

570_reproducibility

2023-10-20

```
#importing the data  
library(readr)  
gdp_52_67 <- read.csv("inputs/gdp_52-67.csv")  
le_52_67 <- read.csv("inputs/le_52-67.csv")
```

```
#Merging the data based on column names  
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'  
  
## The following objects are masked from 'package:stats':  
##  
##   filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
merged_data <- gdp_52_67 %>%  
  full_join(le_52_67)
```

```
## Joining with 'by = join_by(continent, country, year)'
```

Calculate the GDP per capita by dividing the GDP(gdp) column by the population(pop) column

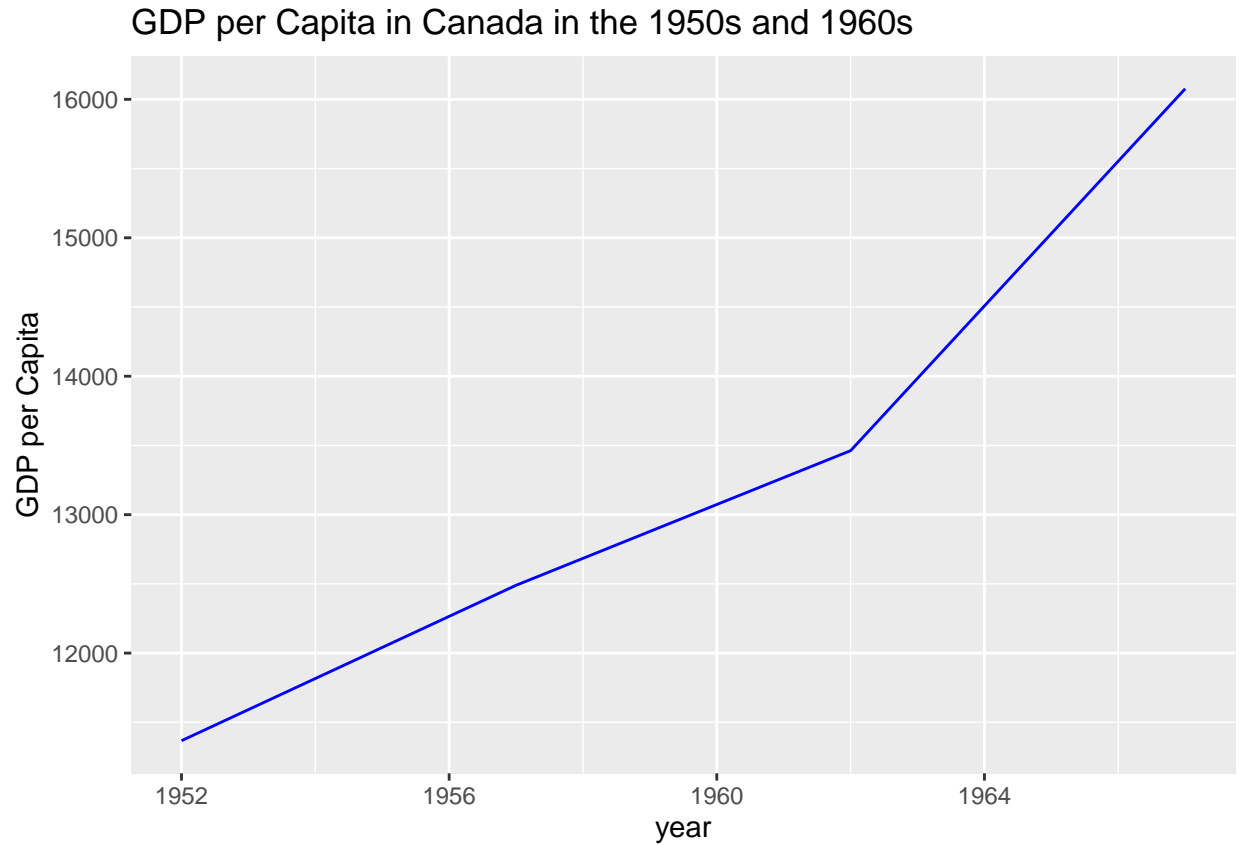
```
# calculating GDP per capita  
merged_data$gdpPercap <- merged_data$gdp / merged_data$pop
```

