

## II. BÖLÜM

### 2. İSTATİSTİK SERİLERİ VE FREKANSLAR

#### 2.1. İstatistik Serileri

İstatistik serileri gözlem değerlerinin büyüklüklerine göre sıralanmasıyla oluşturulur. İstatistik serileri ilgili yığın olayın kavranması açısından etkin bir araç olup, istatistik analizlere temel oluşturur. Değişik ölçütler temel alınarak istatistik serisiyle ilgili farklı sınıflandırmalar yapılabilir.

##### 2.1.1. Zaman ve Mekan Serileri

Eğer gözlem sonuçları yıl, ay, hafta, gün yada saat gibi bir zaman değişkeninin şıklarına göre sıralanırsa oluşturulan seriye **zaman serisi** adı verilir. Yıllara göre ülke nüfusları, günün belli saatlerindeki trafik yoğunluğu, aylara göre ortalama sıcaklık değerleri zaman serisine örnek verilebilir.

**Örnek 2.1.** Eskişehir ili aylara göre 2000 yılı 6 aylık ortalama sıcaklık değerleri (C°) aşağıdaki gibidir.

Ocak	Şubat	Mart	Nisan	Mayıs	Haziran
-1,5	1,3	4,9	10,4	15,1	18,8

Eğer gözlem sonuçları ülke, bölge, şehir yada köy gibi mekan değişkeninin şıklarına göre sıralanırsa, elde edilen seriye **mekan serisi** adı verilir. Şehirlere göre elektrik tüketimi, bölgelere göre tahıl üretimi bu tür serilere örnek verilebilir.

**Örnek 2.2.** Bazı illerin denizden yükseklikleri (m)

Ankara	İzmir	Kars	Muğla	Rize	Sivas	Trabzon	Van	Gaziantep
891	29	1755	646	9	1285	30	1661	855

##### 2.1.2. Dağılım Serileri

Gözlem sonuçlarının maddesel bir değişkenin (zaman ve mekan değişkenleri dışındaki değişkenler) şıklarına göre sıralanmasıyla oluşturulan serilere dağılım serileri adı verilir. Dağılım serileri Basit Seri, Frekans Serisi ve Gruplandırılmış (Sınıflandırılmış) Seriler olarak üç farklı biçimde gösterilir.

### i) Basit Seri

İlgilenilen değişkenin değerlerinin basit (sade) biçimde gösterilmesidir. Eğer derlenen veriler ilgilenilen konunun dışında başka bir yönde örneğin gözlem sırasına göre sıralanmış ise, bu sıralamaya **liste** adı verilir.

**Örnek 2.3.** 20 öğrencinin istatistik notları aşağıdaki gibidir.

70,60,45,85,72,61,38,90,75,78,38,45,45,72,72,90,72,90,45,70

Eğer liste belirtilen amaçlar doğrultusunda düzenlenirse sonuçlara daha kısa sürede ulaşılır. Böyle bir sıralama (küçükten büyüğe doğru sıralama gibi) sonucu elde edilen seriye **Basit Seri** adı verilir. Verilen notları daha kolay anlaşılması için liste halinde küçükten büyüğe doğru yazalım.

38	38	45	45	45	45	60	61	70	70	72	72	72	72	75	78	85	90	90	90
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

**Soru.** Bir bölgedeki 5 binanın değerleri (TL) olarak aşağıdaki gibi verilmiş olsun.

Bina değeri (TL): 125000 130000 135000 145000 160000

Bu seri basit seridir, çünkü değişkenimiz olan bina değerinin sadece fiyatı verilmiştir.

### ii) Frekans Seri

Verilerin daha kolay anlaşılması için gözlem değerlerinin yanına kaç kez tekrarlandığı kaydedilerek oluşturulan seriye **frekans serisi**, tekrarlara da frekans denir.

**Örnek 2.4.** Örnek 2.3'deki basit seriyi frekans serisi biçiminde gösterelim.

Notlar	38	45	60	61	70	72	75	78	85	90	Toplam
Frekans	2	4	1	1	2	4	1	1	1	3	20

Frekans tablosu ile kolay yorum yapma sağlanır. Mesela sınavda 45 alan 4 kişi vardır.

**Soru.** Bir bölgedeki 20 binanın değerleri (TL) olarak aşağıdaki gibi verilmiş olsun.

Bina değeri (TL)	125000	130000	135000	145000	160000	Toplam
f	3	4	6	5	2	20

Bu seri frekans serisidir. 125000 TL fiyata sahip 3 bina, 130000 TL fiyata sahip 4 bina vardır. Bir frekans serisi benzer biçimde okunabilir.

Frekans serisi basit seriye dönüştürülebilir.

125000	125000	125000	130000	130000	130000	130000	135000	135000	135000
135000	135000	135000	145000	145000	145000	145000	145000	160000	160000

### iii) Gruplanmış Seri

Deney yada gözlem sonuçları çok iken, deney yada gözlem sonuçlarının belirli aralıklar (sınıflar) içinde kalan şıklara göre düzenlenmesiyle oluşturulan istatistik serilerine **gruplandırılmış (sınıflandırılmış) seri** denir.

Eğer bir değişkene ait çok sayıda veri varsa bu veriler gruplanarak, verilerin daha kolay anlaşılması yoluna gidilir.

Not Sınıfı	Frekans
30-45	6
45-60	1
60-75	8
75-90	5
Toplam	20

Bir sınıfın alt ve üst sınırları arasındaki farka **sınıf aralığı** yada **sınıf büyüklüğü** adı verilir ve **h** ile gösterilir. Yukarıdaki tabloda 30 alt sınıf, 45 üst sınıf, 30-45 aralığı 15 olup bu değer sınıf aralığıdır.

**Not:** Gruplanmış bir seride bir sınıfın üst değeri diğer sınıfın alt değerine eşit verilmişse, alt sınır değeri ilgili sınıfa dahil, üst sınır değeri ise hariç tutulur.

**Soru.**Bir bölgeye ait arsa fiyatları (TL) aşağıdaki gibi olsun.

Arsa Fiyatları (TL)	f
20.000-30.000 den az	4
30.000-40.000 den az	6
40.000-50.000 den az	10
50.000-60.000 den az	3
60.000-70.000 den az	7
Toplam	30

Bu tabloda 20.000 TL ile 30.000 TL arasında 4 arsa, 30.000 TL ile 40.000 TL arasında 6 arsa, 40.000 TL ile 50.000 TL arasında 10 arsa, 50.000 TL ile 60.000 TL arasında 3 arsa ve 60.000 TL ile 70.000 TL arasında 7 olmak üzere toplam 30 arsa vardır.

