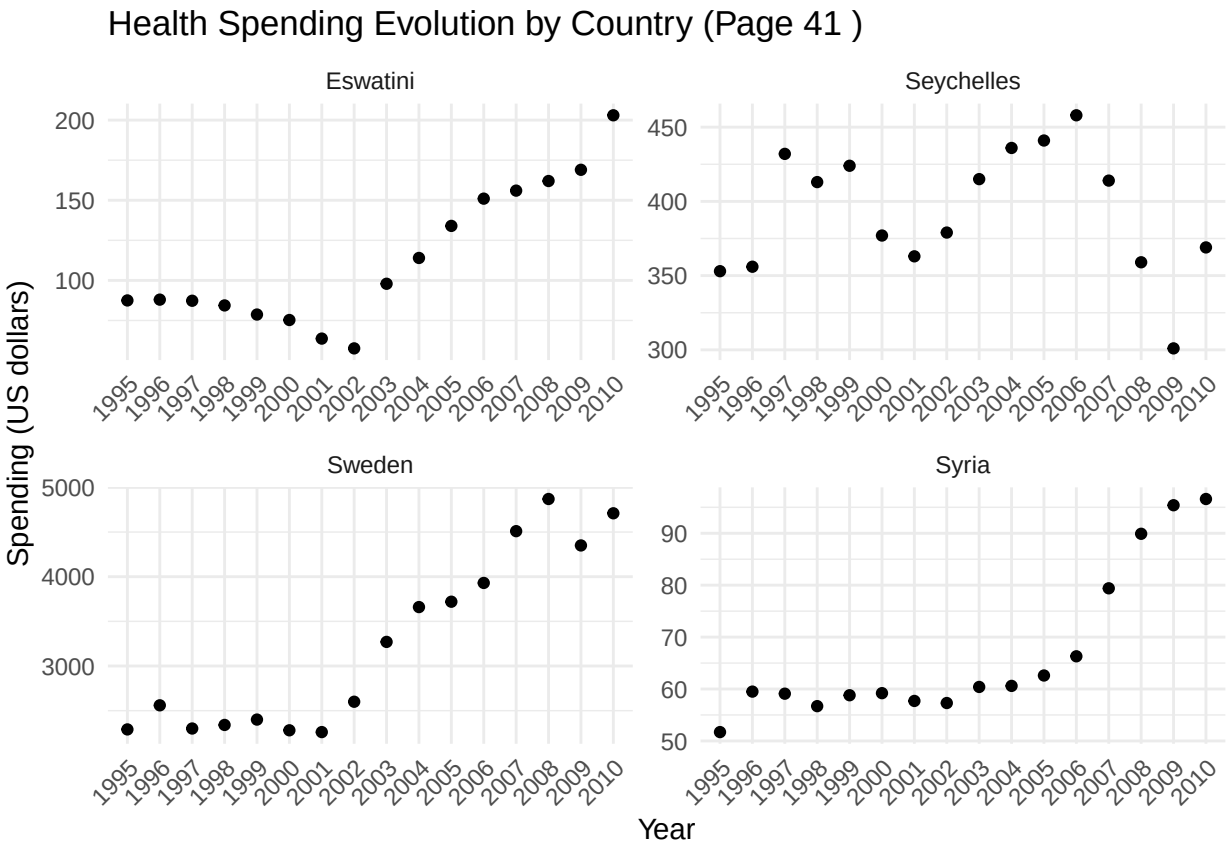

Health Spending Evolution by Country (Page 36)
Paraguay
Portugal
Qatar
Romania
Spending (US dollars)
Year
150
125
100
75
2400
2000
1600
1200
800
1500
1250
1000
750
500
400
300
200
100
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010

