Zeke Hunter-Green

Investigating British royals' wealth using geospatial tools

The Guardian

Links

Slides: https://bit.ly/43iNb9y
(don’t look at solutions!!)

DBeaver: https://dbeaver.io/download/

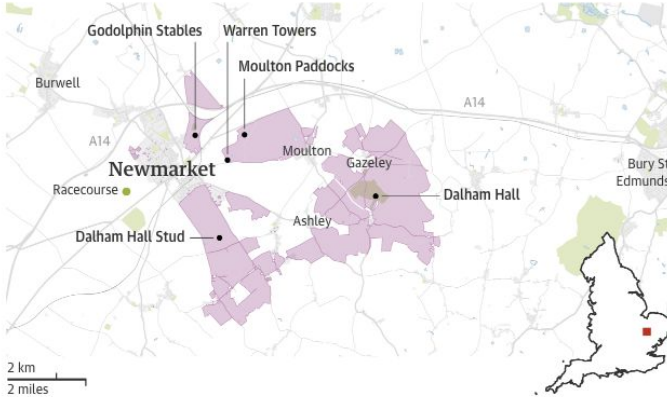
Guardian Investigations & Reporting projects involving geospatial data

Revealed: the huge British property empire of Sheikh Mohammed

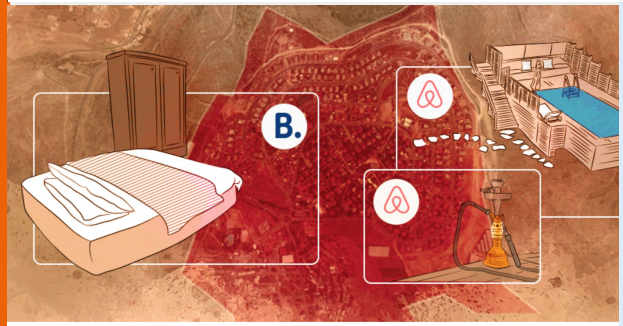
Holdings of more than 40,000 hectares in London, Scotland and Newmarket make Dubai ruler one of UK's biggest landowners

by David Conn, Harry Davies and Sam Cutler

Sheikh Mohammed's Newmarket properties include historic studs, stables and training paddocks



Guardian graphic. Source: Land Registry

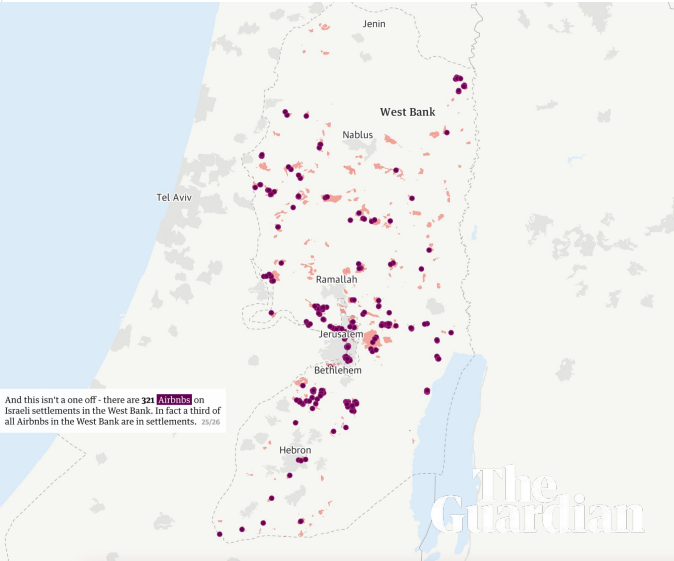


West Bank

Seized, settled, let: how Airbnb and Booking.com help Israelis make money from stolen Palestinian land

As Israel deploys tanks in the West Bank for the first time in 20 years, we reveal how two of the world's biggest travel companies are helping settlers commercialise stolen land

Zeke Hunter-Green, Carmen Aguilar García, Anna Leach, Mark Townsend, Pamela Duncan *and* Prina Shah



Contents

Basics of Geodesy

science of modelling the Earth and positions on its surface

The Guardian

Coordinate Systems

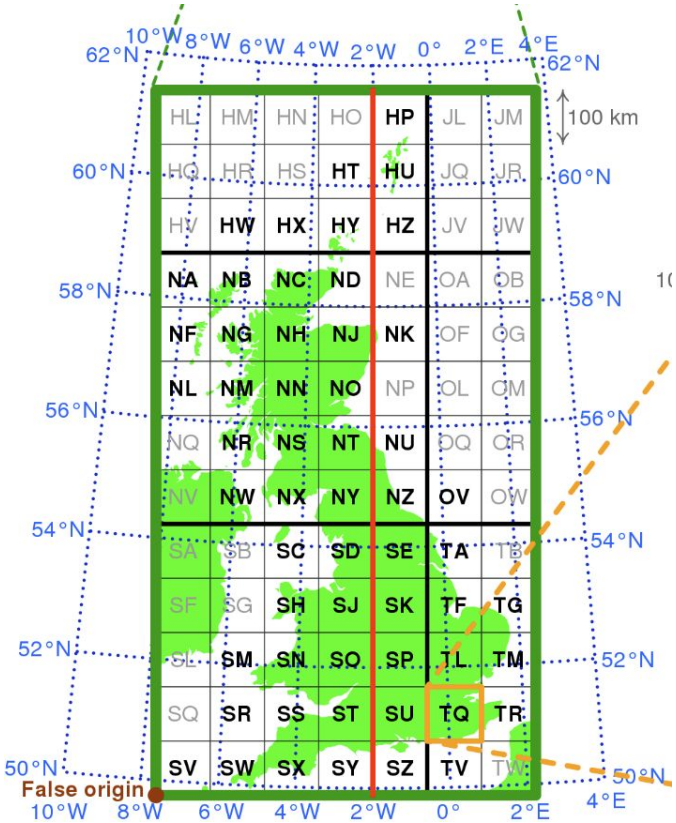
Geographic coordinate system

- describes position in terms of angles
  - North / South of equator (latitude)
  - East / West of Prime Meridian (longitude)
- simplest, oldest, most widely used

