KONU ANLATIMLI SORU ÇÖZÜMLERİ

Sonuçları Yorumlama, Değişebilirlik ve Yanlılık

© Buse Hoca · Matematik Sınıfım (busehoca.com). Bu belge **CC BY-NC 4.0** lisanslıdır: ticari olmayan amaçlarla, kaynak gösterilerek çoğaltılabilir ve paylaşılabilir.

ÇÖZÜMLÜ SORU 1 Uç değerin etkisi ve doğru ölçüyü seçme

Bir bakkalın bir haftalık günlük müşteri sayıları şöyledir:
4, 5, 6, 6, 7, 8, 55
a) Veri setinin **aritmetik ortalamasını** bulunuz.
b) **Ortancasını** ve **tepe değerini** bulunuz.
c) **Açıklığı** hesaplayıp yorumlayınız.
d) Bu veri setini en iyi hangi ölçü temsil eder? Gerekçelendiriniz.

1 a) Aritmetik ortalama

Aritmetik ortalama = Verilerin toplamı : Veri sayısı

Toplam: 4 + 5 + 6 + 6 + 7 + 8 + 55
= 9 + 6 + 6 + 7 + 8 + 55
= 15 + 6 + 7 + 8 + 55
= 21 + 7 + 8 + 55
= 28 + 8 + 55
= 36 + 55 = **91**

Veri sayısı: 7

Ortalama = 91 : 7 = **13**

2 b) Ortanca ve tepe değer

Ortanca: Veriler zaten küçükten büyüğe sıralı:
4, 5, 6, **6**, 7, 8, 55

7 veri var; 7 tek sayı olduğundan tam ortadaki değer alınır. Ortadaki, **4. sıradaki** değerdir:
Ortanca = **6**

Tepe değer: En çok tekrar eden değere bakalım. Yalnızca **6** iki kez, diğerleri birer kez geçiyor.
Tepe değer = **6**

3 c) Açıklık ve yorumu

Açıklık = En büyük veri − En küçük veri
= 55 − 4 = **51**

Yorum: Açıklık 51 gibi çok büyük bir değerdir; oysa verilerin altı tanesi 4 ile 8 arasında sıkışmıştır. Bu kadar büyük bir açıklığın tek sebebi **55** değeridir.

55 sayısı diğer verilerden çok uzakta olduğu için bir **uç değerdir**. (Belki o gün bir indirim vardı ya da bir okul gezisi geldi.)

devamı →

ÇÖZÜMLÜ SORU 1 **Uç değerin etkisi ve doğru ölçüyü seçme** (devam)

4 **d) En iyi temsil eden ölçü**

Cevap: ORTANCA (6)

Gerekçe: Veri setinde **55** gibi bir **uç değer** vardır. Uç değerler aritmetik ortalamayı kendine doğru çeker.

Kanıtlayalım: Ortalama **13** çıktı. Oysa **yedi günün altısında** müşteri sayısı 13'ten **az**! Yani ortalama, haftanın tipik bir gününü hiç temsil etmiyor.

Ortanca ise **6**'dır ve verilerin çoğunluğu (4, 5, 6, 6, 7, 8) bu değerin çevresindedir. Uç değer ortancayı **etkilemez**; çünkü ortanca sıralamadaki **konuma** bakar, değerin büyüklüğüne değil.

Tepe değer de 6'dır ama yalnızca iki tekrara dayandığı için güçlü bir ölçü sayılmaz. Bu veri setinde **ortanca** en güvenilir ölçüdür.

Sonuç: a) 13 b) Ortanca 6, tepe değer 6 c) Açıklık 51; uç değerden kaynaklanıyor d) Ortanca; çünkü 55 bir uç değerdir

Dikkat: Ortalamayı bulduktan sonra "verilerin kaç tanesi bu değerin altında?" diye sor. Çoğu altındaysa ortalama temsil etmiyordur.

ÇÖZÜMLÜ SORU 2 **Ortalamalar eşitken açıklıkla karar verme**

Bir çiftçi, iki farklı tarlada yetişen domateslerin kütlelerini karşılaştırmak istiyor. Her tarladan rastgele seçilen 5 domatesin gram cinsinden kütleleri şöyledir:
Tarla A: 150, 160, 145, 155, 150
Tarla B: 130, 170, 140, 180, 140
a) İki tarlanın **aritmetik ortalamalarını** bulunuz.
b) İki tarlanın **açıklıklarını** bulunuz.
c) Hangi tarladaki domateslerin **değişkenliği** daha fazladır?
d) Çiftçi, kütlesi birbirine yakın domates yetiştirmek istiyorsa hangi tarlayı tercih etmelidir?

1 **a) Aritmetik ortalamalar**

Tarla A:

Toplam = 150 + 160 + 145 + 155 + 150

= 310 + 145 + 155 + 150

= 455 + 155 + 150

= 610 + 150 = **760**

Ortalama = 760 : 5 = **152 g**

Tarla B:

Toplam = 130 + 170 + 140 + 180 + 140

= 300 + 140 + 180 + 140

devamı →

ÇÖZÜMLÜ SORU 2 **Ortalamalar eşitken açıklıkla karar verme** (devam)

= 440 + 180 + 140

= 620 + 140 = **760**

Ortalama = 760 : 5 = **152 g**

İki tarlanın ortalaması TAM OLARAK AYNI: 152 g

Bu noktada duran biri "iki tarla arasında fark yok" derdi. Ama bu yanlış olurdu.

2 b) Açıklıklar

Açıklık = En büyük veri − En küçük veri

Tarla A:

En büyük: **160** · En küçük: **145**

Açıklık = 160 − 145 = **15 g**

Tarla B:

En büyük: **180** · En küçük: **130**

Açıklık = 180 − 130 = **50 g**

Aradaki fark çok büyüktür: B'nin açıklığı A'nınkinin **üç katından** fazladır.

3 c) Hangi tarlada değişkenlik fazla?

Tarla B'deki domateslerin değişkenliği daha fazladır.

Gerekçe: Açıklığı daha büyük olan veri seti **daha geniş bir dağılım** gösterir.

Tarla A'daki domatesler **145 g ile 160 g** arasında, yani dar bir aralıkta toplanmıştır. Tarla B'deki domatesler ise **130 g ile 180 g** arasında, çok daha geniş bir aralığa yayılmıştır.

Ortalamalar eşit olduğu hâlde dağılımlar tamamen farklıdır. **Ortalama merkezi, açıklık yayılımı** anlatır; ikisi birlikte tam bir resim verir.

4 d) Çiftçinin tercihi

Çiftçi Tarla A'yı tercih etmelidir.

Gerekçe: Çiftçi kütlesi **birbirine yakın** domates istiyor. Bu, verilerin **az değişkenlik** göstermesi demektir.

Tarla A'nın açıklığı **15 g** ile çok küçüktür; domatesler neredeyse aynı boydadır. Bu, paketleme ve satışta büyük kolaylık sağlar.

Tarla B'de ise 130 g'lık bir domatesle 180 g'lık bir domates yan yana çıkabilir; ürün **öngörülemez**.

Genel kural: İki veri grubunun ortalamaları eşitse karar, **açıklığa** bakılarak verilir. Tutarlılık isteniyorsa **açıklığı küçük** olan grup seçilir.

Sonuç: a) İkisi de 152 g b) A: 15 g, B: 50 g c) Tarla B d) Tarla A

Dikkat: Ortalamalar eşit çıktığında hemen açıklığa geç; asıl farkı orası gösterir.

ÇÖZÜMLÜ SORU 3 **Yanlı soru ve yanlı örneklem**

Bir öğrenci, okulundaki öğrencilerin spor alışkanlıklarını araştırmak istiyor. Şu adımları izliyor:
1. Anket sorusu olarak "Spor yapmayan öğrenciler neden bu kadar isteksiz?" sorusunu yazıyor.
2. Anketi yalnızca **okul basketbol takımındaki** 12 öğrenciye uyguluyor.
3. Sonuçta "Okulumuzdaki öğrencilerin çoğu düzenli spor yapıyor." diye yazıyor.
a) 1. adımdaki **sorunun** sorunu nedir? Tarafsız hâlini yazınız.
b) 2. adımdaki **örneklemin** sorunu nedir?
c) 3. adımdaki **sonucun** neden güvenilir olmadığını açıklayınız.
d) Araştırmanın doğru yapılması için **ne yapılmalıydı**? İki öneri yazınız.

1 **a) Sorudaki yanlılık**

Sorun: Soru yanlıdır (taraflıdır).

