**2.2) Merkezi Eğilim Ölçüleri**

**Merkezi eğilim ölçüsü**, bir veri kümesinin “merkezini” ya da “tipik” değerini belirlemek amacıyla kullanılan istatistiksel bir kavramdır.

Bu ölçüler, verilerin genel bir eğilim veya ortalama davranışını anlamaya yardımcı olur.

Merkezi eğilim ölçüsü, verilerin çoğunlukla hangi noktada yoğunlaştığını ve hangi değerin tipik olduğunu belirlemeye çalışır.

**En sık kullanılan Merkezi Eğilim Ölçüleri**

1. Aritmetik Ortalama

2. Medyan (Ortanca)

3. Mod (Tepe Değer)

4. Kartiller (Çeyreklikler)

1. **Aritmetik Ortalama**

Aritmetik Ortalama en çok kullanılan merkezi eğilim ölçüsüdür.

Birimlerin belirli bir değişken bakımından aldıkları değerler toplamının birim sayısına bölünmesiyle hesaplanmaktadır.

Aritmetik ortalama, veri seti dağılımının simetrik olduğu bilindiğinde kullanılabilir.

Bir veri setinin **simetrik olması**, verilerin dağılımının ortalama etrafında dengeli bir şekilde dağılması anlamına gelir. Eğer veri setinizde aykırı değerler var ise **aritmetik ortalama** kullanmak tavsiye edilmez.

**Aritmetik Ortalama Örneği**

metin, yazı tipi, ekran görüntüsü, makbuz içeren bir resim

Açıklama otomatik olarak oluşturulduAritmetik Ortalama Örneği

**Aritmetik Ortalama Örnek Senaryo**

Web sitemizde kullanıcıların geçirdiği sürenin kaydı tutulmaktadır.

Kullanıcı sayısı ve toplam geçirilen süre verileri kullanarak, kullanıcıların ortalama ne kadar zaman geçirdiklerini hesaplayabiliriz.

(Kullanıcı Sayısı / Kullanıcıların Toplam Süresi)

Bunun sonucunda web sitemizde üretilen ve tüketilen içerikler için optimizasyon işlemi gerçekleştirebiliriz.

**2.** **Medyan (Ortanca)**

Bir veri setindeki değerler küçükten büyüğe doğru sıralandığında tam ortadaki değer medyan (ortanca) değerdir.

Ana kitledeki birimlerin sayısı çok fazla ise verilerin özetlenmesinde merkezi eğilim ölçüsü olarak medyan kullanılabilir.

Medyan, sınıflama ölçme düzeyi ile ölçülen değişkenler için kullanılmaz.

Aralıklı (interval), oran (ratio) ve sıralama ölçme düzeyinde ölçülen değişkenler için kullanılır.

(Örn. Renk — Cinsiyet vb. veriler sıralanamadığı için medyan kullanılamaz.)

 —

Birim sayısının tek veya çift olmasına göre medyanın bulunması değişmektedir.

İki durumda da ilk olarak eldeki veriler büyüklük sırasına göre (küçükten büyüğe veya büyükten küçüğe) sıraya konulur.

Birim sayısı N ile gösterilmek üzere, medyan hesaplaması:

* Eğer Veri sayısı **tekse**, medyan, ortada yer alan değeri alır.

yazı tipi, metin, beyaz, simge, sembol içeren bir resim

Açıklama otomatik olarak oluşturulduVeri setindeki terim sayısı **TEK** ise

* Eğer veri sayısı **çiftse**, ortadaki iki sayının ortalaması medyanı verir.

metin, yazı tipi, çizgi, beyaz içeren bir resim

Açıklama otomatik olarak oluşturulduVeri setindeki terim sayısı **ÇİFT** ise

**Medyan Örnekleri**

metin, yazı tipi, makbuz, beyaz içeren bir resim

Açıklama otomatik olarak oluşturulduTek Terimli Medyan Örneği

metin, yazı tipi, ekran görüntüsü, çizgi içeren bir resim

Açıklama otomatik olarak oluşturulduÇift Terimli Medyan Örneği

**2.1 Medyan ve Ortalama Karşılaştırılması**

Aritmetik ortalama, veri seti dağılımının simetrik olduğu durumlarda kullanılabilir.

Eğer veri setinde aykırı değerler olmasına rağmen aritmetik ortalama kullanılırsa işlem hatalı çıkacaktır.

**Örnek Senaryo**

metin, yazı tipi, ekran görüntüsü, beyaz içeren bir resim

Açıklama otomatik olarak oluşturulduAykırı Değer Bulunan Veri Seti

Elimizde bu şekilde bir veri seti bulunuyorsa **Medyan** kullanarak işlemler yapılmalıdır.

Veri setimiz içerisinde aykırı değer (190) olduğu için

Medyan: 13 (Doğru Sonuç),

Ortalama: 28,5 değerini almıştır.

**BİLGİ**: Eğer bu veri setinde aritmetik ortalama alarak işlemler yaparsak veri sonuçlarımız yanıltıcı olacaktır.

Bu konuyu karşımızdaki insanlara özetlerken ise “Veri setinin dağılımı simetrik olmadığı için Medyan kullanarak bu sonuç elde edilmiştir.” denilmelidir.

Ayrıca incelenmek üzere olan bir veri seti için araştırma gerçekleştirirken karşımızdaki insana mutlaka “Veri setinin dağılımı simetrik mi?” şeklinde sorulmalıdır.

Aksi halde Veri Bilimi Süreci tamamen hatalı olarak ortaya çıkacaktır.

**3. Mod (Tepe Değer)**

Bir veri grubunda en çok tekrarlanan değere mod (tepe değer) denir.

Tepe değerin hesaplanmasında birimlerin büyüklük sırasına konulması şartı olmamakla birlikte, bu işlemin yapılması tepe değerin bulunmasında kolaylık sağlar.

**NOT**: Bu tür değerler hesap makineleri ve/veya algoritmalarla hesaplanmaktadır.

**Örnek Soru**

metin, yazı tipi, ekran görüntüsü, beyaz içeren bir resim

Açıklama otomatik olarak oluşturulduMod Örneği

**Aritmetik Ortalama — Medyan(Ortanca)– Mod(Tepe Değer)**

metin, çizgi, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturulduVerilerin Dağılımını Gösteren Grafik

4. **Kartiller (Çeyreklikler)**

Küçükten büyüğe doğru sıralanmış verileri dört eşit parçaya bölen değerlere Kartiller (Çeyreklikler) denir.

Kartiller hem merkezi eğilim hem de dağılım ölçülerinde kullanılmaktadır.

**NOT**: Veri setinin ortalama değerinin 25% ilerisi ve 25% gerisini hesaplamak ve daha detaylı bilgi edinmek için kullanılmaktadır.

**Kartiller Gösterimi**

Veri setindeki değerlerin tamamını 100% olarak ele alırsak

Q1 = 25% (Birinci çeyrek)

Q2 = 50% (Medyan (Ortanca))

Q3 = 75% (Üçüncü çeyrektir)

metin, yazı tipi, el yazısı, beyaz içeren bir resim

Açıklama otomatik olarak oluşturulduKartiller Gösterimi

**Kartiller Örneği**

yazı tipi, ekran görüntüsü, metin, beyaz içeren bir resim

Açıklama otomatik olarak oluşturuldu

**KAYNAKLAR:**

* [Turkcell Academy / Geleceği Yazanlar](https://gelecegiyazanlar.turkcell.com.tr/user/login)
* Acikders.ankara.edu.tr