



**ÇEVİRİMİÇİ SEYAHAT ACENTELEİNDE  
ODA FİYATLARINA ETKİ EDEN  
FAKTÖRLERİN ARAŞTIRILMASI**

**Yılmaz AĞCA**

**Doktora Tezi**

**YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI**

**Doç. Dr. Abdulkadir ÖZDEMİR**

**2019**

**Her Hakkı Saklıdır**

**ATATÜRK ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI**

**Yılmaz AĞCA**

**ÇEVİRİMİÇİ SEYAHAT ACENTELERİNDE ODA FİYATLARINA ETKİ EDEN**  
**FAKTÖRLERİN ARAŞTIRILMASI**

**DOKTORA TEZİ**

**TEZ YÖNETİCİSİ**  
**Doç. Dr. Abdulkadir ÖZDEMİR**

**Erzurum - 2019**

**İÇİNDEKİLER**

<b>İÇİNDEKİLER</b> .....	<b>I</b>
<b>ÖZET</b> .....	<b>V</b>
<b>ABSTRACT</b> .....	<b>VI</b>
<b>SİMGELER VE KISALTMALAR DİZİNİ</b> .....	<b>VII</b>
<b>ŞEKİLLER DİZİNİ</b> .....	<b>VIII</b>
<b>TABLolar DİZİNİ</b> .....	<b>X</b>
<b>ÖNSÖZ</b> .....	<b>XII</b>
<b>GİRİŞ</b> .....	<b>1</b>

**BİRİNCİ BÖLÜM****TURİZM, TURİZM SEKTÖRÜNÜN YAPISI VE TÜRKİYE’DE TURİZM**

<b>1.1. TURİZM KAVRAMI</b> .....	<b>7</b>
<b>1.2. TURİZM SEKTÖRÜNÜN YAPISI</b> .....	<b>9</b>
1.2.1. Tur İşletmeleri .....	11
1.2.1.1. Tur Operatörleri veya Tur Toptancıları .....	12
1.2.1.2. Seyahat Acenteleri .....	12
1.2.1.3. Çevrimiçi Seyahat Acenteleri .....	13
1.2.2. Konaklama İşletmeleri.....	17
1.2.3. Yiyecek-İçecek İşletmeleri .....	17
1.2.4. Ulaştırma İşletmeleri .....	20
<b>1.3. TÜRKİYE’DE TURİZM SEKTÖRÜ</b> .....	<b>20</b>
1.3.1. Türkiye’de Turizmin Gelişimi.....	23
1.3.2. Türkiye’de Turizm Sektörü .....	25
1.3.2.1. Türkiye’de Konaklama Sektörü.....	26
1.3.2.2. Türkiye’de Tur/Seyahat Sektörü.....	29

## İKİNCİ BÖLÜM

### WEB MADENCİLİĞİ VE BİLİŞİM KAVRAMLARI

<b>2.1. İNTERNET VE İNTERNETİN GELİŞİMİ .....</b>	<b>33</b>
2.1.1. İnternetin Gelişmesinde Etkili Olan Teknolojiler.....	34
2.1.1.1. Paket Anahtarlama.....	35
2.1.1.2. TCP/IP .....	35
2.1.1.3. DNS .....	36
2.1.1.4. Web/Www .....	37
<b>2.2. BÜYÜK VERİ .....</b>	<b>38</b>
2.2.1. Depolama Teknolojisinin Gelişmesi.....	38
2.2.2. Bilgisayarların İşlem Gücünün Artması.....	39
2.2.3. Mobil Teknolojilerin Gelişmesi .....	41
2.2.4. Büyük Verinin Kullanım Alanları.....	43
2.2.4.1. Müşteri Davranışlarının Araştırılmasında .....	43
2.2.4.2. Dolandırıcılık ve Güvenlik Açıklarının Tespitinde .....	44
2.2.4.3. Bilimsel Araştırmalarda.....	44
2.2.4.4. Sağlık ve Halk Sağlığının İyileştirilmesinde .....	44
<b>2.3. VERİ BİLİMİ VE VERİ MADENCİLİĞİ .....</b>	<b>45</b>
2.3.1. Veri Madenciliği Aşamaları.....	46
2.3.1.1. Problemin Tanımlanması.....	46
2.3.1.2. Veriyi Çözümleme ve Temizleme .....	46
2.3.1.3. Veri Dönüşümü (Data Transformation).....	47
2.3.1.4. Veri Madenciliği.....	47
2.3.1.5. Örüntü Değerlendirme (Pattern Evaluation).....	47
2.3.1.6. Bilginin Raporlanması .....	48
2.3.2. Veri Madenciliği Yöntemleri.....	48
2.3.2.1. İlişki Analizi ve Birliktelik Kuralları.....	48
2.3.2.2. Kümeleme Analizi .....	49
2.3.2.3. Sınıflandırma Analizi.....	50
2.3.2.3.1. Karar Ağaçları (Decision Trees).....	51

2.3.2.3.2. Yapay Sinir Ağları (Artificial Neural Networks, ANN) .....	52
2.3.2.3.3. Lojistik Regresyon.....	54
2.3.2.3.4. Naïve Bayes Sınıflandırma .....	54
<b>2.4. WEB MADENCİLİĞİ.....</b>	<b>55</b>
2.4.1. Web Madenciliğinin Ortaya Çıkmasına Etki Eden Gelişmeler .....	55
2.4.1.1. Gelişen Bilgisayarlar ve Artan İşlem Gücü .....	55
2.4.1.2. Depolama Donanımlarının Ucuzlaması ve Kapasitelerinin Artması ....	56
2.4.1.3. İnternet Hız ve Bant Genişliğinin Artması .....	56
2.4.1.4. İnternetin Genişlemesi ve Üretilen Veri Miktarının Artması .....	57
2.4.2. Web Madenciliğinde Kullanılan İçerik Yayınlama Teknolojileri.....	57
2.4.2.1. Web Siteleri .....	57
2.4.2.2. CSS .....	59
2.4.2.3. XML .....	61
2.4.2.4. JSON.....	62
2.4.2.5. AJAX .....	62
2.4.3. Web Madenciliği Yöntemleri.....	63
2.4.3.1. Python İle Web Madenciliği.....	64
2.4.3.2. R İle Web Madenciliği .....	68
2.4.3.3. Web Madenciliğinde XPath Kullanımı .....	71
2.4.3.4. Knime İle Web Madenciliği .....	79
<b>2.5. LİTERATÜR ÖZETİ .....</b>	<b>86</b>

### ÜÇÜNCÜ BÖLÜM

## ÇEVİRİMİÇİ SEYAHAT ACENTELERİNDE ODA FİYATLARINA ETKİ EDEN FAKTÖRLERİN ARAŞTIRILMASI

<b>3.1. ARAŞTIRMA YÖNTEMİ .....</b>	<b>91</b>
3.1.1. Web Madenciliği ile Verinin Elde Edilmesi.....	91
3.1.2. Verinin Analizi ve Hipotez Testleri.....	93
<b>3.2. BULGULAR .....</b>	<b>95</b>
3.2.1. Regresyon Modeline İlişkin Bulgular .....	108
<b>SONUÇ VE TARTIŞMA.....</b>	<b>119</b>

<b>KAYNAKÇA.....</b>	<b>123</b>
<b>EKLER.....</b>	<b>140</b>
<b>ÖZGEÇMİŞ.....</b>	<b>144</b>



**ÖZET****DOKTORA TEZİ****ÇEVİRİMİÇİ SEYAHAT ACENTELERİNDE ODA FİYATLARINA ETKİ EDEN  
FAKTÖRLERİN ARAŞTIRILMASI****Yılmaz AĞCA****Tez Danışmanı: Doç. Dr. Abdulkadir ÖZDEMİR****2019, 122 sayfa**

**Jüri:** Prof. Dr. Abdullah NARALAN  
Doç. Dr. Abdulkadir ÖZDEMİR  
Dr. Öğr. Üyesi Mustafa KESKİNKILIÇ  
Dr. Öğr. Üyesi Ahmet Kamil KABAKUŞ  
Dr. Öğr. Üyesi Ahmet İlker AKBABA

İnternet ve bu altyapıyı kullanan diğer servisler çok büyük miktarda veri barındırmaktadır. Gelişen depolama teknolojisi, hızlı işlem yapan bilgisayarlar, artan bant genişlikleri; sunucular ve web siteleri üzerinde bulunan bu veriyi işleme ve kullanma olanağı yaratmıştır. Web madenciliği, webi büyük bir veri kaynağı şeklinde kullanarak, bu sistem üzerindeki web sayfalarından bilgi derlemeyi olanaklı hale getiren işlemler bütünüdür. Bir anketle veya yüz yüze bir görüşmeyle günlerce sürecektir bir veri toplama işlemi, eğer veri web üzerinde varsa web madenciliği ile çok kısa sürede yapılabilmektedir.

Bu çalışmada, web madenciliği kullanılarak, birçok seyahat acentesine ait oda fiyatlarının da dâhil olduğu, çeşitli bilgileri üzerinde barındıran bir web sitesinden, Türkiye'deki 6558 konaklama tesisine ait veri elde edilmiştir. Geliştirilen algoritma yardımı ile 2018 Mayıs-Eylül ayları arasındaki fiyatlardan tesislerin ortalama oda fiyatları hesaplanmış, elde edilen 49 değişkenden hangilerinin ortalama fiyat üzerinde etkisi olduğu araştırılmıştır. Bu değişkenler kullanılarak otel oda fiyatı bağımlı değişken olacak şekilde sıralı lojistik regresyon modeli oluşturulmuştur. Regresyon modelinde oda fiyatını etkileyen değişkenler *yorum sayısı*, *puan*, *tesis tipi*, *deniz kıyısı olup olmaması*, *kapalı havuz*, *çamaşır yıkama hizmeti*, *klima*, *mini bar*, *ücretsiz otopark*, *mikrodalga fırın*, *spa* olarak belirlenmiştir. Ek olarak bu veri üzerinde çeşitli istatistiksel analizler yapılmış; fiyatların bölgesel ve tesis türlerine göre farklılıkları, tesis puanı ile fiyat arasındaki ilişki gibi pek çok bulguya ayrıca yer verilmiştir. Kullanılan yöntemin ve elde edilen bulguların araştırmacılara ve sektöre katkı sağlayacağı umut edilmektedir.

**Anahtar Kelimeler:** web madenciliği, web kazıma, büyük veri, otel oda fiyatları, çevrimiçi seyahat acenteleri, turizm

**ABSTRACT****Ph. D. DISSERTATION****INVESTIGATION OF FACTORS AFFECTING ROOM PRICE IN ONLINE TRAVEL AGENCIES****Yılmaz AĞCA****Advisor: Assoc. Prof. Dr. Abdulkadir ÖZDEMİR****2019, Page: 122**

**Jury:** **Prof. Dr. Abdullah NARALAN**  
**Assoc. Prof. Dr. Abdulkadir ÖZDEMİR**  
**Assis. Prof. Dr. Mustafa KESKİNKILIÇ**  
**Assis. Prof. Dr. Ahmet Kamil KABAKUŞ**  
**Assis. Prof. Dr. Ahmet İlker AKBABA**

The Internet and other services that use this infrastructure contain large amounts of data. Developing storage technology, fast processing computers and increased bandwidths provided opportunity to process and use the data on servers and web sites. Using the web as a large source of data, web mining is a set of processes which makes it possible to compile data from the web pages on this system. A data collection process that can last for days with a questionnaire or face to face interview can be completed with web mining in a very short time if the data is available on the web.

In this study, the data of 6558 accommodation premises in Turkey were obtained from a web site that has a variety of information including room prices of many travel agencies using web mining. By means of algorithm developed, average room prices of the accommodation premises were calculated based on prices between May and September 2018 and it was investigated which of the 49 variables obtained had an effect on the average price. By using these variables and considering the hotel room price as dependent variable, a sequential logistic regression model was formed. In the regression model, *number of reviews*, *star rating/review score*, *type of property*, *distance from the sea*, *indoor pool*, *laundry service*, *air conditioner*, *mini bar*, *free parking*, *microwave* and *spa* were found as the variables affecting the room price. In addition, based on these data, various statistical analyses were employed and a variety of findings such as differences between prices according to regions and accommodation types and the relationship between accommodation ratings and the room rate were presented in this study. It was expected that the method used and the findings obtained would contribute to the researchers and the sector.

**Keywords:** web mining, web scraping, big data, hotel room prices, online travel agencies, tourism



**SİMGELER VE KISALTMALAR DİZİNİ**

- AI : All Inclusive (Herşey Dâhil Ücret)
- ANN : Artificial Neural Network
- BB : Bed and Breakfast (Oda ve Kahvaltı Ücreti)
- CDMA : Code Division Multiple Access
- CERN : Conseil Européen pour la Recherche Nucléaire (Avrupa Nükleer Araştırma Merkezi)
- CRS : Computer Reservations Systems
- CSS : Cascading Style Sheets
- ÇSA : Çevrimiçi Seyahat Acentesi
- DARPA : Defense Advanced Research Projects Agency
- FB : Full Board (Tam Pansiyon Ücreti)
- GDS : Global Distribution Systems
- GSM : Global System for Mobile
- HB : Half Board (Yarım Pansiyon Ücreti)
- HTML : HyperText Markup Language
- NFC : Near Field Communication
- OB : Only Bed (Sadece Oda/Yatak Ücreti)
- OTA : Online Travel Agency
- SGML : Standard Generalized Markup Language
- TÜRSAB: Türkiye Seyahat Acenteleri Birliği
- URL : Uniform Resource Locator
- YSA : Yapay Sinir Ağı

## ŞEKİLLER DİZİNİ

<b>Şekil 1.1.</b> Turizm Dağıtım Sistemi .....	11
<b>Şekil 1.2.</b> Yiyecek Sektörünün Yapısı .....	18
<b>Şekil 2.1.</b> 1970'lerin Başlı ARPAnet ve Bağlı Üniversite ve Şirketler .....	33
<b>Şekil 2.2.</b> DNS Hiyerarşik Yapısı .....	37
<b>Şekil 2.3.</b> Bilgisayar İşlem Gücündeki Artış .....	40
<b>Şekil 2.4.</b> K-ortalama Algoritması İle Kümeleme .....	50
<b>Şekil 2.5.</b> Karar Ağacı Diyagramı.....	51
<b>Şekil 2.6.</b> Yapay Sinir Ağı .....	53
<b>Şekil 2.7.</b> HTML Etiketinin Yapısı.....	58
<b>Şekil 2.8.</b> HTML'nin Hiyerarşik Yapısı.....	59
<b>Şekil 2.9.</b> CSS Etiketlerinin Sorgulanması .....	60
<b>Şekil 2.10.</b> XML Örnek Veri Kaydı.....	61
<b>Şekil 2.11.</b> JSON Örnek Veri Kaydı.....	62
<b>Şekil 2.12.</b> Python BeautifulSoup Kütüphanesi İle Web'den İçerik Alma .....	65
<b>Şekil 2.13.</b> Hedef Web Sayfasının Kaynak Kodları .....	66
<b>Şekil 2.14.</b> Otel Adının Bulunduğu Div Nesnesi.....	67
<b>Şekil 2.15.</b> BeautifulSoup Kütüphanesi İle Web'den İçerik Alma .....	67
<b>Şekil 2.16.</b> Pandas Kütüphanesi İle Veri Çerçevesi Oluşturma.....	67
<b>Şekil 2.17.</b> Python ve BeautifulSoup İle Elde Edilen Verinin Çıktısı .....	68
<b>Şekil 2.18.</b> R'da rvest Kütüphanesi İle Web'den İçerik Alma .....	69
<b>Şekil 2.19.</b> Otel Adının Kaynak Kodu .....	69
<b>Şekil 2.20.</b> Otel Fiyatının Kaynak Kodu.....	70
<b>Şekil 2.21.</b> R ve rvest İle Elde Edilen Verinin Çıktısı .....	71
<b>Şekil 2.22.</b> Otel Adı Nesnesinin HTML/CSS Kodları.....	72
<b>Şekil 2.23.</b> Fiyat Nesnesinin HTML/CSS Kodları.....	73
<b>Şekil 2.24.</b> Otel Web Sayfası ve Otelin Çeşitli Özellikleri.....	74
<b>Şekil 2.25.</b> Çoklu Düğüm Seçimi Operatörü Kullanımı HTML/CSS Kodları .....	75
<b>Şekil 2.26.</b> Substring-after ve Substring-before Kullanımı HTML/CSS Kodları.....	77
<b>Şekil 2.27.</b> KNIME Analytics Platform Çalışma Alanı.....	80
<b>Şekil 2.28.</b> Knime'da Web Madenciliği Modülleri .....	81
<b>Şekil 2.29.</b> XPath Sorgu Ekleme Ekranı .....	82

<b>Şekil 2.30.</b> Knime İle Elde Edilen Verinin Çıktısı .....	83
<b>Şekil 2.31.</b> Knime'da Düğümlerin Kullanımı .....	84
<b>Şekil 3.1.</b> İllere Göre Konaklama Tesisi Sayıları .....	101
<b>Şekil 3.2.</b> İllere Göre Ortalama Oda Fiyatları.....	103
<b>Şekil 3.3.</b> R'da Modele İlişkin Tahminin Yapılması.....	118



## TABLOLAR DİZİNİ

<b>Tablo 1.1.</b> Türkiye'deki Turizm İşletme Belgeli Konaklama Tesislerinin Sayısı.....	25
<b>Tablo 1.2.</b> Türkiye'deki Konaklama Tesislerinin Statüsü ve Sayısı (2002-2018) .....	27
<b>Tablo 1.3.</b> Türkiye'deki Turizm Belgeli Konaklama Tesisleri ve Türleri .....	28
<b>Tablo 1.4.</b> Türkiye'deki Seyahat Acenteleri ve Grupları.....	31
<b>Tablo 1.5.</b> Türkiye'de En Çok Yolcu Ağırlayan 10 Havalimanı (2018).....	32
<b>Tablo 2.1.</b> IPv4 ve IPv6 Adresleme Sistemleri .....	36
<b>Tablo 2.2.</b> Sabit Diskler ve Zaman İçerisindeki Gelişimleri.....	39
<b>Tablo 2.3.</b> Mobil İletişimin Gelişimi ve Karşılaştırılması.....	42
<b>Tablo 2.4.</b> Yaygın Kullanılan Bazı XPath İfadeleri ve İşlevleri .....	71
<b>Tablo 2.5.</b> Xpath Operatörleri ve İşlevleri .....	75
<b>Tablo 2.6.</b> Xpath Metin Fonksiyonları ve İşlevleri .....	76
<b>Tablo 2.7.</b> Xpath Nümerik Fonksiyonları ve İşlevleri.....	78
<b>Tablo 3.1.</b> Ortalama Oda Fiyatı Değişkeninin Gruplandırılmasında Kullanılan Yöntemler....	94
<b>Tablo 3.2.</b> Ortalama Oda Fiyatı Değişkeninin Gruplandırılması .....	95
<b>Tablo 3.3.</b> Tesis Türlerine Ait Bazı İstatistikler .....	96
<b>Tablo 3.4.</b> Konaklama Tesisinin Türüne Göre Tek Faktörlü Varyans Analizi Tablosu	96
<b>Tablo 3.5.</b> Tesis Türü Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları .	97
<b>Tablo 3.6.</b> Coğrafi Bölgelere Göre Ortalama Oda Fiyatları .....	98
<b>Tablo 3.7.</b> Tesisin Bulunduğu Bölgeye Göre Tek Faktörlü Varyans Analizi Tablosu ..	98
<b>Tablo 3.8.</b> Tesisin Bulunduğu Bölge Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları .	99
<b>Tablo 3.9.</b> İllere Göre Ortalama Oda Fiyatları .....	100
<b>Tablo 3.10.</b> Ortalama Oda Fiyatı En Yüksek Olan On İl .....	101
<b>Tablo 3.11.</b> Ortalama Oda Fiyatı En Düşük Olan On İl.....	102
<b>Tablo 3.12.</b> Tesisin Bulunduğu İle Göre Tek Faktörlü Varyans Analizi Tablosu .....	103
<b>Tablo 3.13.</b> Tesisin Deniz Kıyısında Olup Olmamasına İlişkin Bazı İstatistikler .....	104
<b>Tablo 3.14.</b> Tesisin Deniz Kıyısında Olmasına Dair Bağımsız Örneklem T Testi Sonuçları	104
<b>Tablo 3.15.</b> Tesislerin Konuk Puanları ve Puanlara Dair Bazı İstatistikler.....	104
<b>Tablo 3.16.</b> Yıldız/Puan ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları ..	105
<b>Tablo 3.17.</b> Tesislerin Konuk Puanlarına Göre Tek Faktörlü Varyans Analizi Tablosu .....	105
<b>Tablo 3.18.</b> Konuk Puanları Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları	106
<b>Tablo 3.19.</b> Bölgelere Göre Tesislerin Aldıkları Yorumlara İlişkin İstatistikler .....	107

<b>Tablo 3.20.</b> Yorum Sayısı ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları	107
<b>Tablo 3.21.</b> Sıralama Puanı ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları .....	108
<b>Tablo 3.22.</b> Regresyon Modeli Belirlenmesine Yönelik Aşamalı Regresyon Sonuçları .....	110
<b>Tablo 3.23.</b> Bağımlı Değişkene (Ortalama Oda Fiyatı) Ait Bazı İstatistikî Değerler ..	110
<b>Tablo 3.24.</b> Ortalama Oda Fiyatı Değişkeninin Gruplandırılması .....	111
<b>Tablo 3.25.</b> Modelde Yer Alan Değişkenler ve Değişkenlere Ait Frekanslar (N=6558).....	111
<b>Tablo 3.26.</b> Modelin Paralel Doğrular Testine Ait Sonuçları .....	114
<b>Tablo 3.27.</b> Lojistik Regresyon Sonucu ve Değişkenlerin Anlamlılıkları .....	114
<b>Tablo 3.28.</b> Regresyon Modeline Ait Pseudo R <sup>2</sup> Değerleri .....	115



**ÖNSÖZ**

Çağımızda internet günlük hayatın önemli parçalarından birisi haline gelmiştir. Örneğin haber okumak, alışveriş yapmak, film izlemek, haberleşmek gibi birçok sosyal aktivite internet aracılığıyla yapılmakta olup, kullanım alanı her geçen gün daha da genişlemektedir. Üstelik içerdiği veri miktarı ile en büyük veri tabanı olarak kabul edilebilir. Bu veriler; klasik veri tabanlarından farklı olarak binlerce web sitesi üzerinde yer almaktadır. İnternet aracılığıyla elde edilen veriler işlendiğinde, geçmişte mümkün olmayacak pek çok araştırma yapılabilir. Bu çalışma, internetin bu zengin veri kaynağından faydalanılarak yapılmıştır. Bu amaçla; çeşitli konaklama tesislerinin oda satışının yapıldığı ve tesislere ait bilgilerin yer aldığı bir web sitesi, geliştirilen bir algoritma yardımı ile taranmış, elde edilen veriler ile bir veri tabanı oluşturulmuş ve bu veri tabanı oda fiyatını etkileyen faktörlerin belirlenmesinde kullanılmıştır. Buna ek olarak, hem yöntem olarak hem de elde edilen bulgular bağlamında literatüre katkı sağlanması amaçlanmıştır.

Yönetim Bilişim Sistemleri alanına yönelmemde etkisi olan bölümün değerli hocalarına ve mesai arkadaşlarımdan Öğr. Gör. Ekrem Akbulut'a teşekkürle bu kısmı sonlandırmak istiyorum. Doktora sürecinde pek çok değerli akademisyenin tarafıma katkısı ve emeği olmuştur. Herkesi burada yazmak mümkün olmasa da, YBS alanında devam etmemde büyük katkısı olan Dr. Abdullatif Kaban, dersini almayı çok istediğim ve eksikliğini yaşadığım Prof. Dr. Uğur Yavuz hocalarıma teşekkür etmek isterim. Yoğunluğundan dolayı ders dışında çok görüşme olanağımız olmayan, fakat o yoğunluğu dolu dolu kullanan, üniversitemizde büyük emeği olan bölüm başkanımız Prof. Dr. Üstün Özen ve bölüm dışından tez izleme komisyonumda yer alan, vaktini bana ayıran Doç. Dr. M. Kürşad Timuroğlu hocalarıma ayrıca teşekkür ederim.

Atatürk Üniversitesinde tanıştığım başta Muzaffer Atasoy olmak üzere tüm arkadaşlarıma, benimle birlikte yorulan sevgili ailem, kardeşim Yıldız Ağca'ya teşekkür ederim. Eğitime ve bilime olan inançla beni yetiştiren sevgili babam Derviş Ağca'yı bu vesileyle anıyorum.

**Erzurum, 2019**

**Yılmaz AĞCA**

## GİRİŞ

Gelişen bilgi ve iletişim teknolojileri insan hayatı üzerinde pek çok köklü değişime neden olmuştur. Birçok sektörün yapısı ve işleyişi bu gelişmeler ile birlikte yeniden şekillenmiştir. Bazı iş alanları ortadan kalkarken, gelişmelere paralel olarak geçmişte var olmayan yeni birçok iş alanı da ortaya çıkmıştır. Turizm sektörü teknoloji yoğun bir sektör olmamasına rağmen, diğer birçok iş alanı gibi teknolojik gelişmelerden oldukça etkilenmiş, özellikle sektörün satış ve pazarlama kısmında önemli değişimler meydana gelmiştir. Bu gelişmeler ve internetin insan hayatına girmesi çevrimiçi seyahat acentesi olarak adlandırılan ve internet üzerinden doğrudan tüketicilere oda, paket tur satışı yapan sistemlerin ortaya çıkmasına neden olmuştur.

Turizm, sürekli kalıfa dönüşmemek ve gelir getirici hiçbir uğraşta bulunmamak koşulu ile bireylerin turizm işletmelerinin ürettiği mal ve hizmetleri tüketerek gerçekleştirdikleri geçici konaklamalardır. Turizm, dünyada en önemli ticari sektörlerden bir tanesidir. Dünyada işgücünün yaklaşık %10'u turizm sektöründe çalışmaktadır. Benzer şekilde turizm sektörünün dünya'daki GSMH içerisindeki payı yaklaşık %10'dur (Sezgin M. , 2015). Turizm, ödemeler dengesine yapmış olduğu katkılar ve istihdam yaratıcı etkisinden dolayı, yapılabirliği açısından potansiyele sahip birçok ülkede önemli bir sektördür.

Sanayi devriminden sonra insanoğlunun hayatı, hiç olmadığı kadar köklü değişikliklere uğramıştır. Günümüze değin, bu değişim hızla devam etmiştir. Bilginin en önemli kaynak olduğu düşünülen bu çağa, bilişim çağı adı verilmektedir (Webster, 2006; Halladay, 2018). Bilişim çağında pek çok sektör gibi turizm sektörü de bu değişimden payını almıştır.

İnternetin insan hayatına girişi teknolojik gelişmelere yeni bir boyut kazandırmıştır. İnternette önce, tatil paketi veya oda kiralama işlemi, daha çok tüketicinin seyahat acenteleri şubelerine giderek basılmış olan kataloglardan tatil yapmak istediği konaklama tesisini ve hizmetlerini seçmesi ile olmaktadır. Günümüzde internet olarak adlandırılan küresel iletişim ağı, insanların hayatını pek çok alanda olduğu gibi, tatil paketi satın alma işleminde de önemli ölçüde değişikliğe uğratmıştır.

Bugün tatile gitmek isteyen bir kişi, kolaylıkla internet üzerinden bu alanda hizmet vermekte olan sitelerden birçok seçenek arasından tatil yapmak istediği yeri ve tesisi seçebilmektedir. Çevrimiçi seyahat acenteleri (ÇSA) olarak adlandırılan internet hizmetleri-siteleri, tüketicilere tatil paketlerini ve turizm hizmetlerini internet üzerinden satın alma olanağı sunmaktadır.

ÇSA'ların ortaya çıkışından önce rezervasyon amaçlı kullanılan bilgisayarlı rezervasyon sistemleri (computer reservations systems), küresel dağıtım sistemleri (global distribution systems) gibi hali hazırda seyahat acenteleri ile mal-hizmet üreticilerini bir araya getiren çevrimiçi sistemler kullanılmaktaydı (Schulz, 1997: 67-70; Kärcher, 1996: 20). İnternetin kullanılmaya başlanması ile birlikte sadece seyahat acenteleri ile mal ve hizmet üreticileri arasındaki bu ağ; genişleyerek ürünlerin doğrudan tüketici ile buluşmasını sağlamıştır. Bu sistemler hem tüketicilerin üreticilere ulaşmasını kolaylaştırmış, hem de üreticilere ürün ve hizmetlerini daha kolay tüketicilere ulaştırma olanağı sunmuştur.

İnternet kullanımı, turizm dağıtım sistemini önemli ölçüde değiştirdiği gibi, Booking.com ve Tripadvisor gibi çevrimiçi sistemlerin dünyanın en büyük seyahat acenteleri haline gelmelerine de neden olmuştur. Bugün Booking.com, Tripadvisor, Rocketmiles ve Expedia gibi çevrimiçi rezervasyon siteleri, 1,3 Trilyon \$ civarındaki turizm pazarından, paylarını her geçen gün artırmaktadır. Bu sitelerden sadece Tripadvisor aylık ortalama 390 milyon kişi tarafından ziyaret edilmektedir (www.tripadvisor.com, 2017). Bu rezervasyon siteleri uluslararası yazında ÇSA'lar (online travel agencies, OTAs) olarak adlandırılmaktadır (Dev ve O'Connor, 2015: 119). Konunun devamında kendi aralarındaki sınıflandırmalar, dikkate alınmadan bu sistemler için ÇSA kısaltması kullanılacaktır.

Tatil paketi satın alırken ÇSA'ların kullanımı son yıllarda artış göstermeye devam etmektedir. İnsanların turizm olayına katılma oranı artarken, satın alma işlemini yaparken çevrimiçi kanalları kullanma payı da sürekli artmaktadır. Dünyada tatil rezervasyonları içerisinde çevrimiçi kanallar ile yapılan harcamalar 2014'de 470 Milyar \$ iken, 2017'de öngörülen rakam yaklaşık 630 Milyar \$'dır. 2020'de ise bunun yaklaşık 820 Milyar \$ olması beklenmektedir (www.statista.com, 2017).



İnsanlar ÇSA'ları birçok nedenden dolayı tercih etmektedir. Bu tercih, yukarıda da bahsedildiği gibi, bu sektörü turizm sisteminin önemli bir aktörü haline getirmiştir. Büyük lüks otellerden aile pansiyonlarına kadar birbirinden farklı çeşitlilikte birçok konaklama tesisi bu sistemler üzerinden oda ve hizmetlerini tüketicilere sunmaktadır. Oteller ise, satışlarını artırmak ve rekabet avantajı sağlamak gibi birçok nedenden dolayı bu sistemler üzerinden oda satışına sıcak bakmaktadırlar (Çalışkan, Kutlu ve Kımiloğlu, 2013). Geçmişte tur operatörleri ve tur toptancılarının yaptıkları işleri, günümüzde Booking.com, Expedia gibi çevrimiçi sistemler yapmaktadır.

ÇSA'lar tüketicilerin seçimlerini kolaylaştırmak ve kendilerine en uygun oteli seçmelerini sağlamak için, benzer olan bazı özelliklerin olup olmamasına göre tesisleri oluşturdukları veri tabanlarına kayıt etmektedirler. Tüketiciler için, aradıkları özellikteki konaklama tesisini bulmada bu veri tabanları büyük kolaylık sağlamaktadır. Bu veri tabanları tüketiciler için olduğu kadar araştırmacılar için de kolaylıklar sağlamaktadır.

Gelişen bilgi ve iletişim teknolojisi araştırma yöntemleri açısından da yeni yöntemleri ortaya çıkarmıştır. Bu yöntemlerden biri web madenciliğidir. Web madenciliği webi büyük bir veri kaynağı olarak ele alarak, bu sistem üzerindeki web sayfalarından bilgi derlemeyi olanaklı hale getiren yöntemler bütünüdür.

Araştırmacılar geçmişte uzun uğraşlar sonucu elde edebilecekleri bilgilere artık web ve buna ilave yöntemler ile kolaylıkla ulaşmaktadır (Schenker, Bunke, Last ve Kandel, 2005: 1-3; Chakrabarti, 2003: 11-12; Han, Kamber ve Pei, 2012: 571; Markov ve Larose, 2007: 7-12). Web madenciliği, web siteleri üzerinde bulunan bilgilerin filtre edilerek sadece gerekli olan bilgilerin çıkarıldığı bir yöntemler bütünüdür.

Web siteleri üzerinde yer alan bilgiler genellikle düzensiz veya yarı düzenli bir halde bulunmaktadır. Web madenciliği ile bu bilgilerin sadece gerekli olanlarından sistemli veri tabanları oluşturulabilmektedir. Gelişen teknoloji, daha önce olmayan sosyal ağlar, elektronik ticaret, bloglar gibi servisleri ortaya çıkarmıştır. Bu servislerin bir kısmında kullanıcılar içerik üreticileri haline gelmiştir. Bu durum hem internetin daha da hızlı gelişmesine hem de milyonlarca farklı kişinin özgün içerik oluşturmasına neden olmuştur.

Hali hazırda servis sağlayıcıların kullanıcıların bilgilerini depoladıkları ve bunu çeşitli amaçlarla kullandıkları bilinmektedir. Günümüzde bu veriler oldukça değerli

sayılmaktadır. Bunun en büyük nedenlerinden biri, bu verinin bilgi çıkarımında bulunmayı sağlayacak nitelikte oluşudur. Fakat bu şirketler çoğu zaman bu veriyi bağımsız kişiler veya kurumlarla paylaşmamaktadır. Web madenciliği, web üzerinde dağınık halde bulunan ya da yarı sistemli olan veriden istifade etmeyi amaçlayan bir yöntemdir.

Web, belki de günümüzdeki en büyük veri tabanıdır. Binlerce işletmenin veri tabanlarından yayınlanmış veriler, milyarlarca kişiye ait çeşitli bilgiler e-ticaret siteleri, bloglar, sosyal ağlar, video yayın siteleri, haber siteleri üzerinde tamamen kamuya açık bir şekilde bulunmaktadır. Dağınık halde olan ve bilgi çıkarımında bulunması sadece insan çabası ile zor olan bu veri, sistemli hale getirildiği zaman çok değerli olabilmektedir.

Geçmişte genellikle perakendeci seyahat acenteleri kataloglarından yapılan tatil paketi ve oda satışı artık internet üzerinden, yani çevrimiçi bir kanaldan yapılmaktadır. Katalogların sahip olduğu kıt imkânların aksine web, oldukça zengin pazarlama, kıyaslama imkânlarına sahiptir. Bir acenta katalogunda bir konaklama tesisine bir ya da birkaç sayfa ayrılabilirken, bir web sitesinde tek bir sayfa üzerinde yüzlerce fotoğraf, yorum, video, tesis planı gibi bilgiler bulunmaktadır.

Bu çalışmada, web madenciliği yöntemi ile internetin zengin bilgi kaynağından nasıl faydalanılabileceği verilmektedir. Örnek veri kaynağı olarak üzerinde otel oda fiyatlarını ve otele ait çeşitli özellikleri barındıran bir ÇSA kullanılmıştır. Araştırmada hem web madenciliği ile internetten nasıl bilgi çıkarımında bulunulabileceği açıklanacak hem de elde edilen veri ve analizlerle Türkiye’de otel oda fiyatlarına etki eden faktörler belirlenecektir.

Bu çalışma, web ve web madenciliği imkânlarından istifade ederek Türkiye’deki otellerin fiyatları üzerinde etkili olan faktörlerin neler olduğunu belirlemeyi amaçlamaktadır. Çalışma ayrıca fiyatların ve diğer özelliklerin bölgesel farklılığını ve coğrafi dağılımını da saptamayı amaçlamaktadır.

Bu amaçla internet üzerinde hizmet veren ve birçok seyahat acentasına ait oda fiyatlarını bir araya getiren bir seyahat sitesinden, geliştirilen bir web madenciliği algoritması ile elde edilmiş veri, çeşitli istatistikî yöntemler ile analiz edilerek oda fiyatları üzerinde etkili olan faktörlerin neler olduğu belirlenmeye çalışılacaktır.

Veri, ÇSA fiyatlarını tek bir platformda toplayan, bir siteden web madenciliği/kazıma (web scraping) yöntemleri ile elde edilmiştir. Verinin elde edildiği sitede konaklama tesisinin çeşitli özelliklere sahip olup olmadığını gösteren değişkenler, çeşitli sıralama değerleri, bulunduğu bölge, otelin türü, oda sayısı ve daha önce orada kalan konukların verdiği puanlar gibi bilgiler yer almaktadır. Araştırma sonucunda konaklama tesisinin sahip olduğu 50'ye yakın değişkenin hangilerinin fiyat üzerinde etkisinin olup olmadığı belirlenecek, yapılacak olan sıralı lojistik regresyon analizi ile fiyat değişkeninin etkileyen değişkenlere ait model veya modeller ortaya konulacaktır.

Araştırma sonucunda Türkiye'de otel oda fiyatlarını etkileyen faktörleri gösteren anlamlı bir model ortaya konulması amaçlanmaktadır. Ortaya konulacak model, araştırmacılar için pek çok çalışmada kullanılacak bilgiler verecektir. Sektör açısından da hangi özelliklerin aslında fiyat açısından belirleyici olduğu ortaya konulacaktır.

Veri elde etmede kullanılan web madenciliği yöntemi Python ve R Programlama dillerinde örnekler üzerinde açıklanarak, gelecekte birçok veri analisti ve araştırmacı için yöntemin uygulanmasında kaynak bir materyal oluşturacaktır. Çalışma bu açılarından da sonraki bilimsel araştırmalara, turizm sektörüne ve turizm ile ilişkili diğer iş sektörlerine katkı sağlamayı amaçlamaktadır.

Çalışma üç bölümden oluşmaktadır. Birinci bölümde turizm ve turizm sektörünün yapısı, dünya ekonomisindeki yeri ve önemi, tarihsel süreç içerisinde turizm sektörünün nasıl geliştiği, modern turizm ve günümüz turizmi hakkında açıklayıcı bilgiler verilecektir. Ayrıca son yıllarda ÇSA'ların ortaya çıkması ile radikal bir değişime giren turizm dağıtım sisteminin yapısı, bu yapıyı oluşturan alt sektörler ve bu sektörlerin yapı içerisindeki rolleri hakkında bilgiler verilecektir. Birinci bölümün son kısmında Türkiye'deki turizm sektörünün gelişimi, dünyadaki yeri, sahip olduğu arz potansiyeli ve yapısı hakkında bilgiler verilecektir.

Çalışmada ikinci bölüm, internet ve internetin gelişmesinde etkili olan teknolojiler, web madenciliği ve web madenciliğini gelişimi konusunda açıklamalar ile başlamaktadır. Sonrasında son yıllarda pek çok alanda bilgi çıkarımının en önemli kavramı büyük veri açıklanmaktadır. Web madenciliği büyük veriyi elde etmenin en önemli yöntemlerinden olduğundan bu konuya değinilmiştir. Web madenciliği büyük veriyi elde etme yöntemi

olarak kullanılırken, elde edilen verinin işlenmesinde yaygın bir şekilde veri madenciliği kullanılmaktadır. Bundan dolayı sonrasında veri madenciliği ve veri madenciliği yöntemlerine kısaca değinilmiştir. İkinci bölümün devamından araştırma verisinin elde edilmesinde kullanılan, araştırmanın önemli bir kısmını oluşturan web madenciliği kavramı, kavramın ortaya çıkmasına neden olan gelişmeler, web madenciliği yapılırken kullanılan içerik yayınlama teknolojileri hakkında bilgi verilmektedir. Sonrasında web madenciliği yöntemi gerçek bir örnek üzerinde, web madenciliğindeki popüler diller olan Python ve R'da gösterilmektedir. Ayrıca araştırmada veri elde etmede kullanılan Knime yazılımında da bu örnekler tekrarlanmaktadır. İkinci bölümün son kısmında yazın taraması sonucunda ilgili konularda daha önce yapılan araştırmalara ve bu araştırmaların bulgularına yer verilmektedir.

Üçüncü bölümde araştırma yöntemi iki kısımda ele alınmakta, birinci kısımda verinin nasıl elde edildiği, Knime yazılımı hakkında çeşitli bilgiler ve kullanılan düğümler açıklanmaktadır. Araştırma yönteminin ikinci kısmında verinin analiz edilmesinde kullanılan istatistiki yöntemler, kullanılan hipotez testleri, araştırma sırasında geliştirilen normalleştirme yöntemi açıklanmaktadır.

Üçüncü bölümün son kısmı, araştırma bulgularından oluşmaktadır. Bulgular bölümünde öncelikle veriye ait tanımlayıcı istatistikler verilmektedir. Lojistik regresyon modelinden önce, hipotez testleri ile gruplar arası farklılıkların anlamlılığına ait sonuçlar verilmektedir. Devamında sürekli haldeki normal dağılmayan ve bu nedenle sıralı kategorik hale dönüştürülen bağımlı değişken, iki farklı gruplandırma yöntemi ile regresyon modeli içerisinde test edilmekte ve bulgularına yer verilmektedir.

Son olarak sonuç ve tartışma bölümünde kısaca araştırmanın konusu, yapılanlar ve elde edilen önemli sonuçlar hakkında bilgi verilecek. Araştırma bulguları geçmişte yapılan benzer çalışmaların sonuçları kapsamında değerlendirilecektir.

## BİRİNCİ BÖLÜM

### TURİZM, TURİZM SEKTÖRÜNÜN YAPISI VE TÜRKİYE'DE TURİZM

Bu bölümde turizm kavramı, turizm sektörünün genel yapısı ve turizm dağıtım sistemi hakkında bilgi verilecektir. Her ne kadar turizm, günlük hayatta insanların yabancı olmadığı bir kavram olsa da, sektörün yapısı ve işleyişi hakkında açıklayıcı bilgiler vermek konunun daha iyi anlaşılması açısından ek katkı sağlayacaktır.

#### 1.1. TURİZM KAVRAMI

Turizm “*tourism*” kelimesinin ilk kullanılmaya başlanması 17.yy’da İngiliz varlıklı sınıfının çocuklarını eğitim amaçlı Avrupa turlarına çıkarması ile başlasa da (Patterson, 2018: 11), günümüzde kullanıldığı biçimiyle turizm sanayi devriminden sonra ortaya çıkmış (Gherasim ve Tanese, 2012), modern çağa ait bir kavramdır. Sanayi devriminden sonra kent nüfusundaki artış ve insanların bu yorucu, kirli ve stresli şehir ortamından bir nebze olsun uzaklaşmak için (Holloway ve Taylor, 2006: 33) kırsalın doğal, temiz ve dinlendirici ortamına gitme ihtiyaçları, modern turizmin ortaya çıkmasına neden olan faktörlerden biridir. Turizm, günümüzde neredeyse dünya nüfusunun 1/6’sının katıldığı (International Tourism, Number of Arrivals, 2016; World Tourism Organization, 2018) bir seyahat türü haline gelmiştir.

İnsanoğlu tarih boyunca seyahat etse de (Beech ve Chadwick, 2006: 3-4; Edgell, Allen, Smith ve Swanson, 2008: 38-39), yapılan bu seyahatlerin büyük bir iş sektörü haline gelmesi için, içinde bulunduğu sosyo-ekonomik yapının değişmesi gerekiyordu. Bu nedenle sanayi devriminden sonraki değişimler ve gelişmeler günümüzde kullanıldığı anlamıyla turizm sektörünün ortaya çıkmasına neden olmuştur.

19.yy başlarında modern anlamda turizm sektörü artık bilinmekteydi. Buhar makinasının getirdiği teknolojik gelişmeler ile insanlar farklı ülkelere ve bölgelere gezi, eğlence amaçlı seyahatler yapmaktaydı (Elliott J. , 2002: 23-24). Bu gelişmeler yaşanırken turizm ile ilgili olarak çeşitli tanımlamalar ortaya atılmaktaydı. Bunlardan biri Guyer-Feuler’in (1905) turizm tanımıdır. Bu tanımda turizm, insanın değişen hava koşulları ve dinlenme ihtiyacının artmasından dolayı, doğa ve sanatsal alanları tanıma isteği, teknolojinin gelişmesi ile ortaya çıkan ve ulusları birbirine yaklaştıran bir olay

olarak tanımlanmıştır (Kozak, Evren ve Çakır, 2013: 15). Bu tanımda teknolojinin hem turizmin gelişmesine etkisine hem de turizmin ortaya çıkmasına neden olan sosyo-ekonomik koşullardaki değişimlerine vurgu yapılmaktadır. Ayrıca tanım turizmin sosyolojik bir etkisi olan insanları tanıma olgusuna da vurgu yapmaktadır. Giderek tanımlar çeşitlenmektedir. Örneğin bugünkü Birleşmiş Milletlerin temeli olan Milletler Cemiyeti (League of Nations), yabancı turisti tanımlarken; kökeni oldukları ülkeleri dışına 24 saati aşacak şekilde seyahat eden kimse olarak tanımlamıştır (Beech ve Chadwick, 2006: 4). Bu tanımda da kalış zamanı ile ilgili duruma bir vurgu yapılmaktadır. Turizmi tanımlama girişimlerinden bir diğeri olan Hunziker ve Krapf'ın (1942) tanımı; turizmi sürekli ikamete dönüşmemek koşuluyla ve herhangi bir kar amacı gütmeyen yapılan seyahatler ve ilişkiler bütünü olarak açıklamıştır (Holloway ve Taylor, 2006: 4-5). Burada da önceki tanımlara ilave olarak herhangi bir maddi kazanç olmaması durumuna vurgu yapılmaktadır.

Turizmin gelişmesine paralel olarak daha pek çok tanımlama yapılmıştır. Bu tanımların birçoğu bugünkü turizm tanımının temelini oluşturmaktadır. Bazıları ise yeni turizm çeşitlerinin ortaya çıkması ile birlikte geçerliliğini yitirmiştir. Bu tanımların standart hale gelmesi, istatistiklerin doğru tutulması, bütçenin düzenlenmesi (Elliott J. , 2002: 20) gibi ekonomik ve sosyal pek çok açıdan sektör ve hükümetler için gerekli hale gelmiştir.

Dünya Turizm Örgütü'nün (World Tourism Organization, WTO) 1991'de düzenlemiş olduğu konferansta, turizm için daha önce yapılan tanımlama çalışmaları gözden geçirilmiş; turizm, turist, yolcu kavramları konusunda tanım önerisinde bulunulmuştur (Goeldner ve Ritchie, 2006: 7). Bu tanım günümüzde çoğu akademik çalışma da dâhil (Beech ve Chadwick, 2006: 4-5) kabul görmüş bir tanımdır. Buna göre *turizm*; tatil, iş veya başka amaçlarla sürekli ikamet ettiği yer dışına bir yılı aşmamak şartıyla yapılan seyahatlerdir (WTO, 1994). Bu tanımlamaya göre, bir olayın turizm olarak kabul edilmesi için, aşağıdaki kriterlerin gerçekleşmesi gerekmektedir (Holloway ve Taylor, 2006: 5-8; Page, 2007: 10-13; Elliott J. , 2002: 20-21; Beech ve Chadwick, 2006: 4-5);

- Kişilerin sürekli ikamet ettikleri yerin dışına yolculuk etmeleri,

- Sürekli ikamet ettikleri yer dışına çıkan kişilerin, bunu sürekli kalışa dönüştürmemeleri veya bu kalışın bir yılı aşmaması,
- Kişilerin ikamet ettiği yer dışına seyahat etseler dahi, transit yolcu olmamaları,
- Öğrenci iseler evleri ve okul arasında seyahat etmiyor olmamaları (Goeldner ve Ritchie, 2006: 11),
- Diplomat, sınır görevlisi, asker, konsolosluk görevlisi, geçici ve kalıcı göçmen, mülteci, göçebe olmamaları gerekir.

Bütün bunların tanımlanmaya çalışılmasındaki amaç, yukarıda da belirtildiği gibi turizm ile ilgili olarak yapılan araştırma, istatistik vb. konularda çelişkili durumlara yol açmamaktır. Dikkat edilmesi gereken önemli hususlardan biri turizm istatistikleri içerisine dâhil edilseler bile kurvaziyer yolcuları, günlük ziyaretçiler ve görevlilerin ziyaretçi olarak adlandırılmalarıdır (Holloway ve Taylor, 2006: 5-7). Bu tartışılabilir bir durumdur. Çünkü kurvaziyer yolcuları, günlük ziyaretçiler ve bu araçların personelleri ziyaret ettikleri yerlerde turizm ürünü tüketmekte ve bölgeye ekonomik katkı sağlayabilmektedir. Fakat yine de turist ile ziyaretçi arasında bir ayrıma gidilmiştir.

Bütün bunlardan hareketle *turist* ise şöyle tanımlanabilir, sürekli ikamet ettiği yer dışına bir yılı aşmamak koşuluyla tatil, iş, sağlık, din ve spor gibi eğlence, dinlenme, kültürel amaçlarla seyahat eden, gezip görmeyi hedefleyen kişilere verilen addır. Türkçe’de gezgin veya seyyah bu anlamda kullanılabilir kelimelerdir. Avrupa dillerinden dilimize geçen turizm kelimesi köken olarak Latince (tornare) kökenlidir. Ayrıca Yunanca’da (tornos, τόρνος) aynı manada kullanılmaktadır. Turist ve turizm kelimesi *dönmek*, *çevirmek* anlamına gelen bu sözcükten türemiş (Theobald, 2005: 9), diğer Avrupa dillerine de benzer şekliyle girmiştir.

## 1.2. TURİZM SEKTÖRÜNÜN YAPISI

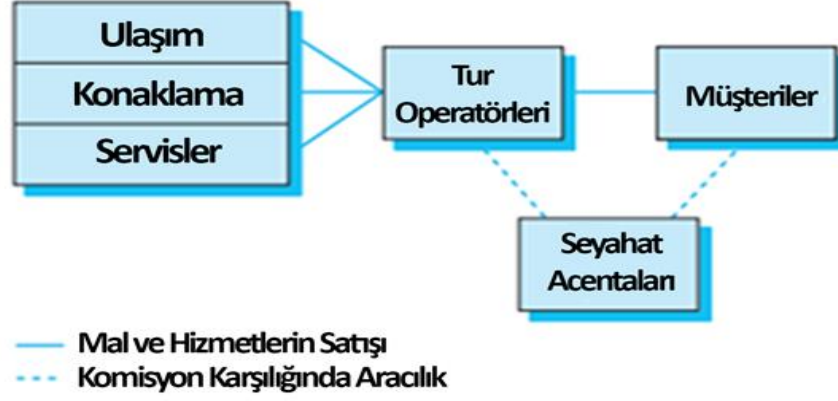
Turizm, günümüz verilerine göre yakıt ve kimyadan sonra dünyanın en büyük ihracat ürünüdür (Inkson ve Minnaert, *Tourism Management: An Introduction*, 2018: 164). Turizm ürünü, sanayi dallarındaki ürünlerden farklı olarak somut bir yapıya sahip değildir. Bu ürün, bir çeşit mal ve hizmetler bütünüdür. Bir hediyelik eşya dükkanındaki tek bir üründen oluşabileceği gibi, turistik bir bölgeye yapılan ve içerisinde ulaşım, yemek

ve konaklama gibi birçok ürün ve hizmeti bir arada barındıran bileşik bir yapıda da olabilmektedir. Bu açıdan düşünüldüğünde turizm yer değiştirmeden dolayı ortaya çıkan eğlenme, dinlenme, kültürel ve dinsel amaçlı tecrübeler, hizmetler, tüketim bütünü olarak düşünülebilir.

Turizm olayında turist geçici olarak yer değiştirmektedir. Bu yer değiştirme boyunca kişinin hem yemek ve konaklama gibi fizyolojik ihtiyaçları karşılanmalı, hem de kişiye dinlenmesini ve geziden keyif almasını sağlayacak aktiviteler sunulmalıdır (Devashish, 2011: 151). Bu ihtiyaçları karşılamaya ya da dinlendirme ve eğlendirmeye yönelik her geçen gün yeni hizmet ve ürünler turizm ürününe dâhil olmaktadır. Bütün bu hizmet ve ürünleri tek bir yapı altında sunmak mümkün olsa da, çoğu zaman tercih edilen bir yöntem değildir. İşlerin benzerliklerine veya uzmanlıklara göre alt sektörler oluşmakta, bu alt sektörlerin oluşturdukları ürün ve hizmetler bir paket ürün haline getirilerek tüketiciye sunulmaktadır.

Turizm sektörü hem yatay hem de dikey olarak geniş bir hiyerarşik yapıya sahiptir. Öyle ki dünyada toplam işgücünün %9.1'i doğrudan veya dolaylı olarak turizm sektöründe istihdam edilmektedir (Fletcher, Fyall, Gilbert ve Wanhill, 2018: 2). Sektörü oluşturan bu işletmeler otel, restoran, seyahat acentası gibi doğrudan turizm ile ilişkili olabilecekleri gibi, inşaat firmaları, turistlere ve sektör çalışanlarına hizmet veren çeşitli iş kolları da olabilmektedir.





Şekil 1.1. Turizm Dağıtım Sistemi (Holloway ve Taylor, 2006, s. 468)

Bütün bu işletmeler hem kendi içerisinde hem de kendi aralarında koordineli bir şekilde çalışmak zorundadır. Bir otel işletmesi konuklarına hizmet verebilmek için önbüro, kat hizmetleri, yiyecek-içecek, satın alma vb. birçok bölümün sistemli bir ilişkisine ihtiyaç duyarken, aynı işletme dışarıda ise, seyahat acenteleri, rehberler, tur operatörleri gibi kurum ve kişilerin işbirliği ile hizmet üretmektedir. Şekil 1.1'de turizm sektöründe alt sektörlerin birbirleriyle ilişkileri ile mal ve hizmetleri tüketiciye ulaştırmaları verilmektedir.

Şekil 1.1'den de anlaşılacağı üzere tur operatörleri veya tur toptancıları paket tur olarak adlandırılan çeşitli hizmet ve ürünleri bir araya getirerek tatil paketi hazırlamaktadırlar. Thomas Cook, daha önce de belirtildiği üzere bunu ilk kez yaparak hem modern turizm anlayışını başlatmış hem de ilk tur operatörü olmuştur (Hamilton J. , 2013). Konunun daha iyi kavranması için ileriki kısımda turizm sektörünü meydana getiren alt sektörler daha detaylı bir şekilde verilmektedir.

### 1.2.1. Tur İşletmeleri

Tur işletmeleri turizm sektörü içerisinde birçok farklı alanda hizmet vermektedirler. Bunlar, tur operatörleri gibi sektördeki birçok hizmeti bir araya getiren büyük işletmeler olabilecekleri gibi, yerel turlar düzenleyen küçük seyahat acenteleri de olabilmektedirler. Tur işletmeleri kavram olarak çoğu zaman seyahat acentası olarak tek bir tanım altında ifade edilmesine rağmen (Ahipaşaoğlu, 2002: 42-45), gerçekte birbirlerinden farklı işler

yapan işletmelerdir. Konunun devamında yaygın olarak sektör içerisinde yer alan bazı tur işletmeleri verilmektedir.

### **1.2.1.1. Tur Operatörleri veya Tur Toptancıları**

Tur operatörleri sektörün çekirdeğini oluşturan işletmelerdir. Tur operatörleri olmadan da insanlar turizm olayını gerçekleştirebilirler fakat kitle turizmini büyük ölçüde tur operatörleri veya benzer bir iş yapan tur toptancıları yüklenmektedir (Sharpley, Travel and Tourism, 2006: 71-72). Tur operatörü ve tur toptancıları çoğu zaman aynı anlamda kullanılmasına karşın (Goeldner ve Ritchie, 2006: 190-191; Holloway ve Humphreys, 2016: 194-195; Medlik, 2003: 165), yazında kavramlar arasındaki farka yönelik çeşitli görüşler mevcuttur.

Tur operatörleri turlar veya transferler esnasında turistlere kendi araç veya çalışanları ile eşlik ederken, tur toptancıları bu işi de diğer birçok iş gibi başka bir sağlayıcıdan satın almaktadırlar (Lubbe, 2005: 237). Sonuç olarak hem tur operatörleri hem de tur toptancıları turizm sektöründe diğer sağlayıcılar tarafından sunulan konaklama, ulaşım, transfer ve eğlence gibi ürün ve hizmetleri bir araya getirerek tek bir ürün olarak tüketiciye, dolaylı veya doğrudan sunan turizm işletmeleridir (Middleton, Fyall, Morgan ve Ranchhod, 2009: 11-12; Holloway ve Humphreys, 2016: 193-195). Diğer bir yaygın tanım, tur toptancısını, bir araya getirdiği turistik ürünlerden paket tur/ürün oluşturan bunu toptan yani başka işletmeler aracılığıyla satan işletmeler olarak tanımlamaktadır (Albustanlıoğlu, 2009). Tur operatörü ve tur toptancısı için şu ayırım yapılabilir. Tur operatörü ulaşım, transfer, konaklama, rehberlik gibi turistik ürünleri bir araya getirerek oluşturduğu tur paketini kendi sahip olduğu perakende acenta şubeleri veya diğer satış kanalları ile tüketiciye satan işletmelerdir. Tur toptancıları ise, benzer şekilde turistik ürünleri bir araya getirerek oluşturdukları paket turları bu işi perakende olarak yapan diğer işletmelere satan, nihai tüketiciyle doğrudan satmayan işletmelerdir.

Bu tanımlardan da anlaşılacağı üzere turizm sektöründe birçok kavramı kesin çizgilerle ayırmak mümkün değildir. Bazı kavramlar veya uygulamalar ülkelere hatta bölgelere göre farklılık gösterebilmektedir.

### **1.2.1.2. Seyahat Acenteleri**

Seyahat acentası kavramı, genellikle başka üreticilerinin hazırladığı ulaşım, konaklama, tur gibi hizmetleri belirli bir komisyon karşılığında satan (Medlik, 2003: 170)

turizm işletmeleri için kullanılmaktadır. Bu kavram ayrıca, kişi ve gruplar için çeşitli turlar düzenleyen turizm işletmeleri için de kullanılmaktadır (Bhatia, 2006). Genel ve yaygın kullanımı perakende olarak turistik ürünü satan genellikle işlek konum ve alışveriş merkezi gibi yerlerdeki temsilcilikler için kullanılsa da, Türkiye’de olduğu gibi özelleştirilmiş çeşitli turlar düzenleyen işletmeler için de bu kavram kullanılmaktadır (Archer ve Syrratt, 2003: 6; Mancini, 2013: 105-112).

Türkiye’de seyahat acentası kavramının tanımı 1618 sayılı Seyahat Acenteleri ve Seyahat Acenteleri Birliği Kanunu ile yapılmıştır. Bu kanunun 1’nci maddesi (e) bendinde seyahat acentası; *“kâr amacı ile turistlere turizm ile ilgili bilgiler vermeye, paket turları ve turları oluşturmaya, turizm amaçlı konaklama, ulaştırma, gezi, spor ve eğlence hizmetlerini görmeye yetkili olan, oluşturduğu ürünü kendi veya diğer seyahat acenteleri vasıtası ile pazarlayabilen ticarî kuruluş”* olarak tanımlanmıştır. Bu tanım hem *seyahat acentası* hem de *tur operatörü* kavramını içine almaktadır. Fakat kanunun 3’ncü maddesi (b) bendindeki B grubu seyahat acenteleri için yapılan *“uluslararası kara, deniz ve hava ulaştırma araçları ile (A) grubu seyahat acentelerinin düzenleyecekleri turların biletlerini satarlar”* şeklinde bir tanım ile bu işi perakende yapan ve tur düzenleyen seyahat acenteleri birbirinden ayrılmaktadır.

### 1.2.1.3. Çevrimiçi Seyahat Acenteleri

Gelişen bilgi ve iletişim teknolojisinin turizm sektörüne kazandırdığı önemli kavramlardan biri çevrimiçi (online) seyahat acenteleridir. Bu gelişme turizm sektöründe önemli yapısal değişimlere neden olmuştur. Öyle ki kısa bir sürede ÇSA’lar sektörün büyük aktörleri haline gelmiştir (Keskinliğiç, Ağca ve Karaman, 2016). E-ticaret kavramının turizm sektörüne yansımaları bu işletmeler olmuştur. Guo, Zheng, Ling ve Yang’ın (2014) yaptıkları çalışmaya göre, 30 büyük otel zincirinin toplam rezervasyonları içerisinde, çevrimiçi rezervasyonlar %57 oranındadır. Sadece 2015 yılında havayolu, otel ve seyahat acentelerinin çevrimiçi satış oranları yaklaşık %6 oranında artmıştır. Havayolu, otel ve seyahat acentelerinin çevrimiçi satışlarındaki bu büyümenin 2020 yılına kadar yaklaşık 820 Milyar doları bulması beklenmektedir (Online travel market - Statistics ve Facts, 2018). Bu hızlı büyüme, ÇSA’ların bu işi doğrudan perakende şubeler aracılığı ile yapan işletmeleri, bir süre sonra ortadan kaldıracağını göstermektedir.

ÇSA'lar tüketicilerin bilgisayar, mobil telefon vb. çevrimiçi yöntemler ile konaklama, ulaştırma ve tatil paketi rezervasyonları yapmalarını olanaklı hale getirmektedir. Bu kolaylık kişilerin internetin bulunduğu her yerden, herhangi bir seyahat acentası bürosuna gitmeden turistik ürünlere ulaşmalarını sağlamaktadır. Geleneksel seyahat acentelerinde basılı çeşitli kataloglar yardımı ile konaklama ve tatil paketleri tanıtımı yapılırken, ÇSA'larda bu iş, web siteleri yardımı ile yapılmaktadır. ÇSA'ların web siteleri, bu ürünler hakkında yazılı ve görsel çokça bilgi barındırmaktadır.

Geleneksel seyahat acentelerinde mümkün olmayan, daha önceki konukların o hizmet hakkındaki görüşlerini öğrenmek, çevrimiçi sistemler sayesinde mümkün olabilmektedir. Bu durum ÇSA'ların geleneksel acentelerin yerine tercih edilmelerindeki en büyük nedenlerden biridir (Keskinkılıç, Ağca ve Karaman, 2016). ÇSA'ların tercih edilme nedeni sadece bununla sınırlı değildir. Pek çok nedenden dolayı diğer e-ticaret sistemleri gibi çevrimiçi olmayan sistemlere tercih edilmektedir. Bunlardan bazıları şu şekilde sıralanabilir (Keskinkılıç, Ağca ve Karaman, 2016; Țugulea, Bobâlcă, Maha ve Maha, 2014: 193);

*Fiyat ve ürünleri karşılaştırabilmek kolaylaşmıştır* (Hobica, 2006); Katalog ile tanıtımı yapılan geleneksel acentalarda, tüketicinin ürünleri fiyat ve özellik olarak birbirleriyle kıyaslaması oldukça güçtür. Bölgelere veya fiyatlara göre düzenlenmiş konaklama tesislerinin bulunduğu broşürler tatile gitmek isteyen müşteriye sunulur, bunlardan birini seçmesi beklenir. Bu kataloglarda tesisleri kıyaslamak çevrimiçi sistemlere kıyasla zordur. Genellikle satış yetkilisinin yardımı ile müşteri tercihini yapmaktadır. Bu seçim işlemine müşterilerin uzun vakitler ayırması gerekmektedir. Tesislerin kataloglara basılan fotoğraf sayısı da kısıtlı sayıdadır. Fakat ÇSA'larda bu işlemler oldukça kolay ve hızlıdır. Veri tabanlarına yüklenmiş olan tesis bilgileri çeşitli kriterlere göre tesislerin sınıflandırılmasını çok hızlı bir şekilde sağlamaktadır. Tatile çıkmak isteyen kişiler bu sitelerdeki çeşitli filtreler yardımı ile bu veri tabanları üzerinde sorgular yapabilmekte, istediği özelliklere göre tesisleri sınırlandırarak tercih yapabilmeyi kolaylaştırmaktadır. ÇSA'ya göre farklılaşmakla birlikte, genellikle, fiyat aralığı, yıldız sayısı, bulunduğu şehir, bölge, havuz, spa, spor salonu, havaalanına uzaklık gibi onlarca kriterlere göre tesis filtre edebilmektedir. Havayolu bilet satış sistemlerinde bu kriterler daha az olmakla birlikte konaklama tesislerindeki gibi önemli kolaylıklar sağlamaktadır.

*Ürünler hakkında daha detaylı bilgi almak mümkündür* (Kuom ve Oertel, 1999); Geleneksel perakendeci seyahat acentelerinde tatil paketi veya konaklama tesisi hakkında bilgiler müşterilere bu bilgilerin basılı olduğu kataloglar yardımı ile verilmektedir (Doswell, 1997: 77). Günümüzde bazı perakendeci acentalarda bu bile artık monitörler ve dijital fotoğraflar yardımı ile yapılmaktadır. Bu şekilde sunulan bilgi konu turizm ürünü gibi tecrübe edilerek tüketilen bir ürün olduğunda, oldukça yetersiz kalabilmektedir. ÇSA'larda genellikle tesis için bir web sayfası tasarlanmakta, tesis hakkında tanıtıcı materyal ve bilgi bu sayfa üzerinde yer almaktadır. Bu web sayfası her ne kadar ÇSA'nın tasarımı doğrultusunda düzenlense de basılı kataloglar kadar sınırlayıcı olmamaktadır. Web'in sağladığı imkânlar sayesinde yüzlerce fotoğraf, odaların şekli, otelin haritadaki konumu, çeşitli tanıtıcı videolar bu sayfalarda yer alabilmektedir. Hatta istenirse, odaların içerisinde sanal bir gezinti yapmak bile mümkündür. Müşteriler haritalandırma servisleri sayesinde, gitmeyi düşündükleri tesisin konumu, doğal çevresi, çevresinde nelerin bulunduğu, başka bir yere olan uzaklığı hakkında da bilgi alabilmektedirler.

*Yorum sistemi sayesinde tesisin önceki konuklarının görüşleri öğrenilebilir* (Gaggioli, Cloudbeds, 2015); Geleneksel seyahat acentelerinde mümkün olmayan hizmetlerden biri de satışa sunulan ürün veya hizmet hakkında daha önce bunu satın almış olan kişilerin görüşlerini öğrenmektir. Bu teorik olarak mümkün olsa da, tercih edilen bir durum değildir. Diğer birçok e-ticaret yöntemi gibi ÇSA'lar için de önemli özelliklerden biri o ürünü daha önce satın almış olan kişilerin o ürün hakkında yorum yapabilmesi, puan verebilmesidir. Bu yorum ve puanlar sayesinde hem oraya rezervasyon yapmak isteyen kişi için bir değerlendirme, hem de tesis yönetici ve çalışanları için olumlu ve olumsuz taraflarını görebilecekleri bir geri bildirim olmaktadır. Yapılan çalışmalar, çevrimiçi satışlarda ürün hakkında daha önce satın alanların yapmış olduğu yorumların oldukça dikkate alındığını göstermektedir (Keskinkılıç, Ağca ve Karaman, 2016; Eddy, 2016).

ÇSA'ları ilk olarak interneti günlük hayata girmesiyle birlikte 1990'ların ortalarında sektörde görülmeye başlamıştır. Bunların öncülleri Sabre, Worldspan GDS gibi küresel dağıtım sistemi şirketleri, Microsoft gibi bilişim şirketleri veya United, Delta, Continental, Northwest gibi havayolu şirketleri olmuştur (Cook ve Billig, 2017: 274-277). Pazarın ilk ÇSA'ları internetin de kurucusu olan ABD'de ortaya çıkmıştır. Bugün dünyada yerel veya küresel pazarda hizmet veren yüzlerce ÇSA bulunmaktadır. Bunlar,

sadece fiyat olarak değil, sundukları hizmetler konusunda da farklılaşma ile pazardaki paylarını artırmaya çalışmaktadırlar.

Airbnb, geleneksel pansiyonculuk sistemine benzeyen, otel olmayan, ev veya oda kiralama hizmeti verirken, Hostelworld.com gibi sistemler öğrenciler gibi düşük fiyatlı kalacak yer arayan kişilere, dünya genelindeki hostellere rezervasyon yapmayı olanaklı hale getirmektedir. Bütün bu sistemlerin belki de en güzel tarafı artık mobil telefonlara kadar girmiş olan internet sayesinde, internetin bulunduğu her yerden rezervasyon yapılabiliyor oluşudur. Android, iOS gibi mobil platformlarda bu sistemlerin neredeyse hepsinin mobil yazılımları da bulunmaktadır. Kullanıcılar bu yazılımlar ile bir bilgisayar ve web tarayıcıya ihtiyaç duymadan rezervasyon sistemine erişebilmekte, doluluk oranları, fiyat değişimleri hakkında bildirim almak gibi birçok faydalı özelliği kullanabilmektedir.

Sojern Global Travel Insights'ın yapmış olduğu bir araştırmada, 2017 yılının son çeyreğinde sadece ABD'de seyahat hakkında yapılan aramaların %26'sının mobil cihazlardan gerçekleştiği görülmüştür. Bu oran dünyanın diğer bölgelerinde daha yüksektir. Avrupa'da %47, Asya'da %40, Orta Doğu ve Afrika'da %38 ve Latin Amerika'da %34 (Kutschera, 2018).

ÇSA pazarı oldukça dinamik bir yapıya sahiptir. Sektörde satın almalar, birleşmeler ve alt markalar sıkça görülmektedir (Gaggioli, Cloudbeds, 2015). 2017 verilerine göre sektörün dört büyük markasından iki büyük aktörü Priceline ve Expedia (Trivago ile birlikte) gelir baz alındığında pazarın %75'ine hakimdir. Sektörün ilklerinden olan Expedia Grup; Expedia.com, Expedia Affiliate Network, trivago, HomeAway, Egencia, Orbitz, Travelocity, Hotwire.com, Wotif Group, CheapTickets, ebookers, CarRentals.com, Classic Vacations, Expedia Local Expert, Expedia CruiseShipCenters, SilverRail, ALICE ve AirAsia Expedia gibi birçok markayı bünyesinde bulundurmaktadır (Prieto, 2017). Priceline ise; Booking.com, Agoda.com, Kayak.com, Cheapflights, Rentalcars.com, Momondo ve OpenTable ile birlikte Booking Holding bünyesi altındaki markaları oluşturmaktadır (Josephs, 2018). Pazardaki bir diğer büyük aktör Çin pazarının en büyük ÇSA'sı olan Ctrip'dir. Ctrip yaklaşık 90 milyon üyesi ve listelediği 70 bin otel ile Asya pazarının en büyük ÇSA'sıdır (Gaggioli, Analysis of Major Online Travel Agencies – OTAs – Update August 2017, 2017). Son olarak pazarda büyük paya sahip

olan şirket Tripadvisor'dur. Tripadvisor diğer ÇSA'lardan farklı olarak bir arama motoru gibi hizmet vermekte, kendi sunduğu fiyatların yanında diğer ÇSA'ların verdiği fiyatları da arama sonuçlarında listelemektedir. Dolayısıyla ziyaretçiler bu tercihleri doğrultusunda Hotels.com, Booking.com ve Agoda.com gibi sitelere yönlendirilmektedir (Delgado, 2018). Ziyaretçi sayısı dikkate alındığında Booking.com, Tripadvisor ve Expedia.com ilk üç sırayı almaktadır (www.alexa.com, 2018).

### **1.2.2. Konaklama İşletmeleri**

Turizm endüstrisinin en temel iş kollarından biri konaklama sektörüdür. Konaklama sektörü denildiğinde ise, akla oteller gelmektedir. Otel, müşterilerine belirli bir ücret karşılığında konaklama ve yeme-içme hizmeti sunan, bu amaç için tasarlanmış; bu hizmetleri sunabilecek kişileri istihdam eden, artan rekabet ve ihtiyaçlar doğrultusunda hizmet portföyünü günün koşullarına uyduran turizm işletmesidir. Günümüzde konaklama işletmeleri birçok ayrı isim ile sınıflandırılırken, günlük kullanımda genellikle sadece otel olarak adlandırılmaktadır. Konaklama tesislerinin sınıflandırılmaları bazen uluslararası adı olurken, bazen de bulunduğu ülkenin yasalarının belirlediği kısıtlara göre belirlenen isim olmaktadır (Holloway ve Humphreys, 2016; Goeldner ve Ritchie, 2006; Beech ve Chadwick, 2006: 378-386; Medlik, 2003: 86).

Seyahat etmenin en temel ihtiyaçlarından biri konaklamadır. Bu nedenle antik çağlardan beri yolcuların konaklayabileceği veya çeşitli ihtiyaçlarını giderebileceği bu tarz tesisler görülmektedir. Roma'da, Osmanlı'da Japonya'da antik dönemlerde günümüzdekinden farklılıkları olsa da gezginler ve tüccarlar için yapılmış bu tarz tesisleri görmek mümkündür (O'Connell J. C., 2017; Freely, 2011: 25).

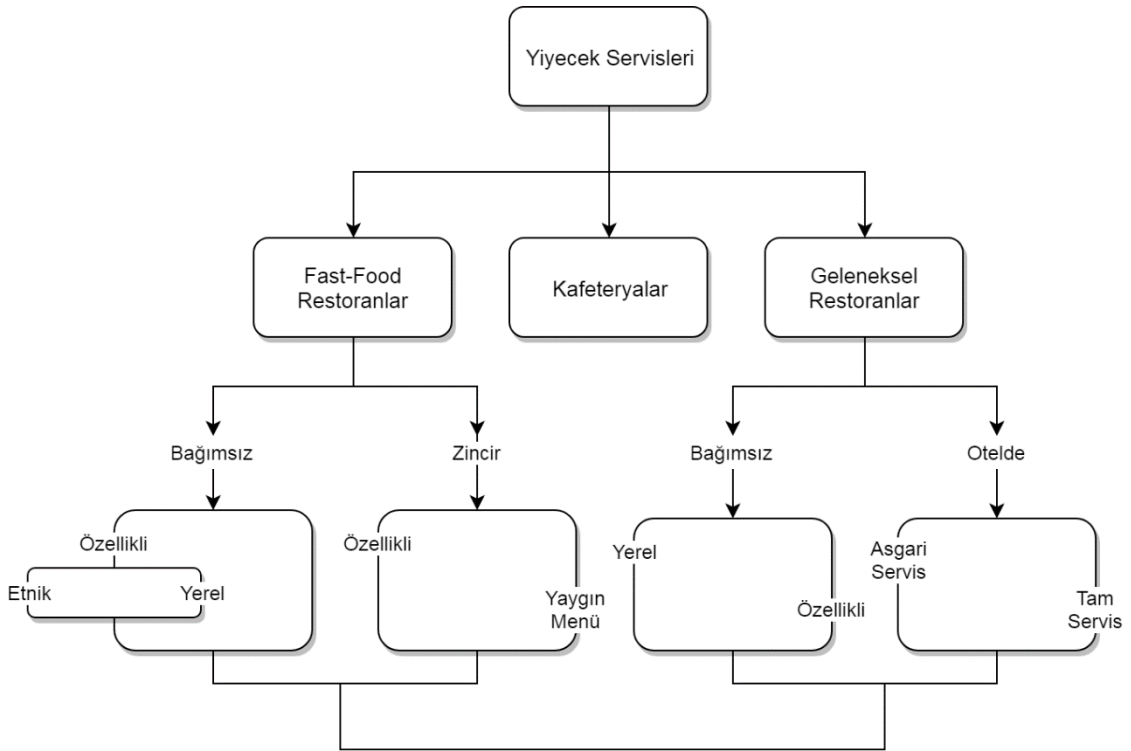
Günümüz otelleri geçmişteki örneklerinden daha karmaşık bir yapıya sahiptir. Konuğun sadece konaklama ihtiyacını değil, otelde kaldığı süre zarfında ihtiyaç duyabileceği daha pek çok ihtiyacını da karşılamaktadır. Bu tesisler sadece konaklama amacıyla değil, tatil ve eğlence amacıyla da kullanıldığından buna hitap edecek çeşitli eğlence, dinlenme ve rekreasyon faaliyetlerini de sunmaktadır. Bu hizmetler günün rekabet şartlarına ve konukların ihtiyaçlarına göre şekillenmektedir.

### **1.2.3. Yiyecek-İçecek İşletmeleri**

Turizm sektörünü oluşturan bir diğer ana sektör yiyecek-içecek sektörüdür. Turistler, yiyecek-içecek ihtiyaçlarını konakladıkları tesiste karşılayabildiği gibi, ziyaret

ettikleri bölgedeki herhangi bir restoran ve benzeri yiyecek-içecek işletmesinde de karşılayabilmektedir. Her iki şekilde de yemekler turistler için önemli bir çekim kaynağıdır (Hall ve Mitchell, 2006; Björk ve Kauppinen-Räisänen, 2017).

Yiyecek-içecek faktörünün turistlerin çekimi ve tutundurulması açısından öneminin farkında olan işletmeler ve kurumlar, pazarlama aracı olarak bunu kullanmaktadırlar (Sparks, Bowen ve Klag, 2003; Boyne, Hall ve Williams, 2003). Şekil 1.2’de yiyecek sektörünün yapısı genel hatlarıyla verilmektedir.



Şekil 1.2. Yiyecek Sektörünün Yapısı (Goeldner & Ritchie, 2006, s. 167)

Şekil 1.2’de de görüleceği üzere turistlere yönelik yiyecek-içecek hizmeti hem otel içerisinde hem de otel dışarısından sağlanabilmektedir. Konuklar otel seçimlerini yaparken otelin sunduğu pansiyon çeşitlerine göre seçim yaparak, farklı yiyecek-içecek hizmetlerinden birini tercih etmektedirler. Konaklama işletmesinde sunulan yiyecek-içecek hizmeti ülkeye veya işletmeye göre farklılık gösterse de, bu işletmeler genel olarak aşağıdaki pansiyon çeşitlerine göre rezervasyon almaktadırlar (Pizam, 2010; Abbot ve Lewry, 1999).



*Sadece Oda (Only Bed, OB):* Ödenen ücretin sadece oda ücreti olduğu, yemek servisinin ekstra ücrete tabi olduğu konaklama şeklidir. “*European Plan*” olarak ta adlandırılmaktadır. Konuk böyle bir rezervasyon yaptığı takdirde ödediği ücret sadece oda ücretini içermektedir. Otel içerisinde yapacağı diğer yeme-içme işleri için ödediği fiyatın üzerine ekstra ücret ödeyecektir.

*Oda ve Kahvaltı (Bed and Breakfast, BB):* Adından da anlaşılacağı üzere ödenen ücretin oda ve sabah kahvaltısını içermesi durumunda kullanılan terimdir. “*Continental Plan*” olarak ta adlandırılmaktadır.

*Yarım Pansiyon (Half Board, HB):* “*Demi Pension*”, “*Half Board*” ve “*Modified American Plan*” olarak adlandırılan bu konaklama şeklinde, ödenen ücret oda, kahvaltı ve bir öğün yemeği kapsamaktadır. Bu öğün genellikle akşam yemeği olmaktadır (Medlik, 2003: 52).

*Tam Pansiyon (Full Board, FB):* 17.yy ABD’deki oteller uzun süreli kalan konuklarına sıklıkla bu sistemi uygulamaktaydılar. Bu nedenle Amerikan otelcilik sisteminde “*American Plan*” olarak adlandırılmaktadır. Bu pansiyon şekli Avrupa otelcilik sisteminde “*En Pension*” veya “*Full Board*” olarak kullanılmaktadır. Bu kalış şeklinde konuğun ödediği fiyata oda ve bütün öğün yemekleri dâhildir (Abbot ve Lewry, 1999).

*Her şey Dâhil (All Inclusive, AI):* Bir standardının olmamasıyla birlikte (Grant, 2016), en az üç öğün yemek, alkolsüz içecekler, çoğu alkollü içeceği içeren pansiyon tipidir (Harrison, 2018). Fransız şirketi Club Méditerranée ve onun kurucusu Gérard Blitz tarafından ilk kez 1950’lerde Mallorca’da uygulanan bu sistem daha sonra aralarında Türkiye’nin de olduğu pek çok ülkede uygulanmıştır (Vanhove, 2017). Günümüzde özellikle sahil otellerinde yaygın şekilde uygulanmaktadır.

Konaklama işletmelerinde genel olarak uygulanan pansiyon çeşitleri bu şekildedir. Bir otelde bu yöntemlerden biri uygulanabileceği gibi, birden çok pansiyon tipi de farklı gruplara uygulanabilmektedir.

Türkiye’deki resort olarak adlandırılan tatile yönelik otellerde genellikle çok öğün içeren her şey dâhil veya tam pansiyon gibi pansiyon tiplerine ağırlık verilirken, şehir otellerinde daha çok oda kahvaltı tarzı bir uygulamanın olduğu görülmektedir (www.otelz.com, 2018).

Otel dışında yiyecek-içecek hizmeti sunan işletmeler oldukça çeşitlidir. Sürekli değişen ihtiyaçlar, zevkler ve çevresel koşullarla birlikte yeni işletme türleri ortaya çıkmaktadır. Bu işletmeler turistler tarafından kullanıldığı gibi yerel halk tarafından da kullanılabilir. Ülkemizde 4533 sayılı Turizm Tesislerinin Belgelendirilmesine ve Niteliklerine İlişkin Yönetmeliği'nin Yeme-İçme ve Eğlence Tesisleri bölümü altında bu tesislerin tanımlanması ve sahip olması gereken özellikler hakkında düzenleme yapılmıştır.

#### **1.2.4. Ulaştırma İşletmeleri**

Turizm bir yer değiştirme faaliyeti olduğundan ulaştırmanın ve bunu sağlayan araçların önemi büyüktür. Bu nedenle motorlu araçların icat edilmesi, uçakların geliştirilmesi ve sivil ulaşım için kullanılması gibi gelişmelerin akabinde turizmin hızla gelişmesi tesadüf değildir.

Ulaştırma işletmeleri turizm sektörünün ve operasyonlarının birçok kısmında kullanılmaktadır. Turistlerin yaşadıkları yerlerden turizm destinasyonuna taşınması, havaalanı, gar gibi yerlerden konaklama tesisine transferler, başka çekim yerlerine yapılan geçici gezilerin transferlerinde, bu ulaştırma işletmeleri devreye girmektedir. Thomas Cook 1841'de bir grup turist ile ilk modern turizm gezisini gerçekleştirirken ulaşım aracı olarak o günün koşullarında en hızlı ve konforlu araç olan treni kullanmıştır (Holloway ve Taylor, 2006: 35). Günümüzde turistik amaçlarla karayolu, havayolu, denizyolu ve demiryolu araçlarının hepsi kullanılmaktadır. Bu ulaşım yöntemlerinin her biri benzer veya farklı amaçlarla turistik seyahatlerde kullanılmaktadır.

### **1.3. TÜRKİYE'DE TURİZM SEKTÖRÜ**

Bir bölgenin turist çekebilmesi için çeşitli özelliklere sahip olması gerekmektedir. Bu özellikler ister yapay olsun isterse doğal, turistlerin merak duygularını harekete geçirmeli, onlarda ziyaret etme isteği uyandırmalıdır. Bu istek, o bölgeyi görmek için harcanacak maddi ve fiziki çabanın üzerinde olmalıdır.

Turizm, yapılan bölge için önemli ekonomik getirileri olan bir sektördür. Fazla bilgi ve teknoloji altyapısı gerektirmeden yapılabilmesi, hizmet ağırlıklı olması, dolayısıyla üretim için çok hammaddeye ihtiyaç duymaması, pek çok bölge için cazip bir sektör olarak görülmesine neden olmaktadır. Fakat bahsedildiği gibi turist çekmek için bir yerin,

potansiyel turistlerin gözünde, oraya ulaşmak için gerekli olan çabaya değmesi gerekmektedir.

Türkiye önemli sayılabilecek turistik çekiciliğe sahip ülkelerden biridir. Bundan dolayıdır ki Dünya’da en çok turist çeken on beş ülke arasına girmektedir (International Tourism, Receipts, 2017). Türkiye birçok medeniyete ev sahipliği yapmış dolayısıyla sahip olduğu kültürel değerleri, uzun güneşlenme mevsimi, kıyıları ve daha birçok özelliği ile turizm pastasında önemli bir paya sahiptir.

Aşağıda turizm talebine neden olan çekim faktörleri önem sırası gözetilmeden verilmektedir (Ateşoğlu ve Bayraktar, 2011; Birdir ve Akgöl, 2015; Yalav-Heckerroth, 2017; Waltner, 2018; Elliott H. , 2013; Lee, 2016).

*Gelişmiş Avrupa ülkelerine yakındır;* Endüstri devriminin başladığı yer olan Avrupa, aynı zamanda modern turizm anlayışının da başladığı yerdir. Turizm olayının gerçekleşmesi için gerekli olan sosyo-ekonomik şartlar ilk olarak Avrupa’da oluşmuştur. Geçmişten günümüze Dünya Savaşları gibi kesintiler dışında bu şartların gelişmesi devam etmiştir. Avrupa ülkeleri dünyada gelişmişlik ve refah seviyesi bakımından ilk sıralarda yer almaktadır. Bu durum turizm için önemlidir. Turizme katılacak kişiler belirli miktarda boş zaman ve maddi olanaklara sahip olmalıdır. Boş zamanın gelişmemiş ülkelerde daha fazla olacağı düşünülse de sosyal güvenlik ve sosyal demokrasinin gelişmediği bu ülkelerde çalışma koşulları çok daha ağırdır. Bu nedenle Avrupa’da sanayi devriminin akabinde kazanılan bu haklara Dünya’da hala bazı gelişmemiş ülkelerin vatandaşları sahip değildir. Bütün bu faktörler önemli bir turizm pazarı olan Avrupa’ya yakın olan Türkiye için önemli bir turizm çekim merkezi olması konusunda avantaj sağlamaktadır.

*Önemli tarihi ve arkeolojik eserlere sahiptir;* Türkiye dünyada ilk yerleşik hayata geçildiği tahmin edilen Mezopotamya ve Anadolu toprakları üzerinde kurulmuştur. Hititler, Sümerler, Asurlar gibi tarihin ilk devletleri bu topraklar üzerinde hüküm sürmüştür. Hüküm sürdükleri dönemde dünyanın en gelişmiş devletleri olarak nitelendirilebilecek Yunan site devletleri, Romalılar, Bizans ve Osmanlılar yine bu topraklarda hüküm sürmüştür. Bu uygarlıklar günümüzde keşfedilen veya edilmeyen binlerce eser bırakmışlardır. Buralarda yaşamış olan bilim adamlarının, düşünürlerin batı dünyası üzerinde önemli etkileri olmuştur. Bütün bunlar kültür turizmi açısından çok

önemli kaynaklardır. Her yıl binlerce insan bunlardan dolayı Türkiye'yi ziyaret etmektedir.

*Kıyılar ve uygun iklim koşulları;* Modern turizmin başladığı 19.yy sonralarından beri kıyılar ve plajlar önemli turizm ve eğlence yerleridir (Holloway ve Humphreys, 2016: 44). Türkiye üç tarafı denizle çevrili yarım ada şeklinde bir ülkedir. Bu kıyılardan özellikle uzun güneşlenme mevsimine sahip Akdeniz ve girintili yapıda olan Ege kıyıları kıyı ve deniz turizmi açısından oldukça elverişlidir. Kuzey Avrupa ülkelerinin birçoğu deniz kıyısına sahip olsa da güneşli gün sayısının azlığı, soğuk denizler, deniz turizmi açısından uygun koşulları sağlamamaktadır. Öte yandan Baltık Denizi kıyılarında hobi balıkçılığı, yelkencilik, su sporları gibi aktiviteler yapıyor olsa da, benzer şekilde deniz suyu sıcaklığı yılın büyük bölümünde insanların yüzmesi için çok uygun değildir. Bu nedenle Rusya ve Kuzey Avrupa ülkelerinden her yıl milyonlarca turist kıyı turizmi için Türkiye'ye gelmektedir. Türkiye ekvatora olan yakınlığı nedeniyle özellikle Akdeniz kıyılarında yılın altı ayında insanların yüzebileceği uygun deniz suyu sıcaklığına sahiptir.

*Siyasi istikrar ve güvenlik;* Bir bölgede turizm yapılabilmesi için gerekli olan faktörlerin başında belki de siyasi istikrar ve güvenlik gelmektedir. Can ve mal güvenliği olmayan bir ülke, turistler için cezbedici fiziksel özelliklere sahip olsa bile çekici olmayacaktır. Buna en iyi örnek yakın geçmişte Kuzey Afrika'da yaşanan toplumsal olaylar gösterilebilir (Ghanem, 2016). Türkiye'dekine benzer turizm arzı gerçekleştiren bu ülkelerdeki güvenlik sıkıntısı, bu ülkelerin talebini İspanya ve Türkiye gibi rakiplerine kaydırmıştır (Perles-Ribes, Ramón-Rodríguez, Moreno-Izquierdo ve Martí, 2018). Türkiye'nin bulunduğu bölgede, benzer fiziksel ve iklimsel özelliğe sahip birçok ülke olsa da, güvenlik ve hukuk açısından birçok rakibinin önüne geçmektedir. Türkiye'de sorun yaşayan bir turist uluslararası anlaşmalarla garanti altına alınan yasalar çerçevesinde hakkını arayabilmektedir.

*Altyapı ve tesis olanakları;* Türkiye sahip olduğu tarihi ve doğal güzellikler ile 20.yy ortalarından itibaren hem ülke idarecilerinin hem de batılı turistlerin ilgisini çekmiştir (Gürel, 2018). Turizm, kalkınma planları döneminde yatırım yapılması ve geliştirilmesi gereken sektörlerden biri olarak belirlenmiştir (Dinçer ve Çetin, 2015). O dönemde gerekli olan altyapı eksikliğinin giderilmesi ve tesisleşme hedefler arasında yer almıştır. Fakat turizm sektöründe tesisleşme adına asıl büyük gelişmeler 1980'lerde

başlamıştır. Bunda turizm teşvik kanununun etkisi büyüktür (Şanlıoğlu ve Özcan, 2017). Bugün Türkiye’de hatırı sayılır nitelikte ve nicelikte turizm altyapısı ve tesisi bulunmaktadır. Kırk yıla yaklaşan tecrübe ile birlikte bu işletmeler her yıl dünyanın çeşitli yerlerinden milyonlarca turisti ağırlamaktadır. Sahip olunan bu imkânlar turizmin yapılabilirliği açısından oldukça önemlidir.

*Kültür çeşitliği;* Türkiye birçok ülkenin sahip olmadığı bir kültür çeşitliliğine sahiptir. Bu çeşitlilik birçok kültürel öğeye yansımaktadır. Özellikle yemek kültürü sonradan geliştirilen mutfakların aksine, geçmiş medeniyetlere ve ülkedeki kültür çeşitliliğine dayanmaktadır. Türkiye mutfağı turistleri ülkeye çeken en önemli öğelerdendir.

### **1.3.1. Türkiye’de Turizmin Gelişimi**

Türkiye ve özellikle İstanbul 16.yy’dan sonra batılı seyyahların durak yerlerinden olmuştur (Hentsch, 1992: 84-86). Şehrin sahip olduğu tarihsel miras, doğu ve batının kültürel sentezi olması, batılıların ziyaretlerindeki nedenlerden olmuştur (MacLean, 2004). Osmanlı’nın son dönemlerinde ise Avrupa’daki çağdaşları gibi devletin aristokrat sınıfı arasında deniz sahilindeki yazlık yerleri ziyaret etme davranışları görülmektedir (Kervankıran ve Çuhadar, 2017: 3). Cumhuriyet döneminde ise Atatürk’ün isteği ile daha sonra adı Türkiye Turing ve Otomobil Kurumu olan, Türkiye Seyyahin Cemiyeti kurulmuştur. Kurum 1930’larda gümrük, trafik mevzuatı gibi konularda özel yetkiler olarak faaliyetlerini sürdürmüştür. Ayrıca Türkiye’yi tanıtıcı ilk broşürler, ilk turistik rehberler ve karayolları haritası gibi faaliyetler bu kurum tarafından gerçekleştirilmiştir.

İkinci Dünya Savaşının sona ermesinden sonra Dünya’da turizmin tekrar canlanmasına paralel olarak Türkiye’de de çeşitli faaliyetler gerçekleştirilmiştir. 1953 yılında turizmi desteklemek için Turizm Endüstrisi Teşvik Kanunu yürürlüğe girmiştir. Akabinde, özel sektörün gelişmediği ülkede, sektörün ihtiyaç duyduğu sermayeyi sağlamak için 1955’te T.C Turizm Bankası A.Ş kurulmuştur. Yine aynı yıl Emekli Sandığı’nın yatırımı ile yapımına 1952’de başlanan ünlü Hilton otel zincirinin İstanbul şubesi hizmete girmiştir. Batı otelcilik anlayışına göre tasarlanan ve hizmet veren bu işletme, yetiştirdiği pek çok personel ile daha sonra açılacak olan işletmelere kalifiye personel sağlamıştır (Özdemir, 2013).

Bu dönemde devlet içerisinde turizmle ilgili teşkilatlanmada da düzenlemelere gidilmiştir. 1949-1957 tarihleri arasında Başbakanlığa bağlı çalışan ve turizm konusunda ülkedeki en yetkili organ olan Basın-Yayın ve Turizm Genel Müdürlüğü 1957’de Basın-Yayın ve Turizm Bakanlığı’na dönüştürülmüştür (Evcin, 2017). Bu bakanlık 1963’te Turizm ve Tanıtım bakanlığı olacaktır (Ormankıran, Bayram ve Karaçar, 2015). Kabul edilen yeni anayasayla birlikte 1961’den itibaren ülke *Planlı Kalkınma Dönemi* olarak adlandırılan ve 5’er yıllık kalkınma planlamaların yapıldığı bir döneme girmiştir (Akça, 2016: 722). Bu kalkınma planları içerisinde turizm sektörü de dâhil edilmiştir.

1982’de kabul edilen 2634 sayılı Turizmi Teşvik Yasası Türkiye turizmi için önemli dönem noktalarından biridir. Liberalleşme eğiliminde olan mevcut dönem hükümeti tarafından *24 Ocak Kararları* olarak adlandırılan ekonomik dönüşüm sonrasında çıkartılmıştır (Saygılıoğlu, 2018). 24 Ocak Kararları ile devletin ekonomi içerisinde payını azaltacak önlemler alınmıştır. Döviz kullanımının serbestleştirilmesi, ithalatta kolaylık sağlanması, üreticilerde gümrük muafiyeti gibi mevcut ekonomik yapıyı değiştirecek kararlar alınmıştır (Çelik H. , 2006: 60-81). 2634 sayılı Turizmi Teşvik Yasası’nda da bunun etkileri görülmektedir.

İlk halinde yer almamasına rağmen 1985-1989 yılları arasında hibe türü teşviklerin verilmesi sektördeki yatırımların artmasında önemli rol oynamıştır (Şanlıoğlu ve Özcan, 2017: 101). Bu teşviklerden biri turizm işletmelerinde kullanılan makine ve teçhizatın fiyatının belirli bir miktarının devlet tarafından işletmeye geri ödenmesidir. 1984’te uygulamaya konulan bu teşvik oranı %6 ile başlamış sonraki yıllarda %25’e kadar çıkmıştır (Şahin, 1989: 310). Bu kanun sonrasında sektör yatırımlarına gümrük muafiyeti, yatırım indirimi, çalışanların ücretlerinin vergilendirilmesinde indirim, yatırım finansman fonundan faydalanma, bina inşaat harcı istisnası, vergi resim ve harç istisnası, döviz tahsisi, K.D.V ertelemesi, arsa ve arazi tahsisi, ithal hakkı, teşvik primi ve teşvik kredisi gibi destekler verilmiştir (Önen, 2000: 20-21). Turizm sektörü 1985’te *kalkınmada özel önem taşıyan sektör* kapsamına alınmıştır. Bunun devamında yatırımların daha da arttığı görülmektedir (Toker, 2007). Türkiye’de turizmin gelişimi turizm arzına göre değerlendirildiğinde **Tablo 1.1** ortaya çıkmaktadır. 2634 sayılı Turizmi Teşvik Kanunu öncesinin de dâhil olduğu 1970-2018 yılları arasında Türkiye’deki *Turizm İşletme Belgeli* konaklama tesis ve yatak sayısı **Tablo 1.1**’de verilmektedir.

**Tablo 1.1.** Türkiye'deki Turizm İşletme Belgeli Konaklama Tesislerinin Sayısı

Yıl	Tesis Sayısı	Oda Sayısı	Yatak Sayısı	Yıl	Tesis Sayısı	Oda Sayısı	Yatak Sayısı
1970	292	15.243	28.354	1994	1.729	128.065	265.136
1971	337	17.282	32.114	1995	1.793	135.436	286.463
1972	363	18.578	34.628	1996	1.866	145.493	301.524
1973	388	20.517	38.528	1997	1.933	151.055	313.298
1974	400	21.699	40.895	1998	1.954	151.397	341.215
1975	421	23.860	44.957	1999	1.907	153.749	319.313
1976	439	24.983	47.307	2000	1.824	156.367	325.168
1977	446	26.496	50.379	2001	1.998	177.371	368.819
1978	473	27.233	52.385	2002	2.124	190.327	396.148
1979	494	28.013	53.956	2003	2.240	202.339	420.697
1980	511	28.992	56.044	2004	2.357	217.664	454.290
1981	529	30.050	58.242	2005	2.412	231.123	483.330
1982	569	32.011	62.372	2006	2.475	241.702	508.632
1983	611	33.694	65.934	2007	2.514	251.987	532.262
1984	642	34.666	68.266	2008	2.566	268.633	567.470
1985	689	41.351	85.995	2009	2.625	289.383	608.765
1986	731	44.342	92.129	2010	2.647	299.621	629.465
1987	834	51.040	106.214	2011	2.783	319.319	668.829
1988	957	58.914	122.306	2012	2.870	336.447	706.019
1989	1102	70.603	146.086	2013	2.982	357.440	749.299
1990	1260	83.953	173.227	2014	3.131	384.454	807.316
1991	1404	97.260	200.678	2015	3.309	404.462	850.089
1992	1498	148.017	219.940	2016	3.641	426.981	899.881
1993	1581	113.995	235.238	2017	3.770		935.000
				2018	3.925		974.574

\***Kaynak:** (Turizm İstatistikleri, 2019; Türkiye Otel İstatistikleri, 2015)

**Tablo 1.1'**de sadece *turizm işletme belgeli* tesisler verilmektedir. Bakanlık tarafından yatırım aşamasında verilen *turizm yatırım belgesi* ve işletme belgesi yerel yönetimler tarafından verilen *belediye belgeli* tesisler tabloda yer almamaktadır.

### 1.3.2. Türkiye'de Turizm Sektörü

Türkiye bahsedildiği üzere turizm arzında kullanabileceği önemli doğal, tarihi ve kültürel öğelere sahip bir ülkedir. Kitle turizminin henüz başlamadığı yıllarda bu öğelerden haberdar turistler tarafından bireysel olarak ziyaret edilmekteydi. 1950'lerden sonra Dünya'da turizmin gelişmesine paralel olarak Türkiye'de gerek kamu gerekse özel sektör tarafından turizm arzı sürekli artış göstermiştir. Günümüzde en çok turist çeken ilk 15 ülke arasında yer almaktadır (International Tourism, Receipts, 2017). Kalkınma planı

döneminde turizm yatırımları yapılırken turizm bölgesi olarak belirlenen yerlere ağırlık verilmesi konusunda karar alınmış, sonraki dönemlerde bu karara sadık kalınmıştır. Böylece özellikle denize kıyı Akdeniz ve Ege bölgelerinde tesisçe zengin turizm bölgeleri oluşmuştur. İstanbul gibi zengin kültürel ve tarihi öğelere sahip şehirler de hem marka olarak hem de turizm arzı olarak önemli konumdadır.

Turizm dinamik yapıya sahip bir sektördür. Sektörleştiği günden günümüze temelinde dinlenmek, eğlenmek gibi faaliyetler bulunsa da zamanla farklı türde pek çok turizm çeşidi ortaya çıkmıştır. Türkiye geniş yüzölçümü, barındırdığı farklı kültür ve tarihi zenginlikleri ile dünyada bilinen birçok turizm çeşidinin uygulandığı bir ülkedir.

Ege ve Akdeniz kıyıları hem uzun güneşlenme mevsimi hem de antik dönemlerden kalma zengin tarihi dokusuyla *kıyı turizmi* ve *kültür turizmi* açısından önemli çekim merkezleridir. Sadece Antalya'ya 2018'de gelen yabancı ziyaretçi sayısı yaklaşık 12,5 milyon kişidir (www.turob.com, 2018). İstanbul ise, sahip olduğu tarihi öğeleri, Doğu Roma ve Osmanlı gibi devrinin önemli devletlerine başkentlik yapması, Batı ve Doğu kültürünü birleştiren yapısı, Anadolu'daki onlarca kültürün bir arada yaşadığı önemli bir kültür turizmi merkezidir. İstanbul'a gelen yabancı ziyaretçi sayısı 2018'de yaklaşık 12,3 milyon kişidir (Çelik A. , 2018).

Türkiye, büyük bölümünde kış mevsiminde kar yağışının görüldüğü dağlık bir ülkedir. Bu dağlardan bazıları, kayak ve kar sporları, dağcılık, trekking, yayla turizmi gibi birçok turizm çeşidine ev sahipliği yapmaktadır.

Turizm sektörü iki büyük alt sektörden oluşmaktadır. Bunlar *konaklama sektörü* ve *tur/seyahat sektörüdür*. Ulaştırma ve yiyecek-içecek gibi sektörler önemli olmakla birlikte konaklama ve tur/seyahat sektörünün faydalandığı yan sektörlerdir.

### 1.3.2.1. Türkiye'de Konaklama Sektörü

İfade edilen bütün bu turizm çeşitleri turistlerin konaklayacağı ve günlük ihtiyaçlarını karşılayacağı tesislere ihtiyaç duymaktadır. Turizmin maliyet unsurlarının başında altyapı ve bu tesisler gelmektedir. **Tablo 1**'de bu tesislerin sadece *turizm işletme belgeli* olanlarına yer verilmiştir. Konaklama işletmeleri bununla sınırlı değildir. Bakanlık tarafından yatırım aşamasında verilen *turizm yatırım belgeli* tesisler, yerel yönetimler tarafından işletme izni verilen *belediye belgeli* tesisler de bulunmaktadır. **Tablo 1.2**'de bu üç tesis statüsüne ait tesis sayıları verilmektedir.



**Tablo 1.2.** Türkiye'deki Konaklama Tesislerinin Statüsü ve Sayısı (2002-2018)

Yıl	Bakanlık				Belediye Belgeli		Toplam	
	Yatırım Belgeli		İşletme Belgeli		Tesis Sayısı	Yatak Sayısı	Tesis Sayısı	Yatak Sayısı
	Tesis Sayısı	Yatak Sayısı	Tesis Sayısı	Yatak Sayısı				
2002	1.138	222.876	2.124	396.148	7.772	408.005	9.896	804.153
2003	1.130	242.603	2.240	420.692	7.637	399.369	9.877	820.061
2004	1.151	259.424	2.357	454.290	7.514	392.582	9.871	846.872
2005	1.039	278.255	2.412	483.330	7.494	392.338	9.906	875.668
2006	869	274.687	2.475	508.742	7.495	391.228	9.970	899.970
2007	776	254.191	2.514	532.262	7.073	399.110	9.587	931.372
2008	772	258.287	2.566	567.470	7.064	397.684	9.630	965.154
2009	754	231.456	2.625	608.765	7.115	402.289	9.740	1.011.054
2010	877	252.984	2.647	628.565	9.185	527.712	11.832	1.156.277
2011	922	267.900	2.783	668.829	8.893	504.877	11.676	1.173.706
2012	960	273.877	2.870	706.019	8.988	512.462	11.858	1.218.481
2013	1.056	301.862	2.982	749.299	9.196	497.728	12.178	1.247.027
2014	1.113	308.944	3.125	803.664	9.188	496.697	12.313	1.300.361
2015	1.125	314.194	3.289	850.089	9.187	496.574	12.476	1.346.663
2016	1.135	312.912	3.641	899.881	9.186	496.538	12.827	1.396.419
2017	1.051	263.033	3.770	935.000	7.607	506.934	11.377	1.441.934
2018	981	225.421	3.925	974.574	7.671	511.076	11.596	1.485.650

\***Kaynak:** (Turizm İstatistikleri, 2018: 14)

Tablo incelendiğinde 2018 verilerine göre Türkiye 11596 konaklama tesisine ve yaklaşık 1.485 bin yatak kapasitesine sahiptir. Türkiye turizm sektöründen 2018 verilerine göre yaklaşık 29,5 Milyar \$ gelir elde etmiştir (Turizm İstatistikleri, 2018: 10).

Turizm Bakanlığı belgelendirdiği tesisleri sahip olduğu çeşitli özelliklere göre sınıflandırmaktadır. Bir kısmı uluslararası isimlendirme ile örtüşen bu sınıflandırma yöntemi tesis özellikleri, bulunduğu yer, amacı hakkında bazı bilgiler vermektedir. Tablo 1.3'de *turizm belgeli* konaklama tesisleri ve türleri hakkında bazı istatistikler verilmektedir.

**Tablo 1.3.** Türkiye'deki Turizm Belgeli Konaklama Tesisleri ve Türleri

<i>Tür</i>	<i>Sınıfı</i>	<i>Turizm Yatırım Belgeli</i>			<i>Turizm İşletme Belgeli</i>		
		<i>Tesis Sayısı</i>	<i>Oda Sayısı</i>	<i>Yatak Sayısı</i>	<i>Tesis Sayısı</i>	<i>Oda Sayısı</i>	<i>Yatak Sayısı</i>
<i>Toplam</i>		981	104.910	225.421	3.925	464.927	974.574
<i>Otel</i>	Toplam	787	84.562	178.730	2.941	392.658	818.011
	5 Yıldızlı	147	40.861	89.189	633	205.280	436.075
	4 Yıldızlı	237	26.427	54.688	810	112.611	232.463
	3 Yıldızlı	302	15.191	30.668	1.045	59.156	118.996
	2 Yıldızlı	39	1.036	2.115	401	14.166	27.682
	1 Yıldızlı	62	1.047	2.070	52	1.445	2.795
<i>Motel</i>	Toplam	1	59	118	6	230	467
	-	1	59	118			
	2. Sınıf				1	65	134
	Motel				5	165	333
<i>Tatil Köyü</i>	Toplam	31	7.809	18.153	78	28.085	61.697
	1. Sınıf	14	3.310	7.167	68	26.520	58.415
	2. Sınıf	17	4.499	10.986	10	1.565	3.282
<i>Termal Otel</i>	Toplam	21	6.698	14.890	88	14.047	29.713
	5 Yıldızlı	12	5.459	12.027	39	9.492	20.169
	4 Yıldızlı	5	856	1.932	30	3.370	7.068
	3 Yıldızlı	4	383	931	19	1.185	2.476
<i>Termal Tatil Köyü</i>	Toplam	1	129	310			
	5 Yıldızlı	1	129	310			
<i>Termal Müstakil Apart Otel</i>	Toplam				2	81	171
	-				2	81	171
<i>Pansiyon</i>	Toplam	1	50	100	108	1.344	2.717
	-	1	50	100	108	1.344	2.717
<i>Kamping</i>	Toplam	7	929	2.144	7	847	2.425
	-	7	929	2.144	7	847	2.425
<i>Oberj</i>	Toplam				3	408	890
	-				3	408	890
<i>Apart Otel</i>	Toplam	1	128	256	187	7.466	17.250
	-	1	128	256	187	7.466	17.250
<i>Özel Tesis</i>	Toplam	29	736	1.473	377	11.631	23.787
	-	29	736	1.473	377	11.631	23.787
<i>Müstakil Apart Otel</i>	Toplam				9	158	328
	-				9	158	328
<i>Golf Tesisi</i>	Toplam	1	60	228	3	531	1.348
	-	1	60	228	3	531	1.348

**Tablo 1.4.** Türkiye'deki Turizm Belgeli Konaklama Tesisleri ve Türleri (Devam)

<i>Turizm Kompleksi</i>	Toplam	1	300	912	4	3.164	6.668
	-	1	300	912	4	3.164	6.668
<i>Butik Otel</i>	Toplam	85	2.798	5.873	92	3.559	7.191
	-	85	2.798	5.873	92	3.559	7.191
<i>B Tipi Tatil Sitesi</i>	Toplam	1	298	1.526	3	266	984
	-	1	298	1.526	3	266	984
<i>Butik Tatil Villası</i>	Toplam				1	70	140
	-				1	70	140
<i>Dağ Evi</i>	Toplam	6	204	440	3	129	341
	-	6	204	440	3	129	341
<i>Çiftlik Evi / Köy Evi</i>	Toplam	6	98	164	9	68	132
	-	6	98	164	9	68	132
<i>Yayla Evi</i>	Toplam	2	52	104	2	30	90
	-	2	52	104	2	30	90
<i>Hostel</i>	Toplam				1	123	166
	-				1	123	166
<i>Butik Termal Otel</i>	Toplam				1	32	58
	-				1	32	58

**\*Kaynak:** (Turizm İstatistikleri Yıllık Bültenler, 2018)

Tablo 1.3 incelendiğinde Türkiye’de bakanlık tarafından belgeli toplam 4906 konaklama tesisinin olduğu görülmektedir. Bu tesisler *Turizm Tesislerinin Belgelendirilmesine ve Niteliklerine İlişkin* yönetmeliğe göre sınıflandırılmaktadır.

### 1.3.2.2. Türkiye’de Tur/Seyahat Sektörü

Turizm sektörünün diğer önemli kısmını tur/seyahat sektörü oluşturmaktadır. Yazında genellikle seyahat sektörü olarak ifade edilen bu sektör ulaştırma sektörü ile karıştırılmaması için bu şekilde ifade edilmiştir. Bundan sonra seyahat sektörü olarak ifade edilecek olan bu kavram ulaştırma sektörünü kapsamamaktadır. Ulaştırma sektörünün durumu hakkında bu bölümde ayrıca bilgi verilecektir.

Seyahat sektörü seyahat acenteleri, tur operatörleri, tur toptancılarını kapsayan paket tur ürününü bir araya getiren veya perakende olarak pazarlayan, çeşitli turlar düzenleyen, paket tur operasyonlarını gerçekleştiren işletmelerden oluşmaktadır (Bkz: Tur İşletmeleri).

Türkiye'nin turizm ve tur alanında ilk kurumlarından biri, Türkiye Turing ve Otomobil Kurumu olmuştur. Atatürk'ün talimatı ile 6 Kasım 1923'de kurulan kurum, önce *Türk Seyyahin Cemiyeti* sonra da *Türkiye Turing Klöbü* adlarını almıştır (Bayer, 2012; Turing Tarihçe, 2019). Bu kurum, 1930 yılında kamu yararına çalışan dernek olarak tanınmış, gümrük ve trafik mevzuatı içerisinde, özel yetkiler ve görevler almıştır.

Turizmin gelişmeye başlamasından sonra seyahat, turist taşıma, sigorta işleri, sorumluluklar ve denetimler ile ilgili olarak yeni bir düzenleme gerekli hale gelmiş ve 1972 yılında 1618 sayılı *Seyahat Acenteleri ve Seyahat Acenteleri Birliği Kanunu* çıkartılmıştır. Türkiye'de çeşitli güncellemelerle birlikte günümüze kadar seyahat acenteleri ile ilgili düzenlemeler bu kanuna göre yapılmaktadır. 1618 sayılı kanun ayrıca Türkiye Seyahat Acenteleri Birliği'nin (TÜRSAB) kuruluşunu sağlamıştır. Kanunun 32'nci maddesinde TÜRSAB'ın kuruluş amacıyla ilgili olarak *seyahat acenteleri, seyahat acenteciliği mesleğinin yurt ekonomisi ve turizmine uygun surette gelişmesini sağlayıcı tedbirleri almak ve mesleki ahlak ve tesanütü korumak amacıyla Seyahat Acenteleri Birliği adı altında tüzel kişiliği haiz bir birlik kurarlar. Seyahat acentelerinin bu birliğe üye olmaları zorunludur, bilgisi verilmektedir.*

1618 sayılı kanun faaliyet gösteren seyahat acentelerinin sınıflandırılması konusunu da netleştirmiştir. Kanuna göre seyahat acenteleri üç sınıfa ayrılmaktadır. Bunlardan A grubu seyahat acenteleri; *kâr amacı ile turistlere turizm ile ilgili bilgiler vermeye, paket turları ve turları oluşturmaya, turizm amaçlı konaklama, ulaştırma, gezi, spor ve eğlence hizmetlerini görmeye yetkili olan, oluşturduğu ürünü kendi veya diğer seyahat acenteleri vasıtası ile pazarlayabilen ticarî kuruluş*, olarak ifade edilmektedir. B grubu seyahat acenteleri; *uluslararası kara, deniz ve hava ulaştırma araçları ile (A) grubu seyahat acentelerinin düzenleyecekleri turların biletlerini satan işletmeler* olarak tanımlanmıştır. Bu tanıma göre B grubu seyahat acenteleri *perakendeci seyahat acenteleri* olarak ifade edilebilir. Kanuna göre C grubu seyahat acenteleri ise; *yalnız Türk vatandaşları için yurt içi turlar düzenleyebilen, seyahat işletmeleridir*. Kanunun 3'ncü maddesinde; *(B) ve (C) grubu seyahat acenteleri kendi hizmetleri dışında kalan diğer seyahat acenteciliği hizmetlerini göremezler. Ancak kendilerine (A) grubu seyahat acentelerinin verecekleri hizmetleri görürler*, ifadesi B ve C grubu acentelerin A grubu acentelerin kendilerine yetki tanınan işler dışında bir iş yapamamakla birlikte, A grubu acentelerin bazı işlerini onlar için yapabilecekleri anlamı çıkmaktadır. Bu dış kaynak

kullanımı sektör içerisinde çeşitli işler için yapılmaktadır. Türkiye’de 2018 verilerine göre 10297 seyahat acentesi bulunmaktadır. Bu acentelerin gruplara göre dağılımı **Tablo 1.5**’de verilmektedir.

**Tablo 1.5.** Türkiye'deki Seyahat Acenteleri ve Grupları

<i>Grup</i>	<i>Merkez</i>	<i>Şube</i>	<i>Toplam</i>
A	7672	2436	10108
B	61	3	64
C	107	18	125
Toplam	7840	2457	10297

\***Kaynak:** (Turizm İstatistikleri, 2018: 22)

**Tablo 1.5** incelendiğinde Türkiye’deki seyahat acentelerinin büyük çoğunluğunun A grubu seyahat acentesi olduğu görülmektedir.

Turizm sektörü içerisinde önemli konumda yer alan bir diğer sektör *ulaştırma* sektörüdür. Turizm olayı çoğu zaman sürekli ikamet edilen yerin dışına olduğu için ulaştırma işletmeleri sıklıkla kullanılmaktadır. Ulaştırma işletmeleri turizm sistemi içerisinde birçok farklı alanda yer almaktadır. Destinasyona gitmek, terminaller ile konaklama tesisi arasındaki transferler, konaklama tesisi dışına yapılan turlar gibi işlerde bu işletmeler kullanılmaktadır. Turizm içerisindeki ulaştırma işletmelerinin en önemlileri havayolu şirketleridir. Uluslararası gelen turistlerin yaklaşık %55’i bu yolla tatil yerlerine ulaşmaktadır (International inbound tourism by mode of transport in 2016, 2016). Gidilecek yerin kolay ulaşılabilir olması turizm açısından önemli bir gerekliliktir. Bu nedenle havayolu gibi hızlı ve konforlu ulaşım aracının turizm destinasyonuna seyahat imkânının olması oradaki turizmin gelişmesini oldukça etkilemektedir. Bunun bilincinde olan ülkeler uçaklar için havalimanlarına gerekli yatırımı yapmaktadırlar. Türkiye’de 51 adet havalimanı bulunmaktadır. Bunlardan sadece Atatürk Havalimanına 2018’de yaklaşık 68 milyon yolcu giriş-çıkış yapmıştır (www.dhmi.gov.tr, 2018). Atatürk Havalimanı bu haliyle Avrupa’nın en çok yolcu ağırlayan 5’nci havalimanıdır (Hudson, 2018). Bu havalimanlarından 37 tanesinde uluslararası uçuşlar gerçekleştirilmektedir (Türkiye'deki havaalanları listesi, 2018). 2018 verilerine göre Türkiye’de en çok yolcu ağırlayan 10 havalimanı **Tablo 1.6**’de verilmektedir.

**Tablo 1.6.** Türkiye’de En Çok Yolcu Ağırlayan 10 Havalimanı (2018)

<i>Havalimanı</i>	<i>Yolcu Sayıları</i>		
	<i>İç Hat</i>	<i>Dış Hat</i>	<i>Toplam</i>
İstanbul Atatürk	19.170.141	48.811.305	67.981.446
İstanbul Sabiha Gökçen	22.468.820	11.667.696	34.136.516
Antalya	7.563.107	24.003.110	31.566.217
Ankara Esenboğa	14.483.453	2.249.198	16.732.651
İzmir Adnan Menderes	10.688.828	2.736.719	13.425.547
Adana	4.927.862	699.134	5.626.996
Muğla Dalaman	1.587.234	2.972.012	4.559.246
Muğla Milas-Bodrum	2.672.257	1.513.704	4.185.961
Trabzon	3.755.574	268.766	4.024.340
Gaziantep	2.362.415	292.921	2.655.336

**\*Kaynak:** (www.dhmi.gov.tr, 2018)

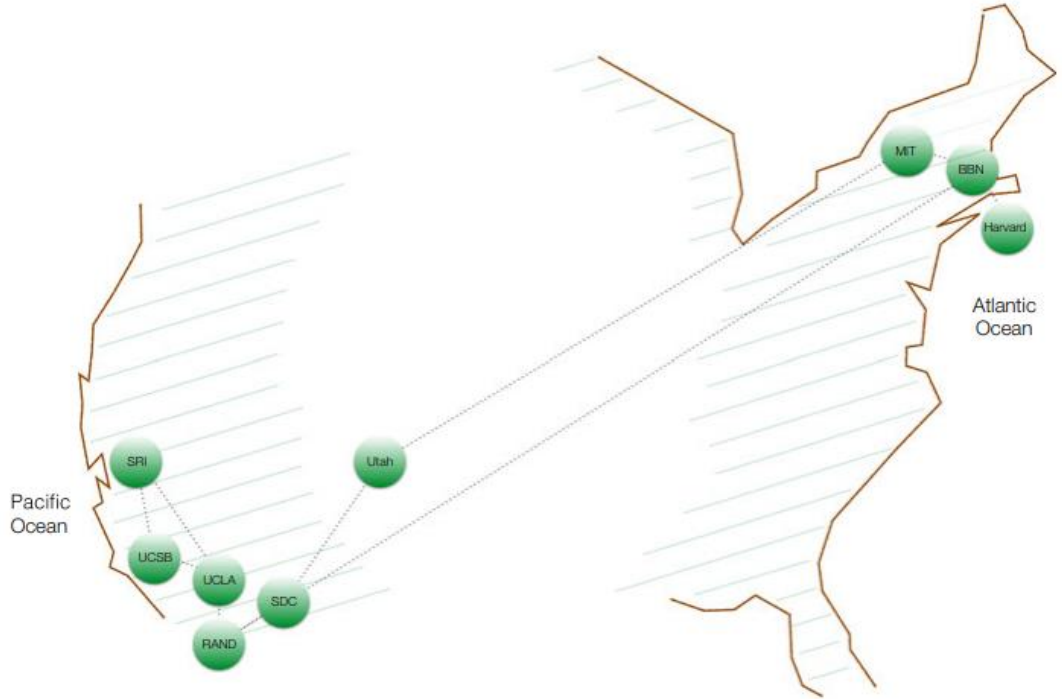
Turizm bölgelerine yakın olan Antalya, İzmir, Muğla gibi havalimanlarının ilk sıralarda yer aldığı görülmektedir.

## İKİNCİ BÖLÜM

### WEB MADENCİLİĞİ VE BİLİŞİM KAVRAMLARI

#### 2.1. İNTERNET VE İNTERNETİN GELİŞİMİ

İnternet çağımızın en önemli icatlarından biridir. İlk olarak ABD hükümetinin bir ordu projesi olarak başlayan bu haberleşme ağı daha sonra üniversitelerin dahil olduğu bir yapı haline gelmiştir. İlk dönemlerinde The ARPAnet olarak adlandırılan bu ağ olası bir nükleer saldırı sonrasında ordudaki haberleşmeyi sağlamak amacıyla geliştirilmiştir. Bu ağ mevcut telefon hatları üzerinden veri alıp göndermeyi sağlayan paket anahtarlama (paket switching) adında yeni bir teknoloji kullanmaktaydı (Skrabec, 2012: 219; Vousinas, 2012). 1979'da ARPAnet'in kurucusu olan DARPA tarafından bu ağ bazı radyo ve uydu tabanlı ağlar ile birleştirildi. Bu dönemden sonra ağ *internet* olarak adlandırılmaya başlandı (Lambert, 2005).



Şekil 2.1. 1970'lerin Başı ARPAnet ve Bağlı Üniversite ve Şirketler (Aiello, 2018)

Bir düşünce olarak internetin kökeni çok daha eskiye dayanmaktadır. 1962’de Linkliler, *Galactic Network* başlığı altında bilgisayarların birbiriyle iletişimde olduğu bir ağı tartışmıştır. Linkliler aynı zamanda DARPA’nın da ilk başkanıdır (Vousinas, 2012: 3). İnternet her ne kadar askeri amaçlı bir proje olarak başlasa da temelleri üniversitelerdeki araştırmalara dayanmaktadır. Bu çalışmalardan en önemlilerinden biri Leonard Kleinrock’un paket anahtarlamasının temeli olan doktora çalışmasıdır (Shelly ve Campbell, 2012: 12). Temeli Kleinrock tarafından atılan bu yöntem daha sonra Thomas Merrill ve Lawrence G. Roberts’ın katkılarıyla paket anahtarlama halini alacaktır. İnternetin gelişmesine katkısı olan bir diğer gelişme TCP/IP teknolojisidir. 1973’te Robert Kahn ve Vinton Cerf ağları birbirine bağlamak için bu sistemin temellerini atmışlardır (Engard, 2016).

1969’un sonuna kadar, ARPAnet’e yalnızca dört bilgisayar bağlanmıştı. Fakat 1970’lerde ağ sürekli olarak büyüdü. 1971’de Hawaii Üniversitesi’ndeki ALOHAnet ağı dâhil oldu. İki yıl sonra Londra’daki University College ve Norveç’teki Royal Radar Kuruluşu da ağı eklendiler (Whittaker, 2002: 16). Hawaii Üniversitesi’nin ALOHAnet ağı veri alışverişinde için radyo benzeri veri iletimini kullanıyordu. Bu yöntem yavaştı fakat ilerde *Ethernet* teknolojisinin geliştirilmesinde önemli katkıları olacaktı (Saunders, 2011: 46).

1980’lerde Delaware Üniversitesi’nden Dave Farber düşük maliyetli bir ağ oluşturmak için telefon hatları üzerinden çevirmeli bağlantı (dial-up) teknolojisini kullanmıştır (Abbate, 1999: 184). PhoneNet olarak adlandırılan bu ağ ARPANET ağına dâhil olmuş ilk ticari ağıdır (Engard, 2016). 1990’ların sonlarına kadar bu teknoloji internet ağına dâhil olmak için yaygın kullanılan yöntemlerden biri olmuştur.

### **2.1.1. İnternetin Gelişmesinde Etkili Olan Teknolojiler**

ARPAnet’den günümüze kadar internet birçok teknolojinin birleştirilmesi ile oluşturulmuş bir sistemdir. Pek çok teknoloji internetin gelişmesinde etkili olmuştur. Bunlardan birçoğu internetin önündeki darboğazı aşmak için geliştirilen teknolojilerdir. Bunlar hakkında bilgi vermek internetin çalışma mantığının kavranması açısından da önemli olacaktır.



### 2.1.1.1. Paket Anahtarlama

Massachusetts Teknoloji Enstitüsü'nden (MIT) Leonard Kleinrock'un *paket anahtarlama*lı ağlar (packet switched networks) üzerine yaptığı çalışma internetin temelini oluşturan sistemlerden biridir (Shelly ve Campbell, 2012). Devre anahtarlama (circuit switching) veri iletiminde iki aygıt arasında bir hat kurulur, bu hat bağlantı boyunca sürdürülür. Bağlantı sonlandırıldığında hat kesilmiş olur. Telefon hatları bu şekilde çalışmaktadır (Shinder, 2002: 170-171). Fakat paket anahtarlama da bu şekilde değildir. Paket anahtarlama da veri paketler halinde bölünür, her bir paket veri; gönderici, alıcı vb. bilgiler içerir. Bu paketler bir hat üzerinde birden çok bağlantı olduğu durumlarda sistemli bir şekilde sıraya konulur. Böylece hat tek bir bağlantıya tahsis edilmemiş olur. İnternetteki veri alışverişi bu temelde sağlanmaktadır.

### 2.1.1.2. TCP/IP

1973'te ağlar arasında ortak kullanılacak standartları belirlemek için yapılan toplantı sonrasında yapılan çalışmalar neticesinde ARPAnet'te NCP protokolünün yerine getirilmiştir. TCP/IP iki farklı protokolden oluşmaktadır (Civelek, 2009: 9-10). Bunlar TCP ve IP protokolleridir.

TCP (internet transmission control program) protokolü paket anahtarlama ağ iletişiminde kayıpsız veri gönderimi yapmak için geliştirilen bir sistemdir. Bob Kahn ve Vinton G. Cerf tarafından geliştirilen bu sistemde veri parçalara ayrılır her bir paketin bir sıra numarası vardır. Doğru iletilmeyen veri kontrol edilerek tekrar gönderilir. Paketin doğru olduğunu kontrol etmek için hata kontrol bitleri (checksum) vardır. TCP paketleri kaynak port (source port), hedef port (destination port), sıra numarası (sequence number), onay numarası (ACK number), başlık uzunluğu (header length/data offset) gibi pek çok bilgiyi barındırır (Miller, 2009: 17-19; Dostálek ve Kabelová, 2006). İnternette üzerindeki cihazlar birbirleriyle iletişim kurabilmek için bir adrese ihtiyaç duyarlar. Bu adres sisteminin adı IP'dir (internet protocol address). Bugün kullanımda olan IPv4. IP adresleri nokta ile ayrılmış 4 adet 8 bitlik sayıdan oluşmaktadır. Bu da toplam  $2^{32}$  adresleme yapılabileceği anlamına gelmektedir. İnternete bağlı aygın sayısındaki artış yaklaşık 4.3 milyar adresleme yapabilen bu sistemin bir süre sonra yetmeyeceği anlamına gelmektedir. Bu öngörü yeni bir adresleme sistemini geliştirmeyi gerekli kılmıştır. Bu darboğazı aşmak için 1999'da IPv6 geliştirilmiştir. 128 bit yapıya sahip bu sistem  $2^{128}$

adresleme yapabilmektedir. IPv4’de ondalık bir adrese izin verirken IPv6’da adresler onaltılık sayı sistemindedir (Huque, 2019; Ellis, 2019).

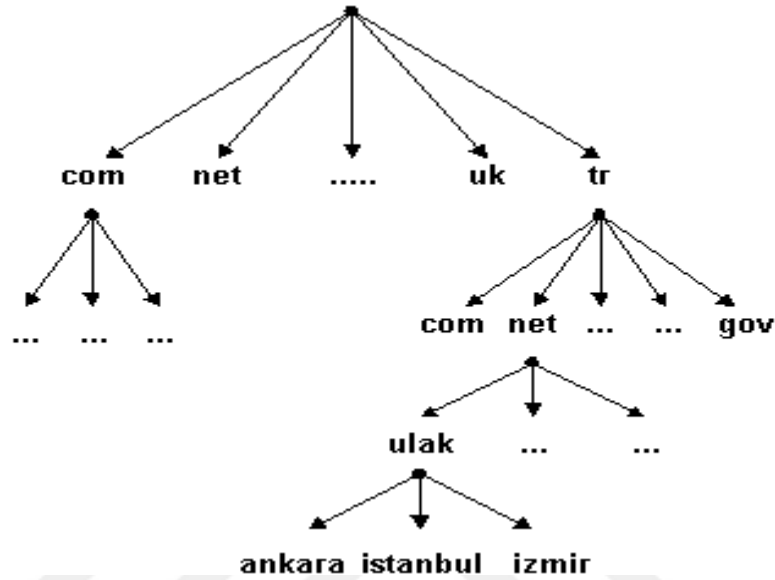
**Tablo 2.1.** IPv4 ve IPv6 Adresleme Sistemleri

	IPv4	IPv6
Yürürlük Tarihi	1981	1999
Adres Boyutu	32-bit	128-bit
Adres Formatı	Ondalık	Onaltılık
Örnek Gösterim	192.168.0.0/31	3EF3:F002:0144::/31
Adres Sayısı	$2^{32}$	$2^{128}$

### 2.1.1.3. DNS

Alan adı sistemi (domain name system) alan adlarına karşılık gelen IP adreslerini tutan veri tabanına ve sistemine verilen addır. ARPAnet’in ilk dönemlerinde ağa bağlı bilgisayar sayısı fazla olmadığından adresler SRI-NIC adlı bir bilgisayarda ve tek bir dosyada tutulmaktaydı. ARPAnet yöneticileri adreslerde bir değişiklik yapacakları zaman bunu NIC’ye (network information center) bildirmekteydiler. Ayrıca adreslerin kayıt edildiği HOSTS.TXT dosyası, istemci bilgisayara FTP aracılığıyla periyodik olarak indiriliyordu. TCP/IP geçildikten sonra bu sistem yeterli gelmemeye başladı. Bu problemin giderilmesi için Paul Mockapetris görevlendirildi. 1984’te DNS sistemi yayınlandı (Liu ve Albitz, 2003: 2-4).

DNS, alan adlarının IP adreslerinin kayıt edildiği yüzlerce sunucunun yer aldığı bir ağdan oluşmaktadır. Bağlanılmak istenen web adresi için ilk istek 13 taneden oluşan kök IP adresine yapılır. Bu adresteki DNS sunucusu adresin uzantısına göre (.com, .net, .tr, .jp) alt DNS sunucularına yönlendirme yapar. Sunucu-istemci arasındaki bu haberleşme sonucunda alan adına tanımlı sunucuya ulaşılır (www.iana.org, 2019; Bakır, 2016).



Şekil 2.2. DNS Hiyerarşik Yapısı (www.belgeler.org, 2019)

#### 2.1.1.4. Web/Www

İnternetin yaygınlaşmasına en çok neden olan teknolojilerin başında *web/www* gelmektedir. Webden önce dosya ve bilgi alışverişi için FTP, WAIS, Telnet ve Gopher gibi sistemler kullanılmaktaydı. Bu araçlar bilgiye erişmek için ya tek başına yeterli değildi ya da belirli bir uzmanlık bilgisi gerektirmekteydi. WWW’de ise bağlanılan sitede sayfa üzerinde metin ve resim verilebiliyordu daha da önemlisi sayfa üzerinde linklere tıklanarak başka sayfalara erişim sağlanıyordu (Murtagh, 1994). İlk web sitesinin yayınlandığı 1990’ dan bu yana (Press, 2015) bu sistem hızla yaygınlaşmış, aradan geçen 29 yıl içerisinde dünyadaki en önemli iletişim, eğlence ve bilgi kaynaklarından biri haline gelmiştir.

Paket Anahtarlama kavramından Ethernet teknolojisine kadar pek çok icat internetin gelişmesine katkı sağlamıştır. Burada sadece belli başlı bazı teknolojiler verilmiştir.

## 2.2. BÜYÜK VERİ

*Büyük veri* (big data) son yıllarda ortaya çıkan bilişim teknolojileri kavramlarından biridir. Gelişen teknolojinin ortaya çıkardığı bu kavram; hacim (volume), çeşitlilik (variety) ve hız (velocity) açısından büyük olan veriyi ifade etmektedir (Jain, 2017: 2). Aletler, nesnelere, doğa ve insanlar her zaman veri üretmektedir. Çoğu zaman bu veri muazzam büyüklük ve çeşitliliktedir. Son yıllara kadar, bu veriyi elde etmek, depolamak ve işlemek mümkün değildi. Gelişen teknoloji ile bu mümkün hale gelmiştir. Hükümetlerden, teknoloji şirketlerine kadar pek çok kurum ve araştırmacı bilgi çıkarımında bulunmada kullanılabilecek bu veriden faydalanmak için çalışmalar yapmaktadır. Hakkında detaylı bilgi vermeden önce büyük veri kavramının ortaya çıkmasına yol açan teknolojik gelişmeleri açıklamak faydalı olacaktır.

### 2.2.1. Depolama Teknolojisinin Gelişmesi

İlk depolama teknolojilerinden biri manyetik banttır. Alman mühendis Fritz Pfleumer tarafından patenti alınan bu icat Vlademar Poulsen'in manyetik tel icadına dayanmaktaydı. Daha sonra Pfleumer'in ilkelerini temel alan çeşitli kayıt ortamları AEG ve BASF firmaları tarafından geliştirildi (Kimizuka, 2012: 191-193). Bu teknoloji geliştirilerek uzun yıllar kullanılmıştır.

Bugünkü bilgisayarların ataları olarak nitelendirilebilecek bilgisayarlar 1950'ler ortaya çıkmaya başlamıştır. Manyetik bantlar yedekleme yapmak için kullanışlı olsa da, sürekli veri alışverişi yapacak bir iş için verimli değildi. IBM şirketi bu problemi çözmek için günümüzdeki sabit diskin atası olan RAMAC 305'i geliştirdi (Gildred, 2018). Elli adet manyetik disklerden oluşan bu aygıt 5 milyon 6 bit karakter yani yaklaşık 3,7 MB veri saklayabiliyordu. Asıl önemli olan şeyse veriye rastgele erişim (random access method) sağlayabiliyordu. Bu disk günümüzdeki disklere benzer bir yapıda olmasına rağmen aralarında büyük bir fark vardır. RAMAC 305 yaklaşık 1 ton ağırlığındaydı. Maliyet olarak karşılaştırıldığında ise MB başına yaklaşık \$10.000'lık bir maliyete sahipti (Clements, 2014: 27). Aşağıda yaklaşık 10'ar yıl arayla sabit diskler ve bu disklere ait bazı bilgiler verilmektedir.

**Tablo 2.2.** Sabit Diskler ve Zaman İçerisindeki Gelişimleri

<i>Yıl</i>	<i>\$/MB</i>	<i>Boyut</i>	<i>Tip</i>	<i>Kapasite MB</i>	<i>Fiyat \$</i>
1956	9.200	24"	Single Head	4	34.500
1966	1.047	14"	Pack 20 Surface	29	30.555
1982	289,21		HD	19	5.495
1986	24,45		HD	20	489
1996	0,204	3,5"	IDE	1.270	259
2006	0,000324	3,5"	SATA-150	250.000	81
2016	0,0000312	3,5"	SATA-3	4.000.000	125
2019	0,0000187	3,5"	SATA-3	8.000.000	150

**\*Kaynak:** (McCallum, Disk Drive Prices (1955-2019), 2019)

**Tablo 2.2** incelendiğinde sabit disklerin zaman içerisinde nasıl bir gelişme gösterdikleri görülebilmektedir. Tabloda verilmeyen SSD, RAID, Cloud gibi teknolojiler de depolama konusunda önemli katkılar sağlamıştır.

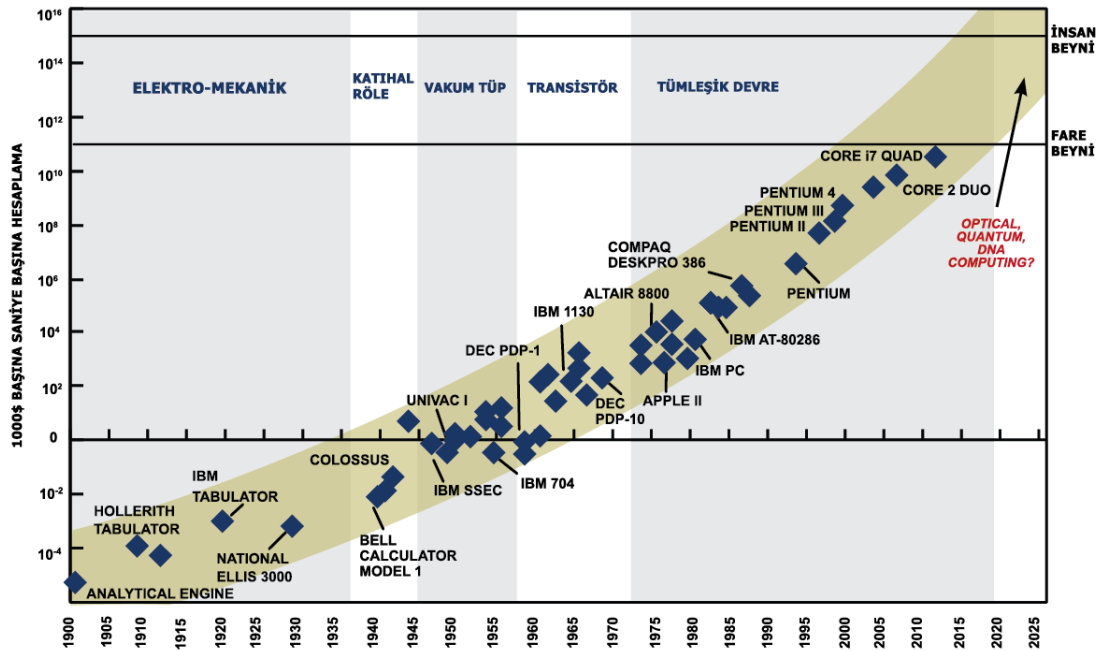
Büyük veri kavramının oluşması için en önemli faktörlerden biri üretilen verinin saklanabiliyor olmasıdır. Belirtildiği üzere günümüzde veri başına depolama maliyeti oldukça düşmüştür. Sadece depolama maliyetinin düşüşü değil, depolama aygıtlarının daha az enerji harcaması, birden çok aygıtın birleştirilerek tek bir aygıt gibi kullanılması, dünyanın herhangi bir yerinden kiralanabilen bulut (cloud) servisler, büyük veri kavramının gelişmesine önemli katkılar sağlamıştır. Sadece Youtube'a dakikada yaklaşık 300 saat video yüklenmektedir (Aslam, 2018; Chi, 2019). Bu miktarda büyük bir veriyi kayıt edecek imkân gelişmiş bir teknoloji ile mümkündür.

### 2.2.2. Bilgisayarların İşlem Gücünün Artması

Sadece veriyi depolayabilmek o verinin kullanılabilmesi için yeterli değildir. Veri miktarı arttıkça o veriyi işlemek ve bilgi çıkarımında bulunmak daha da güçleşmektedir. İlk bilgisayarlar olarak nitelendirilen cihazlar vakum tüpleri ile çalışan çok büyük aygıtlardı. Örneğin 1944'de icat edilen ENIAC 18000 vakum tüpüne sahipti yaklaşık 167 m<sup>2</sup> yer kaplıyordu (Zimmermann, 2017; Soriano, 2012: 22). Bu cihaz delikli kartlar (punch card) ile programlanıyordu ve bugünkü bilgisayarlar ile kıyaslanınca hem

kullanması çok zordu hem de çok yüksek bir maliyete sahipti. ENIAC saniyede 0.05 MIPS (million instructions per second) komut çalıştırıyordu ve bunu yapmak için ortalama 150 kilovat elektrik tüketmesi gerekiyordu (Haigh, Priestley, Priestley ve Rope, 2016: 69). Bilgisayar dünyasının asıl devrim transistörün icadından sonra gerçekleşmiştir. Vakum tüplerinin aksine transistörler hem daha küçük hem de daha az enerji harcamaktaydı (DuCastel ve Jurgensen, 2008: 64). Meydana gelen gelişme açısından değerlendirildiğinde 1946'da bir süper bilgisayar olan ENIAC'ın işlem gücü yaklaşık 500 FLOPS iken, günümüzün süper bilgisayarları yaklaşık 93 petaFLOPS hızında işlem yapabilmektedir. Kıyaslanmanın daha iyi anlaşılması için şöyle bir örnek verilebilir. Günümüzde birkaç yüz gram ağırlığındaki standart bir cep telefonu teorik olarak yaklaşık 200 gigaFLOP işlem yapabilmektedir (Frumusanu, 2018) ve bunu yapmak için gerekli olan enerjiyi ise üzerinde taşıdığı 3000mAh kapasiteli bir pilden sağlamaktadır.

Şekil 2.3'de bilgisayar teknolojisinin 1900'lerden bu yana evrimi ve bir öngörü de olsa geleceği hakkında bazı bilgiler verilmektedir.



Şekil 2.3. Bilgisayar İşlem Gücündeki Artış (Templeton, 2015; Kurzweil, 2008, s. 68)

Gelecekte kuantum bilgisayarlar gibi teknolojiler ile bilgisayarların sahip olduğu işlem gücünün, daha da artacağı tahmin edilmektedir. Hızlı işlem yapan bilgisayarlar büyük verinin işlenmesi için elzem olan teknolojilerden biridir.

### 2.2.3. Mobil Teknolojilerin Gelişmesi

Özellikle son on yılda mobil teknoloji çok önemli gelişmeler göstermiştir. Günümüzde mobil aygıt denildiğinde cep telefonu, taşınabilir bilgisayarlar, tablet ve giyilebilir teknolojiler gibi bir amaç için üretilmiş ve belirli bir hesaplama gücüne sahip aygıtlar anlaşılmaktadır. Bu aygıtların ortak özellikleri, taşınabiliyor oluşları ve çoğu zaman internet veya hücresel bir ağın parçası olmalarıdır.

Hücresel bağlantı kullanan cihazların yaygınlaşması mobil aygıtlar çağının başlangıcı olmuştur. Her ne kadar geçmişi 20.yy başlarına kadar dayansa da bu teknolojinin yaygınlaşması hücresel bağlantı üzerinden çalışan mobil telefonlar ile gerçekleşmiştir. 1990'larda Avrupa'da GSM şebekeleri ABD'de CDMA şebekeleri hızla yaygınlaşmıştır (Andrew, 2018; Dudley, 2018). Teknoloji geliştikçe hem bu aygıtların boyutları küçülmüş hem de telefon özelliklerinin yanına yeni işlevler eklenmiştir. Günümüzde mobil telefonların vazgeçilmez donanımlarından olan kamerayı Japon Sharp şirketi 2000'lerin başında ilk kez bir modelinde piyasaya sürmüştür (Ha, 2010). Mobil telefonların sahip olduğu donanımlar ve işlevler çoğalınca, bu cihazlar artık akıllı telefon (smart phone) olarak adlandırılmaya başlanmıştır.

Günümüzde akıllı telefon olarak adlandırılan cihazlar telefon özelliklerinin yanında onlarca başka işlevi de kullanıcılarına sunmaktadır. Bu cihazlar kablosuz internet, mobil internet, bluetooth gibi kablosuz bağlantı özelliklerine sahip, müzik dinlenebilen, internette gezinebilen, fotoğraf ve video çekebilen mini bilgisayarlar haline gelmişlerdir. Üzerlerinde taşıdıkları onlarca sensör ile coğrafi konum, yükseklik, hava sıcaklığı, ses ve ışık şiddeti gibi değişkenlere ait verileri toplayabilmektedir. Ayrıca kullanıcı bir akıllı telefon ile nabzını ve adımlarını ölçebilmekte, üzerinde taşırken yaptığı işleri öğretebilmekte, NFC gibi teknolojiler ile ödemelerini yapabilmektedir. Ünelere göre geçiş tarihleri farklılık gösterse de mobil şebekelerin gelişim seyri ve özellikleri **Tablo 2.3'**de verilmektedir.

**Tablo 2.3.** Mobil İletişimin Gelişimi ve Karşılaştırılması

<i>Teknoloji</i>	<i>1G</i>	<i>2G/2.5G</i>	<i>3G</i>	<i>4G</i>	<i>5G</i>
Tarih	1970-1980s	1990-2004	2004-2005	2010-	2019 Sonu
Bant genişliği	2 Kbps	14-64 Kbps	2 Mbps	200 Mbps	1 Gbps <
Servis	Mobil Telefon	Dijital Ses, SMS	Entegre yüksek kaliteli ses, video ve veri	Dinamik bilgi erişimi, değişken cihazlar	Dinamik bilgi erişimi, yapay zekâ yetenekleri ile değişken cihazlar
Çoğaltma	FDMA	TDMA/CDMA	CDMA	CDMA	CDMA
Anahtarlama	Devre	Devre/Erişim ağı ve hava arayüzü için devre	Hava arayüzü hariç paket	Tüm paket	Tüm paket
Çekirdek Ağ	PSTN	PSTN	Paket ağ	Internet	Internet
Transfer	Yatay	Yatay	Yatay	Yatay ve Dikey	Yatay ve Dikey

**\*Kaynak:** (Mukhopadhyay, Agarwal, Sharma ve Gupta, 2016; Rajiv, 2018)

Mobil aygıt denildiğinde sadece telefonlar akla gelmemelidir. Günümüzde bazı mobil aygıtlar yukarıda bahsedilen işlerin birçoğunu veya bazısını yerine getirebilmektedir. Giyilebilir teknoloji (wearable technology) ile saat, bileklik, ayakkabı gibi pek çok ürünün içerisinde elektronik aygıt kullanımını görmek mümkündür.

Bütün bu gelişmeler ile insanlar üzerinde taşıdıkları aygıtlar ile neredeyse her yerden internete bağlanabiliyor, içerik veya veri üretebiliyor hale gelmişlerdir. Mobil cihazlar vasıtasıyla üretilen veri, büyük veri kavramının önemli kaynaklarından biridir. Bu veri kullanılarak üretilebilecek bilgi geleceği şekillendirecek niteliktedir.

Burada büyük veri kavramının ortaya çıkmasına zemin hazırlayan bazı teknolojik gelişmeler verilmiştir.

Büyük verinin insanların ve kurumların bu denli ilgisini çekmesinin birçok nedeni vardır. Bazı değişkenler arasında ilişkinin yönünü, büyüklüğü veren veya bazı değişkenlerin diğer değişkenler üzerindeki etkisini gösteren hali hazırda birçok istatistiki yöntem olmasına rağmen, büyük veri sadece büyük miktarda veri kullanıldığında anlamlı olan yöntemlerin kullanılmasına yol açmıştır.

Büyük veri geleneksel yazılım araçları ile işlenmesi mümkün olmayan büyüklükte veriden oluşmaktadır. Burada sadece hacim olarak çok büyük veri düşünülmemelidir. Bazen 50 GB'lık bir veri seti bile büyük veri olarak kabul edilebilmektedir. Önemli olan



verinin ilişkisel bir veri tabanı için karmaşık bir yapıya sahip olmasıdır (Jain, 2017: 2). Büyük verinin tanımlanabilmesi için yazında 3v, 5v gibi kıstaslar belirlenmiştir (Aggarwal A. , 2016). Bunlardan 3v modeline göre;

*Hacim (volume)*: Veri hacmi büyük olmalıdır. Bu konuda farklı görüşler ve tartışmalar mevcuttur (Jain, 2017: 2-3). İlişkisel veri tabanlarının izin verdiği ötesinde, yapısal veya yapısal olmayan karmaşık veri olmalıdır.

*Çeşitlilik (variety)*: Geleneksel veri tabanlarında veri tipleri belirli bir yapıdadır. Fakat büyük veride, metin, ses, video gibi değişkenler de yer alır. Veri yapısal olmayan veya yarı yapısal olabilir (Rouse, 2018; www.oracle.com, 2019; Zyl, 2014).

*Hız (velocity)*: Veri sürekli değişmelidir. Örneğin bir e-ticaret sitesi düşünüldüğünde veri sürekli değişmektedir (Kale, 2017). Fiyatlar, ürünlere yapılan yorumlar, ziyaretçilerin ürün arama sıralaması gibi değişkenler sürekli değişmektedir. Twitter, Facebook gibi sosyal ağlarda bu hız çok daha yüksektir.

Kurumlar sahip oldukları veri tabanlarını veya elde ettikleri verileri çoğu zaman kurum dışında paylaşmamaktadırlar. Çeşitli araçlar yardımı ile bu veriler işlenmekte elde edilen bilgiler ise karar verme, pazarlama, ürün geliştirme, istihbarat gibi birçok alanda kullanılmaktadır. Büyük veri ilk olarak DNA gibi biyolojik araştırmalarda ve iklimsel olayların araştırmalarında kullanılmıştır. Fakat günümüzde daha önce bahsedilen teknolojik gelişmelerle birlikte birçok alanda kullanılmaktadır. Büyük verinin kullanım alanları şu şekildedir (Buluswar, 2016; Kleinberg, 2013; Fung, 2013; Stephenson, 2018).

## **2.2.4. Büyük Verinin Kullanım Alanları**

### **2.2.4.1. Müşteri Davranışlarının Araştırılmasında**

Müşteri bilgilerinin veri tabanlarında tutulması yeni bir kavram değildir. Fakat gelişen teknoloji çok daha büyük veriyi kayıt etmeyi olanaklı hale getirmiştir. İşletmeler ister kendi kaynaklarında olsun ister sosyal medya üzerinden olsun, potansiyel müşterilerinin davranışlarını büyük veri yoluyla analiz edebilmektedir. Bu davranışların analizi, ürün geliştirme, pazarlama ve reklam, eğilimler gibi birçok alanda önemli bilgiler sağlamaktadır.

Benzer şekilde seçimlerden önce seçmen davranışları da bu yolla analiz edilebilmektedir. Bu analizler politikacılara önemli avantajlar sağlamaktadır.

#### **2.2.4.2. Dolandırıcılık ve Güvenlik Açıklarının Tespitinde**

Özellikle finans şirketleri için büyük veri dolandırıcılığın önlenmesi ve güvenlik açıklarının tespitinde önemli katkılar sağlamaktadır. Kredi başvurusu yapan bir müşteri için gerçek zamanlı risk oranı büyük veri ile tespit edilebilmektedir. Benzer şekilde kredi kartı kullanım davranışından o alışverişin dolandırıcılık olup olmadığı önceden tespit edilebilmektedir. Bunları gerçekleştirmek için daha önce örnek vakalardan toplanan verilerden oluşturulan veri ambarı kullanılabilir.

#### **2.2.4.3. Bilimsel Araştırmalarda**

Pek çok bilimsel araştırmada büyük veri uygulamaları kullanılabilir. Örneğin CERN laboratuvarlarında büyük hadron çarpıştırıcısından (LHC) elde edilen veriler bilimsel bir problemin çözümü için kullanılmaktadır. CERN’de her saniyede LHC’den 300 GB veri elde edilmektedir. Bunun sadece 300 MB kullanışlı veridir (Marr, 2016: 14). Örneğin dünyanın önde gelen teknoloji şirketlerinden IBM, Weather Company’i satın almıştır. IBM dünyanın çeşitli yerlerinde bulunan şirketin sahip olduğu binlerce sensörden gelen verileri işleyerek daha doğru hava tahminleri yapmayı amaçlamaktadır (Bridgwater, 2015; Boulton, 2015). IBM’in bu işbirliği ile büyük veri alanında tecrübesini artıracığı da aşikârdır.

#### **2.2.4.4. Sağlık ve Halk Sağlığının İyileştirilmesinde**

Büyük verinin ilk kullanıldığı alanlardan biri sağlık alanıdır. 1990’larda DNA’ların dizilişini çözümlmek için paralel hesaplardan faydalanılmıştır (Patnala ve Kumar, 2018). Günümüzde birçok hastalığın kalıtsal olduğu yapılan bu tarz çalışmalar ile belirlenebilmektedir (Hoffman, 2016). Büyük verinin kullanımı sadece DNA çözümlmesine dayanan çalışmalarla sınırlı değildir. En çok veri toplanan alanlardan biri de sağlık kurumları ve hastanelerdir. Bu kurumlara başvuran hastalardan insan sağlığı ile ilgili pek çok değişken içeren veri toplanmaktadır. Bu veri insan yaşamını iyileştirmeye veya çeşitli sağlık sorunlarını çözmeye yönelik kullanılabilir niteliktedir.

Burada bahsedilmeyen daha pek çok alanda büyük veriden faydalanılmaktadır. Büyük veri sadece kurumların veri tabanlarında yer alan bir kaynak değildir. Günümüzde büyük verinin en büyük kaynaklarından birini, belki de en büyüğünü web oluşturmaktadır. Sosyal ağlar, bloglar, forumlar, çevrimiçi veri tabanları, haber siteleri, e-ticaret siteleri üzerinde herkesin erişebildiği çok kıymetli veriler barındıran

sistemlerdir. Twitter veya Facebook gibi bir sosyal ağ sitesinde internet kullanıcılarının oluşturduğu metin, fotoğraf ve videolardan oluşan terabaytlarca veri, insanların davranışları, eğilimleri, zevkleri ve gündem gibi daha birçok konuda bilgi barındırmaktadır. Benzer şekilde bir e-ticaret sitesinde ürünler hakkında detaylı bilgi bulunmakta, neredeyse her e-ticaret sitesinde o ürünü satın almış kişilerin ürün hakkındaki yorumlarına yer verilmektedir.

Web madenciliği veya web kazıma web üzerinde yer alan büyük ve değerli veriyi elde etme ve düzenleme yöntemlerinden oluşmaktadır. Hali hazırda bu bilgilere veri tabanlarında sahip olan şirketlerin bu bilgileri kullandıkları bilinmektedir. Web madenciliği sadece şirketlere değil, o veriye dışarıdan ulaşabilen herkese bu imkânı tanımaktadır. Bugün çoğunlukla kurumsal anlamda kullanılan bu veriler, yakın zamanda web madenciliğinin yaygınlaşması ile kişiler ve araştırmacılar için de öncelikli kaynaklardan olacak niteliktedir.

### **2.3. VERİ BİLİMİ VE VERİ MADENCİLİĞİ**

Dünya büyük veri çağına girdiğinde, verinin depolanması için gereken teknoloji işletmeler için önemli endişe kaynaklarından biriydi. 2010'dan sonra şirketler veri depolanmasındaki bu soruna odaklandılar. Hadoop ve benzeri yapılar bu probleme karşı geliştirilen teknolojiler oldu. Daha sonraki odak noktası bu verinin işlenmesi olmuştur. Böylelikle veri bilimi yapay zekânın geleceği haline gelmiştir (Sharma H. , 2019). *Veri bilimi*, bilgi edinmek, tahminde bulunma gibi bir iş için bir kaynaktan veriyi elde etme, düzenleme, analiz etme süreç, metot, algoritma ve yöntemler bütünüdür (Davenport ve Patil, 2012; Johansson, 2019; Corea, 2019: 1; Grus, 2019). Veri bilimi, büyük veri, web madenciliği, veri madenciliği, gibi kavramları içine alan, matematik, istatistik, bilgisayar bilimi ve bilişim biliminden faydalanan son yıllarda ortaya çıkan bir kavramdır. Veri bilimi kavramı ilk kez 1974'te Peter Naur tarafından kullanılmıştır (Voulgaris, 2014: 19; Ratner, 2017: 15-16). Fakat bu kavramın yaygınlaşması için bilgisayar ve bağlantılı teknolojilerin gelişmesi gerekecekti.

Veri madenciliği, veri bilimi olarak nitelendirilen yöntemlerin bir kısmını oluşturmaktadır. Veri madenciliği, büyük veri setleri içerisinde anormallikler, örüntü (pattern) ve ilişkiler bulmaya yönelik yöntemlerden oluşmaktadır (www.sas.com, 2019).

Dünya çapındaki işletmeler, satış işlemleri, hisse senetleri, ürün bilgileri, promosyonlar, şirket ve müşteri profilleri gibi devasa büyüklükte veri tabanlarına sahiptir (Han, Kamber ve Pei, 2012). Daha önce de açıklandığı üzere, web muazzam büyüklükte bir veri barındırmaktadır. Bütün bu gelişmelerin sonucunda bu veriyi kullanma ve bu veriden önemli, işe yarar bilgi çıkarma yöntemi olan *veri madenciliği* ortaya çıkmıştır.

Veri madenciliği, istatistik bilimini kullanıyor olsa da sadece bununla sınırlı değildir. Veri madenciliğinde istatistik dışında çeşitli matematiksel yöntemler ve algoritmalar da kullanılmaktadır (www.oracle.com, 2019; Galetto, 2015; Koshy, www.promptcloud.com, 2017). Ayrıca verinin analiz kısmı, veri madenciliğinin sadece bir ayağını oluşturmaktadır. Veri madenciliği, verinin elde edilmesi, temizlenip düzenlenmesi, görselleştirilmesi gibi basamaklara sahiptir.

### **2.3.1. Veri Madenciliği Aşamaları**

Veri madenciliği tek bir işlem değildir. Birçok süreçten oluşan bazen farklılaşan işlemlerden oluşan bir yöntemler bütünüdür. Veri madenciliğinde hedef karar vermede ve seçim yapmada kullanılabilecek değerli veriyi ortaya çıkarmaktadır. Bu nedenle bu hedefe ulaşmak çeşitli basamaklardan oluşmaktadır. Yazında farklı yaklaşımlar olsa da bu aşamalar temelde birbirlerine benzemektedir (Dontha, 2018; Rajput, 2019; García, Luengo ve Herrera, 2015: 1-2; Han, Kamber ve Pei, 2012: 5-7). Aşağıda genel olarak veri madenciliği aşamaları verilmektedir.

#### **2.3.1.1. Problemin Tanımlanması**

Kurum veya araştırmacı için amacın ne olduğunun belirlenmesi aşamasıdır. Teknik bir aşama olmasa da üzerinde durulması gereken bir basamaktır. Amaç bazen bir problemin çözülmesi, müşteri sayısını arttırma, maliyetleri düşürme, satışları artırma gibi belirgin bir hedef olurken, bazen de belirli bir veri içerisindeki gizli değişkenleri, değişkenler arasındaki ilişkiyi bulmaya yönelik te olabilmektedir.

#### **2.3.1.2. Veriyi Çözümleme ve Temizleme**

Bu aşamada elde edilen veri çözümlenmeli ve sonrasında gereksiz olan kısımlar temizlenmelidir. Veri içerisinde amaç için kullanılacak değişkenler hangileri olmalıdır, hangi veriler birleştirilmeli, gibi sorular bu basamakta cevaplandırılmalıdır. Bazen tek bir veri kaynağından faydalanılırken, bazen de birden çok veri kaynağı tek bir veri tabanı

veya ambarı haline getirilebilir. Sonrasında gereksiz olan, tekrarlanan, hatalı ve eksik veriler temizlenmelidir. Bu aşamada araştırmanın sonucu üzerinde yıkıcı etkisi olabilecek düzenlemeler kaçınılmalıdır. Veri temizleme aşaması sonraki aşamaların doğru bir şekilde yürütülmesi için üzerinde önemle durulması gereken bir aşamadır.

### 2.3.1.3. Veri Dönüşümü (Data Transformation)

Bu aşama, veri temizlendikten sonra analizlere uygun hale getirilmesi için gereken işlemleri içermektedir. Verinin normalizasyonu (istatistiksel), değişken türlerinin belirlenmesi veya değiştirilmesi, gruplandırma (binning) gibi işlemler veri dönüşümü işlemlerinden bazılarıdır. Örneğin yapay sinir ağları yöntemi verinin normalizasyonunu gerektirirken, lojistik regresyon modellerinde bağımlı değişken kategorik (binned, nominal) olmalıdır.

### 2.3.1.4. Veri Madenciliği

Bu kısımda veriden bilgi çıkarımda bulunmayı sağlayacak analizler yapılmaktadır. Burada amaca ve verinin yapısına göre yöntem veya yöntemler seçilmelidir. Bu yöntemler bazen korelasyon gibi değişkenler arasındaki ilişkiyi belirlemeye yönelik oluyorken, bazen de yapay sinir ağları, regresyon gibi tahmine yönelik de olabilmektedir. Başka bir tabirle bu yöntemler bazen sınıflandırma, zaman serisi analizleri gibi tahmin amaçlı olabiliyorken, bazen sınıflandırma, kümeleme gibi eldeki veriyi tanımlamaya yönelik de olabilmektedir (Shylaja ve Muralidharan, 2017).

