Crime and Punishment

This activity explores how census data can be used to predict where crime will occur while examining how real crime data helps to inform police strategies and operations

🕐 45 – 55 mins

Newcastle University

Introduction

GIS is increasingly becoming a useful tool for law enforcement across the world. Many of the decisions that authorities make are influenced, in some capacity, by location. Depending on the location of a reported incident, nearby roads may need to be closed, people may need to be evacuated or patrols may need to be diverted along specific routes. Without the ability to determine the location of incidents, many questions will be left unanswered resulting in ill-informed decisions that lack credibility.

Consider the police information recording model (shown below) in terms of the board game Cluedo. Without knowing the room in which the event occurred, there is a degree of risk and guesswork in your accusation. In this practical we will explore the relationship between 'location' and 'event' to see if events can be predicted using location. More information about how GIS is being used to inform police strategies can be found here: https://www.geospatialuk.org/post/data-in-the-line-of-duty-psga-data-keeping-us-safe.
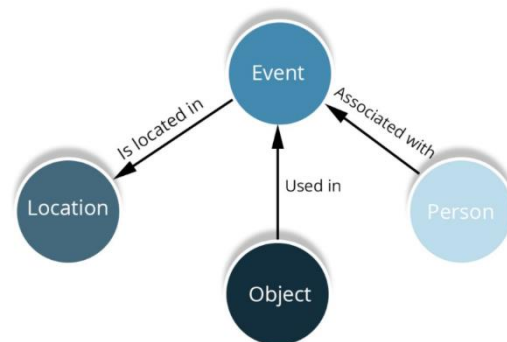
Police Information Recording Model (source: BGS)

The Census consists of a lot of different types of data that provide the government and police forces with information about various socio-economic variables, many of which can be used to inform how certain services are provided (https://census.gov.uk/census-stories). Example variables include employment status, education level, population and housing type. Although the census does not contain information that directly relates to crime, many of the elements within the database can be used as indicators of crime. This is an increasingly common method being used across the world to help authorities predict the general areas where crimes are more likely to occur, based on a variety of contributing factors.

What types of information may you expect to be useful in predicting the location of crime hotspots? Is crime more likely to occur in denser, urban populations or sparsely populated rural areas? Does somebody's education or level of qualifications affect the probability of them living in neighbourhoods with high crime rates? During this exercise, we will use GIS to begin to find answers to these questions and support our answers with data.

Sunderland is a city in the North East of England, with a population of around 340,00. The city is ranked within the top 20 for highest crime rates out of all UK major towns and cities (CrimeRate). Therefore, it will be used as an example throughout this practical as the corresponding police data will allow for a variety of analysis and mapping to be undertaken.

Learning Outcomes

In this practical, students will learn how to:

- Use free online maps
- Access real life online databases
- Manage CSV files
- Import CSV files into ArcGIS Online
- Create a heat map

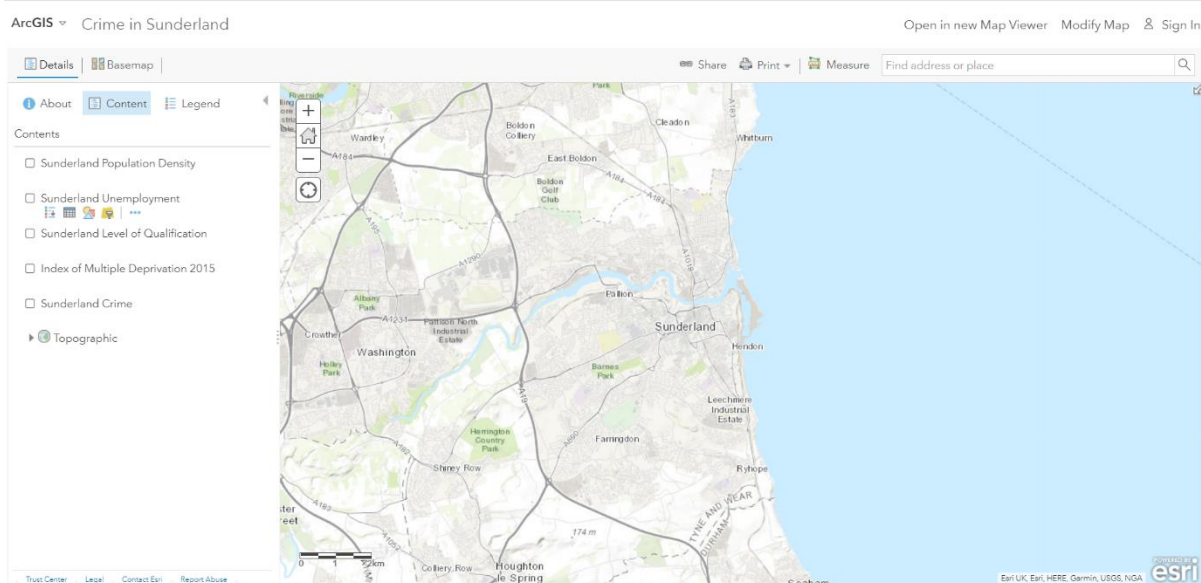This activity is aimed at KS4 – KS5 students, however students of all ages are welcome to complete the activity.

Contents

Section 1: Predicting Crime

This section will introduce using density mapping to identify spatial patterns and predict outcomes.

1.1. Changing the Basemap

a. Go to: https://arcg.is/05KXKv0.

    This will bring up a map which looks like this:

