

**VERİ MADENCİLİĞİ YÖNTEMLERİ İLE ANA HARCAMA
GRUPLARININ PAYLARININ TAHMİNİ**

**ESTIMATION OF MAIN EXPENDITURE GROUPS'
PORTION WITH DATA MINING METHODS**

Levent AHİ

Prof. Dr. Özgür YENİAY
Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2015

Levent AHI'nin hazırladığı “**Veri Madenciliği Yöntemleri İle Ana Harcama Gruplarının Paylarının Tahmini**” adlı bu çalışma aşağıdaki jüri tarafından **İSTATİSTİK ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. H. Hasan ÖRKCÜ
Başkan

.....

Prof. Dr. Özgür YENİAY
Danışman Üye

.....

Yrd. Doç. Dr. Murat Büyükyazıcı
Üye

.....

Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından **YÜKSEK LİSANS TEZİ** olarak onaylanmıştır.

Prof. Dr. Fatma SEVİN DÜZ
Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada;

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversite ya da başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

22/06/2015

Levent AHİ

ÖZET

VERİ MADENCİLİĞİ YÖNTEMLERİ İLE ANA HARCAMA GRUPLARININ PAYLARININ TAHMİNİ

Levent AHI

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Özgür YENİAY

Haziran 2015, 157 sayfa

Bir ülkenin sosyal istatistik altyapısını gelir, eğitim, sağlık, işgücü ve tüketim alanında yapılan araştırmalar oluşturmaktadır. Kurumların ve işletmelerin veri tabanlarında bu alanlarda yapılmış yüzlerce araştırma verisi bulunmaktadır. Kurumlar ve işletmeler, sayısal teknolojilerin gelişmesiyle birlikte her kayıtlarını elektronik olarak biriktirmeye başlamış ve bunun sonucunda da akıl almaz derecede veri birikimi meydana gelmiştir. Veri tabanlarında ve veri ambarlarında depolanan bu veri yığınları arasından anlamlı ilişkilerin, kalıpların ve eğilimlerin ortaya çıkartılması ihtiyacı doğmuş, gelecek adına doğru tahminlerin yapılması önem kazanmıştır.

Veri madenciliği uygulamaları işte bu noktada, istatistiksel analiz ve modellemeler ile makine öğrenimi yöntemlerinin kullanılması ile devreye girmiştir.

Veri madenciliği, eldeki verilerden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılmasıdır.

Veri madenciliğini bu anlamda istatistiksel bir yöntemler serisi olarak görmek mümkün olabilir. Ancak veri madenciliği, geleneksel istatistikten birkaç yönden farklılık göstermektedir. Veri madenciliğinde amaç, kolaylıkla mantıksal kurallara ya da görsel sunumlara çevrilebilecek nitel modelleri çıkarmaktır. Bu bağlamda, veri madenciliği insan merkezlidir ve bazen insan – bilgisayar arayüzü birleştirilir.

Her ne kadar kurumların ve işletmelerin veri tabanlarında yüzlerce araştırma verisi bulunsa da elde bulunan araştırmalar istenilen tüm değişkenlere sahip olmayabilir. Bazı durumlarda istenilen değişkenleri aynı anda barındıran veri setlerine ihtiyaç duyulmaktadır. İstenilen değişkenlerin tümünü aynı araştırmadan elde edebilmek için yeni araştırma yapmak maliyetli olduğu için tercih edilmeyen bir yöntemdir. Son yıllarda bu aşamadaki eksikliği gidermek için model bazlı bazı yaklaşımlar sıklıkla kullanılmaya başlanmıştır.

Model bazlı yaklaşıma örnek olarak istatistiksel eşleştirme gösterilebilir. İstatistiksel eşleştirme, iki veya daha fazla kaynaktan toplanan değişken veya göstergeler için birleşik istatistiksel bilgi üretmek için kullanılan bir yaklaşımdır. Bu yaklaşımın faydası, araştırma maliyeti ve cevaplayıcı yükü yaratmadan mevcut veriyi kullanma ve farklı araştırmalar arasında tutarlılık kontrolü yapabilme imkanı sağlamasıdır.

Bu tez kapsamında, öncelikle Türkiye İstatistik Kurumu'nun (TÜİK) yaptığı Hanehalkı Bütçe Araştırması (HBA) ve Gelir ve Yaşam Koşulları Araştırması (GYKA) çalışmaları hakkında detaylı bilgiler verilmiş, veri madenciliği süreci hakkında genel tanımlamalar yapılmış ve bu kapsamda kullanılan yöntemlerden bahsedilmiştir. Çalışmada veri madenciliği yazılımı IBM SPSS Modeler 16.0 ile TÜİK'in 2011 yılında yapmış olduğu HBA kapsamına dahil olan 9 918 hanehalkının, yine TÜİK'in 2012 yılında yapmış olduğu GYKA kapsamındaki 17 562 hanehalkının ortak değişkenleri kullanılarak GYKA haneleri için Amaca Göre Bireysel Tüketimin Sınıflandırılması (COICOP) 12 ana harcama grubu bazında harcama değişkenlerini tahmin etmeye yönelik veri madenciliği yöntemleri kullanılarak istatistiksel eşleştirme ile ilgili bir uygulama gerçekleştirilmiştir.

Anahtar Kelimeler: Veri madenciliği, impütasyon, hanehalkı bütçe araştırması, gelir ve yaşam koşulları araştırması, istatistiksel eşleştirme

ABSTRACT

ESTIMATION OF MAIN EXPENDITURE GROUPS' PORTION WITH DATA MINING METHODS

Levent AHİ

Master of Science, Department of Statistics

Supervisor: Prof. Dr. Özgür YENİAY

June 2015, 157 pages

A country's social statistical infrastructure constitute research conducted in the field of income, education, health, labour force and consumption. Institutions and enterprises have hundreds of research done in these areas in their database. Institutions and enterprises with the development of digital technologies have begun to accumulate all records electronically and as a result, in the incredibly accumulation of data has occurred. Stored in databases and data warehouses meaningful relationships among the data stack, patterns and trends revealed necessitated, making accurate predictions for the future, has gained importance.

Data mining applications at this point, statistical analysis and machine learning by modeling "artificial intelligence" has been activated through the use of technology.

Data mining is the process of extracting potentially useful information, not clear and previously unknown, from the available data.

Data mining in this sense can be seen as a series of statistical methods. However, data mining, differs in several ways from traditional statistics. Purpose of data mining is to extract qualitative models that can easily be translated into logical rules or visual presentation. In this context, data mining is human -centered and human - computer interface are sometimes combined.

Although institutions and enterprises have hundreds of survey data in the database available survey data may not contain all desired variables. In some cases, variables in studies conducted in these areas are needed to contain in one data set simultaneously. Making a new research to obtain the desired variable from the same study is not preferred because of its cost. In recent years, some model-based approaches have been used frequently to eliminate shortcomings in this stage.

An example of a model-based approach can be shown statistical matching. Statistical matching is a model-based approach for providing joint statistical information based on variables and indicators collected through two or more sources. The potential benefits of this approach lie in the possibility to enhance the complementary use and analysis of existing data sources, without further increasing costs and response burden.

In this thesis, first detailed information about Household Budget Survey (HBS) and Survey of Income and Living Conditions (SILC) conducted by Turkish Statistical Institute (TURKSTAT) are presented, general descriptions are made about the data mining process and methods used in this context are mentioned. Common variables included 9 918 households in the scope of HBS conducted by TURKSTAT in 2011 and 17 562 households in the scope of SILC conducted by TURKSTAT in 2012 are used with the help of data mining software IBM SPSS Modeler 16.0. An application is conducted by data mining methods relation with statistical matching to estimate variables on the basis of Classification of Individual Consumption According to Purpose (COICOP) 12 main expenditure groups for households in the scope of SILC.

Keywords: Data mining, imputation, household budget survey, survey of income and living conditions, statistical matching

TEŐEKKÜR

Çalıőmamın her aőamasında beni yönlendiren, destekleyen ve cesaretlendiren Sayın Prof. Dr. Özgür YENİAY'a, çalıőmam boyunca gösterdikleri destek ve yardımları için Daire Başkanım Sayın Meral DAŐKIRAN ve iő arkadaşım Kadir Korhan BABADAĐ'a, varlıkları ile bana güç veren sevgili eőim Nazan AHİ, biricik ođlum Yiđit Efe AHİ'ye, bana her konuda őartsız, koőulsuz destek veren ve beni bugünlere getiren deđerli aileme ve burada adını yazamadıđım diđer bütün arkadaş ve dostlarıma teőekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER	vi
ŞEKİLLER	x
ÇİZELGELER	xv
KISALTMALAR	xvii
1. GİRİŞ	1
2. GENEL BİLGİLER	6
2.1. Gelir ve Yaşam Koşulları Araştırması	6
2.1.1. Açıklama	6
2.1.2. Kapsam	7
2.1.2.1. Coğrafi Kapsam	7
2.1.2.2. Kapsanan Kitle	7
2.1.2.3. Kapsanan Konular	7
2.1.3. Örneklem Tasarımı	8
2.1.3.1. Örneklem Yöntemi, Birimi ve Tahmin Boyutu	8
2.1.3.2. Örneklem Büyüklüğü ve Cevapsızlık Oranı	8
2.1.3.3. Panel Anket Uygulaması	8
2.1.4. Referans Dönemleri	9
2.1.5. Veri Derleme Yöntemi	9
2.1.6. Tanım ve Kavramlar	10
2.1.6.1. Hanehalkı İle İlgili Tanımlar	10
2.1.6.2. Gelir İle İlgili Tanımlar	11
2.2. Hanehalkı Bütçe Araştırması	15
2.2.1. Açıklama	15
2.2.2. Anketin Örneklem Planı ve Tahmin Yöntemi	17
2.2.2.1. Coğrafi Kapsam	17
2.2.2.2. Kapsanan Kitle	17
2.2.2.3. Tüketim Harcamaları ve Gelirin Kapsamı	17
2.2.2.4. Örneklem Çerçevesi, Birimi ve Yöntemi	17
2.2.3. Tanım ve Kavramlar	18

2.2.4. Hanehalkı Tüketim Harcamalarının Kapsamı	19
2.2.5. Alan Uygulaması Yöntemi	20
3. VERİ MADENCİLİĞİ	22
3.1. Tanımlar	22
3.1.1. Bilgi.....	22
3.1.2. Veri	23
3.1.3. Veri Tabanı	23
3.1.4. Veri Ambarı.....	24
3.2. Veri Tabanında Bilgi Keşfi Süreci	24
3.2.1. Uygulama Alanının Anlaşılması ve Geliştirilmesi.....	25
3.2.2. Bir Veri Seti Seçme ve Oluşturma	25
3.2.3. Önışleme ve Temizleme	25
3.2.4. Veri İndirgeme	25
3.2.5. Uygun Bir Veri Madenciliği Programı ve Algoritması Çalıştırmak	25
3.2.6. Yorumlama, Doğrulama ve Bilgi Keşfini Kullanmak.....	25
3.3. Veri Madenciliği Tarihçesi.....	25
3.4. Veri Madenciliği Tanımı	26
3.5. Veri Madenciliği Süreci	27
3.5.1. İşı Anlamak	28
3.5.2. Veriyi Anlamak.....	29
3.5.3. Veri Hazırlığı.....	29
3.5.3.1. Veri Temizleme.....	29
3.5.3.2. Veri Birleştirme	31
3.5.3.3. Veri Dönüştürme.....	31
3.5.3.4. Veri İndirgeme	32
3.5.4. Modelleme	34
3.5.5. Deęerlendirme	34
3.5.6. Uygulama	34
3.6. Veri Madenciliği Ne Deęildir?	34
3.7. Tahminleme.....	35
3.7.1. Tahmin Kavramı	35
3.7.2. Tahmin Yöntemleri	35
3.7.2.1. Nitel Tahmin Yöntemleri	35
3.7.2.2. Nicel Tahmin Yöntemleri	36

3.7.3. Tahmin Modellerinin Doğruluklarının Ölçülmesi.....	37
4. VERİ MADENCİLİĞİ MODELLERİ	38
4.1. Sınıflama ve Regresyon	39
4.1.1. Karar Ağaçları.....	40
4.1.2. Yapay Sinir Ağları	41
4.1.3. Genetik Algoritmalar	42
4.1.4. K-En Yakın Komşu	43
4.1.5. Bellek Temelli Nedenleme	43
4.1.6. Naif-Bayes	44
4.2. Kümeleme	44
4.3. Birlikte Kuralları ve Ardışık Zamanlı Örüntüler	46
5. UYGULAMA	48
5.1. Veri Madenciliği Süreçleri	48
5.1.1. İş Anlamak	48
5.1.2. Veriyi Anlamak.....	48
5.1.3. Veri Hazırlığı.....	49
5.1.4. Modelleme	53
5.1.4.1. Gıda ve Alkolsüz İçecekler	54
5.1.4.2. Alkollü İçecekler, Sigara ve Tütün	57
5.1.4.3. Giyim ve Ayakkabı	60
5.1.4.4. Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar	63
5.1.4.5. Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri	66
5.1.4.6. Sağlık	69
5.1.4.7. Ulaştırma	72
5.1.4.8. Haberleşme	75
5.1.4.9. Eğlence ve Kültür	78
5.1.4.10. Eğitim Hizmetleri.....	81
5.1.4.11. Lokanta ve Oteller	84
5.1.4.12. Çeşitli Mal ve Hizmetler	87
5.1.4.13. GYKA Haneleri İçin Harcama Değerlerinin Tahmin Edilmesi.....	91
5.1.5. Değerlendirme	103
5.1.6. Uygulama	103
6. SONUÇ VE TARTIŞMA.....	105
KAYNAKLAR	108

EKLER.....	111
ÖZGEÇMİŞ	135

ŞEKİLLER

	<u>Sayfa</u>
Şekil 1.1.	İstatistiksel Eşleştirme Süreci.....4
Şekil 2.1.	Rotasyonel Panel Anket Uygulaması9
Şekil 3.1.	Veri Tabanında Bilgi Keşfi Süreci.....24
Şekil 3.2.	CRISP DM (Cross Industry Standard Process Model for Data Mining) Süreci28
Şekil 4.1.	Yapay Sinir Ağlarının Katmanları42
Şekil 5.1.	Hellinger Uzaklığı52
Şekil 5.2.	% 10'luk Harcama Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (% 50 örneklemeli, TL/ay)54
Şekil 5.3.	% 10'luk Harcama Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay).....55
Şekil 5.4.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay).....56
Şekil 5.5.	Hanehalkı Tipine Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay)56
Şekil 5.6.	% 10'luk Harcama Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (% 50 örneklemeli, TL/ay).....57
Şekil 5.7.	% 10'luk Harcama Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)58
Şekil 5.8.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)59
Şekil 5.9.	Hanehalkı Tipine Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)59
Şekil 5.10.	% 10'luk Harcama Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (% 50 örneklemeli, TL/ay).....60
Şekil 5.11.	% 10'luk Harcama Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay)61
Şekil 5.12.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay).....62
Şekil 5.13.	Hanehalkı Tipine Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay).....62

Şekil 5.14.	% 10'luk Harcama Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (% 50 örnekleme, TL/ay)	63
Şekil 5.15.	% 10'luk Harcama Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay).....	64
Şekil 5.16.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay).....	65
Şekil 5.17.	Hanehalkı Tipine Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay).....	65
Şekil 5.18.	% 10'luk Harcama Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (% 50 örnekleme, TL/ay).....	66
Şekil 5.19.	% 10'luk Harcama Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)	67
Şekil 5.20.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)....	68
Şekil 5.21.	Hanehalkı Tipine Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)	68
Şekil 5.22.	% 10'luk Harcama Gruplarına Göre Ortalama Sağlık Harcamaları (% 50 örnekleme, TL/ay)	69
Şekil 5.23.	% 10'luk Harcama Gruplarına Göre Ortalama Sağlık Harcamaları (TL/ay).....	70
Şekil 5.24.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Sağlık Harcamaları (TL/ay)	71
Şekil 5.25.	Hanehalkı Tipine Göre Ortalama Sağlık Harcamaları (TL/ay).....	71
Şekil 5.26.	% 10'luk Harcama Gruplarına Göre Ortalama Ulaştırma Harcamaları (% 50 örnekleme, TL/ay)	72
Şekil 5.27.	% 10'luk Harcama Gruplarına Göre Ortalama Ulaştırma Harcamaları (TL/ay).....	73
Şekil 5.28.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Ulaştırma Harcamaları (TL/ay).....	74
Şekil 5.29.	Hanehalkı Tipine Göre Ortalama Ulaştırma Harcamaları (TL/ay)	74
Şekil 5.30.	% 10'luk Harcama Gruplarına Göre Ortalama Haberleşme Harcamaları (% 50 örnekleme, TL/ay).....	75
Şekil 5.31.	% 10'luk Harcama Gruplarına Göre Ortalama Haberleşme Harcamaları (TL/ay)	76

Şekil 5.32.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Haberleşme Harcamaları (TL/ay).....	77
Şekil 5.33.	Hanehalkı Tipine Göre Ortalama Haberleşme Harcamaları (TL/ay)	77
Şekil 5.34.	% 10'luk Harcama Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (% 50 örnekleme, TL/ay).....	78
Şekil 5.35.	% 10'luk Harcama Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay)	79
Şekil 5.36.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay)	80
Şekil 5.37.	Hanehalkı Tipine Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay).....	80
Şekil 5.38.	% 10'luk Harcama Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (% 50 örnekleme, TL/ay).....	81
Şekil 5.39.	% 10'luk Harcama Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay)	82
Şekil 5.40.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay)	83
Şekil 5.41.	Hanehalkı Tipine Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay).....	83
Şekil 5.42.	% 10'luk Harcama Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (% 50 örnekleme, TL/ay).....	84
Şekil 5.43.	% 10'luk Harcama Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay)	85
Şekil 5.44.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay)	86
Şekil 5.45.	Hanehalkı Tipine Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay).....	86
Şekil 5.46.	% 10'luk Harcama Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (% 50 örnekleme, TL/ay).....	88
Şekil 5.47.	% 10'luk Harcama Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay)	89
Şekil 5.48.	% 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay).....	90

Şekil 5.49.	Hanehalkı Tipine Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay).....	90
Şekil 5.50.	Gıda ve Alkolsüz İçecek Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	91
Şekil 5.51.	Alkollü İçecekler, Sigara ve Tütün Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	91
Şekil 5.52.	Giyim ve Ayakkabı Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	92
Şekil 5.53.	Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	92
Şekil 5.54.	Mobilya, Ev Aletleri ve Ev Bakım Hizmetler Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	93
Şekil 5.55.	Sağlık Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	93
Şekil 5.56.	Ulaştırma Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	94
Şekil 5.57.	Haberleşme Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	94
Şekil 5.58.	Eğlence ve Kültür Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	95
Şekil 5.59.	Eğitim Hizmetleri Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	95
Şekil 5.60.	Lokanta ve Oteller Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	96
Şekil 5.61.	Çeşitli Mal ve Hizmetler Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	96
Şekil 5.62.	Toplam Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)	97
Şekil 5.63.	HBA ve GYKA Harcama Sınıflarının Karşılaştırılması (%).....	98
Şekil 5.64.	Hanehalkı Tiplerine Göre Gıda ve Alkolsüz İçecekler Harcamalarının Dikey ve Yatay % Dağılımı.....	99
Şekil 5.65.	Hanehalkı Tipine Göre Alkollü İçecekler, Sigara ve Tütün Harcamalarının Dikey ve Yatay % Dağılımı	99

Şekil 5.66.	Hanehalkı Tipine Göre Giyim ve Ayakkabı Harcamalarının Dikey ve Yatay % Dağılımı	99
Şekil 5.67.	Hanehalkı Tipine Göre Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamalarının Dikey ve Yatay % Dağılımı	100
Şekil 5.68.	Hanehalkı Tipine Göre Mobilya, Ev Aletleri ve Ev Bakım Hizmetler Harcamalarının Dikey ve Yatay % Dağılımı	100
Şekil 5.69.	Hanehalkı Tipine Göre Sağlık Harcamalarının Dikey ve Yatay % Dağılımı.....	100
Şekil 5.70.	Hanehalkı Tipine Göre Ulaştırma Harcamalarının Dikey ve Yatay % Dağılımı.....	101
Şekil 5.71.	Hanehalkı Tipine Göre Haberleşme Harcamalarının Dikey ve Yatay % Dağılımı	101
Şekil 5.72.	Hanehalkı Tipine Göre Eğlence ve Kültür Harcamalarının Dikey ve Yatay % Dağılımı	101
Şekil 5.73.	Hanehalkı Tipine Göre Eğitim Hizmetleri Harcamalarının Dikey ve Yatay % Dağılımı	102
Şekil 5.74.	Hanehalkı Tipine Göre Lokanta ve Oteller Harcamalarının Dikey ve Yatay % Dağılımı	102
Şekil 5.75.	Hanehalkı Tipine Göre Çeşitli Mal ve Hizmetler Harcamalarının Dikey ve Yatay % Dağılımı.....	102
Şekil 5.76.	Toplam Harcama Medyan Değerinin %40'ına Göre Hesaplanan Yoksul Hane Oranı (%)	103
Şekil 5.77.	Toplam Harcama Medyan Değerinin % 50'sine Göre Hesaplanan Yoksul Hane Oranı (%).....	103
Şekil 5.78.	Toplam Harcama Medyan Değerinin %60'ına Göre Hesaplanan Yoksul Hane Oranı (%)	103
Şekil 5.79.	Toplam Harcama Medyan Değerinin %70'ine Göre Hesaplanan Yoksul Hane Oranı (%)	103

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 1.1. GYKA ve HBA Karşılaştırması	3
Çizelge 2.1. Hanehalkı Kullanılabilir Net Gelirin Hesaplanması	14
Çizelge 3.1. Veri Madenciliği Ne Değildir?	35
Çizelge 3.2. Tahmin Doğruluğunu Ölçme İstatistikleri	38
Çizelge 5.1. GYKA ve HBA'da Yer Alan ve Türetilen Ortak Değişkenler	49
Çizelge 5.2. % 50 Örneklemeli Model Performansları.....	55
Çizelge 5.3. Model Performansları.....	55
Çizelge 5.4. Hanehalkı Kullanılabilir Geliri ve Gıda ve Alkolsüz İçecekler Harcamaları Sınıfları Arasındaki Korelasyon	56
Çizelge 5.5. % 50 Örneklemeli Model Performansları.....	58
Çizelge 5.6. Model Performansları.....	58
Çizelge 5.7. Hanehalkı Kullanılabilir Geliri ve Alkollü İçecekler, Sigara ve Tütün Harcamaları Sınıfları Arasındaki Korelasyon	59
Çizelge 5.8. % 50 Örneklemeli Model Performansları.....	61
Çizelge 5.9. Model Performansları.....	61
Çizelge 5.10. Hanehalkı Kullanılabilir Geliri ve Giyim ve Ayakkabı Harcamaları Sınıfları Arasındaki Korelasyon.....	62
Çizelge 5.11. % 50 Örneklemeli Model Performansları	64
Çizelge 5.12. Model Performansları	64
Çizelge 5.13. Hanehalkı Kullanılabilir Geliri ve Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları Sınıfları Arasındaki Korelasyon	65
Çizelge 5.14. % 50 Örneklemeli Model Performansları	67
Çizelge 5.15. Model Performansları	67
Çizelge 5.16. Hanehalkı Kullanılabilir Geliri ve Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları Sınıfları Arasındaki Korelasyon.....	68
Çizelge 5.17. % 50 Örneklemeli Model Performansları	70
Çizelge 5.18. Model Performansları	70
Çizelge 5.19. Hanehalkı Kullanılabilir Geliri ve Sağlık Harcamaları Sınıfları Arasındaki Korelasyon	71
Çizelge 5.20. % 50 Örneklemeli Model Performansları	73
Çizelge 5.21. Model Performansları	73

Çizelge 5.22. Hanehalkı Kullanılabilir Geliri ve Ulaştırma Harcamaları Sınıfları Arasındaki Korelasyon	74
Çizelge 5.23. % 50 Örneklemeli Model Performansları	76
Çizelge 5.24. Model Performansları	76
Çizelge 5.25. Hanehalkı Kullanılabilir Geliri ve Haberleşme Harcamaları Sınıfları Arasındaki Korelasyon	77
Çizelge 5.26. % 50 Örneklemeli Model Performansları	79
Çizelge 5.27. Model Performansları	79
Çizelge 5.28. Hanehalkı Kullanılabilir Geliri İle Eğlence ve Kültür Harcamaları Sınıfları Arasındaki Korelasyon.....	80
Çizelge 5.29. % 50 Örneklemeli Model Performansları	82
Çizelge 5.30. Model Performansları	82
Çizelge 5.31. Hanehalkı Kullanılabilir Geliri ve Eğitim Hizmetleri Harcamaları Sınıfları Arasındaki Korelasyon.....	83
Çizelge 5.32. % 50 Örneklemeli Model Performansları	85
Çizelge 5.33. Model Performansları	85
Çizelge 5.34. Hanehalkı Kullanılabilir Geliri İle Lokanta ve Oteller Harcamaları Sınıfları Arasındaki Korelasyon.....	86
Çizelge 5.35. % 50 Örneklemeli Model Performansları	88
Çizelge 5.36. Model Performansları	89
Çizelge 5.37. Hanehalkı Kullanılabilir Geliri İle Çeşitli Mal ve Hizmetler Harcamaları Sınıfları Arasındaki Korelasyon.....	90
Çizelge 5.38. Hanehalkı Kullanılabilir Gelir ve Toplam Harcama Sınıfları Arasındaki Korelasyon	98

KISALTMALAR

TÜİK	Türkiye İstatistik Kurumu
HBA	Hanehalkı Bütçe Araştırması
GYKA	Gelir ve Yaşam Koşulları Araştırması
COICOP	Amaca Göre Bireysel Tüketimin Sınıflandırılması
CRISP DM	Cross Industry Standard Process Model for Data Mining
CAPI	Computer Assisted Personel Interview

1. GİRİŞ

Bir ülkenin sosyal istatistik altyapısını gelir, eğitim, sağlık, işgücü ve tüketim alanında yapılan araştırmalar oluşturmaktadır. Son yıllarda özellikle gelir konusunda yapılan araştırmalar ön plana çıkmaktadır. Ülke ekonomisinin ve insanların yaşam seviyelerini belirleyebilmek, çeşitli sosyal sistemlerin sağlıklı bir şekilde değerlendirmesini yapabilmek için (sosyal adalet ve eşitlik kriterlerine göre) sadece ekonomik açıdan değil sosyal açıdan da gelir dağılımı göstergelerine ihtiyaç duyulmaktadır.

TÜİK, 2006 yılından itibaren, Avrupa Birliği'ne uyum çerçevesinde, amacı gelir dağılımı yanında, yaşam koşulları, sosyal dışlanma ve görece gelir yoksulluğu gibi konularda veri üretmek olan, "panel anket" yönteminin kullanıldığı GYKA'yı uygulamaya başlamıştır.

Bu araştırmadan elde edilen bilgilerle aşağıdaki sorulara cevap aranmaktadır:

- Ülkede gelir ne kadar eşit dağılıyor ve önceki yıllara göre ne tür değişiklik gösteriyor?
- Ne kadar yoksul insan bulunuyor ve bunun bölgesel dağılımı nedir? Bunun önceki yıllara göre durumu nedir?
- Kimler yoksul? Zaman itibarıyla bir değişiklik söz konusu mu?
- Yoksul ve zenginler arasındaki bu farklılık zamanla nasıl bir değişim gösteriyor?
- Kişisel gelirler nasıl bir değişim veya geçiş gösteriyor? Bu değişimin yönü karakteristiklere ve koşullara bağlı olarak nasıl değişiyor, azalıyor mu, artıyor mu?
- Kimler sosyal dışlanma sorunu ve sürekli yoksulluk riski ile karşı karşıya?
- Maddi yoksunluğun Türkiye'deki boyutu nedir?
- İnsanlar hangi koşullarda yaşamaktadırlar?

Araştırmada, gelir, yoksulluk, sosyal dışlanma ve yaşam koşullarına ilişkin göstergeler hesaplamak üzere aşağıdaki kategorilerde bilgi derlenmektedir:

- Konut,
- Ekonomik durum,
- Sosyal dışlanma,

- Gayrimenkul sahipliđi,
- Eđitim,
- Demografi,
- Sađlık durumu,
- İřgücü durumu,
- Gelir durumu.

Gelir bilgilerinin referans dönemi “bir önceki takvim yılı”dır. Yani, 2012 uygulamasındaki gelir bilgileri 2011 yılına aittir.

Yıllık kesit araştırma sonuçlarından, Türkiye, kent, kır ve Ek.2’de verilen İstatistiki Bölge Birimleri Sınıflaması Düzey-1 için tahminler üretilmesi amaçlanmaktadır.

TÜİK’in, 2002 yılından itibaren her yıl düzenli olarak uygulamaya başladığı HBA ise hanelerin sosyo-ekonomik yapıları, yaşam düzeyleri, tüketim kalıpları hakkında bilgi veren ve toplumun ihtiyaçlarının belirlenmesi, kullanılabilir gelirin haneler ya da fertler arasında ne şekilde dağıldığının bilinmesi ve uygulanan sosyo-ekonomik politikaların geçerliliğinin test edilmesi amacıyla kullanılan en önemli kaynaklardan biridir.

HBA’nın örnekleme yapısı gereğince; yıllık anket sonuçlarından Türkiye geneli ile kırsal ve kentsel yerler ayrımında; her yılın anket sonuçlarının son iki yılın anket bilgileriyle birleştirilmesi ile de bölgesel düzeyde sadece oransal tahminler üretilmesi mümkün olmaktadır.

Buna göre; 2009, 2010 ve 2011 yıllarında yapılan HBA veri setleri bir araya getirilerek, sonuçları üç yılı kapsayan ancak 2009 ve 2010 yıllarına ait tüketim harcamaları enflasyona göre 2011 yılına artırıldıktan sonra elde edilen 2009-2010-2011 HBA sonuçları İstatistiki Bölge Birimleri Sınıflaması Düzey-1 ve Düzey-2 bazında Ek.1’de verilen haber bülteni ile kamuoyuna sunulmuştur. 2009-2010-2011 birleşik veri seti, kırsal yerleşim yerinden 9 450, kentsel yerleşim yerlerinden 20 596 örnek hanehalkı olmak üzere toplam 30 046 örnek hanehalkının tüketim harcaması bilgilerini içermektedir. Çizelge 1.1’de GYKA ve HBA bilgilerine karşılaştırmalı olarak yer verilmiştir.

Çizelge 1.1. GYKA ve HBA Karşılaştırması

	GYKA	HBA
Hane Sayısı	19 300	13 248
Cevaplı Hane Sayısı	17 562	9 918
Fert Sayısı	65 472	37 121
Yıllık Anket Sonuçlarından Yapılan Tahminler	Türkiye, kent, kır, İstatistiki Bölge Birimleri Sınıflaması Düzey-1	Türkiye, kent, kır (3 yıllık verinin birleştirilmesi ile İstatistiki Bölge Birimleri Sınıflaması Düzey-1 ve Düzey-2 bazında oransal tahminler)

Bu tez kapsamında, GYKA veri seti kullanılarak tüketim harcamalarının İstatistiki Bölge Birimleri Sınıflaması Düzey-1'e göre oransal dağılımının tahmin edilmesi ve bulunan bu oransal dağılımın 22 Kasım 2012 tarihinde "Hanehalkı Tüketim Harcaması (Bölgesel), 2011" konulu Ek.1'de verilen haber bülteni ile açıklanan sonuçlar ile karşılaştırılması amaçlanmıştır. Ancak GYKA veri seti tüketim harcaması değişkenine sahip değildir. Veri setleri istenilen tüm değişkenleri aynı anda içermeyebilir. Bazen iki, üç veya daha fazla veri setinden bu değişkenlerin bir araya getirilerek kullanılması gerekebilir. İşte bu aşamada istatistiksel eşleştirme kavramı devreye girmektedir.

İstatistiksel eşleştirme, iki veya daha fazla veri kaynağını birleştirmek için kullanılan model bazlı bir yaklaşımdır. Bu yaklaşımın faydası, ek araştırma maliyeti ve cevaplayıcı yükü yaratmadan mevcut veri kaynaklarındaki bilginin kullanılmasını ve analiz edilmesini sağlıyor olmasında yatmaktadır.

İstatistiksel eşleştirmede kullanılacak yaklaşımlar üç kategoriye ayrılmaktadır [1]:

- **Parametrik olmayan yöntemler**

Parametrik olmayan yöntemlere hot-deck yöntemi verilebilir. Verici veri setindeki kayıtlar ile alıcı veri setindeki kayıtlar bir uzaklık fonksiyonu aracılığıyla eşleştirilir. Bu işlemin zayıf yönü, verici veri setindeki kayıt sayısının alıcı veri setindeki kayıt sayısından daha küçük olması durumunda ortaya çıkmaktadır. Bu durumda verici veri setinde bulunan bir kayıt, alıcı veri

setindeki birden fazla kayıta karşılık gelebilmektedir. Bu ise eşleştirme sonucu oluşan değerlerin dağılımının gerçek dağılımlarını yansıtmama riskini artıracaktır.

- **Parametrik yöntemler**

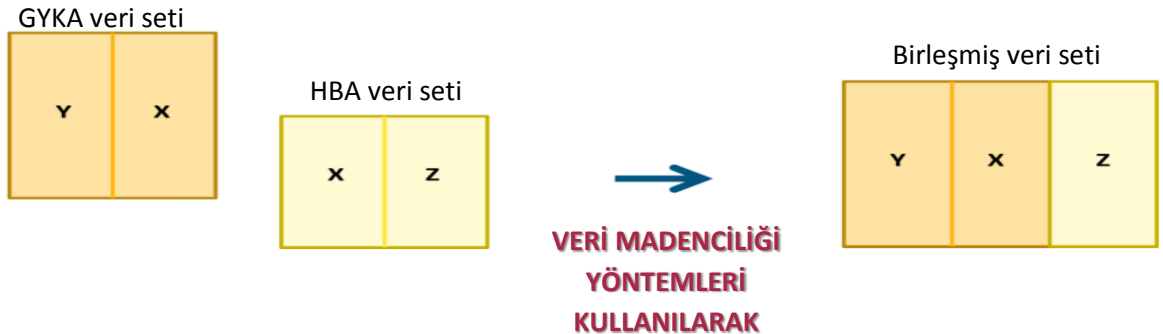
Parametrik yöntemler, bir regresyon modeli ile değerlerinin impütasyonunu içermektedir. Bu yöntemlerin güvenilirliği kurulan regresyon modelinin doğruluğuna çok bağlıdır. Buna ek olarak, ortalamaya doğru kurulan regresyon modeli, bu yöntemlerde potansiyel bir sorun olarak karşımıza çıkabilecektir.

- **Karma yöntemler**

Adından da anlaşılacağı gibi karma yöntemler, parametrik ve parametrik olmayan yöntemlerin bir kombinasyonunu içermektedir. İlk olarak eşleştirilecek değişkenin orta değerini tahmin edecek bir model kurulmaktadır. Daha sonra bir uzaklık fonksiyonu ile, ilk aşamada tahmin edilen orta değere verici veri setindeki en çok benzeyen kayıt eşleştirilmektedir.

Yukarıda bahsedilen her üç yöntemin de birbirine göre üstün veya zayıf yönleri bulunmaktadır. Bu tez çalışmasında parametrik olmayan, parametrik ve karma yöntemin bir arada kullanılması tercih edilmiş olup yöntemler arasındaki farklılıkların gözlenmesi sağlanmıştır.

İstatistiksel eşleştirme yapabilmek için bir veri setinin alıcı, diğer veri seti/setlerinin verici olarak tanımlanması gerekmektedir. Bu kapsamda, GYKA veri seti alıcı, HBA veri seti verici olarak tanımlanmıştır.



Şekil 1.1. İstatistiksel Eşleştirme Süreci

Şekil 1.1’de görüldüğü üzere, GYKA’da bulunan X değişken seti ile HBA’da bulunan X değişken seti ortak değişkenlerden oluşmaktadır. Ek.4’de verilen her iki veri setinde de yer alan ortak değişkenler kullanılarak veri madenciliği yöntemleri ile her üç istatistiksel eşleştirme yaklaşımına göre GYKA hanelerine Amaca Göre Bireysel Tüketimin Sınıflandırılması (COICOP) 12 ana harcama grubu bazında harcama değişkenleri tahmin edilmiş ve bu işlemin sonunda Y,X ve Z değişken setlerine sahip bir birleşmiş veri setine sahip olunmuştur. Böylece, üç yılın verisi bir araya getirilirken yaşanan operasyonel maliyet en aza indirilmiş, farklı anketler arasında tutarlılık kontrolü yapılmış ve aynı zamanda detaylı gelir ve harcama değişkenini içeren birleşik bir veri setine sahip olunmuştur.

İstatistiksel eşleştirmenin başarıya ulaşabilmesi için değişkenlerin ve dağılımlarının karşılaştırılabilir olması gerekmektedir. Bunun için alıcı ve verici veri setlerinin sağlanması gereken bazı şartlar bulunmaktadır [1]:

- Birimlerin tanımının uyumu,
- Referans dönemlerinin uyumu,
- Hedef nüfusun uyumu,
- Değişkenlerin dağılımlarının uyumu,
- Sınıflandırmaların uyumu.

HBA ve GYKA’da kullanılan tanım, kavram ve sınıflandırmalar aynıdır. Her iki çalışmada da Türkiye Cumhuriyeti sınırları içindeki hanelerde yaşayan tüm hanehalkı fertleri kapsama alınmıştır. Ancak çalışmada kurumsal nüfus olarak tanımlanan üniversite yurtları, misafirhane, çocuk yuvası, yetiştirme yurdu, huzurevi, özel nitelikli hastane, hapisane, kışla ve ordu evlerinde yaşayanlar kapsam dışı tutulmuştur.

2. GENEL BİLGİLER

2.1. Gelir ve Yaşam Koşulları Araştırması

2.1.1. Açıklama

Ülke ekonomisinin ve insanların yaşam seviyelerini belirleyebilmek, çeşitli sosyal sistemlerin sağlıklı bir şekilde değerlendirmesini yapabilmek için (sosyal adalet ve eşitlik kıstaslarına göre) sadece ekonomik açıdan değil sosyal açıdan da gelir dağılımı göstergelerine ihtiyaç duyulmaktadır. Gelir dağılımındaki adaletsizlik ve yoksulluk dünyanın karşılaştığı en önemli sorunlardan biri olup, gelir dağılımı sorunu artık sadece sıradan bir ekonomi sorunu olmaktan ziyade politik ve sosyal bir sorun olarak görülmeye başlanmıştır. Gelir eşitsizliğinin değişimini izlemeye yönelik çalışmalar yanında, gelir dağılımı sorununun yoksulluk sorununa indirgenmiş olmasıyla “gelir yoksulluğu”, “sosyal imkan yoksulluğu” gibi yeni kavramlara ilişkin veri üretme ihtiyacı da doğmuştur [2].

TÜİK, ilki 1987 yılında olmak üzere, 2005 yılına kadar gelir dağılımı istatistiklerini HBA'dan üretmiştir. 1994 yılında gerçekleştirilen ikinci uygulamanın ardından 2002 yılından itibaren yıllık düzenli anket uygulamasına geçilmiştir. HBA sonuçlarını içeren son gelir dağılımı haber bülteni 2005 yılına aittir [2].

2006 yılından itibaren, Türkiye’de gelirin hanehalkları ve fertler arasındaki dağılımını ortaya koymak, insanların yaşam koşulları, sosyal dışlanma ve yoksulluğu gelir boyutu ile ölçmek ve profilini belirlemek üzere periyodik olarak her yıl uygulanan ve “panel anket” yönteminin kullanıldığı GYKA uygulanmaya başlanmıştır. Avrupa Birliği uyum çalışmaları kapsamında uygulanmaya başlanan araştırma ile Avrupa Birliği ülkeleri ile karşılaştırılabilir gelir dağılımı, gelire göre göreceli yoksulluk, yaşam koşulları ve sosyal dışlanma konularında veri üretmek amaçlanmaktadır. Ayrıca araştırmaya her yıl yaşam koşullarına yönelik Avrupa Birliği İstatistik Ofisi’nin (EUROSAT) belirlediği konuda modül sorular ilave edilmektedir. 2012 uygulamasında konut koşulları modül anketi uygulanmıştır [2].

Alan uygulamasının her yıl düzenli olarak gerçekleştirildiği ve panel anket yönteminin kullanıldığı araştırmada, örnek fertler 4 yıl boyunca izlenmektedir. Araştırmadan, kesit ve panel veri olmak üzere her yıl iki veri seti elde edilmektedir [2].

2.1.2. Kapsam

2.1.2.1. Coğrafi Kapsam

Türkiye Cumhuriyeti sınırları içinde bulunan tüm yerleşim yerleri kapsama dahil edilmiştir. Bu yerleşim yerleri, TÜİK'in diğer anket çalışmalarında da kullanılmakta olan kent-kır tanımı dikkate alınarak iki tabakaya ayrılmıştır. Buna göre [2]:

Kentsel Yerler: Nüfusu 20 001 ve daha fazla olan yerleşim yerleri,

Kırsal Yerler: Nüfusu 20 000 ve daha az olan yerleşim yerleri olarak tanımlanmıştır.

2.1.2.2. Kapsanan Kitle

Türkiye Cumhuriyeti sınırları içindeki hanelerde yaşayan tüm hanehalkı fertleri kapsama alınmıştır. Ancak çalışmada kurumsal nüfus olarak tanımlanan üniversite yurtları, misafirhane, çocuk yuvası, yetiştirme yurdu, huzurevi, özel nitelikli hastane, hapisane, kışla ve ordu evlerinde yaşayanlar kapsam dışı tutulmuştur [2].

2.1.2.3. Kapsanan Konular

Araştırmada, gelir, yoksulluk, sosyal dışlanma ve diğer yaşam koşullarına ilişkin göstergeleri hesaplamak üzere aşağıdaki kategorilerde bilgi derlenmektedir [2]:

- Konut
- Ekonomik durum
- Sosyal dışlanma
- Gayrimenkul sahipliği
- Eğitim
- Demografi
- Sağlık durumu
- İşgücü durumu
- Gelir durumu

2.1.3. Örnekleme Tasarımı

2.1.3.1. Örnekleme Yöntemi, Birimi ve Tahmin Boyutu

Örnekleme yöntemi; tabakalı, iki aşamalı, küme örneklemesidir [2].

Nihai örnekleme birimi olarak hanehalkı tanımlanmıştır [2].

Yıllık kesit araştırma sonuçlarından, Türkiye, kent, kır ve İstatistiki Bölge Birimleri Sınıflaması Düzey-1, panel araştırma kapsamında 2, 3 ve 4 yıllık ülke geneli için tahminler üretilmesi amaçlanmaktadır [2].

2.1.3.2. Örneklem Büyüklüğü ve Cevapsızlık Oranı

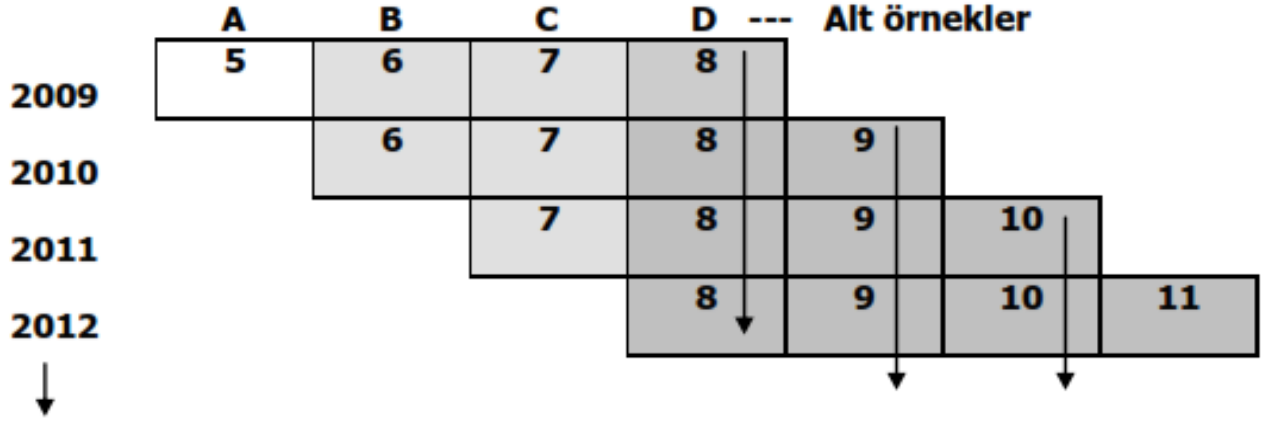
Yıllık örneklem büyüklüğü, araştırmanın tahmin boyutu, amaçları ve hedeflenen değişkenleri itibariyle örnekteki yıpranmaları da dikkate alarak başlangıçta yaklaşık 12 800 hanehalkı civarında belirlenmiştir. Ancak, 2014 yılında İstatistiki Bölge Birimleri Sınıflaması Düzey-2 bazında tahmin üretilmesi hedeflenmiş ve 2011 yılından itibaren örnek hacmi örneğe yeni eklenen her alt örnek ile birlikte arttırılmaya başlanmıştır. Dolayısı ile, 2012 GYKA'da kentsel yerlerden 13 106, kırsal yerlerden 6 194 olmak üzere toplam 19 300 hanehalkı ziyaret edilmiştir. Bu hanehalklarından; 17 562'si ile anket yapılmış, geri kalan 1 738 hanehalkı ile çeşitli nedenlerle anket yapılamamış ve cevaplamama formu düzenlenmiştir. Buna göre, 2012 yılında cevaplamama oranı Türkiye genelinde %9 olarak gerçekleşmiştir [2].

Araştırmada, örnek hanehalklarında yaşayan 15 ve daha yukarı yaştaki fert sayısı toplam 47 526 kişi olup, bu fertlerden 22'si ile çeşitli nedenlerle görüşülememiş veya ankete cevap vermemiştir. Dolayısı ile araştırmada fert cevapsızlık oranı %0,05'dir [2].

2.1.3.3. Panel Anket Uygulaması

Şekil 2.1'deki gibi rotasyonel tasarımın kullanıldığı araştırmada, metodoloji gereği bir kısım hanehalkı bir yıldan diğer yıla örnekte kalırken, diğer kısım örnekten çıkmakta ve yerine yeni örnek hanehalkları girmektedir. Bu araştırmada, yıllık örnek hacminin %75'inin panel kapsamda örnekte bırakılması öngörülmüştür. Bu doğrultuda panel uygulama, hedef kitleyi temsil edebilecek temel örnek seçimi ile başlamaktadır. Seçilen bu temel örnek hanehalkındaki 13 ve daha yukarı yaştaki tüm fertler belirlenmiş izleme kuralları doğrultusunda 4 yıl takip edilmektedir. Takip

kurallarının amacı, hedef nüfusta olabilecek değişiklikleri temel örneklerle yansıtmak ve fertleri belirli bir zaman takip etmektir [2].



Şekil 2.1. Rotasyonel Panel Anket Uygulaması

2.1.4. Referans Dönemleri

Referans dönemi; bilginin ait olduğu zaman süreci olarak tanımlanmakta olup, bu araştırmada farklı bilgiler için farklı referans dönemlerinin kullanımı söz konusudur [2].

Gelir bilgilerinin referans dönemi, "bir önceki takvim yılı"dır. Yani, 2012 uygulamasındaki gelir bilgileri 2011 yılına aittir. İstihdam bilgileri anketin uygulandığı tarihten bir önceki hafta ile anketin yapıldığı tarih dikkate alınarak sorgulanmıştır [2].

Yaşam koşulları göstergeleri anketin yapıldığı andaki durum itibariyle sorgulanmıştır [2].

2.1.5. Veri Derleme Yöntemi

Uygulamasına 2006 yılında başlanan GYKA her yıl düzenli olarak gerçekleştirilmektedir. Veri derleme çalışması Nisan-Temmuz ayları arasında yapılmakta olup yaklaşık üç aylık bir periyoda yayılmaktadır. Veri derleme çalışması iki aşamada gerçekleştirilmektedir [2].

1. aşamada, bir önceki uygulamadaki aynı adresinde oturan hanehalkları ve örneğe yeni eklenen hanehalkları ile görüşülmektedir [2].
2. aşamada, 1. aşamada taşındığı için ulaşılamayan hanehalkları ile hanehalkından ayrılarak yeni hanehalkı oluşturmuş örnek fertler izlenmektedir [2].

Anket uygulaması yüz yüze ve dizüstü bilgisayarlara yüklenmiş veri giriş programları (CAPI – computer assisted personel interview) aracılığıyla yapılmaktadır. Veri giriş programlarına eklenmiş olan sorular arası tutarlılık kontrolleri sayesinde verinin alanda daha tutarlı ve kontrollü girilmesi sağlanmaktadır. Veri derleme çalışmasını takiben fertlerin anket çalışmasından elde edilen T.C. kimlik numaraları kullanılarak kurumların idari kayıtlarından; ücret/müteşebbis geliri, emekli/dul-yetim aylığı, işsizlik ödenekleri, sakatlık/yaşlılık/gazilik vb. aylıkları, aldıkları sosyal yardımlar, gelir/varlık vergileri ile ödedikleri sigorta primlerine ilişkin bilgiler talep edilmekte ve anket verileri ile kontrol edilmektedir. Anketten gelen eksik/hatalı bilgiler idari kayıtlar yardımıyla düzeltilmektedir ve kontrol edilmektedir [2].

2.1.6. Tanım ve Kavramlar

2.1.6.1. Hanehalkı İle İlgili Tanımlar

Hanehalkı: Aralarında akrabalık bağı bulunsun ya da bulunmasın aynı konutta veya aynı konutun bir bölümünde yaşayan, temel ihtiyaçlarını birlikte karşılayan, hanehalkı hizmet ve yönetimine iştirak eden bir veya birden fazla kişiden oluşan topluluktur [2].

Kurumsal nüfus: Günlük yaşam gereksinimleri, yasal bir düzenlemeye dayalı olarak kurulan özel ya da tüzel kurum/kuruluşlarda kısmen ya da tamamen karşılanan, bireysel karar ve davranışlarında yetkili otoritenin kurallarına kısmen ya da tamamen bağımlı olarak hareket eden, ancak bireysel harcamalarına karar verebilen kişilerin yaşamlarını sürdürdükleri alanlarda yaşayan nüfusa kurumsal nüfus denir. Bu tanıma göre; üniversite yurtları, misafirhane, çocuk yuvası, yetiştirme yurdu, huzurevi, özel nitelikteki hastane, hapisane, kışla ve ordu evlerinde yaşayanlar kurumsal nüfusa dahildir [2].

Hanehalkı ferdi: Hanehalkını meydana getiren topluluğun bir üyesidir. Daimi ikameti örnek hanehalkı adresi olan fertler, görüşme anında geçici olarak hanede bulunmasalar da hanehalkı ferdi olarak kabul edilirler. Bununla birlikte, kurumsal nüfus kapsamında yer alan fertler (zorunlu askerlik görevini yapmakta olan er ve erbaşlar, hapiste olanlar, huzurevinde kalan yaşlılar ve yurttan kalan öğrenciler vb.) hanehalkı ferdi olarak kabul edilmezler. *Daimi ikamet*, 12 ay ve daha uzun süredir oturulan veya oturulma niyetinde olunan adrestir [2].

Referans kiři: Hanehalkının sosyo-ekonomik durumu ve hanede yařayan tüm fertlerin kiřisel özellikleri hakkında en doęru bilgiye sahip, hanenin yönetim veya geçiminden sorumlu yetişkin hanehalkı üyesidir [2].

2.1.6.2. Gelir İle İlgili Tanımlar

Gelir: Bir milli ekonomide belirli bir dönemde genellikle bir yılda yeniden yaratılan mal ve hizmet değerlerinin toplamıdır. Buna milli gelir veya toplam hasıla denir. Hanehalkı ya da fert açısından gelir, üretime yapılan katkı karşılığında belirli bir sürede elde edilen değerlerin toplamı olarak tanımlanabilir. Hanehalkı geliri, emek karşılığı alınan ücret ve maařlar şeklinde, toprak kirası (rant), sermaye geliri (faiz) veya teşebbüs geliri (kar) şeklinde olabilir [2].

Faaliyet gelirleri:

Maař, ücret, yevmiye gelirleri: Gelir referans dönemi içerisinde ücret karşılığı bir işte çalışma sonucunda elde edilen nakdi veya ayni gelirlerdir. Gelir referans dönemi içerisinde alınan ikramiye, iş riski, iş güçlüğü teminindeki günlük zammı, fazla mesai, prim ve temettüleri ve ek görevlerden elde edilen gelirler maař, ücret veya yevmiye gelirine dahil edilmiştir [2].

Müteşebbis gelirleri: Gelir referans dönemi içerisinde müteşebbis olarak çalışan fertlerin (işteki durumu kendi hesabına veya işveren olarak çalışanlar) faaliyetleri sonucu elde ettiği gayri safi hasılatından, gelir referans dönemi içerisinde yaptığı tüm giderler (işyeri masrafları, vergiler, sosyal güvenlik kesintileri vb.) düşüldükten sonra elde edilen gelir müteşebbis gelirdir [2].

Faaliyet dışı gelirleri:

Gayrimenkul geliri: Apartman dairesi, dükkan, depo vb. gayrimenkuller ile tarla, bağ, bahçe, arsa gibi mülklerin referans dönemi içerisinde kiraya verilmesi sonucu elde edilen hasılatın, bu hasılatın sağlanması için yapılan tüm giderler (sigorta, vergi, tamir ve bakım giderleri vb.) düşülerek elde edilen gelirdir [2].

Menkul kıymet geliri: Banka hesabından elde edilen faiz, finans kurumlarından veya sermaye şirketlerinden elde edilen temettü veya kar payı vb. gelirlerdir. Menkul kıymet gelirlerinde, vergiler gelir elde edilmeden kaynaktan kesilmektedir. Ancak, menkul kıymet geliri elde eden hanehalkı

fertlerinin, bu geliri elde etmek için, menkul kıymetlerin korunması ile ilgili sigorta, kiralık kasa, tahvil giderleri, vergi vb. giderler düşülerek, elde ettikleri net gelirler menkul kıymet geliri olarak kayıt edilmiştir [2].

Karşılıksız (transfer) gelirler: Gelir referans dönemi içerisinde devlet, özel kurum/kuruluş veya başka kişi/hanehalklarından alınan karşılıksız transferlerdir (emekli maaşı, yaşlılık maaşı, işsizlik maaşı, karşılıksız burslar, düzenli olarak alınan aynı veya nakdi yardımlar vb. şeklindeki ödemelerdir.) İki gruba ayrılabilir [2];

Sosyal transferler; Hanelerin bazı ihtiyaçlarını karşılamak veya parasal sıkıntılarından dolayı bazı risklerle karşı karşıya kalmaları durumunda gelir referans döneminde devletten veya çeşitli kurum/kuruluşlardan aldıkları aynı veya nakdi karşılıksız yardımlar olarak nitelendirilmektedir (Aile-çocuk yardımı, konut yardımı (kira vb.), yoksul hanehalklarına yapılan diğer sosyal yardımlar) .

Hanelerarası transferler; Hanelerin başka kişi veya hanehalklarından düzenli olarak aldıkları karşılıksız aynı ya da nakdi yardımlardır.

İzafi kira: Konuta mülkiyet şekli “ev sahibi”, “lojman” veya “diğer (babasının, akrabasının vb. evinde ikamet edip hiç bir şekilde ücret ödemedi oturanlar)” şeklinde olan hanehalklarının ikamet ettiği konutun kira değeri izafi kira olarak değerlendirilmektedir [2].

İzafi kira değeri iki şekilde hesaplanmaktadır [2];

- i. Hanehalklarından, anket yoluyla, benzer özelliklere sahip olan aynı semtte ve yerleşim yerinde kira ile oturlan konutun kira değeri emsal alınarak saptanmak üzere beyan yoluyla alınmaktadır.
- ii. Regresyon modeli kurularak, açıklayıcı değişkenlerin yardımıyla tahmin edilmektedir.

Hanehalkı kullanılabilir geliri hesaplanırken ikinci hesaplama kullanılmaktadır [2].

Nakdi Gelir: Nakit; para, özellikle kullanılmaya hazır para anlamına gelmektedir. Faaliyet dışı faktörler veya faaliyet karşılığında gelirlerin nakit (para) olarak elde edilmesidir [2].

Aynı Gelir: Aynı, mal ile ödenen anlamına gelmektedir. Aynı gelir, emek veya sermaye karşılığında gelirin mal olarak elde edilmesidir. Burada en önemli husus, aynı gelirin hanede tüketilmesi gerekliliğidir. Aynı geliri oluşturan mal, satılıp nakde çevriliyorsa bu aynı gelir olarak değil, nakdi gelir olarak değerlendirilmektedir [2].

Brüt gelir: Emeklilik ve sosyal sigortalar keseneği ile vergiler düşülmeden önceki gelirdir [2].

Net gelir: Kişilerin ödediği emeklilik ve sosyal sigortalar keseneği ile vergiler hariç, eline geçen gelir net gelirdir [2].

Hanehalkı kullanılabilir net geliri: Hanehalkındaki her bir ferdin elde ettiği kişisel yıllık kullanılabilir gelirlerin (maaş-ücret, yevmiye, müteşebbis geliri ile emekli maaşı, dul-yetim aylıkları ve yaşlılara yapılan ödemeler, karşılıksız burs vb. aynı veya nakdi gelirlerin toplamı) toplamı ile hane bazında elde edilen yıllık gelirlerin (gayrimenkul kira geliri, haneye yapılan karşılıksız yardımlar, 15 yaşın altındaki fertlerin elde ettiği gelirler vb.) toplamından, gelir referans döneminde ödenen vergiler ve diğer hane veya kişilere yapılan düzenli transferler düşüldükten sonra Çizelge 2.1.'deki hanehalkı kullanılabilir gelirine ulaşılmaktadır [2].

Eşdeğerlik ölçüğü / eşdeğer hanehalkı geliri: Günümüzde gelir eşitsizliği bireysel gelirler arasındaki farklardan yola çıkılarak ölçülmektedir. Bu nedenle de hanehalkı düzeyinde toplanan gelirlerin bireysel gelirlere dönüştürülmesi gerekmektedir. Bunu, toplam hane gelirini, haneyi oluşturan fert sayısına bölerek yapmak doğru olmayacaktır. Bu hesaplamada, hanelerin yetişkin-çocuk bileşimlerindeki farklılıkları dikkate almak gerekmektedir. Çünkü bilimsel olarak çocuklar yetişkinlere göre daha az tüketmektedirler. Buna göre, eşdeğerlik ölçüğü olarak adlandırılan katsayılar kullanılarak her bir hanehalkı büyüklüğünün kaç yetişkine (eşdeğer fert sayısına) denk olduğu hesaplanmaktadır. Hanehalkı toplam yıllık kullanılabilir geliri hanehalkının eşdeğer hanehalkı büyüklüğüne bölünerek, o hanehalkı için eşdeğer fert başına düşen, diğer ifadeyle eşdeğer hanehalkı kullanılabilir geliri hesaplanmaktadır [2].

Bu arařtırmada, hanehalkındaki referans kiři için "1", 14 yař ve üzerindeki tüm fertler için "0,5" ve 14 yařın altındaki tüm fertler için "0,3" olarak dikkate alınarak hesaplanan OECD ölçeęi kullanılmıřtır [2].

Çizelge 2.1. Hanehalkı Kullanılabilir Net Gelirin Hesaplanması

(1)	Maař, Ücret veya Yevmiye Net Nakdi Geliri	
(2)	Maař, Ücret veya Yevmiye Ayni Geliri	
(3)	Maař, Ücret veya Yevmiye Toplam Net Geliri	(1)+(2)
(4)	Müteřebbis Net Nakdi Geliri (Tarım / Tarım Dıřı)	
(5)	Müteřebbis Ayni Geliri (Tarım / Tarım Dıřı)	
(6)	Müteřebbis Toplam Net Geliri (Tarım / Tarım Dıřı)	(4)+(5)
(7)	Gayrimenkul Net Nakdi Geliri	
(8)	Gayrimenkul Ayni Geliri	
(9)	Gayrimenkul Toplam Net Geliri	(7)+(8)
(10)	Menkul Kıymet Geliri	(11+12+13+14+15+16+17+18+19+20+21+22+23+24)
(11)	İřsizlik Ödenekleri	
(12)	Emekli Maařı	
(13)	Yařlılık Maařı	
(14)	İsteęe Baęlı Emeklilik	
(15)	Dul/Yetim Maařı	
(16)	Ölüm Tazminatı	
(17)	Hastalık Yardımı	
(18)	Sakatlık Aylıęı	
(19)	Malüllük Aylıęı	
(20)	Gazilik Aylıęı	
(21)	Eęitim İle İlgili Karřılıksız Burs ve Baęıřlar	
(22)	Aile ve Çocuk Yardımları	
(23)	Konut Yardımları	
(24)	Dięer Sosyal Yardımlar	
(25)	Sosyal Transferler	
(26)	15 Yařından Küçük Fertlerin Gelirleri	
(27)	Hanelerarası Transferler (Alınan) (Nafaka Dahil)	
(28)	Hanenin Kendi Tüketimi İin Ürettięi Ürünlerin Deęeri	
(29)	İzafi Kira Geliri (Yıllık)	
(30)	Hanelerarası Transferler (Verilen) (Nafaka Dahil)	
(31)	Ödenen Vergiler (Kira, Otomobil) ve Sosyal Güvenlik Primleri	
(32)	Hanehalkı Yıllık Kullanılabilir Net Geliri	(3+6+9+10+25+26+27+28+29)-(30+31)

2.2. Hanehalkı Bütçe Araştırması

2.2.1. Açıklama

HBA, hanelerin sosyo-ekonomik yapıları, yaşam düzeyleri, tüketim kalıpları hakkında bilgi veren ve toplumun ihtiyaçlarının belirlenmesi, kullanılabilir gelirin haneler ya da fertler arasında ne şekilde dağıldığının bilinmesi ve uygulanan sosyo-ekonomik politikaların geçerliliğinin test edilmesi amacıyla kullanılan en önemli kaynaklardan biridir. TÜİK, 2002 yılından itibaren her yıl düzenli olarak HBA uygulamaya başlamıştır [3].

HBA'dan elde edilen tüketim harcamaları bilgileri ile [3];

- Tüketici fiyat endekslerinde kullanılacak maddelerin seçimi ve temel yıl ağırlıkların elde edilmesi,
- Hanelerin tüketim kalıplarında zaman içinde meydana gelen değişikliklerin izlenmesi,
- Milli gelir hesaplamalarında özel nihai tüketim harcamaları tahminlerine yardımcı olacak verilerin derlenmesi,
- Yoksulluk sınırı ve yoksul fert sayısının belirlenmesi vb. diğer sosyo-ekonomik göstergelerin elde edilmesi,
- Asgari ücret tespit çalışmaları için gerekli verilerin derlenmesi

amaçlanmaktadır.

TÜİK, önceki yıllarda adı "Hanehalkı Gelir ve Tüketim Harcamaları Anketi" olan bu araştırmanın ilkini, 1954 yılında, Ankara'da sadece memur ailelerini kapsayacak şekilde uygulamıştır. Daha sonra, 1964-70 yılları arasında, 1973-74 ve 1978-79 yıllarında Hanehalkı Gelir ve Tüketim Harcamaları Anketi belirli kapsamlarla uygulanmış ve bu yıllar baz alınarak Tüketici Fiyat Endeksleri kurulmuştur. TÜİK, Türkiye genelini kapsayan ilk çalışmayı 1987 yılında gerçekleştirmiş ve bu anketin sonuçlarını "Tüketim Harcamaları" ve "Gelir Dağılımı" olmak üzere iki yayın halinde kullanıcılara sunmuştur. 1987 yılından sonra uygulanan 1994 Hanehalkı Gelir ve Tüketim Harcamaları Anketi, daha önce uygulanan anketlerden farklı bir yöntemle; tüketim harcamaları ve gelir dağılımı amaçlarına yönelik olarak iki ayrı anket şeklinde düzenlenmiştir. Bu uygulamada, 1 Ocak – 31 Aralık 1994 tarihleri arasında Türkiye genelinde 62 kentsel ve 174 kırsal yerleşim yerinde her ay değişen, 1 522'si kentsel, 666'sı kırsal yerleşim yerinde olmak üzere toplam 2 188 örnek haneye

Hanehalkı Tüketim Harcamaları Anketi uygulanmıştır. Bu anketin sonuçları, 1994=100 bazlı Tüketici Fiyatları Endekslerinin hesaplanmasında kullanılmış olup, bu çalışmadan elde edilen sonuçlardan Türkiye, kent, kır ve 7 coğrafi bölge ile seçilmiş 19 il merkezi için tahminler elde edilmiştir. İkinci aşamada ise Şubat – Mayıs 1995 tarihleri arasında “Hanehalkı Tüketim Harcamaları Anketi” kapsamında olan toplam 26 256 örnek haneye “Hanehalkı Gelir Dağılımı Anketi” uygulanmıştır [3].

TÜİK, 2002 yılından itibaren daha küçük bir örnek hacmi ile her yıl düzenli olarak HBA uygulamayı planlamıştır. Buna göre; 2002 HBA, 1 Ocak – 31 Aralık 2002 tarihleri arasında bir yıl süreyle kentsel kesimden her ay değişen yaklaşık 650, kırsal kesimden her ay değişen yaklaşık 150 hanehalkı olmak üzere aylık 800, yıllık ise 9 600 örnek hanehalkına uygulanarak, Türkiye geneli, kentsel ve kırsal yerler ayrımında tüketim harcaması, gelir dağılımı ve yoksulluk göstergeleri elde edilmiş ve sonuçları kamuoyuna duyurulmuştur. 2003 HBA, 1 Ocak–31 Aralık 2003 tarihleri arasında bir yıl süre ile kentsel kesimden her ay değişen 1 512 ve kırsal kesimden her ay değişen 648 hanehalkı olmak üzere aylık toplam 2 160, yıllık toplam 25 920 örnek hanehalkına uygulanmış; anketten tüketim harcamaları, gelir dağılımı ve yoksulluk konularında temel göstergeler elde edilmiştir. 2004-2008 döneminde, HBA her yıl için her ay değişen aylık toplam 720, yıllık toplam 8 640 örnek hanehalkıyla; 2009 anketi her ay değişen aylık ortalama 1 050, yıllık toplam 12 600 örnek hanehalkıyla; 2010 ve 2011 anketleri ise her ay değişen 1 104, yıllık toplam 13 248 örnek hanehalkıyla yürütülmüştür [3].

HBA'nın örnekleme yapısı gereğince; yıllık anket sonuçlarından Türkiye geneli ile kırsal ve kentsel yerler ayrımında; her yılın anket sonuçlarının son iki yılın anket bilgileriyle birleştirilmesi ile de bölgesel düzeyde sadece oransal tahminler üretilmesi mümkün olmaktadır. Buna göre; 2009, 2010 ve 2011 yıllarında yapılan HBA veri setleri bir araya getirilerek, sonuçları üç yılı kapsayan ancak 2009 ve 2010 yıllarına ait tüketim harcamaları enflasyona göre 2011 yılına artırıldıktan sonra elde edilen 2009-2010-2011 HBA sonuçları İstatistikî Bölge Birimleri Sınıflaması Düzey-1 ve Düzey-2 bazında bu haber bülteni ile kamuoyuna sunulmuştur. 2009-2010-2011 birleşik veri seti, kırsal yerleşim yerinden 9 450, kentsel yerleşim yerlerinden 20 596 örnek hanehalkı olmak üzere toplam 30 046 örnek hanehalkının tüketim harcaması bilgilerini içermektedir [3].

2.2.2. Anketin Örnekleme Planı ve Tahmin Yöntemi

2.2.2.1. Coğrafi Kapsam

Türkiye Cumhuriyeti sınırları içinde bulunan tüm yerleşim yerleri kapsama dahil edilmiştir. Bu yerleşim yerleri iki tabakaya ayrılmıştır. Buna göre [3];

Kentsel Yerler: Nüfusu 20 001 ve daha fazla olan yerleşim yerleri,

Kırsal Yerler: Nüfusu 20 000 ve daha az olan yerleşim yerleri olarak tanımlanmıştır.

2.2.2.2. Kapsanan Kitle

Anket çalışmasında, Türkiye Cumhuriyeti sınırları içinde yaşayan tüm hanehalkı fertleri kapsama alınmıştır. Ancak, çalışmada kurumsal nüfus olarak tanımlanan yaşlılar evi, huzur evleri, hapishane, askeri kışla, özel nitelikli hastane, otel, çocuk yuvalarında bulunan nüfus ile pratik nedenlerden dolayı göçer nüfus kapsam dışı tutulmuştur [3].

2.2.2.3. Tüketim Harcamaları ve Gelirin Kapsamı

Tüketim harcaması; hanehalklarının anket ayı içinde yaptığı tüketim amaçlı satın alışları, kendi üretiminden tüketimi, çalışan fertlerin hanehalkı tüketimi için işyerinden getirdiği mal ve hizmetleri, hanehalkının hediye/yardım etmek amacıyla satın aldıkları ile son bir yıl içinde satın aldığı otomobil, beyaz eşya, boya-badana, mobilya, bilgisayar, televizyon, kamera, mobilya, özel okul, dersane vb. dayanıklı tüketim mal ve hizmetleri veya seyrek yapılan harcamaların aylık ortalaması; gelir kapsamında ise son bir yılda elde edilen kullanılabilir gelir bilgileri kapsanmıştır [3].

2.2.2.4. Örnekleme Çerçevesi, Birimi ve Yöntemi

2009 yılından itibaren HBA'da birinci aşama örnekleme birimi olan blokların seçiminde kullanılan temel örnekleme çerçevesi, 2007 yılında oluşturulan ve belli dönemlerde güncellenen Ulusal Adres Veri Tabanıdır [3].

Yukarıda verilen çerçeve kullanılarak bloklar oluşturulmuş, kentsel yerlerden ve belediye teşkilatı olan kırsal yerler ve köylerden yerleşim yerinin büyüklüğüne orantılı olasılıkla bloklar belirlenmiş ve her bloktan nihai örnekleme birimi olan hanehalkları sistematik olarak seçilmiştir [3].

Nihai örnekleme birimi olarak hanehalkı tanımlanmıştır [3].

Örnekleme yöntemi olarak, tabakalı iki aşamalı küme örnekleme yöntemi kullanılmıştır [3].

2.2.3. Tanım ve Kavramlar

Hanehalkı: Aralarında akrabalık bağı bulunsun ya da bulunmasın aynı konutta yaşayan, kazanç ve masraflarını ayırmayan, hanehalkı hizmet ve yönetimine katılan bir veya birden fazla kişiden oluşan topluluktur [3].

Hanehalkı ferdi: Hanehalkını meydana getiren topluluğun her bir üyesidir. Bununla birlikte askerde, hapiste olanlar ve huzurevlerinde kalan yaşlılar ve yurttan kalan üniversite öğrencileri, hanede kalış süresi ne olursa olsun misafirler ve hanehalkından evlenme, askere gitme, çalışmaya gitme gibi sebeplerle kesin olarak ayrılanlar hanehalkı ferdi olarak kapsamamaktadır [3].

Tüketim: Mal ve hizmetlerin insan ihtiyaçlarını ve arzularını karşılamak üzere kullanılmasıdır [3].

Harcama: Bir mal veya hizmetli satın almak için yapılan parasal ödeme veya fedakarlıkların tümüdür [3].

Tüketim harcaması bileşenleri [3]:

- Satın alışların değeri,
- Kendi üretiminden tüketilen ürünlerin değeri,
- Müteşebbis olan fertlerin anket ayında işyerinden haneye getirdiği ürünlerin değeri,
- Ücretli çalışan fertlere anket ayında işvereni tarafından verilen mal ve hizmetlerin değeri,
- Hanehalkının özel kişi veya kuruluşlara vermek üzere hediye veya yardım şeklinde satın aldığı mal veya hizmetlerin değeri,
- İzafe kira,
- Son bir yıl içinde satın alınan otomobil, beyaz eşya, bilgisayar, televizyon, kamera, mobilya, ısıtma ve soğutma sistemi gibi dayanıklı tüketim mallarına yapılan harcamaların aya düşen değeri.

Satın alış: Herhangi bir mal ya da hizmete belli bir ücret ödeyerek sahip olmaktır [3].

Kendi üretiminden tüketim: Hanelerin kendi üretimleri sonucunda elde ettikleri maddeleri, ürettikleri gibi ya da işleyerek tüketmeleridir. Örneğin; hanehalkının kendi ürettiği sütün süt olarak ya da peynir yapılarak tüketilmesi gibi [3].

Aynı gelirlerden tüketim: Hanede iktisaden faal olan fertlerin çalıştıkları işyerinde üretilen ya da satışa sunulan mal ve hizmetlerden hanede tüketilenler, aynı gelirlerden tüketim olarak alınmıştır. Örneğin; ücretli olarak çalışan bir ferde işyerince verilen yiyecek, giyecek vb. yardımlar, servis vb. hizmetler ücretlinin elde ettiği aynı gelirden tüketimi olarak; işveren ya da kendi hesabına çalışan bir ferdin de kendisine ait işyerinde ürettiği ya da sattığı mallardan hanesine tüketmek üzere getirdikleri aynı gelirden tüketim olarak değerlendirilmiştir [3].

Hanehalkının özel kişi veya kuruluşlara vermek üzere hediye veya yardım şeklinde satın aldığı mal veya hizmetler: Hanehalkı fertleri tarafından anket ayı süresince, özel kişi veya kuruluşlara verilmek üzere (hediye etmek veya yardımda bulunmak amacıyla) satın alınan mal ve hizmetlerin değeri kapsamıştır [3].

İzafi kira: Konuta mülkiyet şekli ev sahibi, lojman ve diğer (babasına, akrabasına vb. kişilere ait yerlerde ikamet edip hiç bir şekilde ücret ödemedi ya da çok düşük değerinde kira ödeyenler) şekilde olan hanehalkının ikamet ettiği konutun kira değeri izafi kira olarak değerlendirilmiştir. İzafi kira, tüketim harcaması içinde kapsamaktadır. İzafi kira değeri, benzer özelliklere sahip olan, aynı semtte ve yerleşim yerinde kira ile oturlan konutun kira değeri emsal alınarak saptanmıştır [3].

2.2.4. Hanehalkı Tüketim Harcamalarının Kapsamı

- Harcamaların hanehalkı için yapılmış olması gerekmektedir [3].
- Harcamaların tüketim amacıyla yapılmış olması gerekmektedir [3].
- Tüketim harcamaları aşağıdaki gruplar ayrımında kapsamıştır [3]:
 1. Gıda ve alkolsüz içecekler,
 2. Alkollü içecekler, sigara ve tütün,
 3. Giyim ve ayakkabı,
 4. Konut, su, elektrik, gaz ve diğer yakıtlar,
 5. Mobilya, ev aletleri ve ev bakım hizmetleri,

6. Sağlık,
7. Ulaştırma,
8. Haberleşme,
9. Kültür, eğlence,
10. Eğitim hizmetleri,
11. Otel, lokanta ve pastane,
12. Çeşitli mal ve hizmetler.

- Hanehalkı tarafından yapılan tüketim harcamaları, anket dönemi içinde ölçüm kriterlerine (sahiplenme, kullanım, ödeme) uygun bir şekilde alınmıştır [3].
- Anket döneminde yapılan tüketim harcamaları kapsanmıştır. Araştırmada örnek hanehalkı tarafından anket dönemi içinde [3],
 - Hangi tüketim mal ve hizmetlerinin,
 - Nasıl (peşin veya taksitli satın alış, kendi üretimi, aynı gelen vb. şeklinde),
 - Ne kadar (ne miktarda),
 - Ne değerinde,
 - Nereden alındığı şeklinde bilgiler derlenmiştir.

2.2.5. Alan Uygulaması Yöntemi

HBA'da bilgiler; görüşme, kayıt ve gözlem yöntemleri kullanılarak derlenmiştir. Her anketör, her ay ortalama 6 örnek hanehalkını anket ayı öncesi 1 kez, 1. ve 2.hafta 2'şer kez, 3. ve 4.hafta 1 kez ve anket ayı bitiminde de 1 kez olmak üzere her bir hanehalkını ayda ortalama 8 defa ziyaret ederek tüketim harcamaları ve son ziyarette gelir bilgilerini kayıt etmiştir. Anket ayı öncesinde örnek hanehalkına yapılan ilk ziyarette, hanehalkı ile tanışma sağlandıktan sonra hanehalkının sosyo-ekonomik durumuna ait bilgiler alınmakta, harcama kayıt defterinin nasıl doldurulacağı açıklanmaktadır [3].

Anket ayı içinde yapılan ziyaretler sırasında, örnek hanenin anket ayında yaptığı gıda, giyim, sağlık, ulaştırma, haberleşme, eğitim, kültür, eğlence, konut, ev eşyası vb. gibi tüketim harcamaları hanehalkına bırakılan kayıt defterinden kayıtlama ve görüşme yöntemleri ile alınmıştır. Anket ayı bitiminde yapılan son görüşmede ise,

hanehalkı fertlerinin anket ayı içindeki istihdam durumu, iktisadi faaliyeti, mesleđi, işteki durumu, anket ayı ve son bir yıl içerisinde elde ettiđi gelir bilgileri derlenmiştir [3].

3. VERİ MADENCİLİĞİ

Günümüzde meydana gelen gelişmelerin temelini veri, bilgi ve enformasyon kavramları oluşturmaktadır. Sistem politikalarının ve stratejik kararların temelini veri ve veriden sağlanan bilgi oluştururken, politika ve kararların amacına uygun ve etkili olabilmesi için güvenilir, güncel, doğru ve zamanında ulaşan veriye ihtiyaç vardır. Bu kavramlar hayatımızın her evresinde karşımıza çıkarken, sağlık verileri, alışveriş verileri, otomasyon verileri, oldukça hızlı bir şekilde artarken oluşan veri tabanlarını saklamak da işletmeler açısından büyük bir külfet haline gelmiştir. Bu verilerin depolanması, saklanması belli başlı bir problem haline almış ve araştırmacılar bu problemlere veri tabanları ve dosya sistemlerindeki gelişmelerle çözüm üretme yoluna gitmişlerdir. Araştırmacılar bu muazzam büyüklükteki veri tabanlarını herhangi bir araç kullanmadan analize tabi tutularak karar destek aşamasında kullanımının imkansız olduğunu görmüşler ve bu aşamada veri tabanından bilgi keşfi kavramı ortaya çıkmıştır. Veri tabanından bilgi keşfi; veriden yararlı bilgi çıkarma sürecidir [4, 5].

Veri tabanında bilgi keşfinin uygulanabilmesi için verilerin veri ambarı gibi bir veri tabanı düzeninde olması gerekmektedir. Veri ambarı, sorgulama, raporlama ve analizler için oluşturulan yapılardır. Veri ambarları işlemsel olduğu kadar işlemsel olmayan verileri de içermektedir. Bu yapılar veri tabanında bilgi keşfinin altyapısını oluşturmaktadır. Veri ambarının düzenli olması, bilgi keşfi uygulamalarını oldukça kolaylaştıracak bir etkidir [4, 6].

3.1. Tanımlar

3.1.1. Bilgi

Bilgi, verinin karar verici için anlamlı ve kullanışlı biçime dönüşmesidir. Bu dönüşüm süreci aşağıdaki işlemlerden oluşur [7].

- Toplama,
- Sınıflandırma,
- Yeniden düzenleme,
- Özetleme,
- Saklama,

- Yeniden elde etme,
- İletme,

Veri madenciliğinde dört farklı bilgi çeşidi yer almaktadır. Veri tabanına kolayca depolanan ve işlenen *sığ bilgi*, verideki düzenlilikleri ve örüntüleri ifade eden ve kolay ulaşılamayan *gizli bilgi*, çok boyutlu bir formata depolanan on-line analitik süreç araçlarının faydalandığı *çok boyutlu bilgi* ve veri tabanına depolanan, talimatlar verildiği zaman kolay ulaşılabilen *derin bilgidir* [4, 8].

3.1.2. Veri

Veri, ham gerçeklerin işlenmemiş halidir. Veriler işlenerek ve düzenlenerek bilgiye dönüşür. Veri işleme, temel bir veriyi organize edebilmek ve bu verilerden örüntüler, karmaşık tahminler veya istatistiksel modellerden faydalanılarak sonuçlar ortaya çıkarmaktır. Gerçek anlamda bilgi, kaynak gerektirmektedir. Kesin, ilişkili ve zamanında bilgi, doğru karar vermenin anahtarıdır [4, 9].

Tek başına enformasyon işi eyleme dönüştürmede yeterli değildir. Enformasyon mutlaka bir yönetici tecrübesi ile birleştirilmelidir. O halde *enformasyon* dolaylı bilgidir diyebiliriz. Bunun yanında enformasyon etkili bir eylem için bilgi sağlamaktadır [4, 10].

3.1.3. Veri Tabanı

Veri tabanı, herhangi bir konuda birbirleri ile ilişkili verilerin sistematik olarak oluşturduğu yapılardır. Veri tabanı, sistematik erişim imkanı olan, yönetilebilir, güncellenebilir, taşınabilir, birbirleri arasında tanımlı ilişkiler bulunabilen düzenli bilgiler topluluğudur. Bir bilgisayarda sistematik şekilde saklanmış, programlarca istenebilecek veri yığınları da veri tabanıdır [4, 11].

Veri tabanı, veri merkezini güçlendirmenin ve kontrol etmenin en iyi yoludur. Veri oldukça geniş bir alana düz dosyalar şeklinde yayıldığı için bu verilerin kontrolü ve yönetimi oldukça zor bir hal alır ve bu tür işlemsel verileri çok daha iyi kontrol eden veri tabanına ihtiyaç duyulur [4, 12].

Veri tabanı yönetim sistemleri veriye erişmek için bir program seti ve birbiriyle alakalı veri birikimini içerir. Veri tabanı yönetim sistemlerinin temel amacı; alınan ve depolanan veri tabanı bilgisini etkili ve elverişli bir şekilde kullanacak ortamı sağlamaktır [4, 13].

Veri tabanı yönetim sistemi, veri tabanındaki bütün girişlerin kontrolünü sağlamak için oluşturulan bir yazılımdır. Veri tabanı yönetim sistemleri dört aşamada işlem yapar [4, 12].

- 1- Kullanıcının kayıtlara gerekli veri işleme dilini kullanarak ulaşmasını sağlar.
- 2- Veri tabanı yönetim sistemi, kullanıcının ihtiyacı olan tüm yorumları sağlar.
- 3- Veri tabanı yönetim sistemi, işlemleri denetler.
- 4- Veri tabanı yönetim sistemi, depolanmış veri tabanı üzerinden gerekli uygulamaların yapılmasını sağlar.

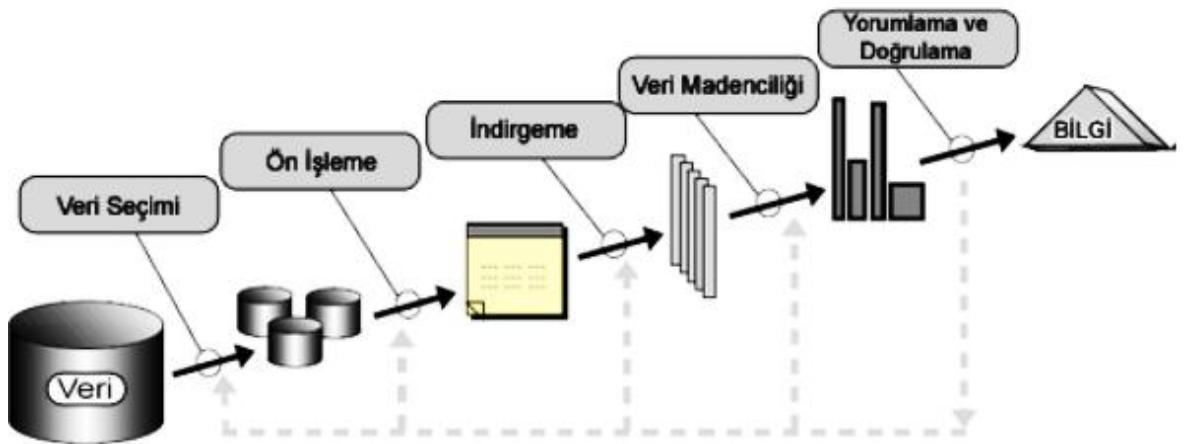
3.1.4. Veri Ambarı

Veri ambarı, çeşitli kaynaklardan periyodik olarak kopyalanan verinin ve tasarımların birleştirilmesidir. İşletmenin hem içinde hem dışında analitik ve bilgilendirici bir süreç için optimize edilmiş bir çevredir [4, 14].

3.2. Veri Tabanında Bilgi Keşfi Süreci

Veri madenciliği, bilgi keşfi ve veri tabanlarında bilgi keşfi terimleri bazı araştırmacılar tarafından karıştırılmaktadır. Birçok araştırmacı ve uygulayıcı veri madenciliği ve bilgi keşfi kavramlarını benzer anlamda kullanmaktadır. Fakat veri madenciliği, bilgi keşfi sürecinin bir aşamasıdır. Veri tabanında bilgi keşfini kısaca tanımlarsak; verideki anlamlı, faydalı, orijinal ve belli bir değeri olan örüntüleri ortaya çıkarma işlemidir [4, 15].

Şekil 3.1.'de görüldüğü gibi veri tabanında bilgi keşfi süreci altı aşamadan oluşmaktadır:



Şekil 3.1. Veri Tabanında Bilgi Keşfi Süreci

3.2.1. Uygulama Alanının Anlaşılması ve Geliştirilmesi

Bu aşamada uygulama alanı anlamaya çalışılarak çeşitli hazırlıklar yapılır ve bilgi keşfindeki amaç geliştirilir [4].

3.2.2. Bir Veri Seti Seçme ve Oluşturma

Bu aşama veri kütlelerini birleştirip sorgu için uygun örneklem kümesini oluşturmayı amaçlar [4].

3.2.3. Önleme ve Temizleme

Bu aşama seçilen örnekte bulunan hatalı ve eksik nitelikteki verilerin temizlendiği ve değiştirildiği bir aşamadır. Ayrıca bu işlem keşfedilen bilginin değerini ve kalitesini artırmaya yöneliktir [4].

3.2.4. Veri İndirgeme

Seçilen örnekte ilgisiz nitelikler ve tekrarlanan veriler temizlenir. Bu aşamanın amacı, veri madenciliği sorgusunun çalışmasını verimli hale getirmektir [4].

3.2.5. Uygun Bir Veri Madenciliği Programı ve Algoritması Çalıştırmak

Bu aşamada hangi veri madenciliği türü kullanılacağına karar verilir (Örneğin; sınıflandırma, regresyon veya kümeleme gibi). Bu genelde veri tabanında bilgi keşfi amacına ve önceki aşamalarına bağlıdır. Bu aşamada ayrıca strateji belirlenmesi, kullanılacak taktiklere karar verilmesi ve örüntü araştırmak için kullanılacak kesin bir yöntemin seçilmesi gerekir. Burada amaç tatmin edici sonuca ulaşana kadar algoritmayı çalıştırmaktır [4].

3.2.6. Yorumlama, Doğrulama ve Bilgi Keşfini Kullanmak

Bu aşamada oluşan örüntüler değerlendirilerek birinci aşamada belirlenen amaç ile ilişki kurulmalı ve elde edilen bilgi amaca yönelik kullanılmalıdır [4].

3.3. Veri Madenciliği Tarihçesi

Günümüzde neredeyse her eve bilgisayar girmiştir ve internet erişimi hemen hemen her yerden sağlanmaktadır. Disk kapasitelerinin artması, her yerden bilgiye ulaşma olasılığı, bilgisayarların çok büyük miktarlarda veri saklamasına ve daha kısa sürede işlem yapmasına olanak sağlamıştır. Geçmişten günümüze veriler her zaman yorumlanmış, bilgi edinmek istenmiş ve bunun için de donanımlar oluşturulmuştur. Bu sayede bilgi, geçmişten günümüze taşınabilir hale gelmiştir [16].

1950'li yıllarda ilk bilgisayarlar sayımlar için kullanılmaya başlanmıştır. 1960'larda ise veri tabanı ve verilerin depolanması kavramı teknoloji dünyasında yerini almıştır. 1960'ların sonunda bilim adamları basit öğrenmeli bilgisayarlar geliştirebilmişlerdir. Minsky ve Papert¹, günümüzde sinir ağları olarak bilinen perseptron'ların sadece çok basit olan kuralları öğrenebileceğini göstermişlerdir. 1970'lerde ilişkisel veri tabanı yönetim sistemleri uygulamaları kullanılmaya başlanmıştır. Bilgisayar uzmanları bununla beraber basit kurallara dayanan uzman sistemler geliştirmişler ve basit anlamda makine öğrenimini sağlamışlardır [16, 17].

1980'lerde veri tabanı yönetim sistemleri yaygınlaşmış ve bilimsel alanlarda uygulanmaya başlanmıştır. Bu yıllarda şirketler, müşterileri, rakipleri ve ürünleri ile ilgili verilerden oluşan veri tabanları oluşturmuşlardır. Bu veri tabanlarının içerisinde çok büyük miktarlarda veri bulunmaktadır ve bunlara SQL veri tabanı sorgulama dili ya da benzeri diller kullanarak ulaşılabilir. 1990'larda artık içindeki veri miktarı katlanarak artan veri tabanlarından, faydalı bilgilerin nasıl bulunabileceği düşünölmeye başlanmıştır. Bunun üzerine çalışmalara ve yayınlara başlanmıştır. 1989 yılında gerçekleştirilen Veri Tabanlarında Bilgi Keşfi Çalışma Grubu toplantısı ve 1991 yılındaki "Knowledge Discovery in Real Databases: A Report on the IJCAI-89 Workshop" başlıklı makalesinin bilgi keşfi ve veri madenciliği ile ilgili temel tanım ve kavramları ortaya koyması ile süreç daha da hızlanmış ve nihayet 1992 yılında veri madenciliği için ilk yazılım gerçekleştirilmiştir [16].

2000'li yıllarda veri madenciliği sürekli gelişmiş ve hemen hemen tüm alanlara uygulanmaya başlanmıştır. Alınan sonuçların faydaları göröldükçe, bu alana ilgi artmıştır [16].

Günümüzde veri madenciliği birçok alanda başarıyla kullanılmaktadır. Teknolojinin hızla gelişmesi yeni yöntemlerin oluşmasına zemin hazırlamıştır.

3.4. Veri Madenciliği Tanımı

Basit bir tanım yapmak gerekirse veri madenciliği, büyük ölçekli veriler arasından bilgiye ulaşma, bilgiyi madenleme işidir. Diğer bir ifadeyle büyük veri yığınları içerisinde gelecekle ilgili tahminlerde bulunabilmemizi sağlayabilecek bağıntıların bilgisayar programı kullanarak aranmasıdır. Veri madenciliği kavramına eş değer olarak veri tabanlarında bilgi madenciliği (knowledge mining from databases), bilgi

¹ *Perceptrons: An Introduction to Computational Geometry* isimli kitabın yazarlarıdır.

çıkarımı (knowledge extraction), veri ve örüntü analizi (data and pattern analysis) ve veri arkeolojisi gibi kavramlar da literatüre geçmiştir [18].

Bunların arasındaki en yaygın kullanım veri tabanlarında bilgi keşfidir. Alternatif olarak veri madenciliği aslında bilgi keşfi sürecinin bir parçası şeklinde kabul görmektedir [18].

Veri madenciliği üzerinde fikir birliğine varılmış ortak bir tanım yoktur. Veri madenciliği ile ilgili bazı tanımlar şöyledir:

- Veri madenciliği, büyük hacimli verilerden bilginin çıkarılması sürecidir. Bir başka ifade ile veri madenciliği büyük ve karmaşık verilerde beklenmeyen patikaların, değerli yapıların ve ilginç ilişkilerin keşfedilmesi bilimidir (sanatıdır) [19].
- Veri madenciliği, büyük veri tabanlarındaki gizli bilgi ve yapıyı açıklamak için, çok sayıda veri analizi aracını kullanan bir süreçtir [20].
- Veri madenciliği, iş kararlarının alınabileceği doğru, alışılmamış, faydalı ve anlaşılabilir örüntüler veya modellerdir [21].
- 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 ve daha fazla bilgiye dayalı karar vermesinde kullanabilmesi amacıyla önceden bilinmeyen, gizli, örtük, klasik yöntemlerle ortaya çıkarılması güç, faydalı, ilginç, anlaşılabilir; ilişki, örüntü, bağıntı veya trendlerin otomatik veya yarı otomatik bir şekilde ortaya çıkarılması sürecidir [22].

3.5. Veri Madenciliği Süreci

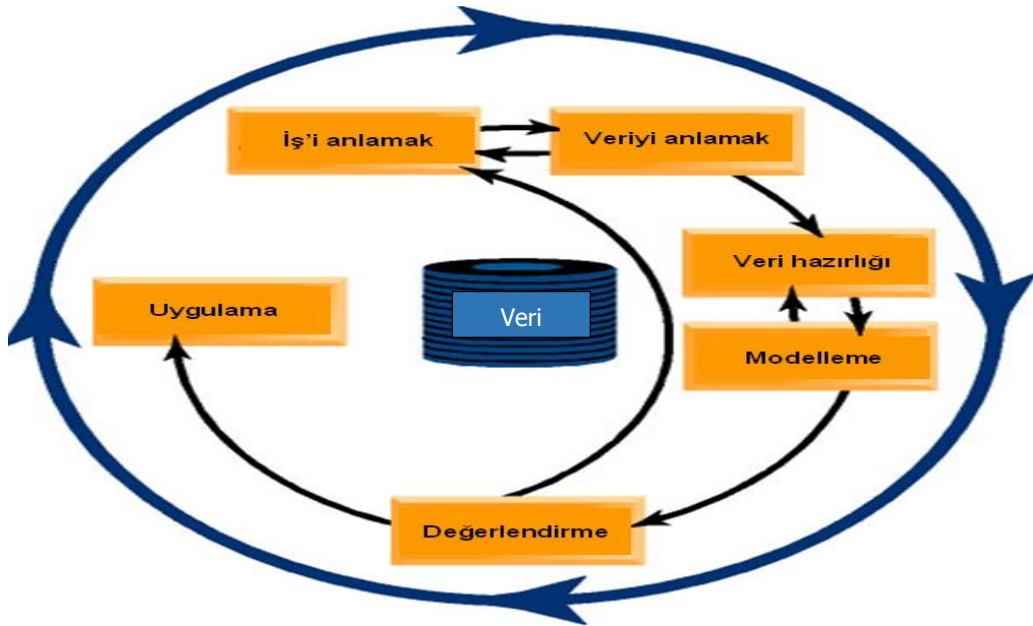
Birçok işletme kendi problemlerine, verilerine göre kendi veri madenciliği sürecini oluşturmaktadır. Ancak veri madenciliği sürecinin oluşturulmasında yapılan yanlışlıklar, sürecin etkinliğine zarar vermektedir. Veri madenciliği sürecinin standartlaştırılması konusunda farklı grup, kurum ve şirketler çeşitli standartlar oluşturmuşlardır. Bunlardan en çok takip edileni Daimler Chrysler ve SPSS tarafından 1996 yılında oluşturulan süreçtir [23].

Veri madenciliği süreci, projenin hayat döngüsüne genel bir bakış sağlar. Projenin birbirini takip eden aşamalarını, her aşamanın kendine ait görevlerini ve bu görevler

arasındaki ilişkileri kapsar. Bu tanımlama aşamasında bütün ilişkileri belirlemek mümkün değildir. Bütün veri madenciliği görevlerinde amaçlara, geçmişte yapılanlara, kullanıcının yararına ve en önemlisi veriye dayanan ilişkiler vardır [24].

Veri madenciliği projesinin hayat döngüsü 6 aşamadan oluşur. Aşamalar çok katı bir sıralama izlememektedir. Her zaman aşamalar arasında geriye ve ileriye hareketlilik gerekmektedir. Hareketlilik, her aşamanın sonucunda bundan sonra hangi aşamanın ya da aşamadaki hangi işin yapılması gerektiğine bağlıdır. Oklar, aşamalar arasındaki sıkça gerçekleşen en önemli bağımlılıkları göstermektedir [24].

Şekil 3.2.'de dıştaki halka veri madenciliğinin döngüsel doğasını sembolize etmektedir. Bir veri madenciliği süreci bir çözüm uygulandıktan sonra devam eder. Süreç sırasında öğrenilen dersler, işle ilgili yeni ve çoğunlukla daha odaklanmış soruların tetikleyicisi olur. Takip eden veri madenciliği süreçleri, bir öncekilerin tecrübelerinden faydalanır [24].



Şekil 3.2. CRISP DM (Cross Industry Standard Process Model for Data Mining) Süreci

3.5.1. İş'i Anlamak

Bu ilk aşama proje amaçlarını ve şartlarını işe yönelik bakış açısıyla anlamaya ve daha sonra bu bilgiyi veri madenciliği problem tanımına çevirmeye odaklanır. Amaçları gerçekleştirmek için bir başlangıç planı hazırlanır [24].

3.5.2. Veriyi Anlamak

Veriyi anlama aşaması, ilk veriyi toplama ile başlar ve veriye aşina olma, veri kalitesi ile ilgili problemlerin belirlenmesi, verideki ilk anlamları farketme ya da gizlenmiş bilgiye hipotez oluşturmak için ilginç altkümeleri belirleme faaliyetleri ile devam eder [24].

3.5.3. Veri Hazırlığı

Veri hazırlığı aşaması, ham veriden, en son veri setini (modelleme araç(lar)ına gönderilecek veri) oluşturma faaliyetlerinin hepsini kapsar [24].

Veri kalitesi, veri madenciliğinde anahtar bir konudur. Veri madenciliğinde güvenilirliğin artırılması için, veri ön işleme yapılmalıdır. Aksi halde hatalı girdi verileri bizi hatalı çıktıya götürecektir. Veri madenciliğinde veri kümesinin büyüklüğünden kaynaklanan en fazla zaman alıcı aşama, verilerin ön işlemden geçirilmesi aşamasıdır. Veri madenciliği uygulamalarında kaynakların %80'i verilerin ön işlemden geçirilmesi ve temizlenmesi süreçleri için harcanmaktadır. Çok sayıda veri ön işleme yöntemi mevcuttur [25, 26].

Veri ön işleme yöntemleri şu şekilde sıralanabilir:

1. Veri temizleme,
2. Veri birleştirme,
3. Veri dönüştürme,
4. Veri indirgeme.

Yukarıda verilen veri ön işleme yöntemleri verilerin kalitesini artıracaktır. Doğru kararların kaliteli veriler gerektirdiği bir gerçektir. Bu sebeple, veri ön işleme süreci veri madenciliğinin önemli bir adımıdır [25]. Bundan sonraki bölümlerde bu veri ön işleme yöntemleri açıklanacaktır.

3.5.3.1. Veri Temizleme

Veri temizleme, eksik verilerin tamamlanması, aykırı değerlerin belirlenmesi amacıyla gürültünün düzeltilmesi ve verilerdeki tutarsızlıkların giderilmesi gibi işlemleri gerektirmektedir [25]. Bu başlık altında veri temizleme için temel yöntemlere kısaca değinilecektir.

Herhangi bir değişkene ilişkin eksik değerlerin doldurulması için farklı yollar

vardır. Bunlardan bazıları aşağıda kısaca açıklanmaktadır [8, 25]:

1. Eksik değer içeren kayıt veya kayıtlar atılabilir.
2. Değişkenin ortalaması eksik değerlerin yerine kullanılabilir.
3. Aynı sınıfa ait tüm örneklem için değişkenin ortalaması kullanılabilir.

Örneğin aynı kredi risk kategorisine giren müşteriler için ortalama gelir değeri eksik değerler yerine kullanılabilir.

4. Var olan verilere dayalı olarak en uygun değer kullanılabilir. Burada sözü edilen en uygun değer belirlenmesi için regresyon veya karar ağacı gibi yöntemler kullanılabilir. Örneğin yaşı x , eğitim düzeyi y olan bir kişi için ücret durumu, mevcut verilerden yukarıdaki yöntemlerden birinin kullanılmasıyla tahmin edilebilir.

Veri temizleme yönteminin kullanılması gereken bir diğer problem ise gürültülü verilerdir. Gürültü, ölçülen değişkendeki varyans veya rassal bir hatadır. Gürültülü verilerin belirlenmesi amacıyla histogram, kümeleme analizi ve regresyon gibi yöntemler kullanılabilir. Eğer ele alınan değişken fiyat gibi sürekli bir değişken ise, gürültülü verilerin düzeltilmesi (smoothing) gerekmektedir. Aşağıda veri düzeltme yöntemlerinden bazıları açıklanmaya çalışılmıştır [25]:

Sepetlere ayırma (Binning): Sepetlere ayırma yöntemi, küçükten büyüğe veya büyükten küçüğe sıralanmış verileri düzeltmek için kullanılır. Yöntemde öncelikle sıralanmış veriler eşit büyüklükteki bin'lere ayrılır. Daha sonra sepetler, sepet ortalamaları, sepet medyanları veya sepet sınırları yardımıyla düzeltilir.

Örneğin değerler 4, 8, 15, 21, 21, 24, 25, 28 ve 34 olsun.

1. sepet: 4, 8, 15

1. sepet: 21, 21, 24

1. sepet: 25, 28, 34 olur.

Örneğin sepet ortalaması ile düzeltme yapılırsa,

Bin 1: 9, 9, 9

Bin 2: 22, 22, 22

Bin 3: 29, 29, 29 olur.

Kümeleme: Aykırı değerler kümeler ile belirlenebilir. Benzer değerler aynı grup veya

küme içinde yer alırken, aykırı değerler kümelerin dışında yer alacaktır.

Regresyon: Veriler regresyon ile verilere bir fonksiyon uydurularak düzeltilebilir. Uydurulan fonksiyona uymayan noktalar aykırı değerlerdir.

Veri temizlemeyi gerektiren bir diğer durum ise verilerdeki tutarsızlıklardır. Bazı veri tutarsızlıkları dışsal referansların kullanımıyla düzeltilebilir. Örneğin kodların kullanımındaki tutarsızlıklar düzeltilebilir. Verilerin birleştirilmesinden kaynaklanan tutarsızlıklar da olabilir. Bir değişken ismi farklı veri tabanlarında farklı şekillerde olabilir [25].

3.5.3.2. Veri Birleştirme

Veri madenciliğinde genellikle farklı veri tabanlarındaki verilerin birleştirilmesi gerekmektedir. Bu farklı veri tabanlarındaki veriler tek bir çatı altındaki veri ambarında birleştirilirler. Farklı veri tabanlarındaki verilerin tek bir veri tabanında birleştirilmesiyle şema birleştirme hataları (schema integration errors) oluşur. Örneğin, bir veri tabanında girişler “tüketici ID” şeklinde yapılmışken, bir diğerinde “tüketici numarası” şeklinde olabilir. Bu tür şema birleştirme hatalarından kaçınmak için meta veriler kullanılır. Veri tabanları ve veri ambarları genellikle meta veriye sahiptirler. Meta veri, veriye ilişkin veridir. Veri birleştirmede önemli bir konu da indirgemedir. Bir değişken, başka bir tablodan türetilmişse fazlalık olabilir. Değişkendeki tutarsızlıklar da, sonuçta elde edilen veri kümesinde fazlalıklara neden olabilir. Bu fazlalıklar korelasyon analizi ile araştırılabilir. Örneğin yukarıda da bahsedilen “tüketici ID” ile “tüketici numarası” korelasyon katsayısı bulunabilir. Eğer bulunan korelasyon katsayısı yüksek ise, değişkenlerden biri veri tabanından çıkarılarak indirgeme yapılır [25].

3.5.3.3. Veri Dönüştürme

Veri dönüştürme ile veriler, veri madenciliği için uygun formlara dönüştürülürler. Veri dönüştürme; düzeltme, birleştirme, genelleştirme ve normalleştirme gibi değişik işlemlerden biri veya bir kaçını içerebilir. Veri normalleştirme en sık kullanılan veri dönüştürme işlemlerinden birisidir. Veri normalleştirme yöntemlerinden bazıları aşağıdaki biçimde sıralanabilir [8, 25].

Min-Max

Min-max normalleştirme ile orijinal veriler yeni veri aralığına doğrusal dönüşüm ile

dönüştürülürler. Bu veri aralığı genellikle 0-1 aralığıdır.

Z Skor

Z skor normalleştirmede (veya 0 ortalama normalleştirme) ise değişkenin herhangi bir y değeri, değişkenin ortalaması ve standart sapmasına bağlı olarak bilinen Z dönüşümü ile normalleştirilir:

Ondalık Ölçekleme

Ondalık ölçekleme ile normalleştirmede ise, ele alınan değişkenin değerlerinin ondalık kısmı hareket ettirilerek normalleştirme gerçekleştirilir. Hareket edecek ondalık nokta sayısı, değişkenin maksimum mutlak değerine bağlıdır.

Örneğin 900 maksimum değer ise, $n=3$ olacağından 900 sayısı 0,9 olarak normalleştirilir.

3.5.3.4. Veri İndirgeme

Veri indirgeme yöntemleri, daha küçük hacimli olarak ve veri kümesinin indirgenmiş bir örneğinin elde edilmesi amacıyla uygulanır. Bu sayede elde edilen indirgenmiş veri kümesine veri madenciliği yöntemleri uygulanarak daha etkin sonuçlar elde edilebilir [25].

Veri indirgeme yöntemleri aşağıdaki biçimde özetlenebilir [25]:

1. Veri birleştirme veya veri küpü (Data aggregation or data cube)
2. Boyut indirgeme (Dimension reduction)
3. Veri sıkıştırma (Data compression)
4. Kesikli hale getirme (Discretization)

Veri birleştirme veya veri küpü yapılacak 2000-2003 yılları için çeyrek dönemlik satış tutarlarından oluşan bir veri kümesinin bulunduğunu varsayalım. Bu yıllar için yıllık satış tutarları tek bir tabloda toplandığında veri birleştirme yapılmış olur. Sonuç olarak elde edilen veri kümesinin hacmi daha küçüktür fakat yapılacak analiz için bir bilgi kaybı söz konusu değildir. Veri küpleri ise çok değişkenli birleştirilmiş bilginin saklandığı küplerdir. Örneğin bir firmanın satış tutarları yıllar, satışı yapılan ürünler ve firmanın farklı satış yerleri için aynı küp üzerinde gösterilebilir. Veri küpleri, özet bilgiye herhangi bir hesaplama yapmadan hızlı bir biçimde erişilmesini sağlarlar [25].

Veri madenciliği yapılacak veri kümesi bazen gereksiz olarak yüzlerce değişken içerebilir. Örneğin bir ürünün satışına ilişkin olarak düzenlenen bir veri kümesinde, tüketicilerin telefon numaraları gereksiz bir değişken olarak yer alabilir. Bu tür gereksiz değişkenler elde edilecek örüntüleri kalitesizleştirebileceği gibi veri madenciliği sürecinin yavaşlamasına da yol açacaktır. Gereksiz değişkenlerin elenmesi amacıyla ileri veya geri yönlü olarak sezgisel seçimler yapılabilir. İleri yönlü sezgisel seçimde orijinal değişkenleri en iyi temsil edecek değişkenler belirlenir. Ardından her bir değişken veya değişkenler grubunun, bu kümeye dahil edilip edilmeyeceği sezgisel olarak belirlenir. Geri yönlü sezgisel seçimde ise öncelikle değişkenlerin tüm kümesi ele alınır. Daha sonra gereksiz bulunan değişkenler kümeden dışlanarak, en iyi değişken kümesi elde edilmeye çalışılır. Boyut indirgeme amacıyla kullanılacak bir diğer yöntem ise karar ağaçlarıdır. Karar ağaçları ele alınacak çıktı değişkenini en iyi temsil edecek değişken kümesini verecektir [25].

Veri sıkıştırımda ise orijinal verileri temsil edebilecek indirgenmiş veya sıkıştırılmış veriler, veri şifreleme veya dönüşümü ile elde edilirler. Bu şekilde indirgenmiş veri kümesi, orijinal veri kümesini bir bilgi kaybı olacak biçimde temsil edebilecektir. Bununla beraber bilgi kaybı olmaksızın indirgenmiş veri kümesi elde edilmesine yarayacak bir takım algoritmalar da mevcuttur. Bu algoritmalar bir takım sınırlamalara sahip olduklarından sıkça kullanılamamaktadır. Bununla beraber temel bileşenler analizi gibi yöntemler, bir bilgi kaybına göz yumularak sıkıştırılmış veri kümesi elde edilmesinde kullanışlıdır [25].

Kesikleştirme, bazı veri madenciliği algoritmaları yalnızca kategorik değerleri ele aldığından, sürekli verilerin kesikli değerlere dönüştürülmesini içerir. Bu şekilde sürekli verilerin kesikli değer aralıklarına dönüştürülmesiyle elde edilen kategorik değerler, orijinal veri değerlerinin yerine kullanılırlar. Bir kavram hiyerarşisi (concept hierarchy), verilen sürekli değişken için, değişkenin ayrıştırılması olarak tanımlanabilir. Kavram hiyerarşileri, düşük düzeyli kavramların yüksek düzeyli kavramlarla değiştirilmesiyle verilerin indirgenmesinde kullanılır. Örneğin yaş değişkeni 1-15, 16-40, 40+ olacak biçimde daha yüksek kavram düzeyinde ifade edilebilir. Bu şekilde veri indirgemede detay bilgiler kayboluyorsa da, genelleştirilmiş veriler daha anlamlı olacak, daha kolay yorumlanabilecek ve orijinal verilerden daha düşük hacim kaplayacaktır [25].

Veri madenciliği programları veri ön işleme yöntemlerinden birçoğunu

gerçekleştirmektedir. Bununla beraber veri işleme ile ilgili özel programlar veya veri ön işleme açısından güçlü bazı özel programlar vardır.

3.5.4. Modelleme

Bu aşamada, farklı modelleme yöntemleri seçilir ve uygulanır ve parametreleri optimum değerlere ayarlanır. Tipik olarak aynı veri madenciliği problem türü için bir kaç yöntem vardır. Bazı yöntemler verinin formu için belli şartlar gerektirir. Bu nedenle çoğunlukla veri hazırlığı aşamasına dönmek gerekmektedir [24].

3.5.5. Değerlendirme

Uygulama aşamasına geçmeden, modeli baştan aşağı değerlendirmek ve iş amaçlarını tam olarak yerine getirdiğinden emin olmak için modeli oluştururken uygulanan adımları gözden geçirmek önemlidir. Başlıca nokta önemli iş konularının yeterince dikkate alınıp alınmadığının kontrol edilmesidir. Bu aşamanın sonunda veri madenciliği sonuçlarının kullanımı ile ilgili bir sonuca varılmalıdır [24].

3.5.6. Uygulama

Modeli oluşturmak genelde projenin sona erdiği anlamına gelmez. Modelin amacı verinin bilgisini, değerini arttırmak olsa da, elde edilen bilgi müşterinin kullanabileceği şekilde düzenlenmeli ve sunulmalıdır. Şartlara bağlı olarak uygulama aşaması, rapor üretmek kadar basit ya da tekrarlanabilen veri madenciliği süreci oluşturmak kadar karmaşık olabilir. Çoğu zaman uygulama adımlarını yürüten kişi veri analisti değil müşteridir. Uygulamayı analist yürütmeyecek bile olsa, müşterinin yaratılacak modellerden faydalanmak için hangi adımları atması gerektiğini anlaması önemlidir [24].

3.6. Veri Madenciliği Ne Değildir?

Toplanan verilerden yapılacak sorgulamalar ve detaylı analizler ile elde edilen sonuçlar veri madenciliği olarak değerlendirilmemelidir. Örneğin; bir süpermarket zincirinde, şubelerin cirolarını ve hangi ürünlerin hangi şubede daha fazla satıldığını sorgulamak, bir satış şirketinde hangi müşterilerin süreklilik gösterdiğini belirlemek tam bir veri madenciliği olarak değerlendirilemez. Aynı şekilde yalnızca regresyon analizi yaparak gelir ile cinsiyet arasındaki ilişkiyi modellemek de veri madenciliği değildir. Veri madenciliğinin ne olmadığı ve ne olması gerektiği Çizelge 3.1.'de birkaç örnekle gösterilmiştir [27, 28].

Çizelge 3.1. Veri Madenciliği Ne Değildir?

NE DEĞİLDİR	NE OLMALIDIR
İnternette ayrıntılı bilgi araştırmak.	İnternette aynı içerikteki benzer bilgileri gruplamak.
Aynı hastalığa sahip hasta kayıtlarını sorgulamak.	Benzer semptomlar görülen aynı hastalığa sahip hastaları gruplamak.
Yer listesinden termal otellerin yerini sorgulamak.	Termal otelleri, hangi hastalığın tedavisi ile ilgili olduğuna göre gruplamak.
Şirketlerin finansal raporlarından tabloları analiz etmek.	Şirketlerin satış ile ilgili veri tabanlarından müşteri profillerini ortaya çıkarmak.

3.7. Tahminleme

3.7.1. Tahmin Kavramı

Tahmin, gündelik dilde, bilinmeyen bir şey hakkında rastgele veya belirsiz bilgi kullanarak bir sonuca varmaktır [29].

Tahmin, özellikle geleceğe dönük olarak yapılıyorsa bu amaçla “öngörü” ya da “kestirim” kavramları kullanılabilir. İngilizce’de, mevcut duruma yönelik tahminler için “estimation” ve “prediction” kavramları, geleceğe dönük tahminler için ise “forecasting” ve “prediction” kavramları kullanılmaktadır. “Prediction” kavramına benzer şekilde, Türkçe’de de tahmin kavramı hem mevcut durumlara hem de geleceğe yönelik varsayımlarda genel amaçlı kullanılmaktadır [30].

3.7.2. Tahmin Yöntemleri

Kullanılacak tahmin yöntemi belirlenirken; verilere erişebilme, elde edilen verilerin niteliği ve seyri, kullanılan kaynakların maliyeti, hata payı, anlama ve uygulama kolaylığı, tahminlerin kapsadığı zaman aralığı ve tahminlerin hazırlanması için gerekli olan süre gibi pek çok faktörün dikkate alınması gerekmektedir. Tahmin yöntemleri en genel itibarıyla “nitel” (kalitatif) ve “nicel” (kantitatif) olarak iki ana grupta ayrıntılı olarak incelenebilir [30].

3.7.2.1. Nitel Tahmin Yöntemleri

“Öznel”, “sübjektif” ya da “yargısal” yöntemler olarak da adlandırılabilen nitel tahmin yöntemleri, tahmin edilmeye çalışılan olgunun anlaşılmasının zor olduğu, bu olguya ilişkin sayısal verilerin elde edilemediği ya da olgunun sayısal olarak ifade edilemeyen bilgiler tarafından etkilendiği, belirsizliğin ve verilerin değişkenliğinin çok

fazla olduđu durumlarda kullanılırlar. Bu yöntemler, mevcut durumlara ve gelecekle ilgili plânlara ilişkin bilgisi olması beklenen kişilerin görüşlerinin toplanmasını gerektirmektedir [30].

Kalitatif yöntemler soyut faktörlerin ve subjektif deneyimlerin ele alınabilmesini sağlamasına karşın, karar verme sürecinde kişisel önyargı ve eğilimlerin yer almasından dolayı genellikle daha düşük performansa sahip tahminlerle sonuçlanmaktadır [31].

3.7.2.2. Nicel Tahmin Yöntemleri

Analizlerde girdi olarak soyut verileri değil, sayısal verileri kullanan yöntemler nicel tahmin yöntemleri olarak adlandırılır. Bu yöntemlerde, verilerin geçmişteki sayısal değerleri kullanılarak elde edilen şablon yardımıyla uygun bir matematiksel model kurulur ve geleceğe yönelik tahminlerde bulunmak için bu modelin tahmin denklemleri kullanılır. Günümüzde nicel yöntemlerde kullanılan sayısal veriler; barkod okuyucuları, çevrimiçi işlemler veya gözlem sonuçlarından elde edilen verilerin bilgisayar yardımıyla saklandığı veri tabanlarından alınmaktadır [30, 31].

Nicel tahmin yöntemleri; nedensel yöntemler ve zaman serileri olmak üzere iki ana gruba ayrılmaktadır [30, 31].

Nedensel ya da ilişki tahmin modelleri (diğer bir ifadeyle regresyon analizi), değerleri tahmin edilecek değişkenle ilişkili olan diğer değişkenlerin belirlenmesini içermektedir. Bu değişkenler belirlendikten sonra geliştirilen istatistiksel model, tahmin edilecek değişken ile diğer değişkenler arasındaki ilişkiyi tanımlamakta ve ele alınan değişkene ilişkin tahminler yapılmasında kullanılmaktadır. Bu modellerin özellikle işletmelerde yoğun olarak kullanılmasının nedeni, yönetimin çeşitli alternatif politikaların etkilerini değerlendirmesine imkân tanınmasıdır. Fakat modelin geliştirilmesinin zor olması, tüm değişkenlere ilişkin geçmiş verilere gereksinim duyulması ve bunun gerektireceği zaman ve maliyet nedeniyle bazı zayıflıklara sahip olduğu unutulmamalıdır [30, 31].

Zaman serileri analizinde, tahmin edilecek değişkene ilişkin geçmiş veriler belirli bir veri seyri elde etmek üzere analiz edilmektedir. Bu nedenle tahminleme sadece geçmiş verilerin bu amaçla analiz edilmesine ve yapılacak tahminlerde kullanılmasına dayanmaktadır. Bu özelliğinden dolayı zaman serileri analizi, değişmeyen koşullar altında daha etkin olmaktadır [30, 31].

Veri madenciliği yöntemleri arasında yer alan yapay sinir ağları, hem nedensel modellerin hem de zaman serilerine dayanan modellerin kurulmasında başarıyla kullanılabilir [30, 31].

3.7.3. Tahmin Modellerinin Doğruluklarının Ölçülmesi

Bir tahmin modelinin doğruluğu, tahmin değerlerinin gerçek gözlem değerlerine ne kadar yakın olduğuna bağlıdır. Bu doğruluğu ölçmek için, ortaya konan model kullanılarak gerçek gözlem değerleri bilinmiyormuşçasına tekrar hesaplanarak tahmin edilir. Ardından bu tahmin değerleri ile gerçek gözlem değerleri arasındaki farklar, çeşitli istatistikler yardımıyla karşılaştırılır. Tahmini gözlem değerleri ile gerçek gözlem değerleri arasındaki bu farklılara tahmin hatası adı verilmektedir [30].

Tahmin modellerinin doğruluğunu ölçmek için yaygın olarak aşağıdaki istatistikler kullanılmaktadır [29]:

- Ortalama mutlak hata (Mean absolute error, MAE),
- Ortalama mutlak yüzde hata (Mean absolute percentage error, MAPE),
- Hata kareler ortalamasının karekökü (Root mean squared error, RMSE),
- Hata kareler ortalaması (Mean squared error, MSE),
- Ortalama yüzde hata (Mean percentage error, MPE),
- Ortalama mutlak sapma (Mean absolute deviation, MAD),
- Yüzde ortalama mutlak sapma (Percent mean absolute deviation, PMAD),
- Tahmin kabiliyeti (Forecast skill, SS),
- Ortalama hata (Average of errors, E).

Bu istatistikler arasında en çok kullanılanların MAE, MAPE, RMSE ve MSE olduğu görülmüştür [30].

Tahmin hatası ile tahmin doğruluğu ters orantılıdır. Tahmin hatası küçüldükçe modelin doğruluk derecesi artmaktadır. Farklı tahmin modellerini karşılaştırırken en iyi model ideal olarak, kullanılan istatistiklerin tamamında en küçük değeri veren modeldir. Ancak, verilerin yapısına bağlı olarak, özellikle aşırı sapmalar nedeniyle bu ideallik her zaman sağlanamaz ve bazı istatistikler abartılı sonuçlar verebilir. Bu tür durumlarda, hataları yüzde olarak ifade etmesi nedeni ile tek başına da bir anlam taşıyan MAPE istatistiği referans alınıp doğrulukları karşılaştırmada tercih edilebilir. MAPE değeri % 10'un altında olan modeller "çok iyi", % 10 ile % 20 arasında olan modeller "iyi", % 20 ile % 50 arasında olan modeller "kabul edilebilir"

ve % 50'nin üzerinde olan modeller ise "yanlış ve hatalı" olarak kabul edilmektedir [30].

En iyi tahmin modelinin belirlenmesinde, tahmin hatası istatistiklerinin yanında aşağıdaki kriterlerin de göz önünde bulundurulmasında fayda vardır [30]:

- Veriye ulaşılabilirlik,
- Veri toplama maliyeti,
- Modelleri çalıştıracak bilgisayar yazılımlarının varlığı ve maliyeti,
- Modelin kurulduğu zaman çerçevesi,
- İstenilen tahmin çeşidi (nokta/aralık).

Çalışmada kullanılan tahmin doğruluğunu ölçme istatistiklerine ait formüller Çizelge 3.2'de verilmiştir:

Çizelge 3.2. Tahmin Doğruluğunu Ölçme İstatistikleri

Ortalama Mutlak Hata <i>Mean Absolute Error (MAE)</i>	$MAE = \frac{\sum_{t=1}^N E_t }{N}$
Ortalama Mutlak Yüzde Hata <i>Mean Absolute Percentage Error (MAPE)</i>	$MAPE = \frac{\sum_{t=1}^N \left \frac{E_t}{Y_t} \right }{N}$
Hata Kareler Ortalamasının Karekökü <i>Root Mean Squared Error (RMSE)</i>	$RMSE = \sqrt{\frac{\sum_{t=1}^N E_t^2}{N}}$
Hata Kareler Ortalaması <i>Mean Squared Error (MSE)</i>	$MSE = \frac{\sum_{t=1}^N E_t^2}{N}$

Formüllerdeki,

$$E_t = Y_t - F_t \text{ 'dir.}$$

E= t dönemindeki tahmin hatası,

Y= t dönemindeki gerçek gözlem değeri,

F= t dönemindeki tahmin gözlem değeridir.

4. VERİ MADENCİLİĞİ MODELLERİ

Veri madenciliğinde kullanılan modeller, tahmin edici (*Predictive*) ve tanımlayıcı (*Descriptive*) olmak üzere iki ana başlık altında incelenmektedir [32].

Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır. Örneğin, bir banka önceki dönemlerde vermiş olduğu kredilere ilişkin gerekli tüm verilere sahip olabilir. Bu verilerde bağımsız değişkenler kredi alan müşterinin özellikleri, bağımlı değişken değeri ise kredinin geri ödenip ödenmediğidir. Bu verilere uygun olarak kurulan model, daha sonraki kredi taleplerinde müşteri özelliklerine göre verilecek olan kredinin geri ödenip ödenmeyeceğinin tahmininde kullanılmaktadır [32, 33].

Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. X/Y aralığında geliri ve iki veya daha fazla arabası olan çocuklu aileler ile, çocuğu olmayan ve geliri X/Y aralığından düşük olan ailelerin satın alma örüntülerinin birbirlerine benzerlik gösterdiğinin belirlenmesi tanımlayıcı modellere bir örnektir [32, 34].

Veri madenciliği modellerini gördükleri işlemlere göre,

- Sınıflama ve Regresyon
- Kümeleme
- Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

olmak üzere üç ana başlık altında incelemek mümkündür. Sınıflama ve regresyon modelleri tahmin edici, kümeleme ve birliktelik kuralları modelleri tanımlayıcı modellerdir [34].

4.1. Sınıflama ve Regresyon

Sınıflama ve regresyon, önemli veri sınıflarını ortaya koyan veya gelecek veri eğilimlerini tahmin eden modelleri kurabilen iki veri analiz yöntemidir. Sınıflama kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır. Örneğin, bir sınıflama modeli banka kredi uygulamalarının güvenli veya riskli olmalarını kategorize etmek amacıyla kurulurken, regresyon modeli geliri ve mesleği verilen potansiyel müşterilerin bilgisayar ürünleri alırken yapacakları harcamaları tahmin etmek için kurulabilir [35].

Sınıflama ve regresyon modellerinde kullanılan başlıca yöntemler şunlardır [34]:

4.1.1. Karar Ağaçları

Karar ağaçları, veri madenciliğinde kuruluşlarının ucuz olması, yorumlanmalarının kolay olması, veri tabanı sistemleri ile kolayca entegre edilebilmeleri ve güvenilirliklerinin iyi olması nedenleri ile sınıflama modelleri içerisinde en yaygın kullanıma sahip yöntemdir [33].

Karar ağacı, adından da anlaşılacağı gibi bir ağaç görünümünde, tahmin edici bir yöntemdir. Ağaç yapısı ile, kolay anlaşılabilen kurallar yaratabilen, bilgi teknolojileri işlemleri ile kolay entegre olabilen en popüler sınıflama yöntemidir [36].

Karar ağacı karar düğümleri, dallar ve yapraklardan oluşur. Karar düğümü, gerçekleştirilecek testi belirtir. Bu testin sonucu ağacın veri kaybetmeden dallara ayrılmasına neden olur. Her düğümde test ve dallara ayrılma işlemleri ardışık olarak gerçekleşir ve bu ayrılma işlemi üst seviyedeki ayrımlara bağlıdır. Ağacın her bir dalı sınıflama işlemini tamamlamaya adaydır. Eğer bir dalın ucunda sınıflama işlemi gerçekleşemiyorsa, o dalın sonucunda bir karar düğümü oluşur. Ancak dalın sonunda belirli bir sınıf oluşuyorsa, o dalın sonunda yaprak vardır. Bu yaprak, veri üzerinde belirlenmek istenen sınıflardan biridir. Karar ağacı işlemi kök düğümünden başlar ve yukarıdan aşağıya doğru yaprağa ulaşana dek ardışık düğümleri takip ederek gerçekleşir [33, 35].

Karar ağacı yöntemini kullanarak verinin sınıflanması iki basamaklı bir işlemdir. İlk basamak öğrenme basamağıdır. Öğrenme basamağında önceden bilinen bir eğitim verisi, model oluşturmak amacıyla sınıflama algoritması tarafından analiz edilir [33, 35].

Öğrenilen model, sınıflama kuralları veya karar ağacı olarak gösterilir. İkinci basamak ise sınıflama basamağıdır. Sınıflama basamağında test verisi, sınıflama kurallarının veya karar ağacının doğruluğunu belirlemek amacıyla kullanılır. Eğer doğruluk kabul edilebilir oranda ise, kurallar yeni verilerin sınıflanması amacıyla kullanılır [33].

Test verisine uygulanan bir modelin doğruluğu, yaptığı doğru sınıflamanın test verisindeki tüm sınıflara oranıdır. Her test örneğinde bilinen sınıf, model tarafından tahmin edilen sınıf ile karşılaştırılır. Eğer modelin doğruluğu kabul edilebilir bir değer ise model, sınıfı bilinmeyen yeni verileri sınıflama amacıyla kullanılabilir [33].

En yaygın kullanılan karar ağacı algoritmaları aşağıda verilmiştir [23].

- CHAID (Chi-Squared Automatic Interaction Detector),
- C&RT (Classification and Regression Trees),
- ID3 (Iterative Dichotomiser 3),
- C4.5 (Extension of ID3 Algorithm).

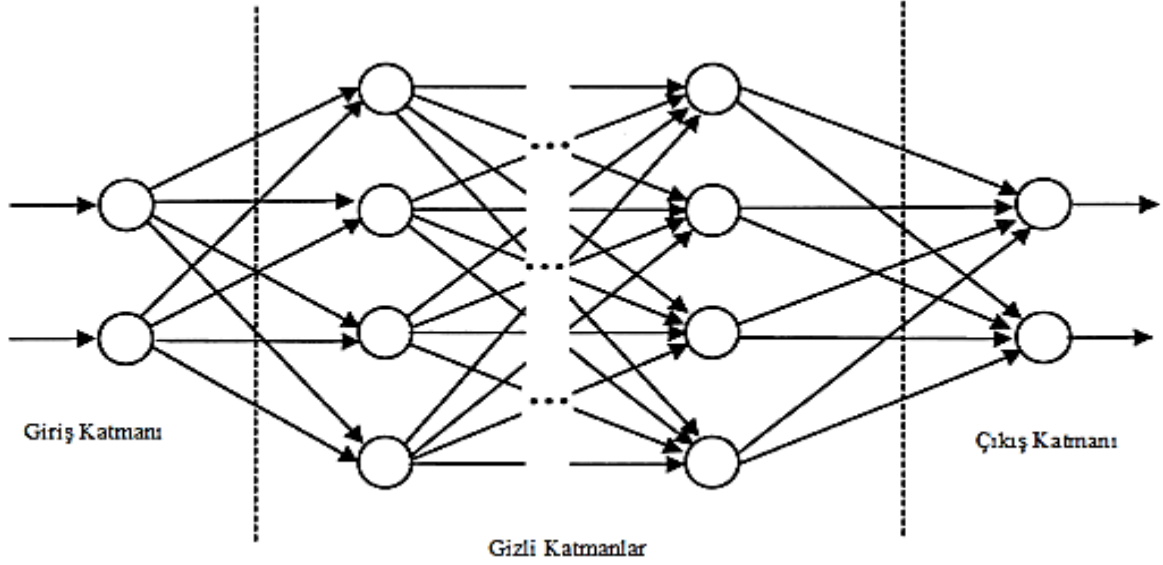
4.1.2. Yapay Sinir Ağları

Yapay sinir ağları (YSA) öğrenme yeteneğine sahip, gelişmiş matematiksel yapıların hesaplanmasını içeren bir yaklaşımdır. Bu yöntem, sinir sisteminin öğrenmesini model alan akademik araştırmaların bir sonucu olarak ortaya çıkmıştır. Sinir ağları karmaşık ve anlaşılması çok güç olan yapılardan anlam türetme becerisine sahip, dikkate değer yeteneklere sahiptir.

YSA, insan beyninin çalışma ilkelerinden ilham alınarak geliştirilmiştir. Ağırlıklı bağlantılar denilen tek yönlü iletişim kanalları vasıtası ile birbirleriyle haberleşirler ve her biri kendi hafızasına sahip birçok işlem elemanından (nöronlardan) oluşurlar. YSA'lar gerçek dünyaya ait ilişkileri tanıyabilir, sınıflandırma, kestirim ve işlev uydurma gibi görevleri yerine getirebilirler [38].

YSA genellikle bir giriş katmanı, gizli katmanlar ve bir çıkış katmanından oluşmaktadır. Basit şekliyle her bir nöron bir önceki katmanlardaki diğer nöronlara, sinaptik ağırlıkları yoluyla bağlanmaktadır [39].

Biyolojik sistemlerde öğrenme, nöronlar arasındaki sinaptik bağlantıların ayarlanması ile gerçekleşmektedir. YSA'da ise öğrenme, girdi ve çıktı verilerinin işlenmesiyle, yani eğitme algoritmasının bu verileri kullanarak bağlantı ağırlıklarını bir yakınsama sağlanana kadar, tekrar tekrar ayarlamasıyla gerçekleşmektedir. YSA'nın katmanları ve işleyişi Şekil 4.1.'de gösterilmiştir [39, 40].



Şekil 4.1. Yapay Sinir Ağlarının Katmanları

4.1.3. Genetik Algoritmalar

Darwin'in "en iyi olan yaşar (survival of the fittest)" prensibine dayalı olarak bir popülasyonu oluşturan bireylerin rekabet etmelerini ve rekabet sonucu elenmelerini sağlayan, evrimsel süreci simüle eden Genetik Algoritmalar (GA), ilk olarak John Holland tarafından ortaya atılmıştır [37].

GA'lar, çözümlerin kodlanması, uygunlukların hesaplanması, çoğalma, çaprazlama ve mutasyon işlemlerinin uygulanması gibi aşamaları içermektedir. Çözümlerin kodlanması aşamasında, tüm çözümlerin aynı boyutlara sahip bitler dizisi biçiminde gösterilmektedir. Popülasyondaki her üyenin uygunluk değeri hesaplanarak çoğalma aşamasına geçilmekte ve mevcut kuşaktan yeni bir popülasyon yaratılmaktadır. Mevcut gen havuzunun potansiyelini araştırmak için, bir önceki kuşaktan daha iyi nitelikler içeren yeni kromozomlar yaratmak amacıyla çaprazlama operatörü kullanılmakta ve genetik çeşitliliği korumak amacıyla mutasyon işlemi uygulanmaktadır. Tüm bu işlemlerden sonra yeni kuşak oluşturulmakta ve döngü durdurulmaktadır [40].

GA'lar, problemlerin çözümü için evrimsel süreci bilgisayar ortamında taklit ederler. Çözüm için tek bir yapının geliştirilmesi yerine, böyle yapılardan meydana gelen bir küme oluştururlar. Problem için olası pek çok çözümü temsil eden bu küme genetik algoritma terminolojisinde nüfus adını almaktadır. Nüfuslar vektör, kromozom veya birey adı verilen sayı dizilerinden oluşmaktadır. Birey içindeki her bir elemana gen

denir. Nüfustaki bireyler evrimsel süreç içinde genetik algoritma işlemcileri tarafından belirlenmektedirler. Genetik algoritmalar yapısı gereği, kötü bireyleri yani uygun olmayan çözümleri, operatörleri sayesinde elemektedir. Bu işlemler bir döngü içerisinde durdurma kriteri sağlanana kadar devam etmektedir [42].

GA'lar, çizelgeleme, tesis yerleşimi, hat dengeleme, atama ve optimizasyon problemlerinin çözümü ile finans, pazarlama ve üretim gibi alanlarda uygulanmaktadır [41].

4.1.4. K-En Yakın Komşu

Veri uzayında birbirine yakın olan aynı tür kayıtlar, birbirlerinin komşusu durumundadırlar. Bu anlayış doğrultusunda, çok kolay fakat güçlü olan k-en yakın komşu algoritması geliştirilmiştir. K-en yakın komşu algoritmasının temel felsefesi komşunun yaptığını yaptırmaktır. Belirli bir bireyin (kayıtın) davranışını (özelliğini) tahmin etmek istenirse, veri uzayında o bireye yakın olan örneğin 10 bireyin davranışına bakılabilir. Bu 10 komşunun davranışının ortalaması hesaplanır ve bu hesaplanan ortalama bireylerin tahmini olur. K-en yakın komşudaki k harfi, araştırdığımız komşu sayısıdır. Örneğin, 5-en yakın komşuda 5 komşuya bakılır [43].

4.1.5. Bellek Temelli Nedenleme

İnsanlar kararlarını genellikle daha önce yaşadıkları deneyimlere göre verirler. Örneğin doktorlar bir hastayı incelerken, elde ettiği bulguları daha önce tedavi ettiği benzer hastalığa yakalanmış hastalar üzerindeki deneyimlerini kullanarak değerlendirirler. Bellek tabanlı yöntemler de benzer şekilde deneyimleri kullanmaktadır. Bu yöntemlerde, bilinen kayıtların bulunduğu bir veri tabanı oluşturulur ve sistem yeni gelen bir kayda komşu olan diğer kayıtları belirler ve bu kayıtları kullanarak tahminde bulunur ya da bir sınıflama işlemi uygular. Bellek tabanlı yöntemlerin en önemli özelliği veriyi olduğu gibi kullanabilme yeteneğidir. Diğer VM yöntemlerinin aksine bellek tabanlı yöntemler, kayıtların şekli (format) yerine sadece iki işlemin varlığı ile ilgilenir. Bu işlemler, iki kayıt arasındaki uzaklığı belirleyen bir uzaklık fonksiyonu ve komşu kayıtları işleyerek bir sonuç üreten kombinasyon fonksiyonudur [23].

4.1.6. Naif-Bayes

Naif-Bayes (NB), temeli Bayes teorisine dayanan, verileri istatistiksel sınıflandırma yöntemlerinden biridir. Veri madenciliği sınıflandırma algoritmalarından olan Bayes, uygulanabilirliği ve hızlı hesaplama performansı ile araştırmacılar tarafından öne çıkan bir algoritmadır. Sınıflandırılacak olayları birbirinden bağımsız olarak ele almaktadır [44].

NB algoritmasının uygulanmasında bir takım kabuller yapılmaktadır. Bunların en önemlisi niteliklerin birbirinden bağımsız olmasıdır. Nitelikler birbirini etkilediği takdirde, burada olasılık hesaplamak zorlaşacağı için niteliklerin hepsinin aynı derecede önemli olduğu kabul edilmektedir [40].

NB, sürekli veri ile çalışmaz. Bu nedenle sürekli değerleri içeren bağımlı ya da bağımsız değişkenler kategorik hale getirilmelidir. NB, modelin öğrenilmesi esnasında, her çıktının öğrenme kümesinde kaç kere meydana geldiğini hesaplamaktadır. Bulunan bu değer, öncelikli olasılık olarak adlandırılır. NB aynı zamanda her bağımsız değişken / bağımlı değişken kombinasyonunun meydana gelme sıklığını bulmaktadır. Bu sıklıklar öncelikli olasılıklarla birleştirilmek suretiyle tahminde kullanılır [45].

4.2. Kümeleme

Kümeleme analizi bir veri kümesindeki bilgileri belirli yakınlık kriterlerine göre gruplara ayırma işlemidir. Bu grupların her birine “küme” adı verilir. Kümeleme analizine kısaca “kümeleme” adı verilir. Kümeleme işleminde küme içindeki elemanların benzerliği fazla, kümeler arası benzerlik ise az olmalıdır. Kümeleme veri madenciliği yöntemlerinden tanımlayıcı modellere yani gözetimsiz sınıflandırmaya girer [46].

Gözetimsiz sınıflamada amaç, başlangıçta verilen ve henüz sınıflandırılmamış bir küme, veriyi anlamlı alt kümeler oluşturacak şekilde öbeklemektir. Kümeleme işlemi tamamen gelen verinin özelliklerine göre yapılır [47].

Kümeleme analizinin kullanılmasında benzer uzaklıklar dikkate alınarak yararlanılabilecek alternatif ölçüt ve yöntemler bulunmaktadır. Birimler arası uzaklıklar için öklid, Manhattan, Mahalanobis, Minkowski veya Canberra uzaklık ölçütleri kullanılabilir. Bu da kümeleme analizinin uygulamada kullanılmasında dikkatli davranmayı zorunlu kılmaktadır. Kümeleme algoritması veri

tabanını alt kümelere ayırır. Her bir kümede yer alan elemanlar dâhil oldukları grubu diğer gruplardan ayıran ortak özelliklere sahiptir. Kümeleme modellerinde amaç, küme üyelerinin birbirlerine çok benzediği, ancak özellikleri birbirlerinden çok farklı olan kümelerin bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir [48].

Kümeleme uygulamalarında üç görevi yerine getirmiş oluruz [49]:

1. Veri setlerini kümeler içinde bölümlere ayırma,
2. Kümeleme sonuçlarını doğrulama,
3. Kümeleri yorumlama.

Kullanıcının amacına ve kullanım alanına göre kümeleme analizinin amaçları aşağıdaki gibi sıralanabilir [50].

1. Doğru türlerin belirlenmesi,
2. Model oluşturmak,
3. Gruplara dayalı tahmin,
4. Hipotez testi,
5. Veri araştırma (inceleme),
6. Hipotez oluşturma,
7. Veri indirgeme.

Kümeleme yöntemleri hiyerarşik ve hiyerarşik olmayan (bölümlemeli) kümeleme yöntemleri olmak üzere iki bölüme ayrılırken bu konuda yapılan araştırmalar bu algoritmaların daha alt bölümlere ayrılabilceğini göstermektedir. Hiyerarşik kümeleme yönteminde özellikle işleyişin daha kolay anlaşılabilmesi için dendogram (ağaç grafiği) dan yararlanılır. Hiyerarşik kümeleme yönteminde anlatılan işlemlere dayalı olarak kullanılan hiyerarşik yöntemlerden en çok kullanılanları; tek bağlantılı, tam bağlantılı, ortalama bağlantı, merkezi ve ward yöntemleridir. Hiyerarşik olmayan kümeleme yöntemi ise küme sayısı hakkında bir ön bilginin olması ya da araştırmacının anlamlı olacak küme sayısına karar vermiş olması durumunda tercih edilmektedir. Hiyerarşik olmayan kümeleme yönteminde en çok tercih edilen iki yöntem k-ortalama ve en çok olabilirlik yöntemleridir [46].

4.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

Birliktelik kuralı, geçmiş verilerin analiz edilerek bu veriler içindeki birliktelik davranışlarının tespiti ile geleceğe yönelik çalışmalar yapılmasını destekleyen bir yaklaşımdır [40].

Bir alışveriş sırasında veya birbirini izleyen alışverişlerde müşterinin hangi mal veya hizmetleri satın almaya eğilimli olduğunun belirlenmesi, müşteriye daha fazla ürünün satılmasını sağlama yollarından biridir. Satın alma eğilimlerinin tanımlanmasını sağlayan birliktelik kuralları ve ardışık zamanlı örüntüler, pazarlama amaçlı olarak “pazar sepeti analizi” adı altında veri madenciliğinde yaygın olarak kullanılmaktadır. Bununla birlikte bu yöntemler, tıp, finans ve farklı olayların birbirleri ile ilişkili olduğunun belirlenmesi sonucunda değerli bilgi kazanımının söz konusu olduğu ortamlarda da önem taşımaktadır [34].

Sepet analizinde amaç alanlar arasındaki ilişkileri bulmaktır. Bu ilişkilerin bilinmesi şirketin kârını arttırmak için kullanılabilir. Eğer X malını alanların Y malını da yüksek olasılıkla aldıklarını biliyorsanız ve eğer bir müşteri X malını alıyor ama Y malını almıyorsa o potansiyel bir Y müşterisidir [51].

Örneğin internet üzerinden kitap satan Amazon şirketi (<http://www.amazon.com>) *BookMatcher* adlı programıyla müşterilerine okudukları ve sevdikleri kitaplara göre satın almaları için kitap tavsiye etmektedir [51].

Eğer elimizdeki veride mallar için sadece satın alındı/alınmadı bilgisi varsa, sepet analizinde mallar arasındaki bağıntı, destek ve güven kıstasları aracılığıyla hesaplanır. İki mal, X ve Y, için destek (*support*) ve güven (*confidence*) tanımları şöyledir [51]:

Destek:

$P(X \text{ ve } Y) = X \text{ ve } Y \text{ mallarını satın almış müşteri sayısı} / \text{Toplam müşteri sayısı}$

Güven:

$P(X|Y) = P(X \text{ ve } Y) / P(Y) = X \text{ ve } Y \text{ mallarını satın almış müşteri sayısı} / Y \text{ malını satın almış müşteri sayısı}$

Destek veride bu bağıntının ne kadar sık olduğunu, *güven* de Y malını almış bir kişinin hangi olasılıkla X malını alacağını söyler. Bağıntının önemli olması için her iki değer de olabildiğince büyük olması gerekir [51].

Birliktelik kuralları, eş zamanlı olarak gerçekleşen ilişkilerin tanımlanmasında kullanılmaktadır. Birliktelik kurallarını aşağıdaki gibi örneklerle ifade etmek mümkündür [40]:

- Müşteriler bira satın aldıklarında %75 olasılıkla kuruyemiş de satın alırlar.
- Düşük yağlı peynir ve yağsız süt alan müşteriler %85 olasılıkla yağsız yoğurt da satın alırlar.

Ardışık zamanlı örüntüler ise birbirleri ile ilişkisi olan ancak birbirini izleyen dönemlerde gerçekleşen ilişkilerin tanımlanmasında kullanılmaktadır. Ardışık zamanlı örüntüleri aşağıdaki gibi örneklerle ifade etmek mümkündür [40]:

- Eşofman satın alan bir müşteri, ilk üç ay içerisinde %55 olasılıkla spor ayakkabı alacaktır.
- X ameliyatı yapıldığında, 20 gün içinde % 45 olasılıkla Y enfeksiyonu oluşacaktır.

5. UYGULAMA

5.1. Veri Madenciliği Süreçleri

5.1.1. İşi Anlamak

Bu ilk aşama proje amaçlarını ve şartlarını işe yönelik bakış açısıyla anlamaya ve daha sonra bu bilgiyi veri madenciliği problem tanımına çevirmeye odaklanır.

Bu tez kapsamında, GYKA veri seti kullanılarak tüketim harcamalarının İstatistiki Bölge Birimleri Sınıflaması Düzey-1'e göre oransal dağılımının tahmin edilmesi ve bulunan bu oransal dağılımın 22 Kasım 2012 tarihinde "Hanehalkı Tüketim Harcaması (Bölgesel), 2011" konulu Ek.1'de verilen haber bülteni ile açıklanan resmi sonuçlar ile karşılaştırılması amaçlanmıştır.

5.1.2. Veriyi Anlamak

Veriyi anlama aşaması ilk veriyi toplama ile başlar ve veriye aşina olma, veri kalitesi ile ilgili problemlerin belirlenmesi, verideki ilk anlamları farketme ya da gizlenmiş bilgiye hipotez oluşturmak için ilginç altkümeleri belirleme faaliyetleri ile devam eder.

TÜİK tarafından 2011 yılında yapılmış olan HBA Mikro Veri CD'si ve 2012 yılında yapılmış olan GYKA Mikro Veri CD'si kullanılmıştır.

HBA Mikro Veri CD'sinin içeriğinde;

- Hane Veri Seti,
- Fert Veri Seti,
- Tüketim Harcaması Veri Seti (HBS-Kod5'li),
- HBS-Kod Madde Tanımları (5'li)

bulunmaktadır.

GYKA Mikro Veri CD'sinin içeriğinde;

- Hane Veri Seti,
- Fert Kayıt Veri Seti,
- Fert Veri Seti,
- Modül Fert Veri Seti,
- Modül Hane Veri Seti

bulunmaktadır.

5.1.3. Veri Hazırlığı

Veri hazırlığı aşaması ham veriden, modelleme araçlarına gönderilecek en son veri setini oluşturma faaliyetlerinin hepsini kapsar.

Bu amaçla öncelikle HBA ve GYKA'da bulunan ortak değişkenlerin belirlenmesi için TÜİK'in 2011 yılında yapmış olduğu HBA kapsamında 9 918 hanenin verileri ve yine TÜİK'in 2012 yılında yapmış olduğu GYKA kapsamında 17 562 hanenin verileri incelenmiştir. Her iki çalışmadaki ortak değişkenler aşağıdaki Çizelge 5.1'de, GYKA ve HBA ortak değişkenlerin yalın ve faktörlü frekans değerleri Ek.3'de, standartlaştırılmış değişkenlerin yalın ve faktörlü frekans değerleri Ek.4'de verilmiştir.

Çizelge 5.1. GYKA ve HBA'da Yer Alan ve Türetilen Ortak Değişkenler

Değişken	Değişken Tipi
Hanehalkı tipi	Mevcut Değişken
Oturulan konutun tipi	Mevcut Değişken
Oturulan konuttaki mülkiyet şekli	Mevcut Değişken
*Oturulan konuta sahip olunan veya oturmaya başlanan yıl	Mevcut Değişken
Son 12 ay içinde EV KİRASININ, FAİZLİ BORÇ GERİ ÖDEMESİNİN veya KONUT KREDİSİNİN planlandığı gibi ödenememe durumu	Mevcut Değişken
Konutta oturma süresi	Mevcut Değişken
Oturulan konuttaki (mutfak, banyo ve tuvalet hariç) oda sayısı	Mevcut Değişken
Oturulan konuttaki faydalanılan alanın metrekare olarak büyüklüğü (m ²)	Mevcut Değişken
Oturulan konutun ısıtma sistemi	Mevcut Değişken
*Oturulan konutta ısıtma amacıyla en fazla kullanılan yakıt türü	Mevcut Değişken
Oturulan konutta BANYO veya DUŞ'un mevcut olma durumu	Mevcut Değişken
Oturulan konutta TUVALET'in (ev içinde) mevcut olma durumu	Mevcut Değişken
Oturulan konutta BAĞIMSIZ MUTFAĞIN mevcut olma durumu	Mevcut Değişken
Oturulan konutta BORULU SU SİSTEMİ'nin mevcut olma durumu	Mevcut Değişken
Oturulan konutta SICAK SU SİSTEMİ'nin mevcut olma durumu	Mevcut Değişken
Hanehalkının TELEFON HATTI'na (sabit) sahip olma durumu	Mevcut Değişken
Hanehalkının CEP TELEFONU'na sahip olma durumu	Mevcut Değişken
Hanehalkının TELEVİZYON'a sahip olma durumu	Mevcut Değişken
Hanehalkının BİLGİSAYAR'a sahip olma durumu	Mevcut Değişken
Hanehalkının İNTERNET'e sahip olma durumu	Mevcut Değişken
Hanehalkının BUZDOLABI'na sahip olma durumu	Mevcut Değişken

Değişken	Değişken Tipi
Hanehalkının BULAŞIK MAKİNASI'na sahip olma durumu	Mevcut Değişken
Hanehalkının OTOMATİK ÇAMAŞIR MAKİNASI'na sahip olma durumu	Mevcut Değişken
Hanehalkının KLİMA'ya sahip olma durumu	Mevcut Değişken
Hanehalkının OTOMOBİL'e (iş amaçlılar hariç) sahip olma durumu	Mevcut Değişken
Hanehalkının yaşadığı yer	Mevcut Değişken
Hane büyüklüğü	Türetilmiş Değişken
Hane büyüklüğü (OECD)	Türetilmiş Değişken
Beş yaşından küçük çocuk sayısı	Türetilmiş Değişken
Dört yaş ile 18 yaş arasındaki çocuk sayısı	Türetilmiş Değişken
17 yaş ile 40 yaş arasındaki erkek sayısı	Türetilmiş Değişken
39 yaş ile 66 yaş arasındaki erkek sayısı	Türetilmiş Değişken
65 yaşından büyük erkek sayısı	Türetilmiş Değişken
17 yaş ile 40 yaş arasındaki kadın sayısı	Türetilmiş Değişken
39 yaş ile 66 yaş arasındaki kadın sayısı	Türetilmiş Değişken
65 yaşından büyük kadın sayısı	Türetilmiş Değişken
İşsiz fert sayısı	Türetilmiş Değişken
İstihdamda yer alan fert sayısı	Türetilmiş Değişken
İşgücüne dahil olmayan fert sayısı	Türetilmiş Değişken
Esas işindeki durumu ücretli, maaşlı olan fert sayısı	Türetilmiş Değişken
Esas işindeki durumu yevmiyeli olan fert sayısı	Türetilmiş Değişken
Esas işindeki durumu işveren olan fert sayısı	Türetilmiş Değişken
Esas işindeki durumu kendi hesabına olan fert sayısı	Türetilmiş Değişken
Esas işindeki durumu ücretsiz aile işçisi olan fert sayısı	Türetilmiş Değişken
Okur-yazar olmayan fert sayısı	Türetilmiş Değişken
Okur-yazar olup, bir okul bitirmeyen fert sayısı	Türetilmiş Değişken
İlkokul mezunu fert sayısı	Türetilmiş Değişken
Ortaokul, mesleki ortaokul ve ilköğretim mezunu fert sayısı	Türetilmiş Değişken
Genel lise mezunu fert sayısı	Türetilmiş Değişken
Mesleki veya teknik lise mezunu fert sayısı	Türetilmiş Değişken
Yüksekokul, fakülte ve üzeri mezun fert sayısı	Türetilmiş Değişken
Oturulan konut için ödenen aylık kira tutarı (TL)	Mevcut Değişken
*Oturulan konutun yıllık izafi kirası (TL)	Mevcut Değişken
Gayrimenkul (kira) geliri (TL)	Mevcut Değişken
Toplam hanehalkı kullanılabilir geliri (TL)	Mevcut Değişken
*Aylık ortalama konut masrafı (TL)	Mevcut Değişken

* ile işaretli değişkenler analizlerde kullanılmamıştır.

Veri setlerindeki değişkenlerin dağılımının benzerliğini ölçebilmek için literatürde iki yöntem öne çıkmaktadır. Birinci yöntem, iki veri setindeki değişkenlerin ağırlıklı frekans dağılımlarını karşılaştırmaktır. İkinci yöntem ise, Hellinger Uzaklığı gibi bir ölçüm yöntemi kullanmaktır. İki değişkenin dağılımındaki benzerliği ölçebilmek için

tek bir rakam sunduğu için Hellinger Uzaklığı uygun bir yöntemdir. Hellinger Uzaklığı'nın genellikle %5'ten büyük çıkması dağılımlardaki benzerlik konusunda tereddütler oluşturmaktadır. Bu durumun tersi olarak Hellinger Uzaklığı'nın %5'ten küçük çıkması dağılımlar arasındaki benzerliğe işaret etmektedir. Hellinger Uzaklığı aşağıdaki formül ile hesaplanmaktadır.

$$HD (V, V') = \sqrt{\frac{1}{2} \sum_{i=1}^K \left(\sqrt{\frac{n_{oi}}{N_o}} - \sqrt{\frac{n_{pi}}{N_p}} \right)^2}$$

Formülde yer alan;

V: verici veri setini,

V': alıcı veri setini,

K: değişkenin düzey sayısını,

n_{oi} : verici veri setindeki i. hücredeki frekans değerini,

n_{pi} : alıcı veri setindeki i. hücredeki frekans değerini ve

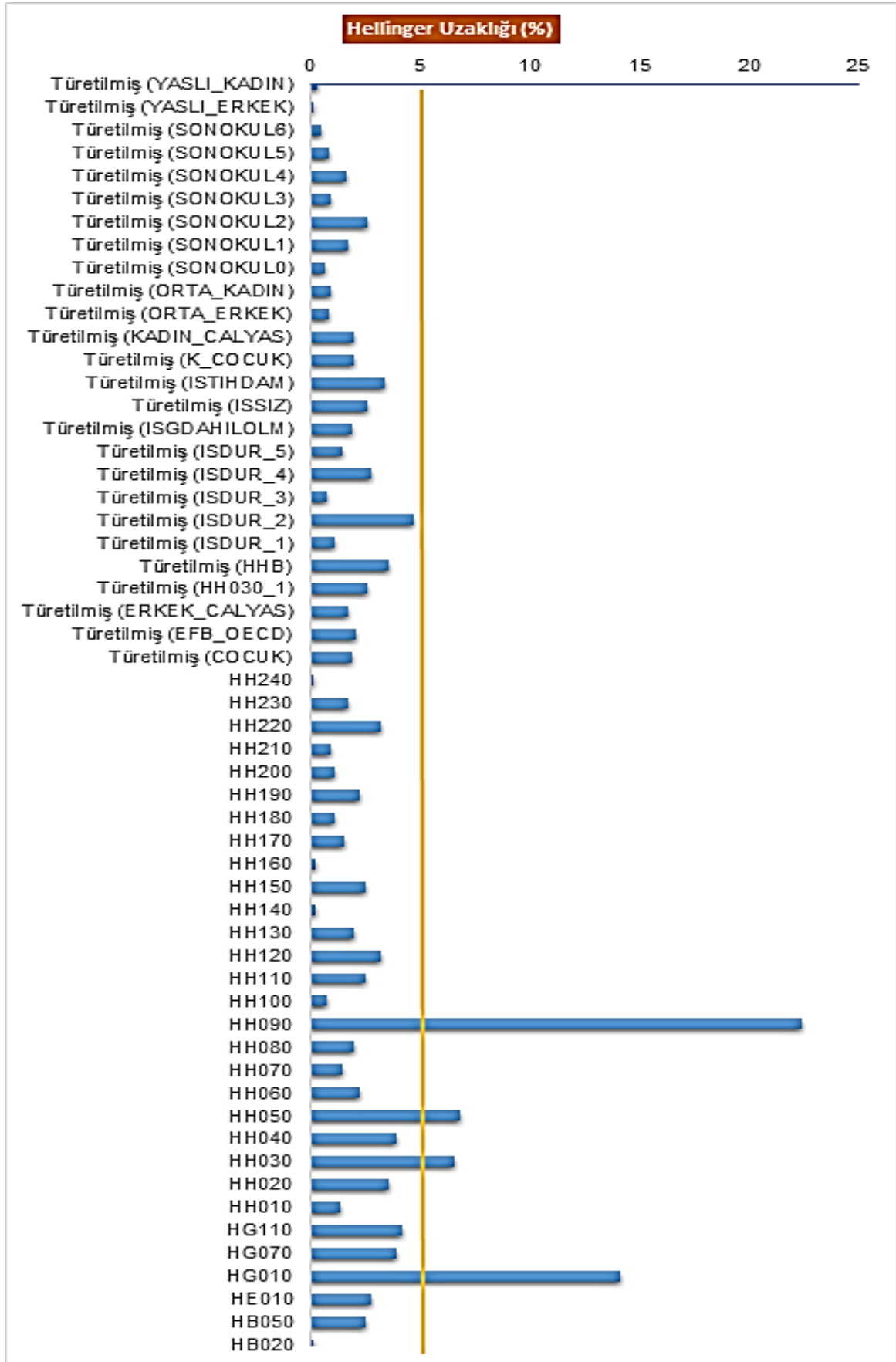
N: örneklem büyüklüğünü

göstermektedir.

Ek.4'te yer alan tablo, HBA ve GYKA veri setlerinde bulunan ortak değişkenlerin ağırlıklı frekans değerlerini ve Hellinger Uzaklığı değerlerini vermektedir.

HH090 (Oturulan konutta ısınma amacıyla en fazla kullanılan yakıt türü), HH050 (Aylık ortalama konut masrafı), HH030 (Oturulan konuta sahip olunan veya oturmaya başlanan yıl) ve HG010 (Oturulan konutun yıllık izafi kirası) değişkenlerinin Hellinger Uzaklığı değeri %5'ten büyük olduğu için bu değişkenler analiz dışı bırakılmıştır.

Şekil 5.1'de Hellinger Uzaklığı değerleri bir arada verilmiştir.



Şekil 5.1. Hellinger Uzaklığı

5.1.4. Modelleme

Modellemede IBM SPSS Modeler 16.0 programı kullanılmış olup parametrik olmayan yöntem olarak KNN algoritması, parametrik yöntem olarak Regresyon algoritması ve karma yöntem olarak ise önce Regresyon algoritması daha sonra KNN algoritması kullanılmıştır.

Hanehalkları tarafından yapılan tüketim harcamaları, amaç (gıda, giyim, konut, ulaşım vb.) ve madde çeşidi (dayanıklı mal, yarı dayanıklı veya dayanıksız mal veya hizmet) esas alınarak uluslararası kıyaslamaya imkan sağlayan Amaca Göre Bireysel Tüketimin Sınıflandırılmasına (COICOP) göre gruplandırılmaktadır. Sınıflama, 12 ana gruptan oluşmaktadır.

1. Gıda ve alkolsüz içecekler,
2. Alkollü içecekler, sigara ve tütün,
3. Giyim ve ayakkabı,
4. Konut, su, elektrik, gaz ve diğer yakıtlar,
5. Mobilya, ev aletleri ve bakım hizmetleri,
6. Sağlık,
7. Ulaştırma,
8. Haberleşme,
9. Eğlence ve kültür,
10. Eğitim hizmetleri,
11. Lokanta, yemek hizmetleri ve oteller,
12. Çeşitli mal ve hizmetler.

Uygulama kapsamında model kurulumu ve analizler 12 ana harcama grubuna göre ayrı ayrı yapılacaktır.

5.1.4.1. Gıda ve Alkolsüz İçecekler

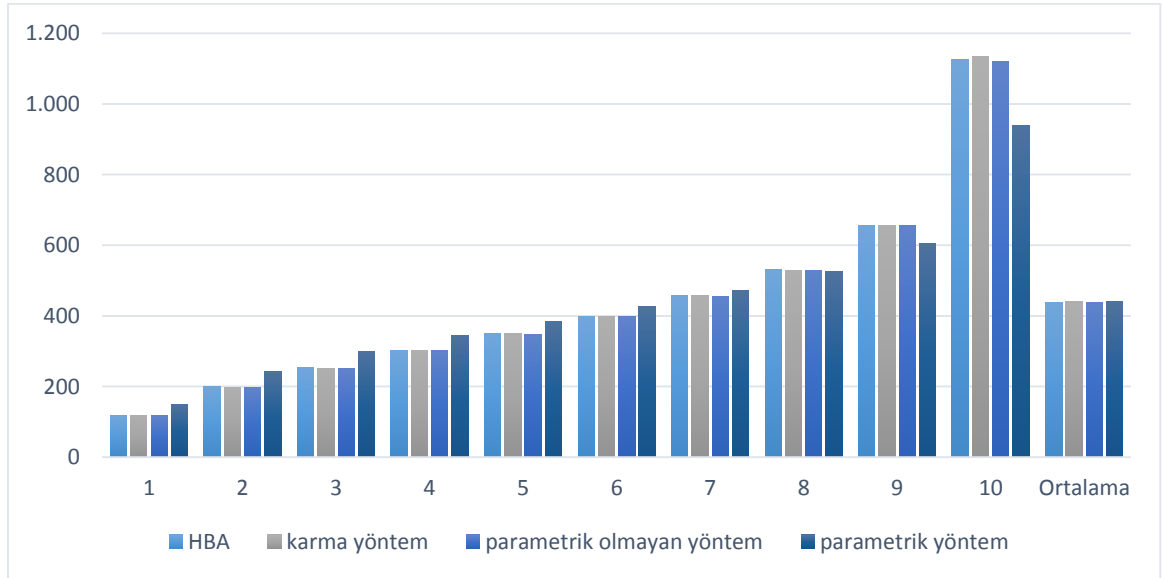
Bu grupta pirinç, ekmek, makarna, et, balık, süt, yoğurt, peynir, diğer süt ürünleri, yumurta, tereyağı, katı ve sıvı yağlar, meyveler, sebzeler, şeker, reçel, bal, çikolata ve şekerlemeler, dondurma, tuz, baharatlar, kahve, çay ve kakao, maden suları, meyve ve sebze suları ve diğer alkolsüz içeceklere yapılan harcamalar bulunmaktadır.

HBA hanelerinin gıda ve alkolsüz içecekler grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak kurulan modeller aracılığıyla gıda ve alkolsüz içecekler harcamaları tekrar tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin gıda ve alkolsüz içecekler harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.2'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.2. % 10'luk Harcama Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (% 50 örnekleme, TL/ay)

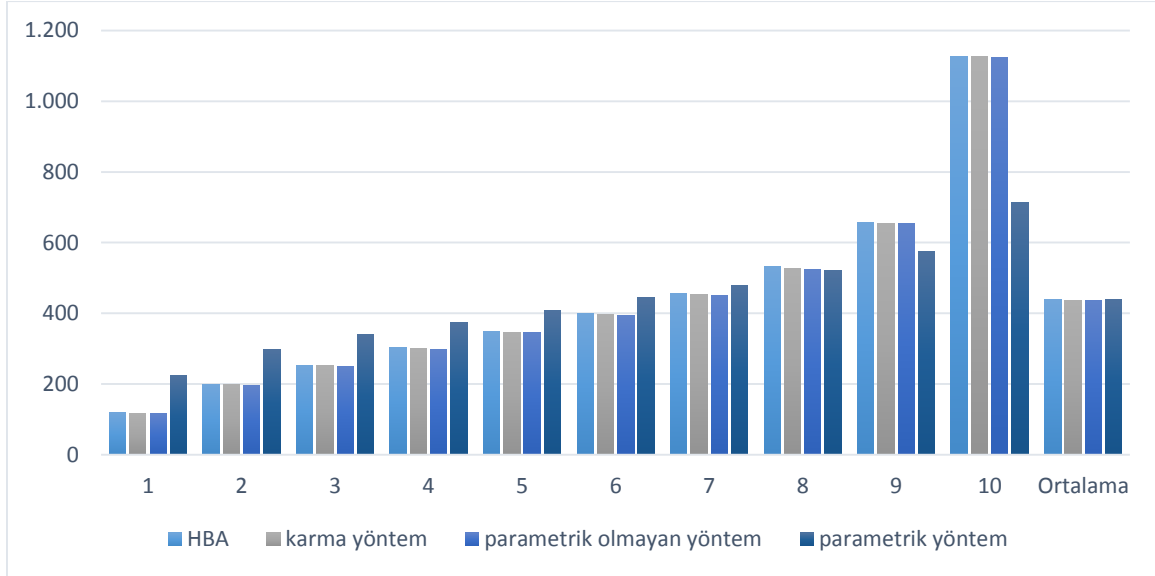
Bu durumu, Çizelge 5.2'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.2. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	24,90	0,09	116,83	13648,03
<i>parametrik olmayan yöntem</i>	24,06	0,07	122,26	14947,88
<i>parametrik yöntem</i>	84,78	0,27	191,82	36795,58

İkinci olarak, kurulan modeller aracılığıyla tüm HBA hanelerinin gıda ve alkolsüz içecekler harcamaları yeniden tahmin edilmiştir.

