

Çoklu Veri Seti Dağılımları

Emre Söylemez

17.11.2022

ÖZET

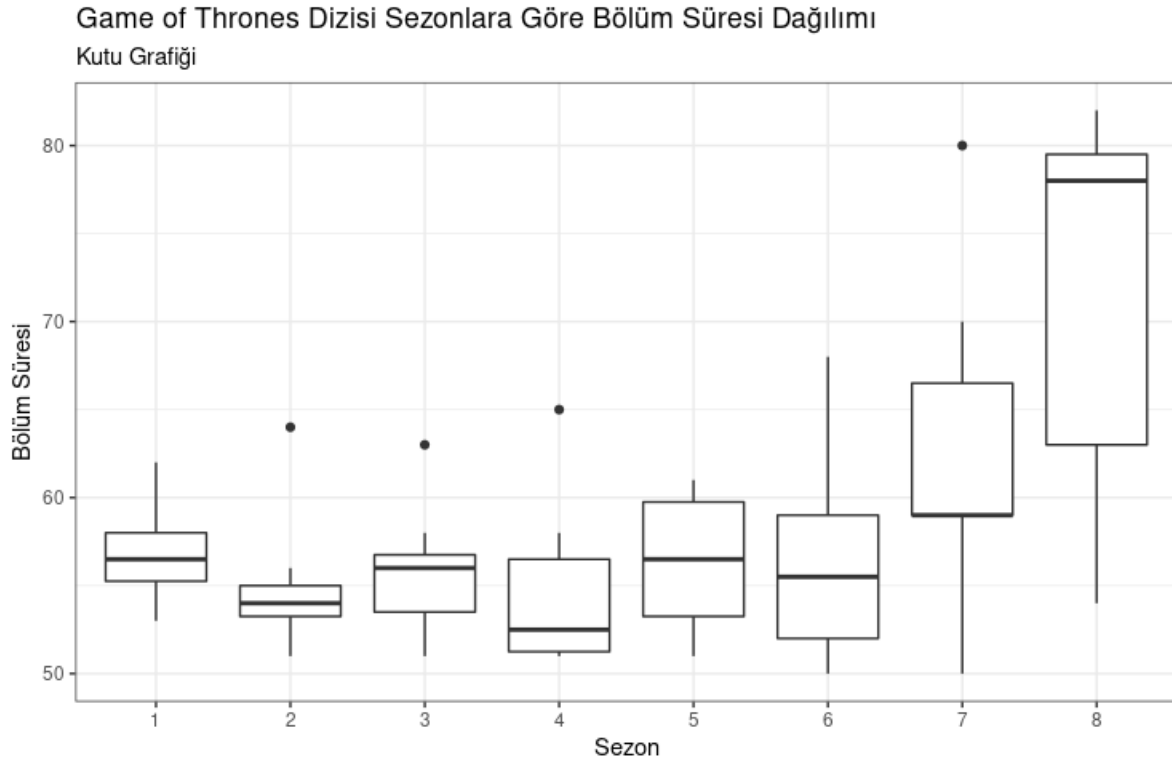
Bu raporda önceden belirlenmiş olan 3 veri setine ait veri görselleştirme çalışmaları yer almaktadır. Öncelikle veri setleri tanıtılmış, sonrasında ise bu veri setleri ile alakalı olan araştırma sorularına ilişkin her biri ayrı bölümde yer alacak şekilde grafikler oluşturulmuş ve yorumlanmıştır.

1.Game of Thrones Dizisi Veri Seti

Bu veri setinde Game of Thrones dizisine ait birçok veri bulunmaktadır. 18 farklı değişken ve bu değişkenler için 73 farklı gözlem yer almaktadır. Bu raporda incelenecek olan değişkenler ise şunlardır:

- Sezon
- Bölüm süresi
- Reyting
- Yazar

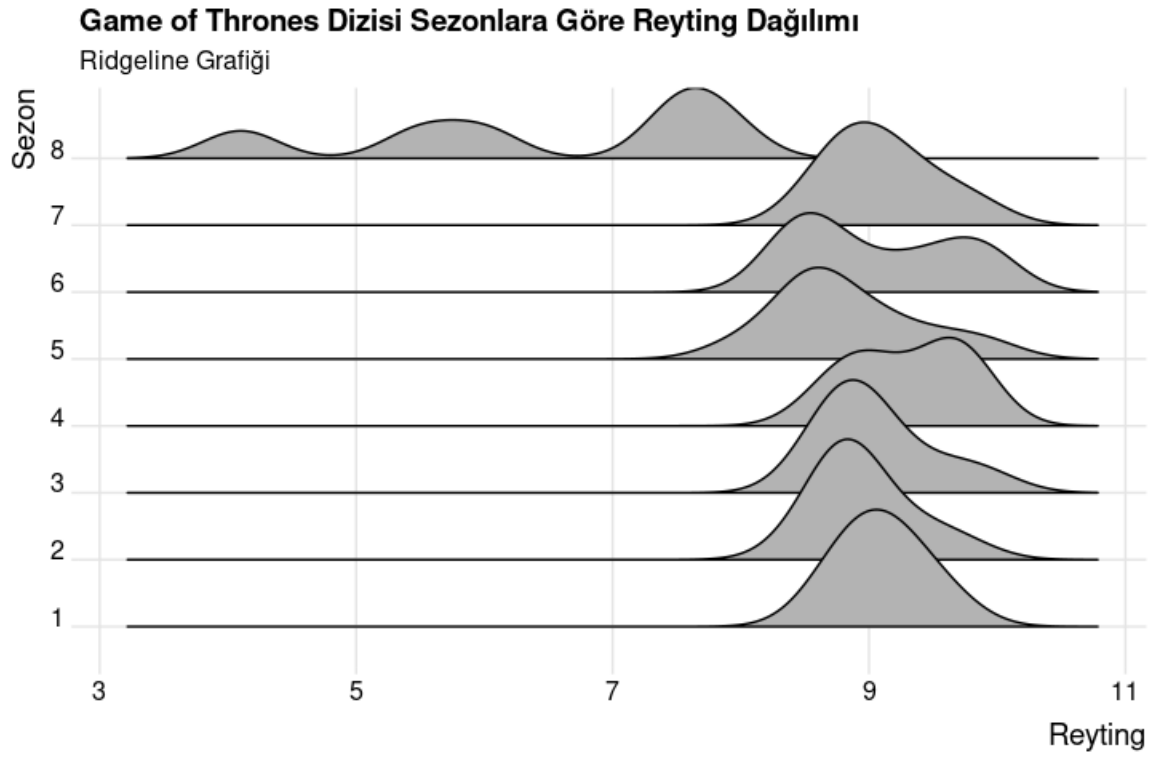
1.1.Sezonlara göre bölüm süresi dağılımlarını inceleyiniz. Bölüm sürelerindeki en yüksek değişimin gözleendiği sezonu belirleyiniz.



Yukarıda sezonlara göre bölüm sürelerinin dağılımını içeren kutu grafiği kullanılmıştır. 2. Sezondan itibaren bölüm süreleri değişimi ile alakalı artış trendi görülmekte olup **en yüksek değişimin** 8. Sezonda olduğu görülmektedir.

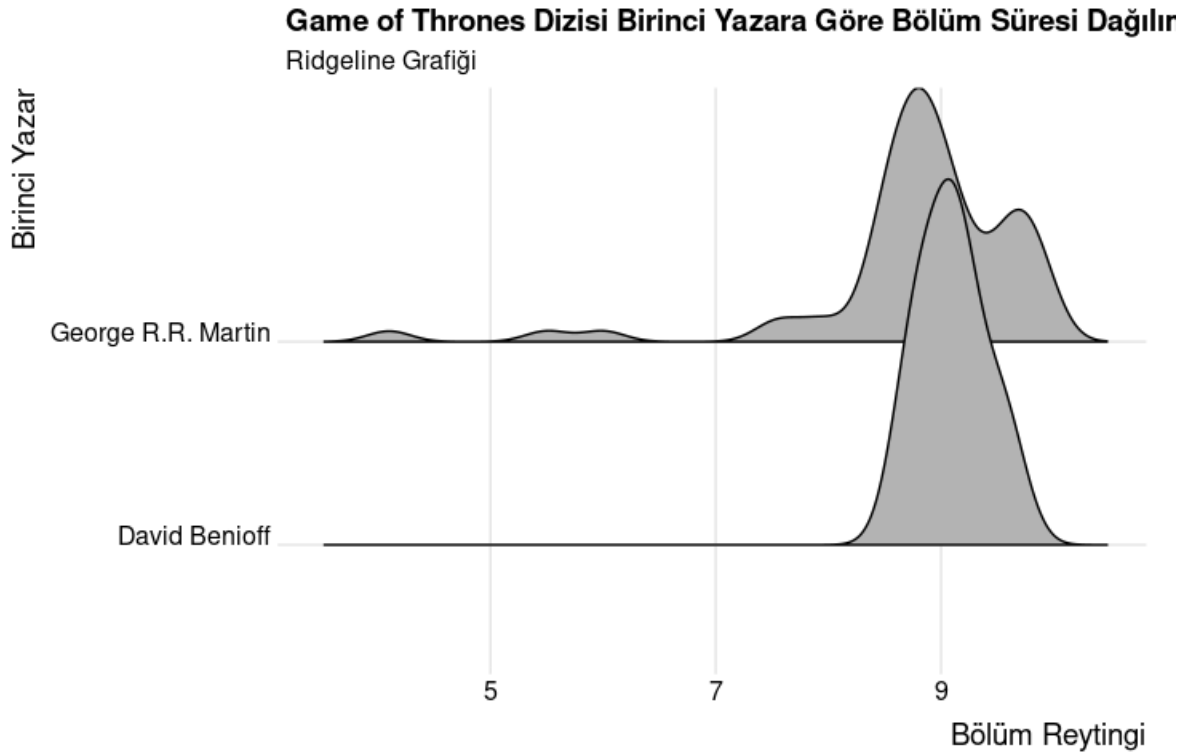
Not: Son 2 sezonda bölüm sayısındaki azalmanın bu değişimdeki artışa etkisini göz ardı etmemek gerekir.

1.2.Sezonlara göre rating dağılımlarını araştırınız. Rating değişiminin en düşük olduğu sezonu belirleyiniz.



Yukarıda sezonlara göre bölüm reytinglerinin dağılımını içeren ridgeline grafiği kullanılmıştır. En yüksek değişiminin 8.sezonda olduğunu görüldüğü grafikte 1-4 arası sezonların ve 5-7 arası sezonların reytinglerinin benzer değişim aralıklarında seyrettiği görülmektedir. **En düşük değişim** ise 1. Sezonda gerçekleşmiştir.

1.3.Birinci yazara göre rating dağılımını araştırınız. Hangi yazarın senaryosunun daha yüksek rating aldığını belirtiniz.



Yukarıda 1. Yazara göre bölüm reytinglerinin dağılımını içeren ridgeline grafiği kullanılmıştır. David Benioff adlı yazarın yazdığı bölüm reytinglerinin daha dar bir aralıkta değişim gösterdiği görülmektedir. George R.R. Martin adlı yazarın yazdığı bölüm reytinglerinin her ne kadar çok daha geniş aralıkta değişim gösterdiği görülse de **daha yüksek reytingleri** alan bölümler kendisine aittir.

Not: Bu veri setindeki Birinci yazar değişkenine ait gözlem değerlerinde 1. Sezona ait bölümler David Benioff, sonraki sezonlara ait bölümler ise George R.R. Martin adlı yazara ait olduğu belirtilse de buradaki gözlem değerleri yanlış veri içermektedir.

George R.R. Martin adlı yazarın ilk 4 sezonda her sezon birer bölüm olmak üzere yazdığı bölümler şunlardır:

- The Pointy End
- Blackwater

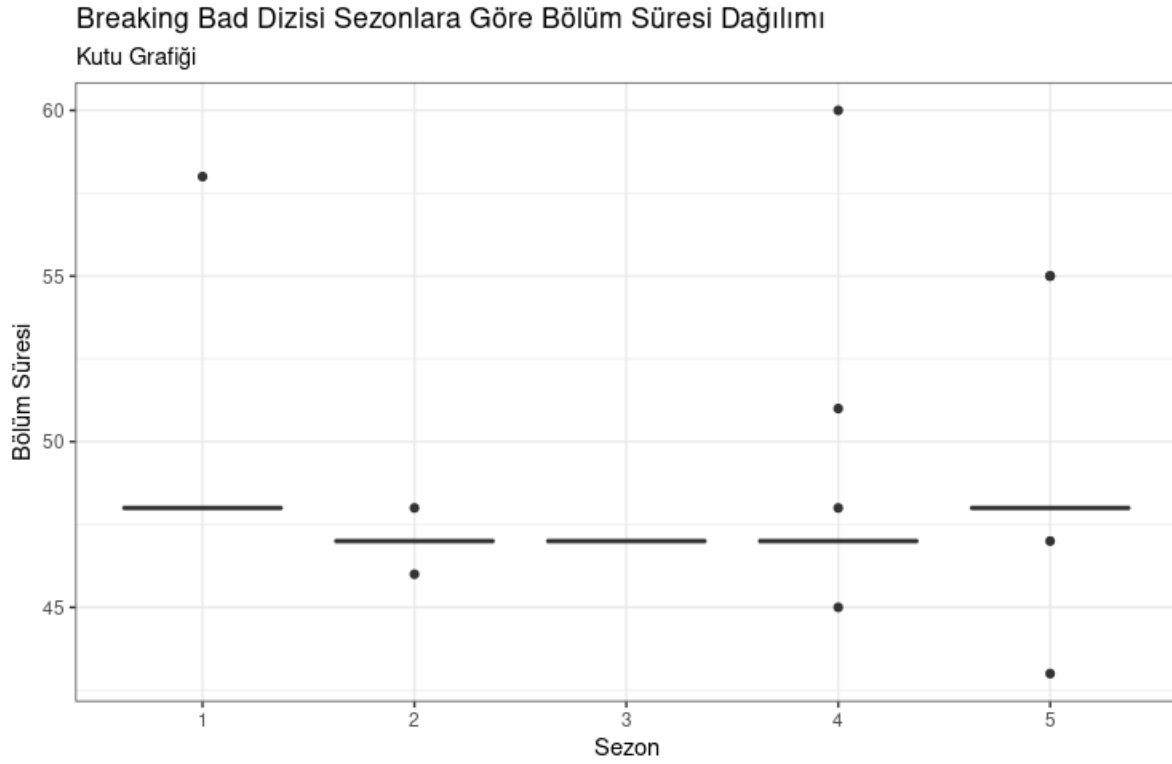
- The Bear and the Maiden Fair
- The Lion and the Rose

2.Breaking Bad Dizisi Veri Seti

Bu veri setinde Breaking Bad dizisine ait birçok veri bulunmaktadır. 10 değişken ve bu değişkenler için 62 farklı gözlem yer almaktadır. Bu raporda incelenecek olan değişkenler ise şunlardır:

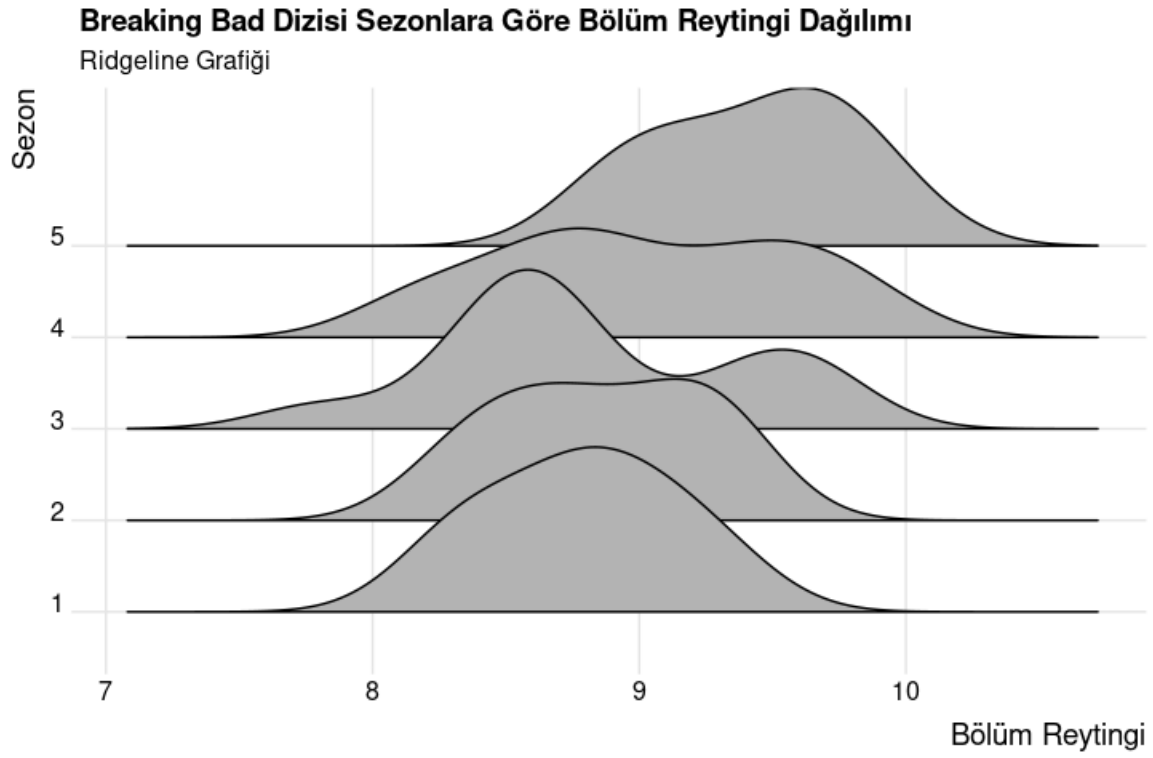
- Sezon
- Bölüm süresi
- Reyting
- İzlenme sayısı(milyon)

