

# How can I implement systematic sampling in R? Can you provide examples?

Authored by  
**stats writer**

April 18, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I implement systematic sampling in R? Can you provide examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136762>

Systematic sampling is a type of probability sampling method that involves selecting a random starting point and then systematically choosing every  $n$ th element from a population. This method ensures that every element in the population has an equal chance of being selected. In order to implement systematic sampling in R, the following steps can be followed:

1. Import the necessary packages: First, you need to import the "sampling" package in R, which contains the function for systematic sampling.
2. Define the population: Next, you need to define the population from which you want to take a sample.
3. Determine the sampling interval: Calculate the sampling interval by dividing the population size by the desired sample size. This will give you the number of elements to skip between each selection.
4. Randomly select the starting point: Use the "sample" function in R to select a random starting point within the population.
5. Use the "seq" function: The "seq" function in R allows you to create a sequence of numbers at a specified interval. Use this function to select every  $n$ th element from the population, where  $n$  is the sampling interval calculated in step 3.
6. Store the selected sample: Finally, store the selected sample in a new variable for further analysis.

Example:

Let's say we have a population of 1000 students and we want to select a sample of 100 students using systematic sampling.

1. Import the "sampling" package: `library(sampling)`
2. Define the population: `population <- 1:1000`
3. Determine the sampling interval: `sampling_interval <- length(population)/100 = 10`
4. Randomly select the starting point: `starting_point <- sample(population, 1)`
5. Use the "seq" function: `sample <- seq(starting_point, length = 100, by = sampling_interval)`
6. Store the selected sample: `selected_sample <- population`

Now, the variable "selected\_sample" contains a sample of 100 students selected using systematic sampling from the population of 1000 students. This sample can be used for further analysis.

## Systematic Sampling in R (With Examples)

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.

This tutorial explains how to perform systematic sampling in R.

**Example: Systematic Sampling in R**

Suppose a superintendent wants to obtain a sample of 100 students from a school that has 500 total students. She chooses to use systematic sampling in which she places each student in alphabetical order according to their last name, randomly chooses a starting point, and picks every 5th student to be in the sample.

The following code shows how to create a fake data frame to work with in R:

```
#make this example reproducible  
set.seed(1)
```

```
#create simple function to generate random last names  
randomNames <- function(n = 5000) {  
do.call(paste0, replicate(5, sample(LETTERS, n, TRUE),  
FALSE))  
}
```

```
#create data frame  
df <- data.frame(last_name = randomNames(500),  
gpa = rnorm(500, mean=82, sd=3))
```

```
#view first six rows of data frame  
head(df)
```

```
last_name gpa  
1 GONBW 82.19580  
2 JRRWZ 85.10598  
3 ORJFW 88.78065  
4 XRYNL 85.94409  
5 FMDCE 79.38993
```

## 6 XZBJC 80.49061

And the following code shows how to obtain a sample of 100 students through systematic sampling:

```
#define function to obtain systematic sample
```

```
obtain_sys = function(N,n){
```

```
k = ceiling(N/n)
```

```
r = sample(1:k, 1)
```

```
seq(r, r + k*(n-1), k)
```

```
}
```

```
#obtain systematic sample
```

```
sys_sample_df = df
```

```
#view first six rows of data frame
```

```
head(sys_sample_df)
```

```
last_name gpa
```

```
3 ORJFW 88.78065
```

```
8 RWPSB 81.96988
```

```
13 RACZU 79.21433
```

```
18 ZOHKA 80.47246
```

```
23 QJETK 87.09991
```

```
28 JTHWB 83.87300
```

```
#view dimensions of data frame  
dim(sys_sample_df)
```

```
100 2
```

**Notice that the first member included in the sample was in row 3 of the original data frame. Each subsequent member in the sample is located 5 rows after the previous member.**

**And from using dim() we can see that the systematic sample we obtained is a data frame with 100 rows and 2 columns.**

## **Types of Sampling Methods**

### **Stratified Sampling in R**

### **Cluster Sampling in R**