Strategic Data Science (SDS)

Spatial Data

Karl Ho
School of Economic, Political and Policy Sciences
University of Texas at Dallas

What is Spatial Data?

What is Spatial Data?

1. [data structures] Information about the locations and shapes of geographic features and the relationships between them, usually stored as coordinates and topology.

What is Spatial Data?

1. [data structures] Information about the locations and shapes of geographic features and the relationships between them, usually stored as coordinates and topology.
2. [data models] Any data that can be mapped.

What is Spatial Data?

1. [data structures] Information about the locations and shapes of geographic features and the relationships between them, usually stored as coordinates and topology.
2. [data models] Any data that can be mapped.

-ESRI GIS Dictionary

What is Spatial Data?

What is Spatial Data?

The two major types of GIS file formats

What is Spatial Data?

The two major types of GIS file formats

1. Raster - grids of cells or pixels. Raster formats are useful for storing GIS data that vary, such as elevation (as seen on this gorgeous Martian landscape map).

What is Spatial Data?

The two major types of GIS file formats

1. Raster - grids of cells or pixels. Raster formats are useful for storing GIS data that vary, such as elevation (as seen on this gorgeous Martian landscape map).
2. Vector - polygons that use points (called nodes) and lines. Vector formats are useful for storing GIS data with firm borders, such as school districts or streets.

What is Spatial Data?

The two major types of GIS file formats

1. Raster - grids of cells or pixels. Raster formats are useful for storing GIS data that vary, such as elevation (as seen on this gorgeous Martian landscape map).
2. Vector - polygons that use points (called nodes) and lines. Vector formats are useful for storing GIS data with firm borders, such as school districts or streets.

- National Geographic Society

What is Spatial Data?

What is Spatial Data?

Raster Model

area is covered by grid with (usually) equal-sized, square cells attributes are recorded by assigning each cell a single value based on the majority feature (attribute) in the cell, such as land use type.

Image data is a special case of raster data in which the “attribute” is a reflectance value from the geomagnetic spectrum cells in image data often called pixels (picture elements).

What is Spatial Data?

What is Spatial Data?

Vector Model

What is Spatial Data?

Vector Model

1. Points or dots (nodes): trees, poles, fire plugs, airports, cities

What is Spatial Data?

Vector Model

1. Points or dots (nodes): trees, poles, fire plugs, airports, cities
2. Lines (arcs): streams, streets, sewers,

What is Spatial Data?

Vector Model

1. Points or dots (nodes): trees, poles, fire plugs, airports, cities
2. Lines (arcs): streams, streets, sewers,
3. Areas (polygons): land parcels, cities, counties, forest, rock type

What is GIS?

What is GIS?

1. GIS is '. . . a powerful set of tools for collecting, storing, retrieving at will, transforming, and displaying spatial data from the real world for a particular set of purposes'.

   - Burrough and McDonnell (1998, p. 11)

What is GIS?

What is GIS?

1. Another definition (from same authors) is "checking, manipulating, and analysing data, which are spatially referenced to the Earth".

What is GIS?

What is GIS?

1. GI
   - Geographic/Geospatial

What is GIS?

1. GI
   - Geographic/Geospatial
2. I
   -Information

What is GIS?

1. GI
   - Geographic/Geospatial
2. I
   -Information
3. S

What is GIS?

1. GI
   - Geographic/Geospatial
2. I
   -Information
3. S

- Systems: the technology

What is GIS?

1. GI
   - Geographic/Geospatial
2. I
   -Information
3. S

- Systems: the technology
- Science: the concepts and theory

What is GIS?

1. GI
   - Geographic/Geospatial
2. I
   -Information
3. S

- Systems: the technology
- Science: the concepts and theory
- Studies: the societal context

GIS and Spatial data

GIS and Spatial data

1. Global Positioning System (GPS)

GIS and Spatial data

1. Global Positioning System (GPS)
    - Satellites provide data of locations on the earth's surface

GIS and Spatial data

1. Global Positioning System (GPS)
    - Satellites provide data of locations on the earth's surface
2. Remote Sensing
    - Use of satellites or aircraft to capture information about the earth's surface

GIS and Spatial data

1. Global Positioning System (GPS)
    - Satellites provide data of locations on the earth's surface
2. Remote Sensing
    - Use of satellites or aircraft to capture information about the earth's surface
3. GIS database
    - Storage of multiple layers of spatial data

Why using spatial data?

Why using spatial data?

Applications:

Why using spatial data?

Applications:

Does the spatial patterning of disease incidences give rise to the conclusion that they are clustered, and if so, are the clusters found related to factors such as age, relative poverty, or pollution sources?

Why using spatial data?

Why using spatial data?

Applications:

Why using spatial data?

Applications:

Given a number of observed soil samples, which part of a study area is polluted?

Why using spatial data?

Applications:

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. PM10),1 and where do they live?

Why using spatial data?

Why using spatial data?

Applications:

Why using spatial data?

Applications:

Do governments tend to compare their policies with those of their neighbours, or do they behave independently?

Spatial data and Attribute data

1. Spatial data specify location stored in a shape file, geodatabase or similar geographic file
2. Attribute (descriptive) data specify characteristics at that location (what, how much, when)

GIS systems (e.g. ArcGIS) traditionally maintain spatial and attribute data separately, then “join” them for display or analysis.

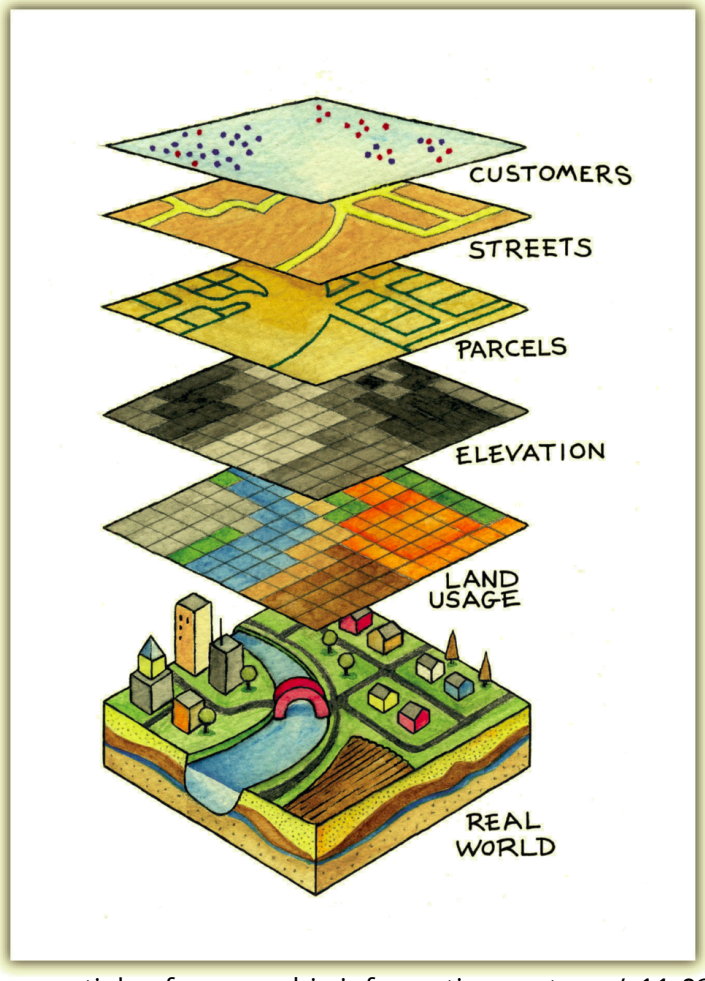
Spatial data: Geocoding

The process of identifying locations to be placed on a map is called geocoding, that is coding the location of an object, a place, an event, a building, or an address where something of interest took place.

Spatial data: Geocoding

At its most basic, something can be geocoded by knowing its latitude and longitude; for example, the capital of Texas, Austin, is located at 30.274694°N latitude and -97.74036°W longitude.

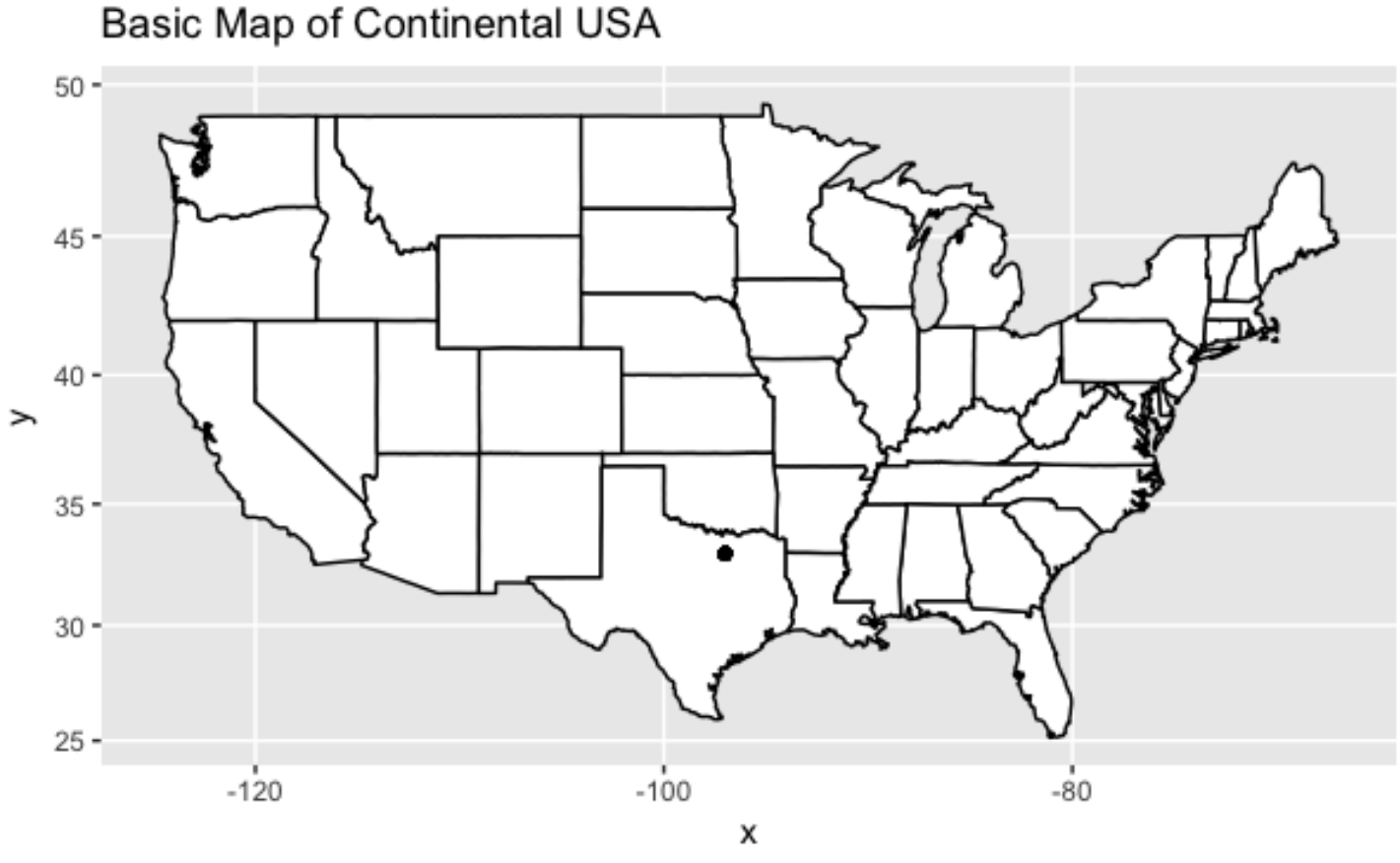
Spatial data layers



source: https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/s11-02-multiple-layer-analysis.html

Workshop: Spatial data in R

Map of the United States



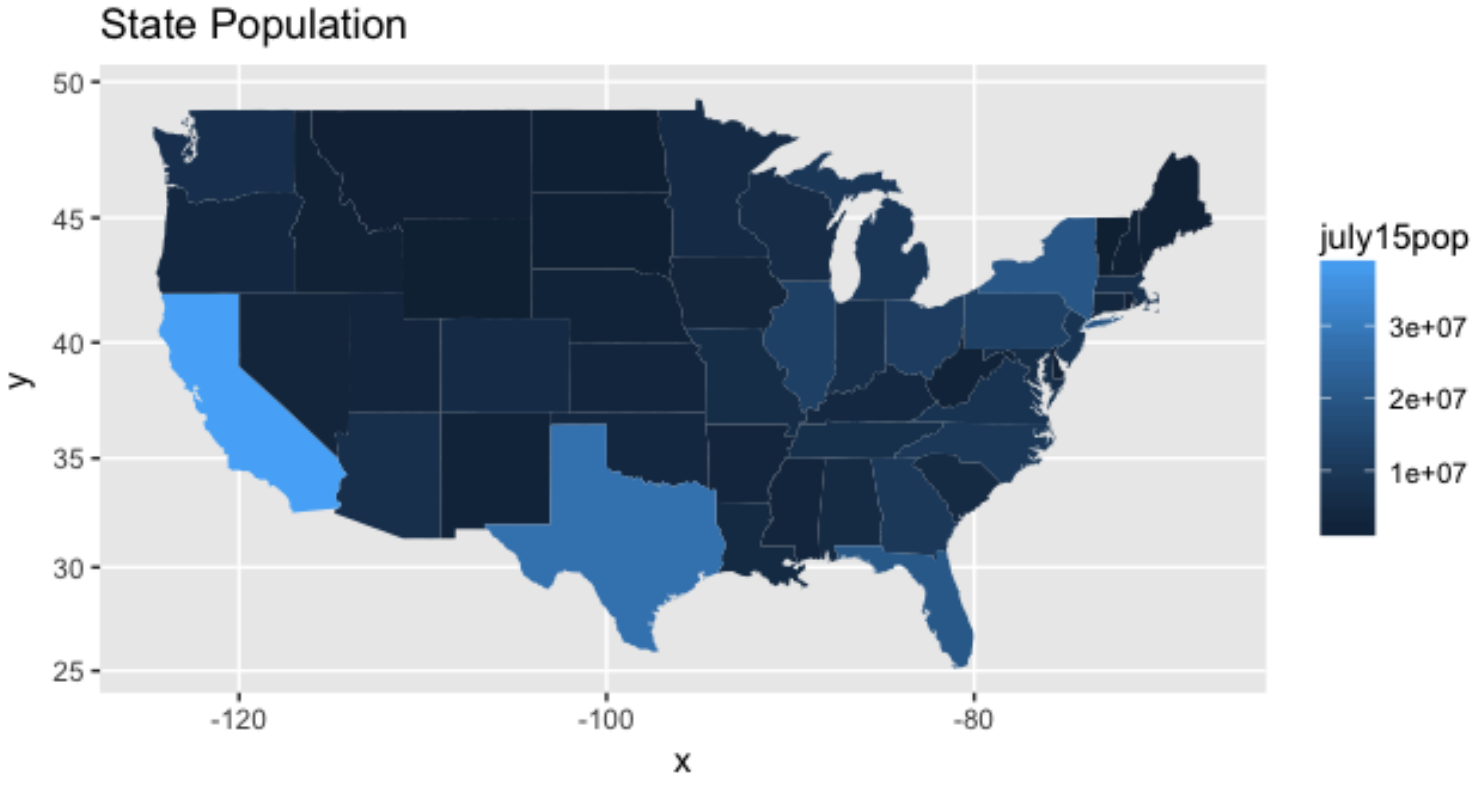
Map of the United States

```
library(ggplot2)
library(ggmap)

# Create US map data frame
USmapDF <- data.frame(state.name, stringsAsFactors=F)
USmapDF$state=tolower(USmapDF$state.name)

# Create Black and White US map
map.us=ggplot(USmapDF,aes(map_id=state))
map.us=map.us+geom_map(map=us,fill="light yellow",
color="black")
map.us=map.us+expand_limits(x=us$long,y=us$lat)
map.us=map.us+coord_map()+ggtitle("Basic Map of
Continental USA")
map.us
```

Map of the United States



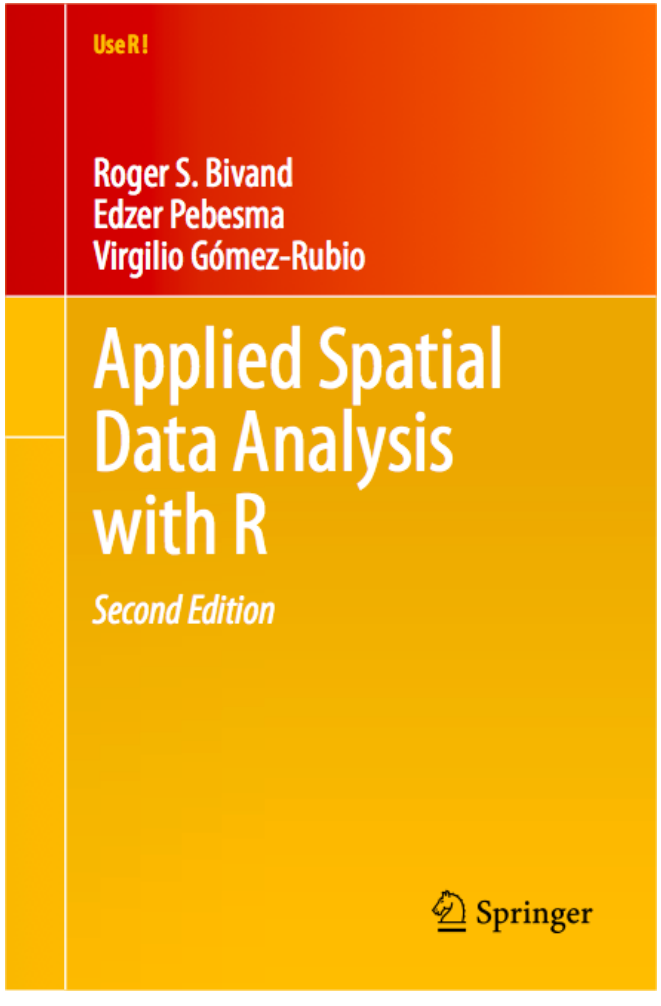
Map of the United States

```
# Read in 2015 Census population data by states
dfStates <- readCensus()

# make sure everything is lowercase
dfStates$state <- tolower(dfStates$stateName)

# Create Color US Population map
map.popColor <- ggplot(dfStates, aes(map_id = state))
map.popColor <- map.popColor + geom_map(map = us,
aes(fill=july15pop))
map.popColor <- map.popColor + expand_limits(x = us$long,
y = us$lat)
map.popColor <- map.popColor+ coord_map() +
ggtitle("State Population")
map.popColor
```

Further reading



Bivand, R.S., Pebesma, E.J., Gomez-Rubio, V. and Pebesma, E.J., 2008. *Applied spatial data analysis with R* (Vol. 747248717). New York: Springer.