SATIŞ VERİ ANALİZİ VE KÜMELEME PROJESİ

# 1. Giriş

Bu projede, bir satış verisi seti kullanılarak temel veri analizi, görselleştirme ve veri madenciliği teknikleri uygulanmıştır.   
Veri seti, farklı ürün kategorilerine, müşterilere, şehir ve eyaletlere ait satış bilgilerini içermektedir.   
Projenin amacı, bu veriler üzerinde anlamlı desenler bulmak, kümeleme yapmak ve belirli gruplar için gelecek dönem satış tahminlerinde bulunmaktır.

# 2. Veri Seti Hakkında

Kullanılan veri seti, Amerika Birleşik Devletleri’ndeki satışları temsil etmektedir ve şu sütunları içermektedir:  
- Order Date (Sipariş Tarihi)  
- Ship Date (Gönderim Tarihi)  
- Sales (Satış Tutarı)  
- Ship Mode, Segment, Region, Category, Sub-Category gibi kategorik sütunlar.  
Veri setinde toplam 9800 satır veri bulunmaktadır.

# 3. Veri Temizleme ve Ön İşleme

Veri seti incelendiğinde, sadece Postal Code sütununda 11 eksik veri gözlemlenmiştir.   
Bu sütun analize dahil edilmediği için eksik veriler problem yaratmamıştır. Tarih sütunları uygun formatta datetime'a çevrilmiştir.  
Kategorik sütunların benzersiz değerleri analiz edilerek veri setinin yapısı daha iyi anlaşılmıştır.

# 4. Görselleştirmeler

Satışların zamana göre dağılımı incelenmiş, toplam satışların yıllar içinde nasıl değiştiği analiz edilmiştir.   
Ayrıca, farklı kategorilere göre toplam satışlar bar grafik yardımıyla gösterilmiştir.   
Bu sayede en çok satış yapılan kategori ve ürün grupları görsel olarak anlaşılır hale getirilmiştir.

# 5. K-Means Kümeleme Analizi

Category ve Sub-Category bazında ortalama satışlar hesaplanarak bu değerler StandardScaler ile ölçeklendirilmiştir.  
Elbow yöntemi ile ideal küme sayısı belirlenmiş ve KMeans algoritması ile ürün grupları 3 kümeye ayrılmıştır.  
Elde edilen kümeler scatterplot ile görselleştirilmiştir. Bu sayede benzer ortalama satış düzeylerine sahip ürün grupları birlikte incelenmiştir.

# 6. Gelecek Dönem Satış Tahmini

Segment, Kategori ve Bölge bazında geçmiş veriler üzerinden satışların zaman serisi kullanılarak bir sonraki döneme (bir ay sonrası) dair tahminler yapılmıştır.   
Tahminler grafikler ile desteklenmiş ve yorumlanabilir hale getirilmiştir. Örnek grafikler ekte sunulmuştur.

# 7. Alt Gruplar Bazında Analiz

Yalnızca toplam satışlara değil, alt segmentlere (örneğin Consumer, Corporate, Home Office), farklı kategorilere (Office Supplies, Technology, Furniture)   
ve bölgelere (East, West, Central, South) göre ayrı ayrı analizler yapılmıştır.  
Her bir grup için satış trendi ve tahmin edilen satış miktarları grafiklerle sunulmuştur.  
Bu sayede detaylı bir öngörü sağlanmış ve projenin kapsamı genişletilmiştir.

# 8. Sonuç

Bu proje kapsamında yapılan analizlerle satış verisinin yapısı anlaşılmış, ürün ve müşteri segmentleri kümeleme yoluyla gruplandırılmış,   
gelecek dönem için satış tahminleri gerçekleştirilmiş ve bu tahminler segment, kategori ve bölgelere göre ayrı ayrı görselleştirilmiştir.  
Proje, veri madenciliği uygulamaları açısından temel teknikleri başarılı şekilde içermektedir.