

How do I perform stratified sampling in Excel using a step-by-step approach?

Authored by
stats writer

May 6, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I perform stratified sampling in Excel using a step-by-step approach?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143213>

Stratified sampling is a statistical technique used to ensure that a sample accurately represents the population it is drawn from. This method involves dividing the population into smaller, more homogeneous groups or strata, and then selecting a sample from each stratum. In Excel, you can perform stratified sampling using a step-by-step approach by following these simple steps:

Step 1: Prepare your data

Arrange your data in columns, with each column representing a different variable. Make sure to label each column and include all relevant data.

Step 2: Identify your strata

Determine the different groups or categories within your data that you want to use as strata. These could be based on demographic factors, geographic location, or any other relevant criteria.

Step 3: Calculate the population size

Use the COUNT function in Excel to determine the total number of units in your population.

Step 4: Calculate the sample size

Based on the desired level of precision and confidence, calculate the sample size for each stratum using the appropriate formula.

Step 5: Create a random number generator

Using the RANDBETWEEN function in Excel, generate a random number for each unit in your population.

Step 6: Sort the data

Sort your data based on the random numbers generated in the previous step. This will ensure that your sample is selected randomly from each stratum.

Step 7: Select your sample

Using the sample size calculated in Step 4, select the required number of units from each stratum to create your sample.

Step 8: Analyze your sample

Use the data from your sample to perform any necessary analysis or draw conclusions about the entire population.

By following this step-by-step approach, you can easily perform stratified sampling in Excel and obtain a representative sample from your population. This technique can be useful in various fields such as market research, social sciences, and quality control.

Perform Stratified Sampling in Excel (Step-by-Step)

In statistics, we often take samples from a population and use the data from the sample to draw conclusions about the population as a whole.

One commonly used sampling method is stratified random sampling, in which a population is split into groups and a certain number of members from each group are randomly selected to be included in the sample.

The following step-by-step example shows how to perform stratified random sampling in Excel.

Step 1: Enter the Data

First, let's enter the following dataset into Excel:

	A	B	C	D	E	F	G
1	Player ID	Team	Points	Rebounds			
2	1	A	12	5			
3	2	A	16	6			
4	3	A	17	6			
5	4	A	17	3			
6	5	A	19	8			
7	6	B	22	9			
8	7	B	25	5			
9	8	B	19	4			
10	9	B	13	7			
11	10	B	9	8			
12	11	C	12	8			
13	12	C	7	12			
14	13	C	17	15			
15	14	C	30	7			
16	15	C	12	6			
17							
18							
19							
20							
21							
22							
23							

Next, we'll perform stratified random sampling in which we randomly select two players from each basketball team to be included in the final sample.

Step 2: Enter Random Values for Each Row

Next, let's create a new column titled Random and type in `=RAND()` for the first value:

E2					=RAND()		
	A	B	C	D	E	F	G
1	Player ID	Team	Points	Rebounds	Random		
2	1	A	12	5	0.235311		
3	2	A	16	6			
4	3	A	17	6			
5	4	A	17	3			
6	5	A	19	8			
7	6	B	22	9			
8	7	B	25	5			
9	8	B	19	4			
10	9	B	13	7			
11	10	B	9	8			
12	11	C	12	8			
13	12	C	7	12			
14	13	C	17	15			
15	14	C	30	7			
16	15	C	12	6			
17							
18							
19							
20							
21							
22							
23							

This generates a random value between 0 and 1.

Next, hover over the bottom right corner of the cell until a tiny cross (+) appears and double click it to paste the =RAND() formula to all remaining cells in the column.

E2					0.720302797390689		
	A	B	C	D	E	F	G
1	Player ID	Team	Points	Rebounds	Random		
2	1	A	12	5	0.720303		
3	2	A	16	6	0.699921		
4	3	A	17	6	0.437896		
5	4	A	17	3	0.644913		
6	5	A	19	8	0.48128		
7	6	B	22	9	0.242952		
8	7	B	25	5	0.987388		
9	8	B	19	4	0.419407		
10	9	B	13	7	0.202296		
11	10	B	9	8	0.983954		
12	11	C	12	8	0.504556		
13	12	C	7	12	0.934046		
14	13	C	17	15	0.95429		
15	14	C	30	7	0.640191		
16	15	C	12	6	0.401028		
17							
18							
19							
20							
21							
22							
23							

