Week 3, Class 6

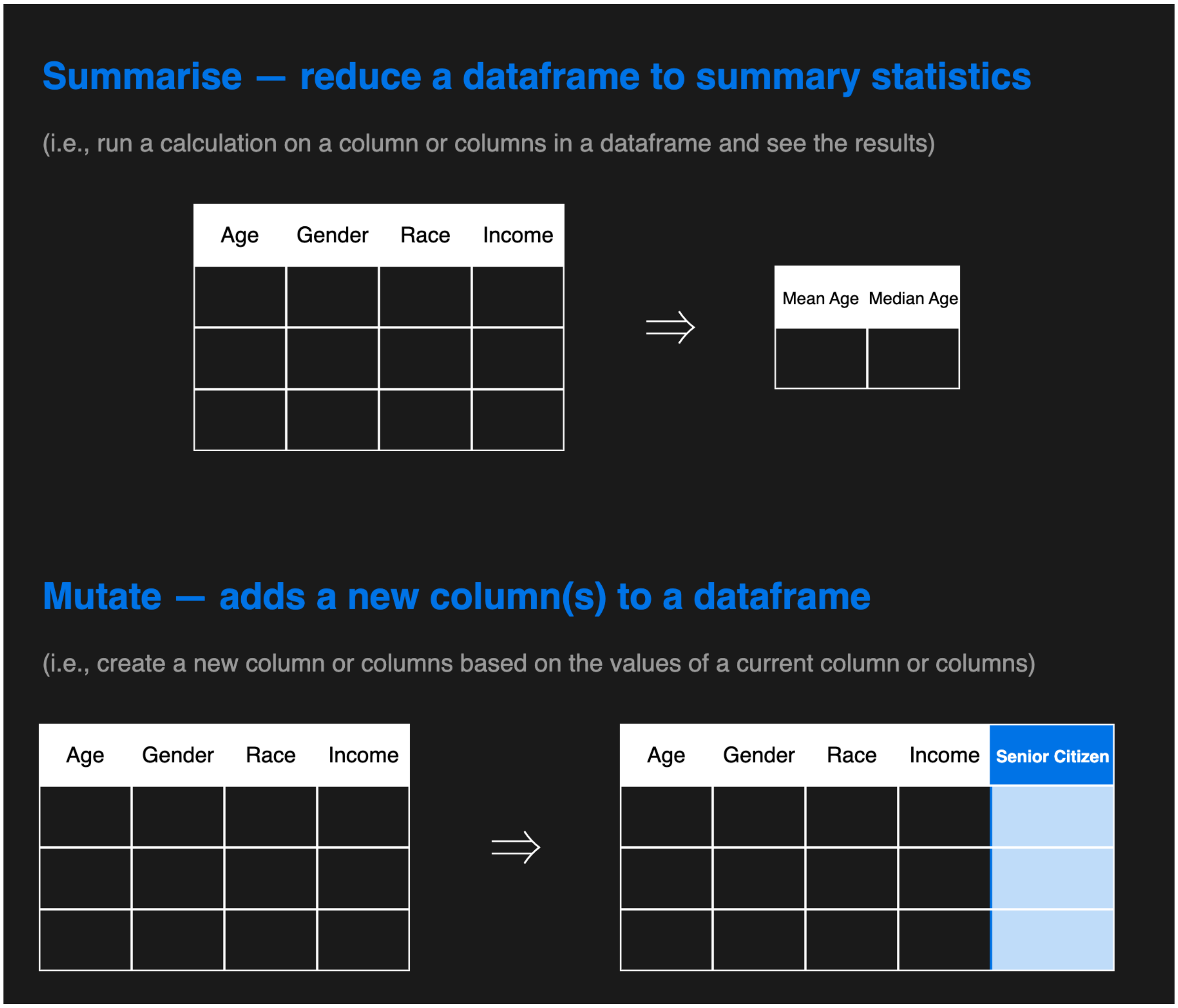
Research Designs

Sean Westwood

Today

- Understand experimental and observational research designs
- Understand natural experiments and their advantages
- Recognize the difference between cross-sectional and longitudinal data
- Evaluate the tradeoffs between different research designs

Recap: Summarise v. Mutate



What Determines A Research Design?

Research Questions

Consider this question: "Does Trump's anti-vaccine rhetoric reduce vaccination rates among his supporters?"

