Do disease clusters have a common cause?

Jonathan Auerbach

9/28/2018

How do scientists determine whether a cluster of diseases have a common cause?

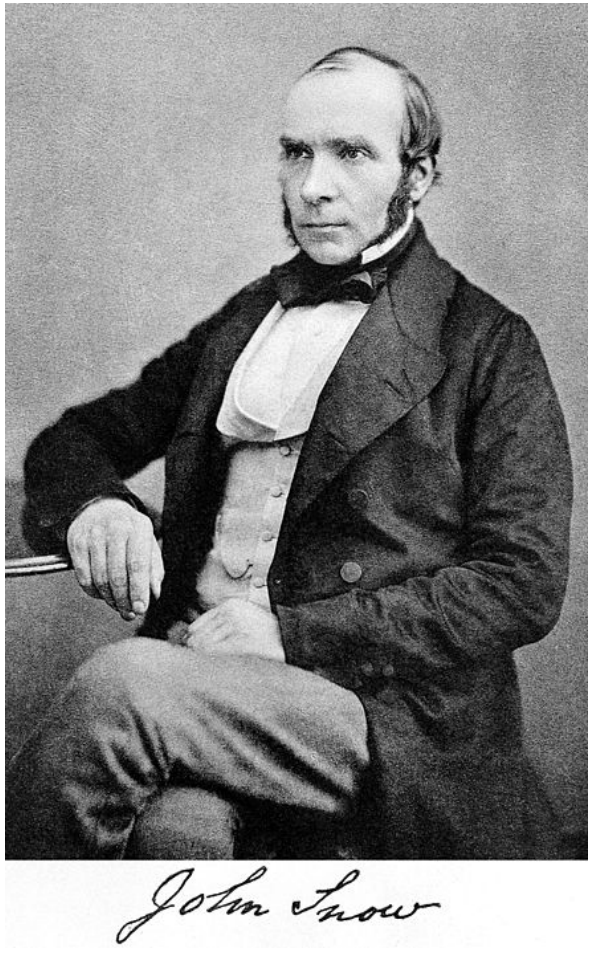
Setup:

```
library("knitr")
library("HistData")
library("tidyverse")
library("ggmap")
library("sp")
theme_set(theme_bw())
```

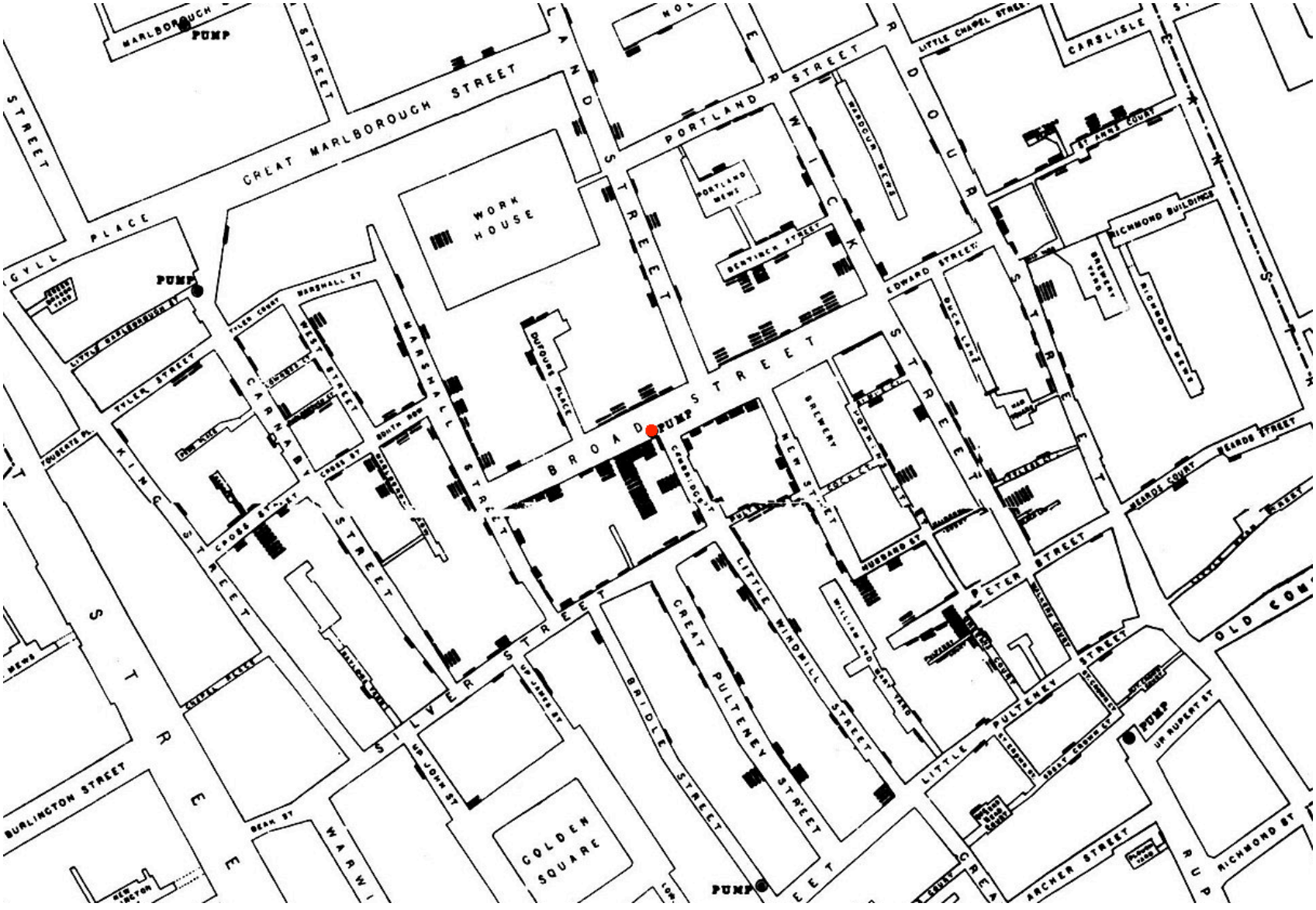
How convincing is a cluster of individuals with the same disease?

- In 1853, a cholera outbreak killed more than ten thousand people.
- Scientists disagreed on whether the cause was airborne or waterborne.
- Snow (1886) mapped the location of every documented cholera case.
- He noticed that the cases concentrated around the Broad Street (water) pump

John Snow (1856)

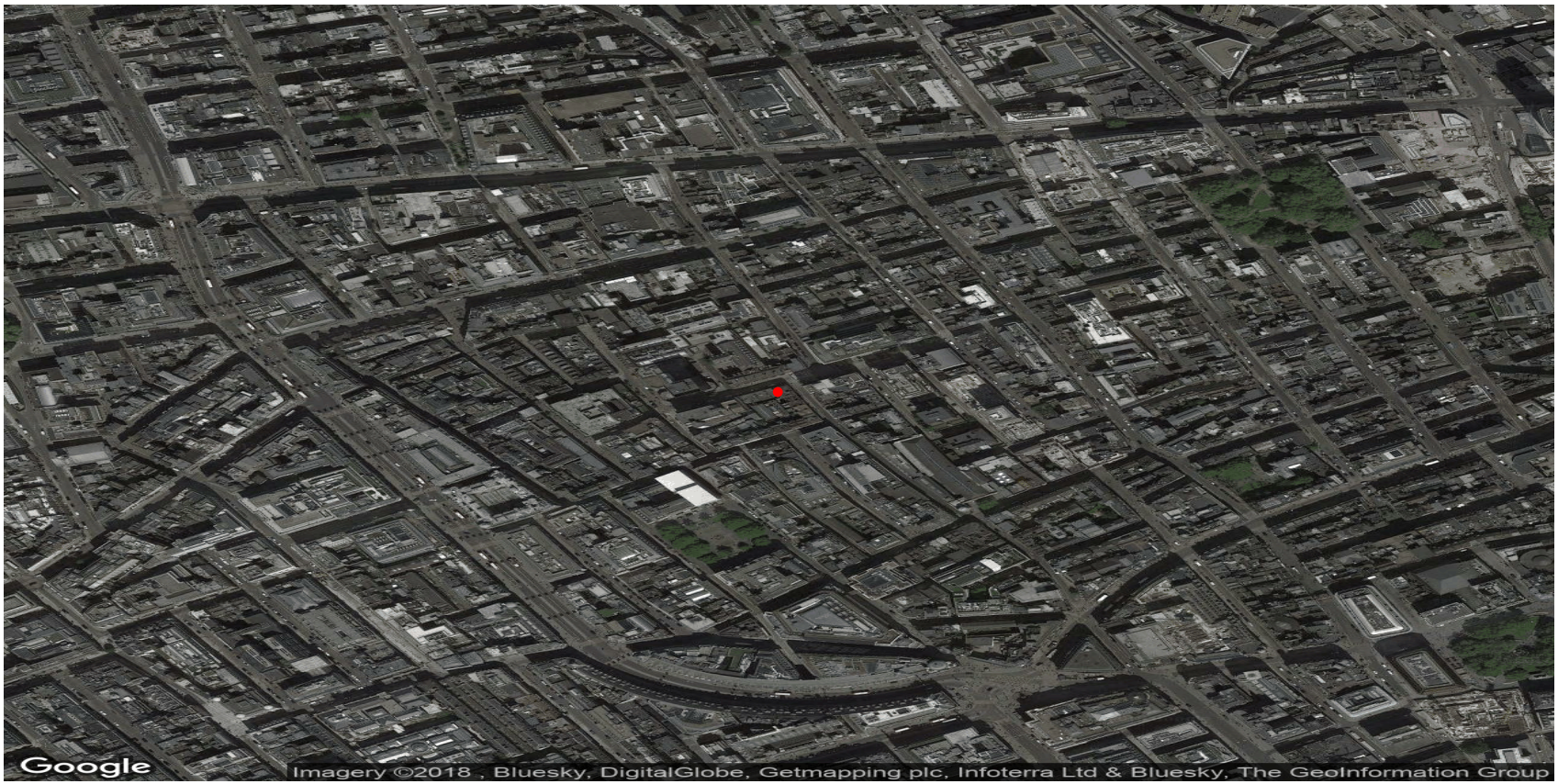


Broad Street Pump and Cholera Cases (London, 1854)



Broad Street Pump (London, 2018)

```
(map <- ggmap(get_map("Broad St Pump, London",
                zoom = 16, maptype = "satellite")) +
          geom_point(aes(X1, X2), color = "red",
                  data = pump))
```



Broad Street Pump (2018) and Snow Map

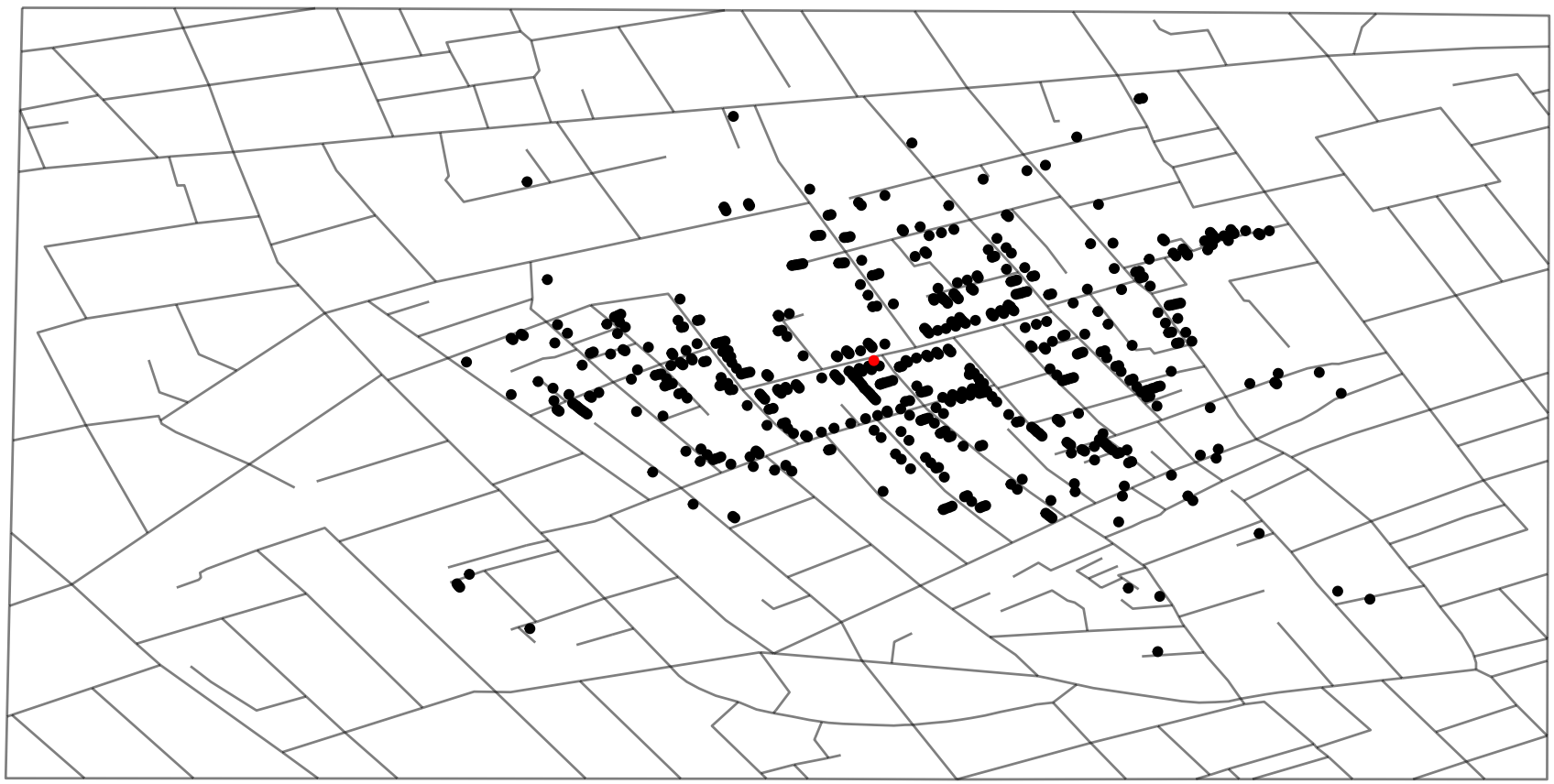
```
map +
  geom_path(aes(x=long, y=lat, group=group),
      color = "white", alpha = .5, size = 2,
      data = Snow_df)
```

Broad Street Pump (1854)

```
(map <- ggplot() + theme_nothing() +
    geom_path(aes(x=long, y=lat, group=group),
                color = "black", alpha = .5,
                data = Snow_df))
```

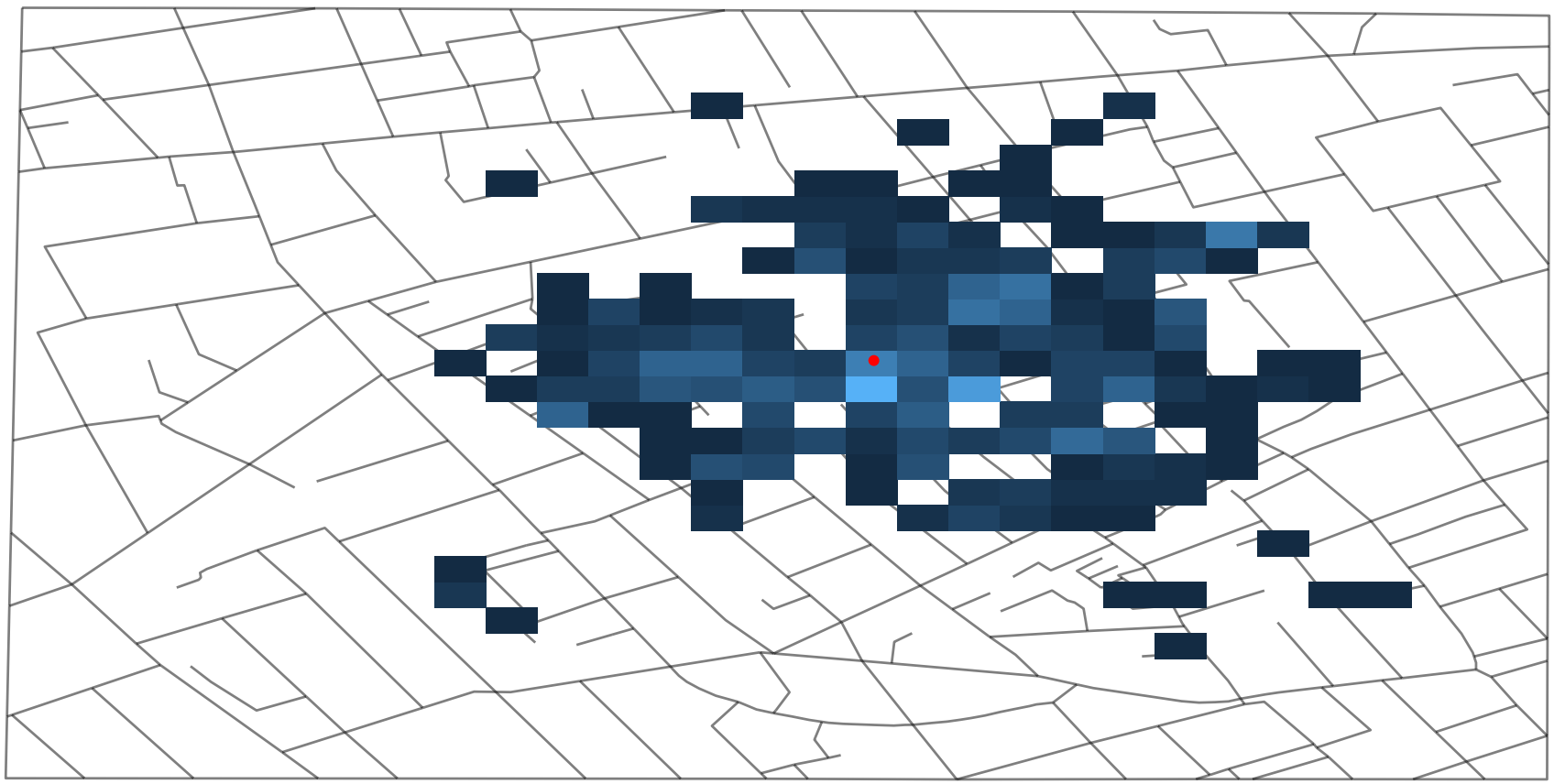
Broad Street Pump and Cholera Cases (1854)

```
map + geom_point(aes(x=long, y=lat),
                 data = Snow_deaths, color = "black") +
    geom_point(aes(X1, X2), color = "red",
               data = pump)
```



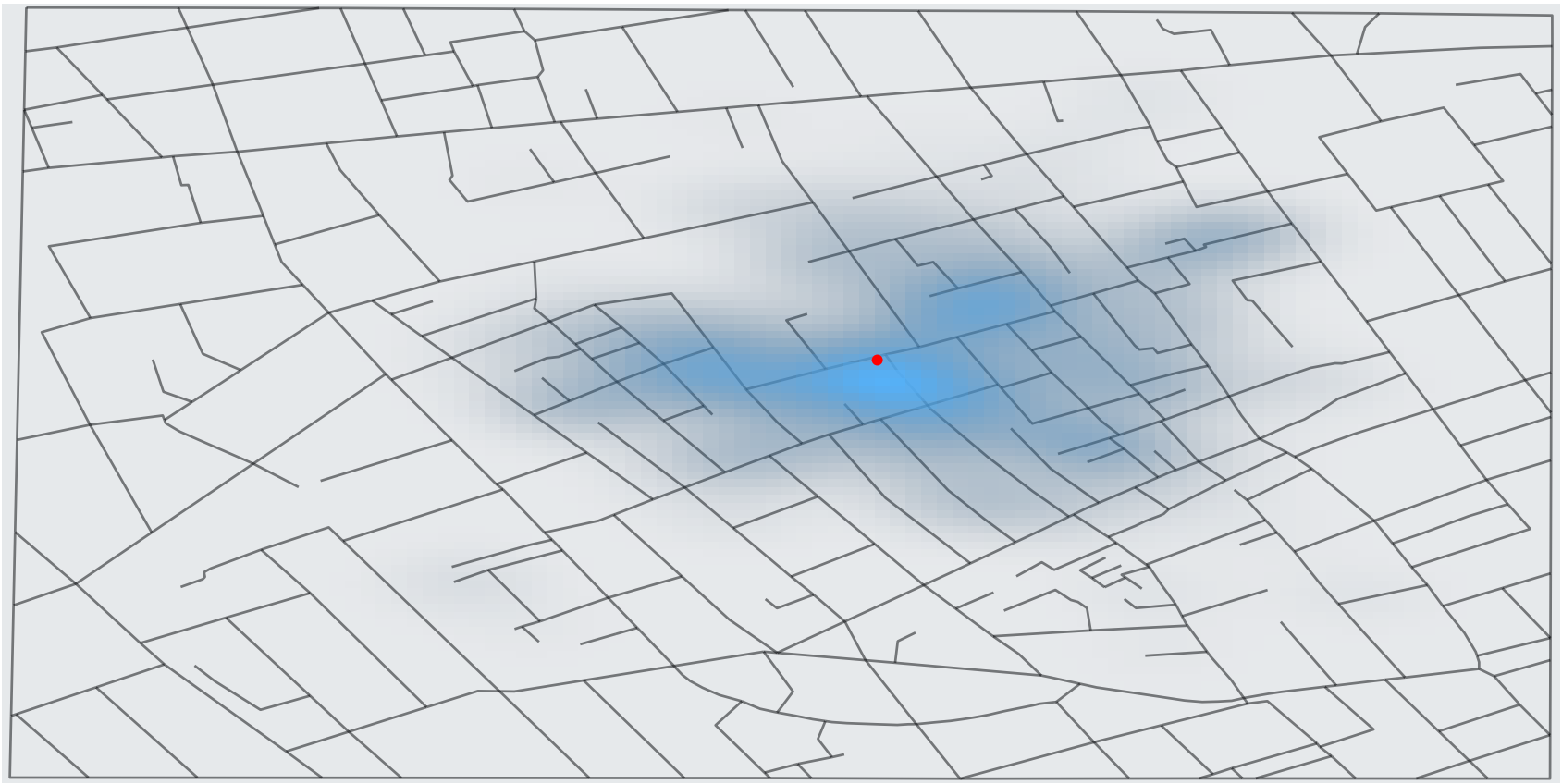
Broad Street Pump and 2D Histogram

```
map + stat_bin_2d(aes(x=long, y=lat),
                  data =Snow_deaths) +
  geom_point(aes(X1, X2), color = "red",
             data = pump)
```



Broad Street Pump and 2D Kernel Density Estimate

```
map + stat_density_2d(aes(x=long, y=lat,
  fill = stat(density), alpha = stat(density)),
  data =Snow_deaths,geom = "raster",contour = FALSE)+
  geom_point(aes(X1, X2), color = "red", data = pump)
```



What caused the Cholera Outbreak?

- The epidemic ended after Snow convinced the City to remove the handle to the Broad Street pump.
- A Cholera Inquiry Committee eventually identified patient zero: a five month old baby.
- Snow's study was historic. It is considered the classic example of good epidemiology.
- Similar observational evidence links cigarette smoking to lung cancer. (Cornfield et al. 1959)

Do all clusters have a cause?

- Movies popularize clusters as strong evidence of misconduct. e.g. Erin Brockovich (Hinkley, CA) and Lois Gibbs (Love Canal, NY)
- CDC is skeptical in general: "the likelihood of establishing a definitive cause-and-effect relationship between the health event and an exposure is slight"
- A 1989 national conference on disease clusters found that cluster studies rarely produce important findings.
- Goodman et al. reviewed over 500 cancer cluster investigations and found only one was able to identify a cause with certainty.

Most investigated clusters have no cause (Goodman et al. 2012)

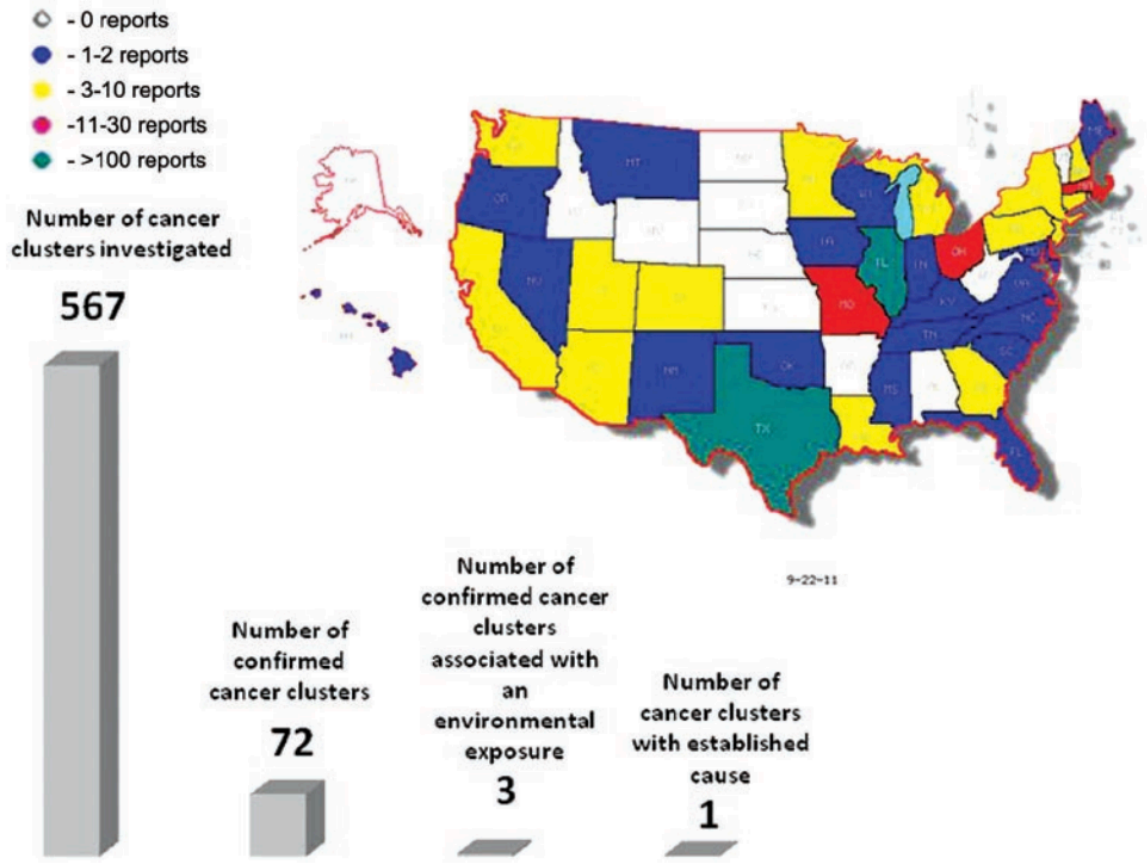
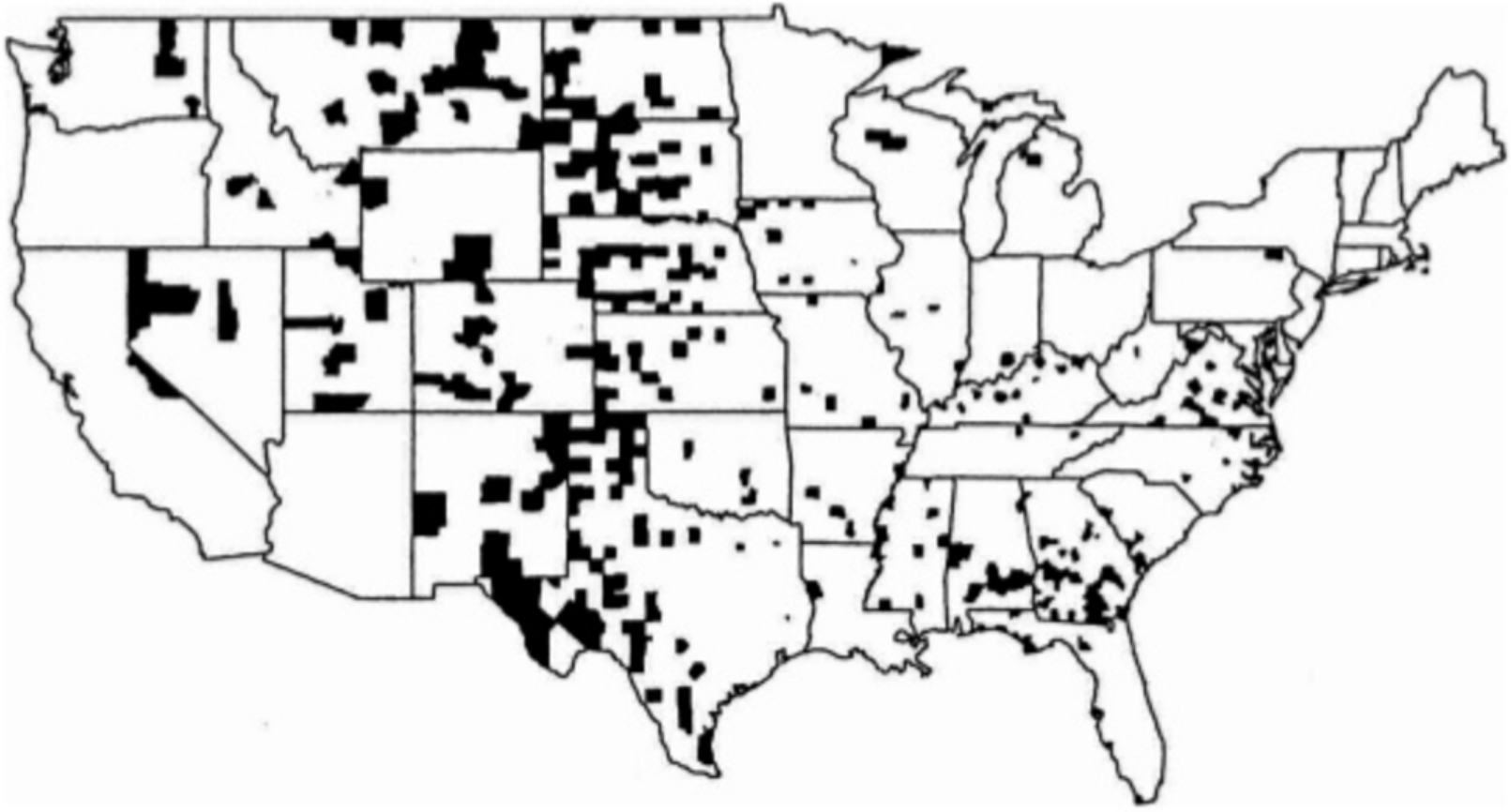


Figure 1. Numbers of publicly available cancer cluster investigation reports by state and comparison of numbers of investigated cancer clusters, confirmed cancer clusters (e.g. investigated clusters where number of cancer cases is greater than expected), clusters linked to an environmental exposure, and cancer clusters with an established cause. Although some of the cluster investigations may have been described in several reports, the numbers in this figure represent unique reported clusters. (Map generated from data in Table 1 using Map-Maker Utility, http://monarch.tamu.edu/~maps2/us_12.htm)

Highest Kidney Cancers Counties (Gelman and Nolan, 2017)

Lowest Kidney Cancer Counties (Gelman and Nolan, 2017)



References

1. Cornfield, Jerome, et al. "Smoking and lung cancer: recent evidence and a discussion of some questions." Journal of the National Cancer institute 22.1 (1959): 173-203.
2. Freedman, David A. "Statistical models and shoe leather." Sociological methodology (1991): 291-313.
3. Gelman, Andrew, and Deborah Nolan. Teaching statistics: A bag of tricks. Oxford University Press, 2017.
4. Goodman, Michael, et al. "Cancer clusters in the USA: what do the last twenty years of state and federal investigations tell us?." Critical reviews in toxicology 42.6 (2012): 474-490.
5. Paneth, Nigel, et al. "A rivalry of foulness: official and unofficial investigations of the London cholera epidemic of 1854." American Journal of Public Health 88.10 (1998): 1545-1553.
6. Snow, John. On the mode of communication of cholera. John Churchill, 1855.