

## Yolculuk Değerlerini Etkileyen Faktörlerin Veri Madenciliği Analizi ile İncelenmesi: Kahramanmaraş Örneği

\*<sup>1</sup>İbrahim Can Turan, <sup>2</sup>Hakan Aslan, <sup>3</sup>Zeliha Çağla Kuyumcu

<sup>1</sup>Fen Bilimleri Enstitüsü, İnşaat Mühendisliği Ulaştırma Programı, Sakarya Üniversitesi, Türkiye  
<sup>2,3</sup>Mühendislik Fakültesi, İnşaat Mühendisliği Bölümü, Sakarya Üniversitesi, Türkiye

### Özet

Ulaşım; toplumsal, sosyal ve ekonomik açıdan önemli bir role sahiptir. Bu nedenle, ekonomik kaynakların kullanımı, dış ilişkileri arttırmak, mal akışlarının düzenli olması, mevcut veya yeni yerleşimlerin gelişmesi, ulaştırma süreçlerinin devamlılığı; kısaca sürdürülebilir ve dengeli kalkınma sağlanabilmesi için ulaşım sektörünün gelişimi de oldukça önemlidir. Ulaşımın düzenlenmesi ve gelişmesi de ancak mevcut durumun analizi neticesinde mümkündür. Veri madenciliği bu analizlerin yapılması ve yorumlanmasını mümkün kılmaktadır. Bu çalışmada, seyahat edicilerin yaşları, cinsiyetleri, otomobil sahiplikleri, ekonomik durumları, seyahatlerine başladıkları ve seyahatlerini bitirdikleri coğrafi noktalar ve seyahat için kullandıkları ulaşım modları (türleri) dikkate alınarak toplanan Kahramanmaraş Büyükşehir Belediyesi Ulaşım Ana Planı Hane Halkı Anketi verileri, zon bazlı olarak incelenmiş; veri madenciliği araçlarından biri olan RapidMiner kullanılarak birliktelik kuralları ile analiz edilerek değişkenler arasındaki ilişkiler belirlenmiş ve yorumlanmıştır.

**Anahtar kelimeler:** Veri madenciliği, birliktelik kuralları, kent içi ulaşım, ulaşım ana planı, RapidMiner

### Abstract

Transportation plays a significant role in the social, economic and social aspects of the community. For this reason, the efficient use of economic resources, increasing foreign relations, regularity of goods flows, development of existing or new settlements, continuity of transportation processes are all related to the quality of available network systems. In other words, the development of the transportation sector is especially important for sustainable and balanced development. The regulation and development of transportation system is possible through the analysis of the current situation. Data mining makes it possible for us to take and interpret these analyses. The data collected by considering the age-gender, vehicle ownership level, economic status (income level) of the participants along with the geographical points regarding the start and finish points of the trips and the transportation modes) used etc. have been obtained through Kahramanmaraş Metropolitan Municipality Transportation Master Plan Household Survey. All these data categorized in terms of zone-base are analyzed accordingly by using RapidMiner, one of the data mining tools, to reveal the relations between the variables through the association rules. Finally, the results of this analysis were interpreted.

**Key words:** Data mining, association rules, urban transportation, transportation master plan, RapidMiner

## **1. Giriş**

Yeni Bilgilerin Toplanması ve Değerlendirilmesi sürecindeki çalışmaların ilki, hane halkı ulaşım araştırmasıdır. Bu çalışmalarda hane halkı, genellikle kurumsal olmayan nüfusu (üniversite yurtları, yetiştirme yurtları, huzurevi, özel nitelikli hastane, cezaevi, kışla vb. yerlerde ikamet edenler dışında kalan nüfusu) temsil etmektedir.

Bu çalışmada kullanılan hane halkı ulaşım araştırması, Kahramanmaraş Ulaşım Ana Planı kapsamında yapılan ulaşım modeli çalışmalarına altlık oluşturan verileri içermektedir.

