

## Gövde-Yaprak Gösterimi

Verilerin çeşitli özelliklerine doğrudan dikkat çekecek şekilde sayıları grafiksel olarak düzenlememizi sağlar. Bir başka deyişle, veri kümesini bir bütün olarak görmemizi ve aşağıdaki özellikleri fark etmemizi sağlar.

- Verinin dağılımının simetrik olup olmadığını
- Verinin yayılımını
- Veri setinde aykırı değer olup olmadığını

Gövde-yaprak gösterimi, veri kümesinin dağılımı ve yukarıda belirtilen özellikleri belirlenirken yaygın olarak kullanılan bir tekniktir, bkz Tukey (1970, 1972)

### Örnek:

$i$	$X_i$	$i$	$X_i$
1	20.9	12	27.6
2	25.2	13	27.7
3	25.7	14	28.2
4	25.9	15	28.4
5	26.1	16	28.6
6	26.4	17	29.5
7	26.7	18	29.6
8	26.8	19	29.9
9	27.0	20	30.1
10	27.3	21	30.2
11	27.5	22	30.6
		23	31.1

Yukarıdaki veri setinin gövde-yaprak gösterimi için öncelikle rakamları

Veri	Bölünmüş veri	Gövde	Yaprak
20.9	20   9	20	9

şeklinde göstermemiz gerekmektedir.

Bu veri setinin gövde-yaprak gösterimi aşağıda gösterildiği gibi elde edilir.

derinlik

1	20	9
	21	
	22	
	23	
	24	
4	25	2 7 9
8	26	1 4 7 8
(5)	27	0 3 5 6 7
10	28	2 4 6
7	29	5 6 9
4	30	1 2 6
1	31	1

Görüldüğü gibi aralık genişliği 1, birimi ise 0.1 dir. Dikkat edilmelidir ki, veri setindeki sayılarda virgülden sonra bir basamak kullanılmıştır.

20.9'a kadar olan gözlem sayısı 1'dir. O yüzden derinlik 1 olarak gösterilir. 25.9'a kadar olan gözlem sayısı 4 tür. Benzer şekilde derinlik 4 olarak gösterilir.

Toplam gözlem sayısı 23 olduğu için medyan

$$X_{\left(\frac{n+1}{2}\right)} = X_{\left(\frac{23+1}{2}\right)} = X_{(12)}$$

olarak hesaplanır. Bir başka deyişle medyan 12. gözlem olarak belirlenir.

Medyanın bulunduğu satırdaki (5) rakamı parantez içine alınmıştır. Bu rakam medyan satırında yer alan gözlem sayısını gösterir.

**Not:** Veri setinde 22 gözlem olsaydı ve 11. gözlem ile 12. gözlemler ayrı ayrı satırlara düşseydi medyan satırında parantez kullanmaya gerek kalmayacaktı.

Derinlik deęerleri bulunurken medyandan önceki satırlardaki gözlem sayıları yukarıdan aşağıya; medyandan sonraki satırlardaki gözlem sayıları aşağıdan yukarıya doğru toplanarak yazılır.

**Not:** Medyan satırındaki, medyan satırının bir üst ve bir alt satırındaki derinlik deęerlerinin toplamının gözlem sayısına eşit olması gerekir. Yukarıdaki örnekte  $8+5+10=23$  olduęu görölmektedir.

Gövde-yaprak gösterimine göre, bu veri seti için 20.9' un aykırı deęer olduęu görölmektedir. Çünkü, 20.9 ile 25.2 gözlemleri arasında hiçbir veri yoktur, bir başka deyişle boşluk vardır.

Gövde-yaprak gösteriminin histograma göre en büyük avantajı aykırı deęerleri daha net bir şekilde görmemizi sağlamasıdır.

Gövde-yaprak gösteriminde satır sayısının etkin seçimi

- Veri setindeki gözlem sayısı
- Kapsanacak aralık
- Bazı bireysel kararlar

göz önüne alınarak yapılır.

### **Maksimum satır sayısının belirlenmesi**

Maksimum satır sayısı aşağıdaki formül yardımıyla hesaplanır

$$L = [10 \log_{10} n].$$

Burada,

$n$ : Gözlem sayısı ve

$[x]$ :  $x$ 'i geçmeyen en büyük tam sayı olarak tanımlanır.

### **Genişliğin belirlenmesi**

Genişlik (range) aşağıdaki formül yardımıyla hesaplanır

$$R = \max - \min.$$

## Aralık genişliğinin belirlenmesi

Aralık genişliği aşağıdaki formül yardımıyla hesaplanır

$$\frac{R}{L}$$

Yukarıda verilen örnek için

$$L = [10 \log_{10} 23] = [13.6] = 13$$

olarak hesaplanır. Bu nedenle, gövde-yaprak gösteriminde maksimum satır sayısı 13 olarak bulunur. Genişlik

$$R = 31.1 - 20.9 = 10.2$$

olarak hesaplanır.

Aralık genişliği ise

$$\frac{R}{L} = \frac{10.2}{13} = 0.78$$

olarak hesaplanır. Bunu da 10'un en yakın üssüne ( $1 \times 10^0$ ) yuvarlarsak aralık genişliği 1 olur.

## Diğer özellikler:

Gövde-yaprak gösteriminin en basit şekli

0 |  
1 |  
2 |  
3 |

olarak ifade edilir. Her bir satırda 0'dan 9'a kadar değerler yer alır.

Bazen satırlardaki gözlem sayısı çok fazla olabilir. Bu problemi çözmenin en etkin yolu her bir satırı

“\*” satırında 0 ile 4 arasındaki değerler

“.” satırında 5 ile 9 arasındaki değerler

yer alacak şekilde ikiye bölmektir.

0\*  
0.  
1\*  
1.  
2\*  
2.

**Not:** Bu durumda aralık genişliği  $5 \times 10$ 'un en yakın üssüne yuvarlanarak bulunur.

**Örnek:** Aşağıdaki veri seti için gövde-yaprak gösterimini oluşturalım.

32.0, 32.7, 32.9, 34.5, 35.1, 35.7, 35.9, 36.8, 37.3, 38.4, 38.7, 42.3, 43.6, 43.8, 46.5, 46.8,  
53.2, 53.8, 57.8, 59.3, 62.5, 64.3, 67.8, 68.7, 71.2, 73.5, 74.6, 77.5, 77.8, 79.2

Gözlem sayısı  $n = 30$

Maksimum satır sayısı  $L = [10 \log_{10} 30] = [14.7] = 14$

Genişlik  $R = 79.2 - 32.0 = 47.2$

Aralık genişliği ise

$$\frac{R}{L} = \frac{47.2}{14} = 3.37 \approx 10$$

olarak hesaplanır.

Buradan yukarıdaki veri seti için gövde-yaprak gösterimi

derinlik

11	3		2 2 2 4 5 5 5 6 7 8 8
(5)	4		2 3 3 6 6
14	5		3 3 7 9
10	6		2 4 7 8
6	7		1 3 4 7 7 9

ya da

3*		2 2 2 4
3.		5 5 5 6 7 8 8
4*		2 3 3
4.		6 6
5*		3 3
5.		7 9
6*		2 4
6.		7 8
7*		1 3 4
7.		7 7 9

şeklinde elde edilir.

**Not:** Bu şekilde de satırlarda hala çok fazla gözlem değeri olursa her bir satır 5'e bölünerek bu sorun giderilmeye çalışılır. Satırlar 5'e bölünürken

“\*” satırında 0,1,

“t” satırında 2,3,

“f” satırında 4,5,

“s” satırında 6,7 ve

“.” satırında 8,9 değerleri yer alır.