

Pandas Kütüphanesi - Veri Çerçevesi (Data Frame) Özet Bilgileri (CheatSheet)

:: AÇIKLAMA ::

```
import pandas as pd
```

pd : pandas kütüphanesi için atanmış kısaltmayı temsil eder.

```
df = pd.DataFrame(Veri_Yapısı)
df = pd.read_excel("excel.xlsx")
```

df : Oluşturulan **Veri Çerçevesine (Data Frame'e)** atanan kısaltmayı temsil eder. Bu kısaltmalar aşağıda kullanılmıştır.

:: VERİ ÇERÇEVESİ OLUŞTUR ::

Fonksiyon / Metot / Parametre

Açıklama

```
pd.DataFrame(dosya, nesne ya da sözlük)
```

Bir Dosyadan, İterable Nesneden ya da Sözlük yapısından Veri Çerçevesi (Data Frame) Oluştur.

```
pd.read_csv("dosya_adı.csv")
```

CSV dosyasını içe aktararak Veri Çerçevesi Oluştur.

```
pd.read_excel("dosya_adı.uzantı")
```

Excel, Calc dosyalarını içe aktararak Veri Çerçevesi Oluştur. Bu fonksiyon **xls**, **xlsx**, **xlsm**, **xlsb**, **odf**, **ods** ve **odt** uzantılı dosyaları destekler.

```
pd.read_sql(query,connection_object)
```

Bir SQL veritabanını içe aktararak Veri Çerçevesi Oluştur.

```
pd.read_html(url)
```

Bir URL (websitesi) ya da yerel html dosyasından tablo içe aktararak Veri Çerçevesi Oluştur.

```
pd.read_json(dosya_adı.json)
```

JSON biçimindeki dosyayı içe aktararak Veri Çerçevesi Oluştur.

```
pd.read_table(dosya_adı.csv)
```

```
pd.read_table("dosya_adı.csv", delimiter=",")
```

CSV dosyasını tablo olarak içe aktararak Veri Çerçevesi Oluştur. **delimiter** parametresi ile, tablo verilerini sütunlara bölebiliriz. Aksi halde her satırdaki veri bir sütuna yazılır.

```
pd.read_clipboard()
```

CTRL + C ile Panoya kopyaladığımız veriyi (örneğin bir Excel tablosu ya da websitesindeki tabloyu) Veri Çerçevesine dönüştürür.

Pandas Kütüphanesi - Veri Çerçevesi (Data Frame) Özet Bilgileri (CheatSheet)

:: VERİ ÇERÇEVESİNİ KEŞFET ::

| Fonksiyon / Metot / Parametre | Açıklama |
|--|--|
| <code>df.head()</code> / <code>df.head(n)</code> | Veri Çerçevesinin ilk n satırını görüntüler. Parantez içerisine sayı yazmazsak ilk 5 satırı görüntüler. Böylece Veri Çerçevesi hakkında ön bilgi edinmemizi sağlar. |
| <code>df.tail()</code> / <code>df.tail(n)</code> | Veri Çerçevesinin son n satırını görüntüler. Parantez içerisine sayı yazmazsak ilk 5 satırı görüntüler. |
| <code>df.columns</code> | Veri Çerçevelerinin başlık satırını çıktı olarak verir / görüntüler. |
| <code>len(df)</code> | Veri Çerçevesinin kaç satırdan oluştuğu bilgisini döndürür. |
| <code>df.info()</code> | Veri Çerçevesinin satır ve sütun sayısı, başlık tipleri (sayı, metin, ...vb) ve doluluk oranı (boş olmayan hücre sayısı) hakkında bilgi görüntüler. |
| <code>df.dtypes</code> | Veri çerçevesinin başlık tiplerini görüntüler. |
| <code>df.describe()</code> | Sayısal veri barındıran sütunlar hakkında detaylı matematiksel bilgiler verir. (Kaç adet veri olduğu, ortalamaları, en küçük ve en büyük sayı,...vb) |
| <code>df.shape</code> | Veri Çerçevesinin satır ve sütun bilgisini döndürür/görüntüler. |
| <code>df.index</code> | Veri Çerçevesinin indeks isimlerini döndürür/görüntüler. |
| <code>df.isnull().any()</code> | Veri Çerçevesinin herhangi bir boş değer içerip içermediğini görüntüler. |