Çalışma alanı sınırları içerisinde ikamet eden hane halkı bireylerinden “hafta içi bir günde” yaptıkları yolculuklar çerçevesinde;

- Son 24 saat içinde gerçekleştirdikleri seyahat davranışları (yolculuk amaçları, yolculuk zamanı, süresi, ulaşım türü, aktarmalar, beklemler, başlangıç ve bitiş noktaları vb.) hakkında bilgi toplamak,
- Hane ve birey için cinsiyet, yaş, eğitim, meslek ve gelir durumu gibi özelliklerine göre seyahat davranışlarını tespit etmek,
- Araç sahipliği ve kullanım alışkanlıkları ile bu alışkanlıklara etkisini tespit etmek,
- Katılımcıların seyahat alışkanlıklarının uyduğu olasılık dağılım parametrelerini belirlemek,

amacıyla veriler elde edilmiştir.

Bu çalışma kapsamında elde edilen veriler ışığında, kentin mevcut sosyal, ekonomik ve demografik göstergeleri bir arada değerlendirilerek Kahramanmaraş’ın mevcut yolculuk yapısına etki eden ana faktörler belirlenmiş ve kent hareketlilik yapısı birliktelik kuralları ile analiz edilmiştir.

Çalışma kapsamında, Kahramanmaraş merkezinde bulunan Dulkadiroğlu ve Onikişubat ilçe sınırları içerisinde mahalleler, sosyoekonomik yapı ve kentle olan ulaşım ilişkileri açısından irdelenerek kentsel dokuya sahip olan mahalleler örneklem alanı olarak değerlendirilmiştir. Kırsal nitelikte olan yerleşim alanları, yerleşik nüfus barındırmayan organize sanayi bölgeleri, merkezi iş alanları, otogar vb. kullanım alanları ise örneklem dâhilinde değildir.

## **2. Yöntem**

Bu çalışmada, olayların birlikte gerçekleşme durumlarını çözümlen veri madenciliği yöntemlerinden biri olan birliktelik kuralları ile çalışılmıştır.

### **2.1. Veri madenciliği**

Veri madenciliği, bilgisayar teknolojisi kullanılmadan değerlendirilemeyecek kadar çok verilerin farklı bir bakış açısı ile analiz edilmesi ve kullanışlı bilgi olarak özetlenmesi sürecidir. Veri madenciliği teknik olarak, büyük ve birbiriyle ilişkili veri tabanları içinde bulunan düzinelere

değişkenler arasındaki korelasyonları ve düzenleri bulma sürecidir. Veri madenciliği üzerinde fikir birliğine varılmış ortak bir tanım yoktur. Bu genel yapıyla ilişkin olarak literatürde veri madenciliği ile ilgili bazı tanımlar şöyle verilmektedir;

- Veri madenciliği; büyük hacimli verilerden öz bilginin çıkarılması sürecidir [1]. Bir başka ifade ile veri madenciliği büyük ve karmaşık verilerde beklenmeyen patikaların, değerli yapıların ve ilginç ilişkilerin keşfedilmesi bilimidir.
- Veri madenciliği büyük veri tabanlarındaki gizli bilgi ve yapıyı açıklamak için, çok sayıda veri analizi aracını kullanan bir süreçtir [2].
- Bilgisayar teknolojilerinin sağlamış olduğu çok hızlı veri işleme ve yüksek hacimde veri depolama imkânları yardımıyla ve farklı disiplinlerin katkısıyla sağlanan araçlarla, sahip olunan çok büyük hacimlerdeki veriden, karar vericinin etkin, öz ve daha fazla bilgiye dayalı karar vermesinde kullanabilmesi amacıyla önceden bilinmeyen, gizli, örtük, klasik metotlarla ortaya çıkarılması güç, faydalı, ilginç, anlaşılabilir; ilişki, örüntü, bağıntı veya trendlerin otomatik veya yarı otomatik bir şekilde ortaya çıkarılması olarak tanımlanır [3].
- Veri madenciliği genel anlamda, büyük miktarda veri içerisinde, gizli kalmış, değerli, kullanılabilir bilgilerin açığa çıkarılmasıdır [4].
- Veri madenciliği ve öz bilgi keşfi, verilerde daha önceden bilinmeyen, anlamlı ve değerli bilgiler elde etme işlemidir [5].

## ***2.2. Veri madenciliği süreci ve aşamaları***

Veri madenciliği bir süreçtir. Veri yığınları arasında soyut kazılar yaparak veriyi ortaya çıkarmanın yanı sıra, bilgi keşfi sürecinde örüntüleri ayrıştırarak süzmek ve bir sonraki adıma hazır hale getirmek de bu sürecin bir parçasıdır. Veri tabanlarında öz bilgi keşfi sürecinde sırasıyla 5 temel aşama izlenmektedir [6].

### ***2.2.1. Problemin tanımlanması***

Analizin hangi amaçla yapılacağına açık bir şekilde ifade edilmesidir. Bu anlamda problem açıkça ifade edilmeli ve ana amaç üzerine odaklanılmalıdır.

### ***2.2.2. Verilerin hazırlanması (Ön işleme)***

Her veri analizi işi yeni bir veri setlerinin toplanması, betimlenmesi ve temizlenmesiyle başlar. Bu süreçten sonra, veriler analiz edilebilir ve sonuçlara ulaşılabilir [7]. Veri kalitesi veri madenciliğinde anahtar bir konudur. Veri madenciliğinde güvenilirliğin artırılması için, veri ön işleme yapılmalıdır [8]. Verilerin hazırlanması aşaması şu adımlardan oluşmaktadır:

- Verilerin toplanması;  
Tanımlanan problem için gerekli olduğu düşünülen verilerin ve bu verilerin toplanacağı veri kaynaklarının belirlenmesi adıdır [6].
- Verilerin temizlenmesi;  
Eksik verilerin tamamlanması, aykırı değerlerin tespit edilmesi ve verilerdeki tutarsızlıkların giderilmesi gibi işlemlerden oluşmaktadır.
- Verilerin birleştirilmesi;  
Farklı veri tabanlarındaki verilerin (veri ambarında) birleştirilmesi işlemidir. En genel anlamında, veri ambarı farklı kaynaklarda tutulan verilerin ortak bir çatı altında birleştirilerek, verilerin zaman boyutunda birbiri ile ilişkilendirilmesini sağlayan, tutarlı ve doğru verilerin yer aldığı sistemdir [3].
- Verilerin dönüştürülmesi;  
Veri madenciliği çalışmasında kullanacağımız model, teknik ve algoritmalar belirli türdeki verilerle çalışmaktadır. Veri dönüştürme; düzeltme, birleştirme, genelleştirme ve normalleştirme gibi işlemleri ifade eder. Bu aşamada, veriler veri madenciliği analizi için uygun formlara dönüştürülür.
- Verilerin indirgenmesi;  
Büyük hacimli veri kümesinden daha küçük hacimli veri kümesinin elde edilmesi (kümeleme, sınıflandırma) işlemidir.

### ***2.2.3. Modelin kurulması ve değerlendirilmesi***

Tanımlanan problem için en uygun modelin bulunabilmesi için, çok sayıda modelin denenmesi gerekebilir. Bu nedenle, veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varıncaya kadar yinelenen bir süreçtir.

### ***2.2.4. Modelin kullanılması***

Kurulan ve geçerliliği kabul edilen model ile doğrudan bir uygulama yapılabileceği gibi, bir başka uygulamanın alt parçası olarak da kullanılabilir.

### ***2.2.5. Modelin izlenmesi***

Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modelin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmesini gerektirebilir.

### **2.3. Birliktelik kuralları**

Veri madenciliği modellerinden biri olan birliktelik kuralları; büyük veri kümeleri arasında birliktelik davranışlarını bulmaya yaramaktadır. Bu kurallar, bilinmeyen ilişkilerin tespit edilmesine ve daha etkin sonuçların belirlenmesi için uygun karar vermeye imkân sağlamaktadır [9].

Birçok birliktelik kuralları keşif sistemleri, Apriori algoritmasında olduğu gibi sık öge kümesi stratejisini kullanır. Bu yaklaşım belirli bir minimum destek gerektirir. Destek eşliğine eşit ya da destek eşliğini aşan tüm veri setleri keşfedildikten sonra birliktelik kuralları oluşturulur. Bu strateji düşük (seyrek) yoğunluktaki veri setleri için çok etkili olup analiz süreci dışında tutulmalarına imkân sağladığı için, gereksiz veri setleri ile zaman kaybedilmemektedir.

Birliktelik kurallarında ilk adım sık öge setlerini tespit etmektir. Bu öge setleri sıklıkla bir araya gelen ve araştırılması gereken öge setleridir. Arama uzayının üstel bir ölçek olmasından dolayı bu aşama, hesaplama gücü açısından çok zorlu olup etkili bir algoritma ve veri yapılarının kullanımına gereksinim duyulur. Gerçek zamanlı bir veri ile uğraşırken bu faktörler gerçekten çok önemli bir hale gelmektedir [13].

Birliktelik kuralları, bir “A” hareketi (öncül olay) meydana geldiğinde “B” hareketinin (sonuç) de meydana gelme olasılığını bularak iki hareket arasındaki ilişkiyi tespit etmeye çalışmaktadır [12]. “A” değerini içeren olayların, “B” değerini de içermeye durumu, yani birliktelik kuralı,  $A \Rightarrow B$  biçiminde gösterilmektedir [11].

Bir birliktelik kuralının kural desteği ( $s$ ),  $A \cup B$ 'yi kapsayan hareketlerin sayısının, veri tabanındaki toplam hareketlerin sayısına ( $N$ ) yüzde cinsinden oranıdır. Kural desteği aşağıdaki şekilde formüle edilir [12]:

$$\text{Kural desteği } (s) = \frac{n(A \cup B)}{n(N)}$$

Güven ( $\alpha$ ) ise,  $A \cup B$ 'yi kapsayan hareketlerin sayısının,  $A$ 'yı kapsayan hareketlerin sayısına yüzde cinsinden oranı yani  $A$ 'yı kapsayan hareketlerin yüzde kaçının  $B$ 'yi de kapsadığının ifadesidir. Güven aşağıdaki şekilde formüle edilir [12]:

$$\text{Güven } (\alpha) = \frac{n(A \cup B)}{n(A)}$$

Kural desteği ( $s$ ) ne kadar yüksekse kural o kadar sık tekrar etmektedir; güven ( $\alpha$ ) ne kadar yüksekse kuralın istisnası o kadar az demektir. Bir diğer deyişle; güven seviyesi, kuralın gücünü ifade etmektedir.

### 3. Analiz

Hane halkı anket çalışmalarında, çalışma alanındaki toplam nüfus ile görüşmek şüphesiz en doğru yöntem olacaktır. Fakat bu durum gerek bütçe gerekse zaman açısından imkânsızdır. Çok az kişi ile görüşme yapmak ise hata payının artmasına sebep olmaktadır. Bu nedenle örneklem büyüklüğünün dengeli olması oldukça önemlidir. Örneklem büyüklüğü için kullanılan klasik yöntemler yeterli olmamaktadır. Ayrıca alt tabaka olarak belirlenen analiz bölgelerine ilişkin ayrı ayrı örneklem büyüklüğü belirlenmesi kendi içinde belirli bir örneklem ile analiz edilmesi gerektiğinden bu yöntemde ortaya çıkacak örneklem sayısı bütçe ve zaman açısından değerlendirildiğinden oldukça yüksek olacaktır. Bu yöntem yerine örneklem hata paylarının çok yüksek olmayacağı en uygun örneklem büyüklüğü belirlenmeli ve her bölgede analiz için minimum bir sayıda gözlem yaparak analiz edilebilir sayıya ulaşmak hedeflenmelidir.

Analiz verileri yolculuk bazlı ele alınmış olup tanımlayıcı istatistikler aşağıdaki gibidir: Tablo 1’de, yapılan yolculukların ilçelere göre dağılımı yüzdelerle ifade edilmiştir. İncelenen 25116 yolculuğun %33,9’u Dulkadiroğlu ilçesine aitken %66,1’i Onikişubat ilçesine aittir. Yapılan yolculukların %82,2’si zon dışına yapılmaktadır (Tablo 2). Yolculukların büyük kısmı (%96,6’sı) ev bazlı (başlangıç veya bitiş noktası ev) yolculukları temsil etmektedir (Tablo 3). Gelir aralıklarına ait %95,6 geçerli veri bulunuyorken, yapılan yolculukların 4,4’ünde gelir aralığı ile ilgili veri bulunmamaktadır (Tablo 4). Yolculukların ulaşım tiplerine göre dağılımları Tablo 5’te belirtilmiştir. İncelendiğinde; kamyon, tır vb. taşıtların frekansının çok az olduğu gözlemlenmiştir. Tablo 6’da görüldüğü üzere, yolculukların %35,2 si araçsız olarak gerçekleşmiştir. Bu sonuca, Tablo 5’te yaya olarak yapılan yolculukları incelediğimizde de ulaşabiliyoruz. Yolculukların süre aralıkları 10 dakika olacak şekilde 13 kategoriye ayrılmıştır (Tablo 7).

