# 

**İzmir Bakırçay Üniversitesi**

**Veri Analitiğinin Temelleri**

**Yüksek Lisans Çalışması**

**Emine Kocabay**

**Nisan - 2025**

# 1. Giriş

Bu çalışma, Amerika Birleşik Devletleri’nde 2015–2018 yılları arasında satılan avokado ürünlerine ilişkin fiyat ve satış hacmi verilerini analiz ederek, ürün türü (organik/konvansiyonel), bölge ve zaman faktörlerine göre değişim gösteren pazar eğilimlerini ortaya koymayı amaçlamaktadır. Son yıllarda sağlıklı beslenme trendleri doğrultusunda popülaritesi artan avokadonun pazardaki davranışı, tüketici tercihleri ve fiyat oluşum süreçlerine dair anlamlı ipuçları sunmaktadır.  
  
Bu kapsamda çalışma, R programlama dili kullanılarak gerçekleştirilmiş ve farklı analiz teknikleriyle desteklenmiştir. Görsel öğelerle desteklenen bu analizler, avokado pazarındaki fiyat dalgalanmaları, satış hacmi eğilimleri ve bölgesel farklılıklar gibi önemli verileri yorumlamaya olanak sağlamaktadır.

## 1.1 Çalışmanın Amacı

Bu araştırmanın temel amacı, avokado fiyatlarının zaman içerisindeki değişimini, ürün türüne (organik/konvansiyonel) ve satış bölgelerine göre analiz etmektir. Ayrıca mevsimsel etkilerin fiyatlara yansıması da değerlendirilmektedir. Elde edilen bulgular, tedarik zinciri yönetimi, pazarlama stratejileri ve tüketici davranışları açısından çeşitli çıkarımlar yapılmasına olanak sağlayacaktır.

## 1.2 Kapsam ve Sınırlılıklar

Çalışma, yalnızca 2015–2018 yılları arasında Amerika’da gerçekleşen avokado satış verilerini kapsamaktadır. Türkiye verisi içermemektedir. Kullanılan veri seti Kaggle üzerinden sağlanmış, ancak bu proje kapsamında tüm analizler özgün olarak R dili ile gerçekleştirilmiştir. Veri setinde yalnızca perakende satışlar yer almakta olup, toptan satış veya ihracat verileri analize dahil edilmemiştir.

## 1.3 Yöntem

Araştırmada nicel analiz yöntemi benimsenmiştir. Tanımlayıcı istatistiklerin yanı sıra grafiksel gösterimler ve zaman serisi analizleri ile veri yorumlanmıştır. Grafikler aracılığıyla bölgesel fiyat dağılımları, yıllara göre fiyat eğilimleri ve ürün türlerine göre karşılaştırmalar sunulmuştur. Tüm analizler R programlama dili ile yapılmış ve ileri düzey veri görselleştirme tekniklerinden yararlanılmıştır.

1. **Veri Seti ve Yöntem**

Bu bölümde çalışmada kullanılan veri seti tanıtılmakta ve analizde izlenen yöntem açıklanmaktadır. Veri ön işleme, değişken seçimi ve analiz süreci detaylandırılmıştır.

**2.1 Veri Setinin Özellikleri**

Araştırmada kullanılan veri seti, 2015–2018 yılları arasında Amerika Birleşik Devletleri’nde satılan avokadolarla ilgili satış verilerini kapsamaktadır. Veri seti 18.249 satır ve 14 sütundan oluşmaktadır. Her bir satır belirli bir hafta, bölge ve ürün tipi (organik/konvansiyonel) için toplam satış hacmini ve ortalama fiyatı içermektedir. Veri setinde kullanılan temel sütunlar şu şekildedir:

* **Date**: Satışın gerçekleştiği hafta
* **AveragePrice**: O haftadaki ortalama satış fiyatı (ABD doları cinsinden)
* **Total Volume**: Satılan toplam avokado miktarı
* **Small, Large, XLarge**: Satılan avokado boyutlarına göre miktarlar
* **Type**: Ürünün organik mi konvansiyonel mi olduğunu belirtir
* **Region**: Satışın yapıldığı bölge
* **Year**: Satışın gerçekleştiği yıl

"Small Bags, Large Bags, XLarge Bags" gibi ambalajlara yönelik sütunlar analiz öncesi çıkarılmıştır. Ayrıca 4046, 4225, 4770 olarak adlandırılmış orijinal sütunlar, daha anlaşılır olması adına sırasıyla Small, Large ve XLarge olarak yeniden adlandırılmıştır.

**2.2 Veri Temizleme ve Ön İşleme**

Analize başlamadan önce veri seti detaylı bir şekilde gözden geçirilmiş ve ön işleme adımları uygulanmıştır. R programlama dili kullanılarak yapılan analizde dplyr, tidyr, lubridate gibi paketlerden faydalanılmıştır. İlk olarak tarih formatı dönüştürülmüş, yıl ve ay gibi zaman bileşenleri ayrı sütunlara ayrılmıştır. Bu adım, zaman serisi analizlerini kolaylaştırmak amacıyla yapılmıştır.

Veri setinde yer alan ancak analiz açısından anlamlı katkı sağlamayan ambalaj türlerine ilişkin sütunlar (Small Bags, Large Bags, XLarge Bags) veri setinden çıkarılmıştır. Ayrıca 4046, 4225, 4770 etiketleri ile anılan ürün türleri daha açıklayıcı hale getirilmiş ve Small, Large, XLarge olarak yeniden adlandırılmıştır. Bu sayede veri seti hem okunabilirlik hem de analiz kolaylığı açısından optimize edilmiştir.

Veri temizleme sürecinde eksik veya aykırı veri gözlemlenmemiştir. Bu nedenle doğrudan analiz aşamasına geçilebilmiştir. Bununla birlikte, bazı kategorik değişkenlerde (örneğin, ürün tipi veya bölge gibi) gruplaştırmalar yapılmıştır. Böylece belirli bölgelerin ya da ürün türlerinin karşılaştırmalı analizleri yapılabilir hale gelmiştir.

**2.3 Analiz Yöntemi**

Analiz aşamasında hem tanımlayıcı istatistikler hem de görsel analiz teknikleri bir arada kullanılmıştır. R programlama dilinin güçlü veri işleme ve görselleştirme paketleri olan ggplot2, dplyr ve lubridate gibi kütüphaneler yardımıyla veri önce sınıflandırılmış, ardından grafiklerle desteklenerek yorumlanabilir hale getirilmiştir. Zaman serisi analizlerinde ortalama fiyat değişimleri aylık bazda incelenmiş; bu inceleme organik ve konvansiyonel ürün ayrımıyla birlikte yapılmıştır.

Analiz kapsamında ilk olarak genel fiyat değişim eğilimleri gözlemlenmiş, ardından en yüksek satış hacmine sahip ilk beş bölge seçilerek bu bölgelerdeki fiyat dağılımları kutu grafikleri ile görselleştirilmiştir. Bu görseller aracılığıyla bölgesel farklılıklar daha net bir şekilde ortaya konmuştur. Ayrıca ürün türüne göre fiyat farklılıkları da detaylandırılmış ve organik ürünlerin sistematik olarak daha yüksek fiyatlardan satıldığı sonucu elde edilmiştir.