b. Before progressing, it is useful familiarise yourself with the Sunderland City region. We can use satellite imagery to help you understand what the data is showing. To do this, select the 'Imagery' option from the 'Basemap' dropdown menu.
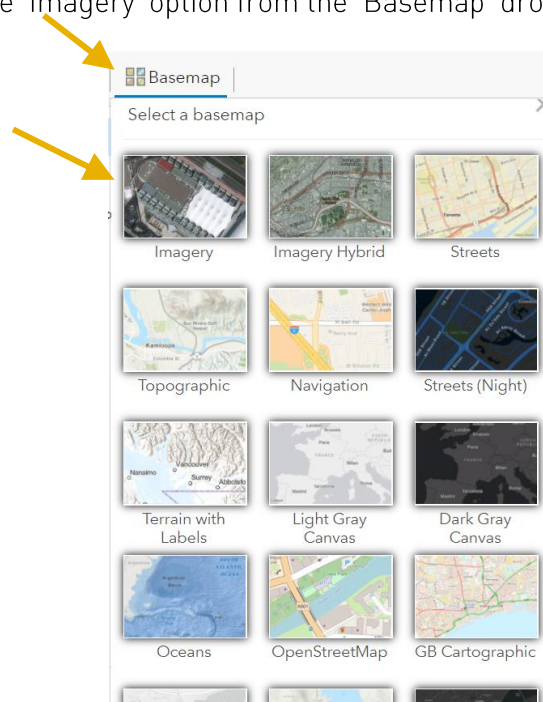
c.   Your map should now look something like this:
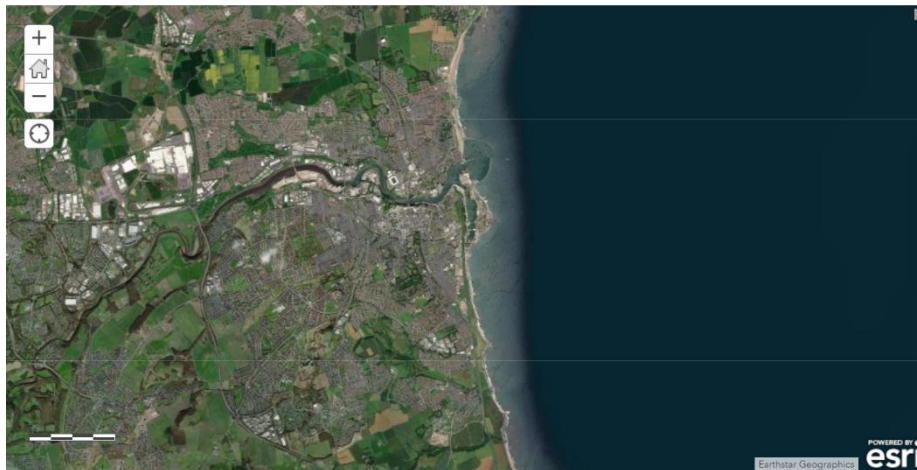
d.   With the imagery basemap loaded, explore the map to become familiar with the area by zooming in and out of Sunderland and the surrounding areas. Try and find Sunderland city centre.

1.2. Crime and Population Density

a.   Census data does not contain information that directly relates to crime however many elements of the census data can be used as indicators of crime that are being increasingly used to better predict where crimes are more likely to occur.

What types of information may you expect to be useful in predicting the location of crime hotspots?

b.   Click on the 'Content' tab in the top left corner. This will show you what 'layers' or datasets are included on the map.
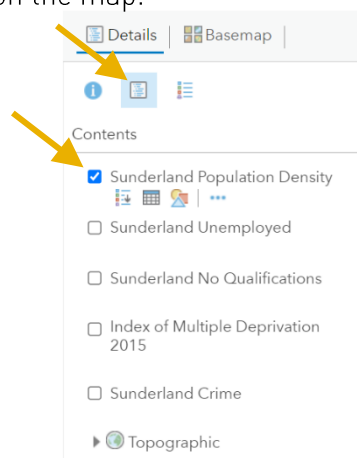
c.   Ensure that the 'Sunderland Population Density' layer is visible by clicking the box to the left of the layer name so it turns 'on'.

d. The map shows the population densities of Sunderland on a blue to beige colour ramp, with areas of navy and lighter areas of beige, but how can we find out what the colours mean?
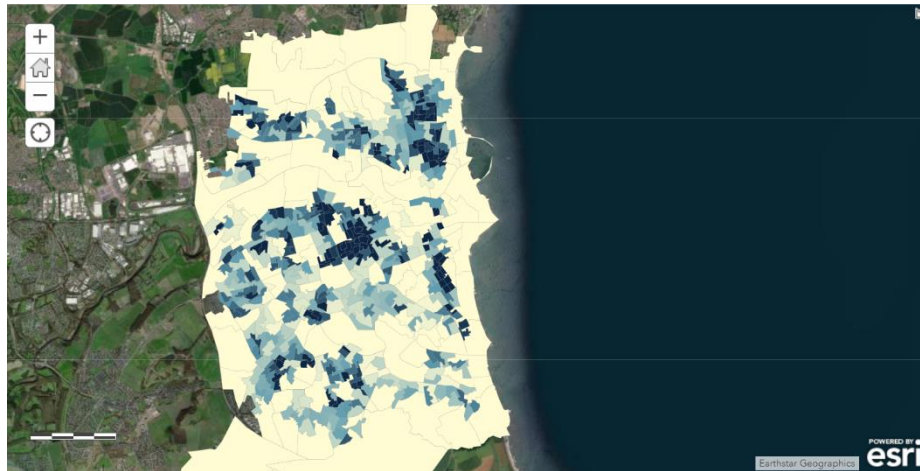

e. Click on the 'Show Map Legend' tool to find out what the colours represent.
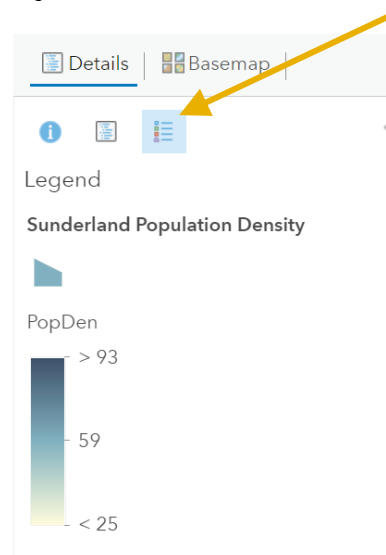
f. Now we can see that the dark blue areas show that there are more than 93 people per hectare living there, while areas of beige indicate a density of less than 25 people per hectare.

g. By turning the 'Sunderland Population Density' layer 'on' and 'off', it is apparent that the areas of highest population density are located in urban areas. Do you expect the most crime to occur in the most or least populated areas?

h. Do you think population density will be a good indicator of crime? Why or why not?

Newcastle University

1.3. Crime and Unemployment

a. As well as population data, the census provides information about the number of unemployed people in an area, a potentially useful indicator of crime.

b. Turn all the layers off. Then turn on 'Sunderland Unemployment' to view this dataset.
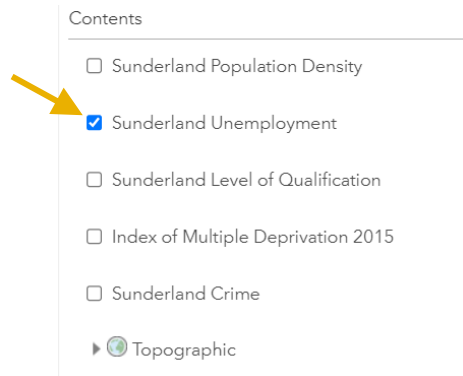
c. The dataset will look something like this:
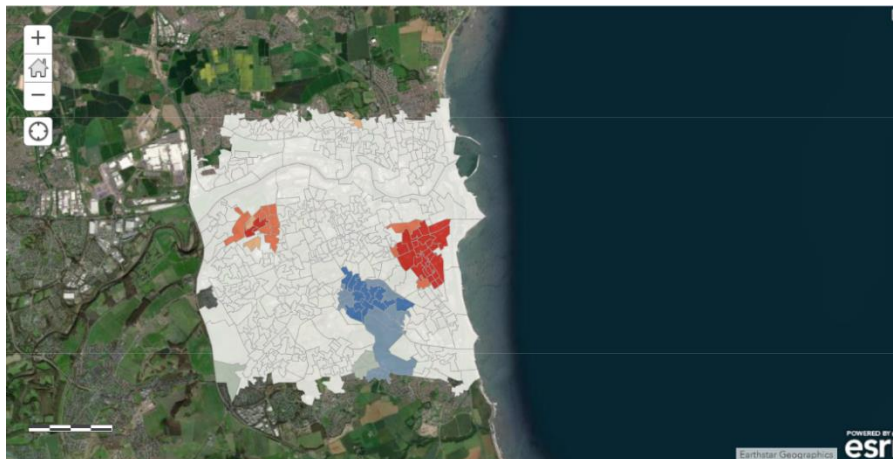
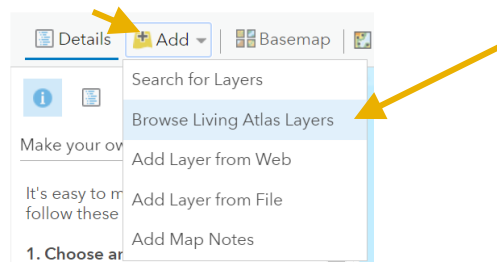d. Repeat the steps in Section 1.2 d – e to view the legend to see what the red and blue areas show.

e. Hotspot analysis is an extremely effective GIS technique. Confusingly, when we refer to hotspots we are not only referring to clusters of high (hotspot) values but also to areas of low (coldspot) values. The clusters of either high or low values are calculated using a set of robust statistical probability calculations. Fortunately, the computer can compute these statistics for us which allows us to focus on analysing the results.

Calculating the statistically significant high and low values is useful for identifying fundamental spatial patterns within the dataset.

f. We can see that these 'hotspots' are areas with either high rates of unemployment (red) or low rates of unemployment (blue).

g. How useful will this data be in predicting crime?

1.4. Crime and Qualifications

a. Another indicator of where crimes may occur is the number of people without qualifications in an area.

b. Turn all the layers off. Then turn on 'Sunderland Level of Qualification' to view this dataset.
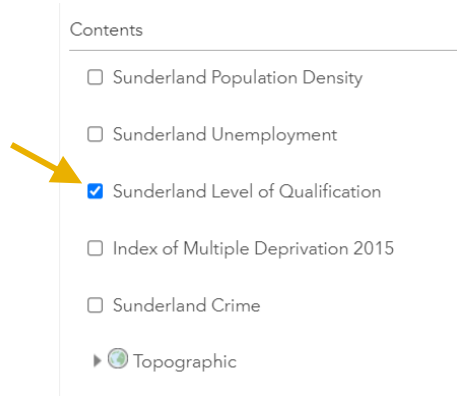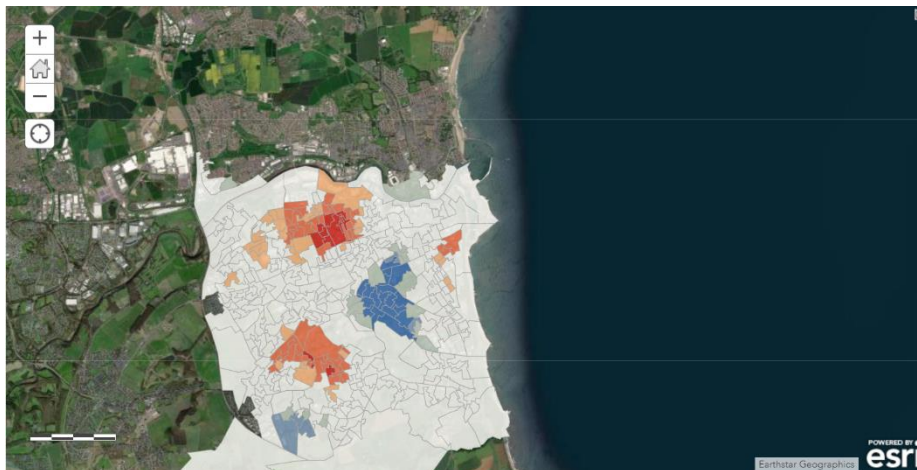
c. The dataset will look something like this:

d. Repeat the steps in Section 1.2 d - e to view the legend to see what the red and blue areas show.

e. We can see that these 'hotspots' are areas where people have no qualifications (red) and areas where people have some qualifications (blue).

f. Do you think this a better or worse indicator of crime?

1.5. Crime and Index of Multiple Deprivation

a. So far, we have only looked at individual census variables to attempt to understand the population characteristics and how these might or might not help us to predict where crimes may occur. It is also possible to combine different census variables to create an index that can attempt to characterise a population based on different census data, giving a single value for each area which we can then use alongside other datasets, such as the crime data.

a. The UK government created an Index of Deprivation which combines a range of census data to produce an index that describes the different amounts of deprivation in each area. This is calculated using a formula that combines several types of data:
    - Income
    - Employment status
    - Education level
    - Health
    - Crime
    - Barriers to Housing and Services
    - Living Environment

    We can use this to try and better predict the areas where crimes are potentially more likely to occur.

b. To access the Index of Deprivation, use the 'Add' tab at the top. Then select 'ArcGIS Living Atlas'. The ArcGIS Living Atlas is a collection of datasets that can be accessed and edited by the user community.

c. Search for 'Index of Deprivation 2015'. When Selected click 'Add to Map', this will load the data onto your map.
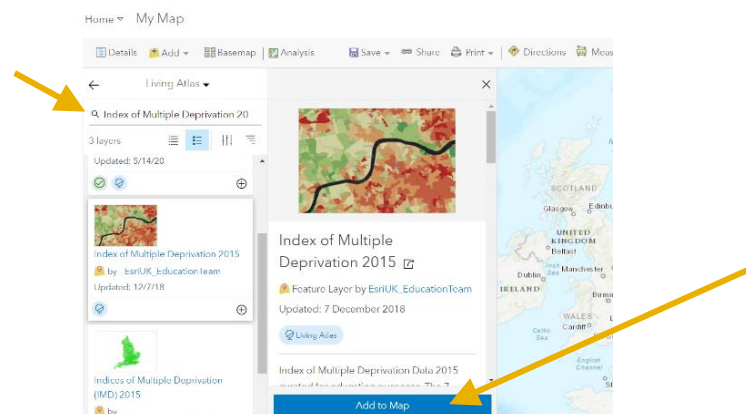
d. Repeat the steps in Section 1.2 d to see what the red and green values mean.

e. The lower values (red, <1) show the most deprived areas and the higher values (green, >10) show the least deprived areas.
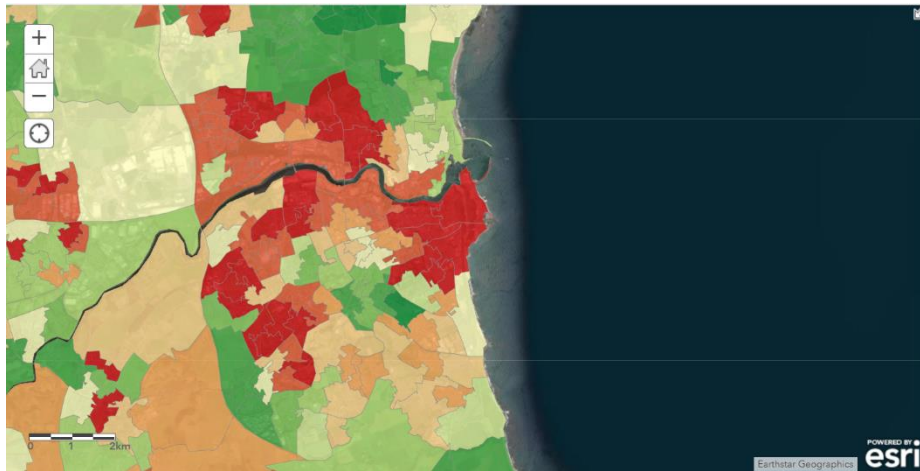
f. Where are the most deprived areas located? Where are the least deprived areas located?

g. This layer covers the whole of England so navigate using the mouse to your home area and check if the data matches with your perception of which of your local areas would be more or less deprived.

h. Each of the above data types are useful indicators of deprivation although for the rest of this practical we will focus mainly on crime data. We will access actual crime data and perform some basic analysis before comparing findings with the multiple deprivation index to see if see if there is indeed a relationship between the two datasets which allows us to predict areas where crimes may be more likely.
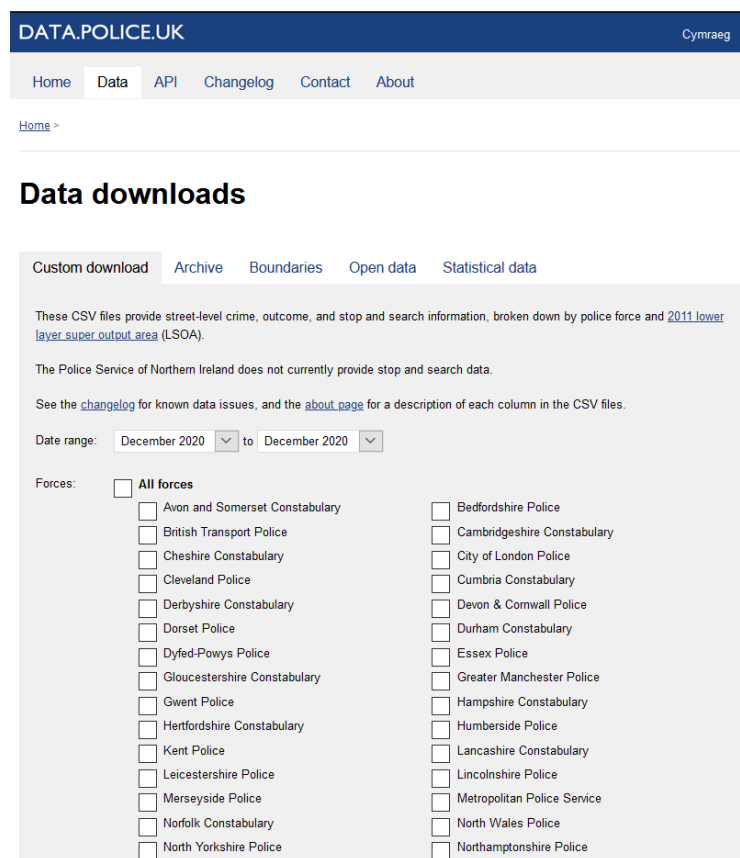
Section 2: Real Time Crime

The Police publish data each month for the preceding months crimes. This is easily accessible and can be downloaded for each borough across the UK.

In this exercise we have used Sunderland in North East England, which falls into the Northumbria Police's area of responsibility, as our example. You can download Police Crime Data relating to your home area by following the same steps below.

2.1. Police Crime Database

a. Go to https://data.police.uk/data/. The page should look like this:

b. Set the date range from January 2021 to January 2021. This means that only one month's data will be shown, which helps to limit the size of the dataset.

If you are choosing your own dataset, make sure to limit the dataset to span one month. If you are having difficulty choosing, you could choose your birthday month.

Geospatial UK

c. Select 'Northumbria Police' to access data for Sunderland.

If you are looking at your own area, select the Constabulary that corresponds to your local area.

d. Select the following option of 'Include Crime Data' and click 'Generate File' (you may need to scroll down to see these options).

e. Once processed, a download option should appear on the following page. Click 'Download now'.

f. Save the file and give it a sensible name such as 'NorthumbriaPoliceDataJanuary2021'.

