

SPSS İLE İSTATİSTİKSEL VERİ ANALİZİ

Statistical Packages for the Social Sciences



PROF.DR.YÜKSEL TERZİ

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN-EDEBİYAT FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

SAMSUN

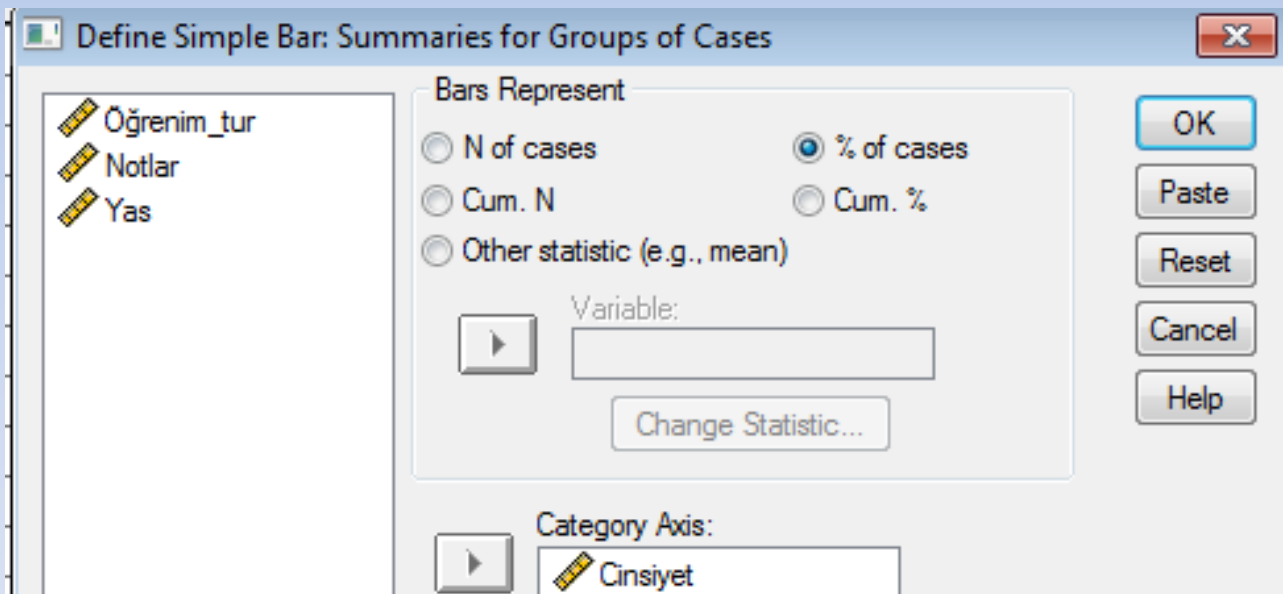
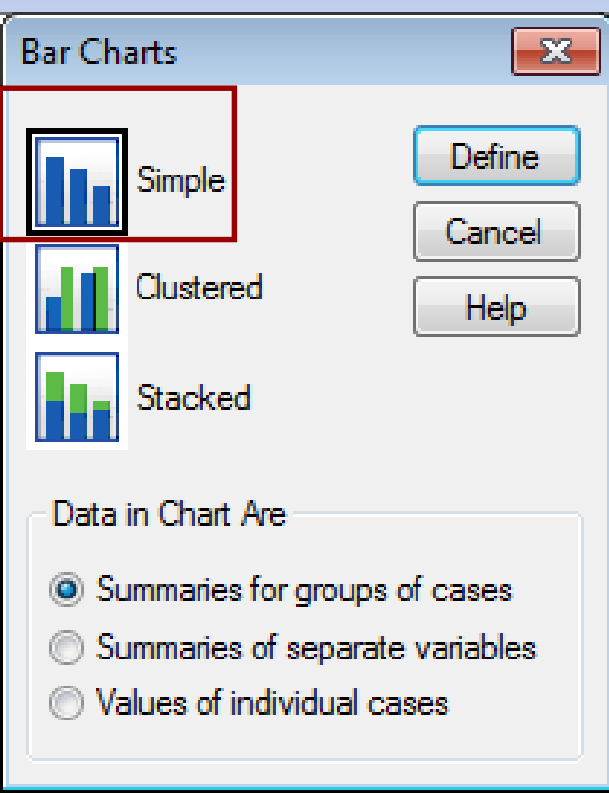
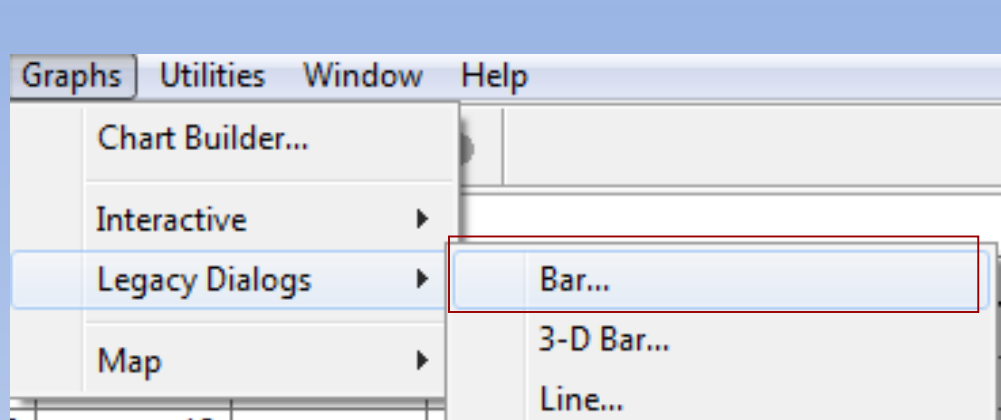
2019

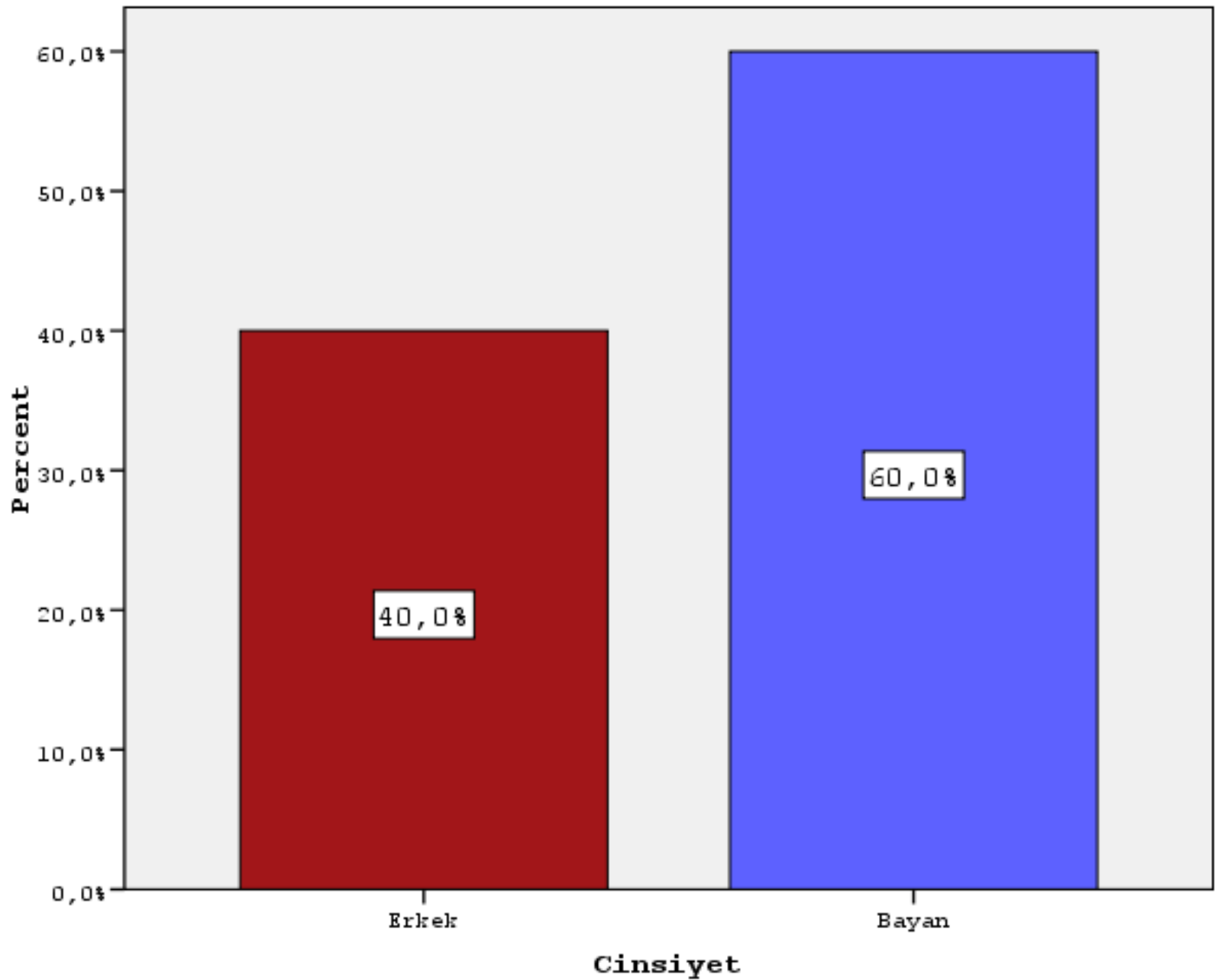
GRAFİKLER

Grafikler gözlem sonuçlarının anlaşılması kolaylaştırır. Çünkü ilk bakışta görsel kolaylık sağlarlar. Ancak grafikler rakamlar derecesinde önemli değildir. Çünkü grafiksiz istatistik olabilir ancak rakamsız grafik asla olamaz. Ayrıca grafikler büyük farkları ve değişimleri gösterebilir. Ancak küçük farklar ve değişimleri gösteremez. Çeşitli grafik türleri vardır.


Bar (Çubuk) Grafiği: Nominal ve ordinal verilerin grafiklerinin yapılması için uygundur. Bu grafik yatay ekseninde kategori etiketleri belirtilirken, her kategori ile frekans veya yüzdeyi temsil eden dikey çubuklar çiziminde oluşur.


