

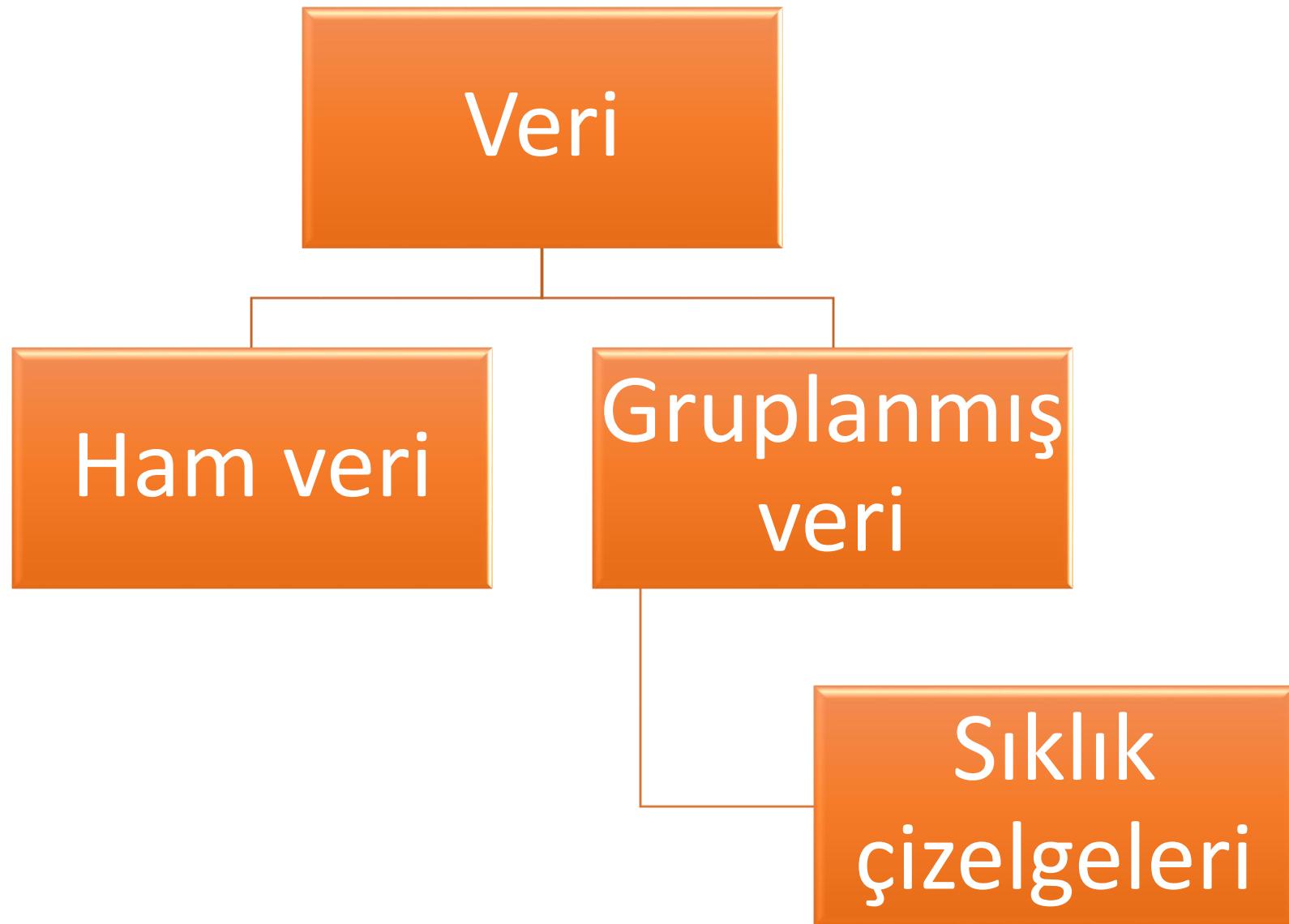


HACETTEPE  
ÜNİVERSİTESİ  
İSTATİSTİK BÖLÜMÜ

## İST155 İSTATİSTİĞE GİRİŞ I

### DERS 3 –SIKLIK ÇİZELGELERİ

**Ders sorumluları: Prof.Dr.Serpil AKTAŞ ALTUNAY (01 Şubesi)  
Doç.Dr. Ayten YİĞİTER (02 Şubesi)**



## Hangisi daha açıklayıcı?



Bir havaalanından 6 saat içinde kalkan uçaklardaki gecikme süresi (dk)

