Are we more civilized today?

Alina Kereszt

created 04.10.2021, updated 05.10.2022

This exercise is based on the dataset provided by OurWorldInData project (https://ourworldindata.org/homicides) based at the Oxford University.

The long-term trend in Homicides in Western Europe

Understanding how homicide rates have changed prior to the modern era requires the help of historians and archivists. Manuel Eisner, a criminology professor at the University of Cambridge, and his colleagues published the Historical Violence Database a compilation of data on long-term trends in homicide rates, in addition to qualitative information such as the cause of death, perpetrator and victim. This database is limited to countries with relatively complete historical records on violence and crime – mainly Western Europe and the US.

Starting in the second half of the nineteenth century, these European regions have consistent police records of those accused of murder or manslaughter and annual counts of homicide victims. To go back further in time, reaching as far back as the thirteenth century, Eisner collected estimates (from historical records of coroner reports, court trials, and the police) of homicide rates made in over ninety publications by scholars.

Homicide rates – measured as the number of homicides per 100,000 individuals – up to 1990 are sourced from Eisner's (2003) publication and the Historical Violence Database.

Are homicide rates in Europe today lower or higher than in the past? Using the provided dataset, display and describe the long-run homicide rates for the five European regions: Italy, England, Germany, Netherlands and Scandinavia.

```
library(tidyverse)
```

Load the available data from ourworldindata.org

You should always interrogate the source of your data. Who compiled it, from where, what is missing, how representative the data are? Check the data/Metadata.txt to learn about the data provenance.

```
Western_Europe <- read_csv("data/homicide-rates-across-western-europe.csv")
```

Inspect the data

How clean and analysis-ready is the dataset? Do you understand what the column names represent? What is the difference between rate and homicide number?

```
head(Western_Europe)
```

```
## # A tibble: 6 × 4
##    Entity  Code   Year Homicide rate in Europe over long-term (per 100,000) (ho…¹
##    <chr>   <chr> <dbl>                                                     <dbl>
## 1 England <NA>   1300                                                        23
## 2 England <NA>   1550                                                         7
## 3 England <NA>   1625                                                         6
## 4 England <NA>   1675                                                         4
## 5 England <NA>   1725                                                         2
## 6 England <NA>   1775                                                         1
## # … with abbreviated variable name
## #   ¹`Homicide rate in Europe over long-term (per 100,000) (homicides per 100,000
## people)`
```

- **Entity:** name of the region
- **Code:** a short code assigned to the region
- **Year:** the year for which we have the homicide rate
- **Homicide rate:** the number of murders per 100,000 population

The difference between homicide rate and homicide number is that the homicide number is the total number of murders that took place in the region, while the homicide rate is this number adjusted for the region's population.

Ok, the data look good except for the column
`Homicide rate in Europe over long-term (per 100,000)` which is not very easy to work with.

- Use the `names()` function and assignment key to relabel this column to `homicides_per_100k`

```
names(Western_Europe)[4] <- "homicides_per_100k"
```

Now, that you have looked at what the data looks like and what it represents, and streamlined it, let's see what big picture it contains.

Let's see what the long-term trend is in homicides

- use `ggplot()` function and remember the `+` at the end of the line
- chose a meaningful `geom_......()` for geometry (hint: points are not great)
- load `Year` on the `x` axis and `homicides_per_100k` column in y axis
- to color individual country entries consistently, assign the country column to the argument

color .
- provide meaningful title and axis labels
- remember to change the `eval` flag so that the code chunk renders when knitted

```
ggplot(data = Western_Europe,
       aes(x = Year,
           y = homicides_per_100k,
           color = Entity)) +
   geom_smooth() +
   labs(title = "Homicide rates in Western Europe 1300-2016",
        x = "Year",
        y = "Homicides per 100.000 population") +
   theme_minimal()
```
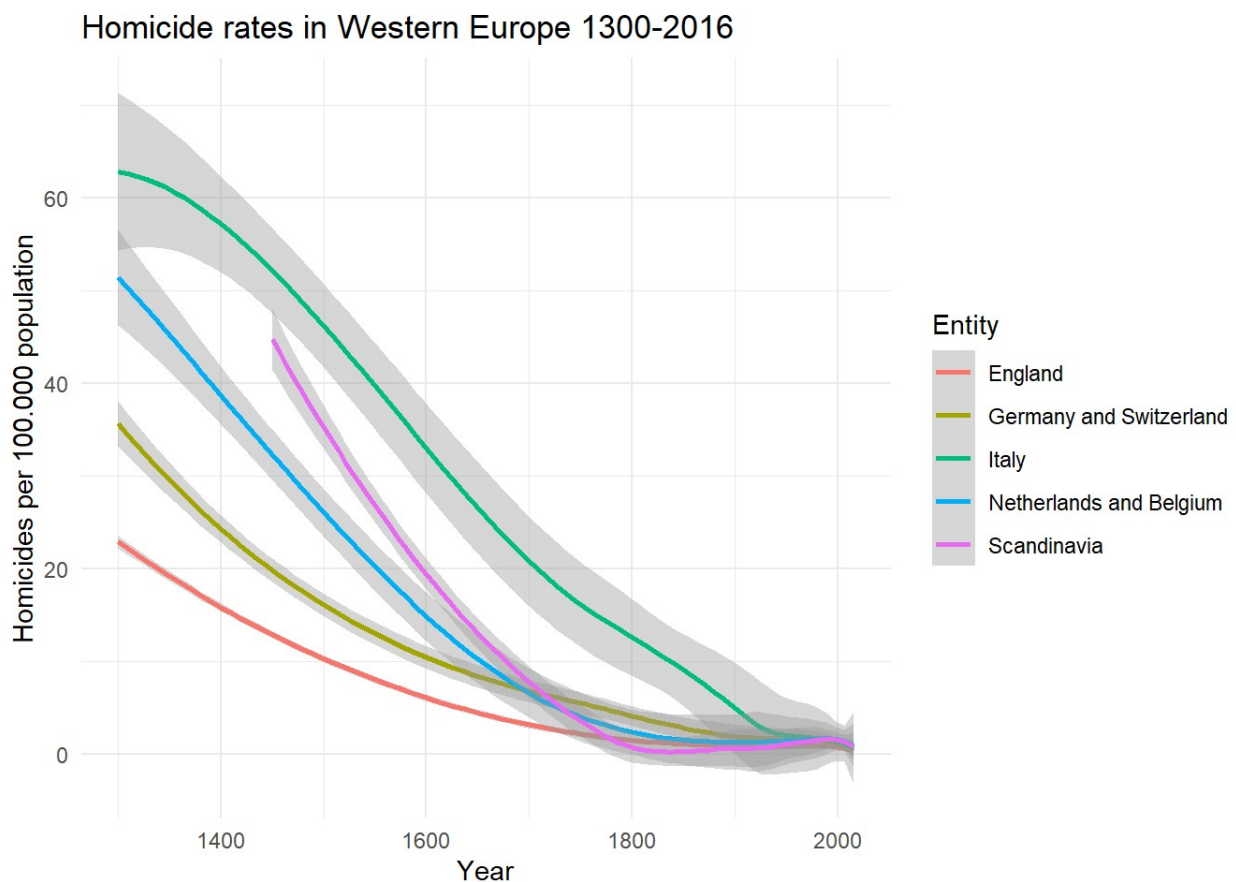

The homicide rate has descended over time for all regions, converging to a very similar homicide rate in the present.

Let's check the rates for individual countries.

Uncouple the homicides of individual countries for easier view

You can visualize each country's trend separately by adding an extra argument to the ggplot, the `facet_wrap()` and feeding it the country column. If in doubt, check your ggplot tutorial and your country column name for exact usage.

- reuse the ggplot from the chunk above
- insert `facet_wrap()` after the specification of geometry to split countries in separate charts
- change the facet "layout" to two columns and three rows so that the trends are easier to see in

horizontal layout.

