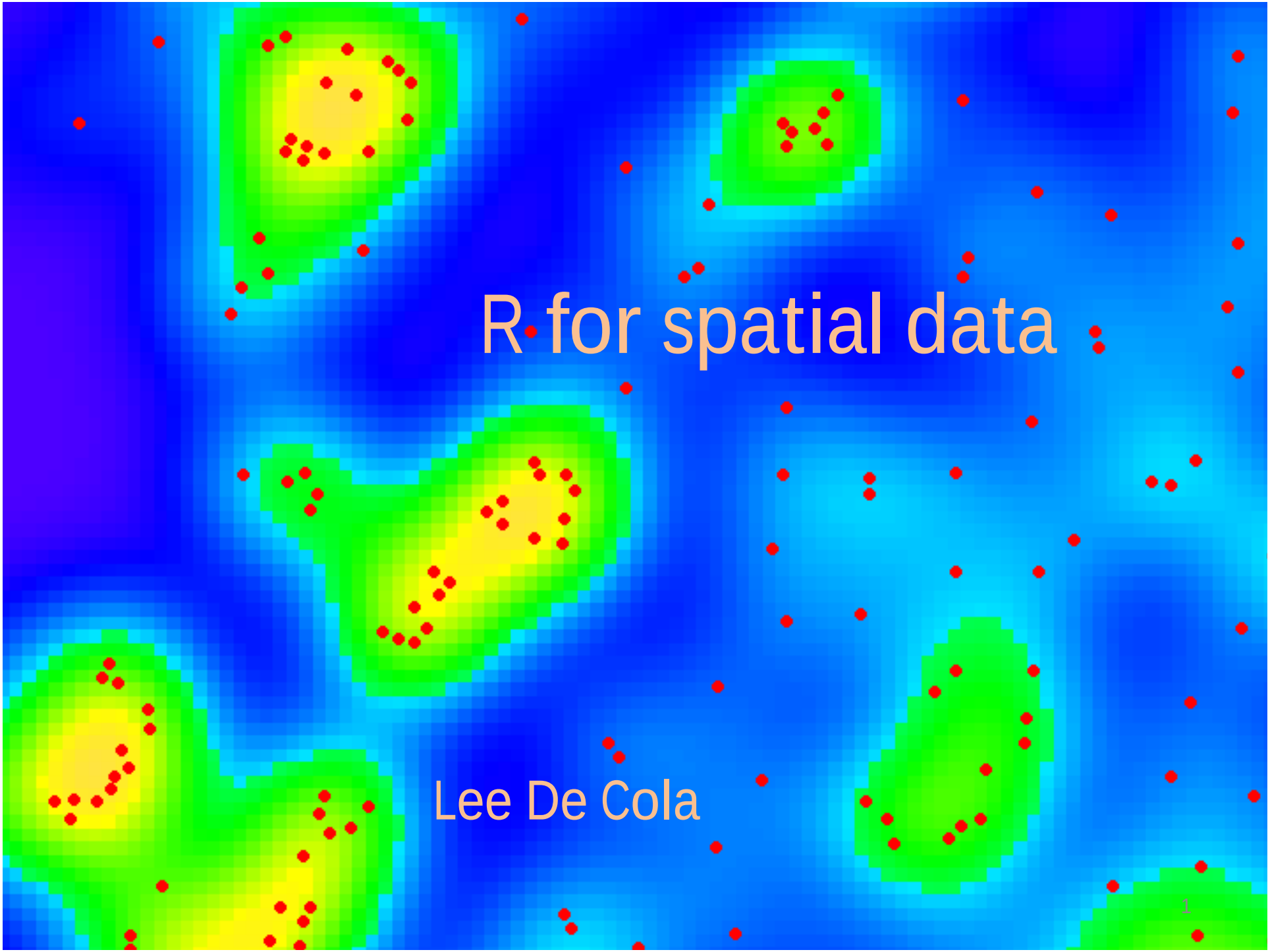
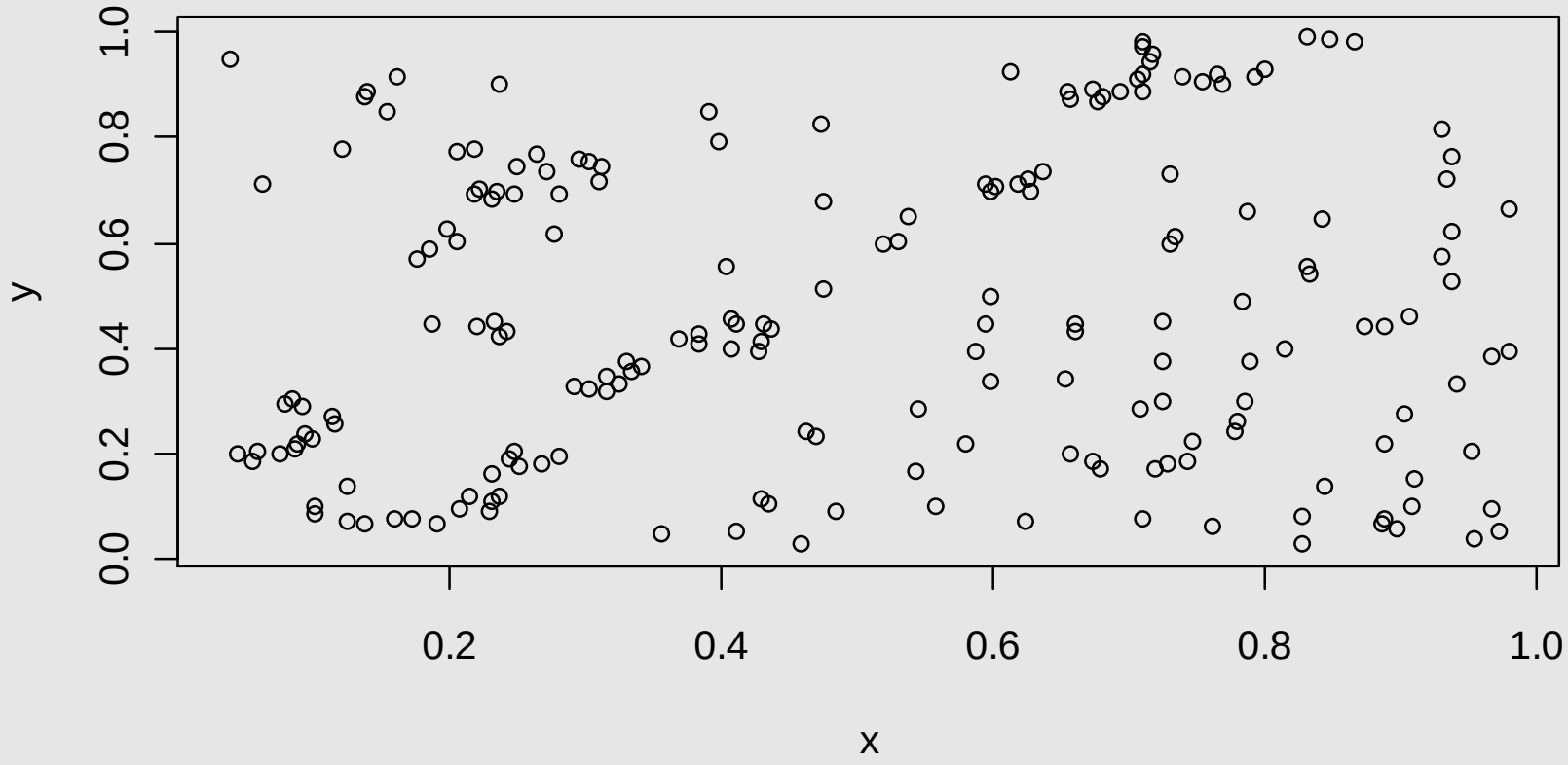
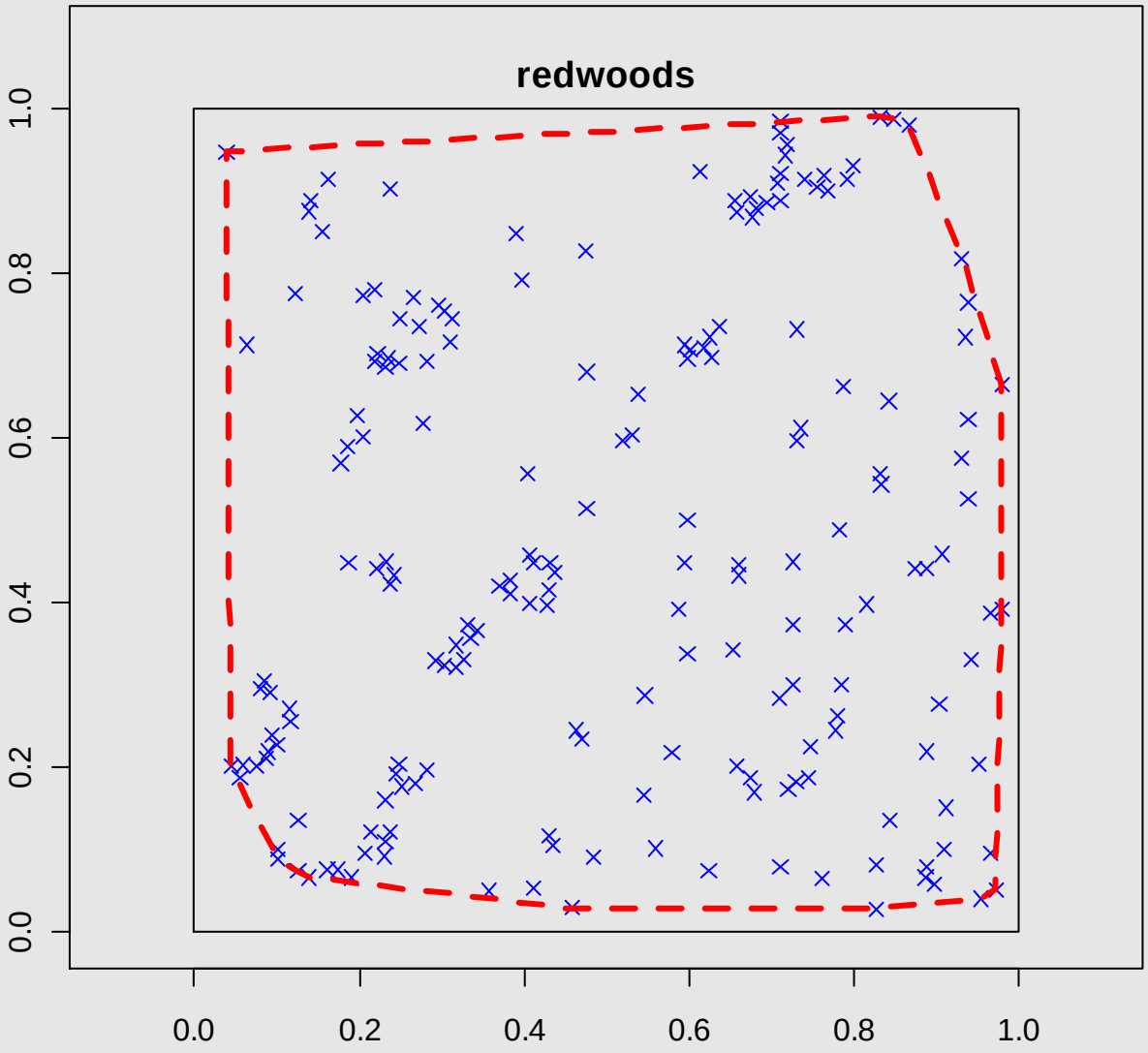