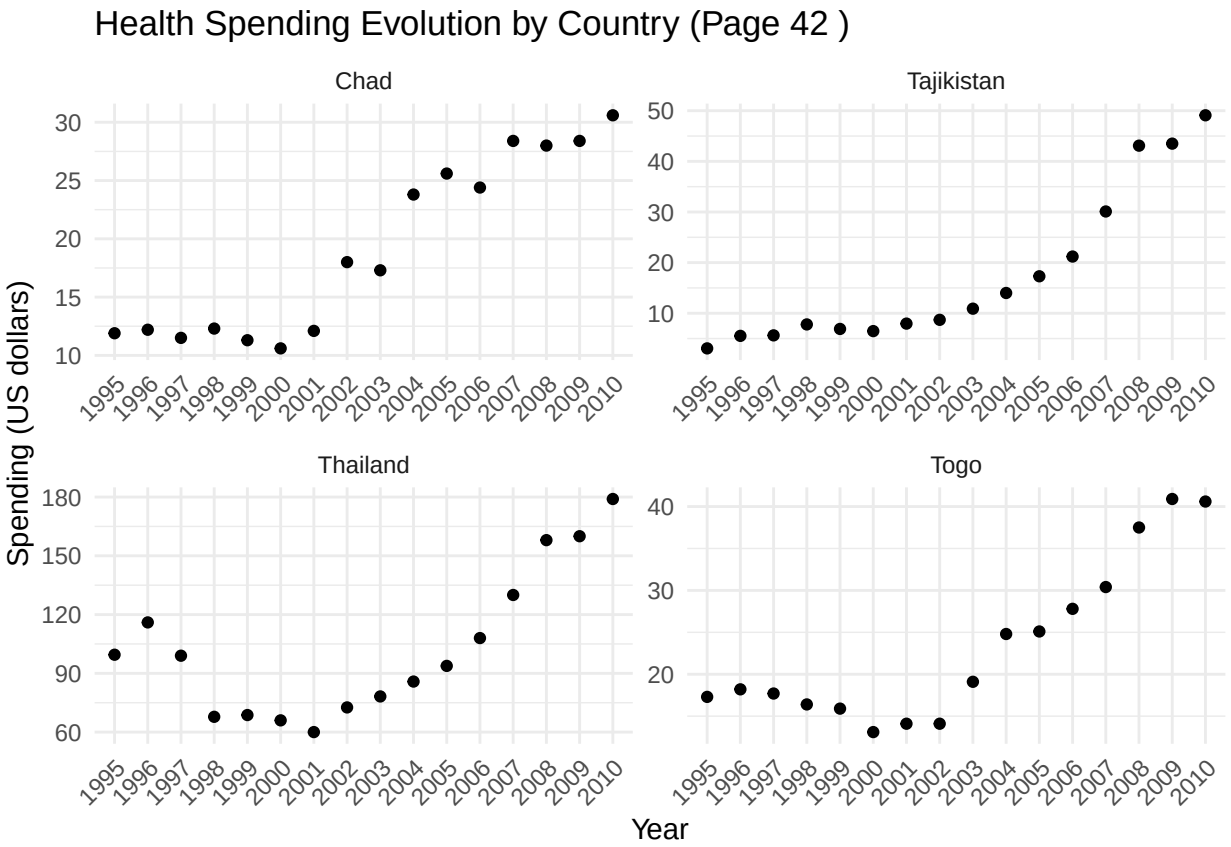

Health Spending Evolution by Country (Page 37)
Russia
Rwanda
Saudi Arabia
Sudan
Spending (US dollars)
Year

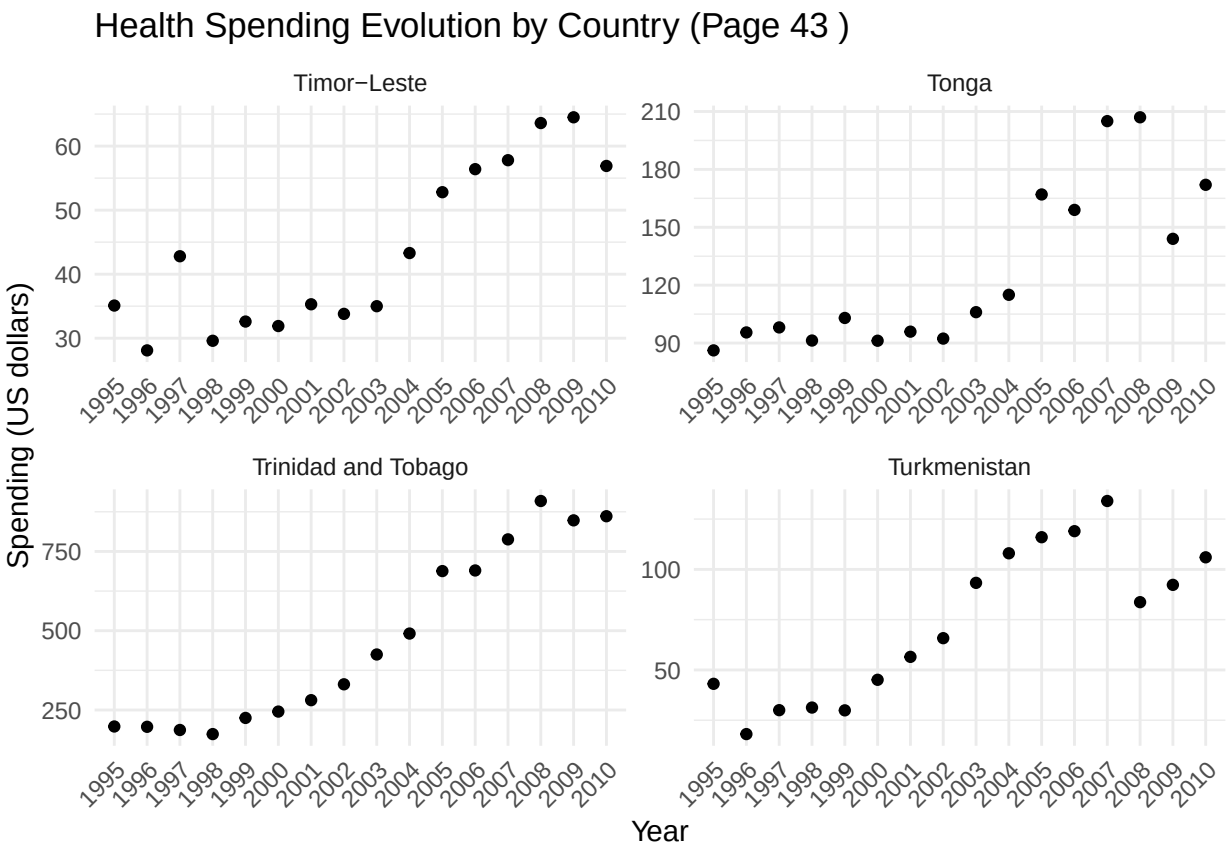

Health Spending Evolution by Country (Page 38)
Senegal
Sierra Leone
Singapore
Solomon Islands
Spending (US dollars)
Year