**Örnek 2.5.** Aşağıda verilen frekans dağılımında ilgili sınıflara karşı gelen sınıf orta değerlerini bulunuz?



Sınıflar	f	Sınıf Orta Noktası
0-4	4	$(0+4)/2=2$
4-8	10	$(4+8)/2=6$
8-12	17	$(8+12)/2=10$
12-16	25	$(12+16)/2=14$
16-20	14	$(16+20)/2=18$
20-24	6	$(20+24)/2=22$
24-28	4	$(24+28)/2=26$
Toplam	80	

### 2.1.3. Birikimli Seriler

Bir frekans dağılımında, her sınıfın frekansına bir önceki sınıfın frekansı eklenerek oluşturulan seriye birikimli seri, bu tür oluşturulan frekanslara da birikimli frekanslar adı verilir. Eğer birikimli seriler küçükten büyüğe doğru oluşturulmuşsa

**–den az**, büyükten küçüğe doğru oluşturulmuşsa **–den çok** olarak isimlendirilir.

**Örnek 2.6.** Bir doğum evinde doğan 100 bebeğe ilişkin sınıflandırılmış seriyi ele alarak –den az ve –den çok birikimli serilerini bulunuz?

Ağırlık (kg)	Frekans	-den az	-den çok
1,50-1,75	5	5	95+5=100 
1,75-2,00	6	6+5=11	89+6=95
2,00-2,25	10	10+11=21	79+10=89
2,25-2,50	10	10+21=31	69+10=79
2,50-2,75	35	35+31=66	34+35=69
2,75-3,00	15	15+66=81	19+15=34
3,00-3,25	13	13+81=94 	6+13=19
3,25-3,50	2	2+94=96	4+2=6
3,50-3,75	4	4+96=100	4
Toplam	100		

Yukarıdaki tablodan –den az serisi yardımıyla 66 bebeğin ağırlıklarının 2,75kg az olduğu, -den çok serisi yardımıyla da 34 bebeğin 2,75 kg dan fazla olduğu görülebilir.

**Örnek 2.7.** Aşağıda verilen frekans serisi için –den az ve –den çok serilerini oluşturunuz?

X	5	10	15	20	25	30	Toplam
f	3	5	8	6	3	5	30

X	5	10	15	20	25	30
f	3	5	8	6	3	5
-den az	3	8	16	22	25	30
-den çok	30	27	22	14	8	5

#### 2.1.4. Bileşik Seriler

Birimlerin birden fazla değişkene göre dağılımlarını bir arada gösteren serilerdir.

**Örnek 2.8.** Bir sınıftan rasgele seçilen 5 öğrencinin boy uzunlukları ve ağırlıkları aşağıdaki gibidir.

Öğrenci No	Uzunluk (m)	Ağırlık (kg)
	X	Y
1	1,72	68
2	1,68	70
3	1,80	76
4	1,74	73
5	1,76	71

Verilen örnekte birim öğrencidir. Boy uzunluğu ve ağırlık ise aynı birim üzerinde tanımlanmış iki farklı değişkendir. Eğer uzunluk bağımsız, ağırlık bağımlı değişken olarak alınırsa frekans dağılımı aşağıdaki gibi olur.

Uzunluk (m)	Ağırlık (kg)
X	Y
1,68	70
1,72	68
1,74	73
1,76	71
1,80	76

## 2.2. Frekans Tabloları

İstatistiksel analizlerden önce istatistiksel varsayımların değerlendirilmesi gerekir. Bunun için aşağıdaki adımlar izlenir:

- ✓ Her değişken ile başlangıç analizi
- ✓ Verilerin temizlenmesi
- ✓ Veri kalitesinin gözden geçirilmesi
- ✓ Verilerin değişiminin belirlenmesi
- ✓ Örneklemin belirlenmesi

Verilerin büyük değerler etrafında mı toplandığını yoksa dengeli mi dağıldığı hususu, frekans dağılımı ile ortaya konabilir. Hatalı veri girişi, hatalı şık kodlamaları ve eksik veri girişi frekans tablolarında rahatlıkla görülebilir.

Veri seti değerler veya kategoriler halinde organize edilir ve bunlar başlıklarla açıklandığında elde edilen veri gösterimine Tablo denir.

**Ham veri** : Çeşitli yöntemlerle toplanan ve sayısal olarak düzenlenmemiş verilere ham veriler denir.

**Sıralanış** : Ham sayısal verilerin artan veya azalan büyüklük sırasına konulmuş durumudur. En büyük ve en küçük sayılar arasındaki farka verilerin açıklığı denir. Örneğin bir sınıfta ki erkek öğrencilerin en uzun boyu olan 190 cm ve en kısa olan 165 cm ise, açıklık  $190-165=25$  cm dir.

Ham veriler yığını özetlenirken, sıklıkla yararlı bir işlem olan bunların sınıflara veya kategorilere ayrılması ve her sınıfa ait birey sayılarının belirlenmesi sınıf frekansı olarak adlandırılır. Verilerin sınıflara göre ve ilgili sınıfın frekansları ile birlikte bir tablo halinde düzenlenmesine frekans dağılımı veya frekans tablosu adı verilir.

Frekans tabloları ele alınan bir değişkene ilişkin ölçüm değerlerinden istenen sınıflara düşen değerlerin sayısını veya frekansını, oran veya % cinsinden belirlenmiş dağılımlarını gösteren tablolardır. Frekans tablolarının hazırlanmasında aşağıdaki esaslar dikkate alınmalıdır.

Herhangi bir gözlem değeri ancak bir sınıfta yer almalı, sınıflar çakışmamalıdır.

Hiçbir gözlem değeri belirlenen sınıfların dışında kalmamalı, tüm gözlemler sınıflarda yer almalıdır.

Frekans dağılımında kullanılan sınıf aralıkları mümkün olduğunca eşit olmalıdır. Eşit olmayan sınıf aralıkları dağılımların grafiksel olarak gösterilmesinde sorunlar yaratır. Ancak bazen çok sayıda boş sınıfın ayıklanması için eşit olmayan aralıkların kullanılması gerekebilir.

### **Basit bir serinin gruplanmış seriye dönüştürülmesi:**

Frekans dağılımlarının yer aldığı frekans tablolarının oluşturulmasında izlenecek adımlar ve her adımda yapılması gereken işlemler aşağıdaki şekilde özetlenebilir:

1. Sınıf sayısı belirlenir (k).

Kaç sınıf yapılacağını yaklaşık hesaplamak için aşağıdaki formül kullanılır (Lind ve Mason, 1997):

$$Sınıf\ sayısı = 1 + 3,322 \times \log(N)$$

2. Her sınıfın sınıf aralığı veya sınıf genişliği belirlenir.

$$C = \frac{X_{max} - X_{min}}{1 + 3,322 \times \log(N)}$$

$X_{max}$  : En büyük gözlem değeri

$X_{min}$  : En küçük gözlem değeri

k: Sınıf sayısı

3. Frekans dağılımlarının kolay yorumlanması için bütün sınıfların aynı genişlikte olması önerilir.

4. Her sınıfın tanımında kullanılacak **sınıf orta değeri (SOD)** bulunur.

$$SOD_i = \frac{ASD_i - ÜSD_i}{2}$$

$ASD_i$  : i. sınıfa ait alt sınır değeri

$ÜSD_i$  : i. sınıfa ait üst sınır değeri

5. Gözlem değerleri tek tek sınıflara birer çizgi çizilerek dağıtılır ve her sınıfa ait frekanslar bu çizgiler sayılarak bulunur.

