In [1]:
```
import pandas as pd

# Membuat DataFrame kosong
df = pd.DataFrame()

# Menampilkan DataFrame
print(df)
```

```
Empty DataFrame
Columns: []
Index: []
```

In [2]:
```
# Membuat DataFrame dengan data
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)

# Menampilkan DataFrame
print(df)
```

```
   Nama  Umur
0  Andi    20
1  Budi    25
2  Cici    30
```

In [3]:
```
# Membuat DataFrame dengan data
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)

# Menampilkan tipe data kolom 'Umur'
print(df['Umur'].dtype)
```

```
int64
```

In [4]:
```
# Membaca data dari file CSV
df = pd.read_csv('Heart_Disease_Prediction.csv')

# Menampilkan DataFrame
print(df)
```

```
     Age  Sex  Chest pain type   BP  Cholesterol  FBS over 120  EKG results  \
0     70    1                4  130          322             0            2
1     67    0                3  115          564             0            2
2     57    1                2  124          261             0            0
3     64    1                4  128          263             0            0
4     74    0                2  120          269             0            2
..   ...  ...              ...  ...          ...           ...          ...
265   52    1                3  172          199             1            0
266   44    1                2  120          263             0            0
267   56    0                2  140          294             0            2
268   57    1                4  140          192             0            0
269   67    1                4  160          286             0            2

     Max HR  Exercise angina  ST depression  Slope of ST  \
0       109                0            2.4            2
1       160                0            1.6            2
2       141                0            0.3            1
3       105                1            0.2            2
4       121                1            0.2            1
```

```
..      ...                ...            ...           ...
265     162                  0            0.5             1
266     173                  0            0.0             1
267     153                  0            1.3             2
268     148                  0            0.4             2
269     108                  1            1.5             2

     Number of vessels fluro  Thallium Heart Disease
0                          3         3      Presence
1                          0         7       Absence
2                          0         7      Presence
3                          1         7       Absence
4                          1         3       Absence
..                       ...       ...           ...
265                        0         7       Absence
266                        0         7       Absence
267                        0         3       Absence
268                        0         6       Absence
269                        3         3      Presence

[270 rows x 14 columns]
```

In [5]:
```
df.describe()
```

Out[5]:

	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max H
count	270.000000	270.000000	270.000000	270.000000	270.000000	270.000000	270.000000	270.0000
mean	54.433333	0.677778	3.174074	131.344444	249.659259	0.148148	1.022222	149.6777
std	9.109067	0.468195	0.950090	17.861608	51.686237	0.355906	0.997891	23.1657
min	29.000000	0.000000	1.000000	94.000000	126.000000	0.000000	0.000000	71.0000
25%	48.000000	0.000000	3.000000	120.000000	213.000000	0.000000	0.000000	133.0000
50%	55.000000	1.000000	3.000000	130.000000	245.000000	0.000000	2.000000	153.5000
75%	61.000000	1.000000	4.000000	140.000000	280.000000	0.000000	2.000000	166.0000
max	77.000000	1.000000	4.000000	200.000000	564.000000	1.000000	2.000000	202.0000

In [6]:
```
df.sort_values(by=['EKG results']).head(3)
```

Out[6]:

	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	Number of vessels fluro	T
178	67	1	4	125	254	1	0	163	0	0.2	2	2	
100	44	0	3	108	141	0	0	175	0	0.6	2	0	
99	50	0	2	120	244	0	0	162	0	1.1	1	0	

In [7]:
```
df[['Age', 'EKG results']].groupby(df['Sex']).max()
```

Out[7]:

Sex	Age	EKG results
0	76	2
1	77	2

In [8]:

```
# Membuat DataFrame dengan data
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)

# Menghapus kolom 'Umur'
df = df.drop('Umur', axis=1)

# Menampilkan DataFrame
print(df)
```

```
   Nama
0  Andi
1  Budi
2  Cici
```

In [9]:

```
# Membuat DataFrame dengan data
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)

# Menghitung rata-rata umur
rata_rata_umur = df['Umur'].mean()

# Menampilkan rata-rata umur
print(rata_rata_umur)
```

```
25.0
```

In [10]:

```
data1 = {'Nama': ['Andi', 'Budi'],
        'Umur': [20, 25]}
df1 = pd.DataFrame(data1)

data2 = {'Nama': ['Cici', 'Dedi'],
        'Umur': [30, 35]}
df2 = pd.DataFrame(data2)

# Menggabungkan DataFrame
df = pd.concat([df1, df2])

# Menampilkan DataFrame
print(df)
```

```
   Nama  Umur
0  Andi    20
1  Budi    25
0  Cici    30
1  Dedi    35
```

In [11]:

```
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)
```

```
df['Umur'] = df['Umur'] + 5
print(df)
```

```
   Nama  Umur
0  Andi    25
1  Budi    30
2  Cici    35
```

In [12]:

```
data = {'Nama': ['Andi', 'Budi', 'Cici'],
        'Umur': [20, 25, 30]}
df = pd.DataFrame(data)
df.plot(
    x='Umur',
    y='Nama',
    kind='scatter'
);
```