

Eksik değerler (missing values), veri analizi sürecinde sıkça karşılaşılan bir problemdir. Doğru tespit ve uygun şekilde doldurulmadıkları takdirde, analiz sonuçlarını ciddi şekilde yanıltabilirler. Bu nedenle öncelikle eksik değerlerin tespit edilmesi, ardından uygun bir yöntemle ele alınması gerekir.

1. Eksik Değerlerin Tespiti

Pandas kütüphanesi kullanılarak şu yöntemlerle eksik veriler tespit edilir:

```
import pandas as pd
```

```
# Örnek veri seti
```

```
data = {  
    'Ad': ['Ali', 'Veli', 'Ayşe', 'Fatma', None],  
    'Yaş': [25, 30, None, 22, 28],  
    'Maaş': [4000, 5000, 4500, None, None]  
}  
df = pd.DataFrame(data)
```

```
# Tüm veri setinde eksik değer var mı?
```

```
print(df.isnull().sum())
```

```
# Hangi satırlarda eksik değer var?
```

```
print(df[df.isnull().any(axis=1)])
```

2. Eksik Değerleri Doldurma Yöntemleri

Eksik değerleri doldurmak (imputation), bağlama göre dikkatle yapılmalıdır. Sık kullanılan bazı yöntemler:

a. Ortalama ile doldurma (sayısal değişkenler için)

```
df['Yaş'] = df['Yaş'].fillna(df['Yaş'].mean())
```

b. Medyan ile doldurma (aşırı uçlara duyarlılığı azaltmak için)

```
df['Maaş'] = df['Maaş'].fillna(df['Maaş'].median())
```

c. Mod (en sık görülen değer) ile doldurma (kategorik değişkenler için)

```
df['Ad'] = df['Ad'].fillna(df['Ad'].mode()[0])
```

d. İleri (forward) veya geri (backward) doldurma

```
df.fillna(method='ffill', inplace=True) # ileri doldurma
```

```
df.fillna(method='bfill', inplace=True) # geri doldurma
```

e. Tahmin (Modelleme) ile Doldurma

Daha sofistike bir yöntem olarak, eksik veriyi tahmin etmek için regresyon veya sınıflandırma modelleri kullanılabilir. Ancak bu yöntem dikkatli seçilmeli, çünkü veriye fazla anlam yükleyebilir.

Özetle:

- Eksik veriler `isnull()` veya `info()` ile tespit edilir.
- Sayısal verilerde ortalama, medyan; kategorik verilerde mod kullanılır.
- Çok eksik değer varsa, satır/sütun silmek de bir seçenek olabilir.
- Veri setinin doğasına göre yöntem seçilmelidir.

Dilersen bir modelleme ile doldurma örneğini de detaylandırabilirim.