

- ARAŞTIRMA MAKALESİ -

**ALTERNATİF VERİ ELDE YÖNTEMİ WEB MADENCİLİĞİ: OTEL ODA
FİYATLARININ ZAMANSAL ANALİZİ**

Yılmaz AĞCA¹

Öz

Turizm sektöründe oda fiyatları birçok faktörlere bağlı olarak sezon içerisinde değişkenlik göstermektedir. Fiyatlar talebin düşük olduğu dönemde düşük olurken, talebin yüksek olduğu dönemde daha yüksek olmaktadır. Fiyatlardaki zamansal bu değişimi detaylı olarak incelemek, gerekli verinin elde edilmesinin zorluğu nedeniyle çoğu zaman mümkün olmamaktadır. Bu çalışmanın amacı, alternatif bir veri elde yöntemi kullanarak, veri elde etme zorluğunun önüne geçmek ve otel oda fiyatlarının zamansal değişimlerini belirlemektir.

Veri elde etmenin üstesinden gelmek için web madenciliği/kazıma olarak adlandırılan alternatif bir yöntem kullanılmıştır. Bu amaçla, geliştirilen bir algoritması ile, bir tatil/konaklama arama motorundan Türkiye'nin Ege Bölgesinde bulunan toplam 675 kıyı otelinin 15 online seyahat acentesine ait 184 günlük fiyat verisi elde edilmiştir. Bu veri, oda fiyatlarının zamansal değişimini belirlemek için kullanılmıştır.

Analiz sonucunda ortalama oda fiyatları, aylara göre anlamlı farklılık gösterdiği görülmektedir ($\chi^2=164,58$; $p<0,05$). Fiyatların en yüksek olduğu ay Ağustos ayıdır. Fiyatlar haftalık olarak analiz edildiğinde de anlamlı farklılık göstermiştir ($\chi^2=144,43$; $p<0,05$). Buna göre sıralama ortalama değerlerin, en yüksek 31, 32, 33 ve 34. haftada olduğu belirlenmiştir. Bu dönem 30.07.2018-26.08.2018 aralığıdır. Fiyatlar hafta sonu ve hafta içi olarak anlamlı bir farklılık göstermemektedir ($U = 3243$, $p>0.05$). Resmi tatil olan dini bayramlarda ortalama oda fiyatlarında gözle görülür yükselişlerin olduğu tespit edilmiştir.

Bu araştırma ile, gerek veriyi elde etmede web madenciliğinin kullanılması, gerekse otel oda fiyatlarının 675 oteli içine alan tam bir bölgeyi kapsayacak biçimde analiz edilmesinin yazına katkı sağlayacağı düşünülmektedir. Ayrıca çalışmada, hem sektör yönetici ve yatırımcıları hem de turistler açısından, karar vermeyi destekleyici nitelikte, Ege bölgesine ait fiyat değişimleri ortaya konulmuştur.

Anahtar Kelimeler: Otel Oda Fiyatları, Web Madenciliği, Zamansal Değişim.

JEL Kodu: L83, C80.

Başvuru: 15.01.2021

Kabul: 01.10.2021

¹ Dr. Öğr. Üyesi, Tokat Gaziosmanpaşa Üniversitesi, Niksar Meslek Yüksekokulu, agca.yilmaz@gmail.com, Tokat, Türkiye, ORCID No: 0000-0002-5912-0977

ALTERNATIVE DATA COLLECTION METHOD WEB MINING: SEASONAL ANALYSIS OF HOTEL ROOM PRICE ²

Abstract

Hotel room prices vary during the season depending on various factors. Prices are low during the period of low demand, while they are higher during the period of high demand. It is often not possible to examine this temporal change in prices in detail due to the difficulty of obtaining the necessary data. The aim of this study is to overcome the difficulty of obtaining data and to determine changes in hotel room prices by using an alternative method of obtaining data.

An alternative method called web mining has been used to overcome the data collecting. For this purpose, a web-mining/scraping algorithm was used to collect data on a travel search engine. The room prices given by 15 travel agencies of the 675 hotels in the Aegean region coast of Turkey were collected by this algorithm. These 184 days of data were used to analyze room prices.

As a result of the analysis, it is seen that average room prices differ significantly by months ($\chi^2=164,58$; $p<0,05$). Prices are the highest in August. Prices also differed significantly when analyzed weekly ($\chi^2=144,43$; $p<0,05$). Accordingly, the ranking average values are highest at 31, 32, 33 and 34 weeks. This period is between 30.07.2018-26.08.2018. Prices did not differ significantly on weekends and weekdays ($U = 3243$, $p>0.05$). It has been determined that there are noticeable increases in average room prices during religious holidays, which are official holidays in Turkey.

With this research, it is thought that both the use of web mining in obtaining the data and the analysis of hotel room rates to cover a full region covering 675 hotels will contribute to the literature. In addition, the findings about Aegean Coast will be supportive for decide both sector administrators, investors and to tourists.

Keywords: *Hotel Room Prices, Web Mining, Seasonal Changes.*

JEL Code: *L83, C80.*

“Bu çalışma Araştırma ve Yayın Etiğine uygun olarak hazırlanmıştır.”

² The Extended English Summary is located the end of the article.

1. GİRİŞ

Ticari değer taşıyan ürünlerde pek çok değişkene bağlı olarak fiyat değişimlerini görmek mümkündür. Bu değişkenlerden biri de zamandır. Zamana göre ürün fiyatlarında değişim meydana gelir ve bu durum modern ekonominin temel konularından biridir (Gupta, 2004, s. 44; Baumol & Blinder, 2016, s. 478). Zaman, bu durumda bağımsız değişken gibi gözükse de, asıl fiyatı etkileyen, fiyat ile bağlantılı arz, talep, ürünü oluşturan şeylerin fiyatları, enflasyon vb. çeşitli değişkenlerin zaman içerisinde değişim göstermesidir.

Fiyat üzerinde etkisi olabilecek bu değişkenlerde, *-ki çoğu zaman bunlar işletme girdileridir*, meydana gelen değişimler, arzı gerçekleştirenler tarafından yapılan düzenlemelerle fiyatlara yansıtılmaktadır. Bu nedenle rekabet ön planda olacak şekilde otel oda fiyatları değişkenlik göstermektedir.

Turizm ürünü, turistik işletmelerin turistlere yönelik ürettiği mal ve hizmetlerden oluşan, üretilen bölgenin doğal ve kültürel güzellikleri ile desteklenen bileşik yapıdaki üründür. Bu ürün bir hediyeelik eşya gibi meta olabileceği gibi, rehberlik gibi soyut bir ürün de olabilmektedir. Turistik ürünün en önemli kısımlarından birini, otel hizmeti, yani konaklama oluşturmaktadır. Günümüzdeki turistik konaklama işletmeleri, konukların tesiste kaldıkları süre boyunca ihtiyaç duyabilecekleri pek çok hizmeti bünyesinde barındırmaktadır. Bunlara ek olarak, bu tesisler konukların tatil/konaklama süresi boyunca eğlendirecek çeşitli aktiviteleri de konuklarına sunmaktadırlar (Holloway & Taylor, 2006).

Turizm talebi, bir dönem içerisinde yeterli satın alma gücüne ve zamanına sahip, turizm faaliyetlerine katılma isteğine sahip olma veya katılan kişi sayısıdır (Frechtling, 2011, s. 4-5; Dwyer, Forsyth, & Dwyer, 2010, s. 5). Dolayısıyla, konaklama işletmelerinin doluluk oranları turizm talebinin bir kısmını oluşturmaktadır. Turizm talebi, ülkeye, bölgeye, otelin bulunduğu konuma ve diğer birçok faktöre göre yıl içerisinde değişkenlik gösterebilir. Dolayısıyla bu; konaklama tesislerinin doluluk oranlarının da sezon içerisinde değişkenlik göstermesi anlamına gelmektedir. Doluluk oranları, *yüksek sezon* denilen zamanlarda en üst seviyeye ulaşırken (Weiermair, Pechlaner, & Bieger, 2006, s. 131), sezon başlangıç ve bitiş dönemlerinde ise düşük olmaktadır. Bu, verimlilik ve kârlılık açısından istenen bir durum değildir. Konaklama işletmeleri, durumu dengelemek için sezonu genişletme çabalarının yanı sıra, yüksek sezonda fiyatları yüksek tutarak kârlılığı maksimize etmeye çalışmaktadırlar. Bir nevi ekonominin en temel unsurlarından birini uygulamakta, talep arttıkça fiyatları da artırmaktadırlar.

Türkiye, en çok turist çeken ülkeler sıralamasında Dünya'da ilk sıralarda yer almaktadır. Türkiye'ye gelen turistlerin önemli bir kısmı, kıyı turizmi olarak adlandırılan, güneş ve denizin çekim faktörü olduğu bir pazarlama stratejisi sonucu ülkeye gelmektedir (OECD Tourism Trends and Policies, 2018). Türkiye'nin özellikle Akdeniz ve Ege kıyıları, kıyı turizmi açısından hem yabancı hem de yerli turistler tarafından büyük talep gören bölgelerdir. Bu talepte, bölgenin sahip olduğu doğal ve kültürel öğelerin yanı sıra, bölgenin uzun güneşlenme süresinin, deniz suyu sıcaklığının da payı vardır.

