

Ece Teker

Mi-POWER Women Empowerment Program

Final Case

## Restoran Gelir Tahmini

### Problem Tanımı ve Literatür Taraması (1. Hafta)

İlk hafta, ekip olarak odaklanacağımız problemi belirlemek amacıyla çeşitli platformlardan birçok veri setini inceledik. Bu süreçte, dikkatimizi çeken bazı örnek veri setleri şunlardır:

[Deprem Senaryosu Analiz Sonuçları](#)

[Yemek Siparişi Süresi Tahmini](#)

[Restoran Gelir Tahmini](#)

[Disney Filmleri Gelir Tahmini](#)

Yapılan tartışmalar sonucunda, restoran gelir tahmininin hem uygulanabilirlik açısından pratik olduğunu hem de örneklerle erişilebilirlik açısından daha uygun olduğunu belirledik. Bu bağlamda, öncelikle Kaggle platformunda bulunan restoran gelir tahmini veri setini kullanarak yapılmış çalışmaları inceledik. Ardından, derste paylaşılan Titanik veri seti üzerindeki Keşifsel Veri Analizi (EDA) kodlarını inceleyerek, kullanacağımız programlama dili ve kütüphaneler üzerine karar verdik.

Bu süreç, projemizin temellerini atmamıza yardımcı oldu ve ileriye dönük adımlarımızı planlamamızı sağladı.

### Veri Toplama ve Ön İşleme (2. ve 3. Hafta)

Verimizi Kaggle'dan indirip github üzerinden bir [repository](#) oluşturduk. Burada bir Jupiter Notebook açarak veriyi incelemeye başladık.

Çıktılara dayanarak dataframe sınıfına ait bu nesne:

- 1000 satır gözlem içerir. index 0 ile 999 arasındadır ( RangeIndex: 1000 entries , 0 to 999 )

- 8 sütun içerir.

1. Number\_of\_Customers (Müşteri Sayısı)

- Restoranı ziyaret eden müşterilerin sayısını belirtir.

- Yüksek müşteri sayısı genellikle daha yüksek gelirle ilişkilidir. Müşteri sayısı, restoranın popülaritesini ve müşteri tabanının genişliğini gösterir.

2. Menu\_Price (Menü Fiyatı)

- Restorandaki ortalama menü fiyatlarını gösterir.
- Menü fiyatları, gelir tahmininde önemli bir faktördür. Yüksek fiyatlar daha yüksek gelir sağlayabilir, ancak müşteri sayısını da etkileyebilir.

### 3. Marketing\_Spend (Pazarlama Harcaması)

- Pazarlama faaliyetlerine ilişkin harcamaları belirtir.
- Pazarlama harcamaları, müşteri çekme ve gelir artırma stratejilerinde kritik bir rol oynar. Daha yüksek pazarlama harcamaları genellikle daha fazla müşteri çekebilir.

### 4. Cuisine\_Type (Mutfak Türü)

- Restoranın sunduğu mutfağın türünü belirtir (örneğin, İtalyan, Meksika, Japon, Amerikan).
- Mutfak türü, belirli müşteri segmentlerini çekebilir ve restoranın popülaritesini etkileyebilir. Farklı mutfak türleri farklı fiyatlandırma stratejilerine sahip olabilir.

### 5. Average\_Customer\_Spending (Ortalama Müşteri Harcaması)

- Müşteri başına ortalama harcamayı gösterir.
- Ortalama müşteri harcaması, gelir tahmini için kritik bir faktördür. Müşterilerin ne kadar harcadığı, toplam geliri doğrudan etkiler.

### 6. Promotions (Promosyonlar)

- Promosyonların yapıp yapılmadığını gösteren ikili bir göstergedir (0 veya 1).
- Promosyonlar, müşteri çekmek ve satışları artırmak için yaygın olarak kullanılır. Promosyonların varlığı (1), satışlarda geçici bir artışa neden olabilir.

### 7. Reviews (Yorumlar)

- Restoranın aldığı yorumların sayısını belirtir.
- Yorumlar, restoranın itibarını ve müşteri memnuniyetini yansıtır. Daha fazla olumlu yorum, daha fazla müşteri çekebilir ve geliri artırabilir.

### 8. Monthly\_Revenue (Aylık Gelir)

- Restoranın simüle edilmiş aylık gelirini gösterir.
- Bu, tahmin için hedef değişkendir ve yukarıdaki faktörlerin etkilediği sonucu gösterir. Amaç, diğer faktörlere dayanarak bu değeri tahmin etmektir.

DataFrame'in hiçbir sütununda eksik(null) veri bulunmamaktadır. Veri analizi ve modelleme işlemi için eksik veri imputation (doldurma) veya veri temizleme adımlarına gerek yok.

```
Aykırı Değerler - Monthly_Revenue:
54      -28.977809
84      563.381332
104     -27.610276
110     -25.453430
605     -28.920890
Name: Monthly_Revenue, dtype: float64
```

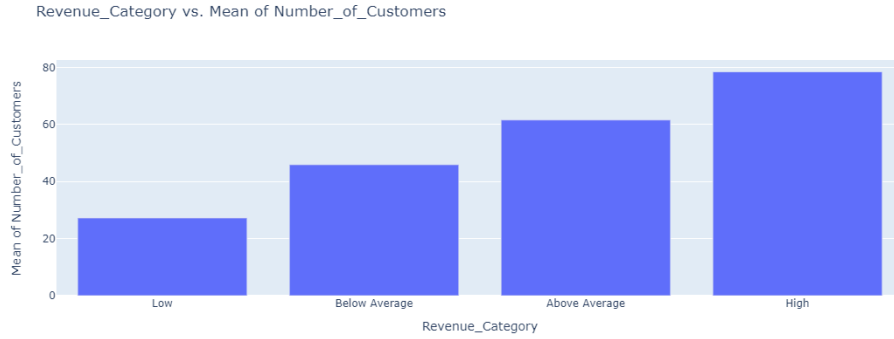
*Resim 1: Aykırı değer araştırması*

Resim 3'te görülen aykırı değerleri tespit edip veri setimizden çıkardık.

