EC316 Assignment 1

Othar Zaldastani II

February 27, 2026

Setup Chunk

```
knitr::opts_chunk$set(fig.pos = "H", message = FALSE, warning = FALSE)
library(dplyr)
library(readr)
library(tidyverse)
knitr::opts_chunk$set(fig.pos = "H")
sweden <- read_csv("sweden_merged.csv")

# I can do this for future assignments if needed, but I got it to work my own way if that's alright.
# setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
```

Section 1

Question 1

Data Confirmation

```
head(sweden)
```

```
## # A tibble: 6 x 8
##    year age    female_deaths male_deaths total_deaths female_pop male_pop
##   <dbl> <chr>          <dbl>       <dbl>        <dbl>      <dbl>    <dbl>
## 1  1751 0               5988        6902        12890     28596.   28609.
## 2  1751 1               1286.       1360.        2646.    26192.   25873.
## 3  1751 2                835.        882.        1717.    26508.   26127.
## 4  1751 3                622         655.        1277.    22054.   21638.
## 5  1751 4                471.        498.         969.    18244.   17682.
## 6  1751 5                339         365.         704.    18297.   17825.
## # i 1 more variable: total_pop <dbl>
```

Question 2

Cleaning Data

```
sweden = filter(sweden, age != "110+")
sweden = sweden |>
  filter(total_pop > 0)
```

Question 3

Life expectancy in 1751

Part A   Mutating Data

```
sweden.1751 = filter(sweden, year == 1751)
sweden.1751 = mutate(sweden.1751,
      mx = total_deaths/total_pop,
      px = (1 - mx),
      px_lagged = lag(px, default = 1),
      lx = cumprod(px_lagged)
      )
```

Part B   Sum of life expectancy

```
sum(sweden.1751$lx)
```

The sum of `lx` is 37.2, meaning life expectancy for people in Sweden in 1751 was 37.2 years.

Question 4

Life expectancy across 4 years.

```
# Adding in some of my own years to compare
years <- c(1751, 1850, 1918, 2020)
for (yr in years) {
    sweden.year <- sweden |>
        filter(year == yr) |>
        mutate(
            mx = (total_deaths/total_pop),
            px = (1 - mx),
            px_lagged = lag(px, default = 1),
            lx = cumprod(px_lagged))
    cat(yr, ":", round(sum(sweden.year$lx, na.rm = TRUE), 2), "\n")}
```

```
## 1751 : 37.16
## 1850 : 43.83
## 1918 : 49.82
## 2020 : 82.63
```

The increase in life expectancy seems relatively constant between the first three years measured, but there is a significant increase between 1918 and 2020. This is likely due to modern medicine such as antibiotics and greatly decreased infant mortality which massively reduced life expectancy.

Question 5

```r
years <- c(1918)
for (yr in years) {
    sweden.year <- sweden |>
        filter(year == yr) |>
        mutate(
            mx = (male_deaths/male_pop),
            px = (1 - mx),
            px_lagged = lag(px, default = 1),
            lx = cumprod(px_lagged))
    cat(yr, ":", round(sum(sweden.year$lx, na.rm = TRUE), 2), "\n")}
```

```
## 1918 : 48.15
```

```r
years <- c(1918)
for (yr in years) {
    sweden.year <- sweden |>
        filter(year == yr) |>
        mutate(
            mx = (female_deaths/female_pop),
            px = (1 - mx),
            px_lagged = lag(px, default = 1),
            lx = cumprod(px_lagged))
    cat(yr, ":", round(sum(sweden.year$lx, na.rm = TRUE), 2), "\n")}
```

```
## 1918 : 51.51
```

In 1918, there was a life expectancy gender gap of approximately 3.3 years favoring women. This may be due to males working in more dangerous occupations such as manufacturing, mining, and other jobs requiring physical labor with suboptimal working conditions.

Section 2

Question 6

The pattern of life expectancy in the data shows a modest, constant increase for the first three years measured, with a massive jump in the final 100 year segment. Infectious disease often kills the very young and the very old. If a population's youngest are dying at a high rate, it will decrease the life expectancy by a large amount because of the amount of leverage they have on affecting the average life expectancy relative to one with healthier young. Some medical advances were the cause of the increased life expectancy in the first 200 or so years of the data set, but the near eradication of infectious disease death in developed countries such as Sweden meant that its population lives long enough to suffer from chronic disease, which we are not yet as good at treating, as opposed to infectious disease. The major gains in healthcare, and thus lifespan, were in the 20th century with the advent of greatly improved medical techniques and the introduction of treatments such as antibiotics and even NSAIDs.

Question 7

```
# Adding in some of my own years to compare
years <- c(1917, 1918, 1920, 2020, 2021, 2024)
for (yr in years) {
    sweden.year <- sweden |>
        filter(year == yr) |>
        mutate(
            mx = (total_deaths/total_pop),
            px = (1 - mx),
            px_lagged = lag(px, default = 1),
            lx = cumprod(px_lagged))
    cat(yr, ":", round(sum(sweden.year$lx, na.rm = TRUE), 2), "\n")}
```

```
## 1917 : 58.83
## 1918 : 49.82
## 1920 : 58.34
## 2020 : 82.63
## 2021 : 83.18
## 2024 : 83.98
```

1918 was the first year of the Influenza pandemic, also known as the Spanish flu. It killed people of all demographics, but had a particularly strong impact on young adults. While other countries were devastated by World War I, Sweden remained neutral, giving us a good baseline to look at in the years preceding and following the Great War. Populations are always vulnerable to epidemic disease, with the characteristics of some countries making them more or less susceptible to outbreaks than others. The life expectancy decreased by nine years during the pandemic, temporarily undoing decades of progress. Epidemics, specifically viral ones are still a massive threat to societies today. A virus has the capability to mutate and multiply faster than we can track and treat it, as proven by the COVID-19 Pandemic of 2020. Populations will likely remain vulnerable to pandemics going forward, at least until we develop much more advanced technology with the capability to quickly neutralize outbreaks.

Question 8

By describing the majority of life expectancy estimates as period measures, Deaton notes that the life expectancy numbers are more meaningful for taking a snapshot of a population at a moment in time. If

an individual is younger, as society changes, they may have a much different life expectancy than could be estimated at the time. One effective tool when looking backwards is the lifespan data for all individuals born in a specific year. For example, life expectancy in Russia in 1940 was 41.4 years, but by 1945, 80% of males born in 1923 were dead and Russian life expectancy was just 23.6 years [Statista, BBC]. Life expectancy is not a gauge of how long people born today, but rather a projection of the current risks and typical lifespans of all people alive at a certain point.