İşletme tarafındaki durum, genel olarak yukarıda bahsedildiği şekildedir. Fiyatların bir de müşteri açısından durumunun değerlendirilmesi faydalı olacaktır. Müşteri doğası gereği gideceği yere, yerin tatil yapmak için *en uygun* olan mevsiminde, kendisi için *en uygun zamanda* ve *en düşük fiyatı* ödeyerek gitmek isteyecektir. Fakat bu üç arzusun birlikte gerçekleşmesi çoğu zaman mümkün olmamaktadır. Birçok kişinin aynı anda ürün talebinde bulunması, arzı kısa dönemde esnek olmayan ürünlerdeki turizm ürünü böyle bir üründür (Tribe, 2016, s. 93-107), fiyatların artmasına neden olmaktadır.

Hem işletme hem de müşteri açısından fiyatlardaki bu zamansal değişimi bilmek önemli bir husustur. Müşterinin tatil planı yaparken bu değişimi bilmesi, dönem ve fiyat çerçevesinde daha mantıklı bir karar vermesine katkı sağlayacaktır. İşletme açısından ise, her ne kadar her işletmeler bu tarz istatistikleri tutuyor olsa da, kendi bölgesi ve alternatif bölgeler arasında kıyaslama yapabilmesine, buna göre bir fiyat politikası oluşturabilmesine olanak sağlayacaktır.

Bu çalışmada, temel olarak *otel oda fiyatlarının zamana göre değişimi* araştırılmıştır. Bu kapsamda, *otel oda fiyatlarının aylara ve haftalara göre değişimi*, *hafta içi hafta sonu fiyatlarının farklılık gösterip göstermediği*, *Ege Bölgesi*’nde bulunan oteller üzerinde *illere göre istatistiki yöntemlerle analiz edilmiştir*.

Fakat bu tür bir araştırmayı yapmak için yeterli miktar veriye sahip olmak önemli faktörlerin başında gelmektedir. Klasik veri toplama yöntemleri ile bir coğrafi bölgenin otel oda fiyatlarının analizini yapmak oldukça zaman ve maliyet gerektiren güç bir iştir. Bu nedenle verinin elde edilmesinde son yıllardaki teknolojik gelişmelerin bilime kazandırdığı web madenciliği yönteminden faydalanılmıştır.

Web kazıma/madenciliği (web scraping/mining) web sitelerinde dağınık halde bulunan veri içerisinden istenen verinin alınması ve bunun otonom olarak yapılması yöntemlerine dayanmaktadır (Ağca, 2019, s. 55). Bu çalışmada bu yöntemin kullanılmasındaki amaç, klasik veri toplama yöntemi ile elde edilmesi zor olan verinin hızlı, doğru ve güvenilir bir şekilde elde edilmesidir. Web madenciliği yönteminin sağladığı olanaklar ile bir sezon boyunca Ege Bölgesi’ndeki 5 ilde yer alan 675 otelin 15 farklı seyahat acentesinden alınmış 184 günlük 72.104 satırdan oluşan fiyat verisi elde edilmiştir.

Bilindiği üzere mikroişlemcilerin icadı, bilgisayarlar, mobil aygıtlar ve internet gibi teknolojiler insan hayatı üzerinde büyük gelişmelere ve değişimlere neden olmuştur. Günümüzde neredeyse bütün bilim alanları, iş sektörleri ve kurumlar bu gelişmeleri kendi amaçları doğrultusunda kullanılmaktadırlar. *Yönetim Bilişim Sistemleri* ve *İşletme Enformatiği* olarak bilinen bilim dallarının varlığı, onlarca yıl geriye gitmesine rağmen gelişen bilgi ve iletişim teknolojileri sayesinde, günümüzde oldukça yaygınlaşmış, iş ve bilim alanında çokça faydalanan alanlar haline gelmiştir.

Yukarıda bahsedilen gelişmeler, turizm alanında da önemli değişimlerin meydana gelmesine yol açmıştır. Hatta daha önce olmayan iş alanlarının ortaya çıkmasına neden olmuştur. Bu yeni sektörlerden biri *online seyahat acenteleridir*. İnternetin icadı ile günlük hayatta

kullanılmaya başlanan bu kavram, turistik ürünlerin pazarlamasında radikal değişikliklere neden olmuştur.

Daha önce seyahat acentesi kataloglarından yapılan tatil/paket tur, konaklama tesisi pazarlaması, büyük oranda web siteleri üzerine, yani internete taşınmıştır. Bu web siteleri, acente kataloglarında olması mümkün olmayan tesislere ait pek çok bilgiyi üzerinde barındırmaktadır. Bu durum tüketici açısından büyük kolaylık ve avantaj sağlamaktadır. Bu nedenle de günümüzde online seyahat acenteleri pazar paylarını hızla artırmaktadırlar.

Online seyahat acenteleri tüketiciye tatil tercihinde sağladığı avantajların yanı sıra başka olanaklar da sağlamaktadır. Bu siteler pazarlamasını veya tanıtımını yaptıkları tesislere ait yüzlerce özelliğe ait bilgiyi veri tabanlarında tutmakta ve web üzerinden yayınlamaktadırlar. Online seyahat alanında uluslararası hizmet veren tripadvisor.com sitesi, sadece Türkiye’den bir sezonda 7.000’in üzerinde konaklama tesisine ait yaklaşık 50 adet değişikliğe ait veriyi yayınlamaktadır. Bu siteler bu verileri konaklama tesisine sağladıkları yönetici paneli vasıtasıyla elde etmekte ve veri tabanlarında saklamaktadırlar. Hali hazırda veriyi barındıran web siteleri çeşitli amaçlarla bu veriyi işlemekte ve kullanmaktadır. Fakat işledikleri bu verinin çıkarımlarını çoğu zaman kamuoyu ile paylaşmamaktadırlar (Ağca, 2019).

Fakat web madenciliği, belirli asgari düzeyde bant genişliğine sahip, internete bağlı her kullanıcıya, site üzerinde yayınlanmış olan veriye erişim ve kullanma olanağı sunmaktadır.

Bu araştırma web madenciliğinin sağladığı olanaklardan istifade etmeyi amaçlamaktadır. Keza bir sezon boyunca bir bölgeye ait fiyat verisinin eğilimini araştırmak, klasik veri toplama yöntemleri ile oldukça güçtür. Bu nedenle geliştirilen bir web madenciliği algoritması ile, bir tatil/konaklama arama motorundan toplam 675 otelin 15 online seyahat acentesine ait 184 günlük fiyat verisi elde edilmiş ve bu veri; *oda fiyatlarının zamansal değişimini* belirlemek için kullanılmıştır.

Bu çalışmanın veri kaynağını oluşturan Ege Bölgesi’nde 600’ün üzerinde otel/konaklama tesisi bulunmaktadır. Ayrıca fiyatlardaki zamansal değişimi belirlemek için belirli bir aralıkta fiyat verisinin tekrar tekrar alınması gerekmektedir. Bu çalışmada 184 günlük fiyat her tesis için tekrar alınmıştır. 675 konaklama tesisine ait 15 acentenin verdiği 184 günlük veri yaklaşık 1,85 milyon satır tutmaktadır. Bu büyüklükte veriyi elde etmek web madenciliği sayesinde mümkün olmuştur.

Sadece kullanılan veri elde etme yöntemi ile değil, aynı zamanda bu büyüklükte bir veri üzerinde yapılan istatistiksel analizler turizm sektörü açısından önemli olabilecek bulguların elde edilmesini sağlamıştır. Bulguların gerek araştırmacılar gerekse sektör yöneticileri ve yatırımcıları açısından faydalı olacağı umut edilmiştir.

Çalışmada öncelikle bu alanda daha önce yapılan benzer araştırmalara yer verilecek, sonrasında araştırma yöntemi, verinin işlenmesinde ve analizinde kullanılan yöntemler açıklanacaktır. Bulgular bölümünde Ege Bölgesi’ndeki konaklama tesislerine ait analiz sonuçları il ve ilçelere göre verilecektir. Burada, fiyatlardaki zamansal değişim, fiyatların hafta içi ve hafta sonu olarak farklılık gösterip göstermediği, bölgeler arası değişimlere ilişkin

istatistiki test sonuçları sunulacaktır. Sonuç ve tartışma kısmında araştırma sonucu çeşitli açılardan değerlendirilecektir.

1.1. Literatür Taraması

Literatürde bir bölgenin bu kadar kapsamlı zamansal fiyat değişimini araştırmaya yönelik çok fazla araştırmaya rastlanmamasına rağmen, geçmişten günümüze değin araştırmacılar için merak edilen konulardan biri olmuştur.