6. Her sınıftaki frekansların oransal frekansları bulunur.

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

7. Oransal frekans değerleri 100 ile çarpılarak her sınıfa ait yüzde frekanslar bulunur.

8. Kümülatif frekanslar ve yüzdeleri bulunur.



**Örnek 2.9.** Bir araştırmada rasgele seçilen 35 otomobilin yaşları aşağıdaki gibidir. Otomobil dağılımlarını gösteren frekans tablosunu oluşturunuz?

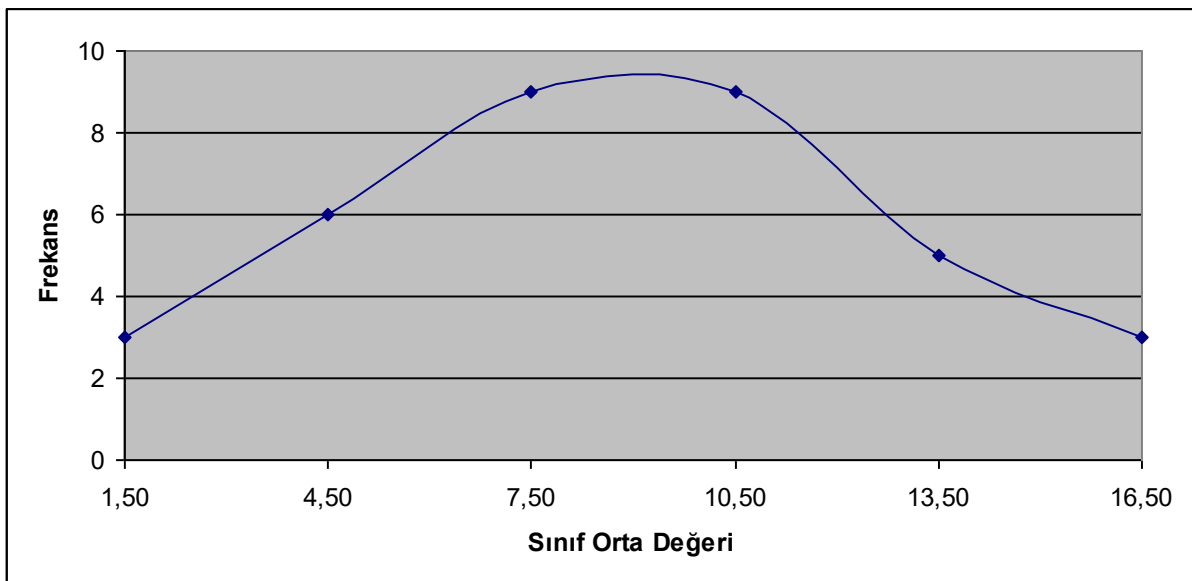
Yaşlar: 17, 14, 9, 8, 4, 9, 5, 6, 7, 12, 17, 10, 9, 9, 1, 11, 3, 2, 6, 7, 12, 13, 15, 5, 8, 7, 9, 12, 11, 6, 1, 3, 7, 4, 10

**Sınıf Sayısı (k)=1 + 3.322\*log<sub>10</sub>(35)=6.13 ~6**

$$C = \frac{17 - 1}{6} = 3$$

Sınıf No	Sınıf Sınırları *	Sınıf Değeri	Frekans			Kümülatif Frekans		
			Sayısal	Oran	%	Sayısal	Oran	%
1	0-3	1.5	///=3	0.09	9	3	0.09	9
2	3-6	4.5	/////=6	0.18	18	9	0.27	27
3	6-9	7.5	/////////=9	0.25	25	18	0.52	52
4	9-12	10.5	/////////=9	0.25	25	27	0.77	77
5	12-15	13.5	/////=5	0.14	14	32	0.91	91
6	15-18	16.5	///=3	0.09	9	35	1.00	100
Toplam			35	1.0	100			

\* Alt sınır değeri ilgili sınıfa dahil, üst sınıf değeri ise dahil değildir.



Şekil 2.1. Çizgi grafiği

**Örnek 2.10.** Yüz yeni doğan çocuğun doğum ağırlıkları (kg) aşağıdaki gibi bulunmuştur.

2,6	2,6	3,0	2,7	3,8	2,7	3,1	2,7	4,8	4,9
3,6	4,0	4,8	2,7	3,9	2,4	2,0	3,4	2,0	3,9
3,6	4,7	2,3	2,4	3,7	2,7	1,7	2,9	2,4	2,1
5,1	3,2	3,1	4,1	3,0	4,3	3,8	3,3	2,4	3,2
<b>1,7</b>	5,0	4,4	2,6	3,0	3,6	4,7	2,6	2,8	2,7
3,1	2,0	2,4	3,5	2,2	3,9	2,9	3,6	2,6	4,0
2,4	3,3	4,1	2,4	2,2	3,1	4,2	3,9	2,4	3,3
3,8	3,5	<b>5,2</b>	2,5	3,4	3,2	3,8	4,1	2,7	3,7
3,3	4,3	3,4	2,1	2,9	3,8	2,2	3,2	2,8	2,9
2,7	5,0	3,1	2,9	2,7	3,5	2,0	2,5	4,7	2,5

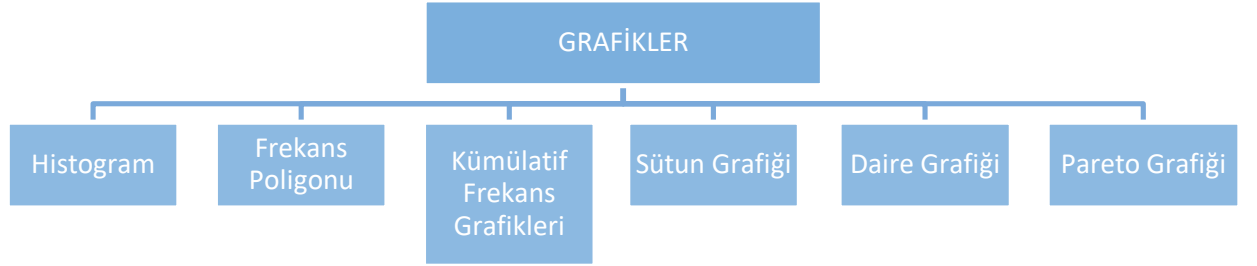
Sınıf sayısı=  $1+3.322.(\log 100)=1+6,644=7.644 \approx 8$  sınıf yapılabilir.

$$C = \frac{X_{max} - X_{min}}{1 + 3,322 \times \log(N)} = \frac{5,2 - 1,7}{8} = 0,5$$

Sınıflar oluşturulurken belirli kurallara uymak gerekir. Çok fazla sayıda sınıf yapılırsa veri iyi özetlenmemiş olur. Az sınıf yapılırsa bilgi kaybı çok fazla olur. Bu nedenle optimum sayıda sınıf oluşturmalıdır.

Sınıf genişliği 0.5 kg olacak şekilde doğum ağırlıklarını sıklık(frekans) tablosu halinde özetlemek için 8 sınıf yapılabilir.

Sınıflar	Sıklık(frekans)
1.5 - 1.9	2
2.0 - 2.4	18
2.5 - 2.9	24
3.0 - 3.4	19
3.5 - 3.9	18
4.0 - 4.4	9
4.5 - 4.9	6
5.0 - 5.4	4
Toplam	100





Grafik, gözlem değerlerinin matematiksel ve bilimsel temellere sahip şekiller hâlinde ifade edilmesidir.

## GİRİŞ

Bilindiği üzere istatistiğin uğraş alanlarından biri de özetleme ve görselleştirme. Frekans dağılımlarının bir tablo şeklinde sunulması bir özetleme şeklidir. Oysa tablolar her zaman kolay anlaşılır olmayabilir. Grafikler gözlem değerlerinin görselleştirilmesinde kullanılır. *Grafik, “gözlem değerlerinin matematiksel ve bilimsel temellere sahip şekiller hâlinde ifade edilmesi” şeklinde tanımlanabilir.*

*Grafikler gözlem değerlerinin daha kolay anlaşılmasını sağlamaktadır.* Çünkü grafikler temsil ettikleri olayların bileşimini ve değişimelerindeki ana eğilimi tüm canlılığı ile ilk bakışta ortaya koymaktadırlar. Bu nedenle rakamlarla uğraşmak istemeyen veya rakamlardan anlam çıkarmakta zorluk çekenler için grafikler daha fazla yardımcı olmaktadır.

Grafik verilmeksizin yalnız rakamlar hâlindeki sonuçları inceleyen bir şahıs, olayın genel eğiliminden çok ayrıntısına takılacak ve bu yüzden olayın genel özelliklerini kavrayamayacaktır. Oysa iyi çizilmiş bir grafik üzerinde gösterilen sonuçlar görsel olduğundan şahıslar tarafından daha anlaşılır olacaktır. Örneğin; yıllar itibarıyla bir ülkenin nüfusunun incelendiği bir araştırmada rakamlardan çok, grafikler daha anlaşılır olacaktır. Grafik üzerinde yıllar itibarıyla nüfusun artıp artmadığı veya nasıl değiştiği ilk bakışta kolaylıkla anlaşılabilir. Oysa rakamlarla bu durum ifade edilmeye çalışıldığında durumun anlaşılması grafiklerde olduğu gibi kolay olmayacaktır [1].

Bu bölümde grafik çeşitlerinden sırasıyla histogram, frekans poligonu, kümülatif frekans dağılımları, daire grafiği, sütun grafiği, pareto grafiği ve serpilme diyagramı anlatılacaktır.

## HİSTOGRAM

*Histogram, gruplandırılmış seriler için oluşturulan bir dikkörtgenler dizisidir. Histogram; yatay eksen de değişkenin aldığı değerlerin, dikey eksen de ise frekansların bulunduğu ve her aralığın frekansı ile orantılı yüksekliklere sahip dikkörtgenlerin gösterildiği bir yoğunluk grafiğidir.* Bu dikkörtgenlerin tabanları gruplandırılmış serideki her bir sınıfın sınıf büyüklüğünü, yükseklikleri ise sınıf frekansını göstermektedir.

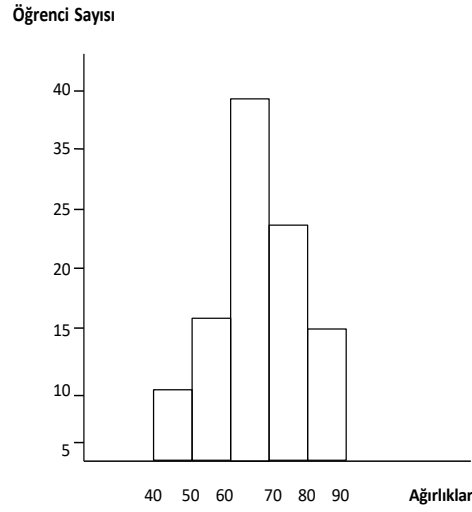
*Histogram çizilirken yatay eksen de gruplandırılmış serinin sınıf sınırları, dikey eksen de ise frekanslar yer almaktadır.* Sınıf aralıkları ve frekans değerleri eksenlerde belirlendikten sonra sınıf sınırlarının alt ve üst sınırlarından frekans değerlerine kadar birer dikme çizilir. Gruplandırılmış serilerde sınıfların frekanslarının sınıf sınırları içerisinde düzgün dağıldığı kabul edildiğinden, çizilen dikmeler yatay eksene paralel bir çizgi ile birleştirilerek dikkörtgenler elde edilir. Bu dikkörtgenlerin tamamı histogramı oluşturmaktadır [2]. Aşağıdaki örnekte 100 öğrenciye ait ağırlık dağılımları verilmiş ve bu dağılıma ait histogram Şekil 4.1'de gösterilmiştir.



Histogram, gruplandırılmış seriler için oluşturulan bir dikkörtgenler dizisidir.

Örnek 4.1. Bir sınıftaki 100 öğrencinin ağırlıklarına göre dağılımları aşağıda verilmiştir. Ağırlıkların dağılımının histogramını çiziniz.

Ağırlıklar	Öğrenci sayısı
40 - 50	10
50 - 60	16
60 - 70	38
70 - 80	22
80 - 90	14



**Şekil 4.1.** Ağırlık Dağılımı İçin Oluşturulan Histogram

Grafik incelendiğinde; en yüksek olan dikdörtgen en fazla öğrencinin olduğu 60 – 70 sınıf aralığını, en kısa dikdörtgen ise en az öğrencinin olduğu 40 – 50 sınıf aralığını göstermektedir. Buradan öğrencilerin en fazla 60 – 70 kilo aralığında, en az ise 40 – 50 kilo aralığında oldukları anlaşılmaktadır.

## FREKANS POLİGONU

*Frekans poligonu, histogramdaki sınıf sınırlarının orta noktalarını apsis, sınıf frekanslarını ordinat kabul eden noktaların doğru parçaları ile birleştirilmesi ile elde edilen grafik çeşididir.* Kısacası histogramı oluşturan dikdörtgenlerin üst kenarlarının orta noktaları birleştirilmek suretiyle elde edilen grafiğe frekans poligonu denir. *Frekans poligonu, sınıf sınırlarının mümkün olduğunca küçültülmesi durumunda bir eğriye dönüşür. Söz konusu eğriye ise frekans eğrisi adı verilir.*

Frekans poligonunun yatay eksen üzerindeki başlangıç noktası ilk sınıftan bir önceki farazi sınıfın orta noktası iken, bitiş noktası ise son sınıftan bir sonraki farazi sınıfın orta noktasıdır. Histogramların kapladığı alan ile poligonun altında kalan alan birbirine eşittir [3]. Aşağıdaki örnekte 100 öğrencinin istatistik dersinden aldıkları notlar kullanılarak elde edilen histogram ile frekans poligonu Şekil 4.2’de gösterilmiştir.



Histogramı oluşturan dikdörtgenlerin üst kenarlarının orta noktaları birleştirilmek suretiyle elde edilen grafiğe frekans poligonu denir.

