



# Lecture 2

---

Cause and Effect

# Last Lecture

---

- **Intro to Data Science**

- Inference

- Given the data of a sample of data8 students, what conclusions can you draw about all the undergrad students at Cal?

- Prediction

- Can you make an informed guess about the final scores of different students based on the data from last semesters?

# Observational Studies

---

- Attempt to make conclusions on the basis of data
  - No rule in generation of data
  - **Observations**
    - Measurements of data
  - **Treatment**
    - Factor of interest
  - **Outcome**
- **Whether treatment had an effect on outcome or not?**
  - Association ( any relationship)
  - Causality ( Treatment actually caused the outcome)

# Regularly Eating Chocolate Is Linked to 8 Percent Lower Heart Attack Risk



By [Lisa Rapaport](#)

Reviewed: July 23, 2020

Report on a July 2020 article  
in the European Journal of  
Preventive Cardiology



While researchers didn't account for people's overall diet and exercise habits, prior research suggests the treat, particularly the dark variety, has disease-fighting compounds.

<https://journals.sagepub.com/doi/full/10.1177/2047487320936787>

<https://www.everydayhealth.com/diet-nutrition/eating-chocolate-regularly-linked-to-lower-heart-attack-risk/>

# Observation

---

- **individuals**, study subjects, participants, units
  - *336,289 US, Swedish, and Australian adults in several studies*
- **treatment**
  - *chocolate consumption*
- **outcome**
  - *heart disease*

# The first question

---

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

- **association**
  - any relation
  - link

Answer: Yes, because those who ate chocolate had less heart disease.

# A Stronger Link?

---

Other headlines about the same article:

## Chocolate is good for the heart

22 Jul 2020

European Society of Cardiology Press Release

## Is eating chocolate heart-healthy? Study says 'yes'

August 26, 2020

No Comments



Family Safety and Health, National Safety Council

<https://www.escardio.org/The-ESC/Press-Office/Press-releases/Chocolate-is-good-for-the-heart>

<https://www.safetyandhealthmagazine.com/articles/20257-is-eating-chocolate-heart-healthy-study-says-yes>

# The next question

---

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

- **causality**

This question is often harder to answer.

“Dr. Alice Lichtenstein, an American Heart Association volunteer and professor of nutrition science and policy at Tufts University, was more skeptical of the findings.”

Market Watch

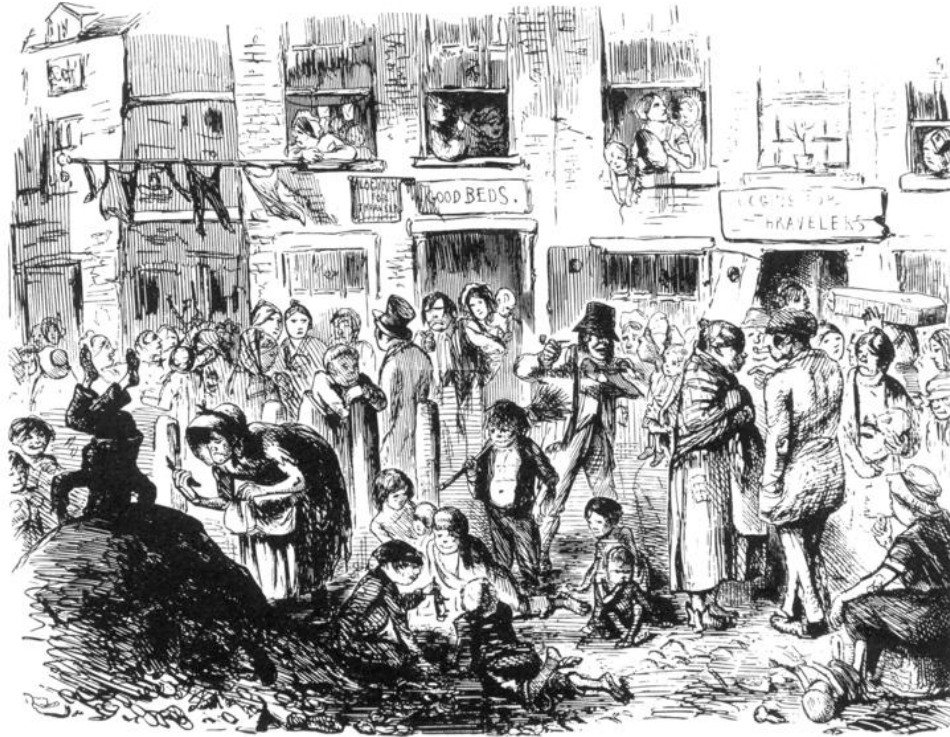
<https://www.marketwatch.com/story/eating-chocolate-once-a-week-can-lower-your-risk-of-heart-disease-study-2020-07-23>

---



# Association

# London, early 1850's



A COURT FOR KING CHOLERA.

Illustration from *Punch*  
(1852)

# Miasma

---

- **Bad smell** given off by waste and rotting matter
  - **Believed to be the main source of disease**
  
  - Believers included:
    - Florence Nightingale
    - Edwin Chadwick, Commissioner of the General Board of Health
-

# Suggested Remedies

---

## Cholera, around 1850

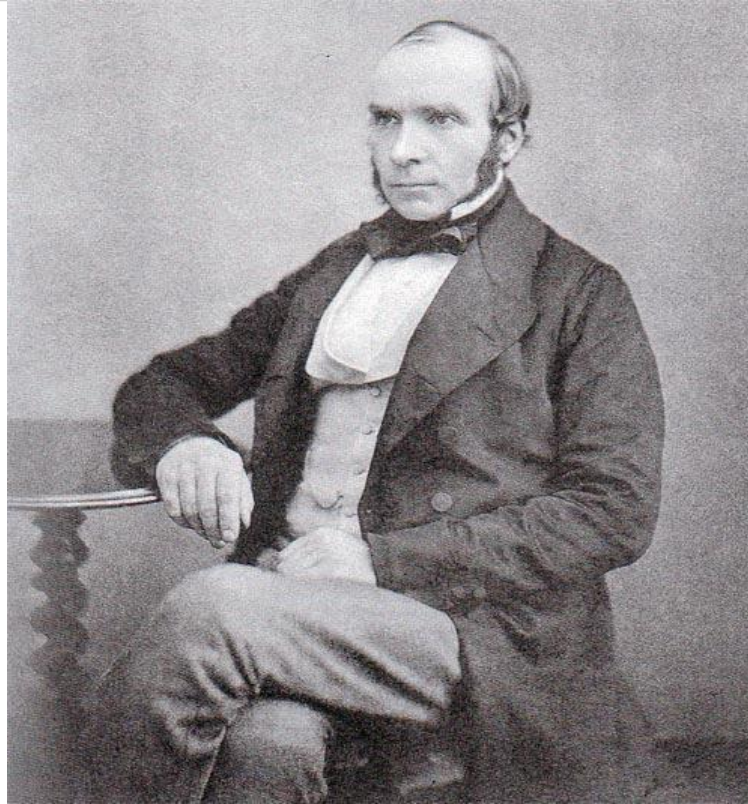
- “fly to clean air”
- “A pocket full o’posies”
- Fire off barrels of gunpowder
- Fetch clean air from the top of the Eiffel Tower in Paris

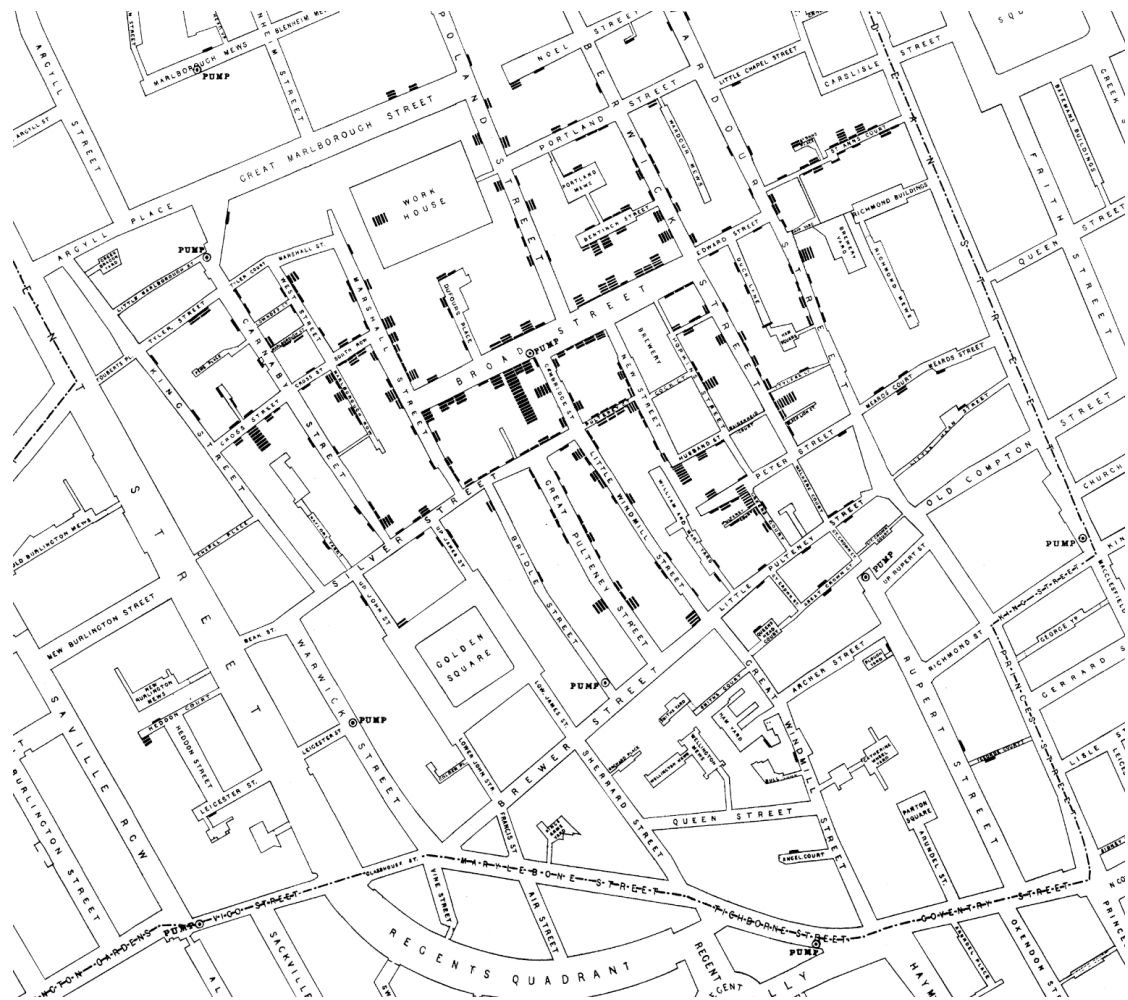
## Covid19, 2020

- Inject disinfectant
  - Sunlight
  - Hydroxychloroquine
  - Take 6 deep breaths, then cough while covering mouth
  - Cannabis, cocaine, mangoes, onion, garlic, drinking water every 15 minutes, tea, eating ice cream, avoiding ice cream
-

# John Snow, 1813-1858

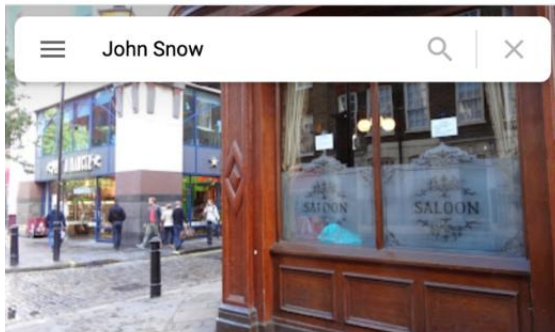
---











## John Snow

4.2 ★★★★★ 1,191 reviews · \$

Pub



Directions



Save



Nearby



Send to your  
phone



Share

Dark-wood saloon bar serving Yorkshire ales, named after doctor who traced London cholera outbreak.

✓ Dine-in · ✗ Takeout



39 Broadwick St, Carnaby, London W1F 9QJ,  
United Kingdom



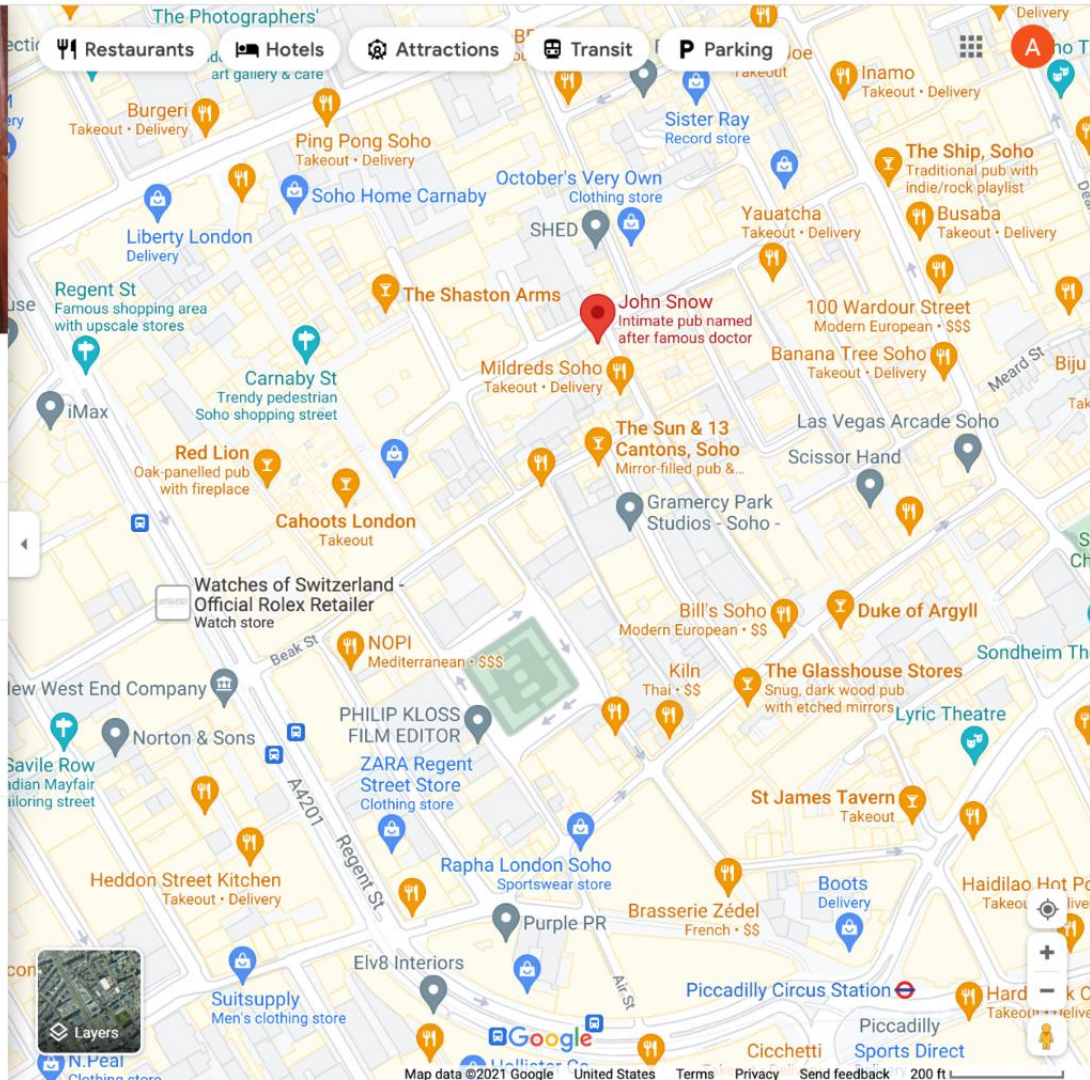
**Closed.** Opens at 12:00 PM



Health & safety: Mask required · Staff required to  
disinfect surfaces between visits



m.facebook.com

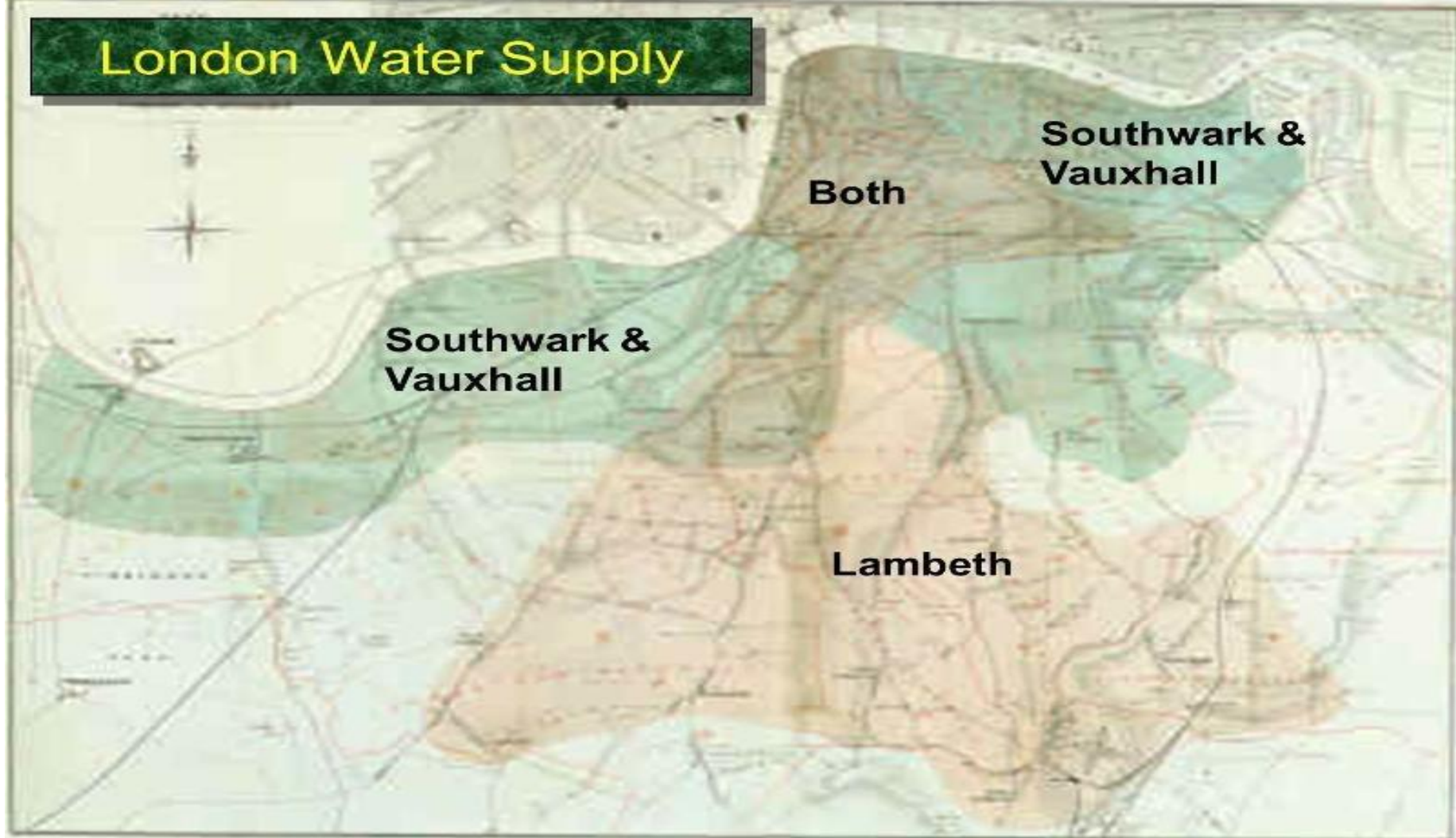






# Causation

# London Water Supply



# Comparison

---

- **treatment group**
- **control group**
  - does not receive the treatment
- No systematic difference between the two groups

# 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                       |

# Key to establishing causality

---

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

# Confounding



# Trouble

---

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**.

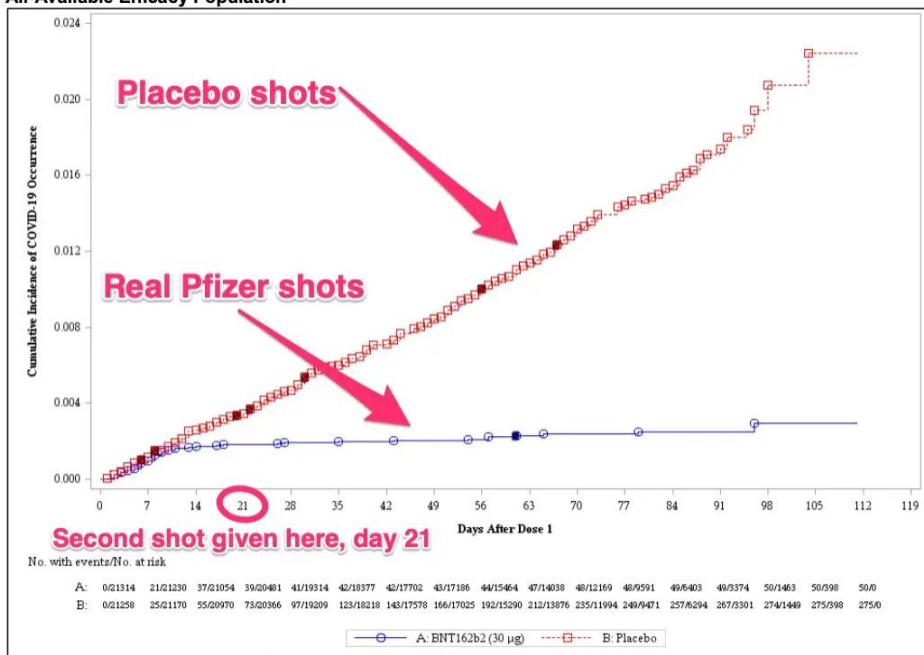
# Randomize!

---

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

# Recent Example

Figure 2. Cumulative Incidence Curves for the First COVID-19 Occurrence After Dose 1, Dose 1 All-Available Efficacy Population



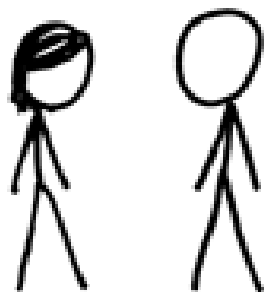
# Careful ...

---

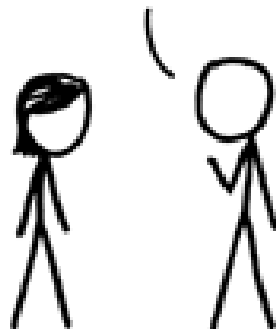
Regardless of what the dictionary says,  
in probability theory

**Random  $\neq$  Haphazard**

I USED TO THINK  
ASSOCIATION IMPLIED  
CAUSATION.



THEN I TOOK A  
STATISTICS CLASS.  
NOW I DON'T.



SOUNDS LIKE THE  
CLASS HELPED.



WELL, MAYBE.