

1996-1998 YILLARI ARASINDA GÜNEY CAROLINA'DA OKUL İÇİ ŞİDDET İSTATİSKLERİ ANALİZİ

(Bir Önceki Projeden Devam Edilecektir)

Hazırlayan : Süleyman Öğrekçi

No :29054041

1996 ve 1998 yılları arasında Güney Carolina'da resmi kayıtlara geçen okul içi şiddet olaylarını ele alan bir çalışma, *South Carolina Department of Public Safety Office of Justice Programs* kurumu tarafından hazırlamış ve kamuoyuyla paylaşılmıştır. Yapılan çalışmanın sonuç belgesi (final report) http://www.jrsa.org/ibrrc/background-status/South_Carolina/sc_report.pdf adresinden ulaşılabilir durumdadır. hazırladığım bu projede bu çalışmanın 381 denegi üzerinden veriler toplanmış olup, bazı incelemeler yapılmış sonuç istatistikler paylaşılmıştır. Bu çalışmada toplanan istatistiklerin bir kısmı yer almaktadır. Örneğin tüm denekler yerine 381 denek, tüm değişkenler yerine belli sayıda değişken (örneğin suçun hangi ayda işlendiğine dair değişken gibi bazı değişkenler yoktur) kullanılmıştır.

Proje hakkında kısa bir bilgi vermem gerekirse, projede 381 denek 8 değişken üzerinden değerlendirilmiştir. Aşağıda değişkenlerin detaylarını kabaca görebileceğimiz görüntü vardır.

1.sav [DataSet1] - SPSS Statistics Data Editor											
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	
1	suc	String	3	0	Suç	{09A, Cinay...}	None	4	Left	Nominal	
2	yas	Numeric	9	0	Yaş	None	None	8	Right	Scale	
3	etnik	String	1	0	Etnik Kökeni	{H, İspanyol...}	None	6	Left	Nominal	
4	irk	String	1	0	Irki	{A, Asyalı...}	None	5	Left	Nominal	
5	cins	String	1	0	Cinsiyeti	{F, Bayan...}	None	4	Left	Nominal	
6	silah	Numeric	9	0	Kullandığı Silah	{1, Silahsız...}	None	9	Right	Scale	
7	yer	Numeric	9	0	Yakalandığı Yer	{2, Bar/Gec...}	None	8	Right	Scale	
8	yil	Numeric	8	0	Yakalanma Yılı	None	None	8	Right	Scale	
9											

Deneklere ilişkin bilgilerin de bir örnek görüntüsü aşağıda verilmiştir.

1.sav [DataSet1] - SPSS Statistics Data Editor

	suc	yas	etnik	irk	cins	silah	yer	yil	
1	13B	17	N	W	M	95	22	1998	
2	13A	16	N	B	F	95	22	1998	
3	13C	15	N	B	F	95	22	1998	
4	13C	13	N	B	F	95	22	1998	
5	13C	34	N	W	M	95	22	1998	
6	120	12	N	W	M	95	22	1998	
7	13A	14	N	W	M	95	22	1998	
8	13B	15	N	B	M	95	22	1998	
9	13B	13	U	B	M	95	22	1998	

Aşağıdaki verilen iki tablo ile de çalışma ile elde edilen istatistiklerin bir özeti bulunmaktadır. Tablolarda değişkenlerin ve durumların sayısı, ortalama değerleri, standart sapmaları, maksimum ve minimum değerleri, kurtosis ve skewness değerleri gibi bası sayısal bilgiler verilmiştir.

Descriptive Statistics

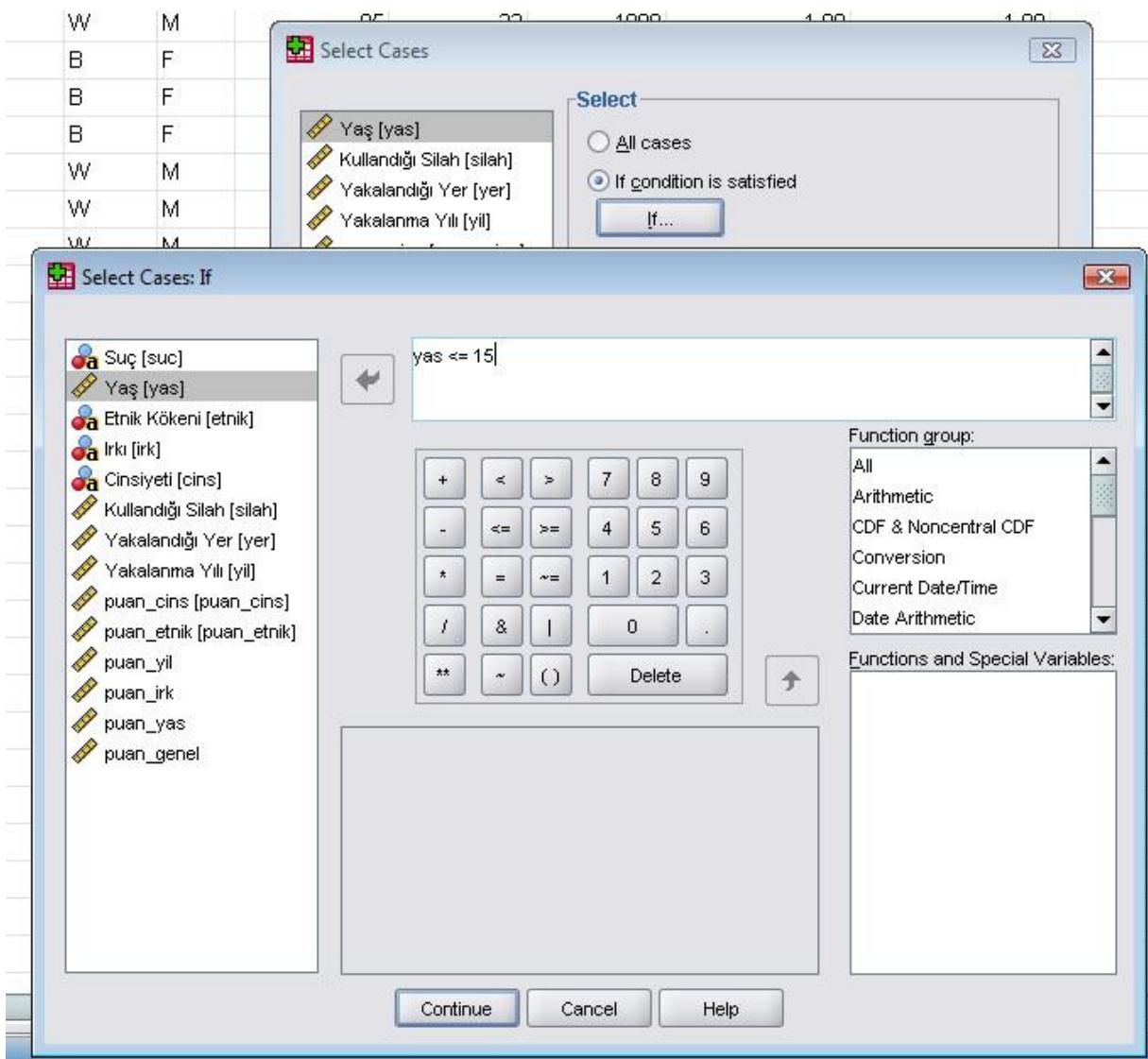
	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Yaş	381	66	8	74	16,08	,351	6,849	46,907	5,026	,125	30,302	,249
Kullandığı Silah	381	0	95	95	95,00	,000	,000	,000				
Yakalandığı Yer	381	9	13	22	21,97	,026	,504	,254	-16,174	,125	274,797	,249
Yakalanma Yılı	381	2	1996	1998	1997,11	,041	,802	,642	-,207	,125	-1,415	,249
Valid N (listwise)	381											

	Suç	Yas	Etnik Kökeni	Irki	Cinsiyeti	Kullandığı Silah	Yakalandığı Yer	Yakalanma Yılı
N	Valid	381	381	381	381	381	381	381
	Missing	0	0	0	0	0	0	0
Mean		16,08				95,00	21,97	1997,11
Std. Error of Mean		,351				,000	,026	,041
Median		15,00				,95,00	,22,00	,1997,00
Mode		15				,95	,22	,1998
Std. Deviation		6,849				,000	,504	,802
Variance		46,907				,000	,254	,642
Skewness		5,026					-16,174	-,207
Std. Error of Skewness		,125				,125	,125	,125
Kurtosis		30,302					274,797	-1,415
Std. Error of Kurtosis		,249				,249	,249	,249
Range		66				0	9	2
Minimum		8				95	13	1996
Maximum		74				95	22	1998
Sum		6125				36195	8369	760900

Yukarıdaki tablolara bakarak istatistik açıdan bazı yorumlar yapılabilir. Örneğin kurtosis değerleri dikkate alınarak yaş değişkeninin dağılım grafiğinin dik, yakalanma yılının dağılımının grafiğinin ise basık yapıda olduğu söylenebilir. Maksimum ve minimum değerleri dikkate alınarak da deneklerin yaşılarının 8 ile 74 arasında değiştiği söylenebilir.

Şimdi analizimize bir önceki projeden kaldığımız yerden devam edelim. Bazı SPSS özelliklerini nasıl kullanacağımıza bakalım.

Örneğin bazı değerlendirmeleri tüm denekler üzerinde değil de belirli özellikleri sağlayan denekler üzerinde yapmamız gerekiyor. Örneğin bir yaş grubu üzerinde değerlendirme yapmak istiyoruz. Bu durumda **data/select cases** komutu kullanılabilir. Aşağıda bu komutun nasıl kullanılacağı örneklenmiştir. 15 yaşından küçük deneklerin değerlendirmesi yapılmıştır.



Bu işlem yapıldıktan sonra SPSS veri ekranı aşağıdaki gibi bir hal alır. Üzeri çizilmiş denekler bu şartta alınmadıkları için dikkate alınmamaktadır.

	suc	yas	etnik	irk	cins	silah	yer
1	13B	17	N	W	M	95	22
2	13A	16	N	B	F	95	22
3	13C	15	N	B	F	95	22
4	13C	13	N	B	F	95	22
5	13C	34	N	W	M	95	22
6	120	12	N	W	M	95	22
7	13A	14	N	W	M	95	22
8	13B	15	N	B	M	95	22
9	13B	13	U	B	M	95	22
10	13B	17	N	B	M	95	22
11	13B	18	N	W	M	95	22
12	13B	15	N	B	F	95	22

Şimdi belirlediğimiz bu özellikler çerçevesinde bazı değişkenlerin frekans tablosuna bakalım.

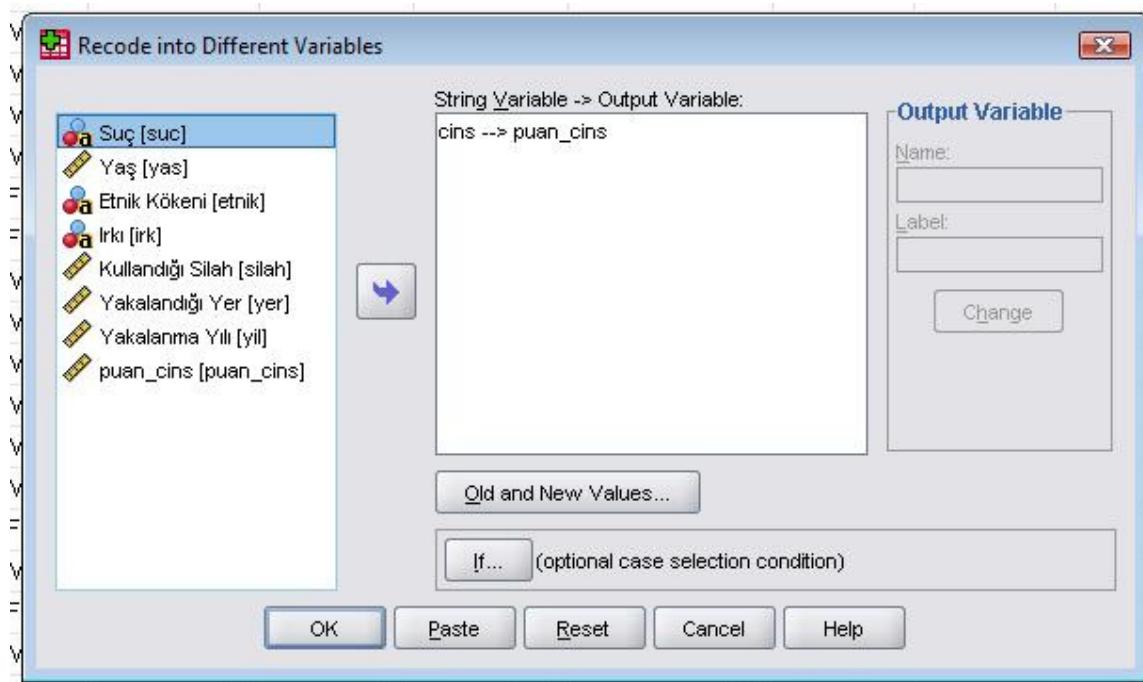
Statistics

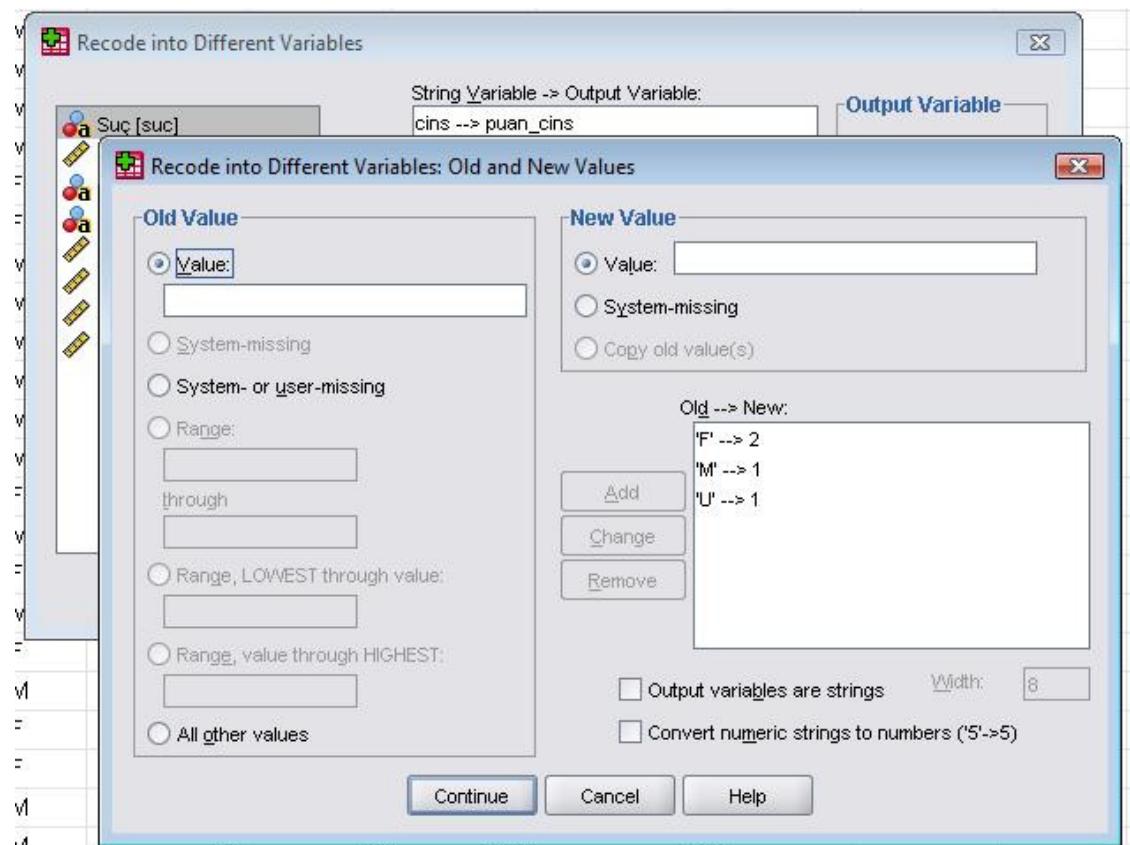
	Suç	Yaş	Etnik Kökeni	Irki	Cinsiyeti
N	Valid	229	229	229	229
	Missing	0	0	0	0
Mean		13,56			
Std. Error of Mean		,102			
Median		14,00			
Std. Deviation		1,540			
Variance		2,370			
Skewness		-1,254			
Std. Error of Skewness		,161			
Kurtosis		1,340			
Std. Error of Kurtosis		,320			
Range		7			
Minimum		8			
Maximum		15			

Şekilde gösterildiği gibi bu istatistikte yaşları 15 ve daha küçük olan denekler dikkate alınmıştır. Bu tabloya bakılarak daha önce yaptığım yorumların benzerini bu yaş grubu için de yapabiliriz.

Şimdi bazı komutların nasıl kullanıldığını göstermek amacıyla değişkenlerimiz üzerinde bazı işlemler yapacağız. SPSS de mevcut olan değişkenler kullanılarak nasıl yeni değişkenler yaratabileceğimizi göstereceğiz. Ve bu yeni değişkenlerle ilerde bazı analizler yapacağız.

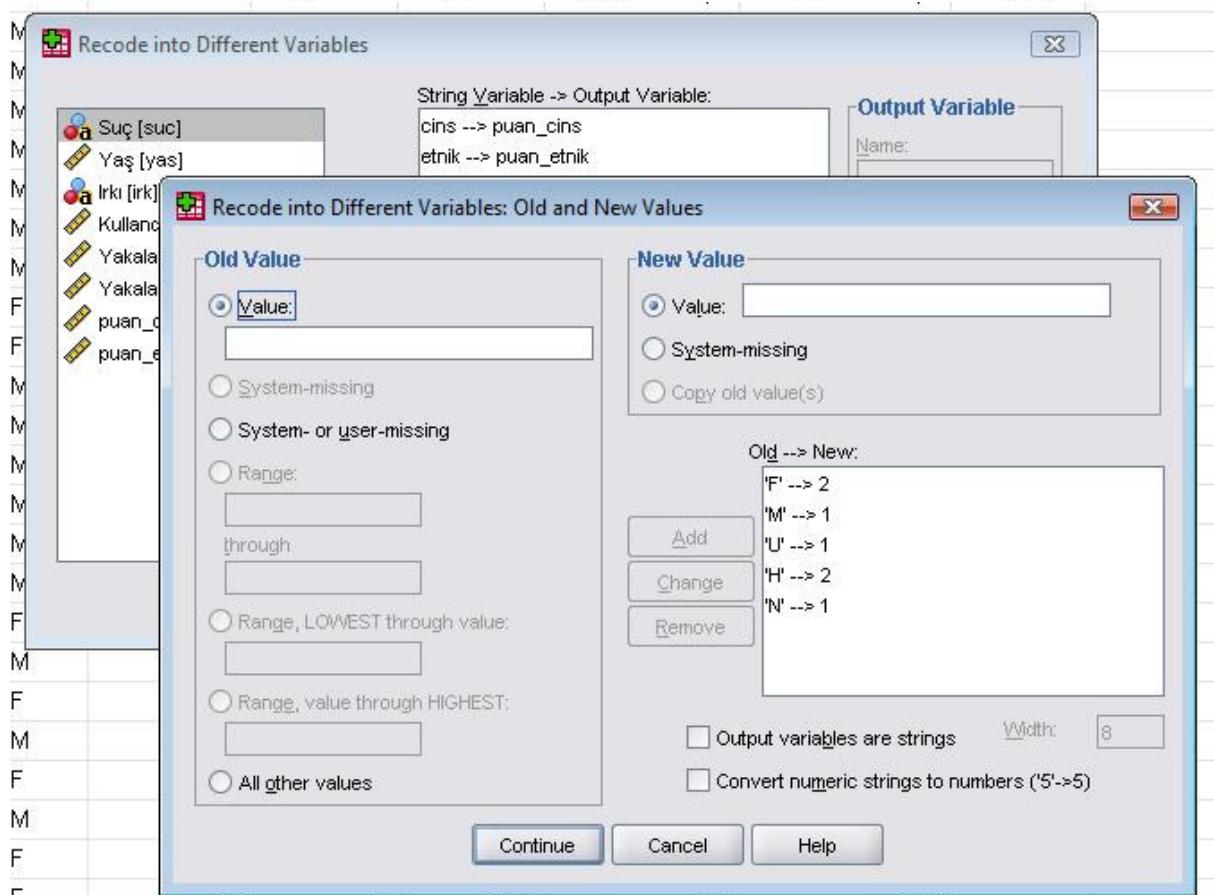
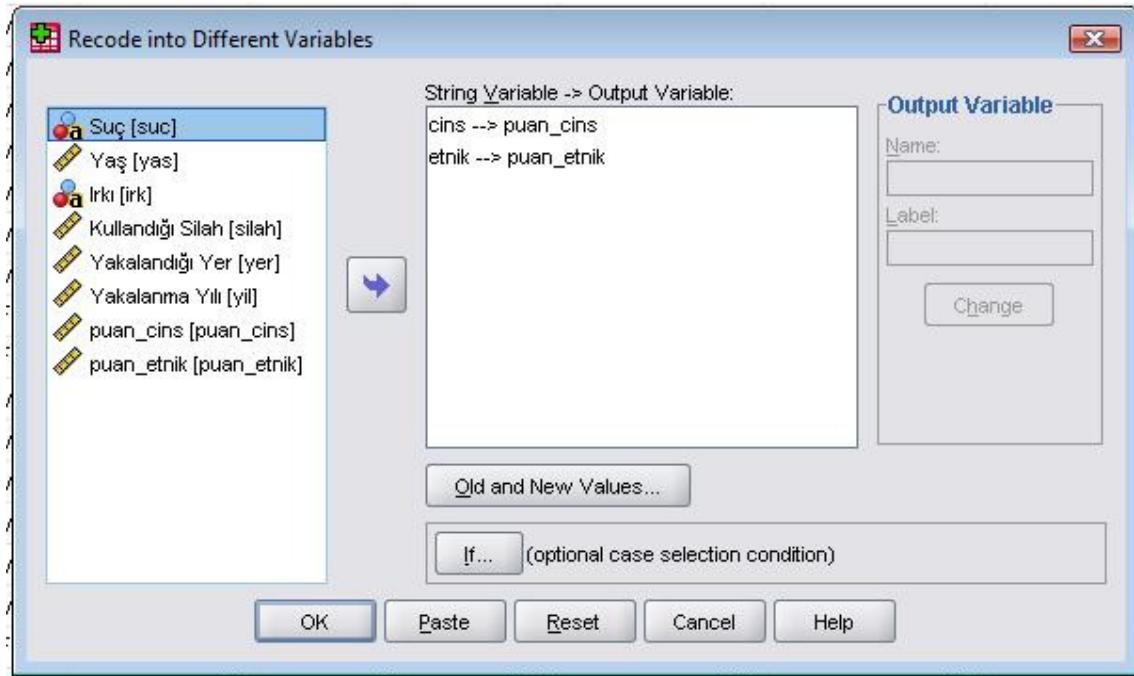
Şimdi bir an için şöyle düşünelim. Her denek istediği suç itibarı ile değerlendirilip yargılanacaktır. Bu yargılama yapılrken deneğin yaşı, cinsiyeti, etnik kökeni, ırkı ve yakalanma yılı gibi bazı etkenler dikkate alınacaktır. Bu maksatla her deneğin yaşı, cinsiyeti, ırkı, etnik kökeni ve yakalanma yılı gibi bazı değişkenleri sayısal veriye dönüştürülüp bunlar dikkate alınarak deneğe bir puan verilecek ve yargılama bu puan da etkili olacaktır. Bu maksatla bu değişkenleri SPSS de bu değişkenleri sayısal verilere dönüştürmeliyiz. Bunun için **transform/recode** ve **transform/compute** komutlarını kullanabiliriz.





suc	yas	etnik	irk	cins	silah	yer	yil	puan_cins
13B	17 N	W	M		95	22	1998	1,00
13A	16 N	B	F		95	22	1998	2,00
13C	15 N	B	F		95	22	1998	2,00
13C	13 N	B	F		95	22	1998	2,00
13C	34 N	W	M		95	22	1998	1,00
120	12 N	W	M		95	22	1998	1,00
13A	14 N	W	M		95	22	1998	1,00
13B	15 N	B	M		95	22	1998	1,00
13B	13 U	B	M		95	22	1998	1,00
13B	17 N	B	M		95	22	1998	1,00
13B	18 N	W	M		95	22	1998	1,00
13B	15 N	B	F		95	22	1998	2,00
13B	17 N	B	F		95	22	1998	2,00
13C	10 N	W	M		95	22	1998	1,00
13A	14 N	B	M		95	22	1998	1,00
13A	14 N	B	M		95	22	1998	1,00
13B	14 N	B	M		95	22	1998	1,00

Yukarıdaki üç şekilde, cinsiyet değişkenleri sayısal veriye çevrilmiştir. Bunun için yeni bir değişken tanımlanmış ve bu değişkene, cinsiyet bayan ise 2 diğer durumlarda ise 1 değeri atanmıştır.



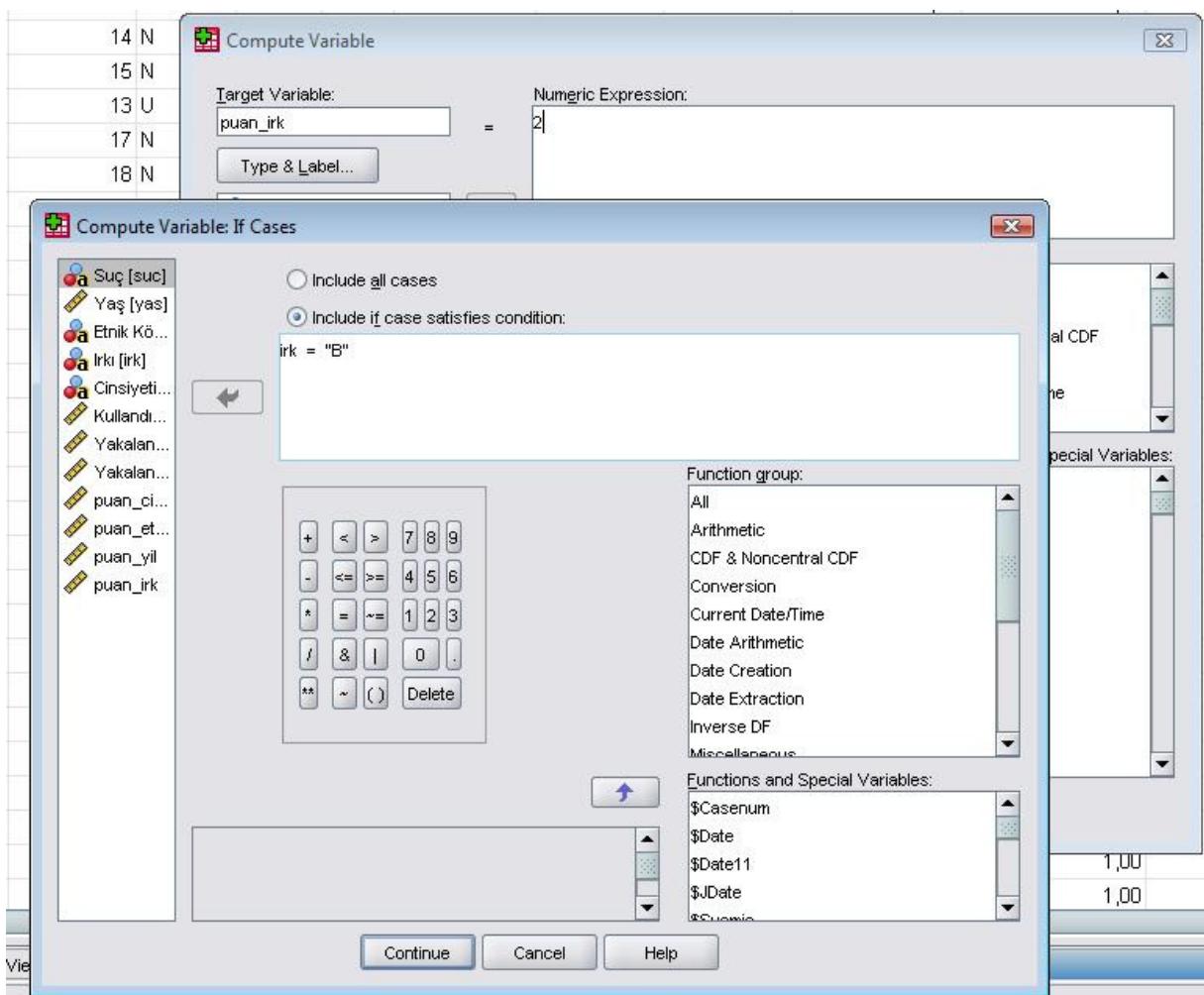
lah	yer	yil	puan_cins	puan_etnik
95	22	1997	2,00	1,00
95	22	1997	1,00	1,00
95	22	1997	2,00	2,00
95	22	1997	1,00	1,00
95	22	1997	2,00	1,00
95	22	1997	2,00	1,00
95	22	1997	1,00	1,00
95	22	1997	1,00	1,00
95	22	1997	1,00	1,00

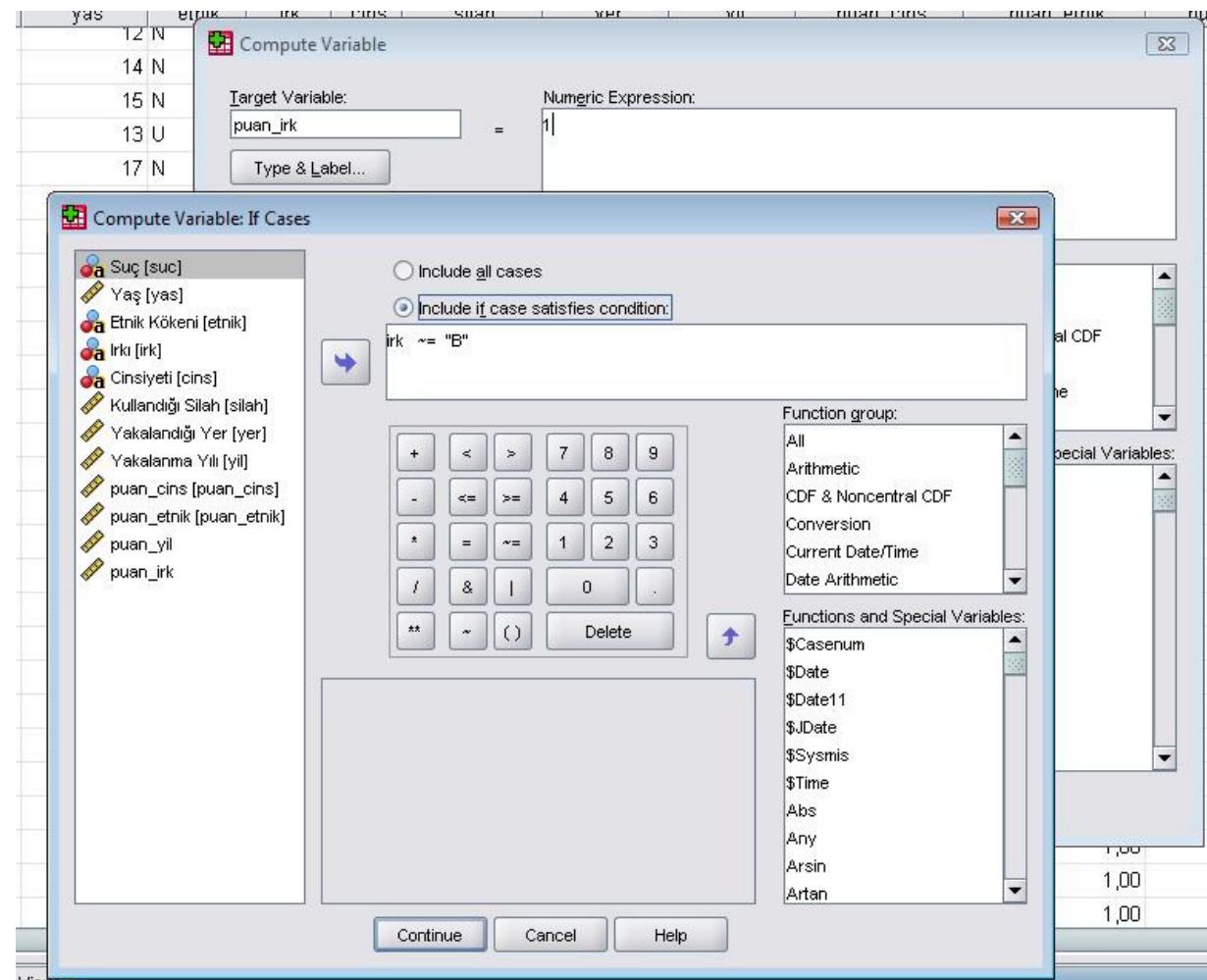
Yukarıdaki resimlerde ise, önceki örnekle benzer yolla, etnik köken bilgisi sayısal bilgiye dönüştürülmüştür. İspanyollar için 2, diğerleri için 1 değeri yeni değişkene atanmıştır.

The screenshot shows the 'Compute Variable' dialog box in SPSS. The 'Target Variable:' field contains 'puan_yil'. The 'Numeric Expression:' field contains '2010 - yil'. On the left, a list of variables is shown: Suc [suc], Yaş [yas], Etnik Kökeni [etnik], Irk [irk], Cinsiyeti [cins], Kullandığı Silah [silah], Yakalandığı Yer [yer], Yakalanma Yılı [yıl], puan_cins [puan_cins], and puan_etnik [puan_etnik]. A 'Type & Label...' button is available for the target variable. Below the expression field is a numeric keypad with operators (+, -, *, /, **, <, <=, =, ~, &, |, (,), Delete) and numbers (0-9). To the right, a 'Function group:' dropdown lists categories like All, Arithmetic, CDF & Noncentral CDF, etc., and a 'Functions and Special Variables:' list is partially visible. At the bottom, there's an 'If...' button for optional case selection and standard dialog buttons: OK, Paste, Reset, Cancel, and Help.

yer	yıl	puan_cins	puan_etnik	puan_yıl
22	1998	2,00	1,00	12,00
22	1998	1,00	1,00	12,00
22	1998	1,00	1,00	12,00
22	1998	1,00	1,00	12,00
22	1998	1,00	1,00	12,00
22	1997	2,00	1,00	13,00
22	1997	1,00	1,00	13,00
22	1997	1,00	1,00	13,00
22	1997	1,00	1,00	13,00
22	1997	2,00	1,00	13,00

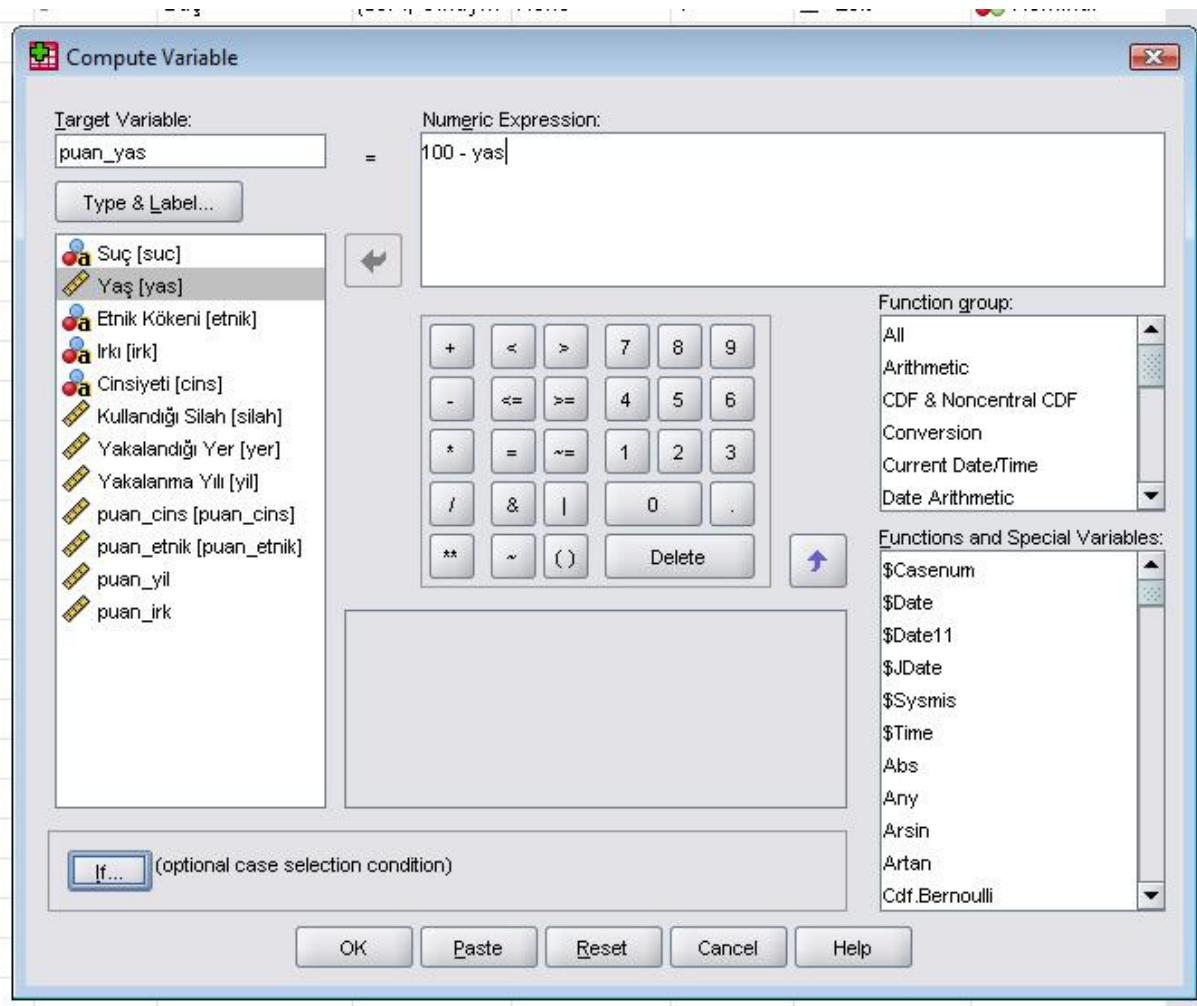
Yukarıdaki iki şekilde gösterilen ise yıl değişkeninin yeni bir değişkene dönüştürülmesi işlemidir. Bu işlem için compute komutu kullanılmıştır. Her deneğin yakalanma yılının 2010 yılı ile farkı alınıp bu değer yıl puanı olarak atanmıştır.





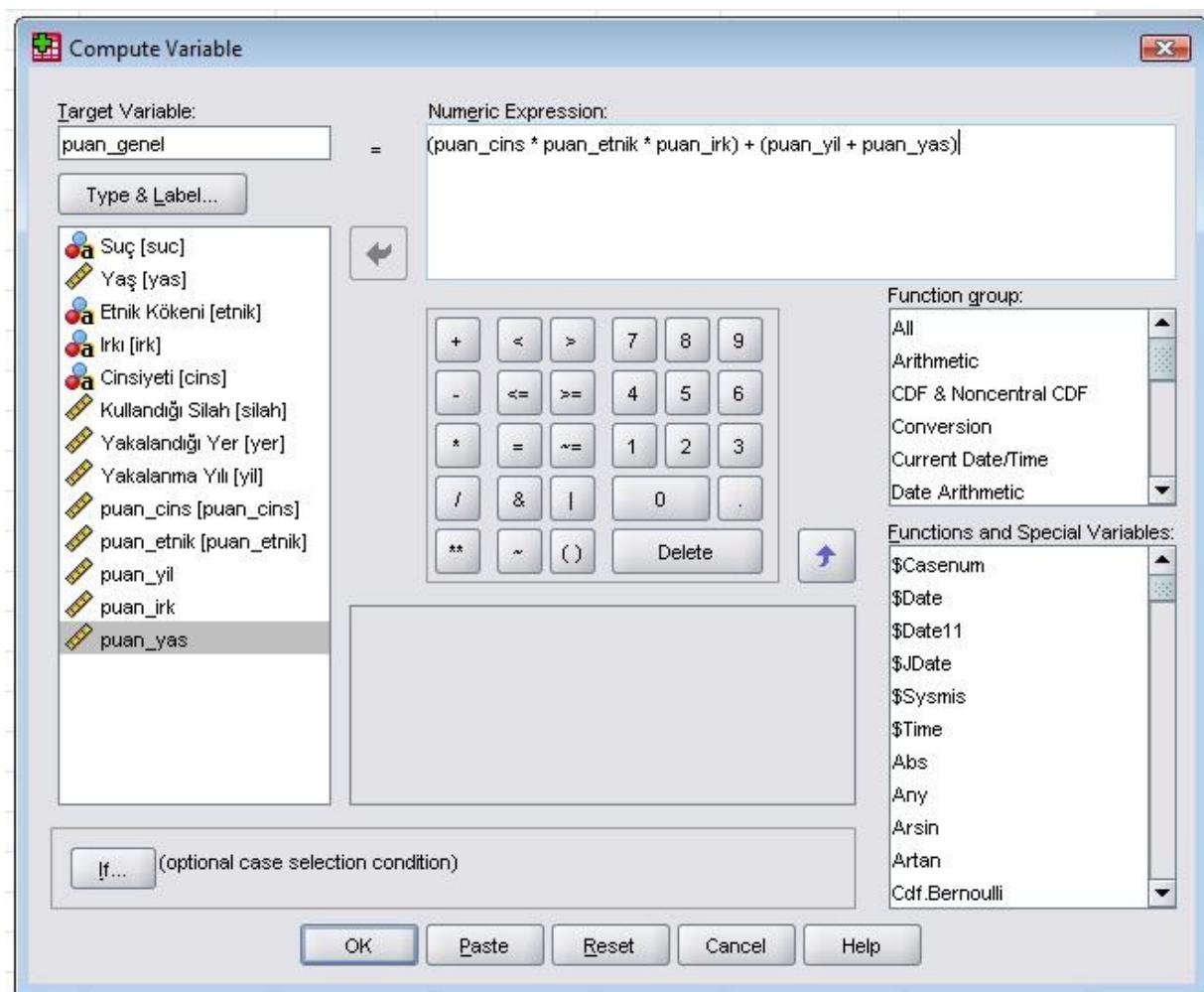
	puan_cins	puan_etnik	puan_yil	puan_irk
8	2,00	1,00	12,00	2,00
8	1,00	1,00	12,00	1,00
8	1,00	1,00	12,00	2,00
8	1,00	1,00	12,00	2,00
8	1,00	1,00	12,00	2,00
8	1,00	1,00	12,00	1,00
8	1,00	1,00	12,00	2,00
8	2,00	1,00	12,00	2,00
8	1,00	1,00	12,00	2,00
8	2,00	1,00	12,00	2,00
8	1,00	1,00	12,00	2,00

Yukarıdaki üç şekilden de anlaşılacağı gibi, compute komutu ve if case özelliği kullanılarak ırk değişkeni sayısal veriye dönüştürülmüştür. Buna göre siyahilerin ırk puanı 2, diğerlerinin ise 1 olarak tanımlanmış ve yeni değişkene atanmıştır.



var	puan_yas	puan_irk	puan_yil	ian_etnik
	83,00	1,00	12,00	1,00
	84,00	2,00	12,00	1,00
	85,00	2,00	12,00	1,00
	87,00	2,00	12,00	1,00
	66,00	1,00	12,00	1,00
	88,00	1,00	12,00	1,00
	86,00	1,00	12,00	1,00
	85,00	2,00	12,00	1,00
	87,00	2,00	12,00	1,00
	83,00	2,00	12,00	1,00
	82,00	1,00	12,00	1,00

Yukarıdaki şekilde ise benzer şekilde her denegen yaşının 100 ile farkı alınıp bu değer yaş puanı olarak yeni bir değişkene atanmıştır.



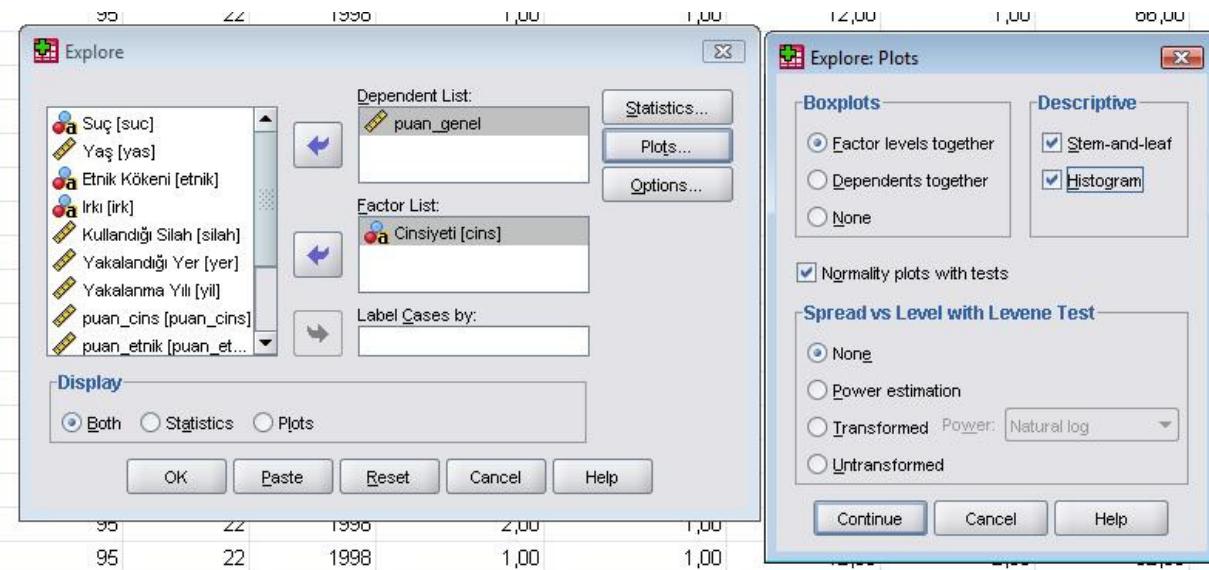
puan_yil	puan_irk	puan_yas	puan_genel	var
12,00	1,00	83,00	96,00	
12,00	2,00	84,00	100,00	
12,00	2,00	85,00	101,00	
12,00	2,00	87,00	103,00	
12,00	1,00	66,00	79,00	
12,00	1,00	88,00	101,00	
12,00	1,00	86,00	99,00	
12,00	2,00	85,00	99,00	
12,00	2,00	87,00	101,00	
12,00	2,00	83,00	97,00	
12,00	1,00	82,00	95,00	
12,00	2,00	85,00	101,00	

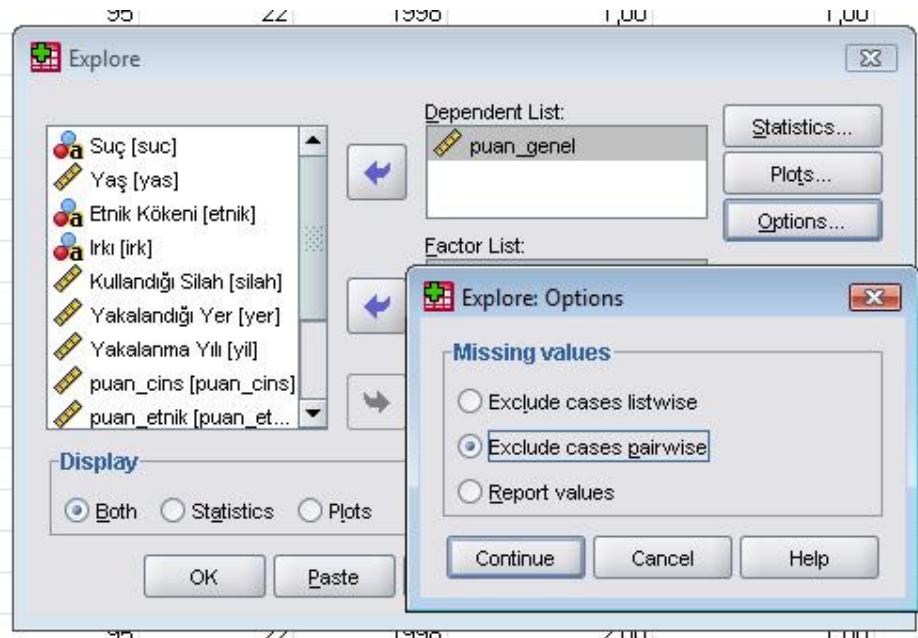
Yukarıda ise, daha önce bazı değişkenlerin sayısal verilere dönüştürülmesi ile tanımlanan yeni değişkenlerin tamamı dikkate alınarak her denegen değerlendirme puanı hesaplanmıştır. Her denek için değerlendirme puanı: $cins_puani * etnik_puani * ırk_puani + (yıl_puani + yaş_puani)$ formülü kullanılarak hesaplanmıştır.

Statistics						
	Cinsiyeti	Irki	Etnik Kökeni	Yakalama Yılı	Yaş	puan_genel
N	Valid	381	381	381	381	381
	Missing	0	0	0	0	0
Mean				1997,11	16,08	99,0787
Std. Error of Mean				,041	,351	,35455
Median				1997,00	15,00	100,0000
Std. Deviation				,802	6,849	6,92053
Variance				,642	46,907	47,894
Skewness				-,207	5,026	-4,850
Std. Error of Skewness				,125	,125	,125
Kurtosis				-1,415	30,302	28,834
Std. Error of Kurtosis				,249	,249	,249
Range				2	66	65,00
Minimum				1996	8	42,00
Maximum				1998	74	107,00

Yukarıdaki tablodan da görüleceği gibi deneklerin genel puanı 42 ile 107 arasında değişmektedir. Bu değerler kullanarak bazı analizler yapacağız.

İstatistiksel analiz yapılrken verilerin dağılımı çok önemlidir. Biz de şimdî verilerimizin normal dağılıma sahip olup olmadıklarını inceleyeceğiz. Bunun farklı yolları vardır. Şimdi bir yolunu inceleyeceğiz. Descriptive statistics menüsünde explore komutuyla aşağıdaki gibi komut verelim.





Yukarıdaki seçenekleri seçip OK a tıklarsak aşağıdaki çıktıları alırız:

Case Processing Summary

Cinsiyeti	Cases						
	Valid		Missing		Total		
	N	Percent	N	Percent	N	Percent	
puan_genel	Bayan	109	100,0%	0	,0%	109	100,0%
	Erkek	272	100,0%	0	,0%	272	100,0%

Descriptives

Cinsiyeti				Statistic	Std. Error
puan_genel	Bayan	Mean		99,4220	,83880
		95% Confidence Interval for Mean	Lower Bound	97,7594	
			Upper Bound	101,0847	
		5% Trimmed Mean		100,9393	
		Median		101,0000	
		Variance		76,691	
		Std. Deviation		8,75732	
		Minimum		42,00	
		Maximum		106,00	
		Range		64,00	
		Interquartile Range		3,00	
		Skewness		-4,303	,231
		Kurtosis		21,381	,459
	Erkek	Mean		98,9412	,36646
		95% Confidence Interval for Mean	Lower Bound	98,2197	
			Upper Bound	99,6626	
		5% Trimmed Mean		99,6716	
		Median		100,0000	
		Variance		36,528	
		Std. Deviation		6,04383	
		Minimum		49,00	
		Maximum		107,00	
		Range		58,00	
		Interquartile Range		3,00	
		Skewness		-5,148	,148
		Kurtosis		33,379	,294

Yukarıdaki iki tabloda verilerin bir özeti verilir.

Tests of Normality

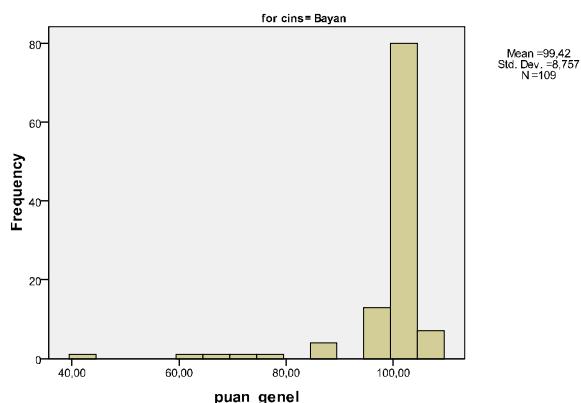
Cinsiyeti	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
puan_genel	.344	109	,000	,482	109	,000
	,267	272	,000	,521	272	,000

^a Lilliefors Significance Correction

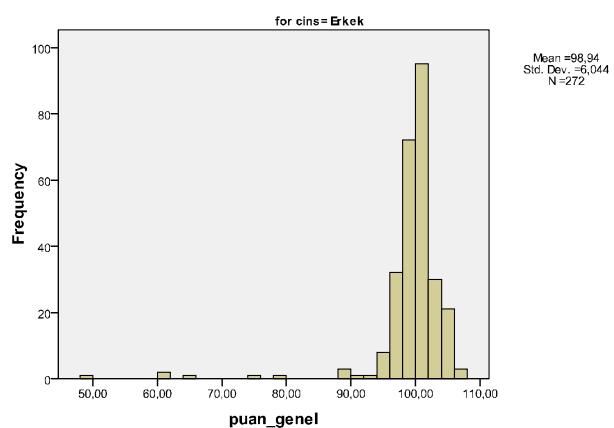
Burada verilerin normal dağılıma uyup uymadığı konusunda iki test yöntemi yardımıyla bilgi verilmiştir. Bu tabloyu yorumlayalım. Burada sig değerleri sözkonusu olduğuna göre H_0 ve alternatif hipotezlerimiz vardır.

Burada **H_0 : veriler normal dağılıma uyar**, ve alternatif hipotez de veriler normal dağılıma uymaz şeklindedir. Tablodan da anlaşılmacağı gibi, her iki test sonucuna göre de sig değerleri 0.05 den küçük olduğu için H_0 hipotezini reddederiz. Yani verinin normal dağılıma uygun olmadığını söyleyiz.

Histogram



Histogram



Yukarıdaki iki grafik ile erkek ve bayanların genel puanlara göre dağılımı gösterilmiştir. Bu grafiklerin yorumlanması ileride detaylı olarak vereceğiz. Bu yüzden şimdi üzerinde durmayacağız.

```

puan_genel Stem-and-Leaf Plot for
cins= Bayan

Frequency      Stem & Leaf

9,00 Extremes    (=<89,0)
1,00          96 . 0
,00          97 .
5,00          98 . 00000
7,00          99 . 0000000
14,00         100 . 0000000000000000
24,00         101 . 00000000000000000000000000
15,00         102 . 0000000000000000
16,00         103 . 0000000000000000
11,00         104 . 000000000000
5,00          105 . 00000
2,00          106 . 00

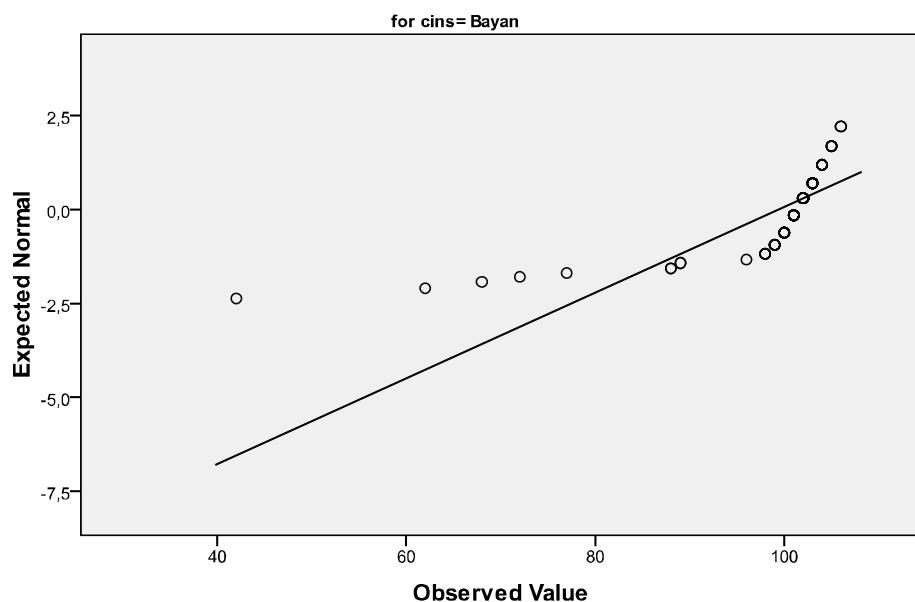
Stem width:      1,00
Each leaf:      1 case(s)

```

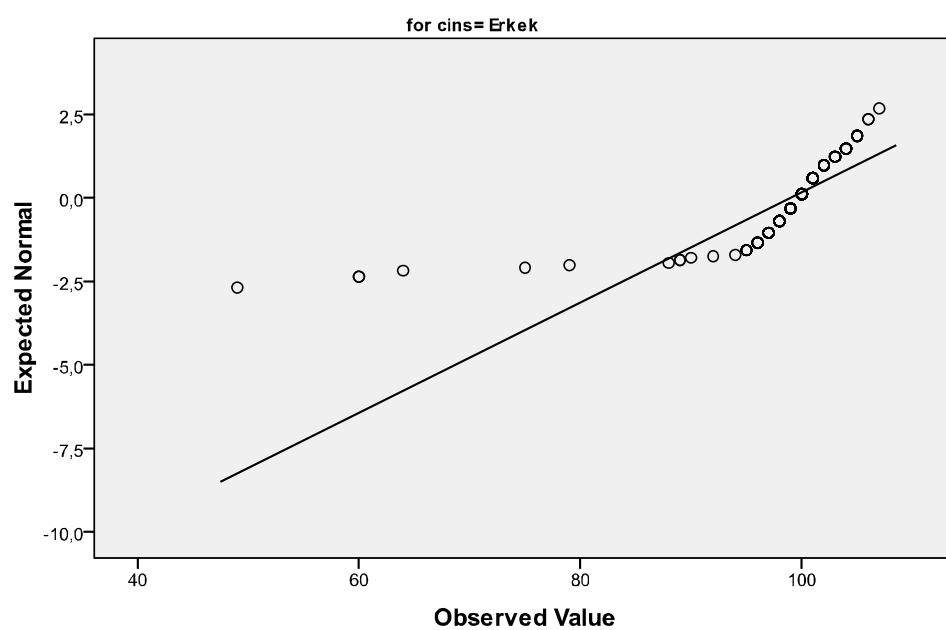
Yukarıdaki iki çizelge ise dal-yaprak grafiği olarak adlandırılır. Şimdi bunun nasıl yorumlanacağına bakalım. Ortadaki sütun dal, sağdakiler ise yaprak olarak adlandırılır. Dal değeri ile yaprak değeri birleştirilerek oluşturulan sayı yaprak değeri sayısı ile yaprak durum çarpanı (leaf case) ile çarpılarak toplam durum hesaplanır.

Bunlar da sol sütunda verilmiştir. Örneğin puanı 95 olan deneklerin sayısını bulalım. $95+0=95$ olup karşısında 7 tane yaprak vardır. Yaprak durum çarpanı 1 olduğu için $1*7=7$ olup toplamda 7 tane deneğin puanı 95 tir deriz.

Normal Q-Q Plot of puan_genel

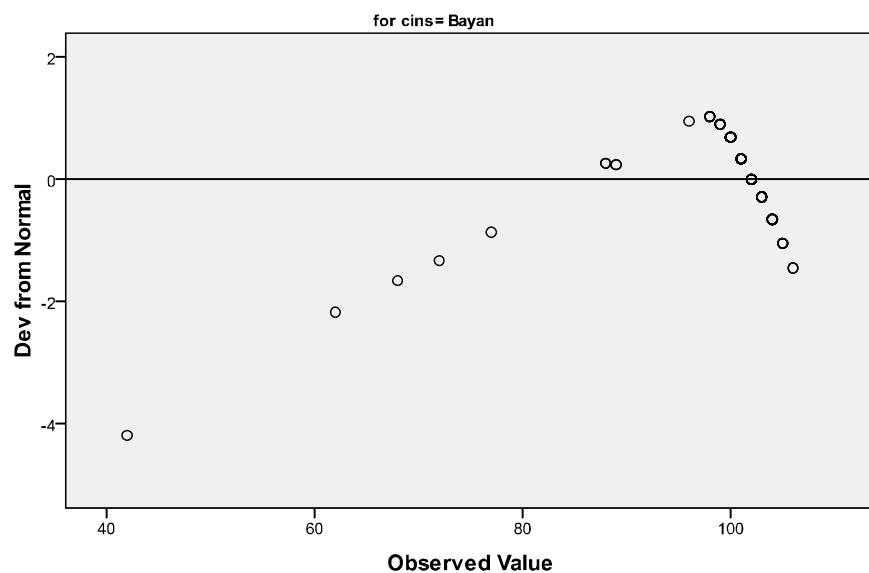


Normal Q-Q Plot of puan_genel

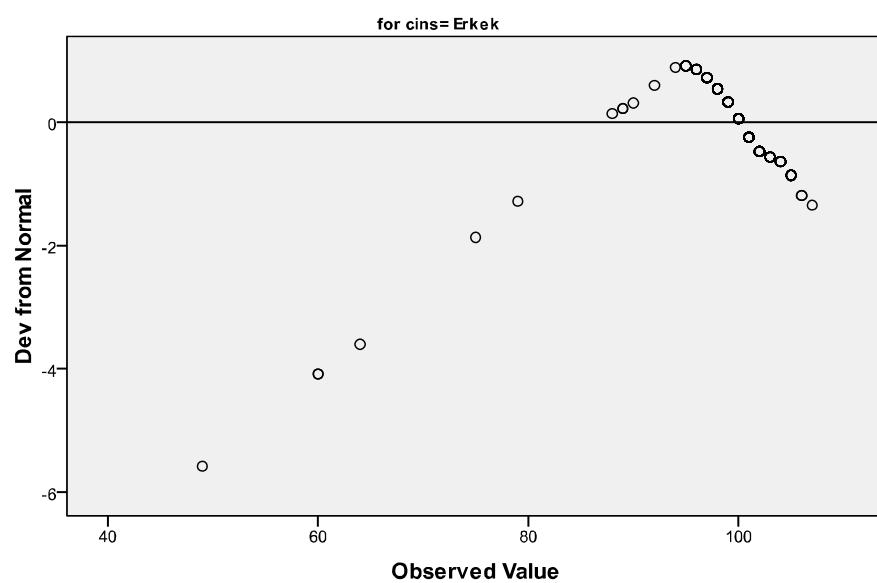


Yukarıdaki iki grafikte erkeklerin ve bayanların puanlarının normal dağılıma uyup uymadıkları görsel olarak verilmiştir. Yuvarlak ile gösterilen puan değerleri, çizgi ile verilen grafik ise normal dağılım grafiğidir. Görüldüğü gibi puan değerleri normal dağılım grafiğiyle çakışmamaktadır. Böylece, az önce testler yardımıyla verdığımız sonuç gibi, bu dağılımin normal dağılıma uymadığı sonucuna varız.

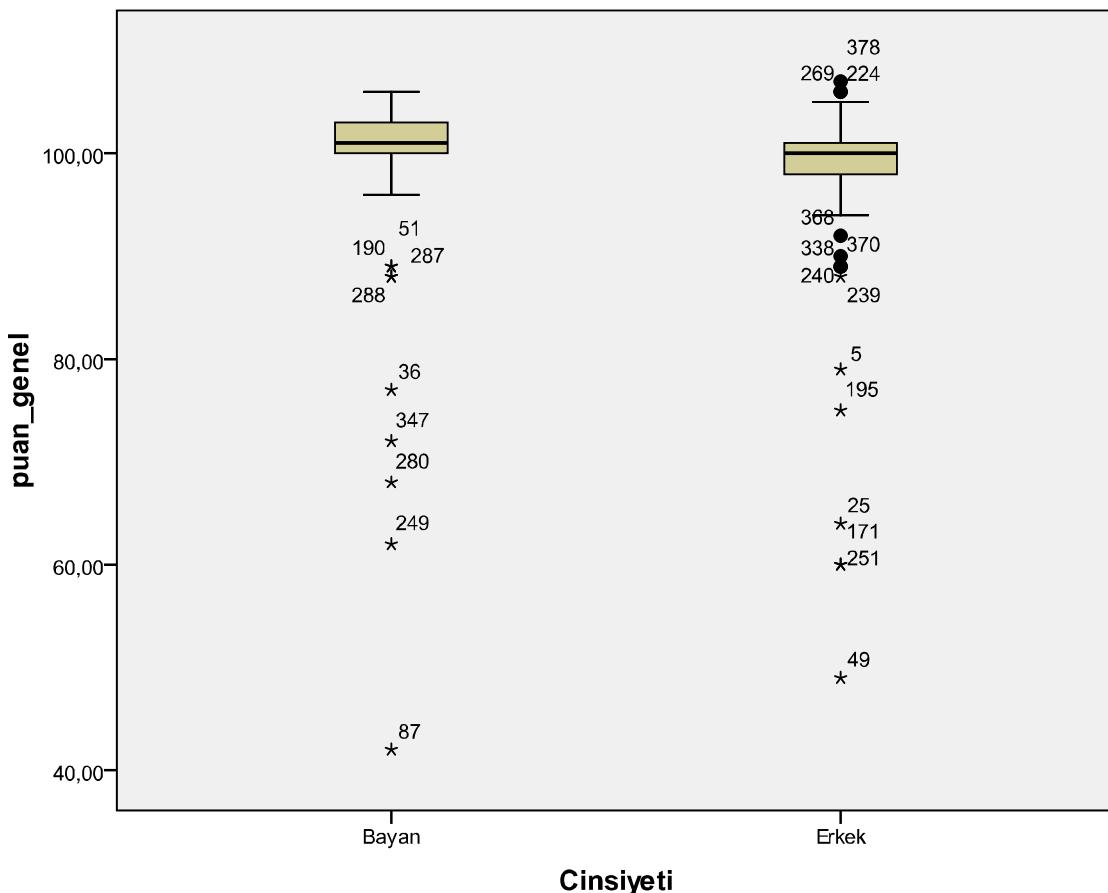
Detrended Normal Q-Q Plot of puan_genel



Detrended Normal Q-Q Plot of puan_genel



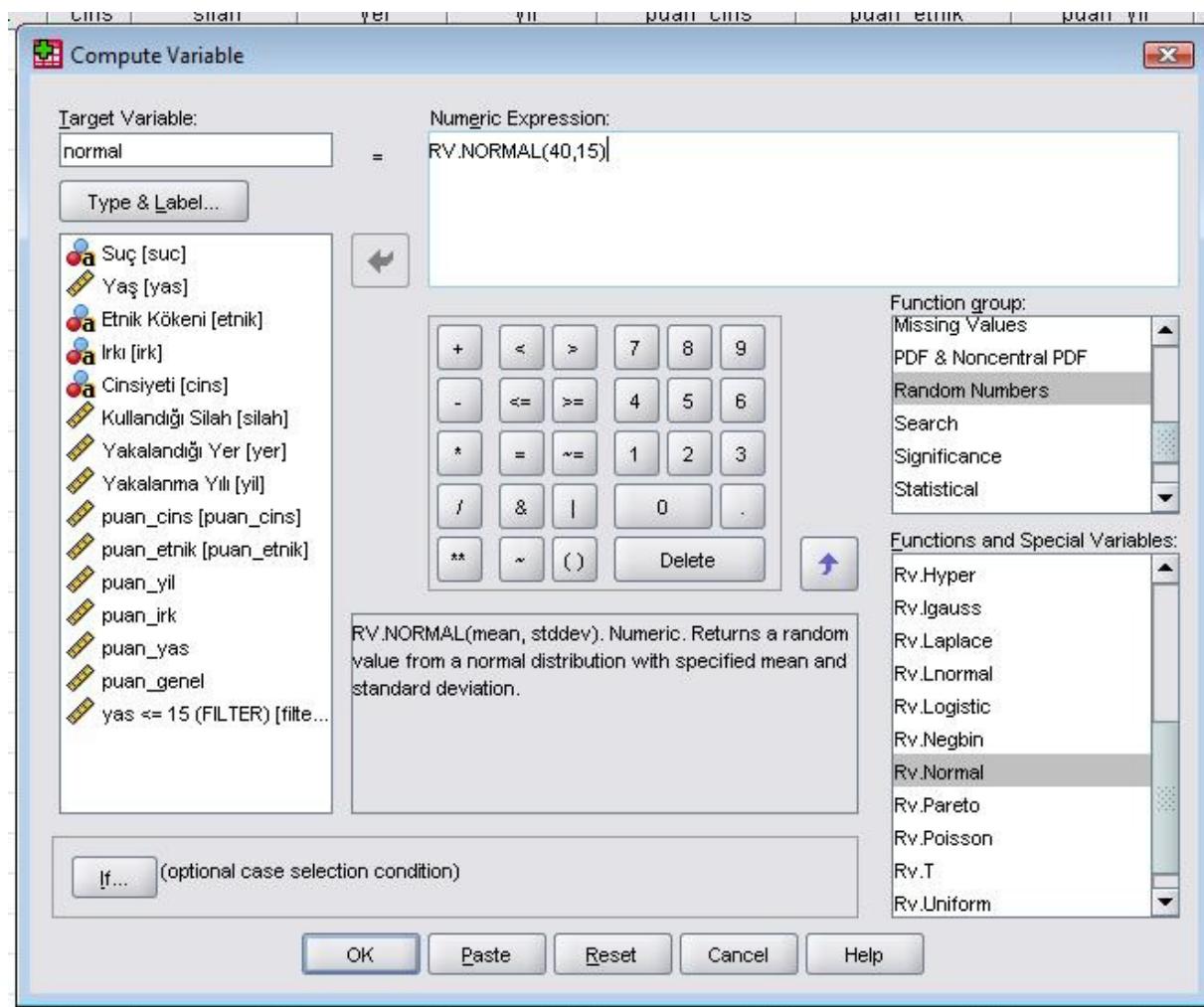
Yukarıdaki iki grafik ile de puan verilerinin normal dağılımdan ayrıldığı noktalar verilmiştir. Bunların toplamı sıfır oluyorsa yada bu noktalar sıfır civarında kümelenmişse dağılımin normal dağılıma uyduğu söylenir. Görüldüğü gibi normal dağılıma uymamaktadır.



Yukarıdaki grafik ise boxplot grafiği olarak adlandırılır. Veriler hakkında oldukça kapsamlı bilgi verir. Bu grafiğin yorumlanması daha sonra detaylı olarak deşineceğiz. Şimdi burayı geçiyoruz.

Gördüğümüz gibi puan değişkenine ait veriler normal dağılıma uymamaktadır. Şimdi kendimiz bir değişken tanımlayıp normal dağılıma uygun veriler girelim ve tekrar bu yöntemle kontrol edelim.

Bunun için tekrar compute komutu ve rv.normal fonksiyonunu kullanalım. Aşağıdaki işlemleri yaparak komut



verelim.

Bu komut ile aşağıdaki sonuç elde edilir:

1_yas	puan_genel	filter_\$	normal
83,00	96,00	0	26,35
84,00	100,00	0	39,44
85,00	101,00	1	35,87
87,00	103,00	1	34,59
66,00	79,00	0	12,16
88,00	101,00	1	13,49
86,00	99,00	1	35,24
---	---	---	---

Şimdi explore komutu ile elde edilecek normalilik testi sonucunu ve grafikleri verelim.

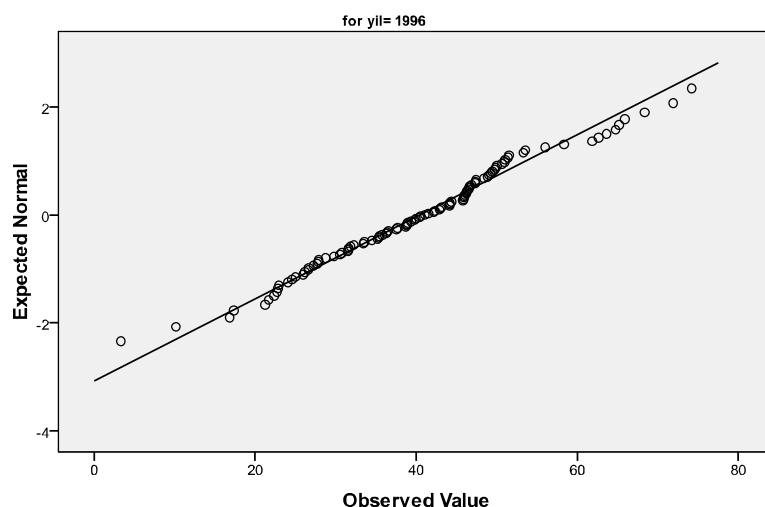
Tests of Normality

Yakalanma Yili	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
normal	1996	,075	103	,175	,988	103	,470
	1997	,041	132	,200*	,993	132	,806
	1998	,044	146	,200*	,994	146	,776

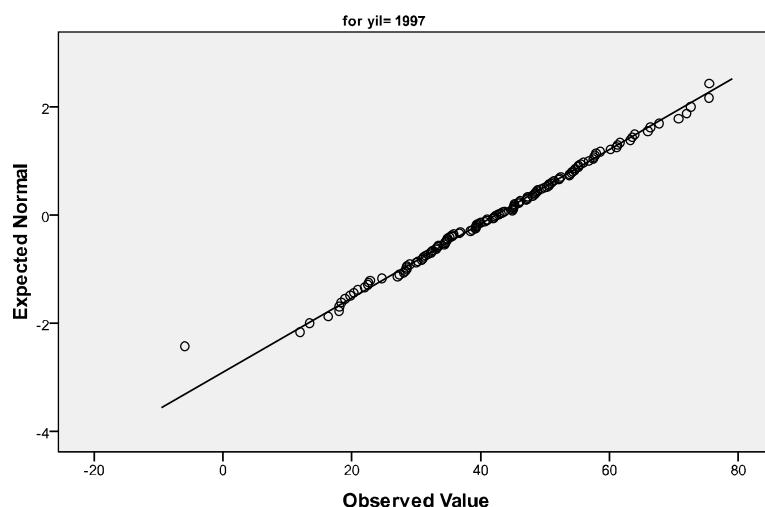
a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

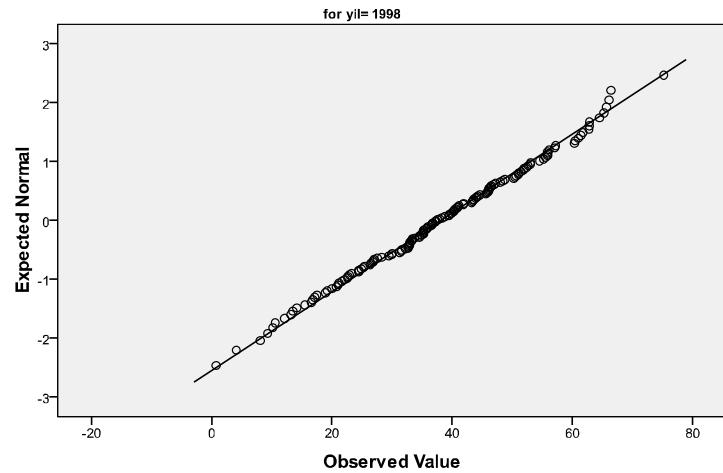
Normal Q-Q Plot of normal



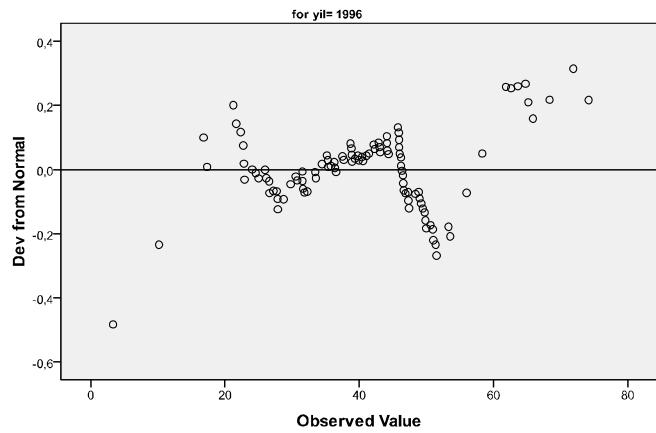
Normal Q-Q Plot of normal



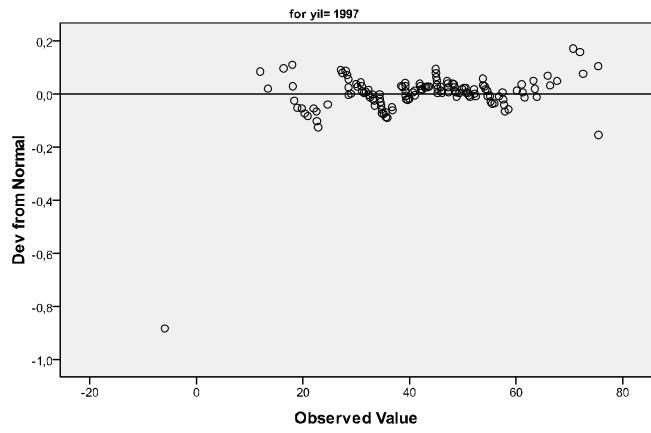
Normal Q-Q Plot of normal



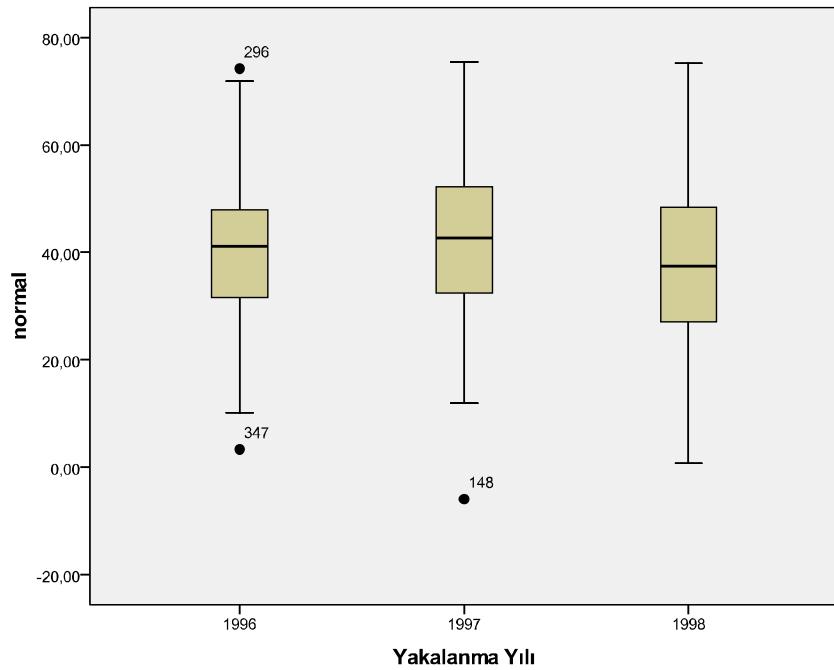
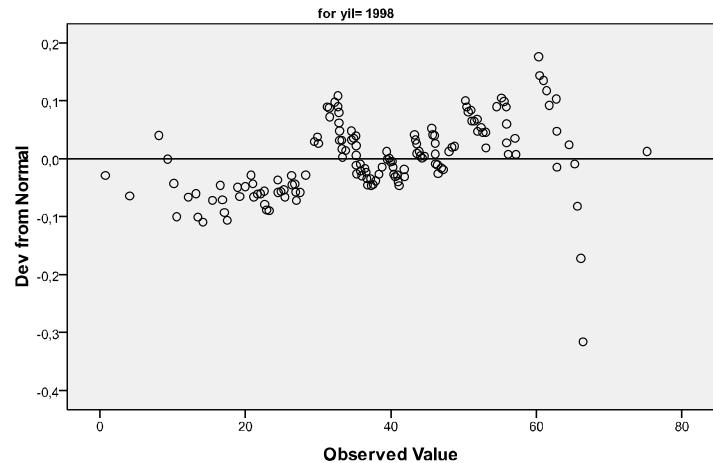
Detrended Normal Q-Q Plot of normal



Detrended Normal Q-Q Plot of normal



Detrended Normal Q-Q Plot of normal



Yukarıdaki tablo ve şekillerden açıkça anlaşılacağı gibi normal değişkenine ait veriler normal dağılıma uymaktadır. Hem sig değeri dolayısıyla H_0 i reddederek, hem de grafiklerde verilerin normal dağılım eğrisiyle çakışmasından bunu söyleyebiliriz.

Şimdi de değişkenlerimizin bağımsızlığını inceleyelim. Bu analizle iki değişkenin birbirine bağlı olduğunu yada birbirinden bağımsız olduğunu söyleyebiliriz. Bunun da birkaç yolu vardır. Biz şimdi birini inceleyeceğiz. Descriptive statistics menüsünde crosstabs komutunu verip aşağıdaki şeillerde gösterilen işlemleri yapalım.

Crosstabs

Row(s): Cinsiyeti [cins] Column(s): Yakalanma Yılı [yıl]

Crosstabs: Statistics

Chi-square Correlations

Nominal

- Contingency coefficient
- Phi and Cramer's V
- Lambda
- Uncertainty coefficient

Nominal by Interval

- Eta
- Kappa
- Risk
- McNemar

Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals: 1

OK Paste Reset Cancel

1,00	1,00	12,00	100,00
2,00	1,00	12,00	101,00
1,00	1,00	12,00	103,00
2,00			79,00
			101,00

Crosstabs

Row(s): Cinsiyeti [cins] Column(s): Yakalanma Yılı [yıl]

Crosstabs: Cell Display

Counts

- Observed
- Expected

Percentages

- Row
- Column
- Total

Residuals

- Unstandardized
- Standardized
- Adjusted standardizer

Noninteger Weights

- Round cell counts Round case weights
- Truncate cell counts Truncate case weights
- No adjustments

OK Paste Reset Cancel

1,00	1,00	12,00	100,00
2,00	1,00	12,00	101,00
1,00	1,00	12,00	103,00
2,00	1,00	12,00	79,00
			101,00

Yukarıdaki seçenekleri seçip komut verirsek aşağıdaki çıktıları alırız:

Cinsiyeti * Yakalanma Yılı Crosstabulation

Cinsiyeti	Bayan	Yakalanma Yılı			Total	
		1996	1997	1998		
Cinsiyeti	Bayan	Count	28	36	45	109
		Expected Count	29,5	37,8	41,8	109,0
		% within Cinsiyeti	25,7%	33,0%	41,3%	100,0%
		% within Yakalanma Yılı	27,2%	27,3%	30,8%	28,6%
		% of Total	7,3%	9,4%	11,8%	28,6%
Erkek	Count	75	96	101	272	
		Expected Count	73,5	94,2	104,2	272,0
		% within Cinsiyeti	27,6%	35,3%	37,1%	100,0%
		% within Yakalanma Yılı	72,8%	72,7%	69,2%	71,4%
		% of Total	19,7%	25,2%	26,5%	71,4%
Total	Count	103	132	146	381	
		Expected Count	103,0	132,0	146,0	381,0
		% within Cinsiyeti	27,0%	34,6%	38,3%	100,0%
		% within Yakalanma Yılı	100,0%	100,0%	100,0%	100,0%
		% of Total	27,0%	34,6%	38,3%	100,0%

Chi-Square Tests

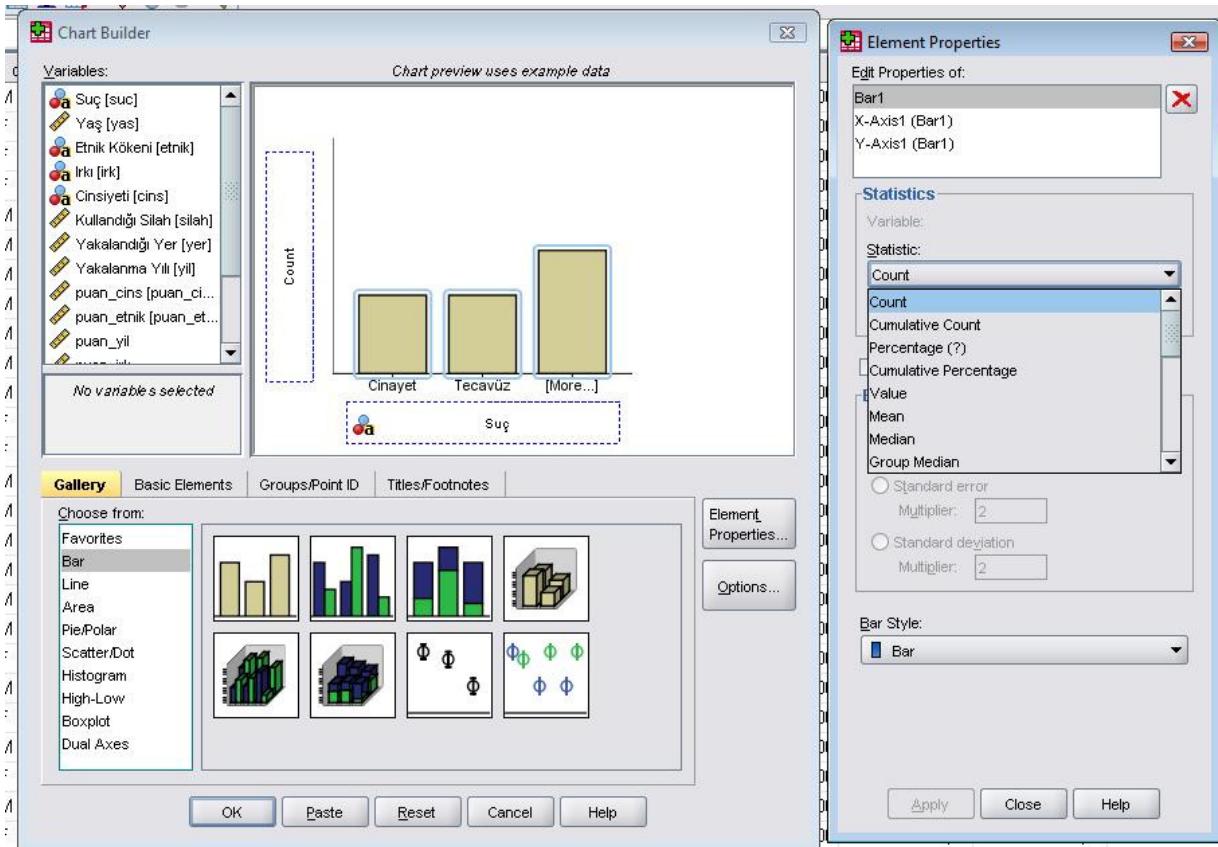
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,568 ^a	2	,753
Likelihood Ratio	,565	2	,754
N of Valid Cases	381		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 29,47.

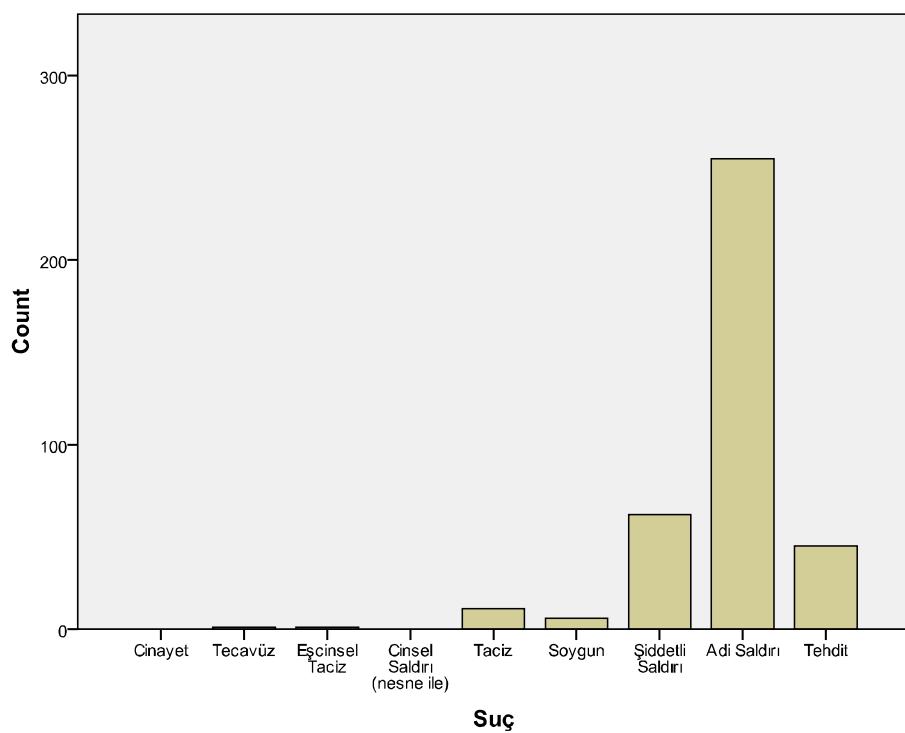
Yukarıdakilerden ikinci tabloyu yorumlayalım. Bu tabloda chi-square testi olarak bilinen test yardımıyla değişkenlerin bağımsılığı incelenmiştir. Buradaki H_0 hipotezimiz: **verilen değişkenler birbirinden bağımsızdır** şeklindedir. Alternatif ise değişkenler bağımlıdır şeklindedir. Sig değeri 0,05 ten büyük olduğu için H_0 hipotezini reddedemiyoruz. Yani deneklerin yakalanma yılları cinsiyetlerine bağımlıdır şeklinde yorumlayabiliriz. Yukarıdaki tablolardan ilkini yorumlayacak olursak, tabloda değişkenlere ait bazı istatistik ve değerlendirmelere yer verildiği görülür. Örneğin; erkeklerden 1996 yılında yakalananların sayısı 73,5 olarak beklenmekteyken 75 olarak gerçekleşmiştir. Ayrıca görmekteyiz ki erkeklerin %27,6 sı 1996, %35,3 ü 1997 ve %37,1 i de 1998 yılında yakalanmıştır. Ayrıca, 1996 yılında yakalananların %72,8 inin erkek olduğu bilgisi de ilk tabloda verilmiştir. benzer şekilde yorumlar yapmak mümkündür.

Şimdi biraz da SPSS programında grafik özelliklerini inceleyelim. SPSS kullanarak verilerin grafiklerle nasıl temsil edileceğini inceleyelim. SPSS birçok grafik çeşidi vasıtasyyla verileri göstermeyi sağlar. Şimdi bazı

grafik türlerinin SPSS ile nasıl elde edileceğini görelim. İlk önce bar tipi denilen çubuk grafiklerinin yapımını inceleyelim.



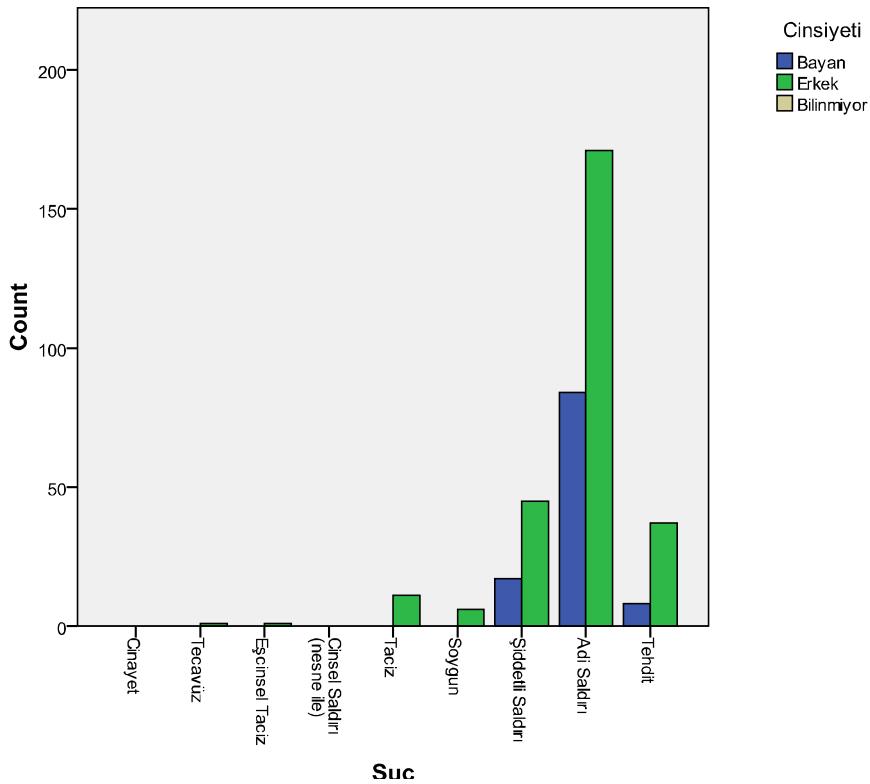
Yukarıdaki şekilde bar tipi grafiğin nasıl yapılabileceği gösterilmiştir. Kısaca özetlersek, SPSS de charts/chart builder komutu verilince yukarıdaki pencere ekranına gelir. Bu pencerede gallery sekmesindeki listeden istediğimiz grafik tipini seçiyoruz. Biz bar tipini seçtik. Sağda açılan alt grafik tiplerinden istediğimizi çift tıklayarak chart preview penceresinde özniteliklere alırız. Biz ilk tip bar grafiğini seçtik. Daha sonra variables listesinden temsil etmek istediğimiz değişkeni preview de istediğimiz aks üzerine bırakırız. Biz burada suç değişkenini atadık. Diğer aksın neyi temsil edeceğini element properties penceresinden statistics açılır listesinden seçebiliriz. Burada da count bilgisini seçtik. Yukarıdaki komutları verince aşağıdaki grafiği elde ederiz. Bu grafiğin suç tiplerinin işleme sayılarını verir. Çünkü y aksisinde count değerini seçmiştık.



