Raster & Vector GIS

John H. Porter

GIS Data

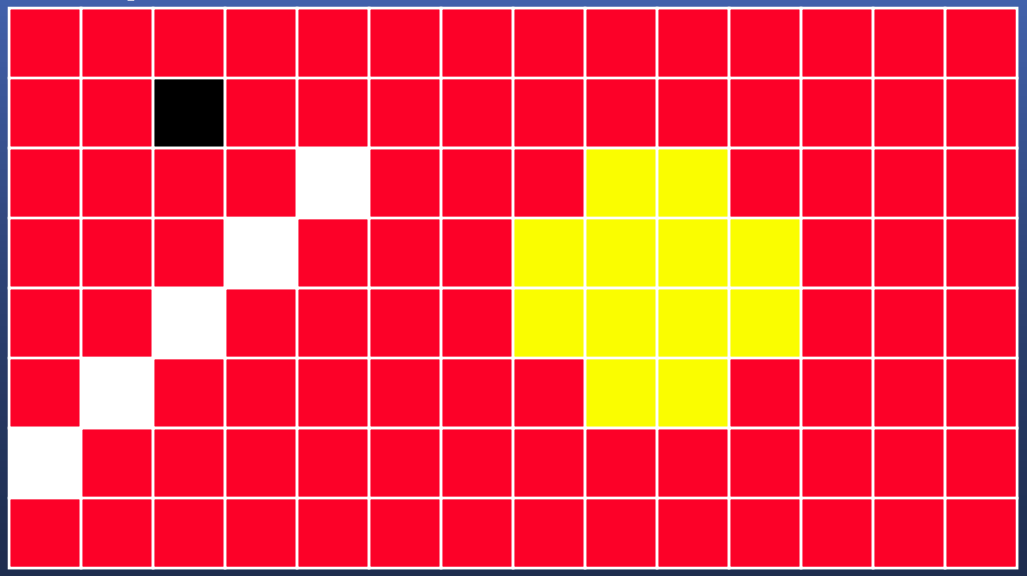
- What types of data does a GIS need to represent?
- Continuous (phenomena)
  - Data takes on a wide range of values
  - Data values do not represent themes or classes, but rather a specific variable
    - E.g. Elevation= 4.534 m
- Discrete/Thematic/Categorical (objects)
  - Feature represents a discrete class
    - E.g. Interstate highway, Forest, Tree stem
  - Data stored is a code that represents a class
    - E.g. 1= oak, 2=maple, 3=hickory

GIS Data

- Images
  - Form of continuous data where the variable being represented is brightness
  - Can include multiple “bands” representing brightness in different spectral ranges
- Digital Elevation Models
  - Continuous data where the variable being represented is elevation

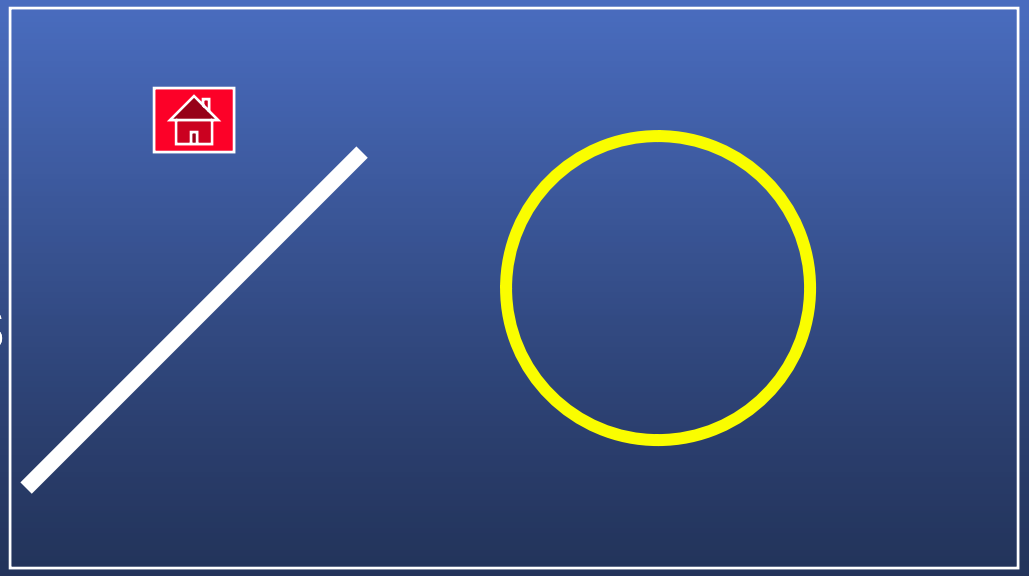
How Can GIS Represent Data?

- Raster
  - Break the area being represented into “pixels” (picture elements)
  - Assign each pixel a value that may represent continuous or discrete values

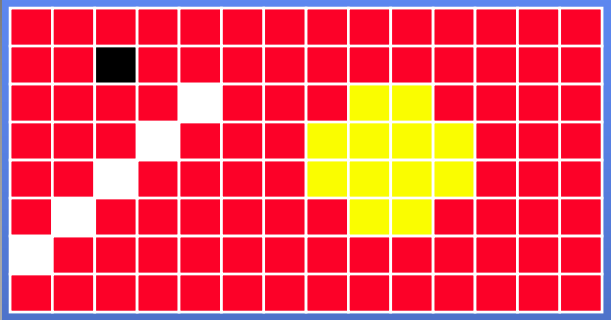


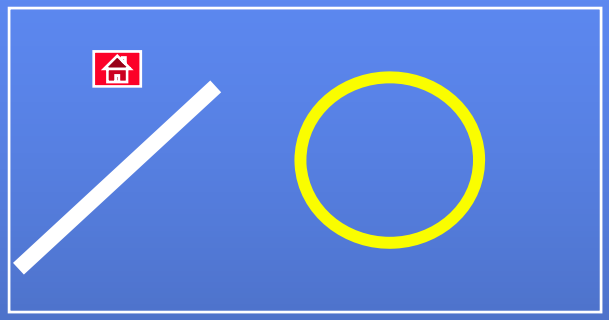
Vector Representations

- Use points and lines to represent features. Polygons are represented by boundaries
- Vector
  - Points
  - Lines
  - Polygons/Areas
  - Regions



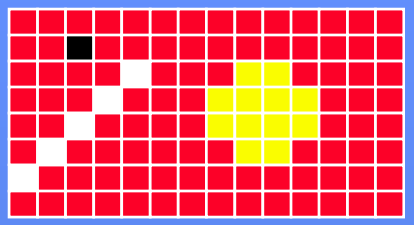
Raster vs. Vector

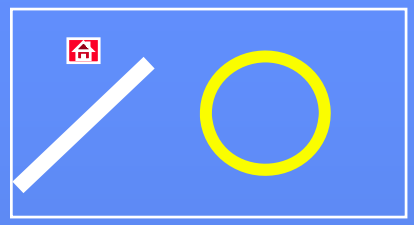




Type	Points	Lines	Polygons
Vector	Dimension-less	Length	Boundary Only
Raster	Single Pixel	Cluster of pixels	Cluster of pixels

Raster vs. Vector





- Raster
  - Resolution determined by pixel size
  - Efficiently represents dense data.
    - E.g. Elevation
- Vector
  - Resolution determined by precision of coordinates
  - Efficiently represents sparse data
    - E.g. House locations

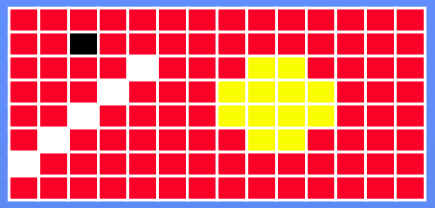
Spatial Resolution & Extent

- Raster
  - Size depends on pixel size and extent of the area covered
  - The size of a pixel should be less than ½ the size of the smallest object to be mapped (the “minimum mapping unit”)
  - “The curse of dimensionality” – storage required for a raster increases product of its linear dimensions in the units of its pixel size

Spatial Resolution and Extent

- Example: We have a square 1 km on a side
  - If each pixel is 10 m on a side we have 1000/10 * 1000/10 = 10,000 pixels
  - If each pixel is 1 m on a side we have 1000/1 * 1000/1 =1,000,000 pixels
  - If each pixel is 1 cm on a side we have
  - 1000/0.01 * 1000/0.01 = 10,000,000,000 – ten billion!
- Example: We change to a square 2 km on a side
  - If each pixel is 10 m on a side: we have 40,000 pixels
  - If each pixel is 1 m on a side we have 4,000,000 pixels
  - If each pixel is 1 cm on a side we have 40,000,000,000 pixels
  - Total size increases 2x2 - if went to 10 km, would be 100x larger!

Raster

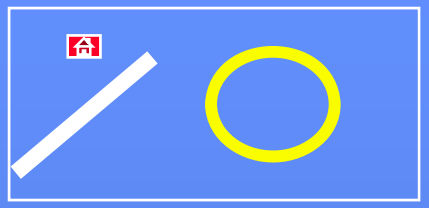


- Advantages
  - “Easy” creation from image data
  - Easy to overlay
  - Efficient storage for dense, heterogeneous data
- Disadvantages
  - Must pre-define spatial resolution
  - Requires large amounts of storage space
  - Inefficient when data is sparse or homogeneous
  - Deals poorly with linear features

Spatial Resolution & Extent

- For Vector data the storage required is not directly dependent on extent, but rather the number of features and their complexity and the level of spatial resolution
- Precision
  - Single Precision – Typically 7 digits – 76543.21
    - Bad for UTM which has lots of digits
  - Double Precision – Typically 15 digits – 54321098765.4321
    - Much better for UTM or where high precision is needed
    - Requires double the storage than single precision
- Shapefiles use double precision for coordinates

Vector



- Advantages
  - High resolution
  - Works well with boundaries
  - Explicit representation of linear features
  - Efficient storage of sparse data
- Disadvantages
  - Manipulations require sophisticated algorithms
  - Processing can require lots of computer time
  - Inefficient storage of dense data

Tools for Raster/Grid Data

- Spatial Analyst provides raster processing capabilities in ARCGIS
  - Uses “GRID “ format raster files
- The Raster Calculator is a powerful tool for combining data layers to produce new raster data layers
  - Has functions for hydrologic and least-cost path analyses

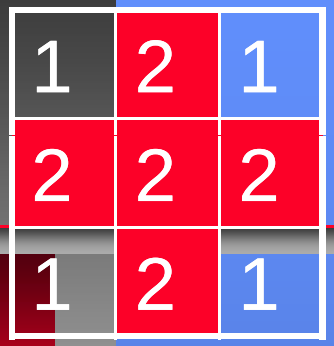
Raster Calculation Example

1	2	1
2	2	2
1	2	1

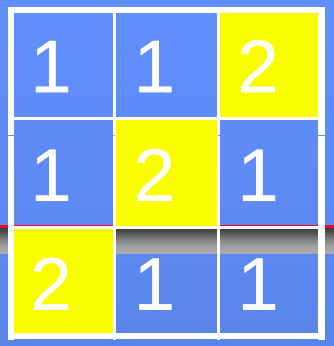
HABITAT

1	1	2
1	2	1
2	1	1

SOILTYPE



HABITAT



SOILTYPE

- We want to calculate an index of habitat suitability for *Multipodes chickini* – the Multilegged Chicken
  - We believe the density of chickens in habitat 2 would be 2 times the density in habitat 1
  - We believe that the density of chickens in soil type 2 will be 3 times the density in soil type 1
- How can we convert these into a habitat suitability score?

1	2	1
2	2	2
1	2	1

HABITAT

1	1	2
1	2	1
2	1	1

SOILTYPE

- In raster calculator we can use the equation
- [Suitability]= 2*([HABITAT]-1) + 3*([SOILTYPE]-1) to produce

0	2	3
2	5	2
3	2	0

Square [brackets] indicate grid layer names

Choosing between Raster and Vector

- How dense is your collection of measurements?
- Are they regularly distributed in a grid?
- Continuous vs categorical?
- Do you need to calculate any statistics based on surrounding areas?
- Do you need to represent lines?