**Mining Retail Behavior in Istanbul Malls Using Consumer Sentiment Indices**

Bu projede, İstanbul’daki 10 AVM’den 2021–2023 yılları arasında toplanan alışveriş verileri ile TÜİK ve TCMB’nin yayımladığı ekonomik güven endeksi verileri birlikte analiz edilecektir. Alışveriş verileri; müşteri kimliği, alışveriş tarihi, ürün kategorisi, miktar, fiyat, ödeme türü gibi bilgileri içermektedir. Bu verilere ek olarak çalışmada kullanılmak üzere ek özellikler türetilecektir.

Ön analizlerde veri setinin boyutu ve değişken yapısı incelenecek, temel istatistikler ve frekans analizleri yapılacak, sonuçlar görselleştirilecektir. Harcamaların gün, ay, mevsim ve tatil durumuna göre dağılımı, günlük ve haftalık harcama trendleri, ödeme türlerinin demografik farklılıkları, AVM’ler arasında müşteri sayısı, toplam harcama ve kategori bazlı farklılıklar analiz edilecektir. Fatura başına ortalama harcama ve ürün sayısı, yüksek harcamalı alışverişler de ayrıca incelenecektir. Ekonomik göstergelerin zaman içindeki değişimi de görselleştirilerek analiz edilecektir.

Bunlara ek olarak, ürün kategorisi, ödeme yöntemi, AVM ve müşteri bazında ekonomik göstergelerden etkilenme analizleri yapılacak; müşteri kümeleri üzerinden ekonomik güvene dayalı tahmin modelleri geliştirilecektir.

Hipotezler:  
• Ekonomik güven arttıkça harcama ve alışveriş sıklığı artar.  
• Güven düştüğünde lüks harcamalar azalır, temel ihtiyaçlar sabit kalır.  
• Erkekler ekonomik dalgalanmalara daha duyarlıdır.  
• Gençler değişimlere daha duyarlıdır.  
• Güven azaldığında kredi kartı kullanımı artar.

Müşteri segmentasyonu (K-means) yapılarak, her profil için ekonomik güvene dayalı tahmin modelleri (Seq2Seq LSTM, N-BEATS) kurulacak ve modeller R², RMSE, MAPE ile değerlendirilecektir.

**Referans Makale ve Veri Kaynağı:**

* Alışveriş Merkezi Veri Seti:  
  <https://www.kaggle.com/datasets/mehmettahiraslan/customer-shopping-dataset>
* Ekonomik Güven Endeksleri:  
  <https://data.tuik.gov.tr/Bulten/Index?p=Ekonomik-Guven-Endeksi-Mayis-2025-54101>

**Proje Planlaması:**

* 26 Mayıs – 1 Haziran: Proje planlaması, veri toplama
* 2 – 8 Haziran: Veri temizliği, bütünleştirme
* 9 – 15 Haziran: Veri dönüştürme, tanımlayıcı istatistikler ve görselleştirme
* 16 – 22 Haziran: Tanımlayıcı istatistikler ve görselleştirme, ara rapor yazımı
* 23 – 29 Haziran: Tanımlayıcı istatistikler ve görselleştirme, ara rapor yazımı
* 30 Haziran – 6 Temmuz: Ara rapor yazımı, modelleme
* 7 – 13 Temmuz: Modelleme, sonuç değerlendirme ve analizi
* 14 – 20 Temmuz: Sonuç değerlendirme ve analizi, sonuçların görselleştirilmesi, sözlü sunum hazırlığı
* 21 – 27 Temmuz: Sözlü sunum hazırlığı, son rapor yazımı
* 28 – 31 Temmuz: Son rapor yazımı
* 2 Haziran – 27 Temmuz (paralel): Literatür taraması (2 haftada bir sistematik okuma)

Bu projede, İstanbul’daki alışveriş merkezlerinde müşteri alışveriş davranışları; tarih, demografi, kategori ve ödeme tercihleri gibi faktörler incelenerek TÜİK ve TCMB’den alınan ekonomik güven endeksleriyle ilişkilendirilecektir. Amaç, ekonomik göstergelerin müşteri harcamaları, alışveriş sıklığı, kategori ve ödeme tercihlerine etkisini ortaya koymak; ayrıca müşteri kümeleri oluşturarak farklı müşteri profilleri için tahmin modelleri geliştirmektir.

