Report_Melanie Rees-Roberts

Melanie Rees-Roberts

2024-11-14

Title: Understanding health and wealth in Scottish Council areas

Background

Over the last 100 years, life expectancy in the UK has increased significantly with advances in living standards and advancing healthcare. However, over the last 10 years this has stalled and even declined in some instances alongside significant inequalities arising across the country. In Scotland, disparities in life expectancy are evident where the most deprived exhibit approximately 10 years difference in life expectancy compared to their more affluent neighbourhoods. In recent years, the cost of living crisis has also impacted on peoples health and wellbeing but the extent to which this has impacts on different income levels is unknown and thereafter how this links to health. The extent to which people feel able to manage their own money may act as a useful proxy measure for financial stress within a cost of living crisis regardless of income levels and be a useful indicator for health status.

Research questions

1. What is the variation in health and mental health across different Scottish council areas?

2. How is population income associated with the perception of being able to manage money well?

3. How does the ability to manage money well impact on health and mental health in Scotland?

Packages and libraries

For this work, the following packages and libraries were used:

#install.packages("tidyverse")

```r
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.3.3
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
## Warning: package 'tibble' was built under R version 4.3.3

## Warning: package 'tidyr' was built under R version 4.3.3

## Warning: package 'readr' was built under R version 4.3.3

## Warning: package 'purrr' was built under R version 4.3.3

## Warning: package 'dplyr' was built under R version 4.3.3

## Warning: package 'stringr' was built under R version 4.3.3

## Warning: package 'forcats' was built under R version 4.3.3

## Warning: package 'lubridate' was built under R version 4.3.3

## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- 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(ggplot2)
```

###Data acquisition # In this work, we used the datasets in Table 1 below. These data were downloaded from https://scotland.shinyapps.io/ScotPHO_profiles_tool/ in a single file and then tidy datasets created for analysis.

#Import the data and create tidy datasets as each variable is listed in a single column with the name of the indicator in the column 'indicator'

```r
Data <- read_csv("ScotPHO_data.csv")
```

```
## Rows: 224 Columns: 11
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (6): 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.
```

```
glimpse(Data)
```

```
## Rows: 224
## Columns: 11
## $ area_code                <chr> "S12000005", "S12000006", "S12000008", "S120~
## $ area_type                <chr> "Council area", "Council area", "Council are~
## $ area_name                <chr> "Clackmannanshire", "Dumfries & Galloway", "~
## $ year                     <dbl> 2020, 2020, 2020, 2020, 2020, 2020, 2020, 20~
## $ period                   <chr> "2019-2021 (3 year aggregate)", "2019-2021 (~
## $ type_definition          <chr> "Years", "Years", "Years", "Years", "Years",~
## $ indicator                <chr> "Healthy life expectancy, females", "Healthy~
## $ numerator                <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ measure                  <dbl> 59.6, 59.2, 59.9, 68.4, 67.2, 64.0, 59.2, 65~
## $ upper_confidence_interval <dbl> 63.3, 61.9, 62.4, 71.0, 69.9, 68.2, 62.1, 69~
## $ lower_confidence_interval <dbl> 55.9, 56.6, 57.4, 65.8, 64.4, 59.8, 56.2, 62~
```

Data cleaning and preparation

##Dataset 1 - Healthy life expectancy. # Create tidy data with helathy life expectancy for males and females

```
hle <- Data %>%
  filter(grepl('Healthy', indicator)) %>%
  select(area_code,area_name,indicator,measure,upper_confidence_interval,lower_confidence_interval) %>%
  rename(sex = indicator,
         hle = measure) %>%
  mutate(sex = gsub('Healthy life expectancy, females','0', sex,)) %>%
  mutate(sex = gsub('Healthy life expectancy, males','1', sex,))

head(hle)
```

```
## # A tibble: 6 x 6
##   area_code area_name  sex     hle upper_confidence_int~1 lower_confidence_int~2
##   <chr>     <chr>      <chr> <dbl>                  <dbl>                  <dbl>
## 1 S12000005 Clackmann~ 0      59.6                   63.3                   55.9
## 2 S12000006 Dumfries ~ 0      59.2                   61.9                   56.6
## 3 S12000008 East Ayrs~ 0      59.9                   62.4                   57.4
## 4 S12000010 East Loth~ 0      68.4                   71                     65.8
## 5 S12000011 East Renf~ 0      67.2                   69.9                   64.4
## 6 S12000013 Na h-Eile~ 0      64                     68.2                   59.8
## # i abbreviated names: 1: upper_confidence_interval,
## #   2: lower_confidence_interval
```

```
glimpse(hle)
```

```
## Rows: 64
## Columns: 6
## $ area_code                <chr> "S12000005", "S12000006", "S12000008", "S120~
## $ area_name                <chr> "Clackmannanshire", "Dumfries & Galloway", "~
## $ sex                      <chr> "0", "0", "0", "0", "0", "0", "0", "0", "0",~
## $ hle                      <dbl> 59.6, 59.2, 59.9, 68.4, 67.2, 64.0, 59.2, 65~
## $ upper_confidence_interval <dbl> 63.3, 61.9, 62.4, 71.0, 69.9, 68.2, 62.1, 69~
## $ lower_confidence_interval <dbl> 55.9, 56.6, 57.4, 65.8, 64.4, 59.8, 56.2, 62~
```