"Spor yapmayan öğrenciler neden bu kadar **isteksiz**?" sorusu iki hata birden yapıyor:

1. Spor yapmayanların **isteksiz olduğunu** daha baştan kabul ediyor. Oysa bunun araştırılması gerekir; sağlık sorunu, imkân yokluğu gibi başka sebepler olabilir.
2. "Bu kadar" ve "isteksiz" ifadeleri **suçlayıcı** bir dil taşıyor; katılımcıyı belirli bir yönde düşünmeye itiyor.

Tarafsız hâli:

"Haftada kaç gün spor yapıyorsunuz?"

ya da

"Spor yapma sıklığınızı etkileyen etkenler nelerdir?"

Tarafsız soru, cevabı **yönlendirmez**; yalnızca bilgi ister.

2 **b) Örneklemdeki yanlılık**

Sorun: Örneklem temsil edici değildir (seçim yanlılığı).

Anket yalnızca **okul basketbol takımındaki** öğrencilere uygulanmış. Bu grup, doğası gereği **zaten spor yapan** öğrencilerden oluşur.

Araştırma "okuldaki öğrencilerin" spor alışkanlığını ölçmek istiyorsa örneklemde **hiç spor yapmayan**, **ara sıra yapan** ve **düzenli yapan** öğrencilerin hepsi bulunmalıdır.

Ayrıca **12 öğrenci**, bir okulun tamamını temsil etmek için oldukça **küçük** bir sayıdır.

3 **c) Sonucun güvenilmezliği**

"Okulumuzdaki öğrencilerin çoğu düzenli spor yapıyor." cümlesi, **12 sporcunun** verisinden **tüm okula** yapılan bir **genellemedir**.

Bu sonuç güvenilir değildir; çünkü:

- Veriler **yanlı bir soruyla** toplanmıştır — cevaplar zaten etkilenmiştir.
- Veriler **yanlı bir gruptan** toplanmıştır — basketbol takımı okulu temsil etmez.
- Örneklem **çok küçüktür**.

devamı →

ÇÖZÜMLÜ SORU 3 **Yanlı soru ve yanlı örneklem** (devam)

Sonuç muhtemelen **olduğundan yüksek** çıkmıştır. Basketbol takımına sorulduğunda herkesin spor yaptığı görülür; bu, okulun tamamı için doğru değildir.

Önemli: Hesaplamalar doğru yapılmış olsa bile, **yanlı veriden doğru sonuç çıkmaz**.

4 **d) Ne yapılmalıydı?**

Öneri 1 — Soruyu tarafsızlaştır.

Yönlendirici, olumsuz ve yargı içeren kelimeler kullanılmamalıydı. "Haftada kaç gün spor yapıyorsunuz?" gibi **ölçülebilir** ve **tarafsız** bir soru sorulmalıydı.

Öneri 2 — Örneklemi genişlet ve çeşitlendir.

Anket, okuldaki **bütün sınıf düzeylerinden**, **rastgele seçilmiş** ve **yeterli sayıda** öğrenciye uygulanmalıydı. Örneğin her sınıf seviyesinden rastgele 30'ar öğrenci.

Ek öneri: Katılımcıların isimleri gibi **kişisel veriler** paylaşılmamalı; sonuçlar yalnızca sayılarla sunulmalıdır.

Bu düzeltmelerden sonra elde edilen sonuç, okulun geneli hakkında **güvenilir** bir bilgi verir.

Sonuç: a) Soru yanlı; tarafsızı "Haftada kaç gün spor yapıyorsunuz?" b) Örneklem temsil edici değil c) Yanlı veriden yapılan genelleme d) Tarafsız soru + geniş ve rastgele örneklem

Dikkat: Bir araştırmayı üç yerde denetle: SORU nasıl sorulmuş, KİMLERE sorulmuş, SONUÇ ne kadar geniş bir gruba genellenmiş.

ÇÖZÜMLÜ SORU 4 **Grafik ölçeği ve daire grafiği dilimleri**

a) Bir sütun grafiğinin dikey ekseni **0'dan değil 20'den** başlamaktadır. Bir öğrenci, yüksekliği 25 çizgisine kadar çıkan sütun için "Bu değer 25 − 20 = 5'tir." diyor. Bu yorum doğru mudur?
b) Ölçeklendirme hatası olan bir grafik hangi tür yanlış yorumlara yol açar? Bir örnek veriniz.
c) Bir kütüphanede haftalık ödünç verilen kitap sayıları: Pazartesi 9, Salı 6, Çarşamba 3, Perşembe 2'dir. Her günün **yüzdesini** hesaplayınız.
d) Bu veriler için çizilen daire grafiğinde Pazartesi **%45**, Salı **%30**, Çarşamba **%15**, Perşembe **%10** gösterilmiştir. Hangi dilimler **hatalı** çizilmiştir?

