

SOC 3811/5811:
BASIC SOCIAL STATISTICS

Sampling Error

Sampling

FiveThirtyEight

Polling averages are adjusted based on state and national polls, which means candidates' averages can shift even if no new polls have been added to this page. [Read more about the methodology.](#)

FILTER BY POLLSTER GRADE



Search polls

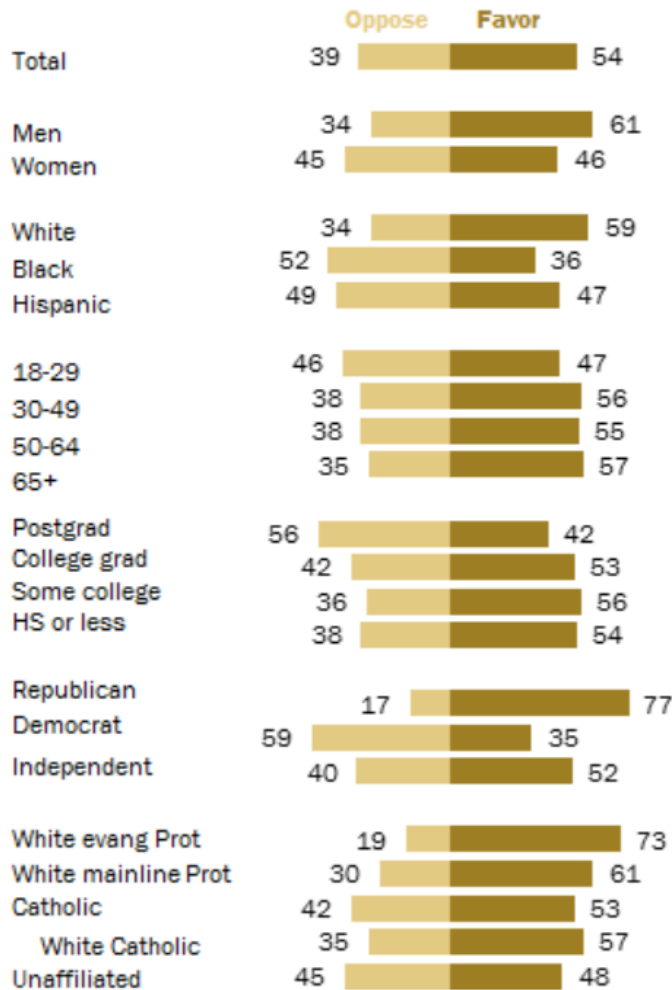
Added today

	DATES	POLLSTER	SAMPLE	RESULT			NET RESULT		
President: general election	Minn.	SEP 8-13, 2020	A+ ABC News/The Washington Post	615 LV	Biden	57%	41%	Trump	Biden +16
	Minn.	SEP 8-13, 2020	A+ ABC News/The Washington Post	705 RV	Biden	57%	40%	Trump	Biden +17

KEY A = ADULTS RV = REGISTERED VOTERS V = VOTERS LV = LIKELY VOTERS

Gender, racial differences in opinions about the death penalty

% who ___ the death penalty for persons convicted of murder



Note: Whites and blacks include only those who are not Hispanic; Hispanics are of any race. Don't know responses not shown.
Source: Survey of U.S. adults conducted April 25-May 1, 2018.

PEW RESEARCH CENTER

“The analysis in this report is based on telephone interviews conducted April 25-May 1, 2018 among a national sample of 1,503 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia”

How is this possible?

Vocabulary

Population

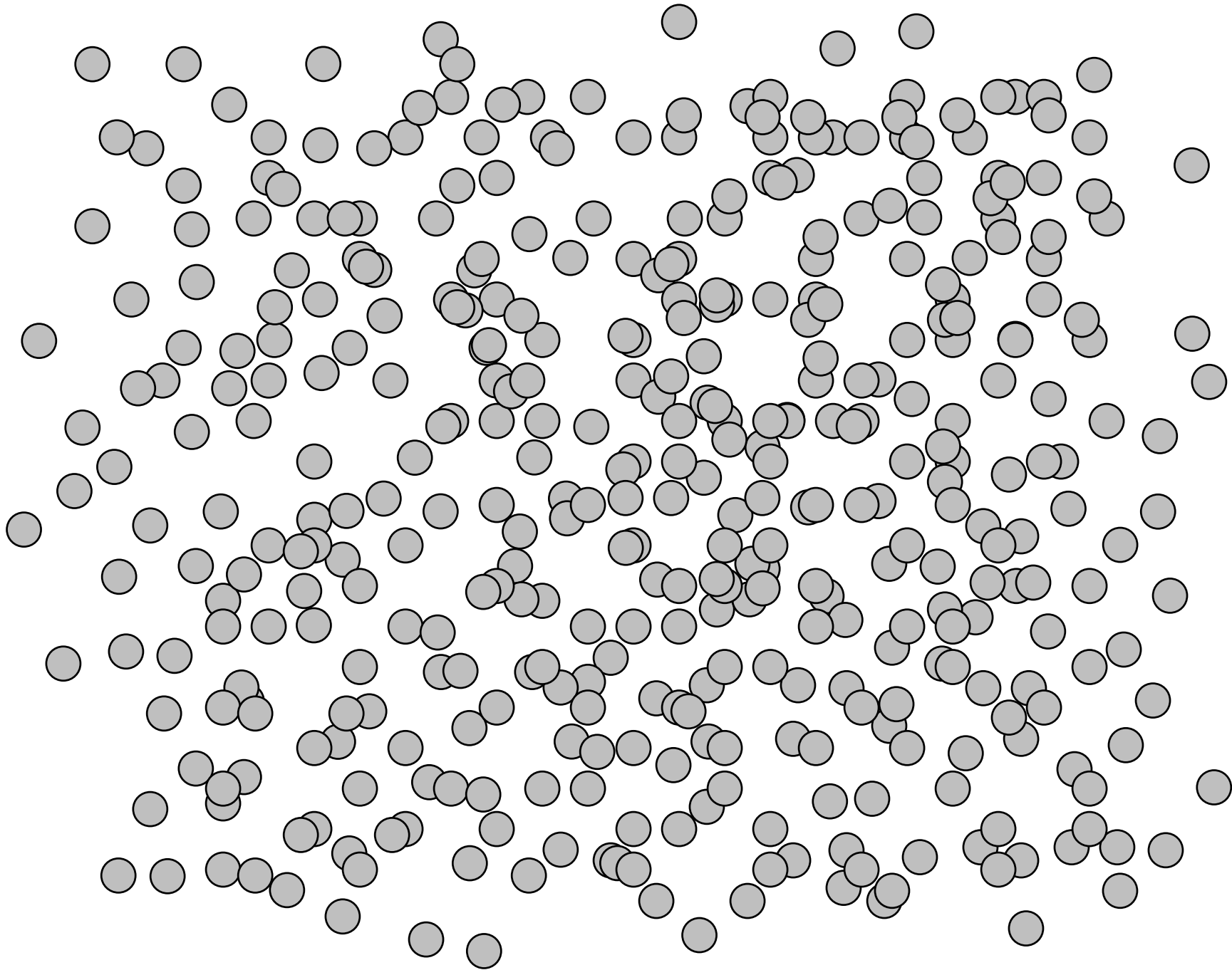
All of the individuals (or “units”) about whom we wish to make conclusions

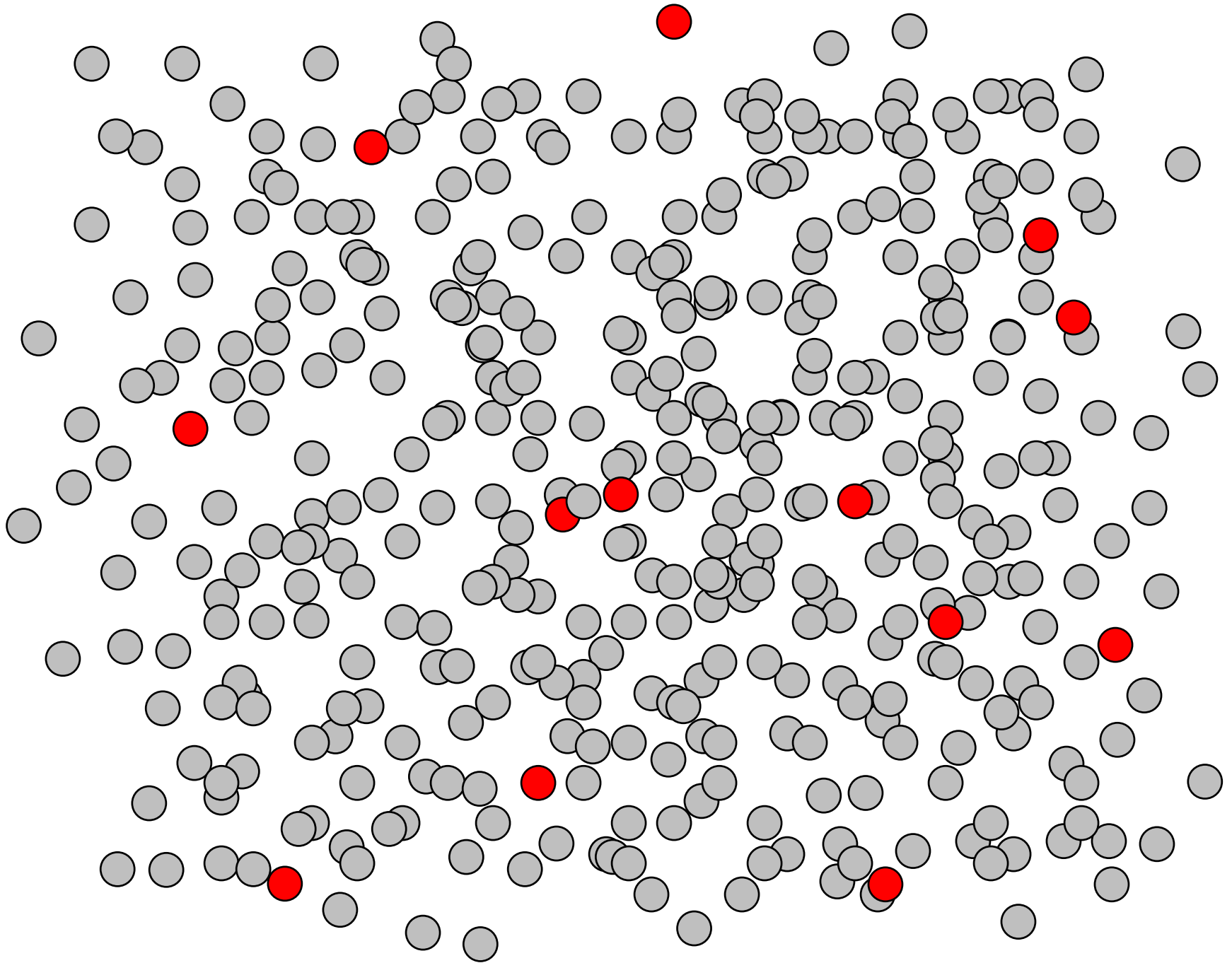
Sample

A subset of this population; it is this subset that we actually observe (or survey or interview or whatever)

Sampling

The process of selecting individuals from the population for inclusion in the sample





Vocabulary

Population Parameter

An attribute of the entire population. For example, the mean or variance of some variable in the full population

Sample Statistic (or Estimate)

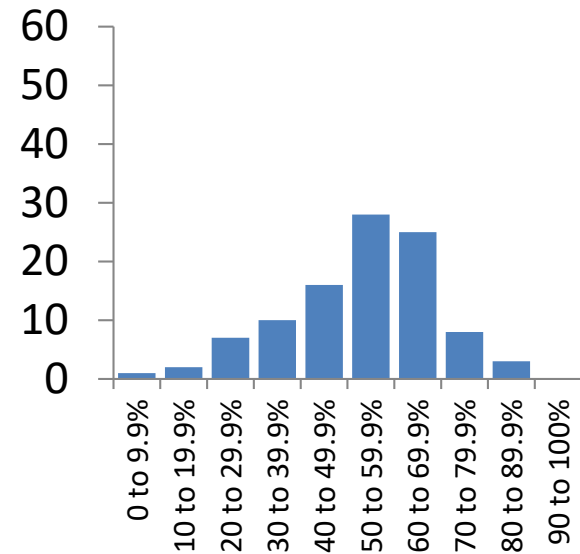
An attribute of a sample, sometimes used to make inferences about the corresponding population parameter

Sampling Error

The difference between a population parameter and the sample statistic being used to estimate it

Sampling Error

Question: If you flip a fair coin, what percentage of the time will it come up “heads?”

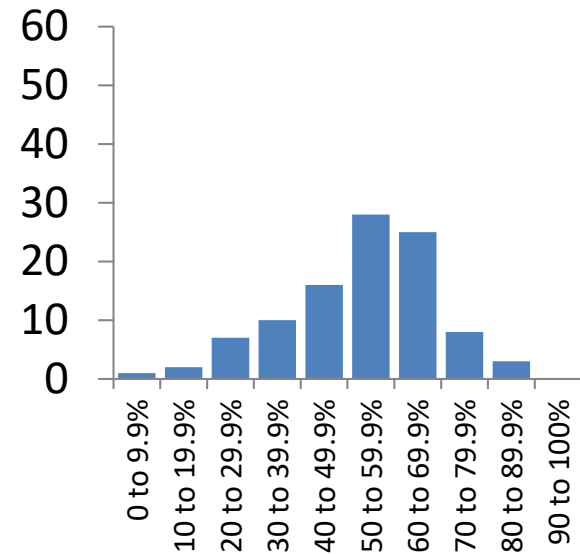


Results from flipping a
fair coin **10** times;
trial repeated 100 times

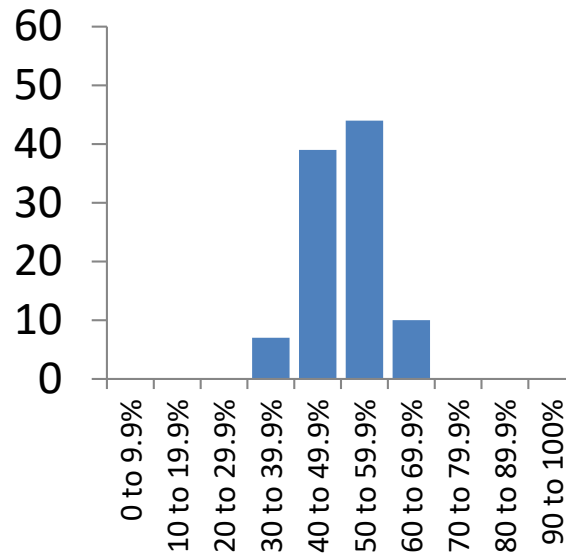
Population Parameter: Percentage of times “heads” comes up among all fair coins
Sample Statistic: Percentage of times “heads” comes up out of 10 flips of a fair coin

Sampling Error

Question: If you flip a fair coin, what percentage of the time will it come up “heads?”



Results from flipping a fair coin **10** times; trial repeated 100 times

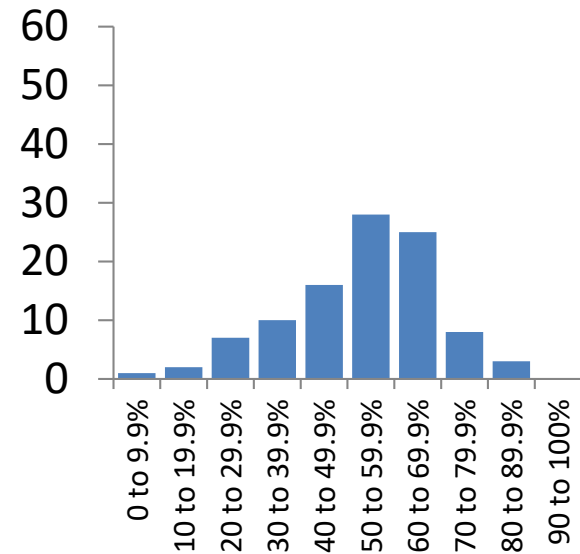


Results from flipping a fair coin **50** times; trial repeated 100 times

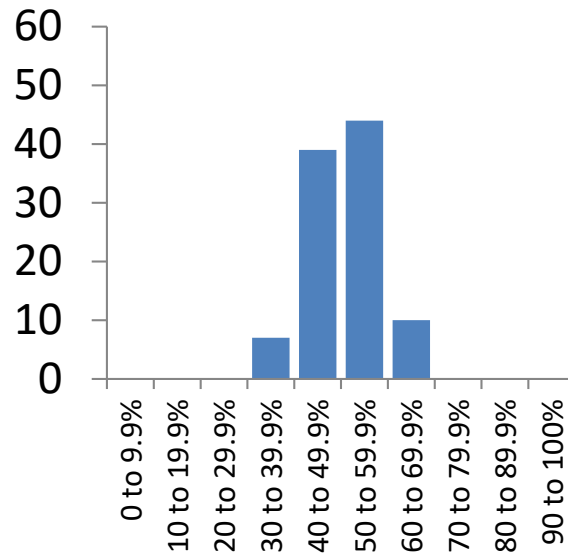
Population Parameter: Percentage of times “heads” comes up among all fair coins
Sample Statistic: Percentage of times “heads” comes up out of 10 flips of a fair coin

Sampling Error

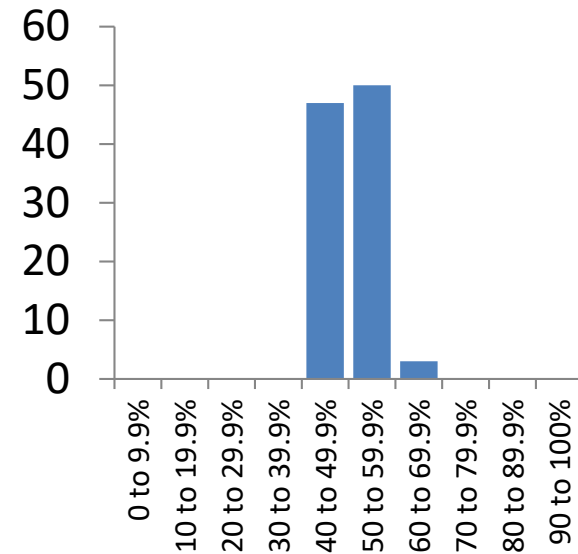
Question: If you flip a fair coin, what percentage of the time will it come up “heads?”



Results from flipping a fair coin **10** times; trial repeated 100 times



Results from flipping a fair coin **50** times; trial repeated 100 times

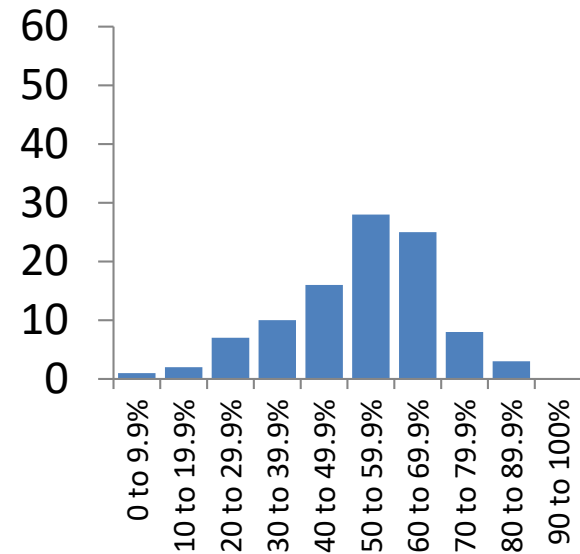


Results from flipping a fair coin **100** times; trial repeated 100 times

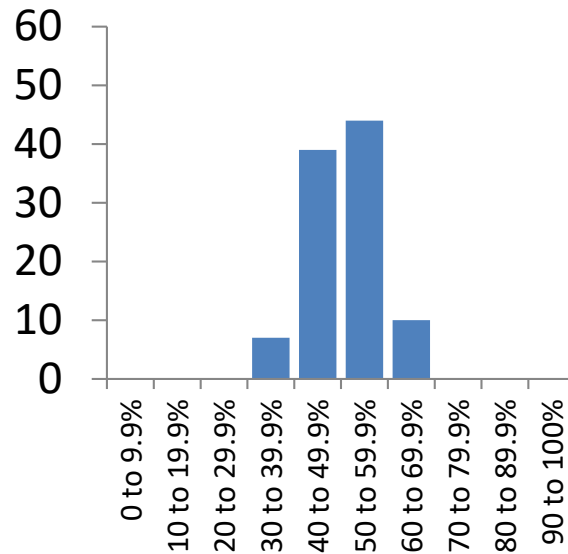
Population Parameter: Percentage of times “heads” comes up among all fair coins
Sample Statistic: Percentage of times “heads” comes up out of 10 flips of a fair coin

Sampling Error

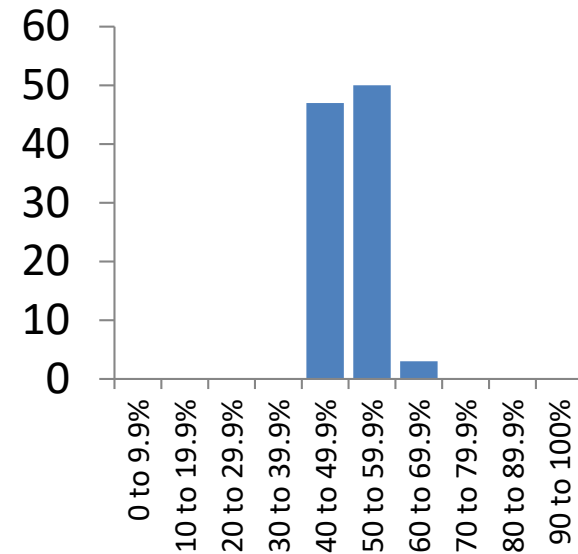
Question: If you flip a fair coin, what percentage of the time will it come up “heads?”



Results from flipping a fair coin **10** times; trial repeated 100 times



Results from flipping a fair coin **50** times; trial repeated 100 times

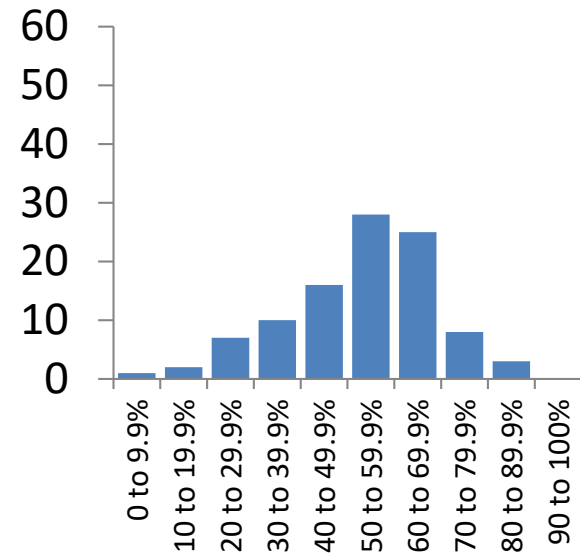


Results from flipping a fair coin **100** times; trial repeated 100 times

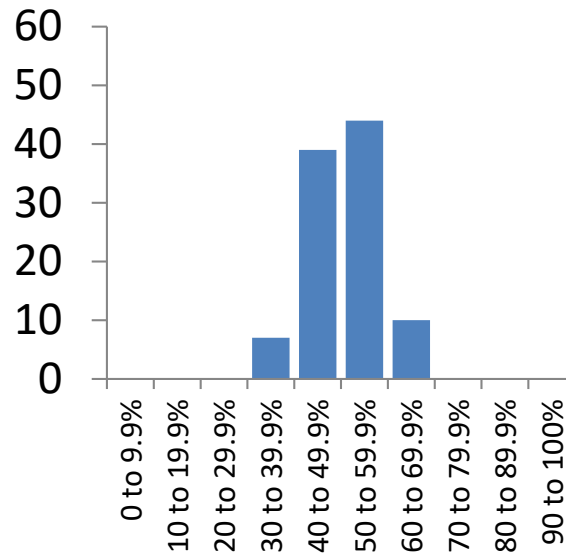
This is a dumb example because we know the population parameter ... it's 50% (It's also too convenient, because all fair coins are the same)

Sampling Error

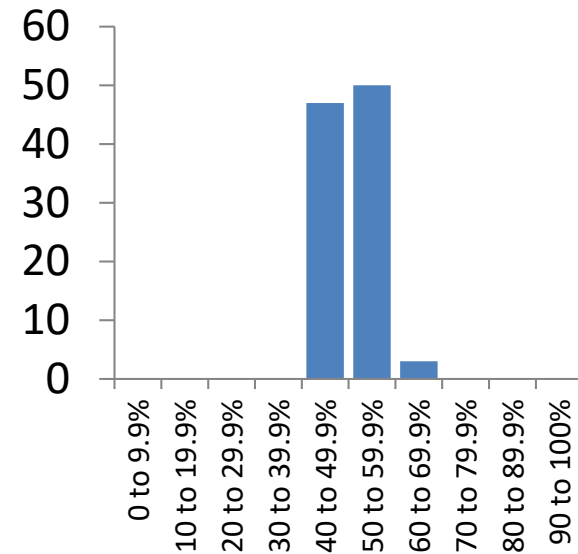
Question: If you flip a fair coin, what percentage of the time will it come up “heads?”



Results from flipping a fair coin **10** times; trial repeated 100 times



Results from flipping a fair coin **50** times; trial repeated 100 times

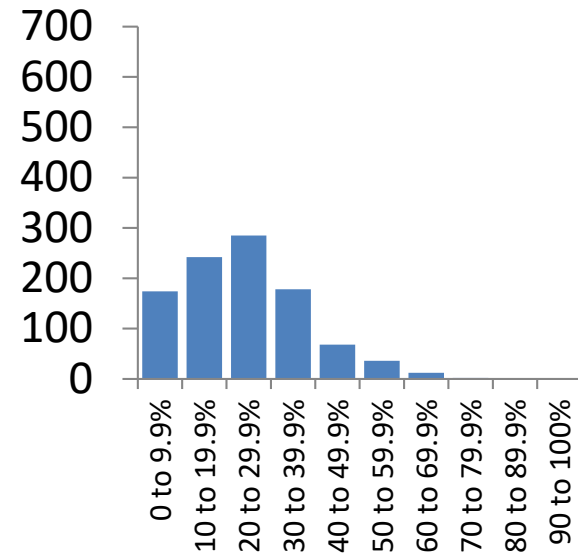


Results from flipping a fair coin **100** times; trial repeated 100 times

Note that (a) there is sampling error and (b) the size of the sampling error goes down as the sample size (i.e., number of flips) goes up

Sampling Error

Question: What percentage of Americans were high school graduates in 1940?



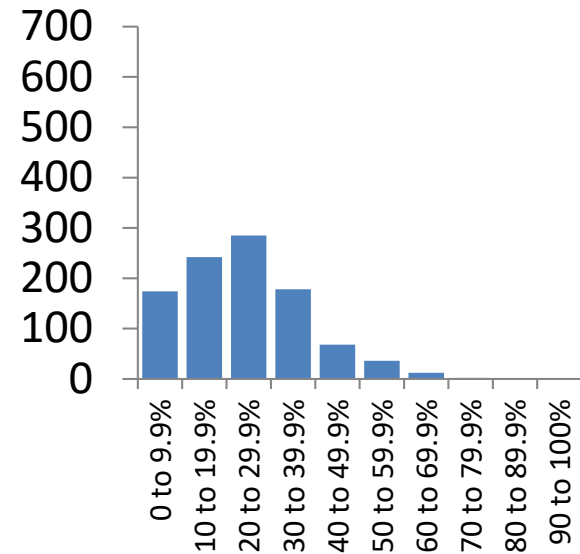
Results from asking **10**
randomly selected people;
trial repeated 1,000 times

Population Parameter: % of ALL Americans who were high school graduates in 1940

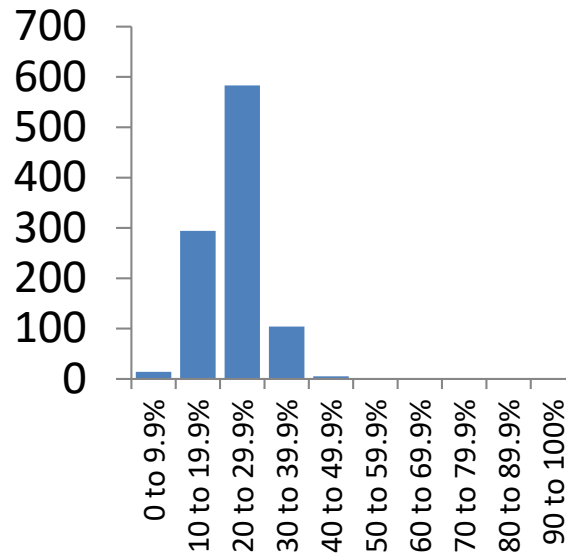
Sample Statistic: % of people in the sample who were high school graduates

Sampling Error

Question: What percentage of Americans were high school graduates in 1940?



