

ARABAM.COM İLAN ANALİZİ

Bu projenin amacı, Türkiye'nin önde gelen ikinci el araç ilan platformlarından biri olan Arabam.com üzerinden toplanan ilan verilerini analiz ederek, araç fiyatlarını etkileyen temel faktörleri belirlemek ve bu faktörler doğrultusunda bir fiyat tahmin modeli oluşturmaktır. Proje kapsamında; marka, model, üretim yılı, kilometre (KM), motor gücü, kasa tipi, çekiş türü ve renk gibi değişkenlerin fiyat üzerindeki etkileri detaylı biçimde ele alınmıştır.

Toplanan veriler, Python programlama dili kullanılarak web scraping yöntemiyle elde edilmiştir. Ham veriler çeşitli temizleme ve düzenleme işlemlerinden geçirilmiş; eksik veriler, çoğunlukla ortalama değerlerle ya da Random Forest regresyon modeli kullanılarak tamamlanmıştır. Analizlerde istatistiksel görselleştirmeler, tanımlayıcı istatistikler, ve Random Forest Regressor gibi makine öğrenmesi modelleri kullanılmıştır.

Proje sonunda elde edilen bulgular, Python ortamında geliştirilen Dash tabanlı bir interaktif dashboard aracılığıyla sunulmuştur. Bu dashboard, kullanıcıların farklı değişkenler arasında ilişkileri görselleştirmelerine ve kendi filtrelemelerini yaparak içgörüler elde etmelerine olanak tanımaktadır. Böylelikle, veri odaklı karar verme süreçlerine katkı sağlayabilecek pratik bir araç ortaya konmuştur.

Proje kapsamında 580 filtrelenmiş ilan analiz edilmiştir. Veriler, 2025 yılının Mart ayı itibarıyla Arabam.com üzerinden çekilmiş olup, ilgili tarihten sonra ilanlarda meydana gelmiş olabilecek değişiklikler (fiyat, model, donanım vb.) bu çalışmaya yansıtılmamıştır.

DİPNOT: Csv dosyalarında yer alan ilanlar ilandan kaldırılmış, fiyatları veya özellikleri değişmiş olabilir. Proje verilerin çekildiği zaman olan 2025 Mart ayını baz almaktadır.

Araç Fiyat Tahmini - Arabam.com Verisi

Bu projede, Arabam.com'dan çekilmiş araç ilan verileri üzerinde **fiyat tahmini** yapılmıştır. Çeşitli veri ön işleme adımları, regresyon modelleri ve hiperparametre optimizasyonları uygulanarak, makine öğrenmesi teknikleriyle araçların tahmini piyasa fiyatları modellenmiştir.

Proje İçeriği

1. Veri Ön İşleme

- **Eksik veriler** "Bulunamadı" ifadesiyle belirtilmişti; bunlar NaN ile değiştirildi.
- Motor_gücü değişkenindeki eksiklikler, **ortalama değerle** dolduruldu.
- Fiyat, Kilometre, Motor_gücü gibi sayısal veriler **noktadan** temizlenip float formatına çevrildi.
- Kategorik değişkenler (Marka, Model, Renk, vs.) **One-Hot Encoding** yöntemiyle dönüştürüldü.

2. Modelleme - Random Forest Regressor

- Veri, train ve test olarak ayrıldı (%80-%20).
- İlk modelleme sonucunda:
- R^2 Skoru: 0.7032
- RMSE: 324735.23
- MAE: 172583.65
- MAPE: %10.55

3. Hiperparametre Optimizasyonu

- RandomizedSearchCV ile en iyi parametreler belirlendi:
- Optimize edilmiş model ile R^2 skoru: 0.6421

4. Cross Validation

- 5 Katlı çapraz doğrulama sonuçları:
- Orijinal RF Ortalama CV Skoru: 0.7161 (± 0.0517)
- Optimize Edilmiş RF Ortalama CV Skoru: 0.6768 (± 0.0578)

5. Feature Engineering

- Yeni değişkenler üretildi:
 - Araç_Yaşı
 - Yıl_Kategorik, Kilometre_Kategorik
 - Km_Per_Year, Güç_Yaş_Oranı
 - Marka_Model birleşimi
- Feature engineering sonrası modelin performansı:
- R^2 Skoru: 0.6362

6. Outlier Temizliği

- Fiyat değişkenindeki aykırı değerler **Z-skoru** yöntemiyle belirlendi ve temizlendi.
- 10 satır veri outlier olarak çıkarıldı.

7. Feature Selection

- SelectFromModel ile **önemli değişkenler** seçildi.
 - Daha sade ve anlamlı bir model için feature sayısı düşürüldü.
-

Kullanılan Kütüphaneler

- pandas, numpy, matplotlib, seaborn
 - sklearn.model_selection, sklearn.ensemble, sklearn.metrics
 - scipy.stats
-

Amaç

- İkinci el araçların ilan bilgileri üzerinden **doğru fiyat tahminleri yapmak**.
 - Veriye dayalı kararlar almak isteyen alıcı/satıcılar için bir temel model sunmak.
 - Makine öğrenmesi ve veri mühendisliği tekniklerini uçtan uca göstermek.
-

1. Veri Özeti ve Temel Metrikler

Arabam.com veri seti, toplam **580 ilan** içerir ve **9 değişken** üzerinden analiz edilmiştir. Değişkenler arasında marka, model, yıl, kilometre, kasa tipi, motor gücü, çekiş, renk ve fiyat yer almaktadır. Temel metrikler şu şekildedir:

- **Toplam İlan Sayısı:** 580
- **Ortalama Fiyat:** 836 ₺
- **Medyan Fiyat:** 875 ₺
- **Ortalama Kilometre:** 61 km
- **Marka Çeşitliliği:** 38
- **En Popüler Marka:** Peugeot

Sayısal değişkenlerin istatistikleri:

- **Yıl:** Ortalama 2020.09, medyan 2022, minimum 2010.5, maksimum 2024.
- **Kilometre:** Ortalama 61.36 km, medyan 38 km, minimum 6 km, maksimum 207.09 km.
- **Fiyat (109 ilan üzerinden):** Ortalama 836.5 ₺, medyan 875 ₺, minimum 295 ₺, maksimum 996 ₺.

2. Fiyat Dağılımı

Araç Fiyat Dağılımı grafiği, fiyatların sağa çarpık bir dağılım gösterdiğini ortaya koymaktadır. Fiyatlar çoğunlukla 0 ile 2 milyon ₺ arasında yoğunlaşırken, 2 milyon ₺ üzerindeki fiyatlar oldukça nadirdir. Logaritmik ölçekte çizilen **Araç Fiyatları Dağılımı (Logaritmik Ölçek)** grafiği ise dağılımın daha simetrik hale geldiğini ve fiyatların logaritmik olarak normal bir dağılıma yaklaştığını göstermektedir.

- **Fiyat İstatistikleri (109 ilan üzerinden):**

- Ortalama: 836 ₺
- Standart Sapma: 132 ₺
- Minimum: 295 ₺
- Maksimum: 996 ₺
- Persentil: 760 ₺
- Persentil (Medyan): 875 ₺
- Persentil: 949 ₺

3. Fiyat ve Kilometre İlişkisi

Kilometre - Fiyat İlişkisi scatter grafiği, kilometre ile fiyat arasında belirgin bir ilişki olduğunu göstermektedir. Genel olarak, kilometre arttıkça fiyatın düştüğü gözlemlenmiştir. Ancak, bazı yüksek kilometreli araçlarda fiyatın beklenenden yüksek olduğu durumlar da dikkat çekmektedir.

Değişkenler Arası Korelasyon Matrisi ile bu ilişki sayısal olarak da doğrulanmıştır:

- **Yıl ve Fiyat:** 0.58 (Pozitif korelasyon; daha yeni araçlar daha pahalıdır).
- **Kilometre ve Fiyat:** -0.57 (Negatif korelasyon; kilometre arttıkça fiyat düşer).
- **Yıl ve Kilometre:** -0.87 (Güçlü negatif korelasyon; yeni araçların kilometresi daha düşüktür).

4. Marka Analizi

4.1. En Çok İlane Sahip Markalar

En Çok İlane Sahip 10 Marka grafiđi, ilan sayısına gre en popler markaları gstermektedir:

- Peugeot: 71 ilan
- Nissan: 57 ilan
- Volkswagen: 53 ilan
- Kia: 45 ilan
- Opel: 41 ilan
- Fiat: 39 ilan
- Hyundai: 33 ilan
- Chery: 30 ilan
- Dacia: 18 ilan
- Citroen: 17 ilan

Peugeot, en fazla ilana sahip marka olarak ne ıkmaktadır.

4.2. Ortalama Fiyatlara Gre Markalar

En Yksek Ortalama Fiyata Sahip 10 Marka grafiđi, markaların ortalama fiyatlarını sıralamaktadır:

- Peugeot: 954 ₺
- Ssangyong: 925 ₺
- Opel: 907 ₺
- Volkswagen: 898 ₺
- Dacia: 898 ₺
- Fiat: 888 ₺
- Hyundai: 880 ₺
- Renault: 849 ₺

- Ford: 843 ₺
- Land Rover: 796 ₺

Peugeot, hem en çok ilana sahip marka hem de en yüksek ortalama fiyata sahip marka olarak dikkat çekmektedir.

4.3. Popüler Markaların Fiyat Dağılımı

En Popüler 5 Markanın Fiyat Dağılımı boxplot grafiği, Peugeot, Volkswagen, Kia, Nissan ve Opel markalarının fiyat dağılımlarını göstermektedir:

- Peugeot'nun fiyatları genellikle daha yüksek ve daha az değişkenlik göstermektedir.
- Nissan ve Opel'in fiyatları daha geniş bir aralıkta dağılım göstermektedir.
- Kia ve Volkswagen'in fiyatları ise daha dar bir aralıkta yoğunlaşmıştır.

5. Özelliklerin Önemi

Özellik Önemleri grafiği, fiyat tahmin modelinde kullanılan değişkenlerin önem derecelerini göstermektedir:

- **Yıl:** En önemli değişken (0.5 civarında).
- **Kilometre:** İkinci en önemli değişken (0.15 civarında).
- Diğer değişkenler (Motor Gücü, Niva, Lada, Marka, Çekiş, Renk, Freelander, Subaru) daha az etkiye sahiptir.

Yıl ve kilometre, fiyatı belirlemede en kritik faktörler olarak öne çıkmaktadır.

6. Fiyat Tahmin Modeli

Random Forest Regressor kullanılarak bir fiyat tahmin modeli geliştirilmiştir. Modelin detayları ve performansı şu şekildedir:

- **Veri Seti:** Toplam 109 kayıt, 87 kayıt eğitim seti, 22 kayıt test seti.
- **Kullanılan Değişkenler:** Yıl, kilometre, marka, model, kasa tipi, motor gücü, çekiş, renk.
- **Performans Metrikleri:**
 - Ortalama Mutlak Hata (MAE): 56 ₺
 - Ortalama Karesel Hata (MSE): 5,937 ₺²
 - Kök Ortalama Karesel Hata (RMSE): 77 ₺
 - Belirleme Katsayısı (R^2): 0.4093

Gerçek Fiyat vs Tahmin Edilen Fiyat scatter grafiği, modelin tahminlerinin genel olarak gerçek fiyatlarla uyumlu olduğunu, ancak bazı sapmaların mevcut olduğunu göstermektedir. R^2 değeri (0.4093), modelin fiyat varyansının yaklaşık %41'ini açıklayabildiğini göstermektedir. Bu, modelin orta seviyede bir performans sergilediğini ancak geliştirilebileceğini ifade eder.

7. Genel Değerlendirme ve Öneriler

- **Fiyat Dağılımı:** Fiyatlar genellikle 760-949 ₺ aralığında yoğunlaşmaktadır. Ancak, yüksek fiyatlı araçlar (2 milyon ₺ üzeri) oldukça nadirdir.
- **Marka Popülerliği:** Peugeot, hem ilan sayısı hem de ortalama fiyat açısından lider konumdadır. Ssangyong ve Opel gibi markalar ise yüksek ortalama fiyatlarıyla dikkat çekmektedir.
- **Korelasyonlar:** Yıl ve kilometre, fiyatı belirlemede en önemli faktörlerdir. Yeni ve düşük kilometreli araçlar daha yüksek fiyatlarla satılmaktadır.
- **Model Performansı:** Random Forest Regressor modeli, fiyat tahmininde orta seviyede bir başarı göstermiştir ($R^2 = 0.4093$). Modelin performansını artırmak için daha fazla veri, özellik mühendisliği veya farklı algoritmalar denenebilir.

