discussion2

Michelle Vered

2023-01-24

# Code

# read in data  
life\_expectancy <- read.csv('UNdata.csv')  
  
# drop rows with na since the only nas are 2 footnote rows incorrectly imported into the data  
life\_expectancy <- life\_expectancy %>%  
 drop\_na()  
  
# summary statistics  
summary(life\_expectancy)

## Country.or.Area Subgroup Year Source   
## Length:80 Length:80 Length:80 Length:80   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## Unit Value Value.Footnotes  
## Length:80 Min. :53.00 Min. : 1.0   
## Class :character 1st Qu.:70.00 1st Qu.: 3.0   
## Mode :character Median :75.00 Median : 5.5   
## Mean :73.88 Mean : 5.5   
## 3rd Qu.:79.00 3rd Qu.: 8.0   
## Max. :84.00 Max. :10.0

# mean life expectancy for men vs. women  
tapply(life\_expectancy$Value, life\_expectancy$Subgroup, summary)

## $Female  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 57.00 74.00 79.00 76.58 81.00 84.00   
##   
## $Male  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 53.00 69.00 73.00 71.17 75.00 79.00

# mean life expectancy for by date range  
tapply(life\_expectancy$Value, life\_expectancy$Year, summary)

## $`1985-1990`  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 53.00 69.00 73.00 71.90 78.25 80.00   
##   
## $`1990-1995`  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 56.00 69.75 74.00 73.15 79.25 81.00   
##   
## $`1995-2000`  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 59.00 71.50 75.00 74.55 80.25 83.00   
##   
## $`2000-2005`  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 61.00 73.25 76.50 75.90 81.25 84.00

# number of countries in the data set  
length(unique(life\_expectancy$Country.or.Area))

## [1] 10

# Analysis

## Links

Github Link: <https://github.com/mvered/DMA3300/tree/main/week_2>

Data Source Link: <https://www.kaggle.com/datasets/mayurmatang/lifeexpectancy>

## Background on Data Set

This data set was sourced from the UN Population Division - World Population Prospects Estimates. That source is a statistics gathering effort by the United Nations Statistics Division. It covers life expectancy data for different countries and for men and women between the years of 1985 - 2005.

## Data Description

The data contains 1,571 observations, each representing life expectancy for a particular demographic group (Male or Female) in a particular 5 year period and in a particular country or territory.

Here are some key columns from the data set:

* Country.or.Area: The country (or territory) that the life expectancy is for.
* Subgroup: Male or Female
* Year: The data is grouped into four different 5-year periods: 1985-1990, 1990-1995, 1995-2000, and 2000-2005.
* Value: The life expectancy value in years, stored as an integer.

There are two rows of data which are actually just footnotes, incorrectly imported as their own rows. None of the rows with actual data have missing values in any of the key columns.

## Descriptive Statistics

The average life expectancy value across all observations in the data set is 73.875 years. The minimum value is 53 years and the maximum is 84 years. The median life expectancy value is 75.

We can also look at average life expectancy for men versus women and for each of the different date ranges. The average life expectancy for women is 76.58 years, compared to 71.17 years for men. The average life expectancy in each of the four 5-year periods was:

* 1985-1990: 71.90 years
* 1990-1995: 73.15 years
* 1995-2000: 74.55 years
* 2000-2005: 75.90 years

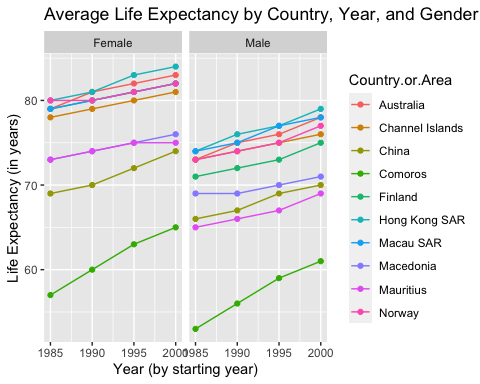
There are 10 unique countries/territories in the data set.

## Research Question & Motivation

We’re interested in understanding whether life expectancy is higher for men or for women, in certain countries versus in others, and whether life expectancy has gone up or down over time.

This information can helpful for people who work in international development to understand the differences between countries and where there is room for improvement. It’s also useful for those working in public health to understand where there may be gaps in healthcare provided in some countries or to certain demographic groups. It would be an encouraging trend to see life expectancy going up over time.

## Data Visualizations





## Findings

1. Life expectancy has gone up over time. This is demonstrable by looking at the line plot above which shows that across gender groups and countries, life expectancy consistently went up from 1985 - 2005.
2. Life expectancy is higher for men than for women. Across all countries and years, the average life expectancy for women was 76.58 years while it was 71.17 years for men.
3. The most common life expectancy value in our data set was between 70-75 years. This is shown in the second plot, the histogram. There are some countries and demographic groups with life expectancy as low as 53 years and as high as 84 years.