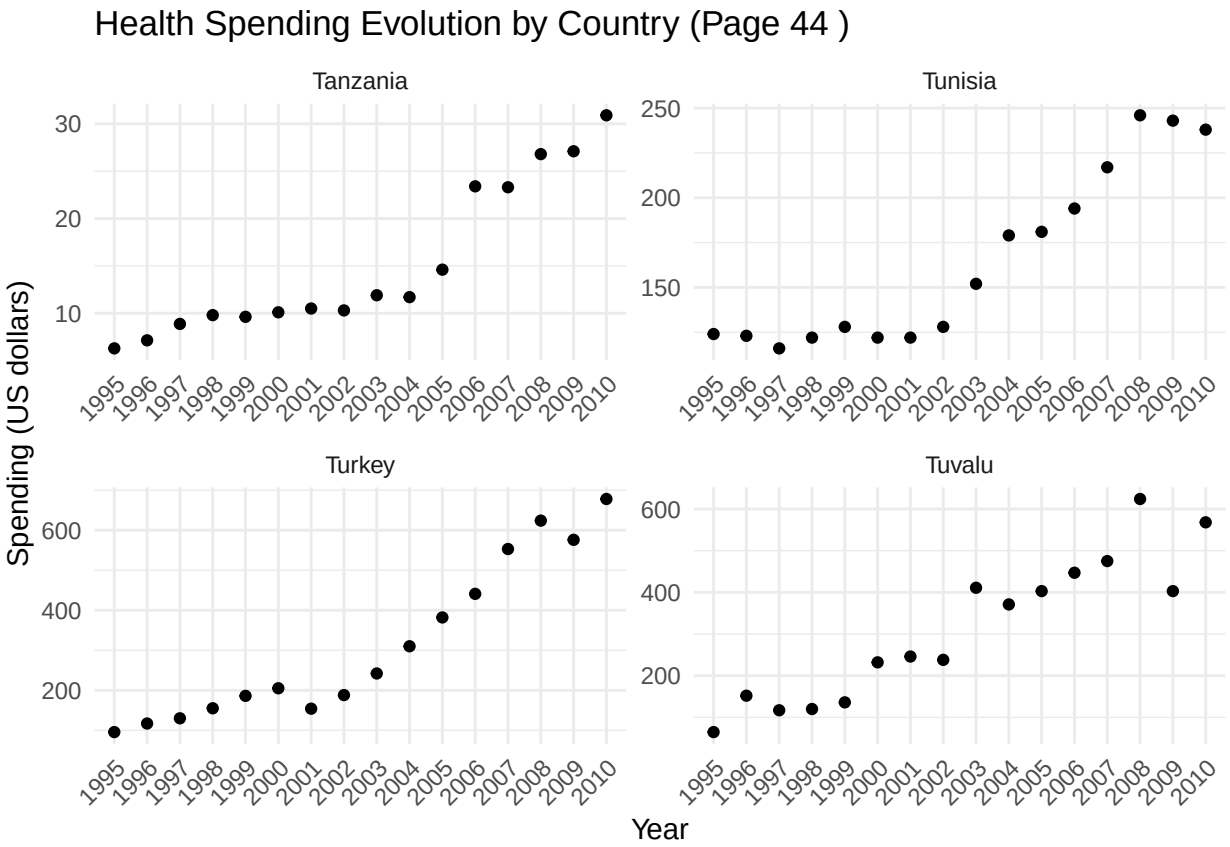

Health Spending Evolution by Country (Page 39)
El Salvador
San Marino
Serbia
Somalia
Spending (US dollars)
Year

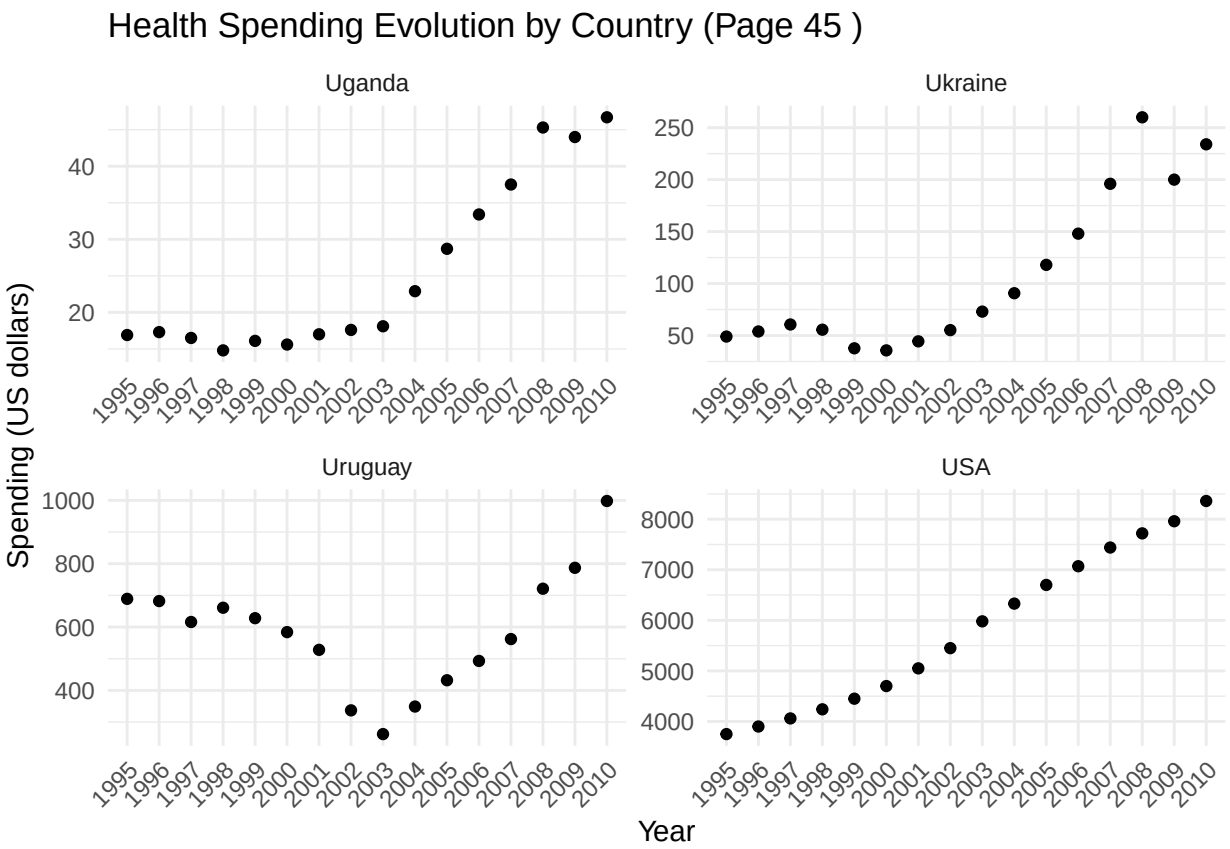

Health Spending Evolution by Country (Page 40)
Sao Tome and Principe
Slovak Republic
Slovenia
Suriname
Spending (US dollars)
Year

