

Spatial Data Types and Formats

Class #6 | GEOG 215
Introduction to Spatial Data Science
Spring 2020

Today's Class

- Spatial data types
 - Vector and Raster
- Data scale and resolution
- Lab 2

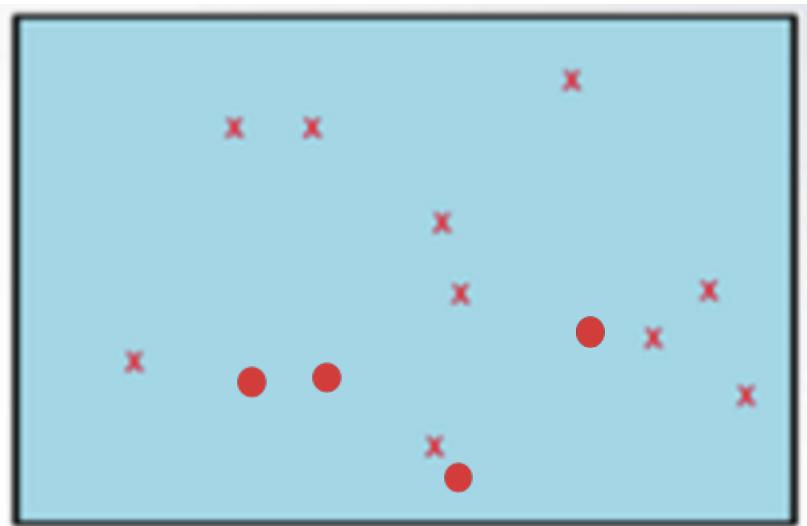
Recap

- Every location needs a reference system
 - Geographic Coordinate System
 - Projected Coordinate System
- **Abstraction**
 - Simplification of the complex world to understand various processes
 - Real World ---> Data Model
 - Assumptions must be made
 - Finding a *signal* from the noise

Conceptual Models of Space

- **Objects** present a view of the world in which the Earth surface is populated by a number discrete entities (e.g., houses, streets, lakes)
 - Objects can be 0-, 1-, 2-, or 3-dimensional
 - Object types include points, lines, polygons, networks

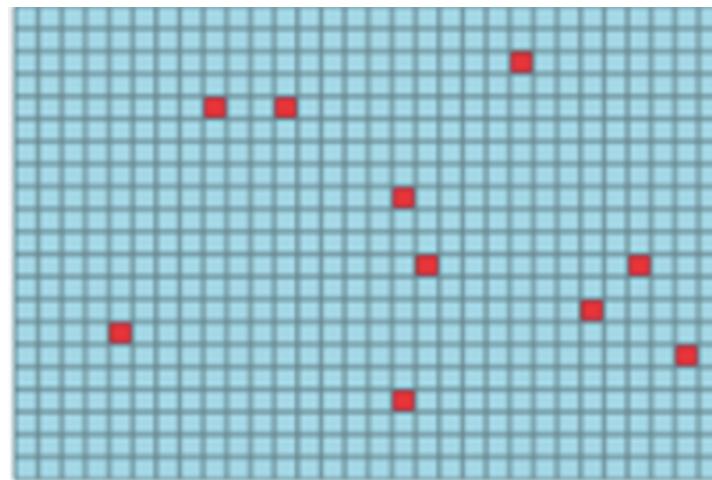
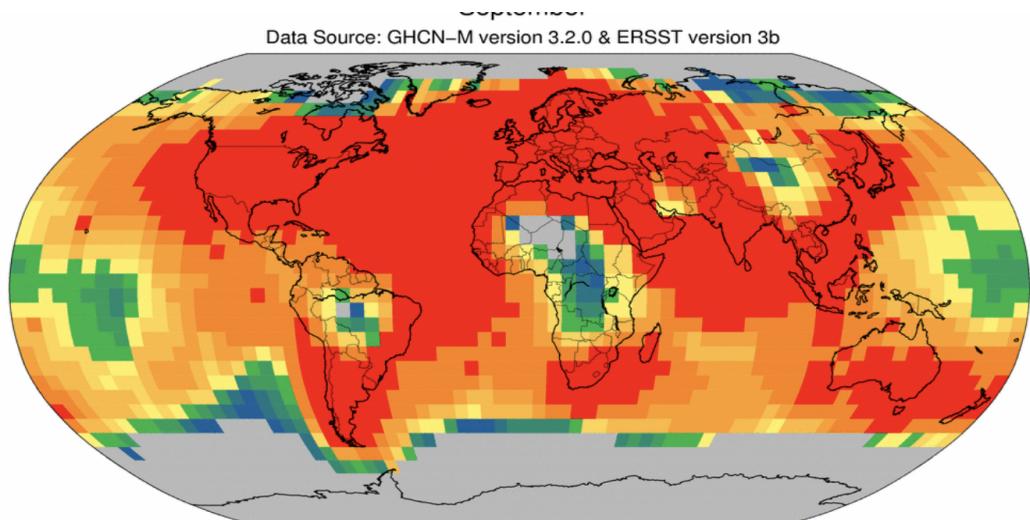
Conceptual Models of Space



