



HACETTEPE ÜNİVERSİTESİ
İSTATİSTİK BÖLÜMÜ

İST155 İSTATİSTİĞE GİRİŞ I

UYGULAMA 4

Ar. Gör. Dr. Derya Turfan – Ar. Gör. Dr. Murat Arat

1) 25 kişinin tuttuğu takımlar aşağıda verilmiştir. Verilerin sıklık çizelgesini oluşturunuz ve uygun grafiği çiziniz.

Kişi	Takım	Kişi	Takım	Kişi	Takım	Kişi	Takım	Kişi	Takım
1	GS	6	GS	11	FB	16	GS	21	GS
2	FB	7	FB	12	GS	17	GS	22	BJK
3	BJK	8	BJK	13	BJK	18	GS	23	FB
4	BJK	9	GS	14	TS	19	FB	24	GS
5	GS	10	GS	15	BJK	20	TS	25	GS

İlk olarak, *Variable View* penceresinde “takım” değişkeni ve ağırlıklandırma işlemi için ise “frekans” değişkeni tanımlanır.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	takım	Numeric	8	0		{1, bjk}...	None	8	Right	Nominal	Input
2	frekans	Numeric	8	0		None	None	8	Right	Scale	Input

Takım değişkenine ait etiketleme işlemi (1→bjk, 2→fb, 3→gs, 4→ts) yapılır. *Data View* penceresinde bu değişkenlere ait gözlem değerleri girilir.

Value Labels

Value:

Label:

Add Change Remove

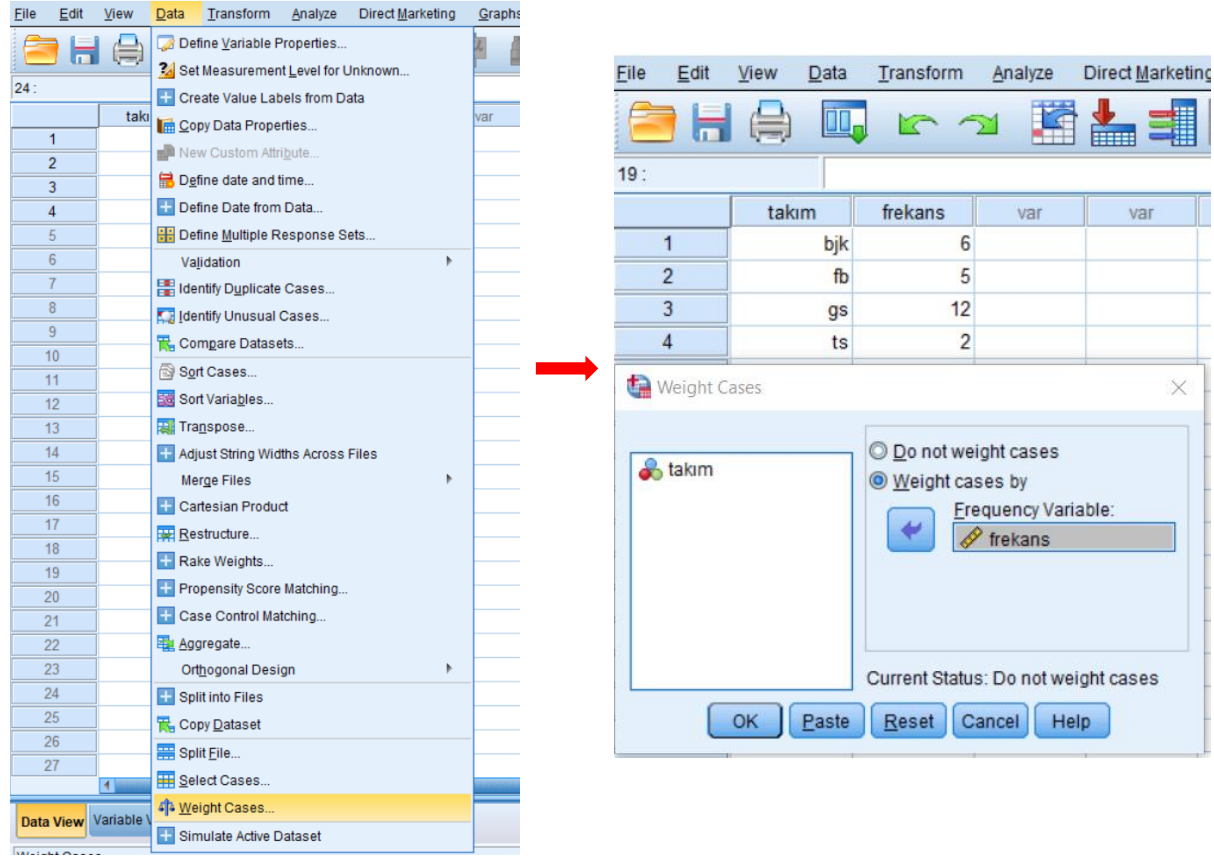
1 = "bjk"
2 = "fb"
3 = "gs"
4 = "ts"

OK Cancel Help



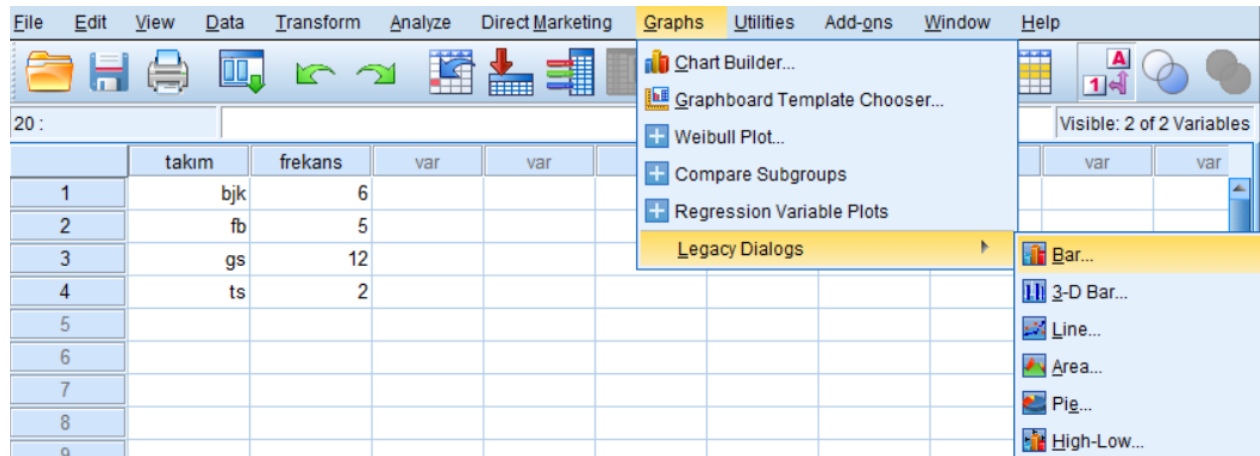
	takım	frekans
1	bjk	6
2	fb	5
3	gs	12
4	ts	2

