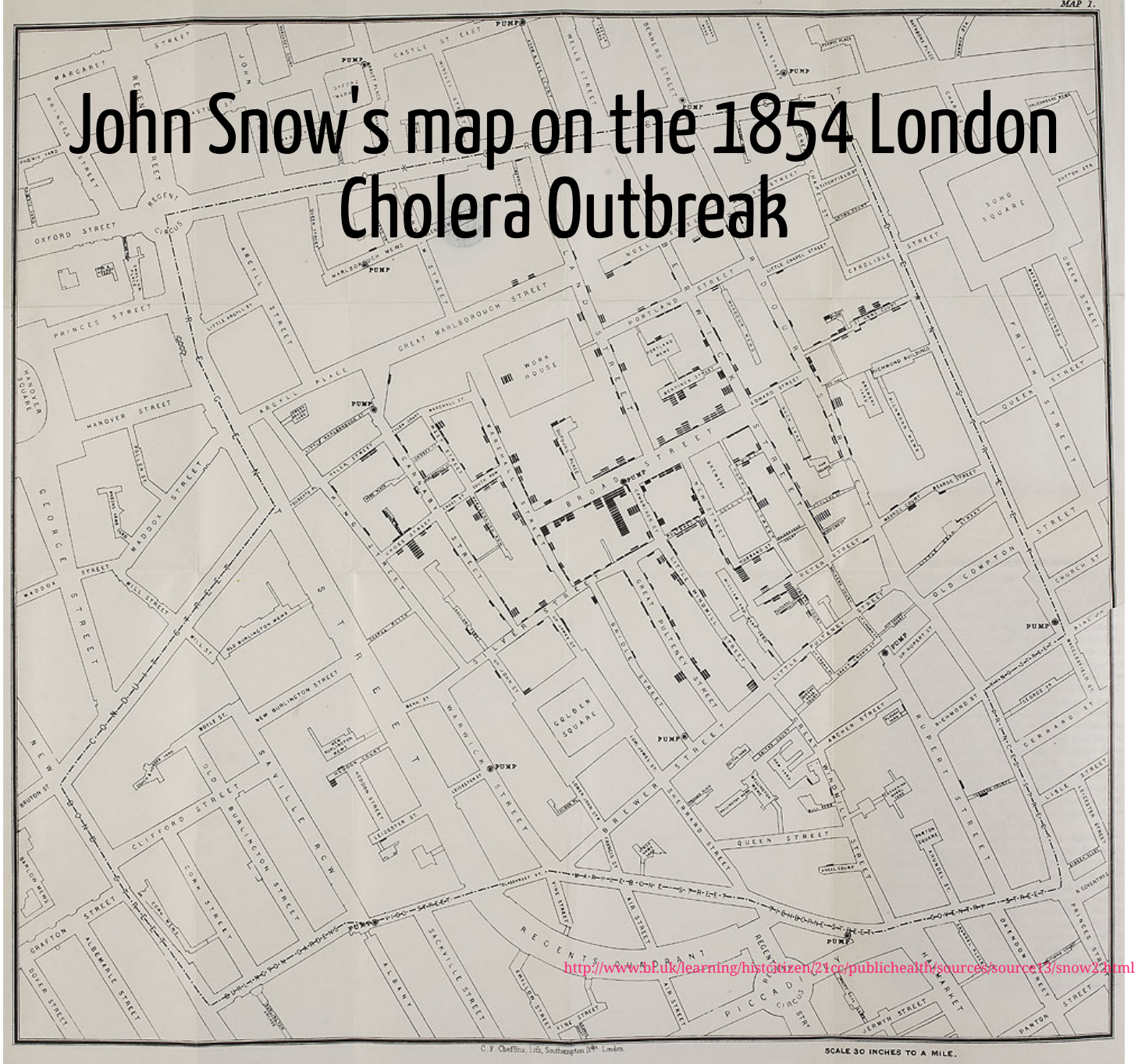
Spatial aggregation and cluster detection

Facundo Muñoz
facundo.munoz@cirad.fr
famuvie
cirad

ASTRE
AnimalS health
Territories Risks Ecosystems

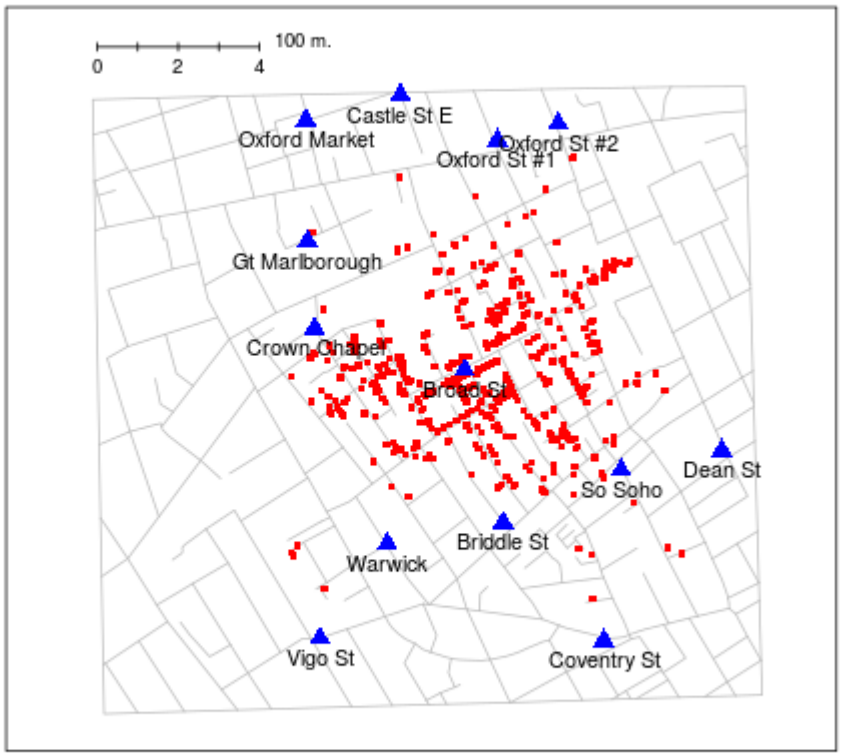
Biostatistics for epidemiology
December 2018

John Snow's map on the 1854 London Cholera Outbreak

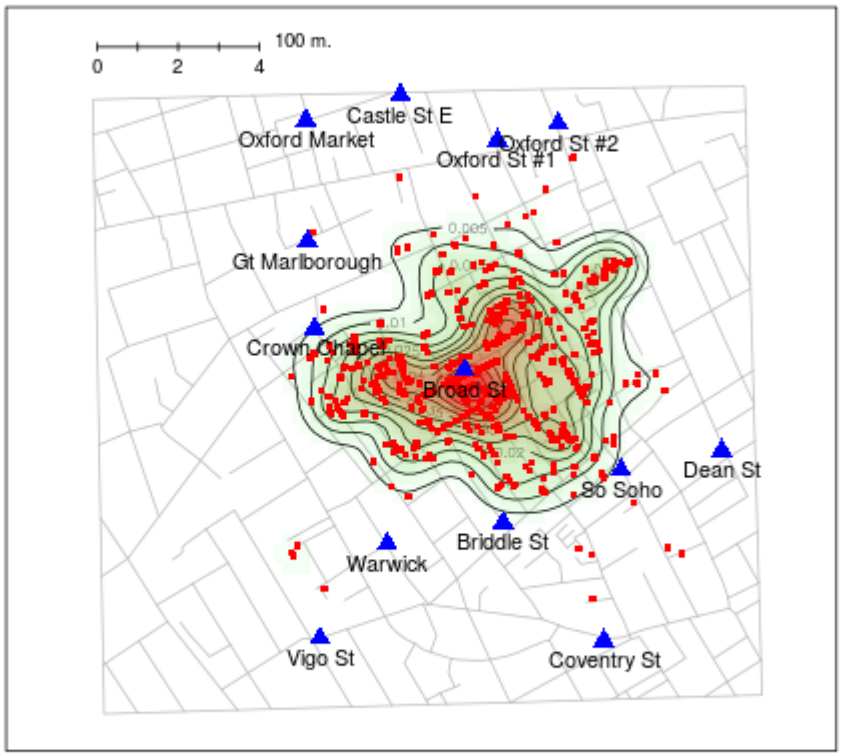


http://www.bl.uk/learning/histcitizen/21cc/publichealth/sources/source13/snow2.html

Water pumps and deaths



Latent *density* of deaths



Key questions

- Do cases tend to occur in **clusters**?

  (assessment of **aggregation**)

- Is there an **unusual group** of cases?

  (cluster **detection**)

- Is the case occurrence **related** with something else?

  (risk factor **identification**)

Key questions

- Do cases tend to occur in **clusters**?

  (assessment of **aggregation**)

- Is there an **unusual group** of cases?

  (cluster **detection**)

- Is the case occurrence **related** with something else?

  (risk factor **identification**)

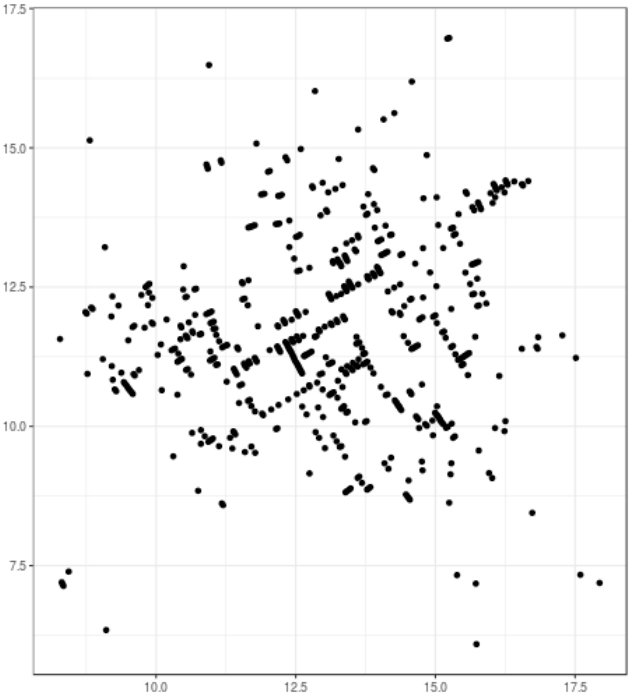
Spatial(-temporal) distribution of point patterns

in the previous terms,

What do you think Snow did?

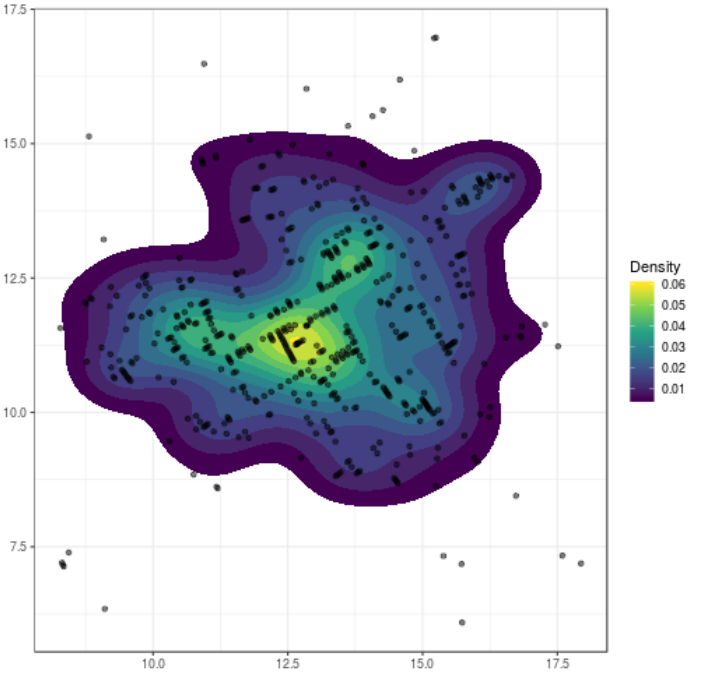
Visualisation

```
Snow.deaths %>%
  ggplot(aes(x, y)) +
  geom_point() +
  coord_fixed() +
  labs(x = NULL, y = NULL)
```



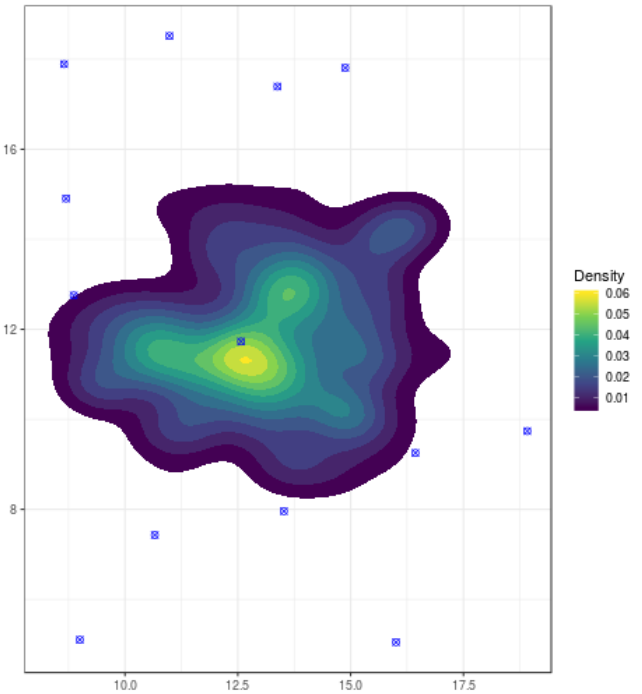
Point density estimation

```
Snow.deaths %>%
  ggplot(aes(x, y)) +
  stat_density_2d(
    aes(fill = stat(level)),
    geom = "polygon"
  ) +
  geom_point(alpha = .5) +
  scale_fill_viridis_c(
    name = "Density"
  ) +
  coord_fixed() +
  labs(x = NULL, y = NULL)
```

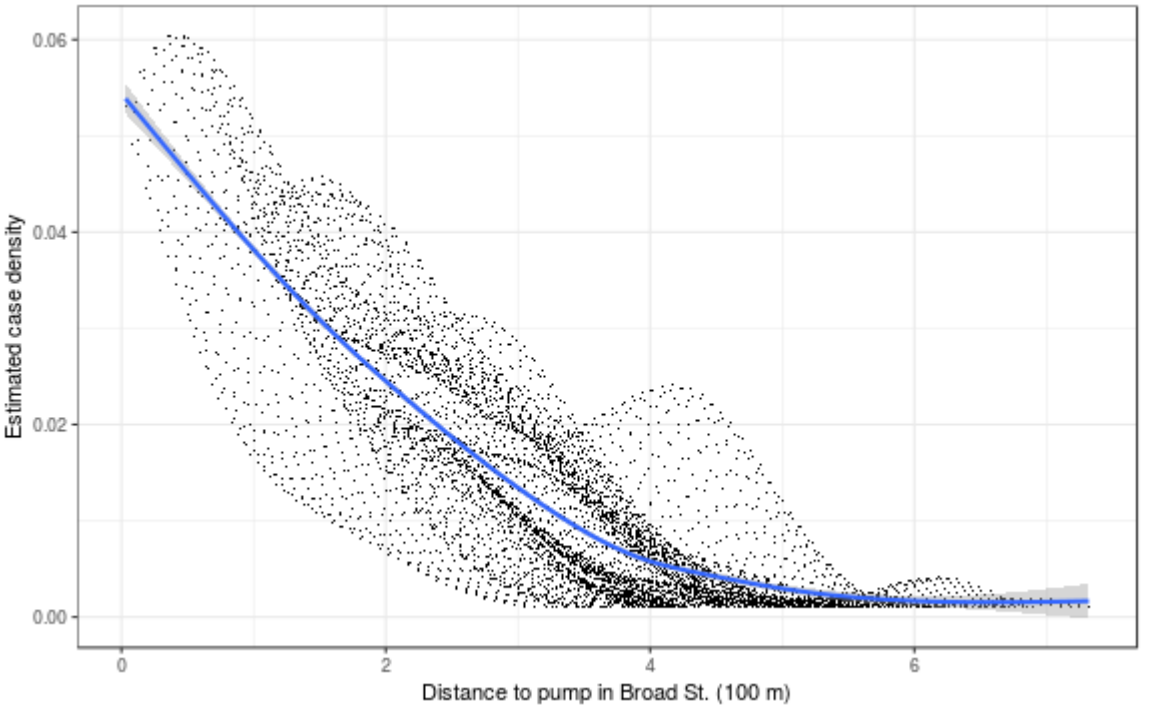


Risk factor hypothesis

```
Snow.deaths %>%
  ggplot(aes(x, y)) +
  stat_density_2d(
    aes(fill = stat(level)),
    geom = "polygon"
  ) +
  # geom_point(alpha = .5) +
  geom_point(
    data = Snow.pumps,
    shape = 13,
    size = 2,
    col = "blue"
  ) +
  scale_fill_viridis_c(
    name = "Density"
  ) +
  coord_fixed() +
  labs(x = NULL, y = NULL)
```



Risk factor assessment



Note: despite the *significant* effect, it does not **prove** causality (right?)

Objectives for the session

- Discuss **methods** and **tools** for quantification and detection of global aggregation, local clusters and risk factor evaluation

Objectives for the session

- Discuss **methods** and **tools** for quantification and detection of global aggregation, local clusters and risk factor evaluation
- Both in **spatial** and **spatio-temporal** settings

Objectives for the session

- Discuss **methods** and **tools** for quantification and detection of global aggregation, local clusters and risk factor evaluation
- Both in **spatial** and **spatio-temporal** settings
- For both **areal** and **point** data

Objectives for the session

- Discuss **methods** and **tools** for quantification and detection of global aggregation, local clusters and risk factor evaluation
- Both in **spatial** and **spatio-temporal** settings
- For both **areal** and **point** data
- **Operational approach**: work out 3 case-studies and discussing topics as needed