Veri madenciliği kısmında önemle üzerinde durulması gereken başka bir konu da her veri tipine her analiz metodunun uygun olmamasıdır. Örneğin *ki-kare*, *spearman* gibi yöntemler kategorik değişkenler kullanıldığında anlamlı oluyorken, *yapay sinir ağları* daha geniş bir değişken türüne hitap etmektedir. Bu nedenlerle, bu aşama veri madenciliği yönteminin çekirdeğini oluşturmaktadır.

### 2.3.1.5. Örüntü Değerlendirme (Pattern Evaluation)

Bu kısımda analiz sonrasında ortaya çıkan sonuçlar değerlendirilir ve önemli varsayılanlardan bir özet çıkarılır. Bu basamak, ilk basamağa benzer şekilde araştırılan konuda uzmanlar tarafından değerlendirilmelidir.

### 2.3.1.6. Bilginin Raporlanması

Bütün bu aşamalardan sonra elde edilen bilgi, ilgili olan kişilerin kolaylıkla anlayabileceği şekilde raporlanmalıdır. Bu aşama bu teknikleri içermektedir. Her zaman çıktığı konunun ilgilileri için anlamlı olmamaktadır. Bu nedenle bu aşamada bu durum göz önünde bulundurularak analiz sonucunu doğru şekilde içeren, gerekirse tablo, grafik gibi öğelerle desteklenen bir rapor hazırlanmalıdır.

Veri madenciliği sonucunda çeşitli amaçlarla kullanılacak değerli bilgi ortaya çıkmaktadır. Elde edilen bilgi, bazen değişkenlerin aldıkları değerlere göre gruplandırılmalarına olanak tanırken bazen de hangi değişkenlerin birbirleriyle birlikte görüldüğünü ortaya koymaktadır.

Veriden örüntü çıkarmak için pek çok yöntem vardır. Bilimin elverdiği ölçüde bu yöntem ve algoritmalar gün geçtikçe daha da geliştirilmektedir. Aşağıda belli başlı bazı veri madenciliği yöntemleri verilmektedir.

### 2.3.2. Veri Madenciliği Yöntemleri

Veri madenciliği regresyon, korelasyon gibi uzun yıllardır sürülegelen klasik istatistiki yöntemlerle birlikte derin öğrenme, genetik algoritmalar gibi kısmen daha yeni ve sürekli gelişen yöntemleri de içermektedir. Veri madenciliği yöntemlerinin sınıflandırılması birçok farklı şekilde ele alınmaktadır (Roiger, 2017; Han, Kamber ve Pei, 2012). Uygulamada veri madenciliği yöntemleri iki temel amaçla kullanılmaktadır. Bunlar *tanımlama* (description) ve *tahminde bulunmadır* (prediction). Veri madenciliği metotları bu iki amacı yerine getirmek için kullanılmaktadır (Fayyad, Piatetsky-Shapiro ve Smyth, 1996: 44). Örneğin regresyon bir tahmin yöntemi iken, kümeleme bir tanımlama yöntemidir (Dunham, 2003).

Aşağıda çeşitli veri madenciliği yöntemleri çeşitli değerlendirmelerden, işlevlerine göre derlenerek verilmektedir (Roiger, 2017; Fayyad, Piatetsky-Shapiro ve Smyth, 1996; Dunham, 2003; Han, Kamber ve Pei, 2012).

#### 2.3.2.1. İlişki Analizi ve Birliktelik Kuralları

İstatistikte iki değişken arasındaki doğrusal ilişkinin varlığı korelasyon (correlation) terimi ile ifade edilmektedir. Korelasyon katsayısı olarak hesaplanan değer, -1 ile +1 arasında bir değer almaktadır. Bu değer 0'a yaklaştıkça ilişkinin gücü

azalmaktadır. Birliktelik (association) analizi kavramı da buna benzemektedir. Birliktelik kuralı, bir X değişkenin meydana gelmesi karşısında bir Y değişkenin de meydana gelme olasılığını ifade etmektedir.

Veri madenciliğinde birliktelik analizi değişkenler ve bir veri setindeki özelliklerin birlikte meydana gelme olasılığını hesaplamak için kullanılmaktadır (Roiger, 2017: 47). Birliktelik kuralları korelasyondan farklı olarak daha çok değişkene sahip birliktelikler için de kullanılmaktadır. İki'den çok değişkenin birlikte meydana gelme olasılığı birliktelik kuralları ile belirlenebilmektedir (Goh ve Ang, 2007). Birliktelik kuralları büyük veri setlerinde ve geniş değişken aralığında olumlu sonuçlar vermektedir (Agrawal, Imielinski ve Swami, 1993). Bu çalışmaların öncüllerinden, Agrawal, Imielinski ve Swami, (1993) *Apriori* olarak adlandırdıkları algoritmaları ile sepet analizi yapmış, hangi ürünlerin birlikte satıldığını araştırmışlardır.

En çok bilinen birliktelik algoritmalarına Apriori, AIS, SETM örnek verilebilir. İkili birlikteliklerde ise Pearson, Spearman's rho, Kendall's tau korelasyon yöntemleri yaygın kullanılmaktadır (Berrar, Dubitzky ve Granzow, 2003: 291; Dunham, 2003: 191).

### 2.3.2.2. Kümeleme Analizi

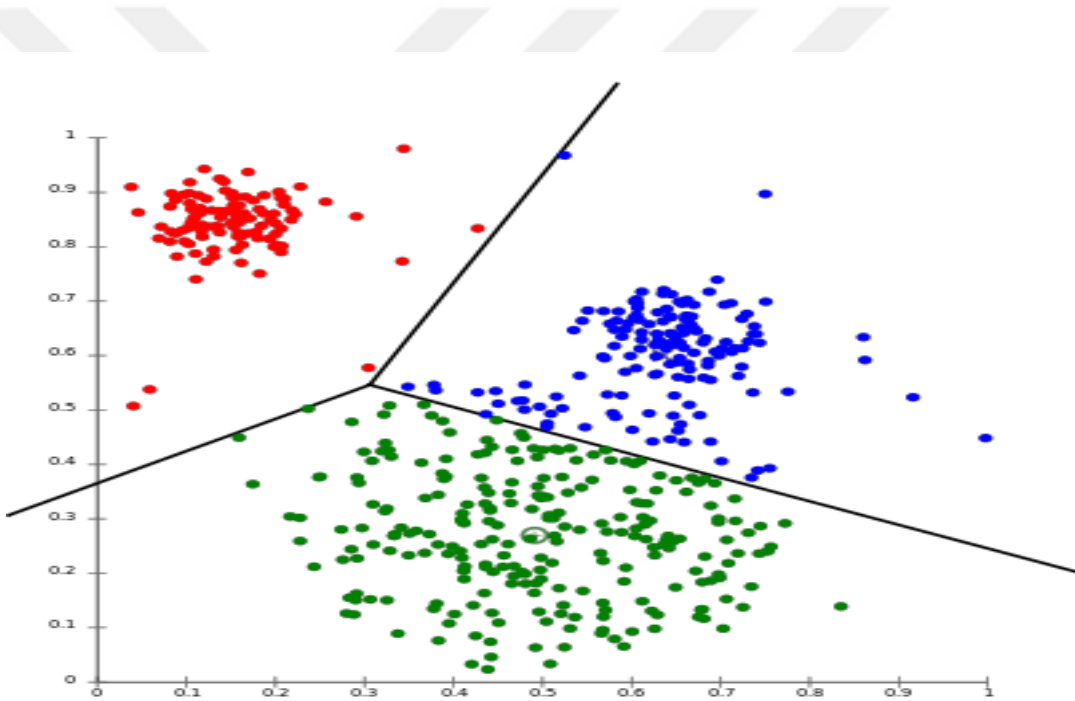
Kümeleme (clustering) analizi veya kümeleme bir veri setindeki birimlerin birbirlerine benzemelerine göre gruplandırılmaları anlamına gelmektedir. Burada amaçlanan aynı küme içerisinde yer alan birimlerin olabildiğince benzer olmaları, farklı kümelerde yer alan birimlerin de olabildiğince benzemez olmalarıdır (Akpınar, 2014: 284-285). Örnek vermek gerekirse, elmalarla dolu bir sepette elmaların renklerine, boyutlarına, tatlarına göre sınıflandırılmaları fakat hangi elmanın hangi sınıftan olduğunun önceden bildirilmemesidir. Kümeleme analizinin sınıflandırma analizlerinden önemli farkı budur. Sınıflandırma analizlerinde çeşitli özelliklere sahip birimlerin hangi sınıfın elemanı olduğu önceden bilinmektedir. Fakat kümeleme analizinde bunu keşfetmek, daha doğrusu yeni bir sınıf tanımlamak algoritmanın işidir.

Kümeleme algoritmalarında birimlerin benzerlikleri belirlenirken çoğu zaman mesafe ölçülerinden faydalanılmaktadır. Kümeleme ayrıca aykırı değerlerin (outlier) tespitinde de kullanılmaktadır (Han, Kamber ve Pei, 2012: 383).

Kümeleme algoritmalarından en bilinenlerden biri k-ortalama (k-mean) algoritmasıdır. İlk kez 1970'lerin ortalarında sunulan bu yöntem kümeleme işlemini

yaparken genellikle Euclid, Kareli Euclid, Chebyshev gibi uzaklık ölçülerini kullanmaktadır (Nisbet, Miner ve Yale, 2018: 145). Buradaki  $k$  verinin ayrılacağı küme sayısını belirtmektedir. Bu yöntemde önceden küme sayısını belirtmek gerekmektedir. Algoritma bu sayıyı temel alarak birimleri benzerliklerine göre gruplandırmaktadır.

Kümeleme algoritmaları hiyerarşik ve bölümlenmeli olarak iki grupta toplanmaktadır. Hiyerarşik yöntemlerde kümeler ağaç şeklinde, iç içe kümeler şeklinde üretilir verinin kendisi dışında bir ayrıca bir parametreye ihtiyaç duyulmamaktadır. Bölümlenmeli yöntemler ise küme sayısı gibi kümelere ait bazı özelliklerin önceden verildiği yöntemlerdir (Sammur ve Webb, 2017: 190-191; Han, Kamber ve Pei, 2012: 408).



Şekil 2.4. K-ortalama Algoritması İle Kümeleme (Chadha, Das, & Karnin, 2018)

### 2.3.2.3. Sınıflandırma Analizi

Yaygın kullanılan bir başka veri madenciliği yöntemi de sınıflandırma (classification) analizidir. Sınıflandırma bir nesnenin sahip olduğu özelliklere göre daha önce belirlenen sınıflardan birine atanması olarak düşünülebilir. Sınıflandırma analizinde kümeleme analizinden farklı olarak verinin (eğitim) ait olduğu bir sınıf önceden bilinmektedir. Bilinen bu sınıflara yeni veri sahip olduğu özelliklere göre atanmaktadır.

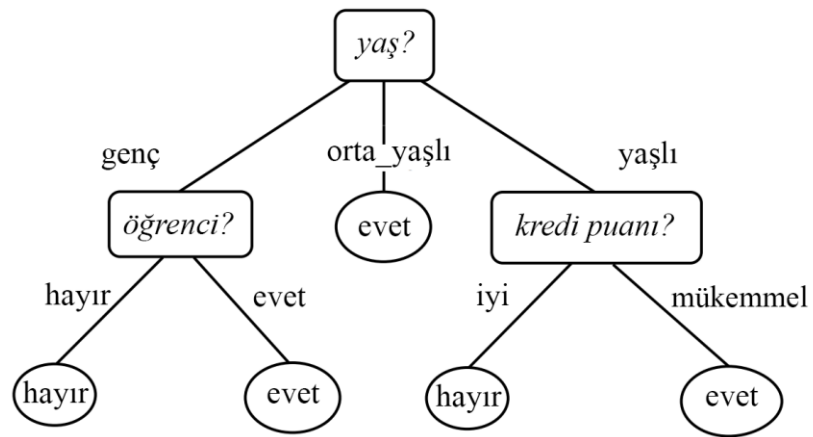


Sınıflandırma analizlerinin kullanım alanı oldukça geniştir. Finansal piyasa eğilimleri, endüstri uygulamalarındaki hataları tespit etme, tıbbi teşhis, imaj ve örüntü tanıma gibi pek çok alanda kullanılmaktadır (Dunham, 2003: 75). Sınıflandırma analizi üç özelliğe sahiptir. Bunlar, öğrenme denetimlidir, bağımlı değişken kategoriktir ve amaç var olan sınıflara yeni örneği en doğru şekilde atamaktır (Roiger, 2017: 34).

Lojistik regresyon, yapay sinir ağları, karar ağaçları ve destek vektör makineleri gibi birçok yöntem, sınıflandırma analizi için kullanılabilir. Bu yöntemler her ne kadar ortak özelliklere sahip olsa da aralarında önemli farklılıklar bulunmaktadır. Bu yöntemlerin daha geniş açıklanması sınıflandırma analizinin anlaşılması açısından faydalı olacaktır.

### 2.3.2.3.1. Karar Ağaçları (Decision Trees)

Karar ağaçları en yaygın kullanılan veri madenciliği yöntemlerinden biridir. Karar ağaçları algoritmaları bir veri setini bir ağacın dallarına benzer şekilde sınıflandırmak için kullanılmaktadır. İlk sınıflandırma kökten başlar, ilerledikçe sınıflandırma sayısı artar. Verinin her sınıflandırıldığı noktaya *düğüm* (node) denilmektedir. Karar ağaçları görselleştirilebilen analizlerdir. Bu akış diyagramları üzerindeki yapraksız her düğüm bir özellik üzerindeki bir testi, her bölünme bir testin çıktısını, her *yaprak düğüm* (leaf node, terminal node) sınıf etiketi ifade etmektedir (Han, Kamber ve Pei, 2012: 291). Örnek bir karar ağacı diyagramı Şekil 2.5’de verilmektedir.



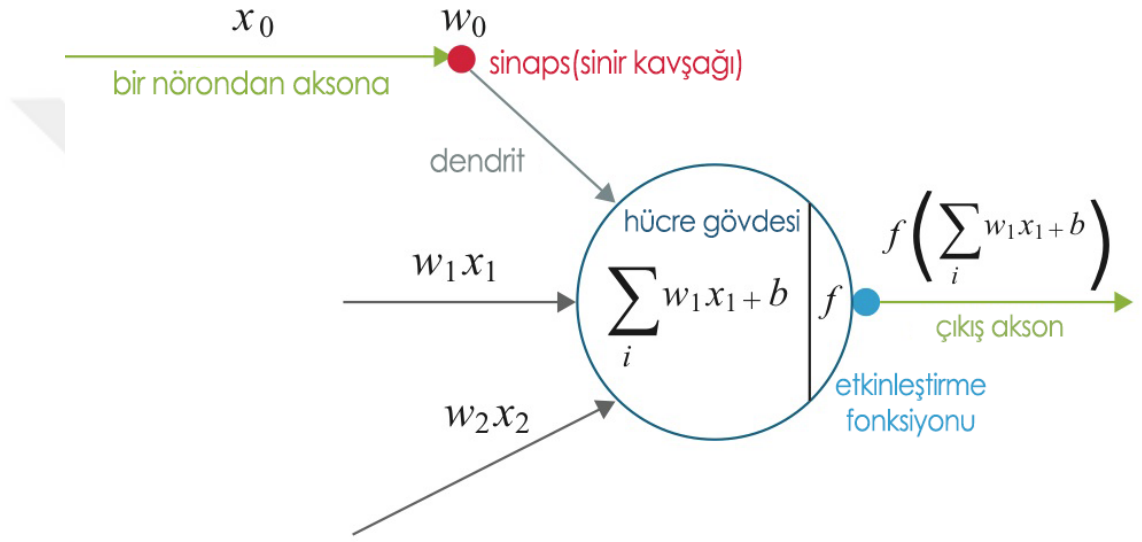
Şekil 2.5. Karar Ağacı Diyagramı (Han, Kamber, & Pei, 2012)

Şekil 2.5’de verilen örnek karar ağacı diyagramında bir müşterinin bilgisayar satın alıp almayacağı analiz edilmiştir. Bu diyagrama göre *orta yaşlı* bir müşteri her halükarda bilgisayar satın almaktadır. Fakat genç bir müşteri eğer öğrenci ise bilgisayar satın almakta, öğrenci değil ise satın almamaktadır.

### 2.3.2.3.2. Yapay Sinir Ağları (Artificial Neural Networks, ANN)

Yapay sinir ağları (YSA), kendisine verilen örnek öğrenme verisinden öğrenme yaparak, daha sonra karşılaştığı veri üzerinde karar verebilen bir makine öğrenmesi yöntemidir. Matematiksel kökenleri 1940'lara dayanan bu yöntem (Dreyfus, 2005: 2), 1969'da Marvin Minsky tarafından tek katmanlı perceptron'un XOR fonksiyonunu öğrenemeyeceği gösterilince YSA'ya olan ilgi azalmıştır. Oysaki bu döneme kadar çeşitli gelişmeler göstermiştir. Örneğin, 1957'de Sovyet matematikçi Kolmogorov, üç katmanlı bir sinir ağının, yeterli sayıda saklı nöron ve uygun doğrusal olmayan aktivasyon fonksiyonları verildiğinde herhangi bir sürekli (continuous) fonksiyonu gösterebildiğini ispatlamıştır. Yine aynı dönemde bir yapay nöron modeli olan perceptron geliştirilmiş, tek katmanlı bir perceptron'un birçok Bool fonksiyonunu öğrenebildiği görülmüştür. 1990'lara doğru geriyayılım yöntemi (back propogation algorithm) ile XOR problemini aşabileceği anlaşılrsa da (Filiz, 2017) destek vektör makinesi (support vector machines) yönteminin iyi sonuçlar vermesi, o dönemde veri setlerinin küçük olması, bilgisayarların işlem gücünün düşük olması gibi nedenlerle tekrar gözden düşmüştür (Akbaş, 2018). Son on yılda ise gelişen iletişim ve bilgisayar teknolojisi ile birlikte pek çok alanda kullanılmaya başlanmıştır.

YSA'lar insan beyninin çalışma sistemine benzer bir çalışma mantığına sahiptir. Ağ, birbirine bağlı yapay nöron grubundan oluşmaktadır. YSA'lar öğrenme aşamasında ağ boyunca akan içsel ve dışsal bilgiye dayalı, yapısı değişen ve uyarlanabilen bir yapıya sahiptir (Sammur ve Webb, 2017: 65; Haykin, 2001: 23-24). YSA'lar biyolojik nöronun girdi, işlem, çıktı sistemini taklit eden birçok basit, genellikle uyarlanabilen işlem birimlerine (yapay nöron) sahiptir. Bu nöronlar değişik etki seviyelerinde birbirlerine bağlanarak ağı oluşturmaktadır (Filiz, 2017). Şekil 2.6'da YSA'nın yapısı verilmektedir.



Şekil 2.6. Yapay Sinir Ağı (Karpathy, 2015)

$X$  değerleri burada bağımsız değişkenler gibi düşünülebilir.  $W$  ile girdilerin ağırlıkları ifade edilmektedir. Girdi değerleri ağırlık değerleri ile çarpılarak hücre gövdesine aktarılmaktadır. Hücre gövdesinde toplanan değerler çıkış fonksiyonuna girdi olmaktadır. Bu kısımda *step*, *sigmoid*, *relu* gibi fonksiyonlar kullanılmaktadır (Aggarwal C. C., 2018: 119). Bu kısımda üretilen değer ağın yapısına göre ya tekrar ağa girdi ya da çıktı olarak verilmektedir.

YSA ve onun özel bir hali olan *Derin Öğrenme (Deep Learning)* aynı zamanda yapay zekâ alanında da yaygınlıkla kullanılan yöntemlerden biridir. Bu kullanım alanıyla sürücüsüz araba, ses tanıma ve çeviri, görüntü tanıma gibi pek çok alanda kullanılmaktadır.

### 2.3.2.3.3. Lojistik Regresyon

Bir başka sınıflandırma yöntemi de lojistik regresyondur. Doğrusal regresyonda, bağımlı değişken numerik iken, lojistik regresyonda kategorik (nominal) veya sıralı kategorik (ordinal) olmaktadır. Doğrusal regresyonda  $Y$  bağımlı değişkeni  $-\infty$  ile  $+\infty$  değer alabiliyorken, lojistik regresyonda tahmin değeri 0 ile 1 arasında olmaktadır. Bu değer birimin sınıflara ait olma oranını vermektedir (Lomax ve Hahs-Vaughn, 2012: 709-710; Tabachnick ve Fidell, 2013: 439-442; Dunham, 2003: 86). Ayrıca lojistik regresyon, normal dağılım şartına da gereksinim duymadığından daha esnek bir yapıya sahiptir. Bütün bu nedenlerden dolayı lojistik regresyon sınıflandırma amaçlı olarak yaygın şekilde kullanılmaktadır.

Bu çalışmada araştırma modellerinin test edilmesinde de lojistik regresyon kullanılmıştır. Otel oda fiyatlarına etki eden değişkenler araştırılırken, bir nevi konaklama tesisinin sahip olduğu özelliklere göre, bağımlı değişken olan oda fiyat kategorilerinden, hangi kategoriye dâhil olduğunun olasılığı belirlenmektedir.

### 2.3.2.3.4. Naïve Bayes Sınıflandırma

Regresyon analizi ile birlikte en eski sınıflandırma yöntemlerinden biridir. Temeli *Bayes Teoremine* dayanmaktadır. Bu teoreme göre (Han, Kamber ve Pei, 2012: 311);

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

$P(A/B)$ , B olayının olması durumunda A'nın gerçekleşme olasılığını,  $P(A/B)$ , B olayının olma durumunda A'nın olma olasılığını  $P(A)$ , A'nın olma olasılığını,  $P(B)$ , B'nin olma olasılığını ifade etmektedir. *Naïve* terimi bağımsız değişkenlerin kendi aralarında bağımsız oldukları yani aralarında herhangi bir korelasyonun olmadığı anlamına gelmektedir (Roiger, 2017: 317).

Naïve Bayes sınıflandırılmasının *Multinomial* (çokterimli), *Bernoulli*, *Gaussian* gibi çeşitli türleri bulunmaktadır.

## 2.4. WEB MADENCİLİĞİ

Madencilik denildiğinde insanların aklına değerli bir maddenin toprak altından çıkarılması gelmektedir. Web madenciliği ise, isminden de anlaşılacağı üzere doğal kaynak madenciliği ile ortak özellikler taşımaktadır. Web madenciliği, webde yayınlanmış olan düzensiz veri veya bilgiyi, çeşitli yollarla elde etmek, düzenlemek ve bilgi çıkarımında bulunabilecek hale getirmek olarak düşünülebilir. Geçmiş yıllarda çok gündemde olmayan bu yöntem günümüzde pek çok farklı alanda kullanılmaktadır.

Oren Etzioni (1996) "*The World Wide Web: Bataklık mı? Yoksa Altın Madeni mi?*" başlıklı makalesinde, webin büyük miktarda yapılandırılmamış veri içerdiğini ve bunun bir veri kaynağı olarak kullanılıp kullanılmayacağını tartışmıştır. Bu çalışma web madenciliği kavramını, ilk dile getiren çalışmalardan biri olarak kabul edilmektedir (Sumathi ve Sivanandam, 2006).

İnternetin ilk yaygınlaşmaya başladığı yıllarda web madenciliğinin çok kullanılmamasının veya bilinmemesinin altında birçok neden yatmaktadır. Başka bir ifade ile web madenciliğinin kullanılmaya başlanması için bilgisayar, internet ve depolama birimleri gibi birçok alanda gelişmelerin olması gerekmektedir. Web madenciliğinin gelişmesine etki eden gelişmeler açıklandığında konu daha iyi kavranacaktır.

### 2.4.1. Web Madenciliğinin Ortaya Çıkmasına Etki Eden Gelişmeler

Web madenciliğinin gelişmesi, hem teknolojik hem de sosyolojik bazı gelişmelerin sonucunda olmuştur. Doğal olarak gelişmeler olduktan sonra, bilim adamları, araştırmacılar ve programcılar bu imkânlar ile bu yöntemi kullanmaya başlamıştır.

#### 2.4.1.1. Gelişen Bilgisayarlar ve Artan İşlem Gücü

Dünyanın ilk mikro işlemcisi 1971 yılında Intel firmasının piyasaya sunduğu Intel 4004 işlemcisidir. Intel 4004, 740 kHz saat hızına sahipti. Fakat bu işlemci kişisel bilgisayarlarda kullanılmamıştır. Kişisel bilgisayarlarda kullanılan ardılı Intel 8008 bir yıl sonra piyasaya çıkmıştır. Bu işlemci 500 kHz saat hızına sahipti ve sadece 16 KB hafıza adresleyebiliyordu (Reilly, 2003: 139; Deshmukh, 2005: 5-7). Günümüzde ortalama bir ev kullanıcısı saat hızı 4 GHz yaklaşan bir işlemciyi temin edebilmektedir. Kıyaslamayı daha açık vermek gerekirse Intel'in 1990'da ürettiği 512 işlemci içeren süper bilgisayarı

Touchstone Delta 32 Gflops işlem gücüne sahipken (Touchstone Delta, 2019), bugün bir kişisel bilgisayar bu hızlara çıkmakta, hatta aşmaktadır (CPU performance, 2019). Grafik işlemcilerde işlem gücü bunun çok daha ötesindedir. Bu gelişmeler daha önce gerçekleşmesi mümkün olmayan işlemleri yapabilmeyi olanaklı hale getirmiştir.

#### **2.4.1.2. Depolama Donanımlarının Ucuzlaması ve Kapasitelerinin Artması**

Veriyi işlemek kadar depolamak ta bir problemdir. 1990'ların ortalarında ortalama bir kişisel bilgisayarın sabit diski 1-2 GB iken bugün terabytelar ölçülebilecek bir kapasiteye ulaşmıştır. Bu da depolama birimlerinin 20 yıllık bir sürede yaklaşık 1000 kat arttığını göstermektedir. Ayrıca maliyetler de tam tersi yönde düşüş göstermiştir. Örneğin 1995'de GB başına maliyet 273\$ civarındayken günümüzde bu 0,025\$ civarındadır (McCallum, Disk Drive Prices (1955-2018), 2018).

Gelişmeler bununla sınırlı değildir. O yıllarda bilinmeyen bulut depolama, flaş disk, SSD gibi teknolojiler bugün hemen hemen herkesin kullandığı teknolojiler haline gelmiştir. Depolama birimlerinin gelişmesi muhtemel büyük bir darboğazın önüne geçmiş, günümüzde kullanılan pek çok teknolojinin önünü açmıştır. İnternetin ilk dönemlerinden farklı olarak günümüzde web siteleri yüksek çözünürlüklü resimlerden, videolardan ve interaktif uygulamalardan oluşmaktadır. Bütün bunlar depolama teknolojisinin gelişmesinin paralelinde olmuştur.

#### **2.4.1.3. İnternet Hız ve Bant Genişliğinin Artması**

Bant genişliği (bandwith) bir hat üzerinde bir saniyelik zaman diliminde iletilebilecek maksimum veri miktarını ifade etmektedir (Fisher, 2019). Dijital sistemlerde genellikle bit veri ölçü birimi ile ifade edilmektedir. Büyük veri miktarlarını iletmek için bant genişliği gerekli bir kavramdır. Öyle ki, Youtube, Netflix gibi servislerin sunulması yüksek bant genişliği ile mümkün hale gelmiştir. Bu tür servisler bant web trafiğinin önemli bir bölümünü kullanmaktadır (Laudon ve Laudon, 2016: 302). Web madenciliğinde, web siteleri/sunucuları üzerindeki veriyi indirme/çözümleme için bir parçasıdır. Bu nedenle bant genişliğinin yüksekliği ve yüksek internet hızı gerekli bir şeydir. İnternet hızının gelişmesine örnek vermek gerekirse, 1990'larda kullanılan çevirmeli ağ (dial-up, 36.6 bps) bir sistemde 700 MB bir dosya yaklaşık 48 saatte indirilirken günümüzde 24 Mbit/s ADSL bir bağlantıda yaklaşık 4 dakika sürmektedir.

Web madenciliğinde internetteki sunucu/web sitelerinden veri elde etme sürecin ilk ve elzem parçasıdır. Bu nedenle bunun hızlı ver sorunsuz bir şekilde gerçekleşmesi önemli bir gerekliliktir.

#### **2.4.1.4. İnternetin Genişlemesi ve Üretilen Veri Miktarının Artması**

Bugün yaklaşık dünya nüfusunun %40'nın internet kullandığı düşünülmektedir. Bu oran yaklaşık 4,1 Milyar kişiye karşılık gelmektedir. 1995'de bunun sadece %1 olduğu düşünülürse ne kadar hızlı bir şekilde yayıldığı anlaşılabilir (www.internetlivestats.com, 2019). İnternetin ilk ortaya çıktığı günlerdekine aksine günümüzde boyut olarak çok küçük cihazlar bile internete girebilmektedir. Örneğin bir cep telefonu, bir tablet, hatta bir kol saati bile internete girebilmektedir. Mobil bağlantılar ile kişiler kapsama alanı içerisinde bulunan herhangi bir yerden internete erişebilmektedir. Bu durum insanların üzerlerinde kolayca taşıyabildikleri aygıtlarla diledikleri zaman internete girebilmelerini olanaklı hale getirmiştir. Günümüzde yaygın kullanılan bloglar, sosyal ağlar ve çeşitli mobil uygulamalar ile kişisel kullanıcılar bile kolaylıkla internet üzerinde bilgi paylaşımında bulunabilmektedirler. Bu veri işlendiğinde çok önemli bilgilerin elde edilebileceği değerli bir maden haline gelmiştir. Pazarlama, sosyoloji, psikoloji, politika, ürün geliştirme, istihbarat gibi pek çok alanda kullanılabilir nitelikteki bu veri madeni, günümüzde etik anlamda da tartışılan bir konu haline gelmiştir.

#### **2.4.2. Web Madenciliğinde Kullanılan İçerik Yayınlama Teknolojileri**

Webde veriyi yayınlamak için HTML temel olmak üzere birçok yöntem vardır. Web madenciliği algoritmaları, bu diller ve kayıt şekilleri ile entegreli çalışarak veriyi elde etmektedir. Bu diller ve kayıt formatlarının en yaygın kullanılanları aşağıda açıklanmaktadır (Munzert, Rubba, Meißner ve Nyhuis, 2015: 9-10).

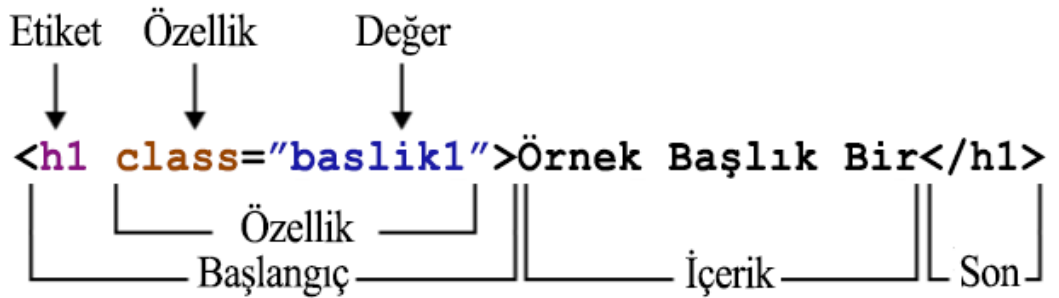
##### **2.4.2.1. Web Siteleri**

Webin ortaya çıkması Tim Berners-Lee'nin 1980'lerin sonunda IBM bilgisayarlar üzerinde çalışan SGML dilini temel alarak bugün de kullanılmakta olan HTML (HyperText Markup Language) dilini geliştirmesi ile olmuştur (Berners-Lee, Cailliau ve Groff, 1992). Bu dil bu gün de web sayfalarını yayınlamak için kullanılmaktadır. Geçen süre içerisinde PHP, ASP gibi diller ortaya çıksa da istemci üzerinde çalışan web tarayıcısının icra ettiği dil hala 5'ci sürümüne ulaşan HTML'dir.

Bir web tarayıcısı hedef URL adresine bağlandığında buradaki HTML kodlarını istemci aygıtına indirerek icra eder. Tim Berners-Lee'nin geliştirdiği *Nexus* ilk web tarayıcısıdır. Web tarayıcılar günümüzde en çok kullanılan yazılımlardandır.

HTML dili basit metin tabanlı kodlardan oluşan web sayfaları oluşturmak için kullanılan ve kendisinden sonra gelen nesnenin yapısını düzenleyen bloklardan oluşan bir dildir. HTML *element* veya *etiket (tag)* olarak adlandırılan bloklardan meydana gelmektedir. Bu elementler < karakteri ile > karakteri arasına yazılmaktadır. Tarayıcı, bu etiketleri tanıyarak kendisinden sonra gelen metin, grafik, video ve ses düzenlemekte veya çalıştırmaktadır. Benzer şekilde web madenciliği uygulamaları veya robotları bir web tarayıcı gibi çalışarak bu etiketler arasındaki bilgiyi okumaktadır. Web tarayıcıdan farklı olarak web madenciliği uygulamaları bu veriyi kullanıcının istediği biçimde ve oranda almaktadır.

Bir web sayfası <html> kodu ile başlamakta, </html> kodu ile bitmektedir. Bu etiketler tarayıcıya bu sayfanın bir web sayfası olduğu bilgisini vermektedir. Web sayfaları genel olarak iki kısımdan oluşmaktadır. Bunlar web sayfası hakkında çeşitli bilgilerin bulunduğu <head> kısmı ve web sayfasının içeriğinin yer aldığı <body> kısmıdır. Örneğin web sayfasının başlık bilgisi <head> bölümünde <title></title> etiketleri arasında yer almaktadır. <title> Web Sayfası 1 </title> şeklinde bir kod sayfayı icra eden tarayıcıya sayfanın başlığının *Web Sayfası 1* olduğu bilgisini vermektedir. Şekil 2.7'de örnek bir HTML etiketinin yapısı verilmektedir.



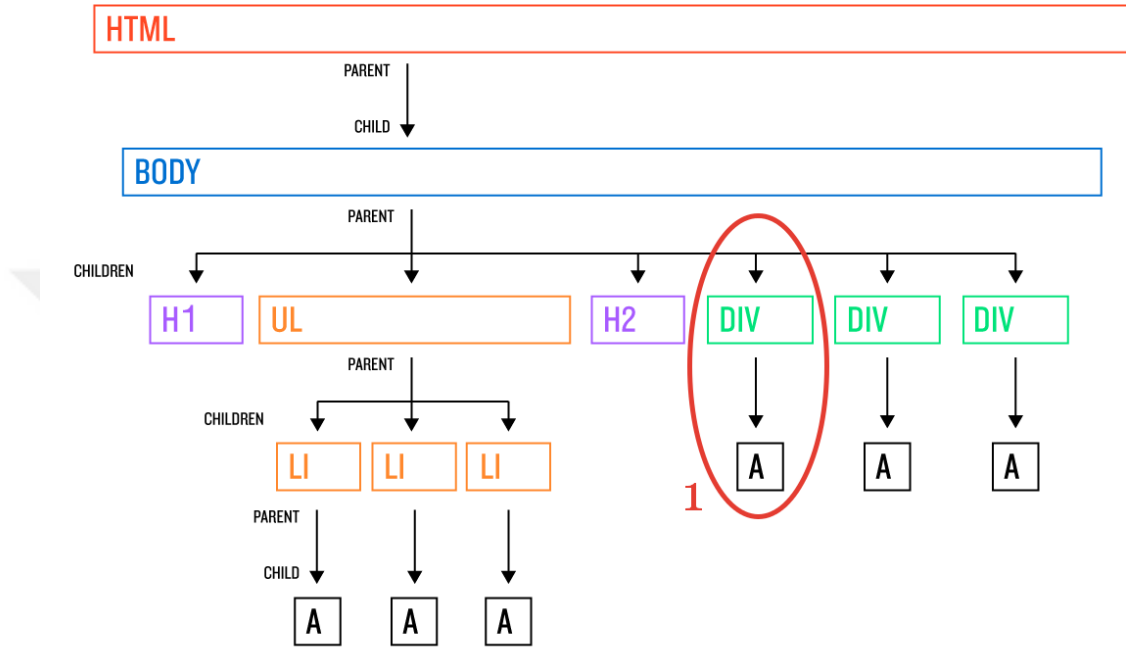
Şekil 2.7. HTML Etiketinin Yapısı

Web madenciliği uygulaması veya algoritması bir web tarayıcısı gibi çalışmaktadır. Bu uygulamalar Şekil 2.7'de örneği verilen *etiket* veya *özellik*e göre arama/filtreleme



yaparak sadece uygun olan içeriği almakta, bunu bir tablo veya veri tabanına kaydetmektedir.

HTML bazen iç içe bazen de alt alta bloklardan oluşmaktadır. Web madenciliği için bir algoritma oluştururken bu hiyerarşik yapıya dikkat etmek gerekmektedir. Document Object Model (DOM) olarak adlandırılan bu yapı Şekil 2.8’de verilmektedir.



Şekil 2.8. HTML'nin Hiyerarşik Yapısı (Donahoe & Hartl, 2019)

Web madenciliği algoritmaları bir sorgu dili gibi çalışmaktadır. Şekil 2.8’de verilen web sayfasında A nesneleri alınmak istendiğinde *Xpath* için *//a* şeklinde bir sorgu yapmak gerekmektedir. Fakat bu sorgu sayfadaki bütün A nesnelere yani linkleri seçecektir. HTML’nin hiyerarşik yapısından faydalanarak daha net bir sorgulama yapılabilir. Örneğin *Xpath* için *//a/div[1]/a* sorgusu sayfadaki bütün A nesnelere altındaki *birinci DIV* nesnesinin altındaki A nesnelere seçecektir [Şekil 2.8-1]. Web madenciliğinde HTML sorgularında kullanılan *Xpath* hakkında detaylı bilgi Web Madenciliğinde *XPath* Kullanımı bölümünde verilmektedir.

#### 2.4.2.2. CSS

Basamaklı Stil Şablonu anlamına gelen (Cascading Style Sheet, CSS) bu yapı 1996’da W3C tarafından yayınlanmıştır. CSS web sayfalarındaki nesnelere nasıl görüneceği bilgisini içeren veya nesnelere düzenleyen, HTML’ye ek bir dildir.

Günümüzdeki web sayfalarının büyük bir kısmında HTML ile birlikte bu dil kullanılmaktadır.

CSS kodları metin tabanlı ayrı bir dosya içerisinde tanımlanabileceği gibi, HTML dosyası içerisinde `<style> </style>` etiketleri kullanılarak tanımlanabilmektedir. Ayrıca bir HTML nesnesine `style="color: green;"` özelliği içerisinde istenilen CSS kodu yazılabilmektedir.

HTML kodlarında olduğu gibi CSS kodları için de sorgulamalar yapılabilmektedir. Şekil 2.9'da örnek CSS kodları verilmektedir.

```

▶ <div class="col-sm-7 col-xs-12">...</div>
▼ <div class="col-sm-5 col-xs-12">
  ▶ <span class="night-info night-info-right">...</span>
  ▼ <div class="rsi-prices-box">
    ▼ <div class="rsi-prices">
      <!---->
      <span class="rsi-calc">621 TL</span>
      <span class="rsi-discount hidden-item">Kazancınız : 0 TL</span>
    </div>
  </div>

```

**Şekil 2.9.** CSS Etiketlerinin Sorgulanması

Şekil incelendiğinde `class` özelliği `rsi-prices` olan bir `div` nesnesi görülmektedir. Buradaki `rsi-prices` aslında bir CSS stildir. HTML kodları için yapılan sorguya benzer şekilde CSS kodları için de sorgular yapılabilmektedir. Günümüzdeki web sayfalarının çoğu CSS stilleri kullandığından web madenciliği sorgularında bu stil etiketleri sıklıkla kullanılmaktadır.

### 2.4.2.3. XML

Geniřletilebilir İřaretleme Dili (eXtensible Markup Language) kısaca XML web üzerinde veri alışveriři için kullanılan en yaygın formatlardan biridir. HTML web sayfasındaki bilginin nasıl görüneceđi ile ilgiliyken XML'nin temel amacı veriyi depolamaktır. Son yıllarda XML ve türevleri web uygulamaları arasında veri alışverişinde sıklıkla kullanılmaktadır (Munzert, Rubba, Meißner ve Nyhuis, 2015: 10). Bu nedenle web madenciliđi için önemli bir kavramdır. Hatta web madenciliđi uygulamalarında kullanılan Xpath dili, ilk XML sorgulamaları için geliştirilmiştir. Ařađıda Őekil 2.10'da örnek bir XML kaydı verilmektedir.

```

1 <?xml version="1.0" encoding="UTF-8" ?>
2 <root>
3   <name>Murat</name>
4   <surname>Aslan</surname>
5   <age>28</age>
6   <address>
7     <city>Malatya</city>
8     <county>Turkey</county>
9   </address>
10  <contact>
11    <type>Telephone</type>
12    <number>+905115604344</number>
13  </contact>
14  <contact>
15    <type>e-mail</type>
16    <address>murat-aslan@walkastory.com</address>
17  </contact>
18 </root>

```

**Őekil 2.10.** XML Örnek Veri Kaydı

Dikkat edileceđi üzere sözdizimi (syntax) HTML'ye benzemektedir. Etiketler kayıt edilecek içeriđe göre kullanıcı tarafından tanımlanmıştır.

#### 2.4.2.4. JSON

JSON (JavaScript Object Notation) XML'nin amacına benzer şekilde, web uygulamaları ve sunucular arasında veri alışverişi ve veri kaydı amacıyla XML'ye bir alternatif olarak geliştirilmiştir (Bolton, 2019). JSON'un da sözdizimi kolay anlaşılır yapıdadır. Daha önce XML için verilen örnek JSON'da Şekil 2.11'deki gibi olacaktır.

```

1 {
2   "name": "Murat",
3   "surname": "Aslan",
4   "age": 28,
5   "address": {
6     "city": "Malatya",
7     "county": "Turkey",
8   },
9   "contact": [{
10    "type": "Telephone",
11    "number": "+905115604344"
12  },
13  {
14    "type": "e-mail",
15    "address": "murat-aslan@walkastory.com"
16  },
17  ]
18 }

```

Şekil 2.11. JSON Örnek Veri Kaydı

XML ile kıyaslandığı zaman JSON kodları daha kısa ve anlaşılması daha kolaydır (Mitchell, 2018: 178-179).

Her iki kayıt türü de web madenciliği için önemli kavramlardır. Verisini paylaşan pek çok web sitesi artık birer standart olan bu dosya türlerinde paylaşım yapmaktadır. Örneğin Twitter'daki eğilimleri (trends) analiz etmek isteyen biri Twitter'ın JSON formatında yayınlamış olduğu verileri alabilir (Munzert, Rubba, Meißner ve Nyhuis, 2015: 10). XML ve JSON formatındaki verileri toplamak, bu veriler sistemli ve düzenli bir yapıda olduğu için HTML'den çok daha kolaydır.

#### 2.4.2.5. AJAX

AJAX (Asynchronous JavaScript and XML) HTML, CSS, XML, JSON, JavaScript gibi pek çok teknolojiyi bir arada kullanan, bir yapıdır. AJAX yapısı XMLHttpRequest nesnesini kullanarak arka planda az miktarda veri değişimi ile sayfada değişiklik yapmayı sağlamaktadır. Bu teknik sayfaya yeni içerik eklenirken eski içeriğin

yeniden sunucudan indirilmesini önlemektedir. Böylece hem hız konusunda hem de bant kullanımını konusunda fayda sağlamaktadır.

Günümüzde AJAX yöntemi kullanılan web siteleri yaygındır. Web madenciliği yaparken bu sayfalara uygun algoritma ve kütüphaneler kullanmak gerekmektedir. AJAX yöntemini kullanan bu dinamik web sayfaları, doğru yöntem kullanılmadığı takdirde eksik bir şekilde indirilmektedir.

Facebook, Twitter, Yelp gibi bilinen pek çok site bu sistemi sayfaların o an görünen kısmından geri kalan kısmını gizlemek için kullanmaktadır. Sayfanın sonuna doğru inildiğinde gizli olan içerik sunucuya istek yapılarak ekrana getirilmektedir. Daha önce var olan kısımlar için ise yeniden istekte bulunulmamaktadır.

Verimli bir web madenciliği için webde içerik ve veri yayınlamak için kullanılan HTML, CSS, XML, JSON ve AJAX gibi teknolojiler hakkında yeterli bilgiye sahip olmak gerekmektedir.

#### **2.4.3. Web Madenciliği Yöntemleri**

Web sitelerini görüntülemek için kullanılan tarayıcılar erişilen sitenin içeriğini görüntülemek için yüklü olduğu ağıta indirmektedir. Fakat bu veri içerisinden sadece önemli olan veya istenen kısımları almak zahmetli, fazla vakit alan ve gereksiz büyüklükteki veri ile uğraşmayı gerektiren bir iştir. Web madenciliği hedef adresteki veri içerisinden sadece istenilen kısmı almayı sağlayan bir yöntemdir.

Web madenciliği birçok farklı şekilde yapılabilir. Bunlardan ilki, Python veya R gibi programlama dilleri aracılığı ile bazı kütüphaneler ve kodlar kullanılarak yapılan web madenciliğidir. Bu tür web madenciliği bir oranda programlama bilgisi gerektirse de, araştırmacıya oldukça esnek bir yapı sunmaktadır. Veri madenciliği yapan kişi, programlama dilinin gücünü kullanarak veriyi istediği şekilde elde edebilmekte ve işleyebilmektedir.

Bir diğer yöntem çeşitli yazılımlar aracılığı ile yapılan web madenciliğidir. Bu yöntem kullanıcıya programlama dili bilmeden bazen çok basit komutlar ile veri elde etmeyi sağlamaktadır. Fakat programlama dilleri kadar esnek değildir. Web madenciliği yazılımının imkân verdiği ölçüde sayfalar taranmakta ve veri elde edilmektedir.

Başka bir yöntemde bu hizmeti sağlayan servisler aracılığı ile web madenciliği yapmaktır. Mozenda, Octoparse ve Webscraper gibi servisler web üzerinden bu hizmeti sağlamaktadır. Çalışma mantıkları web madenciliği yazılımlarına benzemektedir. Fakat bu hizmeti web siteleri veya tarayıcılara yüklenen eklentiler aracılığıyla vermektedirler.

#### 2.4.3.1. Python İle Web Madenciliği

Python programlama dili 1980'lerin sonlarına doğru Guido van Rossum tarafından geliştirilen bir dildir. Ancak 2018'de 3.x sürümüne erişen Python günümüzde yaygın kullanılan dillerden biri haline gelmiştir (Gowrishankar ve Veena, 2019: 7-9). Python günümüzde web madenciliği konusunda en yaygın kullanılan dildir. Bunun pek çok nedeni vardır (Koshy, Best Programming Languages for Web Scraping, 2017). Python için yazılmış BeautifulSoup ve Scrapy gibi kütüphaneler web madenciliğini oldukça kolay ve verimli hale getirmiştir.

Python'da web madenciliği için BeautifulSoup (Richardson, 2015) ve Scrapy gibi kütüphaneler kullanılmaktadır. Bu kütüphaneler bir web sitesinden belirtilen HTML ve CSS etiketine sahip olan nesnelere filtrelemek için kullanılmaktadır. Bu iki kütüphanenin birbirine karşı avantaj ve dezavantajları vardır (Yin, 2019). Aşağıda BeautifulSoup kütüphanesi kullanılarak yapılan bir web madenciliği uygulaması verilmektedir.

```

1 import requests
2 from bs4 import BeautifulSoup
3 import pandas

4 istek = requests.get("https://www.tripadvisor.com.tr/Hotels-
  g293969-Turkey-Hotels.html")
5 icerik = istek.content

6 duzenli = BeautifulSoup(icerik, "html.parser")
7 hepsi = duzenli.find_all("div", {"class": "prw_rup
  prw_meta_hsx_responsive_listing ui_section listItem"})

8 liste = []

9 for oge in hepsi:
10     veri = {}
11     veri["Otel Adı"] = oge.find("div", {"class":
  "listing title"}).text
12     veri["Fiyat"] = oge.find("div", {"class": "price
  autoResize"}).text
13     veri["Yorum Sayısı"] = oge.find("a", {"class":
  "review_count"}).text
14     veri["Link"] = oge.find("a", {"class": "property_title
  prominent"}).attrs['href']

```

```
15 liste.append(veri)
```

### Şekil 2.12. Python BeautifulSoup Kütüphanesi İle Web'den İçerik Alma

Şekil 2.12’de BeautifulSoup kütüphanesi kullanılarak web’den içerik alma örneği verilmektedir. Kaynak kodlar incelenecek olursa *requests*, *BeautifulSoup* ve *pandas* kütüphaneleri kullanılmak üzere projeye dâhil edilmiştir. *BeautifulSoup*’u verimli bir şekilde kullanabilmek için bu kütüphanelere ihtiyaç duyulmaktadır. *Scrapy*’de ek bir kütüphaneye ihtiyaç yoktur. Bu kütüphanelerden *requests*, web sayfasından kaynak kodları, çerezleri (cookies) ve daha birçok şeyi almak için kullanılmaktadır. Python’un standart kütüphanesi *urllib2* de bu amaç için kullanılabilir alternatiflerdendir (Mitchell, 2018: 151-152). *pandas* kütüphanesi ise elde edilmiş veriden veri çerçevesi (data frame) oluşturmak için gereklidir.

*Requests* kütüphanesinin *get()* fonksiyonu kullanılarak *tripadvisor.com* sitesindeki Türkiye otellerinin bulunduğu sayfa *istek* değişkenine aktarılmaktadır. Sonraki satır *content* komutu ile içeriği *byte* cinsine çevirmektedir (Ronquillo, 2019). Değişken bu şekilde görüntülendiğinde düzenlenmemiş *HTML* kodları elde edilecektir. 6’ncı satırda *BeautifulSoup* kütüphanesi kullanılarak kodlar düzenlenerek *duzenli* adlı değişkene aktarılmaktadır.

*BeautifulSoup*’da kaynak kodlar içerisinden istenilen bir değeri almak için *find* veya *find\_all* fonksiyonu kullanılmaktadır. *find* belirtilen özelliklere sahip ilk değeri getirirken, *find\_all* o özellikteki bütün değerleri getirmektedir. Burada asıl önemli olan nokta elde edilmek istenen verinin *HTML* ve *CSS* etiketini doğru bir şekilde belirtmektir. Bu komutlar *duzenli* değişkeni içerisine aktarılmış olan *HTML* kodları içerinden belirtilen özelliğe sahip olan veriyi almaktadır. Şekil 2.12’deki örnekte 4’ncü satırdaki *tripadvisor.com* adresinden sayfada otel isimlerini içeren *div* nesnesinin içeriği *hepsi* değişkenine aktarılmaktadır. Bu *div* nesnesinin sınıfı (*class*) “*prw\_rup prw\_meta\_hsx\_responsive\_listing ui\_section listItem*” dir.

Hedef sayfada elde edilmek istenen nesnenin üzerine sağ tıklanıp (Chrome: İncele;/ Opera: Öğeyi Denetle; Firefox: Öğeyi İncele) tarayıcının ilgili komutu tıklandığında açılan yeni pencerede sayfanın kaynak kodları o nesne seçili bir şekilde ekrana gelmektedir. Şekil 2.13’de içeriği alınmak istenen web sayfasının kaynak kodları verilmiştir.

The screenshot displays a web browser interface with a list of hotels. On the left, there are filter options under 'İlanaklar' (Amenities) including 'Ücretsiz Wifi', 'Kahvaltı dahil', 'Havuz', and 'Ücretsiz Otopark'. Below these, there is a 'Fizgin puanı' (Rating) section with five green circles and the text 've yukarısı' (and above). The main content area shows two hotel cards. The first card is for 'Hotel Amira Istanbul', featuring a night photograph of the hotel building, a 'Çok Satan' (Best Seller) badge, the Agoda logo, and a price of '₺831'. A yellow button labeled 'Fiyata Bakın' (View Price) is visible. The second card is for 'Akra Otel'. The bottom part of the image shows the browser's developer tools with the source code for the selected hotel card highlighted. The code includes various attributes like 'class', 'data-placement-name', 'data-prwidget-name', and 'data-impkey'.

Şekil 2.13. Hedef Web Sayfasının Kaynak Kodları

Şekilde `<div class="prw_rup prw_meta_hsx_responsive_listing ui_section listItem">` elde edilmek istenen *otel adlarını* içinde barındırmaktadır. `<div></div>` etiketi HTML belgesinde bir bölümü belirtmek için kullanılan bir yapıdır.

