



Biyostatistik

Dr. Öğr. Üyesi Necip Ozan Tiryakioğlu

İzlençe

- Ders İçeriđi: Temel biyoistatiksøl kavramlar anlatılır. Popölasyon ortalamaları ve oranları, deđişkenlik, veri bildirimini, hipotez testleri ve grup karşılaştırmaları öğrenilir.
- Notlandırma: %40 Vize, %60 Final
- Ofis Saatleri: Çarşamba günleri 10:00-16:00

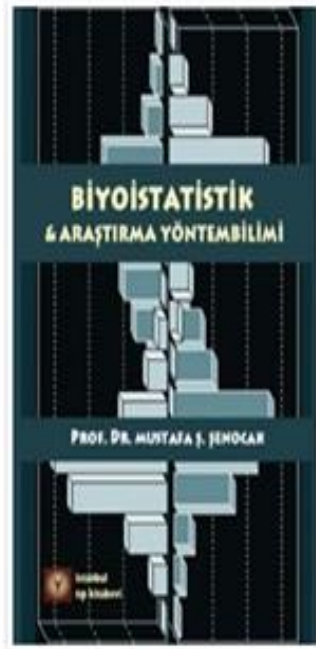
Konu Dağılımı

1	Tanışma-Biyoistatistiğe Giriş
2	İstatistiki kavramlar ve Verilerin Düzenlenmesi
3	Tanımlayıcı İstatistik
4	Merkezi dağılım ölçüleri
5	Olasılık Dağılımları
6	Normal Dağılım
7	Güven Aralıkları ve Örnekleme Yöntemleri
8	Nokta ve Aralık Tahminlenmesi
9	Ortalamaların Karşılaştırılması
10	Oranların Karşılaştırılması
11	Hipotez testleri

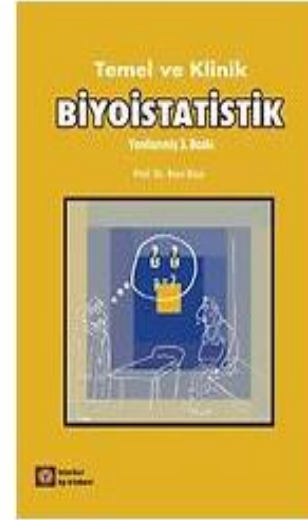
Kaynaklar

Biyoistatistik & Araştırma Yöntembilimi

Yazar Mustafa Ş. Şenocak



Temel ve Klinik Biyoistatistik 3. Baskı



İstatistik Nedir?

Giriş

- Belirli olayların gözlemlenmesi
- Gözlemle ile elde edilen verilerin toplanması
- Düzenlenmesi
- Uygun yöntemlerle analiz edilmesi ile ilgilenen bir bilim dalıdır.

İstatistik

- **Olasılığa dayalı olayların** incelenmesi
- Kuramsallaştırılması
- Aralarında bağıntı kurulması yöntemlerinin tümü → İstatistik bilimini oluşturur.

- Verileri derleme, deney tasarımı, yorumlama ve genelleme yöntemlerini veren bilim → İstatistik
- Canlı özelliklerinin ve bunları etkileyen tüm öğelerin belirlenmesi, sayılması, düzenlenmesi ve değerlendirilmesi sonucu nedensel bağlantıları tespit eden bilim → Biyoistatistik

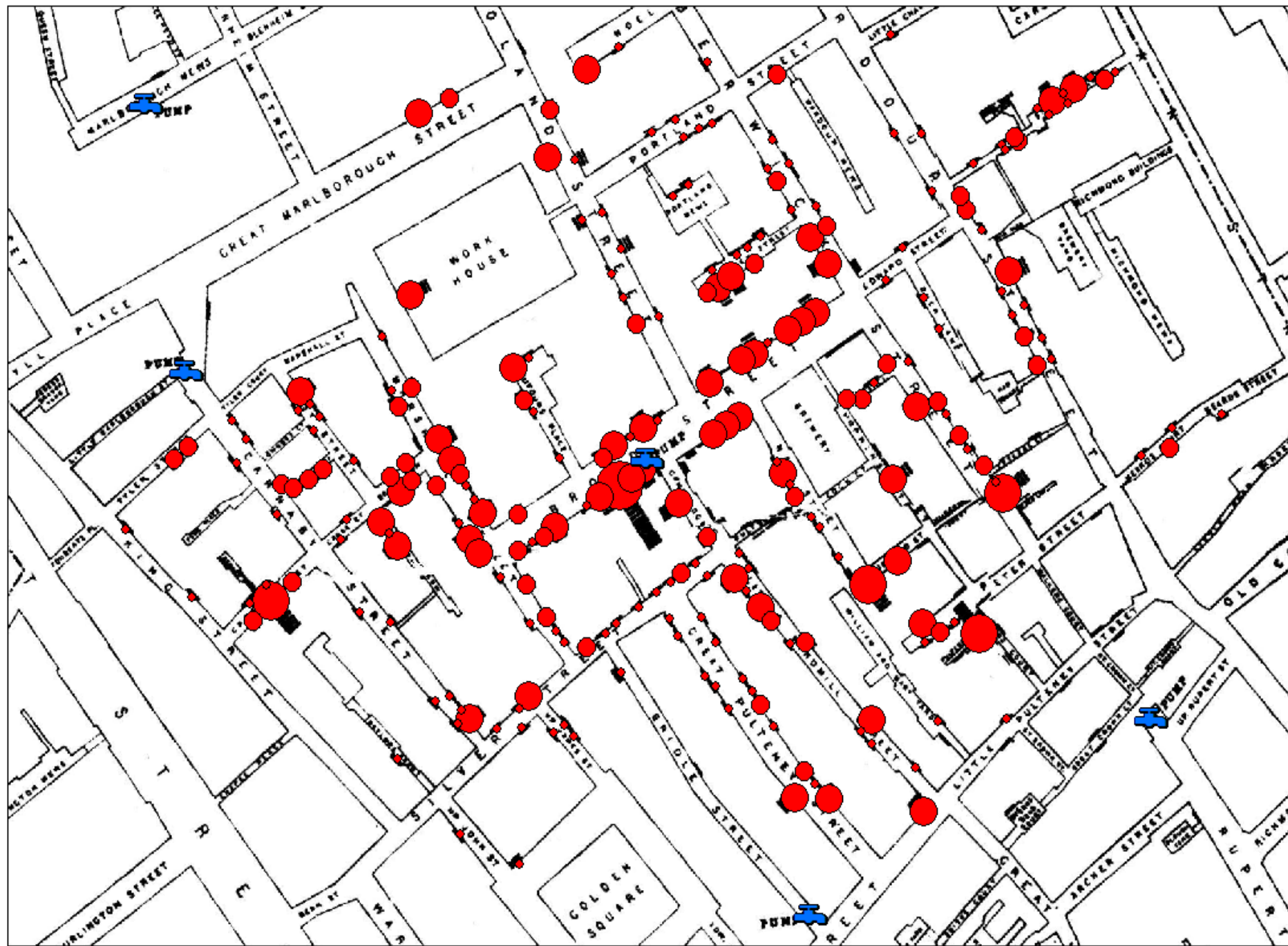
Biyostatistik Kullanım Alanları

- Koruyucu Tıpta
 - Genel ve bölgesel sağlık durumu saptama ve planlanması
 - Toplumsal Özelliklerin değerlendirilmesi
 - Farklı toplumların kıyaslanması
 - Tarama ,tedavi ve önlemlerin değerlendirilmesinde
- Klinik Tıpta
 - Klinik tanı yöntemlerinin yargılanması-kıyaslanması
 - Yeni tedavi yöntemlerinin yargılanması, kıyaslanması
 - Sonuçları etkilediği düşünülen farklı etkenlerin etki derecelerinin ortaya konması

- Tıp ve genel biyoloji → Temel biyoistatistik
- Toplum çapında → Epidemiyoloji
- Doğrudan hasta ve hastalıklara ilişkin → Klinik biyoistatistik

- Ebu Yusuf → De Gradibus, ilaç etkilerini ölçmek için matematiksel metod (850)
- İbni Sina → El Kanun fi el Tıbb, ilaç kullanımı (1025)
- John Baptiste → deney tasarımı, randomizasyon

- Armand Trousseau → Plasebo etkisi (1834)
- John Snow → 1854 kolera salgını, salgın kaynağının tespiti



- Karl Pearson → Korelasyon, regresyon, ki kare
- William Gosset → T testi
- Ronald Fisher → varyans çözümlemesi, popülasyon genetiği

Belirsizlik

- Bilimsel alanlarda bile bilgiye yeterli ve doğru erişilmesi yönünde büyük bir sorun bulunmaktadır: Belirsizlik.
- Belirsizlik: Açık ve net bilgi yoksunluğu

Belirsizliğin Kaynakları

- Kişiler arası farklar: Genetik, Yaş, Cinsiyet, Boy, Kilo
- Yetersiz Bilgi:
 - Hasta kaynaklı: Hatırlayamamak. Doğru tanımlayamamak.
 - Tanı ve Prognoza ilişkin: Olanakların yetersizliği

“Tıp bir belirsizlik ve olasılık kullanma sanatıdır.”

Sir William Osler

Nedensellik

- İki ayrı durumun özellikleri arasındaki birliktelik.
- Sebep → Sonuç

- Nedensellik bağlantısı → Olay ve olguların bağlantılı olması.
- Daha önce var olan bir olayın daha sonra meydana gelen bir olaya neden olması veya etkilemesi.