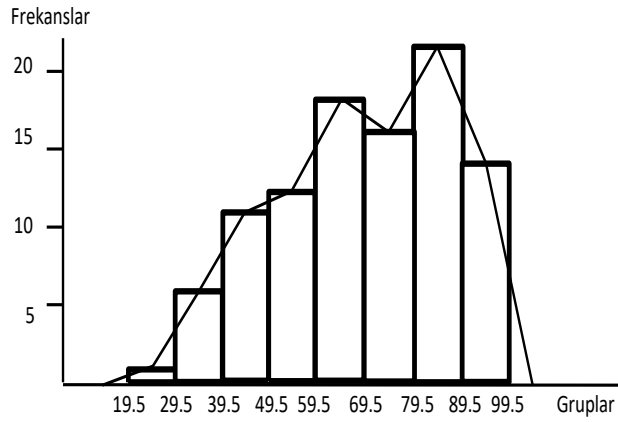
Örnek 4.2. İktisadi ve İdari Bilimler Fakültesi'nde okuyan 100 öğrencinin istatistik dersinden aldığı notları sorularak aşağıdaki verilere ulaşılmıştır. Öğrencilerin istatistik dersinden aldıkları notları kullanarak histogram ve frekans poligonunu çiziniz.

Notlar	Öğrenci Sayısı
20-29	1
30-39	6
40-49	11
50-59	12
60-69	18
70-79	16
80-89	22
90-99	14

Burada verilen seri kesikli gruplandırılmış seri olduğu için öncelikle serinin sınıf sınırları ve sınıf değerleri elde edilmiş ve aşağıda verilmiştir:

Gruplar	Alt Sınır	Üst Sınır	Sınıf Değeri
20 – 29	19.5	29.5	24.5
30 – 39	29.5	39.5	34.5
40 – 49	39.5	49.5	44.5
50 – 59	49.5	59.5	54.5
60 – 69	59.5	69.5	64.5
70 – 79	69.5	79.5	74.5
80 – 89	79.5	89.5	84.5
90 – 99	89.5	99.5	94.5

Sınıf sınırları ve sınıf değerleri kullanılarak serinin histogramı ve frekans poligonu aşağıdaki gibi elde edilmiştir:



Şekil 4.2. İstatistik Notlarının Histogram ve Frekans Poligonu



Kümülatif frekans poligonlarına oжив eğrileri de denir.

## KÜMÜLATİF FREKANS GRAFİKLERİ

“...den az kümülatif frekans dağılımları” ve “...den çok kümülatif frekans dağılımları”nın grafikleri koordinat sistemi üzerinde çizilebilir. Grafiklerin çiziminde yatay eksen de değişken değerleri, dikey eksen ise kümülatif frekans değerleri bulunur. Değişkenin aldığı değerler ile kümülatif frekansların kesiştiği noktaların birleştirilmesi ile kümülatif frekans dağılımlarının grafiği çizilmiş olur. Nispi ve kümülatif frekans dağılımlarına ait grafikleri elde etmek için dikey eksene nispi veya kümülatif frekans değerleri yerleştirilmelidir. *Kümülatif frekans poligonlarına oжив eğrileri de denir* [2].

Gruplandırılmış serilerin ...den az ve ...den çok kümülatif frekans dağılımlarına ait grafiklerin çiziminde yatay eksen de sınıf sınırları işaretlenir. “...den az kümülatif frekans dağılımı” için üst sınıf sınırları apsis, frekans değerleri ordinat olacak şekilde düzlemdeki noktalar belirlenir. Belirlenen noktalar çizgilerle birleştirilir. “...den çok kümülatif frekans dağılımları”na ait grafiklerin çiziminde ise gruplandırılmış seride alt sınıf sınırları kullanılmaktadır. Alt sınıf sınırları apsis, frekans değerleri ordinat olacak şekilde düzlemdeki noktalar çizgilerle birleştirilir. Gruplandırılmış serideki son sınıfın üst sınırında son nokta bulunur ve burada grafik yatay eksene değecek şekilde çizim yapılır. Kısaca gruplandırılmış serilerin “...den az kümülatif frekans dağılımları”na ait grafikler çizilirken üst sınıf sınırları “...den çok kümülatif frekans dağılımları”na ait grafikler çizilirken ise alt sınıf sınırları kullanılmaktadır [4]. Aşağıdaki örnekte Örnek 4.2’de verilen öğrencilerin not dağılımları kullanılarak oluşturulan “...den az ve ...den çok kümülatif frekans dağılımları”nın grafikleri sırasıyla Şekil 4.3 ve Şekil 4.4’te gösterilmiştir.

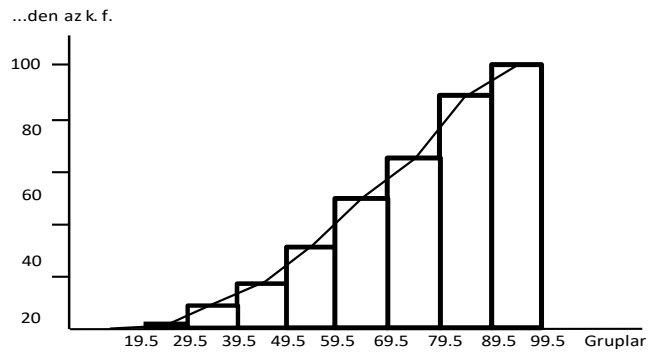
Örnek 4.3. Örnek 4.2’de verilen öğrencilerin not dağılımını kullanarak, ...den az ve ...den çok kümülatif frekans dağılımlarının grafiklerini çiziniz.

Notlar	Öğrenci Sayısı
20-29	1
30-39	6
40-49	11
50-59	12
60-69	18
70-79	16
80-89	22
90-99	14

“...den az kümülatif frekans dağılımı”nın grafiğini çizmek için öncelikle “...den az kümülatif frekans değerleri” elde edilir:

Gruplar	Frekanslar	...den az k. f.
20-29	1	1
30-39	6	7
40-49	11	18
50-59	12	30
60-69	18	48
70-79	16	64
80-89	22	86
90-99	14	100

Bu dağılıma göre hazırlanan “...den az kümülatif frekans dağılımı”nın grafiği aşağıdaki gibidir:



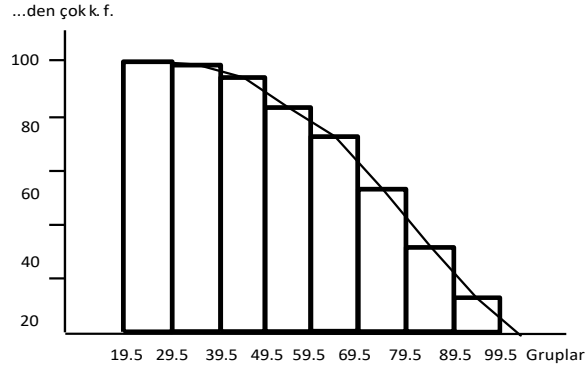
**Şekil 4.3.** “...den Az Kümülatif Frekans Dağılımı”nın Grafiği

“...den çok kümülatif frekans dağılımı”nın grafiğini çizmek için öncelikle “...den çok kümülatif frekans değerleri” elde edilerek aşağıda verilmiştir:

Gruplar	Frekanslar	...den çok k.f.
20-29	1	100
30-39	6	99
40-49	11	93
50-59	12	82
60-69	18	70
70-79	16	52
80-89	22	36
90-99	14	14

Bu dağılıma göre hazırlanan “...den çok kümülatif frekans dağılımı”nın grafiği aşağıdaki gibidir:





**Şekil 4.4.** “...den Çok Kümülatif Frekans Dağılımının” Grafiği

...den az ve ...den çok kümülatif frekans dağılımları aynı grafik üzerinde gösterilebilir. Grafikteki iki kümülatif frekans eğrisinin kesiştiği noktadaki X değeri medyan değerini vermektedir.

