

# Ödev: Veri Analizi ve Manipülasyonu

## Amaç

Bu ödevde öğrenciler, karmaşık veri manipülasyonu, veri temizleme ve analiz becerilerini geliştirecekler. Çeşitli zorluk seviyelerinde alt görevler içeren bir veri analizi süreci üzerinden gidilecektir.

## Gerekli Kütüphaneler

- Pandas
- NumPy
- Matplotlib veya Seaborn (Görselleştirme için)

## Veri Setleri

Öğrenciler aşağıdaki veri setlerini bulup kullanabilir veya kendi verilerini oluşturabilirler:

1. **Satış Verisi:** Her bir satırda farklı ürünlerin satıldığı bir veri seti.
  - o Örnek kolonlar: `tarikh`, `ürün_kodu`, `ürün_adi`, `kategori`, `fiyat`, `adet`, `toplam_satis`
2. **Müşteri Verisi:** Müşterilerle ilgili verileri içerir.
  - o Örnek kolonlar: `musteri_id`, `isim`, `cinsiyet`, `yas`, `sehir`, `harcama_miktari`

## Görevler

### Görev 1: Veri Temizleme ve Manipülasyonu (%25)

1. Eksik verileri ve aykırı (outlier) verileri analiz edip temizleyin. Eksik verileri tamamlamak için çeşitli yöntemleri (ortalama, medyan gibi) kullanarak eksiklikleri doldurun.
2. Fiyat ve harcama gibi değişkenler için aykırı değerleri tespit edip verisetinden çıkarın veya aykırı değerleri belirli bir aralık içine çekin.
3. Müşteri verisi ile satış verisini `musteri_id` üzerinden birleştirerek geniş bir veri seti oluşturun.

### Görev 2: Zaman Serisi Analizi (%25)

1. Satış verisi üzerinde haftalık ve aylık bazda toplam satış ve ürün satış trendlerini analiz edin.
2. `tarikh` sütununu kullanarak, her ayın ilk ve son satış günlerini bulun. Ayrıca, her hafta kaç ürün satıldığını hesaplayın.
3. Zaman serisindeki trendleri tespit etmek için grafikler çizdirin (örneğin: aylık satış artışı veya düşüşü).

### Görev 3: Kategorisel ve Sayısal Analiz (%25)

1. Ürün kategorilerine göre toplam satış miktarını ve her kategorinin tüm satışlar içindeki oranını hesaplayın.

2. Müşterilerin yaş gruplarına göre satış eğilimlerini analiz edin. (Örnek yaş grupları: 18-25, 26-35, 36-50, 50+)
3. Kadın ve erkek müşterilerin harcama miktarlarını karşılaştırın ve harcama davranışları arasındaki farkı tespit edin.

#### Görev 4: İleri Düzey Veri Manipülasyonu (%25)

1. Müşterilerin şehir bazında toplam harcama miktarını bulun ve şehirleri en çok harcama yapan müşterilere göre sıralayın.
2. Satış verisinde her bir ürün için ortalama satış artışı oranı hesaplayın. Bu oranı hesaplamak için her bir üründe önceki aya göre satış değişim yüzdesini kullanın.
3. Pandas `groupby` ile her bir kategorinin aylık toplam satışlarını hesaplayın ve değişim oranlarını grafikte gösterin.

#### Görev 5: Ekstra (BONUS)

1. **Pareto Analizi:** Satışların %80'ini oluşturan ürünleri belirleyin (80/20 kuralını uygulayın). Bu ürünleri grafikte gösterin.
2. **Cohort Analizi:** Müşterilerin satın alım alışkanlıklarını analiz etmek için Pandas ile cohort analizi yapın. Örneğin, ilk kez satın alan müşterilerin tekrar alım oranlarını inceleyin.
3. **Tahmin Modeli:** Aylık veya haftalık satış miktarlarını tahmin etmek için basit bir regresyon modeli (örneğin Linear Regression) uygulayın. `sklearn` kullanarak `train/test split` işlemi ile modeli eğitin ve modelin doğruluğunu ölçün.

#### Teslim

- Yapılan analizlerin ve görevlerin kodlarını açıklayıcı yorumlarla birlikte teslim edin.
- Grafiklerin ve bulguların yorumlarını bir rapor halinde ekleyin.
- Verilerin ve sonuçların görselleştirildiği bir sunum dosyası hazırlayın (isteğe bağlı).