

Applied Spatial Data Analysis (ASDA)

Course Information

Lecturers

Course Outline

Format

Philosophy

Schedule

See course outline

Course Information

[Lecturers](#)[Course Outline](#)[Format](#)[Philosophy](#)[Schedule](#)

Course outline and related documents available on VULA under resources

Course Information

[Lecturers](#)[Course Outline](#)[**Format**](#)[Philosophy](#)[Schedule](#)

Face-to-face lectures

- Wednesdays 09:00 - 11:00 (Honours Lab)

Course Information

[Lecturers](#)[Course Outline](#)[Format](#)[Philosophy](#)[Schedule](#)

The idea is for the course to be very applied and hands-on, and based around worked examples and case studies.

The course is entirely in R, and the goal of the course is as much to cover the theory of ASDA – which is after all relatively simple.

You are going to learn some of the key tools and frameworks for dealing with spatial data.

Emphasis will be on:

- how to choose an appropriate analysis
- applying the available method(s) to a data set in R
- interpreting the results

Course Information

Lecturers
TBC

Course Outline

Format

Philosophy

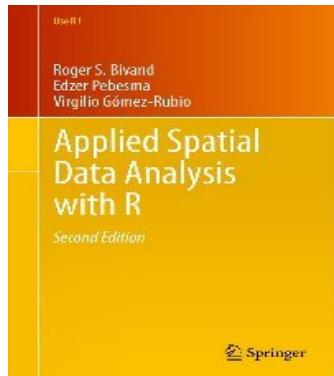
Schedule

Course Material

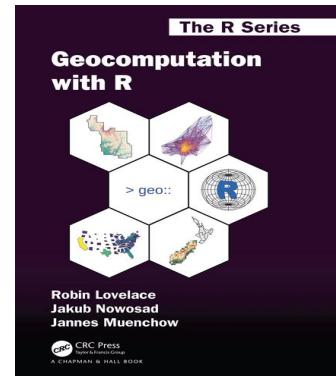
Textbooks

Other Resources

Lecture notes/slides will be provided in addition to some videos. Lectures will be primarily based on material presented in the following textbooks::



E-copy Available via UCT's Library Website
Codes and Datasets can be obtained from:
<https://asdar-book.org/>



Online version of this book is free to read at:
<https://bookdown.org/robinlovelace/geocompr/>

Course Material

Texbooks

Other Resources

- Anselin, Luc, (1999). Spatial Econometrics: Methods and Models. Luwer Academic Publishers.
- Arbia, Giuseppe, (2014). A Primer for Spatial Econometrics : With Applications in R. Palgrave Macmillan UK. (available at UCT library online resource)
- Baddeley, A., Rubak, E., Turner, R. (2016). Spatial Point Patterns Methodology and Applications with R. Chapman & Hall/CRC *
https://azaieznotesblog.files.wordpress.com/2017/08/spatial-point-patterns_chapman-hall-crc-2016.pdf
- Elhorst, Paul. J. (2014). Spatial Econometrics : From Cross-Sectional Data to Spatial Panels. Springer Berlin
- Gelfand, A.E., Diggle, P., Guttorp, P., Fuentes, M. and Fitzmaurice, G. (2010). Handbook of Spatial Statistics. Chapman & Hall/CRC. (available at UCT library online resource) Brilliant resource to have. Part 4 is dedicated to SPPs*
- Haining, Robert, (2004). Spatial Data Analysis, Theory and Practice. Cambridge University Press. *
- Wiegand, T., Moloney, K. A. (2014). Handbook of Spatial Point-Pattern Analysis in Ecology. Chapman & Hall/CRC (available at UCT library online resource)*
- *Geodacenter*

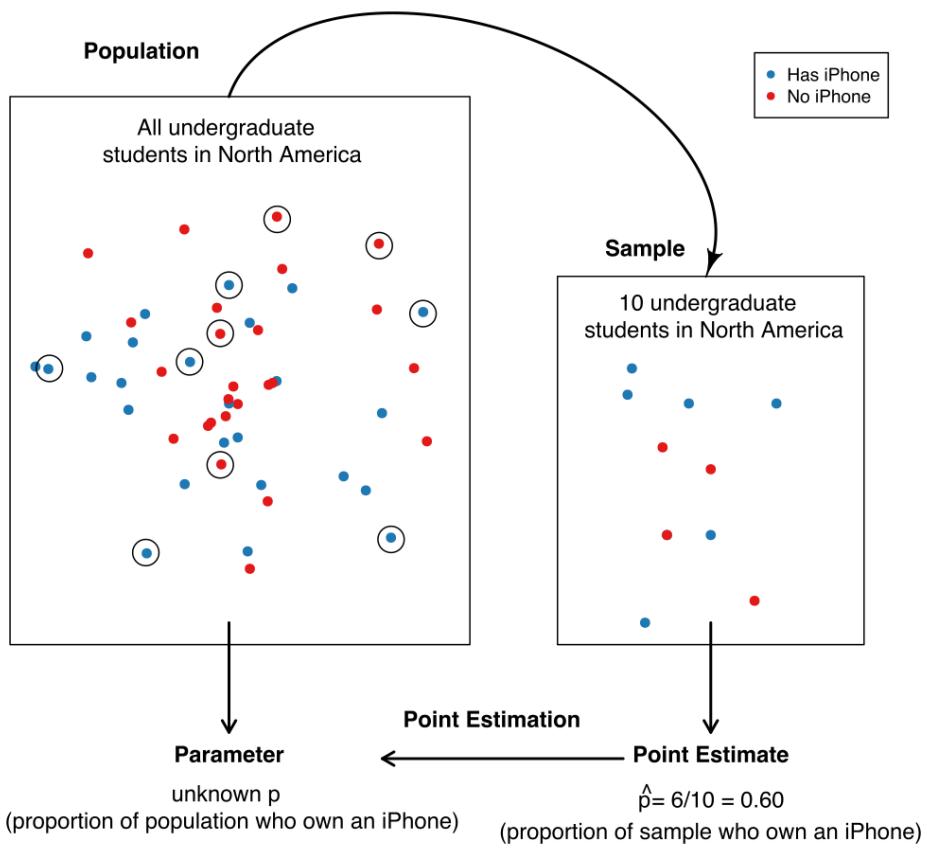
Classical Data Analysis

Data Data... Data... Requirements



Classical Data Analysis

Data Data... Data... Requirements



Classical Data Analysis

Data Data... Data...

Requirements

id	status	wt	ht
101	a	100	5.3
101	b	200	5.2
101	c	100	5.0
101	d	105	5.1
202	a	20	4.3
202	b	22	4.2
202	c	25	4.1

Classical Data Analysis

Data Data... Data...

Requirements

- Identifiability
- Independence
- Normality
- Linearity
- Stationary
- e.t.c.

Spatial Data

What is? Examples E.g.(2) E.g.(2) Why? ?(2) ?(3)

- Data about the locations and shapes of geographic features and the relationships between them. These are typically stored as coordinates and topology
 - Coordinates are usually two-dimensional (x,y) but can also be three-dimensional (x,y,z)
 - A coordinate reference system (CRS) is typically attached to describe which location on Earth the coordinates refer to
- Data can be:
 - observational or experimental,
 - continuous or categorical,
 - univariate or multivariate

Spatial Data

What is?

