Empirical Techniques in Spatial Economics: Lecture 1: Introduction

This Version: August 25, 2021

noeldjohnson@mac.com

Introduction

About me

- Noel Johnson (GMU, Mercatus, CESifo)
- Email: njohnsoL@gmu.edu

Course Website

- **Blackboard**

Recommended material

- (new sf-paradigm) Lovelace, R., Nowosad, J., and J. Muenchow. 2018. Geocomputation with R. CRC Press.
- (for R in general) Grolemund, G. and H. Wickham. 2018. R for Data Science. O'Reilly.
- (useful intro to basic R and econometrics using R) Nick Huntington-Klein's videos at: `http://nickchk.com/videos.html#rstats`

Organization of Course

Open the Syllabus

The geospatial revolution I

Why use GIS in economics?

- Without GIS the unit of analysis is limited to countries, states, or places with administrative/ survey data
- With GIS the unit of analysis can be *any* level of spatial aggregation (cities, grid cells, ethnic homelands)

With satellite data, we can now

- access information difficult to obtain by other means
- at an unusually high spatial resolution
- with wide geographic coverage

The geospatial revolution II

Spatial identification techniques are a lot more credible

- Compare close-by places using differences-in-differences (e.g. distance to historical Protestant missions in SSA)
- Design new instruments based on space (e.g. distance to coast for slave exports in SSA)
- Use spatial discontinuities (e.g. split ethnic groups in SSA or borders of historical states in Viet Nam)

Use space as a transmission channel (e.g. oil prices and the road network in SSA).

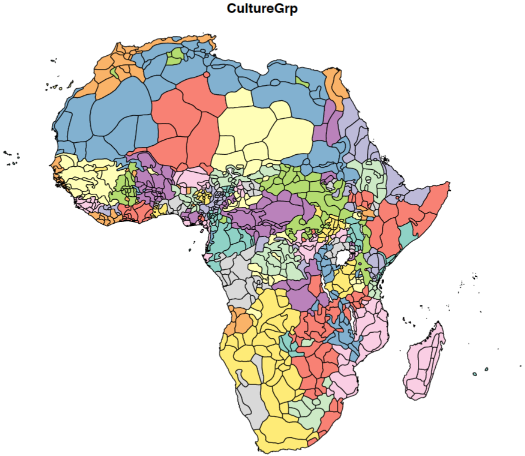
Ask which is the "right" level of spatial aggregation (e.g. cell, state, homeland) – although too little is done in this direction.

Some early applications: Murdock's map

Nunn (2008) The long-term effects of Africa's slave trades *AER*

- Studies the shadow of Africa's long history of slave exports
- Digitizes Murdock's map of ethnic homelands in Africa and matches these with detailed records on slaves and their ethnicities shipped from Africa
- Aggregates these data at the country level for each century from 1400 to 1900
- Uses sailing distances to slave ports as an instrument
- Shows that the African countries that are the poorest today are the ones from which the most slaves were taken

Murdock's ethnolinguistic map (1959)

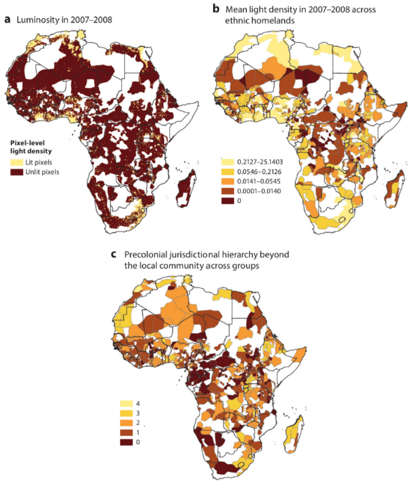


Some early applications: Night lights

Michalopoulos and Papaioannou (2013) Pre-colonial Ethnic Institutions and Contemporary African Development *Econometrica*

- Study how pre-colonial ethnic institutions shape comparative regional development within SSA countries
- Use Murdock's map together with information on pre-colonial centralization of ethnic groups from the *Ethnographic Atlas*
- Merge these data with nighttime lights from NOAA
- Examine the data at the country-group, pixel and adjacent-group level
- Show that regional development is higher in the homelands of ethnicities with centralized, hierarchical, pre-colonial political institutions.

Ethnic institutions and night lights



Where is the research frontier?

The lowest hanging fruits have been picked:

- X or Y with lights will only get you a decent publication today.