Raya (2010) çalışmasında Katalanya kıyılarında 111 otele ait, 2008 Ağustos'un ilk haftasına ait fiyat verisini, Ocak'ın ikinci haftasından Haziran'ın dördüncü haftasına kadar izlemiştir. Fakat burada oda fiyatını tutarı değil sadece değişip değişmediği takip edilmiştir. Yani fiyatın alındığı bir önceki haftadan o haftanın fiyatı farklıysa bağımlı değişken "1" kabul edilmiş, değilse "0" olarak alınmıştır. Çalışmadaki amaç, bölgedeki fiyatın zamansal değişimini belirlemekten ziyade, otellerin doğru fiyat politikaları uygulayıp uygulamadığına dair bulguları ortaya koymaktır.

Yazında otel oda fiyatını etkileyen faktörleri bulmaya yönelik çalışmaları görmek mümkündür (Carvell & Herrin, 1990; Bull, 1994; Papatheodorou, 2002; Pawlicz & Napierala, 2017; Taylor, 1995; Thrane, 2005). Çoğu *hedonik fiyatlandırma modeli* temelli olan bu çalışmalar, bağımsız değişken olarak modellerine *zamanı* dahil etmemişlerdir. Coenders, Espinet ve Marc Saez (2003) ise oluşturdukları tahmin modelinde birçok değişkene ilaveten *zaman değişkenini* de eklemişlerdir. Araştırma verisi İspanya'nın beş turizm bölgesinde bulunan toplam 471 adet otelden günlük alınmıştır. Veri kaynağı olarak seyahat acenteleri katalogları kullanılmıştır. Kapsamlı olan bu çalışmada yüksek sezon ile düşük sezon fiyatları arasında %50'ye yakın bir fiyat farkı belirlenmiştir. Oluşturulan grafikte, havalanın sıcak geçtiği yaz aylarında fiyatların yüksek olduğu görülmektedir. Fiyat zaman grafiği incelendiğinde, tam olmasa da çan eğrisine benzer bir yayılım olduğu görülmektedir (Coenders, Espinet, & Saez, 2003). Bu araştırma ile Coenders, Espinet ve Marc Saez (2003) çalışmalarının ortak ayları alındığında, benzer bir eğri görülmektedir. Araştırmanın yapıldığı yerlerin Akdeniz kıyısında ve benzer iklim özelliklerine sahip olmasının rolü olduğu düşünülebilir.

Sezonluk fiyat değişimi üzerine odaklanan bir diğer çalışma yine İspanya'da yapılmıştır (Saló, Garriga, Rigall-I-Torrent, Vila, & Sayeras, 2012). Araştırma verisi broşürlerden ve belediyeden, haftalık olarak alınmıştır. Burada veri setine otellere ilaveten turistlere kiralanan apartman daireleri de dahil edilmiştir. Araştırma sonucunda kiralık dairelerin fiyatlarındaki değişimin otellere kıyasla daha yumuşak olduğu görülmüştür. Bu çalışmada da Coenders, Espinet ve Marc Saez'in (2003) çalışmalarındakine benzer şekilde yüksek sezon (peak periods) Ağustos ayında yaşanmaktadır.

Bu alanda yapılan bir diğer araştırma Çin'in Hainan bölgesinde yapılmıştır (Wang, Sun, & Wen, 2019). Bu bölge deniz-kum-güneş turizmine dayalı bir turizm arzına sahiptir. Bu açıdan İspanya ve Türkiye'deki yaygın turizm arzına benzerdir. Fakat turizm talebinin bütün bir yıla yayıldığı görülmektedir. Yüksek sezonun Kasım-Nisan aralığında olduğu görülmektedir. Araştırmacılar bu durumu, o aylarda ülkenin geri kalanındaki soğuk hava koşulları dolayısıyla, turistlerin daha sıcak bölgeleri talep etmesi şeklinde yorumlamışlardır. Araştırma verisi bir online seyahat acentesinden alınmış, 346 otele aittir. Oda fiyatlarındaki dağılım da

talebe benzer şekilde çok keskin değildir. Örneğin Ocak ve Şubat ayında ortalama oda fiyatı yaklaşık %0,15 gibi küçük bir yükseliş göstermiştir.

Çalışma bir bölgenin bir sezon boyunca ki verilerini zamansal değişim açısından ele alarak fiyat konusunda elde edilen bulgular ile literatüre katkı sağlamayı amaçlamaktadır. Bu açıdan ele alındığında büyük bir veri seti ile yapılan bir çalışmalardan olacaktır. Yazında yapılan araştırmada Türkiye’de bu çapta bir araştırmaya rastlanmamıştır. Bu açıdan, yazındaki boşluğu dolduracağı düşünülmektedir.

2. YÖNTEM

Otel oda fiyatlarının zamansal olarak incelenmesi üzerine yapılan bu araştırmada veri toplama yöntemi olarak veri madenciliği kullanılmıştır. Otel, konaklama ve meta arama alanlarında hizmet vermekte olan bir arama motorundan veriler veri madenciliği kullanılarak elde edilmiştir. Veri, 20-21 Mart 2018 tarihinde, bir algoritma yardımı ile sadece Ege bölgesindeki *kıyı otelleri* olacak şekilde indirilmiştir. Bu tercihin nedeni, tatil amaçlı otellerde şehir otellerinden daha fazla mevsimsel fiyat dalgalanmalarının olduğunun düşünülmesidir.

Sitede filtreleme yapılarak sadece *otel* ve *resort* olan konaklama tesisleri seçilmiştir. Kendisini apart otel, ev/daire, pansiyon olarak nitelendiren işletmeler araştırmaya dâhil edilmemiştir. Otele ilk giriş tarihi *26 Mart 2018* ve son giriş tarihi *29 Eylül 2018* olacak şekilde sadece *bir gecelik, iki kişi* konaklama fiyatları alınmıştır. 26 Mart 2018’den başlayacak 30 Eylül 2018’de bitecek şekilde, Ege bölgesi kıyılarında bulunan, otel ve resort otel olarak kendisini nitelendiren tesislerin bir gecelik iki kişi konaklama fiyatları alınarak bir veri tabanı oluşturulmuştur. Oluşturulan veri tabanındaki veriler, analize uygun olacak biçimde bir ön işlemden geçirilmiştir. Bu ön işleme işleminde, hatalı veriler düzeltilmiş veya temizlenmiştir.

Araştırmada beş ilde bulunan 675 otelin 15 online seyahat acentesinden alınmış 184 günlük 72.104 fiyat verisi kullanılmıştır.

Veriler üzerindeki istatistiksel analizler için IBM SPSS v.23³ paket programı grafiksel gösterim için MS Excel 2016⁴ ve veri madenciliği için R Studio 1.0.153⁵, kullanılmıştır. Verilerin normal dağılıma uygunluğu için Kolmogorov-Smirnov ve Shapiro-Wilk testi yapılmış, normal dağılım veriler için parametrik, normal dağılmayan veya kategorik olan değişkenler için karşılıkları parametrik olmayan testler uygulanmıştır (Field, 2009, s. 179-180).

Verinin alındığı sitede tek bir otel oda fiyatı bulunmamaktadır. Site, otel için birçok online seyahat acentesinin vermiş olduğu fiyatları ziyaretçilere toplu halde sunmaktadır. Kullanılan algoritma, bu fiyatlar içerisinde en düşük fiyatı baz almaktadır. Site otelleri il ve yer olarak

³ <https://www-01.ibm.com/software/tr/analytics/spss/>

⁴ <https://products.office.com/en-us/excel>

⁵ <https://www.rstudio.com/>

(genellikle bulunduğu ilçe, mahalle ve köy) kategorize etmiştir. Çalışmada, sitede yer aldığı biçimde, bu yer ve bölgeler sınıflandırılmıştır.

Aylık fiyat değişim aralığı hesaplanırken o ayki en yüksek oda fiyatı ile o ayki en düşük oda fiyatının farkı alınmıştır.

3. BULGULAR

Yapılan araştırmaya, dâhil edilen otellerin illere ve ilçelere göre dağılımları, misafirler tarafından yapılan ortalama yorumlar ve yine misafirler tarafından verilen ortalama puanlara ilişkin veriler Tablo 1’de verilmektedir.

Tablo 1: Otel Sayısı 50’nin Üzerinde Olan İl ve İlçelerdeki Otellerin Dağılımı

Bölge	Otel Sayısı		Ortalama	
	N	%	Yorum	Puan
Aydın Çevresi	95	14,1	407,47	7,61
Kuşadası	64	67,37	516,82	7,67
Balıkesir Çevresi	58	8,59	96,82	7,76
İzmir Çevresi	201	29,8	192,12	8,13
Alaçatı	84	41,79	93,99	8,4
Çeşme	75	37,31	255,06	7,82
Muğla Çevresi	284	42,1	294,43	7,95
Bodrum	112	39,44	359,64	7,99

Çıktılara göre; verinin elde edildiği online seyahat acentesi sitesinde 26.03.2018-30.09.2018 tarihleri arasında Ege bölgesinde en çok otel, sırası ile Muğla (284), İzmir (201) ve Aydın (95) illerinde bulunmaktadır.