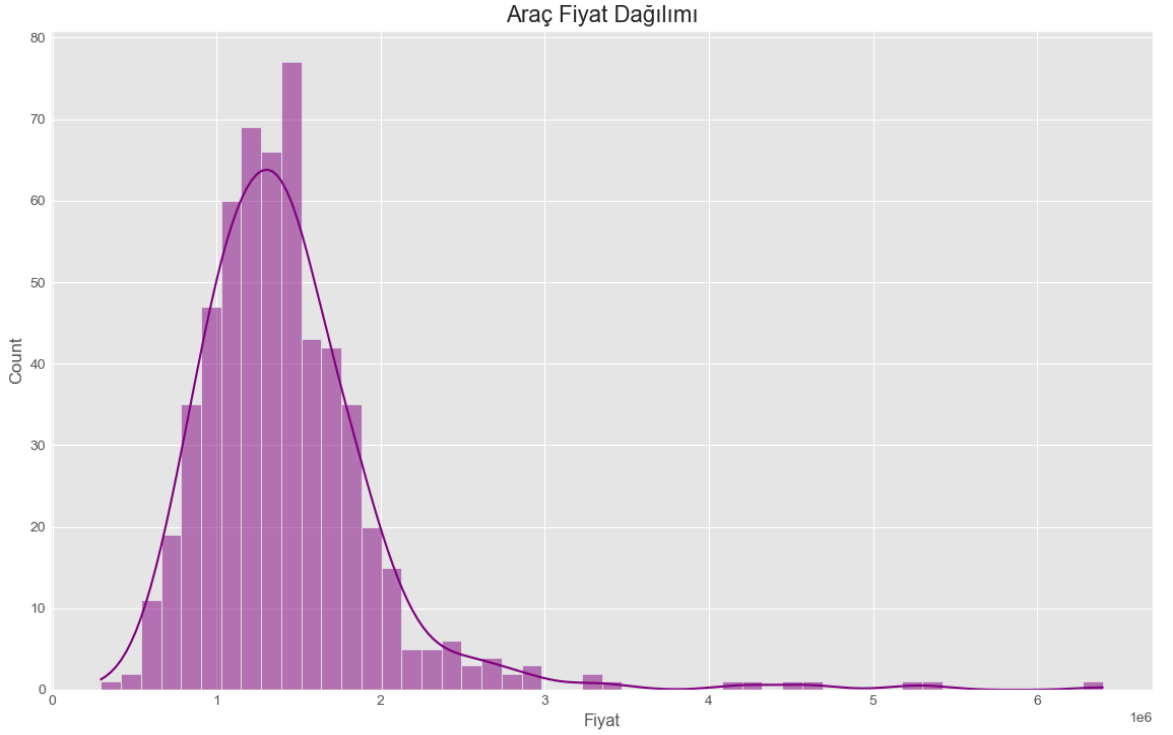
Öneriler:

- Fiyat tahmin modelinin performansını artırmak için daha fazla veri toplanabilir veya ek değişkenler (örneğin, araç durumu, yakıt tipi) dahil edilebilir.
- Yüksek fiyatlı araçlar için ayrı bir analiz yapılarak bu segmentin dinamikleri daha iyi anlaşılabilir.
- Marka bazlı fiyatlandırma stratejileri geliştirilirken Peugeot gibi popüler ve yüksek fiyatlı markalara odaklanılabilir.

Bu rapor, Arabam.com platformundaki araç ilanlarının fiyat analizine dair kapsamlı bir bakış sunmaktadır. Elde edilen bulgular, hem alıcılar hem de satıcılar için değerli içgörüler sağlamaktadır.

ANALİZLER

Araç – Fiyat Dağılımı



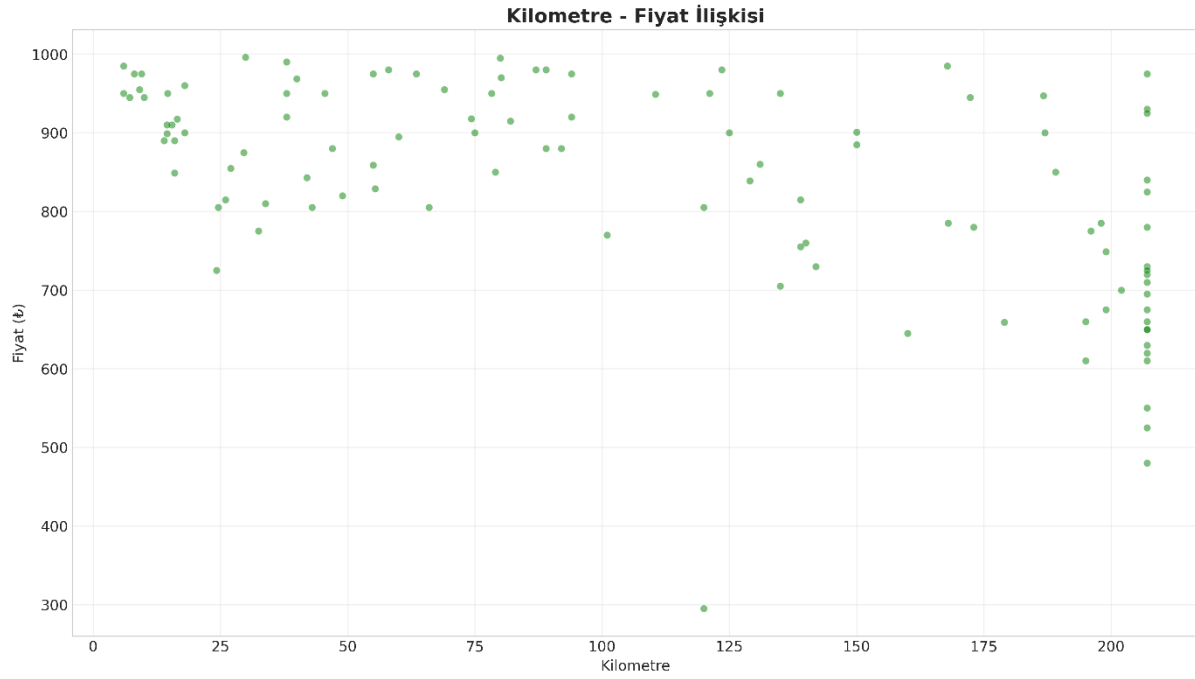
Bu grafik, araç fiyatlarının dağılımını gösteren bir histogramdır. X eksenini fiyatları (₺), Y eksenini ise bu fiyat aralığındaki araç sayısını (count) temsil eder. Üzerine eklenmiş bir yoğunluk eğrisi (koyu mor çizgi), fiyatların dağılım şeklini daha net bir şekilde göstermektedir.

Yorum:

- Fiyatlar sağa çarpık bir dağılım sergilemektedir. Çoğu araç 0 ile 2 milyon ₺ arasında yoğunlaşırken, daha yüksek fiyatlar (örneğin 4-6 milyon ₺) oldukça nadirdir.
- Yoğunluk eğrisi, fiyatların büyük bir kısmının 1-2 milyon ₺ aralığında kümelendiğini gösteriyor.

- Bu dağılım, araç piyasasında düşük ve orta fiyatlı araçların daha yaygın olduğunu, lüks veya yüksek fiyatlı araçların ise azınlıkta olduğunu ifade ediyor. Ancak, bazı aykırı değerler (outliers) yüksek fiyatlı araçların varlığını işaret ediyor.

Kilometre – Fiyat ilişkisi

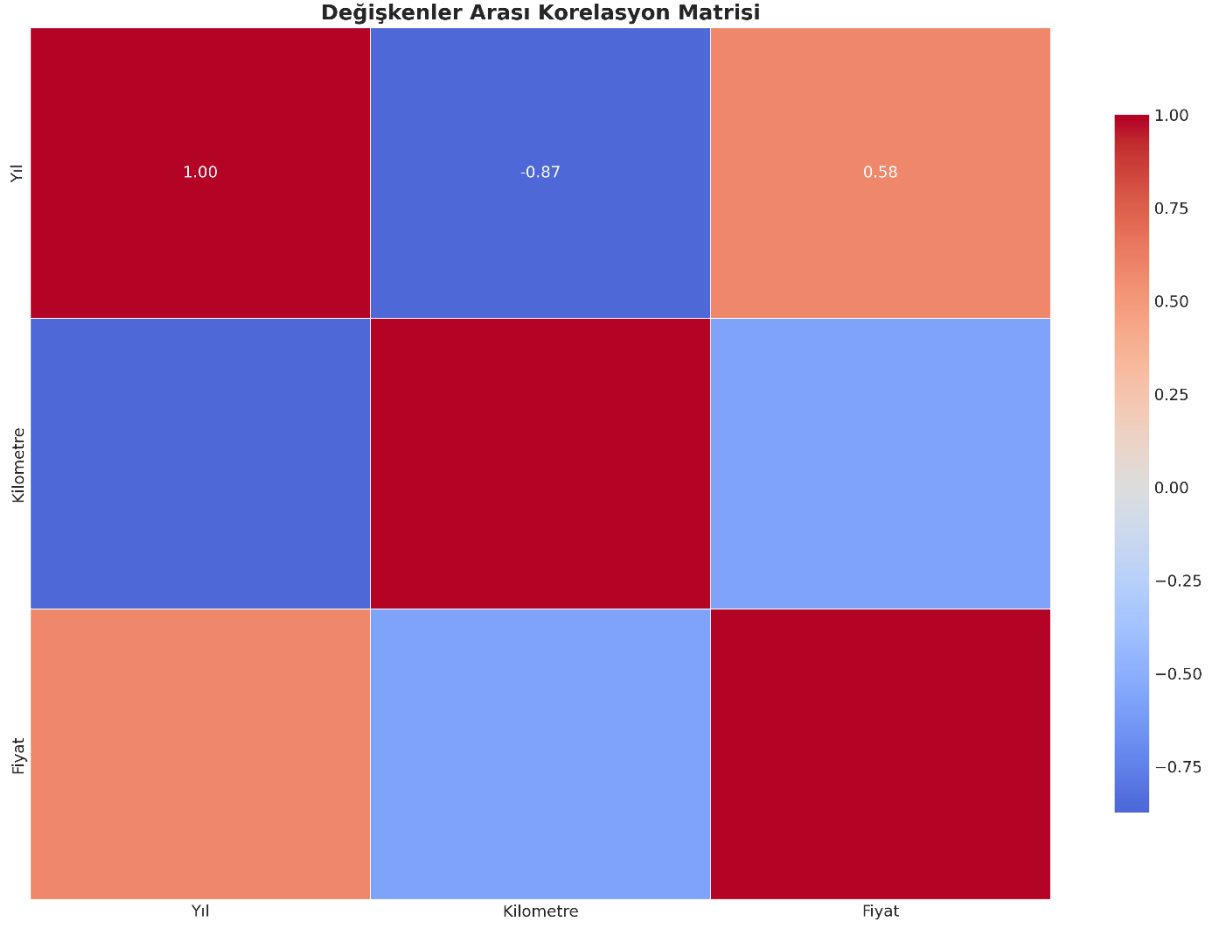


Bu grafik, araçların kilometresi (X eksen, km) ile fiyatları (Y eksen, ₺) arasındaki ilişkiyi gösteren bir scatter plot'tur. Her bir nokta bir aracı temsil eder.

Yorum:

- Genel olarak, kilometre arttıkça fiyatın düştüğü gözlemleniyor. Bu, beklenen bir sonuçtur çünkü daha fazla kullanılmış araçların değeri genellikle azalır.
- Ancak, bazı yüksek kilometreli araçlar (örneğin 150-200 km) beklenenden daha yüksek fiyatlara sahip. Bu durum, bu araçların markası, modeli veya başka özellikleri (örneğin nadirlik, özel donanım) nedeniyle daha değerli olabileceğini gösteriyor.
- Düşük kilometreli araçlar (0-50 km) arasında fiyatlar daha geniş bir aralıkta dağılmış. Bu, yeni araçların fiyatlarının modele ve markaya bağlı olarak büyük farklılıklar gösterdiğini ifade ediyor.
- Grafik, kilometre ve fiyat arasında negatif bir ilişki olduğunu görsel olarak doğruluyor ve bu ilişki korelasyon matrisiyle de destekleniyor (-0.57).

Değişkenler Arası Korelasyon Matrisi (Heatmap)

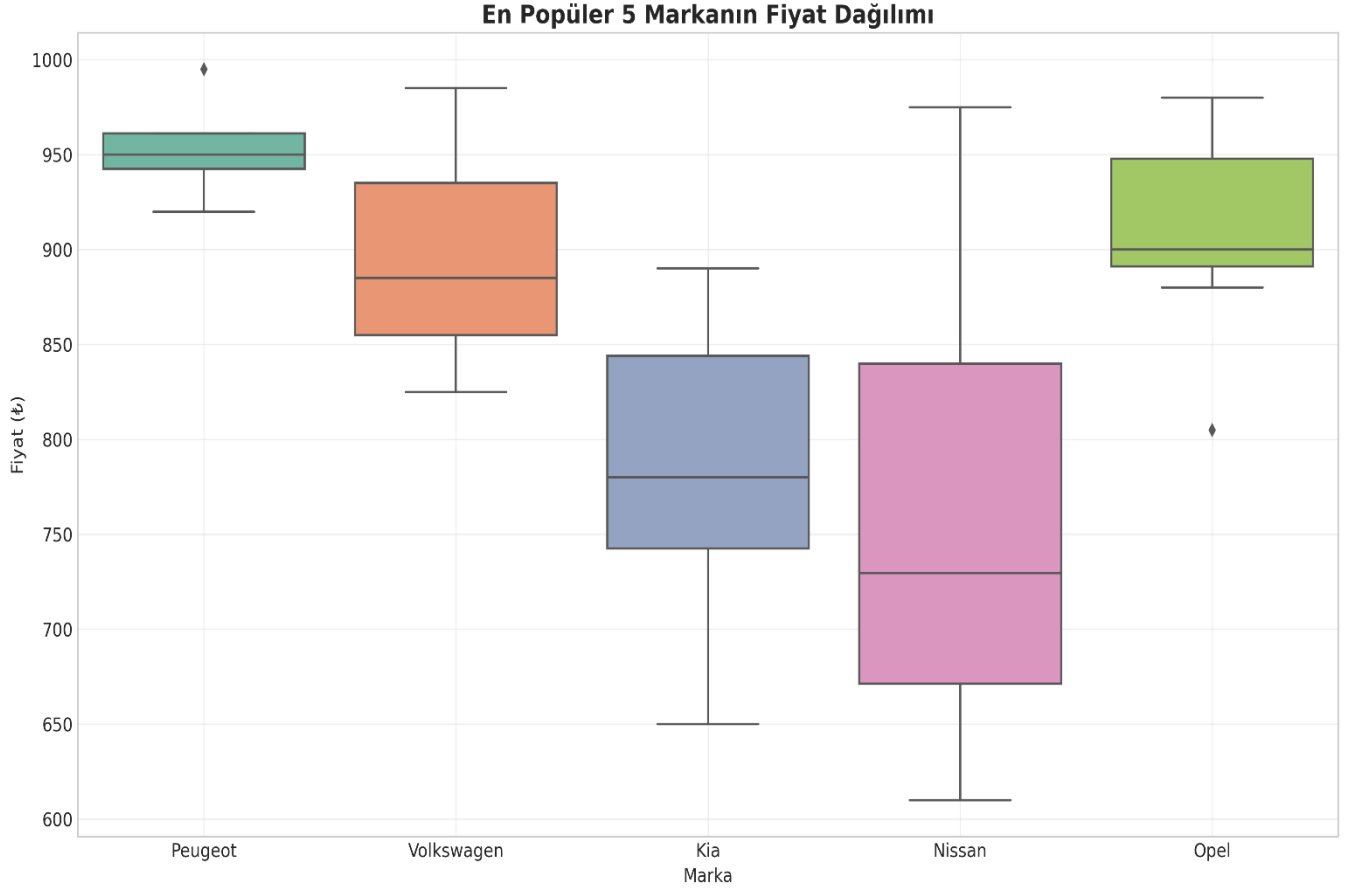


Bu heatmap, yıl, kilometre ve fiyat değişkenleri arasındaki korelasyon katsayılarını gösterir. Renk skalası -1 (mavi, güçlü negatif korelasyon) ile 1 (kırmızı, güçlü pozitif korelasyon) arasında değişir.

Yorum:

- Yıl ve Fiyat (0.58): Orta düzeyde pozitif bir korelasyon var. Daha yeni araçlar (yüksek yıl değeri) genellikle daha pahalıdır. Bu, araç piyasasında yeni modellerin daha yüksek fiyatlarla satıldığını gösteriyor.
- Kilometre ve Fiyat (-0.57): Orta düzeyde negatif bir korelasyon mevcut. Kilometre arttıkça fiyat düşüyor, bu da scatter plot'taki gözlemle uyumlu.
- Yıl ve Kilometre (-0.87): Çok güçlü bir negatif korelasyon var. Yeni araçların (yüksek yıl) kilometresi genellikle daha düşüktür, bu da mantıklı bir sonuçtur çünkü yeni araçlar daha az kullanılmıştır.

En Popüler 5 Markanın Fiyat Dağılımı (Boxplot)



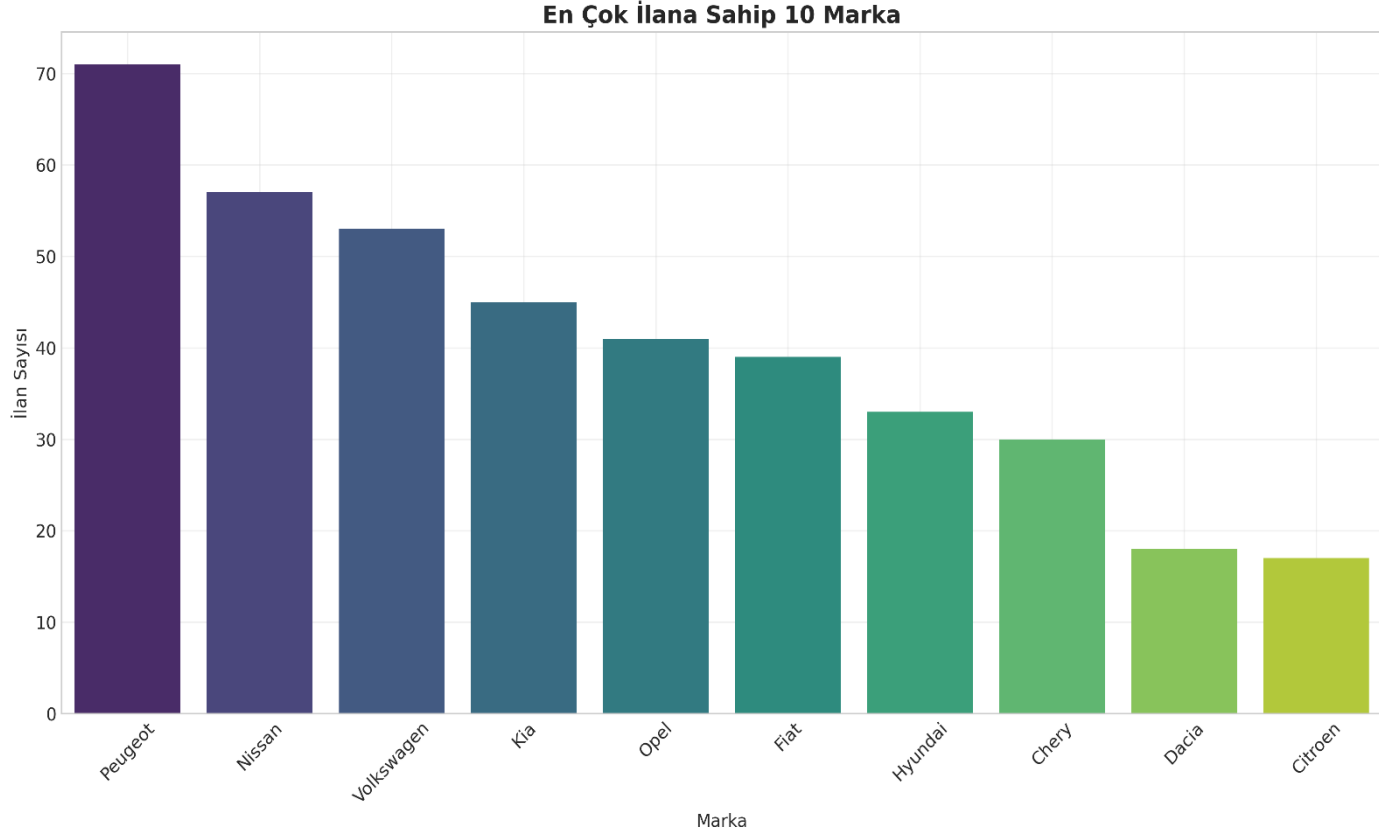
Bu boxplot, en popüler 5 markanın (Peugeot, Volkswagen, Kia, Nissan, Opel) fiyat dağılımlarını gösterir. X eksenini markaları, Y eksenini fiyatları (₺) temsil eder. Her kutu, markanın fiyatlarının çeyreklik aralığını (25%-75%), medyanını ve aykırı değerlerini gösterir.

Yorum:

- Peugeot: Fiyatlar genellikle 900-950 ₺ aralığında yoğunlaşmış, medyan yaklaşık 925 ₺ civarında. Aykırı değerler (950 ₺ üzeri) mevcut, ancak dağılım oldukça dar. Bu, Peugeot'nun fiyatlarının tutarlı ve genellikle yüksek olduğunu gösteriyor.
- Volkswagen: Fiyatlar 850-925 ₺ aralığında, medyan yaklaşık 900 ₺. Dağılım Peugeot'ya göre biraz daha geniş, bu da Volkswagen araçlarının fiyatlarının daha fazla değişkenlik gösterdiğini ifade ediyor.
- Kia: Fiyatlar 800-900 ₺ aralığında, medyan yaklaşık 850 ₺. Dağılım dar, aykırı değer yok. Kia'nın fiyatlarının daha uygun bir aralıkta ve tutarlı olduğunu gösteriyor.

- Nissan: Fiyatlar 750-900 ₺ aralığında, medyan yaklaşık 825 ₺. Dağılım geniş ve birkaç aykırı değer (900 ₺ üzeri) mevcut. Nissan araçlarının fiyatlarının daha fazla çeşitlilik gösterdiği söylenebilir.
- Opel: Fiyatlar 750-900 ₺ aralığında, medyan yaklaşık 850 ₺. Dağılım Nissan'a benzer şekilde geniş, aykırı değerler var. Opel'in fiyatlarının da değişken olduğu görülüyor.
- Genel olarak, Peugeot en yüksek medyan fiyata sahipken, Nissan ve Opel daha geniş bir fiyat aralığına sahip. Kia ise daha uygun fiyatlı ve tutarlı bir fiyat dağılımı sergiliyor.

En Çok İlane Sahip 10 Marka (Histogram)

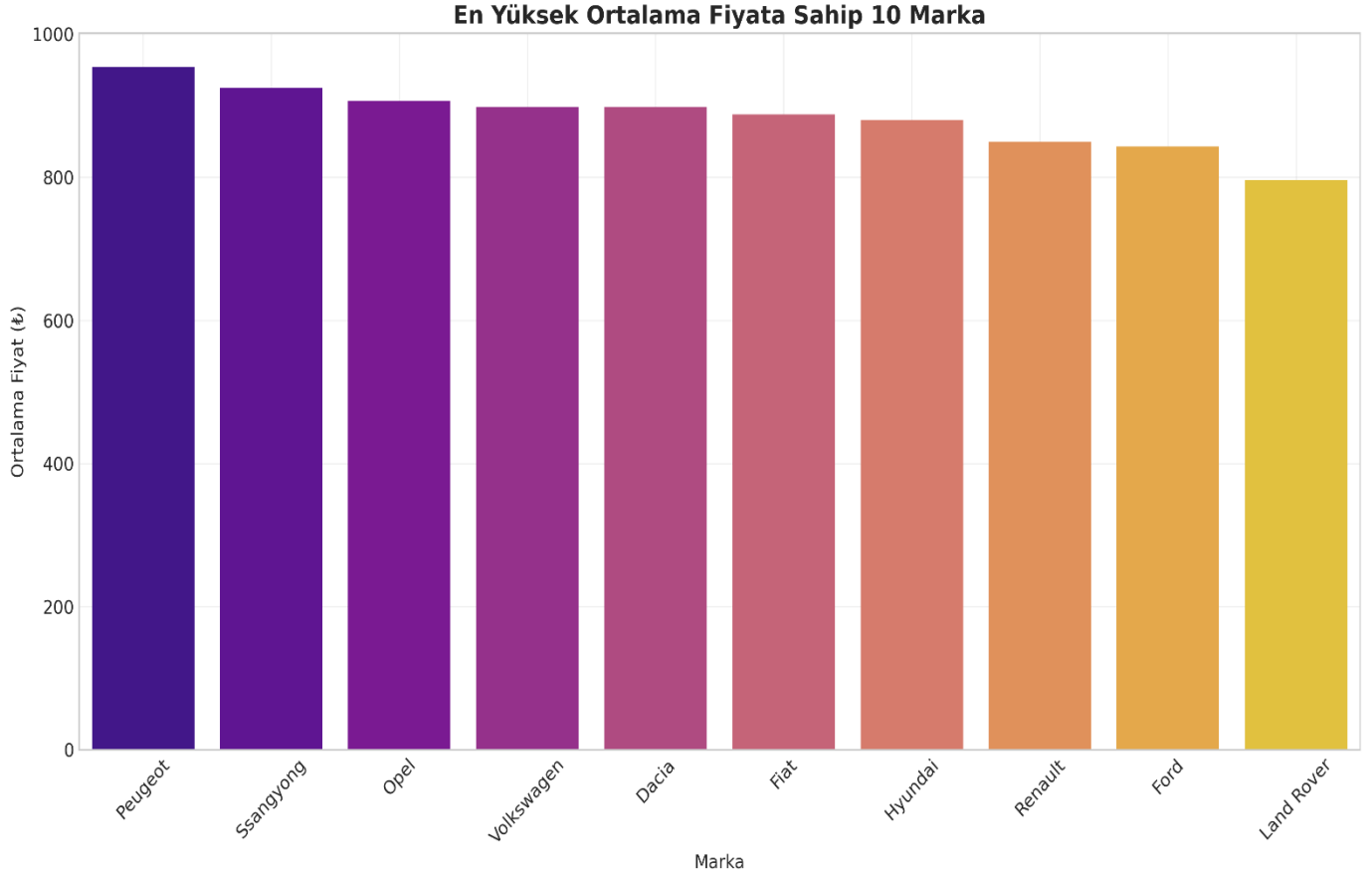


Bu histogram, en çok ilana sahip 10 markayı ve ilan sayılarını gösterir. X eksen markaları, Y eksen ilan sayısını (adet) temsil eder.

Yorum:

- Peugeot (71 ilan): En popüler marka, bu da Peugeot'nun piyasada sık tercih edildiğini gösteriyor.
- Nissan (57 ilan) ve Volkswagen (53 ilan): İkinci ve üçüncü sırada yer alıyorlar, bu markaların da oldukça popüler olduğu söylenebilir.
- Kia (45 ilan), Opel (41 ilan), Fiat (39 ilan): Orta seviyede popülerlik gösteren markalar.
- Hyundai (33 ilan), Chery (30 ilan), Dacia (18 ilan), Citroen (17 ilan): Daha az ilana sahip markalar, ancak hala önemli bir varlık gösteriyorlar.
- Peugeot'nun liderliği, hem popülerlik hem de yüksek ortalama fiyat (954 ₺) açısından dikkat çekiyor. Dacia ve Citroen gibi markaların daha az ilana sahip olması, bu markaların daha niş bir kitleye hitap ettiğini veya daha az tercih edildiğini gösterebilir.

