



Creating Competitive Advantage from Non Structured Data: A Text Mining Approach in Turkish Hospitality Sector

Dr. Taşkın DİRSEHAN*

* Marmara Üniversitesi, İstanbul, Türkiye*, İşletme Fakültesi, İngilizce İşletme Bölümü, Üretim Yönetimi ve Pazarlama ABD, E-mail: taskin.dirsehan@marmara.edu.tr,

Copyright © 2016 Taşkın DİRSEHAN. This is an open access article distributed under the Eurasian Academy of Sciences License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

In recent years, data are exploded and they become more unstructured as a result of increasing web pages, e-mails, social media, photos and video contents. Traditional databases are not enough now to manage such data. However, organizations should manage them to gain competitive advantages. In other terms, hospitality managers should focus on satisfying customers by listening to them. In today's information age, customers can be heard via their comments on travel websites that provide "big data." As unstructured data are increased now, companies face with 'big data' and data mining tools should be applied to transform mass data into information and even into knowledge. The main purpose of this paper is to present a text-mining procedure to be used in marketing strategies to discover competitive elements in the hospitality sector from different locations in Turkey. Text mining is a powerful tool to analyze texts as main unstructured data. They are mostly available in the tourism sector in the form of tourists' satisfaction and dissatisfaction stories. In this study, firstly, the mostly repeated words in the positive and negative reviews are revealed and then, the relationship of these words with the total points of the facility is investigated and finally, the revealed characteristics are discussed with Herzberg's two-factor motivation theory. They are grouped into two categories: (1) motivators: words correlated with total review scores only in the positive reviews; and (2) hygiene factors: words negatively correlated with total review scores only in the negative reviews. In addition, the motivation-related attributes are discussed in terms of their differences between different locations in Turkey.

Keywords: Tourism Marketing, Data Mining, Text Mining

JEL: M31, C80

Yapısal Olmayan Veriden Rekabet Avantajı Yaratma: Türkiye Konaklama Sektörü'nde Metin Madenciliği Yaklaşımı

ÖZET

Son yıllarda, veri patlaması yaşanmış ve web sayfaları, e-postalar, sosyal medya, fotoğraflar ve video içeriklerinin sayısının artışıyla git gide yapısal olmayan veri miktarı artmıştır. Geleneksel veritabanları, artık bu verileri yönetmekte yeterli olamamaktadır. Ancak firmalar, rekabet avantajı kazanmak için bu verileri yönetebilmelidir. Bir diğer ifadeyle, konaklama sektöründeki yöneticiler, tüketicileri dinleyerek onları memnun etmeye odaklanmalıdır.



Günümüz bilgi çağında tüketiciler, web sayfalarına bıraktıkları yorumlar vasıtasıyla dinlenebilir, bu yorumlar da büyük veriyi oluşturur. Günümüzde yapısal olmayan veri artıkça firmalar, artık büyük veriyle daha çok karşılaşmaktadır ve kitlesel veriyi bilgiye hatta malumata dönüştürebilen veri madenciliği tekniklerini uygulamalıdır. Bu çalışmanın temel amacı, Türkiye'nin farklı bölgelerindeki konaklama sektöründeki rekabet unsurlarını bulmak için, pazarlama stratejilerinde kullanılabilecek metin madenciliği yöntemini sunmaktır. Metin madenciliği, yapısal olmayan temel veri kaynağı olan metinlerin analizinde kullanılabilecek güçlü bir araçtır. Bu metinler, daha çok turizm sektöründe turistlerin olumlu ve olumsuz yorumları olarak bulunurlar. Bu çalışmada önce olumlu ve olumsuz yorumlarda en sık tekrarlanan kelimeler bulunmuş, ardından bu kelimelerin tüketicilerin konaklama tesisine verdiği genel puanla ilişkisi incelenmiş ve ortaya çıkan özellikler, Herzberg'in çift-faktör motivasyon teorisiyle paralel olarak tartışılmıştır. Bu özellikler iki kategoride gruplanmıştır: (1) motive ediciler: olumlu yorumlarda toplam yorum sonuçlarıyla korelasyonu olan kelimeler; ve (2) hijyen faktörleri: sadece olumsuz yorumlarda toplam yorum sonuçlarıyla korelasyonu olan kelimeler. Bunlara ek olarak, motivasyon temelli özellikler, Türkiye'nin çeşitli bölgeleri arasındaki farklara dayalı olarak tartışılmıştır.