İlçeler bazında bakıldığında Bodrum 112 otel ile Ege bölgesindeki toplam otel sayısı içerisinde yaklaşık %16.60’lık bir paya sahiptir. İlçelerde sıralama Alaçatı (84), Çeşme (75) ve Kuşadası (64) şeklinde devam etmektedir. Burada dikkat edilecek husus, otel sayılarının verinin alındığı web sitesinde 26.03.2018-30.09.2018 tarihleri yer alan otel sayılarına ait oluşudur. Turizm arzının kısa dönemde esnek olmayışından hareketle yine de bölgenin otel sayısının oransal dağılımı hakkında önemli olabilecek bilgiler vermektedir.

Bulgulardan elde edilebilecek bir diğer önemli husus yorum sayılarıdır. Her ne kadar yorum sayısı görünürde çok önemli gibi görünmese de otelin ve yerin tercih edilme oranı hakkında fikir verebilecek bir değişkendir (Memarzadeh, Blum, & Adams, 2015; Keskinçelik, Ağca, & Karaman, 2016). Konukların tesislere verdiği puanlar değerlendirildiğinde, Alaçatı (Ort: 8,4) ve İzmir çevresinin (Ort: 8,13) diğer bölgelere göre daha yüksek puan aldıkları belirlenmiştir. Ortalama yorum sayısına bakıldığında Aydın Çevresinin (N=95, Ort.=407,47) en yüksek yorum ortalamasına sahip olduğu görülmektedir. İkinci sırada Muğla Çevresi (N=284, Ort.=294,43) üçüncü sırada ise İzmir (N=201, Ort.=192,12) yer almaktadır.

İlçeler en çok yorum almalarına göre sıralandığında, sırası ile Karaburun (N=2, Ort.=1.853,76), Özdere (N=3, Ort.=1.059,53) ve Seferihisar (N=4, Ort.=818,25) gelmektedir. Fakat bu durumun böyle olmasında Karaburun, Özdere ve Seferihisar'da yüksek yorum olan az otelin olmasının payı olabilir. Örneğin Karaburun'un araştırmaya dâhil edilen iki otelinden biri 1.957 yorum almıştır.

Tablo 2'de Ege bölgesindeki illerin Mart-Eylül dönemindeki ortalama oda fiyatları verilmektedir.

Tablo 2: Oda Fiyatlarının Aylara ve Bölgelere Göre Ortalamaları

Bölge	Mar	Nis	May	Haz	Tem	Ağu	Eyl
Aydın Çevresi	238,80	245,15	242,98	294,05	380,03	374,74	244,56
Balıkesir Çevresi	194,40	227,04	227,15	291,30	364,50	385,56	270,81
Çanakkale Çevresi	434,57	620,89	367,28	450,21	433,53	583,07	434,05
İzmir Çevresi	265,62	325,41	304,29	391,72	462,85	486,55	360,41
Muğla Çevresi	210,73	255,42	298,97	409,61	511,25	529,15	354,89
Toplam	241,50	292,02	289,54	380,81	465,17	487,66	415,67

* Fiyatlar 20-21 Mart 2018 tarihinde iki kişi bir gecelik çift kişilik (*double*) oda olacak şekilde alınmıştır.

Bu tarihler arasında Türkiye Cumhuriyeti Merkez Bankası'na (TCMB) ait USD döviz kuru ortalaması 3,9446 TL'dir.

MB'nin sağladığı ve Türkiye İstatistik Kurumu'nun (TÜİK) tüketici fiyat endeksine (TÜFE) göre hazırladığı enflasyon hesaplayıcısına göre fiyatların alındığı 2018 Mart ayı ile araştırmanın hazırlandığı Aralık 2020 arasındaki enflasyon katsayısı 1,5'dir. Buna göre 2018 Mart ayındaki 100 TL'nin Aralık 2020 karşılığı yaklaşık 150 TL'dir.

Bu oranlar kullanılarak güncel fiyatlar hesaplanabilir.

Kaynak:

http://www3.tcmb.gov.tr/enflasyonecalc/enflasyon_hesaplayici.html

https://www.tcmb.gov.tr/kurlar/kurlar_tr.html

Sonuçlar incelendiğinde fiyatların en yüksek olduğu ayların Temmuz, Ağustos ve Eylül ayları olduğu görülmektedir. Burada dikkat çeken başka bir husus Çanakkale ilindeki otellerin ortalama oda fiyatlarının Nisan ayında diğer aylardan görece yüksek oluşudur. Bu dönemde Çanakkale'deki bazı otellerin özellikle hafta sonu yüksek fiyat uygulamasının bunda payı yüksektir. Aşağıda verilen Tablo 3'de bu ildeki fiyatların değişim aralığının da yüksek olduğu görülmektedir. Daha net sonuçların ortaya çıkarılması için fiyatların daha dar aralıklarda incelenmesi faydalı olacaktır.

Tablo 3'de ve Tablo 4'de oda fiyatlarının *aylara ve bölgelere göre değişim aralıkları ve standart sapmaları* verilmektedir. Değişim aralığı ildeki o ay içerisindeki en yüksek oda fiyatından en düşük oda fiyatının çıkarılması ile elde edilmiştir. Bu sonuçlar o bölgedeki fiyat çeşitliliği hakkında bilgi vermektedir.

Tablo 3: Oda Fiyatlarının Aylara ve Bölgelere Göre Değişim Aralıkları ve Standart Sapmaları (1)

Bölge	Mart		Nisan		Mayıs	
	Mak-Min	Std.S	Mak-Min	Std.S	Mak-Min	Std.S
Aydın Çevresi	2.367,00	343,54	2383,00	348,05	2.361,00	310,25
Balıkesir Çevresi	334,00	71,81	689,00	99,08	743,00	114,10
Çanakkale Çevresi	2.275,00	590,07	4677,00	1.058,37	4.491,00	534,57
İzmir Çevresi	2.357,00	250,39	19.524,00	518,95	2.029,00	148,09
Muğla Çevresi	998,00	130,42	4320,00	254,66	3.435,00	264,73
Toplam	2.382,00	249,54	19.550,00	437,87	4.626,00	251,49

Sonuçlar incelendiğinde *Çanakkale* her ay için *en yüksek standart sapma* oranına sahiptir. Nisan ayı için İzmir’de (D.A = 19.524,00), Çanakkale’de (D.A = 4.677,00) ve Muğla’da (D.A = 4.320,00); Mayıs ayı için Çanakkale (D.A = 4491,00) ve Muğla’da (D.A = 3.435,00) değişim aralıkları diğer aylara oranla daha yüksektir. Burada Mart ayına ait değişim aralığı ve standart sapma sadece altı günlük veri olduğu için verilmemiştir.

Tablo 4: Oda Fiyatlarının Aylara ve Bölgelere Göre Değişim Aralıkları ve Standart Sapmaları (2)

Bölge	Haziran		Temmuz		Ağustos	
	Mak-Min	Std.S	Mak-Min	Std.S	Mak-Min	Std.S
Aydın Çevresi	2.365,00	325,39	2.340,00	369,81	2.342,00	358,35
Balıkesir Çevresi	1.244,00	167,64	1.340,00	229,28	1.357,00	239,87
Çanakkale Çevresi	4.516,00	700,32	4.554,00	637,45	4.576,00	935,66
İzmir Çevresi	1.684,00	194,19	1.540,00	215,52	4.295,00	286,56
Muğla Çevresi	5.981,00	379,98	5.980,00	508,51	5.972,00	489,49
Toplam	6.013,00	338,99	5.991,00	416,36	5.983,00	438,50

Haziran, Temmuz ve Ağustos ayı için değişim aralıkları ve standart sapmalar incelendiğinde genel olarak değişim aralığının Mart, Nisan ve Mayıs aylarına oranla daha yüksek olduğu görülmektedir. Daha yüksek standart sapmalar ve değişim aralıkları fiyatların daha geniş aralıkta olduğunu göstermektedir. Bu üç ay incelendiğinde en yüksek değişim aralıkları Aydın, Çanakkale ve Muğla’da görülmektedir. Bu durum bu illerde fiyatları yüksek ve düşük tesislerin bulunabileceğini göstermektedir.

Oda fiyatlarının aylara göre anlamlı bir farklılık gösterip göstermediğini bulmak için veri üzerinde Kruskal-Wallis testi yapılmıştır. Bu analize ve veriye sadece 6 günlük fiyat verisi içerdği için Mart ayı dâhil edilmemiştir. Buna göre yapılan analize ait sonuçlar Tablo 5’de verilmektedir.

Tablo 5: Ege Bölgesi Oda Fiyatlarının Aylara Göre Kruskal-Wallis Testi Sonuçları

Ay	N	Sıralama Ortalaması	χ^2	df	p
Nisan	30	40,52	144,43	5	0,001
Mayıs	31	42,95			
Haziran	30	103,73			
Temmuz	31	140,16			
Ağustos	31	160,90			
Eylül	30	61,47			

Toplamda 183 gün için oda fiyatı analize dâhil edilmiştir. Bu verinin ortalaması = 375,95; standart sapması = 87,39'dır. Elde edilen sonuçlara göre oda fiyatları aylara göre anlamlı farklılık göstermektedir ($\chi^2=144,43$; $p<0,05$). Sıralı ortalama değerlerine bakıldığında en yüksek değerler Ağustos (160,90) ve Temmuz (140,16) aylarında görülmektedir.