Newcastle University

Change the name of the file depending on the police force and month you have chosen.

2.2. Managing a CSV File

a. Navigate to the file you have just downloaded using the file explorer. Within the folder is another folder which contains an Excel (spreadsheet) document.

b. To use the file, you will need to extract it or 'unzip' it. To do this, right click on the file, then click on 'Extract All'/

c. In the dialogue box that appears, choose an appropriate place to save the file. When the file has been extracted, open it.

d. The Excel file might also be called a CSV file having a '.csv' at the end of their name, which stands for a comma-separated values file, which allows data to be saved in a table format. CSVs are a text file, but can be loaded into Excel and look just like a spreadsheet.

Open this and it should have a similar structure to the image below, but with different contents relating to the area and month you have chosen.
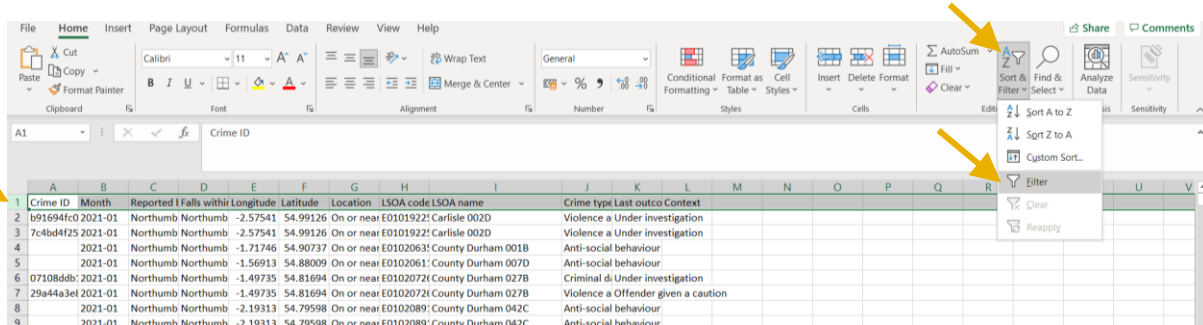
Geospatial UK

e. It may look confusing, however, the file contains data relating to the location, date and type of crime.

f. There may be too many data records to upload directly to ArcGIS online so we must reduce the size of the file.

   We will do this by selecting only one area of interest. Scroll down the Excel worksheet until you find an area/local district that you recognise within the Police area you have chosen (look in the LSOA name column). LSOA stands for Lower Layer Super Output Area and is a way of dividing the country up into manageable sections, with each having a population of around 3000 people.

   In this example, look for 'Sunderland'. You will find lots of rows where the LSOA name contains the name 'Sunderland' as Sunderland is split into many LSOAs, like all towns and cities.
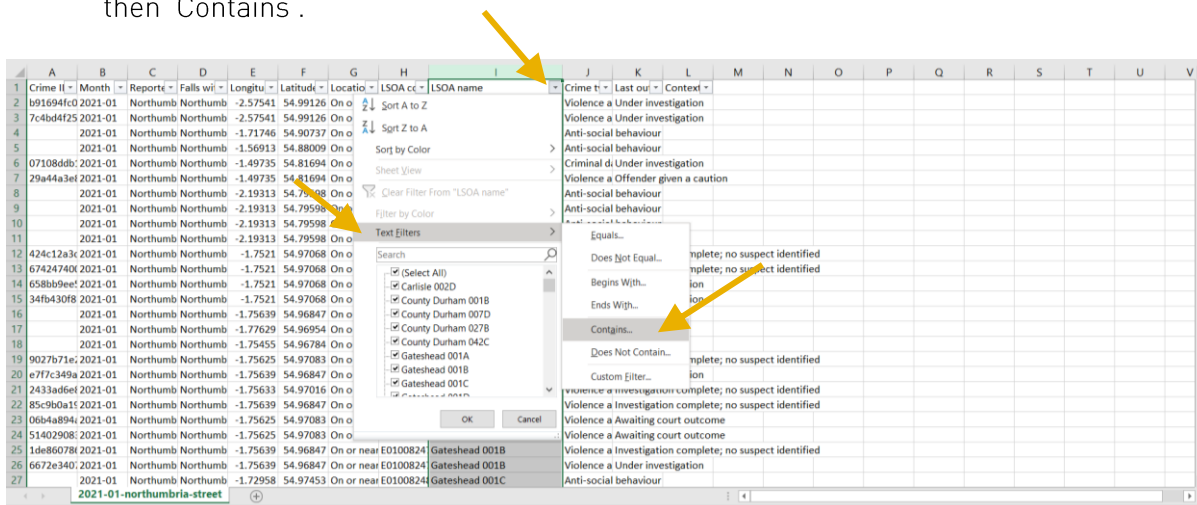
g. Once you have found a relevant area you will need to select all the crimes within this area of interest. To do this, go back up to the top of the file and click on row number 1, to highlight the entire row.

h. The row top row should now look like this (note the drop-down menus in each of the cells):
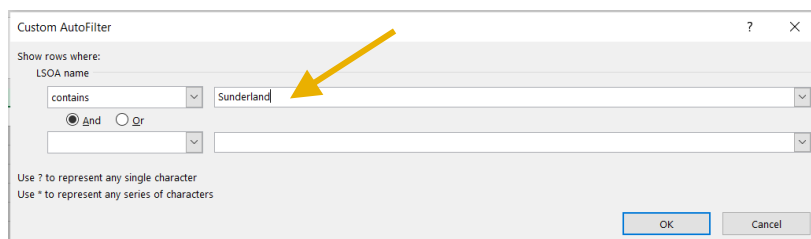
Newcastle University

i. Click on the arrow/drop-down menu for the 'LSOA name' cell. Select 'Text Filters' then 'Contains'.

j. In the subsequent pop up, enter the name of the area that you would like to use, for example 'Sunderland'. This must be spelt correctly. Click OK.
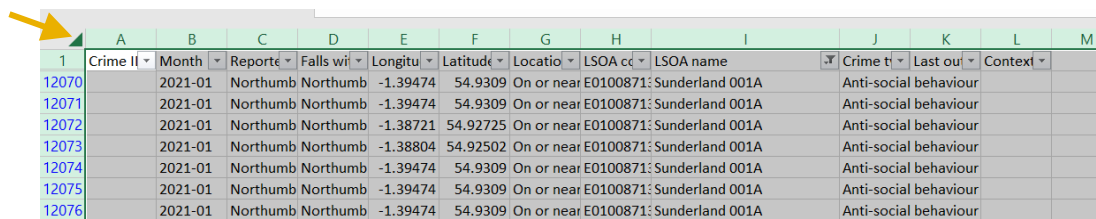

k. Only the rows containing the relevant LSOA name are now selected.

