



2022 USAID/Jordan Annual MEL Conference Basic Sampling– Breakout Session

Tuesday May 24th, 2022

Presenters:

Majd Haddad – Operations & Field Manager

Anas Al Masri – Technical Advisor

USAID – MEL Conference – Basic Sampling Breakout Session



Sampling Introduction



Sampling Importance

- Everyone who has ever worked on a research project knows that resources are limited; time, money and people never come in an unlimited supply. For that reason, most research projects aim to gather data from a sample of people, rather than from the entire population (the census being one of the few exceptions). Sampling allows researchers to:
 - Save Time
 - Save Money
 - Collect Richer Data



Sampling Terms

Universe

Sample

Representative
sample

Sample
Frequency

Sample Frame

Weight

Sampling Error

Confidence
Level

Sampling Bias



Representative Sample

- The representative sample is part of the indigenous community
- It has some equal qualities
- The sample retained all the terms of population science and demographic concepts
- Has the same characteristics as the original community



Sampling Methodologies



Sampling Methodologies

Probability Sampling

- Probability samples rely on a specific sampling method based on some criteria set by the researcher, in which the chosen person (respondent) is randomly selected, and all the people within the same universe have an equal opportunity to be selected.

Non-probability Sampling

- Non-probability samples, this type of sample, the researcher selects people in non-static ways, and in fact, this makes it difficult that the samples to have equal opportunities, as happens in probability sampling.



Non-probability Sampling

Judgmental sampling

Quota sampling

Convenience sampling

Snowball sampling

Probability Sampling

Simple Random Sampling

Cluster Sampling

Stratification

Systematic Selection

Stratified Multistage Sampling

For Discussion

In an organization of 500 employees, if the HR team decides to conduct team-building activities, it is very likely that they will prefer to make a draw.

What is the opportunity for each employee to be selected?



Simple Random Sampling

A group of subjects is selected from a larger group. Each subject has an equal chance of being chosen to be part of the sample.



Advantages

- Generalizable results possible, random sampling, the sampling frame is the whole population, every participant has an equal probability of being selected

Disadvantages

- Less precise than stratified method, less representative than the systematic method

Simple Random Sampling



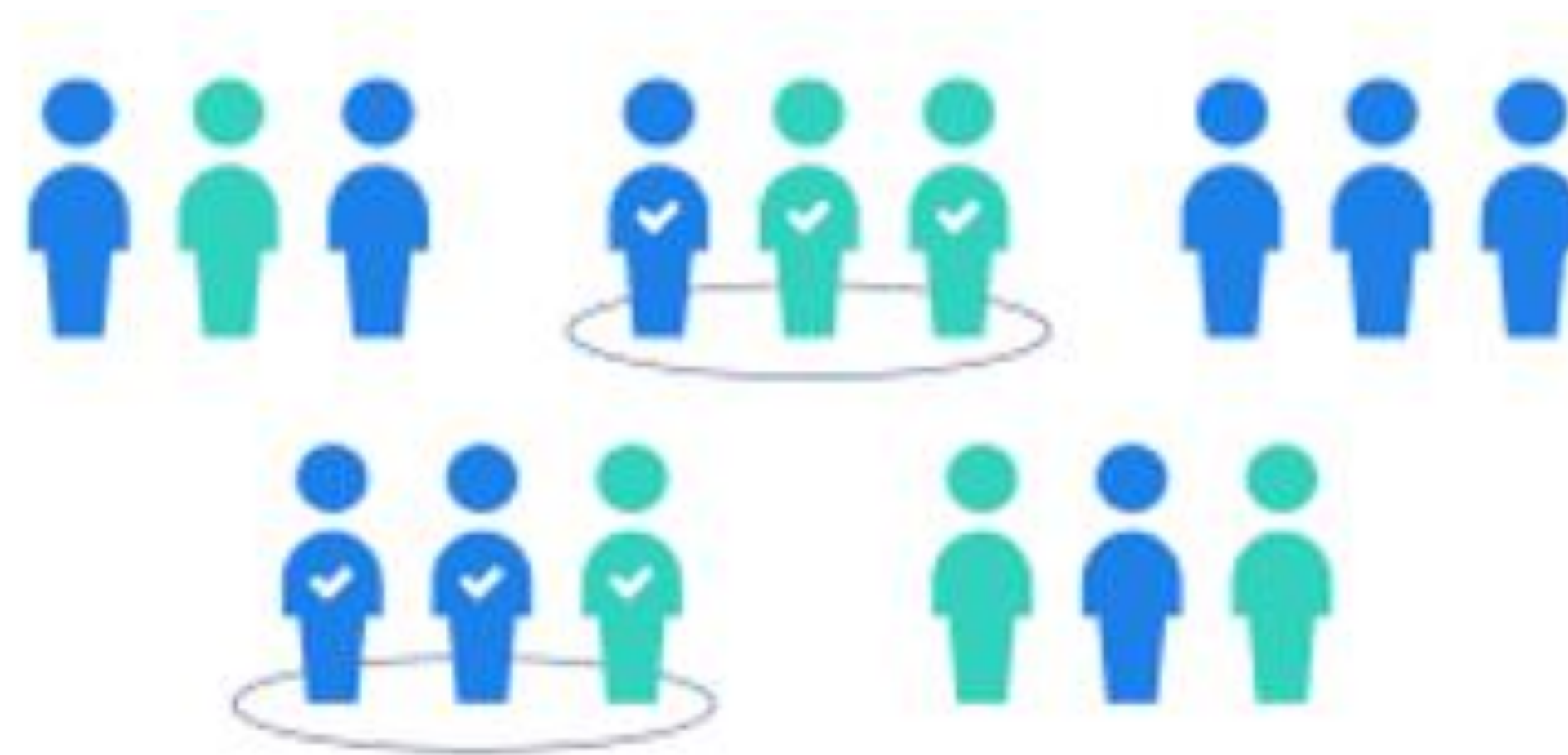
For Discussion

If the government wants to assess the number of immigrants living in it, the Government can divide the kingdom into groups based on governorates such as Amman, Zarqa, ... Etc..

Based on this the results would be read at what level?

Cluster Sampling

A larger population is divided into groups, otherwise known as clusters, commonly by geographic data. The groups are randomly selected to be part of the sample.



