Downloading the Data

UKDS, Nadia Kennar

March 2025 Workshop

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 the git *link*

### Crime Data

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

* Select ‘Downloads’
* In date range select ‘December 2023 – December 2024’
* Select ‘Surrey Police’ and
* select ‘Include Crime Data’.
* Download and unzip the data into your working directory

We will be using just the 02/2024 reported crime.

crime01\_19 <- read\_csv("Data/2024-02/2024-02-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, 2021 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’ (this will be saved as ‘England’ but pay attention to the file name when downloading to ensure you are only using the Surrey Heath files.

shp\_file <- st\_read("Data/Shapefile/england\_lsoa\_2021.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

The CKAN link to the UKDS page. This is the updated InFuse website that allows you to choice specific tables and boundary data.

* <https://statistics.ukdataservice.ac.uk/dataset/?vocab_Area_type=Lower%20Super%20Output%20Areas>

We downloaded our data from the nomisweb from the office for national statistics. (<https://www.nomisweb.co.uk/sources/census_2021>)

* TS001 - Number of usual residents in households and communal establishments
* (<https://www.nomisweb.co.uk/datasets/c2021ts001>)
* WD102EW - Population density (<https://www.nomisweb.co.uk/query/construct/submit.asp?menuopt=201&subcomp=> )
* Click on "Query data" to customize your data selection. In the "Geography" section, choose "Lower Layer Super Output Areas (2011)". Use the search function to find and select all LSOAs within "Surrey Heath".
* Ensure that the variables related to workday population and area measurements are selected.