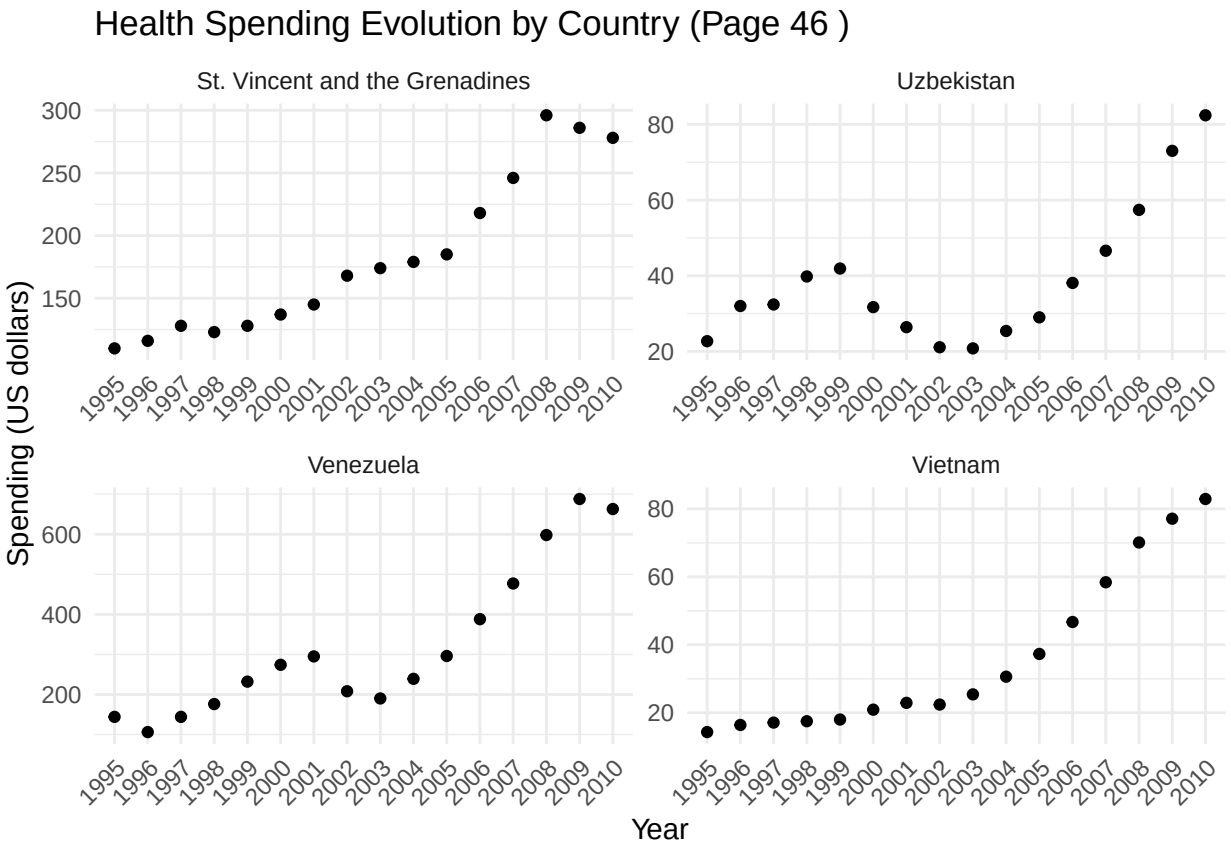

Health Spending Evolution by Country (Page 41)
Eswatini
Seychelles
Sweden
Syria
Spending (US dollars)
Year

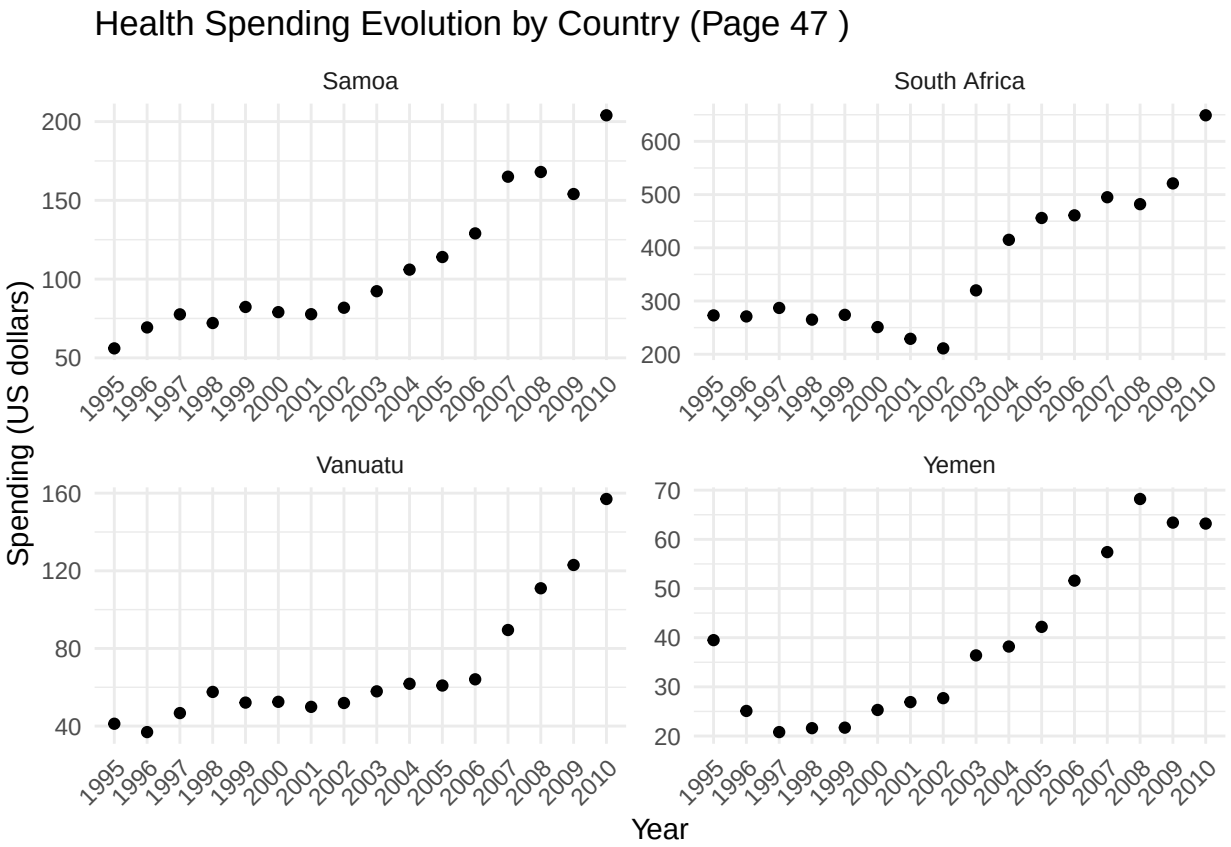

Health Spending Evolution by Country (Page 42)
Chad
Tajikistan
Thailand
Togo
Spending (US dollars)
Year

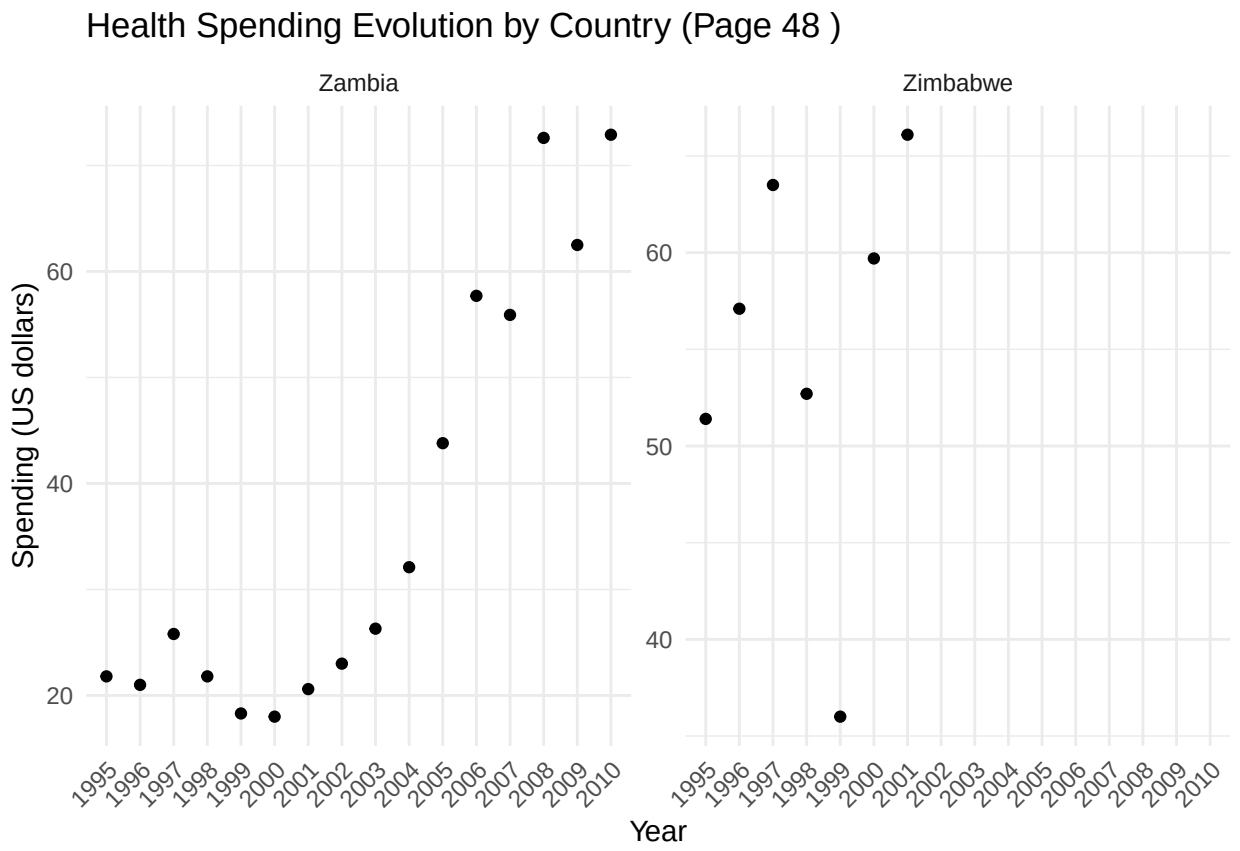

Health Spending Evolution by Country (Page 43)
Timor−Leste
Tonga
Trinidad and Tobago
Turkmenistan
Spending (US dollars)
Year


Health Spending Evolution by Country (Page 44)
Tanzania
Tunisia
Turkey
Tuvalu
Spending (US dollars)
Year


Health Spending Evolution by Country (Page 45)
Uganda
Ukraine
Uruguay
USA
Spending (US dollars)
Year


Health Spending Evolution by Country (Page 46)
St. Vincent and the Grenadines
Uzbekistan
Venezuela
Vietnam
Spending (US dollars)
Year


Health Spending Evolution by Country (Page 47)
Samoa
South Africa
Vanuatu
Yemen
Spending (US dollars)
Year




```r
# Save plots into a single PDF file
pdf("health_spending_evolution.pdf", width = 12, height = 8)

for (i in seq_along(plot_pages_arranged1)) {
  print(plot_pages_arranged1[[i]])
}
dev.off()  # Close the PDF device
```

Now, I will group countries by continent and create line plots with their health spending. Obviously, the representation of the average spending is severely skewed, since population is not being taken into account. A weighted average considering population would result in a more thorough analysis, but for the sake of the exercise I will continue.

