

In [16]:

```
import pandas as pd
df = pd.read_csv('./data/scores.csv')
df.head()
```

Out[16]:

|   | name    | kor   | eng   | math  |
|---|---------|-------|-------|-------|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  |
| 1 | Charles | 90.0  | 80.0  | 75.0  |
| 2 | Danial  | 95.0  | 100.0 | 100.0 |
| 3 | Evan    | 100.0 | 100.0 | 100.0 |
| 4 | Henry   | NaN   | 35.0  | 60.0  |

## 열 추가 / 수정하기

컬럼이 존재하면 수정, 존재하지 않으면 추가한다.

```
df['컬럼명'] = 추가 / 수정할 데이터
df['컬럼명'] = 컬럼 간 연산
```



### 1.1 열 추가하기

In [3]:

```
len(df)
```

Out[3]:

30

In [17]:

```
# 학생 번호 추가하기
df['no'] = range(1, len(df) + 1, 1)
df.head()
```

Out[17]:

|   | name    | kor   | eng   | math  | no |
|---|---------|-------|-------|-------|----|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  | 1  |
| 1 | Charles | 90.0  | 80.0  | 75.0  | 2  |
| 2 | Danial  | 95.0  | 100.0 | 100.0 | 3  |
| 3 | Evan    | 100.0 | 100.0 | 100.0 | 4  |
| 4 | Henry   | NaN   | 35.0  | 60.0  | 5  |

In [18]:

```
# sum 추가
df['sum'] = df['kor'] + df['eng'] + df['math']
df.head()
```

Out[18]:

|   | name    | kor   | eng   | math  | no | sum   |
|---|---------|-------|-------|-------|----|-------|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  | 1  | 285.0 |
| 1 | Charles | 90.0  | 80.0  | 75.0  | 2  | 245.0 |
| 2 | Danial  | 95.0  | 100.0 | 100.0 | 3  | 295.0 |
| 3 | Evan    | 100.0 | 100.0 | 100.0 | 4  | 300.0 |
| 4 | Henry   | NaN   | 35.0  | 60.0  | 5  | NaN   |

## 1.2 열 수정하기

In [19]:

```
# 학생 번호 수정하기(100부터 시작하여 1씩 증가하기)
df['no'] = df['no'] + 100
df.head()
```

Out[19]:

|   | name    | kor   | eng   | math  | no  | sum   |
|---|---------|-------|-------|-------|-----|-------|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  | 101 | 285.0 |
| 1 | Charles | 90.0  | 80.0  | 75.0  | 102 | 245.0 |
| 2 | Danial  | 95.0  | 100.0 | 100.0 | 103 | 295.0 |
| 3 | Evan    | 100.0 | 100.0 | 100.0 | 104 | 300.0 |
| 4 | Henry   | NaN   | 35.0  | 60.0  | 105 | NaN   |

## 2. 열 삭제하기

```
df.drop(columns = 삭제할 컬럼 리스트, inplace = True)
```

- 존재하지 않는 **column** 명을 사용하면 에러 발생
- inplace = True** : 삭제된 결과 적용

In [20]:

```
df.drop(columns = ['no', 'sum']).head() # 여기에서는 미리보기만 해주는 것, 그러나 실제 삭제는 되지 않았다.
```

Out[20]:

|   | name    | kor   | eng   | math  |
|---|---------|-------|-------|-------|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  |
| 1 | Charles | 90.0  | 80.0  | 75.0  |
| 2 | Danial  | 95.0  | 100.0 | 100.0 |
| 3 | Evan    | 100.0 | 100.0 | 100.0 |
| 4 | Henry   | NaN   | 35.0  | 60.0  |

In [21]:

```
# 삭제된 내용 실제 데이터에 적용
df.drop(columns = ['no', 'sum'], inplace = True)
```