l.   Click on the triangle in the top right-hand corner to select all the data.

m. Then copy it by either right click and 'copy' or pressing ctrl + c on your keyboard.

n.   Click the plus icon at the bottom beside the sheet name. This will create a blank new sheet in the same document.
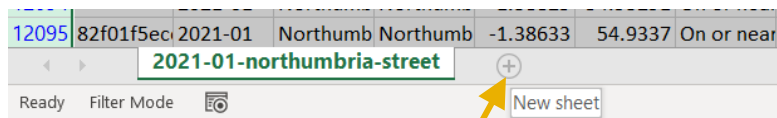
o.   Paste the cells into a new sheet by either clicking 'Paste' at the top or pressing ctrl + v on your keyboard. The populated sheet should now contain only the crimes for the relevant LSOA areas. The example contains only data relating to 'Sunderland'.

p.   Right click on the name of the unfiltered (original) sheet and click delete.
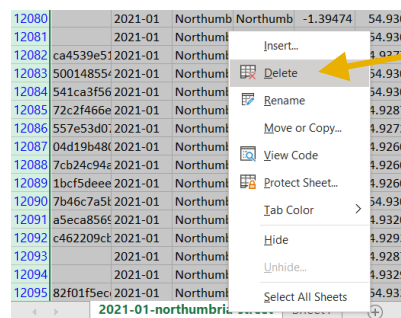
q.   Save the new sheet with a new, sensible name such as 'SunderlandPoliceDataJanuary2021_filtered', making sure it's saved as a '.csv' file.

2.3. Importing a CSV File into ArcGIS

a. Go to the original ArcGIS map online: https://arcg.is/05KXKv0.

b. Click 'Modify Map' in the top right corner.
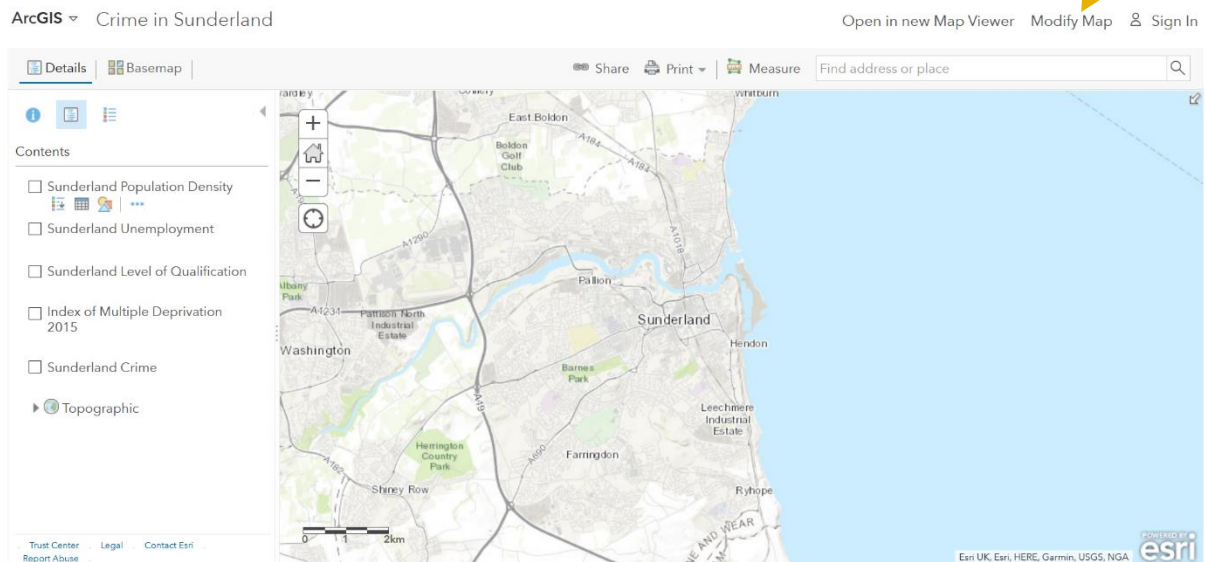
c. Click on 'Add'. Then click 'Add layer from file' from the resulting drop-down menu.
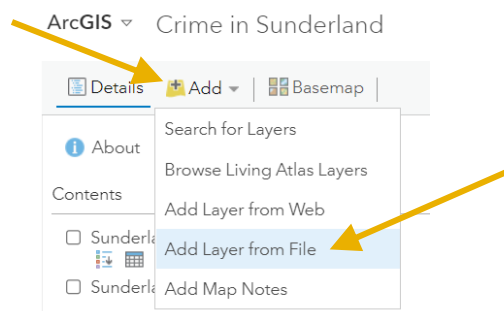
d. Select 'Choose file' from the dialogue box that opens. Navigate to the filtered Excel Document that you have just saved. Then select 'Import Layer'.