**Örnek 4.4.** Bir pil fabrikasında üretilen piller arasından tesadüfi olarak seçilen 400 pilin dayanma süreleri aşağıdaki gibi tespit edilmiştir. ...den az ve ...den çok kümülatif frekans dağılımlarına ait ojiv eğrilerini aynı grafik üzerinde gösteriniz.

Pillerin Dayanma Süresi	Pil Sayısı
300 – 399	14
400 – 499	46
500 – 599	58
600 – 699	76
700 – 799	68
800 – 899	62
900 – 999	48
1000 – 1099	22
1100 – 1199	6

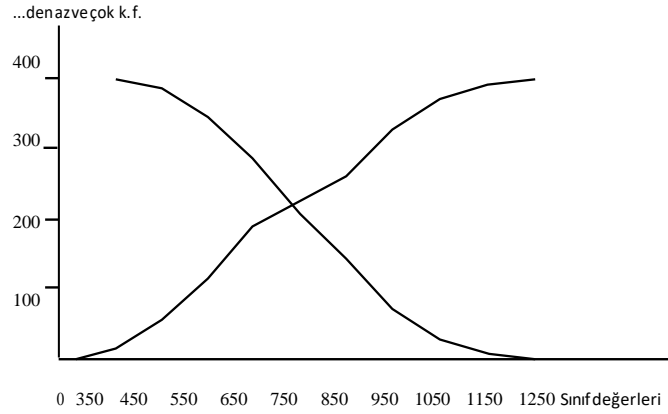
Burada seri kesikli gruplandırılmış seri olduğu için öncelikle serinin sınıf sınırları ve sınıf değerleri elde edilmiş ve aşağıda verilmiştir:

Dayanma Süresi	Pil Sayısı	Alt Sınır	Üst Sınır	Sınıf Değeri
300 – 399	14	299.5	399.5	349.5
400 – 499	46	399.5	499.5	449.5
500 – 599	58	499.5	599.5	549.5
600 – 699	76	599.5	699.5	649.5
700 – 799	68	699.5	799.5	749.5
800 – 899	62	799.5	899.5	849.5
900 – 999	48	899.5	999.5	949.5
1000 – 1099	22	999.5	1099.5	1049.5
1100 – 1199	6	1099.5	1199.5	1149.5

“...den az kümülatif frekans dağılımları” ve “...den çok kümülatif frekans dağılımları”na ait ojiv eğrilerini aynı grafik üzerinde çizmek için öncelikle kümülatif frekans değerleri elde edilerek aşağıda verilmiştir:

Dayanma Süresi	Sınıf Değeri	Pil Sayısı	..den az k.f.	..den çok k.f.
299.5 – 399.5	349.5	14	14	400
399.5 – 499.5	449.5	46	60	386
499.5 – 599.5	549.5	58	118	340
599.5 – 699.5	649.5	76	194	282
699.5 – 799.5	749.5	68	262	206
799.5 – 899.5	849.5	62	324	138
899.5 – 999.5	949.5	48	372	76
999.5 – 1099.5	1049.5	22	394	28
1099.5 – 1199.5	1149.5	6	400	6

Frekans dağılımındaki sınıf değerlerine karşı gelen ...den az ve ...den çok kümülatif frekans değerleri koordinat sisteminde işaretlendikten sonra bu noktalar birleştirilerek ojiv eğrileri aşağıdaki gibi elde edilir. [2]



Şekil 4.5. “...den Az ve ...den Çok Frekans Dağılımlarının” Ojiv Eğrileri

## SÜTUN GRAFİĞİ

Nitel veya başka bir ifade ile kategorik veriler değişkenin vasfı ile ilgili sayısal olmayan verilerdir. Özellikle sosyal bilimlerde araştırma konuları gereği daha fazla nitel veriler kullanılmaktadır. Nitel veride birimlere ait sıfatlar ya da durumlar tespit edilir. Nitel verilerin sunumu grafiklerle yapılabilir. Örneğin; ülkelerin nüfus dağılımı, televizyon kanallarının prime-time izlenme oranları, siyasi partilerin seçimlerde aldıkları oy oranları, meslek gruplarına göre dağılım grafiklerle gösterilebilir. Bu tip verilerin görsel olarak sunulması için çeşitli grafikler mevcuttur. Bu grafik çeşitlerinden biri de sütun grafiğidir.

*Sütun grafiği oluşturulurken, yatay eksen de kategoriler, dikey eksen de ise bu kategorilerin frekansları gösterilerek oluşturulur. Çizilen sütunların genişlikleri*



Sütun grafiği; yatay eksen de kategorilerin, dikey eksen de ise bu kategorilerin frekansları gösterilerek elde edilir.

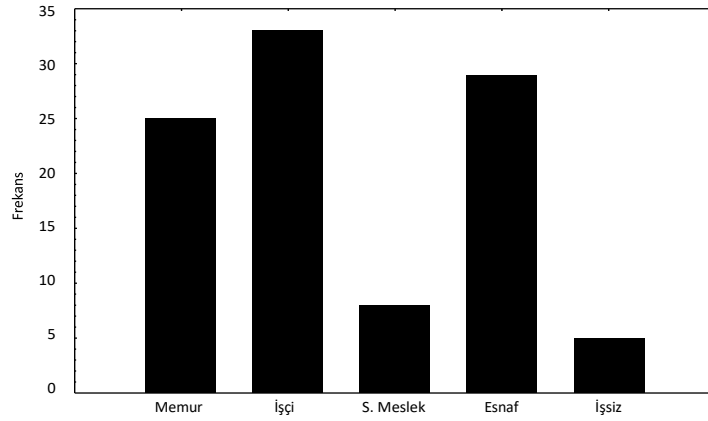
*birbirine eşittir. Çizilen sütunlar birbirine bitişik olabileceği gibi birbirinden ayrık da olabilir [5].*

Aşağıdaki örnekte 100 kişiye ait meslek grupları verilmiş ve bu dağılım kullanılarak oluşturulan sütun grafiği Şekil 4.6'da gösterilmiştir.

Örnek 4.5. Bir ilde yaşayan 100 şahsın meslekleri aşağıdaki gibi tespit edilmiştir. Bu verileri kullanarak sütun grafiğini çiziniz.

