



A.Ü. Beypazarı MYO İstatistik Dersi



Ünite 5

Verilerin sınıflandırılması

Ünitede Ele Alınan Konular

5. Verilerin Sınıflandırılması

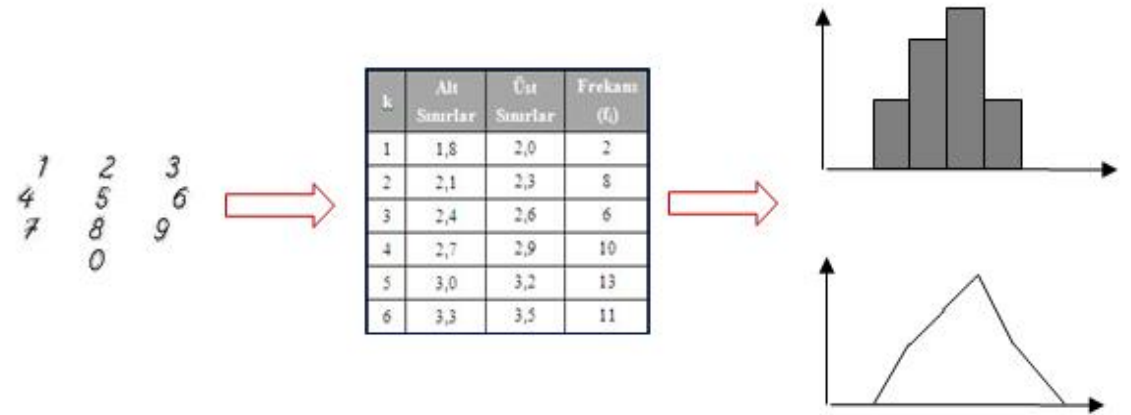
- Nicel verilerin Sınıflandırılması
- Histogram ve Frekans Poligonu
- Nitel verilerin Sınıflandırılması



Verilerin Sınıflandırılması

Uygulamada veri setleri çok sayıda gözlem içerebilmektedir. Bu nedenle gözlem sayısının **çok olduğu** bir seride verilerin tek tek temsil edilmeleri yerine sınıflandırılmaları ve sınıflandırılan bu verilerin tablo ya da grafik yöntemlerle sunulmaları tercih edilir.

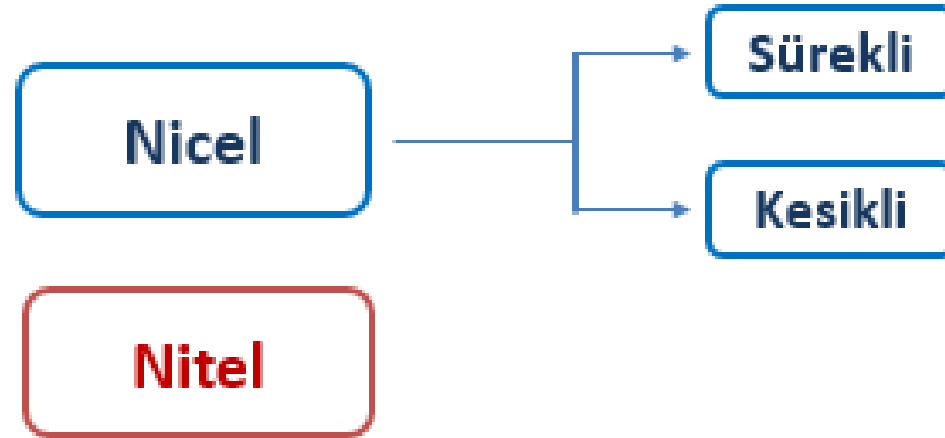
Verinin, oluşturulan sınıflarda, bu sınıflara düşen gözlem sayıları (Frekansları) ile birlikte gösterildiği özetleme yöntemine sınıflandırılmış frekans dağılım tablosu ya da kısaca Frekans dağılım Tablosu denmektedir.



Çok sayıda veriyi özetlemenin en etkin yolu verinin frekans dağılımını oluşturmaktır.

Veriler için Frekans Dağılım Tablosunun Oluşturulması

Bir araştırmada incelenen konuya ilişkin elde edilen veriler, ölçüldükleri ölçü birimi bakımından



1. Sürekli verilerin sınıflandırılması
2. Kesikli verilerin sınıflandırılması
3. Nitel verilerin sınıflandırılması

Sürekli verilerin sınıflandırılması

- 1 Değişim aralığını belirle ($D.A = X_{\max} - X_{\min}$)
- 2 Sınıf sayısı belirlenir ($k = 5 - 20$ arasında olabilir)
- 3 Sınıf Aralığı (c) hesaplanır.
$$c = \frac{\text{Değişim Aralığı}}{k}$$
- 4 Sınıflar oluşturulur.
- 5 Serideki gözlem değerleri hangi sınıfa düşüyorsa o sınıfa kaydedilerek her sınıftaki veri sayısı yani **sınıf frekansları** (f_i) elde edilir.
- 6 Sınıf Değerleri (x_i) Hesaplanır.
- 7 Sınıflandırılmış frekans (sıklık) dağılımlarının grafik gösteriminde **Histogram** yada **Poligon** kullanılır.

Örnek 1

Bir bölgede faaliyet gösteren bilgisayar donanım ve yazılım şirketlerinin ciroları ile ilgili bir araştırmada, 40 firmanın bir önceki yıla ait gelirleri (1000 TL) elde edilmiştir.

Veriler aşağıdaki tabloda verilmiştir. Bu verileri 5 sınıfta özetleyiniz.

Yıllık Cirolar(X)	133	84	105	90	92	97	66	106	101	71
	80	79	68	95	81	89	99	69	75	111
	82	80	77	74	65	95	82	69	72	82
	96	72	88	113	85	94	105	69	65	75

Çözüm

Verilenler:

$$n= 40$$

$$k=5$$

Yıllık Cirolar(X)	65	65	66	68	69	69	69	71	72	72
	74	75	75	77	79	80	80	81	82	82
	82	84	85	88	89	90	92	94	95	95
	96	97	99	101	105	105	106	111	113	133

1

$$\text{Değişim Aralığı}_x = X_{\max} - X_{\min}$$

$$X_{\max} = 133$$

$$X_{\min} = 65$$

$$\text{Değişim Aralığı}_x = 133-65$$

$$\text{Değişim Aralığı}_x = 68$$

2

k= 5 verilmiştir.

3

$$\text{Sınıf Aralığı (c)} = \frac{\text{Değişim Aralığı}}{\text{Sınıf Sayısı}}$$

$$\text{Sınıf Aralığı (c)} = \frac{68}{5}$$

$$\text{Sınıf Aralığı (c)} = 13,6 = \mathbf{14} \text{ (Bin) TL}$$

4

Sınıflar oluşturulur.

 $k=5$ ve $c=14$, $X_{\min}=65$ ve $X_{\max}=133$

Yıllık Cırolar(X)	65	65	66	68	69	69	69	71	72	72
	74	75	75	77	79	80	80	81	82	82
	82	84	85	88	89	90	92	94	95	95
	96	97	99	101	105	105	106	111	113	133

k	Alt Sınırlar
1	65
2	79
3	93
4	107
5	121

k	Alt Sınırlar(A.S)	Üst Sınırlar (Ü.S)
1	65	78
2	79	92
3	93	106
4	107	120
5	121	134

5

Sınıfların frekansları hesaplanır.

Yıllık Cirolar(X)	65	65	66	68	69	69	69	71	72	72
	74	75	75	77	79	80	80	81	82	82
	82	84	85	88	89	90	92	94	95	95
	96	97	99	101	105	105	106	111	113	133

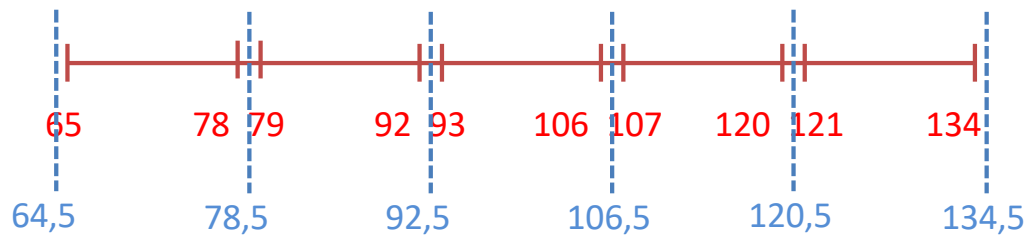
k	Alt Sınırlar	Üst Sınırlar	f
1	65	78	14
2	79	92	13
3	93	106	10
4	107	120	2
5	121	134	1

Sınıf Değerleri Hesaplanır.

k	Alt Sınırlar	Üst Sınırlar	f_i	Sınıf Değerleri (x_i)	Nispi Frekans (% f)
1	65	78	14	71,5	0,350
2	79	92	13	85,5	0,325
3	93	106	10	99,5	0,250
4	107	120	2	113,5	0,050
5	121	134	1	127,5	0,025

Frekans Dağılım Tablosu

k	Alt Sınırlar	Üst Sınırlar	f	Sınıf Değerleri	Gerçek Sınırlar
1	65	78	14	71,5	64,5 – 78,5
2	79	92	13	85,5	78,5 – 92,5
3	93	106	10	99,5	92,5 – 106,5
4	107	120	2	113,5	106,5 – 120,5
5	121	134	1	127,5	120,5 – 134,5

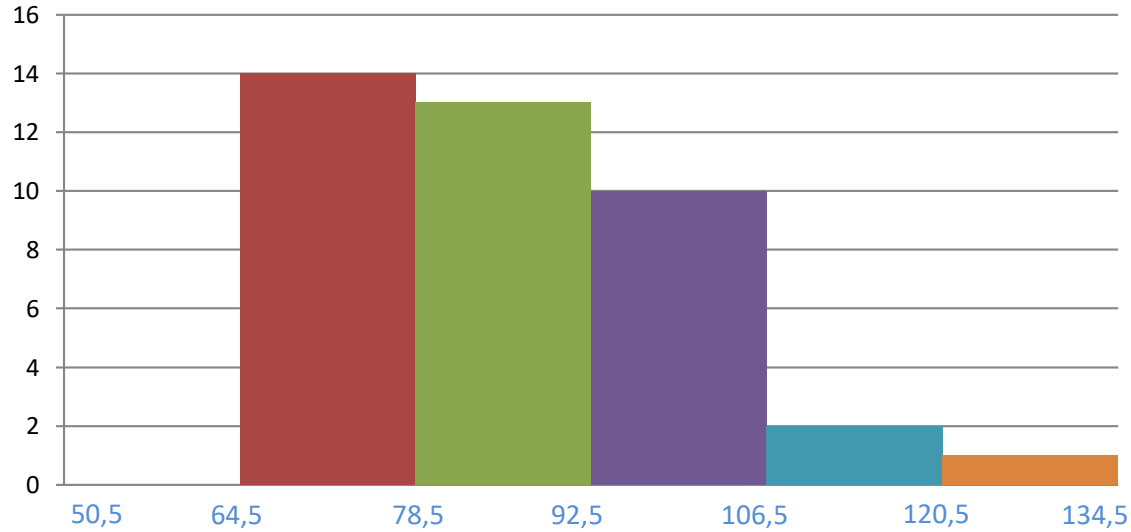


7

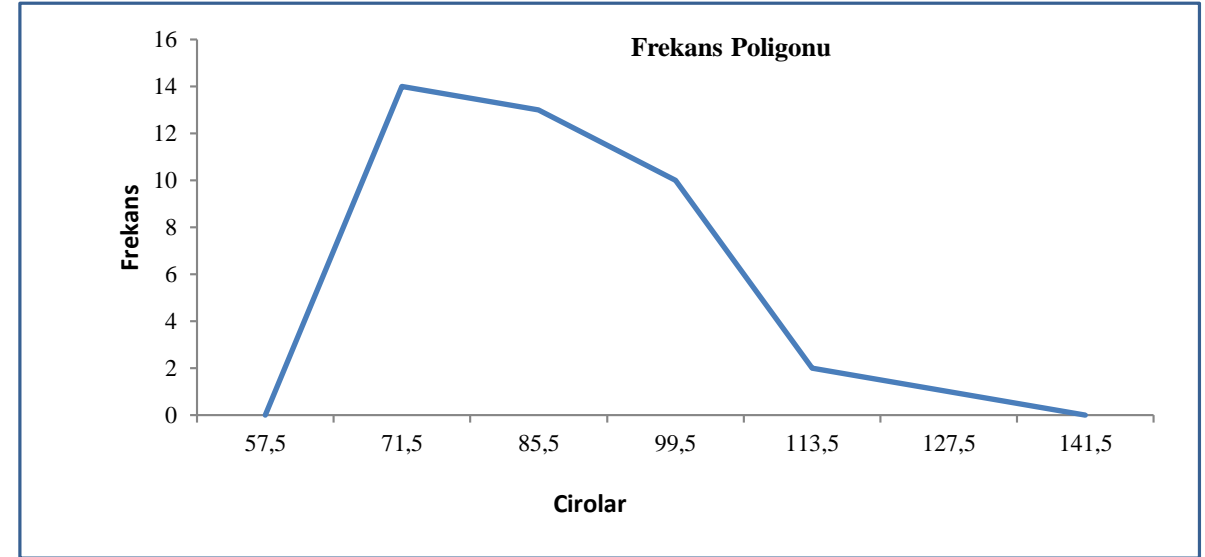
Histogram ve Poligon

k	Alt Sınırlar	Üst Sınırlar	f	Sınıf Değerleri	Gerçek Sınırlar
1	65	78	14	71,5	64,5 – 78,5
2	79	92	13	85,5	78,5 – 92,5
3	93	106	10	99,5	92,5 – 106,5
4	107	120	2	113,5	106,5 – 120,5
5	121	134	1	127,5	120,5 – 134,5

Histogram



Frekans Poligonu



Örnek 2

Bir bankanın havale servisine herhangi bir günde gelen müşterilerin işlemlerini bitirme süreleri **dakika olarak** ölçülmüştür.

Sınıf sayısını (k) **6** olarak verilerin frekans dağılım tablosunu oluşturunuz. Histogram ve frekans poligonunu çiziniz.

Havale İşlem Süreleri (x)									
3,3	2,3	2,3	3,4	2,7	2,3	3,3	2,5	3,5	3,2
2,1	3,4	3,1	3,1	2,9	3,4	2,4	2,3	2,4	2,7
3,1	2,4	2,8	1,8	3,2	2,1	3,1	2,7	3,5	2,5
3,1	3,2	2,8	3,1	2,2	3,5	3,2	2,9	2,5	3,0
2,8	3,4	2,0	3,0	2,8	2,1	3,1	3,4	3,5	2,9

Çözüm

Verilenler

n= 50

k=6

1,8	2,0	2,1	2,1	2,1	2,2	2,3	2,3	2,3	2,3
2,4	2,4	2,4	2,5	2,5	2,5	2,7	2,7	2,7	2,8
2,8	2,8	2,8	2,9	2,9	2,9	3,0	3,0	3,1	3,1
3,1	3,1	3,1	3,1	3,1	3,2	3,2	3,2	3,2	3,3
3,3	3,4	3,4	3,4	3,4	3,4	3,5	3,5	3,5	3,5

1

$$\text{Değişim Aralığı}_x = X_{\max} - X_{\min}$$

$$X_{\max} = 3,5$$

$$X_{\min} = 1,8$$

$$\text{Değişim Aralığı}_x = 3,5 - 1,8$$

$$\text{Değişim Aralığı}_x = 1,7 \text{ dakika}$$

2

k= 6 verilmiştir.

3

$$\text{Sınıf Aralığı (c)} = \frac{\text{Değişim Aralığı}}{\text{Sınıf Sayısı}}$$

$$\text{Sınıf Aralığı (c)} = \frac{1,7}{6}$$

$$\text{Sınıf Aralığı (c)} = 0,283 = \mathbf{0,3 \text{ dakika}}$$

4

Sınıflar oluşturulur.

 $k=6$ $c=0,3$ dak $X_{\min}=1,8$ $X_{\max}=3,5$

1,8	2,0	2,1	2,1	2,1	2,2	2,3	2,3	2,3	2,3
2,4	2,4	2,4	2,5	2,5	2,5	2,7	2,7	2,7	2,8
2,8	2,8	2,8	2,9	2,9	2,9	3,0	3,0	3,1	3,1
3,1	3,1	3,1	3,1	3,1	3,2	3,2	3,2	3,2	3,3
3,3	3,4	3,4	3,4	3,4	3,4	3,5	3,5	3,5	3,5

k	Alt Sınırlar
1	1,8
2	2,1
3	2,4
4	2,7
5	3,0
6	3,3

k	Alt Sınırlar	Üst Sınırlar
1	1,8	(2,1 - 0,1) = 2,0
2	2,1	2,3
3	2,4	2,6
4	2,7	2,9
5	3,0	3,2
6	3,3	3,5

Sınıfların frekansları hesaplanır.

Verilenler: $n= 50$ ve $k=6$

1,8	2,0	2,1	2,1	2,1	2,2	2,3	2,3	2,3	2,3
2,4	2,4	2,4	2,5	2,5	2,5	2,7	2,7	2,7	2,8
2,8	2,8	2,8	2,9	2,9	2,9	3,0	3,0	3,1	3,1
3,1	3,1	3,1	3,1	3,1	3,2	3,2	3,2	3,2	3,3
3,3	3,4	3,4	3,4	3,4	3,4	3,5	3,5	3,5	3,5

k	Alt Sınırlar	Üst Sınırlar	Frekans (f_i)
1	1,8	2,0	2
2	2,1	2,3	8
3	2,4	2,6	6
4	2,7	2,9	10
5	3,0	3,2	13
6	3,3	3,5	11
			$\Sigma f_i=50$

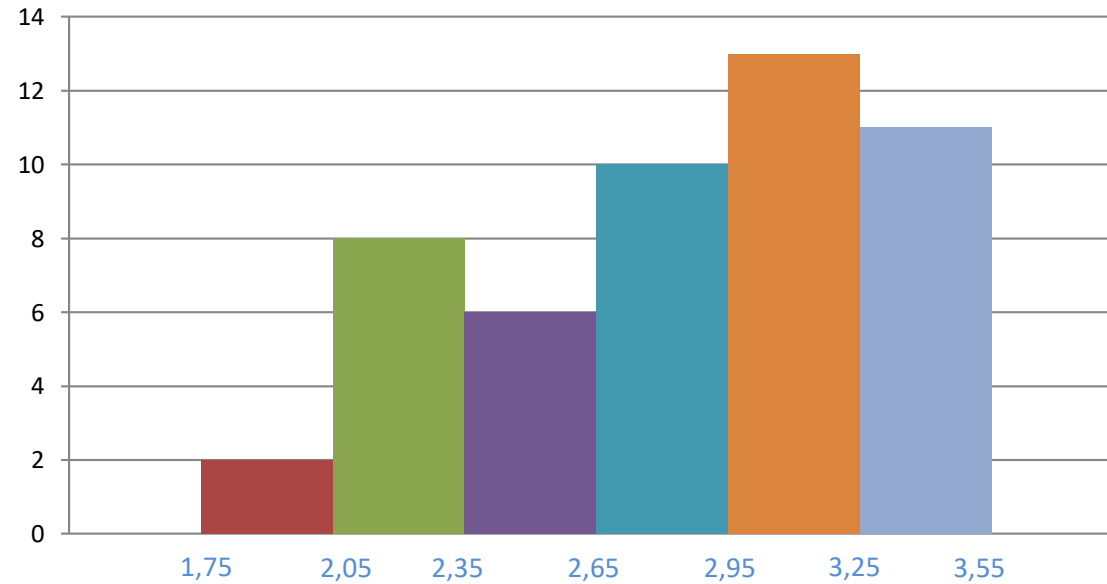
6

Sınıf Değerleri Hesaplanır.

k	Sınıflar		Frekans (f_i)	Sınıf Değeri(X_i)	Oransal Frekans(%f)
	A.S	Ü.S			
1	1,8	2,0	2	1,9	0,04
2	2,1	2,3	8	2,2	0,16
3	2,4	2,6	6	2,5	0,12
4	2,7	2,9	10	2,8	0,20
5	3,0	3,2	13	3,1	0,26
6	3,3	3,5	11	3,4	0,22
			$\Sigma f_i=50$		1

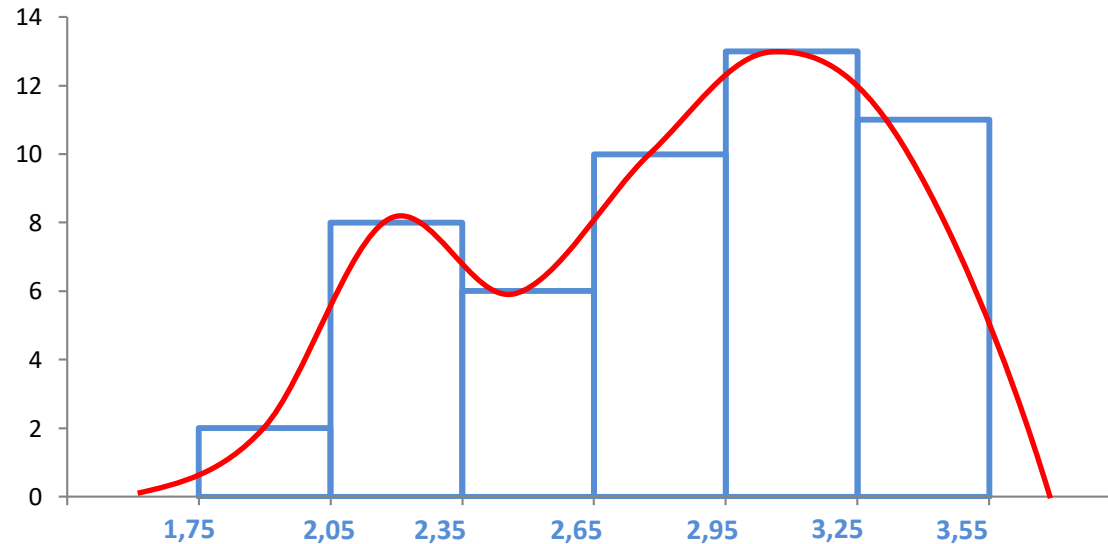
Histogram

k	Sınıflar		Gerçek Sınırlar	Frekans (f_i)	Sınıf Değeri(X_i)	Oransal Frekans(%f)
	A.S	Ü.S				
1	1,8	2,0	1,75-2,05	2	1,9	0,04
2	2,1	2,3	2,05-2,35	8	2,2	0,16
3	2,4	2,6	2,35-2,65	6	2,5	0,12
4	2,7	2,9	2,65-2,95	10	2,8	0,20
5	3,0	3,2	2,95-3,25	13	3,1	0,26
6	3,3	3,5	3,25-3,55	11	3,4	0,22
				$\Sigma f_i=50$		1



Frekans Poligonu

k	Sınıflar		Gerçek Sınırlar	Frekans (f_i)	Sınıf Değeri(X_i)	Oransal Frekans(%f)
	A.S	Ü.S				
1	1,8	2,0	1,75-2,05	2	1,9	0,04
2	2,1	2,3	2,05-2,35	8	2,2	0,16
3	2,4	2,6	2,35-2,65	6	2,5	0,12
4	2,7	2,9	2,65-2,95	10	2,8	0,20
5	3,0	3,2	2,95-3,25	13	3,1	0,26
6	3,3	3,5	3,25-3,55	11	3,4	0,22
				$\Sigma f_i=50$		1



Kesikli verilerin sınıflandırılması

Araştırmaya konu olan bazı özelliklere ilişkin veriler sayılarak elde edilir.

- ✓ Bir hanedeki elektronik cihazların sayısı
- ✓ Bir markete gelen günlük müşteri sayısı
- ✓ Bir günde Kaza yapan araç sayısı
- ✓ Bir ailedeki çocuk sayısı

gibi bazı özelliklere ilişkin değerler sayılarak elde edilir. Gözlem birimlerinde bu özelliklerin küsuratlı değer alması beklenemez. Bu özellikler tamsayı değeri alırlar.

Kesikli verilere ilişkin frekans dağılım tablolarının oluşturulması sürekli verilere nazaran daha kolaydır.

Bu tür veriler için Sınıf sınırları birbirinden **doğal olarak** ayrılmaktadır.

Ayrıca veriler sürekli olmadığından **gerçek sınırların belirlenmesine de gerek yoktur.**

Kesikli veriler için grafik gösterimlerde sütun (bar) grafiği ya da çizgi grafiği kullanılır.

Kesikli verilerin sınıflandırılmasında izlenecek adımlar

- 1 Değişim aralığını belirle ($D.A = X_{\max} - X_{\min}$)
- 2 Sınıf sayısı belirlenir ($k = 5 - 20$ arasında olabilir)
- 3 Sınıf Aralığı (c) hesaplanır.
$$c = \frac{\text{Değişim Aralığı}}{k}$$
- 4 Sınıflar oluşturulur.
- 5 Serideki gözlem değerleri hangi sınıfa düşüyorsa o sınıfa kaydedilerek her sınıftaki veri sayısı yani **sınıf frekansları** elde edilir.
- 6 Sınıf Değerleri Hesaplanır.
- 7 Sınıflandırılmış frekans (sıklık) dağılımlarının grafik gösteriminde **Bar grafiği** kullanılır.

Örnek 3

Bir bölgede faaliyet gösteren KOBİ'lerde çalışan işçi sayılarını araştırmak amacıyla **30 işletme incelenmiş** ve bu işletmelerde çalışan sayısı tespit edilerek aşağıdaki tabloda verilmiştir.

Verileri **4 sınıfta** özetleyiniz.

Firma no	İşçi sayısı	Firma no	İşçi sayısı	Firma no	İşçi sayısı
1	21	11	30	21	27
2	15	12	17	22	13
3	18	13	30	23	16
4	27	14	31	24	13
5	23	15	29	25	32
6	15	16	32	26	32
7	25	17	31	27	33
8	25	18	16	28	35
9	15	19	19	29	30
10	17	20	35	30	27

Çözüm

Verilenler: $n= 30$ ve $k=4$

Verileri sıralayalım

$Z=$ Bir İşletmede çalışan işçi sayısını gösterebilir.

İşçi sayısı	İşçi sayısı	İşçi sayısı
13	19	30
13	21	30
15	23	31
15	25	31
15	25	32
16	27	32
16	27	32
17	27	33
17	29	35
18	30	35

1

$$\text{Değişim Aralığı}_z = Z_{\max} - Z_{\min}$$

$$Z_{\max} = 35$$

$$Z_{\min} = 13$$

$$\text{Değişim Aralığı}_z = 35-13$$

$$\text{Değişim Aralığı}_z = 22$$

2

$k= 4$ verilmiştir.

3

$$\text{Sınıf Aralığı } (c) = \frac{\text{Değişim Aralığı}}{\text{Sınıf Sayısı}}$$

$$\text{Sınıf Aralığı } (c) = \frac{22}{4}$$

$$\text{Sınıf Aralığı } (c) = 5,5 = 6$$

4

Sınıflar oluşturulur.

 $k=4$ ve $c=6$, $Z_{\min}=13$ ve $Z_{\max}=35$

k	Alt Sınırlar
1	13
2	19
3	25
4	31

k	Alt Sınırlar	Üst Sınırlar
1	13	$(19-1) = 18$
2	19	24
3	25	30
4	31	36

5

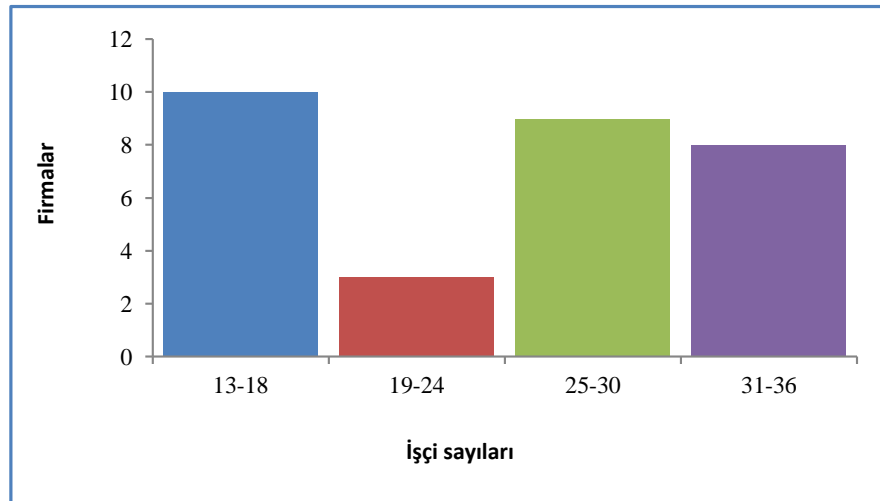
Sınıfların frekansları hesaplanır. Verilenler: $n= 30$ ve $k=4$

13	13	15	15	15	16	16	17	17	18
19	21	23	25	25	27	27	27	29	30
30	30	31	31	32	32	32	33	35	35

6

k	Alt Sınırlar	Üst Sınırlar	Sınıf Değerleri	Frekans (f_i)	Oransal Frekans ($\%f_i$)
1	13	18	15,5	10	$10/30 = 0,33$
2	19	24	21,5	3	$3/30 = 0,10$
3	25	30	27,5	9	$9/30 = 0,30$
4	31	36	33,5	8	$8/30 = 0,267$
				$\Sigma f_i=30$	1

7



Nitel verilerin sınıflandırılması

Araştırmaya konu olan bazı özelliklere ilişkin veriler sözle ifade edilir.

- ✓ Sahip olunan aracın markası
- ✓ Eğitim Düzeyi
- ✓ Cinsiyet
- ✓ Medeni Hal

gibi bazı özelliklere ilişkin değerler sözle ifade edilir.

Kategorilere ayrılan verilerde, sınıfların elde edilmesi için işlem yapmaya gerek kalmaz çünkü her **bir kategori bir sınıfa** temsil eder.

Her sınıfa düşen gözlem sayısı ile frekans dağılım tablosu oluşturulur.

Frekans dağılım tablosunda sadece **sınıflar** (kategoriler), **frekans** ve **oransal frekans** sütunları yer alır.

Örnek 4

Bir sınıfta öğrencilere babalarının eğitim düzeyi sorulmuş ve cevaplar aşağıdaki tabloda verilmiştir.

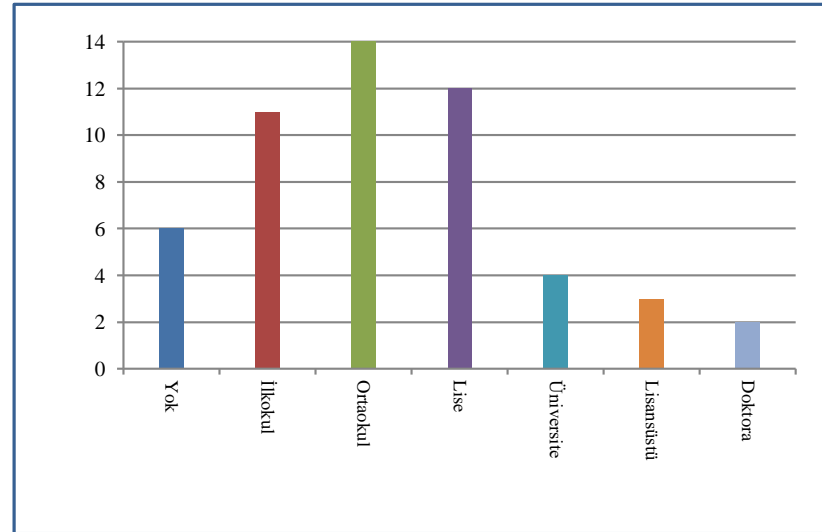
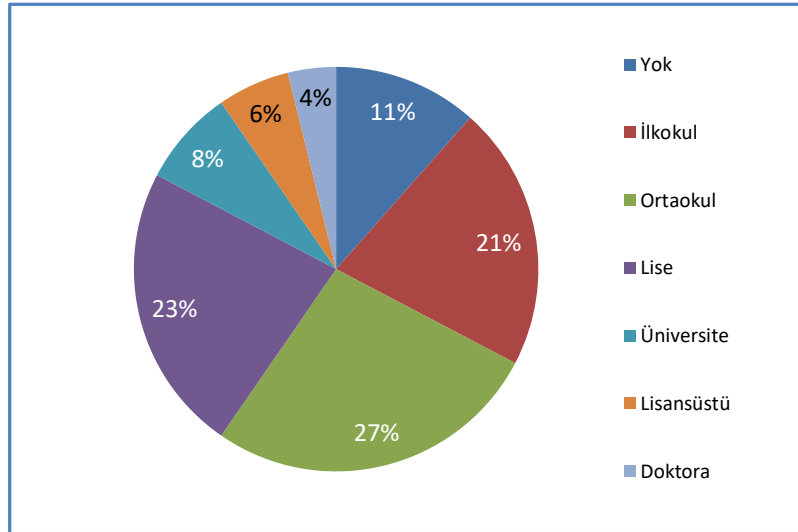
Verileri sınıflandırınız ve uygun bir grafik ile gösteriniz.

Yok	Yok	Yok	Ortaokul	İlkokul	İlkokul	İlkokul	Ortaokul
İlkokul	İlkokul	İlkokul	Lise	Ortaokul	Ortaokul	Ortaokul	Lise
Ortaokul	Ortaokul	Ortaokul	İlkokul	Lise	Lise	Lise	Ortaokul
Lise	Lise	Lise	Doktora	Üniversite	İlkokul	İlkokul	İlkokul
Üniversite	Üniversite	Üniversite	Lise	İlkokul	Ortaokul	Ortaokul	Doktora
Lisansüstü	Lisansüstü	Lisansüstü	Lise	Ortaokul	Lise	Ortaokul	Lise
Yok	Yok	Yok	Ortaokul				

4

Sınıflar oluşturulur. Her kategori bir sınıf olarak ele alınır.

Sınıflar	Frekans	Oransal Frekans(%f)
Yok	6	0,115
İlkokul	11	0,212
Ortaokul	14	0,269
Lise	12	0,231
Üniversite	4	0,077
Lisansüstü	3	0,058
Doktora	2	0,038
Toplam	52	1



Değerlendirme Soruları

Question 6 of 10 ▾ ▶

Point Value: 10 | Total Points: 0 out of 100

🕒 14:21

Yandaki Resimde verilen Frekans Dağılım Tablosuna göre verilerin % kaç 9,5 ile 10,9 aralığında yer alır?

k	Sınıflar		Frekans
	Alt Sınır	Üst sınır	
	6,5	7,9	14
	8	9,4	13
	9,5	10,9	10
	11	12,4	2
	12,5	13,9	1
	Toplam		

- %25
- %100
- %15
- %30
- %20

Submit All

Previous

Next