Examples

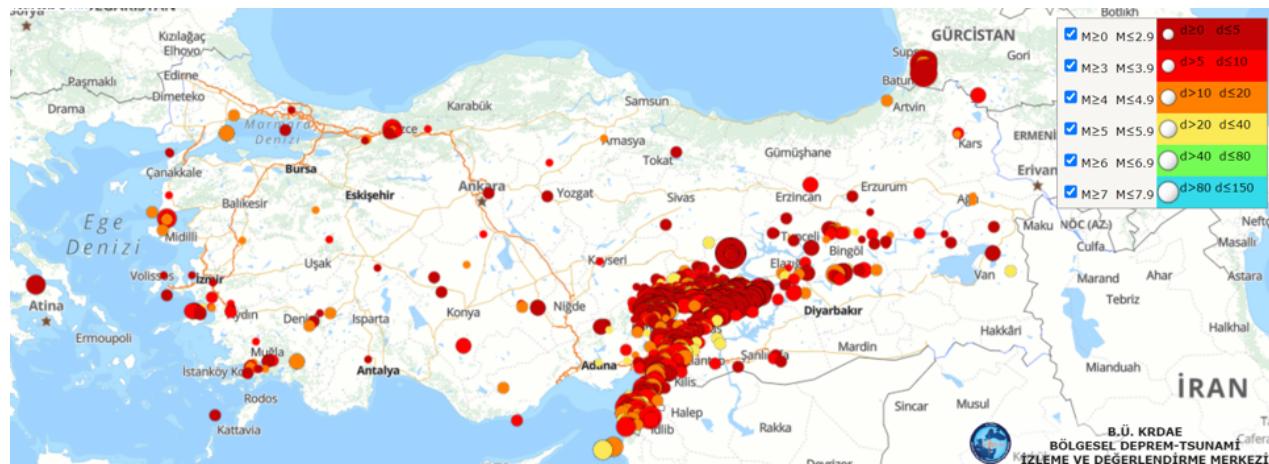
E.g.(2)

E.g.(2)

Why?

?(2) ?(3)

Consider the locations of various earthquakes that happened in/around Turkey on 6 Feb 2023 and onwards!



- We could list the coordinates for all known earthquakes as pairs of longitude/latitude decimal degree values with respect to the prime meridian at Greenwich and zero latitude at the equator!
- The World Geodetic System (WGS84) is a frequently used representation of the Earth!
- If we also have the magnitude of the last measured earthquake, this information is called an attribute: it is non-spatial in itself, but this attribute information exists for each spatial entity (earthquake)!
- Another attribute is the date and time of the last measured earthquake!
- Another attribute is the depth of the earthquake!

Spatial Data

What is?

Examples

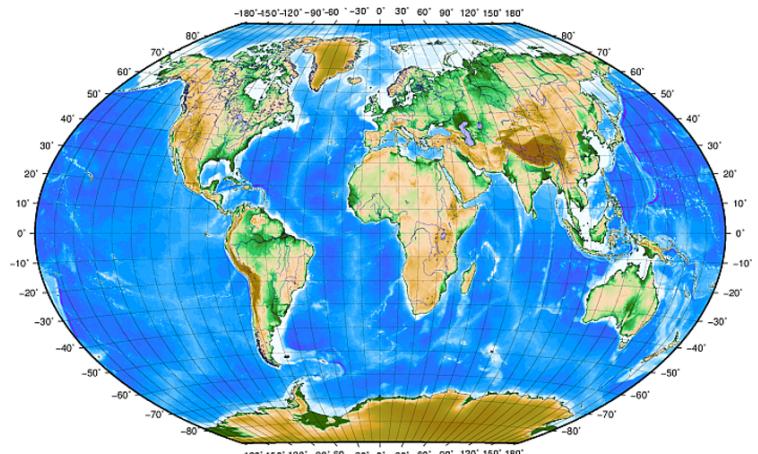
E.g.(2)

E.g.(2)

Why?

?(2)

?(3)



- Lines of latitude run parallel to the equator and divide the earth into 180 equally spaced sections from North to South (or South to North). The reference line for latitude is the equator and each hemisphere is divided into ninety sections, each representing one degree of latitude
- Lines of longitude, on the other hand, do not stand up so well to the standard of uniformity. Lines of longitude run perpendicular to the equator and converge at the poles. The reference line for longitude (the prime meridian) runs from the North pole to the South pole through Greenwich, England

Spatial Data

What is? Examples E.g.(2) E.g.(2) Why? ?(2) ?(3)

- Administrative data
- Animal movement
- Cellphones and GPS enabled devices
- Remote sensing (i.e. satellites)
- Social media
- Surveys
- Transportation
- Climate
- etc.

Spatial Data

What is? Examples E.g.(2) E.g.(2) **Why?** ?(2) ?(3)

Tobler's first law of geography

"everything is related to everything else, but near things are more related than distant things" Waldo R. Tobler (Tobler 1970)

With spatial data

- Spatial effects present
 - **Spatial autocorrelation / dependence** systematic spatial variation that results in observable clusters or a systematic spatial pattern
 - Mechanisms that can cause spatial autocorrelation (Voss et al. 2004): Feedback, Grouping forces, Grouping responses, Nuisance autocorrelation
 - **Spatial heterogeneity:** structural relations that vary over space
 - **Spatial nonstationarity:** variations in the relationship between an outcome variable and a set of predictor variables across space
- Independence assumption not valid
- Omitted variable bias

Spatial Data

What is? Examples E.g.(2) E.g.(2) Why? ?(2) ?(3)

Potential insights provided by the examination of the locational attributes of data

- Distance
- Proximity
- Adjacency
- Neighbourhood
- Interaction

Spatial Data

What is? Examples E.g.(2) E.g.(2) Why? ?(2) ?(3)

Bivand et. al (2013):

Spatial data analysis is concerned with questions about the hypothetical processes that generate the observed data

Possible questions that may arise include the following:

- Does the spatial patterning of disease incidences give rise to the conclusion that they are clustered, and if so, are the clusters found?
- Given a number of observed soil samples, which part of a study area is polluted?
- Given scattered air quality measurements, how many people are exposed to high levels of black smoke or particulate matter (e.g. PM 10), and where do they live?
- Do governments tend to compare their policies with those of their neighbours, or do they behave independently?

Spatial Data ...

[CRS](#) [CRS Types](#) [Map projections](#) [projs \(2\)](#) [projs \(3\)](#)

- A coordinate reference system (CRS) is a coordinate-based system that defines how the 2D projected map relates to real positions on the earth's surface
- A CRS also provides information about the shape of the earth, the size and position of the grid used to locate positions, and the orientation of that grid
- The CRS is important for GIS software because it provides a common language for the software to interpret and display spatial data

Spatial Data ...

CRS

CRS Types

Map projections

projs (2)

projs (3)

Projected CRS: uses a two-dimensional Cartesian coordinate system to represent the surface of the earth.

Geographic CRS: uses a model of the earth's surface to describe the location of positions.

The use of Geographic Coordinate Reference Systems is very common. They use degrees of latitude and longitude and sometimes also a height value to describe a location on the earth's surface. The most popular is called WGS 84.

- WGS84 is defined and maintained by the United States National Geospatial-Intelligence Agency (NGA). It is consistent, to about 1cm, with the International Terrestrial Reference Frame (ITRF).

Spatial Data ...

[CRS](#)[CRS Types](#)[**Map projections**](#)[projs \(2\)](#)[projs \(3\)](#)

- Map projections try to transform the surface of the earth from its spherical shape (3D) to a planar shape (2D).

Spatial Data ...