Anahtar Kelimeler: Turizm Pazarlaması, Veri Madenciliği, Metin Madenciliği

1. GİRİŞ

Firmalar için rekabet avantajı, insanlık tarihiyle bağlantılıdır. Avcı-toplayıcı döneminde yaşayan insanların birbirlerine karşı rekabet avantajları kas gücüydü. Tarım toplumunda, verimli topraklar ve onları ekme yöntemleri rekabet avantajıyken, sanayi toplumunda tarım ve toprağın yerini makineler almıştır. Kitlesel üretime geçilen bu dönemde makine sayısı ve üretim miktarı, firmalar için rekabet avantajı yaratmıştır. Günümüz bilgi toplumunda ise, rekabet avantajı artık veriye dayanmaktadır. Fazla veri toplama, her zaman rekabet avantajına doğrudan katkıda bulunmaz. Bu verinin işlenip, faydalı bilgilerin elde edilmesi, bu bilgilerin de yönetici tecrübelerine de bağlı olarak doğru stratejiler üretilmesinde kullanılması, günümüz rekabet yapısında önemli bir avantaj sağlamaktadır.

Diğer taraftan, özellikle teknolojinin gelişmesiyle ve tüketicinin içerik oluşturduğu yorum, tecrübe içerikli web siteleri ve sosyal medya gibi platformlar, metin tabanlı veri miktarını önemli ölçüde artırmıştır. Bu veriler bir yığın olarak firmaların önünde durmaktadır ve nasıl ki değerli madenlere ulaşmak için madencilik çalışmaları yapılıyorsa, firma için de değerli bilgiye ulaşmak için bu veri yığının güncel tekniklerle işlenebiliyor olması gerekir. Günümüzde de firmalar tüketiciye değer sunabilmek için ve sunduğu değerler bakımından diğer firmalardan farklılaşabilmek için, bu verileri yorumlayabilmelidirler. Çünkü firmalar, iki tip rekabet avantajından yararlanabilirler (Porter, 1985): Maliyet avantajı ve farklılaşma. Maliyet avantajının temeli, müşteriye aktarılan değerın daha ucuz bir fiyatla sunulmasıdır, ancak bu durumda rakiplerin aynı stratejiyi kopyalaması kolaydır. Diğer taraftan farklılaşma, müşteriye, başka firmaların sunmadığı benzersiz bir değer yaratılmasıyla mümkündür. Bu değere ulaşmak ise, veriden anlamlı bilgiye ulaşılmasıyla mümkündür. Bu işlem için gerekli olan en önemli unsur ise büyük veridir. Çünkü geleneksel veritabanı sistemlerinde işlenemeyen bu büyük veri, anlamlı bilgi üretmeye ve dolayısıyla rekabet avantajı yaratmayı amaçlar (McAfee and Brynjolfsson, 2012). Turizm sektöründe de eWOM maçı, çok sayıda



yorum bulmak mümkündür. Bu metin formatındaki verilerden işletmeler için rekabet avantajı sağlayacak bilgiyi üretmek için metin madenciliği kullanılabilir. Bu çalışmada da, müşteri yorumlarını olumlu ve olumsuz olarak toplayan “www.booking.com” internet sitesi incelenmiş ve sonuçlar, Herzberg’in çift faktör motivasyon teorisine göre yorumlanmıştır.

2. YAZIN TARAMASI

Son yıllarda internetin kullanımının artması, makineyle sensörler yardımıyla toplanan verinin artması, cep telefonlarının GPS sinyalleri, mesajlar, güncellemeler, sosyal medyaya yüklenen resimler, eskiye kıyasla günümüzde, veri miktarının ne kadar arttığını, ne kadar çeşitlilik gösterdiğini ve günlük yaşamda ne kadar hızlı değiştiğini göstermektedir. (McAfee and Brynjolfsson, 2012).

Verinin bu özellikleri (miktarı, hızı ve çeşitliliği), “Büyük Veri” devrimi olarak düşünülebilir (Erevelles vd., 2015). Büyük Veri, işletmeler için artık yeni bir sermaye türü olarak görülmektedir.

Büyük Veri, rekabet avantajı sağlamak için turizm pazarlaması alanında da dikkat çekmektedir (Dolnicar ve Ring, 2014).

Büyük Veri, geleneksel veritabanı yapılarıyla toplanıp işlenemeyecek kadar büyük olan verileri ifade eder. Bu verilerin içerisinde bulunan gizli ilişkileri ortaya çıkarmak, sınıflandırmalar ve tahminler yapmak için veri madenciliği teknikleri kullanılır (Laudon and Laudon, 2014).

Diğer taraftan; sosyal medya, e-postalar ve müşteri yorumları, metin verilerini oluşturur. Metin verileri, yapısal olmayan verilerdir ve bir firmandaki kullınılı bilginin %80’ini oluşturduğu belirtilir (Laudon and Laudon, 2014). Bu durumda, inovatif bir metodoloji olarak, metin madenciliği teknikleri; çeşitli istatistik teknikleriyle birlikte kullanılabilir (Gassiot and Coromina, 2013).

Metin madenciliği, içeriğindeki zengin metin verilerinden ötürü, ziyaretçi yorumlarına uygulanabilir. Örneğin, Xiang vd. (2015), ABD oteli tercihlerini incelemek için Expedia.com internet sitesindeki yorumları incelemişlerdir.

Bu çalışmada veri kaynağı olarak, olumlu ve olumsuz yorumlarını ayrı ayrı topladığı için ve bu ayrım Herzberg’in çift faktörlü motivasyon teorisine göre yorumlanabileceği için “booking.com” internet sitesi tercih edilmiştir. Bu teori; motivasyon faktörlerini iki sınıfa ayırmaktadır (Herzberg vd., 1962; Herzberg, 1975):

a. Motivasyonel faktörler (var olduklarında, çalışanlar memnundur, eksiklikleri mutlaka memnuniyetsizlik yaratmayabilir.)

b. Hijyen faktörleri (Eksikliklerinde, çalışanlar memnuniyetsiz olurlar, mevcudiyetlerinde ise çalışanlar mutlaka memnun olmayabilirler).

Turizm, birçok gelişmekte olan ülkelere büyüme fırsatı sunar. Özellikle Akdeniz ülkeleri için önemli bir sektördür (Kushi and Caca, 2012), çünkü gelişmiş ülkeler sahip oldukları avantajlar sayesinde kalkınma düzeylerini gerçekleştirmiş, ancak gelişmekte olan ülkeler, kalkınma ve az gelişmişlik sorunlarını yenmek için, sanayileşme olmadığı için turistik potansiyellerini değerlendirebilmektedir. Çoban ve Özcan (2013), 1963-2010 yılları arasında Türkiye’de turizm ve ekonomik büyüme ilişkisini incelemişler ve özellikle uzun dönemde



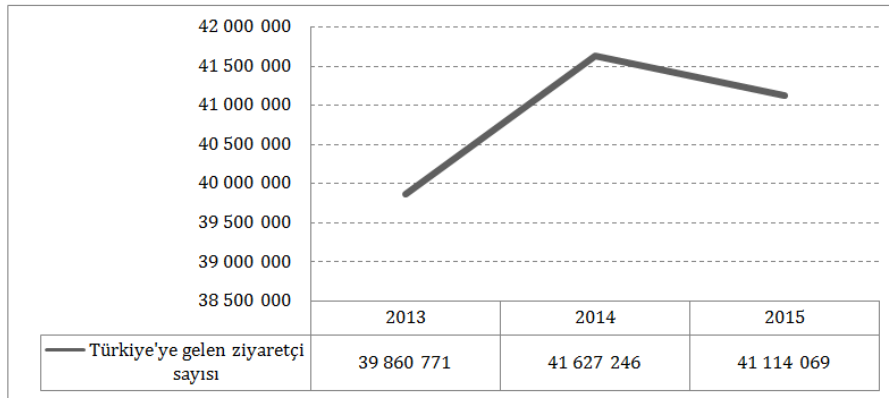
turizmin, ekonomik büyümenin önemli bir nedeni olduğu sonucunu bulmuşlardır. Dolayısıyla turizm, Türkiye'nin önemli büyüme kaynaklarından biri olup dış ticaret açığını kapatmada önemli rol oynamaktadır (Çoban ve Özcan, 2013). Bu yüzden genel olarak, Türkiye 'de konaklama işletmelerine yönelik yabancı turistlerin yorumlarına dayanılarak yapılan bu çalışma, hem işletmelere rekabet avantajı sağlamaları için bir metodoloji sunma amacı taşımakta, hem de konaklama sektöründe ziyaretçilerin sorunlarını ve beğenilerini ortaya çıkarmayı hedeflemektedir.

