Introduction to spatial data analysis

(Nangalatra t@ Amy Wesolowski)
Johns Hopkins Bloomberg School of Public Health

Introduction to spatial data analysis

(Nangalatra t@ Amy Wesolowski)
Johns Hopkins Bloomberg School of Public Health

Ino no asantsika niany?

Voalohany: Manazava karazan'ny 'spatial data' maromaro

Identify different types of spatial data.

Identifier différents types de données spatiales

Ino no asantsika niany?

Voalohany: Manazava karazan'ny 'spatial data' maromaro
Identify different types of spatial data.
Identifier différents types de données spatiales

Faharoa: Mianatra mamaky ireo karazan'ny 'spatial data' @ R
Learn how to read in these types of data in R.
Apprenez à lire dans ces types de données dans R.

Ino no asantsika niany?

Voalohany: Manazava karazan'ny 'spatial data' maromaro
Identify different types of spatial data.
Identifier différents types de données spatiales

Faharoa: Mianatra mamaky ireo karazan'ny 'spatial data' @ R
Learn how to read in these types of data in R.
Apprenez à lire dans ces types de données dans R.

Fahatelo: Mianatra manambatra ireo karazan'ny 'spatial data'
Learn how to combine these data.
Apprenez à combiner ces données.

Ino no asantsika niany?

Voalohany: Manazava karazan'ny 'spatial data' maromaro
Identify different types of spatial data.
Identifier différents types de données spatiales

Faharoa: Mianatra mamaky ireo karazan'ny 'spatial data' @ R
Learn how to read in these types of data in R.
Apprenez à lire dans ces types de données dans R.

Fahatelo: Mianatra manambatra ireo karazan'ny 'spatial data'
Learn how to combine these data.
Apprenez à combiner ces données.

Fahefatra: Mampiharana 'spatial statistical model' tsotra iraike
Learn one basic spatial statistical model.
Apprendre un modèle statistique spatial de base.

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

Voalohany: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

Where are Liantso's tortoises more common?"

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

Where are Liantso's tortoises more common?"

Where are Robuste's nutrition indicators worse?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

Where are Liantso's tortoises more common?"

Where are Robuste's nutrition indicators worse?

Where are Vero's forest sampling sites?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

- Where are Liantso's tortoises more common?"
- Where are Robuste's nutrition indicators worse?
- Where are Vero's forest sampling sites?
- Where are Tanjona's sites with low treatment rates?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Exploratory Data Analysis → **Better describe or visualize a system**

Ohatra: Akeza no tena misy _____?

Implications:
E.g., where to do future studies or where to direct more resources?

Où faire des études futures ou où diriger plus de ressources?

Where are Liantso's tortoises more common?"

Where are Robuste's nutrition indicators worse?

Where are Vero's forest sampling sites?

Where are Tanjona's sites with low treatment rates?

Voalohany: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Spatial modeling → **Finding relationships between spatial variables**

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Spatial modeling → **Finding relationships between spatial variables**

Ohatra: Are _____ and _____ non-randomly distributed?

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Spatial modeling → **Finding relationships between spatial variables**

Ohatra: Are _____ and _____ non-randomly distributed?

Deforestation Rates

Malaria

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Spatial modeling → **Finding relationships between spatial variables**

Ohatra: Are _____ and _____ non-randomly distributed?

- **Deforestation Rates**
- **Malaria**
- **Ants**
- **Fragmented Forest**

Voalohany: Introduction - Why do spatial analysis?

Ino râ azontsika avy amin'ny 'spatial analysis'?

Spatial modeling → **Finding relationships between spatial variables**

Ohatra: Are _____ and _____ non-randomly distributed?

- **Deforestation Rates**
- **Malaria**
- **Ants**
- **Fragmented Forest**

Implications:
There is a (spatial) relationship between A and B

Il y a une relation entre A et B

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

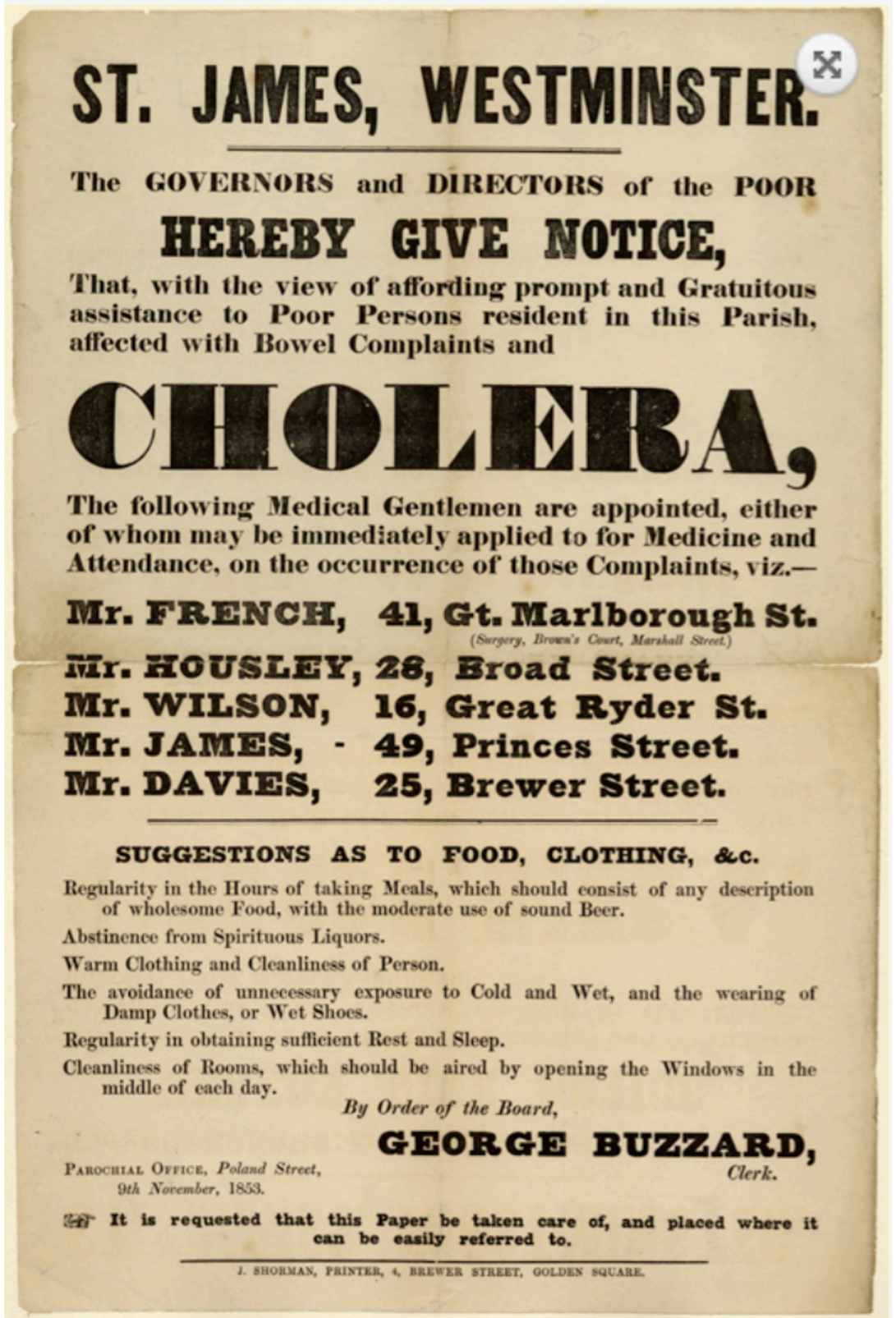
3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

John Snow

London cholera outbreaks: 1832, 1848-49, 1853-54

Officials thought cholera was spread of ‘bad air’



ST. JAMES, WESTMINSTER.

The GOVERNORS and DIRECTORS of the POOR

HEREBY GIVE NOTICE,

That, with the view of affording prompt and Gratuitous assistance to Poor Persons resident in this Parish, affected with Bowel Complaints and

CHOLERA,

The following Medical Gentlemen are appointed, either of whom may be immediately applied to for Medicine and Attendance, on the occurrence of those Complaints, viz.—

Mr. FRENCH, 41, Gt. Marlborough St.
(Surgery, Brown's Court, Marshall Street.)
Mr. HOUSLEY, 28, Broad Street.
Mr. WILSON, 16, Great Ryder St.
Mr. JAMES, - 49, Princes Street.
Mr. DAVIES, 25, Brewer Street.

SUGGESTIONS AS TO FOOD, CLOTHING, &c.

Regularity in the Hours of taking Meals, which should consist of any description of wholesome Food, with the moderate use of sound Beer.

Abstinence from Spirituous Liquors.

Warm Clothing and Cleanliness of Person.

The avoidance of unnecessary exposure to Cold and Wet, and the wearing of Damp Clothes, or Wet Shoes.

Regularity in obtaining sufficient Rest and Sleep.

Cleanliness of Rooms, which should be aired by opening the Windows in the middle of each day.

By Order of the Board,

GEORGE BUZZARD,
Clerk.

Parochial Office, *Poland Street,*
9th November, 1853.

☞ It is requested that this Paper be taken care of, and placed where it can be easily referred to.

J. SHORMAN, PRINTER, 4, BREWER STREET, GOLDEN SQUARE.

ON THE

MODE OF COMMUNICATION

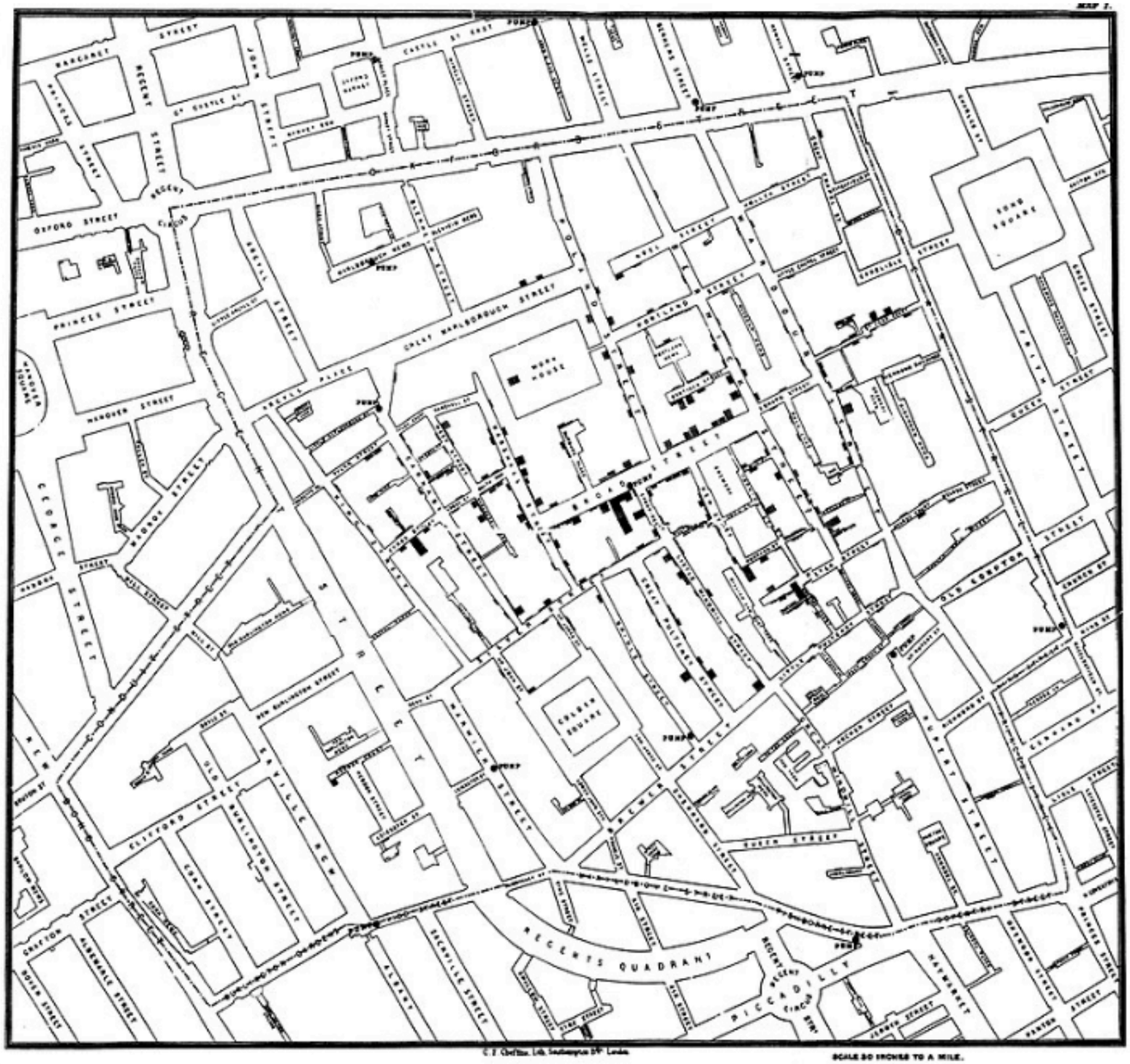
OF

CHOLERA.