Haftalık temelde fiyatları değerlendirmek ve haftalara göre fiyatların aralarında anlamlı bir fark olup olmadığını belirlemek için veri üzerinde analiz yapılmıştır. Yapılan Kruskal-Wallis testi sonucu Tablo 6'de verilmektedir.

Tablo 6: Ege Bölgesi Oda Fiyatlarının Haftalara Göre Kruskal-Wallis Testi Sonuçları

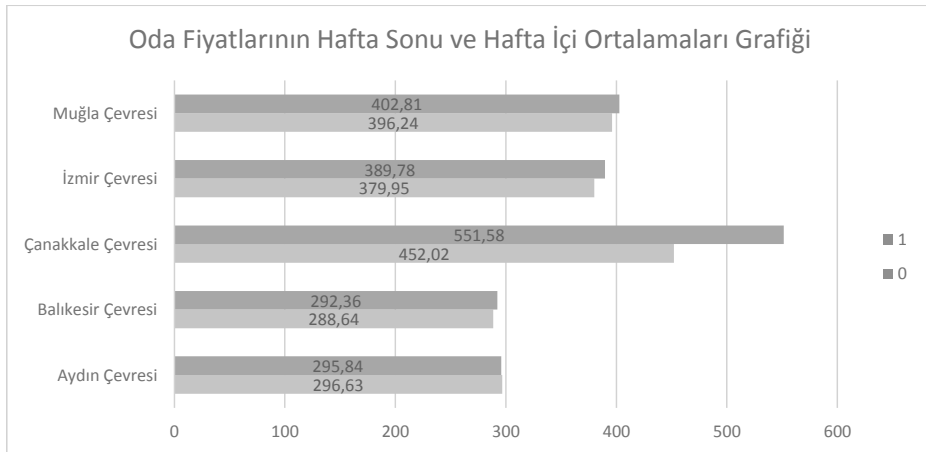
No	Hafta Aralık	N	Sıralama Ortalaması	χ^2	df	p
14	02.Nis-08.Nis	7	49,00	164,582	25	0,001
15	09.Nis-15.Nis	7	22,86			
16	16.Nis-22.Nis	7	46,79			
17	23.Nis-29.Nis	7	47,86			
18	30.Nis-06.May	7	26,21			
19	07.May-13.May	7	26,29			
20	14.May-20.May	7	43,57			
21	21.May-27.May	7	55,57			
22	28.May-03.Haz	7	60,43			
23	04.Haz-10.Haz	7	79,86			
24	11.Haz-17.Haz	7	115,29			
25	18.Haz-24.Haz	7	113,14			
26	25.Haz-01.Tem	7	110,00			
27	02.Tem-08.Tem	7	129,86			
28	09.Tem-15.Tem	7	132,86			
29	16.Tem-22.Tem	7	143,71			
30	23.Tem-29.Tem	7	147,00			
31	30.Tem-05.Ağu	7	157,14			
32	06.Ağu-12.Ağu	7	165,14			
33	13.Ağu-19.Ağu	7	161,00			
34	20.Ağu-26.Ağu	7	178,43			
35	27.Ağu-02.Eyl	7	124,00			
36	03.Eyl-09.Eyl	7	97,00			
37	10.Eyl-16.Eyl	7	86,86			
38	17.Eyl-23.Eyl	7	43,14			
39	24.Eyl-30.Eyl	7	16,00			

Analiz sonucunda haftalara göre ortalama oda fiyatları Ege bölgesi için anlamlı farklılık göstermiştir ($X^2=164,58$; $p<0,05$). Buna göre sıralama ortalama değerleri en yüksek 31, 32, 33 ve 34. Haftada görülmektedir. Tabloda verilmeyen bu haftaların ortalama oda fiyatları ise 31. haftada = 481,02; 32. haftada = 488,27; 33. haftada = 492,93 ve 34. haftada = 532,59'dır. Bu dönem (30.07.2018-26.08.2018) Ege bölgesinin en yüksek oda fiyatlarına sahip dönemidir. Tablo 7'de oda fiyatlarının *hafta sonu* ve *hafta içi* ortalamaları verilmektedir.

Tablo 7: Oda Fiyatlarının Hafta Sonu ve Hafta İçi Ortalamaları

Bölge	Hafta	Nis	May	Haz	Tem	Ağu	Eyl	Toplam
Aydın Çevresi	Hafta İçi	247,14	243,76	293,54	379,83	372,76	241,23	296,63
	Hafta Sonu	240,69	240,75	295,26	380,53	380,66	250,54	295,84
Balıkesir Çevresi	Hafta İçi	222,97	225,68	292,81	367,12	379,66	267,61	288,64
	Hafta Sonu	236,37	231,32	287,62	358,06	404,73	276,37	292,36
Çanakkale Çevresi	Hafta İçi	619,32	342,96	418,01	416,09	541,50	388,55	452,02
	Hafta Sonu	625,49	456,16	547,86	483,53	728,00	522,42	551,58
İzmir Çevresi	Hafta İçi	312,70	305,09	391,45	463,61	484,89	354,80	379,95
	Hafta Sonu	357,21	301,86	392,43	460,82	492,23	370,55	389,78
Muğla Çevresi	Hafta İçi	252,36	297,45	410,50	511,13	525,01	347,86	396,24
	Hafta Sonu	262,17	303,30	407,49	511,54	541,55	367,25	402,81
Toplam	Hafta İçi	287,86	288,34	380,16	464,67	483,20	326,14	371,04
	Hafta Sonu	301,73	293,06	382,40	466,44	501,60	344,25	379,29

Sonuçlar incelendiğinde Aydın hariç diğer bütün illerde hafta sonu fiyatlarının hafta içi fiyatlardan daha yüksek olduğu görülmektedir. Sonuçların bir grafik halinde incelenmesi faydalı olacaktır. *Şekil 1*'de Ege bölgesinin oda fiyatlarının hafta sonu ve hafta içi ortalamaları, sütun grafik şeklinde verilmektedir.



Şekil 1: Oda Fiyatlarının Hafta Sonu ve Hafta İçi Ortalamaları Grafiği

Şekil 1 incelendiğinde Çanakkale çevresinin hafta sonu (1) ortalama oda fiyatlarının hafta içi (0) ortalama oda fiyatlarından gözle görülür biçimde farklı olduğu sonucu çıkmaktadır. Diğer bölgelerde ortalamalar birbirlerine yakındır.

Bu durumun istatistiki olarak anlamlı olup olmadığını test etmek için Mann-Whitney U testi yapılmıştır. Oda fiyatının o günün hafta sonu veya hafta içi olması durumuna göre farklılık gösterip göstermediğini test etmek için yapılan analiz sonuçları Tablo 8’de verilmektedir.

Tablo 8: Ege Bölgesi Oda Fiyatlarının Hafta İçi-Hafta Sonu Durumları Mann-Whitney Testi Sonuçları

Günler	N	Sıralama Ortalaması	Mann-Whitney U Değeri	p
Hafta İçi	130	90,45	3.243	0,534
Hafta Sonu	53	95,81		

Analiz sonucunda oda fiyatlarının ortalaması = 375,95 standart sapması 87,40 bulunmuştur. Elde edilen bulgular, hafta sonu oda fiyatları ile hafta içi oda fiyatları arasında anlamlı bir farklılığın olmadığını göstermektedir (U = 3.243, p>0.05). Hesaplanan sıralı ortalama değerlerine bakıldığında, hafta sonu oda fiyatının (95,81) hafta içindeki oda fiyatından (90,45) yüksek olduğu görülmektedir.

3.1. Zaman Serisi

Zamana göre bağımlı değişkenin değerini tahmin etmek için birçok yöntem kullanılmaktadır. Bu yöntemlerden biri de ARIMA modelleridir. ARIMA, tek değişkenli zaman serisi tahmini yapmaya yarayan Box-Jenkins modelleri olarak da adlandırılan istatistiki yöntemlerdir (Erden, 2020). Bu kısaltma, Auto Regressive (AR), Integrated (I) ve Moving Average(MA) kısımlarından oluşmaktadır. AR seviyesi p , I seviyesi yani kaç fark alındığı d ve MA seviyesi q ile ifade edilmektedir (Kara, 2020). ARIMA modeli aşağıdaki denklem ile ifade edilmektedir (Prabhakaran, 2020).

