

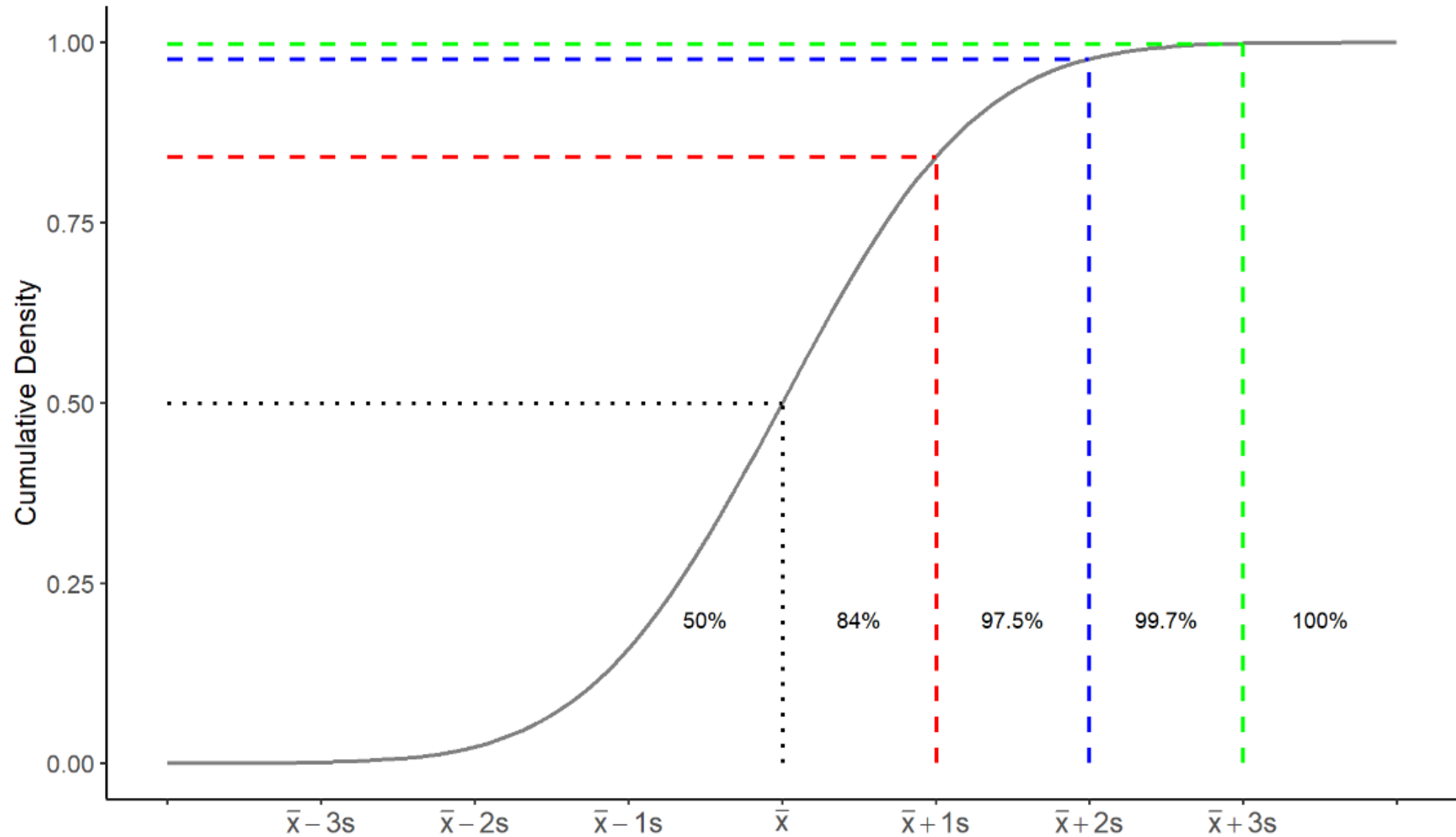
Lecture 8

Types of Studies

Review From Monday:

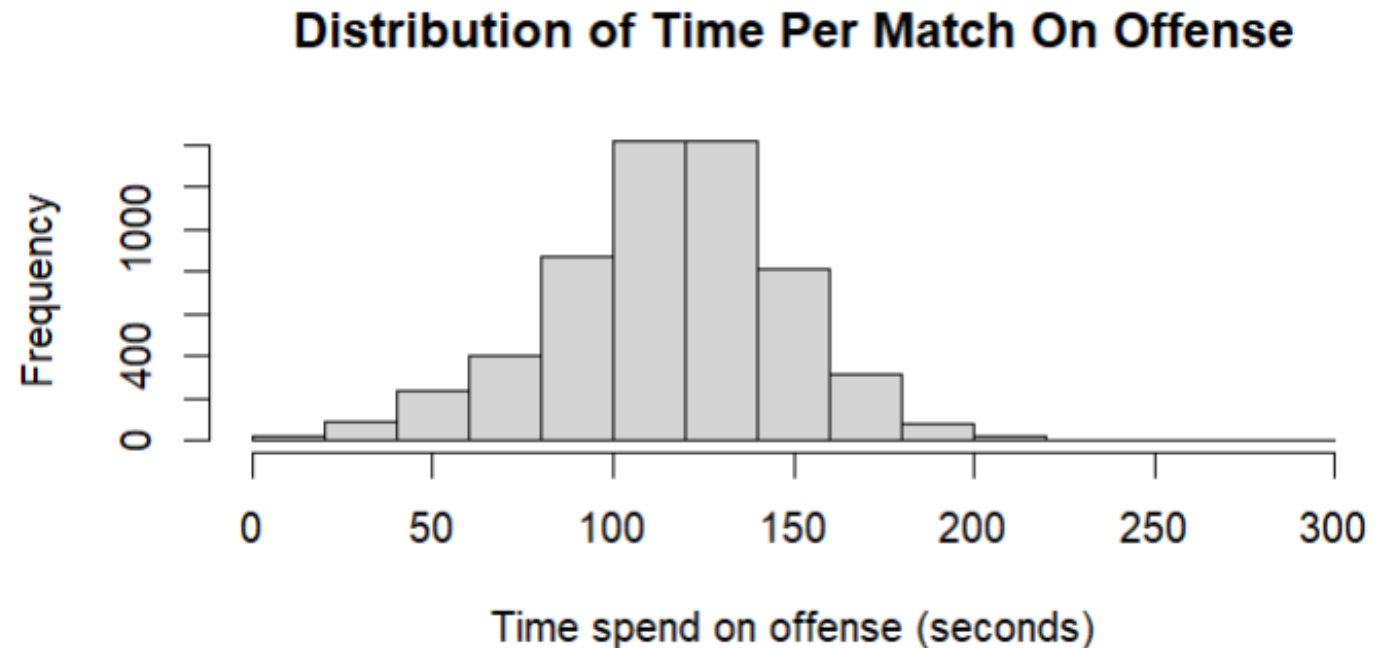
- Density Curve
 - What is a density curve?
 - What types of variables have density curves?
- Normal Distribution
 - What is the shape of normal distribution?
- Empirical Rule
 - Applies to what kind of distribution?
 - What proportion of data will have a value in the interval $\bar{x} \pm 1\sigma$?
 - What proportion of data will have a value in the interval $\bar{x} \pm 2\sigma$?
 - What proportion of data will have a value in the interval $\bar{x} \pm 3\sigma$?
- Z-scores
 - What is a z-score?
 - What does a z-score tell us about an observation?
 - Does z-score convert a variables distribution to the normal distribution?

Cumulative Density Function: Normal Distributions



Practice: This Is Rocket League!!!

- Rocket League is a popular online video game and E-sport that emerged in 2015. The game enjoys a healthy following of around 93 million players per month. The game features players from around the world who compete in sports like soccer (football), basketball and hockey while controlling RC-like vehicles.

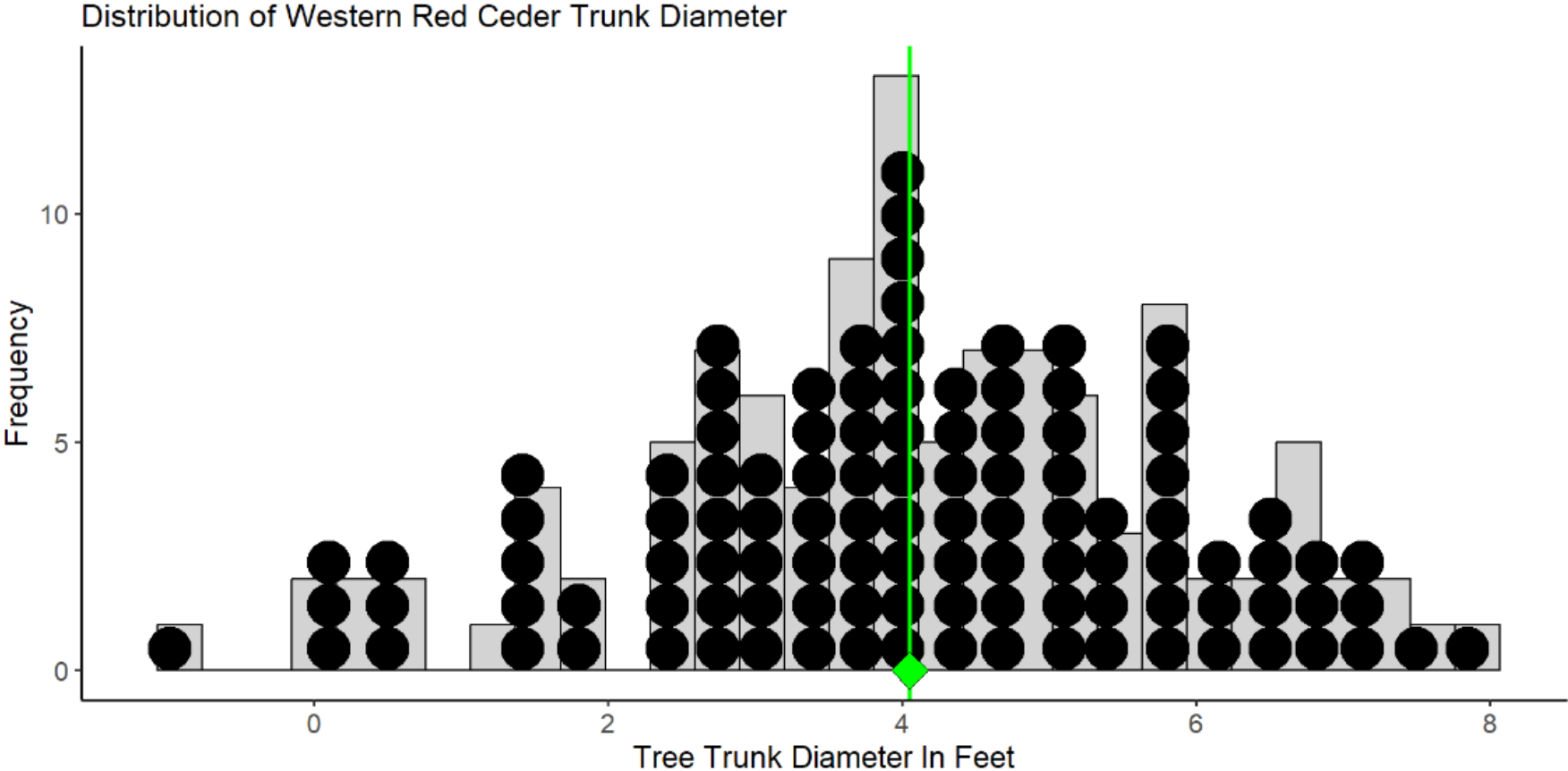


$$\bar{x} = 115, \quad s = 33$$

Estimating Diameter of Western Red Ceders

Consider a small study to estimate the average trunk diameter (in feet) of Western Red Ceders at the idlers rest nature preserve just outside of Moscow.

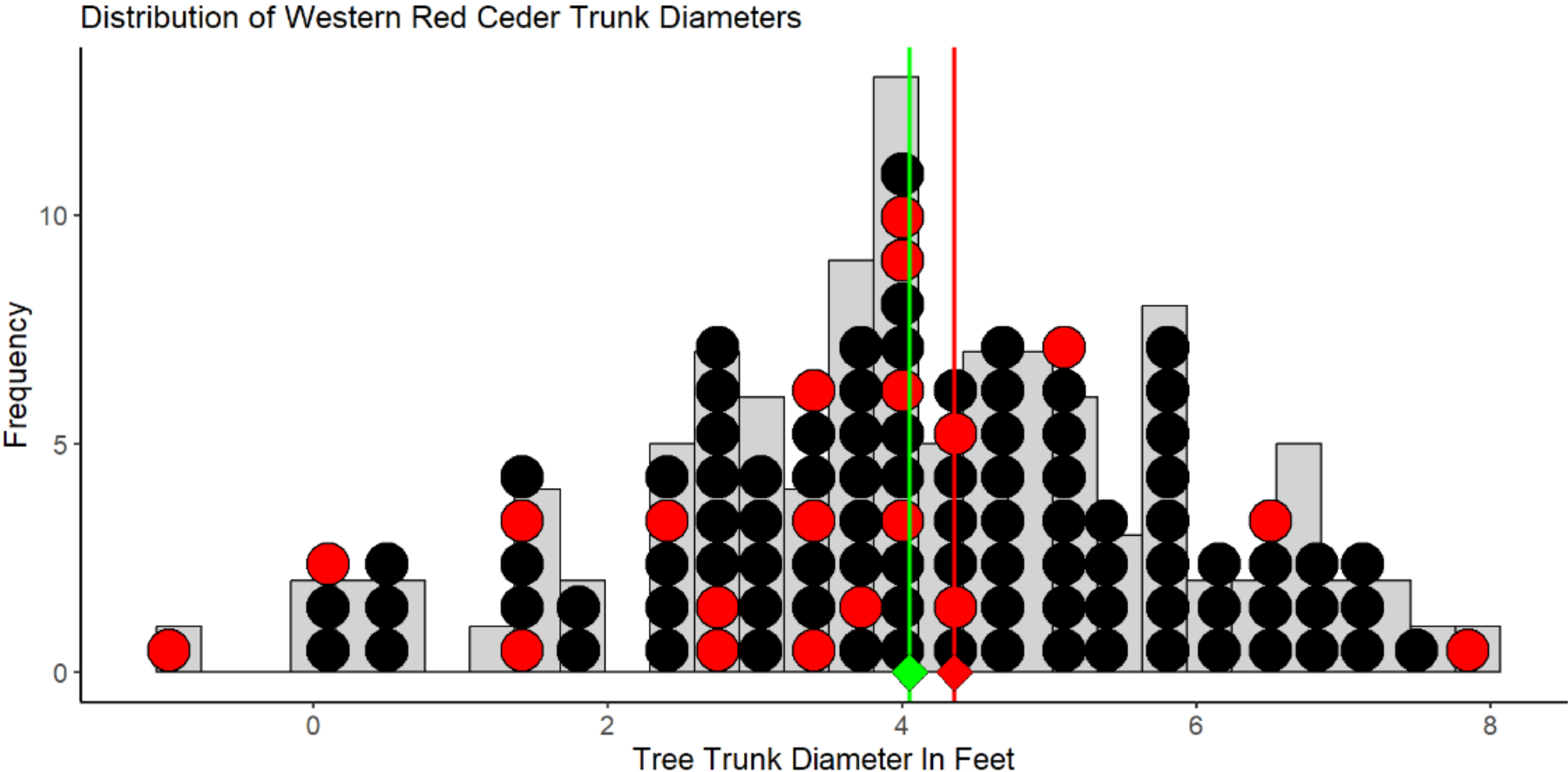
$$N = 109$$
$$\mu = 4.05$$
$$\sigma = 1.81$$



Estimating Diameter of Western Red Ceders

Consider a sample of $n = 20$ trees to estimate the mean trunk diameter...

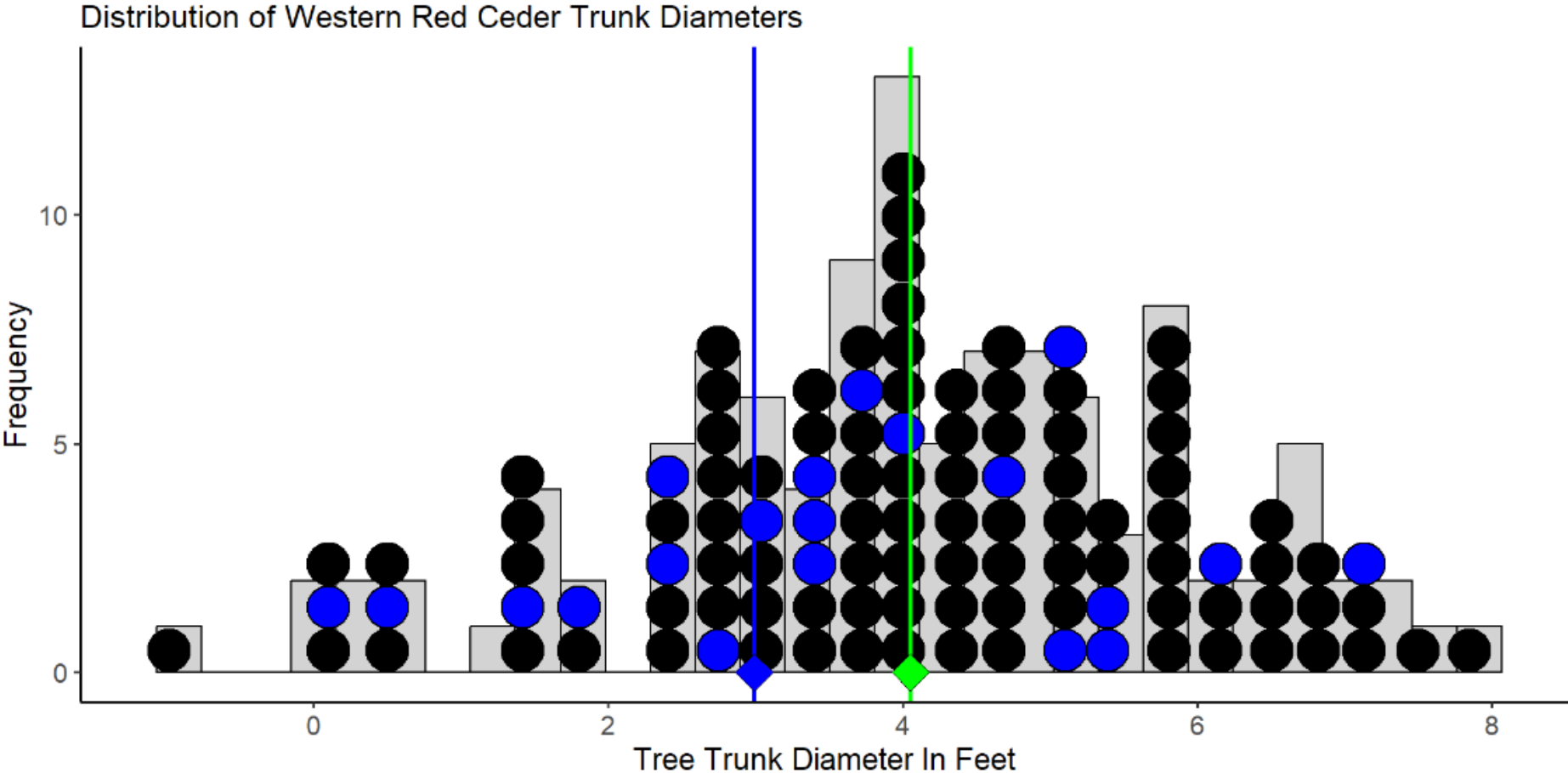
$$\bar{x}_1 = 4.35$$



Estimating Diameter of Western Red Ceders

Consider another sample of $n = 20$ trees to estimate the mean trunk diameter...

