In [7]:

```
import pandas as pd
import numpy as np
```

In [13]:

```
df = pd.read_csv('Data.csv')
```

In [14]:

```
df
```

Out[14]:

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	No
1	Spain	27.0	48000.0	Yes
2	Germany	30.0	54000.0	No
3	Spain	38.0	61000.0	No
4	Germany	40.0	NaN	Yes
5	France	35.0	58000.0	Yes
6	Spain	NaN	52000.0	No
7	France	48.0	79000.0	Yes
8	Germany	50.0	83000.0	No
9	France	37.0	67000.0	Yes

In [16]:

```
df.describe()
```

Out[16]:

	Age	Salary
count	9.000000	9.000000
mean	38.777778	63777.777778
std	7.693793	12265.579662
min	27.000000	48000.000000
25%	35.000000	54000.000000
50%	38.000000	61000.000000
75%	44.000000	72000.000000
max	50.000000	83000.000000

In [21]:

```
df.shape
```

Out[21]:

```
(10, 4)
```

In [22]:

```
df.size
```

Out[22]:

```
40
```

In [23]:

```
df.count()
```

Out[23]:

```
Country      10
Age           9
Salary        9
Purchased    10
dtype: int64
```

In [26]:

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

Out[26]:

```
Country      0
Age          1
Salary       1
Purchased    0
dtype: int64
```

In [27]:

```
df.Salary.values
```

Out[27]:

```
array([72000., 48000., 54000., 61000.,    nan, 58000., 52000., 79000.,
       83000., 67000.])
```

In [28]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 4 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Country    10 non-null     object
 1   Age        9 non-null      float64
 2   Salary     9 non-null      float64
 3   Purchased  10 non-null     object
dtypes: float64(2), object(2)
memory usage: 448.0+ bytes
```

In [31]:

```
df.ndim
```

Out[31]:

```
2
```

In [33]:

```
df.Salary.sum()
```

Out[33]:

```
574000.0
```

In [35]:

```
df.dtypes
```

Out[35]:

```
Country       object
Age          float64
Salary       float64
Purchased     object
dtype: object
```

In [36]:

```
df.replace({'Purchased':{'Yes':1,'No':0}})
```

Out[36]:

	Country	Age	Salary	Purchased
0	France	44.0	72000.0	0
1	Spain	27.0	48000.0	1
2	Germany	30.0	54000.0	0
3	Spain	38.0	61000.0	0
4	Germany	40.0	NaN	1
5	France	35.0	58000.0	1
6	Spain	NaN	52000.0	0
7	France	48.0	79000.0	1
8	Germany	50.0	83000.0	0
9	France	37.0	67000.0	1

In [38]:

```
df['Salary'].value_counts()
```

Out[38]:

```
72000.0    1
48000.0    1
54000.0    1
61000.0    1
58000.0    1
52000.0    1
79000.0    1
83000.0    1
67000.0    1
Name: Salary, dtype: int64
```

In [39]:

```
df.replace({'Age':int()},inplace=True)
```

In [41]:

```
df.dtypes
```

Out[41]:

```
Country       object
Age          float64
Salary       float64
Purchased     object
dtype: object
```

In [ ]: