In [1]:
```
#Experiment no.7
```

In [2]:
```
#Aim : To perform simple Linear regression
```

In [3]:
```
#Name:Srushti Bawane
#Roll no.:5
#sec:B
#sub:ET 1
#date:23-09-2025
```

In [4]:
```
#importing the basic library
import pandas as pd
```

In [5]:
```
import os
```

In [6]:
```
os.getcwd()
```

Out[6]:
```
'C:\\Users\\RH'
```

In [18]:
```
os.chdir('C:\\Users\\RH\\Downloads')
```

In [19]:
```
data=pd.read_csv("Salary_Data.csv")
```

In [20]:
```
data.head()
```

Out[20]:

	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2.0	43525
4	2.2	39891

In [21]:
```
data.shape
```

Out[21]:
```
(30, 2)
```

In [22]:
```
data.size
```

Out[22]:
```
60
```

In [23]:
```
data.ndim
```

Out[23]:
```
2
```

In [24]:
```
data.info
```

Out[24]:
```
<bound method DataFrame.info of     YearsExperience  Salary
0               1.1   39343
1               1.3   46205
2               1.5   37731
3               2.0   43525
4               2.2   39891
5               2.9   56642
6               3.0   60150
7               3.2   54445
8               3.2   64445
9               3.7   57189
10              3.9   63218
11              4.0   55794
12              4.0   56957
13              4.1   57081
14              4.5   61111
15              4.9   67938
16              5.1   66029
17              5.3   83088
18              5.9   81363
19              6.0   93940
20              6.8   91738
21              7.1   98273
22              7.9  101302
23              8.2  113812
24              8.7  109431
25              9.0  105582
26              9.5  116969
27              9.6  112635
28             10.3  122391
29             10.5  121872>
```

In [25]:
```
data.columns
```

Out[25]:
```
Index(['YearsExperience', 'Salary'], dtype='object')
```

In [26]:
```
data.describe()
```

Out[26]:

	YearsExperience	Salary
count	30.000000	30.000000
mean	5.313333	76003.000000
std	2.837888	27414.429785
min	1.100000	37731.000000
25%	3.200000	56720.750000
50%	4.700000	65237.000000
75%	7.700000	100544.750000
max	10.500000	122391.000000

In [27]:
```
data.isnull().sum()
```

Out[27]:
```
YearsExperience    0
Salary             0
dtype: int64
```

Independent and Dependent Variable

In [28]:
```
x=data.drop('Salary',axis=1)
```

In [29]:
```
x.head()
```

Out[29]:

	YearsExperience
0	1.1
1	1.3
2	1.5
3	2.0
4	2.2

In [30]:
```
y=data.Salary
```

In [31]:
```
y.head()
```

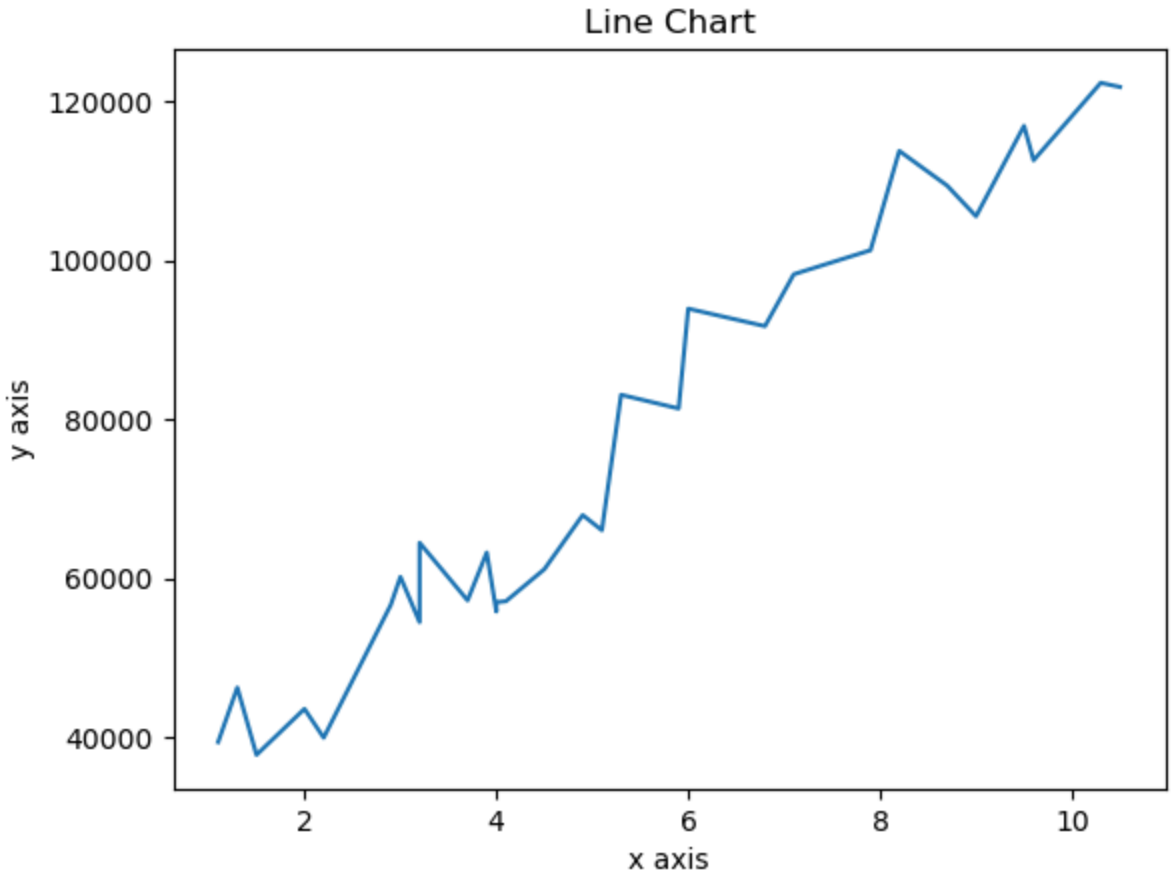
Out[31]:
```
0    39343
1    46205
2    37731
3    43525
4    39891
Name: Salary, dtype: int64
```

Line chart

In [32]:
```
#import  library
import numpy as np
from matplotlib import pyplot as plt
```

In [33]:
```
plt.plot(x,y)
plt.title("Line Chart")
plt.xlabel("x axis")

plt.ylabel("y axis")
plt.show()
```


Line Chart
120000
100000
80000
60000
40000
y axis
2
4
6
8
10
x axis

In [ ]: