

# Geography for Data Science and Analysis

**Donna Viney and Debbie Rhodes**  
ONS Geospatial

©Photo by NASA on Unsplash

Please use  
Slido  
**#DSGP**  
for questions

# Who are we?



# ONS Geospatial

Data Capability (Dcap) →

Data Growth and Operations (DGO) →

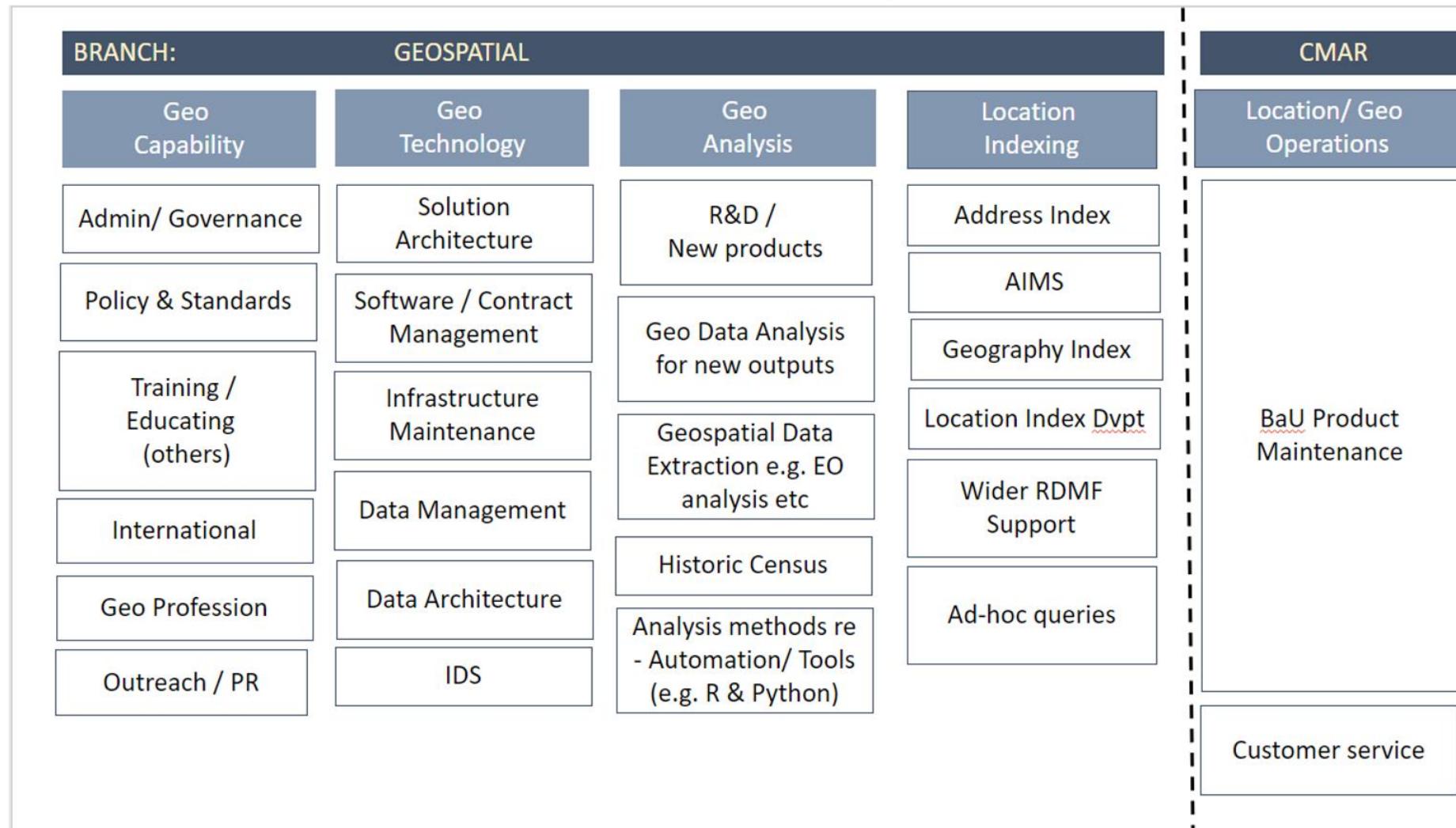
Data Architecture, Location and Integration (DALI) →

**ONS Data Architecture,  
Location and Integration**



**Geospatial**

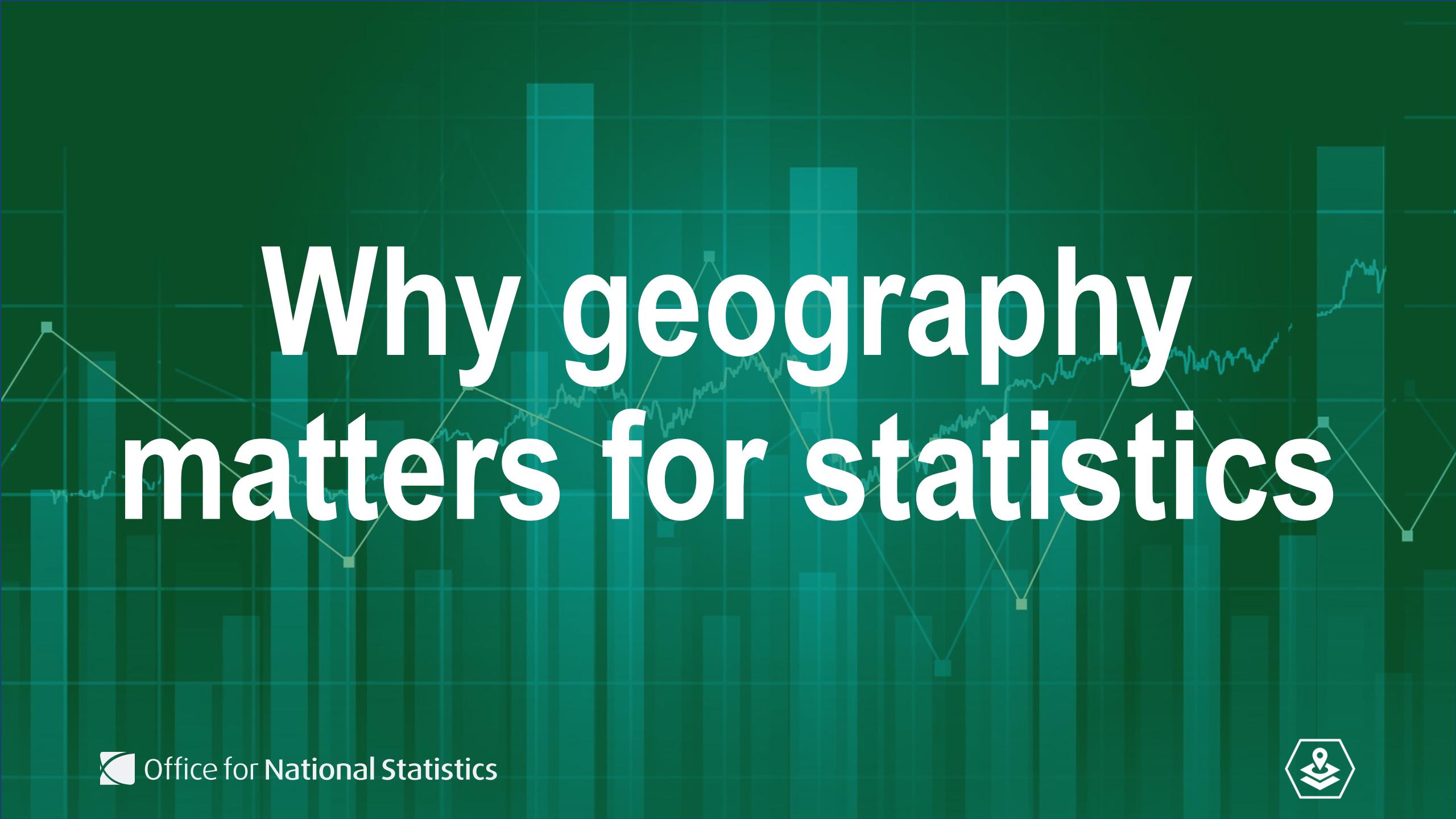
# Areas of responsibility



# Outline for today

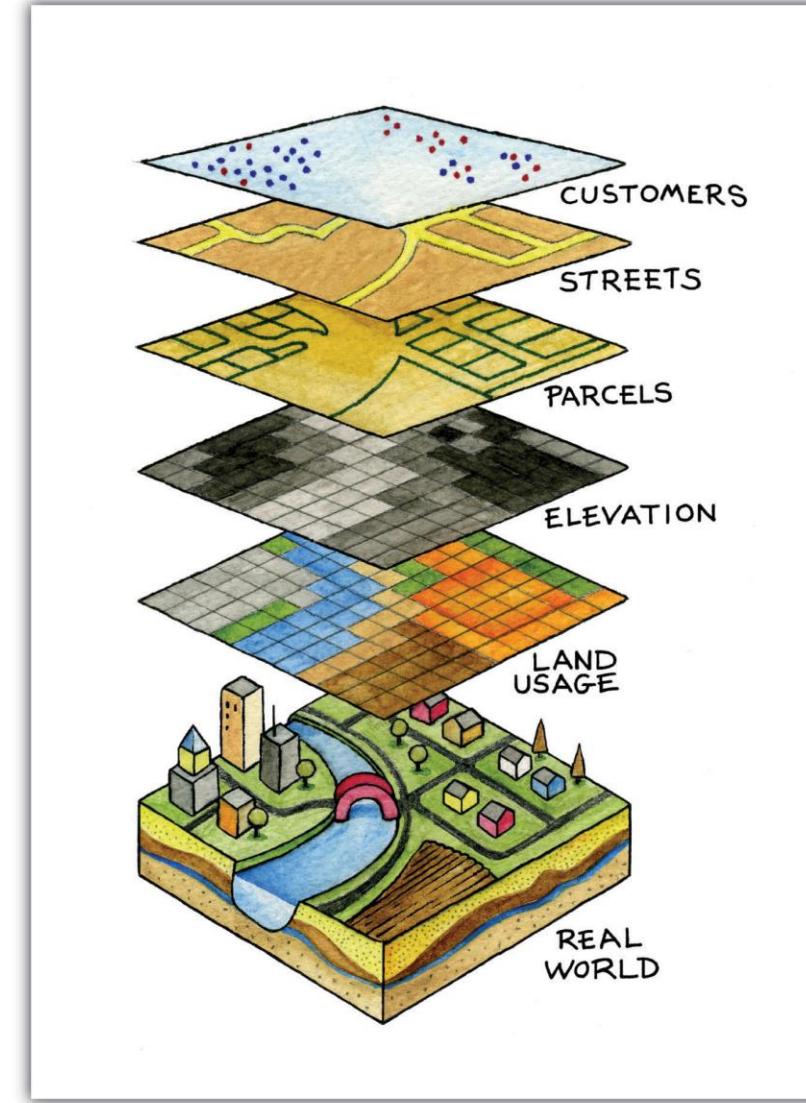
# Topics to be covered

- Why geography matters for statistics
- Geography policies
- Working with geographic data
- Geographic products from ONS
- Geospatial techniques and mapping
- Geographic fallacies
- Geospatial ethics



# Why geography matters for statistics

# Everything happens somewhere...



[https://saylordotorg.github.io/text\\_essentials-of-geographic-information-systems/section\\_11/ca6ce94cdd2e09a1da8aa6ec22336835.jpg](https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/section_11/ca6ce94cdd2e09a1da8aa6ec22336835.jpg)

# The great integrator



# The complexity of UK geography

- Lots of change
- Many organisations involved



Ministry of Housing,  
Communities &  
Local Government



Department  
for  
Business, Energy  
& Industrial Strategy



Department  
of Health &  
Social Care



Home Office



Boundary  
Commission  
for England

The  
Local Government  
Boundary Commission  
for England

NHS



Royal Mail

StatsWales



Scottish Government  
Riaghaltas na h-Alba  
gov.scot

NISRA  
Northern Ireland  
Statistics and Research Agency  
Gníomhaireacht Thúisneachta Éireann  
um Statistíod agus Taighde

National  
Parks



Office for National Statistics



# Using geography requires care

- Time series is challenging
- Quality and integrity of statistics

# Geography Policies



# GSS Geography Policy

- Policies and principles
- GSS geography policy



**GSS Geography  
Policy**

March 2015



# Official Statistics for each UK geography

✓ Consistent

✓ Comparable

✓ Non-disclosive

## Statistics

### REFERENCING

Reference source data at lowest possible geographical level using a standard identifier.

1

### NAMING AND CODING

Use GSS standard codes and names for UK statistical geographies.

2

### MANAGING CHANGE

Apply changes to geographies once a year.

3

### BUILDING BLOCKS

Build official statistics for any geography from whole statistical building blocks.

4

### AREA MEASUREMENT

Apply the right area measurements for each statistical geography e.g. for population density.

5

### CLASSIFICATIONS

Use the right type and currency of geographical classification for your statistical outputs.

6

### PRESENTATION

Use standard order for presenting geographical areas in tables, and best practice for mapping and presentation.

7

## Standards

## Data

### Reference data, metadata and best practice

Products on Open Geography portal at: <https://geoportal.statistics.gov.uk>

Linked Data at <http://statistics.data.gov.uk>



Office for National Statistics



# Some key geography policies (1)

- **REFERENCING:** Lowest possible geographical level
- **NAMING AND CODING:** Use GSS codes and standard names
- **PRESENTATION ORDER:** Follow the GSS guidelines

# Some key geography policies (2)

- **STANDARD AREA MEASUREMENT:** use the official area measurements
- **CLASSIFICATIONS:** Use the geographic classifications

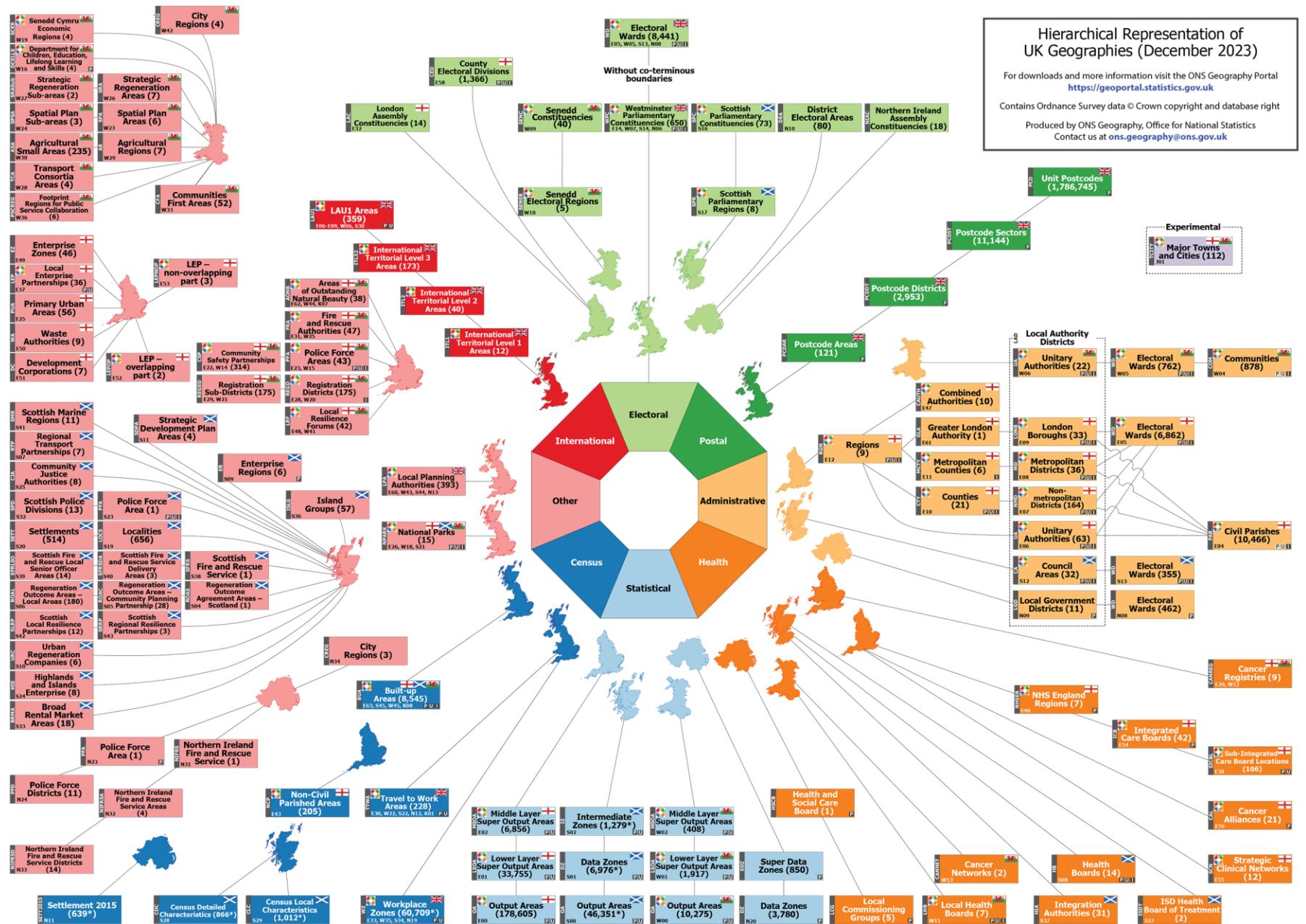
# Questions?

Please use Slido #DSGP for questions



# Geographic Products from ONS

# Hierarchical Representation of UK Geographies



# Administrative Geographies

# Regions

- 9 regions in England
- No regions in Scotland, Wales or Northern Ireland
- Local Authority Districts nest within the regions



# Local government - 2023

- Local government structure varies across England
- Local Authority Districts = LTLA

## Lower Tier Local Authorities (LTLA)

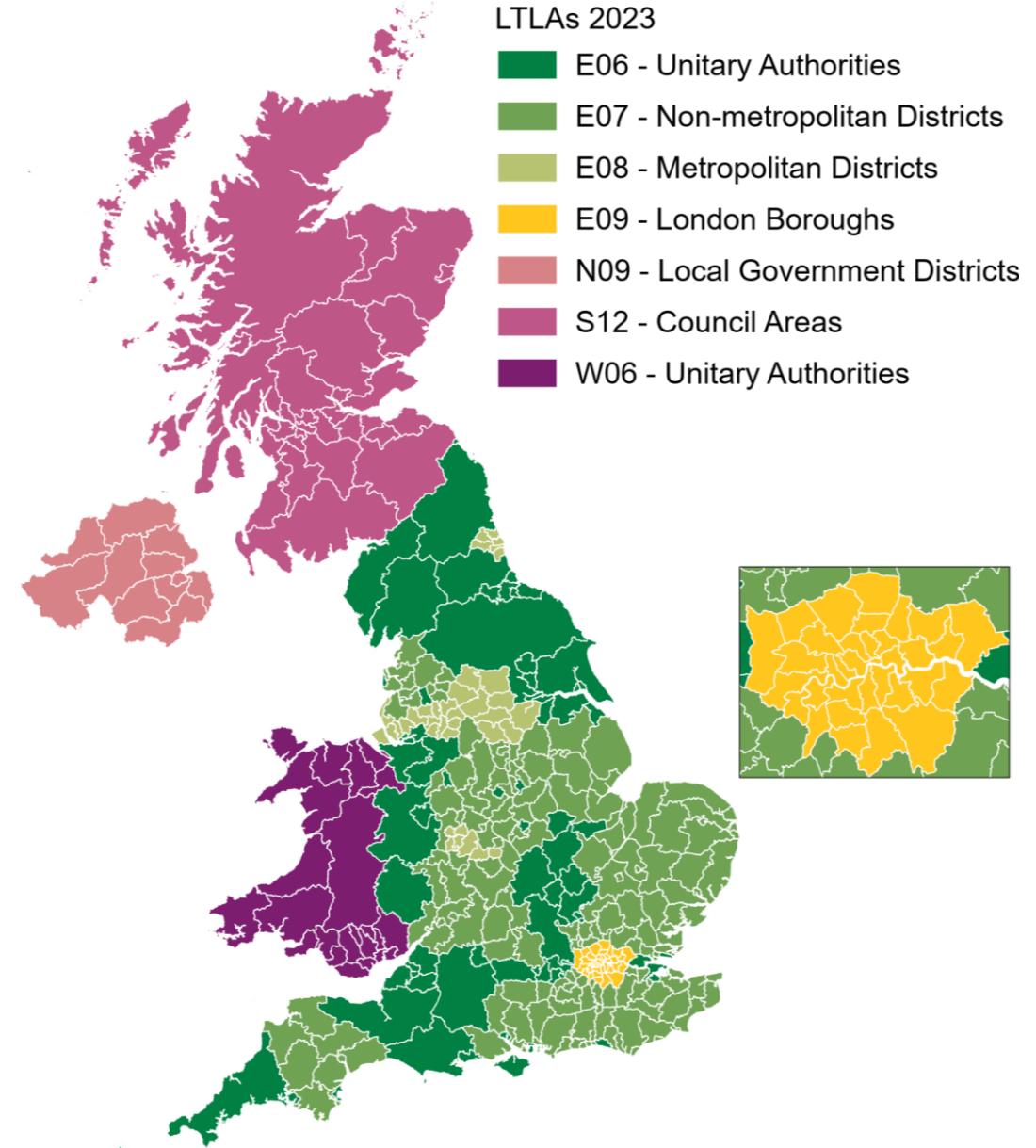
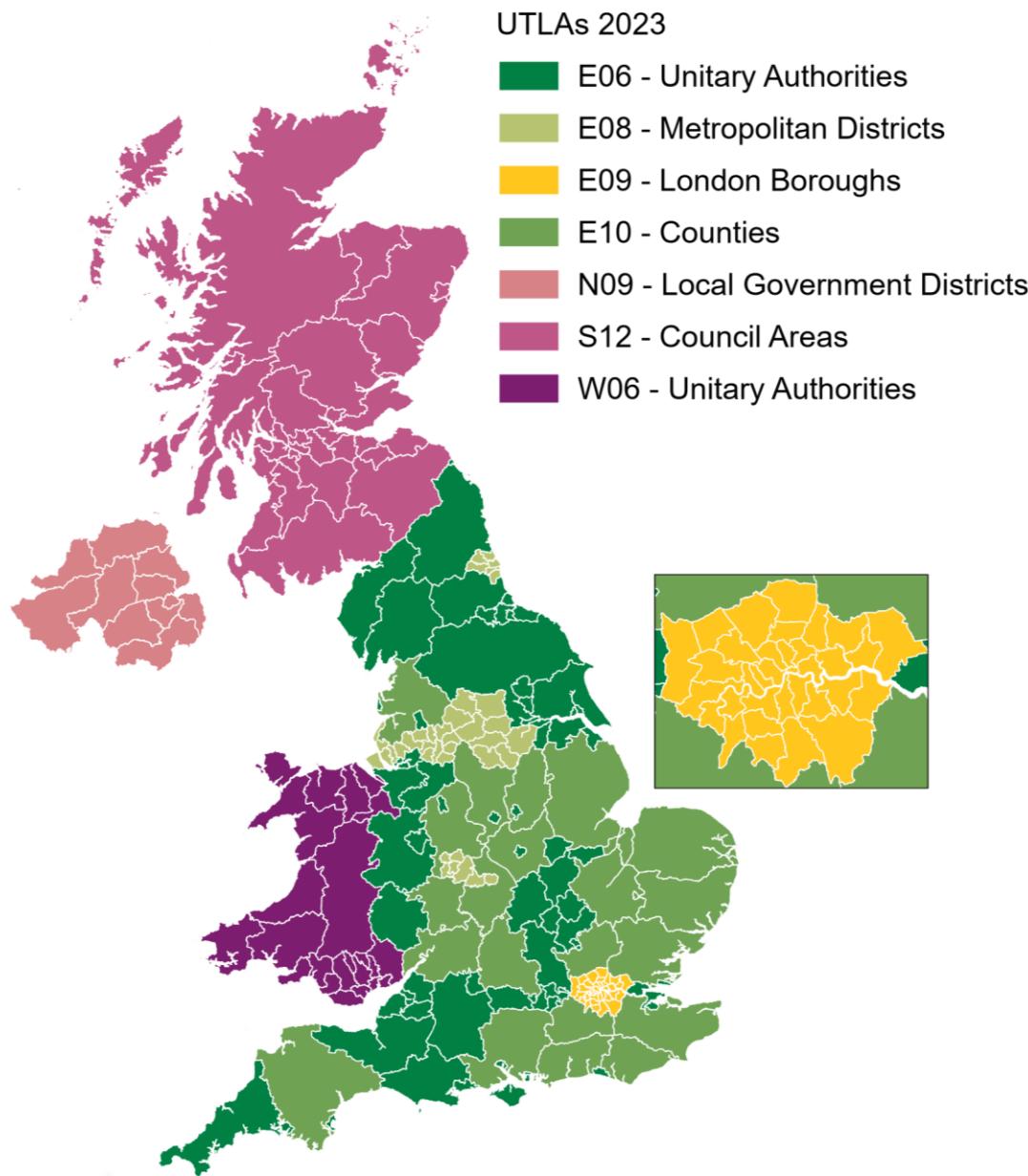
Total: 361

- England – 296
- Wales: 22
- Scotland: 32
- Northern Ireland: 11

## Upper Tier Local Authorities (UTLA)

Total: 218

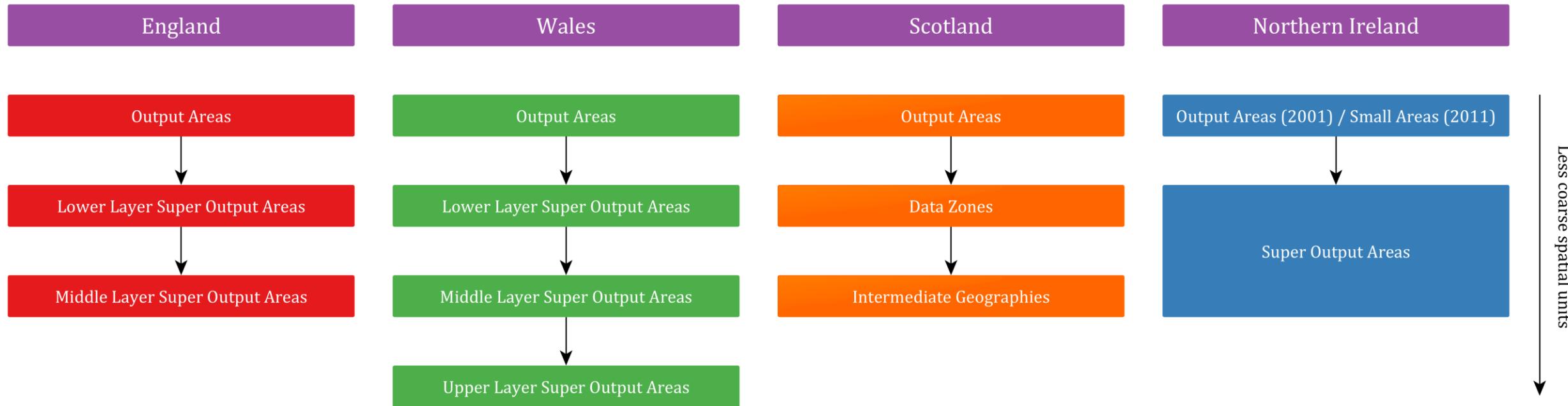
- England: 153
- Wales: 22
- Scotland: 32
- Northern Ireland: 11



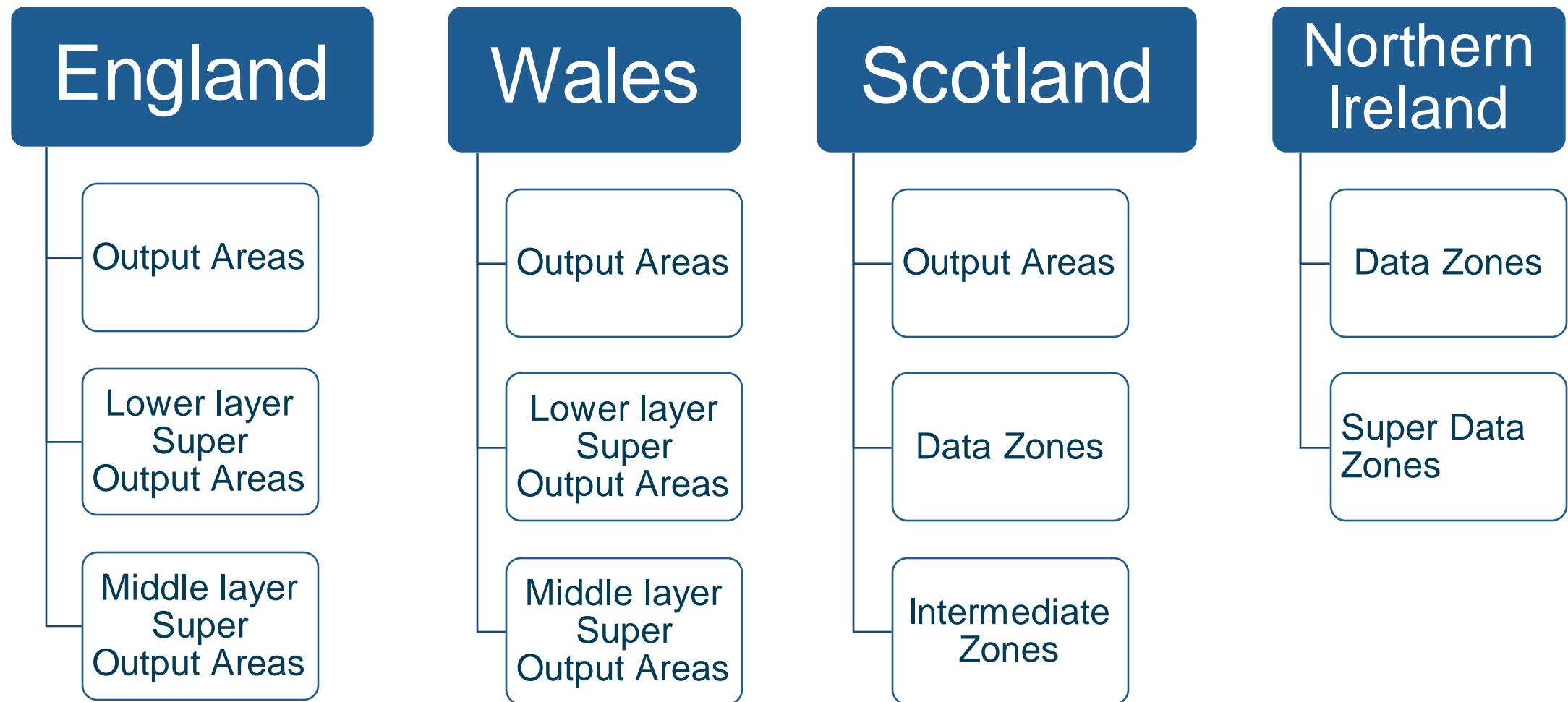
# Statistical Geographies



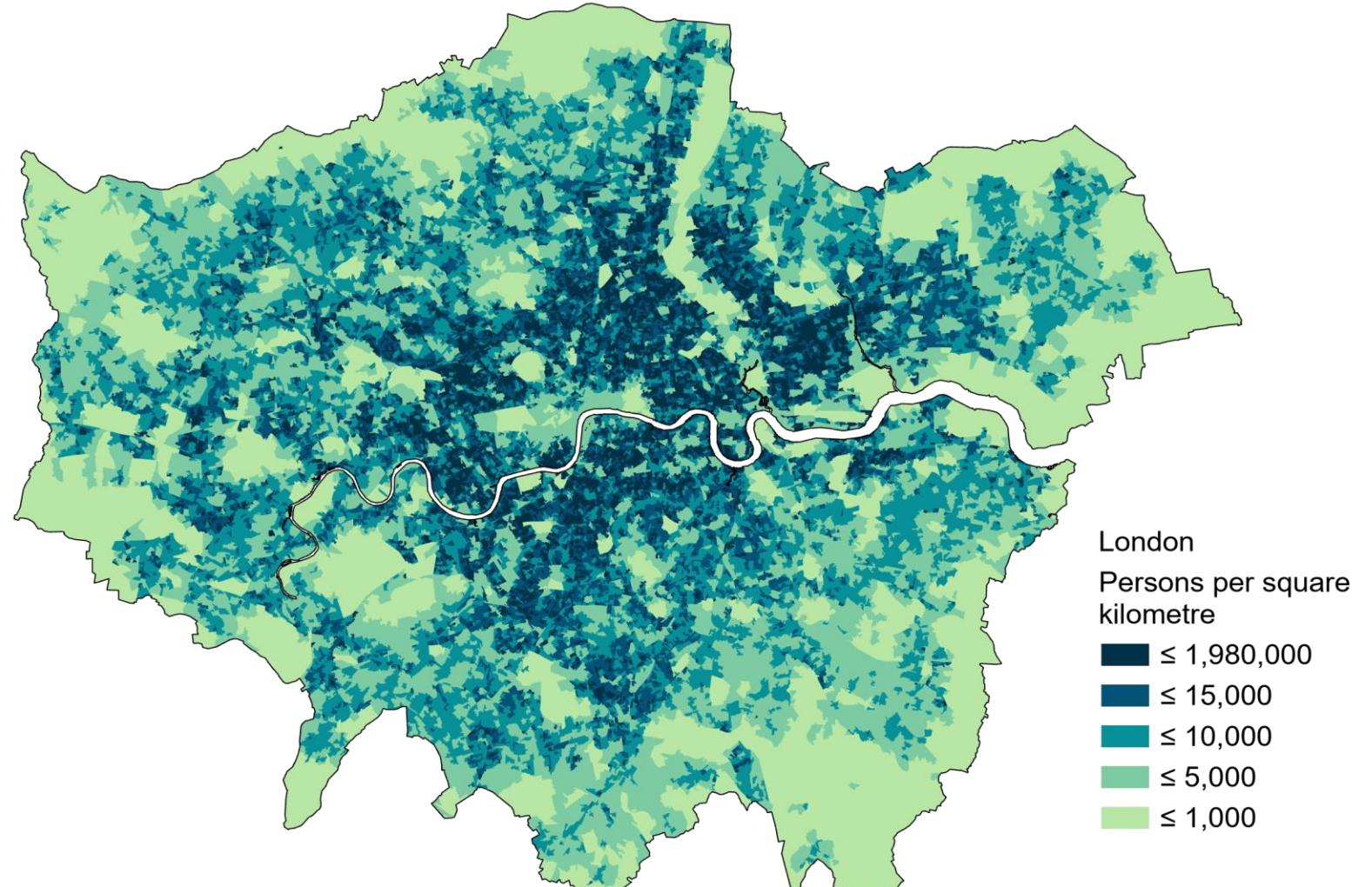
# Statistical Geographies for 2001 and 2011



# Statistical Geographies for 2021

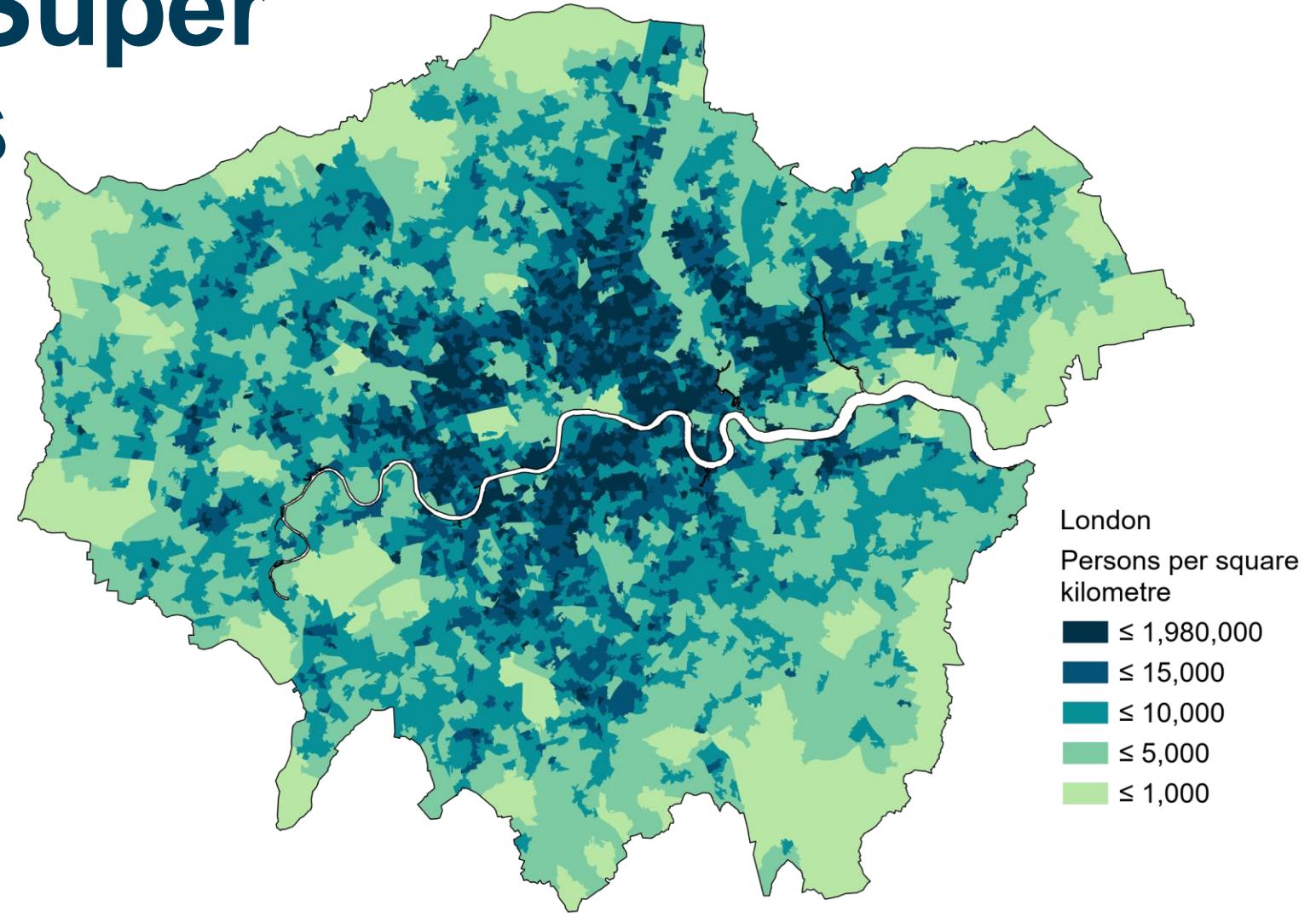


# Output Areas



Contains OS data © Crown copyright and database right 2024  
Source: Office for National Statistics licensed under the Open Government Licence v.3.0

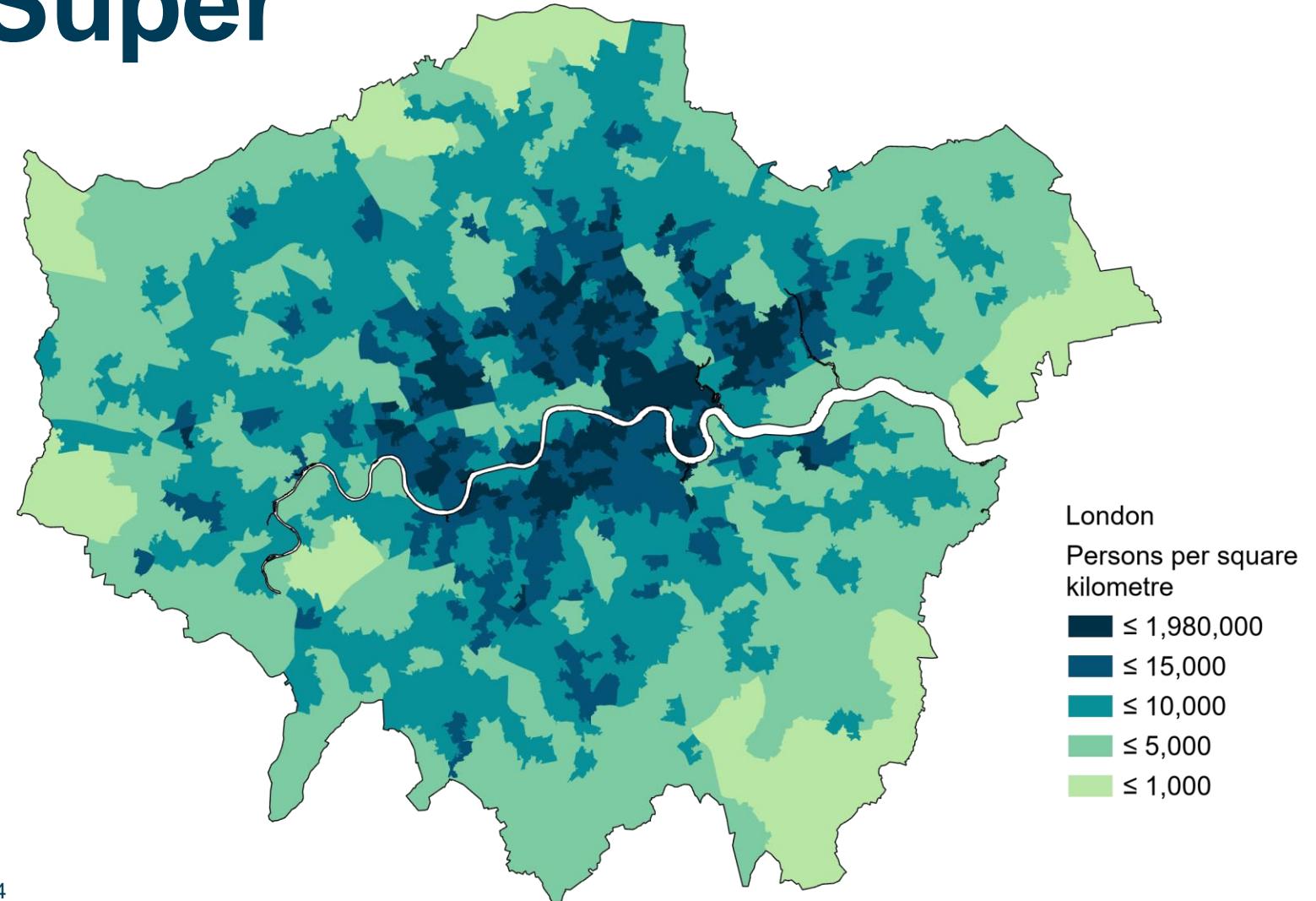
# Lower layer Super Output Areas



Contains OS data © Crown copyright and database right 2024

Source: Office for National Statistics licensed under the Open Government Licence v.3.0

# Middle layer Super Output Areas



Contains OS data © Crown copyright and database right 2024

Source: Office for National Statistics licensed under the Open Government Licence v.3.0

# Output Areas and Workplace Zones

## Output Areas/NI Data Zones    Workplace Zones

- |   |  |
|---|--|
| <ul style="list-style-type: none"><li>• Small areal units of UK Census Geography</li><li>• Residential population</li><li>• Building blocks</li></ul> | <ul style="list-style-type: none"><li>• Workplace statistics</li><li>• Workplace population/workplaces</li></ul> |
|---|--|

# Postcodes and UPRN

# Postcode Directories

## National Statistics Postcode Lookup (NSPL)

- Geography information for each postcode
- Best-fit method
- **Statistical production**

## ONS Postcode Directory (ONSPD)

- Geography information for each postcode
- Direct point-in-polygon assignment
- **Operational use only**

# UPRN

- Unique Property Reference Number
- 12 digit reference number
- Assigned by Ordnance Survey



# UPRN Directories

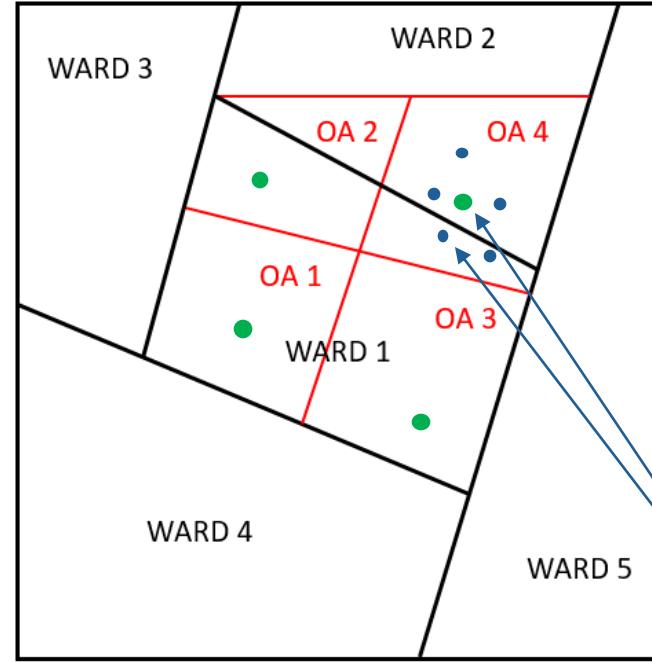
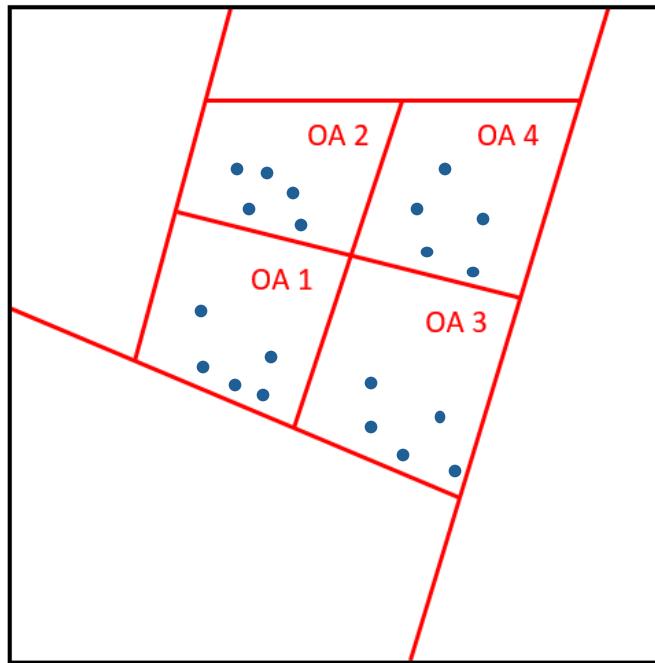
## National Statistics UPRN Lookup (NSUL)

- UPRN to statistical geographies
- Best-fit method
- **Statistical production**

## ONS UPRN Directory (ONSUD)

- UPRN to statistical geographies
- Direct point-in-polygon assignment
- **Operational use only**

# Explaining Best-Fit



Best-fit assignment

OA-Ward Lookup	
OA	WARD
OA 1	WARD 1
OA 2	WARD 1
OA 3	WARD 1
OA 4	WARD 2

NSPL/NSUL: ward 2  
ONSPL/ONSUD: ward 1

# Other Geographical Products

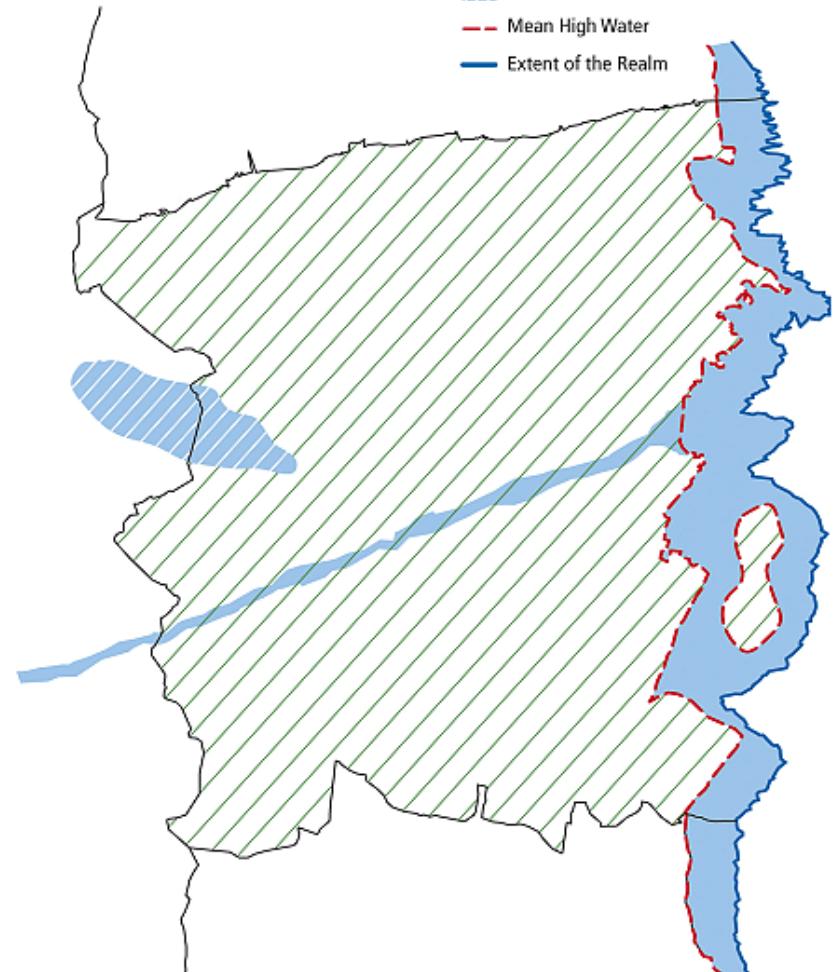
# Look-ups and code lists

- Links between different geographies
- GSS geography codes
- Postcode or UPRN Directories
- How different entities nest or relate to each other

# Standard Area Measurements

DIAGRAMMATIC REPRESENTATION  
OF THE FOUR TYPES OF AREA MEASUREMENTS

KEY  
Land Area  
Inland Water  
Mean High Water  
Extent of the Realm



EXAMPLE OF THE LAYOUT OF THE AREA MEASUREMENTS  
AVAILABLE

WD11CD	WD11NM	AREAEHECT	AREACHECT	AREAIHECT	AREALHECT
E05000193	Bowes	148.34	148.34	0	148.34
E05000194	Bush Hill Park	252.45	252.45	0	252.45
E05000195	Chase	1693.54	1693.54	0	1693.54
E05000196	Cockfosters	1041.28	1041.28	0	1041.28
E05000197	Edmonton Green	312.64	312.64	0	312.64

WD11CD - Ward (as at end December 2011) code  
WD11NM - Ward (as at end December 2011) name

AREAEHECT - Area measurement to extent of the realm boundaries (mean low water mark) in hectares

AREACHECT - Area measurement to coastline feature boundaries (mean high water mark) in hectares

AREAIHECT - Area measurement of inland water features larger than 1km<sup>2</sup>

AREALHECT - Area measurement of land area only (to coastline features and excluding inland water features larger than 1km<sup>2</sup>)

**AREALHECT**  
recommended  
Eurostat & ON  
denominator for  
calculating stat  
population den



# Rural-Urban and Built-Up Areas

- Multiple methods for defining ‘urban’ and ‘rural’.
  - Rural Urban Classification (RUC)
  - Built-up Area (BUA)
  - Major Towns & Cities

# Area classifications

- Where people live or work
- Characteristics and attitudes
- Similar people are likely to live within the same locality
- Area types will be distributed in different locations

Rural Residents

Cosmopolitans

Ethnicity Central

Multicultural  
Metropolitans

Urbanites

Suburbanites

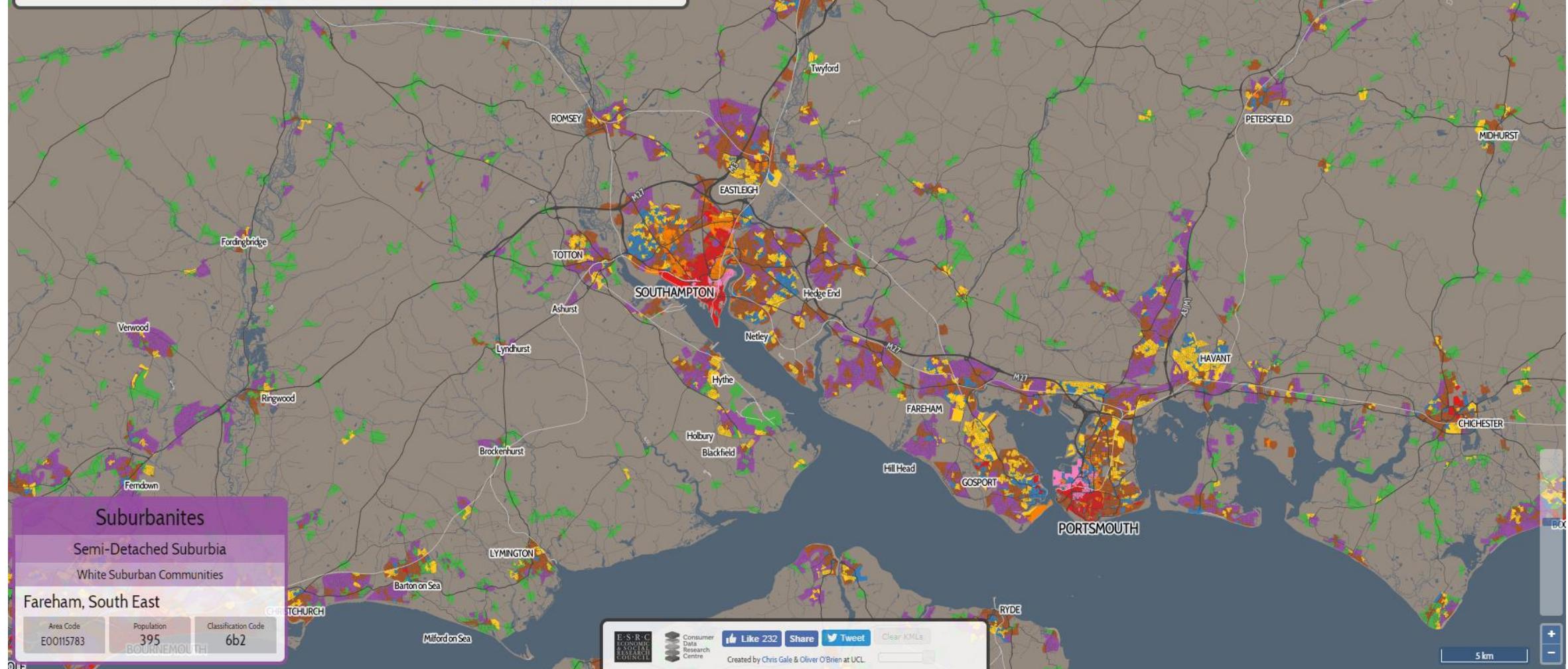
Constrained City  
Dwellers

Hard-Pressed  
Living

Important note: Classifications are an average across the local area, rather than for individual houses, therefore the colour coding on a building is not necessarily indicative of that building.

See pen portraits and radial chart profiles from the ONS for each of the classifications used on this map.

Don't agree with your local classification? You can choose a better one at Open Geodemographics.



# Questions?

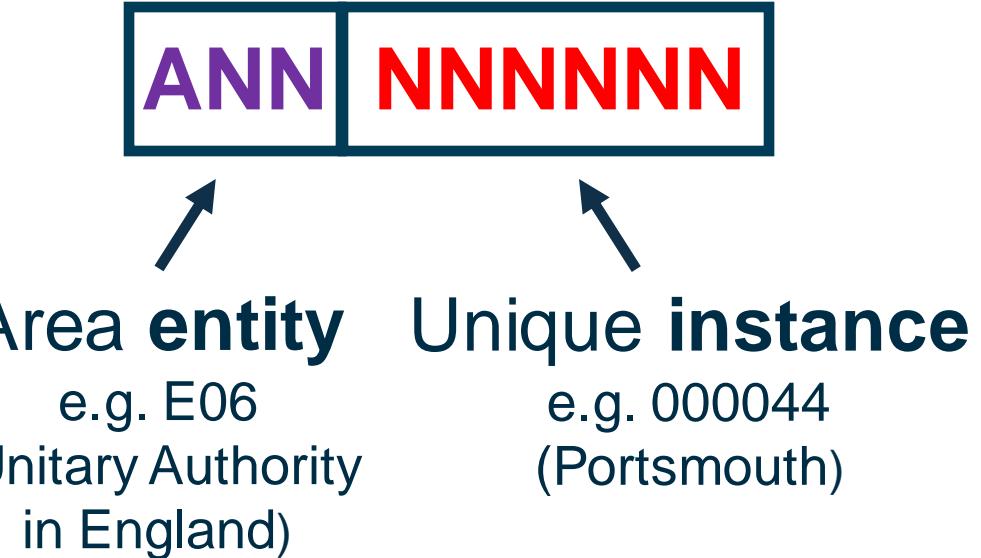
Please use Slido #DSGP for questions



# ONS Geography product features

# Standard names and codes

- GSS codes
- Register of Geographic Codes and the Code History Database
- Uniform Resource Identifier – e.g. <http://statistics.data.gov.uk/id/statistical-geography/E06000044>)

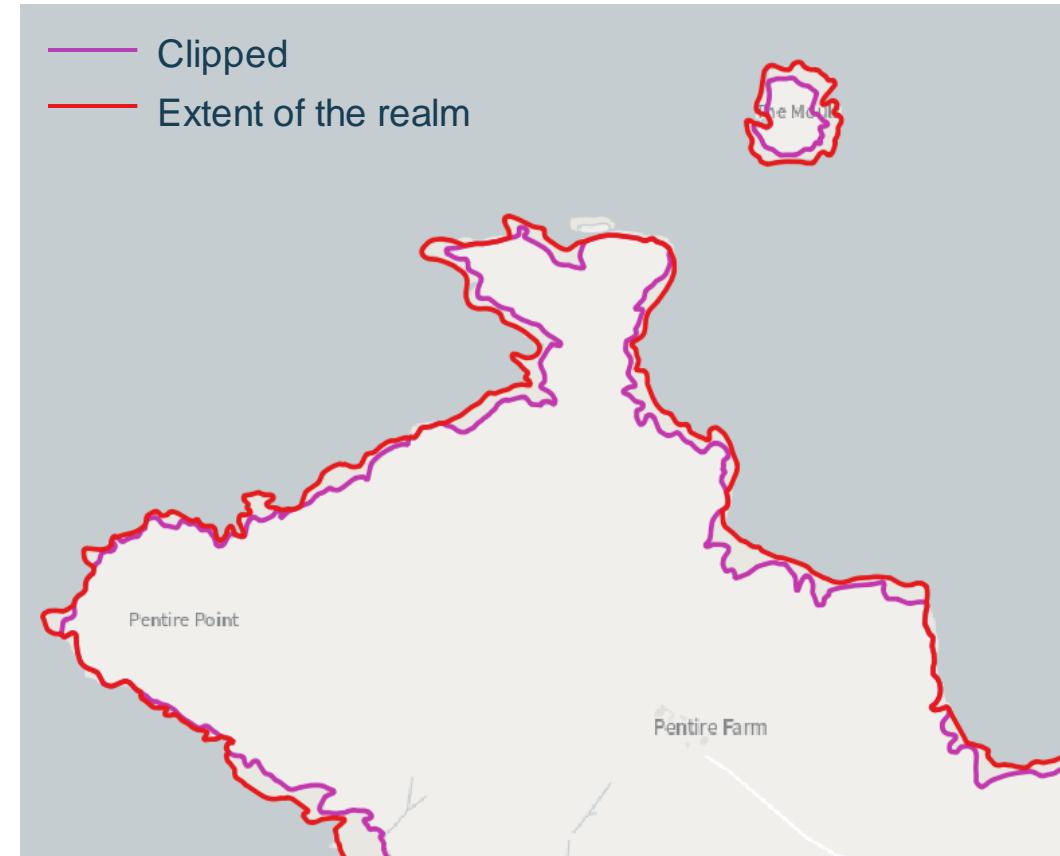


# Referencing areas correctly

- Link using codes
- Beware of linking data using names
  - There are several Newport's in the UK
  - Hundreds of 'High Street'
  - Several 'Castle' Wards

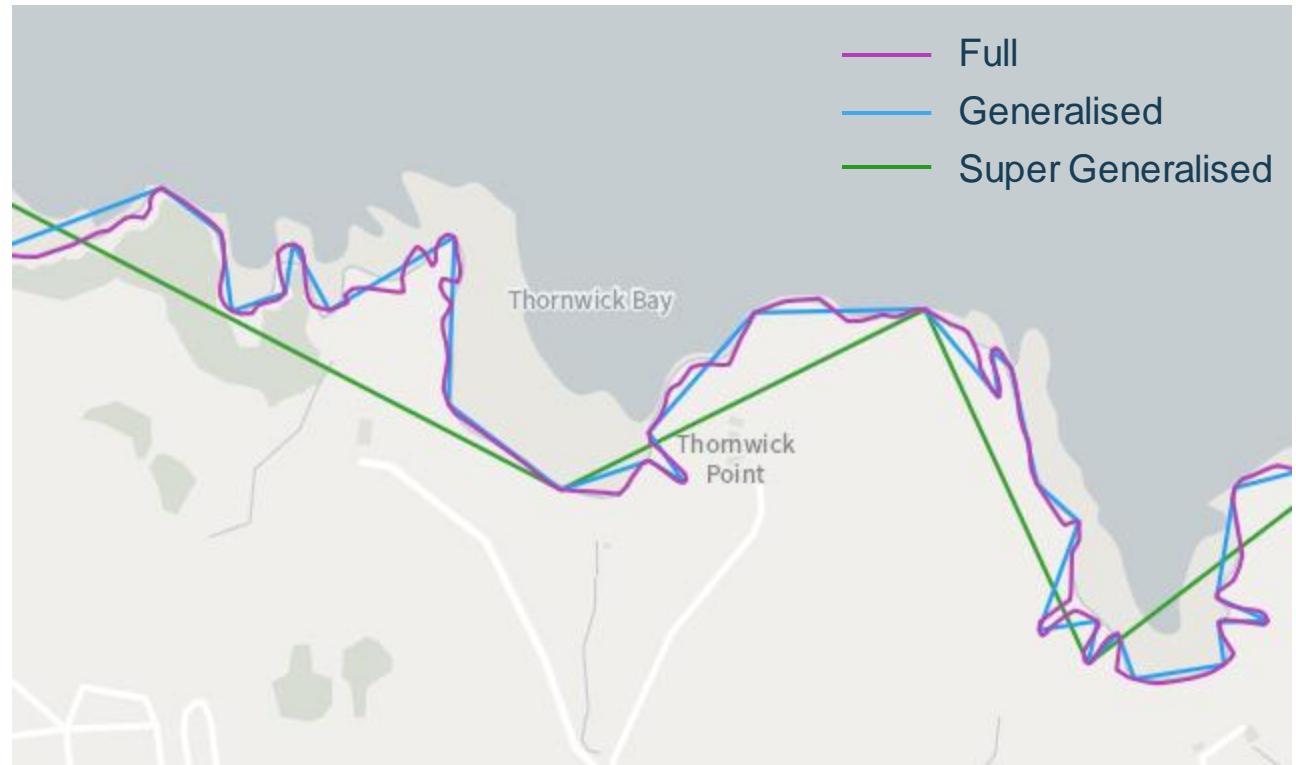
# Geographic Extent

- ONS boundaries are created to two extents:
  - C: clipped to coastline
  - E: extent of the realm



# Resolution

- Generalising - Removing points
- Four boundary resolutions are available
  - F: Full resolution
  - G: Generalised – 20m
  - S: Super generalised – 200m
  - U: Ultra generalised – 500m



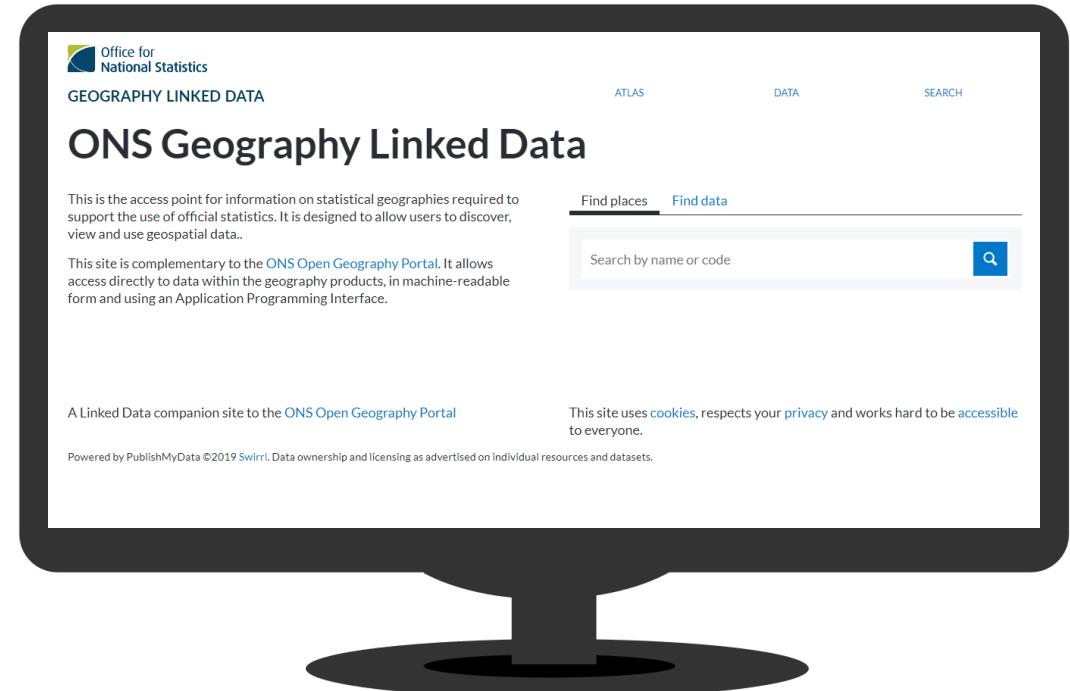
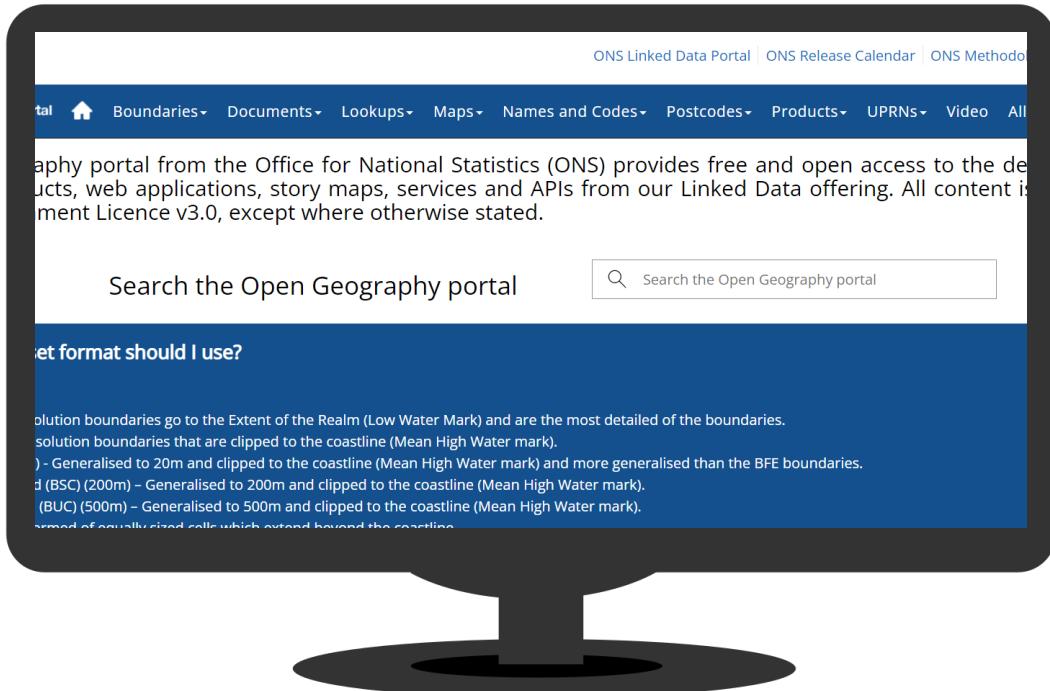
# Combining Resolution and Extent

- Different versions of each boundary set
- For analysis use the most detailed set
- For mapping or visualisation use a more generalised set

	Clipped	Extent of the Realm
Full	BFC	BFE
Generalised	BGC	-
Super Generalised	BSC	-
Ultra Generalised	BUC	-



# Where to get data from



<http://geoportal.statistics.gov.uk/>

<http://statistics.data.gov.uk/>

# Ordnance Survey – Open Data

- Open data and premium data
- Open data can be downloaded here:

<https://osdatahub.os.uk/>

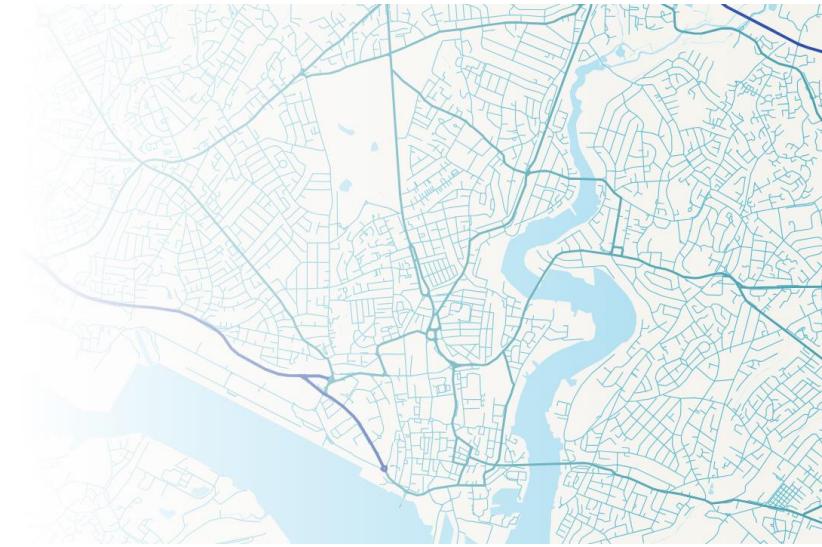
## Welcome to the OS Data Hub

Great Britain's Geospatial Data platform.

Use our APIs and data downloads to integrate OS data in your work and applications.

[Sign up for free](#)

[Explore and access OpenData](#)



Source: Ordnance Survey

# Ordnance Survey – Premium Data

- Public Sector Geospatial Agreement (PSGA)

- Liaison with OS - contact  
[onsgeography@ons.gov.uk](mailto:onsgeography@ons.gov.uk)



OS MasterMap Highways Network - Routing and Asset Management

Great Britain's most complete, detailed and accurate road network dataset.

Networks



OS MasterMap Highways Network - Roads

OS MasterMap® Highways Network is the most complete, detailed and accurate navigable road network dataset for Great Britain. It records the dimensions and accessibility of roads.

Networks



Code-Point with polygons

Code-Point® with polygons shows the notional shape of every postcode unit in Great Britain, and includes major buildings with multiple postcodes.

Addressing & location



OS MasterMap Water Network Layer

OS MasterMap® Water Network Layer offers one of the world's most detailed, heighted water networks, showing the flow and precise course of every river, stream, lake and canal in Great Britain.



AddressBase Premium

Our flagship address database, AddressBase® Premium gives you the most up to date, accurate information about addresses, properties and land areas.



AddressBase Plus

Along with Royal Mail PAF data, AddressBase® Plus gives you up-to-date local authority addresses, multi-occupancy addresses and OS MasterMap TOID references.

Source: Ordnance Survey

# Questions?

Please use Slido #DSGP for questions



# Types of geographic data

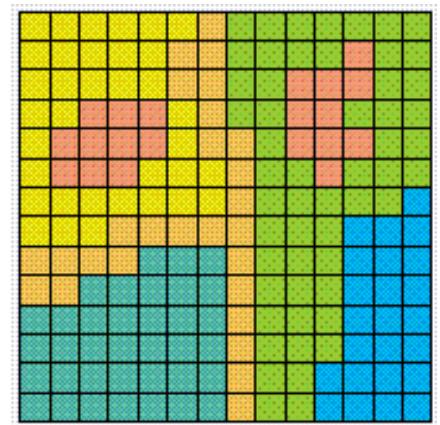
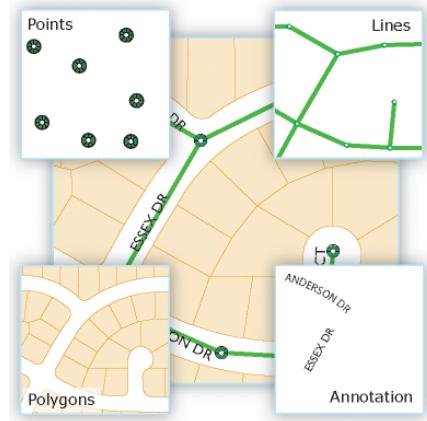
# GIS (Geographic Information System)

- Computer system
- Layer
- Patterns and relationships
- Problem solving and decision-making

# Geographic data types

Two primary types of geospatial data:

- **Vector**
  - Points, lines and polygons
  - Attributes (eg. statistics)
- **Raster**
  - Pixels (also referred to as grid cells).
  - Each pixel has its own value or class.



Images: [arcgis.com](http://arcgis.com)

# Points

- Individual XY coordinates
- Lots of ONS data is point data



Images: [arcgis.com](http://arcgis.com)

# Polygons

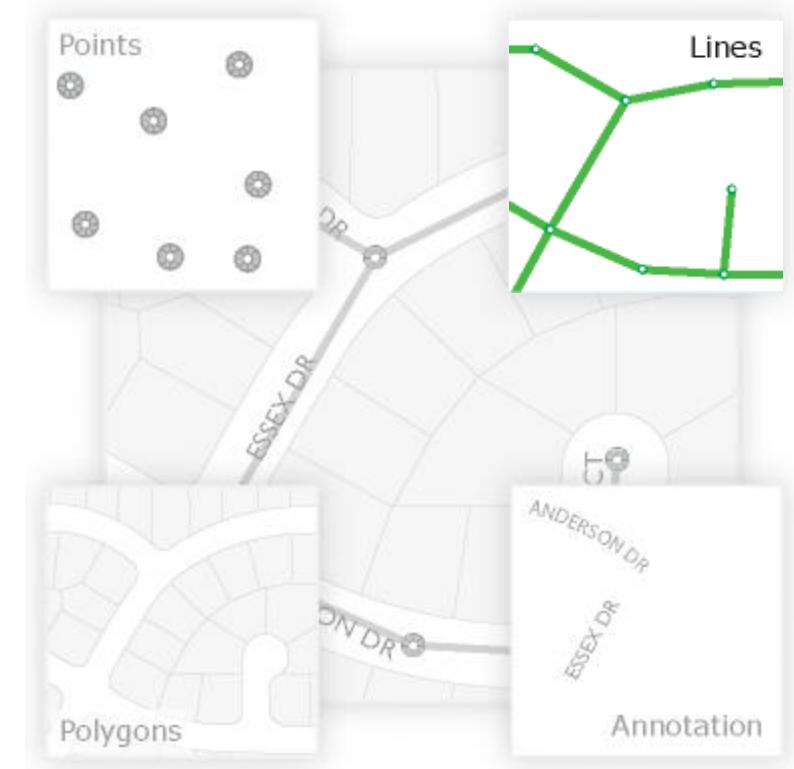
- Polygons represent areas
- Point data can be aggregated to areas



Images: [arcgis.com](http://arcgis.com)

# Lines

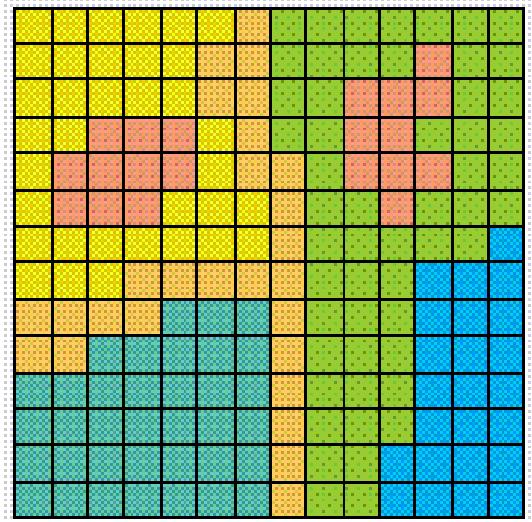
- Roads and rivers
- Road networks
- Very little of the data ONS uses is line data



Images: [arcgis.com](http://arcgis.com)

# Raster

- Grid with a value for each cell
- Remote sensing or Earth Observation
- Environmental data
- Large data sizes



Images: [arcgis.com](http://arcgis.com)

# Tabular data

- Geographical reference
- Join to a geographical point or area

Area of usual residence of mother <sup>1</sup>			Number of live births								
Code	Name	Geography	Total	Male	Female	Within marriage or civil partnership <sup>4</sup>	Outside marriage or civil partnership <sup>4</sup>	Joint registrations same address <sup>4</sup>	Joint registrations different address <sup>4</sup>	Sole registrations <sup>4</sup>	
K02000001	UNITED KINGDOM	Country	731,213	375,529	355,684	377,106	354,107	236,732	80,026	37,349	
K04000001, J99000001	ENGLAND, WALES AND ELSEWHERE <sup>6</sup>	Country	657,076	337,584	319,492	339,267	317,809	214,026	69,719	34,064	
K04000001	ENGLAND AND WALES	Country	656,925	337,506	319,419	339,183	317,742	213,995	69,706	34,041	
E92000001	ENGLAND	Country	625,651	321,513	304,138	326,565	299,086	201,428	65,418	32,240	
E12000001	NORTH EAST	Region	26,684	13,703	12,981	10,539	16,145	9,938	4,416	1,791	
E06000047	County Durham	Unitary Authority	4,868	2,584	2,284	1,785	3,083	1,957	791	335	
E06000005	Darlington	Unitary Authority	1,112	573	539	465	647	383	181	83	
E06000001	Hartlepool	Unitary Authority	1,032	531	501	281	751	441	236	74	
E06000002	Middlesbrough	Unitary Authority	1,868	928	940	687	1,181	655	361	165	
E06000057	Northumberland	Unitary Authority	2,638	1,345	1,293	1,100	1,538	1,024	365	149	
E06000003	Redcar and Cleveland	Unitary Authority	1,366	682	684	372	994	625	249	120	
E06000004	Stockton-on-Tees	Unitary Authority	2,112	1,060	1,052	872	1,240	735	370	135	

# Geographic Data Formats

# Common file formats

## Shapefiles

- Most common file type
- One layer requires 3 files (but can have more):
  - .SHP is the feature geometry
  - .SHX is the shape index position
  - .DBF is the attribute data

## GeoPackages

- GeoPackages - OGC open standard.
- SQLite databases

# Some other file formats

## GeoJSON

- An open standard format
- Represents vector data as text
- Uses longitude/latitude so caution required with projections!

## KML & KMZ

- OGC standard format
- An XML representation of vector data
- Mainly used in Google Earth

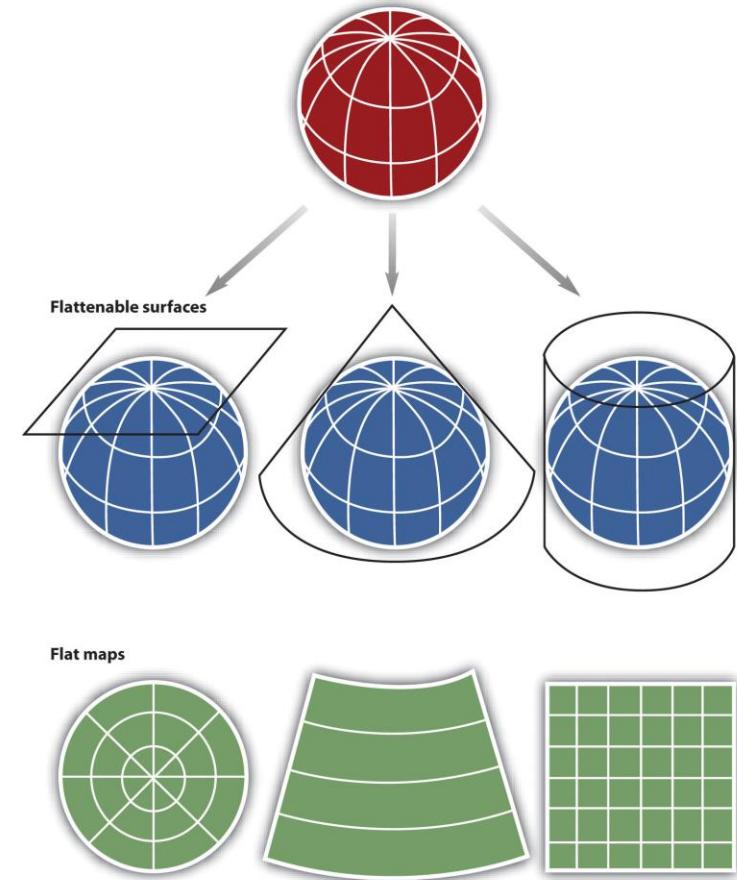
## GeoParquet

- Cloud native vector data
- Can handle very large datasets
- Can run in the browser
- Highly compressed

# Locating Spatial Data

# Coordinate Systems

- Geographical coordinate system - Longitude and latitude
- Projected coordinate system – e.g. British National Grid



From Location/Map Projections And Types - Lessons - Tes Teach

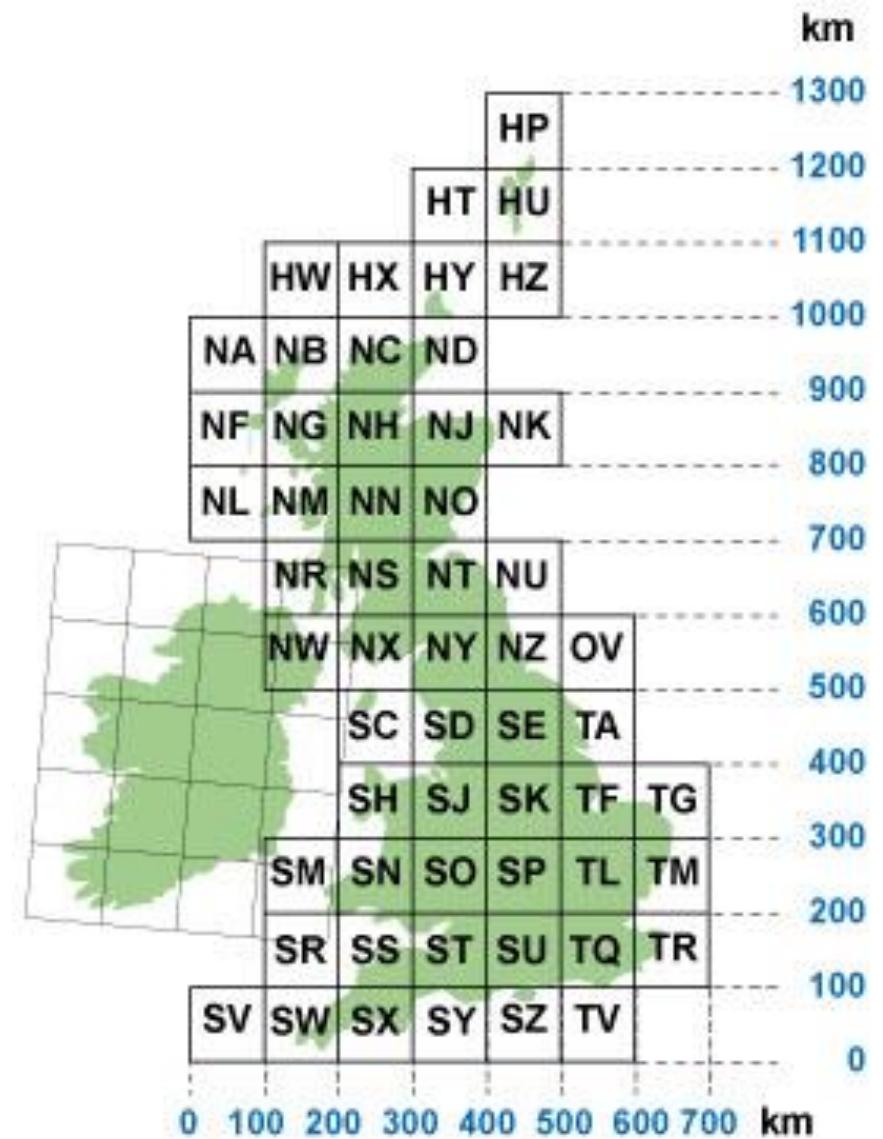
# Coordinate Reference Systems

- Coordinate reference systems (CRS)
- EPSG codes
- Pick carefully!
- All data in the same CRS



# British National Grid

- In Great Britain, British National Grid (BNG) is the coordinate reference system.
- Almost all of the data you will deal with will be in this projection.
- It is maintained by Ordnance Survey



# Common CRS

## In the UK:

- British National Grid (27700)
  - Covers Great Britain and NI
- Irish Transverse Mercator (2157)
  - Covers Ireland and NI
  - There are also older Irish CRSs in use so take care!

## In Europe:

- ETRS89 (4936)
  - Used by Eurostat

## Globally:

- WGS84 (4326) – Mercator
  - The standard world map
- WGS84 (3857) – Web-Mercator
  - Specialised for web mapping



# Questions?

Please use Slido #DSGP for questions



# Break

# Geospatial Software and Tools

# Open Source Geospatial Tools



# Licenced Software



ArcGIS Pro

**FME** :  
by Safe Software

# Spatial Techniques

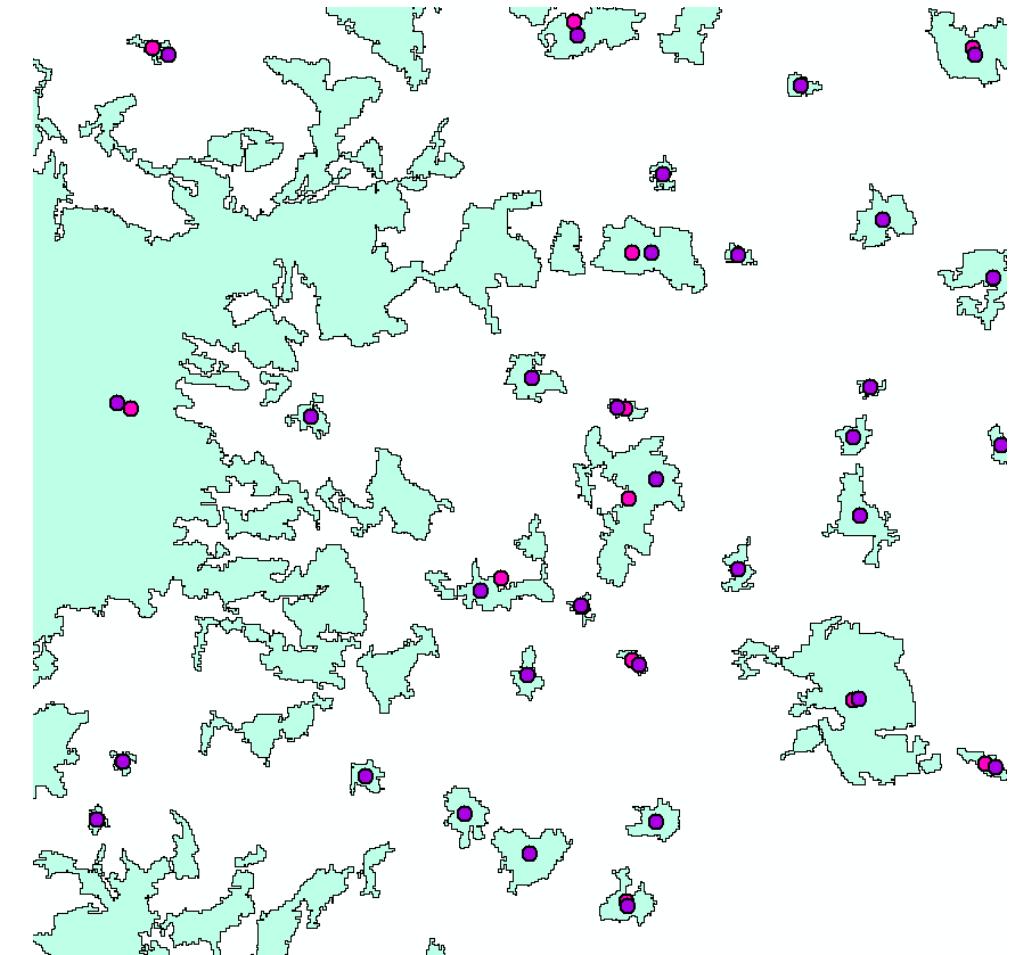
# Joining data to boundaries

- Use GSS codes
- Check results

	A	B	E
	LSOA code (2011)	LSOA name (2011)	Index of Multiple Deprivation (IMD) Score
1			
2	E01000001	City of London 001A	6.208
3	E01000002	City of London 001B	5.143
4	E01000003	City of London 001C	19.402
5	E01000005	City of London 001E	28.652
6	E01000006	Barking and Dagenham 016A	19.837
7	E01000007	Barking and Dagenham 015A	31.576
8	E01000008	Barking and Dagenham 015B	37.596
9	E01000009	Barking and Dagenham 016B	32.844
10	E01000010	Barking and Dagenham 015C	32.437
11	E01000011	Barking and Dagenham 016C	30.679
12	E01000012	Barking and Dagenham 015D	34.714
13	E01000013	Barking and Dagenham 013A	38.997
14	E01000014	Barking and Dagenham 013B	35.67
15	E01000015	Barking and Dagenham 009A	33.632

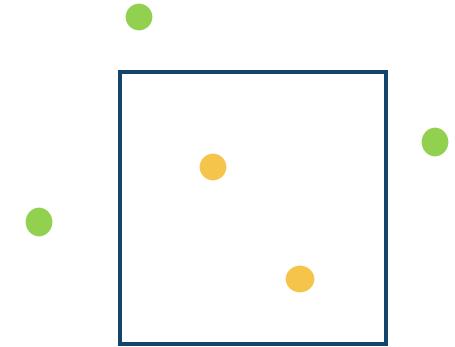
# Point in polygon

- Join using location
- Appends the polygon's attributes to the point dataset
- Aggregate points to statistical geographies

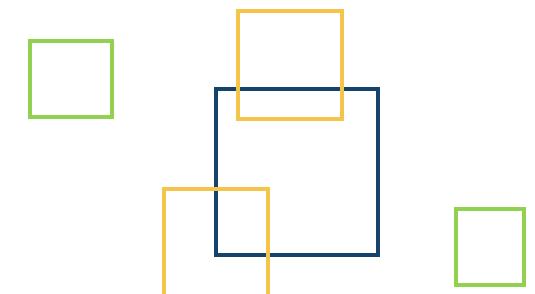


# Select by location

- Select features based on their location relative to other features.
- E.g. Homes were affected by a recent flood



Select points within the blue polygon

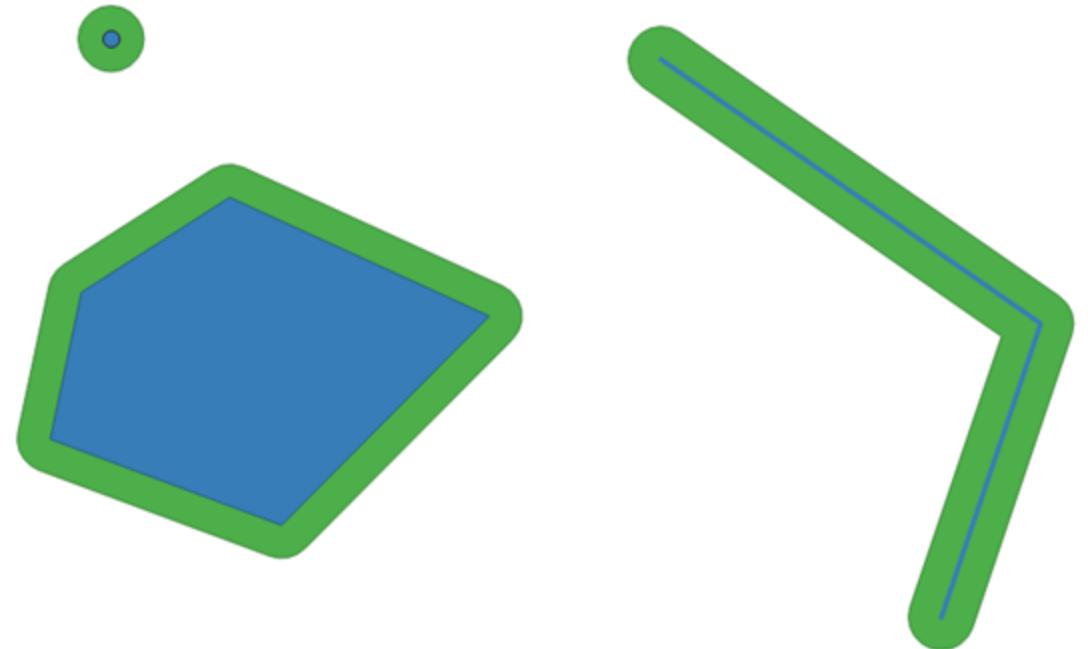


Select polygons which intersect the blue polygon



# Buffer

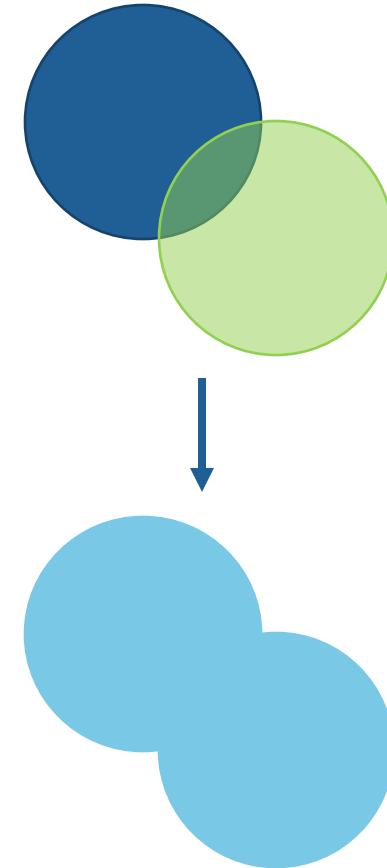
- Distance outward from an object.
- Vector data - points, lines and polygons.



# Dissolve

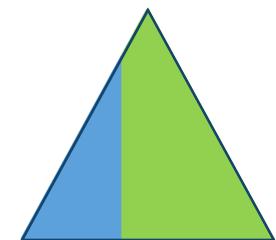
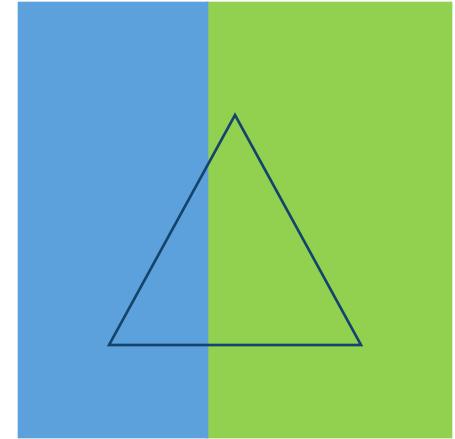
Overlap or share a common boundary or common attribute

Merge together



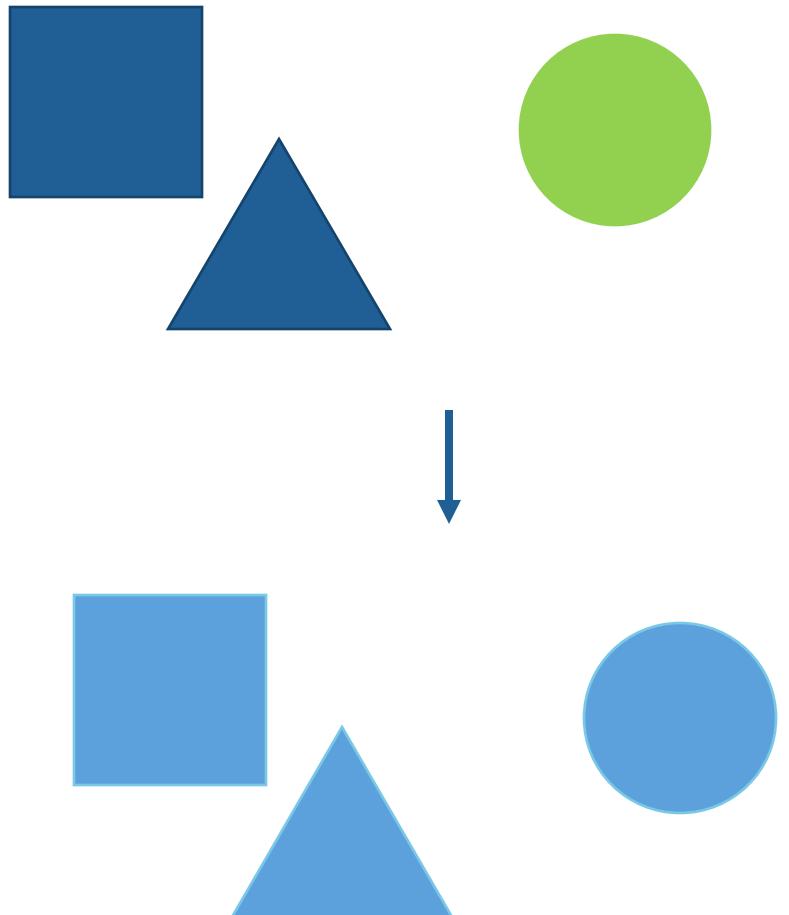
# Clipping

- Clip one feature extent to another
- Overlapping portion is retained



# Merge

Combines two or more layers into a single layer.



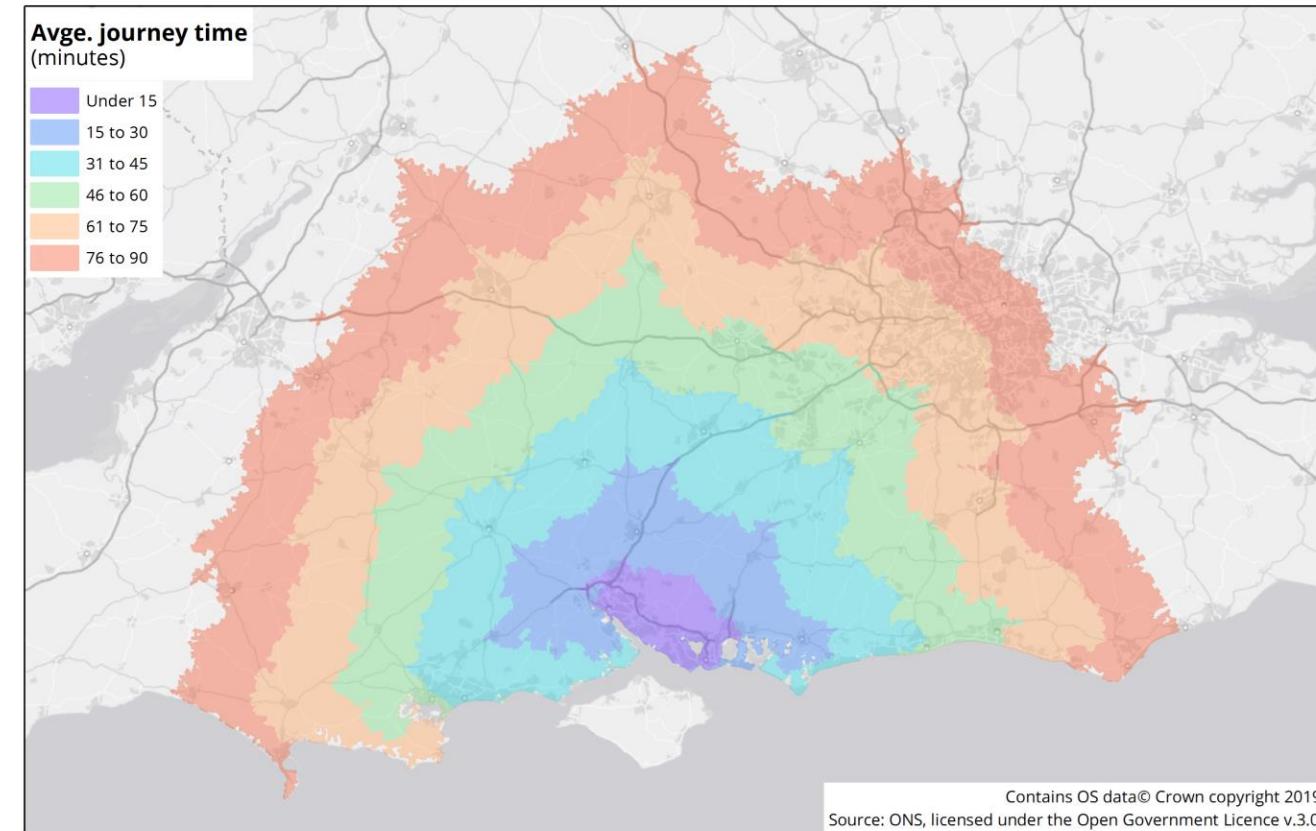
# Advanced spatial techniques to be aware of

# Networks, drivetime and zoning

- Complex analysis of networks
- Fastest route
- Zoning techniques

Average journey times to the ONS Titchfield Office

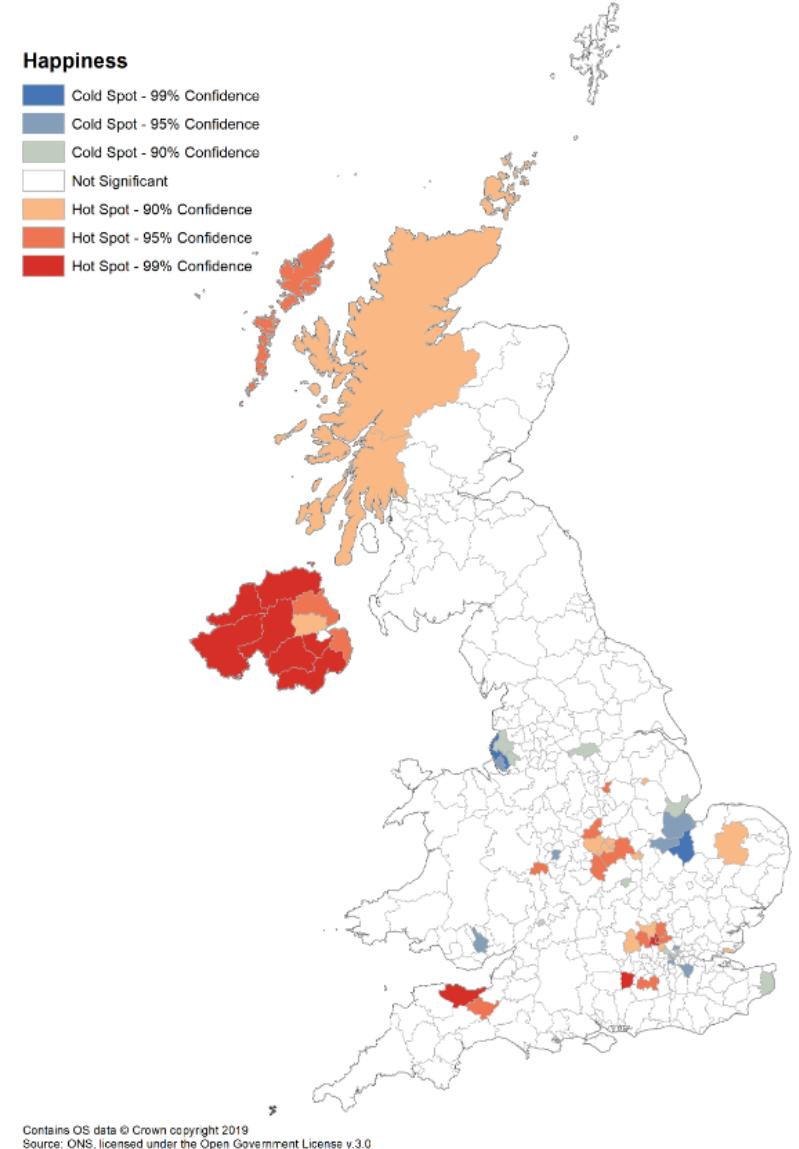
Based on a staff survey and OS Integrated Transport Network (ITN)



# Clusters and hotspots

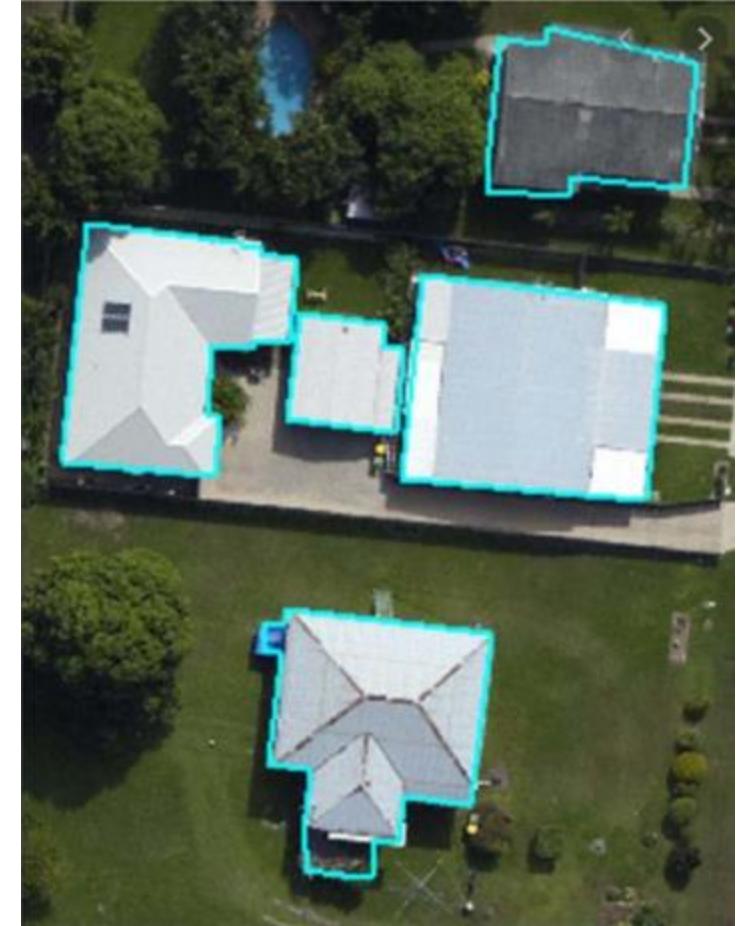
- Expose patterns
- Classify or group objects
- Add statistical rigour

