In [3]:
```
import pandas as pd
data=pd.read_csv('adult.csv',sep=",",names=['age','work','fnlwgt','education','edu-num','mar','occup','rela','race','sex','cg','cl','hou','native','
data
```

Out[3]:

	age	work	fnlwgt	education	edu-num	mar	occup	rela	race	sex	cg	cl	hou	native	sal
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=50K
1	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=50K
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=50K
3	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=50K
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	<=50K
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
32556	27	Private	257302	Assoc-acdm	12	Married-civ-spouse	Tech-support	Wife	White	Female	0	0	38	United-States	<=50K
32557	40	Private	154374	HS-grad	9	Married-civ-spouse	Machine-op-inspct	Husband	White	Male	0	0	40	United-States	>50K
32558	58	Private	151910	HS-grad	9	Widowed	Adm-clerical	Unmarried	White	Female	0	0	40	United-States	<=50K
32559	22	Private	201490	HS-grad	9	Never-married	Adm-clerical	Own-child	White	Male	0	0	20	United-States	<=50K
32560	52	Self-emp-inc	287927	HS-grad	9	Married-civ-spouse	Exec-managerial	Wife	White	Female	15024	0	40	United-States	>50K

32561 rows × 15 columns

In [4]:
```
data['sex'].value_counts()
```

Out[4]:
```
 Male      21790
 Female    10771
Name: sex, dtype: int64
```

In [21]:
```
data[data['sex'].str.contains('Female')]['age'].mean()
```

Out[21]:
```
36.85823043357163
```

In [28]:
```
data[data['native'].str.contains('Germany')]['native'].value_counts()/len(data)*100
```

Out[28]:
```
 Germany    0.420749
Name: native, dtype: float64
```

In [23]:
```
data[data['sal'].str.contains('>50K')]['age'].mean()
```

Out[23]:
```
44.24984058155847
```

In [24]:
```
data[data['sal'].str.contains('>50K')]['age'].std()
```

Out[24]:
```
10.51902771985177
```

In [27]:
```
data[data['sal'].str.contains('<=50K')]['age'].mean()
```

Out[27]:
```
36.78373786407767
```

In [25]:
```
data[data['sal'].str.contains('<=50K')]['age'].std()
```

Out[25]:
```
14.020088490824813
```

In [29]:
```
data[data['sal'].str.contains('>50K')]['education'].unique()
```

Out[29]:
```
array([' HS-grad', ' Masters', ' Bachelors', ' Some-college',
       ' Assoc-voc', ' Doctorate', ' Prof-school', ' Assoc-acdm',
       ' 7th-8th', ' 12th', ' 10th', ' 11th', ' 9th', ' 5th-6th',
       ' 1st-4th'], dtype=object)
```

In [ ]:
```

```