**Veri Analizi ve Manipülasyonu Ödev Raporu**

**Giriş**

Bu rapor, veri analizi sürecinde gerçekleştirilen adımları, kullanılan yöntemleri ve elde edilen sonuçları detaylı bir şekilde açıklamaktadır. Satış ve müşteri verileri üzerinde veri temizleme, analiz ve görselleştirme adımlarını içermektedir. Ödev kapsamında dört ana görev ve ek olarak bir bonus görev gerçekleştirilmiştir.

**Görev 1: Veri Temizleme ve Manipülasyonu**

* Satış ve müşteri veri setlerine genel bir bakış yapılmış ve değişkenlerin tip bilgileri, satır sütun sayıları, eksik değerler, aykırı değerler incelenmiştir.
* **Eksik Veriler:** Veri setlerinde eksik değer bulunmamıştır.

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

* **Aykırı Değerler:**

Aykırı değerlerin tespiti için boxplot ve IQR (çeyreklik aralığı) yöntemi kullanılmıştır. Harcama miktarı ve fiyat değişkenlerinde aykırı değer bulunmamış; ancak toplam satış değişkeninde 35 aykırı değer tespit edilmiştir.

metin, diyagram, ekran görüntüsü, dikdörtgen içeren bir resim

Açıklama otomatik olarak oluşturuldu

Aykırı değerler alt ve üst sınırların dışına çıkmayacak şekilde sınır değerleriyle değiştirilmiştir. Winsorization yöntemi ile düzeltilmiştir. İşlemden sonra yeniden aykırı değerlerin kontrolü yapılmış ve yöntem doğrulanmıştır.

metin, ekran görüntüsü, diyagram, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

* **Birleştirme:** Satış ve müşteri verileri musteri\_id anahtarına göre birleştirilmiş ve yeni bir veri seti oluşturulmuştur.

**Sonuç:** Aykırı değerlerin düzeltilmesi ve veri setinin birleştirilmesi, daha tutarlı bir analiz yapılması için önemli bir adımdır.

**Görev 2: Zaman Serisi Analizi**

* **Trend Analizleri:**

Haftalık ve aylık bazda toplam satış ve ürün adet trendleri analiz edilmiştir. Bu analizin daha kapsamlı olması için hem kategori bazında hem ürün bazında incelemeler gerçekleştirilmiştir. Bu analizin görselleştirilmesi Grafikler kısmında bulunmaktadır.

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Haftalık ve aylık bazda kategorilere ait toplam satış ve adet bilgilerinin ilk 10 verisi bu şekildedir. Veriyi yorumlayacak olursak 2022 Kasım ayının 3. Haftasında Ev aletleri satışı bir önceki haftaya göre ani şekilde artmış, aynı şekilde Kırtasiye kategorisindeki satışlar da Kasım ayının 2. Haftası ani şekilde artmıştır. Aylık veriyi incelediğimizde de bu artışların yansıdığını kasım ayında en çok ev aletleri ve kırtasiye kategorisinde satış yapıldığını gözlemleyebilmekteyiz.

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

Açıklama otomatik olarak oluşturuldu metin, ekran görüntüsü, menü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Her ayın ilk ve son satış günleri bulunmuş ve haftalık satış adetleri hesaplanmıştır.

Verilerin ilk 5 satırı bu şekildedir.

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldumetin, yazı tipi, ekran görüntüsü, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldumetin, yazı tipi, ekran görüntüsü, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

* **Grafikler:**

Haftalık ve aylık bazda toplam satış ve ürün adet trendlerinin sonuçları çizgi grafiği ile görselleştirilmiştir.

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

Açıklama otomatik olarak oluşturuldu

ekran görüntüsü, metin, renklilik, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

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

Açıklama otomatik olarak oluşturuldu

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

Açıklama otomatik olarak oluşturuldu

Aylık ve haftalık olarak toplam satış ve ürün satış trendleri benzer şekilde analiz edilmiştir.

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

Açıklama otomatik olarak oluşturuldu

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

Açıklama otomatik olarak oluşturuldu

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

Açıklama otomatik olarak oluşturuldu

metin, diyagram, ekran görüntüsü, daire içeren bir resim

Açıklama otomatik olarak oluşturuldu metin, diyagram, ekran görüntüsü, daire içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Önemli Bulgular:**

1. Aylık bazda belirgin bir yükseliş veya düşüş trendi gözlemlenmiştir.
2. Bazı haftalarda satış adetlerinde belirgin dalgalanmalar mevcuttur.

**Görev 3: Kategorisel ve Sayısal Analiz**

* **Kategorilere Göre Satışlar:**
  + Ürün kategorilerinin toplam satışlar içindeki oranları hesaplanmıştır.
  + Örneğin, bazı kategoriler toplam satışların %40’ını oluşturmuştur.
* **Yaş Gruplarına Göre Satış Eğilimleri:**
  + En yüksek satış oranı **26-35 yaş grubu** müşterilerinden elde edilmiştir.
* **Cinsiyetlere Göre Harcama Davranışları:**
  + Kadın ve erkek müşterilerin harcama davranışları analiz edilmiş ve cinsiyete göre belirgin farklılıklar tespit edilmiştir.

**Grafikler:**

* Yaş grupları ve kategorilere göre pasta grafikleri oluşturulmuştur.
* Cinsiyetlerin kategori bazında harcama davranışları bar grafik ile görselleştirilmiştir.

**Görev 4: İleri Düzey Veri Manipülasyonu**

* **Şehir Bazında Harcama:**
  + Şehirler toplam harcama miktarına göre sıralanmıştır. En çok harcama yapan şehirler belirlenmiştir.
* **Ürünlerin Ortalama Satış Artışı Oranı:**
  + Aylık satış artış oranları hesaplanmış ve ortalama değerler ürün bazında raporlanmıştır.
* **Kategorilerin Aylık Toplam Satış Değişim Oranı:**
  + Her bir kategori için değişim oranları grafikle gösterilmiştir.

**Sonuç:**

* Şehir bazında harcama analizleri, bölgesel satış stratejilerinin geliştirilmesine yönelik içgörüler sunmuştur.
* Ürün ve kategori bazlı değişim oranları, hangi alanlarda daha fazla odaklanılması gerektiğini göstermiştir.

**Görev 5: Ekstra (BONUS)**

* **Pareto Analizi:**
  + Satışların %80’ini oluşturan ürünler belirlenmiş ve bir bar grafikte görselleştirilmiştir.
* **Cohort Analizi:**
  + Müşterilerin ilk alışveriş aylarına göre tekrar alım oranları analiz edilmiştir. Cohort analizi, müşteri davranışlarının sürekliliği hakkında bilgi vermiştir.
* **Tahmin Modeli:**
  + Aylık toplam satış tahmini için **Linear Regression** kullanılmıştır.
  + Modelin başarım metrikleri:
    - **MAE:** 5.32
    - **MSE:** 36.71
    - **R²:** 0.89

**Sonuç:** Tahmin modeli, gelecekteki satış miktarlarını öngörmek için temel bir başlangıç sağlamıştır.