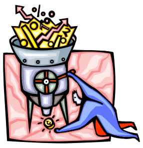
DEĞİŞKENLİK ÖLÇÜLERİ



Dr. Mehmet AKSARAYLI
D.E.Ü. İ.İ.B.F.
EKONOMETRİ BÖLÜMÜ
mehmet.aksarayli@deu.edu.tr

Bölümün Amaçları

Bu Bölümü tamamladıktan sonra neleri yapabileceksiniz:

- Box ve Whisker grafiğini okuma ve yorumlama,
- Değişkenlik kavramını anlama ve verilerin değişkenliğini yorumlama,
- Değim Aralığı, varyans, standart sapma, değişim (varyasyon) katsayısını tanıma ve kullanma,
- Verilerin çarpıklığını hesaplama ve yorumlamak.

Tanımlayıcı İstatistikler

Yer Ölçüleri (Merkezi Eğilim Ölçüleri)
- Duyarlı Ortalamalar
  - Aritmetik ort.
  - Tartılı Aritmetik
  - Geometrik ort.
  - Kareli ort.
  - Harmonik ort.
- Duyarlı Olmayan Ort.
  - Mod
  - Medyan
  - Kartiller

Değişkenlik Ölçüleri
- Değişim Aralığı (Range)
- Standart sapma
- Varyans
- Mutlak sapma
- Değişkenlik katsayısı
- Kartil sapma katsayısı
- Ortalama sapma katsayısı

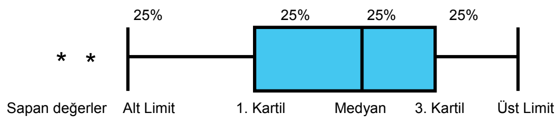
Çarpıklık Ölçüleri
- Bowley asimetri ölçüsü
- Pearson asimetri ölçüsü

Basıklık Ölçüleri

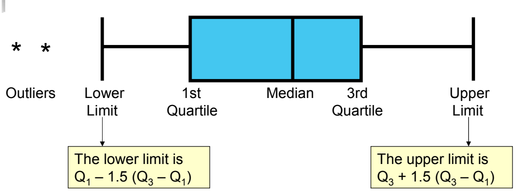
Box ve Whisker Grafiği

- Box and Whisker grafiği verileri merkezi dağılımına göre gösteren bir grafiktir.

Example:



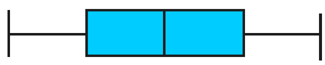
Box and Whisker Grafiğini Yapılandırma



- Merkez kutu Q_1 to Q_3 arasındadır
- Kutu ortasındaki çizgi medyandır,
- Whiskers hesaplanan limitlerde en küçük ve en büyük verileri gösterir,
- Sapan değerler grafiğin dışında gösterilmiştir.

Box ve Whisker Grafiğinin Sekli (Shape)

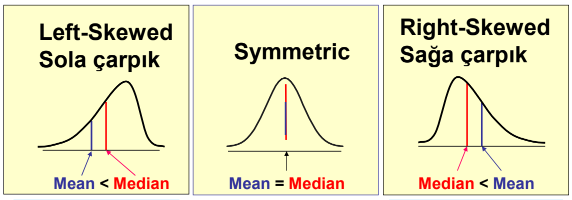
- Eğer veriler medyan etrafına simetrik yayılmışlarsa; medyan kutuyu ortalamıştır.



- (Box and Whisker grafiği hem yatay hem de dikey şekilde gösterilebilir)

Shape of a Distribution (Ortalama ve Medyana göre)

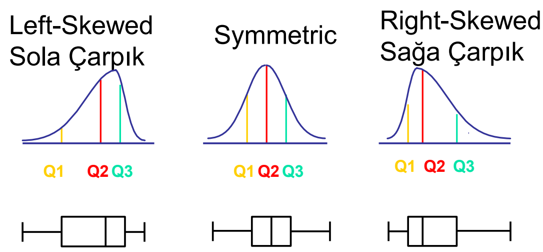
- Verilerin dağılımı nasıl tanımlanır?
- Simetriklik veya Çarpıklık