$$\bar{x}_2 = 2.98$$



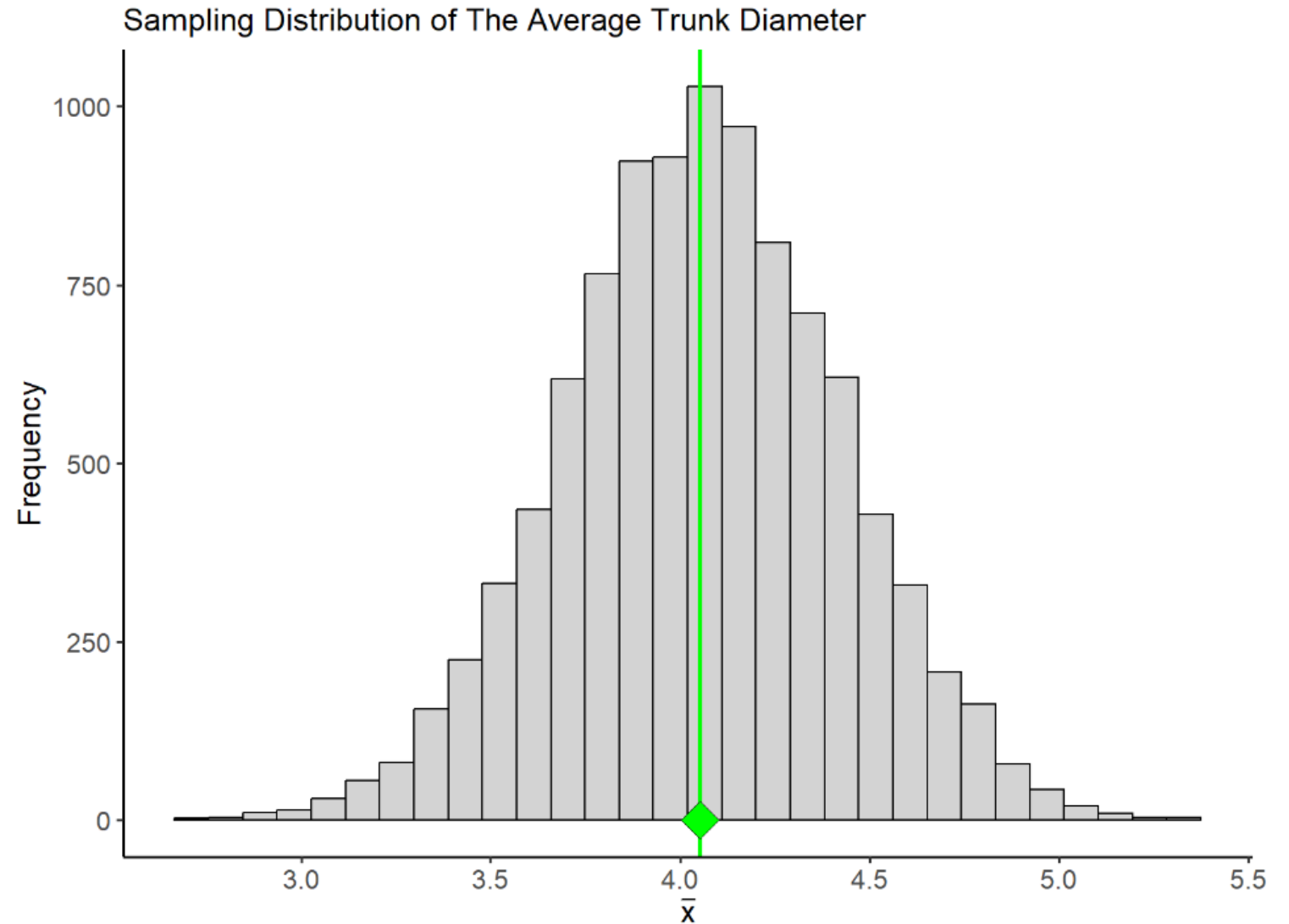
Sampling Distributions

A **sampling distribution** is the distribution of a statistic.

- It arises from repeatedly sampling and estimation

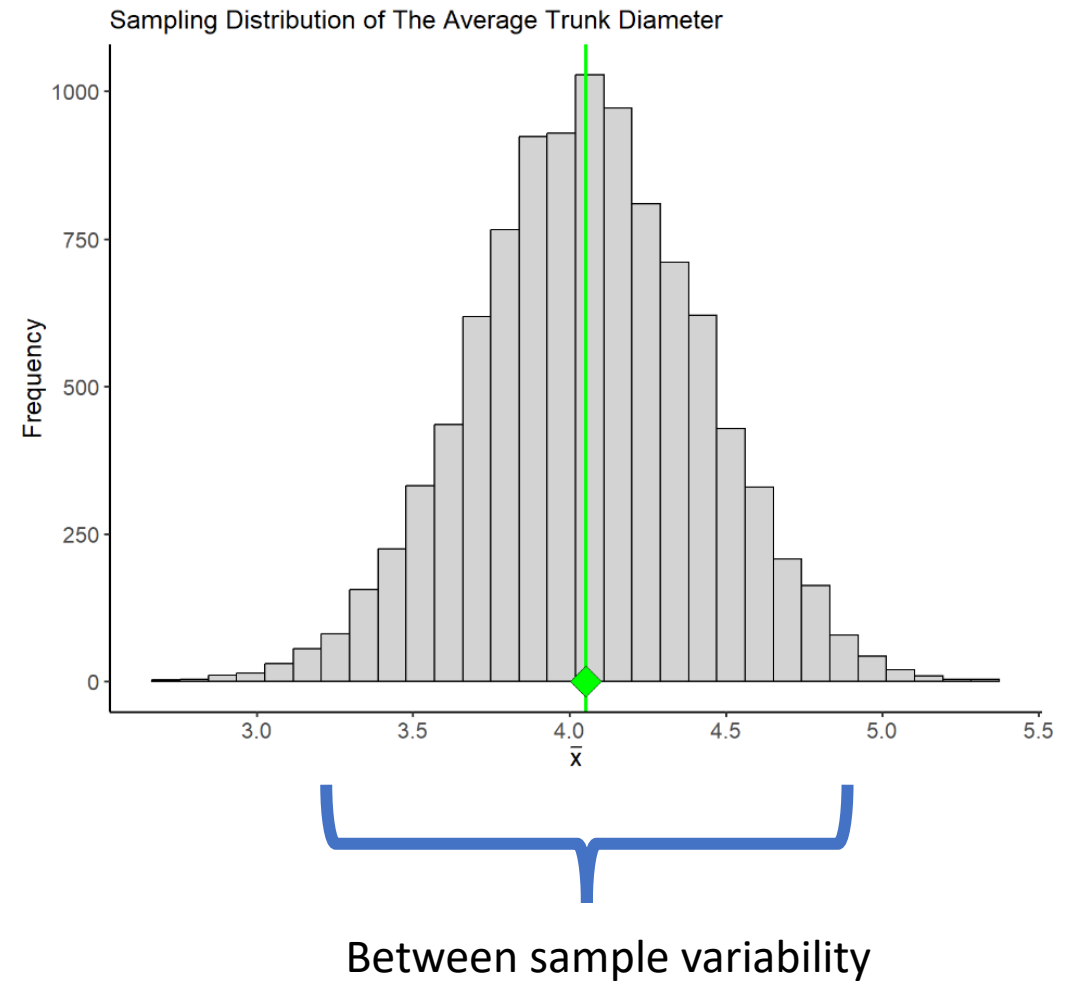
- it is usually a theoretical distribution

