

## Chapter 126

# Stratified Random Sampling

## Introduction

In stratified random sampling, samples are drawn from a population that has been partitioned into subpopulations (or strata) based on shared characteristics (e.g., gender, age, location, etc.). Each individual stratum is sampled independently of all other strata. The Stratified Random Sampling tool in **NCSS** can be used to quickly generate  $K$  independent stratified random samples from a dataset, where each random sample has  $N$  items divided among  $H$  strata. Each random sample is generated without replacement. The user may create a sample based on the data in entire rows or simply sample values from a single column. The random sample selection data may be written to a report in the Output window or stored in user-specified columns of the Data Table.

Random numbers for sampling are generated using the Mersenne Twister algorithm. The user may enter a random seed to replicate previous sampling results or generate a random seed based on the computer's internal clock.

The Stratified Random Sampling tool can be accessed from the Data or Tools menu on the Data window.

*Example – Result of randomly selecting 8 rows from a dataset with S1 and S2 as stratification columns and sample allocation among strata proportional to stratum size*

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Selected
3	A	1	40	66		Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Not Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Selected
16	B	2	16	18	51	Selected
17	B	2	3			Not Selected
18	B	2	73	61		Not Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Not Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Not Selected

# Random Sampling Options

The methods in this procedure that may be used to sample the dataset are described below.

## Dataset Sampling Type

This option specifies the method that will be used to sample the dataset. The options are

- Randomly Select Rows from the Dataset**

In this sampling type, the data in all non-empty columns is used to determine eligibility for sampling. Only rows that are completely empty are ignored. Rows that have some missing values may still be selected in the random sample. This sampling type stores either row numbers (collapsed or expanded to the corresponding rows) or expanded binary selection values that indicate whether each non-empty row was selected.

*Example – Result of randomly selecting 6 rows from a dataset with S1 as the stratification column and sample allocation among strata proportional to stratum size*

Row	S1	X	Y	Z	Random Sample (Binary Format)	Random Sample	Value from S1 (Collapsed Format)	Random Sample (Collapsed Format)
1	A	43			Not Selected		A	2
2	A	2	3		Selected	2	A	3
3	A	40	66		Selected	3	A	7
4	A	92		30	Not Selected		A	8
5	A	45	0	12	Not Selected		B	12
6	A	29	46	44	Not Selected		B	13
7	A	57	44		Selected	7		
8	A	48	12	38	Selected	8		
9	B	3	34		Not Selected			
10								
11	B	78		66	Not Selected			
12	B	52	58	59	Selected	12		
13	B	16	18	51	Selected	13		

## Stratified Random Sampling

- **Randomly Select Values from a Column**

In this sampling type, only the data in the chosen column is used to determine eligibility for sampling. Only non-missing values in the column may be selected. This sampling type stores either actual values (collapsed or expanded to the corresponding rows) or expanded binary selection values that indicate whether each non-missing value was selected.

*Example – Result of randomly selecting 6 values from column Y with S1 as the stratification column and sample allocation among strata proportional to stratum size*

Row	S1	X	Y	Z	Random Sample (Binary Format)	Random Sample	Value from S1 (Collapsed Format)	Random Sample (Collapsed Format)
1	A	43					A	0
2	A	2	3		Not Selected		A	44
3	A	40	66		Selected	66	A	46
4	A	92		30			A	66
5	A	45	0	12	Selected	0	B	34
6	A	29	46	44	Selected	46	B	58
7	A	57	44		Selected	44		
8	A	48	12	38	Not Selected			
9	B	3	34		Selected	34		
10								
11	B	78		66				
12	B	52	58	59	Selected	58		
13	B	16	18	51	Not Selected			

## Stratification Columns

You may select up to 8 categorical data columns for stratification. The strata will be created as the combination of all unique values in the stratification columns. For example, if stratification column 1 has three unique values (A, B, and C) and stratification column 2 has two unique values (1 and 2), then the six resulting strata would be A1, A2, B1, B2, C1, and C2.

Random samples are generated independently within each stratum. Any rows for which any of the stratification values are missing will not be eligible for sampling.

The number of items sampled from each stratum is controlled by the *Sample Allocation among Strata* and *Sample Size* options.

In the example below, S1 (with values A, B, and C) is entered as the stratification column. There are twice as many A's as B's and C's. Since the sample allocation among strata is proportional to sample size, there are twice as many values selected from stratum A than from B and C.

Stratified Random Sampling

Example – Result of randomly selecting 8 rows from a dataset with S1 as the stratification column and sample allocation among strata proportional to stratum size

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Selected
3	A	1	40	66		Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Not Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Not Selected
16	B	2	16	18	51	Selected
17	B	2	3			Not Selected
18	B	2	73	61		Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Not Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Not Selected

Stratified Random Sampling

In the example below, S1 (with values A, B, and C) and S2 (with values 1 and 2) are entered as the stratification columns. There are 6 resulting strata (A1, A2, B1, B2, C1, and C2). With equal sample allocation among strata, two values are selected from each stratum when randomly sampling 12 rows.

*Example – Result of randomly selecting 12 rows from the dataset with S1 and S2 as stratification variables and equal sample allocation among strata*

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Selected
3	A	1	40	66		Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Selected
16	B	2	16	18	51	Selected
17	B	2	3			Not Selected
18	B	2	73	61		Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Selected

## Sample Allocation among Strata

This option is used along with the sample size to determine how many items will be sampled from each stratum. If the chosen sample size is not divisible by the number of strata, then some of the strata may have more or fewer items than expected.

The options are

- **Equal**

An equal number of items is selected from each stratum.

In the example below, S1 (with values A, B, and C) is entered as the stratification column. Even though there are twice as many A's as B's and C's three rows are selected from each stratum since the sample allocation among strata is set to "Equal".

*Example – Result of randomly selecting 9 rows from the dataset with S1 as the stratification column and Sample Allocation among Strata = "Equal"*

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Selected
3	A	1	40	66		Not Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Not Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Not Selected
16	B	2	16	18	51	Selected
17	B	2	3			Selected
18	B	2	73	61		Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Not Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Selected

## Stratified Random Sampling

- **Proportional to Stratum Size**

The relative strata sample size allotments are determined by the size of each stratum. The relative sample size is larger for larger strata.

In the example below, S1 (with values A, B, and C) is entered as the stratification column. There are twice as many A's as B's and C's. Since the sample allocation among strata is proportional to sample size, there are twice as many values selected from stratum A than from B and C.

*Example – Result of randomly selecting 8 rows from the dataset with S1 as the stratification column and sample allocation among strata proportional to stratum size*

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Selected
3	A	1	40	66		Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Not Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Not Selected
16	B	2	16	18	51	Selected
17	B	2	3			Not Selected
18	B	2	73	61		Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Not Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Not Selected

Stratified Random Sampling

- Custom**

The relative strata allotments are determined by user-entered allocation ratios. The first value is the ratio for the first stratum, the second for the second stratum, and so on. The values need not sum to one since they will be scaled to relative percentages, so to achieve 1:2:1 stratum allocation among 3 strata you could enter *1 2 1*, *0.25 0.5 0.25*, or *25 50 25*. All would give the same results for the stratum allocation percentages.

In the example below, S1 (with values A, B, and C) is entered as the stratification column. There are twice as many A's as B's and C's. The custom sample allocation among strata is "1 2 1", such that there are twice as many rows sampled from stratum B (4) as from strata A and C (each with 2).

*Example – Result of randomly selecting 8 rows from the dataset with S1 as the stratification column and sample allocation among strata set to custom (1 2 1)*

Row	S1	S2	X	Y	Z	Random Sample (Binary Format)
1	A	1	43			Not Selected
2	A	1	2	3		Not Selected
3	A	1	40	66		Not Selected
4	A	1	92		30	Not Selected
5	A	1	45	0	12	Not Selected
6	A	1	29	46	44	Not Selected
7	A	2	57	44		Selected
8	A	2	48	12	38	Not Selected
9	A	2	20		2	Not Selected
10	A	2	27			Not Selected
11	A	2	28	84	98	Not Selected
12	A	2	35	97	86	Selected
13	B	1	3	34		Not Selected
14	B	1	78		66	Not Selected
15	B	1	52	58	59	Selected
16	B	2	16	18	51	Selected
17	B	2	3			Selected
18	B	2	73	61		Selected
19	C	1	79	55	88	Not Selected
20	C	1	95	65	28	Selected
21	C	1	67			Not Selected
22	C	2	61	38	26	Not Selected
23	C	2	25		31	Selected
24	C	2	13	79	99	Not Selected



## Random Sample Storage

The random sample may be stored on the Data Table using the options on the Storage tab. The various methods that may be used to store the selected data are described below.

### Selection Data to Store

This option specifies how selection data will be stored in the Data Table. The options are

- Binary Selection Indicators (Dummy Values)**

The stored selection data will appear as binary indicator (i.e., “dummy”) values. The random sample data retains all information contained in other columns of the dataset.

You may enter custom values for the binary indicators of selected and not selected.

*Example – Result of randomly selecting 6 rows from a dataset with S1 as the stratification column and sample allocation among strata proportional to stratum size with Selected = “1” and Not Selected = “0”*

Row	S1	X	Y	Z	Random Sample (Binary Format)
1	A	43			0
2	A	2	3		1
3	A	40	66		1
4	A	92		30	0
5	A	45	0	12	0
6	A	29	46	44	0
7	A	57	44		1
8	A	48	12	38	1
9	B	3	34		0
10	B	78		66	0
11	B	52	58	59	1
12	B	16	18	51	1

*Example – Result of randomly selecting 6 values from column Y with S1 as the stratification column and sample allocation among strata proportional to stratum size with Selected = “Yes” and Not Selected = “[Blank]”*

Row	S1	X	Y	Z	Random Sample (Binary Format)
1	A	43			
2	A	2	3		
3	A	40	66		Yes
4	A	92		30	
5	A	45	0	12	Yes
6	A	29	46	44	Yes
7	A	57	44		Yes
8	A	48	12	38	
9	B	3	34		Yes
10	B	78		66	
11	B	52	58	59	Yes
12	B	16	18	51	

Stratified Random Sampling

- Selected Rows/Values**

The stored selection data (either row numbers or actual data values depending on your selection for Dataset Sampling Type) will appear in place on corresponding rows. The random sample data retains all information contained in other columns of the dataset.

*Example – Result of randomly selecting 6 rows from a dataset with S1 as the stratification column and sample allocation among strata proportional to stratum size*

Row	S1	X	Y	Z	Random Sample
1	A	43			
2	A	2	3		2
3	A	40	66		3
4	A	92		30	
5	A	45	0	12	
6	A	29	46	44	
7	A	57	44		7
8	A	48	12	38	8
9	B	3	34		
10	B	78		66	
11	B	52	58	59	11
12	B	16	18	51	12

*Example – Result of randomly selecting 6 values from column Y with S1 as the stratification column and sample allocation among strata proportional to stratum size*

Row	S1	X	Y	Z	Random Sample
1	A	43			
2	A	2	3		
3	A	40	66		66
4	A	92		30	
5	A	45	0	12	0
6	A	29	46	44	46
7	A	57	44		44
8	A	48	12	38	
9	B	3	34		34
10	B	78		66	
11	B	52	58	59	58
12	B	16	18	51	

Stratified Random Sampling

• **Selected Rows/Values (Collapsed)**

The stored selection data (either row numbers or actual data values depending on your selection for Dataset Sampling Type) will appear in collapsed format. The stored data may be sorted if desired with sorting controlled by the options on the Reports tab. The stored data will appear the same as that displayed in the Random Sample List(s) in the output report.

When sampling values from a column, the original row numbers of selected values may also be stored in an adjacent column to retain a reference to information contained in other columns of the dataset.

*Example – Result of randomly selecting 6 rows from a dataset with S1 as the stratification column, sample allocation among strata proportional to stratum size, and the random sample not sorted*

Row	S1	X	Y	Z	Value from S1 (Collapsed Format)	Random Sample (Collapsed Format)
1	A	43			A	7
2	A	2	3		A	2
3	A	40	66		A	3
4	A	92		30	A	8
5	A	45	0	12	B	12
6	A	29	46	44	B	11
7	A	57	44			
8	A	48	12	38		
9	B	3	34			
10	B	78		66		
11	B	52	58	59		
12	B	16	18	51		

*Example – Result of randomly selecting 6 values from column Y with S1 as the stratification column, sample allocation among strata proportional to stratum size, output sorted by selected value, and with corresponding stratification values and row numbers stored in adjacent columns*

Row	S1	X	Y	Z	Value from S1 (Collapsed Format)	Original Row Number	Random Sample (Collapsed Format)
1	A	43			A	5	0
2	A	2	3		A	7	44
3	A	40	66		A	6	46
4	A	92		30	A	3	66
5	A	45	0	12	B	9	34
6	A	29	46	44	B	11	58
7	A	57	44				
8	A	48	12	38			
9	B	3	34				
10	B	78		66			
11	B	52	58	59			
12	B	16	18	51			

## Example 1 – Selecting a Stratified Random Sample of Rows from the Resale Dataset

This example will demonstrate how to take a stratified random sample of 30 rows (across all strata) from a dataset. The sampling will be stratified by State, with sample allocation among strata proportional to stratum size. The sample selection data will be stored as indicator or dummy values (0's and 1's) in the Data Table on the Data Window, with a 1 indicating that the row was selected and a 0 indicating that it was not selected.

This example will also cover how to save the randomly selected rows to a new subset data file using the filter.

For reproducibility, we will use a random seed of 60677.

### Setup

To run this example, complete the following steps:

#### 1 Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

#### 2 Specify the Stratified Random Sampling procedure options

- Find and open the **Stratified Random Sampling** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

##### Random Sampling Tab

Sample Size Entry Type ..... **Enter the Number of Rows/Values to Select**  
 Number of Rows/Values to Select ..... **30**  
 Dataset Sampling Type ..... **Randomly Select Rows from the Dataset**  
 Random Seed..... **60677** (for reproducibility)  
 Stratification Column(s) ..... **State**  
 Sample Allocation among Strata..... **Proportional to Stratum Size**

##### Report Options (*in the Toolbar*)

Variable Labels ..... **Column Names**

#### 3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

## Random Sampling Summary Report

### Stratified Random Sampling Summary

Number of Rows Selected	30 of 150 (20%)
Dataset Sampling Type	Rows Randomly Selected
Number of Random Sample Replicates	1
User-Entered Random Seed*	60677
Stratification Column	State
Number of Strata	2
Sample Allocation among Strata	Proportional to Stratum Size
Number of Rows Selected by Stratum	State="Nev": 17 of 88 (19.318%) State="Vir": 13 of 62 (20.968%)

Storage: The random sample selection data was stored in column 21 of the Data Table.

\* You can use this random seed to replicate this random sample, but the original dataset, number of random samples to generate, and sampling options must be the same.

This report gives a summary of the random sample that was generated. The sample size of 30 corresponds to 20% of all 150 available rows with data. 17 of 88 rows were selected from State = "Nev" and 13 of 62 rows were selected from State = "Vir". The random sample itself was stored in the Data Table in column 21.

## Random Sample List Report

### Stratified Random Sample 1 of 1

Number of Rows Selected	30 of 150 (20%)
Dataset Sampling Type	Rows Randomly Selected
User-Entered Random Seed	60677
Stratification Column	State

Selected Item	Value from Stratification Column State	Selected Row
1	Nev	18
2	Nev	23
3	Nev	35
4	Nev	36
5	Nev	38
6	Nev	43
7	Nev	48
8	Nev	49
9	Nev	50
10	Nev	59
.	.	.
.	.	.
.	.	.

This report lists in collapsed format the 30 rows that were selected.

## Stored Random Sample Data

Row	State	Sample_1
1	Nev	0
2	Nev	0
3	Vir	0
.	.	.
.	.	.
16	Nev	0
17	Vir	1
18	Nev	1
19	Vir	0
20	Vir	1
.	.	.
.	.	.
147	Vir	0
148	Vir	1
149	Nev	0
150	Vir	0

In column 21 of the Data Table on the Data Window (now named "Sample\_1"), you'll find the random sample stored using binary indicator (dummy) values. You can now set a filter on this column to restrict further analyses to just this sampled set or for saving or exporting the random subset to a new file.

The newly created Note for column "Sample\_1" contains the following information about the sampled data.

Stratified Random Sample Storage Column 1 of 1

---

Number of Rows Selected: 30 of 150 (20%)  
 Dataset Sampling Type: Rows Randomly Selected  
 User-Entered Random Seed\*: 60677

Stratification Column: State  
 Number of Strata: 2  
 Sample Allocation among Strata: Proportional to Stratum Size  
 Number of Rows Selected by Stratum: State="Nev": 17 of 88 (19.318%); State="Vir": 13 of 62 (20.968%)

\* You can use this random seed to replicate this random sample, but the original dataset, number of random samples to generate, and sampling options must be the same.

---

## Saving the Random Sample Subset to a New Data File using the Data Filter

To save the random sample data to a new subset file using the data filter, do the following:

### 4 Set a filter on column 21

- Click the **Filter** button in the toolbar to load the **Filter Editor**.
- Click **New List Condition**, select **column 21 "Sample\_1"** from the column selection window, and click **OK**.
- Check only the value **1** in the list under **Sample\_1 Condition** and click **OK**.
- The dataset will now be filtered with only the randomly selected rows highlighted.

### 5 Save the filtered data to a new file

- On the Data Window menu, select **File** then **Save Subset As**.
- Under Rows to Save, select Save Only Those Rows that Pass the Active Filter.
- Leave Columns to Save set to Save All Columns.
- Click **OK** to save the subset to a new data file.
- Name the new subset data file and click Save to save it.
- Open the saved data file by selecting it from the **File** menu.
- The subset file contains only the randomly selected rows and can be used for further analyses.

## Example 2 – Selecting a Stratified Random Sample of 20% of Values from a Column in the Resale Dataset

This example will demonstrate how to take a stratified random sample from a column in a dataset. We'll select 20% of the values in the column called Price. The sampling will be stratified by State, with equal sample allocation among strata. The sample selection data will be stored in collapsed format in the Data Table with stratification values and row numbers also stored for reference to the original dataset. The resulting sample data will be sorted by value within each stratum.

For reproducibility, we'll use a random seed of 22077.

This example will also demonstrate how to randomly assign selected values to groups with sequential balance using the Block Randomization method in the Randomization Lists procedure.

**If you want to randomly assign the selected values to G groups within strata without regard for sequential balance over time, you can accomplish this all in one step using the Stratified Random Sampling with Group Assignment procedure.**

### Setup

To run this part of the example, complete the following steps:

**1 Open the Resale example dataset**

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

**2 Specify the Stratified Random Sampling procedure options**

- Find and open the **Stratified Random Sampling** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

<b>Random Sampling Tab</b>	
<hr/>	
Sample Size Entry Type .....	<b>Enter the Percent of Rows/Values to Select</b>
Percent of Rows/Values to Select.....	<b>20</b>
Dataset Sampling Type .....	<b>Randomly Select Values from a Column</b>
Column to Sample .....	<b>Price</b>
Random Seed.....	<b>22077</b> (for reproducibility)
Stratification Column(s) .....	<b>State</b>
Sample Allocation among Strata.....	<b>Equal</b>
<b>Reports Tab</b>	
<hr/>	
Sort Random Sample List(s).....	<b>Checked</b>
Sort Type .....	<b>By Selected Row/Value</b>



Stratified Random Sampling

Storage Tab

---

Store the Random Sample(s) in the Data Table ..... **Checked**  
 Selection Data to Store ..... **Selected Rows/Values (Collapsed)**  
 Store the stratification values ..... **Checked**  
 Store the original row number of each selected ..... **Checked**  
 value

Report Options (*in the Toolbar*)

---

Variable Labels ..... **Column Names**

**3 Run the procedure**

- Click the **Run** button to perform the calculations and generate the output.

---

**Random Sampling Summary Report**

**Stratified Random Sampling Summary**

---

Number of Values Selected	30 of 150 (20%)
Dataset Sampling Type	Values Randomly Selected from Price
Number of Random Sample Replicates	1
User-Entered Random Seed*	22077
Stratification Column	State
Number of Strata	2
Sample Allocation among Strata	Equal
Number of Values Selected by Stratum	State="Nev": 15 of 88 (17.045%) State="Vir": 15 of 62 (24.194%)

Storage: The random sample selection data was stored in columns 21 through 23 of the Data Table.

---

\* You can use this random seed to replicate this random sample, but the original dataset, number of random samples to generate, and sampling options must be the same.

This report gives a summary of the random sample that was generated by selecting values from Price. A sample size of 30 corresponds to 20% of all 150 available rows with data. The random sample itself was stored in the Data Table in columns 21 through 23.

## Random Sample List Report

### Stratified Random Sample 1 of 1

---

Number of Values Selected 30 of 150 (20%)  
 Dataset Sampling Type Values Randomly Selected from Price  
 User-Entered Random Seed 22077  
 Stratification Column State

---

Selected Item	Value from Stratification Column State	Original Row	Value Selected from Price
1	Nev	79	26300
2	Nev	71	53300
3	Nev	57	102200
4	Nev	138	133400
5	Nev	6	147500
6	Nev	123	150100
7	Nev	49	167000
8	Nev	84	183300
9	Nev	108	198000
10	Nev	140	233200
11	Nev	29	238900
12	Nev	139	247000
13	Nev	129	318500
14	Nev	77	360100
15	Nev	8	395700
16	Vir	55	32900
17	Vir	20	101800
18	Vir	70	119200
19	Vir	69	141200
20	Vir	61	158400
21	Vir	47	165900
22	Vir	119	195400
23	Vir	32	208900
24	Vir	74	220600
25	Vir	54	241300
26	Vir	122	241800
27	Vir	98	247500
28	Vir	86	291700
29	Vir	116	328300
30	Vir	34	335300

---

This report lists in collapsed format the 30 values from Price (20% of all values) that were selected along with their corresponding strata and original row numbers. The report is sorted by value with each stratum.

## Stored Random Sample Data

Row	Strat_Col_1_State	Row_1	Sample_1
1	Nev	79	26300
2	Nev	71	53300
3	Nev	57	102200
4	Nev	138	133400
5	Nev	6	147500
6	Nev	123	150100
7	Nev	49	167000
8	Nev	84	183300
9	Nev	108	198000
10	Nev	140	233200
.	.	.	.
.	.	.	.
.	.	.	.
21	Vir	47	165900
22	Vir	119	195400
23	Vir	32	208900
24	Vir	74	220600
25	Vir	54	241300
26	Vir	122	241800
27	Vir	98	247500
28	Vir	86	291700
29	Vir	116	328300
30	Vir	34	335300

In columns 21 through 23 of the Data Table on the Data Window, you'll find the random sample data stored in collapsed format. The stored data includes the sampled values ("Sample\_1"), the stratification values ("Strat\_Col\_1\_State"), and the original row numbers ("Row\_1") of the values selected from Price.