(Longer tail extends to left)
Uzun kuyruk sola doğru

(Longer tail extends to right)
Uzun kuyruk sağa doğru

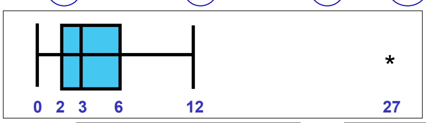
Distribution Shape and Box and Whisker Plot (Kartillere göre)



Box ve Whisker Grafiği Örneği

- Below is a Box-and-Whisker plot for the following data:

Min		Q_1			Q_2			Q_3		Max
0	2	2	2	3	3	4	5	6	11	27



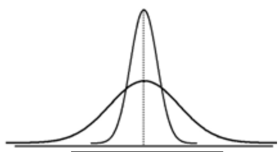
Upper limit = $Q_3 + 1.5\,(Q_3 - Q_1)$
$= 6 + 1.5\,(6 - 2) = 12$

27 is above the upper limit so is shown as an outlier

- Bu veriler sağa çarpıktır.

Değişkenlik - Varyasyon (Variation)

- Değişkenlik ölçümü; verilerin **yayılımı (spread)** veya **değişkenliği (variability)** hak



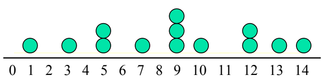
Aynı merkez, Farklı değişkenlik

Değişim Aralığı (Range)

- En basit değişkenlik ölçüsüdür.
- En büyük ve en küçük gözlemler (observations) arasındaki farktır:

D.A.=Range = $x_{maximum} - x_{minimum}$

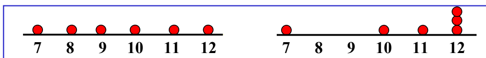
Örnek:



Range = 14 - 1 = 13

Değişim Aralığının Zayıf Yönleri

- Verilerin dağılımını göz ardı eder.



Range = 12 - 7 = 5   **Range = 12 - 7 = 5**

- Sapan değerler için hassastır.

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,5

Range = 5 - 1 = 4

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,4,120

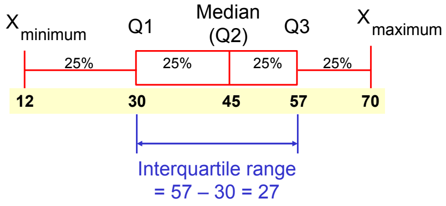
Range = 120 - 1 = 119

Kartiller Arası Değişim Aralığı (Interquartile Range)

- Bazı sapan değer problemlerine uygulanabilir.
- Yüksek ve düşük sapan değerleri dikkate almadan kalan verilerle değişim aralığını hesaplar.
- Interquartile range = 3rd quartile – 1st quartile

Kartiller Arası Değişim Aralığı Örneği:

Örnek:



Ortalama Mutlak Sapma

Ortalama Mutlak Sapma (MAD) : Bir gözlemin ortalamadan ortalama olarak ne kadar saptığının ölçüsüdür.

Basit veriler için: $$MAD = \frac{\sum |x_i - \bar{x}|}{n}$$

Gruplandırılmış veriler için: $$MAD = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

Varyans

- Ortalama mutlak sapmada kullanılan mutlak değerli ifadeler ile işlem yapmanın zor hatta bazı durumlarda imkansız olması sebebiyle yeni değişkenlik ölçüsüne ihtiyaç bulunmaktadır.
- Mutlak değer ifadesindeki zorluk aritmetik ortalamadan farkların karelerinin alınmasıyla ortadan kalkmaktadır.
- ***Veri setindeki her bir gözlem değerinin aritmetik ortalamadan farklarının karelerinin toplamının örnek hacminin bir eksiğine bölünmesinden elde edilen değişkenlik ölçüsüne örnek varyansı adı verilir.***

Varyans (Variance)

- Verilerin aritmetik ortalamadan sapmasının ortalama ölçüsüdür.
  - **Anakitle varyansı**: $$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$$
  - **Örnek(Sample) varyansı**: $$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Basit seriler İçin:

Populasyon Varyansı: $$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

μ : Populasyon Ortalaması N : Populasyon Hacmi

Örnek Varyansı : $$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Gruplanmış Seriler İçin: $$s^2 = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1}$$

Sınıflanmış Seriler İçin : $$s^2 = \frac{\sum_{i=1}^{k} f_i (m_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1}$$

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$$

ifadesi istatistikte bir çok formülde kullanılır ve ***kareler toplamı*** olarak adlandırılır.