	Etken var (E+)	Etken yok (E-)	
Sonuç var (S+)	A	B	Toplam S+
Sonuç yok (S-)	C	D	Toplam S-
	Toplam E+	Toplam E-	

	Ebola enfeksiyonu var	Ebola enfeksiyonu yok	
Ölüm	69	0	69
Sağlıklı	7	44	51
	76	44	

- Ebola ile ölümlü nedensellik bağlantısı;
- Ebola olmayanlar arasında ölen yok (B boş)
- Ölüm için ebola gerekli
- Bütün ebola olanlar ölmemiş (C dolu)
- Ebola ölüm için yeterli değil

	Kuduz enfeksiyonu var	Kuduz enfeksiyonu yok	
Ölüm	69	0	69
Sağlıklı	0	44	44
	69	44	

- Kuduz ölüm için gerekli, B boş.
- Kuduz ölüm için yeterli, C boş.

- Sebep sonuç için hem gerekli hem de yeterliye sebep sonuç arasında kesin bağıntı vardır.
- Bu tip direkt bağlantılara 'tipik olay' adı verilir.
- Tipik olaylara çok nadir rastlanır.

	Sigara (+)	Sigara (-)	
Akciğer Ca +	75	18	93
Akciğer Ca -	425	1782	2207
	500	1800	

- $B > 0 \rightarrow$ Her AK. Ca sigara içmiyor
- $C > 0 \rightarrow$ Her sigara içen AK. Ca değil
- Sigara AK Ca için gerekli veya yeterli değil.

- Sigaranın etkisi nedir?

Sigara içip kanser olanların oranı;

$$75/500 \rightarrow \%15$$

Sigara içmeyip kanser olanların oranı;

$$18/1800 \rightarrow \%1$$

- $\%15 / \%1 \rightarrow 15$.
- Sigara içenler 15 kat daha fazla AK Ca oluyor.
- Odds Ratio \rightarrow Oranlar oranı
- $(\text{Sonuç} / \text{Etkenin Görüldüğü Vaka}) / (\text{Sonuç} / \text{Etkenin Olmadığı Vaka})$

- Sigara içmek kanser riskini arttırır. Yani sigara içmenin kansere sebep olma ihtimali vardır.
- Atipik olay. Olasılığa bağlı süreç.

	Erkek(+)	Kadın(-)	
Mesane Ca +	75	18	93
Mesane Ca -	1425	1782	3207
	1500	1800	

- $(75 / 1500) / (18/1800) = 5$
- Erkeklerin mesane kanserine yakalanma riski 5 kat fazla

	BRCA1(+)	BRCA1(-)	
Meme Ca +	150	90	159
Meme Ca -	1350	1710	3060
	1500	1800	

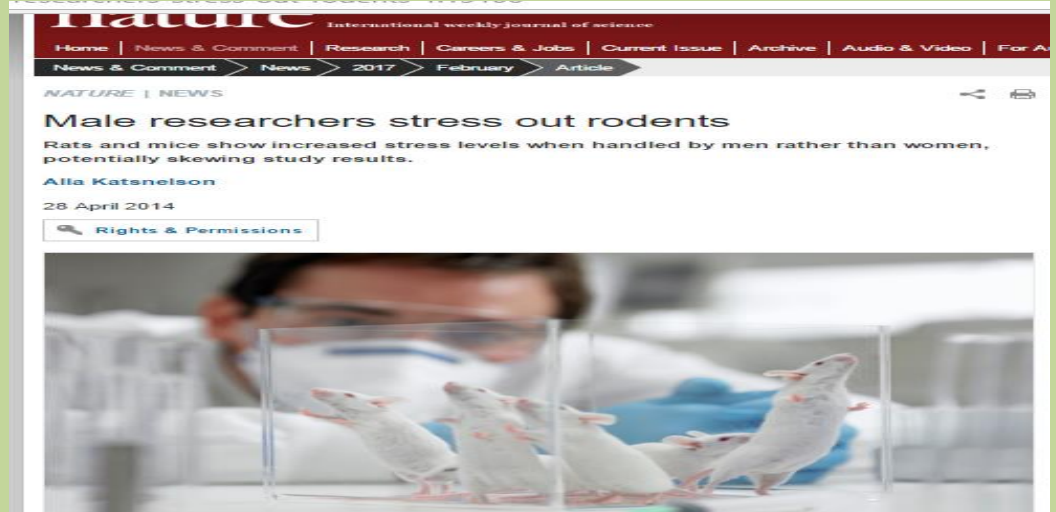
	BRCA1(+)	BRCA1(-)	
Meme Ca +	150	90	159
Meme Ca -	1350	1710	3060
	1500	1800	

- Toplam BRCA1 mut. Kadınların (1500) 150 'si meme ca. $(150/1500)=0.1$ (%10)
- Toplam BRCA1 mut. olmayan kadınların (1800) 9'u meme ca. $(90/1800)=0.05$ (%5)
- %10 / %5 =2 kat daha fazla

	Takım Tutmak (+)	Takım Tutmak (-)	
Mesane Ca +	12	11	23
Mesane Ca -	189	188	377
	201	189	

- $(12/201)/(11/189)=1,02$
- Bir bağlantı yok.

- Gerçek Sebep?



- Matematiksel bağlantı olsa bile fizyolojik bir bağlantı olmadan gerçek bir nedensellik önerilemez

' 3 Çeşit yalan vardır; yalan, kuyruklu yalan ve istatistik'

Mark Twain

- B brek taşı hastalarında kemik yoęunluęu az.
- B brek taşı hastalarında oxf. kolonizasyonu az.
- O zaman oxf kolonizasyonunun az olması kemik yoęunluęunu azaltır?

Current Urology

Original Paper

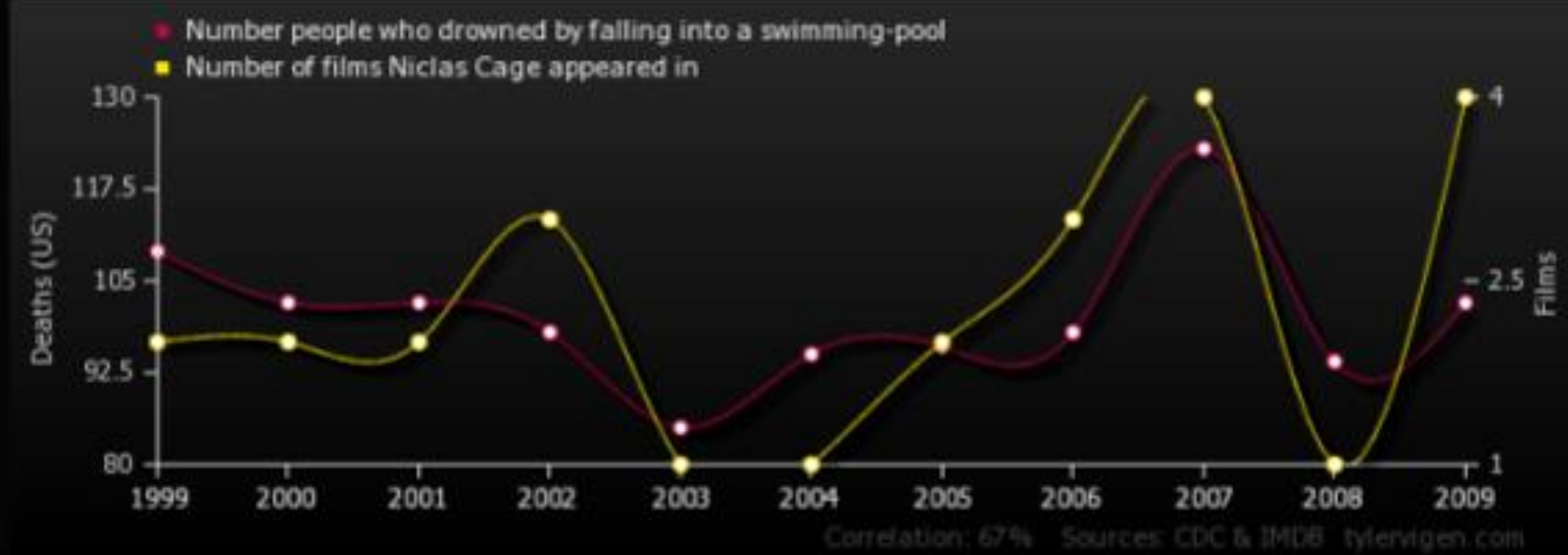
Low *Oxalobacter Formigenes* Colonization is Associated with Reduced Bone Mineral Density in Urinary Stone Forming Patients

Pence S.^a · Ikizceli I.^b · Ozbek E.^c · Tiryakioglu N.O.^{a, d} · Eren H.^{a, e, h} · Polat E.C.^f · Pence H.H.^g

 Author affiliations

Keywords: Urolithiasis · Bone mineral density · Oxalobacter formigenes · Oxalate degradation

- Korelasyon, nedensellik değildir.
- Korelasyon tamamen rastlantısal olabilir.
- Hipotez dışında başka gerçek bir bağlantı olabilir.



Estrogen replacement therapy and coronary heart disease: A quantitative assessment of the epidemiologic evidence ☆☆☆

Meir J. Stampfer M.D. *, †, Graham A. Colditz M.D. *, ‡

 [Show more](#)

- HRT alan kadınlarda kalp hastalığı daha az.
- HRT, kalp hastalıklığı riskini azaltır.

- Sonular tekrarlanamıyor.
- Gerek sebep HRT alan kadınlar ekonomik olarak daha üst sınıftan → Daha ok egzersiz, daha iyi beslenme, sonuta kalp hastalığı riski daha az.

- Deney düzeneyinde hata
- Hawthorne etkisi;
 - Deneyin farkında olan deneklerin farklı davranması
- Örnek sayısı azlığı
 - Erkek vs Kadın boy uzunluğu

- Kanser vaka artışı
 - Tanı yöntemi gelişimi? Gerçek vaka artışı?
- Beyaz önlük etkisi
 - Nabız artışı, gerginlik

İstatistiksel Olaylar

- Tipik olaylar; dış etki yok, kesin sebep sonuç ilişkisi
- Atipik olaylar, yığın olaylar; rastlantısal nedenlerle belirli sınırlar dahilinde farklılaşan olaylar.
- Biyolojik olaylar → Atipik

- Toplum, popülasyon → Aynı özellik açısından ele alınan nesnelere.
 - Tip-2 diyabet hastaları, 30 yaş üzeri erkekler vs.
- Ana özellik → Toplumun incelendiği temel özellik.
 - AKŞ, Kan grubu vs.

- Örneklem → Belirli bir ana özelliğe sahip gruba toplum içinden seçme
- Örneklem → Örneklem ile seçilmiş toplum grubu

- Örneklemeden elde edilen sonuçlar iki şekilde kullanılabilir:
- Tanımlayıcı istatistik → Bir değişkenin nicel olarak ortaya konması
 - Total kolesterol, yaş ortalaması, cinsiyet dağılımı

- Çıkarım → Örnekleme gözlemleyerek toplum için bir sonuca ulaşmak.
 - Kanser riski, yaşa bağlı hastalık gelişimi

- Çıkarımlar kestirimler veya varsayım testleri şeklinde olabilir.
1. Kestirim → Bir parametrenin toplum için en olası düzeyinin belirlenmesi
 1. Nokta → Tek değer, aritmetik ortalama
 - AKŞ 70
 2. Aralık → İki değer arasında
 - Nabız 60-100 bpm

- Varsayım testleri → Bir hipotezin sınanması.
 - Sigara akciğer ca. riskini arttırır, HRT alan kadınlarda meme kanseri riski artar, soya filizi yemek meme kanserine karşı koruyucudur.
 - PSA oranları prostat kanseri açısından belirteçtir.

Hata

- Rassal hatalar → Rastlantılara bağlı olarak ortaya, ortalama gerçek değere yakın ama kestirim aralığı geniş.
 - Boy uzunluğu, cüce + basketbolcu
- Dizgesel hatalar → Yöntemsel hata
 - Hep fazla tartan tartı vs.

- Rassal hata → Kesinlik azaltır
- Dizgesel hata → Geçerlilik azaltır

Kesinlik ve Geçerlilik

- Kesinlik → Rassal değişim sınırları ne denli darsa toplum kestirimi o kadar kesin olur.
 - Erişkin erkeklerin HDL düzeyleri 13-90 mg/dL arasındadır
 - Erişkin erkeklerin HDL düzeyleri 34-48 mg/dL arasındadır

- Kesinliđi etkileyen faktörler:
 1. Örneklem Büyüklüđü (Boy uzunluđu grubu)
 2. Örnekleme Yöntemi (Hassas ölçüm)
 3. Özelliđin ölçümsel deđişkenliđi (Varyans)

- Geçerlilik, araştırma sonucunun yanıtı aranan soruya doğru yanıt verebilme yeteneği.
- Hasta veya sağlam tanısı koydurabilme özelliği
 - PSA seviyeleri, AKŞ, karaciğer enzimleri, enfeksiyon belirteçleri

- Geçerlilik toplumun sadece bir grubuyla kısıtlı olabilir.
 - 40 yaş üzeri erkekler, sadece kadınlar vs.

- Kesinlik için;
 - Sayıca ve simgeleyicilik olarak uygunluk
 - Uç değerlerin denetimi, sınır belirleme

- Geçerlilik için;
 - Varsayıma uygun test
 - T-testi, 2 grup ortalama, normal dağılım

Değişken

- Değişken → Rastlantısal etkilere bağlı ölçümsel farklılaşmalar gösteren 'birim'.
- 2 temel özelliğe sahiptir;
 - Nicelik
 - Nitelik

- Nicelik belirten deęişkenler, ölçülebilen, birimler ile ifade edilen nesnel özellikler.
 - Boy, kilo, yaş, çocuk sayısı
 - Kesintisiz → Boy, kilo
 - Kesintili → Çocuk sayısı

- Nitelik belirten deęişkenler, belirli bir özellięe sahip olma durumunu belirtirler.
 - Kan grubu, Hasta-Saęlıklı, Evli-Bekar.

- Nominal(Kategorik)
 - Cinsiyet, Ağrı
- Ordinal (Sıralı)
 - Tümör derecesi, Ağrı
- Oran (Ölçüm)
 - Tümör büyüklüğü, Boy, Ağrı

Nedensellik

- İki ayrı durumdan birinin diğ̈erinin varlığı sebebiyle ortaya çıkması veya uyumlu olarak deđiřmeleri durumuna 'nedensellik bađlantısı' adı verilir.
- 2 tip nedensellik bađlantısı vardır

1. Nedensel Baę; $A \rightarrow B$; A'nın varlığı B'ye sebep olur.

- Diyabet nöropatik ağrı yapar, güneşte yanmak/solaryum cilt kanserine neden olur.

- Nedensel bağ iki özelliğe sahip olmalıdır:
 1. Öncüllük → Neden, sonuçtan önce meydana gelir.
 2. Bitişiklik → Neden ve sonuç doğrudan temasla veya aracı bağlarla bağlanmıştır.

Bakteri → Bulaşma → Hastalık

2. Birlikte deęişim: Bir olayın seviyesinin deęişmesi, dięer bir olayın seviyesinin deęişmesine sebep olur.

- Mutlak neden sonuç ilişkisi gerekmez,.
- Bir deęişkendeki deęişim ile dięerini tahmin etmeyi saęlar.
 - CRP düzeyleri, PSA seviyeleri

III. Hafta

Temel Kavramlar

- Değişken → Bir örnekten diğerine farklılık gösteren özellikler
 - Yaş
 - Cinsiyet
 - Kilo

- Çocukların, annelerinin özelliklerine bağlı olarak zihinsel sağlıklarını inceleyen bir araştırmada değişkenler neler olabilir?
 - Annenin yaşı, eğitim durumu, sağlık durumu, akıl sağlığı

Gözlem birimi

- Daha önce verilen araştırma örneğine göre gözlem birimi nedir?
 - Anne ve çocuklar
- Biyoistatistikte genellikle gözlem birimi insandır.
 - Bakteri, virüs, hücre

- Hipotez, Varsayım → Olaylar arasında ilişkiler kurmak ve olayları bir nedene bağlamak üzere tasarlanan ve geçerli sayılan bir önermedir.
- Teori → Deney ve testler sonucunda "sürekli olarak" varsayılan sonucu veren hipotez.

- Genellikle iki veya daha fazla değişkenin arasındaki ilişki üzerinden ifade edilir.
 - Bağımlı değişken
 - Bağımsız değişken

- Birey içi değişken → Aynı kişinin sabah – akşam vücut ısısı
- Bireyler arası → İki kişinin yaş farklılığı
- Küme içi → 20 öğrencinin sınav öncesi ve sonrası tansiyonunun değişmesi
- Kümeler arası → Hasta grubuyla, sağlıklı kişiler grubu arasında kan şekeri seviyesi farkı

- Hipotez → Oral kontraseptif kullanımının yol açtığı hormonal değişimler çocukların otizmlili doğmasına sebep olur.
- Bağımlı değişken → Otizm
- Bağımsız değişken → Oral kontraseptif kullanımı

- Hipotez → Hepatit aşularına azalmış yanıt, otoimmünite göstergesidir.
- Bağımlı değişken → Otoimmünite varlığı
- Bağımsız değişken → Hepatit aşularına azalmış yanıt

Araştırma Aşamaları

1. Konu, test edilebilir bir hipotez haline getirilir.
2. Uygun araçlar seçilir → Anket, biyokimya testleri
3. Veriler toplanır ve kaydedilir.
4. Veri analiz edilir → İstatistik testleri
5. Sonuçlar yorumlanır → İstatistiksel çıkarım

Ölçüm Seviyeleri

- Ölçüm → Değişkenin düzeyini saptama işlemi
 - Nominal/Kategorik/Nitel
 - Ordinal/Sıralı
 - Nicel/Ölçü

Kategorik

- Kategorik seviyede ölçümleme → Etiketleme, Gruplama
 - Sayısal olmayan bir özelliği belirtir.
Cinsiyet, Sağlık durumu, Medeni hali

Medeni Durumu	Frekans
Evli	4
Bekar	6
Toplam	10

- Kategorik verilerde her olgu tek bir grubu dahil edilir.
 - Kesişme olmamalı
 - Geniş kapsamlı olmalı → Evli, Bekar. Dul?
 - Ölçekli, dereceli değildir. Daha az veya çok yoktur.

- Kategorik deęişkenler sayılarla gösterilebilir. Bu durum veriler arasında bir seviye farklılığı belirtmez.
- Arabanın rengi → Beyaz=1, Siyah=2, Mavi=3.

Sıralı Değişkenler

- Derece açısından sıralıdır. Ama kesin bir ölçü yoktur.
 - Sosyoekonomik durum

İsim	Sosyoekonomik durum
Hasta1	2
Hasta2	3
Hasta3	1

- Hasta2'nin sosyoekonomik durumu Hasta3'ten daha iyidir. Ama 3 kat iyi değildir.

Nicel/Ölçüsel Değişkenler

- Sıralı ölçümün tersine nicel ölçü, sıranın yanı sıra doğrudan matematiksel bir fark belirtir.
- Doğal nicel ölçüm seviyeleri:
 - Çocuk sayısı
 - Kilo,boy
 - Yaş

- Nicel deęişkenler:
 1. Yalın (Oransal) Deęişkenler: Doęal ölçümler sonucu elde edilen, ondalık deęer alabilen deęişkenlerdir.
 - 186 cm, 187 cm, 186,4 cm
 2. Yapay Deęişkenler: Ondalık deęer alamayan doğrudan bir nicelik belirten deęişkenlerdir.
 - Çocuk sayısı, geçirilen operasyon sayısı

- Yalın deęişkenler genellikle bir sıfır noktası olmaz.
 - Boy \rightarrow 0 cm?
- Yapay deęişkenler ise '0' deęeri alabilir.
 - Çocuk sayısı \rightarrow 0, yok

Kesintili ve Sürekli Nicel Değişkenler

- Kesintili değişkenler tam sayı olarak belirtilir ve sonlu bir sayıyla ifade edilir.
 - Operasyon sayısı $\rightarrow 1,2$
- Sürekli değişken herhangi iki tam sayı arasında herhangi bir değeri alabilir.
 - Kilo $\rightarrow 67, 71.2, 78, 55.32$ kg

- Çocuk sayısı:

- Çocuk sayısı \rightarrow Nicel, Kesintili

- IQ puanı:

- IQ puanı → Nicel, Kesintili

- Cinsiyet:

- Cinsiyet → Kategorik

- Boy →

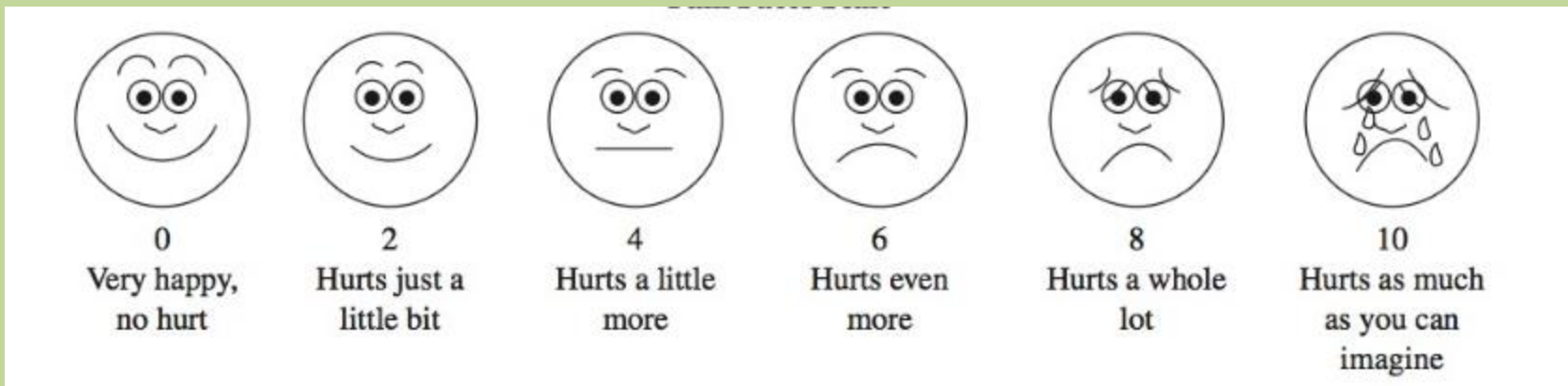
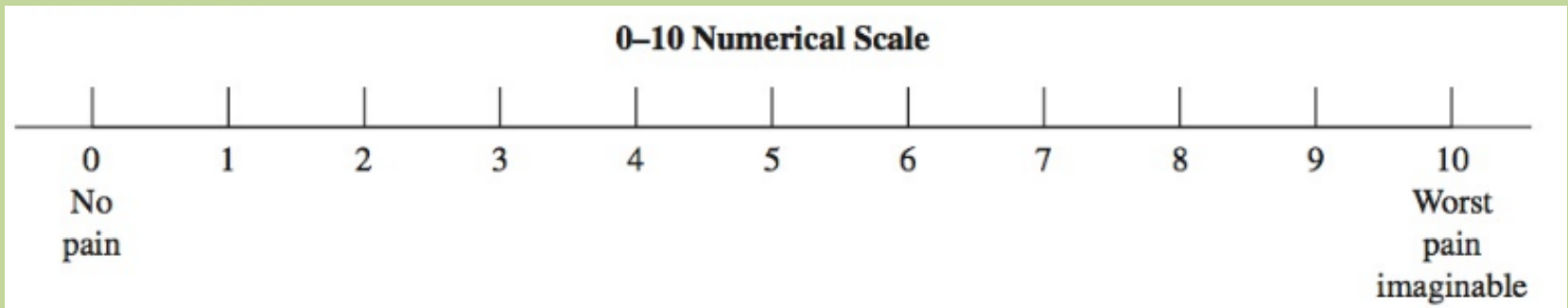
- Boy \rightarrow Nicel, sürekli

Bir deęişken için farklı ölçümler

- Bazı deęişkenler için ölçüm tipleri sabittir.
 - Cinsiyet
- Diğer deęişkenler farklı biçimlerde ölçülebilir
 - Ağrı seviyesi

- Şu anda ağrın var mı?
 - Evet / Hayır → Nominal (Kategorik)
- Nasıl bir ağrı hissediyorsun?
 - Zonklama, Keskin, Sürekli → Kategorik
- Şu anda ağrın ne durumda?
 - Yok, az, çok → Sıralı
- Düne göre ağrın nasıl?
 - Aynı, daha az, daha çok → Sıralı

- 1-10 arası bir ölçekte ağrının ne seviyede?
 - Nicel, kesintili



İsim	Cinsiyet	Bölüm	Not
Hakan	Erkek	Sosyal Hiz.	89
Ayşe	Kadın	Hemşirelik	67
Mehmet	Erkek	FTR	49

- Değişkenler?

- İstatistiğin 2 temel kullanımı vardır
 - Betimsel → Özet tablolar, ortalamalar, grafikler.
 - Çıkarımsal → Karar vermek ve öngöründe bulunmak.
(Risk gruplarının belirlenmesi)

Betimsel

TABLE 5 *Examination Grades
for 80 Students*

72	49	81	52	31
38	81	58	68	73
43	56	45	54	40
81	60	52	52	38
79	83	63	58	59
71	89	73	77	60
65	60	69	88	75
59	52	75	70	93
90	62	91	61	53
83	32	49	39	57
39	28	67	74	61
42	39	76	68	65
58	49	72	29	70
56	48	60	36	79
72	65	40	49	37
63	72	58	62	46

- Notları birkaç kelime ile özetleyebilir misiniz?
- Genel olarak sınıfın notları yüksek mi düşük mü?

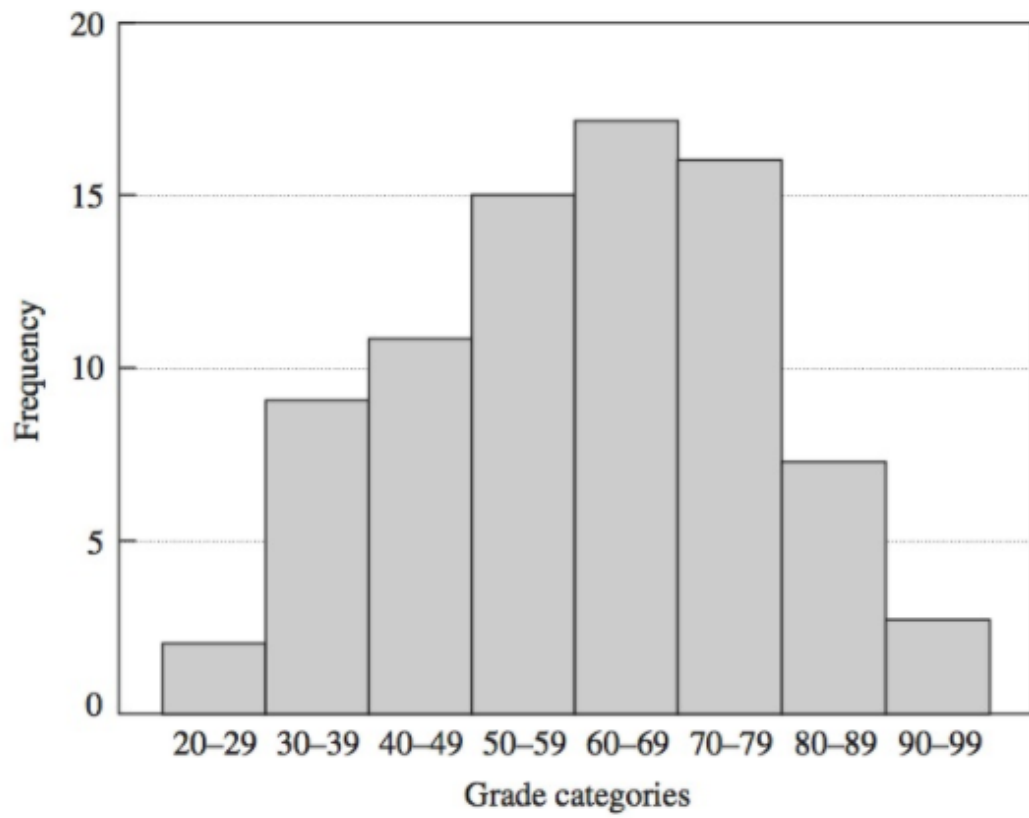
- Betimsel istatistik ile bir grubun özellikleri kolayca görülebilir
- Notlar büyükten küçüğe sıralanabilir.
- Notlar aralık halinde daha küçük gruplara ayrılabilir.

TABLE 6 *Examination Grades for 80 Students: A Grouped Frequency Distribution*

Grades	f
90–99	3
80–89	7
70–79	16
60–69	17
50–59	15
40–49	11
30–39	9
20–29	2

TABLE 5 *Examination Grades for 80 Students*

72	49	81	52	31
38	81	58	68	73
43	56	45	54	40
81	60	52	52	38
79	83	63	58	59
71	89	73	77	60
65	60	69	88	75
59	52	75	70	93
90	62	91	61	53
83	32	49	39	57
39	28	67	74	61
42	39	76	68	65
58	49	72	29	70
56	48	60	36	79
72	65	40	49	37
63	72	58	62	46



- Sınıf genel not ortalaması nedir?
- Bütün notlar toplanıp sınıf mevcuduna bölünerek aritmetik ortalama bulunmasıyla, sınıf diğer sınıflarla kıyaslanabilir.