Coordinate Systems

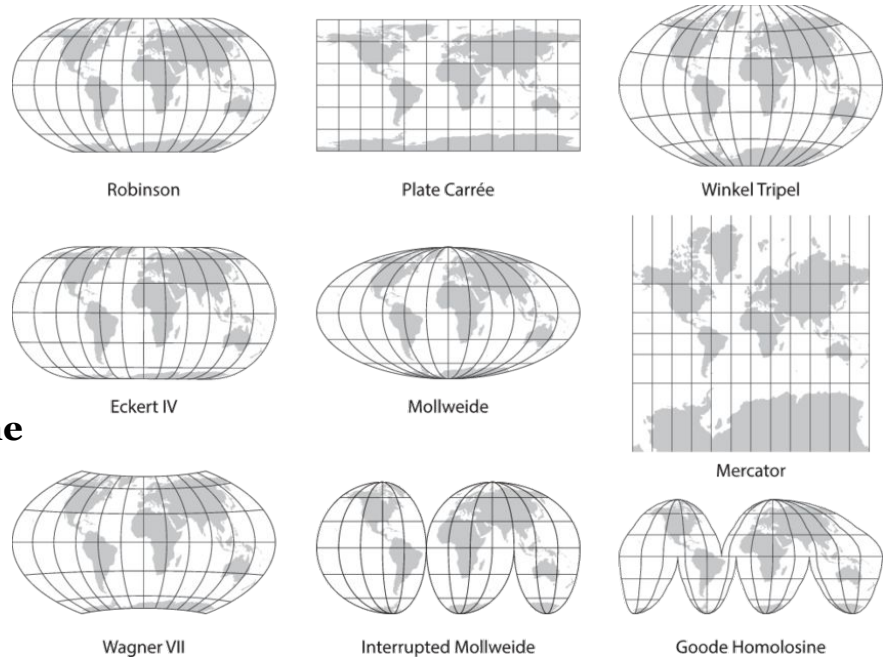
Projected coordinate system

- represents locations on Earth using Cartesian coordinates (x, y)
- used to represent Earth on flat surfaces e.g. computer screens, maps
- projected coordinate systems defined by
  - a geodetic datum
  - origin point
  - unit of measure
  - a **projection**

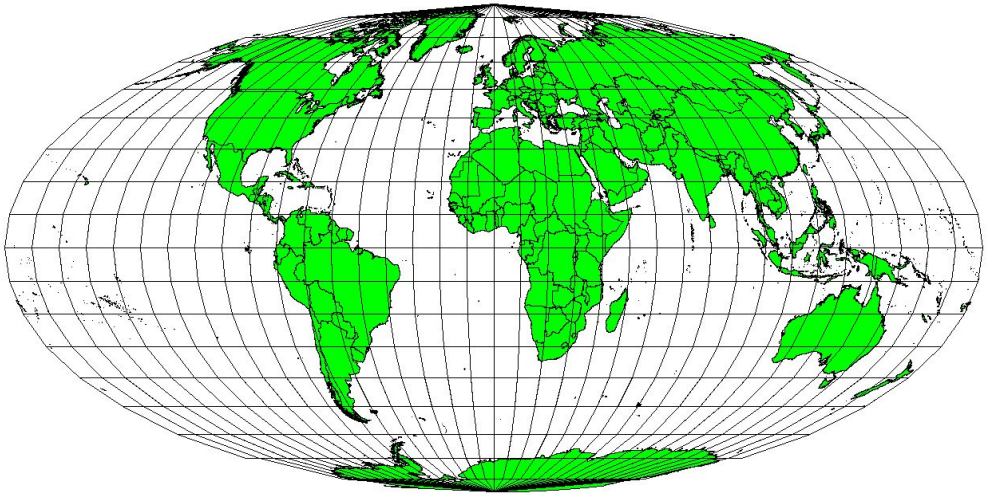


Projections

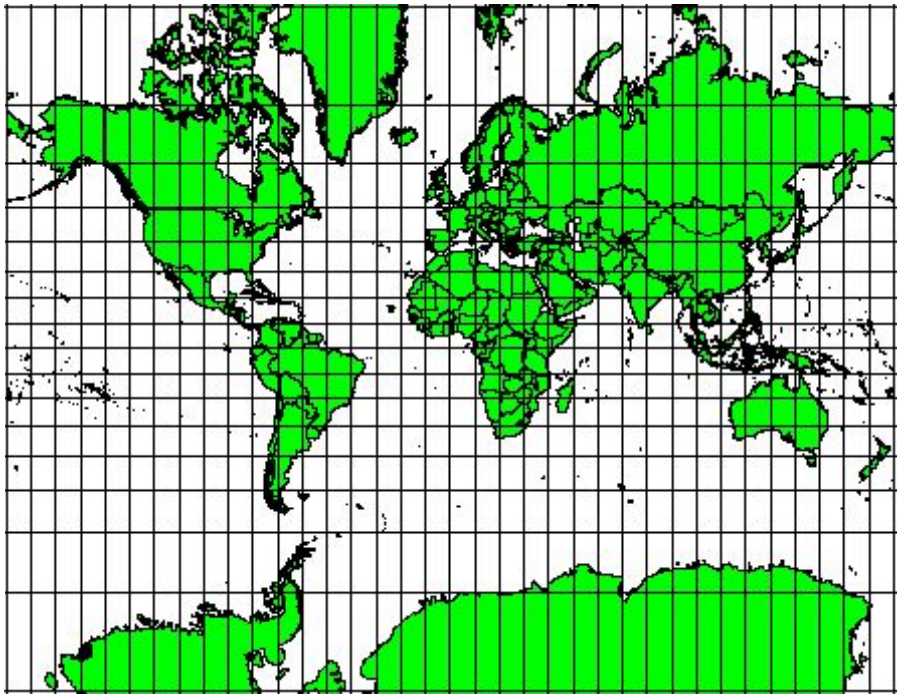
- formula to compute projected coordinates from lat / long coordinates
- never accurate representations of 3D earth. They show distortions in
  - area
  - distance
  - angular conformity
- One of these properties can be preserved at the expense of the others
  - Equal area projections
  - Equidistant projections
  - Projections with angular conformity
  - compromises
- **Important to choose a projection that preserves the property you're interested in!**