We now need to combine males and females to create a single mean healthy life expectancy value for each Council area

```
mean_hle <- hle %>%
  group_by(area_code) %>%
  mutate(mean_hle = mean(hle)) %>%
  select
```

```
## function (.data, ...)
## {
##     UseMethod("select")
## }
## <bytecode: 0x00000163f47b2d30>
## <environment: namespace:dplyr>
```

We have duplicate mean healthy life expectancy values, so now we need to slice the rows containing sex = 0 and remove the column sex

```
mean_hle1 <- mean_hle %>%
  filter(sex != 1) %>%
  select(area_code, area_name, mean_hle)

head(mean_hle1)
```

```
## # A tibble: 6 x 3
## # Groups:   area_code [6]
##   area_code area_name            mean_hle
##   <chr>     <chr>                   <dbl>
## 1 S12000005 Clackmannanshire         60.4
## 2 S12000006 Dumfries & Galloway      60.8
## 3 S12000008 East Ayrshire            58.4
## 4 S12000010 East Lothian             64.6
## 5 S12000011 East Renfrewshire        67.0
## 6 S12000013 Na h-Eileanan Siar       65.7
```

##Dataset 2 - Population income deprived

```
pid <- Data  %>%
  drop_na()

pid <- pid %>%
  select(area_code,numerator,measure) %>%
  rename(population = numerator,
         pid = measure)
```

##Dataset 3 - Mental health score deprived

```
mental <- Data %>%
  filter(grepl('Mental', indicator)) %>%
  select(area_code,measure) %>%
  rename(mental_score = measure)
```

##Dataset 4 - Households managing financially well

```
finance <- Data %>%
  filter(grepl('Households', indicator)) %>%
  select(area_code,measure) %>%
  rename(finance = measure)
```

##Combine data into a single table named data_final.

```
data_final <- inner_join(mean_hle1, pid, by = 'area_code')

data_final <- inner_join(data_final, mental, by = 'area_code')

data_final <- inner_join(data_final, finance, by = 'area_code')
```

Data analysis and visualisation

1. What is the variation in health and mental health across different Scottish council areas?

```
data_final %>%
  ggplot() +
  geom_col(aes(x = reorder(area_name, pid), y = mean_hle)) +
  geom_point(aes(x=area_name, y = mental_score)) +
  coord_cartesian(ylim = c(40, NA)) +
  ggtitle("        Health and mental health across Scottish Councils") +
  labs(x="Area name by increasing % of population income deprived",
       y = "Mean Healthy life expectancy in years (bars)
            Mental Health Score   (dots)") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```
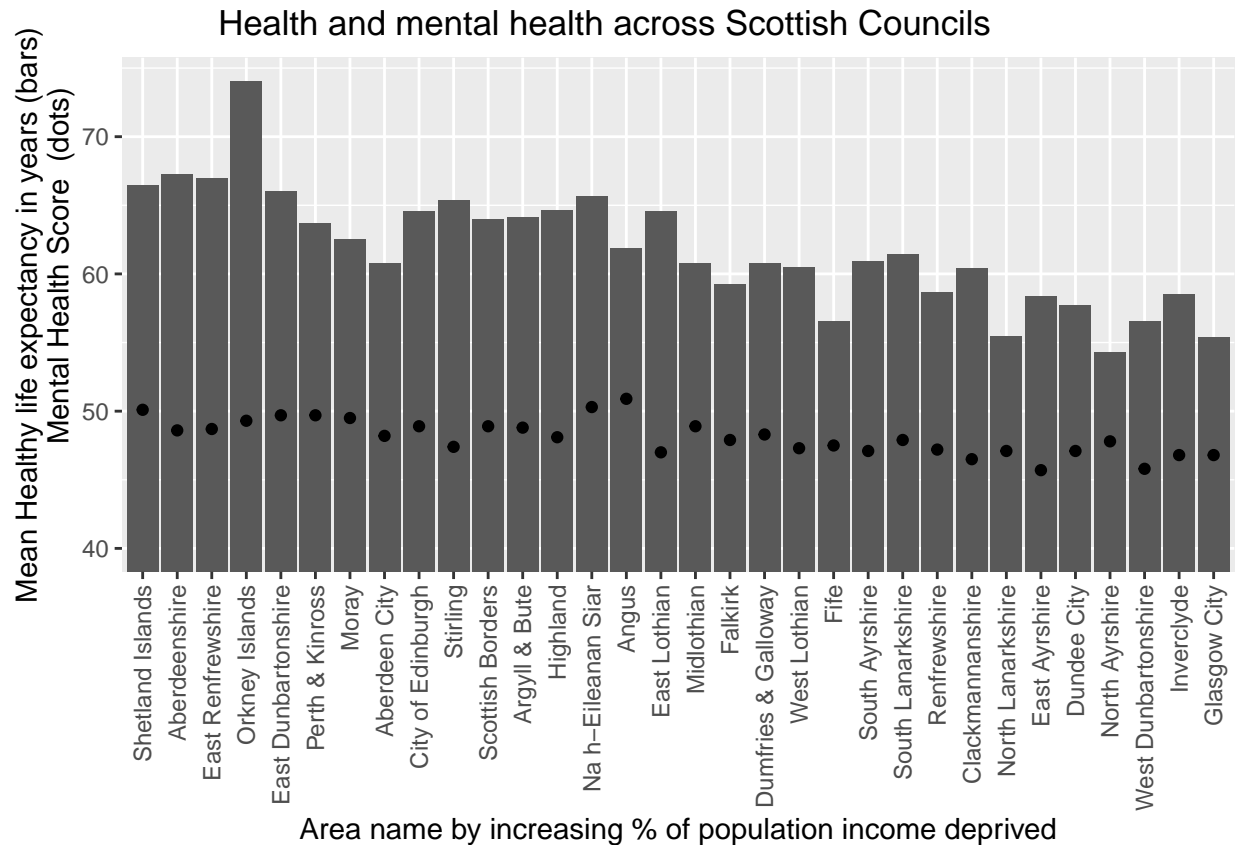
Health and mental health across Scottish Councils

2. **How is population income associated with the perception of being able to manage money well?**