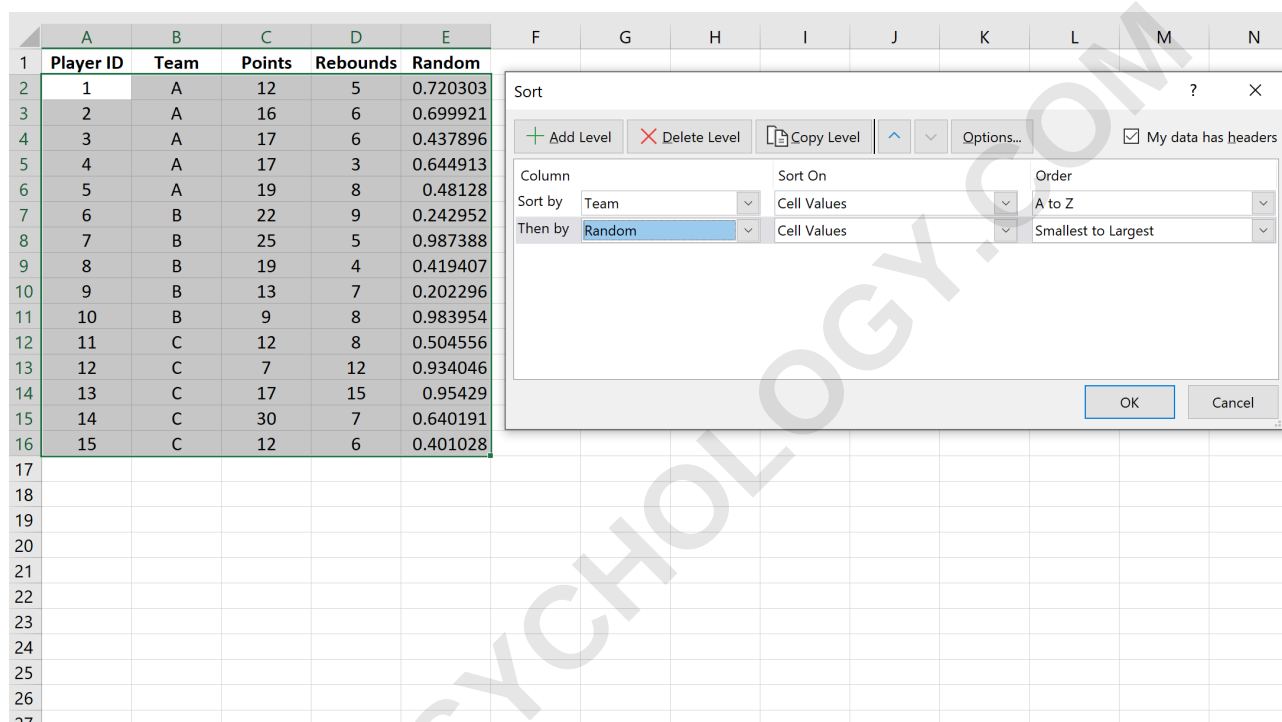
Unfortunately, each time we hit Enter the random cell values will change. To prevent this, copy every value in column E then right click and choose Paste Values into the same column so that the random values will no longer change.

Step 3: Sort Data Values

Next, highlight all of the data. Then click the Data tab along the top ribbon. Then click the Sort button within

the Sort & Filter group.

In the new window that appears, sort first by Team from A to Z, then sort by Random from Smallest to Largest.



	A	B	C	D	E
1	Player ID	Team	Points	Rebounds	Random
2	1	A	12	5	0.720303
3	2	A	16	6	0.699921
4	3	A	17	6	0.437896
5	4	A	17	3	0.644913
6	5	A	19	8	0.48128
7	6	B	22	9	0.242952
8	7	B	25	5	0.987388
9	8	B	19	4	0.419407
10	9	B	13	7	0.202296
11	10	B	9	8	0.983954
12	11	C	12	8	0.504556
13	12	C	7	12	0.934046
14	13	C	17	15	0.95429
15	14	C	30	7	0.640191
16	15	C	12	6	0.401028

Sort dialog box settings:

- Sort by: Team (Cell Values) Order: A to Z
- Then by: Random (Cell Values) Order: Smallest to Largest
- My data has headers: ☒

Once you click OK, the data will be sorted accordingly.

Step 4: Select the Final Sample

The final sample will simply be the first two rows from each team:

	A	B	C	D	E	F	G
1	Player ID	Team	Points	Rebounds	Random		
2	3	A	17	6	0.437896		
3	5	A	19	8	0.48128		
4	4	A	17	3	0.644913		
5	2	A	16	6	0.699921		
6	1	A	12	5	0.720303		
7	9	B	13	7	0.202296		
8	6	B	22	9	0.242952		
9	8	B	19	4	0.419407		
10	10	B	9	8	0.983954		
11	7	B	25	5	0.987388		
12	15	C	12	6	0.401028		
13	11	C	12	8	0.504556		
14	14	C	30	7	0.640191		
15	12	C	7	12	0.934046		
16	13	C	17	15	0.95429		
17							
18							
19							
20							
21							
22							
23							

The final sample will include the following player ID's from each team:

Team A: 3, 5 Team B: 9, 6 Team C: 15, 11

Our stratified random sampling is complete because we have now chosen two players from each team.

The following tutorials explain how to select other types of samples from a population using Excel:

How to Perform Cluster Sampling in Excel

ARABPSYCHOLOGY.COM