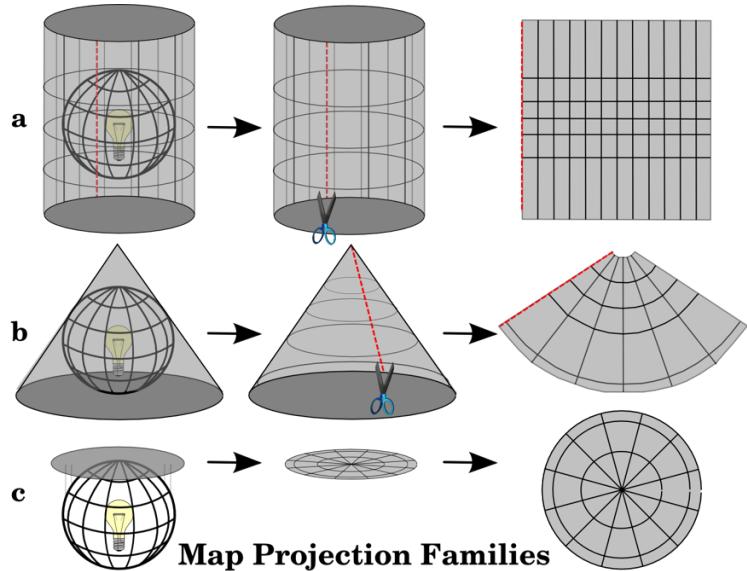
CRS

CRS Types

Map projections

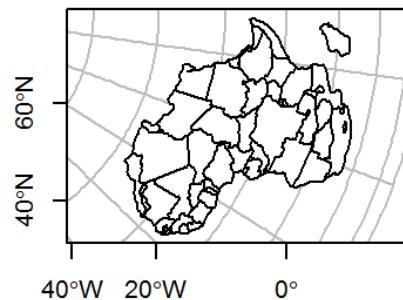
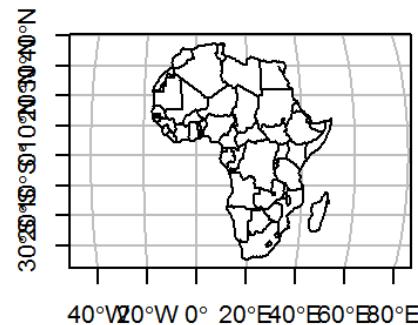
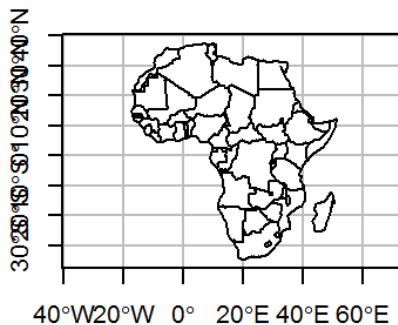
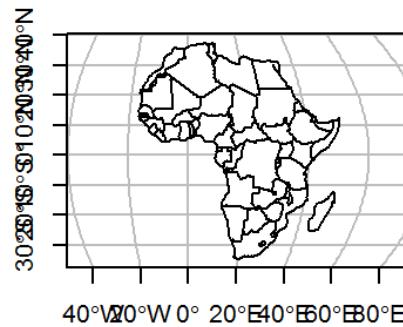
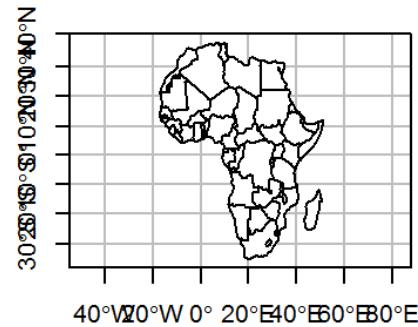
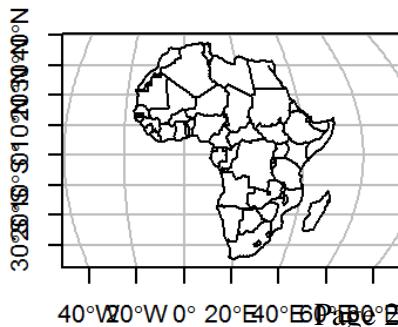
projs (2)

projs (3)



- The process of creating map projections is best illustrated by positioning a light source inside a transparent globe on which opaque earth features are placed.
- Then project the feature outlines onto a two-dimensional flat piece of paper.
- Different ways of projecting can be produced by surrounding the globe in a cylindrical fashion, as a cone, or even as a flat surface. Each of these methods produces what is called a map projection family.

Spatial Data ...

[CRS](#)[CRS Types](#)[Map projections](#)[projs \(2\)](#)[projs \(3\)](#)**Albers Equal Area****Robinson****Plate Carree****Van Der Grinten****Gall-Peters****Goode Homolosine**

Spatial Data ...

Type	Main Goal	Vectors(1)	(2)	Raster(1)	(2)	(3)	(4)
------	-----------	------------	-----	-----------	-----	-----	-----

Various classifications

- Vector
- Raster
- Tabular

Alternatively

- Geostatistical
- Point process
- Areal

We can go from one data type to another depending on the objective of the analysis

Spatial Data ...

Type	Main Goal	Vectors(1)	(2)	Raster(1)	(2)	(3)	(4)
------	-----------	------------	-----	-----------	-----	-----	-----

For each type of data, we wish to :

- Summarise / Visualize the data (mapping the data to show things cannot otherwise - Bivand et. al, 2013)
- Evaluate and describe spatial pattern
 - 1st order properties (mean, intensity)
 - 2nd order properties (covariance/semivariance etc.)
- Simulate spatial data

Spatial Data ...

Type	Main Goal	Vectors(1)	(2)	Raster(1)	(2)	(3)	(4)
------	-----------	------------	-----	-----------	-----	-----	-----

Geographic data (points, lines, polygons).

- Typically come in Shapefiles (*.shp*, *.shx*, *.prj* and *.dbf*)
- Other formats *.gpkg*, *.kml*, *.json* or *.geojson*, *.gpx*

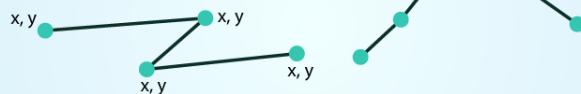
POINTS: Individual **x, y** locations.

ex: Center point of plot locations, tower locations, sampling locations.



LINES: Composed of many (at least 2) vertices, or points, that are connected.

ex: Roads and streams.



POLYGONS: 3 or more vertices that are connected and **closed**.

ex: Building boundaries and lakes.

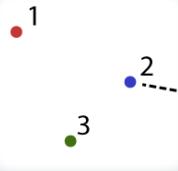


neon®

Spatial Data ...

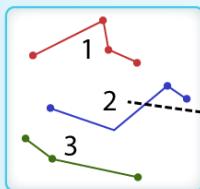
Type Main Goal Vectors(1) (2) Raster(1) (2) (3) (4)

Example Attributes for Point Data



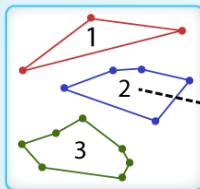
ID	Plot Size	Type	VegClass
1	40	Vegetation	Conifer
2	20	Vegetation	Deciduous
3	40	Vegetation	Conifer

Example Attributes for Line Data



ID	Type	Status	Maintenance
1	Road	Open	Year Round
2	Dirt Trail	Open	Summer
3	Road	Closed	Year Round

Example Attributes for Polygon Data



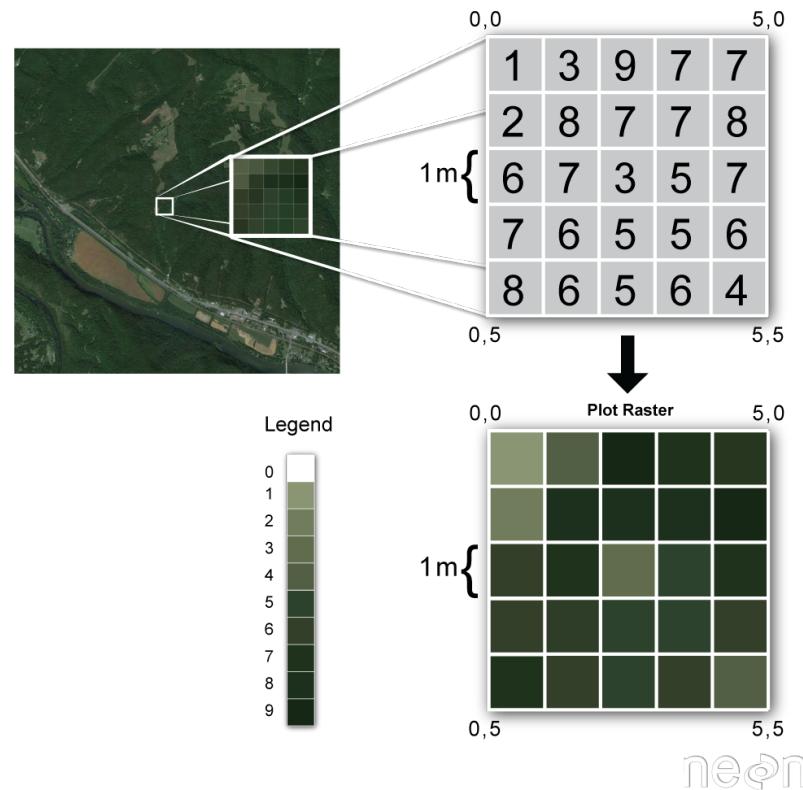
ID	Type	Class	Status
1	Herbaceous	Grassland	Protected
2	Herbaceous	Pasture	Open
3	Herbaceous / Woody	Grassland	Protected

neqr

Spatial Data ...

Type	Main Goal	Vectors(1)	(2)	Raster(1)	(2)	(3)	(4)
------	-----------	------------	-----	-----------	-----	-----	-----

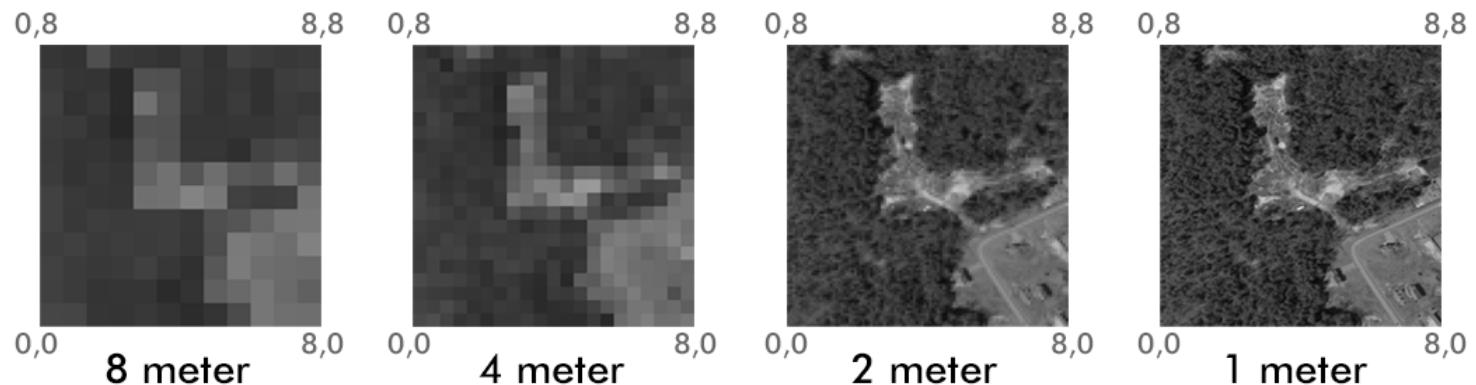
Pixelated data that can be added to a map. Typical file formats .jpg, .tif



Spatial Data ...

Type	Main Goal	Vectors(1)	(2)	Raster(1)	(2)	(3)	(4)
------	-----------	------------	-----	-----------	-----	-----	-----

Raster over the same extent, at 4 different resolutions



NEON

Spatial Data ...

Type

Main Goal

Vectors(1)

(2)

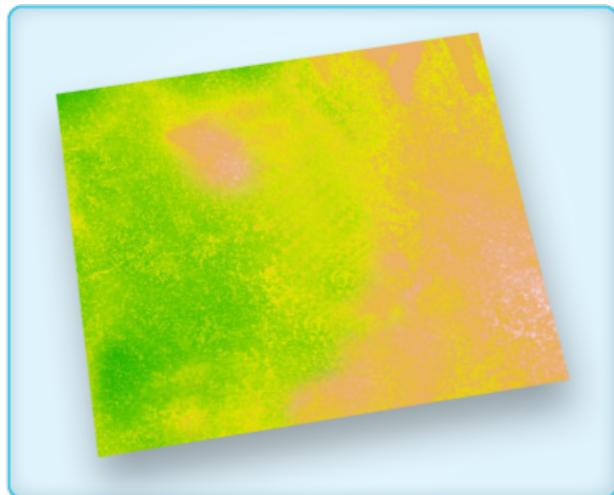
Raster(1)

(2)

(3)

(4)

Single Band Raster



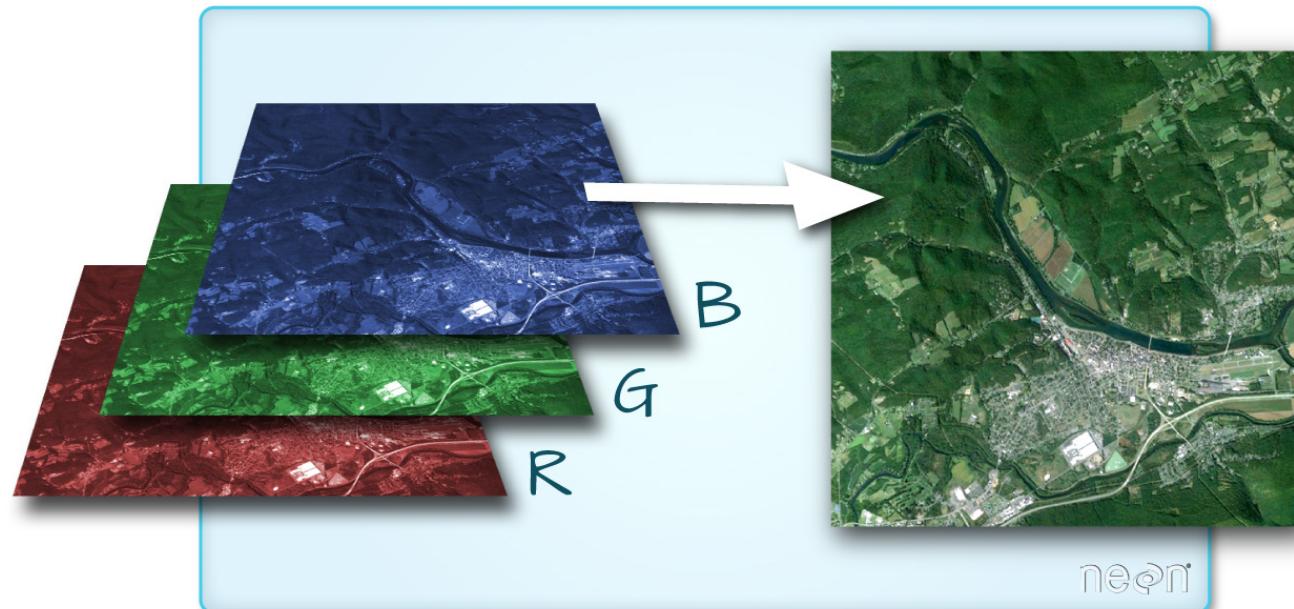
Multi Band Raster



neon°

Spatial Data ...

Type Main Goal Vectors(1) (2) Raster(1) (2) (3) **(4)**



Spatial Areal/Lattice Data

Areal Data

Examples(1)

Examples(2)

Spatial Econometrics

Involves measured quantities for each areal unit ("cells") in a study area

- The associated values are generated according to a statistical process
- Data often aggregated to the areal units
- Involves a finite (countable) number of units
- Units are non-overlapping
- Units can be regular or irregular areal units

The focus is on the relationship between values of neighbouring locations

Spatial Areal/Lattice Data

Areal Data

Examples(1)

Examples(2)

Spatial Econometrics

- Number of deaths due to Septicaemia in the municipalities of South Africa
- Presence or absence of an invasive plant species in square quadrats over a study area
- Pixel values from remote sensing