Newcastle University
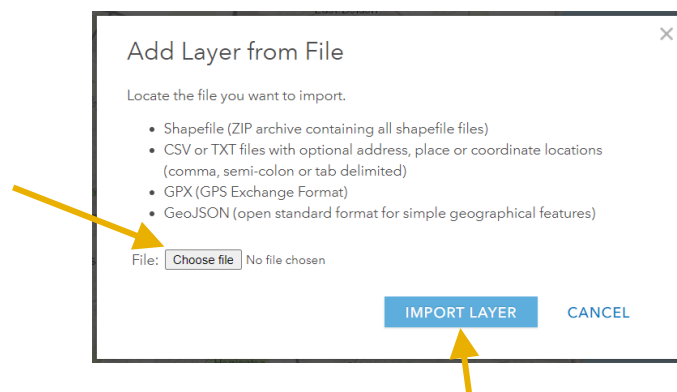
e. The data from the Excel spreadsheet has now been successfully uploaded to ArcGIS online. With the data imported, select the attribute 'Crime Type' to be displayed. Click Done.
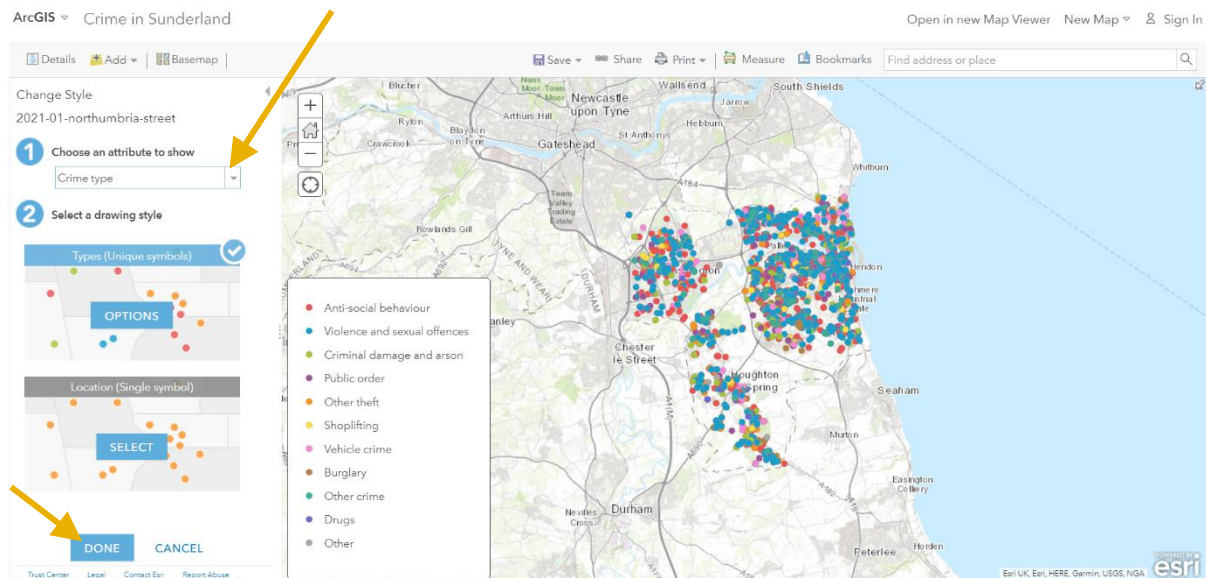
f. We will now explore the details of the crimes more thoroughly in Section 3.

Section 3: Exploring the Scene of the Crime

This section will explore how to add crime data to our map to find out the location and type of crimes committed.

3.1. Using Symbology

a. We can find out more about each dot on the map.

   Each of the points show where the crime occurred. It is important to recognise that the crime locations are published at the postcode level which usually covers a number of streets. This means that even if a crime appears to be located at someone's door, it is simply the middle of the postcode area, rather than the exact location of the crime.

b. Explore the dataset by clicking on any of the coloured dots, a pop up will appear that provides detailed information about the crime.
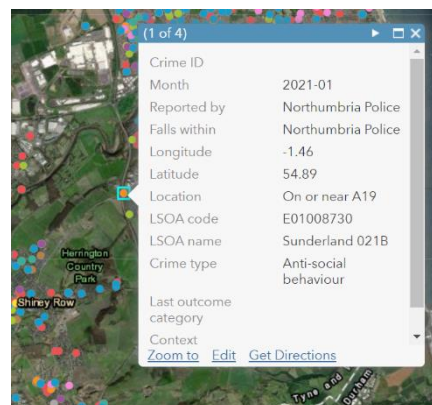
c. You can see that there are a lot of different colours of dots on the map, but what do each of the colours mean? To find out, click on the 'Show Map Legend' tab.

d. Using this legend, which crime do you think is most common? Do certain crimes correspond with certain locations, for example, shoplifting might be more common in town centres?

3.2. Individual Incidents

a. Each dot on the map represents one incident that a Police Officer has responded to.

b. To find out more about each incident, click on the 'Show Table' button.

c. You can find out more about each case using the attribute table. The attribute table is useful for selecting only certain types of crime or for sorting data, making analysis easier.

d. A table will appear in the map window and you can extend the table to look at the details by dragging the small horizontal bar upwards.

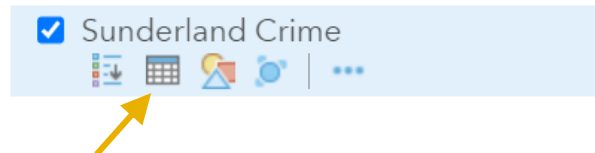| Crime ID	Month	Reported by	Falls within	Longitude	Latitude	Location
	2021-01	Northumbria Police	Northumbria Police	-1.39	54.93	On or near Sports/Recrea Area
	2021-01	Northumbria Police	Northumbria Police	-1.39	54.93	On or near Sports/Recrea Area
	2021-01	Northumbria Police	Northumbria Police	-1.39	54.93	On or near Ho Drive
	2021-01	Northumbria Police	Northumbria Police	-1.39	54.93	On or near Ull Grove

e. Investigate the attribute table. We can now see when the incident was reported, who it was reported by, the location of the incident, the type of crime and the last outcome if any.

f. GIS is not only about visually mapping areas, but we can also examine data and tables known as 'attribute tables', this provides us with additional information about each of the crimes that is included within the dataset.

g. Press the X to close the table when you are finished.

3.3. Heat Maps

a. When you hear the words 'heat map' you might think about Police Officers wearing thermal-sensing goggles, however when we say 'heat map' it means a map which shows the areas which have the highest crime rates.