Şekil 5.3'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



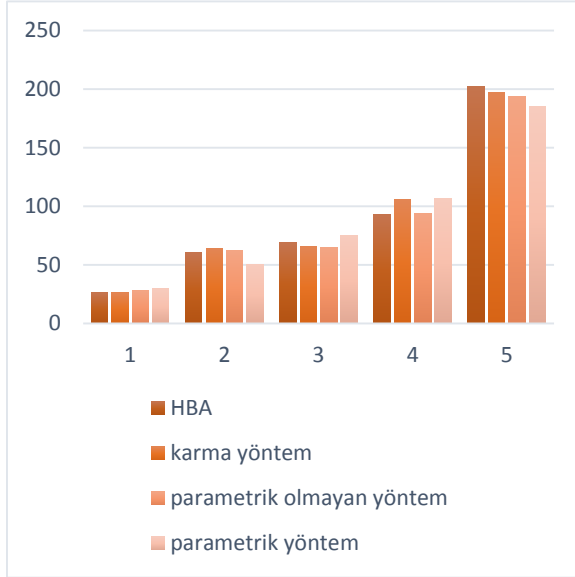
Şekil 5.3. % 10'luk Harcama Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay)

Bu durumu, Çizelge 5.3'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

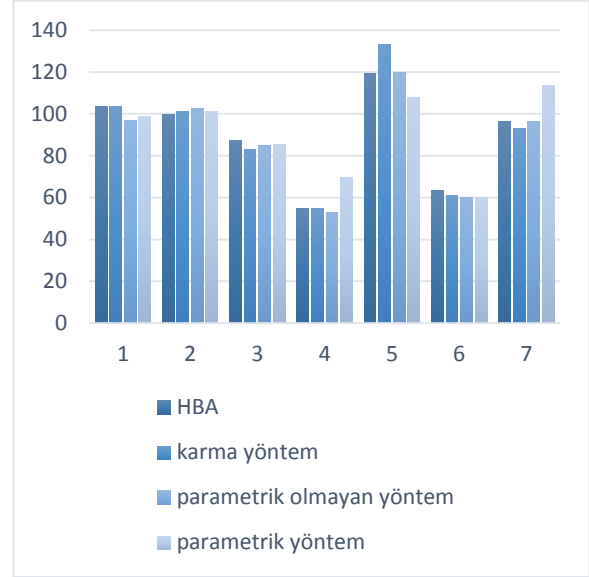
Çizelge 5.3. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	49,58	0,16	169,79	28827,52
<i>parametrik olmayan yöntem</i>	49,06	0,15	173,36	30054,51
<i>parametrik yöntem</i>	171,31	0,56	269,55	72657,85

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama gıda ve alkolsüz içecekler harcamaları sırasıyla kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.4 ve Şekil 5.5'deki grafiklerde verilmiştir.



Şekil 5.4. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay)



Şekil 5.5. Hanehalkı Tipine Göre Ortalama Gıda ve Alkolsüz İçecekler Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.4 ve Şekil 5.5'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının iyi olduğu sonucuna varılabilir. Çizelge 5.4'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.4. Hanehalkı Kullanılabilir Geliri ve Gıda ve Alkolsüz İçecekler Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,410	0,421	0,341
Karma yöntem	0,437	0,444	0,358
Parametrik olmayan yöntem	0,225	0,287	0,212
Parametrik yöntem	0,623	0,619	0,481

5.1.4.2. Alkollü İçecekler, Sigara ve Tütün

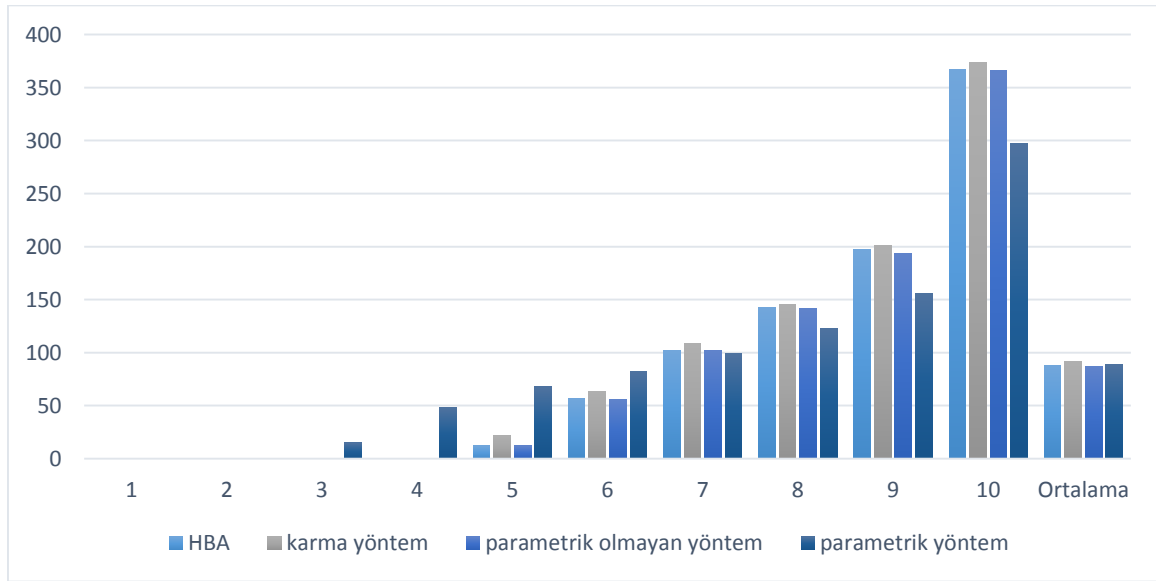
Bu grupta içki, likör, şarap, bira, sigara, puro ve tütüne yapılan harcamalar bulunmaktadır.

HBA hanelerinin alkollü içecekler, sigara ve tütün grubundaki maddelere yapılan harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla alkollü içecekler, sigara ve tütün harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık alkollü içecekler, sigara ve tütün harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.6'daki grafikte genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.6. % 10'luk Harcama Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (% 50 örnekleme, TL/ay)

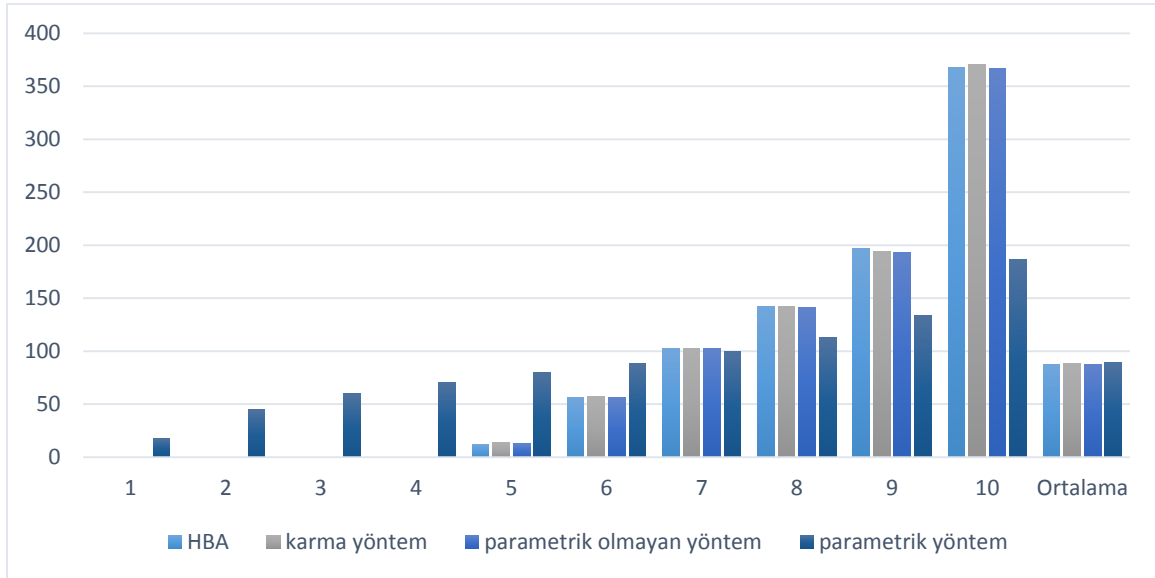
Bu durumu, Çizelge 5.5'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.5. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	7,38	0,07	42,58	1812,69
<i>parametrik olmayan yöntem</i>	10,64	0,11	50,60	2560,59
<i>parametrik yöntem</i>	39,83	0,40	77,87	6063,43

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin alkollü içecekler, sigara ve tütün harcamaları tahmin edilmiştir.

Şekil 5.7'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



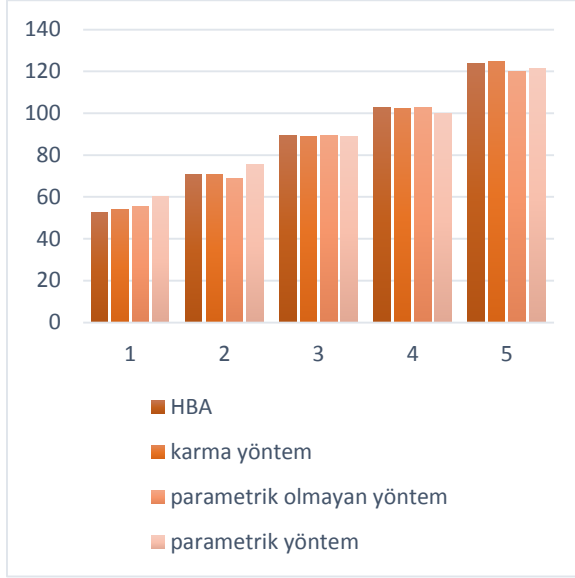
Şekil 5.7. % 10'luk Harcama Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)

Bu durumu, Çizelge 5.6'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

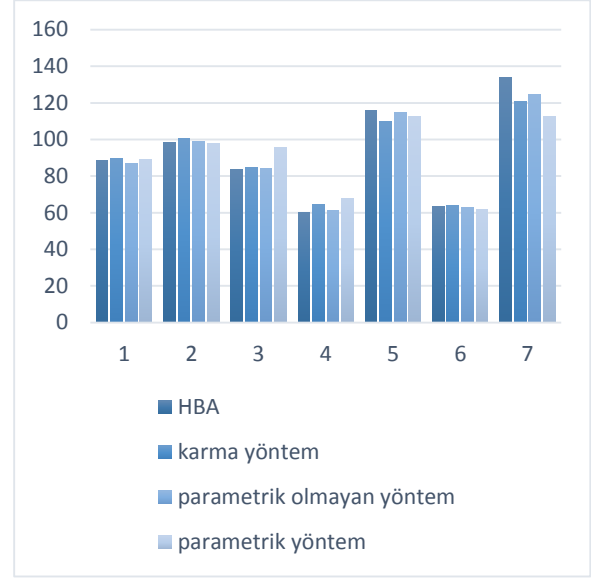
Çizelge 5.6. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	21,39	0,19	73,14	5350,02
<i>parametrik olmayan yöntem</i>	20,77	0,20	71,26	5077,93
<i>parametrik yöntem</i>	81,14	0,77	114,17	13034,12

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama alkollü içecekler, sigara ve tütün harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.8 ve Şekil 5.9'daki grafiklerde verilmiştir.



Şekil 5.8. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)



Şekil 5.9. Hanehalkı Tipine Göre Ortalama Alkollü İçecekler, Sigara ve Tütün Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.8 ve Şekil 5.9'daki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.7'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.7. Hanehalkı Kullanılabilir Geliri ve Alkollü İçecekler, Sigara ve Tütün Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,181	0,179	0,155
Karma yöntem	0,171	0,126	0,099
Parametrik olmayan yöntem	0,102	0,100	0,079
Parametrik yöntem	0,479	0,473	0,354

5.1.4.3. Giyim ve Ayakkabı

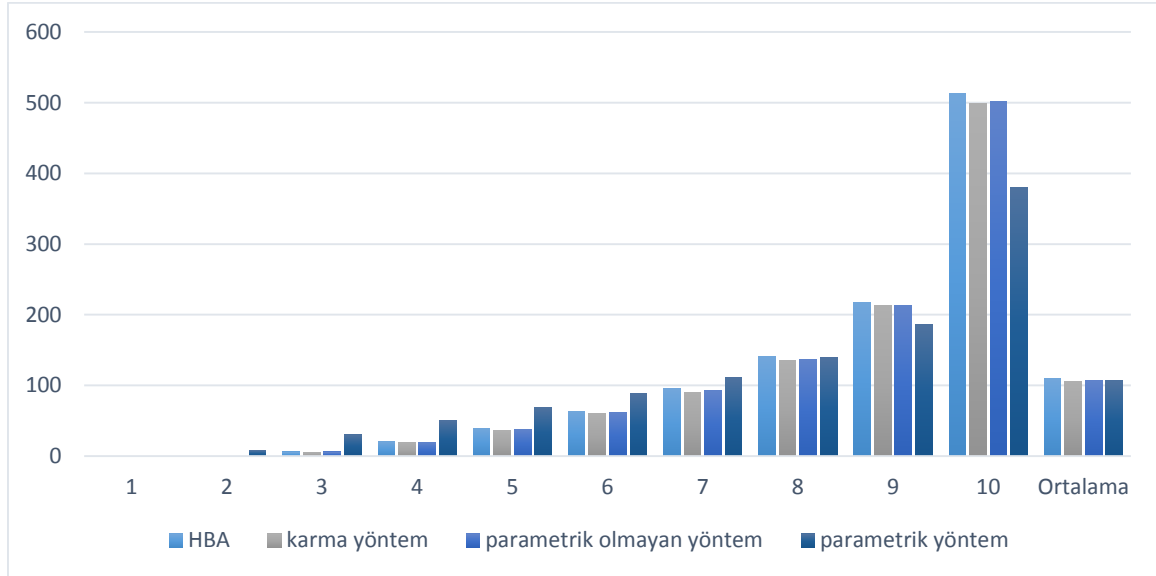
Bu grupta kumaşlar, her türlü erkek, kadın, çocuk ve bebek giyimi, giysi materyalleri ve aksesuarları, kuru temizleme, giysi onarımı, dikişi ve kiralanması, erkek, kadın, çocuk ve bebek ayakkabısı, ayakkabı tamiri ve kiralanması ile ilgili harcamalar bulunmaktadır.

HBA hanelerinin giyim ve ayakkabı grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla giyim ve ayakkabı harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık giyim ve ayakkabı harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.10'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.10. % 10'luk Harcama Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (% 50 örnekleme, TL/ay)

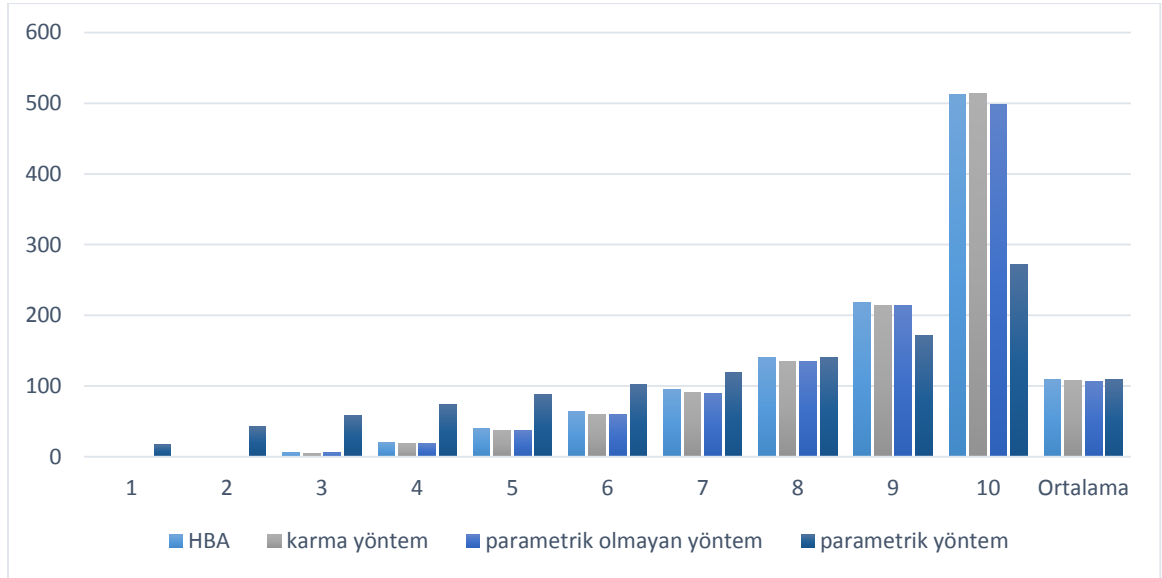
Bu durumu, Çizelge 5.8'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.8. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	24,75	0,52	97,37	9481,72
<i>parametrik olmayan yöntem</i>	12,07	0,22	67,27	4524,83
<i>parametrik yöntem</i>	48,49	1,28	125,76	15816,48

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin giyim ve ayakkabı harcamaları tahmin edilmiştir.

Şekil 5.11'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



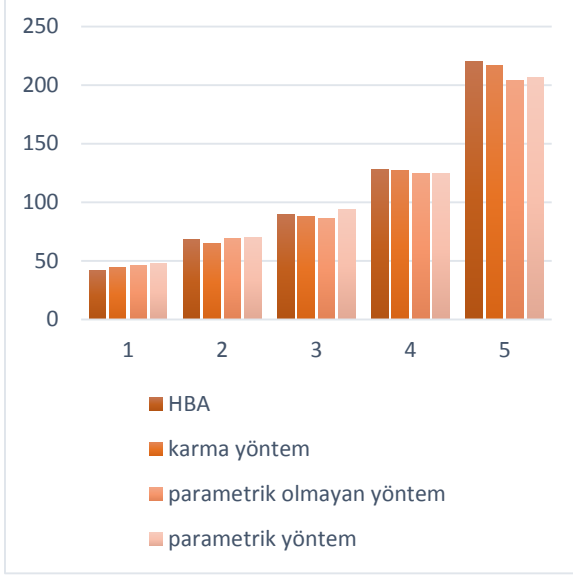
Şekil 5.11. % 10'luk Harcama Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay)

Bu durumu, Çizelge 5.9'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

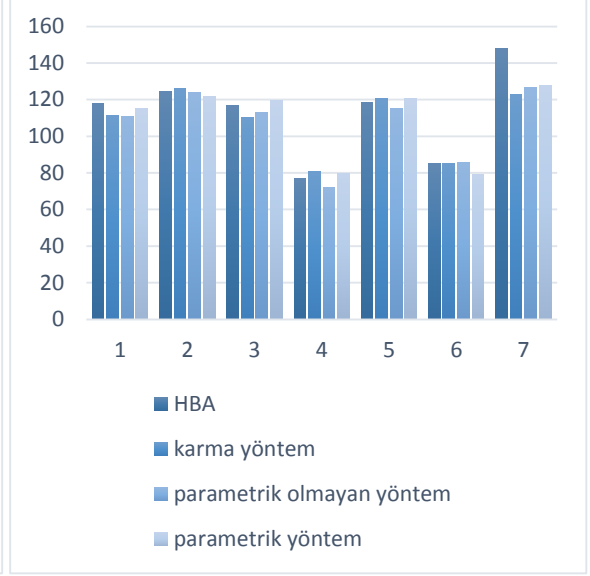
Çizelge 5.9. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	25,50	0,56	102,03	10409,11
<i>parametrik olmayan yöntem</i>	24,75	0,52	97,37	9481,72
<i>parametrik yöntem</i>	93,50	2,61	165,39	27352,62

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama giyim ve ayakkabı harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.12 ve Şekil 5.13'deki grafiklerde verilmiştir.



Şekil 5.12. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay)



Şekil 5.13. Hanehalkı Tipine Göre Ortalama Giyim ve Ayakkabı Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.12 ve Şekil 5.13'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.10'da hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.10. Hanehalkı Kullanılabilir Geliri ve Giyim ve Ayakkabı Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,352	0,338	0,292
Karma yöntem	0,381	0,386	0,291
Parametrik olmayan yöntem	0,222	0,238	0,178
Parametrik yöntem	0,744	0,737	0,591

5.1.4.4. Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar

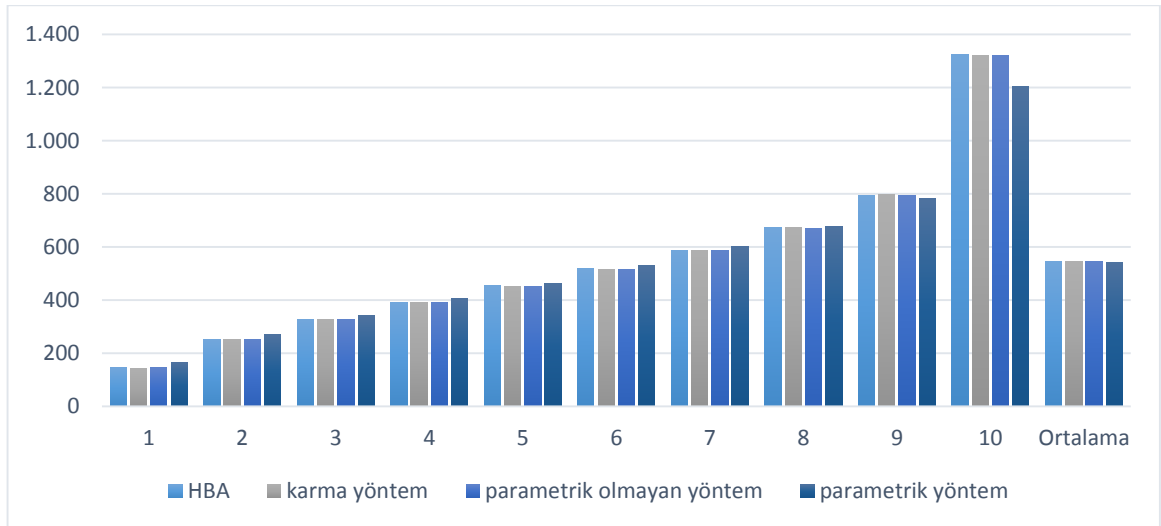
Bu grupta gerçek kira ücreti, izafi kira ücreti (ev sahibi olanlar ve ücretsiz olarak oturanlar için belirlenen), konutun bakım ve onarımına ilişkin hizmetler ve buna ilişkin olarak kullanılan materyaller (badana, boya, yer döşemeleri, pencere, kapı, banyo ve tuvalet teçhizatı, sıhhi tesisat, elektrik tesisatı vb.), çöp toplama ve kanalizasyon hizmetleri, konutla ilgili diğer hizmetler, şebeke suyu, elektrik, gaz, sıvı ve katı yakıtlarla ilgili harcamalar bulunmaktadır.

HBA hanelerinin konut, su, elektrik, gaz ve diğer yakıtlar grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla konut, su, elektrik, gaz ve diğer yakıtlar harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık konut, su, elektrik, gaz ve diğer yakıtlar harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.14'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.14. % 10'luk Harcama Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (% 50 örnekleme, TL/ay)

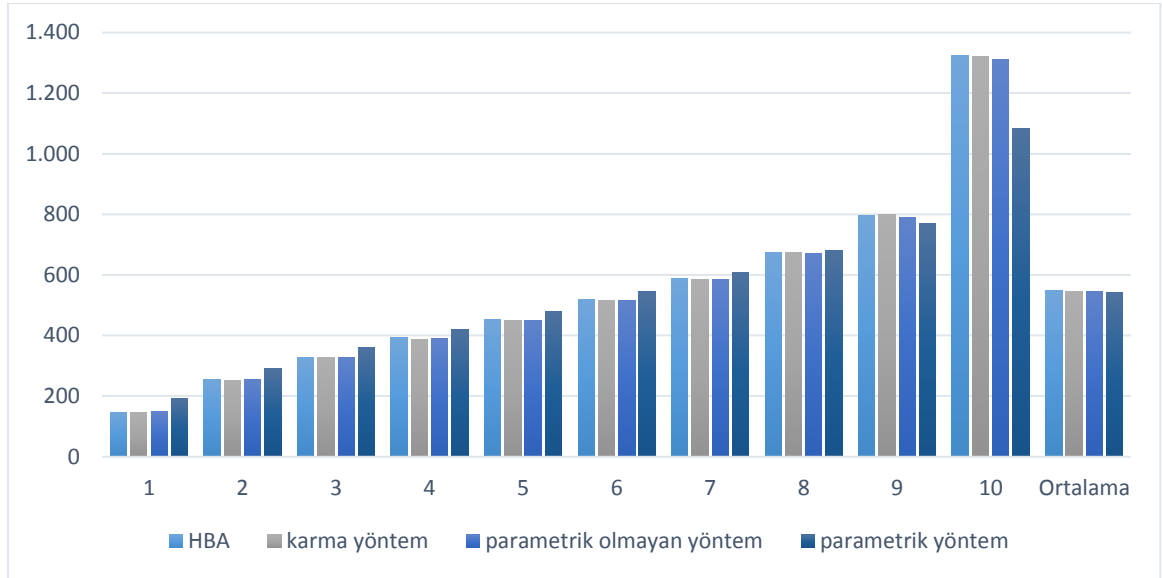
Bu durumu, Çizelge 5.11'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.11. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	22,65	0,05	104,38	10894,70
<i>parametrik olmayan yöntem</i>	21,08	0,04	117,71	13856,65
<i>parametrik yöntem</i>	75,23	0,17	181,04	32776,82

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin konut, su, elektrik, gaz ve diğer yakıtlar harcamaları tahmin edilmiştir.

Şekil 5.15'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



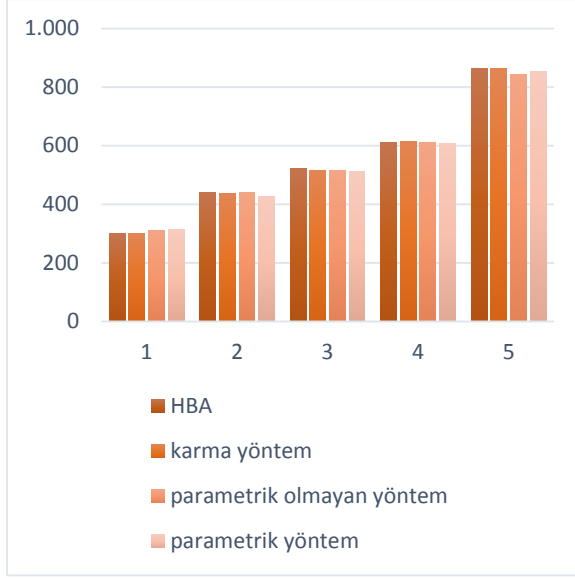
Şekil 5.15. % 10'luk Harcama Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay)

Bu durumu, Çizelge 5.12'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

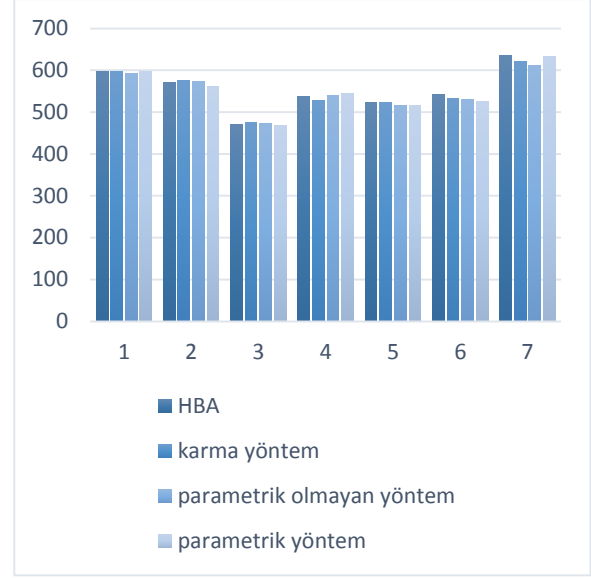
Çizelge 5.12. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	44,35	0,10	149,42	22326,27
<i>parametrik olmayan yöntem</i>	42,83	0,09	168,09	28254,84
<i>parametrik yöntem</i>	151,72	0,34	259,62	67402,40

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama konut, su, elektrik, gaz ve diğer yakıtlar harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.16 ve Şekil 5.17'deki grafiklerde verilmiştir.



Şekil 5.16. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay)



Şekil 5.17. Hanehalkı Tipine Göre Ortalama Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.16 ve Şekil 5.17'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.13'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.13. Hanehalkı Kullanılabilir Geliri ve Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,591	0,590	0,488
Karma yöntem	0,580	0,603	0,462
Parametrik olmayan yöntem	0,421	0,480	0,361
Parametrik yöntem	0,726	0,719	0,573

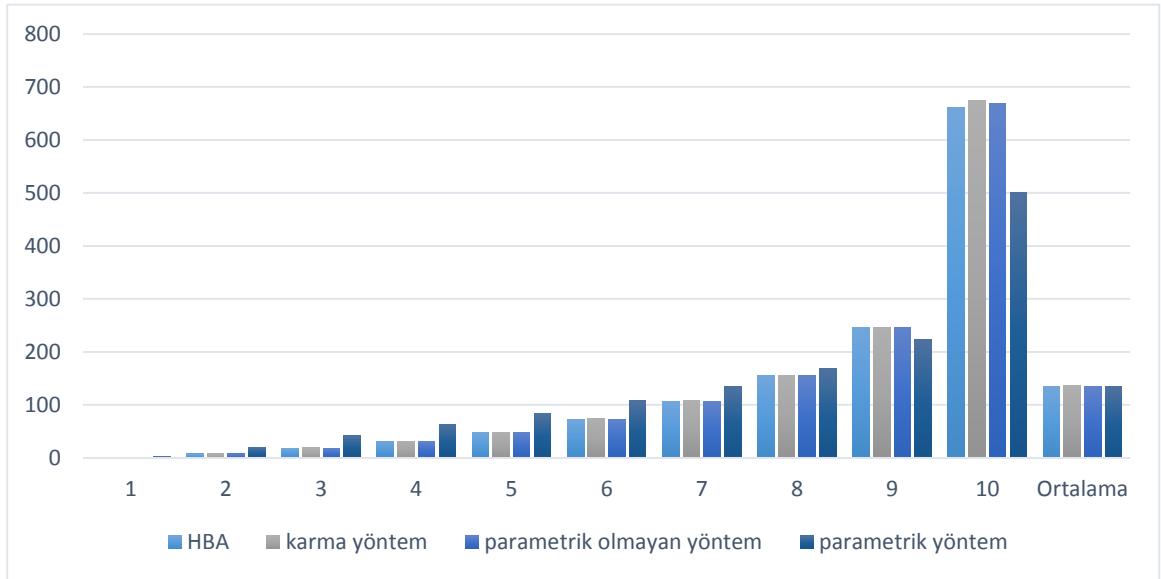
5.1.4.5. Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri

Bu grupta mobilya, halı ve diğer yer döşemeleri, mobilya ve ev döşemelerinin onarımı, ev içi tekstil ürünleri, buzdolabı, dondurucu, çamaşır ve bulaşık makinesi, fırın, ısıtıcı, klima, dikiş makinası, elektrikli küçük ev aletleri, cam ve kristal eşyalar ile çatal, bıçak ve sofrta takımları vb. mutfak eşyaları, ev gereçlerinin onarımı, deterjan vb. ev temizlik ürünleri, gündelikçi, aşçı, vb. konut ve hanehalkıyla ilgili hizmetlere ilişkin harcamaları bulunmaktadır.

HBA hanelerinin mobilya, ev aletleri ve ev bakım hizmetleri grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur. İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla mobilya, ev aletleri ve ev bakım hizmetleri harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık mobilya, ev aletleri ve ev bakım hizmetleri harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.18'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.18. % 10'luk Harcama Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (% 50 örnekleme, TL/ay)

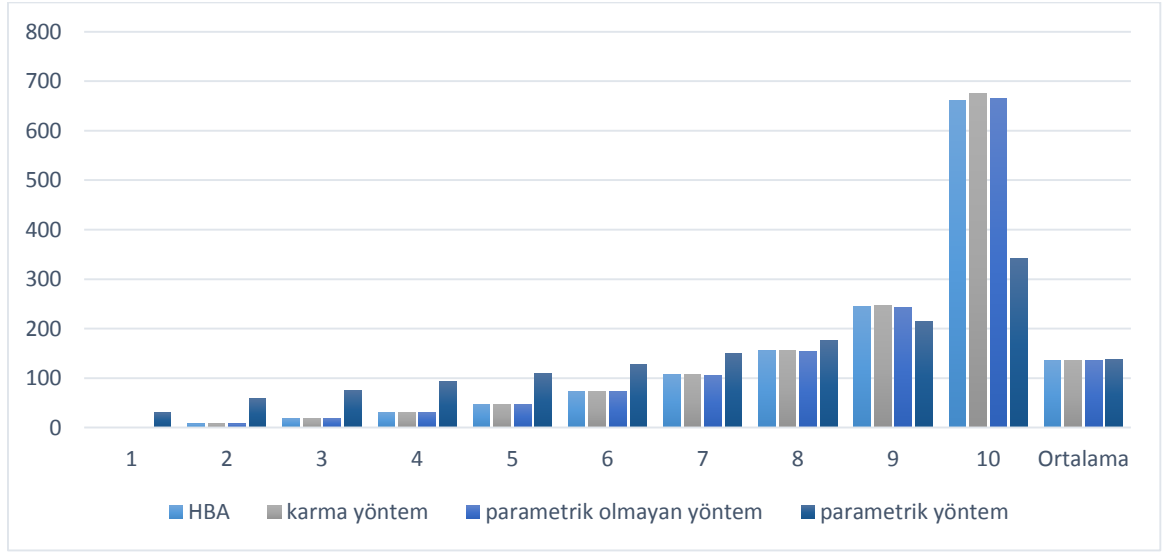
Bu durumu, Çizelge 5.14'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.14. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	15,11	0,46	97,44	9493,60
<i>parametrik olmayan yöntem</i>	15,31	1,15	98,55	9711,63
<i>parametrik yöntem</i>	58,22	2,30	158,26	25045,15

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin mobilya, ev aletleri ve ev bakım hizmetleri harcamaları tahmin edilmiştir.

Şekil 5.19'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



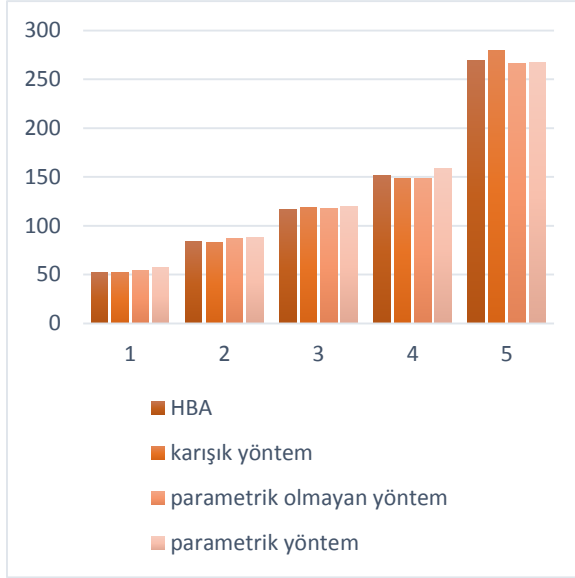
Şekil 5.19. % 10'luk Harcama Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)

Bu durumu, Çizelge 5.15'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

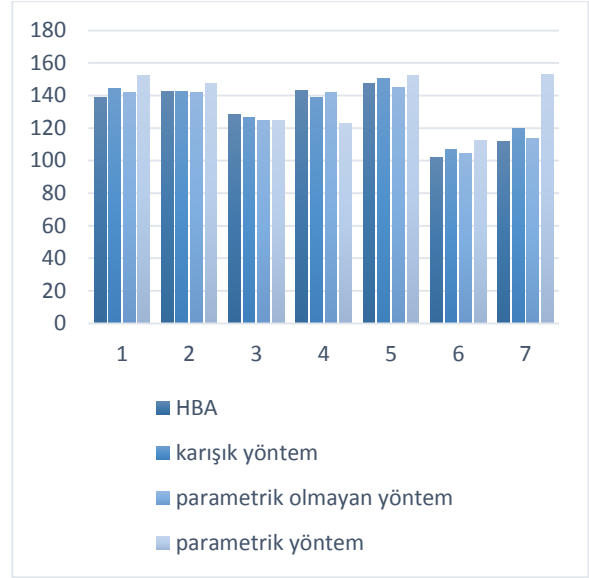
Çizelge 5.15. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	31,97	1,08	145,20	21082,67
<i>parametrik olmayan yöntem</i>	30,71	1,82	141,51	20024,64
<i>parametrik yöntem</i>	116,87	4,90	217,70	47393,65

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama mobilya, ev aletleri ve ev bakım hizmetleri harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.20 ve Şekil 5.21'deki grafiklerde verilmiştir.



Şekil 5.20. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)



Şekil 5.21. Hanehalkı Tipine Göre Ortalama Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.20 ve Şekil 5.21'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.16'da hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.16. Hanehalkı Kullanılabilir Geliri ve Mobilya, Ev Aletleri ve Ev Bakım Hizmetleri Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,340	0,338	0,289
Karma yöntem	0,390	0,436	0,327
Parametrik olmayan yöntem	0,222	0,312	0,230
Parametrik yöntem	0,817	0,844	0,706

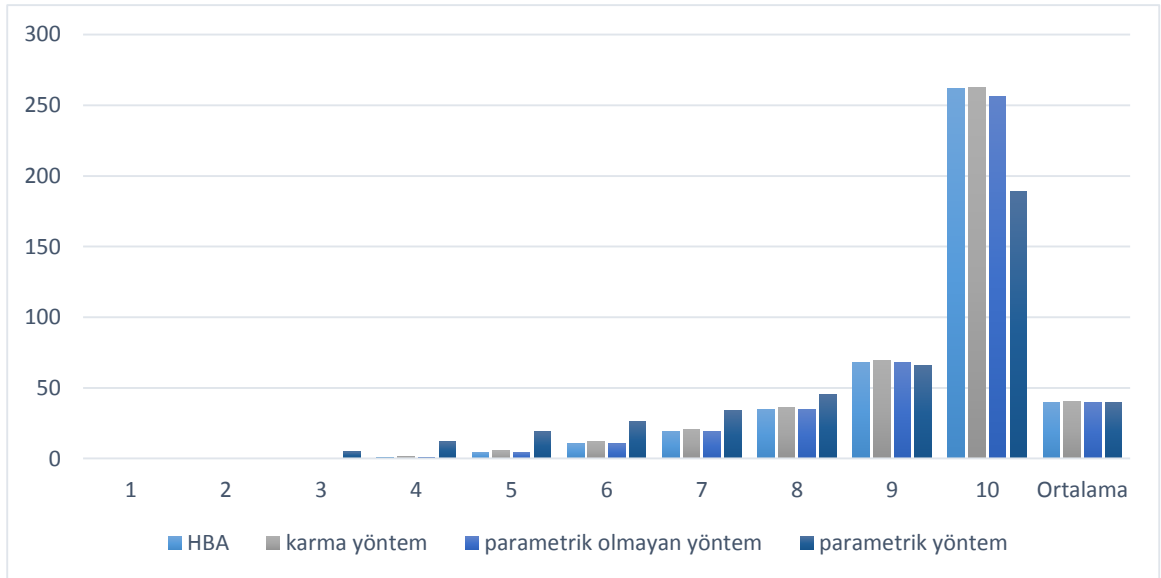
5.1.4.6. Sağlık

Bu grupta ilaçlar, sağlık ürünleri, tedavi edici alet ve malzemeler (gözlük, lens, boyunluk, işitme cihazı, yürüteç vb.), sağlıkla ilgili aletler (tansiyon aleti, şeker ölçme aleti vb.) hastane ve hastane dışı tıbbi hizmetler (doktor muayene ücreti, dişçilik hizmetleri, röntgen, ultrason, tomografi, tahlil vb.), hastane yatak, ameliyat, doğum, fizik tedavi, ambulans vb. ile ilgili harcamalar bulunmaktadır.

HBA hanelerinin sağlık grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur. İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla sağlık harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık sağlık harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.22'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.22. % 10'luk Harcama Gruplarına Göre Ortalama Sağlık Harcamaları (% 50 örnekleme, TL/ay)

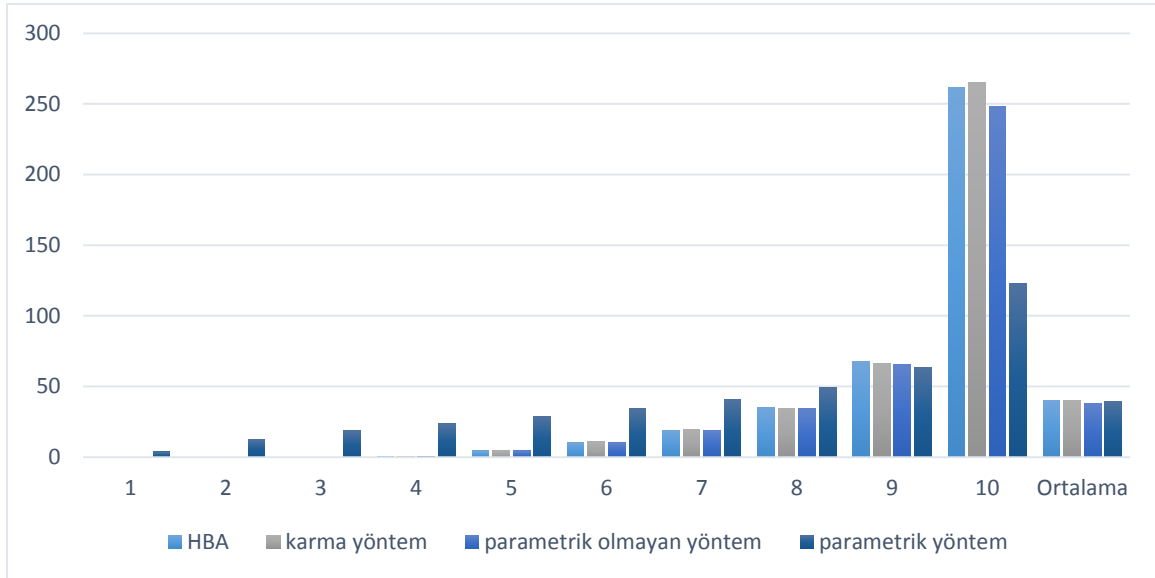
Bu durumu, Çizelge 5.17'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.17. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	5,11	0,26	49,56	2456,61
<i>parametrik olmayan yöntem</i>	5,80	0,28	47,82	2286,47
<i>parametrik yöntem</i>	24,46	1,37	88,17	7773,24

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin sağlık harcamaları tahmin edilmiştir.

Şekil 5.23'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



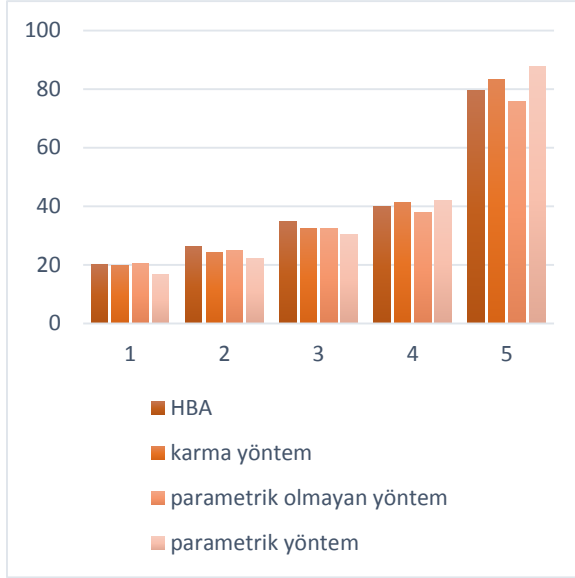
Şekil 5.23. % 10'luk Harcama Gruplarına Göre Ortalama Sağlık Harcamaları (TL/ay)

Bu durumu, Çizelge 5.18'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

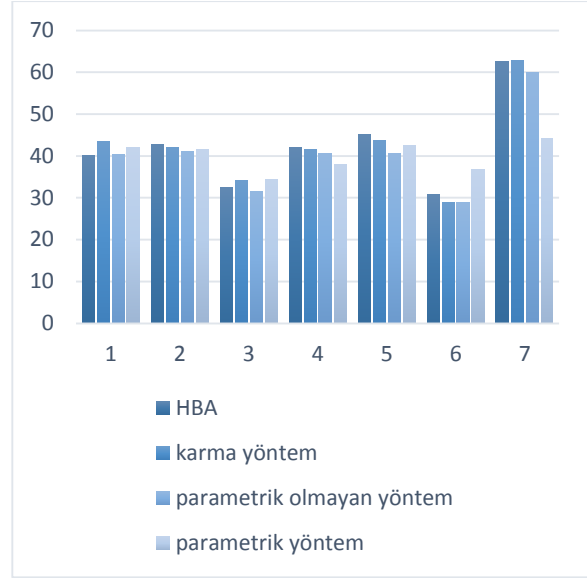
Çizelge 5.18. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	12,95	0,67	71,71	5142,67
<i>parametrik olmayan yöntem</i>	11,13	0,61	62,66	3925,95
<i>parametrik yöntem</i>	48,16	2,83	112,34	12620,73

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama sağlık harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.24 ve Şekil 5.25'deki grafiklerde verilmiştir.



Şekil 5.24. Hanehalkı Kullanılabilir Gelire Göre Sıralanmış % 20'lik Grupların Ortalama Sağlık Harcamaları (TL/ay)



Şekil 5.25. Hanehalkı Tipine Göre Ortalama Sağlık Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.24 ve Şekil 5.25'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.19'da hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.19. Hanehalkı Kullanılabilir Geliri ve Sağlık Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,175	0,172	0,152
Karma yöntem	0,220	0,186	0,142
Parametrik olmayan yöntem	0,162	0,142	0,109
Parametrik yöntem	0,685	0,723	0,579

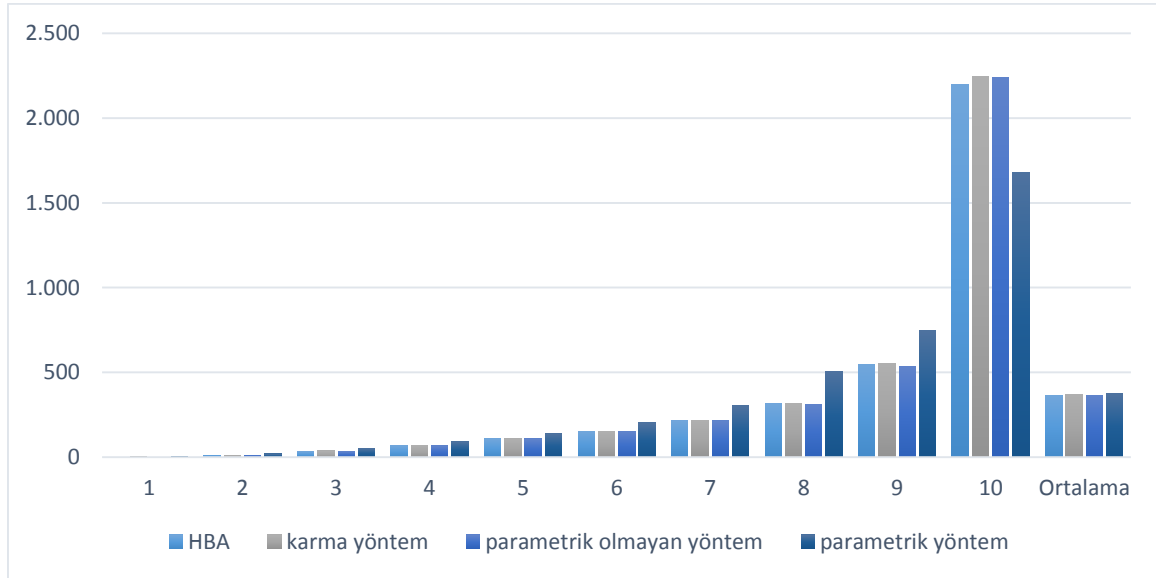
5.1.4.7. Ulaştırma

Bu grupta ulaşım aracı satın alınması (birinci ve ikinci el), motosikletler, bisikletler, yedek parça ve aksesuarlar, yakıt ve yağlar, bakım ve tamirler, yolcu taşımacılığı (demiryolu, karayolu, havayolu, deniz ve su yoluyla), ulaşım araçlarının tamirata, diğer taşıma (nakliye, kargo vb. ulaşım) hizmetleri ile ilgili harcamalar bulunmaktadır.

HBA hanelerinin ulaştırma grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla ulaştırma harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır. Hanelerin aylık ulaştırma harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.26'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.26. % 10'luk Harcama Gruplarına Göre Ortalama Ulaştırma Harcamaları (% 50 örnekleme, TL/ay)

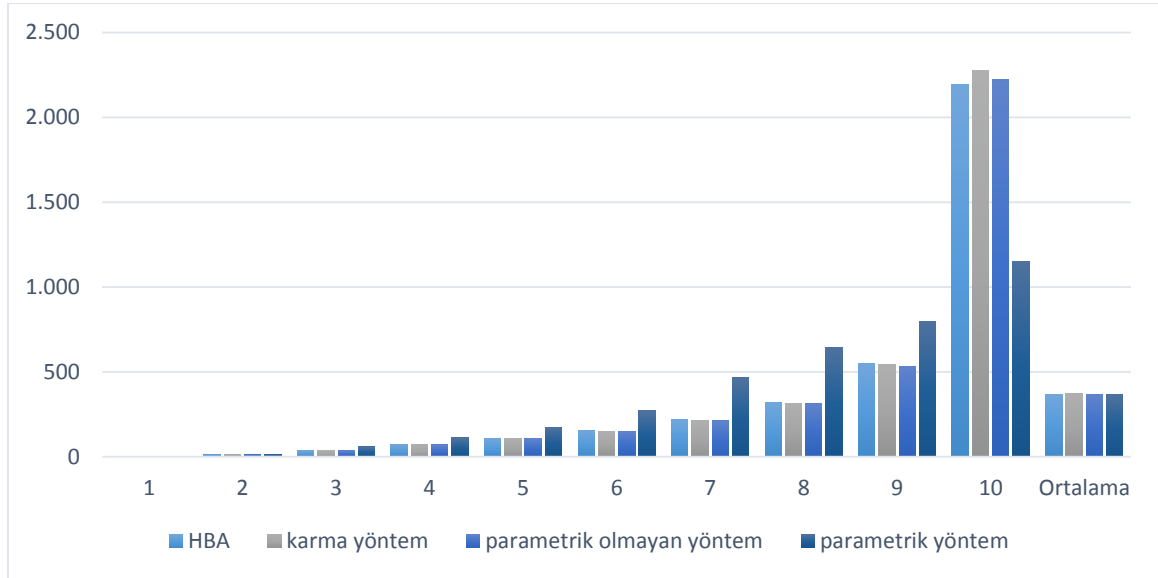
Bu durumu, Çizelge 5.20'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.20. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	41,13	0,32	435,90	190011,99
<i>parametrik olmayan yöntem</i>	39,46	0,23	424,03	179803,14
<i>parametrik yöntem</i>	156,88	1,30	483,42	233693,92

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin ulaştırma harcamaları tahmin edilmiştir.

Şekil 5.27'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



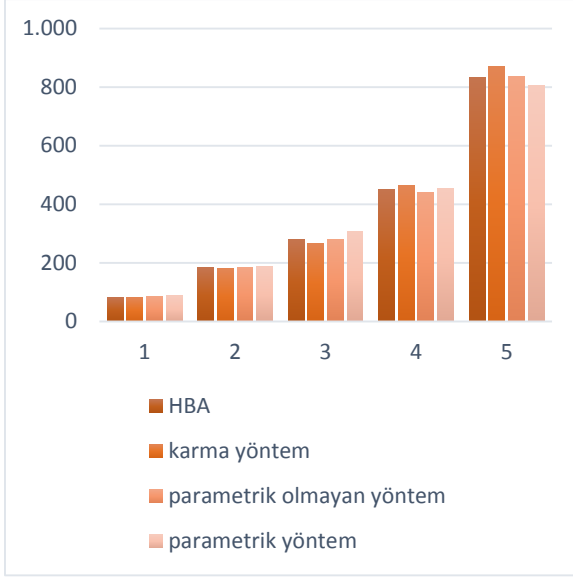
Şekil 5.27. % 10'luk Harcama Gruplarına Göre Ortalama Ulaştırma Harcamaları (TL/ay)

Bu durumu, Çizelge 5.21'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

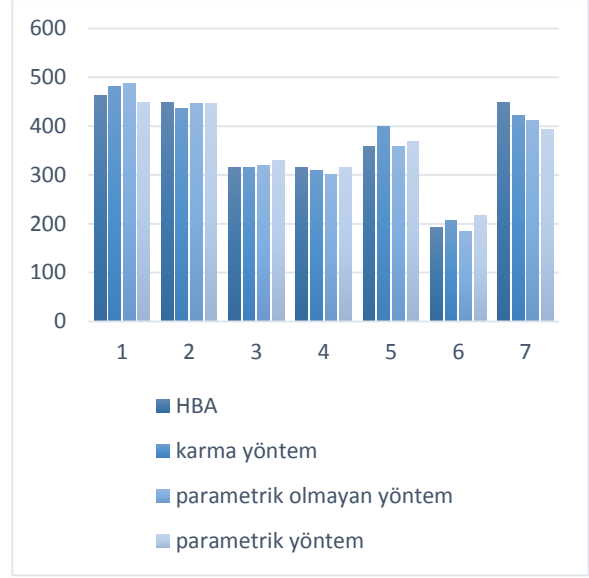
Çizelge 5.21. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	88,66	0,75	613,81	376767,22
<i>parametrik olmayan yöntem</i>	77,25	0,51	512,11	262252,85
<i>parametrik yöntem</i>	328,54	2,75	733,34	537781,35

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama ulaştırma harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.28 ve Şekil 5.29'daki grafiklerde verilmiştir.