Results from asking **10**
randomly selected people;
trial repeated 1,000 times



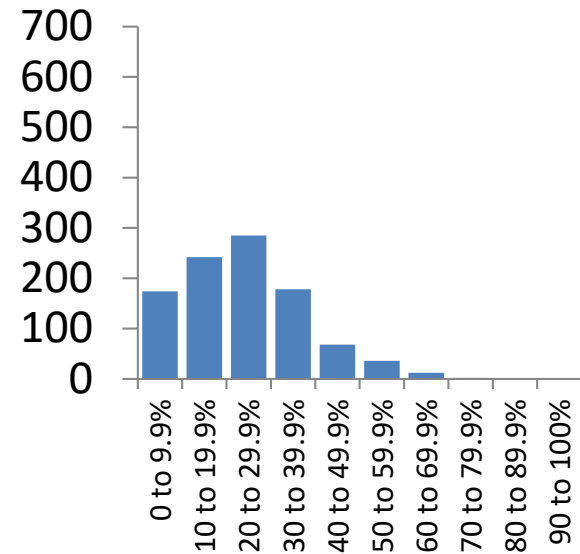
Results from asking **50**
randomly selected people;
trial repeated 1,000 times

Population Parameter: % of ALL Americans who were high school graduates in 1940

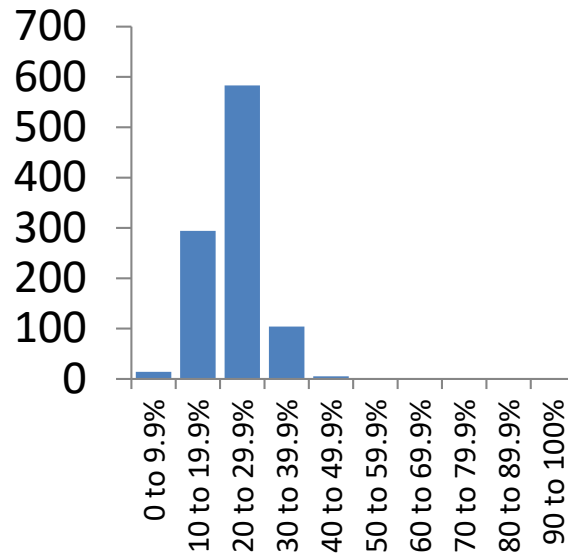
Sample Statistic: % of people in the sample who were high school graduates

Sampling Error

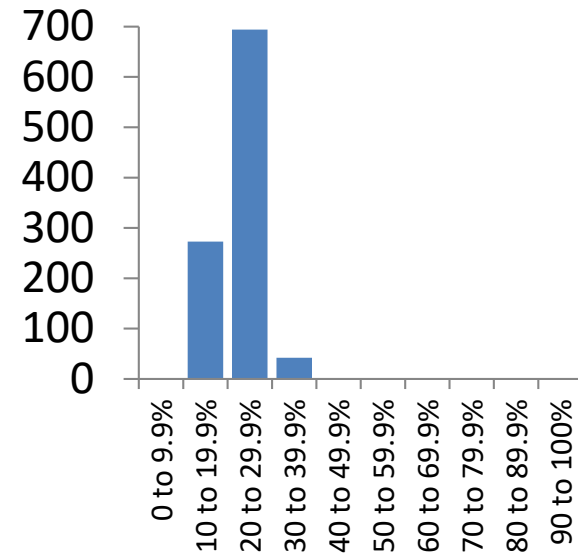
Question: What percentage of Americans were high school graduates in 1940?



Results from asking **10** randomly selected people; trial repeated 1,000 times



Results from asking **50** randomly selected people; trial repeated 1,000 times



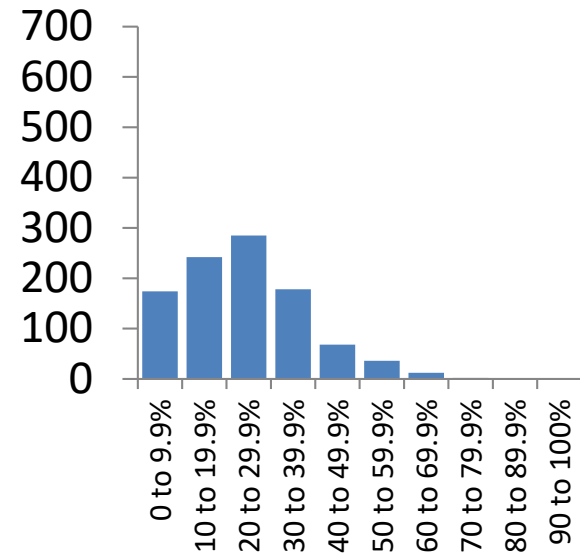
Results from asking **1,000** randomly selected people; trial repeated 1,000 times

Population Parameter: % of ALL Americans who were high school graduates in 1940

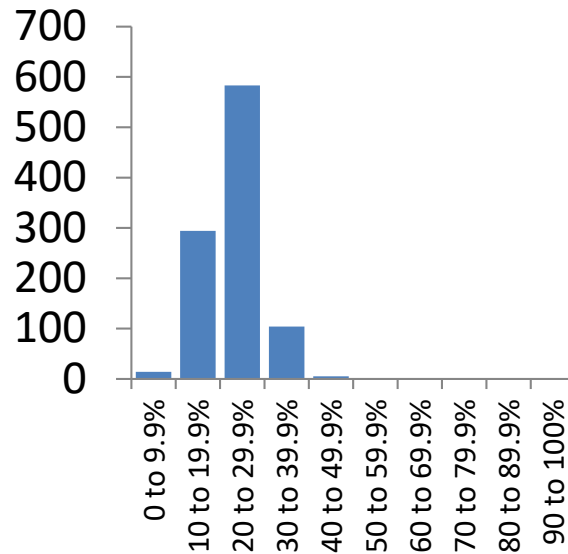
Sample Statistic: % of people in the sample who were high school graduates

Sampling Error

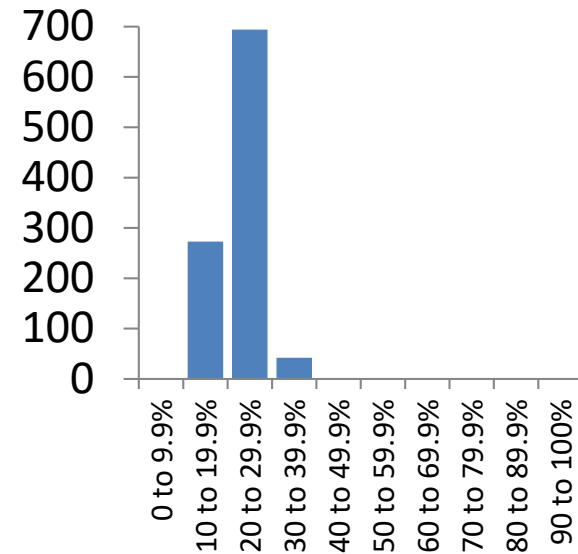
Question: What percentage of Americans were high school graduates in 1940?



Results from asking **10** randomly selected people; trial repeated 1,000 times



Results from asking **50** randomly selected people; trial repeated 1,000 times

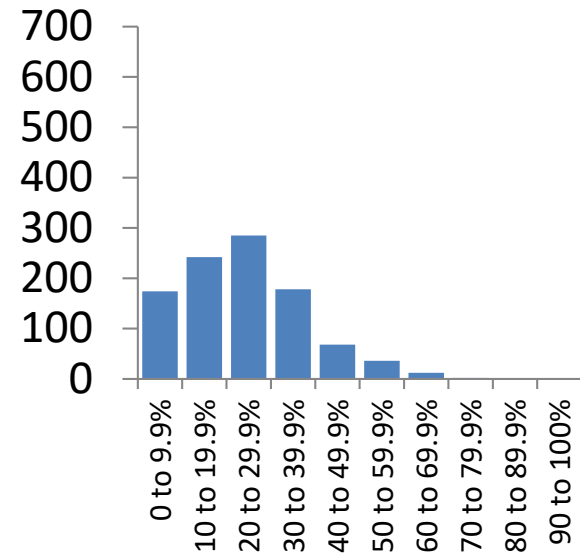


Results from asking **1,000** randomly selected people; trial repeated 1,000 times

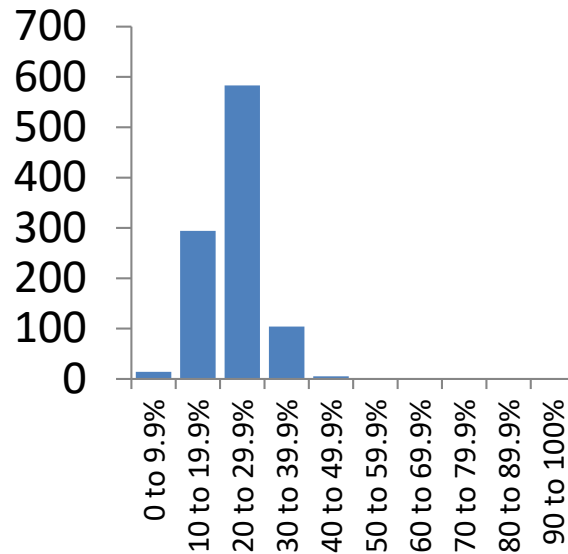
This is *also* a dumb example because we know the population parameter ... it's 23% (I know this because we have access to the full 1940 U.S. Census data)

Sampling Error

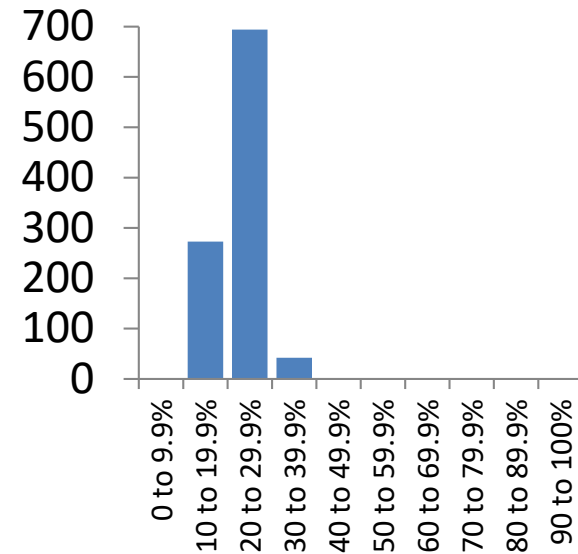
Question: What percentage of Americans were high school graduates in 1940?