Projections

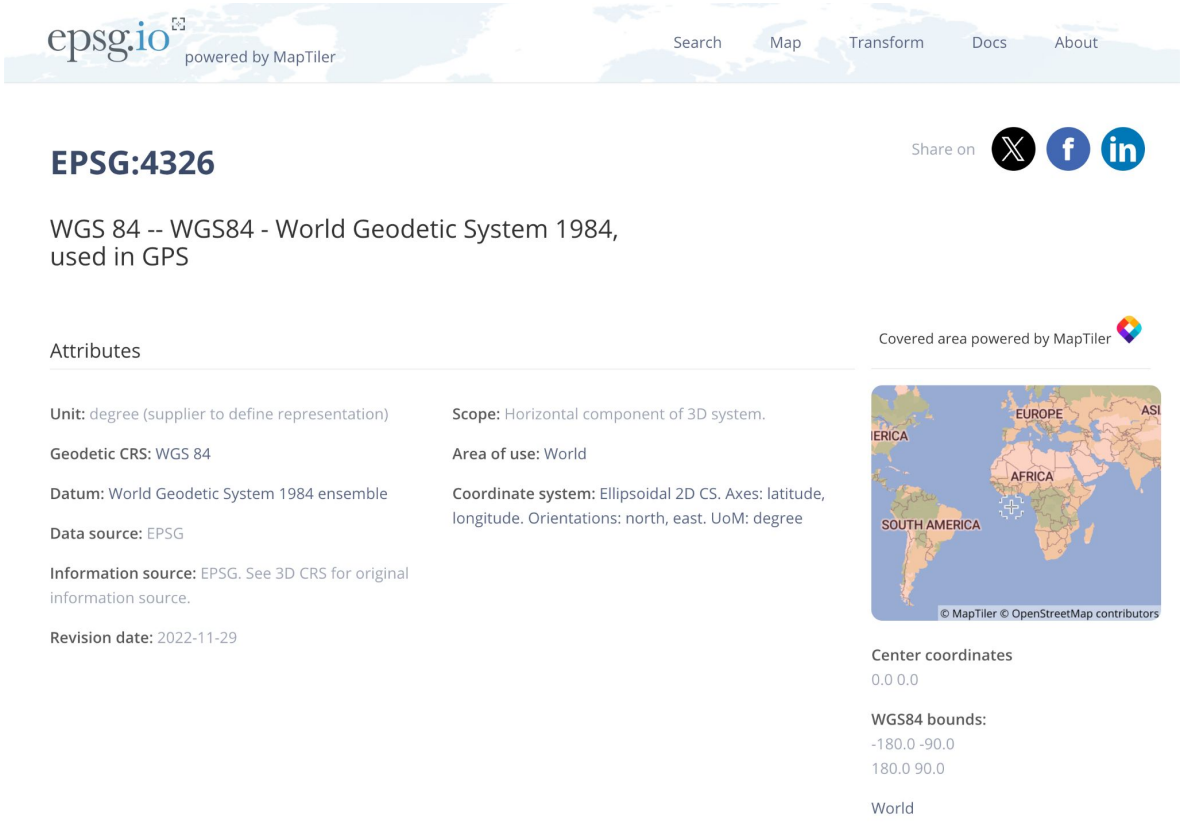


Mollweide Equal Area Cylindrical projection



Mercator projection
(angular conformity)

Coordinate Reference Systems (CRS)



Coordinate Reference Systems (CRS)

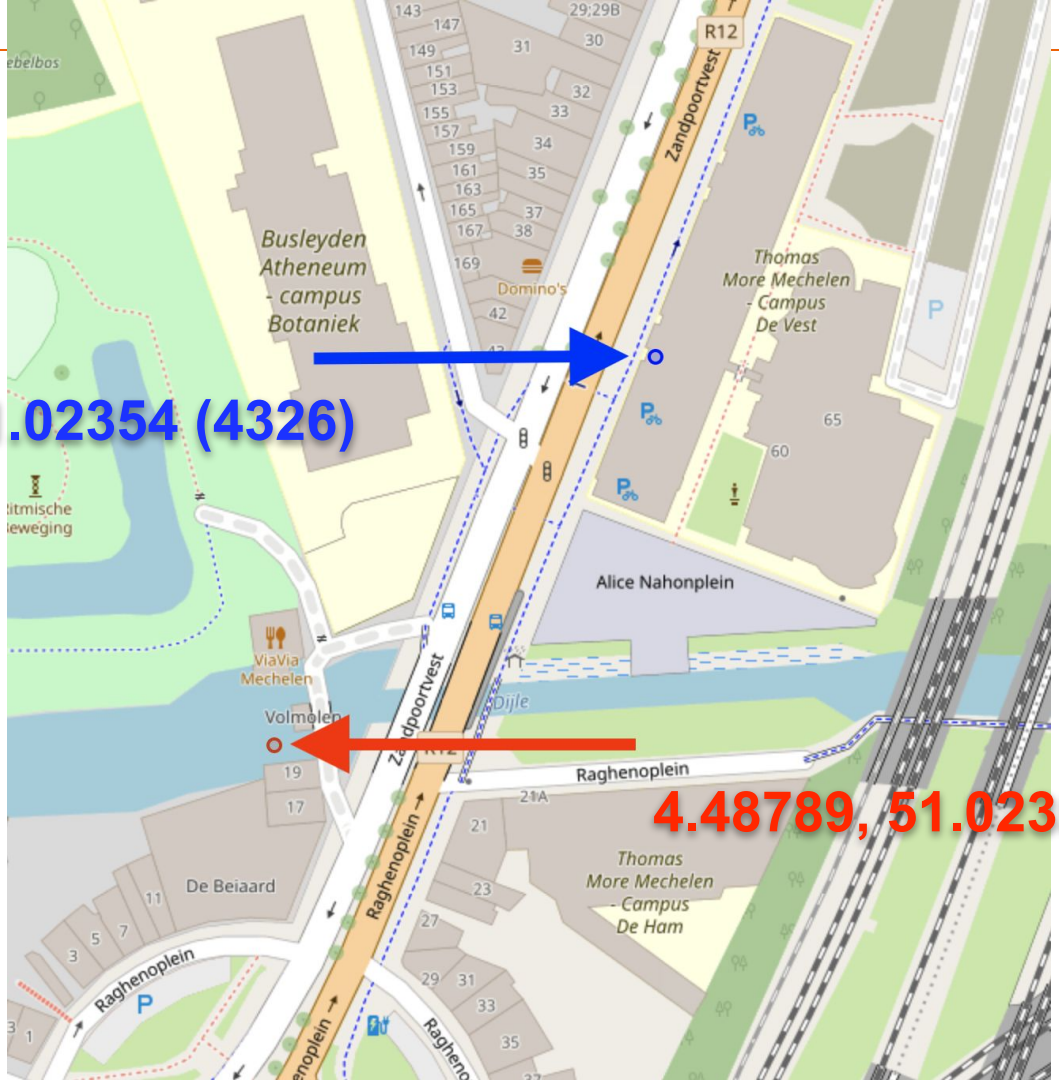
EPSG Code	Name	Area of use	Coordinate System	Projection	Origin	Unit
4326	World Geodetic System 1984 (WGS84)	world used in GPS, Google maps	geographic (lat, lon)	N/A	equator/prime meridian	degree of arc
27700	British National Grid	Britain, Northern Ireland	cartesian (x,y)	Transverse Mercator (**not equal area**)	49.75, -9.0 (WGS 84)	meters
3035	European Terrestrial Reference System 1989 ensemble	Europe	cartesian (x,y)	Lambert azimuthal **equal-area** projection	32.88, -16.1 (WGS 84)	meters

Coordinate Reference Systems (CRS)

EPSG Code	Name	Area of use	Coordinate System	Projection	Origin	Unit
4326	World Geodetic System 1984 (WGS84)	world used in GPS, Google maps	geographic (lat, lon)	N/A	equator/prime meridian	degree of arc
27700	British National Grid	Britain, Northern Ireland	cartesian (x,y)	Transverse Mercator (**not equal area**)	49.75, -9.0 (WGS 84)	meters
3035	European Terrestrial Reference System 1989 ensemble	Europe	cartesian (x,y)	Lambert azimuthal **equal-area** projection	32.88, -16.1 (WGS 84)	meters

Coordinate Reference Systems (CRS)