Şekil 5.28. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Ulaştırma Harcamaları (TL/ay)



Şekil 5.29. Hanehalkı Tipine Göre Ortalama Ulaştırma Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.28 ve Şekil 5.29'daki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.22'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.22. Hanehalkı Kullanılabilir Geliri ve Ulaştırma Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,451	0,491	0,409
Karma yöntem	0,492	0,555	0,425
Parametrik olmayan yöntem	0,274	0,432	0,324
Parametrik yöntem	0,689	0,696	0,555

5.1.4.8. Haberleşme

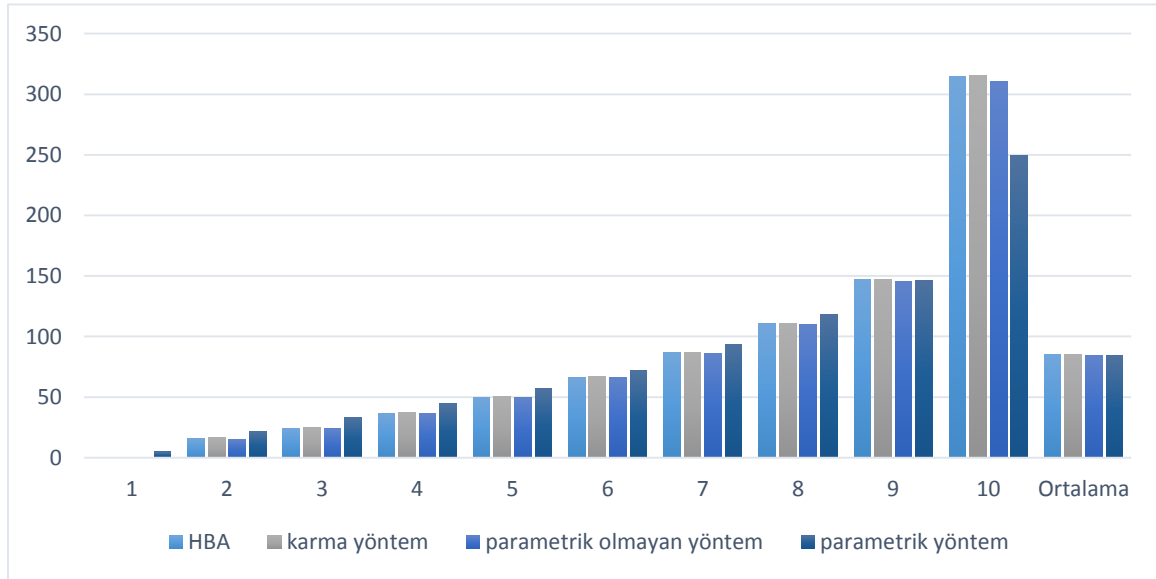
Bu grupta posta hizmetleri, telefon ve telefaks ekipmanları ve onarımı, telefon ve telefaks hizmetleri (telefon görüşme ücreti, telefon kart ücreti, faks çekme ücreti vb.) ile ilgili harcamalar bulunmaktadır.

HBA hanelerinin haberleşme grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla haberleşme harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık haberleşme harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.30'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.30. % 10'luk Harcama Gruplarına Göre Ortalama Haberleşme Harcamaları (% 50 örnekleme, TL/ay)

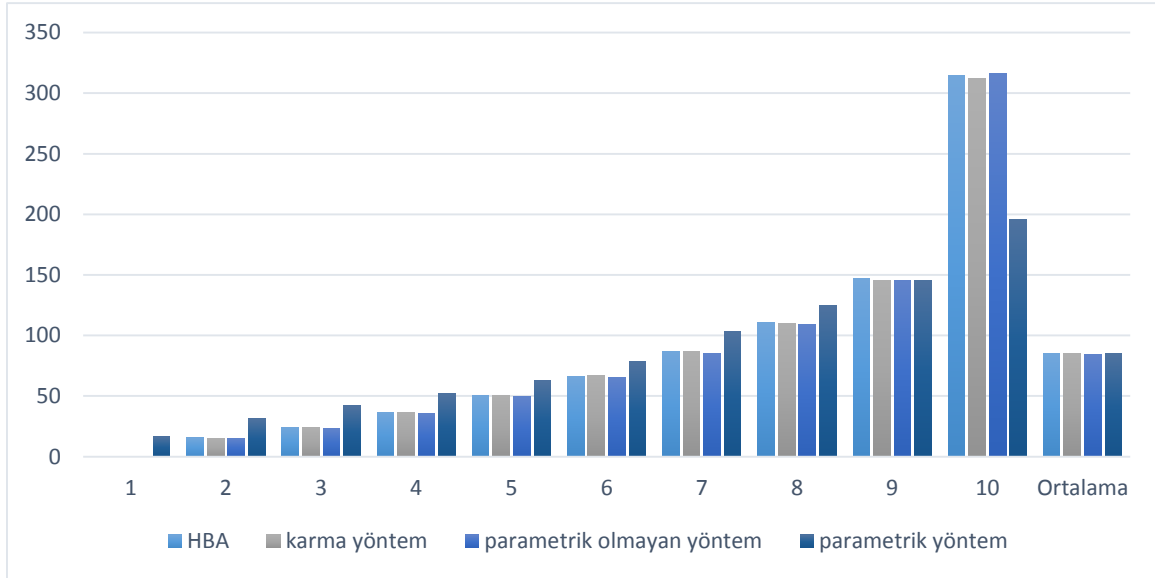
Bu durumu, Çizelge 5.23'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.23. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	6,35	0,09	38,56	1486,87
<i>parametrik olmayan yöntem</i>	6,40	0,09	47,82	2287,13
<i>parametrik yöntem</i>	22,59	0,32	76,61	5869,75

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin haberleşme harcamaları tahmin edilmiştir.

Şekil 5.31'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



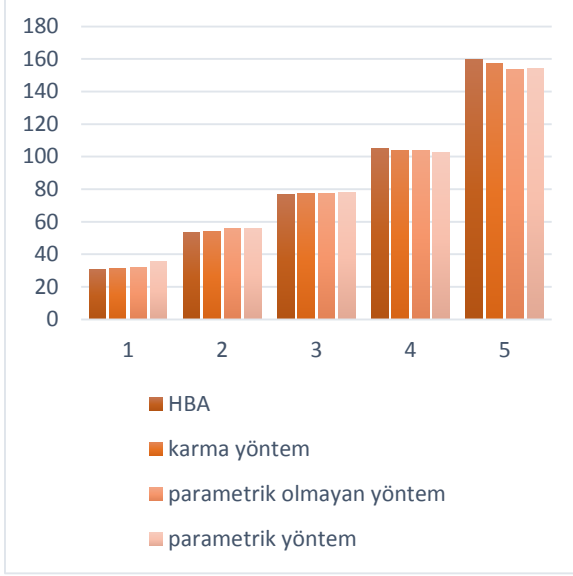
Şekil 5.31. % 10'luk Harcama Gruplarına Göre Ortalama Haberleşme Harcamaları (TL/ay)

Bu durumu, Çizelge 5.24'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

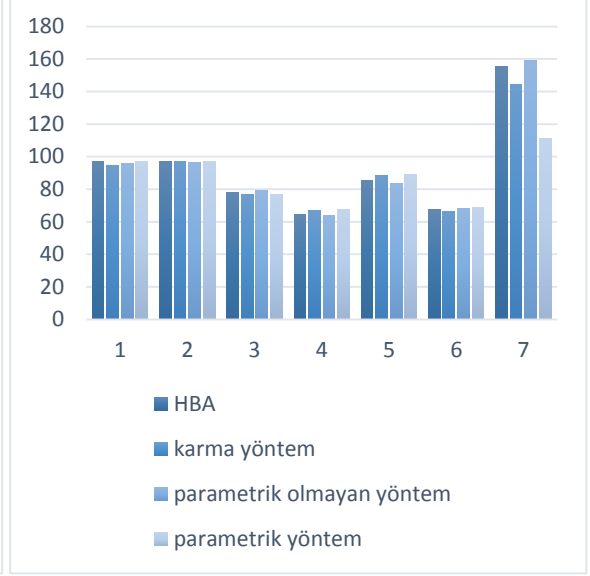
Çizelge 5.24. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	13,45	0,20	57,09	3259,26
<i>parametrik olmayan yöntem</i>	13,08	0,19	66,82	4465,04
<i>parametrik yöntem</i>	45,19	0,68	105,49	11127,50

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama haberleşme harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.32 ve Şekil 5.33'deki grafiklerde verilmiştir.



Şekil 5.32. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Haberleşme Harcamaları (TL/ay)



Şekil 5.33. Hanehalkı Tipine Göre Ortalama Haberleşme Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.32 ve Şekil 5.33'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.25'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.25. Hanehalkı Kullanılabilir Geliri ve Haberleşme Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,332	0,383	0,335
Karma yöntem	0,392	0,535	0,406
Parametrik olmayan yöntem	0,284	0,466	0,353
Parametrik yöntem	0,766	0,769	0,623

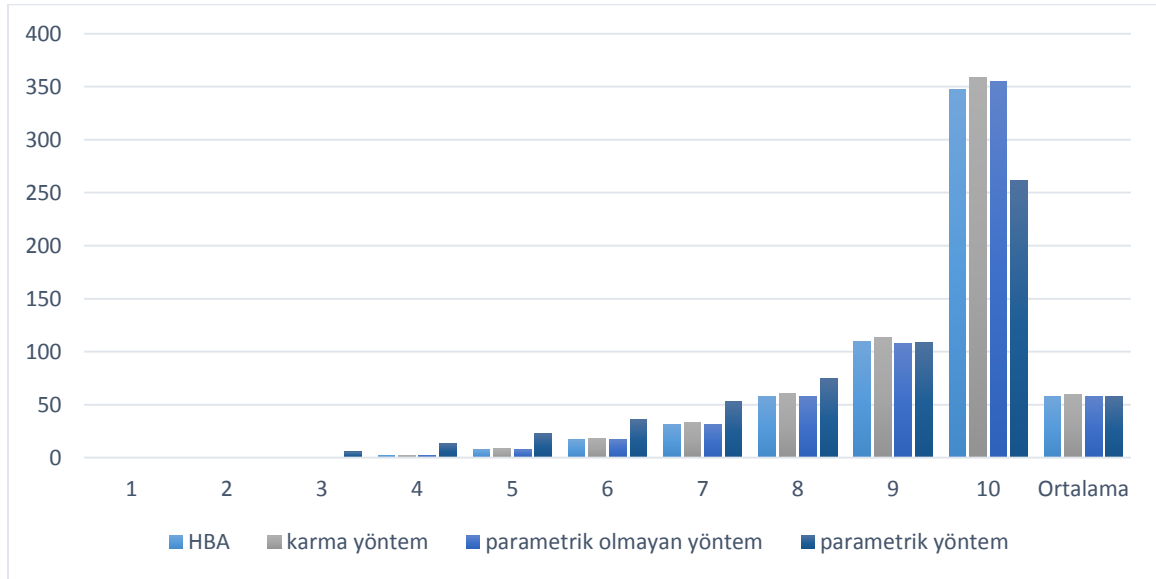
5.1.4.9. Eğlence ve Kültür

Bu grupta televizyon, DVD, VCD Player, ev sineması sistemi, çanak anten, televizyon anteni, radyo, müzik setleri, video, kaset çalar, CD çalar, kamera, fotoğraf makinesi vb. aletler, bilgisayar, fotokopi makinesi vb. veri işlem ekipmanı, optik aletler, müzik aletleri, oyuncaklar, hobi, spor, av ve kamp malzemeleri, bahçe, bitki ve evcil hayvanlara yönelik ürünler, sinema, tiyatro ve konser, müze, spor müsabakalarına, hayvanat bahçelerine giriş ücretleri, şans oyunları, kitap, gazete ve dergiler, kırtasiye materyalleri (kalem, defter, sulu boya, vb.), paket ve yurtdışı turlar ile ilgili olarak yapılan harcamalar bulunmaktadır.

HBA hanelerinin eğlence ve kültür grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur. İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla eğlence ve kültür harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık eğlence ve kültür harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.34'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.34. % 10'luk Harcama Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (% 50 örnekleme, TL/ay)

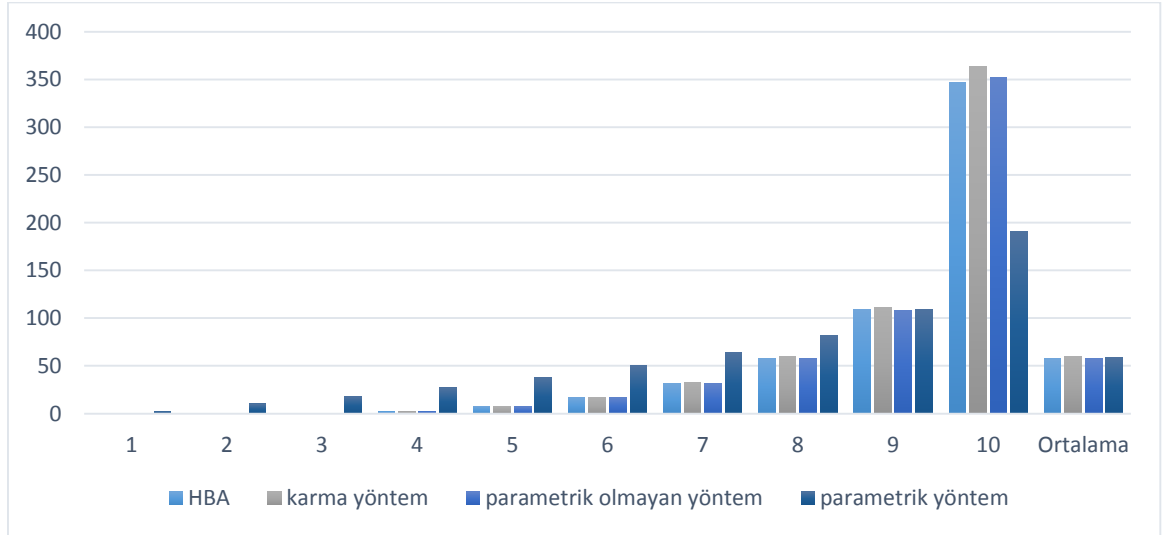
Bu durumu, Çizelge 5.26'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.26. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	6,11	0,53	53,04	2813,01
<i>parametrik olmayan yöntem</i>	6,73	0,29	55,82	3115,64
<i>parametrik yöntem</i>	29,46	1,98	104,58	10937,50

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin eğlence ve kültür harcamaları tahmin edilmiştir.

Şekil 5.35'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



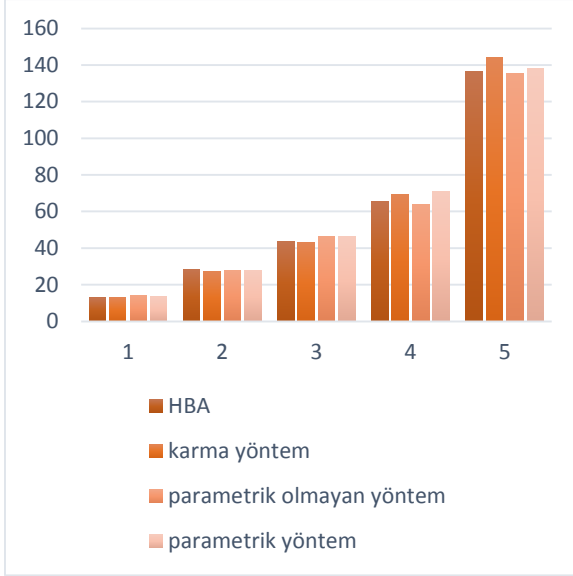
Şekil 5.35. % 10'luk Harcama Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay)

Bu durumu, Çizelge 5.27'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

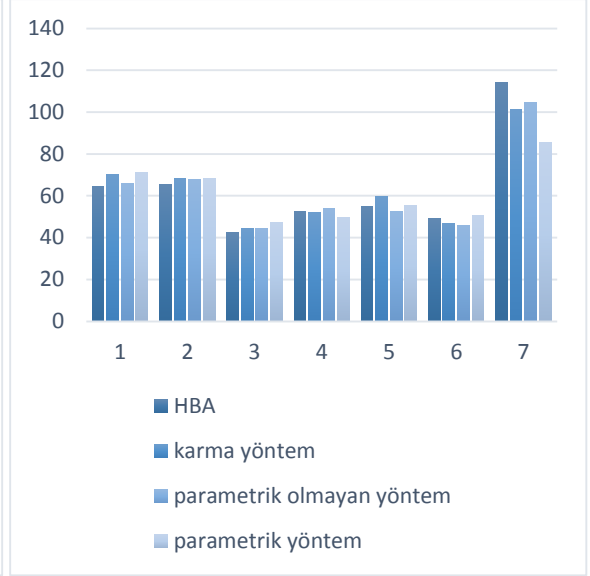
Çizelge 5.27. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	14,69	0,97	80,63	6500,51
<i>parametrik olmayan yöntem</i>	13,50	0,66	75,47	5696,36
<i>parametrik yöntem</i>	56,84	3,94	132,25	17488,65

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama eğlence ve kültür harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.36 ve Şekil 5.37'deki grafiklerde verilmiştir.



Şekil 5.36. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay)



Şekil 5.37. Hanehalkı Tipine Göre Ortalama Eğlence ve Kültür Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.36 ve Şekil 5.37'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.28'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.28. Hanehalkı Kullanılabilir Geliri İle Eğlence ve Kültür Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,301	0,319	0,280
Karma yöntem	0,323	0,409	0,313
Parametrik olmayan yöntem	0,218	0,324	0,248
Parametrik yöntem	0,785	0,819	0,677

5.1.4.10. Eğitim Hizmetleri

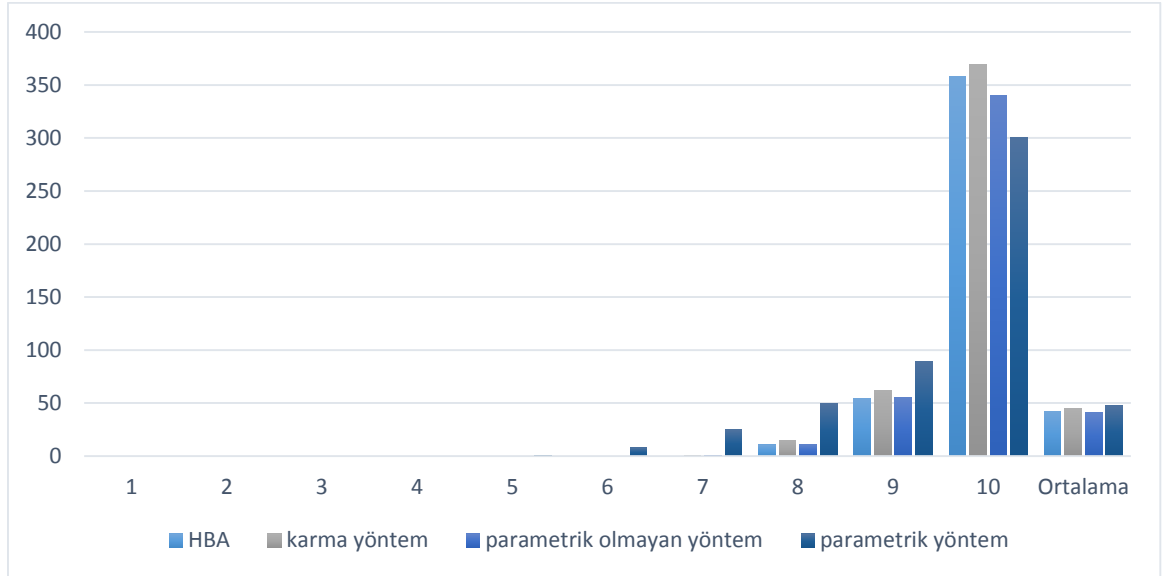
Bu grupta eğitim hizmetleri (okul öncesi, ilköğretim, ortaöğretim, ortaöğretim sonrası ve üniversite öncesi eğitim, diğer eğitimler), düzeyi belirlenemeyen eğitim programları (bilgisayar, kursu, yabancı dil kursu, müzik ve resim kursu vb.), sınav form ücreti ve özel ders ücreti ile ilgili harcamalar bulunmaktadır.

HBA hanelerinin eğitim hizmetleri grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla eğitim hizmetleri harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık eğitim hizmetleri harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.38'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.38. % 10'luk Harcama Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (% 50 örnekleme, TL/ay)

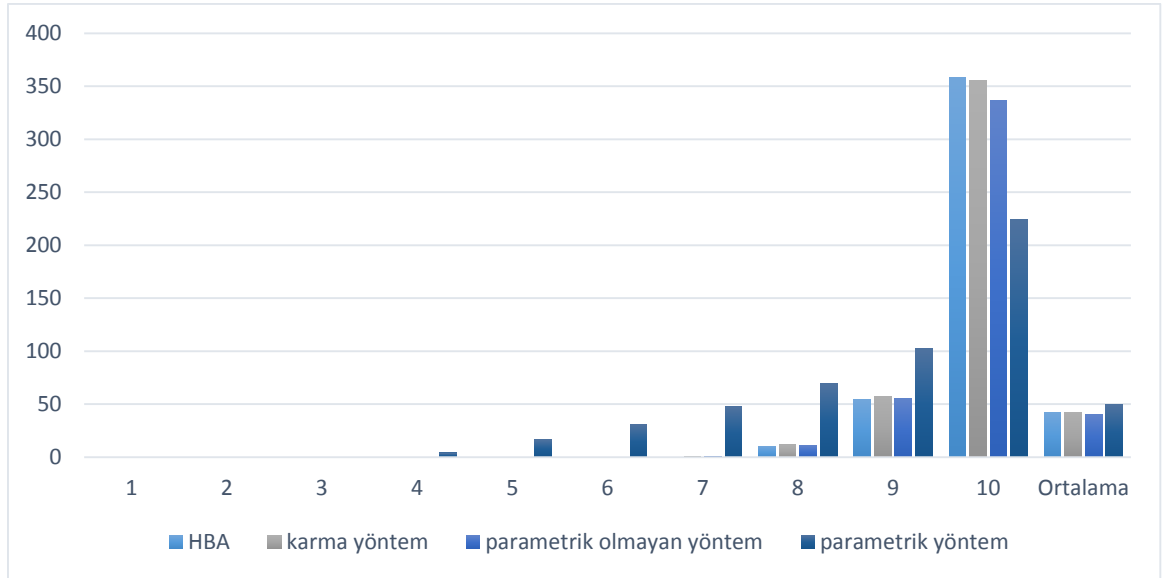
Bu durumu, Çizelge 5.29'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.29. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	5,02	0,05	59,11	3493,92
<i>parametrik olmayan yöntem</i>	6,12	0,08	63,42	4022,25
<i>parametrik yöntem</i>	26,35	0,35	101,14	10228,74

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin eğitim hizmetleri harcamaları tahmin edilmiştir.

Şekil 5.39'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



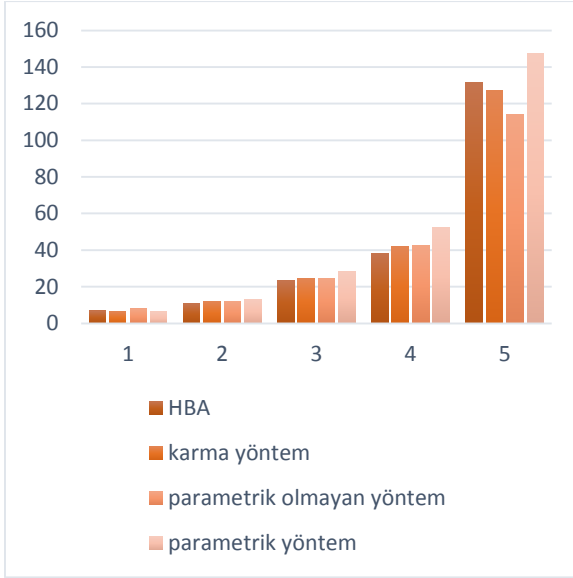
Şekil 5.39. % 10'luk Harcama Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay)

Bu durumu, Çizelge 5.30'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

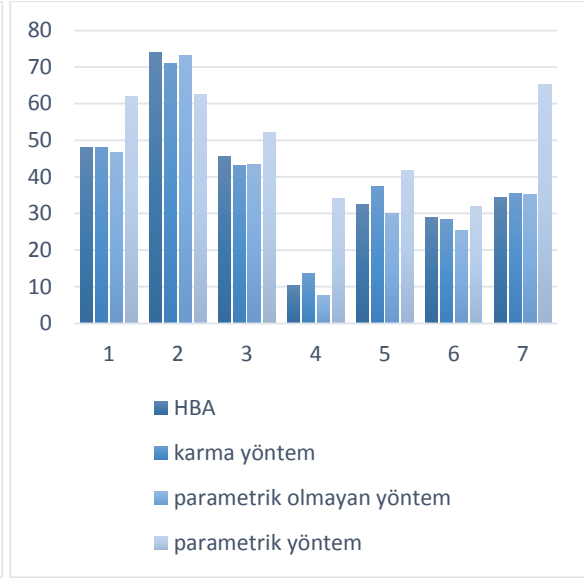
Çizelge 5.30. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	14,16	0,14	104,94	11013,23
<i>parametrik olmayan yöntem</i>	12,25	0,17	92,99	8647,61
<i>parametrik yöntem</i>	65,95	0,93	165,68	27450,59

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama eğitim hizmetleri harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.40 ve Şekil 5.41'deki grafiklerde verilmiştir.



Şekil 5.40. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay)



Şekil 5.41. Hanehalkı Tipine Göre Ortalama Eğitim Hizmetleri Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.40 ve Şekil 5.41'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.31'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.31. Hanehalkı Kullanılabilir Geliri ve Eğitim Hizmetleri Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,288	0,258	0,228
Karma yöntem	0,289	0,305	0,252
Parametrik olmayan yöntem	0,223	0,266	0,222
Parametrik yöntem	0,669	0,692	0,569

5.1.4.11. Lokanta ve Oteller

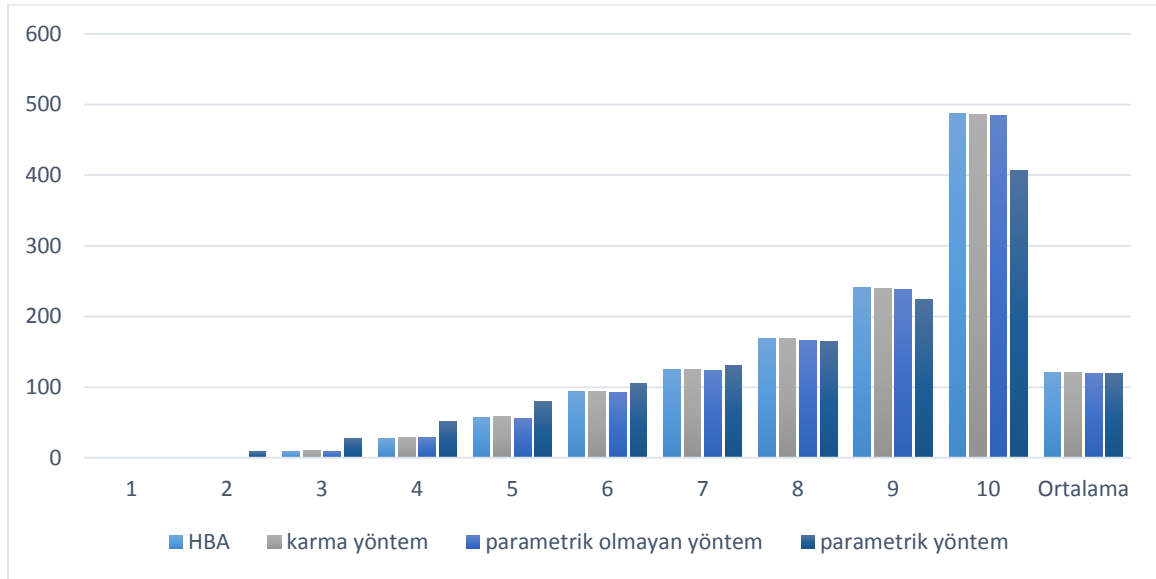
Bu grupta restoran, lokanta, pastane, kafe ve barlarda yenilen yemek ve içecekler, otel, motel, pansiyon, tatil köyleri, yurt ve misafirhaneler ile ilgili olarak yapılan harcamalar bulunmaktadır.

HBA hanelerinin lokanta ve oteller grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla lokanta ve oteller harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık lokanta ve oteller harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.42'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.42. % 10'luk Harcama Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (% 50 örnekleme, TL/ay)

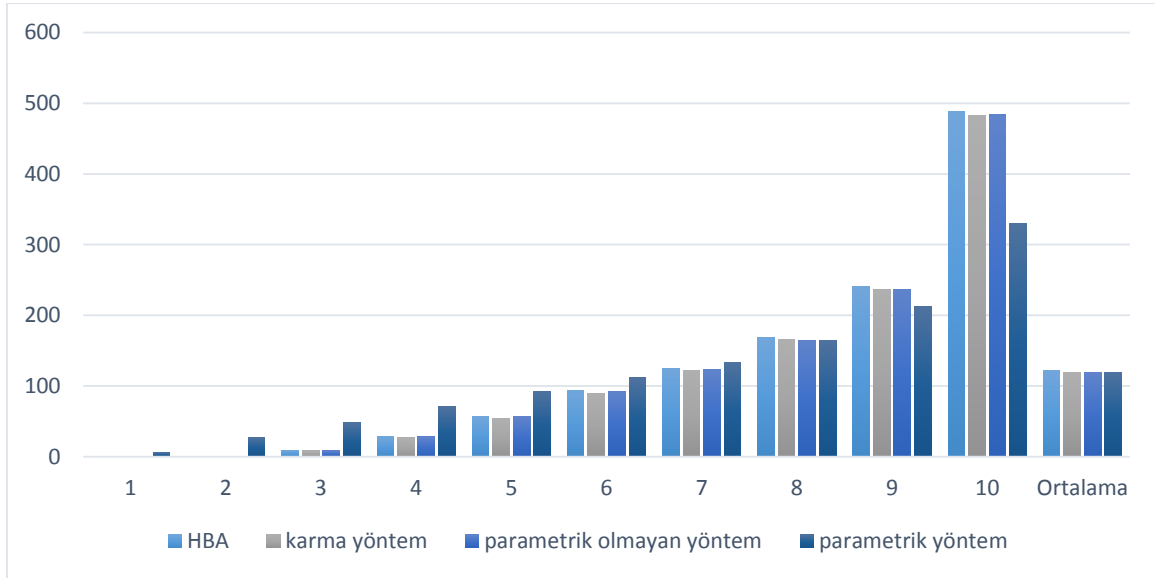
Bu durumu, Çizelge 5.32'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.32. % 50 Örneklemeli Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	9,93	0,27	66,24	4388,04
<i>parametrik olmayan yöntem</i>	10,92	0,39	53,45	2857,14
<i>parametrik yöntem</i>	36,86	1,33	96,31	9275,84

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin lokanta ve oteller harcamaları tahmin edilmiştir.

Şekil 5.43'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



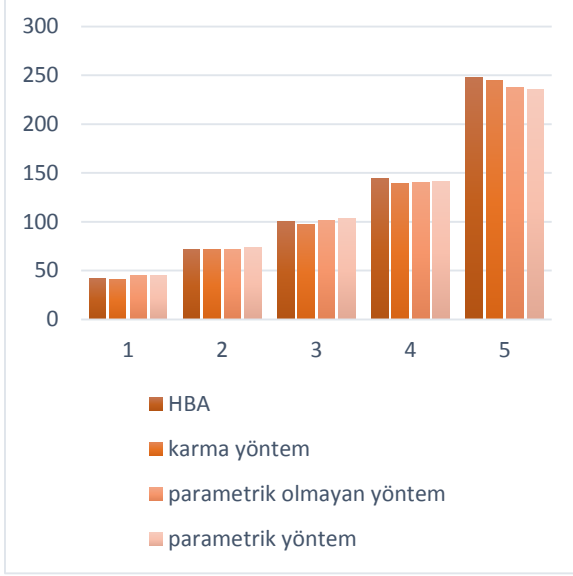
Şekil 5.43. % 10'luk Harcama Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay)

Bu durumu, Çizelge 5.33'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

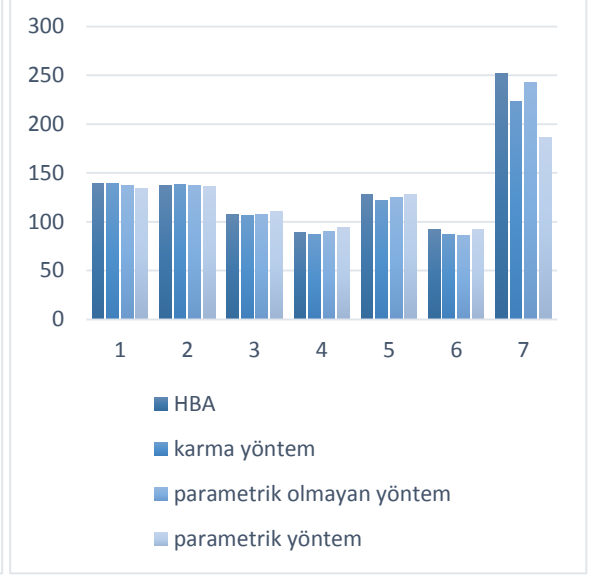
Çizelge 5.33. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	22,02	0,61	85,37	7288,30
<i>parametrik olmayan yöntem</i>	20,97	0,73	74,20	5505,64
<i>parametrik yöntem</i>	74,54	2,62	138,77	19257,00

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama lokanta ve oteller harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.44 ve Şekil 5.45'deki grafiklerde verilmiştir.



Şekil 5.44. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay)



Şekil 5.45. Hanehalkı Tipine Göre Ortalama Lokanta ve Oteller Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.44 ve Şekil 5.45'deki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.34'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.34. Hanehalkı Kullanılabilir Geliri İle Lokanta ve Oteller Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,431	0,420	0,365
Karma yöntem	0,375	0,410	0,309
Parametrik olmayan yöntem	0,261	0,307	0,231
Parametrik yöntem	0,700	0,683	0,536

5.1.4.12. Çeşitli Mal ve Hizmetler

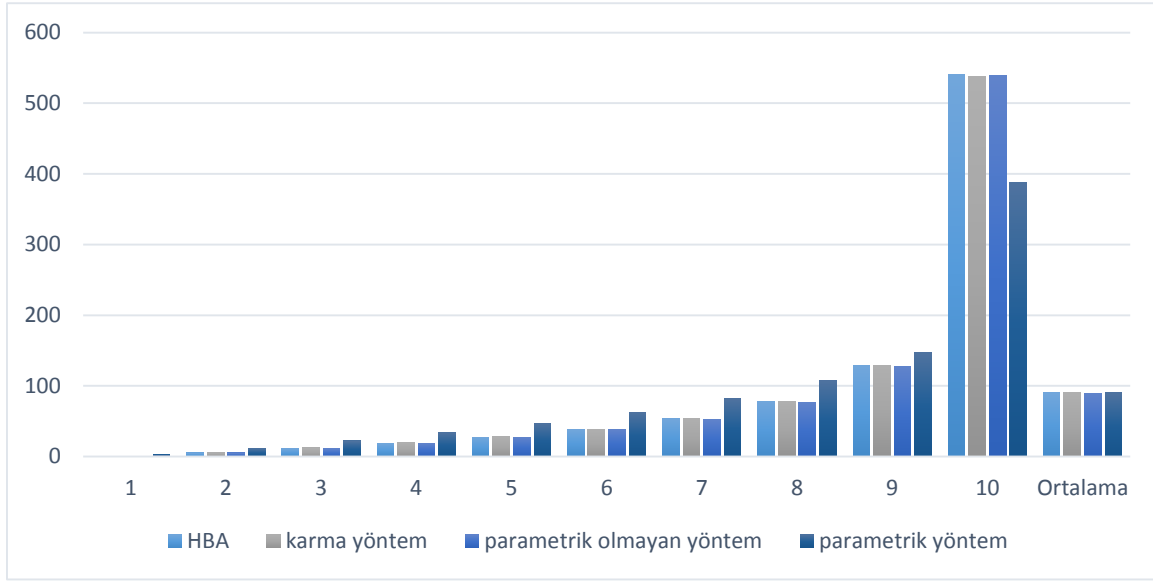
Bu grupta kişisel bakım (saç kesimi, saç boyama, cilt bakımı vb. kuaför hizmetleri, sauna ve hamam harcamaları), kişisel bakım için kullanılan elektrikli aletler (traş makineleri, saç kurutma makinesi, zayıflama aletleri, vb.), kişisel bakımda kullanılan diğer alet, ürün ve eşyalar (sabun, parfüm, tarak, tuvalet kağıdı, şampuan, pamuk, makyaj malzemeleri, diş macunu, deodorant, krem, bebek bakım ürünleri, vb.), mücevherler, saatler, seyahat eşyaları (valiz, vb.), kişisel aksesuarlar (çanta, cüzdan, şemsiye, çakmak, pipo, vb.), bebek için gerekli araçlar (ana kucağı, bebek koltuğu, vb.) kişisel aksesuarların tamir ve bakımı, sosyal hizmetler (huzurevleri, kreşler, vb.), sigortalar (konut, sağlık, ulaştırma, diğer), mali hizmetler ile ilgili olarak yapılan harcamalar (ehliyet alım masrafları, noter ücreti, avukatlık ücreti, emlakçı ücreti, cenaze masrafları, vestiyer ücreti vb.) bulunmaktadır.

HBA hanelerinin çeşitli mal ve hizmetler grubundaki maddelere yaptığı harcamalar baz alınarak üç yöntemle modeller kurulmuştur.

İlk olarak, HBA hanelerinden % 50 örnekleme alınarak bu haneler için kurulan model aracılığıyla çeşitli mal ve hizmetler harcamaları tahmin edilmiş ve kalan hanelerin kendi harcamaları dikkate alınmıştır.

Hanelerin aylık çeşitli mal ve hizmetler harcama değerleri, küçükten büyüğe doğru sıralandıktan sonra haneler 10 eşit parçaya bölünerek % 10'luk gruplar oluşturulmuş ve her bir grubun ortalama harcama tutarları hesaplanmıştır.

Şekil 5.46'daki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



Şekil 5.46. % 10'luk Harcama Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (% 50 örnekleme, TL/ay)

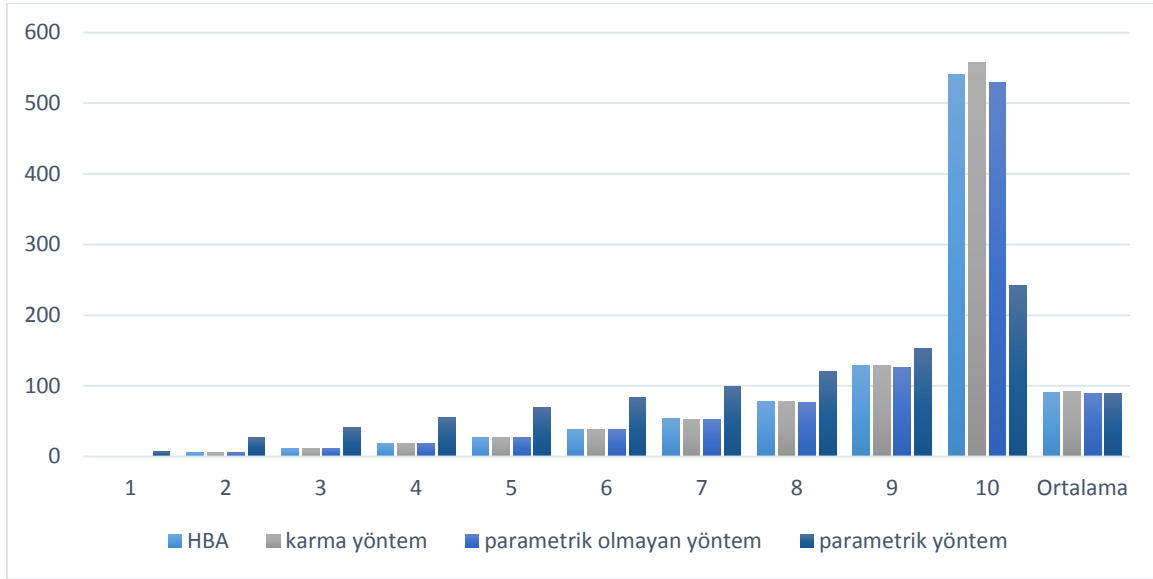
Bu durumu, Çizelge 5.35'deki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

Çizelge 5.35. % 50 Örnekleme Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	9,88	0,35	87,63	7679,39
<i>parametrik olmayan yöntem</i>	9,19	0,40	67,78	4594,52
<i>parametrik yöntem</i>	41,55	1,58	171,89	29545,56

İkinci olarak, kurulan model aracılığıyla tüm HBA hanelerinin çeşitli mal ve hizmetler harcamaları tahmin edilmiştir.

Şekil 5.47'deki grafikten de görüleceği üzere, genel olarak tüm %10'luk dilimlerde gerçek HBA değerleri ile karma ve parametrik olmayan yöntem ile kurulan model sonuçları yakın sonuçlar vermiştir.



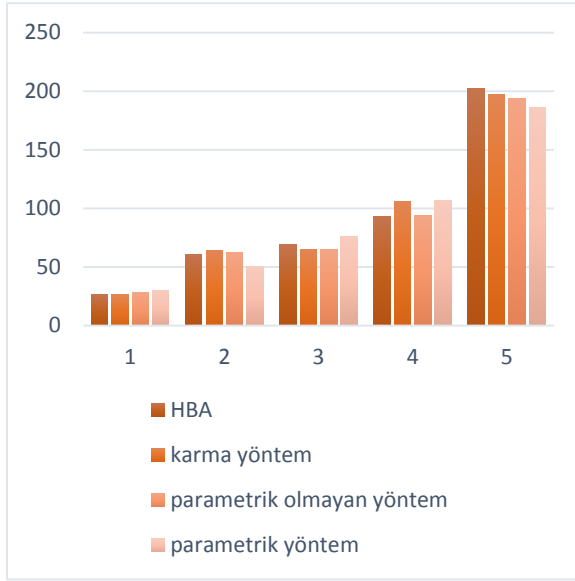
Şekil 5.47. % 10'luk Harcama Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay)

Bu durumu, Çizelge 5.36'daki model performanslarını değerlendirme amaçlı hesaplanan göstergeler de doğrulamıştır.

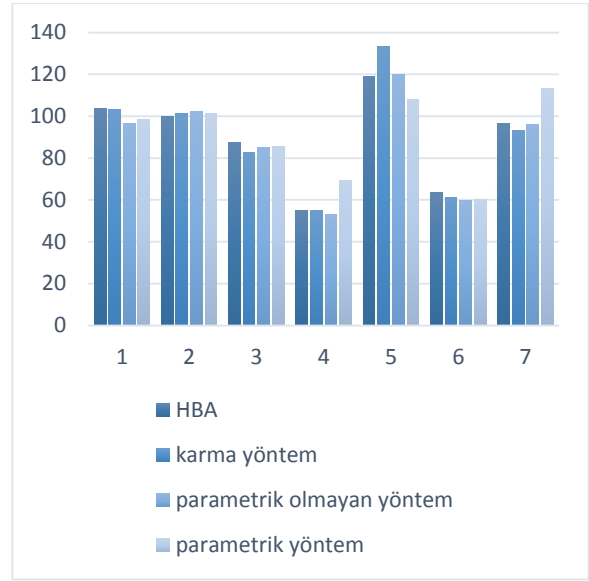
Çizelge 5.36. Model Performansları

Yöntem	MAE	MAPE	RMSE	MSE
<i>karma yöntem</i>	22,00	0,80	219,05	47984,15
<i>parametrik olmayan yöntem</i>	19,45	0,70	121,49	14759,18
<i>parametrik yöntem</i>	83,04	3,19	282,57	79843,20

Üçüncü olarak, HBA hanelerinin hanehalkı kullanılabilir gelire göre sıralanmış % 20'lik grupların ve hanehalkı tipine göre aylık ortalama çeşitli mal ve hizmetler harcamaları kurulan modellerle tahmin edilmiştir. Sonuçlar, Şekil 5.48 ve Şekil 5.49'daki grafiklerde verilmiştir.



Şekil 5.48. % 20'lik Hanehalkı Kullanılabilir Gelir Gruplarına Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay)



Şekil 5.49. Hanehalkı Tipine Göre Ortalama Çeşitli Mal ve Hizmetler Harcamaları (TL/ay)

Gerçek HBA hanelerinin değerleri kullanılarak hesaplanan göstergeler ile her üç yöntemle kurulan modellerle tahmin edilmiş değerler kullanılarak hesaplanan göstergelerin birbirleriyle uyumlu sonuçlar verdiği Şekil 5.48 ve Şekil 5.49'daki grafiklerden görülmektedir. Bu anlamda kurulan modellerin performanslarının genel olarak iyi olduğu sonucuna varılabilir. Çizelge 5.37'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

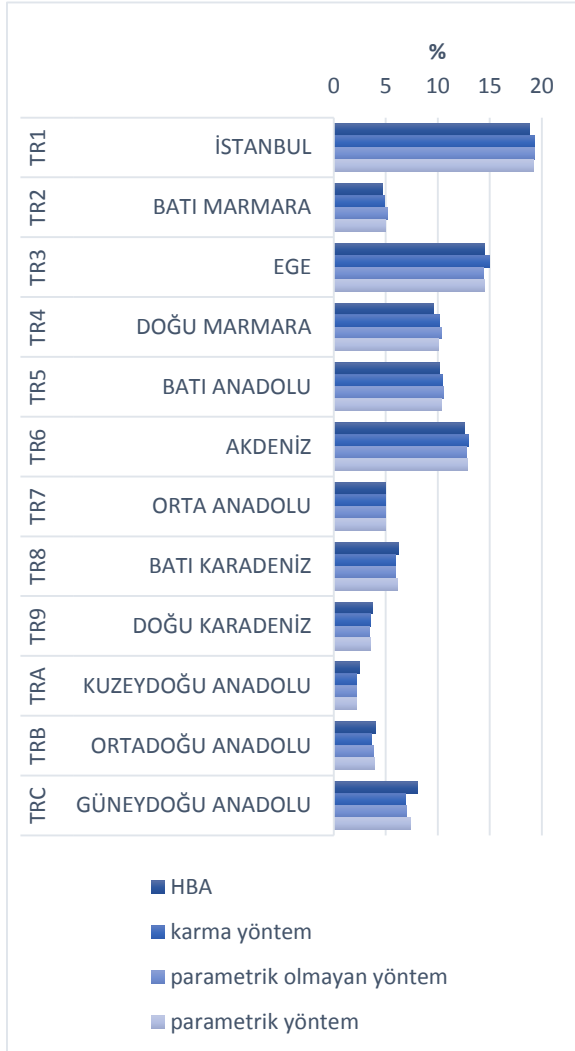
Çizelge 5.37. Hanehalkı Kullanılabilir Geliri İle Çeşitli Mal ve Hizmetler Harcamaları Sınıfları Arasındaki Korelasyon

	Pearson	Spearman	Kendall
HBA	0,299	0,315	0,275
Karma yöntem	0,338	0,459	0,344
Parametrik olmayan yöntem	0,100	0,349	0,260
Parametrik yöntem	0,800	0,803	0,658

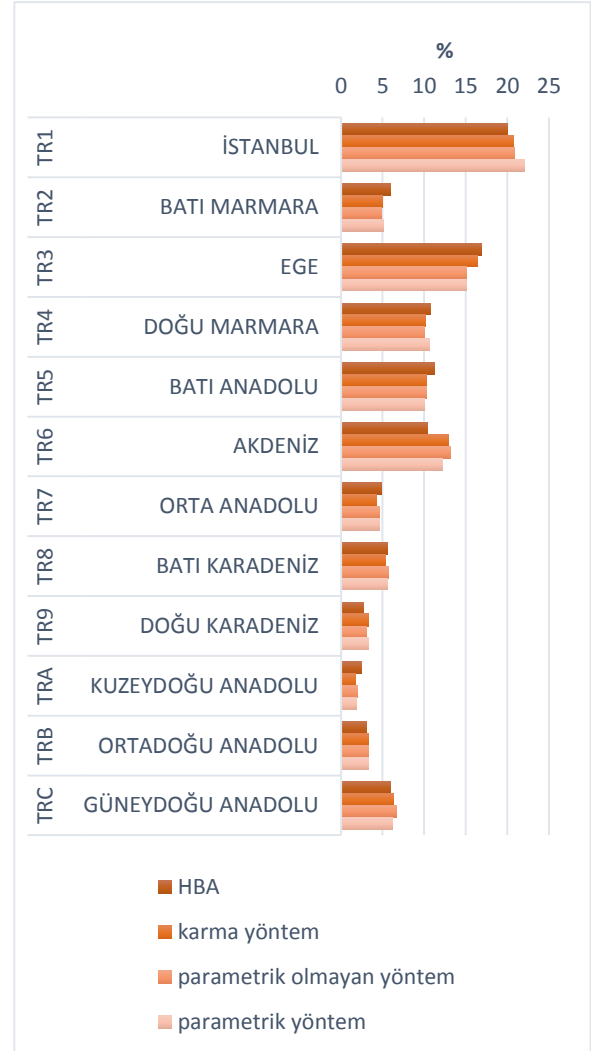
5.1.4.13. GYKA Haneleri İçin Harcama Değerlerinin Tahmin Edilmesi

Önceki bölümlerde her bir harcama grubu bazında model performansları test edilen modeller ile GYKA haneleri için 12'li grup bazında harcama değerleri tahmin edilerek, İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları haber bülteni ile açıklanan HBA dikey yüzdesel dağılımları ile birlikte aşağıda verilmiştir.

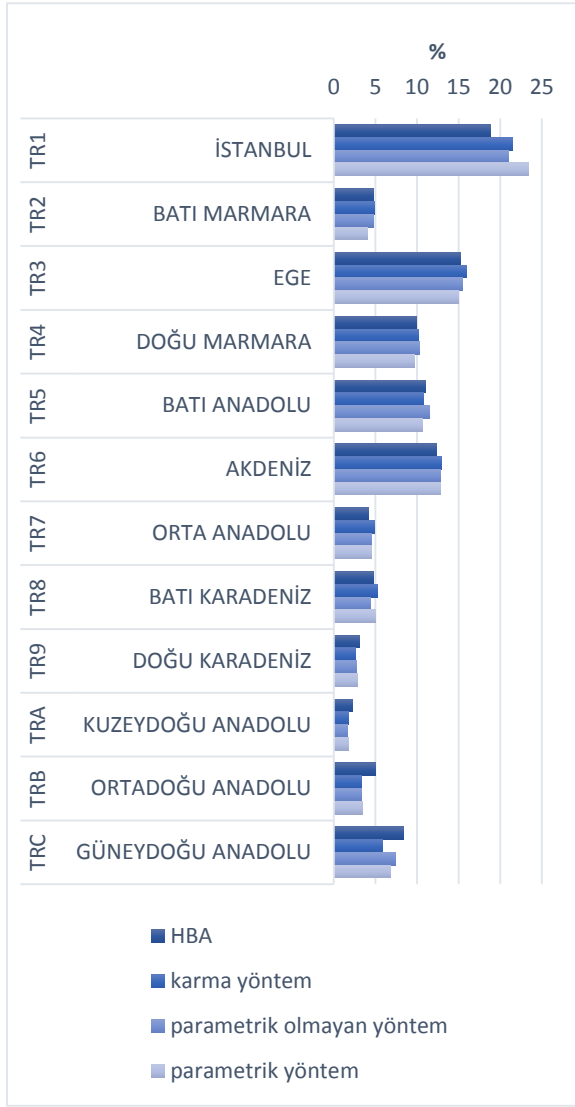
Şekil 5.50 ve Şekil 5.51'deki grafiklerden görüleceği üzere, gıda ve alkolsüz içecek grubuna ve alkollü içecekler, sigara ve tütün grubuna yapılan harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir.



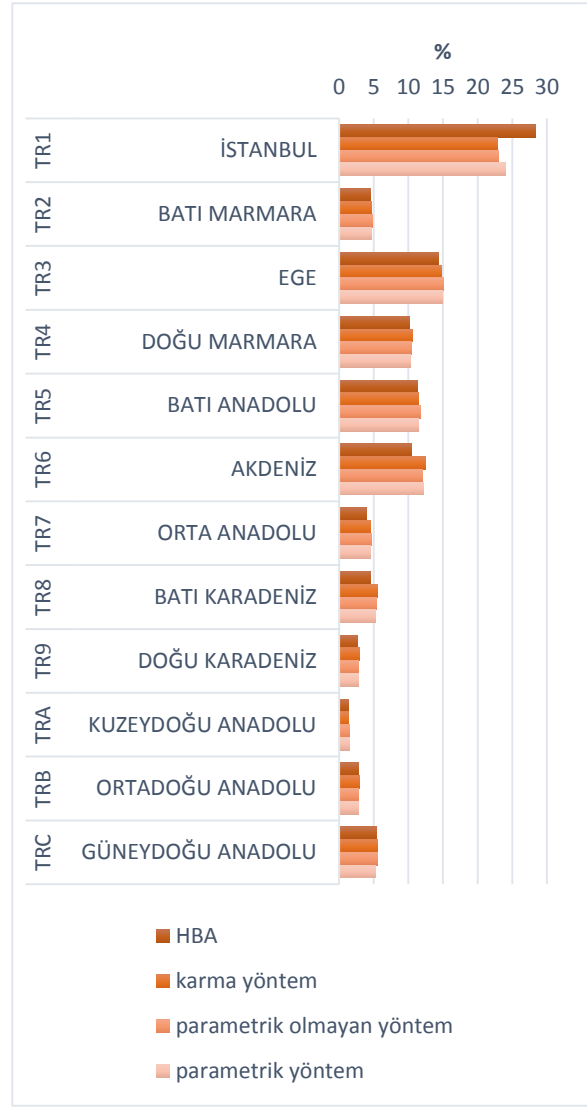
Şekil 5.50. Gıda ve Alkolsüz İçecek Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)



Şekil 5.51. Alkollü İçecekler, Sigara ve Tütün Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

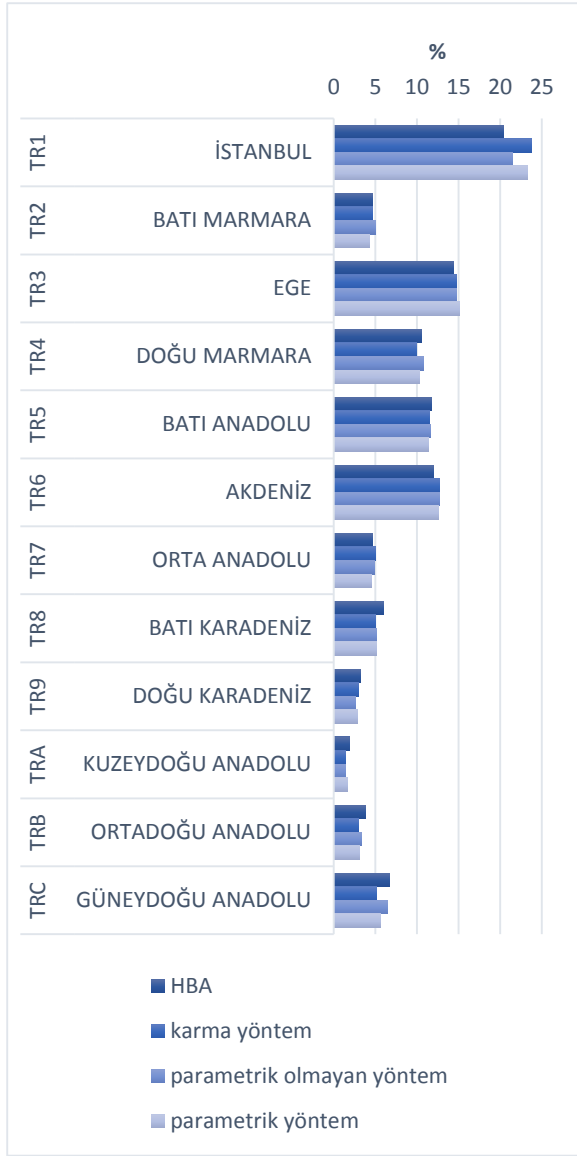


Şekil 5.52. Giyim ve Ayakkabı Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

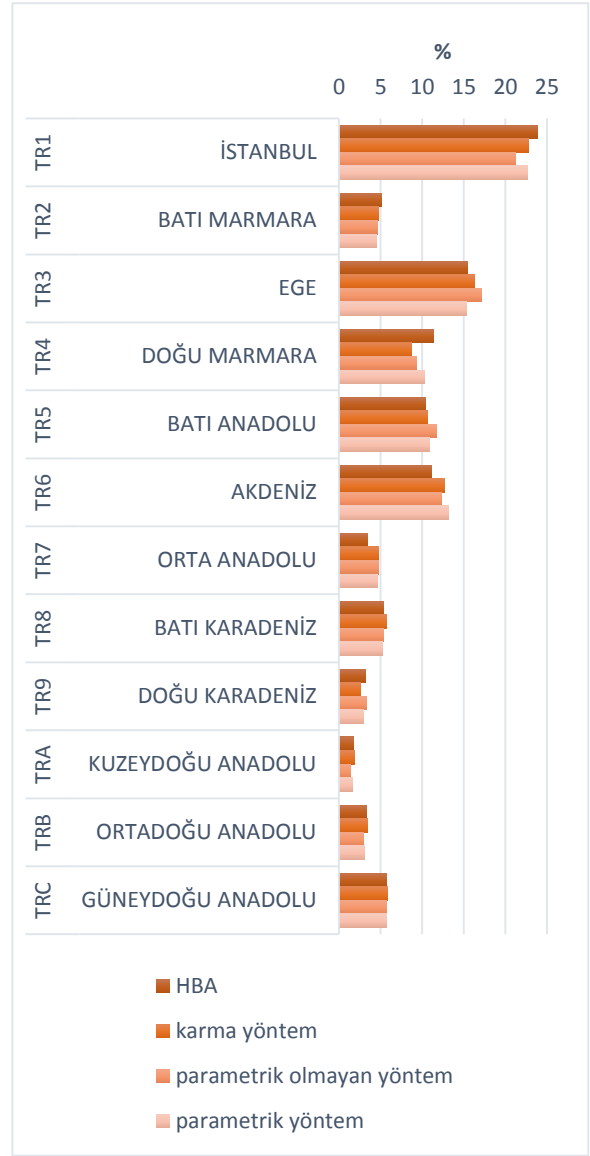


Şekil 5.53. Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.52 ve Şekil 5.53'deki grafiklerden görüleceği üzere, giyim ve ayakkabı grubuna ve konut, su, elektrik, gaz ve diğer yakıtlar grubuna yapılan harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir. Yüzdesel olarak sapmanın en çok olduğu İstanbul Bölgesinde her iki harcama grubu sırasıyla HBA sonuçlarına göre bölgeler içinde %19 ve %28 pay alırken, bu kapsamda GYKA haneleri ile yapılan en iyi tahmin değerleri sırasıyla %21 ve %24'dür.