2.1.Sezonlara göre bölüm süresi dağılımlarını inceleyiniz. Bölüm sürelerindeki en yüksek değişimin gözlemlendiği sezonu belirleyiniz.



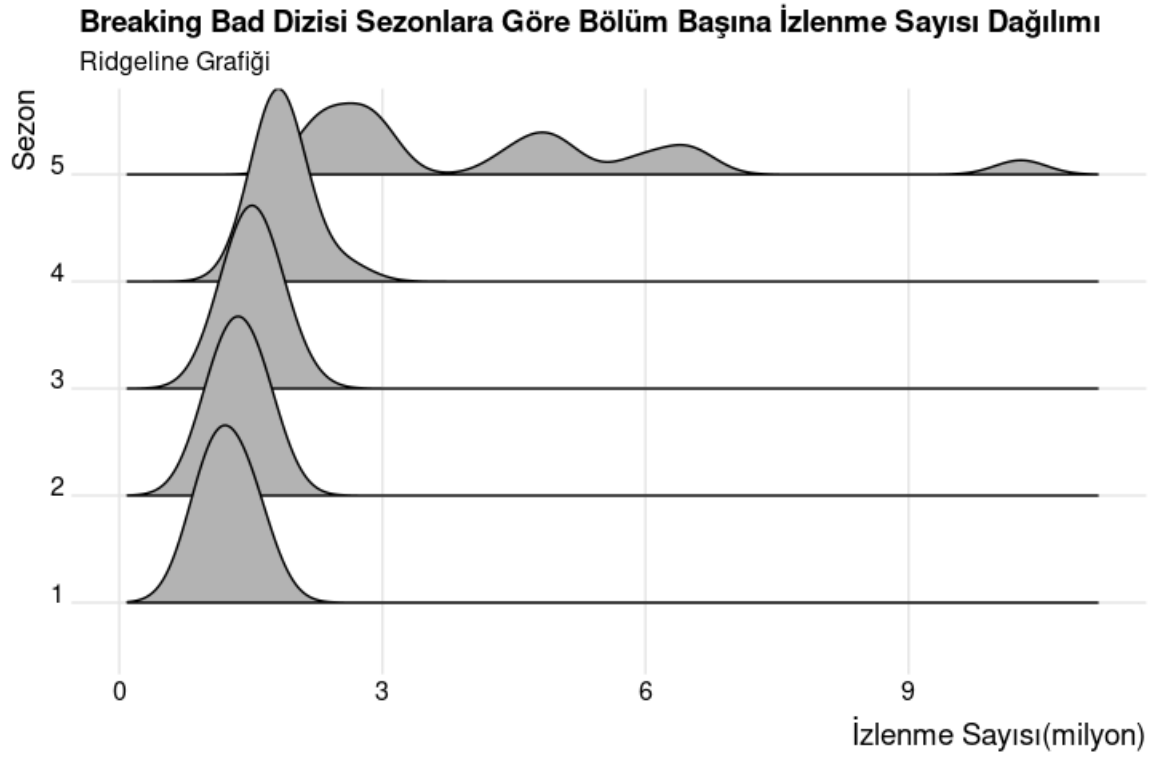
Yukarıda sezonlara göre bölüm sürelerinin dağılımını içeren kutu grafiği kullanılmıştır. Bu grafikte en dikkat çeken durum, her sezon için bölümlerin genellikle belirli bir süreye sadık kalınarak çekildiği görülmektedir. Grafikte her sezon için bazı uç değerler görülmekte olup, bu uç değerler en çok 4. Sezonda görüldüğü için **en yüksek değişimin** gözlemlendiği sezon da bu sezon olmaktadır.

