

TEMEL İSTATİSTİK

Frekans Dağılımı, Frekans Tabloları ve Dağılımın Grafikle Gösterilmesi

Prof. Dr. Ezel Tavşancıl

Geçen haftadan hatırlatma...

- İki tür istatistik var: betimsel ve çıkarımsal
- **Betimsel istatistik:** Toplanan verilerin tablolar haline getirilerek ve grafikleri çizilerek düzenlenmesi ve bu yolla kolay anlaşılır hale getirilmesi; gruplandırılarak özetlenmesi; veri grubunun tamamını belli özellikler açısından açıklayacak merkezî yığılma ve değişme ölçülerinin hesaplanması gibi metotların tamamını içermektedir.

Ham Veri

- Araştırmacı tarafından gözlenerek ya da kaydedilerek elde edilen, işlenmemiş, anlamlı hale getirilmemiş sayılar yığınıdır (Balcı, 2015).
- İlk elde edildikleri haliyle düzensiz yığınlar halindedir (Baykul, 1999).
- Düzenlenmemiş verilerdir (Büyüköztürk, Çokluk ve Köklü, 2013).
- Verinin farklı gösterimi süreci aşağıdaki aşamaları içermektedir:
 - Verilerin sıraya konulması
 - Verilerin tablolanması
 - Verilerin gruplandırılması (Frekans [Tabloları](#) için)
 - Grup sayısının saptanması
 - Grup aralık katsayısının hesaplanması

Verinin Farklı Yollarla Gösterimi

- Belirli bir amaç için toplanmış veriyi anlamlı hale getirmenin değişik yolları vardır (Arıcı, 1998):
 - Verileri sözel ifadelerle açıklamak
 - Verileri tablolar halinde düzenleme
 - Verileri grafiklerle gösterme
 - Veriler üzerinde hesaplamalar yaparak istatistiksel ölçüler bulma
 - Yukarıdaki yolların birkaçını birlikte uygulama
- [Hangisinin kullanılacağı, toplanan verinin durumuna ve kullanım amacına bağlıdır.]

BAZI KAVRAMLAR

Frekans Tabloları

- Düzenlenmesinin en basit ve kullanışlı yolu düzenlemek ve frekans tabloları oluşturmaktır. Tablolar, verinin bir bütün olarak okuyucuya sunulmasını sağlar; anlamlandırılmasını kolaylaştırır (Arıcı, 1998; Büyüköztürk ve diğerleri, 2013).
- **Frekans (frequency)**: Bir değerin gözlenme sıklığı, tekrar sayısıdır.
- İstatistiksel verilerin oluşturduğu tablolara çoğu kez **frekans tablosu** denir. Frekans tablosu, ne gibi ölçümlerin gözlendiğini ve bunların her birinin tekrar sayısını, hangi ölçümlerin kaç birey tarafından alındığını gösterir (Arıcı, 1998).

Frekans Tabloları

- **Üç ana sütun var:**

1. Gözlem sonuçları
2. Gözlemlerin gözlenme sıklıkları
3. Değişkenin bir değerine ait yığılmalı frekans ile önceki tüm değerlere ait frekanslar toplanarak bulunur.

! Yığılmalı frekansların toplamı, tüm gözlem sayısına (toplam frekansa) eşit olmalıdır.

! Herhangi bir değere kadar olan toplam frekansı verdiğiinden, değişkenin belli bir değerine göre bireylerin grup içindeki yerini yerlerinin diğerlerine göre karşılaştırılması olanak verir.

Frekans Tablosuna Geçmeden; Oran ve Yüzde

- Araştırmacılar, frekans (f) ve oran (p) kullanmanın yanı sıra puan dağılımlarını yüzde ile de betimlemektedir.

- Oran: $p = \frac{f}{n}$

- Yüzde: $p(100) = \frac{f}{n}100$

Değerin gözlenme sıklığını, toplam gözlem sayısına bölerek oran elde ediyoruz. Bu değeri 100 ile çarparak da yüzde değerine ulaşıyoruz.

ÖRNEK I (Gruplandırılmamış Veriler)

- **Veri Seti (S):** 20 kişilik bir sınıftaki öğrencilerin matematik test puanları

Ham Veri																			
96	90	80	67	60	51	40	30	51	60	60	67	80	90	51	60	67	60	60	51
Büyükten Küçüğe Sıralanmış Veri																			
96	90	90	80	80	67	67	67	60	60	60	60	60	60	51	51	51	51	40	30

- **Öğrencilerin Matematik Puanlarının Frekans, Toplamalı Frekans, Göreli Frekans ve Toplamalı Göreli Frekans Değerleri**

Ölçüm (X)	Frekans (f)	Toplamalı Frekans (t_f)	Göreli Frekans (rel.f)	Toplamalı Göreli Frekans ($t_{rel.f}$)
96	1	20	0.05	1.00
90	2	19	0.10	0.95
80	2	17	0.10	0.85
67	3	15	0.15	0.75
60	6	12	0.30	0.60
51	4	6	0.20	0.30
40	1	2	0.05	0.10
30	1	1	0.05	0.05

NASIL YORUMLANIR?

Grubun %15'i 67 puan almıştır.

Grubun %75'inin puanı 67 ve altındadır.
67 puandan daha yüksek alan $20-(15)=5$ kişi var.

ÖRNEK II...

(Gruplandırılmamış Veriler)

- Öğrencilerin okuma hızı

Tablo 2.1: Öğrencilerin Okuma Hızı (Dakika olarak)

Ham veri

2 6 1 7 5 4 3 3 7 5 5 3 4
2 6 4 4 6 4 5 4 4 5 2 3

Sıralanmış veri

7 7 6 6 6 5 5 5 5 5 4 4 4 4 4
4 3 3 3 3 2 2 2 1

...ÖRNEK II

- Yığmal Frekans Tablosu ($S = \{7, 7, 6, 6, 6, 5, 5, 5, 5, 5, 4, 4, 4, 4, 4, 4, 4, 3, 3, 3, 3, 2, 2, 2, 1\}$, $n=25$)

Tablo 2.2: Öğrencilerin Okuma Hızı Ölçümlerine Ait Frekans, Toplamalı Frekans, Görel Frekans Ve Toplamalı Görel Frekans

Ölçüm (X)	Frekans (f)	Toplamalı Frekans (t_f)	Görel Frekans (rel.f)	Toplamalı Görel Frekans ($t_{rel.f}$)
7	2	25	.08	1.00
6	3	23	.12	.92
5	5	20	.20	.80
4	7	15	.28	.60
3	4	8	.16	.32
2	3	4	.12	.16
1	1	1	.04	.04

Gruplandırılmış Veriler

- Toplanan verilerin fazla ve ranjinin geniş olması durumunda ham puana dayalı bir frekans tablosu hazırlaması zordur.
- Bu durumda frekans tablosunda veriler gruplandırılarak düzenlenir.
- Veriler gruplandırıldığında dağılımın özgünlüğü bozulmakta ve veri kaybı olmaktadır. Ancak bilgisayar yazılımlarının kullanılmadığı durumlarda kullanılması pratik nedenlerle önemlidir.
- Verilerin Gruplandırılmasında iki varsayım söz konusudur:
 - Bir aralığa rastlayan ölçümler aralığın gerek sınırları boyunca eşit olarak dağılır.
 - Bir aralığa rastlayan ölçümlerin tümü aralığın orta noktası üzerinde yığılır (Arıcı, 1998).

Gruplandırılmış Veriler

- Gruplama, verileri aralıkları eşit olacak şekilde düzenleyip bu aralıklara düşen frekansları saptayarak özetlemektir. Aslında frekans tablosu da özel bir gruptur.
- Gruplandırma için yapılması gereken üç işlem vardır:
 - 1. Grup sayısını saptamak:** Frekans tablosundaki en büyük ve en küçük değer dikkate alınarak önce verilerin kaç grupta toplanacağı belirlenir. Bu karar kesin değildir, aralık katsayısına bağlı olarak değiştirilebilir.

2. Aralık katsayısını hesaplamak: Grup sayısı saptandıktan sonra aralıkların genişliği bulunur. Bütün grupların aralık genişliği eşit olmalıdır.