Homicide rates in Western Europe 1300-2016
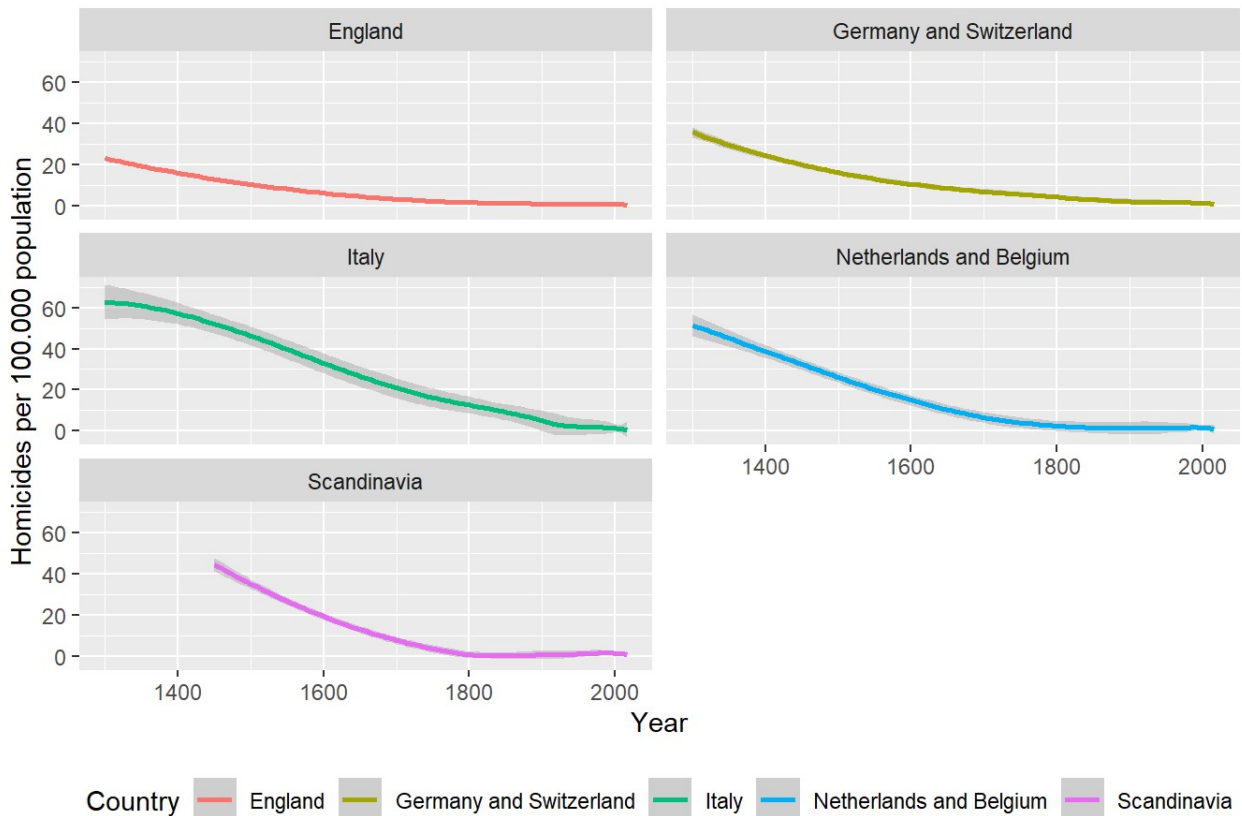
Compare the trends in homicide with the pattern of reign duration among Danish rulers through time.

- Load your Danish king dataset. Hopefully it is tidy and your years and duration of reign are all numeric.
- You need to have a consistent way of plotting the rulers' reign on the x axis, so I recommend you create a midyear column by calculating the middle of each monarch's rule (Hint: `midyear = endyear - ( endyear - startyear )/2`)
- Start a ggplot plotting midyear on x axis and duration on y axis
- Try `geom_smooth()` for geometry
- Provide meaningful labels and a title
- How would you characterize the trend compared to the homicides above?

```
# load data
kings <- read_csv2("data/kings.csv")

# create variable for reign length
kings$reign_length <- kings$end_reign - kings$start_reign

# find midpoint of reign for each monarch and round year to a whole number
kings$midyear <- round(kings$end_reign - (kings$reign_length / 2),
                       digits = 0)

# plot
ggplot(kings,
       aes(x = midyear,
           y = reign_length))+
  geom_smooth()+
  labs(title = "Duration of reign of Danish monarchs",
       x = "Midpoint of reign",
       y = "Duration of reign") +
  theme_minimal()
```