Results from asking **10** randomly selected people; trial repeated 1,000 times



Results from asking **50** randomly selected people; trial repeated 1,000 times

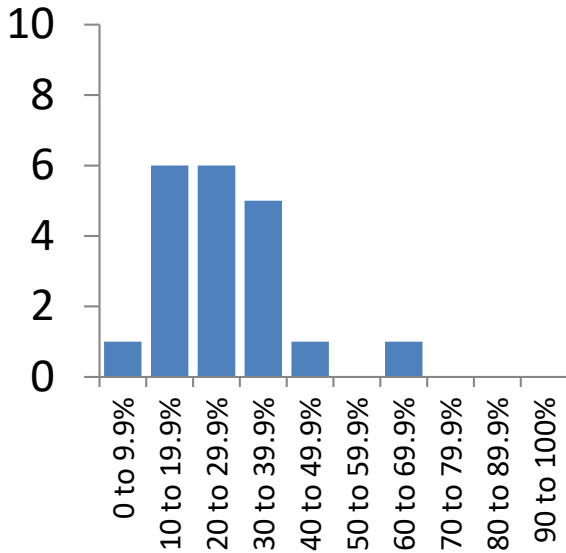


Results from asking **1,000** randomly selected people; trial repeated 1,000 times

Note again that (a) there is sampling error and (b) the size of the sampling error goes down as the sample size (i.e., number of people sampled) goes up

Sampling Error

Question: What percentage of Americans say it is true that humans evolved from an earlier species?

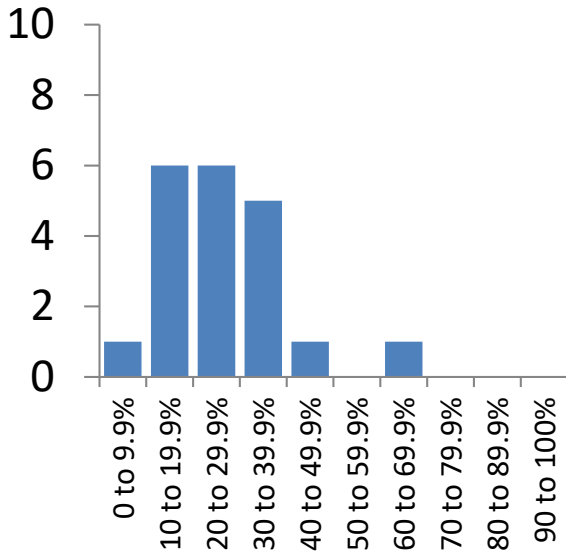


Results from asking **10** randomly selected people; trial repeated 20 times

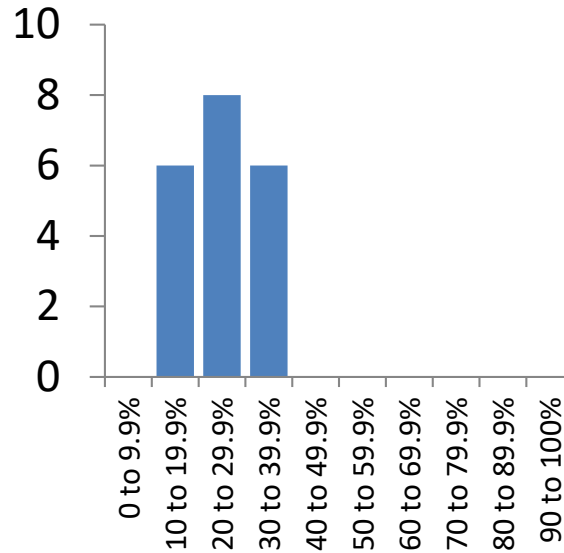
This is a *not* a dumb example ... we do not know the population value.

Sampling Error

Question: What percentage of Americans say it is true that humans evolved from an earlier species?



Results from asking **10** randomly selected people; trial repeated 20 times

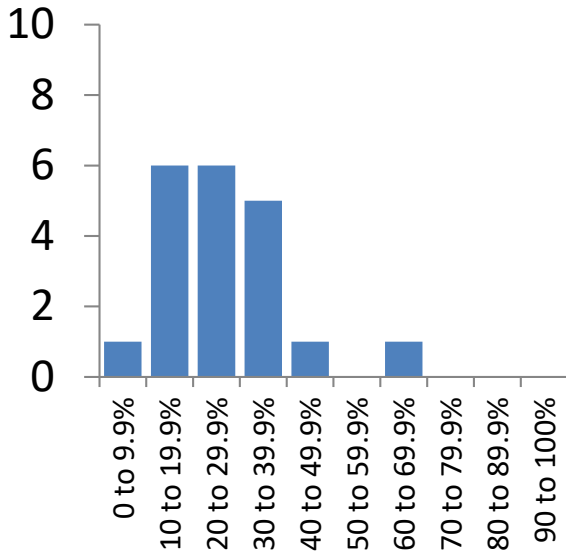


Results from asking **20** randomly selected people; trial repeated 20 times

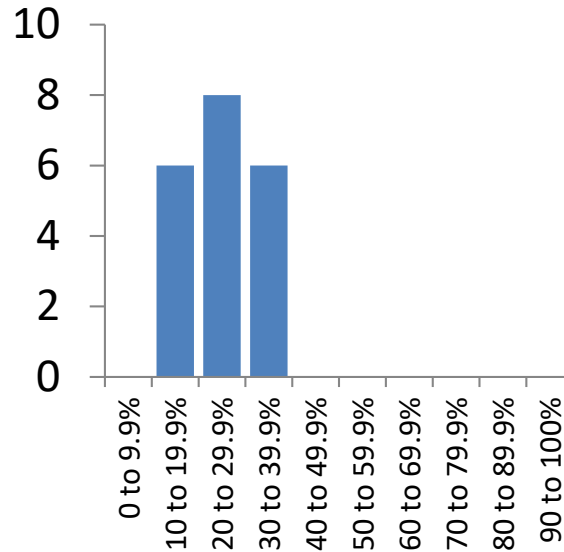
This is a *not* a dumb example ... we do not know the population value.

Sampling Error

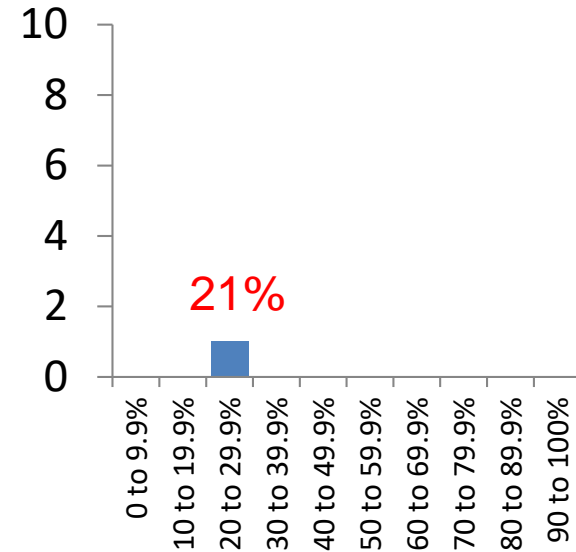
Question: What percentage of Americans say it is true that humans evolved from an earlier species?



Results from asking **10** randomly selected people; trial repeated 20 times



Results from asking **20** randomly selected people; trial repeated 20 times

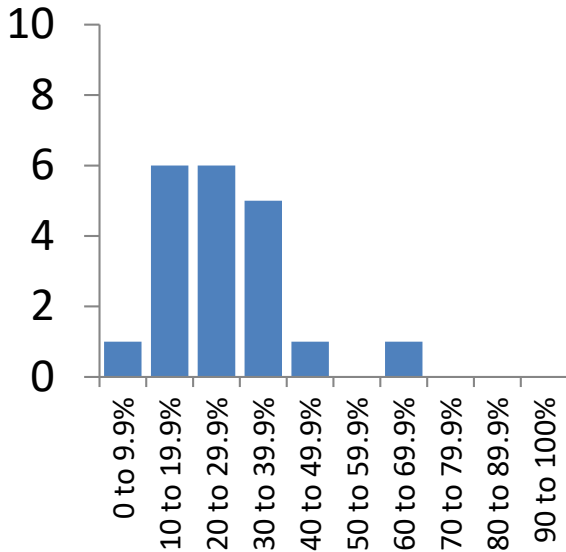


Results from asking **434** randomly selected people; trial repeated **once**

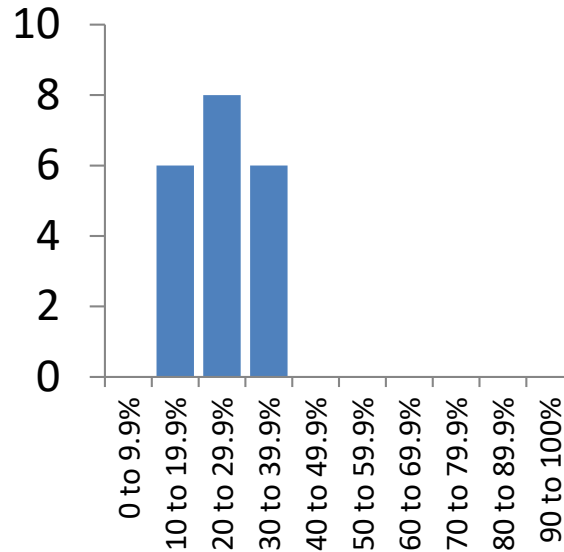
This is a *not* a dumb example ... we do not know the population value. However, if the sampling procedures are sound, the value is certainly close to 21%

Sampling Error

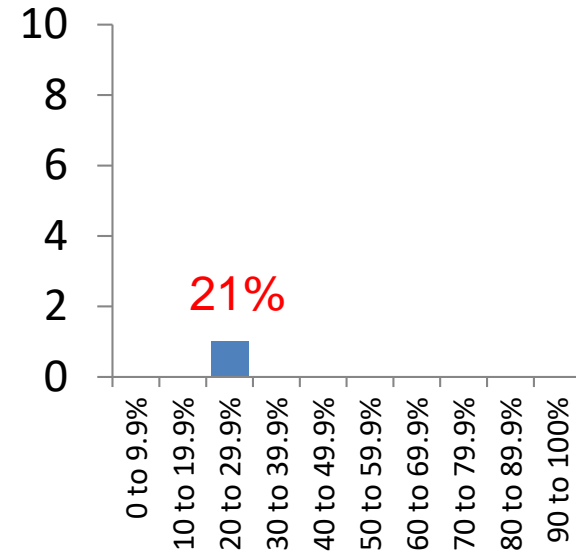
Question: What percentage of Americans say it is true that humans evolved from an earlier species?



