

**T.C.
ERZİNCAN BİNALİ YILDIRIM ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANA BİLİM DALI**

**SAĞLIK İŞLETMELERİNDE VERİ MADENCİLİĞİ
UYGULAMALARI**

Yüksek Lisans

**Hazırlayan
Murat ÇETİN**

**Danışman
Dr. Öğr. Üyesi Salih AKA**

TEMMUZ 2024, ERZİNCAN

BİLİMSEL ETİĞE UYGUNLUK

“Sağlık İşletmelerinde Veri Madenciliği Uygulamaları” isimli “**Yüksek Lisans**” tezim tarafımda incelenmiştir. Buna göre tezimde bilimsel etik ihlali ve intihal olarak nitelendirilebilecek herhangi bir durum olmadığını taahhüt ederim. Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir biçimde elde edildiğini; aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi beyan ederim.

Murat ÇETİN

T.C.
ERZİNCAN BİNALİ YILDIRIM ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü

Ana Bilim Dalı : İşletme Ana Bilim Dalı
Program Adı : Tezli Yüksek Lisans
Tez Başlığı : Sağlık İşletmelerinde Veri Madenciliği Uygulamaları

Yukarıda bilgileri verilen tez çalışmasının a) Giriş b) Ana bölümler ve c) Sonuç kısımlarından oluşan (Kapak, Ön söz, Özet, İçindekiler ve Kaynakça hariç) toplam 73 sayfalık kısmına ilişkin 03/07/2024 tarihinde **Turnitin** intihal programından aşağıda belirtilen filtreleme uygulanarak alınmış olan özgünlük raporuna göre tezin benzerlik oranı % 8'dir.

Filtrelemeye tırnak içerisindeki alıntılar dahil edilmiştir. Filtrelemede yedi (7) kelimeden daha az örtüşme içeren metin kısımları hariç tutulmuştur.

Erzincan Binali Yıldırım Üniversitesi Sosyal Bilimler Enstitüsü Tez İntihal Raporu Uygulama Esaslarını inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına göre tez çalışmasının herhangi bir intihal içermediğini, aksinin tespit edilmesi durumunda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini bilgilerinize arz ederim. 03/07/2024

Danışman: Dr. Öğr. Üyesi Salih AKA

Öğrenci: Murat ÇETİN

KILAVUZA UYGUNLUK

“Sağlık İşletmelerinde Veri Madenciliği Uygulamaları” başlıklı Yüksek Lisans Tezi Erzincan Binali Yıldırım Üniversitesi Lisansüstü Tez Yazım Kılavuzuna uygun olarak hazırlanmıştır.

Hazırlayan

Murat ÇETİN

Danışman

Dr. Öğr. Üyesi Salih AKA



KABUL VE ONAY TUTANAĞI

Dr. Öğr. Üyesi Salih AKA danışmanlığında **Murat ÇETİN** tarafından hazırlanan “**Sağlık İşletmelerinde Veri Madenciliği Uygulamaları**” adlı çalışma jürimiz tarafından Erzincan Binali Yıldırım Üniversitesi Sosyal Bilimler Enstitüsü İşletme Ana Bilim Dalı’nda **Yüksek Lisans Tezi** olarak oy birliği/oy çokluğu kabul edilmiştir.

03/07/2024

JÜRİ:

Danışman	: Dr. Öğr. Üyesi Salih AKA	(İmza)
Üye	: Doç. Dr. Salim Sercan SARI	(İmza)
Üye	: Dr. Öğr. Üyesi Gazi Bilal YILDIZ	(İmza)

ONAY:

Bu tezin kabulü Enstitü Yönetim Kurulu'nun /.... /2024 tarih ve sayılı kararı ile onaylanmıştır.

...../...../2024

Prof. Dr. Necdet TOZLU

Enstitü Müdürü

ÖNSÖZ

Bu tezi hazırlarken desteklerini, bilgi ve tecrübelerini esirgemeyen ve sürekli olarak beni yeni bilgiler öğrenmeye teşvik eden danışman hocam Sayın Dr. Öğr. Üyesi Salih AKA'ya, tez sürecimde yönlendirmeleri ve desteklerinden dolayı Doç. Dr. Zekeriya Fatih İNEÇ, Doç. Dr. Gözde MERT, Dr. Öğr. Üyesi Ömer DEMİRCİ hocalarıma teşekkürlerimi sunarım.

Hayatım boyunca yanımda olan, aldığım kararları her zaman destekleyen, cesaretlendiren ve moral veren, tez sürecimin sonuna kadar her türlü fedakarlıkta bulunan aileme sonsuz sunar teşekkür ederim.

Murat ÇETİN, Erzincan 2024

SAĞLIK İŞLETMELERİNDE VERİ MADENCİLİĞİ UYGULAMALARI

Murat ÇETİN

Erzincan Binali Yıldırım Üniversitesi

Sosyal Bilimler Enstitüsü

Yüksek Lisans Tezi, Temmuz 2024

Tez Danışmanı: Dr. Öğr. Üyesi Salih AKA

ÖZET

Bu tez çalışması, veri madenciliği tekniklerinin sağlık işletmelerinin performansını iyileştirmedeki rolünü incelemektedir. Veri madenciliği, büyük ve karmaşık veri setlerinden anlamlı bilgiler elde etme süreci olarak tanımlanmakta ve sağlık sektörü için büyük bir potansiyel taşımaktadır. Sağlık işletmeleri, hasta bilgilerini analiz ederek ve görünüşte ilgisiz veriler arasındaki bağlantıları keşfederek hizmet kalitesini ve operasyonel verimliliği artırabilir. Bu çalışmanın temel amacı, veri madenciliği tekniklerini kullanarak sağlık işletmelerinde hasta yoğunluğunu azaltmak ve poliklinik sevklerinde zaman kaybını en aza indirmektir. Çalışmada, sağlık sektöründe veri madenciliğinin kullanımına yönelik literatür taraması yapılmış, özellikle Apriori algoritması gibi sık kullanılan veri madenciliği teknikleri incelenmiş ve bu tekniklerin sağlık işletmelerinde nasıl uygulanabileceği araştırılmıştır. Bu doğrultuda, veri madenciliği tekniklerinin sağlık işletmelerinde çeşitli avantajlar sağladığı tespit edilmiştir. Bu avantajlar arasında hasta memnuniyetinin artırılması, operasyonel maliyetlerin düşürülmesi ve sağlık hizmetleri kalitesinin iyileştirilmesi yer almaktadır. Özellikle birliktelik kuralı kullanımına dayalı veri madenciliği tekniklerinin, hastaların uyguladıkları hizmetler arasında bağlantı kurallarının belirlenmesinde etkili olduğu görülmüştür. Veri madenciliği tekniklerinin sağlık işletmelerinde etkin bir şekilde kullanımının, işletmelerin karar alma süreçlerini iyileştirecek ve hasta memnuniyetini iyileştireceği düşünülmektedir. Gelecekteki araştırmaların, daha geniş veri setleri ve farklı veri madenciliği algoritmaları kullanarak bu alandaki bilgiyi derinleştirilmesi önerilmektedir.

Anahtar Kelimeler: Sürdürülebilirlik, Sağlıkta Sürdürülebilirlik, Sağlık Yönetimi, Veri Analizi, Veri Madenciliği

DATA MINING APPLICATIONS IN HEALTHCARE BUSINESSES

Murat ÇETİN

Erzincan Binali Yıldırım University

Graduate School of Social Sciences

Master Thesis, July 2024

Supervisor: Assist. Prof. Dr. Salih AKA

ABSTRACT

This thesis examines the role of data mining techniques in improving the performance of healthcare businesses. Data mining is defined as the process of obtaining meaningful information from large and complex data sets and carries great potential for the healthcare industry. Healthcare businesses can improve service quality and operational efficiency by analyzing patient information and discovering connections between seemingly unrelated data. The main purpose of this study is to reduce patient density in healthcare institutions and minimize time loss in outpatient clinic referrals by using data mining techniques. In the study, a literature review was conducted on the use of data mining in the healthcare sector, frequently used data mining techniques, especially the Apriori algorithm, were examined and how these techniques could be applied in healthcare businesses was investigated. In this regard, it has been determined that data mining techniques provide various advantages in healthcare businesses. These advantages include increasing patient satisfaction, reducing operational costs and improving the quality of healthcare services. In particular, data mining techniques based on the use of association rules have been shown to be effective in determining the connection rules between the services applied by patients. It is thought that the effective use of data mining techniques in healthcare businesses will improve the decision-making processes of businesses and increase patient satisfaction. It is recommended that future research deepen knowledge in this area by using larger data sets and different data mining algorithms.

Keywords: Sustainability, Sustainability in Health, Health Management, Data Analysis, Data Mining

İÇİNDEKİLER

SAĞLIK İŞLETMELERİNDE VERİ MADENCİLİĞİ UYGULAMALARI

BİLİMSEL ETİĞE UYGUNLUK	i
TEZ ÖZGÜNLÜK SAYFASI	ii
KILAVUZA UYGUNLUK	iii
KABUL VE ONAY TUTANAĞI.....	iv
ÖN SÖZ	v
ÖZET	vi
ABSTRACT	vii
İÇİNDEKİLER	viii
TABLOLAR LİSTESİ	xii
ŞEKİLLER LİSTESİ.....	xiii
GİRİŞ	1

BİRİNCİ BÖLÜM

SAĞLIK SEKTÖRÜ

1.1. Sağlık Kavramı	3
1.2. Sağlık Hizmetleri	4
1.3. Sağlık Sektöründe Kalite	4
1.4. Sürdürülebilirlik	4
1.5. Sürdürülebilir Sağlık Hizmetleri	5
1.6. Sağlıkta Veri Madenciliği.....	6
1.7. Sağlık Yönetimi ve Veri Madenciliği İlişkisi.....	9

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

2.1. Veri Madenciliği Teknikleri	15
--	----

2.2. Bilgi Keşfi	16
2.3. Veri Madenciliği Uygulama Alanları	17
2.4. Diğer Uygulamalar	17
2.5. Sağlıkta Veri Madenciliği Kullanımı	18
2.6. Veri Madenciliği Modelleri	19
2.6.1. Tahmin Edici Modeller	20
2.6.1.1 Sınıflandırma.....	20
2.6.1.2. Regresyon (Regression)	21
2.6.2. Tanımlayıcı Modeller	21
2.7. Birliktelik Kuralları Analizi (Association Rules Analysis).....	22
2.7.1. Birliktelik Kuralları Analizi Algoritmaları	23
2.7.1.1. Apriori Algoritması.....	23
2.7.1.2. AIS Algoritması	26
2.7.1.3. SETM Algoritması.....	26
2.7.1.4. FP-Büyüme Algoritması	27

ÜÇÜNCÜ BÖLÜM

SAĞLIK İŞLETMELERİNDE VERİ MADENCİLİĞİ UYGULAMALARI

3.1. Çalışmanın Amacı ve Önemi	28
3.2. Problemin Tanımlanması.....	29
3.3. Kullanılan Algoritma ve Programlar	29
3.3.1. MySQL Programı	29
3.3.2. MySQL Front Programı.....	30
3.3.3. Python Programı	31
3.4. Verilerin Hazırlanması	32
3.4.1. Verilerin Kapsamı.....	34

3.4.2. Veri Seçme ve Dönüştürme	34
3.4.3. Veri Temizleme	36
3.5. Bulgular	41
3.5.1. Verilerin Tanımlayıcı İstatistikleri.....	41
3.5.2. Hasta Demografik Bilgileri.....	41
3.5.2.1. Cinsiyet Bazlı Hasta Dağılım Bilgiler.....	42
3.5.2.2. Cinsiyet ve Yaş Gruplarına Göre Hasta Dağılım Bilgiler.....	42
3.5.2.3. Sağlık Güvence Kurumuna Göre Hasta Dağılım Bilgiler.....	44
3.5.2.4. İkamet Adreslerine Göre Hasta Dağılım Bilgiler.....	45
3.5.2.5. Vatandaşlık Bilgilerine Göre Hasta Dağılım Bilgiler	46
3.5.2.6. Aylara Göre Hasta Dağılım Bilgiler	47
3.5.2.7. Saatlere Göre Hasta Dağılım Bilgiler	48
3.5.3. Birliktelik Kuralları ile Poliklinik Analizi	49
3.5.3.1 Poliklinik Genel Analizi.....	49
3.5.3.1.1. Polikliniklere Göre Vizit Sayısı Bilgileri	50
3.5.3.1.2. Erkek Hastaların Vizit Sayısı Bilgileri.....	51
3.5.3.1.3. Kadın Hastaların Vizit Sayısı Bilgileri.....	51
3.5.3.1.4. Aylara Göre Hastaların Vizit Sayısı Bilgileri.....	52
3.5.3.1.5. Sosyal Güvenlik Kurumlarına Göre Hastaların Vizit Sayısı Bilgileri.....	55
3.5.3.1.6. Yaş Gruplarına Göre Hastaların Vizit Sayısı Bilgileri.....	56
3.5.3.1.7. İkamet Edilen Şehirlere Göre Hastaların Vizit Sayısı Bilgileri.....	58
3.5.3.1.8. İkamet Edilen İlçelere Göre Hastaların Vizit Sayısı Bilgileri	59
3.5.3.2 Birliktelik Kuralı Kullanarak Poliklinik Analizi.....	60
3.5.3.2.1. Verilerin Birliktelik Kuralına Uygun Hale Getirilmesi.....	60

SONUÇ VE TARTIŞMA	72
KAYNAKÇA.....	74
ÖZGEÇMİŞ	80



TABLÖLAR LİSTESİ

Tablo 2.1: Veri Madenciliğinin Modelleri (Akgün vd. 2022)	19
Tablo 3.1: Cinsiyet değişkenine göre veri	42
Tablo 3.2: Yaş gruplarına göre cinsiyet sayılarının veri seti dağılımı	43
Tablo 3.3: 45 ve Üzeri yaştaki hastaların cinsiyete göre sayı dağılımı	43
Tablo 3.4: Sağlık Güvencesi Kurumu değişkenine göre veri setinin dağılımı	44
Tablo 3.5: Hastaların ikamet bilgilerine göre veri setinin dağılımı	45
Tablo 3.6: Hastaların vatandaşlık bilgilerine göre veri setinin dağılımı	46
Tablo 3.7: Aylara göre hastaların veri setinin dağılımı	47
Tablo 3.8: Saatlere göre hastaların veri setinin dağılımı	48
Tablo 3.9: Polikliniklere göre hastaların veri setinin dağılımı	50
Tablo 3.10: Aylara göre hastaların ziyaret hasretleri veri setinin dağılımı	52
Tablo 3.11: Sosyal Güvenlik Kurumuna göre ziyaret sayılarının dağılımı	55
Tablo 3.12: Ziyaret hareketlerinin yaş gruplarına göre veri setinin dağılımı	56
Tablo 3.13: İkamet edilen şehirlere göre ziyaret hareketlerinin veri setinin dağılımı	59
Tablo 3.14: İkamet edilen ilçelere göre ziyaret hareketlerinin veri setinin dağılımı	59
Tablo 3.15: Birliktelik kuralı analizine uygun veri seti örneği	60
Tablo 3.16: Apriori algoritmasına uygun veri seti örneği	61
Tablo 3.17: Apriori Analiz Sonucu Veri Seti	63
Tablo 3.18: Güven Değeri Sonucu Veri Seti	64
Tablo 3.19: Acil ve Diğer Birimler Olmadan Çıkan Sonuçlar Veri Seti	65
Tablo 3.20: Acil ve Diğer Birimler olmadan Güven Destek Değerine Göre Çıkan Sonuç Veri Seti	66
Tablo 3.21: Projenin Python Kodlarının Tamamı	70

ŞEKİLLER LİSTESİ

Şekil 1.1:	Konu Üzerine Öne Çıkan Araştırma Alanları	10
Şekil 1.2:	Yıllara Göre Yayın Sayısı.....	10
Şekil 1.3:	Ülke Bazında Konu Çalışma Ağırlığı.....	11
Şekil 1.4:	Konu Çalışma Alanları	12
Şekil 2.1:	Veri Madenciliği ve İş Zekâsı Piramidi.....	14
Şekil 3.1:	Mysql Front Programı	31
Şekil 3.2:	Visual Studio Code Programına Ait Görsel.....	32
Şekil 3.3:	Aylık Excel Verilerinin MYSQL Veri Tabanında Birleştirilmesi	33
Şekil 3.4:	Ocak Ayına Ait Verilerin EXCEL Görüntüsü.....	34
Şekil 3.5:	Poliklinik Analizi İçin Kullanılacak Veri Seti MySQL Görseli.....	36
Şekil 3.6:	CSV Dosyasında Bulunan Verilerin Pandas Kütüphanesi ile Dataframe Olarak Önbelleğe Alınıp, Listelendiği Jupyter Notebook Sonuç Ekranı	39
Şekil 3.7:	Veri Setimizdeki Boş Kayıtlar Kodları ve Sonuçları	40
Şekil 3.8:	Veri Setimizdeki Veri Temizlik Kodları ve Sonuçları	41
Şekil 3.9:	Veri Setimizdeki Veri Temizlik Kodları ve Sonuçları	46
Şekil 3.10:	Sağlık Hizmeti Alan Kişilerin Vatandaşlık Sayı Grafiği	47
Şekil 3.11:	Aylara Göre Hasta Sayısı Grafiği	48
Şekil 3.12:	Saatlere Göre Hasta Sayısı Grafiği	49
Şekil 3.13:	Veri Setinde Veri Temizliği Sonuç Görseli.....	50
Şekil 3.14:	Erkek Hastaların Vizit Hareketleri Grafiği.....	51
Şekil 3.15:	Kadın Hastaların Vizit Hareketleri Grafiği.....	51
Şekil 3.16:	Aylara Göre Hastaların Birimlere Başvuru Hareketleri Grafiği.....	54
Şekil 3.17:	Kurumlara Göre Vizit Hareket Sayıları Grafiği	56
Şekil 3.18:	Apriori Algoritmasına Uygun Hale Dönüştürülmüş Veri Seti Çıktısı	62

GİRİŞ

Sağlık sektöründe, artan hasta yoğunluğu ve zaman kayıpları, sağlık hizmetlerinin etkin ve verimli bir şekilde sunulmasını engellemektedir. Hastaların poliklinik sevklerinde yaşanan gecikmeler, hem hasta memnuniyetini düşürmekte hem de sağlık kuruluşlarının operasyonel verimliliğini olumsuz etkilemektedir. Bu bağlamda, sağlık işletmelerinin performansını artırmak amacıyla veri madenciliği tekniklerinin kullanımı önemli bir araştırma konusu olarak ortaya çıkmaktadır.

Geleneksel tıp yöntemlerinin ötesine geçerek, veri tabanlı ve analitik yaklaşımlarla desteklenen yeni sağlık modelinin merkezinde hastane veri analitiği yer almakta olup, veri analitiği sağlık hizmetlerinin kalitesini artırmada ve operasyonel verimliliği sağlamada kritik bir rol oynamaktadır. Hastaneler, veri analitiği yöntemlerini kullanarak hastaların tıbbi kayıtlarını daha iyi analiz etmekte ve bu sayede daha etkili hizmet sunmaktadırlar. Veri madenciliği, sağlık kuruluşlarında yeni trendlerin ve gizli bilgilerin keşfedilmesinde önemli bir araç olarak öne çıkmaktadır. Sağlık sektöründe büyük veri setlerinin hızla artması, bu verilerin anlamlı hale getirilmesi için gelişmiş analiz tekniklerinin kullanımını zorunlu kılmaktadır. Veri madenciliği, bu bağlamda tahmin ve karar destek sistemlerini destekleyerek sağlık hizmetlerinde önemli iyileştirmeler sağlamaktadır.

Birliktelik kuralı madenciliği, sağlık sektöründe sıkça kullanılan veri madenciliği yöntemlerinden biridir. Bu çalışmada, birliktelik kuralına dayalı veri madenciliği tekniği kullanarak hastane poliklinik birimleri için en uygun yerleşim düzenini belirlemeye odaklanılmaktadır. Çalışmanın temel amacı, veri madenciliği tekniklerini kullanarak sağlık işletmelerinde hasta yoğunluğunu ve poliklinik sevklerinde yaşanan zaman kaybını azaltmaktır. Bu doğrultuda, mevcut ve yeni açılacak sağlık kuruluşlarında hasta yoğunluğuna göre poliklinik kat planlarının yapılmasında yöneticilere yardımcı olunması hedeflenmektedir. Çalışmada tıbbi verilerden yararlanılarak hastalara uygulanan hizmetler arasında bağlantı kuralları olup olmadığı incelenmiştir. Ayrıca yapılan ayrıntılı veri analizleri ile poliklinik birimlerine yönelik durum tespiti yapılmıştır.

Çalışmanın birinci bölümünde sağlık sektörü ve veri madenciliği ilişkisi sürdürülebilirlik çerçevesi içerisinde incelenmiştir. Bu kapsamda sağlık sektörü ve veri

madencilięi iliřkisini konu alan literatür, bibliyometrik analiz yöntemiyle araştırılmıřtır. İkinci bölümde veri madencilięi kavramsal olarak açıklanmıř, uygulama alanlarına değinilmiř ve özellikle çalışmada kullanılan birliktelik kuralı hakkında bilgiler sunulmuřtur. Üçüncü bölümde ise çalışmanın amacı, kapsamı ve analiz yöntemi açıklanmıř, edilen bulgular ayrıntılı olarak tartıřılmıřtır.

Çalışmanın sonucunda, hastane veri madencilięinin saęlık sektöründeki önemi, uygulama alanları ve geleceęe yönelik olası etkileri vurgulanmıřtır. Çalışmanın sonuçlarının, saęlık yöneticilerine, veri madencilięi ve veri analitięi konularında yol gösterici bilgiler sunacaęı ve saęlık hizmetlerinin kalitesinin artırılmasına katkıda bulunacaęı düşünölmektedir.

BİRİNCİ BÖLÜM

SAĞLIK SEKTÖRÜ

1.1. Sağlık Kavramı

Dünya Sağlık Örgütü sağlık kavramını “sadece hastalık veya sakatlığın olmayışı değil, fiziksel, zihinsel ve sosyal olarak tam bir iyilik hali” olarak tanımlamaktadır (World Health Org., 2022). Sağlık alanında hastalara verilen hizmet, bir ülkenin sosyo-ekonomik düzeyinin belirlenmesinde önemli bir göstergedir. Bir sağlık işletmesinin sunduğu toplumsal sağlık ihtiyaçlarının ve hasta memnuniyetinin doğru analiz edilmesi, işletmenin başarısı ve rekabet gücü açısından büyük önem taşımaktadır. Dolayısıyla sağlık sektörü gelişim ve çözüm odaklı her türlü çabaya büyük ihtiyaç duymaktadır. Bu bakış açısıyla, sağlık işletmeleri, araştırma ve geliştirme faaliyetlerini yoğunlaştırarak sağlık alanındaki pazar paylarını giderek artırmaktadır. Bireysel sağlığa odaklanmanın yanı sıra özel sağlık tesislerinin milli katkısını ve mali gücünü korumaya yönelik çalışmalar da işletmelerin amaçları içerisinde. Sağlık işletmeleri bir yandan bireysel sağlık hizmetlerine odaklanarak toplumsal yaşam standardının gelişmesine katkı sağlarken diğer yandan ekonomik değer yaratarak ülkenin mali gücünün korunmasına yardımcı olmaktadır (Karabulut, 2021).

Sağlık Bakanlığı, 2023 yılında tüm branşlarda 804 Milyon 131 Bin 471 muayene gerçekleştirirken acil servislerde 150 milyon 523 bin 502 başvuruyu karşılamıştır. Ağız ve diş sağlığı hastanelerinde ise 39 milyon 252 bin 502 hizmet verilmiş. 11 milyon 875 bin 556 adet ameliyat yapılmıştır. Evde sağlık merkezlerinde ise 6 milyon 591 bin 114 hizmet verilmiştir. Bu bilgiler Türkiye’deki sağlık işletmelerinin yoğunluğu hakkında fikir vermektedir (Sağlık Bakanlığı, 2024).

1.2. Sağlık Hizmetleri

Ülkemizde Sağlık hizmetleri, 224 sayılı sağlık hizmetlerinin toplumsallaştırılması hakkındaki kanunda belirtildiği üzere; “Sağlık hizmetleri, insan sağlığını bozan çeşitli etkenleri ortadan kaldıran, toplumu bu etkenlerin etkilerinden koruyan, hastaları tedavi eden ve yürütülen tıbbi faaliyetler” olarak tanımlanmaktadır (Türkiye Cumhuriyeti Resmi Gazete [TCRG], 1961). Toplumsal uyum için, tedavi hizmetleri, güvenli barınma, sağlıklı beslenme, eğitim ihtiyacının karşılanması, sosyal yaşam, ekonomik refah, gelir düzeyi gibi konuların yeterli koşullar altında sağlanması gerekmektedir. Bu unsurlar içerisinde sağlık hizmetleri öne çıkmaktadır. İnsan ihtiyaçları arasında biyolojik ihtiyaçlar ilk sırada yer alırken, biyolojik ihtiyaçlar arasında ise sağlık ihtiyaçları en önemli konu olarak görülmektedir (Karabulut,2021).

Sağlık hizmetlerinin geliştirilmesi konusunda teknoloji uygulamalarından faydalanılmaktadır. Teknoloji uygulama tekniklerinden biri olan veri madenciliği sağlık kurumlarının hasta demografik bilgilerinin analizi, kurumların hasta yoğunluğuna göre planlanması ve maliyetlerinin düşürülmesinden, makine öğrenimi teknikleri ile erken hastalık teşhisine ve genetik bilgilerin ortaya çıkarılmasına kadar birçok alanda kullanılmaya başlanmıştır (Karaibrahimoğlu, 2014).

1.3. Sağlık Sektöründe Kalite

Kaliteli sağlık hizmetlerine ihtiyaç duyan bireyler sağlık hizmeti sunan kuruluşlardan belirli beklentilere sahiptirler. Sağlık kuruluşları bireylerin sağlığını korumaya, doğru tedavi yöntemlerini uygulamaya ve bireylerin iyilik halini geliştirmeye odaklanmaktadır. Kuruluşların bu yükümlülükleri tam olarak yerine getirebilmeleri ve bireyleri memnun edebilmeleri büyük ölçüde verecekleri hizmetin kalitesine bağlıdır. Sağlık koşullarının geliştirilmesi için sağlık hizmetleri kalitesinin artırılması gerekmektedir. Bu nedenle ülkedeki sağlık hizmetlerinin kalite koşulları, sağlık bakanlığının sağlıkta kalite standartları ile tüm sağlık kurum ve kuruluşlarını hedefleyecek şekilde geliştirilmektedir (Demirbilek ve Çolak, 2008).

1.4. Sürdürülebilirlik

Teknoloji ve toplumsal örgütlenmenin sürekli genişlemesi olgusunun bir ulusun mevcut ve gelecekteki ihtiyaçlarını karşılama becerisine dayattığı sınırlamalar nedeniyle

ekonomik ve sosyal kalkınma hedeflerinin her açıdan sürdürülebilirlik perspektifinden tanımlanması gerekmektedir. Sürdürülebilir kalkınma gelişmiş veya gelişmekte olan ülkeler açısından pazar odaklı veya merkezi olarak planlanmıştır. Bu kavram, stratejik açıdan ekonominin ve toplumun kademeli dönüşümü anlamına gelmektedir. Kalkınmaya giden yollar, fiziksel olarak sürdürülebilir olmalı ve teorik olarak sağlıklı bir sosyal ve politik ortam sağlamalıdır. Ancak, fiziksel sürdürülebilirlik yalnızca kaynaklara erişim, maliyet-fayda dağılımındaki değişiklikler gibi konulara dikkat edilmesiyle sağlanabilir. Fiziksel sürdürülebilirlik, nesiller boyunca sosyal adalete yönelik bir kaygıyı içerir ve bu kaygı, her nesildeki adaleti de kapsar. Sürdürülebilir kalkınma politikaları, yalnızca çevresel değil, aynı zamanda sosyal ve ekonomik boyutları da göz önünde bulundurarak, gelecek nesillerin ihtiyaçlarını karşılayabilme kapasitesini korumayı hedeflemelidir. Bu kapsamda, fiziksel sürdürülebilirlik ve sosyal adalet, birbiriyle doğrudan ilişkilidir ve sürdürülebilir kalkınmanın temel taşlarını oluşturur. (Brundtland,1987).

1.5. Sürdürülebilir Sağlık Hizmetleri

Sağlığın Dünya Sağlık Örgütü'ne göre tanımına göre ulaşılabilir en yüksek sağlık düzeyi: Sağlık, ırk, din, siyasi düşünce, ekonomik veya sosyal statü ayrımı yapılmaksızın insanların sahip olması gereken en temel haklardan biri olarak kabul edilmektedir. Sağlık aynı zamanda bireylerin, grupların ve toplumun fizyolojik, psikolojik ve sosyal yönlerini de ilgilendiren bir kavramdır. Dolayısıyla sağlık, yalnızca hastalık veya sakatlığın olmaması durumundan çok daha geniş bir tanıma sahiptir. Sağlığın tanımında insanların günlük yaşamlarında sürekli olarak çevre ile etkileşim içinde oldukları ve çevrenin özelliklerinden önemli ölçüde etkilendikleri gerçeği dikkate alınmalıdır (Çelik, 2006).

Sağlık hizmetleri maliyetleri her geçen gün yükselmekte olduğundan doğal ve finansal kaynakların verimli kullanımının önemi artmaktadır. Sağlık şirketlerinin mümkün olduğu kadar az girdiyle daha uzun ömürlü sağlık hizmeti sunarak rekabet avantajı kazanmaları gerekmektedir. Dolayısıyla sağlık yatırımcılarının sürdürülebilir, süreç odaklı, çevreyle uyumlu, maliyetleri düşüren ve ekonomik ihtiyaçları dengeli bir şekilde karşılayan bir sağlık altyapısına ihtiyaç duyacağı açıktır (Karaca vd., 2018).

Sürdürülebilir sağlık hizmetleri, ilk yatırım maliyetleri açısından sağlık hizmeti sağlayıcıları açısından dezavantajlı gibi görünse de, uzun vadeli maliyetleri düşürmesi,

işletme ve bakımının daha ucuz olması ve katma değere sahip olması nedeniyle yatırımcılar için caziptir.

Çevrenin korunması, kalite ve verimlilik esaslarına dayanan sürdürülebilir sağlık hizmetleri, yeşil hastane konsepti çerçevesinde ekonomik, çevresel ve toplumsal amaçlara hizmet etmektedir. Bu konsept dahilinde düşük enerji tüketimine sahip bina altyapısı oluşturulmakta, mevcut yapılar bu doğrultuda yeniden tasarlanmakta, enerji tasarrufu ve düşük karbon emisyonuna olanak sağlayan yeni nesil tıbbi görüntüleme sistemleri kullanılmakta, doğal kaynaklar ve çevre korunmakta, bir başka deyişle çevre dostu çözümler sağlanmaktadır. Fakat sürdürülebilir verimlilik, binaların ve enerji kaynaklarının doğru kullanımının yanında çeşitli maliyet kalemlerinden kaynaklı olumsuzlukların azaltılmasına da ihtiyaç duymaktadır. Zaman, yer ve ekonomik tasarruf sağlamak için sağlık bilgi teknolojisinin yönetim sistemlerine entegre edilmesi ve iş akışında, operasyonlarda ve diğer tüm planlama süreçlerinde kullanılması gerekmektedir.

Sürdürülebilir hastaneler verimli ve çevre dostu olmasının yanı sıra yüksek kalite ve hasta bakım hizmetlerine de odaklanmaktadır. Özellikle hasta bakımı ve hasta bilgileri ile ilgili süreçleri doğru tanımlayabilen, bunları güvenli bir şekilde saklayan ve gerektiğinde bu bilgilere kolay ve hızlı bir şekilde ulaşabilen hastaneler, doğru bilgiyi muhafaza edebilecek, hasta bekleme sürelerini azaltabilecek ve hasta memnuniyetini artıracaklardır. Ayrıca sağlık bilgi çözümlerini etkin bir şekilde kullanan hastaneler, teşhis ve ileri tedaviler için görüntü işleme yoluyla kişiselleştirilmiş tıbbi hizmetler sunarak bakım kalitesini artırabilmektedirler (Karaca vd., 2018).

1.6. Sağlıkta Veri Madenciliği

Sağlık hizmetlerinde yürütülen veri madenciliği çalışmalarında, hastaların elektronik tıbbi kayıtlarından yararlanılmaktadır. Bu verilerden çeşitli tahminler yapmak mümkündür (Yıldırım vd., 2008). Örneğin, bunlardan bazıları şunlardır:

- Belirli bir hastalıktan muzdarip kişilerin genel özelliklerini tahmin etmek
- Tıbbi tedavi sonrasında hastanın durumunu tahmin etmek
- Hastane maliyetlerini tahmin etmek
- Ölüm oranlarını ve salgın hastalıkları tahmin etmek

Hastalık yönetiminde veri madenciliği araştırması, hastalığın tanımlanmasının yanı sıra hastalık durumları ve maliyet modellemesi gibi çalışmaları da içermektedir (Yıldırım vd., 2008). Sağlık kurumlarında veri birikimine neden olan uygulamalar ayrıntılı olarak ele alınırsa şu ana başlıklar altında toplanabilir (Çarklı, 2010):

- Görüntüleme
- Tanı
- Tedavi
- Prognoz
- Hastalık evrelerini kontrol etmek
- Biyomedikal ve biyolojik analiz
- Epidemiyolojik araştırmalar
- Hastane yönetimi
- Tıbbi rehberlik ve eğitim

Tıp alanında biyolojik verilerin yanı sıra hastane yönetimi ve çeşitli teşhis, tanı ve tedavi süreçlerine dair veriler kullanılarak analizler gerçekleştirilebilmektedir (Çarklı, 2010). Demirci vd. (2021) çalışmalarında sağlık hizmetleri ve toplum sağlığı alanındaki performansı ölçmek için stokastik sınır analizi (SSA), veri zarflama analizi (DEA) ve Malmquist toplam faktör verimlilik endeksi (MTFV) analizlerini kullanmışlardır. Ulusal çalışmalarda kullanılan girdi ve çıktı değişkenlerinin sıklığını araştırmak, sağlık hizmetlerinin hem kurumsal hem de ulusal düzeyde etkililiğini/verimliliğini belirlemek istenmiştir. Yetginler (2019) çalışmasında dünya çapında sık görülen kanserlerden rahim ağzı kanserine yönelik olarak erken teşhis imkânı sağlayacak araştırmalar yapmak istemiştir. Çalışmada Naive Bayes algoritması, karar ağacı, destek vektör makinesi gibi veri madenciliği teknikleri kullanılarak elde edilen sonuçlar analiz edilmiştir. Çataloluk (2012) çalışmalarında K-en yakın komşu (KNN) ve k-means algoritmaları kullanarak veri madenciliği yapmış, dermatolojik hastalıkların teşhisi için tahmin yapabilen ve hasta kayıtlarının nitelikleri arasındaki ilişkileri analiz etme imkânı sunabilen yardımcı bir karar verme sistemi tasarlamıştır. Irmak vd. (2012) çalışmalarında ise hasta yoğunluğunu tahmin etmek için çeşitli veri madenciliği tekniklerini kullanmışlardır ve hastane veri

tabanlarından elde edilen sonuçları karşılaştırmalı olarak sunmuşlardır. Kullanılan hastane veri tabanı üzerinde veri aktarımı, filtreleme ve veri ön işleme faaliyetleri gerçekleştirilmiş, zaman serileri ve yapay sinir ağı teknikleri kullanılarak birçok veri madenciliği tahmin modeli oluşturularak üstel yumuşatma, ARIMA ve yapay sinir ağı teknikleri öncelikle kendi içlerindeki farklı modellerle karşılaştırılmıştır. Söz konusu yöntemlerin en uygun modeli belirlenerek birbiriyle karşılaştırılmış ve en iyi sonuçları veren modeller belirlenmiştir.

Veri madenciliği tekniklerinden birliktelik kuralını kullanarak hastane veri tabanlarından elde edilen verilerin incelendiği çalışmalar da bulunmaktadır. Alan vd. (2019) çalışmalarında, toplam 28.738 hasta verisi üzerinde uyguladığı analiz sonucunda hasta profillerini belirleyerek sonuçların hastane yönetimine daha hızlı ve güvenilir kararlar almasında katkı sağlamasını amaçlamışlardır. Oğuz vd. (2009) çalışmalarında, metin madenciliği yöntemlerine apriori birliktelik kuralını entegre ederek kulak burun boğaz epikriz notlarını analiz etmiş ve bu doğrultuda hastaların kanser durumlarını tahmin etmeyi, doktorların karar verme süreçlerinde harcanan zamanı kısaltmayı ve hızlı karar vermeyi hedeflemişlerdir. Çarklı (2010) çalışmasında olası migren teşhisi için Sakarya Üniversitesi çalışanları arasında yapılan anket sonuçlarını analiz etmek amacıyla Apriori algoritmasını kullanarak korelasyon kuralları oluşturmuştur. Çalışma kapsamında, Apriori algoritmasını uygulayan bir web tabanlı yazılım .NET platformunda geliştirilmiş ve bu yazılım sayesinde Apriori algoritmasının işlevselliği anlık olarak takip edilmiştir. Doğrul vd. (2015) çalışmalarında birliktelik kuralları algoritmalarından faydalanarak trafik kazası verisini analiz edilmiştir. Çalışmasında, yer ve zaman faktörüne bağlı olarak kaza sıklıkları tespit edilmekte ve önleyici tedbirlerin artırılmasıyla kazaların azaltılabileceği belirtilmiştir.

Veri madenciliği teknikleri sağlık hizmetleri kalitesinin geliştirilmesinde başvuru yöntemlerinden biri olmaktadır. Altındış (2018) çalışmasında verimliliği ve kaliteyi artırmayı esas alarak müşteri memnuniyetini etkileyen genel faktörleri değerlendirmeyi amaçlamıştır. Çalışma sonucunda büyük veriyi kullanarak istatistiksel veri hesaplamanın oldukça zor olmasına rağmen elde edilen sonuçların daha doğru ve kapsamlı karar alınmasına yol açabileceği savunulmuştur. Bu sonuca dayanarak büyük veriye erişim için teknolojik gereksinimlerin oluşturulması amacıyla gerekli yatırımların

yapılması önerilmektedir. Karabulut (2021) çalışmasında, özel bir sağlık kuruluşunda tedavi gören hastalara ait verileri kullanarak veri analizi kullanarak, sonuç odaklı iyileştirme önerileri sunmayı amaçlamıştır. Bu çalışma kapsamında, bakım alan bireylerin ve bakım hizmeti sağlayan kurumun özellikleri kategorize edilerek hasta profilleri çıkarılmıştır. Hasta profillerinin belirlenmesinin, hastanelerin hasta ihtiyaçlarını karşılamada daha doğru bir yol izlemesine yardımcı olacağı düşünülmüştür.

1.7. Sağlık Yönetimi ve Veri Madenciliği İlişkisi

Bibliyometri, yayımlanmış kitapları, dergileri ve makaleleri incelemek için matematiksel ve istatistiksel yöntemler kullanan saymaya dayalı analitik bir alandır. Bir belgenin/yayının dergi, konu, yazar sayısı, yayın bilgileri gibi belirli özelliklerinin analizine bibliyografik analiz denilmektedir. Belirli bir alandaki akademik çalışmaların ve araştırma sonuçlarının özelliklerini niceliksel olarak analiz eden bir araştırma yöntemi olan bibliyometrik analizin birçok alanda kullanımı giderek artmaktadır (Ayaz vd. 2018). Çalışmamızda, sağlık yönetimi ve veri madenciliği ilişkisine daha sistematik bakabilmek amacıyla bibliyometrik analiz kullanılmıştır. Bibliyometrik analiz adımları şu şekildedir (Bamakan vd., 2022):

Adım 1- Veri tabanı seçimi: Kapsam dikkate alınarak Web of Science veritabanı tercih edilmiştir.

Adım 2- Anahtar kelimelerin tanımlanması: Analiz için, “health management” ve “data mining” kelime öbekleri kullanılmıştır.

Adım 3- Bulgu filtreleme: Anahtar kelimelerce yapılan aramada 5925 adet yayın bulunmuştur. İngilizcede ve makale formatında olan çalışmalar filtrelendiğinde bu sayı 4311 olarak güncellenmiştir.

Adım 4- Analiz: Sonuçlar, çalışma alanları, trend, yazarların ülkeleri gibi alanlarda sınıflandırılmıştır.

Adım 5- İlgili literatür sentezi oluşturulmuştur.

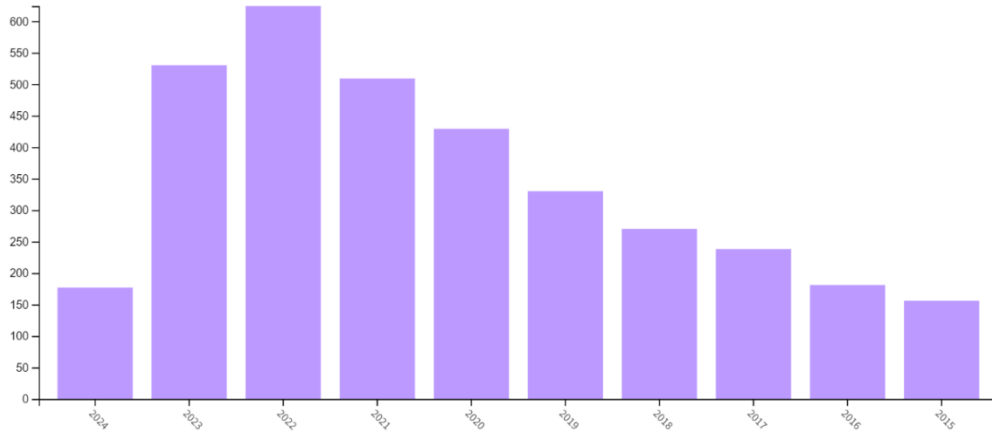
Web of Science veri tabanından elde edilen bilgiler için platformun ön analiz kısmından sağlanan bilgiler doğrultusunda Şekil 1.1’de görüleceği üzere konunun öne çıkan araştırma alanları aşağıdaki şekildedir.

Şekil 1.1: Konu Üzerine Öne Çıkan Araştırma Alanları



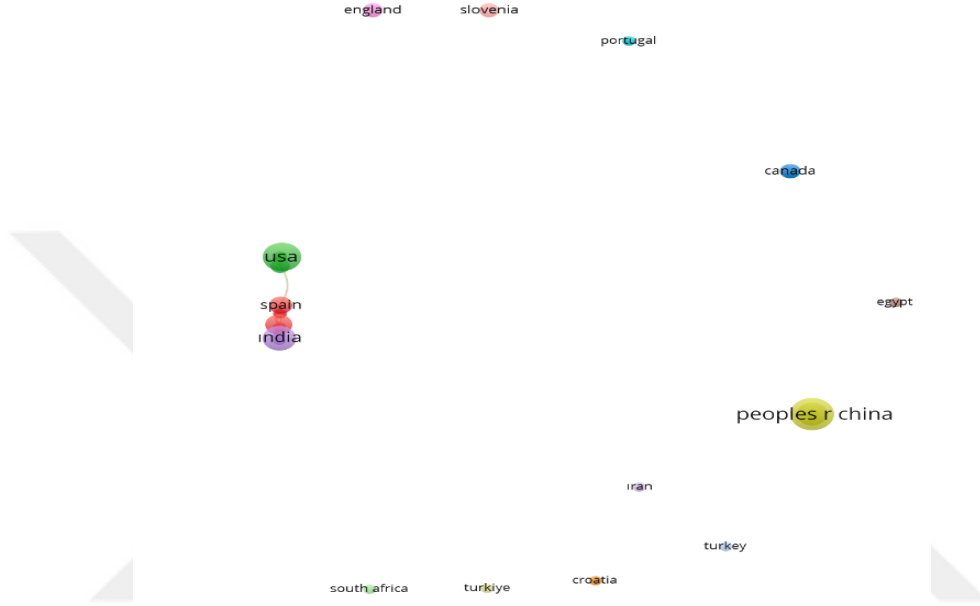
Şekil 1.1'e göre Sağlık yönetimi ve veri Madenciliği konusundaki çalışmaların ağırlıklı olarak çevresel bilimler, kamu çevre iş sağlığı, bilgisayar bilimi bilgi sistemleri, tıbbi enformatik bilimleri gibi alanlarda olduğu görülmektedir. Sağlık bilimi hizmetleri de yine önde gelen çalışma alanlarından biri olarak dikkat çekmektedir.

Şekil 1.2: Yıllara Göre Yayın Sayısı



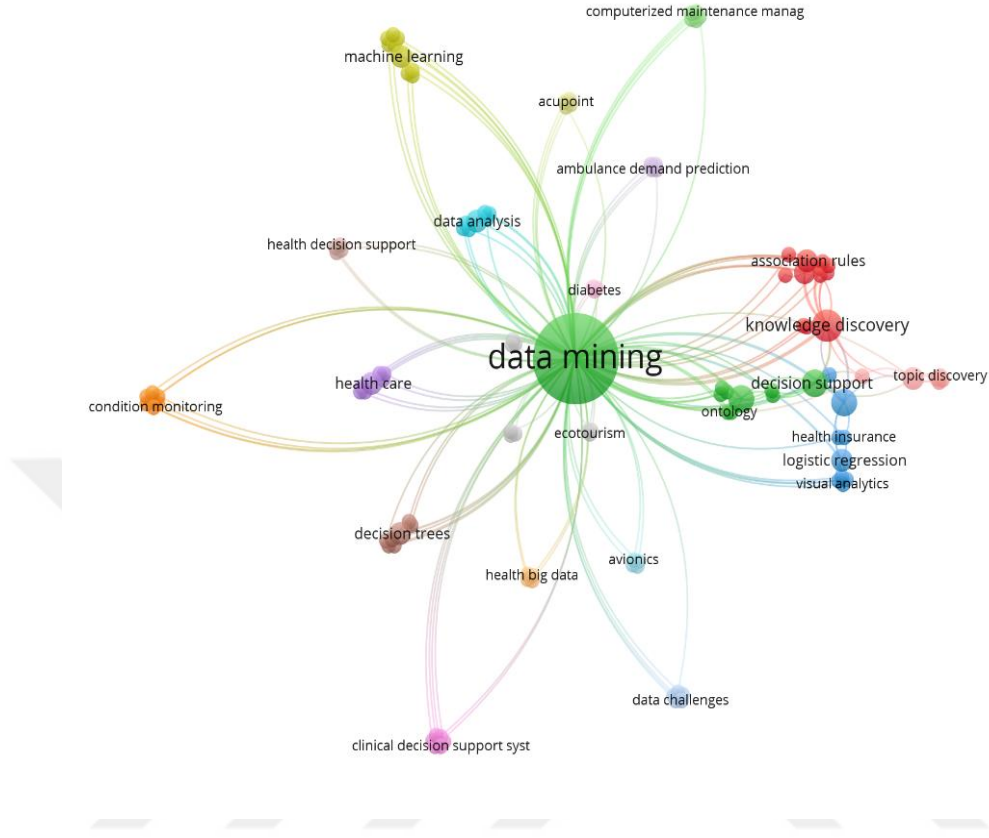
Şekil 1.2.' ye göre 2015 sonrası araştırmalarda artış trendi görülürken 2024 yılı itibari ile toplamda 5925 çalışmaya ulaşılmıştır. Sadece “makale” ve dil olarak “İngilizce” filtreleri tercihi sonrası çalışma sayısı 4311’e düşmüştür. 2023 yılında sınırlı bir düşüş olsa da son yıllarda artış trendinin devam ettiği görülmektedir.

Şekil 1.3: Ülke Bazında Konu Çalışma Ağırlığı



Şekil 1.3'e göre çalışmaların yapıldığı ülkeler araştırıldığında Amerika, Hindistan, İspanya ve Çin gibi ülkelerde yapılan araştırmaların yoğunluğu gözlenmektedir. Avrupa'da ise özellikle İspanya, Slovenya ve İngiltere'de araştırmacılar konuya ilgi göstermektedir. Konu hakkında başat araştırmacılara sahip ülkeler içerisinde Türkiye de bulunmaktadır. Araştırmacıların Amerika ve Çin'de yoğunlaşmasının sebebinin, bu ülkelerdeki bilgi teknolojisi altyapısında yaşanan hızlı değişimler olabileceği düşünülmektedir.

Şekil 1.4: Konu Çalışma Alanları



Sağlık yönetimi ve veri madenciliği ile ilgili olarak Şekil 1.4'te görüleceği üzere bilgi keşfi, makine öğrenmesi, birliktelik kuralları, karar destek sistemleri, sağlık hizmetleri, sağlık karar destek sistemleri gibi konularda çalışmalar olduğu görülmektedir

Bibliyometrik analiz, anahtar kelimeler yoluyla veri tabanlarından elde edilen bilginin derlenip yorumlanmasında kullanılmaktadır. Analiz ile çalışmaların hangi bölgelerde yoğunlaştığı ve konunun boyutları ortaya konulmuş, sağlık alanında veri madenciliği yöntemlerinin giderek daha fazla kullanıldığını ve değerli bir araştırma konusu olduğunu gösteren sistematik bir yaklaşım önerilmiştir.

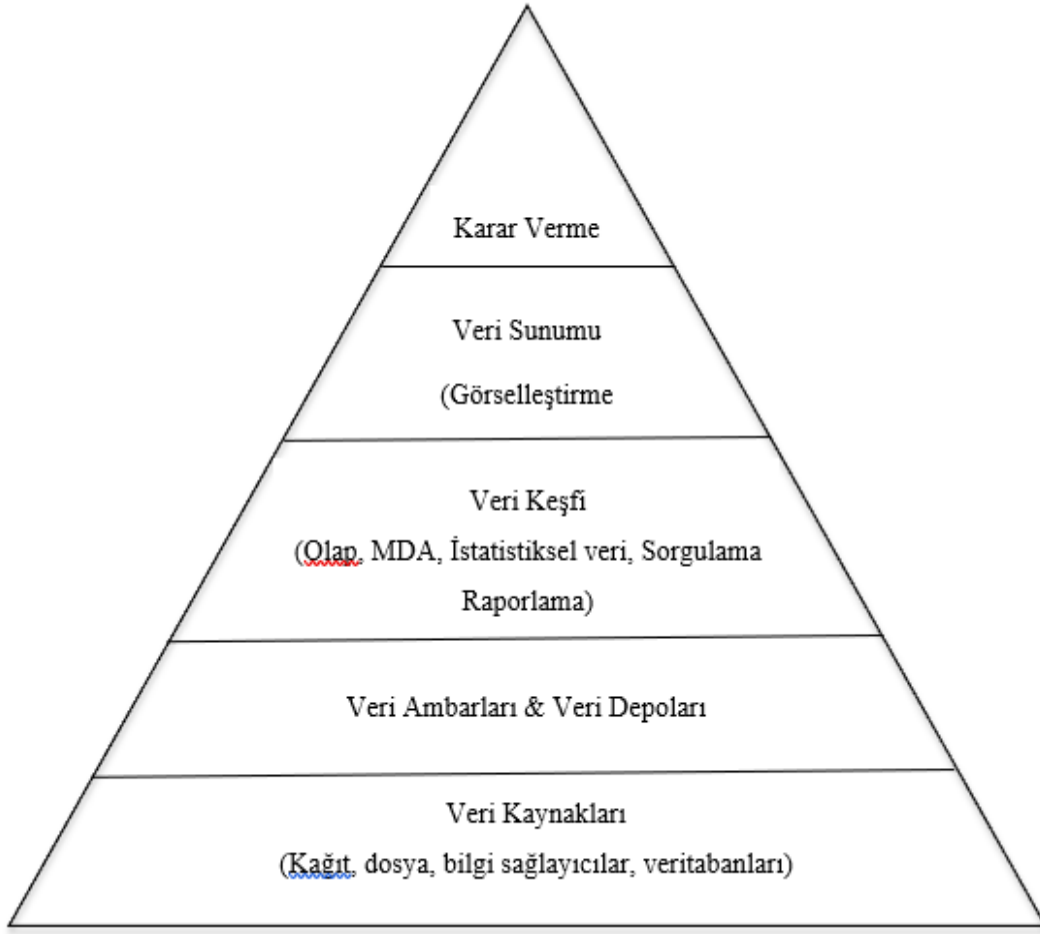
İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

Veri madenciliği, Tukey (1977) tarafından tanımlanan keşfedici veri analizi araştırmalarına dayanmaktadır. Klasik istatistiksel yöntemlerin katı varsayımları, veri setlerinin tam olarak analiz edilmesine izin vermediğinden, veri örüntüleri arasındaki ilişkileri ortaya çıkarmak için yeni teknikler tanıtılmakta ve bu alanda her geçen gün keşfedici ve sağlam yöntemler geliştirilmektedir (Zerman, 2018). Veri madenciliği kavramı, doksanlı yıllarda bilişim teknolojilerinin hızla gelişmesiyle beraber bilgisayar mühendisleri tarafından ortaya atılmıştır (Karaibrahimoğlu, 2014).

Veri madenciliği, 1990'lı yıllardan bu yana paralel veri depolama araçları, manyetik etiketler, barkodlar ve RFID teknolojisinin geliştirilip kullanıldığı bir alandır. Bu nedenle, kullanım yeri ve zamanına bağlı olarak tanımlar farklılık göstermektedir. Alandaki gelişmeler her geçen gün devam ederken, bugün oluşturulan bir tanım yarı yetersiz kalabilmektedir. Veri madenciliği beklenen veya hesaplanabilir sonuçları ortaya koymak için kullanılabilecek bir çözüm süreci olmaktan çok, önceden hesaplanamamış veya düşünülmemiş sonuçlar üretmesi açısından diğer yöntemlerden farklıdır. Kısaca verilerden bilgi keşfetme sürecine veri madenciliği denilmektedir (Jothi vd, 2015). Rekabet ortamında şirketler, rakiplerinin daha önce düşünmediği şeyleri düşünerek rakiplerinin önünde kalmaya çalışmaktadır (Karaibrahimoğlu, 2014). Veri Madenciliği ve iş zekası genel olarak, piramitte aşağıdan yukarıya çıkıldıkça karar vermeyi destekleyen enformasyonun değeri artmaktadır. Bu Şekil 2.1'de görülmektedir.

Şekil 2.1: Veri Madenciliği ve İş Zekâsı Piramidi



Kaynak: Koyuncugil, (2009)

Veri madenciliği araçları, sonuçların planlı uygulanmasında yöneticiler, karar vericiler ve paydaşlar için, iyi yapılandırılmış süreç destek sistemleridir. Veri madenciliğinin amacı, çeşitli sayısal teknikler kullanılarak toplanan bilgileri ilgili kurumsal ve idari destek sistemlerinde faydalanmak üzere araştırmak ve değerlendirmektir.

Veri madenciliği nispeten yeni bir alan olmasına rağmen, kökleri yaklaşık 30 yıl öncesine dayanan bir araştırma ve uygulama geleneğine sahiptir. Bu dönemde SAS, SPSS ve IBM gibi şirketler istatistiksel analize öncülük etmişlerdir. Günümüzde bu şirketler, veri madenciliği alanında oldukça aktiftir ve uzun yıllara dayanan deneyimlerine dayanarak yüksek itibarlı ürünler geliştirmişlerdir. Bu ürünler, çeşitli sektörlerde veri analizi ve iş zekâsı çözümleri sunarak büyük bir değer yaratmaktadırlar (Koyuncugil, 2009).

Sumathi ve Sivanandam (2006)'a göre veri madenciliği aşağıdaki tanımlara uymaktadır:

- Değerli iş verilerinde yeni ilişkilerin ve durumların otomatik olarak aranmasıdır.
- Büyük veri tabanlarından önceden bilinmeyen değerli ve faydalı bilgilerin keşfedilmesi ve önemli iş kararlarının alınmasında kullanılması sürecidir.
- Düzenli tümevarım yoluyla, veri tabanındaki anlaşılabilir kalıpların ortaya çıkarılmasıdır.

2.1. Veri Madenciliği Teknikleri

Veri madenciliği teknikleri veri tabanlarındaki büyük miktarlardaki verilerden yararlı bilgilerin keşfedilmesini amaçlamaktadır. Terabaytlarla ölçülen günümüz veri tabanlarındaki büyük miktarda veri, stratejik öneme sahip hassas bilgiler içermektedir. Ancak en önemli soru, bu kadar büyük miktarda verinin içerdiği önemli bilgilerin nasıl derleneceğidir. Geliri artırıcı ve maliyetleri azaltıcı perspektife sahip veri madenciliği, büyük miktarda gömülü veriden mantıksal bilgiyi hızlı bir şekilde elde edebilmektedir. Veri madenciliği aynı zamanda makine öğrenimi, istatistik, yapay zeka ve çeşitli görselleştirme teknikleri ile de kullanılabilir (Ertuğrul vd. 2013).

Veri madenciliği teknikleri, ilgisiz verileri çeşitli algoritmalar aracılığıyla birbirine bağlamaktadır. Bu yönüyle veri madenciliği, özel sektörde müşteri analizleri ihtiyacına cevap verebilmektedir. İşletmelerde veri madenciliği sistematiğine ait adımlar takip edildiğinde, etkili geri bildirimler elde edilerek çeşitli projeler planlanabilmektedir. Büyük veri setlerini anlamlı bilgilere dönüştürmek için kullanılan istatistiksel veri çıkarma yöntemleri, verilerin sorgulanması ve raporlanması olarak görülmemelidir. Veri madenciliği, karmaşık veri kümelerinden arzu edilen sonuçları elde etme sürecidir. Bir veri kümesi analiz edilirken, amacın mevcut verilerden içgörüler elde etmek olduğu unutulmamalıdır. Aynı şekilde, gelir ve cinsiyet arasındaki ilişkinin basit regresyon analizi ile incelenmesi de veri madenciliği olarak kabul edilmemelidir (Karabulut, 2021). Veri madenciliği astronomi, biyoloji, finans, pazarlama, sigorta, tıp ve daha birçok alanda başarıyla kullanılmaktadır (Akpınar, 2000).

2.2. Bilgi Keşfi

Bilgisayar destekli analiz, analistin toplanan verileri bir istatistiksel pakete veya elektronik tabloya girmesini gerektirmektedir. Kaydedilen verilerden anlamlı bilgilerin çıkarılması sürecine ise veri tabanı bilgisi keşfi adı verilmektedir (Taş, 2018). Veri tabanlarında bilgi keşfi, saklanan verilerden yeni, geçerli, kullanışlı ve anlaşılır kalıpların sonuçlanması sürecidir. Bu süreçteki ilk adım uygulama alanını anlamakla başlamaktadır. Bir veri tabanındaki bilgiyi keşfetmenin son adımı, ortaya çıkan bilgiyi, bilgi görüntüleme ve sunum yöntemlerini kullanarak kullanıcıya sunmaktır. Veri madenciliği, bilgi keşif sürecinin önemli bir unsurudur (Taş, 2018). Bilgi keşfi süreçleri temel olarak şu beş unsurdan oluşmaktadır: Veriler, ön işleme, dönüşüm, yorumlama/değerlendirme ve veri seçimi.

- **Veriler:** Veriler, veri tabanlarından veya analiz yapılmak istenen diğer veri kaynaklarından derlenerek anlamlı veri dosyası içerisinde oluşturulmuş olmalıdır.
- **Veri Seçimi:** Veri seti içinde analize uygun olan verilerin seçilerek, analize uygun hale getirilmesi sürecidir.
- **Ön İşleme:** Anormal verilerin çıkarılması, eksik veya hatalı verilerin düzeltilmesi gibi işlemleri içermektedir. Bu aşamada gerekli olmayan bilgiler kaldırılmaktadır. Bu işlem veri temizleme olarak da adlandırılabilir (Ahmad vd., 2015).
- **Dönüşüm:** Verilerin sınıflandırılması, ilgili özelliklerin seçilmesi veya boyutluluğun azaltılmasıdır.
- **Yorumlama/Değerlendirme:** Verilerin analiz edilerek yorumlanması ve sonucun ortaya çıkması sürecidir.

Veri madenciliği, bilgi keşfi araştırmasının bir parçasıdır ve istatistik, makine öğrenimi ve örüntü tanıma gibi araştırma alanlarıyla iş birliği içerisinde. Veri madenciliği modelleri genel olarak tahmine dayalı ve tanımlayıcı olmak üzere ikiye ayrılmaktadır. Tahmin modelleri, bilinen verileri kullanarak bilinmeyen bir değerini tahmin etmeye çalışmaktadır. Tanımlayıcı modeller ise algoritmalar vasıtasıyla veride gizli olan ortak özellikleri ve ilişkileri aramaktadır (Yıldırım, vd., 2008). Bu süreç istatistiksel ve matematiksel prensipler, çeşitli teknikler ve uygulamalar kullanılarak gerçekleştirilmektedir (Arslan, 2008: 3).

2.3. Veri Madenciliği Uygulama Alanları

Birçok sektörde kullanılmakta olan veri madenciliği, e-ticaret, bankacılık, pazarlama, sigorta ve sağlık hizmetleri gibi alanlarda kullanılmaktadır. Veri madenciliğinin kullanılması büyük veri ambarlarının oluşturulmasına olanak sağlarken, daha doğru sonuçlara ulaşmak için büyük ölçüde yardımcı olmakta ve bu alanlarda giderek daha popüler hale gelmektedir (Çerkezi 2013). Aşağıda veri madenciliğinin farklı alanlardaki kullanım tarzlarından bahsedilmiştir.

- **Pazarlama Yönetimi:** Veri madenciliği pazarlama amaçlarına yönelik olarak müşteri satın alma davranışının belirlenmesi, demografik bilgiler, kampanya ürünleri, mevcut müşterileri kaybetmeden yeni müşteri kazanımı, alışveriş sepeti analizi, müşteri ilişkileri yönetimi, müşteri tahmin analizi, satış tahmini, internet sitesi ziyaret analizleri gibi konularda sıklıkla kullanılmaktadır (Poyraz, 2012; Zerman, 2018).
- **Bankacılık Yönetimi:** Finansal açıdan çapraz satış, risk değerlendirmesi, maliyetleri düşürme, kayıpları ve sızıntılarını önleme, alternatif kanallar oluşturma açısından veri madenciliği yöntemleri kullanışlı araçlardır. Hangi müşteri profilinin neyi, ne zaman ve neden tercih ettiğini anlayabilen kuruluşlar, talep yaratma ve doğru talebi doğru zamanda karşılayıp sunma avantajına sahiptir. Bankalar, rakip bankaya geçme ihtimali olan mevcut müşterilerini ve kaybedilen müşterilerin neden sistemden ayrıldığını belirlemek gibi spesifik hususlarda dahi veri madenciliği yöntemlerinden faydalanabilmektedirler (Çerkezi, 2013).
- **Sigortacılık Yönetimi:** Sigortacılık alanında veri madenciliği, sözleşmelerini yenileyecek veya yeni bir sözleşme imzalaması muhtemel olan müşterileri belirlenmesi, sigorta dolandırıcılığının tespit edilmesi, risk altındaki müşterilerin belirlenmesi gibi birçok konuda çözüm üretmek amacı ile kullanılmaktadır (Karabulut, 2021).

2.4. Diğer Uygulamalar

Veri madenciliği haberleşme, ulaşım, sağlık vb. birçok alanda kullanılmaktadır.

- **Sinyal işleme:** Telefon konuşmalarına müdahale ve ses gürültüsü nedeniyle gelecekte oluşabilecek kayıpların ortadan kaldırılması ve veri analizi yapabilmek için kullanılmaktadır.
- **Medikal:** Hasta adayı verilerinin veri madenciliği algoritmaları ile analiz edilerek tanı ve tedavi alanında araştırmalar yapmak amacı ile kullanılmaktadır.
- **Tıp:** Ameliyat riski taşıyan hastalar için, ameliyat öncesinde ameliyat gerekip gerekmediğini kolayca belirlemek için veri madenciliği teknikleri için kullanılmaktadır.
- **Kimya:** Yeni ilaçların, yeni kimyasal maddelerin keşfinde ve çeşitli analizlerde kullanılmaktadır.
- **Uzay bilimi ve teknolojisi:** Gezegenlerin takibi ve takibinden elde edilen veriler gibi birçok alanda kullanılır. Gezegenlerin yüzeylerinin şeklini, yaşam alanlarını, konumlarını ve hatta yeni galaksilerin keşfini keşfetmek için kullanılmaktadır.

Ayrıca elektronik cihaz için parmak izi tanıma, yüz şekline dayalı tanımlama, insan sesinin komut olarak kullanılması gibi alanlarda kullanılan yapay zekâ teknikleri de veri madenciliğinde uygulanmaktadır (Zerman, 2018).

2.5. Sağlıkta Veri Madenciliği Kullanımı

Veri madenciliği, kronik hastalıklara ilişkin tıbbi ilaçların tüketim alışkanlıklarının ve risklerinin belirlenmesi, tıbbi süreçlerin optimize edilmesi, ilaç hataları durumunda hata sinyallerinin sağlanması gibi teknik sağlık problemlerinin yanı sıra sağlık personelinin performansının izlenmesi, sağlık hizmetleri verimliliğinin incelenmesi gibi yönetsel kararlarda da kullanılabilir. Sağlık sektöründeki veri madenciliği aynı zamanda organizasyonel maliyetlerin etkili bir şekilde azaltılmasına da yardımcı olmaktadır. Sağlık sektöründe hastane hizmeti almak isteyen hastaların sistemde kalış sürelerini kısaltarak hastane maliyetlerini düşüren işletmeler, maliyet kontrolünden faydalanarak bu alandaki rekabetçi güçlerini artırabilmektedir (Karabulut, 2021).

2.6. Veri Madenciliği Modelleri

Tahmine dayalı ve tanımlayıcı modeller, veri madenciliğinde yaygın olarak kullanılan sınıflandırma türleridir. Bu veri madenciliği modelleri, farklı amaçlarla kullanılmakta olup birçok algoritma ve teknik içermektedir. Tanımlayıcı modeller, veriler arasındaki gizli ilişkileri ortaya çıkarırken; tahmine dayalı modeller, gelecekteki olayları tahmin etmek için kullanılır. Tahmine dayalı modeller, "Bu işlemde dolandırıcılık var mı?" veya "Bu müşteriden ne kadar gelir elde edebiliriz?" gibi sorulara yanıt aramaktadır. Bu modeller, mevcut verilerden yararlanarak gelecekteki sonuçları tahmin etmeye çalışmakta ve genellikle makine öğrenmesi algoritmaları ile desteklenmektedirler.

Tablo 2.1: Veri Madenciliğinin Modelleri

Veri Madenciliği		
Tahmin Edici Modeller		Tanımlayıcı Modeller
Sınıflandırma	Regresyon	Kümeleme
Karar Ağaçları	Lojistik	Birliktelik Kuralları
Yapay Sinir Ağları	Doğrusal	
K-En Yakın Komşu		

Kaynak: (Akgün vd. 2022)

Tanımlayıcı modeller ise, veriler arasındaki gizli ilişkileri ve örüntüleri ortaya çıkarmayı amaçlamaktadırlar. Tanımlayıcı bir model, "Bebek bezi alan müşterinin bebek maması satın alma olasılığı 3 kat artmaktadır" gibi sonuçlara ulaşılmasını sağlamaktadır. Bu modeller, veri kümesinde mevcut olan bilgileri analiz ederek anlamlı ve kullanışlı bilgiler sunmaktadırlar. Veri madenciliğinde kullanılan algoritmalar ve teknikler hem tanımlayıcı hem de tahmine dayalı modeller içerebilmekte ve bu modellerin kombinasyonları ile daha kapsamlı analizler yapılabilirler. Bu sayede, işletmeler ve kuruluşlar, verilerden elde edilen içgörülerle stratejik kararlar alabilmektedirler. (Zerman, 2018). Tablo 2.1’de veri madenciliği tekniklerinin sınıflandırılmasını göstermektedir.

2.6.1. Tahmin Edici Modeller

Tahmine dayalı modeller mevcut verileri kullanarak modeller oluşturmakta, ardından bu modelleri büyük, keşfedilmemiş veri kümelerine uygulamakta ve sonuçlara bağlı kalarak tahmine dayalı analizler yapmaktadırlar (Karabulut, 2021).

Zerman (2018)'a göre tahmin edici modeller şu şekilde ayrılmaktadır:

- Sınıflandırma
- Regresyon
- Karar ağaçları
- Yapay sinir ağları
- K-en yakın komşuluğu
- Lojistik
- Doğrusal

Genel olarak sınıflandırma ve regresyon modelleri, diğer tahmin edici modeller içerisinde sıklıkla kullanılmaktadırlar.

2.6.1.1 Sınıflandırma

Sınıflandırma işlemi insanın düşünce yapılarına en uygun veri madenciliği teknikleri arasındadır. İnsanlar çevrelerindeki nesneleri ve olayları daha iyi anlamak ve başkalarına anlatmak için hemen hemen her şeyi sınıflandırma eğilimindedirler, veri madenciliğinde ise sınıflandırma tekniği, veri setini anlamak ve yorumlamak açısından çok fazla kullanılmaktadır (Ahmad vd., 2015; Jothi vd., 2015). İnsanların davranışlarına göre, hayvanların türlerine göre, evlerin ise görünüşlerine göre benzer gruplar içerisinde toplanması çevredeki unsurların farklı özellikleri bakımından sınıflandırılabilirliğini göstermektedir (Poyraz, 2012). Tıp alanında sınıflandırmalar, semptomlara ve sağlık koşullarına dayalı olarak tıbbi teşhisleri ve prognozları tanımlamak için kullanılabilmektedirler (Yoo vd., 2012).

Veri madenciliğinde sınıflandırma, mevcut verilerin belirlenen özelliklere göre sınıflara bölünmesi ve yeni verilerin hangi sınıfa dahil edilmesi gerektiğinin belirlenmesini içermektedir. Bir başka deyişle bu işlem yeni bulunan bir girdinin hangi

sınıfa ait olduğunun belirlenmesi olarak da ifade edilebilmektedir. Veri madenciliği sınıflandırma uygulamalarına örnek olarak bankaların kredi başvurularını düşük, orta ve yüksek riskli olarak sınıflandırması ve yeni bir öğrencinin okula hangi sınıfa gitmesi gerektiğinin belirlenmesi verilebilir (Poyraz, 2012).

2.6.1.2. Regresyon (Regression)

Regresyon, veri madenciliğinde fonksiyonları tanımlamayı kolaylaştıran önemli bir tekniktir. Temelde matematiksel bir araç olmasının yanında farklı varyasyonlar arasındaki ilişkileri göstermekte oldukça faydalıdır (Ahmad vd., 2015). Regresyon analizi, tahminlerde bulunmak amacıyla matematiksel formüllerin mevcut verilere uygulanması için kullanılmaktadır. Doğrusal regresyon veya lojistik regresyon teknikleri kullanılarak mevcut veriden bir fonksiyon belirlenmekte ve yeni veriler, mevcut işlevine uygulanarak tahminlerde bulunmaktadır. Değerleri bilinen değişkenler diğer değişkenleri tahmin etmek için kullanılmaktadır. Regresyon terminolojisinde tahmin edilen değişkene "bağımlı değişken", bağımlı değişkeni tahmin etmek için kullanılan değişkene ise "bağımsız değişken" adı verilmektedir (Zerman, 2018).

2.6.2. Tanımlayıcı Modeller

Tanımlayıcı modeller, verilerdeki keşfedilmemiş unsurları ortaya çıkarmayı ve bir veri setindeki değişkenler arasında analiz sonuçlarını etkileyen ilişkilerin olup olmadığını belirlemeyi amaçlamaktadırlar. Tanımlayıcı modeller birliktelik kuralları ve kümeleme modelleri olarak sınıflandırılmaktadırlar (Karabulut, 2021).

- **Birliktelik Kuralları:** Birliktelik kuralları, bir veri tabanındaki kayıtlar arasındaki ilişkileri inceleyerek hangi olayların bir arada meydana gelebileceğini keşfetmeye çalışan bir veri madenciliği tekniğidir. Birliktelik kurallarını kullanmanın en yaygın örneği alışveriş sepeti analizidir. Bu süreç, satın alınan ürünler arasındaki ilişkiyi keşfederek müşterilerin satın alma alışkanlıklarını analiz etmektedir. Bu tür birlikteliğin keşfedilmesi, müşterilerin hangi ürünleri birlikte satın aldığı hakkında bilgi sağlamakta ve pazar yöneticilerinin bu bilgilere dayanarak daha etkili satış stratejileri geliştirmesine olanak tanımaktadır (Çerkezi 2013).

- **Kümeleme Modelleri:** Kümeleme analizinin amacı veriyi benzer özelliklere göre gruplandırmaktır. Eldeki veriler incelenerek benzerlik kümeleri oluşturulmakta ve birbirine benzemeyen veriler başka bir kümeye yerleştirilmektedir. Bir grup fiziksel veya soyut verinin, benzer verilerin aynı sınıfta kalacak şekilde daha küçük gruplara bölünmesi işlemine kümeleme denilmektedir. Birçok uygulama aynı kümedeki veri kümelerini kullanarak daha kolay ve hızlı çalışabilmektedir.

Özellikle kümeleme günlük hayatta sıklıkla karşılaşılan bir tekniktir. Küme analizi, bilginin daha erişilebilir olmasını sağlamak için belgeleri World Wide Web (WWW) üzerinde gruplamak da dahil pazar analizi, örüntü tanıma, görüntü işleme ve veri analizi gibi alanlarda yaygın olarak kullanılmaktadır. Sınıflandırmadan farklı olarak kümeleme, sınıfın veya eğitim verilerinin önceden belirlenmesini gerektirmemektedir (Zerman, 2018).

2.7. Birliktelik Kuralları Analizi (Association Rules Analysis)

Birliktelik kuralı teknikleri, bilgisayar bilimi alanında geliştirilmiş ancak pazar sepeti analizi, belirli satırlar ve web tıklama akışı analizi gibi alanlarda daha yaygın olarak uygulanmıştır. Genel olarak amaç, bir grup sürecinde sıklıkla bir arada görünen parça gruplarını vurgulamaktır. Başka bir deyişle birliktelik kuralları, büyük veri kümelerindeki değişkenlerin özel değerleri arasındaki ilişkilerin ve bağlantıların keşfedilmesidir. Teknik, analistlerin ve araştırmacıların büyük veri setlerindeki gizli kalıpları bulmasına olanak tanımaktadır. Tekniğin güçlü yönleri kullanışlılığı, anlaşılabilirliği ve olası tüm modelleri keşfetme yeteneğidir ancak tüm olasılıkları ortaya çıkarması aynı zamanda onun zayıflığı olarak görülebilmektedir. Karar vericinin olasılıkları değerlendirmek için büyük miktarda bilgiyi yönetmesi zor ve zaman alıcı olabilmektedir (Alan, vd. 2019).

Birliktelik kuralı tekniğini uygularken iki hususun dikkate alınması gerekmektedir. İlk olarak, büyük işlemsel veri kümelerinden kalıpların çıkarılması, bilgisayar kaynakları ve hesaplama açısından çok maliyetli ve zaman alıcıdır. İkinci husus ise, bulunan bazı desenler tesadüfen oluşmuş ve sahte olabilmektedir. Bu amaçla keşfedilen örüntülerin de bir değerlendirme sürecine tabi tutulması gerekmektedir. Birliktelik kuralı tekniklerinin bir sonucu olarak çok sayıda kural üretildiğinden, buradaki amaç, bir dizi öncül ve ardıl işlemler arasında güçlü bağımlılıklar sergileyen kuralları

belirlemektir. Destek düzeyi, güvenilirlik ve büyüme oranı gibi metrikler, kuralların ifade ettiği bağıllık düzeyini ölçmek için kullanılmaktadır. Birliktelik kuralındaki güven ölçüsü, istatistik biliminde genel olarak kullanılan güven aralıkları ve güven kavramlarından bağımsızdır. Aynı şekilde kaldırma oranı kavramı da finanstaki kaldırma oranı kavramından farklı ve bağımsızdır (Irmak, S. 2009).

Birliktelik kurallarında; nesne, müşterinin satın aldığı üründür, işlem ise birlikte satın alınan tüm nesnelerin kümesidir. Birliktelik kurallarında birçok önemli terim sıklıkla kullanılmaktadır. Sonuçlar bu kuralın sol tarafını, önceki kuralın sağ tarafını temsil eder, destek değeri, güven değeri, minimum destek değeri Min_Support olarak görüntülenir, minimum güven değeri Min_Trust değeri olarak görüntülenir (Çarklı, 2010).

2.7.1. Birliktelik Kuralları Analizi Algoritmaları

Birliktelik kuralları Agrawal vd. (1993) tarafından ortaya konulduktan sonra, yöntem süreci içerisindeki işlemleri gerçekleştirmek için çeşitli algoritmalar geliştirilmiştir. Bu algoritmalarından bazıları aşağıda sunulmuştur. Bu algoritmaların en temeli AIS, en ünlüsü ise Apriori algoritmasıdır (Taş, 2018; Ahmad vd., 2015)

2.7.1.1. Apriori Algoritması

Apriori algoritması yinelemeli yapısıyla davranışsal bilgileri içeren bir veri tabanında sıklıkla tekrarlanan öge kümelerini keşfetmek için kullanılmaktadır. Bu algoritmaya dair veri madencilik süreci iki aşamadan oluşmaktadır. İlk olarak sık sık meydana gelen tüm öğeler bulunmakta, sonrasında ise bu sık görülen öğelerden güçlü birliktelik kuralları oluşturulmaktadır. Birliktelik kurallarının ilk aşamasında kullanılan Apriori algoritması, sık öge madenciliğinde kullanılan en yaygın ve klasik algoritmadır. Bu algoritmada özellikler ve veriler "Boolean" ilişki kuralları kullanılarak değerlendirilmektedir (Apriori, 2024, <https://cotur.medium.com/apriori-algoritması-a90a38856373>)

Apriori Algoritması içerisinde doğru ve yanlış değerlerinin olduğu Boolean veri tipleri için sık öge kümelerini çıkarmak amacıyla kullanılmaktadır. Sık öge kümesi üretimi sırasında ortaya çıkan aday öge kümesi sayısını azaltmak için destek önlemleri uygulanmaktadır. Bu prensibe göre, eğer bir element kümesi ortaksa, bu element

kümesinin tüm alt kümeleri de eşit derecede ortaktır (Irmak, 2009). Algoritma, yaygın olarak kullanılan alt kümeleri her seferinde bir öge olarak genişletmektedir. Bu adıma aday oluşturma süreci denilmekte ve grup adayları daha sonra verilere göre test edilmektedir. Apriori, aday kayıtlarını verimli bir şekilde saymak için birinci sınıf bir arama yöntemi ve karma ağaç yapısı kullanmaktadır. Bu algoritmada kullanılan verilerde her işleme ait veri kodu sayısaldır ve veri kodları en düşük değerden en yüksek değere doğru sıralanmaktadır (Zerman, 2018). Algoritma adayları veritabanına başvurmadan önceki bir dizi ortak ögeye dayalı olarak oluşturduğu durumlarda zaman kaybına sebep olan oldukça kısıtlı verim sergilemektedir (Ceylan, 2023).

Apriori algoritması, onu takip eden birçok algoritmayı etkilemiştir. Aşağıda Apriori algoritmasının kullanıldığı farklı uygulama alanlarından örnekler yer almaktadır (Demirok, 2019):

- Kredi kartı işlemleri,
- Telekomünikasyon hizmetleri,
- Bankacılık hizmetleri,
- Sigorta,
- Müşteri hizmetleri ve pazarlama hizmetleri,
- Tıp ve sağlık hizmetleri ilişkilendirme kuralı madenciliği

Birliktelik kuralları, veri madenciliğinde sıkça kullanılan yöntemlerden biridir. Bu kuralların başlangıçta kullanılan iki temel istatistiği destek (support) ve güven (confidence) değerleridir. Bu değerler sayısal terimlerdir ve tanımlanmaları için bazı temel kavramların açıklanması gerekmektedir. Birliktelik kurallarının tanımlanmasında kullanılan terimler şunlardır:

D: İşlemler veritabanı

N: Veritabanındaki işlem sayısı

I: Bir işlem, yani bir ürün kümesi

Bir ürün kümesinin destek değeri (support), o ürün kümesini içeren işlemlerin toplam işlem sayısına oranıdır. Destek değeri, X ürün kümesi için aşağıdaki şekilde hesaplanır:

$$\text{Destek}(X) = |\{I \mid I \in D \wedge I \supseteq X\}| / N$$

Burada, I bir işlem kümesi, $|\{I\}|$ ise bu kümenin eleman sayısını ifade etmektedir.

Birliktelik kurallarının destek değeri, hem kuralın öncülü (antecedent) hem de sonucu (consequent) olan işlemlerin toplam işlem sayısına oranıdır. Güven değeri ise, öncül kümesini içeren işlemlerin aynı zamanda sonuç kümesini de içermesi oranıdır. Bir $A \Rightarrow C$ birlikteliği için destek ve güven değerleri şu şekilde tanımlanır (Webb, 2003):

$$\text{Destek}(A \Rightarrow C) = \text{Destek}(A \cup C)$$

$$\text{Güven}(A \Rightarrow C) = \text{Destek}(A \cup C) / \text{Destek}(A)$$

Eğer destek değeri yeterince yüksekse (ve işlemler, gelecekteki işlemler ile aynı veri dağıtımından rastgele bir örneği temsil ediyorsa), bu durumda güven seviyesi, kuralın öncül kümesini içeren herhangi bir gelecekteki işlemin, aynı zamanda sonuç kümesini de içereceği ihtimalinin makul bir tahminidir (Alan, vd. 2019).

Birliktelik kurallarında destek (support) ve güven (confidence) değerlerinin yanı sıra önemli bir diğer hesaplama ise "lift" değeridir. Lift değeri, bir ürün kümesinin diğer bir ürün kümesiyle olan ilişkisinin gücünü ölçer ve aşağıdaki şekilde hesaplanır:

$$\text{Lift}(A \Rightarrow C) = \text{Güven}(A \Rightarrow C) / \text{Destek}(C)$$

Burada, $\text{Güven}(A \Rightarrow C)$, A kümesini içeren işlemlerin aynı zamanda C kümesini de içermesi oranıdır ve $\text{Destek}(C)$ ise C kümesini içeren işlemlerin toplam işlem sayısına oranıdır.

Lift değerinin yorumlanması şu şekildedir:

- Lift > 1: A kümesinin görülmesinin C kümesinin görülmesi üzerinde pozitif korelasyona sahip olduğunu gösterir. Bu durum, A ve C kümelerinin birlikte görülme olasılığının bağımsız olmaktan daha yüksek olduğunu ifade eder.

- Lift < 1: A kümesinin görülmesinin C kümesinin görülmesi üzerinde negatif korelasyona sahip olduğunu gösterir. Bu durumda, A ve C kümeleri birlikte görölme olasılığı açısından bağımsız olmaktan daha düşüktür.

- Lift = 1: A ve C kümelerinin birlikte görölme olasılığının birbirinden bağımsız olduğunu ifade eder. Yani, A ve C kümelerinin birlikte görülmesi, rastgele bir dağılımla aynıdır.

Lift değeri, birliktelik kurallarının analizinde önemli bir metriktir çünkü iki ürün kümesi arasındaki bağımlılık derecesini ve yönünü belirler (Taş, 2018:37-38).

2.7.1.2. AIS Algoritması

AIS algoritmasının odak noktası, veri tabanının kalitesinin ve karar destek isteklerini işlemek için gereken işlevselliğin iyileştirilmesidir. Bu algoritma yalnızca bir ögesinden kaynaklanan bir birliktelik kuralı oluşturmaktadır. Başka bir deyişle, bu kurallar yalnızca bir kayıtla sonuçlanmaktadır. AIS algoritması, aday küme tanımlama adımını gerçekleştirmek için budama tekniğini kullanmaktadır. Budama tekniğinin temeli aday kümeler arasında yer alan gereksiz kümenin tespit edilerek ortadan kaldırılmasıdır. Bu adım tamamlandıktan sonra, her aday kümesi için destek hesaplanmaktadır. Minimum destek seviyesinin üzerinde hesaplanan destek seviyeleri, yayılmış küme olarak işaretlenmektedir. Algoritma bir sonraki taramada, sonraki aday kümeleri belirlemek için bu ortak imzaya sahip kümeleri kullanmaktadır (Taş, 2018).

2.7.1.3. SETM Algoritması

SETM algoritması, büyük miktarlarda veriyi işlemek için SQL altyapısından faydalanmaktadır. Bu algoritmada büyük veri setinin her elemanı iki parametreden oluşmaktadır. Biri verinin adı, diğeri ise bu veriyi ayırt etmeye yardımcı olan benzersiz özellik numarasıdır. SETM algoritması, veritabanı taranır taranmaz aday kayıtları oluşturmakta ancak bunlar geçişin sonunda sayılmaktadır. Daha sonra AIS algoritmasıyla aynı şekilde yeni bir aday kayıt oluşturulmakta, ancak üretim prosesinin proses ID'si, TID dizi yapısı içerisinde bir aday kayıt olarak kaydedilmektedir. Bu işlemler aday oluşturma sürecini, sayımdan ayırmaktadır. Çalıştırma sonunda sıralı yapıların birleştirilmesiyle desteklenen aday kayıt sayısı belirlenmektedir (Zerman, 2018)

2.7.1.4. FP-Büyüme Algoritması

FP büyüme algoritması, sistem kaynaklarını verimli bir şekilde kullanabilen ve büyük veri kümelerini hızlı bir şekilde işleyebilen ortak nesne kümelerini bulmak için kullanılan bir birliktelik kuralı algoritmasıdır. Algoritma, veri tabanının tamamının daha yoğun ve daha küçük bir veri yapısı olan frekans desen ağacında tutulması sayesinde önceki muadillerinden daha verimli performans gösterebilmektedir. FP büyüme algoritması, öncelik tabanlı algoritmalarından farklı olarak tüm veri tabanını yalnızca iki kez taramaktadır. İlk tarama tüm öğeler için destek değerini hesaplarken, ikinci tarama bir ağaç veri yapısı oluşturmaktadır. Bu algoritma büyük veritabanları için kullanışlıdır çünkü her seferinde aday nesneden oluşan yeni bir küme oluşturarak veri tabanını tarama ihtiyacını ortadan kaldırmaktadır (Taş, 2018).

ÜÇÜNCÜ BÖLÜM

SAĞLIK İŞLETMELERİNDE VERİ MADENCİLİĞİ

UYGULAMALARI

3.1. Çalışmanın Amacı ve Önemi

Sağlık alanında veri madencilik uygulamalarının kullanılması, hastaların hastane kayıtlarındaki verilerin yorumlanması sonucunda hastaya ve doktora karar destek konusunda yardımcı olmaktadır. Geçmiş ve yakın zaman tüm veriler yeni bilgi keşiflerini kolaylaştırmakta ve hekimlere hızlı karar vermelerinde destek olmaktadır. Bu çalışmada, 2023 yılında Erzincan Binali Yıldırım Üniversitesi Mengücek Gazi Eğitim ve Araştırma Hastanesini ziyaret eden hastaların vizit hareketleri kayıt bilgileri kullanılarak veri madenciliği algoritmaları ile analiz uygulaması gerçekleştirilmiştir. Hastanenin hasta demografik yapısının ortaya konulması ve verilen hizmetler doğrultusunda ilişkili birimlerin birbirleri ile yakın konumlandırılmasının hastane hizmet performansı ve süreç verimliliği üzerinde olumlu etkisinin olacağı düşünülmüştür. Hasta demografik bilgileri doğrultusunda ihtiyaç duyulacak poliklinik sayısının belirlenmesi bir kapasite problemi olarak da değerlendirilebilir. Bu açıdan bakıldığında çalışmanın çok yönlü bir çözüm perspektifine sahip olduğu düşünülmektedir.

Çalışmanın amacı; sağlık uygulamaları için veri madenciliği ve birliktelik kuralları ile istatistiksel çözümleme sağlamak ve hastanenin poliklinik hizmetleri için daha sürdürülebilir ve verimli sağlık hizmeti sağlamaktır. Özellikle Apriori algoritması kullanılarak veri setleri analiz edilmiş ve poliklinik ilişkileri hakkında anlamlı desenler elde edilmiştir.

3.2. Problemin Tanımlanması

Çalışmada hastane veritabanından sağlanan veriler ışığında öncelikle MySQL sorgulamaları ile poliklinik hizmetlerine dair demografik tespitler ortaya konulmuştur. İkinci aşamada ise Apriori uygulamasından faydalanarak hastaların farklı polikliniklerde aldığı sağlık hizmetleri arasındaki keşfedilmemiş ilişki bağlantıları tespit edilmiştir. Araştırma modeli, betimsel ve analitik bir desene dayanmaktadır. Betimsel model, sağlık işletmelerindeki mevcut durumu ortaya koyarken, analitik model veri madenciliği tekniklerinin uygulanması ve sonuçların yorumlanmasını kapsamaktadır. Kuruma ait otomasyon sisteminden, 2023 yılına ait ziyaret hareketleri, bu hareketlerin tanısı, demografik bilgiler alınmıştır. Elde edilen bu veriler, veri madenciliği teknikleri kullanılarak analiz edilmiştir. Analiz sürecinde, Python programlama dili ve veri analizi yazılımları kullanılmıştır. Bu yazılımlar aracılığıyla, Apriori algoritması gibi veri madenciliği teknikleri uygulanmış ve elde edilen sonuçlar yorumlanmıştır. Ayrıca, veri görselleştirme araçları kullanılarak analiz sonuçları görsel olarak sunulmuştur. Araştırmanın yürütülmesi için kullanılan materyaller şunlardır:

- Sağlık işletmesinden elde edilen elektronik hasta kayıtları ve yönetimsel veriler
- Python programlama dili ve ilgili kütüphaneler (pandas, numpy, matplotlib, seaborn)
- Veri madenciliği algoritmaları ve yazılımları
- Sağlık Bakanlığı raporları
- Veri görselleştirme araçları (ör. Matplotlib)

3.3. Kullanılan Algoritma ve Programlar

3.3.1. MySQL Programı

MySQL, web tasarım ve programlama sektöründe sıklıkla kullanılan açık kaynak kodlu veritabanı yönetim sistemidir. SQL, MySQL veri tabanının çekirdek alt yapısını oluşturmaktadır. MySQL ve SQL birbirine sıkça karıştırılmasına rağmen iki kavram aynı olmayı ifade etmemektedir. Aşağıda MySQL ve SQL arasındaki bazı farklar sunulmuştur:

- SQL, temel olarak kullanılan veri tabanlarının sorgulama dilidir.
- MySQL, ilk açık kaynak kodlu veritabanı dilidir.
- MySQL, içerisinde bulunan verilerin korunmasını amaçlayan kapsamlı bir ilişkisel veri tabanıdır.
- SQL, yapısal bir sorgu içeren dildir.
- MySQL, içerisindeki verileri silmeye, düzenlemeye, eklemeye izin verir.
- MYSQL veritabanı yazılımı dili iken SQL ise sorgu dilidir.

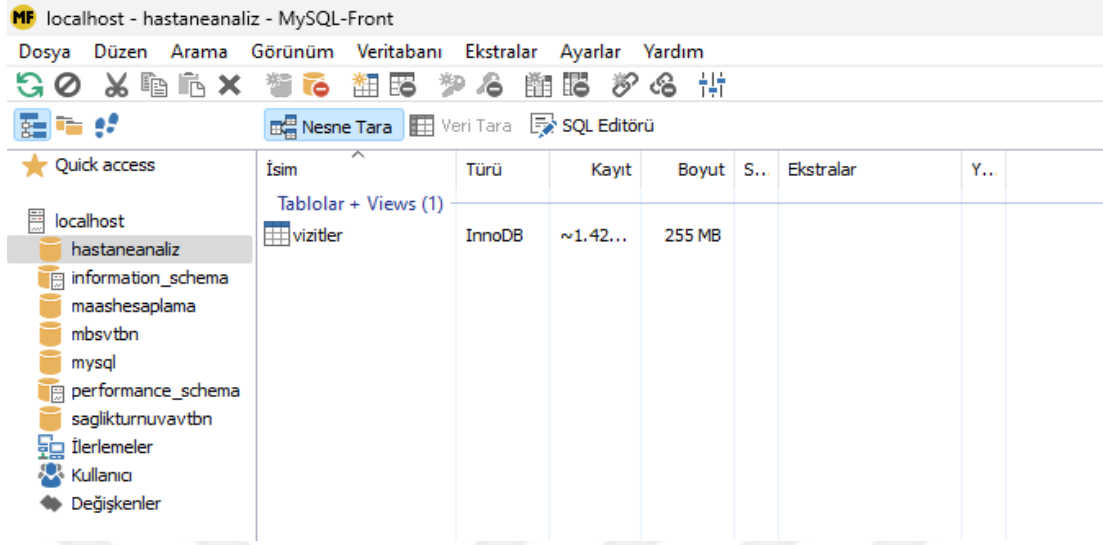
MySQL, ücretsiz olan kullanılabilen açık kaynak bir veri tabanıdır. Bununla birlikte, gelişmiş özellikleri bulunan ücretli bazı sürümleri de bulunmaktadır. Microsoft SQL Server ve Prophet gibi diğer veri tabanlarına kıyasla kullanımı basittir. Tüm programlama dilleriyle kullanılabilen MySQL, büyük verilerde bile yüksek performans sunabilmektedir. Tablo ve veritabanı boyutunda da herhangi bir limit bulunmamaktadır (MYSQL. (b.t.) 2024, <https://www.protan.com.tr/MySQL-nedir/>).

3.3.2. MySQL Front Programı

MySQL bir veritabanı olduğundan dolayı veri oluşturma, kayıt, silme, düzenleme gibi birçok işlemi SQL kodları ile yapmak gerekmektedir. SQL veritabanları yönetim programları ile daha hızlı bir şekilde kullanabilmekte ve bu birliktelik ile etkili işlemler yapılabilmektedir. MySQL veritabanı için MySQL Front programından yararlanılmaktadır. Program yakın zaman içerisinde güncelleme ve destek hizmetleri almamasına rağmen 6.1 sürümü ile SQL komutları kullanılmaktadır.

Çalışmamızda sağlık kurumundan Excel dosyası olarak temin ettiğimiz veri setinin, hızlı bir şekilde ön işleme, veri temizleme gibi bilgi keşfi aşamalarını gerçekleştirmek için MySQL ve SQL komutları kullanılmıştır. MySQL veri tabanında işlemler yapabilmek için MySQL Front SQL editörü mevcuttur. SQL editörü kullanarak verilerimizin istatistiki bilgileri alınmış ve verilerin istenilen formata çevrilmesi sağlanmıştır.

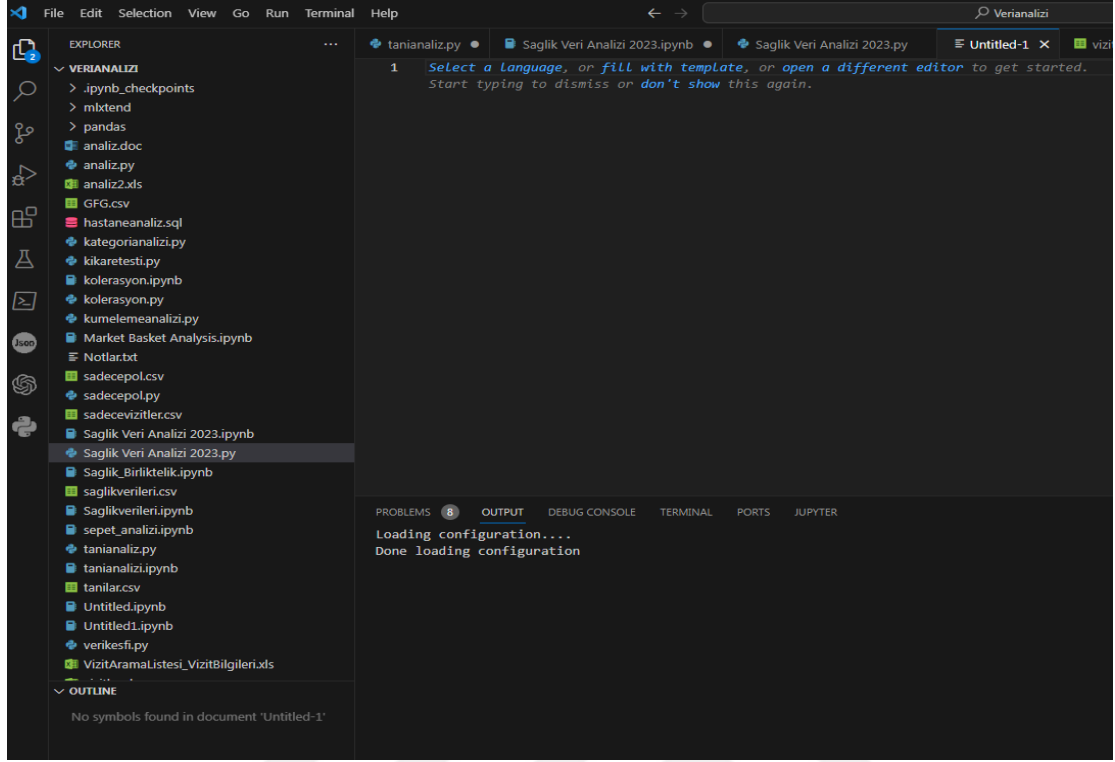
Şekil 3.1: Mysql Front Programı



3.3.3. Python Programı

Python nesneye dayalı, bilimsel ve etkileşimli üst düzey bir programlama dilidir. Girintilere dayalı söz dizimi kodlanma şeklinden dolayı, program akılda kalıcıdır ve program dilinin öğrenilmesi oldukça kolaydır. Özellikle her türlü platformda (Unix, Linux, Windows, Amiga) çalışmasından dolayı en çok tercih edilen diller arasındadır. Python ile veritabanı yazılımı, ağ programlama, web tasarım, veri analizi gibi birçok alanda yazılım geliştirilebilmektedir. Aynı zamanda siber güvenlik alanında çok kullanışlı bir programlama dilidir. Veri analizi konusunda NumPy, keos, matplotlib gibi kütüphaneler Python dilinin tercih edilmesinde en büyük neden olarak görülmektedir. Python dilini kullanabilmek ve geliştirmek için birçok editör mevcuttur. Pycharm, Eclipse, Jupyter, Visual Studio bu editörlerden bazılarıdır (Python. (b.t.) 2024, <https://tr.wikipedia.org/wiki/Python>). Veri analizlerini yapmak için Visual Studio ve Jupyter editörlerinden faydalanılmıştır.

Şekil 3.2: Visual Studio Code Programına Ait Görsel



3.4. Verilerin Hazırlanması

Verilerin hazırlanması sürecinde, hastane bilgi yönetim sisteminde bulunan hasta kayıtlarının analiz edilmesi için 1 Ocak 2023 ile 31 Aralık 2023 yılında Erzincan Binali YILDIRIM Üniversitesi Mengücek Gazi Eğitim ve Araştırma Hastanesine başvuran tüm hastaların verileri, Excel formatında etik kurul izni ve sağlık kurumundan alınan izin ile temin edilmiştir. Veriler KVKK uygun şekilde hastanın kişisel verileri olmadan toplanmıştır. Tüm verilerin tek bir Excel dosyası ile verilmesinin zor olmasından dolayı, aylık Excel dosyaları şeklinde veriler teslim alınmış ve bu veriler MYSQL veri tabanında birleştirilmiştir.

Bu işlem için MYSQL programında 1 adet vizitler isminde bir tablo oluşturulmuş ve içerisine ID, dosyano, birim, kurum, vizittarihi, vizitsaati, uyruk, cinsiyet, yas, il, ilce, tanisi, birimbrans isminde sütunlar (field) açılmıştır. MySQL programında içe aktar menüsünden faydalanılarak Excel dosyaları tek tek aylık olarak vizitler tablosuna aktarılmıştır. Tarafımızca oluşturulan sütun isimleri ve veri seti eş anlamlıdır. Bu şekilde verilerimizi analiz ederken hangi verinin hangi sütunda olduğunu daha rahat görmemiz sağlanmıştır.

Vizitler tablosunda kullanılan sütun adları şu şekildedir:

- ID: Otomatik sayı kayıtların sıra numarası olarak kullanılmaktadır.
- dosyano: Hastaneye başvuran her hastaya 1 adet dosya açılır. Bu hastanın benzersiz kayıt numarasıdır.
- Kurum: Hastanın sağlık güvencesi olduğu kurum bilgisinin tutulduğu alandır.
- Vizittarihi: Hastanın ilgili poliklinik başvuru tarihidir.
- Vizitsaati: Hastanın ilgili poliklinik başvuru saatidir.
- Uyruk: Hastanın hangi ülke vatandaşı olduğunun tutulduğu alandır.
- Yaş: hastanın yaş bilgisinin tutulduğu alandır.
- İl: Hastanın hangi il de ikamet ettiğinin tutulduğu alandır.
- İlçe: Hastanın ikamet ettiği ilçenin tutulduğu alandır.
- Tanisi: Hastanın başvurduğu güne ait tanısının tutulduğu alandır.
- Birimbrans: Hastanın başvurduğu poliklinik birim brans adının tutulduğu alandır.

Şekil 3.3:Aylık Excel Verilerinin MYSQL Veri Tabanında Birleştirilmesi

ID	dosyano	birim	kurum	vizittarihi	vizitsaati	uyruk	cinsiyet	yas	il	ilce	tanisi	birimbrans
1647208	15402406	ACIL SERVİS (SARI ALAN)	SGK (Bağkur)	2023-01-01	00:08	TÜRKİYE CUMH.	Erkek	20	ERZİNCAN	MERKEZ	R07.4 GÖĞÜS AC	Acil
1647209	16340689	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	2023-01-01	00:12	TÜRKİYE CUMH.	Erkek	2	ERZİNCAN	MERKEZ	J22 AKUT ALT SC	Acil
1647210	16630059	ACIL SERVİS (SARI ALAN)	SGK (Bağkur)	2023-01-01	01:17	TÜRKİYE CUMH.	Erkek	18	OSMANİYE	KADIRLI	R11 BULANTI VE	Acil
1647211	16444357	ACIL ÇOCUK POLİKLİNİĞİ	SGK (601c1-c3-c9)	2023-01-01	01:28	TÜRKİYE CUMH.	Kadın	1	ERZİNCAN	MERKEZ	L50.0 ALLERJİK (	Acil
1647212	16095588	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	2023-01-01	01:35	TÜRKİYE CUMH.	Kadın	7	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647213	15620535	ACIL SERVİS (SARI ALAN)	SGK (S.G.K.)	2023-01-01	03:23	TÜRKİYE CUMH.	Erkek	19	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647214	16623949	ACIL SERVİS (SARI ALAN)	SGK (S.G.K.)	2023-01-01	03:57	TÜRKİYE CUMH.	Erkek	25	ERZİNCAN	MERKEZ	M79.1 MIYALJİ	Acil
1647215	15213740	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	2023-01-01	05:36	TÜRKİYE CUMH.	Erkek	49	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647216	16128124	ACIL ÇOCUK POLİKLİNİĞİ	SGK (601c1-c3-c9)	2023-01-01	07:24	TÜRKİYE CUMH.	Kadın	6	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647217	15751955	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	2023-01-01	07:24	TÜRKİYE CUMH.	Erkek	62	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647218	16300627	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	2023-01-01	08:58	TÜRKİYE CUMH.	Kadın	3	ERZİNCAN	MERKEZ	J39 ÜST SOLUNU	Acil
1647219	15129842	ACIL SERVİS (YEŞİL ALAN)	SGK (S.G.K.)	2023-01-01	08:59	TÜRKİYE CUMH.	Erkek	39	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST:	Acil
1647220	15718260	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	2023-01-01	09:25	TÜRKİYE CUMH.	Erkek	10	ERZİNCAN	MERKEZ	J39 ÜST SOLUNU	Acil
1647221	16533764	ACIL ÇOCUK POLİKLİNİĞİ	T.C. ZİRAAT BANK	2023-01-01	09:28	TÜRKİYE CUMH.	Erkek	0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNU	Acil
1647222	16182410	ENJEKSİYON POLİKLİNİĞİ	SGK (Emekli San	2023-01-01	09:37	TÜRKİYE CUMH.	Kadın	27	ERZİNCAN	MERKEZ	Z00.0 GENEL TIB	Acil
1647223	16061481	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	2023-01-01	09:42	TÜRKİYE CUMH.	Erkek	36	ERZİNCAN	MERKEZ	R51 BAŞ AĞRISI	Acil
1647224	15109600	ACIL SERVİS (YEŞİL ALAN)	SGK (S.G.K.)	2023-01-01	09:50	TÜRKİYE CUMH.	Kadın	40	ERZİNCAN	MERKEZ	M54.9 BEL AĞRI	Acil
1647225	15918447	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	2023-01-01	09:52	TÜRKİYE CUMH.	Erkek	16	ERZİNCAN	MERKEZ	H10.9 KONJONK	Acil
1647226	15608614	ACIL ENJEKSİYON POL. - GÖZ EK BİNASI	SGK (Bağkur)	2023-01-01	10:10	TÜRKİYE CUMH.	Kadın	41	ERZİNCAN	MERKEZ	Z00.0 GENEL TIB	Acil

Tanı ve poliklinik analizlerinin sağlıklı olması için boş verilerin temizlenmesi veya yerlerine ortalama medyan verilerin aktarılması ile veri analizinin sağlıklı ve kaliteli

sonuçlar vermesi amaçlanmıştır. Analizin bu kısmında yaşanan sorunların tüm süreci etkileyeceği ve sürecin yanlış sonuçlar vermesine sebep olacağı düşünülmüştür.

3.4.1. Verilerin Kapsamı

Sağlık sektöründe hizmet veren hastaneye ait, 1.447.705 veri seti kullanılmıştır. Hastaların her bir poliklinik başvuruları bir hareket olarak kayıtlarda tutulmuştur. Birliktelik kuralı gereğince bu kayıtlar *dosyano* değişkeni sayesinde gruplandırılarak, veriler analiz edilecek bir veri seti haline getirilmiştir. Veri madenciliğinin ilk aşaması her zaman veri temizleme aşamasıdır. Bu nedenle temizlenen veriler doğru sonuç verme konusunda yardımcı olmuştur.

3.4.2. Veri Seçme ve Dönüştürme

Verilerin seçilme ve dönüştürülme sürecinde, verilerin birleştirilmesi ve yine verilerin EXCEL’ den MYSQL’ e aktarımı yapılmıştır.

Şekil 3.4: Ocak Ayına Ait Verilerin EXCEL Görüntüsü

	A	B	C	D	E	F	G	H	I	J	K	L
1	dosyano	birim	kurum	vizittarihi	vizitsaati	uyruk	cinsiyet	yas	il	ilce	tanisi	birimbrans
2	15402406	ACIL SERVİS (SARI ALAN)	SGK (Bağkur)	1.01.2023	00:08	TÜRKİYE CUMHURİYETİ	Erkek	20.0	ERZİNCAN	MERKEZ	R07.4 GÖĞÜS AĞRISI, TANIMLANMIŞ	Acil
3	16340669	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	00:12	TÜRKİYE CUMHURİYETİ	Erkek	2.0	ERZİNCAN	MERKEZ	J22 AKUT ALT SOLUNUM YOLU ENF	Acil
4	16530059	ACIL SERVİS (SARI ALAN)	SGK (Bağkur)	1.01.2023	01:17	TÜRKİYE CUMHURİYETİ	Erkek	18.0	OSMANİYE	KADIRLI	R11 BULANTI VE KUSMA	Acil
5	16444357	ACIL ÇOCUK POLİKLİNİĞİ	SGK (60/c1-c3-c9)	1.01.2023	01:28	TÜRKİYE CUMHURİYETİ	Kadın	1.0	ERZİNCAN	MERKEZ	L50.0 ALLERJİK ÜRTİKER	Acil
6	16095588	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	01:35	TÜRKİYE CUMHURİYETİ	Kadın	7.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
7	15620535	ACIL SERVİS (SARI ALAN)	SGK (S.G.K.)	1.01.2023	03:23	TÜRKİYE CUMHURİYETİ	Erkek	19.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
8	16523949	ACIL SERVİS (SARI ALAN)	SGK (S.G.K.)	1.01.2023	03:57	TÜRKİYE CUMHURİYETİ	Erkek	25.0	ERZİNCAN	MERKEZ	M79.1 MİYALJİ	Acil
9	15213740	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	1.01.2023	05:36	TÜRKİYE CUMHURİYETİ	Erkek	49.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
10	16128124	ACIL ÇOCUK POLİKLİNİĞİ	SGK (60/c1-c3-c9)	1.01.2023	07:24	TÜRKİYE CUMHURİYETİ	Kadın	6.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
11	15751955	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	1.01.2023	07:24	TÜRKİYE CUMHURİYETİ	Erkek	62.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
12	16300627	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	08:58	TÜRKİYE CUMHURİYETİ	Kadın	3.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
13	15129842	ACIL SERVİS (YEŞİL ALAN)	SGK (S.G.K.)	1.01.2023	08:59	TÜRKİYE CUMHURİYETİ	Erkek	39.0	ERZİNCAN	MERKEZ	J06.9 AKUT ÜST SOLUNUM YOLU	Acil
14	15718260	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	09:25	TÜRKİYE CUMHURİYETİ	Erkek	10.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
15	16533764	ACIL ÇOCUK POLİKLİNİĞİ	T.C. ZİRAAT BANKASI A.Ş. MENŞİ	1.01.2023	09:28	TÜRKİYE CUMHURİYETİ	Erkek	0.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
16	16182410	ENJEKSİYON POLİKLİNİĞİ	SGK (Emekli Sandığı)	1.01.2023	09:37	TÜRKİYE CUMHURİYETİ	Kadın	27.0	ERZİNCAN	MERKEZ	Z00.0 GENEL TIBBİ MUAYENE	Acil
17	16061481	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	1.01.2023	09:42	TÜRKİYE CUMHURİYETİ	Erkek	36.0	ERZİNCAN	MERKEZ	R51 BAŞ AĞRISI	Acil
18	15109600	ACIL SERVİS (YEŞİL ALAN)	SGK (S.G.K.)	1.01.2023	09:50	TÜRKİYE CUMHURİYETİ	Kadın	40.0	ERZİNCAN	MERKEZ	M54.5 BEL AĞRISI	Acil
19	15918447	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	1.01.2023	09:52	TÜRKİYE CUMHURİYETİ	Erkek	16.0	ERZİNCAN	MERKEZ	H10.9 KONJONKTİVİT, TANIMLANMIŞ	Acil
20	15608614	ACIL ENJEKSİYON POL. - GÖZ EK	SGK (Bağkur)	1.01.2023	10:10	TÜRKİYE CUMHURİYETİ	Kadın	41.0	ERZİNCAN	MERKEZ	Z00.0 GENEL TIBBİ MUAYENE	Acil
21	15695815	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	10:18	TÜRKİYE CUMHURİYETİ	Erkek	10.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
22	16413878	ACIL ÇOCUK POLİKLİNİĞİ	SGK (S.G.K.)	1.01.2023	10:22	TÜRKİYE CUMHURİYETİ	Erkek	2.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
23	15985334	ACIL POLİKLİNİĞİ - GÖZ EK BİNASI	SGK (S.G.K.)	1.01.2023	10:23	TÜRKİYE CUMHURİYETİ	Kadın	9.0	ERZİNCAN	MERKEZ	J39.9 ÜST SOLUNUM YOLUNUN	Acil
24	16261128	ACIL ÇOCUK POLİKLİNİĞİ	SGK (Bağkur)	1.01.2023	10:26	TÜRKİYE CUMHURİYETİ	Kadın	4.0	ERZİNCAN	MERKEZ	J39 ÜST SOLUNUM YOLU DİĞER	Acil
25	15705093	ACIL SERVİS (YEŞİL ALAN)	SGK (S.G.K.)	1.01.2023	10:33	TÜRKİYE CUMHURİYETİ	Erkek	41.0	ERZİNCAN	MERKEZ	J39.9 ÜST SOLUNUM YOLUNUN	Acil
26	15107358	ACIL CERRAHİ	SGK (60/c1-c3-c9)	1.01.2023	10:42	TÜRKİYE CUMHURİYETİ	Erkek	26.0	ERZİNCAN	MERKEZ	Z00.0 GENEL TIBBİ MUAYENE	Acil

Ocak 2023 ile Aralık 2023’ e ait verilerin aylık EXCEL kayıtları MYSQL veritabanı editörü MYSQL Front Programı kullanılarak oluşturulan ziyaretler tablosuna ay ay olarak aktarılarak tek bir tabloda birleştirilmiştir. Tek bir tabloda birleştirilmesi ve MYSQL veritabanı seçilmesinin en büyük sebebi, analiz yapılacak formatların daha hızlı oluşturulması ve bazı istatistik bilgilerin SQL komutları ile hızlıca alınmasını sağlamaktır.

MySQL Front programı ile MySQL veritabanına aktarılmış ve birleştirilmiş veriler Şekil 3.5’de görülmektedir. 1.447.705 satır veriye bu şekilde SQL komutlar ile daha hızlı işlemler yapılmıştır. Veri analizlerinde kullanmak için öncelikle tüm tablo CSV formatına dönüştürülmüştür. Verilerin Python’ da, CSV formatında oluşturulan dosya kullanılarak analiz edilmesi gerekmektedir.

Ayrıca hasta tanılarında birliktelik kuralı algoritmalarını kullanmak için verileri gruplandıran ve tanıları birleştiren bir sorgu çalıştırılmıştır.

```
Select dosyano, kurum, cinsiyet, yas, il, ilce, uyruk, vizittarihi, vizitsaati,  
GROUP_CONCAT(v.tanisi SEPARATOR " - ") as tanisi from vizitler v group  
by v.dosyano
```

Yukarıdaki SQL kodu ile vizitler tablosundaki tanıları, *dosyano* sütununa göre gruplandırılarak Dosya Numarası sütunu, hastaların ortak kimlik numarası olarak kullanılmıştır. Yukarıdaki komut çalıştırıldıktan sonra 248.924 satır veri seti oluşmuştur.

```
Select dosyano, kurum, cinsiyet, yas, il, ilce, uyruk, vizittarihi, vizitsaati,  
GROUP_CONCAT(v.birimbrans SEPARATOR " - ") as birimbrans from vizitler  
v group by v.dosyano
```

Şekil 3.5: Poliklinik Analizi İçin Kullanılacak Veri Seti MySQL Görseli

```

1 select dosyano, kurum, cinsiyet, yas, il, ilce, uyruk, vizittarihi, vizitsaati, GROUP_CONCAT(v.birimbrans SEPARATOR " - ") as birimbrans
2
3 group by v.dosyano

```

* dosyano	kurum	cinsiyet	yas	il	ilce	uyruk	vizittarihi	vizitsaati	birimbrans
***	SGK (S.G.K.)	Kadin	15	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-12-15	19:20	<MEMO>
15101311	SGK (S.G.K.)	Kadin	50	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-03-31	11:26	<MEMO>
15101312	SGK (SSK)	Kadin	40	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-11-10	09:01	<MEMO>
15101313	SGK (Bağkur)	Kadin	58	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-03-26	09:25	<MEMO>
15101314	SGK (SSK)	Erkek	51	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-06-09	08:05	<MEMO>
15101316	SGK (SSK)	Kadin	52	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-02-17	15:02	<MEMO>
15101317	SGK (S.G.K.)	Kadin	49	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-05-22	08:54	<MEMO>
15101321	SGK (60/c1-c3-c9)	Kadin	96	ERZINCAN	KEMAH	TÜRKİYE CUMHURİYETİ	2023-12-15	11:40	<MEMO>
15101322	SGK (SSK)	Kadin	62	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-05-30	14:37	<MEMO>
15101324	SGK (S.G.K.)	Kadin	51	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-07-27	21:53	<MEMO>
15101325	SGK (Bağkur)	Kadin	67	ERZINCAN	ÜZÜMLÜ	TÜRKİYE CUMHURİYETİ	2023-12-18	09:02	<MEMO>
15101330	SGK (SSK)	Kadin	36	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-01-13	19:21	<MEMO>
15101333	SGK (SSK)	Kadin	54	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-05-26	22:54	<MEMO>
15101334	SGK (Emekli San.)	Erkek	53	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-01-05	18:33	<MEMO>
15101335	SGK (Emekli San.)	Kadin	50	İSTANBUL	ÜSKÜDAR	TÜRKİYE CUMHURİYETİ	2023-09-01	13:17	<MEMO>
15101336	SGK (Bağkur)	Kadin	42	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-03-12	22:33	<MEMO>
15101338	SGK (Emekli San.)	Erkek	53	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-01-19	15:59	<MEMO>
15101339	SGK (SSK)	Kadin	55	ERZINCAN	ÜZÜMLÜ	TÜRKİYE CUMHURİYETİ	2023-12-08	08:10	<MEMO>
15101340	SGK (S.G.K.)	Kadin	76	TOKAT	NIKSAR	TÜRKİYE CUMHURİYETİ	2023-11-20	11:13	<MEMO>
15101344	SGK (SSK)	Kadin	41	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-03-15	13:02	<MEMO>
15101346	SGK (S.G.K.)	Kadin	70	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-09-18	16:18	<MEMO>
15101347	SGK (Bağkur)	Kadin	50	ERZINCAN	KEMAH	TÜRKİYE CUMHURİYETİ	2023-03-14	04:42	<MEMO>
15101348	SGK (SSK)	Erkek	35	YOZGAT	SARAYKENT	TÜRKİYE CUMHURİYETİ	2023-12-23	22:02	<MEMO>
15101349	SGK (SSK)	Kadin	74	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-11-08	17:11	<MEMO>
15101351	SGK (Emekli San.)	Erkek	73	İSTANBUL	ÜMRANİYE	TÜRKİYE CUMHURİYETİ	2023-08-22	08:01	<MEMO>
15101352	SGK (SSK)	Erkek	70	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-12-20	13:08	<MEMO>
15101353	SGK (S.G.K.)	Erkek	62	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-01-04	09:31	<MEMO>
15101354	SGK (SSK)	Erkek	54	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-02-15	17:02	<MEMO>
15101355	SGK (S.G.K.)	Kadin	70	ERZINCAN	MERKEZ	TÜRKİYE CUMHURİYETİ	2023-05-23	09:35	<MEMO>

Yukarıdaki SQL kodu ile vizitler tablosundaki *birimbrans* yani ziyaret hareketine ait poliklinik birim adına göre gruplandırdık. Burada amaç, her bir satırda bulunan poliklinik hareketlerini, dosya numarası sütununa göre tek bir satırda birleştirerek birliktelik kuralına uygun bir yapıda CSV formatına aktarabilmektir. Yukarıdaki kodu çalıştırdıktan sonra 248.924 satır veri setine ulaşılmıştır. Dosya Numarasına göre gruplandırıldığında 248.924 farklı kişinin hastaneye başvurduğu görülmektedir.

3.4.3. Veri Temizleme

Veri madenciliğinde veriler ne kadar temiz ve formatları o kadar analiz yapılacak algoritmaya uygunsa o kadar sağlıklı sonuçlar elde edilebilmektedir. Veri temizleme sürecinin diğer aşamasında Python programlama dili kullanılmıştır. Veri seti MySQL programında öncelikle EXCEL'den aktarıldığı için *birimbrans* ve *tanisi* sütunlarındaki verilerin sağında ve solunda boşluk olma ihtimalinden dolayı, MySQL komutlarından TRIM ile boşluk temizleme yapılmış, UPDATE komutu ile veriler temizlenmiş ve veriler güncellenmiştir.

```
UPDATE vizitler SET tanisi = TRIM(tanisi), birimbrans = TRIM(birimbrans)
```

Veri temizliğinin diğer aşamasında boş kayıtlar silinmiş ya da boş olan kayıtlara ortalama medyan, genel metinler veya 0 verilerek analizin daha doğru çıkması amaçlanmıştır. Bu işlemler MySQL üzerinden veya Python ile yapılabilmektedir. Sağlıklı veri analizi yapmak için Python programı tercih edilmiştir. Python programı veri temizliğini gerçekleştirmenin avantajlarından biri, Python ‘un CSV dosyasını ön belleğine almasıdır. Bu sayede temizleme işlemlerinin sonucunda oluşacak kodların hatalı olması durumunda CSV dosyasına diğer bir deyişle veri setine bir zarar olmadan dosyanın bozulma ihtimali en aza indirilmiştir.

Python ile veri analizi yapabilmek için *numpy*, *pandas*, *matplotlib*, *seaborn* kütüphanelerinin yüklenmesi gerekmektedir. Python programlama dilini kullanmak için bir editöre ihtiyaç bulunmaktadır. Bu nedenle Visual Studio Code programı üzerinden işlemler gerçekleştirilmiş ve program üzerinden TERMINAL menüsünden aşağıdaki kodları yazılarak ilgili kütüphanelerin yüklenmesi sağlanmıştır.

```
# Windows Kullanıcıları için
pip install numpy
pip install pandas
pip install matplotlib
pip install seaborn

# Linux Kullanıcıları için
sudo apt-get install python-numpy
```

- Numpy: Tam adı “Numeric Python” olan hesaplamalar yapmak için Python program dilinde kullanılan kütüphanedir. Büyük boyutlu veri setlerinin bilimsel hesaplamalarını çok hızlı bir şekilde yapmak için kullanılmaktadır. Mühendislik, veri analizi, yapay zeka ve bir çok alanda kullanılmaktadır (NumPy. 2024).
- Pandas: Sayısal tablolar ve zaman etiketli veri setlerinin analizinde Python programlama dilinde kullanılan bir kütüphanedir. Veri setlerinin tekrar birleştirilmesi, temizlenmesi, birbirine eklenmesi gibi birçok işlemi yaparak geri

veri setine döndürmektedir. Veri filtrelemesi yaparak basit bir veritabanı olarak kullanılabilir.

- Dataframe (Veri İskeletleri) yapısı ile içerisine EXCEL, Txt, CSV formatlarında veriler alabilmektedir (Pandas. 2024, <https://tr.wikipedia.org/wiki/Pandas>).
- matplotlib: Numpy için görselleştirme ve grafik çizim kütüphanesidir. İşlevsellik olarak MATLAB programına çok benzeyen bir şekilde tasarlanmıştır. Temel haritalama, ısı haritaları, 3 boyutlu çizimler yapabilen ücretsiz ve açık kaynak kodlu bir kütüphanedir (Matplotlib. 2024, <https://tr.wikipedia.org/wiki/Matplotlib>).
- Seaborn: Matplotlib kütüphanesinin arayüz kütüphanesidir. Bu kütüphane ile matplotlib ile çok fazla satırda yapılacak grafik görsellerinin birkaç satırda yapılması sağlanmaktadır. Göz alıcı renk tonları ile estetik grafikler, renk paleti kütüphanesi, dataframeler içindeki verilerin hızlı bir şekilde görselleştirilmesinde kullanılmaktadır. Grafik temaları ile daha etkili grafikler oluşturmaya imkân sağlamaktadır (seaborn. 2024, <https://www.veribilimiokulu.com/seaborn-ile-veri-gorsellestirmesi>).

```
import numpy as np  
  
import pandas as pd  
  
import matplotlib.pyplot as plt  
  
import seaborn as sns
```

Visual Studio ekranında yeni bir Python dosyası oluşturulmuş ve dosya adı analiz.py olarak kaydedilmiştir. Py ise Python dosya uzantısıdır. Python'da veri temizleme ve analiz işlemleri için yukarıdaki kodlar yazılarak ilgili kütüphaneler, uygulamaya eklenmiştir. CSV dosyasındaki veri seti, *Pandas* kütüphanesi kullanılarak dataframe olarak önbelleğe alınmıştır. Aşağıdaki kodlar yardımı ile CSV dosyası önbelleğe yüklenmiştir. Bu işlem sonrası analizde kullanılacak veri seti ana kaynak üzerinden değil, önbellekte bulunan veri seti üzerinden çekileceğinden, veri kaybının önüne geçilmiştir.

```
Data=pd.read_csv("tumvizitler.csv",usecols=['dosyano','cinsiyet','yas','tanisi'],sep=';',low_memory=False)

print("Veriler Listeleniyor\n", Data) #verileri ekrana yazarak kontrol ediyoruz
```

Yukarıdaki kod ile tumvizitler.csv dosyasının içindeki 2023 yılına ait MySQL den aktarılan ziyaret hareketlerinin önbelleğe aktarılması sağlanmıştır. Print komutu ile veri setinin başarılı bir şekilde önbelleğe yüklenip yüklenmediği ve analiz kodları için hazır olup olmadığı kontrol edilmiştir.

Şekil 3.6:CSV Dosyasında Bulunan Verilerin Pandas Kütüphanesi ile Dataframe Olarak Önbelleğe Alınıp, Listelendiği Jupyter Notebook Sonuç Ekranı

```
[3]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import os
import csv

import seaborn as sns

from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, association_rules

[4]: print("Veriler Çekiliyor \n ")
tumdata= pd.read_csv("tumvizitler.csv", usecols = ['dosyano','cinsiyet','yas','tanisi'],sep=';',low_memory=False)
eskidata=tumdata.query("tanisi != ','")
print("Veriler Listeleniyor\n",tumdata)

Veriler Çekiliyor

Veriler Listeleniyor
dosyano cinsiyet yas \
0      15402406 Erkek  20
1      16340669 Erkek   2
2      16530059 Erkek  18
3      16444357 Kadın   1
4      16095588 Kadın   7
...
1447700 15118321 Erkek  15
1447701 16023893 Erkek  10
1447702 16487022 Erkek   2
1447703 15991243 Erkek  10
1447704 15555682 Kadın  35

tanisi
0      R07.4 GÖĞÜS AĞRISI, TANIMLANMAMIŞ
1      J22 AKUT ALT SOLUNUM YOLU ENFEKSİYONU, TANIMLA...
2      R11 BULANTI VE KUSMA
3      L50.0 ALLERJİK ÜRTİKER
4      J06.9 AKUT ÜST SOLUNUM YOLU ENFEKSİYONU, TANIM...
...
1447700 J06.9 AKUT ÜST SOLUNUM YOLU ENFEKSİYONU, TANIM...
1447701 B30.9 VİRAL KONJONKTİVİT, TANIMLANMAMIŞ
1447702 J06.9 AKUT ÜST SOLUNUM YOLU ENFEKSİYONU, TANIM...
1447703 J06.9 AKUT ÜST SOLUNUM YOLU ENFEKSİYONU, TANIM...
1447704 J06.9 AKUT ÜST SOLUNUM YOLU ENFEKSİYONU, TANIM...
```

[1447705 rows x 4 columns]

Veri temizleme işlemlerine başlarken .isna().sum() komutu ile veri seti içerisinde kaç tane boş kayıt olduğu tespit edilmiş ve Print komutu kullanarak işlem yapılacak sütunların ekran kaydı alınmıştır.

Şekil 3.7: Veri Setimizdeki Boş Kayıtlar Kodları ve Sonuçları

```
Boskayitlar = tumdata.isna().sum()

print(Boskayitlar)

dosyano          0
kurum            0
vizittarihi      0
vizitsaati       0
cinsiyet         75
yas              0
il              1678
ilce             1686
tanisi          132143
birimbrans       68527
dtype: int64
```

Şekil 3.7’de gözüktüğü gibi *cinsiyet* alanında 75 adet, *il* alanında 1678 adet, *ilce* alanında 1686 adet, *tanisi* alanında 132.143 adet, *birimbrans* alanında 68.527 adet boş kayıt kayıt olduğu tespit edilmiştir. Veri temizleme işlemleri, verinin olduğu satırın silinmesi veya yerine 0, medyan veya ortak değerlerden birinin atanması şeklinde yapılmaktadır. Analizler ve istatistik kodlarından önce, veri temizliği yapmanın çok daha sağlıklı sonuçlar verdiği görülmüştür. Veri setimizi incelediğimizde *Tanisi*, *birimbrans*, *cinsiyet* sütunundaki değerlerin, aynı satırlarında boş veri olup olmadığını bilmediğimiz için bu kayıtları silmemiz durumunda veri setinden çok fazla veri kaybı olacağı ve analiz sonuçlarımız doğru çıkmayacağı öngörülmüştür.

Veri temizliği yaparken sırasıyla; *Tanisi* içindeki verilerin silinmemesi için ‘Z00.0 GENEL TIBBİ MUAYENE’ değeri ile doldurulmuştur. Bu sayede boş olan kayıtların silinmesi engellenerek veri bütünlüğü sağlanmıştır. Diğer değerler de aşağıdaki şekilde güncellenmiştir:

- Cinsiyet içindeki verileri ‘ERKEK’ değeri ile
- İl içindeki verileri ‘ERZİNCAN’ değeri ile
- İlçe içindeki verileri ‘MERKEZ’ değeri ile
- BirimBrans içindeki verileri ‘DİĞER’ değeri

Veri setinde bulunan kayıtlar temizledikten sonra istatistiki bilgiler alınıp kontrolleri sağlanmış ve veri seti analiz yapılabilir duruma getirilmiştir.

Şekil 3.8: Veri Setimizdeki Veri Temizlik Kodları ve Sonuçları

Verilerin Hazırlanması

Öncelikle, veri setindeki tanı (tanisi) alanının string türünde olduğundan emin olmak için aşağıdaki kod kullanılmıştır:

```
tumdata["tanisi"] = tumdata["tanisi"].astype(str) # Verilerin string türünde olduğundan emin ol
```

Eksik ve Boş Değerlerin Temizlenmesi

Veri setindeki boş veya eksik değerler, belirli varsayılan değerlerle doldurulmuştur. Bu işlemde kullanılan kod aşağıda verilmiştir:

```
# Temizleme İşlemi: Yaş sütunundaki boş veya eksik değerleri 0 olarak doldurabiliriz.
tumdata["cinsiyet"].fillna('Erkek', inplace=True) # Cinsiyet alanına Erkek atadık
tumdata["tanisi"].fillna('Z00.0 GENEL TIBBİ MUAYENE', inplace=True) # Yanılara Genel muayene atadık

tumdata["il"].fillna('ERZİNCAN', inplace=True) # İl alanına Erzincan atadık
tumdata["ilce"].fillna('MERKEZ', inplace=True) # İlçe alanına Merkez atadık

tumdata["birimbrans"].fillna('DİĞER', inplace=True) # Birimbranş alanına DİĞER atadık
```

3.5. Bulgular

3.5.1. Verilerin Tanımlayıcı İstatistikleri

Veri setinde seçme, dönüştürme ve temizlik aşaması sonrası genel istatistik süreçlerine geçilmiştir. Bu bölümde veri setindeki istatistiki tespitlere odaklanılarak, kuruluşa 2023 yılında başvuran **248.924** kişinin **1.447.705** adet hareketinden, hastaların demografik yapıları, polikliniklerin yaş, cinsiyete göre tercih oranları belirlenmiştir.

3.5.2. Hasta Demografik Bilgileri

Çalışmanın bu kısmında hastaların demografik bilgileri aşağıda yer alan tablo ve şekillerde sunulmuştur. Verilerin analiz sonuçları, MySQL ile birleştirilen ve CSV formatı haline getirilen dosyanın Python kodları ile kullanılarak yapılmıştır.

3.5.2.1. Cinsiyet Bazlı Hasta Dağılım Bilgiler

```
# "cinsiyet" sütunundaki farklı değerlerin dağılımını inceleyin
cinsiyet_dagilim = tumdata['cinsiyet'].value_counts()
print(cinsiyet_dagilim)
# Dağılımları ve istatistikleri görselleştirin
plt.pie(cinsiyet_dagilim, labels=cinsiyet_dagilim.index)
```

Yukarıdaki kodlar yardımıyla poliklinik veri setinde hastaneyi 2023 yılında ziyaret eden kişilerin cinsiyet dağılımı sayısal ve yüzdesel olarak Tablo 3.1’de verilmiştir.

Tablo 3.1: Cinsiyet Değişkenine Göre Veri

Cinsiyet	Sayı (Kişi)	Yüzde %
Kadın	126.583	%50,85
Erkek	122.341	%49,15
Toplam	248.924	% 100

Tablo 3.1’e göre Cinsiyet veri setinin dağılımına göre 126.583 kişi ve %50,85 oran ile en çok 2023 yılında kadınlar polikliniklere başvuruda bulunmuşlardır. Erkekler ise 122.341 kişi ile toplam hasta sayısının %49,15’i oranında bir paya sahip olarak poliklinik hizmeti almışlardır.

3.5.2.2. Cinsiyet ve Yaş Gruplarına Göre Hasta Dağılım Bilgiler

```
# Yaş gruplarını tanımla
def yas_grubu(yas):
    if yas <= 7:      return '0-7 arası'
    elif yas <= 14:
        return '8-14 arası'
    elif yas <= 25:
        return '15-25 arası'
    elif yas < 45:
        return '26-45 arası'
```

```

else:
    return '45 ve üzeri'

# Yaş gruplarını hesapla
tumdata['yas_grup'] = tumdata['yas'].apply(yas_grubu)

# Cinsiyet ve yaş gruplarını grupla ve say
cinsiyet_yas_grup = tumdata.groupby(['cinsiyet',
'yas_grup']).size().unstack().fillna(0)

print(cinsiyet_yas_grup)

```

Yukarıdaki Python kodları kullanılarak veri seti, yaş gruplarına göre gruplandırılmış ve Print komutu ile ekrana yazdırılarak Tablo 3.2'deki sonuçlar elde edilmiştir.

Tablo 3.2: Yaş Gruplarına Göre Cinsiyet Sayılarının Veri Seti Dağılımı

Cinsiyet	0 – 7 Arası	07 – 14 Arası	15 – 25 Arası	26 – 45 Arası	45 ve Üzeri
Erkek	14.070	9.772	23.391	32.067	43.041
Kadın	12.546	8.978	22.830	32.728	49.501
Toplam	26.616	18.750	46.221	64.795	92.542

Tablo 3.2'e göre 2023 yılında sağlık kurumuna, sağlık hizmeti almak için başvuran 248.924 adet hastanın cinsiyet ve yaş dağılımı incelendiğinde polikliniğe en fazla başvuran gruptaki 45 yaş ve üzeri hastaların 43.041 kişinin erkek hasta, 49.501 kişinin kadın hasta olduğu görülmektedir. Bunun yanı sıra 07-14 yaş arası çocuk hasta sayısı grubu en küçük paya sahip grup olarak ortaya çıkmıştır. Sağlık kurumuna başvuran hastaların cinsiyet ve yaş dağılımı Tablo 3.3'de verilmiştir.

Tablo 3.3: 45 ve Üzeri Yaştaki Hastaların Cinsiyete Göre Sayı Dağılımı

Cinsiyet	45 – 55 Arası	55 – 65 Arası	65 – 75 Arası	75 ve Üzeri
Erkek	15.176	13.162	10.067	4.636
Kadın	16.780	14.930	11.036	6.755
Toplam	31.956	28.092	21.103	11.391

Tablo 3.3’e göre hastalar içerisinde en fazla sayıya sahip olan 45 yaş ve üzeri hasta grubu cinsiyet ve yaş dağılımı incelendiğinde, en fazla sağlık kuruluşuna başvuranların 31.956 kişi ile 45 – 55 Yaş arası hastalar olduğu; bunların 15.176 adet kişinin erkek hasta, 16.780 adet kişinin kadın hasta olduğu görülmektedir.

3.5.2.3. Sağlık Güvence Kurumuna Göre Hasta Dağılım Bilgiler

Hastaneyi 2023 yılında ziyaret eden kişilerin sağlık güvence kurumuna ait bilgileri Tablo 3.4’de sunulmuştur.

Tablo 3.4: Sağlık Güvencesi Kurumu değişkenine Göre Veri Setinin Dağılımı

Kurum Adı	Hasta Sayısı (Kişi)	Kurum Adı	Hasta Sayısı (Kişi)
Sgk (Ssk)	91.935	Bkk Borclular	3.265
Sgk (S.G.K.)	57.735	Ücretli Hastalar	3.222
Sgk (Emekli Sandığı)	35.123	Kapalı ve Açık Cezaevi	2.002
Sgk (Bağkur)	19.948	Sgk (Er ve Erbaş)	1.985
Sgk (60/C1-C3-C9)	15.734	Diğer Kurumlar	13.826
Acil Hal	4.149		
	Toplam	248.924	

Tablo 3.4’e göre en çok SGK (SSK) kurumu sağlık güvencesi olan hastalar polikliniklere başvuruda bulunmuştur. Yine 15.734 kişi, Sgk (60/C1-C3-C9) kodu ile Genel Sağlık Sigortası kapsamında hizmet almıştır. Ücretli Hastalar sınıfındaki 3.222 kişinin sağlık kurumundan ücret karşılığında poliklinik hizmeti aldığı görülmektedir.

3.5.2.4. İkamet Adreslerine Göre Hasta Dağılım Bilgiler

```
# "il" sütunundaki farklı değerlerin dağılımını inceleyin

dagilim = tumdata['il'].value_counts()

top_dagilim = dagilim.head(10)

print(top_dagilim)

# Dağılımları ve istatistikleri görselleştirin

plt.figure(figsize=(12, 6))

plt.subplot(2, 2, 1)

top_dagilim.plot(kind='bar', rot=45)

plt.title("İkamet Ettikleri İllere Göre Dağılımı")
```

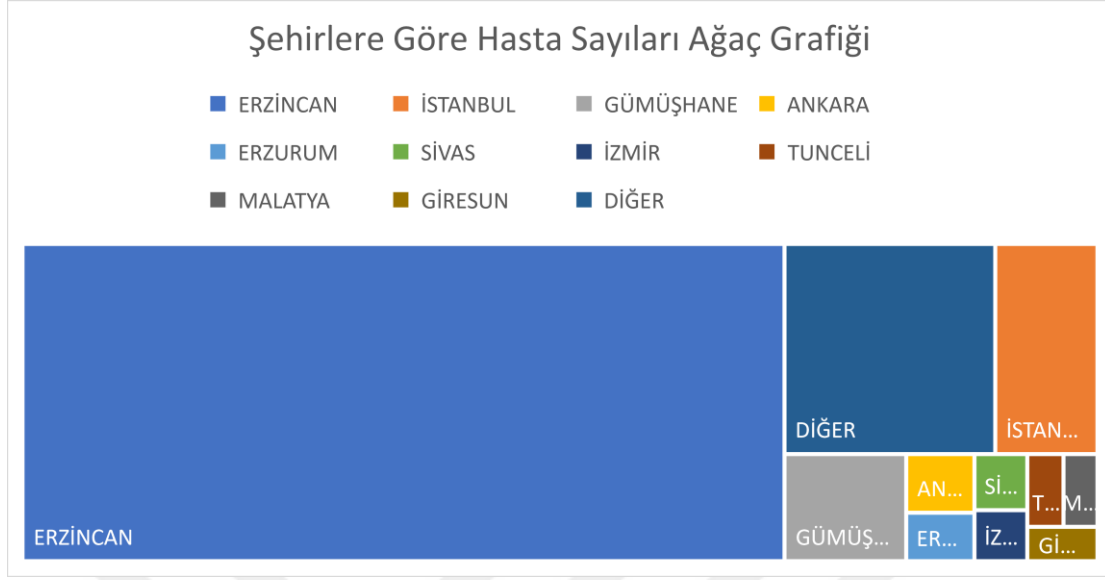
Yukarıdaki Python kodları hastaların ikamet ettiği şehirlere göre gruplandırılması amacıyla kullanılmış ve Print komutu ile Tablo 3.5’deki sonuçlar elde edilmiştir.

Tablo 3.5: Hastaların İkamet Bilgilerine Göre Veri Setinin Dağılımı

Şehir	Hasta Sayısı	Şehir	Hasta Sayısı
Erzincan	176.506 kişi	İzmir	1.981 kişi
İstanbul	15.649 kişi	Tunceli	1.917 kişi
Gümüşhane	9.511 kişi	Malatya	1.803 kişi
Ankara	2.921 kişi	Giresun	1.760 kişi
Erzurum	2.422 kişi	Diğer	32.274 kişi
Sivas	2.180 kişi		

Tablo 3.5’e göre 2023 yılında sağlık kurumunda sağlık hizmeti alanların ikamet ettiği şehirlerin grafiği Şekil 3.9’de gösterilmiştir.

Şekil 3.9:Veri Setimizdeki Veri Temizlik Kodları ve Sonuçları



Şekil 3.9’a göre 2023 yılında sağlık kurumuna sağlık hizmeti almak için başvuran 248.924 adet hastanın ikamet ettiği şehir dağılımı incelendiğinde 176.506 kişinin Erzincan’da ikamet ettiği, yoğunluk sıralamasının ise İstanbul, Gümüşhane, Ankara ve Erzurum illeri olarak gerçekleştiği görülmektedir. Diğer şehirler sınıfında 32.274 kişi yer almaktadır. Bu grafiğe göre İstanbul’da yaşayan 15.649 kişi tedavi olmak amaçlı olarak analizi gerçekleştirilen sağlık kurumunu tercih etmişlerdir. Çevre illerin yanı sıra Ankara ve İzmir’de yaşayan kişiler de tedavi amaçlı olarak analizi gerçekleştirilen sağlık kurumuna başvurmuşlardır.

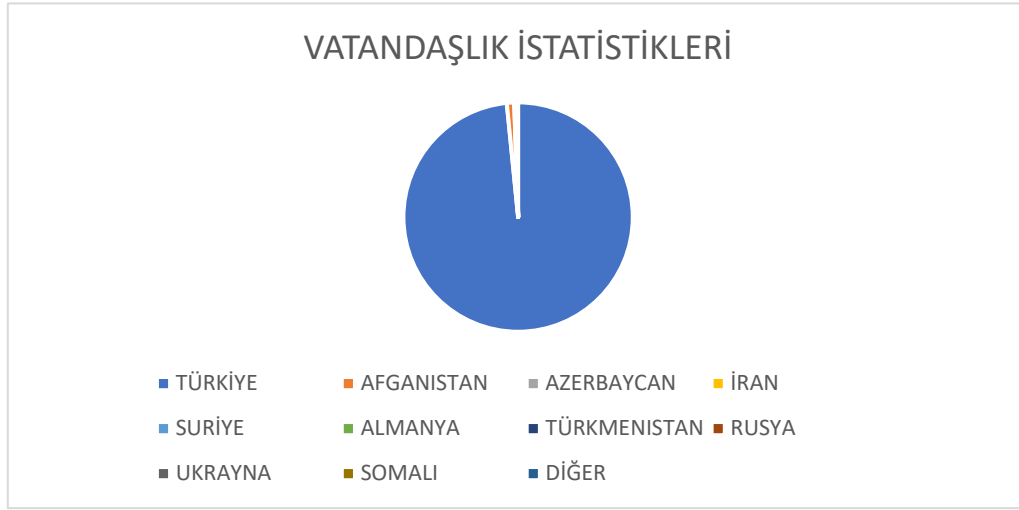
3.5.2.5. Vatandaşlık Bilgilerine Göre Hasta Dağılım Bilgiler

Tablo 3.6: Hastaların Vatandaşlık Bilgilerine Göre Veri Setinin Dağılımı

ÜLKE	SAYI	ÜLKE	SAYI
Türkiye	244.992	Türkmenistan	46
Afganistan	2.567	Rusya	43
Azerbaycan	231	Ukrayna	42
İran	175	Somalı	40
Suriye	175	Diğer	451
Almanya	162		

Tablo 3.6’ya göre 2023 Yılında sağlık kurumunda sağlık hizmeti alanların vatandaşlık bilgileri grafiği Şekil 3.10’da gösterilmiştir.

Şekil 3.10:Sağlık Hizmeti Alan Kişilerin Vatandaşlık Sayı Grafiği



Şekil 3.10'a göre sağlık hizmeti almak için başvuran 248.924 adet hastanın vatandaşlık dağılımını incelediğimizde 244.992 kişinin ülkemizden olduğu daha sonra ise Afganistan, Azerbaycan, İran, Suriye ülke vatandaşları olarak sıralandığı görülmektedir. 451 kişi ise diğer ülkelere göre dağılım göstermektedir.

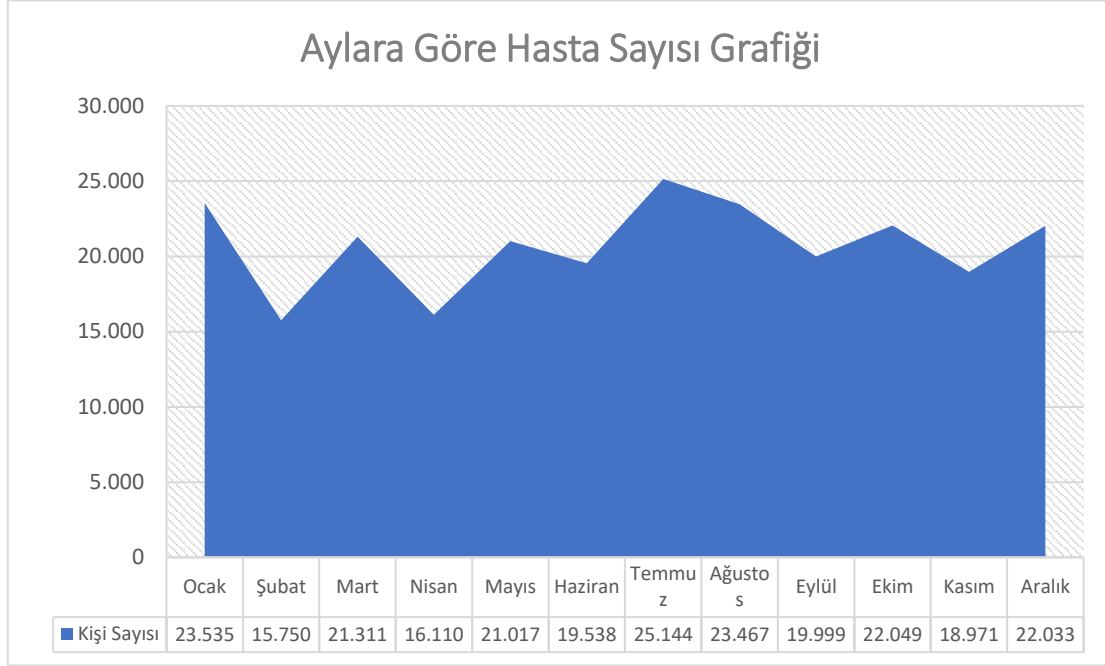
3.5.2.6. Aylara Göre Hasta Dağılım Bilgiler

Sağlık kurumuna başvuran hastaların aylık bazdaki yoğunluk durumu Tablo 3.7 ve Şekil 3.11'de sunulmaktadır.

Tablo 3.7: Aylara Göre Hastaların Veri Setinin Dağılımı

Ay	Kişi Sayısı	Ay	Kişi Sayısı
Ocak	23.535	Temmuz	25.144
Şubat	15.750	Ağustos	23.467
Mart	21.311	Eylül	19.999
Nisan	16.110	Ekim	22.049
Mayıs	21.017	Kasım	18.971
Haziran	19.538	Aralık	22.033
Toplam		248. 924	

Şekil 3.11: Aylara Göre Hasta Sayısı Grafiği



Şekil 3.11'e göre ay bazında hasta başvuru durumu incelendiğinde en fazla başvurunun 25.144 kişi ile temmuz ayında, en az başvurunun ise 15.750 kişi ile şubat ayında gerçekleştiği görülmektedir. Başvuru sayılarında genel olarak yatay bir seyir izlenmektedir.

3.5.2.7. Saatlere Göre Hasta Dağılım Bilgiler

2023 Yılında sağlık kurumunda sağlık hizmeti alanların saatlere göre dağılımı Tablo 3.8 ve Şekil 3.12'de gösterilmiştir.

Tablo 3.8: Saatlere Göre Hastaların Veri Setinin Dağılımı

Saat	Kişi Sayısı	Saat	Kişi Sayısı
00:00	2.530	12:00	7.675
01:00	1.455	13:00	29.344
02:00	905	14:00	22.269
03:00	727	15:00	17.412
04:00	588	16:00	7.237
05:00	653	17:00	9.056
06:00	664	18:00	7.522
07:00	1.756	19:00	8.048
08:00	28.595	20:00	7.762
09:00	31.059	21:00	6.787
10:00	28.269	22:00	5.321
11:00	19.708	23:00	3.582

Şekil 3.12: Saatlere Göre Hasta Sayısı Grafiği



Şekil 3.12’ e göre hastaların saat bazlı dağılımını incelendiğinde 31.059 kişi ile yoğunluğun en fazla olduğu saat diliminin 9:00, yoğunluğun en az olduğu saat diliminin ise 588 kişi ile 04:00 olduğu görülmektedir. Saat 00:00 ile 08:00 arasında hasta yoğunluğu en az düzeyde iken 08:00 ile 16:00 arasında en yoğun olarak gerçekleşmiştir.

3.5.3. Birliktelik Kuralları ile Poliklinik Analizi

Çalışmanın bu aşamasında hastaların başvurduğu polikliniklere (vizit hareketleri) ait veriler, birliktelik kuralı kullanılarak analiz edilmiştir. Analiz süreci poliklinik hareketlerinin bulunduğu CSV dosyası Python ile okutularak ve bu bilgiler Pandas kütüphanesinin dataframe içerisine yüklenerek başlamıştır.

3.5.3.1 Poliklinik Genel Analizi

Poliklinik analizi yapılmadan önce Şekil 3.13’de görüldüğü gibi veri setinde boş veri olmayacak şekilde kayıtlar temizlenmiştir. Bu bölümde polikliniklerin yaş, cinsiyet, kurum, yaş, uyruk, aylara, saatlere göre istatistikleri görülmektedir. Öncelikle veri setinin demografik bilgileri analiz edilmiştir.

Şekil 3.13: Veri Setinde Veri Temizliği Sonuç Görseli

```
Boskayitlar = tumdata.isna().sum()

print(Boskayitlar)

dosyano      0
kurum        0
cinsiyet     0
yas          0
il           0
ilce         0
uyruk        0
vizittarihi  0
vizitsaati   0
birimbrans   0
yas_grup     0
date         0
date_formatted 0
time         0
time_formatted 0
dtype: int64
```

Şekil 3.13'e göre 2023 yılında hastanede bulunan 300 adet polikliniğin 46 adet birim kategorisi ile hizmet verdiği görülmektedir. Çalışmanın bu aşamasında veri analizi birim bazlı olarak yapılmıştır.

3.5.3.1.1. Polikliniklere Göre Vizit Sayısı Bilgileri

Tablo 3.9'da poliklinikler bazında vizit sayıları yer almaktadır.

Tablo 3.9: Polikliniklere Göre Hastaların Veri Setinin Dağılımı

Birim Adı	Vizit Sayısı
Acil	496.884
Göz Hastalıkları	87.820
İç Hastalıkları	80.754
Ortopedi ve Travmatoloji	73.260
Tanımsız Birim	68.527
Çocuk Sağlığı ve Hastalıkları	61.203
Kardiyoloji	58.098
Kulak-Burun-Boğaz Hastalıkları	48.498
Kadın Hastalıkları ve Doğum	47.882
Deri ve Zührevi Hastalıkları	43.538
Diğer Poliklinikler	381.241
Toplam	1.447.705

Tablo 3.9'a göre Hastaneye başvuran hastaların en çok başvurduğu birim Acil olarak görülmektedir, daha sonra ise Göz Hastalıklarına 87.820 kez ve İç Hastalıklarına 80.754 kez başvuru gerçekleşmiştir.

3.5.3.1.2. Erkek Hastaların Vizit Sayısı Bilgileri

Şekil 3.14: Erkek Hastaların Vizit Hareketleri Grafiği

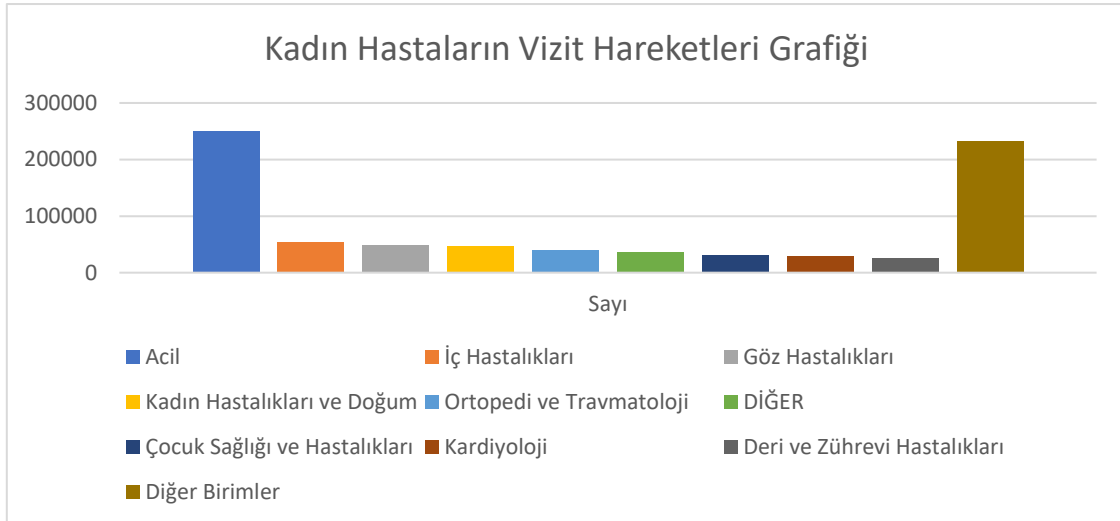


Şekil 3.14'e göre Erkek hastalar en çok 247.400 kez Acile başvurmuşlardır. Acil ve Diğer Birimler sonrası en çok başvuru 39.675 ile Göz Hastalıkları ve 33.256 ile Ortopedi ve Travmatoloji birimlerinde gerçekleşmiştir.

3.5.3.1.3. Kadın Hastaların Vizit Sayısı Bilgileri

Şekil 3.15'de kadın hastaların vizit hareketleri grafiği görülmektedir.

Şekil 3.15: Kadın Hastaların Vizit Hareketleri Grafiği



Şekil 3.15’e göre kadın hastalar en çok 249.484 kez Acile başvurmuşlardır. Acil ve Diğer Birimler dışında 53.809 kez İç Hastalıkları, 48.145 kere Göz Hastalıkları ve 47.882 kez ise Kadın Hastalıkları ve Doğum birimlerine başvuru bulunmaktadır.

3.5.3.1.4. Aylara Göre Hastaların Vizit Sayısı Bilgileri

Aylara göre hasta dağılımını görmek için aşağıda yer alan SQL kod ile MySQL veri tabanında bulunan verilerden Tablo 10’daki veri seti dağılımı elde edilmiştir. “Count(CASE WHEN month (vizittarihi)=1 THEN 0 END) AS 'Ocak' ...” şeklindeki kod yardımı ile veriler SQL penceresinde yan yana ve ay ay yazmak için kullanılmıştır. Her ay için bu kod çoğaltılarak 2023 yılı için aylık veri seti alınmıştır.

```
SELECT
    birimbrans,
    Count(CASE WHEN month (vizittarihi)=1 THEN 0 END) AS 'Ocak',
    Count(CASE WHEN month (vizittarihi)=2 THEN 0 END) AS 'Şubat',
    Count(CASE WHEN month (vizittarihi)=3 THEN 0 END) AS 'Mart',
    Count(CASE WHEN month (vizittarihi)=4 THEN 0 END) AS 'Nisan',
    Count(CASE WHEN month (vizittarihi)=5 THEN 0 END) AS 'Mayıs',
    Count(CASE WHEN month (vizittarihi)=6 THEN 0 END) AS 'Haziran',
    Count(CASE WHEN month (vizittarihi)=7 THEN 0 END) AS 'Temmuz',
    Count(CASE WHEN month (vizittarihi)=8 THEN 0 END) AS 'Ağustos',
    Count(CASE WHEN month (vizittarihi)=9 THEN 0 END) AS 'Eylül',
    Count(CASE WHEN month (vizittarihi)=10 THEN 0 END) AS 'Ekim',
    Count(CASE WHEN month (vizittarihi)=11 THEN 0 END) AS 'Kasım',
    Count(CASE WHEN month (vizittarihi)=12 THEN 0 END) AS 'Aralık'
FROM vizitler v
group by birimbrans
```

Tablo 3.10 ve Şekil 3.16’da hastaların 2023 yılında ay ay hangi birimlere gittiğine dair sayısal dağılım verilmiştir.

Tablo 3.10: Aylara Göre Hastaların Vizit Hasretleri Veri Setinin Dağılımı

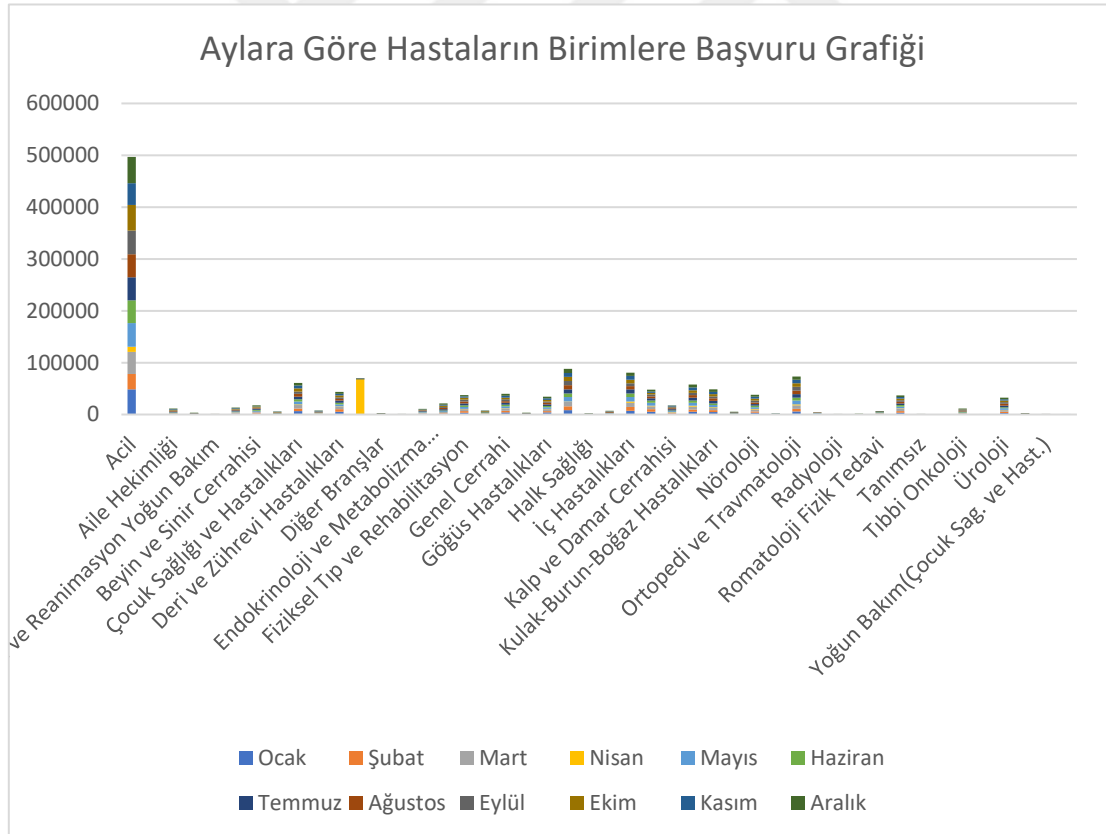
Birim Adı	O	Ş	M	N	M	H	T	A	E	E	K	A
Acil	48513	29413	42818	9638	46120	43600	44354	44628	45543	49879	41624	50754
Adli tıp	0	1	0	0	0	0	0	0	0	7	31	43
Aile hekimliği	754	585	859	139	1122	767	1222	1173	1341	1260	1148	1167
Algoloji	365	304	341	104	325	261	190	127	276	217	210	228

Tablo 3.10: Aylara Göre Hastaların Vizit Hasretleri Veri Setinin Dağılımı (Devamı)

Anestezi ve Reanimasyon Yoğun Bakım	47	37	33	16	35	26	30	40	43	42	33	35
Anesteziyoloji ve Reanimasyon	1502	945	1277	414	1266	797	1381	1167	1118	1142	1131	1116
Beyin ve Sinir Cerrahisi	1721	1603	1947	310	1805	1473	1747	1762	876	1662	1115	1578
Çocuk cerrahisi	474	437	521	107	497	410	580	541	469	465	522	409
Çocuk Sağlığı ve Hastalıkları	6394	5390	6474	1102	5735	4370	4749	5802	4383	6152	5565	5087
Çocuk ve Ergen Ruh Sağlığı ve Hastalıkları	284	604	941	117	805	552	635	619	632	911	804	780
Deri ve Zührevi Hastalıkları	5292	4773	4975	796	3712	3280	3846	3316	3114	3840	3062	3532
Diğer	1	0	580	67936	1	3	2	1	0	0	2	1
Diğer branşlar	155	129	295	81	155	128	183	85	207	192	185	608
Diş Hastalıkları ve Tedavisi	35	32	30	2	36	18	20	24	18	32	36	32
Endokrinoloji ve Metabolizma Hastalıkları	985	897	1119	238	1309	900	1184	954	741	704	887	834
Enfeksiyon Hastalıkları ve Klinik Mikrobiyoloji	1665	1270	1475	462	1607	1477	2188	2308	2151	2251	2076	2260
Fiziksel Tıp ve Rehabilitasyon	3386	3153	3707	865	3760	2787	3628	3680	3620	3766	2683	2597
Gastroenteroloji	733	540	888	237	770	624	767	290	780	734	633	900
Genel cerrahi	4107	3046	4259	892	3706	3013	3621	3191	3309	3507	3400	3870
Göğüs cerrahisi	267	150	249	65	281	223	202	240	285	339	351	364
Göğüs hastalıkları	4000	2842	3166	735	3162	2395	3278	2973	2712	3037	2982	3096
Göz hastalıkları	8458	7242	9000	684	8560	6696	8162	7779	7878	7962	7710	7689
Halk sağlığı	113	61	90	0	86	70	91	79	113	89	124	114
Hematoloji	567	461	614	189	723	597	616	680	638	616	718	689
İç hastalıkları	7084	7243	8539	2273	8438	6807	7864	6451	5407	6952	7224	6472
Kadın Hastalıkları	5413	4605	5830	1513	5494	4049	4418	3497	3680	3122	2855	3406
Kalp ve Damar Cerrahisi	1400	1162	1622	528	1693	1393	2233	2167	1772	1468	1318	1204
Kardiyoloji	5387	4155	5520	2130	5226	4485	5359	4979	4663	5050	5327	5817
Kulak-burun-boğaz hastalıkları	4798	3457	4709	864	4378	3435	4908	4542	3807	4588	4398	4614
Nefroloji	440	368	491	40	502	413	437	560	387	384	398	359
Nöroloji	3291	2949	3667	687	3654	2647	3133	3088	3617	4053	3623	3743
Nükleer tıp	6	58	72	17	138	112	141	118	0	123	118	141
Ortopedi ve Travmatoloji	6255	4530	6937	1828	7204	5766	6928	6705	7098	6999	6659	6351
Plastik, Rekonstrüktif ve Estetik Cerrahi	629	518	651	129	701	461	690	391	0	0	0	0

Tablo 3.10: Aylara Göre Hastaların Vizit Hasretleri Veri Setinin Dağılımı (Devamı)

Radyoloji	6	29	16	342	5	2	0	3	9	1	5	9
Romatoloji (İç hastalıkları)	0	0	0	0	0	0	0	0	0	8	667	990
Romatoloji fizik tedavi	975	852	905	337	790	794	3	506	0	171	180	1
Ruh Sağlığı ve Hastalıkları	3785	2996	4781	370	3399	2895	3289	3003	2987	3413	2735	2996
Tanımsız	15	20	51	26	21	18	23	10	10	11	13	7
Tıbbi biyokimya	0	0	0	0	0	0	0	0	0	0	0	2
Tıbbi onkoloji	958	756	965	480	1109	927	861	1117	1162	1111	985	1045
Tıbbi patoloji	0	0	0	0	0	0	0	2	17	37	31	13
Üroloji	3105	2287	2955	701	2922	2232	3154	3233	3077	3203	2782	3084
Yoğun bakım	108	86	120	88	107	111	119	109	82	69	64	82
Yoğun Bakım(Çocuk Sag. Ve Hast.)	51	52	56	0	59	46	48	48	29	38	38	35
Yoğun bakım(göğüs hast.)	12	5	9	1	14	10	10	6	11	9	5	10

Şekil 3.16: Aylara Göre Hastaların Birimlere Başvuru Hareketleri Grafiği

Şekil 3.16'ya göre veri setimizi ay bazlı incelediğimizde, Acil poliklinikleri en fazla başvurulmuş poliklinikler olarak gözükülmektedir. Acil ve Diğer birimlerini çıkardığımızda Ocak ayında 8.458 kişi Göz Hastalıklarına, 7.084 kişi ise İç hastalıkları

birimine müracaat ettiği görülmektedir. Şubat ayında 7.243 kişi İç hastalıklarına, 7.242 kişi ise Göz Hastalıkları birimine başvuru gerçekleşmiştir. Ağustos döneminde 6.705 kişi Ortopedi ve Travmatoloji bölümüne başvurmuştur. Bir önceki aylara göre Ortopedi ve Travmatoloji bölümüne başvuranların sayısı artmış, bu artış trendi Eylül, Ekim, Kasım aylarında da devam etmiştir.

3.5.3.1.5. Sosyal Güvenlik Kurumlarına Göre Hastaların Vizit Sayısı Bilgileri

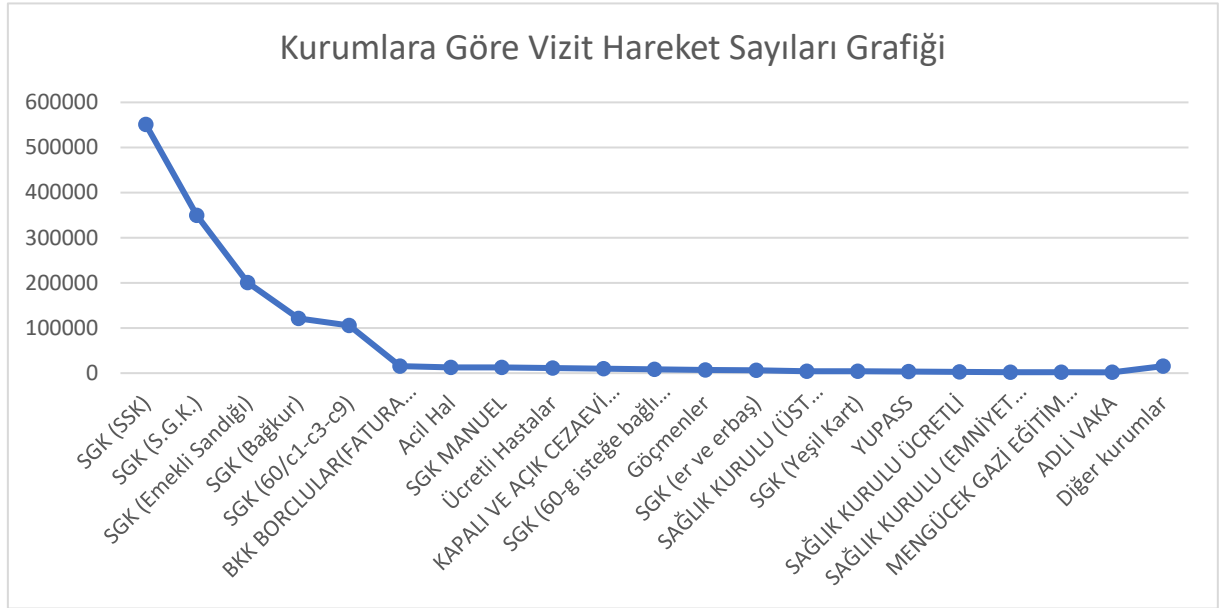
Tablo 3.11’de vizit başvurularının sosyal güvenlik kurumları içerisindeki dağılımı sunulmaktadır. Bu tablo incelendiğinde Sosyal Güvenlik Kurumu kendi içinde S.S.K, S.G.K, Emekli Sandığı gibi provizyon türleri olduğu görülmüştür.

Tablo 3.11: Sosyal Güvenlik Kurumuna Göre Vizit Sayılarının Dağılımı

Kurum	Vizit Sayısı	Kurum	Vizit Sayısı
Sgk (ssk)	550.756	SGK (60-g isteğe bağlı sigortalılar)	8.320
Sgk (s.g.k.)	349.547	Göçmenler	6.864
Sgk (emekli sandığı)	200.716	SGK (er ve erbaş)	6.127
Sgk (bağkur)	121.275	Sağlık kurulu (üst yazı+engelli)	4.311
SGK (60/c1-c3-c9)	105.611	Sgk (yeşil kart)	3.854
Bkk borçlular	15.304	Yupass	3.406
Acil hal	12.832	Sağlık kurulu ücretli	2.960
Sgk manuel	12.752	Sağlık kurulu (emniyet personeli)	2.377
Ücretli hastalar	10.940	Mengücek gazi eğitim araştırma hastanesi protokol hastası	2.249
Kapalı ve açık cezaevi mahkumlar	10.104	Diğer kurumlar	1.5640

Tablo 3.11’e göre Toplam 1.447.705 vizit hareketi incelendiğinde genel olarak Sosyal Güvenlik Kurumuna ait hastaların hastaneye daha çok başvurduğu görülmektedir. BKK Borçlular şeklinde listelenen hastaların herhangi bir sağlık güvencesi bulunmamaktadır. Mahkumlar statüsünden hizmet alan hastalar için 10.104 vizit hareketi görülmektedir. Şekil 3.17’de bu dağılım görselleştirilmektedir.

Şekil 3.17: Kurumlara Göre Vizit Hareket Sayıları Grafiği



Ayrıca ek bilgi olarak 3.406 ziyaret hareketi YUPASS (Yurtdışı Pasaportlular) statüsündeki hastalara, 6.864 ziyaret hareketi de Göçmenler statüsünde sağlık hizmeti alan hastalara aittir.

3.5.3.1.6. Yaş Gruplarına Göre Hastaların Ziyaret Sayısı Bilgileri

Ziyaret hareketlerinin yaş grupları içerisindeki dağılımına dair bilgiler Tablo 3.12’de yer almaktadır.

Tablo 3.12: Ziyaret Hareketlerinin Yaş Gruplarına Göre Veri Setinin Dağılımı

Birim Adı	0-7 arası	8-14 arası	15-25 arası	26-45 arası	45 ve üzeri
Acil	65452	39365	94872	132778	164417
Adli Tıp	5	3	24	38	12
Aile Hekimliği	192	198	2885	3234	5028
Algoloji (Anesteziyoloji ve Reanimasyon)	2	0	102	775	2069
Anestezi ve Reanimasyon Yoğun Bakım	2	1	4	15	395
Anesteziyoloji ve Reanimasyon	1403	508	1177	3141	7027
Beyin ve Sinir Cerrahisi	175	145	1391	5467	10421
Deri ve Zührevi Hastalıkları	1757	2503	13880	12116	13282
Diğer Branşlar	93	37	504	1077	692
Diş Hastalıkları ve Tedavisi	249	18	16	22	10

Tablo 3.12: Vizit Hareketlerinin Yaş Gruplarına Göre Veri Setinin Dağılımı (Devamı)

Endokrinoloji ve Metabolizma	4	10	1076	3741	5921
Enfeksiyon Hastalıkları	12	0	3286	7273	10619
Fiziksel Tıp ve Rehabilitasyon	183	375	2820	9921	24333
Gastroenteroloji	2	0	767	2290	4837
Genel Cerrahi	3	16	3985	13287	22630
Göz Hastalıkları	5926	5623	10514	14569	51188
Göğüs Cerrahisi	71	76	466	774	1629
Göğüs Hastalıkları	75	123	3546	8380	22254
Halk Sağlığı	0	0	88	418	524
Hematoloji	0	0	569	1765	4774
Kadın Hastalıkları ve Doğum	4	47	11685	27928	8218
Kalp ve Damar Cerrahisi	17	44	810	4228	12861
Kardiyoloji	2	0	2817	8723	46556
Kulak-Burun-Boğaz Hastalıkları	5658	3260	7743	12430	19407
Nefroloji	0	19	250	643	3867
Nöroloji	10	68	4117	9928	24029
Nükleer Tıp	7	8	39	141	849
Ortopedi ve Travmatoloji	3067	4815	10405	18511	36462
Plastik, Rekonstrüktif ve Estetik Cerrahi	165	159	968	1411	1467
Radyoloji	4	2	21	89	311
Romatoloji (İç hastalıkları)	0	0	110	595	960
Romatoloji Fizik Tedavi	2	0	389	1858	3265
Ruh Sağlığı ve Hastalıkları	14	35	6415	13706	16479
Tanımsız	23	4	25	54	119
Tıbbi Biyokimya	0	0	0	2	0
Tıbbi Onkoloji	0	0	72	1110	10294
Tıbbi Patoloji	20	3	15	24	38
Yoğun Bakım	18	9	78	151	889
Yoğun Bakım(Göğüs Hast.)	0	0	2	2	98
Yoğun Bakım(Çocuk Sag. ve Hast.)	500	0	0	0	0
Tanımsız	7595	4187	10366	17905	28474
Çocuk Cerrahisi	3753	1211	467	1	0
Çocuk Sağlığı ve Hastalıkları	41784	14493	4876	50	0
Çocuk ve Ergen Ruh Sağlığı ve Hastalıkları	2121	3502	2060	1	0
Üroloji	625	498	2224	6605	22783
İç Hastalıkları	35	38	9642	22344	48695

Tablo 3.12'e göre 0 ile 7 Yaş arasındaki hastalar en çok Acil polikliniklerine 65.452 kez başvururken Çocuk Sağlığı ve Hastalıkları polikliniklerine 41.784 kez başvurmuşlardır. 3.753 kez Çocuk Cerrahisi polikliniklerine başvurdukları görülmüştür.

8 ile 14 yaş arasındaki hastalar ise 39.365 kez Acil polikliniklerine başvuruda bulunurken 14.493 kez ise Çocuk Sağlığı ve Hastalıkları polikliniklerine başvurmuşlardır. Diğer poliklinik sayılarının düşük olma sebebi 18 yaşında kadar çocuk hastaların Çocuk Sağlığı ve Hastalıkları birimleri dışına, doktor sevki olmadan gidememesinden kaynaklı olduğu düşünülmektedir.

15 ile 25 yaş arasındaki hastaların ise 94.872 kez Acil polikliniklerine başvurdukları görülürken Acil dışında 13.880 kez Deri ve Zührevi Hastalıkları, 11.685 kez ise Kadın Hastalıkları ve Doğum polikliniklerine başvurdukları görülmektedir. Çocuk Sağlığı ve Hastalıkları poliklinikleri 15 yaşından önceki yaş gruplarında, on binden fazla başvuru alırken bu yaş grubunda sayı 4.876'a kadar düşmektedir.

26 ile 45 yaş arasındaki hastalarda ise 132.778 kez Acil polikliniklerine, 27.928 kez Kadın Hastalıkları ve Doğum polikliniklerine, 22.344 kez ise İç Hastalıkları polikliniklerine başvurmuşlardır. Bu yaş gurubunda Ruh Sağlığı ve Hastalıkları ve Göz Hastalıkları polikliniklerine başvuru sayısının arttığı gözlenmektedir.

45 yaş ve üzerindeki hastalarda ise 164.417 kez Acil polikliniklerine, 51.188 kez Göz Hastalıkları polikliniklerine, 48.695 kez İç Hastalıkları polikliniklerine başvurdukları görülmektedir. Bu yaş grubunda Kardiyoloji poliklinikleri ve Ortopedi ve Travmatoloji polikliniklerine başvuruda artış yaşanmıştır.

3.5.3.1.7. İkamet Edilen Şehirlere Göre Hastaların Vizit Sayısı Bilgileri

Tablo 3.13'de, 2023 yılının vizit hareketlerinde ikamet edilen şehirlerden gelen hastaların vizit sayıları sunulmuştur. Tabloda vizit sayılarının en fazla olduğu 10 adet şehir ve en az olan 5 şehrin veri seti dağılımı gözükmemektedir.

Tablo 3.13: İkamet Edilen Şehirlere Göre Vizit Hareketlerinin Veri Setinin Dağılımı

Şehir Adı	Vizit Sayısı	Şehir Adı	Vizit Sayısı
ERZİNCAN	1.205.964	İZMİR	6.126
İSTANBUL	47.576	GİRESUN	5.505
GÜMÜŞHANE	33.688		
TUNCELİ	9.748	BURDUR	302
ANKARA	8.537	KARAMAN	272
ERZURUM	7.690	DÜZCE	269
SİVAS	6.982	BİLECİK	263
MALATYA	6.735	KASTAMONU	262

Tablo 3.13'e göre Erzincan ilinden gelen 1.205.964 vizit hareketi görülmekte iken İstanbul ilinde ikamet eden hastaların 47.576 vizit hareketi gerçekleşmiştir. Giresun ilinden sonraki veriler ise en az vizit hareketi olan illeri göstermektedir. En az olarak Kastamonu ilinde ikamet eden hastaların 262 kez vizit hareketi olduğu görülmektedir.

3.5.3.1.8. İkamet Edilen İlçelere Göre Hastaların Vizit Sayısı Bilgileri

Tablo 3.14'te 2023 yılının vizit hareketlerinde ikamet edilen ilçelerden gelen hastaların vizit sayılarının en fazla olduğu 18 adet ilçenin veri seti dağılımı görülmektedir.

Tablo 3.14: İkamet Edilen İlçelere Göre Vizit Hareketlerinin Veri Setinin Dağılımı

İlçe	Vizit Sayısı	İlçe	Vizit Sayısı
MERKEZ	1.028.080	KELKİT	58
ÜZÜMLÜ	57.486	PATNOS	26
TERCAN	35.875	EYÜPSULTAN	21
KEMAH	21.563	YAKUTİYE	21
ÇAYIRLI	20.984	SERİK	18
REFAHİYE	20.900	GÜRÜN	15
İLİÇ	12.781	YENİŞEHİR	14
OTLUKBELİ	5.012	ÇAT	14
KEMALİYE	2.829	ŞİRAN	14

Tablo 3.14’e göre Erzincan iline ait ilçeler sol sütunda sunulmuştur. Polikliniklere Erzincan Merkez ilçesinden 1.028.080 hasta başvururken, bu sayıyı 57.486 ile Üzümlü, 35.875 ile Tercan ilçesinde yaşayan hastaların başvuruları takip etmiştir. Kelkit, Patnos, Eyüp Sultan, Yakutiye gibi komşu illere ait ilçelerin de polikliniklere en fazla başvuru yapan ilçeler arasında olduğu veri setinden anlaşılmaktadır.

3.5.3.2 Birliktelik Kuralı Kullanarak Poliklinik Analizi

3.5.3.2.1. Verilerin Birliktelik Kuralına Uygun Hale Getirilmesi

Birliktelik kuralı ile poliklinik veri analizi yapabilmek için öncelikle ziyaret hareketlerinin, bu kurala göre düzenlenmesi gerekmektedir. Veriler Apriori algoritmasına uygun olacak şekilde her bir hastaya göre gruplandırılmıştır.

```
SELECT `dosyano`, `kurum`, `cinsiyet`, `yas`, `il`, `ilce`, `uyruk`, `vizittarihi`,  
`vizitsaati`, GROUP_CONCAT(`birimbrans` SEPARATOR ' ', ') AS 'Poliklinikler'  
FROM `vizitler` GROUP BY `dosyano`
```

Yukarıdaki SQL kodu sağlık kurumundan alınan Excel verilerini vizitler tablosuna aktarılan veri seti formatına dönüştürmüştür. Hastaya ait demografik bilgiler ve polikliniklerin birliktelik kuralı analizine uygun biçimde düzenlenmiş veri seti aşağıda Tablo 3.15’da gösterilmektedir.

Tablo 3.15: Birliktelik Kuralı Analizine Uygun Veri Seti Örneği

Hasta No	Poliklinikler
1	Acil, Diğer, Kardiyoloji, Diğer, Kardiyoloji, Yoğun Bakım
2	İç Hastalıkları, İç Hastalıkları, Göz Hastalıkları, Fiziksel Tıp ve Rehabilitasyon, Göz Hastalıkları, Endokrinoloji ve Metabolizma Hastalıkları
3	Göz Hastalıkları, Kulak-Burun-Boğaz Hastalıkları
4	Kardiyoloji

Tablo 3.16’ya göre hastaların geldiği her bir ziyaret hareketi, tek bir satırda gruplandırılmıştır.

Tablo 3.16: Apriori Algoritmasına Uygun Veri Seti Örneği

Hasta No	Acil	Diğer	Kardiyoloji	Yoğun Bakım	İç Hastalıkları
1	1	1	1	1	0
2	0	0	0	0	1
3	0	0	0	0	0

Bazı hastaların 1 ziyaret hareketi varken bazı hastaların ise daha fazla ziyaret hareketi mevcuttur ve ziyaret hareketleri SQL de GROUP_CONCAT komutu ile birleştirilmiş ve analizin ilk aşaması tamamlanmıştır. Tablo 3.16’ de Apriori algoritmasına uygun veri seti örneği görülmektedir.

Tablo 3.16’ dan anlaşılacağı üzere verinin, Apriori algoritmasına uygun format olan Boolean veri tipine yani 0 veya 1 değerlerinin olduğu ya da True veya False değerlerinin olduğu veri setine dönüştürülmesi gerekmektedir. Boolean tablosu, tüm ziyaretleri ayrı ayrı olarak açıklayan ve her ziyaretin hangi hastalarda mevcut olup olmadığını gösteren bir tablodur.

Frekans (yineleme sayısı) ise bir polikliniğin toplam ziyaret sayılarını gösterir iken *Destek (Support) değeri*, verideki öğeler arasındaki bağlantının ne kadar çok olduğunu göstermektedir. Destek değeri 0-1 arasında iken Min Support değerini ise tarafımızdan belirlenir. Örneğin;

Kardiyoloji Frekans Değeri = 25465

Acil Frekans Değeri = 135142

Destek değeri ise her poliklinik için frekans değerinin toplam giriş sayısına bölünmesiyle elde edilmektedir. Bu nedenle limiti 0 ile 1 arasında olması beklenmektedir.

Destek = Frekans / Toplam Başvuru Minimum destek değeri tarafımızdan belirlenir ve değiştirilebilir.

```

data["birimbrans"].replace(" ","_",regex=True,inplace=True)

datax = list(data["birimbrans"].apply(lambda x:x.split(',')))

#Tüm Tanılar için analiz

te = TransactionEncoder()

te_data = te.fit(datax).transform(datax)

df = pd.DataFrame(te_data,columns=te.columns_)

```

Yukarıdaki Python kodu, verilerin Tablo 3.17’deki şekilde Apriori Algoritmasına uygun hale getirilmesini sağlamaktadır. Öncelikle ‘**data["birimbrans"].replace(" ","_",regex=True, inplace=True)**’ kodunu kullanarak *birimbrans* değişkeninde bulunan poliklinik adları boşluklardan temizlenmiş, daha sonra ‘**datax = list(data["birimbrans"].apply(lambda x:x.split(',')))**’ kodu ile her bir satırda virgül (,) ile ayrılan poliklinik başvuruları bir listeye daha aktarılmıştır. Her bir poliklinik adından 1 sütun (columns) oluşturulmuş ve apriori için uygun biçimine dönüştürülmüştür. Şekil 3.18’de Apriori algoritmasına uygun hale dönüştürülmüş veri seti çıktısı örneği görülmektedir.

Şekil 3.18: Apriori Algoritmasına Uygun Hale Dönüştürülmüş Veri Seti Çıktısı

	Acil	Aile Hekimliği	Algoloji (Anesteziyoloji ve Reanimasyon)	...	Üroloji	İç Hast	İç Hastalıkları
0	False	False	False	...	False	False	True
1	False	False	False	...	False	False	False
2	False	False	False	...	False	False	False
3	True	False	False	...	False	False	False
4	False	False	False	...	False	False	True
...	...	...	...	...	...	...	...
99369	False	False	False	...	False	False	False
99370	False	True	False	...	False	False	False
99371	False	False	False	...	False	False	True
99372	False	True	False	...	False	False	False
99373	False	True	False	...	False	False	False

Analizin il adımında öncelikle destek değeri bulunmalıdır. Bu amaçla Python programlama dili kullanılmış ve Apriori algoritması kodları çalıştırılarak veri seti destek değerleri bulunmuştur. Bu sayede daha önce belirlenen minimum destek değerinden büyük veya ona eşit destek değerlerine sahip poliklinikleri içeren yeni bir tablo

oluşturulmuştur. Minimum destek değeri %10 olarak tanımlanmıştır. Destek değeri bu değerin altında olan poliklinikler aşağıdaki tabloda yer almamaktadır, Minimum destek değeri verilerin çok fazla olmasından kaynaklı olarak %10 seçilmiştir. Tablo 3.17’de Apriori analiz sonucunda polikliniklere dair elde edilen destek değerleri görülmektedir.

Tablo 3.17: Apriori Analiz Sonucu Veri Seti

Sıra No	Destek Değeri	Poliklinikler
1	0.662483	(Acil)
2	0.210815	Göz Hastalıkları
3	0.190488	(DİĞER)
4	0.189066	İç Hastalıkları
5	0.156048	Ortopedi ve Travmatoloji
6	0.141646	(Acil, DİĞER)
7	0.132129	(Göz Hastalıkları, Acil)
8	0.128501	(Acil, İç Hastalıkları)
9	0.126332	Kulak Burun Boğaz Hastalıkları
10	0.118390	(Deri ve Zührevi Hastalıkları)
11	0.117606	(Kardiyoloji)
12	0.108174	(Ortopedi ve Travmatoloji, Acil)

Sonuçlara göre destek değerleri en büyük olan iki poliklinik %66 ile Acil, %21 ile Göz Hastalıkları olmuştur. Tablo 3.17’de 6, 7 ve 8 sıra no ile yer alan çoklu poliklinik gruplarının birliktelik halindeki destek değerlerinin minimum destek değeri olan %10 dan fazla olduğu görülmektedir. Bu adımdan sonra ise Güven değeri hesaplanmıştır. Destek ve güven değerleri ne kadar büyük olursa birliktelik kuralı da o kadar güçlü olmaktadır.

$$\text{Güven (X -> Y)} = \text{Frekans(X, Y)} / \text{Frekans(X)}$$

Örneğin, X ve Y poliklinikleri için Destek değerinin %10 ve Güven değerinin %50 olduğunu varsayarak Python programlama dilinde bir güven analizi kodu çalıştırıldığında Tablo 3.18’deki sonuçlar elde edilmektedir (destek ve guven, 2024, <https://www.veribilimiokulu.com/associationrulesanalysis>).

Tablo 3.18: Güven Değeri Sonucu Veri Seti

Antecedents	consequents	Antecedent support	Consequent support	support	confidence	lift
Göz Hastalıkları	Acil	0.210815	0.662483	0.132129	0.626751	0.946063
İç Hastalıkları	Acil	0.189066	0.662483	0.128501	0.679663	1.025.933
Ortopedi	Acil	0.156048	0.662483	0.108174	0.693209	1.046.379
DİĞER	Acil	0.190488	0.662483	0.141646	0.743594	1.122.434

Lift, bir kuralın öncülleri ile sonuçları arasındaki bağımlılığı ölçer. Lift değeri 1'den büyükse, öncüller ve sonuçlar arasında pozitif bir bağımlılık vardır. Lift değeri 1'den büyük olan kurallar, öncüller ve sonuçlar arasında anlamlı bir ilişki olduğunu gösterir. Lift değeri 1'den küçükse, bu bağımlılık negatif veya tesadüfi olabilir. *Consequents (Sonuçlar)* ise Birliktelik kuralının sağ tarafındaki öğeler veya olaylardır. Güven değeri %50 den fazla olan polikliniklerin birliktelik kuralına göre analiz edildiği durumda Acil polikliniklerinin en çok eşleşen poliklinik olarak görülmüştür. Yine sonuçlara göre, Diğer (Tanımsız veya Boş) polikliniği ile Acil poliklinikleri %74 oranında birbirleri ile bağlantılıdır. Ortopedi ve Travmatoloji %69, İç Hastalıkları %67 ve Göz Hastalıkları ise %62 oranında Acil poliklinikleri ile bağlantılı olduğu görülmektedir. Bağlantı durumları Acil polikliniğine giden hastaların aynı zaman dilimi içerisinde diğer polikliniklere de gitme durumunun oldukça sık olduğunu ifade etmektedir. Hastalar Acil polikliniğine çok fazla başvurduklarından dolayı diğer polikliniklerin birliktelik kuralı sonuçları Acil Poliklinikleri ile bağlantılı gerçekleşmektedir.

Acil ve DİĞER poliklinikler veri analizine dahil edilmeden bir Apriori analizi yapıldığında daha farklı sonuçlar ortaya çıkacağı düşünülmüştür. Buradan hareketle Tablo 3.17 ve Tablo 3.18'deki sonuçlar için yapılan işlemler, Acil ve DİĞER polikliniklere ait veriler, veri setinden çıkarılarak tekrarlanmış ve Tablo 3.19 ve Tablo 3.20'deki sonuçlar elde edilmiştir.

Bu analizde farklı olarak minimum destek değeri (%) 4,5 olarak tanımlanmıştır. Destek değeri bu değerin altında olan poliklinikler Tablo 3.19’da yer almamaktadır. Buda değerin (%) 4.5 olması daha öncede bahsettiğimiz gibi tamamen bizim tarafımızdan belirlenmektedir. Bu değer sabit olmamak ile beraber veri setinin büyüklüğüne göre değişmektedir. Analiz yaparken minimum destek değerini farklı vererek en uygun sonuca ulaşılmıştır.

Tablo 3.19: Acil ve Diğer Birimler Olmadan Çıkan Sonuçlar Veri Seti

Sıra No	Destek Değeri	Poliklinikler
1	0.267602	Göz Hastalıkları
2	0.239994	İç Hastalıkları
3	0.198082	Ortopedi ve Travmatoloji
4	0.160361	Kulak Burun Boğaz Hastalıkları
5	0.150280	Deri ve Zührevi Hastalıkları
6	0.149285	Kardiyoloji
7	0.112845	Çocuk Sağlığı ve Hastalıkları
8	0.107343	Fiziksel Tıp ve Rehabilitasyon
9	0.107021	Nöroloji
10	0.106211	Göğüs Hastalıkları
11	0.095762	Kadın Hastalıkları ve Doğum
12	0.090091	Üroloji
13	0.089862	Genel Cerrahi
14	0.078699	Ruh Sağlığı ve Hastalıkları
15	0.069826	Göz Hastalıkları, İç Hastalıkları
16	0.059398	Kardiyoloji, İç Hastalıkları
17	0.056960	Beyin ve Sinir Cerrahisi
18	0.055833	Ortopedi ve Travmatoloji, İç Hastalıkları
19	0.052590	Enfeksiyon Hastalıkları ve Klinik Mikrobiyoloji
20	0.049765	Ortopedi ve Travmatoloji, Göz Hastalıkları
21	0.048271	Kulak Burun Boğaz Hastalıkları, İç Hastalıkları
22	0.046104	Göz Hastalıkları, Kulak Burun Boğaz Hastalıkları

Sonuçlara göre %26 ile Göz Hastalıkları, %23 ile İç Hastalıkları polikliniklerinin destek değerleri birliktelik kurallarına uygundur. 15.inci sıradan sonra yer alan gruplarda birden fazla polikliniğin birliktelik halindeki destek değerlerinin minimum destek değeri

olan %4,5 dan fazla olduğu görülmektedir. Bu adımdan sonra ise Güven değerini hesaplanmıştır. Destek değeri %4,5 ve Güven değeri %10 seçilerek Python programlama dili çerçevesinde güven analizi çalıştırıldığında Tablo 3.20’deki sonuçlar elde edilmiştir.

Tablo 3.20: Acil ve Diğer Birimler olmadan Güven Destek Değerine Göre Çıkan Sonuç Veri Seti

Antecedents	consequents	Antecedent support	Consequent support	support	confidence	lift
Fiziksel Tıp ve Rehabilitasyon	İç Hastalıkları	0.107343	0.239994	0.045966	0.428219	1.784.291
Kardiyoloji	İç Hastalıkları	0.149285	0.239994	0.059398	0.397882	1.657.886
Kulak Burun Boğaz Hasta..	İç Hastalıkları	0.160361	0.239994	0.048271	0.301014	1.254.260
İç Hastalıkları	Göz Hastalıkları	0.239994	0.267602	0.069826	0.290950	1.087.251
Kulak Burun Boğaz Hastalıkları	Göz Hastalıkları	0.160361	0.267602	0.046104	0.287500	1.074.356
Ortopedi ve Travmatoloji	İç Hastalıkları	0.198082	0.239994	0.055833	0.281871	1.174.494
Göz Hastalıkları	İç Hastalıkları	0.267602	0.239994	0.069826	0.260933	1.087.251
Ortopedi ve Travmatoloji	Göz Hastalıkları	0.198082	0.267602	0.049765	0.251236	0.938841
İç Hastalıkları	Kardiyoloji	0.239994	0.149285	0.059398	0.247498	1.657.886
İç Hastalıkları	Ortopedi ve Travmatoloji	0.239994	0.198082	0.055833	0.232646	1.174.494
İç Hastalıkları	Kulak Burun Boğaz Hastalıkları	0.239994	0.160361	0.048271	0.201135	1.254.260
İç Hastalıkları	Fiziksel Tıp ve Rehabilitasyon	0.239994	0.107343	0.045966	0.191531	1.784.291
Göz Hastalıkları	Ortopedi ve Travmatoloji	0.267602	0.198082	0.049765	0.185967	0.938841
Göz Hastalıkları	Kulak Burun Boğaz Hastalıkları	0.267602	0.160361	0.046104	0.172285	1.074.356

Tablo 3.20’e göre, poliklinikler arasındaki ilişkiyi değerlendirmek amacıyla birliktelik kuralları analizi kullanılarak elde edilen sonuçları incelemektedir. Analiz sonucunda elde edilen veriler, her bir kuralın öncül ve sonuç destek değerleri, kuralın genel destek değeri, güven ve lift değerleri ile açıklanmıştır.

1. Fiziksel Tıp ve Rehabilitasyon → İç Hastalıkları

Öncül Desteği (Antecedent Support): 0.107343 Fiziksel Tıp ve Rehabilitasyon bölümünün tüm verilerde görülme oranı %10.73'tür. Sonuç Desteği (Consequent Support): 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur. Destek (Support): 0.045966 Hem Fiziksel Tıp ve Rehabilitasyon hem de İç Hastalıkları bölümlerinin birlikte görülme oranı %4.60'tır. Güven (Confidence): 0.428219 Fiziksel Tıp ve Rehabilitasyon bölümünün bulunduğu durumların %42.82'sinde İç Hastalıkları da bulunmaktadır. Lift (Kaldırma): 1.784291 Fiziksel Tıp ve Rehabilitasyon ve İç Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.784 kat daha fazladır. Bu, iki bölüm arasında pozitif bir bağımlılık olduğunu göstermektedir.

2. Kardiyoloji → İç Hastalıkları

Öncül Desteği: 0.149285 Kardiyoloji bölümünün tüm verilerde görülme oranı %14.93'tür. Sonuç Desteği: 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur. Destek: 0.059398 Hem Kardiyoloji hem de İç Hastalıkları bölümlerinin birlikte görülme oranı %5.94'tür. Güven: 0.397882 Kardiyoloji bölümünün bulunduğu durumların %39.79'unda İç Hastalıkları da bulunmaktadır. Lift: 1.657886 Kardiyoloji ve İç Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.658 kat daha fazladır. Bu, iki bölüm arasında pozitif bir bağımlılık olduğunu göstermektedir.

3. Kulak Burun Boğaz Hastalıkları → İç Hastalıkları

Öncül Desteği: 0.160361 Kulak Burun Boğaz Hastalıkları bölümünün tüm verilerde görülme oranı %16.04'tür. Sonuç Desteği: 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur. Destek: 0.048271 Hem Kulak Burun Boğaz Hastalıkları hem de İç Hastalıkları bölümlerinin birlikte görülme oranı %4.83'tür. Güven: 0.301014 Kulak Burun Boğaz Hastalıkları bölümünün bulunduğu durumların %30.10'unda İç Hastalıkları da bulunmaktadır. Lift: 1.254260 Kulak Burun Boğaz Hastalıkları ve İç Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.254 kat daha fazladır. Bu, iki bölüm arasında pozitif bir bağımlılık olduğunu göstermektedir.

4. İç Hastalıkları → Göz Hastalıkları

Öncül Desteği: 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur. Sonuç Desteği: 0.267602 Göz Hastalıkları bölümünün tüm verilerde görülme oranı %26.76'dır.

Destek: 0.069826 Hem İç Hastalıkları hem de Göz Hastalıkları bölümlerinin birlikte görülme oranı %6.98'dir. Güven: 0.290950 İç Hastalıkları bölümünün bulunduğu durumların %29.10'unda Göz Hastalıkları da bulunmaktadır. Lift: 1.087251 İç Hastalıkları ve Göz Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.087 kat daha fazladır. Bu, iki bölüm arasında hafif bir pozitif bağımlılık olduğunu göstermektedir.

5. Kulak Burun Boğaz Hastalıkları → Göz Hastalıkları

Öncül Desteği: 0.160361 Kulak Burun Boğaz Hastalıkları bölümünün tüm verilerde görülme oranı %16.04'tür. Sonuç Desteği: 0.267602 Göz Hastalıkları bölümünün tüm verilerde görülme oranı %26.76'dır. Destek: 0.046104 Hem Kulak Burun Boğaz Hastalıkları hem de Göz Hastalıkları bölümlerinin birlikte görülme oranı %4.61'dir. Güven: 0.287500 Kulak Burun Boğaz Hastalıkları bölümünün bulunduğu durumların %28.75'inde Göz Hastalıkları da bulunmaktadır. Lift: 1.074356 Kulak Burun Boğaz Hastalıkları ve Göz Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.074 kat daha fazladır. Bu, iki bölüm arasında hafif bir pozitif bağımlılık olduğunu göstermektedir.

6. Ortopedi ve Travmatoloji → İç Hastalıkları

Öncül Desteği: 0.198082 Ortopedi ve Travmatoloji bölümünün tüm verilerde görülme oranı %19.81'dir. Sonuç Desteği: 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur. Destek: 0.055833 Hem Ortopedi ve Travmatoloji hem de İç Hastalıkları bölümlerinin birlikte görülme oranı %5.58'dir. Güven: 0.281871 Ortopedi ve Travmatoloji bölümünün bulunduğu durumların %28.19'unda İç Hastalıkları da bulunmaktadır. Lift: 1.174494 Ortopedi ve Travmatoloji ve İç Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.174 kat daha fazladır. Bu, iki bölüm arasında pozitif bir bağımlılık olduğunu göstermektedir.

7. Göz Hastalıkları → İç Hastalıkları

Öncül Desteği: 0.267602 Göz Hastalıkları bölümünün tüm verilerde görülme oranı %26.76'dır.

Sonuç Desteği: 0.239994 İç Hastalıkları bölümünün tüm verilerde görülme oranı %23.99'dur.

Destek: 0.069826 Hem Göz Hastalıkları hem de İç Hastalıkları bölümlerinin birlikte görülme oranı %6.98'dir. Güven: 0.260933 Göz Hastalıkları bölümünün bulunduğu durumların %26.09'unda İç Hastalıkları da bulunmaktadır. Lift: 1.087251 Göz Hastalıkları ve İç Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından 1.087 kat daha fazladır. Bu, iki bölüm arasında hafif bir pozitif bağımlılık olduğunu göstermektedir.

8. Ortopedi ve Travmatoloji → Göz Hastalıkları

Öncül Desteği: 0.198082 Ortopedi ve Travmatoloji bölümünün tüm verilerde görülme oranı %19.81'dir. Sonuç Desteği: 0.267602 Göz Hastalıkları bölümünün tüm verilerde görülme oranı %26.76'dır. Destek: 0.049765 Hem Ortopedi ve Travmatoloji hem de Göz Hastalıkları bölümlerinin birlikte görülme oranı %4.98'dir. Güven: 0.251236 Ortopedi ve Travmatoloji bölümünün bulunduğu durumların %25.12'sinde Göz Hastalıkları da bulunmaktadır. Lift: 0.938841 Ortopedi ve Travmatoloji ve Göz Hastalıkları bölümlerinin birlikte görülme olasılığı, bu bölümlerin bağımsız olarak görülme olasılığından %93.88'dir. Bu, iki bölüm arasında bağımsız bir ilişki olduğunu göstermektedir.

Birliktelik kuralı analizinde kullanılan Python Kodlarının tamamı Tablo 3.21'de sunulmuştur.

Tablo 3.21: Projenin Python Kodlarının Tamamı

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import os
import csv
import datetime
import docx

import seaborn as sns

from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, association_rules

print("Veriler Çekiliyor \n ")
data= pd.read_csv("poliklinikanaliz.csv", usecols =
['birimbrans'],sep=';',low_memory=False)

#print("Veriler Listeleniyor\n",data)

data["birimbrans"].replace(" ","_",regex=True, inplace=True)
selRows = data[data['birimbrans'] == 'Acil'].index
data = data.drop(selRows, axis=0)

selRows = data[data['birimbrans'] == 'DİĞER'].index
data = data.drop(selRows, axis=0)

datax = list(data["birimbrans"].apply(lambda x:x.split(',')))

#Tüm Tanılar için analiz
te = TransactionEncoder()
te_data = te.fit(datax).transform(datax)
df = pd.DataFrame(te_data,columns=te.columns_)

print (df.head(10))

#df.to_csv('gxd.csv')

# Apriori Algoritması Çalışıyor

df1 = apriori(df, min_support=0.05, use_colnames=True, verbose = 1)
df1 =df1.sort_values(by="support", ascending=False)
print("apriori Çalıştı \n",df1)

doc = docx.Document()
table = doc.add_paragraph(str(df1))

doc.add_page_break()
```

```

doc.save('apriori.docx')

# Apriori Algoritması Çalıştı

# confidence Çalışıyor
confidence = association_rules(df1, metric = "confidence", min_threshold = 0.10)
confidence.sort_values('confidence', ascending = False, inplace = True)
print("confidence Çalıştı \n",confidence)
#f.write("confidence Çalıştı \n"+confidence.to_string()+"\n")

doc = docx.Document()
table = doc.add_paragraph(str(confidence))

doc.add_page_break()

doc.save('confidence.docx')

# confidence Çalıştı

# lift Çalışıyor
lift = association_rules(df1, metric = "lift", min_threshold = 0.5)
lift.sort_values('lift', ascending = False, inplace = True)
print("Lift Çalıştı \n",lift)
#f.write("Lift Çalıştı \n"+lift.to_string()+"\n")

doc = docx.Document()
table = doc.add_paragraph(str(lift))

doc.add_page_break()

doc.save('lift.docx')

# lift Çalıştı

```

SONUÇ VE TARTIŞMA

Veri madenciliği, büyük veri tabanlarından daha önce keşfedilmemiş, kullanılabilir ve geçerli bilgilerin keşfedilmesi süreci olarak ifade edilebilmektedir. Birlikte kuralı madenciliği, ilişkisel veritabanları, işlemsel veritabanları ve diğer veri deposu biçimleri de dahil olmak üzere çeşitli veri tabanlarında ortak kalıpları, korelasyonları, ilişkileri veya nedensel yapıları bulmaya yönelik bir yöntemdir. Sağlık alanındaki birçok sorunun çözümünde veri madenciliği teknikleri kullanılmaya başlanmıştır. Özellikle sağlık alanında verilerin düzenli ve çok büyük boyutlarda olması bu alanı veri madenciliği için elverişli kılmaktadır.

Bu çalışmada Erzincan Binali Yıldırım Üniversitesi Mengücek Gazi Eğitim ve Araştırma hastanesine ait 2023 yılı ziyaret hareketleri veri seti incelenmiştir. Çalışmada, birlikte kurallarından Apriori algoritması kullanılarak poliklinik hareketleri incelenmiştir. Hastaneye ait veriler analiz edilirken Python ve kullanılan kütüphaneler ile birlikte MYSQL ve SQL Sorgulama dilinden faydalanılmış hastanenin 2023 yılı hasta profili çıkarılmış ve hastaların hangi polikliniklere daha çok başvurdukları görülmüştür. Toplamda 248.924 farklı hasta kaydına ait olan 1.447.705 ziyaret hareketi birlikte kuralı ile yapılan analiz sonucunda incelendiğinde en çok hasta başvurusunun 496.884 ile Acil polikliniklerine olduğu görülmüştür. Göz Hastalıkları, İç Hastalıkları ve Ortopedi ve Travmatoloji poliklinikleri ise en fazla başvuru alan diğer poliklinikler olarak ön çıkmaktadır. Analiz çerçevesinde hastaların cinsiyet, yaş, uyruk, ikamet ettiği il ve ikamet ettiği ilçeler baz alınarak hastanenin hasta profilleri çıkarılmıştır. Birlikte kuralları analizi sonuçlarına göre Acil polikliniklerine gelen hastaların daha sonra Göz Hastalıkları, İç Hastalıkları ve Ortopedi ve Travmatoloji polikliniklerine en çok başvurdukları ortaya çıkmıştır.

İlk analizimizde belirgin oranda tüm poliklinikleri Acil ve tanımlayan poliklinik başvuruları etkilediği görülmüştür bu nedenle Acil ve Diğer (tanımsız,boş) poliklinikler veri setinden çıkarıldıktan sonra en çok birbirleri ile eşleşen polikliniklerin analizi yapıldığında Fiziksel Tıp ve Rehabilitasyon, Kardiyoloji, Kulak Burun Boğaz Hastalıkları polikliniklerine gelen hastaların İç Hastalıkları polikliniklerine başvurma olasılıkları yüksek olduğu görülmüştür.

Birliktelik kuralları analizi sonuçlarının, hasta yoğunluđuna göre poliklinik planlaması yapılmasında faydalı olacađı düşünölmektedir. Bu süreç çerçevesinde birbirleri ile ilişkili olan polikliniklerin yerleşim planlaması içerisinde olabildiğince yakın konumlandırılması sonucu tesis içerisindeki hasta ve donanım hareketliliğinin asgari düzeylere indirilebilmesi mümkün olabilecektir. Böylesi yerleşim kararlarının, sağıık işletmesi için uzun vadede maliyetlerin azaltılması ve verimliliğinin artırılması amaçlarına hizmet edeceđi öngörölmektedir. Ayrıca veri seti neticesinde yapılan demografik analizler, sağıık işletmesinin sahip olduđu kısıtlı kaynakların hangi alanlara yoğunlaştırılması gerektiđi konusunda genel bir fikir sunmaktadır. Analiz ile elde edilen verilere göre MHRS randevu yoğunluđunu planlamak amacı ile polikliniklerin aylık, günlük ve saatlik vizit sayıları dikkate alınmasının randevu planlama açısından faydalı olacađı öngörölmektedir.

KAYNAKÇA

- Ahmad, P., Qamar, S., & Rizvi, S. Q. A. (2015). Techniques of Data Mining in Healthcare: A Review. *International Journal of Computer Applications*, 120(15).
- Akgün, E., & Maral, Ö. (2022). Uzaktan Eğitimde Veri Madenciliği Yöntemi Kullanılarak Yapılmış Araştırmalarda Öğrenme Çıktıları Üzerine Sistemik Bir İnceleme. *Bilgi Ve İletişim Teknolojileri Dergisi*, 4(2), 197-226.
- Akpınar, H. (2000). Veri Tabanlarında Bilgi Keşfi ne Veri Madenciliği. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29(1), 1-22.
- Alan, M. A., & Yeşilyurt, C. (2019). Birliktelik Kuralları Madenciliği İle Yatan Hasta Profiline Çıkarılması. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 23(4), 1917-1926.
- Altındış, S. (2018). Büyük Verinin Sağlık Hizmetleri Kalitesindeki Rolü. *Sakarya Tıp Dergisi*, 8(2), 205-213.
- Apriori. (2024). <https://cotur.medium.com/apriori-algoritması-a90a38856373>
- Aşık, D. (2019). *Böbrek Nakli Yapılan Hastalarda Veri Madenciliği Yöntemleri İle Akut Rejeksiyon Durumunun İncelenmesi*. (Yayınlanmamış Yüksek Lisans Tezi), Akdeniz Üniversitesi, Sağlık Bilimleri Enstitüsü, Antalya.
- Ayaz, N., & Türkmen, B. M. (2018). Yöresel Yiyecekleri Konu Alan Lisansüstü Tezlerin Bibliyometrik Analizi. *Gastroia. Journal of Gastronomy And Travel Research*, 2(1), 22-38.
- Babacan, A. (2023). *Sağlık Bakanlığı Hastanelerinde Finansal Performans Analizi: Türkiye Örneği*, (Yayınlanmamış Yüksek Lisans Tezi). Kırıkkale Üniversitesi, Sağlık Bilimleri Enstitüsü, Kırıkkale
- Bamakan, S. M. H., Malekinejad, P., & Ziaei, M. (2022). Towards Blockchain-Based Hospital Waste Management Systems; Applications And Future Trends. *Journal Of Cleaner Production*, 349, 131440.
- Başak, O., Uğur, B., & Samur, M. K. (2009). *Kulak Burun Boğaz Epikriz Notlarından Birliktelik Kurallarının Çıkarılması*.
- Brundtland, G. H. (1987). What Is Sustainable Development. *Our Common Future*, 8(9).
- Cansever, İ. H. (2021). Sürdürülebilir Kalkınma Ve Sağlık: Türkiye'nin 2023 Hedefleri İle Karşılaştırmalı Bir Değerlendirme. *Hacettepe Sağlık İdaresi Dergisi*, 24(3), 633-650.
- Ceylan, S. (2023). *Temel Veri Madenciliği Algoritmalarının Başarımlarının Endokrin Veri Seti Üzerinde Karşılaştırılması* (Yayınlanmamış Yüksek Lisans Tezi). Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü, Denizli.
- Cinaroglu, S., & Avcı, K., (2014). Dünya Sağlık Örgütü'ne Üye Ülkelerin Doğumda Beklenen Yaşam Süresi Bakımından Veri Madenciliği Yöntemleri Kullanılarak Sınıflandırılması. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri Ve İstatistik Dergisi*, 20, 88-103.
- Çarklı, B. (2010). *Sağlık Sektöründe Apriori Algoritması İle Bir Veri Madenciliği Uygulaması*. (Yayınlanmamış Yüksek Lisans Tezi), Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, Sakarya.

- Çarklı, B. (2010). *Sağlık Sektöründe Apriori Algoritması İle Bir Veri Madenciliği Uygulaması* (Yayınlanmamış Doktora Tezi). Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya.
- Çataloluk, H. (2012). *Gerçek Tıbbi Veriler Üzerinde Veri Madenciliği Yöntemlerini Kullanarak Hastalık Teşhisi* (Yayınlanmamış Yüksek Lisans Tezi). Bilecik Üniversitesi, Fen Bilimleri Enstitüsü, Bilecik.
- Çelik, Y. (2006). Sürdürülebilir Kalkınma Kavramı Ve Sağlık. *Hacettepe Sağlık İdaresi Dergisi*, 9(1), 19-37.
- Çerkezi, M. (2013). *Veri Madenciliği Yöntemlerini Kullanarak Diyabetik Retinopati Hastalığının Teşhisi* (Yayınlanmamış Doktora Tezi). Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.
- Demirci, Ş., Konca, M., & İlgün, G. (2021). Sağlık Hizmetlerinin Performansının Değerlendirilmesinde Kullanılan Değişkenlerin Birliktelik Kuralları Analizi İle İncelenmesi. *Hacettepe Sağlık İdaresi Dergisi*, 24(3), 507-532.
- Demirok, Y. (2019). *Birliktelik Kuralı Yöntemleri İle E-Ticaret Satışlarının Analizi*. (Yayınlanmamış Yüksek Lisans Tezi). Maltepe Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Deniz, N. (2016). *Sağlık sektöründe hizmet kalite algısının ölçümüne yönelik ölçek geliştirme çalışması*. (Yayınlanmamış doktora tezi). Eskişehir Osmangazi Üniversitesi, Sosyal Bilimler Enstitüsü, Eskişehir.
- Destek ve Guven, (Şubat, 2024), <https://www.veribilimiokulu.com/associationruleanalysis/>
- Doğrul, G., Akay, D., & Kurt, M. (2015). Trafik Kazalarının Birliktelik Kuralları İle Analizi. *Gazi Mühendislik Bilimleri Dergisi*, 1(2), 265-283.
- Dünya Sağlık Örgütü (18.03.2022). <https://www.who.int/about/governance/constitution>.
- Ertuğrul, İ., Organ, A., & Şavlı, A. (2013). Veri Madenciliği Uygulamasına İlişkin PAÜ Hastanesinde Hasta Profiline Belirlenmesi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 19(2), 97-103.
- Gedik, Y. (2021). Sürdürülebilir Tedarik Zinciri Yönetimi Ve Sürdürülebilirliğin Tedarik Zincirleri Üzerindeki Etkileri: Kavramsal Bir Değerlendirme. *Uluslararası Yönetim İktisat Ve İşletme Dergisi*, 17(3), 830-860.
- Irmak, S., Köksal, C. D., & Asilkan, Ö. (2012). Hastanelerin Gelecekteki Hasta Yoğunluklarının Veri Madenciliği Yöntemleri İle Tahmin Edilmesi. *Uluslararası Alanya İşletme Fakültesi Dergisi*, 4(1), 101-114.
- Jothi, N., & Husain, W. (2015). Data Mining In Healthcare—A Review. *Procedia Computer Science*, 72, 306-313.
- Karabulut, (2021). *Hastane Bilgi Yönetim Sistemlerinde Veri Madenciliği: Hasta Profil Tahmini*. (Yayınlanmamış Yüksek Lisans Tezi). Karabük Üniversitesi, Lisansüstü Eğitim Enstitüsü, Karabük.
- Karaca, P.Ö., Atılgan, E., & Zekioğlu, A. (2018). Sağlık Hizmetlerinde Sürdürülebilirlik Bağlamında İnovatif Bir Uygulama: Yeşil Hastaneler. *Ejovoc (Electronic Journal Of Vocational Colleges)*, 8(2), 77-87.

- Kavurkacı, Ş., Aydın, Z. K., & Şamlı, R. (2011). Büyük Ölçekli Veri Tabanlarında Bilgi Keşfi. *Akademik Bilişim Konferansları*, 2-4.
- Koh, H. C., & Tan, G. (2011). Data Mining Applications In Healthcare. *Journal Of Healthcare Information Management*, 19(2), 65.
- Koyuncugil, A., & Özgülbaş, N. (2009). Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları. *Bilişim Teknolojileri Dergisi*, 2(2).
- Matplotlib. 29.01.2024, <https://tr.wikipedia.org/wiki/matplotlib>
- MYSQL. 29.01.2024, <https://www.protan.com.tr/mysql-nedir>
- Numpy. 29.01.2024, <https://www.pythonkod.com/python-numpy-kutuphanesi-kurulumu>
- Özbay, Ö. (2015). Veri Madenciliği Kavramı Ve Eğitimde Veri Madenciliği Uygulamaları. *Uluslararası Eğitim Bilimleri Dergisi*, 5, 262-272.
- Pandas. (29.01.2024), <https://tr.wikipedia.org/wiki/pandas>
- Poyraz, O. (2012). *Tıpta Veri Madenciliği Uygulamaları: Meme Kanseri Veri Seti Analizi*. (Yayınlanmamış Yüksek Lisans Tezi). Trakya Üniversitesi Fen Bilimleri Enstitüsü, Edirne.
- Python. 29.01.2024, <https://tr.wikipedia.org/wiki/python>
- R. Veloso, F. Portela, M. F. Santos, Á. Silva, F. Rua, A. Abelha, And J. Machado, (2014). "A Clustering Approach For Predicting Readmissions In Intensive Medicine," *Procedia Technol.*, 16, 1307–1316.
- Sağlık Bakanlığı, (18.03.2024), <https://www.saglik.gov.tr/tr-101548/bakan-koca-2023-yili-saglik-verilerini-acikladi.html>
- Sağlık Hizmetlerinin Sosyalleştirilmesi Hakkında Kanun (1961,12 Ocak), 18.03.2022, <https://www.mevzuat.gov.tr/mevzuatmetin/1.4.224.pdf>
- Sargutan, A. E. (2005). Sağlık Sektörü Ve Sağlık Sistemlerinin Yapısı. *Hacettepe Sağlık İdaresi Dergisi*, 8(3), 400-428.
- Seaborn. (29.01.2024), <https://www.veribilimiokulu.com/seaborn-ile-veri-gorsellestirmesi>
- Sumathi, S., Sivanandam, S.N. (2006). *Introduction to Data Mining And Its Applications*. Springer, Berlin.
- Taş, Y. (2018). *Birlikte Kuralları Madenciliği Ve Bir Uygulama*. (Yayınlanmamış Yüksek Lisans Tezi). "Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, Sivas.
- Webb, G. I. (2003). Discovering associations with numeric variables. In KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 383-388)
- Yetginler, B. (2019). *Rahim Ağzı Kanserinin Veri Madenciliği Yöntemleri İle Sınıflandırılması*. (Yayınlanmamış Yüksek Lisans Tezi). Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, Kırıkkale.
- Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J. F., & Hua, L. (2012). Data Mining In Healthcare And Biomedicine: A Survey Of The Literature. *Journal Of Medical Systems*, 36, 2431-2448.
- Zakur, Y., & Flaih, L. (2023). Apriori Algorithm And Hybrid Apriori Algorithm In The Data Mining: A Comprehensive Review. In *E3S Web Of Conferences* (Vol. 448, P. 02021). EDP Sciences.

Zerman, M. (2018). *BirlikteK Kuralı Algoritmaları İle Büyük Veriler Üzerinde Analitik Analizler: Havaalanı Örneđi*. (Yayınlanmamış Yüksek Lisans Tezi). Haliç Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.



EKLER

EK 1. Erzincan İl Sağlık Müdürlüğü İzin Belgesi

T.C.
SAĞLIK BAKANLIĞI
ERZİNCAN İL SAĞLIK MÜDÜRLÜĞÜ
BİLİMSEL ARASTIRMA İZİN TALEPLERİ DEĞERLENDİRME KOMİSYONU

ARAŞTIRMACIYA AİT BİLGİLER	
Adı Soyadı	Murat ÇETİN
Kurumu / Üniversitesi	Erzincan Binali Yıldırım Üniversitesi Sağlık Bilimleri Fakültesi
Araştırma yapılacak yer	Erzincan Binali Yıldırım Üniversitesi Mengücek Gazı Eğitim Ve Araştırma Hastanesi
İletişim Bilgileri	05362545753
Araştırmanın konusu	Var / Madencilik ve Sağlık İşletmeleri Performansının İyileştirilmesi
Üniversite / Kurum onayı	Var / Yok
	Var
KOMİSYON GÖRÜŞÜ	
Karar / Görüş	Oybirliği / Oyçokluğu ile kabul edilmiştir
	Oybirliği ile kabul edilmiştir

EK 2. Etik Kurul İzin Belgesi



T.C
ERZİNCAN BİNALİ YILDIRIM ÜNİVERSİTESİ
İNSAN ARAŞTIRMALARI SOSYAL VE BEŞERİ BİLİMLER
ETİK KURULU KARARI

Etik Kurul Toplantı Tarihi	24/01/2024
Protokol No	01/03
Araştırma Başlığı	Veri Madenciliği ve Sağlık İşletmeleri Performansının İyileştirilmesi
Araştırma Türü	Nicel-Tarama (tanımlayıcı)
Araştırmacılar	Dr. Öğr. Üyesi Salih AKA (Sorumlu Araştırmacı) Murat ÇETİN (Araştırmacı)
Karar	Başvuru dosyanıza ait araştırmanız etik açıdan uygun bulunmuştur.
Açıklama: - Etik Kurul Onayı, uygulama ve/veya veri toplama için araştırmacının ilgili kurum veya kuruluşlardan izin alma sorumluluğunu ortadan kaldırmaz. - Kurul üyelerine ait araştırma önerileri görüşülürken, ilgili yönerge gereğince, öneri sahibi üye görüşmelere katılmamış ve oy kullanmamıştır.	

e-imzalıdır

Prof. Dr. Haydar EFE
İnsan Araştırmaları Sosyal ve Beşeri Bilimleri
Etik Kurul Başkanı

belge, güvenli elektronik imza ile imzalanmıştır.

ÖZGEÇMİŞ

Ad Soyad: Murat ÇETİN	
Eğitim Bilgileri	
Lisans	
Üniversite	Anadolu Üniversitesi
Fakülte	İşletme Fakültesi
Bölümü	İşletme
Yüksek Lisans	
Üniversite	
Enstitü Adı	
Ana Bilim Dalı	
Programı	
Makale ve Bildiriler (Varsa)	