Visualize the data with line graphs

```
#visualizing life expectancy and gdp per capita for Canada in the 1950s and 60s  
library(ggplot2)  
  
#filtering the data for years 50-60s and Canada  
filtered_data <- merged_data %>%  
  filter(country == "Canada", year >= 1950, year <= 1969)  
  
#creating a line plot for gdp per Capita and year  
gdpPercap_plot <- ggplot(filtered_data, aes(x = year, y = gdpPercap)) +  
  geom_line(color = "blue") +
```

```
labs(y = "GDP per Capita") +
ggtitle ("GDP per Capita in Canada in the 1950s and 1960s")

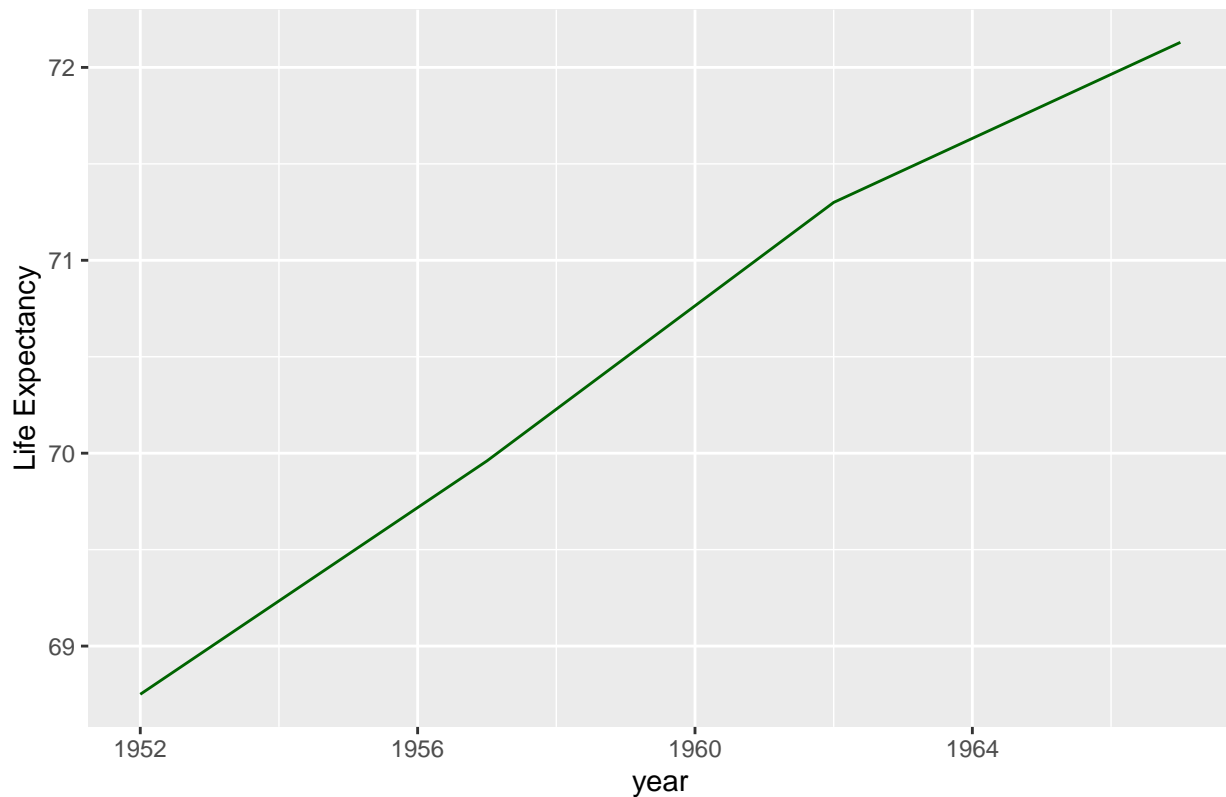
print(gdpPercap_plot)
```



```
#creating a line plot for life Expectancy and year
lifeExp_plot <- ggplot(filtered_data, aes(x = year, y = lifeExp)) +
  geom_line(color = "dark green") +
  labs(y = "Life Expectancy") +
  ggtitle ("Life Expectancy in Canada in the 1950s and 1960s")

print(lifeExp_plot)
```

Life Expectancy in Canada in the 1950s and 1960s



From the graphs above, it is evident that both GDP per capita and life expectancy has increased over time in Canada from 1950 -1970. Both increase at a steady rate until 1965, in which life expectancy becomes less steep and GDP per Capita becomes more steep.

```
#regressing life expectancy on GDP per Capita
model <- lm(lifeExp ~ gdpPercap, data = filtered_data)

summary(model)
```

```
##
## Call:
## lm(formula = lifeExp ~ gdpPercap, data = filtered_data)
##
## Residuals:
##      1      2      3      4
## -0.39996  0.02538  0.68572 -0.31114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.121e+01  2.333e+00  26.240  0.00145 **
## gdpPercap    6.989e-04  1.733e-04   4.033  0.05633 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6032 on 2 degrees of freedom
## Multiple R-squared:  0.8905, Adjusted R-squared:  0.8358
## F-statistic: 16.27 on 1 and 2 DF, p-value: 0.05633
```