



# Rating Product & Sorting Reviews in Amazon



# İş Problemi

E-ticaretteki en önemli problemlerden bir tanesi ürünlere satış sonrası verilen puanların doğru şekilde hesaplanmasıdır. Bu problemin çözümü e-ticaret sitesi için daha fazla müşteri memnuniyeti sağlamak, satıcılar için ürünün öne çıkması ve satın alanlar için sorunsuz bir alışveriş deneyimi demektir. Bir diğer problem ise ürünlere verilen yorumların doğru bir şekilde sıralanması olarak karşımıza çıkmaktadır. Yanıltıcı yorumların öne çıkması ürünün satışını doğrudan etkileyeceğinden dolayı hem maddi kayıp hem de müşteri kaybına neden olacaktır. Bu 2 temel problemin çözümünde e-ticaret sitesi ve satıcılar satışlarını arttırırken müşteriler ise satın alma yolculuğunu sorunsuz olarak tamamlayacaktır.





# Veri Seti Hikayesi

Amazon ürün verilerini içeren bu veri seti ürün kategorileri ile çeşitli metadataları içermektedir. Elektronik kategorisindeki en fazla yorum alan ürünün kullanıcı puanları ve yorumları vardır.

12 Değişken

4915 Gözlem

71.9 MB

<b>reviewerID</b>	Kullanıcı ID'si
<b>asin</b>	Ürün ID'si
<b>reviewerName</b>	Kullanıcı Adı
<b>helpful</b>	Faydalı değerlendirme derecesi
<b>reviewText</b>	Değerlendirme
<b>overall</b>	Ürün rating'i
<b>summary</b>	Değerlendirme özeti
<b>unixReviewTime</b>	Değerlendirme zamanı
<b>reviewTime</b>	Değerlendirme zamanı Raw
<b>day_diff</b>	Değerlendirmeden itibaren geçen gün sayısı
<b>helpful_yes</b>	Değerlendirmenin faydalı bulunma sayısı
<b>total_vote</b>	Değerlendirmeye verilen oy sayısı



# Proje Görevleri

---

**Görev 1:** Average Rating’i güncel yorumlara göre hesaplayınız ve var olan average rating ile kıyaslayınız.

Paylaşılan veri setinde kullanıcılar bir ürüne puanlar vermiş ve yorumlar yapmıştır. Bu görevde amacımız verilen puanları tarihe göre ağırlıklandırarak değerlendirmek. İlk ortalama puan ile elde edilecek tarihe göre ağırlıklı puanın karşılaştırılması gerekmektedir.

**Adım 1:** Ürünün ortalama puanını hesaplayınız.

**Adım 2:** Tarihe göre ağırlıklı puan ortalamasını hesaplayınız.

**Adım 3:** Ağırlıklandırılmış puanlamada her bir zaman diliminin ortalamasını karşılaştırıp yorumlayınız.

## Görev 2: Ürün için ürün detay sayfasında görüntülenecek 20 review'ı belirleyiniz.

**Adım 1:** helpful\_no değişkenini üretiniz.

- total\_vote bir yoruma verilen toplam up-down sayısıdır.
- up, helpful demektir.
- Veri setinde helpful\_no değişkeni yoktur, var olan değişkenler üzerinden üretilmesi gerekmektedir.
- Toplam oy sayısından (total\_vote) yararlı oy sayısı (helpful\_yes) çıkarılarak yararlı bulunmayan oy sayılarını (helpful\_no) bulunuz.

**Adım 2:** score\_pos\_neg\_diff, score\_average\_rating ve wilson\_lower\_bound skorlarını hesaplayıp veriye ekleyiniz.

- score\_pos\_neg\_diff, score\_average\_rating ve wilson\_lower\_bound skorlarını hesaplayabilmek için score\_pos\_neg\_diff, score\_average\_rating ve wilson\_lower\_bound fonksiyonlarını tanımlayınız.
- **score\_pos\_neg\_diff**'a göre skorlar oluşturunuz. Ardından; df içerisinde **score\_pos\_neg\_diff** ismiyle kaydediniz.
- **score\_average\_rating**'a göre skorlar oluşturunuz. Ardından; df içerisinde **score\_average\_rating** ismiyle kaydediniz.
- **wilson\_lower\_bound**'a göre skorlar oluşturunuz. Ardından; df içerisinde **wilson\_lower\_bound** ismiyle kaydediniz.

**Adım 3:** 20 Yorumu belirleyiniz ve sonuçları Yorumlayınız.

- **wilson\_lower\_bound**'a göre ilk 20 yorumu belirleyip sıralayınız.
- Sonuçları yorumlayınız.

**miuul**

miuul.com