

NumPy Nedir?

NumPy, Python programlama dili için bilimsel hesaplamaların temelini oluşturan güçlü bir kütüphanedir.

Kısaltısı, Python'da *büyük, çok boyutlu diziler (arrayler) ve matrisler* üzerinde hızlı ve verimli işlemler yapmamızı sağlar.

Temel Özellikleri:

1. ndarray (*N-dimensional Array*):

- NumPy'ın kalbi, *çok boyutlu dizi* anlamına gelen ndarray nesnesidir.
- Bunlar, standart Python listelerinden *çok daha hızlı* ve *daha az bellek* kullanarak çalışır, çünkü C veya C++ dillerinde yazılmış kısımları vardır.
- Bu diziler *homojen* olmak zorundadır, yani tüm elemanları aynı veri tipinden (tam sayı, ondalık sayı vb.) olmalıdır.

2. Yüksek Seviye Matematiksel İşlevler:

- Doğrusal cebir, Fourier dönüşümü, rastgele sayı üretimi ve istatistiksel işlemler (ortalama, standart sapma vb.) gibi karmaşık matematiksel işlemleri kolayca yapabileceğiniz bir dizi fonksiyon sunar.

Neden NumPy Kullanılır?

Python'ın standart listeleri yerine NumPy kullanmanın ana nedenleri şunlardır:

- **HIZ** = Özellikle büyük veri kümeleriyle çalışırken, NumPy işlemleri standart Python listelerine göre kat kat *daha hızlıdır* (bazen 50 kata kadar).
- **VERİMLİLİK** = Bellek kullanımında daha etkilidir.
- **FONKSİYONELLİK** = Veri bilimi, makine öğrenimi ve bilimsel hesaplamalar için vazgeçilmez olan gelişmiş array ve matris operasyonlarını sağlar.

Kullanım Alanları

NumPy, veri bilimi ve makine öğrenimi ekosisteminin temel taşıdır. Genellikle şu alanlarda kullanılır:

- **VERİ BİLİMİ VE ANALİZLERİ** = Veri temizleme, işleme ve manipülasyonu.
- **MAKİNE ÖĞRENİMİ** = Özellikle model oluşturma, matris işlemleri ve tensör gösterimleri için temel bir araçtır.
- **İSTATİSTİK** = İstatistiksel hesaplamalar ve analizler.

- *Görüntü İşleme*: Görüntülerini matrisler olarak temsil etme ve işleme.
-

Özetle, NumPy, Python'da hızlı ve verimli bilimsel hesaplamalar yapmanın anahtarıdır.

Pandas Nedir?

Pandas, Python programlama dili için geliştirilmiş, *veri analizi* ve *manipülasyonu* konusunda tartışmasız en güçlü kütüphanedir. Adı, ekonometri terimi olan "Panel Data"dan (Panel Veri) gelir.

Eğer elinde tablolar, Excel dosyaları, SQL veritabanlarından çekilmiş veriler gibi yapılandırılmış bir veri varsa, Pandas onu bir çalışma masasına (DataFrame) dönüştürür ve üzerinde dilediğin gibi çalışmasını sağlar.

Temel Veri Yapıları:

1. *DataFrame* (Veri Çerçevesi):

- Pandas'in en temel ve en çok kullanılan yapısıdır. Tıpkı bir *Excel tablosu* veya bir *SQL tablosu* gibi düşünebilirsin; satırları ve isimlendirilmiş sütunları (kolonları) vardır. Farklı veri tiplerini (sayı, metin, tarih vb.) tek bir tabloda tutabilir.

2. *Series* (Seri):

- Tek boyutlu bir veri yapısıdır. Tıpkı bir DataFrame'in tek bir sütunu gibi düşünülebilir. Her bir elemanının bir etiketi (index) vardır.

Neden Pandas Kullanılır?

Pandas, veri bilimi projelerinde "veri hazırlama" aşamasının neredeyse tamamını halledebilir:

- *Veri Okuma/Yazma*: Farklı formatlardaki (CSV, Excel, JSON, SQL) verileri kolayca okuyup yazabilir.
 - *Veri Temizleme*: Eksik verileri doldurma, hatalı verileri düzeltme, veri tiplerini dönüştürme gibi temizlik işlemlerini çok hızlı ve kolay hale getirir.
 - *Veri Manipülasyonu*: Veriyi filtreleme, gruplama (groupby), birleştirme (merge), yeniden şekillendirme (pivot etme) gibi karmaşık işlemleri tek bir komutla yapabilirsin.
 - *Istatistiksel Analiz*: Veri seti hakkında özet istatistikler (ortalama, medyan, standart sapma vb.) hızlıca alınabilir.
-

Matplotlib Nedir?

Matplotlib, Python'da *veri görselleştirme* için kullanılan temel çizim kütüphanesidir. Veri analizinden elde ettiğimiz sonuçları, insan beyninin daha kolay anlayabileceği *grafiklere* ve *çizimlere* dönüştürmek için kullanılır.

Temel Özellikleri:

1. Geniş Grafik Çeşitliliği:

- Çizgi grafikleri (Line Plot)
- Çubuk grafikleri (Bar Chart)
- Dağılım grafikleri (Scatter Plot)
- Histogramlar
- Pasta grafikleri (Pie Chart)
- 3D grafikler (ek uzantılarla)
- Ve daha pek çok özelleştirilmiş görselleştirmeyi oluşturabilir.

2. Yüksek Özelleştirme:

- Matplotlib, grafiğin neredeyse her yönünü (renkler, eksen etiketleri, başlıklar, yazı tipleri, çözünürlük) detaylıca özelleştirme imkanı sunar. Bu sayede, bilimsel yayınlar veya profesyonel raporlar için yüksek kaliteli çıktılar elde edilebilir.

3. Temel Oluşturucu:

- Diğer popüler görselleştirme kütüphaneleri (örneğin Seaborn), Matplotlib'in üzerine inşa edilmiştir ve Matplotlib'in işlevsellliğini daha basit bir sözdizimi ile sunar.

Neden Matplotlib Kullanılır?

Matplotlib, analizin hikayesini anlatmak için kullanılır:

- *Trendleri Keşfetme*: Veri setindeki değişimleri ve eğilimleri (örneğin zaman içindeki hisse senedi fiyatı değişimi) kolayca gösterir.
- *İlişkileri Anlama*: İki değişken arasındaki ilişkiyi (örneğin gelir ve harcama) dağılım grafiği ile görselleştirir.
- *Dağılımı Gösterme*: Verilerin nasıl yayıldığını (örneğin bir sınavdaki notların dağılımı) histogramlarla ortaya koyar.
- *Raporlama ve Sunum*: Analiz sonuçlarını profesyonel ve anlaşılır görsellerle destekler.

Özetle Üçü Bir Arada:

- *NumPy*: Sayısal işlemler ve hızlı dizi (array) oluşturma. *Temel matematik*.
- *Pandas*: Yapılandırılmış veriyi düzenleme, temizleme ve analiz etme. *Veri masası*.
- *Matplotlib*: Analiz sonuçlarını görselleştirme. *Veri çizimi*.

Bu üç kütüphane, genellikle bir veri bilimi projesinde *birlikte ve sırayla* kullanılır.