Proof of concept for health indicators

Julian Flowers

2024-07-12

Table of contents

# 1. Introduction

Outline an end-to-end process for creating public health indicators and generating public health profiles.

## 1.1 Global workflow

|  |
| --- |
| Figure 1.1: Workflow |

# 2. Indicator definitions

These are as extracted from the definitions as agreed for the proof of concept

## 2.1 Anti-microbial resistance

Percentage of bloodstream infections due to selected antimicrobial-resistant organisms (MRSA, Escherichia coli resistant to 3rd-generation cephalosporin (e.g., ESBL- E. coli))

### 2.1.1 Staphylococcus aureus

| title | denominator | numerator | metric | notes |
| --- | --- | --- | --- | --- |
| Percentage of bloodstream infections due to selected antimicrobial-resistant organisms: MRSA | Total No. of patients with growth of S. aureus in tested blood samples | No. of patients with growth of methicillin-resistant S. aureus in tested blood samples | Proportion |  |

### 2.1.2 Escherichia coli

| title | denominator | numerator | metric | notes |
| --- | --- | --- | --- | --- |
| Percentage of bloodstream infections due to selected antimicrobial-resistant organisms: E.coli | Total No. of patients with growth of E. coli in tested blood samples | No. of patients with growth of E.coli resistant to 3rd generation cephalosporins in tested blood samples | Proportion | Resistance to ceftazidine, cefotaxine |

## 2.2 Smoking

Current cigarette smoking among women aged 18-44 years

| title | denominator | numerator | metric | notes |
| --- | --- | --- | --- | --- |
| PCurrent cigarette smoking among women aged 18-44 years | Women aged 18-44 years who information about cigarette smoking (excluding unknowns and refusals). | Women aged 18-44 years who reported that they smoked ≥100 cigarettes in their lifetime and currently smoke every day or some days. | Rate with 95% confidence intervals | Data provided is attendances at smoking clinics rather than survey based smoking prevalence |

## 2.3 Injury

Nonfatal hospitalizations for all injuries

| title | denominator | numerator | metric | notes |
| --- | --- | --- | --- | --- |
| Nonfatal hospitalizations for all injuries | Midyear population for the calendar year under surveillance obtained from the KSA Census or suitable alternative. | Nonfatal hospitalizations with any of the following ICD-10-CM diagnostic codes in the principal diagnosis field | Rate with 95% confidence intervals | Data provided is attendances at smoking clinics rather than survey based smoking prevalence |

## 2.4 Flu vaccination coverage

### 2.4.1 Paediatric

### 2.4.2 65+

### 2.4.3 At-risk groups

# 3. Data summaries

## 3.1 AMR

The dummy data is per patient sample based.

### 3.1.1 E. coli

|  |
| --- |
| Table 3.1: E. coli resistance rates to 3rd generation cephalosporins |

| age\_band | resistant | sensitive | total | sum\_res | sum\_tot | mean\_rate | value | lowercl | uppercl | confidence | method |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [0,5) | 8 | 6 | 14 | 83 | 152 | 0.546 | 0.571 | 0.3259 | 0.786 | 95% | Wilson |
| [10,15) | 1 |  |  | 83 | 152 | 0.546 |  |  |  | 95% | Wilson |
| [15,20) | 1 | 2 | 3 | 83 | 152 | 0.546 | 0.333 | 0.0615 | 0.792 | 95% | Wilson |
| [20,25) | 2 | 1 | 3 | 83 | 152 | 0.546 | 0.667 | 0.2077 | 0.939 | 95% | Wilson |
| [25,30) | 1 |  |  | 83 | 152 | 0.546 |  |  |  | 95% | Wilson |
| [30,35) | 3 | 3 | 6 | 83 | 152 | 0.546 | 0.500 | 0.1876 | 0.812 | 95% | Wilson |
| [35,40) | 2 | 4 | 6 | 83 | 152 | 0.546 | 0.333 | 0.0968 | 0.700 | 95% | Wilson |
| [40,45) | 1 | 3 | 4 | 83 | 152 | 0.546 | 0.250 | 0.0456 | 0.699 | 95% | Wilson |
| [45,50) | 1 | 2 | 3 | 83 | 152 | 0.546 | 0.333 | 0.0615 | 0.792 | 95% | Wilson |
| [50,55) | 7 | 3 | 10 | 83 | 152 | 0.546 | 0.700 | 0.3968 | 0.892 | 95% | Wilson |
| [55,60) | 9 | 6 | 15 | 83 | 152 | 0.546 | 0.600 | 0.3575 | 0.802 | 95% | Wilson |
| [60,65) | 11 | 9 | 20 | 83 | 152 | 0.546 | 0.550 | 0.3421 | 0.742 | 95% | Wilson |
| [65,70) | 9 | 7 | 16 | 83 | 152 | 0.546 | 0.562 | 0.3318 | 0.769 | 95% | Wilson |
| [70,75) | 5 | 8 | 13 | 83 | 152 | 0.546 | 0.385 | 0.1771 | 0.645 | 95% | Wilson |
| [75,80) | 5 | 4 | 9 | 83 | 152 | 0.546 | 0.556 | 0.2667 | 0.811 | 95% | Wilson |
| [80,85) | 10 | 6 | 16 | 83 | 152 | 0.546 | 0.625 | 0.3864 | 0.815 | 95% | Wilson |
| [85,90) | 2 | 5 | 7 | 83 | 152 | 0.546 | 0.286 | 0.0822 | 0.641 | 95% | Wilson |
| [90,95) | 5 | 2 | 7 | 83 | 152 | 0.546 | 0.714 | 0.3589 | 0.918 | 95% | Wilson |
| [95,100) |  | 1 |  | 83 | 152 | 0.546 |  |  |  | 95% | Wilson |

|  |
| --- |
| Figure 3.1: e. coli resistance |

### 3.1.2 S.aureus

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3.2: MRSA   | age\_band | resistant | sensitive | total | sum\_res | sum\_tot | mean\_rate | value | lowercl | uppercl | confidence | method | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | [0,5) | 4 |  |  | 33 | 50 | 0.66 |  |  |  | 95% | Wilson | | [10,15) | 1 | 1 | 2 | 33 | 50 | 0.66 | 0.500 | 0.0945 | 0.905 | 95% | Wilson | | [15,20) |  | 2 |  | 33 | 50 | 0.66 |  |  |  | 95% | Wilson | | [20,25) |  | 1 |  | 33 | 50 | 0.66 |  |  |  | 95% | Wilson | | [25,30) |  | 2 |  | 33 | 50 | 0.66 |  |  |  | 95% | Wilson | | [30,35) | 3 | 3 | 6 | 33 | 50 | 0.66 | 0.500 | 0.1876 | 0.812 | 95% | Wilson | | [35,40) | 4 | 1 | 5 | 33 | 50 | 0.66 | 0.800 | 0.3755 | 0.964 | 95% | Wilson | | [40,45) |  | 1 |  | 33 | 50 | 0.66 |  |  |  | 95% | Wilson | | [45,50) | 1 | 2 | 3 | 33 | 50 | 0.66 | 0.333 | 0.0615 | 0.792 | 95% | Wilson | | [50,55) | 1 | 1 | 2 | 33 | 50 | 0.66 | 0.500 | 0.0945 | 0.905 | 95% | Wilson | | [55,60) | 2 | 6 | 8 | 33 | 50 | 0.66 | 0.250 | 0.0715 | 0.591 | 95% | Wilson | | [60,65) | 6 | 1 | 7 | 33 | 50 | 0.66 | 0.857 | 0.4869 | 0.974 | 95% | Wilson | | [65,70) | 1 | 1 | 2 | 33 | 50 | 0.66 | 0.500 | 0.0945 | 0.905 | 95% | Wilson | | [70,75) | 4 | 1 | 5 | 33 | 50 | 0.66 | 0.800 | 0.3755 | 0.964 | 95% | Wilson | | [75,80) | 1 | 2 | 3 | 33 | 50 | 0.66 | 0.333 | 0.0615 | 0.792 | 95% | Wilson | | [80,85) | 4 | 1 | 5 | 33 | 50 | 0.66 | 0.800 | 0.3755 | 0.964 | 95% | Wilson | | [85,90) | 1 | 1 | 2 | 33 | 50 | 0.66 | 0.500 | 0.0945 | 0.905 | 95% | Wilson | |

|  |
| --- |
| Figure 3.2: Staph aureus resistance rates |

## 3.2 Flu vaccination coverage

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3.3: Counts of flu vaccination by age group, gender and region   | age\_groups | gender | Al Baha | Al Jawf | Al Qassim | Asir | Hail | Jazan | Makkah Al Mukarramah | Najran | Northern Frontier | Riyadh | Sharqiya | Tabuk | madina | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 0-18 | female | 38 | 1 | 50 | 161 | 34 | 37 | 282 | 18 | 1 | 361 | 31 | 19 | 140 | | 0-18 | male | 2 |  | 18 | 33 | 6 | 20 | 188 | 2 |  | 122 | 23 | 45 | 91 | | 65+ | female | 1 |  |  | 49 | 3 | 2 | 64 | 3 |  | 19 | 1 | 6 | 24 | | 65+ | male |  |  |  | 10 |  | 4 | 29 | 5 |  | 13 |  | 2 | 14 | | other | female | 69 |  | 157 | 753 | 167 | 105 | 1,182 | 48 | 2 | 1,260 | 102 | 127 | 765 | | other | male | 9 | 1 | 74 | 295 | 37 | 120 | 938 | 50 | 1 | 536 | 199 | 229 | 443 | |

## 3.3 Smoking

Counts of female smoking clinic attendees by directorate for 15+ and 18-44 years

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3.4: Counts of female attendees at smoking cessation clinics   | age\_1844 | directorate\_name | female | | --- | --- | --- | | 1 | AlAhsa | 73 | | 1 | Asir | 9 | | 1 | Baha | 28 | | 1 | Bisha | 6 | | 1 | Eastern | 551 | | 1 | Hafer AlBatin | 60 | | 1 | Hail | 22 | | 1 | Jazan | 145 | | 1 | Jeddah | 1,569 | | 1 | Jouf | 23 | | 1 | Madinah | 205 | | 1 | Makkah | 578 | | 1 | Najran | 82 | | 1 | Northern Borders | 4 | | 1 | Qassim | 22 | | 1 | Qunfotha | 13 | | 1 | Riyadh | 1,830 | | 1 | Tabuk | 291 | | 1 | Taif | 100 |  | age\_15 | directorate\_name | female | | --- | --- | --- | | 1 | AlAhsa | 93 | | 1 | Asir | 84 | | 1 | Baha | 28 | | 1 | Bisha | 6 | | 1 | Eastern | 579 | | 1 | Hafer AlBatin | 60 | | 1 | Hail | 22 | | 1 | Jazan | 158 | | 1 | Jeddah | 1,788 | | 1 | Jouf | 28 | | 1 | Madinah | 303 | | 1 | Makkah | 636 | | 1 | Najran | 85 | | 1 | Northern Borders | 7 | | 1 | Qassim | 48 | | 1 | Qunfotha | 13 | | 1 | Qurayyat | 4 | | 1 | Riyadh | 1,862 | | 1 | Tabuk | 436 | | 1 | Taif | 136 | |

## 3.4 Injury

|  |
| --- |
| Figure 3.3: Alluvial plot of counts of hospitalisation for injury be age, gender and ICD10 code |

# 4. Getting started

## 4.1 Getting started

These pages set out the process for loading, cleaning, reshaping and recoding provided datasets for calculating indicator values as a proof of concept, and creating a health profile from those indicators.

To facilitate reproducibility the analyses are written in R code in the form of a notebook and R scripts which do much of the pre-processing. The full code is made available as a Github repository at <https://github.com/julianflowers/poc>. The files can be cloned (downloaded)

* In RStudio, go to “File > New Project”
* Click on “Version Control:
* Checkout a project from a version control repository”
* Click on “Git:
* Clone a project from a repository”
* Fill in the info: URL: use HTTPS address
* Create as a subdirectory of: Browse to where you would like to create this folder

### 4.1.1 The indicators

4 sets of indicators are used for this PoC:

* Methicillin resistant Staph aureus and 3rd generation cephalosporin resistant E.coli - proportion of samples tested which are resistant
* Flu vaccination coverage rates
* Smoking rates in women
* Injury admission rates

Fully specifying the indicators

1. The proportion of blood samples which grow either Staph aureaus or E. coli, which are tested for antibiotic sensitivity and which are found to be resistant to oxacllin or 3rd generation cephalosporins respectively for time period X to Y, stratified by [area] / [time period] / [age group] / [gender]
2. The proportion of the population which has been vaccinated against flu for the time period X to Y stratified by [area] / [time period] / [age group] / [gender]
3. The rate of smoking in women aged 15+ / 18-44, per 100,000 female population for the time period X to Y stratified by [area] / [time period] / [age group] / [gender]
4. The rate of hospitalisation for injury for the time period X to Y stratified by [area] / [time period] / [age group] / [gender]

### 4.1.2 Pre-processing script

As a first step we will run a script which does a number of things.

1. Imports the datasets for each indicator
2. Loads KSA 2022 census data by age, gender and region[[1]](#footnote-61)
3. Makes variable names consistent
4. Recodes region names so that they match the names used in the Census 2022 data
5. Maps directorate names (used in smoking data) to region[[2]](#footnote-63)
6. Ensures numeric variables (e.g. age) are converted to numbers
7. Creates a set of intermediate data tables
8. Saves file of reshaped population data,

The intermediate tables can be reused for further analysis and generating profiles.

The script can be run by typing

```  
`source("~/proof-of-concept/scripts/pre-process.R")`  
  
```

at the prompt in the console

This generates a set of objects in the R environment

objects()

[1] "age\_names" "area\_names"   
 [3] "census\_names" "cnames"   
 [5] "csv" "csvs"   
 [7] "data" "dfs"   
 [9] "dfs\_3" "dfs\_4"   
[11] "dfs1" "dfs2\_agg"   
[13] "dfs3" "f\_r"   
[15] "final\_poc\_data" "gender\_names"   
[17] "get\_coordinates\_from\_google\_maps" "m\_r"   
[19] "pop\_agg" "pops"   
[21] "reg\_dir\_lu" "regional\_counts"   
[23] "regional\_counts\_complete" "RETICULATE\_PYTHON"   
[25] "sa\_bound" "sa\_shp"   
[27] "sa\_shp\_1" "sc\_coords"   
[29] "sc\_dir\_links" "sc\_dir\_names"   
[31] "sc\_ll" "sc\_ll\_sf"   
[33] "sc\_loc" "scc\_dir"   
[35] "shps" "smok\_1"   
[37] "smok\_agg" "tmpd"   
[39] "url" "xl"   
[41] "xls"

Objects called dfs reflect pre-processing of the original datasets. Objects containing names, contain the different variables used for age, gender and area variables across datasets. Those called sc are used to map locations of smoking clinics as part of recoding directorates to regions for the smoking data (see below).

The object regional\_counts\_complete is a data frame of regional age\_band, gender specific counts for each indicator and forms the basis of indicator generation. Note, this includes region-age-band-gender combinations for which there is no data (because these combinations are not present in the original data - although they maybe in the full datasets).

regional\_counts\_complete |>  
 head() |>  
 flextable::flextable()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Region | age\_band | sum\_f\_Smoking | sum\_f\_flu | sum\_f\_injury | sum\_m\_Smoking | sum\_m\_flu | sum\_m\_injury | | --- | --- | --- | --- | --- | --- | --- | --- | | `Asir | [ 0, 5) |  | 140 | 160 |  | 26 | 0 | | `Asir | [ 5, 10) | 0 | 7 |  | 1 | 1 |  | | `Asir | [ 10, 15) | 5 | 4 |  | 355 | 2 |  | | `Asir | [ 15, 20) | 39 | 13 |  | 2,744 | 4 |  | | `Asir | [ 20, 25) |  | 79 |  |  | 25 |  | | `Asir | [ 25, 30) | 1 | 76 |  | 21 | 22 |  |   Table 4.1: Region age-band specific counts by gender and indicator |

### 4.1.3 Mapping health directorates to regions for smoking data

The smoking data provided is clinic based data disaggregated at the level of health directorate. There are 20 directorates in KSA, and 13 regions.

Available population data is at regional level, so to generate population denominators for the smoking data I have mapped directorates to regions as follows:

* I used a dataset which contained spatial locations of healthcare facilities (smoking cessation clinics) at health directorate level[[3]](#footnote-66)
* For each location I extracted spatial coordinates (longitude and latitude)
* I obtained a regional boundary file (shape file) from <https://data.humdata.org/dataset/41ce9023-1d21-4549-a485-94316200aba0/resource/99834c81-ad34-415e-91c5-af053d8e55b4/download/sau_capp_adm1_1m_ocha.zip>
* I spatially joined the clinic location and boundary files (see
* This created a lookup table for directorates and regions and enabled the smoking data to be recoded to regions and calculation of rates using the census regional population estimates

sc\_ll\_sf |>  
 ggplot() +  
 geom\_sf(data = sa\_bound) +  
 geom\_sf(aes(colour = name))

|  |
| --- |
| Figure 4.1: Location map of smoking clinics mapped to KSA regional boundaries |

# 5. Population demography

## 5.1 Population pyramids

|  |  |  |
| --- | --- | --- |
| |  | | --- | | (a) Absolute population |  |  | | --- | | (b) Relative population |   Figure 5.1: Population pyramids |

# 6. AMR walkthrough

## 6.1 Introduction

This document outlines a stepwise approach to calculating AMR indicators from dummy data kindly supplied by PHA.

There are x steps

1. EDA (exploratory data analysis of raw data) - this involves cleaning, visualisation and creation of relevant variables.
2. Review of indicator definitions
   * Numerator
   * Denominator
3. Method for calculating numerator and denominator values from dataset. The outline uses R code for reproducibility and flexibility.
4. Calculating indicator values and uncertainty intervals
5. Suggested indicator visualisations (if appropriate).

## 6.2 AMR indicators

### 6.2.1 MRSA

Percentage of bloodstream infection due to methicillin-resistant Staphylococcus aureus (MRSA)

Numerator: No. of patients with growth of methicillin-resistant S. aureus in tested blood samples

Denominator: Total No. of patients with growth of S. aureus in tested blood samples

### 6.2.2 E. coli

Percentage of bloodstream infection due to 3rd-generation cephalosporin resistant E. coli

Numerator: No. of patients with growth of 3rd-generation cephalosporin resistant E. coli in tested blood samples

Denominator: Total No. of patients with growth of E. coli in tested blood samples

### 6.2.3 Import data

df <- amr

334 observations

## 6.3 Data preparation

### 6.3.1 calculate 5-year age bands

amr <- amr[, `:=` (age\_band = cut(age\_year, breaks = seq(0, 100, 5), right = FALSE))][]  
  
head(amr)

record\_number sample\_no patient\_mrn location  
 <num> <char> <char> <char>  
1: 1 ###### ##### Outpatient  
2: 17 ###### ##### Inpatient  
3: 20 ###### ##### Inpatient  
4: 25 ###### ##### Inpatient  
5: 43 ###### ##### Outpatient  
6: 63 ###### ##### Outpatient  
 patient\_hospitalized  
 <char>  
1: Patient had NOT been admitted for more than 2 days in the past 30 days  
2: Patient has been hospitalized for 2 days or less  
3: Patient has been hospitalized for more than 2 days  
4: Patient has been hospitalized for 2 days or less  
5: Patient had NOT been admitted for more than 2 days in the past 30 days  
6: Patient had NOT been admitted for more than 2 days in the past 30 days  
 specific\_location age\_year community\_origin site first\_name second\_name  
 <char> <num> <char> <char> <char> <char>  
1: Emergency Room 0 Community Origin Blood #### #####  
2: Intensive Care Unit 71 Community Origin Blood #### #####  
3: Intensive Care Unit 44 Hospital Origin Blood #### #####  
4: Intensive Care Unit 67 Community Origin Blood #### #####  
5: Emergency Room 67 Community Origin Blood #### #####  
6: Emergency Room 92 Community Origin Blood #### #####  
 family\_name national\_iqama\_id nationality pathogen\_name minocycline  
 <char> <char> <char> <char> <lgcl>  
1: #### ########## ##### Escherichia coli NA  
2: #### ########## ##### Escherichia coli NA  
3: #### ########## ##### Escherichia coli NA  
4: #### ########## ##### Escherichia coli NA  
5: #### ########## ##### Escherichia coli NA  
6: #### ########## ##### Escherichia coli NA  
 tigecycline ampicillin penicillin\_g oxacillin cefoxitin cefotaxime  
 <lgcl> <char> <lgcl> <char> <char> <char>  
1: NA R NA <NA> <NA> R  
2: NA R NA <NA> <NA> NA  
3: NA S NA <NA> <NA> S  
4: NA R NA <NA> <NA> R  
5: NA R NA <NA> <NA> NA  
6: NA R NA <NA> <NA> NA  
 ceftazidime ceftriaxone cefixime cefepime doripenem ertapenem imipenem  
 <char> <char> <lgcl> <char> <char> <char> <char>  
1: R R NA R NA S S  
2: S S NA S NA S S  
3: S S NA S NA S S  
4: R R NA R NA S S  
5: I S NA S NA S S  
6: R R NA R R S S  
 meropenem co\_trimoxazole azithromycin amikacin gentamicin ciprofloxacin  
 <char> <char> <lgcl> <lgcl> <lgcl> <char>  
1: S S NA NA NA S  
2: S S NA NA NA S  
3: S S NA NA NA S  
4: S R NA NA NA S  
5: S S NA NA NA S  
6: S R NA NA NA R  
 levofloxacin colistin spectinomycin age\_band  
 <char> <char> <lgcl> <fctr>  
1: S NA NA [0,5)  
2: S S NA [70,75)  
3: S NA NA [40,45)  
4: S NA NA [65,70)  
5: S S NA [65,70)  
6: R S NA [90,95)

### 6.3.2 remove non-relevant data

This step removes identifiers (names, record IDs)

amr <- amr |> select(-c(family\_name, first\_name, sample\_no, patient\_mrn, second\_name, national\_iqama\_id, nationality))

### 6.3.3 create per test file (long data)

* this create a *per test* dataset rather than a per patient sample dataset

amr\_long <- amr |>  
 pivot\_longer(names\_to = "antibiotic", values\_to = "resistance", cols = minocycline:spectinomycin) |> setDT()

### 6.3.4 Recode 3rd generation cephalosporins

* this step adds a new variable which labels 3rd generation cephalosporins

amr\_long <- amr\_long[, gen\_3 := case\_when(str\_detect(antibiotic, "cef") ~ "3rd-gen", TRUE ~ "other")][]

## 6.4 Data summarisation and description (EDA)

* first generate a high level tabular summary

gtsummary::tbl\_summary(amr)

* represent this visually - we’ll use decomposition trees

amr\_freq <- amr\_long[pathogen\_name == "Escherichia coli", .N, by = .(age\_band, gen\_3, resistance, pathogen\_name, community\_origin)]  
  
collapsibleTreeSummary(amr\_freq,   
 c( "community\_origin", "gen\_3","resistance"),   
 root = "E. coli",   
 nodeSize = "N",   
 attribute = "N",   
 fontSize = 16,   
 collapsed = FALSE)

|  |
| --- |
| Figure 6.1: Decomposition tree for E. coli |

## 6.5 Numerators and denominators

To calculate indicators we need to calculate

* patients with blood stream infection
* samples with antibiotic resistance

amr\_long

Index: <pathogen\_name>  
 record\_number location  
 <num> <char>  
 1: 1 Outpatient  
 2: 1 Outpatient  
 3: 1 Outpatient  
 4: 1 Outpatient  
 5: 1 Outpatient  
 ---   
7678: 1210 Inpatient  
7679: 1210 Inpatient  
7680: 1210 Inpatient  
7681: 1210 Inpatient  
7682: 1210 Inpatient  
 patient\_hospitalized  
 <char>  
 1: Patient had NOT been admitted for more than 2 days in the past 30 days  
 2: Patient had NOT been admitted for more than 2 days in the past 30 days  
 3: Patient had NOT been admitted for more than 2 days in the past 30 days  
 4: Patient had NOT been admitted for more than 2 days in the past 30 days  
 5: Patient had NOT been admitted for more than 2 days in the past 30 days  
 ---   
7678: Patient has been hospitalized for more than 2 days  
7679: Patient has been hospitalized for more than 2 days  
7680: Patient has been hospitalized for more than 2 days  
7681: Patient has been hospitalized for more than 2 days  
7682: Patient has been hospitalized for more than 2 days  
 specific\_location age\_year community\_origin site pathogen\_name  
 <char> <num> <char> <char> <char>  
 1: Emergency Room 0 Community Origin Blood Escherichia coli  
 2: Emergency Room 0 Community Origin Blood Escherichia coli  
 3: Emergency Room 0 Community Origin Blood Escherichia coli  
 4: Emergency Room 0 Community Origin Blood Escherichia coli  
 5: Emergency Room 0 Community Origin Blood Escherichia coli  
 ---   
7678: Non Intensive Unit 96 Hospital Origin Blood Staphylococcus aureus  
7679: Non Intensive Unit 96 Hospital Origin Blood Staphylococcus aureus  
7680: Non Intensive Unit 96 Hospital Origin Blood Staphylococcus aureus  
7681: Non Intensive Unit 96 Hospital Origin Blood Staphylococcus aureus  
7682: Non Intensive Unit 96 Hospital Origin Blood Staphylococcus aureus  
 age\_band antibiotic resistance gen\_3  
 <fctr> <char> <char> <char>  
 1: [0,5) minocycline <NA> other  
 2: [0,5) tigecycline <NA> other  
 3: [0,5) ampicillin R other  
 4: [0,5) penicillin\_g <NA> other  
 5: [0,5) oxacillin <NA> other  
 ---   
7678: [95,100) gentamicin <NA> other  
7679: [95,100) ciprofloxacin R other  
7680: [95,100) levofloxacin R other  
7681: [95,100) colistin NA other  
7682: [95,100) spectinomycin <NA> other

## 6.6 Calculate resistance rates

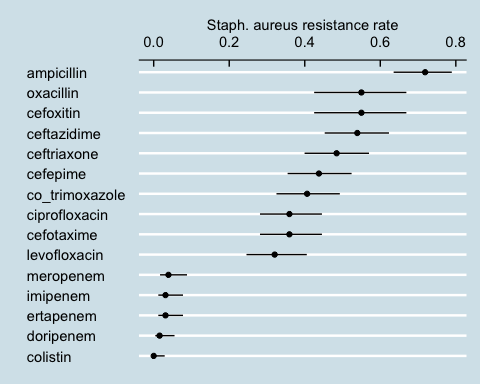
* calculate proportion of tests resistant
* calculate confidence interval (using Wilson’s score method for proportions via the PHEindicatormethods R package)

amr\_long[pathogen\_name == "Escherichia coli" & !is.na(resistance), .N, by = .(resistance, gen\_3)] |>  
 pivot\_wider(names\_from = resistance, values\_from = N) |>  
 rowwise() |>  
 mutate(total\_tests = sum(c\_across(R:I), na.rm = TRUE),   
 resistance\_rate = R / total\_tests) |>  
 flextable::flextable()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 6.1: E. coli resistance rates to 3rd generation cephalosporins   | gen\_3 | R | NA | S | I | total\_tests | resistance\_rate | | --- | --- | --- | --- | --- | --- | --- | | other | 300 | 193 | 646 | 7 | 1,146 | 0.2617801 | | 3rd-gen | 269 | 56 | 192 | 4 | 521 | 0.5163148 | |

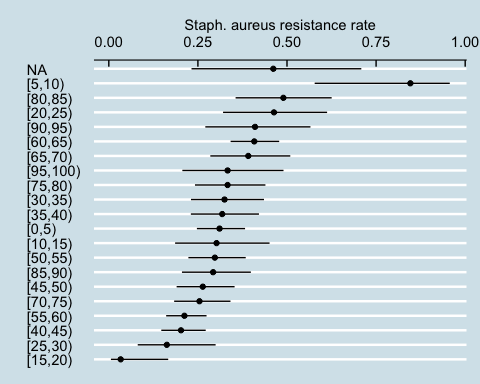
### 6.6.1 by antibiotic

options(digits = 2)  
  
amr\_res\_ci\_sa <- amr\_long[pathogen\_name == "Staphylococcus aureus" & !is.na(resistance), .N, by = .(antibiotic, resistance)] |>  
 pivot\_wider(names\_from = resistance, values\_from = N, values\_fill = 0) |>  
 rowwise() |>  
 mutate(total\_tests = sum(c\_across(S:I), na.rm = TRUE),   
 resistance\_rate = R / total\_tests)  
  
phe\_proportion(amr\_res\_ci\_sa, R, total\_tests) |>  
 bind\_cols(amr\_res\_ci\_sa) |>  
 ggplot() +  
 geom\_point(aes(reorder(antibiotic, value), value)) +  
 geom\_linerange(aes(antibiotic, ymin = lowercl, ymax = uppercl)) +  
 coord\_flip() +  
 labs(y = "Staph. aureus resistance rate", x = "") +   
 scale\_y\_continuous(position = "right")



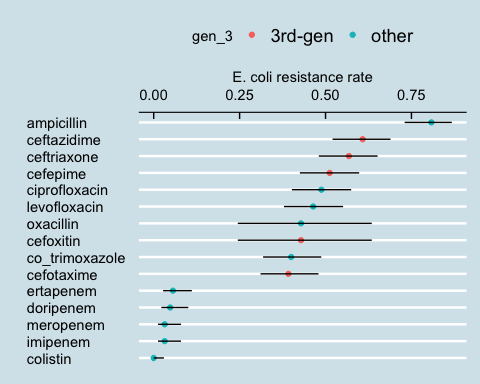
### 6.6.2 by age

amr\_res\_ci\_age <- amr\_long[pathogen\_name == "Staphylococcus aureus" & !is.na(resistance), .N, by = .(age\_band, resistance)] |>  
 pivot\_wider(names\_from = resistance, values\_from = N, values\_fill = 0) |>  
 rowwise() |>  
 mutate(total\_tests = sum(c\_across(S:I), na.rm = TRUE),   
 resistance\_rate = R / total\_tests)  
  
phe\_proportion(amr\_res\_ci\_age, R, total\_tests) |>  
 bind\_cols(amr\_res\_ci\_age) |>  
 ggplot() +  
 geom\_point(aes(reorder(age\_band, value), value)) +  
 geom\_linerange(aes(age\_band, ymin = lowercl, ymax = uppercl)) +  
 coord\_flip() +  
 labs(y = "Staph. aureus resistance rate", x = "") +   
 scale\_y\_continuous(position = "right")



## 6.7 E. coli

amr\_res\_ci\_ec <- amr\_long[str\_detect(pathogen\_name, "coli") & !is.na(resistance), .N, by = .(antibiotic, resistance, gen\_3)] |>  
 pivot\_wider(names\_from = resistance, values\_from = N, values\_fill = 0) |>  
 rowwise() |>  
 mutate(total\_tests = sum(c\_across(R:I), na.rm = TRUE),   
 resistance\_rate = R / total\_tests)  
  
phe\_proportion(amr\_res\_ci\_ec, R, total\_tests) |>  
 bind\_cols(amr\_res\_ci\_ec) |>  
 ggplot() +  
 geom\_point(aes(reorder(antibiotic, value), value, colour = gen\_3)) +  
 geom\_linerange(aes(antibiotic, ymin = lowercl, ymax = uppercl)) +  
 coord\_flip() +  
 labs(y = "E. coli resistance rate", x = "") + scale\_y\_continuous(position = "right")



# 7. Smoking

## 7.1 Workflow

|  |
| --- |
| Figure 7.1: Proposed workflow for calculating smoking rates |

### 7.1.1 Load census population and create age bands

## create age band  
  
pops <- setDT(pops)[, `:=` (`18-44` = dplyr::between(age, 18, 44), `15+` = age >= 15, `80+` = age >= 80, age\_band = cut(age, seq(0, 110, 5), right = FALSE))][]  
  
pop\_age <- pops[, .(n = .N, sumpops = sum(Population, na.rm = TRUE)), by = .(Region, age\_band, Gender)][order(age\_band)]  
  
pop1844F <- pops[, .(n = .N, sumpops = sum(Population, na.rm = TRUE)), by = .(Region, `18-44`, Gender)][`18-44` == "TRUE" & Gender == "Female",]  
  
pop15F <- pops[, .(n = .N, sumpops = sum(Population, na.rm = TRUE)), by = .(Region, `15+`, Gender)][`15+` == "TRUE" & Gender == "Female",]

### 7.1.2 load smoking data and calculate age bands

smok\_age <- smoking[, `:=` (`18-44` = dplyr::between(age, 18, 44), `15+` = age >= 15, `80+` = age >= 80, age\_band = cut(age, seq(0, 110, 5), right = FALSE))][]  
  
smok\_age\_bands <- smok\_age[, .(n = .N, smokers = sum(n, na.rm = TRUE)), by = .(Region, age\_band, Gender)][order(age\_band)]  
  
smok1844F <- smok\_age[, .(n = .N, smokers = sum(n, na.rm = TRUE)), by = .(Region, `18-44`, Gender)][`18-44` == "TRUE" & Gender == "female",]  
  
smok15F <- smok\_age[, .(n = .N, smokers = sum(n, na.rm = TRUE)), by = .(Region, `15+`, Gender)][`15+` == "TRUE" & Gender == "female",]

### 7.1.3 Join datasets and calculate rates

This step uses Byar’s method for confidence interval for rates

smok\_rates <- pop1844F |>  
 left\_join(smok1844F, by = "Region") |>   
 phe\_rate(x = smokers, n = sumpops)   
  
smok\_rates |>  
 dplyr::select(Region, smokers, population = sumpops, value, lowercl, uppercl) |>  
 flextable::flextable()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Region | smokers | population | value | lowercl | uppercl | | --- | --- | --- | --- | --- | --- | | 'Asir |  | 366,324 |  |  |  | | Al Bahah | 28 | 61,969 | 45.183882 | 30.017729 | 65.30573 | | Al Hudud ash Shamaliyah | 4 | 70,149 | 5.702148 | 1.553643 | 14.59976 | | Al Jawf | 23 | 103,536 | 22.214495 | 14.077587 | 33.33412 | | Al Madinah al Munawwarah |  | 387,944 |  |  |  | | Al Mintaqah ash Sharqiyah |  | 877,403 |  |  |  | | Al Qasim |  | 249,081 |  |  |  | | Ar Riyadh |  | 1,609,493 |  |  |  | | Ha'il | 22 | 136,642 | 16.100467 | 10.086571 | 24.37741 | | Jazan |  | 267,342 |  |  |  | | Makkah al Mukarramah |  | 1,507,301 |  |  |  | | Najran | 82 | 105,915 | 77.420573 | 61.572876 | 96.10050 | | Tabuk | 291 | 163,453 | 178.032829 | 158.162585 | 199.70848 |   Figure 7.2: Smoking clinic attendance rates (18-44) |

smok\_rates |>  
 write\_csv("data/smok\_rates.csv")  
   
smok\_rates |>  
 ggplot() +  
 geom\_col(aes(reorder(Region, value), value)) +  
 geom\_linerange(aes(x = Region, ymin = lowercl, ymax = uppercl)) +  
 coord\_flip() +  
 labs(x = " ",  
 y = "Rate per 100.000")

|  |
| --- |
| Figure 7.3: Smoking clinic attendance rates (18-44) |

# 8. Flu vaccination coverage

## 8.1 Workflow

|  |
| --- |
| Figure 8.1: Proposed workflow for calculating flu vaccination coverage |

## 8.2 Clean data

Recode region names Add new age bands

flu\_data <- flu\_data |>  
 mutate(paed = dplyr::between(age, 0, 18),   
 older = age >= 65) |>  
 mutate(Region = case\_when(str\_detect(Region, "Baha") ~ "Al Bahah",   
 str\_detect(Region, "Qu?assim") ~ "Al Qasim",  
 str\_detect(Region, "Asir") ~ "`Asir",  
 str\_detect(Region, "Hail") ~ "Ha'il",  
 str\_detect(Region, "[Mm]adin") ~ "Al Madinah al Munawwarah",  
 str\_detect(Region, "Jizan") ~ "Jazan",  
 str\_detect(Region, "Makka") ~ "Makkah al Mukarramah",  
 str\_detect(Region, "Sharqiya" ) ~ "Al Mintaqah ash Sharqiyah",  
 str\_detect(Region, "Northern Frontier") ~ "Al Hudud ash Shamaliyah",  
 str\_detect(Region, "Riyad") ~ "Ar Riyadh",  
 TRUE ~ Region ))

## 8.3 Calculate regional numerators and denominators

* Calculate region-by-gender populations for 0-18 and 65+
* Calculate region-by-gender vaccinations for 0-18 and 65+

pop\_paed <- popns |>  
 group\_by(Region, Gender, paed) |>  
 reframe(pop = sum(Population)) |>  
 filter(paed == "TRUE") |>  
 mutate(Gender = recode(Gender, "Female" = "female", "Male" = "male"),   
 Region = recode(Region, "`Asir" = "'Asir"))  
  
pop\_older <-popns |>  
 group\_by(Region, Gender, older) |>  
 reframe(pop = sum(Population)) |>  
 filter(older == "TRUE") |>  
 mutate(Gender = recode(Gender, "Female" = "female", "Male" = "male"),   
 Region = recode(Region, "`Asir" = "'Asir"))  
  
## fills all region-gender categories and remove unknown  
paed\_flu <- flu\_data |>  
 filter(paed == "TRUE") |>  
 count(Region, Gender) |>  
 complete(Region, Gender) |>  
 filter(Gender != "unknown") |>  
 mutate( Region = recode(Region, "`Asir" = "'Asir"))  
  
older\_flu <- flu\_data |>  
 filter(older == "TRUE") |>  
 count(Region, Gender) |>  
 complete(Region, Gender) |>  
 filter(Gender != "unknown") |>  
 mutate( Region = recode(Region, "`Asir" = "'Asir"))

## 8.4 Join populations and vaccination data

paed\_nd <- paed\_flu |>  
 left\_join(pop\_paed) |>  
 select(-paed)  
  
older\_nd <- older\_flu |>  
 left\_join(pop\_older) |>  
 select(-older)

## 8.5 Calculate coverage by age group

paed\_coverage <- paed\_nd |>  
 phe\_proportion(x = n, n = pop) |>  
 mutate(age = "0-18")   
  
older\_coverage <- older\_nd |>  
 phe\_proportion(x = n, n = pop) |>  
 mutate(age = "65+")  
  
bind\_rows(paed\_coverage, older\_coverage) |>  
 write\_csv("data/flu\_coverage.csv")

## 8.6 Tabulate

options(digits = 3)  
  
paed\_coverage |>  
 select(-c(statistic, method)) |>  
 head() |>  
 flextable()  
older\_coverage |>  
 select(-c(statistic, method)) |>  
 head() |>  
 flextable()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 8.1: Flu vaccination coverage   | Region | Gender | n | pop | value | lowercl | uppercl | confidence | age | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 'Asir | female | 161 | 335,011 | 0.0004806 | 0.00041187 | 0.0005607 | 95% | 0-18 | | 'Asir | male | 33 | 347,751 | 0.0000949 | 0.00006757 | 0.0001333 | 95% | 0-18 | | Al Bahah | female | 38 | 57,230 | 0.0006640 | 0.00048382 | 0.0009112 | 95% | 0-18 | | Al Bahah | male | 2 | 59,077 | 0.0000339 | 0.00000928 | 0.0001234 | 95% | 0-18 | | Al Hudud ash Shamaliyah | female | 1 | 70,350 | 0.0000142 | 0.00000251 | 0.0000805 | 95% | 0-18 | | Al Hudud ash Shamaliyah | male |  | 73,068 |  |  |  | 95% | 0-18 |  | Region | Gender | n | pop | value | lowercl | uppercl | confidence | age | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 'Asir | female | 49 | 32,224 | 0.001521 | 0.0011505 | 0.002010 | 95% | 65+ | | 'Asir | male | 10 | 36,288 | 0.000276 | 0.0001497 | 0.000507 | 95% | 65+ | | Al Bahah | female | 1 | 7,791 | 0.000128 | 0.0000227 | 0.000727 | 95% | 65+ | | Al Bahah | male |  | 6,752 |  |  |  | 95% | 65+ | | Al Madinah al Munawwarah | female | 24 | 31,099 | 0.000772 | 0.0005187 | 0.001148 | 95% | 65+ | | Al Madinah al Munawwarah | male | 14 | 31,231 | 0.000448 | 0.0002671 | 0.000752 | 95% | 65+ | |

## 8.7 Visualise

paed\_coverage |>  
 ggplot() +  
 geom\_col(aes(Region, value, fill = Gender)) +  
 geom\_linerange(aes(Region, ymin = lowercl, ymax = uppercl)) +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1) ) +  
 facet\_wrap(~ Gender) +  
 scale\_fill\_manual(values = c("blue", "red"))  
  
  
older\_coverage |>  
 ggplot() +  
 geom\_col(aes(Region, value, fill = Gender)) +  
 geom\_linerange(aes(Region, ymin = lowercl, ymax = uppercl)) +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1) ) +  
 facet\_wrap(~ Gender) +  
 scale\_fill\_manual(values = c("blue", "red"))

|  |  |  |
| --- | --- | --- |
| |  | | --- | | (a) Paediatric coverage |  |  | | --- | | (b) 65+ coverage |   Figure 8.2: Flu vaccination coverage |

# 9. Injury

## 9.1 Injury

options(digits = 2)  
summary(injury)

DateOfBirth code diagnosis   
 Min. :1939-08-16 00:00:00.000 Length:10380 Length:10380   
 1st Qu.:1970-10-10 00:00:00.000 Class :character Class :character   
 Median :1985-11-10 00:00:00.000 Mode :character Mode :character   
 Mean :1984-01-01 12:27:30.296   
 3rd Qu.:1996-07-05 00:00:00.000   
 Max. :2022-01-30 00:00:00.000   
 NA's :773   
 Gender Region   
 Length:10380 Length:10380   
 Class :character Class :character   
 Mode :character Mode :character

obs <- nrow(injury)  
na\_dob <- injury[is.na(DateOfBirth), ]  
na\_dob\_n <- nrow(na\_dob)

Overall there are 10380 records, of which 773 (7.45% ) have missing dates of of birth.

injury[, .N, by = .(diagnosis, code, Gender, Region )]