Şekil 2.12’de 7’nci satırdaki komutla hedef sayfada sadece içeriği alınmak istenen kısım *hepsi* adlı değişkene aktarılmaktadır. 8’nci satırda daha sonra *for* döngüsü ile elde edilecek olan veri için bir liste oluşturulmaktadır. *For* döngüsünde hedef sayfadan elde edilmek istenen veri için tek tek CSS etiketleri bulunmuştur. *Otel Adı* class değeri *listing\_title* olan *div* nesnesi içerisindedir. Şekil 2.14’de kaynak kodları içerisinde bu gösterilmektedir.



```

▼<div class="prw_rup prw_meta_hsx_listing_name listing-title" data-prwidget-name="meta_hsx_listing_name" data-prwidget-init>
  ▼<div class="listing_title">
    <a target="_blank" href="/Hotel_Review-g293974-d1674691-Reviews-Hotel_Amira_Istanbul-Istanbul.html" id="property_1674691"
      class="property_title prominent" data-clicksource="HotelName" onclick="return false;" dir="ltr">Hotel Amira Istanbul</a>
  </div>
</div>

```

**Şekil 2.14.** Otel Adının Bulunduğu Div Nesnesi

Aratılmak istenen nesne için sadece *class* özelliği değil, diğer özellikler de parametre olarak kullanılabilir.

```

15 for oge in hepsi:
16     linkler = oge.find("a", {"target":
    "_blank"}).attrs["href"]
17     print(linkler)

```

**Şekil 2.15.** BeautifulSoup Kütüphanesi İle Web'den İçerik Alma

Şekil 2.15'de *target* özelliği *\_blank* olan *a* nesnelerinin yani linklerin *href* özellikleri *linkler* değişkenine aktarılmaktadır. For döngüsü ile *hepsi* değişkeni içerisinde bu özelliklere sahip bütün değerler tek tek alınmaktadır.

Şekil 2.16'da elde edilen ve listeye dönüştürülen veri Şekil 'da veri çerçevesine dönüştürülmektedir.

```

18 df = pandas.DataFrame(liste)
19 df = df[["Otel Adı", "Fiyat", "Yorum Sayısı", "Link"]]
20 df

```

**Şekil 2.16.** Pandas Kütüphanesi İle Veri Çerçevesi Oluşturma

Elde edilen *df* adındaki veri çerçevesinin ekran çıktısı şekildeki gibi olacaktır. 19'ncü satırda varsayılan (default) olarak alfabetik sıralanan sütunlar otel adından başlayacak şekilde yeniden sıralanmıştır.

	Otel Adı	Fiyat	Yorum Sayısı	Link
0	Radisson Blu Hotel & Spa - Istanbul Tuzla	₺385	671 yorum	/Hotel_Review-g293974-d5281394-Reviews-Radisso...
1	Idylle Hotel	₺596	165 yorum	/Hotel_Review-g293974-d661625-Reviews-Idylle_H...
2	Crowne Plaza Hotel Antalya	₺313	2.661 yorum	/Hotel_Review-g297962-d1465019-Reviews-Crowne_...
3	FLORYA HOUSE HOTEL	₺196	65 yorum	/Hotel_Review-g293974-d14397053-Reviews-Florya...
4	Radisson Blu Hotel Istanbul Ottomare	₺507	354 yorum	/Hotel_Review-g293974-d9459106-Reviews-Radisso...
5	Venus Suite Hotel	₺245	359 yorum	/Hotel_Review-g297992-d7998780-Reviews-Venus_S...
6	Wyndham Grand İstanbul Kalamış Marina Hotel	₺818	1.006 yorum	/Hotel_Review-g293974-d3588557-Reviews-Wyndham...
7	Pera Palace Hotel	₺621	1.442 yorum	/Hotel_Review-g293974-d294949-Reviews-Pera_Pal...
8	Wyndham Grand İzmir Özdilek	₺371	884 yorum	/Hotel_Review-g298006-d301057-Reviews-Wyndham_...
9	Akra Otel	₺486	1.434 yorum	/Hotel_Review-g297962-d295121-Reviews-Akra-Ant...
10	Hyatt Regency İstanbul Ataköy	₺709	526 yorum	/Hotel_Review-g293974-d7262350-Reviews-Hyatt_R...
11	Park Inn by Radisson İstanbul Atasehir	₺274	157 yorum	/Hotel_Review-g293974-d13221269-Reviews-Park_I...
12	Titanic Business Kartal	₺293	1.741 yorum	/Hotel_Review-g293974-d1059302-Reviews-Titanic...
13	Radisson Blu Hotel, İstanbul Sisli	₺1.586	1.356 yorum	/Hotel_Review-g293974-d4562615-Reviews-Radisso...
14	Ephesus Palace	₺116	97 yorum	/Hotel_Review-g293976-d10492835-Reviews-Ephesu...
15	Hilton İstanbul Bosphorus	₺1.132	3.065 yorum	/Hotel_Review-g293974-d294614-Reviews-Hilton_I...
16	Regie Ottoman İstanbul	₺615	307 yorum	/Hotel_Review-g293974-d7266122-Reviews-Regie_O...
17	Bellamaritimo Hotel	₺138	285 yorum	/Hotel_Review-g297992-d1768828-Reviews-Bellama...

**Şekil 2.17.** Python ve BeautifulSoup İle Elde Edilen Verinin Çıktısı

Şekil 2.17’de hedef sayfadan elde edilmiş *otel adı*, *fiyat*, *yorum* ve *link* verisi *pandas* kütüphanesi aracılığıyla veri çerçevesine dönüştürülmüştür. Şekil 2.17’de verilmeyen her sayfada toplam 30 adet otel bilgisi yer almaktadır.

Python ve BeautifulSoup web madenciliği için yaygın kullanılan yöntemlerden biridir. Bir programlama dilinin sağladığı esneklik bu yöntemde mevcuttur. Yukarıdaki örnekte basit bir şekilde hedef sayfadan bazı verilerin alınması gösterilmektedir. BeautifulSoup ve Requests kütüphaneleri ile sadece içerik alma dışında GET, OPTIONS, POST, PUT gibi komutlar da kullanılabilir (Reitz, Quickstart: Requests, 2018).

#### 2.4.3.2. R İle Web Madenciliği

R özellikle veri analizinde ve istatistiksel hesaplamalarda son yıllarda oldukça yaygınlaşan bir programlama dilidir. Açık kaynak kodlu oluşu, birçok programcının yeni kütüphaneler yazarak programın gelişmesine destek vermesini sağlamıştır. R, veri analizi, istatistik vb. hesaplamalar konusunda kullanıcının işini kolaylaştırıcı birçok kütüphaneye sahiptir (Kabacoff, 2011: 3-5). CRAN olarak adlandırılan ağdan bu kütüphaneler ücretsiz bir şekilde indirilebilmektedir. Daha önce Python ile yapılan örnek web madenciliği işlemi aşağıdaki örnekte R’da yapılmaktadır. R’da web kazıma için bu örnekte *rvest*

kütüphanesi kullanılmaktadır. *XML* gibi alternatif kütüphaneler de kullanılabilir (Boehmke, 2015).

```

1 library("rvest")
2 site_url <- read_html("https://www.tripadvisor.com.tr/Hotels-
g293969-Turkey-Hotels.html")
3 otel <- site_url %>% html_nodes(".property_title") %>%
html_text()
4 fiyat <- site_url %>% html_nodes(".price-wrap .price") %>%
html_text()
5 fiyat <- gsub("\u20ba", "", fiyat)
6 yorum_sayisi <- site_url %>% html_nodes(".review_count") %>%
html_text()
7 yorum_sayisi <- gsub(" yorum", "", yorum_sayisi)
8 link <- site_url %>% html_nodes(".prominent") %>%
html_attr("href")
9 link <- paste("https://www.tripadvisor.com.tr", link, sep = "")
10 data.frame(otel, fiyat, yorum_sayisi, link, stringsAsFactors =
FALSE) %>% View()

```

Şekil 2.18. R'da rvest Kütüphanesi İle Web'den İçerik Alma

Şekil 2.18 incelendiğinde 1'nci satırda *rvest* kütüphanesinin projeye dâhil edildiği görülmektedir. Daha önce bahsedildiği gibi *XML* kütüphanesi de web sayfalarından veri aktarmak için kullanılmaktadır. Eğer *rvest* yüklü değil ise *install.packages("rvest")* komutu ile yüklenmelidir. *site\_url* değişkenine hedef adresteki kaynak kodlar aktarılmaktadır. 3'ncü satırda *site\_url* içerisinde *.property\_title* etiketine sahip nesnelere aranmaktadır. Kaynak kod içerisinde ilgili nesne Şekil 2.19'da verilmektedir.

```

▼<div class="listing_title">
  <a target="_blank" href="/Hotel_Review-g293974-d5281394-Reviews-Radisson_Blu_Hotel_Spa_Istanbul_Tuzla-
Istanbul.html" id="property_5281394" class="property_title prominent" data-clicksource="HotelName" onclick=
"return false;" dir="ltr">Radisson Blu Hotel & Spa - Istanbul Tuzla</a>
</div>
</div>

```

Şekil 2.19. Otel Adının Kaynak Kodu

Örneğin devamında *html\_text()* fonksiyonu ile *.property\_title* özelliğine sahip nesnelerdeki metin (text) *otel* değişkenine alınmaktadır. İşlem tamamlandıktan sonra *print(otel)* komutu işletildiğinde hedef sayfada bulunan toplam 30 otelin adları listelenecektir. 4'ncü satırda otel adlarının alınmasına benzer bir şekilde sitede otele karşılık verilen fiyat değerleri alınmaktadır. Otel oda fiyatının kaynak kodları Şekil 2.20'deki gibidir.

```

▼<div class="priceBlock ui_column is-12-tablet">
  ▼<div class="price-wrap ">
    ▶<div class="xwrap">...</div>
      <div class="price __resizeWatch" data-sizegroup="mini-meta-price" data-clickpart="chevron_price" data-index="0">₺386</div>
    </div>

```

**Şekil 2.20.** Otel Fiyatının Kaynak Kodu

Kaynak kod incelendiğinde otelin oda fiyatı *price-wrap* sınıfındaki *div* nesnesinin içindeki *price* sınıfından bir *div* nesnesinin içerisinde. Doğru veriyi alabilmek için tarayıcılara eklenti olarak eklenen araçlar bulunmaktadır. Bu araçlar yardımı ile doğru etiketi (tag) bulmak daha kolay bir hale gelmektedir. Örnekte her iki etiket te *html\_nodes(".price-wrap .price")* şeklinde arka arkaya kullanılarak fiyat değeri elde edilmiştir.

Fiyat metnindeki *TL* simgesi *gsub()* fonksiyonu ile veriden temizlenmiştir. Nesnedeki metin dışında bir özellik alınmak isteniyorsa *html\_attr("href")* fonksiyonu kullanılmaktadır. 8'nci satırda *href* özelliği yani *a* nesnesinin linki alınmaktadır. Son olarak vektör halindeki otel, fiyat, yorum sayısı ve link değişkenlerinden bir veri çerçevesi oluşturulmaktadır. Veri çerçevesi oluşturulduktan sonra verinin ekran çıktısı Şekil 'de verilmektedir.

otel	fiyat	yorum_sayisi	link
1 FLORYA HOUSE HOTEL	197	65	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
2 Basileus Otel	344	2.009	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
3 Radisson Blu Hotel & Spa - Istanbul Tuzla	388	671	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
4 Aydınli Cave Hotel	432	1.372	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g297983-d...">https://www.tripadvisor.com.tr/Hotel_Review-g297983-d...</a>
5 Sura Hagia Sophia Hotel	854	2.191	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
6 Idylle Hotel	586	172	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
7 Crowne Plaza Hotel Antalya	316	2.666	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g297962-d...">https://www.tripadvisor.com.tr/Hotel_Review-g297962-d...</a>
8 Akra Otel	488	1.436	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g297962-d...">https://www.tripadvisor.com.tr/Hotel_Review-g297962-d...</a>
9 Hilton Istanbul Bosphorus	948	3.071	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
10 The Marmara Taksim	709	1.532	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
11 Kayakapi Premium Caves - Cappadocia	800	1.532	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g297989-d...">https://www.tripadvisor.com.tr/Hotel_Review-g297989-d...</a>
12 Hotel Amira Istanbul	708	3.752	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
13 Walton Hotels Galata	334	346	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
14 Cappadocia Cave Suites	500	1.086	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g297983-d...">https://www.tripadvisor.com.tr/Hotel_Review-g297983-d...</a>
15 Radisson Blu Hotel Istanbul Ottomare	509	359	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
16 Pera Palace Hotel	831	1.443	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>
17 Park Inn by Radisson Istanbul Atasehir	270	157	<a href="https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...">https://www.tripadvisor.com.tr/Hotel_Review-g293974-d...</a>

Şekil 2.21. R ve rvest İle Elde Edilen Verinin Çıktısı

### 2.4.3.3. Web Madenciliğinde XPath Kullanımı

XPath, XML dosyaları içerisinde sorgulamalar yapmak için geliştirilmiş bir dildir. İlk etapta XML için geliştirilse de daha sonra pek çok dil için uyarlanmıştır. XPath, XSLT'nin bir parçası olsa da tek başına da kullanılabilir (Kay, 2008: 8). XPath kullanım kolaylığı ve kullanılabilirliği nedeniyle web sayfalarından içerik elde etmek için yaygın şekilde kullanılmaktadır. Daha önce örnekleri verilen Python ve R'da da XPath kullanılmaktadır (Horner, 2015). Aşağıda **Tablo 2.4**'de en çok kullanılan bazı XPath ifadeleri ve bu ifadelerin işlevleri verilmektedir (Pham, 2016).

**Tablo 2.4.** Yaygın Kullanılan Bazı XPath İfadeleri ve İşlevleri

İfade	Açıklama
nodename	"nodename" adındaki elementi seçer.
/	Eğer ilk ifade olarak / işareti ile başlıyorsa kök dizinden aramaya başlar. Sonraki her / bir alt düğümü ifade eder.
//	Belgenin içerisinde düğüm nerede olursa olsun arama/seçme yapar.
.	Geçerli düğümü seçer.
..	Bir üst düğümü seçer.
@	Düğümün herhangi bir özelliğini seçer.
[@attribute = 'value']	Belirli bir özelliğe sahip olan düğümü seçer
text()	Düğümdeki metni seçer.
	Birden çok sorguyu birleştirir.

Hedef sayfada elde edilmek istenen nesnenin üzerine sağ tıklanıp *incele* komutu seçildiğinde ilgili nesne merkezde olacak şekilde web sayfasının kaynak kodları ekrana gelecektir. Burada otel adı için bu işlem yapılmıştır. Kaynak kodlar **Şekil** verilmektedir.

Kodlar incelendiğinde *otel adının* bir *a* nesnesi altında olduğu görülmektedir. Bunu

```
<a target="_blank" href="/Hotel_Review-g298033-d5975414-Reviews-Green_Nature_Diamond_Hotel-Marmaris_Marmaris_District_Mugla_Province_Turkish_Aegean_Co.html" id="property_5975414" class="property_title prominent" data-clicksource="HotelName" onclick="return false;" dir="ltr">Green Nature Diamond Hotel
</a>
</div>
```

### Şekil 2.22. Otel Adı Nesnesinin HTML/CSS Kodları

XPath ile alabilmek için bu nesneni seçilmesi gerekmektedir. *//a* şeklinde bir sorgu web sayfasındaki bütün *a* nesnelerini seçecektir. Bu nedenle doğru seçimi yapabilmek için daha özel bir özellik belirtmek gerekir. *//a[@class="property\_title prominent"]* şeklindeki bir sorgu *class* özelliği *property\_title prominent* olan *a* nesnelerini seçecektir. Otel adına sağ tıklayıp *incele* komutu seçilerek açılan tarayıcının *elements* sekmesinde *Ctrl+F* tuş kısayolu ile bu komut işletildiğinde sayfada yer alan 30 otele ait isim bilgisinin seçildiği görülecektir.

Aynı örnekte *a* nesnesinin *href* özelliği olarak otelin linkinin olduğu görülmektedir. Eğer bu link alınmak isteniyorsa, *//a[@class="property\_title prominent"]/@href* şeklinde bir sorgu çalıştırmak gerekmektedir. Bu sorgu açıklanacak olursa, belge içerisinde yer alan bütün *a* nesnelerinde *class* özelliği *property\_title prominent* olanların *href* özelliği seçilmektedir.

Aynı sayfada oteller için verilen gecelik konaklama ücretleri de yer almaktadır. Bunlar alınmak isteniyor olsun. Fiyat üzerine sağ tıklayıp kaynak kod ekranı açıldığında **Şekil** ekrana gelecektir.

```

▼ <div class="priceBlock ui_column is-12-tablet">
  ▼ <div class="price-wrap">
    ▶ <div class="xwrap">...</div>
      <div class="price __resizeWatch" data-sizegroup="mini-
        meta-price" data-clickpart="chevron_price" data-index="0">
        ₺456</div>
      </div>
    ▼ <div class="provider __resizeWatch" data-sizegroup="mini-
      meta-provider" data-clickpart="chevron_provider" data-index=
        "0">
      
      ▶ <span class="provider_text">...</span>
  
```

**Şekil 2.23.** Fiyat Nesnesinin HTML/CSS Kodları

Kaynak kodlar incelendiğinde 456 TL olan fiyatın *price \_\_resizeWatch* class etiketine sahip bir *div* nesnesinin altında olduğu görülmektedir. Fakat *//div[@class="price \_\_resizeWatch"]* şeklinde bir sorgu ile bu alınmak istendiğinde her konaklama tesisi için birden çok fiyat alınmaktadır. Bunlar içerisinde en iyi fiyatı veren sadece alınmak isteniyor olsun;

```

//div[@class="priceBlock ui_column is-12-tablet"]//div[@class="price
  __resizeWatch"]

```

Şeklinde bir sorgu *priceBlock ui\_column is-12-tablet* altında bulunan *price \_\_resizeWatch* class etiketine sahip olan *div* nesnesini seçecektir. Böylece sadece en düşük fiyat elde edilmiş olacaktır. Örneğin en düşük fiyatı veren acentenin adı da alınmak istensin, dikkat edilirse acentenin logosunun *alt* özelliği bunu vermektedir. *//div[@class="provider \_\_resizeWatch"]//img/@alt* şeklinde bir sorgu *alt* özelliğinin değerinin yani en iyi fiyatı veren acentenin adının alınmasını sağlayacaktır.

Bir otele ait özellikler otel ait web sayfasında yer almaktadır. Bu özelliklerden bir veri tabanı oluşturulmak istendiğinde algoritmanın önce otelin web sayfasının linkini bulup sonra tek tek bu sayfaları ziyaret ederek özelliğin olup olmamasına göre veri toplaması gerekmektedir. Şekil 2.24’de örnek bir otel web sayfası verilmektedir.

Genel Bilgi	Odalar	Yorumlar	Hakkında	Fotoğraflar	Yakınlarda	Sorular ve Yanıtlar	Gezinlerd
✓ Restoran		✓ Bar/Oturma Salonu		✓ Spa		★★★★★	
✓ Spor Salonu/Spor Odalı Fitness Merkezi		✓ Havuz		✓ Oda servisi		ÖDÜLLER VE SERTİFİKALAR	
✓ Deniz kıyısı		✓ Ücretsiz İnternet		✓ Ücretsiz otopark		🏆 Mükemmellik Sertifikası	
✓ Kahvaltı dahil		✓ Açık havuz		✓ Bebek Bakımı		OTEL TARZI	
✓ Birden Çok Dil Bilen Personel		✓ Concierge		✓ Çamaşır Yıkama Hizmeti		Plaj, Lüks, Spa	
✓ Çocuk Aktiviteleri (Çocuğa/Aileye Uygun)		✓ Havaalanı ulaşım hizmeti		✓ İnternet Bağlantılı İş Amaçlı Bölüm		ODA TÜRLERİ	
✓ Kahvaltı Mevcut		✓ Kapalı havuz		✓ Klima		Sigara içilmeyen odalar, Engelliler için uygun odalar, Stütlar, Aile Odaları	
✓ Konferans Salonu		✓ Kapalı havuz		✓ Mini bar		ODA SAYISI	
✓ Tekerlekli sandalye		✓ Toplantı odaları		✓ Ziyafet Salonu		634	
⊗ Ücretsiz Yüksek Hızda İnternet (WiFi)						FİYAT ARALIĞI	
						₺464 - ₺2.343 (Standart Oda için Ortalama Fiyatlar temel alınmıştır)	

**Şekil 2.24.** Otel Web Sayfası ve Otelin Çeşitli Özellikleri

Şekil incelendiğinde otelde olan olanak veya donanımın karşısında ✓ işareti olmayan özelliğin karşısında ise ⊗ işareti yer almaktadır. Bu incelendiğinde olmayan özelliğin *class* özelliğinin *textitem unavailable* olduğu görülmektedir.

```
//div[@class="entry ui_column is-4-tablet is-6-mobile is-4-desktop"]/div[@class="textitem unavailable"]
```

Şeklinde bir sorgu sadece bulunmayan özelliği listeleyecektir. Şekil 2.24’deki otel için bu komut işletildiğinde sadece Ücretsiz Yüksek Hızda İnternet (WiFi) listelenecektir.

```
//div[@class="entry ui_column is-4-tablet is-6-mobile is-4-desktop"]/div[@class="textitem"]
```

Şeklinde bir komut işletildiğinde ise sadece otelin sahip olduğu özellikler elde edilecektir.

XPath’de mantıksal sorgulamalar da yapmak mümkündür. Tablo 2.5’de XPath operatörleri ve işlevleri verilmektedir (XPath Operators, 2019).



**Tablo 2.5.** Xpath Operatörleri ve İşlevleri

<i>İfade</i>	<i>Açıklama</i>
	İki düğümü seçmek için kullanılmaktadır.
+	Toplama işlemi.
-	Çıkarma işlemi.
*	Çarpma işlemi.
div	Bölme İşlemi.
=	Eşittir.
!=	Eşit değildir.
<	Küçüktür.
<=	Küçük ve eşittir.
>	Büyüktür.
>=	Büyük ve eşittir.
or	Veya
and	Ve
mod	Kalanı verir.

Birden çok değer alınmak istendiğinde | operatörü kullanılmaktadır. Örneğin **Şekil** 'de verilen kaynak kod içerisinde *fiyat* ve *fiyatı veren acente* seçilmek isteniyor olsun.

```

▼ <div class="price-wrap">
  <div class="price __resizeWatch" data-sizegroup="mini-
  meta-price" data-clickpart="chevron_price" data-index="0"
  style">₺198</div>
</div>
▼ <div class="provider __resizeWatch" data-sizegroup="mini-
meta-provider" data-clickpart="chevron_provider" data-index=
"0" xpath="1" style">
  

```

**Şekil 2.25.** Çoklu Düğüm Seçimi Operatörü Kullanımı HTML/CSS Kodları

Şekil 2.25'deki kaynak kodlarda *fiyat* bilgisinin *price-wrap* class etiketine sahip *div* nesnesi altındaki *price \_\_resizeWatch* etiketine sahip başka bir *div* nesnesinin altında olduğu görülmektedir. Burada *price \_\_resizeWatch* class etiketine sahip nesne aratılabilir fakat farklı bir kullanım yöntemi göstermek için `//div[@class="price-wrap"]/div[1]` şeklinde bir komutla *fiyat* bilgisine ulaşılmıştır. Bu komut *price-wrap* div nesnesi

altındaki *birinci* div nesnesini seçmektedir. Köşeli parantezin içindeki rakam kaçınıcı nesne olduğunu ifade etmektedir.

```
//div[@class="provider __resizeWatch"]//div[@class="price-wrap"]/div[1]
```

Aynı sorgu içerisinde birden fazla nesneyi seçmek için | operatörü kullanılmaktadır. Yukarıda verilen örnekte, kaynak kod içerisinde önce *fiyatı veren acente* sonra *fiyat* bilgisi elde edilmektedir.

XPath'deki aritmetik operatörlerin kullanımları diğer programlama dillerine benzemektedir. Örneğin `//div[@class="price-wrap"]/div[1]<500` şeklindeki bir sorgu 500'den küçük olan değerleri verirken `//div[@class="price-wrap"]/div[1]*0.82` sorgu sonucundaki değeri 0,82 ile çarpılmaktadır. Fakat kaynak kodlardaki değerler ön ek olarak TL simgesine sahip olduğu için sonuç çıkmayacaktır. Bu tür durumlar için XPath metin fonksiyonlarına sahiptir. Aşağıdaki tabloda XPath'de kullanılan metin fonksiyonları ve işlevleri verilmektedir (MDN web docs: XPath, 2018; Using XPath Functions, 2013).

**Tablo 2.6.** Xpath Metin Fonksiyonları ve İşlevleri

<i>İfade</i>	<i>Açıklama</i>
string()	Sayıyı metine dönüştürür.
concat(string, ...)	Metinleri birleştirir.
starts-with(string, string)	Bir metin belirtilen metinle mi başlıyor?
contains(string, string)	Bir metin belirtilen bir metni içeriyor mu?
substring(string, number, number)	Sadece belirtilen karaktere kadar metni alır.
string-length(string)	Metnin uzunluğunu verir.
substring-after(string, string)	Belirtilen metin ya da karakterden sonrasını alır.
substring-before(string, string)	Belirtilen metin ya da karakterden öncesini alır.
translate(string1, string2, string3)	Bir metin içerisinde aratılan metni başka bir metin ile değiştirir.
normalize-space(string)	Kelimeler arasında bir boşluk olacak şekilde düzeltir.

Yukarıda verilen fonksiyonlarla ilgili örneklerin verilmesi kullanımları açısından daha açıklayıcı olacaktır.

Şekil 2.25'deki kaynak kodlar incelendiğinde *fiyat* değerinin önünde TL simgesi olduğu görülmektedir. Bu durum sonrasında bir sayı var olsa da yorumlayıcının (interpreter) bunu metin (string) olarak algılamasına neden olacaktır. TL simgesinden oluşan bu ilk karakteri silmek için `substring(//div[@class="price __resizeWatch"], 2)` komutu yeterli olacaktır. Burada `//div[@class="price __resizeWatch"]` sorgusu ile

döndürülen ₺657 değeri 2'nci karakterden itibaren seçilmeye başlandığı için 657 olarak alınacaktır. Tam tersi bir işlem yapmak için `string(substring(//div[@class="price__resizeWatch"], 2))` komutu işletilmelidir. Burada nümerik bir değer olan 657 sayısı, metin haline dönüştürülmektedir.

Yine bir önceki örnek ile devam edildiğinde `starts-with(//div[@class="price__resizeWatch"], "₺")` komutu döndürülen değer TL simgesi ile başlayıp başlamadığını sorgulayacak ve başladığı için de `1 (true)` değeri olarak döndürülecektir. `substring-after` ve `substring-before` kullanımı ile ilgili örnek için Şekil 2.26 Şekil 'deki kaynak kodlar verilmektedir.

```
▶ <a class="ui_bubble_rating bubble_45" alt="5 baloncuk
üzerinden&nbsp;4,5" href="/Hotel_Review-g293974-d1776549-Reviews-
Crowne_Plaza_Istanbul_Asia-Istanbul.html#REVIEWS" target="_blank"
data-clicksource="BubbleRating" onclick="return false;">...</a>
<a class="review_count" href="/Hotel_Review-g293974-d1776549-
Reviews-Crowne_Plaza_Istanbul_Asia-Istanbul.html#REVIEWS"
target="_blank" onclick="return false;" data-clicksource=
"ReviewCount">493 yorum</a>
```

**Şekil 2.26.** Substring-after ve Substring-before Kullanımı HTML/CSS Kodları

Şekilde `yorum sayisi` bir `a` nesnesi içerisinde verilmektedir. `//a[@class="review_count"]` şeklindeki bir sorgu `493 yorum` çıktısını verecektir. Burada sadece sayı kısmı alınmak istendiğinde `substring-before(//a[@class="review_count"], " ")` şeklinde bir sorgu işletmek gereklidir. Bu sorgu ile “ “ içerisinde yer alan boşluk karakterinden önceki metin yani `493` sayısı alınacaktır. `substring-after` tam tersi sonuç söz konusudur. Aynı kod ile bu komut işletildiğinde `yorum` şeklinde bir çıktı alınacaktır.

Tablo 2.6'daki fonksiyonlar içerisinde en çok kullanılanlardan biri `normalize-space(string)` fonksiyonudur. Bazı web sayfaları metin içerisinde gereksiz boşluk karakteri içermektedir. Normal olması gereken kelimeler arasında bir boşluk ve noktalama işaretlerinden sonra bir boşluktur. Böyle olmayan durumlarda `normalize-space(string)` fonksiyonu işletilerek içeriğindeki metin olması gereken haline getirilmektedir. Örneğin `normalize-space("4.118 işletmeden 173 tanesi")` şeklindeki bir metin fonksiyon çalıştırıldıktan sonra `"4.118 işletmeden 173 tanesi"` haline getirilecektir.

Bir metin içerisinde aratılan bir metnin başka bir metin ile değiştirilmesi işlemi *translate(string1, string2, string3)* fonksiyonu ile yapılmaktadır. Örnek sitede binlik ayracı olarak nokta (".") kullanılmaktadır. Bazı durumlarda bu istenmeyebilir. Örneğin *translate(//div[@class="price \_\_resizeWatch"],".", "")* şeklindeki komut *price \_\_resizeWatch* class etiketli metin içerisinde noktaları bulup, temizleyecektir. Böylece normalde *1.402* şeklindeki bir sayı *1402* şekline dönüştürülecektir.

Tablo 2.7’de XPath’de kullanılan bazı nümerik fonksiyonlar ve işlevleri verilmektedir.

**Tablo 2.7.** Xpath Nümerik Fonksiyonları ve İşlevleri

<i>İfade</i>	<i>Açıklama</i>
<i>count()</i>	Kaç tane nesne seçildiyse sayısını verir.
<i>sum()</i>	Seçili sayıların toplamlarını verir.
<i>avg()</i>	Seçili sayıların ortalamasını verir.
<i>ceiling()</i>	En küçük tamsayıya yukarı doğru yuvarlar.
<i>floor()</i>	En küçük tamsayıya aşağı doğru yuvarlar.
<i>round()</i>	Bir sayının ondalık olmayan kısmını alır.

Web madenciliği yapılmak istenen hedef her bir web sayfasında 30 adet otele ait isim, fiyat, yorum sayısı, puan gibi bazı bilgiler yer almaktadır. Daha önce otel adını seçmek için kullanılan XPath sorgusu *count(//div[@class="listing\_title"])* şeklinde *count()* fonksiyonu içerisine alındığında sonuç 30 olarak döndürülecektir. Diğer fonksiyonların kullanımı da benzer şekildedir. 19,83 değeri *ceiling()* ile 20, *floor()* ile 19, *round()* ile yine 19 olarak döndürülecektir.

XPath web madenciliği için önemli araçlar sağlamaktadır. Daha önce web madenciliği örnekleri verilmiş olan Python (Reitz, 2019) ve R’da (Sanchez, 2014) XPath kullanmak mümkündür. Python’da bunun için *lxml* kütüphanesi kullanılırken, R’da *XML* kütüphanesi kullanılmaktadır. XPath kullanımı her iki programlama dilinde de web madenciliği açısından zaten güçlü olan bu dillere daha da işlevsellik kazandırmaktadır. Yine web madenciliği örneği verilen Knime’da XPath bu yöntemin önemli bir parçasıdır.

Bu yöntemlerin kullanılması, elde edilmesi zor ve maliyetli olan veri elde etme yöntemlerine devrimsel bir çözüm üretmektedir. Web olarak adlandırılan devasa veri madeninde bilimsel araştırmalar için çok önemli verileri bulmak gelişen teknoloji ve bu tarz yöntemler ile geçmiştekenden çok daha kolaydır.

Bütün bu yöntemlerin webden veri çıkarımında bulunması webin ilk yıllarından bu güne kadar yerini korumuş bir yapıya dayanmaktadır. Bu webin mucidi Tim Berners-Lee'nin geliştirdiği HTML dilidir. Daha sonra CSS gibi yapılar da buna dâhil olsa da HTML etiketleri web madenciliğinin temelini oluşturmaktadır. İster bir yazılım isterse bir programlama dili aracılığıyla web madenciliği yapılıyor olsun, HTML ve CSS etiketlerini (tags) kullanmak zorundadır.

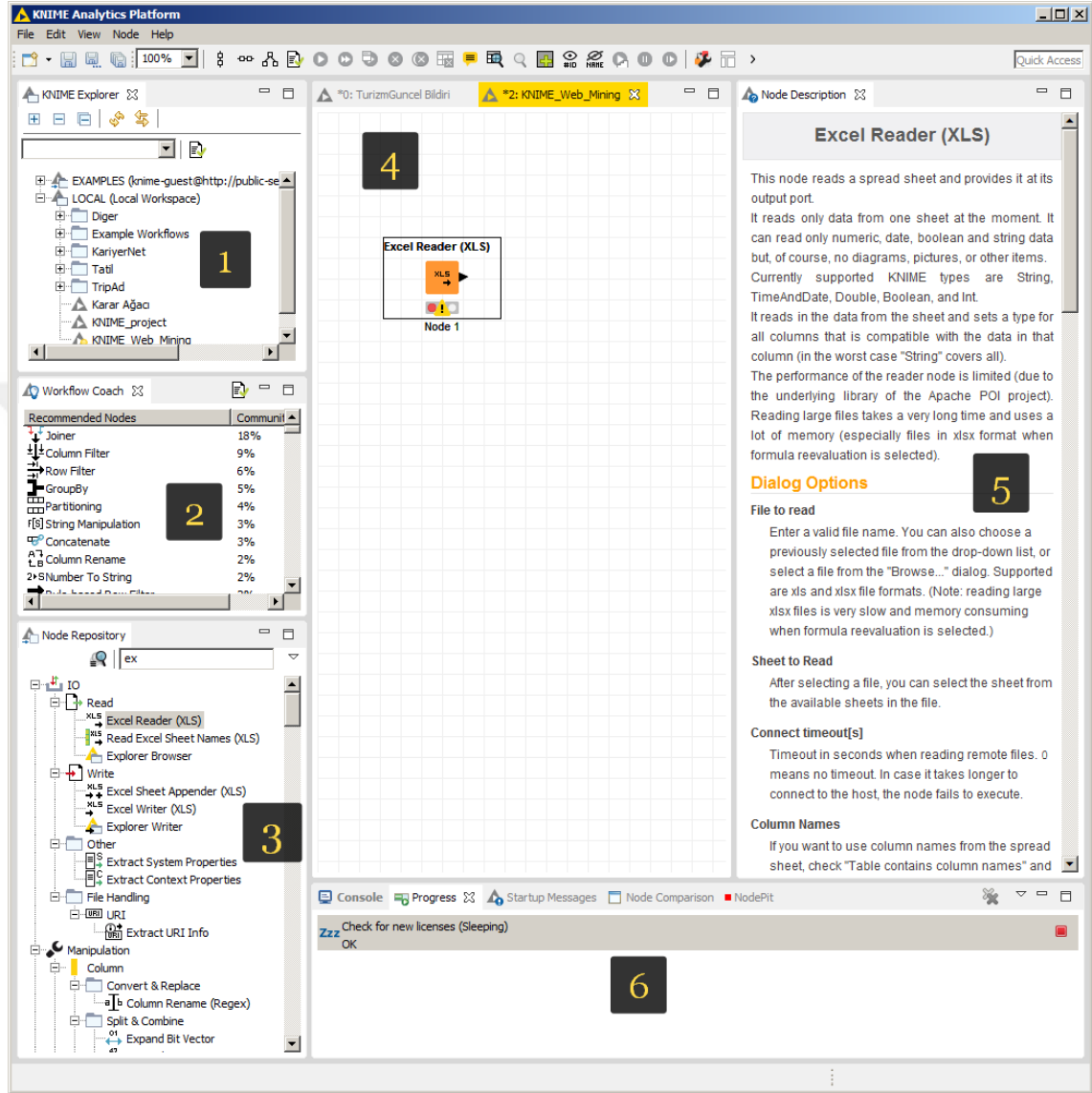
#### **2.4.3.4. Knime İle Web Madenciliği**

Programlama dilleri dışında sadece web madenciliği için hazırlanmış veya içerisinde web madenciliği araçları barındıran yazılımlar da bulunmaktadır. Parsehub, Mozenda ve Octoparse gibi yazılımlar özellikle web madenciliği için hazırlanmış yazılımlarken, Knime içerdiği bazı araçlarla web madenciliği için de kullanılmaktadır.

Knime Analytics Platform hızlı, basit ve verimli bir şekilde veri bilimi yapma üzerine tasarlanmış bir yazılımdır. KNIME açık kaynak kodlu bir yazılımdır, sadece ticari amaçlarla kullanıldığında yıllık lisans ücreti talep edilmektedir. KNIME geleceği tahmin etmek, yeni fikirler için madencilik yapmak veri içerisindeki gizli potansiyeli keşfetmek için tasarlanmıştır (www.knime.com, 2019).

KNIME *node (düğüm)* olarak adlandırılan işlem birimlerinin akış alanında (workflow) birbirlerine bağlanması ile çalışan bir platformdur. Bir düğüme gelen veri işlenerek başka bir düğüme girdi olmaktadır.

KNIME bir ana pencere ve ana pencerenin altında çeşitli amaçlar için tasarlanmış alt pencerelerden oluşmaktadır. Şekil 2.27’de KNIME yazılımının görünümü ve çalışma alanı verilmektedir.



Şekil 2.27. KNIME Analytics Platform Çalışma Alanı

Şekilde verilen ekran görüntüsü incelendiğinde KNIME iş akış alanı [4] pencerenin ortasında yer almaktadır. Node (düğüm) olarak adlandırılan işlem birimleri bu alana taşınarak kullanılmaktadır. *Node Repository* [3] penceresinde çeşitli başlıklar altında verilen düğümler yer almaktadır. İstendiği takdirde buradaki arama kutusu düğüm bulmak için kullanılmaktadır. Bu kısımda *IO*, *Manipulation*, *Views*, *Analytics*, *Database* gibi üst başlıklar yer almaktadır. Bunlar açıldığında çeşitli amaçlarla kullanılacak düğümler bulunmaktadır.

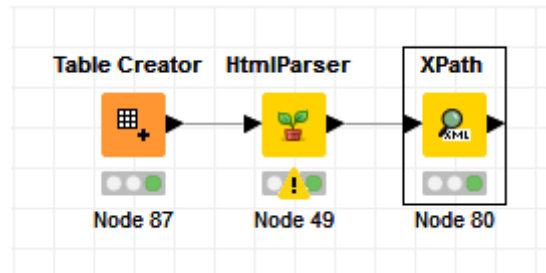
*Node Repository* penceresinde yer alan düğümler iş akış alanına eklendiğinde düğüm seçiliyken *Node Description* penceresi aktif hale gelmektedir [5]. Bu pencere düğüm kullanım bilgisi, giriş-çıkış portları varsa ek bilgi hakkında bilgi içermektedir.

*Workflow Coach* penceresi [2] eklenen bir düğümden sonra gelebilecek düğümleri önermektedir. Bu öneriler program kullanıcılarından toplanan verilerden derlenmektedir. Şekil 2.27'deki ekran görüntüsünde bir *Excel Reader* düğümünden sonra %18 oranında bir *Joiner* düğümü eklendiği önerisinde bulunmaktadır.

*Knime Explorer* penceresi [1] KNIME'da daha önce kayıt edilmiş projeleri göstermektedir. Bu kısımda iki ana dizin bulunmaktadır. LOCAL dizini adından da anlaşılacağı üzere KNIME'in yüklü olduğu bilgisayarda yer alan projeleri göstermektedir. EXAMPLES dizini ise, KNIME sunucularına bağlanarak örnek proje dosyaları indirmeye olanak vermektedir. Bu sunuculara *üye* olarak veya *misafir* olarak bağlanılabilir.

*Progress* [6] penceresi programın o an yaptığı işlemleri ve bu işlemlerin ilerleme durumunu göstermektedir. Programın işleyişinde meydana gelen herhangi bir hata bu pencerede gösterilmektedir.

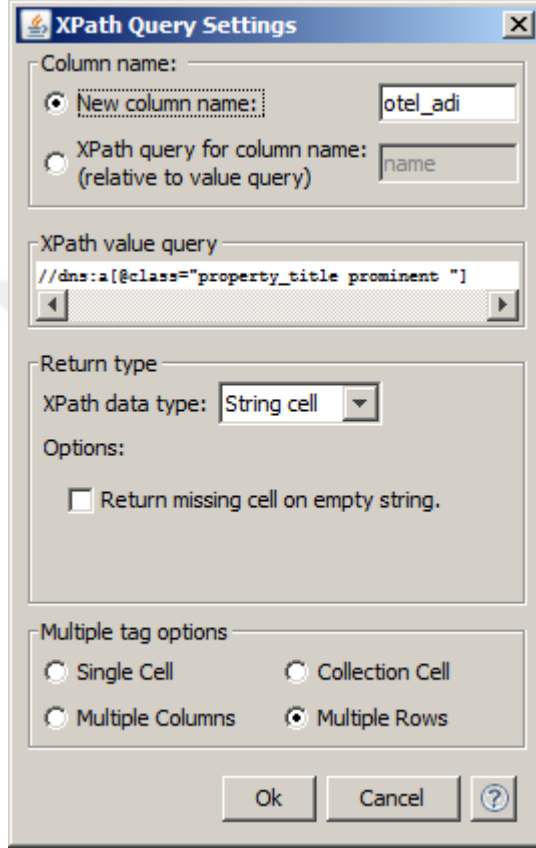
Knime'da web sayfalarından içerik almak için *HtmlParser* modülü kullanılmaktadır. Fakat daha öncede bahsedildiği gibi her modülün farklı bir işlevi bulunmaktadır. *HtmlParser* ile sayfanın kaynak kodları hafızaya alınmaktadır. Bu kaynak kodlar içerisinde istenilen verinin alınabilmesi için *XPath* modülü kullanılmaktadır. Şekil 2.28'de Knime'da basit bir web madenciliği örneği verilmektedir.



Şekil 2.28. Knime'da Web Madenciliği Modülleri

Örnek incelendiğinde *Table Creator* modülü içerisinde yazılmış olan hedef web sayfası *HtmlParser* modülü tarafından ziyaret edilerek kaynak kodları elde edilmektedir.

Kaynak kodlar içerisinde sadece istenilen verinin elde edilmesi için ise, XPath modülü kullanılmaktadır. XPath, XML dokümanları içerisinde sorgulama/arama yapmak için geliştirilmiş bir dildir. XPath, günümüzde JavaScript, Java, XML Schema, PHP, Python, C ve C++ ve daha birçok dilin kaynak kodları içinde sorgulama yapmak için kullanılmaktadır (www.w3schools.com, 2019).



Şekil 2.29. XPath Sorgu Ekleme Ekranı

XPath modülü seçiliyken F6 klavye kısayol tuşu ile XPath ayarlarının yapıldığı diyalog penceresi ekrana gelmektedir. Buradan *Add XPath* tuşu kullanılarak yeni sorgu eklenmektedir. Şekil 2.29'da görüleceği üzere *XPath value query* kısmına kaynak kodlar içerisinde alınmak istenen nesne için XPath sorgusu girilmektedir. *Multiple Rows*'un seçilmesi verinin tek bir hücre olmadığı birden fazla satır olduğu anlamına gelmektedir.

XPath sorgusu, *property\_title prominent* sınıfında *a* yani link nesnelere aramaktadır. Burada herhangi ek bir sorgu girilmediğinden varsayılan olarak o nesne içerisindeki *metin* elde edilecektir. Bu *metin* ise otel adını içermektedir. Sorgunun

devamında */@href* şeklinde bir istek olsaydı, o nesnenin *href* özelliği yani linki alınacaktı.



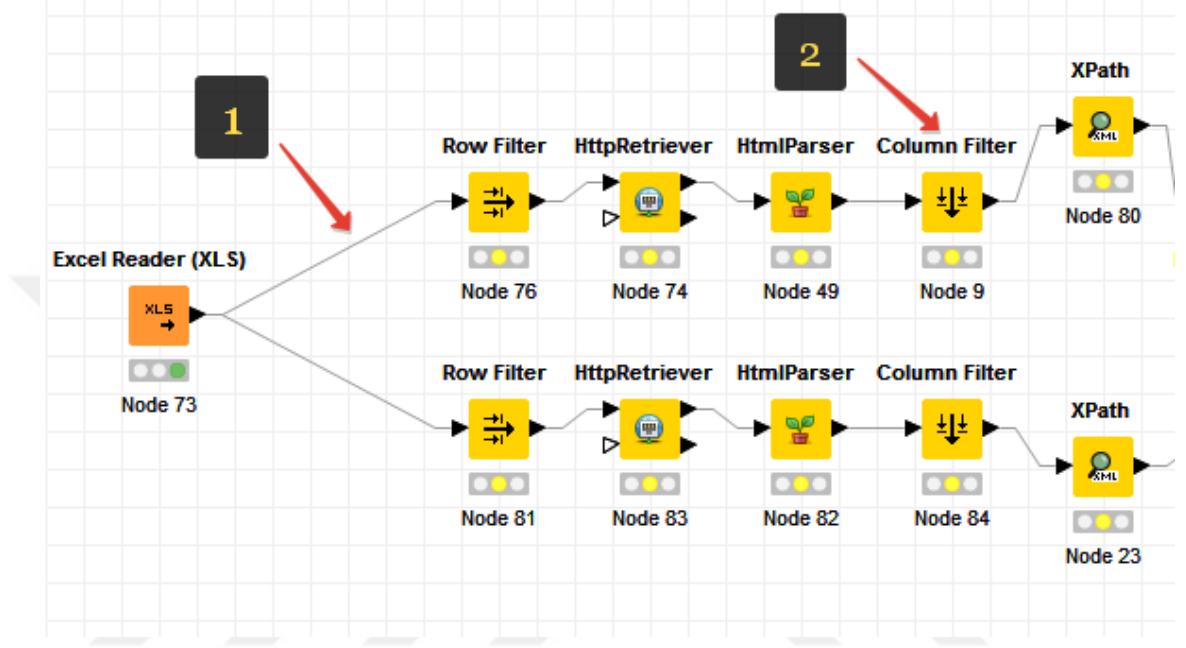
Sorgular sonucunda elde edilen çıktı Şekil 2.30’da verilmektedir.

S	adres	S	otel_adi	S	fiyat	S	yorum_...	S	link
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		FLORYA HOUSE HOTEL		€198		65 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Basileus Otel		€344		2.009 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Radisson Blu Hotel & ...		€388		671 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Aydinli Cave Hotel		€433		1.372 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Sura Hagia Sophia Hotel		€856		2.191 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Idylle Hotel		€588		172 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Crowne Plaza Hotel A...		€316		2.667 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Akra Otel		€489		1.436 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Hilton İstanbul Bosph...		€948		3.071 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		The Marmara Taksim		€709		1.532 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Kayakapi Premium Ca...		€802		1.532 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Hotel Amira İstanbul		€710		3.754 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Cappadocia Cave Sui...		€502		1.086 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Radisson Blu Hotel Is...		€511		359 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Pera Palace Hotel		€898		1.443 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		Mövenpick Hotel İsta...		€511		1.362 yorum		https://www.tripadvisor.com.tr/Hotel_Rev
	https://www.tripadvisor.com.tr/Hotels-g293969-Turkey-Hotels.html		White House Hotel Is...		€859		4.211 yorum		https://www.tripadvisor.com.tr/Hotel_Rev

Şekil 2.30. Knime İle Elde Edilen Verinin Çıktısı

Xpath ile ilgili detaylı bilgi Web Madenciliğinde XPath Kullanımı bölümünde detaylı olarak verilmektedir. Sorgu sonucu elde edilen çıktı çoğu zaman bu şekilde düzenli olmamaktadır. Output’da istenmeyen karakterler, kod artıkları olabilmektedir. Knime’da *Math Formula* gibi düğümler bu fazlalıkları temizlemek için kullanılmaktadır. Her düğümün kullanımı daha önce de bahsedildiği gibi *Node Description* penceresinde verilmektedir.

Knime'daki düğümler (araçlar) kullanım alanlarına göre kategoriler halinde *Node Repository* penceresinde verilmektedir. Bu düğümler çalışma alanına sürüklenerek kullanılmaktadır. Bir düğüm tek başına kullanılmamaktadır. Bir düğümdeki işlem sonucu elde edilen çıktı (output) bir *ok* yardımı ile diğer düğüme bağlanmalıdır. Aşağıdaki **Şekil** şekilde bununla ilgili örnek verilmektedir.



**Şekil 2.31.** Knime'da Düğümlerin Kullanımı

Şekil incelendiğinde *Excel Reader* düğümü ile bir *excel* dosyasından okunan veri *Row Filter* düğümüne [1] girdi olmaktadır. Çıktı, düğümün özelliğine göre bir veya birden çok düğüme girdi olabilmektedir. Şekilde *Excel Reader* çıktısı *Node 76* ve *Node 81* adında iki düğüme girdi olmaktadır. Adından da anlaşılacağı üzere *Row Filter* düğümleri satırları filtrelemek için kullanılmaktadır. Her düğümün altında o düğümün adı yer almaktadır. İstenirse bu ad değiştirilebilir. Veri üzerinde uygulanan filtreler de düğüm mantığı ile çalışmaktadır. Satır filtresi uygulandığı gibi gerekli görülmeyen sütunlar için de *Column Filter* kullanılmaktadır.

Knime'da çeşitli amaçlarla kullanılan yüzlerce düğüm bulunmaktadır. Bunlar kategoriler halinde *Node Repository* penceresinde listelenmektedir. Bu kategoriler ve altında bulunan düğümler şu şekildedir.

## IO

Bu kısımda veriyi dosyadan okumak (Input) veya dosyaya yazmak (Output) için kullanılan düğümler bulunmaktadır. Yaygın kullanılan birçok dosya türü Knime tarafından desteklenmektedir. CSV, ARFF, XLS, PMML, Image okumak veya yazmak için desteklenen dosya formatlarından bazılarıdır.

### *Manipulation*

Bu kısımda tablo halinde olan veya bu hale dönüştürülen veri üzerinde çeşitli dönüştürme işlemleri için kullanılabilir düğümler bulunmaktadır. Satır, sütunlar ve tablolar üzerinde uygulanabilecek bu işlemler oldukça zengindir. Örneğin bir satırdaki metinsel (string) değerler sayısal (number) değerlere dönüştürülecekse, kullanılacak düğüm bu kısımda bulunmaktadır. Ayrıca daha önce de bahsedilen satır, sütun filtreleri de bu kategori altında bulunmaktadır.

### *Views*

Veriyi görselleştirmek için kullanılan araçlar bu kısımda bulunmaktadır. Örneğin bir *karar ağacı* algoritmasının görselleştirilen sonucunu görmek için bu kısımdaki ilgili düğümler kullanılmaktadır. Ayrıca tanımlayıcı istatistikler için *histogram*, *çubuk*, *pasta grafik* gibi grafik araçları bu kısımdadır.

### *Analytics*

Bu kısımda veri üzerinde analiz yapmak için kullanılan düğümler bulunmaktadır. Bu düğümler *madencilik (mining)*, *istatistik* ve *mesafe hesaplama (distance calculation)* olarak üç ayrı kategori altında yer almaktadır. Veri madenciliği için kullanılabilir bayes, kümeleme, karar ağaçları, lojistik regresyon, yapay sinir ağları gibi algoritmaların yer aldığı zengin bir kütüphane yer almaktadır.

### *Database*

Knime veri tabanlarına bağlanabilen ve bu veri tabanları ile veri alışverişi yapabilen bir yazılımdır. MySQL, MSSQL, SQLite gibi veri tabanları desteklenmektedir. Bu veri tabanları üzerinde çeşitli sorgular da yapılabilmektedir.

### *Other Data Types*

Bu kısımda metin madenciliği, metin işleme, zaman serileri ve *web log* işlemlerinde kullanılabilir düğümler bulunmaktadır. *Doğal Dil İşleme (NLP)* işlemleri bu araçlar ile kolaylıkla yapılabilmektedir.

### *Structured Data*

Web dosya kayıt yapılarından olan JSON ve XML işlemleri ile ilgili düğümler bu kısımda bulunmaktadır. Buradaki araçlar ile JSON ve XML ile ilgili okuma, yazma, dönüştürme ve manipülasyon işlemleri yapılmaktadır.

### *Scripting*

Knime R ve Java kodlarını da desteklemektedir. Bu kısımda bulunan düğümler ile bu dillere ait kaynak kodlar kullanılabilir. Ayrıca kurulan bazı eklentiler ile Weka kütüphaneleri de kullanılmaktadır.

### *Community Nodes*

Palladian ve Vernalis gibi üçüncü parti uygulamalar bu kısımda yer almaktadır. Özellikle Palladian web madenciliği açısından oldukça faydalı araçlar içermektedir.

Node Repository kısmında bunlar dışında birçok düğüm bulunmaktadır. Ayrıca *Help/Install New Software* komutu ile Knime veya diğer geliştiriciler tarafından yayınlanan birçok eklenti de yüklenebilmektedir.

## **2.5. LİTERATÜR ÖZETİ**

Bu kısmı iki farklı şekilde değerlendirmek gerekmektedir. Yazında otel oda fiyatlarını etkileyen faktörlerin belirlenmesinde web madenciliği ile elde edilen verinin kullanılmasına dair bir araştırmaya rastlanmamıştır. Bu nedenle otel oda fiyatlarını etkileyen faktörlerin belirlenmesine yönelik çalışmalar web madenciliğinden bağımsız olarak verilecektir. Web madenciliği ile veri elde etme ve bu verinin bilimsel araştırmalarda kullanımı ile ilgili çalışmalar ayrı olarak değerlendirilecektir.

Öncelikle Türkiye’de otel oda fiyatlarını etkileyen faktörleri belirlemek için yapılan araştırmaları belirlemek için yapılan literatür taramasında, bu araştırmaların *Hedonik Fiyatlandırma* olarak adlandırılan ekonometrik bir yöntemle yoğunlaştığı görülmektedir.

*Hedonik Fiyatlandırma*; bir malın fiyatını malın sahip olduğu çeşitli özellikler ile tahmin etmeye çalışan bir regresyon modelidir (Xiao, 2016: 12-13). Hedonik fiyatlandırma modeli, bir malın fiyatının aslında o malın sahip olduğu özelliklerinin değerinin toplamına eşit olduğunu varsayar (Rosen, 1974). Bu model, regresyon analizi kullanarak malın fiyatı üzerinde etkisi olan, mala ait değişkenlerin, toplam fiyata etkisini bulmayı amaçlar (Monson, 2009; Löch ve Axhausen, 2010; Yayar ve Gül, 2014; Andersson, 2008: 231; Castro, Ferreira ve Ferreira, 2016). Örneğin bir otel işletmesi ürün olarak düşünüldüğünde; bu otelin plajının, spa, spor salonunun olması, havaalanına yakınlık, manzarası gibi pek çok özellik alıcı tarafından bu otelin tercih nedeni olabilir. Regresyon analizi ile bir model oluşturularak buradaki her bir özelliğin otelin toplam oda fiyatı üzerindeki etkisi ortaya konulabilir. Hedonik fiyatlandırma modeli, fiyat için doğru tahmin modeli oluşturmak için (Monson, 2009: 64) doğrusal model, tam logaritmik model, doğrusal logaritmik model, logaritmik doğrusal model (Kaya, 2012: 54-58), gibi regresyon temelli yöntemler kullanır.

Hedonik fiyat varsayımından hareketle, Kesbiç, Bahar, Baldemir ve İnci (2011) Muğla ili deniz ve kıyı turizminde faaliyet gösteren 4 yıldızlı 147 otel üzerinde yaptıkları çalışmada, otel oda fiyatları üzerinde hangi faktörlerin anlamlı olduğunu araştırmışlardır. Modelde kullanılan 21 değişken içerisinde fiyat üzerinde en çok etkisi olan değişkenler, *otelin yarım pansiyon, oda ve kahvaltı hizmeti vermesi ve self-servis hizmet vermesi* olarak belirlenmiştir. Bu çalışmada otellere ilişkin paket tur fiyatları ve otellere ait özellikler ilgili yerdeki turizm acenteleri kataloglarından elde edilmiştir. Pazarı göre farklı kataloglar basıldığından İngiliz pazarına ilişkin kataloglar tercih edilmiştir.

Kesbiç, Bahar, Baldemir ve İnci'nin (2011) çalışmalarına benzer bir çalışma Işık ve Bilici (2017) tarafından Erzurum ilinde faaliyet gösteren oteller üzerinde yapılmıştır. Işık ve Bilici (2017) Erzurum kent merkezinde bulunan çeşitli özelliklerde 20 adet otele ait veri üzerinde yine Hedonik fiyatlandırma analizi yapmıştır. Modelde 26 değişken kullanılmıştır. Bu değişkenlerden 15'nin oda fiyatı üzerinde pozitif etkisinin 2 tanesinin ise negatif etkisinin olduğu belirlenmiştir. Pozitif etkisi olan değişkenler, otelin zincir otel olması, konferans salonu olması, otel odasında banyo olması, oda fiyatına kahvaltı dâhil olması, oda büyüklüğü, bar olması, restoran olması, gece kulübü olması, fitness olması, odada LCD TV olması, Wi-Fi olması, Fotokopi/faks olması, shuttle olmasıdır. Otelin

kayak imkânının olmaması ve otelin konumunun merkeze uzak olması ise fiyat üzerinde negatif etkiye sahiptir. Otellere ait bilgiler ve fiyatlar, booking.com'dan alınmıştır.

Hem Kesbiç, Bahar, Baldemir ve İnci'nin (2011) çalışması, hem de Işık ve Bilici'nin (2017) çalışması bölgesel veriler kullanılarak yapılmıştır. Kesbiç, Bahar, Baldemir ve İnci (2011) çalışmasında bir turizm sezonu boyunca (Mayıs-Eylül 2008) Muğla'nın ilçelerinde bulunan turistik otellere ait veriyi seyahat acenteleri kataloglarından derlemiştir. Referans fiyatlar ise 1-23 Temmuz'da sezonun yüksek fiyatlı döneminden alınmıştır. Işık ve Bilici (2017) ise bir ÇSA olan booking.com'dan derledikleri verileri kullanarak Erzurum bölgesi otelleri üzerinde analizler yapmışlardır.

Yurtdışında yapılmış fakat Türkiye'deki otelleri de analizlerine dâhil eden çalışmalar da bulunmaktadır. Bunlardan Papatheodorou (2002) çalışmasına, Akdeniz çanağındaki 10 ülkeyi dâhil etmiş, örneklem içerisinde Muğla bölgesinden oteller yer almıştır. Papatheodorou (2002) fiyat modeli kurarken otelin bulunduğu bölgeye olan uçuşları, her şey dâhil (all inclusive) sisteminin olup olmamasını, yıldız sayısını, zincir otel olup olmamasını faktör olarak kullanmıştır. Papatheodorou da (2002) fiyatları ve otel özelliklerini seyahat acentesi kataloglarından almıştır.

Turizm sektöründe ürün özelliklerinin toplam fiyat ile ilişkisi turizm pazarlaması içerisinde vurgulanmaktadır (Abrate, Capriello ve Fraquelli, 2011: 913). Bu turistik ürünün bileşik ürün olma özelliği, otel gibi sadece tek bir ürün gibi görünen ürünlerde de düşünülebilir. Otel ya da konaklama ürünü de aslında birçok ürün ve hizmetten meydana gelmektedir.

Uluslararası yazında turistik ürünü oluşturan ürünlerin toplam fiyat içerisinde etkisini araştıran veya bir bütün olarak düşünüp dış faktörlerin fiyat üzerindeki etkisini araştıran pek çok çalışma yapılmıştır (Carvell ve Herrin, 1990; Bull, 1994; Taylor, 1995; Papatheodorou, 2002; Aguiló, Alegre ve Sard, 2003; Espinet, Saez, Coenders ve Fluvia, 2003; Thrane, 2005; Thrane, 2007; Hung, Shang ve Wang, 2010; Zhang, Zhang, Lu, Cheng ve Zhang, 2011; Pawlicz ve Napierala, 2017).

Bu çalışmaların bazıları oda fiyatının tek bir özellik ile ilişkisi üzerinde dururken (Bull, 1994; Israeli, 2002), diğerleri birçok faktör içerisinde hangilerinin fiyat üzerinde etkisi olduğunu belirlemeyi amaçlamıştır.

Web madenciliği kullanılarak yapılan arařtırmaları son yıllarda farklı pek çok farklı alanda görmek mümkündür. Alanları ne olursa olsun, webin sağladığı zengin kaynağı keşfetmiş arařtırmacılar veri elde etmede web madenciliğini kullanmışlardır. Gyódi (2019) çalışmasında, web üzerinde turistlere konut yada oda kiralama servisi olan Airbnb üzerinde web madenciliği yapmış, Airbnb servisinin geleneksel otel endüstrisi ve yerel halk üzerindeki potansiyel etki bağlamında yapısını incelemektedir. Arařtırmada Paris, Barselona, Berlin ve Varşova otelleri ve Airbnb'nin aynı bölgelerdeki ev ve odaları karşılaştırılmıştır. Verinin elde edilmesinde Python programlama dili ve Selenium Web Sürücüsünden faydalanılmıştır.

Egara, Kerami, Wiryana ve Maulana Kusuma, (2017) bir akademik sosyal ağ olan researchgate.net'i web madenciliği yöntemi ile analiz etmiş, Endonezya'daki akademisyenlerin ne oranda ulusal veya uluslararası yayın yapıp yapmadıklarını değerlendirmişlerdir.

Başka bir sosyal medya çalışmasında Akkineni, Lakshmi ve Vijaya Babu (2016), hükümet politikalarıyla ilgili tweetleri çıkarma metodolojisine değinmektedir. Çalışmada, akış API'si kullanılarak Twitter'dan devlet politikaları ile ilgili verileri çıkarma işlemi yapılmaktadır.

Bakaev ve Avdeenko (2015) işgücü piyasasının yönetimi için karar vermede kullanılacak yapılandırılmamış veriyi web madenciliği yöntemi ile elde etmişlerdir. Rusya'nın Novosibirsk bölgesinde yapılan çalışmada, işgücü piyasasını izlemek için yürütölen bir proje için işle alakalı web sitelerinden veri çekilmiştir.

Petrovic ve Stanisevic (2017) konferans bildirilerinde web üzerinde bulunan ikinci el otomobil pazarına ait verilerin nasıl web madenciliği ile elde edileceğini açıklamaktadırlar.

Polidoro, Giannini, Conte, Mosca ve Rossetti (2015) çalışmalarında web madenciliği tekniğini tüketici fiyat endeksi için (HICP) çeşitli mal ve hizmetlerin fiyatlarını elde etmek için kullanmaktadırlar.

Dikkat edileceği üzere veri toplamada sağlayacağı faydalar yeni yeni fark edilen web madenciliği yöntemi birçok alanda kullanılmaktadır. Web madenciliği klasik veri toplama yöntemlerinden çok daha hızlı ve düşük maliyetlidir. Ayrıca web gibi muazzam büyüklükte bir veri kaynağına sahiptir.

Bu çalışmada, geçmişte fiyat faktörlerini belirlemek amacıyla yapılan araştırmalardan (Carvell ve Herrin, 1990; Bull, 1994; Taylor, 1995; Papatheodorou, 2002; Aguiló, Alegre ve Sard, 2003; Espinet, Saez, Coenders ve Fluvia, 2003; Thrane, 2005; Thrane, 2007; Hung, Shang ve Wang, 2010; Zhang, Zhang, Lu, Cheng ve Zhang, 2011; Pawlicz ve Napierala, 2017) farklı olarak, veri elde etmede kullanılan yöntemin sağladığı avantaj ile neredeyse bir ülkenin bütün otellerinin dâhil olduğu bir veri seti kullanılmıştır. ÇSA'nın web sitesi üzerinde yayınladığı veri tabanı şirketin konaklama tesislerine sağladığı bir giriş paneli üzerinden konaklama tesisi personeli tarafından yapılmaktadır. Bu şekilde düşünüldüğünde yaklaşık 6500 konaklama tesisine ait fiyat, donanım ve hizmetlere ait değişkenleri içeren veri tabanının yayınlanmış hali tam tersi bir işlem yapılarak tekrar veri tabanı haline getirilmektedir.



## ÜÇÜNCÜ BÖLÜM

### ÇEVİRİMİÇİ SEYAHAT ACENTELERİNDE ODA FİYATLARINA ETKİ EDEN FAKTÖRLERİN ARAŞTIRILMASI

#### 3.1. ARAŞTIRMA YÖNTEMİ

Araştırmanın yöntemi veri toplama ve verinin analizi olmak üzere iki kısımdan oluşmaktadır. Birinci kısımda, geliştirilen algoritma yardımı ile otel fiyat verisi ve ilgili değişkenler, Tripadvisor web sitesine bağlanarak bir veri tabanı oluşturulmuştur. Bu sitenin tercih edilmesinin nedeni birçok acenteye ait fiyat verisini üzerinde barındırması, otellere ait özelliklerin detaylı olarak belirtilmesi, robot yazılımı engelleyici bir sistem kullanmamasıdır. Bu tercih araştırma öncesi yapılan bir pilot çalışma ile belirlenmiştir. Ayrıca Tripadvisor dünya genelinde hizmet veren sektörün en büyük ve tanınmış şirketlerinden biridir.

6558 konaklama tesisine ait veriyi elde eden bu algoritmanın her işleyişi 10 saatin üstünde sürdüğünden dolayı haftada bir kez çalıştırılmıştır. Oda fiyatları elde edilirken bir hafta sonraya iki kişi bir gece konaklama olacak şekilde düzenleme yapılmıştır. 6558 konaklama tesisine ait 49 değişken ve fiyat verisi toplam 19 hafta olarak toplanmıştır.

Bu işlemler Knime Analytics Platform v3.6 yazılımı HtmlParser ve XPath düğümleri aracılığıyla yapılmıştır. Veri temizleme, düzenleme ve analizlere hazırlama aşamasında Microsoft Excel v.2013 daha büyük işlem gücü gerektiren uygulamalar için Google Docs Cloud servisi kullanılmıştır.

İkinci kısımda ise elde edilen ve analizler için uygun hale getirilen veriler, uygun yöntemler ile istatistiksel analizlere tabi tutulmuştur. Bu yöntemler uygulanırken, normal dağılım gibi testlerin varsayımları dikkate alınmış ona göre istatistiksel analiz seçilmiştir.

##### 3.1.1. Web Madenciliği ile Verinin Elde Edilmesi

Bu çalışmada web madenciliği için Knime Analytics Platform v3.6 yazılımı kullanılmıştır. Fakat gerek Python gerekse R dili web madenciliği için verimli bir şekilde kullanılabilir. Knime'in tercih edilme sebebi; yazılımın sadece verinin elde edilmesi için değil, temizlenip düzenlenmesi için de araçları içeriyor oluşudur.

Python ve R kullanılarak web madenciliği uygulaması örneği Web Madenciliği Yöntemleri bölümünde verilmektedir. Kullanılan yöntemler arasında büyük farklılıklar bulunmamaktadır. Bu yöntemler daha önce de bahsedilen HTML, CSS, XML, JSON ve AJAX gibi veri yayınlama yöntemlerinin etiketlerini sorgulayarak içerik alma yöntemi ile çalışmaktadır.

Veri elde edilirken Knime yazılımından ve bu yazılımın eklentilerinden olan Palladian araçlarından faydalanılmıştır. Palladian içerisinde bulunan *HtmlParser* düğümü ile hedef web sayfasının kaynak kodları elde edilmiştir. Elde edilen kaynak kodlar *XPath* sorguları kullanılarak sadece gerekli olan verilerin yer aldığı tablolar haline dönüştürülmüştür.

Tripadvisor Türkiye sayfası, *HtmlParser* ve *XPath* düğümleri yardımı ile sadece otel linkleri ve isimlerini alacak şekilde çözümlenmiştir. Elde edilen linklerden bir veri tabanı oluşturulmuştur. Sezon boyunca yapılan web madenciliği işlemi bu kaynak linkler üzerinde yapılmıştır. Böylece sezon boyunca her fiyat alımı aynı veri seti için yapılmıştır.

Konaklama tesislerine ait 6558 adres her hafta kullanılan algoritma ile taranmıştır. Yine önceki işleme benzer şekilde bu sayfalar *HtmlParser* ve *XPath* düğümleri ile çözümlenmiştir. Fakat bu kez değişken sayısı arttığı için daha karmaşık bir algoritma kullanılmış bazı değişkenler için temizleme ve birleştirme işlemi uygulanmıştır. Temizleme ve birleştirme işlemi için *Math Formula* düğümü kullanılmıştır.

Konaklama tesisine ait özelliklerin ve hizmetlerin yer aldığı tablolar özelliğin olması (1) veya olmaması (0) olarak dönüştürme işlemine tabi tutulmuştur. Bu işlem için *One to Many (PMML)* düğümü kullanılmıştır. Bu araç, bir tesis yüzme havuzuna sahipse yüzme havuzu sütununa “1” değerini vermekte, sahip değilse “0” değerini vermektedir. Böylece uygulanacak istatistiksel analizler için veri uygun hale getirilmektedir.

Verinin elde edilmesi, temizlenmesi, dönüştürülmesi işlemlerinden sonra oluşturulan tablolar analizlerde kullanılmak için *Excel Write (XLS)* düğümü ile kayıt edilmektedir. Sezon boyunca bu yolla elde edilen veri sezon sonunda birleştirilmiş böylece fiyatlardaki değişimler ve ortalama fiyatlar hesaplanmıştır.

### 3.1.2. Verinin Analizi ve Hipotez Testleri

2018 yılı Mayıs ayı ilk haftasından itibaren Eylül ayı son haftası da dâhil olacak şekilde toplam 19 hafta boyunca elde edilen fiyatlar, her konaklama tesisi için *ortalama fiyat* olacak şekilde hesaplanmıştır. Veri düzenleme ve temizleme aşamalarında aşağıda verilen işlemler gerçekleştirilmiştir.

Fiyat verisi olmayan oteller veriden silinmiştir. Site Yunanistan, Gürcistan gibi yakın olan bazı ülkeleri de Türkiye otellerine dâhil etmiştir. Bunlar veriden çıkarılmıştır. Çekilen veri ile gerçek veri arasında herhangi bir farklılık olup olmadığı, rastgele bazı konaklama tesislerinin sayfası ziyaret edilerek kontrol edilmiştir. Hafta sonu hafta içi aşırı fiyat farkı olan birkaç konaklama tesisi veriden çıkarılmıştır.

Web sitesi konaklama tesisleri için özgün bir bölge dağılımına sahiptir. Örneğin Kapadokya, Ege Bölgesi, Denizli İli, İstanbul aynı seviyede bölgelerden bazılarıdır. Bunlar hipotez testlerinde göz ardı edilerek Türkiye'nin bölgesel formuna göre değerlendirilmiş ve hipotez testleri o şekilde yapılmıştır.

Araştırma bağımsız değişkenlerinden 43 tanesi konaklama tesisinin sahip olduğu donanım ve hizmetlerden oluşmaktadır. Bunlardan *ücretsiz yüksek hızda internet* değişkeni varyans şişkinlik faktörü ( $5 < VIF$ ) kabul edilebilir değerinin ( $1 < VIF \leq 5$ ) üzerinde çıktığından dolayı veriden çıkarılmıştır. *Ücretsiz internet* adında başka bir değişkenin olması bu durumu açıklar niteliktedir. Bu göz önünde bulundurulduğunda iki değişken birleştirilmiş fakat yeni değişkenin de VIF değeri kabul edilebilir düzeyin üzerinde çıkmıştır (Heiberger ve Holland, 2004: 243). Bu nedenle tutarsızlık olabileceğinden dolayı yeni oluşturulan değişken de dâhil olmak üzere üç değişken de modele dâhil edilmemiştir. Verinin inceleme aşamasında *casino ve talih oyunları* hizmetine sahip 3 tesisin olduğu görülmüştür. 2634 Sayılı Turizmi Teşvik Kanununda Değişiklik Yapılması Hakkında Kanun Tasarısı ile Türkiye'de bu hizmetin verilmesi olanaksızdır. Bu durumun, tesis siteme dâhil edilirken yapılan hatadan kaynaklandığı düşünülmektedir. Bu nedenle bu değişken de veriden çıkarılmıştır.

Hipotez testleri için belirlenen bağımlı değişken *ortalama oda fiyatı* 60,50 TL ile 12279,88 TL arasında değişen sürekli haldeki bir değişkendir ( $N=6558$ ). Yapılan normallik testlerinde verinin normal dağılmadığı görülmektedir. Bu nedenle model test edilirken kullanılacak olan doğrusal regresyon yerine parametrik olmayan *lojistik*

*regresyon* modeli kullanılmıştır. *Ortalama oda fiyatı* değişkeni sıralı (ordinal) hale dönüştürülmeye müsait olduğu için lojistik regresyon modellerinden sıralı *lojistik regresyon* (OLOGREG) tercih edilmiştir.

Ortalama oda fiyatı değişkeni sıralı hale dönüştürülürken yazında yaygın yöntemler ile denemeler yapılmıştır. Bu yöntemler ile yapılan testlerde elde edilen  $R^2$  değerleri ve lojistik regresyon varsayımlarının yerine gelip gelmediğine dair bilgiler Tablo 3.1’de verilmektedir.

**Tablo 3.1.** Ortalama Oda Fiyatı Değişkeninin Gruplandırılmasında Kullanılan Yöntemler

Yöntem	Model Uyum İyiliği (Goodness of Fit)*	Paralel Doğrular Varsayımı (Test of Parallel Lines)*	Pseudo $R^2$ **
Eşit 5’li	+	-	0,36
Eşit 7’li	+	-	0,36
Eşit 9’li	+	-	0,37
Std.Sap $\pm$ 1***	+	+	0,77
Std.Sap $\pm$ 2***	+	+	0,74
Std.Sap $\pm$ 3***	+	+	0,73

Bağlantı Fonksiyonu: Complementary Log-log.

\*  $p > 0,05$

\*\* Cox and Snell  $R^2$

\*\*\* IBM SPSS v.23 /Transform/Visual Binning

Normal dağılmayan *ortalama oda fiyatı* değişkenini *sıralı lojistik regresyon* modeline uygun hale getirmek için sürekli haldeki değişkenin kategorik hale dönüştürülmesi gerekmektedir. Sıralı lojistik regresyonun varsayımları bununla sınırlı değildir. Paralel doğrular varsayımının yerine gelmesi ve model uyum iyiliğinin de uygun olması gerekmektedir. Ortalama oda fiyatı her grup eşit sayıda fiyat içerecek şekilde 5’li, 7’li ve 9’lu olmak üzere gruplandırılmış fakat her üç gruplandırmada da paralel doğrular varsayımı sağlanmamıştır ( $p < 0,05$ ). Standart sapma  $\pm 1$ , 2 ve 3 gruplandırma yöntemlerinden hepsinin lojistik regresyon varsayımları sağlanmıştır. Bağımsız değişkenleri açıklama oranı en yüksek olan gruplandırma Standart sapma  $\pm 1$  yöntemidir ( $R^2 = ,77$ ). Ayrıca gruplandırma için kümeleme algoritmaları (k-Means, k-medoid) kullanılmış fakat modelin  $R^2$  değeri bu değerlerden daha düşük çıkmıştır.

Lojistik regresyon varsayımlarına uyan ve en yüksek bağımsız değişkenleri açıklama oranına ise *standart sapma  $\pm 1$*  gruplandırma yöntemi sahiptir ( $R^2 = 0,95$ ).

Bu nedenle arařtırmada *standart sapma*  $\pm 1$  olarak oluřturulan bağımlı sıralı kategorik deęiřken arařtırmada tercih edilmiřtir.

Birden fazla bağımsız deęiřkeni ieren bir srete, gerekleřme olasılıęı  $\pi(Y = 1|X_1, X_2, \dots, X_p) = \pi(\underline{x})$  ise (bir kategoride olma olasılıęı),  $p$  bağımsız deęiřken iin sıralı lojistik regresyon (OLOGREG) řu řekilde hesaplanmıřtır (O'Connell A. A., 2006: 13).

$$\begin{aligned} \ln(Y') &= \text{logit}[\pi(\underline{x})] = \ln\left(\frac{\pi(\underline{x})}{1 - \pi(\underline{x})}\right) \\ &= \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \end{aligned} \quad (1)$$

Arařtırmada regresyon modelinden bağımsız olarak iki bağımsız grubun ortalamalarının anlamlılıęını test etmek iin bağımsız rneklem t testi (independent samples t test), ikiden fazla grubun ortalamalarının anlamlılıęını test etmek iin *tek faktrl varyans analizi* (oneway ANOVA) yapılmıřtır.

Regresyon modellerinden bağımsız bir řekilde nceki konukların verdięi *yıldız/puan* ile *ortalama oda fiyatı*, *yorum sayısı* ile *ortalama oda fiyatı* ve *sıralama puanı* ile *ortalama fiyat* arasındaki iliřkinin varlıęını test etmek iin korelasyon analizi yapılmıřtır. *Pearson korelasyon analizinden* faydalanılmıřtır.

### 3.2. BULGULAR

Ortalama oda fiyatlarının gruplarına ait fiyat aralıkları ve o aralıkta yer alan ortalama oda fiyatı sayısı Tablo 3.2'de verilmektedir.

**Tablo 3.2:** Ortalama Oda Fiyatı Deęiřkeninin Gruplandırılması

		<i>Frekans</i>	<i>Yzde</i>	<i>Geerli Yzde</i>	<i>Kmlatif Yzde</i>
Gruplar	0 – 395,44	4634	70,7	70,7	70,7
	395,45 – 812,77	1503	22,9	22,9	93,6
	812,78 +	421	6,4	6,4	100,0
	Toplam	6558	100,0	100,0	

Ortalama oda fiyatlarının konaklama tesisinin türüne göre anlamlı farklılık gösterip göstermediği 6558 konaklama tesisinden alınan verilere göre araştırılmıştır. Sonuçlara ilişkin bazı istatistikler Tablo 3.3’de verilmektedir.

**Tablo 3.3.** Tesis Türlerine Ait Bazı İstatistikler

<i>Tesis Türü</i>	<i>N</i>	<i>Ortalama</i>	<i>S.S</i>
Otel	3057	403,86	382,05
Belirtilmemiş	927	561,02	594,12
Pansiyon	1823	311,22	354,06
Özel Konaklamalar	473	374,94	412,07
Apart	148	368,41	285,17
Lodge	86	274,19	232,19
Villa	21	490,12	307,82
Kamping	3	357,99	140,41
Hostel	20	220,94	77,97
<b>Toplam</b>	<b>6558</b>	<b>395,44</b>	<b>417,31</b>

Ortalama oda fiyatlarının tesis türüne göre anlamlı farklılık gösterip göstermediğini test etmek için tek faktörlü varyans analizi (oneway ANOVA) yapılmıştır. Test sonuçları aşağıdaki tabloda verilmektedir.

**Tablo 3.4.** Konaklama Tesisinin Türüne Göre Tek Faktörlü Varyans Analizi Tablosu

<i>Varyans Kaynağı</i>	<i>Kareler Toplamı</i>	<i>s.d.</i>	<i>Kareler Ortalaması</i>	<i>F</i>	<i>p</i>
Gruplar Arası	715,740	8	89,467		
Grup İçi	7975,099	6549	1,218	73,469	0,001
<b>Toplam</b>	<b>8690,839</b>	<b>6557</b>			

Tek faktörlü varyans analizi sonuçlarına göre ortalama oda fiyatları konaklama tesisine göre anlamlı farklılık göstermektedir ( $F(8,6549) = 73,47; p < 0,05$ ). Farklılıkların hangi gruplar arasında olduğunu belirlemek amacıyla çoklu karşılaştırma (post-hoc test) testi yapılmıştır. Grupların varyansları homojenlik göstermediği için *Tamhane T2* testi kullanılmıştır. Buna göre farklılıkların hangi gruplar arasında olduğuna dair tablo aşağıda verilmektedir.

**Tablo 3.5.** Tesis Türü Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları

Tesis Türü	Otel	Belirtilmemiş	Pansiyon	Özel Konaklamalar	Apart	Lodge	Villa	Kamping	Hostel
Otel		*	*	*		*			*
Belirtilmemiş	*		*	*	*	*			*
Pansiyon	*	*		*	*		*		
Özel Konaklamalar	*	*	*			*			*
Apart		*	*			*			*
Lodge	*	*		*	*		*		
Villa			*			*			*
Kamping									
Hostel	*	*		*	*		*		

Tamhane T2 testi değerlendirildiğinde; tesis tipi *otel* (Ort.=403,86; S.S.=382,05) olan işletmelerin *ortalama oda fiyatları*, tesis tipini *belirtmeyen* (Ort.=561,02; S.S.=594,12), *pansiyon* (Ort.=311,22; S.S.=354,06), *özel konaklamalar* (Ort.=374,94; S.S.=412,07), *lodge* (Ort.=274,19; S.S.=232,19) ve *hostellerle* (Ort.=220,94; S.S.=77,97) göre anlamlı farklılık göstermektedir.

Bazı işletmeler tesis tipini belirtmemiştir. Tesis tipini belirtmeyen işletmelerin (Ort.=561,02; S.S.=594,12) anlamlı farklılık gösterdiği işletme türleri şöyledir. *Otel* (Ort.=403,86; S.S.=382,05), *pansiyon* (Ort.=311,22; S.S.=354,06), *özel konaklamalar* (Ort.=374,94; S.S.=412,07), *apart* (Ort.=368,41; S.S.=285,17), *lodge* (Ort.=274,19; S.S.=232,19) ve *hostel* (Ort.=220,94; S.S.=77,97).

*Pansiyon* türü işletmeler, *özel konaklamalar*, *apart* ve *villa* türü işletmeler ile ortalama oda fiyatı bakımından anlamlı farklılık göstermektedir. Diğer tesis türlerinin ortalama oda fiyatı bakımından aralarındaki farklılıkları öğrenmek için Tablo 3.4 ve Tablo 3.5 birlikte incelenebilir.

Türkiye çeşitli coğrafi özelliklere göre yedi bölgeye ayrılmıştır. Konaklama tesisleri de verinin elde edildiği web sitesinde bu bağlamda kategorize edilmese de araştırma sürecinde coğrafi bölgelere göre değerlendirilmiştir. Konaklama tesislerinin coğrafi bölgelere göre dağılımı ve diğer bazı istatistikler Tablo 3.6'da verilmektedir.

**Tablo 3.6.** Coğrafi Bölgelere Göre Ortalama Oda Fiyatları

<i>Bölgeler</i>	<i>N</i>	<i>Gruplandırılmış</i>		<i>Gruplandırılmamış</i>			
		<i>Ort. Fiyat</i>	<i>S.S.</i>	<i>Ort. Fiyat</i>	<i>S.S.</i>	<i>Min.</i>	<i>Mak.</i>
Ege Bölgesi	1765	4,47	1,15	427,69	515,83	80,00	12279,88
Akdeniz Bölgesi	1352	4,46	1,35	475,63	548,24	60,50	9012,60
Marmara Bölgesi	2157	4,39	,96	362,79	257,47	84,33	4301,71
Karadeniz Bölgesi	439	4,29	1,10	366,47	322,16	72,50	3721,00
İç Anadolu Bölgesi	628	3,96	1,13	315,18	275,46	71,50	3373,00
Güneydoğu Anadolu Bölgesi	113	3,49	,94	248,44	230,86	96,00	2272,25
Doğu Anadolu Bölgesi	104	3,51	1,03	249,03	163,19	89,20	1333,33
<b>Toplam</b>	<b>6558</b>	<b>4,35</b>	<b>1,15</b>	<b>395,43</b>	<b>417,33</b>	<b>60,50</b>	<b>12279,88</b>

Tablo incelendiğinde Türkiye’deki konaklama tesislerinin ortalama oda fiyatlarının 60,50 TL ile 12279 TL arasında değiştiği görülmektedir. Ortalama oda fiyatlarının ortalaması ise 395,43 TL’dir. Bölgelere bakıldığında, ortalama oda fiyatı en yüksek olan bölgenin 475,63 TL ile Akdeniz Bölgesi olduğu görülmektedir. Ortalama oda fiyatı en düşük olan bölge ise 248,44 TL ile Güneydoğu Anadolu Bölgesi’dir. Bölgelerde fiyat istatistikleri bu şekildedir. Ortalama oda fiyatlarının bölgelere göre anlamlı farklılık gösterip göstermediğini belirlemek için tek faktörlü varyans analizi (oneway ANOVA) yapılmıştır. Test sonuçları aşağıdaki **Tablo 3.7**’de verilmektedir.

**Tablo 3.7.** Tesisin Bulunduğu Bölgeye Göre Tek Faktörlü Varyans Analizi Tablosu

<i>Varyans Kaynağı</i>	<i>Kareler Toplamı</i>	<i>s.d.</i>	<i>Kareler Ortalaması</i>	<i>F</i>	<i>p</i>
Gruplar Arası	301,714	6	50,286		
Grup İçi	8389,124	6551	1,281	39,286	,001
<b>Toplam</b>	<b>8690,839</b>	<b>6557</b>			

Tek faktörlü varyans analizi sonuçlarına göre ortalama oda fiyatları konaklama tesisine bulunduğu *coğrafi bölgeye* göre anlamlı farklılık göstermektedir ( $F(6,6551) = 39,27$ ;  $p < 0,05$ ). Farklılıkların hangi bölgeler arasında olduğunu belirlemek amacıyla çoklu karşılaştırma (post-hoc test) testi yapılmıştır. Grupların varyansları homojenlik göstermediği için *Tamhane T2* testi kullanılmıştır. Buna göre farklılıkların hangi gruplar arasında olduğuna dair tablo aşağıda verilmektedir.



**Tablo 3.8.** Tesisin Bulunduğu Bölge Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları

<i>Bölge</i>	Ege	Akdeniz	Marmara	Karadeniz	İç Anadolu	Güneydoğu Anadolu	Doğu Anadolu
Ege				*	*	*	*
Akdeniz					*	*	*
Marmara					*	*	*
Karadeniz	*				*	*	*
İç Anadolu	*	*	*	*		*	*
Güneydoğu Anadolu	*	*	*	*	*		
Doğu Anadolu	*	*	*	*	*		

Tamhane T2 testi değerlendirildiğinde; *Ege Bölgesinde* bulunan tesislerin ortalama oda fiyatları (Ort.=427,69; S.S.=515,83), *Karadeniz* (Ort.=366,47; S.S.=322,16), *İç Anadolu* (Ort.=315,18; S.S.=275,46), *Güneydoğu Anadolu* (Ort.=248,44; S.S.=230,86) ve *Doğu Anadolu* (Ort.=249,03; S.S.=163,19) bölgesindekilerden anlamlı farklılık göstermektedir.

**Tablo** incelendiğinde *Ege Bölgesi*, *Marmara Bölgesi* ve *Akdeniz Bölgesinin* ortalama oda fiyatlarının kendi aralarında anlamlı farklılık göstermediği görülmektedir. Bu üç bölge ve *Karadeniz Bölgesi*; *İç Anadolu*, *Güneydoğu Anadolu* ve *Doğu Anadolu bölgeleri* ile anlamlı farklılık göstermektedir.

Ortalama oda fiyatları, tesisin bulunduğu şehre göre de incelenmiştir. Şehirdeki *toplam konaklama tesisi sayısı*, gruplandırılmış ortalama oda fiyatları, en düşük ve en yüksek oda fiyatları gibi bazı istatistikler ile *en çok tesisin bulunduğu* ilk 20 il aşağıdaki tabloda verilmektedir.

**Tablo 3.9.** İllere Göre Ortalama Oda Fiyatları

<i>Şehir</i>	<i>N</i>	<i>Gruplandırılmış</i>		<i>Gruplandırılmamış</i>		<i>Min.</i>	<i>Mak.</i>
		<i>Ortalama Fiyat</i>	<i>S.S.</i>	<i>Ortalama Fiyat</i>	<i>S.S.</i>		
İstanbul	1560	4,44	0,99	376,66	277,51	84,33	4301,71
Antalya	1171	4,57	1,37	510,78	579,97	60,50	9012,60
Muğla	948	4,66	1,16	494,51	645,94	98,25	12279,88
İzmir	547	4,41	1,06	376,76	314,18	80,00	4482,14
Nevşehir	278	4,41	1,13	393,66	332,18	71,50	3373,00
Ankara	191	3,82	1,01	275,18	173,18	95,50	1428,29
Balıkesir	181	4,30	0,80	321,46	173,17	116,63	1840,11
Aydın	156	4,01	1,21	322,67	228,87	86,75	1840,36
Çanakkale	151	4,50	0,80	353,76	147,73	139,25	1307,55
Trabzon	113	5,09	0,96	536,19	370,05	158,00	2664,00
Bursa	93	4,29	0,87	323,75	133,42	131,00	900,43
Denizli	65	3,65	1,11	249,29	125,51	100,00	666,40
Mersin	64	3,94	0,89	277,37	109,13	120,93	558,33
Karabük	52	3,77	1,13	312,55	406,84	72,50	2490,86
Kocaeli	52	3,71	0,94	254,54	97,76	105,25	568,80
Konya	48	3,38	0,94	218,00	73,90	97,29	466,63
Adana	47	3,28	0,83	199,93	61,06	113,00	409,00
Sakarya	43	4,30	1,01	415,05	459,08	150,00	2869,10
Samsun	43	3,47	0,98	230,21	102,89	91,00	538,71
Eskişehir	43	3,05	0,72	186,24	43,50	105,20	361,75

Veriler incelendiğinde en çok otelin sırası ile İstanbul (N=1560), Antalya (N=1171), Muğla (N=948) ve İzmir'de (N=547) bulunduğu görülmektedir. Bu şehirlerden Antalya, Muğla, İzmir şehirleri kıyı otelciliğinin yoğunlukta olduğu yerler

iken, İstanbul ve Nevşehir (N=278) kültürel turizm ile ön plana çıkan şehirlerdir. Aşağıdaki şekilde illere göre konaklama tesisi sayıları grafiksel olarak gösterilmiştir.



**Şekil 3.1.** İllere Göre Konaklama Tesisi Sayıları

Ortalama oda fiyatları yüksekten düşüğe doğru sıralanmış, Türkiye’deki ortalama oda fiyatları en yüksek olan 10 il **Tablo 3.10**’de verilmektedir.

**Tablo 3.10.** Ortalama Oda Fiyatı En Yüksek Olan On İl

Şehir	N	Gruplandırılmış		Gruplandırılmamış		Min.	Mak.
		Ortalama Fiyat	S.S.	Ortalama Fiyat	S.S.		
Karaman	4	4,25	1,89	892,71	1359,01	184,00	2930,50
Trabzon	113	5,09	0,96	536,19	370,05	158,00	2664,00
Antalya	1171	4,57	1,37	510,78	579,97	60,50	9012,60
Muğla	948	4,66	1,16	494,51	645,94	98,25	12279,88
Giresun	18	4,28	1,13	468,57	817,25	155,00	3721,00
Rize	34	4,88	0,88	447,18	206,74	186,70	815,50
Sakarya	43	4,30	1,01	415,05	459,08	150,00	2869,10
Nevşehir	278	4,41	1,13	393,66	332,18	71,50	3373,00
Burdur	5	4,60	1,52	385,62	163,26	135,00	555,20
İzmir	547	4,41	1,06	376,76	314,18	80,00	4482,14

Tablo incelendiğinde ortalama oda fiyatı açısından en yüksek fiyata sahip ilin Karaman (Ort.=892,71; S.S.=1359,01) olduğu görülmektedir. Fakat dikkat edileceği üzere bu ildeki konaklama tesisi sayısı 4’tür. Ayrıca standart sapmanın (S.S.=1359,01) diğer illere nazaran daha yüksek olduğu görülmektedir. Kıyı turizmi açısından önemli illerden olan Antalya (Ort.=510,78; S.S.=597,97) ve Muğla’da (Ort.=494,51; S.S.=645,94) ise ortalama fiyat farkının yaklaşık 16 TL olduğu görülmektedir.

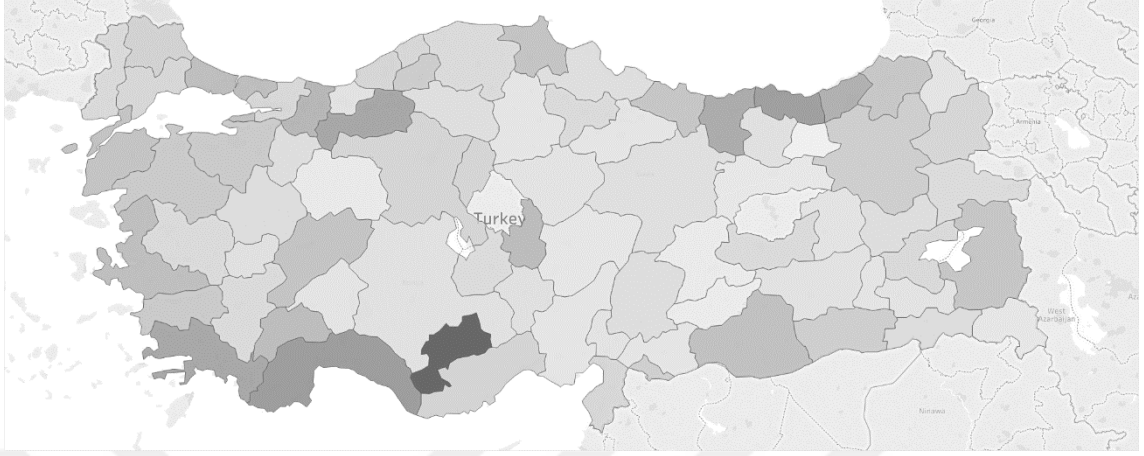
Şehirlere göre ortalama oda fiyatları düşükten yükseğe doğru sıralanmış, buna göre Türkiye’deki en düşük 10 il aşağıdaki Tablo 3.11’de verilmiştir.

**Tablo 3.11.** Ortalama Oda Fiyatı En Düşük Olan On İl

Şehir	N	Gruplandırılmış		Gruplandırılmamış		Min.	Mak.
		Ortalama Fiyat	S.S.	Ortalama Fiyat	S.S.		
Bayburt	1	3,00		157,00		157,00	157,00
Adıyaman	9	3,11	0,33	176,87	25,22	150,00	234,71
Malatya	20	2,90	1,02	180,43	60,55	89,20	329,00
Eskişehir	43	3,05	0,72	186,24	43,50	105,20	361,75
Erzincan	9	3,11	0,60	192,72	36,61	123,50	261,25
Kayseri	27	3,19	1,04	194,56	69,77	99,00	377,17
Hakkâri	1	3,00		195,00		195,00	195,00
Adana	47	3,28	0,83	199,93	61,06	113,00	409,00
Çorum	10	3,30	0,67	199,93	45,27	130,00	278,00
Tokat	6	3,33	0,52	202,22	33,79	168,17	252,29

Tablo incelendiğinde ortalama en düşük oda fiyatının Bayburt (Ort.=157,00) ilinde olduğu görülmektedir. Fakat bu ilde araştırmaya dâhil edilen sadece bir tesis bulunmaktadır. Hakkâri (Ort.=195,00) ilinde de Bayburt ilindeki gibi sadece bir tesis araştırmaya dâhildir. Diğer iller incelendiğinde ortalama oda fiyatları bakımından ortalama en düşük fiyata sahip illerin sırasıyla Adıyaman (Ort.=176,87; S.S.=25,22), Malatya (Ort.=180,43; S.S.=60,55), Eskişehir (Ort.=186,24; S.S.=43,50) ve Erzincan (Ort.=192,72; S.S.=36,61) şeklinde olduğu görülmektedir.

Tüm Türkiye’deki ortalama oda fiyatlarının illere göre grafiksel gösterimi Şekil 3.2’de verilmektedir.



**Şekil 3.2.** İllere Göre Ortalama Oda Fiyatları

Bütün bu betimleyici istatistiklerden sonra ortalama oda fiyatlarının illere göre anlamlı farklılık gösterip göstermediğini belirlemek için tek faktörlü varyans analizi (oneway ANOVA) yapılmıştır. Test sonuçları aşağıdaki tabloda verilmektedir.

**Tablo 3.12.** Tesisin Bulunduğu İle Göre Tek Faktörlü Varyans Analizi Tablosu

<i>Varyans Kaynağı</i>	<i>Kareler Toplamı</i>	<i>s.d.</i>	<i>Kareler Ortalaması</i>	<i>F</i>	<i>p</i>
Gruplar Arası	949,62	77	12,333		
Grup İçi	7741,22	6480	1,195	10,323	0,001
<i>Toplam</i>	<i>8690,84</i>	<i>6557</i>			

Tek faktörlü varyans analizi sonuçlarına göre ortalama oda fiyatları *tesis*in bulunduğu ile göre anlamlı farklılık göstermektedir ( $F(77,6480) = 10,33; p < 0,05$ ). Faktör sayısı 50’den fazla olduğundan dolayı çoklu karşılaştırma testi (post-hoc) yapılmamıştır.

Model içerisinde önemli olabilecek bir diğer değişken konaklama tesisinin *deniz kıyısında* olup olmamasıdır. Yapılan analizde analize dâhil edilen otellerin %10 deniz kıyısında olup olmamasına ilişkin bazı istatistikler tabloda verilmektedir.

**Tablo 3.13.** Tesisin Deniz Kıyısında Olup Olmamasına İlişkin Bazı İstatistikler

<i>Deniz Kıyısı</i>	<i>N</i>	<i>Ort.</i>	<i>S.S.</i>	<i>Min.</i>	<i>Mak.</i>
Hayır	5967	378,46	402,92	60,50	12279,88
Evet	591	566,75	511,55	91,00	4567,69
Toplam	6558	395,43	417,33	60,50	12279,88

Deniz kıyısında olmayan konaklama tesislerinin ortalama oda fiyatları 378,46 TL iken, deniz kıyısında olan konaklama tesislerinin ortalama oda fiyatları 566,75 TL'dir. Bu istatistiklerin anlamlı olup olmadığını belirlemek için Bağımsız Örneklem T Testi (Independent Samples T Test) yapılmıştır. Test sonuçları tabloda verilmektedir.

**Tablo 3.14.** Tesisin Deniz Kıyısında Olmasına Dair Bağımsız Örneklem T Testi Sonuçları

<i>Deniz Kıyısı</i>	<i>N</i>	<i>Ort.</i>	<i>S.S.</i>	<i>S.D.</i>	<i>t</i>	<i>P</i>
Hayır	5967	378,46	402,92	6556	-12,510	0,001
Evet	591	566,75	511,55			

Bulgulara göre ortalama oda fiyatları konaklama tesislerinin deniz kıyısında olup olmaması bakımından anlamlı farklılık göstermektedir ( $t_{(6556)} = -12,510$ ;  $p > 0,05$ ). Deniz kıyısında olan otellerin ortalama oda fiyatları deniz kıyısında olmayanlara nazaran daha yüksektir. Araştırma verisinde tesisi daha önce ziyaret edenler tarafından 1'den 5'e kadar 0,5 artan puanların bulunduğu *yıldız/puan* değişkeni bulunmaktadır. Toplam 9 seviyeden oluşan bu puanlandırmaya ilişkin bazı istatistikler tabloda verilmektedir.

**Tablo 3.15.** Tesislerin Konuk Puanları ve Puanlara Dair Bazı İstatistikler

<i>Yıldız/Puan</i>	<i>N</i>	<i>Ort. Fiyat</i>	<i>S.S.</i>	<i>Min.</i>	<i>Mak.</i>
1,0	39	287,70	263,21	72,50	1514,57
1,5	33	282,83	201,47	120,29	1290,36
2,0	84	283,55	315,27	70,40	2869,10
2,5	205	270,27	209,55	85,88	2043,55
3,0	507	306,71	223,31	69,50	2037,80
3,5	1016	337,26	330,47	89,20	5065,20
4,0	1743	375,49	312,74	80,00	6781,67
4,5	1794	483,24	587,75	71,50	12279,88
5,0	866	439,37	364,75	60,50	4680,00
Boş	271	344,76	432,51	91,00	3721,00
<i>Toplam</i>	<i>6558</i>	<i>395,44</i>	<i>417,34</i>	<i>60,50</i>	<i>12279,88</i>

Yıldız/Puan ile ortalama oda fiyatı arasındaki ilişkinin varlığını test etmek için korelasyon analizi yapılmıştır. Herhangi bir gruplandırma yapılmamış ortalama oda fiyatı ile konukların vermiş olduğu *yıldız/puan* arasındaki ilişkiye dair korelasyon analizi sonucu tabloda verilmektedir.

**Tablo 3.16.** Yıldız/Puan ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları

			Ortalama Fiyat	Yıldız
<i>Pearson</i>	<i>Ortalama Oda</i>	Korelasyon Katsayısı	1,000	0,245*
<i>Korelasyon</i>	<i>Fiyatı</i>	p. (2-tailed)	.	0,001
		N	6558	6287

\*,  $p < 0.01$

Sonuçlara göre konukların verdiği *yıldız/puan* ile ortalama oda fiyatları arasında anlamlı ve pozitif yönlü ilişki görülmektedir ( $r(6287)=0,245$ ;  $p<0,01$ ).

Puanlar 1'den 5'e kadar 0,5 artarak toplam 9 seviyeden oluşmaktadır. Bu puanlar nümerik gibi görünse de aslında sıralı kategorik haldedir. Bu şekilde *yıldız/puan* değişkeni yeniden kodlanmış ve 1'den 9'a kadar bir seviyeden oluşan yeni bir *yıldız/puan* değişkeni oluşturulmuştur. Ortalama oda fiyatının bu yeni *yıldız/puan* değişkenine göre anlamlı farklılık gösterip göstermediğini test etmek için tek faktörlü varyans analizi (oneway ANOVA) yapılmıştır. Bu analiz sonuçlarına ilişkin sonuçlar aşağıdaki tabloda verilmektedir.

**Tablo 3.17.** Tesislerin Konuk Puanlarına Göre Tek Faktörlü Varyans Analizi Tablosu

Varyans Kaynağı	Kareler Toplamı	s.d.	Kareler Ortalaması	F	p
Gruplar Arası	530,80	8	66,351		
Grup İçi	7621,94	6278	1,214	54,651	0,001
Toplam	8152,74	6286			

Tek faktörlü varyans analizi sonuçlarına göre ortalama oda fiyatları tesisin aldığı *yıldız/puana* göre anlamlı farklılık göstermektedir ( $F(8,6278) = 54,65$ ;  $p<0,05$ ). Bu farklılıkların hangi gruplar arasında olduğunu belirlemek için çoklu karşılaştırma testi yapılmıştır. Grup varyansları homojenlik göstermediği için *Tamhane T2* testi tercih edilmiştir. Bu teste ilişkin sonuçlar tabloda verilmektedir.

**Tablo 3.18.** Konuk Puanları Bakımından Çoklu Karşılaştırma Tamhane T2 Testi Sonuçları

<i>Puan</i>	1	2	3	4	5	6	7	8	9
1									*
2								*	*
3						*	*	*	*
4					*	*	*	*	*
5				*			*	*	*
6			*	*			*	*	*
7			*	*	*	*		*	*
8	*	*	*	*	*	*	*		
9	*	*	*	*	*	*	*		

Çoklu karşılaştırma testi sonuçları incelendiğinde düşük puan alan tesislerin ortalama oda fiyatları ile yüksek puan alanların ortalama oda fiyatlarında anlamlı farklılıklar görülmektedir. Örneğin 1 yıldız/puan alan (Ort.=3,64; S.S.=1,442) tesislerin ortalama oda fiyatları 8 yıldız/puan (Ort.=4,65; S.S.=1,148) ve 9 yıldız/puan alanlarla (Ort.=4,65; S.S.=1,044) anlamlı farklılık göstermektedir. Benzer şekilde 8 yıldız/puan ve 9 yıldız/puan alanlarla, 1, 2 (Ort.=3,85; S.S.=1,034), 3 (Ort.=3,63; S.S.=1,170), 4 (Ort.=3,66; S.S.=1,080), 5 (Ort.=3,96; S.S.=1,064), 6 (Ort.=4,14; S.S.=1,073) ve 7 yıldız/puan alan (Ort.=4,34; S.S.=1,101) tesislerin ortalama oda fiyatları da anlamlı farklılık göstermektedir. Bu durum fiyat artığında yıldız/puanın da artığı anlamına gelmektedir. Korelasyon sonucu düşük bir ilişkiyi gösterse de ilişki yönü pozitifdir ( $r(6287)=,245$ ;  $p<0,01$ ).

ÇSA'ların ziyaretçilere sağladığı en büyük avantajlardan biri site üzerinde satışı bulunan tesisler hakkında orayı daha önce ziyaret edenlerin yaptıkları yorumlardır. Turistik ürünün yapısı gereği bu sektöre özellikle müşteri lehine önemli katkılar sağlamaktadır (Keskinçılıç, Ağca ve Karaman, 2016). Diğer pek çok e-ticaret sitesinde de benzer durum söz konusudur. Araştırma verisinde de *yorum* değişkeni bulunmaktadır. Bu değişken o tesisin aldığı yorum sayısını ifade etmektedir. Yorum sayısı belirlenirken sezon sonundaki yorum sayısı dikkate alınmıştır. Her ne kadar yorum sayısının ortalama oda fiyatı ile ilişkili olabileceği düşünülmese de tesisin tercih edilebilirliği üzerinde etkili olabilmektedir. Yorum sayısı o tesisin memnuniyet puanı hakkında bilgi vermese de çok yorum alması kısmen çok tercih edilmiş olduğu anlamı taşımaktadır. Aşağıdaki tabloda bölgelere göre tesislerin aldıkları toplam yorum sayıları ve ortalama yorum sayıları verilmektedir.



**Tablo 3.19.** Bölgelere Göre Tesislerin Aldıkları Yorumlara İlişkin İstatistikler

<i>Bölge</i>	<i>Toplam Yorum</i>	<i>Ortalama Yorum</i>
Akdeniz Bölgesi	326.384	251,26
Doğu Anadolu Bölgesi	3.936	40,58
Ege Bölgesi	388.812	225,53
Güneydoğu Anadolu Bölgesi	6.376	57,96
İç Anadolu Bölgesi	80.268	130,73
Karadeniz Bölgesi	14.870	37,93
Marmara Bölgesi	324.782	158,20
<i>Genel Toplam</i>	<i>1.145.428</i>	<i>182,13</i>

Sonuçlar incelendiğinde en çok yorumu Ege Bölgesinin (388.812) aldığı görülmektedir. İkinci sırada ise birbirlerine çok yakın sayılarla Akdeniz (326.384) ve Marmara (324.782) Bölgeleri gelmektedir. Araştırma verisi göz önünde bulundurulduğunda Türkiye’de araştırmanın yapıldığı tarih itibariyle tesis başına 182,13 yorum yapılmıştır. Yorum ile ortalama oda fiyatı arasında mantıken bir ilişki olmadığı düşünülse de tercih edilen tesislerin daha çok yorum alacağı, dolayısıyla talep yüksekliği ilişkisinden, dolaylı olarak fiyatla ilişkili olabileceği düşünüldüğü için bu iki değişken için korelasyon analizi yapılmıştır. *Yorum sayısı* ile ortalama oda fiyatı arasındaki ilişkinin varlığını ve derecesini ölçmek için yapılan korelasyon analizi sonuçları tabloda verilmektedir.

**Tablo 3.20.** Yorum Sayısı ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları

		<i>Ortalama Fiyat</i>	<i>Yorum Sayısı</i>
<i>Pearson Korelasyon</i>	<i>Ortalama Oda Fiyatı</i>	Korelasyon Katsayısı p. (2-tailed) N	1,000 . 6558
			0,356* 0,001 6289

\*.  $p < 0.01$

Sonuçlar değerlendirildiğinde *yorum sayısı* ile ortalama oda fiyatı arasında pozitif yönlü orta düzey anlamlı bir ilişki görülmektedir ( $r(6289)=0,356$ ;  $p<0,01$ ).

Araştırma verisinde, konaklama tesislerini aldıkları konuk puanlarına göre sıralamak için *sıralama* değişkeni bulunmaktadır. Bu değişken o tesisin bölgedeki tesisler içerisinde kaçınıcı sırada olduğunu belirtmektedir. Bu tek başına toplam veri içerisinde çok anlamlı olmamaktadır. Ayrıca her sıralama o bölge için geçerli olduğundan bu değişkeni tek başına genel sıralama için kullanmak hatalı sonuçlara yol açmaktadır. Örneğin İstanbul ilindeki 1211 konaklama tesisi içerisinde 1’nci sırada olan işletme ile

Muğla/Bozburun'da bulunan 20 konaklama tesis içerisinde 1'nci olan işletmenin sıralaması farklı olmalıdır.

Bu nedenle yeni bir sıralama puanı oluşturmak gerekmektedir. Bunun için şu şekilde bir yol izlenmiştir.

$$\frac{\text{Toplam Otel Sayısı}}{\text{Bölgedeki Sırası}} = x \quad (2)$$

$$\text{Yeni Sıralama Puanı} = \frac{xi - \min(x)}{\text{mak}(x) - \min(x)} \quad (3)$$

Bu yöntem ile *yeni sıralama puanı* hesaplanmıştır. Böylece daha fazla rakip arasında üst sıralarda olan işletmelerin puanları daha yüksek olmuştur. Hesaplanan bu yeni sıralama puanı ile ortalama oda fiyatı arasında ilişki durumunu araştırmak için korelasyon analizi yapılmıştır. Analiz sonuçları tabloda verilmektedir.

**Tablo 3.21.** Sıralama Puanı ile Ortalama Oda Fiyatı Arasındaki Korelasyon Sonuçları

			<i>Ortalama Fiyat</i>	<i>Sıralama Puanı</i>
<i>Pearson Korelasyon</i>	<i>Ortalama Oda Fiyatı</i>	Korelasyon Katsayısı	1,000	0,260*
		p. (2-tailed)	.	0,001
		N	6224	2348

\*. p < 0.01

Sonuçlar değerlendirildiğinde *tesisın sıralama puanı* ile *ortalama oda fiyatı* arasında pozitif yönlü zayıf bir ilişki görülmektedir ( $r(2348)=0,260$ ;  $p<0,01$ ). Bu yöntem ile elde edilen *yeni sıralama puanı* geniş bir araştırma ile test edilmediğinden uygulanacak olan regresyon modellerinde eski hali ile yer almaktadır.

### 3.2.1. Regresyon Modeline İlişkin Bulgular

Araştırmanın verisi toplam 49 adet değişken içermektedir. Bunlardan 43 tanesi tesisin sahip olduğu donanım ya da hizmetlerin olup olmadığını ifade ederken, 2 tanesi tesisin bulunduğu *şehir* ve *bölgeyi* ifade etmektedir. Daha önce tesisi ziyaret edenlerin verdikleri puanlara göre hesaplanan *yıldız/puan* ve o bölgede bulunan toplam tesis içerisinde bu puanlamaya göre kaçınıcı sırada olduğunu gösteren *sırası* değişkeni, bölgedeki *toplam otel sayısı*, *yorum sayısı*, *tesis tipi* veri setindeki diğer değişkenlerdir.

İyi bir regresyon modeli oluşturmak için bu değişkenlerin sayısını azaltma yoluna gidilmiştir.

Daha önce belirtildiği üzere regresyon analizinin koşullarından olan (Heiberger ve Holland, 2004: 243) çoklu doğrusal bağımlılık olmaması kuralına aykırı *ücretsiz yüksek hızda internet* değişkeninin varyans şişkinlik faktörü ( $5 < VIF$ ) kabul edilebilir değerin ( $1 < VIF \leq 5$ ) üzerinde çıkmıştır. *Ücretsiz internet* adında başka bir değişkenin olması bu durumu açıklar niteliktedir. Bu göz önünde bulundurulduğunda iki değişken birleştirilmiş fakat yeni değişkenin de VIF değeri kabul edilebilir düzeyin üzerinde çıkmıştır. Bu nedenle tutarsızlık olabileceğinden dolayı üç değişken de modele dâhil edilmemiştir.

Ayrıca *casino ve talih oyunları* değişkeni 2634 Sayılı Turizmi Teşvik Kanununda Değişiklik Yapılması Hakkında Kanun Tasarısı ile Türkiye’de yasak olmasından dolayı modelden çıkarılmıştır. Toplam veri içerisinde sadece bu hizmete sahip 3 tesis bulunmaktadır. Bu durumun, tesis siteme dâhil edilirken yapılan hatadan kaynaklandığı düşünülmektedir. *Bar/oturma salonu* değişkeni hemen hemen her otelde bulunmasına karşın yapılan incelemede tesislerin sadece %40’ında bu donanımın olduğu görülmektedir. Bu tesislerin büyük bir kısmının ortalama gecelik konaklama ücretinin 1000 TL ve üstü olan işletmeler olduğu görülmüş ve böyle bir olasılığın düşük olacağından bu değişken de modelden çıkarılmıştır. Değişken modelden çıkarıldıktan sonra yapılan regresyon analizinde bunun  $R^2$  değeri üzerinde herhangi bir değişime neden olmadığı görülmüştür.

Bunlar dışında regresyon modeli içerisinde yer alan anlamlı olmayan bağımsız değişkenlerin belirlenebilmesi için *aşamalı regresyon (stepwise regression)* analizi yapılmıştır. Bu yöntemlerden *ileriye doğru seçim (forward selection)* veya *geriye doğru ayıklama (backward elimination)* tipleri arasında  $R^2$  değerleri açısından önemli bir farklılığın olmadığı görülmüştür. Yapılan aşamalı regresyon sonucunda *Yorum, Şehir, Yıldız, Spor Salonlu/Spor Odalı Fitness Merkezi, Tesis Tipi, Deniz Kıyısı, Kapalı Havuz, Çamaşır Yıkama Hizmeti, Mini Bar, Klima, Mikrodalga Fırın, Spa, Ücretsiz Otopark, Bölge* değişkenlerini içeren modelin en yüksek  $R^2$  değerine sahip olduğu görülmektedir. Aşağıdaki tabloda en yüksek  $R^2$  değerine sahip model ve bu modele ilişkin bazı istatistiksel sonuçlar verilmektedir.

**Tablo 3.22.** Regresyon Modeli Belirlenmesine Yönelik Aşamalı Regresyon Sonuçları

Bağımlı Değişken	Bağımsız Değişkenler	Beta	t	p	R <sup>2</sup>	Düzeltilmiş R <sup>2</sup>	F
Ortalama Oda Fiyatı	Yorum	,212	17,397	,000			
	Şehir	-,203	-16,080	,000			
	Yıldız	,196	17,450	,000			
	Spor Salonlu	,068	3,820	,000			
	Tesis Tipi	-,099	-8,639	,000			
	Deniz Kıyısı	,093	7,449	,000			
	Kapalı Havuz	,055	3,522	,002	0,256	0,254	152,23*
	Çamaşır Yıkama Hiz.	-,064	-3,808	,001			
	Mini Bar	,099	6,402	,000			
	Klima	-,104	-5,798	,000			
	Mikrodalga Fırın	,041	3,639	,000			
	Spa	,052	2,898	,004			
	Ücretsiz Otopark	-,036	-2,561	,010			
	Bölge	-,025	-2,417	,016			

\*  $p < 0,05$ ;  $F(14,6209) = 152,23$ ;  $R^2 = 0,256$

Asıl modelin testinde *sıralı lojistik regresyon* (ordinal logistic regression, OLOGREG) yöntemi kullanılmıştır. Bu yöntemin *çoklu lineer regresyon* modeline tercih edilmesinde, lineer regresyonun varsayımlarından olan *normal dağılımın* bağımlı değişkende görülmemesi etkili olmuştur.

Modeldeki bağımlı değişken olan *ortalama oda fiyatı* değişkenine ait bazı istatistiki değerler tabloda verilmektedir.

**Tablo 3.23.** Bağımlı Değişkene (Ortalama Oda Fiyatı) Ait Bazı İstatistiki Değerler

	N	Min.	Mak.	Ort.	Std. Sap.	Çarpıklık		Basıklık	
						Değ.	Std. Hata	Değ.	Std. Hata
Ortalama Oda Fiyatı	6558	60,50	12280	395,43	417,34	9,488	0,030	171,70	0,060

Veriler incelendiğinde sezon boyunca elde edilen ortalama oda fiyatının 60,50 TL ile 12280 TL arasında değiştiği görülmektedir. Normal dağılmayan bu değişkeni lojistik regresyona uygun olacak sıralı kategorik (ordinal) hale getirmek için belirli bir düzene göre gruplandırmak gerekmektedir. Yazında bunun için kullanılan bazı yöntemler mevcuttur. Fakat ortalama oda fiyatlarının yaklaşık %98'i 60,50 TL - 1418 TL arasındadır. Geriye kalan %2'lik kısım ise 1419 TL - 12280 TL aralığındadır. Yazındaki  $(\max(x) - \min(x))/\text{grup sayısı}$  gibi yöntemler hem mantık hatasına hem de model uyumunun olmamasına veya düşük R<sup>2</sup> değerlerine neden olmaktadır. Bu nedenle veri

üzerinde eşit 3'lü, 5'li, 7'li, 9'lu,  $\pm 1$ , 2, 3 standart sapma yöntemleri denenmiş en yüksek  $R^2$  değeri ( $R^2 = ,78$  (Cox and Snell);  $p < 0,05$ ) ve model uyumluluğu  $\pm 1$  standart sapma kullanılarak yapılan grupta görülmüştür. Standart sapma  $\pm 1$  yöntemi kullanılarak oluşturulan grupların frekans değerleri aşağıdaki tabloda verilmektedir.

**Tablo 3.24.** Ortalama Oda Fiyatı Değişkeninin Gruplandırılması

		<i>Frekans</i>	<i>Yüzde</i>	<i>Geçerli Yüzde</i>	<i>Kümülatif Yüzde</i>
Gruplar	0 – 395,44	4634	70,7	70,7	70,7
	395,45 – 812,77	1503	22,9	22,9	93,6
	812,78 +	421	6,4	6,4	100,0
	Toplam	6558	100,0	100,0	

Veri setinin model içerisine alınan değişkenlerine ait bazı istatistikler tabloda verilmektedir.

**Tablo 3.25.** Modelde Yer Alan Değişkenler ve Değişkenlere Ait Frekanslar (N=6558)

<i>Değişkenler</i>	<i>Değerler</i>	<i>N</i>	<i>Yüzde</i>
Şehir	İstanbul	1493	23,70%
	Antalya	1126	17,90%
	Muğla	934	14,90%
	İzmir	530	8,40%
	Nevşehir	276	4,40%
	Ankara	184	2,90%
	Balıkesir	171	2,70%
	Aydın	150	2,40%
	Çanakkale	145	2,30%
	Trabzon	90	1,40%
	Bursa	90	1,40%
	Denizli	62	1,00%
	Mersin	59	0,90%
	Karabük	47	0,70%
	Kocaeli	47	0,70%
	Konya	47	0,70%
	Adana	47	0,70%
	Eskişehir	42	0,70%
	Sakarya	38	0,60%
	Hatay	37	0,60%
	Gaziantep	36	0,60%
Samsun	35	0,60%	
Rize	33	0,50%	
Bolu	31	0,50%	

**Tablo 3.25.** Modelde Yer Alan Değişkenler ve Değişkenlere Ait Frekanslar (Devam)

Kayseri	27	0,40%
Mardin	23	0,40%
Manisa	22	0,30%
Tekirdağ	22	0,30%
Kastamonu	21	0,30%
Yalova	20	0,30%
Ordu	20	0,30%
Edirne	20	0,30%
Bartın	20	0,30%
Amasya	19	0,30%
Sinop	18	0,30%
Malatya	18	0,30%
Kars	17	0,30%
Diyarbakır	17	0,30%
Giresun	16	0,30%
Şanlıurfa	15	0,20%
Isparta	14	0,20%
Elazığ	13	0,20%
Afyonkarahisar	13	0,20%
Erzurum	13	0,20%
Aksaray	12	0,20%
Van	11	0,20%
Sivas	11	0,20%
Düzce	11	0,20%
Zonguldak	9	0,10%
Kahramanmaraş	9	0,10%
Adıyaman	9	0,10%
Çorum	8	0,10%
Erzincan	8	0,10%
Bitlis	7	0,10%
Kütahya	6	0,10%
Uşak	6	0,10%
Tokat	6	0,10%
Batman	6	0,10%
Burdur	5	0,10%
Kırklareli	5	0,10%
Artvin	5	0,10%
Niğde	5	0,10%
Karaman	4	0,10%
Çankırı	3	0,00%
Ağrı	3	0,00%
Ardahan	3	0,00%

**Tablo 3.25.** Modelde Yer Alan Değişkenler ve Değişkenlere Ait Frekanslar (Devam)

	Gümüşhane	2	0,00%
	Muş	2	0,00%
	Şırnak	2	0,00%
	Bilecik	2	0,00%
	Yozgat	2	0,00%
	Kırıkkale	1	0,00%
	Kilis	1	0,00%
	Bingöl	1	0,00%
	Siirt	1	0,00%
	Osmaniye	1	0,00%
	Hakkâri	1	0,00%
	Bayburt	1	0,00%
Bölge	Marmara Bölgesi	2053	32,70%
	Ege Bölgesi	1723	27,40%
	Akdeniz Bölgesi	1298	20,60%
	İç Anadolu Bölgesi	614	9,80%
	Karadeniz Bölgesi	392	6,20%
	Güneydoğu Anadolu Bölgesi	110	1,70%
	Doğu Anadolu Bölgesi	97	1,50%
Tesis Tipi	Otel	3021	48,10%
	Oda ve Kahvaltı / Pansiyon	1704	27,10%
	Belirtilmemiş	902	14,30%
	Özel Konaklamalar	422	6,70%
	Apart	124	2,00%
	Lodge	76	1,20%
	Hostel	18	0,30%
	Villa	17	0,30%
	Kamping	3	0,00%
Spor Salonlu/Spor Odalı Fitness Merkezi	Yok	5614	89,30%
	Var	673	10,70%
Deniz Kıyısı	Yok	5716	90,90%
	Var	571	9,10%
Kapalı Havuz	Yok	5902	93,90%
	Var	385	6,10%
Çamaşır Yıkama Hizmeti	Yok	4312	68,60%
	Var	1975	31,40%
Mini Bar	Yok	4699	74,70%
	Var	1588	25,30%
Klima	Yok	3725	59,20%
	Var	2562	40,80%
Mikrodalga Fırın	Yok	6206	98,70%
	Var	81	1,30%

**Tablo 3.25.** Modelde Yer Alan Değişkenler ve Değişkenlere Ait Frekanslar (Devam)

Spa	Yok	5640	89,70%
	Var	647	10,30%
Ücretsiz Otopark	Yok	4661	74,10%
	Var	1626	25,90%

Bu modelde toplam 49 tane bağımsız değişkenden sadece 12 tanesi yer almıştır. Modeli test etmek için yapılan *complementary log-log* bağlantılı sıralı *lojistik regresyonun* varsayımlarından olan paralellik varsayımı anlamlılık testi sonuçları tabloda verilmektedir.

**Tablo 3.26.** Modelin Paralel Doğrular Testine Ait Sonuçları

Model	-2 Log Likelihood	Ki-Kare	s.d	p
Null Hypothesis	0,001			
General	0,001	0,001	96	1,000

Bağlantı Fonksiyonu: Complementary Log-log.

Bu sonuçlar sıralı lojistik regresyon analizinin varsayımlarından paralellik varsayımının sağlandığını göstermektedir ( $p > 0,05$ ). Ayrıca tablosu verilmeyen model uyumluluğu da anlamlı çıkmıştır ( $p < 0,05$ ; s.d=96,  $\chi^2=9474$ ). TabloAşağıdaki tabloda sadece anlamlı çıkan değişkenlere yer verilmiştir.

**Tablo 3.27.** Lojistik Regresyon Sonucu ve Değişkenlerin Anlamlılıkları

Bağımlı Değişken	$\beta$	Wald	p	Odds ( $e^\beta$ )	%95 Güven Aralığı	
[Ortalama Oda Fiyatı = 1]	2,839	0,001	0,972	17,0986		
[Ortalama Oda Fiyatı = 2]	3,890	0,002	0,962	48,9109		
Bağımsız Değişkenler					Alt Sınır	Üst Sınır
Yorum	0,001	175,541	0,001	1,0010	1,0009	1,0011
Yıldız/Puan	0,219	91,432	0,001	1,2448	1,1900	1,3022
[Tesis_Tipi = Otel]	0,813	5,658	0,017	2,2546	1,1534	4,4076
[Tesis_Tipi = Belirtilmemiş]	1,041	9,101	0,003	2,8320	1,4402	5,5689
[Tesis_Tipi = Apart Otel]	0,854	5,759	0,016	2,3490	1,1691	4,7198
[Tesis_Tipi = Villa]	1,531	11,996	0,001	4,6227	1,9439	10,9937
[Tesis_Tipi = Hostel]	.	.	.	.	.	.
[Deniz K1Y1S1 = 0]	-0,299	23,754	0,001	0,7415	0,6580	0,8357
[Deniz K1Y1S1 = 1]	.	.	.	.	.	.
[Kapalı Havuz = 0]	-0,191	4,222	0,040	0,8261	0,6885	0,9913
[Kapalı Havuz = 1]	.	.	.	.	.	.
[Çamaşır Yıkama Hizmeti = 0]	0,14	7,528	0,006	1,1502	1,0409	1,2712



**Tablo 3.27.** Lojistik Regresyon Sonucu ve Değişkenlerin Anlamlılıkları (Devam)

[Çamaşır Yıkama Hizmeti = 1]	.	.	.	.	.	.
[Mini Bar = 0]	-0,193	14,503	0,001	0,8244	0,7461	0,9112
[Mini Bar = 1]	.	.	.	.	.	.
[Klima = 0]	0,16	9,587	0,002	1,1735	1,0598	1,2994
[Klima = 1]	.	.	.	.	.	.
[Mikrodalga Fırın = 0]	-0,527	15,224	0,001	0,5903	0,4531	0,7692
[Mikrodalga Fırın = 1]	.	.	.	.	.	.
[Spa = 0]	-0,233	8,046	0,005	0,7921	0,6745	0,9303
[Spa = 1]	.	.	.	.	.	.
[Ücretsiz Otopark = 0]	0,135	8,039	0,005	1,1445	1,0438	1,2550
[Ücretsiz Otopark = 1]	.	.	.	.	.	.

Lojistik regresyonun yorumlanması doğrusal regresyondan farklı olmaktadır. Her değişkende bir referans değeri alınmaktadır. Tablo 3.27’de referans değerlerinin karşılıkları boş bırakılmıştır. Anlamlı bulunan değişkenler bu referans değerlerine göre yorumlanmaktadır. *Ücretsiz otopark değişkeninde*, referans değeri ücretsiz otoparkın olmasıdır. Bu sonuçlara göre ücretsiz otopark *bulunmayan tesisin* ortalama oda fiyatı, *bulunan tesisin* 1,14 kat daha yüksektir.

Dikkat edileceği üzere bütün ikili değişkenlerde referans değeri o donanımın veya hizmetin olmasıdır (1). Odds değeri buna göre yorumlanmaktadır. Bu durumda 1’den büyük *odds değerleri* ( $e^{\beta}$ ) ortalama oda fiyatının artmasına; 1’den küçük olan değerler ise azalmasına neden olmaktadır. Diğer anlamlı değişkenler yorumlandığında *spa hizmetinin olmaması* fiyatın *olmasına oranla* 0,79 kat artmasına, mikrodalga fırının olmaması olmasına oranla fiyatın 0,59 kat artmasına neden olmaktadır. Bu durumda *spa hizmeti* ve odalarda *mikrodalga fırının* olması fiyatın artması yönünde etki etmektedir. Bütün bunlar göz önüne alınarak değerlendirildiğinde bu modelde tesisin *deniz kıyısında* olması, *kapalı havuza*, *mini bara*, *mikrodalga fırına* ve *spa’ya* sahip olması ortalama oda fiyatını *pozitif* yönde etkilemektedir. Bağımsız değişkenlerin modeldeki belirsizliği ne oranda açıkladığına dair  $R^2$  sonuçları aşağıdaki tabloda verilmektedir.

**Tablo 3.28.** Regresyon Modeline Ait Pseudo  $R^2$  Değerleri

Cox and Snell	0,778
Nagelkerke	0,993
McFadden	0,985

Bağlantı fonksiyonu: Complementary Log-log.

Model uyumuna ait değerler incelendiğinde sonuçların iyi olduğu görülmektedir. Cox ve Snell testine göre modeldeki bağımsız değişkenler bağımlı değişkeni %77,8 oranında açıklamaktadır. Nagelkerke ve McFadden R<sup>2</sup> daha yüksek orandadır.

Sıralı lojistik regresyonda yeni bir veri ile tahminde bulunmak ta mümkündür. Fakat doğrusal regresyondan farklı olarak girilen değerlere göre bağımlı değişkendeki her gruba göre olasılık hesaplanır. Bu sonuçlara göre en yüksek olasılığa sahip olan grup tahmin grubudur. Aşağıda bunu hesaplamak için kullanılan fonksiyon yer almaktadır (Rawat, 2018; Arnold, 2007; Ford, 2015).

$$\text{logit} [P(Y \leq J)] = \alpha_j - \sum \beta_i X_i \quad (4)$$

Bu formülde J bağımlı değişkeni  $\alpha$  o gruba ait tahmin değerini (estimate)  $\beta$  Bu formüle göre yukarıdaki modelden elde edilen sonuçlar yerine koyulduğunda aşağıdaki denklem ortaya çıkmaktadır.

$$\begin{aligned} \text{logit} [P(Y \leq 1)] = & 2,839 - (Puan \times 0,219 + Yorum \times 0,001 \\ & + Tesis Tipi \beta Puanı \times 1 + Deniz Kıyısı \\ & \times -0,299 + Kapalı Havuz \times -0,191 \\ & + Çamaşır Yıkama Hizmeti \times 0,14 + Mini Bar \\ & \times -0,193 + Klima \times 0,16 + Mikrodalga Fırın \\ & \times -0,527 + Spa \times -0,233 + Ücretsiz Otopark \\ & \times 0,135) \end{aligned} \quad (5)$$

Eşitliğe göre bir örnek hesaplanacak olursa, 4 puan alan, 20 yoruma sahip, tesis tipi otel olan, deniz kıyısında olmayan, kapalı havuzu olan, çamaşır yıkama hizmeti sağlayan, mini barı olan, klimalı, mikrodalga fırını olmayan, spa'sı olmayan, ücretsiz otoparkı olan bir tesisin  $\leq 395,44$  grubunda veya oda fiyatının 395,44'den küçük olma ihtimali şu şekilde hesaplanacaktır.

$$\begin{aligned} \text{logit} [P(Y \leq 1)] = & 2,839 - (4 \times 0,219 + 20 \times 0,001 + 0,813 \\ & \times 1 + 0 \times -0,299 + 1 \times -0,191 + 1 \times 1,14 + 1 \\ & \times -0,193 + 1 \times 0,16 + 1 \times 0,16 + 0 \times -0,527 \\ & + 0 \times -0,233 + 1 \times 0,135) \end{aligned} \quad (6)$$

$$\text{logit} [P(Y \leq 1)] = 1,606 \quad (7)$$

$$P(Y \leq 1) = \exp(1,606) / (1 + \exp(1,606)) \quad (8)$$

$$P(Y \leq 1) = P(Y = 1) = 0,83 \quad (9)$$

Bu sonuçlara göre tesisin  $\leq 395,44$  grubunda olma olasılığı %83'dir. Aynı tesisin ikinci grupta olma yani ortalama oda fiyatının 395,45 – 812,77 TL aralığında olma olasılığı ise şöyledir.

$$\begin{aligned} \text{logit } [P(Y \leq 2)] &= 3,890 - (4 \times 0,219 + 20 \times 0,001 + 0,813 \\ &\times 1 + 0 \times -0,299 + 1 \times -0,191 + 1 \times 1,14 + 1 \\ &\times -0,193 + 1 \times 0,16 + 1 \times 0,16 + 0 \times -0,527 \\ &+ 0 \times -0,233 + 1 \times 0,135) \end{aligned} \quad (10)$$

$$\text{logit } [P(Y \leq 2)] = 2,657 \quad (11)$$

$$P(Y \leq 2) = \exp(2,657) / (1 + \exp(2,657)) \quad (12)$$

$$P(Y \leq 2) = 0,944 \quad (13)$$

İlk gruptan farklı olarak ikinci grup olasılığı hesaplanırken ikinci grubun olasılığından birinci grubun olasılığı çıkartılmaktadır.

$$P(Y = 2) = P(Y \leq 2) - P(Y \leq 1) = 0,944 - 0,85 \quad (14)$$

$$P(Y = 2) = 0,084 \quad (15)$$

Bu sonuca göre verilen özellikteki bir tesisin ikinci grupta olma olasılığı %8,4 olarak hesaplanacaktır.

Yöntemin R dilindeki hali aşağıda verilmektedir. Ayrıca istatistik paket programlarının bazılarında (IBM SPSS v.23 Analyze/Regression/Ordinal/Output/Saved Variables/Estimated response probabilities) bu tahmin yapılabilmektedir.

Şekil 3.3 incelendiğinde 34'ncü satırdan itibaren tahmini yapılmak istenen konaklama tesisinin özellikleri girilmektedir. Örnek tesis, 4 puan/yıldız, 20 yorum almış, tesis tipi 1 olan, 41'nolu şehirde bulunan ve deniz kıyısında olmayan diğer özellikleri devamında verilen özellikte bir tesistir. 48'nci satırdaki komut çalıştırıldığında grup tahmini yapılacaktır.

```

1 # polr fonksiyonu MASS kütüphanesinde
2 library(MASS)
3
4 # bağımlı değişken faktör türünde olmalı
5 veri$NormDistGr <- as.factor(veri$NormDistGr)
6
7 # bağımlı ve bağımsız değişkenler modele dahil ediliyor
8 model_iki <- polr(veri$NormDistGr~
9
10             Puan +
11             Yorum +
12             Tesis_Tipi +
13             Sehir +
14             DenizKıyısı +
15             Klima +
16             MiniBar +
17             ÜcretsizOtopark +
18             MikrodalgaFırın +
19             CamasirYikamaHizmeti +
20             Spa +
21             KapaliHavuz
22
23             , data = veri, Hess = TRUE)
24 # t değerleri hesaplanıyor
25 ozet_tablo <- coef(summary(model_iki))
26
27 # p anlamlılık değerleri hesaplanıyor
28 pdeg <- pnorm(abs(ozet_tablo[, "t value"]), lower.tail =
29 FALSE) * 2
30 ozet_tablo <- cbind(ozet_tablo, "p value" = round(pdeg,3))
31
32 #test verisi giriliyor
33 test <- data.frame(
34     "Puan"=4,
35     "Yorum"=20,
36     "Tesis_Tipi"=1,
37     "Sehir"=41,
38     "DenizKıyısı"=0,
39     "Klima"=1,
40     "MiniBar"=1,
41     "ÜcretsizOtopark"=0,
42     "MikrodalgaFırın"=1,
43     "Spa"=0,
44     "CamasirYikamaHizmeti"=1
45     "KapaliHavuz"=1
46 )
47
48 #olasılık hesaplanıyor
49 round(predict(model_iki, test, type = "p"), 3)

```

Şekil 3.3. R'da Modele İlişkin Tahminin Yapılması

## SONUÇ VE TARTIŞMA

Bu araştırmada otel oda fiyatlarına etki eden değişkenler web madenciliği yöntemi ile elde edilen verilerle belirlenmeye çalışılmıştır. Bu amaçla, geliştirilen bir algoritma yardımı ile, birçok acentenin çevrimiçi oda fiyatlarını bulunduran Tripadvisor.com sitesi üzerinde web madenciliği yapılmış, Türkiye'ye ait veriler elde edilmiştir.

Algoritma, 2018 Mayıs ve Eylül ayları dâhil olmak üzere her hafta çalıştırılarak 19 haftalık toplam 6558 konaklama tesisine ait fiyat, şehir, bölge, donanım ve sağladıkları çeşitli hizmetlere ait bilgiler elde edilmiştir. Elde edilen 49 bağımsız değişken ile *ortalama oda fiyatı* bağımlı değişkeni açıklanmaya çalışılmıştır. En iyi regresyon modelini bulmak için *aşamalı regresyon (stepwise regression)* analizi yapılmıştır. Yapılan aşamalı regresyon sonucunda *Yorum, Şehir, Yıldız, Spor Salonlu/Spor Odalı Fitness Merkezi, Tesis Tipi, Deniz Kıyısı, Kapalı Havuz, Çamaşır Yıkama Hizmeti, Mini Bar, Klima, Mikrodalga Fırın, Spa, Ücretsiz Otopark, Bölge* değişkenlerini içeren modelin en yüksek  $R^2$  değerine sahip olan model olduğu görülmüştür (Bkz: Tablo 3.22).

Bağımlı değişken gruplandırılmış, bu gruplamalardan en yüksek  $R^2$  değerine sahip olan model (ordinal logistic regression, OLOGREG) içerisinde kullanılmıştır. Anlamlı değişkenleri **Tablo**'de verilen, *regresyon modelinin* bağımlı değişkeni açıklama oranı en az %77,8 çıkmıştır. Oluşturulan model ile örnek bir konaklama tesisin hangi fiyat aralığında olacağı hesaplanmış, Şekil 3.3'de bunun nasıl yapılacağı *R Programlama Dili*'nde gösterilmiştir.

Ayrıca regresyon modelleri dışında çeşitli hipotez testleri ile değişkenlerin anlamlılıkları test edilmiştir. *Ortalama oda fiyatlarının* konaklama tesisine göre anlamlı farklılık gösterip göstermediği test edilmiş, testin sonucu anlamlı çıkmıştır ( $(8,6549) = 73,47; p < 0,05$ ). Farklılıkların hangi gruplar arasında olduğu **Tablo** 'da verilmiştir. Tesisin bulunduğu coğrafi bölgeye göre ortalama oda fiyatının anlamlı farklılık gösterip göstermediği test edilmiş, tek faktörlü varyans analizi sonucu anlamlı çıkmıştır ( $(F(6,6551) = 39,27; p < 0,05)$ ). Bu test için hangi bölgeler arasında farklılık olduğu **Tablo** 'de verilmiştir. Benzer şekilde konaklama tesisinin bulunduğu ile göre ortalama oda fiyatları anlamlı farklılık göstermiştir ( $(F(77,6480) = 10,33; p < 0,05)$ ). Ortalama oda fiyatı en yüksek beş il sırası ile Karaman, Trabzon, Antalya, Muğla ve Giresun olurken, ortalama oda fiyatı en düşük olan beş il ise Bayburt, Adıyaman, Malatya, Eskişehir ve Erzincan olmuştur.

Deniz kıyısında olup olmamasına göre tesisler sınıflandırılmış, ortalama oda fiyatlarının farklılığı test edilmiştir. Test sonucuna göre konaklama tesislerinin deniz kıyısında olup olmamaları, ortalama oda fiyatları bakımından anlamlı farklılık göstermiştir ( $t_{(6556)} = -12,510$ ;  $p > 0,05$ ). Deniz kıyısında olan otellerin ortalama oda fiyatları (Ort: 566,75) deniz kıyısında olmayanlara (Ort: 378,46) nazaran daha yüksektir.

Toplanan veri içerisinde daha önce konaklama tesisini ziyaret eden konukların verdikleri puanları içeren bir değişken bulunmaktadır. *Yıldız/Puan* olarak adlandırılan bu değişken ile ortalama oda fiyatı arasındaki ilişkinin tespiti için yapılan korelasyon testinde anlamlı ve pozitif yönlü orta-düşük ilişki görülmektedir ( $r(6287)=0,245$ ;  $p<0,01$ ).

Çalışmanın ikinci bölümünde veri elde edilirken kullanılan yöntem üç farklı yazılımdan örneklerle anlatılmıştır. Bu yöntemler veri elde etmenin zor olduğu klasik veri toplama yöntemlerine hızlı ve verimli bir alternatif olmaktadır. Araştırma sonuçlarının da gösterdiği üzere 6558 konaklama tesisine ait 49 adet değişkene ait veri ve toplam 5 aylık *oda fiyatı verisi* hiç de kolay olmayacak klasik anket veya görüşme gibi yöntemlere kıyasla, web madenciliği ile kısa süre ve düşük maliyetle elde edilmiştir.

Kullanılan web madenciliği yöntemi araştırmalarda veri sağlamak için kullanılacak yeni ve alternatif bir yöntemdir. Bu yöntem ile elde edilen veri, klasik istatistikî yöntemler ile test edilebileceği gibi, veri madenciliği yöntemleri ile de işlenebilmektedir. Ayrıca son yılların güncel metodlarından olan büyük veri, yapay zekâ, sinir ağları gibi modeller için de gerekli olan büyük hacimli veri bu yolla elde edilebilmektedir.

Araştırmada konaklama tesislerinin iki kişilik bir gecelik oda fiyatları her hafta periyodik olarak algoritma ile elde edilmiştir. Bu veri ülke, bölge, şehir, tesis tipi gibi pek çok değişkene göre bu fiyatların değişimini de vermektedir. Fakat bu çalışmada fiyatlardaki değişimlere ait sonuçlar verilmemiştir. Bu haftalık fiyatlar elde edilen fiyat sayısına bölünerek ortalama oda fiyatı elde edilmiştir. Analizlerdeki bağımlı değişkenler bu ortalama oda fiyatlarından oluşmaktadır. Ortalama oda fiyatlarını elde etmek daha uzun zaman alan bir yöntem olsa da daha doğru sonuçların elde edilmesini sağlamıştır.

Kullanılan yöntem gelecekte birçok bilim dalında veya iş araştırmalarında veri elde etmede kullanılabilir maliyeti düşük, hızlı ve verimli bir yöntemdir. Bu duruma örnek teşkil etmesi açısından, bu çalışma, geçmişte meta tarama veya anket yöntemi ile yapılan otel oda fiyatlarını etkileyen faktörleri belirlemeye yönelik (Goldberg, Green ve Wind,

1984; Wu, 1999; Thrane, 2007; Wei-Ting, Jui-Kou ve Fei-Ching, 2010; Taylor, 1995; Bull, 1994) arařtırmaların onlarca katı veri kullanılarak yapılmıřtır. Elde edilen sonuçlar Türkiye'nin otel oda fiyatları hakkında yapılan en kapsamlı çalıřmalardan biri olmuřtur. Web madencilięi kullanılmadan bu çapta bir arařtırmayı yapmak, yüksek maliyetli, uzun süren, büyük iřgücü gerektiren bir çalıřma olacaktır.

Web madencilięi yöntemi web üzerinde yayınlanmış, erişime açık her türlü veri için kullanılabilir bir yöntemdir. Kullanılan algoritma başka bir arařtırmada hedef web sitesinin kaynak kodlarına uyarlanarak istenilen veri çekilebilmektedir. Üstelik bu iřlem internet hızına baęlı olarak günümüz teknolojisi ile oldukça hızlı gerçekleřecektir. Arařtırma yöntemlerinde güvenilir veri elde etmek, geçmiřten günümüze yoğun emek isteyen zorlu bir iřtir. Web madencilięi, geçmiřte veri elde etmede sorun yařayan pek çok arařtırma için önemli kolaylık saęlayacaktır. Sosyal aęlar, e-ticaret siteleri, bloglar gibi web ile ortaya çıkan pek çok veri kaynaęı, web madencilięi ile birlikte, arařtırmacılara yeni çalıřmaların önünü açacaktır.

Daha önce de bahsedildięi gibi, verinin elde edilmesinde, programlama dilleri, kısmen sorgu dilleri kullanan yöntemler ve tamamen otomatik yöntemler kullanılabilir. Burada arařtırmacının karřısına řöyle bir seęenek çıkmaktadır. Programlama dilleri dięer yöntemlerden daha fazla bilgi ve uzmanlık gerektirirken, daha detaya inilebilen esnek bir yapıya sahiptir. Sorgu dili, HTML, CSS bilgisi gerektiren yöntemler, bazı yazılımlar ve tarayıcı eklentileri ile kullanılan, programlama dillerinden daha basit fakat az da olsa bilgi gerektiren yöntemlerdir. Web madencilięi için geliřtirilen otomatik yazılımları kullanan yöntemler dięer iki yöntemden daha basit, öğrenmesi kolay, kod bilgisi gerektirmeyen fakat çoęu zaman arařtırmacıya esneklik saęlamayan yöntemlerdir. Bu yazılımları kullanılarak yapılan web madencilięinin başarısı kullanılan yazılımın becerisi ile doęru orantılıdır. Kullanıcı bu kıstasları göz önünde bulundurarak tercihini yapmalıdır.

Bu çalıřma ileride yapılacak olan arařtırmalara iki kategoride katkı saęlamayı amaçlamaktadır. Bunlardan ilki otel oda fiyatları konusunda elde edilen bulgulardır. Bu konuda daha önce yapılan çalıřmalar (Kesbiç, Bahar, Baldemir ve İnci, 2011; Iřık ve Bilici, 2017; Papatheodorou, 2002) sadece bölgesel temelde olmuřtur. Bu arařtırma farklı olarak bütün ülkeyi kapsayacak řekilde yapılmıřtır. Böylece Türkiye'de otel oda

fiyatlarını etkileyen faktörlerin belirlenmesine yönelik bir model ve değişkenler ortaya konulmuştur. Ayrıca daha önce otel oda fiyatlarını etkileyen faktörlerin belirlenmesi konusunda yapılan çalışmalardan (Carvell ve Herrin, 1990; Bull, 1994; Taylor, 1995; Papatheodorou, 2002; Aguiló, Alegre ve Sard, 2003; Espinet, Saez, Coenders ve Fluvia, 2003; Thrane, 2005; Thrane, 2007; Hung, Shang ve Wang, 2010; Zhang, Zhang, Lu, Cheng ve Zhang, 2011; Pawlicz ve Napierala, 2017) farklı olarak, yeni bir veri elde etme yöntemi kullanılmıştır. Böylece bir örneklem üzerinde yapılan bu çalışmaların aksine veri setinin neredeyse tamamı kullanılmıştır.

Yöntem uygulanırken az da olsa çeşitli problemlerle karşılaşılabilir. Örneğin, her hafta çalıştırılan algoritma bazı problemlere neden olmuştur. Bunlardan en önemlisi hedef sitenin kaynak kodlarının değişmesidir. Kaynak kodlar değiştiğinde buna bağımlı olarak oluşturulan sorguların da değişmesi gerekmektedir. Bu nedenle indirmeden önce küçük bir örnek veri ile algoritmanın doğru çalışıp çalışmadığı test edilmiştir. Kaynak kodlar uzun zaman dilimlerinde değiştiğinden bu tarz bir algoritmanın tamamen otomatik olması şu an için mümkün görünmemektedir.

Bu çalışmanın ileride yapılacak olan araştırmalara sağlamayı amaçladığı ikinci katkı veri elde etme konusudur. Daha önce de bahsedildiği üzere, web muazzam büyüklükte veri içeren, sürekli genişleyen bir veri kaynağıdır. Bu veri yarı sistemli veya dağınık halde web üzerinde yer almaktadır. Bu çalışmada, web üzerinde yarı sistemli halde bulunan verinin nasıl elde edildiği örnek bir algoritma gösterilmiştir. Bu çalışmadaki örnekte veri üzerinde çeşitli hipotez testleri ve istatistikî yöntemler uygulanmış, elde edilen bulgular verilmiştir. Bu yöntem ile elde edilen veri üzerinde doğal dil işleme (NLP), veri madenciliği, yapay sinir ağları, derin öğrenme, yapay zekâ gibi pek çok yöntem uygulanabilmektedir. Daha önce bu yöntemlerin ihtiyaç duyduğu verinin elde edilmesi güçken, web ve web madenciliği bu konuda önemli bir araç olacaktır.



## KAYNAKÇA

- Abbate, J. (1999). *Inventing the Internet*. Cambridge, US.: MIT Press.
- Abbot, P. ve Lewry, S. (1999). *Front Office. Procedures, Social Skills, Yield and Management* (2 b.). Oxford, UK: Butterworth-Heinemann.
- Abrate, G., Capriello, A. ve Fraquelli, G. (2011). When Quality Signals Talk: Evidence From The Turin Hotel Industry. *Tourism Management*(32), 912-921.
- Aggarwal, A. (2016). *Managing Big Data Integration in the Public Sector*. Hershey, US.: IGI Global.
- Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*. Cham, Switzerland: Springer Nature Switzerland AG. doi:<https://link.springer.com/book/10.1007%2F978-3-319-94463-0>
- Agrawal, R., Imielinski, T. ve Swami, A. (1993). Mining Association Rules between Sets of Items in Large Databases. *ACM SIGMOD Record*, 22(2), 487-499. doi:10.1145/170036.170072
- Aguiló, E., Alegre, J. ve Sard, M. (2003). Examining the Market Structure of The German and UK Tour Operating Industries Through An Analysis of Package Holiday Prices. *Tourism Economics*, 9(3), 255-278.
- Ahipaşaoğlu, H. S. (2002). *Seyahat Acentacılığı ve Tur Operatörlüğü*. (İ. Ünlü, Dü.) Eskişehir: Anadolu Üniversitesi.
- Aiello, M. (2018). *The Web Was Done by Amateurs: A Reflection on One of the Largest Collective Systems Ever Engineered*. Cham, Switzerland.: Springer International Publishing AG.
- Akbaş, E. (2018, 11 24). [www.emo.org.tr](http://www.emo.org.tr). 05 11, 2019 tarihinde Derin Yapay Sinir Ağları ve Derin Öğrenme'ye Kısa Bir Giriş: [http://www.emo.org.tr/ekler/b6d1307298e4886\\_ek.pdf](http://www.emo.org.tr/ekler/b6d1307298e4886_ek.pdf) adresinden alındı
- Akça, Y. (2016). Türkiye'nin Kalkınma Planlarında Turizm Politikası. *International Conference on Eurasian Economies* (s. 721-726). Kaposvár, Hungary: Beykent University Publications. <https://www.avekon.org/proceedings/avekon07.pdf> adresinden alındı
- Akpınar, H. (2014). *Data: Veri Madenciliği Veri Analizi*. İstanbul, Türkiye.: Papatya Yayıncılık Eğitim.
- Albustanlıoğlu, T. (2009). *Seyahat Acenteciliği Ders Notları*. 03 28, 2019 tarihinde <http://acikerisim.baskent.edu.tr/bitstream/handle/11727/1921/ff9472e56b30a188a7b69f78aad0e83a.pdf> adresinden alındı
- Andersson, D. E. (2008). Hotel attributes and hedonic prices: an analysis of internet-based transactions in Singapore's market for hotel rooms. *Ann Reg Sci*, 229-240. doi:10.1007/s00168-008-0265-4
- Andrew, O. (2018, 08 28). [www.textrequest.com](http://www.textrequest.com). 04 29, 2019 tarihinde The History and Evolution of the Smartphone: 1992-2018: <https://www.textrequest.com/blog/history-evolution-smartphone/> adresinden alındı

- Archer, J. ve Syrratt, G. (2003). *Manual of Travel Agency Practice* (3 b.). Burlington: Butterworth-Heinemann, Elsevier.
- Arnold, D. (2007). *Probit and Complementary Log-Log Models*. 02 20, 2019 tarihinde stat.ualberta.ca: <http://www.stat.ualberta.ca/~kcarrier/STAT562/daniel.pdf> adresinden alındı
- Aslam, S. (2018, 06 24). *omnicoreagency.com*. 04 28, 2019 tarihinde YouTube by the Numbers: Stats, Demographics ve Fun Facts: <https://www.omnicoreagency.com/youtube-statistics/> adresinden alındı
- Ateşoğlu, İ. ve Bayraktar, S. (2011). Ağızdan Ağıza Pazarlamanın Turistlerin Destinasyon Seçimindeki Etkisi. *Uluslararası Yönetim ve İşletme Dergisi*, 7(14), 95-108. <http://ijmeb.org/index.php/zkesbe/article/view/22> adresinden alındı
- Bakır, M. (2016, 03 09). *www.muminbakir.com*. 04 27, 2019 tarihinde Root DNS (Kök DNS): [http://www.muminbakir.com/root\\_dns.html](http://www.muminbakir.com/root_dns.html) adresinden alındı
- Bayer, Y. (2012, 07 08). *TURING'in malvarlığı iştah kabartıyor*. 04 11, 2019 tarihinde hurriyet.com.tr: <http://www.hurriyet.com.tr/turing-in-malvarligi-istah-kabartiyor-20938773> adresinden alındı
- Beech, J. ve Chadwick, S. (2006). *The Business of Tourism Management*. Harlow: Financial Times Prentice Hall.
- Berners-Lee, T., Cailliau, R. ve Groff, J.-F. (1992). The world-wide web. *Computer Networks and ISDN Systems*, 454-459. <https://web.stanford.edu/class/cs344g/www-1992.pdf> adresinden alındı
- Berrar, D. P., Dubitzky, W. ve Granzow, M. (Dü). (2003). *A Practical Approach to Microarray Data Analysis*. New York, US.: Kluwer Academic Publishers.
- Bhatia, A. K. (2006). *The Business of Tourism: Concepts and Strategies*. New Delhi: Sterling Publishers Pvt. Ltd. 02 2018, 22 tarihinde <https://books.google.com.tr/books?id=Lp5I2UJdGiUC> adresinden alındı
- Birdir, K. ve Akgöl, Y. (2015). Gastronomi Turizmi ve Türkiye'yi Ziyaret Eden Yabancı Turistlerin Gastronomi Deneyimlerinin Değerlendirilmesi. *İşletme ve İktisat Çalışmaları Dergisi*, 3(2), 57-68. [http://www.isletmeiktisat.com/index.php/iicd/article/view/42/pdf\\_23](http://www.isletmeiktisat.com/index.php/iicd/article/view/42/pdf_23) adresinden alındı
- Björk, P. ve Kauppinen-Räsänen, H. (2017). A destination's gastronomy as a means for holiday well-being. *British Food Journal*, 119(7), 1578-1591. doi:10.1108/BFJ-09-2016-0394
- Boehmke, B. (2015, 11 21). *Scraping HTML Tables*. 03 11, 2019 tarihinde bradleyboehmke.github.io: <http://bradleyboehmke.github.io/2015/12/scraping-html-tables.html> adresinden alındı
- Bolton, D. (2019, 01 25). *www.dice.com*. 05 13, 2019 tarihinde XML vs. JSON: What's the Difference for Developers?: <https://insights.dice.com/2019/01/25/xml-vs-json-difference-developers/> adresinden alındı
- Boulton, C. (2015, 10 18). *www.cio.com*. 05 01, 2019 tarihinde Weather Company forecasts more big data for IBM Watson Analytics:

<https://www.cio.com/article/3006300/weather-company-forecasts-more-big-data-for-ibm-watson-analytics.html> adresinden alındı

- Boyne, S., Hall, D. ve Williams, F. (2003). Policy, Support and Promotion for Food-Related Tourism Initiatives. *Journal of Travel & Tourism Marketing*, 14(3-4), 131-154. doi:10.1300/J073v14n03\_08
- Bridgwater, A. (2015, 10 27). *www.forbes.com*. 05 01, 2019 tarihinde IBM Works With The Weather Company To Track Earth's Big Data Atmosphere: <https://www.forbes.com/sites/adrianbridgwater/2015/10/27/ibm-works-with-the-weather-company-to-track-earths-big-data-atmosphere> adresinden alındı
- Bull, A. O. (1994). Pricing A Motel's Location. *International Journal of Contemporary Hospitality Management*, 6(6), 10-15.
- Buluswar, M. (2016). *www.mckinsey.com/*. 05 01, 2019 tarihinde How companies are using big data and analytics: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-companies-are-using-big-data-and-analytics> adresinden alındı
- Carvell, S. A. ve 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
- Castro, C., Ferreira, F. A. ve Ferreira, F. (2016). Trends in hotel pricing Identifying guest value hotel attributes using the cases of Lisbon and Porto. *Worldwide Hospitality and Tourism Themes*, 8(6), 691-698. doi:10.1108/WHATT-09-2016-0047
- Chadha, G., Das, P. ve Karnin, Z. (2018, 10 08). <https://aws.amazon.com>. 05 09, 2019 tarihinde K-means clustering with Amazon SageMaker: <https://aws.amazon.com/tr/blogs/machine-learning/k-means-clustering-with-amazon-sagemaker/> adresinden alındı
- Chakrabarti, S. (2003). *Mining the Web: Discovering Knowledge from Hypertext Data*. San Francisco: Elsevier.
- Chi, C. (2019). *blog.hubspot.com*. 04 28, 2019 tarihinde 51 YouTube Stats Every Video Marketer Should Know in 2019: <https://blog.hubspot.com/marketing/youtube-stats> adresinden alındı
- Civelek, M. E. (2009). *İnternet Çağı Dinamikleri*. İstanbul, Türkiye.: BETA Basım Yayım A.Ş.
- Clements, A. (2014). *Computer Organization & Architecture: Themes and Variations*. Stamford, US.: Cengage Learning.
- Cook, G. N. ve Billig, B. G. (2017). *Airline Operations and Management: A Management Textbook*. Abingdon, UK: Routledge Taylor & Francis Group.
- Corea, F. (2019). *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science*. Cham, Switzerland: Springer Nature Switzerland.
- CPU performance*. (2019). 03 06, 2019 tarihinde <https://asteroidsathome.net>: [https://asteroidsathome.net/boinc/cpu\\_list.php](https://asteroidsathome.net/boinc/cpu_list.php) adresinden alındı

- Çalışkan, S., Kutlu, B. ve Kımlıoğlu, H. (2013). The Contribution of Online Travel Agencies to the Profitability of Hotels: Case of Turkey. *Interdisciplinary Journal of Research in Business*, 3(3), 8-18.
- Çelik, A. (2018). *Turizm İstatistikleri Raporu*. 04 10, 2019 tarihinde Türkiye Otelciler Birliği: <http://www.turob.com/Files/Kasim%202018%20İstanbul.pdf> adresinden alındı
- Çelik, H. (2006). *1980-2000 Yılları Arasında Türkiye’de Uygulanan Neo Liberal Politikaların İşçi Sınıfına Etkileri. (Yayınlanmamış Yüksek Lisans Tezi)*. Denizli: Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü.
- Davenport, T. H. ve Patil, D. (2012, 10). *Harvard Business Review*. 05 04, 2019 tarihinde Data Scientist: The Sexiest Job of the 21st Century: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century> adresinden alındı
- Delgado, P. (2018, 01 26). *Has TripAdvisor become an OTA?* 08 07, 2018 tarihinde Mirai.com: <https://www.mirai.com/blog/has-tripadvisor-become-an-ota/> adresinden alındı
- Deshmukh, A. V. (2005). *Microcontrollers: Theory and Applications*. New Delhi, India: Tata McGraw-Hill Publishing Company.
- Dev, C. S. ve O'Connor, P. (2015). Challenge the middlemen? A Hotelier Considers Backing an Industry Response to Online Travel Agencies. *Harvard Business Review*, 93(12), 119-123. doi:edsgcl.437630100
- Devashish, D. (2011). *Tourism Marketing*. Delhi, India: Pearson Education India.
- Dinçer, M. Z. ve Çetin, G. (2015). Kalkınma Planlarında Turizm. D. Küçükaltan, H. Çeken ve Ş. O. Mercan (Dü) içinde, *Değişik Perspektifleriyle Turizm Politikası ve Planlaması* (s. 171-192). Ankara: Detay Yayınları.
- Donahoe, L. ve Hartl, M. (2019). *www.learnenough.com*. 05 13, 2019 tarihinde Learn Enough CSS & Layout to Be Dangerous: <https://www.learnenough.com/css-and-layout-tutorial/introduction> adresinden alındı
- Dontha, R. (2018, 12 19). *Digital Transformation*. 05 04, 2019 tarihinde Data Mining Steps: <https://digitaltransformationpro.com/data-mining-steps/> adresinden alındı
- Dostálek, L. ve Kabelová, A. (2006). *Understanding TCP/IP: A Clear and Comprehensive Guide to TCP/IP Protocols*. Birmingham, UK.: Packt Publishing Ltd.
- Doswell, R. (1997). *Tourism: How Effective Management Makes the Difference*. Boston: Butterworth-Heinemann.
- Dreyfus, G. (2005). *Neural Networks: Methodology and Applications*. Berlin, Germany: Springer-Verlag.
- DuCastel, B. ve Jurgensen, T. (2008). *Computer Theology: Intelligent Design of the World Wide Web*. Austin, US.: Midori Press LLC.
- Dudley, D. (2018, 11 27). *www.flauntdigital.com*. 04 29, 2019 tarihinde The Evolution of Mobile Phones: 1973 to 2019: <https://flauntdigital.com/blog/evolution-mobile-phones/> adresinden alındı

- Dunham, M. H. (2003). *Data Mining: Introductory and Advanced Topics*. Upper Saddle River, US.: Pearson Education India.
- Eddy, D. (2016, 06 07). *dcsplus.net*. 04 23, 2016 tarihinde Customer Reviews Bring Benefits to Travel Agency Websites: <http://www.dcsplus.net/blog/customer-reviews-bring-benefits-to-travel-agency-websites> adresinden alındı
- Edgell, D. L., Allen, M. D., Smith, G. ve Swanson, J. (2008). *Tourism Policy and Planning: Yesterday, Today and Tomorrow*. Amsterdam: Butterworth-Heinemann.
- Egara, E., Kerami, D., Wiryana, I. ve Maulana Kusuma, T. (2017). Researchgate data analysis to measure the strength of Indonesian research. *Far East Journal of Electronics and Communications*, 17(5), 1177-1183. doi:10.17654/EC017051177
- Elliott, H. (2013, 06 10). *10 Reasons To Visit Turkey Now*. 03 30, 2019 tarihinde forbes.com: <https://www.forbes.com/sites/hannahelliott/2013/07/10/10-reasons-to-visit-turkey-now> adresinden alındı
- Elliott, J. (2002). *Tourism: Politics and Public Sector Management*. London: Routledge.
- Ellis, J. (2019). *www.comms-express.com*. 04 27, 2019 tarihinde Internet Protocol Version 4 Addresses: <https://www.comms-express.com/infozone/article/ipv4/> adresinden alındı
- Engard, B. (2016, 11 22). *www.jefferson.edu*. 04 25, 2019 tarihinde FROM ARPANET TO WORLD WIDE WEB: AN INTERNET HISTORY TIMELINE: <https://online.jefferson.edu/communications/internet-history-timeline/> adresinden alındı
- Espinet, J. M., Saez, M., Coenders, G. ve Fluvia, M. (2003). Effect on Prices of the Attributes of Holiday Hotels: A Hedonic Prices Approach. *Tourism Economics*, 9(2), 1-13.
- Evcin, E. (2017). Basın-Yayın ve Turizm Bakanlığı Döneminde Turistik Tanıtım ve Propaganda (1957-1963). *Atatürk Dergisi*, 6(2), 41-91.
- Fayyad, U., Piatetsky-Shapiro, G. ve Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence*, 37-54.
- Filiz, F. (2017, 07 23). *www.medium.com*. 05 10, 2019 tarihinde Yapay Sinir Ağları: <https://medium.com/@fahrettinf/4-1-1-yapay-sinir-aglari-86f0440f85c2> adresinden alındı
- Fisher, T. (2019, 01 11). *What Is Bandwidth?* 03 08, 2019 tarihinde Lifewire: <https://www.lifewire.com/what-is-bandwidth-2625809> adresinden alındı
- Fletcher, J. E., Fyall, A., Gilbert, D. ve Wanhill, S. (2018). *Tourism: Principles and Practice*. Harlow, UK: Pearson Education Limited.
- Ford, C. (2015, 10 05). *Fitting and Interpreting a Proportional Odds Model*. 02 20, 2019 tarihinde University of Virginia Library: <https://data.library.virginia.edu/fitting-and-interpreting-a-proportional-odds-model/> adresinden alındı
- Freely, J. (2011). *A History of Ottoman Architecture*. Southampton, UK: WIT Press.