• Matematiksel olarak hesaplama kolaylığı sağlaması açısından formüllerde kareler toplamının açılımı olan aşağıdaki eşitlik kullanılabilir.

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2=\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}$$

Basit Seriler İçin:

$$s^2=\frac{\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x\right)^2}{n}}{n-1}$$

Gruplanmış Seriler İçin:

$$s^2=\frac{\sum_{i=}^{k}f_i x_i^2-\frac{\left(\sum_{i=}^{k}f_i x_i\right)^2}{\sum_{i=1}^{k}f_i}}{\sum_{i=1}^{k}f_i-1}$$

Sınıflanmış Seriler İçin :

$$s^2=\frac{\sum_{i=}^{k}f_i m_i^2-\frac{\left(\sum_{i=}^{k}f_i m_i\right)^2}{\sum_{i=1}^{k}f_i}}{\sum_{i=1}^{k}f_i-1}$$

Standart Sapma (Standard Deviation)

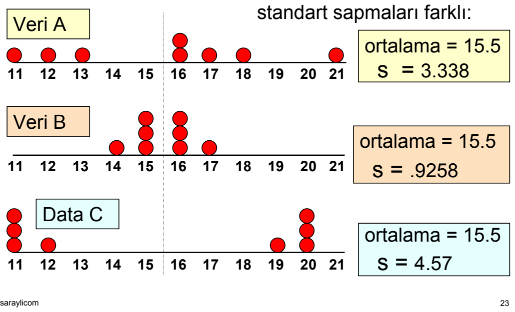
- En yaygın kullanılan değişkenik ölçüsüdür,
- Ortalama değişkenliği gösterir,
- Orijinal verilerin aynı ölçü birimine sahptir.
  - **Anakitle** standard deviation: $\sigma=\sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$
  - **Örne** standard deviation: $s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

Standart sapma: σ

- **Bir dizi ölçümün gösterdiği değişimin en güvenilir ölçüsüdür.**
- **Dağılım fazlaysa standart sapma büyük, dağılım dar alanda ise küçüktür.**

Standart Sapmaları Karşılaştırma

Ortalama aynı, fakat standart sapmaları farklı:



Örnek verileri için standart sapma hesaplama:

Örnek Veriler (X$_i$) : 10 12 14 15 17 18 18 24

n = 8 Ortalama = $\bar{x}$ = 16

$$s=\sqrt{\frac{(10-\bar{x})^2+(12-\bar{x})^2+(14-\bar{x})^2+\cdots+(24-\bar{x})^2}{n-1}}$$

$$=\sqrt{\frac{(10-16)^2+(12-16)^2+(14-16)^2+\cdots+(24-16)^2}{8-1}}$$

$$=\sqrt{\frac{130}{7}}=\boxed{4.3095}$$

Basit seriler İçin:

Populasyon Standart Sapması: $\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$

μ : Populasyon Standart Sapması N : Populasyon Hacmi

Örnek Standart Sapması : $s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$

Gruplanmış Seriler İçin: $s = \sqrt{\frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1}}$

Sınıflanmış Seriler İçin : $s = \sqrt{\frac{\sum_{i=1}^{k} f_i (m_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1}}$

Basit veriler için:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
35	-10	100
40	-5	25
45	0	0
50	5	25
55	10	100
225		250

$\bar{x} = 225/5 = 45$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{250}{5-1}} = 7.9$$

Gruplandırılmış ve sınıflandırılmış veriler için standart sapma:

x_i	f_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
35	1	-10	100	100
40	4	-5	25	100
45	5	0	0	0
50	4	5	25	100
55	1	10	100	100
Top.	15			400

$$s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i - 1}} = \sqrt{\frac{400}{14}} = 5.345$$

Sınıflandırılmış veriler için standart sapma:

Örnek: Aşağıdaki tabloda 30 günlük süre içinde bir restoranın kullandığı et miktarının dağılımı verilmiştir. Günlük kullanılan et miktarının varyansını ve standart sapmasını hesaplayınız.

Sınıflar	f_i	m_i	$f_i(m_i - \bar{x})^2$
30-36'dan az	2	33	$2(33-46,6)^2$
36-42'den az	6	39	$6(39-46,6)^2$
42-48'den az	10	45	$10(45-46,6)^2$
48-54'dan az	7	51	$7(51-46,6)^2$
54-60'dan az	4	57	$4(57-46,6)^2$
60-66'dan az	1	63	$1(63-46,6)^2$
Toplam	30		1579,2

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i} = 46{,}6\,kg.$$

$$s^2 = \frac{\sum_{i=1}^{k} f_i (m_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1} = \frac{1579{,}2}{30-1} \approx 54{,}46$$

$$s = \sqrt{s^2} = \sqrt{54{,}46} \approx 7{,}38\,kg.$$

Varyasyon - Değişkenlik Katsayısı (Coefficient of Variation)

- Göreli (relative) değişkenliği ölçer
- (%) olarak youmlanır
- İki veya daha fazla veri seti için karşılaştırmada kullanılır.
- **Örnek:** İstanbul'da ve Ankara'da yaşayan ailelerin aylık gelirlerinin değişkenliklerinin karşılaştırılması

Population

$$CV = \left(\frac{\sigma}{\mu}\right) \cdot 100\%$$

Sample

$$CV = \left(\frac{s}{\bar{x}}\right) \cdot 100\%$$

Değişkenlik Katsayısı Örneği:

- Hisse senedi A:
  - Geçen yılki ortalama fiyat = \$50
  - Standart sapma = \$5

$$CV_A = \left(\frac{s}{\bar{x}}\right) \cdot 100\% = \frac{\$5}{\$50} \cdot 100\% = 10\%$$

- Hisse senedi B:
  - Geçen yılki ortalama fiyat = \$100
  - Standart sapma = \$5

$$CV_B = \left(\frac{s}{\bar{x}}\right) \cdot 100\% = \frac{\$5}{\$100} \cdot 100\% = 5\%$$

Her iki hisse senedi de aynı standart sapmaya sahip, fakat hisse senedi B daha az göreli değişkenliğe sahiptir.

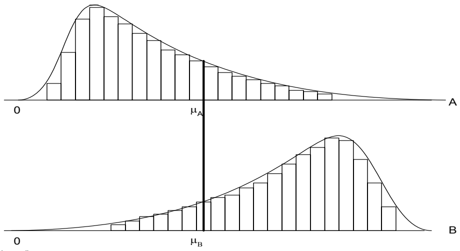
Örnek: Buca ve Alsancak için gelir dağılımıyla ilgili veriler aşağıdaki gibidir:

	Buca	Alsancak
$\overline{X} \rightarrow$	25	50
s $\rightarrow$	5	7
$C_v \rightarrow$	0.2	0.14

Yorum: Buca'daki gelir dağılımı Alsancak'takine göre daha değişkendir.

Çarpıklık (Asimetri) Ölçüleri

- Populasyonları birbirinden ayırmak için her zaman yalnızca yer ve yayılım ölçüleri yeterli olmayabilir. Aşağıda iki farklı populasyondan alınmış örnekler için oluşturulan histogramlar verilmiştir.



- Şekilden görüleceği üzere A ve B örneklerinin aynı ortalamaya ve yaklaşık olarak aynı değişkenliğe sahip olmalarına rağmen bu iki örneğin açıkça aynı populasyondan gelmediği söylenir.

- Asimetri (çarpıklık) ifadesi simetrik olmayan anlamını taşımaktadır.

- Şekillere bakıldığında frekansların A'da daha çok sol tarafta (küçük x_i değerlerinde), B'de ise daha çok sağ tarafta (büyük x_i değerlerinde), toplandığı görülmektedir.