But, the rewards are still huge:

- An innovative story told with spatial data will be *credible*.
- Digitizing historical maps and other data can take you a long way (top 5s!).
- Geo-coding events data takes manpower but provides a huge public good (e.g. AidData).
- Some measurement issues remain (e.g. 'Top Lights') and "big" spatial data is an active research field.

In the words of Dave Donaldson and Adam Storeygard "it seems safe to say that the real excitement lies ahead."

Spatial basics

Sources of Spatial Data

Go to Your Website

Coordinate reference systems

All geographic data have a *coordinate reference system* (CRS). CRSs are often referred to as *geographic* or *projected*.

Why do we need a CRS?

- The Earth is an oblate spheroid but our maps are "flat".
- CRSs are ways to represent (parts of) Earth on a plane.
- There is no unique representation, many CRSs have been developed for different purposes.

The CRS defines the metrics for distances, areas and direction.

Merge data with different CRSs and your GPS may send you off a cliff (remember iMaps?).

Geographic CRSs

Geographic coordinate systems indicate locations using decimal degrees of longitude and latitude (e.g. Hannover is 9.7320° E, 52.3759° N).

Quick reminder:

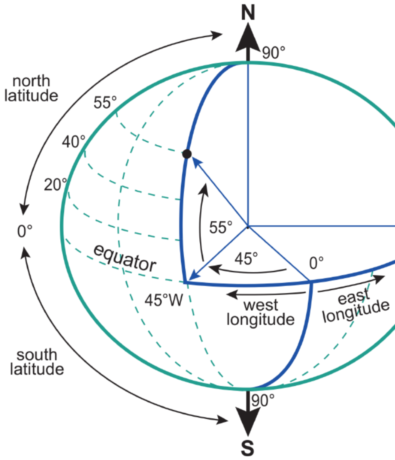
- *Longitude* is the angular distance from the Prime Meridian in East-West distance.
- *Latitude* is the angular distance North or South of the equator.

Careful:

- A 1° longitude arc at the equator is 111 km.
- A 1° longitude arc at 45° latitude is 79 km.

Distances and areas vary in meters when given in degrees!

Earth in angles



Map datums

A map datum is a mathematical representation of Earth.

If Earth were a perfect spheroid, we could use a simple spherical model. But, the equatorial radius is about 6,378 km and the polar radius is about 6,357 km.

Ellipsoidal models are defined by these two parameters. They "compress" the sphere a bit in the polar direction.

There are many map datums, few geocentric, many local:

- World Geodetic System 1984 (WGS84)
- North American Datum 1983 (NAD83)
- European Datum 1950 (ED50)

Most modern data you will encounter uses WGS84 (e.g. any GPS).

Differences in datums

Shouldn't all datums be the same?

- UK maps use the OSGB36 datum. Go to Greenwich observatory on the map and you will see longitude of $0°$.
- Now check your GPS which uses WGS84 and learn that you are at $0.0077°$ W, $51.4826°$ N ($\sim$ 100m off)

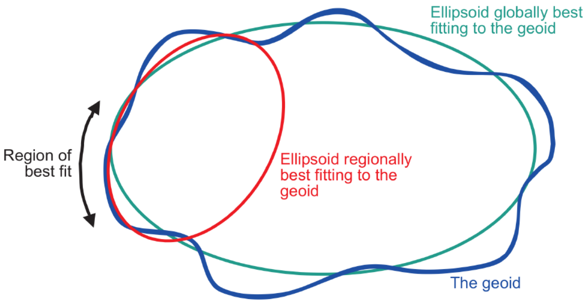
What happened?

- OSGB36 obeys the Greenwich line since 1884 which uses astronomical longitude (w/ gravity anomalies)
- WGS84 uses geodetic latitude (w/o gravity anomalies)

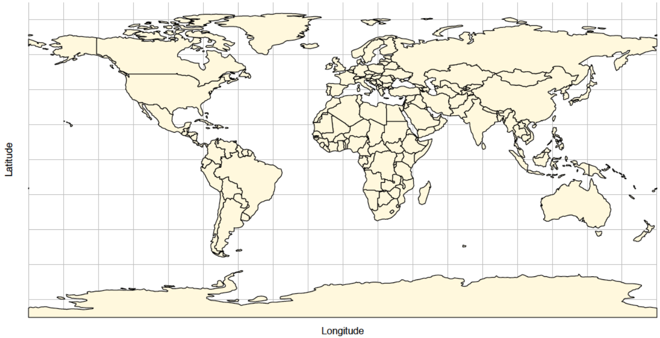
Why does this matter?

