

SPSS İLE İSTATİSTİKSEL VERİ ANALİZİ

Statistical Packages for the Social Sciences



PROF.DR.YÜKSEL TERZİ

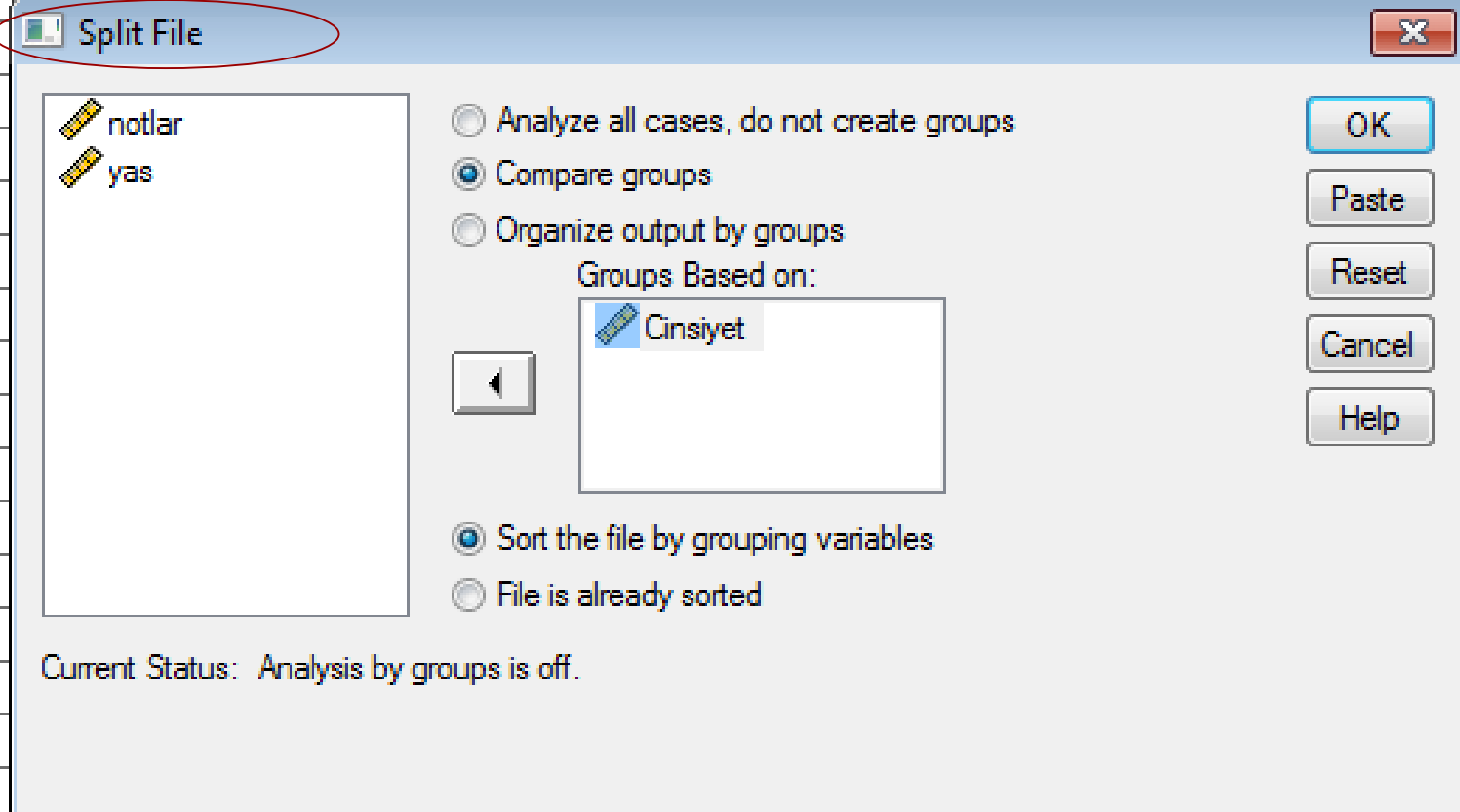
ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN-EDEBİYAT FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

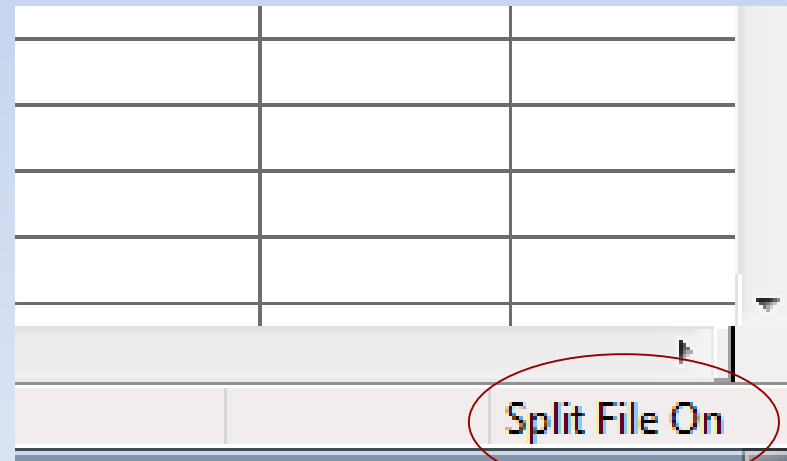
SAMSUN

2019



Split file:(Veri dosyası ayırma)
Veri dosyasını istenilen değişkenin alt seviyelerine göre parçalar. Böylece veri dosyası, değişken kaç farklı değer alıyorsa o sayıda alt dosyalara ayrılır.

Analysis all cases : Grublamanın olmadığı durum

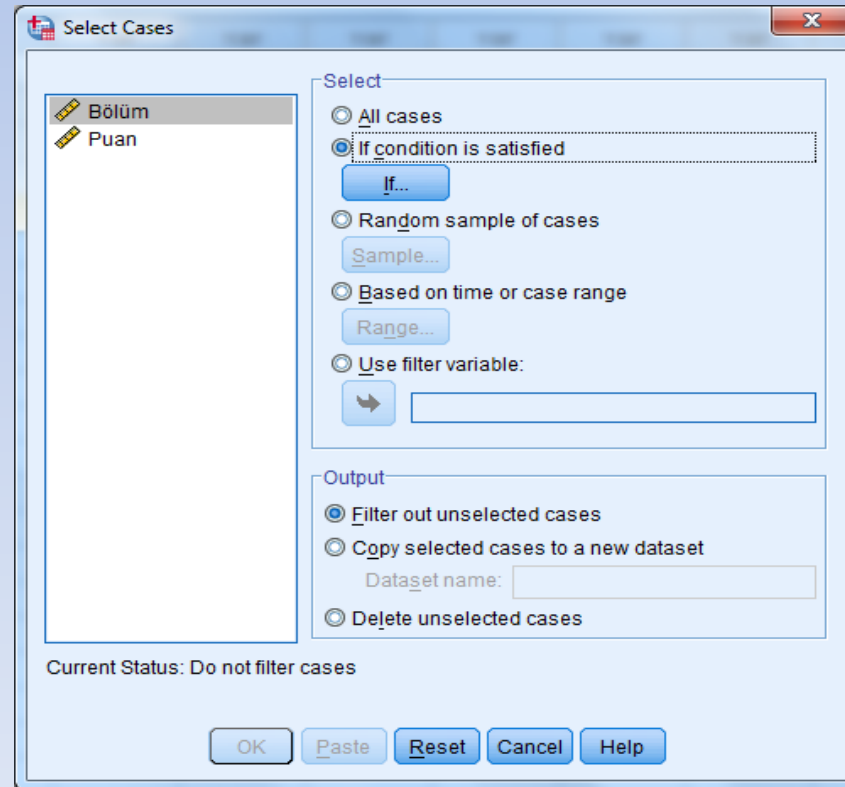


Select Cases (Belirli vakaların seçimi)

