

Pandas-1: Tablo ile Çalışma

DataFrame Mantığı

Veri Okuma, Seçme ve Filtreleme

Analitik Mercek: Bölüm 1

DataFrame Nedir?

Özellik (Feature)

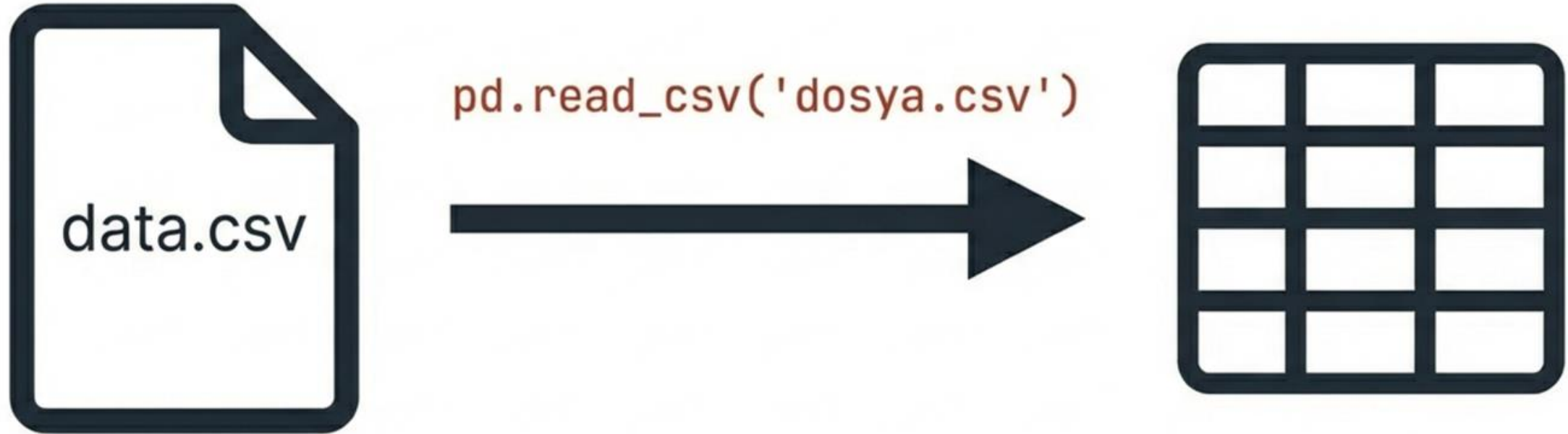
Name	Age	City	Salary
Ali	25	Istanbul	50000
Ayşe	30	Ankara	60000
Mehmet	22	Izmir	45000

Kayıt (Record)

DataFrame = Satır + Sütun Tablosu

Pandas öğrenmek = Tablo düşünmeyi öğrenmek.

Veri Okuma (Acquisition)



CSV dosyasını içeri almanın ve analize başlamanın ilk adımı.

Veriyi Tanıma Protokolü

df.head()

Inter

İlk 5 satır.

Name	Age	City	Salary
Ali	25	Istanbul	50000
Ayşe	30	Ankara	60000
Mehmet	22	Izmir	45000
Fatma	28	Bursa	55000
Can	35	Antalya	70000

df.info()

Inter

Sütun tipleri, boş değer kontrolü.

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 100 entries, 0 to 99  
Data columns (total 4 columns):  
#   Column  Non-Null Count  Dtype  
---  -  
0   Name    100 non-null    object  
1   Age     100 non-null    int64  
2   City    100 non-null    object  
3   Salary  100 non-null    int64  
dtypes: int64(2), object(2)  
memory usage: 3.2+ KB
```

df.describe()

Inter

Sayısal özet (istatistikler).

	Age	Salary
count	100.000000	100.000000
mean	29.500000	58500.000000
std	5.123456	12345.678900
min	20.000000	40000.000000
25%	25.000000	48750.000000
50%	29.000000	57500.000000
75%	34.000000	67500.000000
max	40.000000	85000.000000

Satır ve Sütun Mantığı

Columns (Sütun İsimleri)



	Name	Age	City
0	Ali	25	Istanbul
1	Ayşe	30	Ankara
2	Mehmet	22	Izmir
3	Fatma	28	Bursa

Index
(Satır Kimliği)



DataFrame bu iki temel bileşenden oluşur.

Sütun Seçimi

Tek Sütun (Series)