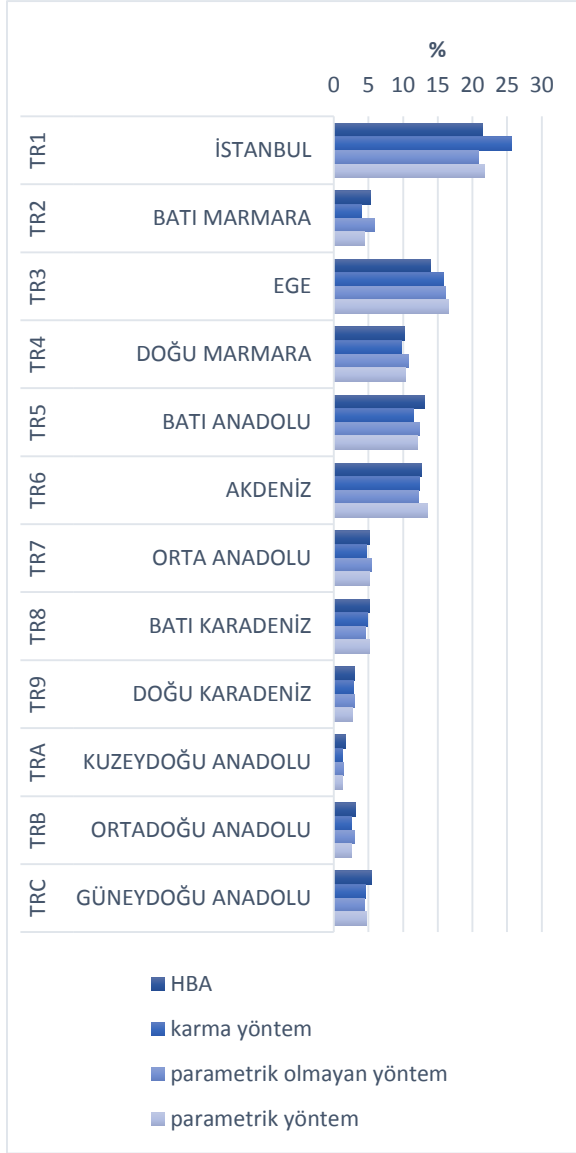


Şekil 5.54. Mobilya, Ev Aletleri ve Ev Bakım Hizmetler Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

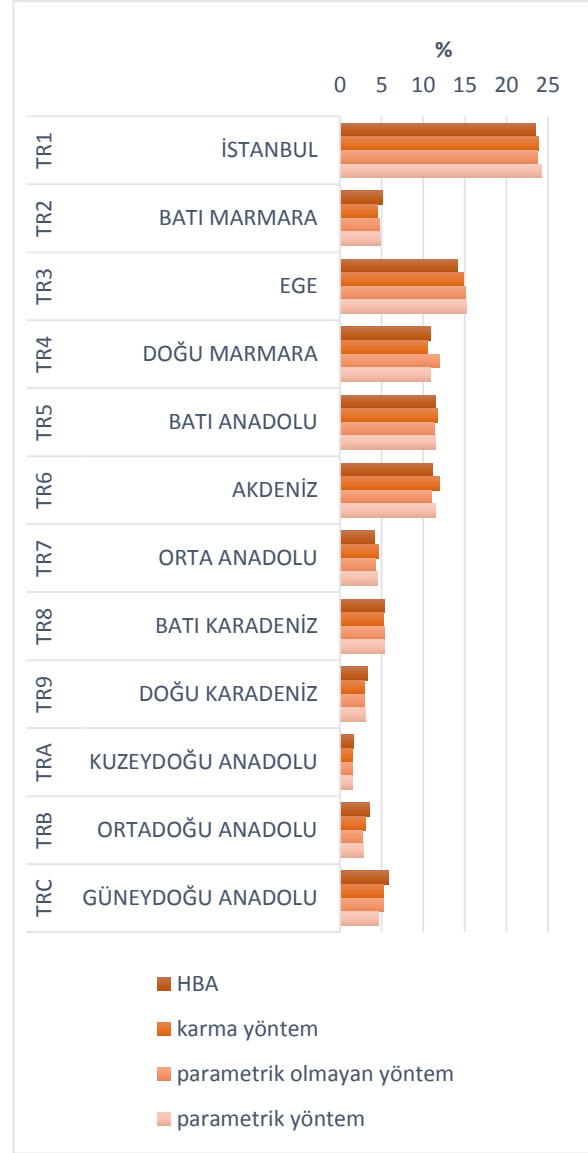


Şekil 5.55. Sağlık Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.54 ve Şekil 5.55'deki grafiklerden görüleceği üzere, mobilya, ev aletleri ve ev bakım hizmetler grubuna ve sağlık grubuna yapılan harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir. Yüzdesel olarak sapmanın en çok olduğu Orta Anadolu Bölgesinde sağlık harcamaları HBA sonuçlarına göre bölgeler içinde %3 pay alırken, bu kapsamda GYKA haneleri ile yapılan en iyi tahmin değeri %5'dir.

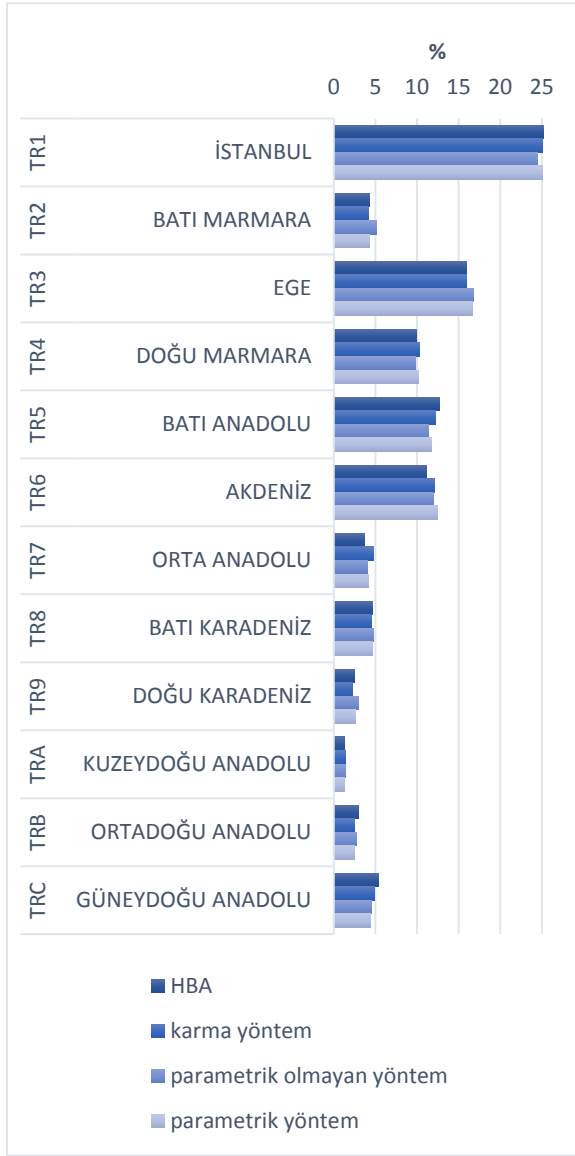


Şekil 5.56. Ulaştırma Harcamalarının İstatistik Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

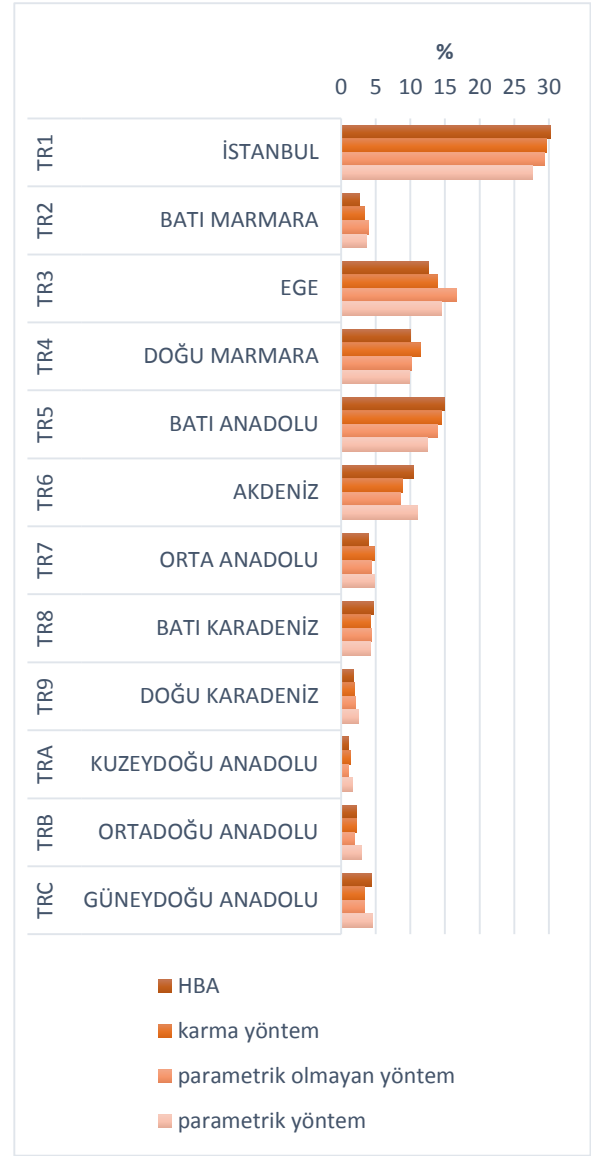


Şekil 5.57. Haberleşme Harcamalarının İstatistik Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.56 ve Şekil 5.57'deki grafiklerden görüleceği üzere, ulaştırma grubuna ve haberleşme grubuna yapılan harcamaların İstatistik Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir. Yüzdesel olarak sapmanın en çok olduğu Ege Bölgesinde ulaştırma harcamaları HBA sonuçlarına göre bölgeler içinde %14 pay alırken, bu kapsamda GYKA haneleri ile yapılan en iyi tahmin değeri %16'dır.

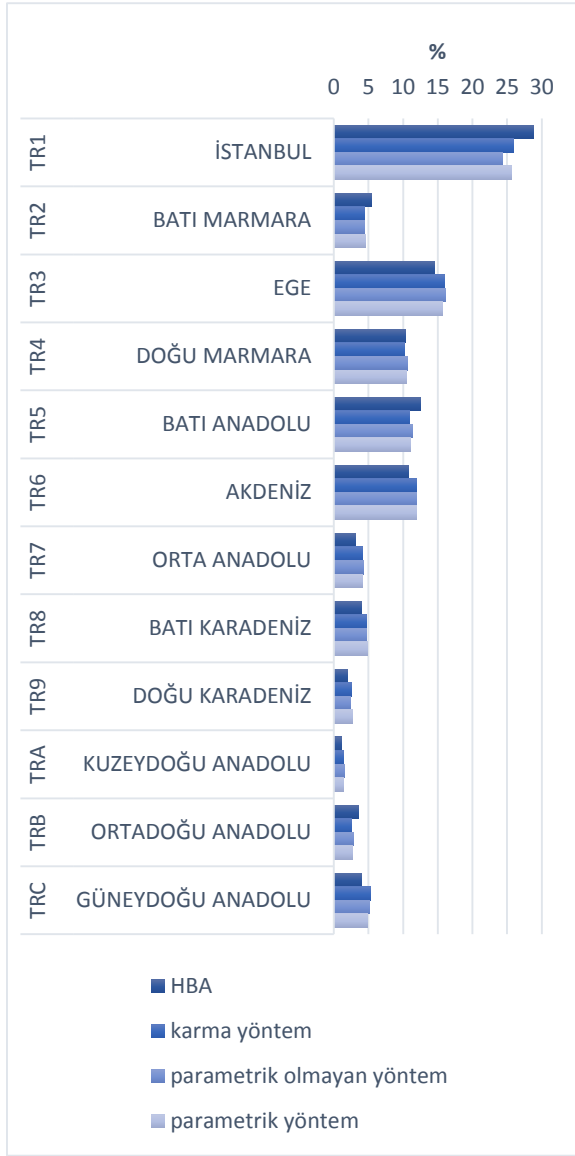


Şekil 5.58. Eğlence ve Kültür Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

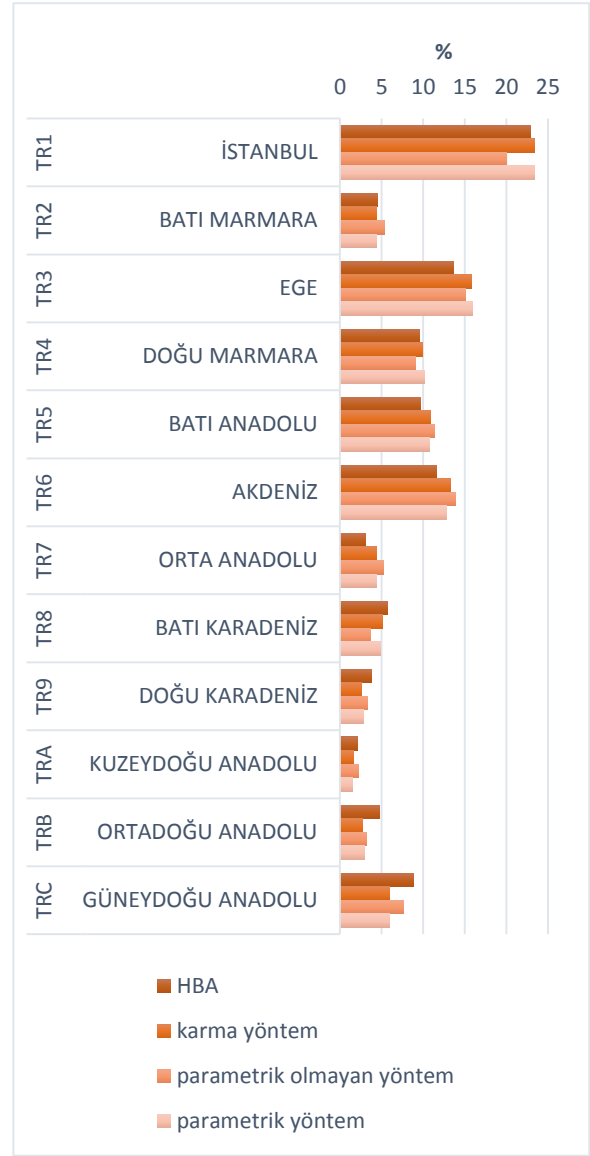


Şekil 5.59. Eğitim Hizmetleri Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.58 ve Şekil 5.59'daki grafiklerden görüleceği üzere, eğlence ve kültür grubu ile eğitim hizmetleri grubuna yapılan harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir.

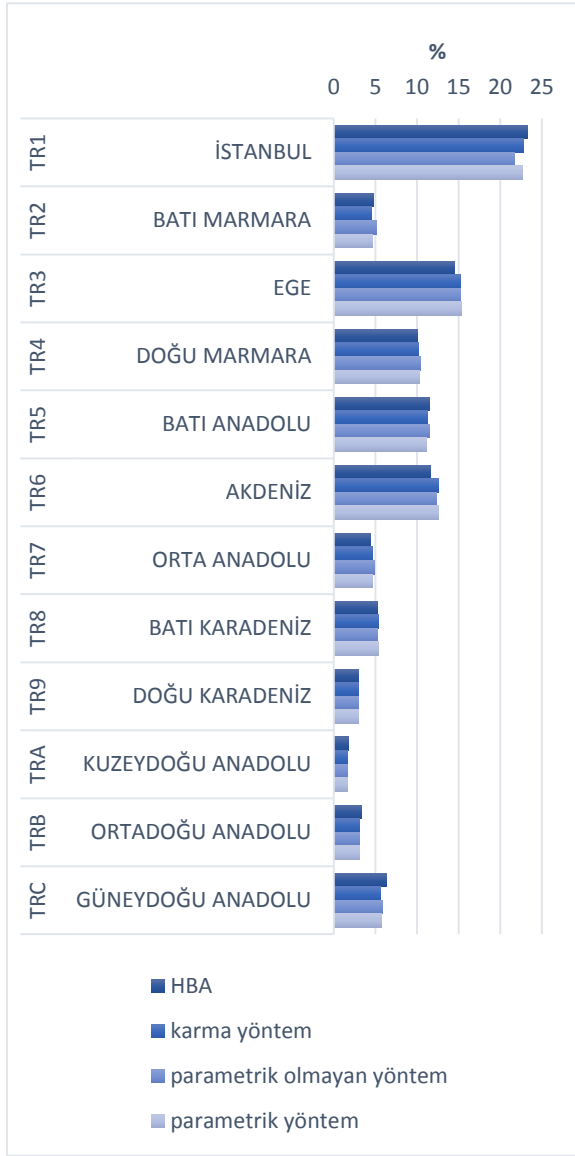


Şekil 5.60. Lokanta ve Oteller Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)



Şekil 5.61. Çeşitli Mal ve Hizmetler Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.60 ve Şekil 5.61'deki grafiklerden görüleceği üzere, lokanta ve oteller grubuna ve çeşitli mal ve hizmetler grubuna yapılan harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir. Yüzdesel olarak sapmanın en çok olduğu İstanbul Bölgesinde lokanta ve oteller harcamaları HBA sonuçlarına göre bölgeler içinde %29 pay alırken, bu kapsamda GYKA haneleri ile yapılan en iyi tahmin değeri %26'dır.



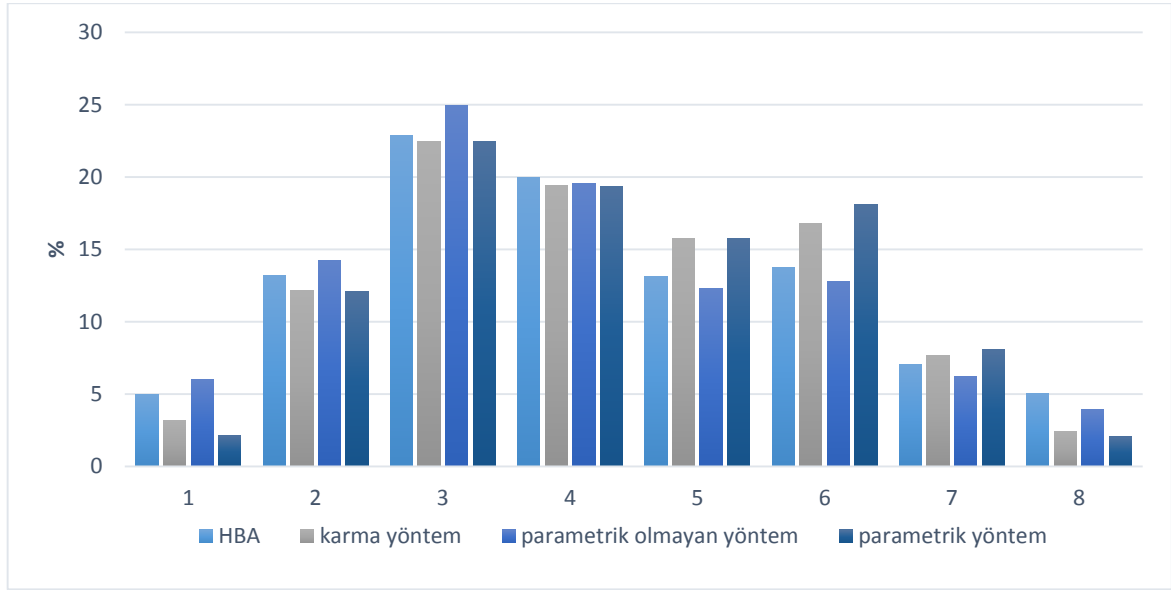
Şekil 5.62. Toplam Harcamalarının İstatistiki Bölge Sınıflaması Düzey-1'e Göre Dağılımı (Dikey %)

Şekil 5.62'deki grafikten görüleceği üzere, toplam harcamaların İstatistiki Bölge Sınıflaması Düzey-1'e göre dikey yüzdesel dağılımları ile HBA dikey yüzdesel dağılımları hemen hemen tüm bölgelerde benzerlik göstermektedir. Ayrıca, Çizelge 5.38'de hanehalkı kullanılabilir geliri ve oluşturulan harcama sınıfları arasında korelasyonlar verilmiştir.

Çizelge 5.38. Hanehalkı Kullanılabilir Gelir ve Toplam Harcama Sınıfları Arasındaki Korelasyon

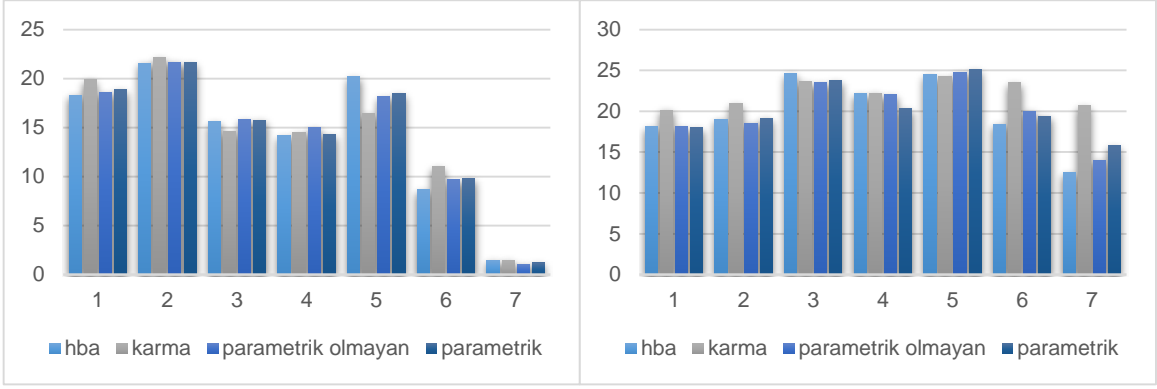
	Pearson	Spearman	Kendall
HBA	0,731	0,731	0,625
Karma yöntem	0,804	0,829	0,688
Parametrik olmayan yöntem	0,455	0,546	0,415
Parametrik yöntem	0,828	0,846	0,709

Şekil 5.63'da HBA ve GYKA harcama sınıflarının karşılaştırılması verilmiştir. Parametrik olmayan yöntem ile yapılan harcama tahminlerinin % 3.48 Hellinger uzaklığı ile HBA sonuçlarına çok benzer olduğu söylenebilir.

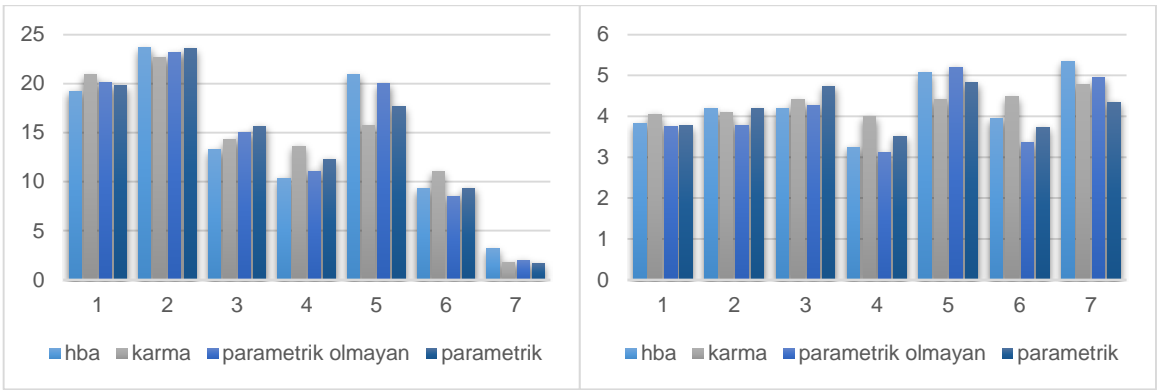


Şekil 5.63. HBA ve GYKA Harcama Sınıflarının Karşılaştırılması (%)

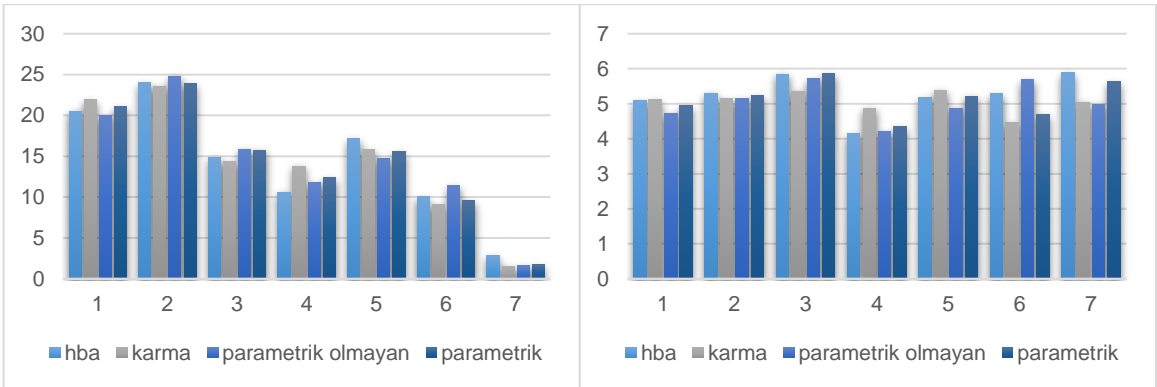
GYKA haneleri için 12'li harcama grubu bazında harcama değerleri tahmin edilerek, Hanehalkı Tipine göre dikey ve yatay yüzdesel dağılımları ile HBA dikey ve yüzdesel dağılımları ile birlikte karşılaştırmalı olarak Şekil 5.64 ve Şekil 5.75 arasında yer alan grafiklerde verilmiştir. Dağılımların birbiri ile uyumlu oldukları görülmektedir.



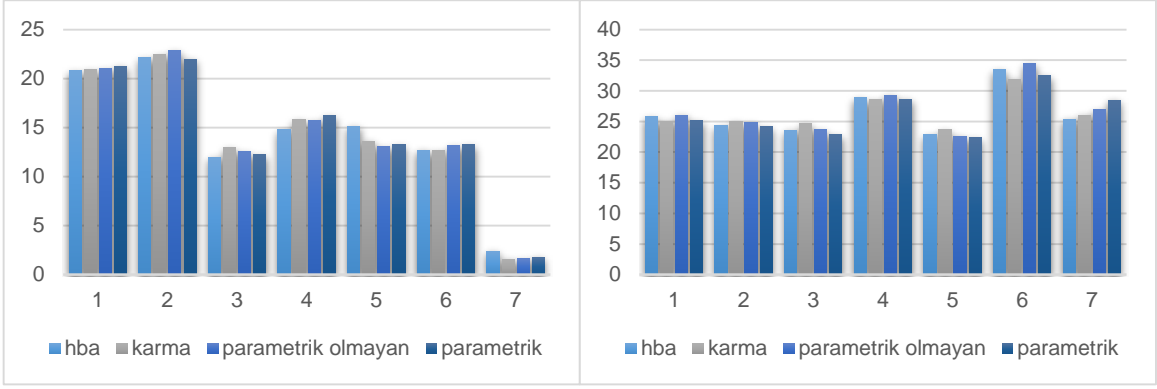
Şekil 5.64. Hanehalkı Tipine Göre Gıda ve Alkolsüz İçecekler Harcamalarının Dikey ve Yatay % Dağılımı



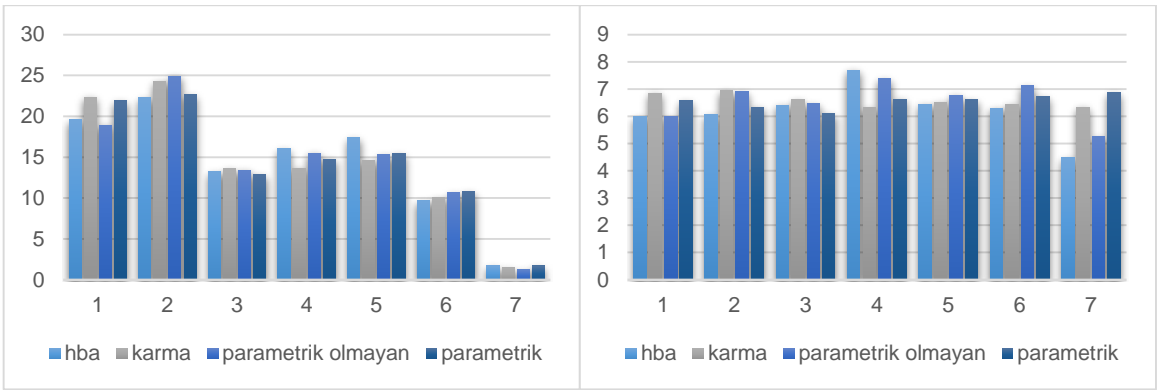
Şekil 5.65. Hanehalkı Tipine Göre Alkollü İçecekler, Sigara ve Tütün Harcamalarının Dikey ve Yatay % Dağılımı



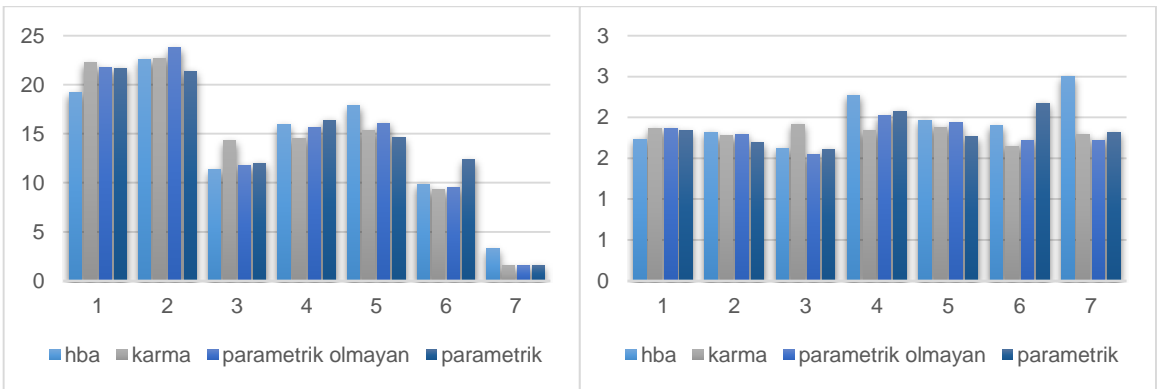
Şekil 5.66. Hanehalkı Tipine Göre Giyim ve Ayakkabı Harcamalarının Dikey ve Yatay % Dağılımı



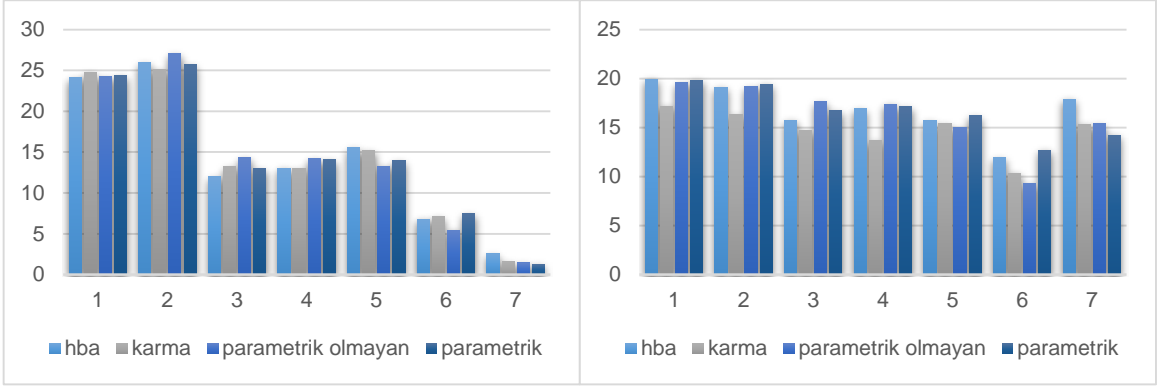
Şekil 5.67. Hanehalkı Tipine Göre Konut, Su, Elektrik, Gaz ve Diğer Yakıtlar Harcamalarının Dikey ve Yatay % Dağılımı



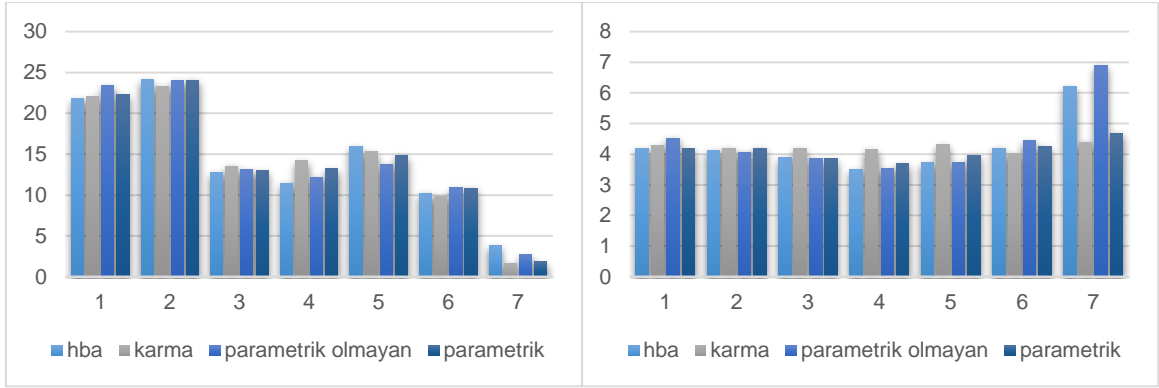
Şekil 5.68. Hanehalkı Tipine Göre Mobilya, Ev Aletleri ve Ev Bakım Hizmetler Harcamalarının Dikey ve Yatay % Dağılımı



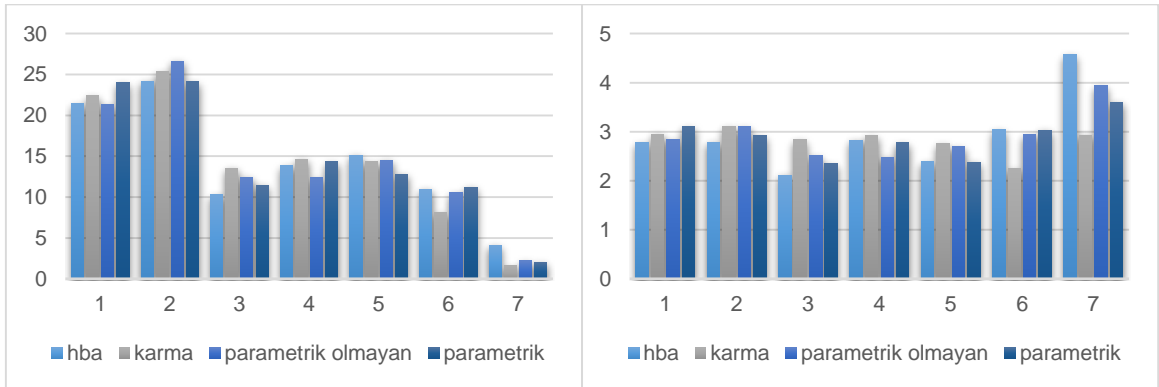
Şekil 5.69. Hanehalkı Tipine Göre Sağlık Harcamalarının Dikey ve Yatay % Dağılımı



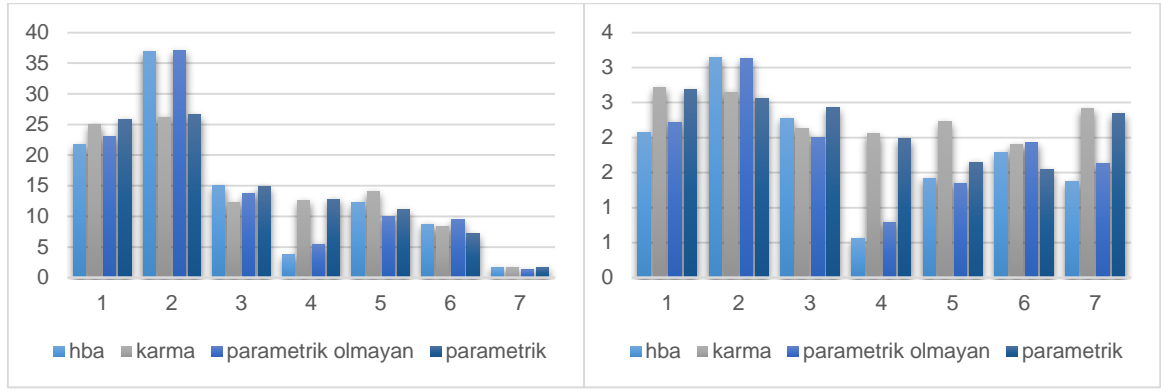
Şekil 5.70. Hanehalkı Tipine Göre Ulaştırma Harcamalarının Dikey ve Yatay % Dağılımı



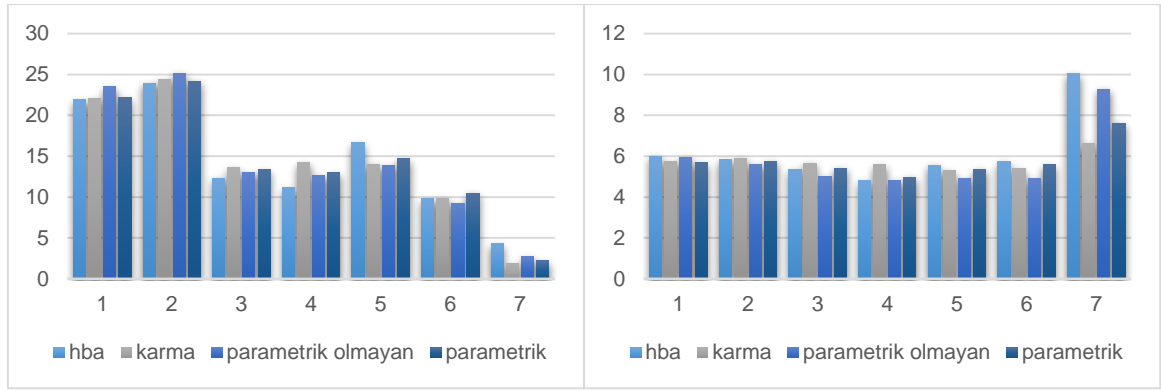
Şekil 5.71. Hanehalkı Tipine Göre Haberleşme Harcamalarının Dikey ve Yatay % Dağılımı



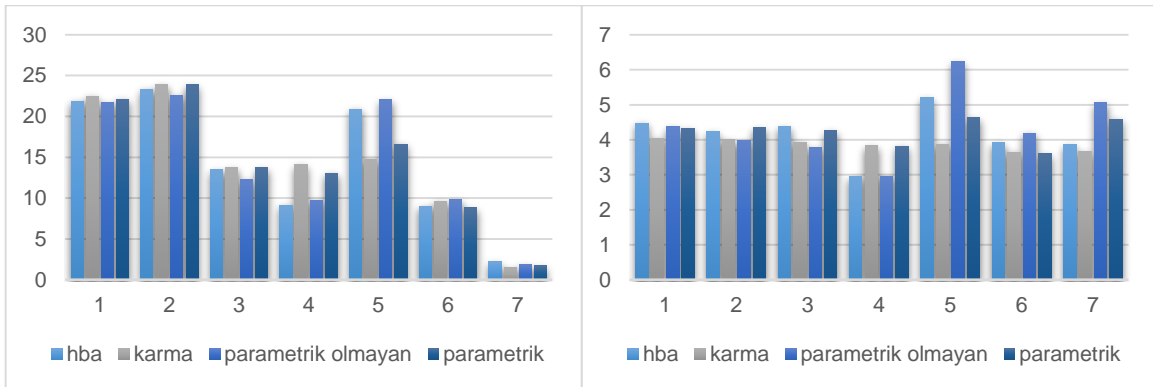
Şekil 5.72. Hanehalkı Tipine Göre Eğlence ve Kültür Harcamalarının Dikey ve Yatay % Dağılımı



Şekil 5.73. Hanehalkı Tipine Göre Eğitim Hizmetleri Harcamalarının Dikey ve Yatay % Dağılımı



Şekil 5.74. Hanehalkı Tipine Göre Lokanta ve Oteller Harcamalarının Dikey ve Yatay % Dağılımı



Şekil 5.75. Hanehalkı Tipine Göre Çeşitli Mal ve Hizmetler Harcamalarının Dikey ve Yatay % Dağılımı

Yoksulluk, bireylerin ve hanehalklarının, toplumun ortalama refah düzeyinin belli bir oranının altında olması durumudur. Buna göre, toplumun genel düzeyine göre belli bir sınırın altında gelir ve harcamaya sahip olan birey veya hanehalkı görece anlamda yoksul olarak tanımlanır. Genellikle gelir veya tüketim harcamaları medyan değerinin %40'ı, % 50'si, %60'ı veya %70'i sınır değer olarak kullanılmaktadır. Karma yöntem ve parametrik olmayan yöntem ile hesaplanan yoksulluk oranlarının

HBA yoksulluk oranları ile uyumlu oldukları Şekil 5.76 ve Şekil 5.79 arasında yer alan grafiklerde görülmektedir.



Şekil 5.76. Toplam Harcama Medyan Değerinin %40'ına Göre Hesaplanan Yoksul Hane Oranı (%)



Şekil 5.77. Toplam Harcama Medyan Değerinin % 50'sine Göre Hesaplanan Yoksul Hane Oranı (%)



Şekil 5.78. Toplam Harcama Medyan Değerinin %60'ına Göre Hesaplanan Yoksul Hane Oranı (%)



Şekil 5.79. Toplam Harcama Medyan Değerinin %70'ine Göre Hesaplanan Yoksul Hane Oranı (%)

5.1.5. Değerlendirme

Son aşama olan uygulama aşamasına geçmeden, modeli baştan aşağı değerlendirmek ve iş amaçlarını tam olarak yerine getirdiğinden emin olmak için modeli oluştururken uygulanan adımları gözden geçirmek önemlidir. Başlıca nokta önemli iş konularının yeterince dikkate alınıp alınmadığının kontrol edilmesidir. Bu aşamanın sonunda veri madenciliği sonuçlarının kullanımı ile ilgili bir sonuca varılmalıdır.

5.1.6. Uygulama

Modeli oluşturmak genelde projenin sona erdiği anlamına gelmez. Modelin amacı verinin bilgisini, değerini arttırmak olsa da, elde edilen bilgi müşterinin kullanabileceği şekilde düzenlenmeli ve sunulmalıdır. Şartlara bağlı olarak uygulama aşaması, rapor üretmek kadar basit ya da tekrarlanabilen veri madenciliği süreci oluşturmak kadar karmaşık olabilir. Çoğu zaman uygulama

adımlarını yrten kiři veri analisti deęil mřteridir. Uygulamayı analist yrtmeyecek bile olsa, mřterinin yaratılacak modellerden faydalanmak iin hangi adımları atması gerektięini anlaması nemlidir.

6. SONUÇ VE TARTIŞMA

İstatistikler ve istatistiksel analizler her alanda her zaman çok önem taşımıştır. Yoğun rekabet ortamında çalışan kurumlar kendilerini farklılaştırabilmek için bilgiyi etkin ve verimli bir biçimde kullanma yoluna gitmektedirler. Veri madenciliği uygulamalarından fayda sağlanmasına neden olan en önemli faktörlerden birisi, çok miktarda veriyi istenilen şekilde işleyebilme olanağı sağlamasıdır. Veri madenciliğini alışılan istatistiksel yöntemlerden ayıran en belirgin yanı, büyük veriler üzerinde bilgisayar ve internet teknolojilerini kullanılarak, yapılması gereken analizleri çok daha düşük maliyet ve sürelerde yapabilmesidir. Dolayısıyla zaten çok önemli olduğu aşikar olan istatistiksel analizlerin bu denli kolay uygulanabilir olması yüzünden, veri madenciliği bir anda dikkati çekmiştir.

Veri madenciliği, eldeki yapısız veriden, anlamlı ve kullanışlı bilgiyi çıkarmaya yönelik çalışmalarının bütünü olmuştur. Veri madenciliği, veriden bilgi üreterek ortalama kararlar yerine özgün kararlar verilmesini destekleyen, satışları, karlılığı, yenilikçiliği ve kaynak kullanımında etkinliği artıran önemli bir yönetim aracıdır. Organizasyonların veri ambarlarındaki çok önemli bilgilere yoğunlaşmasına yardımcı olan, çok büyük veri tabanlarında saklı, akıllı bilgiyi ortaya çıkaran yeni bir teknolojidir. Veri madenciliği, veriyi toplamak ve yönetmekten çok, analiz yapmayı ve tahmini içermektedir.

Bu tez çalışmasında, veri madenciliği yöntemleri ile istatistiksel eşleştirme yaklaşımı kullanılarak bir uygulama gerçekleştirilmiştir. Uygulama kapsamında, TÜİK'ten temin edilen 2012 GYKA ve 2011 HBA verilerine belirlenen yöntem ayrı ayrı uygulanmış ve sonuçlar özetlenmiştir.

İstatistiksel eşleştirme son yıllarda popüler olmuş bir kavramdır. Tüm veri setleri ihtiyaç olan tüm değişkenlere aynı anda sahip olmayabilir. Bu nedenle iki veya daha fazla veri setinde bulunan ihtiyaç olan değişkenleri tek bir veri setinde birleştirmek gerekmektedir. Eğer veri setlerinde aynı birimler varsa, veri setlerini birleştirmek son derece kolay bir iştir. Ancak veri setlerinde aynı birimler yoksa birleştirme işlemi o kadar basit değildir. Farklı veri setlerinde bulunan farklı birimleri eşleştirip birleştirmek için istatistiksel eşleştirme ortaya çıkmıştır. İstatistiksel eşleştirme için kullanılacak parametrik, parametrik olmayan ve iki yöntemin karışımı olan karma yöntemler literatürde bulunmaktadır. Bu üç yöntemin kendine özgü üstün ve zayıf

yönleri bulunmaktadır. Bu tez çalışmasında her üç yöntem de birlikte kullanılarak üstün veya zayıf oldukları yönlerin değerlendirilmesine imkan sağlanmıştır.

Son üç yıllık HBA verileri birleştirilerek İstatistiki Bölge Birimleri Sınıflaması Düzey-1 bazında açıklanan harcama grupları oransal tahminleri ile tek yıllık GYKA verileri kullanılarak açıklanan yine aynı kapsamdaki tahminler arasında yüksek miktarda bir uyumluluk bulunmaktadır. HBA veri setinde 30 046 hanenin tüketim harcaması bilgileri bulunmakta iken aynı sonuca, kapsamında 17 562 hane bulunan GYKA verileri ile de ulaşılmıştır. HBA veri seti ve kullanılan üç yöntemle oluşturulan modelden elde edilen hane halkı tipine göre harcama gruplarının dikey ve yatay yüzde dağılımları da uygulama kapsamında ele alınmıştır. Bunun yanı sıra HBA verileri kullanılarak oluşturulan modellerin tutarlılığı ve performansı, HBA hanelerinin harcamaları oluşturulan modeller kullanılarak yeniden tahmin edilerek test edilmiştir.

Veri madenciliği uygulamaları hem veri madenciliği uzmanlığı hem de üzerinde çalışılan konuyla ilgili uzmanlık gerektirmektedir. Bu iki uzmanlığı bir arada bulunduracak uzmanları bulmak oldukça güç olup, bu durumun sonucu olarak uygulamada birçok zorluklarla karşılaşmaktadır.

Eğer istenen tüm değişkenleri içeriğinde barındıran bir ana veri kaynağınız yoksa cevaplayıcı birimlere yönelmek maliyetli bir seçim olacaktır. Bu maliyetten kurtulmanın yolu model bazlı yaklaşımlar kullanmaktan geçmektedir. İstatistiksel eşleştirme bu kapsamda uygulanabilecek bir yaklaşımdır. Hangi istatistiksel eşleştirme yaklaşımı kullanılırsa kullanılsın cevaplayıcı birimlere ulaşmaktan daha az maliyetli ve etkili bir yol olacağı kesindir. Gelecek çalışmalarda veri madenciliğinin farklı yöntemleri istatistiksel eşleştirme için uygulanabilir.

Veri madenciliği uygulamalarında mutlaka bilgisayar kullanımı gerekmektedir. Bunun nedeni kullanılan algoritmalarla yapılan hesaplamalar oldukça karmaşık ve fazla sayıdadır. Ayrıca veri madenciliğine konu olan uygulama alanlarında çok fazla veri bulunmaktadır. Bu iki durum bilgisayar kullanımını bir zorunluluk haline getirmiştir. Veri madenciliğinin ilk ortaya çıktığı 90'lı yıllarda bilgisayarlar çok pahalı olduğu için veri madenciliği kullanımının yaygınlaşması gecikmiştir. Günümüzde bilgisayar ve bilgisayar programlarının fiyatları daha kabul edilebilir seviyeye inmiştir. Yine de bu konu veri madenciliği uygulamaları için bir kısıt oluşturmaktadır.

Geleceğin, en azından yakın geleceğin, geçmişten çok fazla farklı olmayacağını varsayarsak geçmiş veriden çıkarılmış olan kurallar gelecekte de geçerli olacak varsayımı gelecek için tahmin yapmamızı sağlayacaktır. Bu tahminler ile birlikte hem operasyonel maliyetlerimiz azalacaktır hem de isabetli kararlar vererek geleceğin bilinmeyen riskleri ile bugünden baş edebilme imkanımız olacaktır. Bu tahminlerin tamamında veri madenciliği yöntemlerinin kullanılması analistlerin ve karar vericilerinin hayatını kolaylaştıracaktır.

KAYNAKLAR

- [1] http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-13-07/EN/KS-RA-13-007-EN.PDF, (Ekim, **2014**).
- [2] Tablo-8: Gelir ve Yaşam Koşulları Araştırması Metodolojisi Hakkında Genel Açıklama, <http://tuik.gov.tr/PreHaberBultenleri.do?id=13594>, (Eylül, **2014**).
- [3] Tablo-3: Hanehalkı Bütçe Anketleri Birleşik Veri Setine İlişkin Açıklamalar, <http://tuik.gov.tr/PreHaberBultenleri.do?id=13147>, (Eylül, **2014**).
- [4] Özdemir, A., Aslay, F.Y., Çam, H., Veritabanında Bilgi Keşfi Süreci: Gümüşhane Devlet Hastanesi Uygulaması, *SÜ İİBF Sosyal ve Ekonomik Araştırmalar Dergisi*, Sayı 20, 347-365, **2010**.
- [5] Skillicorn, D., *Knowledge Discovery for Counterterrorism and Law Enforcement*, Taylor and FrancisGroup, USA, **2009**.
- [6] Khan, A., *Data Warehousing 101 Concepts and Implementation*, The Canton Group, USA, **2003**.
- [7] Uma G.G., *Management Information System: A Managerial Perspective*, USA, West Pub. Co., **1996**.
- [8] Roiger, R.J., Geatz, M.W., *Data Mining : A Tutorial - Based Primer*, Library of Congress Cataloging, USA, **2003**.
- [9] Rob, P., Coronel, C., *Database Sytems: Design, Implementation, and Management Eight Edition*, Thomson Course Technology, USA, **2009**.
- [10] Williams, K., *Using Information for Decision Making*, Chartered Management Institute, India, **1997**.
- [11] Kaya, Y., Tekin, R., *Veritabanı ve Uygulamaları*, Papatya Yayıncılık Eğitim, İstanbul, **2007**.
- [12] Narang, R., *Database Manegement Systems*, Prentice-Hall of India Private Limited, New Delhi, **2004**.
- [13] Kidar, S., *Data Management Systems: First Edition*, Technical Publications Pune, India, **2004**.
- [14] Hammergren, T.C., Simon, A.R., *Data Warehousing for Dummies, 2nd Edition*, Wiley Publishing, Canada, **2009**.
- [15] Cios K.J. vd, *Data Mining Knowledge Discovery Approach*, Springer, USA, **2007**.
- [16] Savaş, S., Topaloğlu, N., Yılmaz, M., Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri", *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, No:21, 5, **2012**.
- [17] Adriaans, P., Zantinge, D., *Data Mining*, Boston, MA, USA Addison Wesley Longman Publishing, **1997**.
- [18] http://tr.wikipedia.org/wiki/Veri_madenciliği, (Eylül, **2014**).

