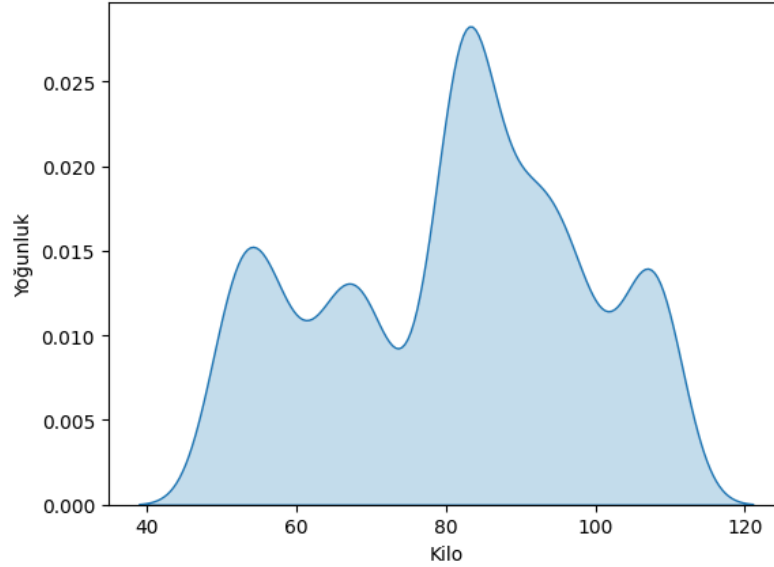
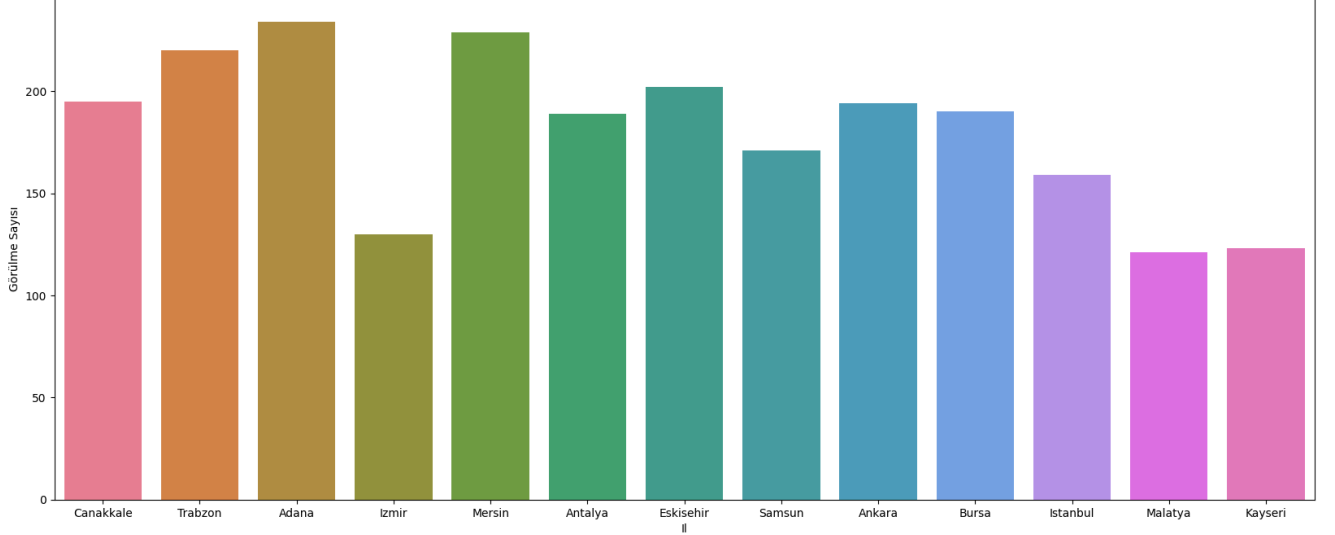


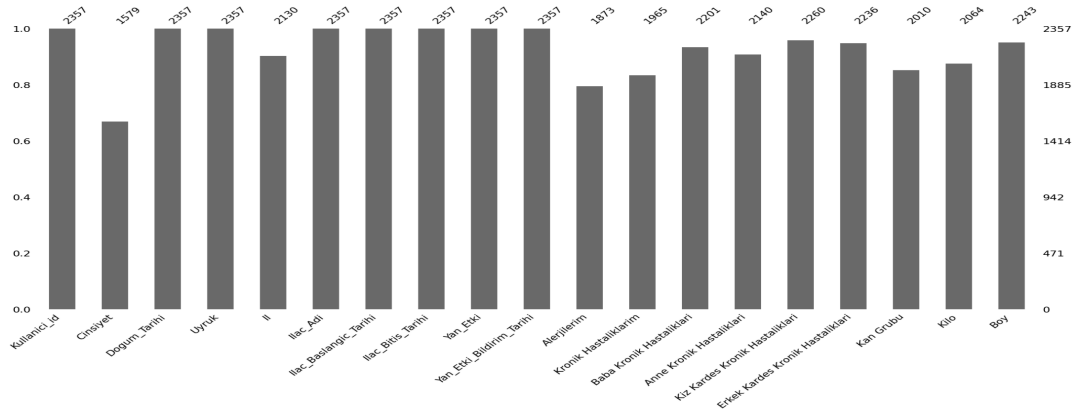
MELİH KARAKAŞ

melihkarakas88@gmail.com

Veri kümesi içerisinde birden çok nitelik var. Niteliklerin daha net biçimde anlaşılabilmesi için çeşitli görselleştirmeleri ve ortalama, standart sapma gibi değerlerin incelenmesi aşamalarını gerçekleştirdim. Grafik çizimleri mevcut. Aşağıda birkaç örneği bulunmaktadır.



Bu grafik çizimlerinin gerçekleştirildiği sırada verilerin dağılımını gösterilirken aynı zamanda eksik verilerin görselleştirilmesi için de bir grafik çizdim. Aşağıda bulunan grafik, niteliklerin ne kadarlık kısmının eksik olduğunu bize açıklar niteliktedir.



Ardından artık eksik verilerin doldurulması, aykırı değer analizi, kategorik verilerin kodlanması ve diğer ön işleme aşamaları gerçekleştirilmiştir. Bu noktada sayısal eksik veriler en çok tekrar eden veriler ile (medyan), kategorik veriler veri kümesinde kendisinden önce gelen değer ile değiştirilmiştir. Aykırı değer analizinde LocalOutlierFactor kullanılmış ve aykırı değerler silinmiştir. Aynı zamanda verilerin normalleştirme aşamaları da gerçekleştirilmekle birlikte bu alanda MinMaxScaler yöntemi kullanılmıştır.