3. METODOLOJİ

Bu araştırmanın amacı, Türkiye turizm sektöründe, yabancı turistlerin yorumlarından faydalanarak konakladıkları işletmelere yönelik mikro ve makro düzeyde beğenilerini ve problemlerini metin madenciliği yaklaşımıyla ortaya çıkarmaktır.

Bu amaçla, "booking.com" internet sitesindeki yorumlar incelenmiştir. Destinasyonlar seçilirken Türkiye'deki en çok turist çeken bölgeler göz önünde bulundurulmuştur. Kültür ve Turizm Bakanlığı verilerine göre, Türkiye'ye gelen ziyaretçi sayısı, 2015 yılında, bir önceki yıla göre %1,23 azalarak 41.11.069 olmuştur (bkz. Şekil 1).

Şekil 1. Yıllar İtibariyle Türkiye'ye Gelen Ziyaretçi Sayısı



Kaynak: Kültür ve Turizm Bakanlığı verilerine göre yazar tarafından hazırlanmıştır.

2015 yılında Türkiye'yi ziyaret eden yabancı sayısına bakıldığında ise, en çok giriş yapılan sınır kapılarının bağlı olduğu ilk beş il, şu şekildedir (oranları için bkz. Şekil 2):

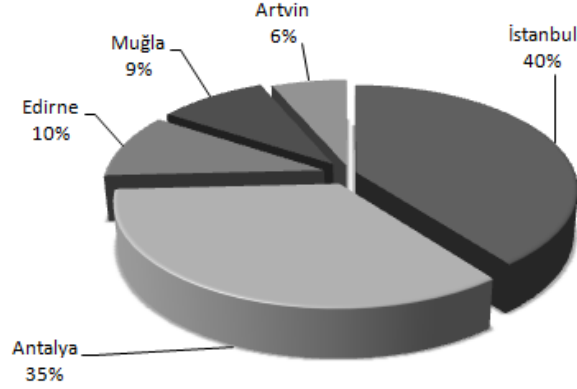
- 1- İstanbul (12.428.733)
- 2- Antalya (10.874.093)
- 3- Edirne (3.162.945)
- 4- Muğla (2.918.000)
- 5- Artvin (2.018.061)

Yurtdışı kaynaklı internet siteleri incelendiğinde, Euromonitor'un 2014 verilerine göre en çok ziyaret edilen 100 destinasyon listesine Türkiye'den ; 9. sıradan İstanbul, 11. Sıradan Antalya, 56. Sıradan Muğla, 60. Sıradan Edirne ve 100. Sıradan Artvin dahil olmuştur. Touropia web sitesi ise, Türkiye'de en çok turist çeken yerler listesinde 10'dan 1'e doğru sırasıyla;



Aspendos Tiyatrosu, Patara Sahili, Pamukkale, Bodrum Kalesi, Nemrut Dağı, Ölüdeniz, Topkapı Sarayı, Efes, Göreme ve Aya Sofya yer almaktadır (<http://www.touropia.com/tourist-attractions-in-turkey/>). Tripadvisor web sitesinde ise Türkiye’de en çok turist çeken yerler; İstanbul, Göreme, Ürgüp, Dalyan, Ölüdeniz ve Marmaris olarak sıralanmaktadır (<https://www.tripadvisor.com.tr/Tourism-g293969-Turkey-Vacations.html>).