Asimetri Ölçüleri

PEARSON ÇARPIKLIK ÖLÇÜSÜ

$$Sk_p = \frac{\bar{x} - \text{mod}}{s} \quad \text{veya} \quad Sk_p = \frac{3(\overline{X} - med)}{s}$$

$Sk_p<|1|$ --- Az çarpık dağılım

$Sk_p>|1|$ --- Çarpık dağılım

$Sk_p < 0 \rightarrow$ Negatif çarpık(Sola)

$Sk_p > 0 \rightarrow$ Pozitif Çarpık(Sağa)

$Sk_p = 0$ ise dağılış simetrik

BOWLEY ÇARPIKLIK ÖLÇÜSÜ

$$Sk_b = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$$

$Sk_b < 0 \rightarrow$ Negatif çarpık(Sola)

$Sk_b > 0 \rightarrow$ Pozitif Çarpık(Sağa)

$Sk_b = 0$ ise dağılış simetrik

Örnek: Aşağıdaki tabloda 30 günlük süre içinde bir restoranın kullandığı et miktarının dağılımından elde edilen bazı tanımlayıcı istatistikler verilmiştir. Buna göre pearson ve bowley asimetri ölçülerini hesaplayıp yorumlayınız.

Aritmetik Ort.	Mod	Medyan	Q_1	Q_3	s^2
46,6	45,4	46,2	41,5	51,9	54,46

$$Sk_p = \frac{3(\overline{X} - med)}{s} = \frac{3(46{,}6 - 46{,}2)}{\sqrt{54{,}46}} \approx 0{,}16 > 0$$

Sağa Çarpık , Pozitif Asimetri

$$Sk_p = \frac{\bar{x} - \text{mod}}{s} = \frac{46{,}6 - 45{,}4}{\sqrt{54{,}46}} \approx 0{,}16 > 0$$

Sağa Çarpık, Pozitif Asimetri

$$Sk_b = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{(51{,}9 - 46{,}2) - (46{,}2 - 41{,}5)}{51{,}9 - 41{,}5}$$

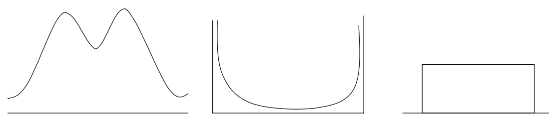
$$= \frac{1}{10{,}4} \approx 0{,}10 > 0$$

Sağa Çarpık , Pozitif Asimetri

Simetrik Dağılım — A.O = Med = Mod

Sağa çarpık dağılım — A.O > Med > Mod

Sola çarpık dağılım — A.O < Med < Mod



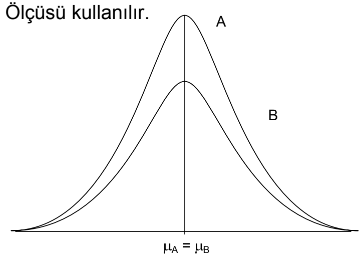
İki modlu simetrik dağılım

Modu olmayan dağılım

Tekdüzen dağılım

Basıklık Ölçüsü

Aşağıdaki A ve B dağılımlarının ortalamaları, değişkenlik ölçülerinin aynı olmasından dolayı ve hatta ikisinin de simetrik olmalarından dolayı bu iki dağılışı ayırt etmek için Basıklık Ölçüsü kullanılır.



Herhangi bir olasılık fonksiyonunun şekli ile ilgili parametrelerden bir tanesi de basıklık ölçüsüdür. **Basıklık Ölçüsü** ortalamaya göre dördüncü momentten gidilerek hesaplanır ve α_4 olarak gösterilir.

$$\alpha_4 = \frac{\mu_4}{\sigma^4}$$

Basit Seri İçin

$$\mu_4 = \frac{\sum_{i=1}^{n}(x_i-\mu)^4}{n}$$

α_4 = 3 ise Seri Normal

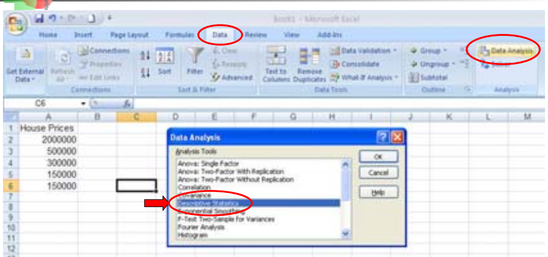
α_4 < 3 ise Seri Basık

α_4 < 3 ise Seri Sivri Ya da Yüksek

Microsoft Excel ile tanımlayıcı istatistikleri hesaplama:

- Descriptive Statistics are easy to obtain from Microsoft Excel
  - Menüden:
    Data / data analysis / descriptive statistics
  - Açılan pencereden gerekli veriler gir:

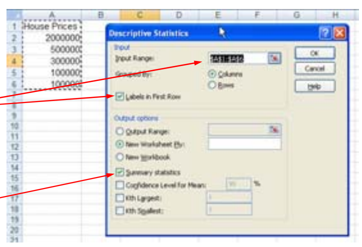
Excel Kullanımı



- Seç:

Data / data analysis / descriptive statistics

Using Excel



- Gerekli veriler gir
- “summary statistics” butonunu seç
- OK butonu…

Excel Çıktısı

	A	B
1	*House Prices*	
2		
3	Mean	600000
4	Standard Error	357770.8764
5	Median	300000
6	Mode	100000
7	Standard Deviation	800000
8	Sample Variance	6.4E+11
9	Kurtosis	4.130126953
10	Skewness	2.006835938
11	Range	1900000
12	Minimum	100000
13	Maximum	2000000
14	Sum	3000000
15	Count	5
16		

Örnekler…

Örnek 1: Gruplanmış Seriler İçin Aritmetik Ortalama

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} \qquad \sum_{i=1}^{k} f_i = n$$

f : frekans
k: grup sayısı
i = 1,2,3,……….,k

Grup	Frekans	$x_i f_i$
51	1	**51**
66	3	**198**
72	4	**288**
82	5	**410**
94	7	**658**
	$\sum f_i$ =20	**1605**

Örnek: Yandaki tabloda bir Samsung bayisindeki LCD televizyonların ekran boyutlarına göre satış miktarları verilmiştir. Frekans dağılımının aritmetik ortalamasını hesaplayınız.

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{51(1)+66(3)+....+94(7)}{1+3+4+5+7}$$

$$= \frac{1605}{20} = 80{,}25$$

Örnek 2 : Aşağıdaki tabloda 30 günlük süre içinde bir restoranın kullandığı et miktarının dağılımı verilmiştir. Günlük kullanılan et miktarının aritmetik ortalamasını hesaplayınız.

Sınıflar	f_i	m_i	$m_i f_i$
30-36'dan az	2	*33*	*66*
36-42'den az	6	*39*	*234*
42-48'den az	10	*45*	*450*
48-54'dan az	7	*51*	*357*
54-60'den az	4	*57*	*228*
60-66'den az	1	*63*	*63*
Toplam	30		***1398***

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i} = \frac{33(2)+39(6)+...+63(1)}{30}$$

$$= \frac{1398}{30} = 46{,}6\ kg.$$

Örnek 3: Aşağıdaki tabloda 30 günlük süre içinde bir restoranın kullandığı et miktarının dağılımı verilmiştir. Günlük kullanılan et miktarının modunu hesaplayınız.

Mod sınıfı → 42-48'den az

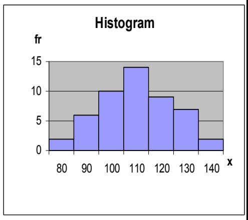
Sınıflar	f_i
30-36'dan az	2
36-42'den az	6
42-48'den az	10
48-54'dan az	7
54-60'den az	4
60-66'den az	1
Toplam	30

$$Mod = L_{mod} + \frac{\Delta_1}{\Delta_1+\Delta_2}.i$$

$$= 42 + \frac{(10-6)}{(10-6)+(10-7)}.6 = 45{,}4\ kg.$$

Örnek 4: Bir tencere pazarlama firmasına bağlı çalışan 50 satış personelinin 1 aylık tencere satışları aşağıdaki gibidir. Bu verileri kullanarak histogramı çiziniz, aritmetik ortalama, mod, medyan ve kartilleri hesaplayarak yorumlayınız.