Spatial Areal/Lattice Data

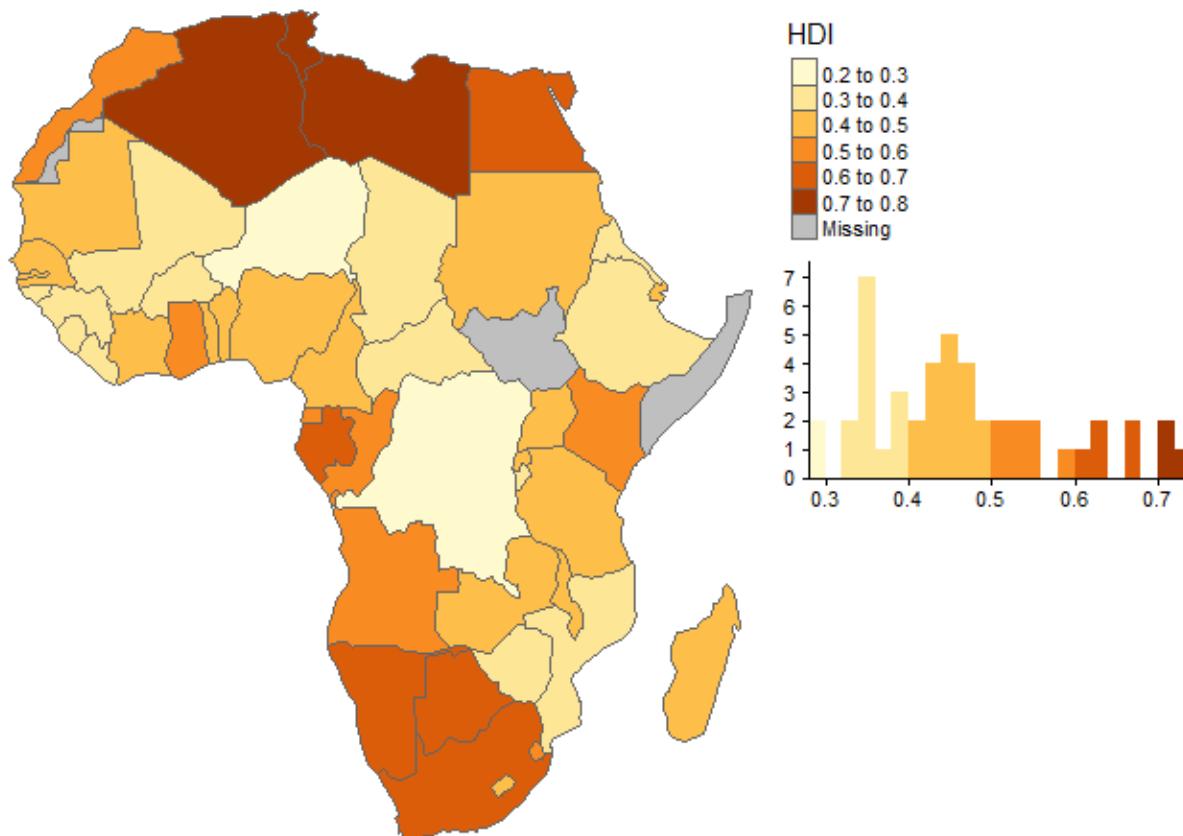
Areal Data

Examples(1)

Examples(2)

Spatial Econometrics

Human Development Index (HDI) in Africa



Spatial Areal/Lattice Data

Areal Data

Examples(1)

Examples(2)

Spatial Econometrics

A subfield of econometric methods concerned with spatial aspects present in cross-sectional and spatio-temporal data sets

- Specifying the structure of spatial effects
- Testing for the presence of spatial effects
- Estimating models with spatial effects
- Spatial prediction

Point Pattern

Point Pattern (1)

Point Pattern (2)

Point Pattern (3)

The main interest is in the locations (points) of all occurrences of some event. E.h.

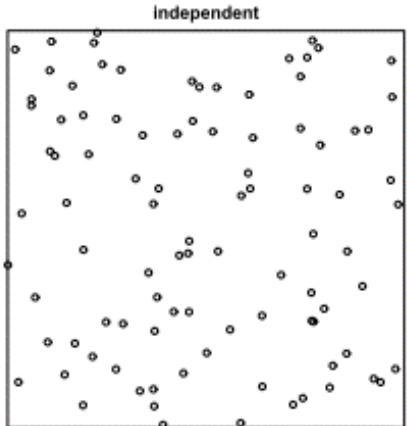
- Earthquake epicentres
- Location of road accidents
- Locations of longleaf pines

The points may also be "marked" (e.g. magnitude of earthquakes as a marker of size)

The question of interest is whether the points exhibits complete spatial randomness, clustering, or regularity

Point Pattern

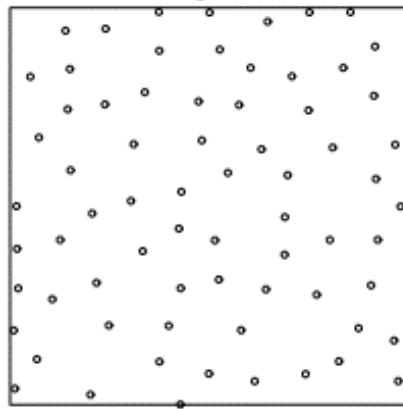
Point Pattern (1)



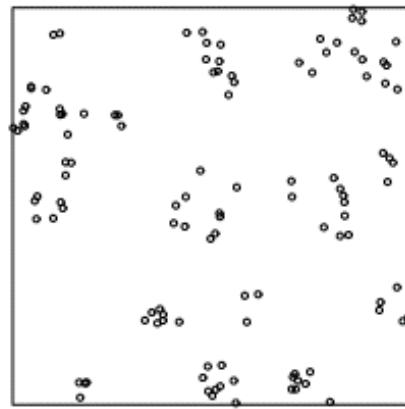
Point Pattern (2)

Point Pattern (3)

regular

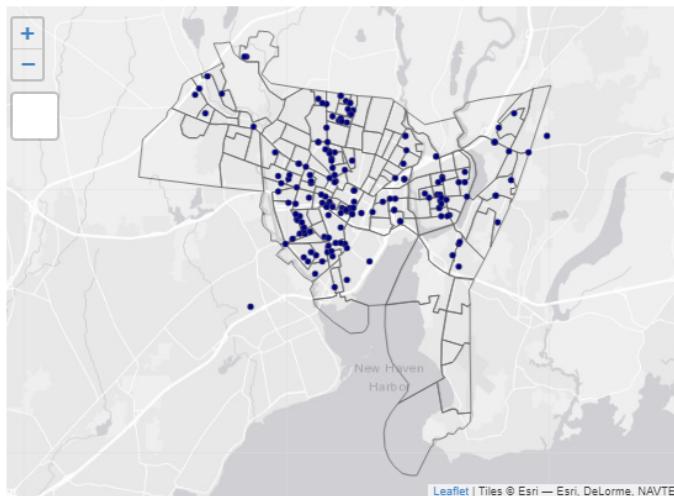


clustered



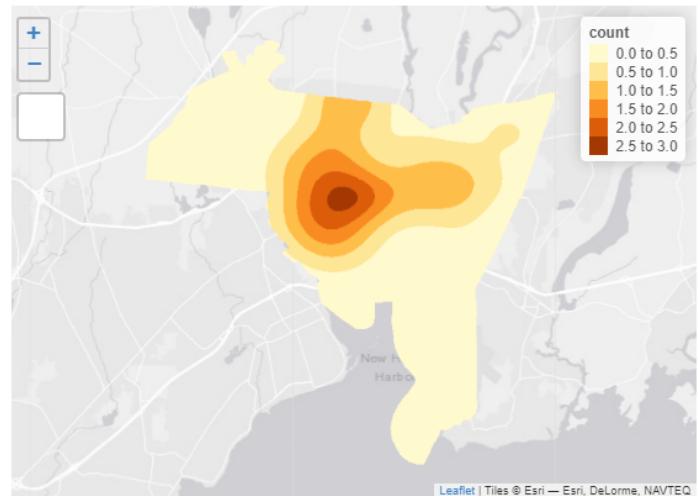
Point Pattern

Point Pattern (1)



Point Pattern (2)

Point Pattern (3)



Geostatistical

Geostatistical

Example (1)

Example (2)

The quantity of interest has a value at any location in a study domain but we only measure/record the quantity of interest at sampled sites

Aim:

- To describe the spatial variability of the phenomenon under study
- To predict the value of a variable at a point where it has not been sampled