GRUP ARALIK KATSAYISI: En büyük ve en küçük gözlem arasındaki farkı grup sayısına bölünmesi ile hesaplanmaktadır.

$$a = \frac{X_{EB} - X_{EK}}{gs} \quad (X_{EB}: \text{En büyük gözlem, } X_{EK} = \text{En küçük gözlem; } gs: \text{ grup sayısı})$$

- Aralık katsayısının tek olması tercih edilen bir durumdur.
- Gereğinden küçük ya da büyük alınmaması önemli:
 - Büyük olursa: Bilgi kaybına yol açar.
 - Küçük olursa: Grup sayısı artar, dağılımın hangi aralıkta dolayısıyla hangi değer etrafında toplandığı görülemez. Verilerin özetlenmesinde beklenen fayda sağlanamaz.

3. Grupların Oluřturulması:

- Gruplar, aralık katsayısı bulunduktan sonra, en küçük veya en büyük gözlemden başlanarak gruplandırılır.
- Üç nokta önemlidir:
 - Grup aralıkları eşitliđi
 - Bir verinin yalnız bir gruba dahil olması
 - Gruplamaya olabildiđince çok verinin aralıđın orta noktasına geleceđi deđerden başlanması

(Baykul, 1999)

- Aşağıdaki veri setini düşünelim:

Ham Veri																			
96	90	80	67	60	51	40	30	51	60	60	67	80	90	51	60	67	60	60	51
Büyükten Küçüğe Sıralanmış Veri																			
96	90	90	80	80	67	67	67	60	60	60	60	60	60	51	51	51	51	40	30

- En küçük ve en büyük değerler: 96 ve 30
- Bu iki değer arasındaki fark (ranj: $96-30=66$)
- 6 gruplu bir sınıflandırma yapmak için: $66/6=11$
- Tam sayı olmasaydı, yuvarlama yapılırdı.
- İlk puan aralığı sınır değerleri: 30 ile 40 [$30+(11-1)$] olacak şekilde örüntü devam ettirilir.
- 30-41 arasında 2 gözlem var: 30 ve 40

ÖRNEK III (Gruplandırılmış Veriler)

Tablo 2.3: Bir Grup Öğrencinin Başarı Testi Puanları

09	35	33	57	56	51	75	87	98	78	77	75	65	67	63	61	68	43	41	48	48	23	25
11	17	10	21	29	27	26	34	35	33	31	52	55	53	58	77	76	71	79	72	73	93	83
87	86	84	88	20	20	21	22	43	44	45	49	55	56	53	51	58	48	42	30	30	31	37
38	39	41	40	69	68	61	60	31	30	33	37	56	66	69	64	59	55	54	56	41	43	45
48	69	64	40	43	53	55	58															

- Veri puanlarını dağılımı $98-9=89$ 'dur. Tahmini grup sayısı 10 olarak belirlenirse ranjı 10'a bölerek aralık katsayısını hesaplayabiliriz.
- $a = \frac{89}{10} = 8.9$; tek sayı tercih edildiğinden > 9

Tablo 2.4: Tablo 2.3'deki Veriler İçin Tahmini Grup Sayısına Karşılık Gelen Aralık Katsayıları

Grup sayısı (g)	Aralık katsayısı (a)	Çarpım (gxa)	(gxa) > Ranj
5	18	90	90>89
10	9	90	90>89
13	7	91	91>89
15	6	90	90>89

Tanis (1987) yaklaşımı: Grup sayısı ile aralık katsayısının çarpımı ranji örten en düşük sayı olmalıdır.

Tablo 2.5: Başarı Testi Puanlarına Ait Gruplandırılmış Frekans Dağılımı

Puan Aralığı	f	rel.f	Orta Nokta	Gerçek sınırlar	Toplamalı f	Toplamalı rel.f
90-98	2	.02	94	89.5-98.5	100	1.00
81-89	6	.06	85	80.5-89.5	98	.98
72-80	9	.09	76	71.5-80.5	92	.92
63-71	12	.12	67	62.5-71.5	83	.83
54-62	17	.17	58	53.5-62.5	71	.71
45-53	13	.13	49	44.5-53.5	54	.54
36-44	15	.15	40	35.5-44.5	41	.41
27-35	14	.14	31	26.5-35.5	26	.26
18-26	8	.08	22	17.5-26.5	12	.12
09-17	4	.04	13	08.5-17.5	4	.04