- Wrong datum, everything gets shifted (e.g. in historical maps)

Local datums illustrated



Geographic CRS in WGS84



Projected CRSs

The Earth surface is projected on a flat plane. All projected CRSs are based on a geographic CRS and then convert the 3d surface of the Earth into the x and y values of a projected CRS.

All projections induce *distortions* of either distance, area or shape ("can crush an orange in many ways"). *No flat map can preserve distances, areas and direction at the same time.*

There are planar, conic and cylindrical projections.

- Our earlier *image* of the geographic CRS in WGS84 uses an equidistant cylindrical (or plate carrée) projection.
- Most projections will provide locations in meters from some origin on the projected surface.

Some projections humor: `http://xkcd.com/977/`

Distortion properties

Conformal map projection:

- The angles between lines in the map are identical to the angles between the original lines on the curved reference surface.
- Angles with short sides and small shapes are displayed correctly.

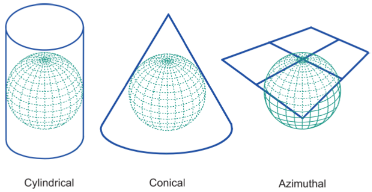
Equal-area (equivalent) map projection:

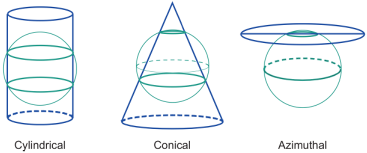
- The areas in the map are identical to the areas on the curved reference surface (taking into account the map scale).
- Areas are represented correctly on the map.

Equidistant map projection:

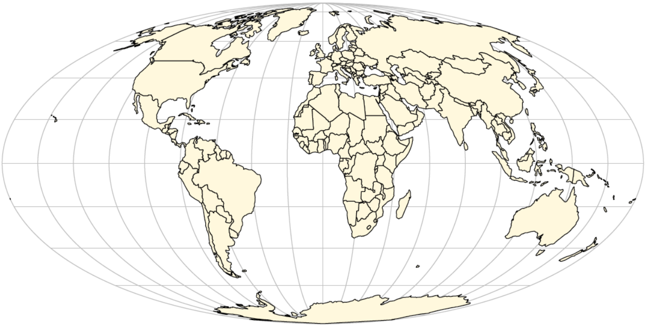
- The length of particular lines in the map are the same as the length of the original lines on the curved reference surface.
- Distances along these lines are correct on the map.

Planar, conic and cylindrical projections





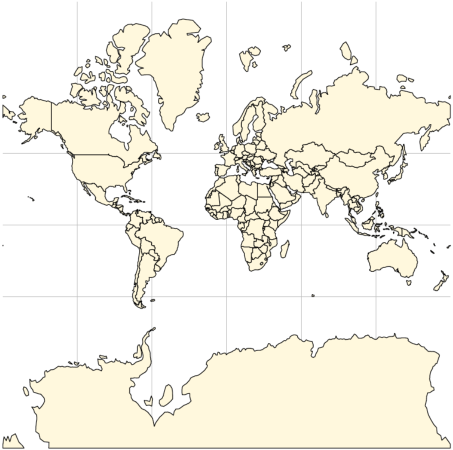
Mollweide equal-area projection



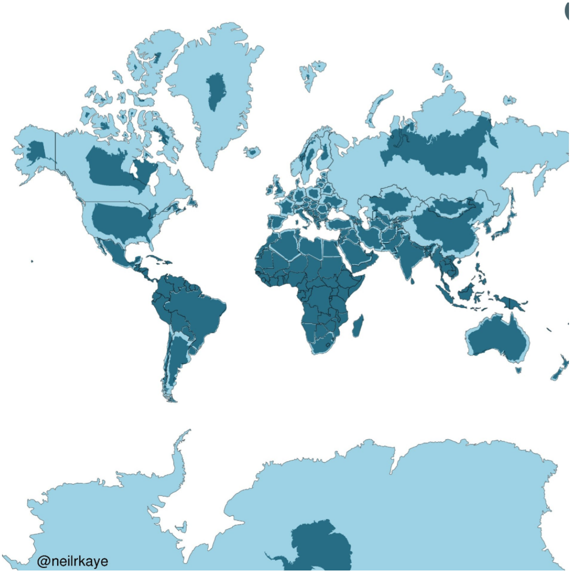
Lambert azimuthal equal-area projection

Azimuthal equal-distant projection

Web Mercator



Perhaps someone should brief the President on the Mercator projection...