- Grup frekans dağılımları, grafikler, ortalamalar kullanılarak belirli yönelimleri ve özellikleri tespit edebiliriz.

Çıkarımsal İstatistik

- Genelleme
- Örnekten topluma
- Hipotez testleri

Erkek üniversite öğrencileri marihuana kullanmaya daha mı yatkın?

Marijuana Use	Gender of Respondent	
	Male	Female
Number who have tried it	35	15
Number who have not tried it	<u>65</u>	<u>85</u>
Total	100	100

Marijuana Use	Gender of Respondent	
	Male	Female
Number who have tried it	30	20
Number who have not tried it	<u>70</u>	<u>80</u>
Total	100	100

Marijuana Use	Gender of Respondent	
	Male	Female
Number who have tried it	26	24
Number who have not tried it	<u>74</u>	<u>76</u>
Total	100	100

- Bu soruya evet yanıtı verebilmek için hangi farkı anlamlı kabul edeceğiz?
- Hangi noktadaki fark bunu söyleyebilmek için yeterli?

P Değeri

- P değeri (Probability) → Hipotezimizin geçersiz olması durumunda bile gözlemlediğimiz farka eşit veya daha yüksek bir sonuç elde etme ihtimali.
- Elde edilen sonucun hipotezden bağımsız olarak tamamen rastgele elde edilebilme ihtimali.

- 0-1 arasında ama 0 veya 1 olamaz.
- $0,05 \rightarrow \%5$
- $0.25 \rightarrow \%25$

	İyileşen	İyileşmeyen	Toplam
A tedavisi	56	15	71
B tedavisi	67	30	97
Toplam	123	45	168
	p=0,1565		

	İyileşen	İyileşmeyen	Toplam
A tedavisi	56	15	71
B tedavisi	67	21	97
Toplam	123	45	168
	p=0,1565		

	İyileşen	İyileşmeyen	Toplam
A tedavisi	112	30	142
B tedavisi	134	60	194
Toplam	246	90	336
	p=0,0451		

- p değerinin belirli bir olasılık düzeyinin altında bulunması varsayımın gerçekliğini belirtir.
- Genellikle bu sınır 0,05 olarak kabul edilir.
 - 0,10 – 0,001
- Son tablodaki sonuçlara göre iki tedavi şekli arasında anlamlı bir fark vardır.

- Total kolesterol ve Hb1Ac düzeylerinin değişimleri bağlantılı mıdır?
 - T. Kol. =214,6 mg/dl
 - Hb1Ac=%4,7
 - $r=0,57$
 - $p=0,000043$

IV.Hafta

Veri Sunumu

- Ham verinin liste halinde sunulması → Açıklayıcı değil
- Değişkenlerin tablo ile özetlenmesi → Hızlı çıkarımlar yapılmasına uygun

- Dağılım aralığı \rightarrow Değişkenin en büyük ve en küçük değeri
- Sınıf genişliği \rightarrow Dağılım aralığı / Sınıf sayısı
 - Sınıf sayısı min 5 maks 15
 - Optimum 10

- Frekans (Sıklık) $(f) \rightarrow$ Görülme miktarı
- Yüzde frekans $\rightarrow (f/n) \times 100$

- Birikimli frekans (bf veya F) \rightarrow her sınıfın frekansına bir önceki sınıfın frekansı eklenerek oluşturulan frekanslar toplamı
- Yüzde birikimli frekans $\rightarrow (bf/n) \times 100$

- Notlar 0-100
- 10 sınıf
- $100/10=10$ sınıf genişliği
 - 0-10,10-20,20-30.....90-100

	f	%f	bf	%bf
90-100	2	4,54	2	4,54
80-90	3	6,81	5	11,35
70-80	6	13,63	11	25
60-70	5	11,36	16	36,36
50-60	11	25	27	61,36
40-50	7	15,9	34	77,26
30-40	3	6,81	37	84,07
20-30	4	9,09	41	93,16
10-20	2	4,54	43	97,7
0-10	1	2,27	44	99,97
	44	99,95		

- Sınıf orta değeri \rightarrow Bir sınıfın üst ve alt sınırlarının ortasında kalan değer.
- $(\text{Alt limit} + \text{üst limit}) / 2 = \text{sınıf orta değeri}$

	Sınıf Orta Deęeri
90-100	95
80-90	85
70-80	75
60-70	65
50-60	55
40-50	45
30-40	35
20-30	25
10-20	15
0-10	5

Tablo 3. 60 Kişide cinsiyet, sigara, kan şekeri ve kan basıncı bilgileri.

Denek no	Cinsiyet	Sigara	Açlık kan şekeri	Sistolik kan basıncı	Denek no	Cinsiyet	Sigara	Açlık kan şekeri	Sistolik kan basıncı
1	1	1	95	110	31	2	1	116	127
2	1	2	99	122	32	2	1	98	118
3	1	2	87	124	33	2	1	89	125
4	1	2	106	106	34	2	1	86	120
5	1	2	105	117	35	2	2	105	119
6	1	2	116	118	36	2	2	100	125
7	1	2	95	115	37	2	1	104	123
8	1	1	96	121	38	2	1	99	120
9	1	1	104	111	39	2	1	87	128
10	1	1	111	125	40	2	2	100	119
11	1	2	85	118	41	2	2	101	120
12	1	2	108	108	42	2	2	110	120
13	1	2	91	111	43	2	1	98	121
14	1	2	102	119	44	2	2	98	118
15	1	2	96	115	45	2	1	117	120
16	1	1	106	120	46	2	1	101	118
17	1	2	96	117	47	2	1	103	121
18	1	1	93	126	48	2	1	106	118
19	1	1	100	119	49	2	1	102	118
20	1	2	95	113	50	2	1	99	120
21	1	2	96	116	51	2	1	87	107
22	1	2	101	118	52	2	1	104	112
23	1	2	98	124	53	2	2	94	132
24	1	2	113	120	54	2	1	95	130
25	1	2	106	114	55	2	1	106	114
26	1	2	100	133	56	2	1	106	119
27	1	2	113	120	57	2	1	94	107
28	1	1	89	116	58	2	1	90	108
29	1	2	107	122	59	2	1	98	132
30	1	2	89	119	60	2	2	96	132

(Cinsiyet : 1 = Kadın 2 = Erkek ; Sigara : 1 = İçiyor 2 = İçmiyor)

Tablo 4. Açlık kan şekeri sıklık tablosu

(Sınıf gerçek sınırları) Sınıflar	Salt sıklık f_i	Birikimli sıklık F_i
85 - 89 'den az	5	5
89 - 93 'den az	5	10
93 - 97 'den az	12	22
97 - 101 'den az	12	34
101 - 105 'den az	9	43
105 - 109 'den az	10	53
109 - 113 'den az	2	55
113 - 117 'den az	4	59
117 - 121 'den az	1	60
Toplam	60	60

Tablo 7. Cinsiyet dağılımı.

Cinsiyet	Salt sıklık	Yüzde sıklık
	n	%
Kadın	30	0.50
Erkek	30	0.50
Toplam	60	1.00

Tablo 8. Cinsiyete göre sigara içme durumu.

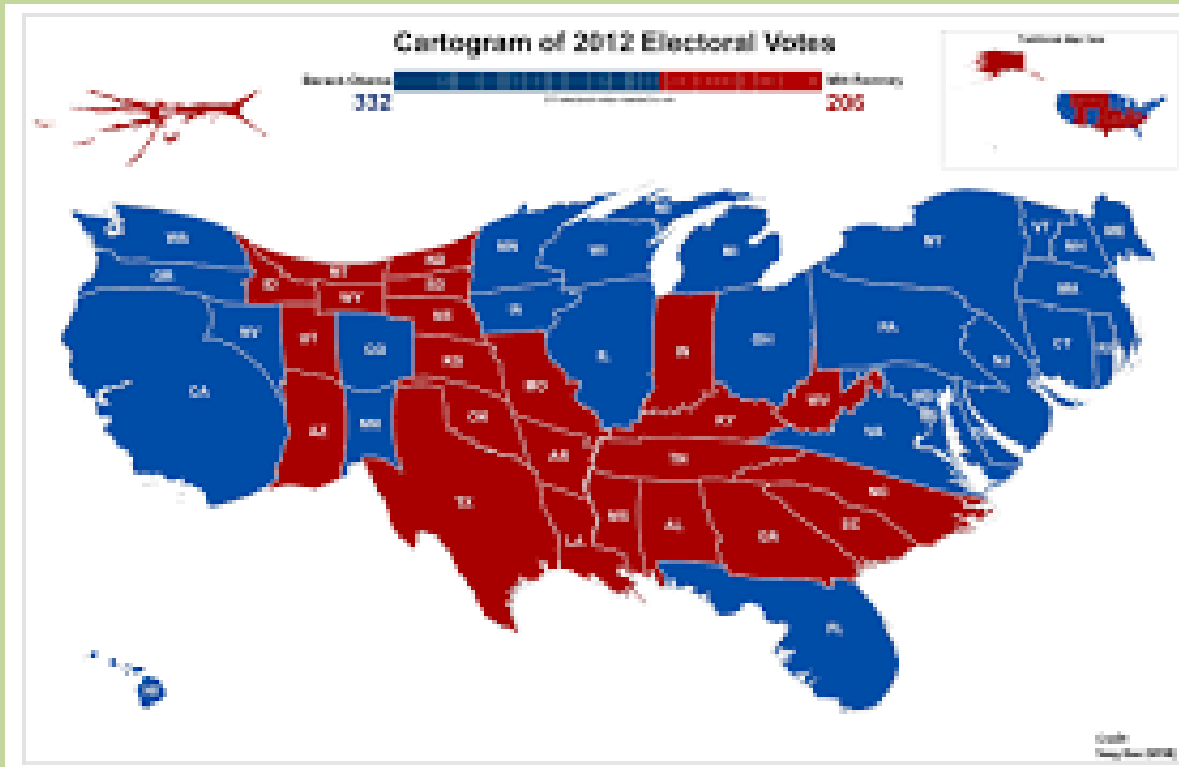
	Sigara içiyor		Sigara içmiyor		Toplam
	n	%	n	%	n
Kadın	8	%26.67	22	%73.33	30
Erkek	22	%73.33	8	%26.67	30
Toplam	30	%50.00	30	%50.00	60