Results from asking **10** randomly selected people; trial repeated 20 times



Results from asking **20** randomly selected people; trial repeated 20 times

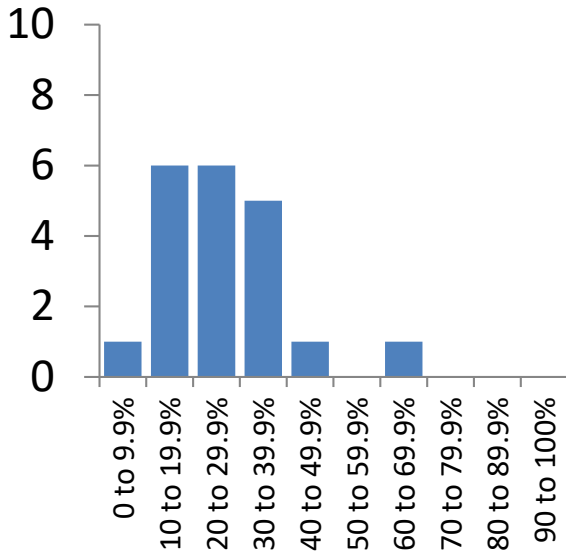


Results from asking **434** randomly selected people; trial repeated **once**

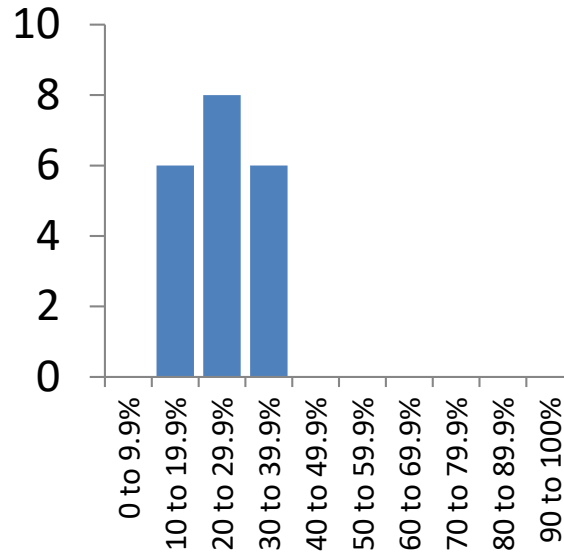
Again, in the left two figures sampling error goes down as sample size goes up

Sampling Error

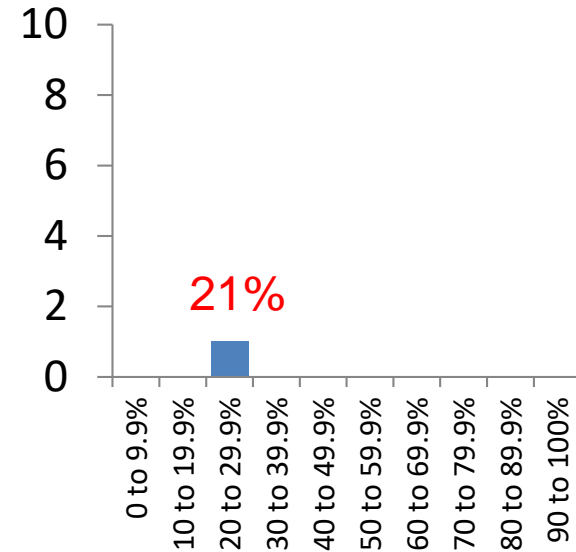
Question: What percentage of Americans say it is true that humans evolved from an earlier species?



Results from asking **10** randomly selected people; trial repeated 20 times



Results from asking **20** randomly selected people; trial repeated 20 times



Results from asking **434** randomly selected people; trial repeated **once**

How confident should we be in the 21% result, obtained from just one sample of size $n=434$?

Sampling Error

How much sampling error should we expect?

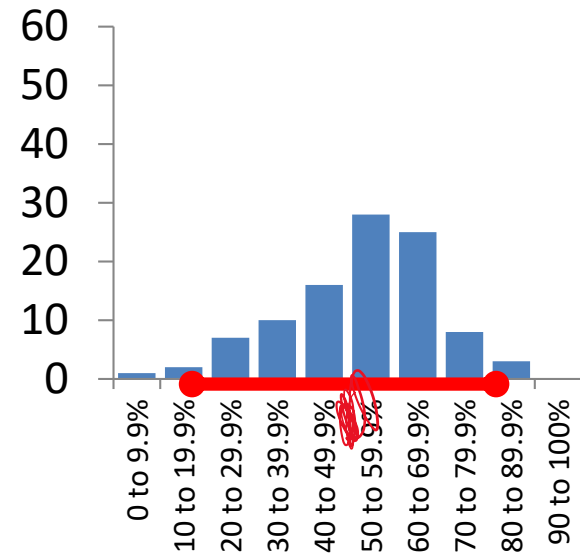
When we are talking about percentages, a good guess is given by the *conservative margin of error*

If the sample is representative of the population from which it was drawn, the sample percentage and the population percentage will differ by less than the conservative margin of error at least 95% of the time

$$\text{conservative margin of error} = \pm \frac{1}{\sqrt{n}} \times 100\%$$

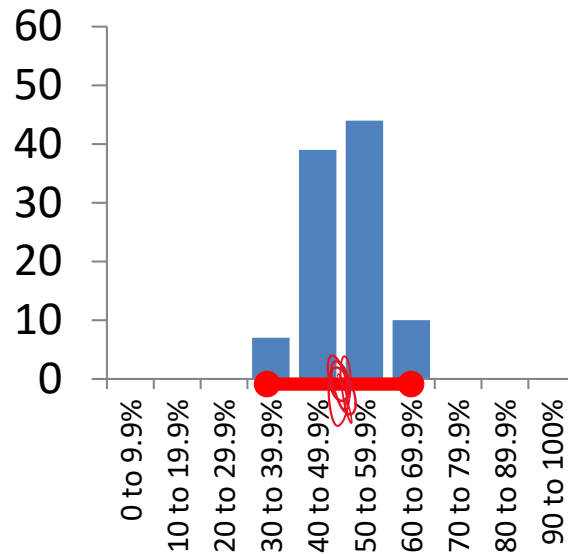
Sampling Error

Question: If you flip a fair coin, what percentage of the time will it come up “heads?”



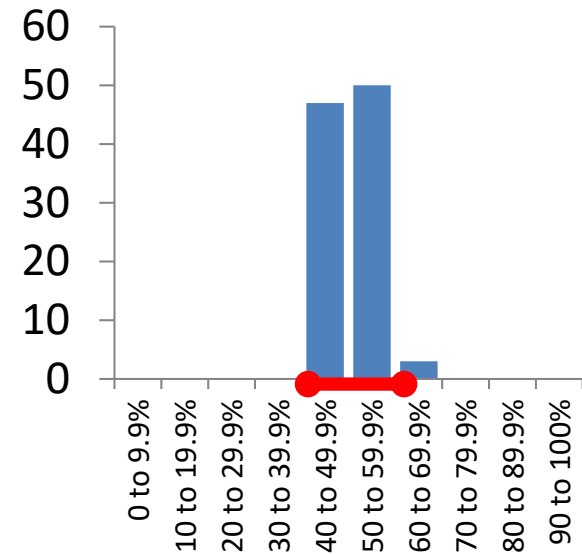
Results from flipping a fair coin **10** times; trial repeated 100 times

$$\pm \frac{1}{\sqrt{10}} \times 100\% = \pm 32\%$$



Results from flipping a fair coin **50** times; trial repeated 100 times

$$\pm \frac{1}{\sqrt{50}} \times 100\% = \pm 14\%$$

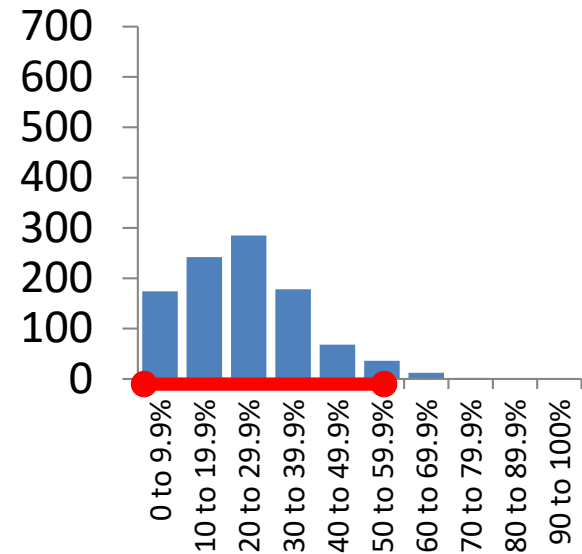


Results from flipping a fair coin **100** times; trial repeated 100 times

$$\pm \frac{1}{\sqrt{100}} \times 100\% = \pm 10\%$$

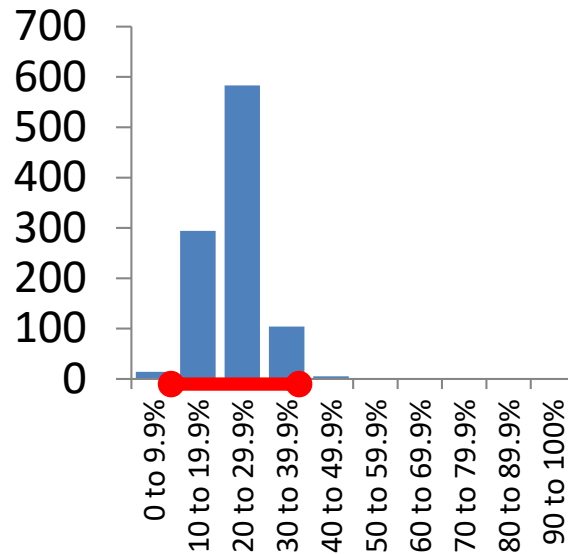
Sampling Error

Question: What percentage of Americans were high school graduates in 1940?