0	12	3	40	41	50
2	3	1	4	2	7
1	25	55	11	2	16
21	9	5	5	6	19
11	12	13	39	14	2
22	30	10	4	3	3
15	26	31	23	33	1
23	10	27	14	1	9
60	43	60	45	26	8
9	24	4	20	10	5
1	2	44	55	1	3
4	1	5	11	17	18
20	2	0	21	20	26
4	1	7	32	15	33
21	2	3	37	5	35
22	24	36	45	11	45
34	5	4	40	20	2
37	5	0	55	21	50
2	9	0	10	33	51
1	12	22	11	34	1
0	11	23	12	5	0
0	12	45	0	10	2
50	7	0	5	0	2

Gecikme	Uçak sayısı	Yüzde
0-10	66	47,8
11-20	24	17,4
21-30	18	13,0
31-40	14	10,1
41-50	11	8,0
51-60	5	3,6
Toplam	138	100

## Sıklık çizelgeleri



Acil polikliniğine başvuran hastaların  
yaşları

Yaş	n
35-40	21
41-45	55
46-50	80
51-55	95

COVID-19 Pozitif olan  
kişilerin son 10 gündeki  
temas ettiği kişi sayısı

Temas	n
0	1
1-5	9
6-10	13
11-15	20
>15	5

## En çok şikayet edilen belirti

Şikayet	n
Yüksek Ateş	60
Öksürük	34
Boğaz Ağrısı	56
Eklem Ağrısı	9
Halsizlik	4
İshal	3

COVID-19 geçirme  
sayısı

Kaç kez	n
1	55
2	6
3	1

## Başvurdukları hastane

Başvuru	n
Şehir hastanesi	101
Üniversite Hastanesi	44
Sempt Polikliniği	10
Özel Hastane	23
Devlet Hastanesi	77

## Sağlık Hizmetlerinden memnuniyet

Memnuniyet	n
Hiç memnun değilim	15
Memnun değilim	22
Memnunun	19
Çok memnunum	13

The height of people in 2 classes:

Height (cm)	People
140 - 159	14
160 - 179	22
180 - 199	28
200 - 219	3

Helen's scores in a game:

Score	Frequency
40 - 49	3
50 - 59	5
60 - 69	6
70 - 79	2

Annual salary (£)	Number of people
0 - 9,999	10
10,000 - 19,999	9
20,000 - 29,999	9
30,000 - 39,999	1
40,000 - 49,999	1

Sports	Tally	Frequency
Baseball		7
Football		6
Soccer		9
Basketball		12

Table 13: How many hours do you spend watching television?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-2 hours	85	39,5	39,5	39,5
	3-4 hours	77	35,8	35,8	75,3
	5-6 hours	49	22,8	22,8	98,1
	Others	4	1,9	1,9	100,0
	Total	215	100,0	100,0	

# Nicel Verilerde Sıklık Çizelgeleri

Sıklık çizelgelerinde bazı tanımlar:

Elimizde  $n$  tane gözlemden oluşan bir veri kümesi olsun.

- **En küçük değer (minimum)** : Verilerdeki en küçük gözlem değeri.
- **En büyük değer (maximum)** : Verilerdeki en büyük gözlem değeri.
- **Dağılım Genişliği (DG)-Range (R)**: Verilerdeki en büyük gözlem ile en küçük gözlem arasındaki fark.

$$DG = \text{Max} - \text{Min}$$

- Örneğin bir sınıfta en büyük yaş 25, en küçük yaş 17 ise

$$DG = 25 - 17 = 8 \text{ 'dir.}$$

**Sınıf (class):** Verilerin ayrıldığı grup sayısıdır.  $k$  ile gösterilir, tam sayıdır.

$k=1+3.3\log(n)$ 'den

ya da

$\sqrt{n}$  'den hesaplanır.  $k$ 'nın genellikle 5 ile 20 arasında olması istenir.

**Alt Sınır (AS) :** Bir sınıftaki aralığın küçük olanıdır.

**Üst Sınır (ÜS) :** Bir sınıftaki aralığın büyük olanıdır.

Örneğin, yaş aralığı «21-35» ise AS=21 ; ÜS=35'dir.

**Sınıf aralığı (c) :** Ard arda gelen iki sınıfın alt ya da üst sınırları arasındaki farktır. Örneğin,

YAŞ

21-35

36-50

olarak iki sınıfa ayrılmışsa  $c=36-21=15$  ,  $c=50-35=15$ 'dir.

$$c = \frac{\text{Dağılım Genişliği (DG)}+a}{k}$$

Eğer veri tam sayıdan oluşuyor ise:  $a=1$

Eğer veri ondalıklı değer içermiyor ise:

noktadan sonra 1 hane olması durumunda  $a=0.1$

noktadan sonra 2 hane olması durumunda  $a=0.01$

Örneğin,  $DG=13$ ,  $k=10$  ise :  $c = \frac{13+1}{10} = 1.4$

$DG=5.7$  ,  $k=5$  ise  $c = \frac{5.7+0.1}{5} = 1.16$

$DG=4.05$  ,  $k=6$  ise  $c = \frac{4.05+0.01}{6} = 0.677$

**Sıklık (frekans) ( $f_i$ ) :** Bir sınıfa kaç gözlem düştüğünü gösterir.

$$f_1 + f_2 + f_3 + \dots + f_k = n$$

$$\sum_{i=1}^k f_i = n$$

**Çetele (tally)**: Gözlemlerin düşüğü sınıfları belirlemek için yapılan işaretlemelerdir. Çetelemeler sayılarak fi'ler elde edilir.

**Göreli frekans (percent) ( $p_i$ )** : Her sınıfı düşen sıklığın toplam denek sayısına oranıdır.

$$p_i = \frac{f_i}{n} , \quad i = 1, 2, \dots, k \quad \sum_{i=1}^k p_i = 1$$

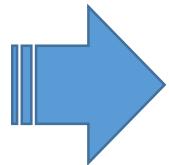
**Sınıf Orta Değeri (Mid Point) ( $S_i$ )**: Sınıfın AS ve ÜS 'nin ortalamasıdır.

$$S_i = \frac{AS_i + ÜS_i}{2} , \quad i = 1, 2, \dots, k$$

Örneğin 21-35 için  $S_1 = (21+35)/2 = 28$ 'dir.

**Birikimli Sıklık ( $F_i$ ) (cumulative frequency)**: Bir sınıfın sıklığı ile önceki sınıfların sıklıklarının toplamı ile elde edilen sıklıklardır.

**Birikimli yüzde (% $F_i$ )** :  $\frac{F_i}{n}$  oranından hesaplanır.



Nicel veriler için düzenlenen sıkılık çizelgelerinde AS, ÜS, fi, pi,  $S_i$ ,  $F_i$  ve  $\%F_i$  sütunları bulunur.

**Örnek 1:** 30 öğrenciye ait dönem sonu akademik ortalama puanları verilmiştir. Sıklık çizelgesini oluşturunuz.

2.12	3.42	2.76	3.18	0.50	2.14
2.35	3.70	2.84	1.96	1.29	2.36
2.78	2.80	2.15	1.65	1.03	3.87
2.90	2.91	2.19	1.72	1.59	2.40
3.05	2.63	3.26	0.98	1.70	2.61

$$\text{Min}=0.50$$

$$\text{Max} = 3.87$$

$$\sqrt{n} = \sqrt{30} = 5.47 \approx 5$$

$$\text{DG} = 3.87 - 0.50 = 3.37$$

$$c = \frac{3.37 + 0.01}{5} = 0.676 \cong 0.7$$

En küçük gözlemden başlanır

Bir alttaki sınıfın AS'sinin son hanesinden 1 çıkartılır

AS	ÜS	Çetele	fi	pi	Si	Fi	% Fi
0.50	1.19	///	3	0.10	0.845	3	0.10
1.20	1.89	///	5	0.167	1.545	8	0.27
1.90	2.59	/// //	8	0.267	2.245	16	0.53
2.60	3.29	/// // /	11	0.366	2.945	27	0.90
3.30	3.99	///	3	0.10	3.645	30	1
			30	1			

$$=(0.50+1.19)/2$$

$$=3/30$$

$$=3+5+8$$

$$=16/30$$

Akademik ortalaması 2.60 ile 3.29 arasında olan kaç öğrenci vardır?

Akademik ortalaması 0.50 ile 1.19 arasında olan öğrencilerin yüzdesi kaçtır?

Akademik ortalaması 2.60 ve daha fazla olan öğrencilerin yüzdesi kaçtır?

Akademik ortalaması 1.89 ve daha az olan öğrencilerin yüzdesi kaçtır?