1 **a) Ölçek 0'dan başlamayınca**

Öğrencinin yorumu YANLIŞTIR.

Dikey eksen 20'den başlasa bile, eksendeki sayılar **gerçek değerleri** gösterir. Sütunun ucu **25** çizgisine geliyorsa o sütunun değeri **25**'tir.

devamı →

ÇÖZÜMLÜ SORU 4 **Grafik ölçeği ve daire grafiği dilimleri** (devam)

Öğrenci, sütunun **görünen yüksekliğini** değerle karıştırmıştır. Ölçek 20'den başladığı için sütunun görünen kısmı gerçekten de 5 birimlik bir yükseklik kadardır — ama bu, **sütunun değeri değil**, yalnızca grafikteki görüntüsüdür.

Doğru değer: 25

2 b) Ölçek hatasının yol açtığı yorumlar

Dikey eksen 0'dan başlamadığında **küçük farklar olduğundan çok büyük görünür**.

Örnek: Bir şirketin kârı Nisan'da 30, Mayıs'ta 25 birim olsun. Aradaki gerçek fark yalnızca **5 birimdir**; yani Mayıs, Nisan'ın yaklaşık **%83**'ü kadardır.

Ancak eksen **20'den** başlatılırsa sütunların görünen yükseklikleri 10 ve 5 olur. Grafiğe bakan biri Mayıs'ın Nisan'ın **yarısı** kadar olduğunu sanar!

Gerçek fark %17 civarındayken grafik %50'lik bir düşüş izlenimi verir. İşte bu yüzden ölçeklendirme hatası, reklamlarda ve yanıltıcı haberlerde sıkça kullanılan bir yöntemdir.

İkinci tip hata — belirsiz ölçek: Eksendeki aralıklar eşit değilse (0, 5, 10, 15, 25, 30 gibi) sütunların değeri **kesin olarak okunamaz**; sadece hangi aralıkta olduğu tahmin edilebilir.

3 c) Yüzdeleri hesaplama

Önce **toplam** kitap sayısını bulalım:

9 + 6 + 3 + 2 = **20**

Her günün yüzdesi = (O günün sayısı / Toplam) × 100

Pazartesi: 9/20 = 45/100 = **%45**

(Payda 20'yi 100'e tamamlamak için 5 ile çarparız: 9 × 5 = 45)

Salı: 6/20 = 30/100 = **%30**

Çarşamba: 3/20 = 15/100 = **%15**

Perşembe: 2/20 = 10/100 = **%10**

Kontrol: 45 + 30 + 15 + 10 = **100** ✓ Yüzdelerin toplamı %100 olmalıdır.

4 d) Dilimleri karşılaştırma

Hesapladığımız değerlerle grafikteki değerleri karşılaştıralım:

Pazartesi: hesaplanan **%45** — grafikte **%45** ✓

Salı: hesaplanan **%30** — grafikte **%30** ✓

Çarşamba: hesaplanan **%15** — grafikte **%15** ✓

Perşembe: hesaplanan **%10** — grafikte **%10** ✓

Cevap: Hiçbir dilim hatalı çizilmemiştir. Daire grafiği verilere tamamen uygundur.

Buradan çıkarılacak ders: Bir grafiğin hatalı olduğunu **varsaymak** yerine her dilimi **tek tek hesaplayıp** karşılaştırmak gerekir. Soru "hangileri hatalı?" diye sorduğu için hata aramaya eğilim duyarız; oysa doğru cevap "hiçbiri" de olabilir.

devamı →

ÇÖZÜMLÜ SORU 4 **Grafik ölçeği ve daire grafiği dilimleri** (devam)

Uyarı: Yüzdelerin toplamının %100 çıkması, grafiğin doğru olduğunu **garanti etmez**. Örneğin Çarşamba %20, Perşembe %5 yazılsaydı toplam yine 100 olurdu ama iki dilim de hatalı olurdu. Bu yüzden toplam kontrolü **yeterli değildir**.

Sonuç: a) Yanlış; değer 25'tir b) Küçük farklar çok büyük görünür c) %45, %30, %15, %10 d) Hiçbiri hatalı değil

Dikkat: Daire grafiğinde toplamın %100 çıkması doğruluğu garanti etmez; her dilimi ayrı ayrı hesapla.