

Bu kod, veri analizi için yaygın olarak kullanılan bazı Python kütüphanelerini kullanarak iki veri seti üzerinde temel veri analizi ve görselleştirme işlemleri gerçekleştirir.

Kütüphaneler:

- pandas: Veri setlerini işleme ve analiz etme için kullanılır.
- numpy: Matematiksel işlemler ve sayısal veri işlemleri için kullanılır.
- seaborn: Veri görselleştirme için kullanılır.
- matplotlib: Veri görselleştirme için kullanılır.

Veri Oluşturma:

- İlk veri seti (df1), numpy kütüphanesi kullanılarak 100 satır ve 3 sütun içeren rastgele verilerle oluşturulur.
- Veri seti, A sütununda 0 ile 100 arasında tamsayılar, B sütununda 50 ortalama ve 10 standart sapmaya sahip normal dağılımlı sayılar ve C sütununda "X", "Y" ve "Z" harflerinden oluşan kategorik veriler içerir.

Veri Çekme:

- İkinci veri seti (df2), seaborn kütüphanesi kullanılarak GitHub'dan bir CSV dosyası olarak çekilir.
- Veri seti, iris çiçeğinin türlerine ve boyutlarına ilişkin veriler içerir.

Veri Setlerinin İlk Bakışta İncelenmesi:

- head() fonksiyonu kullanılarak her iki veri setinin ilk 5 satırı yazdırılır.
- Bu sayede veri setlerinin genel yapısı ve içerdiği değişkenler hakkında bilgi edinilir.

Veri Setlerinin Temel İstatistiklerinin İncelenmesi:

- describe() fonksiyonu kullanılarak her iki veri setinin temel istatistikleri (ortalama, standart sapma, minimum değer, maksimum değer vb.) yazdırılır.
- Bu sayede veri setlerinin genel dağılımı ve değişkenler arasındaki ilişkiler hakkında bilgi edinilir.

Eksik Verilerin İncelenmesi:

- isnull().sum() fonksiyonu kullanılarak her iki veri setinde eksik veri olup olmadığı kontrol edilir.
- Bu sayede veri setlerinde herhangi bir eksik veri olup olmadığı ve eksik verilerin hangi değişkenlerde yoğunlaştığı belirlenir.

Kategorik Değişkenlerin İncelenmesi:

- value_counts() fonksiyonu kullanılarak ilk veri setindeki kategorik değişkenin ("C" sütunu) her bir kategorinin kaç kez tekrarlandığını gösteren bir tablo şeklinde yazdırılır.
- Bu sayede kategorik değişkenin dağılımı ve hangi kategorinin daha yaygın olduğu hakkında bilgi edinilir.

Veri Görselleştirme:

- `sns.histplot()` fonksiyonu kullanılarak ilk veri setindeki her sayısal değişkenin histogramı çizilir.
- Bu sayede her bir değişkenin dağılımı ve ortalama değeri hakkında bilgi edinilir.
- `sns.pairplot()` fonksiyonu kullanılarak ikinci veri setindeki tüm sayısal değişkenler çiftler halinde çizilir.
- Bu sayede değişkenler arasındaki ilişkiler ve korelasyonlar görselleştirilir.
- `sns.heatmap()` fonksiyonu kullanılarak ikinci veri setindeki tüm sayısal değişkenler arasındaki korelasyon katsayılarını gösteren bir ısı haritası çizilir.
- Bu sayede değişkenler arasındaki ilişkilerin gücü ve yönü hakkında bilgi edinilir.