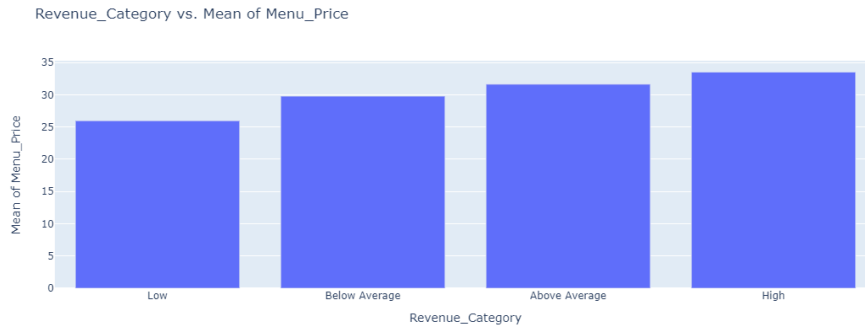
Son olarak da yazdığımız özel fonksiyonu kullanarak Number\_of\_Customers: numeric, Menu\_Price: numeric, Marketing\_Spend: numeric, Cuisine\_Type: categorical, Average\_Customer\_Spending: numeric, Promotions: numeric but categorical, Reviews: numeric, Monthly\_Revenue: numeric şeklinde kategorik ve numerik değer ayrıştırmasını tamamladık.

### Veri Analizi ve Görselleştirme (3. ve 4. Hafta)

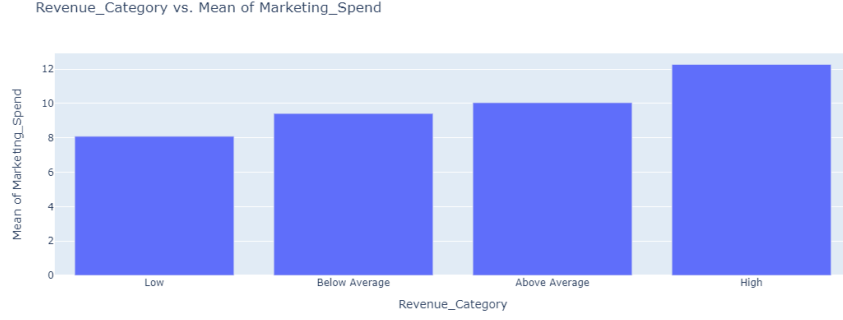
Hedef değişkenimiz sürekli bir değer olduğu için, analizimizi daha iyi yapmak adına restoran gelirlerini dört kategoriye ayırdık: Low (Düşük), Below Average (Ortalama Altı), Above Average (Ortalama Üstü) ve High (Yüksek). Bu kategoriler, gelir dağılımını daha anlaşılır hale getirirken, verilerin yorumlanmasını da kolaylaştırdı. Göze çarpan en önemli bulguları aşağıda bulabilirsiniz:



*Resim 2: Gelir kategorisi vs. Ortalama Müşteri Sayısı*



*Resim 3: Gelir kategorisi vs. Ortalama Menü Fiyatı*



Resim 4: Gelir kategorisi vs. Ortalama Pazarlama Harcaması

Resim 2, 3 ve 4'te görüldüğü üzere ortalama müşteri sayısı, menü fiyatı ve pazarlama harcamaları arttıkça gelirin de arttığını gözlemledik. Buna dayanarak aylık gelir ve bu değişkenler arasında bir bağlantı olduğunu söyleyebiliriz.

### Modelleme ve Makine Öğrenimi (5. ve 6. Hafta)

Veriyi öncelikle standart scaler ve one hot encoder gibi ön işleme teknikleriyle hazırladıktan sonra, eğitim ve test setlerine ayırdım. Daha sonra, RandomForest, XGBoost, Gradient Boosting, SVR, MLPRegressor ve KNN gibi farklı regresyon modellerini kullanarak eğitim gerçekleştirdim. Elde ettiğimiz sonuçları değerlendirmek için MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error) ve doğru tahmin oranlarını hesapladım. Bu değerlendirmeler sonucunda en iyi modelin %36.68'lik oranla MLPRegressor olduğunu belirledim.

### Flask Web Sitesi

En iyi modeli seçtikten sonra restoran gelir tahmini yapabilmek adına basit bir Flask uygulaması geliştirdim. Uygulamanın son haline [github repository'nden](#) ulaşabilirsiniz. Uygulamanın çalışırken ekran görüntüleri aşağıda verilmiştir.

**Restaurant Revenue Prediction**

Customers: 62

Menu Price: 42.69

Marketing: 9.65

Cuisine: American

Avg. Spending: 17.83

Resim 5: Restoran Gelir Tahmini Input Giriş

Marketing:

9,65

Cuisine:

American

Avg. Spending:

17,83

Promotions:

0

Reviews:

308

Predict

Resim 6: Restoran Gelir Tahmini Input Giriş Devamı

Predicted Monthly Revenue

The predicted monthly revenue is: **\$362.25**

[Go back to homepage](#)

Resim 7: Restoran Gelir Tahmini Sonuç