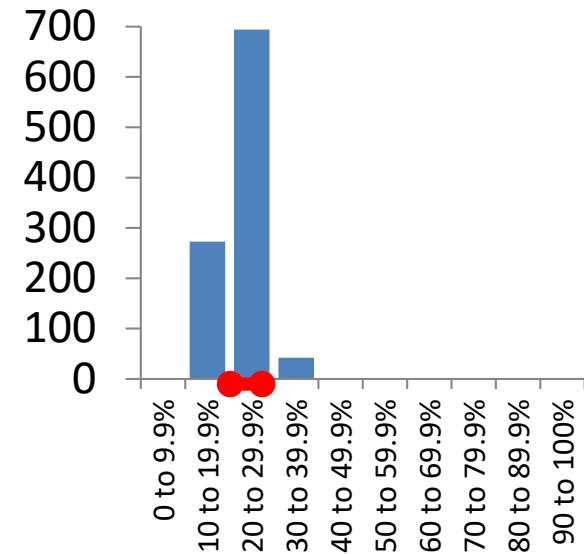
Results from asking **10** randomly selected people; trial repeated 1,000 times

$$\pm \frac{1}{\sqrt{10}} \times 100\% = \pm 32\%$$



Results from asking **50** randomly selected people; trial repeated 1,000 times

$$\pm \frac{1}{\sqrt{50}} \times 100\% = \pm 14\%$$

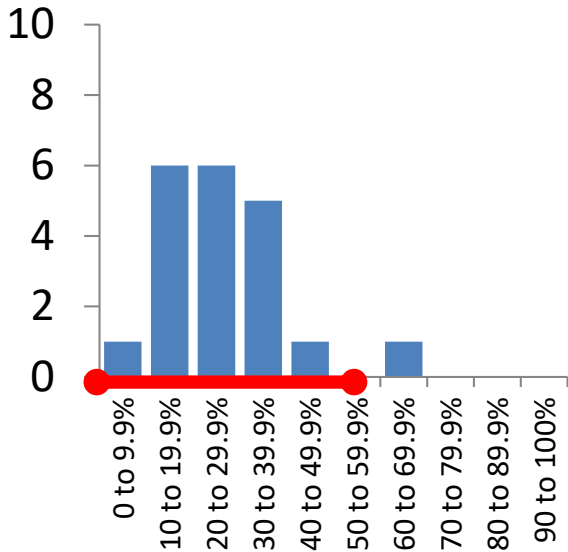


Results from asking **1,000** randomly selected people; trial repeated 1,000 times

$$\pm \frac{1}{\sqrt{1000}} \times 100\% = \pm 3\%$$

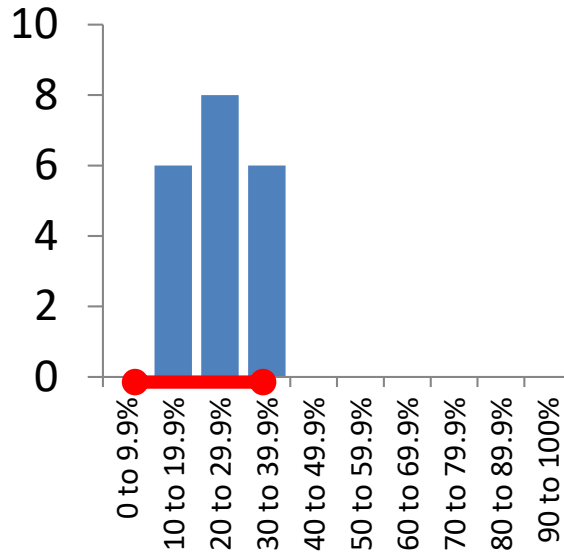
Sampling Error

Question: What percentage of Americans say it is true that humans evolved from an earlier species?



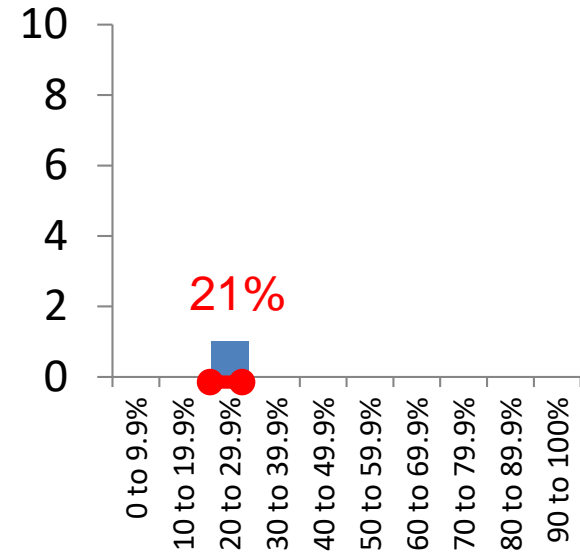
Results from asking **10** randomly selected people; trial repeated 20 times

$$\pm \frac{1}{\sqrt{10}} \times 100\% = \pm 32\%$$



Results from asking **20** randomly selected people; trial repeated 20 times

$$\pm \frac{1}{\sqrt{50}} \times 100\% = \pm 14\%$$

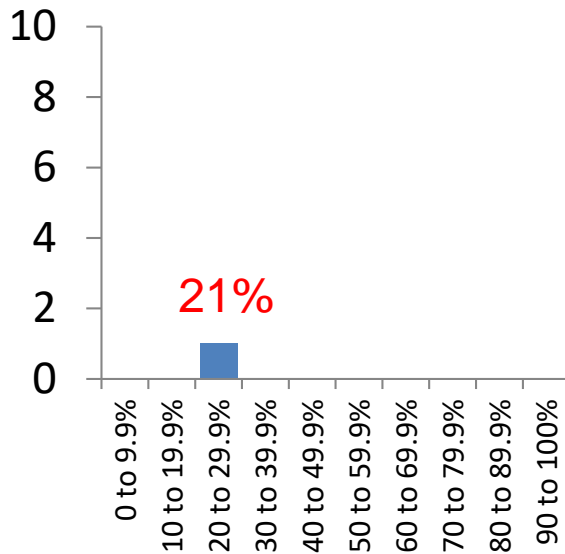


Results from asking **434** randomly selected people; trial repeated **once**

$$\pm \frac{1}{\sqrt{434}} \times 100\% = \pm 4.8\%$$

Sampling Error

Question: What percentage of Americans say it is true that humans evolved from an earlier species?



Results from asking **434** randomly selected people; trial repeated **once**

How confident should we be in the 21% result, obtained from just one sample of size $n=434$?

If the sample is really representative of the population, then we are 95% certain that the population value is $21\% \pm (1/\sqrt{434}) \times 100\% = 21\% \pm 4.8\%$

Sampling Error

$$\text{conservative margin of error} = \pm \frac{1}{\sqrt{n}} \times 100\%$$

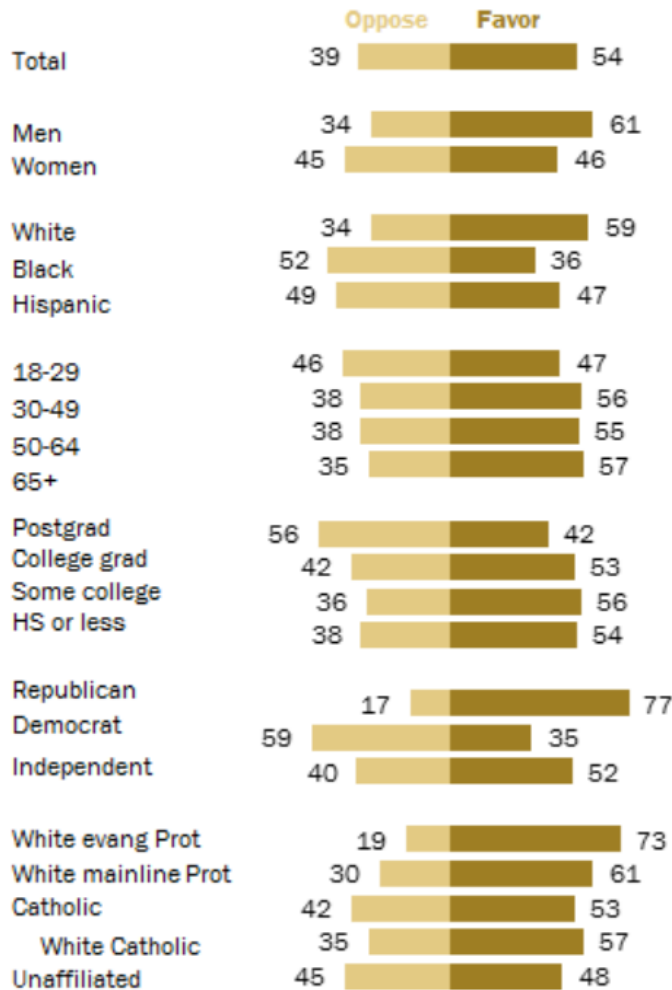
The conservative margin of error only (approximately) works for percentages

To quantify the margin of error, cases must be sampled at random from the population

The larger the sample size, the smaller the margin of error

Gender, racial differences in opinions about the death penalty

% who _____ the death penalty for persons convicted of murder



Note: Whites and blacks include only those who are not Hispanic; Hispanics are of any race. Don't know responses not shown.
Source: Survey of U.S. adults conducted April 25-May 1, 2018.

