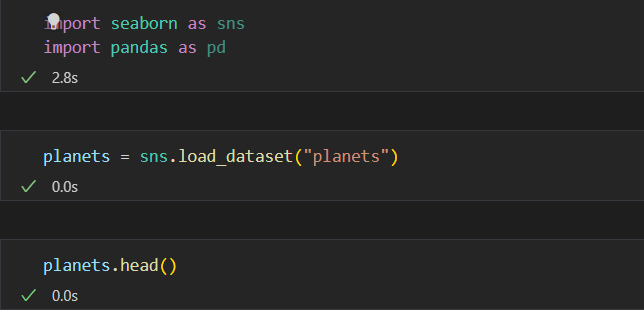
**Veri Setinin Betimlenmesi**

Her zaman ilk iş olarak veri setinin **yapısal işlemlerini** (yedeğini vb. işlemleri) yaparak başlıyoruz

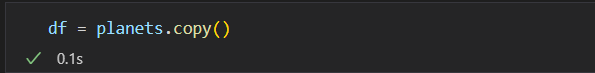


metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu



metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

“Method” isimli değişkeni Kategorik değişken olarak dönüştürelim

metin içeren bir resim

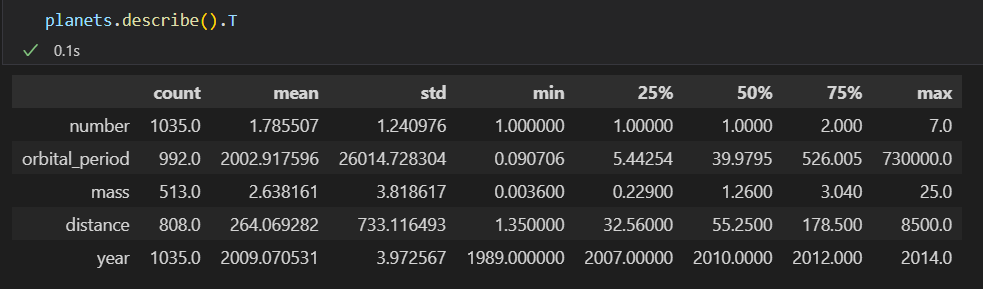
Açıklama otomatik olarak oluşturuldu

…

Bu işlemlerin ardından **Betimsel işlemlere** geçebiliriz

Describe.T alarak başlamaktayız

**NOT:** Describe(betimlemek) metodu eksik değerleri göz ardı eder ve kategorik değişkenleri dışarıda bırakır. (Describe sadece sayısal değişkenleri içerir)



Standart sapması, varyansı ve diğer işlemlerine bakarak buradaki değerleri, az çok veri setini genel olarak yorumlamaya çalışarak devam ediyoruz

1. Number, gezegen adedini belirtiyordu mesela ve görüyoruz ki min 1 tane max ise 7 tane gezegen keşfedilmiş
2. Orbital\_Period, yörünge periyodunu temsil ediyordu. Min ve max arasında çok fark var ve std mean’den daha büyük yani ortalamadan (merkezden) uzak bir değer seti var
3. Mass, gezegen kütlesiydi. Std sapma, mean’den büyük demek ki ortalamadan(merkezden) uzak değerler mevcut - bu bilgiye çeyrekliklerden %25 ve %75 arasındaki farktan da bakarak destekledik elbette-
4. Distance, mesafelerdi. Min değer 1.35 iken max değer 8500. Demek ki dağılımı (standart sapması) çok geniş
5. Year, burada numeric olarak alındığı için bunun type’ını Date olarak dönüştürmeli ve ondan sonra değerlendirmeliyiz.

İnclude (İçermek)

Tüm değişkenleri (sayısallara, kategorik değişkenlerde dahil edince) tabloda değerlendirmek için kullanılabilir fakat bu çok anlamlı olmayacaktır.