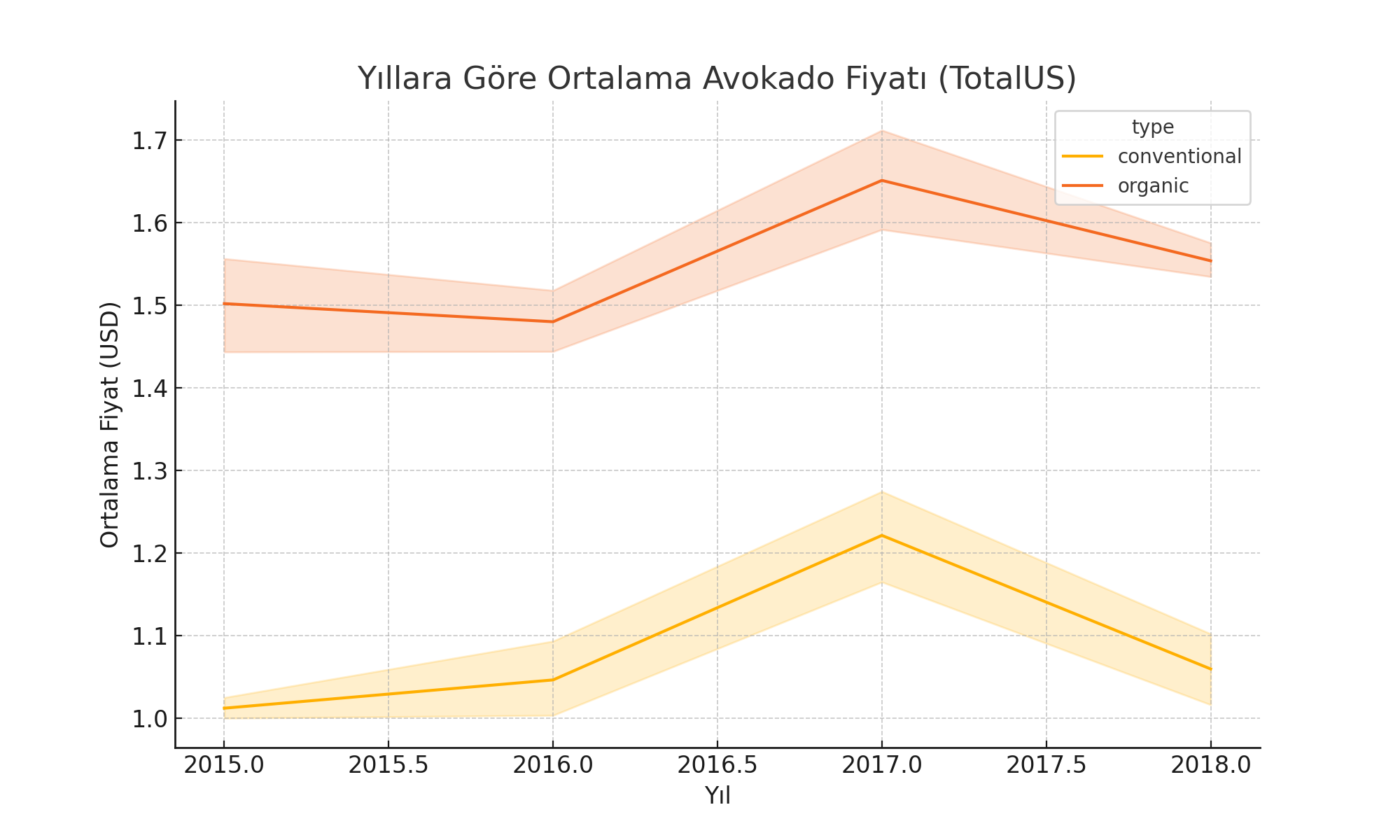
Yapılan tüm analizlerde araştırma sorularına odaklı, açıklayıcı ve sistematik bir yaklaşım benimsenmiştir. Görsel araçlar yalnızca destekleyici değil, aynı zamanda veri içgörülerini somutlaştıran temel unsurlar olarak kullanılmıştır. Bu şekilde analiz süreci sadece teknik değil, aynı zamanda yorumlayıcı bir yapıya kavuşturulmuştur.

1. **Bulgular ve Yorumlar**

Bu bölümde, veri seti üzerinden yapılan analiz sonuçları sunulmakta ve bunlar üzerinden çıkarımlar yapılmaktadır. Görsel analizlerle desteklenen bulgular sayesinde avokado pazarındaki eğilimler detaylı şekilde ele alınmaktadır.

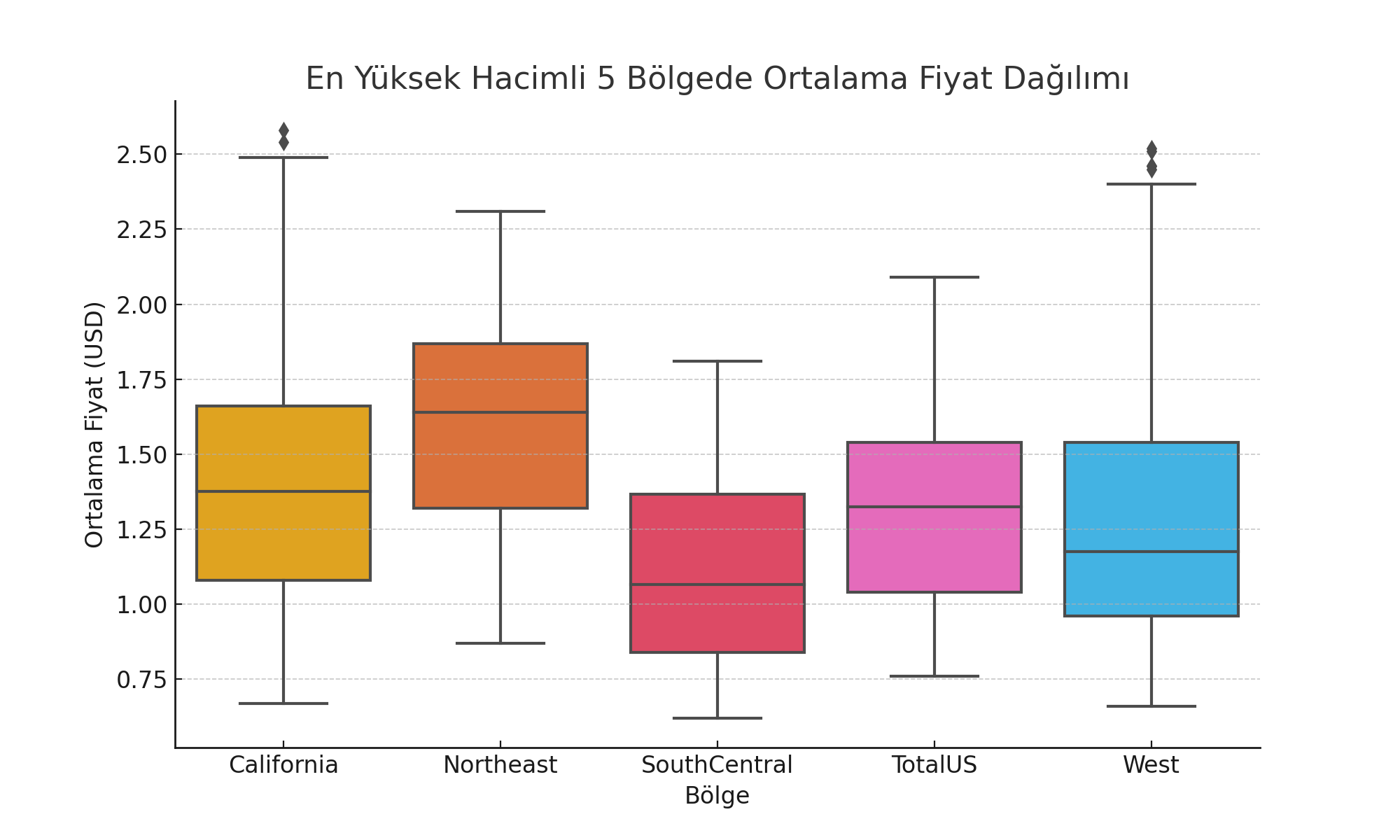
**3.1 Ortalama Fiyatların Zaman İçerisindeki Değişimi**

Organik ve konvansiyonel avokado fiyatları zaman içerisinde benzer dalgalanmalar göstermektedir. Ancak organik ürünlerin ortalama fiyatlarının sistematik olarak daha yüksek olduğu gözlemlenmektedir. Özellikle 2017 ve 2018 yıllarında organik ürünlerde fiyat artışı daha belirgin hale gelmiştir.



**3.2 Satış Hacmi ve Bölgesel Dağılım**

En yüksek satış hacmine sahip bölgeler arasında California, South Central, West ve Great Lakes gibi bölgeler öne çıkmaktadır. Bu bölgelerde tüketimin daha yoğun olması, fiyat üzerinde de etkili olmuştur. Bölgesel satış hacimleri ile ortalama fiyatlar arasında belirli bir korelasyon olduğu gözlemlenmiştir.