Şekil 2. 2015 yılında Türkiye’ye en çok giriş yapılan sınır kapılarının bağlı olduğu iller



Kaynak: Kültür ve Turizm Bakanlığı sınır giriş-çıkış verilerine (<http://yigm.kulturturizm.gov.tr/TR,9854/sinir-giris-cikis-istatistikleri.html>) göre yazar tarafından hazırlanmıştır.

Bu listeler göz önünde bulundurularak, bu çalışmada aşağıdaki bölgelerde yer alan konaklama işletmelerine yönelik yorumlar incelenmiştir:

1. İstanbul
2. Antalya
3. Marmaris
4. Mardin
5. Side
6. Efes
7. Bodrum
8. Dalyan
9. Kuşadası
10. Trabzon
11. Ölüdeniz
12. Ürgüp
13. Edirne
14. Fethiye

Dolayısıyla seçilen bölgeler, Türkiye’nin en çok turist çektiği Ege ve Akdeniz bölgelerinde yoğunluk göstermektedir. Her bölgeden, 100’er yorum toplanmış, dolayısıyla toplamda 1400 yorum analiz edilmiştir. Yorumların toplanırken bazı kriterlere dikkat edilmiştir:

- 1- Hem olumlu hem olumsuz yorum yazan ziyaretçilerin bilgileri dikkate alınmıştır (dolayısıyla 1400 olumlu, 1400 olumsuz yorum toplanmıştır),



2- Yönlendirmeyi engellemek adına, aynı otele ilişkin otuzdan fazla yorum toplanmamıştır,
3- Varyansı artırmak için hem yüksek puanlı hem de düşük puanlı oteller dikkate alınmıştır (Otellerin puanları 2,5 ile 9,6 arasında olup ortancası 8,3'tür).

Diğer taraftan, başka dillerde yazılmış yorumlar; amacın dilbilgisinden ziyade anahtar kelimelerin toplanması olduğu için çevrimiçi tercüme araçları kullanılarak İngilizce'ye çevrilmiştir.

4.ARAŞTIRMANIN BULGULARI

Yorumlarına bakılan 1400 ziyaretçiden, 265'i Türk, 181'i Birleşik Krallık, 72'si Suudi Arabistan, 63'ü ABD, 49'u Avustralya, 25'i Kuveyt, 24'ü Rus, 26'sı Kanada ve kalanlar da diğer ülke vatandaşlarıdır. Çalışmanın adımları şu şekilde özetlenebilir:

(1. adım) Ziyaretçilerin Türkiye'deki konaklama deneyimlerine yönelik olumlu ve olumsuz yorumları ile ziyaretçilerin (adı, soyadı, yaşı, milliyeti, vb.) ve kaldıkları tesislerin özelliklerine göre (yıldız sayısı, genel değerlendirme puanı, vb.), veri madenciliği sürecine uygun olarak önce bir veri ambarı olarak hazırlanmıştır.

(2. adım) Eldeki veri ambarından kullanılacak kısımlar, (örneğin bu çalışmada; bölgelerin isimleri, bölgedeki konaklama tesislerine ziyaretçilerin verdikleri değerlendirme puanları, ziyaretçilerin olumlu ve olumsuz yorumları), bir veri çarşısı olarak ayrıştırılmıştır.

(3. adım) Metin madenciliğiyle, önce incelenen yorumlarda en çok tekrar eden kelimeler incelenmiştir. Olumlu yorumlarda en sık tekrarlanan kelimeler (100 ve daha fazla defa tekrarlananlar) Tablo 1'de, olumsuz yorumlarda en sık tekrarlanan kelimeler (100 ve daha fazla defa tekrarlananlar) ise Tablo 2'de listelenmiştir. Ayrıca, olumlu yorumlardaki kelimelerin, rekabet avantajını geliştirebilme düzeylerine göre de bir sınıflandırma önerilmiştir. Örneğin,

(a)genel değerlendirme konusuyla ilgili olan kelimeler (oteller veya odalar) genel olarak yöneticilere bir sorun olup olmadığı hakkında bilgi verebilir.

