In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
dataset1=pd.read_csv("general_data.csv")
```

Out[1]:

	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EducationField	EmployeeCount
0	51	No	Travel_Rarely	Sales	6	2	Life Sciences	1
1	31	Yes	Travel_Frequently	Research & Development	10	1	Life Sciences	1
2	32	No	Travel_Frequently	Research & Development	17	4	Other	1
3	38	No	Non-Travel	Research & Development	2	5	Life Sciences	1
4	32	No	Travel_Rarely	Research & Development	10	1	Medical	1

5 rows × 24 columns

In [2]:
```
attrition_ds= dataset1.loc[dataset1['Attrition'] == 'Yes']
```

Out[2]:

	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EducationField	EmployeeCou
1	31	Yes	Travel_Frequently	Research & Development	10	1	Life Sciences	
6	28	Yes	Travel_Rarely	Research & Development	11	2	Medical	
13	47	Yes	Non-Travel	Research & Development	1	1	Medical	
28	44	Yes	Travel_Frequently	Research & Development	1	2	Medical	
30	26	Yes	Travel_Rarely	Research & Development	4	3	Medical	
...	...	...	...	...	...	...	...	
4381	29	Yes	Travel_Rarely	Research & Development	7	1	Life Sciences	
4386	33	Yes	Travel_Rarely	Sales	11	4	Marketing	
4388	33	Yes	Travel_Rarely	Sales	1	3	Life Sciences	
4391	32	Yes	Travel_Rarely	Sales	23	1	Life Sciences	
4402	37	Yes	Travel_Frequently	Sales	2	3	Marketing	

711 rows × 24 columns

In [3]:

Out[3]:
```
count    711.000000
mean       9.012658
std        7.772368
min        1.000000
25%        2.000000
50%        7.000000
75%       15.000000
max       29.000000
Name: DistanceFromHome, dtype: float64
```

Hypotheis1

For Sample of 711 employeess mean Distance is found to be 9.01

In [4]:

Out[4]:
```
count       711.000000
mean      61682.616034
std       44792.067695
min       10090.000000
25%       28440.000000
50%       49080.000000
75%       71040.000000
max      198590.000000
Name: MonthlyIncome, dtype: float64
```

Hypotheis2

For Sample of 711 employeess mean monthly incomeis found to be 61682.61

In [7]:

Out[7]:
```
count    711.000000
mean      15.481013
std        3.775289
min       11.000000
25%       12.000000
50%       14.000000
75%       18.000000
max       25.000000
Name: PercentSalaryHike, dtype: float64
```

Hypotheis3

For Sample of 711 employeess mean of employeess who leave the company is found to be 15.4

In [10]:

Out[10]:
```
count    709.000000
mean       8.255289
std        7.164018
min        0.000000
25%        3.000000
50%        7.000000
75%       10.000000
max       40.000000
Name: TotalWorkingYears, dtype: float64
```

Hypotheis4

For Sample of 709 employeess mean of experience of employeess who leave the company is found to be 8.2

In [11]:

Out[11]:

```
count    711.000000
```

Hypotheis5

For Sample of 711 employeess mean of trainig houes is found to be 2.65

In [14]:

Out[14]:
```
count    711.000000
mean       5.130802
std        5.941598
min        0.000000
25%        1.000000
50%        3.000000
75%        7.000000
max       40.000000
Name: YearsAtCompany, dtype: float64
```

Hypotheis6

For Sample of 711 employeess mean experince spent in the company is 5.13

In [16]:

```
count    711.000000
mean       1.945148
std        3.148633
min        0.000000
25%        0.000000
50%        1.000000
75%        2.000000
max       15.000000
Name: YearsSinceLastPromotion, dtype: float64
```

Hypotheis7

For Sample of 711 employeess mean of YearsSinceLastPromotion is 3.14

In [17]:

```
count    711.000000
mean       2.852321
std        3.138918
min        0.000000
25%        0.000000
50%        2.000000
75%        5.000000
max       14.000000
Name: YearsWithCurrManager, dtype: float64
```

Hypotheis8

For Sample of 711 employeess mean of YearsWithCurrManager is 2.85

In [18]:

```
count    711.000000
mean      33.607595
std        9.675693
min       18.000000
25%       28.000000
50%       32.000000
75%       39.000000
max       58.000000
Name: Age, dtype: float64
```

Hypotheis9

For Sample of 711 employeess mean of YearsWithCurrManager is 33.60

In [20]:

```
count    707.000000
mean       2.936351
std        2.678774
min        0.000000
25%        1.000000
50%        1.000000
75%        5.000000
max        9.000000
Name: NumCompaniesWorked, dtype: float64
```

Hypotheis10

For Sample of 711 employeess mean of NumCompaniesWorked is 2.93