**3.3 Ürün Türüne Göre Fiyat Dağılımları**

Organik ve konvansiyonel ürünler arasında önemli fiyat farkları bulunmaktadır. Organik ürünler, hemen her bölgede konvansiyonel ürünlere göre daha yüksek fiyatla satılmaktadır. Bu durum, organik ürünlerin üretim maliyetleri ve tüketici tercihlerinin bir yansıması olarak yorumlanabilir.

### 4. Sonuçlar ve Öneriler

Bu çalışmada, Amerika Birleşik Devletleri’nde 2015–2018 yılları arasında satılan avokadolara ilişkin veriler kullanılarak çeşitli analizler yapılmıştır. Yapılan analizler sonucunda, avokado fiyatlarının zaman içerisinde dalgalı bir seyir izlediği, organik ürünlerin sistematik olarak daha yüksek fiyatlara satıldığı ve bölgesel farklılıkların belirgin olduğu ortaya konmuştur. Fiyatların satış hacmini etkilediği ve küçük boyutlu ürünlerin daha çok tercih edildiği de gözlemlenmiştir.

#### 4.1 Temel Bulguların Özeti

* Avokado fiyatları mevsimsel dalgalanmalar göstermektedir.
* Organik ürünler her bölgede daha pahalıdır.
* Batı yakası bölgelerinde fiyatlar ortalamanın üzerindedir.
* Fiyat arttığında satış hacmi azalma eğilimi göstermektedir.
* En çok tercih edilen ürün boyutu Small’dur.

#### 4.2 Politika ve Pazarlama Önerileri

* **Üreticiler için**: Mevsimsel arz-talep dengesine dikkat edilerek üretim planlaması yapılmalıdır.
* **Perakendeciler için**: Organik ürünlere yönelik kampanyalar, özellikle gelir düzeyi yüksek bölgelerde artırılabilir.
* **Tüketici eğilimleri**: Küçük boyutlu ürünlerin daha çok tercih edilmesi, bu ürünlerin pazarlamada öne çıkarılmasını gerekli kılmaktadır.

#### 4.3 Gelecekteki Araştırmalar İçin Öneriler

Bu çalışmada sadece ABD pazarı ele alınmıştır. Farklı ülkelerdeki fiyat ve tüketim eğilimlerini karşılaştırmak, küresel eğilimlerin belirlenmesi açısından önemlidir. Ayrıca pandemi sonrası dönem gibi dışsal şokların etkileri de gelecekteki çalışmalarda değerlendirilebilir.

## 5. Gelecek Çalışmalar İçin Öneriler

Bu çalışmada analiz, yalnızca 2015–2018 yılları arasındaki ABD avokado piyasası verileriyle sınırlı tutulmuştur. Gelecekteki araştırmalarda daha geniş zaman aralıkları veya farklı ülkeler de değerlendirilerek küresel ölçekte karşılaştırmalı analizler yapılabilir. Ayrıca, tüketici davranışlarını etkileyen sosyal medya etkisi, ekonomik krizler, pandemi gibi dışsal faktörlerin fiyatlara ve tüketim alışkanlıklarına etkisi de detaylı şekilde incelenmelidir.

İleri düzey çalışmalar için önerilebilecek bazı başlıklar şunlardır:

* Makine öğrenmesi algoritmaları kullanılarak fiyat tahmin modelleri oluşturulması
* Mevsimsel üretim verilerinin tarımsal politikalarla birlikte değerlendirilmesi
* Sosyo-ekonomik değişkenlerle (gelir düzeyi, eğitim, şehirleşme oranı) bölgesel analizlerin genişletilmesi
* Avokadonun diğer sağlıklı gıdalarla karşılaştırmalı pazar analizlerinin yapılması

Bu tür çalışmalar, pazarlama stratejileri geliştirmek isteyen işletmelere ve tarım politikası oluşturan karar vericilere yol gösterici olabilir.