The variation between different samples leads to variation in the estimates that are calculated



Margin of Error

- The **margin of error** of an estimate measures how far we expect an estimate to fall from the true value of a population parameter
- It is a measure of the between sample variability in our estimate
- It is the largest distance between the true population parameter and an estimate that is not an outlier



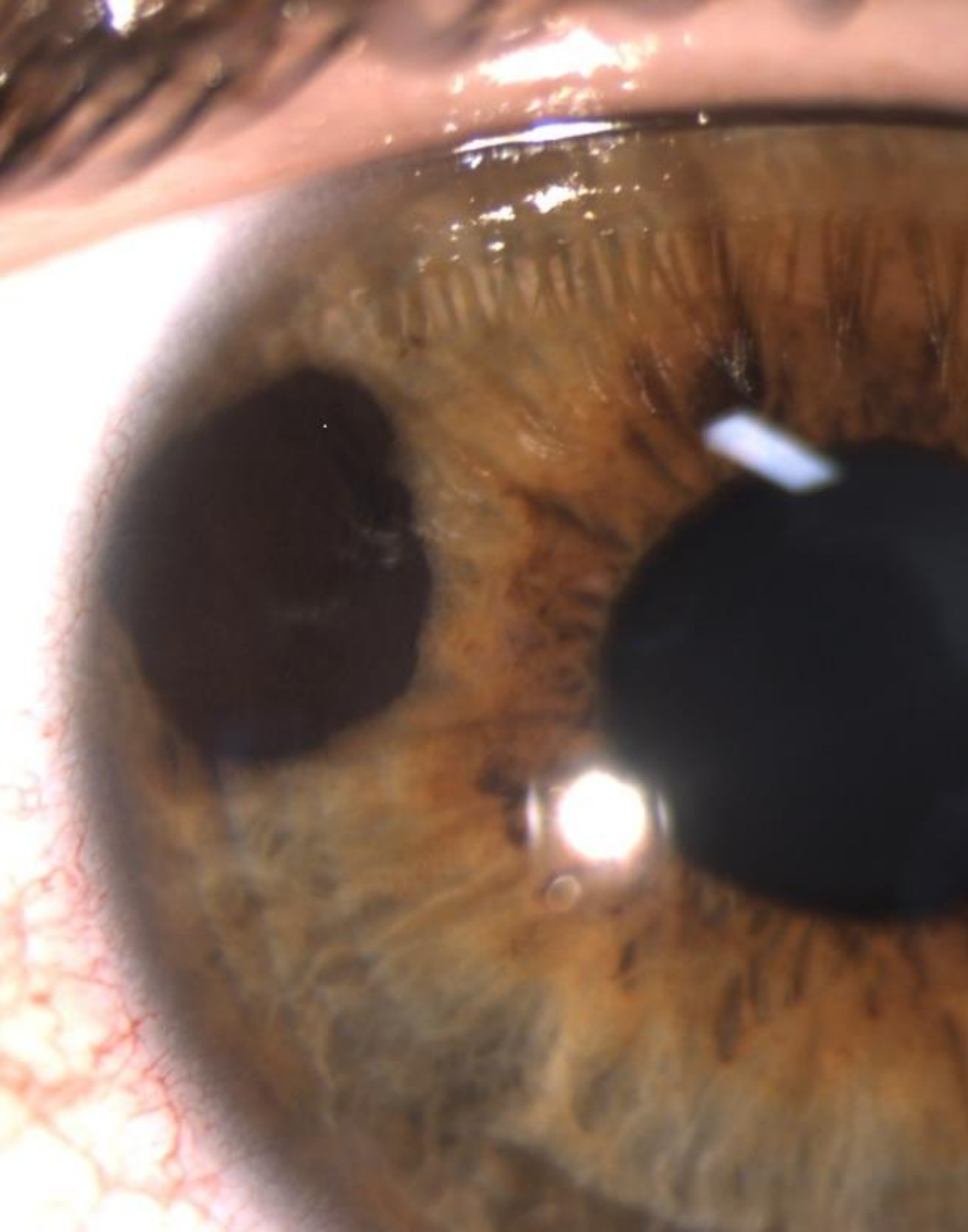
Statistical Significance

- A **Statistically significant** result is one that is decidedly *not* due to ordinary variation - this means a result that is not due to chance or coincidence.
- statistically significant results is which falls outside the margin of error.

Sampling and Experimental Designs

Anecdotal evidence

- Descriptive statistics alone is not enough evidence to make conclusive decisions about a variable or patterns
- Patterns in data can arise from many different sources
- **Anecdotal evidence** - evidence from information or testimony that is based on personal observations, individual experiences, or isolated examples, rather than on systematic and rigorous scientific analysis.
 - Anecdotal evidence often starts with phrases like “In my experience” or “it seems to me”



The German Study (Stang et al., 2001)

- This study compared 118 patients with a rare form of eye cancer called uveal melanoma to 475 healthy patients who did not have this eye cancer.
- Cell phone use was measured using a questionnaire
- Findings: on average, the eye cancer patients used cell phones more often

Study 1: The German Study (Stang et al., 2001)

- What is question the authors are trying to answer with data?