diagnosis  
 <char>  
 1: Crushing injury of ankle and foot  
 2: Fracture of femur  
 3: Superficial injury of forearm  
 4: Dislocation, sprain and strain of joints and ligaments at ankle and foot level  
 5: Intracranial injury  
 6: Fracture of forearm  
 7: Fracture of forearm  
 8: Certain early complications of trauma, not elsewhere classified  
 9: Injury of unspecified body region  
10: Fracture of neck  
11: Fracture of skull and facial bones  
12: Open wound of thorax  
13: Open wound of shoulder and upper arm  
14: Fracture of femur  
15: Fracture of forearm  
16: Fracture of femur  
17: Certain early complications of trauma, not elsewhere classified  
18: Fracture of rib(s), sternum and thoracic spine  
19: Fracture of femur  
20: Burn body region unspecified  
21: Fracture of lower leg, including ankle  
22: Fracture of foot, except ankle  
23: Toxic effect of contact with venomous animals  
24: Other and unspecified injuries of neck  
25: Fracture of lower leg, including ankle  
26: Dislocation, sprain and strain of joints and ligaments at neck level  
27: Burns classified according to extent of body surface involved  
28: Other specified complications of trauma  
29: Fracture of lower leg, including ankle  
30: Burn of shoulder and upper limb, except wrist and hand  
31: Fracture of skull and facial bones  
32: Burn of shoulder and upper limb, except wrist and hand  
33: Open wound of head  
34: Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances  
35: Fracture of lower leg, including ankle  
36: Sequelae of injuries of head  
37: Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances  
38: Burn of wrist and hand  
39: Fracture of shoulder and upper arm  
40: Burns classified according to extent of body surface involved  
41: Fracture of skull and facial bones  
42: Superficial injury of wrist and hand  
43: Fracture of shoulder and upper arm  
44: Burn body region unspecified  
45: Injury of muscle and tendon at wrist and hand level  
46: Intracranial injury  
47: Foreign body in alimentary tract  
48: Other specified complications of trauma  
49: Fracture of lower leg, including ankle  
50: Complications of procedures, not elsewhere classified  
51: Fracture of shoulder and upper arm  
52: Intracranial injury  
53: Crushing injury of ankle and foot  
54: Fracture of femur  
55: Fracture of spine, level unspecified  
56: Dislocation, sprain and strain of joints and ligaments at ankle and foot level  
57: Fracture of lumbar spine and pelvis  
58: Dislocation, sprain and strain of joints and ligaments at wrist and hand level  
59: Injury of other and unspecified intrathoracic organs  
60: Burn body region unspecified  
61: Superficial injury of hip and thigh  
62: Crushing injuries involving multiple body regions  
63: Burns classified according to extent of body surface involved  
64: Open wound of head  
65: Other injuries involving multiple body regions, not elsewhere classified  
66: Fracture of forearm  
67: Intracranial injury  
68: Burn of head and neck  
69: Crushing injury of wrist and hand  
70: Toxic effect of contact with venomous animals  
71: Other and unspecified injuries of neck  
72: Other specified complications of trauma  
73: Open wound of head  
74: Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances  
75: Burn of wrist and hand  
76: Fracture of skull and facial bones  
77: Fracture of lower leg, including ankle  
78: Fracture of spine, level unspecified  
79: Fracture of rib(s), sternum and thoracic spine  
80: Fracture of skull and facial bones  
81: Intracranial injury  
82: Fracture of lumbar spine and pelvis  
83: Certain early complications of trauma, not elsewhere classified  
 diagnosis  
 code Gender Region N  
 <char> <char> <char> <int>  
 1: S97 Male Makkah 140  
 2: S72 Female Makkah 240  
 3: S50 Male Makkah 300  
 4: S93 Male Riyadh 160  
 5: S06 Male Makkah 400  
 6: S52 Male Madinah 258  
 7: S52 Female Makkah 360  
 8: T79 Male Makkah 160  
 9: T14 Male Madinah 160  
10: S12 Male Riyadh 160  
11: S02 Male Riyadh 160  
12: S21 Male Riyadh 160  
13: S41 Male Riyadh 160  
14: S72 Male Madinah 225  
15: S52 Male Makkah 180  
16: S72 Male Makkah 180  
17: T79 Male Madinah 59  
18: S22 Male Madinah 66  
19: S72 Female Riyadh 600  
20: T30 Male Madinah 500  
21: S82 Male Madinah 167  
22: S92 Male Makkah 160  
23: T63 Male Madinah 58  
24: S19 Male Madinah 58  
25: S82 Male Makkah 160  
26: S13 Male Riyadh 280  
27: T31 Male Riyadh 280  
28: T89 Male Riyadh 100  
29: S82 Female Riyadh 280  
30: T22 Male Riyadh 280  
31: S02 Male Makkah 187  
32: T22 Male Makkah 160  
33: S01 Male Northern Frontier 58  
34: T50 Male Northern Frontier 57  
35: S82 Male Sharqiya 160  
36: T90 Female Makkah 160  
37: T50 Female Madinah 160  
38: T23 Male Madinah 58  
39: S42 Male Najran 160  
40: T31 Male Makkah 160  
41: S02 Male Al Baha 58  
42: S60 Female Al Jawf 160  
43: S42 Female Al Qassim 160  
44: T30 Female Asir 160  
45: S66 Male Hail 160  
46: S06 Male Jazan 140  
47: T18 Male Madinah 20  
48: T89 Male Makkah 20  
49: S82 Female Makkah 60  
50: T81 Male Madinah 20  
51: S42 Female Makkah 20  
52: S06 Female Makkah 20  
53: S97 Male Madinah 20  
54: S72 Female Madinah 135  
55: T08 Male Madinah 7  
56: S93 Male Makkah 20  
57: S32 Male Makkah 27  
58: S63 Female Makkah 20  
59: S27 Male Makkah 20  
60: T30 Male Makkah 20  
61: S70 Male Makkah 20  
62: T04 Male Makkah 20  
63: T31 Male Madinah 20  
64: S01 Male Makkah 20  
65: T06 Male Makkah 20  
66: S52 Female Madinah 122  
67: S06 Male Madinah 7  
68: T20 Male Makkah 20  
69: S67 Male Makkah 20  
70: T63 Female Madinah 102  
71: S19 Female Madinah 102  
72: T89 Female Riyadh 180  
73: S01 Female Northern Frontier 102  
74: T50 Female Northern Frontier 103  
75: T23 Female Madinah 102  
76: S02 Female Al Baha 102  
77: S82 Female Madinah 13  
78: T08 Female Madinah 13  
79: S22 Female Madinah 114  
80: S02 Female Makkah 13  
81: S06 Female Madinah 13  
82: S32 Female Makkah 13  
83: T79 Female Madinah 101  
 code Gender Region N

unique(injury$Region)

[1] "Makkah" "Riyadh" "Madinah"   
 [4] "Northern Frontier" "Sharqiya" "Najran"   
 [7] "Al Baha" "Al Jawf" "Al Qassim"   
[10] "Asir" "Hail" "Jazan"

The data contains values for 12 of the 13 regions.

## exclude missing DoB  
injury\_dob <- injury[!is.na(DateOfBirth),]  
  
injury\_dob <- injury\_dob[!is.na(DateOfBirth), age := .(as.POSIXct(ymd("2023-06-01")) - DateOfBirth)][, age\_year := floor(as.numeric(age) / 365)][, age\_band := cut(age\_year, seq(0, 120, 5))][]

## 9.2 Aggregate injury data

injury\_dob[, .N, by = .(Region, diagnosis, code, age\_band, Gender)][order(Region)] |>  
 pivot\_wider(names\_from = Region, values\_from = N) |>  
 arrange(age\_band, code, Gender) |>  
 flextable::flextable()