Examples:

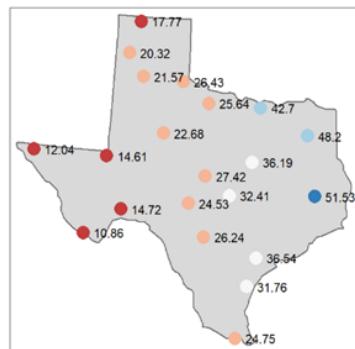
- Weekly concentrations of ozone in Kenya
- Temperature readings at various weather stations in Mali
- Richness of iron ore within an ore body at a mine in Limpopo

Geostatistical

Geostatistical

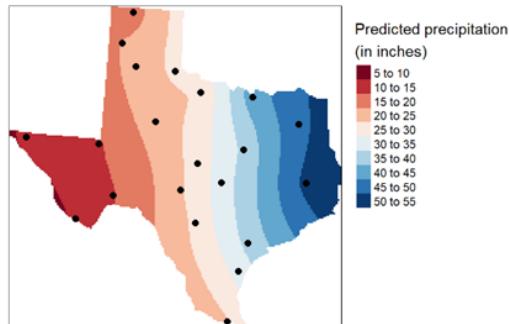
Example (1)

Example (2)



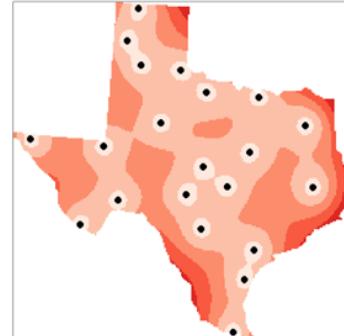
Sampled precipitation
(in inches)

- 10 to 20
- 20 to 30
- 30 to 40
- 40 to 50
- 50 to 60



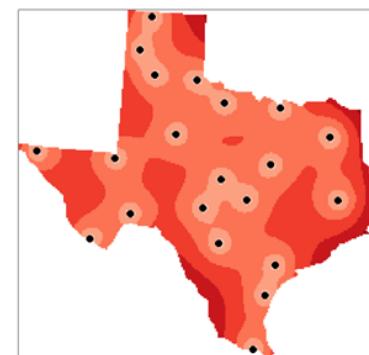
Predicted precipitation
(in inches)

- 5 to 10
- 10 to 15
- 15 to 20
- 20 to 25
- 25 to 30
- 30 to 35
- 35 to 40
- 40 to 45
- 45 to 50
- 50 to 55



Variance map
(in squared inches)

- 0 to 2
- 2 to 4
- 4 to 6
- 6 to 8
- 8 to 10
- 10 to 12



95% CI map
(in inches)

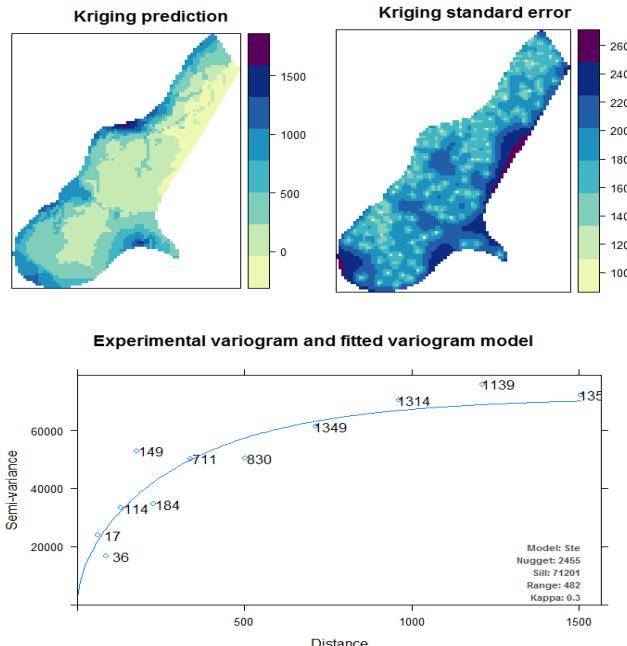
- 0 to 1
- 1 to 2
- 2 to 3
- 3 to 4
- 4 to 5
- 5 to 6
- 6 to 7

Geostatistical

Geostatistical

Example (1)

Example (2)



Interpolation of Zinc concentrations (ppm)

Issues

Issues MAUP Ecological Fallacy Edge Effect Global vs Local

- Modifiable areal unit problem (MAUP)
- Ecological Fallacy
- Edge effect
- Missing data
- Spatial weights
- e.t.c.

Issues

Issues

MAUP

Ecological Fallacy

Edge Effect

Global vs Local

a) Effects of aggregation

2	4	6	1
3	6	3	5
1	5	4	2
5	4	5	4

Mean = 3.75
Std = 2.60

3	3.5
4.5	4
3	3
4.5	4.5

Mean = 3.75
Std = 0.50

3.75	3.75
3.75	3.75

Mean = 3.75
Std = 0.00

b) Effects of zoning systems

2.5	5	4.5	3
3	4.5	4.5	3

Mean = 3.75
Std = 0.93

2.75	4.75	4.5	3.0
------	------	-----	-----

Mean = 3.75
Std = 1.04

4	1
4	3.67

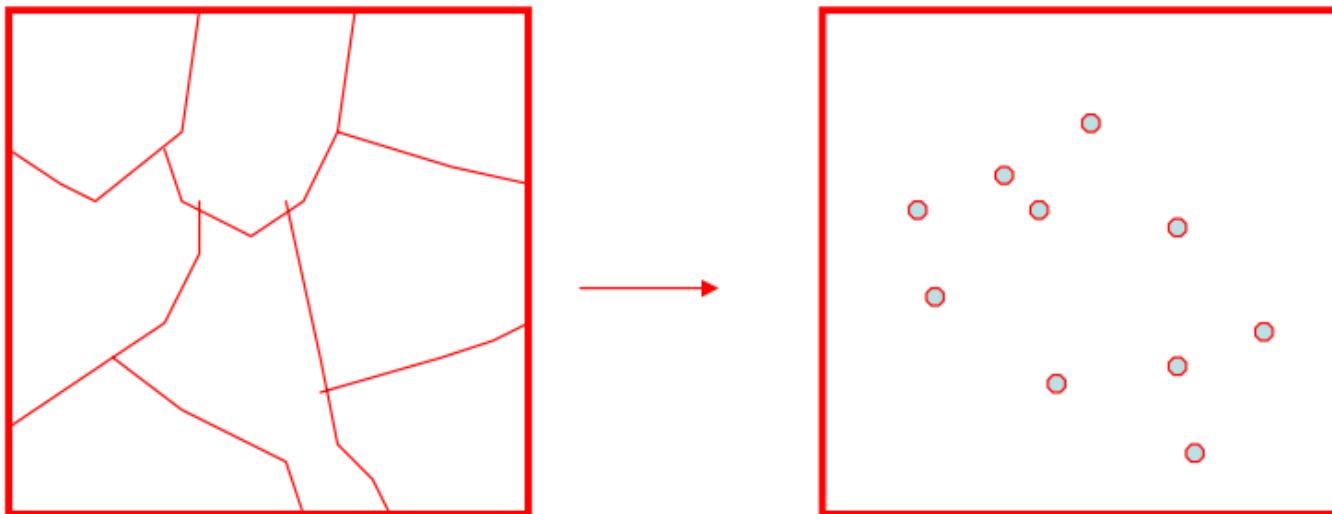
Mean = 3.17
Std = 2.11

Da Re et. al (2020)

Issues

[Issues](#)[MAUP](#)[**Ecological Fallacy**](#)[Edge Effect](#)[Global vs Local](#)