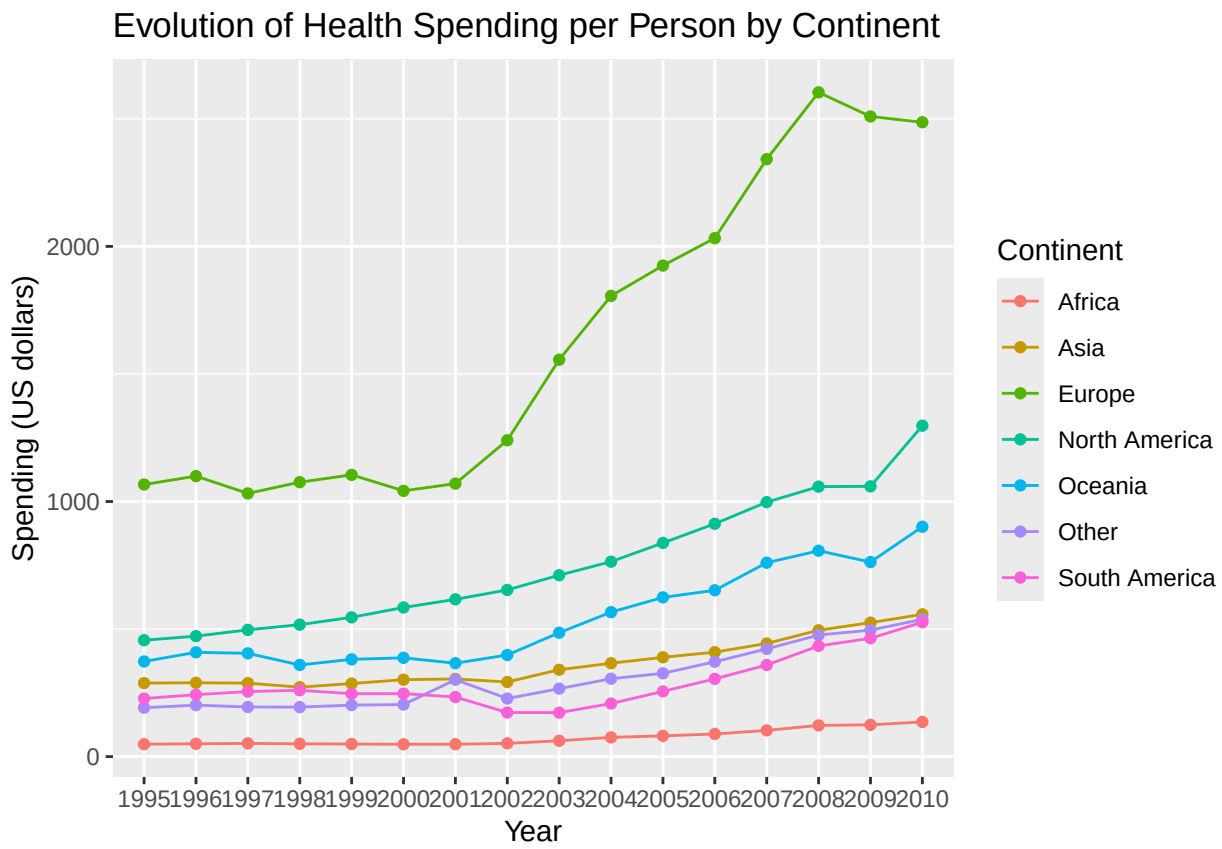
```r
# Assuming health_long is your data frame
countries <- unique(health_long$country)

# Create groups

countries_europe <- c("Albania", "Andorra", "Armenia", "Austria", "Azerbaijan", "Belarus", "Belgium", "

countries_africa <- c("Algeria", "Angola", "Benin", "Botswana", "Burkina Faso", "Burundi", "Cabo Verde"

countries_asia <- c("Afghanistan", "Bahrain", "Bangladesh", "Bhutan", "Brunei", "Cambodia", "China", "Ea

countries_north_america <- c("Antigua and Barbuda", "Bahamas", "Barbados", "Belize", "Canada", "Costa Ri
```

```r
countries_oceania <- c("Australia", "Fiji", "Kiribati", "Marshall Islands", "Micronesia", "Nauru", "New

countries_south_america <- c("Argentina", "Bolivia", "Brazil", "Chile", "Colombia", "Ecuador", "Guyana"
```

```r
# Create a new column 'continent' based on the categorization
health_long <- health_long %>%
  mutate(
    continent = case_when(
      country %in% countries_europe ~ "Europe",
      country %in% countries_asia ~ "Asia",
      country %in% countries_africa ~ "Africa",
      country %in% countries_north_america ~ "North America",
      country %in% countries_south_america ~ "South America",
      country %in% countries_oceania ~ "Oceania",
      TRUE ~ "Other"  # Handle any countries not categorized
    )
  )

# Calculate average spending per continent per year
average_spending <- health_long %>%
  group_by(continent, Year) %>%
  summarize(avg_spending = mean(Spending, na.rm = TRUE))

head(average_spending)
```

```
## # A tibble: 6 x 3
## # Groups:   continent [1]
##   continent Year  avg_spending
##   <chr>     <chr>        <dbl>
## 1 Africa    1995          48.6
## 2 Africa    1996          50.3
## 3 Africa    1997          52.0
## 4 Africa    1998          50.2
## 5 Africa    1999          49.3
## 6 Africa    2000          48.3
```

```r
# Plotting with adjustments
plot2 <- ggplot(average_spending,
                aes(x = Year, y = avg_spending,
                color = continent, group = continent)) +
  geom_line() +
  geom_point() +
  labs(title = "Evolution of Health Spending per Person by Continent",
       x = "Year",
       y = "Spending (US dollars)",
       color = "Continent")

ggsave("evolution_health_spending.png", plot = plot2, width = 8, height = 6)

plot(plot2)
```



Now, I want to know what the median spending per continent has been throughout 1995-2010. Thus, I will plot some box plots that will allow me to see where the median lies at a glance. Additionally, I believe the IQR will provide helpful information regarding how concentrated the data is, which conveniently allows us to determine whether people spend similar amounts of money on health-related services in each continent.

```
by(health_long, health_long$continent, summary)
```