```r
data_final %>%
ggplot(aes(x = pid, y = finance)) +
geom_point() +
geom_smooth(method = lm) +
ggtitle("Correlation of percentage population income deprived with ability
        of population to manage their own money for Scottish Council areas") +
labs(x="% of population that are income deprived",
     y = "Ability to manage their own money effectively (score)") +
coord_cartesian(ylim = c(0, NA))
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

Correlation of percentage population income deprived with ability of population to manage their own money for Scottish Council areas
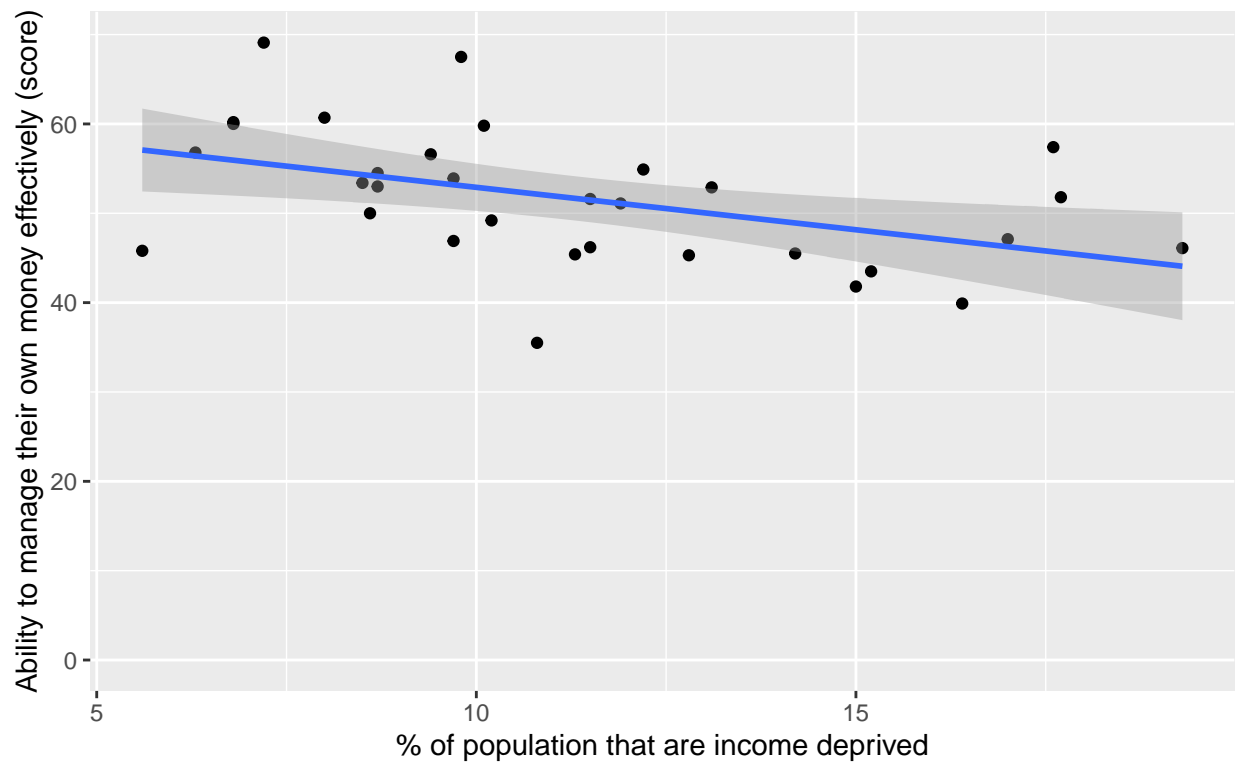
3. How does the ability to manage money well impact on health and mental health in Scotland?

Discussion

##Limitations #different years of data