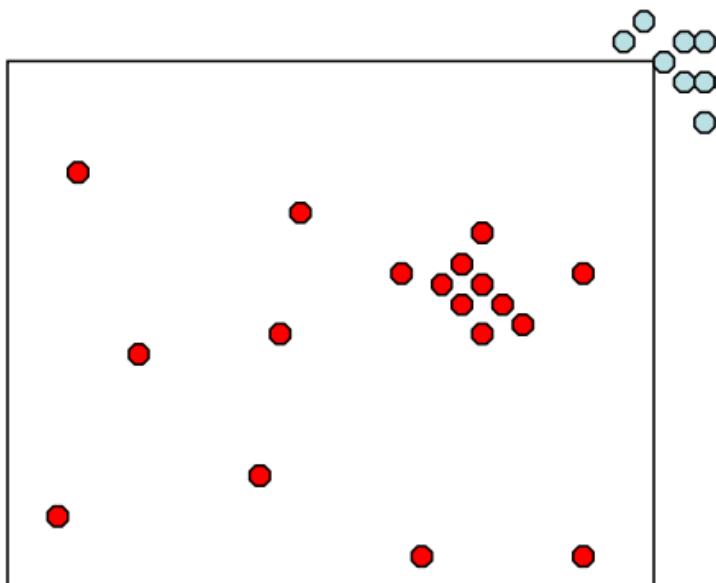
- Error in the interpretation of statistical data
- Enference about individual behaviour based on aggregate data



Issues

[Issues](#)[MAUP](#)[Ecological Fallacy](#)[Edge Effect](#)[Global vs Local](#)

- Arises when an artificial boundary is imposed on a study (often to keep it manageable / for ease of analysis)
- There is a possibility that values outside the study area boundary affect what is inside the boundary



Issues

Issues

MAUP

Ecological Fallacy

Edge Effect

Global vs Local

Global

- Single-valued statistics
- Similarity across space
- Search for regularities

Local

- Multi-valued statistics
- Difference across space
- Search for exceptions

R for Spatial Analysis

[Why R](#) [Why R...](#) [Ecosystem](#) [Framework](#) [Cheats](#)

- *de facto* language for spatial statistics
- Interactive (RStudio IDE)
- Graphics
- Open-source (free!)
- Community
- Ecosystem
- Rapidly evolving

R for Spatial Analysis

Why R

Why R...

Ecosystem

Framework

Cheats

Other options include: ArcGIS (Commercial), GeoBUGS, GeoDa, GRASS, Python?

- Python and R programmers get inspired from each other, i.e. Python's [plotnine](#) inspired by R's [ggplot2](#), and R's [rvest](#) by Python's [BeautifulSoup](#)
- You can also use functions written in Python with [source_python\(\)](#) function in R
- You can run R code from Python with [rp2](#) package, and you can run Python code from R using [reticulate](#). R version of DL package [Keras](#) calls Python
- Though, we do encourage you to learn Python as well. No harm in two languages

Take away: There is no winner, you are here to learn the skills, your focus should be on skills. If you can program in R, you can do it in any other language

R for Spatial Analysis

[Why R](#)[Why R...](#)[Ecosystem](#)[Framework](#)[Cheats](#)

Commonly used libraries when undertaking spatial analysis in R.

<https://cran.r-project.org/web/views/Spatial.html>

CRAN Task View: Analysis of Spatial Data

Maintainer: Roger Bivand

Contact: Roger.Bivand at nhh.no

Version: 2019-07-15

URL: <https://CRAN.R-project.org/view=Spatial>

Base R includes many functions that can be used for reading, visualising, and analysing spatial data. The focus in this view is on "geographical" spatial data, where observations can be identified with geographical locations, and where additional information about these locations may be retrieved if the location is recorded with care.

Base R functions are complemented by contributed packages, some of which are on CRAN, and others are still in development. One location is [Github](#). Some key packages including [sf](#) and [stars](#) are grouped under [r-spatial](#), others including [raster](#) under [rspatial](#). Maintenance of the [sp](#) is continuing here: [sp](#).

Another set of locations for the development and maintenance of packages on [R-Forge](#), which lists "Spatial Data and Statistics" projects in its [project tree](#). Information on R-spatial packages was until 2016 posted on the R-Forge spatial project [website](#), including a visualisation gallery.

The contributed packages address two broad areas: moving spatial data into and out of R, and analysing spatial data in R.

The [R-SIG-Geo](#) mailing-list is a good place to begin for obtaining help and discussing questions about both accessing data, and analysing it. The mailing list is a good place to search for information about relevant courses. Further information about courses may be found under the "Events" tab of [this blog](#).

There are a number of contributed tutorials and introductions; a recent one is [Introduction to visualising spatial data in R](#) by Robin Lovelace and James Cheshire.

The packages in this view can be roughly structured into the following topics. If you think that some package is missing from the list, please fork the [task view repository](#) and provide a pull request in ctv format for the ctv:Spatial.ctv file.

Classes for spatial data and metadata

Because many of the packages importing and using spatial data have had to include objects of storing data and functions for visualising it, an initiative is in progress to construct shared classes and plotting functions for spatial data.

Complementary initiatives are ongoing to support better handling of geographic metadata in R.

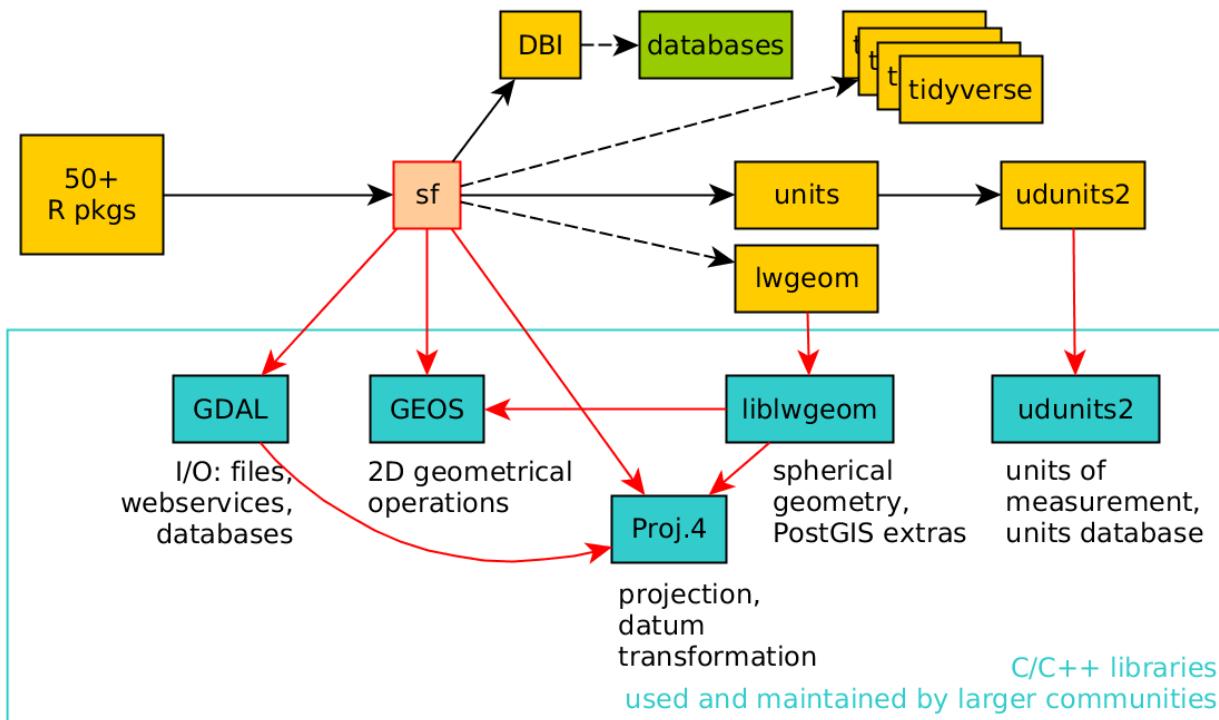
R for Spatial Analysis

Why R Why R... Ecosystem

Framework

Cheats

- **sp** - Classes and Methods for Spatial Data
- **sf** - Improved Classes
- Reading: Bivand 2022_Progress in the R ecosystem for representing and handling



R for Spatial Analysis

[Why R](#)[Why R...](#)[Ecosystem](#)[Framework](#)[Cheats](#)

Useful Cheat Sheets (See **Vula > ASDA2023 > Cheat Sheets**)

- Base R
- Advanced R
- rmarkdown
- ggplot2
- ganimate
- cartography
- Leaflet
- mapsf

Simple features, the sf R package

[sf](#) [Installing sf](#) [sf Cheat sheet](#) [\(2\)](#) [Basics](#)

A package that provides **simple features access** for R.