Grup
(puan)
aralığı

Gruplandırılmış Veri İle Çalışmak

- Orijinal verinin bilgi kaybı söz konusu ama yine başvurmak gerekli.
- Bir grup aralığının teorik olarak alt ya da üst sınırı yoksa buna **açık grup aralığı** denir. (65 yaş ve üzeri)
- Puanlar, en yakın tama tamamlanarak kaydedilmişse, 54-62 grup aralığı, teorik olarak 53.5 ve 62.5'e kadar tüm ölçüleri içerir. Bulunan bu yeni değerler grubuna **gerçek sınır değerleri** denir. Gerçek sınır değerleri, pratikte, bir grup aralığının üst sınır değeri ile bir sonraki grup aralığının alt sınır değeri toplanıp ikiye bölünerek hesaplanır.
- Bir grup aralığının genişliği (aralık katsayısı), o grubun gerçek alt ve üst gerçek sınırı arasındaki farka eşittir.
 - 54-62 grubunun aralık katsayısı, $62.5-53.5=9$ 'dur.
- **Orta nokta**, bir grup dağılımının ortalama değeridir. Orta nokta, bulunduğu grup aralığının üst ve alt sınır değerleri toplamının yarısıdır.

(Büyüköztürk ve diğerleri, 2013)

Verilerin Grafikle Gösterilmesi

- Grafikler, nicel verileri görsel formlara dönüştürerek bu verilerde kolayca fark edilemeyen ilişkilerin görülmesini olanaklı kılar (Ravid, 1994).
- Grafik> Veriyi anlamlı olarak özetlemenin yolu
- Grafiklerle göstermenin de çeşitli yolları var: Histogram, sütun grafiği, poligon, frekans dağılım eğrisi, dal-yaprak diagramı, interpolation vb. (Gravetter & Wallnau, 2012).
- Uygun olarak hazırlanan bir görsel 10 000 kelime kadar değerlidir (Guilford, 1956).

Histogram

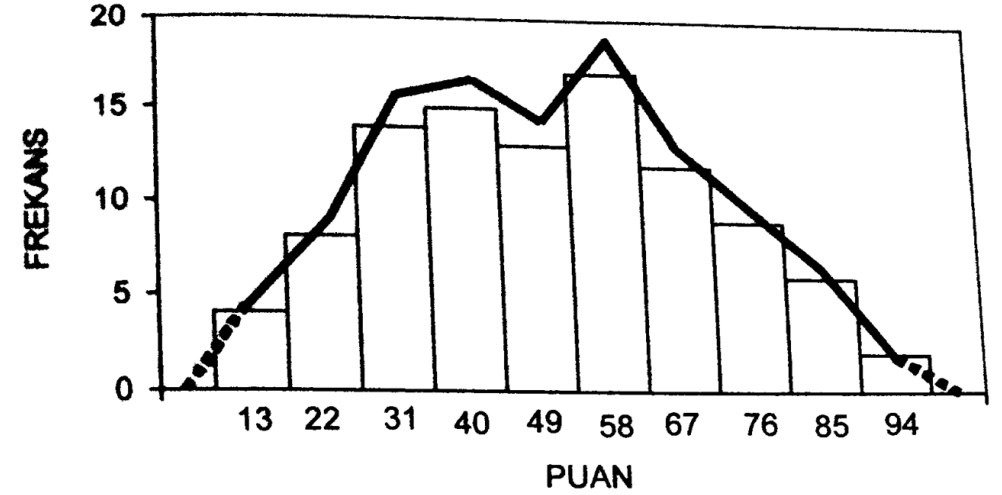
- Dikey eksen frekansları, yatay eksen de ölçümleri/gözlemleri/puanları ya da puan aralıklarını göstermektedir.
- İki eksenin kesiştiği nokta her iki eksen için de en düşük değeri gösterir. Yatay eksen, her zaman 0'dan başlamak zorunda değildir.
- Sürekli değişkenler için uygun olan görsel formdur.
- Histogram, bar diagramı ya da çizgi grafiği çizilirken genellikle bir değişkenin düzeyleri ya da puanlar yatay ekseninde (X), bu düzey ya da puanlara karşılık gelen frekanslar dikey ekseninde (Y) gösterilir.
- Örn: Yatay eksen öğrencilerin test puanları, dikey eksen frekanslar
- Grafikteki sütunların orta noktaları, ilgili puan aralığının orta noktasına karşılık gelir. Bu noktalar birleştirilerek **frekans poligonu** elde edilir. Bu çizginin başlangıç ve bitişi noktalarla gösterilmektedir.