(b)Diğer taraftan müşterileri memnun eden özellikler, kontrol edilmesi güç ve görece daha kolay olanlar şeklinde iki sınıfta toplanabilir:

(i)Kontrol edilmesi güç olan faktörler (konum, havuz ve manzara), genellikle otelin ya da konaklama tesisinin ilk faaliyete geçtiği zaman verilmiş kararlardır. Dolayısıyla, müşteriye memnun etmek için örneğin konumun değiştirilmesi kolay değildir.

(ii) Kontrol edilmesi daha kolay olan faktörler ise, yöneticilerin alacağı kararlarla rekabet avantajı dağılayabilir. Örneğin, personel, ziyaretçiler için bir memnuniyet faktörü ise, personelin hizmetinin iyileştirilmesi için yöneticiler, personellere eğitim verebilirler ya da örneğin kahvaltayı veya yemek çeşitliğini değiştirebilirler.



Diğer taraftan, olumsuz yorumlarda tekrar eden kelime sayısı bu kadar fazla olmadığı için bir sınıflandırma önerilmemiştir.

Tablo 1. Olumlu Yorumlarda En Sık Tekrar Eden Kelimelerin Listesi

Önerilen Sınıflandırma		Kelime	Olumlu Yorumlarda Tekrarlanma Sayısı	Olumsuz Yorumlarda Tekrarlanma Sayısı
Genel Değerlendirme Konusu		Otel (hotel)	396	270
		Oda (room)	404	474
Niteleyici Sıfatlar		Güzel (good)	406	69
		Muhteşem (Great)	217	20
		Hoş (nice)	260	20
		Temiz (Clean)	188	58
Kontrol Edilmesi Güç Faktörler	Konumla Alakalı Özellikler	Konum (location)	461	65
		Havuz (pool)	160	59
		Manzara (view)	212	40
		Sahil (beach)	130	43
Kontrol Edilmesi Daha Kolay Faktörler	Personelle Alakalı Özellikler	Personel (staff)	502	155
		Dostça (friendly-)	251	20
		Yardımsever (helpful)	217	29
	Yemekle Alakalı Özellikler	Yemek (food)	159	64
		Kahvaltı (breakfast)	234	174

Kaynak: Yazar tarafından oluşturulmuştur.

Tablo 2. Olumsuz Yorumlarda En Sık Tekrar Eden Kelimelerin Listesi

Kelime	Olumlu Yorumlarda Tekrarlanma Sayısı	Olumsuz Yorumlarda Tekrarlanma Sayısı
Oda (room)	404	474
Otel (hotel)	396	270
Kahvaltı (breakfast)	234	174
Personel (staff)	502	155
Kötü (poor)	3	120
Wi-fi	27	106
Yatak (bed)	53	114
Küçük (small)	29	101

Kaynak: Yazar tarafından oluşturulmuştur.

(4. adım)İncelenen kelimeler daha sonra, ziyaretçilerin yorumlarında geçip geçmediğine göre “1” veya “0” olarak kodlanmış ve ziyaretçinin konaklama tesisine verdiği puanla ilişkisinin kuvvetine bakılmıştır. İlgili hipotezler aşağıda listelenmiş, detaylı olarak her bir kelimenin tekrarlanma sayısı ile konaklama tesisine verilen puanların ilişkisi üzerine kurulan araştırma hipotezleri Tablo 3’te özetlenmiştir. Bu hipotezleri test etmek amacıyla, korelasyon analizinden faydalanılmıştır. Analiz sonuçlarına göre hipotezlerin kabul durumlarına da Tablo 3’te yer verilmiş olup, kabul edilen hipotezlerdeki ilişki büyüklüğünü gösteren Pearson Katsayıları, Tablo 4’te listelenmiştir.

H₁: Olumlu yorumlarda en sık tekrar eden kelimelerin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.

H₂: Olumsuz yorumlarda en sık tekrar eden kelimelerin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.