## Sürekli Nicel verilerde AS ve ÜS gösterimleri :

Bir spor kulubündeki sporcuların boy uzunlukları

Boy	Sıklık
150-160	14
161-170	50
171-175	32
176-180	19
181-190	8

Boy	Sıklık
$150 < \text{boy} \leq 160$	14
$160 < \text{boy} \leq 170$	50
$170 < \text{boy} \leq 175$	32
$175 < \text{boy} \leq 180$	19
$180 < \text{boy} \leq 190$	8

**Örnek 2:** 43 öğrencinin İstatistik dersinin genel sınavından aldığı puanları verilmiştir. Sıklık çizelgesini oluşturunuz.

45	50	75	60	95	80	100	95	50	65	95
100	70	100	80	75	90	75	85	95	70	100
60	100	85	95	95	100	65	30	65	75	80
65	33	23	28	36	40	45	59	85	46	

$$k=1+3.3\log (43)=6.43 \sim 6$$

$$\text{MIN}=23$$

$$\text{MAX}=100$$

$$\text{DG}=100-23=77$$

$$c = \frac{77 + 1}{6} = 13$$

<b>AS</b>	<b>ÜS</b>	<b>Çetele</b>	<b>fi</b>	<b>Si</b>	<b>pi</b>	<b>Fi</b>	<b>% Fi</b>
23	35	////	4	29	0.09	4	0.093
36	48	///	5	42	0.12	9	0.209
49	61	///	5	55	0.12	14	0.326
62	74	/// /	6	68	0.14	20	0.463
75	87	/// //	10	81	0.23	30	0.698
88	100	/// //	13	94	0.30	43	1
		///	43	1			

## Bir sıklık çizelgesinin sağlaması gereken özellikler

- 1 • Veri kümesindeki en büyük gözlem ile en küçük gözlemi kapsamalıdır.
- 2 • Açıkta gözlem kalmamalıdır.
- 3 •  $k$  ve  $n$  orantılı olmalıdır.
- 4 • Sıfır sıklık olmamalıdır.
- 5 • Yoğunluk sınıfının ortalarında olmalı uçlara doğru azalmalıdır.

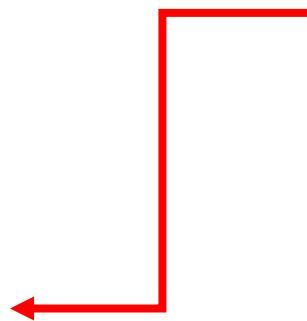
## Nitel Verilerde Sıklık Çizelgeleri

Sıralanabilir nitel veriler için düzenlenen sıklık çizelgelerinde:  
Sınıf,  $f_i$ ,  $p_i$ ,  $F_i$  ve  $\%F_i$  sütunları yer alır.

Sınıflanabilir nitel veriler için düzenlenen sıklık çizelgelerinde:  
Sınıf,  $f_i$  ve  $p_i$  sütunları yer alır.

**Örnek 3:** Yaz okulunda Matematik dersi alan 60 öğrencinin genel sınav harf notları verilmiştir. Sıklık çizelgesini düzenleyiniz.

Verilerin özelliği : Ham veridir.  
Sıralanabilir nitel veridir.



A1	B1	F2
A1	B2	B3
A2	D	B3
A2	D	C3
B1	D	C2
F3	A1	C1
F3	A2	D
F3	C1	D
F2	C2	D
A1	C1	B1
B1	C3	B1
B2	C3	B1
B3	C1	A2
D	A1	D
D	A2	D
D	F3	F3
D	F3	F3
D	F3	F2
F3	C1	C2

<b>NOTLAR</b>	<b>fi</b>	<b>pi</b>	<b>Fi</b>	<b>%Fi</b>
A1	5	0.083	5	0.083
A2	5	0.083	10	0.167
B1	6	0.10	16	0.267
B2	2	0.033	18	0.30
B3	3	0.05	21	0.35
C1	5	0.083	26	0.433
C2	3	0.05	29	0.483
C3	3	0.05	32	0.533
D	14	0.233	46	0.767
F2	3	0.05	49	0.817
F3	11	0.183	60	100
<b>Toplam</b>	<b>60</b>	<b>1</b>		

Sınıfın %23.3'ü (14/60) «D» notu ile geçmiştir.

Sınıfın % 76.7'si (46/60) «D» ve üstü not alarak geçmiştir.

Sınıfın %23.3'ü (1-0.767) «F2 ve F3» notu alarak kalmıştır.

**Örnek 4:** Bir kütüphanede bir gün içinde ödünç verilen 100 kitabın türleri bilgisayar kayıtlarından elde edilmiştir. Sıklık çizelgesini düzenleyiniz.

Edebiyat	Güzel Sanatlar	Güzel Sanatlar	Güzel Sanatlar	Tarih
Tarih	Güzel Sanatlar	Güzel Sanatlar	Güzel Sanatlar	Coğrafya
Güzel Sanatlar	Coğrafya	Bilim ve Teknik	Coğrafya	Edebiyat
Bilim ve Teknik	Coğrafya	Coğrafya	Coğrafya	Edebiyat
Felsefe	Temel Bilimler	Tarih	Tarih	Bilim ve Teknik
Temel Bilimler	Tarih	Güzel Sanatlar	Güzel Sanatlar	Felsefe
Edebiyat	Edebiyat	Edebiyat	Edebiyat	Felsefe
Edebiyat	Edebiyat	Edebiyat	Edebiyat	Edebiyat
Edebiyat	Güzel Sanatlar	Güzel Sanatlar	Güzel Sanatlar	Güzel Sanatlar
Edebiyat	Tarih	Temel Bilimler	Coğrafya	Tarih
Edebiyat	Felsefe	Felsefe	Felsefe	Felsefe
Edebiyat	Edebiyat	Edebiyat	Coğrafya	Coğrafya
Edebiyat	Edebiyat	Bilim ve Teknik	Edebiyat	Coğrafya
Edebiyat	Tarih	Tarih	Edebiyat	Temel Bilimler
Edebiyat	Temel Bilimler	Bilim ve Teknik	Coğrafya	Edebiyat
Edebiyat	Temel Bilimler	Temel Bilimler	Temel Bilimler	Coğrafya
Tarih	Edebiyat	Edebiyat	Coğrafya	Bilim ve Teknik
Tarih	Edebiyat	Edebiyat	Edebiyat	Bilim ve Teknik
Tarih	Edebiyat	Bilim ve Teknik	Edebiyat	Bilim ve Teknik
Edebiyat	Bilim ve Teknik	Güzel Sanatlar	Tarih	Bilim ve Teknik

→ Verilerin özelliği : Ham veridir. Sınıflanabilir nitel veridir.

<b>Kitap Türü</b>	<b>Sıklık</b>	<b>pi</b>
Bilim ve Teknik	11	0.11
Coğrafya	13	0.13
Edebiyat	34	0.34
Felsefe	7	0.07
Güzel Sanatlar	14	0.14
Tarih	13	0.13
Temel Bilimler	8	0.08
Toplam	100	1

En çok ödünç alınan kitap türü %34 ile Edebiyat kitaplarıdır.

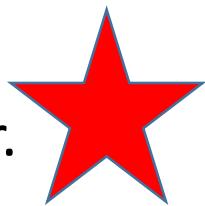
En az ödünç alınan kitap türü %7 ile Felsefe kitaplarıdır.

## Eşit Aralıklı Olmayan ve Açık Uçlu Sıklık Çizelgeleri

**Eşit aralıklı olmayan sıklık çizelgesi:** Bir veri kümesinde bazı sınıflara veri düşmemesi sonucu bu sınıfların birleştirilmesiyle ortaya çıkan sıklık çizelgeleridir. Bu tür sıklık çizelgelerinde sınıf aralığı ( $c$ ) sabit değildir.

**Açık uçlu sıklık çizelgeleri :** Veri kümesinde en büyük ya da en küçük değerin bilinmediği durumda hazırlanan sıklık çizelgeleridir.

Bir sıklık çizelgesi hem eşit aralıklı olmayan hem de açık uçlu olabilir.



## Örnekler

Yaş	Sıklık
10-20	23
21-31	6
32-42	4
43-53	10
54 ve üzeri	8

Bir hastaneye başvuran hastaların yaş dağılımı: üst sınırı açık uçlu sıkılık çizelgesi

Koşu	Sıklık
10sn altında	3
11-14	24
15-18	38
19-22	14
23-26	5

100m koşu şeçmelerine başvuran sporcuların koşu süreleri: alt sınırı açık uçlu sıkılık çizelgesi

Sigara	Sıklık
0	71
1-5	22
6-7	34
9-11	10
15-22	7
23-25	4

Bir günde içilen sigara sayısı: eşit aralıklı olmayan sıkılık çizelgesi

Gelir	Sıklık
$\leq 1500$	101
1501-2000	275
2001-3000	302
3001-5000	99
5001-10000	14
$\geq 10001$	3

Hanehalkı geliri: alt-üst sınırı açık uçlu ve eşit aralıklı olmayan sıkılık çizelgesi

**ÖDEV:** Bir pandemi hastanesine başvuran ve COVID-19 test sonucu pozitif çıkan 160 hastanın yaş dağılımı verilmiştir. Bu veriler için sıkılık çizelgesini düzenleyiniz.

15	5	77	34	39	3	10	55
2	6	34	33	10	9	28	34
8	9	54	30	11	13	39	21
12	10	11	10	23	32	33	22
15	22	22	22	34	20	21	45
67	23	34	34	54	12	45	67
54	34	56	45	53	35	12	78
23	23	65	54	16	38	75	75
12	45	45	65	17	49	70	23
2	67	34	76	76	80	40	20
80	54	42	65	40	39	55	19
27	12	12	43	31	40	51	15
23	13	22	78	22	56	60	14
34	56	23	61	30	29	71	20
45	44	55	60	19	31	32	30
54	56	76	35	10	30	35	32
30	76	66	36	4	22	80	71
13	45	23	37	50	19	65	74
19	12	80	44	45	20	12	5
11	9	65	40	44	10	4	1

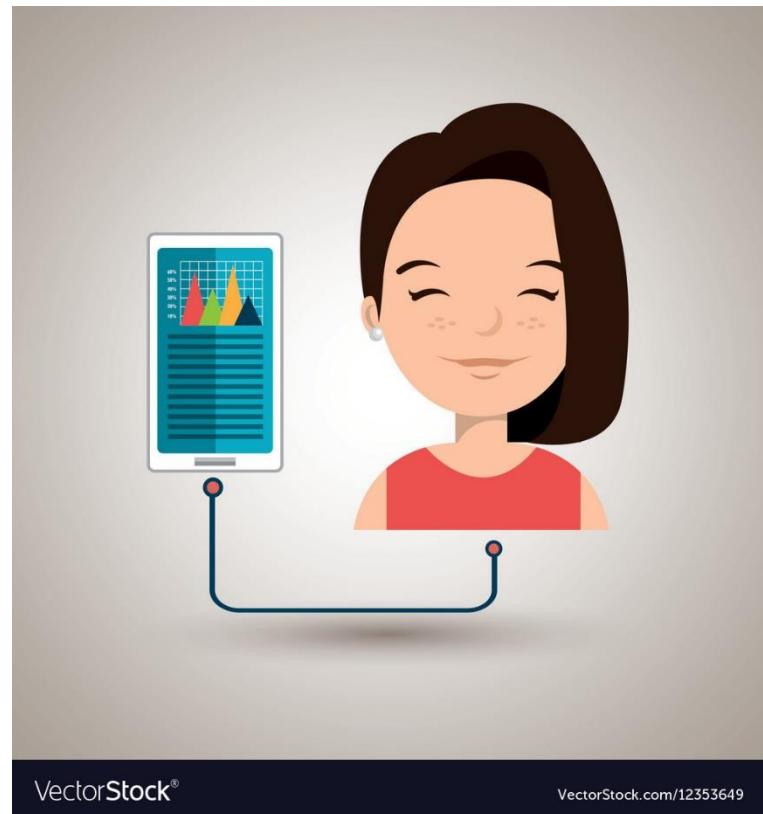
## BIG DATA (BÜYÜK VERİ) NE DEMEK ?

Sosyal medya araçları, arama motorları , internet alışveriş siteleri gibi kaynaklardan elde edilen büyük hacimli verilerdir. Bu veriler işlenerek anlamlı bilgiler elde etme amaçlanır.

Büyük veriye yatırım yapılarak satış, ürün geliştirme ve daha başka farklı alanlarda etkin bir fayda sağlanmaktadır.

Big data'ya yatırım yapan şirketler:

Facebook, Instagram, Amazon, Netflix, Starbucks, General Electrics, Whatsapp



Bir sonraki derste grafikler incelenecek.



## KAYNAKLAR

- 1-) H.Demirhan, C.Hamurkaroğlu ,“İstatistiksel Yöntemlere Giriş”, H.Ü.Yayınları, 2011.
- 2-) Serpil Cula, Zehra Muluk, “Temel İstatistik Yöntemler”, Başkent Üniversitesi yayınları,2006.
- 3-) Levent Özbek, Esin Köksal Babacan, “İstatistiğe Giriş”, TÜBİTAK e-kitap.
- 4-) Birdal Şenoğlu, Mehmet Yılmaz, Sibel Açık Kemaloğlu, İstatistiğe Giriş, TÜBİTAK e-kitap.