- [19] Ganesh, S., Data Mining: Should it be Included in The Statistics Curriculum?, *The Sixth International Conference on Teaching Statistics*, ICOTS-6, Cape Town, South Africa, 7–12 July, **2002**.
- [20] Oğuzlar, A., *Veri Madenciliğine Giriş*, Ekin Kitabevi, Bursa, **2004**.
- [21] Kuonen, D., Data Mining and Statistics: What is the Connection?, <http://www.statoo.com/en/publications/articleDM4TDAN.pdf>, (Eylül, **2014**).
- [22] Şentürk, A., *Veri Madenciliği Kavram ve Teknikler*, Ekin Kitabevi, Bursa, **2006**.
- [23] Ay, D., *Veri Madenciliği ve Apriori Algoritması İle Süpermarket Analizi*, Yüksek Lisans Tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya, **2009**.
- [24] <http://ugurozmen.com/crm/en-iyi-teklif-sss-3>, (Eylül, **2014**).
- [25] Oğuzlar, A., Veri Ön İşleme, *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, Sayı: 21, 67-76, **2003**.
- [26] Piramuthu, S., Evaluating Feature Selection Methods for Learning in Data Mining Applications, *European Journal of Operational Research*, 156, 483-494, **2004**.
- [27] Argüden, Y., Erşahin, B., *Veri Madenciliği: Veriden Bilgiye, Masraftan Değere*, İstanbul, Arge Danışmanlık Yayınları, **2008**.
- [28] Gorunescu, F., *Data Mining Concepts, Models And Techniques*, Berlin, Springer, **2011**.
- [29] <http://tr.wikipedia.org/wiki/Tahmin>, (Ekim, 2014).
- [30] Asilkan Ö., *Veri Madenciliği Kullanılarak İkinci El Otomobil Pazarında Fiyat Tahmini*, Doktora Tezi, Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü, Antalya, **2008**.
- [31] Özdemir A., Özdemir A., Talep Tahminlemede Kullanılan Yöntemlerin Karşılaştırılması: Seramik Ürün Grubu Firma Uygulaması, *Ege Academic Review*, 105-114, **2006**.
- [32] Zhong, N. - Zhou, L., Methodologies for Knowledge Discovery and Data Mining: *Third Pacific-Asia Conference*, Pakdd-99, Beijing, China, April 26-28, 1999 : Proceedings, Springer Verlag, **1999**.
- [33] Özekes, S., Veri Madenciliği Modelleri ve Uygulama Alanları, *İstanbul Ticaret Üniversitesi Dergisi*, 3, 65-82, **2003**.
- [34] Akpınar, H., Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29, 1-22, **2000**.
- [35] Han, J., Kamber, M., *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, 1st Ed., San Francisco, USA, **2000**.
- [36] Berry, M.J.A., Linoff, G.S., *Mastering Data Mining: The Art and Science of Customer Relationship Management*, John Wiley & Sons, 1st Ed., **1999**.
- [37] Holland, J. H., *Adaption in Natural and Artificial Systems*, MIT Press, Cambridge, MA, **1975**.

- [38] Küçüksille E., *Veri Madenciliği Süreci Kullanılarak Portföy Performansının Değerlendirilmesi ve İMKB Hisse Senetleri Piyasasında Bir Uygulama*, Doktora Tezi, Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta, **2009**.
- [39] Kalogirou S. A., Applications of Artificial Neural-Networks for Energy Systems, *Applied Energy*, 67, 17-35, **2000**.
- [40] Çalış, A., *Veri Madenciliği Yaklaşımı İle Bireysel Müşterilerin Kredi Ödeme Performanslarının Değerlendirilmesi*, Yüksek Lisans Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli, **2013**.
- [41] Yeniay, Ö., An Overview of Genetic Algorithms, *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi-A*, 2, 37-49, **2001**.
- [42] Gülçe G., *Veri Ambarı ve Veri Madenciliği Teknikleri Kullanılarak Öğrenci Karar Destek Sistemi Oluşturma*, Yüksek Lisans Tezi, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli, **2010**.
- [43] Dolgun, M. Ö., *Büyük Alışveriş Merkezleri İçin Veri Madenciliği Uygulamaları*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, **2006**.
- [44] Olgun M. O., Özdemir G., İstatistiksel Özellik Temelli Bayes Sınıflandırıcı Kullanarak Kontrol Grafiklerinde Örüntü Tanıma, *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 27, 303-311, **2012**.
- [45] Akbulut S., *Veri Madenciliği Teknikleri İle Bir Kozmetik Markanın Ayrılan Müşteri Analizi ve Müşteri Segmentasyonu*, Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2006**.
- [46] Sarıman, G., Veri Madenciliğinde Kümeleme Teknikleri Üzerine Bir Çalışma: K-Means ve K-Medoids Kümeleme Algoritmalarının Karşılaştırılması, *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15-3,192-202, **2011**.
- [47] Dinçer, E., *Veri Madenciliğinde K-Means Algoritması ve Tıp Alanında Uygulanması*, Yüksek Lisans Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli, **2006**.
- [48] Arslan, H., *Sakarya Üniversitesi Web Sitesi Erişim Kayıtlarının Web Madenciliği İle Analizi*, Yüksek Lisans Tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya, **2008**.
- [49] Ching W. K., Pong M. K., *Advances in Data Mining and Modeling*, 1st ed., World Scientific, Hong Kong, China, **2002**.
- [50] Çakmak Z., Uzgören, N., Keçek, G., Kümeleme Analizi Teknikleri İle İllerin Kültürel Yapılarına Göre Sınıflandırılması ve Değişimlerinin İncelenmesi, *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 12, 15-36, **2005**.
- [51] Alpaydın, E., Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri, *Bilişim 2000 Eğitim Semineri*, **2000**.

EKLER

- EK 1. HABER BÜLTENİ
- EK 2. İSTATİSTİKİ BÖLGE BİRİMLERİ SINIFLAMASI
- EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ
- EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Haber Bülteni

Sayı: 13147
22/11/2012
10:00

HANEHALKI TÜKETİM HARCAMASI(BÖLGESEL), 2011

Ülkemizde hanehalkı tüketim harcamalarının % 23,3'ü İstanbul'da gerçekleşmektedir

Hanehalkı Bütçe Araştırması verilerinden bölge bazında tahmin üretmek üzere 2009, 2010 ve 2011 yılı Hanehalkı Bütçe Araştırmalarının uygulandığı örnek hanelerin tamamı birleştirilerek 2009-2010-2011 yıllarına ait hanehalkı harcamaları elde edilmiş; 2009 ve 2010 yıllarındaki harcama değerleri 2011 yılının ilgili ayındaki fiyatlarına çekilmiştir.

Bu sonuçlara göre toplam tüketim harcamasının % 23,3'ü İstanbul'da oturan hanehalkları tarafından yapılırken, Ege Bölgesi % 14,5'lik oranla ikinci sırada yer almakta, bunu % 11,6 oranı ile Akdeniz Bölgesi izlemektedir. Harcamaların sadece % 1,8'i Kuzeydoğu Anadolu Bölgesi'ne aittir.

Harcama gruplarının İstatistik Bölge Sınıflaması Düzey-1'e göre dağılımı, 2009-2010-2011 (Dikey %)

BÖLGELER (Düzey 1)	Harcama Grupları												
	Toplam	Gıda ve alkolsüz içecekler	Alkollü içecek, sigara, tütün	Giyim ve ayakkabı	Konut ve kira	Ev eşyası, ev bakım hizmet.	Sağlık	Ulaştırma	Haberleşme	Eğlence ve kültür	Eğitim hizmetleri	Lokanta ve oteller	Çeşitli mal ve hizmetler
Toplam	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
İstanbul	23.3	18.8	20.0	18.9	26.3	20.4	23.8	21.4	23.5	25.3	31.1	26.8	22.9
Batı Marmara	4.8	4.7	5.9	4.8	4.6	4.7	5.1	5.3	5.1	4.3	2.8	5.5	4.5
Ege	14.5	14.5	16.9	16.2	14.4	14.4	15.5	14.0	14.1	16.0	12.6	14.5	13.6
Doğu Marmara	10.1	9.6	10.8	9.5	10.2	10.5	11.4	10.2	10.9	9.9	10.0	10.4	9.6
Batı Anadolu	11.5	10.2	11.3	11.0	11.4	11.7	10.4	13.1	11.5	12.7	14.9	12.5	9.7
Akdeniz	11.6	12.6	10.4	12.4	10.4	12.0	11.1	12.7	11.1	11.2	10.4	10.7	11.6
Orta Anadolu	4.4	5.0	4.9	4.2	4.0	4.6	3.4	5.1	4.1	3.7	4.0	3.1	3.1
Batı Karadeniz	5.2	6.2	5.6	4.8	4.5	6.0	5.3	5.1	5.3	4.6	4.7	4.0	5.7
Doğu Karadeniz	3.0	3.7	2.7	3.1	2.6	3.2	3.2	3.0	3.3	2.5	1.8	2.0	3.8
Kuzeydoğu Anadolu	1.8	2.5	2.5	2.2	1.4	1.9	1.8	1.7	1.7	1.3	1.1	1.1	2.1
Orta Anadolu	3.4	4.0	3.1	5.0	2.8	3.8	3.3	3.1	3.5	3.0	2.3	3.5	4.8
Güneydoğu Anadolu	6.3	8.1	5.9	8.4	5.4	6.7	5.7	5.5	5.8	5.4	4.4	3.9	8.8

EK 1. HABER BÜLTENİ

Gıda ve alkolsüz içecek harcamalarına en fazla payı % 29,7 ile Kuzeydoğu Anadolu Bölgesi ayırıyor

2009-2010-2011 birleşik verilerine göre İstatistik Bölge Birimleri Sınıflaması Düzey-1 bazında bölge içindeki harcama gruplarının dağılımına bakıldığında, Kuzeydoğu Anadolu Bölgesi'nde hanehalklarının tüketim amaçlı yaptığı harcamalar içinde en yüksek payı % 29,7 oranıyla gıda ve alkolsüz içecekler harcaması alırken, aynı bölgede harcamalardaki en düşük payı % 1,1 ile eğitim hizmetleri oluşturmaktadır.

İstanbul Bölgesi'nde ise harcama grupları içerisinde en yüksek payı % 32,6 ile konut ve kira harcamaları alırken, gıda ve alkolsüz içecekler grubu % 17,6 oranıyla ikinci sırada yer almaktadır.

Ulaştırma harcamalarında en yüksek payı Orta Anadolu Bölgesi, en düşük payı ise Güneydoğu Anadolu Bölgesi almaktadır. Haberleşme harcamalarına Güney Doğu Anadolu, konut ve kira ile eğlence ve kültür harcamalarına ise Kuzeydoğu Anadolu en az payı ayıran bölgelerdir. Eğitim harcamalarına en düşük payı ayıran bölgeler Batı Marmara ve Kuzeydoğu Anadolu Bölgesi iken en yüksek payı İstanbul ayırmaktadır.

Harcama gruplarının İstatistik Bölge Sınıflaması Düzey 1'e göre dağılımı, 2009-2010-2011 (Yatay %)

BÖLGELER (Düzey 1)	Toplam	Gıda ve alkolsüz içecekler	Alkollü içecek, sigara, tütün	Giyim ve ayakkabı	Konut ve kira	Ev eşyası, ev bakım hizmet.	Sağlık	Ulaştırma	Haberleşme	Eğlence ve kültür	Eğitim hizmetleri	Lokanta ve oteller	Çeşitli mal ve hizmetler
İstanbul	100,0	17,6	3,6	4,1	32,6	5,5	2,0	14,1	4,2	2,9	2,6	6,7	4,0
Batı Marmara	100,0	21,1	5,2	5,1	25,7	6,1	2,1	16,9	4,4	2,4	1,1	6,2	3,8
Ege	100,0	21,9	5,0	5,4	26,7	6,2	2,1	14,8	4,0	3,0	1,7	5,5	3,8
Doğu Marmara	100,0	20,8	4,6	5,0	27,0	6,5	2,2	15,5	4,4	2,6	1,9	5,6	3,8
Batı Anadolu	100,0	19,4	4,2	4,9	26,7	6,4	1,8	17,5	4,1	3,0	2,5	5,9	3,4
Akdeniz	100,0	23,8	3,8	5,5	24,3	6,5	1,9	16,9	3,9	2,6	1,8	5,0	4,0
Orta Anadolu	100,0	25,1	4,7	4,9	24,7	6,6	1,5	17,9	3,9	2,3	1,8	3,8	2,8
Batı Karadeniz	100,0	26,1	4,5	4,7	23,3	7,2	2,0	15,0	4,2	2,4	1,8	4,2	4,5
Doğu Karadeniz	100,0	26,8	3,8	5,2	23,4	6,9	2,1	15,3	4,6	2,3	1,2	3,5	5,1
Kuzeydoğu Anadolu	100,0	29,7	5,8	6,1	21,2	6,6	2,0	13,9	3,9	1,9	1,1	3,3	4,5
Ortadoğu Anadolu	100,0	25,1	3,8	7,4	22,1	6,9	1,9	13,8	4,1	2,4	1,3	5,5	5,7
Güneydoğu Anadolu	100,0	28,1	4,0	6,8	22,9	6,6	1,8	13,3	3,8	2,3	1,4	3,4	5,6

EK 2. İSTATİSTİKİ BÖLGE BİRİMLERİ SINIFLAMASI

KOD - CODE	DÜZEY 1 - LEVEL 1	DÜZEY 2 - LEVEL 2	DÜZEY 3 - LEVEL 3
TR	TÜRKİYE - TURKEY		
TR1 TR10 TR100	Istanbul	Istanbul	Istanbul
TR2 TR21 TR211 TR212 TR213 TR22 TR221 TR222	Bati Marmara - West Marmara	Tekirdag Balıkesir	 Tekirdag Edirne Kırlareli Balıkesir Canakkele
TR3 TR31 TR310 TR32 TR321 TR322 TR323 TR33 TR331 TR332 TR333 TR334	Ege - Aegean	Izmir Aydın Manisa	 Izmir Aydın Denizli Mugla Manisa Afyon Kutahya Usak
TR4 TR41 TR411 TR412 TR413 TR42 TR421 TR422 TR423 TR424 TR425	Doğu Marmara - East Marmara	Bursa Kocaeli	 Bursa Eskisehir Bilecik Kocaeli Sakarya Düzce Bolu Yalova
TR5 TR51 TR510 TR52 TR521 TR522	Bati Anadolu - West Anatolia	Ankara Konya	 Ankara Konya Karaman
TR6 TR61 TR611 TR612 TR613 TR62 TR621 TR622 TR63 TR631 TR632 TR633	Akdeniz - Mediterranean	Antalya Adana Hatay	 Antalya Isparta Burdur Adana Mersin Hatay Kahramanmaraş Osmaniye
TR7 TR71 TR711 TR712 TR713 TR714 TR715 TR72 TR721 TR722 TR723	Orta Anadolu - Central Anatolia	Kırıkkale Kayseri	 Kırıkkale Aksaray Nigde Nevsehir Kırşehir Kayseri Sivas Yozgat

KOD - CODE	DÜZEY 1 - LEVEL 1	DÜZEY 2 - LEVEL 2	DÜZEY 3 - LEVEL 3
TR8 TR81 TR811 TR812 TR813 TR82 TR821 TR822 TR823 TR83 TR831 TR832 TR833 TR834	Bati Karadeniz - West Black Sea	Zonguldak Kastamonu Samsun	 Zonguldak Karabuk Bartın Kastamonu Cankın Sinop Samsun Tokat Corum Amasya
TR9 TR90 TR901 TR902 TR903 TR904 TR905 TR906	Doğu Karadeniz - East Black Sea	Trabzon	 Trabzon Ordu Giresun Rize Artvin Gümüşhane
TRA TRA1 TRA11 TRA12 TRA13 TRA2 TRA21 TRA22 TRA23 TRA24	Kuzey Doğu Anadolu - North East Anatolia	Erzurum Ağrı	 Erzurum Erzincan Bayburt Ağrı Kars Iğdır Ardahan
TRB TRB1 TRB11 TRB12 TRB13 TRB14 TRB2 TRB21 TRB22 TRB23 TRB24	Orta Doğu Anadolu - Central East Anatolia	Malatya Van	 Malatya Elazığ Bingöl Tunceli Van Mus Bitlis Hakkari
TRC TRC1 TRC11 TRC12 TRC13 TRC2 TRC21 TRC22 TRC3 TRC31 TRC32 TRC33 TRC34	Güney Doğu Anadolu - South East Anatolia	Gaziantep Sanlıurfa Mardin	 Gaziantep Adıyaman Kilis Sanlıurfa Diyarbakır Mardin Batman Sınak Siirt

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA						
Kod	Tanım	Frekans			Kod	Tanım	Frekans				
		Yalın	%	Faktörlü	%		Yalın	%	Faktörlü	%	
HB050	Hanehalkı tipi				HHTPI	Hanehalkı tipi					
5	Tek kişilik hanehalkı	1.499	8.54	1.745.001	8.63	1	Tek çocuklu çekirdek aile	1.820	18.35	3.684.320	19.08
6	Bağımlı çocuk olmayan, iki yetişkinli, yetişkinlerin ikisi de 65 yaşın altında	2.194	12.49	2.502.493	12.38	2	İki çocuklu çekirdek aile	2.045	20.62	4.080.906	21.13
7	Bağımlı çocuk olmayan, iki yetişkinli, yetişkinlerden en az biri 65 ve daha fazla yaşta	1.377	7.84	1.321.373	6.53	3	Üç ve daha fazla çocuklu çekirdek aile	1.479	14.91	2.690.411	13.93
8	Bağımlı çocuk olmayan diğer hanehalkı	1.912	10.89	2.483.138	12.28	4	Çocuksuz çift	1.602	16.15	2.916.257	15.10
9	Tek yetişkinli, en az bir bağımlı çocuk olan hanehalkı	416	2.37	398.641	1.97	5	Ataerkil veya geniş aile	1.483	14.95	3.060.649	15.85
10	İki yetişkinli, bir bağımlı çocuk olan hanehalkı	2.271	12.93	2.854.450	14.12	6	Tek yetişkinli aile	1.305	13.16	2.478.711	12.84
11	İki yetişkinli, iki bağımlı çocuk olan hanehalkı	2.656	15.12	3.279.417	16.22	7	Birarada yaşayan kişiler	184	1.86	400.385	2.07
12	İki yetişkinli, üç ve daha fazla bağımlı çocuk olan hanehalkı	1.910	10.88	1.934.826	9.57						
13	Bağımlı çocuk olan diğer hanehalkı	3.308	18.84	3.676.006	18.18						
16	Diğer (hanehalkı tipi tespit edilemeyenler)	19	0.11	25.232	0.12						
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	
HH010	Oturulan konutun tipi				KONUTTP	Oturulan konutun tipi					
1	Müstakil konut	8.186	46.61	8.485.205	41.96	1	Müstakil konut – Bağımsız	3.568	35.97	6.808.087	35.25
2	Kız ya da sıralı ev	660	3.76	748.222	3.70	2	Müstakil konut – Tek duvara bitişik	916	9.24	1.683.834	8.72
3	Apartman (10 daireden az)	4.456	25.37	5.611.792	27.75	3	Apartman - Bodrum/Zemin	613	6.18	1.225.636	6.35
4	Apartman (10 ve daha fazla daire)	4.252	24.21	5.362.456	26.52	4	Apartman - Normal kat	4.680	47.19	9.293.877	48.13
5	Diğer	8	0.05	12.903	0.06	5	Apartman - Dupleks daire	99	1.00	216.939	1.12
					6	Apartman - Çatı katı	31	0.31	60.273	0.31	
					7	Diğer	11	0.11	22.992	0.12	
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	
HH020	Oturulan konuttaki mülkiyet şekli				MÜLKİYET	Konuta mülkiyet durumu					
1	Ev sahibi	10.794	61.46	11.978.158	59.24	1	Ev sahibi	6.086	61.36	11.626.972	60.21
2	Kiracı	3.592	20.45	4.452.235	22.02	2	Kiracı	2.269	22.88	4.619.627	23.92
3	Lojman	302	1.72	319.240	1.58	3	Lojman	215	2.17	393.524	2.04
4	Diğer (Herhangi bir kira ödemeyen)	2.874	16.36	3.470.945	17.17	4	Diğer	1.348	13.59	2.671.513	13.83
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	
HH030	Oturulan konuta sahip olunan veya oturmaya başlanan yıl				KONUTSAH	Konuta sahip olunan yıl					
1		6.768	38.54	8.242.420	40.76	1		3.832	38.64	7.684.665	39.79
2	-1980	2.160	12.30	2.179.707	10.78	2	-1980	1.130	11.39	2.167.695	11.22
3	1981-1990	1.623	9.24	1.814.883	8.98	3	1981-1990	929	9.37	1.811.075	9.38
4	1991-2000	3.012	17.15	3.378.778	16.71	4	1991-2000	1.744	17.58	3.279.517	16.98
5	2001-2010	3.462	19.71	3.928.291	19.43	5	2001-2010	2.199	22.17	4.197.777	21.74
6	2011-	537	3.06	676.499	3.35	6	2011-	84	0.85	170.907	0.88
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	
HE010	Son 12 ay içinde EV KIRASININ, FAİZLİ BORÇ GERİ ÖDEMESİNİN veya KONUT KREDİSİNİN planlandığı gibi ödenememe durumu				KON_BORC	Konutun borcu olup olmadığı					
1	Evet (Bir kez)	986	5.61	1.266.093	6.26			3.832	38.64	7.684.665	39.79
2	Evet (2 veya daha fazla)	1.921	10.94	2.352.093	11.63	1	Evet	709	7.15	1.390.150	7.20
3	Hayır olmadı	2.373	13.51	2.912.040	14.40	2	Hayır	5.377	54.21	10.236.822	53.01
4	Bu tür ödeme yok	12.282	69.94	13.690.353	67.71						
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	
İtfa edilmiş (HH030 -1)	Konuta oturma süresi				OTUR_SUR	Konuta oturma süresi					
1	0-10	9.131	51.99	10.881.888	53.82	1	0-10	5.497	55.42	10.841.011	56.14
2	11-20	4.057	23.10	4.669.577	23.09	2	11-20	2.297	23.16	4.394.433	22.76
3	21-30	1.964	11.18	2.214.171	10.95	3	21-30	1.072	10.81	2.055.780	10.65
4	31-40	1.382	7.87	1.426.534	7.05	4	31-40	624	6.29	1.205.082	6.24
5	41-50	607	3.46	614.682	3.04	5	41-50	252	2.54	482.542	2.50
6	51-60	297	1.69	298.397	1.48	6	51-60	110	1.11	209.868	1.09
7	61-70	97	0.55	91.736	0.45	7	61-70	41	0.41	79.531	0.41
8	71-80	24	0.14	22.620	0.11	8	71-80	24	0.24	42.805	0.22
9	81-90	2	0.01	1.268	0.01	9	81-90	1	0.01	585	0.00
10	91-100	1	0.01	706	0.00						
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637	100.00	

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA				
Kod	Tanım	Frekans			Kod	Tanım	Frekans		
		Yalın	%	Faktörü			Yalın	%	Faktörü
HH060	Oturulan konuttaki (maufak, banyo ve tuvalet hariç) oda sayısı				ODA_SAY	Oda sayısı			
1	1 oda	242	1.38	269.861	1	1 oda	90	0.91	178.606
2	2 oda	1.697	9.66	1.831.204	2	2 oda	804	8.11	1.592.131
3	3 oda	7.137	40.64	8.479.192	3	3 oda	3.976	40.09	7.924.440
4	4 oda	7.387	42.06	8.447.261	4	4 oda	4.439	44.76	8.466.641
5	5 oda	839	4.78	899.768	5	5 oda	467	4.71	861.962
6	6 oda	187	1.06	208.262	6	6 oda	97	0.98	200.027
7	7 oda	42	0.24	43.985	7	7 oda	30	0.30	59.747
8	8 oda	24	0.14	32.637	8	8 oda	11	0.11	22.258
9	9 oda	4	0.02	5.127	9	9 oda	1	0.01	1.519
10 ve üzeri	10 ve üzeri oda	3	0.02	3.281	10 ve üzeri	10 ve üzeri oda	3	0.03	4.306
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH070	Oturulan konuttaki faydalanılan alanın metrekare olarak büyüklüğü (m2)				KON_ALAN	Konutun büyüklüğü (m2)			
1	0-100	10.611	60.42	12.467.619	1	0-100	5.895	59.44	11.573.398
2	101-200	6.816	38.81	7.601.093	2	101-200	3.926	39.58	7.571.762
3	201-300	126	0.72	140.881	3	201-300	91	0.92	160.128
4	301-400	8	0.05	9.022	4	301-400	3	0.03	4.634
5	401-500	1	0.01	1.963	5	401-500	2	0.02	1.132
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH080	Oturulan konutun ısıtma sistemi				ISIT_SIS	Konutun ısıtma sistemi			
1	Soba (Kömür, gaz, doğalgaz, elektrik)	11.175	63.63	12.012.615	1	Soba	6.097	61.47	11.597.424
2	Kalorifer (Müşterek veya merkezi ısıtma)	1.946	11.08	2.041.502	2	Müşterek veya merkezi ısıtma	1.135	11.44	2.054.173
3	Kalorifer (Kat kaloriferi/kombi)	3.869	22.03	5.428.558	3	Kat kaloriferi/kombi	2.269	22.88	4.928.409
4	Klima	550	3.13	715.326	4	Klima	378	3.81	669.175
5	Diğer	22	0.13	22.578	5	Diğer	39	0.39	62.457
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH090	Oturulan konutta ısıtma amacıyla en fazla kullanılan yakıt türü				YAKIT_TURU	Yakıt türü			
1	Odu	3.090	17.59	3.638.565	1	Odu	4.332	43.68	8.213.654
2	Kömür	7.796	44.39	7.909.788	2	Kömür	1.851	18.66	3.346.346
3	Doğalgaz	4.719	26.87	6.720.327	3	Doğalgaz	2.716	27.38	5.923.928
4	Fuel-oil	71	0.40	85.197	4	Fuel-oil	21	0.21	23.478
5	LPG (Tüpgaz)	11	0.06	12.846	5	Motorin-mazot	17	0.17	26.324
6	Elektrik	945	5.38	1.230.602	6	Gazyağı	2	0.02	3.422
7	Tezek	817	4.65	492.730	7	LPG (tüpgaz)	19	0.19	36.919
8	Diğer	113	0.64	130.524	8	Elektrik	710	7.16	1.340.256
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH100	Oturulan konutta BANYO veya DÜŞ'ün mevcut olma durumu				BANYO	Banyo sahipliği			
1	Evet (Sadece kendi kullanımında)	16.361	93.16	19.107.951	1	Yok	339	3.42	611.538
2	Evet (Paylaşılan)	505	2.88	550.798	2	Var	9.579	96.58	18.700.099
3	Havır	696	3.96	561.829					
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH110	Oturulan konutta TUVALET'in (ev içinde) mevcut olma durumu				TUVALET	Tuvalet Sahipliği			
1	Evet (Sadece kendi kullanımında)	15.233	86.74	18.009.812	1	Yok	1.063	10.72	1.924.765
2	Evet (Paylaşılan)	562	3.20	607.641	2	Var	8.855	89.28	17.386.872
3	Havır	1.767	10.06	1.603.125					
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH120	Oturulan konutta BAĞIMSIZ MUTFAĞIN mevcut olma durumu				MUTFAK	Mutfak Sahipliği			
1	Evet	16.884	96.14	19.625.181	1	Yok	174	1.75	306.047
2	Havır	678	3.86	595.398	2	Var	9.744	98.25	19.005.590
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637
HH130	Oturulan konutta BORULU SU SİSTEMİNİN mevcut olma durumu				SU_SIS	Borulu su sistemi sahipliği			
1	Evet	17.149	97.65	19.880.342	1	Yok	126	1.27	198.401
2	Havır	413	2.35	340.236	2	Var	9.792	98.73	19.113.236
	Toplam	17.562	100.00	20.220.578	100.00	Toplam	9.918	100.00	19.311.637

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA				
Kod	Tanım	Frekans			Kod	Tanım	Frekans		
		Yalın	%	Faktörü			Yalın	%	Faktörü
HH140	Oturulan konutta SICAK SU SİSTEM'inin mevcut olma durumu				SICAKSU	Sıcak su sahipliği			
1	Evet	14.016	79.81	16.966,338	0	Yok	1.693	17.07	3.161,713
2	Hayır	3.546	20.19	3.254,240	1	Var	8.225	82.93	16.149,924
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
HH150	Hanehalkının TELEFON HATT'ına (sabit) sahip olma durumu				TELEFON	Telefon adedi			
1	Evet	8.191	46.64	9.824,827	0	Yok	4.665	47.04	9.236,393
2	Hayır, maddi yetersizlik	3.326	18.94	3.467,164	1	1 adet	5.235	52.78	10.035,049
3	Hayır, diğer nedenler	6.045	34.42	6.928,587	2	2 adet	14	0.14	30,592
					3	3 adet	3	0.03	6,942
					4	4 adet	1	0.01	2,661
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
HH160	Hanehalkının CEP TELEFONU'na sahip olma durumu				CEPTEL	Cep telefonu adedi			
1	Evet	16.389	93.32	19.059,189	0	Yok	608	6.13	1.136,534
2	Hayır, maddi yetersizlik	417	2.37	431,995	1	1 adet	1.963	19.79	3.596,845
3	Hayır, diğer nedenler	756	4.30	729,394	2	2 adet	4.373	44.09	8.584,589
					3	3 adet	1.930	19.46	3.775,927
					4	4 adet	821	8.28	1.692,872
					5	5 adet	168	1.69	387,385
					6	6 adet	37	0.37	83,826
					7	7 adet	12	0.12	34,231
					8	8 adet	2	0.02	3,249
					9	9 adet	4	0.04	16,179
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
HH170	Hanehalkının TELEVIZYON'a sahip olma durumu				TELEVIZ	Televizyon adedi			
1	Evet	17.179	97.82	19.829,886	0	Yok	856	8.63	1.781,359
2	Hayır, maddi yetersizlik	227	1.29	230,667	1	1 adet	7.158	72.17	13.979,358
3	Hayır, diğer nedenler	156	0.89	160,025	2	2 adet	1.651	16.65	3.099,655
					3	3 adet	226	2.28	403,278
					4	4 adet	23	0.23	40,157
					5	5 adet	4	0.04	7,830
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
					LCDTELE	LCD veya Plazma türü televizyon adedi			
					0	Yok	8.190	82.58	15.807,012
					1	1 adet	1.563	15.76	3.170,827
					2	2 adet	145	1.46	290,494
					3	3 adet	14	0.14	31,435
					4	4 adet	4	0.04	6,929
					5	5 adet	1	0.01	2,661
					9	9 adet	1	0.01	2,277
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
HH180	Hanehalkının BILGISAYAR'a sahip olma durumu				BILGISAY	Bilgisayar adedi			
1	Evet	7.277	41.44	9.417,835	0	Yok	5.527	55.73	10.626,254
2	Hayır, maddi yetersizlik	4.978	28.35	5.163,992	1	1 adet	4.182	42.17	8.247,383
3	Hayır, diğer nedenler	5.307	30.22	5.638,751	2	2 adet	170	1.71	357,147
					3	3 adet	31	0.31	64,394
					4	4 adet	7	0.07	14,713
					5	5 adet	1	0.01	1,745
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637
HH190	Hanehalkının İNTERNET'e sahip olma durumu				INTERN	İnternet adedi			
1	Evet	5.593	31.85	7.403,750	0	Yok	6.679	67.34	12.833,334
2	Hayır, maddi yetersizlik	5.375	30.61	5.846,619	1	1 adet	3.225	32.52	6.451,468
3	Hayır, diğer nedenler	6.594	37.55	6.970,209	2	2 adet	14	0.14	26,835
	Toplam	17.562	100.00	20.220,578	100.00	Toplam	9.918	100.00	19.311,637

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA					
Kod	Tanım	Frekans			Kod	Tanım	Frekans			
		Yalın	%	Faktörü			Yalın	%	Faktörü	
HH210	Hanehalkının BUZDOLAB'na sahip olma durumu				BUZDOLAB	Buzdolabı adedi				
1	Evet	17,187	97.86	19,885,003	98.34	0	141	1.42	259,168	1.34
2	Hayır, maddi yetersizlik	322	1.83	281,464	1.39	1	9,690	97.70	18,879,647	97.76
3	Hayır, diğer nedenler	53	0.30	54,111	0.27	2	84	0.85	168,837	0.87
						3	3	0.03	3,985	0.02
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
HH220	Hanehalkının BULAŞIK MAKİNASI'na sahip olma durumu				BULAŞIK	Bulaşık makinesi adedi				
1	Evet	8,236	46.90	10,290,349	50.89	0	5,330	53.74	10,362,098	53.66
2	Hayır, maddi yetersizlik	6,578	37.46	6,817,709	33.72	1	4,582	46.20	8,938,636	46.29
3	Hayır, diğer nedenler	2,748	15.65	3,112,520	15.39	2	6	0.06	10,903	0.06
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
HH200	Hanehalkının OTOMATİK ÇAMAŞIR MAKİNASI'na sahip olma durumu				ÇAMAŞIR	Otomatik çamaşır makinesi adedi				
1	Evet	16,323	92.94	19,079,781	94.36	0	535	5.39	954,545	4.94
2	Hayır, maddi yetersizlik	1,011	5.76	883,568	4.37	1	9,366	94.43	18,324,661	94.89
3	Hayır, diğer nedenler	228	1.30	257,229	1.27	2	17	0.17	32,431	0.17
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
HH230	Hanehalkının KLİMA'ya sahip olma durumu				KLİMA	Klima adedi				
1	Evet	2,694	15.34	3,450,924	17.07	0	8,394	84.63	16,372,104	84.78
2	Hayır, maddi yetersizlik	5,504	31.34	6,557,313	32.43	1	1,266	12.76	2,489,886	12.89
3	Hayır, diğer nedenler	9,364	53.32	10,212,341	50.50	2	181	1.82	312,820	1.62
						3	51	0.51	88,667	0.46
						4	18	0.18	33,914	0.18
						5	7	0.07	12,268	0.06
						6	1	0.01	1,978	0.01
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
HH240	Hanehalkının OTOMOBİL'e (iş amaçları hariç) sahip olma durumu				OTOMOBİL	Otomobil olup olmadığı				
1	Evet	5,676	32.32	6,793,667	33.60	0	6,575	66.29	12,863,606	66.61
2	Hayır, maddi yetersizlik	9,029	51.41	10,295,404	50.92	1	3,343	33.71	6,448,031	33.39
3	Hayır, diğer nedenler	2,857	16.27	3,131,507	15.49					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
HB020	Hanehalkının yaşadığı yer				KIRIKTİKİD	Kır ve kent yerleşim yeri kodu				
1	Kır	5,922	33.72	5,999,665	29.67	1	3,045	30.70	5,767,628	29.87
2	Kent	11,640	66.28	14,220,913	70.33	2	6,873	69.30	13,544,009	70.13
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (HKB)	Hane büyüklüğü				HKB	Hanenin büyüklüğü				
1		1,499	8.54	1,745,001	8.63	1	655	6.60	1,219,157	6.31
2		3,774	21.49	4,045,551	20.01	2	2,024	20.41	3,755,201	19.45
3		3,693	21.03	4,582,373	22.66	3	2,219	22.37	4,466,999	23.13
4		3,879	22.09	4,834,438	23.91	4	2,416	24.36	4,848,517	25.11
5		2,131	12.13	2,499,328	12.36	5	1,299	13.10	2,509,889	13.00
6		1,118	6.37	1,197,233	5.92	6	613	6.18	1,200,910	6.22
7		619	3.52	606,781	3.00	7	298	3.00	563,176	2.92
8		347	1.98	300,109	1.48	8	152	1.53	298,399	1.55
9		191	1.09	165,552	0.82	9	99	1.00	196,909	1.02
10		124	0.71	94,638	0.47	10	50	0.50	86,678	0.45
11		72	0.41	51,934	0.26	11	31	0.31	59,014	0.31
12		36	0.20	34,742	0.17	12	20	0.20	39,857	0.21
13		30	0.17	24,000	0.12	13	12	0.12	20,431	0.11
14		18	0.10	13,600	0.07	14	9	0.09	17,035	0.09
15		11	0.06	8,121	0.04	15	6	0.06	8,437	0.04
16		8	0.05	7,287	0.04	16	6	0.06	11,293	0.06
17		2	0.01	1,721	0.01	17	3	0.03	3,889	0.02
18		2	0.01	3,416	0.02	18	1	0.01	1,571	0.01
19		3	0.02	2,477	0.01	19	1	0.01	650	0.00
20		2	0.01	949	0.00	20	1	0.01	1,359	0.01
21		1	0.01	545	0.00	21	-	0.00	-	0.00
22		1	0.01	351	0.00	22	-	0.00	-	0.00
23		-	0.00	-	0.00	23	2	0.02	1,701	0.01
24		-	0.00	-	0.00	24	1	0.01	585	0.00
25		1	0.01	429	0.00	25	-	0.00	-	0.00
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA					
Kod	Tanım	Frekans			Kod	Tanım	Frekans			
		Yalın	%	Faktörü			Yalın	%	Faktörü	
Türetilmiş (EFB OECD)	Hane büyüklüğü				EFB OECD	Hane büyüklüğü				
1	1.0-2.0	8,986	51.17	10,389,946	51.38	1.0-2.0	4,906	49.47	9,456,078	48.97
2	2.1-3.0	6,447	36.71	7,742,036	38.29	2.1-3.0	3,918	39.50	7,708,313	39.92
3	3.1-4.0	1,556	8.86	1,580,776	7.82	3.1-4.0	814	8.21	1,586,307	8.21
4	4.1-5.0	401	2.28	364,092	1.80	4.1-5.0	192	1.94	398,911	2.07
5	5.1-6.0	114	0.65	98,454	0.49	5.1-6.0	59	0.59	115,076	0.60
6	6.1-7.0	40	0.23	30,999	0.15	6.1-7.0	14	0.14	22,062	0.11
7	7.1-8.0	10	0.06	8,178	0.04	7.1-8.0	11	0.11	21,266	0.11
8	8.1-	8	0.05	6,097	0.03	8.1-	4	0.04	3,624	0.02
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (K COÇUK)	Beş yaşından küçük çocuk sayısı				Türetilmiş (K COÇUK)	Beş yaşından küçük çocuk sayısı				
0	0	13,133	74.78	15,245,866	75.40	0	7,572	76.35	14,266,635	73.88
1	1	3,349	19.07	3,872,241	19.15	1	1,878	18.94	4,014,067	20.79
2	2	894	5.09	940,701	4.65	2	404	4.07	906,461	4.69
3	3	143	0.81	125,794	0.62	3	48	0.48	94,935	0.49
4	4	35	0.20	27,607	0.14	4	11	0.11	24,348	0.13
5	5	6	0.03	4,285	0.02	5	1	0.01	1,316	0.01
6	6	1	0.01	705	0.00	6	4	0.04	3,876	0.02
7	7	1	0.01	3,381	0.02					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (COÇUK)	Dört yaş ile 18 yaş arasındaki çocuk sayısı				Türetilmiş (COÇUK)	Dört yaş ile 18 yaş arasındaki çocuk sayısı				
0	0	9,318	53.06	11,055,598	54.67	0	5,024	50.66	10,218,722	52.91
1	1	3,746	21.33	4,479,951	22.16	1	2,257	22.76	4,503,472	23.32
2	2	2,714	15.45	3,102,912	15.35	2	1,648	16.62	2,980,228	15.43
3	3	1,012	5.76	988,101	4.89	3	580	5.85	968,966	5.02
4	4	406	2.31	337,736	1.67	4	231	2.33	387,364	2.01
5	5	212	1.21	156,251	0.77	5	99	1.00	142,269	0.74
6	6	99	0.56	63,016	0.31	6	47	0.47	63,616	0.33
7	7	39	0.22	23,695	0.12	7	19	0.19	32,430	0.17
8	8	9	0.05	9,328	0.05	8	5	0.05	7,010	0.04
9	9	2	0.01	965	0.00	9	4	0.04	3,774	0.02
10	10	1	0.01	335	0.00	10	2	0.02	2,551	0.01
11	11	2	0.01	1,912	0.01	11	-	0.00	-	0.00
13	13	1	0.01	351	0.00	13	2	0.02	1,235	0.01
14	14	1	0.01	429	0.00					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (ERKEK CALYAS)	17 yaş ile 40 yaş arasındaki erkek sayısı				Türetilmiş (ERKEK CALYAS)	17 yaş ile 40 yaş arasındaki erkek sayısı				
0	0	8,776	49.97	9,242,628	45.71	0	4,883	49.23	8,606,450	44.57
1	1	7,620	43.39	9,431,107	46.64	1	4,404	44.40	9,127,003	47.26
2	2	923	5.26	1,227,920	6.07	2	520	5.24	1,265,592	6.55
3	3	200	1.14	270,257	1.34	3	88	0.89	242,976	1.26
4	4	30	0.17	38,276	0.19	4	19	0.19	53,615	0.28
5	5	10	0.06	6,477	0.03	5	4	0.04	16,001	0.08
6	6	2	0.01	1,470	0.01					
7	7	1	0.01	2,442	0.01					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (ORTA_ERKEK)	39 yaş ile 66 yaş arasındaki erkek sayısı				Türetilmiş (ORTA_ERKEK)	39 yaş ile 66 yaş arasındaki erkek sayısı				
0	0	8,976	51.11	10,370,764	51.29	0	4,926	49.67	9,731,860	50.39
1	1	8,512	48.47	9,772,994	48.33	1	4,854	49.95	9,499,680	49.19
2	2	69	0.39	71,560	0.35	2	37	0.37	78,488	0.41
3	3	5	0.03	5,261	0.03	3	1	0.01	1,609	0.01
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (YASLI_ERKEK)	65 yaşından büyük erkek sayısı				Türetilmiş (YASLI_ERKEK)	65 yaşından büyük erkek sayısı				
0	0	15,422	87.81	18,093,035	89.48	0	8,766	88.38	17,302,383	89.60
1	1	2,134	12.15	2,123,396	10.50	1	1,149	11.58	2,004,395	10.38
2	2	6	0.03	4,147	0.02	2	3	0.03	4,859	0.03
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA					
Kod	Tanım	Frekans			Kod	Tanım	Frekans			
		Yalın	%	Faktörü		Yalın	%	Faktörü		
Türetmiş (KADIN_CALYAS)	17 yaş ile 40 yaş arasındaki kadın sayısı				Türetmiş (KADIN_CALYAS)	17 yaş ile 40 yaş arasındaki kadın sayısı				
0	0	7.901	44.99	8.659,156	42.82	0	4.247	42.82	7.896,155	40.89
1	1	8.632	49.15	10.322,072	51.05	1	5.052	50.94	10.163,924	52.63
2	2	860	4.90	1.038,746	5.14	2	509	5.13	1.028,987	5.33
3	3	125	0.71	161,537	0.80	3	81	0.82	158,904	0.82
4	4	32	0.18	30,822	0.15	4	23	0.23	50,142	0.26
5	5	6	0.03	4,922	0.02	5				
6	6	4	0.02	2,267	0.01	6		0.06	13,524	0.07
7	7	2	0.01	1,056	0.01					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (ORTA_KADIN)	39 yaş ile 66 yaş arasındaki kadın sayısı				Türetmiş (ORTA_KADIN)	39 yaş ile 66 yaş arasındaki kadın sayısı				
0	0	8.792	50.06	10.378,574	51.33	0	4.848	48.88	9.701,818	50.24
1	1	8.591	48.92	9.662,132	47.78	1	4.968	50.09	9.412,463	48.74
2	2	169	0.96	169,183	0.84	2	97	0.98	189,513	0.98
3	3	10	0.06	10,689	0.05	3	5	0.05	7,844	0.04
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (YASLI_KADIN)	65 yaşından büyük kadın sayısı				Türetmiş (YASLI_KADIN)	65 yaşından büyük kadın sayısı				
0	0	14.723	83.83	17.433,828	86.22	0	8.563	86.34	16.709,659	86.53
1	1	2.813	16.02	2.763,573	13.87	1	1.343	13.54	2.577,671	13.35
2	2	26	0.15	23,177	0.11	2	12	0.12	24,308	0.13
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (ISSIZ)	İssiz fert sayısı				Türetmiş (ISSIZ)	İssiz fert sayısı				
0	0	16.181	92.14	18.537,721	91.68	0	9.014	90.89	17.492,470	90.58
1	1	1.237	7.04	1.505,085	7.44	1	778	7.84	1.554,599	8.05
2	2	131	0.75	168,306	0.83	2	104	1.05	222,325	1.15
3	3	8	0.05	6,409	0.03	3	19	0.19	34,872	0.18
4	4	2	0.01	694	0.00	4	3	0.03	7,371	0.04
5	5	3	0.02	2,363	0.01					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (ISTHDAM)	İstihdamda yer alan fert sayısı				Türetmiş (ISTHDAM)	İstihdamda yer alan fert sayısı				
0	0	4.103	23.36	4.170,214	20.62	0	1.930	19.46	3.359,373	17.40
1	1	7.446	42.40	8.790,985	43.48	1	4.320	43.56	8.379,197	43.39
2	2	4.382	24.95	5.336,523	26.39	2	2.717	27.39	5.471,491	28.33
3	3	1.101	6.27	1.313,509	6.50	3	663	6.68	1.446,688	7.49
4	4	368	2.10	447,396	2.21	4	204	2.06	456,938	2.37
5	5	104	0.59	115,685	0.57	5	59	0.59	149,925	0.78
6	6	37	0.21	29,712	0.15	6	13	0.13	25,263	0.13
7	7	11	0.06	8,810	0.04	7	8	0.08	17,604	0.09
8	8	7	0.04	4,628	0.02	8	3	0.03	3,186	0.02
9	9	2	0.01	2,693	0.01	9	1	0.01	1,970	0.01
10	10		0.01	422	0.00					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00
Türetmiş (ISGDAHILOLM)	İşgücüne dahil olmayan fert sayısı				Türetmiş (ISGDAHILOLM)	İşgücüne dahil olmayan fert sayısı				
0	0	3.455	19.67	4.248,359	21.01	0	2.118	21.36	4.321,701	22.38
1	1	7.338	41.78	8.627,578	42.67	1	4.156	41.90	8.140,193	42.15
2	2	4.621	26.31	5.006,651	24.76	2	2.519	25.40	4.671,900	24.19
3	3	1.501	8.55	1.653,744	8.18	3	835	8.42	1.598,375	8.28
4	4	423	2.41	461,959	2.28	4	215	2.17	419,731	2.17
5	5	137	0.78	139,792	0.69	5	52	0.52	110,710	0.57
6	6	59	0.34	56,450	0.28	6	15	0.15	34,261	0.18
7	7	17	0.10	17,625	0.09	7	2	0.02	6,027	0.03
8	8	8	0.05	4,883	0.02	8	2	0.02	3,640	0.02
9	9	2	0.01	3,108	0.02	9	3	0.03	3,495	0.02
10	10		0.00	-	0.00	10	1	0.01	1,604	0.01
11	11	1	0.01	429	0.00					
	Toplam	17,562	100.00	20,220,578	100.00	Toplam	9,918	100.00	19,311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA					
Kod	Tanım	Frekans			Kod	Tanım	Frekans			
		Yalın	%	Faktörü			Yalın	%	Faktörü	
Türetmiş (ISDUR 1)					Türetmiş (ISDUR 1)					
0	Esas işindeki durumu ücretli, maaşlı olan fert sayısı	8.934	50,87	9.318,878	46,09	0	4.804	48,44	8.917,786	46,18
1		6.344	36,12	7.759,074	38,37	1	3.738	37,69	7.393,803	38,29
2		1.996	11,37	2.727,294	13,49	2	1.200	12,10	2.556,376	13,24
3		252	1,43	368,335	1,82	3	150	1,51	370,039	1,92
4		30	0,17	37,829	0,19	4	23	0,23	63,577	0,33
5		5	0,03	8,572	0,04	5	3	0,03	10,056	0,05
6		1	0,01	595	0,00					
7										
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00
Türetmiş (ISDUR 2)					Türetmiş (ISDUR 2)					
0	Esas işindeki durumu yevmiyeli olan fert sayısı	15.817	90,06	18.245,205	90,23	0	8.552	86,23	16.619,006	86,06
1		1.504	8,56	1.698,718	8,40	1	1.157	11,67	2.286,248	11,84
2		202	1,15	224,326	1,11	2	175	1,76	346,488	1,79
3		31	0,18	40,267	0,20	3	28	0,28	49,538	0,26
4		5	0,03	7,416	0,04	4	5	0,05	9,675	0,05
5		1	0,01	1,401	0,01	5	1	0,01	682	0,00
6		1	0,01	1,429	0,01					
7		1	0,01	1,815	0,01					
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00
Türetmiş (ISDUR 3)					Türetmiş (ISDUR 3)					
0	Esas işindeki durumu işveren olan fert sayısı	16.750	95,38	19.171,244	94,81	0	9.464	95,42	18.403,618	95,30
1		789	4,49	1.018,958	5,04	1	440	4,44	884,259	4,58
2		23	0,13	30,377	0,15	2	14	0,14	23,760	0,12
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00
Türetmiş (ISDUR 4)					Türetmiş (ISDUR 4)					
0	Esas işindeki durumu kendi hesabına olan fert sayısı	13.462	76,65	15.801,059	78,14	0	7.488	75,50	14.535,770	75,27
1		3.980	22,66	4.281,057	21,17	1	2.324	23,43	4.559,892	23,61
2		114	0,65	131,413	0,65	2	102	1,03	210,903	1,09
3		4	0,02	3,847	0,02	3	3	0,03	2,954	0,02
4		2	0,01	3,202	0,02	4	1	0,01	2,117	0,01
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00
Türetmiş (ISDUR 5)					Türetmiş (ISDUR 5)					
0	Esas işindeki durumu ücretsiz aile işçisi olan fert sayısı	15.088	85,91	17.704,390	87,56	0	8.714	87,86	16.840,691	87,20
1		1.757	10,00	1.792,586	8,87	1	868	8,75	1.746,721	9,04
2		462	2,63	460,080	2,28	2	229	2,31	491,973	2,55
3		174	0,99	187,083	0,93	3	72	0,73	159,237	0,82
4		53	0,30	54,859	0,27	4	20	0,20	42,648	0,22
5		16	0,09	12,345	0,06	5	12	0,12	26,633	0,14
6		4	0,02	2,416	0,01	6	2	0,02	2,632	0,01
7		7	0,04	6,188	0,03	7	1	0,01	1,102	0,01
8		1	0,01	632	0,00					
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00
Türetmiş (SONOKUL0)					Türetmiş (SONOKUL0)					
0	Okur-yazar olmayan fert sayısı	12.742	72,55	15.480,654	76,56	0	7.580	76,43	14.787,689	76,57
1		3.717	21,17	3.768,454	18,64	1	1.849	18,64	3.586,137	18,57
2		896	5,10	789,187	3,90	2	377	3,80	738,243	3,82
3		151	0,86	128,708	0,64	3	79	0,80	138,690	0,72
4		41	0,23	36,248	0,18	4	20	0,20	35,838	0,19
5		8	0,05	10,934	0,05	5	9	0,09	16,380	0,08
6		5	0,03	5,632	0,03	6	3	0,03	7,621	0,04
7		1	0,01	429	0,00	7	1	0,01	1,039	0,01
8		1	0,01	334	0,00					
	Toplam	17.562	100,00	20.220,578	100,00		9.918	100,00	19.311,637	100,00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA						
Kod	Tanım	Frekans				Kod	Tanım	Frekans			
		Yalın	%	Faktörlü	%			Yalın	%	Faktörlü	%
Türetilmiş (SONOKUL1)					Türetilmiş (SONOKUL1)						
0	Okur-yazar olup, bir okul bitirmeyen fert sayısı	14.242	81.10	16.913,723	83.65	0	Okur-yazar olup, bir okul bitirmeyen fert sayısı	8.342	84.11	16.273,860	84.27
1		2.734	15.57	2.773,186	13.71	1		1.280	12.91	2.454,656	12.71
2		454	2.59	416,759	2.06	2		223	2.25	446,034	2.31
3		96	0.55	88,147	0.44	3		49	0.49	88,943	0.46
4		23	0.13	15,155	0.07	4		17	0.17	36,727	0.19
5		11	0.06	12,841	0.06	5		5	0.05	8,663	0.04
6		2	0.01	767	0.00	6		2	0.02	2,754	0.01
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL2)					Türetilmiş (SONOKUL2)						
0	İlkokul mezunu fert sayısı	6.708	38.20	7,717,683	38.17	0	İlkokul mezunu fert sayısı	3.563	35.92	6,938,863	35.93
1		6.270	35.70	6,953,714	34.39	1		3.471	35.00	6,608,026	34.22
2		4.113	23.42	4,964,324	24.55	2		2.586	26.07	5,067,956	26.24
3		370	2.11	465,908	2.30	3		239	2.41	556,298	2.88
4		85	0.48	99,025	0.49	4		52	0.52	123,564	0.64
5		12	0.07	15,325	0.08	5		4	0.04	9,845	0.05
6		2	0.01	868	0.00	6		3	0.03	7,084	0.04
7		1	0.01	3,411	0.02						
8		1	0.01	320	0.00						
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL3)					Türetilmiş (SONOKUL3)						
0	Ortaokul, mesleki ortaokul ve ilköğretim mezunu fert sayısı	10.984	62.54	12,514,384	61.89	0	Ortaokul, mesleki ortaokul ve ilköğretim mezunu fert sayısı	6.205	62.56	12,029,934	62.29
1		4.701	26.77	5,504,722	27.22	1		2.648	26.70	5,160,189	26.72
2		1.506	8.58	1,785,838	8.83	2		869	8.76	1,708,517	8.85
3		300	1.71	340,768	1.69	3		152	1.53	320,619	1.66
4		56	0.32	55,832	0.28	4		36	0.36	79,475	0.41
5		11	0.06	13,896	0.07	5		5	0.05	8,729	0.05
6		4	0.02	5,139	0.03	6		3	0.03	3,173	0.02
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL4)					Türetilmiş (SONOKUL4)						
0	Genel lise mezunu fert sayısı	14.040	79.95	15,807,739	78.18	0	Genel lise mezunu fert sayısı	7.681	77.45	14,806,200	76.67
1		2.928	16.67	3,640,458	18.00	1		1.814	18.29	3,614,636	18.72
2		507	2.89	646,511	3.20	2		366	3.69	759,191	3.93
3		70	0.40	104,079	0.51	3		46	0.46	104,930	0.54
4		13	0.07	18,920	0.09	4		10	0.10	22,805	0.12
5		4	0.02	2,872	0.01	5		1	0.01	3,874	0.02
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL5)					Türetilmiş (SONOKUL5)						
0	Mesleki veya teknik lise mezunu fert sayısı	14.613	83.21	16,500,050	81.80	0	Mesleki veya teknik lise mezunu fert sayısı	8.199	82.67	15,873,119	82.19
1		2.545	14.49	3,189,689	15.77	1		1.492	15.04	2,967,913	15.37
2		379	2.16	498,906	2.47	2		209	2.11	430,786	2.23
3		23	0.13	28,320	0.14	3		15	0.15	33,765	0.17
4		2	0.01	3,613	0.02	4		3	0.03	6,053	0.03
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL6)					Türetilmiş (SONOKUL6)						
0	Yüksek okul, fakülte ve üzeri mezun fert sayısı	14.142	80.53	15,935,267	78.81	0	Yüksek okul, fakülte ve üzeri mezun fert sayısı	7.862	79.27	15,226,263	78.85
1		2.280	12.98	2,821,215	13.95	1		1.380	13.91	2,706,986	14.02
2		1,061	6.04	1,342,663	6.64	2		628	6.33	1,273,969	6.59
3		70	0.40	105,100	0.52	3		42	0.42	94,351	0.49
4		8	0.05	15,738	0.08	4		6	0.06	10,666	0.06
7		1	0.01	595	0.00						
	Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA						
Kod	Tanım	Frekans				Kod	Tanım	Frekans			
		Yalın	%	Faktörlü	%			Yalın	%	Faktörlü	%
HH040	Oturulan konut için ödenen aylık kira tutarı (TL)					KIRA_AY	Aylık ödenen kira tutarı (TL)				
1		13.668	77.83	15.449,103	76.40	1		7.434	74.95	14.298,485	74.04
2	0-100	234	1.33	241.344	1.19	2	0-100	214	2.16	369,253	1.91
3	101-200	666	3.79	673.319	3.33	3	101-200	463	4.67	795,956	4.12
4	201-300	1.091	6.21	1.183,598	5.85	4	201-300	672	6.78	1.312,386	6.80
5	301-400	914	5.20	1.136,871	5.62	5	301-400	572	5.77	1.202,352	6.23
6	401-500	502	2.86	723,802	3.58	6	401-500	310	3.13	713,961	3.70
7	501-600	230	1.31	374,781	1.85	7	501-600	124	1.25	282,793	1.46
8	601-700	99	0.56	161,166	0.80	8	601-700	49	0.49	123,145	0.64
9	701-800	59	0.34	91,855	0.45	9	701-800	19	0.19	46,198	0.24
10	801-900	27	0.15	49,473	0.24	10	801-900	13	0.13	35,626	0.18
11	901-1000	20	0.11	32,686	0.16	11	901-1000	13	0.13	34,141	0.18
12	1001-2000	46	0.26	89,783	0.44	12	1001-2000	29	0.29	81,278	0.42
13	2001-3000	6	0.03	12,819	0.06	13	2001-3000	3	0.03	7,769	0.04
						14	3001-4000	1	0.01	2,642	0.01
						15	4001-5000	2	0.02	5,650	0.03
	Toplam	17.562	100.00	20.220,578	100.00		Toplam	9.918	100.00	19.311,637	100.00
HC010	Oturulan konutun yıllık izafi kirası (TL)					IZAFI_K*12	izafi kira (TL)				
1		3.894	22.17	4.771,475	23.60	1		2.269	22.88	5.013,152	25.96
2	201-300	-	0.00	-	0.00	2	201-300	2	0.02	3,607	0.02
3	301-400	-	0.00	-	0.04	3	301-400	8	0.08	14,703	0.08
4	401-500	1	0.01	369	0.05	4	401-500	10	0.10	20,886	0.11
5	501-600	5	0.03	7,677	0.10	5	501-600	297	2.99	565,377	2.93
6	601-700	12	0.07	10,490	0.37	6	601-700	-	0.00	-	0.00
7	701-800	25	0.14	19,751	0.62	7	701-800	30	0.30	54,077	0.28
8	801-900	84	0.48	74,842	18.32	8	801-900	41	0.41	90,296	0.47
9	901-1000	163	0.93	125,952	16.07	9	901-1000	28	0.28	49,650	0.26
10	1001-2000	3.848	21.91	3.704,507	10.12	10	1001-2000	1.391	14.03	2.564,966	13.28
11	2001-3000	3.042	17.32	3.248,596	10.28	11	2001-3000	1.845	18.60	3.287,174	17.02
12	3001-4000	1.809	10.30	2.046,856	19.28	12	3001-4000	883	8.90	1.583,642	8.20
13	4001-5000	1.784	10.16	2.079,453	1.08	13	4001-5000	1.396	14.08	2.573,503	13.33
14	5001-10000	2.761	15.72	3.898,962	0.07	14	5001-10000	1.497	15.09	2.994,416	15.51
15	10001-20000	128	0.73	218,399	0.00	15	10001-20000	185	1.87	417,645	2.16
16	20001-30000	6	0.03	13,248	0.00	16	20001-30000	24	0.24	57,282	0.30
						17	30001-40000	5	0.05	6,985	0.04
						18	40001-50000	1	0.01	2,092	0.01
						19	50001-60000	4	0.04	6,665	0.03
						20	60001-70000	2	0.02	5,520	0.03
	Toplam	17.562	100.00	20.220,578	100.00		Toplam	9.918	100.00	19.311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA						
Kod	Tanım	Frekans				Kod	Tanım	Frekans			
		Yalın	%	Faktör	%			Yalın	%	Faktör	%
HG070	Gayrimenkul (kira) geliri (TL)					(MUSKDEG * 12) + (APTDEG * 12) + (YAZKDEG * 12) + (ARSAKDEG * 12) + (DUKKDEG * 12) + (OTELKDEG * 12) + (TARKDEG * 12) + (BAGKDEG * 12) + (SERAKDEG * 12)	Gayrimenkul (kira) geliri (TL)				
1	0-100	15,135	86.18	17,408,509	86.03	1	0-100	8,495	85.65	16,691,218	86.43
2	101-200	71	0.40	76,889	0.38	2	101-200	30	0.30	57,800	0.30
3	201-300	63	0.36	63,513	0.31	3	201-300	18	0.18	32,978	0.17
4	301-400	45	0.26	43,518	0.22	4	301-400	13	0.13	20,798	0.11
5	401-500	52	0.30	57,559	0.28	5	401-500	11	0.11	17,286	0.09
6	501-600	45	0.26	43,567	0.22	6	501-600	21	0.21	36,851	0.19
7	601-700	29	0.17	31,776	0.16	7	601-700	3	0.03	7,290	0.04
8	701-800	29	0.17	33,692	0.17	8	701-800	9	0.09	14,853	0.08
9	801-900	25	0.14	27,802	0.14	9	801-900	8	0.08	14,724	0.08
10	901-1000	42	0.24	39,628	0.20	10	901-1000	15	0.15	28,317	0.15
11	1001-2000	313	1.78	340,977	1.69	11	1001-2000	143	1.44	250,331	1.30
12	2001-3000	374	2.13	399,115	1.97	12	2001-3000	240	2.42	427,403	2.21
13	3001-4000	296	1.69	323,995	1.60	13	3001-4000	198	2.00	371,622	1.92
14	4001-5000	297	1.69	354,398	1.75	14	4001-5000	183	1.85	349,452	1.81
15	5001-10000	482	2.74	612,613	3.03	15	5001-10000	307	3.10	556,078	2.88
16	10001-20000	165	0.94	228,684	1.13	16	10001-20000	136	1.37	270,560	1.40
17	20001-30000	56	0.32	72,528	0.36	17	20001-30000	40	0.40	77,155	0.40
18	30001-40000	14	0.08	17,893	0.09	18	30001-40000	13	0.13	23,626	0.12
19	40001-50000	9	0.05	14,531	0.07	19	40001-50000	10	0.10	17,539	0.09
20	50001-60000	5	0.03	7,163	0.04	20	50001-60000	10	0.10	19,939	0.10
21	60001-70000	5	0.03	7,424	0.04	21	60001-70000	4	0.04	6,224	0.03
22	70001-80000	3	0.02	4,993	0.02	22	70001-80000	3	0.03	7,071	0.04
23	80001-90000	-	0.00	-	0.00	23	80001-90000	3	0.03	6,296	0.03
24	90001-100000	2	0.01	2,874	0.01	24	90001-100000	1	0.01	1,250	0.01
25	100001-200000	5	0.03	6,935	0.03	25	100001-200000	3	0.03	2,855	0.01
	Toplam	17,562	100.00	20,220,578	100.00	26	200001-300000	1	0.01	2,119	0.01
HG110	Toplam hanehalkı kullanılabilir geliri (TL)					YILKULGEL	Yıllık hanehalkı kullanılabilir geliri (TL)				
1	0-100	-	0.00	-	0.00	1	0-100	14	0.14	24,032	0.12
2	101-200	1	0.01	717	0.00	2	101-200	1	0.01	1,455	0.01
3	201-300	4	0.02	6,264	0.03	3	201-300	-	0.00	-	0.00
4	301-400	-	0.00	-	0.00	4	301-400	1	0.01	987	0.01
5	401-500	-	0.00	-	0.00	5	401-500	1	0.01	2,093	0.01
6	501-600	1	0.01	715	0.00	6	501-600	1	0.01	2,125	0.01
7	601-700	1	0.01	333	0.00	7	601-700	1	0.01	1,526	0.01
8	701-800	2	0.01	1,358	0.01	8	701-800	2	0.02	2,026	0.01
9	901-1000	2	0.01	1,347	0.01	9	901-1000	1	0.01	1,295	0.01
10	1001-2000	10	0.06	12,117	0.06	10	1001-2000	11	0.11	20,852	0.11
11	2001-3000	90	0.51	88,368	0.44	11	2001-3000	46	0.46	85,419	0.44
12	3001-4000	139	0.79	136,101	0.67	12	3001-4000	52	0.52	94,345	0.49
13	4001-5000	232	1.32	220,945	1.09	13	4001-5000	97	0.98	182,839	0.95
14	5001-10000	2,203	12.54	2,128,560	10.53	14	5001-10000	956	9.64	1,762,128	9.12
15	10001-20000	6,521	37.13	7,170,077	35.46	15	10001-20000	3,495	35.24	6,617,722	34.27
16	20001-30000	4,001	22.78	4,813,905	23.81	16	20001-30000	2,372	23.92	4,652,621	24.09
17	30001-40000	1,979	11.27	2,513,421	12.43	17	30001-40000	1,311	13.22	2,617,087	13.55
18	40001-50000	941	5.36	1,229,712	6.08	18	40001-50000	608	6.13	1,264,456	6.55
19	50001-60000	543	3.09	701,571	3.47	19	50001-60000	360	3.63	751,817	3.89
20	60001-70000	317	1.81	402,752	1.99	20	60001-70000	198	2.00	423,222	2.19
21	70001-80000	168	0.96	218,584	1.08	21	70001-80000	119	1.20	250,050	1.29
22	80001-90000	111	0.63	143,804	0.71	22	80001-90000	77	0.78	165,909	0.86
23	90001-100000	71	0.40	95,665	0.47	23	90001-100000	33	0.33	76,155	0.39
24	100001-200000	195	1.11	282,017	1.39	24	100001-200000	132	1.33	254,727	1.32
25	200001-300000	22	0.13	37,863	0.19	25	200001-300000	21	0.21	44,788	0.23
26	300001-400000	5	0.03	10,281	0.05	26	300001-400000	5	0.05	8,918	0.05
27	400001-500000	3	0.02	4,100	0.02	27	400001-500000	1	0.01	1,893	0.01
	Toplam	17,562	100.00	20,220,578	100.00	28	600001-700000	1	0.01	566	0.00
						29	700001-800000	1	0.01	582	0.00
							Toplam	9,918	100.00	19,311,637	100.00

EK 3. GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

GYKA					HBA							
Kod	Tanım	Frekans				Kod	Tanım	Frekans				
HH050	Aylık ortalama konut masrafı (TL)	Yalın	%	Faktörlü	%	ANAGRUP_4 - ZAFI_K	Harcama değeri (TL)	Yalın	%	Faktörlü	%	
1	0-100	3,342	19.03	3,425,573	16.94	1	0-100	2,041	20.58	3,989,016	20.66	
2	101-200	5,705	32.48	6,351,272	31.41	2	101-200	2,675	26.97	5,136,092	26.60	
3	201-300	3,550	20.21	4,201,297	20.78	3	201-300	1,827	18.42	3,436,544	17.80	
4	301-400	1,498	8.53	1,691,330	8.36	4	301-400	1,155	11.65	2,189,154	11.34	
5	401-500	1,085	6.18	1,267,434	6.27	5	401-500	765	7.71	1,516,336	7.85	
6	501-600	841	4.79	1,041,528	5.15	6	501-600	517	5.21	1,047,533	5.42	
7	601-700	600	3.42	815,891	4.03	7	601-700	376	3.79	793,128	4.11	
8	701-800	349	1.99	511,998	2.53	8	701-800	210	2.12	456,084	2.36	
9	801-900	220	1.25	321,567	1.59	9	801-900	126	1.27	257,019	1.33	
10	901-1000	119	0.68	189,510	0.94	10	901-1000	53	0.53	119,794	0.62	
11	1001-2000	232	1.32	357,114	1.77	11	1001-2000	154	1.55	328,037	1.70	
12	2001-3000	18	0.10	40,657	0.20	12	2001-3000	12	0.12	27,723	0.14	
13	3001-4000	2	0.01	3,053	0.02	13	3001-4000	4	0.04	6,885	0.04	
14	4001-5000	1	0.01	2,356	0.01	14	4001-5000	-	0.00	-	0.00	
						15	5001-10000	3	0.03	8,292	0.04	
		Toplam	17,562	100.00	20,220,578	100.00		Toplam	9,918	100.00	19,311,637	100.00

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans					HBA Frekans				Hellingering Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%	
HB050	Hanehalkı tipi										2.5164
1	Tek çocuklu çekirdek aile	3,011	17.14	3,851,063	3,677,953	19.05	1,820	18.35	3,684,320	19.08	
2	İki çocuklu çekirdek aile	3,296	18.77	4,229,093	4,038,989	20.91	2,045	20.62	4,080,906	21.13	
3	Üç ve daha fazla çocuklu çekirdek aile	2,584	14.71	2,842,462	2,714,690	14.06	1,479	14.91	2,690,411	13.93	
4	Cocuksuz çift	3,048	17.36	3,194,852	3,051,240	15.80	1,602	16.15	2,916,257	15.10	
5	Ataerkil veya geniş aile	2,886	16.43	2,892,545	2,762,522	14.30	1,483	14.95	3,060,648	15.85	
6	Tek yetişkinli aile	2,458	14.00	2,896,491	2,766,290	14.32	1,305	13.16	2,478,711	12.84	
7	Birarada yaşayan kişiler	279	1.59	314,072	299,954	1.55	184	1.86	400,385	2.07	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH010	Oturlan konutun tipi									1.3531	
1	Müstakil konut	8,846	50.37	9,233,427	8,818,373	45.66	4,484	45.21	8,491,921	43.97	
2	Apartman	8,708	49.58	10,974,248	10,480,941	54.27	5,423	54.68	10,796,724	55.91	
3	Diğer	8	0.05	12,903	12,323	0.06	11	0.11	22,992	0.12	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH020	Oturlan konuttaki mülkiyet şekli									3.5509	
1	Ev sahibi	10,794	61.46	11,978,158	11,439,725	59.24	6,086	61.36	11,626,972	60.21	
2	Kiracı	3,592	20.45	4,452,235	4,252,102	22.02	2,269	22.88	4,619,627	23.92	
3	Lojman	302	1.72	319,240	304,889	1.58	215	2.17	393,524	2.04	
4	Diğer	2,874	16.36	3,470,945	3,314,921	17.17	1,348	13.59	2,671,513	13.83	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HE010	Son 12 ay içinde EV KIRASININ, FAİZLİ BORÇ GERİ ÖDEMESİNİN veya KONUT KREDİSİNİN planlandığı gibi ödenememe durumu									2.8069	
.	.	6,768	38.54	8,242,420	7,871,912	40.76	3,832	38.64	7,684,665	39.79	
1	Evet	1,455	8.28	1,840,765	1,758,020	9.10	709	7.15	1,390,150	7.20	
2	Hayır	9,339	53.18	10,137,393	9,681,705	50.13	5,377	54.21	10,236,822	53.01	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (HH030 - 1)	Konutta oturlan süre									2.5761	
1	0-10	9,131	51.99	10,881,888	10,392,733	53.82	5,497	55.42	10,841,011	56.14	
2	11-20	4,057	23.10	4,668,577	4,458,718	23.09	2,297	23.16	4,394,433	22.76	
3	21-30	1,964	11.18	2,214,171	2,114,641	10.95	1,072	10.81	2,055,780	10.65	
4	31-40	1,382	7.87	1,426,534	1,362,409	7.05	624	6.29	1,205,082	6.24	
5	41-50	607	3.46	614,682	587,051	3.04	252	2.54	482,542	2.50	
6	51-60	297	1.69	298,397	284,983	1.48	110	1.11	209,868	1.09	
7	61-70	97	0.55	91,736	87,612	0.45	41	0.41	79,531	0.41	
8	71-80	24	0.14	22,620	21,603	0.11	24	0.24	42,805	0.22	
9	81-90	2	0.01	1,268	1,211	0.01	1	0.01	585	0.00	
10	91-100	1	0.01	706	675	0.00					
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH060	Oturlan konuttaki (mutfak, banyo ve tuvalet hariç) oda sayısı									2.2776	
1	1 oda	242	1.38	269,861	257,730	1.33	90	0.91	178,606	0.92	
2	2 oda	1,697	9.66	1,831,204	1,748,889	9.06	804	8.11	1,592,131	8.24	
3	3 oda	7,137	40.64	8,479,192	8,098,041	41.93	3,976	40.09	7,924,440	41.03	
4	4 oda	7,387	42.06	8,447,261	8,067,546	41.78	4,439	44.76	8,466,641	43.84	
5	5 oda	839	4.78	899,768	859,322	4.45	467	4.71	861,962	4.46	
6	6 oda	187	1.06	208,262	198,901	1.03	97	0.98	200,027	1.04	
7	7 oda	42	0.24	43,985	42,008	0.22	30	0.30	59,747	0.31	
8	8 oda	24	0.14	32,637	31,170	0.16	11	0.11	22,258	0.12	
9	9 oda	4	0.02	5,127	4,897	0.03	1	0.01	1,519	0.01	
10	10 ve üzeri oda	3	0.02	3,281	3,133	0.02	3	0.03	4,306	0.02	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans					HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%	
HH070	Oturulan konuttaki faydalanılan alanın metrekare olarak büyüklüğü (m2)										1.4448
1	0-100	10,611	60.42	12,467,619	11,907,183	61.66	5,895	59.44	11,573,398	59.93	
2	101-200	6,816	38.81	7,601,093	7,259,414	37.59	3,926	39.58	7,571,762	39.21	
3	201-300	126	0.72	140,881	134,548	0.70	91	0.92	160,128	0.83	
4	301-400	8	0.05	9,022	8,617	0.04	3	0.03	4,634	0.02	
5	401-500	1	0.01	1,963	1,875	0.01	2	0.02	1,132	0.01	
6	501-600						1	0.01	582	0.00	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH080	Oturulan konutun ısıtma sistemi										2.0123
1	Soba	11,175	63.63	12,012,615	11,472,632	59.41	6,097	61.47	11,597,424	60.05	
2	Müşterek veya merkezi ısıtma	1,946	11.08	2,041,502	1,949,734	10.10	1,135	11.44	2,054,173	10.64	
3	Kat kaloriferi/Kombi	3,869	22.03	5,428,558	5,184,537	26.85	2,269	22.88	4,928,409	25.52	
4	Klima	550	3.13	715,326	683,171	3.54	378	3.81	669,175	3.47	
5	Diğer	22	0.13	22,578	21,563	0.11	39	0.39	62,457	0.32	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH100	Oturulan konutta BANYO veya DUŞ'un mevcut olma durumu										0.8085
1	Evet	16,866	96.04	19,658,749	18,775,063	97.22	9,579	96.58	18,700,099	96.83	
2	Hayır	696	3.96	561,829	536,574	2.78	339	3.42	611,538	3.17	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH110	Oturulan konutta TUVALET'in (ev içinde) mevcut olma durumu										2.5290
1	Evet	15,795	89.94	18,617,453	17,780,575	92.07	8,855	89.28	17,386,872	90.03	
2	Hayır	1,767	10.06	1,603,125	1,531,062	7.93	1,063	10.72	1,924,765	9.97	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH120	Oturulan konutta BAĞIMSIZ MUTFAĞIN mevcut olma durumu										3.2684
1	Evet	16,884	96.14	19,625,181	18,743,003	97.06	9,744	98.25	19,005,590	98.42	
2	Hayır	678	3.86	595,398	568,634	2.94	174	1.75	306,047	1.58	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH130	Oturulan konutta BORULU SU SİSTEMİ'nin mevcut olma durumu										2.0187
1	Evet	17,149	97.65	19,880,342	18,986,695	98.32	9,792	98.73	19,113,236	98.97	
2	Hayır	413	2.35	340,236	324,942	1.68	126	1.27	198,401	1.03	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH140	Oturulan konutta SICAK SU SİSTEMİ'nin mevcut olma durumu										0.2669
1	Evet	14,016	79.81	16,966,338	16,203,679	83.91	8,225	82.93	16,149,924	83.63	
2	Hayır	3,546	20.19	3,254,240	3,107,958	16.09	1,693	17.07	3,161,713	16.37	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH150	Hanehalkının TELEFON HATTI'na (sabit) sahip olma durumu										2.5345
1	Evet	8,191	46.64	9,824,827	9,383,188	48.59	5,253	52.96	10,075,244	52.17	
2	Hayır	9,371	53.36	10,395,751	9,928,449	51.41	4,665	47.04	9,236,393	47.83	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH160	Hanehalkının CEP TELEFONU'na sahip olma durumu										0.2140
1	Evet	16,389	93.32	19,059,189	18,202,454	94.26	9,310	93.87	18,175,103	94.11	
2	Hayır	1,173	6.68	1,161,389	1,109,183	5.74	608	6.13	1,136,534	5.89	
HH170	Hanehalkının TELEVIZYON'a sahip olma durumu										1.5423
1	Evet	17,179	97.82	19,829,886	18,938,507	98.07	9,784	98.65	19,045,606	98.62	
2	Hayır	383	2.18	390,692	373,130	1.93	134	1.35	266,031	1.38	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
HH180	Hanehalkının BİLGİSAYAR'a sahip olma durumu										1.1359
1	Evet	7,277	41.44	9,417,835	8,994,491	46.58	4,391	44.27	8,685,382	44.97	
2	Hayır	10,285	58.56	10,802,743	10,317,146	53.42	5,527	55.73	10,626,255	55.03	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans					HBA Frekans					Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%		
HH190	Hanehalkının İNTERNET'e sahip olma durumu											2.2740
1	Evet	5,593	31.85	7,403,750	7,070,942	36.61	3,239	32.66	6,478,303	33.55		
2	Hayır	11,969	68.15	12,816,828	12,240,695	63.39	6,679	67.34	12,833,334	66.45		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HH210	Hanehalkının BUZDOLAB'na sahip olma durumu											0.9247
1	Evet	17,187	97.86	19,885,003	18,991,147	98.34	9,777	98.58	19,052,469	98.66		
2	Hayır	375	2.14	335,575	320,490	1.66	141	1.42	259,168	1.34		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HH220	Hanehalkının BULAŞIK MAKİNAS'na sahip olma durumu											3.2178
1	Evet	8,236	46.90	10,290,349	9,827,785	50.89	4,588	46.26	8,949,539	46.34		
2	Hayır	9,326	53.10	9,930,229	9,483,852	49.11	5,330	53.74	10,362,098	53.66		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HH200	Hanehalkının OTOMATİK ÇAMAŞIR MAKİNAS'na sahip olma durumu											1.1043
1	Evet	16,323	92.94	19,079,781	18,222,120	94.36	9,383	94.61	18,357,092	95.06		
2	Hayır	1,239	7.06	1,140,797	1,089,517	5.64	535	5.39	954,545	4.94		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HH230	Hanehalkının KLİMA'ya sahip olma durumu											1.7733
1	Evet	2,694	15.34	3,450,924	3,295,801	17.07	1,524	15.37	2,939,533	15.22		
2	Hayır	14,868	84.66	16,769,654	16,015,836	82.93	8,394	84.63	16,372,104	84.78		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HH240	Hanehalkının OTOMOBİL'e (iş amaçlılar hariç) sahip olma durumu											0.1561
1	Evet	5,676	32.32	6,793,667	6,488,283	33.60	3,343	33.71	6,448,031	33.39		
2	Hayır	11,886	67.68	13,426,911	12,823,354	66.40	6,575	66.29	12,863,606	66.61		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HB020	Hanehalkının yaşadığı yer											0.1508
1	Kır	5,922	33.72	5,999,665	5,729,972	29.67	3,045	30.70	5,767,628	29.87		
2	Kent	11,640	66.28	14,220,913	13,581,665	70.33	6,873	69.30	13,544,009	70.13		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (HHB)	Hane büyüklüğü											3.6022
1	1	1,499	8.54	1,745,001	1,666,561	8.63	655	6.60	1,219,157	6.31		
2	2	3,774	21.49	4,045,551	3,863,698	20.01	2,024	20.41	3,755,201	19.45		
3	3	3,693	21.03	4,582,373	4,376,389	22.66	2,219	22.37	4,466,999	23.13		
4	4	3,879	22.09	4,834,438	4,617,124	23.91	2,416	24.36	4,848,517	25.11		
5	5	2,131	12.13	2,499,328	2,386,980	12.36	1,299	13.10	2,509,889	13.00		
6	6	1,118	6.37	1,197,233	1,143,416	5.92	613	6.18	1,200,910	6.22		
7	7	619	3.52	606,781	579,506	3.00	298	3.00	563,176	2.92		
8	8	347	1.98	300,109	286,619	1.48	152	1.53	298,399	1.55		
9	9	191	1.09	165,552	158,110	0.82	99	1.00	196,909	1.02		
10	10	124	0.71	94,638	90,384	0.47	50	0.50	86,678	0.45		
11	11	72	0.41	51,934	49,600	0.26	31	0.31	59,014	0.31		
12	12	36	0.20	34,742	33,180	0.17	20	0.20	39,857	0.21		
13	13	30	0.17	24,000	22,921	0.12	12	0.12	20,431	0.11		
14	14	18	0.10	13,600	12,989	0.07	9	0.09	17,035	0.09		
15	15	11	0.06	8,121	7,756	0.04	6	0.06	8,437	0.04		
16	16	8	0.05	7,287	6,960	0.04	6	0.06	11,293	0.06		
17	17	2	0.01	1,721	1,643	0.01	3	0.03	3,889	0.02		
18	18	2	0.01	3,416	3,263	0.02	1	0.01	1,571	0.01		
19	19	3	0.02	2,477	2,366	0.01	1	0.01	650	0.00		
20	20	2	0.01	949	906	0.00	1	0.01	1,339	0.01		
21	21	1	0.01	545	521	0.00	-	0.00	-	0.00		
22	22	1	0.01	351	336	0.00	-	0.00	-	0.00		
24	24	-	0.00	-	0	0.00	2	0.02	1,701	0.01		
26	26	-	0.00	-	0	0.00	1	0.01	585	0.00		
30	30	1	0.01	429	410	0.00	-	0.00	-	0.00		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans						HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%		
Türetilmiş (EFB OECD)		Hane büyüklüğü										2.0840
1	1.0-2.0	8,986	51.17	10,389,946	9,922,905	51.38	4,906	49.47	9,456,078	48.97		
2	2.1-3.0	6,447	36.71	7,742,036	7,394,022	38.29	3,918	39.50	7,708,313	39.92		
3	3.1-4.0	1,556	8.86	1,580,776	1,509,718	7.82	814	8.21	1,586,307	8.21		
4	4.1-5.0	401	2.28	364,092	347,726	1.80	192	1.94	398,911	2.07		
5	5.1-6.0	114	0.65	98,454	94,029	0.49	59	0.59	115,076	0.60		
6	6.1-7.0	40	0.23	30,999	29,606	0.15	14	0.14	22,062	0.11		
7	7.1-8.0	10	0.06	8,178	7,810	0.04	11	0.11	21,266	0.11		
8	8.1-..	8	0.05	6,097	5,823	0.03	4	0.04	3,624	0.02		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (K COÇUK)		Beş yaşından küçük çocuk sayısı										1.9591
0	0	13,133	74.78	15,245,866	14,560,544	75.40	7,572	76.35	14,266,635	73.88		
1	1	3,349	19.07	3,872,241	3,698,178	19.15	1,878	18.94	4,014,067	20.79		
2	2	894	5.09	940,701	898,415	4.65	404	4.07	906,461	4.69		
3	3	143	0.81	125,794	120,140	0.62	48	0.48	94,935	0.49		
4	4	35	0.20	27,607	26,366	0.14	11	0.11	24,348	0.13		
5	5	6	0.03	4,285	4,092	0.02	1	0.01	1,316	0.01		
6	6	1	0.01	705	673	0.00	4	0.04	3,876	0.02		
7	7	1	0.01	3,381	3,229	0.02						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (COÇUK)		Dört yaş ile 18 yaş arasındaki çocuk sayısı										1.9368
0	0	9,318	53.06	11,055,598	10,558,635	54.67	5,024	50.66	10,218,722	52.91		
1	1	3,746	21.33	4,479,951	4,278,572	22.16	2,257	22.76	4,503,472	23.32		
2	2	2,714	15.45	3,102,912	2,963,432	15.35	1,648	16.62	2,980,228	15.43		
3	3	1,012	5.76	988,101	943,684	4.89	580	5.85	968,966	5.02		
4	4	406	2.31	337,736	322,554	1.67	231	2.33	387,364	2.01		
5	5	212	1.21	156,251	149,227	0.77	99	1.00	142,269	0.74		
6	6	99	0.56	63,016	60,183	0.31	47	0.47	63,616	0.33		
7	7	39	0.22	23,695	22,630	0.12	19	0.19	32,430	0.17		
8	8	9	0.05	9,328	8,909	0.05	5	0.05	7,010	0.04		
9	9	2	0.01	965	922	0.00	4	0.04	3,774	0.02		
10	10	1	0.01	335	320	0.00	2	0.02	2,551	0.01		
11	11	2	0.01	1,912	1,826	0.01	-	0.00	-	0.00		
13	13	1	0.01	351	335	0.00	2	0.02	1,235	0.01		
14	14	1	0.01	429	410	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ERKEK CALYAS)		17 yaş ile 40 yaş arasındaki erkek sayısı										1.7214
0	0	8,776	49.97	9,242,628	8,827,160	45.71	4,883	49.23	8,606,450	44.57		
1	1	7,620	43.39	9,431,107	9,007,166	46.64	4,404	44.40	9,127,003	47.26		
2	2	923	5.26	1,227,920	1,172,724	6.07	520	5.24	1,265,592	6.55		
3	3	200	1.14	270,257	258,109	1.34	88	0.89	242,976	1.26		
4	4	30	0.17	38,276	36,555	0.19	19	0.19	53,615	0.28		
5	5	10	0.06	6,477	6,186	0.03	4	0.04	16,001	0.08		
6	6	2	0.01	1,470	1,404	0.01						
7	7	1	0.01	2,442	2,332	0.01						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ORTA ERKEK)		39 yaş ile 66 yaş arasındaki erkek sayısı										0.8498
0	0	8,976	51.11	10,370,764	9,904,585	51.29	4,926	49.67	9,731,860	50.39		
1	1	8,512	48.47	9,772,994	9,333,685	48.33	4,954	49.95	9,499,680	49.19		
2	2	69	0.39	71,560	68,343	0.35	37	0.37	78,488	0.41		
3	3	5	0.03	5,261	5,024	0.03	1	0.01	1,609	0.01		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans					HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%	
Türetilmiş (YASLI_ERKEK)	65 yaşından büyük erkek sayısı										0.1777
0	0	15,422	87.81	18,093,035	17,279,730	89.48	8,766	88.38	17,302,383	89.60	
1	1	2,134	12.15	2,123,396	2,027,947	10.50	1,149	11.58	2,004,395	10.38	
2	2	6	0.03	4,147	3,961	0.02	3	0.03	4,859	0.03	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (KADIN_CALYAS)	17 yaş ile 40 yaş arasındaki kadın sayısı										1.9852
0	0	7,901	44.99	8,659,156	8,269,916	42.82	4,247	42.82	7,896,155	40.89	
1	1	8,632	49.15	10,322,072	9,858,081	51.05	5,052	50.94	10,163,924	52.63	
2	2	860	4.90	1,038,746	992,053	5.14	509	5.13	1,028,987	5.33	
3	3	125	0.71	161,537	154,275	0.80	81	0.82	158,904	0.82	
4	4	32	0.18	30,822	29,437	0.15	23	0.23	50,142	0.26	
5	5	6	0.03	4,922	4,701	0.02	6	0.06	13,524	0.07	
6	6	4	0.02	2,267	2,165	0.01					
7	7	2	0.01	1,056	1,009	0.01					
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (ORTA_KADIN)	39 yaş ile 66 yaş arasındaki kadın sayısı										0.9257
0	0	8,792	50.06	10,378,574	9,912,043	51.33	4,848	48.88	9,701,818	50.24	
1	1	8,591	48.92	9,662,132	9,227,807	47.78	4,968	50.09	9,412,463	48.74	
2	2	169	0.96	169,183	161,578	0.84	97	0.98	189,513	0.98	
3	3	10	0.06	10,689	10,208	0.05	5	0.05	7,844	0.04	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (YASLI_KADIN)	65 yaşından büyük kadın sayısı										0.3483
0	0	14,723	83.83	17,433,828	16,650,155	86.22	8,563	86.34	16,709,659	86.53	
1	1	2,813	16.02	2,763,573	2,639,347	13.67	1,343	13.54	2,577,671	13.35	
2	2	26	0.15	23,177	22,135	0.11	12	0.12	24,308	0.13	
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (ISSIZ)	İşsiz fert sayısı										2.5725
0	0	16,181	92.14	18,537,721	17,704,427	91.68	9,014	90.89	17,492,470	90.58	
1	1	1,237	7.04	1,505,085	1,437,430	7.44	778	7.84	1,554,599	8.05	
2	2	131	0.75	168,306	160,740	0.83	104	1.05	222,325	1.15	
3	3	8	0.05	6,409	6,121	0.03	19	0.19	34,872	0.18	
4	4	2	0.01	694	663	0.00	3	0.03	7,371	0.04	
5	5	3	0.02	2,363	2,257	0.01					
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	
Türetilmiş (ISTIHDAM)	İstihdamda yer alan fert sayısı										3.4447
0	0	4,103	23.36	4,170,214	3,982,758	20.62	1,930	19.46	3,359,373	17.40	
1	1	7,446	42.40	8,790,985	8,395,819	43.48	4,320	43.56	8,379,197	43.39	
2	2	4,382	24.95	5,336,523	5,096,640	26.39	2,717	27.39	5,471,491	28.33	
3	3	1,101	6.27	1,313,509	1,254,466	6.50	663	6.68	1,446,688	7.49	
4	4	368	2.10	447,396	427,285	2.21	204	2.06	456,938	2.37	
5	5	104	0.59	115,685	110,485	0.57	59	0.59	149,925	0.78	
6	6	37	0.21	29,712	28,377	0.15	13	0.13	25,263	0.13	
7	7	11	0.06	8,810	8,414	0.04	8	0.08	17,604	0.09	
8	8	7	0.04	4,628	4,420	0.02	3	0.03	3,186	0.02	
9	9	2	0.01	2,693	2,571	0.01	1	0.01	1,970	0.01	
10	10	1	0.01	422	403	0.00					
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00	

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans						HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%		
Türetilmiş (ISGDAHILOLM)	İşgücüne dahil olmayan fert sayısı											1.8717
0	0	3,455	19.67	4,248,359	4,057,390	21.01	2,118	21.36	4,321,701	22.38		
1	1	7,338	41.78	8,627,578	8,239,757	42.67	4,156	41.90	8,140,193	42.15		
2	2	4,621	26.31	5,006,651	4,781,596	24.76	2,519	25.40	4,671,900	24.19		
3	3	1,501	8.55	1,653,744	1,579,406	8.18	835	8.42	1,598,375	8.28		
4	4	423	2.41	461,959	441,193	2.28	215	2.17	419,731	2.17		
5	5	137	0.78	139,792	133,508	0.69	52	0.52	110,710	0.57		
6	6	59	0.34	56,450	53,913	0.28	15	0.15	34,261	0.18		
7	7	17	0.10	17,625	16,833	0.09	2	0.02	6,027	0.03		
8	8	8	0.05	4,883	4,663	0.02	2	0.02	3,640	0.02		
9	9	2	0.01	3,108	2,968	0.02	3	0.03	3,495	0.02		
10	10	-	0.00	-	0	0.00	1	0.01	1,604	0.01		
11	11	1	0.01	429	410	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ISDUR_1)	Esas işindeki durumu ücretli, maaşlı olan fert sayısı											1.1369
0	0	8,934	50.87	9,318,878	8,899,982	46.09	4,804	48.44	8,917,786	46.18		
1	1	6,344	36.12	7,759,074	7,410,294	38.37	3,738	37.69	7,393,803	38.29		
2	2	1,996	11.37	2,727,294	2,604,699	13.49	1,200	12.10	2,556,376	13.24		
3	3	252	1.43	368,335	351,778	1.82	150	1.51	370,039	1.92		
4	4	30	0.17	37,829	36,129	0.19	23	0.23	63,577	0.33		
5	5	5	0.03	8,572	8,187	0.04	3	0.03	10,056	0.05		
6	7	1	0.01	595	569	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ISDUR_2)	Esas işindeki durumu yevmiyeli olan fert sayısı											4.7260
0	0	15,817	90.06	18,245,205	17,425,060	90.23	8,552	86.23	16,619,006	86.06		
1	1	1,504	8.56	1,698,718	1,622,358	8.40	1,157	11.67	2,286,248	11.84		
2	2	202	1.15	224,326	214,242	1.11	175	1.76	346,488	1.79		
3	3	31	0.18	40,267	38,457	0.20	28	0.28	49,538	0.26		
4	4	5	0.03	7,416	7,083	0.04	5	0.05	9,675	0.05		
5	5	1	0.01	1,401	1,338	0.01	1	0.01	682	0.00		
6	6	1	0.01	1,429	1,365	0.01						
7	7	1	0.01	1,815	1,734	0.01						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ISDUR_3)	Esas işindeki durumu işveren olan fert sayısı											0.8063
0	0	16,750	95.38	19,171,244	18,309,472	94.81	9,464	95.42	18,403,618	95.30		
1	1	789	4.49	1,018,958	973,154	5.04	440	4.44	884,259	4.58		
2	2	23	0.13	30,377	29,011	0.15	14	0.14	23,760	0.12		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (ISDUR_4)	Esas işindeki durumu kendi hesabına olan fert sayısı											2.7493
0	0	13,462	76.65	15,801,059	15,090,781	78.14	7,488	75.50	14,535,770	75.27		
1	1	3,980	22.66	4,281,057	4,088,618	21.17	2,324	23.43	4,559,892	23.61		
2	2	114	0.65	131,413	125,505	0.65	102	1.03	210,903	1.09		
3	3	4	0.02	3,847	3,674	0.02	3	0.03	2,954	0.02		
4	4	2	0.01	3,202	3,059	0.02	1	0.01	2,117	0.01		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans						HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%		
Türetilmiş (ISDUR 5)	Esas işindeki durumu ücretsiz aile işçisi olan fert sayısı											1.4653
0	0	15,088	85.91	17,704,390	16,908,555	87.56	8,714	87.86	16,840,691	87.20		
1	1	1,757	10.00	1,792,586	1,712,007	8.87	868	8.75	1,746,721	9.04		
2	2	462	2.63	460,080	439,399	2.28	229	2.31	491,973	2.55		
3	3	174	0.99	187,083	178,673	0.93	72	0.73	159,237	0.82		
4	4	53	0.30	54,859	52,393	0.27	20	0.20	42,648	0.22		
5	5	16	0.09	12,345	11,790	0.06	12	0.12	26,633	0.14		
6	6	4	0.02	2,416	2,307	0.01	2	0.02	2,632	0.01		
7	7	7	0.04	6,188	5,910	0.03	1	0.01	1,102	0.01		
8	8	1	0.01	632	603	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (SONOKUL0)	Okur-yazar olmayan fert sayısı											0.7017
0	0	12,742	72.55	15,480,654	14,784,779	76.56	7,580	76.43	14,787,689	76.57		
1	1	3,717	21.17	3,768,454	3,599,057	18.64	1,849	18.64	3,586,137	18.57		
2	2	896	5.10	789,187	753,712	3.90	377	3.80	738,243	3.82		
3	3	151	0.86	128,708	122,922	0.64	79	0.80	138,690	0.72		
4	4	41	0.23	36,248	34,618	0.18	20	0.20	35,838	0.19		
5	5	8	0.05	10,934	10,442	0.05	9	0.09	16,380	0.08		
6	6	5	0.03	5,632	5,379	0.03	3	0.03	7,621	0.04		
7	7	1	0.01	429	409	0.00	1	0.01	1,039	0.01		
8	8	1	0.01	334	319	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (SONOKUL1)	Okur-yazar olup, bir okul bitirmeyen fert sayısı											1.7170
0	0	14,242	81.10	16,913,723	16,153,429	83.65	8,342	84.11	16,273,860	84.27		
1	1	2,734	15.57	2,773,186	2,648,528	13.71	1,280	12.91	2,454,656	12.71		
2	2	454	2.59	416,759	398,025	2.06	223	2.25	446,034	2.31		
3	3	96	0.55	88,147	84,185	0.44	49	0.49	88,943	0.46		
4	4	23	0.13	15,155	14,474	0.07	17	0.17	36,727	0.19		
5	5	11	0.06	12,841	12,264	0.06	5	0.05	8,663	0.04		
6	6	2	0.01	767	732	0.00	2	0.02	2,754	0.01		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (SONOKUL2)	İlkokul mezunu fert sayısı											2.6563
0	0	6,708	38.20	7,717,683	7,370,763	38.17	3,563	35.92	6,938,863	35.93		
1	1	6,270	35.70	6,953,714	6,641,136	34.39	3,471	35.00	6,608,026	34.22		
2	2	4,113	23.42	4,964,324	4,741,172	24.55	2,586	26.07	5,067,956	26.24		
3	3	370	2.11	465,908	444,965	2.30	239	2.41	556,298	2.88		
4	4	85	0.48	99,025	94,573	0.49	52	0.52	123,564	0.64		
5	5	12	0.07	15,325	14,636	0.08	4	0.04	9,845	0.05		
6	6	2	0.01	868	829	0.00	3	0.03	7,084	0.04		
7	7	1	0.01	3,411	3,258	0.02						
8	8	1	0.01	320	306	0.00						
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
Türetilmiş (SONOKUL3)	Ortaokul, mesleki ortaokul ve ilköğretim mezunu fert sayısı											0.9738
0	0	10,984	62.54	12,514,384	11,951,846	61.89	6,205	62.56	12,029,934	62.29		
1	1	4,701	26.77	5,504,722	5,257,278	27.22	2,648	26.70	5,160,189	26.72		
2	2	1,506	8.58	1,785,838	1,705,562	8.83	869	8.76	1,708,517	8.85		
3	3	300	1.71	340,768	325,450	1.69	152	1.53	320,619	1.66		
4	4	56	0.32	55,832	53,322	0.28	36	0.36	79,475	0.41		
5	5	11	0.06	13,896	13,271	0.07	5	0.05	9,729	0.05		
6	6	4	0.02	5,139	4,908	0.03	3	0.03	3,173	0.02		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans				HBA Frekans				Hellinger Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	Yalın	%	Faktörlü	%	
Türetilmiş (SONOKUL4) Genel lise mezunu fert sayısı										1.6503
0	0	14,040	79.95	15,807,739	15,097,160	78.18	7,681	77.45	14,806,200	76.67
1	1	2,928	16.67	3,640,458	3,476,814	18.00	1,814	18.29	3,614,636	18.72
2	2	507	2.89	646,511	617,450	3.20	366	3.69	759,191	3.93
3	3	70	0.40	104,079	99,401	0.51	46	0.46	104,930	0.54
4	4	13	0.07	18,920	18,069	0.09	10	0.10	22,805	0.12
5	5	4	0.02	2,872	2,743	0.01	1	0.01	3,874	0.02
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL5) Mesleki veya teknik lise mezunu fert sayısı										0.8214
0	0	14,613	83.21	16,500,050	15,758,352	81.60	8,199	82.67	15,873,119	82.19
1	1	2,545	14.49	3,189,689	3,046,308	15.77	1,492	15.04	2,967,913	15.37
2	2	379	2.16	498,906	476,480	2.47	209	2.11	430,786	2.23
3	3	23	0.13	28,320	27,047	0.14	15	0.15	33,765	0.17
4	4	2	0.01	3,613	3,451	0.02	3	0.03	6,053	0.03
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00
Türetilmiş (SONOKUL6) Yüksek okul, fakülte ve üzeri mezun fert sayısı										0.5254
0	0	14,142	80.53	15,935,267	15,218,956	78.81	7,862	79.27	15,226,263	78.85
1	1	2,280	12.98	2,821,215	2,694,397	13.95	1,380	13.91	2,706,986	14.02
2	2	1,061	6.04	1,342,663	1,282,309	6.64	628	6.33	1,273,369	6.59
3	3	70	0.40	105,100	100,376	0.52	42	0.42	94,351	0.49
4	4	8	0.05	15,738	15,031	0.08	6	0.06	10,666	0.06
7	7	1	0.01	595	568	0.00				
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00
HH040 Oturulan konut için ödenen aylık kira tutarı (TL)										3.9553
1	.	13,668	77.83	15,449,103	14,754,646	76.40	7,434	74.95	14,298,485	74.04
2	0-100	234	1.33	241,344	230,495	1.19	214	2.16	369,253	1.91
3	101-200	666	3.79	673,319	643,053	3.33	463	4.67	795,956	4.12
4	201-300	1,091	6.21	1,183,598	1,130,394	5.85	672	6.78	1,312,386	6.80
5	301-400	914	5.20	1,136,871	1,085,768	5.62	572	5.77	1,202,352	6.23
6	401-500	502	2.86	723,802	691,266	3.58	310	3.13	713,961	3.70
7	501-600	230	1.31	374,781	357,934	1.85	124	1.25	282,793	1.46
8	601-700	99	0.56	161,166	153,921	0.80	49	0.49	123,145	0.64
9	701-800	59	0.34	91,855	87,726	0.45	19	0.19	46,198	0.24
10	801-900	27	0.15	49,473	47,249	0.24	13	0.13	35,626	0.18
11	901-1000	20	0.11	32,686	31,216	0.16	13	0.13	34,141	0.18
12	1001-2000	46	0.26	89,763	85,728	0.44	29	0.29	81,278	0.42
13	2001-3000	6	0.03	12,819	12,243	0.06	3	0.03	7,769	0.04
14	3001-4000						1	0.01	2,642	0.01
15	4001-5000						2	0.02	5,650	0.03
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00

EK 4. STANDARTLAŞTIRILMIŞ GYKA VE HBA DEĞİŞKENLERİNİN YALIN VE FAKTÖRLÜ FREKANS DEĞERLERİ

Kod	Tanım	GYKA Frekans						HBA Frekans				Hellingir Uzaklığı (%)
		Yalın	%	Faktörlü	Düzeltilmiş Faktörlü	%	Yalın	%	Faktörlü	%		
HG070	Gayrimenkul (kira) geliri (TL)											3.9278
1	0-100	15,135	86.18	17,408,509	16,625,974	86.09	8,495	85.65	16,691,218	86.43		
2	101-200	71	0.40	76,889	73,433	0.38	30	0.30	57,800	0.30		
3	201-300	63	0.36	63,513	60,658	0.31	18	0.18	32,978	0.17		
4	301-400	45	0.26	43,518	41,562	0.22	13	0.13	20,798	0.11		
5	401-500	52	0.30	57,559	54,972	0.28	11	0.11	17,286	0.09		
6	501-600	45	0.26	43,567	41,609	0.22	21	0.21	36,851	0.19		
7	601-700	29	0.17	31,776	30,347	0.16	3	0.03	7,290	0.04		
8	701-800	29	0.17	33,692	32,178	0.17	9	0.09	14,853	0.08		
9	801-900	25	0.14	27,802	26,552	0.14	8	0.08	14,724	0.08		
10	901-1000	42	0.24	39,628	37,847	0.20	15	0.15	28,317	0.15		
11	1001-2000	313	1.78	340,977	325,650	1.69	143	1.44	250,331	1.30		
12	2001-3000	374	2.13	399,115	381,174	1.97	240	2.42	427,403	2.21		
13	3001-4000	296	1.69	323,995	309,431	1.60	198	2.00	371,622	1.92		
14	4001-5000	297	1.69	354,398	338,467	1.75	183	1.85	349,452	1.81		
15	5001-10000	482	2.74	612,613	585,075	3.03	307	3.10	556,078	2.88		
16	10001-20000	165	0.94	228,684	218,404	1.13	136	1.37	270,560	1.40		
17	20001-30000	56	0.32	72,528	69,267	0.36	40	0.40	77,155	0.40		
18	30001-40000	14	0.08	17,893	17,089	0.09	13	0.13	23,626	0.12		
19	40001-50000	9	0.05	14,531	13,877	0.07	10	0.10	17,539	0.09		
20	50001-60000	5	0.03	7,163	6,841	0.04	10	0.10	19,939	0.10		
21	60001-70000	5	0.03	7,424	7,090	0.04	4	0.04	6,224	0.03		
22	70001-80000	3	0.02	4,993	4,769	0.02	3	0.03	7,071	0.04		
23	80001-90000	-	0.00	-	0	0.00	3	0.03	6,296	0.03		
24	90001-100000	2	0.01	2,874	2,745	0.01	1	0.01	1,250	0.01		
25	100001-200000	5	0.03	6,935	6,623	0.03	3	0.03	2,855	0.01		
26	200001-300000						1	0.01	2,119	0.01		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		
HG110	Toplam hanehalkı kullanılabilir geliri (TL)										4.1620	
1	0-100	-	0.00	-	0	0.00	14	0.14	24,032	0.12		
2	101-200	1	0.01	717	685	0.00	1	0.01	1,455	0.01		
3	201-300	4	0.02	6,264	5,982	0.03	-	0.00	-	0.00		
4	301-400	-	0.00	-	0	0.00	1	0.01	987	0.01		
5	401-500	-	0.00	-	0	0.00	1	0.01	2,093	0.01		
6	501-600	1	0.01	715	683	0.00	1	0.01	2,125	0.01		
7	601-700	1	0.01	333	318	0.00	1	0.01	1,526	0.01		
8	701-800	2	0.01	1,358	1,297	0.01	2	0.02	2,026	0.01		
9	901-1000	2	0.01	1,347	1,286	0.01	1	0.01	1,295	0.01		
10	1001-2000	10	0.06	12,117	11,572	0.06	11	0.11	20,852	0.11		
11	2001-3000	90	0.51	88,368	84,396	0.44	46	0.46	85,419	0.44		
12	3001-4000	139	0.79	136,101	129,983	0.67	52	0.52	94,345	0.49		
13	4001-5000	232	1.32	220,945	211,013	1.09	97	0.98	182,839	0.95		
14	5001-10000	2,203	12.54	2,128,560	2,032,878	10.53	956	9.64	1,762,128	9.12		
15	10001-20000	6,521	37.13	7,170,077	6,847,773	35.46	3,495	35.24	6,617,722	34.27		
16	20001-30000	4,001	22.78	4,813,905	4,597,513	23.81	2,372	23.92	4,652,621	24.09		
17	30001-40000	1,979	11.27	2,513,421	2,400,440	12.43	1,311	13.22	2,617,087	13.55		
18	40001-50000	941	5.36	1,229,712	1,174,435	6.08	608	6.13	1,264,456	6.55		
19	50001-60000	543	3.09	701,571	670,035	3.47	360	3.63	751,817	3.89		
20	60001-70000	317	1.81	402,752	384,648	1.99	198	2.00	423,222	2.19		
21	70001-80000	168	0.96	218,584	208,758	1.08	119	1.20	250,050	1.29		
22	80001-90000	111	0.63	143,804	137,340	0.71	77	0.78	165,909	0.86		
23	90001-100000	71	0.40	95,665	91,365	0.47	33	0.33	76,155	0.39		
24	100001-200000	195	1.11	282,017	269,340	1.39	132	1.33	254,727	1.32		
25	200001-300000	22	0.13	37,863	36,161	0.19	21	0.21	44,788	0.23		
26	300001-400000	5	0.03	10,281	9,819	0.05	5	0.05	8,918	0.05		
27	400001-500000	3	0.02	4,100	3,916	0.02	1	0.01	1,893	0.01		
28	600001-700000						1	0.01	566	0.00		
29	700001-800000						1	0.01	582	0.00		
	Toplam	17,562	100.00	20,220,578	19,311,637	100.00	9,918	100.00	19,311,637	100.00		

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Levent AHI
Doğum Yeri : Ankara
Medeni Hali : Evli
E-Posta : leventahi@tuik.gov.tr
Adresi : Eryaman Mahallesi Dil Devrimi Caddesi Eryaman Evleri
Günötesi Sitesi 60/1 Eryaman Ankara

Eğitim

Lise : Ankara Ticaret Lisesi (1988 - 1991)
Lisans : Hacettepe Üniversitesi İstatistik (1992 - 1996)
Yüksek Lisans : Hacettepe Üniversitesi İstatistik (2012 - 2014)

Yabancı Dil ve Düzeyi

İngilizce - YDS 83,75

İş Deneyimi

(1997 - 1999) Sümerbank A.Ş.
Pazarlama Uzmanı
(2000 - 2000) Bayındırbank A.Ş.
Pazarlama Uzmanı
(2001 - 2003) Toprakbank A.Ş.
Müfettiş Yardımcısı
(2003 - 2003) Dışbank A.Ş.
İç Kontrol Uzmanı
(2003 - 2006) Türk Telekomünikasyon A.Ş.
Telekom Uzman Yardımcısı
(2006 - ...) Türkiye İstatistik Kurumu
İstatistikçi

Deneyim Alanları

Bankacılık, Denetim ve İstatistik

Tezden Üretilmiş Projeler ve Bütçesi

-

Tezden Üretilmiş Yayınlar

-

Tezden Üretilmiş Tebliğ ve /veya Poster Sunumu ile Katıldığı Toplantılar

-