(Baykul, 1999; Büyüköztürk ve diğerleri, 2013)

Histogram Örneği

Tablo 2.5: Başarı Testi Puanlarına Ait Gruplandırılmış Frekans Dağılımı

Puan Aralığı	f	rel.f	Orta Nokta	Gerçek sınırlar	Toplamalı f	Toplamalı rel.f
90-98	2	.02	94	89.5-98.5	100	1.00
81-89	6	.06	85	80.5-89.5	98	.98
72-80	9	.09	76	71.5-80.5	92	.92
63-71	12	.12	67	62.5-71.5	83	.83
54-62	17	.17	58	53.5-62.5	71	.71
45-53	13	.13	49	44.5-53.5	54	.54
36-44	15	.15	40	35.5-44.5	41	.41
27-35	14	.14	31	26.5-35.5	26	.26
18-26	8	.08	22	17.5-26.5	12	.12
09-17	4	.04	13	08.5-17.5	4	.04

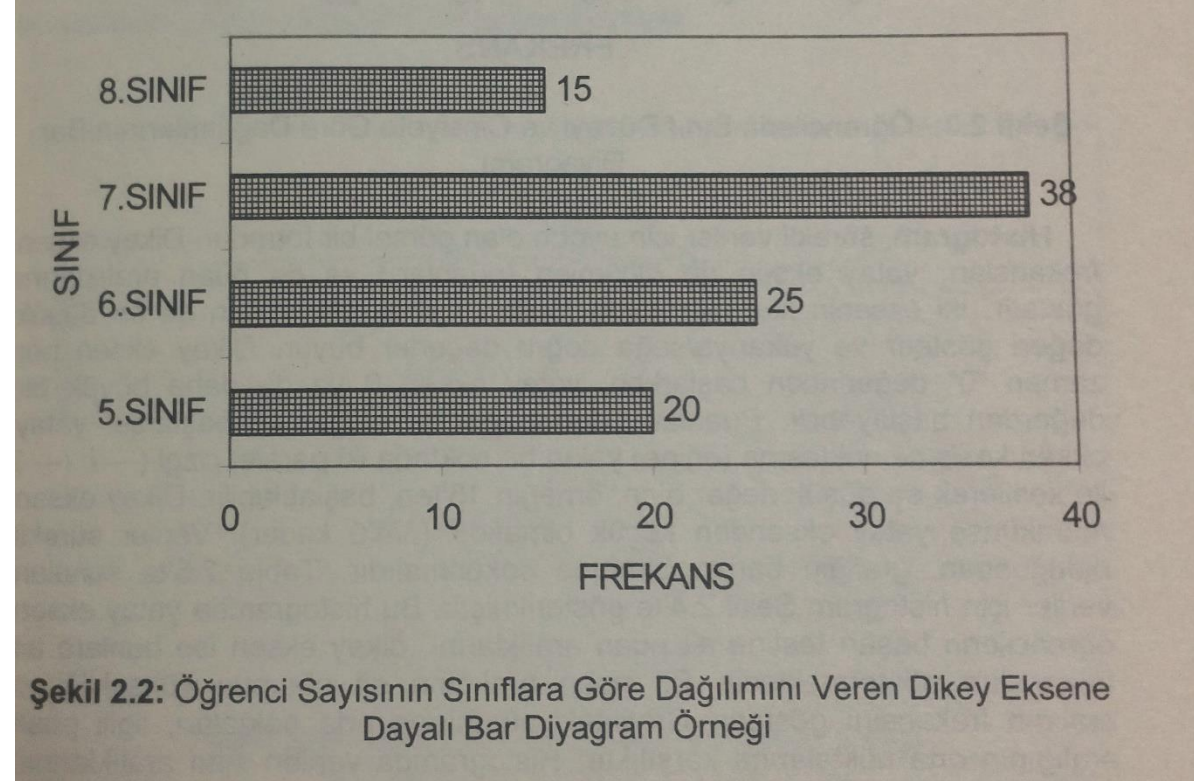


Şekil 2.4 Tablo 2.5'deki Veriler İçin Histogram

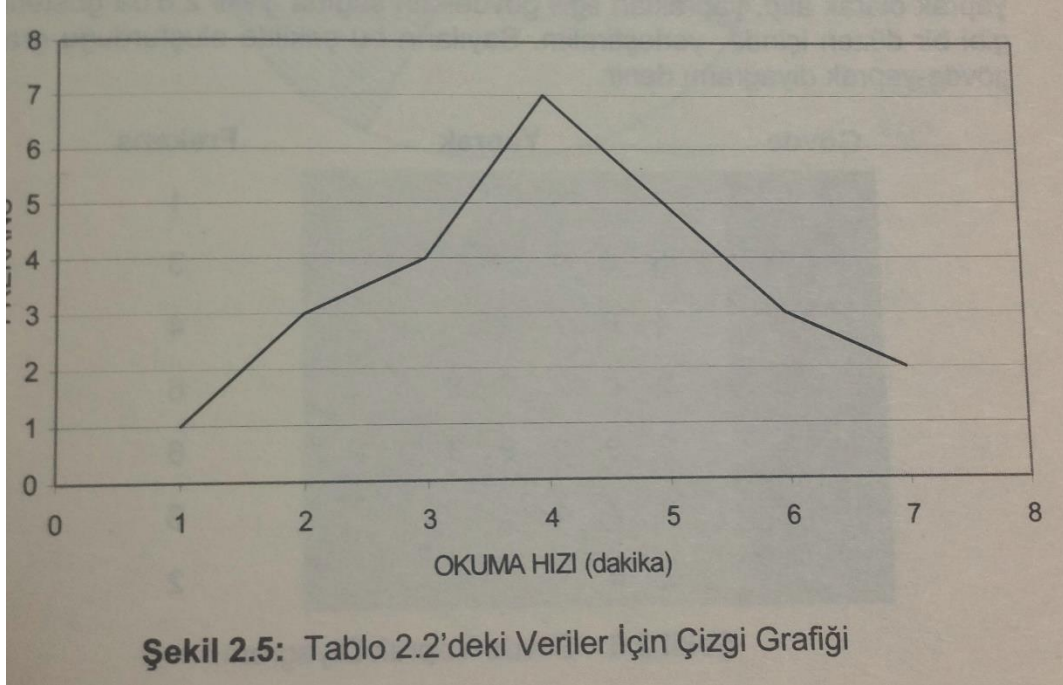
(Büyüköztürk ve diğerleri, 2013)

Sütun Grafiği

- **Sütun grafiği (Bar Diagramı):** Birbirini izleyen barların bir serisini gösterir.
- Barlar birbirine dokunmazlar ve genellikle düşükten yükseğe ya da tersine sıralanırlar.
- Bu tür grafikler süreksiz ya da kategorik değişkenler için uygundur.
- Bir eksen değişkenin düzeyini, diğeri ise frekans ya da yüzdelerini gösterir.



