Announcements

- Assignment: due on Fri March 1st !
- Choose one among
    - Images
    - Audio
    - Text
- Specify in which folder it should run (among the three of above)
- Make sure it runs from beginning to end

- Send your .ipynb to gquer@ucsd.edu

Controlled Experimentation

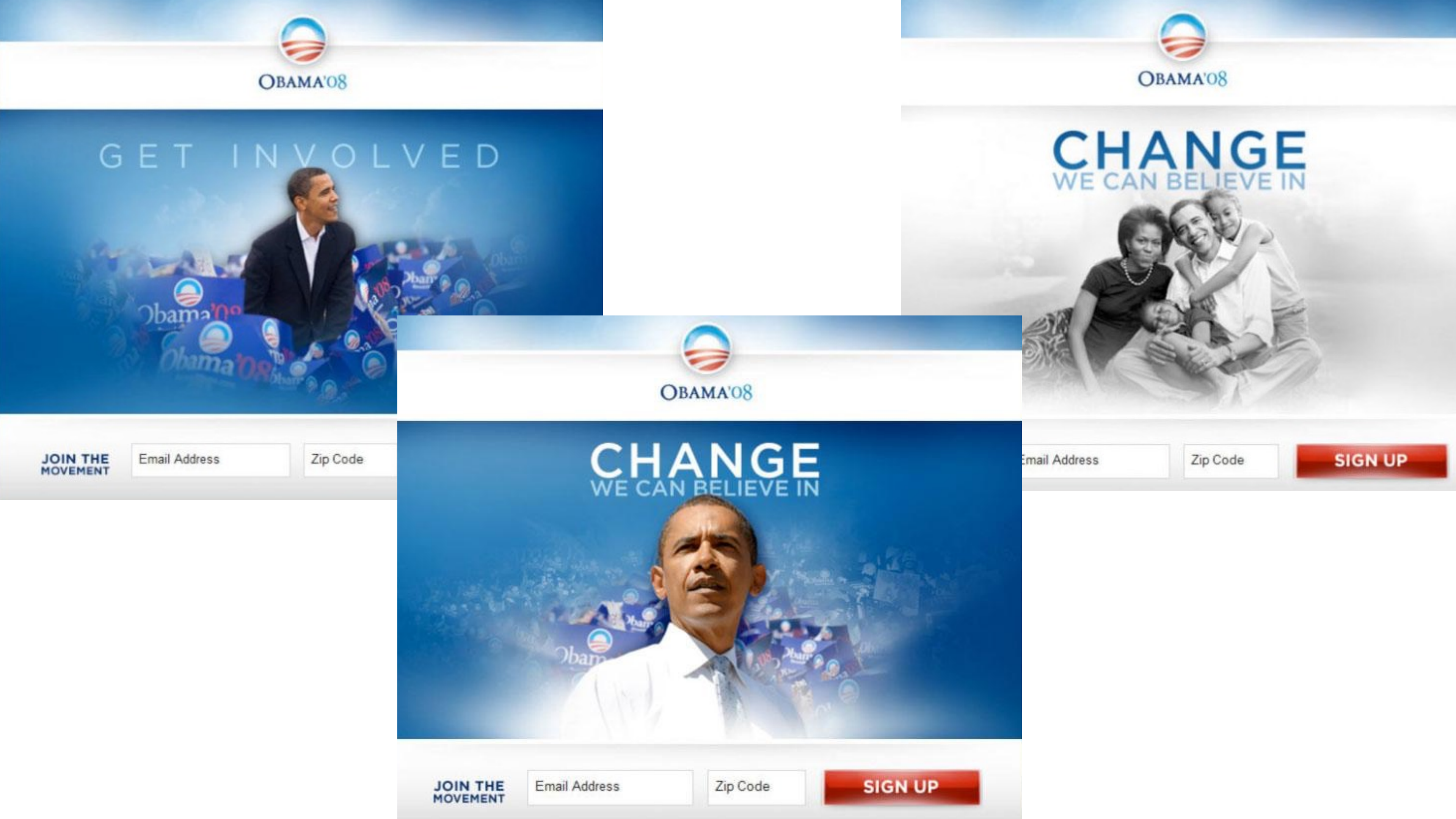
In the 1700s, a British ship's captain observed the lack of scurvy among sailors serving on the naval ships of Mediterranean countries, where citrus fruit was part of their rations.

He then gave half his crew limes (the Treatment group) while the other half (the Control group) continued with their regular diet.

Despite much grumbling among the crew in the Treatment group, the experiment was a success, showing that consuming limes prevented scurvy.

While the captain did not realize that scurvy is a consequence of vitamin C deficiency, and that limes are rich in vitamin C, the intervention worked.

British sailors eventually were compelled to consume citrus fruit regularly, a practice that gave rise to the still-popular label limeys


OBAMA'08
GET INVOLVED
JOIN THE MOVEMENT
Email Address
Zip Code
OBAMA'08
CHANGE
WE CAN BELIEVE IN
mail Address
Zip Code
SIGN UP
OBAMA'08
CHANGE
WE CAN BELIEVE IN
JOIN THE MOVEMENT
Email Address
Zip Code
SIGN UP

A/B tests

- Email A and B: Binary outcomes
  - 0, 0, 1, 1, 0, 1, 0, 0, …
  - 0, 0, 1, 0, 0, 0, 0, 1, …
- Diet A and B: weight
  - 32, 28, 27, 33, 38, 32, 31, …
  - 30, 26, 28, 34, 27, 33, 30, …
- How do we deal with these problems?

A/B test: binary outcomes

- Email 1: n_1 = 605, clicks: c_1 = 351
- Email 2: n_2 = 585, clicks: c_2 = 123
- Click per email: p_1 = 0.58 , p_2 = 0.21
- Is there enough evidence that Email 1 is better than email 2?
- Numbers are large (>100) so we can approximate with a Gaussian
- The null hypothesis is $p_1 = p_2$, we can calculate
- $p = (c_1 + c_2) / (n_1 + n_2)$: the mean click rate in the null hypothesis
- $\sigma^2 = p(1-p)$: the variance of the outcome
-
- If t>1.96, they are actually different (with 95% confidence)

$$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

A/B test: real values

- Diet 1: n_1 = 220, average: μ_1 = 32
- Diet 2: n_2 = 189, average: μ_2 = 30
- Is there enough evidence that Diet 1 is more energetic than Diet 2?
- Numbers are large (>100) so we can approximate with a Gaussian
- The null hypothesis is $\mu_1 = \mu_2$, we can calculate
- σ^2 the variance of the outcome metric (more complicate to derive, but can be calculated from data)
-
- If t>1.96, they are actually different (with 95% confidence)

$$t = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$$

Minimum sample size

$$n = 16\sigma^2/\Delta^2$$

σ^2 is the variance of the outcome metric (in case of binary outcome, p(1-p))

Δ is the sensitivity (amount you want to detect) at 80% power

n is the sample size ($n = n_1+n_2$)

$$\Delta = \hat{\mu}_1 - \hat{\mu}_2$$

$$\Delta = \hat{p}_1 - \hat{p}_2$$

Data ***Science***

“A man conducting a gee-whiz science show with fifty thousand dollars’ worth of Frankenstein equipment is not doing anything scientific if he knows beforehand what the results of his efforts are going to be. A motorcycle mechanic, on the other hand, who honks the horn to see if the battery works is informally conducting a true scientific experiment. He is testing a hypothesis by putting the question to nature.”

- Zen and the Art of Motorcycle Maintenance