The newly created Note for column "Sample\_1" contains the following information about the sampled data.

Stratified Random Sample Storage Column 1 of 1

-----  
 Number of Values Selected: 30 of 150 (20%)  
 Dataset Sampling Type: Values Randomly Selected from Price  
 User-Entered Random Seed\*: 22077

Stratification Column: State  
 Number of Strata: 2  
 Sample Allocation among Strata: Equal  
 Number of Values Selected by Stratum: State="Nev": 15 of 88 (17.045%); State="Vir": 15 of 62 (24.194%)

\* You can use this random seed to replicate this random sample, but the original dataset, number of random samples to generate, and sampling options must be the same.

## Assigning Randomly Selected Values to Groups with Stratification using the Randomization Lists Procedure

Continuing with this example, suppose you want to assign the 30 randomly selected values to groups within each stratum with sequential balance. This can be accomplished using the Block Randomization method in the Randomization Lists procedure.

**If you want to randomly assign the selected values to *G* groups within strata without regard for sequential balance over time, you can accomplish this in one step using the Stratified Random Sampling with Group Assignment procedure.**

**The purpose of this example is to demonstrate how the Randomization Lists procedure may be used in conjunction with the Stratified Random Sampling procedure to assign actual sampled items to groups. The Randomization Lists procedure contains additional randomization algorithms for assigning items to groups that are not available in the Stratified Random Sampling with Group Assignment procedure and may be of interest to the user.**

For reproducibility, we will use a random seed of 60502 in the Randomization Lists procedure.

### Setup

To run this part of the example, complete the following steps:

#### 4 Specify the Randomization Lists procedure options

- Find and open the **Randomization Lists** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 6** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

#### Randomization Tab

Randomization Algorithm.....	<b>Block Randomization</b>
Random Seed.....	<b>60502</b> (for reproducibility)
Block Size Multiplier(s) .....	<b>1 2</b>
Block-Subject Allocation .....	<b>Random</b>
Constrain the actual block allocation... ..	<b>Unchecked</b>
Search for a randomization list... ..	<b>Unchecked</b>
Maximum Search Iterations .....	<b>1000</b>
Sample Size Entry Type .....	<b>Enter the Overall Total Sample Size for the Study</b>
Overall Total Sample Size .....	<b>30</b>
Number of Treatment Groups .....	<b>3</b>
Title or Name .....	<b>Group</b>
Assign Equal Sample Sizes to All Treatment Groups .....	<b>Checked</b>
Treatment Group 1 Label.....	<b>A</b>
Treatment Group 2 Label.....	<b>B</b>
Treatment Group 3 Label.....	<b>C</b>

## Stratified Random Sampling

## Centers (Strata) Tab

---

Number of Centers (Strata) .....**2**  
 Title or Name .....**State**  
 Assign Equal Sample Sizes to All Centers (Strata).....**Checked**  
 Center (Stratum) 1 Label .....**Nev**  
 Center (Stratum) 2 Label .....**Vir**  
 Number of Additional Factors .....**0**

## List Options Tab

---

Include a Column of Sequence Numbers .....**Checked**  
 Include a Column of Subject ID's .....**Unchecked**  
 Include a Column of Block Identifiers.....**Checked**  
 Include a Column of Combined Stratification Codes.....**Unchecked**  
 Include a Column of Abbreviated Treatment Codes .....**Unchecked**  
 Include a Column of Unique Randomization Codes .....**Unchecked**

## Reports Tab

---

Show Summary .....**Checked**  
 Show Randomization List .....**Checked**  
 Show Randomization List Details .....**Unchecked**  
 Show References .....**Checked**

## Storage Tab

---

Store Randomization List.....**Checked**  
 Store Randomization List Data Starting In.....**<Empty>**  
 Update output column names... .....**Checked**

**5 Run the procedure**

- Click the **Run** button to perform the calculations and generate the output.

## Randomization List Summary

### Randomization List Summary

---

Randomization Algorithm	Block Randomization
User-Entered Random Seed	60502
Block Size Multiplier(s)	1 2
Block Size(s)	3 6
Total Number of Blocks	7

Overall Total Sample Size	30
Number of Treatment Groups	3
Number of Centers (Strata)	2

---

### Treatments

---

Group	Group Code	Sample Size	Allocation %	
			Actual	Target
A	A	10	33.33%	33.33%
B	B	10	33.33%	33.33%
C	C	10	33.33%	33.33%

---

### Centers (Strata)

---

State	Blocks	Sample Size	Allocation %	
			Actual	Target
Nev	3	15	50%	50%
Vir	4	15	50%	50%

---

This report displays the summary of the randomization list, including strata and block information.

## Stratified Random Sampling

## Randomization List

**Randomization List**

---

Randomization Algorithm      Block Randomization (Block Sizes = 3 6)  
User-Entered Random Seed    60502

---

<b>Sequence</b>	<b>State</b>	<b>Block</b>	<b>Group</b>
1	Nev	1	A
2	Nev	1	A
3	Nev	1	C
4	Nev	1	C
5	Nev	1	B
6	Nev	1	B
7	Nev	2	C
8	Nev	2	A
9	Nev	2	C
10	Nev	2	B
11	Nev	2	A
12	Nev	2	B
13	Nev	3	A
14	Nev	3	C
15	Nev	3	B
16	Vir	4	C
17	Vir	4	A
18	Vir	4	B
19	Vir	5	C
20	Vir	5	A
21	Vir	5	B
22	Vir	6	B
23	Vir	6	C
24	Vir	6	A
25	Vir	7	C
26	Vir	7	C
27	Vir	7	A
28	Vir	7	B
29	Vir	7	B
30	Vir	7	A

---

The report lists the group to which each item (denoted by Sequence) should be assigned.

Stratified Random Sampling

Stored Randomization List Data

Row	Strat_Col_1_State	Row_1	Sample_1	Sequence	Statex	Block	Group
1	Nev	79	26300	1	Nev	1	A
2	Nev	71	53300	2	Nev	1	A
3	Nev	57	102200	3	Nev	1	C
4	Nev	138	133400	4	Nev	1	C
5	Nev	6	147500	5	Nev	1	B
6	Nev	123	150100	6	Nev	1	B
7	Nev	49	167000	7	Nev	2	C
8	Nev	84	183300	8	Nev	2	A
9	Nev	108	198000	9	Nev	2	C
10	Nev	140	233200	10	Nev	2	B
11	Nev	29	238900	11	Nev	2	A
12	Nev	139	247000	12	Nev	2	B
13	Nev	129	318500	13	Nev	3	A
14	Nev	77	360100	14	Nev	3	C
15	Nev	8	395700	15	Nev	3	B
16	Vir	55	32900	16	Vir	4	C
17	Vir	20	101800	17	Vir	4	A
18	Vir	70	119200	18	Vir	4	B
19	Vir	69	141200	19	Vir	5	C
20	Vir	61	158400	20	Vir	5	A
21	Vir	47	165900	21	Vir	5	B
22	Vir	119	195400	22	Vir	6	B
23	Vir	32	208900	23	Vir	6	C
24	Vir	74	220600	24	Vir	6	A
25	Vir	54	241300	25	Vir	7	C
26	Vir	122	241800	26	Vir	7	C
27	Vir	98	247500	27	Vir	7	A
28	Vir	86	291700	28	Vir	7	B
29	Vir	116	328300	29	Vir	7	B
30	Vir	34	335300	30	Vir	7	A

In columns 21 through 23 of the Data Table on the Data Window, you'll find the random sample data stored in collapsed format from the Stratified Random Sampling procedure. In columns 24 through 27 you'll find the stored list from the Randomization Lists procedure, including the assigned group for each selected value, randomized independently within each stratum.