Rstudioreport_Final

Abhinaya

2025-11-30

Aim

To examine the trend in age–sex standardised Coronary Heart Disease hospitalisation rates in Scotland between 2016/17–2018/19 and 2021/22–2023/24 according to the rolling averages

Research Question

"How have age–sex standardised CHD hospitalisation rates in Scotland changed between 2016/17–2018/19 and 2021/22–2023/24 according to the 3-year rolling averages"

Load Packages

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.5.2

## Warning: package 'ggplot2' was built under R version 4.5.2

## Warning: package 'tibble' was built under R version 4.5.2

## Warning: package 'tidyr' was built under R version 4.5.2

## Warning: package 'readr' was built under R version 4.5.2

## Warning: package 'purrr' was built under R version 4.5.2

## Warning: package 'dplyr' was built under R version 4.5.2

## Warning: package 'forcats' was built under R version 4.5.2

## Warning: package 'lubridate' was built under R version 4.5.2
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.6
## v forcats   1.0.1     v stringr   1.5.2
## v ggplot2   4.0.1     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.2.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(readr)
library(tidyr)
library(dplyr)
library(here)
```

```
## Warning: package 'here' was built under R version 4.5.2
```

```
## here() starts at C:/Users/abhih/OneDrive/Documents/GitHub/Rstudioreport-Abhinaya
```

Read in data

The dataset used for the study is the coronary heart disease (CHD) patient hospitalisations, in Scotland, for the years 2006-2022, from the Scottish Public Health Observatory Online profiles Tool.

Numerator= Number of patients admitted due to coronary heart disease each year

Denominator = Total population each year.

Measure = Rate of admissions per 100,000 persons in the population.

```r
CHD_total <- read_csv("ScotPHO_CHD_total.csv")
```

```
## Rows: 20 Columns: 12
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (7): area_code, area_type, area_name, period, type_definition, indicator...
## dbl (5): year, numerator, measure, upper_confidence_interval, lower_confiden...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
glimpse(CHD_total)
```

```
## Rows: 20
## Columns: 12
## $ area_code                 <chr> "S00000001", "S00000001", "S00000001", "S000~
## $ area_type                 <chr> "Scotland", "Scotland", "Scotland", "Scotlan~
## $ area_name                 <chr> "Scotland", "Scotland", "Scotland", "Scotlan~
## $ year                      <dbl> 2003, 2004, 2005, 2006, 2007, 2008, 2009, 20~
## $ period                    <chr> "2002/03 to 2004/05 financial years; 3-year ~
```

```
## $ type_definition           <chr> "Age-sex standardised rate per 100,000", "Ag~
## $ indicator                 <chr> "Coronary heart disease (CHD) patient hospit~
## $ numerator                 <dbl> 27182.0, 26465.0, 25603.3, 24778.3, 23717.0,~
## $ measure                   <dbl> 637.6, 614.5, 587.3, 561.0, 529.3, 500.4, 47~
## $ upper_confidence_interval <dbl> 645.6, 622.2, 594.8, 568.3, 536.4, 507.2, 48~
## $ lower_confidence_interval <dbl> 629.8, 606.8, 579.8, 553.8, 522.4, 493.7, 47~
## $ data_source               <chr> "Public Health Scotland (SMR01)", "Public He~
```

```
view(CHD_total)
```

Select relevant columns (year, period, numerator, measure)

```
CHD_tidy <- CHD_total %>% select (year, period, numerator, measure) %>%  filter( year >= 2017) %>% renar
```

##Renaming the variables to shorten the names in "Period" column

```
CHD_tidy <- CHD_tidy %>% mutate(period = gsub(" financial years; 3-year aggregates","", period )) %>% mu
```

Summary statistics of the data

```
summary(CHD_tidy)
```

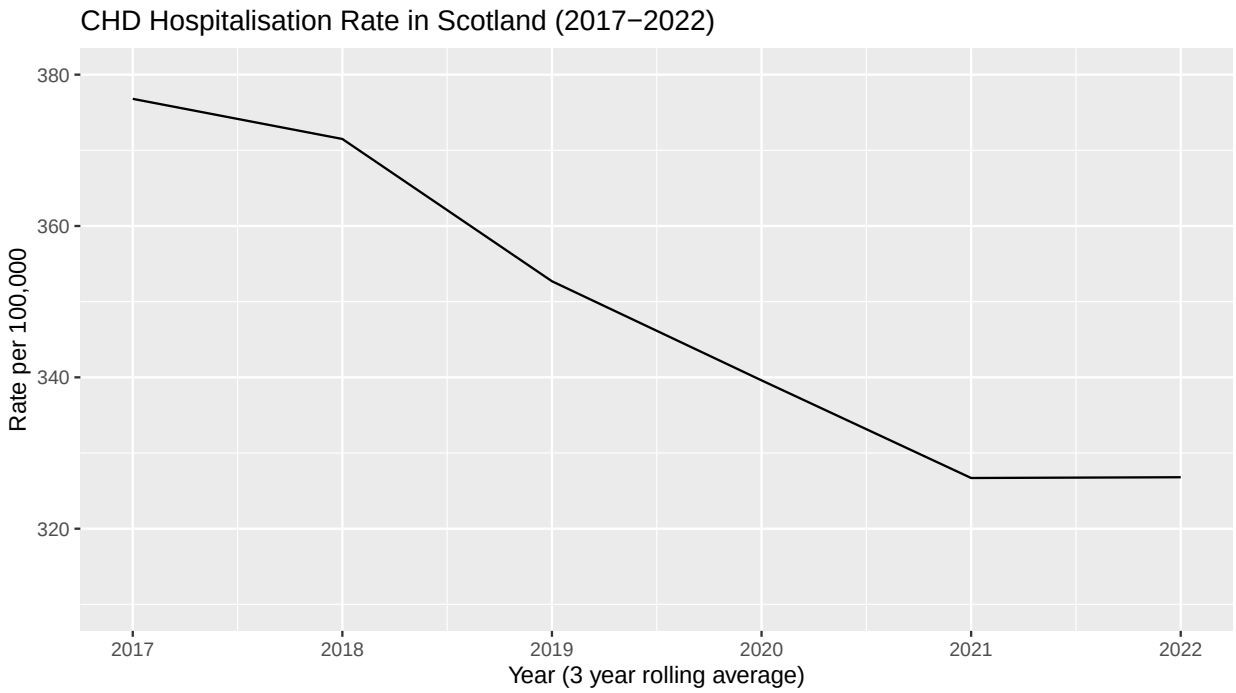
```
##       year         period          Hosp_per_year   HospRate_100k
##  Min.   :2017   Length:6           Min.   :17952   Min.   :326.7
##  1st Qu.:2018   Class :character   1st Qu.:18253   1st Qu.:330.0
##  Median :2020   Mode  :character   Median :18676   Median :346.1
##  Mean   :2020                      Mean   :18801   Mean   :349.0
##  3rd Qu.:2021                      3rd Qu.:19463   3rd Qu.:366.8
##  Max.   :2022                      Max.   :19667   Max.   :376.8
```

```
view(CHD_tidy)
```

Plot line graph

Plotting a line graph to show the trend line of the change in hospitalisation rates over the years.