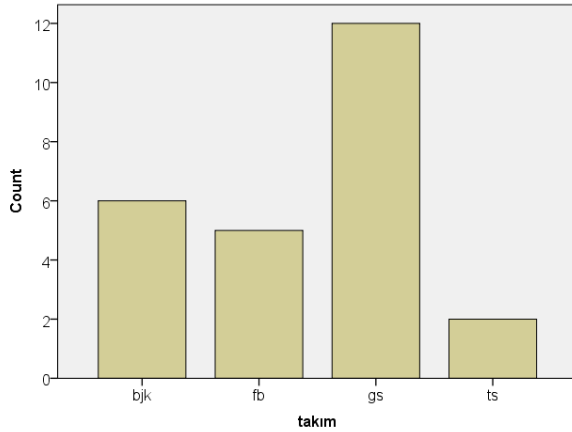
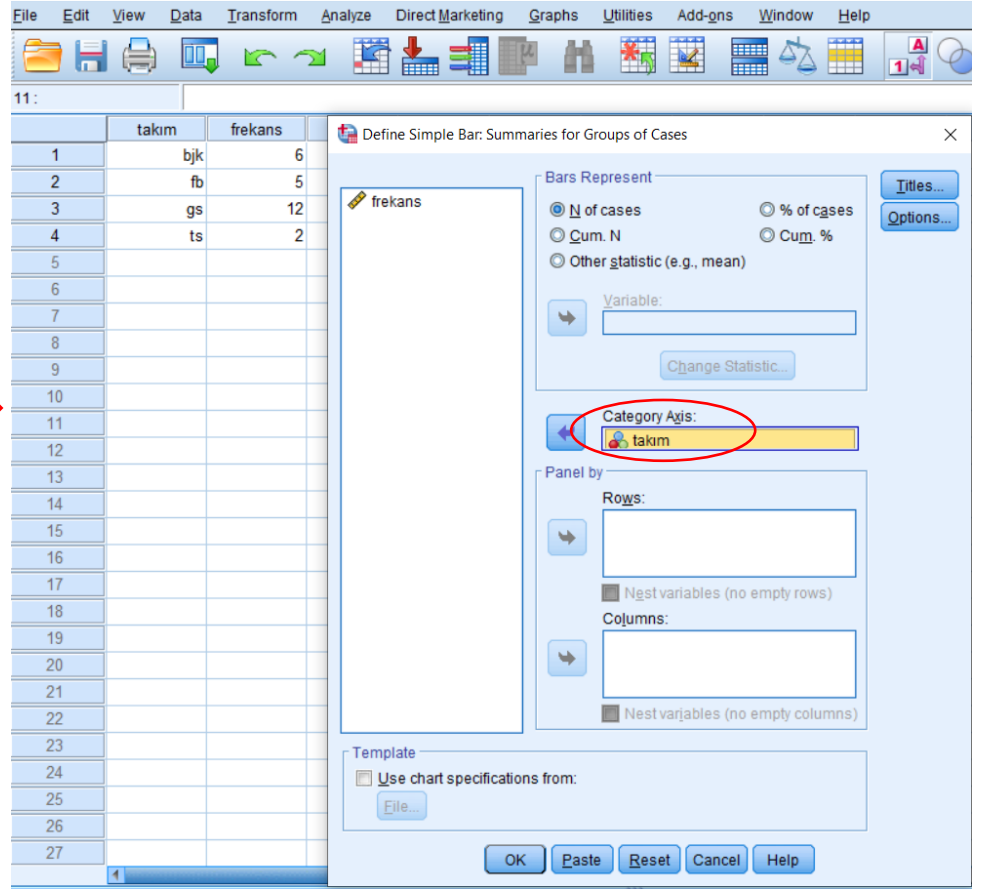
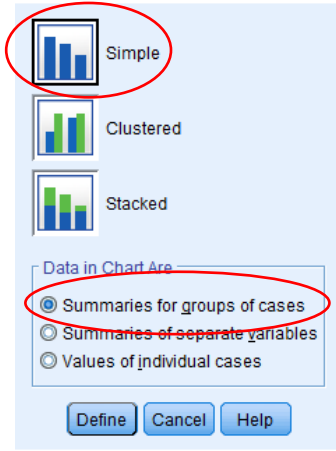
Veri girişinden sonra tanımlanan “frekans” değişkeni üzerinden ağırlıklandırma işlemi yapılır.



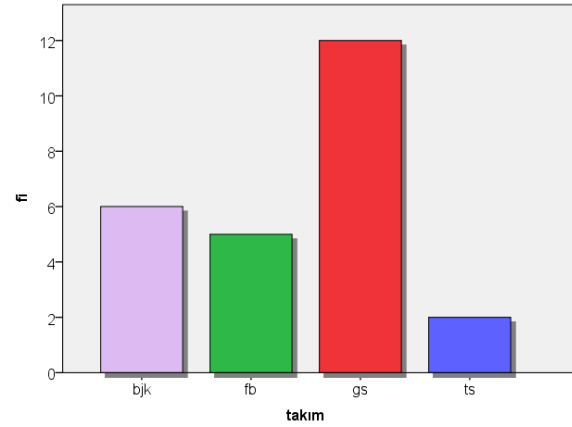
Yapılan ağırlıklandırma işlemi kontrol edildikten sonra değişkene ait uygun olan grafikler belirlenir. “takım” değişkeni sınıflanabilir nitel veri türüne sahip olduğu için çizilebilecek uygun grafikler çubuk ve pasta grafikleridir.

Çubuk (Bar) Grafiği Çizimi



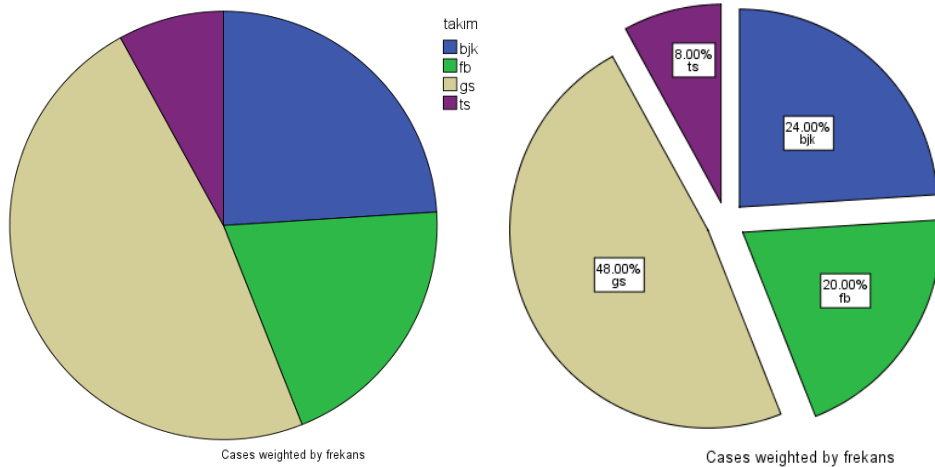
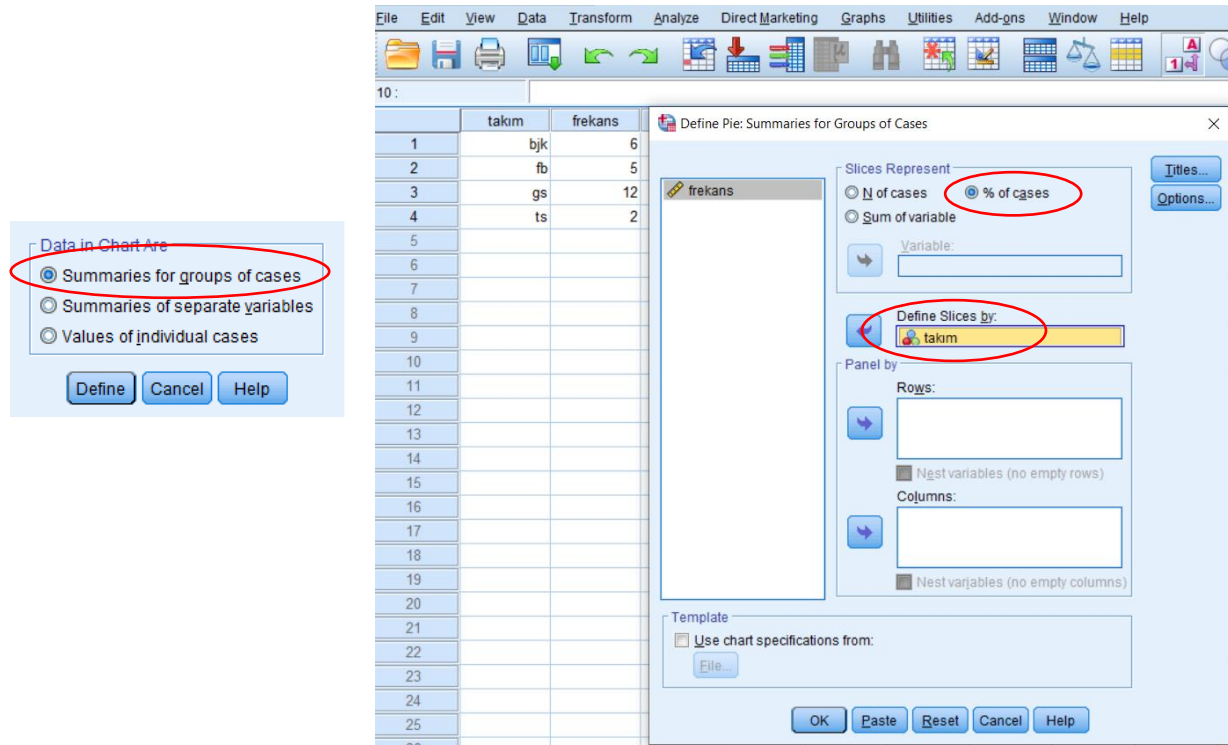
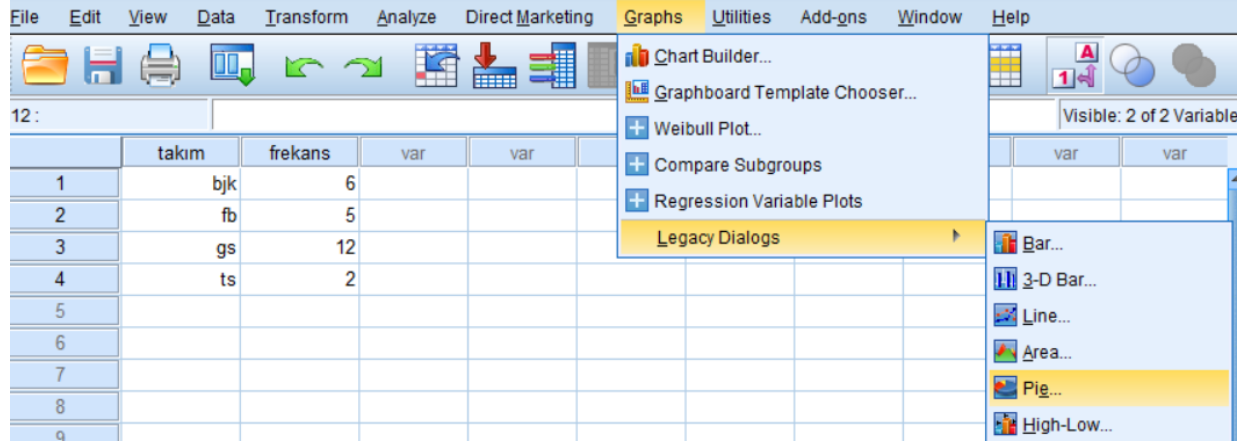


Cases weighted by frekans



Cases weighted by frekans

Pasta (Pie) Grafiği Çizimi:



2) H.Ü. Fen Fakültesi İstatistik Bölümü kız ve erkek öğrencilerinin ileri matematik dersinden aldıkları notlar aşağıda verilmiştir. Verilere uygun grafiği çiziniz.

Kız	10 12 14 48 30 40 34 72 72 20 28 2 44 62 68 40 84 56 74 74 30 50 84 50 74 32 82 72 58 100 58
Erkek	2 2 2 50 2 52 16 46 10 10 12 24 38 54 26 58 58 42 40 24 40 32 52 70 92 10 100 40 20 66 68 74

Variable View penceresinde “kız” ve “erkek” değişkenleri ayrı ayrı tanımlanır ve *Data View* penceresinde bu değişkenlere ait gözlem değerleri girilir.

File	Edit	View	Data	Transform	Analyze	Direct Marketing	Graphs	Utilities	Add-ons	Window	Help
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	kız	Numeric	8	0		None	None	8	Right	Scale	Input

File	Edit	View	Data	Transform	Analyze	Direct Marketing	Graphs	Utilities	Add-ons	Window	Help
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	erkek	Numeric	8	0		None	None	8	Right	Scale	Input

File	Edit	View	Data
9 :			
	kız		
1		10	
2		12	
3		14	
4		48	
5		30	

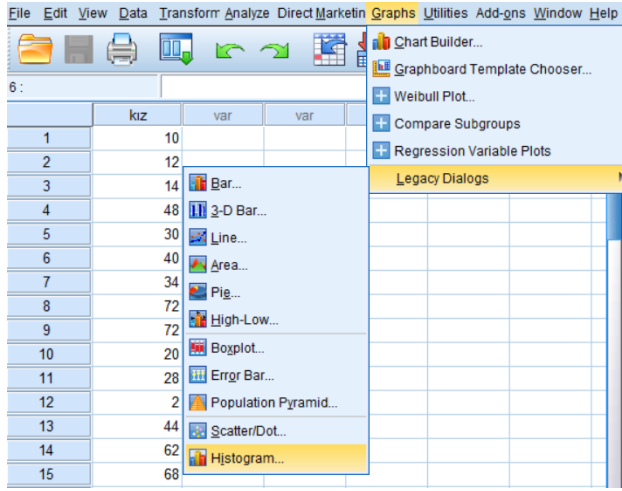
File	Edit	View	Data
	erkek		
1		2	
2		2	
3		2	
4		50	
5		2	

Öğrencilerin ileri matematik dersinden aldıkları notlar sürekli nicel veri türüne sahiptir. Sürekli nicel veriler için histogram grafiği çizdirmek uygun olacaktır. Histogram grafiği SPSS programında 3 farklı yoldan çizdirilebilir. Bu sorunun çözümünde birinci yoldan yararlanalım.

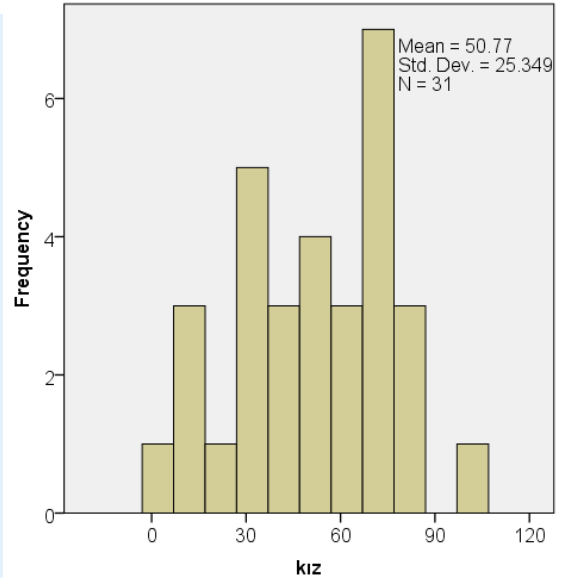
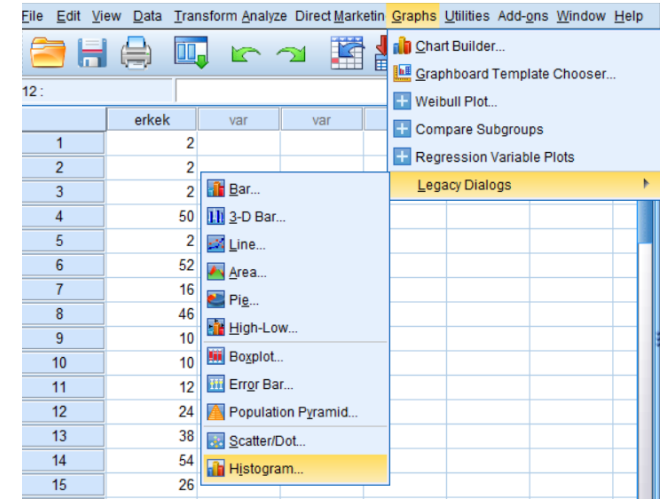
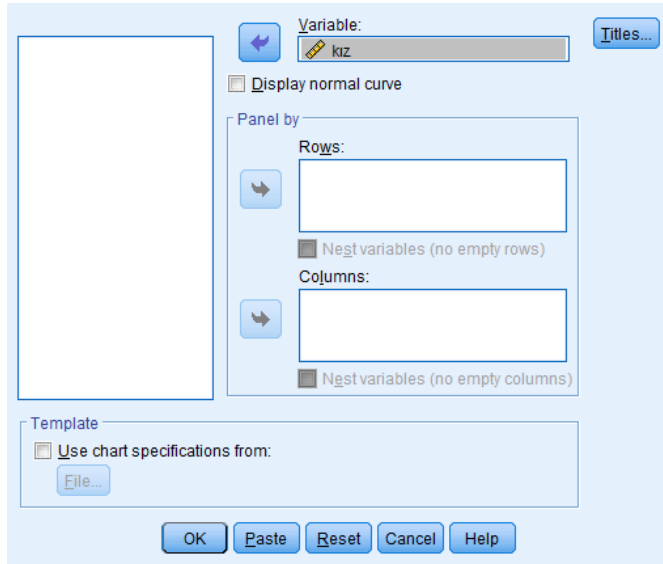
1. yol: Graphs → Legacy Dialogs → Histogram

2. yol: Analysis → Descriptive Statistics → Explore → Plots → Histogram

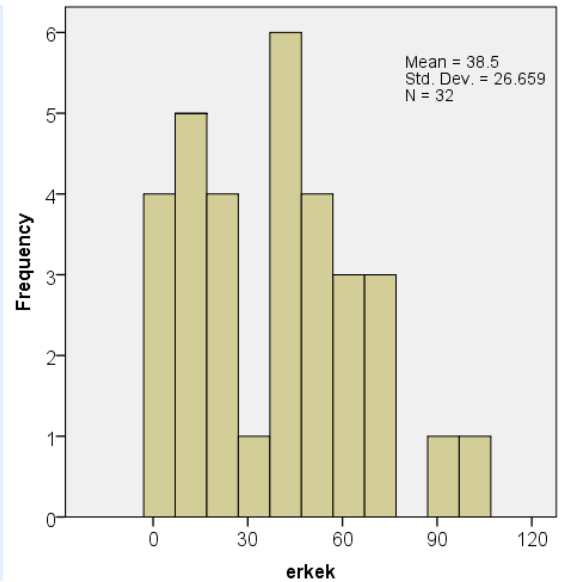
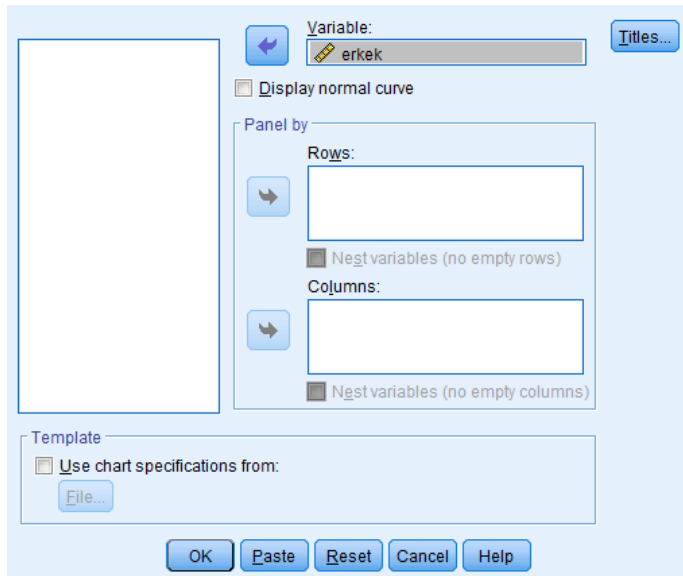
3. yol: Analysis → Descriptive Statistics → Frequencies → Charts → Histogram



“kız” değişkeni için histogram grafiğini çizdirelim.



“erkek” değişkeni için histogram grafiğini çizdirelim.



3) Türkiye’de 2003-2004 yılları arasındaki TEFE ve TÜFE değerleri aşağıda verilmiştir. Verilere uygun grafiği çiziniz.

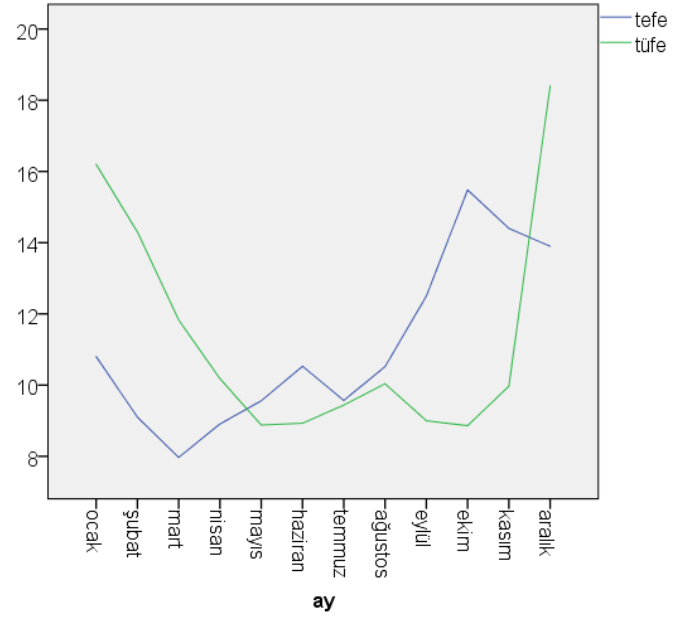
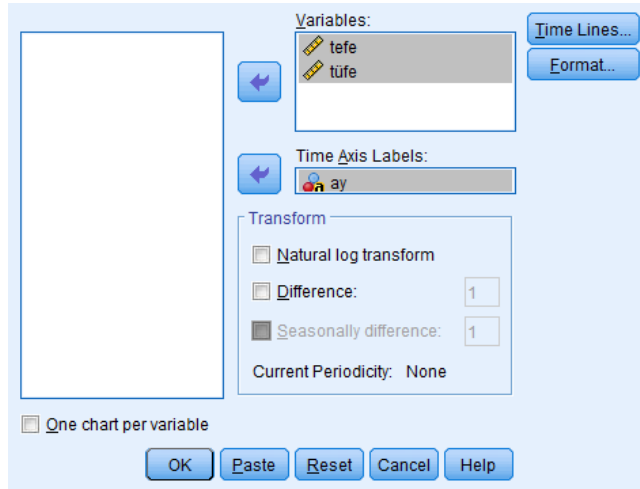
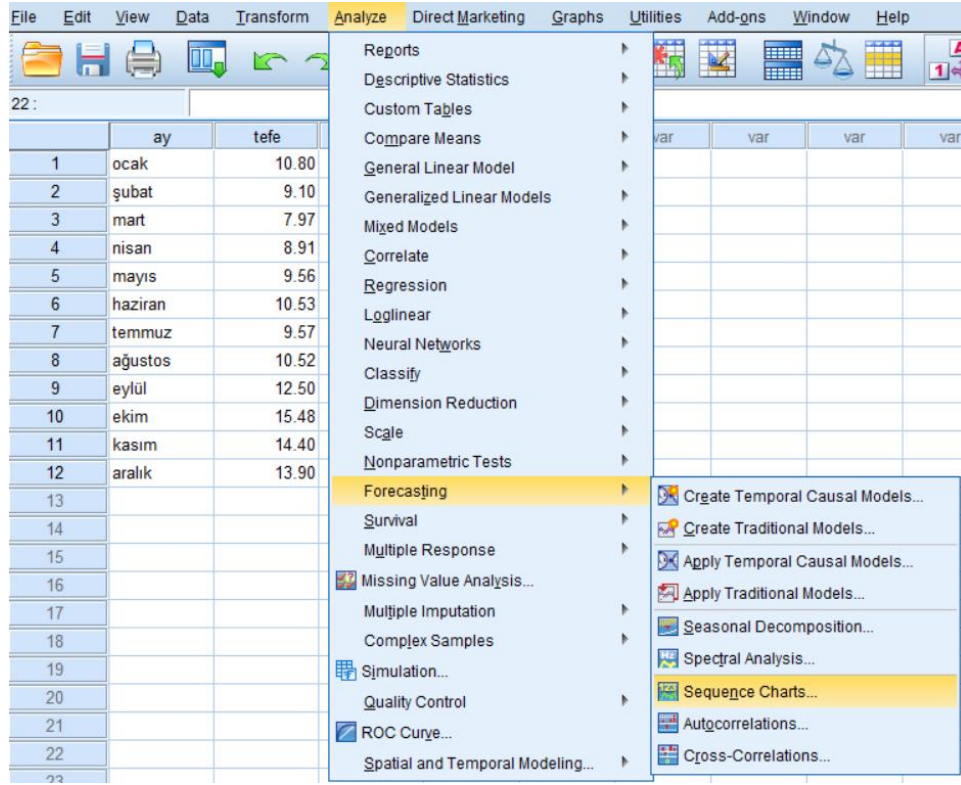
Aylar	Ocak	Şubat	Mart	Nisan	Mayıs	Haziran	Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık
TEFE	10,8	9,1	7,97	8,91	9,56	10,53	9,57	10,52	12,5	15,48	14,4	13,9
TÜFE	16,2	14,3	11,83	10,18	8,88	8,93	9,44	10,04	9	8,86	9,97	18,4

Soruda geçen Tefe ve Tüfe değerleri zamana bağlı (ay) olarak değişmektedir. Bu nedenle uygun grafik türü olarak trend grafiği seçilmelidir. *Variable View* penceresinde ilk olarak “ay”, “tefe” ve “tüfe” değişkenleri tanımlanır. *Data View* penceresinde bu değişkenlere ait gözlem değerleri girilir.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ay	String	8	0		None	None	8	Left	Nominal	Input
2	tefe	Numeric	8	2		None	None	8	Right	Scale	Input
3	tüfe	Numeric	8	2		None	None	8	Right	Scale	Input

	ay	tefe	tüfe
1	ocak	10.80	16.20
2	şubat	9.10	14.30
3	mart	7.97	11.83
4	nisan	8.91	10.18
5	mayıs	9.56	8.88
6	haziran	10.53	8.93
7	temmuz	9.57	9.44
8	ağustos	10.52	10.04
9	eylül	12.50	9.00
10	ekim	15.48	8.86
11	kasım	14.40	9.97
12	aralık	13.90	18.40

Tefe ve tüfe değerleri aylara bağlı olarak değiştiği için ve aynı zamanda aylara bağlı olarak tefe ve tüfe değerlerinin etkisini görmek istediğimiz için trend grafiğini çizdirelim.