2.2.Sezonlara göre rating dağılımlarını araştırınız. Rating değişiminin en düşük olduğu sezonu belirleyiniz.



Yukarıda sezonlara göre bölüm reytinglerinin dağılımını içeren ridgeline grafiği kullanılmıştır. Grafikte en yüksek reytinglerin 5. Sezonda olduğu görülmektedir. 3. ve 4. Sezonlar dizi için reyting değişiminin en yüksek olduğu sezonlar olduğu görülmesine karşın reytinglerin yüksek olduğu dikkat çekicidir. Reyting değişimi 1. Sezonda **en düşük** olarak görülmektedir.

2.3.Sezonlara göre izlenme sayısı dağılımlarını araştırınız. İzlenme sayısı değişiminin en düşük olduğu sezonu belirleyiniz.



Yukarıda sezonlara göre izlenme sayısı(milyon) dağılımını içeren ridgeline grafiği kullanılmıştır. İzlenme sayısındaki en yüksek değişimin 5. Sezonunda olduğu görülmektedir. 1. ve 2. Sezonlar izlenme sayısı değişiminin düşüklüğü bakımından birbirlerine çok yakın olsalar da 1. Sezonunda fark 0.53(milyon) iken 2. Sezonunda ise 0.62(milyon) şeklindedir. Buna göre izlenme sayısındaki değişim 1. Sezonunda **en düşüktür**.

Not: 2. Sezonunda 5 bölüme ait veri olmadığı göz önünde bulundurulmalıdır.

