my\_report

Julian Flowers

21/09/2017

## My simple report

This is my report about diabetes audit data. Firstly I am going to use the fingertipsR package to download some data - I am interested in the recently published diabetes data and seeing if people with type 1 diabetes get the same level of care as people with type 2 diabetes.

I am only going to use 2 R packages for this analysis - tidyverse which includes dplyr containing tools for manipulating our data, and ggplot2 for charting.

To get data from Fingertips you need to know some ID numbers. The first step is to run the profiles() function which we’ll save to an R object called *prof*.

prof <- profiles()

We can look at the *prof* object with the View function (note capital V)

View(prof)

I can see that the ProfileID is 137 and the DomainID is 1938133136.

To download the data all we need to do is pass the DomainID to the fingertips\_data function. We’ll save this to an R object *diabetes\_data*. One important issue is that by default the function downloads data for Upper Tier Local Authorities (ID = 102). If there is no data for UTLAs for a given indicator in Fingertips, no data will be returned. If data is avaialble for other geographies these need to be specficied. In this case - looking at the profile in Fingertips - we can see data is available for CCGs. CCGs have AreaTypeIDs of 152 and 153 - we need to include this in the download function. It’ll take a few seconds to download.

## Download  
diabetes\_data <- fingertips\_data(DomainID = 1938133136, AreaTypeID = c(152, 153))

Now we can examine what we have.

View(diabetes\_data)

We may want to reduce the number of variables, and use only the latest data. The download contains the field TimeperiodSortable. This in included to identify the latest data available in Fingertips - the latest data will always be the maximum value of this field.

We will use the filter function to restrict the dataset to the latest data:

filtered\_data <- filter(diabetes\_data, TimeperiodSortable == max(TimeperiodSortable) )

we’ll recode the data to split out type 1 from type 2 diabetes;

filtered\_data <- filtered\_data %>%  
 mutate(dm\_type = ifelse(stringr::str\_detect(IndicatorName, "type 2"), "type\_2",   
 ifelse(stringr::str\_detect(IndicatorName, "type 1"), "type\_1",   
 "other")))

and we’ll restrict the number of fields using the select function - this allows to restrict our dataset to fewer columns

filtered\_data <- select(filtered\_data, dm\_type, IndicatorName, AreaName, Sex, Age, Value, LowerCI95.0limit, UpperCI95.0limit, Count, Denominator)  
  
  
## Show the top 6 rows  
head(filtered\_data)

## # A tibble: 6 x 10  
## dm\_type IndicatorName AreaName Sex Age Value LowerCI95.0limit  
## <chr> <chr> <chr> <chr> <chr> <dbl> <dbl>  
## 1 other Level of partici… England Pers… All … 95.0 94.5  
## 2 other Level of partici… London NHS… Pers… All … 96.4 NA   
## 3 other Level of partici… Wessex NHS… Pers… All … 90.9 NA   
## 4 other Level of partici… Cheshire a… Pers… All … 93.0 NA   
## 5 other Level of partici… Yorkshire … Pers… All … 94.7 NA   
## 6 other Level of partici… Central Mi… Pers… All … 93.8 NA   
## # ... with 3 more variables: UpperCI95.0limit <dbl>, Count <dbl>,  
## # Denominator <dbl>

## Exploratory analysis

options(digits = 3)  
  
filtered\_data %>%  
 group\_by(AreaName, IndicatorName) %>%  
 count() %>%  
 filter(n >1)

## # A tibble: 4,311 x 3  
## # Groups: AreaName, IndicatorName [4,311]  
## AreaName IndicatorName n  
## <chr> <chr> <int>  
## 1 Central Midlands NHS region Level of participation in the Nation… 4  
## 2 Central Midlands NHS region People with type 1 diabetes who have… 4  
## 3 Central Midlands NHS region People with type 1 diabetes who have… 4  
## 4 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 5 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 6 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 7 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 8 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 9 Central Midlands NHS region People with type 1 diabetes who rece… 4  
## 10 Central Midlands NHS region People with type 1 diabetes whose sm… 4  
## # ... with 4,301 more rows

summary(filtered\_data)

## dm\_type IndicatorName AreaName   
## Length:16940 Length:16940 Length:16940   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Sex Age Value LowerCI95.0limit  
## Length:16940 Length:16940 Min. : 12.8 Min. :10   
## Class :character Class :character 1st Qu.: 70.9 1st Qu.:68   
## Mode :character Mode :character Median : 83.3 Median :81   
## Mean : 78.2 Mean :76   
## 3rd Qu.: 92.6 3rd Qu.:90   
## Max. :100.0 Max. :99   
## NA's :1316   
## UpperCI95.0limit Count Denominator   
## Min. : 15 Min. : 5 Min. : 5   
## 1st Qu.: 73 1st Qu.: 570 1st Qu.: 820   
## Median : 85 Median : 2760 Median : 3870   
## Mean : 80 Mean : 15086 Mean : 18483   
## 3rd Qu.: 94 3rd Qu.: 9740 3rd Qu.: 12145   
## Max. :100 Max. :2593585 Max. :2721580   
## NA's :1316

summary\_stats <- filtered\_data %>%  
 group\_by(dm\_type, IndicatorName) %>%  
 distinct() %>%  
 summarise(n = n(),  
 mean = mean(Value, na.rm = TRUE),   
 median = median(Value, na.rm = TRUE),   
 min = min(Value, na.rm = TRUE),   
 max = max(Value, na.rm = TRUE),   
 q25 = quantile(Value, probs = 0.25, na.rm = TRUE),   
 q75 = quantile(Value, probs = 0.75, na.rm = TRUE)  
 )  
   
 summary\_stats %>%  
 pander::pander()

