

İlaç Yan Etkileri Üzerine Rapor

1. İlaç ve Yan Etki İlişkisi Analizi

Bu çalışmada, ilaçlar ve onların yan etkileri arasındaki ilişkileri inceleyen bir veri seti kullanılarak bir analiz yapılmıştır. Veriler, kullanıcıların farklı ilaçları kullandıktan sonra yaşadıkları yan etkilerle ilgilidir. Bu analizde amaç, hangi ilaçların hangi yan etkilerle daha sık ilişkilendirildiğini bulmak ve elde edilen bilgilerin ilaç-yan etki ilişkilerinde anlamlı desenler olup olmadığını belirlemektir. Veri setinde yer alan değişkenler; ilaç bilgileri, kullanıcıların demografik özellikleri, kronik hastalık bilgileri ve yan etkiler ile ilgilidir.

Bu raporda, veri setindeki eksiklikler, kategorik ve sayısal değişkenlerin dağılımları, korelasyonlar, veri işleme yöntemleri ve sonuçlar detaylı olarak ele alınacaktır.

2. Veri Seti Bilgileri

Veri seti, **2357** gözlem ve **19** değişken içermektedir. Bu değişkenler arasında ilaç isimleri, kullanıcı bilgileri, yan etkiler ve kronik hastalıklar yer almaktadır. Veri setinin boyutu yaklaşık **350 KB** olup, ortalama bir kayıt boyutu **152.1** byte olarak hesaplanmıştır. Aşağıda veri setinin temel özellikleri detaylandırılmıştır.

	Kullanici_id	Cinsiyet	Dogum_Tarihi	Uyruk	İl	Ilac_Adi	Ilac_Baslangic_Tarihi	Ilac_Bitis_Tarihi	Yan_Etki
0	107	Male	1960-03-01	Türkiye	Canakkale	trifluoperazine	2022-01-09	2022-03-04	Kabizlik
1	140	Male	1939-10-12	Türkiye	Trabzon	fluphenazine hcl	2022-01-09	2022-03-08	Yorgunluk
2	2	Female	1976-12-17	Türkiye	Canakkale	warfarin sodium	2022-01-11	2022-03-12	Carpinti
3	83	Male	1977-06-17	Türkiye	Adana	valproic acid	2022-01-04	2022-03-12	Sinirlilik
4	7	Female	1976-09-03	Türkiye	Izmir	carbamazepine extended release	2022-01-13	2022-03-06	Agizda Farkli Bir Tat
5	131	Male	1982-01-05	Türkiye	NaN	fluoxetine dr	2022-01-07	2022-03-19	Agizda Farkli Bir Tat
6	61	NaN	1997-01-10	Türkiye	Mersin	carbamazepine extended release	2022-01-05	2022-03-07	Agizda Farkli Bir Tat
7	95	NaN	1997-01-15	Türkiye	Antalya	thiothixene	2022-01-10	2022-03-07	Deride Morarma
8	3	Female	1973-08-05	Türkiye	NaN	asenapine tablet, sublingual	2022-01-08	2022-03-17	Bas Agrisi
9	53	NaN	1941-10-16	Türkiye	Adana	fluphenazine hcl elixir	2022-01-10	2022-03-13	Gormede Bulaniklik

2.1. Değişkenler

Veri setindeki temel değişkenler şunlardır:

- **Kullanıcı ID:** Her bir kullanıcıya özgü bir kimlik numarası.
- **Cinsiyet:** Kullanıcının cinsiyeti.
- **Doğum Tarihi:** Kullanıcının doğum tarihi.
- **Uyruk:** Kullanıcının milliyeti.
- **İl:** Kullanıcının yaşadığı şehir.
- **İlaç Adı:** Kullanıcının kullandığı ilaç adı.
- **İlaç Başlangıç Tarihi:** İlaç kullanımının başladığı tarih.
- **İlaç Bitiş Tarihi:** İlaç kullanımının sona erdiği tarih.
- **Yan Etki:** Kullanıcının yaşadığı yan etkiler.
- **Yan Etki Bildirim Tarihi:** Yan etkinin raporlandığı tarih.
- **Alerjilerim:** Kullanıcının bildirdiği alerjiler.
- **Kronik Hastalıklarım:** Kullanıcının sahip olduğu kronik hastalıklar.
- **Baba Kronik Hastalıkları:** Kullanıcının babasının sahip olduğu kronik hastalıklar.
- **Anne Kronik Hastalıkları:** Kullanıcının annesinin sahip olduğu kronik hastalıklar.
- **Kız Kardeş Kronik Hastalıkları:** Kullanıcının kız kardeşinin sahip olduğu kronik hastalıklar.
- **Erkek Kardeş Kronik Hastalıkları:** Kullanıcının erkek kardeşinin sahip olduğu kronik hastalıklar.
- **Kan Grubu:** Kullanıcının kan grubu.
- **Kilo:** Kullanıcının ağırlığı (kg).
- **Boy:** Kullanıcının boyu (cm).

Her bir kullanıcının kimlik numarası ve demografik bilgileri (cinsiyet, doğum tarihi, il, uyruk) gibi veriler, sağlık durumlarını anlamak ve olası tedavi planlarını kişiselleştirmek için kullanılabilir. Özellikle ilaç adı, kullanım başlangıç ve bitiş tarihleri gibi bilgiler, spesifik ilaçların belirli hasta gruplarında nasıl yan etkiler yarattığını analiz etmek açısından değerlidir.

3. Eksik veriler

Veri setinde belirli eksiklikler bulunmaktadır. Bu eksikliklerin oranları ve adetleri şu şekildedir:

Matrix Eksik Veriler

Sütun	Eksik Veri (%)	Eksik Veri (Adet)
Cinsiyet	%33.0	778
İl	%9.6	227
Alerjilerim	%20.5	484
Kronik Hastalıklarım	%16.6	392
Baba Kronik Hastalıkları	%6.6	156
Anne Kronik Hastalıkları	%9.2	217
Kız Kardeş Kronik Hastalıkları	%4.1	97
Erkek Kardeş Kronik Hastalıkları	%5.1	121
Kan Grubu	%14.7	347
Kilo	%12.4	293
Boy	%4.8	114

Yüzdelik Eksik Veriler

4. Veri İşleme Adımları

4.1 Eksik Verilerin Doldurulması

Eksik veriler, sayısal veriler için ortalama değerleriyle, kategorik veriler için ise "bilinmiyor" değeri ile doldurulmuştur. Bu işlem, modelin performansını olumsuz etkileyebilecek eksik veri sorununu çözmeye yönelik önemli bir adımdır.

- Sayısal Değişkenler:** Boy, kilo gibi sayısal değişkenlerde eksik değerler ortalama ile dolduruldu. Ortalama ağırlık 80,86 kg, ağırlık sütunundaki minimum değer 50 kg ve maksimum değer değeri 110 kg'dır. Ortalama boy 174,64 cm, minimum değer 145 cm ve maksimum değer 203 cm'dir. Boy dağılımı 160 cm ile 187 cm arasında yoğunlaşmaktadır.

- **Kategorik Değişkenler:** Cinsiyet, il, alerjiler ve kronik hastalıklar gibi kategorik değişkenlerde eksik veriler "bilinmiyor" olarak işaretlendi.

4.2 Kategorik Verilerin Kodlanması

Kategorik veriler, makine öğrenmesi modelleri için sayısal verilere dönüştürülmek zorundadır. Bu çalışmada, kategorik veriler Label Encoding yöntemiyle kodlanmıştır. Bu yöntemde, her bir kategorik değer, sayısal bir değere karşılık gelecek şekilde yeniden kodlanır. Örneğin, belirli bir kategorideki benzersiz değerler 0, 1, 2 gibi sayılarla temsil edilir.

4.3 Sayısal Verilerin Standardizasyonu

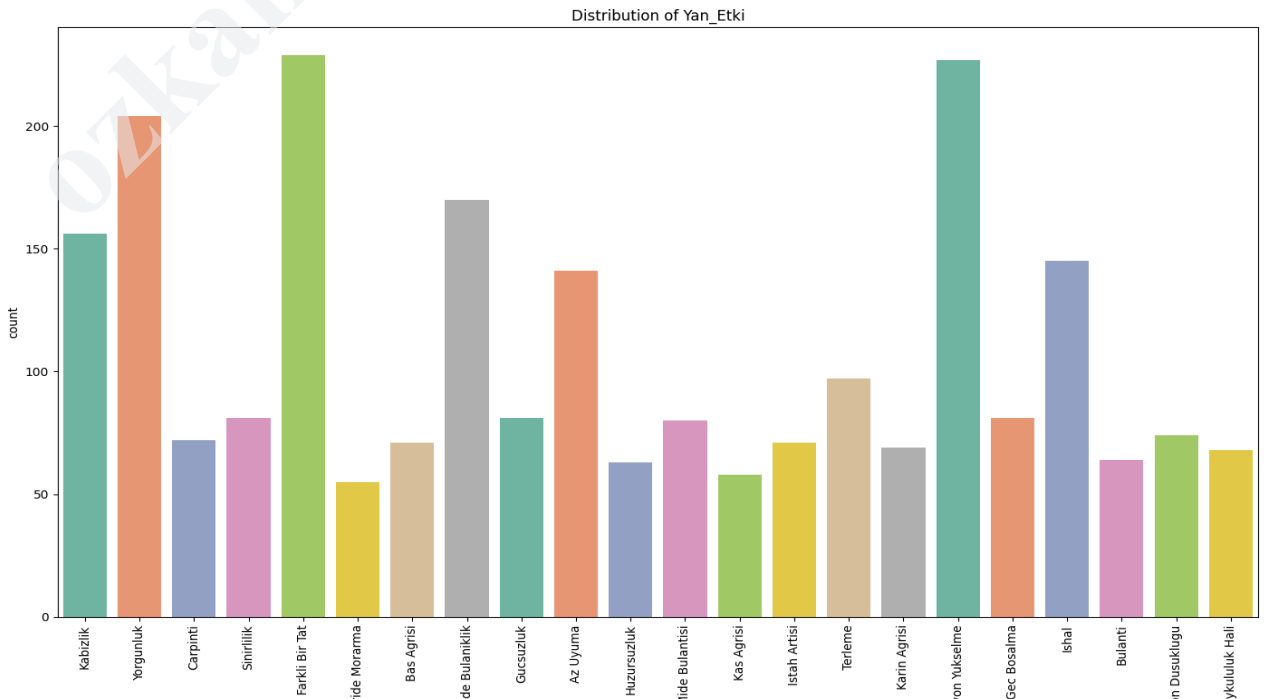
Sayısal değişkenler (boy, kilo gibi) farklı birimlere sahip olduğundan, modelin performansını artırmak için bu veriler standardize edilmiştir. **Standardizasyon**, verilerin ortalaması 0 ve standart sapması 1 olacak şekilde yeniden ölçeklendirilmesi işlemidir. Bu işlem, özellikle mesafe tabanlı algoritmalarda (örneğin, K-Nearest Neighbors) modelin performansını artırır.

5. Analizi

Verilerin yapısı ve dağılımı, görselleştirme teknikleri (histogram, scatter plot, vb.) kullanılarak detaylı bir şekilde incelenmiş ve böylece daha iyi anlaşılmıştır.

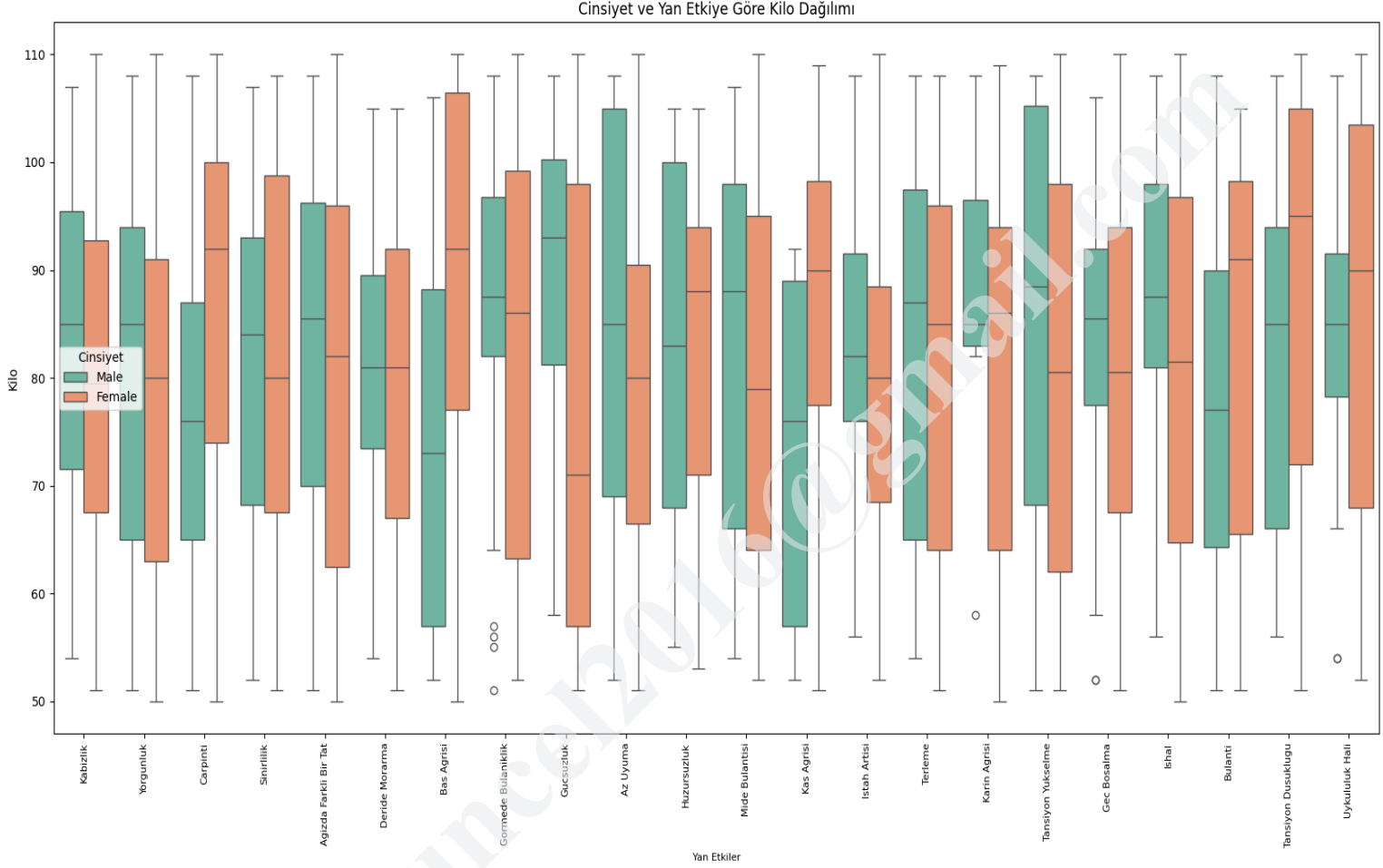
5.1 Genel Yan Etkileri Analizi

İlacı kullanan hastaların büyük bir kısmı, tedavi sürecinde özellikle ağızda farklı bir tat ve yüksek tansiyon gibi yan etkiler yaşadıklarını belirtmişlerdir. Bu durum, ilacın bazı kişilerde bu belirtilere neden olabileceğini göstermektedir.



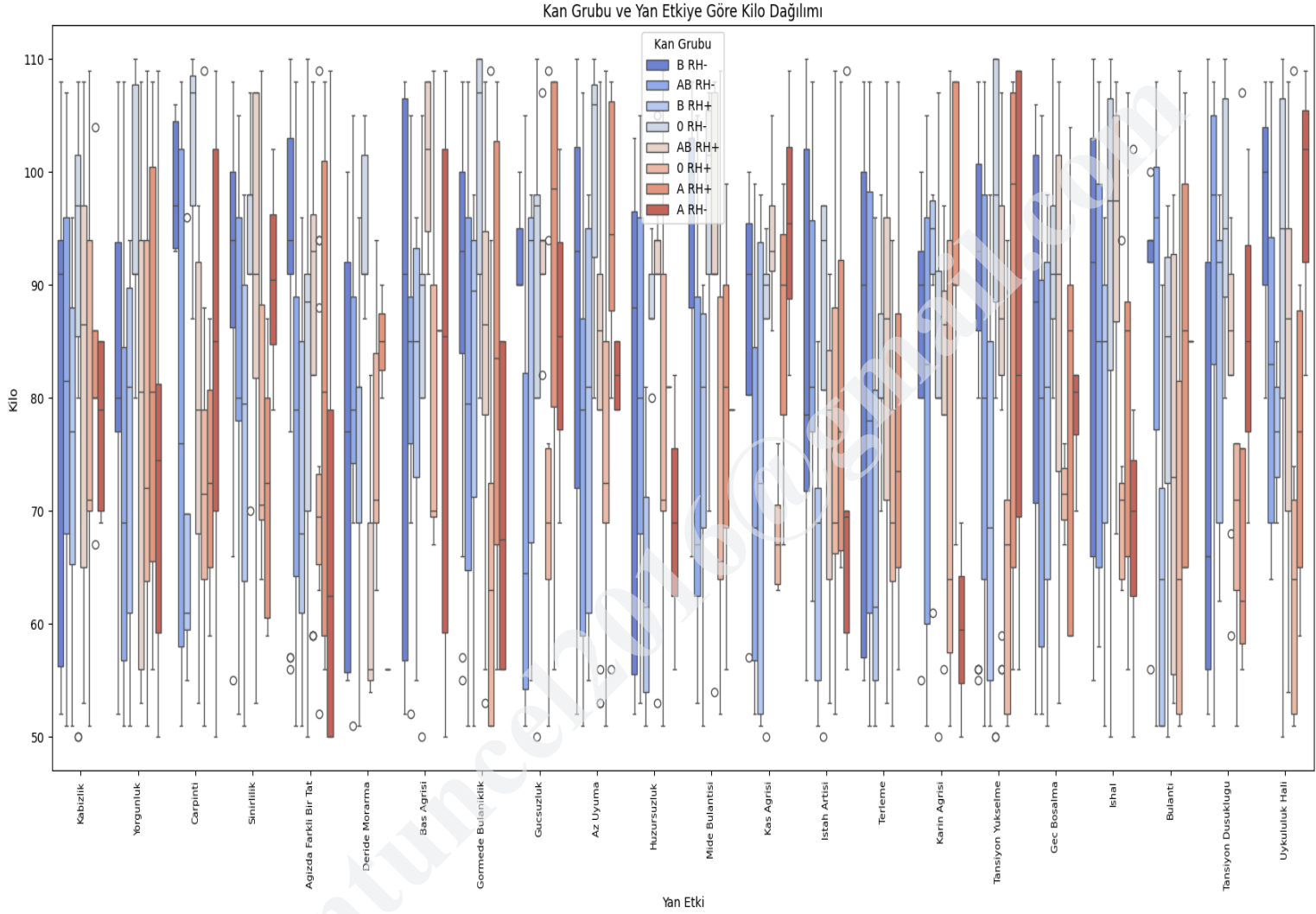
5.2 Cinsiyet ve Yan Etkiye Göre Kilo Analizi

Genel olarak, kadınların ve erkeklerin kilo aralıkları benzer görünmektedir, ancak bazı yan etkilerde cinsiyetler arasında farklılıklar göze çarpmaktadır. Örneğin, bazı yan etkilerde erkeklerin ortalama kilosu daha yüksekken, diğerlerinde kadınların ortalama kilosu daha yüksektir.



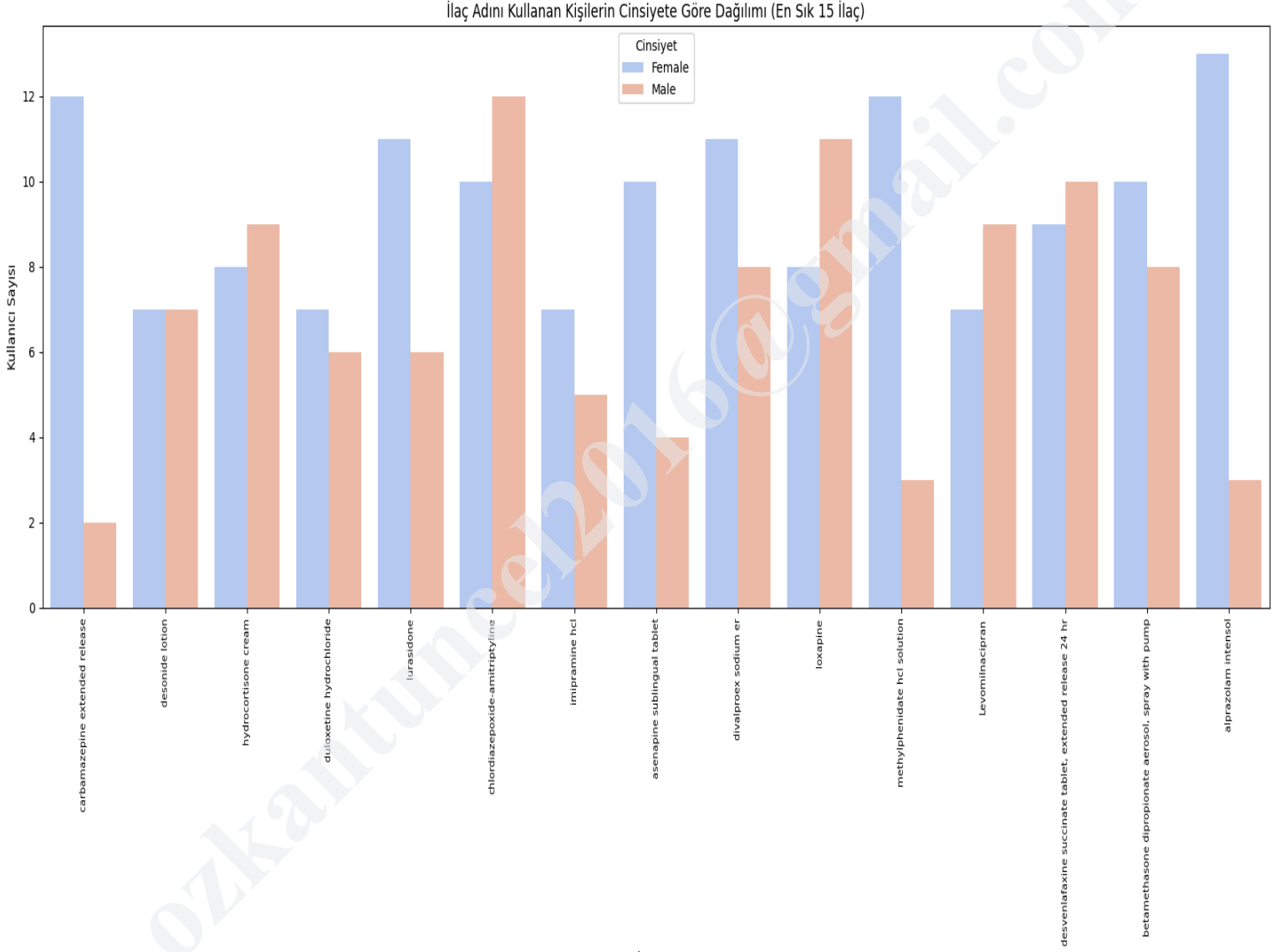
5.3 Kan Grubu ve Yan Etkiye Gore Kilo Analizi

Grafik, bazı yan etkilerde belirli kan gruplarının diğerlerinden daha yüksek veya düşük kilo aralıklarına sahip olabileceğini göstermektedir.



