Homework 6

Cultural Datascience 2022

Jørgen Højlund Wibe

Loading data

```
library(pacman)
pacman::p_load(gapminder, tidyverse)
df <- gapminder
```

Task 1

Task description

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.

Solution

```
# Defining function
find_gdp <- function(df, year=NULL, country=NULL) {
  if (!is.numeric(year)) {
    warning(stop("Year must be of type numeric"))
  }
  if (!is.character(country)) {
    warning(stop("Country must be of type character"))
  }
  if(!is.null(year)) { # if the 'year' variable is not empty...
    df <- df[df$year %in% year, ] # add the specified year to df
  }

  if (!is.null(country)) { # same as above
    df <- df[df$country %in% country, ] # -//-
  }

  gdp <- df$pop * df$gdpPercap # calculating gdp: multiplies population with GDP per capita

  new_df <- cbind(df, gdp=gdp) # combining results into one dataframe
  return(new_df)
}

# Running function, specifying values
find_gdp(df, year = c(1967, 1977, 1987, 1997, 2007, 2017), country = "Denmark" )
```

```
##   country continent year lifeExp     pop gdpPercap          gdp
## 1 Denmark    Europe 1967  72.960 4838800  15937.21  77116977700
## 2 Denmark    Europe 1977  74.690 5088419  20422.90 103920280028
## 3 Denmark    Europe 1987  74.800 5127024  25116.18 128771236166
## 4 Denmark    Europe 1997  76.110 5283663  29804.35 157476118456
## 5 Denmark    Europe 2007  78.332 5468120  35278.42 192906627081
```

Task 2

Task description

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 (https://swcarpentry.github.io/r-novice-gapminder/07-control-flow/index.html))

Solution

```
countries <- unique(df$country)

for (i in countries) {
  life_expectancy <- df %>%
    filter(country == i) %>%
    pull(lifeExp)

  if ((mean(life_expectancy) < 50) && (i %in% grep("^B", countries, value = T))) {
    print(paste0(i, " starts with 'B' and has a average life expectancy less than 50!"))
  }

  else if (between(mean(life_expectancy), 50, 70) && (i %in% grep("^B", countries, value = T))) {
    print(paste0(i, " starts with 'B' and has an average life expectancy between 50 and 70"))
  }

  else if (mean(life_expectancy) > 70 && (i %in% grep("^B", countries, value = T))) {
    print(paste0(i, " starts with 'B' and has an average life expectancy more than 70!"))
  }
}
```

```
## [1] "Bahrain starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Bangladesh starts with 'B' and has a average life expectancy less than 50!"
## [1] "Belgium starts with 'B' and has an average life expectancy more than 70!"
## [1] "Benin starts with 'B' and has a average life expectancy less than 50!"
## [1] "Bolivia starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Bosnia and Herzegovina starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Botswana starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Brazil starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Bulgaria starts with 'B' and has an average life expectancy between 50 and 70"
## [1] "Burkina Faso starts with 'B' and has a average life expectancy less than 50!"
## [1] "Burundi starts with 'B' and has a average life expectancy less than 50!"
```