We have already looked at a number of 'hotspots' earlier in this practical and a heat map is a similar method of displaying data. We will use the heat map to examine the spatial distribution of crime. The hotspot maps are defined by census area boundaries which results in a segmented display. Heat maps are useful for displaying data in a more continuous, aesthetically impactful manner.

b. To do this, click on the 'Change Style' button.

c. Change the attribute to 'Show Location Only' instead of 'Crime Type'.
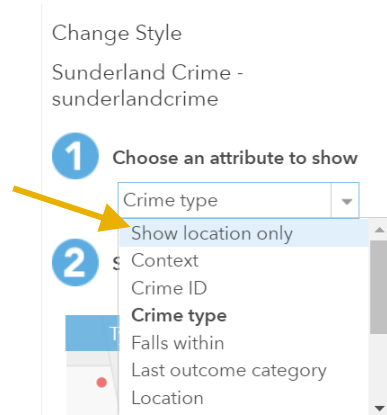
d. To convert these dots into a heat map, select 'Heat Map' as the drawing style instead of 'Location (single symbol)'.

e. Click 'Done' to continue. Your map will now look something like this:

f. To view the legend, repeat Section 3.2 a and b.

g. We can now see that the areas in white have the highest crime rates, while the areas in blue have the lowest crime rates.
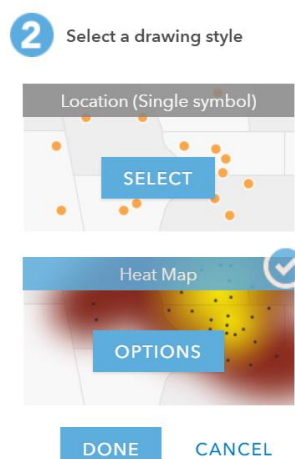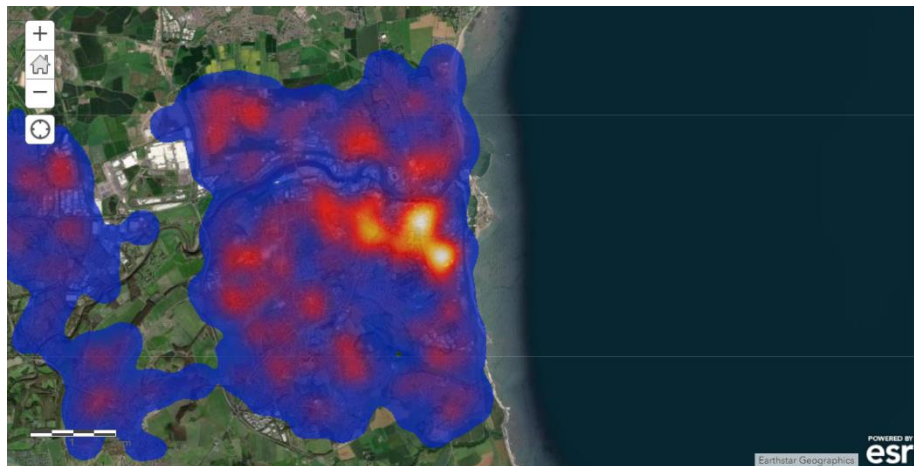
    This data is visualised using a dynamic scale, allowing you to zoom in and out hence providing a better understanding at both small and large map scales.

h. Do the locations of high crime rates match the areas that you predicted using indicators earlier?

i. Turn the other layers 'on' and 'off' or use the transparency slider to compare and contrast the areas of high crime rates against the variables of population, employment, qualifications and the Index of Multiple Deprivation. To change the transparency, click on the three dots under the layer name select 'Transparency' before using the slider to make layers more, or less, visible.
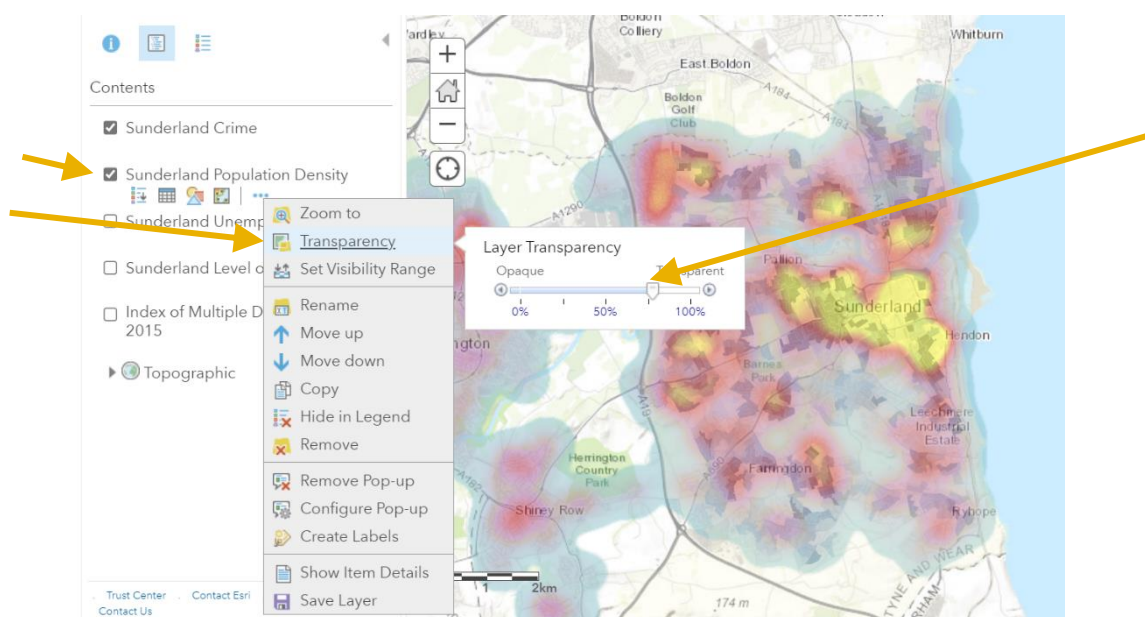
j. Do you feel there is a strong degree of correlation between the various datasets?

Newcastle University

Section 4: Summary

In this exercise we have explored some of the potential analysis techniques available to police forces in helping to make society safer. The two videos below show South Wales Police use GIS to support their work. During this exercise we have used some of the same techniques implemented by the analysts in the videos, most notably hotspot mapping.

- [BBC Two - Mapping Crime, An introduction to GIS](#)

- [BBC Two - Mapping Crime, Using GIS to target vehicle crime in Swansea](#)

Police forces all over the world are constantly adopting GIS methods to support their strategy and policy development. A great example of how GIS can be implemented to successfully reduce crime exists in Connecticut, USA. Hartford police department used similar methods to those described in the videos and those discussed in this practical. The results were dramatic; over just a three-year period GIS contributed to strategies that resulted in a 34% reduction in burglaries, a 35% reduction in car thefts and also supported schemes that led to a 29% reduction in crimes relating to stolen registration plates ([Esri](#)).

If you want to investigate further, repeat this practical with a different borough in the UK. Use (https://data.police.uk/data/) to find this information.

This concludes the exercise.

Geospatial UK

This activity was created by Newcastle University on behalf of Geospatial UK.

For more resources or activities, visit

www.geospatialuk.org

Newcastle University