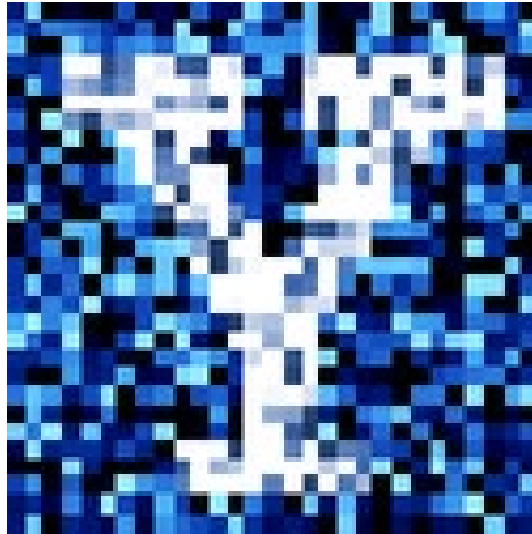# YData: Introduction to Data Science
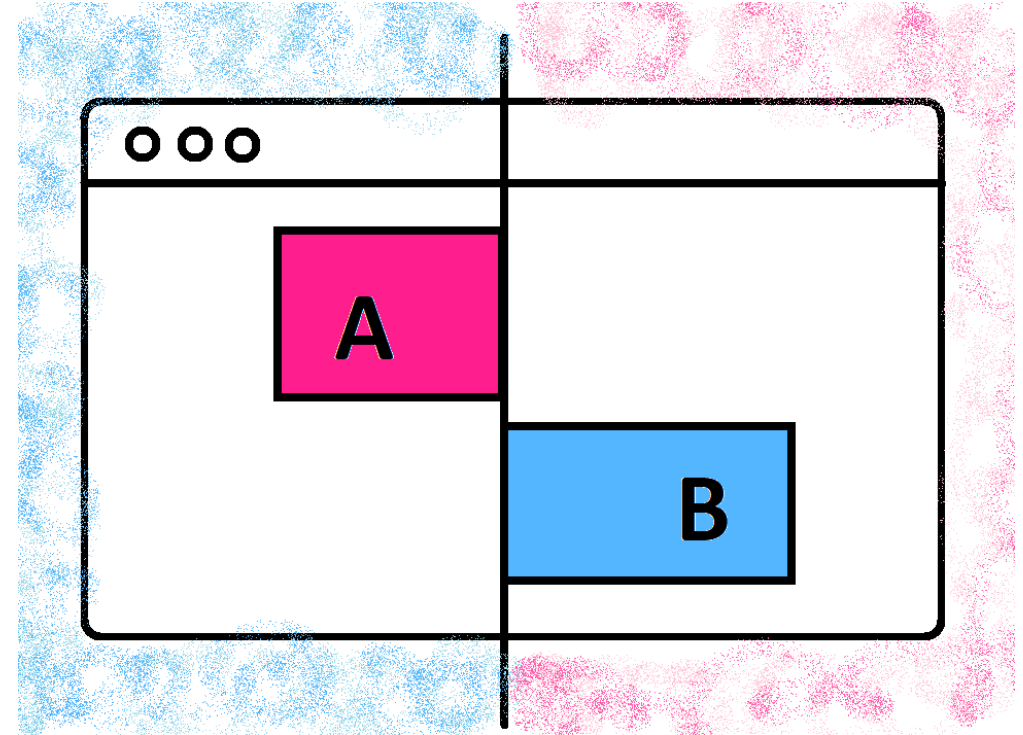

Lecture 20: Causality

Overview

Hypothesis test continued

- "A/B" testing continued
  - Deflategate example

- Causality

Announcements

Homework 6 is due on Sunday at 11pm

Exam information
- 50 minutes
- On paper
- In class on the 18th
- We will provide you with the function sheet that is on Canvas
  - No other "cheat sheet" allowed
- Just bring some pencils/pens
  - And press hard so we can read your writing when it is scanned

Review and continuation of A/B Testing

Comparing two samples

Question: Do the two sets of values come from the same underlying distribution?

Answering this question by performing a statistical test is called:

- By Data Scientists: "A/B testing"
- By Statisticians: "Two independent samples hypothesis test"

"Boy, those French: They have a different word for everything!"

Boy, those Data Scientists: They renamed everything

Side note: Data Scientists renaming things

Statistics	Machine Learning
Data Point	Instance
Covariate	Feature
Parameters	Weights
Estimation / Fitting	Learning
Regression / Classification	Supervised Learning
Clustering / Density Estimation	Unsupervised Learning
Response	Label
Test set performance	Generalization

Statistical Modeling: The Two Cultures

Leo Breiman
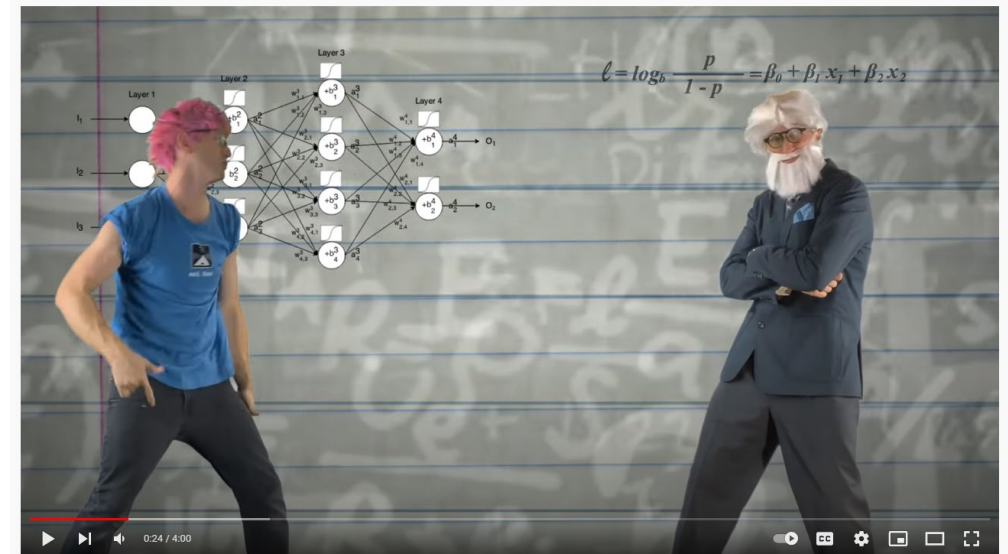
Data Scientist vs. Statistician

A/B testing of web sites

21% CONTROL

34% VARIATION
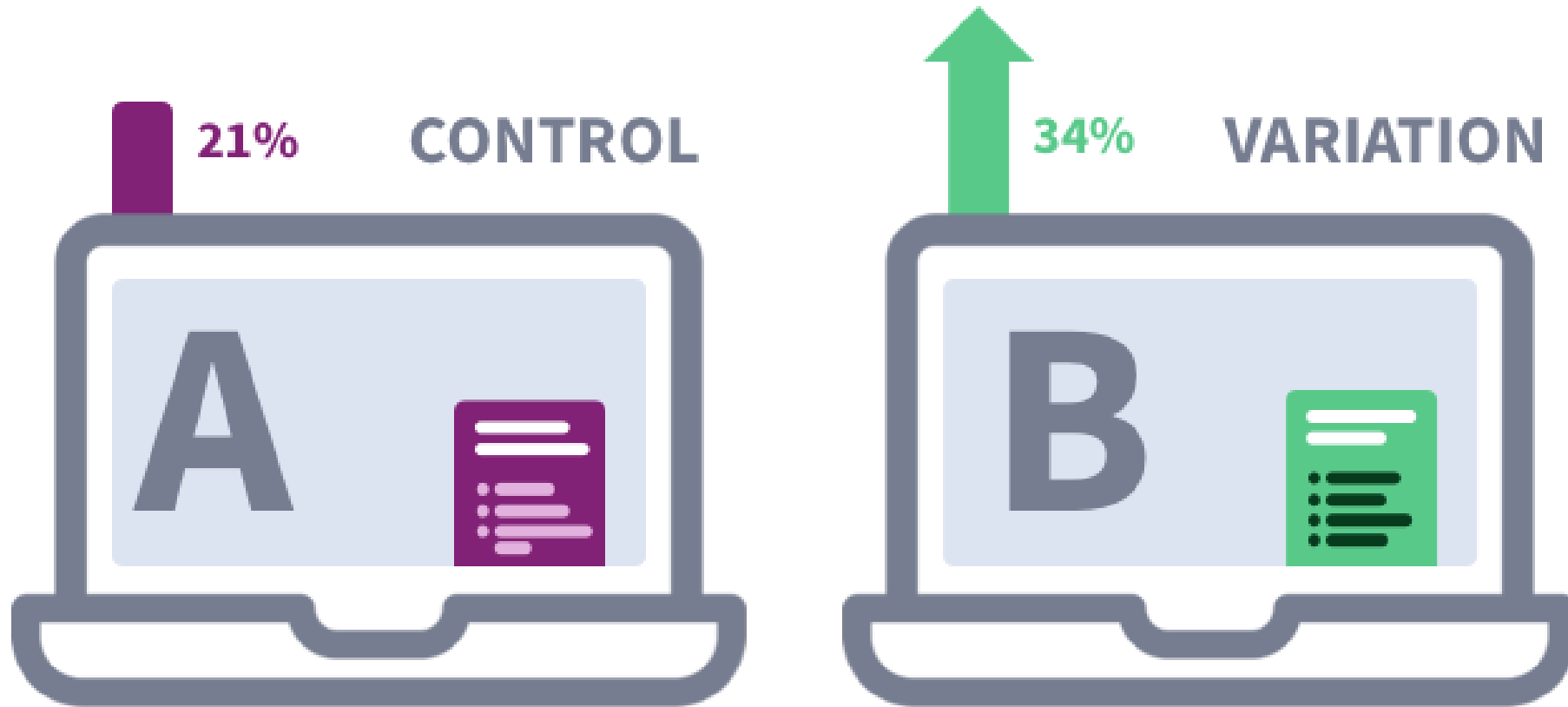
Last class example: baby birth weights

Compare:

- **Group A:** Birth weights of babies of mothers who smoked during pregnancy

- **Group B:** Birth weights of babies of mothers who didn't smoke

Hypotheses:

- **Null hypothesis**: In the population, the distributions of the birth weights of the babies in the two groups are the same

- **Alternative hypothesis**: In the population, the babies of the mothers who didn't smoke were heavier, on average, than the babies of the smokers.

Test statistic: Average Group B - Average Group A

Q: Could we have used these statistics instead?

- Median Group B   - Median Group A

- | Average Group B - Average Group A |

Simulating under the null hypothesis

If the null is true, all rearrangements of the birth weights among the two groups are equally likely
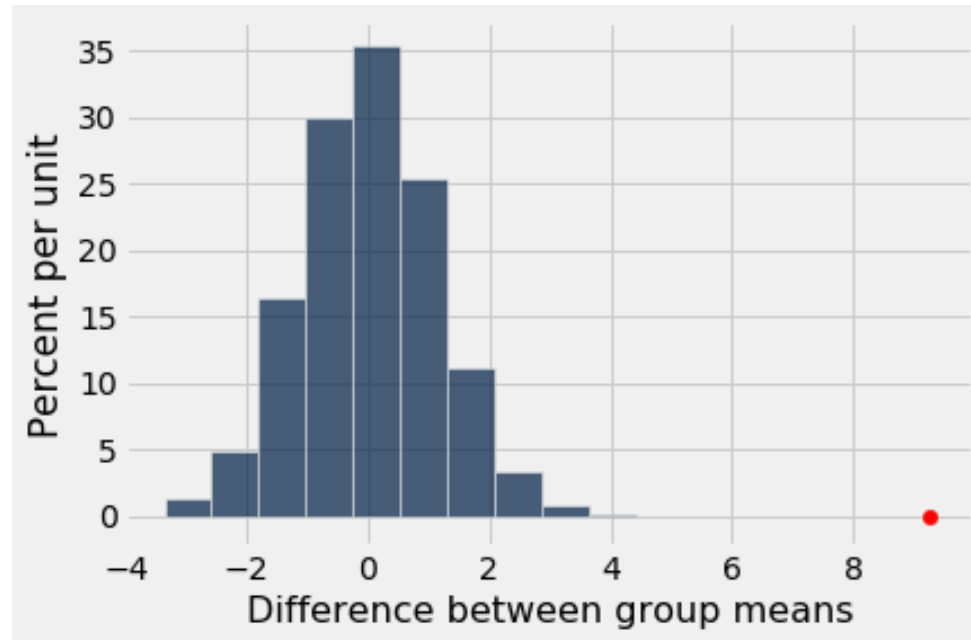
Plan:

- Shuffle all the birth weights
- Assign some to "Group A" and the rest to "Group B", maintaining the two sample sizes
- Find the difference between the averages of the two shuffled groups
- Repeat

Create the null distribution!
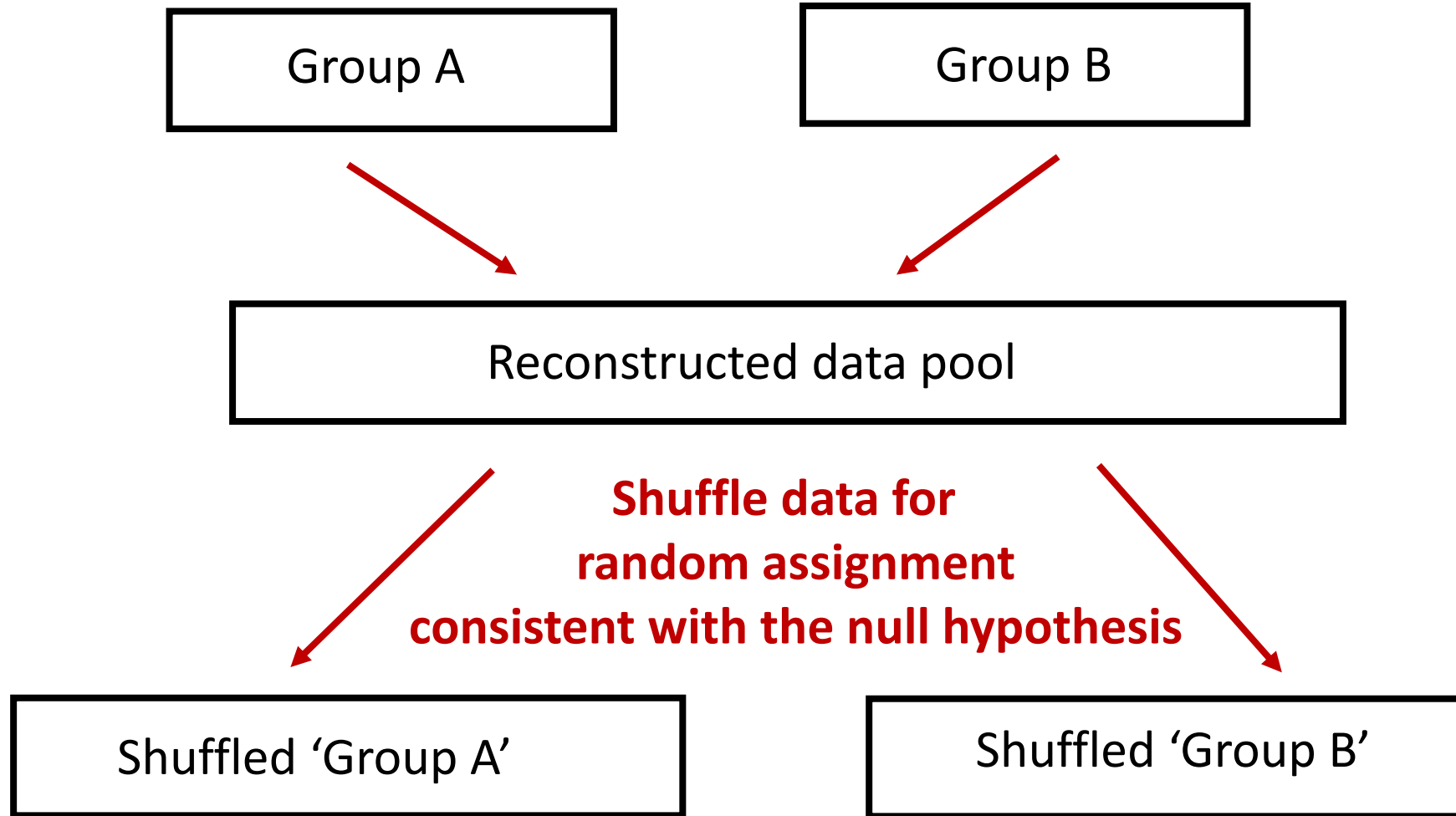
One null distribution statistic: Average B – Average A

Null distribution and p-value

What is the p-value here?
What do we conclude?

Deflategate

2015 AFC Championship Game

Deflategate

'Deflategate' returns, focus on Tom Brady's destroyed cellphone

POSTED 9:54 AM, MARCH 5, 2016, BY CNN WIRE, UPDATED AT 10:33AM, MARCH 5, 2016

Wikipedia:

- The 2015 AFC Championship Game football tampering scandal, commonly referred to as Deategate, or Ballghazi

Claim: New England Patriots deflated footballs

Data:

- Pressure (PSI) drop was measured in 11 Patriots and 4 Colts footballs at halftime

Null hypothesis

What are the null and alternative hypotheses?

- Null: PSI drop is the same for both teams
- Alternative: PSI drop is greater for Patriots

To test this hypothesis, repeat this process:

- Randomly permute all 15 balls
- Label 11 of them "Patriots" and the remaining 4 "Colts"
- Compare the averages of the two groups

Let's explore this in Jupyter!

Causality

Causality

Recall from class 2:

- **An association** is the presence of a reliable relationship between the treatments an outcome

- **A causal relationship** is when changing the value of a treatment variable influences the value outcome variable

Is there an association and/or causal relationship for:

- The example of smoking mothers and baby weights?

What are some confounding variables?

Randomized Controlled Experiment

Sample A: control group

Sample B: treatment group

The treatment and control groups are selected at random; this allows causal conclusions!

Any difference in outcomes between the two groups could be due to:

- Chance
- The treatment

Randomized Controlled Experiment

Take a group of participant and ***randomly assign***:

- Half to a *treatment group* where they get chocolate

- Half in a *control group* where they get a fake chocolate (placebo)

- See if there is more improvement in the treatment group compared to the control group
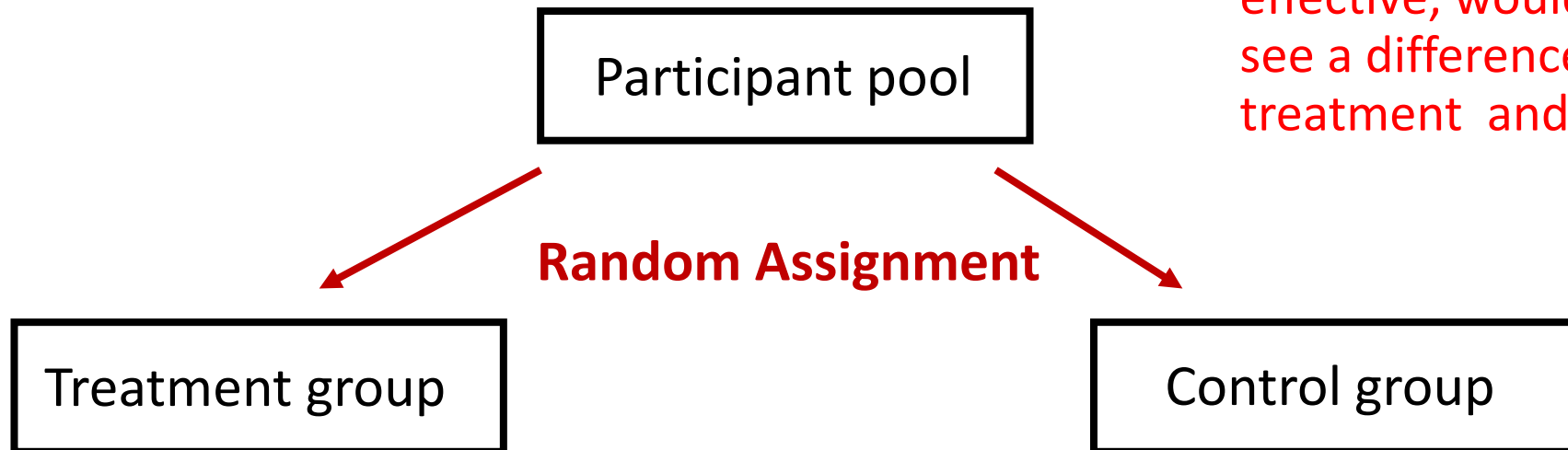
Q: If the treatment was not effective, would we expect to see a difference between the treatment and control groups?

Participant pool

Random Assignment

Treatment group

Control group

Case study

RCT to study Botulinum Toxin A (BTA) as a treatment to relieve chronic back pain
- 15 patients in the treatment group (received BTA)
- 16 in the control group (normal saline)

Trials were run double-blind: neither doctors nor patients knew which group they were in.

Results
- 2 patients in the control group had relief from pain (outcome=1)
- 9 patients in the treatment group had relief.

Can this difference be just due to chance?

The hypotheses

Null:

- BTA does not lead to an increase in pain relief
  - i.e., if many people were to get BTA and saline, the proportion of people who experienced pain relief would be the same in both groups.

Alternative:

- BTA leads to an increase in pain relief
  - i.e., if many people were to get BTA and saline, the proportion of people who experienced pain relief would be higher for those who received BTA

Let's explore this in Jupyter!