



المعهد العربي للتدريب والبحوث الإحصائية  
دورة متوسطة حول تصميم عينات المسوح الإحصائية  
التطبيق العملي باستخدام SPSS الجزء الثاني

سحب العينة العشوائية /الطبقية/العنقودية

إعداد:

نايف عابد

[nayif@pcbs.gov.ps](mailto:nayif@pcbs.gov.ps) [nayifabed@yaoo.com](mailto:nayifabed@yaoo.com)

حزيران - 2022

## 1. العينة الطبقيّة بأكثر من متغير للطبقة:

سيتم استخدام إطار العينة السابق والمسمى Frame1 حيث يحتوي على 439 أشخاص، وكل شخص يوجد عنه مجموعة من المعلومات أهمها الجنس والفئة العمرية والحالة الاجتماعية... الخ.

يراد سحب عينة طبقية عشوائية بسيطة (بدون إرجاع)، مقدارها 60 شخص.

متغيرات الطبقة هي: الجنس والفئات العمرية (3 فئات)

a- توزيع العينة على الطبقات هو توزيع متناسب مع حجم الطبقة.

b- توزيع متناسوي

c- توزيع حسب متوسط التوزيعين المتناسوي والمتناسب مع حجم الطبقة (تمرين بيتي)

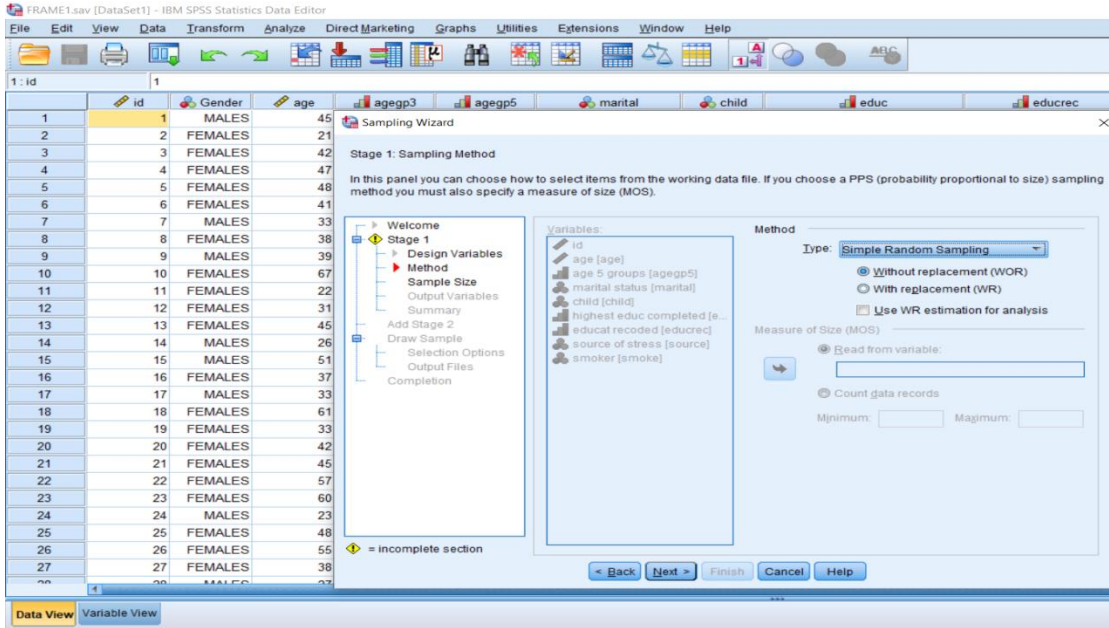
سيتم إعداد ملف لتصميم العينة اسمه plan\_stratified

id	Gender	age	agegp3	agegp5	marital	child	educ	educrec	source	smoke	var
1	MALES	45	45+	41 - 49	MARRIED FIRST TIME	YES	SOME ADDITIONAL TRAINING	some additional training	WORK	NO	
2	FEMALES	21	18 - 29	18 - 24	SINGLE	NO	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	NO	
3	FEMALES	42	30 - 44	41 - 49	MARRIED FIRST TIME	YES	SOME SECONDARY	did not complete high sc...	FAMILY	NO	
4	FEMALES	47	45+	41 - 49	MARRIED FIRST TIME	YES	SOME SECONDARY	did not complete high sc...	CHILDREN	NO	
5	FEMALES	48	45+	41 - 49	MARRIED FIRST TIME	NO	COMPLETED HIGHSCHOOL	completed high school	RELATION...	NO	
6	FEMALES	41	30 - 44	41 - 49	DIVORCED	YES	POSTGRADUATE COMPLETED	completed postgrad uni	WORK	NO	
7	MALES	33	30 - 44	33 - 40	SINGLE	NO	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	NO	
8	FEMALES	38	30 - 44	33 - 40	DIVORCED	YES	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	NO	
9	MALES	39	30 - 44	33 - 40	LIVING WITH PARTN...	YES	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	YES	
10	FEMALES	67	45+	50+	MARRIED FIRST TIME	NO	SOME SECONDARY	did not complete high sc...	SPOUSE ...	NO	
11	FEMALES	22	18 - 29	18 - 24	STEADY RELATIONS...	NO	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	NO	
12	FEMALES	31	30 - 44	25 - 32	MARRIED FIRST TIME	YES	COMPLETED UNDERGRADU...	completed undergrad uni	LIFE IN G...	NO	
13	FEMALES	45	45+	41 - 49	MARRIED FIRST TIME	YES	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	NO	
14	MALES	26	18 - 29	25 - 32	MARRIED FIRST TIME	NO	COMPLETED UNDERGRADU...	completed undergrad uni	WORK	YES	
15	MALES	51	45+	50+	REARRIED	YES	SOME SECONDARY	did not complete high sc...	MONEY/FI...	NO	
16	FEMALES	37	30 - 44	33 - 40	REARRIED	YES	COMPLETED HIGHSCHOOL	completed high school	WORK	NO	
17	MALES	33	30 - 44	33 - 40	SINGLE	NO	SOME ADDITIONAL TRAINING	some additional training	WORK	NO	
18	FEMALES	61	45+	50+	WIDOWED	YES	COMPLETED HIGHSCHOOL	completed high school	WORK	NO	
19	FEMALES	33	30 - 44	33 - 40	SINGLE	NO	SOME ADDITIONAL TRAINING	some additional training	WORK	NO	
20	FEMALES	42	30 - 44	41 - 49	MARRIED FIRST TIME	YES	SOME ADDITIONAL TRAINING	some additional training	MONEY/FI...	NO	
21	FEMALES	45	45+	41 - 49	MARRIED FIRST TIME	YES	POSTGRADUATE COMPLETED	completed postgrad uni	WORK	NO	
22	FEMALES	57	45+	50+	MARRIED FIRST TIME	NO	SOME ADDITIONAL TRAINING	some additional training	MONEY/FI...	NO	

يتم تحديد الطبقات كما هو موضح في الشكل التالي:

The screenshot shows the 'Sampling Wizard' dialog box in IBM SPSS. The 'Stratify By' section is set to 'Gender (Gender)' and 'age 3 groups (agegp3)'. The 'Variables' list includes 'id', 'age (age)', 'age 5 groups (agegp5)', 'marital status (marital)', 'child (child)', 'highest educ completed (educ)', 'educat recoded (educrec)', 'source of stress (source)', and 'smoker (smoke)'. The 'Clusters' section is empty. The 'Input Sample Weight' field is empty. The 'Stage Label' field is empty. The 'Completion' button is highlighted.

ثم اختيار نوع العينة:



لمعرفة حجم العينة في كل طبقة يجب معرفة عدد الطبقات وعدد الأشخاص في كل طبقة، وهذا يتم من خلال إعداد توزيع للطبقات، ويمكن ذلك بعدة طرق مثل عمل crosstab كما يلي:  
وهنا من الواضح أن لدينا 6 طبقات فيها مشاهدات (غير فارغة)

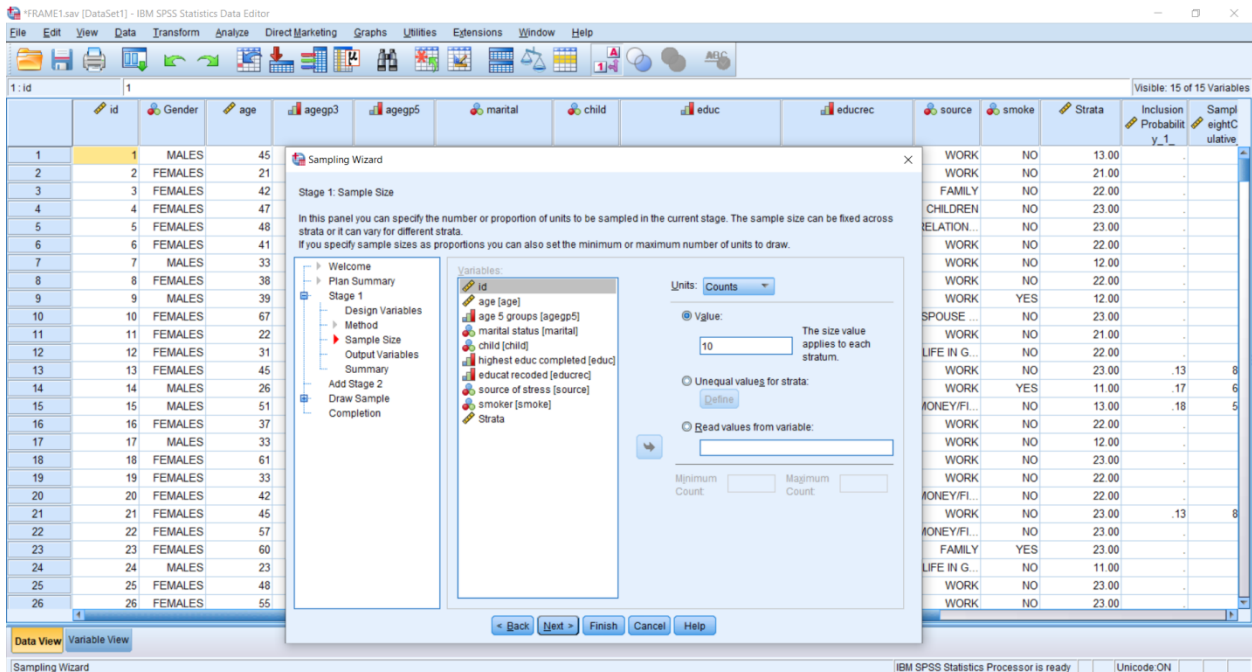
		age 3 groups			Total
		18 - 29	30 - 44	45+	
Gender	MALES	60	68	57	185
	FEMALES	89	85	80	254
Total		149	153	137	439

يمكن أيضاً عمل متغير جديد يتكون من رمز جديد للطبقة كالتالي:

رمز الطبقة = متغير الجنس X 10 + متغير الفئات العمرية

Strata	Frequency	Percent	Valid Percent	Cumulative Percent
11.00	60	13.7	13.7	13.7
12.00	68	15.5	15.5	29.2
13.00	57	13.0	13.0	42.1
21.00	89	20.3	20.3	62.4
22.00	85	19.4	19.4	81.8
23.00	80	18.2	18.2	100.0
Total	439	100.0	100.0	

في حالة التوزيع المتساوي على الطبقات يكون حجم العينة في كل طبقة =  $10 = 60 / 6$  أشخاص في كل طبقة.



وبإكمال بقية الإجراءات يتم الحصول على العينة الطبقيّة العشوائية البسيطة كما هو موضح في الجدول الذي ينتج عن برنامج SPSS

			Number of Units Sampled		Proportion of Units Sampled	
			Requested	Actual	Requested	Actual
Gender = 1	agegp3 =	1	10	10	16.7%	16.7%
		2	10	10	14.7%	14.7%
		3	10	10	17.5%	17.5%
Gender = 2	agegp3 =	1	10	10	11.2%	11.2%
		2	10	10	11.8%	11.8%
		3	10	10	12.5%	12.5%

والعينة المسحوبة على الإطار تظهر كما هو موضح بالشكل أدناه.

في حالة التوزيع المتناسب مع الحجم، يتم توزيع العينة بشكل منفصل ويمكن إعداد ذلك باستخدام برنامج excel كالتالي:

Strata	Frequency	Percent	Sample size = 60*percent	Final sample
11	60	13.7	8.22	8
12	68	15.5	9.3	9
13	57	13.0	7.8	8
21	89	20.3	12.18	12
22	85	19.4	11.64	12
23	80	18.2	10.92	11
Total	439	100.0	60.06	60

أو أن يكون التوزيع على الجدول التقاطعي كما يلي:

		age 3 groups			Total
		18 - 29	30 - 44	45+	
Gender	MALES	8	9	8	25
	FEMALES	12	12	11	35
Total		20	21	19	60

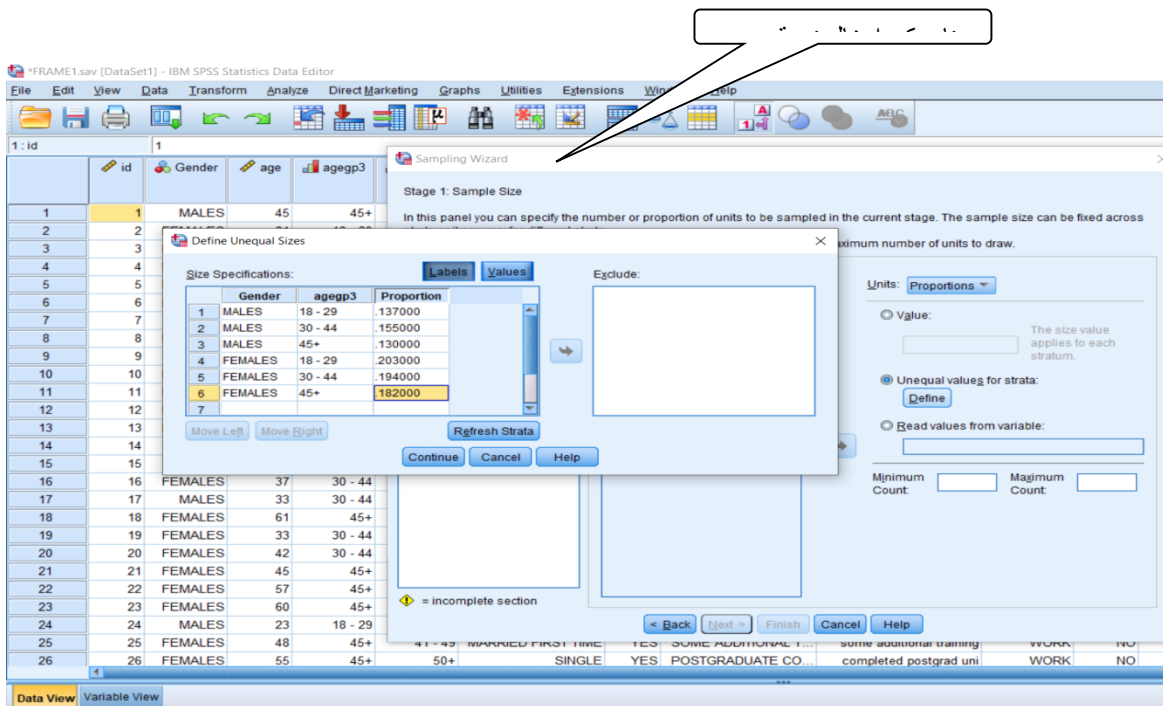
وبذلك يصبح لدينا توزيع متناسب مع حجم الطبقات

ويمكن تحديد حجم العينة في كل طبقة باستخدام Unequal Values for Strata وتحديد حجم العينة بكل طبقة بشكل يدوي أو إدخال النسبة المئوية كما في الشكل التالي والذي يليه:

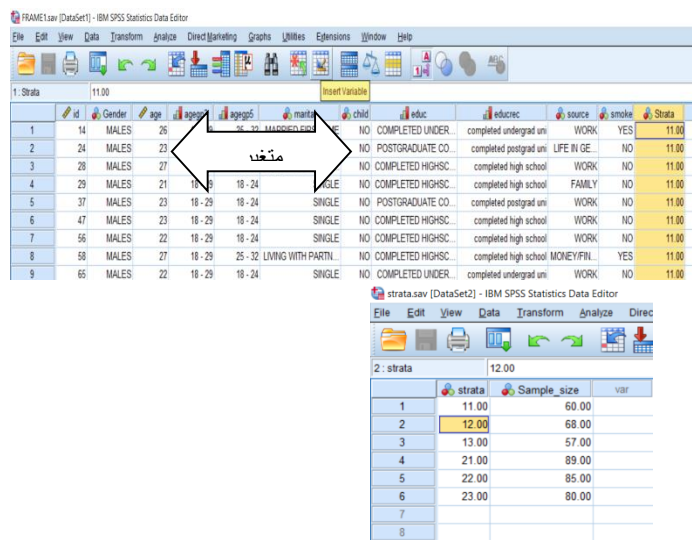
The screenshot shows the 'Sampling Wizard' dialog box in IBM SPSS Statistics. The 'Define Unequal Sizes' tab is selected, showing a table of size specifications for different strata. The table has columns for Gender, agegp3, and count. The 'Define' button is highlighted, and a callout box points to it with the text 'يتم الضغط على هذا الزر...'. The background shows a data table with columns for id, Gender, age, and agegp3.

Gender	agegp3	count
1 MALES	18 - 29	8
2 MALES	30 - 44	9
3 MALES	45+	8
4 FEMALES	18 - 29	12
5 FEMALES	30 - 44	12
6 FEMALES	45+	11





كما يمكن استخدام الأمر read values from variable وذلك من خلال دمج متغير جديد في ملف الإطار يحتوي على حجم العينة،



يتم اضافة المتغير لملف الاطار واستخدامه لتحديد حجم العينة كالتالي:

	Strata	id	Gender	age	agegp3	agegp5	marital	child	educ	educrec	source	smoke	Sample_size	var	var
1	11.00	14	MALES	26	18 - 29	25 - 32	MARRIED FIRST TIME	NO	COMPLETED UNDER...	completed undergrad uni	WORK	YES	8.00		
2	11.00	24	MALES	23	18 - 29	18 - 24	SINGLE	NO	POSTGRADUATE CO...	completed postgrad uni	LIFE IN GE...	NO	8.00		
3	11.00	28	MALES	27	18 - 29	25 - 32	STEADY RELATIONS...	NO	COMPLETED HIGHSC...	completed high school	WORK	NO	8.00		
4	11.00	29	MALES	21	18 - 29	18 - 24	SINGLE	NO	COMPLETED HIGHSC...	completed high school	FAMILY	NO	8.00		
5	11.00	37	MALES	23	18 - 29	18 - 24	SINGLE	NO	POSTGRADUATE CO...	completed postgrad uni	WORK	NO	8.00		
6	11.00	47	MALES	23	18 - 29	18 - 24	SINGLE	NO	COMPLETED HIGHSC...	completed high school	WORK	NO	8.00		
7	11.00	56	MALES	22	18 - 29	18 - 24	SINGLE	NO	COMPLETED HIGHSC...	completed high school	WORK	NO	8.00		
8	11.00	58	MALES	27	18 - 29	25 - 32	LIVING WITH PARTN...	NO	COMPLETED HIGHSC...	completed high school	MONEY/FIN...	YES	8.00		
9	11.00	65	MALES	22	18 - 29	18 - 24	SINGLE	NO	COMPLETED UNDER...	completed undergrad uni	WORK	NO	8.00		

من الشاشة التي تحتوي على حجم العينة يتم اختيار Read values from variables ويوضع فيه المتغير الذي اسمه sample size.

ويتم إكمال بقية الإجراءات لسحب العينة ويتم الحصول على العينة والمخرجات تكون كالتالي:

			Number of Units Sampled		Proportion of Units Sampled	
			Requested	Actual	Requested	Actual
Gender = 1	agegp3 =	1	8	8	13.3%	13.3%
		2	9	9	13.2%	13.2%
		3	8	8	14.0%	14.0%
Gender = 2	agegp3 =	1	12	12	13.5%	13.5%
		2	12	12	14.1%	14.1%
		3	11	11	13.8%	13.8%

والعينة التي تم سحبها هي:



	Strata	id	Gender	age	agegp3	agegp5	marital	child	educ	educrec	source	smoke	Sample_size	Inclusion_Probability_1	SampleWeightCumulative_1	SampleWeight_Final
1	11.00	14	MALES	26	18 - 29	25 - 32	MARRIED ...	NO	COMPLET...	completed ...	WORK	YES	8.00	.13		7.50
2	11.00	24	MALES	23	18 - 29	18 - 24	SINGLE	NO	POSTGRA...	completed ...	LIFE IN G ...	NO	8.00			
3	11.00	28	MALES	27	18 - 29	25 - 32	STEADY R...	NO	COMPLET...	completed ...	WORK	NO	8.00			
4	11.00	29	MALES	21	18 - 29	18 - 24	SINGLE	NO	COMPLET...	completed ...	FAMILY	NO	8.00	.13	7.50	7.50
5	11.00	37	MALES	23	18 - 29	18 - 24	SINGLE	NO	POSTGRA...	completed ...	WORK	NO	8.00			
6	11.00	47	MALES	23	18 - 29	18 - 24	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
7	11.00	56	MALES	22	18 - 29	18 - 24	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
8	11.00	58	MALES	27	18 - 29	25 - 32	LIVING WI...	NO	COMPLET...	completed ...	MONEY/FI...	YES	8.00			
9	11.00	65	MALES	22	18 - 29	18 - 24	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
10	11.00	66	MALES	23	18 - 29	18 - 24	STEADY R...	NO	COMPLET...	completed ...	WORK	NO	8.00			
11	11.00	75	MALES	22	18 - 29	18 - 24	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
12	11.00	76	MALES	25	18 - 29	25 - 32	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
13	11.00	77	MALES	26	18 - 29	25 - 32	SINGLE	NO	COMPLET...	completed ...	WORK	NO	8.00			
14	11.00	78	MALES	22	18 - 29	18 - 24	STEADY R...	NO	COMPLET...	completed ...	WORK	NO	8.00			
15	11.00	79	MALES	24	18 - 29	18 - 24	STEADY R...	NO	COMPLET...	completed ...	MONEY/FI...	NO	8.00	.13	7.50	7.50
16	11.00	115	MALES	22	18 - 29	18 - 24	LIVING WI...	NO	COMPLET...	completed ...	MONEY/FI...	NO	8.00			



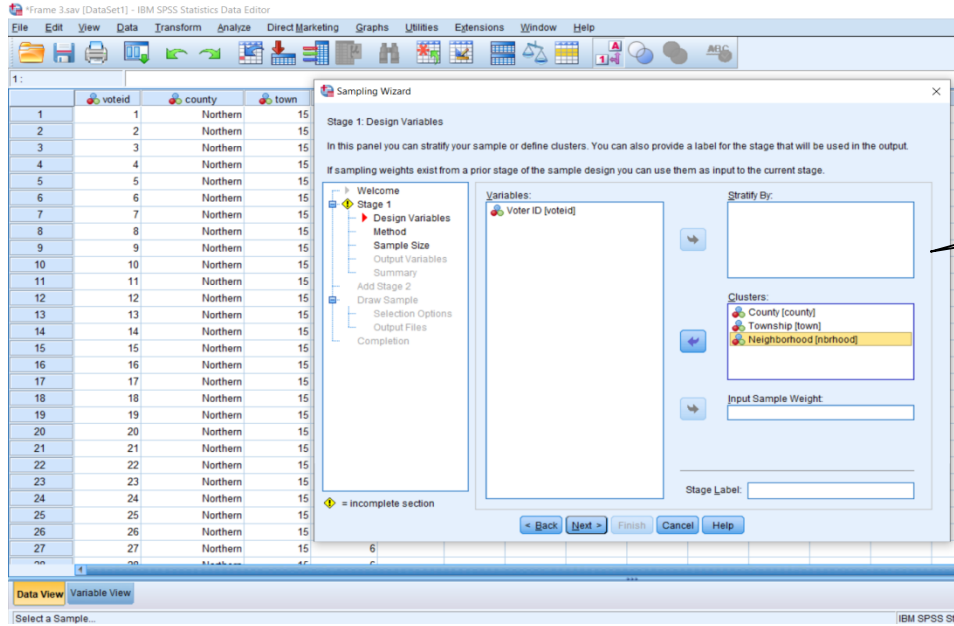
تمرين: سحب عينة طبقية مقدارها 60 موزعة بطريقة المتوسط الحسابي بين التوزيع المتساوي والتوزيع المتناسب مع الحجم.

### 3) العينة العنقودية من مرحلة واحدة:

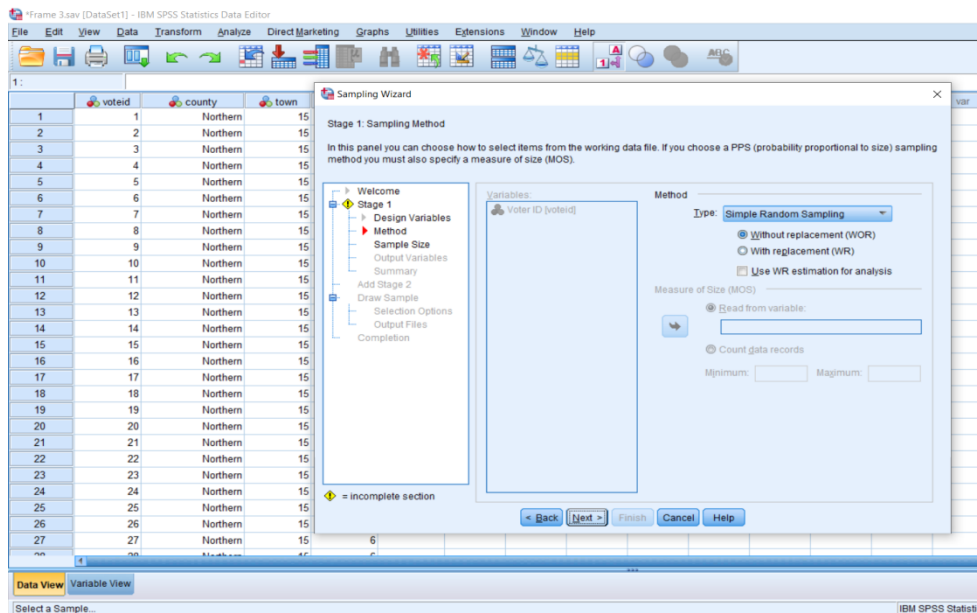
مثال لدينا إطار Frame3 فيه 118186 شخص، ويحتوي على مناطق إقليمية كبيرة (شمال، شرق، غرب، جنوب، وسط)، وكل منطقة تحتوي على عدد من المدن وكل مدينة تحتوي على عدد من الأحياء السكنية التي يوجد فيها عدد معين من السكان.

يراد سحب عينة عنقودية عشوائية بسيطة من مرحلة واحدة، والعنقود هو رمز الحي السكاني، وحجم العينة سيكون 30 حياً سكانيا (عنقوداً).

معنى أنها من مرحلة واحدة أي سيكون ضمن العينة جميع الأشخاص الساكنين في الأحياء ضمن العينة (30 حي)



نوع العينة هو عينة عشوائية بسيطة بدون إرجاع



حجم العينة هو 30 عنقود (حي سكني) في المرحلة الاولى.

The screenshot shows the 'Sampling Wizard' dialog box in IBM SPSS Statistics. The 'Stage 1: Sample Size' panel is active, showing the variable 'Voter ID [voteid]' selected. The sample size is set to 30. The background data table has columns 'voteid', 'county', and 'town'.

voteid	county	town
1	Northern	15
2	Northern	15
3	Northern	15
4	Northern	15
5	Northern	15
6	Northern	15
7	Northern	15
8	Northern	15
9	Northern	15
10	Northern	15
11	Northern	15
12	Northern	15
13	Northern	15
14	Northern	15
15	Northern	15
16	Northern	15
17	Northern	15
18	Northern	15
19	Northern	15
20	Northern	15
21	Northern	15
22	Northern	15
23	Northern	15
24	Northern	15
25	Northern	15
26	Northern	15
27	Northern	15

ويتم إجراء بقية الخطوات بشكل اعتيادي حتى سحب العينة

The screenshot shows the 'InclusionProbability...' dialog box in IBM SPSS Statistics. The 'Inclusion Probability' is set to 0.12. The background data table has columns 'voteid', 'county', 'town', 'nrhood', 'Inclusion Probability\_1', 'SampleWeightCumulative\_1', and 'SampleWeightFinal'.

voteid	county	town	nrhood	Inclusion Probability_1	SampleWeightCumulative_1	SampleWeightFinal
34928	567	Western	4	4		
34929	568	Western	4	4		
34930	569	Western	4	4		
34931	570	Western	4	4		
34932	571	Western	4	4		
34933	572	Western	4	4		
34934	573	Western	4	4		
34935	574	Western	4	4		
34936	575	Western	4	4		
34937	576	Western	4	4		
34938	1	Northern	4	4	.12	8.03
34939	2	Northern	4	4	.12	8.03
34940	3	Northern	4	4	.12	8.03
34941	4	Northern	4	4	.12	8.03
34942	5	Northern	4	4	.12	8.03
34943	6	Northern	4	4	.12	8.03
34944	7	Northern	4	4	.12	8.03
34945	8	Northern	4	4	.12	8.03
34946	9	Northern	4	4	.12	8.03
34947	10	Northern	4	4	.12	8.03
34948	11	Northern	4	4	.12	8.03
34949	12	Northern	4	4	.12	8.03
34950	13	Northern	4	4	.12	8.03
34951	14	Northern	4	4	.12	8.03
34952	15	Northern	4	4	.12	8.03
34953	16	Northern	4	4	.12	8.03

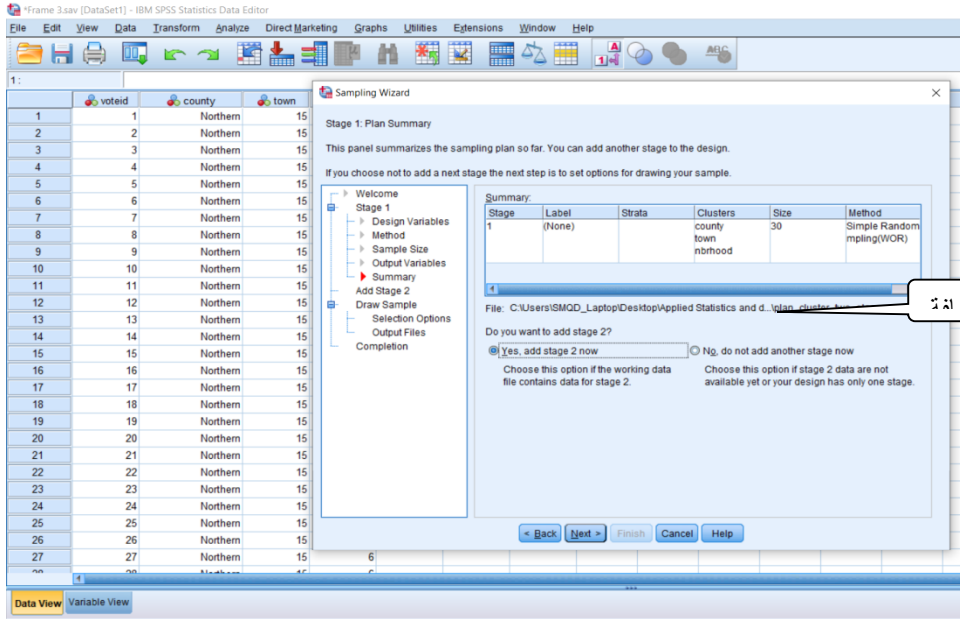
في هذه العينة يتم سحب العينة بتحديد كل الأشخاص الموجودين ضمن العنقود.

4) العينة متعددة المراحل: سحب عينة من مرحلتين، المرحلة الأولى ستكون عينة عشوائية بسيطة بدون إرجاع والمرحلة الثانية ستكون عينة عشوائية بسيطة بدون إرجاع

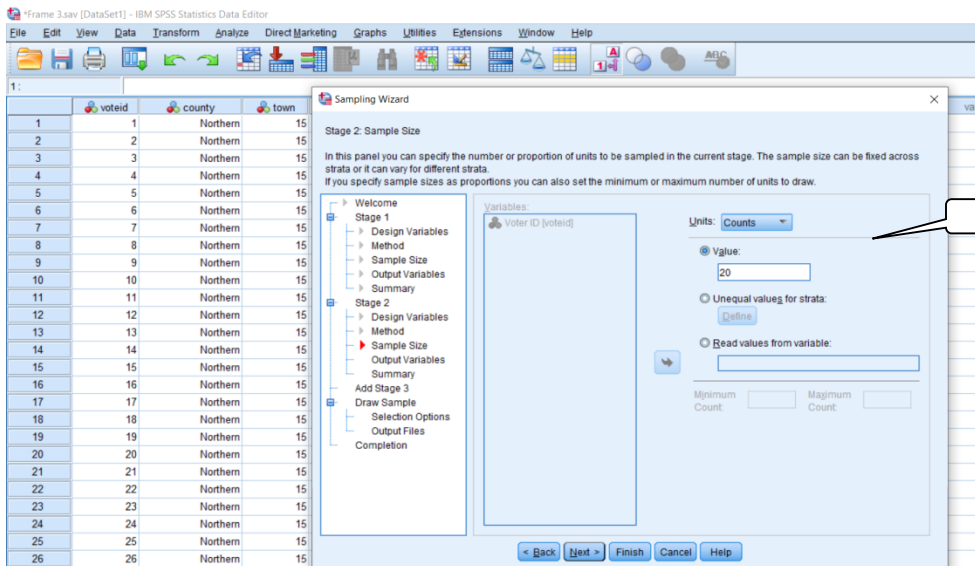
مثال على نفس الإطار 3 Frame .

يراد سحب عينة من الأحياء في المرحلة الأولى ومقدارها 30 وعينة مكونة من 20 شخص في المرحلة الثانية وكلا العينتين هي عينة عشوائية بسيطة بدون إرجاع.

يتم تصميم العينة بنفس الطريقة السابقة ونفس متغير العنقود، ولكن يضاف مرحلة ثانية للعينة يتم فيها سحب عينة من 20 شخص.



حجم العينة في المرحلة الثانية = 20



والعينة الناتجة تكون كالتالي:

voteid	county	town	nbrhood	Inclusion Probability_1	SampleWeightCumulative_1	Inclusion Probability_2	SampleWeightCumulative_2	SampleWeight_Final	var	var	var	var	var	var	var	var	var
1869	120	Eastern	11	6	-	-	-	-									
1870	121	Eastern	11	6	-	-	-	-									
1871	122	Eastern	11	6	-	-	-	-									
1872	123	Eastern	11	6	-	-	-	-									
1873	124	Eastern	11	6	-	-	-	-									
1874	125	Eastern	11	6	-	-	-	-									
1875	126	Eastern	11	6	-	-	-	-									
1876	127	Eastern	11	6	-	-	-	-									
1877	128	Eastern	11	6	.12	8.03	.03	284.78	284.78								
1878	129	Eastern	11	6	-	-	-	-									
1879	130	Eastern	11	6	-	-	-	-									
1880	131	Eastern	11	6	.12	8.03	.03	284.78	284.78								
1881	132	Eastern	11	6	-	-	-	-									
1882	133	Eastern	11	6	-	-	-	-									
1883	134	Eastern	11	6	-	-	-	-									
1884	135	Eastern	11	6	-	-	-	-									
1885	136	Eastern	11	6	-	-	-	-									
1886	137	Eastern	11	6	-	-	-	-									
1887	138	Eastern	11	6	-	-	-	-									
1888	139	Eastern	11	6	-	-	-	-									
1889	140	Eastern	11	6	.12	8.03	.03	284.78	284.78								
1890	141	Eastern	11	6	-	-	-	-									
1891	142	Eastern	11	6	-	-	-	-									
1892	143	Eastern	11	6	-	-	-	-									
1893	144	Eastern	11	6	-	-	-	-									
1894	145	Eastern	11	6	-	-	-	-									

تمرين: سحب عينة طبقية من مرحلتين: المرحلة الأولى 5 احياء سكنية عينة عشوائية منتظمة من كل طبقة ،  
والمرحلة الثانية سحب عينة من 10 اشخاص عينة عشوائية منتظمة، ومتغير الطبقة هو المنطقة.