5.4 İlaçların Kullanan Kişilerin Cinsiyete Göre Analizi

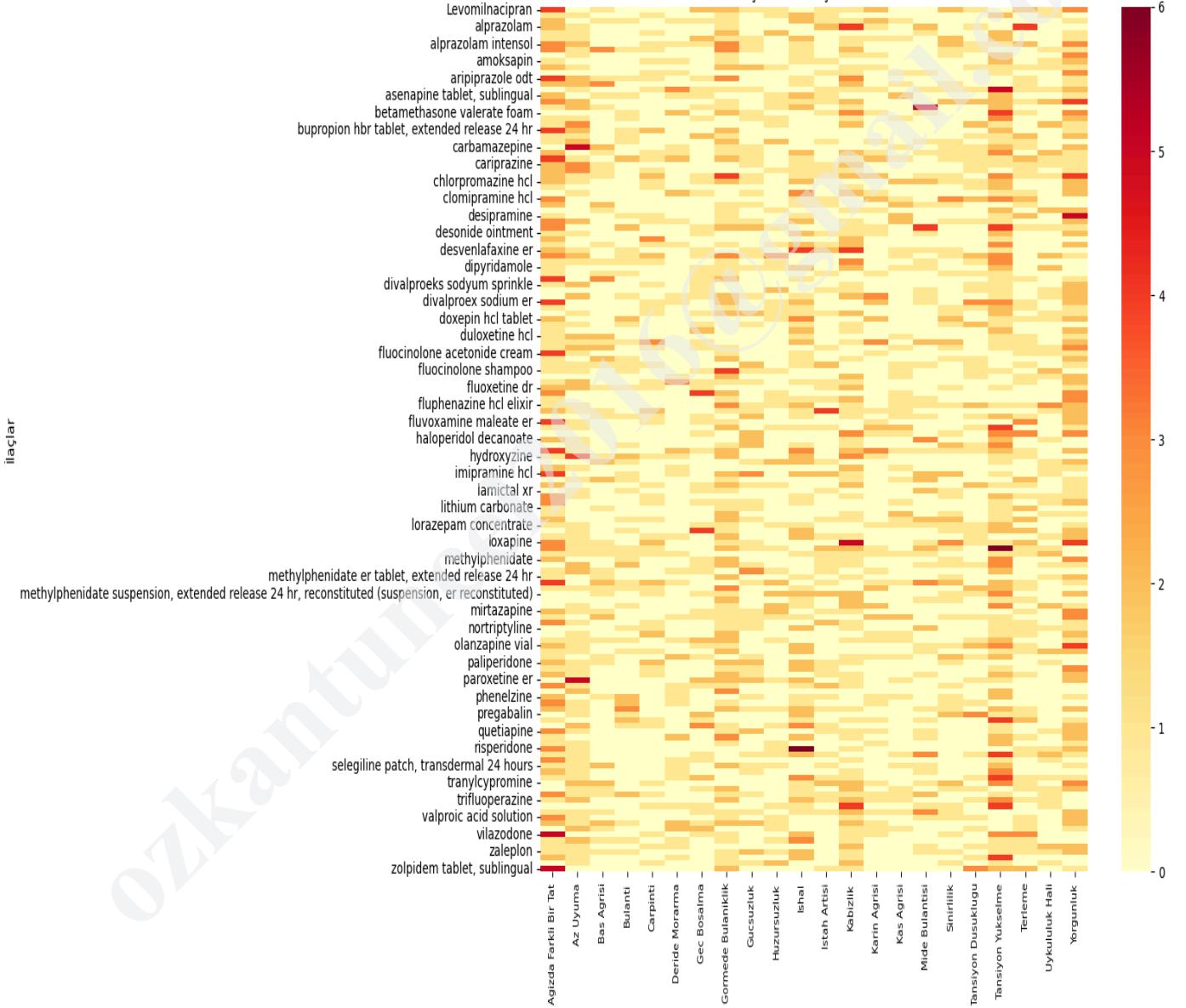
Bazı ilaçlarda belirgin cinsiyet farklılıkları gözlenmektedir. Örneğin, "ceftriaxone, extended release" ve "alprazolam intensol" kadınlar tarafından daha fazla kullanılırken, "chlordiazepoxide-amitriptyline" erkekler tarafından daha çok tercih edilmektedir. Diğer ilaçlarda ise kullanım oranları cinsiyetler arasında daha dengeli görünmektedir



5.5 İlaçlar ve Yan Etkileri İlişkisi Analizi

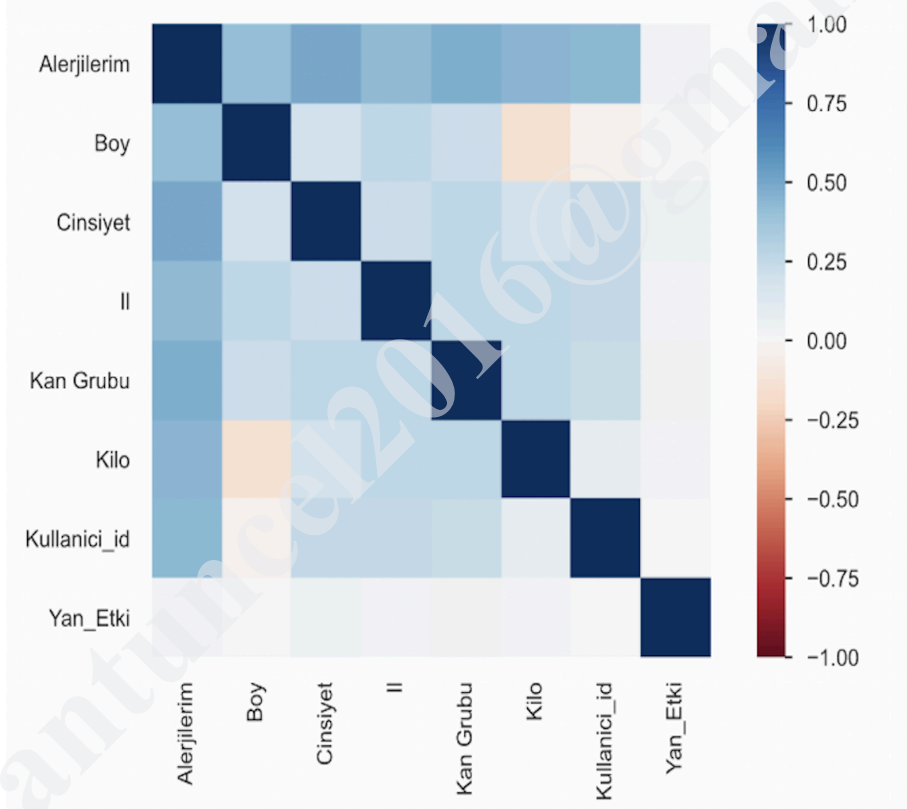
Isı haritasında gösterildiği üzere, ilaçların yan etkileri genel olarak dengeli bir dağılım göstermektedir. Ancak, bazı ilaçların kişiler üzerinde daha fazla yan etki gösterdiği gözlemlenmiştir. Loxapine bu duruma örnek olarak verilebilir. DrugBank veritabanında belirtildiği gibi, Loxapine bir antipsikotik sakinleştiricidir. Bu ilacın etki mekanizması nedeniyle, kullanıcılarda hipotansiyon (düşük tansiyon) gibi yan etkilere neden olabilmektedir. Bu tür bulgular, ilaçların etkileri ve yan etkileri hakkında daha detaylı araştırmaların önemini vurgulamaktadır.

İlaç-Yan Etki İlişki Haritası



5.5 Korelasyon Analizi

Alerjiler ile diğer değişkenler arasında orta düzeyde pozitif korelasyonlar gözlemlenirken, kilo ile boy arasındaki ilişki zayıf negatif bir yönde seyretmektedir. Yan etkiler ile diğer değişkenler arasındaki düşük korelasyonlar, bu değişkenlerin güçlü bir bağa sahip olmadığını ortaya koymaktadır. İlginç bir bulgu olarak, bazı alerjilerin bölgesel özellikler taşıdığı söylenebilir. Ayrıca, cinsiyet ile yan etkiler arasında belirgin bir ilişki bulunması, bu faktörlerin etkisini vurgulamaktadır. Bu bulgular, değişkenler arasındaki dinamikleri anlamak ve potansiyel örüntüleri belirlemek açısından önemli ipuçları sunmaktadır.



	Alerjilerim	Boy	Cinsiyet	II	Kan Grubu	Kilo	Kullanici_id	Yan_Etki
Alerjilerim	1.000	0.410	0.498	0.423	0.476	0.442	0.436	0.021
Boy	0.410	1.000	0.199	0.261	0.220	-0.151	-0.026	0.000
Cinsiyet	0.498	0.199	1.000	0.223	0.267	0.194	0.251	0.050
II	0.423	0.261	0.223	1.000	0.267	0.258	0.251	0.022
Kan Grubu	0.476	0.220	0.267	0.267	1.000	0.267	0.239	0.038
Kilo	0.442	-0.151	0.194	0.258	0.267	1.000	0.071	0.019
Kullanici_id	0.436	-0.026	0.251	0.251	0.239	0.071	1.000	0.000
Yan_Etki	0.021	0.000	0.050	0.022	0.038	0.019	0.000	1.000