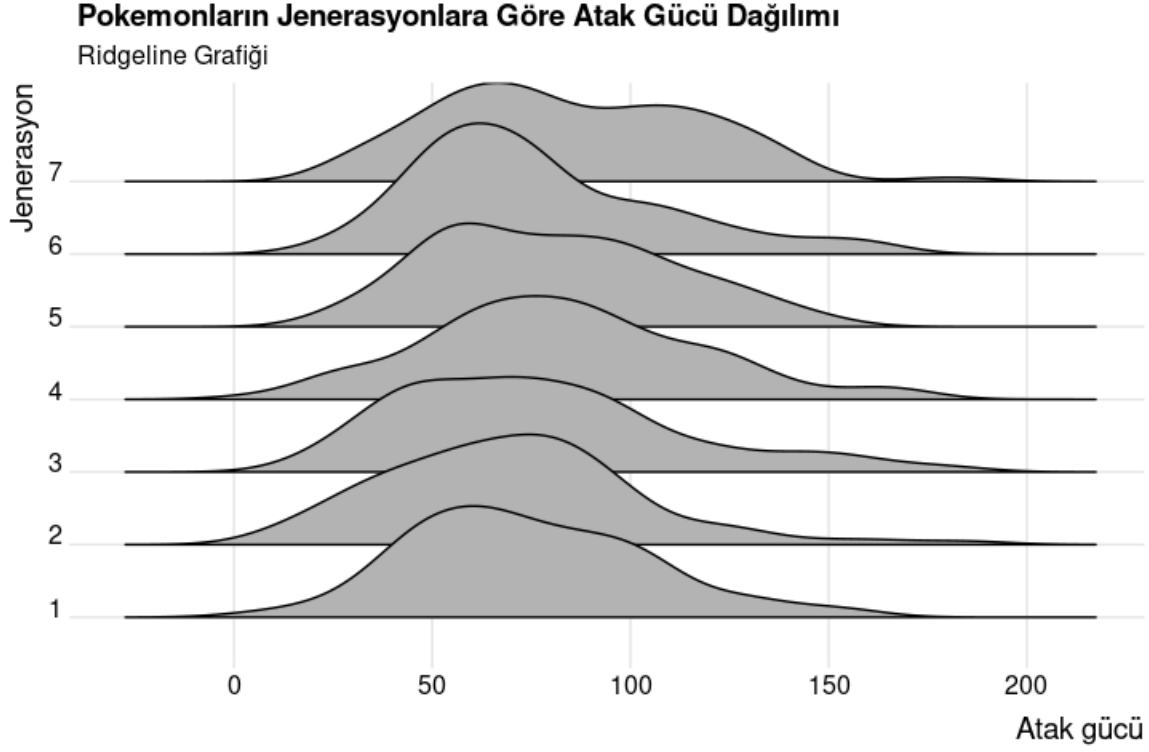
3.Pokemon Veri Seti

Bu veri setinde Pokemon çizgi filmine ait birçok veri bulunmaktadır. 14 değişken ve bu değişkenler için 801 farklı gözlem yer almaktadır. Bu raporda incelenecek olan değişkenler ise şunlardır:

- Jenerasyon

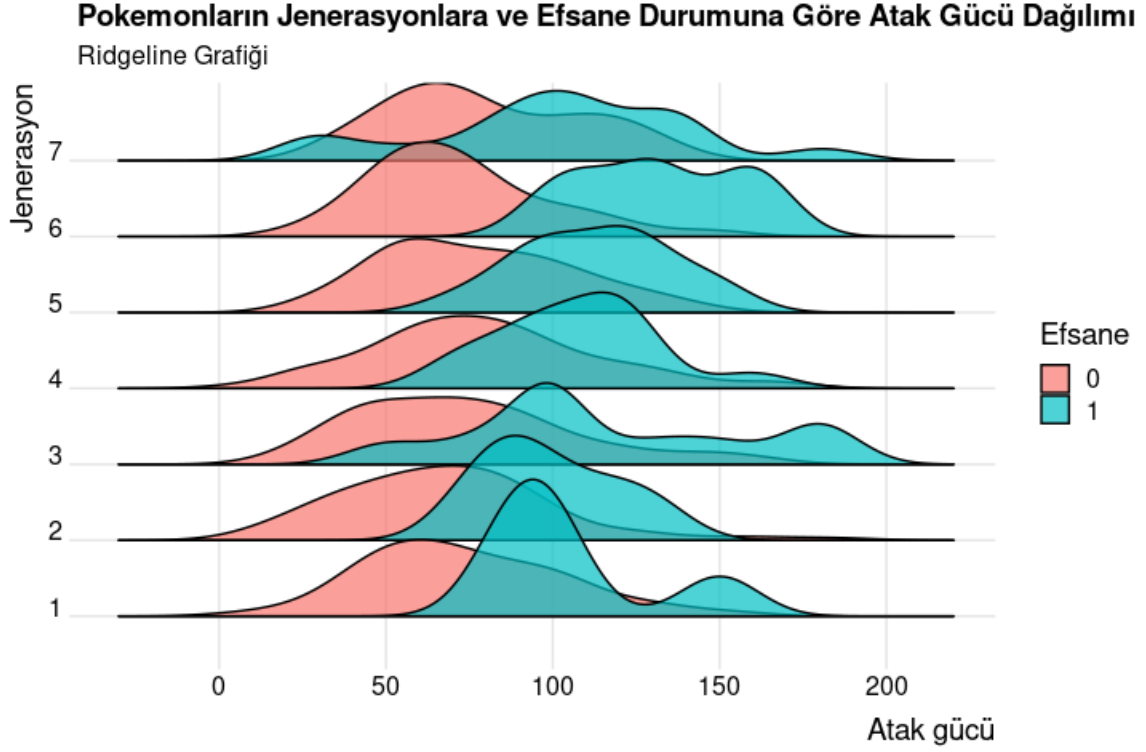
- Atak gücü
- Hız
- Efsane

3.1. Jenerasyonlarına göre pokemonların atak güçlerinin dağılımını araştırınız.



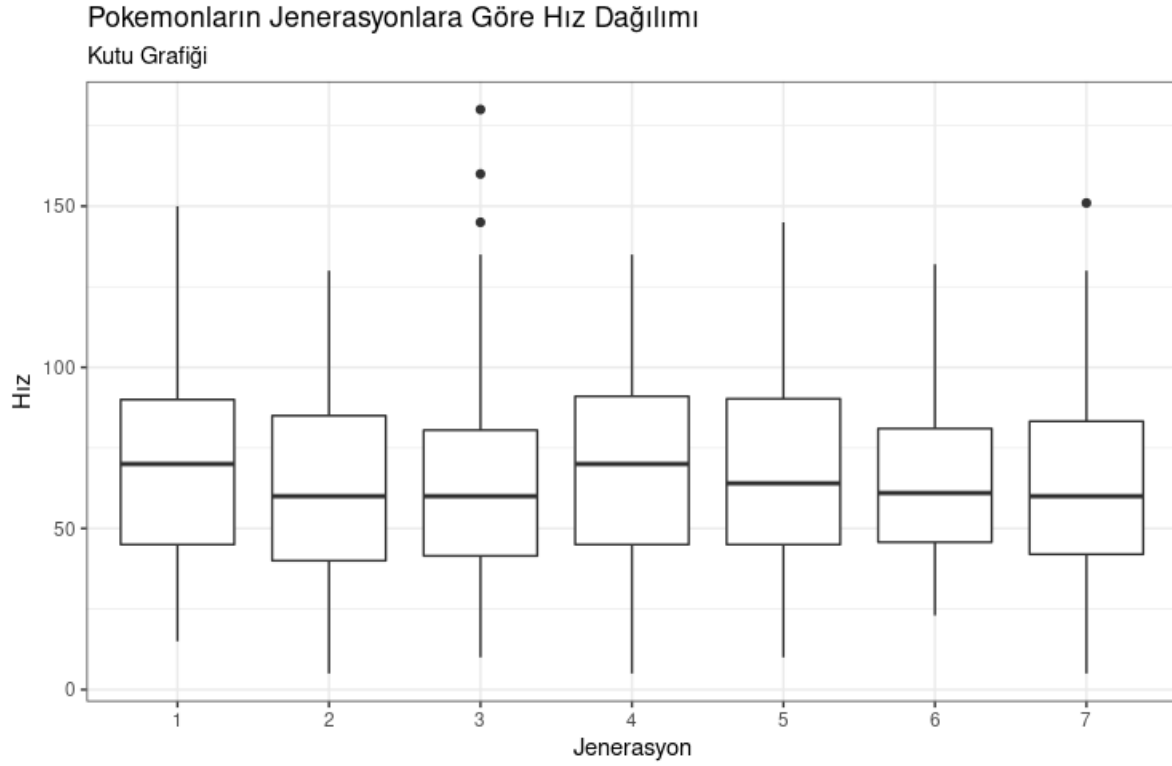
Yukarıda jenerasyonlarına göre pokemonların atak güçlerinin dağılımını içeren ridgeline grafiği kullanılmıştır. Grafikte pokemonların atak güçlerindeki **en yüksek değişimin** 7. Jenerasyonda olduğu görülmektedir. 1. Jenerasyon ise atak güçlerinin **en düşük** olduğu jenerasyondur. Genel olarak her jenerasyonda ortalama güçte yoğunluk olduğu görülse de 7. Jenerasyon **en güçlü** olarak görülmektedir.

3.2. Jenerasyon ve efsane olup olmadıklarına göre atak güçlerinin dağılımını araştırınız.



Yukarıda jenerasyon ve efsane olup olmadıklarına göre atak güçlerinin dağılımını içeren ridge-line grafiği kullanılmıştır. Grafikte efsane olmayan pokemonlar “0”, efsane olan pokemonlar ise “1” olarak ikili sisteme göre kodlanmıştır. Her jenerasyon için efsane pokemonların, diğer Pokemonlardan **daha güçlü** oldukları görülmektedir. Efsane pokemonların atak güçleri arasındaki **en yüksek değişimler** 3. ve 7. Jenerasyonda görülmektedir.

3.3. Jenerasyonlarına göre hızlarının dağılımını araştırınız. En hızlı jenerasyonu belirleyiniz.



Yukarıda jenerasyonlarına göre hızlarının dağılımını içeren kutu grafiği kullanılmıştır. Grafikte 3. Jenerasyonun **en hızlı** olduğu görülmektedir. 4. ve 5. Jenerasyonlar genel Pokemon yoğunluğu açısından daha fazla hızlı Pokemon içermesine rağmen en hızlı Pokemon olarak 3. Jenerasyonu geçememiş olduğu görülmektedir.