**Kullanılacak Veriler:**

* Kaggle’dan alınmış İstanbul’daki 10 alışveriş merkezinden 2021-2023 yılları arasında toplanan müşteri, tarih, kategori, fiyat vb. bilgileri içeren alışveriş verisi
* TÜİK ve TCMB’den alınmış ekonomik güven endeksleri verileri
* Tarih bazlı türetilen özellikler (mevsim, tatil, hafta içi vb.)
* Alışveriş başına hesaplanan toplam harcama (quantity × price)

**Veri Temizleme:**  
Eksik, tutarsız veya hatalı kayıtlar tespit edilip uygun yöntemlerle düzeltilecek veya veri setinden çıkarılacaktır. (uç değer kontrolü, tutarsızlıkların düzeltilmesi, vb.)

**Veri Dönüşümü:**  
Farklı formatlardaki tarih, sayı ve kategori verileri birbirleriyle uyumlu hale getirilecek, analizler ve modelleme için gerekirse encoding ve normalizasyon işlemleri uygulanacaktır.

**Veri Bütünleştirme:**  
Alışveriş verisi ile ekonomik güven endeksleri, tarih bazında birleştirilecek; analizlerde kullanılacak ek özellikler (season, is\_weekday, is\_holiday, total\_price) türetilecektir.

**Ön Analiz (EDA):**

* Veri setinin boyutu ve yapısı incelenecek.
* Değişkenler sayısal ve kategorik olarak sınıflandırılacak.
* Sayısal değişkenler için temel istatistikler, kategorik değişkenler için frekans analizleri yapılacak ve tüm bu analizler görselleştirilecektir.
* Alışverişlerin gün, ay, mevsim ve tatil durumuna göre dağılımı analiz edilecek.
* Günlük ve haftalık harcama trendleri incelenecektir.
* Kategori bazlı harcama ve satış adetleri, demografik ilişkilerle değerlendirilecektir.
* Ödeme türlerinin dağılımı ve demografik/harcama bazlı farklılıklar analiz edilecektir.
* AVM’lerin müşteri sayısı, toplam harcama, ürün miktarı ve popüler kategorileri değerlendirilecektir.
* Fatura başına ortalama harcama ve ürün sayısı, yüksek harcama ve ürün sayısına sahip faturalar incelenecektir.
* Aktif müşteriler ve alışveriş alışkanlıkları (alışveriş sayısı, harcama tutarı, AVM tercihleri, tarihler, kategori tercihleri) belirlenecektir.
* Ekonomik göstergelerin zaman içindeki değişimi analiz edilecektir.
* Tüm bu analizler görselleştirilecektir ancak analiz sayısının fazla olmasından ötürü görsellerin tamamı, çalışmanın devamında yazılacak raporlara dahil edilmeyebilir. Dahil edilmeyen analizler projenin GitHUB repository’sinde “Additional Material” olarak yer alacaktır.

**Müşteri Segmentasyonu:**

* K-means algoritması ile toplam harcama, alışveriş sayısı, tarih, AVM, kategori ve ödeme yöntemi gibi çalışmanın devamında seçilecek değişkenlere göre müşteri kümeleri oluşturulacaktır.
* Küme sayısı Elbow ve Silhouette yöntemleriyle belirlenecek, ardından kümeler yorumlanarak segmentlere göre müşteri profilleri çıkarılacak ve ekonomik göstergelere tepkileri analiz edilecektir.

**Analizler:**

* Ürün kategorisi, ödeme yöntemi, AVM, fatura ve müşteri bazlı ekonomik göstergelerden etkilenme analizleri ve görselleştirilmesi
* Alışverişlerin tatil/gün içi/hafta sonu gibi zaman bazlı ekonomik göstergelerden etkilenme trend analizleri ve görselleştirilmesi
* Oluşan kümeler üzerinden farklı müşteri profilleri için ekonomik göstergelere dayalı tahmin modelleri (Seq2Seq LSTM, N-BEATS) geliştirilmesi
* Sonuçların R2, RMSE ve MAPE metrikleri ile değerlendirilmesi ve görselleştirilmesi

**Hipotezler ve Beklenen Sonuçlar:**

* Ekonomik güven arttıkça harcama ve alışveriş sıklığı artacaktır.
* Ekonomik güven düşüş dönemlerinde lüks ve esnek tüketim azalırken, temel ihtiyaç harcamaları daha sabit kalacaktır.
* Erkeklerin harcama davranışı ekonomik göstergelere daha duyarlı olacaktır.
* Genç tüketiciler ekonomik değişimlere daha hassas tepkiler verecektir.
* Düşen ekonomik güvenle kredi kartı kullanımında artış gözlenecektir.