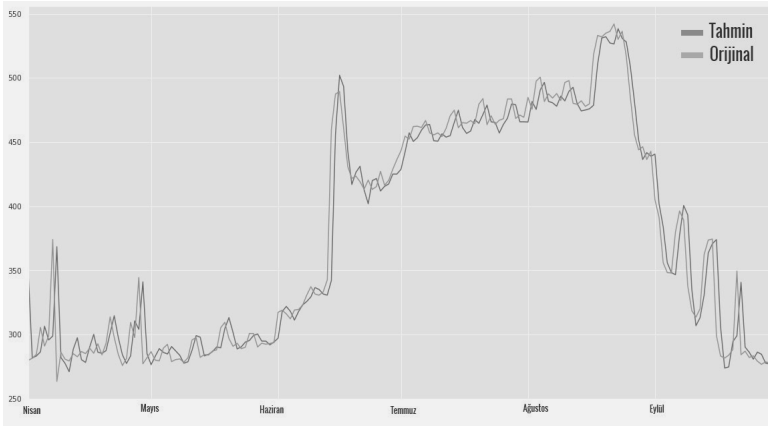
$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} \quad (1)$$

ARIMA modeli varsayım olarak durağanlık (stationary) özelliği taşıyan veya durağanlaştırılan veriler ile yapılmaktadır. Bu nedenle durağan serilere uygulanabilen ARMA modelinin, durağanlaştırılan verilerle de uyumlu bir alternatifi olarak kullanılmaktadır (Valipour, Ebrahim Banihabib, & Behbahani, 2013; Krishna, Iyer, & Sanders, 2015). Araştırma verisi ACF ve PACF grafikleri ve Dickey-Fuller testi ile analiz edilmiş durağana yakın bir sonuç olmasına rağmen durağanlık ön şartını taşımadığı görülmüş, dolayısıyla ARIMA modeli tercih edilmiştir. Verinin durağanlaştırılması için $\log X_t - \text{Rolling_Window}$ yöntemi kullanılmıştır (Nikolopoulos & Thomakos, 2019, s. 65; Maklin, 2019; Sharma, 2019). Buradaki ortalamalar haftalık alınmıştır. Analizler, Python programlama dili ile yapılmıştır. En uygun ARIMA modelini bulmak için pmdarima kütüphanesi altında yer alan auto_arima metodu kullanılmıştır. Analiz sonucunda en uygun modelin ARIMA(0,1,0) olduğu görülmüştür. Buna göre model özeti ’de verilmektedir.

ARIMA Model Results						
Dep. Variable:	D.Toplam		No. Observations:	182		
Model:	ARIMA(0, 1, 0)		Log Likelihood	-810.874		
Method:	css		S.D. of innovations	20.830		
Date:	Wed, 30 Dec 2020		AIC	1625.747		
Time:	16:58:49		BIC	1632.155		
Sample:	04-02-2018		HQIC	1628.345		
	- 09-30-2018					
	coef	std err	z	P> z	[0.025	0.975]
const	0.0046	1.544	0.003	0.998	-3.022	3.031

Şekil 2: ARIMA Modeli Sonucu

Eğer bağımlı değişken durağan değilse bunun için mümkün olan en basit model ARIMA(0,1,0) modelidir ve bu aynı zamanda rastgele değişme (random walks) olarak da adlandırılmaktadır (Nau, 2020). Bu sonuçlara göre veri ve tahmin grafiği çizildiğinde Şekil 3 ortaya çıkmaktadır.



Şekil 3: Zaman Serisi Grafiği

Orijinal veriye ve tahmin verisine ait şekil incelendiğinde tahmin verisinin orijinal veriye oldukça yakın bir doğrulukta olduğu görülmektedir. *Yüksek Sezon* olarak nitelendirilen dönemde fiyatların en yüksek seviyesine ulaştığı görülmektedir. Şekilde haziran ayı ortalarında fiyatlarda ani bir yükseliş dikkat çekmektedir. Bunun nedeninin, verinin toplandığı yıl olan 2018'de *Ramazan Bayramı*'nın 15-17 Haziran tarihleri arasında olmasının ihtimali yüksektir. Benzer şekilde *Kurban Bayramı* tarihleri olan 21-24 Ağustos aralığında da benzer yükseliş görülmektedir. Bu iki yükseliş göz ardı edildiğinde Mayıs-Eylül aralığında normal dağılıma benzer bir fiyat dağılımının olduğu görülmektedir.

4. TARTIŞMA

Bu çalışmada, otel oda fiyatlarının zamansal değişimi araştırılmıştır. Bu amaçla, Ege bölgesindeki 675 kıyı otelinin fiyat verileri, üzerinde birçok online seyahat acentesine ait fiyat verilerini barındıran bir siteden *web madenciliği* kullanılarak elde edilmiştir. Sonuçlar incelendiğinde Ege bölgesindeki oda fiyatların en yüksek olduğu ayların Temmuz (ortalama = 465,17), Ağustos (ortalama = 487,66) ve Eylül (ortalama = 415,67) ayları olduğu görülmektedir.

Fiyatların aylara göre farklılığının anlamlı olup olmadığı araştırıldığında; oda fiyatlarının aylara göre anlamlı farklı olduğu görülmüştür ($X^2=144,43$; $p<0,05$). Veriler haftalık olarak ele alındığında fiyatlar yine anlamlı farklılık göstermiştir ($X^2=164,58$; $p<0,05$). Buna göre sıralama ortalama değerleri en yüksek 31, 32, 33 ve 34. Haftalarda ortalama oda fiyatları; 31. haftada = 481,02; 32. haftada = 488,27; 33. haftada = 492,93 ve 34. haftada = 532,59'dır. Bu dönem (30.07.2018-26.08.2018) Ege bölgesinin *en yüksek oda fiyatlarına* sahip dönemidir. Hafta içi ve hafta sonu oda fiyatlarının anlamlı bir farklılık gösterip göstermediğini araştırmak için veri üzerinde Mann-Whitney Testi yapılmıştır. Ege bölgesinde hafta sonu fiyatları (Sırl.Ort: 95,81) hafta içi fiyatlara (Sırl.Ort: 90,45) oranla daha yüksek olmasına rağmen, test sonucuna göre hafta sonu oda fiyatları ile hafta içi oda fiyatları arasında anlamlı bir farklılık yoktur ($U = 3243$, $p>0.05$).

Son olarak fiyatların zamana göre değişimini ortaya koymak bir eğitim verisi kullanılarak tahminde bulunmak için makine öğrenmesi ve ARIMA modeli kullanılmıştır. Hata oranının en düşük olduğu modeli bulmak için *auto arima* yöntemi kullanılmış ve bunun sonucunda *ARIMA(0,1,0)* modelinin veri üzerinde en yüksek performansı sergilediği görülmüştür.

Araştırma kıyı otelleri üzerinde yapılmıştır. Kıyı otelleri deniz mevsimine bağlı olarak talep gören tesislerdir. Bu nedenle oda fiyatları şehir otellerinden farklı dağılım göstermektedir (Karyopouli & Koutra, 2013). Benzer çalışmaların şehir otelleri üzerinde de yapılması yazına ek katkı sağlayacaktır. Sadece Ege bölgesinde yapılan bu çalışmanın Türkiye'nin tamamına uygulanması turizm endüstrisi açısından değerli sonuçlar verecektir.

SONUÇ

Otel oda fiyatlarını diğer pek çok mal ve hizmet gibi zamansal fiyat değişikliğine uğramaktadır. Fakat turizm ürününde, satış sezonunun kısa oluşu (Goeldner & Ritchie, 2006, s. 358) bu değişimin daha da esnek olmasına neden olmaktadır. Daha net ifadeyle, ürün yılın sadece bir döneminde üretilmekte ve satılmaktadır. Bunun da en büyük nedenlerinden ikisi turizm sektörünün doğal ve çevresel güzelliklere büyük oranda bağımlı olması ve insanların tatile çıkacakları dönemin çoğu zaman belirli olmasıdır (Holloway & Taylor, 2006). Otel oda fiyatlarının zamansal olarak esnek olduğu bilinmesine rağmen (Upton, 1999) bu alanda yapılan çalışma sayısı çok değildir (Lim & Chan, 2011). Bunun da en büyük nedenlerinden biri veri toplamadaki zorluklardır.

Bu araştırma, veri toplama konusunda yeni yöntemlerden biri olan web madenciliği yöntemini kullanarak (Gorunescu, 2011), Ege bölgesindeki 5 ilden toplam 675 otelin 15 seyahat acentesi tarafından verilmiş 184 günlük fiyat verisini sağlamıştır. Gerek büyük veriyi elde etmede web madenciliğinin kullanılması gerekse otel oda fiyatlarının 675 oteli içine alan tam bir bölgeyi kapsayacak biçimde yapılmasının yazına katkı sağlayacağı düşünülmektedir.

Web madenciliği alternatif bir veri elde etme yöntemi olarak turizm konusunda toplanması çok zor olan verilerin elde edilmesini sağlamıştır. Python, R gibi programlama dilleri ve bazı paket programlar ile yapılabilen bu yöntem pek çok alanda değerli verilerin toplanmasına imkân vermektedir. Bu yöntem başka bir çalışmada metin madenciliği ile turizm gündemini bulmak için kullanılmış, önemli bulgular elde edilmiştir (Ağca, Aksoy, & Özdemir, 2018). Bu yöntemin bilimsel araştırma verisi elde etmede yaygınlaşacağı düşünülmektedir.