sınıflar	f_i	m_i	$\Sigma\, m_i f_i$	Σf_i
80-≤90	2	85	170	2
90-≤100	6	95	570	8
100-≤110	10	105	1050	18
110-≤120	14	115	1610	32
120-≤130	9	125	1125	41
130-≤140	7	135	945	48
140-≤150	2	145	290	50
toplam	50		5760	



$$\bar{X} = \frac{\sum f_i m_i}{\sum f_i} = \frac{5760}{50} = 115.2$$

sınıflar	f_i	m_i	$\Sigma\, m_i f_i$	Σf_i
80-≤90	2	85	170	2
90-≤100	6	95	570	8
100-≤110	10	105	1050	18
110-≤120	14	115	1610	32
120-≤130	9	125	1125	41
130-≤140	7	135	945	48
140-≤145	2	145	290	50
toplam	50		5760	

$$Q_2 = Med = L_{Q_2} + \frac{\frac{\sum f_i}{2} - f_l}{f_{Q_2}}.i = 110 + \frac{\frac{50}{2} - 18}{14}.10 = 115$$

$$Mod = L_{mod} + \frac{\Delta_1}{\Delta_1+\Delta_2}.i = 110 + \frac{4}{4+5}.10 = 114.4$$

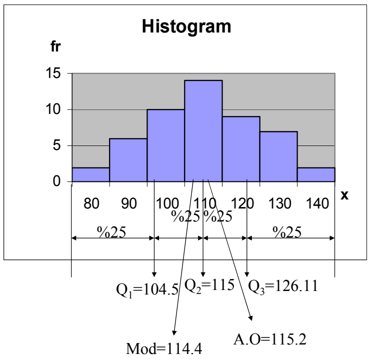
sınıflar	f_i	m_i	$\Sigma\, m_i f_i$	Σf_i
80-≤90	2	85	170	2
90-≤100	6	95	570	8
100-≤110	10	105	1050	18
110-≤120	14	115	1610	32
120-≤130	9	125	1125	41
130-≤140	7	135	945	48
140-≤145	2	145	290	50
toplam	50		5760	

$$Q_1 = L_{Q_1} + \frac{\frac{\sum f_i}{4} - f_l}{f_{Q_1}} .i$$

$$= 100 + \frac{\frac{50}{4} - 8}{10} .10 = 104.5$$

$$Q_2 = Med = 115$$

$$Q_3 = L_{Q_3} + \frac{\frac{3\sum f_i}{4} - f_l}{f_{Q_{31}}} .i = 120 + \frac{\frac{3\times 50}{4} - 32}{9} .10 = 126.11$$



Yorumlama:

Q_1 : Gözlemlerin %25'i 104.5'ten daha küçüktür.

Yorumlama:

Q_2 : Gözlemlerin %50'si 115'ten daha küçük, %50'si 115'ten daha büyüktür.

Yorumlama:

Q_3 : Gözlemlerin %25'i 126.11'den daha büyüktür.

Örnek 5: İstatistik I dersini alan 10 öğrencinin vize notları aşağıdaki gibi sıralanmıştır. Buna göre vize notları için varyans ve standart sapmayı hesaplayınız.

30,41,53,61,68,79,82,88,90,98

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{30+41+....+98}{10} = 69$$

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(30-69)^2 + (41-69)^2 + ... + (98-69)^2}{9}$$

$$= \frac{4538}{9} \approx 504{,}22$$

$$s^2 \approx 504{,}22 \quad \rightarrow s = \sqrt{s^2} = \sqrt{504{,}22} \approx 22{,}45$$

İstatistik I vizesinden alınan notların ortalama etrafında yaklaşık olarak 22 puan değiştiği görülmektedir.

Aynı soru kareler ortalamasının açılımı kullanılarak çözüldüğünde aynı sonuçları verecektir.

30,41,53,61,68,79,82,88,90,98

x	x^2
30	900
41	1681
53	2809
61	3721
68	4624
79	6241
82	6724
88	7744
90	8100

$$\sum_{i=1}^{n} x_i = 690 \quad \sum_{i=1}^{n} x_i^2 = 52148$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x\right)^2}{n}}{n-1} = \frac{52148 - \frac{(690)^2}{10}}{9}$$

$$s^2 \approx 504{,}22$$

$$s = \sqrt{s^2} = \sqrt{504{,}22} \approx 22{,}45$$

Örnek 6 : Yandaki tabloda bir Samsung bayisindeki LCD televizyonların ekran boyutlarına göre satış miktarları verilmiştir. Frekans dağılımının varyans ve standart sapmasını hesaplayınız.

Grup	Frekans	$x_i f_i$	$x_i^2 f_i$
51	1	**51**	**2601**
66	3	**198**	**13068**
72	4	**288**	**20736**
82	5	**410**	**33620**
94	7	**658**	**61852**
	$\sum f_i$ **=20**	**1605**	**131607**

$$s^2 = \frac{\sum_{i=1}^{k} f_i x_i^2 - \frac{\left(\sum_{i=1}^{k} f_i x_i\right)^2}{\sum_{i=1}^{k} f_i}}{\sum_{i=1}^{k} f_i - 1} = \frac{131607 - \frac{(1605)^2}{20}}{19} \approx 147{,}67$$

$$s = \sqrt{s^2} = \sqrt{147{,}67} \approx 12{,}15$$

Örnek 7 : Kuruyemiş satan bir dükkanda bir haftalık sürede satılan leblebi, fıstık ve bademlerin ortalamaları ve standart sapmaları aşağıda verilmiştir. Buna göre kuruyemişleri değişkenlikleri açısından karşılaştırınız ve kuruyemişin değişkenliğinin daha fazla olduğunu belirtiniz.

	$\bar{X}$	s
Leblebi	30 kg.	5 kg.
Fıstık	40 kg.	4 kg.
Badem	10 kg.	3 kg.

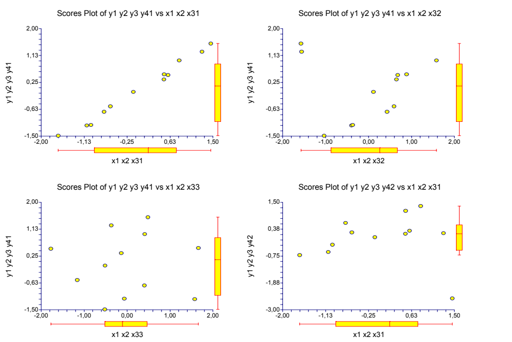
$$C_{V_{leblebi}} = \frac{s}{\bar{X}} * 100 = \frac{5}{30} * 100 = 16{,}67 = \%16{,}67$$

$$C_{V_{fıstık}} = \frac{s}{\bar{X}} * 100 = \frac{4}{40} * 100 = 10 = \%10$$

$$C_{V_{BADEM}} = \frac{s}{\bar{X}} * 100 = \frac{3}{10} * 100 = 30 = \%30$$

Üç kuruyemişin değişkenlikleri karşılaştırıldığında en küçük standart sapma değeri bademde olmasına rağmen en büyük varyasyon katsayısına sahip olduğundan en fazla değişkenliğin bademde olduğu görülür. Aritmetik ortalamalar içerisinde standart sapma yüzdelerine bakıldığında en büyük yüzde bademdedir.

GRAFİKLER


Scores Plot of y1 y2 y3 y41 vs x1 x2 x31
y1 y2 y3 y41
x1 x2 x31
Scores Plot of y1 y2 y3 y41 vs x1 x2 x32
y1 y2 y3 y41
x1 x2 x32
Scores Plot of y1 y2 y3 y41 vs x1 x2 x33
y1 y2 y3 y41
x1 x2 x33
Scores Plot of y1 y2 y3 y42 vs x1 x2 x31
y1 y2 y3 y42
x1 x2 x31