Pandas

Pandas datalarda analiz ve manipulation yapabilmek icin yazilmis ucretsiz bir kutuphanedir. Ozellikle sayisal tablolar ve zaman serileri icin islemler sunar.Pandas cogunlukla DataFrameler bicimindeki ML icin kullanilir.

Pandas excel gibi dosyalardaki verilerin ice aktarilmasina izin verir.Groupby,join,merge,melt gibi islemlerin yani sira bosluklari doldurma ve veri temizleme gibi ozellikleri de vardir.

```
serr = pd.Series([11,21,13,41,15,61])
serr

0    11
1    21
2    13
3    41
4    15
5    61
dtype: int64
```

Pandas int, string, float, objects gibi her turlu veriyi labellayabilen tek boyutlu bir dizidir. Diziler sadece tek bir liste icerebiliyorken Data Frameler birden fazla liste icerebilir.

Data Frame Basics-1 (Attributes)

DataFrame, verilerin satirlar ve sutunlar olarak hizalandigi iki boyutlu bir veri yapisidir.Veriler,satirlar ve sutunlardan olusur.Buyutu degisebilir,potensiyel olarak heterojen tablo datasidir.

Series

Tek boyutlu serilerdir.

sort()

Verileri siralamak icin kullanilir.

```
sort_index,sort_values,
```

isin()

Icerisine yazilan degerin elimizdeki datadaki varligini eslestirme ile bool sonuc olarak donduren bir methoddur.

keys()

Datadaki verilerin start, stop ve step degerlerini veren bir methoddur.

values

Datadaki tipine bagli olarak serileri array olarak dondurur.

items()

Bu method indexleri ve valuelari iterable bir tuple olarak dondurur.

read_csv()

```
dataframe=pd.read_csv("csv name.csv")
```

Klasorumuzdeki belgeyi okumaya yarayan bir methoddur.

head()

Bu method konuma gore ilk n degerdeki satirlari donderir. Dafault degeri 5'tir.

tail()

Bu method konuma gore son n degerdeki satirlari donderir.Dafault degeri 5'tir.

sample()

Bir object axisinden rastgele ogeler donduren bir methoddur.icerisindeki deger kadar dondurur.

shape

DataFrame'in boyutlarini tanimlar bunu dondurur.

Data Frame Basics-2 (Index & Selecting)

DaraFrame verilerin iki boyutlu olarak satirlar ve sutunlar seklinde hizalanan veri yapisidir.

index()

Tum pandas objeleri icin axis etiketleri saqlayan objecttir.

columns

Datanin sutunlarini getirir.

shape

DataFrame'in boyutlarini dondurur.

size

Bu objedeki elementlerin sayisini int olarak dondurur.

ndim

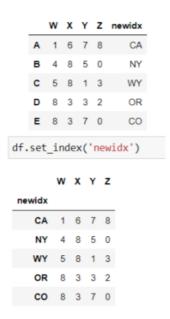
Array boyutlarini veya axis sayisini dondurur.

reset index()

DataFrame indexlerini sifirlar ve default degerleri kullanir.DataFrame birden fazla indexe sahipse MultiIndex ise bu yontem bir veya birden fazla duzeyi kaldirabilir.

set_index()

Mevcut olan sutundaki degerleri kullanarak DataFrame'e yeni index atar.



["col"]

Sutunlari bu sirayla secmek icin parantez icinde belirtmek gerekir.dataFrame'de bir sutun yoksa bir istisna beklenmedik durum olusur.'W' basligindaki degerler olur.

iloc[]

Pozisyona gore gormek icin indexlerle secim yapilir.llk girilen deger (:) indexine gore getirir.ikinci deger ise (:) getirilen datanin icinde nelerin gosterilmesi icin baslik

sayisini gosterir.Eger iki deger arasina virgul konulursa getirilen datanin icindeki degeri direkt getirir.

loc[","]

Ilk girilen degerin datadaki ikinci girilen basliktaki degerini dondurur.

Conditional Indexing

Parantez icinde belirtilen basliktaki yerde kosula uyan degerleri dondurur. df[df["Y"] < 6] Y basliginda 6 dan kucuk degerler

["Baslik"] == "Deger"

Bu baslikta buna esit bir deger var mi diye bakar sonra bool degerler donderir.

df[df["Baslik"] == "Deger"]

Sadece bu kurala uyan degeri donderir.

isin[]

Bu method, data icindeki belli sartlardaki verileri aldigimizda bunlar icinde belli sartlar belirlemek icin kullanilir.Orn;Belli bir seviye uzerinde maas alan kisilerden su mesleklerde calisanlari istedigimizde meslekleri bu method icinde belirtebiliriz.

Data Frame Basics-3 (Properties)

info()

Bu method dizin tipi,sutunlar,bos olmayan degerler ve bellek kullanimi dahil DataFrame hakkinda bir cok bilgi yazdirir.

describe()

eries veya DataFrame'in sayısal değerlerinin yüzdelik, ortalama ve std gibi bazı istatistiksel verilerini hesaplamak için kullanılır.Bir veri kümesinin dağılımının merkezi eğilimini, dağılımını ve şeklini özetleyenleri içeren tanımlayıcı istatistikler oluşturur.

shape

DataFrame'in boyutlarini yazdirir

len()

Bir String'teki uzunluğu (karakter sayısını) döndürür. Sözlükler, listeler veya demetler için giriş sayısını verir.

col.value_counts()

Benzersiz degerlerin sayisini iceren bir seri dondurur.

mean()

Istenen eksen uzerindeki degerlerin ortalamasini dondurur.

col.sum()

Istenen eksen uzerindeki degerlerin toplamini dondurur.

col.unique()

Unique'ler gorunum sirasina gore getirilir.Siralama yapmaz.Benzersiz elemanlari getirir.

isnull().sum()

Null degerleri tespit eder.Her sutundaki toplam deger sayisini dondurur.

drop()

Satirlari veya sutunlari label adlarini,karsilik gelen axis'ini veya sutun adlarini ya da satirlarin indexlerini belirterek kaldirmaya yarar.