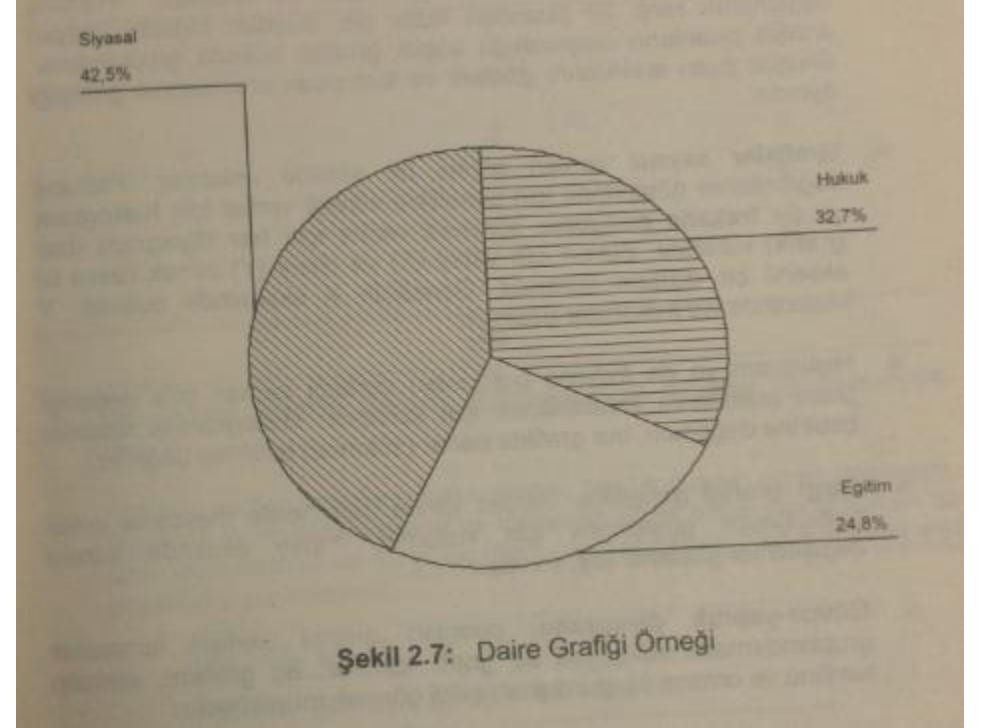
Çizgi Grafiđi



- Srekli bir deđiřkenin deđerlerine karřılık gelen frekansların dađılımını gstermek amacıyla ok yaygın olarak kullanılmaktadır.
- Kırılma noktalarının y deđerleri toplamı?

Daire Grafiđi

- Bir bütnn parçalarının frekans ya da oran olarak gösterildiđi grafik türdr.
- zellikle kamuoyu arařtırmalarının sunulmasında sıklıkla kullanılmaktadır.
- Sınıflamalı deđiřkenin dzeylerini, toplam içindeki ađırlıkları bakımından gstermek amacıyla da kullanılır.
- RN: Arařtırmaya katılan đrencilerin %42.5'i siyasal, %32.7'si hukuk ve %24.8'ieđitim bilimleri alanında đrenim grmektedir.



(Bykztrk ve diđerleri, 2013)

Daire Grafiđi

Yapılandırmasında Açı İlişkisi

(Baykul, 1999)

Örnek: 1986 yılında ülkemizden ABD, Japonya, Almanya, Fransa, İngiltere, İtalya ve diđer ülkelere yapılan ihracatın miktarları ABD doları olarak Tablo 3.13'te görüldüğü gibidir.

Bu verileri gösteren bir daire grafiđi Şekil 3.17'de görüldüğü gibi çizilebilir.

Bu grafiđin çizilmesinde bir örnek olmak üzere Almanya'ya olan ihracata ait merkez açığı hesaplayalım. Bunun için Almanya'ya olan ihracatın toplam ihracata oranını

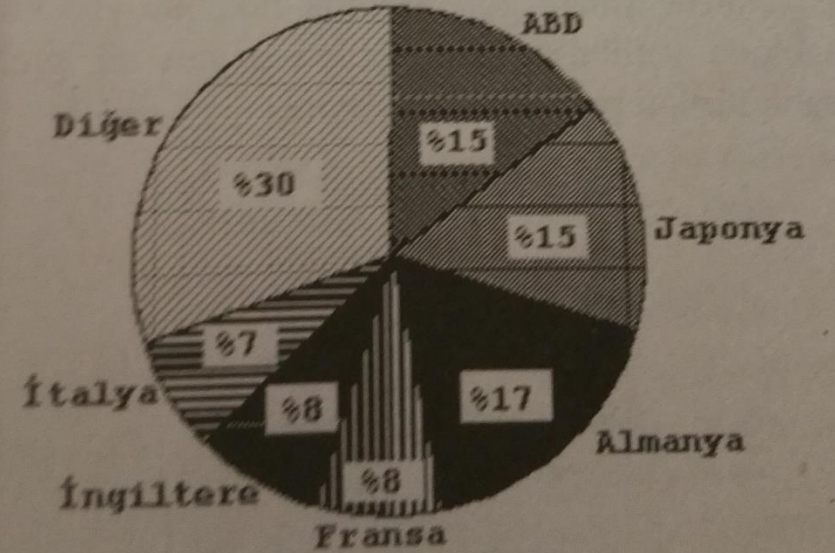
$$\frac{242411}{1422652} = 0,17$$

biçiminde bulup bu değeri 360 ile çarpılarak $17 \times 360 = 61^\circ$ bulunur.

1986 YILI İHRACATIMIZIN
ÜLKELERE GÖRE DAĞILIMI

Ülke	Ihracat (\$)	%	Açı(α)
ABD	217336	15	54°
Japonya	209153	15	54°
Almanya	242411	17	61°
Fransa	119430	08	29°
İngiltere	107013	08	29°
İtalya	97835	07	25°
Diđer	429474	30	108°
Toplam	1422652	100	360°

Kaynak: DİE, Türkiye İstatistik Cep Yıllığı, 1991.