**Tablo 1.** Yolculukların İlçelere Göre Dağılımı

	Frekans	Yüzdelerle Değer	Geçerli Yüzdelerle Değer	Kümülatif Yüzdelerle Değer
DULKADİROĞLU	8514	33,9	33,9	33,9
ONİKİŞUBAT	16602	66,1	66,1	100,0
Toplam	25116	100,0	100,0	

**Tablo 2.** Yolculukların Zon İçi ve Zon Dışı Dağılımları

	Frekans	Yüzdelerle Değer	Geçerli Yüzdelerle Değer	Kümülatif Yüzdelerle Değer
ZON DIŞI	20636	82,2	82,2	82,2
ZON İÇİ	4480	17,8	17,8	100,0
Toplam	25116	100,0	100,0	

**Tablo 3.** Yolculukların Amaçlarına Göre Dağılımı

	Frekans	Yüzdelerik Değer	Geçerli Yüzdelerik Değer	Kümülatif Yüzdelerik Değer
HBW (Ev Bazlı İş Yolculukları)	7258	28,9	28,9	28,9
HBS (Ev Bazlı Okul Yolculukları)	9444	37,6	37,6	66,5
HBO (Ev Bazlı Diğer Yolculuklar)	7572	30,1	30,1	96,6
NHB (Ev Bazlı olmayan)	842	3,4	3,4	100,0
Toplam	25116	100,0	100,0	

**Tablo 4.** Yolculukların Gelir Aralıklarına Göre Dağılımı

	Frekans	Yüzdelerik Değer	Geçerli Yüzdelerik Değer	Kümülatif Yüzdelerik Değer
1.000 TL altı	1464	5,8	6,1	6,1
1.001 - 2.000 TL	3799	15,1	15,8	21,9
2.001 - 3.000 TL	11009	43,8	45,8	67,7
3.001 - 4.000 TL	2888	11,5	12,0	79,8
4.001 - 5.000 TL	2280	9,1	9,5	89,2
5.001 - 7.500 TL	1379	5,5	5,7	95,0
7.500 - 10.000 TL	864	3,4	3,6	98,6
10.001 - 15.000 TL	273	1,1	1,1	99,7
15.001 - 20.000 TL	52	,2	,2	99,9
20.000 TL üstü	15	,1	,1	100,0
Toplam geçerli veri	24023	95,6	100,0	
Kayıp veri	1093	4,4		
Toplam	25116	25116	100,0	

**Tablo 5.** Yolculukların Ulaşım Tipine Göre Dağılımı

	Frekans	Yüzdelerik Değer	Geçerli Yüzdelerik Değer	Kümülatif Yüzdelerik Değer
YAYA	8831	35,2	35,2	35,2
ÖZEL ARAÇ	7178	28,6	28,6	63,7
SERVİS	4275	17,0	17,0	80,8
TOPLU TAŞIMA	4813	19,2	19,2	99,9
DİĞER(KAMYON,TIR,VS)	19	,1	,1	100,0
Toplam	25116	100,0	100,0	

**Tablo 6.** Yolculukların Araçlı Gerçekleşme Dağılımı

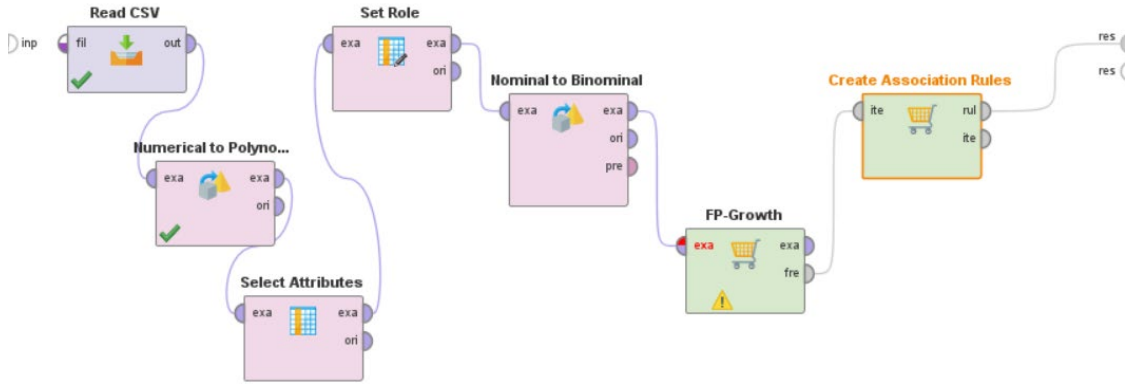
	Frekans	Yüzdelerik Değer	Geçerli Yüzdelerik Değer	Kümülatif Yüzdelerik Değer
ARAÇSIZ	8831	35,2	35,2	35,2
ARAÇLI	16285	64,8	64,8	100,0
Toplam	25116	100,0	100,0	

**Tablo 7.** Yolculukların Süre Aralığına Göre Dağılımı

	Frekans	Yüzdelerik Değer	Geçerli Yüzdelerik Değer	Kümülatif Yüzdelerik Değer
0-10	5705	22,7	22,7	22,7
11-20	6326	25,2	25,2	47,9
21-30	6468	25,8	25,8	73,7
31-40	1567	6,2	6,2	79,9
41-50	1575	6,3	6,3	86,2
51-60	2295	9,1	9,1	95,3
61-70	253	1,0	1,0	96,3
71-80	206	,8	,8	97,1
81-90	382	1,5	1,5	98,7
91-100	50	,2	,2	98,8
101-110	54	,2	,2	99,1
111-120	121	,5	,5	99,5
121 VE ÜZERİ	114	,5	,5	100,0
Toplam	25116	100,0	100,0	

Yukarıdaki tablolarda geçerli yüzdelerik değer, kayıp veri olması halinde, yalnızca mevcut değerler üzerinden hesaplanan yüzdelerik değerleri temsil etmektedir.

Şekil 1’de, RapidMiner programı ile oluşturulan analiz süreci yer almaktadır.



**Şekil 1.** RapidMiner analiz süreci modeli

Veriler, csv dosyası haline getirildikten sonra Rapidminer ile incelenmiştir. Sayısal veriler bu anlamda kategorik hale getirilip, birliktelik kuralı ile analiz edilmiştir. Oluşturulan süreçte, analiz edilecek değerler “Select atributes” komutu ile seçilip, “Set role” komutu ile “zonici” sütunu etiketlenmiş ve bu sonucu etkileyen öncül değerlerin bulunması hedeflenmiştir. Bu anlamda kategorik hale getirilen veriler “polynomial” halden “binomial” hale (doğru – yanlış) dönüştürülerek daha basite indirgenmiştir. Tüm bu işlemleri takiben, son olarak birliktelik kuralı ile analiz yapılması için gereken algoritmalar sürece dahil edilerek analiz süreci başlatılmıştır.



#### 4. Sonuç

Verilerin, birliktelik kuralı ile analiz edilerek incelenmesi sonucunda Tablo 8 elde edilmiş ve yorumlanmıştır.

**Tablo 3.** RapidMiner ile verilerin birliktelik kuralı ile analizi

Öncüller (Premises)	Sonuçlar (Conclusions)	Destek (Support)	Güven (Confidence)
araçlı = 1, yolculuk amacı = 1	ZON DIŞI	0,24	0,97
araçlı = 1, yolculuğun başladığı ilçe = 1	ZON DIŞI	0,23	0,97
araçlı = 1, yolculuğun bittiği ilçe = 1	ZON DIŞI	0,23	0,97
araçlı = 1, gelir aralığı = 3	ZON DIŞI	0,25	0,96
araçlı = 1	ZON DIŞI	0,62	0,96
araçlı = 1, yolculuğun bittiği ilçe = 2	ZON DIŞI	0,37	0,94
araçlı = 1, yolculuğun başladığı ilçe = 2	ZON DIŞI	0,37	0,94
ulaşım tipi = 2	ZON DIŞI	0,27	0,93
araçlı = 1, ulaşım tipi = 2	ZON DIŞI	0,27	0,93
yolculuk amacı = 1	ZON DIŞI	0,27	0,93
araçlı = 1, yolculuğun bittiği ilçe = 2, yolculuğun başladığı ilçe = 2	ZON DIŞI	0,26	0,92
süre aralığı = 3	ZON DIŞI	0,24	0,91
yolculuk amacı = 3	ZON DIŞI	0,25	0,83

Bu veriler ışığında örnek olması açısından birliktelik kuralı çerçevesinde aşağıda örnek bir değerlendirme ve yorumlama yapılmıştır.