Out[21]:

|   | name    | kor   | eng  | math |
|---|---------|-------|------|------|
| 0 | Aiden   | 100.0 | 90.0 | 95.0 |
| 1 | Charles | 90.0  | 80.0 | 75.0 |

| 1  | Charles  | 90.0  | 80.0  | 75.0  |
|----|----------|-------|-------|-------|
| 2  | Danial   | 95.0  | 100.0 | 100.0 |
| 3  | Evan     | 100.0 | 100.0 | 100.0 |
| 4  | Henry    | NaN   | 35.0  | 60.0  |
| 5  | Ian      | 90.0  | 100.0 | 90.0  |
| 6  | James    | 70.0  | 75.0  | 65.0  |
| 7  | Julian   | 80.0  | 90.0  | 55.0  |
| 8  | Justin   | 50.0  | 60.0  | 100.0 |
| 9  | Kevin    | 100.0 | 100.0 | 90.0  |
| 10 | Leo      | 90.0  | 95.0  | 70.0  |
| 11 | Oliver   | 70.0  | 75.0  | 65.0  |
| 12 | Peter    | 100.0 | 95.0  | 100.0 |
| 13 | Amy      | 90.0  | 75.0  | 90.0  |
| 14 | Chloe    | 95.0  | 100.0 | 95.0  |
| 15 | Danna    | 100.0 | 100.0 | 100.0 |
| 16 | Ellen    | NaN   | 60.0  | NaN   |
| 17 | Emma     | 70.0  | 65.0  | 70.0  |
| 18 | Jennifer | 80.0  | 55.0  | 80.0  |
| 19 | Kate     | 50.0  | NaN   | 50.0  |
| 20 | Linda    | 100.0 | 90.0  | 100.0 |
| 21 | Olivia   | 90.0  | 70.0  | 90.0  |
| 22 | Rose     | 70.0  | 65.0  | 70.0  |
| 23 | Sofia    | 100.0 | 100.0 | 100.0 |
| 24 | Tiffany  | 90.0  | NaN   | 90.0  |
| 25 | Vanessa  | 95.0  | 70.0  | 95.0  |
| 26 | Viviana  | 100.0 | 80.0  | 100.0 |
| 27 | Vikkie   | NaN   | 50.0  | 100.0 |
| 28 | Winnie   | 70.0  | 100.0 | 70.0  |
| 29 | Zuly     | 80.0  | 90.0  | 95.0  |

### 3. 컬럼명 바꾸기

#### 3.1. 컬럼명 한번에 바꾸기

`df.columns` = 컬럼명 리스트

- 컬럼명 리스트의 항목 수는 컬럼 수와 동일해야 한다.



`df.rename(columns = [현재 컬럼명: 바꿀 컬럼명]의 쌍으로 된 딕셔너리)`

In [23]:

```
df.columns
```

Out[23]:

```
Index(['name', 'kor', 'eng', 'math'], dtype='object')
```

```
In [27]:
```

```
# '이름', '국어', '영어', '수학'
df.columns = ['이름', '국어', '영어', '수학']
df.head()

# 전체 컬럼 수와 컬럼명 리스트의 항목 수가 다르면 error
```

```
Out[27]:
```

|   | 이름      | 국어    | 영어    | 수학    |
|---|---------|-------|-------|-------|
| 0 | Aiden   | 100.0 | 90.0  | 95.0  |
| 1 | Charles | 90.0  | 80.0  | 75.0  |
| 2 | Danial  | 95.0  | 100.0 | 100.0 |
| 3 | Evan    | 100.0 | 100.0 | 100.0 |
| 4 | Henry   | NaN   | 35.0  | 60.0  |

## 3.2 특정 컬럼명 바꾸기

```
df.rename(columns = {'현재컬럼명1' : '바꿀컬럼명1', '현재컬럼명2' : '바꿀컬럼명2', , , , })
```



```
In [28]:
```

```
df.columns
```

```
Out[28]:
```

```
Index(['이름', '국어', '영어', '수학'], dtype='object')
```

```
In [30]:
```

```
# 이름 -> 성명
df.rename(columns = {'이름' : '성명'}, inplace = True)
```

```
In [31]:
```

```
df.columns
```

```
Out[31]:
```

```
Index(['성명', '국어', '영어', '수학'], dtype='object')
```