Downloading the Data

UKDS, Nadia Kennar

28/10/2021

This document provides some extra information on how to download the datasets required for the project. These include, the crime dataset, the shapefile and population statistics

All data can be found via git *link*

### Crime Data

Follow the link <https://data.police.uk/>

* Select ‘Downloads’
* In date range select ‘January 2019 to December 2019’
* Select ‘Surrey Police’ and
* select ‘Include Crime Data’.
* Download and unzip the data into your working directory

We will be using just the 08/2020 reported crimes rep.

crime01\_19 <- read\_csv("Data/2020-08/2020-08-surrey-street.csv") %>%   
 janitor::clean\_names()

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## `Crime ID` = col\_character(),  
## Month = col\_character(),  
## `Reported by` = col\_character(),  
## `Falls within` = col\_character(),  
## Longitude = col\_double(),  
## Latitude = col\_double(),  
## Location = col\_character(),  
## `LSOA code` = col\_character(),  
## `LSOA name` = col\_character(),  
## `Crime type` = col\_character(),  
## `Last outcome category` = col\_character(),  
## Context = col\_logical()  
## )

## Shapefile

To obtain boundary data we will use the <https://borders.ukdataservice.ac.uk/bds.html>

Steps to download:

* Select; England, Statistical Building Block, 2011 and later
* Click ‘Find’
* Select ‘English Lower Layer Super Output Areas’
* Click ‘List Areas’
* Select ‘Surrey Health’
* Click ‘Extract Boundary Data’

Read in the Shapefile for ‘Surrey Heath’

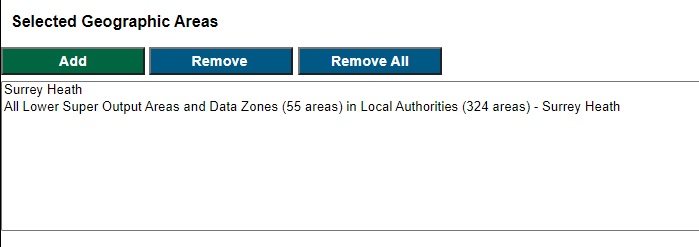
shp\_file <- st\_read("Data/Shapefile/england\_lsoa\_2011.shp")

## Reading layer `england\_lsoa\_2011' from data source `C:\Users\naken\OneDrive\Documents\Research Associate\Work\Crime Mapping\CrimeMapping\Data\Shapefile\england\_lsoa\_2011.shp' using driver `ESRI Shapefile'  
## Simple feature collection with 55 features and 3 fields  
## Geometry type: MULTIPOLYGON  
## Dimension: XY  
## Bounding box: xmin: 485406.9 ymin: 154122.5 xmax: 501181.2 ymax: 166842.9  
## Projected CRS: OSGB 1936 / British National Grid

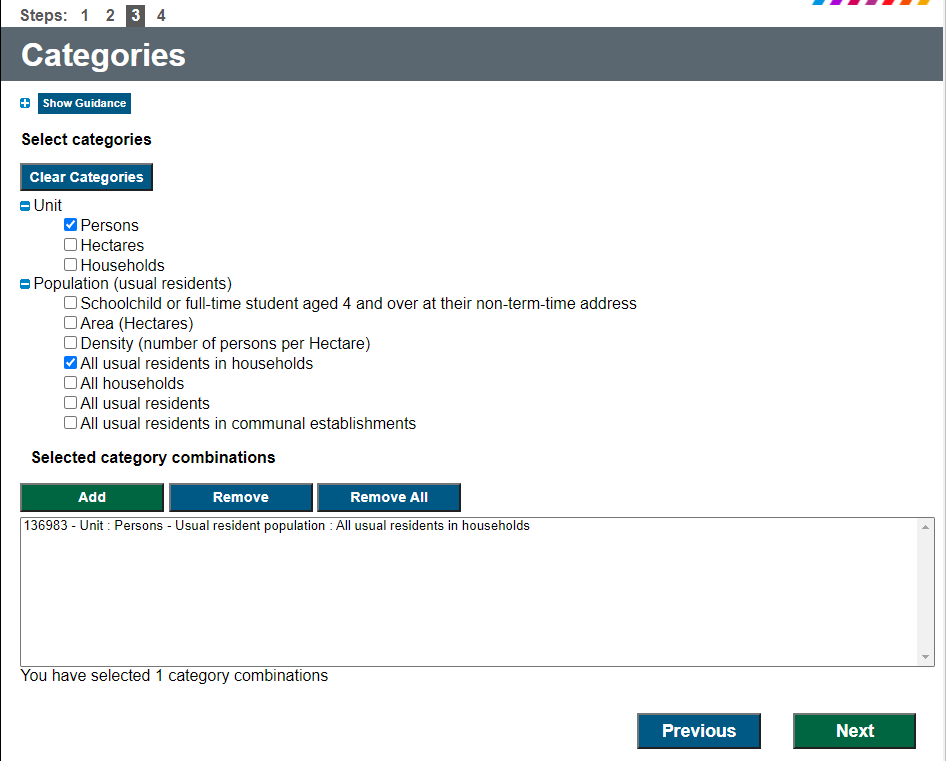
## Census Data

To obtain population statistics use <http://infuse2011gf.ukdataservice.ac.uk/>

* Select ‘Geography’
* Expand the ‘Local Authorities’ tab in England
* Tick and expand ‘Surrey Heath’
* Select ‘Lower Super Output area and Data Zones (55 areas)’
* Select ‘Add’ and ‘Next’



* Click ‘Population (usual residence)’
* Click Next
* Under unit select ‘Persons’ and under ‘Population (usual residents)’ select ‘All usual residets in households’
* Click ‘Add’



* Click ’Next and download the data

To read in the data

pop <- read\_csv("Data/Census Population/Data\_UNIT\_URESPOP.csv") %>%   
 slice(3:47) %>% #removes first and second row  
 select(2,3,6,7) %>% #selects the columns of interest  
 janitor::clean\_names() %>% #cleans the names   
 rename(pop\_count = f136983) %>% #rename the variable  
 mutate(pop\_count = as.numeric(pop\_count)) #turns variable to numeric

## Warning: Missing column names filled in: 'X8' [8]

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## CDU\_ID = col\_double(),  
## GEO\_CODE = col\_character(),  
## GEO\_LABEL = col\_character(),  
## GEO\_TYPE = col\_character(),  
## GEO\_TYP2 = col\_character(),  
## F2383 = col\_character(),  
## F136983 = col\_character(),  
## X8 = col\_logical()  
## )

\*Important Note: the above instructions for the census data only obtain the residential population, the data in the workshop also includes the workday population\*