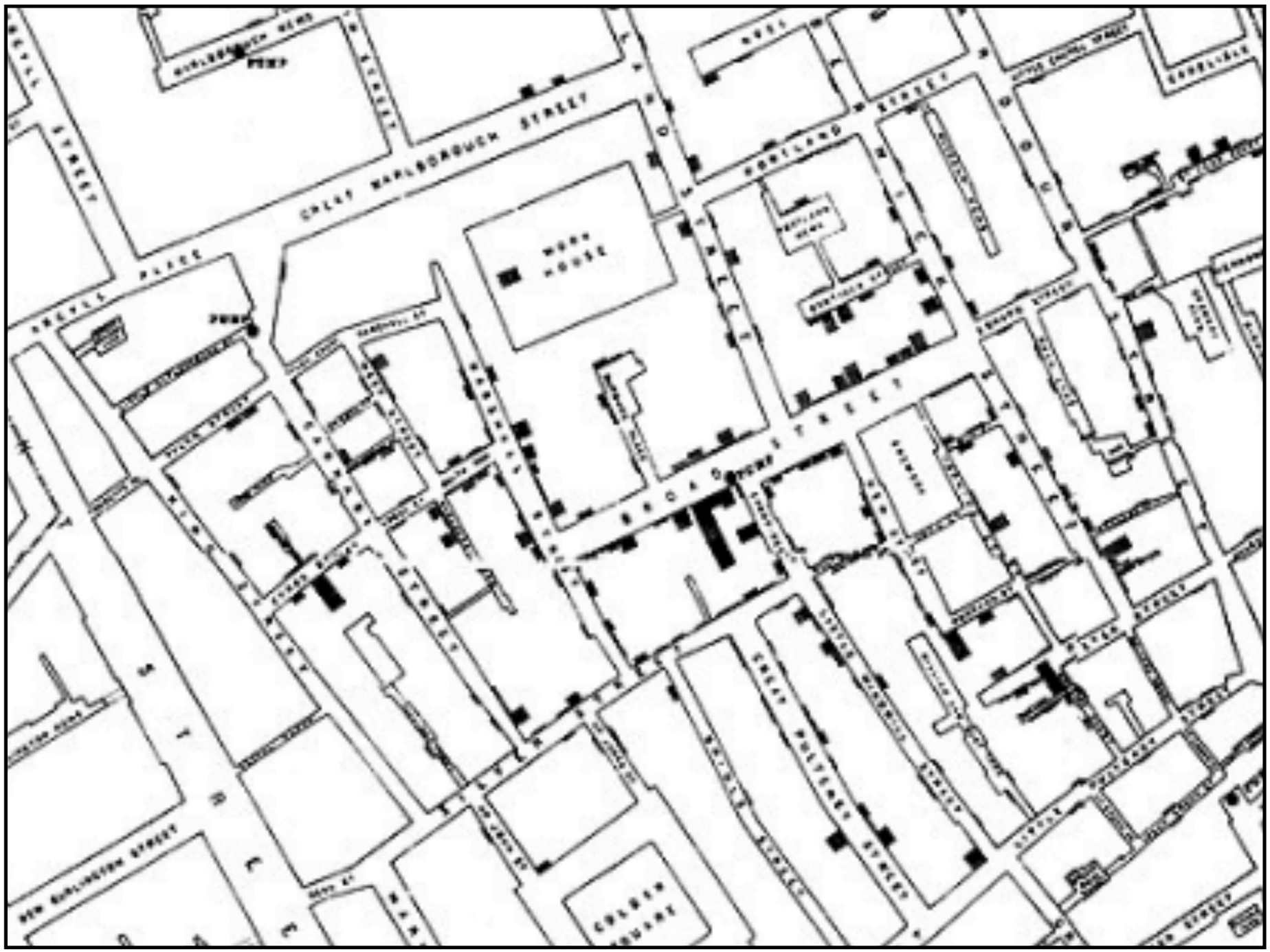
Published in 1849
*update in 1855

During the 1853-54 outbreak, Snow conducted a study in Soho, London where ~600 people died

Used government death-registration data and house-to-house interviews to map the victims' residencies and identified the proximity to water sources



John Snow cholera map (London, 1854)

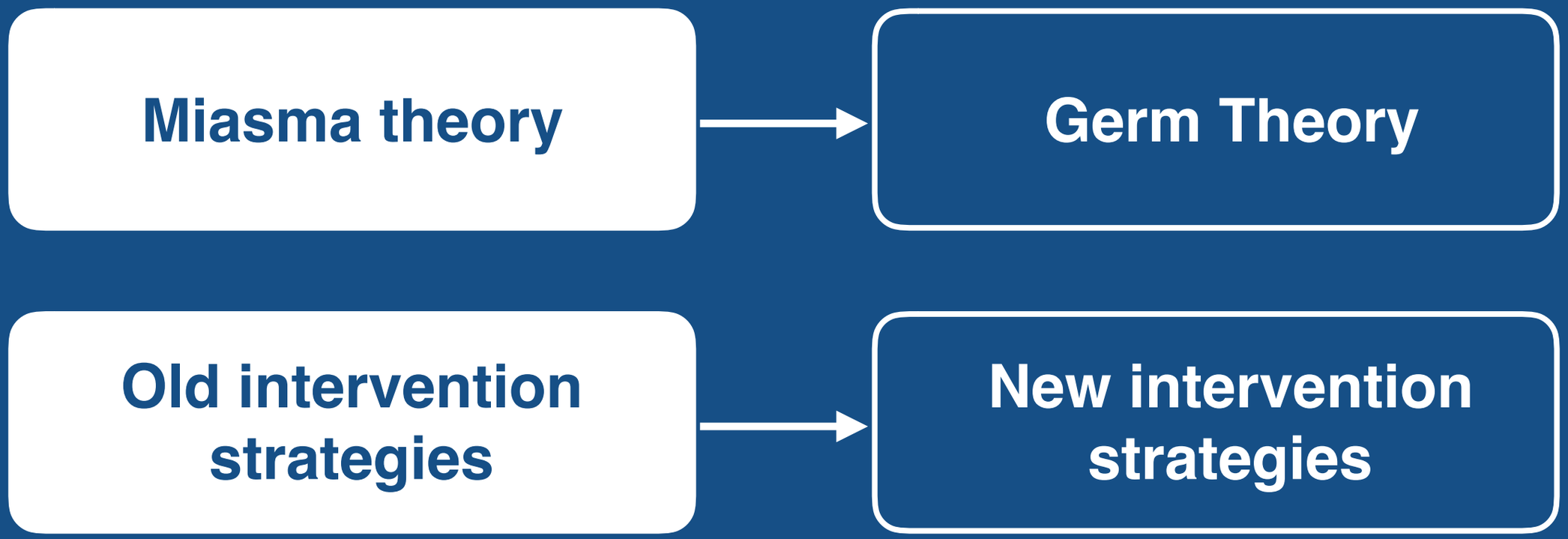


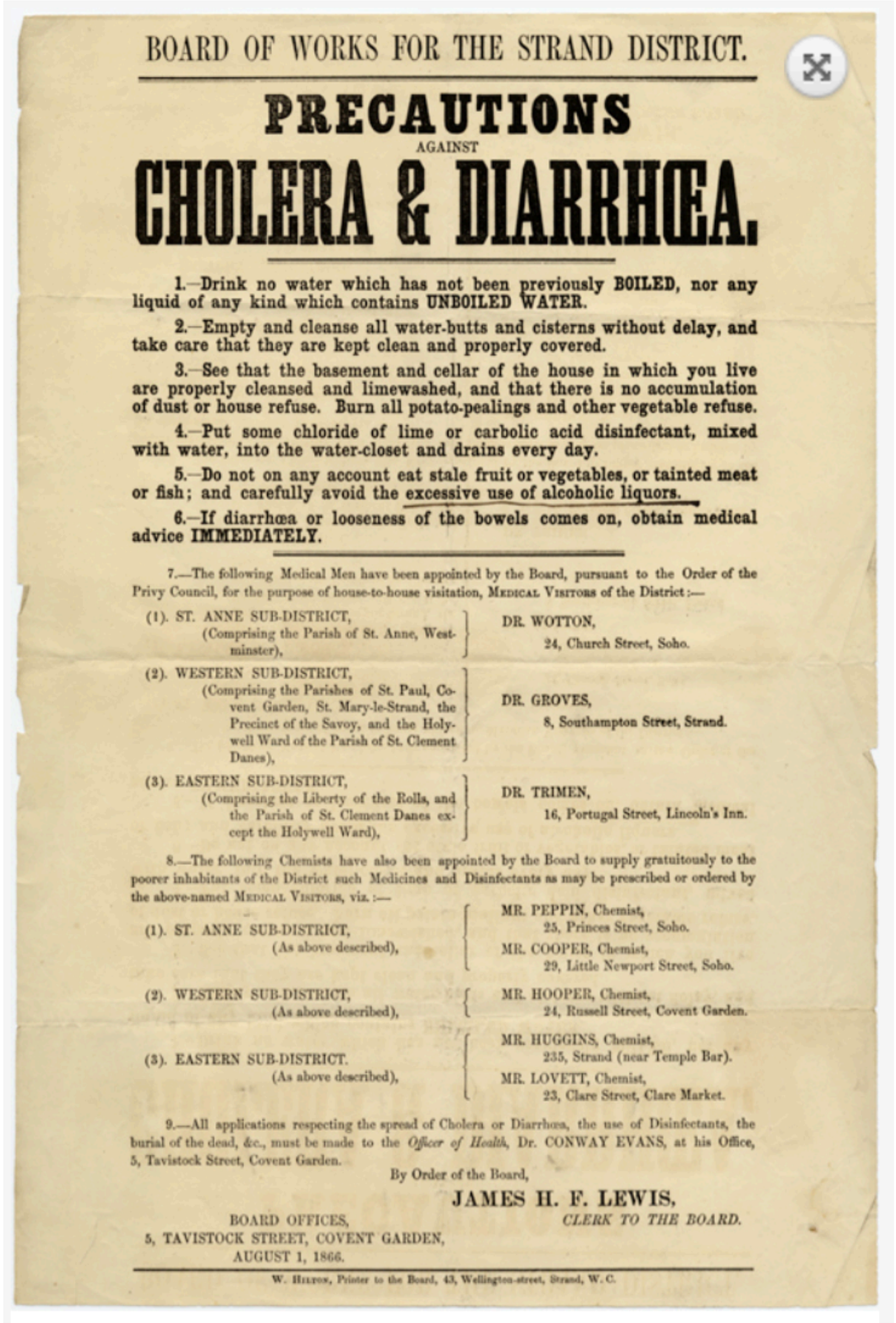
John Snow cholera map (London, 1854)


BROAD STREET
PUMP
CAMBRIDGE ST
NEW STREET
BREWERY
LITTLE

Miasma theory

Germ Theory


Miasma theory
Germ Theory
Old intervention strategies
New intervention strategies



BOARD OF WORKS FOR THE STRAND DISTRICT.

PRECAUTIONS
AGAINST
CHOLERA & DIARRHŒA.

1.—Drink no water which has not been previously BOILED, nor any liquid of any kind which contains UNBOILED WATER.

2.—Empty and cleanse all water-butts and cisterns without delay, and take care that they are kept clean and properly covered.

3.—See that the basement and cellar of the house in which you live are properly cleansed and limewashed, and that there is no accumulation of dust or house refuse. Burn all potato-pealings and other vegetable refuse.

4.—Put some chloride of lime or carbolic acid disinfectant, mixed with water, into the water-closet and drains every day.

5.—Do not on any account eat stale fruit or vegetables, or tainted meat or fish; and carefully avoid the excessive use of alcoholic liquors.

6.—If diarrhœa or looseness of the bowels comes on, obtain medical advice IMMEDIATELY.

7.—The following Medical Men have been appointed by the Board, pursuant to the Order of the Privy Council, for the purpose of house-to-house visitation, MEDICAL VISITORS of the District:—

(1). ST. ANNE SUB-DISTRICT, (Comprising the Parish of St. Anne, Westminster),	DR. WOTTON, 24, Church Street, Soho.
(2). WESTERN SUB-DISTRICT, (Comprising the Parishes of St. Paul, Covent Garden, St. Mary-le-Strand, the Precinct of the Savoy, and the Holywell Ward of the Parish of St. Clement Danes),	DR. GROVES, 8, Southampton Street, Strand.
(3). EASTERN SUB-DISTRICT, (Comprising the Liberty of the Rolls, and the Parish of St. Clement Danes except the Holywell Ward),	DR. TRIMEN, 16, Portugal Street, Lincoln's Inn.

8.—The following Chemists have also been appointed by the Board to supply gratuitously to the poorer inhabitants of the District such Medicines and Disinfectants as may be prescribed or ordered by the above-named MEDICAL VISITORS, viz.:—

(1). ST. ANNE SUB-DISTRICT, (As above described),	MR. PEPPIN, Chemist, 25, Princes Street, Soho. MR. COOPER, Chemist, 29, Little Newport Street, Soho.
(2). WESTERN SUB-DISTRICT, (As above described),	MR. HOOPER, Chemist, 24, Russell Street, Covent Garden.
(3). EASTERN SUB-DISTRICT. (As above described),	MR. HUGGINS, Chemist, 235, Strand (near Temple Bar). MR. LOVETT, Chemist, 23, Clare Street, Clare Market.

9.—All applications respecting the spread of Cholera or Diarrhœa, the use of Disinfectants, the burial of the dead, &c., must be made to the *Officer of Health*, Dr. CONWAY EVANS, at his Office, 5, Tavistock Street, Covent Garden.

By Order of the Board,

JAMES H. F. LEWIS,
CLERK TO THE BOARD.

BOARD OFFICES,
5, TAVISTOCK STREET, COVENT GARDEN,
AUGUST 1, 1866.

W. HEITON, Printer to the Board, 43, Wellington-street, Strand, W.C.

1. **Drink no water** which has not been previously boiled nor any liquid of any kind which contains unboiled water.
2. **Empty and cleanse all water-butts and cisterns** without delay, and take care that they are kept clean and properly covered.
3. See that the basement and cellar of the house in which you live are properly cleansed and lime washed, and that there is no accumulation of dust or house refuse. Burn all potato-peelings and other vegetable refuse.
4. Put some **chloride of lime or carbolic acid disinfectant, mixed with water,** into the water-closet and drains every day.
5. Do not on any account eat stale fruit or vegetables or tainted meat or fish; and carefully avoid the excessive use of alcoholic liquors.
6. If diarrhea or looseness of the bowls comes on, obtain medical advice immediately.