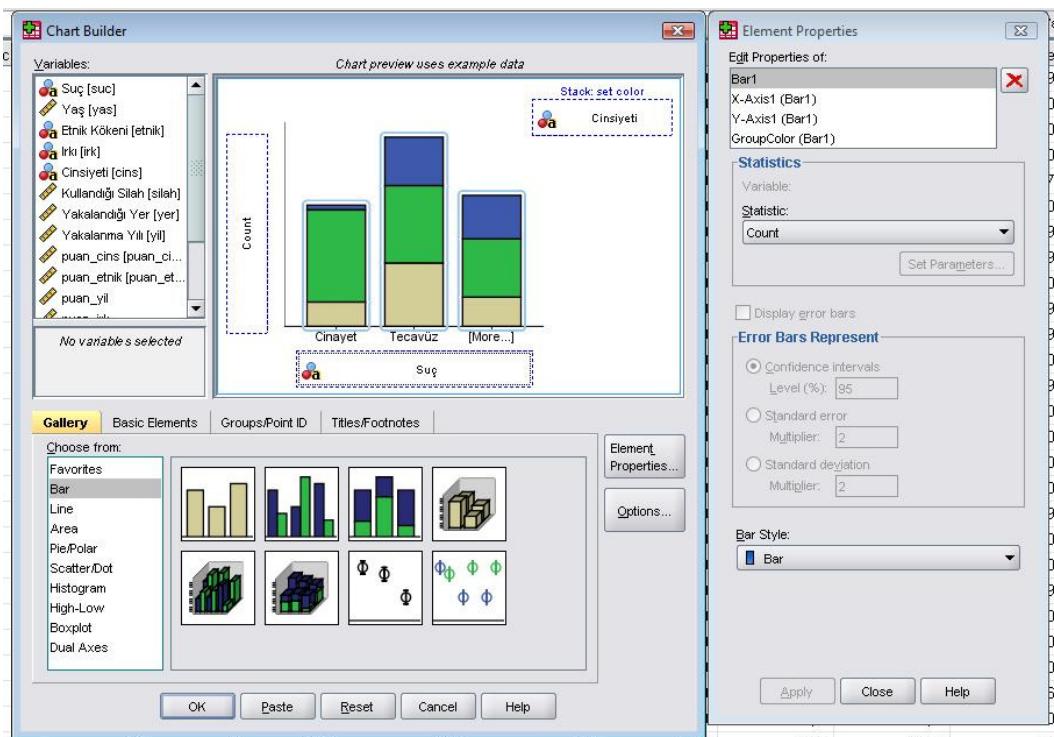
Şimdi diğer bar grafiği tiplerini inceleyelim.

The screenshot shows the SPSS Chart Builder interface. On the left, the 'Variables' list includes 'Suç [suc]', 'Yaş [yas]', 'Etnik Kökeni [etnik]', 'Irki [irk]', 'Cinsiyeti [cins]', 'Kullandığı Silah [silah]', 'Yakalandığı Yer [yer]', 'Yakalanma Yılı [yıl]', 'puan_cins [puan_ci...]', 'puan_etnik [puan_et...]', and 'puan_yıl'. The 'Gallery' tab is selected, showing various chart types like Bar, Line, Area, etc. A preview of a clustered bar chart is shown in the center, with 'Count' on the y-axis and 'Suç' on the x-axis. The 'Element Properties' dialog on the right is set for 'Bar1' with 'Cluster on X: set color' and 'Cinsiyeti' selected. The 'Statistics' section shows 'Variable: Count' and 'Error Bars Represent' set to 'Confidence intervals' at 95% level. Buttons at the bottom include OK, Paste, Reset, Cancel, and Help.

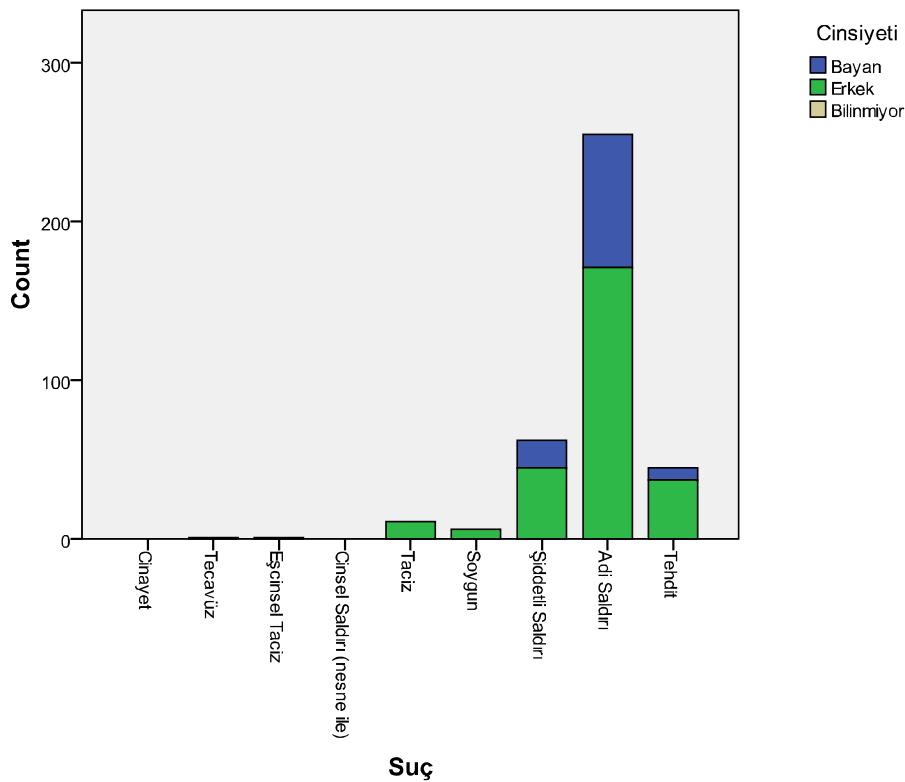
Yukarıdaki şekilde, ikinci tip bar seçik ve previewdeki atamaları yaptık. Bu komutlar ile aşağıdaki çıktıyı alırız.



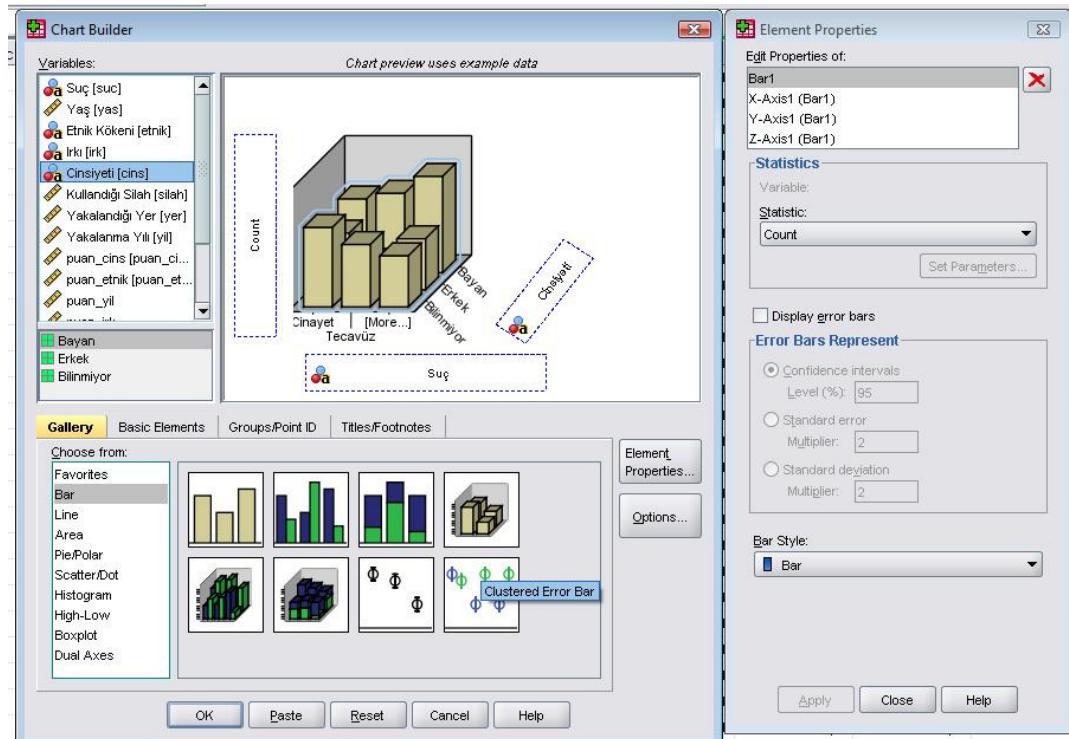
Bu tip grafikle iki farklı değişkenin istatistiklerini tek grafikte görebiliyoruz. Örneğin burada suç tiplerinin cinsiyete göre dağılımını görebiliyoruz.



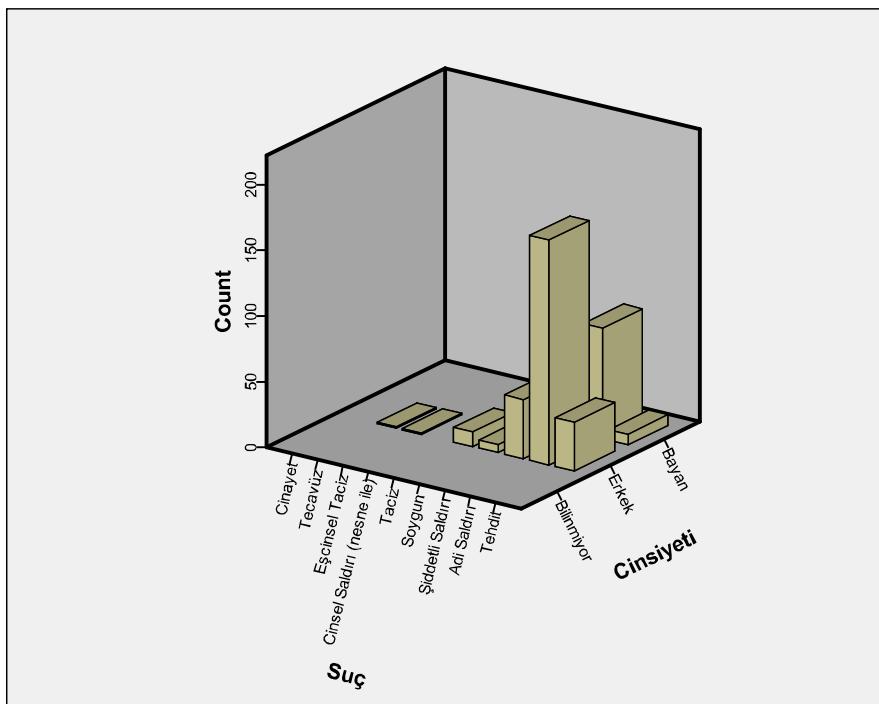
Yukarıdaki komutla da bir önceki grafiğe benzer olan aşağıdaki çıktıyı alırız.



Şimdi kısaca diğer bar tipi grafikleri görelim.



Yukarıdaki komutla öncek grafiklerdeki tür verileri 3 boyutlu olarak görebiliriz. Aşağıdaki çıktı alınır.



Yukarıdaki grafiklerde 2 değişkenin istatistikleri tek grafikte verildi. Üç değişkenin istatistikleri de aşağıdaki grafik tipiyle verilebilir.

Chart Builder

Variables:

- Suç [suc]
- Yaş [yas]
- Etnik Kökeni [etnik]
- İrk [irk]
- Cinsiyeti [cins]
- Kullandığı Silah [silah]
- Yakalandığı Yer [yer]
- Yakalanma Yılı [yil]
- puan_cins [puan_ci...]
- puan_etnik [puan_et...]
- puan_yil
- Ispanyol
- İspanyol Değil
- Bilinmiyor

Gallery Basic Elements Groups/Point ID Titles/Footnotes

Element Properties

Edit Properties of:

- Bar1
- X-Axis1 (Bar1)
- Y-Axis1 (Bar1)
- Z-Axis1 (Bar1)

Statistics

Variable: Statistic: Count

Display error bars

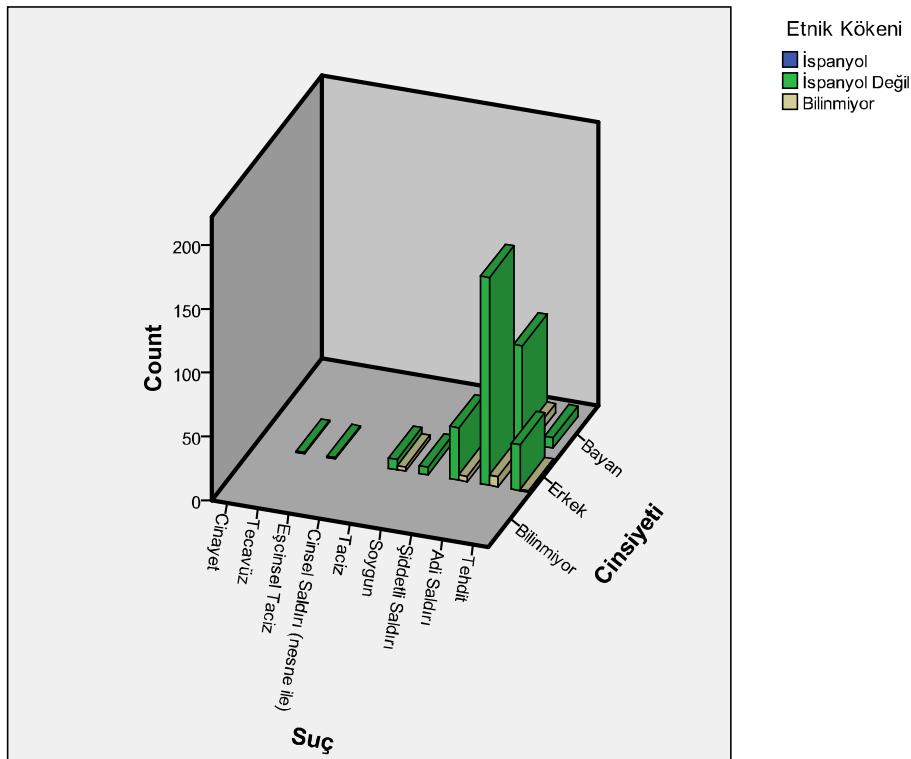
Error Bars Represent

- Confidence intervals Level (%): 95
- Standard error Multiplier: 2
- Standard deviation Multiplier: 2

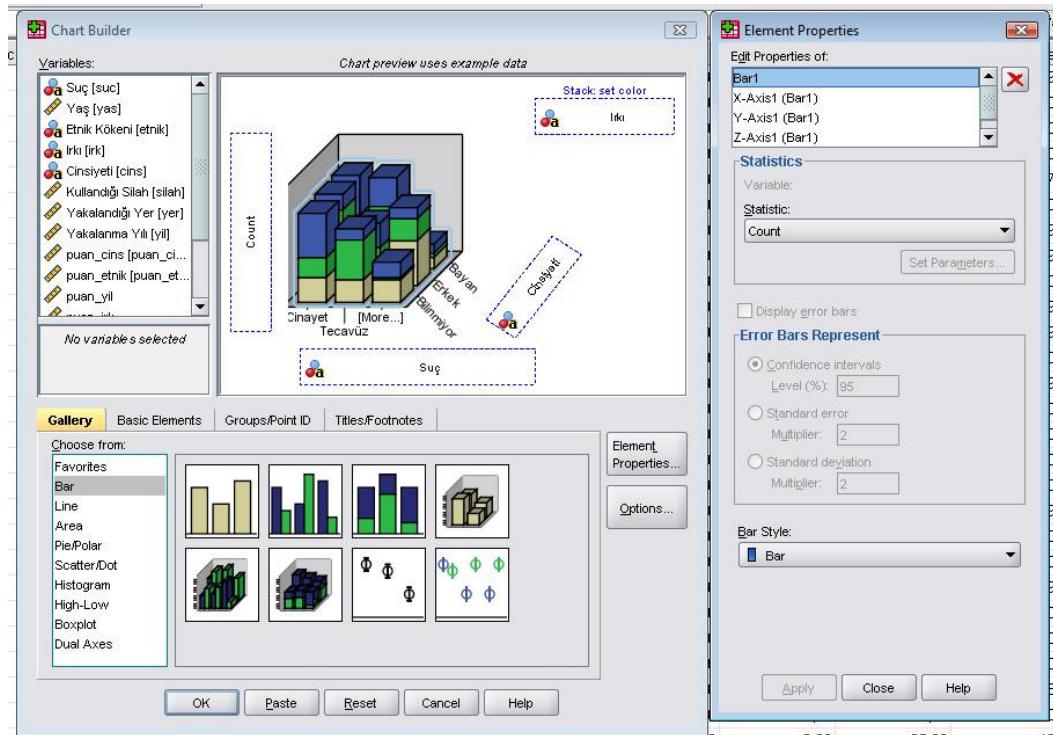
Bar Style: Bar

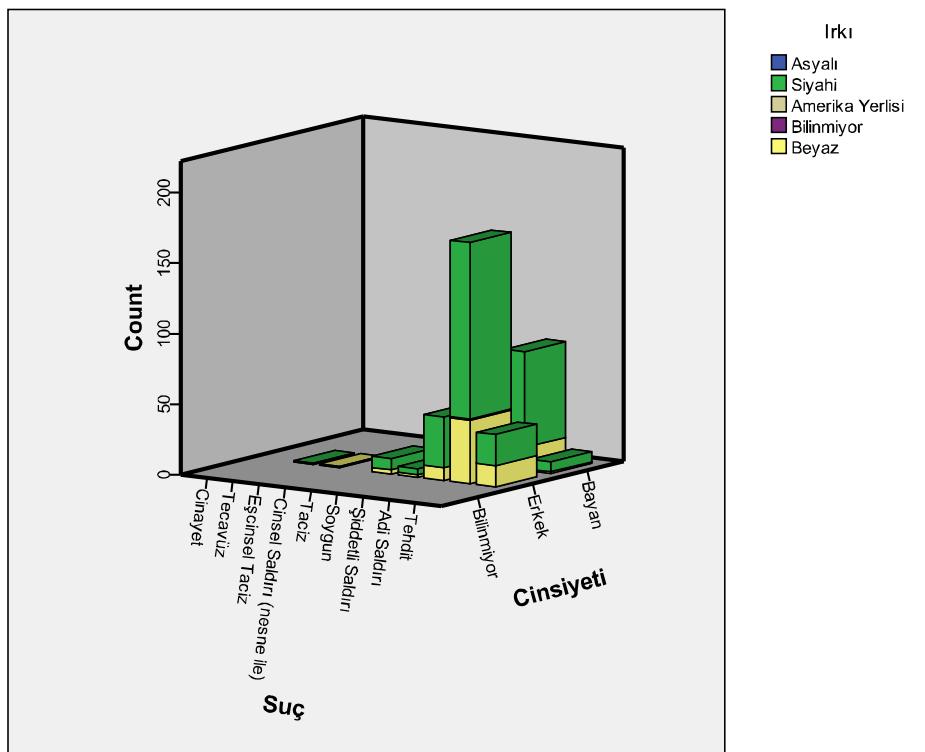
OK Paste Reset Cancel Help Apply Close Help

Bu komutla suçların hem cinsiyete hemde ırka göre dağılımı verilmiştir. Aşağıdaki çıktı alınır.

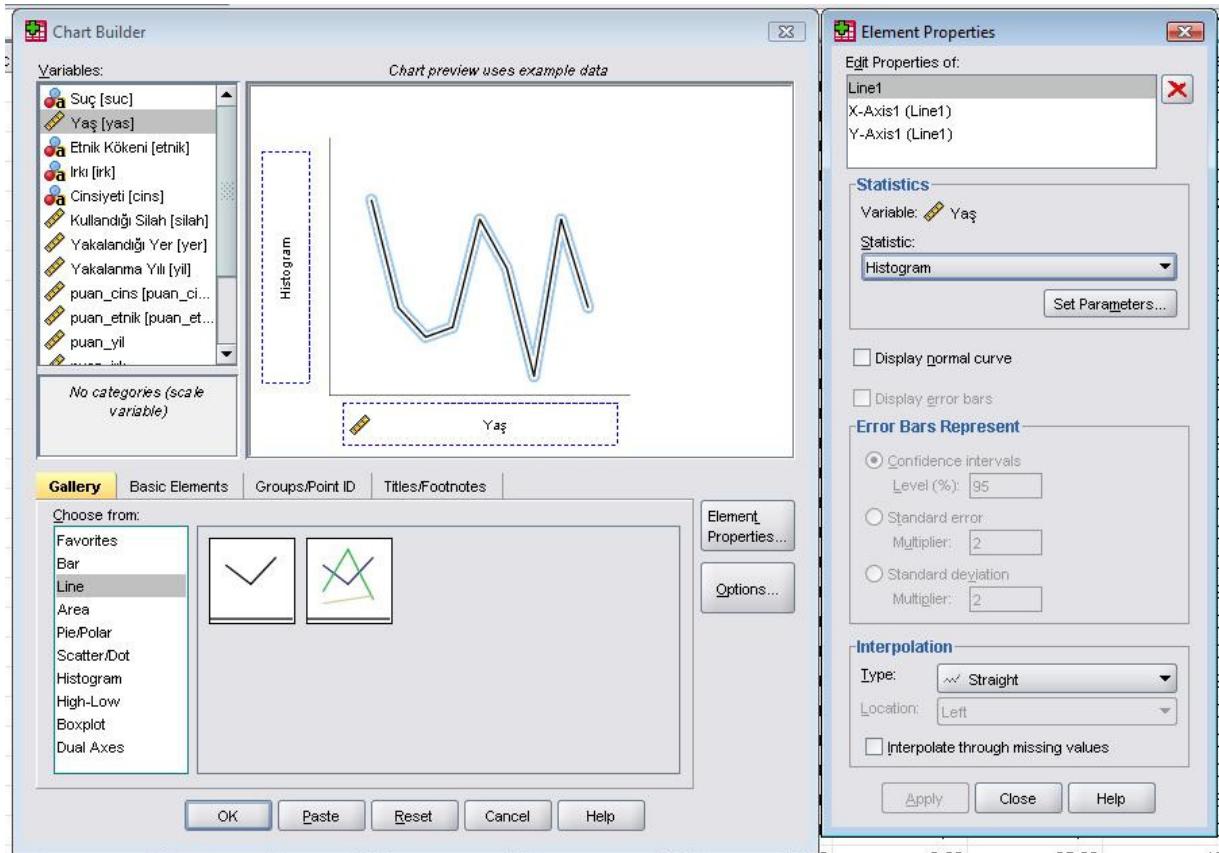


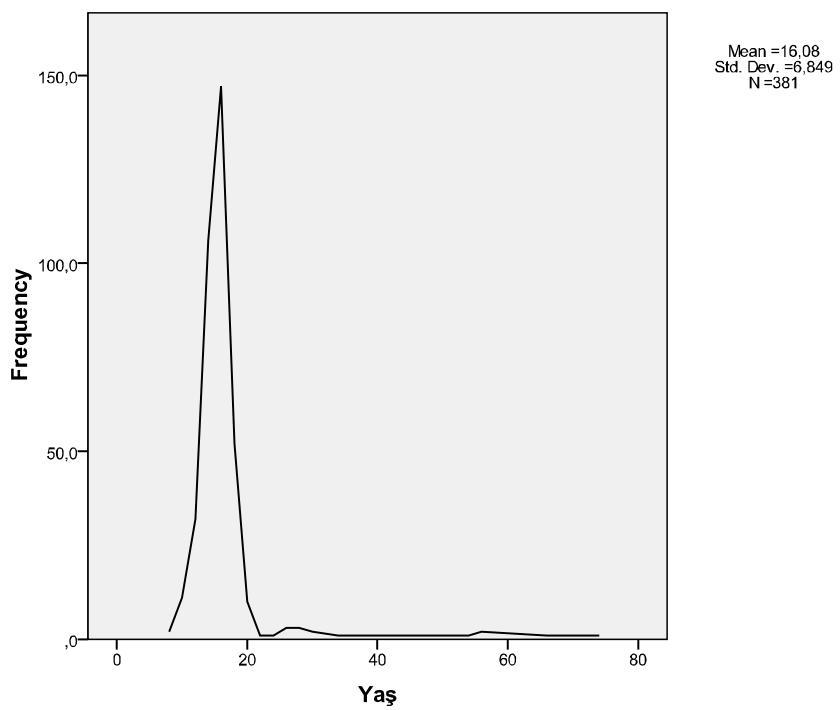
Yukarıdakine benzer şekilde diğer bir grafik seçenekleri de aşağıda verilmiştir.





Diğer bir grafik tipi de line tipi dediğimiz çizgi grafiğidir. Aşağıdaki komut ve çıktıyla örnek verelim.





Gördüğü gibi yaş değişkeninin dağılımı verilmiştir. Range nin geniş olduğu değişkenlerin veya sürekli değişkenlerin istatistiklerini bar yerine line tipi grafik ile göstermek genelde daha avantajlıdır.

Diğer bir line tipi grafiği de aşağıdaki şekilde örnekleylem.

Chart Builder

Variables:

- Suç [suc]
- Yaş [yas]
- Etnik Kökeni [etnik]
- Irkı [irk]
- Cinsiyeti [cins]
- Kullandığı Silah [silah]
- Yakalandığı Yer [yer]
- Yakalanma Yılı [yıl]
- puan_cins [puan_ci...]
- puan_etnik [puan_e...]
- puan_yıl

Gallery

Choose from:

- Bar
- Line**
- Area
- Pie/Polar
- Scatter/Dot
- Histogram
- High-Low
- Boxplot
- Dual Axes

Element Properties

Edit Properties of: Line1

X-Axis1 (Line1)
Y-Axis1 (Line1)
GroupColor (Line1)

Statistics

Variable: Count

Display error bars

Error Bars Represent

Confidence intervals
Level (%): 95

Standard error
Multiplier: 2

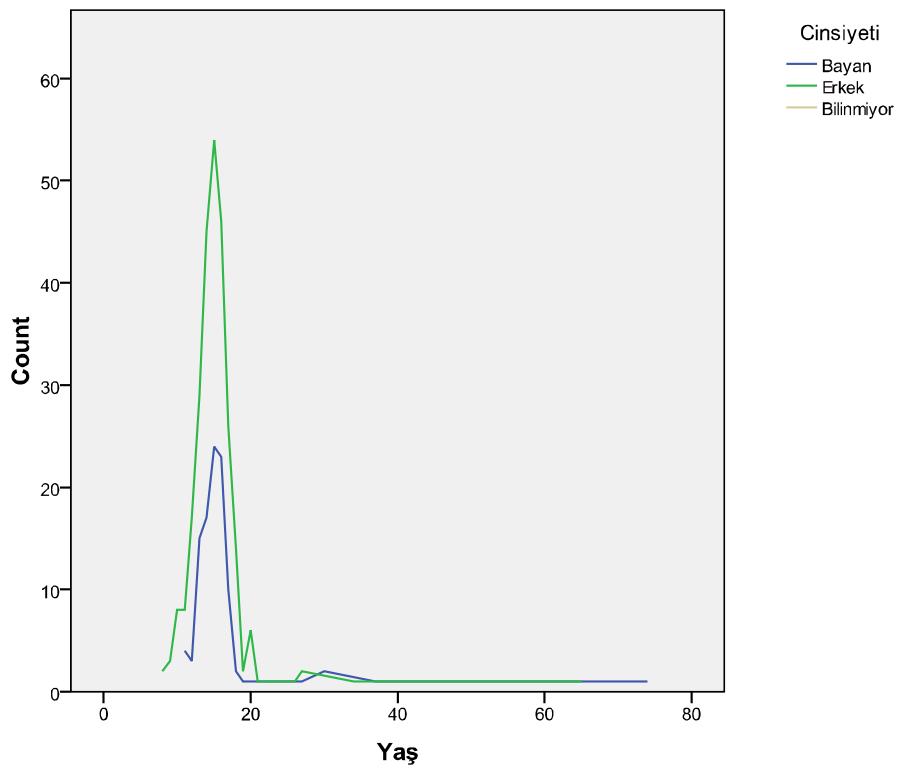
Standard deviation
Multiplier: 2

Interpolation

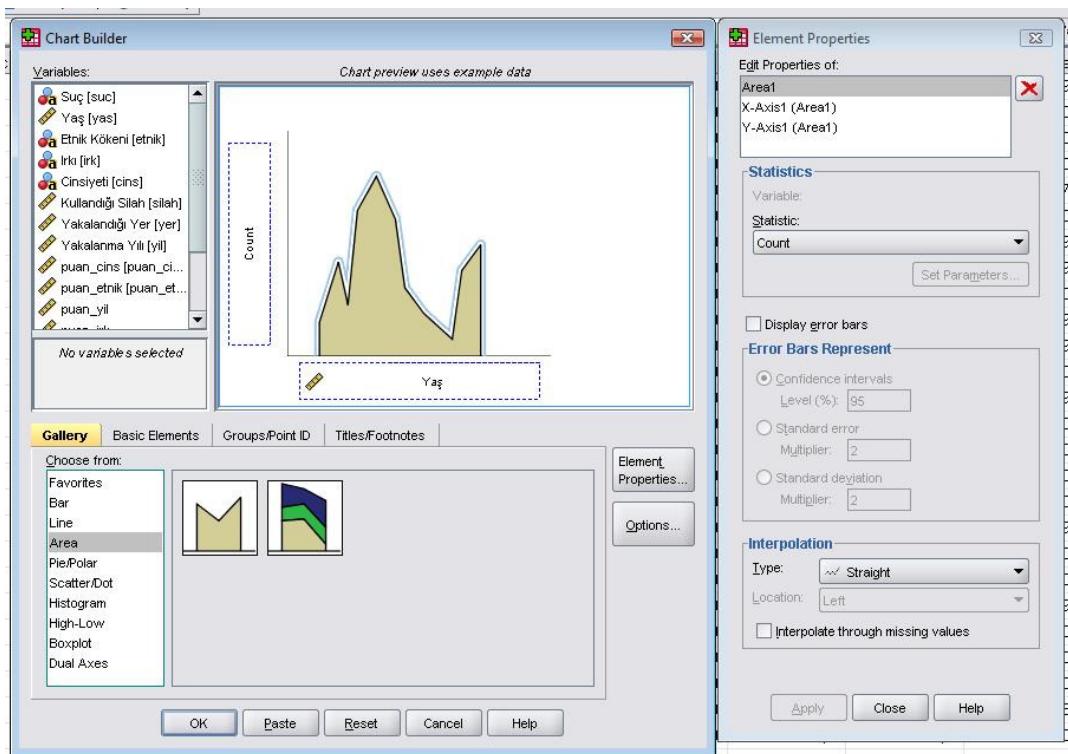
Type: Straight

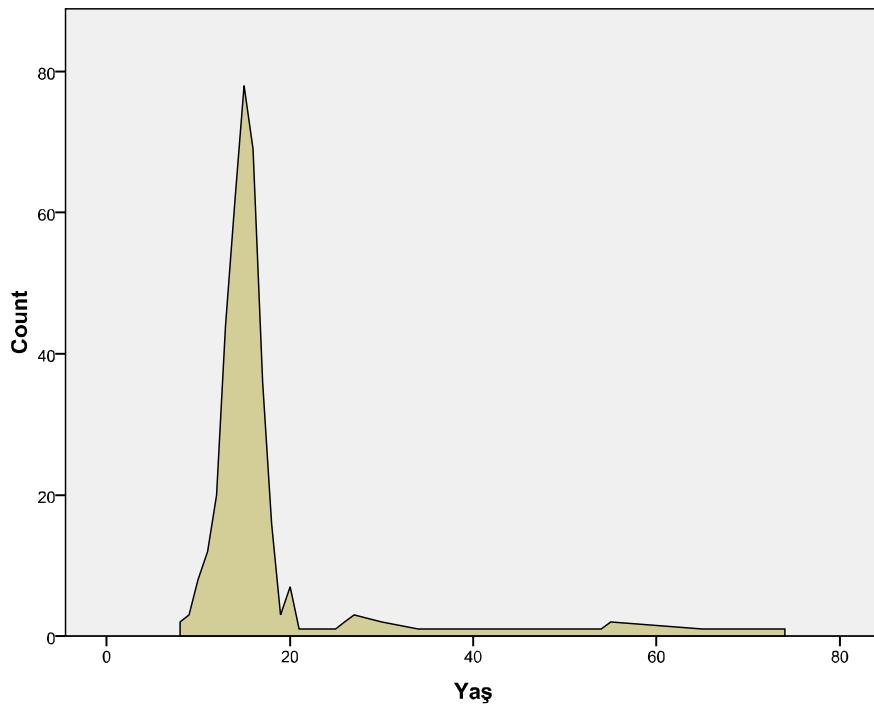
Location: Left

Interpolate through missing values



Yukarıdaki örnekte görüldüğü gibi bu tip bir line grafiğinde yaşın cinsiyete göre dağılımı verilmiştir. Başka bir grafik tipi de area (alan) grafiğidir. Yukarıdakilerle benzer istatistikler bu tip grafiklerle de gösterilebilir. Aşağıda bunlara örnek verelim.





Çıktıdan da anlaşıldığı gibi yaş dağılımı alan grafiğiyle verilmiştir. Birden fazla değişkenin bilgileri de bu grafikle verilebilir. Aşağıdaki örneği verelim.

Chart Builder

Variables:

- Suç [suc]
- Yaş [yas]
- Etnik Kökeni [etnik]
- İrk [irk]
- Cinsiyeti [cins]
- Kullandığı Silah [silah]
- Yakalandığı Yer [yer]
- Yakalanma Yılı [yili]
- puan_cins [puan_ci...]
- puan_etnik [puan_et...]
- puan_yil

Gallery Basic Elements Groups/Point ID Titles/Footnotes

Element Properties

Edit Properties of:

- Area1
- X-Axis1 (Area1)
- Y-Axis1 (Area1)
- GroupColor (Area1)

Statistics

Variable: Count

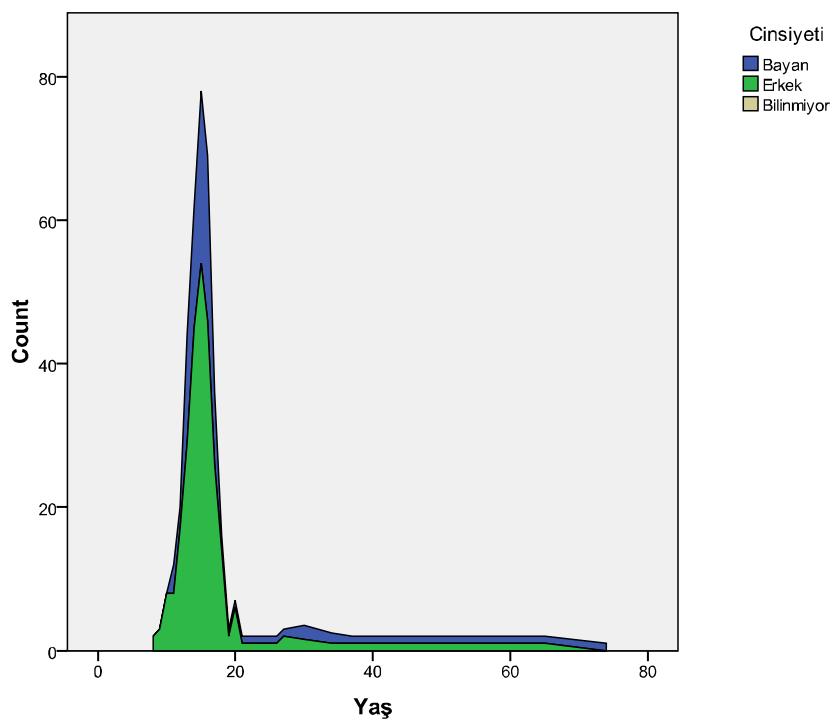
Error Bars Represent

- Confidence intervals Level (%): 95
- Standard error Multiplier: 2
- Standard deviation Multiplier: 2

Interpolation

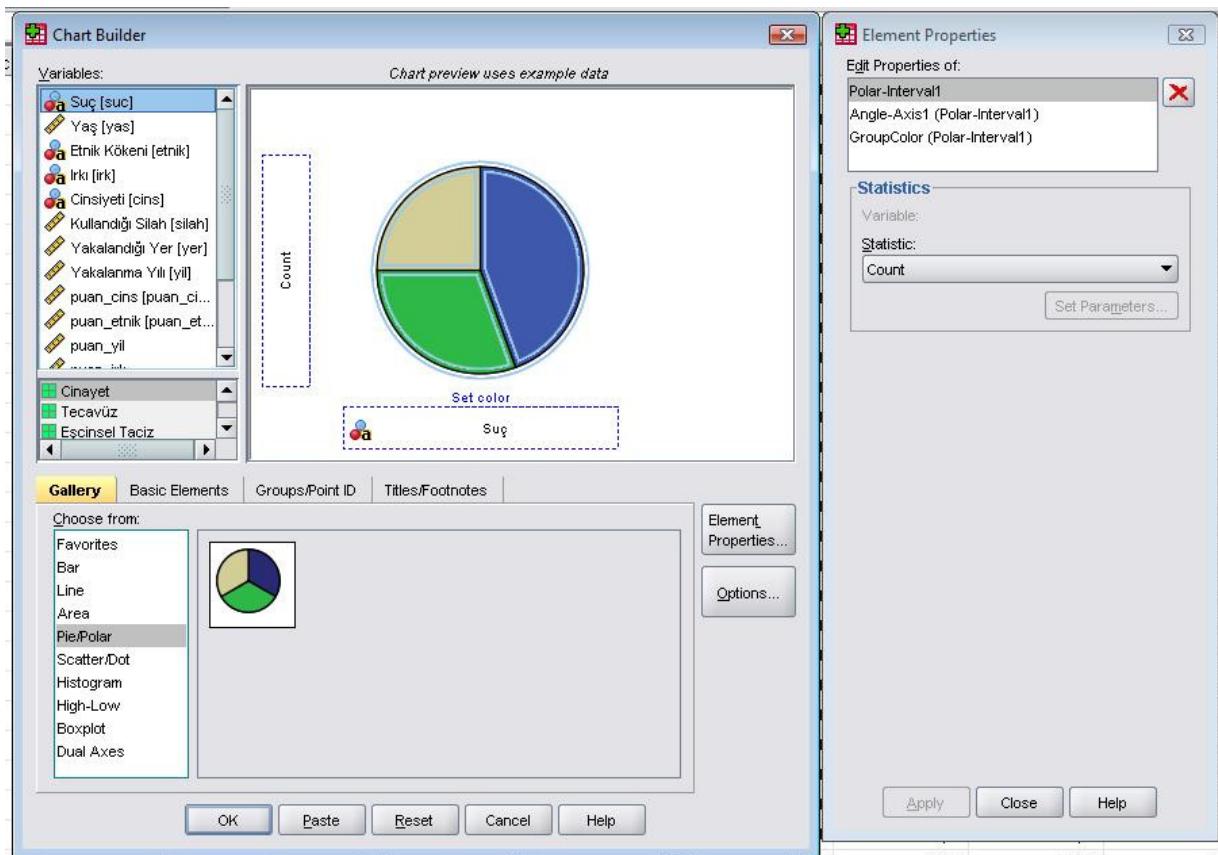
Type: Straight Location: Left

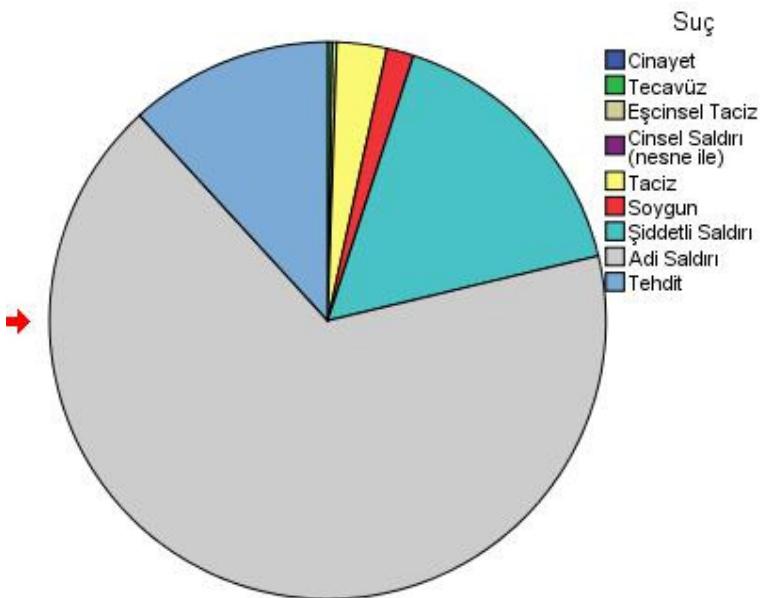
Buttons: OK, Paste, Reset, Cancel, Help, Apply, Close, Help



Yukarıdaki çıktıda yaşın cinsiyete göre dağılımı verildi.

Pie (pasta) grafiği ile de tek değişkene ait veriler gösterilebilir. Aşağıda bunu örnekledik.





Bir diğer önemli grafik tipi de nokta saçılım grafiğidir (dot/scatter). Bu tür grafik ile iki değişken arasındaki ilişki verilebilir. Aşağıdaki örneği inceleyelim.

Chart Builder

Variables:

- Suç [suc]
- Yaş [yas]
- Etnik Kökeni [etnik]
- İrkı [irk]
- Cinsiyeti [cins]
- Kullandığı Silah [silah]
- Yakalandığı Yer [yer]
- Yakalanma Yılı [yili]
- puan_cins [puan_ci...]
- puan_etnik [puan_et...]
- puan_yil

No categories (scale variable)

Chart preview uses example data

Scatter/Dot

Choose from:

- Favorites
- Bar
- Line
- Area
- Pie/Polar
- Scatter/Dot**
- Histogram
- High-Low
- Boxplot
- Dual Axes

Element Properties

Edit Properties of:
Point1
X-Axis1 (Point1)
Y-Axis1 (Point1)

Statistics

Variable: Yaş
Statistic:

Display error bars

Error Bars Represent

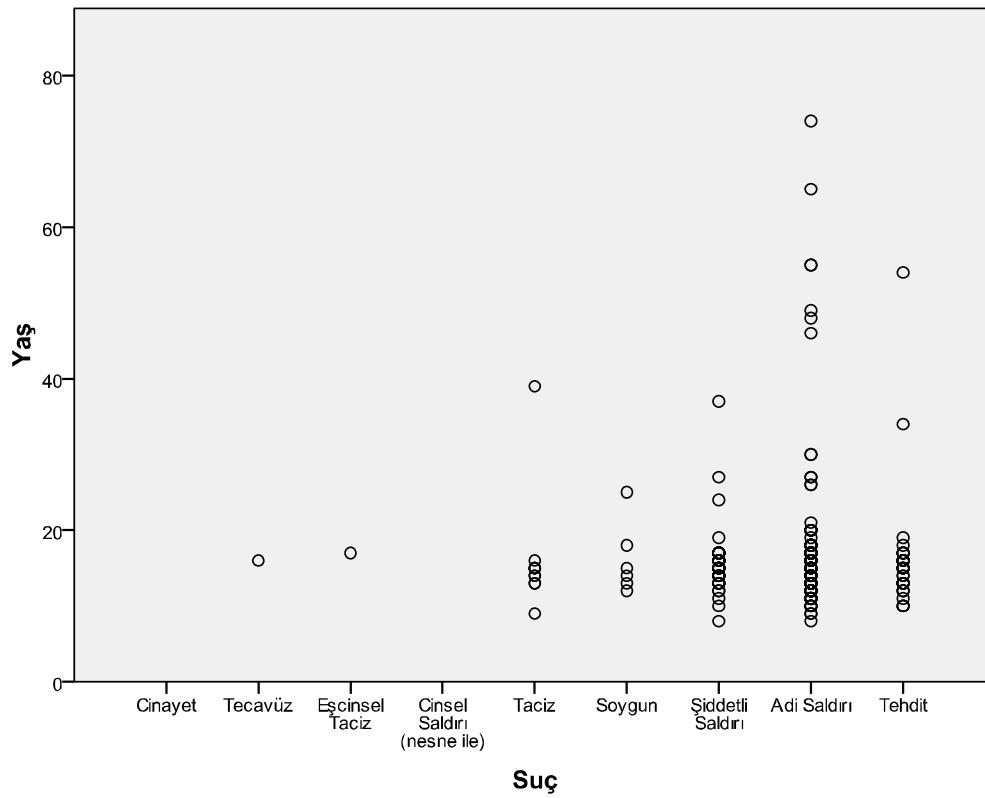
Confidence intervals
Level (%): 95

Standard error
Multiplier: 2

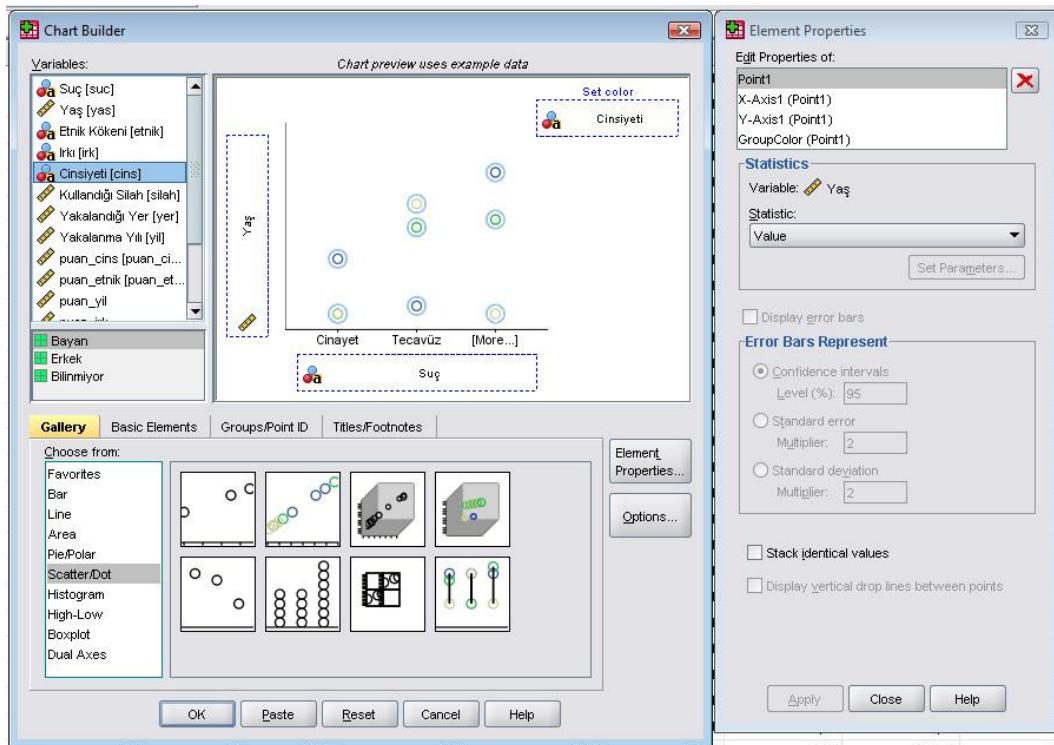
Standard deviation
Multiplier: 2

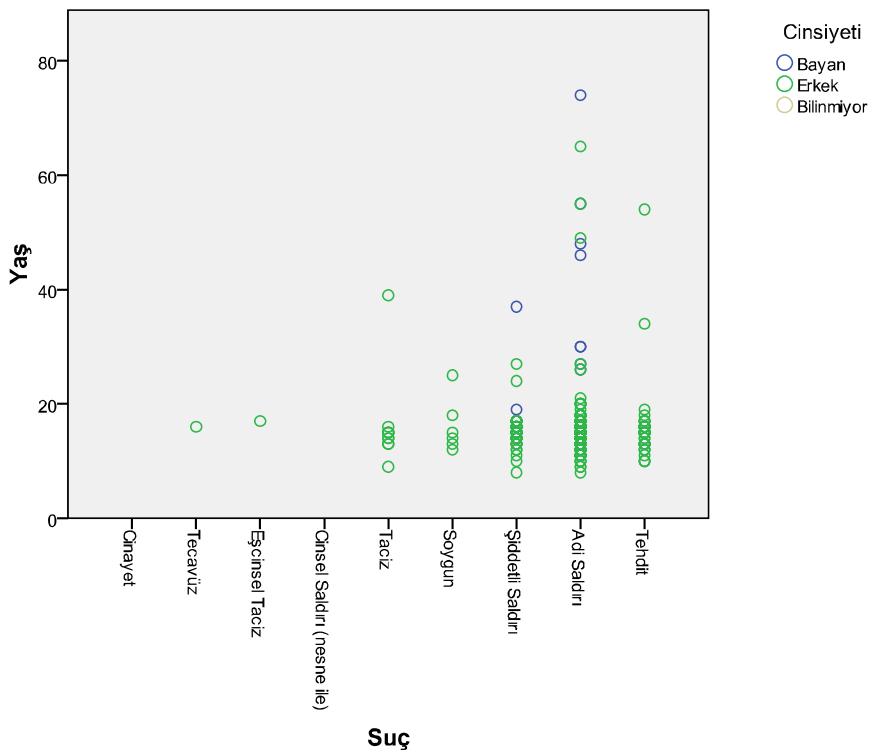
Stack identical values

Display vertical drop lines between points

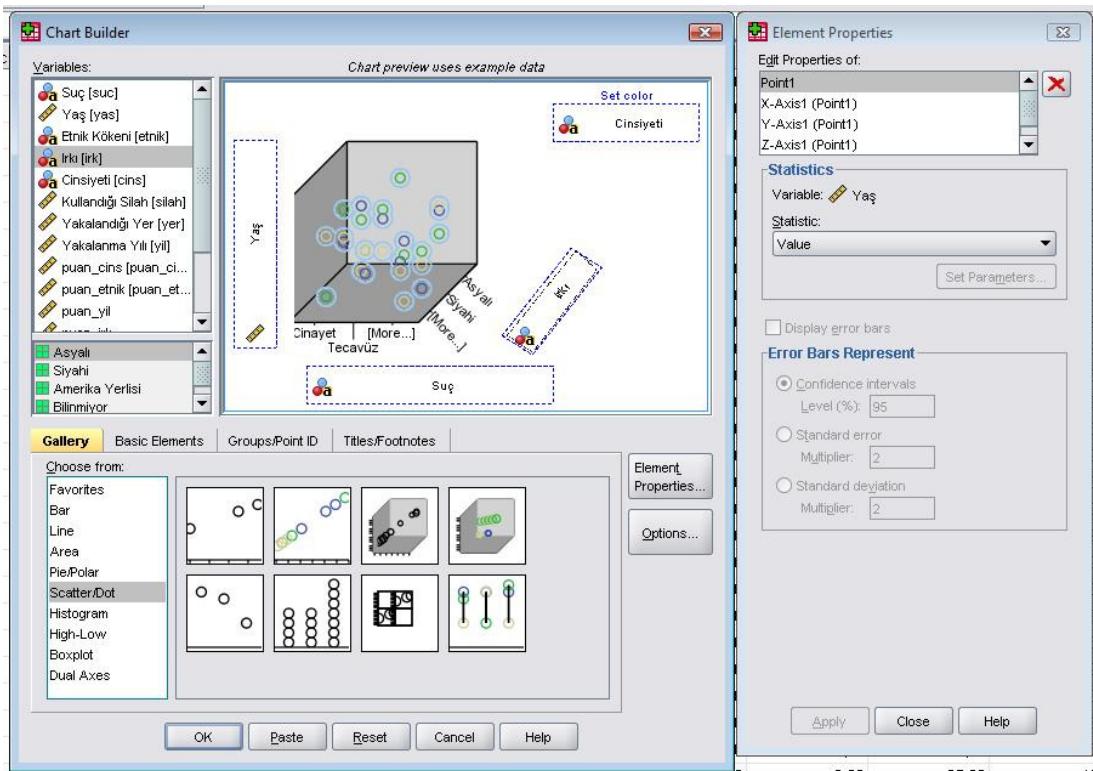


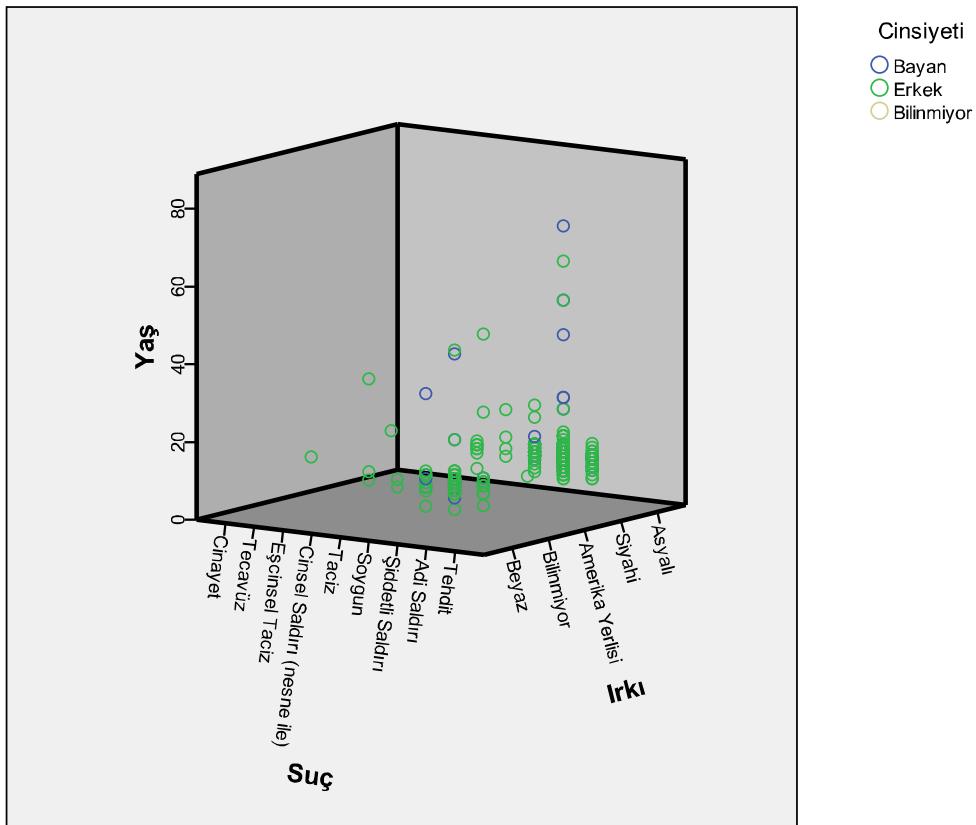
Çıktıdan da anlaşılacağı üzere suç çeşitleri ile bunların işlendiği yaşlar arasındaki ilişki verilmiştir. Daha fazla değişkenler için de bu grafiğin çeşitleri vardır. Buna da bir örnek verelim.





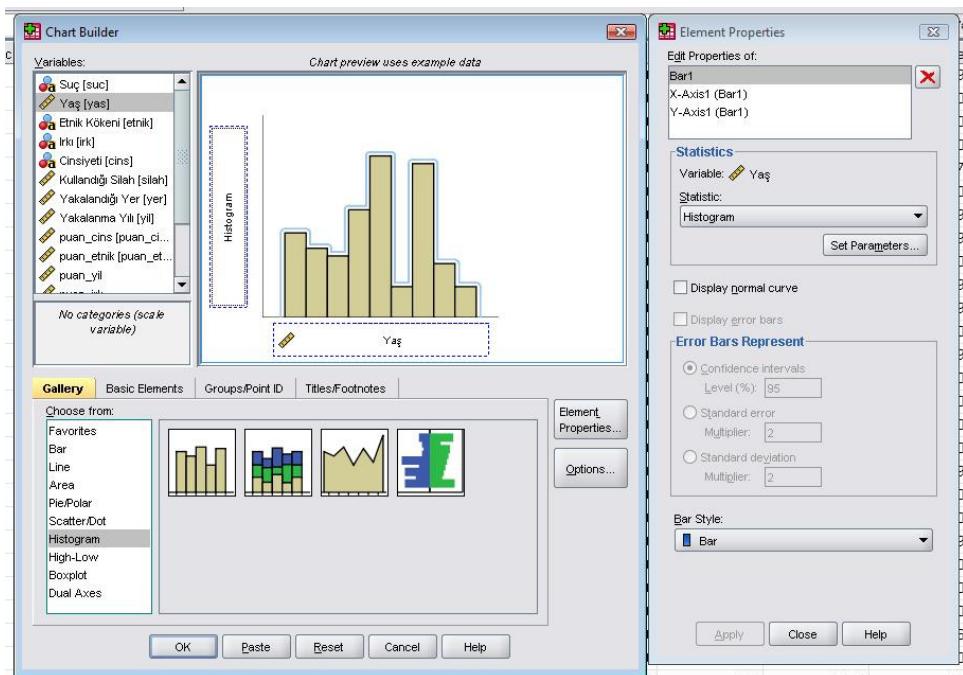
Bu çıktıda bir öncekinden farklı olarak suçları işleyenlerin cinsiyetlerinin yaşa göre dağılım bilgisi de verilmiştir. Bu tip bir grafikle daha fazla değişkenin dağılımı da bir arada gösterilebilir. Buna örnek olarak aşağıdaki şemayı verelim.

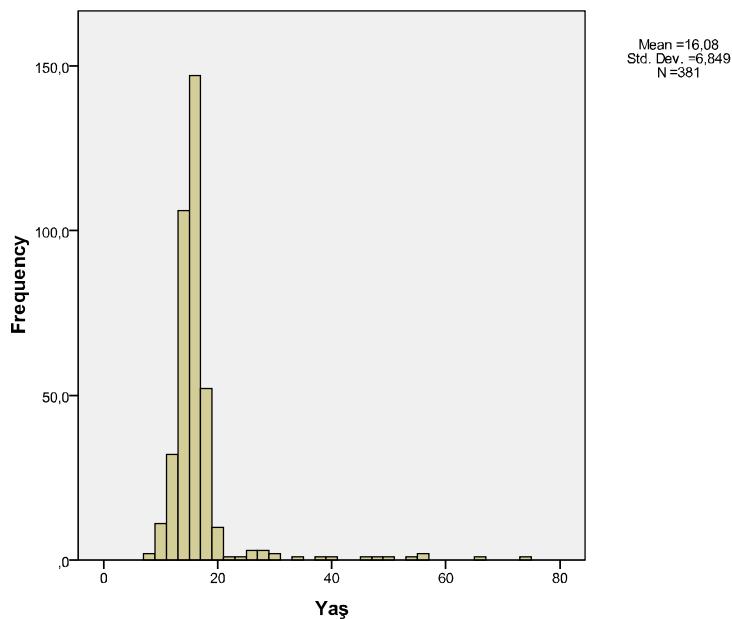




Göründüğü gibi suçu işleyen cinsiyeti ve yaşıının dışında ırkının da yaşlara göre dağılımı verilmiştir.

Bar tipine benzeyen bir grafik tipi olarak histogram tipi grafik de vardır. Bunun çeşitlerini örneklerle gösterelim.





Bu tipin diğer çeşitlerine d değinelim.

Chart Builder

Variables:

- Suç [suc]
- Yaş [yas]
- Etnik Kökeni [etnik]
- İrk [irk]
- Cinsiyeti [cins]
- Kullandığı Silah [silah]
- Yakalandığı Yer [yer]
- Yakalananı Yılı [yili]
- puan_cins [puan_ci...]
- puan_etnik [puan_et...]
- puan_yil

Gallery

Element Properties

Edit Properties of:

- Bar1
- X-Axis1 (Bar1)
- Y-Axis1 (Bar1)
- GroupColor (Bar1)

Statistics

Variable: Yaş
Statistic: Histogram

Display normal curve
 Display error bars

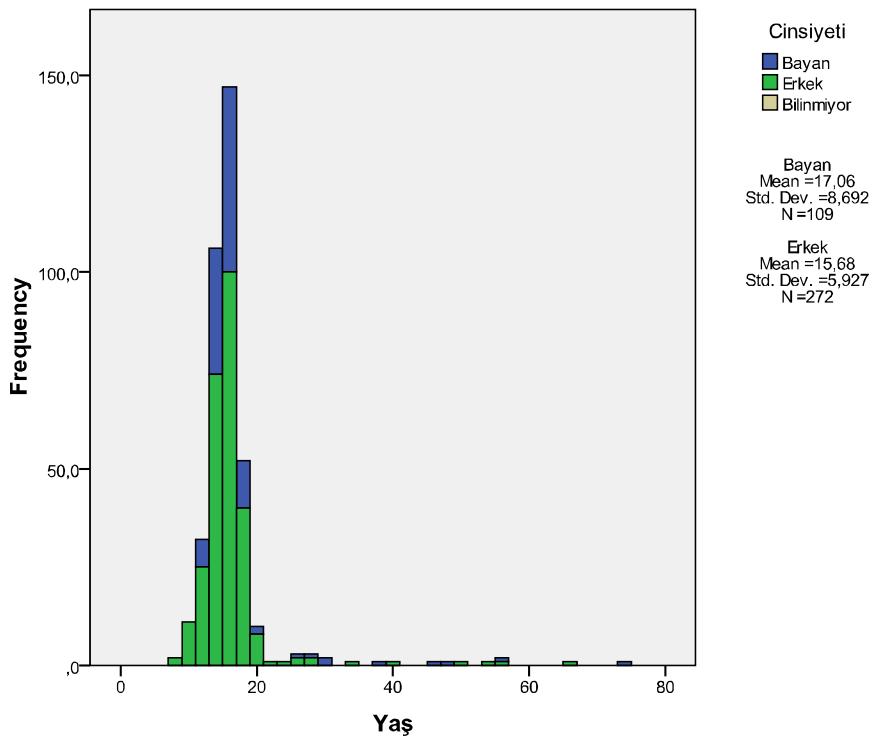
Error Bars Represent

(Confidence intervals
Level (%): 95)
(Standard error
Multiplier: 2)
(Standard deviation
Multiplier: 2)

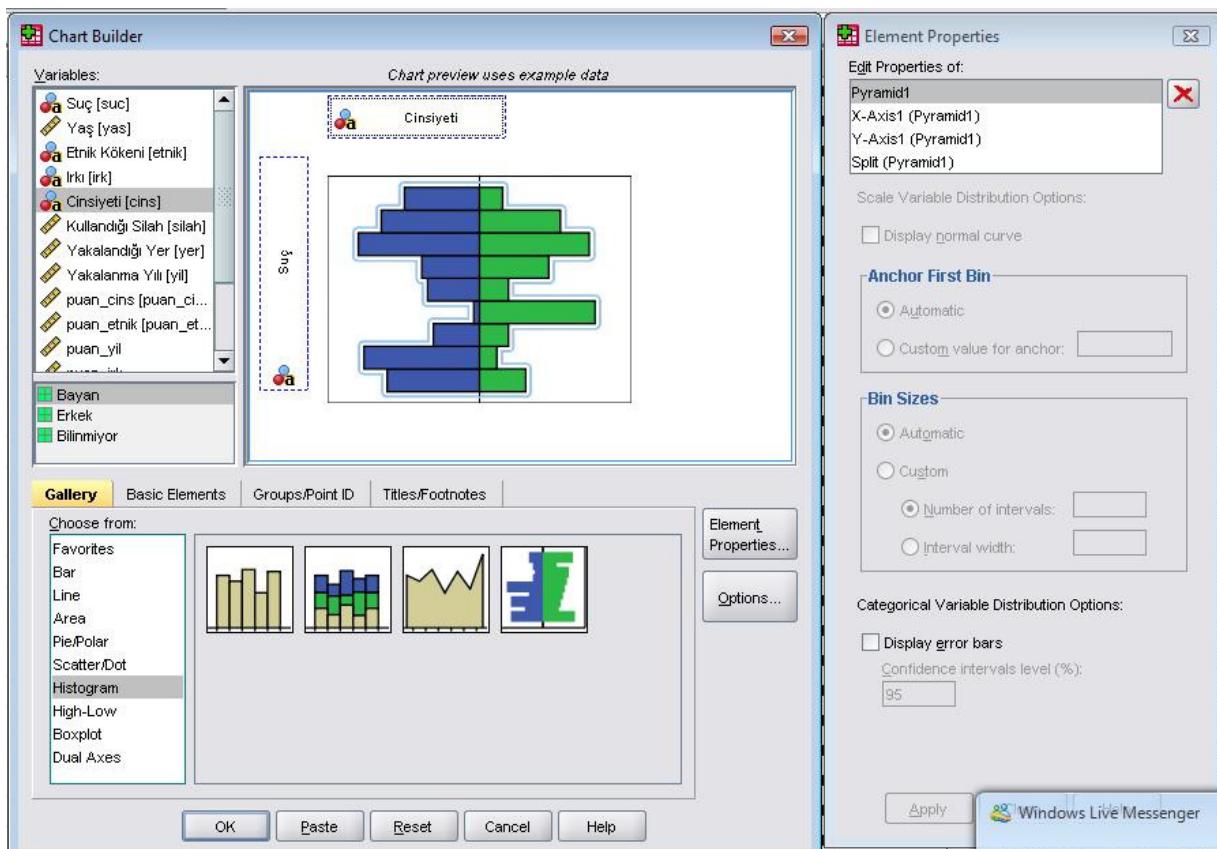
Bar Style:

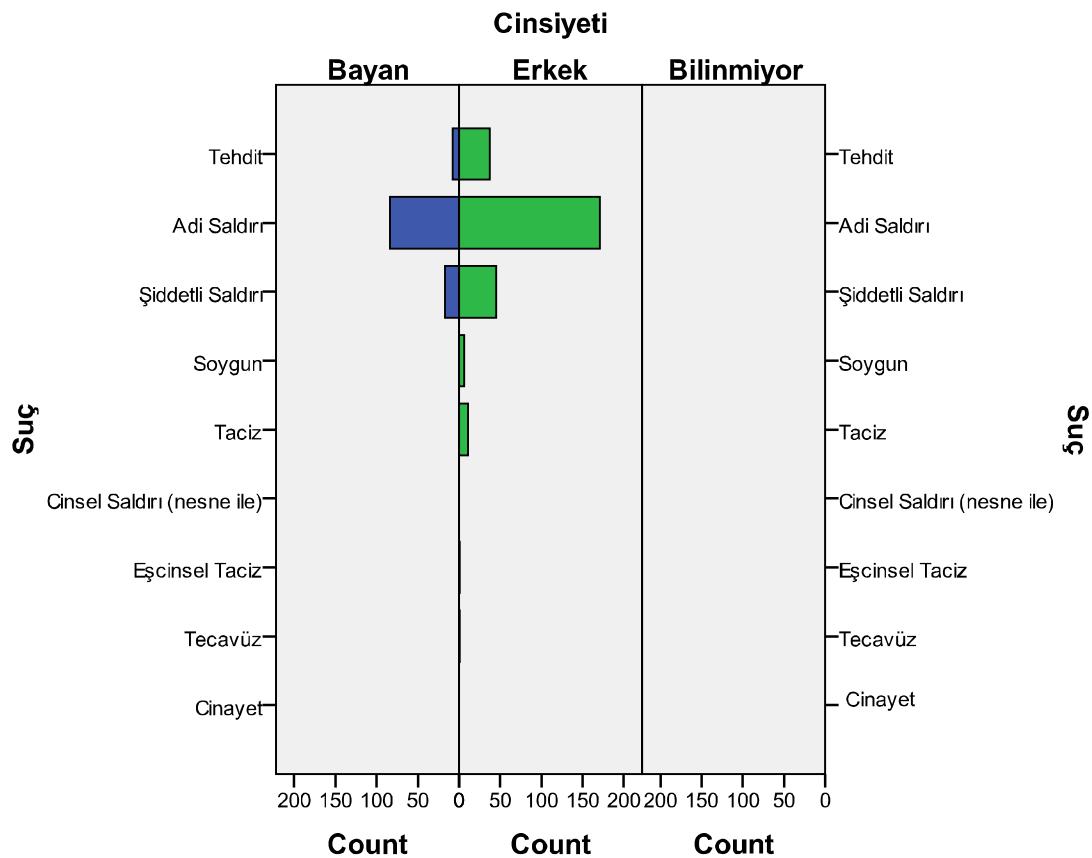
Bar

OK Paste Reset Cancel Help Apply Close Help



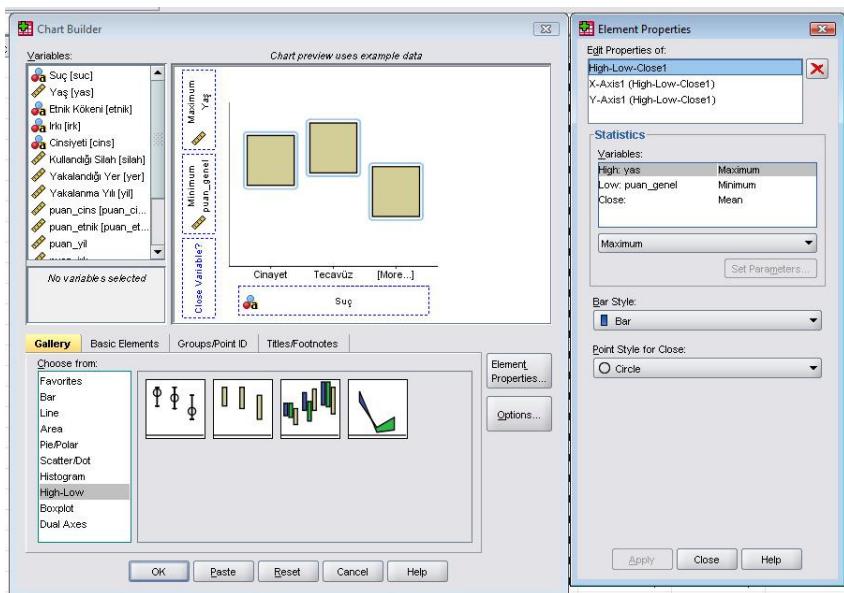
Çıktıda görüldüğü gibi yaşların cinsiyete göre dağılımı da grafikte verilmektedir. Bir diğer tip de aşağıda örneklendirilmiştir.

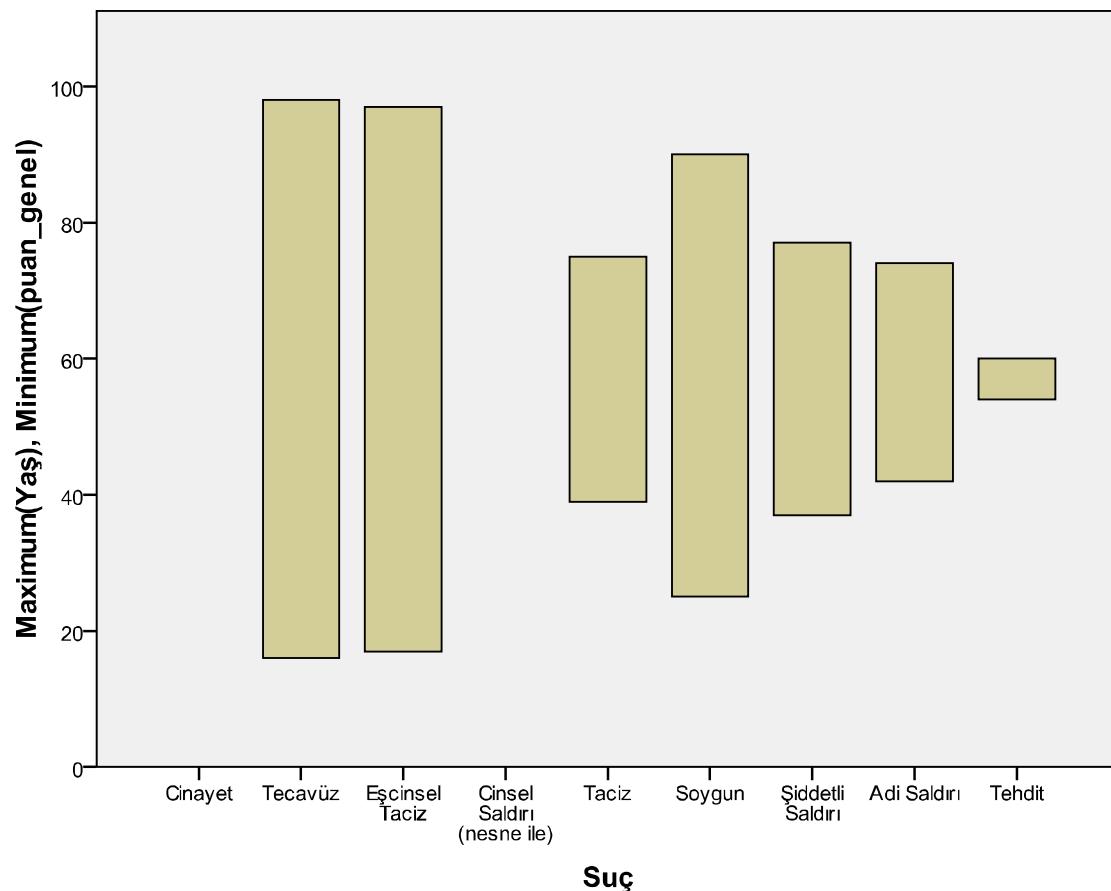




Göründüğü gibi grafikte suçların cinsiyete göre dağılımı farklı bir şekilde verilmektedir.

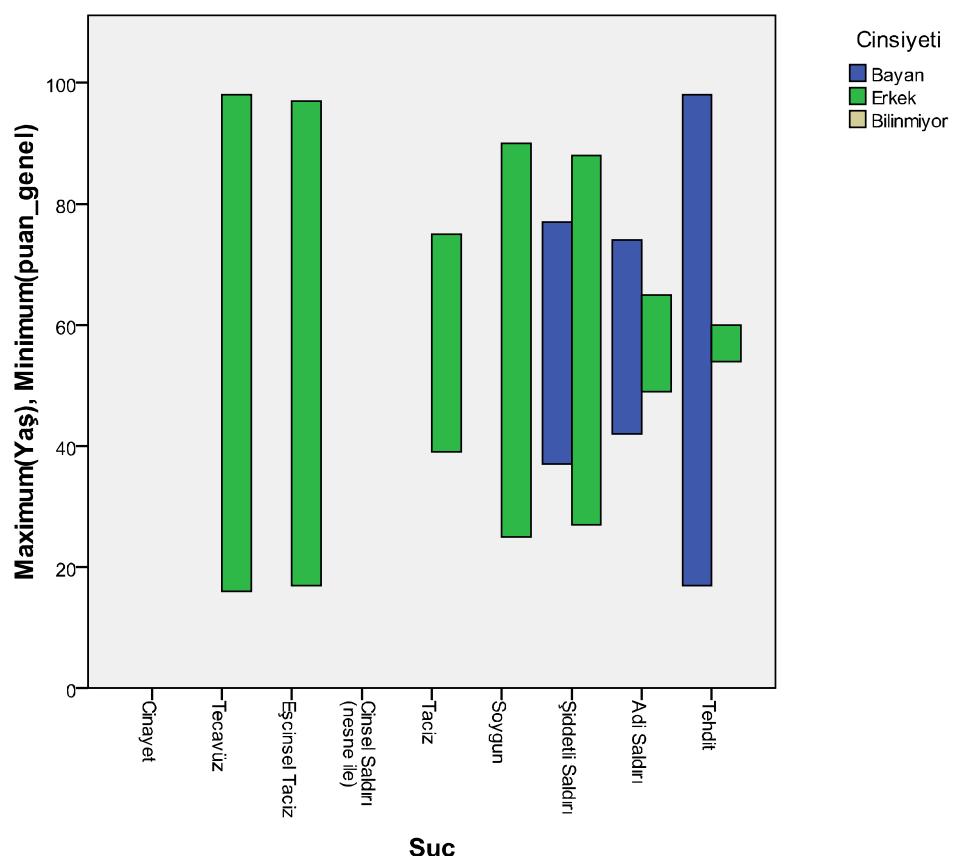
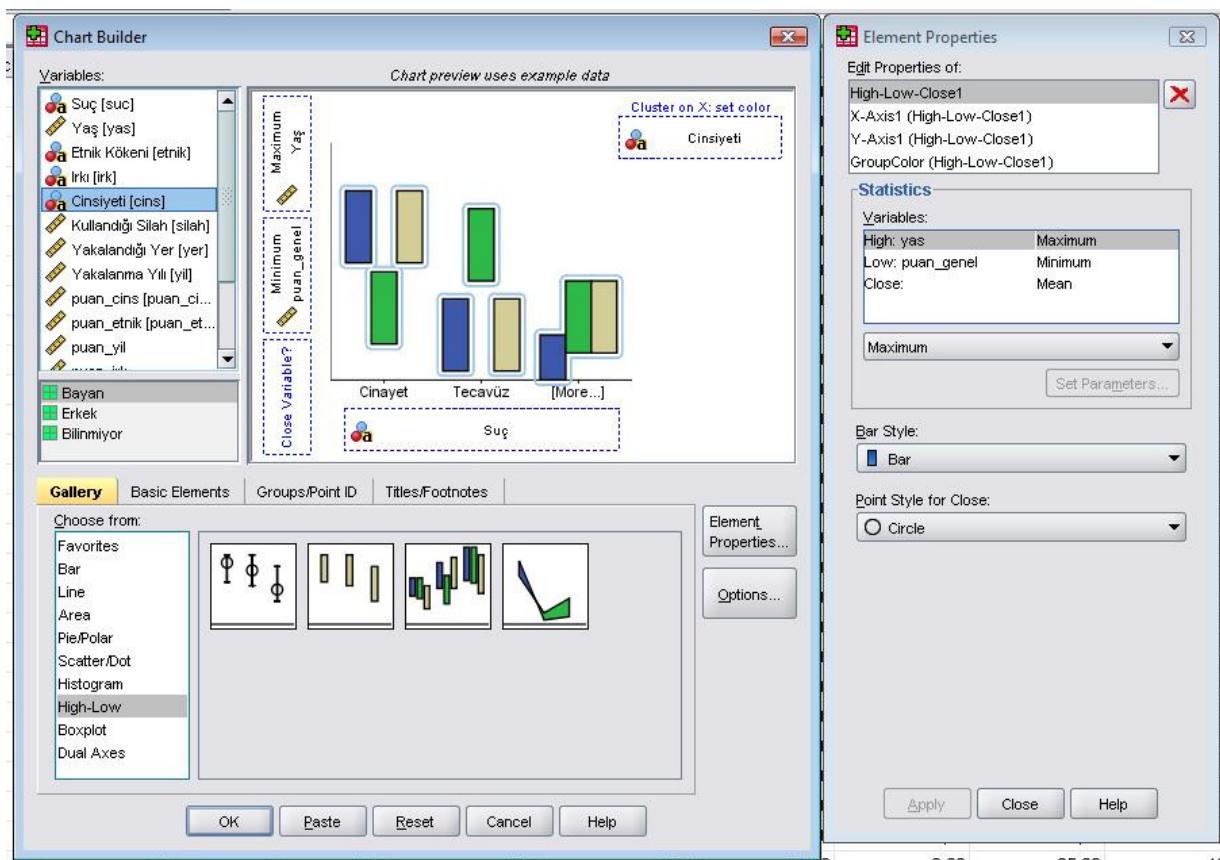
High-low (max-min) grafiği ile de değişkenlerin maximum ve minimum değerleri verilebilir. Bunu aşağıdaki örnek ile açıkladık.





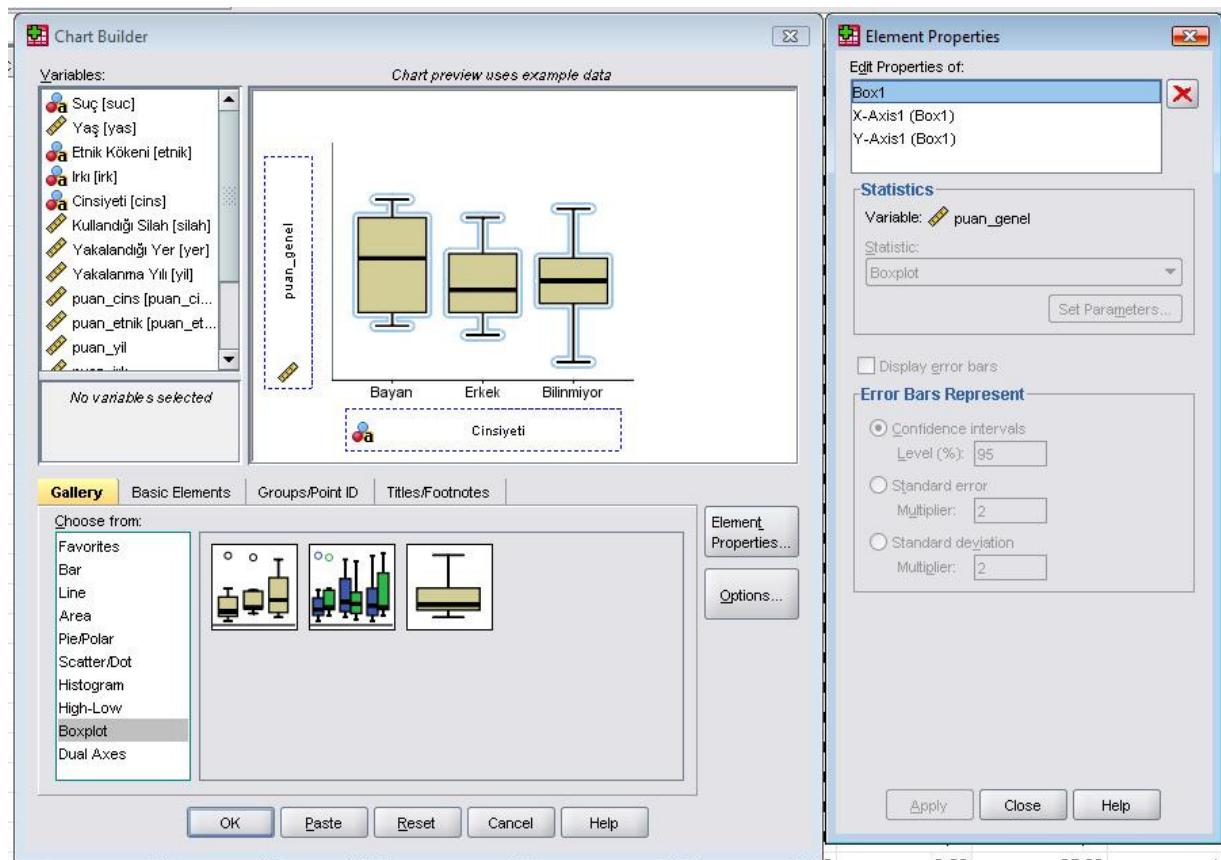
Çıktıdan da anlaşılacağı gibi grafikte yaş değişkenine ait verilerin minimumu ile genel puan değişkeninin maksimum değerlerinin suçlara göre dağılımı verilmiştir. Örneğin, taciz suçunu işleyen deneklerin maksimum yaşı 75 civarında iken bu deneklerin minimum genel puanı 40 civarındadır.

Bu grafiği daha fazla değişken ile de yaratabiliriz. Bunu da aşağıdaki örnek ile inceleyelim.

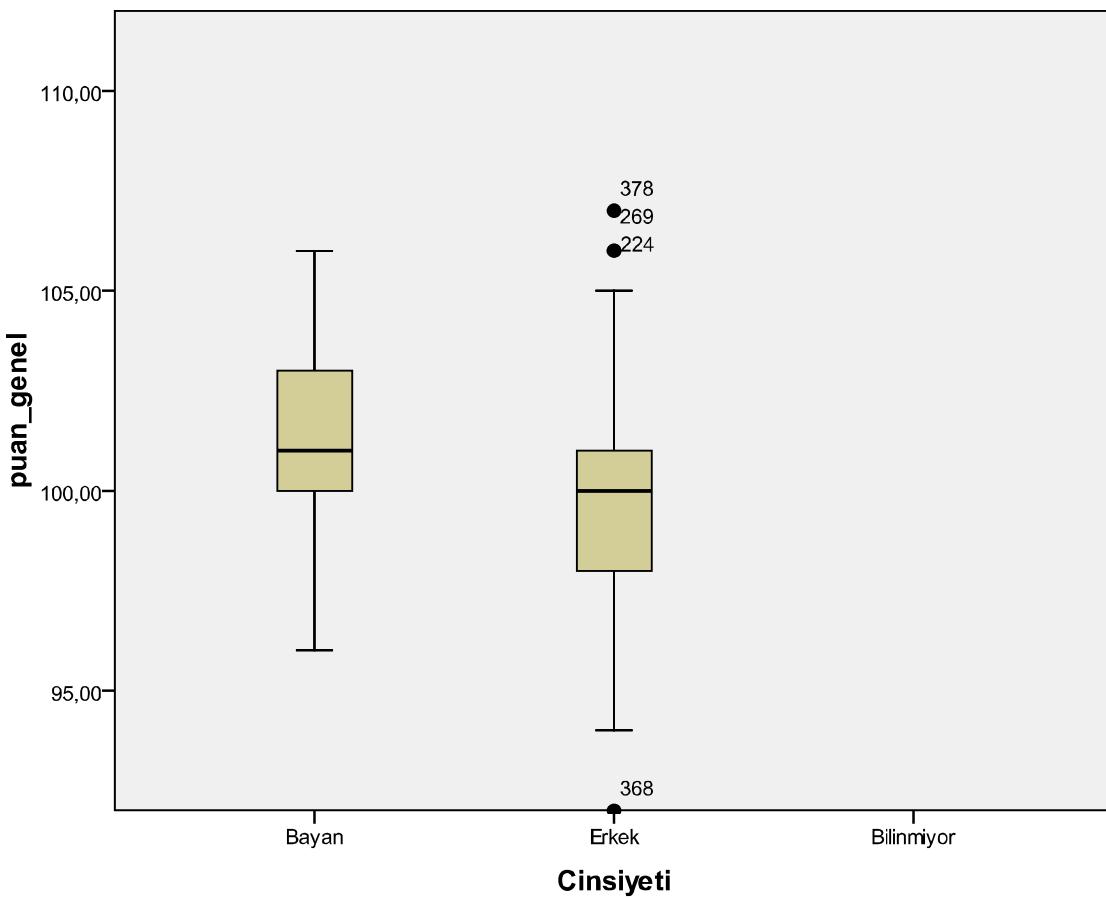


Çıktıdan da anlaşıldığı gibi, bir önceki örnekte verilen değişkenlerin max ve min değerleri sadece suç çeşidine göre değil, suç işleyen deneklerin cinsiyetine göre de verilmektedir. Örneğin şiddetli saldırı suçunu işleyen bireylerden, bayan olanların maximum yaşı 38 civarında iken, minimum genel puanı 76 civarındadır. Benzer şekilde erkeklerin de maximum yaşı 23 civarı iken minimum genel puanı 93 civarındadır.

Şimdi boxplot dediğimiz grafik türünü inceleyelim. Bu grafikle verilere ait istatistikler daha kapsamlı olarak verilebilir. Bir örnek üzerinde inceleyelim.



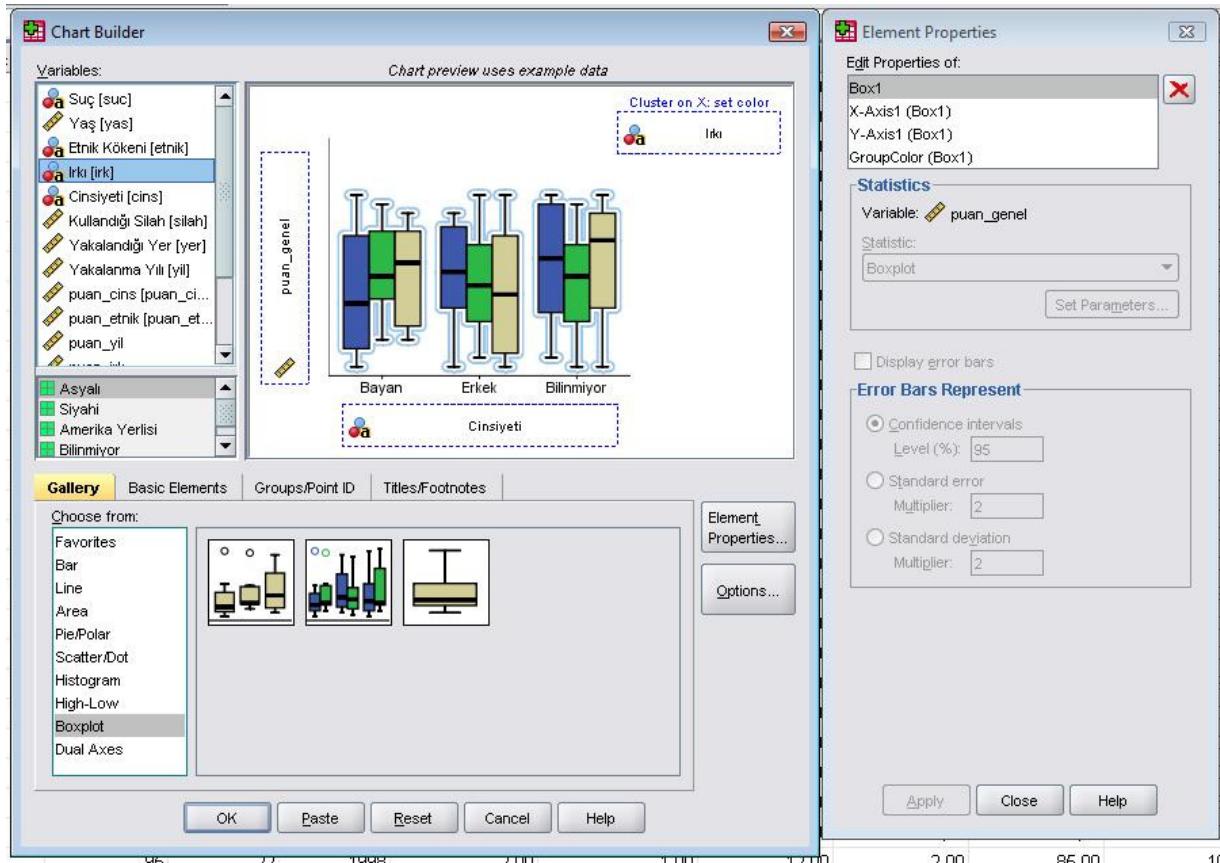
Yukarıdaki komutları verdigimiz taktirde aşağıdaki çıktı elde edilir.



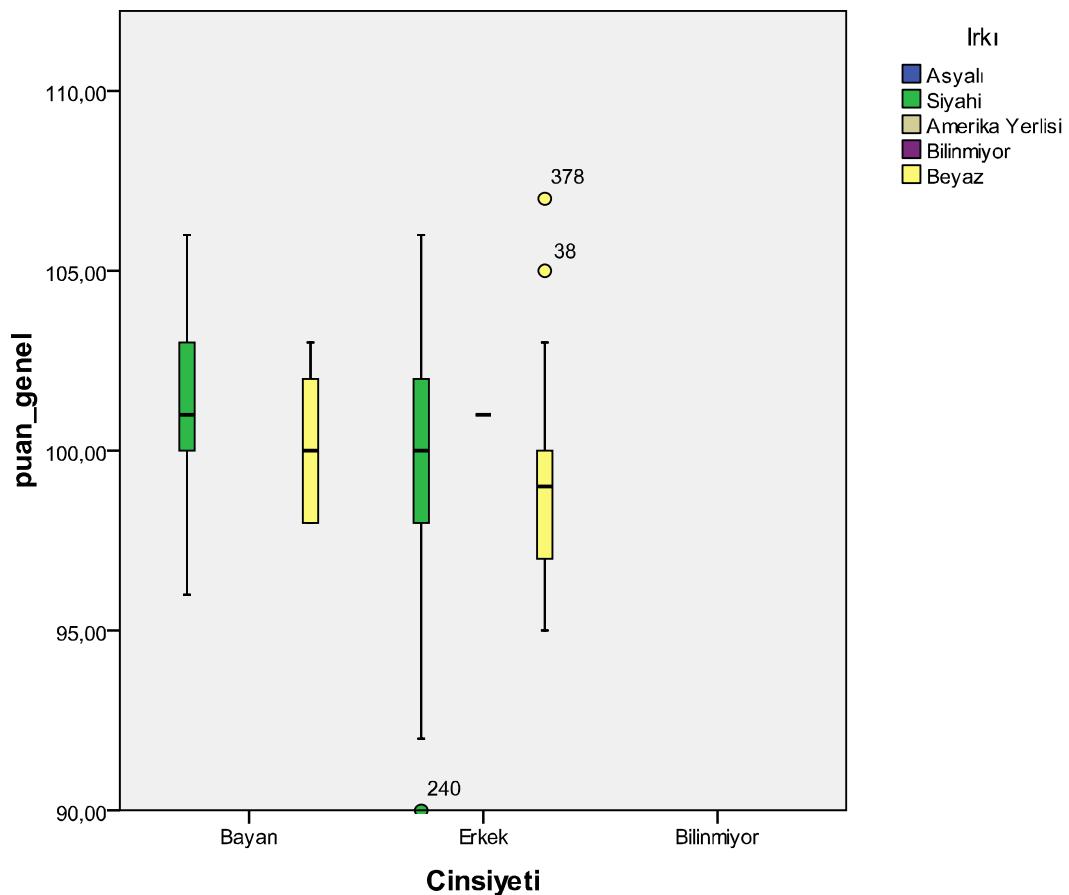
Yukarıdaki çıktıda deneklerin cinsiyetlerinin genel puanına göre dağılımı verilmiştir. Bu grafiği biraz yorumlayalım. Grafik bir kutucuk (box), uçlarında çizgiler (whiskers), kutucuğun ortasında bir çizgi ve alt whiskerslerin altında ve üstünde bazı noktalardan oluşmaktadır. Boxplot grafiğinde whisker in alt ve üst uçlarındaki sınırlar değişkenin maksimum ve minimum değerlerini yani rage ini belirtir. Buna göre erkek deneklerin puanların dağılımı grafikte 92 ile 105 arasında (ortalama) değiştiği verilmiştir. Ortadaki kutunun uzunluğu ise interquartile range (IQR) dediğimiz çeyreklikler arası farkı (3. Ve 1. Çeyreklik arası) verir. Datanın 4 çeyrekten oluştuğu düşünülürse bu farkın tüm deneklerin ortadaki %50 lik diliminin dağılımini verdiği söylenir. Buna göre bayanların puanlarının interquartile range i 100 ile 103 (ortalama) arasında olduğu söylenebilir. Grafikte kutuların ortalarındaki yatay çizgiler ise verilerin ortalama değerini verir. Buna göre erkeklerin puanlarının ortalama değeri 100 civarındadır. Grafikteki noktalar ise outliers dediğimiz aşan değerleri verir. Bu değerler grubu temsil etmeyen değerlerdir. Yanlış girilmiş olabilir. Grafik bizi bu konuda uyarır. Örneğin erkeklerden 368, 224, 269 ve 378 numaralı deneklerin puanları grubun dağılımindan oldukça değişik olup grubu temsil etmemektedirler. Bu grafikte yok fakat boxplot grafiklerinde * ile de bazı denekler belirtilebilir. Bu belirtilen deneklerin değerleri extremum değerlerdir. Yani alttan ve üstten en uç değerlerdir. Bu grafik ayrıca bize dağılımin grafiği hakkında da bilgi verir. Boxları sola doğru yatırıp median çizgisini de tepe noktası gibi düşünüp

çan eğrisi oluşturulabilir. Buna göre bayanların puan dağılımı sağa çarpık, erkeklerin puan dağılımı ise sola çarpıktır.

Daha önceki grafik çeşitlerinde olduğu gibi bu grafikte de birden fazla değişkenin dağılımı sunulabilir. Aşağıdaki örnekte bu durumu gösterelim.



Bu komutları verince aşağıdaki çıktıyı alırız.



Göründüğü gibi grafikte sadece cinsiyete göre puan dağılımı değil aynı zamanda ırklara göre de puan dağılımı verilmiştir. Bir önceki boxplot grafiği örneğinde yaptığımız yorumlar gibi bu grafik için de yorum yapabiliriz.