Sezonluk fiyat dağılımının durağanlığa yakın ve normal dağılıma benzediği görülmüştür. Sezon içerisinde yer alan *dini bayramların* dolayısıyla *resmî tatillerin* fiyatların eğimini yukarı yönde değiştirdiği görülmüştür. Bu durum, kıyı otellerinde fiyatı *yerli turizm talebinin* de etkilediği sonucunu verebilmektedir. Araştırmada sadece bir sezonluk veri kullanılmıştır. Birden çok sezona ait verinin kullanılması, daha net sonuçların ortaya çıkmasına neden olacaktır.

Bölgenin hafta sonu fiyatları ortalama olarak hafta içinden daha yüksek olmasına rağmen bu fark istatistiki olarak anlamlı bulunmamıştır. Hafta sonu ile hafta içi fiyat farklılıklarının en yüksek olduğu yer Çanakkale'dir.

Bulgular hem sektör yönetici ve yatırımcıları hem de turistler açısından değerlendirilebilir. Sektör yöneticileri açısından bakıldığında; bölgedeki fiyat değişimlerini göz önünde bulundurarak daha iyi getiri yönetimi planlaması yapmak mümkündür. Tatil planı yapanlar açısından ise hangi dönemlerin fiyat açısından daha uygun olduğu görülebilir. Her ne kadar 184 günlük fiyatları içerse de bu sadece bir sezona ait veri seti ile yapılan bir çalışmadır. Birkaç sezon arka arkaya yapılacak çalışmalar daha doğru bulgular ortaya koyacaktır. Burada turizm sektörünün talep konusunda oldukça kırılgan yapısını da göz önünde bulundurmak gerekmektedir.

Bir bölgedeki otellerin oda fiyatlarının zamansal değişimleri konusu oldukça geniş kapsamlı ele alınması gereken bir konudur. Bu çalışmada, mevcut veri kullanılarak, bazı genel bulgulara yer verilmiştir. Ayrıca araştırma verisi Ege bölgesindeki kıyı otelleri ile sınırlandırılmıştır. Başka bölgelerde yapılacak çalışmalar, kıyı otelleri ve kıyıda olmayan otellerin fiyat farklılıkları, otellerin turistik bir bölgede olup olması gibi kriterler göz önünde bulundurarak yapılacak çalışmalar; bu çalışmanın da yer alacağı, otel oda fiyatlarının zamansal analizleri hakkında daha detaylı bilgi sunacaktır.

ALTERNATIVE DATA COLLECTION METHOD WEB MINING: SEASONAL ANALYSIS OF HOTEL ROOM PRICE

1. INTRODUCTION

Prices of commercial products vary depending on many variables. One of the factors affecting the prices of the products is time. Time is also an important variable for the price of tourist products. Tourism product is a composite product consisting of goods and services produced by touristic enterprises for tourists and supported by the natural and cultural beauties of the region. One of the most important parts of the touristic product is the hotel service and accommodation. Hotel room rates vary considerably during the tourist season. Knowing this temporal change in prices is important for both customers and hotel managers.

Hotel room rates are usually high during the so-called high season and low during the start and end of the season (Weiermair, Pechlaner, & Bieger, 2006, s. 131). In terms of profitability, it is desirable that the hotel is always at the same level of occupancy. Accommodation businesses are trying to maximize profitability by keeping prices high during high season to offset revenue. In a way, they apply one of the most basic elements of the economy and increase prices as demand increases.

Turkey's Mediterranean and Aegean coasts in particular are in great demand by both foreign and domestic tourists in terms of coastal tourism. In addition to the natural and cultural elements of the region, the long summer period and sea water temperature also have an effect. In this study, the change of hotel room rates over time was investigated. In this context, the change of hotel room prices by months and weeks, whether weekday weekend prices differed, was analyzed by statistical methods. The research was carried out on the prices of coastal hotels in the Aegean Region.

It is difficult and time-taken to obtain data by classical methods for this type of research, so web mining (scraping) method was used to obtain the data. Web scraping/mining is the receipt of the requested data from the data found scattered on the websites. For this purpose, programming languages such as Python and R or web scraping software are often used (Ağca, 2019, s. 55). In this study, we obtained a large amount of data that is difficult to collect using web scraping method. As of a tourist season, price data was obtained from 675 hotels in 5 provinces in the Aegean Region, consisting of 72,104 lines for 184 days, taken from 15 different travel agencies.

2. METHODOLOGY

After determining the website where the data will be obtained, a web scraping script is written using the R programming language. *BeautifulSoup* and *pandas* libraries are used. Hotel room rates were received on March 20, 2018. Prices charged are two people per night. Prices are from 26 March 2018 to 29 September 2018. Only hotel and resort prices are taken. Aparthotel, home/apartment, hostel prices are not included in the data set. Hotels in the data set are coastal hotels in the Aegean region.

The data was first cleaned and preprocessed. Kolmogorov-Smirnov and Shapiro-Wilk were tested for the normal distribution of the data. Parametric tests were performed for normal distributed data. Non-parametric tests were applied for variables that were not normally distributed or categorical (Field, 2009, s. 179-180). There is no single hotel room rate on the site. Therefore, the algorithm used is based on the lowest price among these prices.

3. RESULTS

According to the findings, between 26.03.2018-30.09.2018, the most hotels in the Aegean region are located in Mugla (284), Izmir (201) and Aydın (95), respectively. On the basis of cities, Bodrum has a share of approximately 16.60% of the total number of hotels in the Aegean region with 112 hotels. The number of hotels in the cities is Alacati (84), Cesme (75) and Kusadasi (64).

Customer reviews are a variable that can give an idea of the preferred rate of hotel and place (Memarzadeh, Blum, & Adams, 2015; Keskinçiliç, Ağca, & Karaman, 2016). When the scores given by the guests to the facilities were evaluated, it was determined that Alacati (Avg: 8.4) and İzmir (Avg: 8.13) scored higher than other regions.

When the results are examined, it is seen that the months with the highest prices are July, August and September. Kruskal-Wallis test was performed on the data to see if hotel room rates differed significantly by month. March is not included in this analysis and data because it contains only 6 days of price data. Average of this data = 375.95; standard deviation = 87.39. According to the results obtained, room rates differ significantly by months ($X^2=144.43$; $p<0.05$). When looking at the sequential average values, the highest values are seen in August (160.90) and July (140.16).

The data was analyzed to evaluate prices on a weekly basis and to determine whether there was a significant difference between prices by week. As a result of the analysis, average room rates varied significantly for the Aegean region by week ($X^2=164.58$; $p<0.05$). Accordingly, the average values of the ranking are the highest 31, 32, 33 and 34 weeks. Average room rates for these weeks are = 481.02 in week 31; Week 32 = 488.27; Week 33 = 492.93 and week 34 = 532.59. Between 30.07.2018 and 26.08.2018 period is the highest room rates in the Aegean region.

A Mann-Whitney U test was performed to determine if there was a significant difference between weekday and weekend prices. The findings show that there is no significant difference between weekend room rates and weekday room rates ($U = 3.243$, $p>0.05$).

The research data was analyzed with ACF and PACF graphs and Dickey-Fuller test. It has been observed that it does not see stationarity. Therefore, ARIMA model is preferred. In order to find the most suitable ARIMA model, *auto_arima* method under *pmdarima* library in Python language was used. As a result of the analysis, it was determined that the most suitable model was ARIMA(0,1,0).

4. DISCUSSION

In this research, data was obtained using web mining and tested by various statistical methods to determine the seasonal change of hotel room prices in the Aegean region. The change in hotel room rates is determined by the ARIMA model using machine learning. Prices vary significantly by months ($X^2=144.43$; $p<0.05$) and weeks ($X^2=164.58$; $p<0.05$), but not between weekdays and weekends.

In this research, Web mining was used to obtain data that is difficult to collect. This method, which can be done with programming languages such as Python and R and some package programs, allows valuable data to be collected in many areas.

CONCLUSION

The findings can be evaluated for both industry managers, investors and tourists. From the point of view of industry managers; it is possible to plan for better income management by considering price changes in the region. The findings can be evaluated for both industry managers, investors and tourists. From the point of view of industry managers; it is possible to plan for better income management by considering price changes in the region. For those who plan a holiday, it can be seen which periods are more affordable in terms of price. Although it includes 184-day prices, this is only a study with a data set for one season. Studies for several seasons will produce more accurate findings. Here, it is also necessary to consider the very fragile structure of the tourism sector in terms of demand.

The issue of seasonal changes in room rates of hotels in a region is an issue that needs to be addressed quite extensively. In this research, some general findings are given using existing data. In addition, research data is limited to coastal hotels in the Aegean region. Studies to be carried out in other regions, considering criteria such as price differences of coastal hotels and non-coastal hotels, whether the hotels are in a touristic area; this study will provide more detailed information about temporal analysis of hotel room rates.

