```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
df = pd.read_csv('/content/HR.csv')
df.head()
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Educatic
0	41	Yes	Travel_Rarely	1102	Sales	1	
1	49	No	Travel_Frequently	279	Research & Development	8	
2	37	Yes	Travel_Rarely	1373	Research & Development	2	
3	33	No	Travel_Frequently	1392	Research & Development	3	
4	27	No	Travel_Rarely	591	Research & Development	2	

5 rows × 35 columns

```
df.columns
```

```
Index(['Age', 'Attrition', 'BusinessTravel', 'DailyRate', 'Department',
       'DistanceFromHome', 'Education', 'EducationField', 'EmployeeCount',
       'EmployeeNumber', 'EnvironmentSatisfaction', 'Gender', 'HourlyRate',
       'JobInvolvement', 'JobLevel', 'JobRole', 'JobSatisfaction',
       'MaritalStatus', 'MonthlyIncome', 'MonthlyRate', 'NumCompaniesWorked',
       'Over18', 'OverTime', 'PercentSalaryHike', 'PerformanceRating',
       'RelationshipSatisfaction', 'StandardHours', 'StockOptionLevel',
       'TotalWorkingYears', 'TrainingTimesLastYear', 'WorkLifeBalance',
       'YearsAtCompany', 'YearsInCurrentRole', 'YearsSinceLastPromotion',
       'YearsWithCurrManager'],
      dtype='object')
```

```
#mean of monthly income
df.loc[:,"MonthlyIncome"].mean()
```

```
6502.931292517007
```

```
# mean of age
df.loc[: , "Age"].mean()
```

```
36.923809523809524
```

```
#Meadian of monthly  income
df.loc[: , "MonthlyIncome"].median()
```

```
4919.0
```

```
#Meadian of age
df.loc[: , "Age"].median()
```

```
36.0
```

```
#Mode of monthly income
df.loc[: , "MonthlyIncome"].mode()
```

```
0    2342
Name: MonthlyIncome, dtype: int64
```

```
#mode of age
df.loc[: , "Age"].mode()
```

```
0    35
Name: Age, dtype: int64
```

```
#STD of monthly  income
df.loc[: , "MonthlyIncome"].std()
```

```
4707.956783097994
```

```
#STD of age
df.loc[: , "Age"].std()
```

```
9.135373489136732
```

```
arr1 = np.array(df["MonthlyIncome"])
arr2 = np.array(df['Age'])
```

```
print("Income " , arr1)
print("Age" , arr2)
```

```
Income  [5993 5130 2090 ... 6142 5390 4404]
Age [41 49 37 ... 27 49 34]
```

```
#MAX income
print(max(arr1))

#min income
print(min(arr1))
```

```
19999
1009
```

```
#MAX age
print(max(arr2))

#min age
print(min(arr2))
```

```
60
18
```

```
df["BusinessTravel"].replace({"Travel_Rarely" : 1 , "Travel_Frequently" : 0} , inplace = T
df["Attrition"].replace({"Yes" : 1 , "No" : 0} , inplace = True)
df.head()
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Educatio
0	41	1	1	1102	Sales	1	
1	49	0	0	279	Research & Development	8	
2	37	1	1	1373	Research & Development	2	
3	33	0	0	1392	Research & Development	3	
4	27	0	1	591	Research & Development	2	

5 rows × 35 columns