Gapminder Data Set Exploration

Denitsa Vasileva

2019-09-14

What is the gapminder data set?

- Built-in data set in R
- A Tibble
- Contains GDP of countries over a 12 year period
- To load it we use the following code:

```
library(gapminder)
```

Exploring the gapminder data set: Summary Statistics

To get summary statistics on the gapmind data set run the following code:

```
summary(gapminder)
```

Exploring the gapminder data set: How many countries are included in this data set?

To find how many countries are represented in the data set we use the following code:

```
num_countries <- length(unique(gapminder$country))
print (num_countries)
```

```
## [1] 142
```

142 countries were included in this study.

Exploring the gapminder data set: What was the country with the lowest GDP per capita overall?

```
min_GDP_country <- data.frame(gapminder
[gapminder$gdpPercap==min (gapminder$gdpPercap),])
print (min_GDP_country )
```

The Dem. Republic of Congo was the country with the lowest GDP per capita in this data set.

Exploring the gapminder data set: What was the country with the highest GDP per capita?

```
max_GDP_country <- data.frame(gapminder
[gapminder$gdpPercap==max (gapminder$gdpPercap),])
print (max_GDP_country)
```

```
##    country continent year lifeExp    pop gdpPercap
## 1  Kuwait      Asia 1957  58.033 212846  113523.1
```

Kuwait was the country with the highest GDP per capita in this data set.

Exploring the gapminder data set: Minimum GDP per capita per year

To find the lowest GDP per capita in every year use the following code:

```
df_min_GDP_year <- data.frame()
years <- unique(gapminder$year)
for (year in years){
  new_df<-data.frame(gapminder
                     [gapminder$year==year,])
  x<-new_df[(new_df$gdpPercap==
               min(new_df$gdpPercap)),]
  df_min_GDP_year<-rbind(df_min_GDP_year,x )
}
print(df_min_GDP_year)
```

Exploring the gapminder data set: Maximum GDP per capita per year

To find the highest GDP per capita in every year use the following code:

```
df_max_GDP_year <- data.frame()
years<- unique(gapminder$year)
for (year in years){
  new_df<-data.frame(gapminder
                     [gapminder$year==year,])
  x<-new_df[(new_df$gdpPercap==max
               (new_df$gdpPercap)),]
  df_max_GDP_year<-rbind(df_max_GDP_year,x )
}
print(df_max_GDP_year)
```