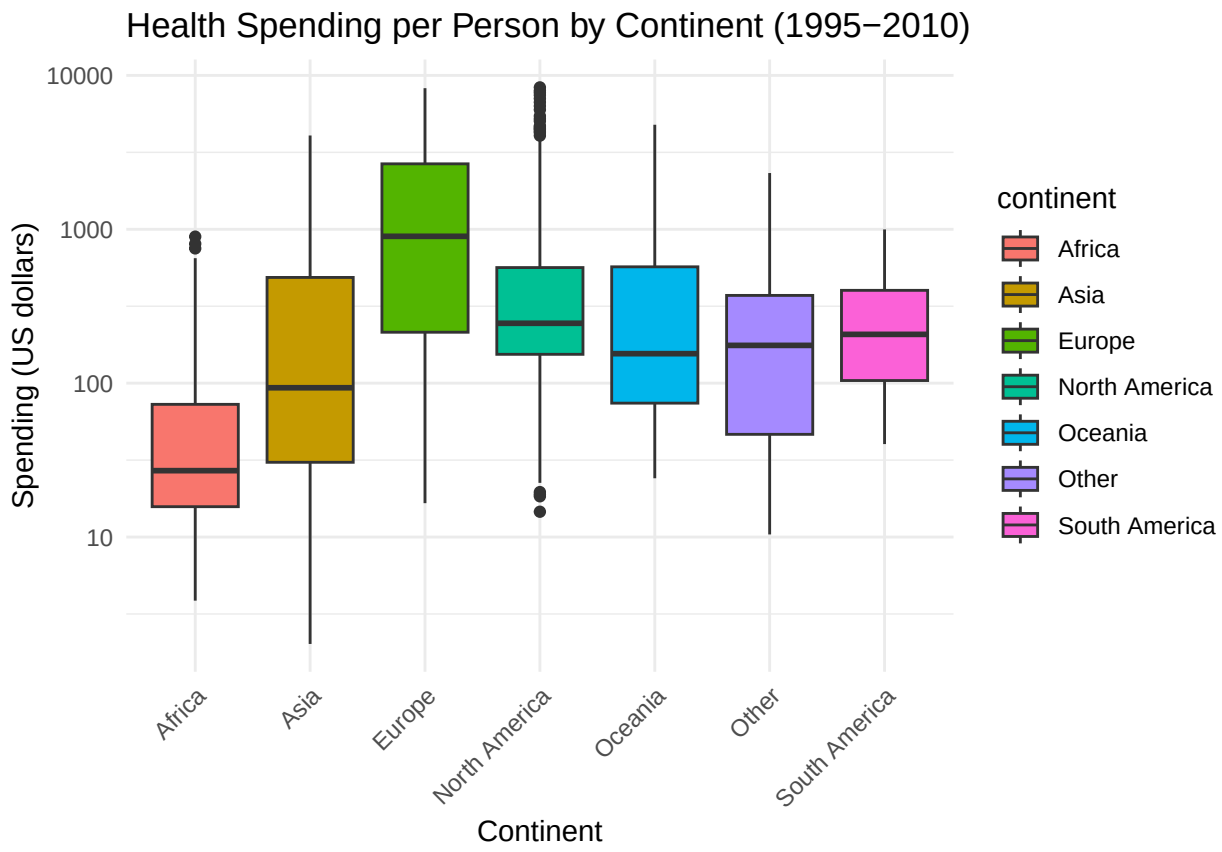
```
## health_long$continent: Africa
##    country              Year             Spending       continent
##  Length:816         Length:816         Min.   :  3.86   Length:816
##  Class :character   Class :character   1st Qu.: 15.75   Class :character
##  Mode  :character   Mode  :character   Median : 27.00   Mode  :character
##                                        Mean   : 74.16
##                                        3rd Qu.: 72.85
##                                        Max.   :896.00
##                                        NA's   :21
## ------------------------------------------------------------
## health_long$continent: Asia
##    country              Year             Spending        continent
##  Length:544         Length:544         Min.   :   2.02   Length:544
##  Class :character   Class :character   1st Qu.:  30.60   Class :character
##  Mode  :character   Mode  :character   Median :  93.35   Mode  :character
##                                        Mean   : 366.77
##                                        3rd Qu.: 486.50
##                                        Max.   :4070.00
##                                        NA's   :8
```

```
## ------------------------------------------------------------
## health_long$continent: Europe
##    country              Year              Spending       continent
##  Length:768         Length:768         Min.   :  16.6   Length:768
##  Class :character   Class :character   1st Qu.: 214.2   Class :character
##  Mode  :character   Mode  :character   Median : 900.0   Mode  :character
##                                        Mean   :1624.4
##                                        3rd Qu.:2662.5
##                                        Max.   :8260.0
## ------------------------------------------------------------
## health_long$continent: North America
##    country              Year              Spending       continent
##  Length:320         Length:320         Min.   :  14.6   Length:320
##  Class :character   Class :character   1st Qu.: 154.0   Class :character
##  Mode  :character   Mode  :character   Median : 245.0   Mode  :character
##                                        Mean   : 743.6
##                                        3rd Qu.: 564.0
##                                        Max.   :8360.0
##                                        NA's   :3
## ------------------------------------------------------------
## health_long$continent: Oceania
##    country              Year              Spending        continent
##  Length:208         Length:208         Min.   :  24.10   Length:208
##  Class :character   Class :character   1st Qu.:  74.17   Class :character
##  Mode  :character   Mode  :character   Median : 155.50   Mode  :character
##                                        Mean   : 539.86
##                                        3rd Qu.: 570.25
##                                        Max.   :4780.00
## ------------------------------------------------------------
## health_long$continent: Other
##    country              Year              Spending       continent
##  Length:192         Length:192         Min.   :  10.4   Length:192
##  Class :character   Class :character   1st Qu.:  46.5   Class :character
##  Mode  :character   Mode  :character   Median : 176.0   Mode  :character
##                                        Mean   : 307.5
##                                        3rd Qu.: 371.8
##                                        Max.   :2320.0
## ------------------------------------------------------------
## health_long$continent: South America
##    country              Year              Spending      continent
##  Length:192         Length:192         Min.   : 40.2   Length:192
##  Class :character   Class :character   1st Qu.:104.0   Class :character
##  Mode  :character   Mode  :character   Median :207.5   Mode  :character
##                                        Mean   :288.1
##                                        3rd Qu.:401.0
##                                        Max.   :998.0
```