| diagnosis | code | age\_band | Gender | Al Baha | Al Jawf | Al Qassim | Asir | Hail | Jazan | Madinah | Makkah | Najran | Northern Frontier | Riyadh | Sharqiya |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Crushing injury of wrist and hand | S67 | (0,5] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Burn body region unspecified | T30 | (0,5] | Female |  |  |  | 160 |  |  |  |  |  |  |  |  |
| Burn body region unspecified | T30 | (0,5] | Male |  |  |  |  |  |  | 320 | 20 |  |  |  |  |
| Fracture of skull and facial bones | S02 | (10,15] | Male |  |  |  |  |  |  |  | 20 |  |  | 160 |  |
| Fracture of forearm | S52 | (10,15] | Male |  |  |  |  |  |  | 199 |  |  |  |  |  |
| Superficial injury of wrist and hand | S60 | (10,15] | Female |  | 160 |  |  |  |  |  |  |  |  |  |  |
| Fracture of femur | S72 | (10,15] | Female |  |  |  |  |  |  | 9 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (10,15] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Crushing injury of ankle and foot | S97 | (10,15] | Male |  |  |  |  |  |  | 20 |  |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (10,15] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (10,15] | Female |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Burn body region unspecified | T30 | (10,15] | Male |  |  |  |  |  |  | 160 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (10,15] | Female |  |  |  |  |  |  |  |  |  |  | 8 |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (15,20] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Crushing injuries involving multiple body regions | T04 | (15,20] | Male |  |  |  |  |  |  |  | 19 |  |  |  |  |
| Open wound of head | S01 | (20,25] | Female |  |  |  |  |  |  |  |  |  | 17 |  |  |
| Fracture of skull and facial bones | S02 | (20,25] | Female | 19 |  |  |  |  |  |  | 1 |  |  |  |  |
| Intracranial injury | S06 | (20,25] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (20,25] | Female |  |  |  |  |  |  | 18 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (20,25] | Female |  |  |  |  |  |  | 19 |  |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (20,25] | Female |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of forearm | S52 | (20,25] | Female |  |  |  |  |  |  | 38 |  |  |  |  |  |
| Fracture of forearm | S52 | (20,25] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Fracture of femur | S72 | (20,25] | Female |  |  |  |  |  |  | 19 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (20,25] | Female |  |  |  |  |  |  | 2 | 20 |  |  |  |  |
| Crushing injuries involving multiple body regions | T04 | (20,25] | Male |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (20,25] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Injury of unspecified body region | T14 | (20,25] | Male |  |  |  |  |  |  | 160 |  |  |  |  |  |
| Foreign body in alimentary tract | T18 | (20,25] | Male |  |  |  |  |  |  | 20 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (20,25] | Female |  |  |  |  |  |  | 19 |  |  |  |  |  |
| Burns classified according to extent of body surface involved | T31 | (20,25] | Male |  |  |  |  |  |  |  |  |  |  | 280 |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (20,25] | Female |  |  |  |  |  |  | 160 |  |  | 18 |  |  |
| Toxic effect of contact with venomous animals | T63 | (20,25] | Female |  |  |  |  |  |  | 19 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (20,25] | Female |  |  |  |  |  |  | 18 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (20,25] | Female |  |  |  |  |  |  |  |  |  |  | 32 |  |
| Other specified complications of trauma | T89 | (20,25] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Open wound of head | S01 | (25,30] | Female |  |  |  |  |  |  |  |  |  | 13 |  |  |
| Open wound of head | S01 | (25,30] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of skull and facial bones | S02 | (25,30] | Female | 12 |  |  |  |  |  |  | 2 |  |  |  |  |
| Intracranial injury | S06 | (25,30] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Intracranial injury | S06 | (25,30] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Fracture of neck | S12 | (25,30] | Male |  |  |  |  |  |  |  |  |  |  | 160 |  |
| Other and unspecified injuries of neck | S19 | (25,30] | Female |  |  |  |  |  |  | 12 |  |  |  |  |  |
| Open wound of thorax | S21 | (25,30] | Male |  |  |  |  |  |  |  |  |  |  | 160 |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (25,30] | Female |  |  |  |  |  |  | 14 |  |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (25,30] | Female |  |  |  |  |  |  |  | 2 |  |  |  |  |
| Fracture of shoulder and upper arm | S42 | (25,30] | Male |  |  |  |  |  |  |  |  | 160 |  |  |  |
| Superficial injury of forearm | S50 | (25,30] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Fracture of forearm | S52 | (25,30] | Female |  |  |  |  |  |  | 12 |  |  |  |  |  |
| Fracture of femur | S72 | (25,30] | Female |  |  |  |  |  |  | 6 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (25,30] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (25,30] | Male |  |  |  |  |  |  | 160 |  |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (25,30] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Burn of head and neck | T20 | (25,30] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Burn of wrist and hand | T23 | (25,30] | Female |  |  |  |  |  |  | 4 |  |  |  |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (25,30] | Female |  |  |  |  |  |  |  |  |  | 12 |  |  |
| Toxic effect of contact with venomous animals | T63 | (25,30] | Female |  |  |  |  |  |  | 12 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (25,30] | Female |  |  |  |  |  |  | 12 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (25,30] | Female |  |  |  |  |  |  |  |  |  |  | 14 |  |
| Open wound of head | S01 | (30,35] | Female |  |  |  |  |  |  |  |  |  | 11 |  |  |
| Open wound of head | S01 | (30,35] | Male |  |  |  |  |  |  |  |  |  | 8 |  |  |
| Fracture of skull and facial bones | S02 | (30,35] | Female | 18 |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of skull and facial bones | S02 | (30,35] | Male | 10 |  |  |  |  |  |  | 1 |  |  |  |  |
| Intracranial injury | S06 | (30,35] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Intracranial injury | S06 | (30,35] | Male |  |  |  |  |  |  | 1 | 17 |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (30,35] | Female |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (30,35] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (30,35] | Female |  |  |  |  |  |  | 13 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (30,35] | Male |  |  |  |  |  |  | 9 |  |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (30,35] | Female |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (30,35] | Male |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Open wound of shoulder and upper arm | S41 | (30,35] | Male |  |  |  |  |  |  |  |  |  |  | 160 |  |
| Superficial injury of forearm | S50 | (30,35] | Male |  |  |  |  |  |  |  | 140 |  |  |  |  |
| Fracture of forearm | S52 | (30,35] | Female |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Fracture of forearm | S52 | (30,35] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Fracture of femur | S72 | (30,35] | Female |  |  |  |  |  |  | 12 |  |  |  |  |  |
| Fracture of femur | S72 | (30,35] | Male |  |  |  |  |  |  | 9 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (30,35] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (30,35] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Fracture of foot, except ankle | S92 | (30,35] | Male |  |  |  |  |  |  |  | 2 |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (30,35] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (30,35] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Burn of shoulder and upper limb, except wrist and hand | T22 | (30,35] | Male |  |  |  |  |  |  |  |  |  |  | 280 |  |
| Burn of wrist and hand | T23 | (30,35] | Female |  |  |  |  |  |  | 10 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (30,35] | Male |  |  |  |  |  |  | 9 |  |  |  |  |  |
| Burns classified according to extent of body surface involved | T31 | (30,35] | Male |  |  |  |  |  |  | 2 | 160 |  |  |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (30,35] | Female |  |  |  |  |  |  |  |  |  | 11 |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (30,35] | Male |  |  |  |  |  |  |  |  |  | 8 |  |  |
| Toxic effect of contact with venomous animals | T63 | (30,35] | Female |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Toxic effect of contact with venomous animals | T63 | (30,35] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (30,35] | Female |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (30,35] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (30,35] | Female |  |  |  |  |  |  |  |  |  |  | 26 |  |
| Other specified complications of trauma | T89 | (30,35] | Male |  |  |  |  |  |  |  |  |  |  | 14 |  |
| Intracranial injury | S06 | (35,40] | Male |  |  |  |  |  | 140 |  |  |  |  |  |  |
| Dislocation, sprain and strain of joints and ligaments at neck level | S13 | (35,40] | Male |  |  |  |  |  |  |  |  |  |  | 280 |  |
| Fracture of lumbar spine and pelvis | S32 | (35,40] | Female |  |  |  |  |  |  |  | 4 |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (35,40] | Male |  |  |  |  |  |  |  | 21 |  |  |  |  |
| Fracture of forearm | S52 | (35,40] | Male |  |  |  |  |  |  |  | 180 |  |  |  |  |
| Injury of muscle and tendon at wrist and hand level | S66 | (35,40] | Male |  |  |  |  | 160 |  |  |  |  |  |  |  |
| Fracture of femur | S72 | (35,40] | Male |  |  |  |  |  |  |  | 19 |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (35,40] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Burn of shoulder and upper limb, except wrist and hand | T22 | (35,40] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Open wound of head | S01 | (40,45] | Female |  |  |  |  |  |  |  |  |  | 8 |  |  |
| Open wound of head | S01 | (40,45] | Male |  |  |  |  |  |  |  |  |  | 2 |  |  |
| Fracture of skull and facial bones | S02 | (40,45] | Female | 1 |  |  |  |  |  |  | 5 |  |  |  |  |
| Fracture of skull and facial bones | S02 | (40,45] | Male | 2 |  |  |  |  |  |  | 1 |  |  |  |  |
| Intracranial injury | S06 | (40,45] | Female |  |  |  |  |  |  | 5 |  |  |  |  |  |
| Intracranial injury | S06 | (40,45] | Male |  |  |  |  |  |  | 1 | 142 |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (40,45] | Female |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (40,45] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (40,45] | Female |  |  |  |  |  |  | 41 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (40,45] | Male |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (40,45] | Female |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of forearm | S52 | (40,45] | Female |  |  |  |  |  |  | 39 |  |  |  |  |  |
| Fracture of forearm | S52 | (40,45] | Male |  |  |  |  |  |  | 9 |  |  |  |  |  |
| Dislocation, sprain and strain of joints and ligaments at wrist and hand level | S63 | (40,45] | Female |  |  |  |  |  |  |  | 19 |  |  |  |  |
| Superficial injury of hip and thigh | S70 | (40,45] | Male |  |  |  |  |  |  |  | 19 |  |  |  |  |
| Fracture of femur | S72 | (40,45] | Female |  |  |  |  |  |  | 11 |  |  |  |  |  |
| Fracture of femur | S72 | (40,45] | Male |  |  |  |  |  |  | 4 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (40,45] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Crushing injury of ankle and foot | S97 | (40,45] | Male |  |  |  |  |  |  |  | 139 |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (40,45] | Female |  |  |  |  |  |  | 3 |  |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (40,45] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (40,45] | Female |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (40,45] | Male |  |  |  |  |  |  | 3 |  |  |  |  |  |
| Burn body region unspecified | T30 | (40,45] | Male |  |  |  |  |  |  | 19 |  |  |  |  |  |
| Burns classified according to extent of body surface involved | T31 | (40,45] | Male |  |  |  |  |  |  | 17 |  |  |  |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (40,45] | Female |  |  |  |  |  |  |  |  |  | 8 |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (40,45] | Male |  |  |  |  |  |  |  |  |  | 3 |  |  |
| Toxic effect of contact with venomous animals | T63 | (40,45] | Female |  |  |  |  |  |  | 7 |  |  |  |  |  |
| Toxic effect of contact with venomous animals | T63 | (40,45] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (40,45] | Female |  |  |  |  |  |  | 38 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (40,45] | Male |  |  |  |  |  |  | 10 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (40,45] | Female |  |  |  |  |  |  |  |  |  |  | 8 |  |
| Other specified complications of trauma | T89 | (40,45] | Male |  |  |  |  |  |  |  |  |  |  | 1 |  |
| Open wound of head | S01 | (45,50] | Female |  |  |  |  |  |  |  |  |  | 33 |  |  |
| Open wound of head | S01 | (45,50] | Male |  |  |  |  |  |  |  |  |  | 7 |  |  |
| Fracture of skull and facial bones | S02 | (45,50] | Female | 33 |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of skull and facial bones | S02 | (45,50] | Male | 6 |  |  |  |  |  |  |  |  |  |  |  |
| Intracranial injury | S06 | (45,50] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Intracranial injury | S06 | (45,50] | Male |  |  |  |  |  |  |  | 41 |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (45,50] | Female |  |  |  |  |  |  | 33 |  |  |  |  |  |
| Other and unspecified injuries of neck | S19 | (45,50] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Fracture of rib(s), sternum and thoracic spine | S22 | (45,50] | Female |  |  |  |  |  |  | 2 |  |  |  |  |  |
| Fracture of lumbar spine and pelvis | S32 | (45,50] | Female |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of shoulder and upper arm | S42 | (45,50] | Female |  |  | 160 |  |  |  |  |  |  |  |  |  |
| Fracture of forearm | S52 | (45,50] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Dislocation, sprain and strain of joints and ligaments at wrist and hand level | S63 | (45,50] | Female |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Superficial injury of hip and thigh | S70 | (45,50] | Male |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of femur | S72 | (45,50] | Female |  |  |  |  |  |  | 34 |  |  |  |  |  |
| Fracture of femur | S72 | (45,50] | Male |  |  |  |  |  |  | 6 | 1 |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (45,50] | Female |  |  |  |  |  |  | 4 |  |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (45,50] | Male |  |  |  |  |  |  |  |  |  |  |  | 160 |
| Fracture of foot, except ankle | S92 | (45,50] | Male |  |  |  |  |  |  |  | 158 |  |  |  |  |
| Dislocation, sprain and strain of joints and ligaments at ankle and foot level | S93 | (45,50] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Crushing injury of ankle and foot | S97 | (45,50] | Male |  |  |  |  |  |  |  | 1 |  |  |  |  |
| Fracture of spine, level unspecified | T08 | (45,50] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (45,50] | Female |  |  |  |  |  |  | 33 |  |  |  |  |  |
| Burn of wrist and hand | T23 | (45,50] | Male |  |  |  |  |  |  | 6 |  |  |  |  |  |
| Burn body region unspecified | T30 | (45,50] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Burns classified according to extent of body surface involved | T31 | (45,50] | Male |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (45,50] | Female |  |  |  |  |  |  |  |  |  | 33 |  |  |
| Poisoning by diuretics and other and unspecified drugs, medicaments and biological substances | T50 | (45,50] | Male |  |  |  |  |  |  |  |  |  | 6 |  |  |
| Toxic effect of contact with venomous animals | T63 | (45,50] | Female |  |  |  |  |  |  | 33 |  |  |  |  |  |
| Toxic effect of contact with venomous animals | T63 | (45,50] | Male |  |  |  |  |  |  | 8 |  |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (45,50] | Female |  |  |  |  |  |  | 1 |  |  |  |  |  |
| Other specified complications of trauma | T89 | (45,50] | Female |  |  |  |  |  |  |  |  |  |  | 57 |  |
| Other specified complications of trauma | T89 | (45,50] | Male |  |  |  |  |  |  |  |  |  |  | 14 |  |
| Fracture of forearm | S52 | (50,55] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of femur | S72 | (50,55] | Female |  |  |  |  |  |  |  | 40 |  |  | 320 |  |
| Fracture of lower leg, including ankle | S82 | (50,55] | Female |  |  |  |  |  |  |  |  |  |  | 280 |  |
| Other injuries involving multiple body regions, not elsewhere classified | T06 | (50,55] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Complications of procedures, not elsewhere classified | T81 | (50,55] | Male |  |  |  |  |  |  | 20 |  |  |  |  |  |
| Sequelae of injuries of head | T90 | (50,55] | Female |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Certain early complications of trauma, not elsewhere classified | T79 | (55,60] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Intracranial injury | S06 | (60,65] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Intracranial injury | S06 | (60,65] | Male |  |  |  |  |  |  |  | 40 |  |  |  |  |
| Fracture of forearm | S52 | (60,65] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of femur | S72 | (60,65] | Female |  |  |  |  |  |  | 20 |  |  |  | 280 |  |
| Fracture of femur | S72 | (60,65] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Fracture of lower leg, including ankle | S82 | (60,65] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Dislocation, sprain and strain of joints and ligaments at ankle and foot level | S93 | (60,65] | Male |  |  |  |  |  |  |  |  |  |  | 160 |  |
| Fracture of skull and facial bones | S02 | (65,70] | Male |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Fracture of femur | S72 | (65,70] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Injury of other and unspecified intrathoracic organs | S27 | (70,75] | Male |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of femur | S72 | (70,75] | Female |  |  |  |  |  |  |  | 160 |  |  |  |  |
| Fracture of shoulder and upper arm | S42 | (75,80] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of femur | S72 | (75,80] | Female |  |  |  |  |  |  |  | 20 |  |  |  |  |
| Fracture of femur | S72 | (75,80] | Male |  |  |  |  |  |  | 160 |  |  |  |  |  |
| Fracture of forearm | S52 | (80,85] | Female |  |  |  |  |  |  |  | 320 |  |  |  |  |

