

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

```
In [2]: employees = pd.read_excel('employee_data.xlsx')
```

```
In [3]: employees.head(10)
```

```
Out[3]:
```

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary |
|----------|----------|--------|-----|------------------|-----------------|-----------|---------------|
| 0 | 1 | 1 | 39 | 5 | 12 | 4 | 57700 |
| 1 | 2 | 0 | 44 | 12 | 8 | 6 | 76400 |
| 2 | 3 | 0 | 24 | 0 | 2 | 4 | 44000 |
| 3 | 4 | 1 | 25 | 2 | 1 | 4 | 41600 |
| 4 | 5 | 0 | 56 | 5 | 25 | 8 | 163900 |
| 5 | 6 | 1 | 41 | 9 | 10 | 4 | 72700 |
| 6 | 7 | 1 | 33 | 6 | 2 | 6 | 60300 |
| 7 | 8 | 0 | 37 | 11 | 6 | 4 | 63500 |
| 8 | 9 | 1 | 51 | 12 | 16 | 6 | 131200 |
| 9 | 10 | 0 | 23 | 0 | 1 | 4 | 39200 |

```
In [4]: emp_sorted = employees.sort_values('Age', ascending=False)
emp_sorted.head()
```

```
Out[4]:
```

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary |
|------------|----------|--------|-----|------------------|-----------------|-----------|---------------|
| 103 | 104 | 1 | 65 | 4 | 9 | 4 | 57800 |
| 93 | 94 | 1 | 64 | 5 | 7 | 4 | 55700 |
| 21 | 22 | 0 | 63 | 16 | 20 | 4 | 140400 |
| 101 | 102 | 0 | 61 | 9 | 15 | 6 | 109100 |
| 77 | 78 | 1 | 61 | 0 | 7 | 4 | 40500 |

```
In [5]: emp_sorted = emp_sorted.reset_index(drop=True)
emp_sorted.head()
```

```
Out[5]:
```

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary |
|----------|----------|--------|-----|------------------|-----------------|-----------|---------------|
| 0 | 104 | 1 | 65 | 4 | 9 | 4 | 57800 |
| 1 | 94 | 1 | 64 | 5 | 7 | 4 | 55700 |
| 2 | 22 | 0 | 63 | 16 | 20 | 4 | 140400 |
| 3 | 102 | 0 | 61 | 9 | 15 | 6 | 109100 |

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary |
|---|----------|--------|-----|------------------|-----------------|-----------|---------------|
| 4 | 78 | 1 | 61 | 0 | 7 | 4 | 40500 |

```
In [7]: emp_female = employees[employees['Gender']==1].reset_index(drop=True)
```

```
In [8]: emp_female.head()
```

```
Out[8]:
```

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary |
|---|----------|--------|-----|------------------|-----------------|-----------|---------------|
| 0 | 1 | 1 | 39 | 5 | 12 | 4 | 57700 |
| 1 | 4 | 1 | 25 | 2 | 1 | 4 | 41600 |
| 2 | 6 | 1 | 41 | 9 | 10 | 4 | 72700 |
| 3 | 7 | 1 | 33 | 6 | 2 | 6 | 60300 |
| 4 | 9 | 1 | 51 | 12 | 16 | 6 | 131200 |

```
In [9]: emp_female.shape
```

```
Out[9]: (119, 7)
```

```
In [10]: emp_female['Age'].dtypes
```

```
Out[10]: dtype('int64')
```

```
In [11]: emp_female['Age'].mean()
```

```
Out[11]: 40.319327731092436
```

```
In [12]: fem_mean = emp_female['Age'].mean()  
round(fem_mean,2)
```

```
Out[12]: 40.32
```

```
In [13]: emp_female['Annual Salary'].min()
```

```
Out[13]: 12400
```

```
In [14]: emp_female['Age'].max()
```

```
Out[14]: 65
```

```
In [15]: emp_female['Age'].median()
```

Out[15]: 42.0

In [16]: `emp_female['Total Experience'] = emp_female['Prior Experience']+emp_female['Beta Experi`

In [17]: `emp_female.head()`

Out[17]:

| | Employee | Gender | Age | Prior Experience | Beta Experience | Education | Annual Salary | Total Experience |
|---|----------|--------|-----|------------------|-----------------|-----------|---------------|------------------|
| 0 | 1 | 1 | 39 | 5 | 12 | 4 | 57700 | 17 |
| 1 | 4 | 1 | 25 | 2 | 1 | 4 | 41600 | 3 |
| 2 | 6 | 1 | 41 | 9 | 10 | 4 | 72700 | 19 |
| 3 | 7 | 1 | 33 | 6 | 2 | 6 | 60300 | 8 |
| 4 | 9 | 1 | 51 | 12 | 16 | 6 | 131200 | 28 |

In []: `emp_female.to_excel('emp_female.xlsx')`

In []: