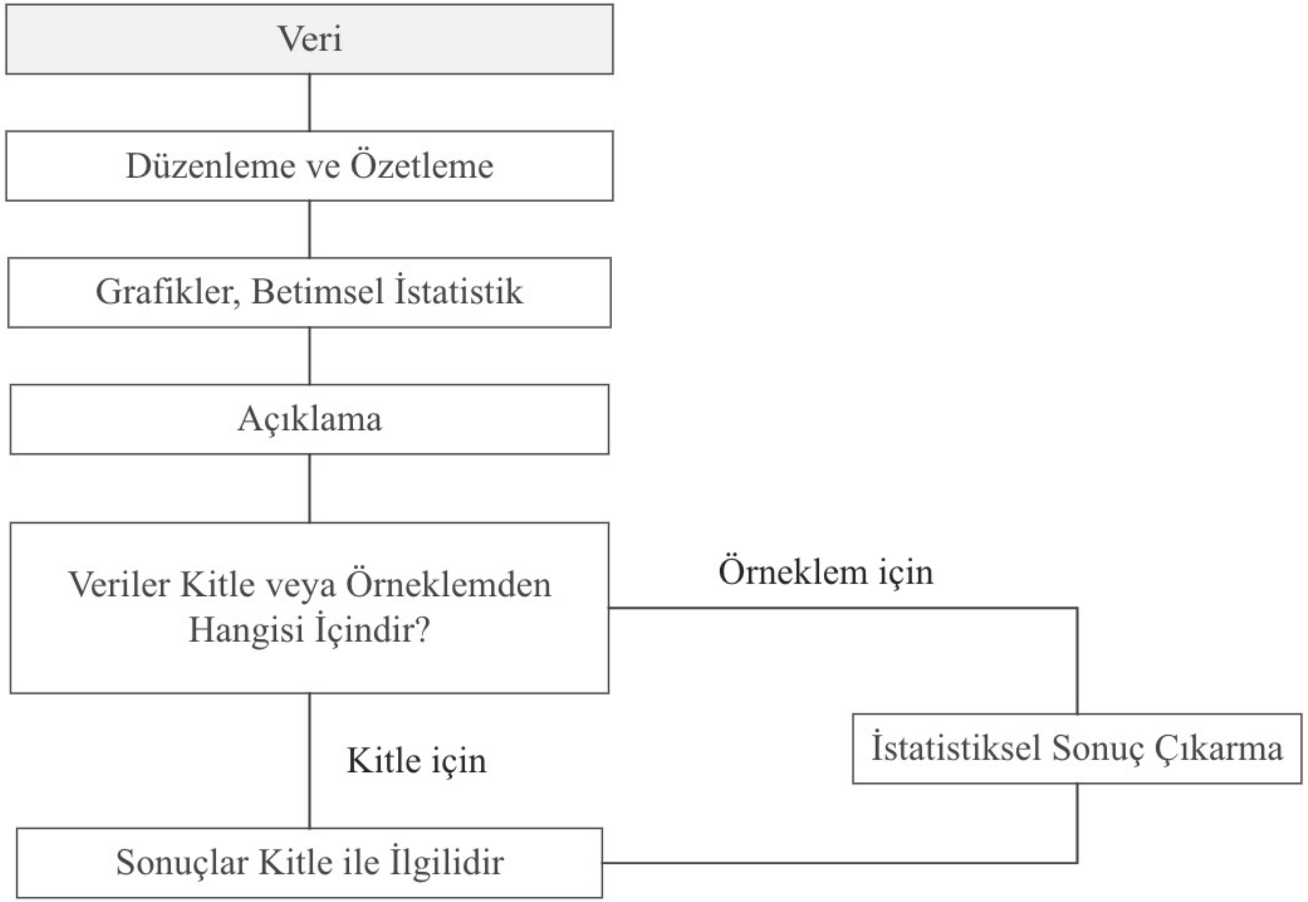
T.C.
KARADENİZ TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ

Mühendislik İstatistiği-II #Veri Düzenlenmesi ve Analizi

DR. ÖĞR. ÜYESİ FATMA BETÜL YENİ

İstatistiksel Analiz



Veri Nedir?

Belirli amaçlar için toplanan bilgilere **veri** (data, bilgi) denir.

Veriler pek çok yol ile toplanabilir:

- Yayınlanmış kaynaklardan,
- Tasarlanmış bir deneyden,
- Anket sonuçlarından,
- Gözlem sonuçlarından,
- Görüşme yoluyla,
- Bilgisayar aracılığı ile…

Veri Toplanması (Derleme)

İstatistiksel araştırma yapılırken, belli bir topluluğa ya da topluluğu oluşturan birimlere ilişkin bilgi toplamaya **derleme (veri toplanması)** denir.

Verilerin toplanması iki yönteme göre yapılmaktadır:

- Tam sayım
- Örnekleme

Verilerin Düzenlenmesi

Verilerin organizasyonu, özetlenmesi ve istatistiksel analizi için yöntemler düşünüldüğünde iki tip verinin varlığından söz edilebilir:

- Sayısal (nicel) veri
- Kategorik (nitel) veri

Birimlerin farklı değerler alabildiği nitelik veya niceliklerine **değişken** denir.

Ölçü birimi bakımından	Değer aralığı bakımından
→ Nicel değişken	→ Kesikli değişken
→ Nitel değişken	→ Sürekli değişken

Verilerin Düzenlenmesi

Cinsiyet	Yaş	Boy	İngilizce Bilgisi
Kadın	19	1.66	Çok iyi
Erkek	21	1.75	İyi
Erkek	20	1.83	İyi
Kadın	19	1.69	Kötü
Erkek	19	1.78	Çok kötü
Kadın	21	1.72	İyi
Kadın	20	1.54	Çok iyi

Frekans Dağılımı

- Derlenen veri üzerinde herhangi bir işlem yapılmamışsa, bunlara "ham veri" ya da sınıflandırılmamış (gruplandırılmamış) veri denir.
- Birim sayısı az olan yığınların çeşitli özellikleri ham veriye dayanarak kolaylıkla belirtilir.
- Yığın çok sayıda birimden oluşuyorsa, bunları sınıflandırmakla yığının çeşitli özelliklerini belirlemek kolaylaşacaktır.
- Sınıflandırmanın en doğru yolu frekans tablosudur. Burada, gözlenen veri sınıflara ayrılır. Sonuçlanan tablo her bir sınıftaki gözlem sayısını verecektir.

Frekans Dağılımı

Meslek	Gözlem Sayısı
Öğretmen	26
İşçi	21
Mühendis	20
Sağlık Personeli	28
Mimar	5
	Σ = 140

Boy (X) [cm]	Sınıf Orta Değeri (X_j) [cm]	Öğrenci Sayısı (Frekans veya Sıklık)
155 ≤ X < 160	(155+160)/2=157,5	12
160 ≤ X < 165	162,5	20
165 ≤ X < 170	167,5	36
170 ≤ X < 175	172,5	24
175 ≤ X < 180	177,5	18
180 ≤ X < 185	182,5	10
		Σ = 120

Frekans Dağılımı

	Değişken Değerleri	Birim Sayısı (F_j)	
Sınıflar	$X_{1a} \leq X \leq X_{1ü}$	F_1	Çokluk (Frekans)
	$X_{2a} \leq X \leq X_{2ü}$	F_2	
	. . . .	. . .	
	$X_{ka} \leq X \leq X_{kü}$	F_k	

n : Toplam birim (gözlem) sayısı

L : En büyük değer

S : En küçük değer

R : Değişim genişliği (L-S)

k : Sınıf sayısı

j : Sınıf sıra no indisi (j=1,...,k)

X_{ja} : j. sınıfın alt değeri

$X_{jü}$: j. sınıfın üst değeri $\left(\bar{X}_j = \frac{X_{ja}+X_{jü}}{2} \right)$

h_j : j. sınıfın genişliği

Frekans Dağılımı – Adımları

- **1. Adım:** Gözlemlerin sayısı belirlenir. n = Gözlem sayısını gösterir ($50 \le n$ tercih edilir)
- **2. Adım:** En büyük değer(L) ve en küçük değer (S) bulunur. (L-S) farkı hesaplanır. R ile gösterilen bu farka değişim genişliği (range) denir.
- **3. Adım:** $\sqrt{n} \le k$ olacak şekilde sınıf sayısı k belirlenir.

Kesin bir yaklaşım olmamakla birlikte sınıf sayısı için $k = 1+3,3*\log n$ formülü de kullanılabilir.

Genel olarak $5 \le k \le 20$ seçilmesi tavsiye edilir.

Örneklem Büyüklüğü	Sınıf sayısı
50'den az	5 – 7
51 – 100 arası	7 – 8
101 – 500 arası	8 – 10
501 – 1000 arası	10 – 11
1000 – 5000 arası	11 – 14
5000'den çok	14 – 20

Frekans Dağılımı – Adımları

- **4. Adım:** $R/k \leq h$ olacak şekilde h sınıf genişliği belirlenir. Başka bir sınıflamaya gerek duyulmadıkça sınıf genişliklerini eşit alma yoluna gidilir.

- **5. Adım:** Sınıfların alt (X_{ja}) ve üst ($X_{jü}$) limitleri belirlenir. Küçük gözlem değerine eşit ya da daha küçük bir değer ilk sınıfın alt limiti (X_{1a}) olarak seçilir. Bu değere ardışık olarak sınıf genişliği h'nin eklenmesiyle diğer sınıfların alt limitleri bulunur. Verinin sürekli ya da kesikli olmasına göre ilk sınıftan başlayarak ($X_{jü}$) üst limitleri de aynı yolla bulunur.

- **6. Adım:** Sınıf frekansı belirlenir. Her sınıf için (sınıf limitleri dahil) o sınıfa düşen gözlem sayısı (sınıf frekansı) saptanır. Frekanslar F_1, F_2, … , F_k ile gösterilir.

Frekans Dağılımı – Adımları

- **7. Adım:** Sınıf limitlerinin ortalaması alınarak her sınıf için sınıf orta noktaları $(\overline{X}_j)$ bulunur.
- **8. Adım:** Ardışık olarak frekanslar F_1 , $F_1 + F_2$, $F_1 + F_2 + F_3$ ... şeklinde toplanarak eklemeli frekans sütunu oluşturulur.

! Sınıflar oluşturulurken her verinin tek bir sınıfa girdiğinden emin olmak gerekir, yanı sınıflar ayrık olmalıdır.

~~4,00 - 4,50~~	4,00 - 4,49
~~4,50 - 5,00~~	4,50 - 5,00

! Sınıflar arası boşluk olmamalıdır. Bazı durumlarda sınıflar birleştirilebilir

Frekans Dağılımı – Örnek

Örnek: Büyük bir şirketin 100 satış elemanı vardır. Her bir satış elemanı tarafından gerçekleştirilen aylık satışlar (*bin lira) aşağıda verilmiştir. Bu veriler için frekans tablosu düzenlenmesi istenmektedir.

23	16	14	20	27	19	14	8	19	27
19	17	17	16	17	22	21	0	9	3
26	14	9	11	14	20	14	6	11	12
11	17	13	19	17	7	20	9	13	20
20	17	20	16	16	10	16	10	19	13
11	24	21	27	5	15	15	14	13	25
17	20	8	16	17	14	9	16	8	16
16	16	14	22	13	7	8	13	5	13
14	27	19	16	20	9	16	19	14	29
16	15	9	17	8	18	14	18	13	10

Frekans Dağılımı – Örnek

- n= 100
- L (en büyük değer)= 29
- S (en küçük değer)= 0
- R = L – S = 29
- $\sqrt{n} \leq$ k olmak üzere $\sqrt{100} \leq$ k ise k=10 alınabilir.
- R/k $\leq$ h olmak üzere 29/10 = 2,9 ise h=3 alınabilir.
- X_{1a}=0 olmak üzere X_{2a}; 0 + 3 = 3 olarak bulunur, bu şekilde her sınıfın alt-üst değerleri belirlenir.
- Son olarak her sınıfın frekans değeri ve birikimli frekans değeri bulunur.

Frekans Dağılımı – Örnek

Sınıf No	Sınıf Limitleri	Sınıf Orta Noktası	Gözlem Sayısı (Frekans)	Eklemeli Frekans
1	0 - 2	(0 + 2) / 2 = 1	1	1
2	3 - 5	(3 + 5) / 2 = 4	3	4
3	6 - 8	(6 + 8) / 2 = 7	8	12
4	9 - 11	(9 + 11) / 2 = 10	13	25
5	12 - 14	(12 + 14) / 2 = 13	20	45
6	15 - 17	(15 + 17) / 2 = 16	25	70
7	18 - 20	(18 + 20) / 2 = 19	17	87
8	21 - 23	(21 + 23) / 2 = 22	5	92
9	24 - 26	(24 + 26) / 2 = 25	3	95
10	27 - 29	(27 + 29) / 2 = 28	5	100

İstatistiksel Verilerin Kullanıma Sunumu

İstatistiksel verilerin görselleştirilmesi tablo ya da grafikler ile yapılabilir.

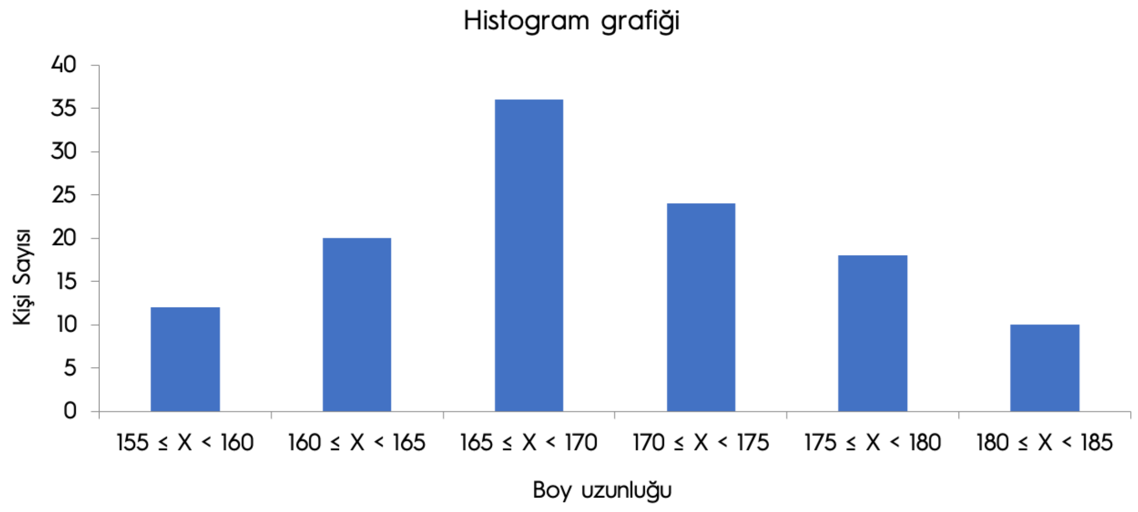
- Tablo: Bir istatistiksel tabloda;
  - Tablo adı
  - Tablo başlığı
  - Ön sütun
  - Gövde kesinlikle bulunmalıdır.

- Grafik: Sayısal bilgilerin değerlendirilmesinde yararlanılan geometrik şekillerdir. Grafikler söz konusu verilerin görsel olarak incelenmesini sağlar.

Histogram

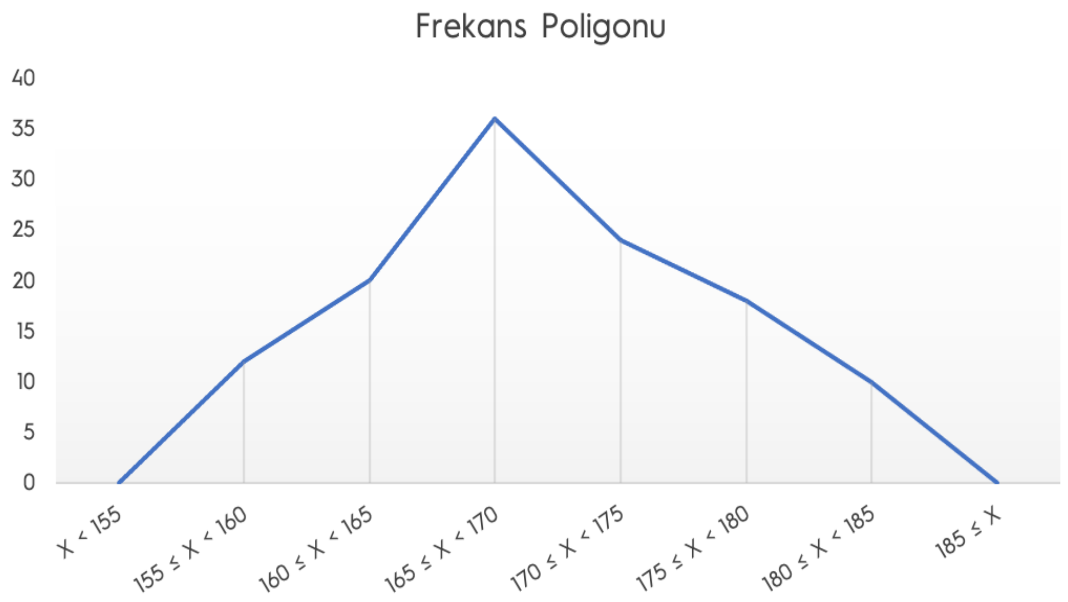
Yatay eksen sınıfları; düşey eksen çokluğu veya frekansı gösterir. Sınıflar için bir birim genişlik seçilir. Her sınıfı temsil eden çoklukları yükseklik kabul edilen dikdörtgenler sınıfların karşılığı olan aralıklar üzerine çizilir.

Boy (X) [cm]	Öğrenci Sayısı (Frekans)
$155 \leq X < 160$	12
$160 \leq X < 165$	20
$165 \leq X < 170$	36
$170 \leq X < 175$	24
$175 \leq X < 180$	18
$180 \leq X < 185$	10



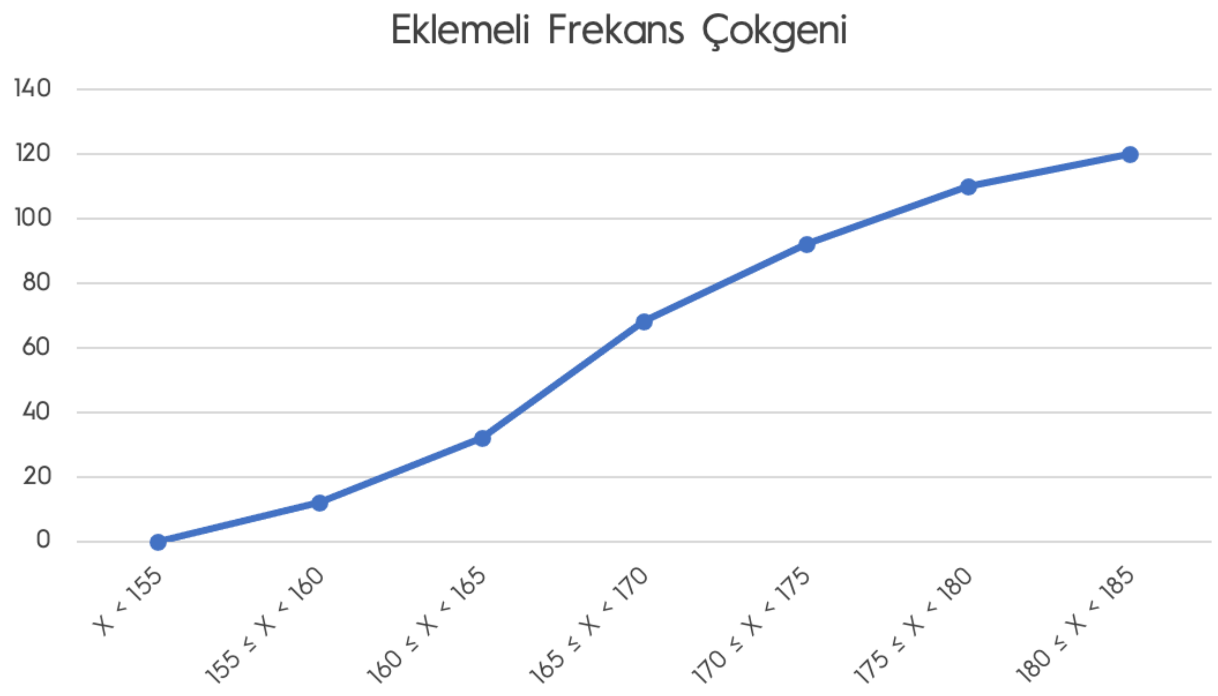
Poligon (Frekans Çokgeni)

Histogramın düz doğru parçalarıyla gösterim şeklidir. Her sınıfı temsil eden dikdörtgenin üst orta noktalarının doğrularla birleştirilmesiyle oluşturulur.



Birikimli Poligon (Frekans Çokgeni)

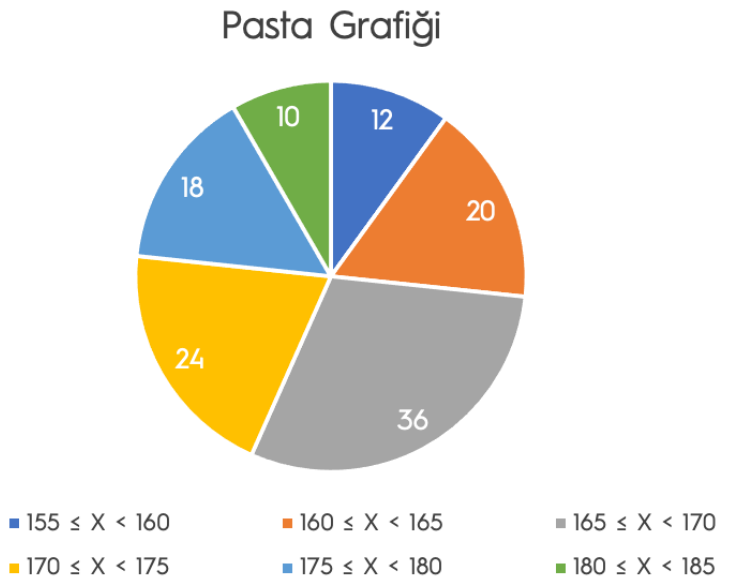
Eklemeli frekans değerlerinin kullanılması ile elde edilir.



Dairesel Grafik

Bütünü oluşturan parçaları göstermek için pasta veya daire grafikleri de kullanılmaktadır. Daire veya pasta, toplamı yansıtır; göbekten kesilmiş parçalar (pasta dilimleri) ise toplam içindeki payları gösterir. Her parçanın alanı frekansla orantılı olarak çizilir.

Boy aralığı	Frekans	Açı
155 ≤ X < 160	12	(12x360)/120=36
160 ≤ X < 165	20	60
165 ≤ X < 170	36	108
170 ≤ X < 175	24	72
175 ≤ X < 180	18	54
180 ≤ X < 185	10	30

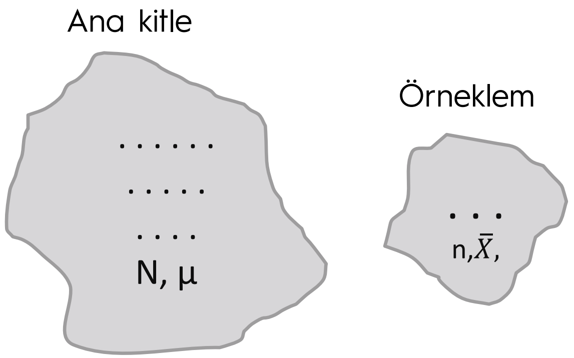


Merkezi Dağılım Ölçüleri

- Merkezi Eğilim Ölçüleri (MEÖ), belli bir değişkene veya özelliğe ilişkin ölçme sonuçlarının veya gözlem değerlerinin hangi değer etrafında toplandığını gösteren ölçülerdir.
- Sınıfların veya grupların ölçülecek özellik bakımından yorumlanmasını kolaylaştırırlar.
- Başlıca merkezi eğim ölçüleri;
  - ✓ Aritmetik ortalama
  - ✓ Ağırlıklı ortalama
  - ✓ Geometrik ortalama
  - ✓ Harmonik ortalama
  - ✓ Medyan (ortanca)
  - ✓ Mod (tepe değer)

Aritmetik Ortalama

İstatistikte en çok kullanılan ortalama türüdür. Gözlem değerleri toplamının gözlem sayısına bölünmesiyle elde edilir.



	Birim Sayısı	Aritmetik Ortalama
Ana kitle	N	μ
Örneklem	n	$\bar{X}$

- Ham verilerde veya az sayıdaki verilerde: $x_1, x_2, \dots, x_n$ gözlem değerleri olmak üzere;

$$\bar{X} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$$

Aritmetik Ortalama

- Sınıflandırılmış verilerde aritmetik ortalama hesaplanırken;

x_j	F_j
x_1	F_1
x_2	F_2
…	…
x_k	F_k

$$n = \sum F_j$$

- Bir değişkenin aldığı değerler $x_1, x_2, \dots, x_n$ ve bu değerlerinin her birinin tekrarlanma sıklığı sırasıyla $F_1, F_2, \dots, F_n$ olmak üzere;

$$\overline{X} = \frac{F_1x_1 + F_2x_2 + \cdots + F_kx_k}{F_1 + F_2 + \cdots + F_k} = \frac{\sum_{j=1}^{k} F_j \, . \, x_j}{\sum_{j=1}^{k} F_j}$$

Aritmetik Ortalama

- Gruplanmış verilerde aritmetik ortalama hesaplanırken;

Gözlem Değerleri	Frekans	Sınıf Orta Değeri: $\bar{X}_j$
$X_{1a} \leq X \leq X_{1ü}$	F_1	$\bar{X}_1$
$X_{2a} \leq X \leq X_{2ü}$	F_2	$\bar{X}_2$
. . .	. . .	. . .
$X_{ka} \leq X \leq X_{kü}$	F_k	$\bar{X}_k$

- $\bar{X}_j = \frac{X_{ja}+X_{jü}}{2}$ j. Sınıfın orta değeri olmak üzere verilerin aritmetik ortalaması;

$$\bar{X} = \frac{\sum_{j=1}^{k} F_j \, . \, \bar{X}_j}{\sum_{j=1}^{k} F_j}$$

Aritmetik Ortalama - Örnek

Örnek: Bir üniversite yerleşkesinde 80 öğrenci ile yapılan ankette günlük harcadıkları miktar TL olarak aşağıdaki gibi bulunmuştur. Bu üniversitedeki öğrencilerin günlük ortalama harcama miktarı nedir?

Günlük Harcama x_j [TL]	Öğrenci Sayısı F_j	$x_j . F_j$
20	10	200
25	18	450
30	25	750
35	16	560
40	11	440
TOPLAM	ΣF_j= 80	$\Sigma F_j x_j$=2400

$$\overline{\mathrm{X}} = \frac{\sum_{j=1}^{k} \mathrm{F_j} \, . \, \mathrm{x_j}}{\sum_{j=1}^{k} \mathrm{F_j}} = \frac{2400}{80} = 30 \text{ TL}$$

Aritmetik Ortalama - Örnek

Örnek: Bir yöredeki çiftçi ailelerinin sahip oldukları araziye göre dağılımları aşağıdaki gibidir. Aile başına düşen ortalama arazi miktarını bulunuz.

Sahip oldukları Arazi [dekar]	Çiftçi Aile Sayısı: F_j	Sınıf Orta Değeri: $\bar{X}_j$	$F_j . \bar{X}_j$
0 ≤ X < 10	2	5	10
10 ≤ X < 20	6	15	90
20 ≤ X < 40	12	30	360
40 ≤ X < 60	29	50	1450
60 ≤ X < 80	21	70	1470
80 ≤ X < 100	6	90	540
100 ≤ X < 140	4	120	480
Toplam	**ΣF_j= 80**		$\Sigma F_j x_j$=4400

$$\bar{X} = \frac{\sum_{j=1}^{k} F_j . \bar{X}_j}{\sum_{j=1}^{k} F_j} = \frac{4400}{80} = 55 \text{ dekar}$$

Aritmetik Ortalamanın Özellikleri

1. En iyi bilinen ve en çok kullanılan merkezi eğilim ölçüsüdür.
2. Tüm gözlem değerleri dikkate alınarak hesaplanır. Bu nedenle hassas bir MEÖ'dür.
3. Uç değerlerden oldukça fazla etkilenir.

- Bir sınıfın bir dersten aldığı notlar şöyle olsun:

  23, 24, 26, 25, 25, 26, 26, 30, 30, 32, 90, 95

  Bu durumda; $\overline{\mathrm{X}} = \frac{\sum \mathrm{X_j}}{\mathrm{n}} = \frac{452}{12} = 37{,}6$ olur.

  Uç değeri çıkarırsak = 90 + 95 = 185;

  $\overline{\mathrm{X}} = \frac{\sum \mathrm{X_j}}{\mathrm{n}} = \frac{452-185}{10} = 26{,}7$ olarak hesaplanır.

Ağırlıklı Ortalama ($\bar{X}_w$)

Bir veri kümesinde yer alan gözlem değerleri $x_1, x_2, \dots, x_n$ ve bu gözlem değerlerinin ağırlıkları da $w_1, w_2, \dots, w_n$ olduğunda bu değerlerin ortalamasına ağırlıklı ortalama denir ve aşağıdaki gibi hesaplanır:

$$\bar{X}_w = \frac{w_1.x_1 + w_2.x_2 + \dots + w_n.x_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum_{j=1}^{n} w_j.x_j}{\sum_{j=1}^{n} w_j}$$

Ağırlıklı Ortalama ($\bar{X}_w$)

Örnek: Bir yem fabrikası; içerisinde arpa, mısır, soya fasulyesi, şeker pancarı küspesi ve diğer besin maddelerinin yer aldığı küçük baş hayvan yemi üretmektedir. Besin maddelerinin katkı oranları sırasıyla %35, %25, %15, %20 ve %5'dir. Birim satın alma fiyatları ise TL/ton olarak yine sırasıyla 150, 120, 200, 100 ve 300 'dür. Bu verilere göre fabrikada üretilen besi yeminin ortalama birim hammadde maliyeti ne olur?

- x_1=150 ; x_2=120 ; x_3=200 ; x_4=100 ve x_5=300 TL/ton
- w_1=%35 ; w_2=%25 ; w_3=%15 ; w_4=%20 ; w_5=0,05 olmak üzere;
- Birim hammadde maliyeti

$$\bar{X}_w = \frac{\sum_{j=1}^{n} w_j . x_j}{\sum_{j=1}^{n} w_j} = \frac{(0,35)150+(0,25)120+(0,15)200+(0,20)100+(0,05)300}{0,35+0,25+0,15+0,20+0,05} = 147,5 \text{ TL/ton olur.}$$

Geometrik Ortalama

Gözlem değerleri arasındaki büyüme oranını ve değişmeyi gösterir. Başlıca iki kullanım alanı vardır:

- Yüzdeler, indeksler ve oranların ortalamasının bulunması,
- Bir zaman periyodundan diğerine geçişte satış, üretim vb. gibi işletme verilerindeki (birimlerindeki) % artışın ortalamalarının bulunması.

Geometrik Ortalama

- Ham verilerde veya az sayıdaki verilerde: Değişkenler $x_1, x_2, \ldots, x_n$ şeklinde ol. üz.

$$\overline{X}_G = \sqrt[n]{x_1 . x_2 \ldots x_n} = (x_1 . x_2 \ldots x_n)^{1/n}$$

$$\overline{X}_G = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

- Sınıflandırılmış / gruplandırılmış frekans dağılımında:

$$\overline{X}_G = \sqrt[n]{x_1^{F_1} . x_2^{F_2} . \ldots x_k^{F_k}} \quad \text{veya}$$

$$\overline{X}_G = \sqrt[n]{\overline{X}_1^{F_1} . \overline{X}_2^{F_2} \ldots \overline{X}_k^{F_k}}$$

Geometrik Ortalamanın Özellikleri

- Serinin (dizinin) veya gözlem değerlerinin her teriminden etkilenir. Ancak uç değerlerin ağırlığı aritmetik ortalamadan daha azdır.
- Sadece pozitif değerlerin ortalamasında kullanılır. Yani gözlem değerleri negatif ve sıfır olamaz.
- Fiyat, nüfus, ulusal gelir vb. değişkenlerin ortalamalarının bulunmasında yaygın olarak kullanılır.
- Geometrik ortalama daima aritmetik ortalamadan küçük çıkar.
- Cebirsel işlemlere uygundur.

Geometrik Ortalama - Örnek

Örnek: Endüstri Mühendisliği öğrencilerinin alt dönemden aldıkları ders sayıları aşağıdaki gibi gruplandırılmıştır. Her bir öğrenci ortalama kaç dersi alt dönemden almıştır?

Ders Sayısı	**Öğr. Sayısı (F_J)**	$\bar{X}_J$
1 - 3	18	2
4 - 6	12	5
7 - 9	4	8
10 - 12	2	11

$$\bar{X}_G = \sqrt[36]{2^{18} \cdot 5^{12} \cdot 8^4 \cdot 11^2} = 3,48$$

Harmonik Ortalama ($\bar{X}_H$):

- Hız, fiyat, verimlilik, zaman oranları gibi değişkenlerin ortalamalarının hesabında kullanılır.
- Bir dizinin, dizi terimlerinin terslerinin aritmetik ortalaması alınarak bulunur.

$x_1, x_2, \dots, x_n$ gözlem değerleri olmak üzere;

- Ham verilerde veya az sayıdaki verilerde; $\bar{X}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$
- Sınıflandırılmış/gruplandırılmış frekans dağılımında;

$$\bar{X}_H = \frac{\sum_{j=1}^{k} F_j}{\sum_{j=1}^{k} \frac{F_j}{x_j}} \quad \text{veya} \quad \bar{X}_H = \frac{\sum_{j=1}^{k} F_j}{\sum_{j=1}^{k} \frac{F_j}{\bar{X}_j}}$$

Harmonik Ortalama

- Serinin her teriminden etkilenir. Ancak uç değerlerden en az etkilenen ortalama türüdür.
- Pozitif değerli seriler için hesaplanır (negatif ve sıfır değerli veri olmamalıdır).
- Diğer ortalamalardan daima daha küçük çıkar.

Harmonik Ortalama

Örnek: Şehirler arası çalışan bir hızlı tren, gittiği mesafenin ilk üçte birinde 300 km/h, ikinci üçte birinde 450 km/h ve son üçte birinde 360 km/h hız yapmıştır. Bu trenin ortalama hızı nedir?

x_1= 300; x_2=450; x_3=360 olmak üzere

$$\overline{X}_H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{3}{\frac{1}{300}+\frac{1}{450}+\frac{1}{360}} = 360 \text{ km/h}$$

Ortanca (Medyan – Orta Değer)

- Bir değişkene ilişkin gözlem değerlerinin küçükten büyüğe doğru sıralanmasında orta sıradaki değere denir (diziyi %50 bir tarafa, %50'sini diğer tarafta olacak şekilde ayıran değerdir).
- Medyan hassas bir merkezi eğilim ölçüsü (ortalama) değildir. Uç değerlerden etkilenmez.
- Çift sayıda gözlem varsa orta değer (medyan), iki merkezi değerin aritmetik ortalamasıdır.

Ortanca (Medyan – Orta Değer)

$x_1, x_2, \dots, x_n$ gözlem değerleri olmak üzere bunlar $x_1 < x_2 < \cdots, < x_n$ şeklinde sıralanırsa;

- Ham verilerde veya az sayıdaki verilerde:

$$M = \begin{cases} X_{\frac{n+1}{2}} \quad , & \text{n tek ise} \\ \dfrac{X_{\frac{n}{2}} + X_{\frac{n+2}{2}}}{2}, & \text{n çift ise} \end{cases}$$

Ortanca (Medyan – Orta Değer)

Örnek: 2, 3, 2, 4, 4, 6, 6, 5, 8, 8, 9 sayıları için orta değeri bulalım.

- Önce sayıları küçükten büyüğe sıralayalım;

  2, 2, 3, 4, 4, 5, 6, 6, 8, 8, 9

  n=11 yani tek sayı olduğu için → $M = x_{11+1/2} = x_6 = 5$ olarak bulunur.

Örnek: 1, 2, 3, 3, 5, 5, 5, 6, 7, 7, 7, 8, 9, 9 biçiminde sıralanmış veriler için orta değeri bulalım.

- Önce sayıları küçükten büyüğe sıralayalım;

  1, 2, 3, 3, 5, 5, 5, 6, 7, 7, 7, 8, 9, 9

  n=14 yani çift sayı olduğu için → $M = (x_{14/2} + x_{14+2/2}) / 2 = (5 + 6) / 2 = 5{,}5$ olarak bulunur.

Ortanca (Medyan – Orta Değer)

- Sınıflandırılmış verilerde; öncelikle birikimli frekans değerleri bulunur. Ardından n çift sayı ise n/2; tek sayısı ise (n+1)/2 hesaplanarak medyanın birikimli frekans değerleri içinde düştüğü aralık bulunur. O sıraya ait gözlem değeri medyan'ı verir.

Gözlem Değeri	Gözlem Sayısı	Birikimli Frekans
x_1	F_1	F_1
$\boxed{x_2}$	F_2	$\boxed{F_1+F_2}$
. . . .	. . . .	. . . .
x_k	F_k	$F_1+F_2+\ldots+F_k$

$\Sigma F_j = n$

n/2 değeri F_1+F_2 aralığına düşüyorsa medyan değeri x_2 olarak bulunuz.

Ortanca (Medyan – Orta Değer)

Örnek: Bir üniversite yerleşkesinde 80 öğrenci ile yapılan ankette günlük harcadıkları para miktarı TL olarak aşağıdaki gibi bulunmuştur. Bu üniversitedeki öğrencilerin günlük ortalama harcama miktarının medyanını bulalım.

Günlük Harcama x_j [TL]	Öğrenci Sayısı F_j	Birikimli Toplam	Aralık
20	10	10	1 - 10
25	18	28	11 - 28
$\boxed{30}$	25	53	$\boxed{29 - 53}$
35	16	69	54 - 69
40	11	80	70 - 80

$\Sigma F_j = 80$

n=80 olduğuna göre n/2=40 'ın düştüğü aralık 3. sınıfta 29-53 aralığıdır.

Buna göre medyan=30'dur.

Ortanca (Medyan – Orta Değer)

- Gruplandırılmış verilerde;

  n: Gözlem sayısı

  L: Medyan sınıfının alt sınır değeri

  h: Medyan sınıfın genişliği

  F_{medyan}: Medyan sınıfın frekansı

  F_b: Medyan sınıfına kadar olan frekansların birikimli toplam değeri olmak üzere;

  $$M = L + \frac{\frac{n}{2} - F_b}{F_{Medyan}} \cdot h$$ ile hesaplanır.

- $f_1+f_2+\ldots+f_i \geq n/2$ (n çift ise) ya da $f_1+f_2+\ldots+f_i \geq n+1/2$ (n tek ise) olan ilk sınıfa medyan sınıfı denir.

Ortanca (Medyan – Orta Değer)

Örnek: Bir yöredeki çiftçi ailelerinin sahip oldukları araziye göre dağılımları aşağıdaki gibidir. Aile başına düşen arazi miktarının medyanını bulalım.

Sahip oldukları Arazi [dekar]	Çiftçi Aile Sayısı: F_j	Birikimli Toplam	Aralık
0 ≤ X < 10	3	3	1 – 3
10 ≤ X < 20	5	8	4 – 8
20 ≤ X < 40	12	20	9 – 20
40 ≤ X < 60	24	44	21 – 44
60 ≤ X < 80	20	64	45 – 64
80 ≤ X < 100	10	74	65 – 74
100 ≤ X	7	81	75 – 81

n = 81 ve n+1/2=41

L = 40

h = 20

F_{Medyan}= 24

F_b = 20 olmak üzere

$$M = 40 + \frac{40{,}5-20}{24} \cdot 20 = 57$$

Tepe Değer (Mod)

- Gözlem değerleri arsında en fazla tekrarlanan değerdir.
- Her değer yalnız bir kez elde edilmişse mod yoktur.

Örnek:	41 41 42 42 43 43 43 44 44	MOD = 43
	41 41 42 42 43 43 44 44	MOD = Yok
	41 41 42 43 44 45 45 46	MOD = 41 ve 45

Tepe Değer (Mod)

- Gruplanmış verilerde öncelikle frekansı en büyük olan sınıf MOD sınıfı olarak belirlenir. Ardından aşağıdaki formülle MOD değeri hesaplanır;

$$\mathrm{MOD} = \mathrm{L} + \frac{\mathrm{F_m} - \mathrm{F_{m-1}}}{(\mathrm{F_m} - \mathrm{F_{m-1}}) + (\mathrm{F_m} - \mathrm{F_{m+1}})} \cdot \mathrm{h}$$

L: Mod sınıfının alt sınır değeri

$\mathrm{F_m}$: Mod sınıfının frekansı

$\mathrm{F_{m-1}}$: Mod sınıfından bir önceki sınıfın frekansı

$\mathrm{F_{m+1}}$: Mod sınıfından bir sonraki sınıfın frekansı

h: Mod sınıfının sınıf genişliği

Tepe Değer (Mod)

Örnek: Bir yöredeki çiftçi ailelerinin sahip oldukları araziye göre dağılımları aşağıdaki gibidir. Aile başına düşen arazi miktarının modunu bulunuz

Sahip oldukları Arazi [dekar]	Çiftçi Aile Sayısı: F_j
$0 \leq X < 10$	3
$10 \leq X < 20$	5
$20 \leq X < 40$	12
$40 \leq X < 60$	24
$60 \leq X < 80$	20
$80 \leq X < 100$	10
$100 \leq X$	7

$L = 40$

$F_m = 24$

$F_{m-1} = 12$

$F_{m+1} = 20$

$h = 20$ olmak üzere;

$$\text{MOD} = 40 + \frac{24-12}{(24-12)+(24-20)} \cdot 20 = 55$$ dekar'dır.

Mod'un Özellikleri

- Ölçülemeyen özelliklerin tek eğilim ölçüsüdür.
- Hassas olmayan bir eğilim ölçüsüdür.
- Dizideki veri sayısı az ise hesaplanması uygun değildir.
- Yoğunluğun en fazla olduğu noktada bulunmasından dolayı verilerin önemli bir bölümünün gerçek değerini gösterir.

Ortalama-Mod-Medyan İlişkisi

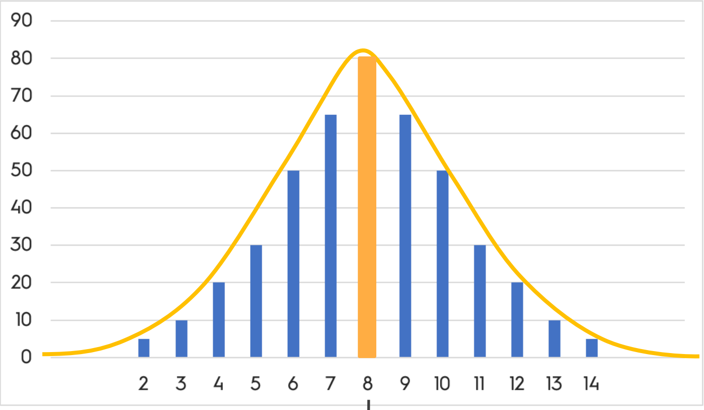
Merkezi Eğilim Ölçüleri

Aritmetik Ortalama	Medyan (Ortanca)	Mod
2	2	$\boxed{2}$
4	2	$\boxed{2}$
2	2	$\boxed{2}$
2	4	4
5	5	5
5	$\boxed{5}$	5
7	7	7
8	8	8
9	9	9
10	10	10
12	12	12

66/11=6

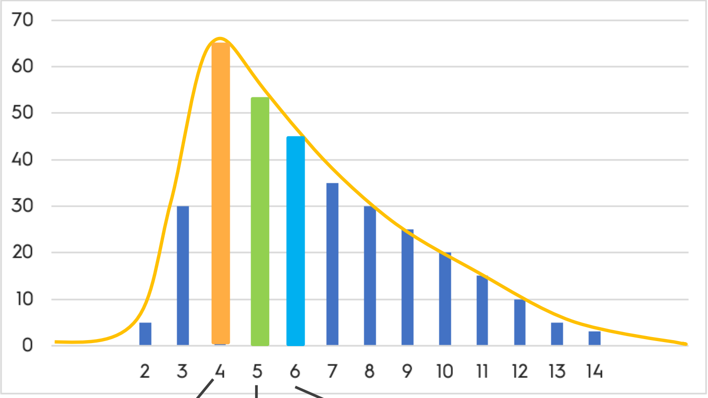
Ortalama-Mod-Medyan İlişkisi

Simetrik dağılım



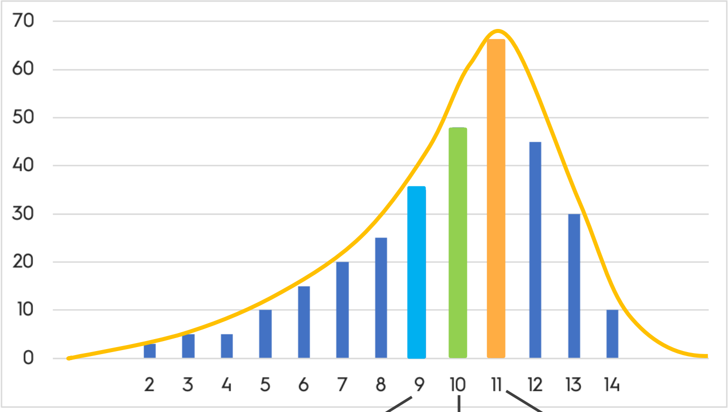
Aritmetik ortalama = Mod = Medyan

Sağa çarpık dağılım



Mod < Medyan < Aritmetik ortalama

Sola çarpık dağılım



Aritmetik ortalama < Medyan < Mod