Spatial Analysis with R: A short guide

PS236B Section

schoi@berkeley.edu

February 28, 2014

This section

- does not cover a lot of things
- will not make you a spatial data analyst

But

- helps you understand how R reads spatial data
- covers (very) basic tools useful for students interested in designs exploiting variation in geography
- includes QnA session.

the bible

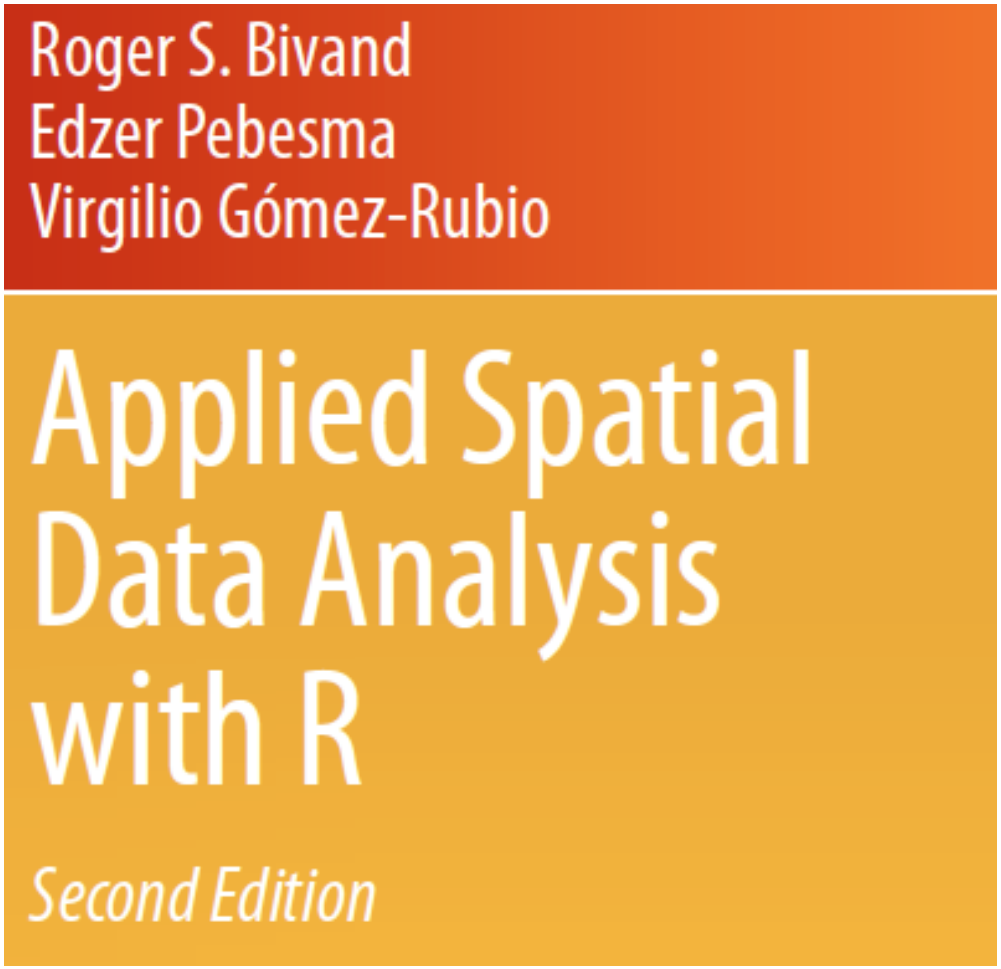


Figure: an electronic version freely available at Oskicat

Why spatial data in R?

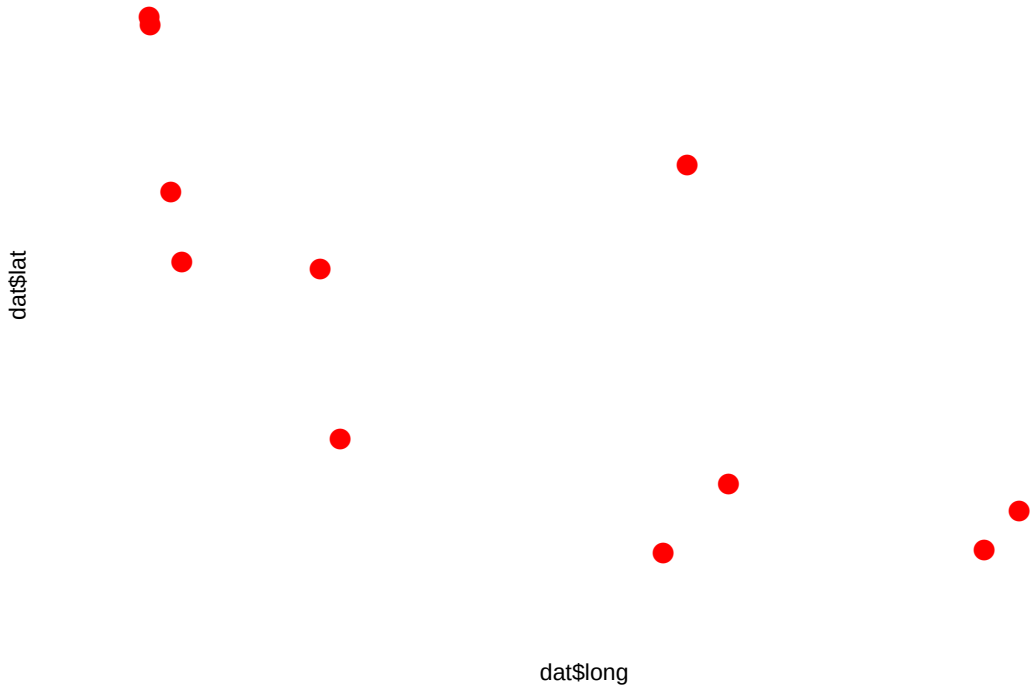
- free and easy
- awesome plots.
- R packages keep developing..
- replication friendly
- statistical analysis can be easily done
- QnA forum available online

Understanding spatial data

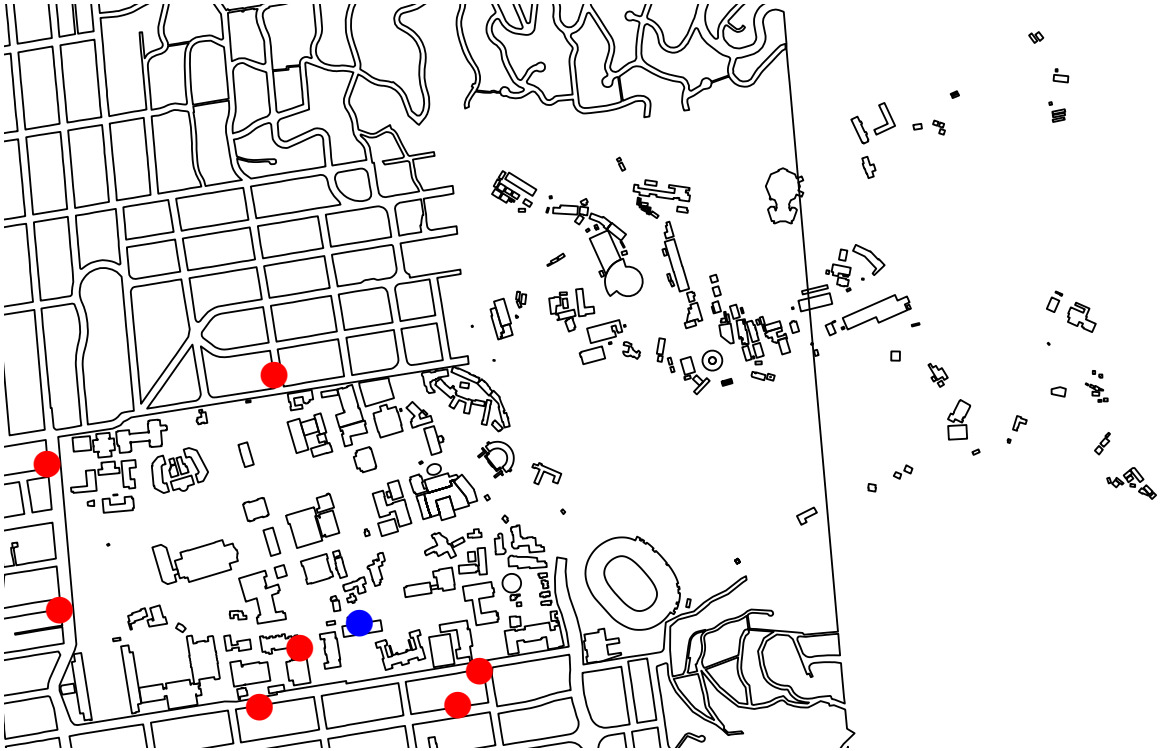
Types of spatial data:

- Point
- Line
- Polygon
- Grid

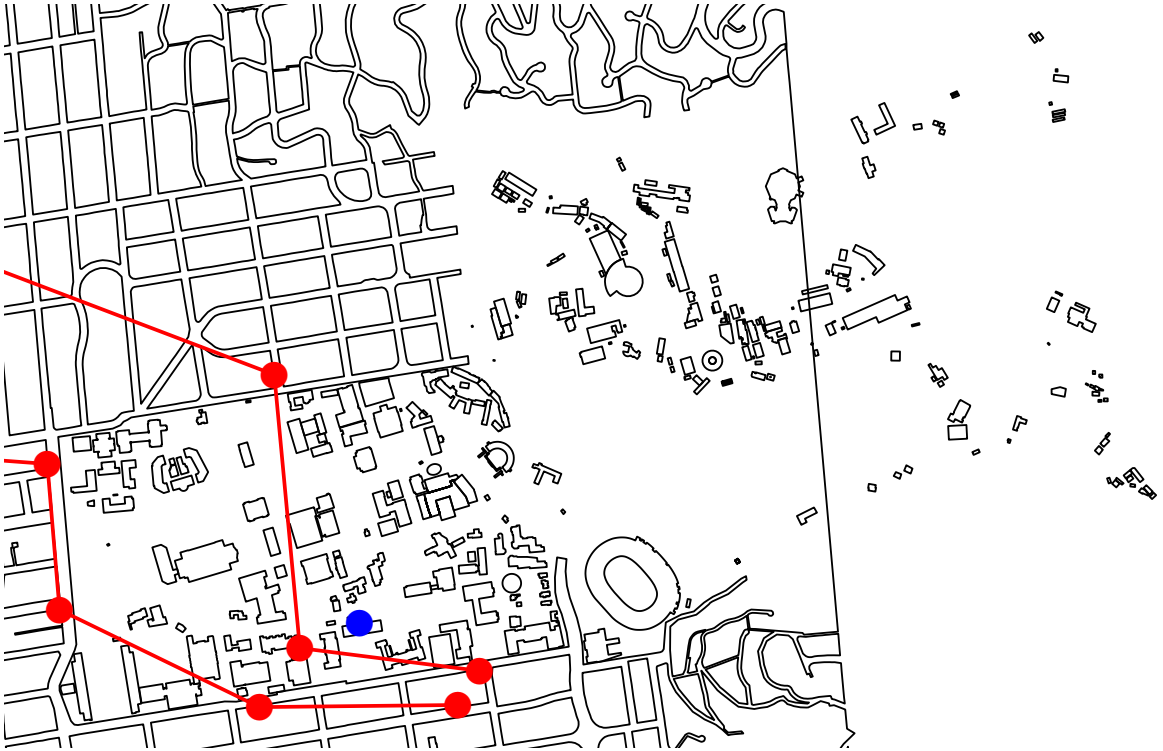
Understanding spatial data



Understanding spatial data



Understanding spatial data



Spatial Class in R

Spatial objects in sp package: two baseline slots

- bbox: min/max coordinates
- proj4string: coordinate reference system (CRS)
- how to represent a point on a globe

```
> slot(uc_bd, "proj4string")
CRS arguments:
 +proj=lcc +lat_1=37.06666666666667 +lat_2=38.43333333333333 +lat_0=36.5
+lon_0=-120.5 +x_0=2000000 +y_0=500000.0000000001 +datum=NAD83 +units=us-ft
+no_defs +ellps=GRS80 +towgs84=0,0,0
```

Spatial Class in R

```
Slot "bbox":
      min     max
x 6051919 6060426
y 2143364 2148529

Slot "proj4string":
CRS arguments:
 +proj=lcc +lat_1=37.06666666666667 +lat_2=38.43333333333333 +lat_0=36.5
+lon_0=-120.5 +x_0=2000000 +y_0=500000.0000000001 +datum=NAD83 +units=us-ft
+no_defs +ellps=GRS80 +towgs84=0,0,0
```

- Any `Spatial*` class is the subclass of `Spatial`.
- Any `Spatial*` class thus has two baseline slots, and more.

Coordinate systems and Projection

- one point on a space does not have any meaning unless it is referenced to a coordinate system
- getting the right projection: essential to spatial analysis
- Geographic Coordinate System: Long/Lat (e.g. WGS 1984)
- Projected coordinate systems: projecting the round surface of the earth onto a flat 2D space.
- Which coordinate system should you use?
- See code.

SpatialPoints

- the first subclass of `Spatial`.
- a single location in a two-dimensional space: a pair of numbers
- two baseline slots + `coords` slot
- See code.

SpatialPointsDataFrame

- attribute data
- unique location : unique row names
- spatial data + attribute data by matching on ID
- See code.

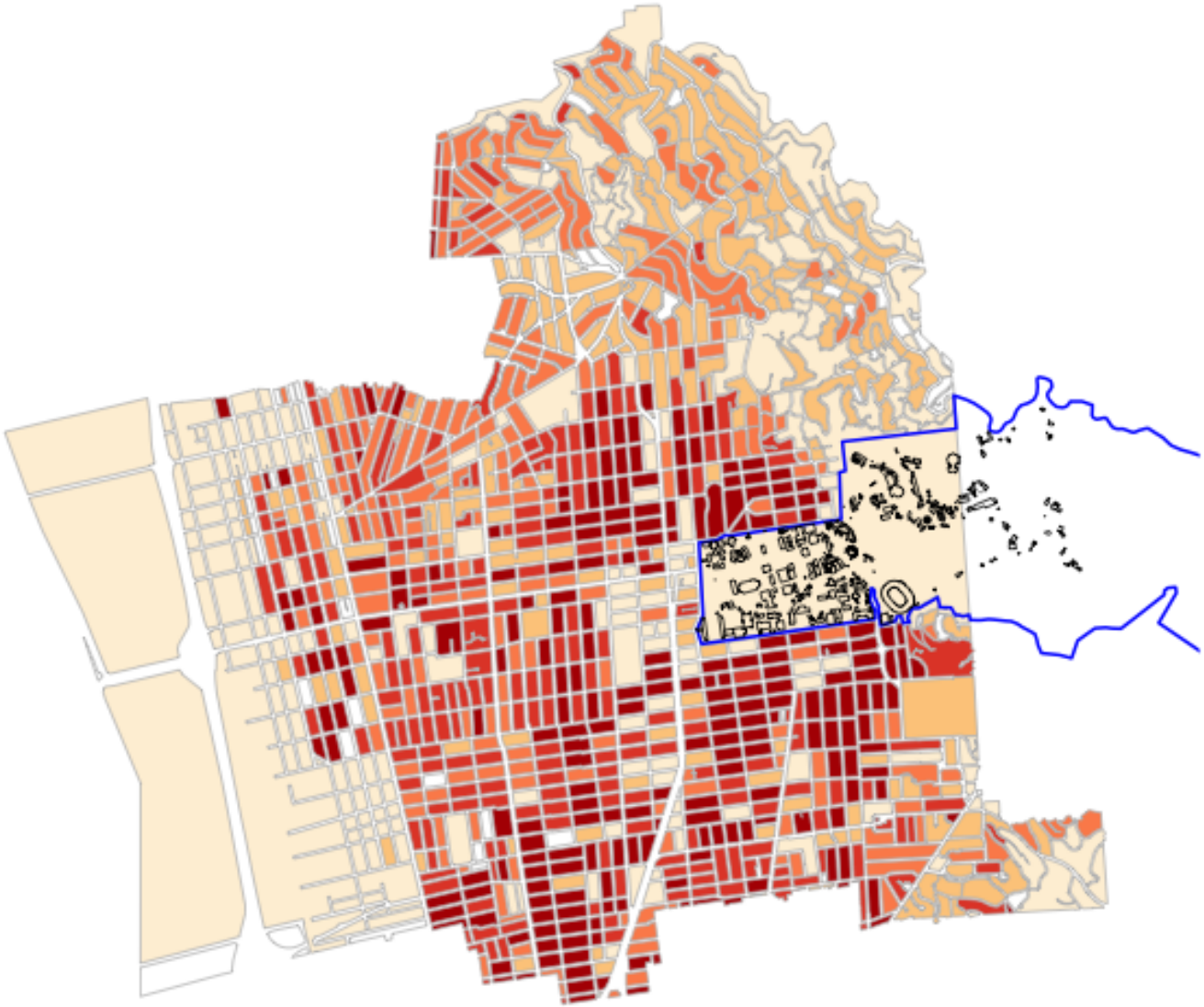
SpatialLines, SpatialLinesDataFrame

- connecting a sequence of points until first `NA`
- Line - Lines - SpatialLines
- See code.

`SpatialPolygons`, `SpatialPolygonsDataFrame`

- connecting a sequence of points
- the end point = the starting point
- Polygon - Polygons - SpatialPolygons
- understanding `list` class in R is important
- See code.

Attribute Data



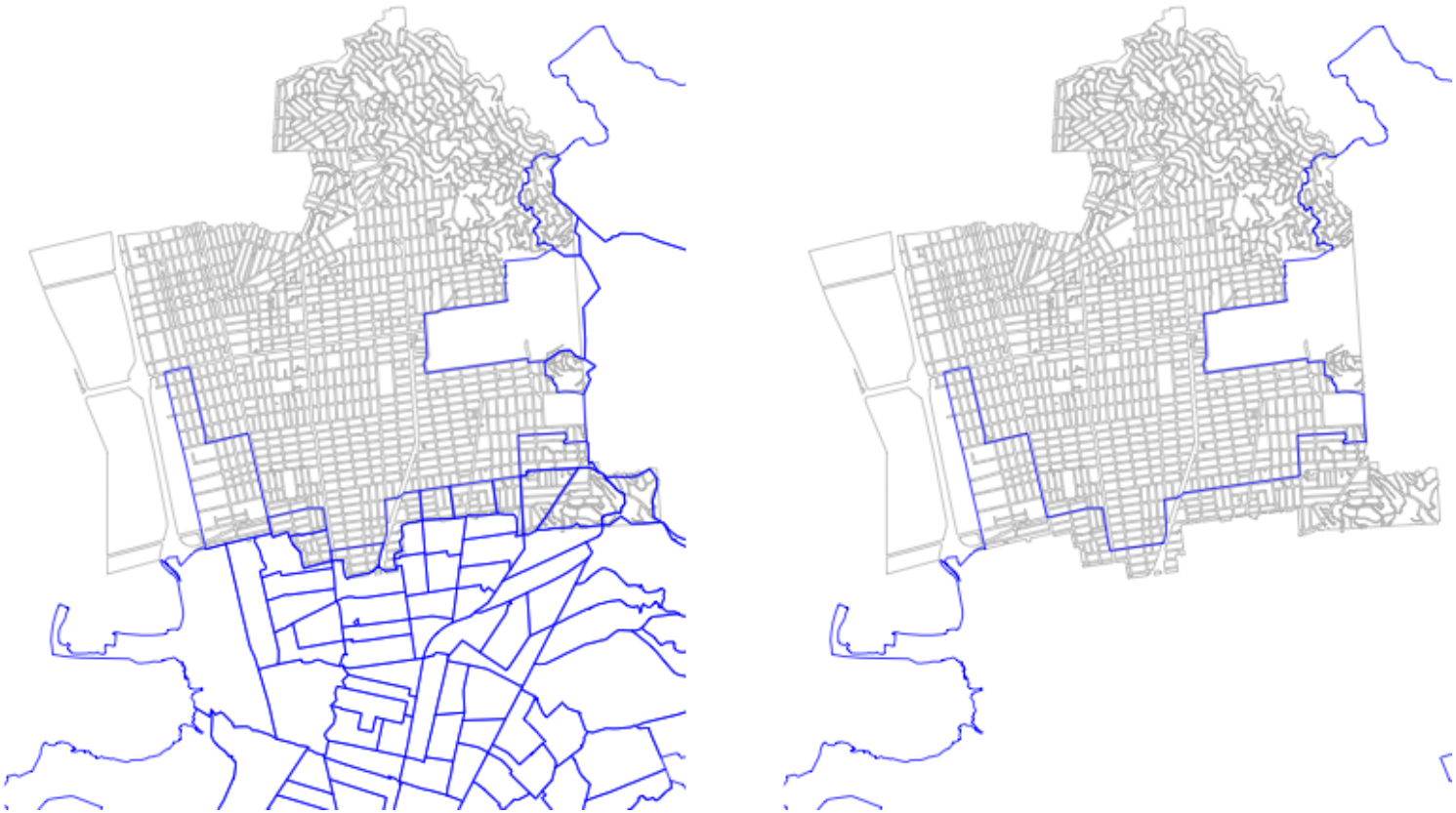
Manipulating Spatial Data

- now it's getting a little complicated
- See code.

Useful functions for spatial data manipulation

- `gDistance, gBuffer, UnaryUnion, overlay`
- See code.

Manipulating SpatialPolygonsData



Geographic Regression Discontinuity

- Define boundary regions
- Do you expect any sorting around the boundary?
- See code.

Q & A

Any questions?