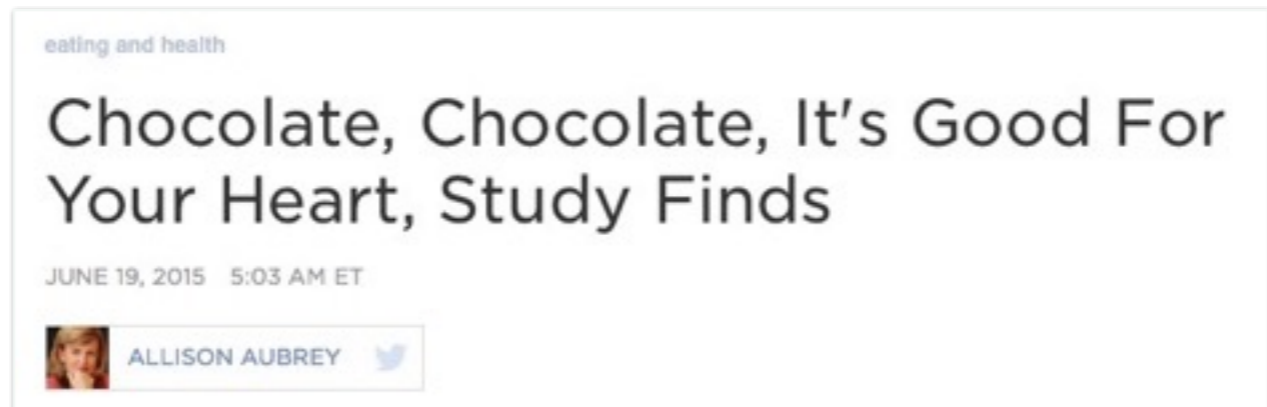


# Causality and Experiments



[npr.org](http://npr.org) (report on a study in [heart.bmj.com](http://heart.bmj.com))

# Observation

- individuals, study subjects, participants, units  
*European adults*

---

- treatment  
*chocolate consumption*
- outcome  
*heart disease*

## The first question

---

Is there **any relation** between chocolate consumption and heart disease?

- **association**  
“any relation”

## An answer

### Some data:

“Among those in the top tier of chocolate consumption, 12 percent developed or died of cardiovascular disease during the study, compared to 17.4 percent of those who didn’t eat chocolate.”

- Howard LeWine of Harvard Health Blog, reported by [npr.org](http://npr.org)

- Yes, this points to an association.

## The next question

Does chocolate consumption **lead to** a reduction in heart disease?

- causality

---

This question is often harder to answer.

“[The study] doesn’t prove a cause-and-effect relationship between chocolate and reduced risk of heart disease and stroke.”

*- JoAnn Manson, chief of Preventive Medicine at Brigham and Women’s Hospital, Boston*

## Miasmas, miasmatism, miasmatists (pre 20th century)

---

Bad smells given off by waste and rotting matter

Believed to be the main source of disease

Suggested remedies:

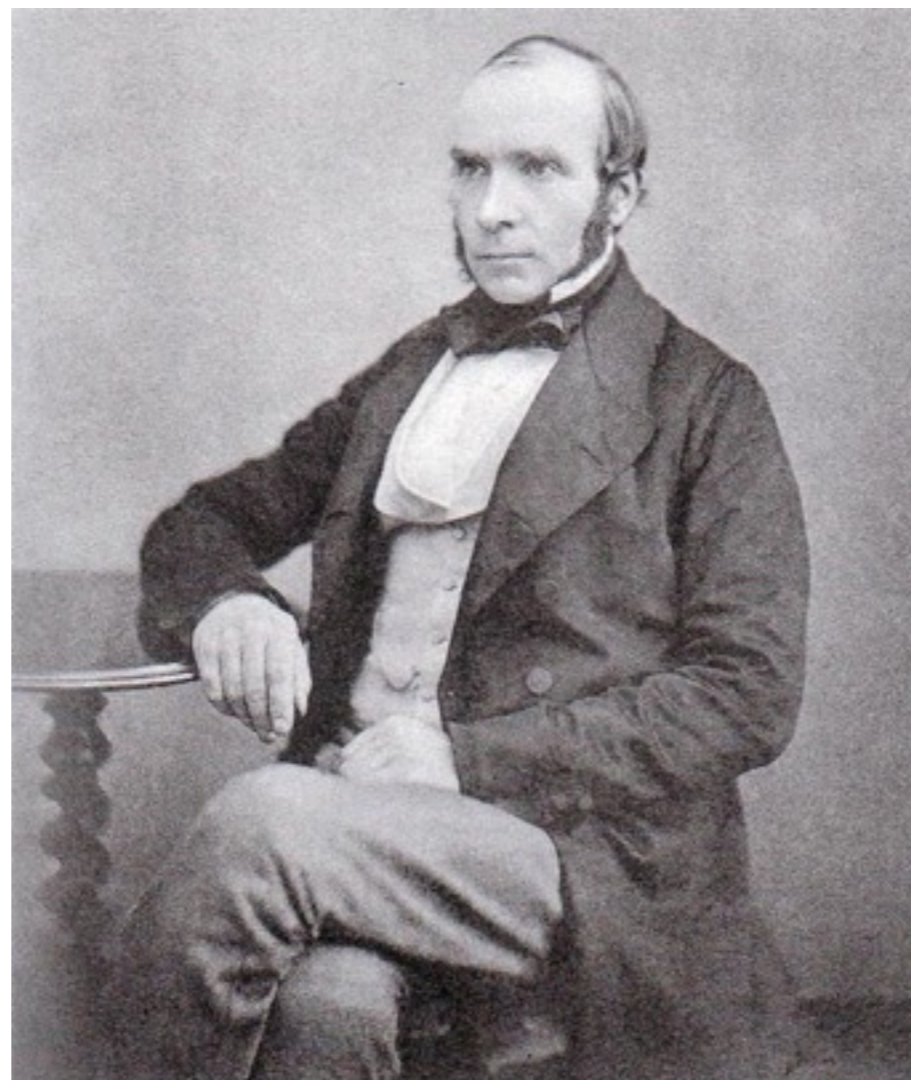
- “fly to clene air”
- “a pocket full o’posies”
- fire off barrels of gunpowder

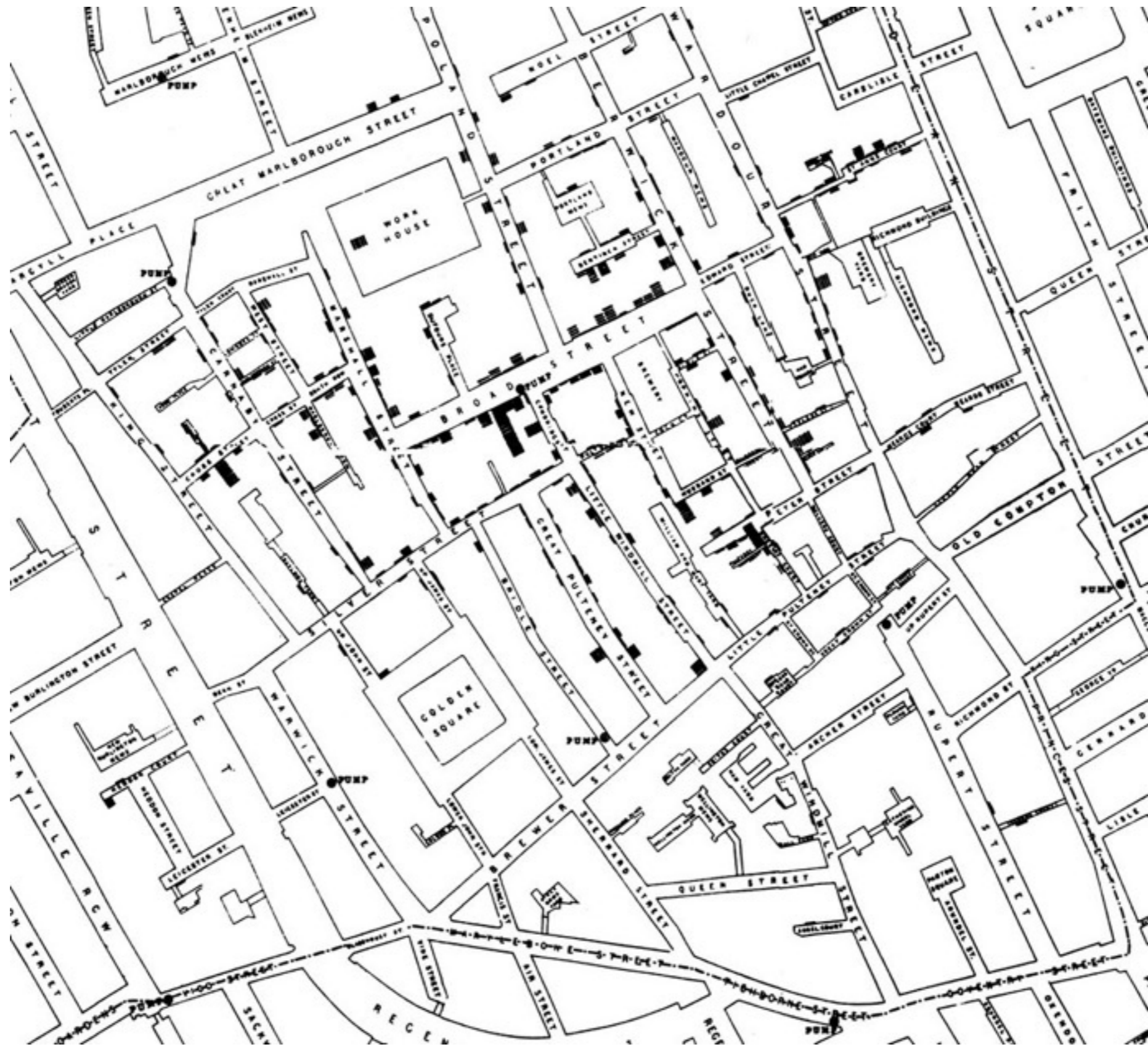
Staunch believers:

- Florence Nightingale
- Edwin Chadwick, Commissioner of the General Board of Health

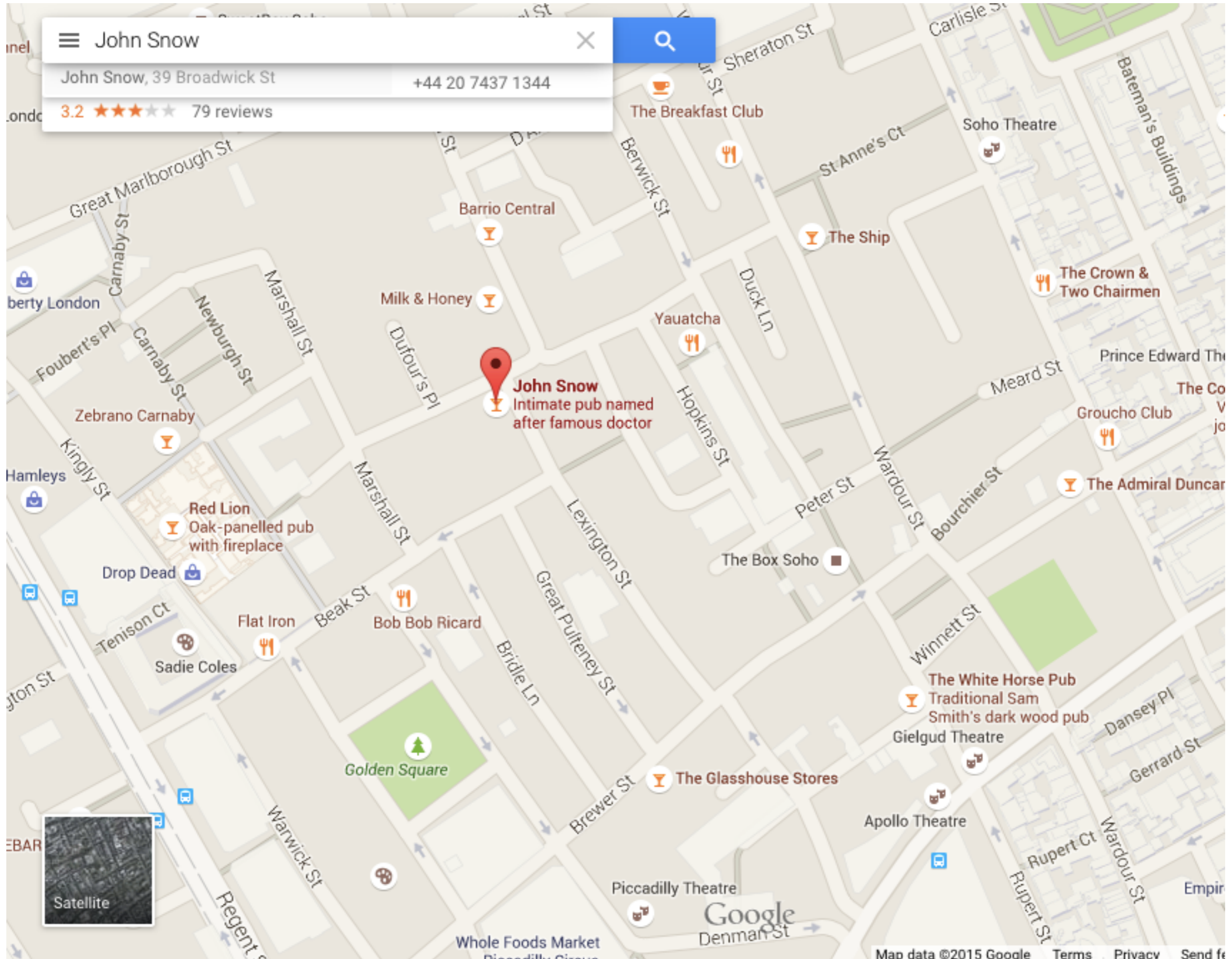
## John Snow, 1813-1858

---









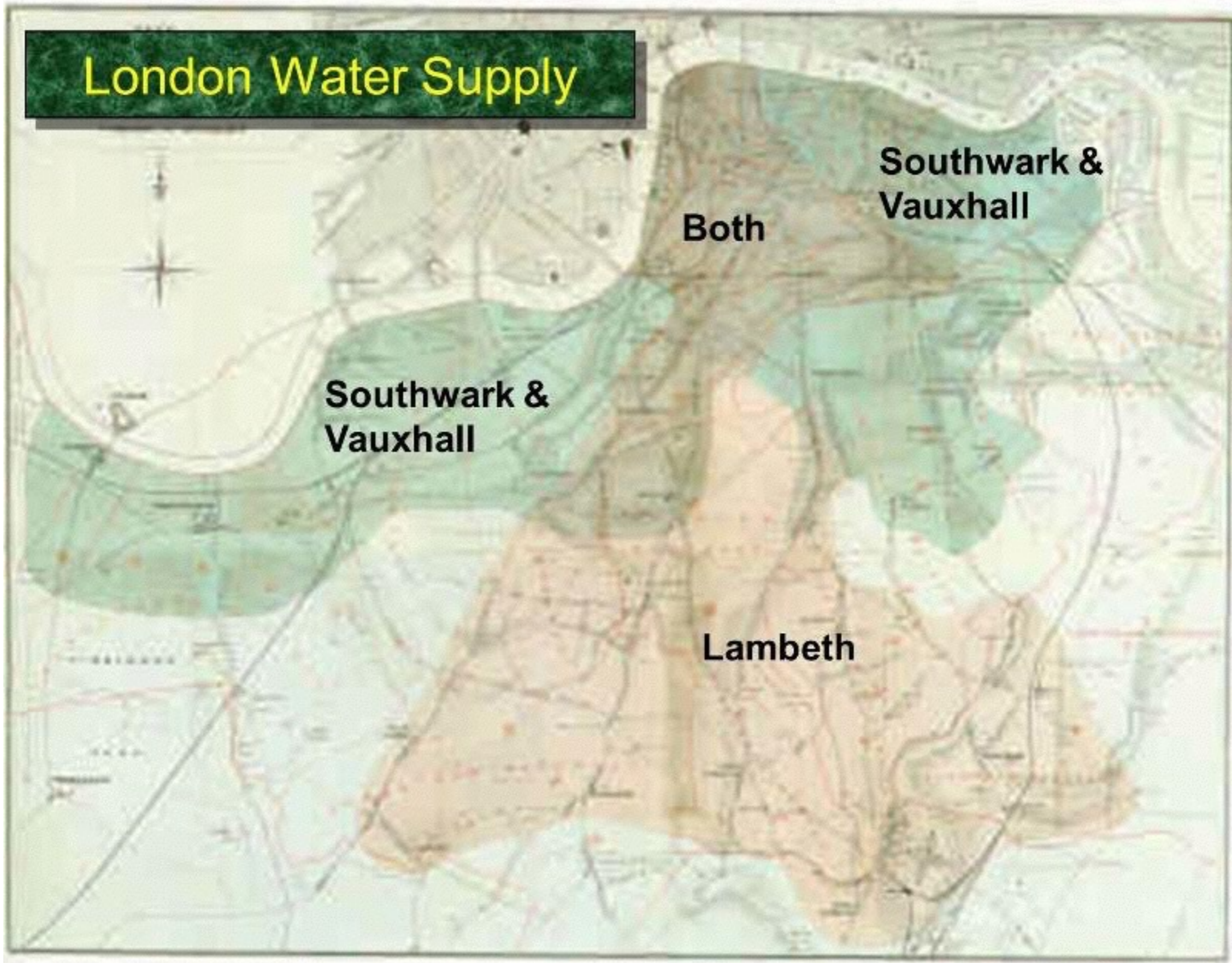


# Comparison

---

- treatment group
- control group  
does not receive the treatment

# London Water Supply



## Snow's "Grand Experiment"

---

"... there is no difference whatever in the houses or the people receiving the supply of the two Water Companies, or in any of the physical conditions with which they are surrounded ..."

The two groups were *similar except for the treatment*.

## Snow's Table

---

Supply Area	Number of houses	Cholera deaths	Deaths per 10,000 houses
S&V	40,046	1,263	315
Lambeth	26,107	98	37
Rest of London	256,423	1,422	59

If the treatment and control groups are **similar apart from the treatment**, then a difference in outcomes can be ascribed to the treatment.

If the treatment and control groups have **systematic differences other than the treatment**, then it might be difficult to identify causality.

Such differences are often present in **observational studies**.  
When they lead researchers astray, they are called **confounding factors**.

# Randomize!

- If you assign individuals to treatment and control **at random**, then the two groups will be similar apart from the treatment.
- You can account – mathematically – for variability in the assignment.

## Randomized Controlled Experiment



## Caution ...

---

Regardless of what the dictionary says,  
in probability theory

**Random  $\neq$  Haphazard**