Advantages

- Readily doable with most budgets, does not require a sampling frame

Disadvantages

- Results may not be reliable nor generalizable

Cluster Sampling



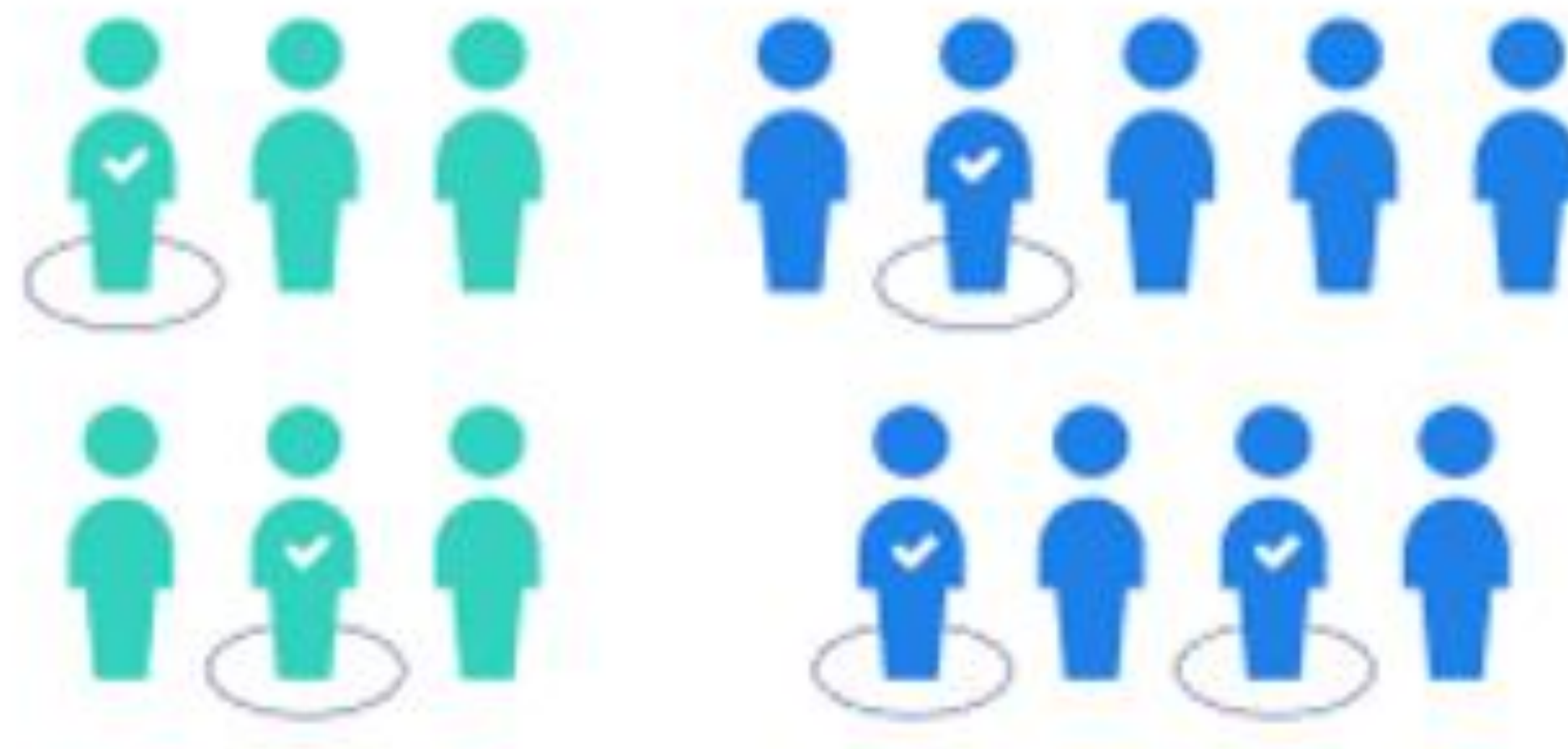
For Discussion

If we need to study each gender group separately.
If we need to study the educational level in each gender.

What's the difference!

Stratification

The population is first divided into strata; these strata are then further divided by diverse characteristics until a satisfactory sample demographics reached



Advantages

- Inclusive of strata (subgroups), reliable and generalizable results

Disadvantages

- Does not work well with multiple variables

Stratification



For Discussion

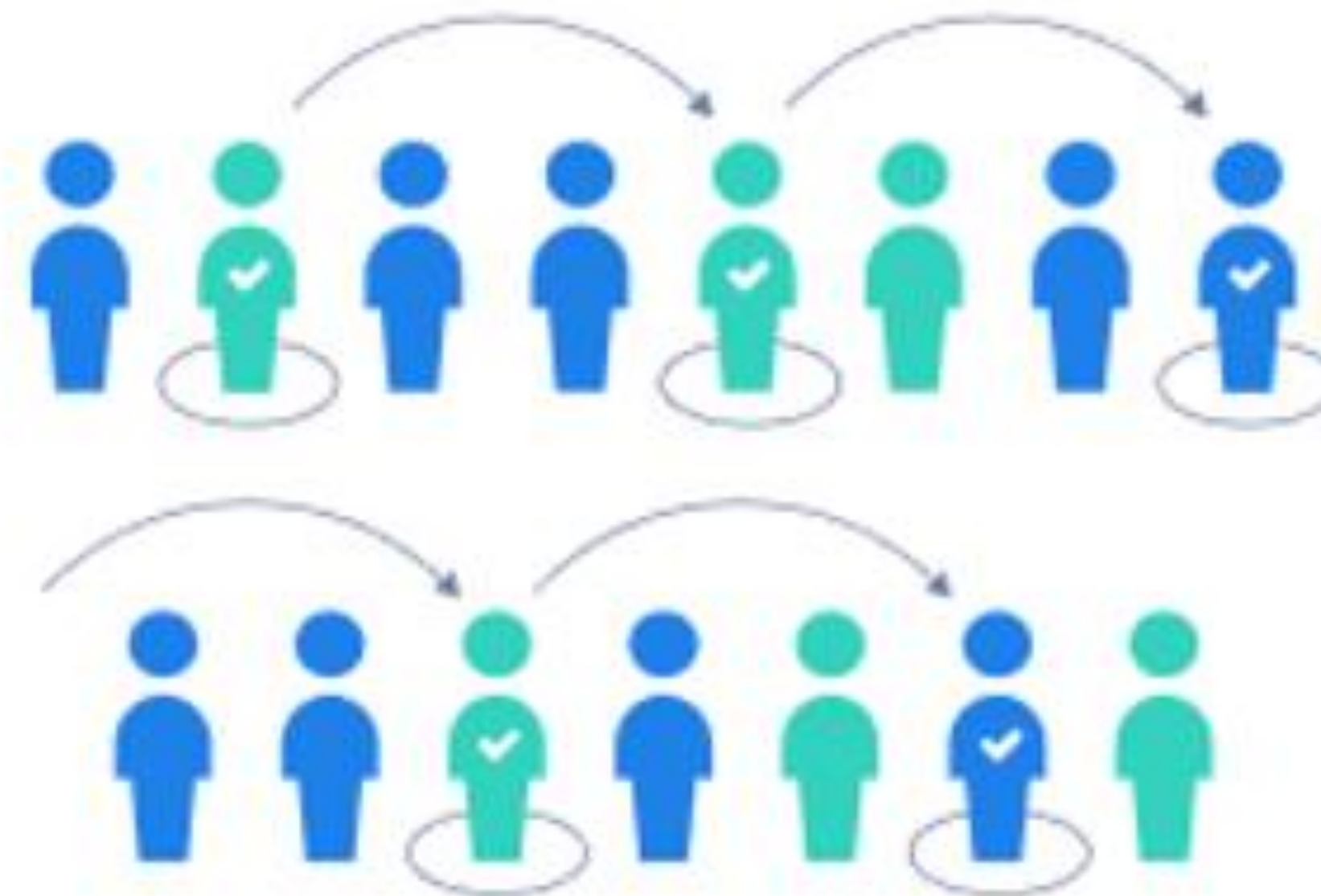
All employees of the company are listed in alphabetical order. From the first 10 numbers, you randomly select a starting point: number 6. From number 6 onwards, every 10th person on the list is selected (6, 16, 26, 36, and so on), and you end up with a sample of 100 people.

What is the interval (frequency) here?



Systematic Selection

Subjects are selected at regular periodic intervals most commonly by a numerical (kth) interval.



Advantages

- More feasible than simple or stratified methods, sampling frame is not always required

Disadvantages

- Generalizability may decrease if baseline characteristics repeat across every nth participant

Systematic Selection



For Discussion

When we say General population study

What is the 1st thing come to your mind?

What is the level of details we can get in the analysis?



Stratified Multistage Sampling

- Multistage sampling is often considered an extended version of cluster sampling.
- the population is divided into clusters and some clusters are selected at the first stage
- At each next stage, the selected clusters are divided into smaller clusters, and the process is repeated until reaching the last step
- At the last step, only some members of each cluster are selected for the sample.



Advantages

- You don't need to start with a sampling frame of your target population.
- Compared to a simple random sample, it's relatively inexpensive and effective when you have a large or geographically dispersed population.
- It's flexible—you can vary sampling methods between stages based on what's appropriate or feasible.

Disadvantages

- Compared to simple random samples, you'll need a larger sample size for a multistage sample to achieve the same statistical inference properties.
- The best choice of sampling method at each stage is very subjective, so you'll need clear reasoning for your decision.
- It can lead to unrepresentative samples because large sections of populations may not be selected for sampling.