Grafik Sunum

- Tablolar, grafikler ile herkesin anlayabileceği hale gelir
- Olayları grafiklerle karşılaştırmak daha kolaydır
- Grafikler daha ilgi çekici ve akılda kalıcıdır

Kartogram

- Bir deęişkene ait yoğunlukların harita üzerinde gösterilmesi.
 - Sık / seyrek nokta, renk yoğunluğu

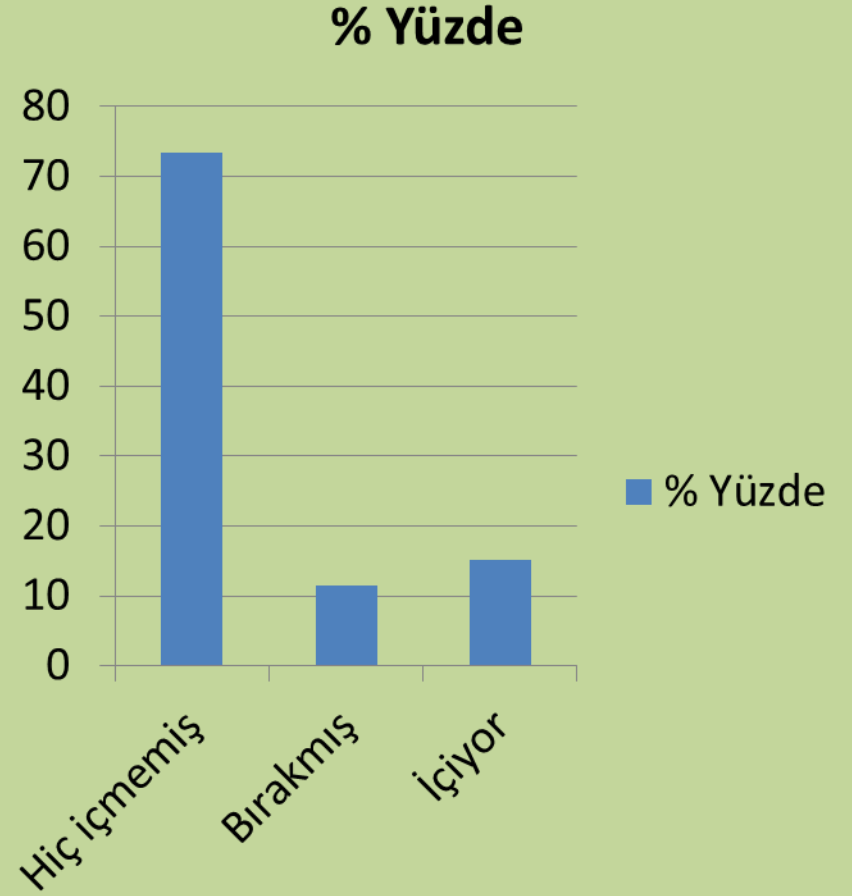


- Nitel ölçüm düzeyi → Çubuk veya pasta grafiği
- Nicel Ölçüm düzeyi → Histogram ve çizgi grafiği
- Sınıf genişlikleri aynıysa → f , % f
- Sınıf genişlikleri farklıysa → F , % F

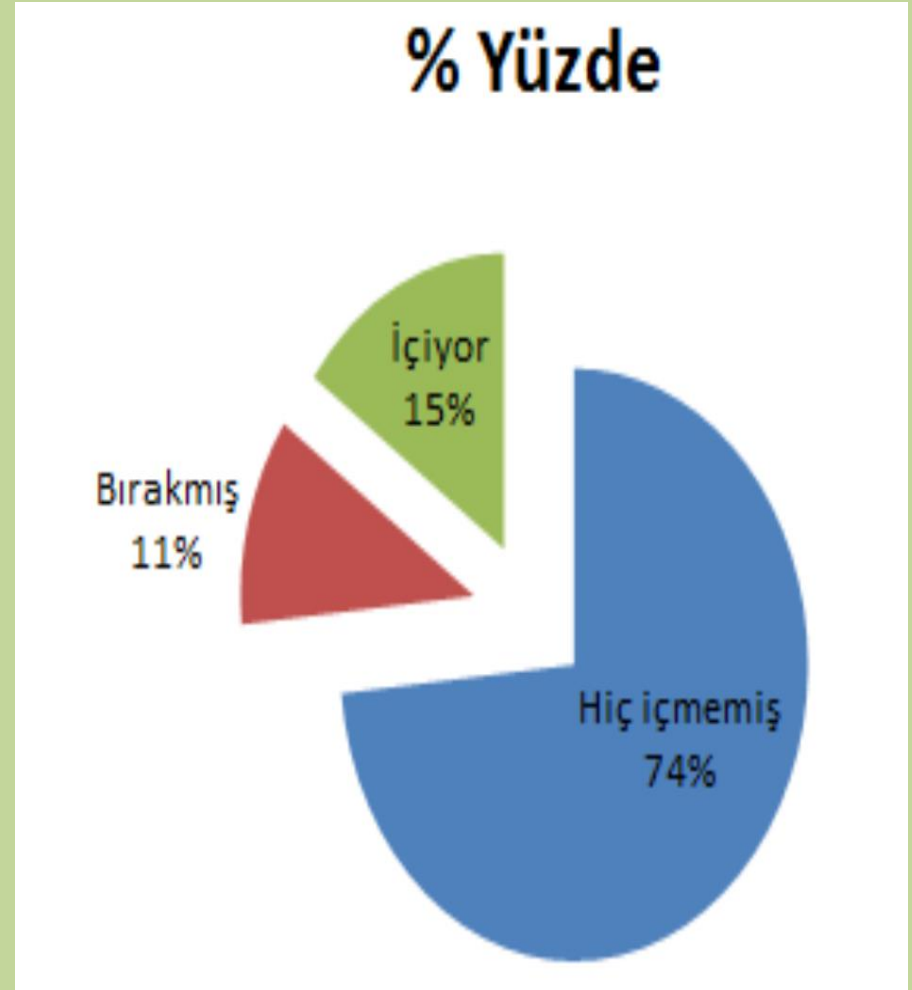
Çubuk Grafikler

- Sınıflamalı (Kategorik) ölçümler için uygundur.
- Her bir kategoriye gösteren çubuklar, kategoriye dahil verilerle orantılı şekilde artar.

Sigara içme durumu	%
Hiç içmemiş	73,46
Bırakmış	11,39
İçiyor	15,15
Toplam	100



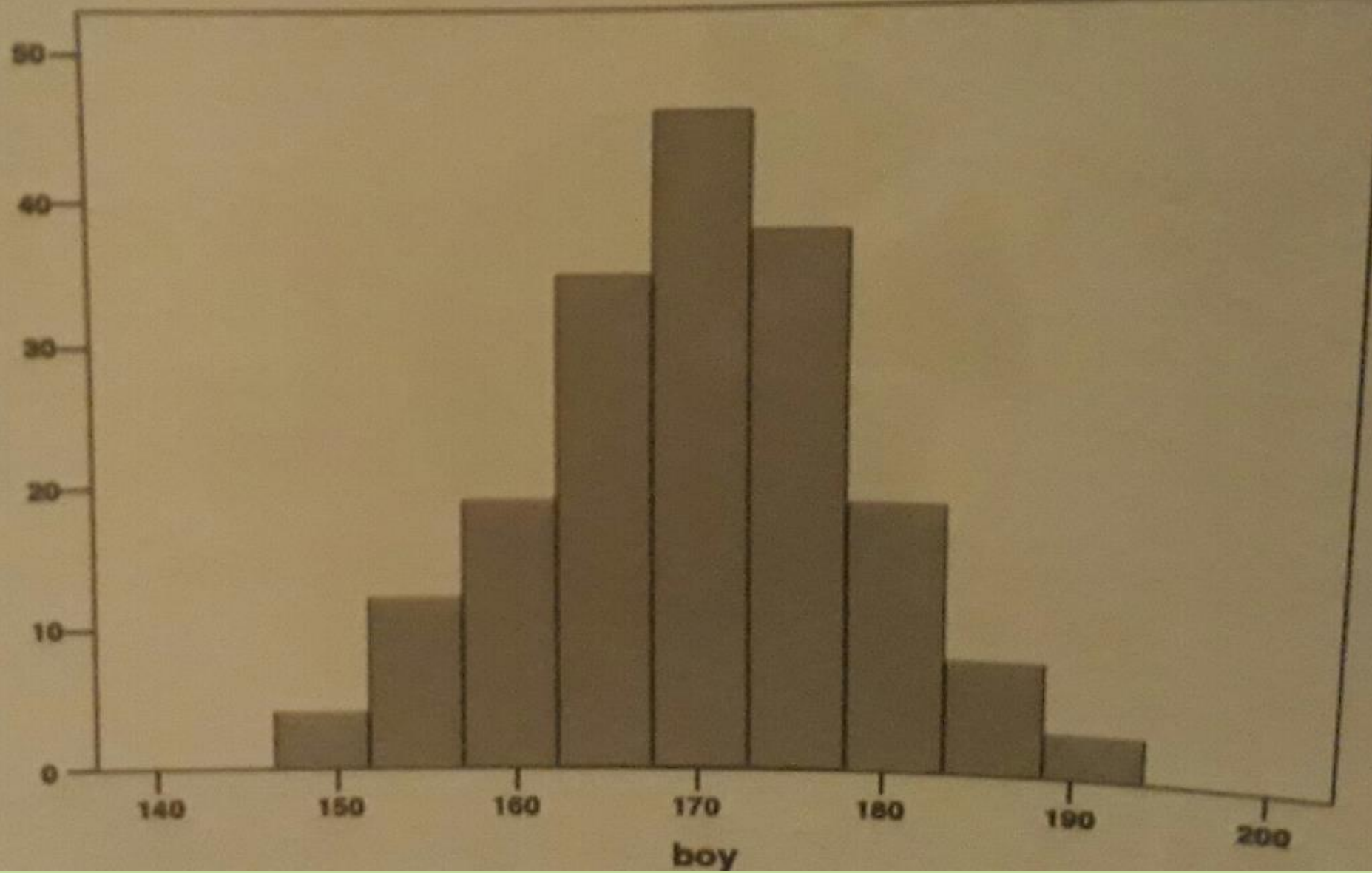
- Aynı veriler pasta grafiđi ile gösterildiđinde, yüzdelere tekabül eden alanlar 360 derece üzerinden oranlanır
- $\%74 \times 360 = 266,40^0$



Tablo 12. Öğrencilerin boy değerleri.