EPSG Code	Name	Area of use	Coordinate System	Projection	Origin	Unit
4326	World Geodetic System 1984 (WGS84)	world used in GPS, Google maps	geographic (lat, lon)	N/A	equator/prime meridian	degree of arc
27700	British National Grid	Britain, Northern Ireland	cartesian (x,y)	Transverse Mercator (**not equal area**)	49.75, -9.0 (WGS 84)	meters
3035	European Terrestrial Reference System 1989 ensemble	Europe	cartesian (x,y)	Lambert azimuthal **equal-area** projection	32.88, -16.1 (WGS 84)	meters



The Guardian

Overview Geographic Information Systems (GIS)

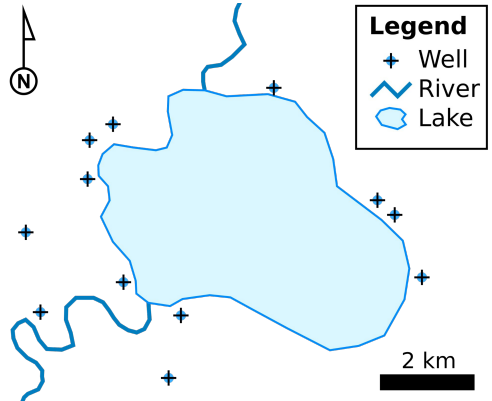
- *spatial data formats*
- *spatial databases / PostGIS*

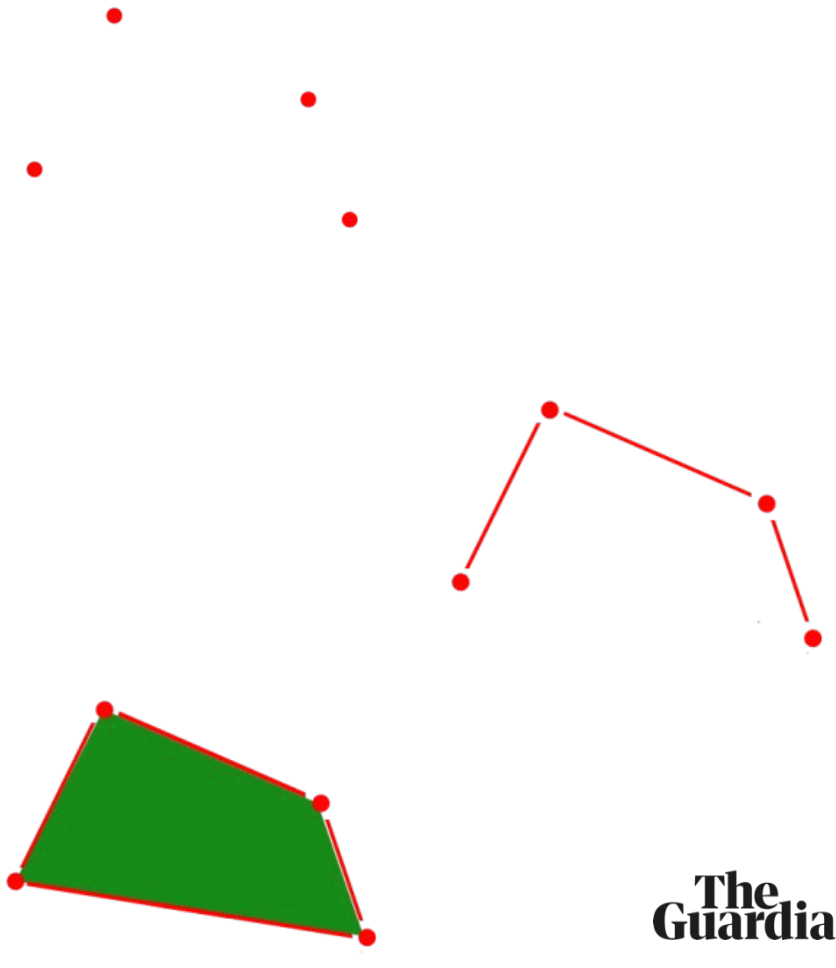
The Guardian

Spatial data formats

Vector Data

- individual points stored as coordinate pairs
- **points** joined to form **lines** or in closed areas to form **polygons**
- represent **discrete** locations

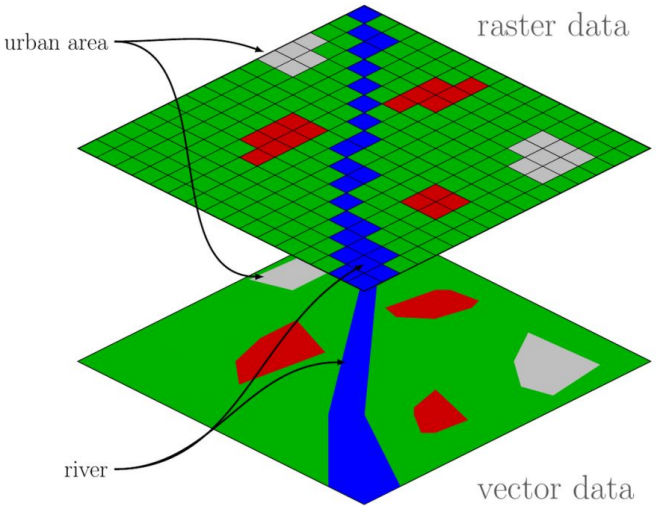




Spatial data formats

Raster Data

- surface divided into a grid
  - each cell has associated values
- represents **continuous** data
  - rainfall, elevation, temperature, land cover etc.



PostGIS

Overview

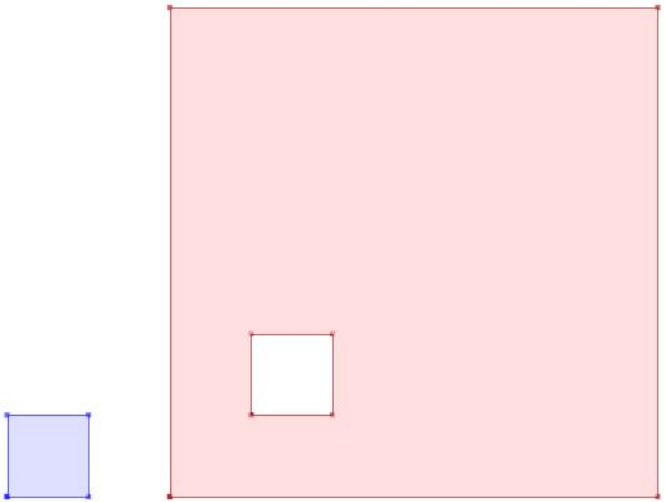
- PostGIS is a **spatial database**
  - store and manipulate spatial objects like any other types of object in the database
- Extension to **Postgres** which provides spatial **data types**, **indexes**, and **functions**
  - MySQL, Neo4j, Elasticsearch, bigquery, SQL Server, SQLite all have geospatial functionality
- raster + vector data

PostGIS

Spatial Data Types

Simple geometries

- `Point`
- `Linestring`
- `Polygon`
  - polygons can have holes in them!



```
POLYGON((0 0, 1 0, 1 1, 0 1, 0 0))
```

```
POLYGON((0 0, 10 0, 10 10, 0 10, 0 0),(1 1, 1 2, 2 2, 2 1, 1 1))
```

