# JohnSnow

*Joshua*

*11 septembre 2018*

## Getting the data from the shapefiles

First, unzip the data

```
unzip('SnowGIS_SHP.zip')
```

Then get the number of deaths and their position

```
raw_deaths <- rgdal::readOGR(dsn = "SnowGIS_SHP/", layer = "Cholera_Deaths")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/joshua/Documents/M2/METH/Sujet_C/JohnSnow/SnowGIS_SHP", layer: "Cholera_Deaths"
## with 250 features
## It has 2 fields
```

```
deaths <- as.data.frame(raw_deaths@coords) #Creates an array with coordinates and counts of deaths
deaths$count <- raw_deaths$Count
```

Then the position of pumps

```
raw_pumps <- rgdal::readOGR(dsn = "SnowGIS_SHP/", layer = "Pumps")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/joshua/Documents/M2/METH/Sujet_C/JohnSnow/SnowGIS_SHP", layer: "Pumps"
## with 8 features
## It has 1 fields
```

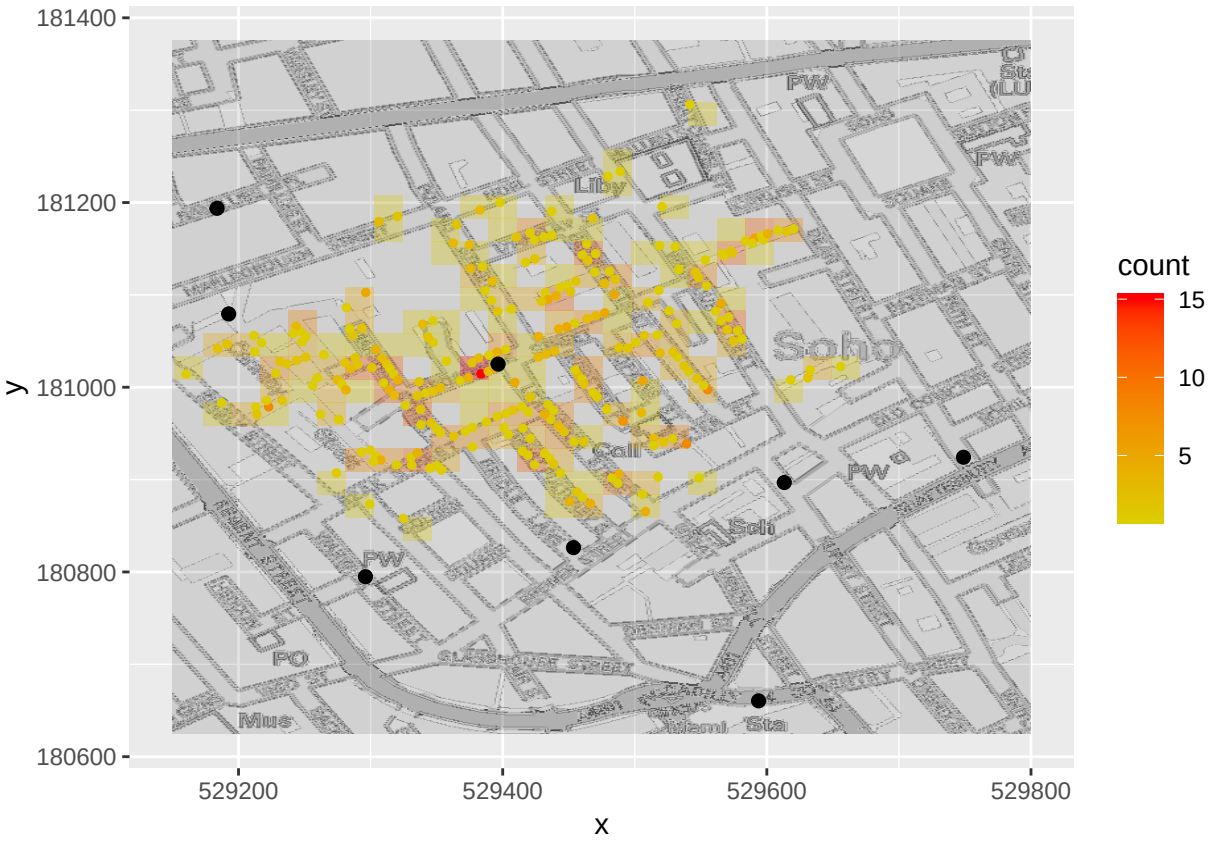
```
pumps <- as.data.frame(raw_pumps@coords)
```

Finally plot everything. The map is taken from an image. We make a 2D histogram as well as plotting for the deaths. Pumps are left as black dots.

```
rel <- raster("SnowGIS_SHP/OSMap.tif")
rel_spdf <- as(rel, "SpatialPixelsDataFrame")
rel <- as.data.frame(rel_spdf)
rel <- subset(rel, x > 529150 & x < 529800 & y > 180625 & y < 181375)

ggplot() +
  geom_raster(data = rel, aes_string(x = "x", y = "y", alpha = "OSMap")) +
  scale_alpha(name = "", range = c(0.9, 0.1), guide = F) +
  scale_colour_gradient(low = "#DDCC00", high = "red") +
  scale_fill_gradient(low = "#DDCC00", high = "red") +
  geom_bin2d(data = deaths, aes(x = coords.x1, y = coords.x2), binswidth = c(500, 500), alpha = 0.3, sh
  geom_point(data = deaths, aes(x = coords.x1, y = coords.x2, colour = count), size = 1) +
  geom_point(data = pumps, aes(x = coords.x1, y = coords.x2), colour = "black", size = 2)
```

```
## Warning: Ignoring unknown parameters: binswidth
```


181400
181200
181000
180800
180600
y
529200
529400
529600
529800
x
count
15
10
5
Soho
Liby
Coll
Pw
PO
Mus