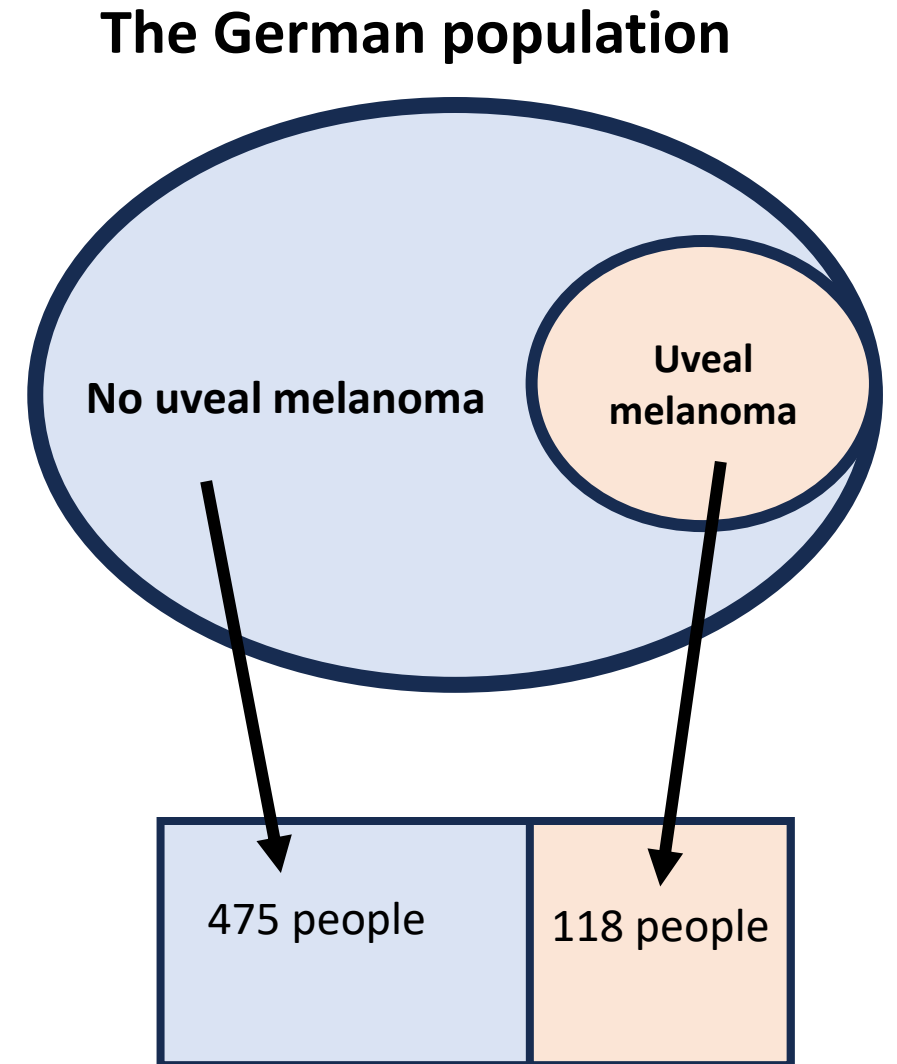
Is cell phone use associated with uveal melanoma?

- What is the population?

All citizens living in Germany

- What is the sample?

Two samples are taken: one taken from the subpopulation of people with uveal melanoma, another taken from the subpopulation of people without uveal melanoma





Study 2: The British Study (Hepworth et al., 2006)

- This study compared 966 patients with brain cancer to 1,716 healthy patients who did not have brain cancer.
- Cell phone use was measured using a questionnaire
- Findings: cell phone use for the two groups was similar

The British Study (Hepworth et al., 2006)

- What is question the authors are trying to answer with data?

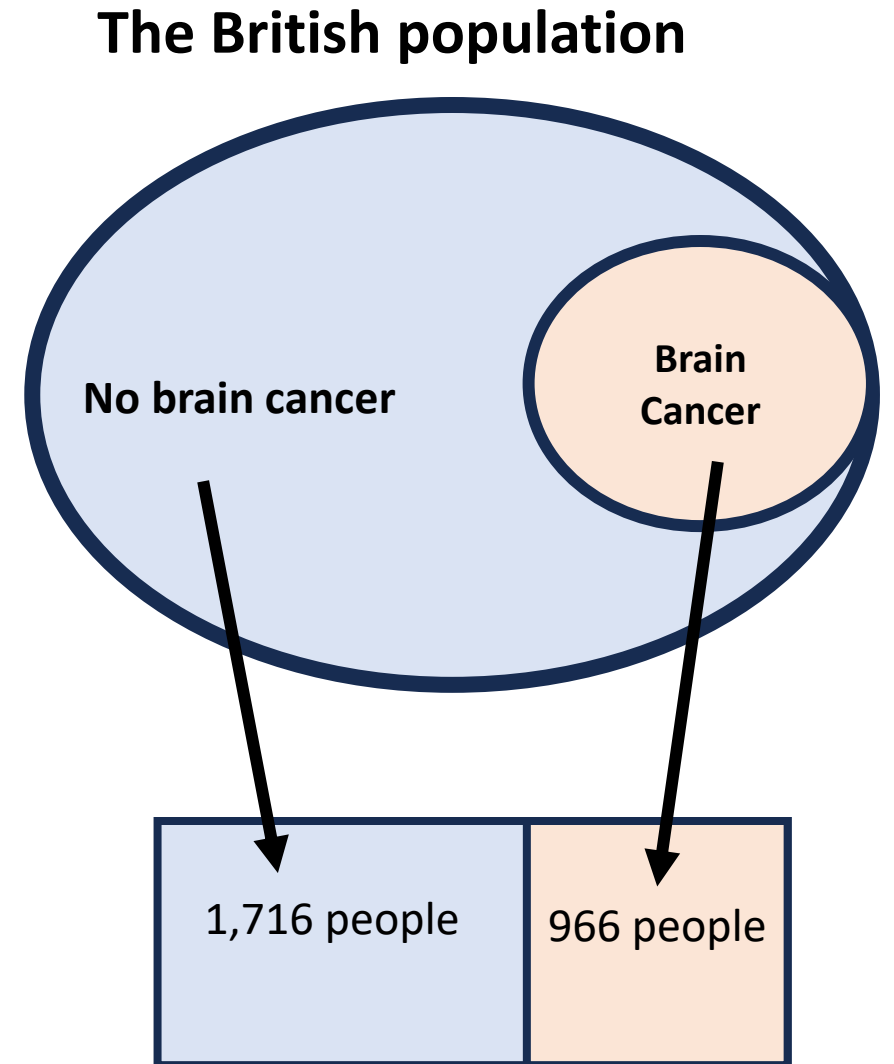
Is cell phone use associated with brain cancer?