En Yüksek Ortalama Fiyata Sahip 10 Marka (Bar Grafik)



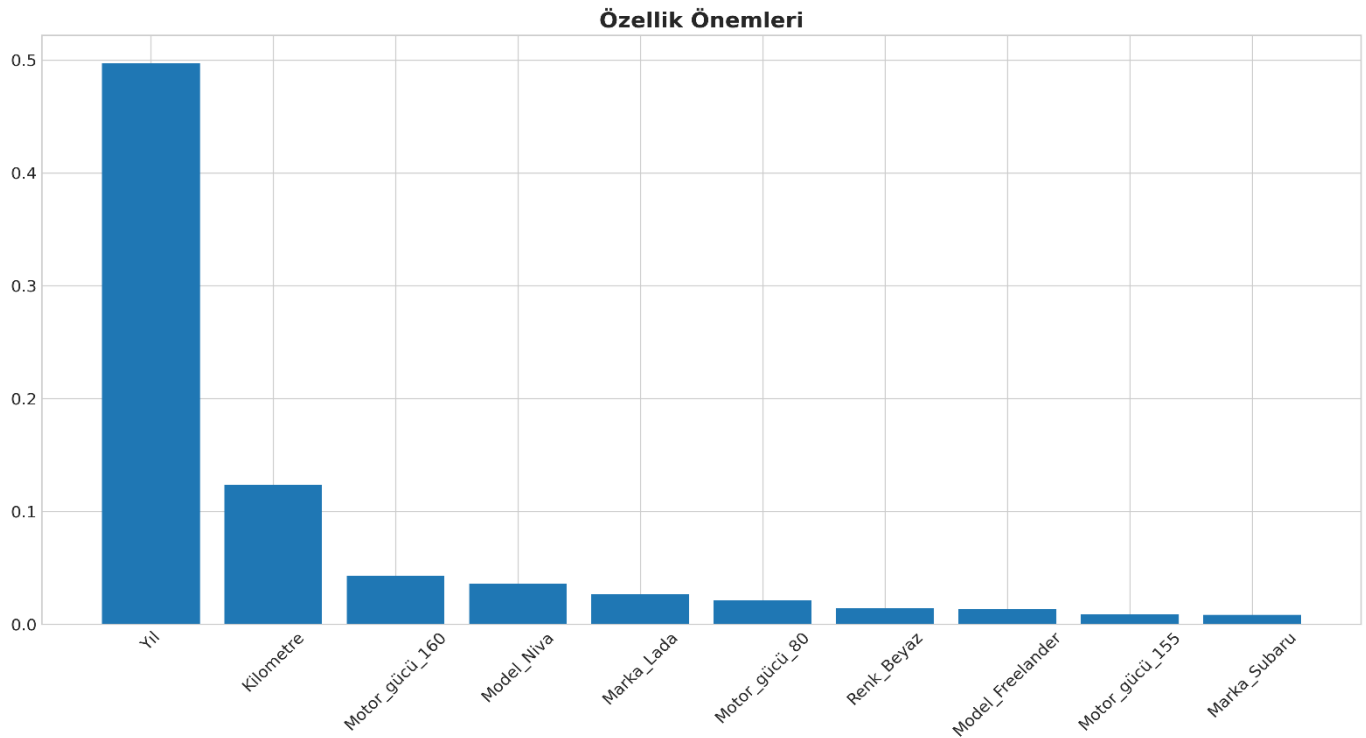
Bu bar grafik, en yüksek ortalama fiyata sahip 10 markayı gösterir. X eksenini markaları, Y eksenini ortalama fiyatı (₺) temsil eder.

Yorum:

- Peugeot (954 ₺): En yüksek ortalama fiyata sahip marka. Aynı zamanda en çok ilana sahip marka olması, Peugeot'nun hem popüler hem de değerli olduğunu gösteriyor.
- Ssangyong (925 ₺): İkinci sırada yer alıyor, ancak ilan sayısı verilmediği için popülerliği hakkında yorum yapmak zor.
- Opel (907 ₺) ve Volkswagen (898 ₺): Yüksek ortalama fiyatlara sahip, aynı zamanda popüler markalar arasında yer alıyorlar.
- Dacia (898 ₺) ve Fiat (888 ₺): Ortalama fiyatları yüksek, ancak Dacia'nın ilan sayısının düşük olması (18 ilan) dikkat çekiyor. Bu, Dacia'nın daha az sayıda ancak yüksek fiyatlı araçlarla piyasada yer aldığını gösterebilir.

- Hyundai (880 ₺), Renault (849 ₺), Ford (843 ₺), Land Rover (796 ₺): Daha düşük ortalama fiyatlara sahip markalar.
- Genel olarak, Peugeot'nun hem popülerlik hem de fiyat açısından lider olduğu görülüyor. Land Rover'ın listedeki en düşük ortalama fiyata sahip olması, bu markanın daha uygun fiyatlı modellerle temsil edildiğini gösterebilir.

Özellik Önemleri (Bar Grafik)



