220_Reproducibility_Part_2

Robert Atkinson

04/11/2020

In this write-up I will be working with the 2011 OAC and the 2015 IMD data.

Loading Packages

```
library(tidyverse)
```

```
## -- Attaching packages ----------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.1     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts -------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Add Data

Introducing the data that we will use for the analysis.

2011 OAC

The **2011 OAC** data is the Geodemographics Classification derived from the UK's 2011 Census. This data is grouped into Ouptut Area levels which are seperated into smaller groups:

- supergroups
- groups
- subgroups

This informaton can be found at the *Datashine Website*

```
leicester_2011OAC <-
  readr::read_csv("2011_OAC_Raw_uVariables_Leicester.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   OA11CD = col_character(),
##   LSOA11CD = col_character(),
##   LSO11ANM = col_character(),
##   MSOA11CD = col_character(),
```

```
##   MSOA11NM = col_character(),
##   LAD11CD = col_character(),
##   LAD11NM = col_character(),
##   supgrpname = col_character(),
##   grpcode = col_character(),
##   grpname = col_character(),
##   subgrpcode = col_character(),
##   subgrpname = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

2015 IMD

The **2015 IMD** data is a series of indexes, representing the reletive deprevation of small areas in England.

This information can be found on the *Government Website*

Load the 2015 IMD data:

```
leicester_IMD2015 <-
  readr::read_csv("IndexesMultipleDeprivation2015_Leicester.csv")
```

```
## Parsed with column specification:
## cols(
##   FeatureCode = col_character(),
##   DateCode = col_double(),
##   Measurement = col_character(),
##   Units = col_logical(),
##   Value = col_double(),
##   IndicesOfDeprivation = col_character()
## )
```

Analysis

First, the IMD data is tidied and converted into a wide format.

- Use filter to retrieve just the Decile data.
- Tidy data by removing brackets and punctuation.
- Convert the table into a wide format.
- Remove columns that are no longer needed.

```
#wider data
leicester_IMD2015_decile_wide <- leicester_IMD2015 %>%
  # Select only Scores
  dplyr::filter(Measurement == "Decile") %>%

  # Trim names of IndicesOfDeprivation
  dplyr::mutate(IndicesOfDeprivation =
                  str_replace_all(IndicesOfDeprivation, "\\s", "")) %>%
  dplyr::mutate(IndicesOfDeprivation =
                  str_replace_all(IndicesOfDeprivation, "[:punct:]", "")) %>%
  dplyr::mutate(IndicesOfDeprivation =
                  str_replace_all(IndicesOfDeprivation, "\\(", "")) %>%
  dplyr::mutate(IndicesOfDeprivation =
                  str_replace_all(IndicesOfDeprivation, "\\)", "")) %>%
```

```
  # Convert data into wide format
  pivot_wider(
    names_from = IndicesOfDeprivation,
    values_from = Value
  ) %>%

  # Drop columns
  dplyr::select(-DateCode, -Measurement, -Units)
```

Join both dataframes

- use an inner join to combine the two dataframes
- use Lower Layer Super Output Area as 2nd colum in new table

```
leicester_2011OAC_IMD2015 <-
  leicester_2011OAC %>%
  inner_join(
    leicester_IMD2015_decile_wide,
    by = c("LSOA11CD" = "FeatureCode")
  )
```

Create table

- Filter data so to retrieve only the Lower Level Super Output Area that contains the University of Leicester (E01013649).
- Select Output Area, Lower Level Output Area, Supergroup name, IMD, and total population.
- Creat a well-formated tavle

```
leicester_2011OAC_IMD2015 %>%
  # Retrieve Lower Level Output Area for area around University of Leicester
  dplyr::filter(LSOA11CD == "E01013649") %>%

  # Select Output Area, Lower Level Output Area, Supergroup name, IMD, and total population.
  dplyr::select(OA11CD, LSOA11CD, supgrpname,
                IndexofMultipleDeprivationIMD, Total_Population) %>%

  # Create table
  knitr::kable()
```

OA11CD	LSOA11CD	supgrpname	IndexofMultipleDeprivationIMD	Total_Population
E00169447	E01013649	Cosmopolitans	5	235
E00168083	E01013649	Cosmopolitans	5	230
E00068893	E01013649	Cosmopolitans	5	289
E00068892	E01013649	Cosmopolitans	5	297
E00068890	E01013649	Cosmopolitans	5	490

Conclusion

To conclude, the 5 Output Areas that make up the Lower Level Super Output Area where the Unversity of Leicester resides are all classified as Cosmopolitan. With regards to thee Index of Multiple Deprivation, all

areas have alocated a score of 5. This score indicates that all these neighbourhoods have a relative deprivation of between 40% and 50% according to the Ministry of Housing, Communities & Local Government.