Meslekler	Frekanslar
Memur	25
İşçi	33
Serbest Meslek	8
Esnaf	29
İşsiz	5

Yatay ekseninde meslek grupları dikey ekseninde ise bu meslek gruplarının frekansları gösterilerek sütun grafiği elde edilmiş ve aşağıda verilmiştir:



**Şekil 4.6.** Meslek Gruplarına Ait Sütun Grafiği

Grafik incelendiğinde; ilk bakışta en fazla sayıda kişinin işçi olduğu, ikinci sırada ise esnafların olduğu ve an az sayıda kişinin ise işsiz olduğu rahatlıkla görülebilir [2].

## DAİRE GRAFİĞİ

*Daire grafiği yukarıda izah edilen sütun grafiği gibi kategorik verilerin sunumunda kullanılan bir grafik türüdür. Daire grafiği, nominal ölçekle elde edilmiş veriler, kategorik veriler veya az sayıda sınıfa ayrılabilen veriler için kullanılabilecek bir grafik türüdür. Daire grafiğinde her sınıf veya kategori sahip olduğu frekansla orantılı büyüklükteki dilimlerle gösterilmektedir. Şekil itibarıyla dilimlenmiş pastaya benzediği için bu grafiğe pasta grafiği de denilmektedir [4].*



Daire grafiği, sütun grafiği gibi kategorik verilerin sunumunda kullanılan bir grafik türüdür.



Daire grafiği, şekil itibarıyla dilimlenmiş pastaya benzediği için bu grafiğe pasta grafiği de denilmektedir.

Daire grafiğinin çizilmesi için öncelikle nispi frekans değerleri elde edilmektedir. Bir dairenin iç açıları toplamı  $360^\circ$  olduğuna göre nispi frekanslar 360 ile çarpılarak her bir kategorinin  $360^\circ$  den alacağı pay tespit edilir. Toplam veri sayısı  $360^\circ$  lik açıya karşılık gelmektedir. Genellikle saatin 12'yi gösterdiği noktadan başlanarak her bir kategorinin belirlediği alanın bulunması için, saatin dönüş istikametine göre, kategoriler için belirlenen açılar kadar ilerlenir [2].

*Yukarıda ifade edildiği gibi daire grafiği, az sayıda sınıfa veya kategoriye ayrılabilen veriler için kullanılabilecek bir grafik türüdür.* Sınıf veya kategori sayısı arttıkça daire grafiğinin çizimi ve anlaşılması zorlaşır. Grafiklerde amaç, verilerin anlaşılır bir şekilde görsel olarak sunulması olduğuna göre grafik karmaşıklıkça amacından uzaklaşmış olacaktır. Dolayısıyla kategorik veriler için grafik çizilirken bu hususun dikkate alınması gerekmektedir [1].

Aşağıdaki örnekte bir ülkedeki çalışma kategorilerine göre çalışanların sayısı verilmiş ve bu veriler kullanılarak oluşturulan daire grafiği Şekil 4.7'de gösterilmiştir.

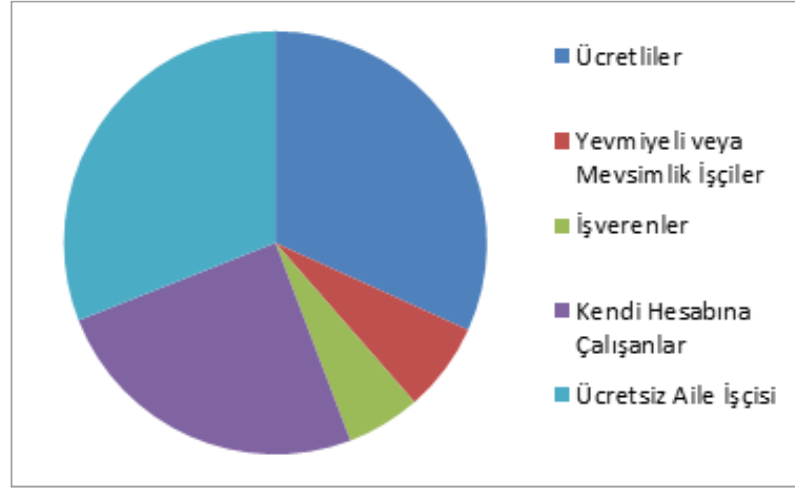
Örnek 4.6. Bir ülkedeki çalışma kategorilerine göre çalışanların sayısı aşağıdaki gibidir. Söz konusu verileri kullanarak daire grafiğini çizeriz.

Çalışma Kategorileri	Çalışan Sayısı
Ücretliler	6283355
Yevmiyeli veya Mevsimlik İşçiler	1347728
İşverenler	1118950
Kendi Hesabına Çalışanlar	4905037
Ücretsiz Aile İşçileri	6129021

Daire grafiğinin çizilmesi için ilk olarak nispi frekans değerleri elde edilmiştir. Daha sonra nispi frekans değerleri 360 ile çarpılarak her bir kategorinin  $360^\circ$  den alacağı paylar tespit edilmiştir. Elde edilen değerler aşağıda gösterilmiştir:

Çalışma Kategorileri	Frekanslar	n.f.	$360^\circ \times n.f$
Ücretliler	6283355	0.318	114
Yevmiyeli veya Mevsimlik İşçiler	1347728	0.068	25
İşverenler	1118950	0.057	20
Kendi Hesabına Çalışanlar	4905037	0.247	89
Ücretsiz Aile İşçileri	6129021	0.310	112
<b>Toplam</b>	<b>19784091</b>		<b>360</b>

Gerekli hesaplamalar yapıldıktan sonra saatin 12'yi gösterdiği noktadan başlanarak her bir kategorinin belirlediği alanın bulunması için saatin dönüş istikametine göre, kategoriler için belirlenen açılar kadar ilerlenir. Böylece daire grafiği aşağıdaki gibi elde edilir:



Şekil 4.7: Çalışma Kategorilerine Ait Daire Grafiği

Grafik incelendiğinde en fazla payı %31.8 ile ücretliler alırken onu %31 ile ücretsiz aile işçileri ve %24.7 ile kendi hesabına çalışanlar takip etmektedir. En az payı ise %5.7 ile işverenler almaktadır [2].

## PARETO GRAFİĞİ

İlk kez İtalyan Ekonomist Pareto (1848-1923) tarafından bulunan Pareto grafiği, söz konusu bilim adamının adıyla anılmaktadır. *Pareto grafiği gelir dağılımları ile ilgili çalışmalarda tespit edilmiştir. Grafik hata ve maliyet analizleri için kullanılan basit bir yöntemdir.*

*Pareto grafiği; bozuk ürünler, tamirler, arızalar, talepler, noksanlıklar veya kazalar ile mali kayıplar ve bunların sebepleri gibi olayların görsel olarak meydana gelme frekanslarını gösteren bir tür frekans dağılım grafiğidir.*

Pareto grafiği bir sorunu oluşturan nedenleri önem sırasına göre sıralayarak, önemlileri önemsizlerden ayırt etmeye ve dikkatleri önemli nedenler üzerinde toplamaya imkân vermektedir. Grafikte soldan sağa doğru gidildikçe küçülen sütunlara sahip veri sınıfları sıralanmaktadır [6].