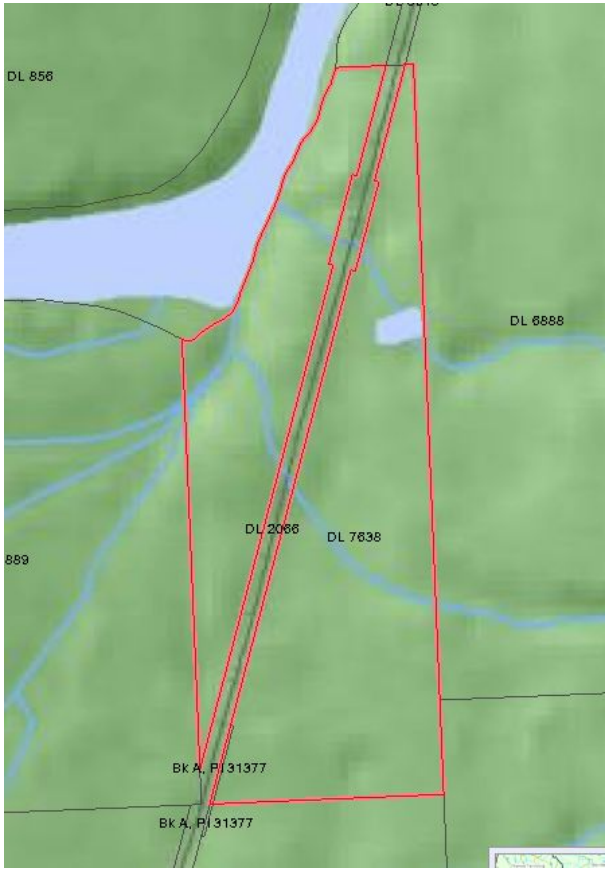
PostGIS

Spatial Data Types

collection types - group multiple simple geometries

- `MultiPoint` - collection of `Point`s
- `MultiLineString` - collection of `Linestring`s
- `MultiPolygon` - collection of `Polygon`s
    - **most useful for land ownership projects**
- `GeometryCollection` - heterogeneous collection of geometries (including other collections)

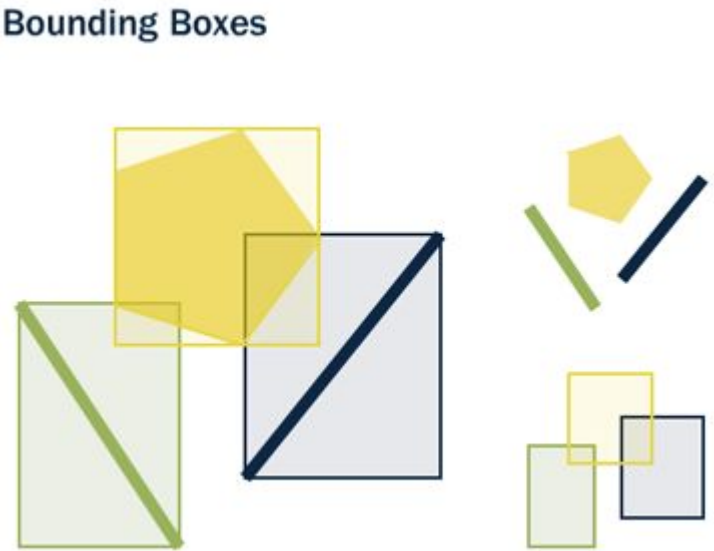
supertype is a `Geometry`



PostGIS

Spatial Indexes

- indexes avoid scanning of every record in table by organizing data into a search tree
- spatial indexes use **bounding box** to index efficiently
  - smallest rectangle parallel to the coordinate axes containing the feature
- determining if e.g. polygon A is inside B is computationally intensive for polygons but fast for rectangles
- spatial indexes provide approximate results first



PostGIS

Spatial Functions

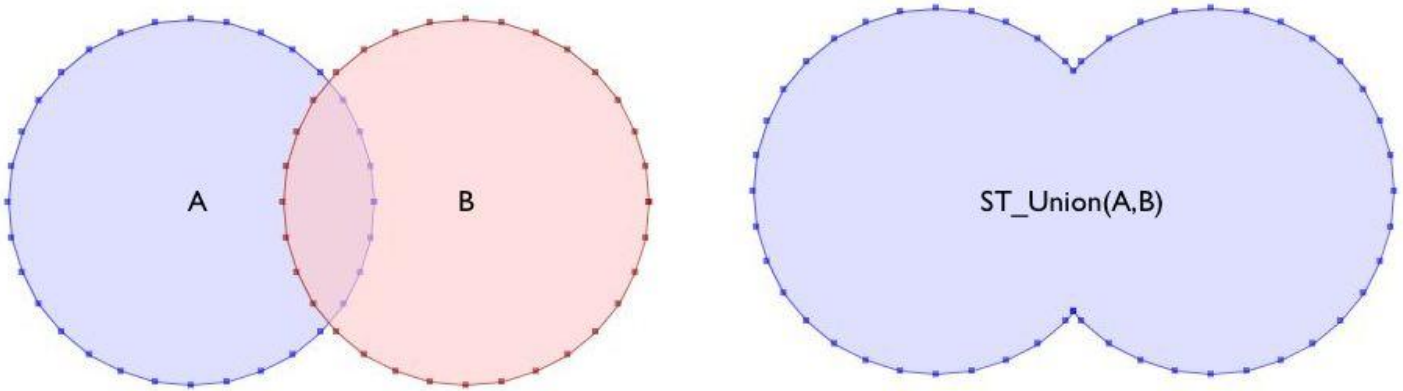
- float **ST_Area**(geometry *g1*);
  - Returns the area of a polygonal geometry. Units specified by the SRID.

- geometry **ST_Transform**(geometry *g1*, integer *srid*);
  - Returns a new geometry with its coordinates transformed to a different spatial reference system.

PostGIS

Spatial Functions

- geometry **ST_Union**(geometry[] *g1_array*);
- geometry **ST_Union**(geometry *g1*, geometry *g2*);
    - Unions the input geometries to produce a result geometry with no overlaps

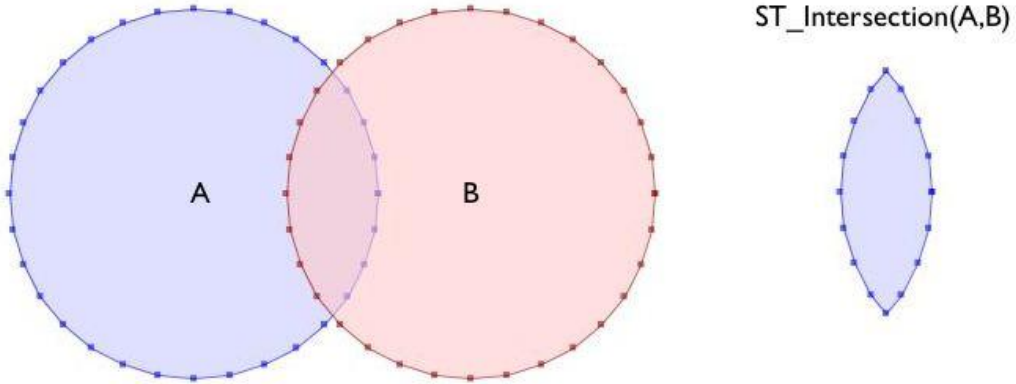


The Guardian

PostGIS

Spatial Functions

- boolean **ST_Intersects**(geometry *geomA* , geometry *geomB*);
  - Returns **true** if two geometries intersect.
  - useful for joining tables
- geometry **ST_Intersection**(geometry *geomA* , geometry *geomB*);
  - Returns a geometry representing the intersection of two geometries.



The Cost of the Crown

The Guardian

Cost of the Crown

- Investigations team scrutinises royal wealth in run up to coronation
  - links to slavery
  - royal collection
  - incomes from crown estate
  - **royal duchies - Duchies of Lancaster and Cornwall**

Royal Duchies

Duchy of Cornwall

- created in 1337
- belongs to male heir to the throne (or sovereign if there isn't one)
- > 240k acres of property
  - ~ 130k acres of land
  - ~ 100k acres of foreshore
  - 14k acres of estuaries, riverbeds, mineral rights
- notable locations
  - Highgrove in Gloucestershire - Prince of Wales' private home
  - 70k acres of Dartmoor, leased to MoD
- legal status
  - doesn't pay corporation tax
  - not subject to FOI requests
  - permission from parliament to sell land
  - **not** included in Land Registry CCOD

arms of Duke of Cornwall

Duchy of Lancaster

- created in 1351
- belongs to sovereign since Henry VIII
- ~ 45k acres of land
- notable locations
  - Queen's chapel of Savoy
  - Village in Cheshire
- legal status
  - doesn't pay corporation tax (King voluntarily pays tax on income)
  - not subject to FOI requests
  - included in Land Registry CCOD

arms of Duke of Lancaster

Bona Vacantia

- bona vacantia =“vacant goods” - assets of people who die without a will and without known blood relatives
- in most of England and Wales, these assets go to the Treasury
- since medieval period, Duchies take *bona vacantia* assets of those who die in respective counties palatine
- Duchy claims to donate *bona vacantia* funds to charity
- Investigations team see documents revealing
    - Duchy of Lancaster uses BV funds to maintain “heritage assets” including rental properties, barns, petrol stations, etc.
    - maintenance includes new roofs, painting, new boilers, new double-glazed windows. etc.
    - duchy invests BV funds in
        - Areas of outstanding natural beauty
        - National heritage conservation area
        - Sites of special scientific interests
        - Special Areas of Conservation

Bona Vacantia

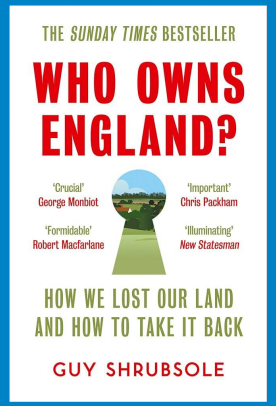
- tl;dr - King Charles secretly profiting from the assets of dead citizens
- want to know the scale of this problem i.e. **What proportion of the Duchy of Lancaster's property falls into these four categories (BV areas)?**

Approach

1. find data representing boundaries of Duchy of Lancaster properties and BV Areas
2. load data into a PostGIS database
3. do some preprocessing in PostGIS
4. **create PostGIS queries to calculate proportion of duchy land that lies inside BV areas**

Sourcing data - Duchy properties

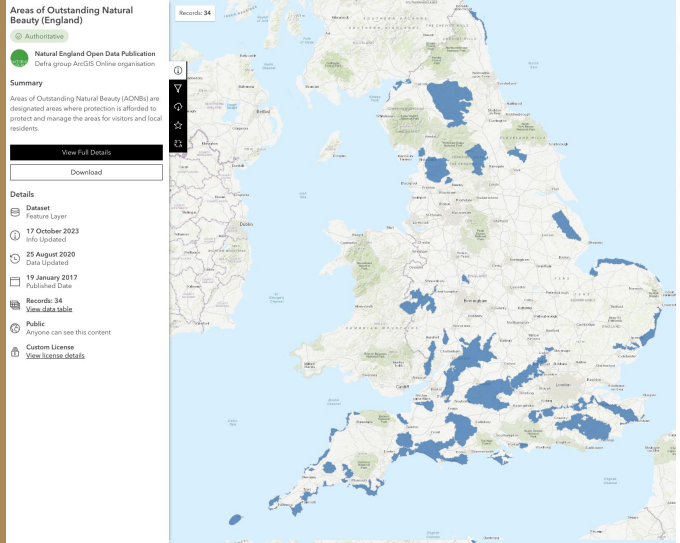
- Shapefiles of Royal Duchies from Guy Shrubsole
- Created in research for chapter on Royal Duchies in Who Owns England https://whoownsengland.org/

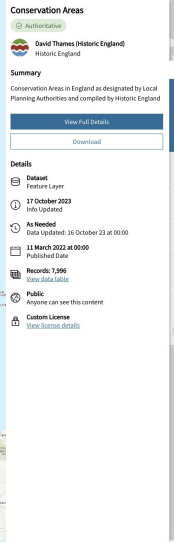


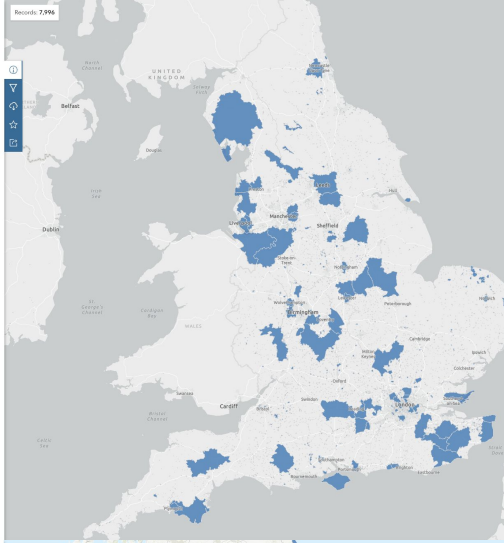
The Guardian

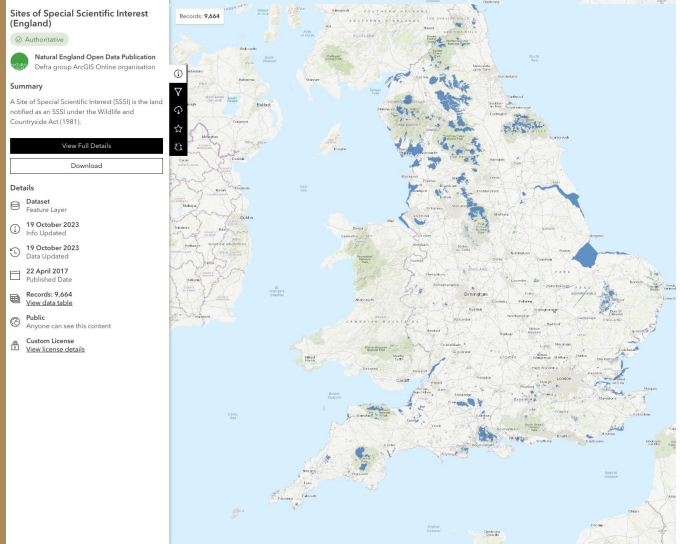
Sourcing data - bona vacantia areas

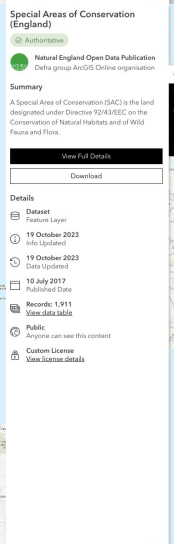
- Boundary data available from Natural England and Historic England data portals
- Can be downloaded as Shapefiles

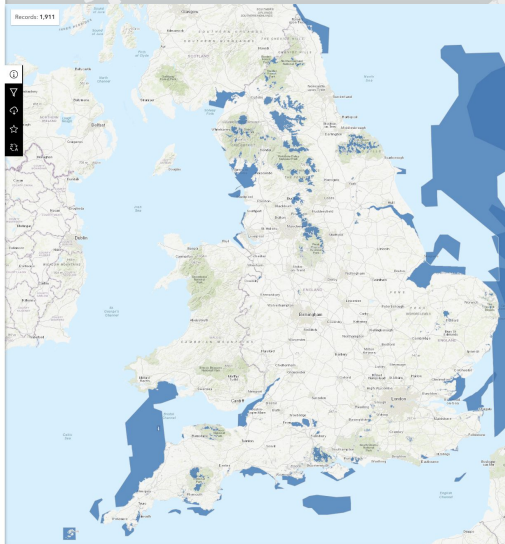












Loading data into PostGIS Database

- 5 shapefiles with same CRS (EPSG 27700 - British National Grid)
- **ogr2ogr** - command line utility for converting data between GIS data formats, including common file formats and common spatial databases

```
ogr2ogr \
 -nln duchy_of_lancaster \ # new layer (table) name
 -nlt PROMOTE_TO_MULTI \ #  automatically promote layers that mix polygon or multipolygons to multipolygons
 -lco GEOMETRY_NAME=geom \ # column to contain geometry data
 -lco FID=gid \ # primary key column name
 Pg:"dbname=lurch host=localhost user=lurch_master password=xyz port=15432" \ # db connection string
 Duchy\ of\ Lancaster\ -\ all\ estates.shp  # path to shapefile
```

- same for 4 BV areas. Table names:
  - aonb (areas of outstanding natural beauty england)
  - conservation_areas (national heritage conservation areas)
  - sssi (sites of special scientific interest)
  - special_areas_of_conservation

The Guardian

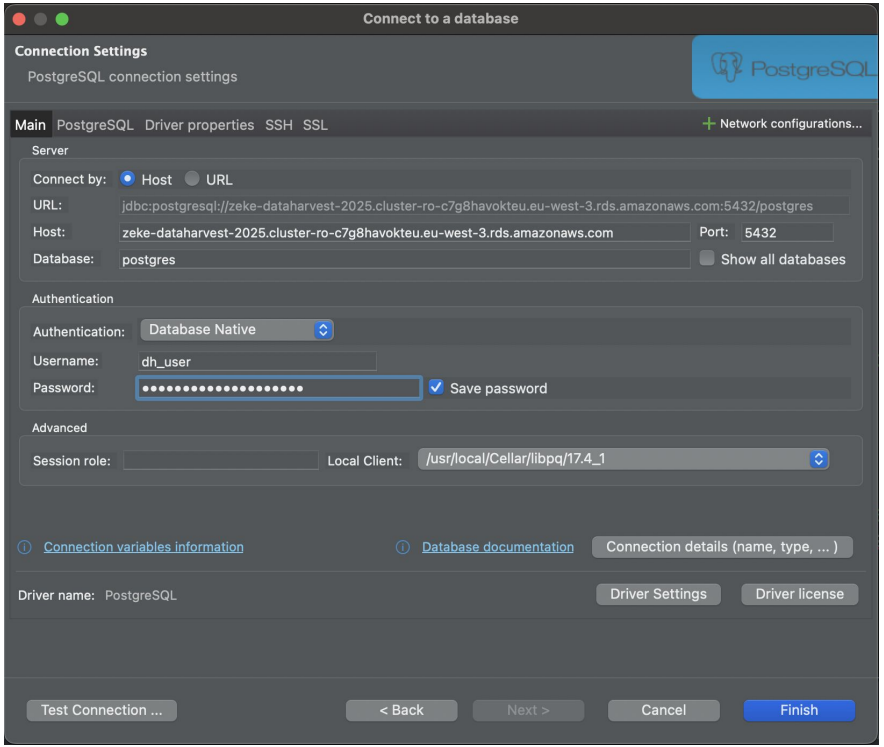
Preprocessing

- data is using CRS EPSG 27700 which **isn't equal area**!
- transform geometries to a CRS using a projection which preserves area
- EPSG 3035 works for this purpose

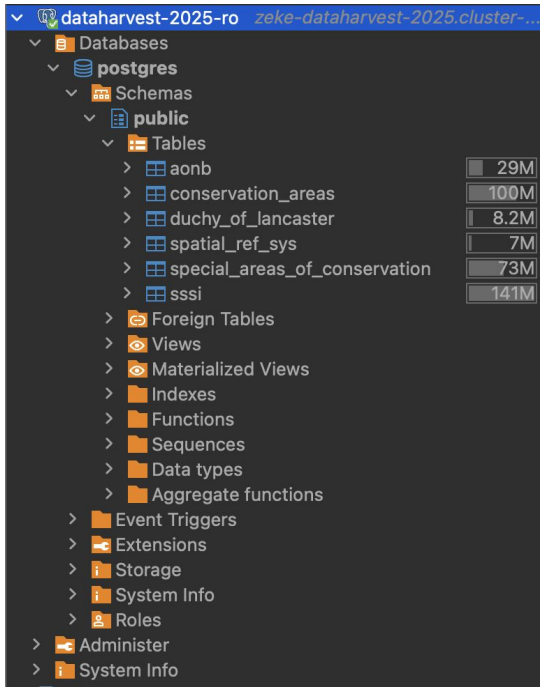
```
alter table duchy_of_lancaster add column geom_3035 geometry(multipolygon, 3035);

update duchy_of_lancaster set geom_3035 = st_transform(geom, 3035);

CREATE INDEX ON duchy_of_lancaster USING GIST (geom_3035);

-- same for 4 other tables
```

Connect to database in DBeaver

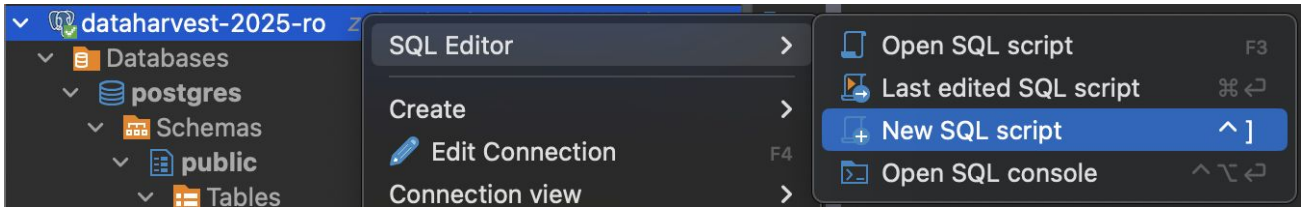
- Top right -> New Database Connection
- Select “Postgres”
    - Host **zeke-dataharvest-2025.cluster-ro-c7g8havokteu.eu-west-3.rds.amazonaws.com**
    - Database **postgres**
    - Port **5432**
    - Username **dh_user**
    - Password **CVQM%aHW%gN987Z*38sv**



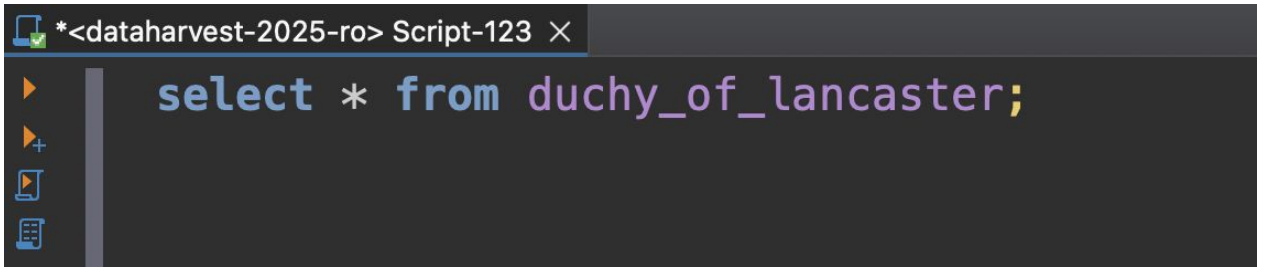
Execute an SQL query



1. Expand connection in Database Navigator to view tables



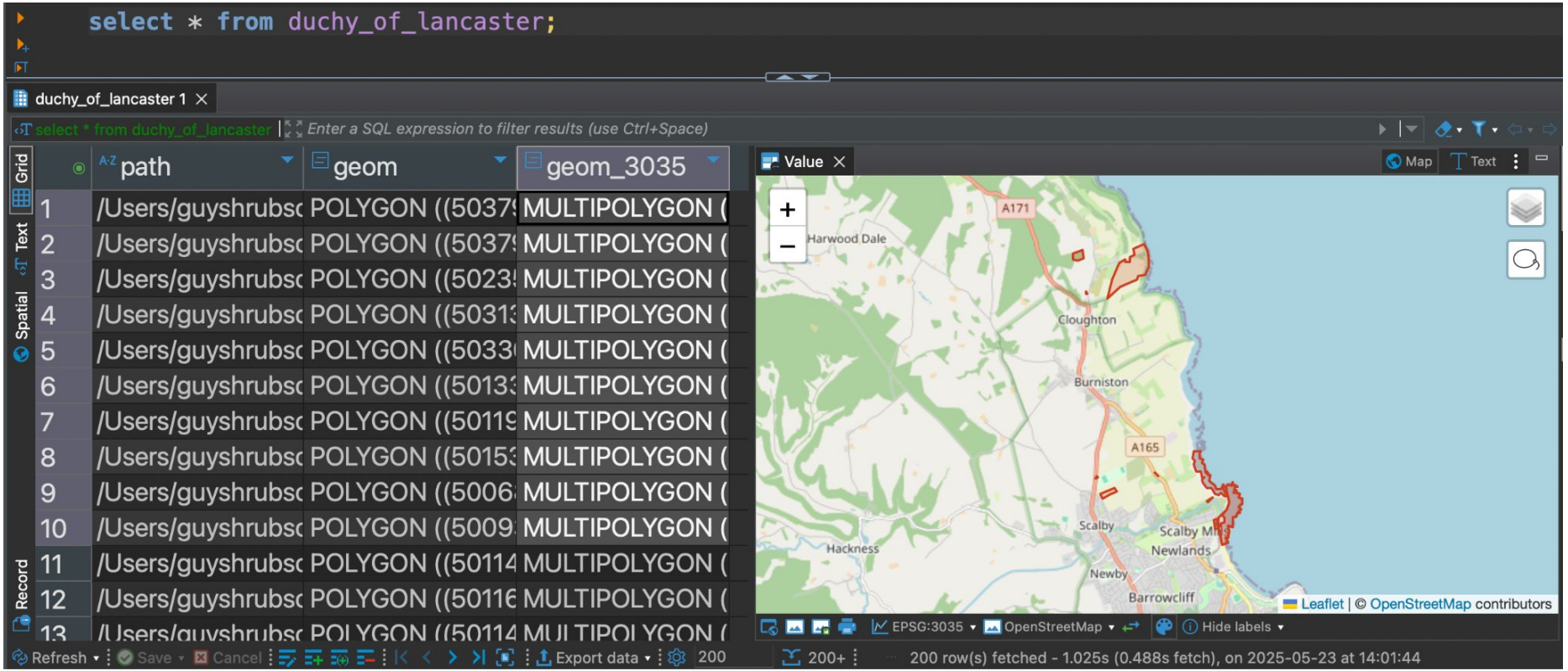
2. Right click on connection > SQL Editor > New SQL script



3. write some SQL - Command + Enter (on mac) or play button to execute

The Guardian

Execute an SQL query

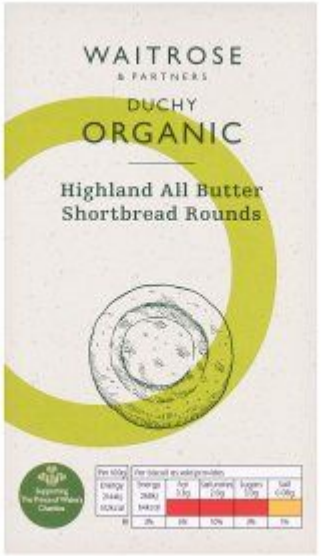


The Guardian

Use PostGIS to calculate the proportion of Duchy of Lancaster land that lies inside BV areas

prizes for

- first correct answer
- fastest query



Use PostGIS to calculate the proportion of Duchy of Lancaster land that lies inside BV areas

- area of duchy of lancaster, unioned
- union of bv areas
- find intersection of duchy with bv areas
- calculate area

Stuck?

Solutions here

Conclusion

- A bit of geodesy background is useful
  - must use appropriate projection for use case
  - must use same CRS across different sets
- PostGIS is amazing
  - use spatial indexes to speed up queries

Thank you

The Guardian