R for spatial data
Lee De Cola

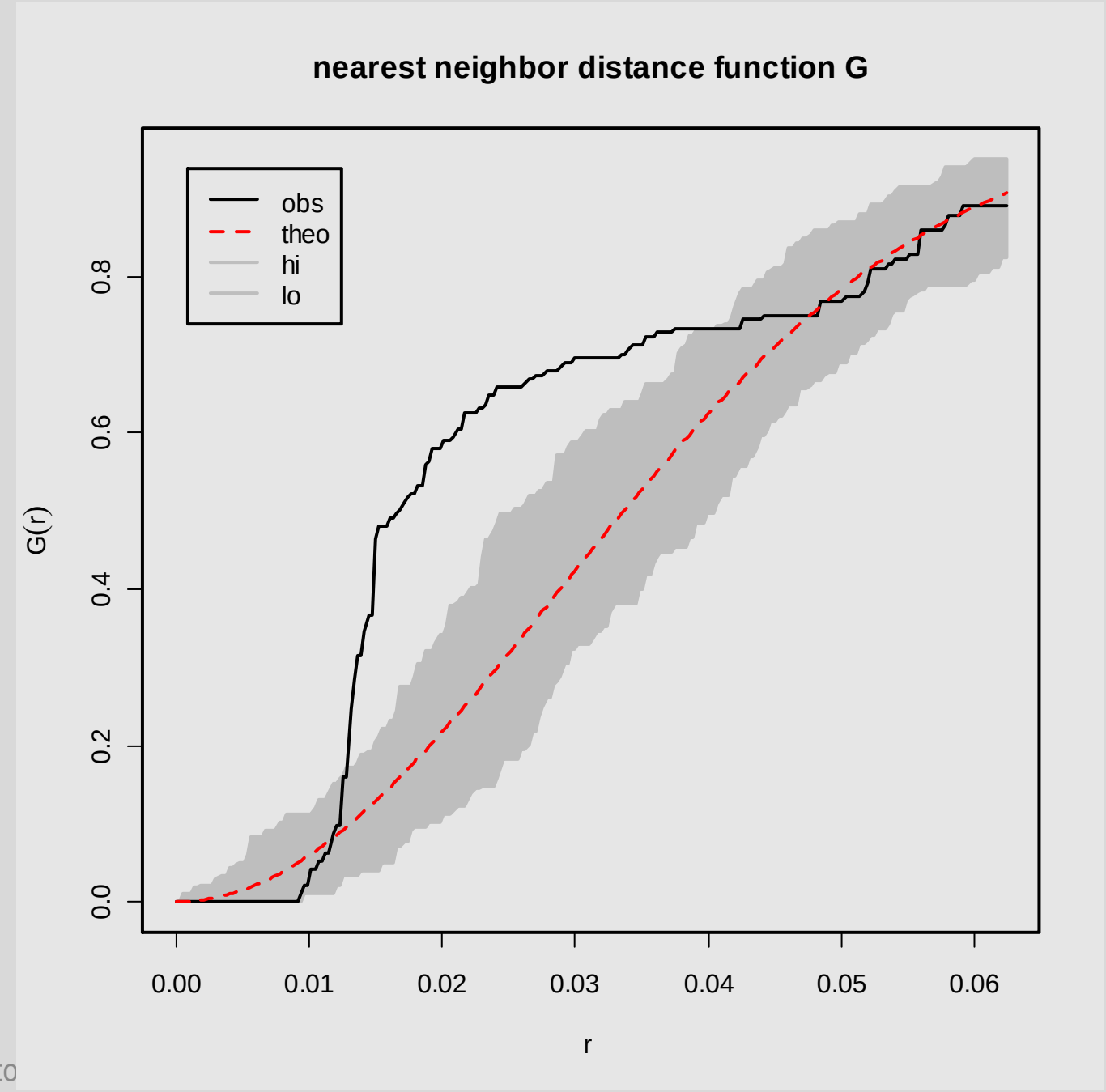
points

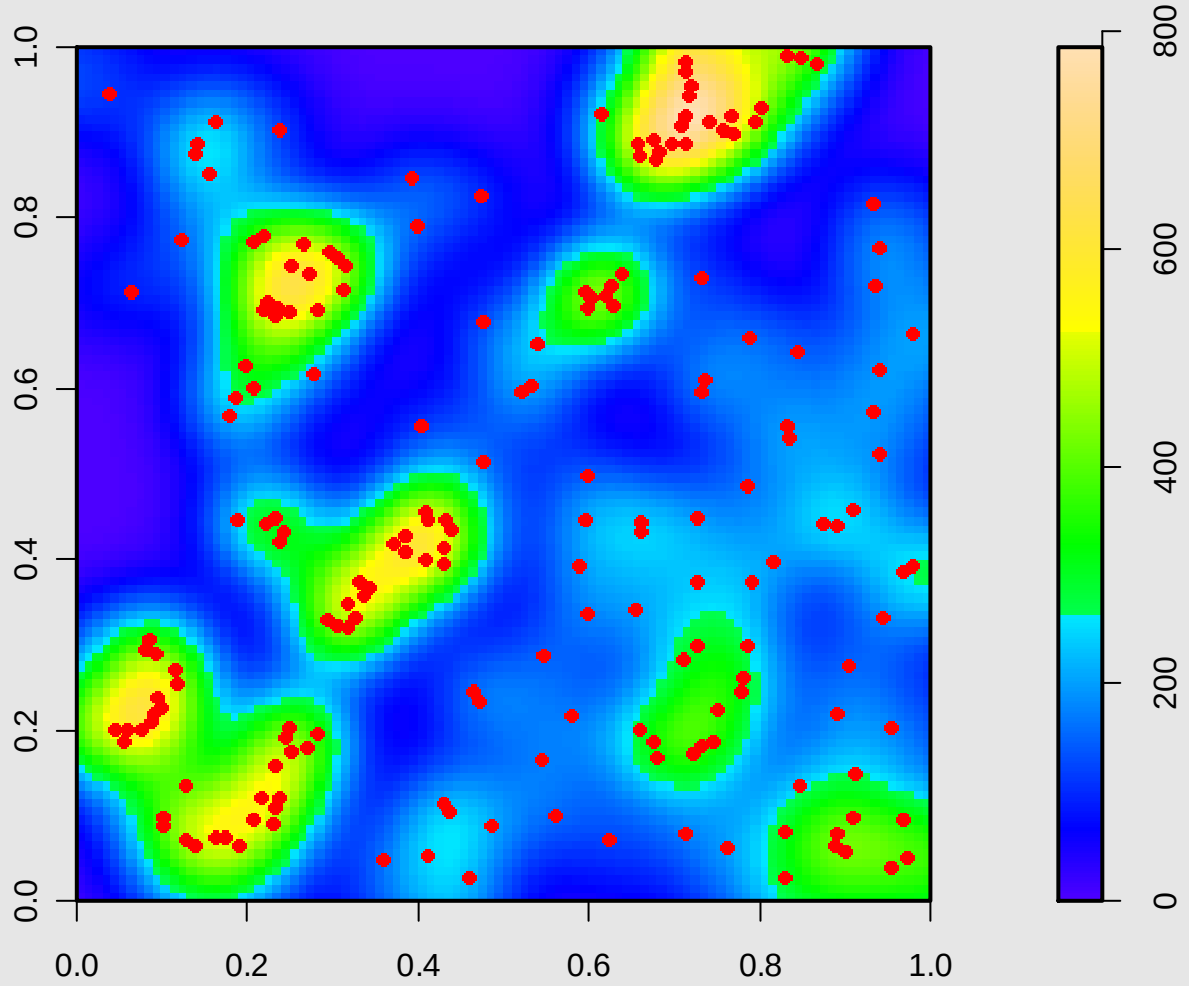
planar point pattern

```
List of 5
 $ window    :List of 4
  ..$ type  : chr "rectangle"
  ..$ xrange: num [1:2] 0 1
  ..$ yrange: num [1:2] 0 1
  ..$ units :List of 3
  .. ..$ singular  : chr "unit"
  .. ..$ plural    : chr "units"
  .. ..$ multiplier: num 1
  .. ..- attr(*, "class")= chr "units"
  ..- attr(*, "class")= chr "owin"
 $ n        : int 195
 $ x        : num [1:195] 0.931 0.939 0.935 0.98 0.787 ...
 $ y        : num [1:195] 0.818 0.764 0.722 0.665 0.661 ...
 $ markformat: chr "none"
 - attr(*, "class")= chr "ppp"
```


y
0.0
0.2
0.4
0.6
0.8
1.0
0.2
0.4
0.6
0.8
1.0
x


redwoods
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0

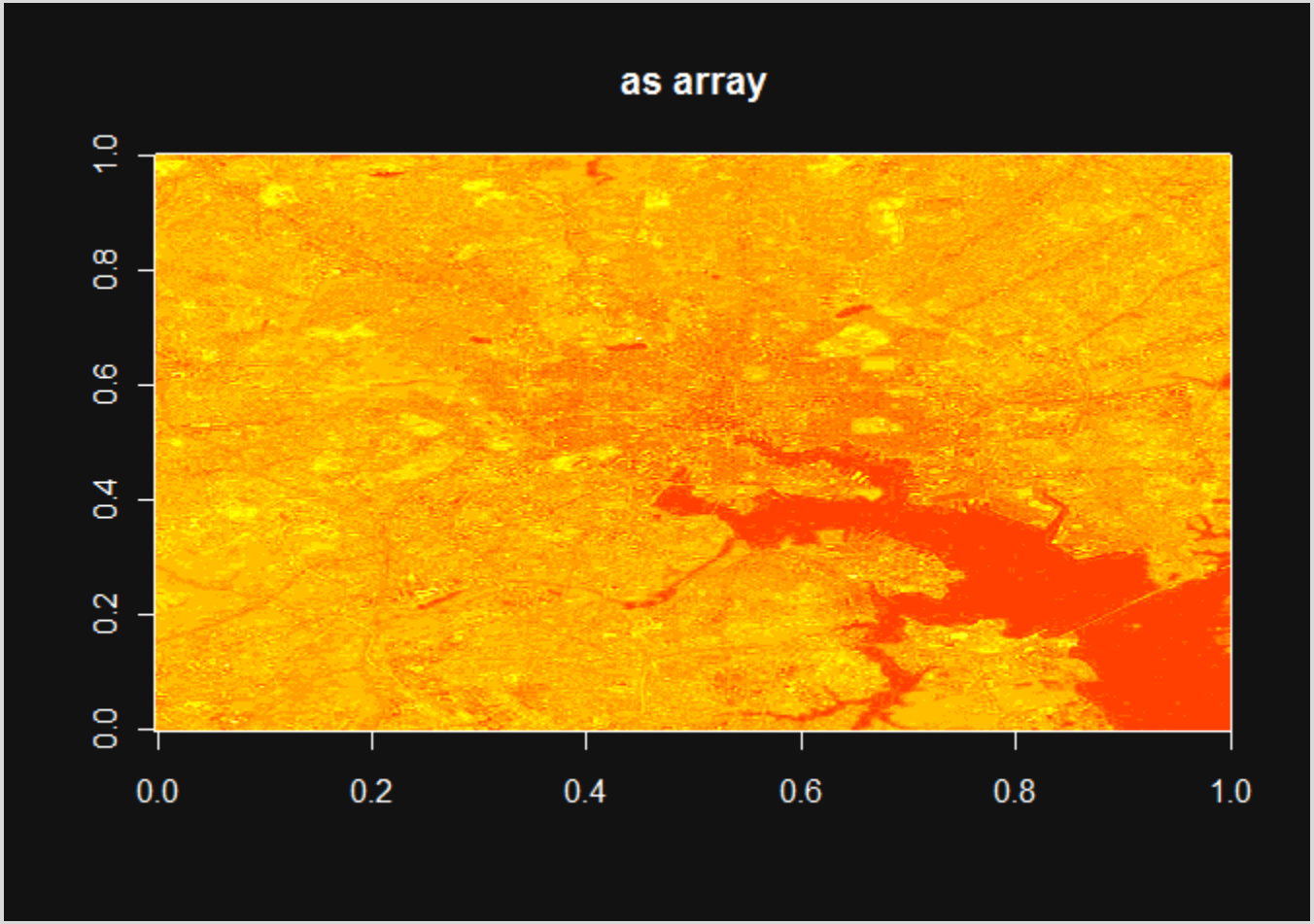

nearest neighbor distance function G
obs
theo
hi
lo
G(r)
0.0
0.2
0.4
0.6
0.8
0.00
0.01
0.02
0.03
0.04
0.05
0.06
r

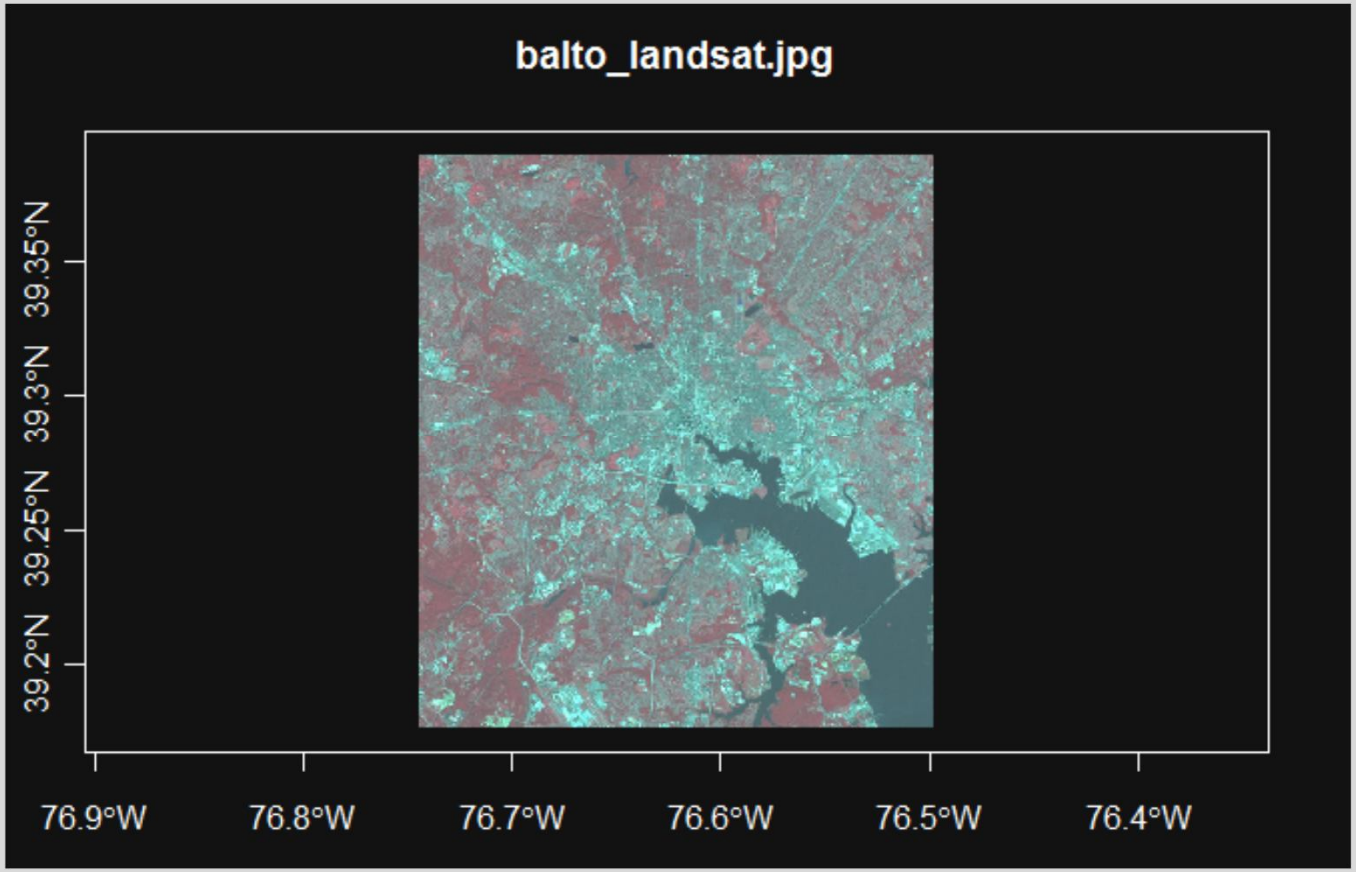

bandwidth 0.05
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800

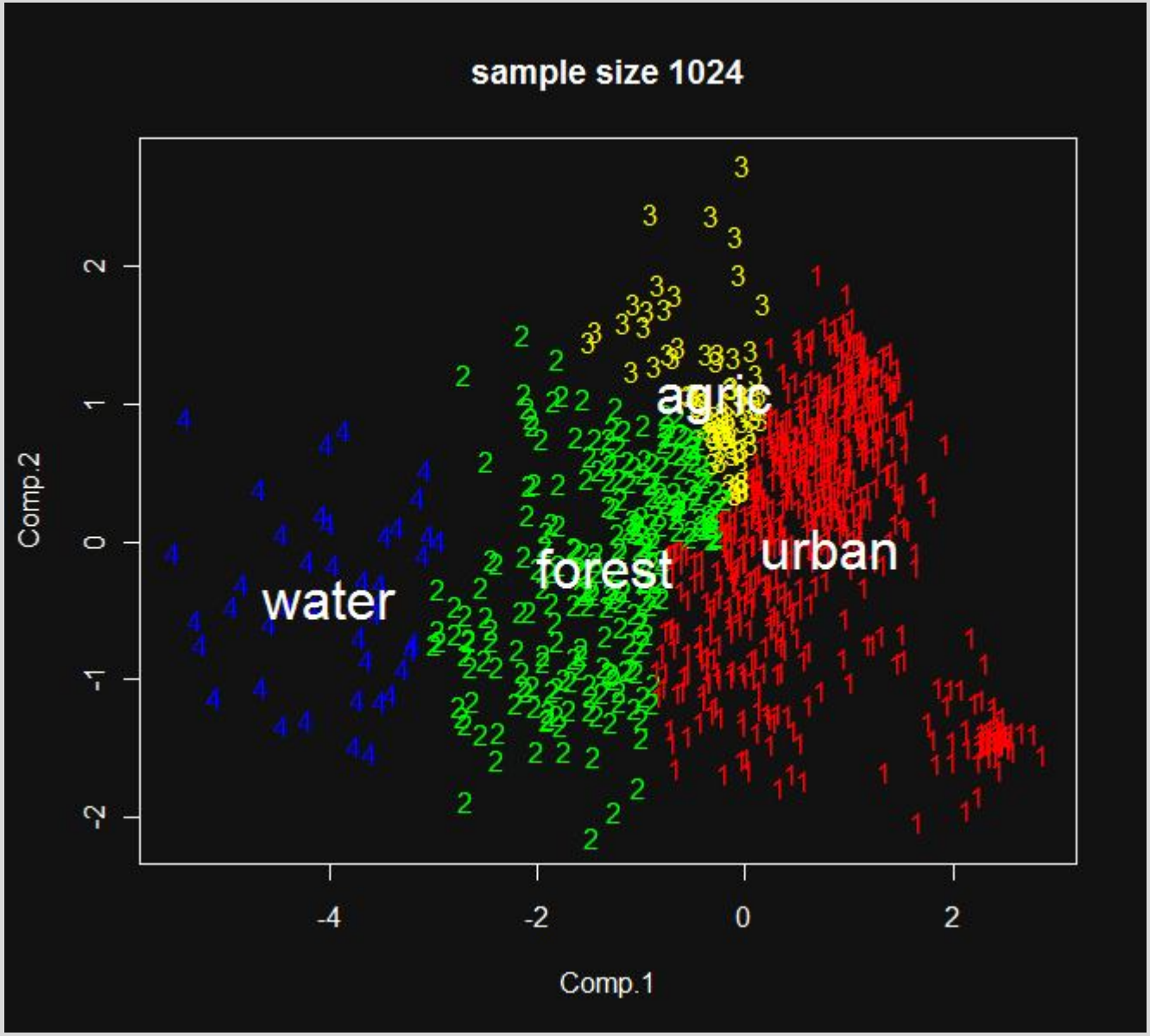
grids

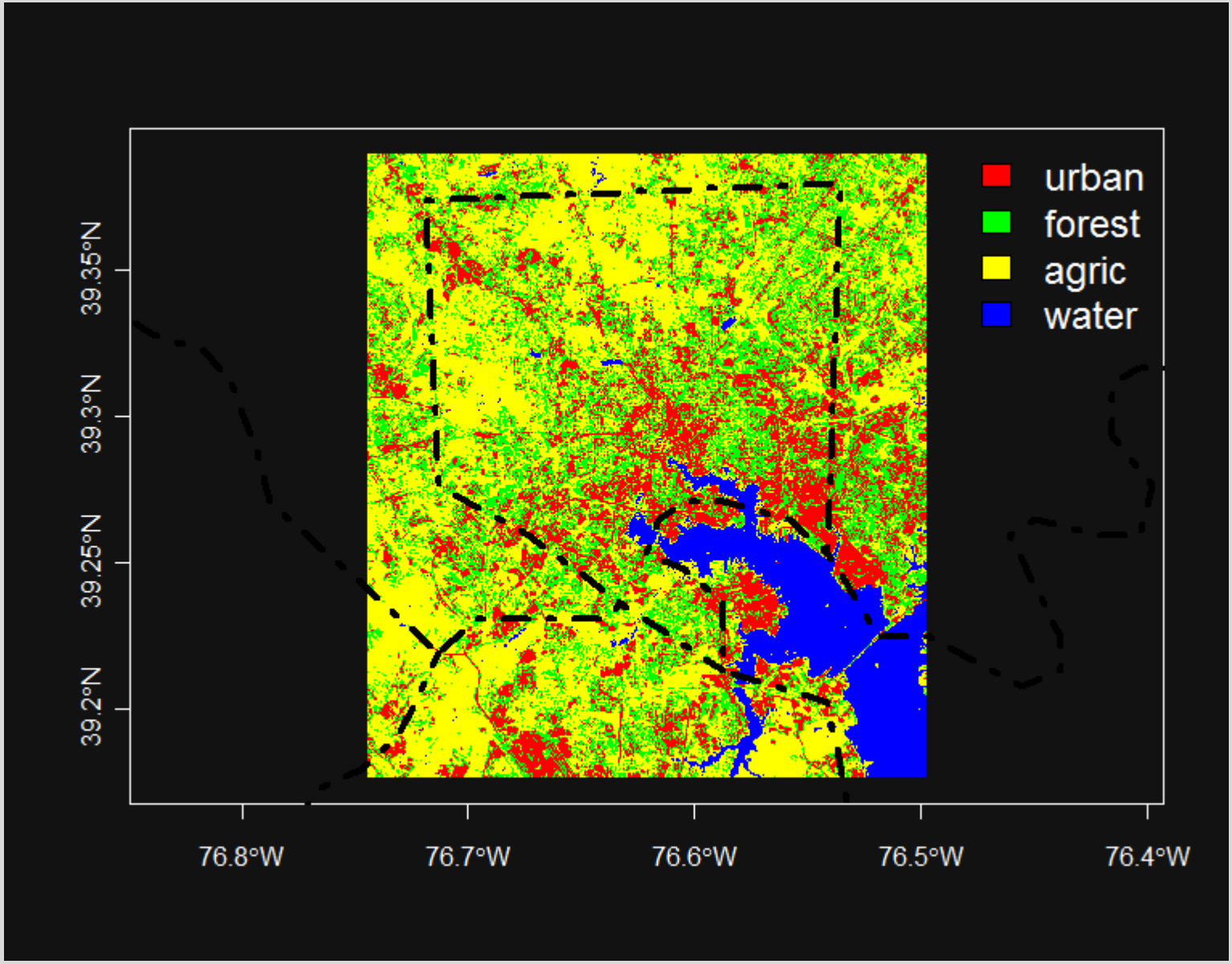
```
Formal class 'SpatialGridDataFrame' [package "sp"] with 6 slots
  ..@ data       :'data.frame': 891667 obs. of  3 variables:
  .. ..$ band1: int [1:891667] 147 148 144 142 152 166 160 146 131 140 ...
  .. ..$ band2: int [1:891667] 124 125 121 119 131 145 140 126 111 120 ...
  .. ..$ band3: int [1:891667] 132 133 131 129 140 154 151 137 122 131 ...
  ..@ grid       :Formal class 'GridTopology' [package "sp"] with 3 slots
  .. .. ..@ cellcentre.offset: Named num [1:2] -76.7 39.2
  .. .. .. ..- attr(*, "names")= chr [1:2] "x" "y"
  .. .. ..@ cellsize         : num [1:2] 0.000278 0.000213
  .. .. ..@ cells.dim        : int [1:2] 889 1003
  ..@ grid.index : int(0)
  ..@ coords     : num [1:2, 1:2] -76.7 -76.5 39.2 39.4
  .. ..- attr(*, "dimnames")=List of 2
  .. .. ..$ : NULL
  .. .. ..$ : chr [1:2] "x" "y"
  ..@ bbox       : num [1:2, 1:2] -76.7 39.2 -76.5 39.4
  .. ..- attr(*, "dimnames")=List of 2
  .. .. ..$ : chr [1:2] "x" "y"
  .. .. ..$ : chr [1:2] "min" "max"
  ..@ proj4string:Formal class 'CRS' [package "sp"] with 1 slots
  .. .. ..@ projargs: chr " +proj=longlat +ellps=GRS80 +datum=NAD83 +no_defs +towgs84=0,0,0"
```

SpatialGridDataFrame


as array
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0


balto_landsat.jpg
39.35°N
39.3°N
39.25°N
39.2°N
76.9°W
76.8°W
76.7°W
76.6°W
76.5°W
76.4°W


sample size 1024
agric
urban
forest
water
Comp.2
Comp.1
-2
-1
0
1
2
-4
-2
0
2

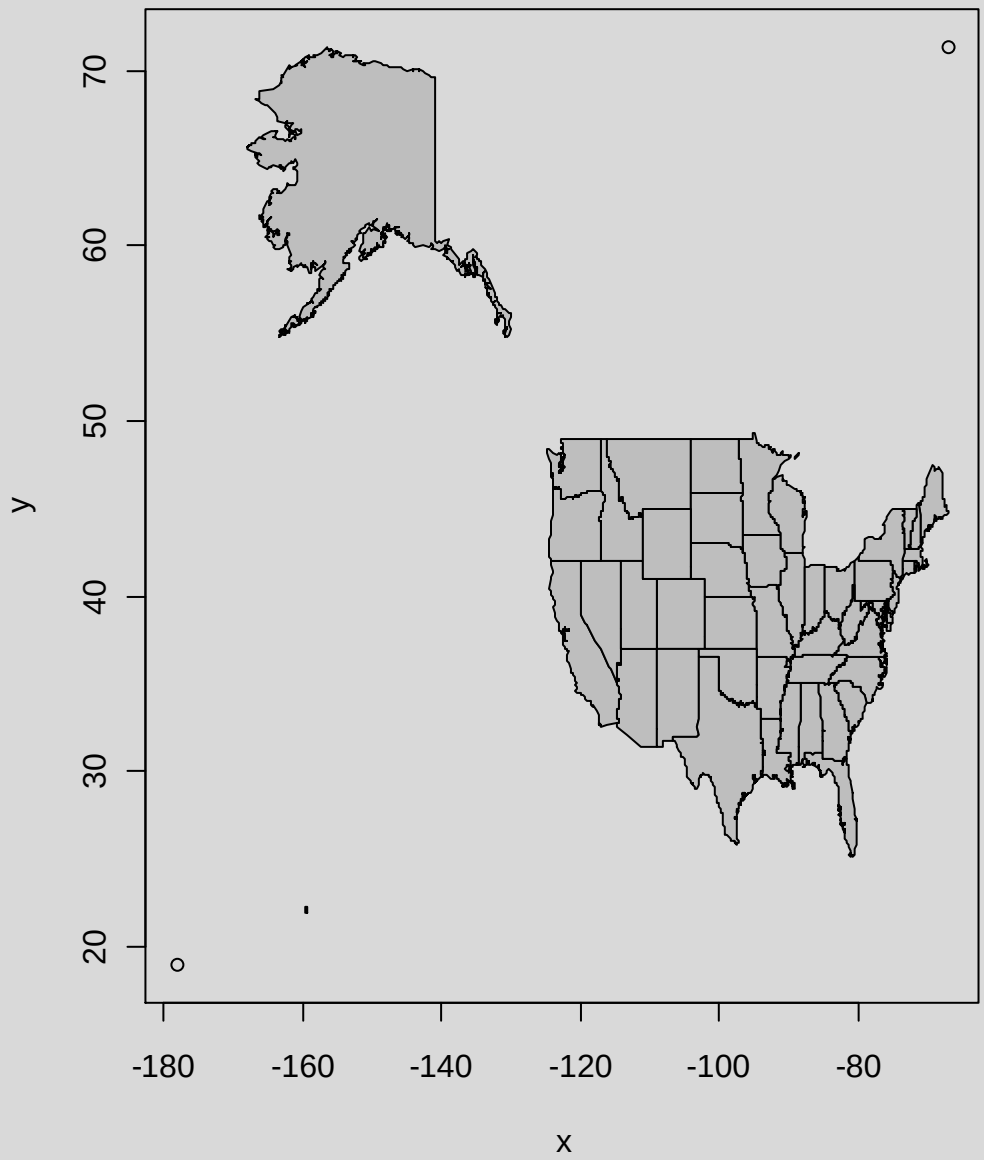

urban
forest
agric
water
39.35°N
39.3°N
39.25°N
39.2°N
76.8°W
76.7°W
76.6°W
76.5°W
76.4°W

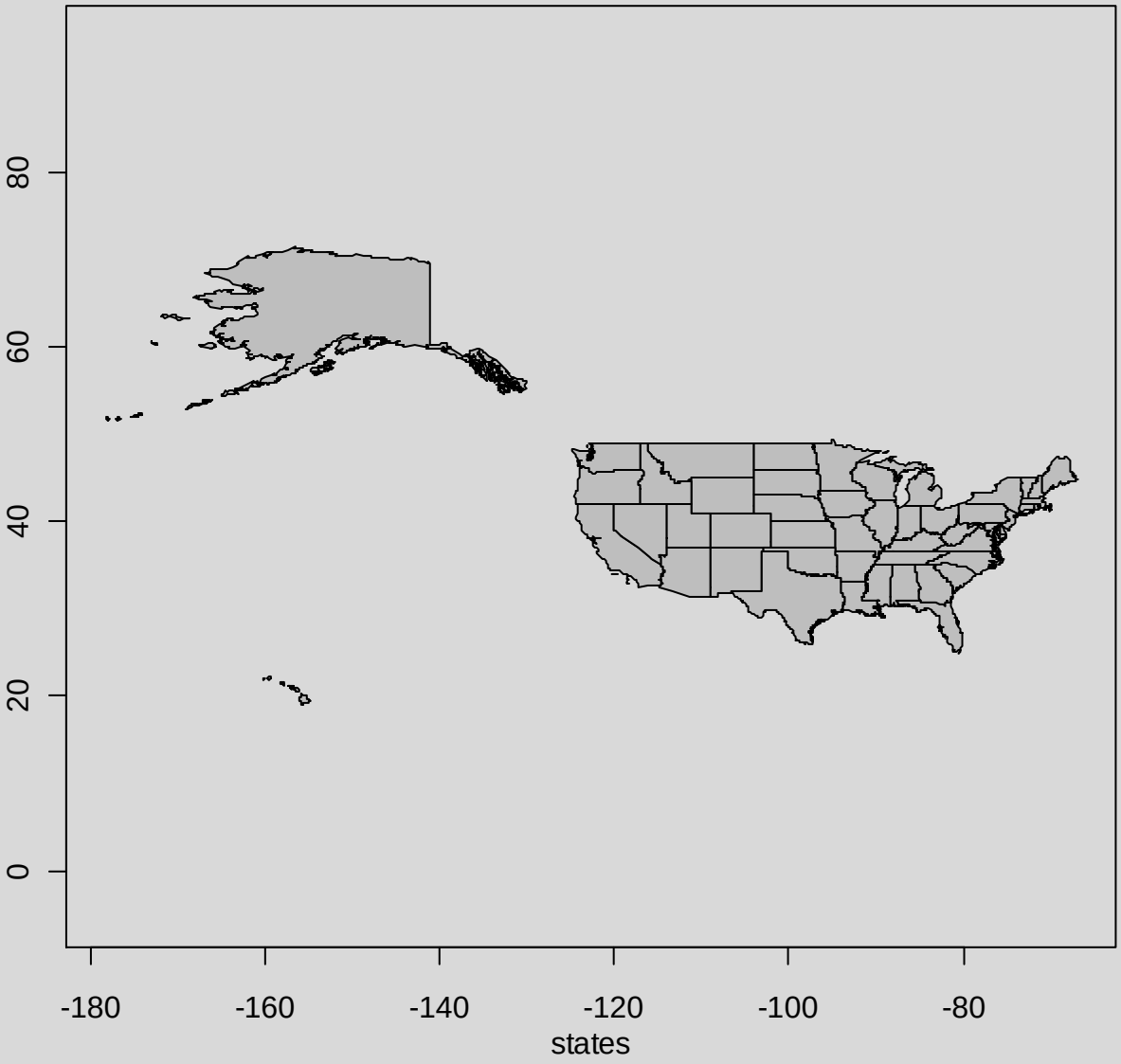
polygons

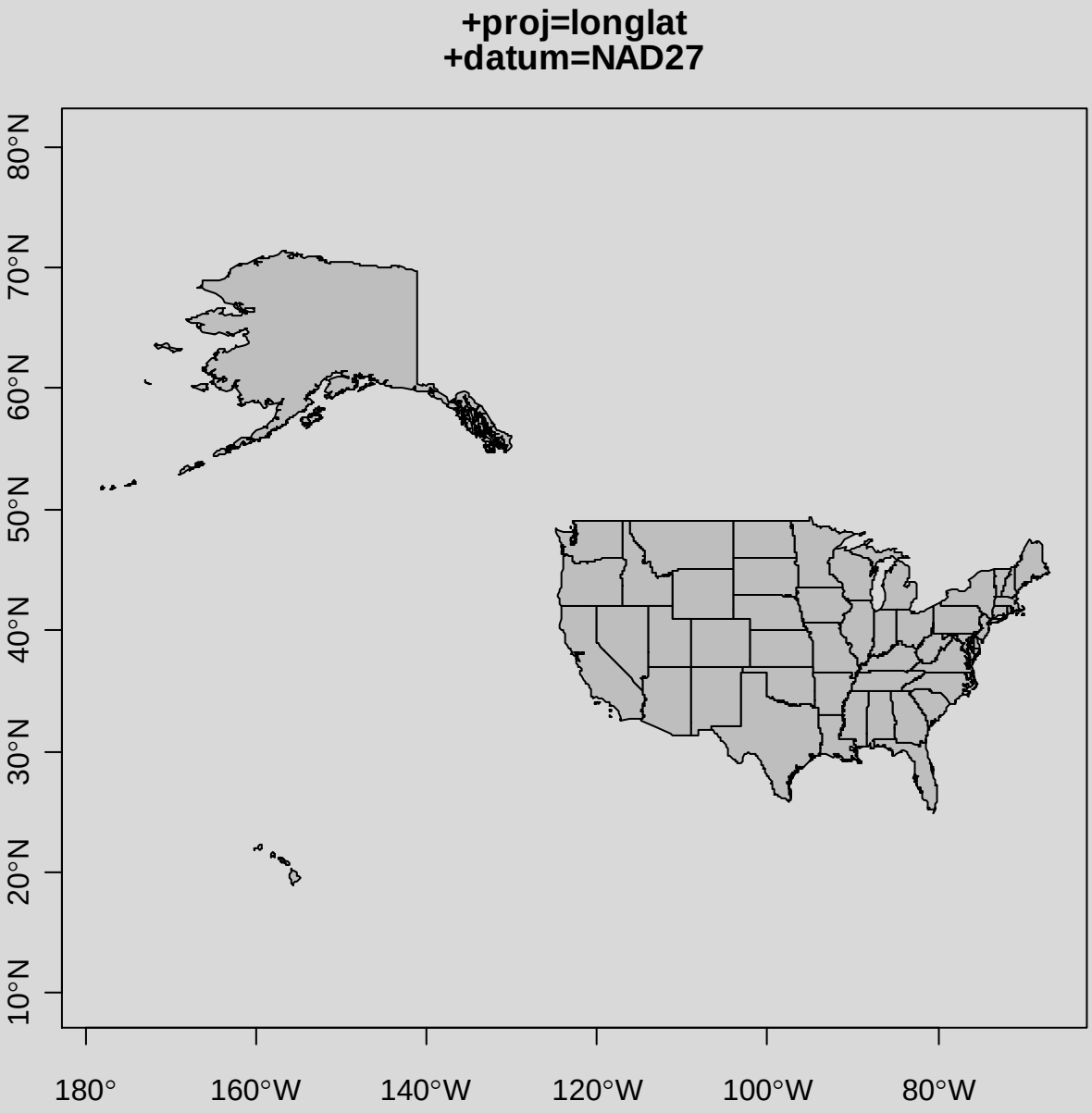
shape file to SpatialPolygonsDataFrame

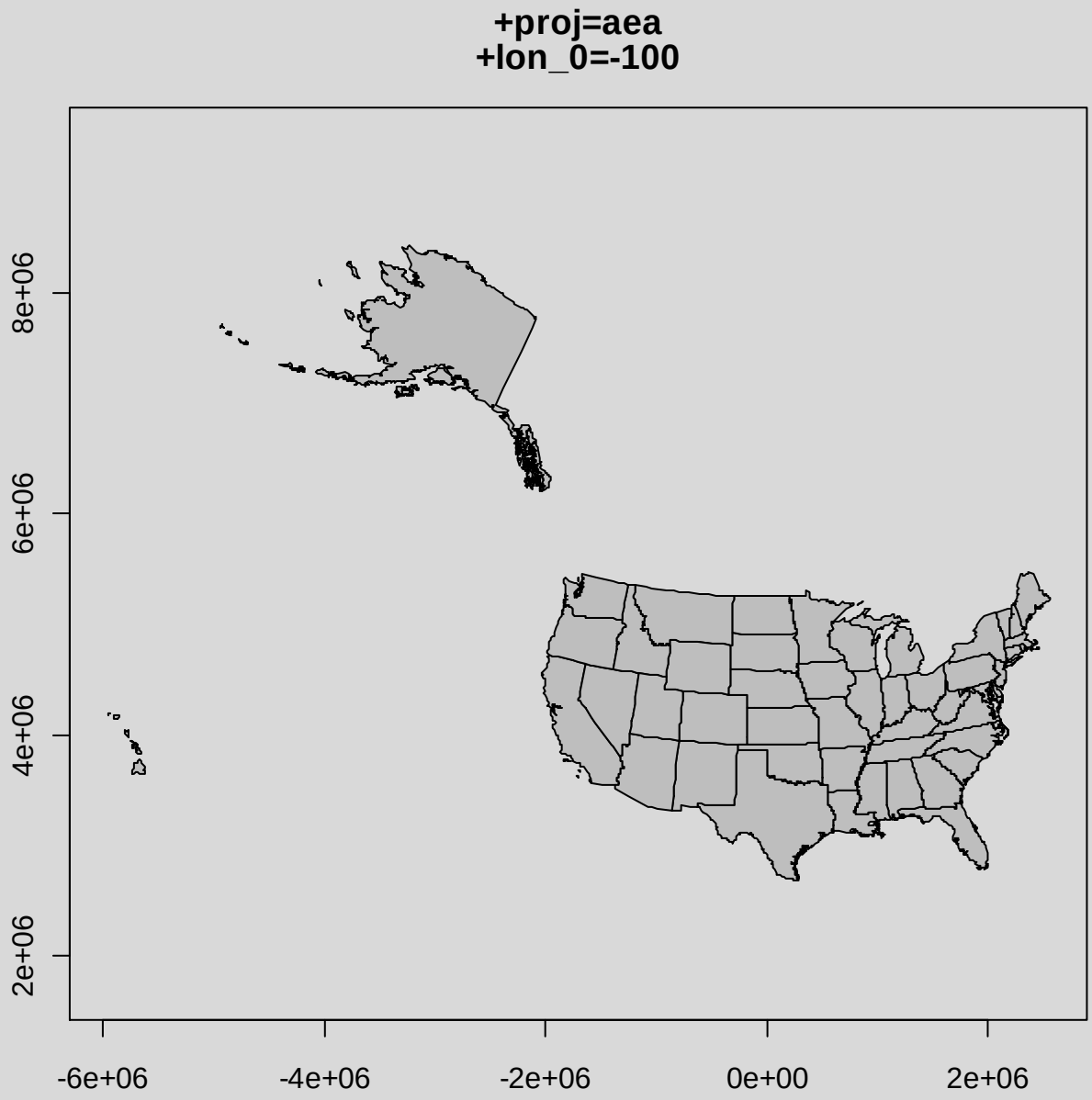
```
> map_sp <- readShapePoly(
+   paste(
+     paste(dirname, filename, sep = '\\'),
+     'shp',
+     sep = '.'
+   ),
+   IDvar = 'STATE_ABBR',
+   verbose = T
+ )
Shapefile type: Polygon, (5), # of Shapes: 51
> class(map_sp)
[1] "SpatialPolygonsDataFrame"
attr(,"package")
[1] "sp"
```

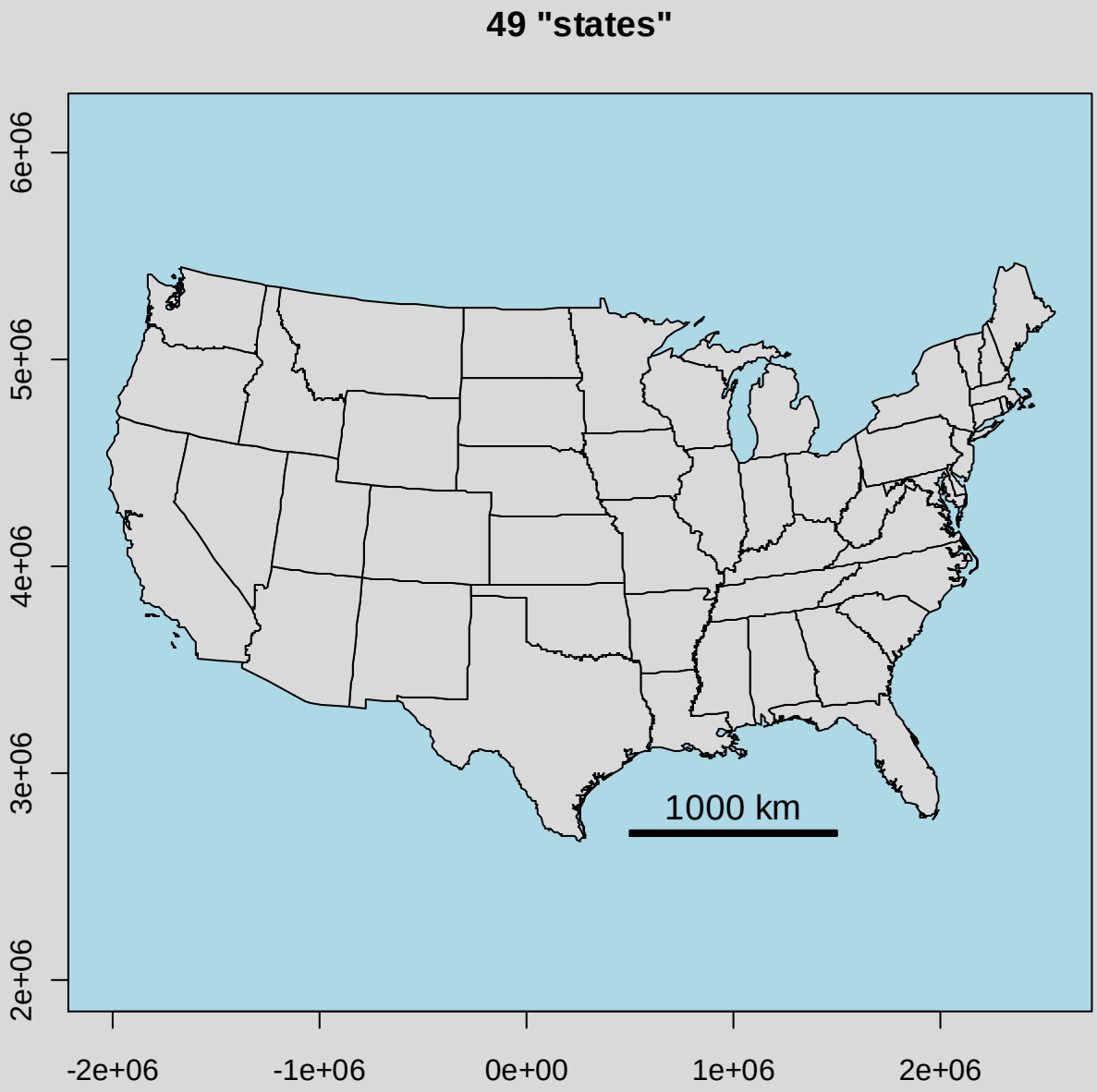
```
> str(map_sp, 2)
Formal class 'SpatialPolygonsDataFrame' [package "sp"] with 5 slots
  ..@ data       :'data.frame': 51 obs. of  10 variables:
  .. ..- attr(*, "data_types")= chr [1:10] "F" "C" "C" "C" ...
  ..@ polygons   :List of 51
  ..@ plotOrder  : int [1:51] 1 44 27 5 33 34 4 38 6 51 ...
  ..@ bbox       : num [1:2, 1:2] -178.2 18.9 -67 71.4
  .. ..- attr(*, "dimnames")=List of 2
  ..@ proj4string:Formal class 'CRS' [package "sp"] with 1 slots
> str(slot(map_sp, 'data'))
'data.frame':   51 obs. of  10 variables:
 $ AREA      : num  576594 51716 52913 113713 157776 ...
 $ STATE_NAME: Factor w/ 51 levels "Alabama","Alaska",..: 2 1 4 3 5 6 7 9 8 10 ...
 $ STATE_FIPS: Factor w/ 51 levels "01","02","04",..: 2 1 4 3 5 6 7 9 8 10 ...
 $ SUB_REGION: Factor w/ 9 levels "E N Cen","E S Cen",..: 6 2 9 4 6 4 5 7 7 7 ...
 $ STATE_ABBR: Factor w/ 51 levels "AK","AL","AR",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ POP1990   : num  550043 4040587 2350725 3665228 29760021 ...
 $ POP2000   : num  624523 4395481 2566938 4894006 33603430 ...
 $ POP90_SQMI: int  1 78 44 32 189 32 661 9187 324 232 ...
 $ Shape_Leng: num  407.6 17.2 20.9 23.3 42.3 ...
 $ Shape_Area: num  277.5 12.9 13.5 28.9 41.5 ...
 - attr(*, "data_types")= chr  "F" "C" "C" "C" ...
```

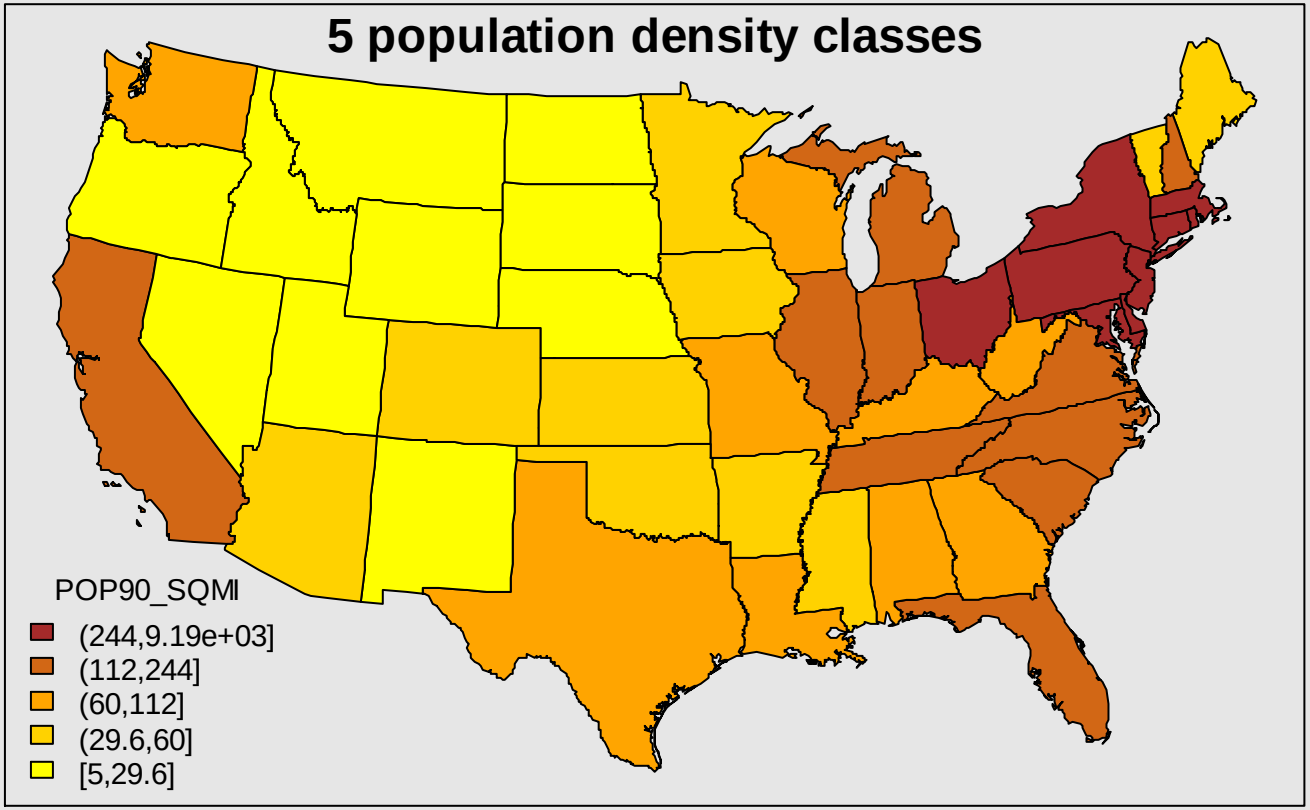

polygons...
70
60
50
40
30
20
y
-180
-160
-140
-120
-100
-80
x


projection: NA
80
60
40
20
0
-180
-160
-140
-120
-100
-80
states


+proj=longlat
+datum=NAD27
80°N
70°N
60°N
50°N
40°N
30°N
20°N
10°N
180°
160°W
140°W
120°W
100°W
80°W


+proj=aea
+lon_0=-100
8e+06
6e+06
4e+06
2e+06
-6e+06
-4e+06
-2e+06
0e+00
2e+06

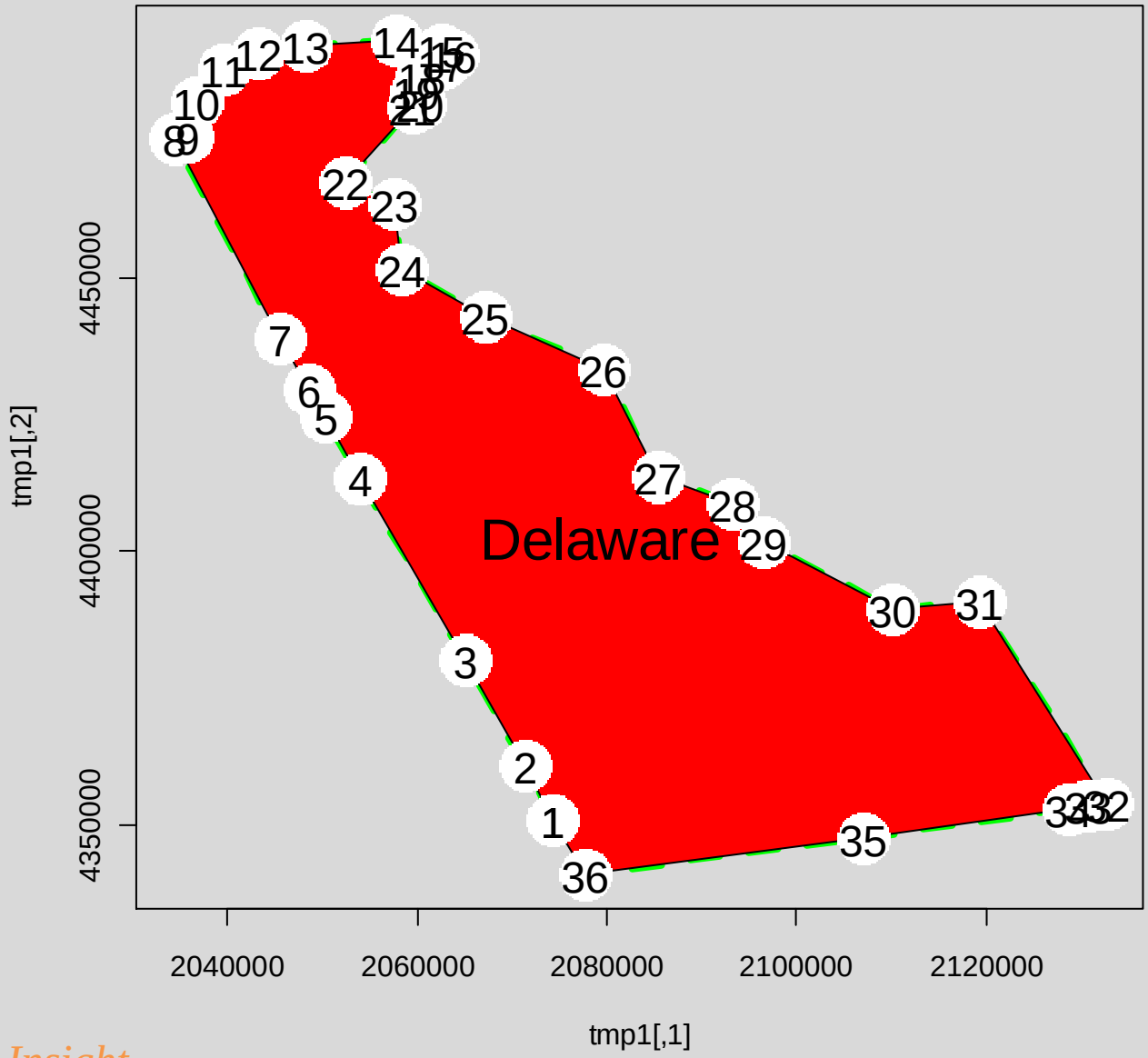

49 "states"
6e+06
5e+06
4e+06
3e+06
2e+06
-2e+06
-1e+06
0e+00
1e+06
2e+06
1000 km

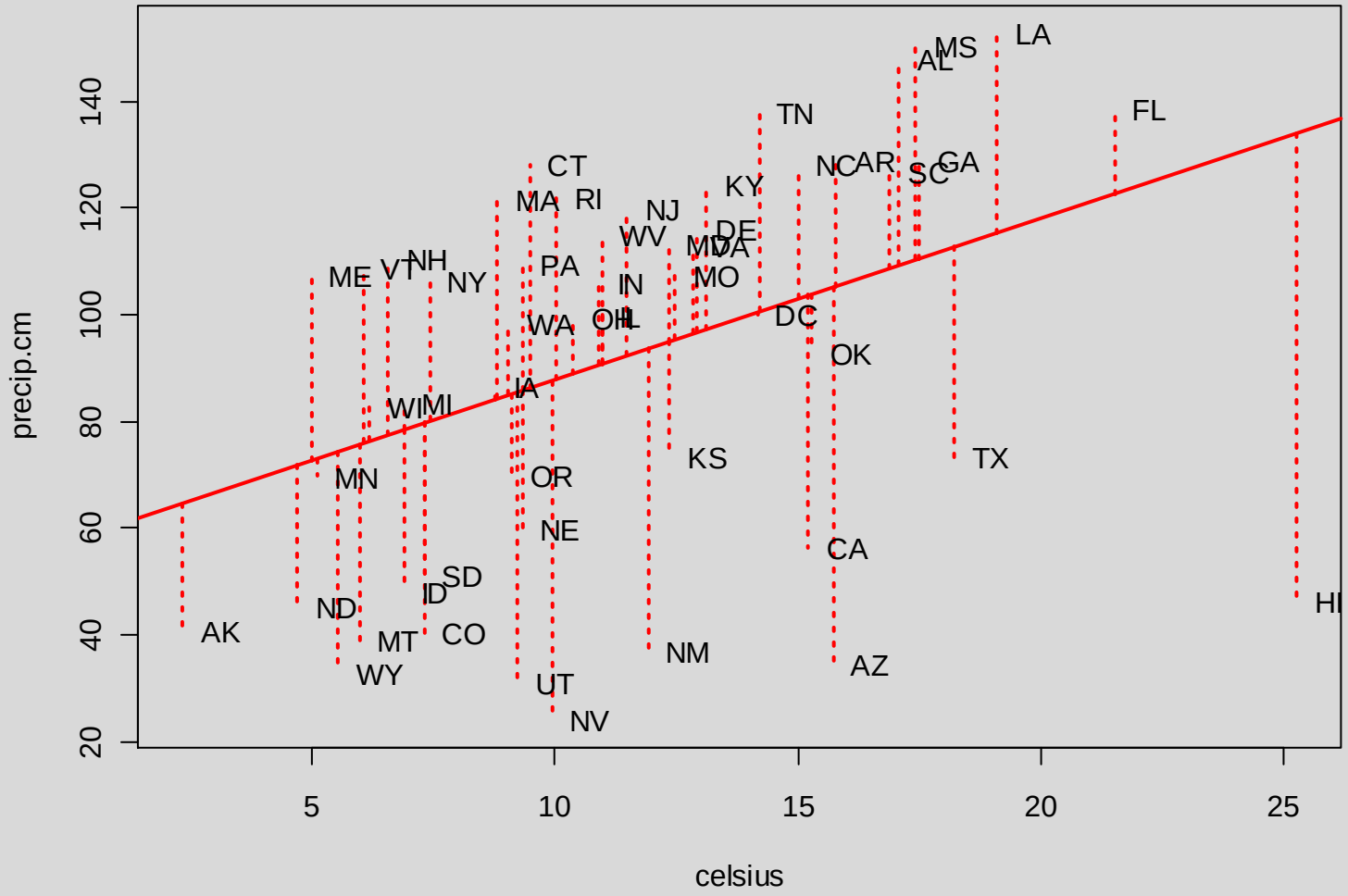

5 population density classes
POP90_SQMI
(244,9.19e+03]
(112,244]
(60,112]
(29.6,60]
[5,29.6]

slots in slots in lists...

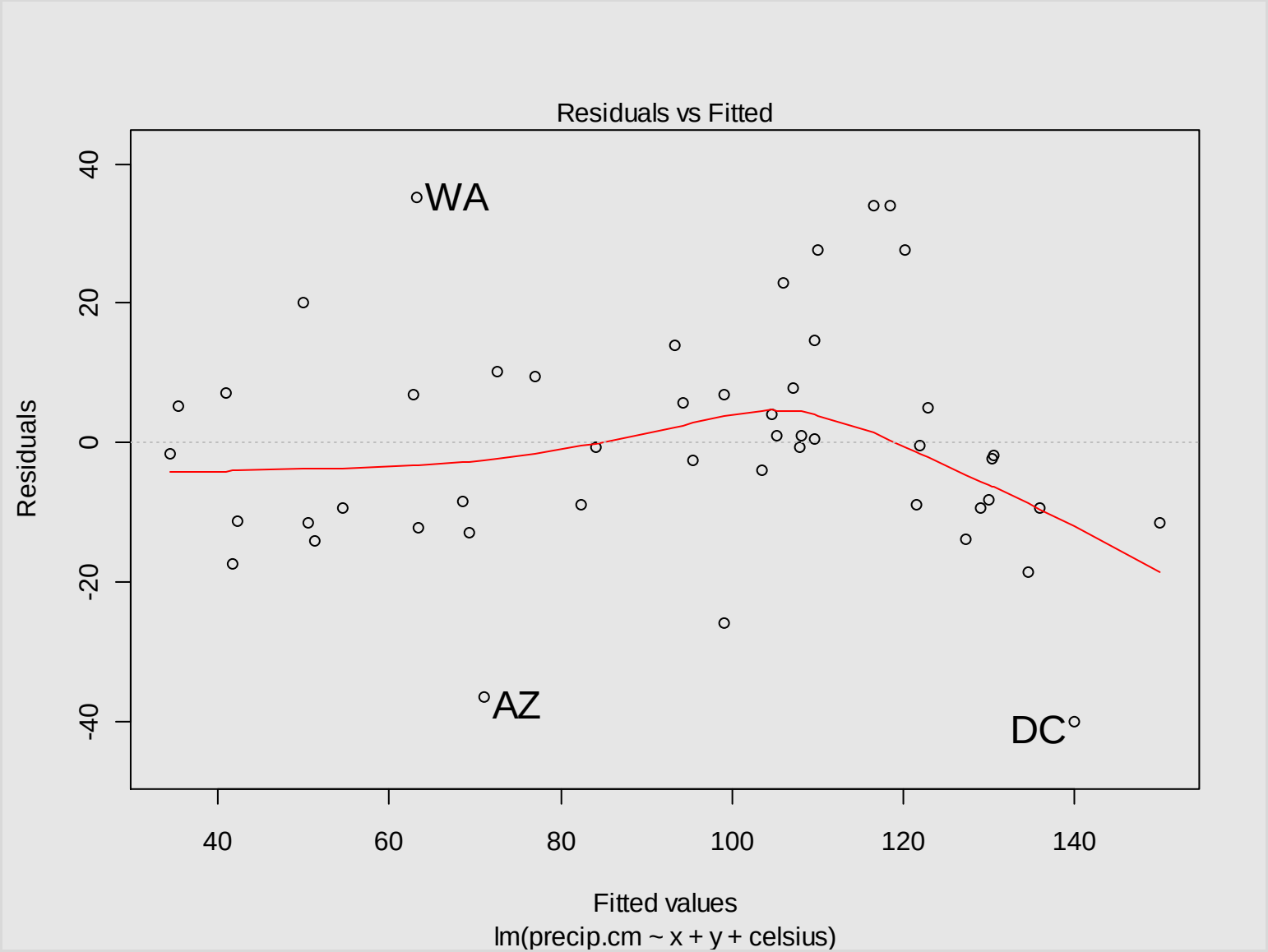
```
> US_map
An object of class "Polygons"
Slot "Polygons":
[[1]]
An object of class "Polygon"
Slot "labpt":
[1] 2079178 4402736

Slot "area":
[1] 5.321e+09

Slot "hole":
[1] FALSE

Slot "ringDir":
[1] 1
```

```
Slot "coords":
        [,1]    [,2]
 [1,] 2074290 4350978
 [2,] 2071377 4360883
 [3,] 2065105 4380169
                . . .
[34,] 2128665 4352966
[35,] 2107051 4347752
[36,] 2077658 4340921
[37,] 2074290 4350978

Slot "plotOrder":
[1] 1

Slot "labpt":
[1] 2079178 4402736

Slot "ID":
[1] "DE"

Slot "area":
[1] 5.321e+09
```

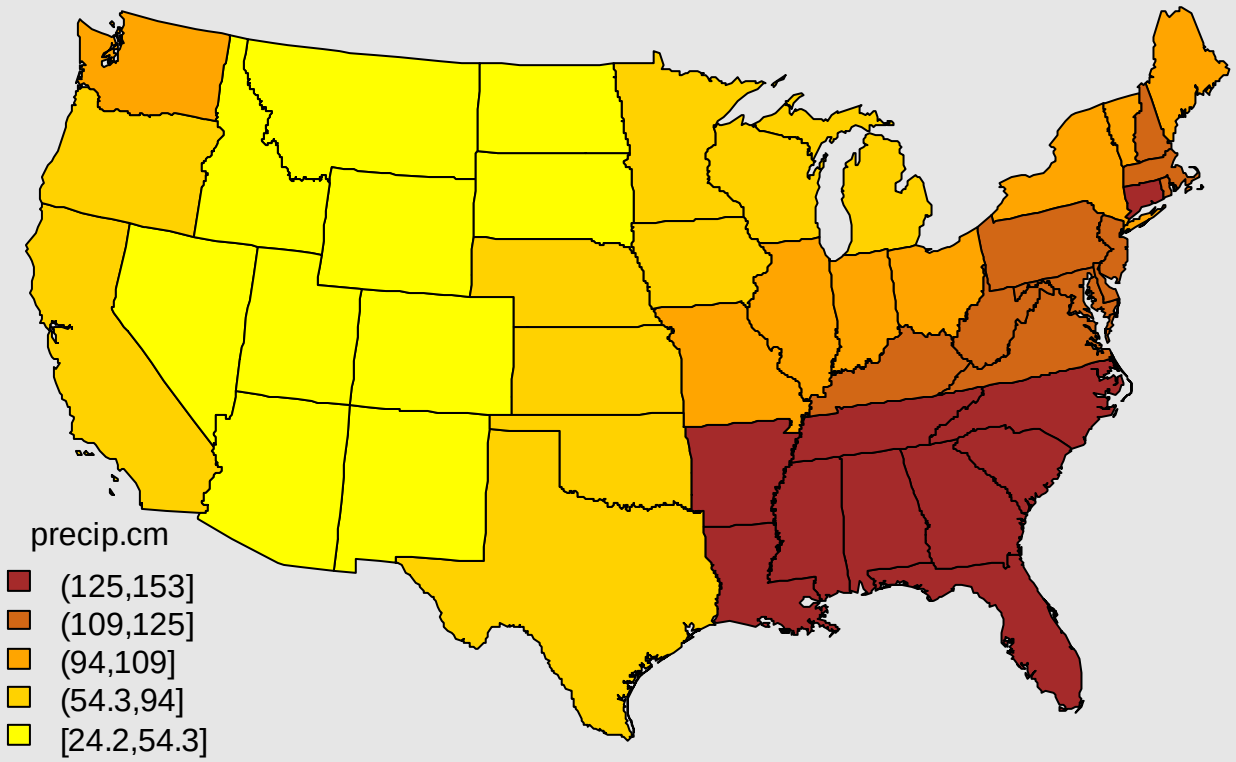

Delaware
tmp1[,2]
tmp1[,1]
4350000
4400000
4450000
2040000
2060000
2080000
2100000
2120000

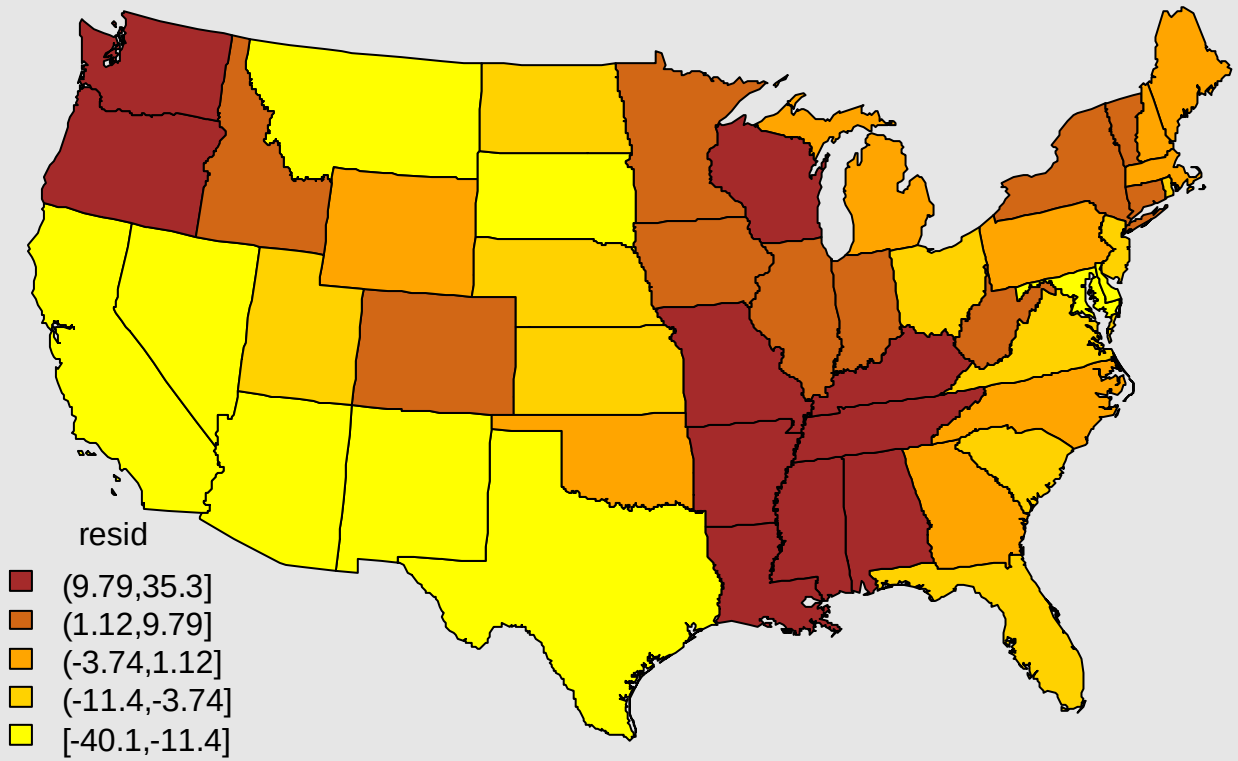

precip.cm
20
40
60
80
100
120
140
5
10
15
20
25
celsius
LA
MS
AL
TN
FL
CT
NC
AR
SC
GA
MA
RI
KY
NJ
WV
DE
MD
VA
NH
VT
ME
NY
PA
IN
MO
WA
OH
IL
DC
OK
IA
WI
MI
KS
TX
MN
OR
NE
CA
SD
ID
ND
HI
AK
CO
MT
NM
WY
AZ
UT
NV

	x	y	celsius	precip.cm	resid
AL	1222691	3523127	17.09	148.03	27.742
AR	684755	3701176	15.79	128.98	22.877
AZ	-1062689	3670290	15.73	34.57	-36.524
CA	-1699595	4105675	15.22	56.39	-12.926
CO	-475778	4147375	7.31	40.56	5.302
CT	2225180	4748798	9.47	127.99	5.102
.	.	.			
VA	1831343	4172315	12.84	112.75	-8.812
VT	2152492	5012496	6.04	108.76	4.187
WA	-1539903	5234050	9.03	98.50	35.298
WI	790232	4803632	6.18	82.91	10.285
WV	1659198	4263318	10.96	115.06	7.828
WY	-611893	4605223	5.54	32.94	-1.504

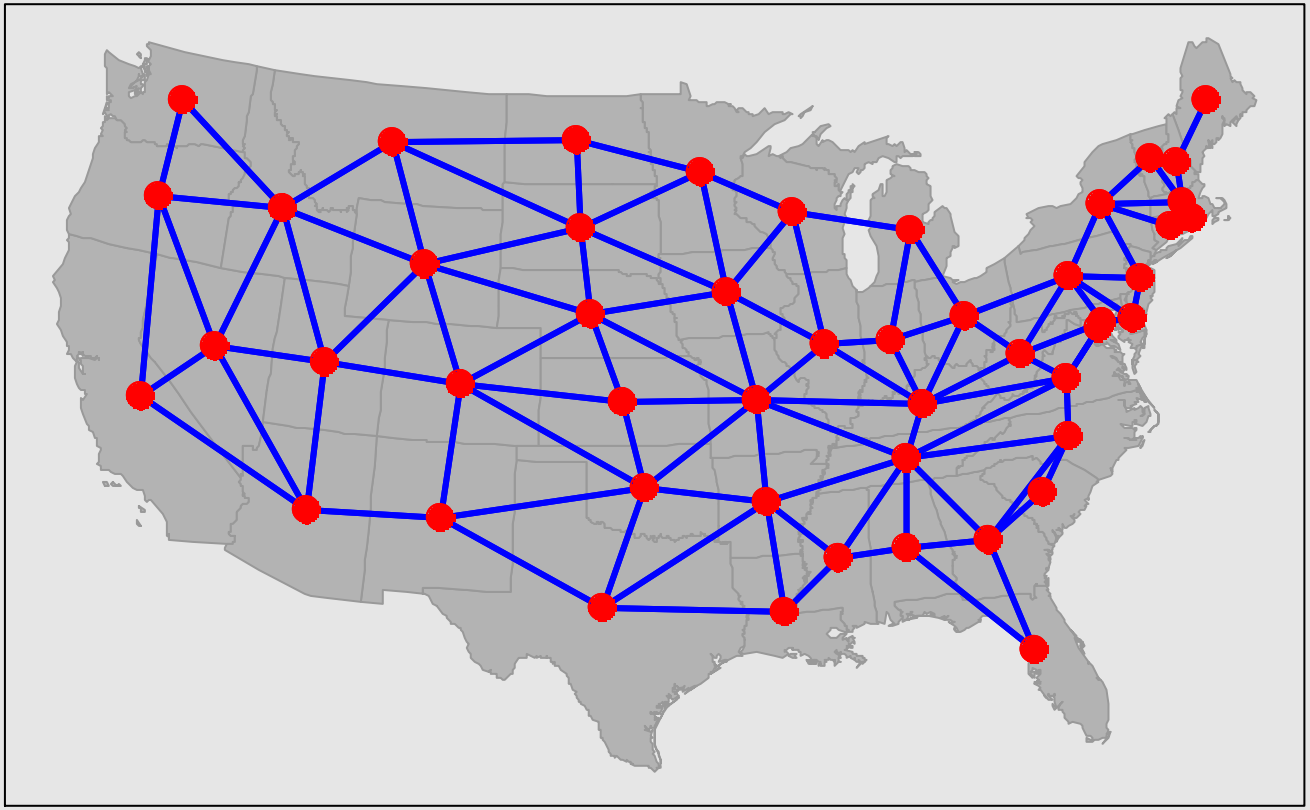

Residuals vs Fitted
WA
AZ
DC
Residuals
40
20
0
-20
-40
40
60
80
100
120
140
Fitted values
lm(precip.cm ~ x + y + celsius)

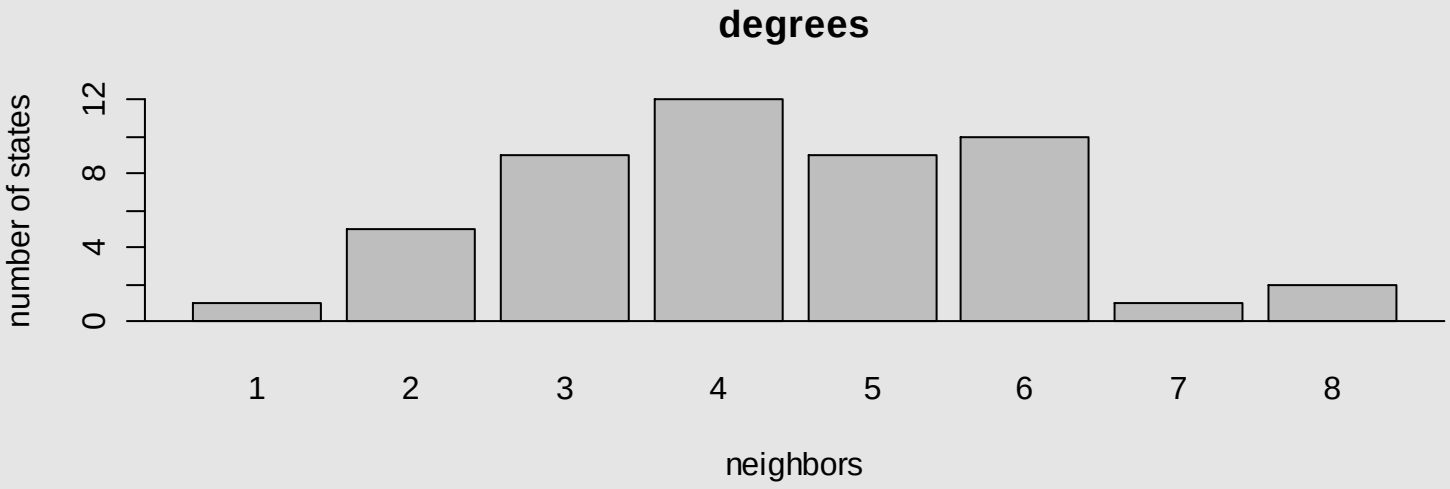
```
Call:
lm(formula = precip.cm ~ x + y + celsius, data = tmp1)

Residuals:
    Min      1Q  Median      3Q     Max
-40.070  -9.374  -0.716   7.157  35.298

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.46e+02   6.38e+01   -2.29   0.0270 *
x            1.92e-05   2.10e-06    9.15  7.9e-12 ***
y            3.27e-05   1.13e-05    2.90   0.0057 **
celsius      7.45e+00   1.47e+00    5.05  7.7e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.9 on 45 degrees of freedom
Multiple R-squared: 0.792,      Adjusted R-squared: 0.778
F-statistic: 57.1 on 3 and 45 DF,  p-value: 2.19e-15
```

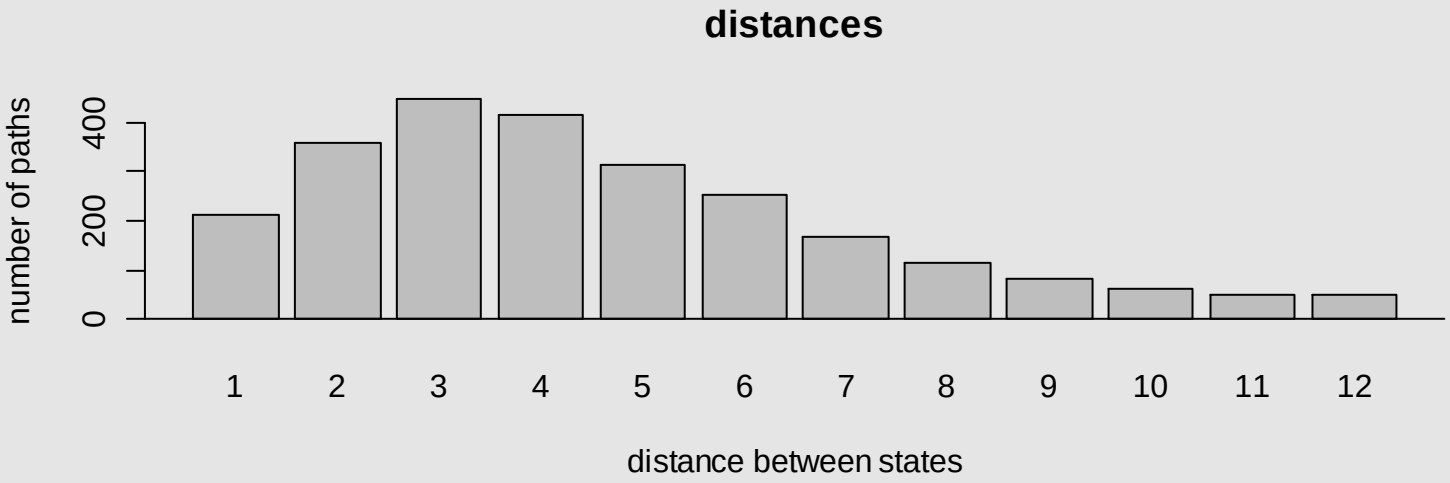

precip.cm
(125,153]
(109,125]
(94,109]
(54.3,94]
[24.2,54.3]

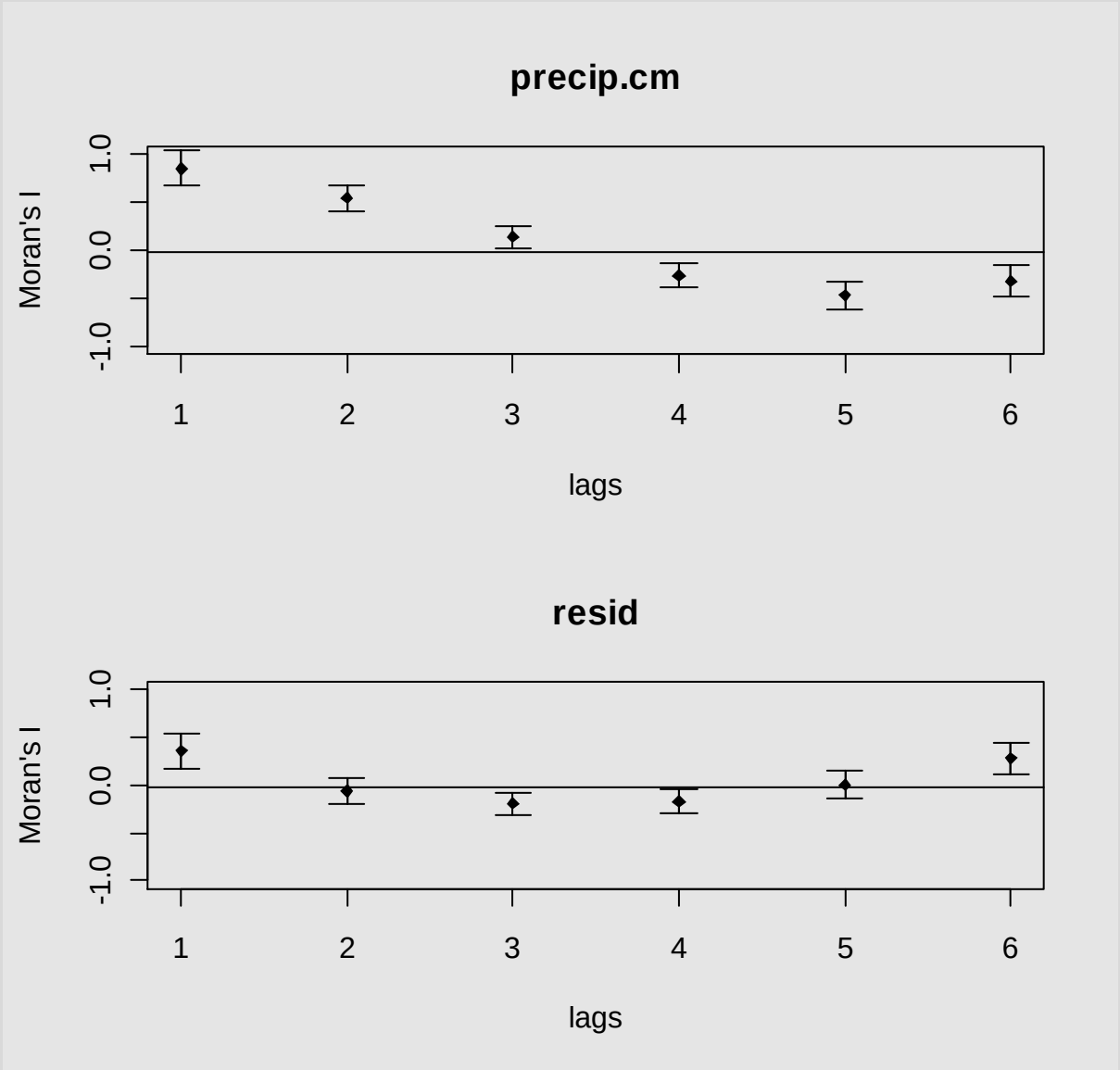

resid
(9.79,35.3]
(1.12,9.79]
(-3.74,1.12]
(-11.4,-3.74]
[-40.1,-11.4]

neighbor graph




degrees
number of states
0
4
8
12
1
2
3
4
5
6
7
8
neighbors


distances
number of paths
0
200
400
1
2
3
4
5
6
7
8
9
10
11
12
distance between states


precip.cm
Moran's I
1.0
0.0
-1.0
1
2
3
4
5
6
lags
resid
Moran's I
1.0
0.0
-1.0
1
2
3
4
5
6
lags

Analysis

REPEAT:

RESIDUAL = DATA – PATTERN

until *RESIDUAL* is random.

R: +/–

Advantages

- free
- access to spatial information and attribute data
- lots of packages
- reusable scripts

Disadvantages

- ‘shallow’ learning curve
- lots of packages (untenured faculty)

A FEW PACKAGES

package	objects	example function()
maps		map()
maptools		readShapePoly()
spatial	pp	surf.ls(), variogram()
rgdal		readOGR()
sp	Spatial...	plot()
spatstat	ppp	
gstat	meuse	krige()
spdep		moran()
splancs		kernel2d()
epitools		poisson.exact()
RColorBrewer		brewer.pal()

TIME

tricky

often regular

directional

SPACE

trickier

rarely regular

dimensional

orthogonal

‘explains’ much but causes nothing...

Lee De Cola
Data to Insight
LDECOLA@COMCAST.NET