Universal Transverse Mercator (UTM)

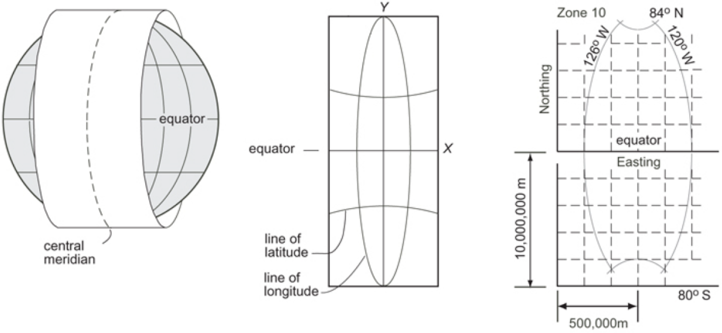
Peal the Earth into 6° zones. A circle has 360°, so there are 60 UTM zones on Earth. UTM numbers them 0-60 (equivalent to longitude) and defines the regions (north and south).

Each zone has its own central meridian. Zones 1N and 1S start at 180° W. The limits of each zone are 84° N and 80° S. The polar regions use the Universal Polar Stereographic coordinate system.

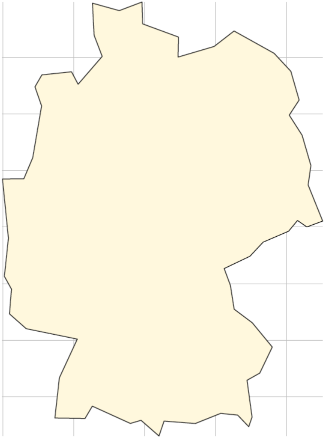
The coordinate system grid for each zone is projected individually using the Mercator projection (a cylindrical projection).

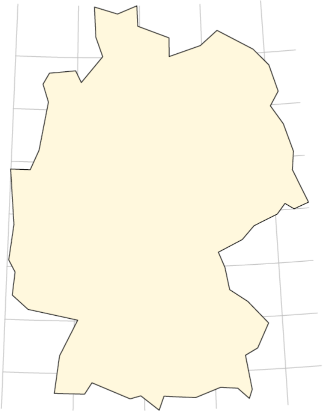
Area and distance distortions are very small with a 6° zone (less than 0.1%) and shapes are accurately displayed. Almost all topographic maps use the UTM projection.

UTM slicing illustrated



Germany in WGS84 and UTM 32





Proj4 strings and EPSG codes for WGS84

Proj4 strings contain all relevant information on the chosen CRS and can be used to create custom reference systems.

Our world map has the following Proj4 string:

```
+proj=longlat +datum=WGS84 +no_defs
```

- `proj=longlat`: the data are in a geographic (latitude and longitude) coordinate system
- `datum=WGS84`: the datum is WGS84
- `+no_defs`: no other defaults are used

This Proj4 string has an EPSG code of 4326 – a standard defined by the European Petroleum Survey Group (EPSG).

You can look up these strings and codes here: `http://epsg.io/`

Which reference sytem to use?

Depends on what you want to do. My personal recommendation:

- *Geographic.* For all computations and display purposes where you do not care about area or distances.
- *Mollweide or Lambert azimuthal equal-area.* For global area calculations.
- *Azimuthal equidistant.* For accurate straight-line distance between a point and the center point of the local projection.
- *Universal Transverse Mercator.* For local area and distance computations. Germany spans about 1.5 UTM Zones (32+33).

Advanced reference materials

On sensors and satellites:

- Harris, A. (2013). Satellite orbits and sensor resolution. In Thermal Remote Sensing of Active Volcanoes: A User's Manual (pp. 113–152). Chapter, Cambridge: Cambridge University Press. `http://doi.org/10.1017/CBO9781139029346.005`

On projections:

- Environmental Systems Research Institute. (2000). Understanding Map Projections.

On distances:

- Banerjee, S. (2005). Geodetic Distance Computations in Spatial Modeling. Biometrics 61, 617–625.