Conceptual Models of Space

- Fields present a view of the world in which Earth is a continuous surface that is characterized by a number of properties
 - e.g., temperature, land use, soil pH

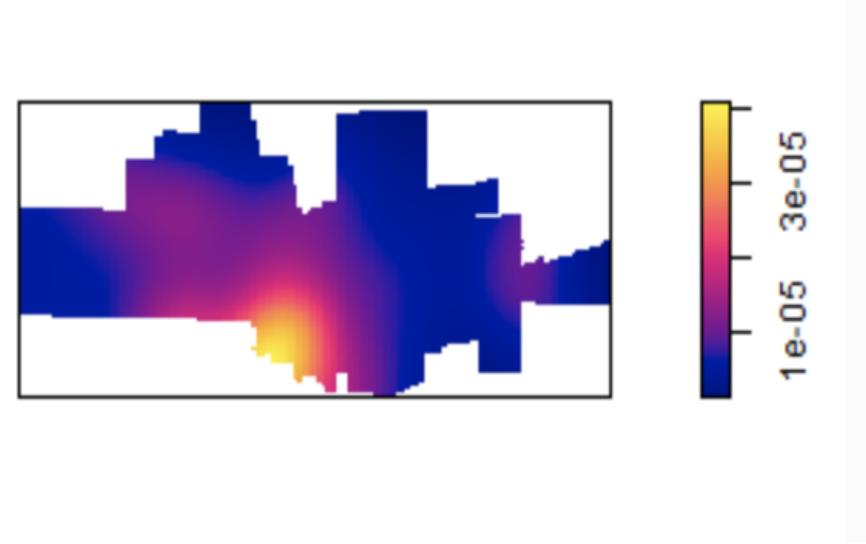
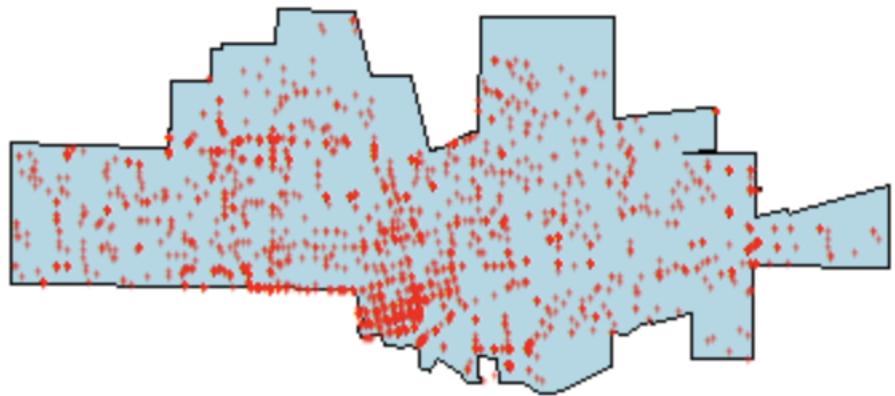
Conceptual Models of Space



Conceptual Models of Space

- Important differences between object and field conceptual models
 - **Observed element**
 - In the object model, focus is on the “object”
 - Object may have multiple attributes
 - In the field model, focus is on the “attribute”
 - Attribute varies over space
 - **Empty space**
 - In the object model, space between entities is “empty”
 - No empty space in the field model

Conceptual Models of Space

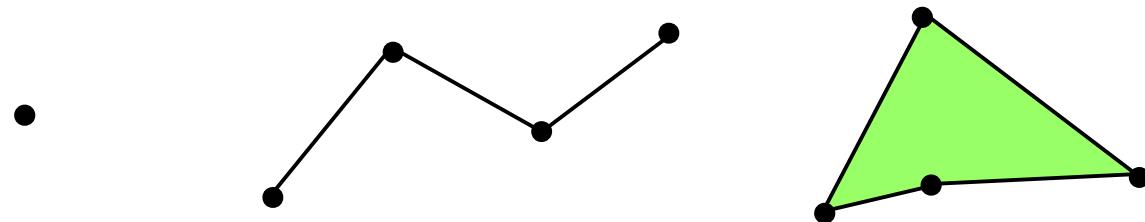


Spatial Data Models/Types

- We use spatial data models to describe the conceptual models for digital representation
 - Two primary models are raster and vector
- **Raster**
 - Uses grid cells
 - Each cell coded with a single value
- **Vector**
 - Uses spatial “features”
 - Points, lines, polygons

Vector Data

- Feature classes:
 - Points, Lines, Polygons
 - Points are the basic building blocks of vector data



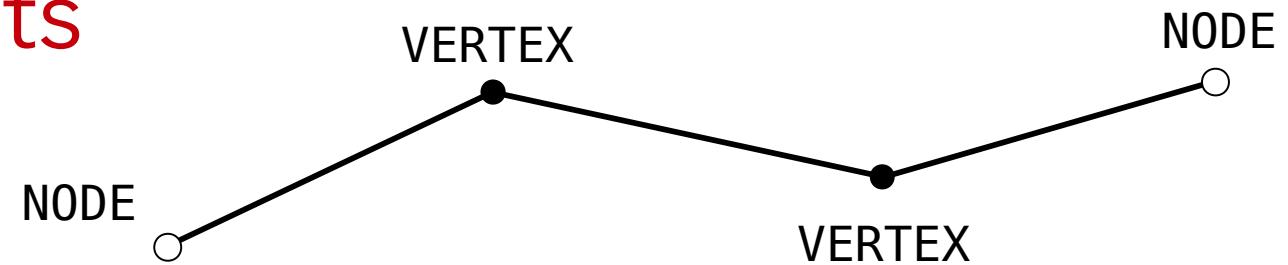
- Generally, a vector data layer contains a single “feature class” or “type”
 - Data layers contain multiple features (or discrete objects)
 - Each object may have multiple attributes

Vector Features

- Point
 - A single coordinate pair of values (x,y)
 - 0 dimensional object
 - No length or area
 - Important to recognize the difference between data representation and data display
 - Points have an “area” when we view them
 - Control over the point size via the symbology
 - Data remains 0-D

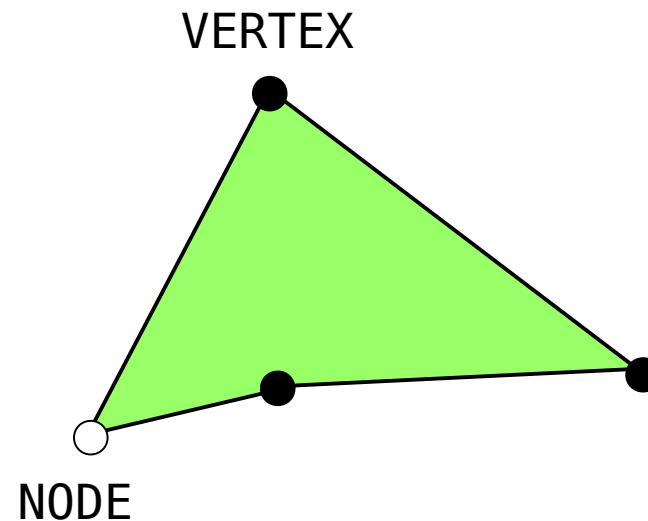
Vector Features

- Line
 - An ordered set of points
 - 1 dimensional object
 - Nodes
 - Starting and ending points
 - From node = Start point, To node = End point
 - Vertices
 - Intermediate points that define the shape of the line