Şekil 3.17: Daire Grafiđi

Yaprak-Gövde Diagramı

- Herhangi bir deęişkene ilişkin olarak elde edilen verilerin genel olarak betimlenmesinde kullanılan bir tekniktir.
- Bu diagram, orijinal veri kaybı olmaksızın verileri gruplamada kullanılan açıklayıcı bir veri analizi teknięi olarak tanımlanabilir (Elifson, Runyon ve Haber, 1990).
- Başta gelen sayı ya da sayılar gövde, daha sonra gelen sayı ya da sayılar yaprak olarak isimlendirilir ve düzenlenir. Böylece sayıların kendisi veri grafięi olur.
- Genelde gözlem deęerlerindeki son sayı yapraklar olur.
- Örneęin, golf oyununda yer alan 25 oyuncunun ařaęıda verilen ortalama vuruř uzunlukları için bir gövde yaprak diagramı çizelim:

- Bu sayıların ilk basamağını, yani 22, 23, 24, 25, 26, 27 ve 28 sayılarını gövde olarak alalım. Daha sonra üçüncü basamaktaki sayıları da yaprak olarak alıp, yaprakları ilgili gövdelerin sağına düzen içinde yerleştirelim. Sayıların oluşturduğu şekil, yaprak-gövde diagramıdır.

227	244	246	278	262	252	269	260	247	277	250	235	274	257	282	269
255	263	236	289	258	231	255	241	261							

Gövde	Yaprak	Frekans
22	7	1
23	5 6 1	3
24	4 6 7 1	4
25	2 0 7 5 8 5	6
26	2 9 0 9 3 1	6
27	8 7 4	3
28	2 9	2

Şekil 2.6: Gövde-Yaprak Diyagramı

(Büyüköztürk ve diğerleri, 2013)

Değerlendirme

1. Eksen göstergelerini dikkate alındığında, *histogram* ile *sütun grafiği* arasındaki farkın ne olduğu söylenebilir?
2. Dikey kullanılan bir histogramda sütunların yüksekliği neyi göstermektedir?
3. Frekans poligonu içeren histogramdan sütunlar kaldırıldığında elde edilen grafik türü nedir?
4. Örneklemdeki alt gruplara dağılımı göstermek için kullanılan daire ve sütun grafiği arasındaki fark için ne söyleyebilirsiniz?
5. Histogramda, her bir sütunun yüksekliği ve genişliği için ne söylersiniz?

Yanıtlar

1. Bar grafiđi daha çok kategorik deđişkenler için uygundur, histogram ise süreklilik gerektirir. Bu nedenle ilkinde sütunlar ayrıken ikincisinde birbirine dokunmalıdır. Başka bir deyişle sütun grafiđinde gruplar aralıkları kesikli; histogramda ise süreklidir.
2. Frekans
3. Çizgi grafiđi
4. İlki, tüm grubu bir bütün olarak deđerlendirip alt grupları tüm grup içindeki oranı hakkında bilgi verirken; ikincisi her bir alt grubu tek başına sıklıkları ele alır, tüm gruba deđil birbirine göre kıyaslama olanađı verir.
5. Dikdörtgen sütunların yüksekliđi frekans deđerlerini gösterir ve her bir deđer aralığında deđişebilir. Genişliđi ise aralık katsayısıdır ve her grup aralıđı için sabittir.

KAYNAKLAR

- Arıcı, H. (1998). *İstatistik: Yöntemler ve uygulama*. Kendi Yayını.
- Balcı, A. (2015). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler*. Ankara: Pegem Akademi.
- Baykul, Y. (1999). *İstatistik: Metodlar ve uygulamalar*. Ankara: Anı Yayıncılık.
- Büyüköztürk, Ş., Çokluk, Ö. ve Köklü N. (2013). *Sosyal bilimler için istatistik*. Ankara: Pegem Akademi.
- Gravetter, F. J., & Wallnau, L. B. (2012). *Statistics for behavioral sciences*. Belmont, CA: Wadsworth.
- Guilford, J. P. (1956). *Fundamental statistics in psychology and education*. U.S.A.: McGraw-Hill.
- Ravid, R. (1994). *Practical statistics for educators*. New York: University of press of America Inc.