- What is the population?

All citizens living in Great Britain

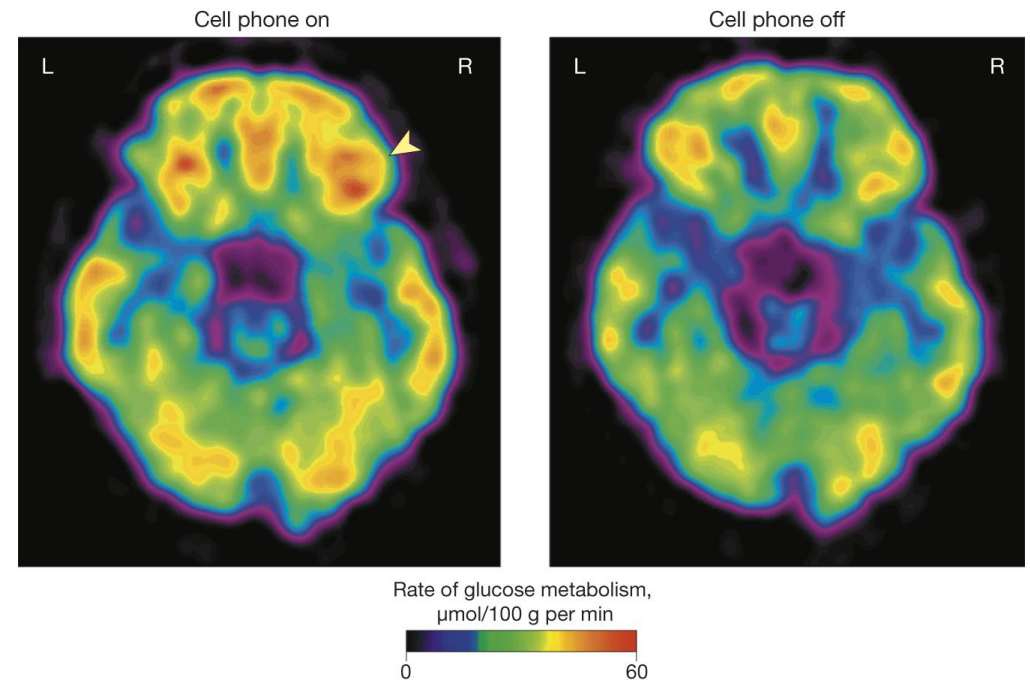
- What is the sample?

Two samples are taken: one taken from the subpopulation of people with brain cancer and another taken from the subpopulation of people without brain cancer



Study 3: (Volkow et al., 2011) – *Journal of the American Medical Association*

- Authors used a randomized crossover study to compare the effects of cell phone use on brain glucose metabolism in 47 individuals.
- Patients were fitted with two cell phone devices, one for each ear
 - Each patient was given two positron emission topography (PET) scans
 - The first was applied when both phones were off
 - The second was applied during a 50-minute muted call to one of the two cell phones (the call was randomly assigned to either the right or left phone)
- Comparison of the PET scans showed significantly increased activity in the part of the brain closest to the phone.



Study 3 (Volkow et al., 2011) – *Journal of the American Medical Association*

- What is question the authors are trying to answer with data?

Does cell phone use cause increased brain activity (measured as glucose metabolism)?

- What is the population?

Any healthy individual

- What is the sample?

47 individuals selected as participants

What is “good” data?

Studies 1 and 3 conflict with the results of study 2 – which should we trust?

Study and sampling designs can have a major impact on results...

Knowledge of different study designs helps guide us in deciding what research we should trust and when we should be skeptical.

Response and Explanatory Variables

- In all 3 studies investigating cell phone use and physiological activity in the brain, the researchers were interested in two variables, a response variable and an explanatory variable
- **Explanatory variable** – this is the variable we manipulate or observe changes in
 - what was the explanatory variable in the German study?
- **Response variable** – this variable measures the outcome of interest. Studies focus on how the outcome “responds” to changes in the explanatory variable
 - what was the response variable in the German study?
- A study can have multiple response variables and multiple explanatory variables

Experimental vs Observational Studies



In an **experimental study**, researchers assign subjects to experimental conditions called **treatments** and then observe outcomes of the response variable(s).



In an **observational study** the researcher observes values of the response and explanatory variables in different subjects without any manipulation of the subjects in the study



Which (if any) of the three studies we examined are experimental? Which are observational?

Association vs Causation

- A **statistical association** between two variables is a numerical measure of their relatedness.
- A **causal** relationship between two variables means that changing one variable causes a proportional change in the other variable (also called a **cause and effect** relationship)
- An association between two variables **does not** imply causal relationship between them
- Example: there is a statistical association between the number of people who drowned by falling into a pool and the number of films Nicolas Cage appeared in a given year. However, there is obviously no causal relationship.