Table continues below

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| dm\_type | IndicatorName | n | mean | median | min | max |
| other | Level of participation in the National Diabetes Audit | 225 | 95.03 | 98.59 | 33.33 | 100 |
| type\_1 | People with type 1 diabetes who have had their BMI recorded | 229 | 75.54 | 76 | 51.11 | 92.66 |
| type\_1 | People with type 1 diabetes who have received an annual foot check | 229 | 70.51 | 71.2 | 53.46 | 82.98 |
| type\_1 | People with type 1 diabetes who received a blood pressure check | 229 | 90.74 | 91.09 | 81.29 | 97.52 |
| type\_1 | People with type 1 diabetes who received a blood test | 229 | 84.83 | 85.12 | 71.34 | 94.02 |
| type\_1 | People with type 1 diabetes who received a cholesterol check | 229 | 80.66 | 81.34 | 65.62 | 88.93 |
| type\_1 | People with type 1 diabetes who received a serum creatinine test | 229 | 82.86 | 83.49 | 64.29 | 92.01 |
| type\_1 | People with type 1 diabetes who received all 8 care processes | 229 | 34.43 | 33.91 | 12.79 | 59.78 |
| type\_1 | People with type 1 diabetes who received urinary albumin test | 229 | 50.59 | 49.36 | 22.44 | 77.27 |
| type\_1 | People with type 1 diabetes whose smoking status is recorded | 229 | 79.93 | 80.17 | 65.05 | 95.05 |
| type\_2 | People with type 2 diabetes who have had their BMI recorded | 229 | 83.28 | 83.96 | 63.01 | 97.15 |
| type\_2 | People with type 2 diabetes who have received an annual foot check | 229 | 79.8 | 83.25 | 56.24 | 95.15 |
| type\_2 | People with type 2 diabetes who received a blood pressure check | 229 | 96.42 | 96.44 | 91.57 | 99.14 |
| type\_2 | People with type 2 diabetes who received a blood test | 229 | 95.32 | 95.49 | 88.5 | 97.99 |
| type\_2 | People with type 2 diabetes who received a cholesterol check | 229 | 93.19 | 93.33 | 85.04 | 97.15 |
| type\_2 | People with type 2 diabetes who received a serum creatinine test | 229 | 95.11 | 95.42 | 88.87 | 97.67 |
| type\_2 | People with type 2 diabetes who received all 8 care processes | 229 | 48.13 | 48.45 | 18.04 | 88.78 |
| type\_2 | People with type 2 diabetes who received urinary albumin test | 229 | 65.78 | 67.69 | 34.16 | 93.51 |
| type\_2 | People with type 2 diabetes whose smoking status is recorded | 229 | 86.03 | 86.94 | 66.04 | 97.48 |

|  |  |
| --- | --- |
| q25 | q75 |
| 92.99 | 100 |
| 70.94 | 81.25 |
| 65.74 | 74.78 |
| 89.2 | 92.45 |
| 82.35 | 87.9 |
| 78.4 | 83.08 |
| 80.62 | 85.71 |
| 28.57 | 39.44 |
| 42.86 | 57.14 |
| 76.76 | 83.2 |
| 80 | 87.35 |
| 72.65 | 87.24 |
| 95.96 | 97.07 |
| 94.55 | 96.47 |
| 92.35 | 94.19 |
| 94.33 | 96.14 |
| 41.65 | 54.42 |
| 60.01 | 72.33 |
| 83.21 | 89.98 |

### Comparing mean levels

filtered\_data %>%   
 filter(IndicatorName != "Level of participation in the National Diabetes Audit") %>%  
 select(IndicatorName, Count, Denominator, AreaName, Value)%>%  
 distinct() %>%   
 spread(IndicatorName, Count)

## # A tibble: 4,035 x 21  
## Denominator AreaName Value `People with type 1 … `People with type 1…  
## <dbl> <chr> <dbl> <dbl> <dbl>  
## 1 210 NHS Warwi… 28.6 NA NA  
## 2 210 NHS Warwi… 59.5 125 NA  
## 3 210 NHS Warwi… 64.3 NA 135  
## 4 210 NHS Warwi… 66.7 NA NA  
## 5 210 NHS Warwi… 71.4 NA NA  
## 6 210 NHS Warwi… 88.1 NA NA  
## 7 215 NHS Warwi… 23.3 NA NA  
## 8 215 NHS Warwi… 72.1 NA NA  
## 9 260 NHS Bradf… 65.4 NA 170  
## 10 260 NHS Bradf… 67.3 NA NA  
## # ... with 4,025 more rows, and 16 more variables: `People with type 1  
## # diabetes who received a blood pressure check` <dbl>, `People with type  
## # 1 diabetes who received a blood test` <dbl>, `People with type 1  
## # diabetes who received a cholesterol check` <dbl>, `People with type 1  
## # diabetes who received a serum creatinine test` <dbl>, `People with  
## # type 1 diabetes who received all 8 care processes` <dbl>, `People with  
## # type 1 diabetes who received urinary albumin test` <dbl>, `People with  
## # type 1 diabetes whose smoking status is recorded` <dbl>, `People with  
## # type 2 diabetes who have had their BMI recorded` <dbl>, `People with  
## # type 2 diabetes who have received an annual foot check` <dbl>, `People  
## # with type 2 diabetes who received a blood pressure check` <dbl>,  
## # `People with type 2 diabetes who received a blood test` <dbl>, `People  
## # with type 2 diabetes who received a cholesterol check` <dbl>, `People  
## # with type 2 diabetes who received a serum creatinine test` <dbl>,  
## # `People with type 2 diabetes who received all 8 care processes` <dbl>,  
## # `People with type 2 diabetes who received urinary albumin test` <dbl>,  
## # `People with type 2 diabetes whose smoking status is recorded` <dbl>