PEW RESEARCH CENTER

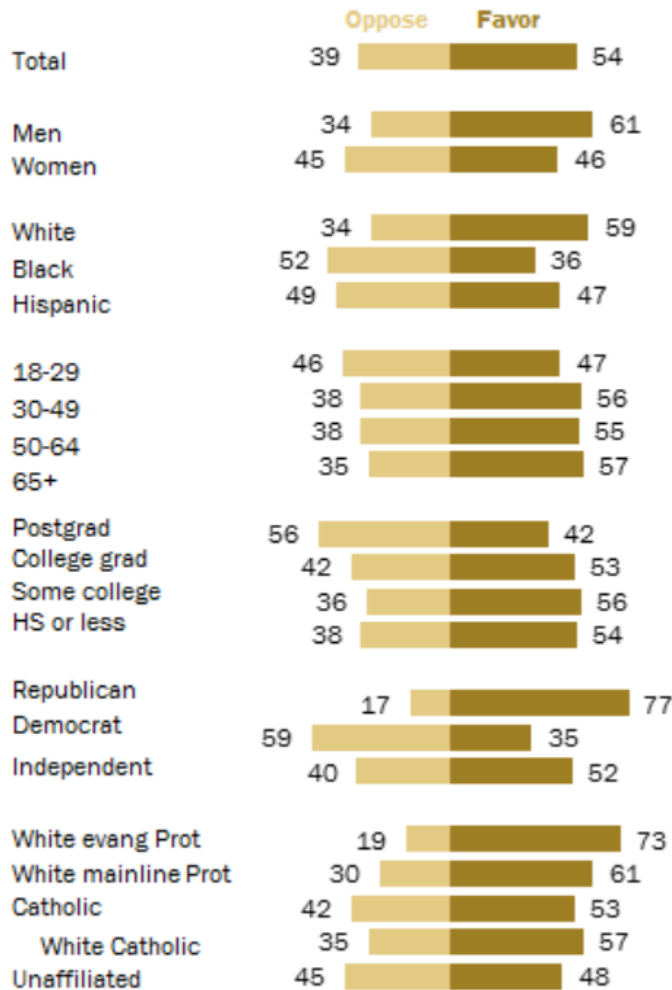
QUESTION

1. Construct a confidence interval for the percentage of all Americans who support the death penalty

“The analysis in this report is based on telephone interviews conducted April 25-May 1, 2018 among a national sample of 1,503 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia”

Gender, racial differences in opinions about the death penalty

% who ___ the death penalty for persons convicted of murder



Note: Whites and blacks include only those who are not Hispanic; Hispanics are of any race. Don't know responses not shown.
Source: Survey of U.S. adults conducted April 25-May 1, 2018.

PEW RESEARCH CENTER

QUESTION

1. Construct a confidence interval for the percentage of all Americans who support the death penalty

$$\text{Margin of error} = \pm \frac{1}{\sqrt{n}} \times 100\%$$

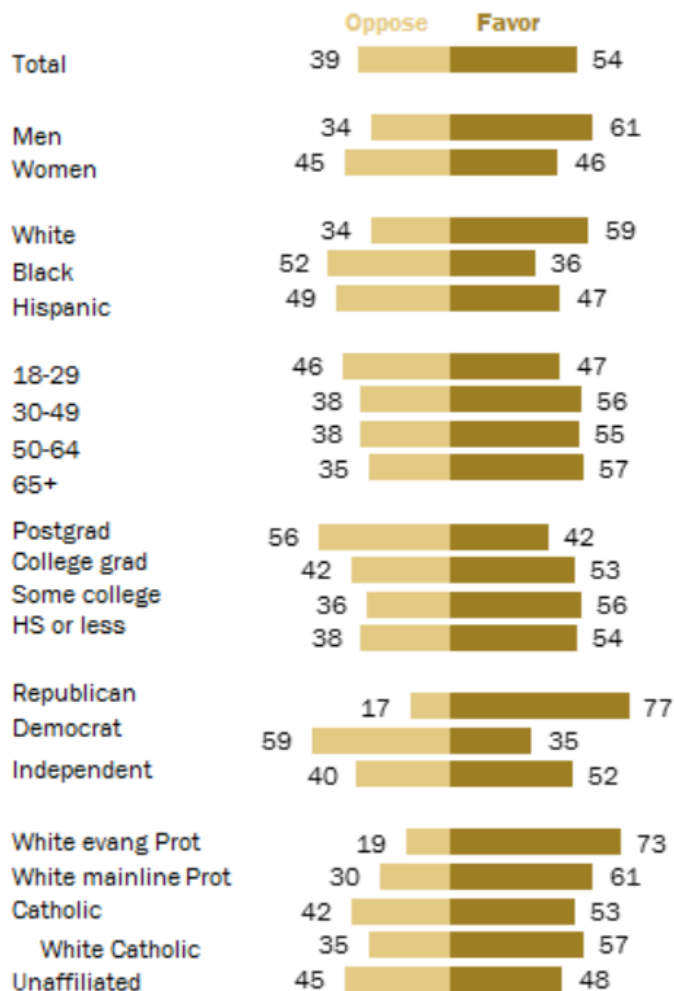
$$\text{Margin of error} = \pm \frac{1}{\sqrt{1,503}} \times 100\%$$

$$\text{Margin of error} = 2.58\%$$

“The analysis in this report is based on telephone interviews conducted April 25-May 1, 2018 among a national sample of 1,503 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia”

Gender, racial differences in opinions about the death penalty

% who ___ the death penalty for persons convicted of murder



Note: Whites and blacks include only those who are not Hispanic; Hispanics are of any race. Don't know responses not shown.
Source: Survey of U.S. adults conducted April 25-May 1, 2018.

PEW RESEARCH CENTER

QUESTION

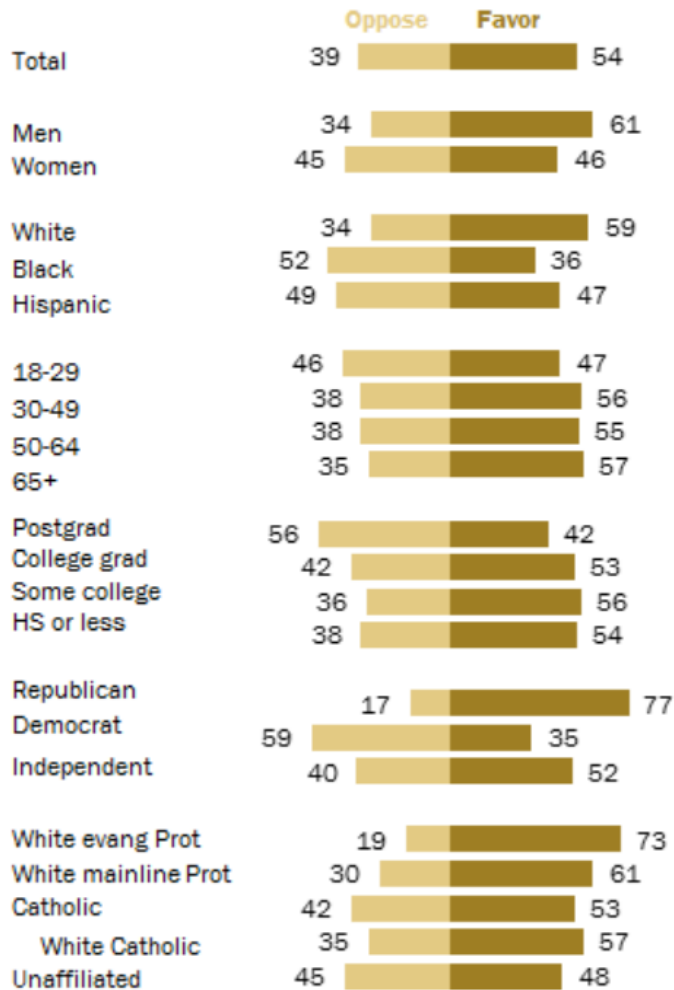
1. Construct a confidence interval for the percentage of all Americans who support the death penalty

2. How large would their sample have to have been for their confidence interval to be $\pm 2\%$?

“The analysis in this report is based on telephone interviews conducted April 25-May 1, 2018 among a national sample of 1,503 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia”

Gender, racial differences in opinions about the death penalty

% who _____ the death penalty for persons convicted of murder



Note: Whites and blacks include only those who are not Hispanic; Hispanics are of any race. Don't know responses not shown.
Source: Survey of U.S. adults conducted April 25-May 1, 2018.

PEW RESEARCH CENTER

QUESTION

1. Construct a confidence interval for the percentage of all Americans who support the death penalty

2. How large would their sample have to have been for their confidence interval to be $\pm 2\%$?

Margin of error = 1%

$$= \pm \frac{1}{\sqrt{n}} \times 100\%$$

$$0.01 = \pm \frac{1}{\sqrt{n}}$$

$$0.01 \times \sqrt{n} = \pm 1$$

$$\sqrt{n} = \pm \frac{1}{0.01}$$

$$\sqrt{n} = \pm 100$$

$$n = \pm 100^2$$

$$n = 10,000$$

“The analysis in this report is based on telephone interviews conducted April 25-May 1, 2018 among a national sample of 1,503 adults, 18 years of age or older, living in all 50 U.S. states and the District of Columbia”

Worksheet

FiveThirtyEight

Polling averages are adjusted based on state and national polls, which means candidates' averages can shift even if no new polls have been added to this page. [Read more about the methodology.](#)

FILTER BY POLLSTER GRADE



Search polls

Added today

	DATES	POLLSTER	SAMPLE	RESULT			NET RESULT
President: general election	Minn.	SEP 8-13, 2020	A+ ABC News/The Washington Post	615 LV	Biden 57%	41% Trump	Biden +16
	Minn.	SEP 8-13, 2020	A+ ABC News/The Washington Post	705 RV	Biden 57%	40% Trump	Biden +17

KEY A = ADULTS RV = REGISTERED VOTERS V = VOTERS LV = LIKELY VOTERS

1. Construct a confidence interval for the percentage of likely voters in Minnesota who would have voted for Trump as of the day this poll was taken
2. How large would their sample have to have been for their confidence interval to be $\pm 1\%$?