Zaman serileri grafiğinden aylara bağlı olarak tefe ve tüfe değerleri aynı grafik içinde elde edilir. Eğer tefe ve tüfe değerleri için zaman serileri grafiğini ayrı grafiklerde görmek istersek izleyeceğimiz adımlar şu şekilde olacaktır:

Variables:

tefe
tüfe

Time Axis Labels:

ay

Transform

☐ Natural log transform

☐ Difference: 1

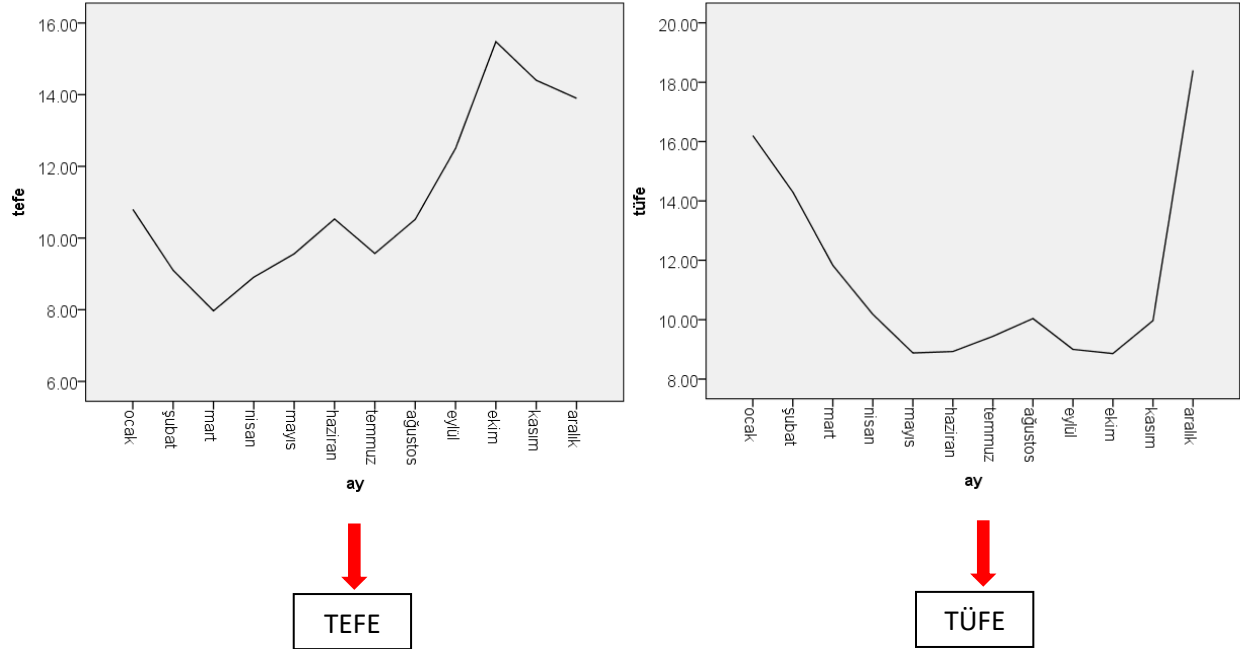
☐ Seasonally difference: 1

Current Periodicity: None

☒ One chart per variable

OK Paste Reset Cancel Help

“One chart per variable” seçeneğini seçmemiz sonucunda elde edeceğimiz grafikler aşağıdaki gibidir.



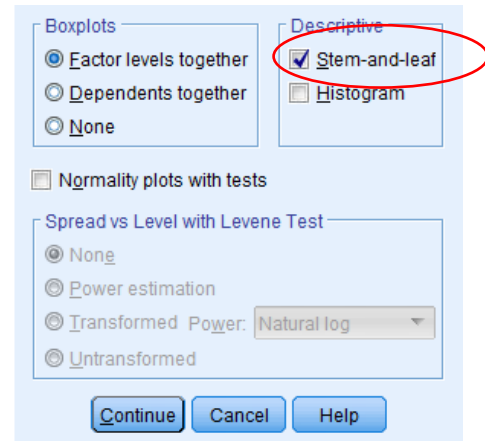
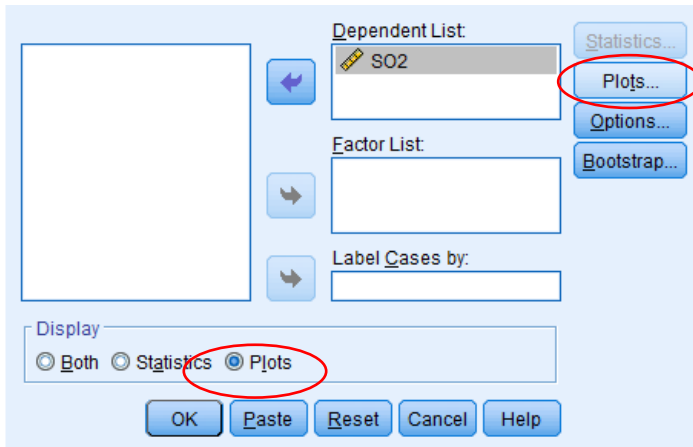
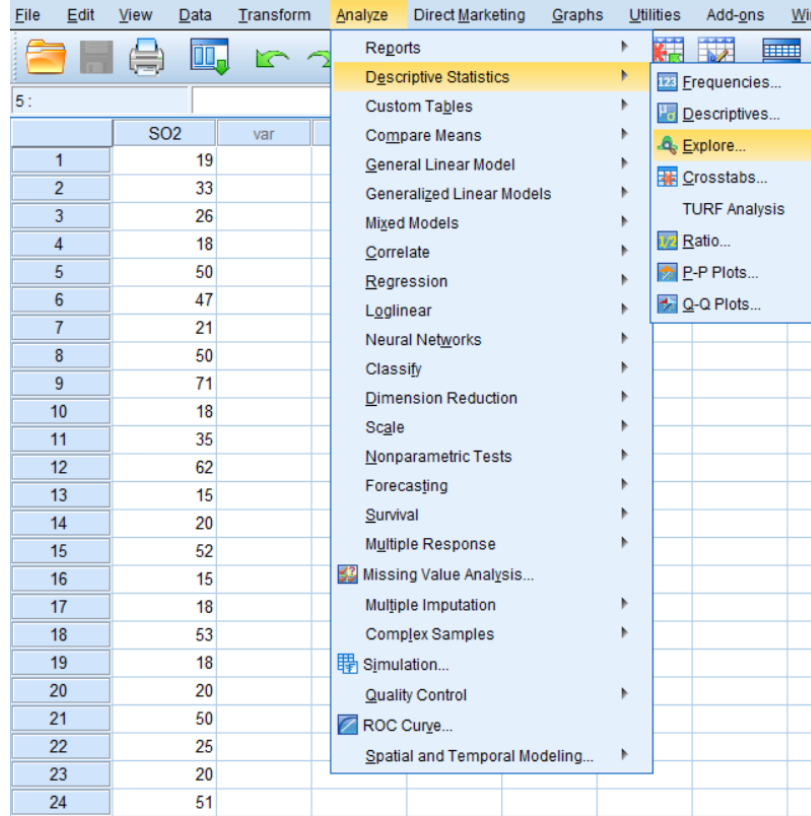
4) T.C. Sağlık Bakanlığı Refik Saydam Hıfzıssıhha Merkezi Başkanlığı tarafından yapılan ölçümlerde Ankara'nın Sıhhiye semtinden kükürt dioksit (SO_2) ölçümleri elde edilmiştir. Dal yaprak grafiğini çiziniz.

19	18	21	18	15	15	18	25
33	50	50	35	20	18	20	20
26	47	71	62	52	53	50	51

Bu soru için *Variable View* penceresinde “SO2” değişkeni tanımlanır ve *Data View* penceresinde bu değişkene ait gözlem değerleri girilir.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	SO2	Numeric	8	0		None	None	8	Right	Scale	Input

Dal yaprak grafiği çiziminde ise izlenecek adımlar şu şekildedir:



Dal-Yaprak grafiğimiz ise aşağıdaki gibi elde edilmektedir.

SO2 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
7.00	1 .	5588889
6.00	2 .	000156
2.00	3 .	35
1.00	4 .	7
6.00	5 .	000123
1.00	6 .	2
1.00	7 .	1

Stem width: 10
Each leaf: 1 case(s)

- Stem width: Dalın genişliği (x10)
- Each leaf: Her bir yaprağın kaç gözeyi ifade ettiğini gösterir.

5) 30 öğrenciye ilişkin akademik ortalama değerleri aşağıda verilmiştir. Verilen veriler için uygun grafiği çiziniz.

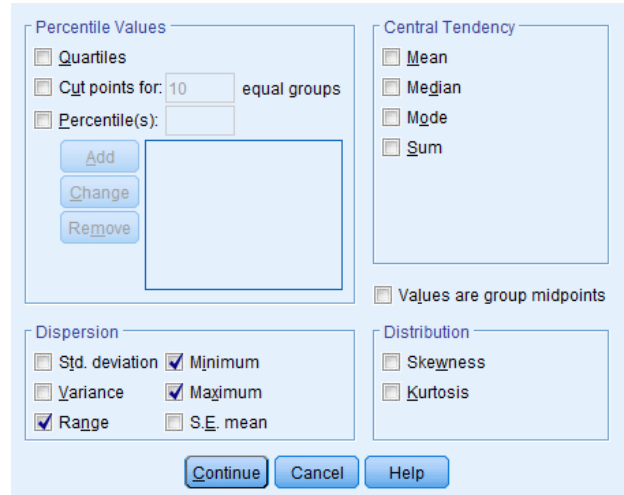
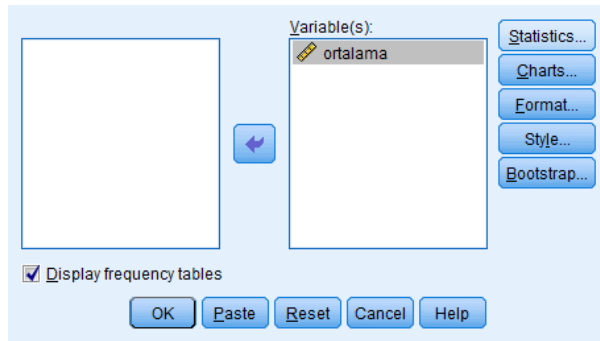
2,12	3,42	3,70	2,80	2,91	2,63
2,35	2,76	2,84	2,15	2,19	3,26
2,78	3,18	1,96	1,65	1,72	0,98
2,90	0,56	1,29	1,03	1,59	1,70
3,05	2,14	2,36	3,85	2,40	2,61

Akademik ortalamalara ait veriler sürekli nicel veri türüne sahiptir. Bu nedenle histogram, birikimli sıklık ve dağılım poligonu grafikleri çizilebilir. İlk olarak “ortalama” değişkeni *Variable View* sayfasında tanımlanır ve *Data View* kısmında ise veri girişi yapılır.

File	Edit	View	Data	Transform	Analyze	Direct Marketing	Graphs	Utilities	Add-ons	Window	Help
1	ortalama	Numeric	8	2	None	None	8	Right	Scale	Input	

Belirtilen grafikleri çizmeden önce gerekli hesaplamaların yapılması gerekmektedir. İlk olarak, veri setindeki minimum değer, maksimum değer ve dağılım genişliğini elde edelim.

File	Edit	View	Data	Transform	Analyze	Direct Marketing	Graphs	Utilities	Add-ons	Win
					Reports					
					Descriptive Statistics				Frequencies...	
					Custom Tables				Descriptives...	
					Compare Means				Explore...	
					General Linear Model				Crosstabs...	
					Generalized Linear Models				TURF Analysis	
					Mixed Models				Ratio...	
					Correlate				P-P Plots...	
					Regression				Q-Q Plots...	
					Loglinear					
	ortalama	var								
1	2.12									
2	2.35									
3	2.78									
4	2.90									
5	3.05									
6	3.42									
7	2.76									



Statistics

ortalama		
N	Valid	30
	Missing	0
Range		3.29
Minimum		.56
Maximum		3.85

Çıktı sayfasında görülen bu değerlerden yararlanılarak sınıf sayısı (k) ve sınıf aralığı (c) elde edilir.

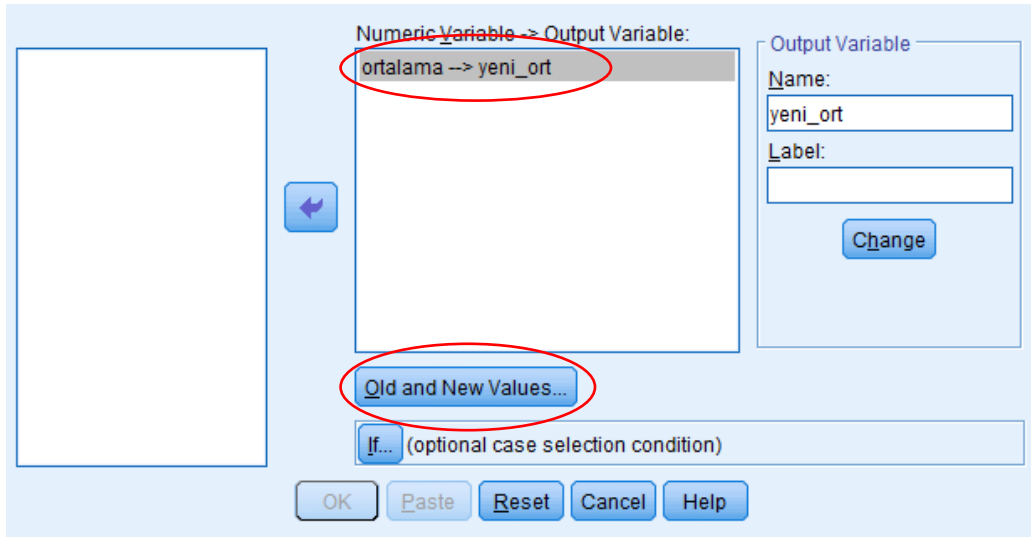
$$\text{sınıf sayısı } (k) = 1 + 3.3 \log(30) \cong 5.87 \cong 6$$

$$\text{sınıf aralığı } (c) = \frac{\text{Dağılım genişliği} + \text{son haneye 1 ekle}}{k} = \frac{3.29 + 0.01}{6} = 0.55$$

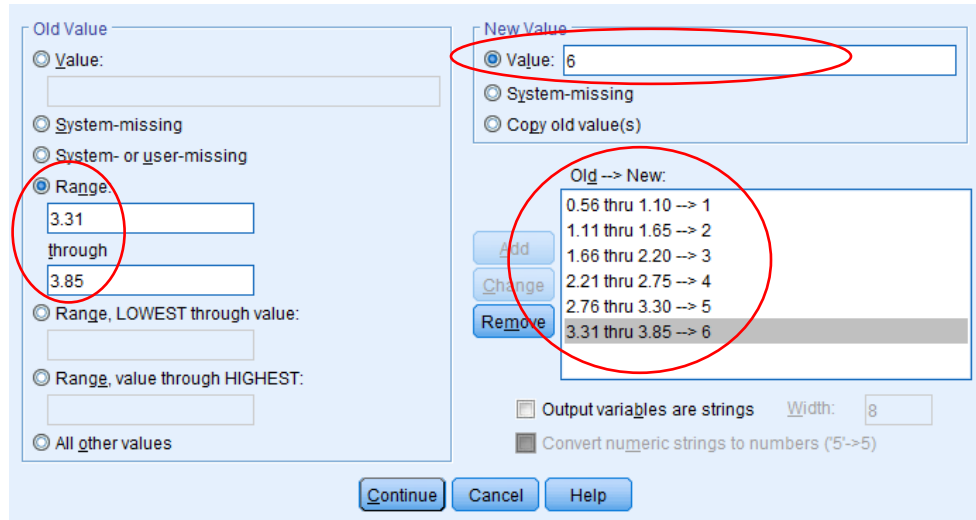
Belirlenen sınıf sayısı ve sınıf aralığı ile sınıfların alt sınır ve üst sınıf değerleri elde edilir:

Sınıf	Alt Sınır	Üst Sınır
1	0.56	1.10
2	1.11	1.65
3	1.66	2.20
4	2.21	2.75
5	2.76	3.30
6	3.31	3.85

Sınıf bazında konuşacağımız için SPSS programında nicel değerlerimiz nitel değişkene yani sınıflara dönüştürülür. Bu işlem için **Transform → Recode into Different Variables** sekmesinde aşağıda yer alan adımlar takip edilmelidir.



Açılan sayfada ise *Range* kısmından yararlanılarak her bir sınıfın alt ve üst değerleri girilir ve belirlenen 6 sınıf elde edilir.



Data View penceresinde oluşturulan yeni değişken görülür.



File Edit View Data Transform An		
4 :		
	ortalama	yeni_ort
1	2.12	3.00
2	2.35	4.00
3	2.78	5.00
4	2.90	5.00
5	3.05	5.00

yeni_ort değişkeni için belirlenen sınıf değerlerine göre etiketleme yapılır.

Value Labels

Value:

Label:

Spelling...

Add

Change

Remove

1.00 = "0.56-1.10"

2.00 = "1.11-1.65"

3.00 = "1.66-2.20"

4.00 = "2.21-2.75"

5.00 = "2.76-3.30"

6.00 = "3.31-3.85"

OK Cancel Help



	ortalama	yeni_ort
1	2.12	1.66-2.20
2	2.35	2.21-2.75
3	2.78	2.76-3.30
4	2.90	2.76-3.30
5	3.05	2.76-3.30

Elde edilen tüm bilgilerle sıklık çizelgesi çizdirilir. Bu işlem için **Analyze→ Descriptive Statistics** sekmesinden *Frequencies* seçeneği seçilir. Açılan pencerede artık *yeni_ort* değişkeni üzerinden işlem yapacağımız için bu değişken seçilir ve sayfada yer alan “*Display frequency tables*” seçeneği işaretlenerek işlem tamamlanır.

Variable(s):

ortalama

yeni_ort

Statistics...

Charts...

Format...

Style...

Bootstrap...

☒ Display frequency tables

OK Paste Reset Cancel Help

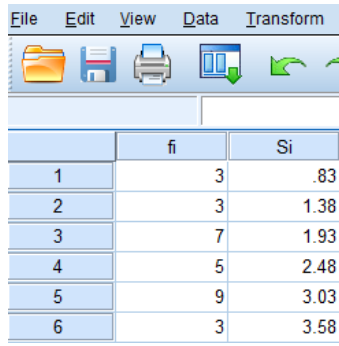
Çıktı sayfasında aşağıda yer alan tablo elde edilir. SPSS programında sıklık çizelgesi için frekans (sıklık), göreceli sıklık, birikimli göreceli sıklık görülebilmektedir.

		yeni_ort			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0.56-1.10	3	10.0	10.0	10.0
	1.11-1.65	3	10.0	10.0	20.0
	1.66-2.20	7	23.3	23.3	43.3
	2.21-2.75	5	16.7	16.7	60.0
	2.76-3.30	9	30.0	30.0	90.0
	3.31-3.85	3	10.0	10.0	100.0
	Total	30	100.0	100.0	

Sınıf	Alt Sınır	Üst Sınır	f_i	p_i	$\%F_i$	S_i	F_i
1	0.56	1.10	3	0.10	0.10	0.83	3
2	1.11	1.65	3	0.10	0.20	1.38	6
3	1.66	2.20	7	0.233	0.433	1.93	13
4	2.21	2.75	5	0.167	0.600	2.48	18
5	2.76	3.30	9	0.30	0.90	3.03	27
6	3.31	3.85	3	0.10	1	3.58	30

Elde edilen bu bilgilerle ilk olarak histogram grafiğini çizdirelim.

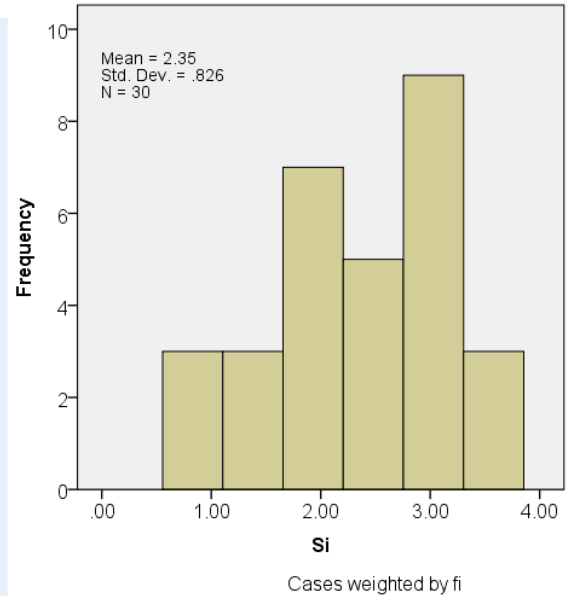
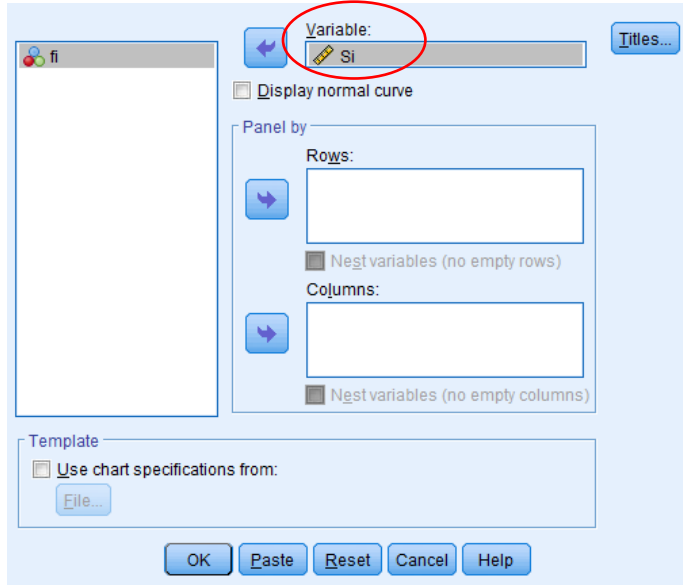
Histogram grafiğinin x ekseninde S_i , y ekseninde ise f_i değerlerimiz bulunmaktadır. Dolayısıyla, *Variable View* kısmında f_i ve S_i değişkenlerimizi tanımlayıp *Data View* penceresinde sıklık çizelgesinden yararlanarak bulduğumuz bu değerlerin girişi yapılır.



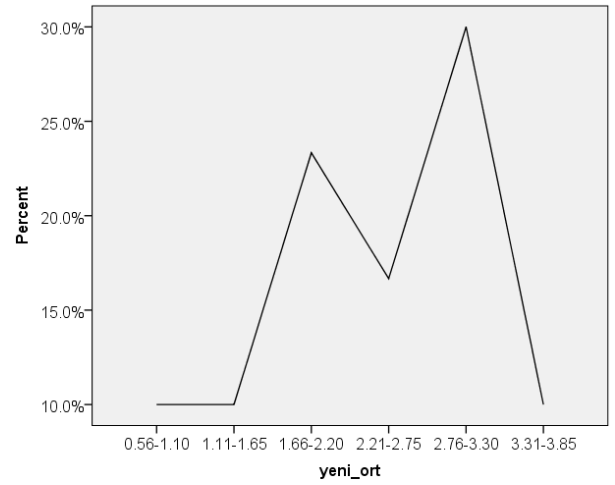
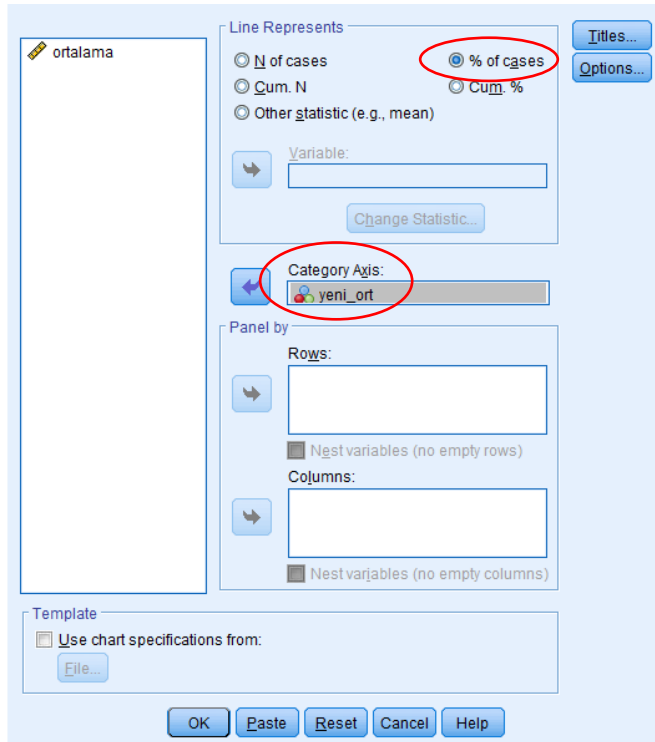
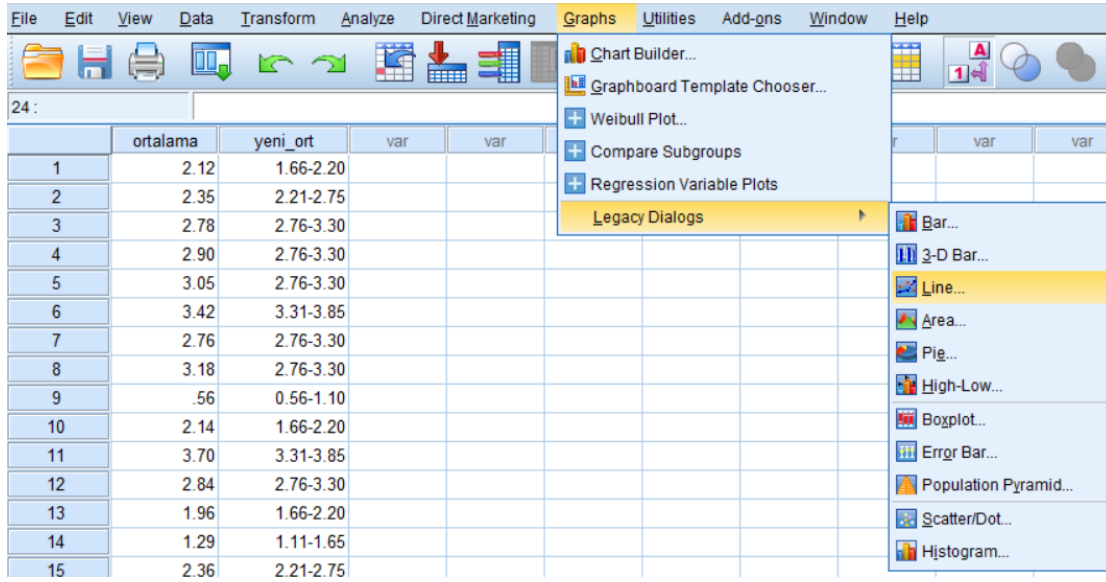
	fi	Si
1	3	.83
2	3	1.38
3	7	1.93
4	5	2.48
5	9	3.03
6	3	3.58

Frekans değerlerimizi belirten f_i 'ler için ağırlıklandırma işlemi yapılır.

f_i - S_i değerlerinden yararlanarak histogram grafiğini çizdirelim:



İkinci olarak, dağılışı poligonu grafiğini çizdirelim:



Son olarak, birikimli sıklık grafiğini çizdirelim.

ortalama

Line Represents

☐ N of cases ☐ % of cases ☒ Cum. %

Variable:

Change Statistic...

Category Axis:

yeni_ort

Panel by

Rows:

Nest variables (no empty rows)

Columns:

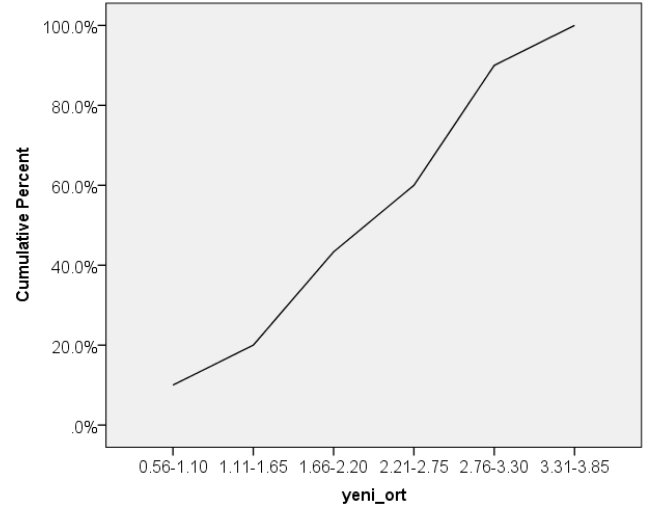
Nest variables (no empty columns)

Template

Use chart specifications from:

File...

OK Paste Reset Cancel Help



6) Aşağıda yer alan veriler için dal-yaprak grafiğini çiziniz.

23,25	24,13	24,76	24,81	24,98	25,31	25,57	25,89	26,28	26,34	27,09
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

4.soruda dal-yaprak grafiğinin nasıl elde edileceğine dair adımlar detaylı olarak verilmişti. Bu soru için adımlar tekrarlanmadan direk sonuç kısmına geçilmektedir.

File	Edit	View	Data	Transform	Analyze	Direct Marketing	Graphs	Utilities	Add-ons	Window	Help
1	gözlem	Numeric	8	2		None	None	8	Right	Scale	Input

gözlem Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	23 .	2
4.00	24 .	1789
3.00	25 .	358
2.00	26 .	23
1.00	27 .	0

Stem width: 1.00
Each leaf: 1 case(s)

File	Edit	View	Data
1	gözlem	23.25	
2	24.13		
3	24.76		
4	24.81		
5	24.98		
6	25.31		
7	25.57		
8	25.89		
9	26.28		
10	26.34		
11	27.09		

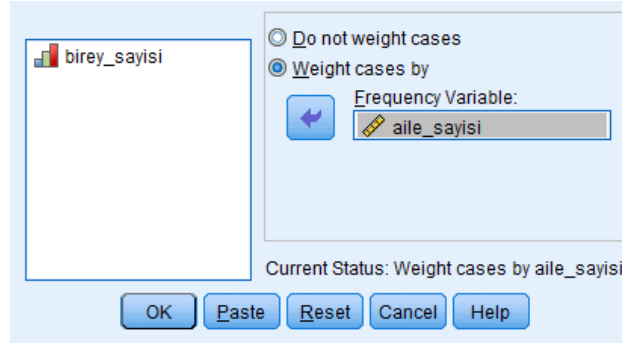
7) 30 aile üzerinden yapılan bir çalışmada ailelerin birey sayılarına göre dağılımları verilmiştir. Bu veriye ait uygun grafiği çiziniz.

Birey Sayısı	Aile Sayısı
4-6	4
7-9	13
10-12	10
13-15	3

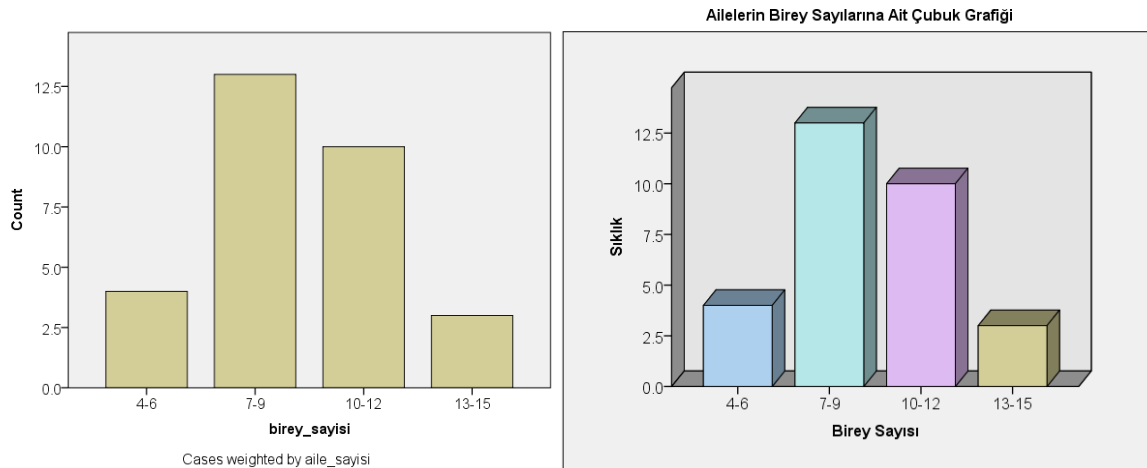
Variable View penceresinde ilk olarak “birey_sayisi” ve “aile_sayisi” değişkeni tanımlanır. Data View penceresinde bu değişkenlere ait gözlem değerleri girilir.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	birey_sayisi	Numeric	8	0		{1, 4-6}...	None	8	Right	Ordinal	Input
2	aile_sayisi	Numeric	8	0		None	None	8	Right	Scale	Input

Birey sayısı için etiketleme işlemi (1 → 4-6, 2 → 7-9, 3 → 10-12, 4 → 13-15) ve aile sayısı için ise ağırlıklandırma işlemi yapılır.



Sıklık çizelgesi şeklinde düzenlenmiş bir soru olduğundan çubuk grafiği çizdirilebilecek uygun bir grafik türüdür. Çubuk grafiğinin nasıl elde edileceğine dair adımlar daha önceki sorularda detaylı bir şekilde verilmiştir.



8) Türkiye'nin yıllara göre (1998-2006) ayçiçeği ihracatı verileri verilmiştir. Verilerin zaman serisi (trend) grafiğini çiziniz.

Yıl	İhracat
1998	1750
1999	2500
2000	3750
2001	4000
2002	5000
2003	7500
2004	10000
2005	12500
2006	16500

İlk olarak, *Variable View* penceresinde ilk olarak “zaman” ve “ihracat” değişkenleri tanımlanır. *Data View* penceresinde bu değişkenlere ait gözlem değerleri girilir.

The image shows two screenshots from the SPSS software interface. The top screenshot is the *Variable View* window, which displays two variables: 'zaman' (Time) and 'ihracat' (Export). Both are defined as Numeric, with a width of 8 and 0 decimal places. The bottom screenshot is the *Time Series Diagnostics* dialog box. In the 'Variables' list, 'ihracat' is selected. In the 'Time Axis Labels' list, 'zaman' is selected. The 'Transform' section shows options for 'Natural log transform', 'Difference' (set to 1), and 'Seasonally difference' (set to 1). The 'Current Periodicity' is set to 'None'. The 'One chart per variable' checkbox is checked. Buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' are at the bottom.

3.soruda zaman serileri grafiğinin nasıl elde edileceğine dair adımlar detaylı olarak verilmişti. Bu soru için adımlar tekrarlanmadan direk sonuç kısmına geçilmektedir.