Stratified Multistage Sampling



How do we choose the appropriate sampling approach?

“Four Considerations”



1) Research Objective

- Defining the population of interest through the research question
- Calculating the sample size from the population of interest

2) Sampling Frame Availability

- Do we have a **simple sample frame** or a list of the population of interest from which a sample can be taken
- **If not**, we may use **stratification** (the population is first divided into strata; these strata are then further divided by diverse characteristics until a satisfactory sample demographics reached)



3) Study Design

- Defining the population of interest through the research question
- Calculating the sample size from the population of interest

4) Random Sampling

- The best sampling method is the one that best answers the research question
- Random sampling allows for generalizability of results
- When random sampling is not feasible, non-random methods can be used



Using Confidence Intervals and Margin of Error



Confidence Interval

An educated guess about some characteristic of the population. A confidence interval contains an initial estimate plus or minus a margin of error

Margin of Error

The amount by which you expect your results to vary, if a different sample were taken.



Calculating the Confidence Interval

- Sample
- Mean
- Standard deviation (SD)
- Decide the confidence interval (most common are 99% and 95%)
- Find the Z value
- **Equation= Mean \pm Z-value (SD/ $\sqrt{\text{sample size}}$)**



Calculating the Margin of Error

- **Step 1:** Find the critical value (a z-score or a t-score)
- **Step 2:** Find the standard deviation or the standard error
- **Step 3:** Multiply the critical value from Step 1 by the standard deviation or standard error from Step 2.



Example: Calculate the confidence interval

$$\text{Equation} = \text{Mean} \pm \text{Z-value} (\text{SD} / \sqrt{\text{sample size}})$$

- Let's imagine a group of researchers that are interested in determining whether or not the oranges grown on a particular farm are large enough to be sold to a prospective grocery chain.
- The researchers randomly select **46** oranges from trees on the farm as their **sample**.
- The researchers then calculate of a **mean** weight of **86** grams from their sample.
- Calculated the **standard deviation = 6.2**
- Confidence level 95% which corresponds to a Z value of **Z = 1.960**

Example: Calculate the margin of error

$$\text{Equation} = \text{Critical value} * \text{Standard Deviation}$$

- 900 students were surveyed and had an average GPA of 2.7 with a **standard error** is 0.013. Calculate the margin of error for a 90% confidence level.
- The **critical value** is 1.645 (see this video for the calculation)

Revisiting our Learning Questions

What is the importance of rigorous sampling?

What are the different sampling methods?

Selecting the appropriate sampling method.

Conclusion



Sampling is necessary for data collection due to its efficiency in representing a larger population.

Probability sampling, while more complex than non-probability sampling, is more utilized in sampling due to its accuracy.

Probability sampling methods include simple random sampling, cluster sampling, stratification, systematic selection, and stratified multistage sampling.

The nature of sampling includes a chance of uncertainty; therefore, confidence intervals and the margin of error must be calculated.

Sampling will empower MEL specialists and experts by providing effective tools for working experts to gauge the needs of the Jordanian population



References

- “Sampling Strategies” Kandace J. Landreneau, RN, PhD, CCTC, PostDoctoral Research Fellow, University of California-San Francisco, Walnut Creek, CA, Research Committee Member
- Stewart, A., 2016. Basic statistics and epidemiology: a practical guide. 4th Ed. Oxford: CRC Press.
- How to choose your sampling strategy to guarantee relevant results. Available at: ([link](#))
- Statistics How To
- Understanding Confidence Intervals | Easy Examples & Formulas Published on August 7, 2020 by Rebecca Bevans. Scribbr Available at: ([link](#))
- Multistage Sampling | An Introductory Guide with Examples. Published on August 16, 2021 by Pritha Bhandari. Scribbr. Available at: ([link](#))
- How to Calculate Confidence Intervals. Alchemer Available at: ([link](#))
- Your guide to margin of error. Qualtrics. Available at: ([link](#))



Thanks!

- The Mindset Team