Sınıflar	Sınıf Değeri	Salt Sıklık
147,5 - 152,5 'den az	150	4
152,5 - 157,5 'den az	155	12
157,5 - 162,5 'den az	160	19
162,5 - 167,5 'den az	165	35
167,5 - 172,5 'den az	170	46
172,5 - 177,5 'den az	175	38
177,5 - 182,5 'den az	180	19
182,5 - 187,5 'den az	185	8
187,5 - 192,5 'den az	190	3
Toplam		184

Sıklık



Çizgi Grafikleri

- Yatay ve dikey eksenlerdeki verilerin kesişim noktalarının birleştirilmesi ile oluşturulur.
- Özellikle zaman bağımlı değişimleri göstermek için uygundur.

Tablo 14. 6-11 yaş çocuklarda ortalama sistolik kan basıncı deęerleri (mm-Hg)

Yaş	Ortalama Kan Basıncı (mm-Hg)	
	Kız	Erkek
6	103,4	102,4
7	103,8	103,0
8	106,9	105,9
9	107,8	105,3
10	110,9	107,6
11	114,1	108,1

116,0
114,0
112,0
110,0
108,0
106,0
104,0
102,0
100,0
98,0
96,0

6

7

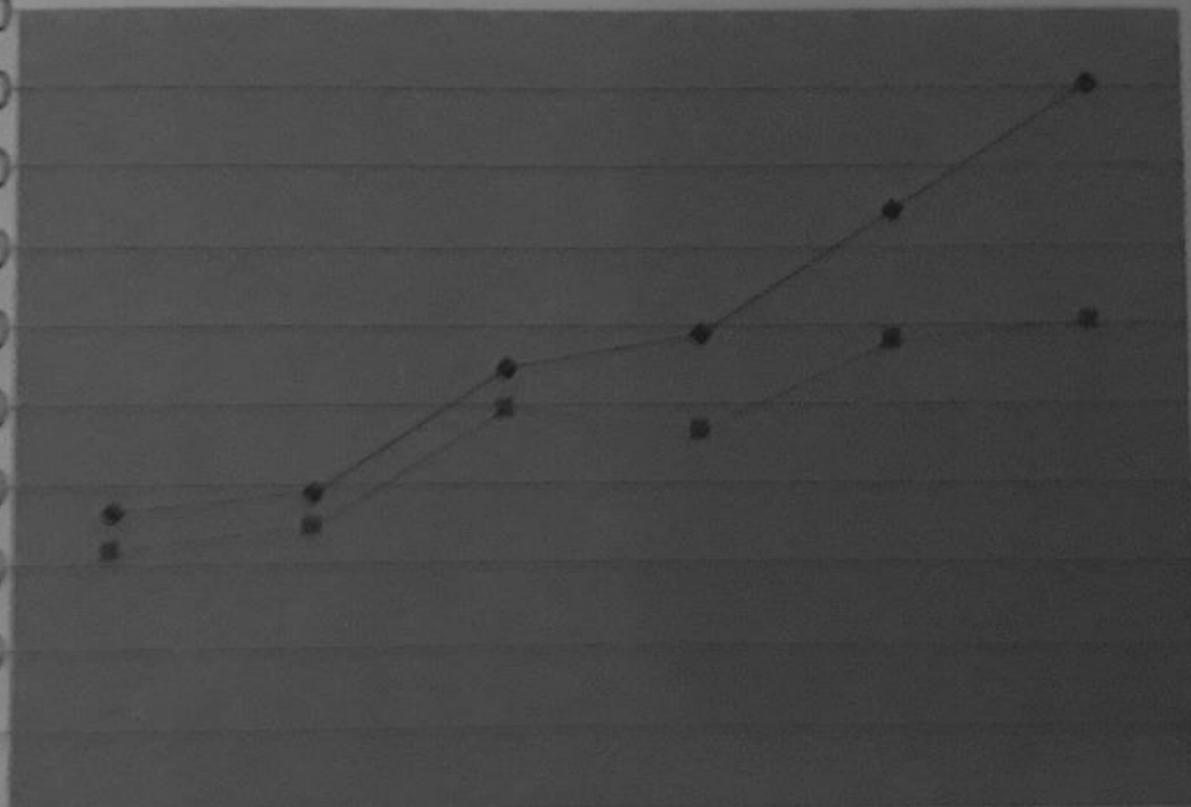
8

9

10

11

◆ Kız
■ Erkek

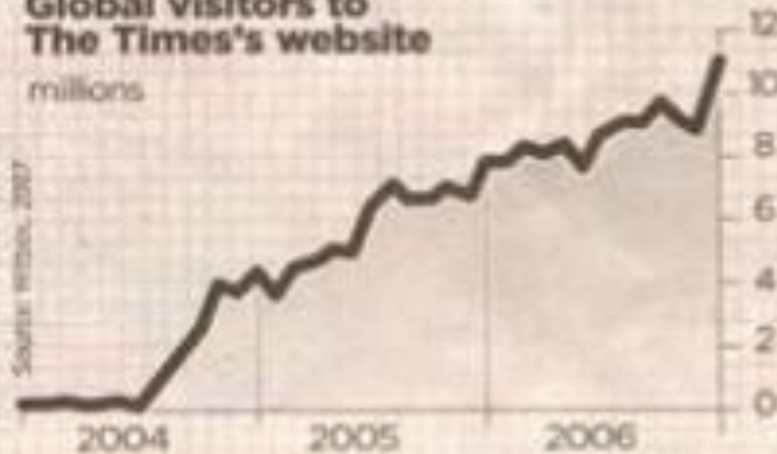


Yanıtıcı Grafikler

The Times leaves the rest behind - in print and online

Global visitors to The Times's website

millions



Full-price sales

485,720



The Times

446,954

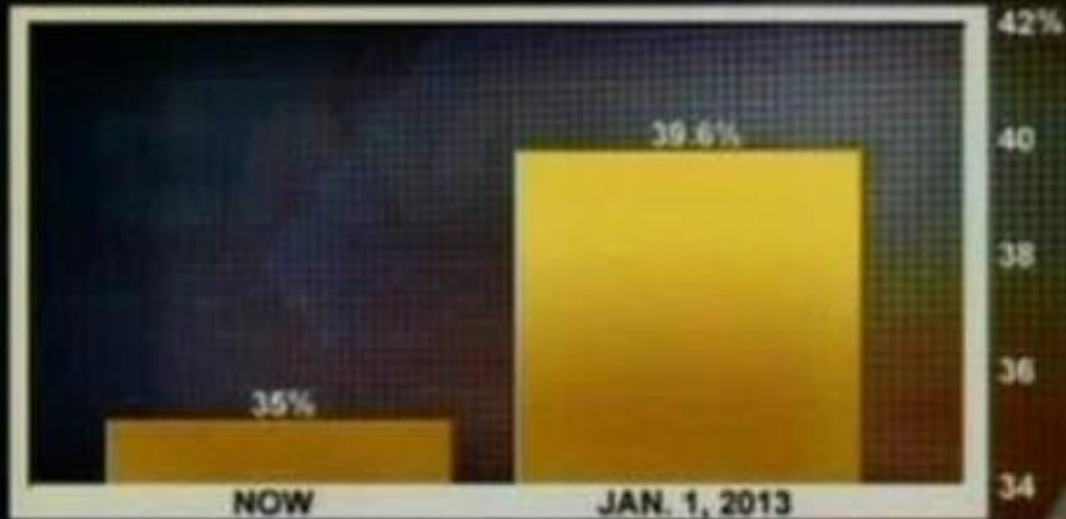
Daily Telegraph

Feb.
row
unio.
staff
unio.
limit
claf.
threat
of the

Jul
A n
inne
the
be a
Plea

IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



8:01 p ET

FOX
BUSINESS

TOP STORIES

TECHNOLOGY

CONSUMER

WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

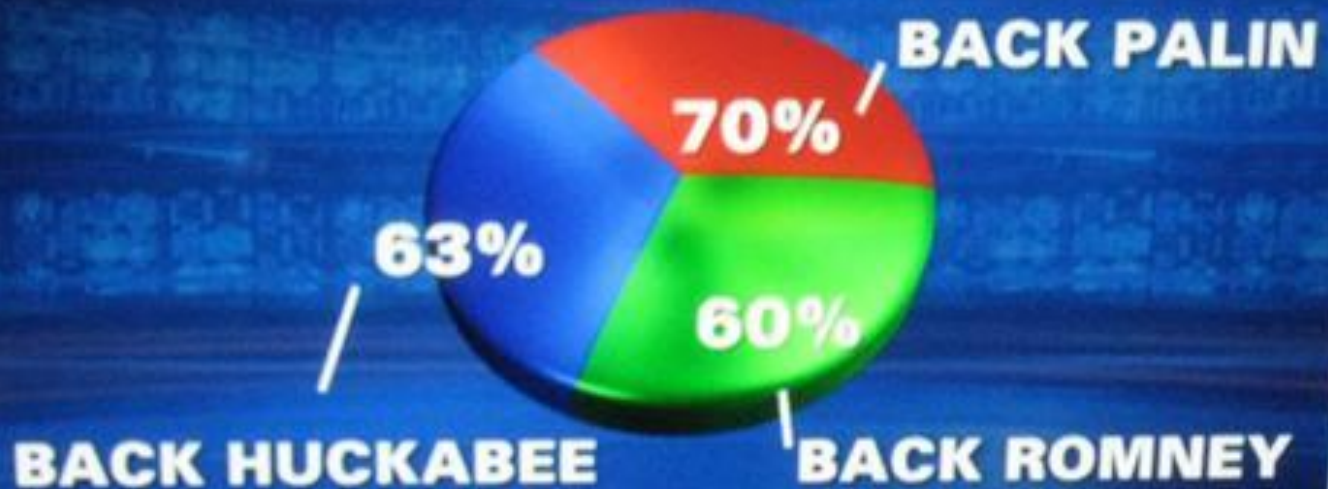
DOW 13008.68 ∇ 64.33

S&P 1379.32 ∇ 5.98

NASDAQ 2939.52 ∇ 6.32

2012 PRESIDENTIAL RUN

GOP CANDIDATES



FOX

47'

SOURCE: OPINIONS

DYNAMIC

Merkezi Eğilim Ölçüleri

- Bir değişkene dair tüm verileri özetleyen değerler.
 - Ortalama,
 - Mod
 - Medyan

Aritmetik Ortalama

- Gözlemlerin sayısal değerlerinin gözlem sayısına bölünmesi ile elde edilir.

\bar{x}

ile gösterilir ve basit serilerde aşağıdaki formül ile hesaplanır.

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

\bar{x} , μ : Aritmetik ortalama.

Σ : Toplama simgesi.

x_i : X deęişkeninin i. gözlem deęeri.

n : Örneklem birim sayısı.

N : Anayığın birim sayısı.

Tablo 1. 13 Kişide sistolik kan basıncı değerleri

Denek no	x_i
1	127
2	118
3	125
4	120
5	119
6	125
7	123
8	120
9	128
10	119
11	120
12	120
13	121

- $127+118+\dots+121=1585$

- $n=13$

- $1585/13=121,92$

Gruplanmış serilerde aritmetik ortalama

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

- Her bir grup için grup orta değeri hesaplanır.
- Orta değer o sınıfın frekansı ile çarpılır.
- Her bir sınıfın orta değer x frekans çarpımı toplanır.
- Toplam denek sayısına (bf, F) bölünür.

Sınıflar ('den az)	x_i	f_i	$f_i x_i$
0 - 10	5	3	15
10 - 20	15	12	180
20 - 30	25	35	875
30 - 40	35	45	1575
40 - 50	45	110	4950
50 - 60	55	45	2475
60 - 70	65	35	2275
70 - 80	75	30	2250
80 - 90	85	15	1275
90 - 100	95	10	950
Toplam		340	16820

- $16820/340=49,47$

Medyan

- Bir seriyi oluşturan elemanlar sıraya dizildiğinde seriyi eleman sayısı açısından ikiye bölen değer. (Med)
- Basit serilerde n tek ise medyan $\rightarrow (n+1)/2$ 'inci terim
- n çift ise $n/2$ 'inci ve $(n+2)$ 'inci terimin aritmetik ortalaması

- 12,24,13,46,23,15,17

- 12,13,15,17,23,24,46 n=7

- $(n+1) / 2 = 4$

- 4. eleman \rightarrow 17

- 12,24,13,46,23,15
 - 12,13,15,23,24,46 $n=6$
 - $n/2 = 3$, $(n+2)/2=4$
 - 3. eleman=15 4.eleman=23 a.o $\rightarrow (15+23)/2=19$
 - Med=19

Gruplanmış Serilerde Medyan

- İlk olarak birikimler frekanslar hesaplanır.
- Ortanca değeri içeren sınıf belirlenir.
 - Örneğin Birikimli frekans toplamı 340 ise yani toplam 340 deney varsa $340/2=170$. deneyi içeren sınıf ortanca sınıfıdır.

Gruplanmış Serilerde Medyan

$$\text{Med} = l + \frac{\frac{n}{2} - F}{f} \cdot xi$$

- l : ortanca sınıfının alt sınırı
- F : l 'den küçük terim sayısı
- f : Ortanca sınıfının frekansı
- i : Ortanca sınıfının genişliği

Sınıflar	Xi	fi	Fi
0 -10	5	3	3
10 -20	15	12	15
20 – 30	25	35	50
30 – 40	35	45	95
40 – 50	45	110	205
50 – 60	55	45	250
60 - 70	65	35	285
70 - 80	75	30	315
80 - 90	85	15	330
90 - 100	95	10	340

$$40 + \frac{\frac{340}{2} - 95}{110} \times 10 = 46,818$$

Mod

- Basit serilerde mod (tepe değeri) en sık tekrarlanan değerdir.
- Sınıflanmış verilerde ilk olarak en sık tekrar eden sınıf belirlenir (frekansı en yüksek) ve aşağıdaki formül kullanılır.

$$\hat{X}$$
$$\text{Mod} = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

l : Mod sınıfının alt sınırı

\triangle_1 : Mod sınıfı frekansı - önceki sınıfın frekansı

\triangle_2 : Mod sınıfı frekansı - sonraki sınıfın frekansı

Alt sınır

Sınıflar ('den az)	x_i	f_i	$f_i x_i$
0 - 10	5	3	15
10 - 20	15	12	180
20 - 30	25	35	875
30 - 40	35	45	1575
40 - 50	45	110	4950
50 - 60	55	45	2475
60 - 70	65	35	2275
70 - 80	75	30	2250
80 - 90	85	15	1275
90 - 100	95	10	950
Toplam		340	16820

Mod sınıfı

Sınıf Aralığı

$$40 + \frac{110-45}{(110-45) + (110-45)} \times 10$$

Mod=45

- Geometrik artış gözlemlenmesi durumunda geometrik ortalama kullanılır.

$$\text{Geo. Ort.} = \sqrt[n]{X_1 X_2 X_3 \dots X_n}$$
$$2, 4, 8 \quad \sqrt[3]{2 \times 4 \times 8} = 4$$

Dağılım Ölçüleri

- Genel olarak bir değişkenin ortalamadan uzaklaşması, değişkenliği gibi durumları belirlemek için kullanılır.
 - Yayılma genişliği
 - Varyans
 - Standart Sapma
 - Değişkenlik Katsayısı

Yayılnma genişliđi

- Basit seri \rightarrow En büyük deđer – en küçük deđer
- Gruplanmıř seri \rightarrow son sınıf OD – ilk sınıf OD

Varyans

- Verilerin ortalamadan sapmalarının karelerinin ortalaması
- Tüm analizlerde deęişkenlik ölçüsü olarak kullanılır.

Basit Seri

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

Gruplanmış Seri

$$s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{n-1}$$

$$(\bar{x} = 49.47).$$

Sınıflar ('den az)	x_i	f_i	$(x_i - \bar{x})$	$f_i (x_i - \bar{x})^2$
0 - 10	5	3	-44,47	5932,7427
10 - 20	15	12	-34,47	14258,1708
20 - 30	25	35	-24,47	20957,3315
30 - 40	35	45	-14,47	9422,1405
40 - 50	45	110	-4,47	2197,8990
50 - 60	55	45	5,53	1376,1405
60 - 70	65	35	15,53	8441,3315
70 - 80	75	30	25,53	19553,4270
80 - 90	85	15	35,53	18935,7135
90 - 100	95	10	45,53	20729,8090
Toplam		340		121804,7060

$$s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{(n - 1)} = \frac{121804.7060}{(340 - 1)} = 359.31$$

Standart Sapma

- Bir veri setinin 'saçılımını' gösteren temel değer
- En sık kullanılan değişkenlik ölçüsü
 - Düşük → Veriler ortalamaya doğru toplanıyor
 - Yüksek → Veriler geniş dağılım gösteriyor
- s ile gösterilir.

$$(\bar{x} = 49.47).$$

Sınıflar ('den az)	x_i	f_i	$(x_i - \bar{x})$	$f_i(x_i - \bar{x})^2$
0 - 10	5	3	-44,47	5932,7427
10 - 20	15	12	-34,47	14258,1708
20 - 30	25	35	-24,47	20957,3315
30 - 40	35	45	-14,47	9422,1405
40 - 50	45	110	-4,47	2197,8990
50 - 60	55	45	5,53	1376,1405
60 - 70	65	35	15,53	8441,3315
70 - 80	75	30	25,53	19553,4270
80 - 90	85	15	35,53	18935,7135
90 - 100	95	10	45,53	20729,8090
Toplam		340		121804,7060

$$s^2 = \frac{\sum_{i=1}^k f_i(x_i - \bar{x})^2}{(n-1)} = \frac{121804.7060}{(340-1)} = 359.31$$

- $S = \sqrt{359,31} = 18,96$

Değişim Katsayısı

- Standart sapmanın aritmetik ortalamaya göre yüzdesel büyüklüğünü gösterir.

$$Dk = \frac{s}{\bar{x}} \times 100 = \frac{18,96}{49,47} \times 100 = 38,33$$