```
df['Salary']
```

Name	Age	City	Salary
Ali	25	Istanbul	50000
Ayşe	30	Ankara	60000
Mehmet	22	Izmir	45000
Fatma	28	Bursa	55000
Can	35	Antalya	70000

Çoklu Sütun (DataFrame)

```
df[['Age', 'Salary']]
```

Name	Age	City	Salary
Ali	25	Istanbul	50000
Ayşe	30	Ankara	60000
Mehmet	22	Izmir	45000
Fatma	28	Bursa	55000
Can	35	Antalya	70000

Navigasyon: loc vs iloc

iloc

POZİSYON (Konum)

	0	1	2
0	Ali	25	İstanbul
1	Ayşe	30	Ankara
2	Mehmet	22	İzmir

iloc = Sayı (Integer)

loc

ETİKET (Label)

	City	Age	Salary
0	Ali	25	50000
1	Ayşe	30	60000
2	Mehmet	22	45000

loc = İsim (Index/Column Adı)

Seim Örnekleri

```
df.iloc[0:5, 0:3]
```

İlk 5 satır, ilk 3 sütun.

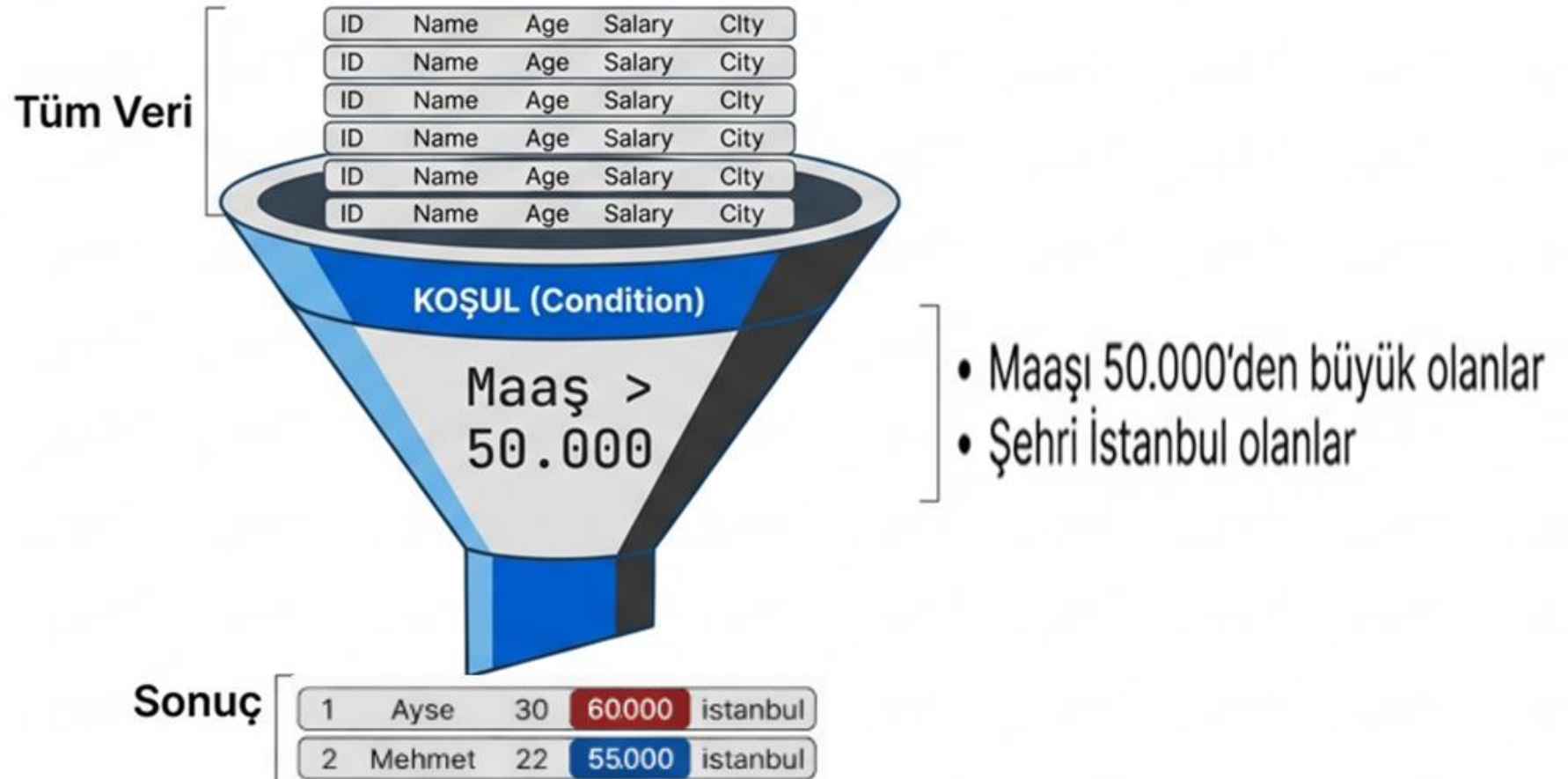
	0	1	2	3	4	5	6	7	8	9
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										

