DS102-09-12

Lesson 9 Hands-On

Heather Walker

submitted: 2022-08-18*

*Hit with a combo: sinus infection led to waiting too long to start, and then had issues with RStudio (with .Rprofile) that delayed finished the work on time. Wanting to start sooner on classwork once I'm feeling better!

Directions:

- use `gapminder` data frame
- compare 5 countries of your choice
- use this command to see all the countries:
  `levels(gapminder$country)`
- TIP! Load library `dplyr`, `ggplot2`, and `gapminder` for this activity

Country selection:

Denmark, Norway, Sweden, Finland, and Iceland
the “Nordic Region”

Comparisons of per capita GDP

1952

Lowest per capita GDP:

Finland

Highest per capita GDP:

Norway

2007

Lowest per capita GDP:

Finland (again)

Lowest per capita GDP:

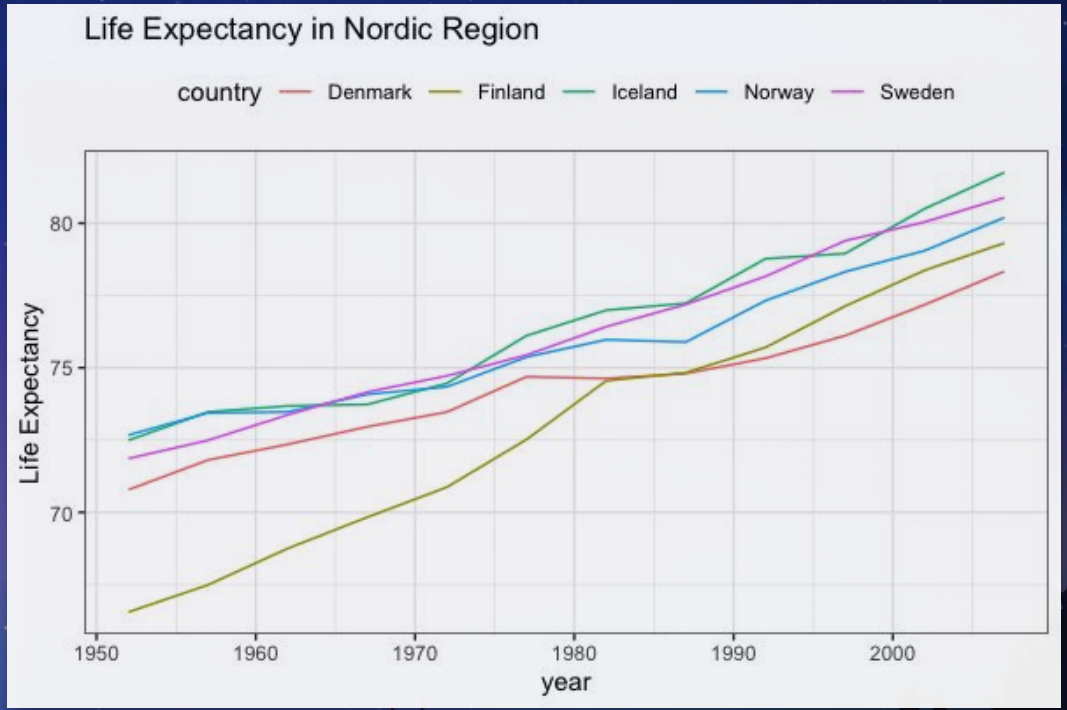
Norway (again)

```
# Plot life expectancy as a function of time:
ggplot(gm_NordicClean)

+ geom_line(aes(x = year, y = lifeExp, color = country))

+ ylab("Life Expectancy")

+ ggtitle("Life Expectancy in Nordic Region")
```



What are the variations in the life expectancy between countries?

- All 5 had a gradual improvement in life expectancy from 1952 to 2007.
- Finland started (1952) with a much lower life expectancy compared to the other 4 Nordic Region countries.
- Finland had a steeper increase in life expectancy from 1952 to just after 1980.
- After 1980 to 2007, Finland's improvement in life expectancy was kept a similar pace as the other 4.

Directions continued:

- Compute the median of `lifeExp` for each year on the entire `gapminder` data frame
- Compare your 5 countries to the entire `gapminder` data frame
- For what years is the life expectancy for your 5 countries ABOVE the median life expectancy for the entire `gapminder` data frame?

Life expectancy comparison

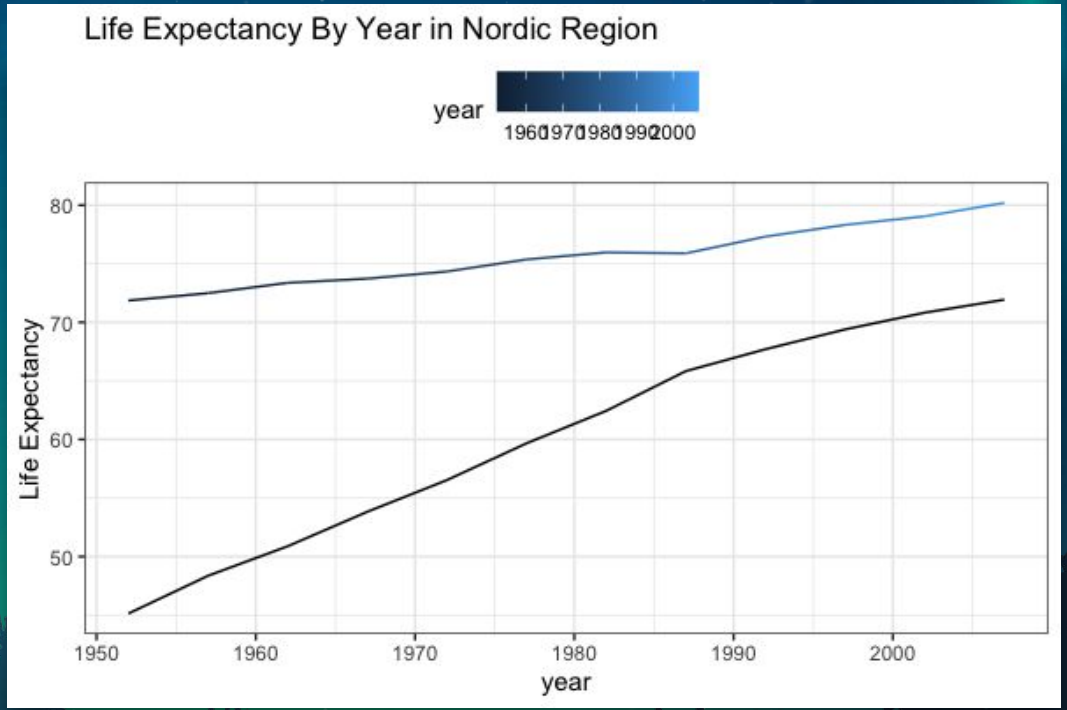
Which year did your selected 5 countries have a median life expectancy ABOVE the median life expectancy for the entire `gapminder` data frame?

- The Nordic Region had a higher median life expectancy ABOVE the entire `gapminder` data frame for the entire time frame (1952 to 2007)
- This is not surprising –
  - From the included data, the Nordic Region has a much higher GDP for those years
  - Outside of the data, the Nordic Region is made up of all 1st world countries (Source)

```
# Combined data plot*

ggplot(plot_mediansTogether)
+ geom_line(aes(x = year, y = life_medNordic, color = year))
+ geom_line(aes(x = year, y = life_medAll))
+ ylab("Life Expectancy")

+ ggtitle("Life Expectancy By Year in Nordic Region")
```



*I am not happy with this solution, but wanted to submit at least something for this part of the assignment.


Thanks !