Hotspot Analysis of Happiness across the UK  
Cluster and outlier analysis for local authority district, 2018



# Earth Observation and Machine Learning

- Satellite data
- Machine learning (AI)



Source: ESRI

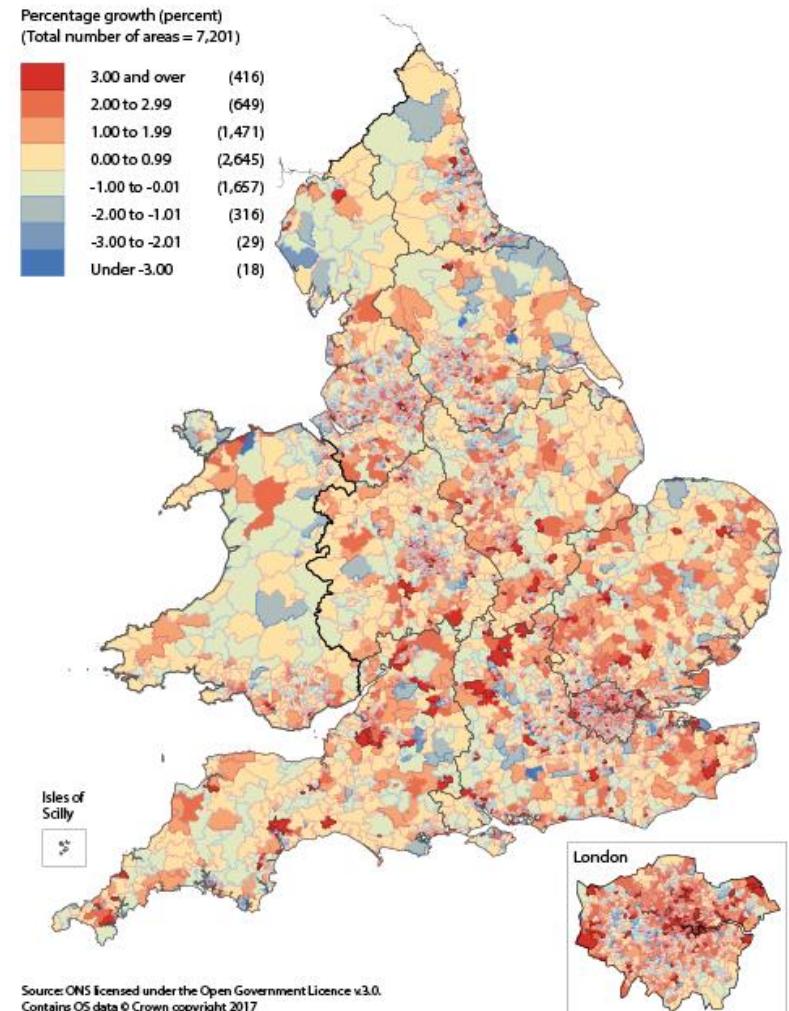
# Mapping your data



# Mapping your data (1)

- Mapping and visualising your data - powerful.
- Interrogate source data
- Investigate possible relationships and patterns
- Present results
- QA.

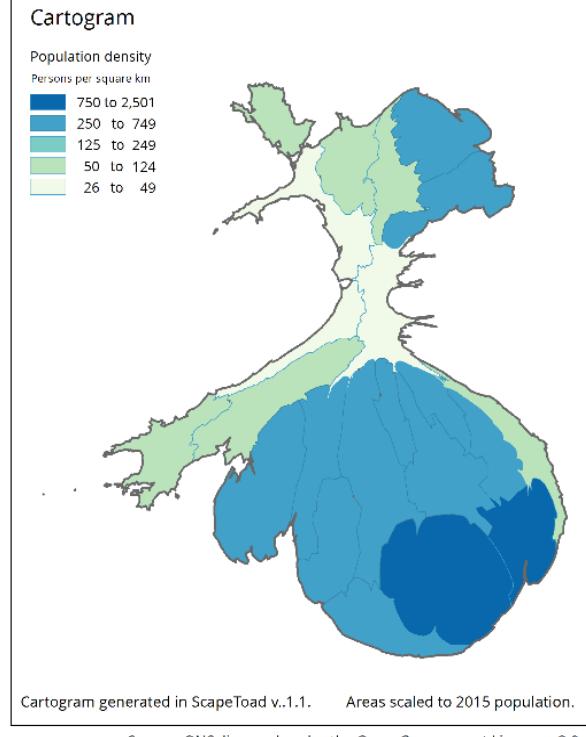
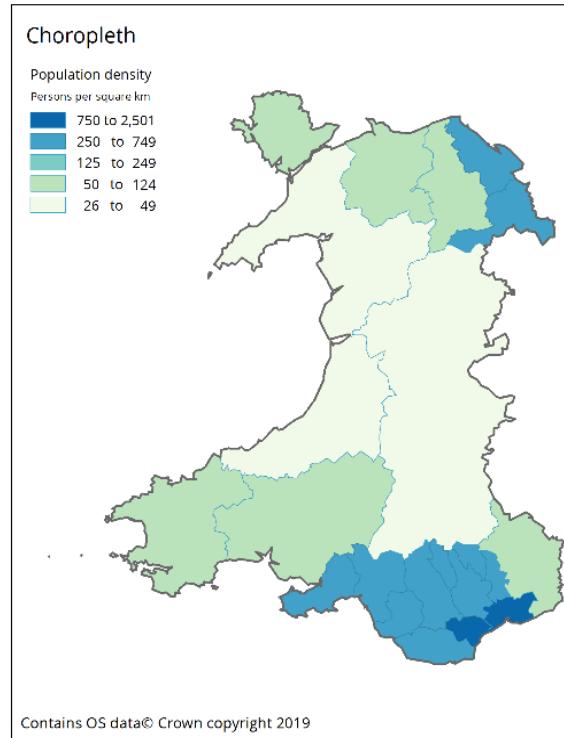
Percentage change in population, by MSOA, England and Wales, mid-2015 to mid-2016



# Mapping your Data (2)

- Range of map types
- Customise ranges, colours and symbols
- Basic rules

**Population density, Wales (2015)**  
Choropleth and cartogram by local (unitary) authorities



# Basic rules

- Understand your data
- Classification methods
- Limit class breaks

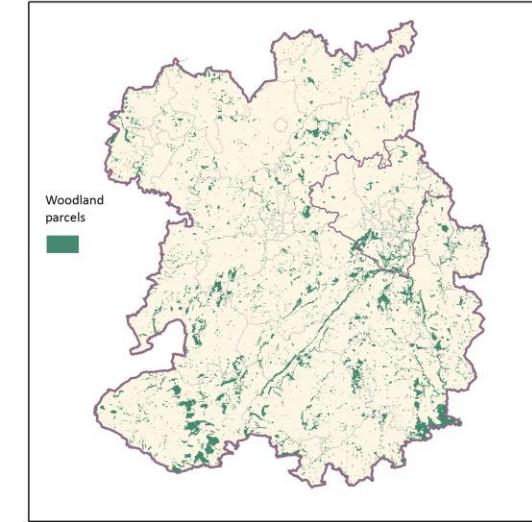
# Normalise your data

- When mapping counts to areas you need to normalise your counts

Woodland areas in Shropshire and Wrekin, 2018

Woodland instances

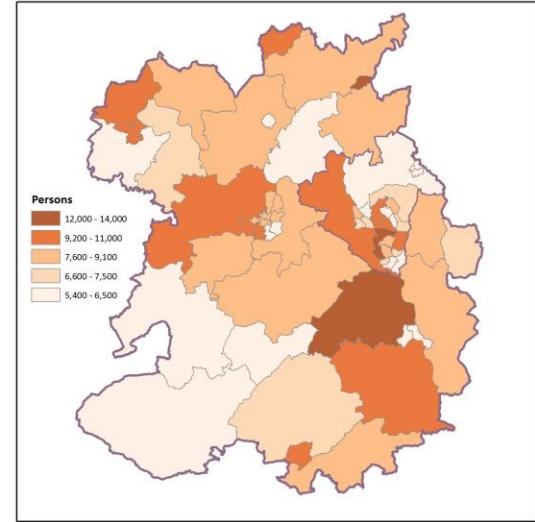
2011 MSOA boundaries



Mid-2018 population estimates, Shropshire

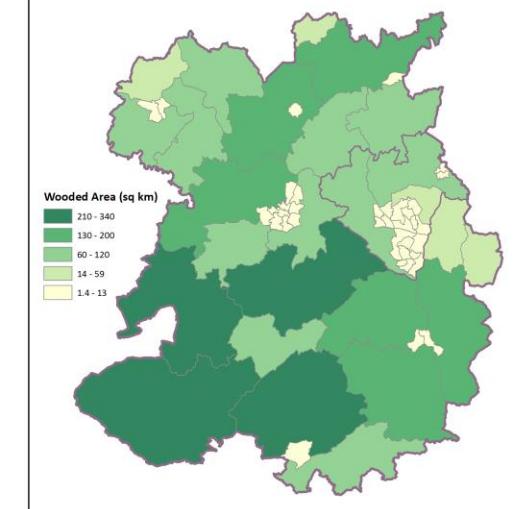
Persons per MSOA

2011 MSOA boundaries



Total woodland per MSOA

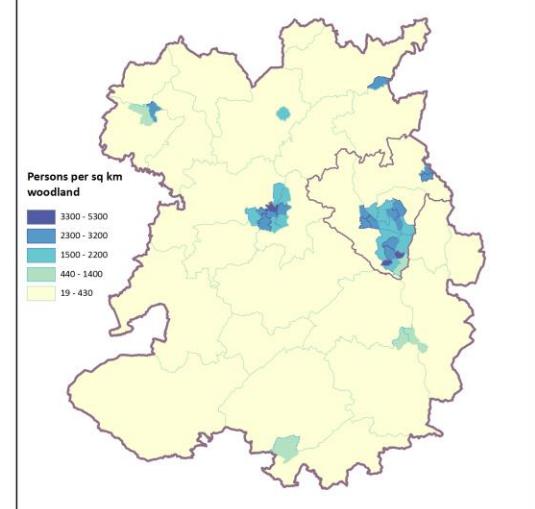
2011 MSOA boundaries



Contains OS data © Crown copyright (2020)

Persons per sq. km woodland

2011 MSOA boundaries



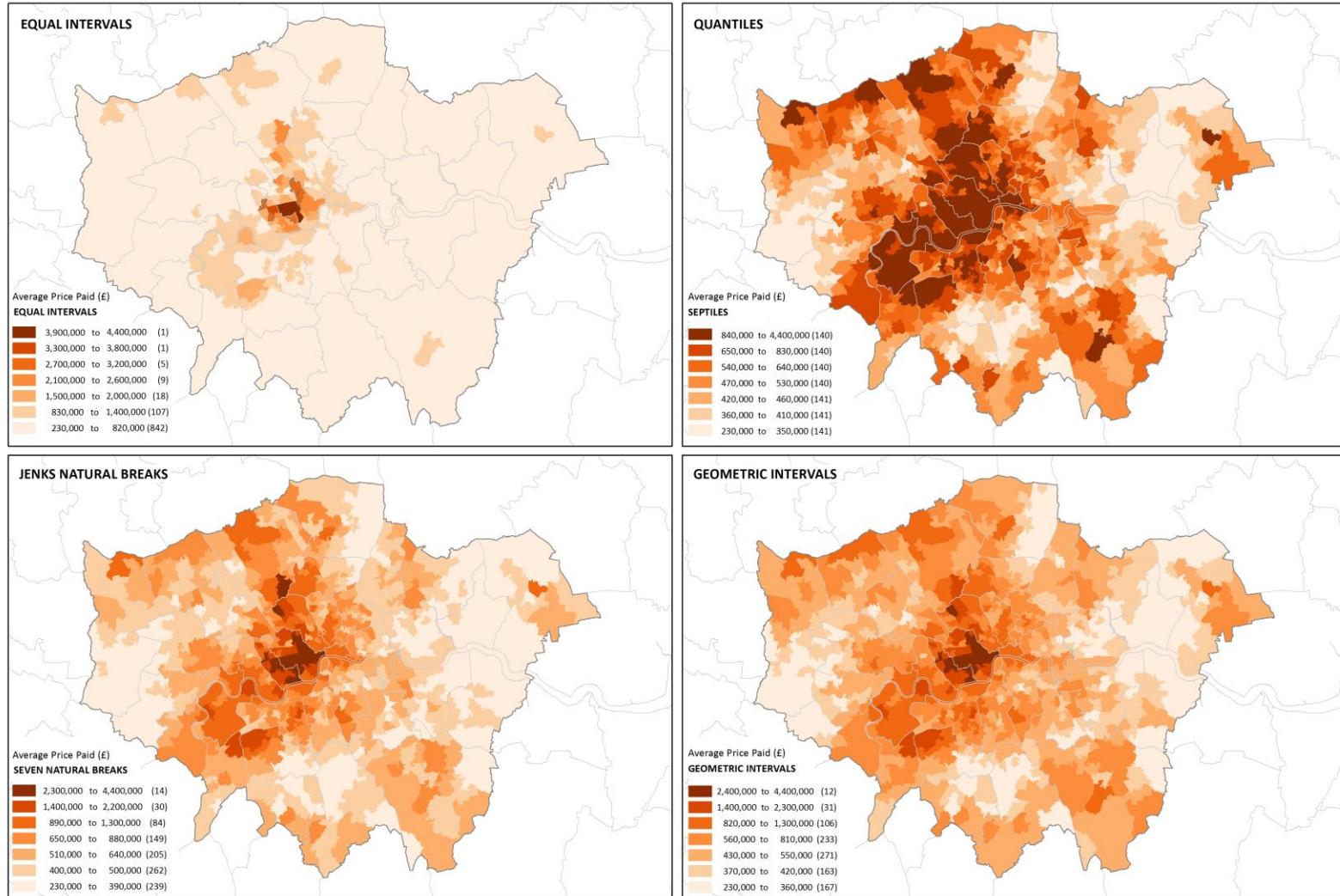
Contains OS data © Crown copyright (2020)

# Selecting breaks

- Pick carefully!
- Different selections highlight trends differently

## Residential property sales, 2017

Average price per 2011 Middle Layer Super Output Area (MSOA)



Source: HM Land Registry, licensed under the Open Government Licence v.3.0. Contains OS data © Crown copyright and database right [2018]

Contains National Statistics data © Crown copyright and database right [2018]



# Questions?

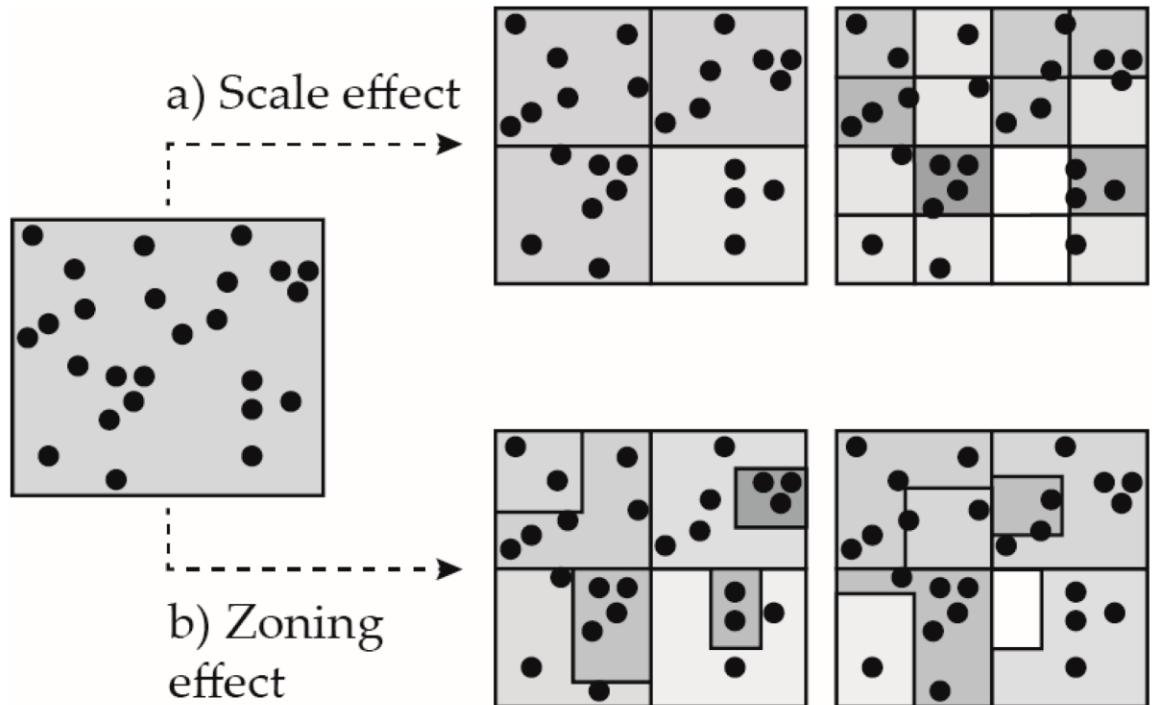
Please use Slido #DSGP for questions



# Geographic fallacies

# Modifiable Areal Unit Problem

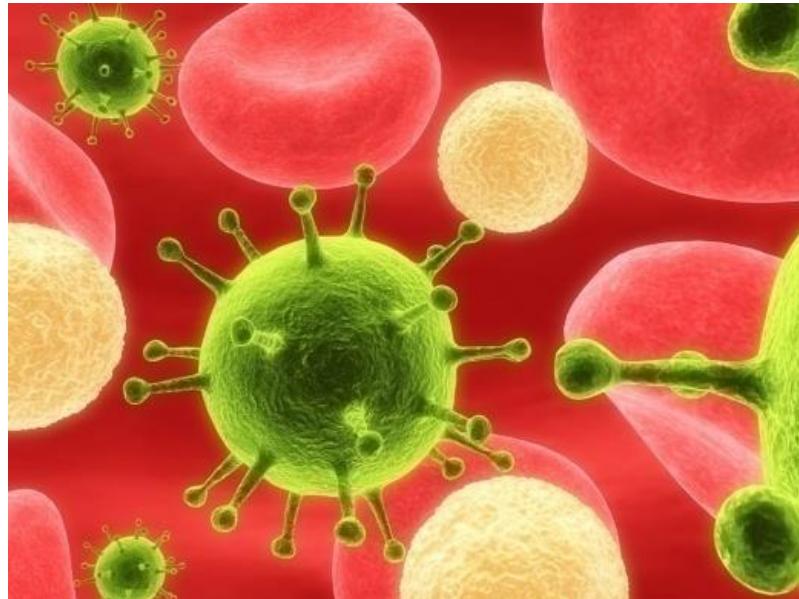
- Same data different results
- Can be a source of statistical bias



[http://www.mdpi.com/safety/safety-02-00017/article\\_deploy/html/images/safety-02-00017-g002.png](http://www.mdpi.com/safety/safety-02-00017/article_deploy/html/images/safety-02-00017-g002.png)

# Locational fallacy

- Spatial characterisation of the data
- ‘Influences’ understanding



Diseases by home address? Why not work address?



Crime data – Where does fraud happen?

# Ecological Fallacy

- Incorrect inferences about individual's from the group value
- Misunderstanding relationships and correlations
- Aggregating data may conceal real variations

People in the combined area earn about £30,000

Combined Area

Average income is  
£30,000

Area 1

Average  
income is  
£10,000

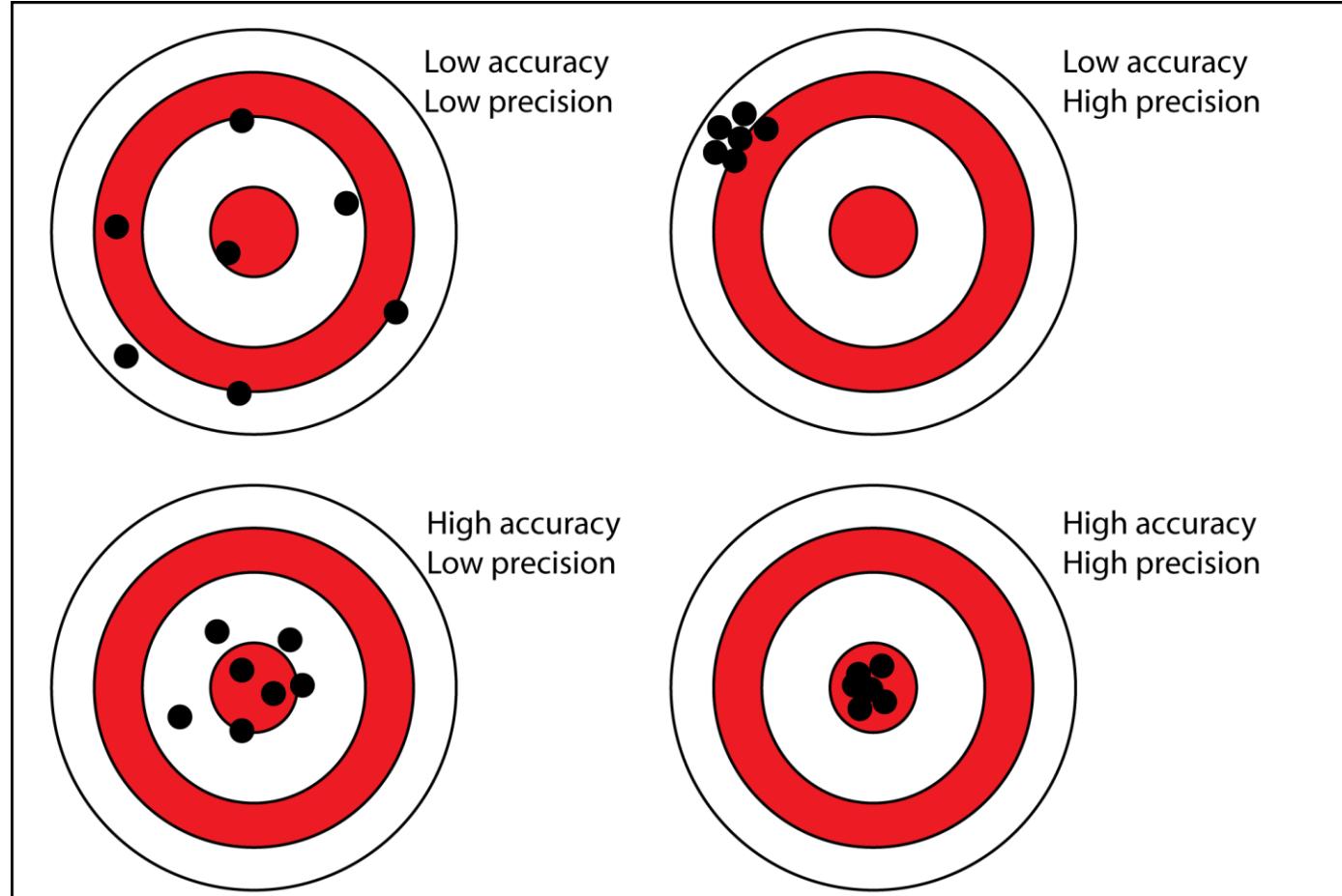
Area 2

Average  
income is  
£50,000

Aggregated data

Real data

# Accuracy vs Precision

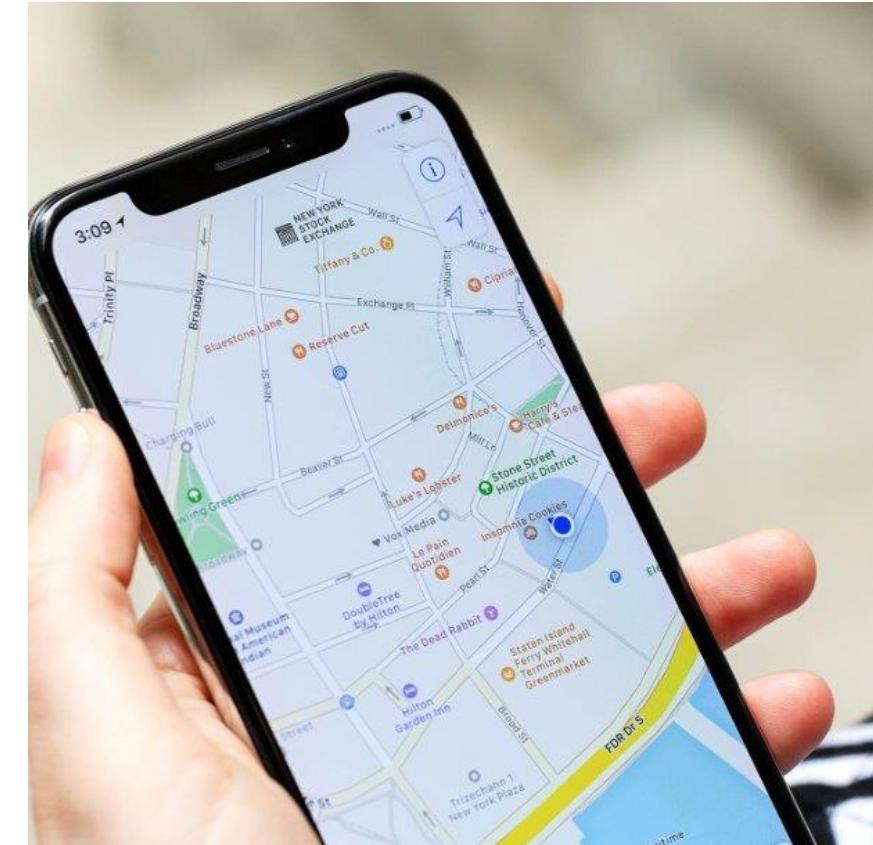


[https://kaiserscience.files.wordpress.com/2015/09/precision\\_accuracy.png](https://kaiserscience.files.wordpress.com/2015/09/precision_accuracy.png)

# Geospatial ethics

# Geospatial data and ethics

- Ethical considerations
- Data from smart devices like phones, car GPS or CCTV
- Huge benefits - but can bring risks.



# Fitness tracking app Strava gives away location of secret US army bases

Data about exercise routes shared online by soldiers can be used to pinpoint overseas facilities

- Latest: Strava suggests military users 'opt out' of heatmap as row deepens



A military base in Helmand Province, Afghanistan with route taken by joggers highlighted by Strava. Photograph: Strava Heatmap

Source: The Guardian

## All Was Quiet at the Birmingham Weather Office. Until a Trump Forecast Brought a Storm.

Tornadoes are the most challenging tempests that the National Weather Service workers in Alabama typically handle. But all that changed with a faulty Hurricane Dorian tweet by the president.



Source: New York Times



# Why ethics matters for research

- Moral and social responsibility
- Public trust and accountability
- Protect participants and researchers

# Ethics are hard!

- Context specific
- Principles
- Competing obligations
- An inconvenient truth!



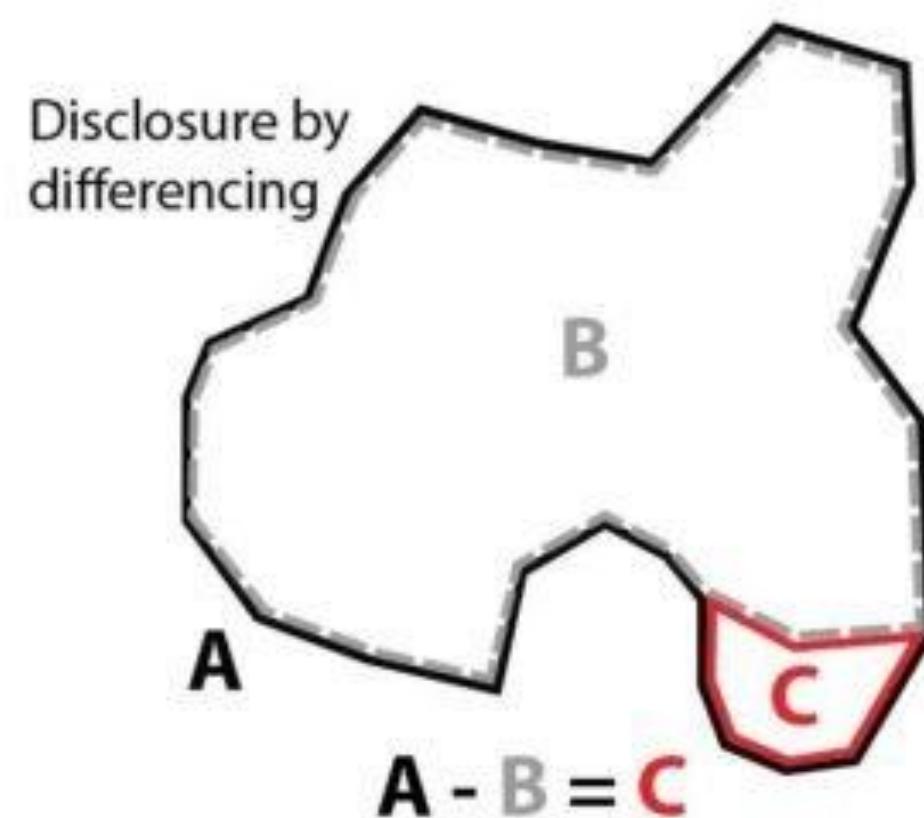
Source: PowerPoint stock images

# Four Key Responsibilities

- Confidentiality
- Inclusivity
- Bias and discrimination
- Presentation

# Confidentiality and Disclosure Risk (1)

- Points or small areas
- Disclosure by differencing
- Output geography



# Confidentiality and Disclosure Risk (2)

- Data minimisation
  - De-identify data
  - Use synthetic data
  - Randomised response approaches
- Provenance and Limitations

# Confidentiality and Disclosure Risk (3)

- Transparency
- Consent
  - What is being collected
  - When
  - Why
  - What is being retained

# Inclusivity

- Exclude individuals/groups
- Unrepresentative
- Engagement
- Additional data sources
- Documentation



Source: PowerPoint stock images

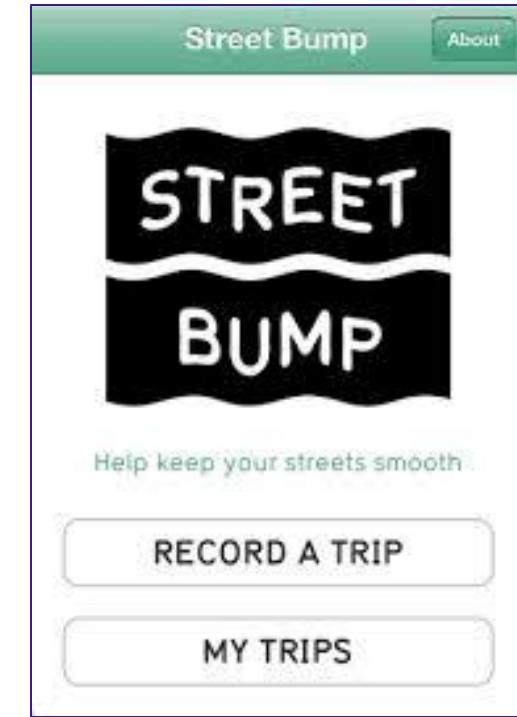
# Bias and Discrimination

- Inherent bias
- Disadvantage an area/community



# Example: Boston Street Bump

- Pothole detection
- Timely reporting
- Biased



From: <https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0>

# Presentation of Geospatial Data

- Fair view of reality
- Storytelling
- Honesty



[buzzfeed.com/lanesainty/viral-australian-bushfire-maps-confusing-false-information](https://buzzfeed.com/lanesainty/viral-australian-bushfire-maps-confusing-false-information)

# Guidance

The screenshot shows the UK Statistics Authority website. At the top, there are links for 'UK Statistical System', 'Code of Practice', and 'Office for Statistics Regulation'. The main header features the 'UK Statistics Authority' logo with the tagline 'Statistics for the Public Good'. A search bar and social media icons for Twitter, LinkedIn, and RSS are also present. Below the header, a navigation menu includes 'Home', 'About us', 'News', 'Boards and committees', 'Correspondence', 'Consultations and reviews', and 'Research guidance'. The current page is 'Ethical considerations in the use of geospatial data for research and statistics', which is under the 'Research guidance' category. The breadcrumb navigation shows 'UK Statistics Authority > Publications > Ethical considerations in the use of geospatial data for research and statistics'.

## Table of contents

- Introduction
- [Geospatial data and ethics](#)
- [Confidentiality and disclosure risk](#)
- [Ensuring inclusivity](#)
- [Bias and discrimination](#)
- [Presentation of geospatial data](#)
- [ETHICS CHECKLIST](#)
- [Where can I find out more?](#)
- [Feedback on this guidance](#)

## Ethical considerations in the use of geospatial data for research and statistics

Published: 18 May 2021

Last updated: 6 September 2021

### Introduction

This guidance document explores ethical considerations in the use of geospatial data for research, analysis and statistics. It has been developed by the UK Statistics Authority's [Centre for Applied Data Ethics](#) in partnership with geospatial colleagues.

It is divided into 8 main parts, providing an initial introduction to geospatial data and ethics, before moving on to consider case examples and potential mitigations related to four key areas, as well as links to further resources.

An ethics checklist is also provided for researchers and analysts using geospatial data, which summarises the main points covered in this guidance.

**Don't have time to read the whole document? Skip to our geospatial [ETHICS CHECKLIST](#)**

**Click here for Guidance!**



Office for National Statistics



**data = can ≠ should**

“(your scientists) were so pre-occupied with whether or not they could, they didn’t stop to think if they *should*”



- Dr Ian Malcolm, Jurassic Park

# Ethics case studies

- 1: Mapping Muslim Neighbourhoods
- 2: Stag and Deer Migration Routes
- 3: Proprietary Software in an Emergency
- 4: Tracking Mobile Phones in Mobility Research
- 5: Collateral Damage Mapping
- 6: Low-level Radioactive Waste Siting Map
- 7: Environmental Justice Web Map

# Resources (1)

- Self-service training - <https://onsgeo.github.io/geospatial-training/> - includes
  - Awareness of Geography and Statistics
  - Practical Geography for Statistics
  - Introduction to QGIS
  - Introduction to GIS in R
  - More GIS in R
  - Introduction to GIS in Python
  - More GIS in Python
  - How to make a good map

# Resources (2)

- GSS Geography Policy - <https://gss.civilservice.gov.uk/wp-content/uploads/2018/03/GSS-Geography-Policy-6.pdf>
- UKSA Ethics Guidance - [Ethical considerations in the use of geospatial data for research and statistics – UK Statistics Authority](#)
- R online resource - <https://r.geocompx.org/>
- Python online resource - <https://py.geocompx.org/>

# Resources (3)

- Open Geography portal - <http://geoportal.statistics.gov.uk/>
- Linked Data portal - <http://statistics.data.gov.uk/>
- Geography enquiries – [ONS.Geography@ons.gov.uk](mailto:ONS.Geography@ons.gov.uk)

# Questions?

Please use Slido #DSGP for questions