Tablo 3. Araştırma Hipotezlerinin Listesi

H₁₋₁: Olumlu yorumlarda “otel” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₂: Olumlu yorumlarda “oda” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₃: Olumlu yorumlarda “güzel” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₄: Olumlu yorumlarda “muhteşem” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₁₋₅: Olumlu yorumlarda “hoş” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₆: Olumlu yorumlarda “temiz” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₇: Olumlu yorumlarda “konum” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₈: Olumlu yorumlarda “havuz” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₉: Olumlu yorumlarda “manzara” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₁₀: Olumlu yorumlarda “sahil” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₁₁: Olumlu yorumlarda “personel” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₁₋₁₂: Olumlu yorumlarda “dostça” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₁₃: Olumlu yorumlarda “yardımsever” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₁₋₁₄: Olumlu yorumlarda “yemek” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₁₋₁₅: Olumlu yorumlarda “kahvaltı” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₂₋₁: Olumsuz yorumlarda “oda” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₂₋₂: Olumsuz yorumlarda “otel” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₂₋₃: Olumsuz yorumlarda “kahvaltı” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₂₋₄: Olumsuz yorumlarda “personel” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₂₋₅: Olumsuz yorumlarda “kötü” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₂₋₆: Olumsuz yorumlarda “wi-fi” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red
H₂₋₇: Olumsuz yorumlarda “yatak” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Kabul
H₂₋₈: Olumsuz yorumlarda “küçük” kelimesinin tekrarlanma sayısı ile ziyaretçilerin konaklama tesisine verdikleri puan arasında anlamlı ilişki vardır.	Red


Tablo4. Ziyaretçilerin Konaklama Tesisine Verdikleri Puanla Anlamlı Korelasyonu Bulunan, Olumlu veya Olumsuz Yorumlarda Sık Tekrar Edilen Kelimeler

		Pearson Korelasyon Katsayısı
Olumlu Yorumlarda Sık Tekrar Edilen Kelimeler	Personel	0,178**
	Muhteşem	0,129*
	Yardımcı	0,121*
Olumsuz Yorumlarda Sık Tekrar Edilen Kelimeler	Oda	-0,140*
	Otel	-0,295**
	Personel	-0,274**
	Kötü	-0,173**
	Yatak	-0,139*

**0,01 düzeyinde anlamlıdır.

*0,05 düzeyinde anlamlıdır.

5. TARTIŞMA VE SONUÇ

Ziyaretçilerin kaldıkları tesisleri değerlendirdikleri, bir nevi ağızdan ağza pazarlama kaynağı olan “booking.com” web sitesinde yer alan, Türkiye’de tatil yapan ziyaretçilerin yorumları üzerine yapılan bu çalışmada, olumlu ve olumsuz yorumlarda en sık tekrarlanan kelimeler metin madenciliği yöntemiyle incelenmiş ve bu kelimelerin yorumlarda tekrar sayısının, ziyaretçilerin değerlendirme puanıyla ilişkilerinin ilişki kuvveti analiz edilmiştir.

Öncelikle yapılan metin madenciliği yöntemi, günümüzde yapısal olmayan verinin incelenmesi için inovatif bir yöntemdir. Bu çalışmada toplanan 1400 olumlu, 1400 olumsuz, toplam 2800 yorumun tek tek anlamlı biçimde okunması mümkün değildir. Dolayısıyla bir veri yığını olarak bulunan bu verinin madenciliği, yorumlardan değerli olan bilgiyi, yani en sık tekrarlanan kelimeleri ortaya çıkartmaktadır. Ardından ortaya çıkarılan kelimelerin yorumlarda geçip geçmemesine göre, ziyaretçilerin değerlendirme puanlarıyla ilişkilerinin anlamlılığı incelenmiştir.

Elde edilen sonuçlar, çalışanlar üzerine geliştirilmiş Herzberg’in çift faktör motivasyon teorisini, turistlere adapte edilerek yorumlanabilir. Yani, ziyaretçilerin olumsuz yorumlarında yer alan (onları demotive eden) bazı özellikler (kelimeler) vardır ve bu kelimelerin yorumlarda geçmesi, ziyaretçilerin değerlendirme puanıyla ters yönde anlamlı ilişkilerin mevcut olduğunu göstermektedir. Ancak bu özelliklerin var olması, müşterilerin memnun olacağı anlamına gelmez, yani bu özellikler, hijyen faktörleri olarak değerlendirilebilir.