# 10. Health profile

## 10.1 Putting it all together

Health profiles combine multiple indicators to allow comparison, usually between population or administrative units e.g. regions, hospitals, ethnic groups, age groups and provide tools to facilitate a range of comparison.

UK health profiles (for example, Public Health Outcome Framework) offer a wide range of comparison:

* Between and within areas
* Between age groups and genders

## 10.2 Visualisation

| Comparison | Visualisation | Comment |
| --- | --- | --- |
| Single area, multiple indicators | Spine or bullet chart | Indicator values are scaled so that they can be directly compared. Single area (unit) values are usually encoded as a dot or line, plotted on a bar representing the range of values for all areas / units. |
| Multiple areas, single indicator | Column / bar chart | Plots can be ordered and plotted horizontally to enable rapid appraisal of relative performance. Confidence intervals can be added |
| Indicator pairs / multiple indicators | Scatter plot or scatter plot matrices | May include smooths (lines representing fit to the data - common options are linear, loess and generalised additive models) |
| Spatial variation | Choropleth maps | Requires boundary files and GIS tools |
| Overview | Heatmap / tile chart | Colour cells with statistical significance categories (e.g. high / low) |
|  |  |  |

We will take the toolkit for plotting provided by the UKHSA which generates the charts used in the online version if the PHOF as a basis for plotting profiles. This is available at https://github.com/ukhsa-collaboration/fingertipscharts.

Note: in the PHOF statistical significance is defined by whether the global average value (e.g. KSA) is contained with the confidence limits of the unit values (e.g. ’Asir region). It does not calculate limits for the global value. The rationale for this is explained in Public Health Data Science (2018)

## install fingertipscharts  
##   
  
if(!require("fingertipscharts")) devtools::install\_github("https://github.com/ukhsa-collaboration/fingertipscharts", force = TRUE)  
  
library(fingertipscharts)

To use fingertipscharts data has to be in the right format.

Bar charts can be created using the compare\_areas function

smok\_rates |>  
 compare\_areas(area = region, value = value, lowerci = lowercl, upperci = uppercl, title = "Rate of attendance at smoking clinics 18-44: F")  
  
flu\_coverage |>  
 filter(gender == "female" & age == "0-18") |>  
 compare\_areas(area = region, value = value, lowerci = lowercl, upperci = uppercl, fill = gender, title = "Flu vaccination coverage 0-18; F")

|  |  |  |
| --- | --- | --- |
| |  | | --- | | (a) Smoking rates |  |  | | --- | | (b) Paediatric flu coverage |   Figure 10.1: Using `compare-areas |

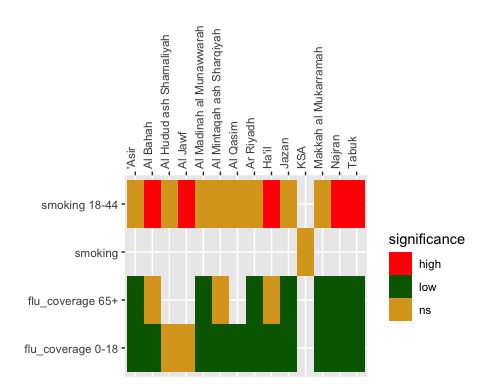
## 10.3 Data overview

PHOF uses a heatmap (known as a “tartan rug” to display area by indicator summaries. The term tartan rug is used to refer to the colour scheme (red-amber-green) to denote statistical significance. The

prof\_data <- bind\_rows(smok\_rates, flu\_coverage) |>  
 select(-c(11:13))   
  
## calculate ksa values  
ksa\_values <- prof\_data |>  
 group\_by(indicator) |>  
 reframe(ksa\_values = 100000 \* sum(n, na.rm = TRUE) / sum(pop))  
  
## join to data  
prof\_data <- prof\_data |>  
 left\_join(ksa\_values)  
  
## code for statistical significance  
  
prof\_data <- prof\_data |>  
 mutate(significance = case\_when(lowercl > ksa\_values ~"high" ,  
 uppercl < ksa\_values~ "low",  
 TRUE ~ "ns"))  
  
head(prof\_data) |>  
 gt::gt()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 10.1: Required data format example   | indicator | region | age | gender | n | pop | value | lowercl | uppercl | timeperiod | ksa\_values | significance | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | smoking 18-44 | 'Asir | 18-44 | female | NA | 366324 | NA | NA | NA | 2023 | 7.618658 | ns | | smoking 18-44 | Al Bahah | 18-44 | female | 28 | 61969 | 45.183882 | 30.017729 | 65.30573 | 2023 | 7.618658 | high | | smoking 18-44 | Al Hudud ash Shamaliyah | 18-44 | female | 4 | 70149 | 5.702148 | 1.553643 | 14.59976 | 2023 | 7.618658 | ns | | smoking 18-44 | Al Jawf | 18-44 | female | 23 | 103536 | 22.214495 | 14.077587 | 33.33412 | 2023 | 7.618658 | high | | smoking 18-44 | Al Madinah al Munawwarah | 18-44 | female | NA | 387944 | NA | NA | NA | 2023 | 7.618658 | ns | | smoking 18-44 | Al Mintaqah ash Sharqiyah | 18-44 | female | NA | 877403 | NA | NA | NA | 2023 | 7.618658 | ns | |

prof\_data |>  
 ggplot() +  
 geom\_tile(aes(region, indicator, fill = significance)) +  
 scale\_fill\_manual(values = c("red", "darkgreen", "goldenrod")) +  
 scale\_x\_discrete(position = "top") +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 0)) +  
 labs(y = "", x = "")



# References

Public Health Data Science. 2018. “PHDS Guidance - Confidence Intervals - Public Health Profiles.” Public Health England.

1. Note. The census data was downloaded from <https://tableau.saudicensus.sa/#/views/TA3-PopulationbydetailedAgebyRegionGovernorateNationalityandGenderAR_16850208449070/PopulationbydetailedAgebyRegionGovernorateNationalityandGenderARCSV.csv> and variable and region names translated to English using ChatGPT4o [↑](#footnote-ref-61)
2. This uses the spatial locations of smoking clinics which include directorate names to map to KSA regional boundaries [↑](#footnote-ref-63)
3. [https://www.moh.gov.sa/en/Ministry/Projects/TCP/Pages/default.aspx](https://www.moh.gov.sa/en/Ministry/Projects/TCP/Pages/default.aspx%22) [↑](#footnote-ref-66)