KAYNAKÇA

- Ağca, Y. (2019). *Çevrimiçi Seyahat Acentelerinde Oda Fiyatlarına Etki Eden Faktörlerin Araştırılması* (Yayınlanmamış Doktora Tezi). Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.
- Ağca, Y., Aksoy, H., & Özdemir, A. (2018). Web Madenciliği Yönteminin Geçmiş Gündemi Bulmak İçin Kullanımı: Türkiye'deki Turizm Gündemi Örneği. *The Journal of Social Science*, 30(30), 344-358. doi:http://dx.doi.org/10.16990/SOBIDER.4562
- Baumol, W. J., & Blinder, A. S. (2016). *Economics: Principles and Policy* (13 b.). Boston, US: Cengage Learning.
- Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control* (4 b.). Hoboken, US.: John Wiley & Sons, Inc.
- Bull, A. O. (1994). Pricing A Motel's Location. *International Journal of Contemporary Hospitality Management*, 6(6), 10-15.
- Carvell, S. A., & Herrin, W. E. (1990). Pricing in the Hospitality Industry: An Implicit Market Approach. *Hospitality Review*, 27-37. <http://scholarship.sha.cornell.edu/articles/194/> adresinden alındı
- Coenders, G., Espinet, J. M., & Saez, M. (2003). Predicting Random Level and Seasonality of Hotel Prices: A Latent Growth Curve Approach. *Tourism Analysis*, 8, 1-17.
- Dwyer, L., Forsyth, P., & Dwyer, W. (2010). *Tourism Economics and Policy*. Bristol, UK: Channel View Publications.
- Erden, C. (2020, 05 06). *Zaman Serisi Tahminleri ve ARIMA Modelleri*. Canererden.com: <http://canererden.com/2020/05/06/zaman-serisi-tahminleri-ve-arima-modelleri> adresinden alındı
- Field, A. (2009). *Discovering Statistics Using SPSS* (3 b.). London: SAGE Publications Ltd.
- Frechtling, D. (2011). *Forecasting Tourism Demand: Methods and Strategies*. New York, US: Routledge.
- George, D., & Mallery, P. (2010). *SPSS for Windows step by step : a simple guide and reference, 17.0 update* (10 b.). Boston: Allyn & Bacon.
- Goeldner, C. R., & Ritchie, R. B. (2006). *Tourism; Principles, Practices, Philosophies* (10 b.). New Jersey: John Wiley & Sons, Inc.
- Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques* (Cilt 12). Berlin: Springer-Verlag. doi:10.1007/978-3-642-19721-5
- Gupta, G. (2004). *Macroeconomics: Theory and Applications* (2 b.). New Delhi: Tata McGraw-Hill Publishing.
- Holloway, J. C., & Taylor, N. (2006). *The Business of Tourism* (7 b.). Harlow: Financial Times Prentice Hall.
- Kara, O. (2020, 02 03). *ARIMA ile Bitcoin Fiyat Tahmini*. Medium.com: <https://medium.com/icrypex/arima-ile-bitcoin-fiyat-tahmini-510dd7a94d97> adresinden alındı
- Karyopoulou, S., & Koutra, C. (2013). Mature Coastal Mediterranean Destinations: Mitigating Seasonality. C. Costa, E. Panyik, & D. Buhalis (Dü) içinde, *Trends in European Tourism Planning and Organisation*. Ontario, Canada: Channel View Publications.
- Keskinkılıç, M., Ağca, Y., & Karaman, E. (2016). İnternet ve Bilgi Sistemleri Kullanımının Turizm Dağıtım Kanallarına Etkisi Üzerine Bir Uygulama. *İşletme Araştırmaları Dergisi*, 8(4), 445-472. doi:10.20491/isarder.2016.227

- Koushik, D., Higbie, J. A., & Eister, C. (2011). Retail Price Optimization at InterContinental Hotels Group. *Interfaces*, 42(1), 45-57. <http://www.jstor.org/stable/41472728> adresinden alındı
- Krishna, V. B., Iyer, R. K., & Sanders, W. H. (2015). ARIMA-Based Modeling and Validation of Consumption Readings in Power Grids. E. Rome, M. Theocharidou, & S. Wolthusen (Dü.), *10th International Conference, CRITIS 2015* içinde (s. 199-210). Berlin, Germany: Springer. <https://link.springer.com/book/10.1007/978-3-319-33331-1> adresinden alındı
- Lim, C., & Chan, F. (2011). An econometric analysis of hotel and motel room nights in New Zealand with stochastic seasonality. *International Journal of Revenue Management*, 5(1), 63-83.
- Maklin, C. (2019, 05 25). *ARIMA Model Python Example — Time Series Forecasting*. Towards Data Science: <https://towardsdatascience.com/machine-learning-part-19-time-series-and-autoregressive-integrated-moving-average-model-arima-c1005347b0d7> adresinden alındı
- Memarzadeh, F., Blum, S. C., & Adams, C. (2015). The impact of positive and negative e-comments on business travelers' intention to purchase a hotel room. *Journal of Hospitality and Tourism Technology*, 6(3), 258-270. doi:10.1108/JHTT-09-2014-0049
- Nau, R. (2020). *Statistical forecasting: notes on regression and time series analysis*. duke.edu: <https://people.duke.edu/~rnau/411arim.htm> adresinden alındı
- Nikolopoulos, K. I., & Thomakos, D. D. (2019). *Forecasting With The Theta Method: Theory and Applications*. Hoboken, US: Wiley.
- (2018). *OECD Tourism Trends and Policies*. Paris, France: OECD Publishing.
- Papatheodorou, A. (2002). Exploring Competitiveness in Mediterranean Resorts. *Tourism Economics*, 8(2), 133-150.
- Pawlicz, A., & Napierala, T. (2017). The Determinants of Hotel Room Rates: An Analysis of The Hotel Industry in Warsaw, Poland. *International Journal of Contemporary Hospitality Management*, 29(1), 571-588.
- Prabhakaran, S. (2020). *ARIMA Model – Complete Guide to Time Series Forecasting in Python*. machinelearningplus.com: <https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/> adresinden alındı
- Raya, J. M. (2010). The effect of time on hotel pricing strategy. *Applied Economics Letters*, 1-5. doi:10.1080/13504851.2010.532091
- Saló, A., Garriga, A., Rigall-I-Torrent, R., Vila, M., & Sayeras, J. M. (2012). Differences in seasonal price patterns among second home rentals and hotels: empirical evidence and practical implications. *Tourism Economics*, 18(4), 731-747. doi:10.5367/te.2012.0141
- Sharma, A. (2019, 06 24). *Moving Averages in Pandas*. Data Camp: <https://www.datacamp.com/community/tutorials/moving-averages-in-pandas> adresinden alındı
- Taylor, P. (1995). Measuring Changes in the Relative Competitiveness of Package Tour Destinations. *Tourism Economics*, 1(2), 169-182.
- Thrane, C. (2005). Hedonic Price Models and Sun-and-Beach Package Tours: The Norwegian Case. *Journal of Travel Research*, 43(3), 302-308. <https://search.proquest.com/docview/217444429?accountid=15959> adresinden alındı
- Tribe, J. (2016). *The Economics of Recreation, Leisure and Tourism* (5 b.). New York, US: Routledge, Taylor & Francis Group.

- Upton, G. (1999, 05 17). A Seasonal Opportunity: BUSINESS TRAVEL HOTEL BARGAINS: If You Go Away In Summer There Are Good Rates To Be Found, Says Gillian Upton: [London Edition]. *Financial Times*. 04 07, 2018 tarihinde <https://search.proquest.com/docview/248682242?accountid=15959> adresinden alındı
- Valipour, M., Ebrahim Banihabib, M., & Behbahani, S. M. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology*, 476, 433-441. doi:<https://doi.org/10.1016/j.jhydrol.2012.11.017>
- Wang, X., Sun, J., & Wen, H. (2019). Tourism seasonality, online user rating and hotel price: A quantitative approach based on the hedonic price model. *International Journal of Hospitality Management*, 79, 140-147.
- Weiermair, K., Pechlaner, H., & Bieger, T. (Dü). (2006). *Time Shift, Leisure and Tourism: Impacts of Time Allocation on Successful Products and Services*. Berlin, Germany: Erich Schmidt Verlag GmbH.

KATKI ORANI / CONTRIBUTION RATE	AÇIKLAMA / EXPLANATION	KATKIDA BULUNANLAR / CONTRIBUTORS
Fikir veya Kavram / <i>Idea or Notion</i>	Araştırma hipotezini veya fikrini oluşturmak / <i>Form the research hypothesis or idea</i>	Yılmaz AĞCA
Tasarım / <i>Design</i>	Yöntemi, ölçeği ve deseni tasarlamak / <i>Designing method, scale and pattern</i>	Yılmaz AĞCA
Veri Toplama ve İşleme / <i>Data Collecting and Processing</i>	Verileri toplamak, düzenlenmek ve raporlamak / <i>Collecting, organizing and reporting data</i>	Yılmaz AĞCA
Tartışma ve Yorum / <i>Discussion and Interpretation</i>	Bulguların değerlendirilmesinde ve sonuçlandırılmasında sorumluluk almak / <i>Taking responsibility in evaluating and finalizing the findings</i>	Yılmaz AĞCA
Literatür Taraması / <i>Literature Review</i>	Çalışma için gerekli literatürü taramak / <i>Review the literature required for the study</i>	Yılmaz AĞCA