Cinsiyet	Öğrenim tur	Notlar	Yas
1	1	45	20
1	1	65	15
1	1	70	18
1	2	90	22
2	2	25	25
2	2	30	19
2	2	60	24
2	2	50	21
2	1	75	23
2	1	65	24






Bar Charts Define Clustered Bar: Summaries for Groups of Cases

 Simple Define

 Clustered Cancel

 Stacked Help

Data in Chart Are

- Summaries for groups of cases
- Summaries of separate variables
- Values of individual cases

Bars Represent

- N of cases
- % of cases
- Cum. N
- Cum. %
- Other statistic (e.g., mean)

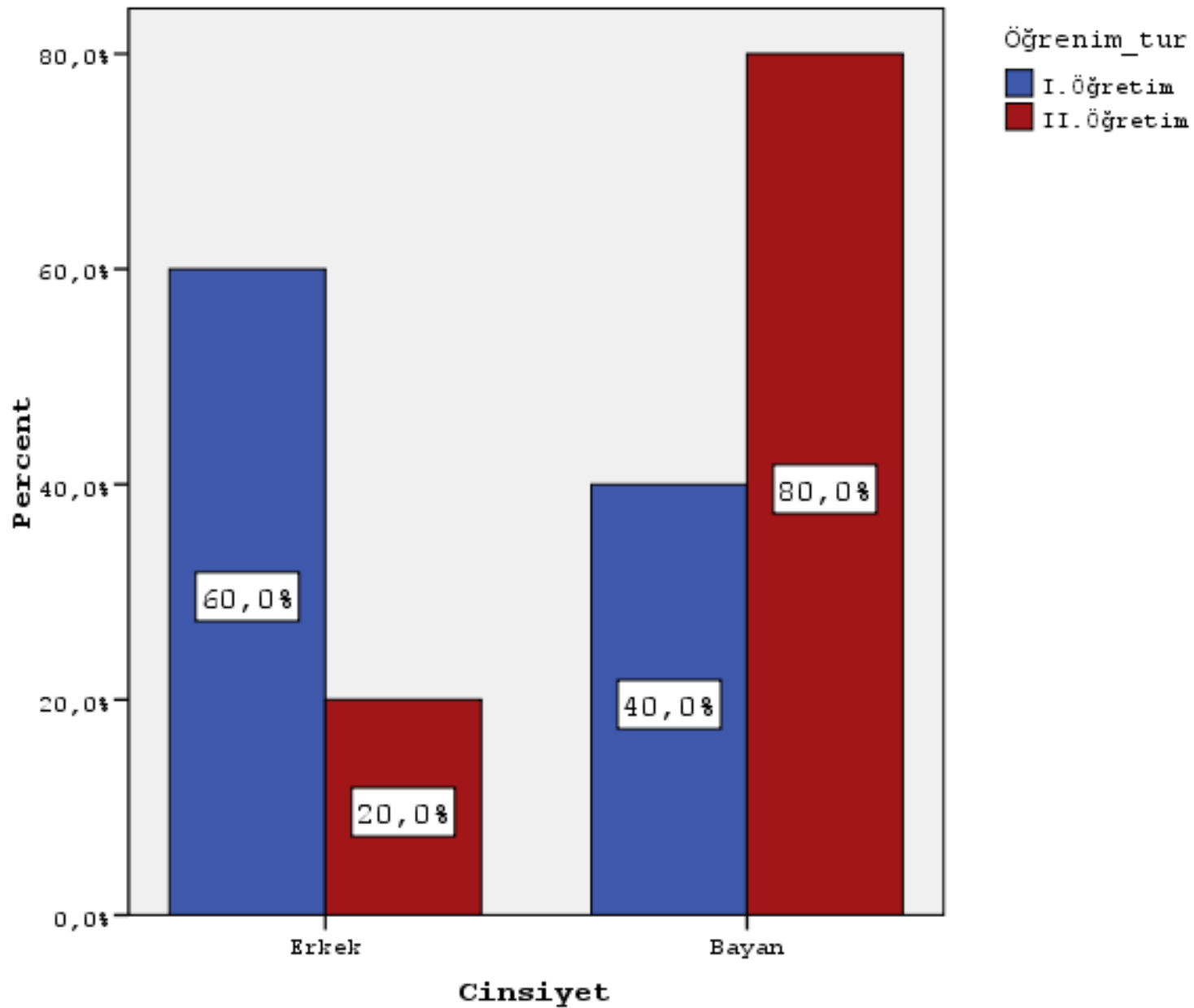
Variable: Change Statistic...

Category Axis:

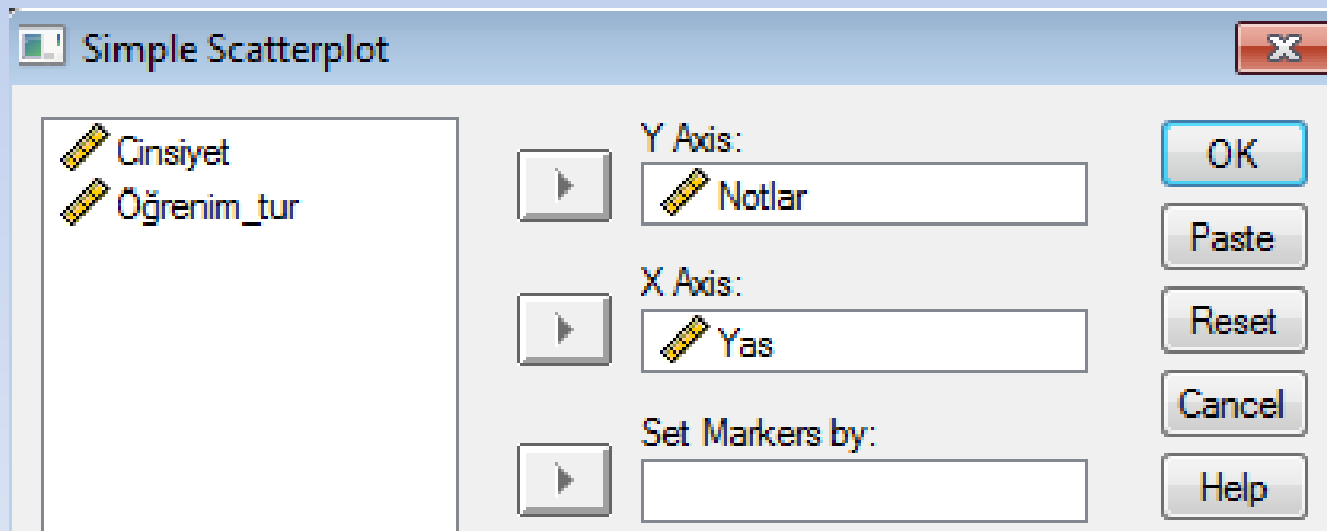
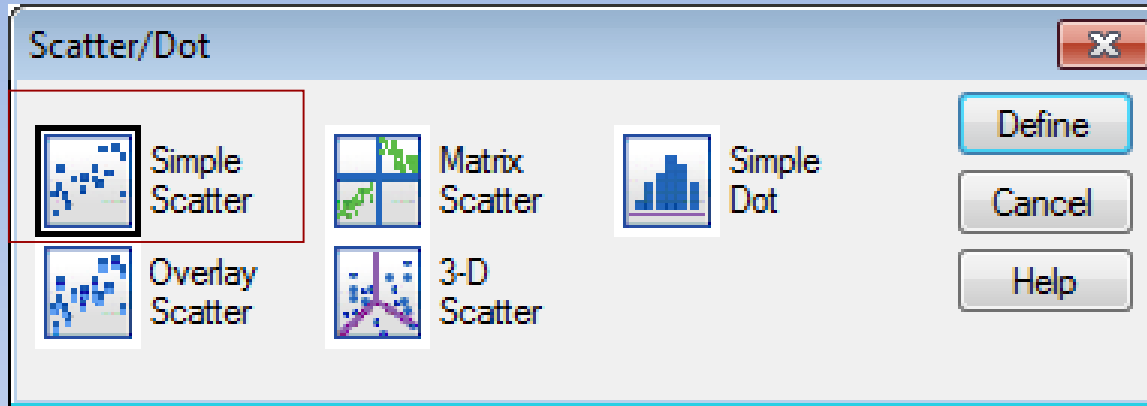
Define Clusters by:

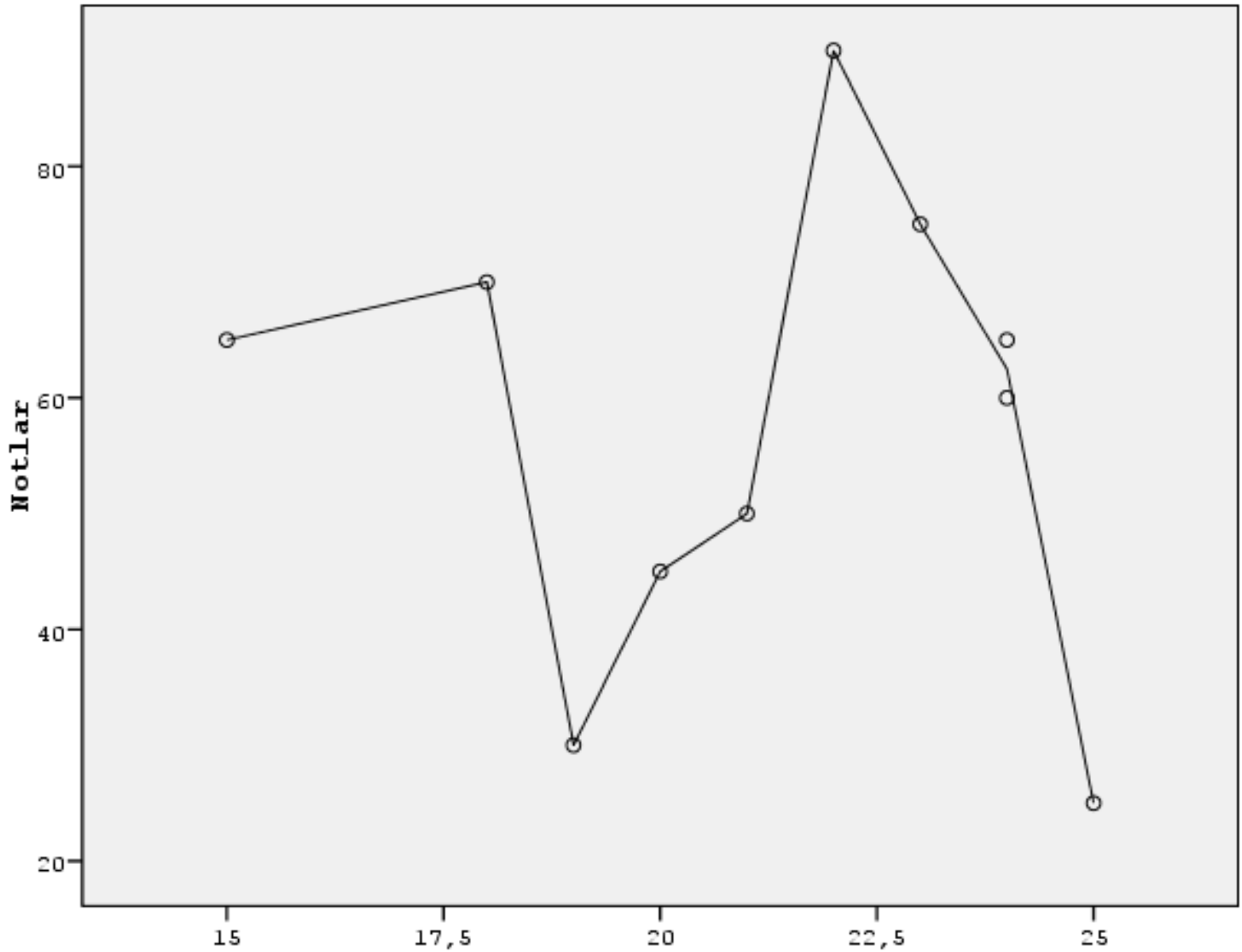
Notlar
Yas

OK Paste Reset Cancel Help

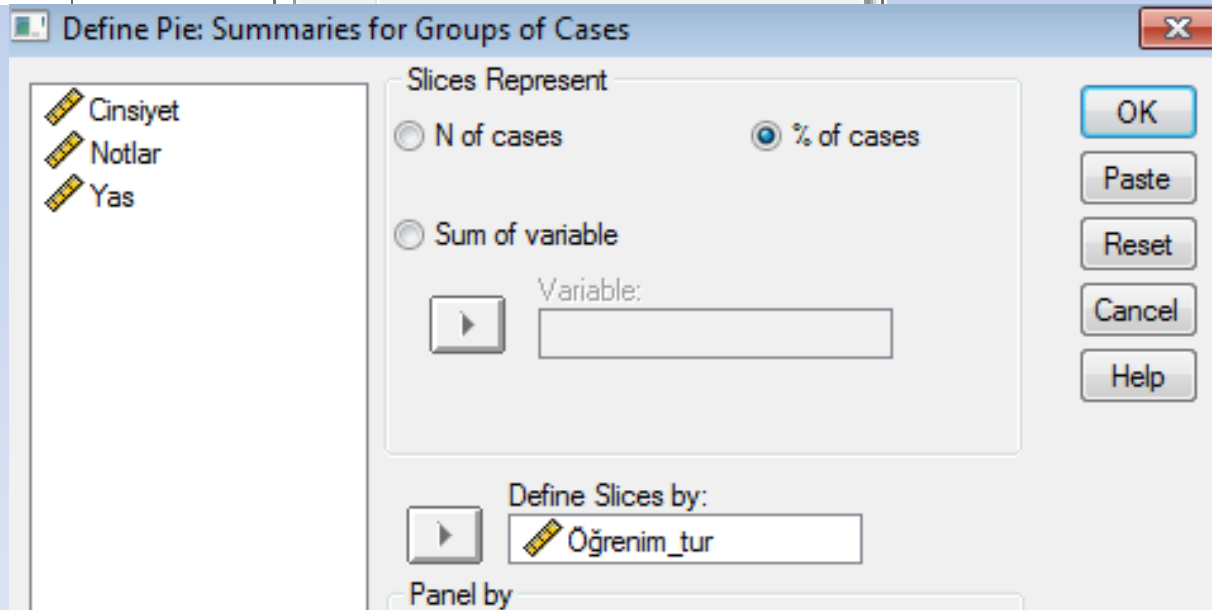
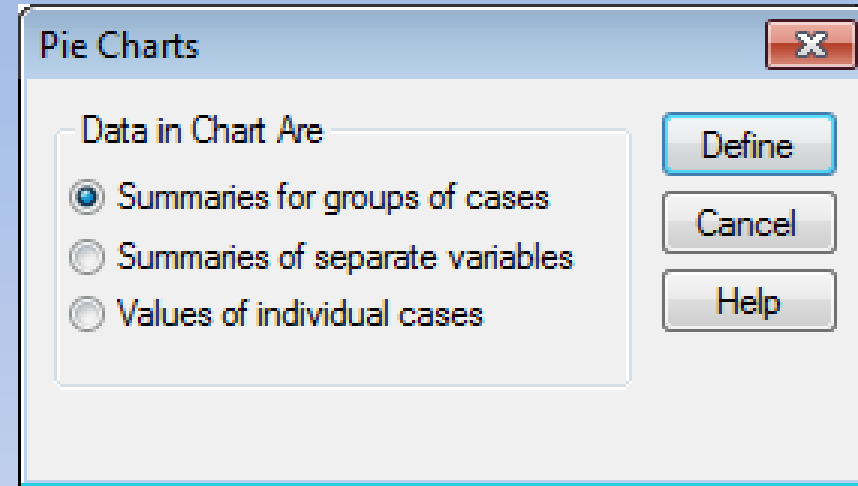
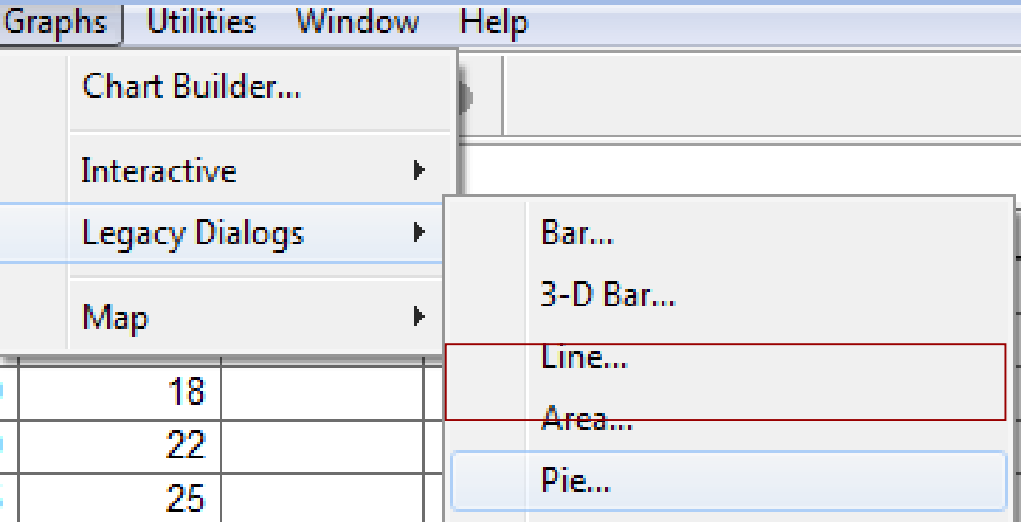


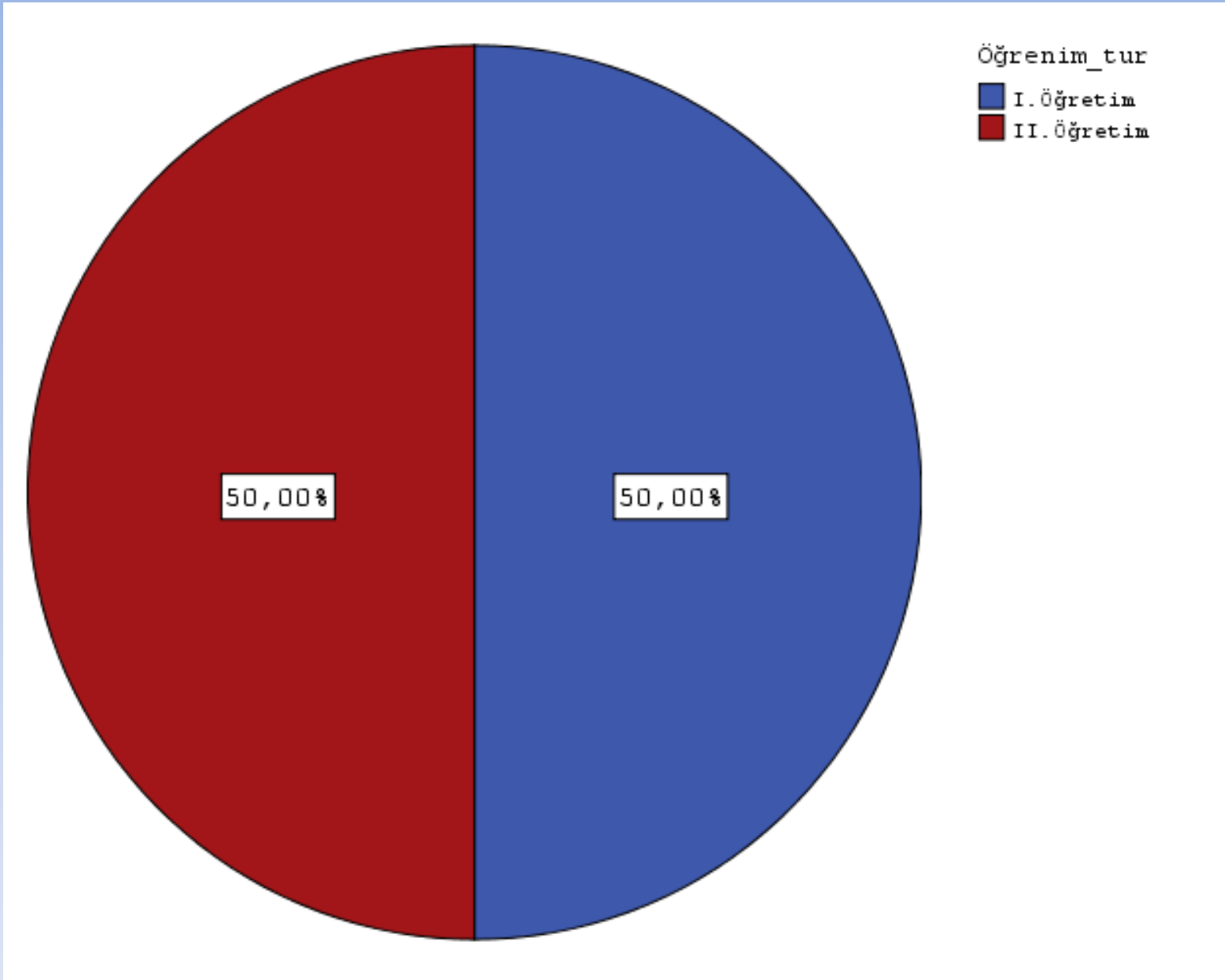
Scatter/Dot Grafiđi








Pie (pasta) Grafikleri





Histogram Grafikleri

Histogram [X]

 Cinsiyet
 Öğrenim_tur
 Yas

▶ Variable:

Display normal curve

Panel by

▶ Rows:

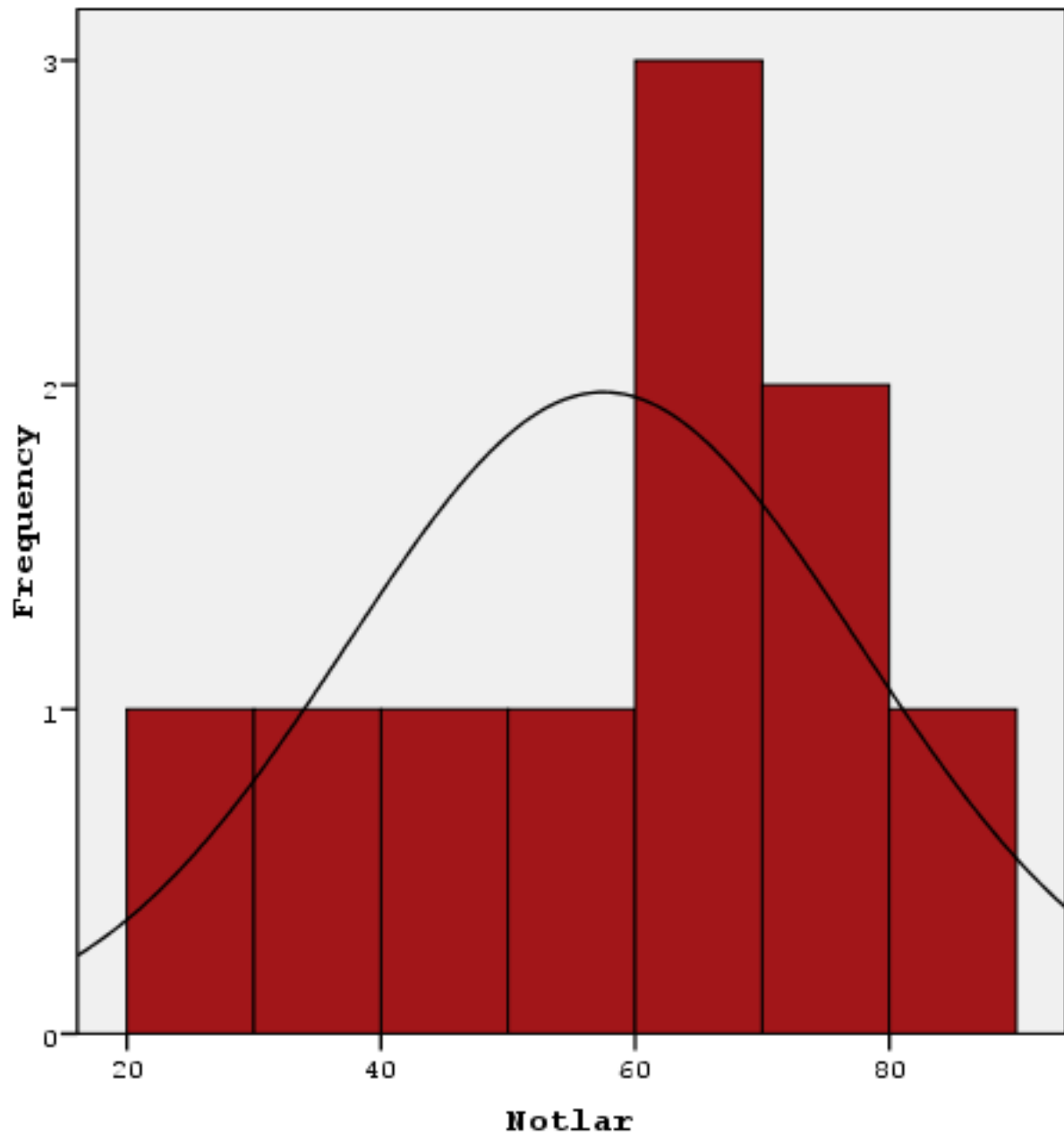
Nest variables (no empty rows)

▶ Columns:

Nest variables (no empty columns)

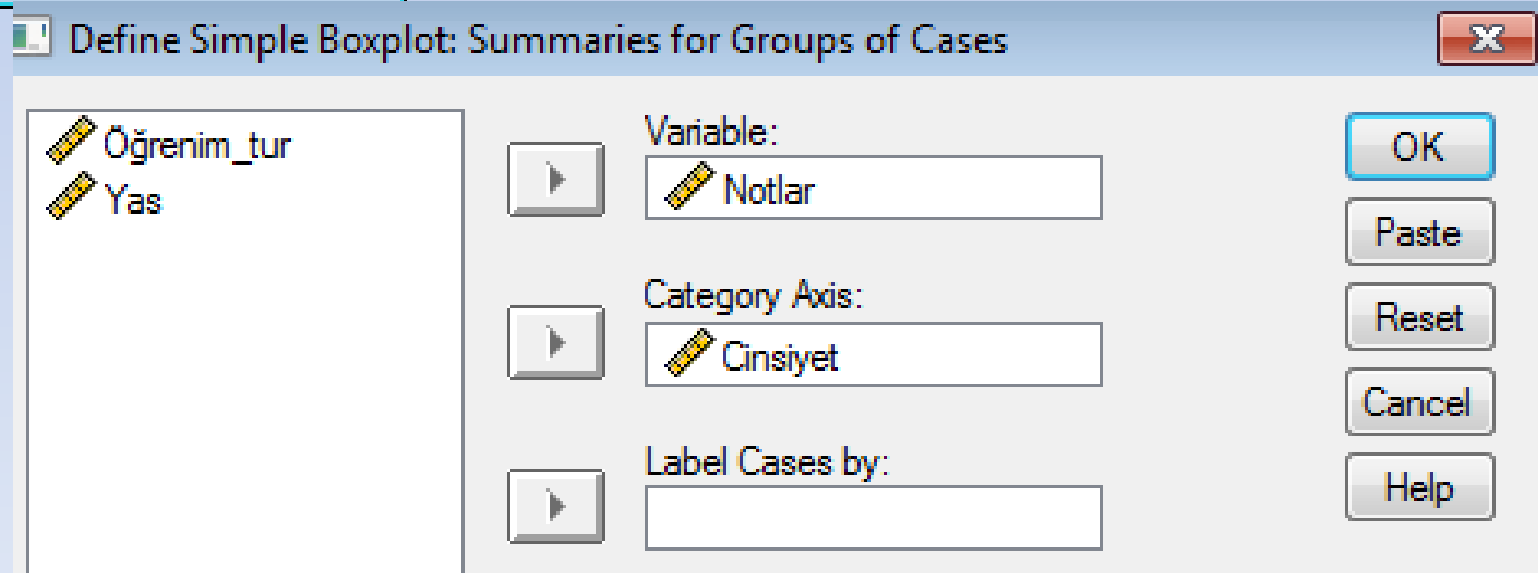
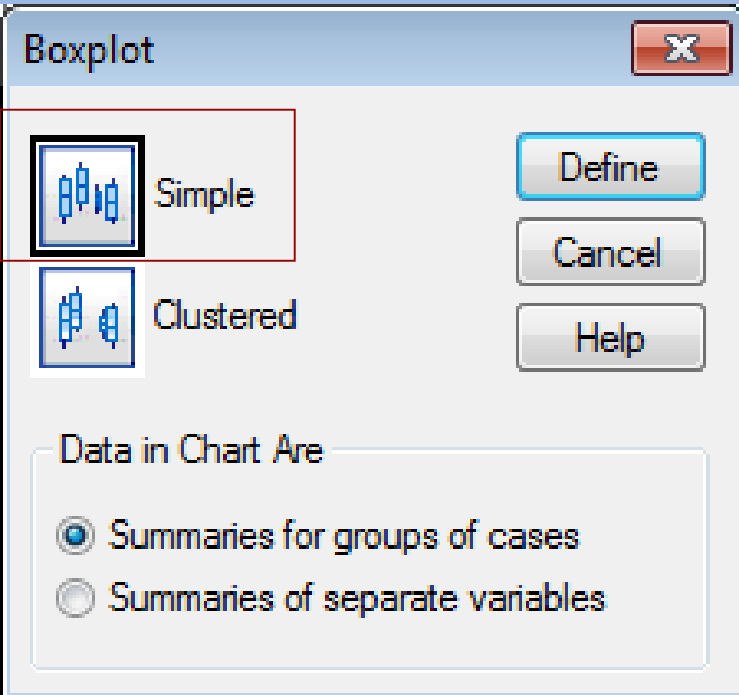
Template

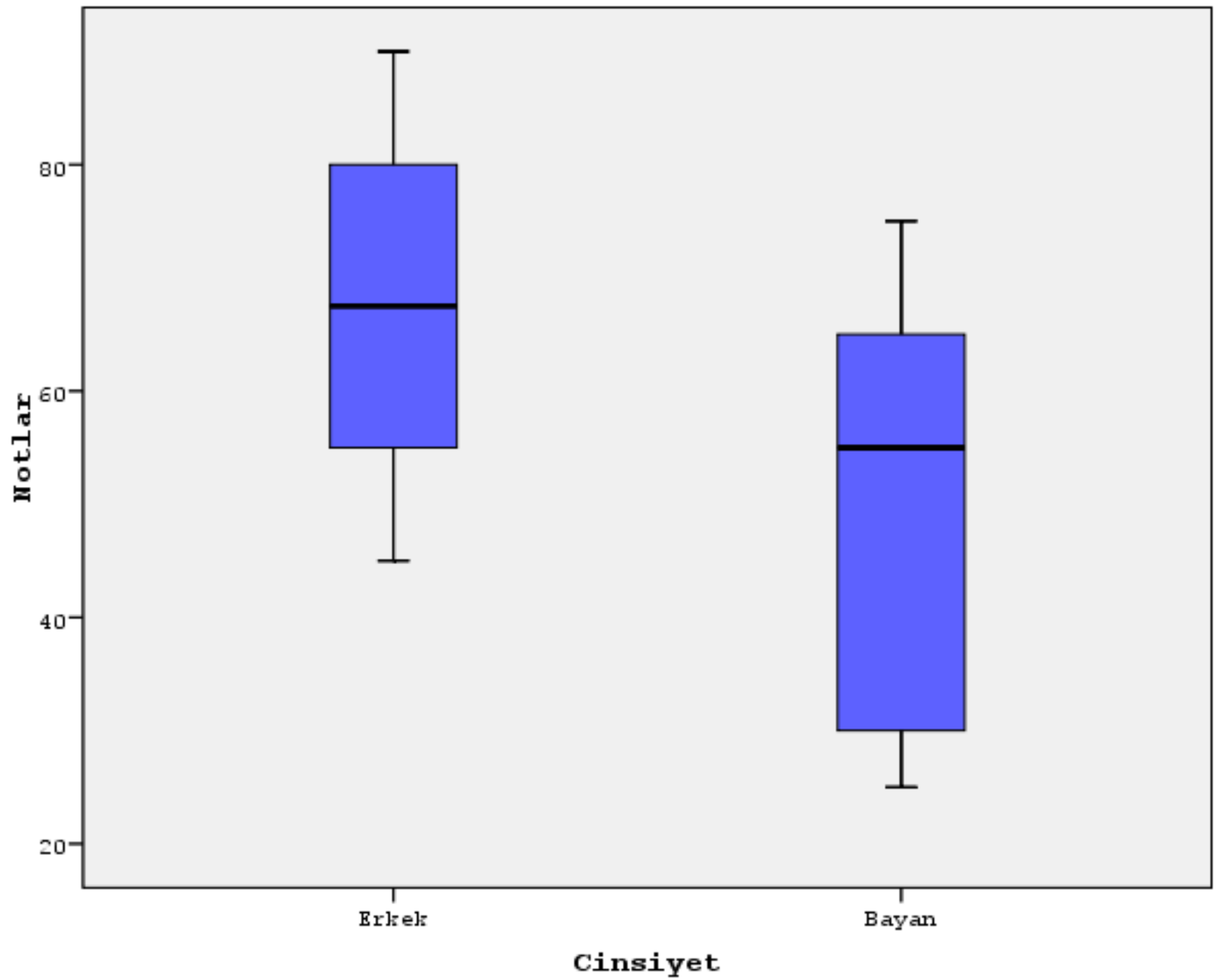
Use chart specifications from:

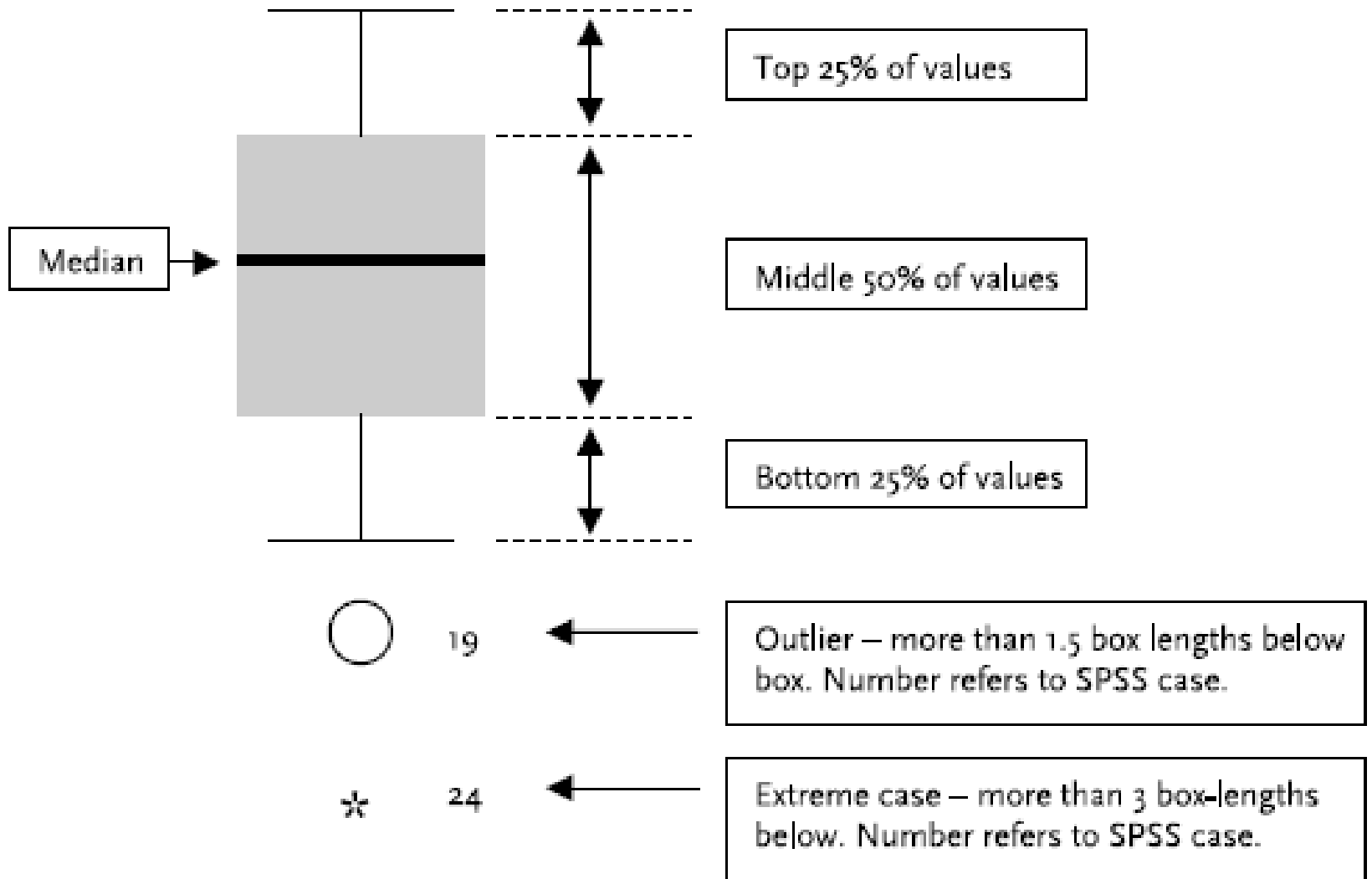


Mean =57,5
Std. Dev. =20,
173
N =10

Boxplot (Kutu) Grafikleri







*** Aşırı gözlem (extreme values)**

o Sapan gözlem (outliers)

Aşırı gözlemler : Veride genel eğilimin oldukça dışında çıkan ya da diğer gözlemlerden oldukça farklı olan gözlemlere aşırı gözlemler denir. Aşırı gözlemler hatalı veri girişi, yanlış kodlama, nadir görülen bir durum, vb. sebeplerden ortaya çıkabilir.

Box plot grafiğinde $Q1-3 \times IQR$ ve $Q3+3 \times IQR$ sınırları dışında kalan gözlemler çok aşırı değer (extreme value) olarak tanımlanır ve * ile gösterilirler. $IQR=Q3-Q1$

Kutunun alt ve üst sınırından 1,5 kutu boyu 3 kutu boyu arasında olan gözlemler ise «o» ile gösterilir ve aşırı değer (outlier) olarak adlandırılır.

«o» değerleri $Q1-1,5 \times IQR$ ve $Q3+1,5 \times IQR$ sınırları dışında kalan değerlerdir.

STEM&LEAF GRAFIĞİ (DAL VE YAPRAK GRAFIĞİ)

YAS Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1,00	2 .	9
4,00	3 .	0234
5,00	3 .	56789
5,00	4 .	01234
5,00	4 .	56789
5,00	5 .	01234
4,00	5 .	6789
4,00	6 .	0123
2,00	6 .	69
3,00	7 .	013
3,00	7 .	578
Stem width:	10,00	
Each leaf:	1 case(s)	

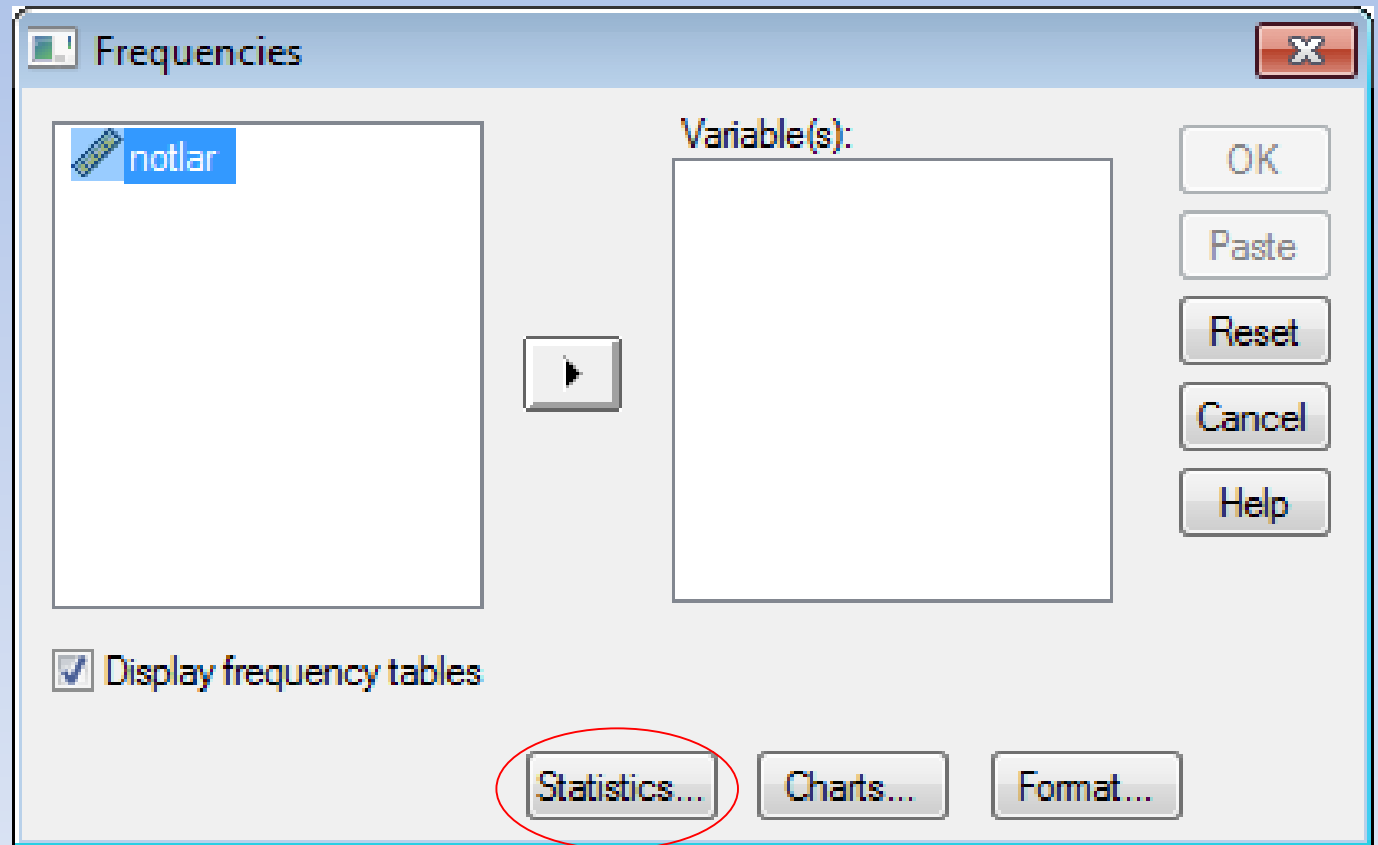
Analyze>Descriptive statistics>Explore yolu ile *Plots* tuşu seçilerek *Stem-and leaf* işaretlenir.

Burada stem (dal) ve leaf (yaprak) anlamındadır.

Yukarıdaki tabloda örneğin “4,00 3.0234” satırı 30 ile 34 değerleri arasında 4 değer olduğunu gösterir. Tablo altındaki **Stem width** : dal genişliğini, **Each leaf** : her yapraktaki deneklerin tek tek gösterildiğini ifade eder.

Frekans, Ortalamalar, Standart Sapma ve Varyansın Bulunuşu

	notlar	
1	30	
2	45	
3	50	
4	50	
5	50	
6	55	
7	55	
8	60	
9	75	
10	80	
11		



Frequencies: Statistics

Percentile Values

Quartiles

Cut points for: equal groups

Percentile(s):

Central Tendency

Mean

Median

Mode

Sum

Values are group midpoints

Dispersion

Std. deviation

Variance

Range

Minimum

Maximum

S.E. mean

Distribution

Skewness

Kurtosis

Quartiles: Kartiller

Mean: Aritmetik ortalama

Median: Medyan

Mode: Mod

Sum: Toplam

Skewness: Çarpıklık

Kurtosis: Basıklık

Std.deviation: Standart sapma

Variance: Varyans

Range: Değişim aralığı (max-min)

S.E.mean (Standard error mean): ortalamanın standart hatası

Statistics

notlar

N	Valid	10
	Missing	0
Mean		55,00
Std. Error of Mean		4,534
Median		52,50
Mode		50
Std. Deviation		14,337
Variance		205,556
Skewness		,318
Std. Error of Skewness		,687
Kurtosis		,587
Std. Error of Kurtosis		1,334
Range		50
Minimum		30
Maximum		80
Sum		550
Percentiles	25	48,75
	50	52,50
	75	63,75

Bir serinin ortalaması verilirken standart hatası ile birlikte verilir. Popülasyondaki değişimi göstermek için standart hata kullanılır. Veri setindeki değişim araştırılıyorsa standart sapma tercih edilir.

Aritmetik Ortalama \pm Standart Hata

$$\bar{X} \pm S_{\bar{X}}$$

Böylece ortalamaya ait değişkenlik yanında ifade edilmiş olur.

Bir serinin ortalaması verilirken standart hatası ile birlikte verilir.

Aritmetik Ortalama \pm Standart Hata

Standart Sapma mı? Standart Hata mı?

Çalışma gruplarına ait veriler, sadece ilgili olduğu grubun özelliğini/özelliklerini gösteriyorsa “*ortalama \pm standart sapma*” tercih edilmelidir.

Verileri birbiri ile karşılaştırarak, gruplar arasında fark olup olmadığı öğrenilmek isteniyorsa; “*ortalama \pm standart hata*” kullanılması daha uygun olacaktır.

Herhangi bir dağılımda gözlem değerlerinin %75'i $AO \pm 2 SS$ aralığında yer alır.

Normal dağılımda ise:

Ortalama ± 1 (standart sapma):

Verilerin % 68,27'inin "ortalama ± 1 (standart sapma)" kadar ortalamanın sağına-soluna yayıldığını gösterir.

Ortalama ± 2 (standart sapma):

Verilerin % 95,45'inin "ortalama ± 2 (standart sapma)" kadar ortalamanın sağına-soluna yayıldığını gösterir.

Ortalama ± 3 (standart sapma):

Verilerin % 99,73'ünün "ortalama ± 3 (standart sapma)" kadar ortalamanın sağına-soluna yayıldığını gösterir.

Aynı durumu standart hata için yapılacak olursa:

Ortalama ± 1 (standart hata):

Verilerin % 68,27 olasılıkla “ ± 1 (standart hata)” kadar sağa-sola yayıldığını gösterir. (Yani, bu çalışma grubu, aynı popülasyondan 100 kez tekrar tekrar oluşturulacak olursa, bunların 68’inin ortalaması, bu sınırlar arasında kalacaktır).

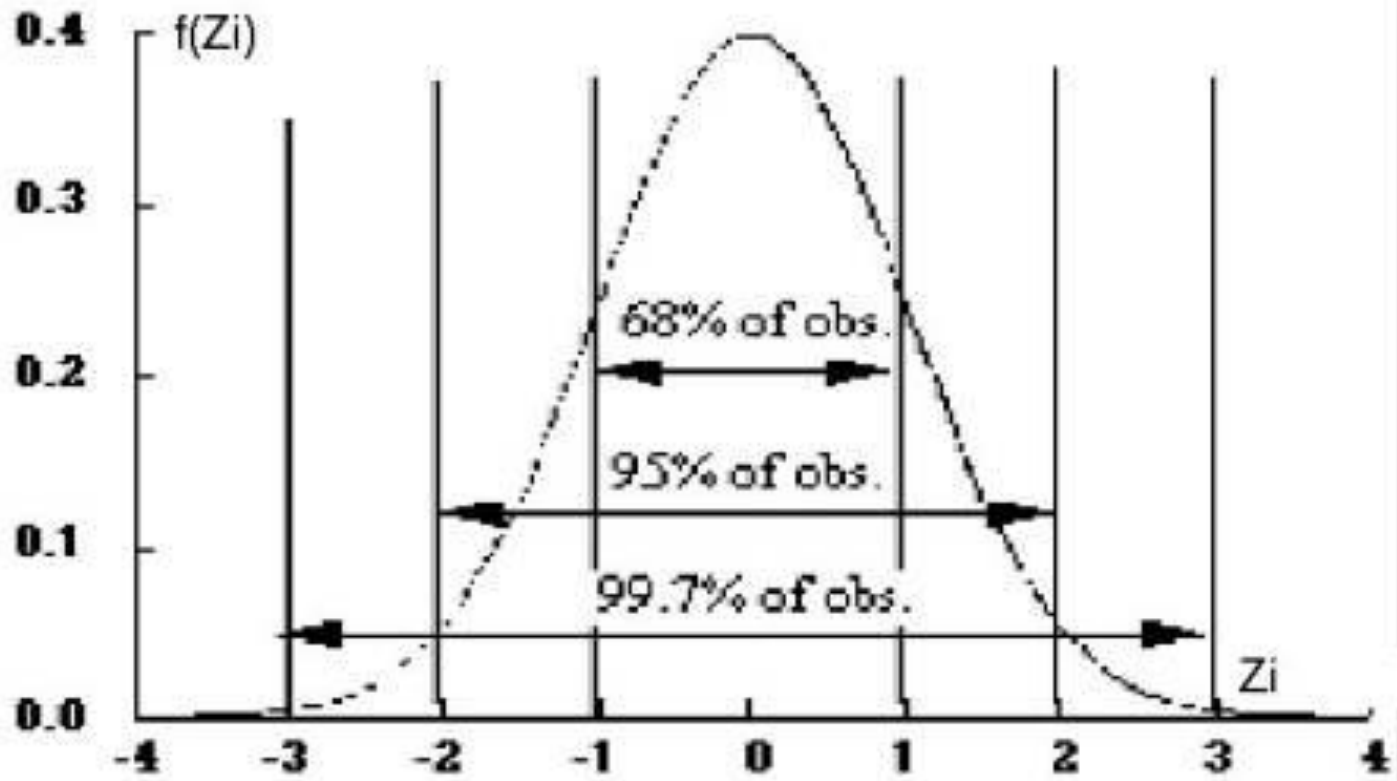
Ortalama ± 2 (standart hata):

Verilerin % 95,45 olasılıkla “ ± 2 (standart hata)” kadar sağa-sola yayıldığını gösterir.

Ortalama ± 3 (standart hata):

Verilerin % 99,73 olasılıkla “ ± 3 (standart hata)” kadar sağa-sola yayıldığını gösterir.

Standart ayrılış ölçüsü ve standart hata Normal dağılış gösteren özellikler için oldukça kullanışlıdır. Normal dağılışta bireylerin %68.27 si ortalamanın bir standart sapma solunda ve bir standart sapma sağındaki noktalar arasında bulunur. Bireylerin %95.45 i ortalamanın iki standart sapma sağında ve solunda bulunur.



FAKTÖRİYEL HESAPLAMA

	veri
1	1
2	2
3	3
4	4
5	5
6	6

Compute Variable

Target Variable: Faktöriyel = Numeric Expression: EXP(LNGAMMA(veri+1))

Type & Label...

veri

Function group:

- All
- Arithmetic
- CDF & Noncentral CDF
- Conversion
- Current Date/Time
- Date Arithmetic
- Date Creation

Functions and Special Variables:

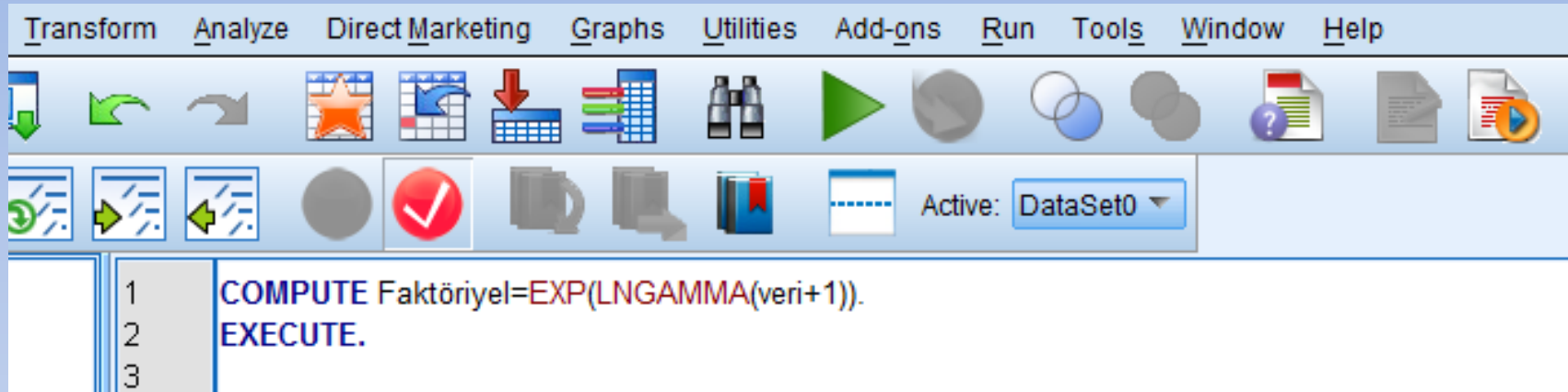
- Abs
- Arsin
- Artan
- Cos
- Exp
- Lg10
- Ln
- Lgamma
- Mod
- Rnd(1)
- Rnd(2)

LNGAMMA(numexpr). Numeric. Returns the logarithm of the complete Gamma function of numexpr, which must be numeric and greater than 0.

If... (optional case selection condition)

OK Paste Reset Cancel Help

Syntax komut ile faktöriyel hesaplama:



The screenshot shows the IBM SPSS Syntax editor interface. The menu bar includes Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Run, Tools, Window, and Help. The toolbar contains various icons for data manipulation and analysis. The active dataset is 'DataSet0'. The Syntax editor contains the following commands:

```
1 COMPUTE Faktöriyel=EXP(LNGAMMA(veri+1)).  
2 EXECUTE.  
3
```

	veri	Faktöriyel
1	1	1
2	2	2
3	3	6
4	4	24
5	5	120
6	6	720

Excel'de faktöriyel hesaplama:

L	M
1	=ÇARPINIM(L1)
2	
3	ÇARPINIM(sayı)
4	
5	

L	M
1	1
2	2
3	6
4	24
5	120

ANALİZ ÖNCESİ VERİ TARAMA

- 1. Verilerin doğruluğu (accuracy)**
- 2. Kayıp değerler (missing value)**
- 3. Aykırı değerler (Outliers)**
- 4. Normallik vasayımı (normality test)**
- 5. Doğrusallık (linearity)**
- 6. Eşvaryanslılık (Homojenlik-homogeneity)**

ANALİZ ÖNCESİ VERİ TARAMA

1. Verilerin doğruluğu (accuracy) incelenmesi gerekir. Analiz sonuçlarının kalitesi, veri kalitesine bağlıdır. Hatalı (uygun olmayan) verilerle yapılan analizlerden elde edilecek sonuçlar ve bu sonuçlara dayalı oluşturulan yorumlar geçerli olmayacaktır.

Nicel veriler için frekanslara bakılarak alt ve üst değerler belirlenebilir. Ortalama ve s.sapma değerlerine bakılır.

Nitel veriler de ise kodlamalara dikkat edilir, yanlış kodlama yapılıp yapılmadığı tespit edilir.

2. Kayıp veri (missing value) verilerin deęerlendirilmesi gerekir. Kayıp veri ölçme aracının başarısızlığından, deneklerin tüm sorulara cevap vermemesinden ya da veri girişı sırasında yapılan hatalardan kaynaklanabilir.

Eđer kayıp veri az ise o silme yapılabilir. Ancak kayıp veri çok ise o deneklerin veya deęişkenin silinmesi ciddi veri kaybına sebep olabilir. Örneklem büyüklüğü düşer ve eđer araştırma için gruplar arası karşılaştırmalar isteniyorsa, bazı gruplar analiz için uygun olmayan örneklem büyüklüğüne inerek karşılaştırmaları tehlikeye sokabilir.

Kayıp veriler için kestirim yapılabilir. Ancak kestirim ya da yaklaşık deęer atama işlemi sadece nicel deęişkenler için yapılabilmektedir.

