# “How to Select a Random Sample in SAS? Can you provide some examples?”

Authored by

**stats writer**

July 1, 2024

Selecting a random sample in SAS is a statistical method used to obtain a representative subset of data from a larger population. This process involves randomly selecting data points from the population, ensuring that each data point has an equal chance of being selected. This can be achieved in SAS through various sampling techniques such as simple random sampling, stratified sampling, and cluster sampling. These techniques can be implemented using SAS programming language and functions, such as the RAND function and PROC SURVEYSELECT. For example, to select a random sample of 100 observations from a dataset in SAS, one could use the code: PROC SURVEYSELECT DATA = dataset OUT = random_sample METHOD = SRS SIZE = 100; RUN; This will create a new dataset called "random_sample" with a random selection of 100 observations from the original dataset. Other sampling techniques and options can be specified within the PROC SURVEYSELECT statement to customize the random sample selection process.

## Select a Random Sample in SAS (With Examples)

## Here are the two most common ways to select a of rows from a dataset in SAS:

## Method 1: Select Random Sample Using Sample Size

## proc surveyselectdata=original_data out=random_sample method=srs /*specify simple random sampling as sampling method*/sampsize=3 /*select 3 observations randomly*/seed=123; /*set seed to make this example reproducible*/run;

## Method 2: Select Random Sample Using Proportion of Total Observations

```
proc surveyselectdata=original_data
out=random_sample
method=srs /*specify simple random sampling as sampling method*/samprate=0.2 /*select 20% of all observations randomly*/seed=123; /*set seed to make this example reproducible*/run;
```

The following examples show how to use each method with the following dataset in SAS:

```
/*create dataset*/
data original_data;
input team $ points rebounds;
datalines;
Warriors 25 8
Wizards 18 12
Rockets 22 6
Celtics 24 11
Thunder 27 14
Spurs 33 19
Nets 31 20
Mavericks 34 10
Kings 22 11
Pelicans 39 23
```

```
;
run;

/*view dataset*/proc printdata=original_data;
```

| Obs | team | points | rebounds |
|---|---|---|---|
| 1 | Warriors | 25 | 8 |
| 2 | Wizards | 18 | 12 |
| 3 | Rockets | 22 | 6 |
| 4 | Celtics | 24 | 11 |
| 5 | Thunder | 27 | 14 |
| 6 | Spurs | 33 | 19 |
| 7 | Nets | 31 | 20 |
| 8 | Maverick | 34 | 10 |
| 9 | Kings | 22 | 11 |
| 10 | Pelicans | 39 | 23 |



### Example 1: Select Random Sample Using Sample Size

The following code shows how to select a random sample of observations from the dataset using a sample size of n=3:

```
/*select random sample*/proc surveyselectdata=original_data
out=random_sample
method=srs
sampsize=3
```

```
seed=123;
run;

/*view random sample*/
proc printdata=random_sample;
```

| Obs | team | points | rebounds |
|---|---|---|---|
| 1 | Warriors | 25 | 8 |
| 2 | Thunder | 27 | 14 |
| 3 | Pelicans | 39 | 23 |

**We can see that three rows were randomly selected from the original dataset.**

### Example 2: Select Random Sample Using Proportion of Total Observations

**The following code shows how to select a random sample of observations from the dataset by using the samprate function to specify that we'd like the random sample to represent 20% of all original observations:**

```
/*select random sample*/
proc surveyselectdata=original_data
out=random_sample
method=srs
samprate=0.2
```

```
seed=123;
run;

/*view random sample*/
proc printdata=random_sample;
```

| Obs | team | points | rebounds |
|---|---|---|---|
| 1 | Warriors | 25 | 8 |
| 2 | Spurs | 33 | 19 |

### Additional Resources

## The following tutorials explain how to perform other common tasks in SAS: