

1. Kütüphanelerin Import Edilmesi ve Verinin Okunması

- **Açıklama:** Proje kapsamında gerekli kütüphaneler (pandas, numpy, seaborn, matplotlib) import edilerek veri seti bir DataFrame'e yüklenmiştir. Veriler incelenmiş ve sütunların veri tipleri aşağıdaki gibi belirlenmiştir:

	Column	Non-Null Count	Dtype
0	Kullanici_id	2357 non-null	int64
1	Cinsiyet	1579 non-null	object
2	Dogum_Tarihi	2357 non-null	datetime64[ns]
3	Uyruk	2357 non-null	object
4	Il	2130 non-null	object
5	Ilac_Adi	2357 non-null	object
6	Ilac_Baslangic_Tarihi	2357 non-null	datetime64[ns]
7	Ilac_Bitis_Tarihi	2357 non-null	datetime64[ns]
8	Yan_Etki	2357 non-null	object
9	Yan_Etki_Bildirim_Tarihi	2357 non-null	datetime64[ns]
10	Alerjilerim	1873 non-null	object
11	Kronik Hastaliklarim	1965 non-null	object
12	Baba Kronik Hastaliklari	2201 non-null	object
13	Anne Kronik Hastaliklari	2140 non-null	object
14	Kiz Kardes Kronik Hastaliklari	2260 non-null	object
15	Erkek Kardes Kronik Hastaliklari	2236 non-null	object
16	Kan Grubu	2010 non-null	object
17	Kilo	2064 non-null	float64
18	Boy	2243 non-null	float64

Şekil 1.1 Veri Bilgisi

2. Eksik Veri Kontrolü

- **Açıklama:** Verideki eksik değerler kontrol edilmiştir. Özellikle **Cinsiyet**, **İl**, **Alerjilerim**, **Kronik Hastalıklarım**, **Kilo**, **Boy** ve **Kan Grubu** sütunlarında eksik veri bulunmuştur. Eksik verilerin oranları ve alınan aksiyonlar aşağıda detaylandırılmıştır:
- **Cinsiyet:** 1579 değer mevcut, 778 eksik.
- **İl:** 2130 değer mevcut, 227 eksik.
- **Alerjilerim:** 1873 değer mevcut, 484 eksik.
- **Kronik Hastalıklarım:** 1965 değer mevcut, 392 eksik.
- **Kan Grubu:** 2010 değer mevcut, 347 eksik.

- **Kilo:** 2064 değer mevcut, 293 eksik.
- **Boy:** 2243 değer mevcut, 114 eksik.

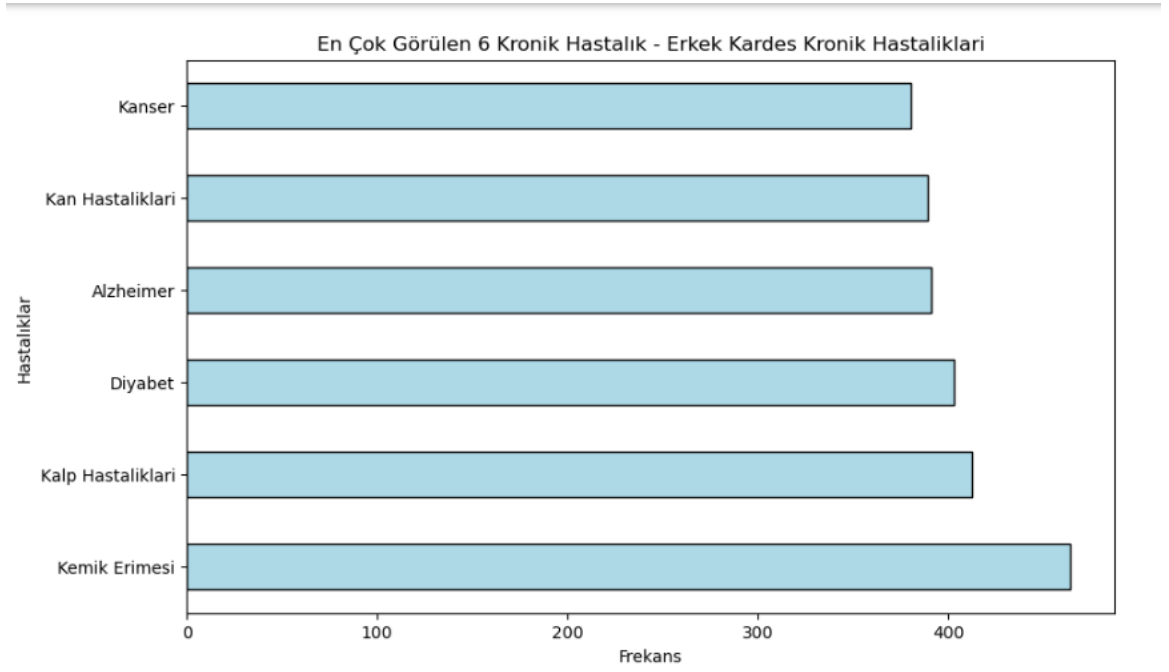
"Eksik veriler SimpleImputer kullanılarak doldurulmuştur. Kategorik sütunlar için en sık görülen değer, sayısal sütunlar için ortalama değer kullanılarak eksik veriler doldurulmuştur.". Örneğin, **boy** ve **kilo** eksik verileri, ortalama değerlerle doldurulmuştur.

"Veri setindeki sayısal veriler, özellikle kilo ve boy gibi değişkenler farklı birimlerde olduğu için StandardScaler ile standartlaştırılmıştır. Bu işlem, sayısal verilerin ortalama 0 ve standart sapma 1 olacak şekilde yeniden ölçeklendirilmesini sağlamaktadır. Bu, makine öğrenimi modellerinin performansını artırır ve öğrenme süreçlerini kolaylaştırır."

3. Veri Analizi

a. En Çok Görülen Kronik Hastalıklar: Kalp, Astım, Alzheimer

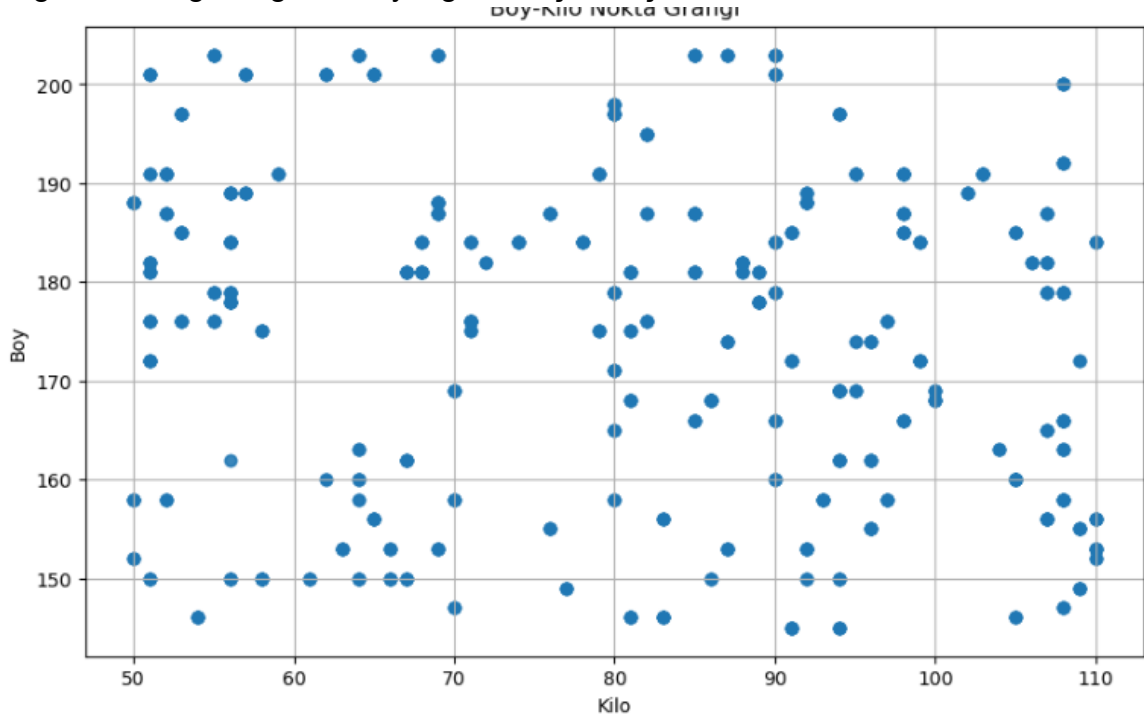
- **Açıklama:** Veride yer alan **Kronik Hastalıklarım** sütunu incelenmiş ve en sık görülen hastalıklar **Kemik Erimesi 1976** , **Alzheimer 1891** ve **Kan Hastalıkları 1812** olarak tespit edilmiştir. Bu hastalıkların frekansları analiz edilerek raporlanmıştır.



Şekil 3.1 Erkek Kardeşte Görülen Kronik Hastalıklar

b. Boy ve Kilo Dağılımları

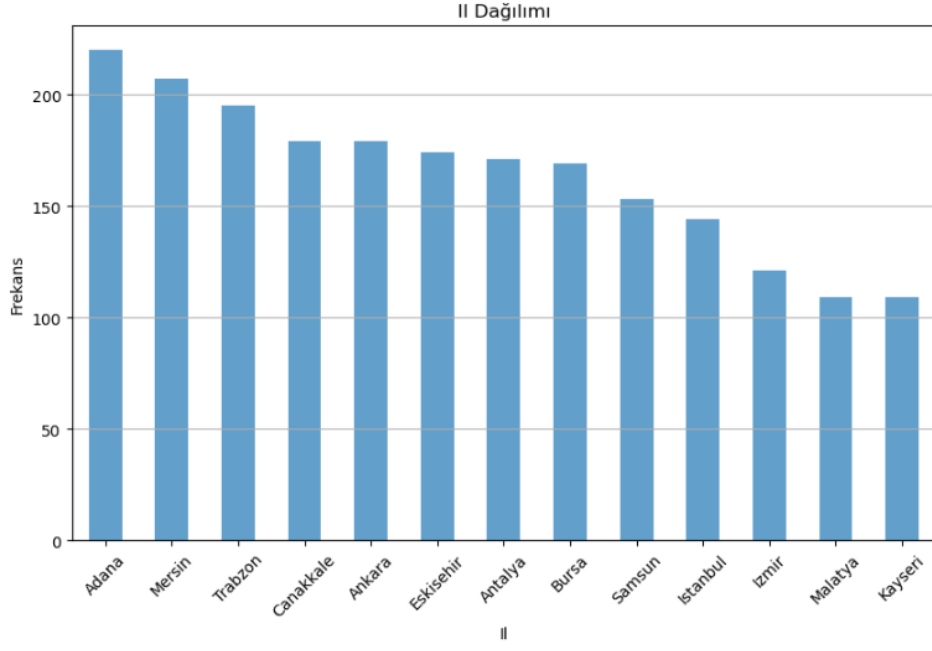
- **Açıklama:** **Boy** ve **Kilo** sütunlarındaki veriler analiz edilmiş ve bu verilerin dağılımı incelenmiştir. Eksik veriler ortalama ile doldurulmuştur ve verilerin dağılımı histogram grafikleriyle görselleştirilmiştir.



Şekil 3.2 Boy ve Kilo Nokta Dağılımları

c. İllerin Dağılımları: Adana, Mersin, Trabzon

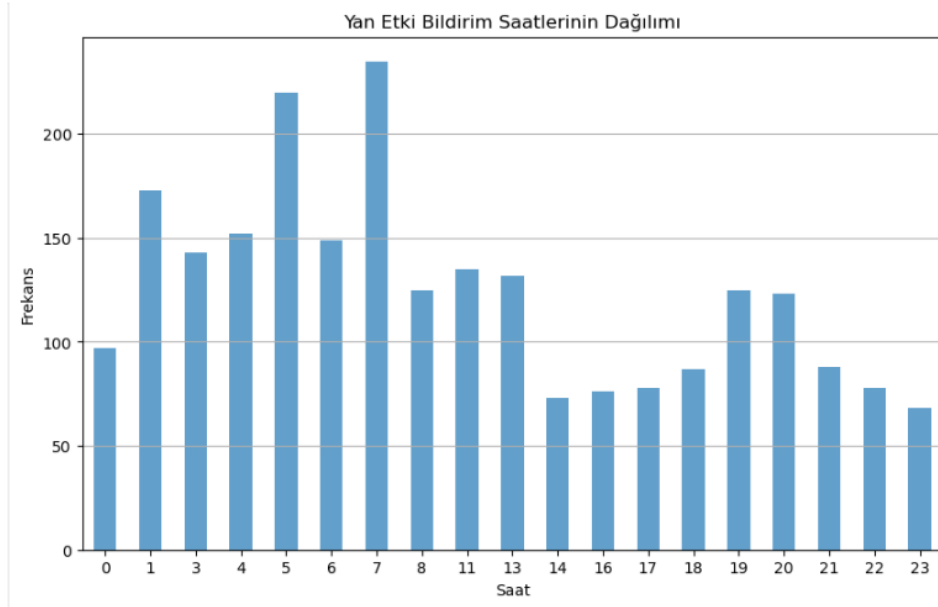
- **Açıklama:** **İl** sütunu üzerinden coğrafi dağılım analiz edilmiştir. Veride en çok görülen iller **Adana**, **Mersin** ve **Trabzon** olarak öne çıkmıştır. Bu iller frekanslarına göre sıralanmıştır.



Şekil 3.3 Yan Etki İl Dağılımları

d. Yan Etki Saatleri: Sabah Kalkış Saatlerinde Artış (5-6-7)

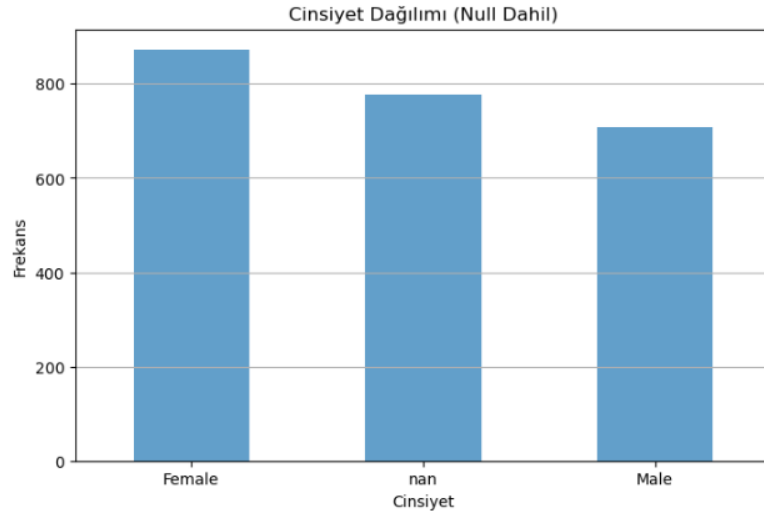
- **Açıklama:** Verideki **Yan Etki Bildirim Tarihi** sütunundan yan etki saatleri elde edilmiştir. Yan etkilerin sabah saatlerinde (özellikle 5-6-7 arası) artış gösterdiği belirlenmiştir. Bu bulgu, yan etki zamanlamasının sabah saatlerinde daha yoğun olduğunu göstermektedir.



Şekil 3.4 Yan Etki Bildirim Saatlerinin Dağılımı

e. Kadın Sayısının Erkek Sayısından Daha Fazla Olduğu

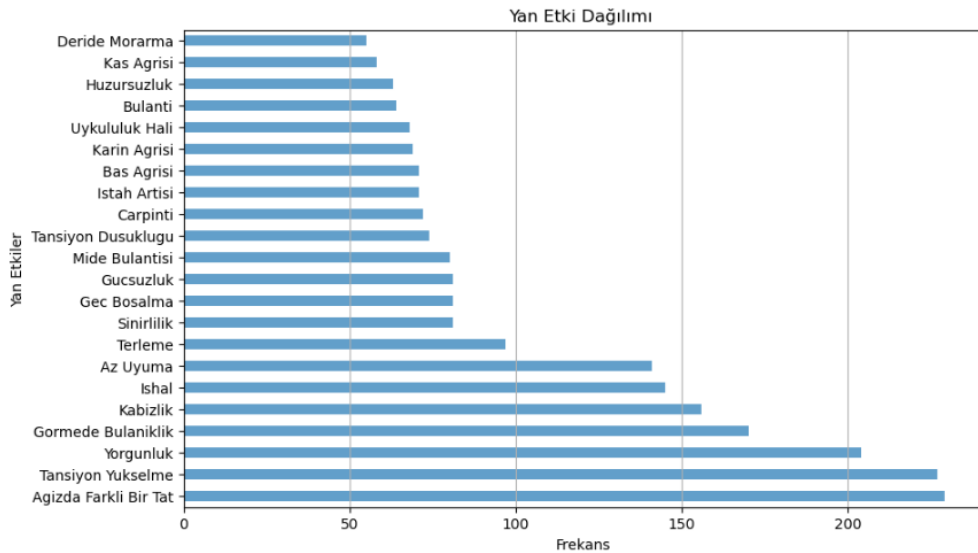
- **Açıklama:** Veride yer alan **Cinsiyet** dağılımı analiz edilmiştir. Kadın sayısının erkek sayısından daha fazla olduğu tespit edilmiştir.



Şekil 3.5 Cinsiyet Dağılımı

f. En Yüksek Görülen Yan Etkiler: Ağızda Farklı Bir Tat, Tansiyon Yükselmesi, Yorgunluk

- **Açıklama:** Veride en sık görülen yan etkiler analiz edilmiştir. En yaygın yan etkiler olarak **Ağızda Farklı Bir Tat 229**, **Tansiyon Yükselmesi 227** ve **Yorgunluk 204** tespit edilmiştir.

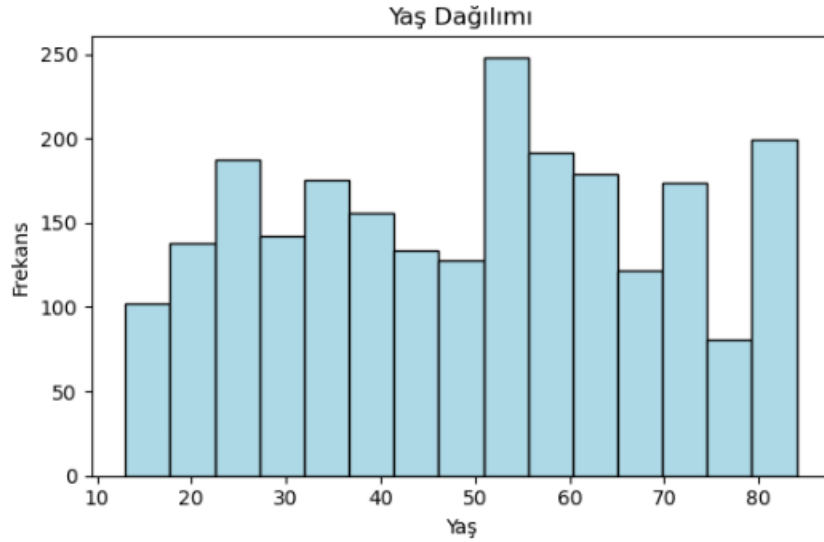


Şekil

3.6 Yan Etki Dağılımları

g. Yaş Dağılımı: En Küçük 13, En Büyük 84, Genel Yoğunluk 50-65 Arasında

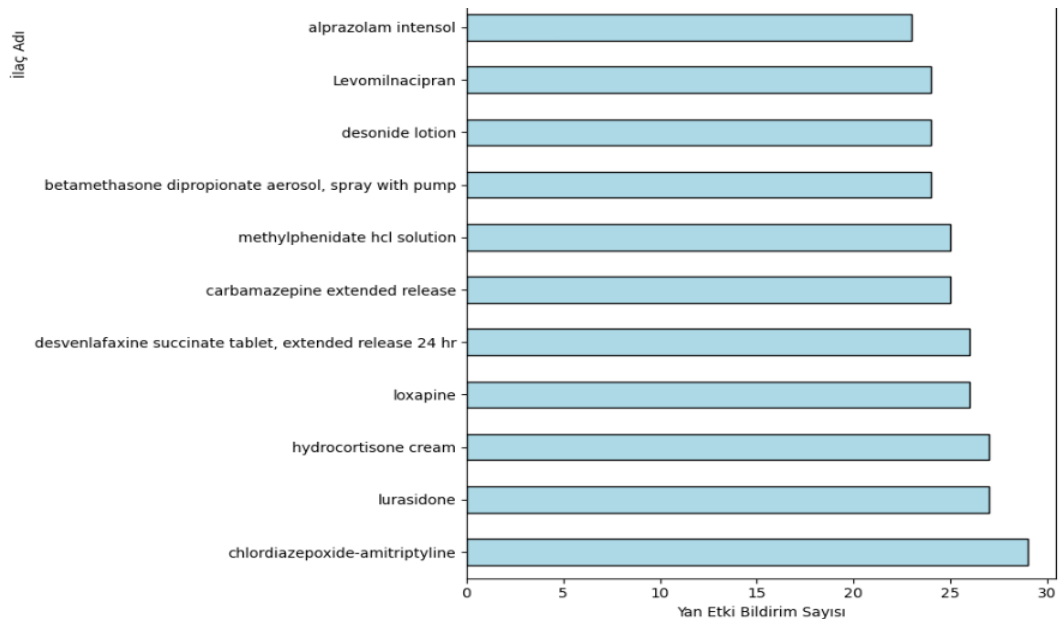
- **Açıklama:** Verideki **yaş** dağılımı analiz edilmiştir. En küçük yaşın 13, en büyük yaşın 84 olduğu ve genel olarak 50-65 yaş aralığında yoğunluk olduğu belirlenmiştir.



Şekil 3.7 Yaş Dağılımları

h. En Çok Etkiye Sahip İlaçlar: Hydrocortisone Cream, Lurasidone, Chlordiazepoxide-Amitriptyline

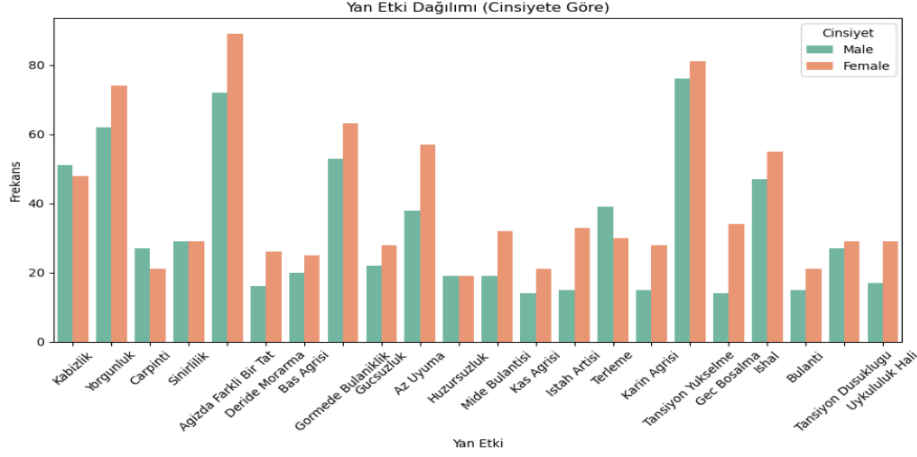
- **Açıklama:** Veride yer alan **ilaç Adı** sütunu analiz edilmiştir. En fazla yan etkiye sahip ilaçlar **Hydrocortisone Cream 27**, **Lurasidone 27** ve **Chlordiazepoxide-Amitriptyline 29** olarak belirlenmiştir



Şekil 3.8 Yan Etki Bildirim Sayısı

i. Kadın ve Erkeklerde Görülen Yan Etkiler Dengeli Dağılımda

- **Açıklama:** Kadın ve erkeklerde görülen yan etkiler incelenmiş ve bu etkilerin dengeli bir şekilde dağıldığı tespit edilmiştir.



Şekil 3.9 Yan Etkilerin Cinsiyete Göre Dağılımı

4. Veri Kolon Ekleme İşlemleri

a. Kronik Hastalıklar Sayısallaştırıldı (LabelEncoder ile)

- **Açıklama:** Kronik Hastalıklarım sütunu, LabelEncoder() kullanılarak sayısal değerlere dönüştürülmüştür.

b. Yaş Hesaplandı ve Kolon Olarak Eklendi

- **Açıklama:** Doğum Tarihi sütunundan yaş hesaplanarak yeni bir kolon olarak eklenmiştir.

c. Yan Etki Saatleri Hesaplandı ve Kolon Olarak Eklendi

- **Açıklama:** Yan Etki Bildirim Tarihi sütunundan saat bilgisi çıkarılarak yan etki saatleri hesaplanmış ve yeni bir kolon olarak eklenmiştir.

d. Boy ve Kilo ile Vücut Kitle İndeksi Hesaplandı

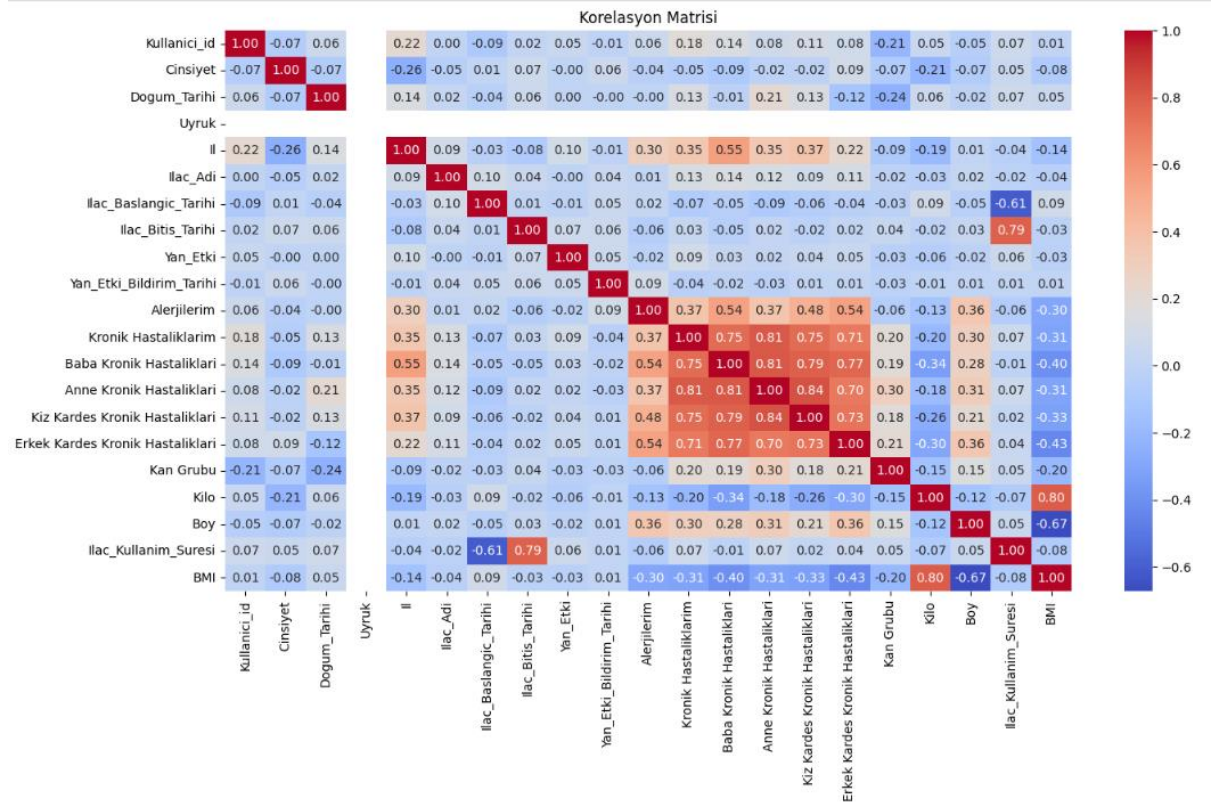
- **Açıklama:** Boy ve kilo verilerinden **Vücut Kitle İndeksi (VKİ)** hesaplanmış ve yeni bir kolon olarak eklenmiştir.

e. Eksik Boy ve Kilo Verileri Ortalama ile Dolduruldu

- **Açıklama:** Eksik **boy** ve **kilo** verileri ortalama değerlerle doldurulmuştur.

e. korelasyon Analizi

- **Açıklama:** Verideki kolonların hangi kolonlar ile ilişkili olduğu ve buna bağlı yan etkide neler değişebileceğini anlamak için korelasyon analizi yapılmıştır.



Şekil 4.1 Korelasyon Analizi

5. Etki Etmeyen Kolonların Kaldırılması

- **Açıklama:** Analize katkı sağlamayacak olan **Uyruk, İlaç Başlangıç Tarihi, İlaç Bitiş Tarihi, Yan Etki Bildirim Tarihi** ve **Doğum Tarihi** gibi kolonlar veri setinden çıkarılmıştır.