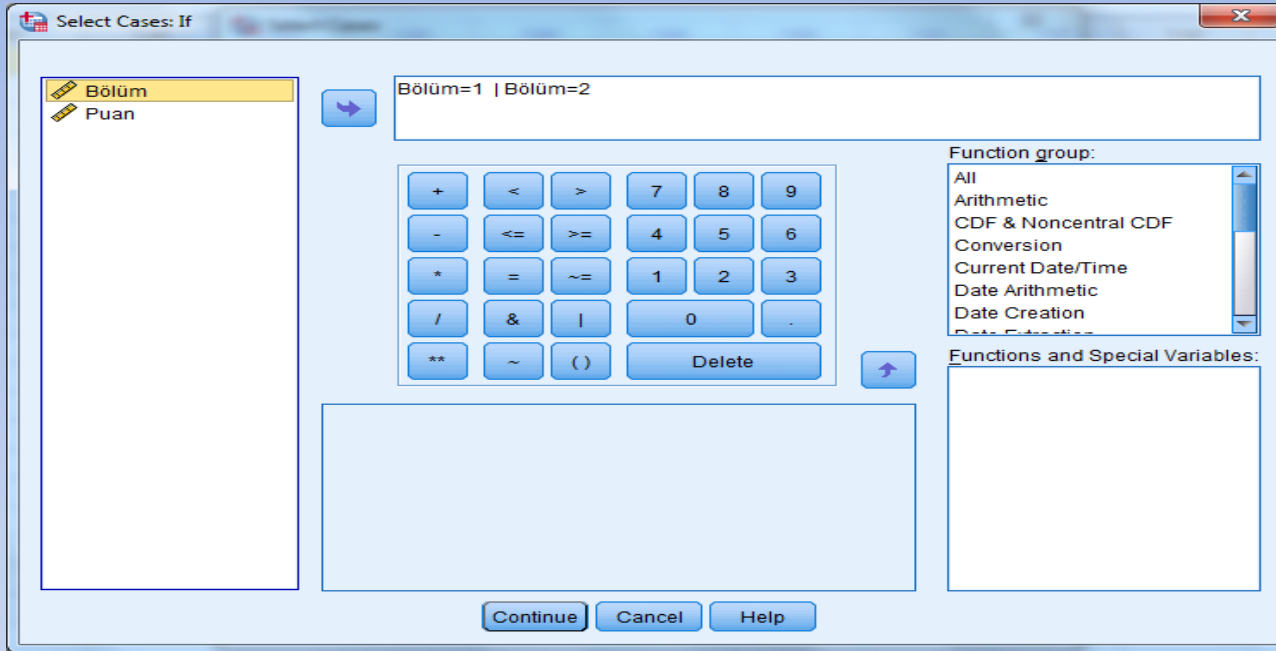
Birden çok kategoriye sahip bir nitel değişkenin bir veya daha çok kategorisi seçilerek, seçilen kategoriler için analiz işlemleri yapma imkanı sağlar.

If condition is satisfied: Belirli koşulu sağlayan vakaların seçimi

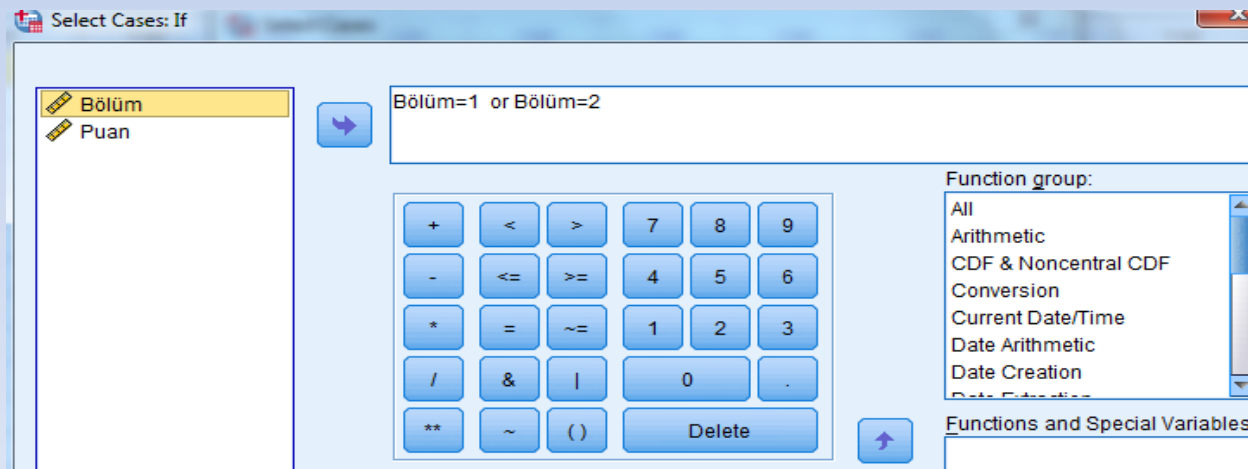
	Bölüm	Puan
1	Kimya	55,0
2	Kimya	50,5
3	Kimya	61,0
4	Kimya	62,0
5	Kimya	65,0
6	İstatistik	70,0
7	İstatistik	75,0
8	İstatistik	75,5
9	İstatistik	82,0
10	İstatistik	85,0
11	Matematik	45,0
12	Matematik	50,0
13	Matematik	40,0
14	Matematik	65,0
15	Matematik	60,0



Bölüm değişkeninden sadece Kimya (kod=1) ve İstatistik (kod=2) bölümlerinin seçilmesi.



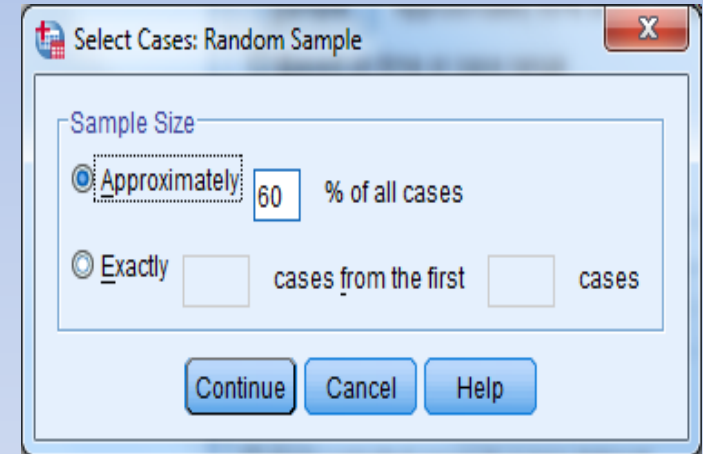
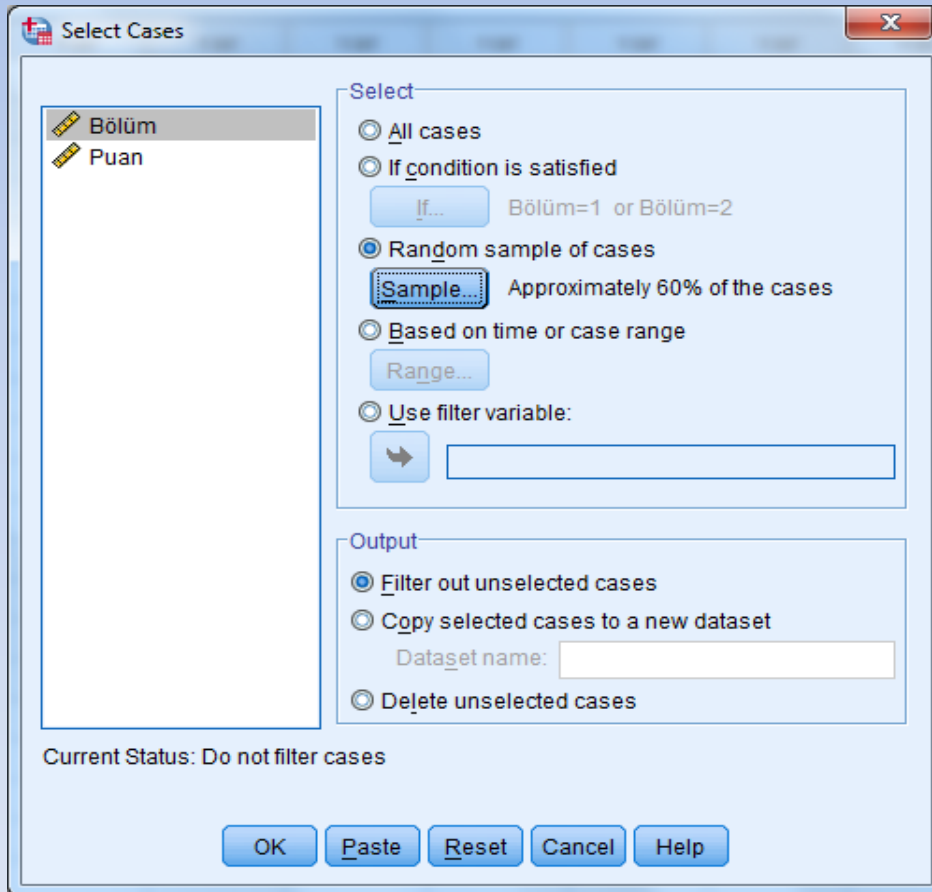
VEYA



	Bölüm	Puan	filter_Ş
1	Kimya	55,0	Selected
2	Kimya	50,5	Selected
3	Kimya	61,0	Selected
4	Kimya	62,0	Selected
5	Kimya	65,0	Selected
6	İstatistik	70,0	Selected
7	İstatistik	75,0	Selected
8	İstatistik	75,5	Selected
9	İstatistik	82,0	Selected
10	İstatistik	85,0	Selected
11	Matematik	45,0	Not Selected
12	Matematik	50,0	Not Selected
13	Matematik	40,0	Not Selected
14	Matematik	65,0	Not Selected
15	Matematik	60,0	Not Selected

Seçilen (select cases yapılan) kategoriler aktif (selected), seçilmeyen kategoriler ise üzeri çizilerek iptal (not selected) edilir. Puan değişkeninin seçilen kimya ve istatistik bölümü öğrencilerinin (toplam 10 kişi) tanımlayıcı istatistikleri aşağıdaki gibi bulunur.

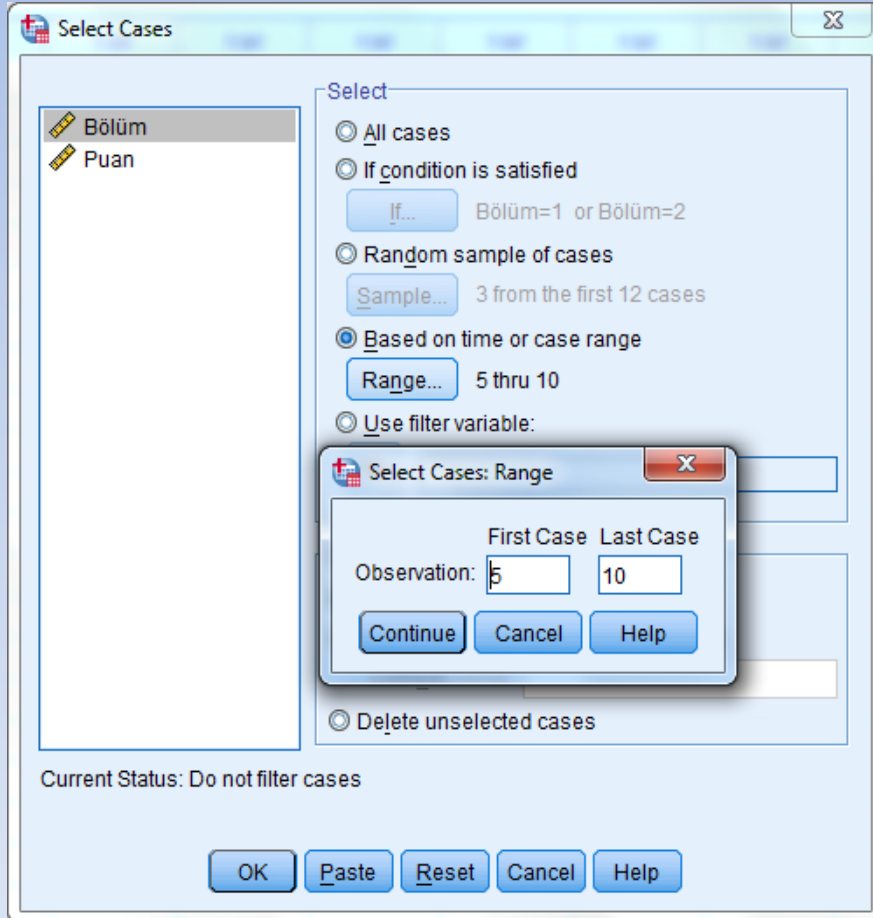
Random sample of cases: Vakalar arasından şansa bağlı olarak bazı vakaların seçilmesi. **Select cases** komutu ile veri setinden rastgele örneklem seçilebilir. Yaklaşık olarak (Approximately) ne kadar veri seçileceği girilen yüzde değere yakın bir sayıda rastgele örneklem seçilir.



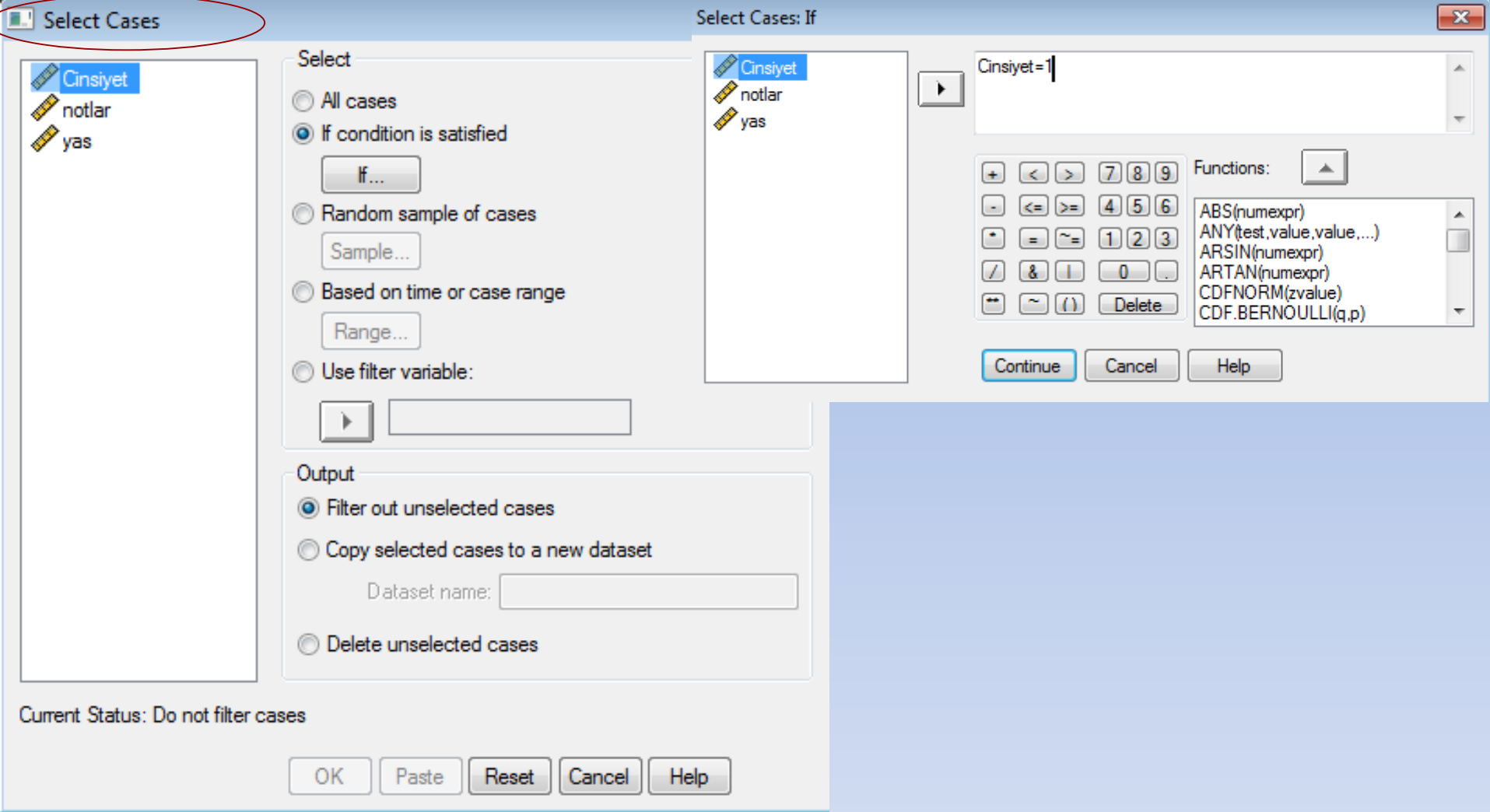
Bir başka seçenekte kesin vaka nımarası belirtilmesi şeklinde seçim yapılabilir. Bunun için Exactly (kaç adet) ve cases from the first (ilk Sayıdaki kayıt içinden) şeklinde tanım yapılır.

	Bölüm	Puan	filter_S
1	Kimya	55,0	1
2	Kimya	50,5	1
3	Kimya	61,0	1
4	Kimya	62,0	1
5	Kimya	65,0	1
6	İstatistik	70,0	0
7	İstatistik	75,0	1
8	İstatistik	75,5	0
9	İstatistik	82,0	1
10	İstatistik	85,0	0
11	Matematik	45,0	0
12	Matematik	50,0	1
13	Matematik	40,0	1
14	Matematik	65,0	1
15	Matematik	60,0	1

Belirli bir vakadan itibaren belirli sayıda örneklemede yapılabilir. Bunun için **Based on time or case range**: İki vaka numarası veya iki tarih arasındaki gözlem (başlangıç ve bitiş) arasında verileri seçme imkanı sağlar.



	Bölüm	Puan
1	Kimya	55,0
2	Kimya	50,5
3	Kimya	61,0
4	Kimya	62,0
5	Kimya	65,0
6	İstatistik	70,0
7	İstatistik	75,0
8	İstatistik	75,5
9	İstatistik	82,0
10	İstatistik	85,0
11	Matematik	45,0
12	Matematik	50,0
13	Matematik	40,0
14	Matematik	65,0
15	Matematik	60,0



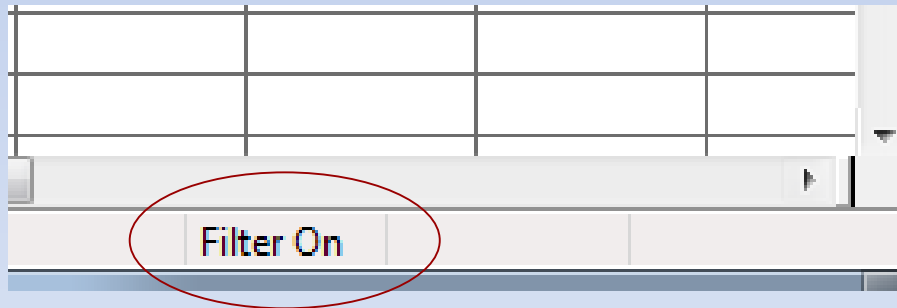
Select cases :Veri süzme veya seçme-

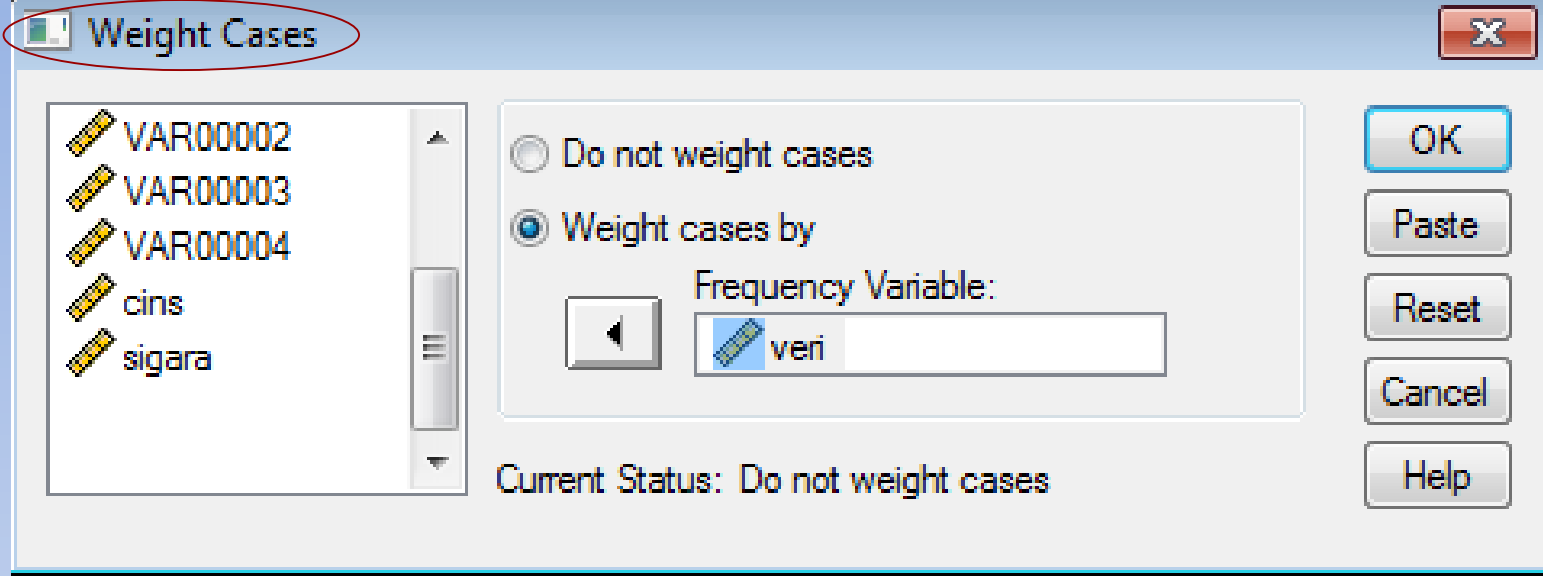
Farklı değerler alan değişkenlerin istenilen değerlerinin analize katılmasını sağlar.

If condition is satisfied : Değişkenlerle ilgili hesap ve fonksiyon işlemleri yapar.

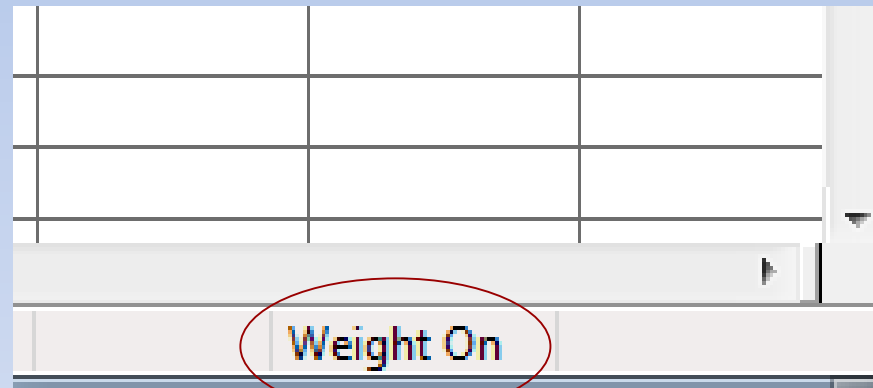
Random sample of cases : Bazı örnekler üzerinde işlem yapar.

	Cinsiyet	notlar	yas	filter \$	
1	1	45	20	1	
2	1	65	15	1	
3	1	70	18	1	
4	1	90	22	1	
5	2	25	25	0	
6	2	30	32	0	
7	2	60	38	0	
8	2	50	43	0	
9					





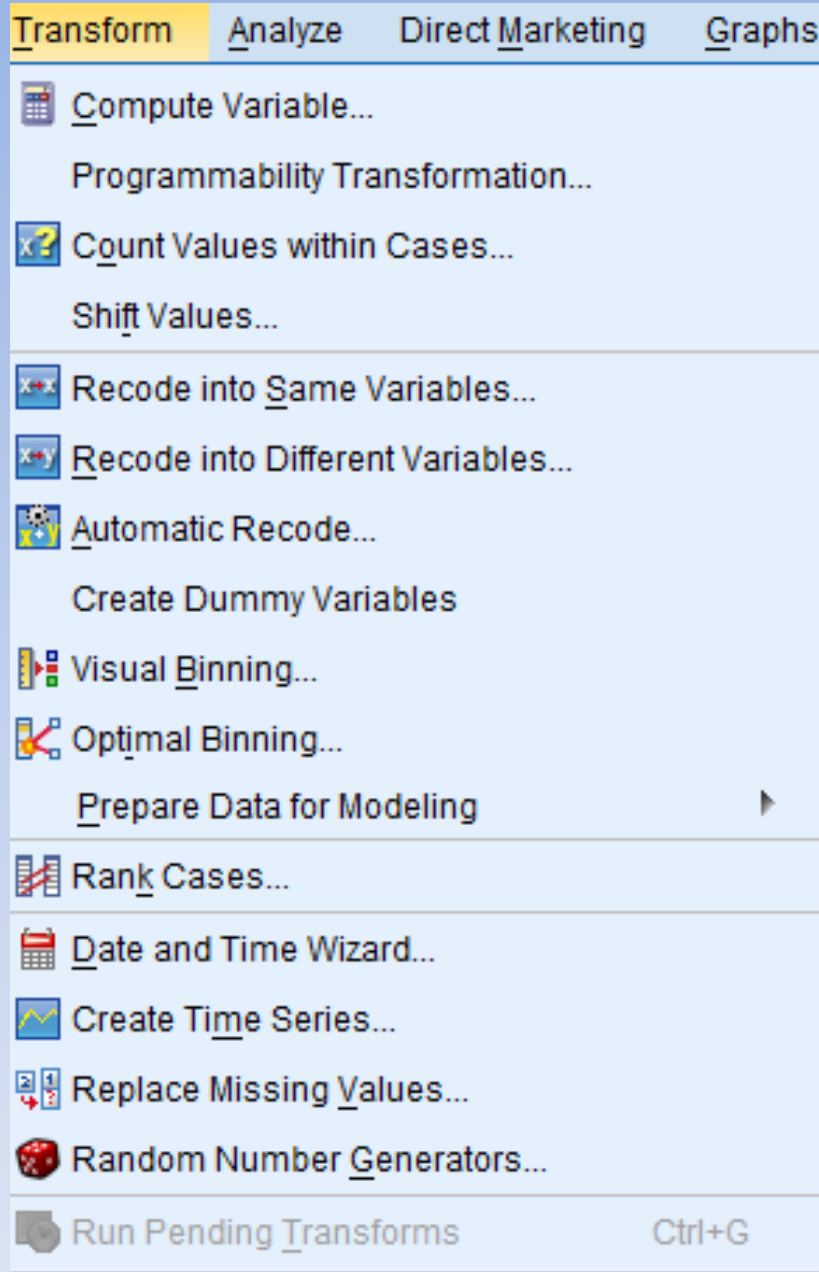
cins	sigara	veri
1	1	45
1	2	50
2	1	60
2	2	30



Do not weight cases : Ağırlıklı durum yok

Weight cases by – Frequency Variable:

Bir sütunda yer alan değerlerin diğer bir değişken için ağırlık olarak tanımlanmasını sağlar.



Compute :Mevcut deęişkenlerle matematiksel işlem yaparak, yeni deę. elde edilir.

Target variable : Yapılan işlem sonunda bulunacak yeni deęişkenin adı.

Random number seed : Rasgele üretilecek olan sayıların ilk deęerini belirler.

Count :Bir deęişkendeki aynı deęere sahip birimleri belirler.

Categorize variables : Verilere sıra numarası verilir.

Rank cases :Deęişkenleri sıralama deęerine dönüştürür ve yeni deęişken olarak atar.

Deęişik sırala türleri vardır.(Blom, Tukey, Rankit gibi)

Automatic Recode : Deęişkenin deęerlerini büyüklük sırasına göre dizerek yeni deęişken 1 den başlayarak kod numarası verir.

Create Time Series : Bir veriyi alarak zaman serisi deęişkeni türetir.

Difference : Farklı **Seasonal** : Mevsimlik

Smooth : Düzgün

Replace Missing Values : Seride yer alan eksik gözlem yerine bir deęer atar. Bu deęer ya serinin ortalaması yada zaman serisi fonksiyonlarından birisidir.

Compute Variable (Yeni hesap deęiřkeni tanımlama)

Hesaplamalarla yeni deęiřkenler oluřturmak için bu ekran kullanılır. Deęiřkenlerle ilgili her türlü matematiksel hesaplamaların yapıldığı bir komuttur.

Örnek . 15 öğrencinin puanları aşağıdaki gibidir. Bu deęerlerin 10 tabanına göre logaritmasını ve kareköklerini özel tanımlı fonksiyonlar (LG10 ve SQRT) kullanılarak alınabilir.

Puan
55,0
50,5
61,0
62,0
65,0
70,0
75,0
75,5
82,0
85,0
45,0
50,0
40,0
65,0
60,0

Compute Variable

Target Variable: logPuan = Numeric Expression: LG10(Puan)

Type & Label...

Bölüm
Puan

Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation

Functions and Special Variables:
Abs
Arasin
Artan
Cos
Exp
Lg10
Ln
Lgamma
Mod
Rnd(1)
Rnd(2)

LG10(numexpr). Numeric. Returns the base-10 logarithm of numexpr, which must be numeric and greater than 0.

if... (optional case selection condition)

OK Paste Reset Cancel Help

Puan	logPuan
55,0	1,74
50,5	1,70
61,0	1,79
62,0	1,79
65,0	1,81
70,0	1,85
75,0	1,88
75,5	1,88
82,0	1,91
85,0	1,93
45,0	1,65
50,0	1,70
40,0	1,60
65,0	1,81
60,0	1,78

Recode Into Same-Different Variables (Değer kodlaması yapmak)

Bir değişkene ait verileri matematiksel işlem yapmadan isteğe göre değiştirmek veya yeni kodlarla ifade etmede kullanılır. Nicel olan bir değişken gruplanarak kategorik hale getirilebilir.

Nicel değişken olan **numeric variable** kısmına girilir ve **output variable** kısmına da yeni oluşturulacak değişkenin adı girilir ve **change** ile tanımlanın hangi değişkenler arasında yapılacağı tanımlanır. Daha sonra eski (old) ve yeni (new) değer tanımı (**Old and New Values**) kısmından kodlamalar tanımlanır.

Recode into same variables seçilirse yeni elde edilen kodlar eski değişkenin bulunduğu sütuna yazılacağı için eski değişken kaybolur. **recode into different variables** seçilirse orijinal değişken aynen kalır, değişiklikler yeni değişken üzerinde tanımlanarak gösterilir.

Örnek. 15 öğrencinin aylık yapmış oldukları masraflar (TL) aşağıdaki gibidir. Bu verileri aşağıdaki gibi kodlayalım.

Masraf:200,225,250,275,300,350,500,650,700,725,750,800,850,900, 1000

≤ 250	(250-500]	(500-750]	> 750
1	2	3	4

	Masraf
1	200
2	225
3	250
4	275
5	300
6	350
7	500
8	650
9	700
10	725
11	750
12	800
13	850
14	900
15	1000

Recode into Different Variables

Numeric Variable -> Output Variable:
Masraf --> Masrafkod

Output Variable
Name:
Masrafkod
Label:

(optional case selection condition)

SPSS Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:

through

Range, LOWEST through value:

250

Range, value through HIGHEST:

All other values

New Value

Value: 1

System-missing

Copy old value(s)

Old --> New:

Add

Change

Remove

Output variables are strings

Convert numeric strings to

Continue Cancel Help

SPSS Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:

250

through

500

Range, LOWEST through value:

Range, value through HIGHEST:

All other values

New Value

Value: 2

System-missing

Copy old value(s)

Old --> New:

Lowest thru 250 --> 1

Add

Change

Remove

Output variables are strings

Convert numeric strings to

Continue Cancel Help

SPSS Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:

through

Range, LOWEST through value:

Range, value through HIGHEST:
1000

All other values

New Value

Value: 4

System-missing

Copy old value(s)

Old --> New:

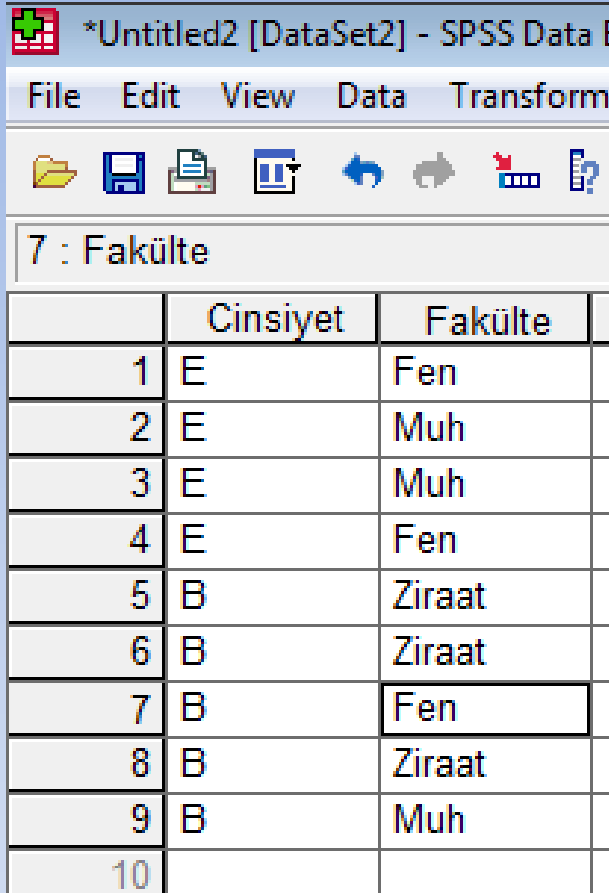
Lowest thru 250 --> 1
250 thru 500 --> 2
500 thru 750 --> 3

Output variables are strings

Convert numeric strings to numbers

	Masraf	Masrafkod
1	200	1
2	225	1
3	250	1
4	275	2
5	300	2
6	350	2
7	500	2
8	650	3
9	700	3
10	725	3
11	750	3
12	800	4
13	850	4
14	900	4
15	1000	4

Automatic Recode

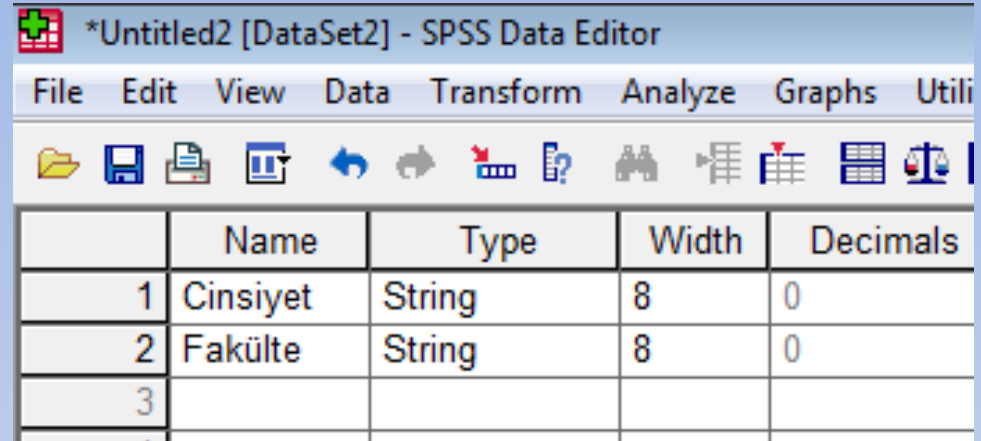


*Untitled2 [DataSet2] - SPSS Data Editor

File Edit View Data Transform

7 : Fakülte

	Cinsiyet	Fakülte
1	E	Fen
2	E	Muh
3	E	Muh
4	E	Fen
5	B	Ziraat
6	B	Ziraat
7	B	Fen
8	B	Ziraat
9	B	Muh
10		



*Untitled2 [DataSet2] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utili

	Name	Type	Width	Decimals
1	Cinsiyet	String	8	0
2	Fakülte	String	8	0
3				

String olarak girilen nitel değişkenleri, Automatic recode komutu ile kodlayarak, Numeric değişkene çevrilir.

Automatic Recode

Variable -> New Name

Cinsiyet -> Cinskod
 Fakülte -> ????????

New Name: Fakkod

Recode Starting from
 Lowest value Highest value

Use the same recoding scheme for all variables
 Treat blank string values as user-missing

Template
 Apply template from: File...
 Save template as: File...

OK
 Paste
 Reset
 Cancel
 Help

Add New Name

Cinskod	Fakkod
2	1
2	2
2	2
2	1
1	3
1	3
1	1
1	3
1	2

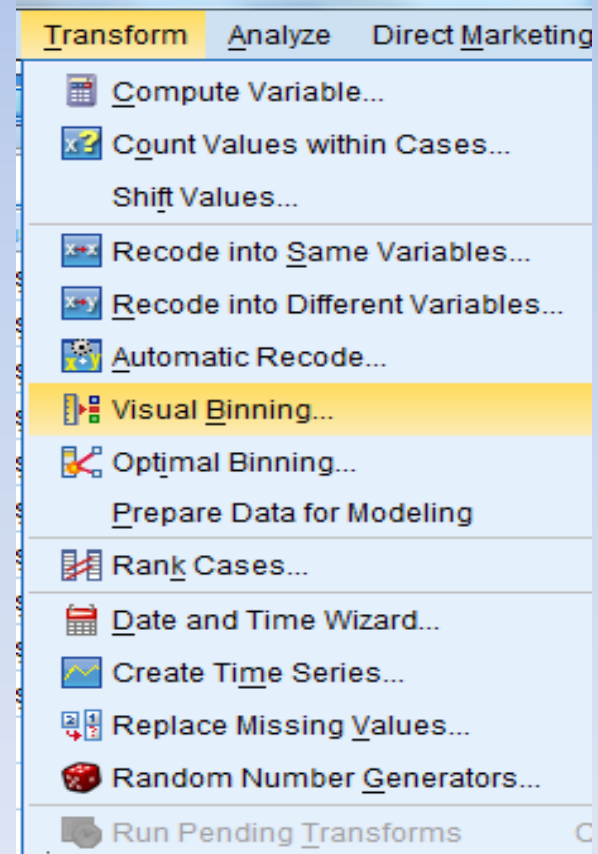
New name kısmına string değişkenlerin yani ismi yazılır ve Add New Name butonuna basılır.

Lowest value işaretlenirse string değişkenin aldığı değerlere baştan sona doğru kod numarası verilir (B=1, E=2 olur). Highest value seçilirse sondan aşağı doğru kod numarası verilir (E=1, B=2).

Visual Binning (Görsel gruplama yapmak)

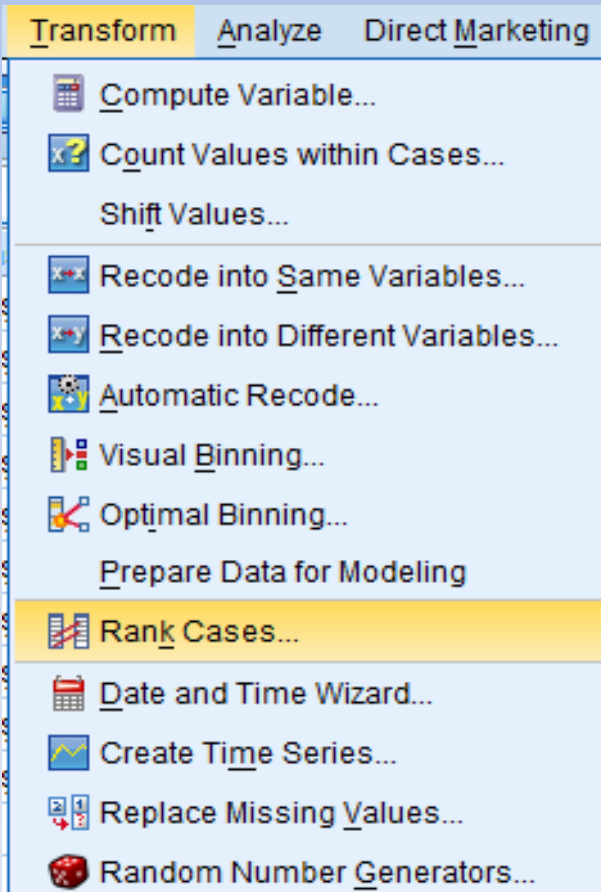
Optimal Binning (Optimum gruplama yapmak)

Prepare Data For Modeling (Modelleme için veri hazırlama)

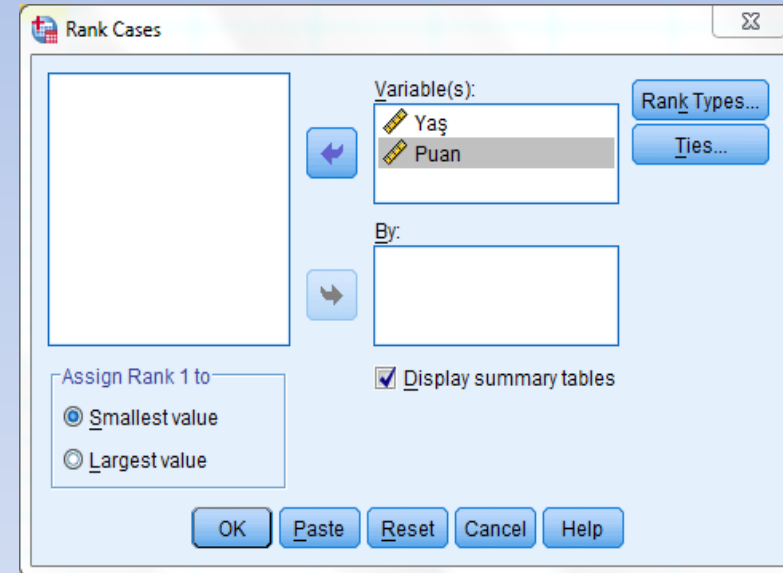


Rank Cases (Vakaların sıralanması)

Gözlem değerlerine sıra puanları verilir. Sıra puanları en küçük gözlemden (smallest value) başlanarak veya en büyük gözlemden başlanarak (largest value) verilebilir. Aynı gözlemden birden fazla varsa **Ties** kısmından girilerek sıra puanlarının ortalaması alınacaksa **Mean** seçilir.



	Yaş	Puan
1	22	55
2	19	65
3	24	70
4	21	80
5	22	90
6	20	95
7	19	70
8	19	50
9	18	45
10	21	40
11	25	50
12	20	60
13	24	75
14	21	70
15	18	80



Rank Cases: Types

Rank Fractional rank as %

Save score Sum of case weights

Fractional rank Ntiles: 4

Proportion estimates Normal scores

Proportion Estimation Formula

Blom Tukey Rankit Van der Waerden

Continue Cancel Help

Rank Cases: Ties

Rank Assigned to Ties

Mean Low High

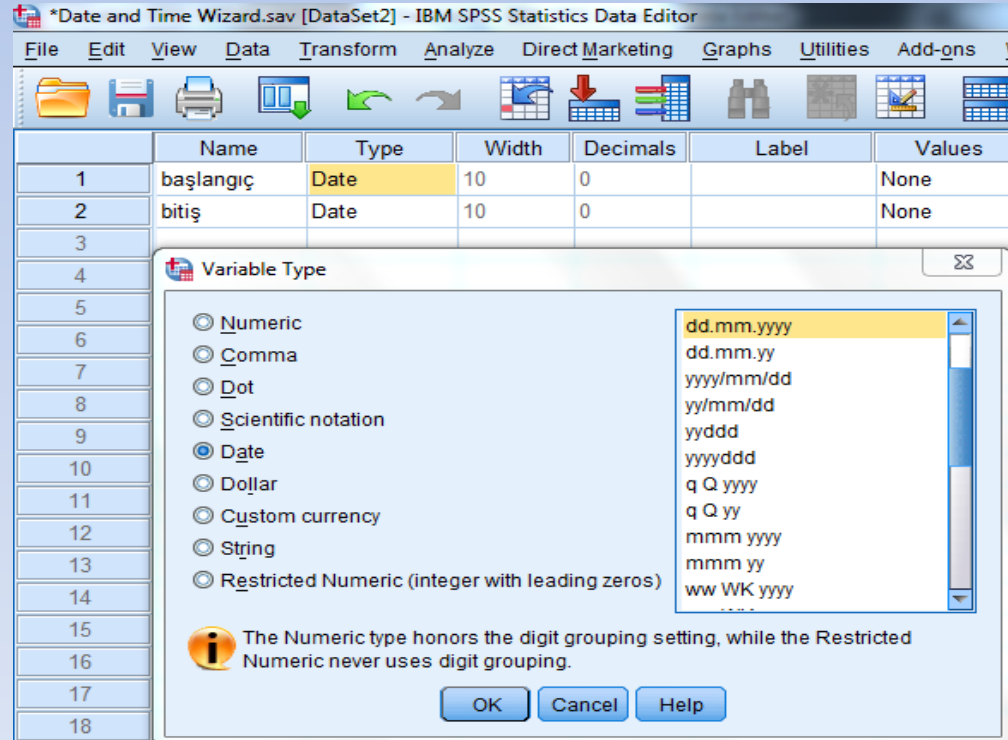
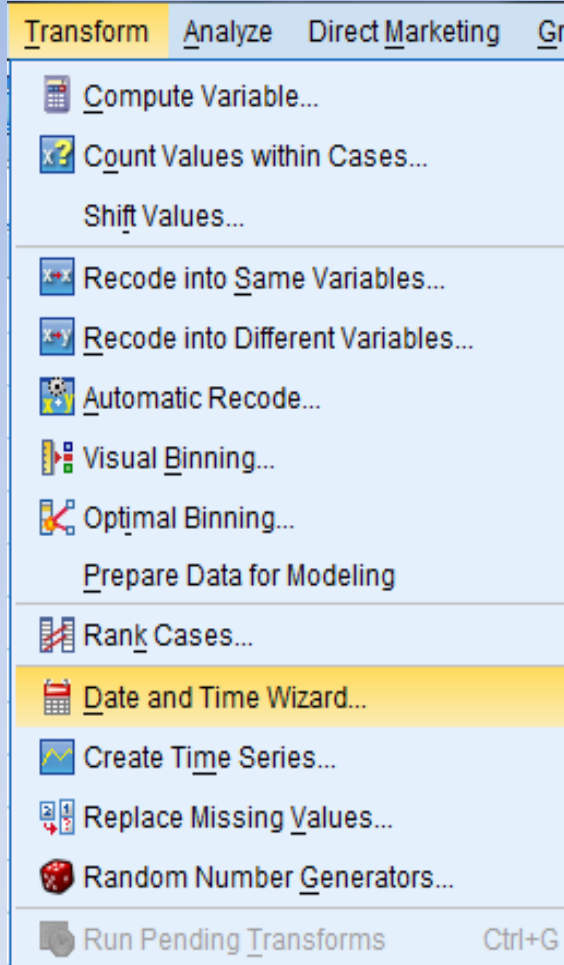
Sequential ranks to unique values

Continue Cancel Help

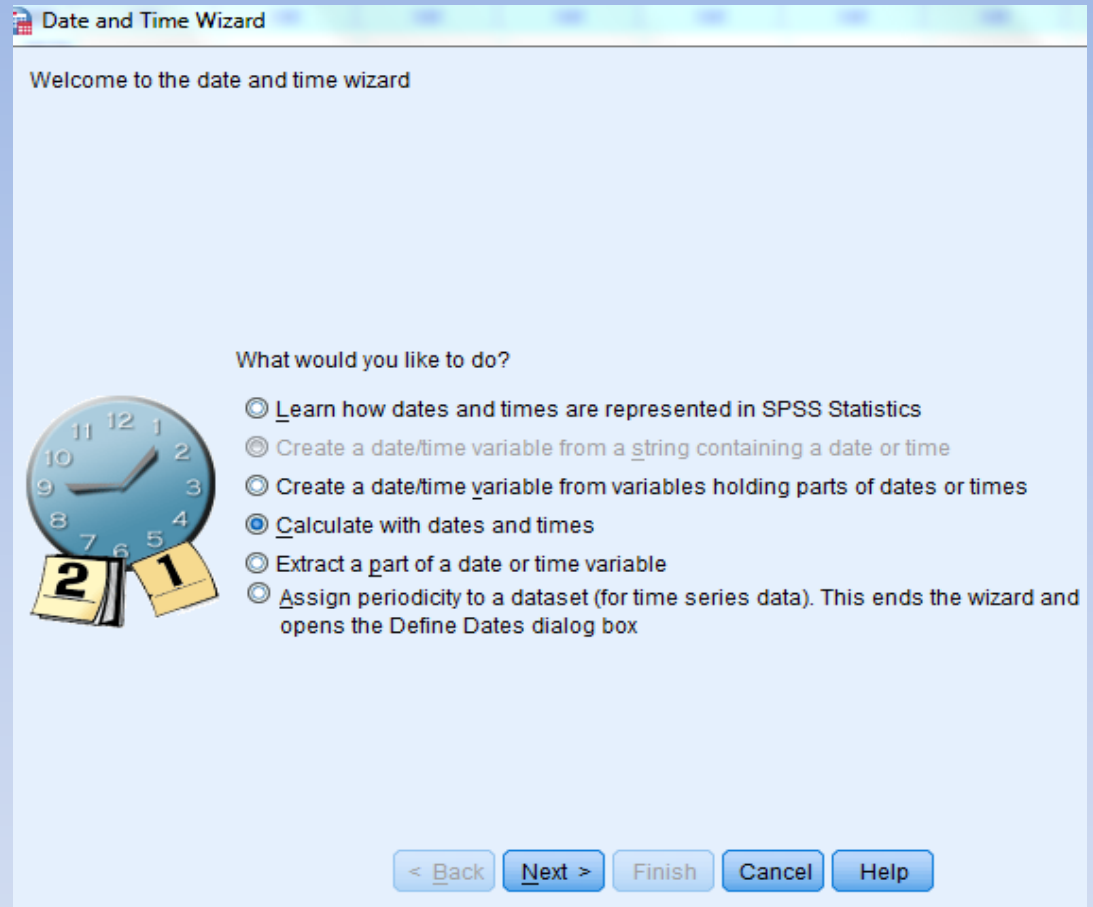
	Yaş	Puan	RPuan	RYaş
1	22	55	5,0	11,5
2	19	65	7,0	4,0
3	24	70	9,0	13,5
4	21	80	12,5	9,0
5	22	90	14,0	11,5
6	20	95	15,0	6,5
7	19	70	9,0	4,0
8	19	50	3,5	4,0
9	18	45	2,0	1,5
10	21	40	1,0	9,0
11	25	50	3,5	15,0
12	20	60	6,0	6,5
13	24	75	11,0	13,5
14	21	70	9,0	9,0
15	18	80	12,5	1,5

Date and Time Wizard (Tarih ve saat düzenlemesi)

Örnek. Tarih değişkeni olarak tanımlanan iki değişkenin başlangıç ve bitiş tarihleri arasında geçen süre yıl, ay, hafta, gün, saat gibi zaman dilimi olarak hesaplayalım.



	başlangıç	bitiş
1	05.02.1998	18.09.2008
2	25.11.2000	16.10.2010
3	27.09.1970	16.11.2012
4	06.02.1985	04.12.2011
5	12.12.2009	14.09.2012



Do Calculations on Dates

Choose one of the following tasks and press Next

- Add or subtract a duration from a date (e.g., add a month to an age or add a time variable to a date/time variable)
- Calculate the number of time units between two dates (e.g., calculate an age in years from a birthdate and another date)
- Subtract two durations (e.g., time worked - time commuting)



< Back Next > Finish Cancel Help

Calculate the number of time units between two date or date/time variables.

The result will be an integer variable. Any fractional part of a unit will be discarded. The result will be a duration variable. Only duration variables are shown in the variables list below.



Variables:

Current date and ti...

Date 1:

bitiş

minus Date2:

başlangıç

Unit:

Days

Result Treatment

- Truncate to integer
- Round to integer
- Retain fractional part

For month and year units, the result is based on average unit length unless truncation is used.

\$TIME is the current date and time.

< Back Next > Finish Cancel Help

Calculation: bitiş - başlangıç

Result Variable:

Units:

Days

Variable Label:

Execution

Create the variable now Paste the syntax into the syntax window



< Back

Next >

Finish

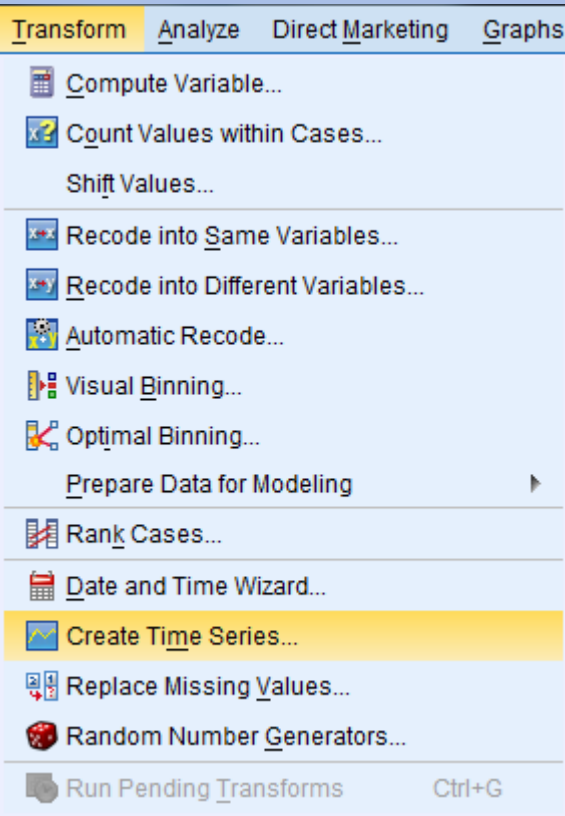
Cancel

Help

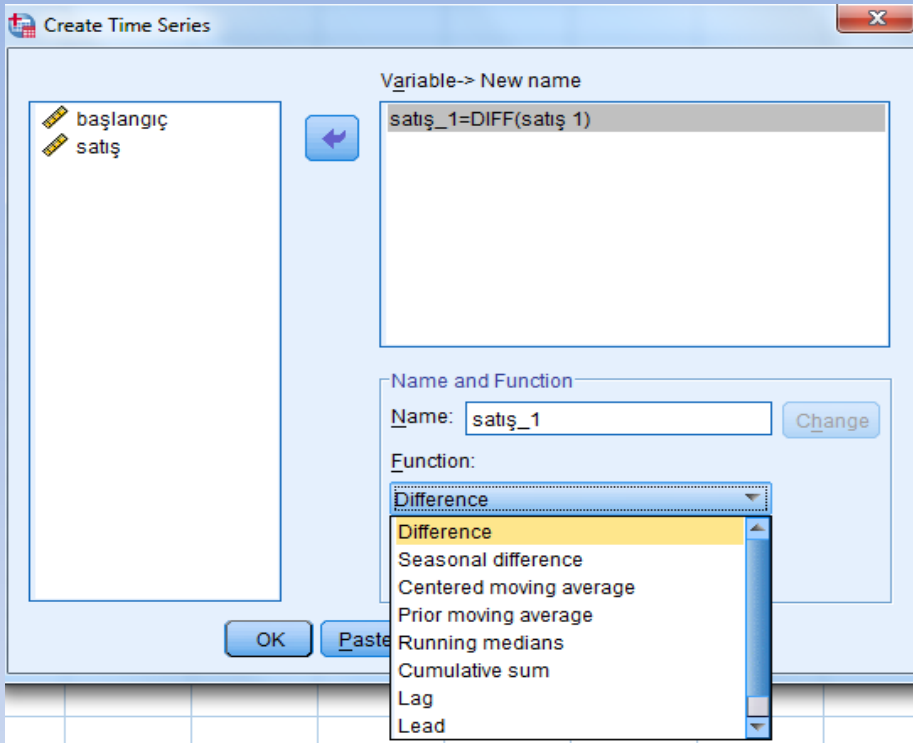
	başlangıç	bitiş	gün_sayısı
1	05.02.1998	18.09.2008	3878
2	25.11.2000	16.10.2010	3612
3	27.09.1970	16.11.2012	15391
4	06.02.1985	04.12.2011	9797
5	12.12.2009	14.09.2012	1007

Date and Time Wizard (Tarih ve saat düzenlemesi)

Zaman serisi oluşturmada kullanılır.



	başlangıç	satış
1	January	45
2	February	50
3	March	58
4	April	65
5	May	70
6	June	73
7	July	77
8	August	83
9	September	86
10	October	90
11	November	92
12	December	98



	başlangıç	satış	satış_1
1	January	45	.
2	February	50	5
3	March	58	8
4	April	65	7
5	May	70	5
6	June	73	3
7	July	77	4
8	August	83	6
9	September	86	3
10	October	90	4
11	November	92	2
12	December	98	6