```
ggplot(CHD_tidy, aes(x = year, y = HospRate_100k)) +
  labs(title = "CHD Hospitalisation Rate in Scotland (2017-2022)",
       x = "Year (3 year rolling average)",
       y = "Rate per 100,000") +
  geom_line() +
  coord_cartesian(ylim = c(310, 380))
```



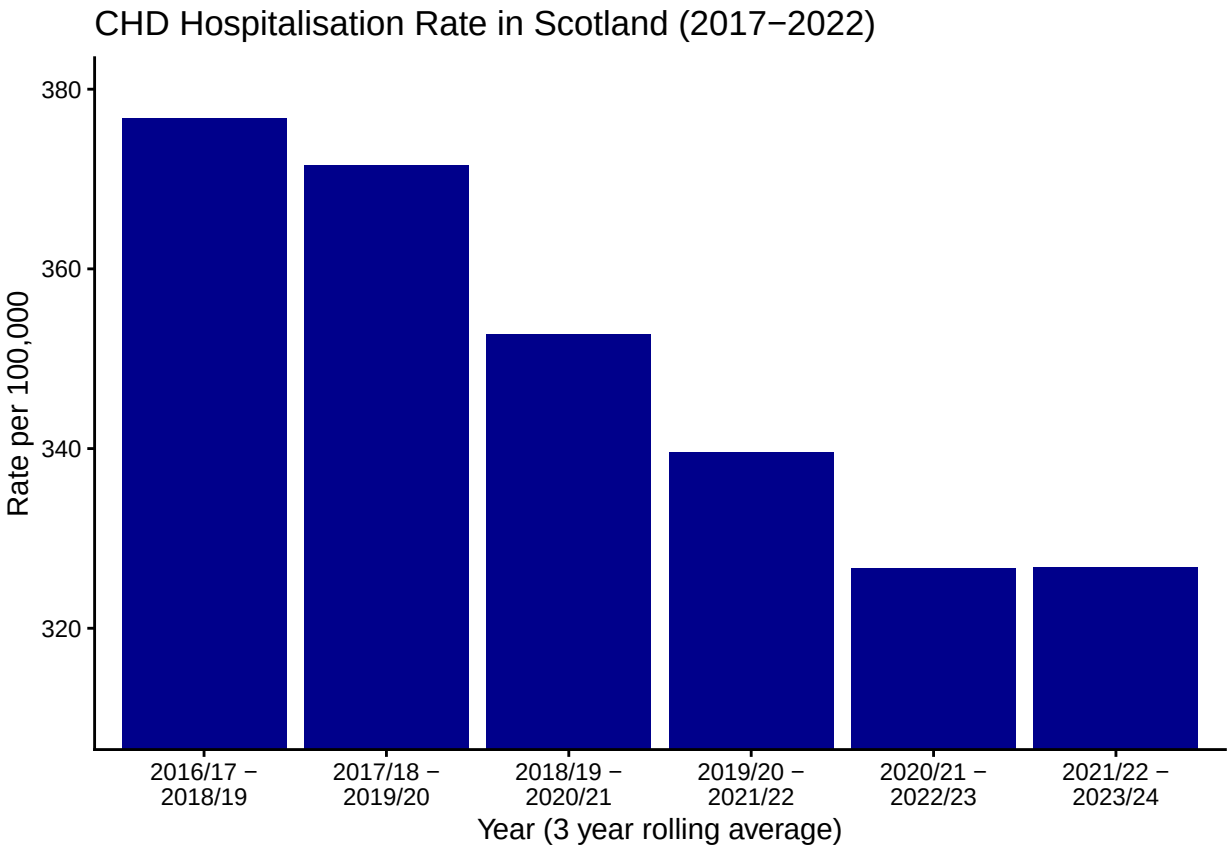

```
data_frame()
```

```
## Warning: ‘data_frame()‘ was deprecated in tibble 1.1.0.
## i Please use ‘tibble()‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```

```
## # A tibble: 0 x 0
```

Plot a bar graph

```
ggplot(CHD_tidy, aes(x= str_wrap(period, 10), y= HospRate_100k)) + geom_col(fill = "blue4") + labs(titl
       x = "Year (3 year rolling average)",
       y = "Rate per 100,000") + theme_classic() +
  coord_cartesian(ylim = c(310, 380))
```


CHD Hospitalisation Rate in Scotland (2017−2022)
Rate per 100,000
380
360
340
320
2016/17 − 2018/19
2017/18 − 2019/20
2018/19 − 2020/21
2019/20 − 2021/22
2020/21 − 2022/23
2021/22 − 2023/24
Year (3 year rolling average)