Pandas Kütüphanesi - Veri Çerçevesi (Data Frame) Özet Bilgileri (CheatSheet)

:: VERİ ÇERÇEVESİNİ DÜZENLE ::

| Fonksiyon / Metot / Parametre | Açıklama |
|---|--|
| <code>df = pd.read_excel("dosya_adı.xls", index_col=0)</code> | Veri Çerçevesi (Data Frame) oluştururken, sütunlardan birini, indeks değeri olarak ayarlamak/atamak için index_col parametresi kullanılır. Parametreye, değer olarak, indis olarak atanacak Sütunun indis değeri yazılır. |
| <code>df.set_index("Sütun_adı")</code> | Veri Çerçevesi oluşturduktan sonra, istediğimiz zaman indeks değerlerini değiştirmek / atamak için set_index() parametresi kullanılır. |
| <code>df = pd.read_excel("dosya_adı.xls", usecols = ["Sütun1", "Sütun5", Sütun9"])</code> | Veri Çerçevesi (DataFrame) oluştururken, sadece istediğimiz sütunları kullanmak istediğimizde usecols parametresini kullanmalıyız. Değer olarak sütun adı ya da sütun indisi yazılabilir. |
| <code>df = pd.read_excel("dosya_adı.xls", usecols = [2,3,8])</code> | |
| <code>df = pd.read_excel("dosya_adı.xls", skiprows = [0,5,6])</code> | Veri Çerçevesi (DataFrame) oluştururken, atlamak / göz ardı etmek Veri Çerçevesine dahil etmemek için, skiprows parametresi kullanmalıyız. |
| <code>df["Yeni_Sütun_Adı"] = Atanacak_Değerler</code> <u>Örnek:</u> <code>df["Toplam"] = df["Sütun1"] + df["Sütun2"]</code> | Veri Çerçevesi (DataFrame) Yeni Sütun eklemek ve bu sütuna Değer Atamak için bu yapıyı kullanabiliriz. Atanacak_Değerler, Liste yapısında olabileceği gibi, Diğer sütunların kopyası, matematik, mantık ya da karşılaştırma operatörleri ile belirlenmiş halleri de olabilir. |
| <code>df.drop("Satır ya da Sütun Adı", axis=0/1, inplace=True/False)</code> | Veri Çerçevesinden (Data Frame'den) Satır ya da Sütun Silmek istediğimizde drop() metodunu kullanabiliriz. axis = 0 (Sıfır) Satırları , axis = 1 (Bir) Sütunları temsil eder. inplace parametresi, gerçekleştirilen silme işleminin kalıcı ya da geçici olmasını ayarladığımız kısımdır. inplace = True yazılırsa, yapılan işlem kalıcı hale gelecektir. |

:: STRING METOTLARI ::

| | |
|--|--|
| <code>df["Sütun1"].str.upper()</code> | Sütun1'deki Metinsel ifadelerin tümünü büyük harf çevirir. |
| <code>df["Sütun1"].str.lower()</code> | Sütun1'deki Metinsel ifadelerin tümünü küçük harf çevirir. |
| <code>df["Sütun1"].str.capitalize()</code> | Sütun1'deki Metinsel ifadelerin sadece baş harflerini büyük harf çevirir. |

:: VERİYİ DIŞA AKTAR / FARKLI BİÇİMDE KAYDET ::

| Fonksiyon / Metot / Parametre | Açıklama |
|---|---|
| <code>df.to_csv("dosya_adı.csv")</code> | Veri Çerçevesini CSV dosyası olarak Dışa Aktar/Kaydet . |
| <code>df.to_excel("dosya_adı.uzantı")</code> | Veri Çerçevesini Excel / Calc olarak Dışa Aktar/Kaydet . xls, xlsx, xlsm, xlsb, odf, ods ve odt uzantılı dosyalar desteklenir. |
| <code>df.to_sql(tablo_adı, bağlantı_nesnesi)</code> | Veri Çerçevesini SQL Veri Tabanı olarak Dışa Aktar/Kaydet . |
| <code>df.to_json("dosya_adı.json")</code> | Veri Çerçevesini JSON dosyası olarak Dışa Aktar/Kaydet . |