Diğer taraftan, bazı kelimelerin olumlu yorumlarda var olmasının, ziyaretçilerin puanıyla ilişkisi pozitif yönde ve anlamlıdır. Yani, bu özelliklerin olmaması, memnuniyetsizliği getirmezken; onların var olması, müşterilerin memnuniyetini getirecektir. Ancak, çalışmanın bulgularına göre, bu durumun tek istisnası “personel”dir. Tablo 3’e bakıldığında, hem olumlu



hem olumsuz yorumlarda personel kelimesinin var olması ile ziyaretçi puanları arasında anlamlı korelasyon bulunmaktadır. Yani ziyaretçiler, personelden memnun kalırsa, onları olumlu yorumlarında anlatmakta ve buna göre otele yüksek puan vermekte, eğer memnun kalmazlarsa da, onlardan olumsuz yorumlarında bahsetmekte ve değerlendirmelerinde düşük puan vermektedirler. Bunun için, otel yöneticileri, mutlaka personelle ilgili özellikleri ön planda tutmalı, onlara gerekli eğitimleri vermelidir.

Bu çalışmada izlenen metin madenciliği süreci; konaklama tesislerinin yöneticileri tarafından, ziyaretçilerinin memnun kalmadığı özellikleri ve onları memnun edecek rekabet avantajlarını keşfetmek için kullanılabilir. Ayrıca, gelecek çalışmalarda akademisyenler de, geleneksel istatistik analiz yöntemlerini, bu çalışmada olduğu gibi “büyük veri”yle ilişkilendirerek anlamlı sonuçlara ulaşabilirler.

KAYNAKLAR

- Çoban, O., & Özcan, C. (2013). Türkiye'de Turizm Gelirleri - Ekonomik Büyüme İlişkisi: Nedensellik Analizi (1963-2010). *Eskişehir Osmangazi Üniversitesi İİBF Dergisi* , 8 (1), 243-261.
- Dolnicar, S., & Ring, A. (2014). Tourism Marketing Research: Past, Present and Future. *Annals of Tourism Research* , 47, 31-47.
- Erevelles, S., Fukawa, N., & Swayne, L. (2015). Big Data Consumer Analytics and The Transformation of Marketing. *Journal of Business Research* , In Press.
- Euromonitor's 100 city destination ranking.
<http://www.traveldailynews.com/news/article/70264/euromonitor-s-100-city-destination-ranking>.
- Gassiot, A., & Coromina, L. (2013). Destination Image of Girona: An Online Text-Mining Approach. *International Journal of Management Cases* , 301-314.
- Herzberg, F. (1975). *Work and the nature of man*. New York: T.Y. Crowell.
- Herzberg, F., Mausner, B., & Snyderman, B. (1962). *The motivation to work*. New York: John Wiley and Sons, Inc.
- <http://www.touropia.com/tourist-attractions-in-tur>, ... <http://www.touropia.com/tourist-attractions-in-turkey/>.
- <https://www.tripadvisor.com.tr/Tourism-g293969-Tur>, ... Popüler Destinasyonlar: Türkiye.
- Kushi, E., & Caca, E. (2012). Informal Sector in the Sun-and-Beach Segment in Albania . *EuroEconomica* , 31 (1), 33-39.
- Laudon, K., & Laudon, J. (2014). *Management Information Systems: Managing the Digital Form* (13th Global Edition b.). USA: Pearson Education Limited.
- McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review* , October, 61-68.
- Porter, M. (1985). *Competitive advantage – creating and sustaining superior performance*. New York: Free Press.
- TC Kültür ve Turizm Bakanlığı, <http://www.kulturturizm.gov.tr/>, erişim tarihi: 01.04.2016.
- Xiang, Z., Schwartz, Z., Gerdes, J., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management* , 44, 120-130.