London's polluted water supply was still the subject of satire in 1866.

Image: MARY EVANS PICTURE LIBRARY

Wikimedia Commons

CHOLERA AND WATER.

BOARD OF WORKS

FOR THE LIMEHOUSE DISTRICT,

Comprising Limehouse, Ratcliff, Shadwell, and Wapping.

The INHABITANTS of the District within which CHOLERA IS PREVAILING, are earnestly advised

NOT TO DRINK ANY WATER WHICH HAS NOT PREVIOUSLY BEEN BOILED.

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

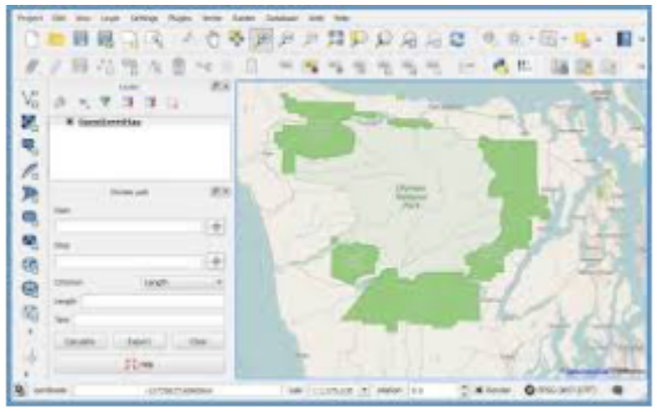
3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

ArcGIS

- Most commonly used mapping software
- Extensive mapping and spatial analysis capabilities
- Expensive software, available only on Windows

QGIS



- Freely available version of ArcGIS
- Fewer capabilities, but can still do most common mapping things

R

- The best!
- Statistical software that increasingly has many mapping and spatial statistical capabilities

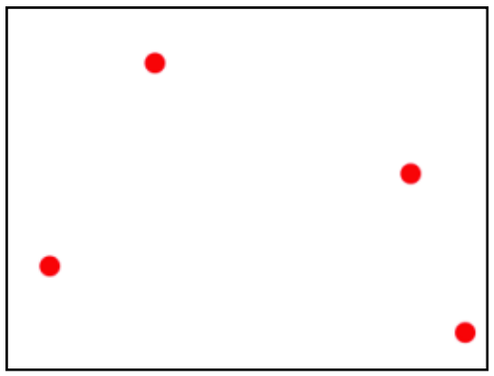
R

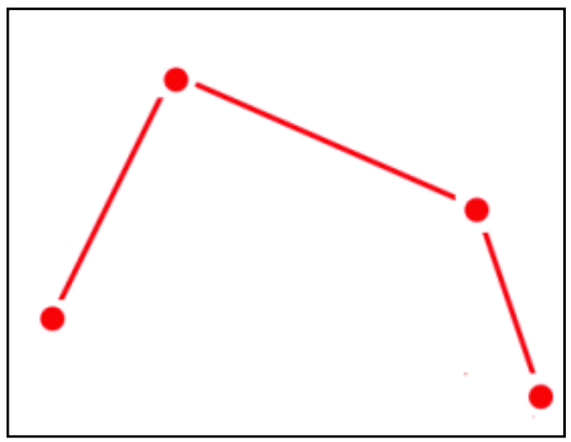
- The best!
- Statistical software that increasingly has many mapping and spatial statistical capabilities

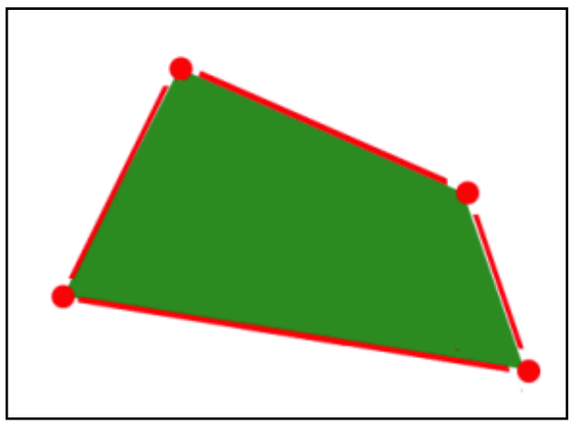
Forms of spatial data

Vector Data

Points, lines, polygons
(“shapefiles”)

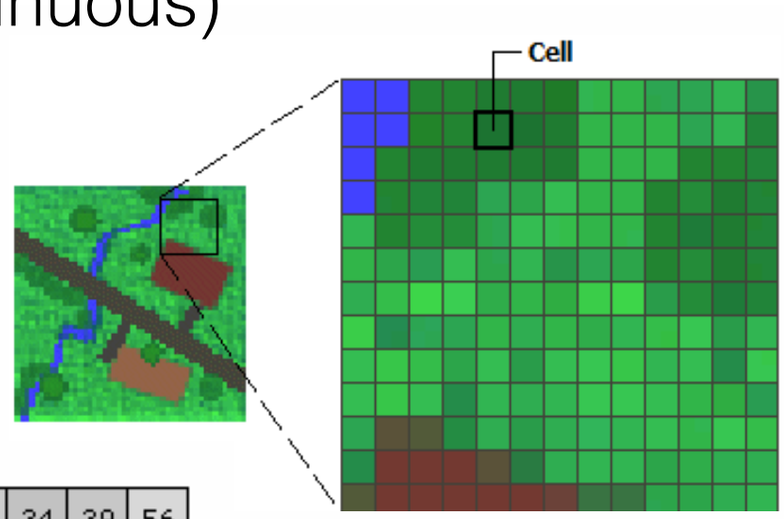




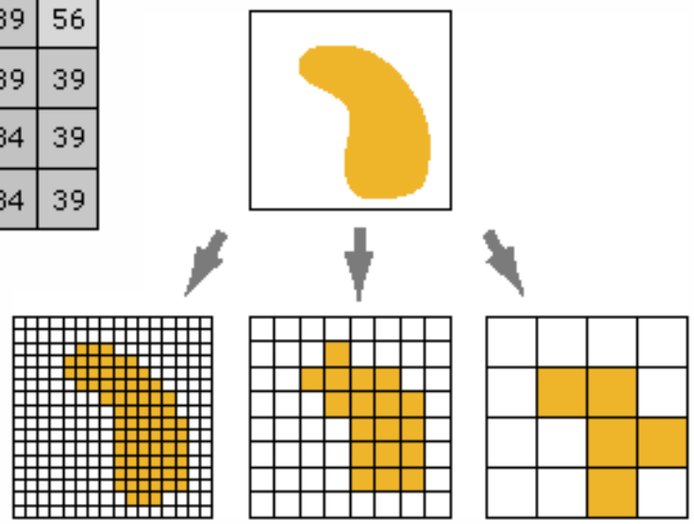


Raster Data

Area is divided into cell that each contain the value of a point (often continuous)



80	74	62	45	45	34	39	56
80	74	74	62	45	34	39	56
74	74	62	62	45	34	39	39
62	62	45	45	34	34	34	39
45	45	45	34	34	30	34	39



Shapefiles data

Raster data

Raster data

Shapefiles data

Shapefiles data

Raster data

Raster data

Shapefiles data

shapefiles

Vector Data

Points, lines, polygons

A standardized data format to use for GIS analysis.

Includes a collection of files

Mandatory files

- `.shp` — shape format; the feature geometry itself
- `.shx` — shape index format; a positional index of the feature geometry to allow seeking forwards and backwards quickly
- `.dbf` — attribute format; columnar attributes for each shape, in dBase IV format

Other files

- `.prj` — projection format; the coordinate system and projection information, a plain text file describing the projection using well-known text format
- `.sbn` and `.sbx` — a spatial index of the features
- `.fbn` and `.fbx` — a spatial index of the features that are read-only
- `.ain` and `.aih` — an attribute index of the active fields in a table
- `.ixs` — a geocoding index for read-write datasets
- `.mxs` — a geocoding index for read-write datasets (ODB format)
- `.atx` — an attribute index for the `.dbf` file in the form of *shapefile.columnname*`.atx` (ArcGIS 8 and later)
- `.shp.xml` — geospatial metadata in XML format, such as ISO 19115 or other XML schema
- `.cpg` — used to specify the code page (only for `.dbf`) for identifying the character encoding to be used
- `.qix` — an alternative quadtree spatial index used by MapServer and GDAL/OGR software

shapefiles

DIVA-GIS

free, simple & effective

Download program | Documentation | Free Spatial Data

Home

Free Spatial Data

Country level

Download country level data for any country in the world: administrative boundaries, roads, railroads, altitude, land cover, population density.

Global level

A new file with the (2011) global country boundaries

Global climate data

See WorldClim or diva-gis specific data here

Species occurrence data

GBIF, HerpNET, MaNIS, OBIS, ORNIS, REMIB

Crop (genebank) collection data

GENESYS

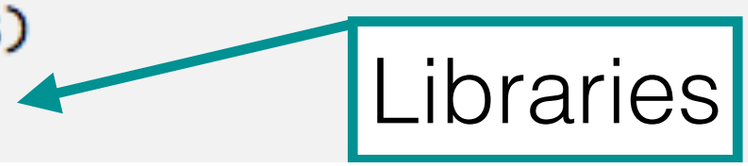
Near global 90 meter resolution elevation data

Download

www.diva-gis.org

reading shapefiles into R

```
library(maptools)
library(rgdal)
```



reading shapefiles into R

```
library(maptools)
library(rgdal)

mdg_admin1_shp<-readShapePoly('~/Dropbox/Teaching/MDG_Shp/MDG_adm1.shp',
proj4string = CRS('+proj=longlat'))
mdg_admin1_shp<-gSimplify(mdg_admin1_shp, tol = 0.01)

par(mfrow=c(1,3))
plot(mdg_admin1_shp)
```

Libraries

Read in file

Plotting

reading shapefiles into R

```
library(maptools)
library(rgdal)

mdg_admin1_shp<-readShapePoly('~/Dropbox/Teaching/MDG_Shp/MDG_adm1.shp',
proj4string = CRS('+proj=longlat'))
mdg_admin1_shp<-gSimplify(mdg_admin1_shp, tol = 0.01)

par(mfrow=c(1,3))
plot(mdg_admin1_shp)
```

Libraries

Read in file

Plotting

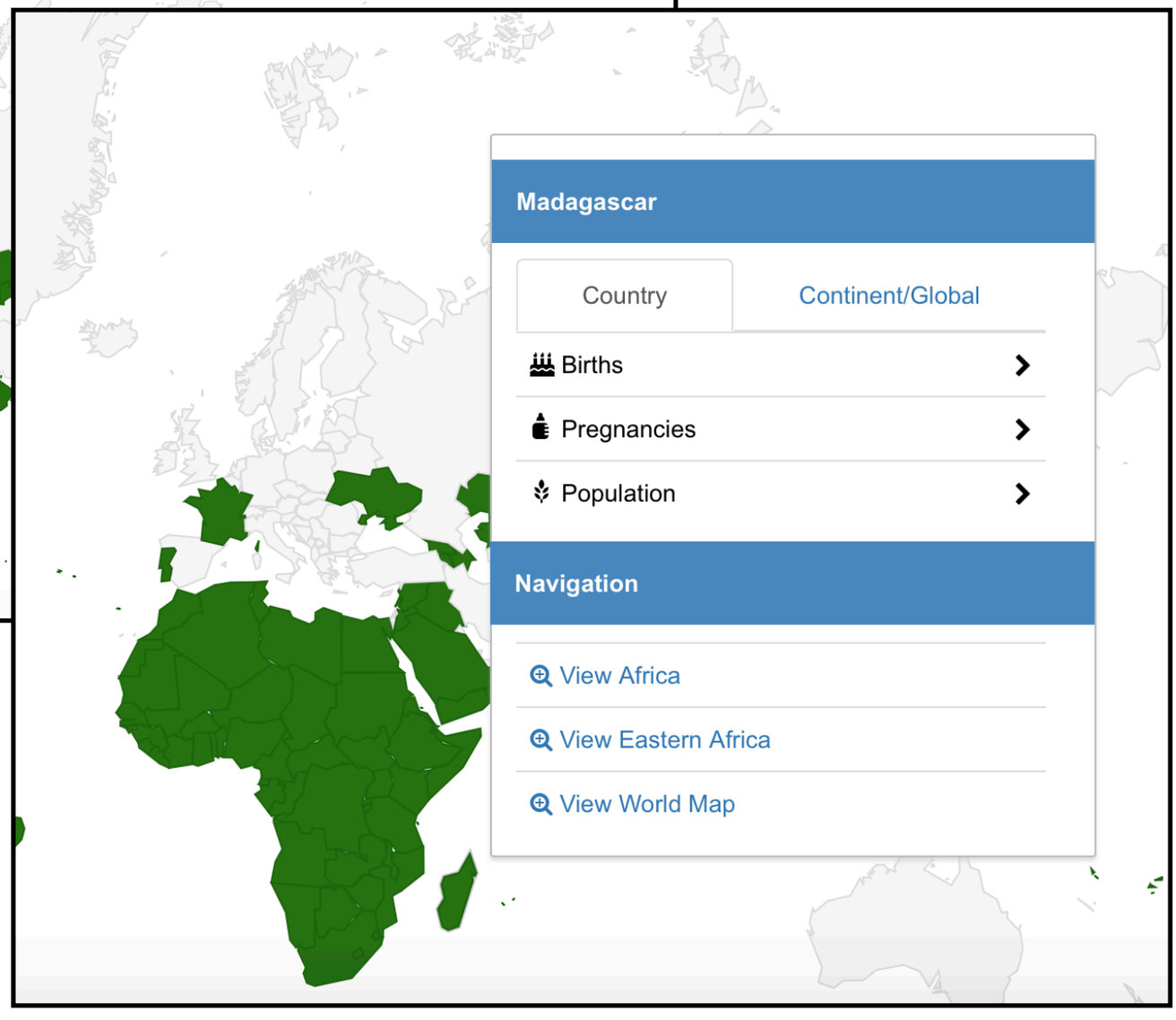
reading population raster data

What is WorldPop?

High spatial resolution, contemporary data on human population distributions are a prerequisite for the accurate measurement of the impacts of population growth, for monitoring changes and for planning interventions. The WorldPop project aims to meet these needs through the provision of detailed and open access population distribution datasets built using transparent approaches.

www.worldpop.org.uk

reading population raster data



reading population raster data

```
library(raster)
library(rgdal)
library(colorRamps)

mdg_preg<-raster('~/Dropbox/Teaching/MDG_Preg/MDG_pregs_pp_v2_2015.tif')
```

Libraries

Read in file

reading population raster data

```
library(raster)
library(rgdal)
library(colorRamps)

mdg_preg<-raster('~/Dropbox/Teaching/MDG_Preg/MDG_pregs_pp_v2_2015.tif')

par(mfrow=c(1,3))
image(mdg_preg, col = blue2red(10))
```

Libraries

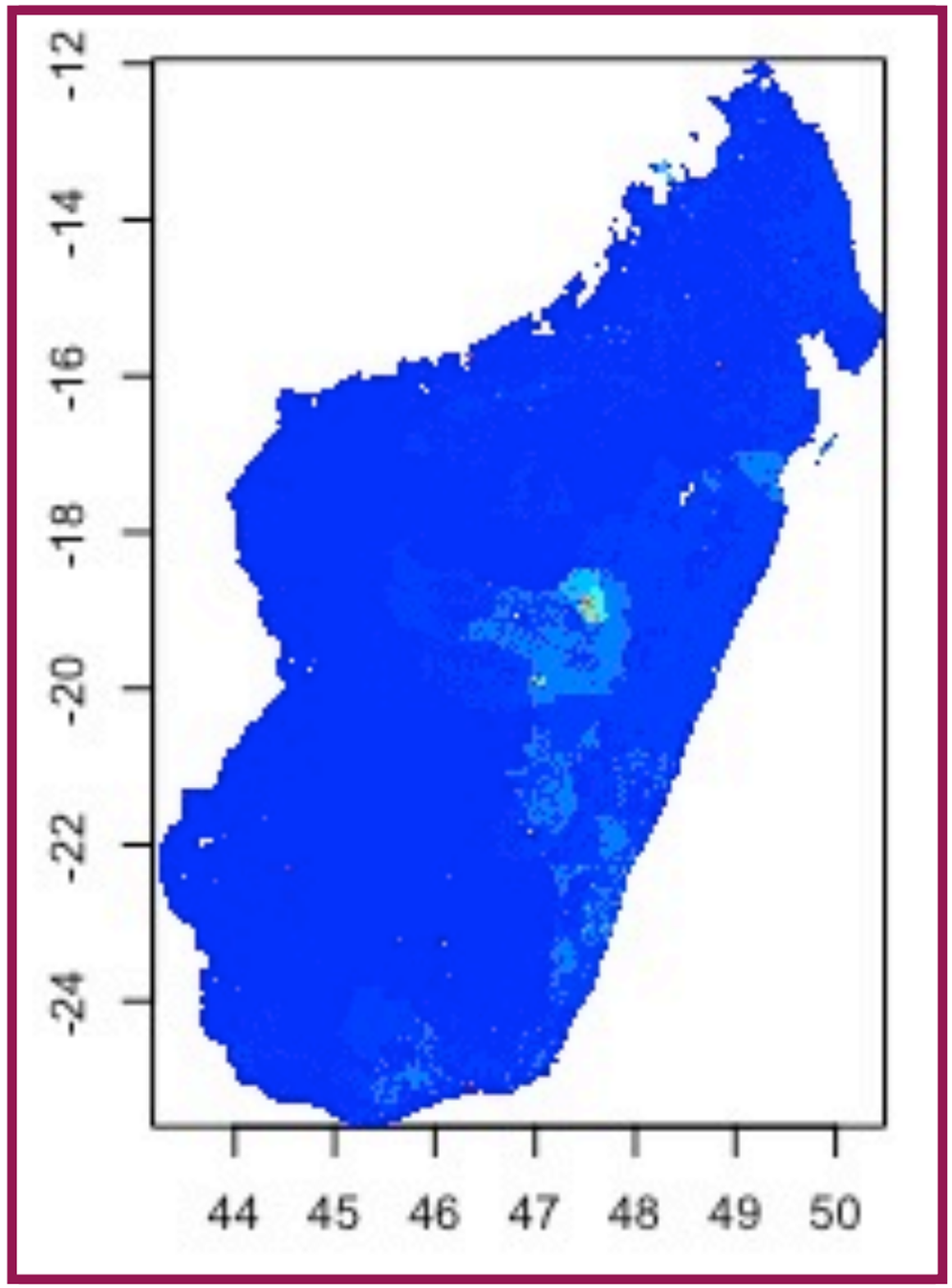
Read in file

Plotting

reading population raster data

```
par(mfrow=c(1,3))
image(mdg_preg, col = blue2red(10))
```

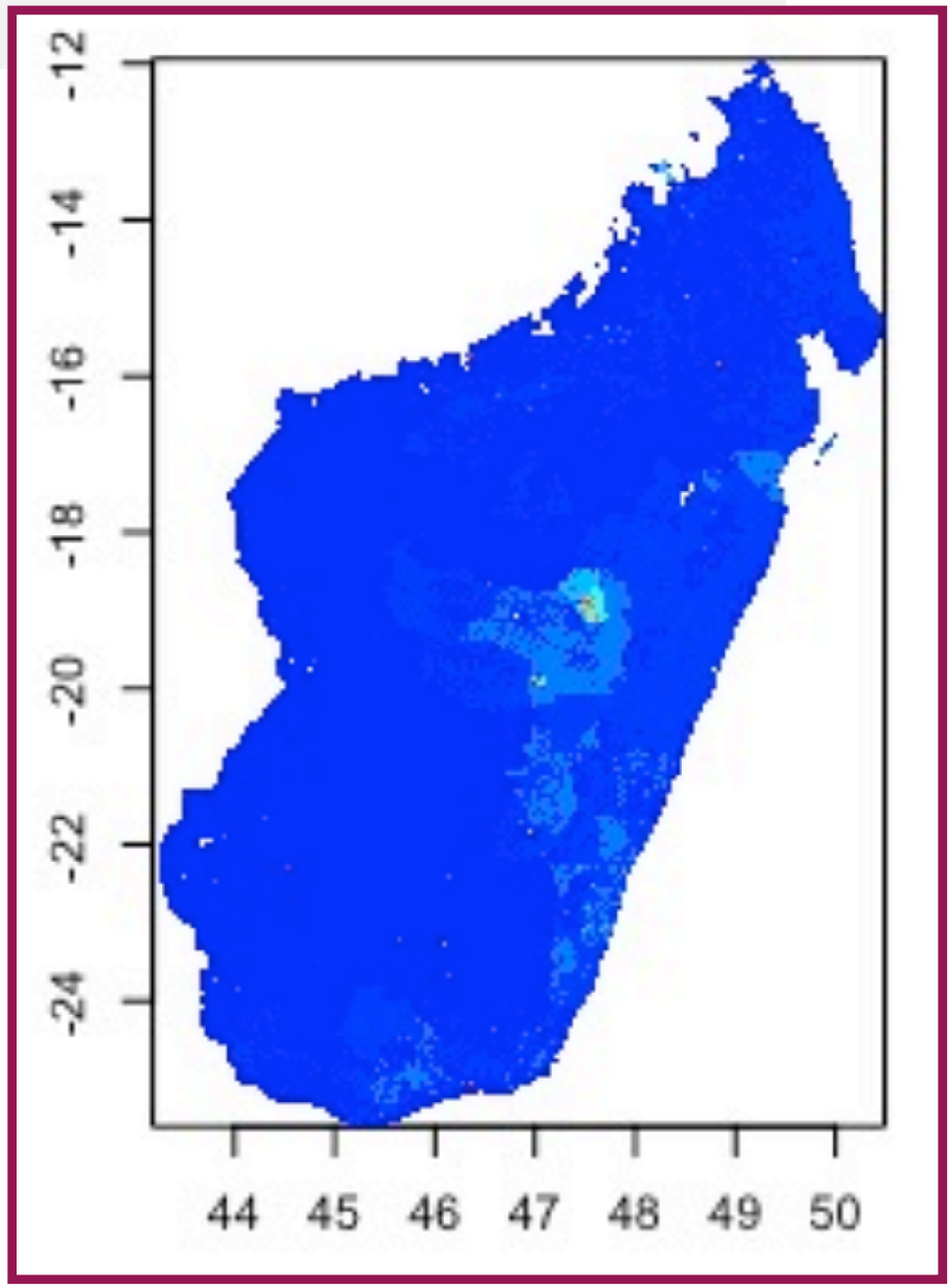
Plotting



reading population raster data

```
par(mfrow=c(1,3))
image(mdg_preg, col = blue2red(10))
image(log(mdg_preg+1), col = blue2red(10))
```

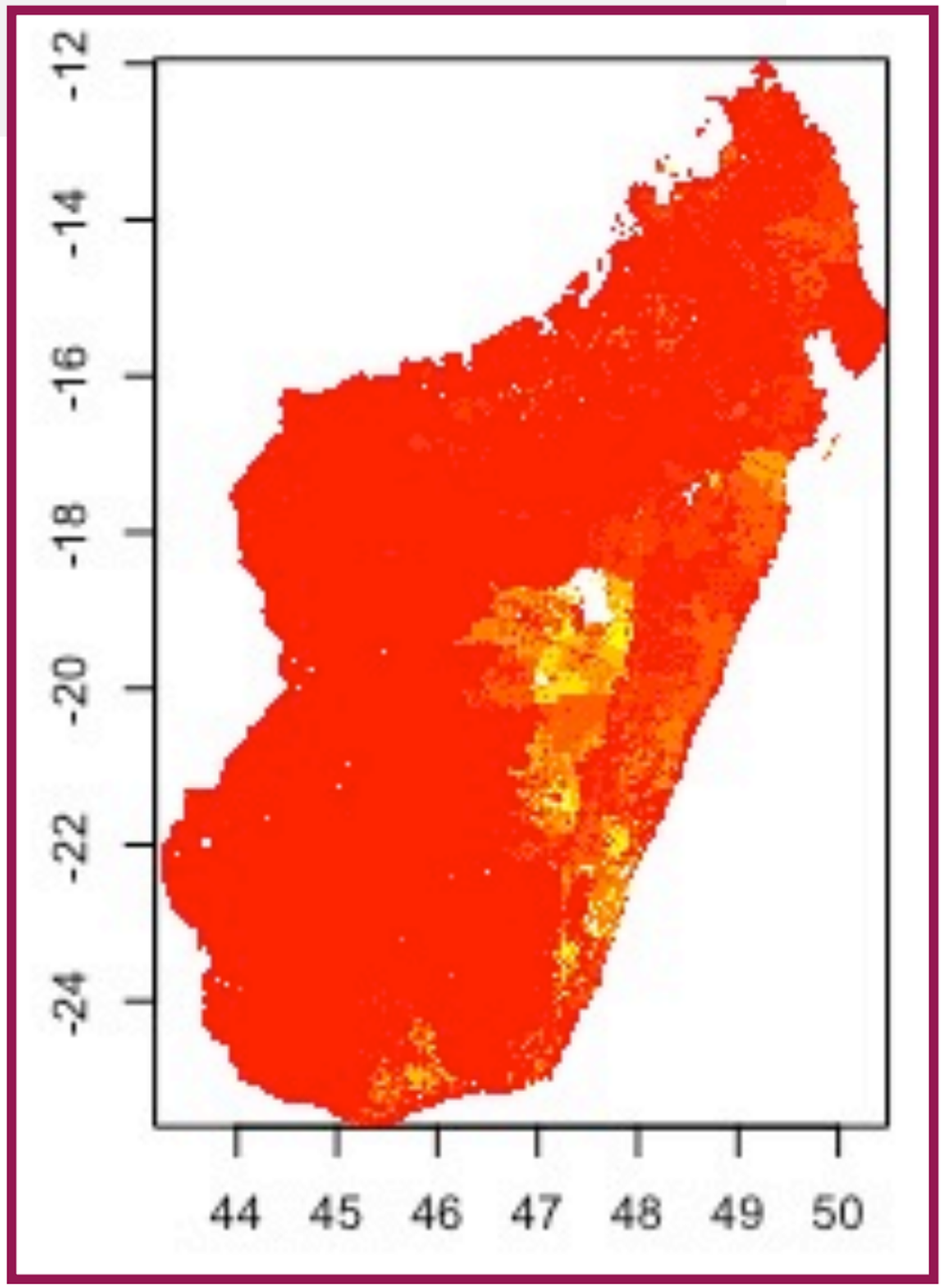
Plotting



reading population raster data

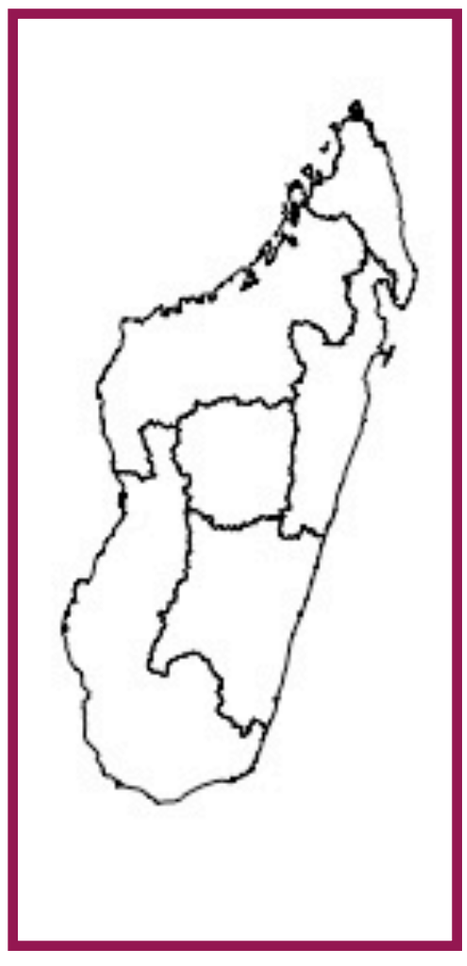
```
par(mfrow=c(1,3))
image(mdg_preg, col = blue2red(10))
image(log(mdg_preg+1), col = blue2red(10))
image(mdg_preg, zlim = c(0,10))
```

Plotting



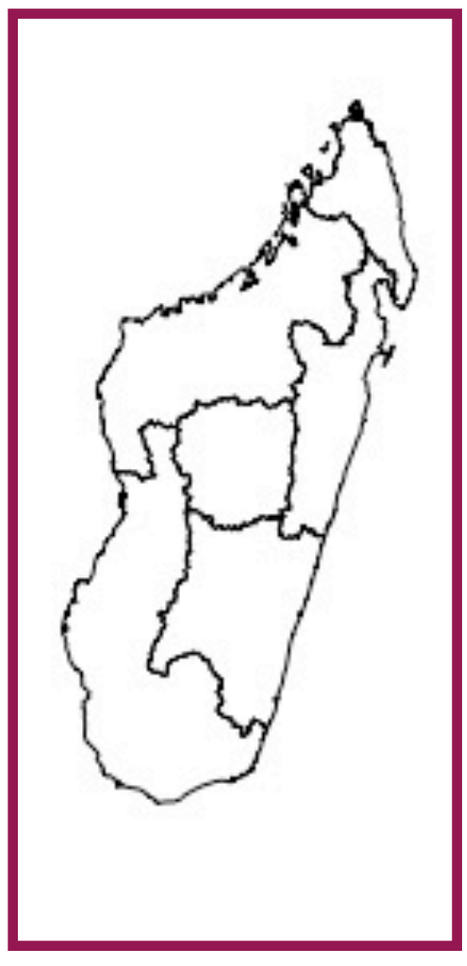
combining raster and shapefiles

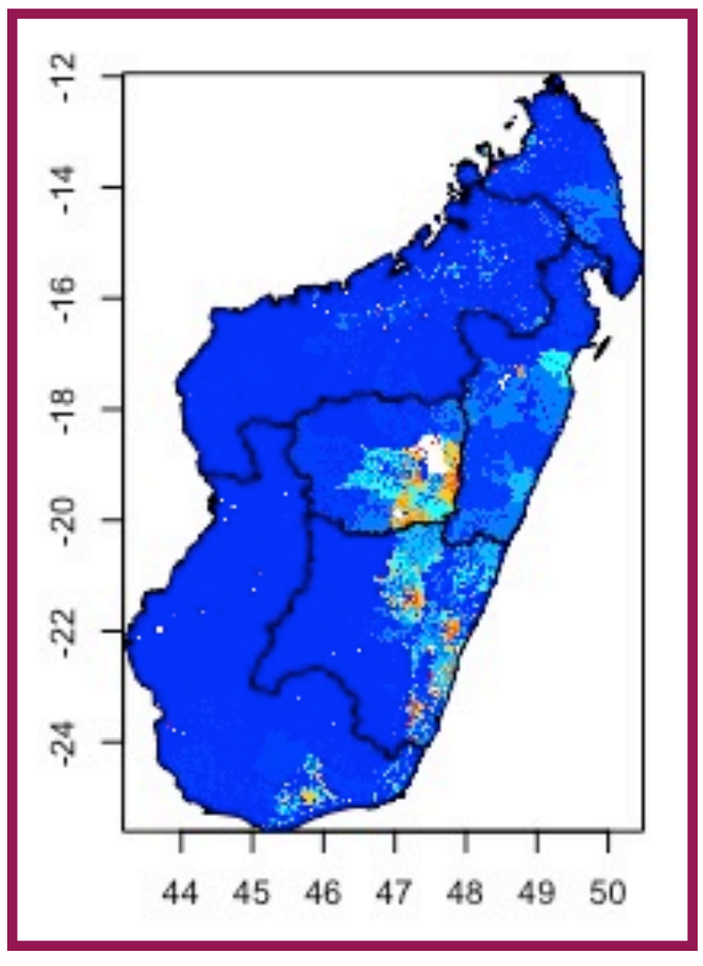
```
par(mfrow=c(1,3))
plot(mdg_admin1_shp)
```



combining raster and shapefiles

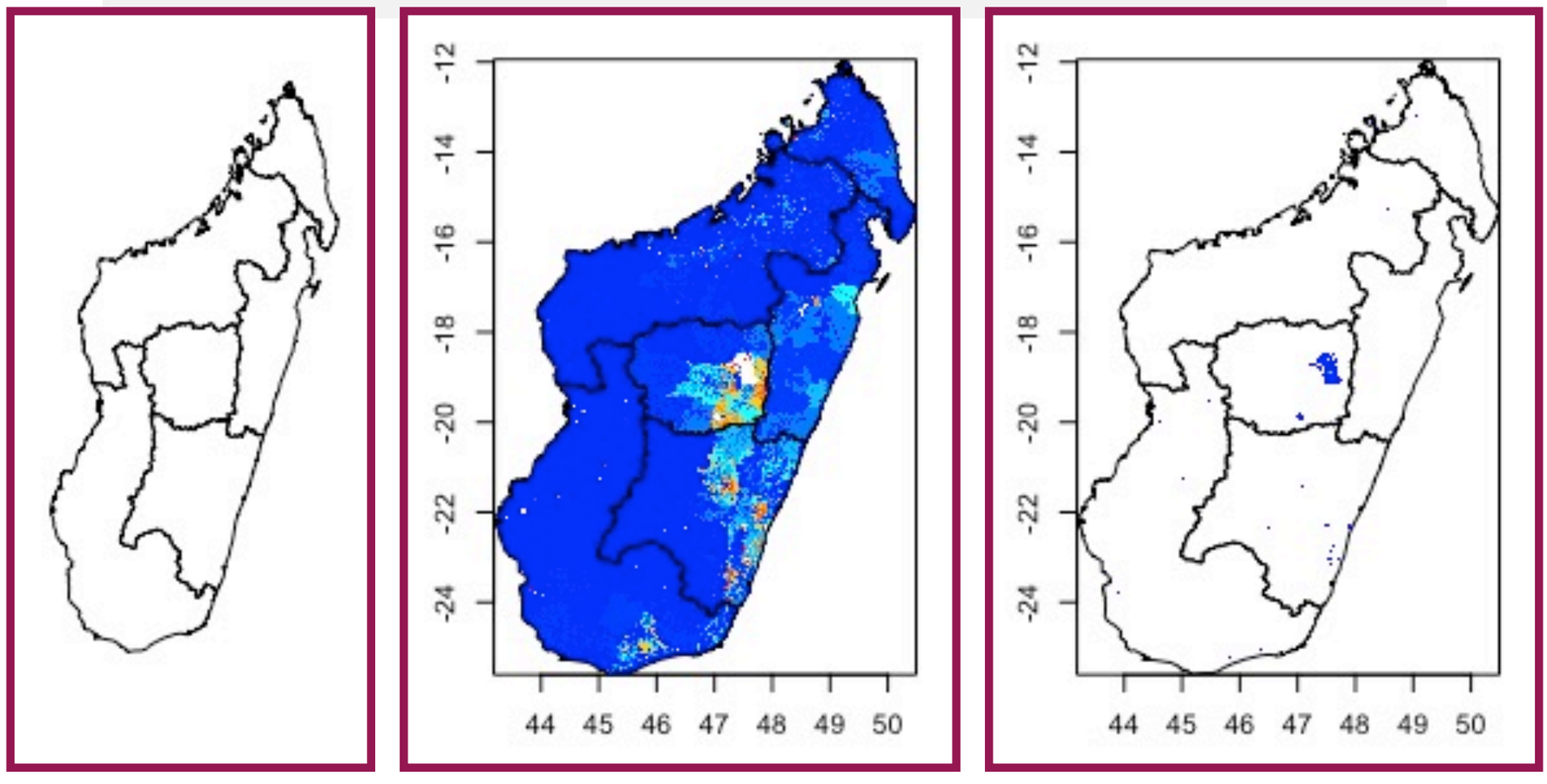
```
par(mfrow=c(1,3))
plot(mdg_admin1_shp)

image(mdg_preg, zlim = c(0,10), col = blue2red(10))
plot(mdg_admin1_shp, add = TRUE)
```



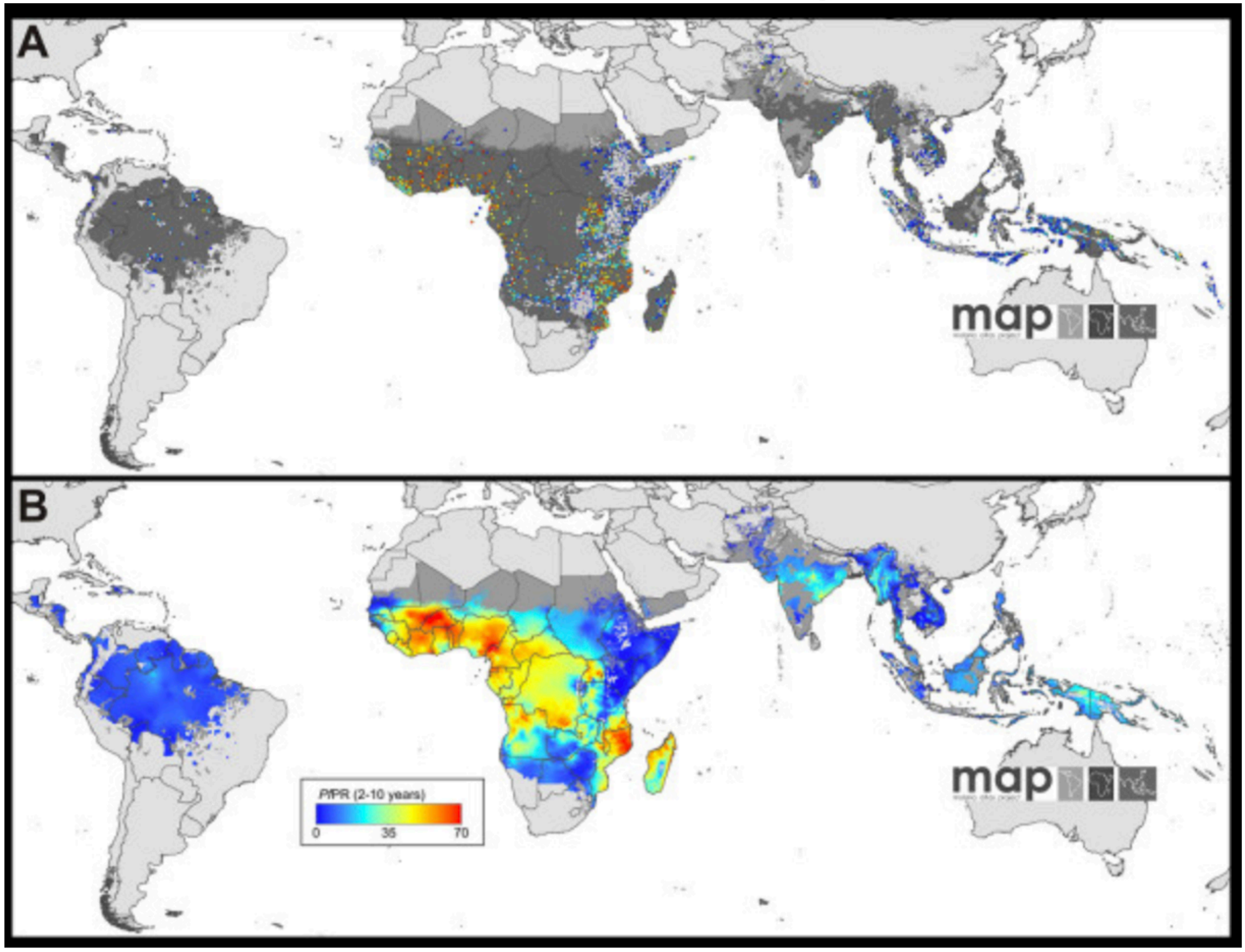


combining raster and shapefiles

```
par(mfrow=c(1,3))
plot(mdg_admin1_shp)

image(mdg_preg, zlim = c(0,10), col = blue2red(10))
plot(mdg_admin1_shp, add = TRUE)

image(mdg_preg, zlim = c(10,3000), col = blue2red(10))
plot(mdg_admin1_shp, add = TRUE)
```



extracting raster data to shapefile polygons



extracting raster data to shapefile polygons

```
extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)
```


```
> extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)
```

extracting raster data to shapefile polygons

```
extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)

mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x,
na.rm=TRUE) else NA ))
```

```
> extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)
> mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x, na.rm=TRUE) el
se NA ))
```

extracting raster data to shapefile polygons

```
extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)

mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x,
na.rm=TRUE) else NA ))

total_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x,
na.rm=TRUE) else NA ))
```

```
> extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)
> mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x, na.rm=TRUE) el
se NA ))
> total_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x, na.rm=TRUE) e
lse NA ))
```

extracting raster data to shapefile polygons

```
extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)

mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x,
na.rm=TRUE) else NA ))

total_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x,
na.rm=TRUE) else NA ))
```

```
> extract_preg_by_admin1<-extract(mdg_preg, mdg_admin1_shp, weights = FALSE)
> mean_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x, na.rm=TRUE) el
se NA ))
> total_value<-unlist(lapply(test, function(x) if (!is.null(x)) mean(x, na.rm=TRUE) e
lse NA ))
> mean_value
[1] 4.4266196 1.4565147 1.7865749 0.7549086 1.9292333 0.8437911
> total_value
[1] 4.4266196 1.4565147 1.7865749 0.7549086 1.9292333 0.8437911
```

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

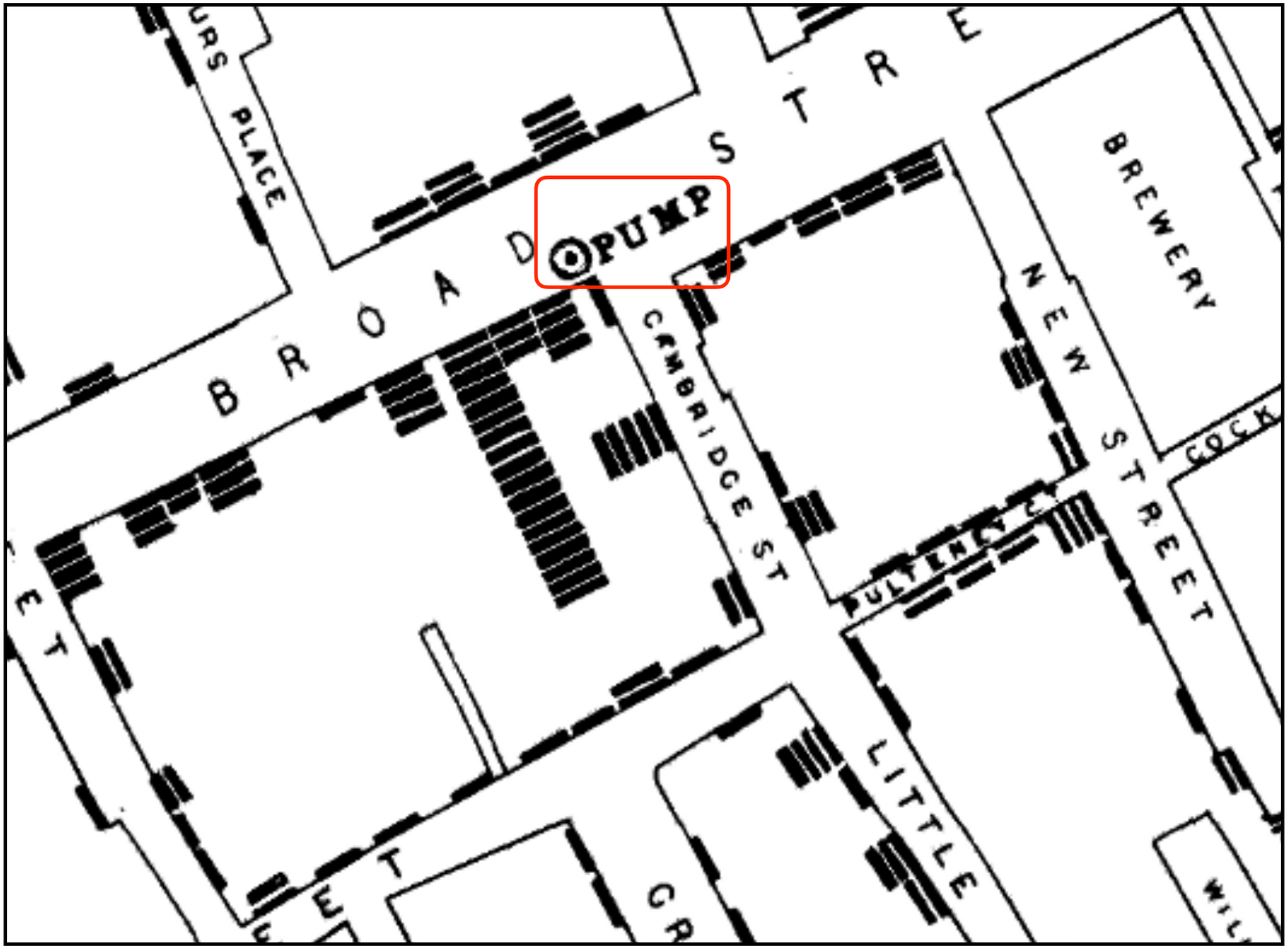
1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

exploratory data analysis of point process data



exploratory data analysis of point process data

```
library(spatstat)

deaths_points_file<-read.csv('Snow_deaths.csv', header = TRUE)

pumps_points_file<-read.csv('Snow_pumps.csv', header = TRUE)

street_points_file<-read.csv('Snow_streets.csv', header = TRUE)
```

exploratory data analysis of point process data

```
library(spatstat)

deaths_points_file<-read.csv('Snow_deaths.csv', header = TRUE)

pumps_points_file<-read.csv('Snow_pumps.csv', header = TRUE)

street_points_file<-read.csv('Snow_streets.csv', header = TRUE)
```

```
> head(deaths_points_file)
  case         x         y
1    1 13.588010 11.095600
2    2  9.878124 12.559180
3    3 14.653980 10.180440
4    4 15.220570  9.993003
5    5 13.162650 12.963190
6    6 13.806170  8.889046
```

exploratory data analysis of point process data

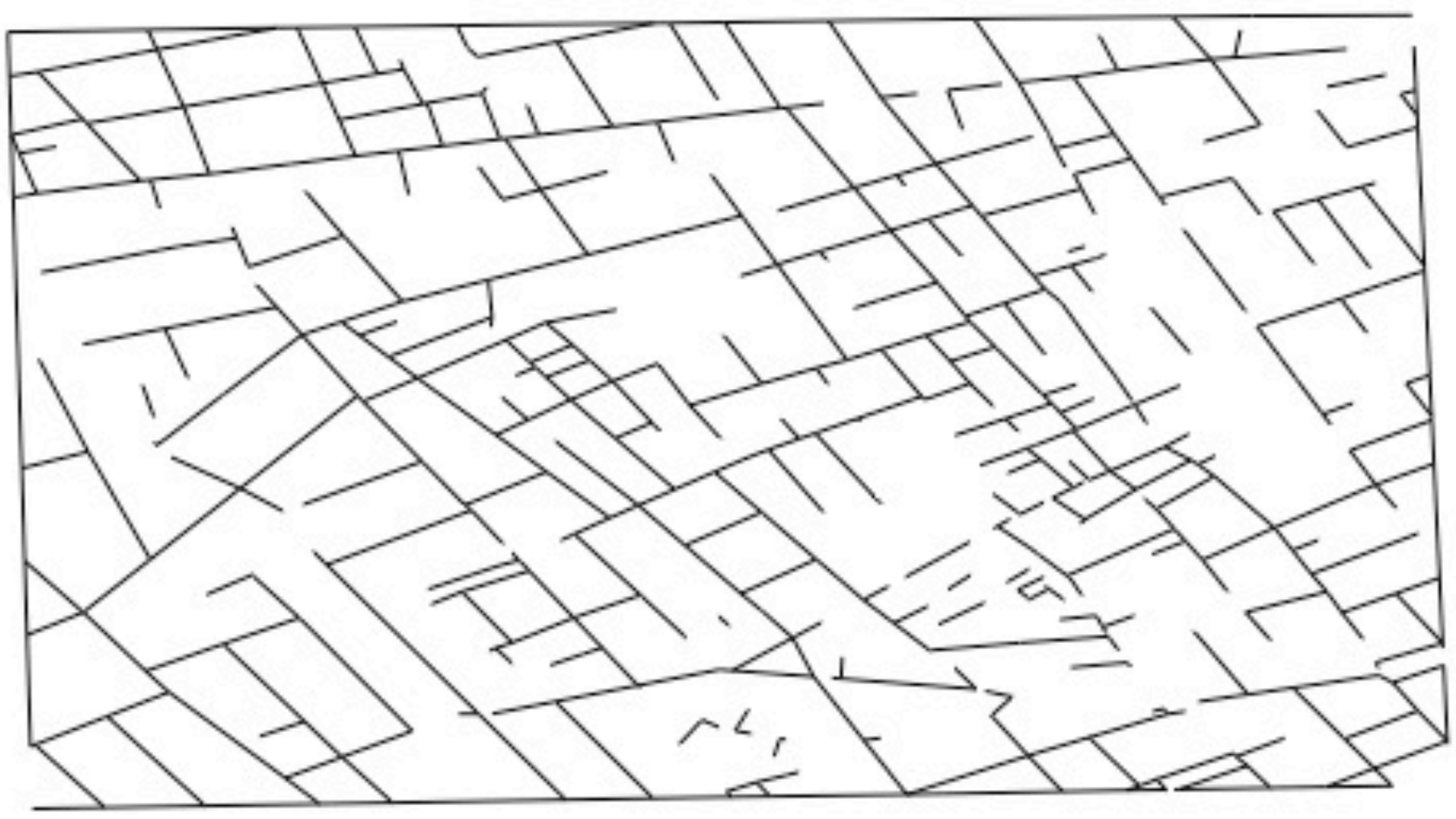
```
library(spatstat)

deaths_points_file<-read.csv('Snow_deaths.csv', header = TRUE)

pumps_points_file<-read.csv('Snow_pumps.csv', header = TRUE)

street_points_file<-read.csv('Snow_streets.csv', header = TRUE)
```

```
> head(deaths_points_file)
  case        x        y
1    1 13.588010 11.095600
2    2  9.878124 12.559180
3    3 14.653980 10.180440
4    4 15.220570  9.993003
5    5 13.162650 12.963190
6    6 13.806170  8.889046
```

```
> head(pumps_points_file)
  pump           label        x        y
1    1  Oxford Market  8.651201 17.89160
2    2    Castle St E 10.984780 18.51785
3    3   Oxford St #1 13.378190 17.39454
4    4   Oxford St #2 14.879830 17.80992
5    5 Gt Marlborough  8.694768 14.90547
6    6   Crown Chapel  8.864416 12.75354
```

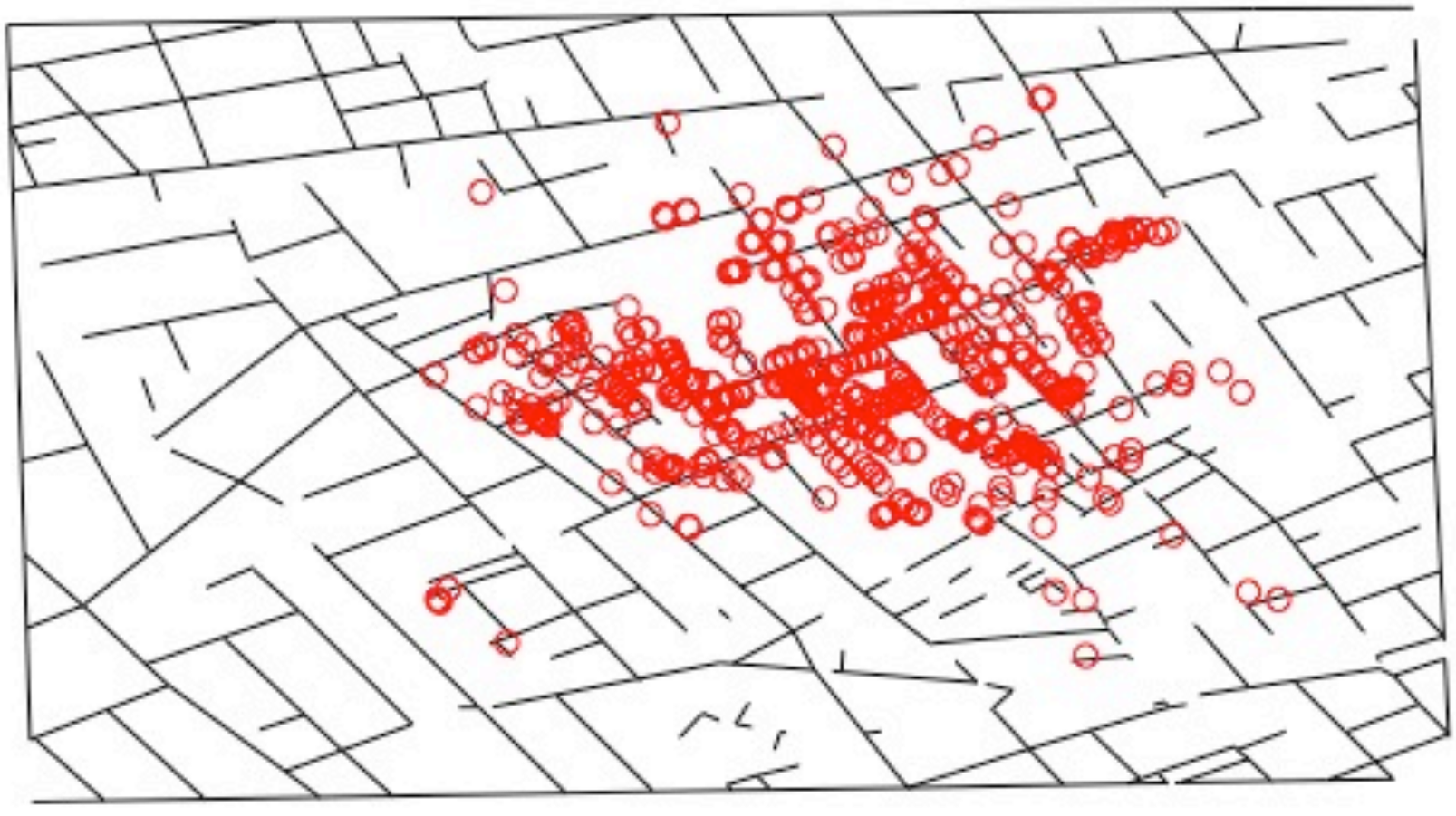
exploratory data analysis of point process data

```
par(mfrow=c(1,1))
plot(NA, NA, xlim = range(street_points_file$x), ylim =
range(street_points_file$y), xlab = '', ylab = '', bty = 'n', xaxt = 'n',
yaxt = 'n')

for(ii in 1:max(street_points_file$street)){
  sub_dat<-street_points_file[which(street_points_file$street == ii),]
  lines(c(sub_dat$x[1], sub_dat$x[2]), c(sub_dat$y[1], sub_dat$y[2]))
}
```



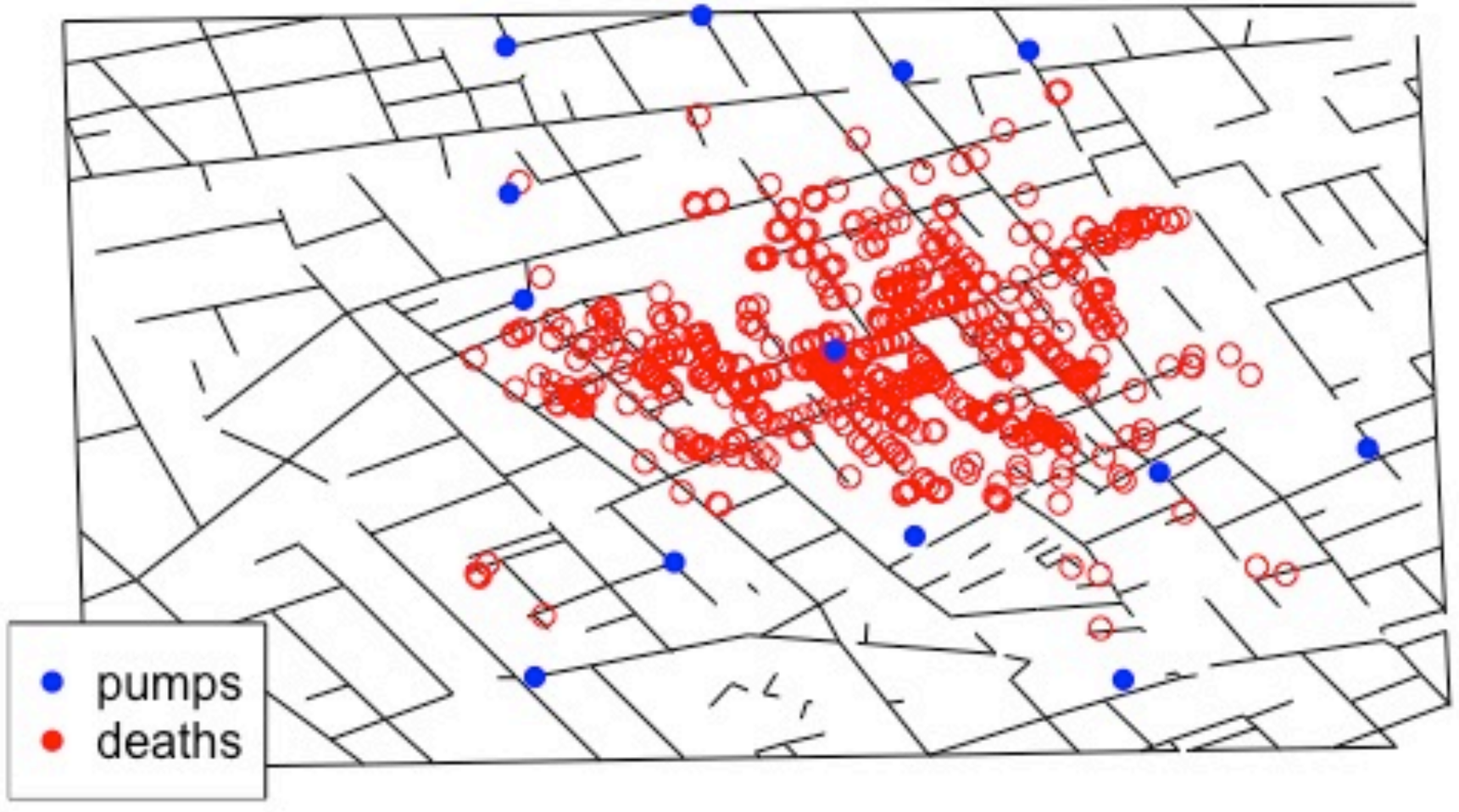
exploratory data analysis of point process data

```
points(deaths_points_file$x, deaths_points_file$y, col = 'red')

points(pumps_points_file$x, pumps_points_file$y, col = 'blue', pch = 16)

legend('bottomleft', legend = c('pumps', 'deaths'), col = c('blue', 'red'),
pch = 16)
```



exploratory data analysis of point process data

```
points(deaths_points_file$x, deaths_points_file$y, col = 'red')

points(pumps_points_file$x, pumps_points_file$y, col = 'blue', pch = 16)

legend('bottomleft', legend = c('pumps', 'deaths'), col = c('blue', 'red'),
pch = 16)
```



exploratory data analysis of point process data

```
r <- raster(nrows = 20, ncol = 10, xmn = min(street_points_file$x), xmx =
max(street_points_file$x), ymn = min(street_points_file$y), ymx =
max(street_points_file$y), values = FALSE)
```

exploratory data analysis of point process data

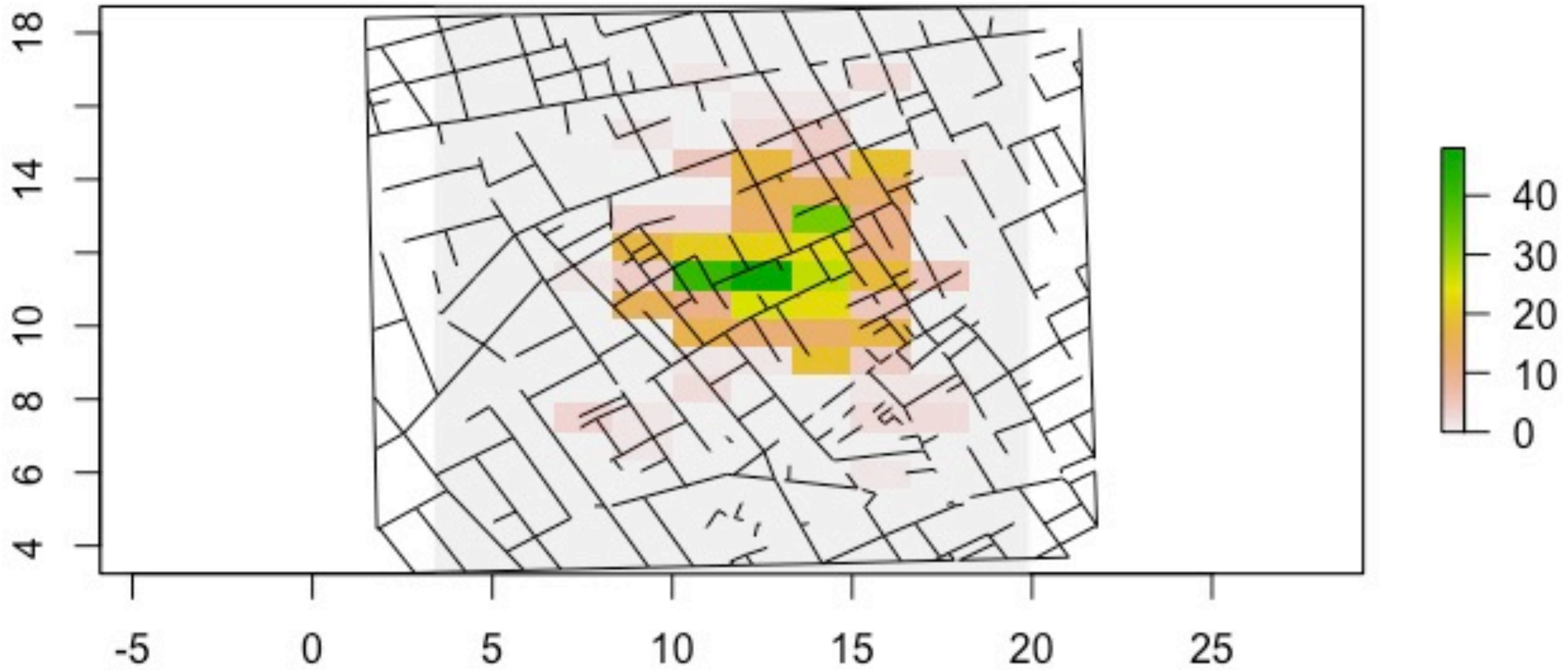
```
r <- raster(nrows = 20, ncol = 10, xmn = min(street_points_file$x), xmx =
max(street_points_file$x), ymn = min(street_points_file$y), ymx =
max(street_points_file$y), values = FALSE)

test.1<-rasterize(deaths_points_file[,c('x', 'y')], r, fun = 'count',
background = 0)
```

exploratory data analysis of point process data

```
r <- raster(nrows = 20, ncol = 10, xmn = min(street_points_file$x), xmx =
max(street_points_file$x), ymn = min(street_points_file$y), ymx =
max(street_points_file$y), values = FALSE)

test.1<-rasterize(deaths_points_file[,c('x', 'y')], r, fun = 'count',
background = 0)
```

```
plot(test.1)
  for(ii in 1:max(street_points_file$street)){
    sub_dat<-street_points_file[which(street_points_file$street == ii),]
    lines(c(sub_dat$x[1], sub_dat$x[2]), c(sub_dat$y[1], sub_dat$y[2]))
  }
```

exploratory data analysis of point process data



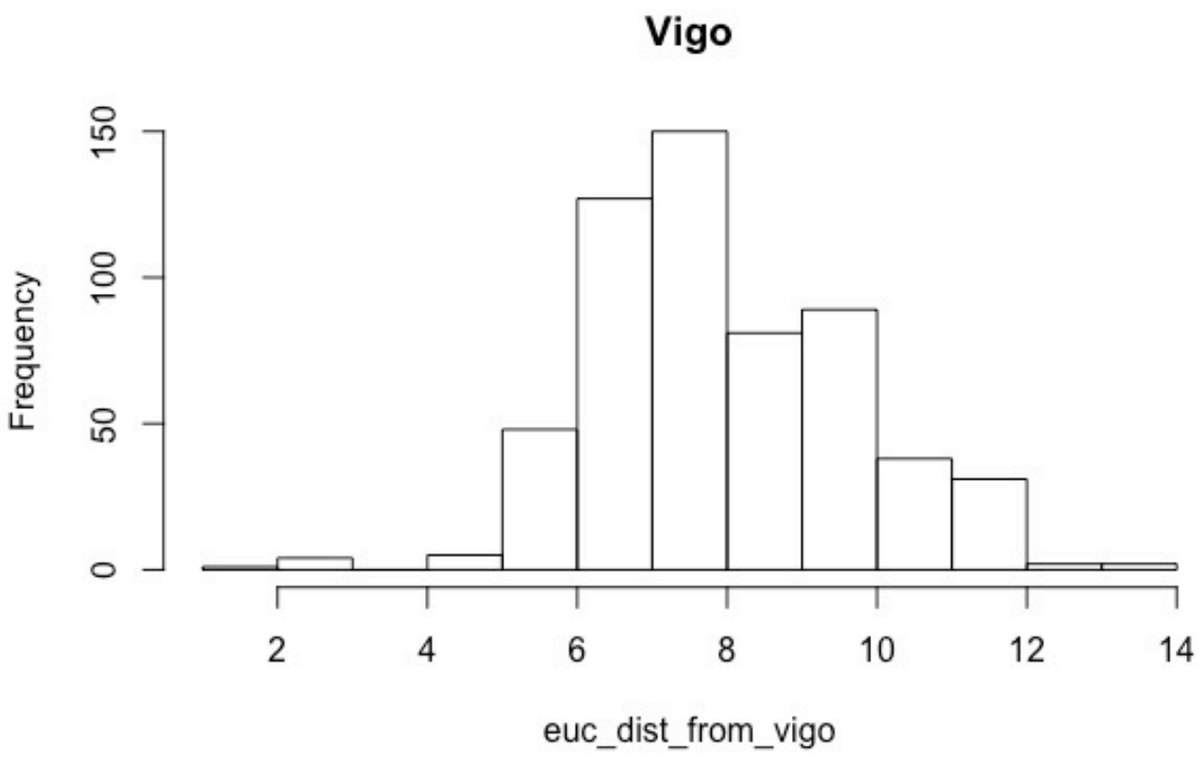
exploratory data analysis of point process data

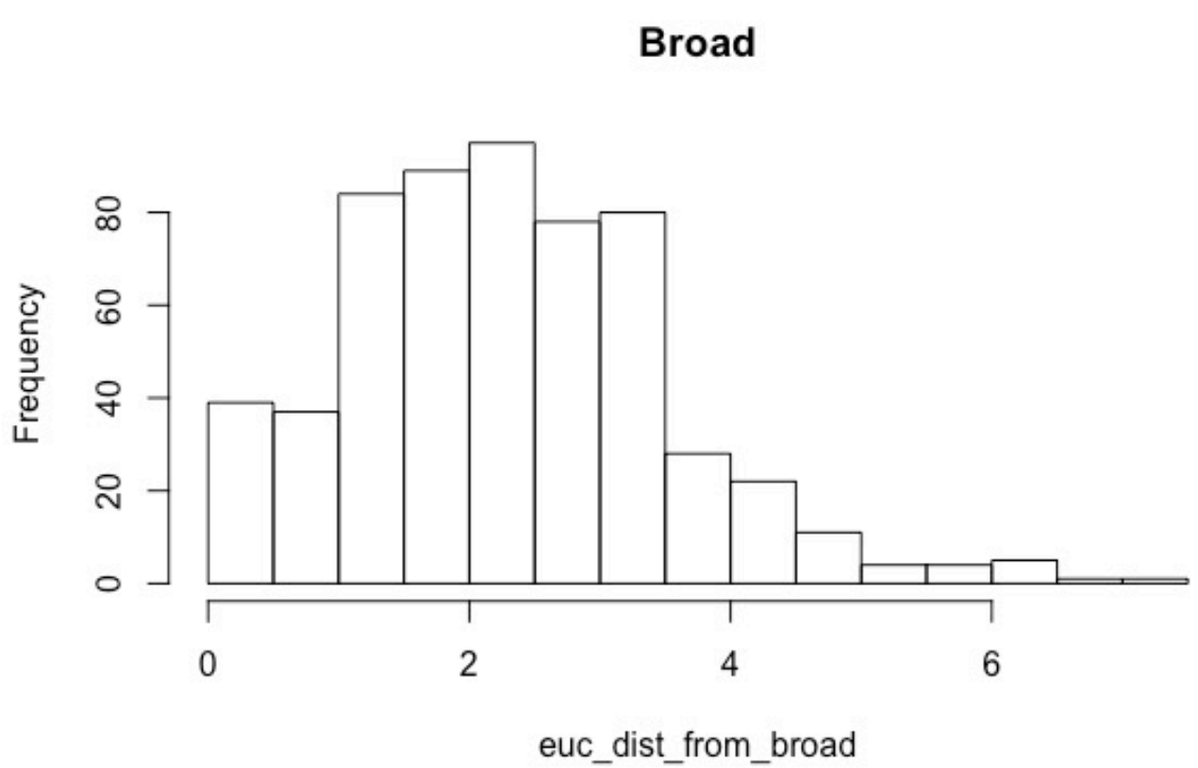
```
broad_coord<-pumps_points_file[which(pumps_points_file$label == 'Broad St'),]
euc_dist_from_broad<-sqrt((deaths_points_file$x - broad_coord$x)^2 +
(deaths_points_file$y - broad_coord$y)^2)
```

```
vigo_coord<-pumps_points_file[which(pumps_points_file$label == 'Vigo St'),]
euc_dist_from_vigo<-sqrt((deaths_points_file$x - vigo_coord$x)^2 +
(deaths_points_file$y - vigo_coord$y)^2)

hist(euc_dist_from_broad, main = 'Broad')
hist(euc_dist_from_vigo, main = 'Vigo')
```

exploratory data analysis of point process data

```
broad_coord<-pumps_points_file[which(pumps_points_file$label == 'Broad St'),]
euc_dist_from_broad<-sqrt((deaths_points_file$x - broad_coord$x)^2 +
(deaths_points_file$y - broad_coord$y)^2)

vigo_coord<-pumps_points_file[which(pumps_points_file$label == 'Vigo St'),]
euc_dist_from_vigo<-sqrt((deaths_points_file$x - vigo_coord$x)^2 +
(deaths_points_file$y - vigo_coord$y)^2)

hist(euc_dist_from_broad, main = 'Broad')
hist(euc_dist_from_vigo, main = 'Vigo')
```





1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

spatial modeling

```
library(mgcv)
dhs<-read.csv('~/Dropbox/Teaching/SpatialStatsMada/Madagascar2008-2009.csv',
header = TRUE)
good<-which(dhs$long != 0 & dhs$lat != 0, arr.ind = TRUE)
```

spatial modeling

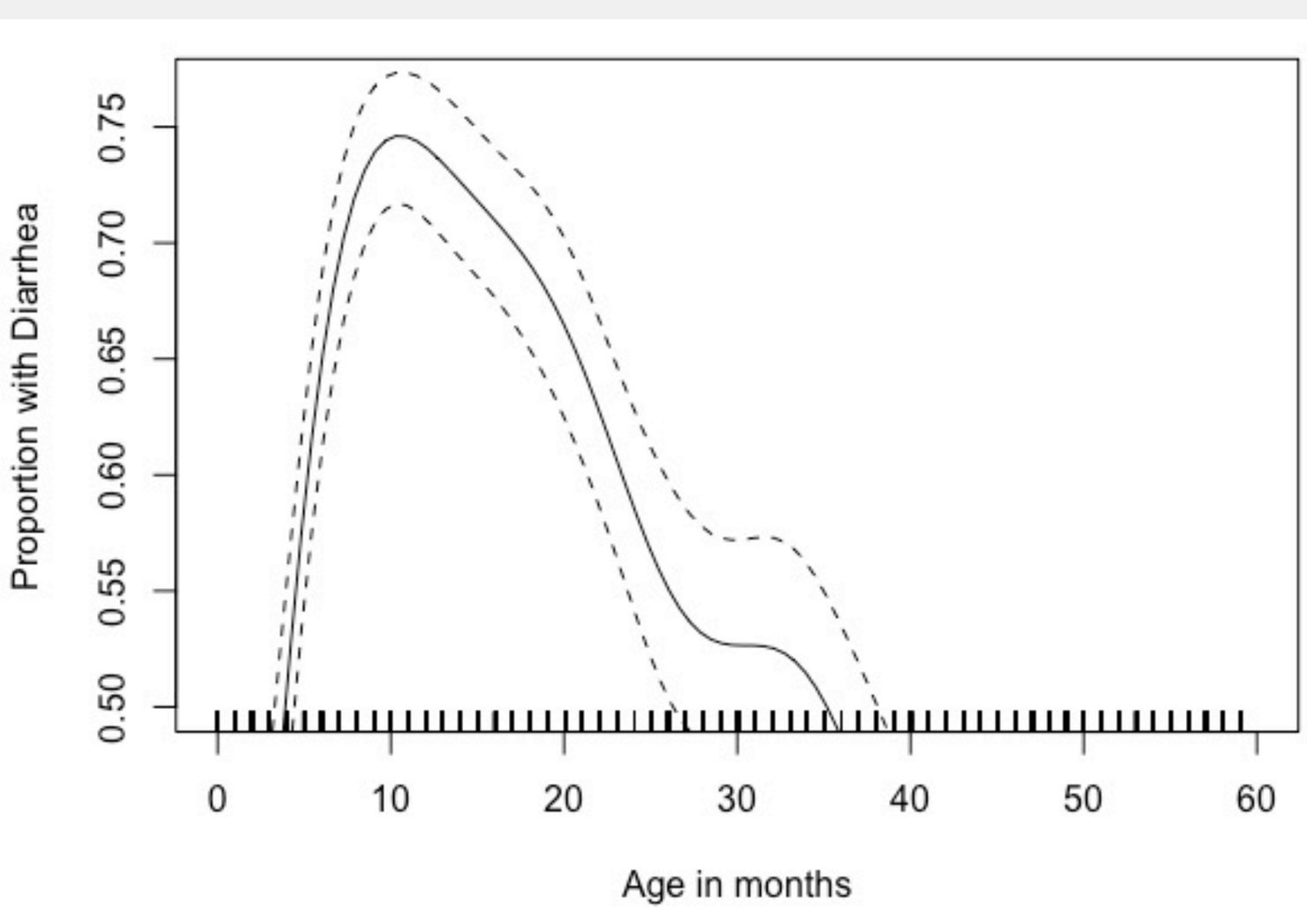
```
library(mgcv)
dhs<-read.csv('~/Dropbox/Teaching/SpatialStatsMada/Madagascar2008-2009.csv',
header = TRUE)
good<-which(dhs$long != 0 & dhs$lat != 0, arr.ind = TRUE)

fit<-gam(diarrhea~s(age.in.months)+s(long,lat), family = 'binomial', data =
dhs[good,])
summary(fit)
```

spatial modeling

```
> summary(fit)

Family: binomial
Link function: logit

Formula:
diarrhea ~ s(age.in.months) + s(long, lat)

Parametric coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.5954     0.0407  -63.77   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                     edf Ref.df Chi.sq p-value
s(age.in.months)  8.003  8.729  332.2  <2e-16 ***
s(long,lat)      25.343 28.168  156.7  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0508   Deviance explained = 8.08%
UBRE = -0.4515  Scale est. = 1         n = 11349
```

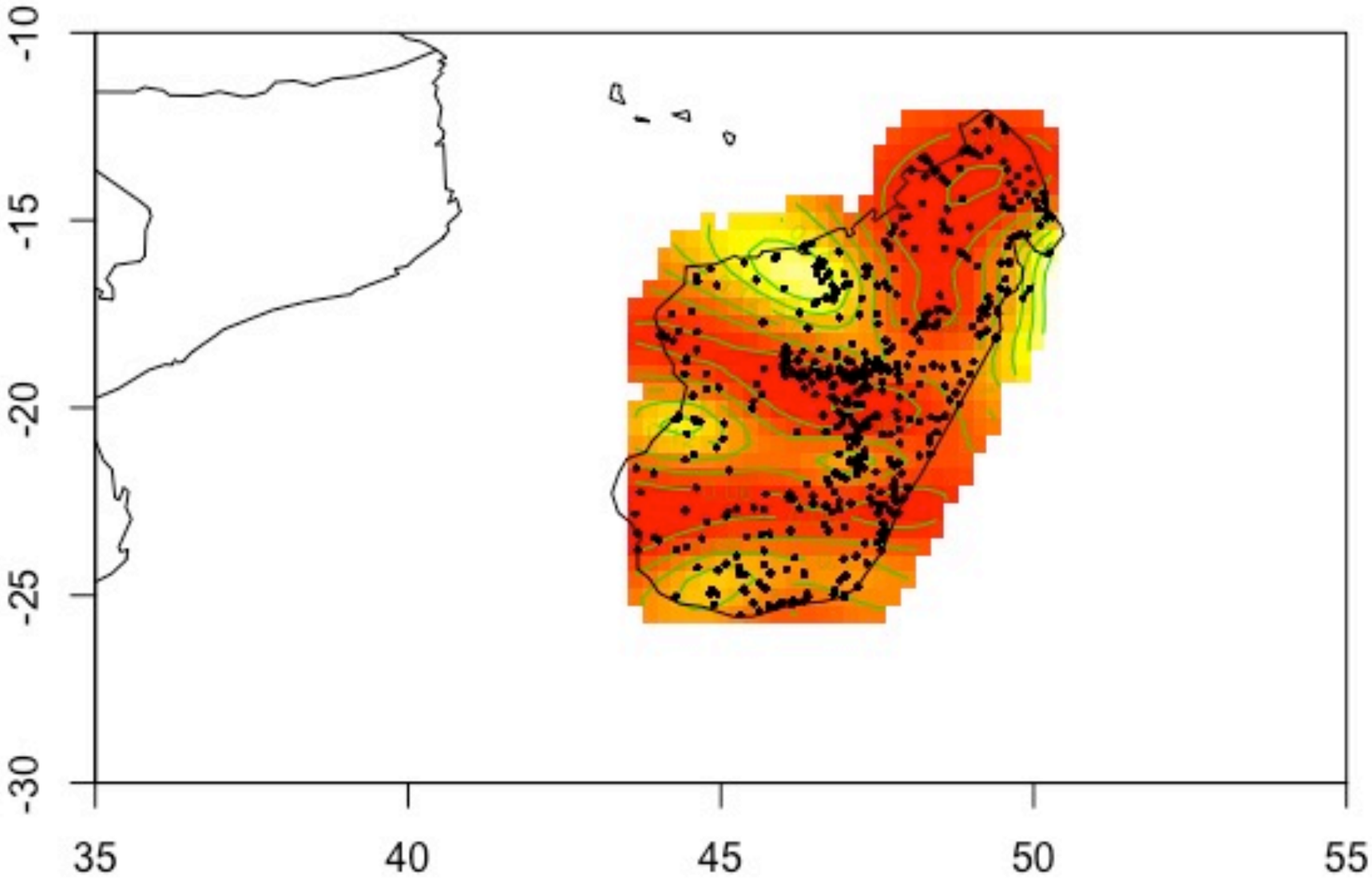
spatial modeling

```
fit<-gam(diarrhea~s(age.in.months)+s(long,lat), family = 'binomial', data =
dhs[good,])
summary(fit)

plot(fit, select = 1, trans = function(x)exp(x)/(1+exp(x)), xlab = 'Age in
months', ylab = 'Proportion vaccinated')
```



spatial modeling

```
vis.gam(fit,view=c("long","lat"),plot.type="contour",too.far=0.1,type="respon
se",color="heat",ylim=c(-30,-10),xlim=c(35,55), xlab="", ylab="", main="")
points(dhs$long[good], dhs$lat[good], pch=19,cex=0.2)
map(add=TRUE)
```



1: Introduction - Why do spatial analysis?
Ino râ azontsika avy amin'ny 'spatial analysis'?

2: Visualizing spatial data sets
Software
Different types of spatial data
Reading in and mapping

3: Exploratory data analysis
Mapping multiple sources of spatial data
Summary statistics

4: Spatial modeling
Finding relationships between spatial variables

Key Concepts

1. There are many different forms of spatial data that can be read into R using specialized functions and packages
   Il existe de nombreuses formes de données spatiales pouvant être lues dans R en utilisant des fonctions et des packages spécialisés.

2. It is important that you understand the form of the data and what the spatial scale and variables represent.
   Il est important que vous compreniez la forme des données et ce que représentent l'échelle spatiale et les variables.

3. Combining these data in the forms of plots or different statistics can be used to analyze and describe the relationship between these data.
   La combinaison de ces données sous la forme de graphiques ou de statistiques différentes peut être utilisée pour analyser et décrire la relation entre ces données.

Additional helpful resources

http://rspatial.org/analysis/rst/8-pointpat.html

http://zevross.com/blog/2015/03/30/map-and-analyze-raster-data-in-r/

https://cran.r-project.org/doc/contrib/intro-spatial-rl.pdf

https://data.cdrc.ac.uk/tutorial/an-introduction-to-spatial-data-analysis-and-visualisation-in-r

http://www.stats.uwo.ca/faculty/kulperger/S9934a/Computing/Spatstat-pn0y.pdf

loading libraries

```
library(maptools)
library(raster)
library(rgdal)
library(rgeos)
library(colorRamps)
library(mgcv)
library(maps)
library(googleway)
library(leaflet)
```