Three possible approaches:

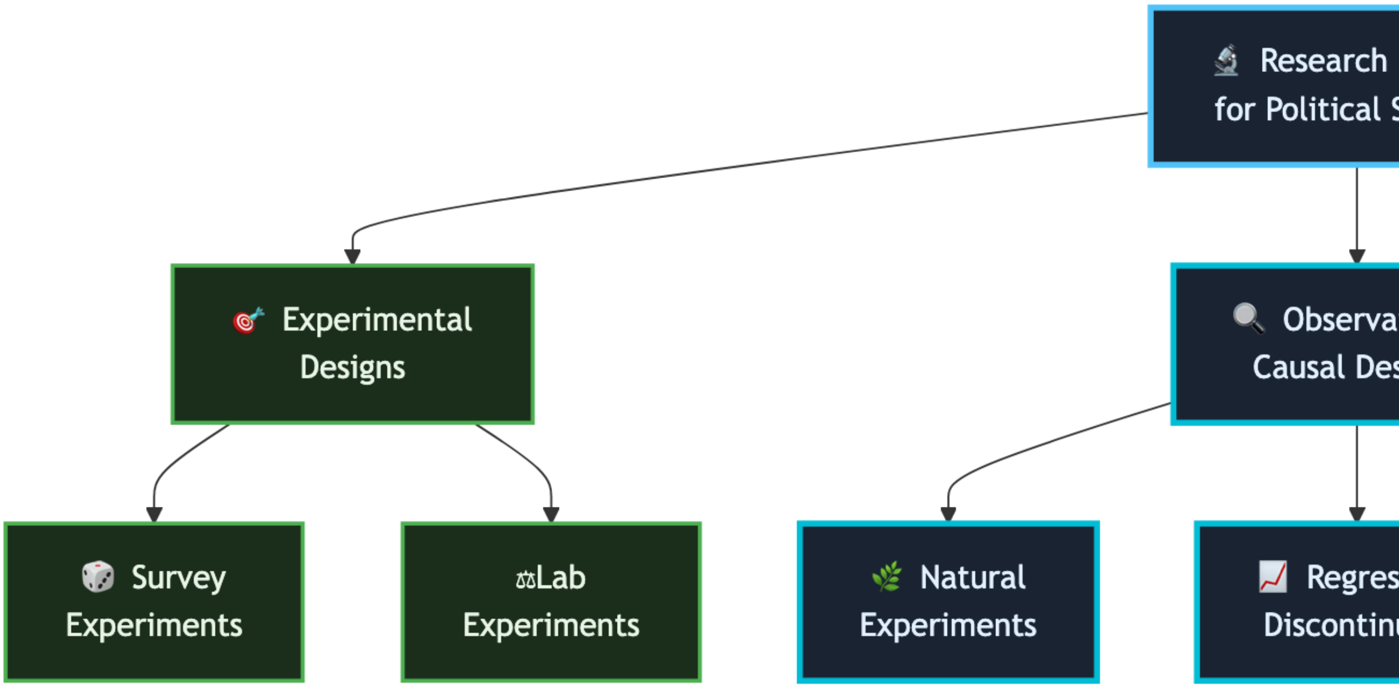
- **Approach 1**: Survey people about Trump and their vaccination status
- **Approach 2**: Compare vaccination rates in Trump vs. Biden counties

- **Approach 3**: Track the same people's vaccination status before and after seeing or not seeing Trump statements on vaccines

The design determines whether we can answer the question!

The Three Main Types of Research Designs

Research Design Taxonomy



Experimental Designs: The Gold Standard

What Makes an Experiment?

Two key features:

1. **Random Assignment**: Researchers control who gets what treatment
2. **Control Groups**: Direct comparison between treated and untreated groups

Example: Testing whether voter reminders increase turnout

- **Random Assignment**: Flip a coin to decide who gets reminder calls
- **Control Group**: People who don't receive calls
- **Treatment Group**: People who receive calls
- **Comparison**: Turnout rates between the two groups

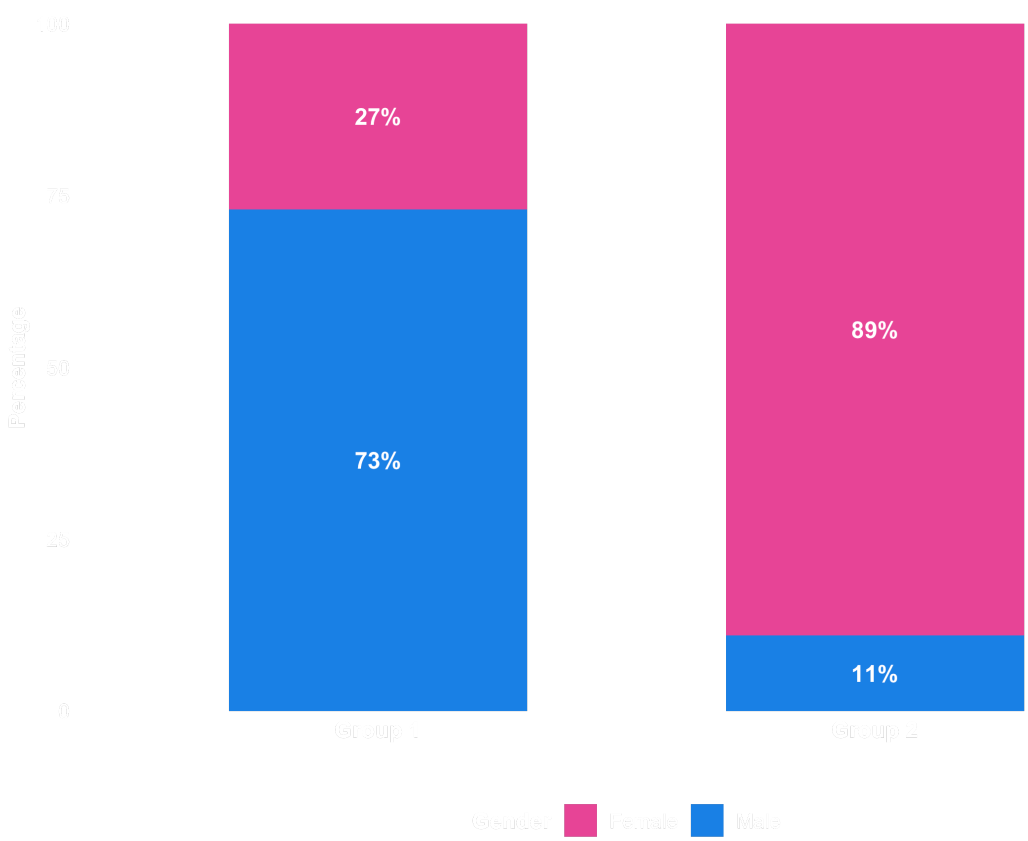
Why Random Assignment Matters

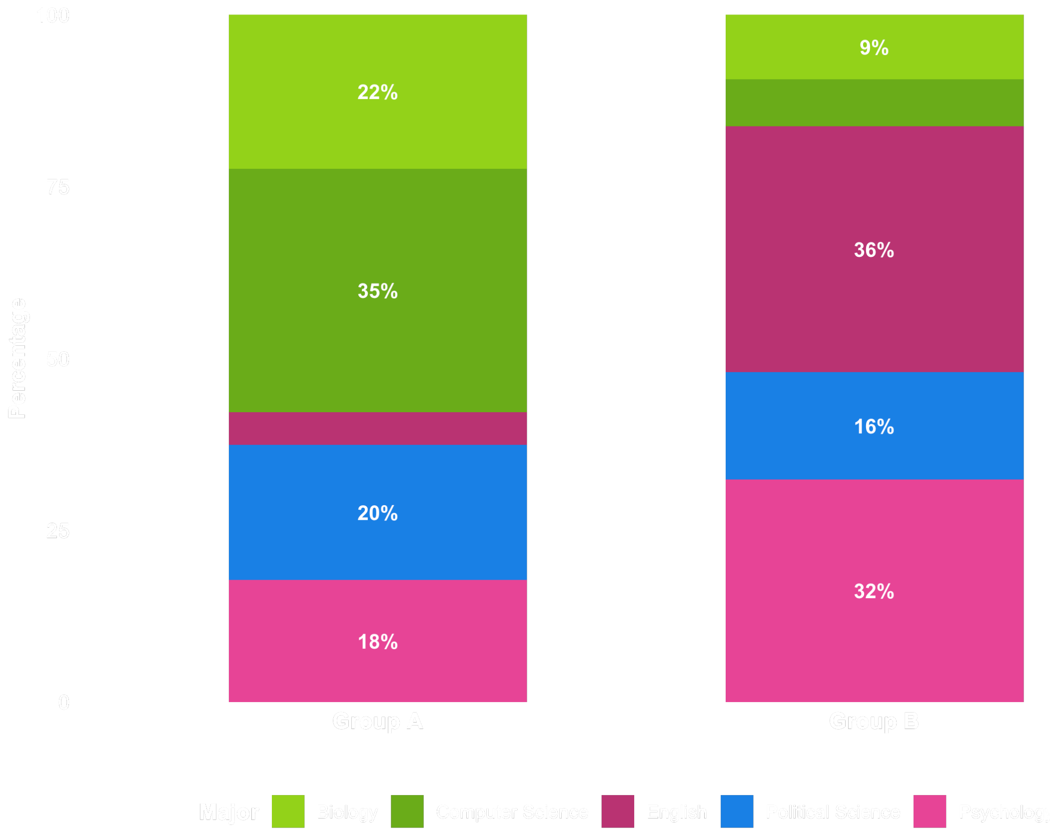
The Problem: People are different in many ways

- Some people always vote, others never do
- Some are more politically engaged
- Some have more flexible schedules

Why Random Assignment

Without random assignment, groups can be very different:





What is Experimental Control?

- **Control group**: A comparison group that doesn't receive the treatment

Without control, we cannot make causal claims

The comparison problem:

- **Without control**: We can't tell if changes are due to our treatment or something else
- **With control**: We can isolate the effect of our specific treatment (everything else should be the same)
- **Example**: If we only study people who received voter mobilization calls, we can't know if their high turnout was due to the calls or because they were already motivated voters (on the voter rolls)

What Are We Controlling For?

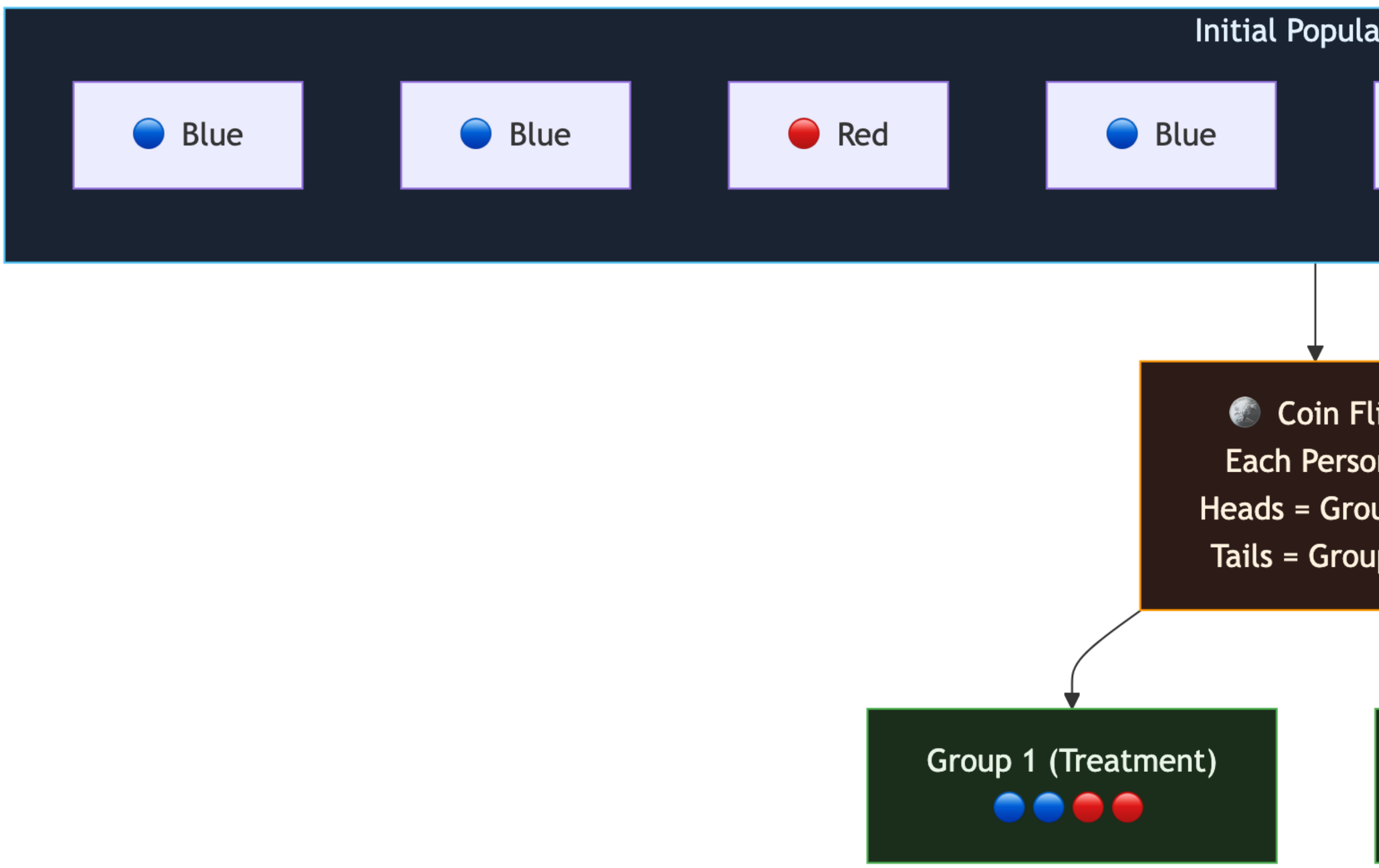
The threats to valid causal inference

Four major threats we control for:

- **Confounding variables**: Other factors that might affect the outcome
- **Selection bias**: Differences between people who do/don't receive treatment
- **Time trends**: Changes that happen regardless of treatment
- **Measurement bias**: Differences in how we observe outcomes

How Random Assignment Works

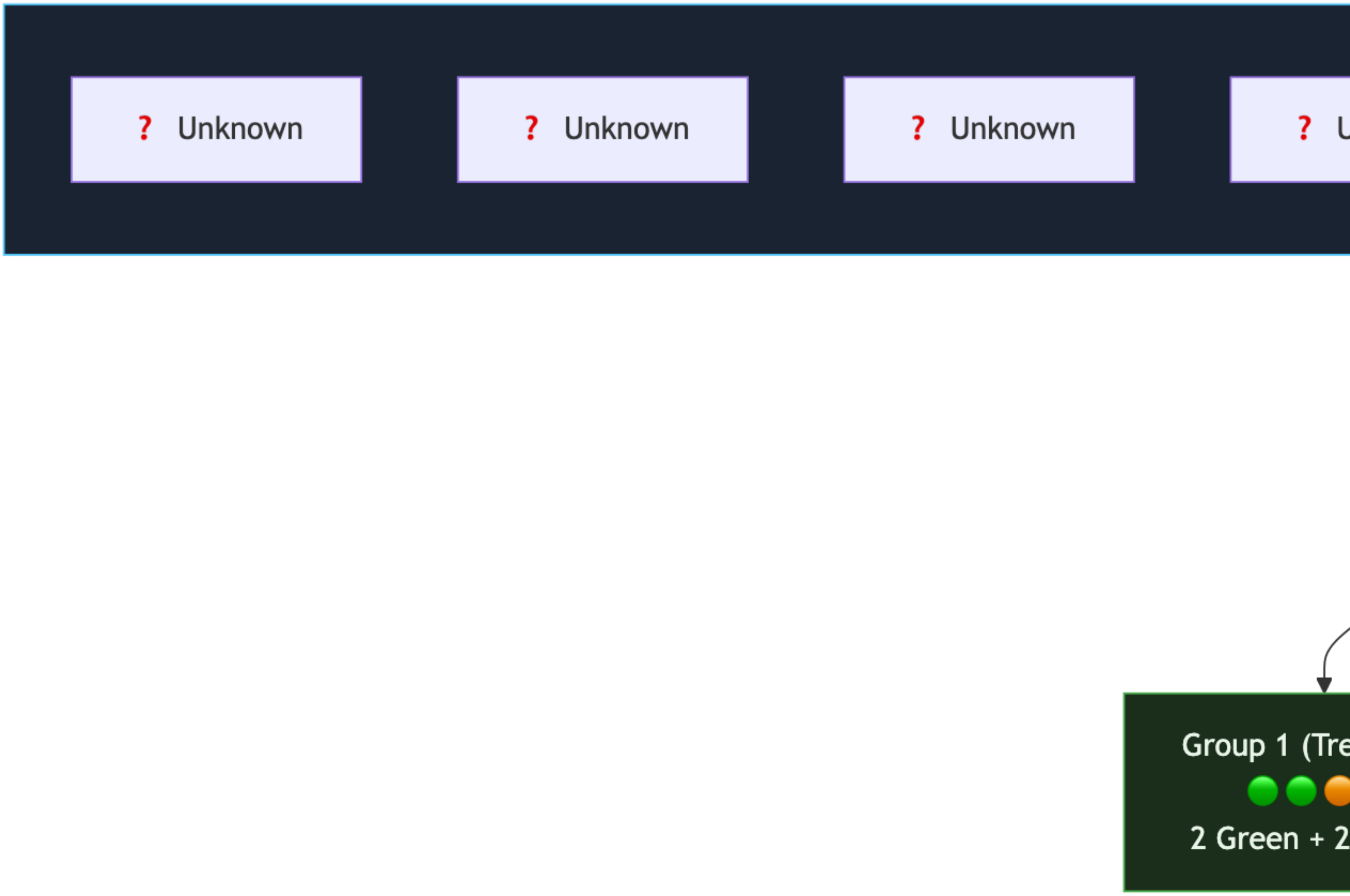
Random assignment: Each subject has an equal chance of being assigned to any group



Key insight: Random assignment makes treatment and control groups statistically identical on average

Random Assignment: Unobserved Characteristics

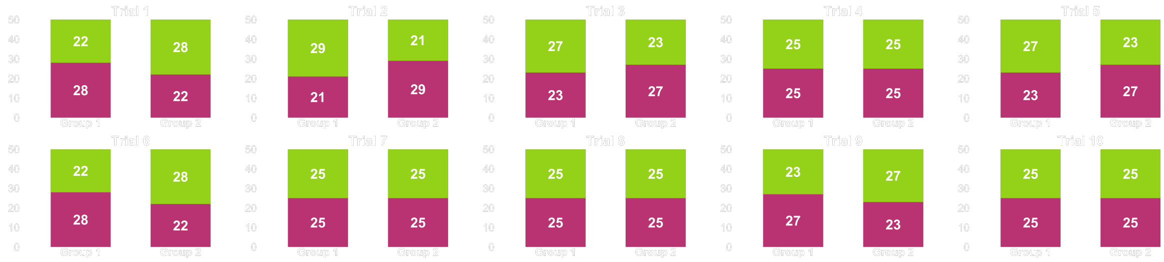
In practice, we often don't know people's characteristics beforehand



Key insight: Random assignment works even when we don't know characteristics in advance - it automatically balances groups across all traits (known and unknown)!

Random Assignment: Individual Trial Variation

Random assignment doesn't *guarantee* perfectly balanced characteristics in any single trial (100 people, 50 purple/50 pink, randomly assigned to groups of 50 each)



Even with equal group sizes (50/50), the purple/pink balance varies by chance across individual trials. Each experiment produces different compositions! But this is okay! On average we will be close to 50/50/

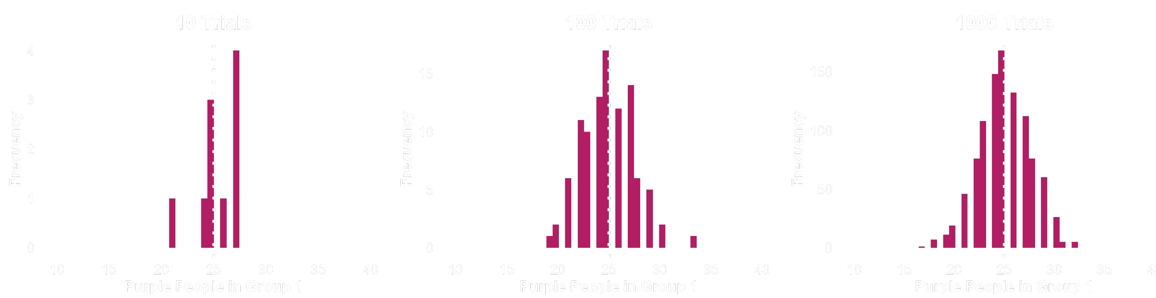
The Law of Large Numbers

What happens when we run many random assignments? (100 people, 50 purple/50 pink, groups of 50 each)

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_bar()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_bar()`).

Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_bar()`).
```



As we increase the number of trials, the distribution converges to the expected value (25 purple per group, white dotted line). Random assignment creates characteristic balance *on average* - the more trials, the closer we get to true balance!

Some Political Science Experiments

Political science has embraced experimental methods in recent decades

Why experiments matter in political science:

- **Causal claims**: Experiments allow us to make definitive statements about cause and effect
- **Policy relevance**: Results directly inform campaign strategies, government programs, and civic interventions
- **Theory testing**: Experiments can definitively test competing theoretical predictions
- **Real-world impact**: Experimental findings shape how campaigns run, how governments communicate, and how civic organizations operate

Experiment 1: Get Out The Vote

Gerber & Green (2000) - Do campaign methods actually work?

Research Question: Does personal canvassing, direct mail, and phone calls increase voter turnout?

Voters randomly assigned to get:

1. Personal canvassing
2. Direct mail

3. Phone calls
4. No contact (control)

Results: Personal canvassing increased turnout by 7-10 percentage points; phone calls and mail had minimal effects

Experiment 2: Media and Reconciliation in Rwanda

Paluck (2009) - Can media promote peace after conflict?

Research Question: Can radio programming reduce intergroup prejudice and promote reconciliation?

Communities randomly assigned to:

1. Educational radio soap opera promoting reconciliation
2. Control radio program about health

Results: Educational radio soap opera significantly changed social norms and behaviors toward reconciliation

Experiment 3: Deep Canvassing and Transgender Rights

Broockman & Kalla (2016) - Can prejudice be changed through conversation?

Research Question: Can brief conversations change deeply held prejudices about transgender people?

Voters randomly assigned to:

1. 10-minute conversation about transgender rights and experiences
2. Control conversation about recycling

Results: 10-minute conversations reduced transgender prejudice for at least 3 months

Observational Designs: When Experiments Aren't Possible

When You Can't Randomly Assign

Some research questions involve things you can't control:

- Constitutional changes (can't randomly assign constitutions to countries)
- Major life events (can't randomly assign parents to divorce)
- Policy implementations (can't randomly assign laws to states)

Solution: Find situations where assignment was "as good as random"

"As good as random" means:

- Assignment to treatment/control wasn't controlled by researchers
- BUT the assignment process was effectively random
- No systematic bias determines who gets treated
- Creates similar groups, just like an experiment would

What Does "As Good As Random" Mean?

When nature or institutions create randomization for us

Examples of "as good as random" assignment:

- **Draft lottery**: Birthdays randomly determine military service
- **Policy rollout**: Budget constraints mean only some areas get new program first
- **Natural disasters**: Random timing affects some regions but not others

- **Administrative cutoffs**: Arbitrary thresholds (age 65 for Medicare) create treatment groups

Key insight: We're looking for situations where treatment assignment is **unrelated to potential outcomes**. If assignment is random (or effectively random), we can make causal claims even without controlling the assignment ourselves.

Natural Experiments

Three examples of "as good as random" assignment in the real world

Natural Experiment 1: The Draft Lottery

Angrist (1990) - Does military service affect lifetime earnings?

Research Question: Does military service hurt or help lifetime earnings and employment?

"As Good As Random" Assignment: Vietnam-era draft lottery used birthdates to determine military service

1. Men born on certain dates were drafted
2. Men born on other dates were not drafted
3. Birthdate is effectively random with respect to earning potential

Results: Military service reduced lifetime earnings by about 15% due to lost work experience

Natural Experiment 2: Rainfall and Economic Voting

Achen & Bartels (2016) - Do voters blame politicians for things beyond their control?

Research Question: Do economic conditions that politicians can't control affect election outcomes?

"As Good As Random" Assignment: Rainfall variation creates random economic conditions

1. Droughts randomly hurt local economies in farming areas
2. Floods randomly damage local economies
3. Weather is unrelated to politician quality or policies

Results: Voters punish incumbent politicians for poor economic conditions caused by weather, not policy

Natural Experiment 3: Female Political Representation

Chattopadhyay & Duflo (2004) - Do female leaders prioritize different policies?

Research Question: When women hold political office, do they invest more in policies that benefit women?

"As Good As Random" Assignment: India's constitutional requirement for random rotation of reserved seats

1. Village councils randomly selected for female leadership requirements
2. Rotation schedule determined by lottery, not local preferences
3. Creates random assignment of female vs. male political leadership

Results: Villages with female leaders invested significantly more in drinking water infrastructure and roads

When to Collect Data

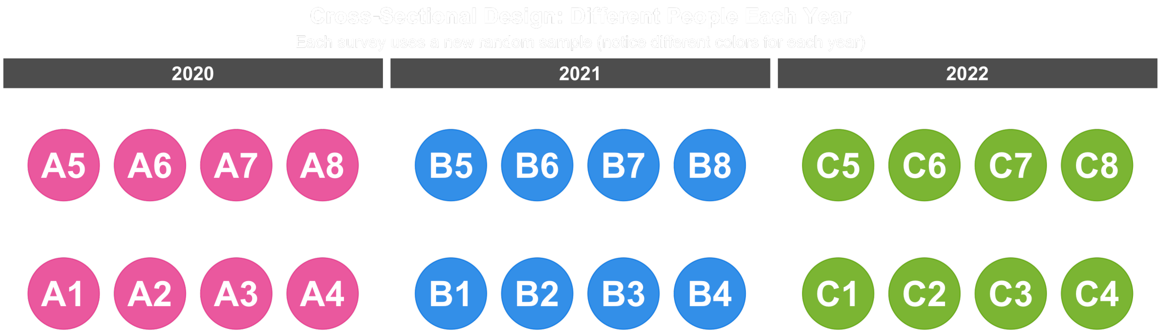
Once? Multiple times?

How often can we collect data: Once (cross-sectional) or multiple times (repeated cross-sectional or longitudinal)?

Examples across research designs:

- **Experiments**: Can measure participants once (post-treatment only) or multiple times (pre- and post-treatment)
- **Natural experiments**: Can compare groups at one time point or track them over multiple time points
- **Observational studies**: Can survey different people each year or follow the same people over time

Cross-Sectional Data: Snapshots in Time



Cross-sectional data limitations:

- **Pros**: Quick, less expensive, good for current conditions
- **Cons**: Can't show causation, timing unclear, can't track individual change

Longitudinal Panel Data: Following the Same People

Longitudinal panel data advantages:

- **Pros**: Can show change, establish timing, stronger causal inference, track individual trajectories
- **Cons**: Expensive, people drop out (attrition), takes years

A Story from 1662: John Graunt

The World's First Data Scientist/ founder of demography

1606.

A TABLE of the

CHRISTENINGS and MORTALITY

For the Year 1605 and 1606.*

Weeks.	Days of the Month.		Chrift.	Bur.	Pla.	Par. infec.	Weeks.	Days of the Month.		Chrift.	Bur.	Pla.	Par. infec.
1	Dec.	26.	100	116	5	5	28	July	3.	109	110	25	12
2	January	2.	117	151	6	5	29		10.	111	134	33	18
3		9.	130	138	4	4	30		17.	115	146	50	22
4		16.	124	138	3	2	31		24.	96	140	46	26
5		23.	143	121	6	4	32		31.	132	178	66	29
6		30.	124	101	3	2	33	Auguft	7.	131	181	67	29
7	Febr.	6.	122	105	5	5	34		14.	141	197	75	33
8		13.	131	118	7	6	35		21.	133	189	85	28
9		20.	126	109	12	6	36		28.	125	207	85	29
10		27.	102	117	9	8	37	Septem.	4.	123	241	116	32
11	March	6.	110	98	7	4	38		11.	134	216	105	28
12		13.	126	137	9	7	39		18.	121	214	92	36
13		20.	123	133	14	11	40		25.	132	204	87	35
14		27.	134	123	17	8	41	October	2.	121	256	141	40
15	April	3.	123	114	13	9	42		9.	134	218	106	38
16		10.	132	145	27	11	43		16.	142	227	117	37
17		17.	139	129	12	8	44		23.	131	224	109	38
18		24.	118	110	11	7	45		30.	124	226	101	34
19	May	1.	92	136	17	10	46	Novem.	6.	136	183	68	27
20		8.	116	103	13	11	47		13.	125	162	41	20
21		15.	128	94	13	8	48		20.	121	145	28	11
22		22.	113	132	14	9	49		27.	143	123	22	13
23		29.	94	98	9	7	50	Decem.	4.	155	160	45	17
24	June	5.	129	112	16	8	51		11.	135	137	38	20
25		12.	127	112	19	14	52		18.	136	132	28	15
26		19.	121	119	15	10	53		25.	134	135	38	19
27		26.	132	126	24	16							

The Totals { Chriftened — — 6614
Buried — — 7920
Whereof of the Plague 2124 }

* BELL's London's Remembrancer.

A TABLE

John Graunt (1620-1674):

London haberdasher who introduced systematic analysis of death and population

The Problem:

Data on death and population was not consistently collected and was error prone.

Graunt:

Comparing years in the Bills of Mortality, he able to make estimates about the size of the population of London and England, birth rates and mortality rates of males and females, and the rise and spread of certain diseases.

What Graunt Discovered

Before Graunt: People thought plague deaths were random acts of God

Simulation of Graunt's Analysis:

```
# A tibble: 4 x 4
  season avg_plague_deaths avg_total_deaths total_parishes
  <fct>              <dbl>            <dbl>          <int>
1 Winter               2.1              7.4             20
2 Spring               3.3             12.4             20
3 Summer               5.3             19.7             20
4 Fall                 3               11               20
```

1. **Patterns exist**: Death rates weren't random - they followed predictable patterns
2. **Seasonal variation**: Clear seasonal patterns with highest deaths in summer (plague season), lowest in winter
3. **Demographic regularities**: Consistent ratios of male to female births across time
4. **Population estimation**: First systematic attempt to count London's population using data

Graunt invented the idea that social phenomena follow discoverable patterns - the foundation of all social science

Choosing the Right Design

Example: Voter Turnout Research

Question 1: "What percentage of Americans vote in midterm elections?"

- **Best Design**: Descriptive survey or administrative data analysis
- **Why**: Simple descriptive question requiring representative data

Question 2: "Do voter reminders increase turnout?"

- **Best Design**: Randomized field experiment
- **Why**: Can randomly assign reminders, ethical to do so, clear causal question

Question 3: "Does early voting increase overall turnout?"

- **Best Design**: Natural experiment or difference-in-differences
- **Why**: Can't randomly assign electoral rules, but can compare before/after adoption

Tradeoffs Between Designs

Internal vs. External Validity

Internal Validity: Are your results correct for your specific study?

- **Experiments**: High internal validity (random assignment eliminates confounding)
- **Observational studies**: Lower internal validity (many potential confounders)

External Validity: Do your results apply to other people, places, times?

- **Experiments**: Often lower external validity (artificial settings, selected populations)
- **Observational studies**: Often higher external validity (real-world settings)

The Fundamental Tradeoff

No design is perfect

- **Experiments**: Great for causation, limited for generalization
- **Natural experiments**: Balance of both, but rare
- **Observational studies**: Great for description, limited for causation

Your choice depends on your research question and what tradeoffs you're willing to accept

Key Concepts to Remember

- **Research design drives everything** - your question determines your approach
- **Random assignment eliminates confounding** - the power of experiments
- **Natural experiments offer a middle ground** - when you can't randomly assign
- **All designs have tradeoffs** - choose based on your priorities
- **History matters** - Gaunt showed us that social patterns are discoverable

Questions?

Key takeaway: The right research design is the one that best matches your research question while acknowledging the tradeoffs you're willing to accept.