```
df.loc[0:5, ['Age', 'Salary']]
```

Belirli satırlar ve ismen seçilen sütunlar.

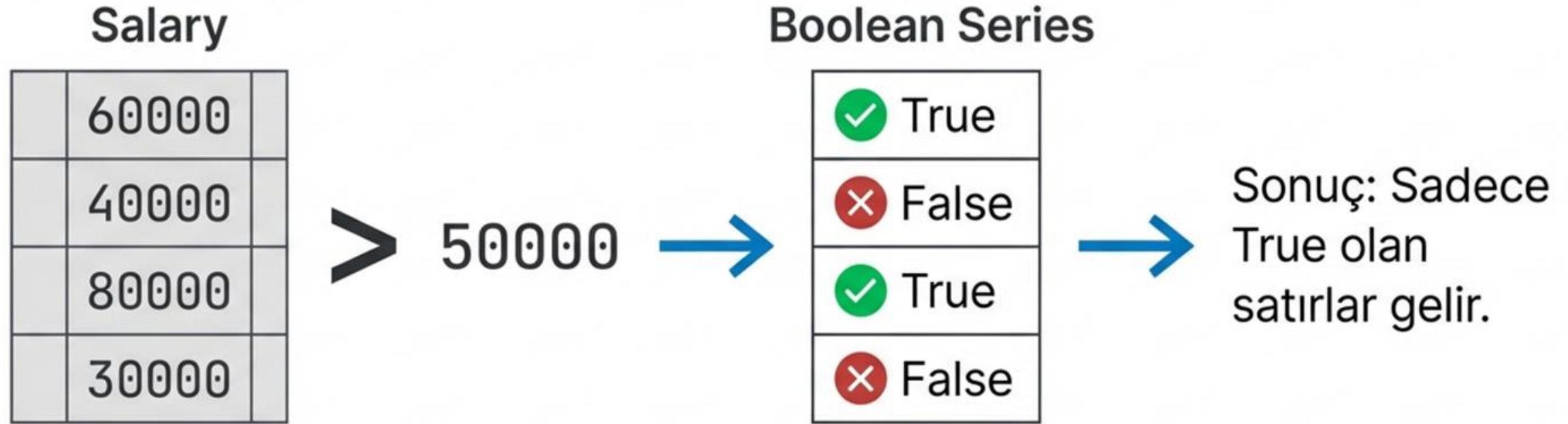
	Name	Age	City	Salary	Dept	YoE	Rating	Bonus	Level	Status
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										

Filtreleme Mantığı

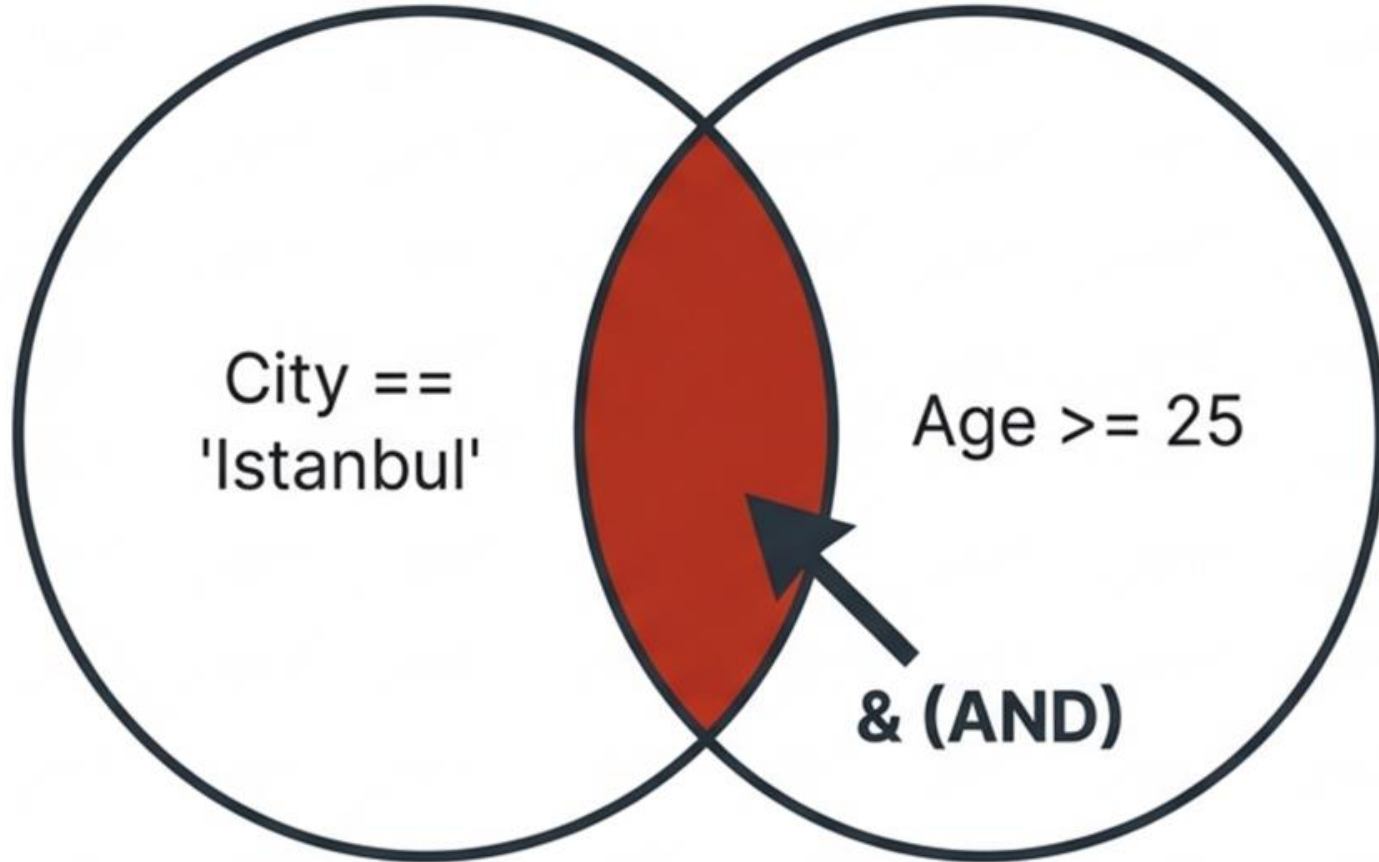


Filtre = Koşula uyan satırları seçmek.

Boolean Mask (Arka Plan İşleyişi)



Çoklu Koşullar (AND / OR)



Dikkat:

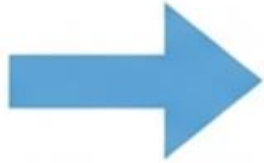
1. 'and' / 'or' kelimelerini kullanma! (&, | kullan)
2. Her koşul parantez içinde olmalı.

```
(df['City'] == 'Istanbul') & (df['Age'] >= 25)
```

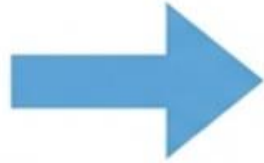
Özet: Analitik Akış



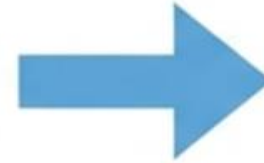
Yükle
(read_csv)



Tanı
(head, info)



Seç
(loc, iloc)



Sorgula
(Filtreleme)

- - - - -