Duration of reign of Danish monarchs
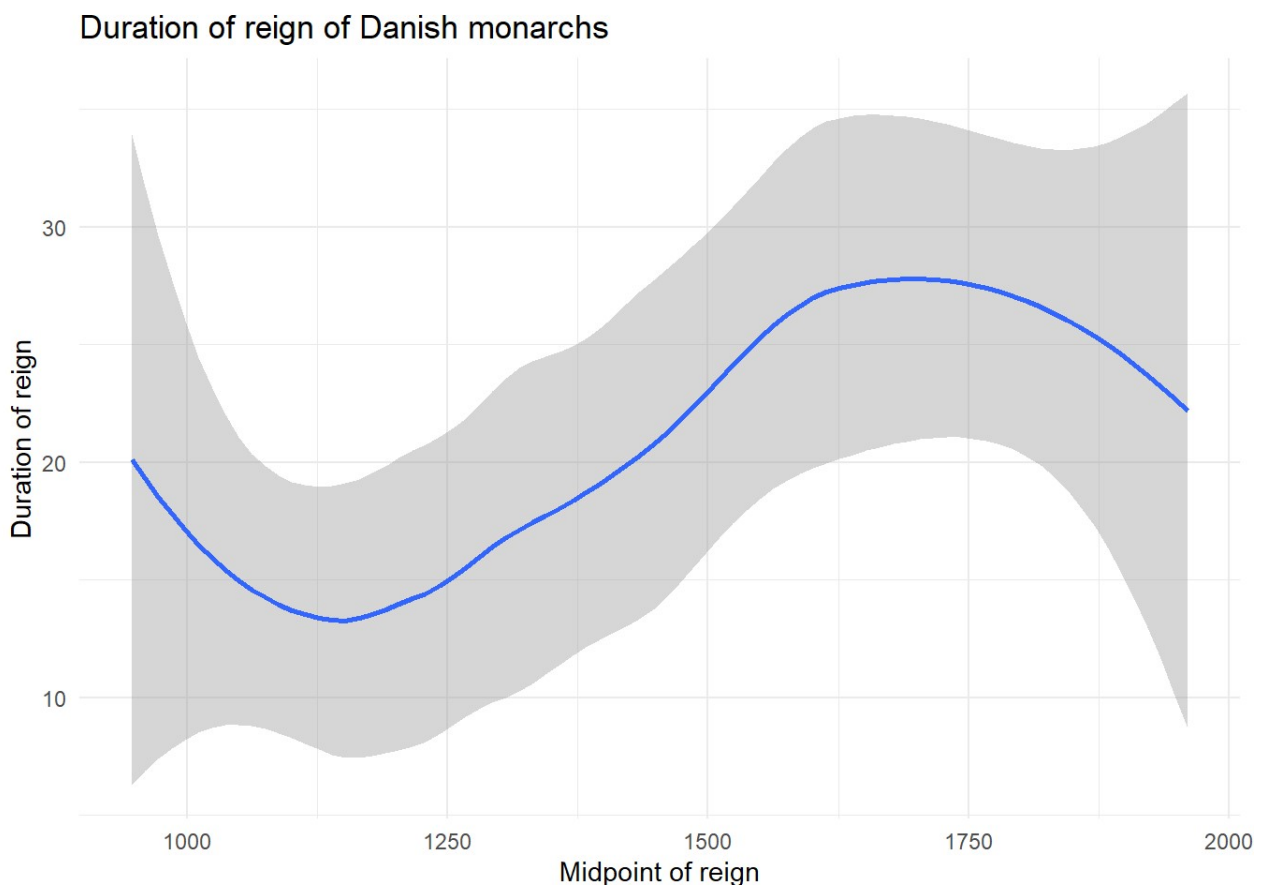
Overall, there seems to be a tendency for Danish monarchs to reign longer as history progresses. Queen Margrethe II, who has reigned for 50 years and will hopefully reign for many more, is not yet included in this plot, seeing as her reign has yet to end - however, she will confirm this trend of longer reign.

Correlation does not equal causation, though, and in order to uncover any potential causal relationships between reign lengths and homicide rate, further scientific inquiry would be necessary.

Final tasks:

1. Plot: In the facetted plot above, move the legend from the current position on the side to below the facets, and label it "Country" instead of "Entity".

2. Rmarkdown:

- edit the author of the document, and convert 'Final Tasks' into heading #2 (like the other headings)
- add a `floating table of contents` to your Rmarkdown document,
- provide informative `chunk-names` and edit flags in your R chunks, and
- automatically generate a `timestamp` to show when the document was last updated. (Hint: check the Rmarkdown episode in our Data Carpentry tutorial)

3. Question: In <250 words articulate your answer on the basis of the data visualisations to the following question: are we more civilized today?

If we define 'civilized' as 'at an advanced stage of social and cultural development' (Oxford Languages, retrieved 05.10.2022), one could argue that indeed we are. The data used in this homework assignment is limited in its scope, but a few inferences are readily available to be made.

The declining homicide rate is in itself a positive development, seeing as one is less likely to be murdered. However, it can also imply a more mentally healthy population as well as the presence of norms and institutions that allow the resolution of interpersonal conflict through non-violent means. Meanwhile, though the continuing existence of monarchies is dubious in its positivity, the lengthening reign of Danish monarchs can imply that people live longer and that subjects are less likely to shorten their sovereign's reign by assassinating them prematurely.