- The `sf` package provides a set of tools for working spatial vector data
- Represents simple features as records in a `data.frame` with a `geometry` column
- `sf` types for all dimensions (XY, XYZ, XYM, XYZM)
- Faster data I/O
- Article in [R Journal article](#) summarising the `sf` package



(Image from: [Allison Horst](#))

Simple features, the sf R package

[sf](#) [Installing sf](#) [sf Cheat sheet](#) (2) [Basics](#)

```
# Installs from CRAN
install.packages("sf")

## If installing development versions from github
devtools::install_github("r-spatial/sf")
```

You may also have to do the following (depending on your OS):

- **Windows:** Install [Rtools](#)
- **Mac:** Install [Homebrew](#)

Simple features, the sf R package

sf Installing sf sf Cheat sheet (2) Basics

- sf cheatsheet and reference website

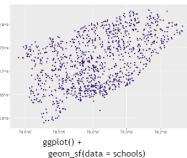
Spatial manipulation with sf: : CHEAT SHEET

The sf package provides a set of tools for working with geospatial vectors, i.e. points, lines, polygons, etc.



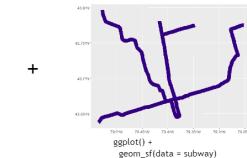
Geometric confirmation

- st_contains(x, y, ...) Identifies if x is within y (i.e. point within polygon)
- st_covered_by(x, y, ...) Identifies if x is completely within y (i.e. polygon completely within polygon)
- st_covers(x, y, ...) Identifies if any point from x is outside of y (i.e. polygon outside polygon)
- st_crosses(x, y, ...) Identifies if any geometry of x have commonalities with y
- st_disjoint(x, y, ...) Identifies when geometries from x do not share space with y
- st_equals(x, y, ...) Identifies if x and y share the same geometry
- st_intersects(x, y, ...) Identifies if x and y geometry share any space
- st_overlaps(x, y, ...) Identifies if geometries of x and y share space, are of the same dimension, but are not completely contained by each other
- st_touches(x, y, ...) Identifies if geometries of x and y share a common point but their interiors do not intersect
- st_within(x, y, ...) Identifies if x is in a specified distance to y



Geometric operations

- st_boundary(x) Creates a polygon that encompasses the full extent of the geometry
- st_buffer(x, dist, nQuadSegs) Creates a polygon covering all points of the geometry within a given distance
- st_centroid(x, ..., of_largest_polygon) Creates a point at the geometric centre of the geometry
- st_convex_hull(x) Creates geometry that represents the minimum convex geometry of x
- st_line_merge(x) Creates linestring geometry from sewing multi linestring geometry together
- st_node(x) Creates nodes on overlapping geometry where nodes do not exist
- st_point_on_surface(x) Creates a point that is guaranteed to fall on the surface of the geometry
- st_polyonize(x) Creates polygon geometry from linestring geometry
- st_segmentize(x, dfMaxLength, ...) Creates linestring geometry from x based on a specified length
- st_simplify(x, preserveTopology, dTolerance) Creates a simplified version of the geometry based on a specified tolerance



Geometry creation

- st_triangulate(x, dTolerance, bOnlyEdges) Creates polygon geometry as triangles from point geometry
- st_voronoi(x, envelope, dTolerance, bOnlyEdges) Creates polygon geometry covering the envelope of x, with x at the centre of the geometry
- st_point(x, c(numeric vector), dim = "XYZ") Creating point geometry from numeric values
- st_multipoint(x = matrix(numeric values in rows), dim = "XYZ") Creating multi point geometry from numeric values
- st_linestring(x = matrix(numeric values in rows), dim = "XYZ") Creating linestring geometry from numeric values
- st_multilinestring(x = list(numeric matrices in rows), dim = "XYZ") Creating multi linestring geometry from numeric values
- st_polygon(x = list(numeric matrices in rows), dim = "XYZ") Creating polygon geometry from numeric values
- st_multipolygon(x = list(numeric matrices in rows), dim = "XYZ") Creating multi polygon geometry from numeric values

Simple features, the sf R package

sf

Installing sf

sf Cheat sheet

(2)

Basics

Spatial manipulation with sf: : CHEAT SHEET

The sf package provides a set of tools for working with geospatial vectors, i.e. points, lines, polygons, etc.

Geometry operations

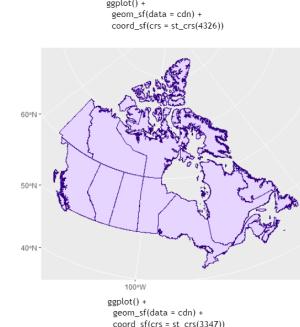
- `st_contains(x, y, ...)` Identifies if x is within y (i.e. point within polygon)
- `st_crop(x, y, ..., xmin, ymin, xmax, ymax)` Creates geometry of x that intersects a specified rectangle
- `st_difference(x, y)` Creates geometry from x that does not intersect with y
- `st_intersection(x, y)` Creates geometry of the shared portion of x and y
- `st_sym_difference(x, y)` Creates geometry representing portions of x and y that do not intersect
- `st_snap(x, y, tolerance)` Snap nodes from geometry x to geometry y
- `st_union(x, y, ..., by_feature)` Creates multiple geometries into a single geometry, consisting of all geometry elements

Geometric measurement

- `st_area(x)` Calculate the surface area of a polygon geometry based on the current coordinate reference system
- `st_distance(x, y, ..., dist_fun, by_element, which)` Calculates the 2D distance between x and y based on the current coordinate system
- `st_length(x)` Calculates the 2D length of a geometry based on the current coordinate system

Misc operations

- `st_as_sf(x, ...)` Create a sf object from a non-geospatial tabular data frame
- `st_cast(x, to, ...)` Change x geometry to a different geometry type
- `st_coordinates(x, ...)` Creates a matrix of coordinate values from x
- `st_crs(x, ...)` Identifies the coordinate reference system of x
- `st_join(x, y, join, FUN, suffix, ...)` Performs a spatial left or inner join between x and y
- `st_make_grid(x, cellsize, offset, n, crs, what)` Creates rectangular grid geometry over the bounding box of x
- `st_nearest_feature(x, y)` Creates an index of the closest feature between x and y
- `st_nearest_points(x, y, ...)` Returns the closest point between x and y
- `st_read(dsn, layer, ...)` Read file or database vector dataset as a sf object
- `st_transform(x, crs, ...)` Convert coordinates of x to a different coordinate reference system



This cheatsheet presents the sf package [Edzer Pebesma 2018] in version 0.6.3. See <https://github.com/r-spatial/sf> for more details.

CC BY Ryan Garnett <http://github.com/ryangarnett>
<https://creativecommons.org/licenses/by/4.0/>

Simple features, the sf R package

sf

Installing sf

sf Cheat sheet

(2)

Basics

- Exploratory Spatial Data Analysis (ESDA)
 - Purpose: Understand data properties and structure in support of:
 - Hypothesis development
 - Analysis and
 - Reporting of results
 - Summary Statistics and distributional analysis
 - Univariate, bivariate, multivariate
 - Maps (Static & Interactive)
 - Graphs
- Model Specification
- Model Estimation
- Model Diagnostics
- Model Prediction

Your Turn

- Think about other examples of spatial data
- Install the `sf`, `tmap` and `tmaptools` packages
- Do some reading and practice on your own. I recommend:
 - Read Chapters 1 and 2 of *Applied Spatial Data*
 - Work-through Sections 1 and 2 of *Geocomputation with R*