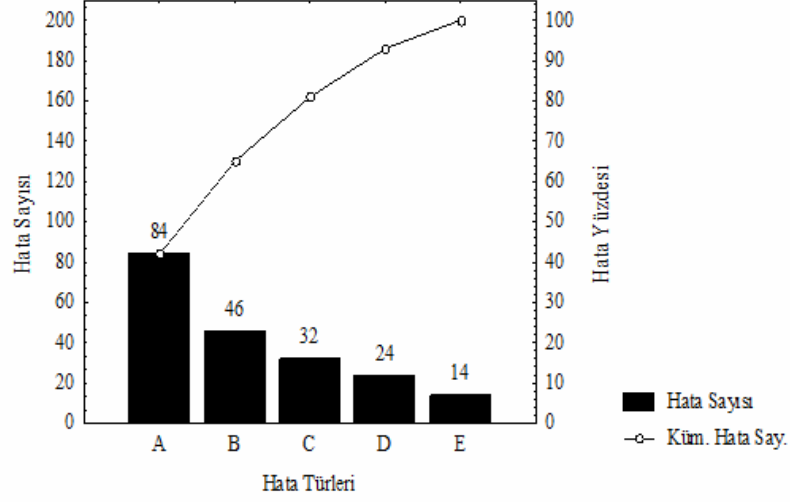
*Pareto grafiğinde en önemli problemler sol tarafta yer alırken, önemsiz olanlar sağ tarafta yer alır.* Bazen “Diğerleri” adı altında çok önemsiz durumlar bir sınıf altında toplanıp birleştirilebilir. “Diğerleri” kategorisi kullanıldığında bu sınıf en sağda yer alır. Maliyet, frekanslar veya % gibi değerler de dikey eksenle gösterilir.

Pareto grafiği bazı özellikleri bakımından histograma benzemektedir. Histogramdan şu özelliği ile ayırt edilebilir: Pareto grafiğinde yatay eksen kategorik verileri gösterirken, histogramda yatay eksen numerik verileri göstermektedir.

Bazı durumlarda pareto grafiğinde kümülatif eğri gösterilmektedir. Söz konusu eğri soldan sağa doğru eklenen veri sınıflarının toplamını göstermektedir [4]. Şekil 4.8.’de örnek bir pareto grafiği verilmiştir.



Pareto grafiği, hata ve maliyet analizleri için kullanılan basit bir yöntemdir.



Şekil 4.8. Örnek Bir Pareto Grafiği

Grafik incelendiğinde en fazla A türü hatanın yapıldığı onu da B türü hatanın takip ettiği görülmektedir. En az hata türünün ise E türü hata olduğu pareto grafiğinden kolaylıkla anlaşılmaktadır.

## SERPİLME DİYAGRAMI

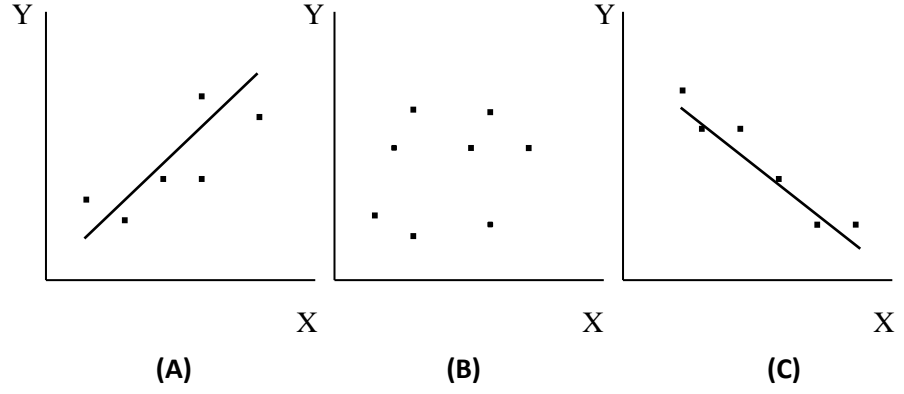
Bilimin en temel amaçlarından birisi, değişkenler arasındaki ilişkileri incelemektir. *Değişkenler arasındaki ilişkilerin grafik aracılığı ile incelenmesinde serpilme diyagramından yararlanılmaktadır. Serpilme diyagramlarında yatay eksen genellikle neden durumunda olan değişken (X) ve dikey eksen de sonuç durumunda olan değişken (Y) kullanılır.* Neden durumunda olan değişkene (X) bağımsız değişken ve sonuç durumunda olan değişkene (Y) de bağımlı değişken adı verilmektedir. Bu iki değişkene ait sıralı ikili çiftleri olan (X, Y) değerler kartezyen koordinat sistemi üzerine yerleştirilir [7].

Boy ile kilo, fiyat ile talep, arz ile talep, gelir ile tasarruf, gelir ile harcama, çalışma süresi ile başarı notu gibi ikili değişkenler arasındaki ilişkileri incelemek için serpilme diyagramı kullanılabilir.

*(X, Y) değerlerinin oluşturduğu noktalara bakılarak ilişkinin yönü ve derecesi göz kararıyla tahmin edilebilir.* Bunun için en dışta kalan noktalar birleştirilerek bir şekil elde edilir. Söz konusu şeklin doğrultusuna bakılarak ilişkinin yönü saptandığı gibi, şeklin durumuna göre ilişkinin derecesi hakkında tahminde bulunulur. Şayet şekil oldukça dar bir elipse benziyorsa ilişkinin kuvvetli olduğu ifade edilir. Elips genişledikçe ilişkinin derecesi zayıflar. Değişkenler arasında ilişki bulunmadığında ise noktaların oluşturduğu şekil bir daireye benzer [4]. Şekil 4.9.'da üç farklı durum için serpilme diyagramı örnekleri verilmiştir.



Değişkenler arasındaki ilişkilerin grafik aracılığı ile incelenmesinde serpilme diyagramından yararlanılmaktadır.



**Şekil 4.9.** Serpilme Diyagramı Örnekleri

Şekil 4.9 incelendiğinde; A ve C grafiklerinde şekil elipse benzediği için X ve Y değişkenleri arasında ilişki olduğuna karar verilir. (A) grafiği X ve Y değişkenleri arasında pozitif ve kuvvetli bir ilişkinin olabileceğini işaret ederken, (C) grafiği X ve Y değişkenleri arasında ters yönlü kuvvetli bir ilişkinin olabileceğini işaret etmektedir. (B) grafiğinde ise şekil daireye benzediği için X ve Y değişkenleri arasında ilişki olmadığına karar verilir [2].



#### Bireysel Etkinlik

- Hangi veri için hangi grafik türünün daha uygun olacağını tartışınız.
- Düşüncelerinizi sistemde ilgili ünite başlığı altında yer alan “tartışma forumu” bölümünde paylaşabilirsiniz.
- Grafik türleri için birer örnek veriniz.
- Seçim ve referandum sonuçları hangi grafik türü ile gösterilir?