



SAPIENZA  
UNIVERSITÀ DI ROMA

## Introduction to Spatial Data Analysis

### 1) Introduction to GIS and geodata

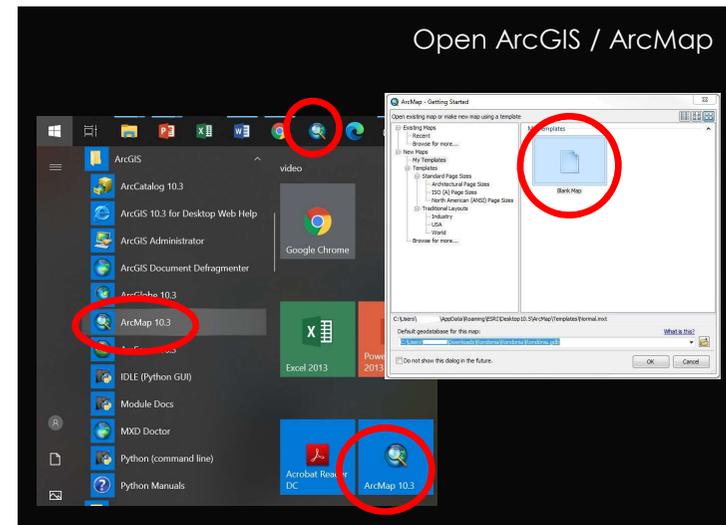
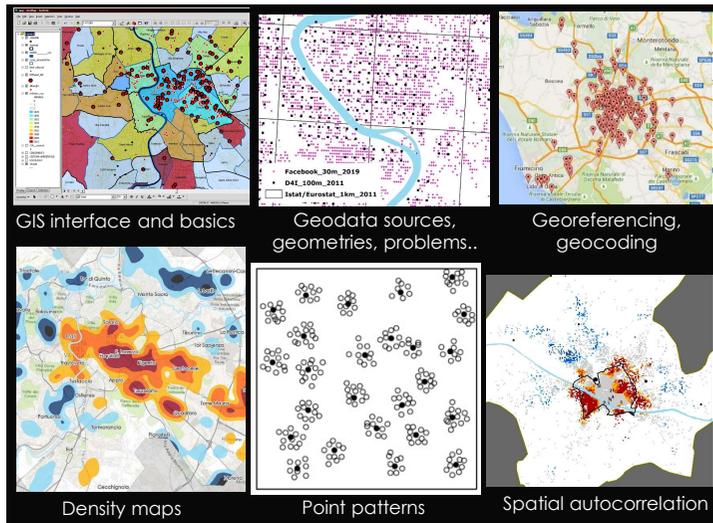
Economic Geography | EPOS | 2022

Filippo Celata (filippo.celata@uniroma1.it)

## Contents

1) **Introduction to GIS and geodata** | Introduction to GIS software (ArcGIS, QGIS). Specificities and sources of spatial data. Spatial coordinates, scales, geometries and partitions. Georeferencing and geocoding. Geoprocessing. Mapping (with ArcGIS).

2) **Spatial statistics** | Introduction to spatial statistics. Point processes, spatial clustering, density maps. Measurement, mapping and interpretation of global and local spatial autocorrelation. Exploratory spatial data analysis with ArcGIS.



Senza titolo - ArcMap

Table of contents (layers)

'list by drawing order' vs. 'list by source'

Layers' overlay and symbology

To remove layers: right click / remove

Add data and basic tools

1) Click add data and add desktop/spatial22/basemap/ (add all shapefiles)

'Connect to folder' (if the data folder is not in the 'look in' window)

zoom

"where is my map!?"

Select features and clear selection

"what is this?"

Find text or X/Y

Measure (in map units)

Tool bar (Customize / Toolbars / Tool bar)

Add data in ArcPRO

Add vector data in QGIS

Spatial data

Points

Lines

Polygons

Surfaces

Vector

Raster

## Typologies of geodata

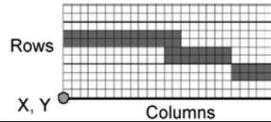
### ◆ Vector formats

#### ◆ Discrete representations of reality



### ◆ Raster formats

#### ◆ Use square cells to model reality



Reality  
(A highway)

Geodata spatial resolution (vs. Scale / Extent)

## Vector data files

Add Data

Look in: usosuolo\_bacino

Name	Type
usosuolo_bacino.shp	Shapefile

Name:

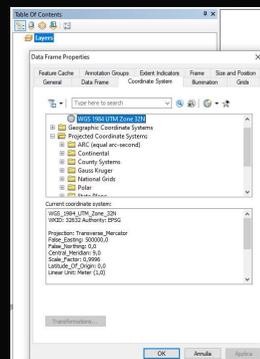
Show of type: Datasets and Layers

Indirizzo: C:\arcgis\tuscia\WEEK1

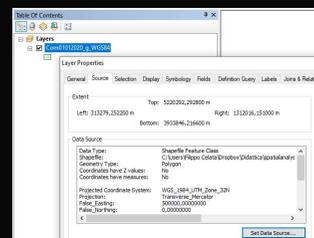
usosuolo_bacino_33N	Table	10.114 KB	File DBF
usosuolo_bacino_33N.prj	Coordinate system	1 KB	File PRJ
usosuolo_bacino_33N.sbn	Spatial index	557 KB	File SBN
usosuolo_bacino_33N.sbx	Spatial index	31 KB	File SBX
usosuolo_bacino_33N.shp	Geometry	134.531 KB	File SHP
usosuolo_bacino_33N.shx	Index of geometry	466 KB	File SHX
usosuolo_bacino_33N	Layer file	38 KB	ArcGIS Layer
usosuolo_VT.shp_VOSTRO.2220.688.sr.lock		0 KB	File LOCK

## Coordinate systems and projections

Coordinate system of the dataframe/map



Coordinate system of each shapefile



To **change the shapefile's coord. system**: set the coord. system of the dataframe and, in ArcGIS, right click on the layer/Data/Export data, in QGIS, 'save vector as', saving with 'the same coordinate system as the dataframe'

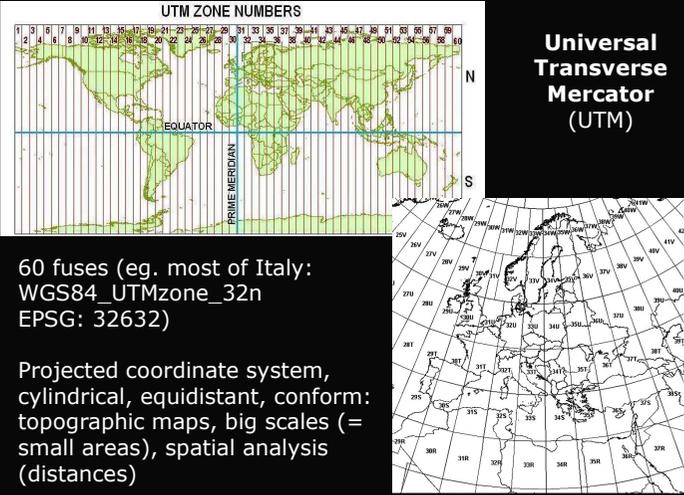
## Coordinate systems and projections

How to represent a spheroid on a bi-dimensional surface minimizing distortions in terms of:

- I. shapes (**conform** projections)
- II. extent (**equivalent** projections)
- III. distance (**equidistant** projections)
- IV. angles/routes (**isogonic** projections)



- **Geographic (DMS) vs. Projected** coordinate systems (Meters)
- **Best case**: coordinate system of the output map = coordinate system of the dataframe = coordinate system of each layer



**UTM ZONE NUMBERS**

**Universal Transverse Mercator (UTM)**

60 fuses (eg. most of Italy: WGS84\_UTMzone\_32n  
EPSG: 32632)

Projected coordinate system, cylindrical, equidistant, conform: topographic maps, big scales (= small areas), spatial analysis (distances)



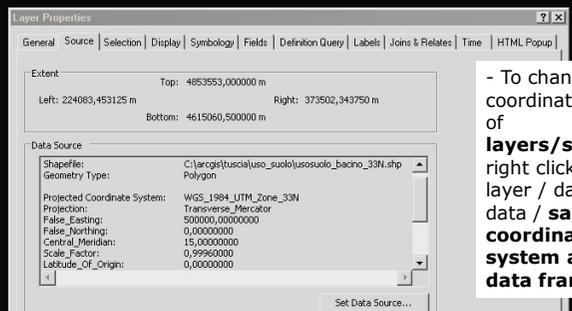
**World Geodetic System 1984**

(WGS84  
EPSG:4326)

**World Geodetic System**, cartographic standard since 1960, used in cartography, geodesy, and satellite navigation (GPS). Last version: 1984, established and maintained by the US National Geospatial-Intelligence Agency, based on a standard reference ellipsoid model

### Set the coordinate system

- To set the dataframe (map) coordinate system: layers / properties / coordinate system -> change the coordinate: right click on 'layers' / dataframe properties / 'geographic coordinate systems' / World / WGS 1984

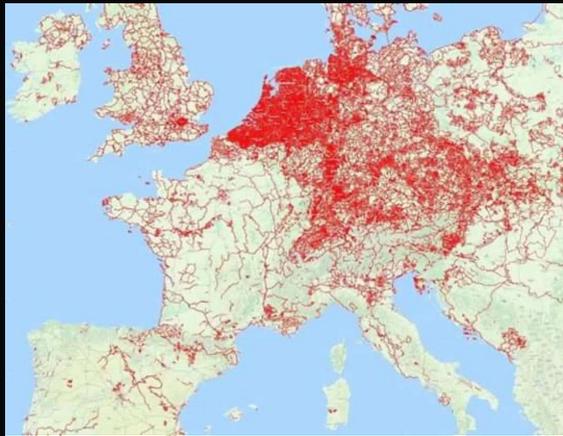


- To change the coordinate system of layers/shapefiles: right click on the layer / data / export data / same coordinate system as the data frame

## Sources and specificities of secondary geodata



## "European bicycle path density" ?



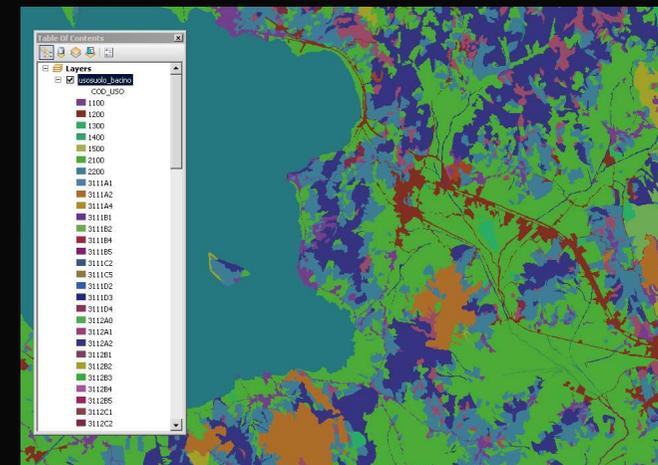
## Openstreetmap layers: buildings



## Openstreetmap layers



## Land use





**Administrative partitions** pros and cons: homogeneity/uniformity (extent/attributes), irregularity/compactness (shape), ..arbitrariness

And non-administrative polygonal spatial data units (regions..): e.g. TRAVEL-TO-WORK AREAS



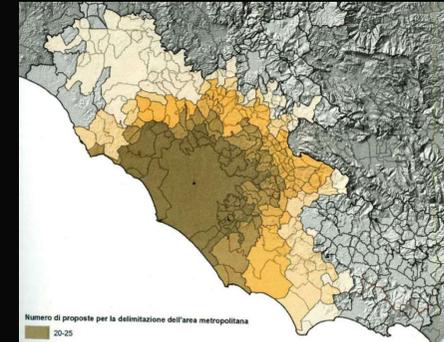
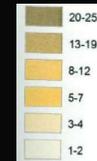
**What is a city? - Spatial units**

Data are collected for several levels:

- A **City** is a local administrative unit (LAU) where the majority of the population lives in an urban centre of at least 50 000 inhabitants.
- The **Functional Urban Area** consists of a city and its commuting zone. (This was formerly known as larger urban zone (LUZ)).
- The **Greater city** is an approximation of the urban centre when this stretches far beyond the administrative city boundaries.

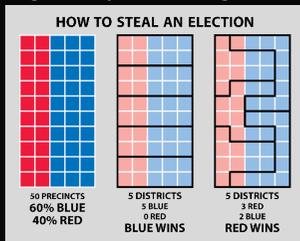
Eurostat / OECD

Number of proposals for the delimitation of the metropolitan area of Rome (Celata-Martellozzo, 2014)

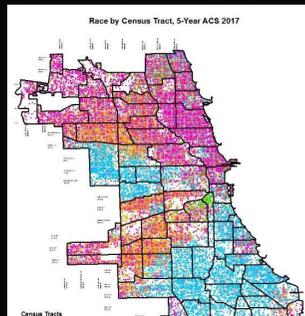


**The modifiable area unit problem (MAUP)**

Eg. Gerrymandering..



Eg. sociospatial segregation



Eg. Concentration

4	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

1	1	0	0	0
1	1	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

1	1	0	0	0	0	0
1	1	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

The **modifiable area unit problem (MAUP)**: any geographical discontinuity is artificial, (more or the less) arbitrary, modifiable, and influences results and explanations

-**Scale problem**,  $f(\text{Spatial resolution})$ . E.g. statistical relations are stronger the lower the spatial resolution, because variance is lower = the more we aggregate data, the stronger they correlate. The higher the spatial resolution, the more the variance, the probability of errors (which are not independent: eg. population is attributed to the neighbor), and of outliers.

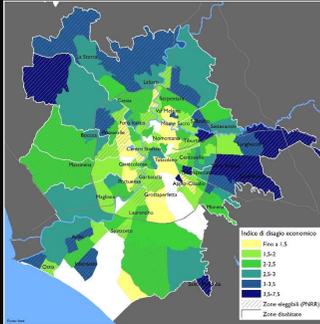
-**Zoning problem**,  $f(\text{Geodata geometry})$ , for any given number of zones, results are influenced by their shape.

**Ecological fallacy**: the results of aggregate analysis cannot be attributed to each individual/to higher scales (the rate of suicides is higher where more catholics live = catholics are more keen to suicide?)

Geodata **quality/accuracy** (depends on spatial resolution), **completeness/exhaustiveness**.

## Sub-municipal areas: 'neighbourhoods'

Roma: zone urbanistiche

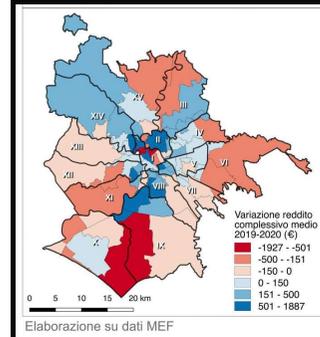


Istat, 'aree sub-comunali' (www.istat.it/it/archivio/104317)



'Circoscrizioni'; 'Municipi', 'zone toponomastiche' (Roma); 'Quartieri' (Venezia, Siracusa); 'Zone decentramento', 'Nuclei identità locale' (Milano)

## Sub-municipal areas: postal codes/areas



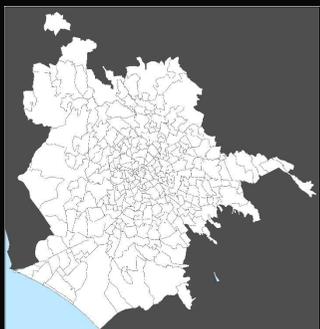
Elaborazione su dati MEF

Roma (per-capita income variation, 2019-2020)



London

## Sub-municipal areas

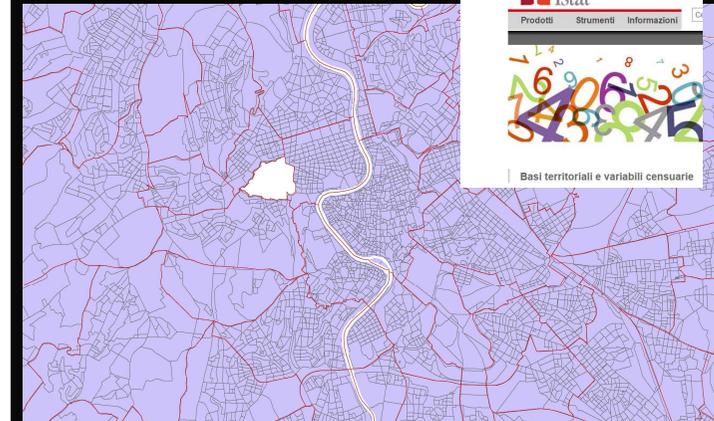


Zone OMI (real estate values)



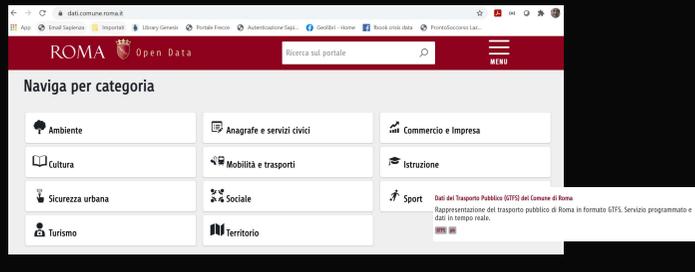
Eg. increasing polarization in 97 over 112 Italian cities 2008-2015 (Antoniucci-Marella 2018)

## Census tracks





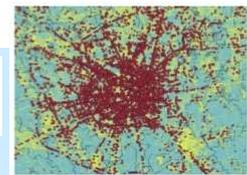
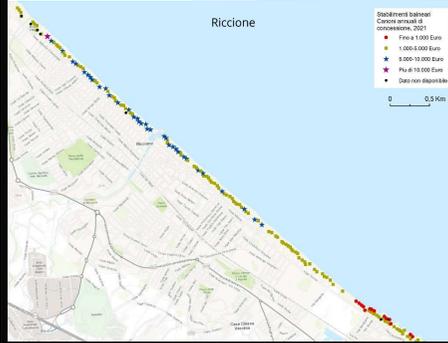
# Open (geo) data



# Open (geo) data

Dall'utilizzo di Open Street Map per gli indicatori di performance sulla sicurezza stradale alla localizzazione geografica degli incidenti stradali

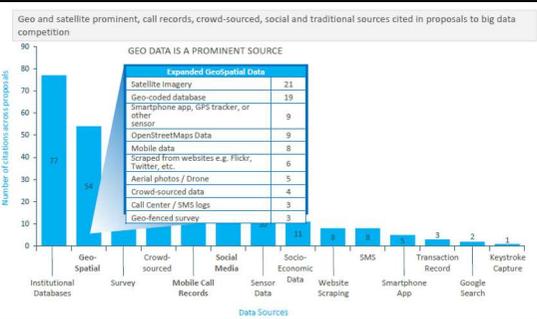
Eg. beach concessions fees



Eg. road accidents

# Big (Geo)Data

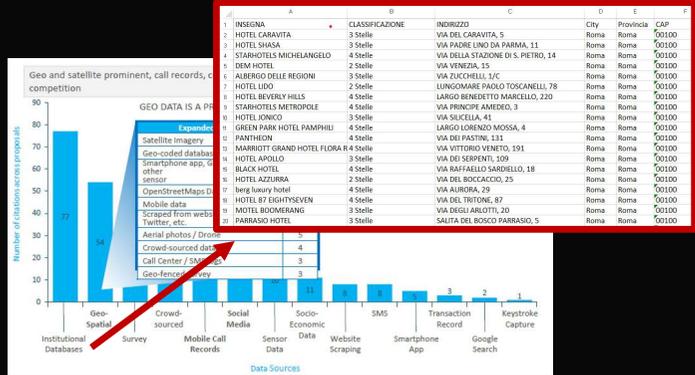
## Sources



# Big (Geo)Data

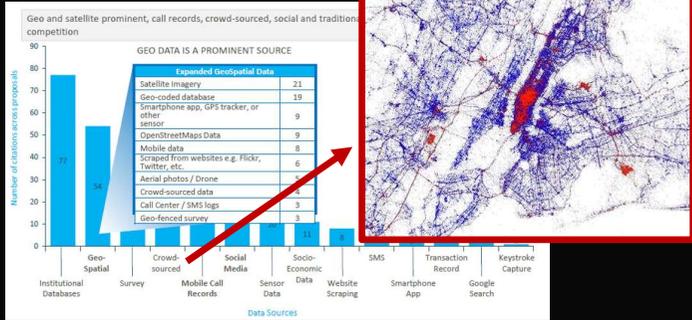
## Public archives

E.g. accommodation facilities



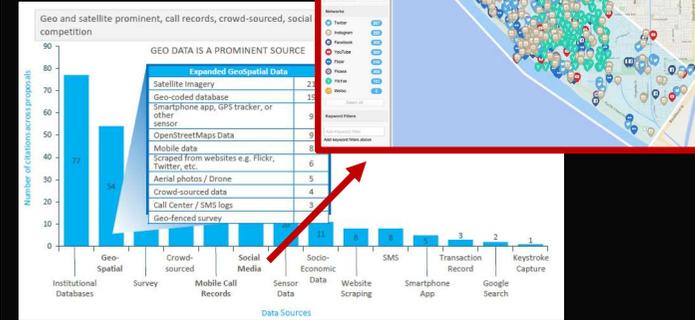
# Big (Geo)Data

Crowdsourced repositories  
Eg. 'locals and tourists' (E. Fischer)



# Big (Geo)Data

Social Media



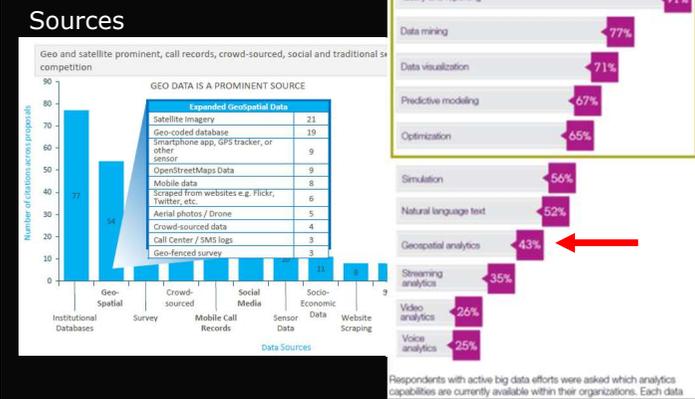
# Big (Geo)Data

Web scraping (eg. Airbnb)



# Big (Geo)Data

...applications



Etc. etc. etc. etc.

Global [ edit ]

GIS data for global datasets

Name	Description
Natural Earth	Public domain vector and raster dataset. Supported by the NACIS. <sup>[1]</sup>
Global Map	Provides consistent coverage of all the Earth's land cover area. Includes different thematic maps such as transportation, elevation, drainage, vegetation, administrative boundaries, land cover, population centres, and land use. <sup>[Registration required]</sup>
UNEP Environmental Data Explorer	Includes global forest cover, global potential evapotranspiration, global average monthly temperatures, dams, watershed boundaries, and much more. Use the advanced search to select geospatial data sets. <sup>[Provided by the United Nations Environment Programme]</sup>
GSNHD	Global Set consistent, Hierarchical, high-resolution Shoreline Database: high quality, and consistent data. <sup>[1]</sup>
ASTER GDEM	30m resolution global elevation data derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer satellite images. <sup>[1]</sup>
OpenTopography	OpenTopography facilitates community access to high-resolution, Earth science-oriented, topography data, and related tools and resources. Find high-resolution, Earth science-based, topography data, and related tools and resources. Available as dense point clouds and DEMs.
NCAR GIS Climate Change Scenarios	Includes data used by the IPCC in their reports. Operated by the National Center for Atmospheric Research. <sup>[Registration required]</sup>
IRULDEO Climate Data Library	Collection of more than 300 datasets from various climate modes and datasets. Operated by The Earth Institute and Lamont-Doherty Earth Observatory.
Global Climate Monitor	Climate web viewer containing accessible climatic information from 1901 that can be downloaded in a variety of GIS formats.
HydroSHEDS	Global hydrological data based on the SRTM elevation data. Includes river networks, watershed boundaries, drainage directions, and flow accumulations.
USGS Land Cover Institute	Set of data from the USGS for numerous land cover datasets. Although primarily US data, further down the list there is data for other continents.
Atlas of the Biosphere: Mapping the Biosphere	Raster maps of environmental variables including soil pH, potential evapotranspiration, average snow depth, and more.
Global 200	Vector data from WWF of "a set of the Earth's terrestrial, freshwater, and marine ecoregions that harbor exceptional biodiversity and are representative of its ecosystems." <sup>[1]</sup>

geodati.gov.it

La ricerca dell'informazione geografica inizia da qui

dati.gov.it

Earth Engine Data Catalog

Canada AAFC Annual Crop Inventory

Allen Coral Atlas (ACA) - Geomorphic Zonation and Benthic Habitat - v1.0

AHN Netherlands 0.5m DEM, Interpolated

AHN Netherlands 0.5m DEM, Non-Interpolated

RESOURCEWATCH

Historical Gridded Population

SOURCE: Ullrich LUPEL

Historical population estimates from 10,000 BC through 2017

ESRI 'basemap' (background)

Editor

1:200,000

Add Basemap

Imagery

Light Gray Canvas

National Geographic

OpenStreetMap

Oceans

Shaded Relief

Streets

Terrain

Topographic

USA Topo Maps

ArcGIS 10

ESRI 'basemap' (background)

ArcGIS Pro - Quick

Project | Map | Insert | Analysis | Edit | Imagery | Share

Basemap

Contents

Map1

Imagery

Imagery with Labels

Streets

Topographic

Dark Gray Canvas

Light Gray Canvas

National Geographic

Terrain with Labels

Oceans

OpenStreetMap

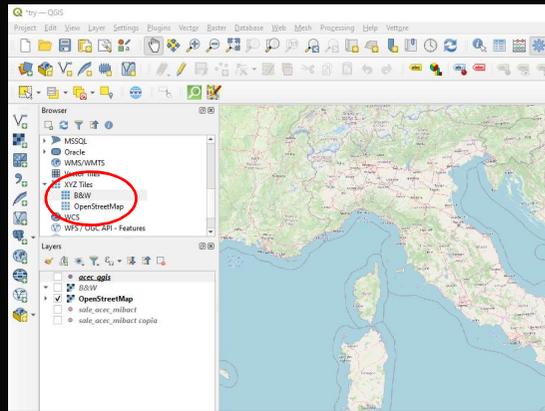
USA Topo Maps

USGS National Map

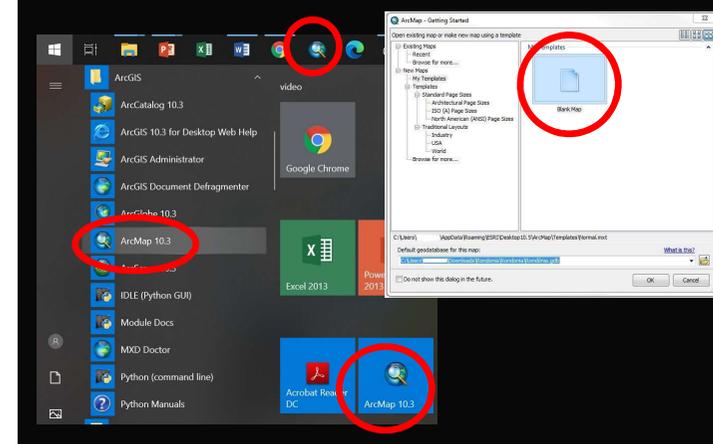
Modern Antique Map

ArcGIS PRO

## Qgis basemaps



## Open ArcGIS / ArcMap



## Lab: Italian gridded population density map

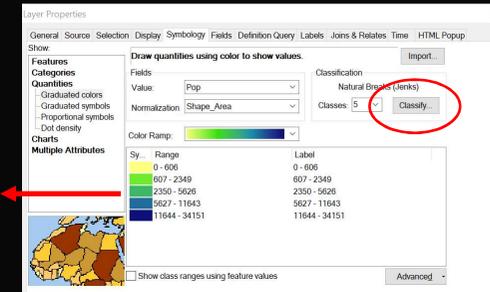
1. On the web, go to: <https://www.istat.it/it/archivio/155162> (STATISTICHE SULLA POPOLAZIONE PER GRIGLIA REGOLARE)
2. Download the 'file geografico':



## Lab: Italian gridded population density map

3. In the 'downloads' folder, unzip the file.
4. Add the shapefile to Arcmap
5. In Arcmap, right click on the layer and click on 'Properties'
6. In the 'symbology' menu, set 'quantities', 'graduated colors'. Set the 'Value' = Pop, and 'Normalization' = Shape\_area

[ Right click on the color ramp, 'properties for all symbols', and make the outline invisible ('outline color' = 'no color' )



## Lab: Italian gridded population density map

7. In the 'Classify' menu, set 'Sampling', 'Maximum sampling size', to 310,000

8. In the 'Classify' menu, number of 'Classes', classes' ranges ('Break values')

9. Overlay the regional partition by adding the shapefile of Italian regions in Desktop/Spatial22/Basemap

10. Check the results on the map, and change the it as you like