- Frumusanu, A. (2018, 03 26). *www.anandtech.com*. 04 28, 2019 tarihinde The Samsung Galaxy S9 and S9+ Review: Exynos and Snapdragon at 960fps: <https://www.anandtech.com/show/12520/the-galaxy-s9-review/6> adresinden alındı
- Fung, K. (2013). *Numbersense: How to Use Big Data to Your Advantage*. New York, US.: McGraw Hill Professional.
- Gaggioli, A. (2015, 06 18). *Cloudbeds*. 04 23, 2016 tarihinde Analysis of Major Online Travel Agencies – OTAs: <https://www.cloudbeds.com/articles/analysis-of-major-online-travel-agencies-otas/> adresinden alındı
- Gaggioli, A. (2017, 08 17). *Analysis of Major Online Travel Agencies – OTAs – Update August 2017*. 08 07, 2018 tarihinde Cloudbeds: <https://www.cloudbeds.com/articles/analysis-major-online-travel-agencies-otas-update-august-2017/> adresinden alındı
- Galetto, M. (2015, 12 11). *www.ngdata.com*. 05 04, 2019 tarihinde What Is Data Mining?: <https://www.ngdata.com/what-is-data-mining/> adresinden alındı
- García, S., Luengo, J. ve Herrera, F. (2015). *Data Preprocessing in Data Mining*. Cham, Switzerland: Springer International Publishing.
- Ghanem, H. (2016). *The Arab Spring Five Years Later: Toward Greater Inclusiveness*. Washington, US: Brookings Institution Press. 03 30, 2019 tarihinde <https://www.jstor.org/stable/10.7864/j.ctt1657tv8> adresinden alındı
- Gherasim, M. E. ve Tanese, G. (2012). The Evolution and Inter-Dependence of Tourism with the Environment. *Contemporary Readings in Law and Social Justice*, 4(1), 458-470.
- Gildred, J. (2018, 03 29). *www.cloudwards.net*. 04 28, 2019 tarihinde History of the Hard Drive: <https://www.cloudwards.net/history-of-the-hard-drive/> adresinden alındı
- Goeldner, C. R. ve Ritchie, R. B. (2006). *Tourism; Principles, Practices, Philosophies* (10 b.). New Jersey: John Wiley & Sons, Inc.
- Goh, D. H. ve Ang, R. P. (2007). An introduction to association rule mining: An application in counseling and help-seeking behavior of adolescents. *Behavior Research Methods*, 39(2), 259-266.
- Gowrishankar, S. ve Veena, A. (2019). *Introduction to Python Programming*. Boca Raton: Taylor & Francis Group.
- Grant, L. (2016, 02 04). *Everything You Need to Know About All-Inclusive Resorts*. 08 09, 2018 tarihinde Oyster.com: <https://www.oyster.com/articles/51469-everything-you-need-to-know-about-all-inclusive-resorts/> adresinden alındı
- Grus, J. (2019). *Data Science from Scratch* (2 b.). Sebastopol, US.: O'Reilly Media, Inc.
- Guo, X., Zheng, X., Ling, L. ve Yang, C. (2014). Online Coopetition Between Hotels and Online Travel Agencies: From The Perspective of Cash Back After Stay. *Tourism Management Perspectives*, 12, 104-112. doi:10.1016/j.tmp.2014.09.005
- Gürel, M. Ö. (2018). *Mid-Century Modernism in Turkey: Architecture Across Cultures in the 1950s and 1960s*. New York, US: Routledge. 03 30, 2019 tarihinde <https://books.google.com.tr/books?id=ivmPCgAAQBAJ> adresinden alındı

- Gyódi, K. (2019). Airbnb in European cities: Business as usual or true sharing economy? *Journal of Cleaner Production*, 221(1), 536-551. doi:<https://doi.org/10.1016/j.jclepro.2019.02.221>
- Ha, P. (2010, 10 25). *www.time.com*. 04 29, 2019 tarihinde All-TIME 100 Gadgets: [http://content.time.com/time/specials/packages/article/0,28804,2023689\\_2023708\\_2023648,00.html](http://content.time.com/time/specials/packages/article/0,28804,2023689_2023708_2023648,00.html) adresinden alındı
- Haigh, T., Priestley, P. M., Priestley, M. ve Rope, C. (2016). *ENIAC in Action: Making and Remaking the Modern Computer*. Cambridge, US.: The MIT Press.
- Hall, C. M. ve Mitchell, R. (2006). Gastronomy, Food and Wine Tourism. D. Buhalis ve C. Costa (Dü) içinde, *Tourism Business Frontiers: Consumers, Products and Industry* (s. 137-147). Burlington, UK: Elsevier Butterworth-Heinemann.
- Halladay, J. (2018, 06 30). *MyBit*. 04 28, 2019 tarihinde Moore's Law and the Future of IoT: <https://medium.com/mybit-dapp/moores-law-and-the-future-of-iot-d9ed7d725f0a> adresinden alındı
- Hamilton, J. (2013). *Thomas Cook: The Holiday Maker*. Gloucestershire: The History Press.
- Han, J., Kamber, M. ve Pei, J. (2012). *Data Mining: Concepts and Techniques* (3 b.). Waltham: Morgan Kaufmann Publishers.
- Haroutunian, S., Mitsis, P. ve Pashardes, P. (2005). Using Brochure Information for the Hedonic Analysis of Holiday Packages. *Tourism Economics*, 11(1), 69-84.
- Haykin, S. (2001). *Neural Networks: A Comprehensive Foundation* (2 b.). Delhi, India: Pearson Education.
- Heiberger, R. M. ve Holland, B. (2004). *Statistical Analysis and Data Display: An Intermediate Course with Examples: An Intermediate Course with Examples in S-Plus, R, and SAS*. New York, US: Springer Science & Business Media.
- Hentsch, T. (1992). *Imagining the Middle East*. (F. A. Reed, Çev.) New York, US: Black Rose Books.
- Hobica, G. (2006, 04 15). *Six reasons why we need online travel agencies*. Airfarewatchdog.com: <http://www.airfarewatchdog.com/blog/6530176/six-reasons-why-we-need-online-travel-agencies/> adresinden alındı
- Hoffman, S. (2016). *Electronic Health Records and Medical Big Data: Law and Policy*. New York, US.: Cambridge University Press.
- Holloway, J. C. ve Humphreys, C. (2016). *The Business of Tourism* (10 b.). Harlow: Pearson Education Limited.
- Holloway, J. C. ve Taylor, N. (2006). *The Business of Tourism* (7 b.). Harlow: Financial Times Prentice Hall.
- Horner, J. (2015, 05 22). *Old is New: XML and rvest*. 03 14, 2019 tarihinde R-bloggers.com: <https://www.r-bloggers.com/old-is-new-xml-and-rvest/> adresinden alındı

- Hudson, F. (2018, 03 14). *Top 8 Busiest Airports in Europe*. 04 12, 2019 tarihinde flight-delayed.co.uk: <https://www.flight-delayed.co.uk/blog/2018/03/14/top-8-busiest-airports-in-europe> adresinden alındı
- Hung, W.-T., Shang, J.-K. ve Wang, F.-C. (2010). Pricing Determinants in The Hotel Industry: Quantile Regression Analysis. *International Journal of Hospitality Management*, 29(3), 378-384. doi:10.1016/j.ijhm.2009.09.001
- Huque, S. (2019). *www.galido.net*. 04 27, 2019 tarihinde IPv4 vs IPv6 – Why is IPv6 so important?: <https://galido.net/blog/ipv4-vs-ipv6/> adresinden alındı
- Inkson, C. ve Minnaert, L. (2018). *Tourism Management: An Introduction* (2 b.). Los Angeles, USA: SAGE Publication Ltd.
- International inbound tourism by mode of transport in 2016*. (2016). 08 11, 2018 tarihinde statista.com: <https://www.statista.com/statistics/305515/international-inbound-tourism-by-mode-of-transport/> adresinden alındı
- International Tourism, Number of Arrivals*. (2016). 02 11, 2018 tarihinde The World Bank: <https://data.worldbank.org/indicator/ST.INT.ARVL> adresinden alındı
- International Tourism, Receipts*. (2017). 03 30, 2019 tarihinde The World Bank: <https://data.worldbank.org/indicator/ST.INT.RCPT.CD> adresinden alındı
- Israeli, A. A. (2002). Star rating and corporate affiliation: their influence on room price and performance of hotels in Israel. *International Journal of Hospitality Management*, 21(4), 405-424. doi:10.1016/S0278-4319(02)00037-3
- Işık, C. ve Bilici, N. (2017). Otel Oda Fiyatlarının Hedonik Model İle Analizi: Erzurum Örneği. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 21(1), 305-315.
- Jain, V. K. (2017). *Big Data and Hadoop*. New Delhi, India.: Khanna Publishing.
- Johansson, R. (2019). *Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib* (2 b.). New York, US.: Apress Media.
- Josephs, L. (2018, 02 21). *Online travel giant Priceline Group changes name to Booking Holdings*. 08 07, 2018 tarihinde cnbc.com: <https://www.cnbc.com/2018/02/21/priceline-group-changes-name-to-booking-holdings.html> adresinden alındı
- Kabacoff, R. I. (2011). *R in Action: Data Analysis and Graphics with R*. Shelter Island, US: Manning Publications Co.
- Kale, V. (2017). *Big Data Computing: A Guide for Business and Technology Managers*. Boca Raton, US.: CRC Press.
- Kärcher, K. (1996). The four Global Distribution Systems in the travel and tourism industry. *Information Technology*, 6(2), 20-22.
- Karpathy, A. (2015). 05 10, 2019 tarihinde Neural Networks Part 1: Setting Up the Architecture: <http://cs231n.github.io/neural-networks-1/> adresinden alındı
- Kay, M. (2008). *XSLT 2.0 and XPath 2.0 Programmer's Reference* (4 b.). Indianapolis: Wiley Publishing.



- Kaya, A. (2012). *Türkiye'de Konut Fiyatlarını Etkileyen Faktörlerin Hedonik Fiyat Modeli İle Belirlenmesi* (Cilt Uzmanlık Yeterlilik Tezi). Ankara: T.C Merkez Bankası, İstatistik Genel Müdürlüğü.
- Kervankıran, İ. ve Çuhadar, M. (2017). Türkiye'de İç Turizmin Gelişimi ve Mekânsal İstatistik Yöntemlerle Analizi. *Turizm Akademik Dergisi*, 4(2), 1-18.
- Kesbiç, C. Y., Bahar, O., Baldemir, E. ve İnci, M. (2011). Muğla Deniz Turizminde Hizmet Niteliğinin Piyasa Fiyatı Üzerindeki Etkileri: Hedonik Fiyatlandırma Modeli. *Finans Politik & Ekonomik Yorumlar*, 48(555), 65-82.
- Keskinkılıç, M., Ağca, Y. ve 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
- Kimizuka, M. (2012). *sts.kahaku.go.jp*. 04 27, 2019 tarihinde Historical Development of Magnetic Recording and Tape Recorder: [http://sts.kahaku.go.jp/diversity/document/system/pdf/073\\_e.pdf](http://sts.kahaku.go.jp/diversity/document/system/pdf/073_e.pdf) adresinden alındı
- Kleinberg, A. (2013, 08 20). *www.entrepreneur.com*. 05 01, 2019 tarihinde 5 Small Ways to Use Big Data to Majorly Improve Business: <https://www.entrepreneur.com/article/227957> adresinden alındı
- Koshy, J. (2017, 08 09). *Best Programming Languages for Web Scraping*. 03 10, 2019 tarihinde [promptcloud.com](https://www.promptcloud.com): <https://www.promptcloud.com/blog/best-programming-language-for-web-scraping/> adresinden alındı
- Koshy, J. (2017, 08 14). *www.promptcloud.com*. 05 04, 2019 tarihinde Should Data Scientists Learn Web Scraping?: <https://www.promptcloud.com/blog/should-data-scientists-learn-web-scraping/> adresinden alındı
- Kozak, M. A., Evren, S. ve Çakır, O. (2013). Tarihsel Süreç İçinde Turizm Paradigması. *Anatolia: Turizm Araştırmaları Dergisi*, 24(1), 7-22.
- Kuom, M. ve Oertel, B. (1999). Virtual Travel Agencies. *NetNomics*(2), 225-235. doi:10.1023/A:1019114208191
- Kurzweil, R. (2008). *The Singularity Is Near: When Humans Transcend Biology* (2 b.). London, UK: Penguin Group (USA) Inc.
- Kutschera, S. (2018, 07 4). *Travel statistics to know about in 2018 and 2019*. 08 06, 2018 tarihinde Trekk Soft: <https://www.trekksoft.com/en/blog/65-travel-tourism-statistics-for-2019> adresinden alındı
- Lambert, L. (2005). *The Internet: A Historical Encyclopedia. Chronology*. (H. W. Poole, Dü.) Santa Barbara, US.: ABC-CLIO. <https://books.google.com.tr/books?id=qi-ItIG6QLwC> adresinden alındı
- Laudon, K. C. ve Laudon, J. P. (2016). *Management Information Systems: Managing the Digital Firm* (4 b.). Harlow, UK: Pearson.
- Lee, D. (2016, 05 30). *11 Reasons Why You Should Visit Turkey*. 03 30, 2019 tarihinde [gobackpacking.com](https://gobackpacking.com/reasons-visit-turkey/): <https://gobackpacking.com/reasons-visit-turkey/> adresinden alındı

- Liu, C. ve Albitz, P. (2003). *DNS and BIND: Help for System Administrators* (5 b.). Sebastopol, US.: O'Reilly Media, Inc.
- Lomax, R. G. ve Hahs-Vaughn, D. L. (2012). *An Introduction to Statistical Concepts* (3 b.). New York, US.: Routledge: Taylor & Francis Group.
- Löch, M. ve Axhausen, K. W. (2010). Modeling Hedonic Residential Rents for Land Use and Transport Simulation While Considering Spatial Effects. *The Journal of Transport and Land Use*, 3(2), 39-63. doi:10.1598/jtlu.v3i2.117
- Lubbe, B. (Dü.). (2005). *Tourism Management in Southern Africa* (2 b.). Cape Town, South Africa: Pearson Education South Africa.
- MacLean, G. M. (2004). *The Rise of Oriental Travel: English Visitors to the Ottoman Empire, 1580 - 1720*. New York, US: Springer. 03 30, 2019 tarihinde <https://books.google.com.tr/books?id=1DUWDAAAQBAJ> adresinden alındı
- Mancini, M. (2013). *Access: Introduction to Travel and Tourism*. New York, USA: Delmar Cengage Learning.
- Markov, Z. ve Larose, D. T. (2007). *Data Mining The Web: Uncovering Patterns in Web Content, Structure, and Usage*. New Jersey: John Wiley & Sons.
- Marr, B. (2016). *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. West Sussex, UK.: John Wiley and Sons Ltd.
- McCallum, J. C. (2018, 11 16). *Disk Drive Prices (1955-2018)*. 03 07, 2019 tarihinde <https://jcmit.net>: <https://jcmit.net/diskprice.htm> adresinden alındı
- McCallum, J. C. (2019). *Disk Drive Prices (1955-2019)*. 04 28, 2019 tarihinde <https://jcmit.net/diskprice.htm> adresinden alındı
- MDN web docs: XPath*. (2018, 11 11). 03 15, 2019 tarihinde [developer.mozilla.org](https://developer.mozilla.org): <https://developer.mozilla.org/en-US/docs/Web/XPath> adresinden alındı
- Medlik, S. (2003). *Dictionary of Travel, Tourism and Hospitality* (3 b.). Burlington: Butterworth-Heinemann, Elsevier Science.
- Miller, P. M. (2009). *TCP/IP: The Ultimate Protocol Guide*. Boca Raton, US.: BrownWalker Press. 04 26, 2019 tarihinde [https://books.google.com.tr/books?id=3JQF1PMY\\_QAC](https://books.google.com.tr/books?id=3JQF1PMY_QAC) adresinden alındı
- Mitchell, R. (2018). *Web Scraping with Python: Collecting More Data from the Modern Web* (2 b.). Sebastopol, US: O'Reilly Media, Inc.
- Monson, M. (2009). Valuation Using Hedonic Pricing Models. *Cornell Real Estate Review*, 7, 62-73.
- Mukhopadhyay, S., Agarwal, V., Sharma, S. ve Gupta, V. (2016). A Study On Wireless Communication Networks Based On Different Generations. *International Journal of Current Trends in Engineering & Research (IJCTER)*, 2(5), 300-304.
- Munzert, S., Rubba, C., Meißner, P. ve Nyhuis, D. (2015). *Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining*. Chichester, UK.: John Wiley & Sons.

- Murtagh, F. (1994, 09 10). *www.astro.auth.gr*. 04 27, 2019 tarihinde Information Discovery Tools: WAIS, Gopher, and the World-Wide Web: [https://www.astro.auth.gr/ANTIKATOPTRISMOI/ASTRO-RESOURCES/subsection3\\_2\\_7.html](https://www.astro.auth.gr/ANTIKATOPTRISMOI/ASTRO-RESOURCES/subsection3_2_7.html) adresinden alındı
- Nisbet, R., Miner, G. ve Yale, K. (2018). *Handbook of Statistical Analysis and Data Mining Applications*. London, UK.: Elsevier Inc. Academic Press.
- O'Connell, A. A. (2006). *Logistic Regression Models for Ordinal Response Variables*. Thousand Oaks, US: Sage Publications, Inc.
- O'Connell, J. C. (2017). *Dining Out in Boston: A Culinary History*. Hanover: University Press of New England.
- Online travel market - Statistics & Facts*. (2018). 08 04, 2018 tarihinde The Statistics Portal: <https://www.statista.com/topics/2704/online-travel-market/> adresinden alındı
- Ormankıran, Y., Bayram, A. T. ve Karaçar, E. (2015). Turizm Bakanları Üzerine Bir İnceleme. *Turizm Akademik Dergisi*, 2(2), 1-14.
- Önen, M. O. (2000). *Türkiye'nin Turizm Sektöründeki Gelişmeler, Dünya Turizmindeki Yeri Ve Türkiye Kalkınma Bankası'nın Rolü. (Araştırma)*. Ankara: Türkiye Kalkınma Bankası. 04 2019, 09 tarihinde [http://www.kalkinma.com.tr/data/file/raporlar/ESA/SA/2000-SA/SA-00-02-05\\_Turizm\\_Sektoru.pdf](http://www.kalkinma.com.tr/data/file/raporlar/ESA/SA/2000-SA/SA-00-02-05_Turizm_Sektoru.pdf) adresinden alındı
- Özdamar, K. (2015). *Paket Programlar ile İstatistiksel Veri Analizi* (10 b., Cilt 1). Eskişehir: Nisan Kitapevi Yayınları.
- Özdemir, M. (2013). Türk Turizminin Temel Taşlarından: Emekli Sandığı Otelleri. *Anatolia: Turizm Araştırmaları Dergisi*, 24(2), 266-275.
- Page, S. J. (2007). *Tourism Management Managing for Change* (2 b.). Burlington: Elsevier.
- Papatheodorou, A. (2002). Exploring Competitiveness in Mediterranean Resorts. *Tourism Economics*, 8(2), 133-150.
- Patnala, B. D. ve Kumar, R. K. (2018). *A Novel Level-Based DNA Security Algorithm Using DNA Codons*. (R. G. Bush, Dü.) Singapore: Springer.
- Patterson, I. (2018). *Tourism and Leisure Behaviour in an Ageing World*. Oxfordshire, UK: CABI.
- Pawlicz, A. ve 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.
- Perles-Ribes, J. F., Ramón-Rodríguez, A. B., Moreno-Izquierdo, L. ve Martí, M. T. (2018). Winners and losers in the Arab uprisings: a Mediterranean tourism perspective. *Current Issues in Tourism*, 21(16). doi:10.1080/13683500.2016.1225697
- Petrovic, D. ve Stanisevic, I. (2017). Web scrapping and storing data in a database, a case study of the used cars market. *25th Telecommunication Forum (TELFOR)*

- Telecommunication Forum (TELFOR)*. Belgrade, Serbia: IEEE. doi:10.1109/TELFOR.2017.8249451
- Pham, K. (2016). *Selecting content on a web page with XPath*. 03 14, 2019 tarihinde labs.timtom.ch: <http://labs.timtom.ch> adresinden alındı
- Pizam, A. (Dü.). (2010). *International Encyclopedia of Hospitality Management (2 b.)*. Oxford, UK: Butterworth-Heinemann Elsevier.
- Polidoro, F., Giannini, R., Conte, R. L., Mosca, S. ve Rossetti, F. (2015). Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation. *Statistical Journal of the IAOS*, 31, 165-176. doi:10.3233/SJI-150901
- Press, G. (2015, 01 02). *www.forbes.com*. 04 27, 2019 tarihinde A Very Short History Of The Internet And The Web: <https://www.forbes.com/sites/gilpress/2015/01/02/a-very-short-history-of-the-internet-and-the-web-2/> adresinden alındı
- Prieto, M. (2017). *10 Online Travel Public Companies*. 08 07, 2018 tarihinde TravelTechMedia: <https://medium.com/traveltechmedia/10-online-travel-public-companies-dee6df73f768> adresinden alındı
- Rajiv. (2018, 05 28). *www.rfpage.com*. 04 29, 2019 tarihinde Evolution of wireless technologies 1G to 5G in mobile communication: <https://www.rfpage.com/evolution-of-wireless-technologies-1g-to-5g-in-mobile-communication/> adresinden alındı
- Rajput, A. (2019). *Data Mining – Knowledge Discovery in Databases(KDD)*. 05 04, 2019 tarihinde GeeksforGeeks: <https://www.geeksforgeeks.org/data-mining-kdd-process/> adresinden alındı
- Ratner, B. (2017). *Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data*. Boca Raton, US.: CRC Press.
- Rawat, A. (2018, 02 19). *Ordinal Logistic Regression: An overview and implementation in R*. 02 20, 2019 tarihinde Towards Data Science: <https://towardsdatascience.com/implementing-and-interpreting-ordinal-logistic-regression-1ee699274cf5> adresinden alındı
- Reilly, E. D. (2003). *Milestones in Computer Science and Information Technology*. Westport, US: Greenwood Press.
- Reitz, K. (2018). *Quickstart: Requests*. 03 10, 2019 tarihinde <http://docs.python-requests.org>: <http://docs.python-requests.org/en/master/user/quickstart/> adresinden alındı
- Reitz, K. (2019). *HTML Scraping*. 03 15, 2019 tarihinde The Hitchhiker's Guide to Python: <https://docs.python-guide.org/scenarios/scrape/> adresinden alındı
- Richardson, L. (2015). *Beautiful Soup Documentation*. 03 10, 2019 tarihinde crummy.com: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/#> adresinden alındı
- Roiger, R. J. (2017). *Data Mining: A Tutorial-Based Primer (2 b.)*. Boca Raton, US.: CRC Press, Taylor & Francis Group.

- Ronquillo, A. (2019, 01 23). *Python's Requests Library*. 03 10, 2019 tarihinde Realpython.com: <https://realpython.com/python-requests/> adresinden alındı
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34-55. <http://www.jstor.org/stable/1830899> adresinden alındı
- Rouse, M. (2018). <https://searchdatamanagement.techtarget.com>. 04 30, 2019 tarihinde Big Data: <https://searchdatamanagement.techtarget.com/definition/big-data> adresinden alındı
- Sammut, C. ve Webb, G. I. (Dü.). (2017). *Encyclopedia of Machine Learning and Data Mining*. New York, US.: Springer. doi:10.1007/978-1-4899-7687-1
- Sanchez, G. (2014). *Getting Data from the Web with R*. (C. Dimitrakakis, Dü.) 03 15, 2019 tarihinde University of Gothenburg: [http://www.cse.chalmers.se/~chrdimi/downloads/web/getting\\_web\\_data\\_r4\\_parsing\\_xml\\_html.pdf](http://www.cse.chalmers.se/~chrdimi/downloads/web/getting_web_data_r4_parsing_xml_html.pdf) adresinden alındı
- Saunders, R. A. (2011). *Ethnopolitics in Cyberspace: The Internet, Minority Nationalism, and the Web of Identity*. Lanham, US.: Lexington Books.
- Saygılıoğlu, N. (2018, 01 24). *24 Ocak Kararları'nı hatırlamak....* 04 04, 2019 tarihinde dünya.com: <https://www.dunya.com/kose-yazisi/24-ocak-kararlarini-hatirlamak/399841> adresinden alındı
- Schenker, A., Bunke, H., Last, M. ve Kandel, A. (2005). *Graph-Theoretic Techniques for Web Content Mining*. Danvers: World Scientific Publishing Co. Pte. Ltd.
- Schulz, A. (1997). Electronic Marketcoordination in the Travel Industry: The Role of Global Computer Reservation Systems. A. M. Tjoa (Dü.) içinde, *Information and Communication Technologies in Tourism* (s. 67-74). Vienna: Springer. doi:10.1007/978-3-7091-6848-6\_8
- Sezgin, M. (2015). *Genel Turizm* (1 b.). (Z. Aslan, Dü.) Ankara: Grafiker Yayınları.
- Sharma, H. (2019, 04 13). [www.edureka.co](http://www.edureka.co). 05 04, 2019 tarihinde What Is Data Science? A Beginner's Guide To Data Science: <https://www.edureka.co/blog/what-is-data-science/> adresinden alındı
- Sharpley, R. (2006). *Travel and Tourism*. London: SAGE Publications.
- Shelly, G. B. ve Campbell, J. (2012). *Discovering the Internet: Brief*. Boston, US: Cengage Learning.
- Shinder, D. L. (2002). *Computer Networking Essentials*. Indianapolis, US.: Cisco Press. BLG433-Bilgisayar Haberleşmesi Ders Notları. adresinden alındı
- Shylaja, S. ve Muralidharan, R. (2017). Application of Data Mining Techniques in Wireless Sensor Networks: A Review. *IRACST – International Journal of Computer Networks and Wireless Communications (IJCNWC)*, 7(3), 1-5.
- Skrabec, Q. R. (2012). *The 100 Most Significant Events in American Business: An Encyclopedia*. Santa Barbara, US: Greewood.
- Soriano, J. L. (2012). *Maximizing Benefits from IT Project Management: From Requirements to Value Delivery*. Boca Raton, US: CRC Press. 04 28, 2019

tarihinde <https://books.google.com.tr/books?id=XWvRBQAAQBAJ> adresinden alındı

- Sparks, B., Bowen, J. ve Klag, S. (2003). Restaurants and the tourist market. *International Journal of Contemporary Hospitality Management*, 15(1), 6-13.
- Stephenson, D. (2018). *Big Data Demystified: How to use big data, data science and AI to make better business decisions and gain competitive advantage*. London, UK: FT Publishing Financial Times.
- Sumathi, S. ve Sivanandam, S. (2006). *Introduction to Data Mining and Its Applications*. Berlin, Germany: Springer. 05 11, 2019 tarihinde [https://books.google.com.tr/books?id=QwveL0jFk\\_gC](https://books.google.com.tr/books?id=QwveL0jFk_gC) adresinden alındı
- Şahin, A. (1989). 21. Yüzyıla Girerken Türkiye'nin Turizm Sektörüne Uyguladığı Teşvik Politikalarının Genel Bir Değerlendirilmesi. *İstanbul Üniversitesi İktisat Fakültesi Mecmuası*, 47(1-4), 307-322. <http://dergipark.gov.tr/iuifm/issue/837/9247> adresinden alındı
- Şanlıoğlu, Ö. ve Özcan, E. Ö. (2017). Türkiye'de Uygulanan Turizm Teşvik Politikaları ve Sonuçları Üzerine Bir Değerlendirme. *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi*, 7(2), 97-118. 03 30, 2019 tarihinde <http://dergipark.gov.tr/kusbd/issue/30549/302983> adresinden alındı
- Tabachnick, B. G. ve Fidell, L. S. (2013). *Using Multivariate Statistics* (6 b.). Boston: Pearson Education.
- Taylor, P. (1995). Measuring Changes in the Relative Competitiveness of Package Tour Destinations. *Tourism Economics*, 1(2), 169-182.
- Templeton, G. (2015, 07 29). *www.extremetech.com*. 04 28, 2019 tarihinde What is Moore's Law?: <https://www.extremetech.com/extreme/210872-extremetech-explains-what-is-moores-law> adresinden alındı
- Theobald, W. F. (2005). *Global Tourism* (3 b.). Burlington: Elsevier.
- 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
- Thrane, C. (2007). Examining the Determinants of Room Rates for Hotels in Capital Cities: The Oslo Experience. *Journal of Revenue and Pricing Management*, 5(4), 315-323.
- Toker, B. (2007). Türkiye'de Turizm Sektörü Teşviklerinin Değerlendirilmesi. *Yönetim ve Ekonomi*, 14(2), 81-92.
- Touchstone Delta*. (2019). 03 06, 2019 tarihinde Computer History Museum: <https://www.computerhistory.org/revolution/supercomputers/10/74/288> adresinden alındı
- Ţugulea, O., Bobâlcă, C., Maha, A. ve Maha, L.-G. (2014). Using Internet And Travel Agencies In Planning A Trip. A Qualitative Approach. *Centre for European Studies (CES) Working Papers*, 6(1), 181-200.

- Turing Tarihçe.* (2019). 04 11, 2019 tarihinde [turing.org.tr](http://www.turing.org.tr): <http://www.turing.org.tr/tarihce/> adresinden alındı
- Turizm İstatistikleri.* (2018). 04 10, 2019 tarihinde T.C. Kültür ve Turizm Bakanlığı: <http://yigm.kulturturizm.gov.tr/Eklenti/62462,2018turizmgenelistatistiklerpdf.pdf?0> adresinden alındı
- Turizm İstatistikleri.* (2019). 04 09, 2019 tarihinde T.C. Kültür ve Turizm Bakanlığı: <http://yigm.kulturturizm.gov.tr/TR-9851/turizm-istatistikleri.html> adresinden alındı
- Turizm İstatistikleri Yıllık Bültenler.* (2018). 04 10, 2019 tarihinde T.C. Kültür ve Turizm Bakanlığı: <http://yigm.kulturturizm.gov.tr/Eklenti/62268,2018-yili-turizm-isletme-ve-yatirim-belgeli-tesis-istat-.xlsx?0> adresinden alındı
- Türkiye Otel İstatistikleri.* (2015). 04 09, 2019 tarihinde Türkiye Otelciler Birliği: <http://www.turob.com/tr/istatistikler/turkiye-otel-istatistikleri-9d67c3c> adresinden alındı
- Türkiye'deki havaalanları listesi.* (2018). 04 12, 2019 tarihinde <https://tr.wikipedia.org/>: [https://tr.wikipedia.org/wiki/Türkiye'deki\\_havaalanları\\_listesi](https://tr.wikipedia.org/wiki/Türkiye'deki_havaalanları_listesi) adresinden alındı
- Using XPath Functions.* (2013). 03 15, 2019 tarihinde [docs.oracle.com](https://docs.oracle.com/): [https://docs.oracle.com/cd/E35413\\_01/doc.722/e35419/dev\\_xpath\\_functions.htm](https://docs.oracle.com/cd/E35413_01/doc.722/e35419/dev_xpath_functions.htm) adresinden alındı
- Vanhove, N. (2017). *The Economics of Tourism Destinations: Theory and Practice* (3 b.). Oxford, UK: Routledge.
- Voulgaris, Z. (2014). *Data Scientist: The Definitive Guide to Becoming a Data Scientist*. Baskin Ridge, US.: Technics Publications.
- Vousinas, G. (2012). *Internet & the birth of InfoCom industry. IT & Economic Performance*. LAP LAMBERT Academic Publishing.
- Waltner, A. (2018, 07 02). *35 Reasons Why You Should Visit Turkey*. 03 30, 2019 tarihinde [swedishnomad.com](http://swedishnomad.com): <https://www.swedishnomad.com/reasons-to-visit-turkey/> adresinden alındı
- Webster, F. (2006). Making Sense of the Information Age. *Sociology and Cultural Studies*, 8(4), 439-458. doi:10.1080/13691180500418212
- Wei-Ting, H., Jui-Kou, S. ve Fei-Ching, W. (2010). Pricing determinants in the hotel industry: Quantile regression analysis. *International Journal of Hospitality Management*, 29(3), 378-384. doi:<https://doi.org/10.1016/j.ijhm.2009.09.001>
- Whittaker, J. (2002). *The Internet: The Basics*. London, UK.: Routledge.
- World Tourism Organization.* (2018, 01). 02 11, 2018 tarihinde Tourism Market Trends UNWTO: <http://mkt.unwto.org/barometer/january-2018-volume-16-advance-release> adresinden alındı
- WTO. (1994). *Recommendations on Tourism Statistics*. Luxembourg, Madrid, New York, Paris: World Tourism Organization. <https://www.e-unwto.org/doi/pdf/10.18111/9789211613629> adresinden alındı

- Wu, L. (1999). The pricing of a brand name product: Franchising in the motel services industry. *Journal of Business Venturing*, 87-102. doi:10.1016/S0883-9026(97)00101-8
- www.alexacom.com*. (2018, 08 07). 08 07, 2018 tarihinde The top 500 sites on the web: <https://www.alexacom.com/topsites/category/Recreation/Travel> adresinden alındı
- www.belgeler.org*. (2019). 04 27, 2019 tarihinde BIND 8 ile DNS Nasıl: [http://www.belgeler.org/howto/bind-ile-dns-howto\\_dns.html](http://www.belgeler.org/howto/bind-ile-dns-howto_dns.html) adresinden alındı
- www.dhmi.gov.tr*. (2018). 04 12, 2019 tarihinde Devlet Hava Meydanları İşletmesi Genel Müdürlüğü: <https://www.dhmi.gov.tr/Lists/IstatistikList/Attachments/339/YOLCU.pdf> adresinden alındı
- www.iana.org*. (2019). 04 27, 2019 tarihinde Root Zone Database: <https://www.iana.org/domains/root/db> adresinden alındı
- www.internetlivestats.com*. (2019, 03 08). 03 08, 2019 tarihinde Internet Live Stats: <http://www.internetlivestats.com/internet-users/> adresinden alındı
- www.knime.com*. (2019). 03 13, 2019 tarihinde About KNIME: <https://www.knime.com/about> adresinden alındı
- www.oracle.com*. (2019). 04 30, 2019 tarihinde What Is Big Data?: <https://www.oracle.com/big-data/guide/what-is-big-data.html> adresinden alındı
- www.oracle.com*. (2019). 05 04, 2019 tarihinde Data Mining Concepts: [https://docs.oracle.com/cd/B28359\\_01/datamine.111/b28129/process.htm](https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm) adresinden alındı
- www.otelz.com*. (2018, 08 10). <https://www.otelz.com/search> adresinden alındı
- www.sas.com*. (2019). 05 04, 2019 tarihinde Data Mining: What it is and why it matters: [https://www.sas.com/en\\_us/insights/analytics/data-mining.html](https://www.sas.com/en_us/insights/analytics/data-mining.html) adresinden alındı
- www.statista.com*. (2017). 11 27, 2017 tarihinde Global digital travel sales 2014-2020. adresinden alındı
- www.tripadvisor.com*. (2017, 11 15). [tripadvisor.com: https://www.tripadvisor.com.tr/pages/about\\_us.html](https://www.tripadvisor.com.tr/pages/about_us.html) adresinden alındı
- www.turob.com*. (2018). 04 10, 2019 tarihinde Türkiye Otelciler Birliği: İstatistikler: <http://www.turob.com/Files/Kasim%202018%20Antalya.pdf> adresinden alındı
- www.w3schools.com*. (2019). 03 13, 2019 tarihinde XML and XPath: [https://www.w3schools.com/xml/xml\\_xpath.asp](https://www.w3schools.com/xml/xml_xpath.asp) adresinden alındı
- Xiao, Y. (2016). Hedonic Housing Price Theory Review. *Urban Morphology and Housing Market. Springer Geography*. (s. 11-40). içinde Singapore: Springer. doi:10.1007/978-981-10-2762-8\_2
- XPath Operators*. (2019). 03 15, 2019 tarihinde [w3schools.com: https://www.w3schools.com/xml/xpath\\_operators.asp](https://www.w3schools.com/xml/xpath_operators.asp) adresinden alındı
- Yalav-Heckerth, F. (2017, 06 19). *10 Reasons Why You Should Visit Turkey at Least Once in Your Lifetime*. 03 30, 2019 tarihinde [theculturetrip.com](http://theculturetrip.com):



<https://theculturetrip.com/europe/turkey/articles/10-reasons-why-you-should-visit-turkey-at-least-once-in-your-lifetime/> adresinden alındı

- Yayar, R. ve Gül, D. (2014). Mersin Kent Merkezinde Konut Piyasası Fiyatlarının Hedonik Tahmini. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 14(3), 87-99.
- Yin, M. (2019, 02 25). *Scrapy Tutorial #1: Scrapy VS Beautiful Soup*. 03 10, 2019 tarihinde [accordbox.com](https://www.accordbox.com/blog/scrapy-tutorial-1-scrapy-vs-beautiful-soup/): <https://www.accordbox.com/blog/scrapy-tutorial-1-scrapy-vs-beautiful-soup/> adresinden alındı
- Zhang, H., Zhang, J., Lu, S., Cheng, S. ve Zhang, J. (2011). Modeling hotel room price with geographically weighted regression. *International Journal of Hospitality Management*, 30, 1036-1043. doi:10.1016/j.ijhm.2011.03.010
- Zimmermann, K. A. (2017, 09 06). [www.livescience.com](http://www.livescience.com). 04 28, 2019 tarihinde History of Computers: A Brief Timeline: <https://www.livescience.com/20718-computer-history.html> adresinden alındı
- Zyl, T. v. (2014). *Big Data: Techniques and Technologies in Geoinformatics*. (H. A. Karimi, Dü.) Boca Raton, US.: Taylor & Franchis Group, LLC.

## EKLER

Ek 1: Türkiye'deki Ortalama Oda Fiyatlarına İlişkin Sayısal Bilgiler

Şehir	N	Gruplandırılmış		Gruplandırılmamış			
		Ort.	Std.Sap.	Ort.	Std.Sap.	Min.	Mak.
Muğla	948	4,66	1,164	494,51	645,94	98,25	12279,88
Antalya	1171	4,57	1,374	510,78	579,97	60,50	9012,60
İzmir	547	4,41	1,059	376,76	314,18	80,00	4482,14
İstanbul	1560	4,44	,987	376,66	277,51	84,33	4301,71
Giresun	18	4,28	1,127	468,57	817,25	155,00	3721,00
Nevşehir	278	4,41	1,126	393,66	332,18	71,50	3373,00
Karaman	4	4,25	1,893	892,71	1359,01	184,00	2930,50
Sakarya	43	4,30	1,013	415,05	459,08	150,00	2869,10
Trabzon	113	5,09	,960	536,19	370,05	158,00	2664,00
Karabük	52	3,77	1,131	312,55	406,84	72,50	2490,86
Şanlıurfa	15	3,73	1,100	364,80	530,59	146,00	2272,25
Aydın	156	4,01	1,213	322,67	228,87	86,75	1840,36
Balıkesir	181	4,30	,804	321,46	173,17	116,63	1840,11
Ankara	191	3,82	1,012	275,18	173,18	95,50	1428,29
Van	11	4,00	1,183	338,97	336,10	164,50	1333,33
Çanakkale	151	4,50	,799	353,76	147,73	139,25	1307,55
Mardin	25	3,88	1,092	304,70	224,67	144,17	1172,20
Sinop	19	4,26	1,046	349,45	197,91	143,14	1037,57
Yalova	21	4,52	,981	366,14	183,71	140,45	973,91
Elazığ	14	3,36	,929	249,78	209,44	140,00	966,20
Bursa	93	4,29	,867	323,75	133,42	131,00	900,43
Rize	34	4,88	,880	447,18	206,74	186,70	815,50
Afyonkarahisar	14	4,36	1,082	353,74	166,19	153,00	778,30
Erzurum	13	4,00	1,155	305,37	174,63	152,50	720,33
Bolu	31	4,35	,709	331,91	115,13	182,00	708,67
Denizli	65	3,65	1,110	249,29	125,51	100,00	666,40
Gaziantep	37	3,19	,995	203,60	96,03	96,00	662,00
Zonguldak	10	3,80	1,229	266,84	152,70	138,50	652,60
Ordu	21	4,57	,811	342,20	127,45	181,09	573,91
Kars	19	3,84	,765	267,77	98,66	179,00	573,00
Kocaeli	52	3,71	,936	254,54	97,76	105,25	568,80
Edirne	21	3,76	,995	252,42	98,52	128,25	568,13
Mersin	64	3,94	,889	277,37	109,13	120,93	558,33
Burdur	5	4,60	1,517	385,62	163,26	135,00	555,20
Kırklareli	5	3,80	,837	283,13	147,67	188,88	539,50
Samsun	43	3,47	,984	230,21	102,89	91,00	538,71

Bitlis	8	3,50	1,309	254,92	131,12	106,00	504,00
Hatay	40	3,95	,876	260,04	82,93	117,75	475,00
Kastamonu	23	3,74	,810	247,65	83,47	104,00	469,67
Konya	48	3,38	,937	218,00	73,90	97,29	466,63
Manisa	23	3,87	,458	248,25	60,82	167,00	465,00
Amasya	19	3,89	,567	248,73	63,40	198,00	464,86
Tekirdağ	28	4,00	,667	260,47	76,78	154,33	464,67
Artvin	5	4,60	,548	337,65	85,67	233,50	448,00
Kütahya	6	3,50	1,049	239,77	109,58	129,00	447,13
Kahramanmaraş	10	3,50	,972	235,01	84,04	130,00	419,71
Adana	47	3,28	,826	199,93	61,06	113,00	409,00
Sivas	12	3,50	,905	226,74	73,15	127,20	401,17
Bartın	20	4,05	,759	260,80	72,26	122,00	394,08
Kayseri	27	3,19	1,039	194,56	69,77	99,00	377,17
Düzce	12	3,67	,778	214,60	61,30	126,10	362,70
Eskişehir	43	3,05	,722	186,24	43,50	105,20	361,75
Kırıkkale	2	4,00	1,414	270,33	127,75	180,00	360,67
Isparta	14	3,21	,893	206,12	69,72	144,60	360,10
Gümüşhane	2	3,00	2,828	223,18	176,32	98,50	347,86
Muş	2	3,50	2,121	238,25	136,12	142,00	334,50
Aksaray	12	3,83	,835	242,80	65,46	123,33	329,50
Malatya	20	2,90	1,021	180,43	60,55	89,20	329,00
Çankırı	4	3,50	,577	230,32	67,51	162,50	308,67
Kilis	1	4,00		301,14		301,14	301,14
Diyarbakır	17	3,47	,624	214,01	46,82	117,00	294,86
Niğde	5	3,60	,548	224,85	50,04	173,00	290,00
Ağrı	3	3,67	,577	243,22	61,23	173,67	289,00
Şırnak	2	3,50	,707	241,83	66,23	195,00	288,67
Çorum	10	3,30	,675	199,93	45,27	130,00	278,00
Ardahan	3	3,33	,577	207,04	52,40	170,00	267,00
Erzincan	9	3,11	,601	192,72	36,61	123,50	261,25
Uşak	6	3,67	,516	225,21	37,53	172,40	258,00
Tokat	6	3,33	,516	202,22	33,79	168,17	252,29
Bilecik	2	4,00	0,000	233,92	18,27	221,00	246,83
Adıyaman	9	3,11	,333	176,87	25,22	150,00	234,71
Yozgat	2	3,50	,707	211,00	29,70	190,00	232,00
Batman	6	3,50	,548	204,36	16,13	183,29	225,57
Bingöl	1	4,00		212,00		212,00	212,00
Siirt	1	4,00		210,00		210,00	210,00
Osmaniye	1	4,00		210,00		210,00	210,00
Hakkâri	1	3,00		195,00		195,00	195,00
Bayburt	1	3,00		157,00		157,00	157,00
<b>Toplam</b>	<b>6558</b>	<b>4,35</b>	<b>1,151</b>	<b>395,44</b>	<b>417,34</b>	<b>60,50</b>	<b>12279,88</b>

**Ek 2: Verinin Elde Edildiđi Dönemde Ortalama Aylık Döviz Kurları**

Aylar	<i>Aylık Ortalama Dolar Kuru</i>		<i>Aylık Ortalama Euro Kuru</i>	
	<i>\$/TL</i>	<i>TL/\$</i>	<i>€/TL</i>	<i>TL/€</i>
Mayıs'2018	4,42	0,226	5,23	0,191
Haziran'2018	4,64	0,216	5,42	0,185
Temmuz'2018	4,76	0,210	5,56	0,180
Ağustos'2018	5,74	0,174	6,62	0,151
Eylül'2018	6,38	0,157	7,44	0,134

### Ek 3: Otel Oda Fiyatlarını Etkileyen Faktörlerin Belirlenmesine Yönelik Geçmiş Yıllarda Yapılan Bazı Çalışmalar

YIL	ÇALIŞMA	YER	KAYNAK	VERİ	MODEL	BAĞIMSIZ DEĞİŞKENLER	BAĞIMLI DEĞİŞKEN
1984	Goldberg, S. M., Green, P. E. ve Wind, Y. (1984). Conjoint Analysis of Price Premiums for Hotel Amenities.	U.S	Telefon İle Anket	180 Otel	Hybrid Conjoint Model	Oda, spor, güvenlik, dinlenme salonu (Lounge)	Oda Fiyatı
1990	Carvell, S. A. ve Herrin, W. E. (1990). Pricing in the Hospitality Industry; An Implicit Market Approach	San Francisco	Seyahat Acentesi Kataloğu	20 Otel (1982-1986)	Hedonic Regression	Konsiyerj, Spor Salonu, Kuru Temizleme, Ücretsiz Yerel Telefon, Kahvaltı, Yıldız, Turistiklik bir yere mesafe (Fisherman's Wharf)	Ortalama Oda Fiyatı
1994	Bull, A. O. (1994). Pricing A Motel's Location.	Avustralya		20 Otel, 15 Motel	Lineer Hedonic	Oda fiyatı, Yıldız, Otel Yaşı, Restoran, Merkeze Mesafe, Yol Kenarı Olma	Oda Fiyatı
1995	Taylor, P. (1995). Measuring Changes in the Relative Competitiveness of Package Tour Destinations.	Akdeniz	Seyahat Acentesi Kataloğu	150 Paket Tur	Hedonic Regression Analysis	Dört Farklı Destinasyon Kıyaslanmış	Paket Tur Fiyatları
1999	Wu, L. (1999). The pricing of a brand name product: Franchising in the motel services industry.	Arkansas ve Kansan	Amerikan Otomobil Derneği (AAA) Kataloğu	155 Otel	Lineer Regression Analysis	AAA Derecesi, Restoran, Havuz, Zincir Otel Büyüklüğü,	Oda Fiyatları (En Düşük ve En Yüksek)
2002	Papathodorou, A. (2002). Exploring Competitiveness in Mediterranean Resorts	Akdeniz, (Türkiye dâhil)	Seyahat Acentesi Kataloğu	240 Paket Tur	Hedonic Regression Analysis	Zincir, Yıldız, Haftalık Uçuş Sayısı, Pansiyon Türü, Uçuş Zamanı, Her Şey Dahil,	Paket tur fiyatları
2003	Aguiló, E., Alegre, J. ve Sard, M. (2003). Examining the Market Structure of The German and UK Tour Operating Industries Through An Analysis of Package Holiday Prices	Balea Adaları Özerk Topluluğu, İspanya	Seyahat Acentesi Kataloğu	693 Otel	Hedonic Regression Analysis	Yıldız, Odadaki Yatak Sayısı, Pansiyon Tipi, Tur Operatörü, Mini Bar, Sauna vb.	Oda Fiyatları
2005	Haroutunian, S., Mitsis, P. ve Pashardes, P. (2005). Using Brochure Information for the Hedonic Analysis of Holiday Packages.	Akdeniz, (Türkiye dâhil)	Seyahat Acentesi Kataloğu	349 Paket Tur	Hedonic Regression Analysis	Ülke, Banyo, Spor, Güzellik Salonu, All Inclusive, Acente, Tenis, Oyun Odası, Restoran,	Paket tur fiyatları
2007	Thrane, C. (2007). Examining the Determinants of Room Rates for Hotels in Capital Cities; The Oslo Experience	Oslo	Online,	74 Otel	Semilog Hedonic	Zincir, mini-bar, park, restoran, oda servisi, yataklar vs.	Fiyat
2010	Hung, W.-T., Shang, J.-K. ve Wang, F.-C. (2010). Pricing Determinants in The Hotel Industry; Quantile Regression Analysis	Tayvan	Hükümet Raporu	58 Otel	Quantile Regression Analysis (ORA)	Oda Sayısı, Otelin Yaşı, Kat Hizmetleri Personeli Sayısı	Ortalama Oda Fiyatları

## ÖZGEÇMİŞ

<b>Kişisel Bilgiler</b>	
Adı Soyadı	Yılmaz AĞCA
Doğum Yeri ve Tarihi	Malatya, 06.04.1983
<b>Eğitim Durumu</b>	
Lisans Öğrenimi	Akdeniz Üniversitesi, Turizm İşletmeciliği ve Otelcilik Y.O, Seyahat İşletmeciliği Bölümü
Y. Lisans Öğrenimi	Gaziosmanpaşa Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme A.B.D
Bildiği Yabancı Diller	İngilizce, Almanca
Bilimsel Faaliyetleri	Yönetim Bilişim Sistemleri, Turizm Enformatiği, Sürdürülebilir Turizm
<b>İş Deneyimi</b>	
Stajlar	Novaraft Seyahat Acentesi Samsun Öğretmenevi Büyük Samsun Oteli
Projeler	Innovative Tourism - Collecting Cultural Heritage through International Walks, EU Erasmus+ K2
Çalıştığı Kurumlar	Corendon Turistic, Operasyon, NL. Enka Reizen, Operasyon, NL. IMC Web Design, Web Tasarım.
<b>İletişim</b>	
E-Posta Adresi	agca.yilmaz@gmail.com
Tarih	