Öncül - 1 : araçlı = 1 (Araçlı)  
 Öncül - 2 : yolculuk amacı = 1 (Ev Bazlı İş Yolculukları)  
 Sonuç : Zon Dışı  
 Destek (Support) : 0,24  
 Güven (Confidence) : 0,97

Ev bazlı iş yolculuklarının araçla gerçekleşmelerinin zon dışı olanları, toplam yolculukların %24'ünü kapsamaktadır ve bu (Ev bazlı iş yolculuklarının araçla gerçekleşmelerinin zon dışı olanlarının) yolculukların, ev bazlı iş yolculuklarının araçla gerçekleşmelerine oranı %97'dir.

Bura da sayısal değerleri ifade edilen Destek ve Güven değerleri aşağıdaki şekilde hesaplanmıştır.

$$\text{Destek} = \frac{\text{Ev bazlı iş yolculuklarının araçla gerçekleşmelerinin zon dışı olanları}}{\text{Toplam yolculuk sayısı}} = \frac{6147}{25116} = 0,24$$

$$\text{Güven} = \frac{\text{Ev bazlı iş yolculuklarının araçla gerçekleşmelerinin zon dışı olanları}}{\text{Ev bazlı iş yolculuklarının araçla gerçekleşmeleri}} = \frac{6147}{6333} = 0,97$$

Bu çalışma kapsamında yapılan veri madenciliği analizlerden elde edilen sonuçlar, farklı kategoriler üzerinde de birliktelik kuralları ile analizlerinin yapılması ile çok daha geniş ve zengin içerikli hale getirilecek olup, elde edilen bu geniş içerikli sonuçların karşılaştırılması ile en anlamlı birliktelikler ortaya konulmuş olacaktır.

## **Kaynakça**

- [1] Ganesh S. Data Mining: Should it be included in the ‘Statistics’ curriculum?, The Sixt International Conference on Teaching Statistics, Cape Town, South Africa, 2002.
- [2] Oğuzlar A. Veri Madenciliğine Giriş, Ekin Kitabevi, Bursa, 2004.
- [3] Şentürk A. Veri Madenciliği Kavram ve Teknikler, Ekin Kitabevi, Bursa, 2006.
- [4] Koyuncugil, AS. Veri Madenciliği ve Sermaye Piyasalarına Uygulaması, Sermaye Piyasası Kurulu Araştırma Raporu, 2007.
- [5] Yıldırım P, Uludağ M, Görür A. Hastane Bilgi Sistemlerinde Veri Madenciliği, Akademik Bilişim Kongresi, Çanakkale Onsekiz Mart Üniversitesi, 2007.
- [6] Akpınar H. Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, İstanbul Üniversitesi İşletme Fakültesi Dergisi, Cilt 29, Sayı 1/Nisan; 2000, s. 1–22.
- [7] Dasu T Johnson T. Exploratory Data Mining and Data Cleaning, John Wiley & Sons Publication, New Jersey, USA; 2003.
- [8] Oğuzlar A. Veri Önleme, Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Sayı 21/ Temmuz-Aralık; 2003, s. 67–76.
- [9] Liao CW Perng YH. Data mining for occupational injuries in the Taiwan construction industry, Safety Science; 2007.
- [10] Mirabadi A Sharifian S. Application of association rules in Iranian Railways (RAİ) accident data analysis. Safety Science, 2010.
- [11] Agrawal R Imielinski T Swami A. Mining association rules between sets of items in large databases, Association for Computing Machinery (ACM) SIGMOD. 1993; 207-216.
- [12] Bothorel G Serrurier M Hurter C. Utilisation d’outils de visual data mining pour l’exploration d’un ensemble de règles d’association, IHM '11 23rd French Speaking Conference on Human-Computer Interaction, 2011.
- [13] Taniar D. Mobile computing: concepts, methodologies, tools, and applications: Concepts, Methodologies, Tools, and Applications (Vol. 1), IGI Global. 2008.