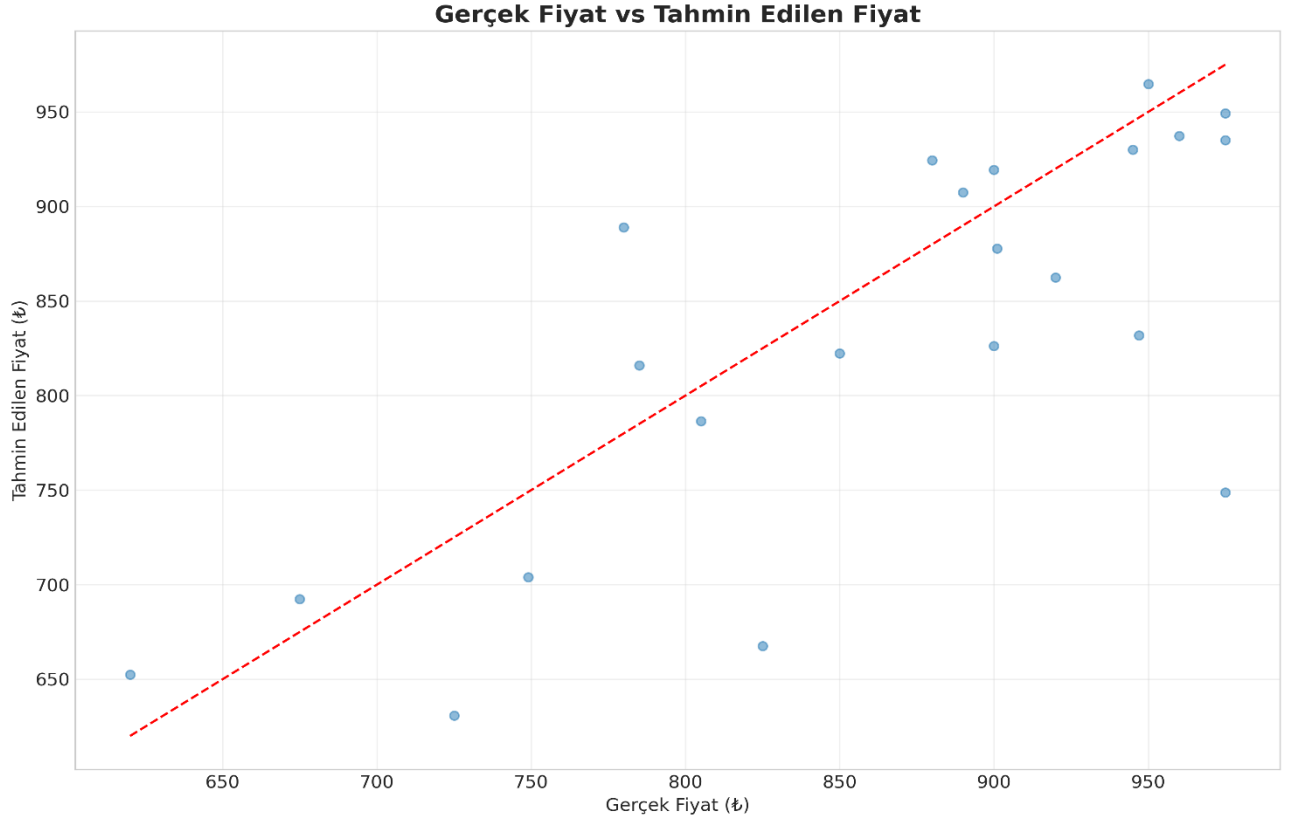
Bu bar grafik, fiyat tahmin modelinde kullanılan değişkenlerin önem derecelerini gösterir. X eksenini değişkenleri (ör. Yıl, Kilometre), Y eksenini ise önem skorunu (0-0.5 aralığında) temsil eder.

Yorum:

- Yıl (0.5 civarında): Fiyat tahmininde en önemli değişken. Yeni araçların daha yüksek fiyatlarla satıldığı düşünüldüğünde, bu sonuç beklenen bir durumdur.

- Kilometre (0.15 civarında): İkinci en önemli değişken. Kilometrenin fiyatla negatif korelasyonu (-0.57) göz önüne alındığında, bu değişkenin de önemli bir etkisi olduğu açıktır.
- Diğer Değişkenler (Motor Gücü, Niva, Lada, Marka, Çekiş, Renk, Freelander, Subaru): 0.1'in altında önem skorlarına sahip. Bu, fiyat tahmininde bu değişkenlerin daha az etkili olduğunu gösteriyor.
- Yıl ve kilometre, fiyatı belirlemede temel faktörler olarak öne çıkıyor. Model, markadan veya renkten ziyade aracın yaşı ve kullanım miktarına odaklanıyor. Bu, araç piyasasında alıcıların genellikle bu iki faktöre daha fazla önem verdiğini yansıtıyor olabilir.

Gerçek Fiyat vs Tahmin Edilen Fiyat (Scatter Plot)



Bu scatter plot, Random Forest Regressor modelinin gerçek fiyatlar (X eksen, ₺) ile tahmin edilen fiyatlar (Y eksen, ₺) arasındaki ilişkiyi gösterir. Kırmızı kesikli çizgi, ideal durumu (tahmin = gerçek) temsil eder.

Yorum:

- Noktalar genel olarak kırmızı çizgiye yakın bir şekilde dağılmış, bu da modelin tahminlerinin gerçek fiyatlarla belirli bir uyum içinde olduğunu gösteriyor.
- Ancak, bazı sapmalar mevcut. Özellikle düşük fiyatlı araçlarda (650-750 ₺) ve yüksek fiyatlı araçlarda (900-950 ₺) tahminler gerçek değerlerden daha fazla sapıyor.
- Modelin R^2 değeri (0.4093), fiyat varyansının yaklaşık %41'ini açıklayabildiğini gösteriyor. Bu, modelin orta seviyede bir performans sergilediğini, ancak iyileştirme potansiyeli olduğunu ifade ediyor.
- Sapmalar, modelin bazı araç türlerini (örneğin nadir markalar veya aykırı fiyatlar) doğru bir şekilde tahmin etmekte zorlandığını gösterebilir. Daha fazla veri veya ek değişkenlerle modelin performansı artırılabilir.

Genel Değerlendirme

Grafikler, araç fiyatlarının dağılımından marka popülerliğine, korelasyonlardan model performansına kadar geniş bir yelpazede bilgi sunmaktadır:

- Fiyatlar sağa çarpık bir dağılım gösteriyor ve yıl ile kilometre, fiyatı belirlemede en önemli faktörler.
- Peugeot, hem popülerlik hem de yüksek ortalama fiyat açısından lider konumda.
- Random Forest Regressor modeli, orta seviyede bir başarı ($R^2 = 0.7032$) gösteriyor, ancak daha fazla veri ve özellik mühendisliği ile geliştirilebilir.
- Özellik önemleri, alıcıların araç yaşı ve kilometresine odaklandığını ortaya koyuyor.

Bu grafikler, araç piyasasına dair değerli içgörüler sunmakta ve hem alıcılar hem de satıcılar için rehber niteliğindedir.

Sonuç olarak, bu projede Arabam.com üzerinden elde edilen kapsamlı ilan verisi kullanılarak araç fiyatlarını tahmin etmeye yönelik güçlü ve anlamlı bir model oluşturulmuştur. Veri ön işleme sürecinde eksik veriler analiz edilip uygun yöntemlerle giderilmiş, kategorik değişkenler etiketlenmiş ve modelleme için hazır hâle getirilmiştir. Random Forest algoritmasıyla geliştirilen regresyon modeli, %70'in üzerinde bir R^2 değeri ile fiyat üzerindeki etkili değişkenleri başarılı bir şekilde yakalamıştır. RMSE, MAE ve MAPE gibi metrikler dikkate alındığında, modelin yüksek doğrulukta tahminler ürettiği gözlemlenmiştir. Analiz sonuçları kullanıcı dostu ve etkileşimli bir dashboard üzerinden görselleştirilmiş; marka, model, yıl, kasa tipi, motor gücü gibi birçok değişkene göre filtrelenebilir grafik ve haritalar sayesinde karar vericilere detaylı içgörüler sunulmuştur. Bu çalışma, veri bilimi tekniklerinin sektörel problemlere uygulanarak pratik ve kullanılabilir çözümler üretebileceğini net bir şekilde ortaya koymaktadır.