Vector Features

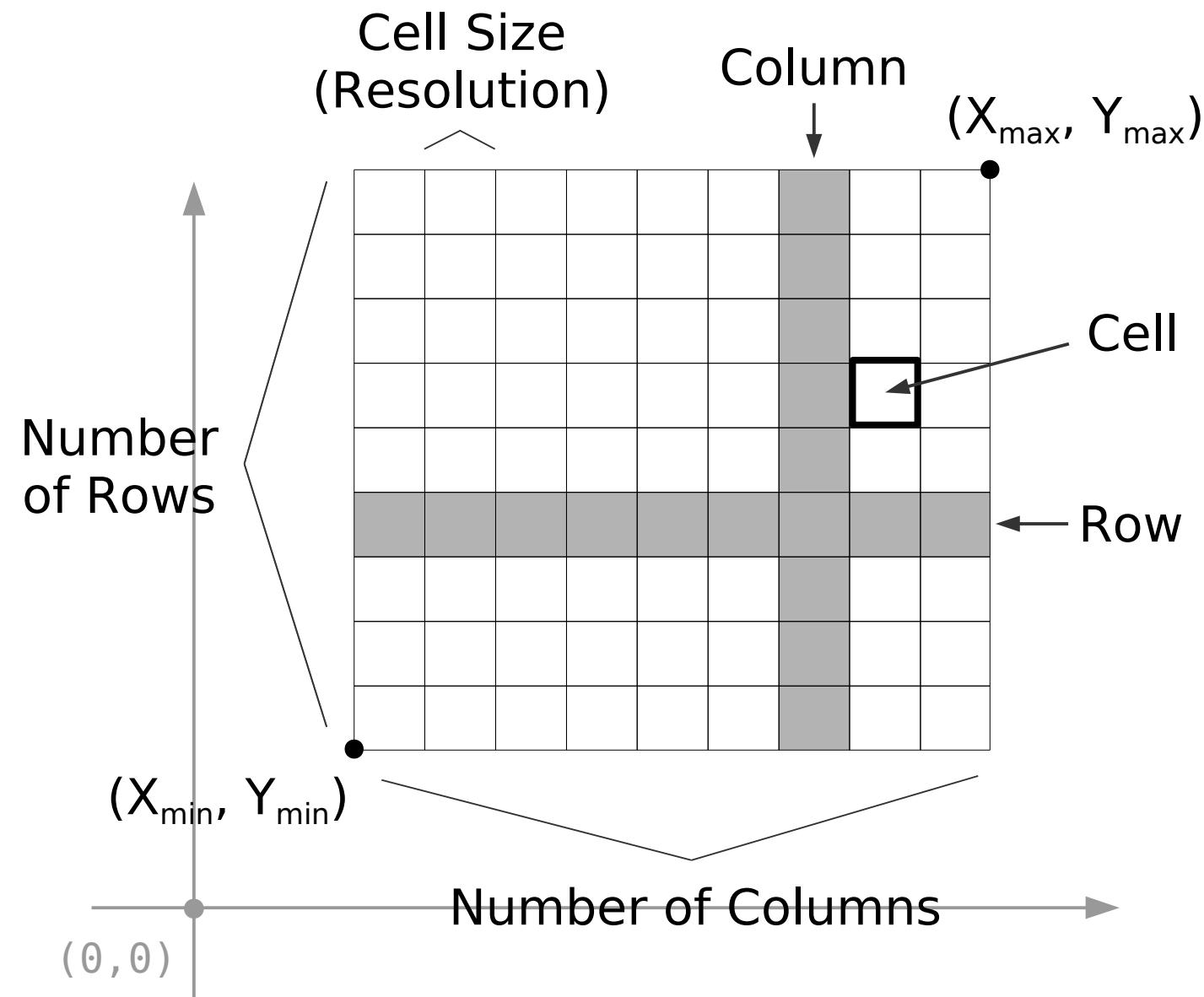
- Polygon
 - A little more complex
 - 2 dimensional object
 - A single, closed line
 - Node
 - Starting and ending points
 - Vertices
 - Intermediate points that define the shape of the line
 - A set of connected lines forming an enclosed area



Field View

- Represented as raster data
 - Attribute based
 - Regularly spaced tessellations
 - Generally, square grid cells (or pixels)
 - Every cell location has a specific value

Raster Data Basics



Basic File Storage

- Matrix

3, 3, 5, 5, 1

3, 3, 5, 1, 1

3, 3, 3, 1, 1

3, 4, 4, 4, 1

4, 4, 4, 4, 4

3	3	5	5	1
3	3	5	1	1
3	3	3	1	1
3	4	4	4	1
4	4	4	4	4

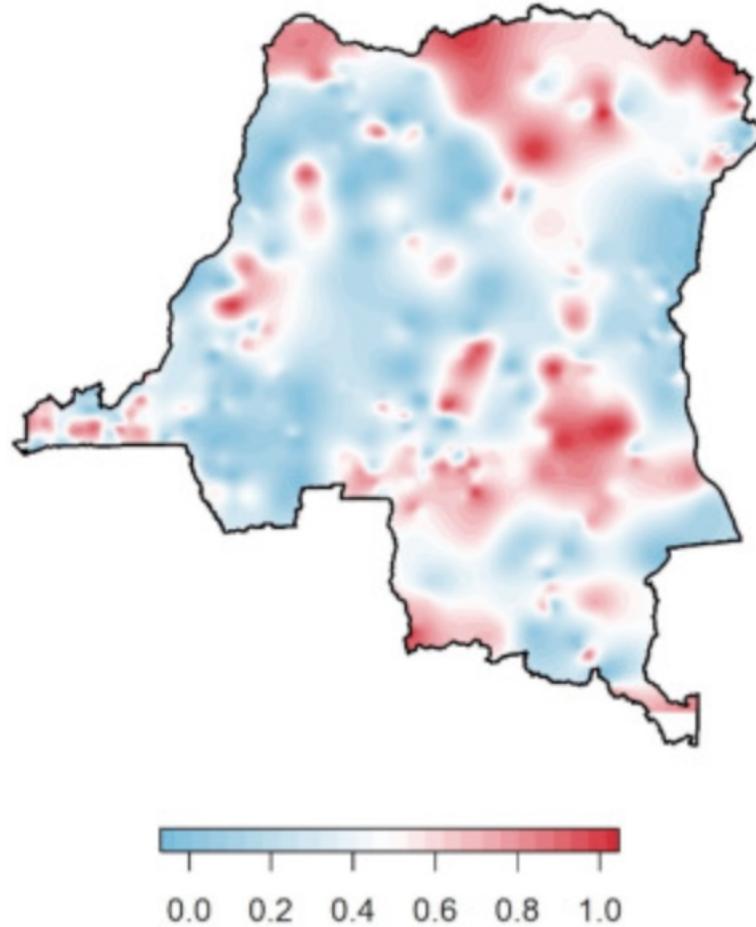
- Linear

3, 3, 5, 5, 1, 3, 3, 5, 1, 1, 3, 3, 3, 1, 1, 3, 4, 4, 4, 1, 4, 4, 4, 4, 4

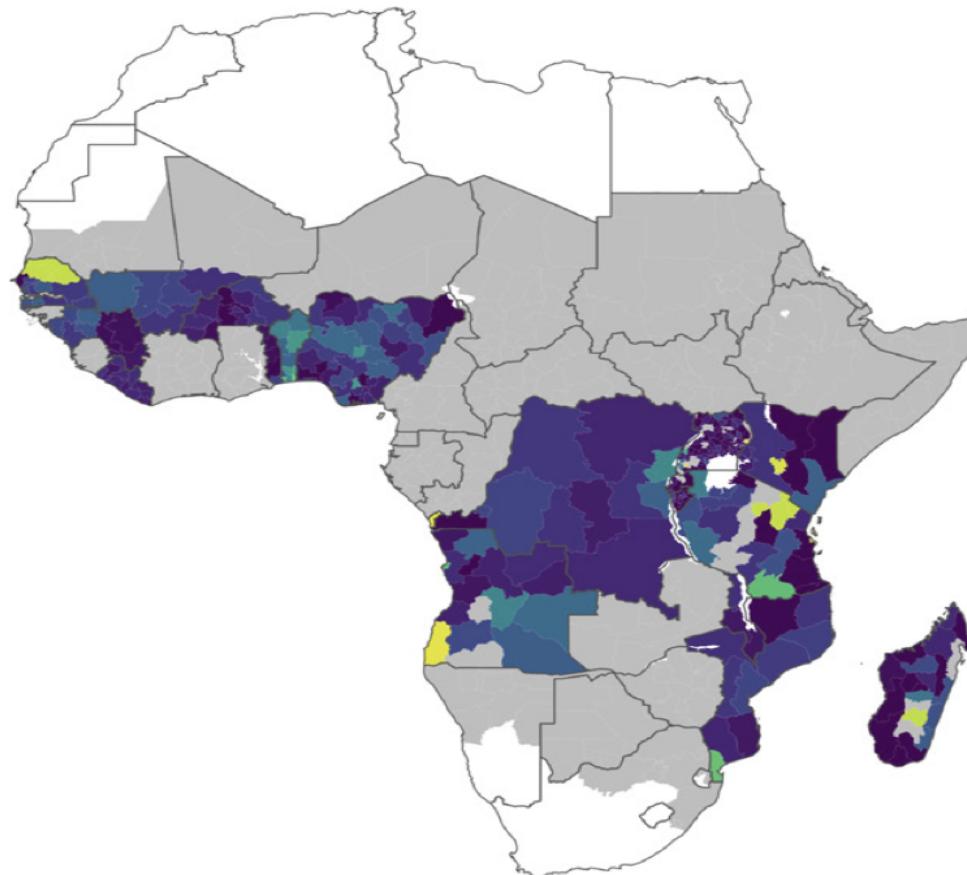
- Requires additional information

- Header (e.g., 5 rows, 5 columns)

Examples



Malaria vector prevalence in the Democratic Republic of Congo



Distribution of Malaria tests by country in Sub-saharan Africa

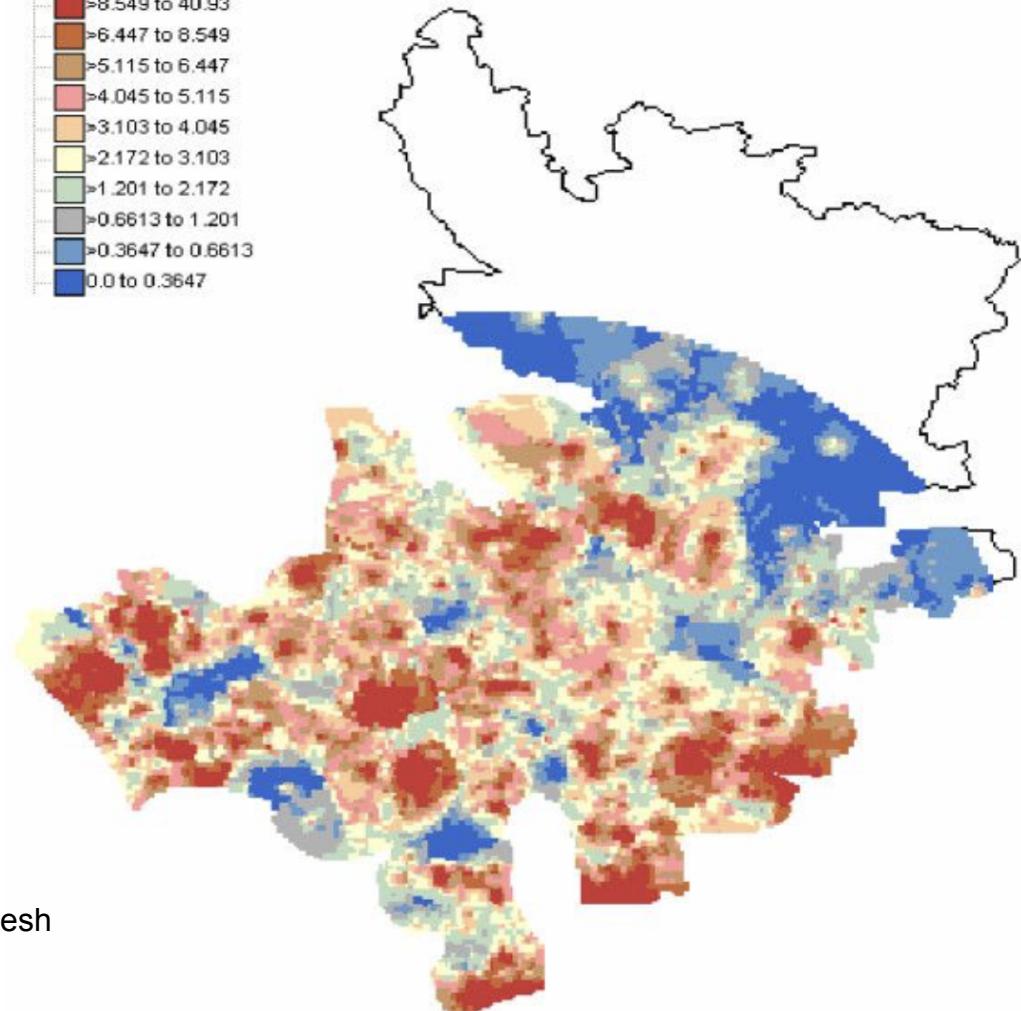
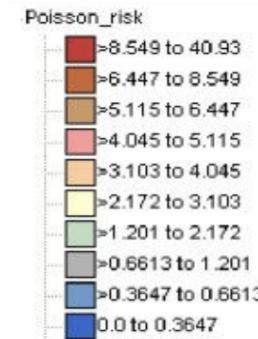
Examples

Figure 1

- Cholera case baris
- Dysentery case baris
- Non-case baris
- Matlab Hospital
- Dhonagoda River
- Study Area



Cholera in Rural Bangladesh



Cholera

Scale of Spatial Data

- Scale is easily changed in digital representations
 - Paper maps are static
 - Consider the ability to “zoom” in and out on a computer display
- For spatial data,
 - **Resolution**
 - Refers to the smallest resolvable unit
 - **Extent**
 - Refers to the size of the area represented

Data Resolution

- Raster data sets
 - Cell size
- Points
 - Dependent on average point spacing

Data Resolution

- Polygon and line features
 - Resolution is often defined by the “minimum mapping unit”
 - Minimum mapping unit is the smallest line or polygon feature that is (or can be) mapped
 - Dependent on scale of original source and goals of interpreters
 - Can also be defined by administrative units
 - e.g., County-level data
 - Values for each county

Attribute Files/Tables

- *Records or features* contain a set of attribute values for a single entity or feature
 - Table rows
- *Items or Fields* are attributes or variables describing the entities
 - Table columns

Record #	PLACENAME	STATE	Item or Field
1	Adrian	MI	22281
2	Albion	MI	10095
3	Alger	MI	2015
4	Algonac	MI	4612
5	Allegan	MI	4592
6	Allen Park	MI	31051
7	Allendale	MI	6950
8	Alma	MI	9075
9	Almont	MI	2375
10	Alpena	MI	11358
11	Ann Arbor	MI	109646
12	Atlanta	MI	2299
13	Atlantic Mine	MI	3379
14	Avoca	MI	2325

Polleverywhere

```
> print(australia_provinces[1,])  
Simple feature collection with 11 features and 3 fields  
geometry type:  MULTIPOLYGON  
dimension:      XY  
bbox:           xmin: 112.9194 ymin: -54.75042 xmax: 159.1065 ymax: -9.240167  
epsg (SRID):    4326  
proj4string:    +proj=longlat +datum=WGS84 +no_defs  
#>
```

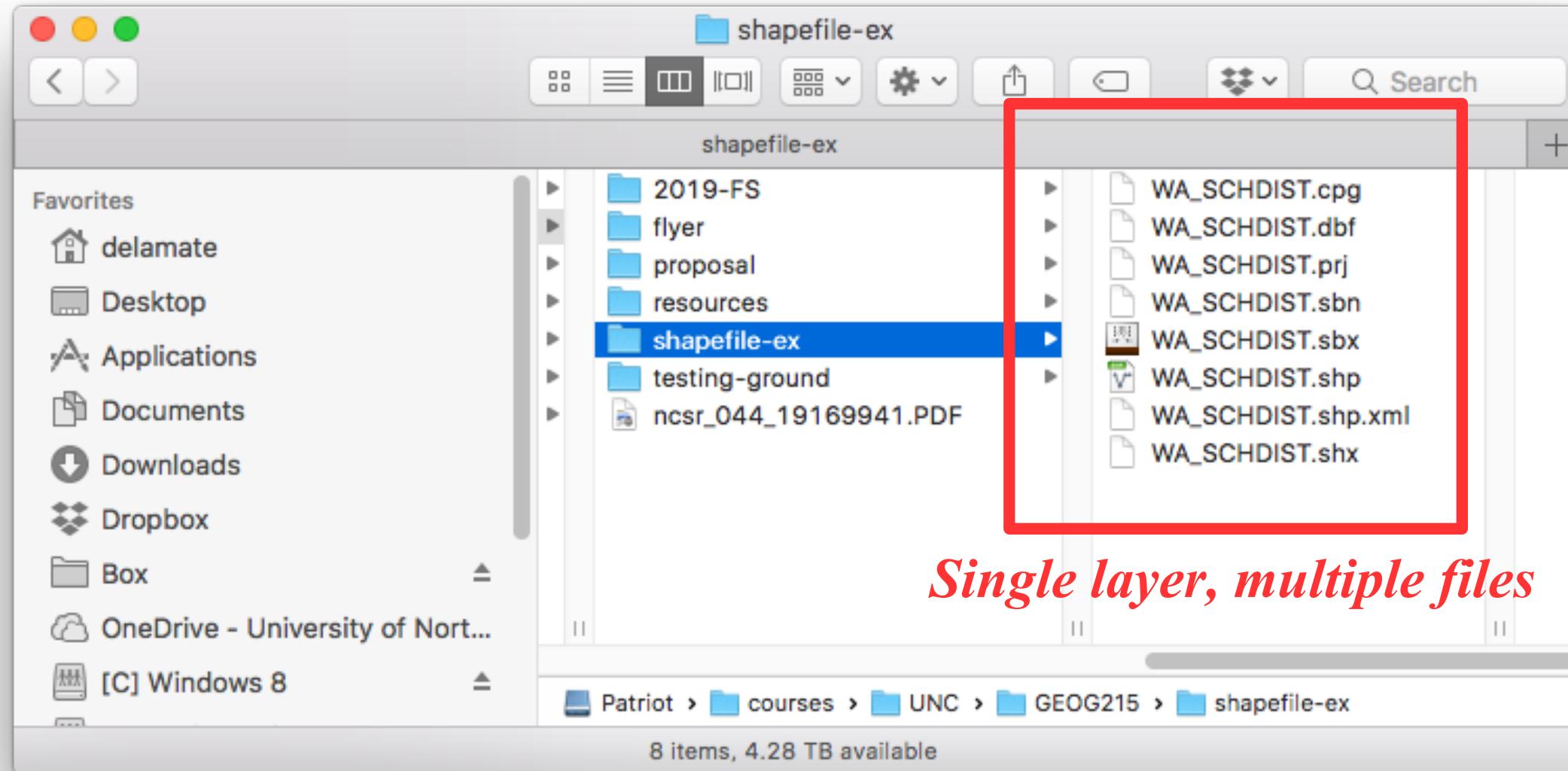
Polleverywhere

```
## dimensions : 5060, 4299, 21752940 (nrow, ncol, ncell)
## resolution : 1, 1 (x, y)
## extent      : 254570, 258869, 4107302, 4112362 (xmin, xmax, ymin,
ymax)
## coord. ref. : +proj=utm +zone=11 +datum=WGS84 +units=m +no_defs
+ellps=WGS84 +towgs84=0,0,0
```

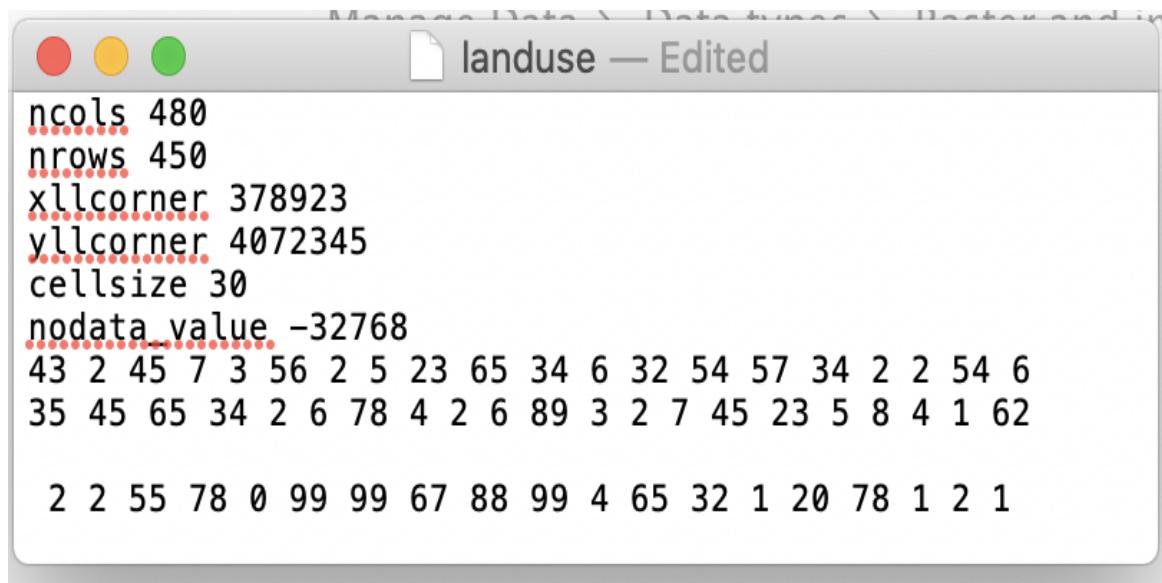
Common Spatial Data Formats

Name	Extension	Info	Type	Model
ESRI Shapefile	.shp (the main file)	Popular format consisting of at least three files. No support for: files > 2GB; mixed types; names > 10 chars; cols > 255.	Vector	Partially open
GeoJSON	.geojson	Extends the JSON exchange format by including a subset of the simple feature representation.	Vector	Open
KML	.kml	XML-based format for spatial visualization, developed for use with Google Earth. Zipped KML file forms the KMZ format.	Vector	Open
GPX	.gpx	XML schema created for exchange of GPS data.	Vector	Open
GeoTIFF	.tif/.tiff	Popular raster format. A TIFF file containing additional spatial metadata.	Raster	Open
Arc ASCII	.asc	Text format where the first six lines represent the raster header, followed by the raster cell values arranged in rows and columns.	Raster	Open
R-raster	.gri, .grd	Native raster format of the R-package raster.	Raster	Open
SQLite/SpatiaLite	.sqlite	Standalone relational database, SpatiaLite is the spatial extension of SQLite.	Vector and raster	Open
ESRI FileGDB	.gdb	Spatial and nonspatial objects created by ArcGIS. Allows: multiple feature classes; topology. Limited support from GDAL.	Vector and raster	Proprietary
GeoPackage	.gPKG	Lightweight database container based on SQLite allowing an easy and platform-independent exchange of geodata	Vector and raster	Open

Shapefile (file structure)



Raster (Data Structure)



A screenshot of a Mac OS X-style text editor window titled "landuse — Edited". The window contains Raster data structure code. The code includes metadata parameters: ncols (480), nrows (450), xllcorner (378923), yllcorner (4072345), cellsize (30), and nodata_value (-32768). Below the metadata is a grid of numerical values representing the raster data.

```
ncols 480
nrows 450
xllcorner 378923
yllcorner 4072345
cellsize 30
nodata_value -32768
43 2 45 7 3 56 2 5 23 65 34 6 32 54 57 34 2 2 54 6
35 45 65 34 2 6 78 4 2 6 89 3 2 7 45 23 5 8 4 1 62

2 2 55 78 0 99 99 67 88 99 4 65 32 1 20 78 1 2 1
```

LAB 2

Wrap Up

- Finish data camp exercises
- Lab 2 Due ,Feb 4 11:59 pm
- Office Hours- Thu,Fri, next Monday
- Next class
 - Importing and handling spatial data in R
 - Finding and Acquiring sources for spatial data
(Think about what kind of data and questions you are interested in- fill in survey):