In [7]:
```
import numpy as np
import pandas as pd
```

In [8]:
```
df = pd.read_csv("Downloads/HR-Employee-Attrition.csv")
```

In [9]:
```
df.head()
```

Out[9]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	Emp
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences	
1	49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences	
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other	
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences	
4	27	No	Travel_Rarely	591	Research & Development	2	1	Medical	

5 rows × 35 columns

In [10]:
```
#Mean of monthly income and age
print("The mean of monthly income is :",df.loc[:,"MonthlyIncome"].mean())
print("The mean of age is :",df.loc[:,"Age"].mean())
```

```
The mean of monthly income is : 6502.931292517007
The mean of age is : 36.923809523809524
```

In [11]:
```
#Mode of monthly income and age
print("The median of monthly income is :",df.loc[:,"MonthlyIncome"].median())
print("The median of age is :",df.loc[:,"Age"].median())
```

```
The median of monthly income is : 4919.0
The median of age is : 36.0
```

In [12]:
```
#Median of monthly income and age
print("The mode of monthly income is :",df.loc[:,"MonthlyIncome"].mode())
print("The mode of age is :",df.loc[:,"Age"].mode())
```

```
The mode of monthly income is : 0    2342
Name: MonthlyIncome, dtype: int64
The mode of age is : 0    35
Name: Age, dtype: int64
```

In [13]:
```
#Standard deviation of monthly income and age
print("The standard deviation of monthly income is
:",df.loc[:,"MonthlyIncome"].std())
print("The standard deviation of age is :",df.loc[:,"Age"].std())
```

```
The standard deviation of monthly income is : 4707.956783097995
The standard deviation of age is : 9.135373489136734
```

In [14]:

```
#Storing age and monthly income in array and then finding maximum and minimum values
array1 = np.array(df['MonthlyIncome'])
array2=np.array(df["Age"])
print("Income",array1)
print("Age array",array2)
print("Maximum income among the employees is :",max(array1))
print("Minimum income among the employees is :",min(array1))
print("Maximum age among the employees is :",max(array2))
print("Minimum age among the employees is :",min(array2))
```

```
Income [5993 5130 2090 ... 6142 5390 4404]
Age array [41 49 37 ... 27 49 34]
Maximum income among the employees is : 19999
Minimum income among the employees is : 1009
Maximum age among the employees is : 60
Minimum age among the employees is : 18
```

In [15]:

```
# Replacing the categorical values by numeric values
df.head()
df["BusinessTravel"].replace({"Travel_Rarely":1, "Travel_Frequently":0},
inplace=True)
df["Attrition"].replace({ "Yes":1, "No":0}, inplace=True)
df.head()
```

Out[15]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	Emplc
0	41	1	1	1102	Sales	1	2	Life Sciences	
1	49	0	0	279	Research & Development	8	1	Life Sciences	
2	37	1	1	1373	Research & Development	2	2	Other	
3	33	0	0	1392	Research & Development	3	4	Life Sciences	
4	27	0	1	591	Research & Development	2	1	Medical	

5 rows × 35 columns

In [16]:

```
df.describe()
```

Out[16]:

	Age	Attrition	DailyRate	DistanceFromHome	Education	EmployeeCount	EmployeeNumb
count	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000	1470.0	1470.0000
mean	36.923810	0.161224	802.485714	9.192517	2.912925	1.0	1024.8653
std	9.135373	0.367863	403.509100	8.106864	1.024165	0.0	602.0243
min	18.000000	0.000000	102.000000	1.000000	1.000000	1.0	1.0000
25%	30.000000	0.000000	465.000000	2.000000	2.000000	1.0	491.2500
50%	36.000000	0.000000	802.000000	7.000000	3.000000	1.0	1020.5000
75%	43.000000	0.000000	1157.000000	14.000000	4.000000	1.0	1555.7500
max	60.000000	1.000000	1499.000000	29.000000	5.000000	1.0	2068.0000

8 rows × 27 columns

In [ ]: