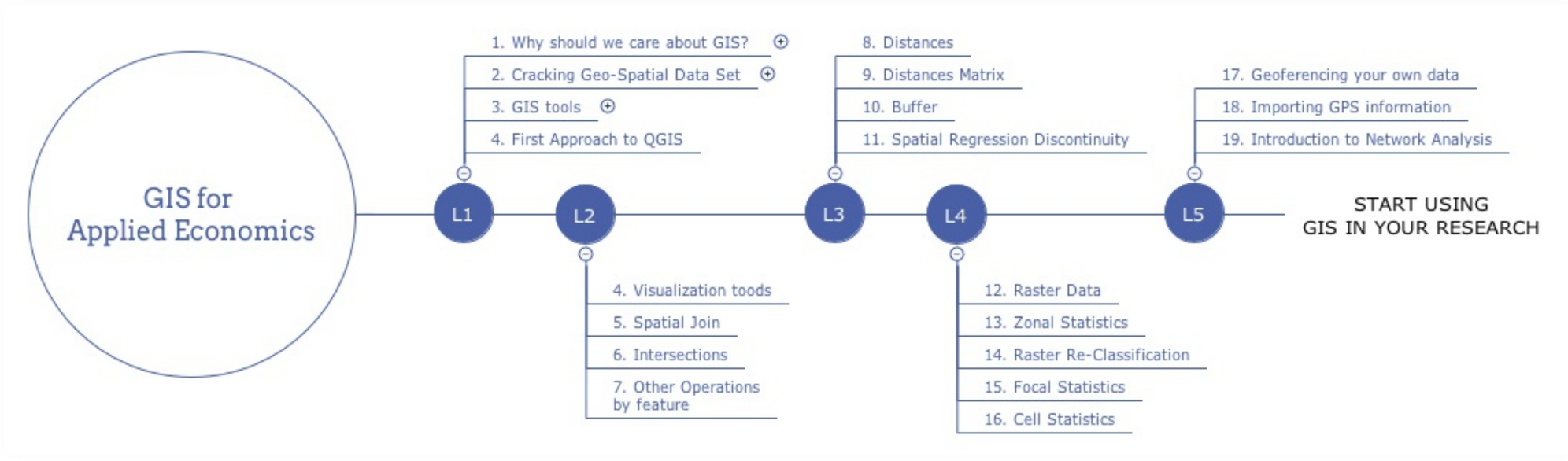
GIS for Applied Economics

Juan Carlos Muñoz-Mora
Univeristat Pompeu Fabra
juancarlos.munoz@upf.edu

Nov 2016

Overview


GIS for
Applied Economics
L1
1. Why should we care about GIS?
2. Cracking Geo-Spatial Data Set
3. GIS tools
4. First Approach to QGIS
L2
4. Visualization toods
5. Spatial Join
6. Intersections
7. Other Operations by feature
L3
8. Distances
9. Distances Matrix
10. Buffer
11. Spatial Regression Discontinuity
L4
12. Raster Data
13. Zonal Statistics
14. Raster Re-Classification
15. Focal Statistics
16. Cell Statistics
L5
17. Geoferencing your own data
18. Importing GPS information
19. Introduction to Network Analysis
START USING
GIS IN YOUR RESEARCH

- Learning by doing approach (½ theory/applications - ½ practice)
- We will replicate published papers, yet we can consider use your own research as practice

- Practice
  - Option 1: A single problem set at the end
  - Option 2: Small task after each session

Why should we care about GIS in Economics?

It is not new concept in Economic thought!

[...] if economics profession has notably neglected it [geography], this is not because economists have been uninterested in the subject, but because they have regarded it as intractable.

Source: Fujita, M; Krugman, P and Venables, A., 1999.

So, what's new?

[1] New amazing GIS data sets available out there

provides a original source of variation

Remotely sensed information

(Donaldson and Storeygard, 2016).

access to information difficult to obtain by other means

unusually high spatial resolution

wide geographic coverage

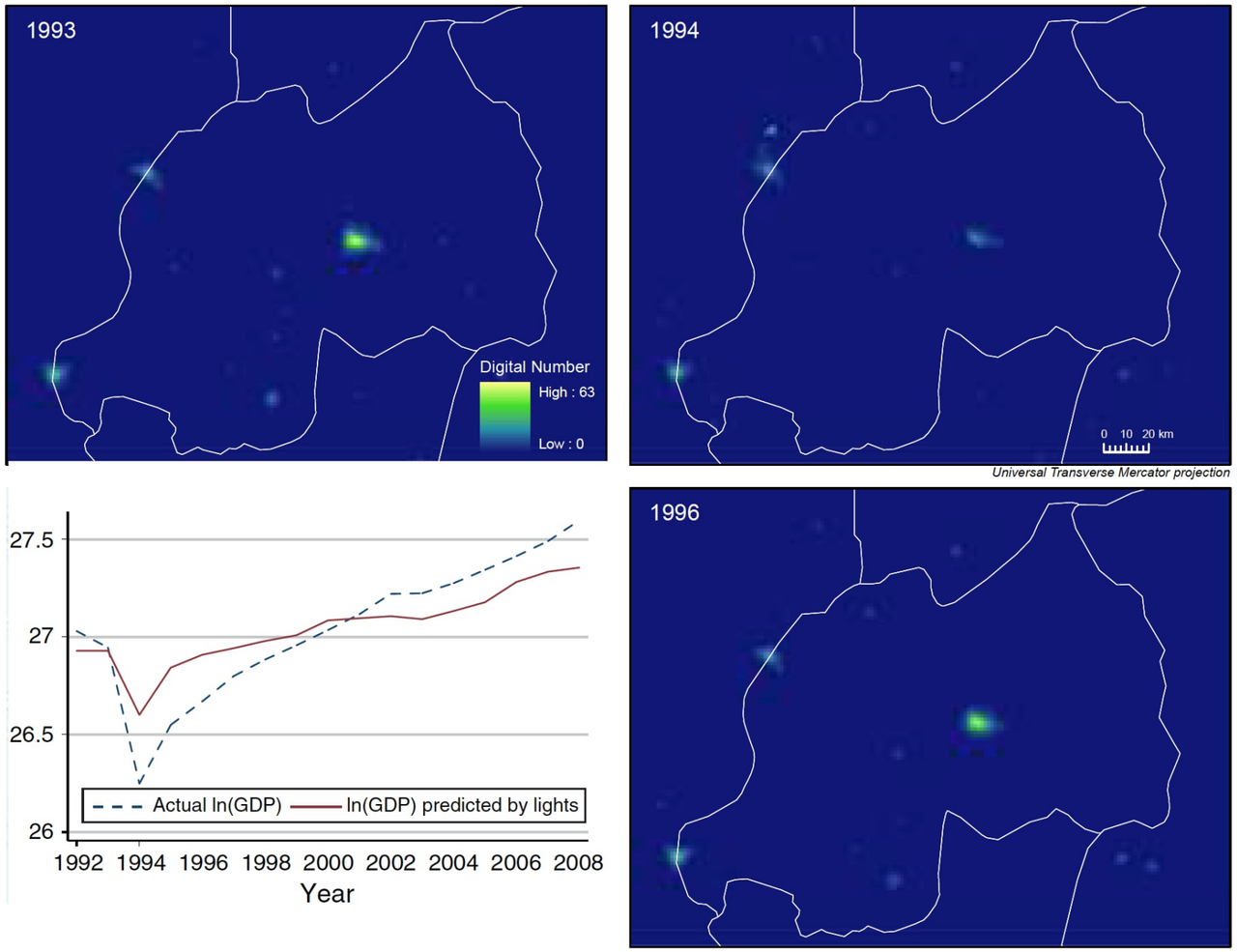


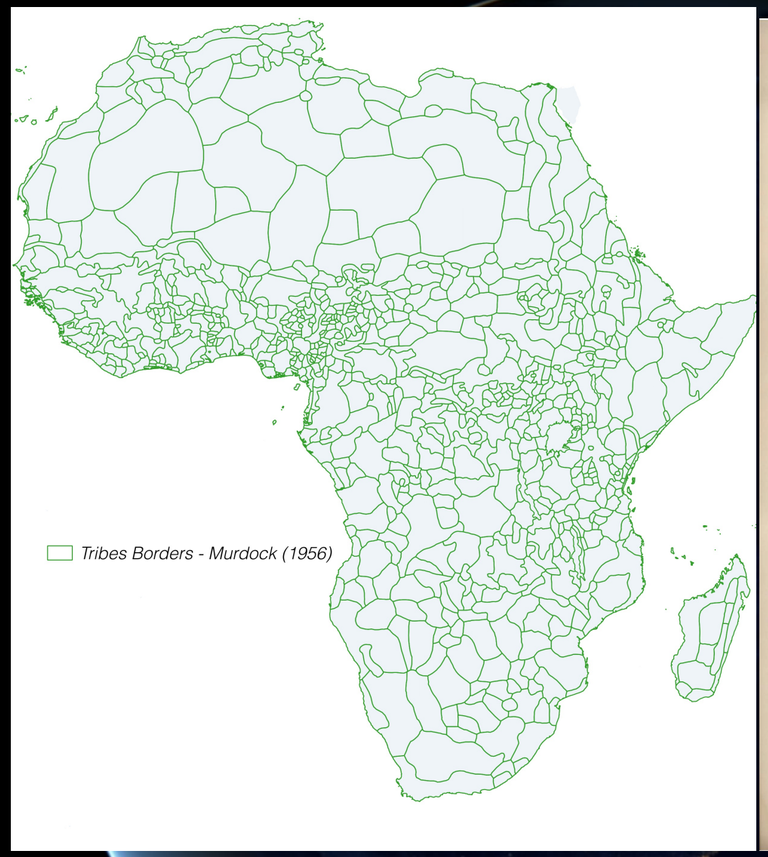
FIGURE 4. GENOCIDE EVENT: RWANDA

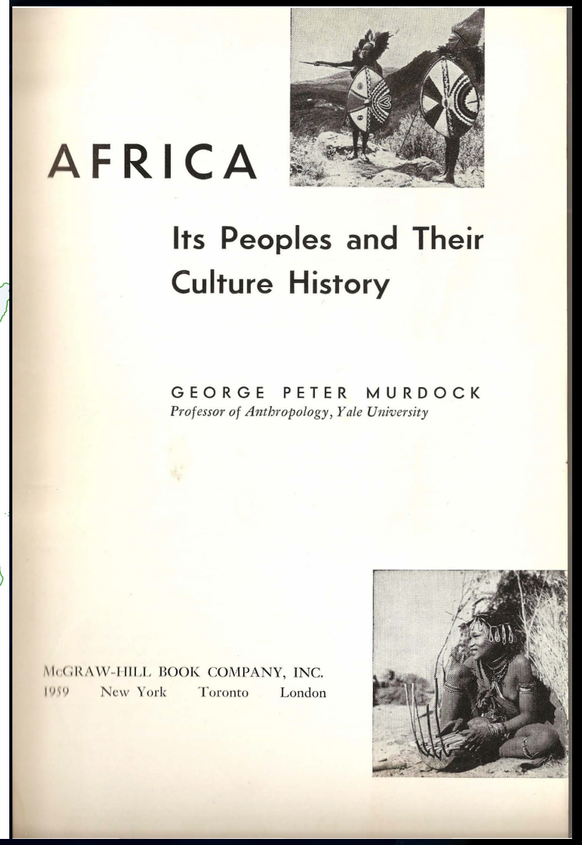
(Vernon Henderson, J.; Storeygard, A. and Weil, D 2012)

Other examples: Henderson et al (2012), Pinkovskiy & Sala-i-Martin (2016), Hodler & Raschky (2014), Michalopoulos & Papaioannou (2013, 2014), Alesina et al (2016), among others.

Historical Maps

recovery old (persistent) characteristics




AFRICA

Its Peoples and Their Culture History

GEORGE PETER MURDOCK
Professor of Anthropology, Yale University

McGRAW-HILL BOOK COMPANY, INC.
1959 New York Toronto London

Digitized by Nunn (2008). He found a negative relationship between the number of slaves exported from a country and current economic performance.

Then use by Nunn & Wantchekon (2011), Michalopoulos & Papaioannou (2013, 2014, 2015), Alsan (2015), Alesina et al. (2016), etc.

Geo-Coded individual-level data sets

exploring geography and location as an important dimension

(a) Basic Pharmaceuticals (SIC2441)

(b) Pharmaceutical Preparations (SIC2442)

(c) Other Agricultural and Forestry Machinery (SIC2932)

(d) Machinery for Textile, Apparel and Leather Production (SIC2954)

FIGURE 1
Maps of four illustrative industries

(Duranton, D. and Overman, H 2005) → location patterns of industries in UK

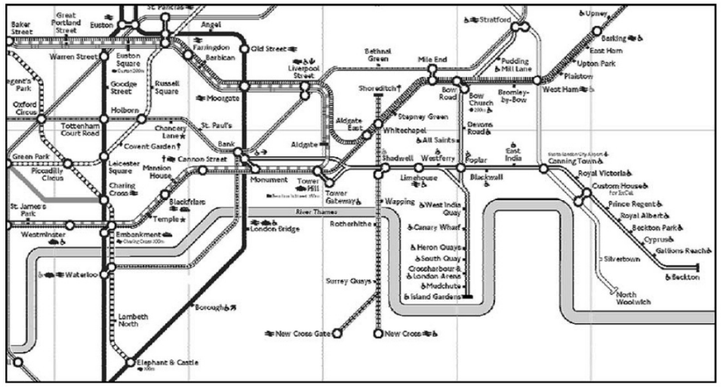
Others examples of geo-coded data sets: .LSMS-ISA (World Bank) -> geo-coded household level 6 African countries

[2] GIS may improve empirical strategy stronger

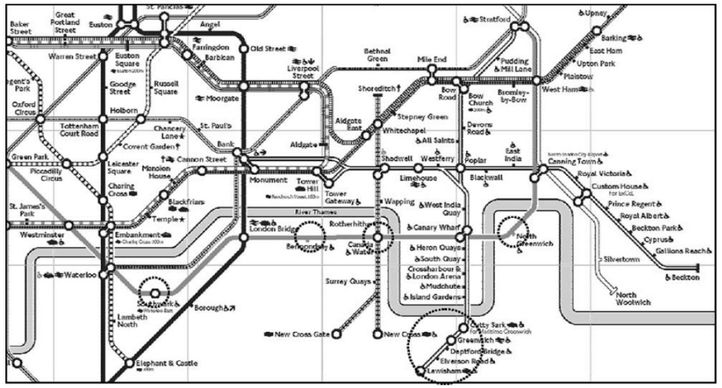
reducing the omitted-variable bias (I)

gathering more information from neighborhood

Fig. 1. The geographical extent of the samples. The larger circle shows the extent of the sample within 30 km of Holborn. The smaller circle shows the extent of the sample within 20 km of Bromley. Greater London boundary shown for comparison.



(a) pre-2000



(b) post-1999

Fig. 2. London Underground Network in Central, and South East London. The Jubilee Line extension is the solid grey line running from Green Park in the West to Stratford in the North East. The Lewisham DLR extension runs South West from Cutty Sark at the bottom right of the diagram. Dotted rings show the relevant new stations.

Gibbons, S. and Machin, S (2005) ↑ transport → ↑ house prices

Other nice examples are: (i) Conley & Udry (2010) social learning in the diffusion Ghana ; (ii) Acemoglu, D; Garcia-Jimeno, C. and Robinson (2014) direct and spillover effects of local state capacity in Colombia

reducing the omitted-variable bias (II)

building instruments

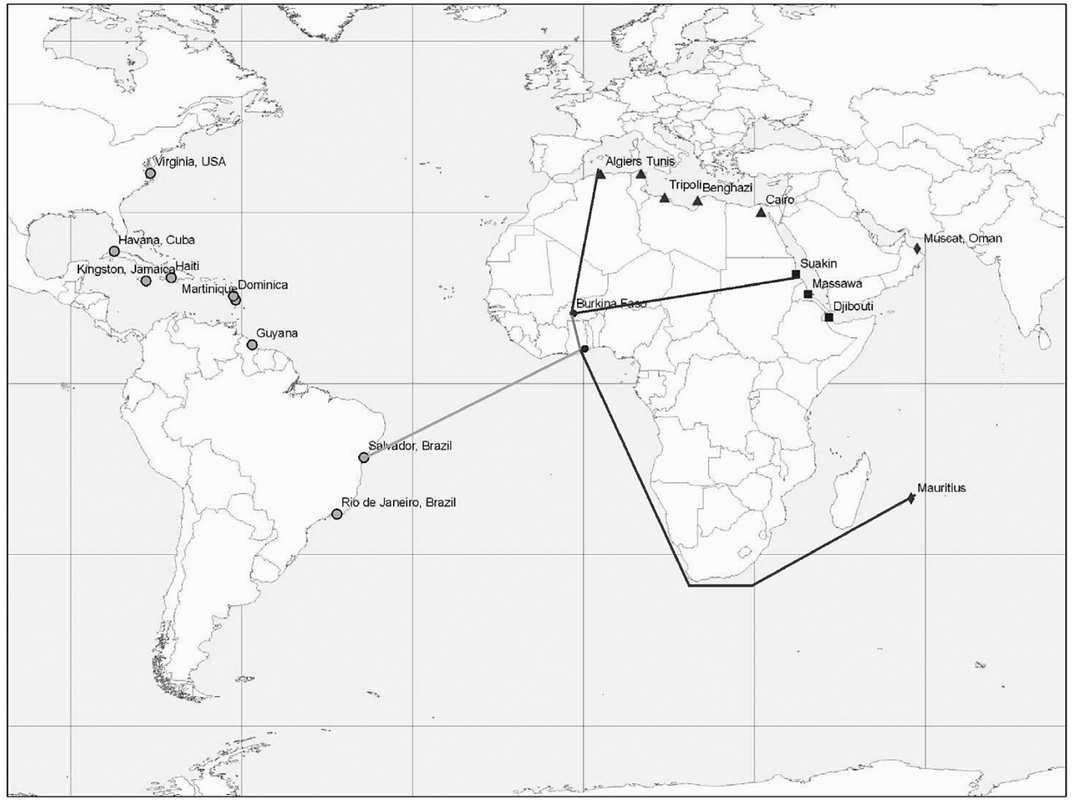


FIGURE V
Example Showing the Distance Instruments for Burkina Faso

Nunn (2008) Current African Countries & Slave Trades

Instrument Slave Export → distances from each African country to the locations where slaves were demanded

New identification strategies

Spatial Regression Discontinuity

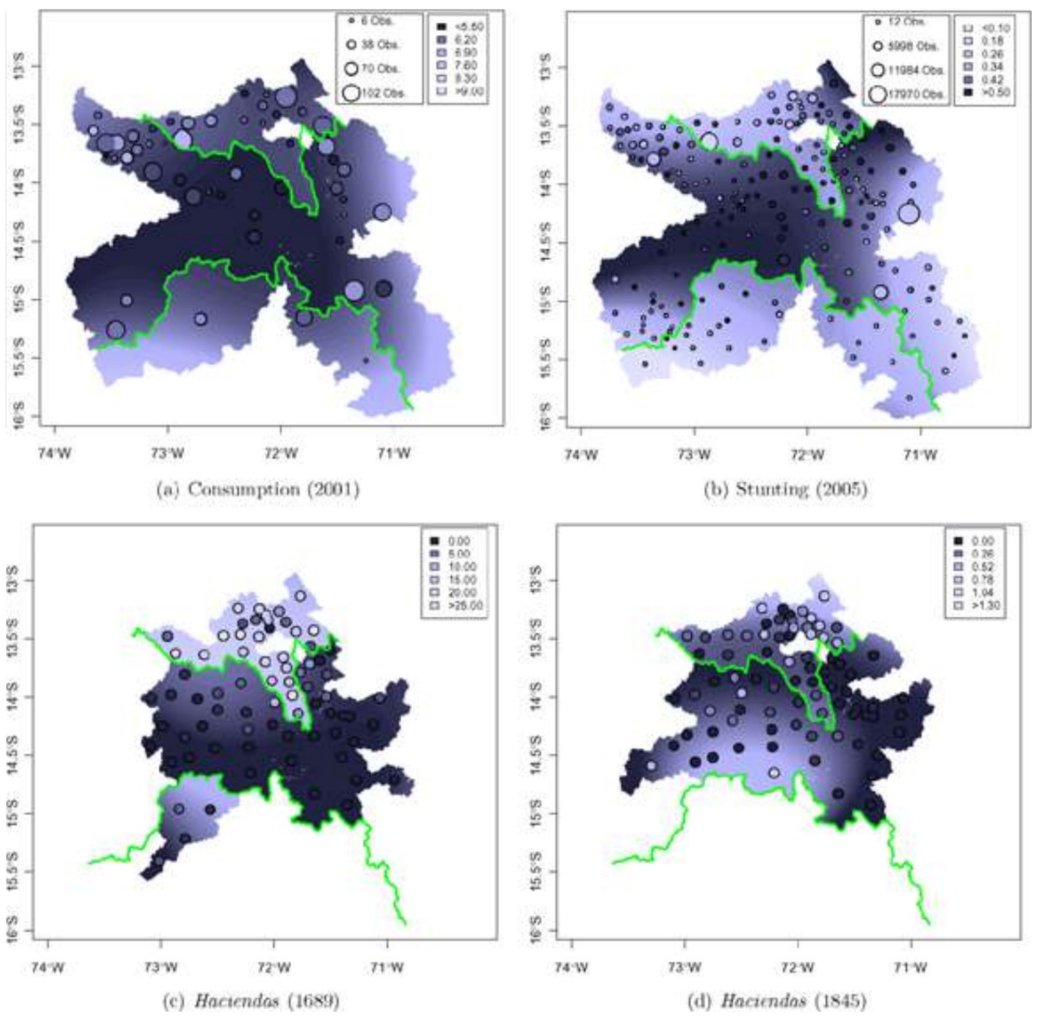


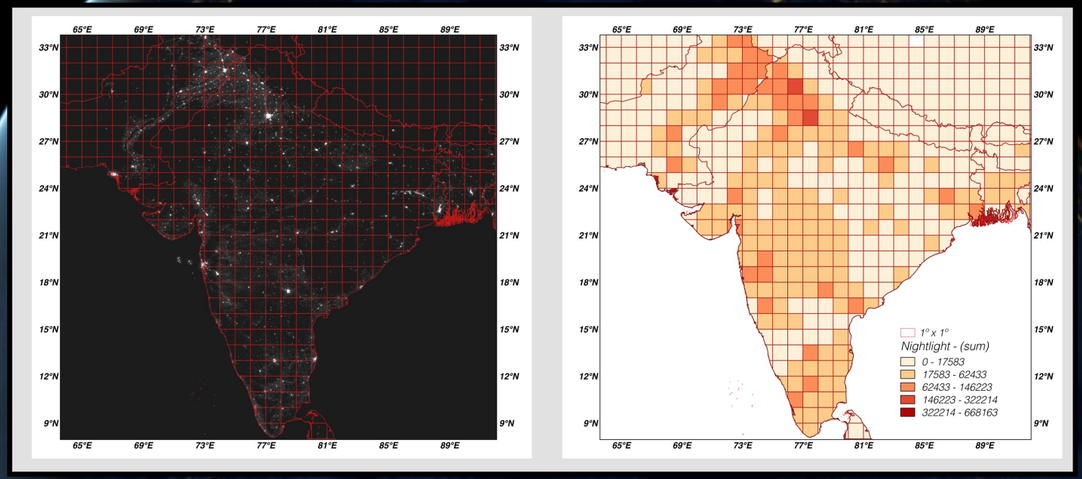
FIGURE 2.—Plots of various outcomes against longitude and latitude. See the text for a detailed description.

Dell (2010) Current African Countries & Slave Trades

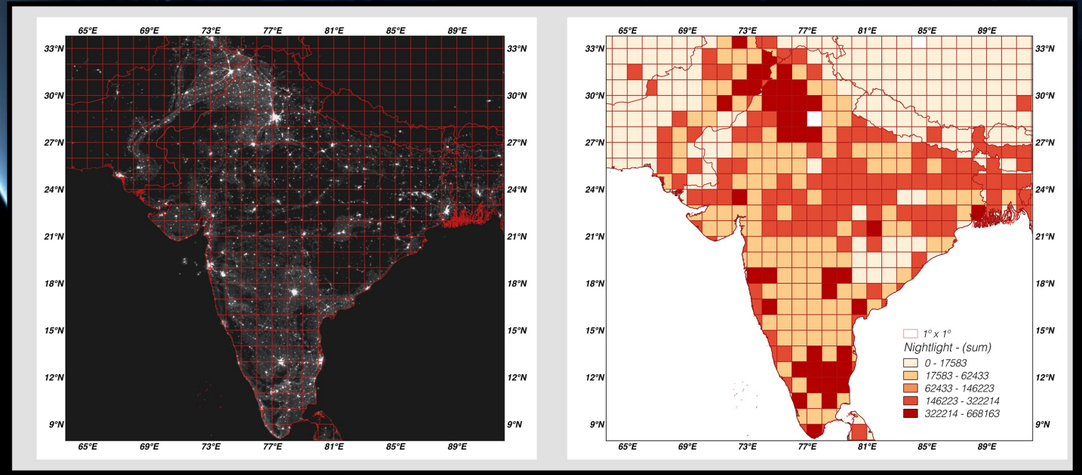
Colonial Institutions (Mita) → current outcomes

Exogenous unit of analysis

1992



2010



Montalvo and Reynal-Querol (2016)

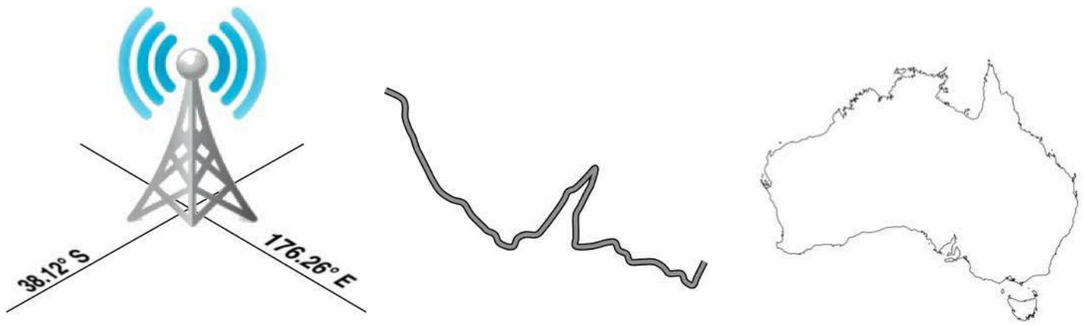
[3] Provides a new dimensions to gathering micro-level data

Improve your data collection

Most collection data software allows to gather GPS location (e.g. ODK Survey, Survey CTO, CAPI)

Cracking Geo-Spatial Datasets

"Geo-spatial" data



The terms "geo-spatial" refers to information that is located on the earth's surface using coordinates.

"Geo-spatial" data

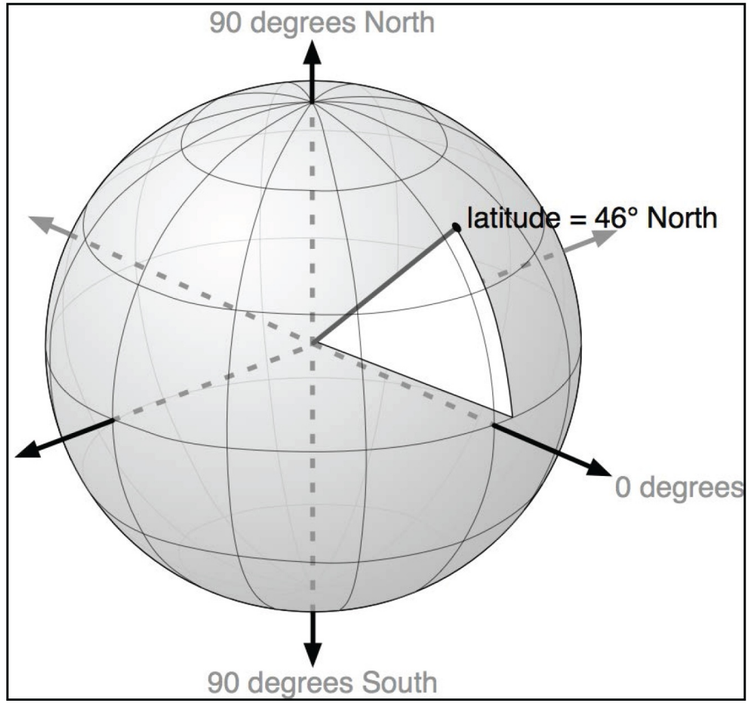
Location (e.g. coordinates)

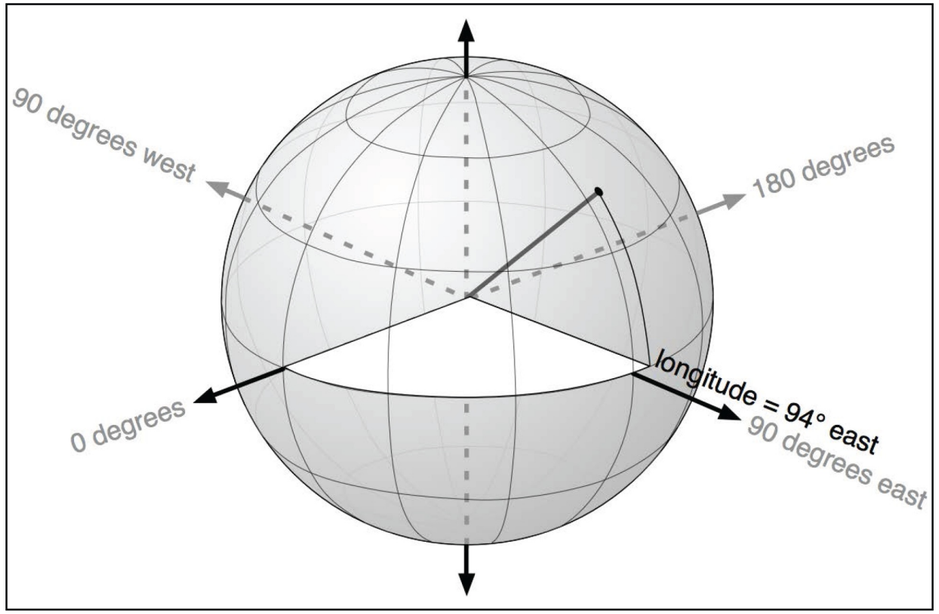
Type (e.g. line,area)

Attributes (e.g. temperature,...)

How can I represent those points, lines, and areas on the surface of the earth?

Geodetic Location

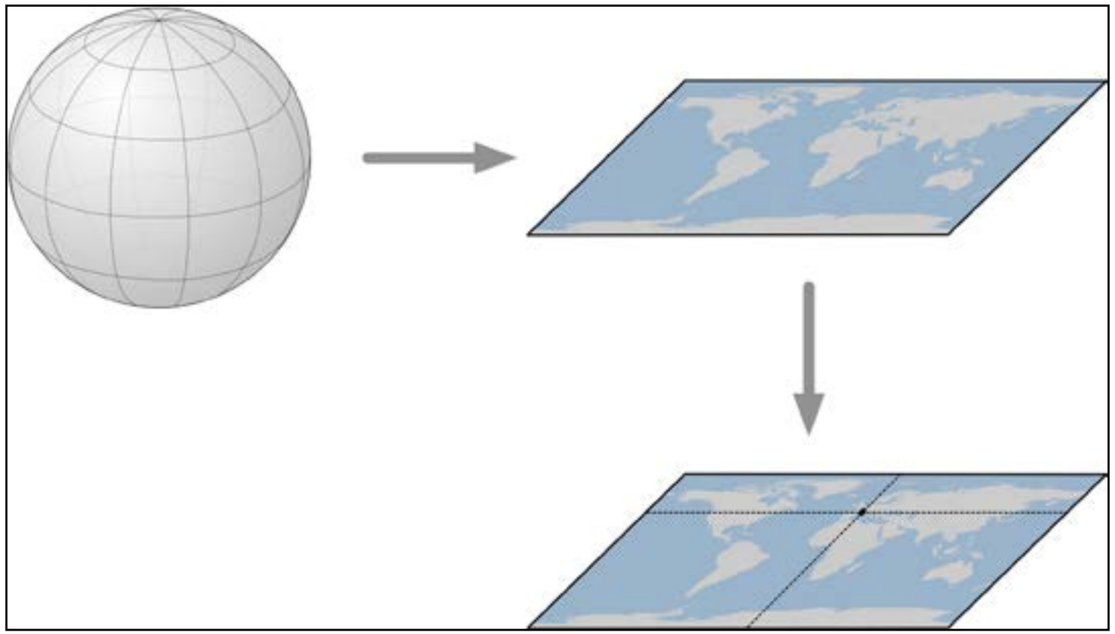




Latitude → north-south direction, relative to the equator

Longitude → east-west direction, relative to an arbitrary point (typically the location of the Royal Observatory in Greenwich - England)

... yet earth is not a perfect sphere, it is more like an oblate spheroid.



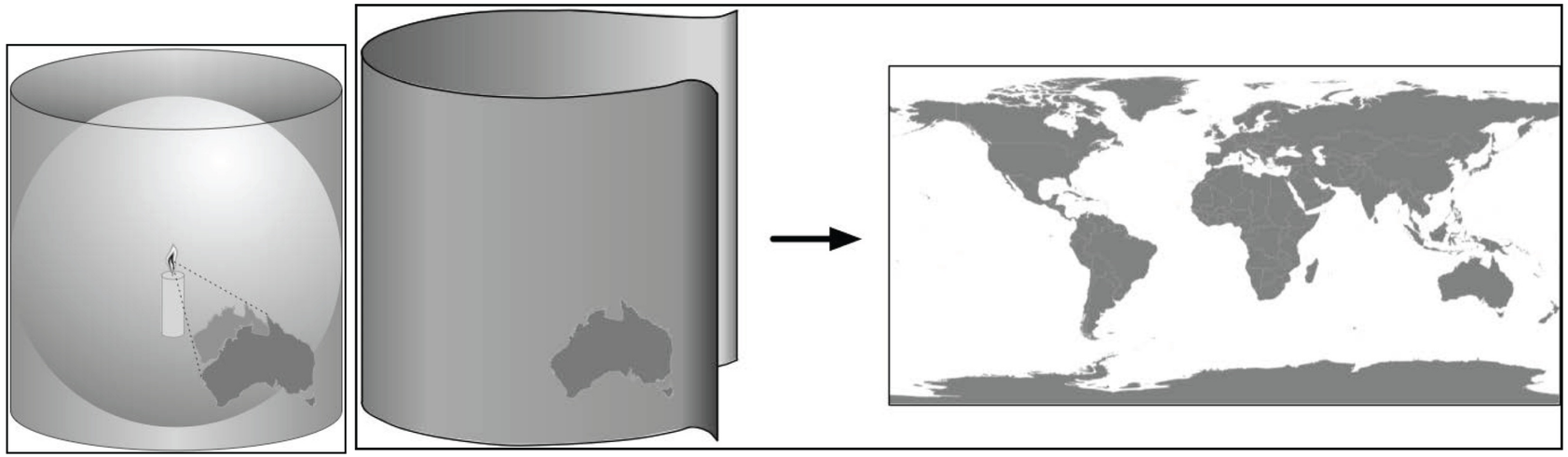
The mathematician Carl Gausse proved that it is mathematically impossible to project R^3 shape (e.g. sphere), onto a flat plane without introducing some sort of distortion.

What should we do?

... we need to choose a distortion that fits your needs

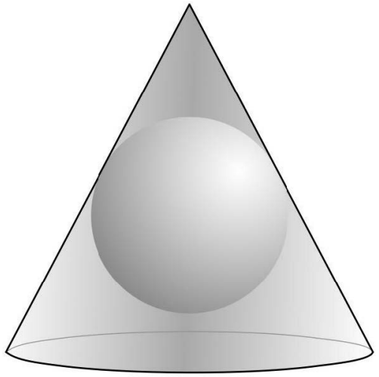
cylindrical
conical
azimuthal

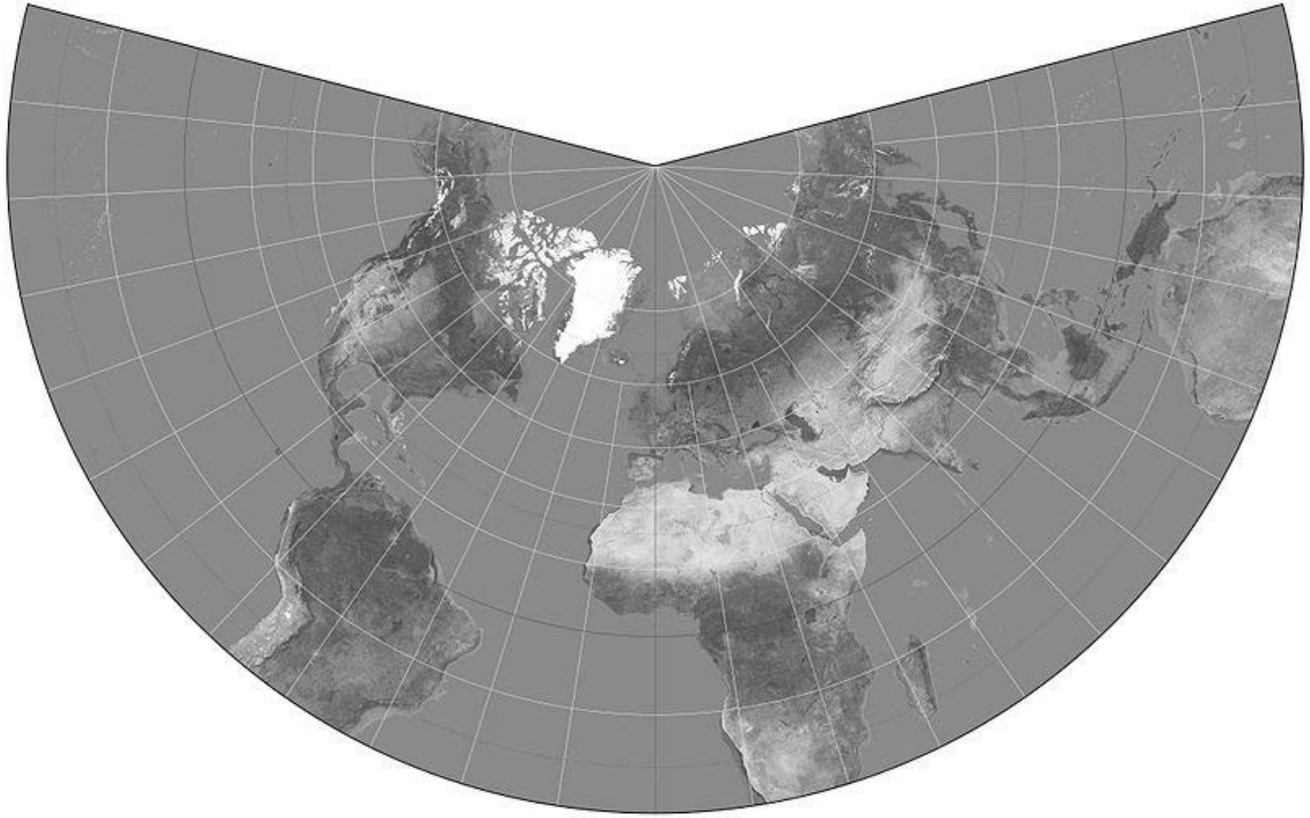
Cylindrical Projections



It include the Mercator Projection, the Equal-Area Cylindrical Projection, and the Universal Transverse Mercator Projection.

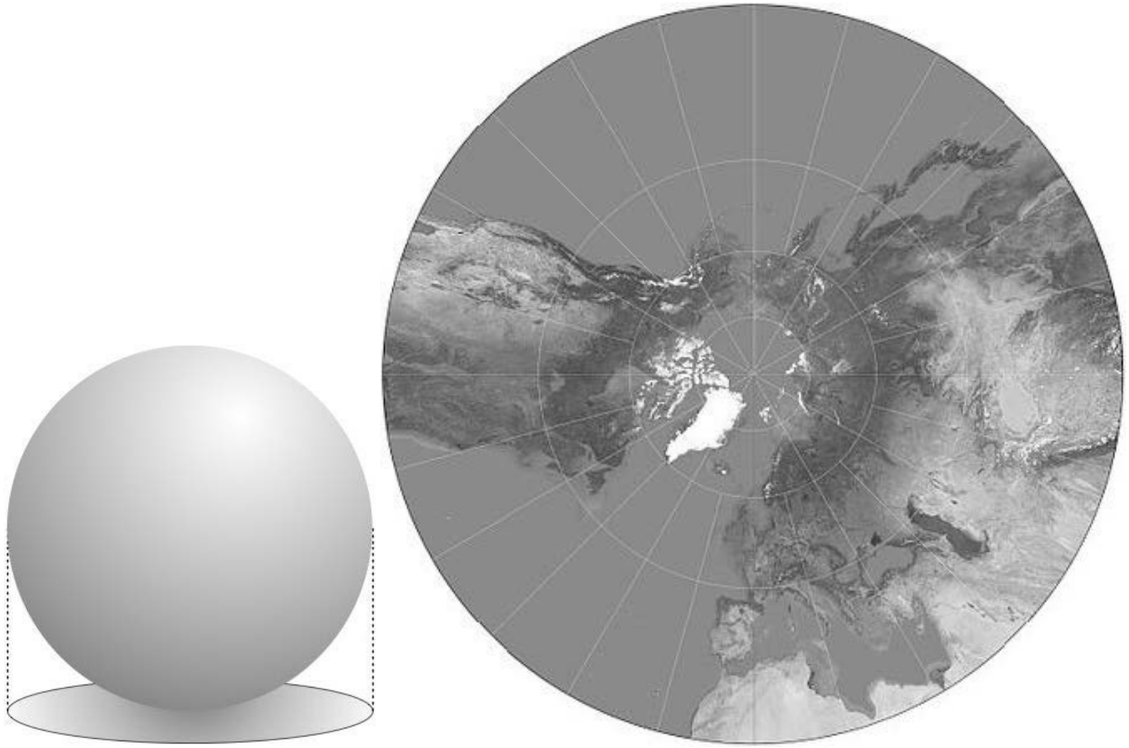
Conical Projections





It includes the Albers Equal-Area, the Lambert Conformal Conic Projection, and the Equidistant Projection.

Azimuthal Projections



It includes the Gnomonic Projection, the Lambert Equal-Area Azimuthal Projection, and the Orthographic Projection.

Coordinate System

Unprojected Coordinate System

(point directly on the earth's surface)

Not Projections $\rightarrow (x_0, y_0)$

Projected Coordinate System

(point on a two-dimensional map that represents the surface of the earth)

Use Projections $\rightarrow (\hat{x_0}, \hat{y_0})$

.. still something missing

yet, we said that the earth is not a perfect sphere

Datum is a mathematical model of the earth used to describe locations on the earth's surface. A datum consists of a set of reference points, often combined with a model of the shape of the earth.

Most common datum: NAD 27, NAD 83 and WGS 84

"Geo-spatial" data

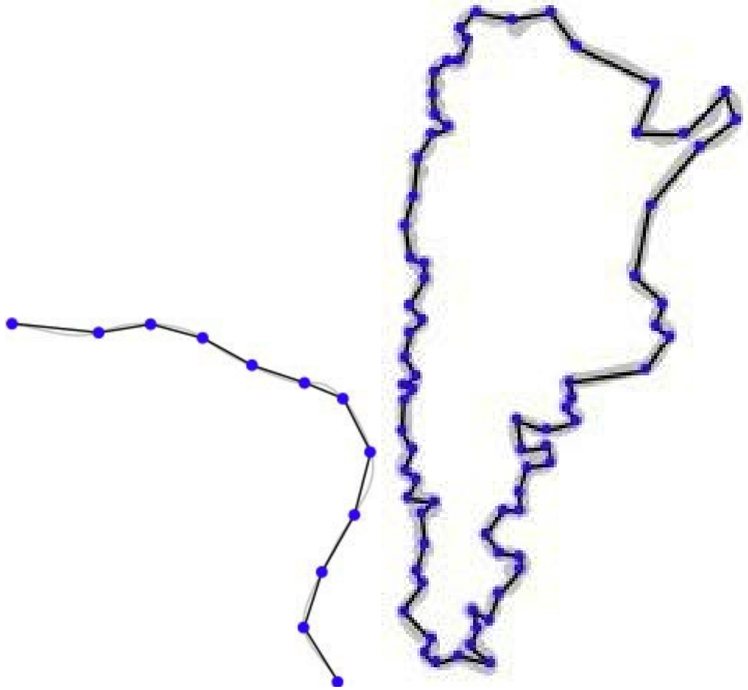
Location (e.g. coordinates)

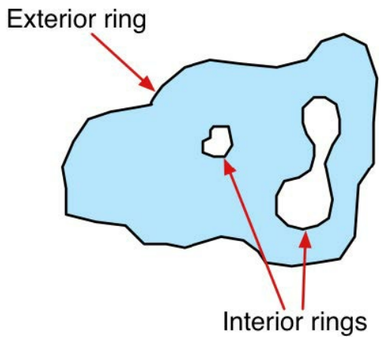
Type (e.g. line,area)

Attributes (e.g. temperature,...)

How can I represent those points, lines, and areas on the surface of the earth?

Vector Data





Point (lat, lon) Linestring Polygons Polygons

Raster data (images as data)

Geo-located map set

Pixel characteristics as main attribute

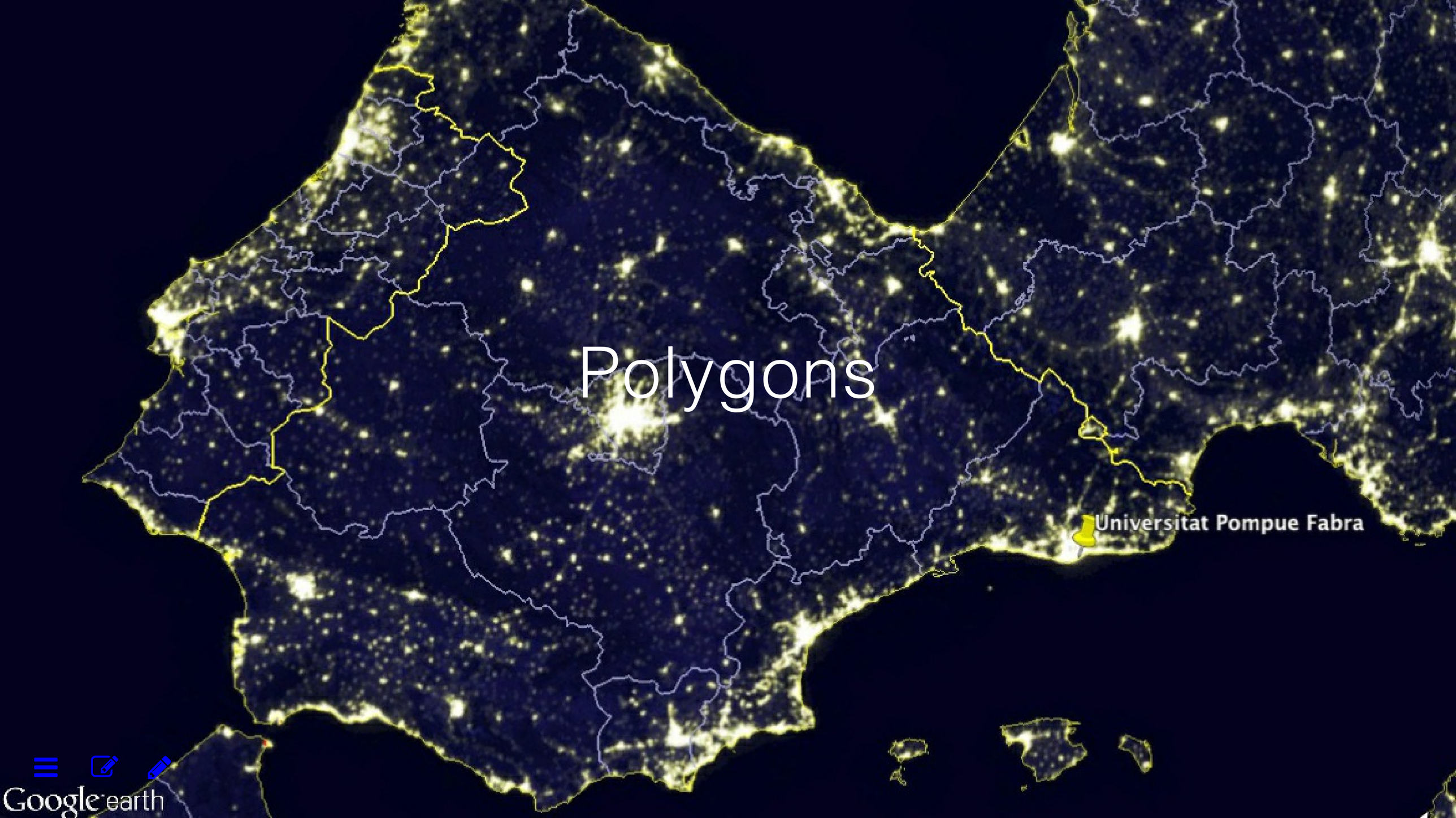
Classified (Remote sensitive data)

Raster Data

Raster Data

Polygons

Polygons



Linestring

E-82
E-80
IP5
AP-1
E5
E7
E-803
AP-15
E80
E9
AP-61
AP-68
E-7
R-5
R-2
R-3
AP-41
R-4
A9
E-9
AP-2
E-90
Universitat Pompue Fab
A-43
AP-36
E-5
E-15
AP-7

Point
UPF

"Geo-spatial" data

GIS data format

Getting geo-spatial dataset into a file (s)

Many formats available

Raster Data

DEM,geo-tiff,band based files
Vector Data

Shapelife, Simple features, TIGER/Line,...

micro-format (WKT,GeoJSON,...)

Shapefile

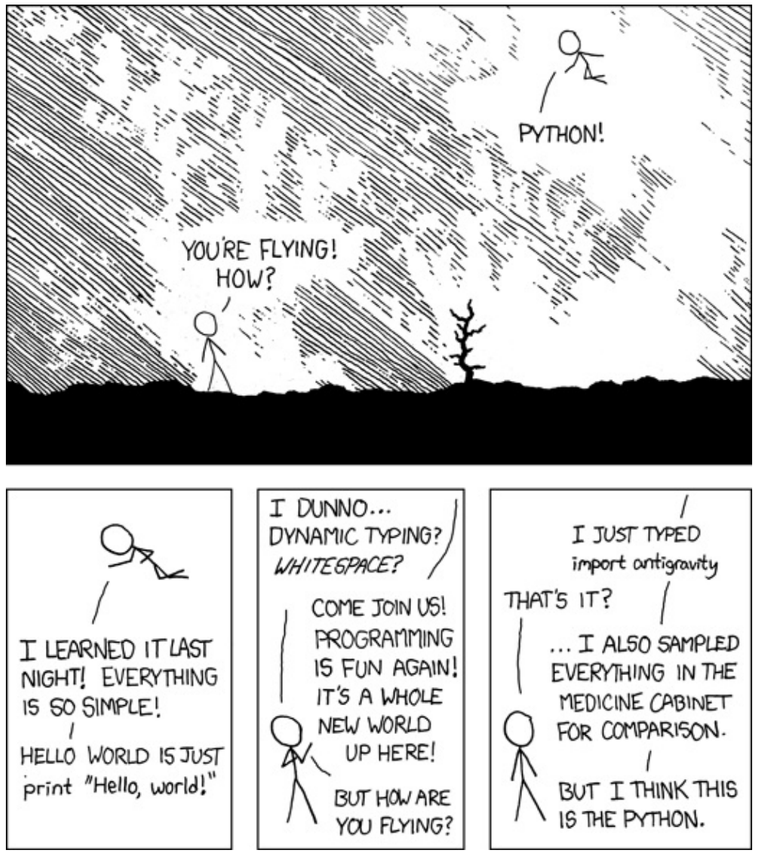
.shp it contains the geometry (type)
.shx fixed size
.dbf attributes (data base)

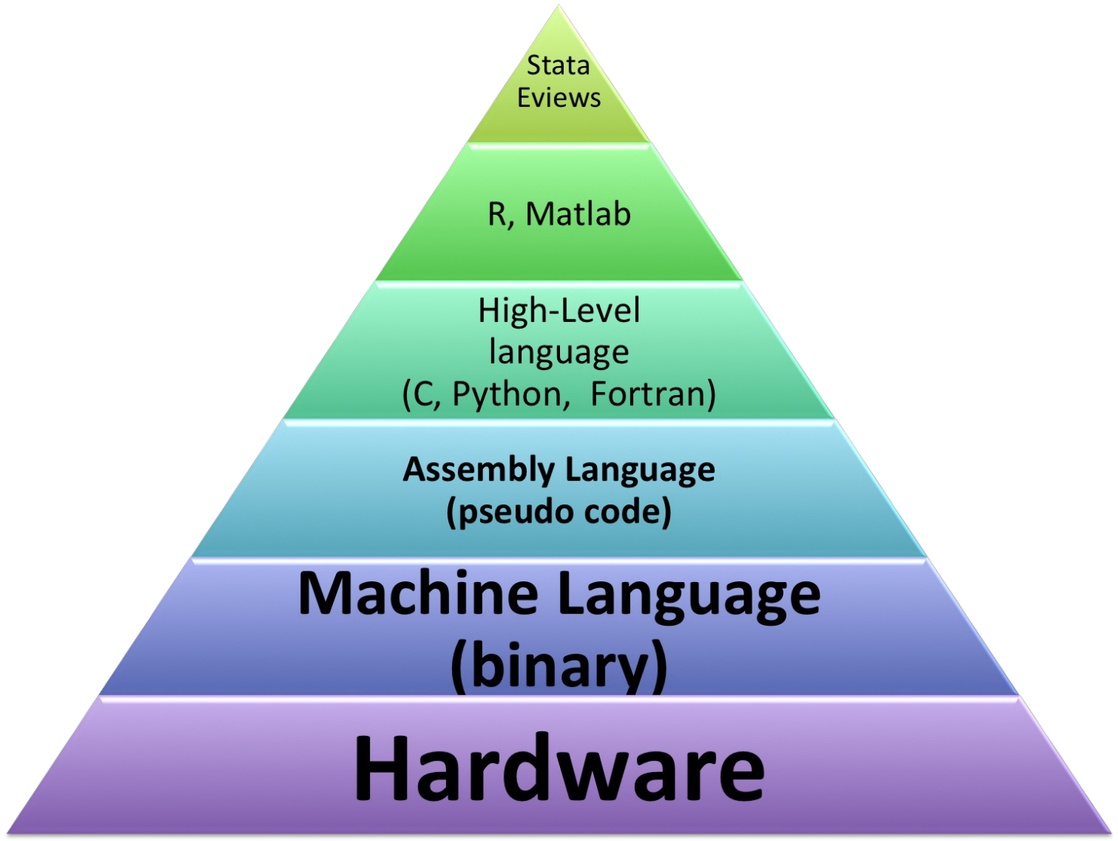
Other files such as: .prj(projections)...

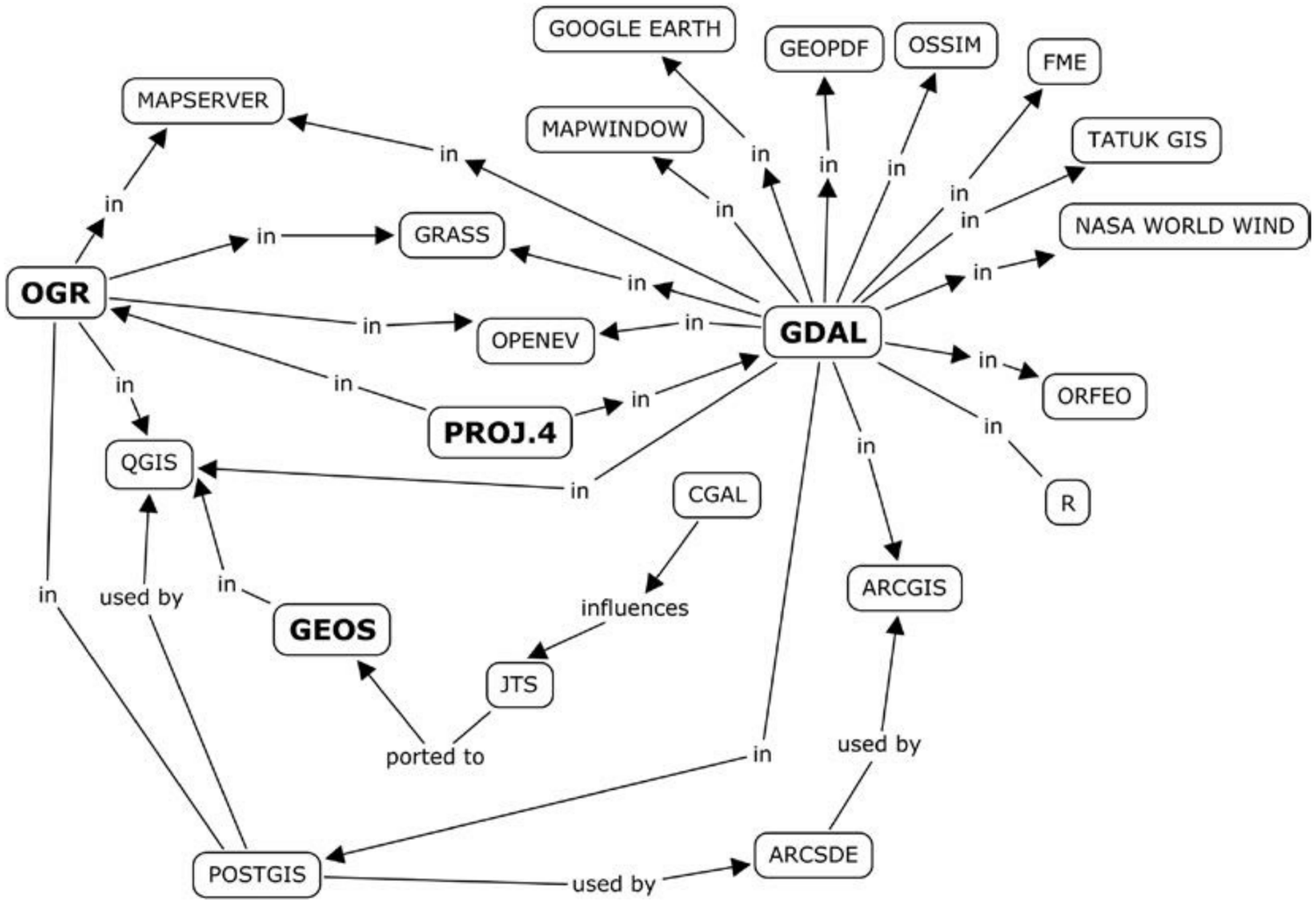
Same file name, yet different extensions

GIS tools

C, C++ and Python based solutions (objected-oriented syntaxis)







Source: Lawhead, 2013

ArcGIS

👎 Commercial Software

👍 Pre-defined python package - ArcPy (+ Model Builder)

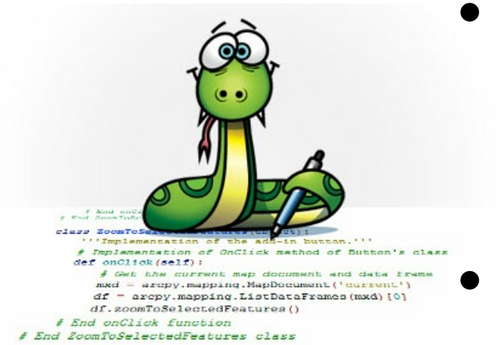
👎 Memory demanding; Windows only

QGIS

👍 Open source software; easy to create map images;
👍 Pre-defined python package - PyQGIS
👍compatible with any OS (even android)

More differences

Standalone solutions



- Complete control of the underlying algorithms, data, and execution (no black box)
- Integration wit others python packages
- Solution no-software depending (replication and collaborations)
- GIS analysis beyond the bottoms
- Integration with new high-performance environment (i.e. cluster,cloud)

Ready for starting?

– sample data sets