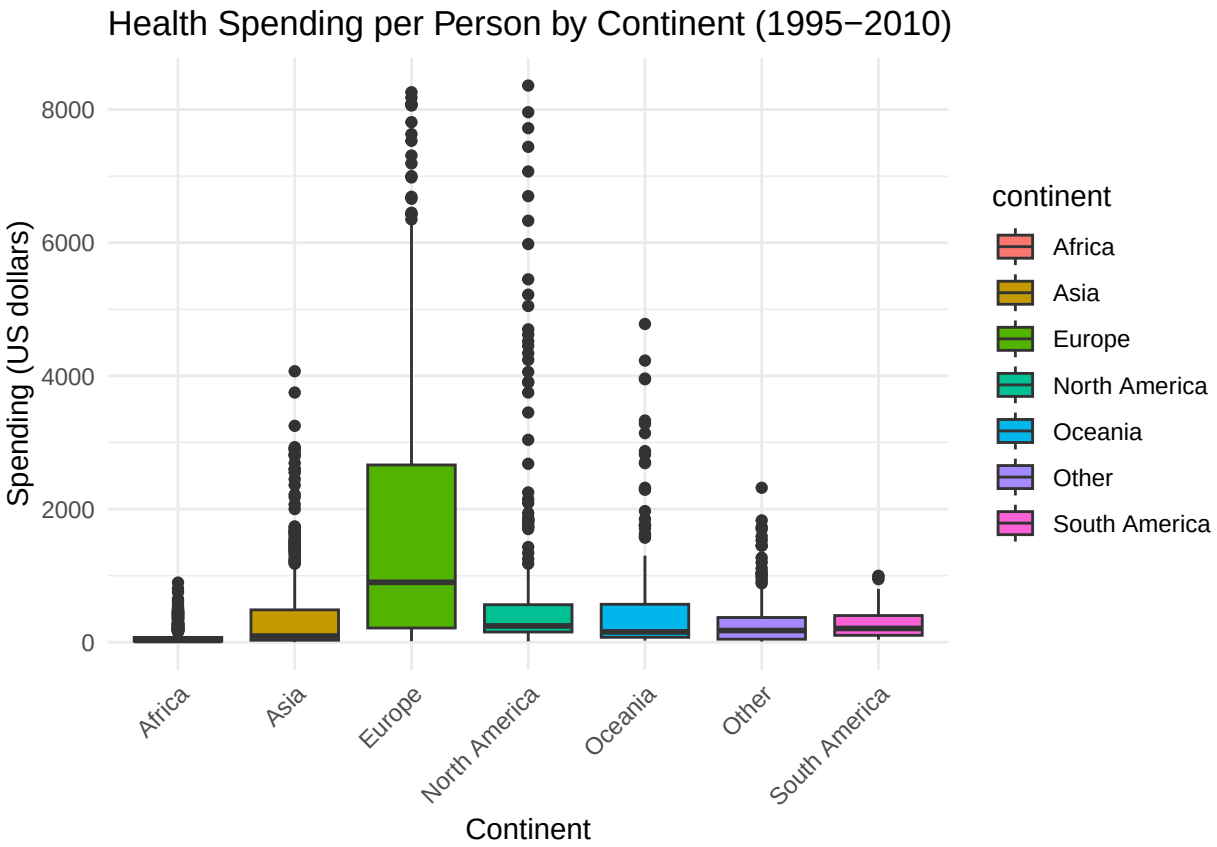
I will be using both logarithmic and non-logarithmic scales to better understand the data.

```
plot3 <- ggplot(health_long, aes(x = continent, y = Spending, fill = continent)) +
  geom_boxplot() +
  labs(title = "Health Spending per Person by Continent (1995-2010)",
       x = "Continent",
       y = "Spending (US dollars)") +
```

```
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_y_continuous(trans = 'log10')  # Using log scale due to large range

# Save the plot
ggsave("health_spendinglog_by_continent.png", plot = plot3, width = 10, height = 6)

# Display the plot
print(plot3)
```




```
plot4 <- ggplot(health_long, aes(x = continent, y = Spending, fill = continent)) +
  geom_boxplot() +
  labs(title = "Health Spending per Person by Continent (1995-2010)",
       x = "Continent",
       y = "Spending (US dollars)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_y_continuous()  # Using log scale due to large range

# Save the plot
ggsave("health_spending_by_continent.png", plot = plot4, width = 10, height = 6)

# Display the plot
print(plot4)
```



Intra-Continental Trends: Europe is the continent with the largest IQR in the non-logarithmic scale, showing absolute data, indicating a significant disparity among European countries. Some countries have high health expenditures while others have much lower ones, contributing to this difference. The presence of diverse healthcare systems throughout the continent may be a direct cause of this phenomenon.

North America's distribution is heavily influenced by high outliers, likely reflecting the USA's exceptionally high healthcare costs and Canada's healthcare system. Additionally, North America presents outliers below the 1st quartile mark, illustrating a severe disparity among countries.

Africa's low variability in absolute terms masks significant relative differences when viewed on a logarithmic scale. This reveals substantial relative differences in spending even among lower-spending countries, a nuance lost in the linear scale.

Lastly, Asia, Oceania, and South America form a middle tier, showing significant internal variability. Most continents display positively skewed distributions, indicating a longer tail of higher spenders.

Global Inequality: The linear scale highlights the enormous gap in health spending between continents. While some countries in North America and Europe spend over $8000 per person annually, most African nations spend less than $100. This extreme disparity underscores a critical global health inequality issue. The data emphasizes the urgent need for increased health funding in Africa and other low-spending regions.

High variability within continents suggests opportunities for knowledge sharing and policy learning among neighboring countries. This raises the question of whether similar countries can apply the same strategies to improve the situation. Similarly, extreme outliers in high-spending regions raise questions about healthcare system efficiency and sustainability (Kruk et al. 2018).

References

Kruk, M. E., Gage, A. D., Arsenault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., Adeyi, O., Barker, P., Daelmans, B., Doubova, S. V., English, M., García-Elorrio, E., Guanais, F., Gureje, O., Hirschhorn, L. R.,

Jiang, L., Kelley, E., Lemango, E. T., Liljestrand, J., Malata, A., ... Pate, M. (2018). High-quality health systems in the Sustainable Development Goals era: time for a revolution. The Lancet. Global health, 6(11), e1196–e1252. https://doi.org/10.1016/S2214-109X(18)30386-3