

How do I perform systematic sampling in Excel?

Authored by
stats writer

May 6, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I perform systematic sampling in Excel?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143203>

Systematic sampling is a statistical sampling method used to select a representative sample from a larger population. This method involves selecting every n th item from the population, where n is a predetermined interval. In Excel, you can perform systematic sampling by using the RAND() function to generate a list of random numbers and then sorting the list in ascending order. Next, you can use the INDEX() function to select every n th number from the sorted list, representing the items to be included in the sample. This process can be repeated multiple times to ensure a more accurate representation of the population. By following these steps, you can easily perform systematic sampling in Excel to gather data for analysis and decision making.

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

Researchers 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 systematic sampling, which is implemented with a simple two step process:

1. Place each member of a population in some order.
2. Choose a random starting point and select every n th member to be in the sample.

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

Step 1: Enter the Data

First, let's enter the values for some dataset in Excel:

	A	B	C	D	E	F	G	H	
1	PlayerID	Points	Rebounds						
2	1	15	6						
3	2	16	7						
4	3	16	7						
5	4	15	6						
6	5	14	5						
7	6	19	4						
8	7	22	5						
9	8	24	8						
10	9	26	9						
11	10	19	11						
12	11	11	4						
13	12	9	12						
14	13	8	5						
15	14	12	13						
16	15	15	13						
17	16	15	9						
18	17	21	4						
19	18	31	4						
20	19	34	3						
21	20	29	7						
22									
23									
24									
25									

Step 2: Enter Parameter Values

Next, we need to decide how many values we'd like to randomly sample. For this example, we'll choose a sample size of $n = 4$.

Next, we'll use the following formulas to calculate the population size, step size for our systematic sampling, and random starting point:

	A	B	C	D	E	F	G	H
1	PlayerID	Points	Rebounds			n	4	
2	1	15	6			N	20	=COUNT(A2:A21)
3	2	16	7			Step Size	5	=ROUNDDOWN(G2/G1, 0)
4	3	16	7			Start	3	=RANDBETWEEN(1, 20)
5	4	15	6					
6	5	14	5					
7	6	19	4					
8	7	22	5					
9	8	24	8					
10	9	26	9					
11	10	19	11					
12	11	11	4					
13	12	9	12					
14	13	8	5					
15	14	12	13					
16	15	15	13					
17	16	15	9					
18	17	21	4					
19	18	31	4					
20	19	34	3					
21	20	29	7					
22								
23								
24								
25								
26								

Step 3: Label Each Data Value

Our RANDBETWEEN() function randomly chose the value 3 as the starting point in our dataset. Thus, give the data value at position 3 a label of 1 and then number down to the step size of 5:

	A	B	C	D	E	F	G	H
1	PlayerID	Points	Rebounds	Labels		n	4	
2	1	15	6			N	20	=COUNT(A2:A21)
3	2	16	7			Step Size	5	=ROUNDDOWN(G2/G1, 0)
4	3	16	7	1		Start	3	=RANDBETWEEN(1, 20)
5	4	15	6	2				
6	5	14	5	3				
7	6	19	4	4				
8	7	22	5	5				
9	8	24	8					
10	9	26	9					
11	10	19	11					
12	11	11	4					
13	12	9	12					
14	13	8	5					
15	14	12	13					
16	15	15	13					
17	16	15	9					
18	17	21	4					
19	18	31	4					
20	19	34	3					
21	20	29	7					
22								
23								
24								
25								

Next, type =D4 in the next available cell:

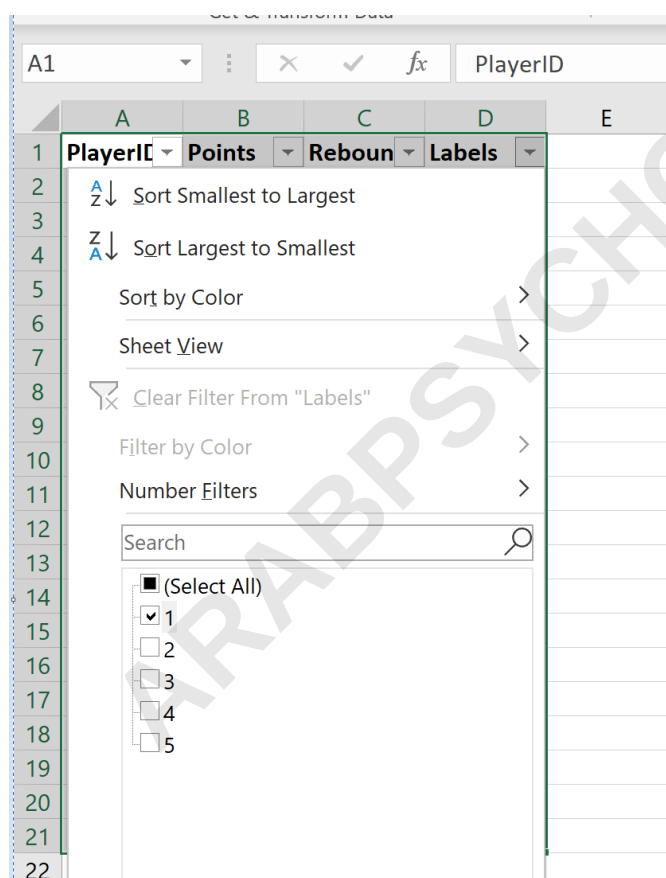
	A	B	C	D	E	F	G	H
1	PlayerID	Points	Rebounds	Labels		n	4	
2	1	15	6			N	20	=COUNT(A2:A21)
3	2	16	7			Step Size	5	=ROUNDDOWN(G2/G1, 0)
4	3	16	7	1		Start	3	=RANDBETWEEN(1, 20)
5	4	15	6	2				
6	5	14	5	3				
7	6	19	4	4				
8	7	22	5	5				
9	8	24	8	=D4				
10	9	26	9					
11	10	19	11					
12	11	11	4					
13	12	9	12					
14	13	8	5					
15	14	12	13					
16	15	15	13					
17	16	15	9					
18	17	21	4					
19	18	31	4					
20	19	34	3					
21	20	29	7					
22								

	A	B	C	D	E	F	G	H
1	PlayerID	Points	Rebounds	Labels		n	4	
2	1	15	6	4		N	20	=COUNT(A2:A21)
3	2	16	7	5		Step Size	5	=ROUNDDOWN(G2/G1, 0)
4	3	16	7	1		Start	3	=RANDBETWEEN(1, 20)
5	4	15	6	2				
6	5	14	5	3				
7	6	19	4	4				
8	7	22	5	5				
9	8	24	8	1				
10	9	26	9	2				
11	10	19	11	3				
12	11	11	4	4				
13	12	9	12	5				
14	13	8	5	1				
15	14	12	13	2				
16	15	15	13	3				
17	16	15	9	4				
18	17	21	4	5				
19	18	31	4	1				
20	19	34	3	2				
21	20	29	7	3				
22								
23								
24								

Step 4: Filter Values

Lastly, we need to filter the dataset to only include data values that have a label of 1.

To do so, click the Data tab along the top ribbon. Then click the Filter icon. Then click the column called Labels and filter to only include rows that have a label with value 1:



Once you do so, the data will be filtered to only include data values that have a Label of 1:

	A	B	C	D	E
1	PlayerID	Points	Rebound	Labels	
4	3	16	7	1	
9	8	24	8	1	
14	13	8	5	1	
19	18	31	4	1	
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					

This represents our sample of size 4.