Portfolio 6

Sigurd Fyhn Sørensen

2022-10-30

# 1

Define a defensive function that calculates the Gross Domestic Product of a nation from the data available in the gapminder dataset. You can use the population and GDPpercapita columns for it. Using that function, calculate the GDP of Denmark in the following years: 1967, 1977, 1987, 1997, 2007, and 2017.

library(gapminder)  
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.4.0 ✔ purrr 0.3.5   
## ✔ tibble 3.1.8 ✔ dplyr 1.0.10  
## ✔ tidyr 1.2.1 ✔ stringr 1.4.1   
## ✔ readr 2.1.3 ✔ forcats 0.5.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

df <- gapminder::gapminder  
  
GDP\_func <- function(pop, GDPpercap){  
 stopifnot(is.numeric(pop))  
 stopifnot(is.numeric(GDPpercap))  
   
 GDP <- pop \* GDPpercap  
   
 if (sum(is.na(GDP))> 0){ warning("Some GDP values returned NA")}  
 if (sum(GDP <= 0)> 0){ warning("GDP can't be negative check your the level of Population & GDP Per Capita")}  
   
 return(GDP)  
   
}  
  
df <- df %>%   
 mutate(GDP = GDP\_func(df$pop, df$gdpPercap))  
  
  
df %>%   
 filter(country == "Denmark" & year %in% c( 1967, 1977, 1987, 1997, 2007,2017))

## # A tibble: 5 × 7  
## country continent year lifeExp pop gdpPercap GDP  
## <fct> <fct> <int> <dbl> <int> <dbl> <dbl>  
## 1 Denmark Europe 1967 73.0 4838800 15937. 77116977700.  
## 2 Denmark Europe 1977 74.7 5088419 20423. 103920280028.  
## 3 Denmark Europe 1987 74.8 5127024 25116. 128771236166.  
## 4 Denmark Europe 1997 76.1 5283663 29804. 157476118456.  
## 5 Denmark Europe 2007 78.3 5468120 35278. 192906627081.

# 2

Write a script that loops over each country in the gapminder dataset, tests whether the country starts with a ‘B’ , and prints out whether the life expectancy is smaller than 50, between 50 and 70, or greater than 70. (Hint: remember the grepl function, and review the Control Flow tutorial)

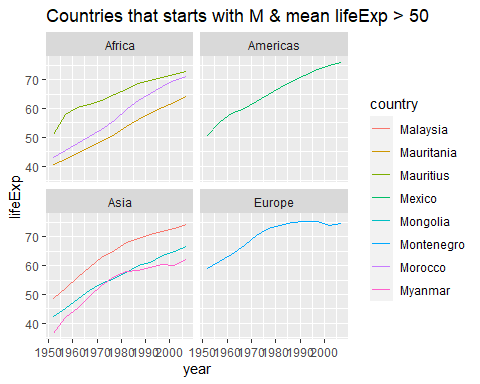
df\_unique <- df %>%   
 distinct(country)  
  
data.frame(Country = df$country , year = df$year,   
 Beings\_with\_b = grepl(df$country, pattern = "^B"),  
 Life\_exp\_bin = if\_else(df$lifeExp > 70, ">70",if\_else(df$lifeExp < 50, "<50", "50-70"))) %>%   
 sample\_n(20) #Just to show a subset

## Country year Beings\_with\_b Life\_exp\_bin  
## 1 Dominican Republic 1952 FALSE <50  
## 2 Finland 2007 FALSE >70  
## 3 Mali 1962 FALSE <50  
## 4 Niger 1957 FALSE <50  
## 5 Japan 1957 FALSE 50-70  
## 6 Japan 2007 FALSE >70  
## 7 Austria 1962 FALSE 50-70  
## 8 Kenya 2002 FALSE 50-70  
## 9 Liberia 2007 FALSE <50  
## 10 Puerto Rico 2002 FALSE >70  
## 11 Costa Rica 2002 FALSE >70  
## 12 Angola 1967 FALSE <50  
## 13 Netherlands 1967 FALSE >70  
## 14 Burkina Faso 1952 TRUE <50  
## 15 Poland 1987 FALSE >70  
## 16 Congo, Rep. 2007 FALSE 50-70  
## 17 Gabon 2002 FALSE 50-70  
## 18 Greece 1962 FALSE 50-70  
## 19 Yemen, Rep. 1992 FALSE 50-70  
## 20 Rwanda 1987 FALSE <50

# 3

Challenge/Optional: Write a script that loops over each country in the gapminder dataset, tests whether the country starts with a ‘M’ and graphs life expectancy against time (using plot() function) as a line graph if the mean life expectancy is under 50 years.

df %>%   
 filter(grepl(country, pattern = "^M") == T) %>%  
 group\_by(country) %>%   
 filter(mean(lifeExp) > 50) %>%   
 ggplot(aes(x = year, y = lifeExp, Group = country, color = country)) + geom\_line() + facet\_wrap(~continent) + labs(title = "Countries that starts with M & mean lifeExp > 50")



df %>%   
 filter(grepl(country, pattern = "^M") == T) %>%  
 group\_by(country) %>%   
 filter(mean(lifeExp) < 50) %>%   
 ggplot(aes(x = year, y = lifeExp, Group = country, color = country)) + geom\_line() + facet\_wrap(~continent) + labs(title = "Countries that starts with M & mean lifeExp < 50")