2. Kayıp veri (missing value)

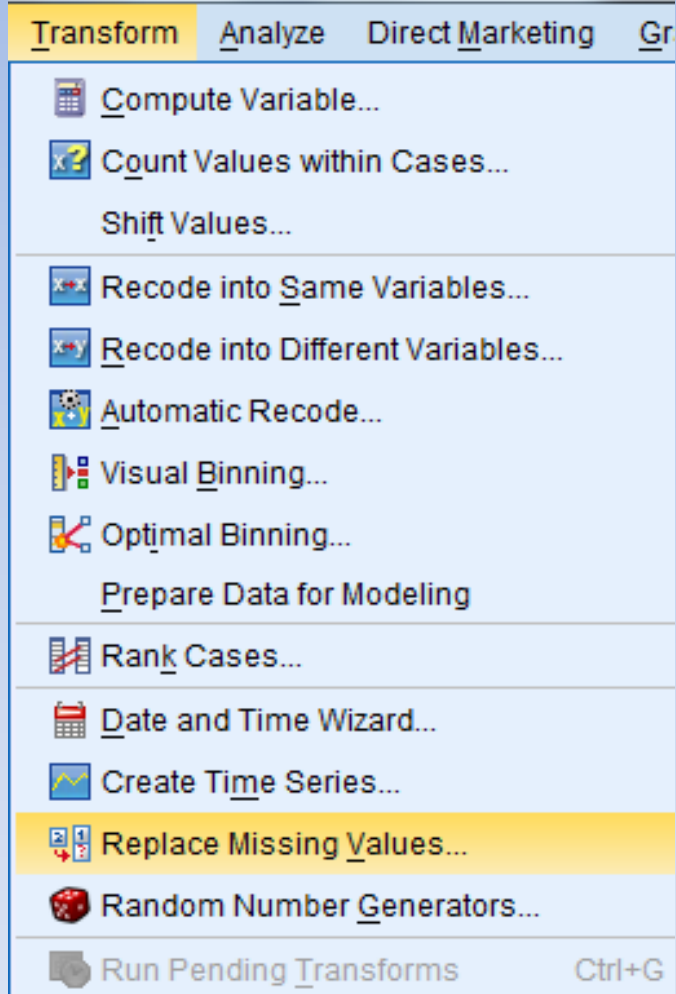
Kayıp değer kestirimi için bir yaklaşım eldeki verilerden ortalama bulup, kayıp değerlere bu değeri atamaktır. Bu durumda genel ortalama değişmezken, varyans bir miktar düşecektir.

Kayıp veri kestirimi için diğer bir yaklaşım regresyon yaklaşımıdır. Kayıp değerler içeren değişken bağımlı değişken alınır. Ortalama yöntemine göre göre daha objektif bir kestirim yapılır.

Araştırmacı hem kayıp gözlemlili durum için hem de kestirim yapıldıktan sonraki durum için analizleri yapmalı ve sonuçları karşılaştırmalıdır (Mertler ve Vannatta, 2005).

Replace Missing Values (Kayıp deęer ikamesi)

Kayıp deęerlerin tahmin edilmesinde kullanılır.



Series Mean: Serinin Ortalamasını kayıp değer olarak ikame etmek

Belirli bir değişkene göre tüm verilerin aritmetik ortalamasıdır.

Mean of Nearby Points : Komşu Noktaların Ortalaması

Eksik gözlemin komşusu olan verilerin ortalamasıdır. Çevreleyen değerlerin sayısı, yakın noktaların uzaklığı (span of nearby points) seçeneği kullanılarak belirlenir. Eğer 2 alınırsa eksik verinin öncesi ve sonrası alınarak ortalama bulunur.

Median of Nearby Points : Komşu Noktaların Medyanı

Eksik verinin komşusu olan verilerin medyanıdır (ortancasıdır). Çevreleyen değerlerin kaç adedinin kullanılacağı araştırmacı tarafından belirlenir.

Linear Interpolation : Doğrusal Değer Kestirimi

Eksik verilerden sonraki gözlem değerlerinden önceki gözlem değeri çıkarılır ve bulunan fark değer, eksik gözlem sayısının bir fazlasına bölünür. Bölümle elde edilen değer, eksik değerlerden önceki gözlem değerine eklenerek ilk eksik gözlem bulunur. Aynı şekilde bölümle bulunan değer, bulunan ilk eksik gözlem değerine eklenerek ikinci eksik gözlem değeri bulunur. Böylece devam edilerek son eksik gözlem değeri bulunur.

Linear Trend of Point : Noktanın Doğrusal Eğimi

Regresyon modelinde yer alan bağımsız değişkenlerden biri kayıp veriler içermektedir. Bu değişken diğer bağımsız değişkenlerle regresyon analizine tabi tutulur. Regresyon analizinden elde edilen tahmin denklemi kullanılarak kayıp veriler yerine atama yapılır.

	Puan
1	50
2	80
3	100
4	70
5	50
6	.
7	30
8	40
9	50
10	60

Replace Missing Values

New Variable(s):
Puan_1 = SMEAN(Puan)

Name and Method
Name: Puan_1
Method: Series mean

Span of nearby points:
 Number: 2 All

OK Paste Reset Cancel Help

	Puan	Puan_1
1	50	50,0
2	80	80,0
3	100	100,0
4	70	70,0
5	50	50,0
6	.	58,9
7	30	30,0
8	40	40,0
9	50	50,0
10	60	60,0

New Variable(s):
Puan_1 = MEAN(Puan 1)

Name and Method
Name: Puan_1
Method: Mean of nearby points

Span of nearby points:
 Number: 1 All

OK Paste Reset Cancel Help

Puan	Puan_1
50	50
80	80
100	100
70	70
50	50
.	40
30	30
40	40
50	50
60	60

Replace Missing Values

New Variable(s):
Puan_1 = MEDIAN(Puan ALL)

Name and Method
Name: Puan_1 Change
Method: Median of nearby points
Span of nearby points:
 Number: 2 All

OK Paste Reset Cancel Help

Puan	Puan_1
50	50
80	80
100	100
70	70
50	50
.	50
30	30
40	40
50	50
60	60

Linear Interpolation

	X1	X2	X3	X4
1	4	2	4	5
2	6	.	6	6
3	7	5	.	8
4	.	7	9	11
5	10	9	12	.
6	12	10	14	15

Replace Missing Values

New Variable(s):

```
X1_1 = LINT(X1)
X2_1 = LINT(X2)
X3_1 = LINT(X3)
X4_1 = LINT(X4)
```

Name and Method

Name: [Change](#)

Method:

Span of nearby points:

Number: All

[OK](#) [Paste](#) [Reset](#) [Cancel](#) [Help](#)

	X1	X2	X3	X4	X1_1	X2_1	X3_1	X4_1
1	4	2	4	5	4,0	2	4,0	5,0
2	6	.	6	6	6,0	4	6,0	6,0
3	7	5	.	8	7,0	5	7,5	8,0
4	.	7	9	11	8,5	7	9,0	11,0
5	10	9	12	.	10,0	9	12,0	13,0
6	12	10	14	15	12,0	10	14,0	15,0

Linear Trend of Point :

	X1	X2	X3	X4
1	4	2	4	5
2	6	.	6	6
3	7	5	.	8
4	.	7	9	11
5	10	9	12	.
6	12	10	14	15
7				
8				
9				
10				

Replace Missing Values
✕

New Variable(s):

X1_1 = TREND(X1)
 X2_1 = TREND(X2)
 X3_1 = TREND(X3)
 X4_1 = TREND(X4)

Name and Method

Name: Change

Method: Linear trend at point

Span of nearby points:

Number: All

OK Paste Reset Cancel Help

	X1	X2	X3	X4	X1_1	X2_1	X3_1	X4_1
1	4	2	4	5	4,0	2	4,0	5,0
2	6	.	6	6	6,0	4	6,0	6,0
3	7	5	.	8	7,0	5	7,8	8,0
4	.	7	9	11	8,7	7	9,0	11,0
5	10	9	12	.	10,0	9	12,0	12,8
6	12	10	14	15	12,0	10	14,0	15,0

3. Uç Değerler (Outliers)

Uç (aykırı) değerlerin etkilerinin değerlendirilmesi gerekir. Alışlageldik değerlerin dışındaki değerlere ya da aşırı değerlere sahip olan denekler uç değerler olarak adlandırılır. Uç değerlerin ortaya çıkma nedenleri şunlardır:

- i) Araştırmacılarca veri girişinde yapılan hatalar,**
- ii) Cevaplayıcının kasıtlı cevaplar vermesi,**
- iii) Örnekleme hataları (deneğin örneklemin alındığı evrenin bir üyesi olmaması),**
- iv) Uygulama hataları,**
- v) Ölçme hataları,**
- vi) Dağılıma ait yanlış varsayımda bulunulması,**
- vii) Verilerin alındığı anakütlenin yapısı,**
- viii) Soruların yanlış anlaşılmasından kaynaklanan hatalar.**

Uç değerler istatistik testlerin sonuçlarını bozabilir. Pek çok istatistiksel işlem ortalamadan olan sapmaların karesine dayanır. Eğer bir gözlem ortalamadan çok uzakta ise sapma değeri büyüyecektir.

Uç Değerlerin Belirlenmesi:

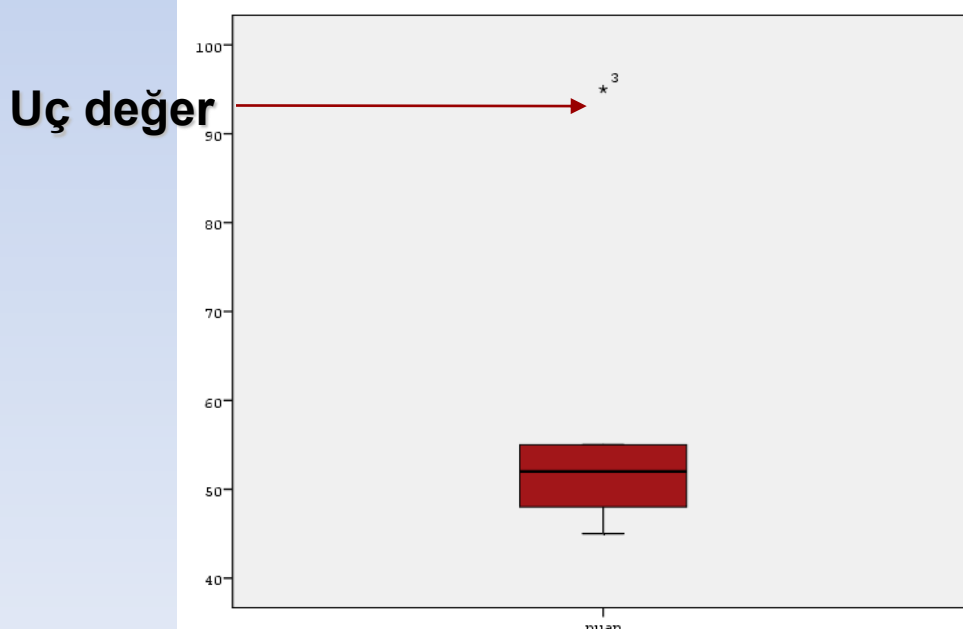
1. Tek Değişkenli Uç değerlerin Belirlenmesi

i) Z Puanı:

Z Puanı ilgili değer in ortalamadan kaç standart sapma uzaklıkta olduğunu göstermektedir. Z puanı (-3,+3) aralığının dışında ise bu değeri uç değer olarak tanımlayan bazı bilim adamları mevcuttur. Ancak buna karşı çıkan bilim adamları mevcuttur (Iglewicz ve Hoaglin, 1993).

ii) Grafik Yöntemi:

Uç değerlerin bulunmasında Box plot (Kutu grafiği) kullanılabilir.



Uç Değerlerin Belirlenmesi:

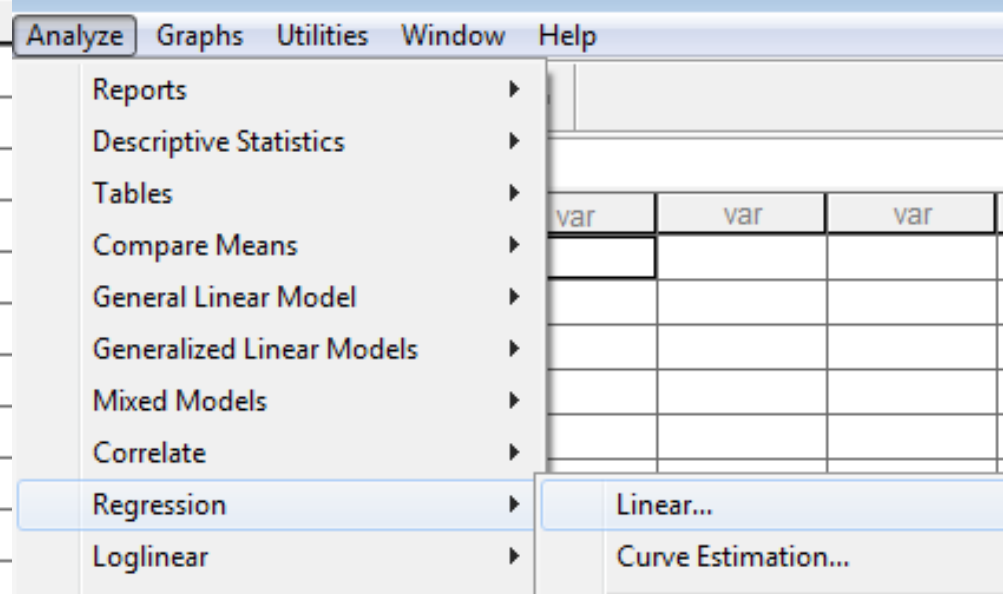
2. Çoklu Değişken İçin Uç Değerlerin Belirlenmesi

i) Mahalanobis Uzaklığı:

Mahalanobis uzaklık değeri hesaplanır ve Ki-kare tablo değeriyle karşılaştırılır. Mahalanobis değeri ki-kare tablo değerinden büyük ise bu değer uç değer olarak kabul edilir. $P=0,01$ alınır.

Çok deęişkenli analizlerde uç deęerler Mahalanobis uzaklığı kullanılarak belirlenir.

	yas	boy	agirlik	kanbasýn
1	51	166	67,00	115
2	64	165	61,00	122
3	46	174	83,00	130
4	39	168	78,90	126
5	58	162	67,00	110
6	54	178	90,00	141
7	31	171	77,70	124
8	67	173	89,30	150
9	48	165	70,00	110
10	78	152	58,00	119
11	39	177	82,50	130
12	51	166	63,00	120
13	73	178	93,10	149
14	53	174	89,00	125
15	56	169	72,00	125
16	47	159	64,00	114
17				



Linear Regression

yas
 boy
 agirik

Dependent:

Block 1 of 1

Independent(s):
 yas
 boy
 agirik

Method:

Selection Variable:

Case Labels:

WLS Weight:

Linear Regression: Save

Predicted Values
 Unstandardized
 Standardized
 Adjusted
 S.E. of mean predictions

Distances
 Mahalanobis
 Cook's
 Leverage values

Prediction Intervals
 Mean Individual
 Confidence Interval: %

Coefficient statistics
 Create coefficient statistics
 Create a new dataset
 Dataset name:

Write a new data file

Residuals
 Unstandardized
 Standardized
 Studentized
 Deleted
 Studentized deleted

Influence Statistics
 DfBeta(s)
 Standardized DfBeta(s)
 DfFit
 Standardized DfFit
 Covariance ratio

Export model information to XML file

Include the covariance matrix

MAH_1	Analyze	Graphs	Utilities	Window	Help
,94168	Reports				
5,27463	Descriptive Statistics				Frequencies...
,75239	Tables				Descriptives...
3,09470	Compare Means				Explore...
,87433	General Linear Model				Crosstabs...
1,86781					
3,31692					
3,27989					
,59264					
7,98991					
2,65863					
2,99060					
5,55067					
1,89301					
,82909					
3,09311					

Explore

yas
boy
agirlik
kanbasin
Mahalanobis Distance

Dependent List:
MAH_1

Factor List:

Label Cases by:

Display
 Both Statistics Plots

OK
Paste
Reset
Cancel
Help
Statistics...
Plots...
Options...

Explore: Statistics ✕

Descriptives
 Confidence Interval for Mean: %

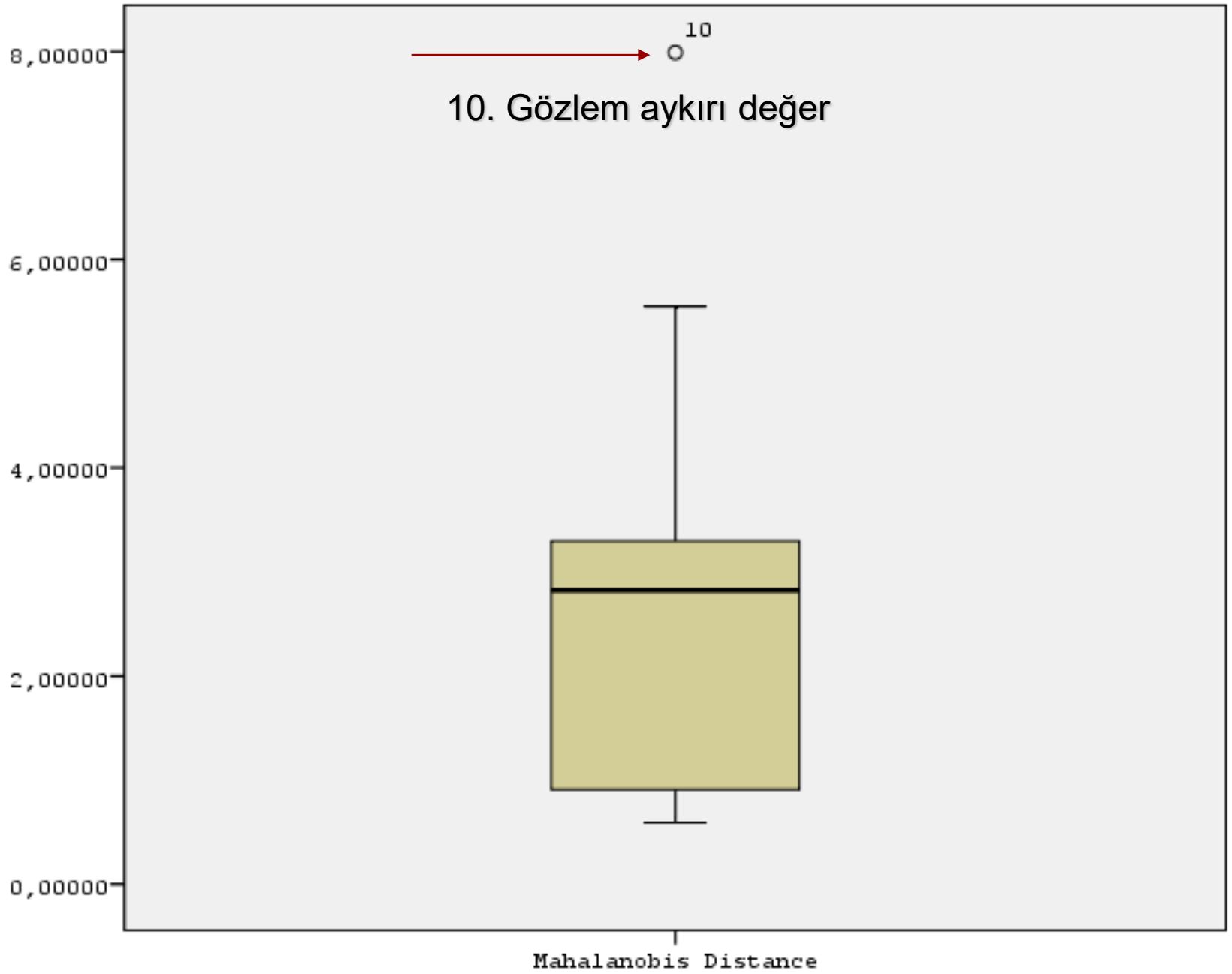
M-estimators

Outliers

Percentiles

Extreme Values

			Case Number	Value
Mahalanobis Distance	Highest	1	10	7,98991
		2	13	5,55067
		3	2	5,27463
		4	7	3,31692
		5	8	3,27989
	Lowest	1	9	,59264
		2	3	,75239
		3	15	,82909
		4	5	,87433
		5	1	,94168



Uç Değerlerin Belirlenmesi:

2. Çoklu Değişken İçin Uç Değerlerin Belirlenmesi

ii) Cook Uzaklığı:

Belirli bir modele göre bir gözlemin etkisini göstermek üzere geliştirilmiştir (Cook, 1977). Uç değer olarak bulunan gözlemin silinmesi yerine, buna neden olan durumlar belirlenir.

Cook değeri 1'den büyük olan gözlem uç değer olarak kabul edilir. Ancak genel olarak Cook değeri $4/n$ 'den büyükse o gözlem uç değer olarak kabul edilmelidir.

y	x1	x2	COO_1
40	4	120	4,52259
52	5	145	,60026
55	8	160	,11351
60	9	170	,11223
68	12	175	,01278
74	15	180	,27068

$$4/n=4/6=0,33$$

1. ve 2. gözlem değerleri 0,33'den büyük olduğundan uç değer olarak kabul edilirler.

The image shows two screenshots of the SPSS Linear Regression dialog boxes. The left screenshot is the main 'Linear Regression' dialog, and the right is the 'Linear Regression: Save' dialog.

Linear Regression Dialog:

- Dependent: y
- Independent(s): x1, x2
- Method: Enter
- Buttons: Statistics..., Plots..., Save..., Options..., Style..., Bootstrap...

Linear Regression: Save Dialog:

- Predicted Values:
 - Unstandardized
 - Standardized
 - Adjusted
 - S.E. of mean predictions
- Residuals:
 - Unstandardized
 - Standardized
 - Studentized
 - Deleted
 - Studentized deleted
- Distances:
 - Mahalanobis
 - Cook's
 - Leverage values
- Influence Statistics:
 - DfBeta(s)
 - Standardized DfBeta(s)
 - DfFit

Uç Değerlerin Belirlenmesi:

2. Çoklu Değişken İçin Uç Değerlerin Belirlenmesi

iii) Leverage Uzaklığı:

Leverage değerli 0-1 aralığında değerler alır. 1'e yaklaştıkça gözlemin uç değer olma olasılığı artar. Leverage değeri $2p/n$ değerinden büyükse gözlem uç değer olarak kabul edilir.

p: Açıklayıcı (bağımsız) değişken sayısı

n: Gözlem sayısı

	y	x1	x2	LEV_1
1	40	4	120	,73368
2	52	5	145	,23031
3	55	8	160	,07792
4	60	9	170	,25126
5	68	12	175	,11870
6	74	15	180	,58814

$2p/n=2*2/6=0,67$
1. Gözlem uç değerdir.

Linear Regression

Dependent:

Block 1 of 1

Independent(s):

Method:

x1

x2

Cook's Distance [C...

Previous

Next

Statistics...

Plots...

Save...

Options...

Style...

Bootstrap...

Linear Regression: Save

Predicted Values

Unstandardized

Standardized

Adjusted

S.E. of mean predictions

Residuals

Unstandardized

Standardized

Studentized

Deleted

Studentized deleted

Distances

Mahalanobis

Cook's

Leverage values

Influence Statistics

DfBeta(s)

Standardized DfBeta(s)

DfFit