In [ ]:
```
#EXP 7
```

In [ ]:
```
#Aim: Linear regression
```

In [1]:
```
# Name:Anisha Yogendra Mahajan
# Roll no.: 34
# Sec:A
# Subject:ET1
# Date: 22/09/2025
```

In [3]:
```
import pandas as pd
```

In [4]:
```
import os
```

In [5]:
```
os.getcwd()
```

Out[5]:
```
'C:\\Users\\USER'
```

In [6]:
```
os.chdir("C:\\Users\\USER\\Desktop")
```

In [13]:
```
data=pd.read_csv("Salary_Data - Salary_Data.csv")
```

In [14]:
```
data.head(20)
```

Out[14]:

	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

	YearsExperience	Salary
15	4.9	67938
16	5.1	66029
17	5.3	83088
18	5.9	81363
19	6.0	93940

In [15]:
```
data.tail()
```

Out[15]:

	YearsExperience	Salary
25	9.0	105582
26	9.5	116969
27	9.6	112635
28	10.3	122391
29	10.5	121872

In [16]:
```
data.describe()
```

Out[16]:

	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 [17]:
```
data.shape
```

Out[17]:
```
(30, 2)
```

In [18]:
```
data.size
```

Out[18]:
```
60
```

In [19]:
```
data.ndim
```

Out[19]:
```
2
```

In [20]:
```
data.columns
```

Out[20]:
```
Index(['YearsExperience', 'Salary'], dtype='object')
```

In [21]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   YearsExperience  30 non-null     float64
 1   Salary           30 non-null     int64  
dtypes: float64(1), int64(1)
memory usage: 608.0 bytes
```

In [23]:
```
data.isnull().sum()
```

Out[23]:
```
YearsExperience    0
Salary             0
dtype: int64
```

Independent and dependent variables

In [25]:
```
X=data.drop('Salary',axis=1)
```

In [26]:
```
X.head()
```

Out[26]:

	YearsExperience
0	1.1
1	1.3
2	1.5
3	2.0
4	2.2

In [27]:
```
y=data.Salary
```

In [28]:
```
y.head()
```

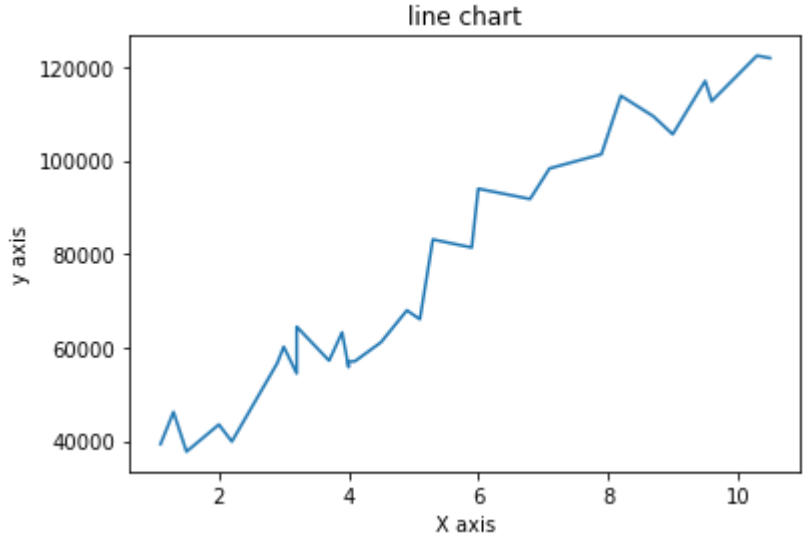
Out[28]:
```
0    39343
1    46205
2    37731
3    43525
4    39891
Name: Salary, dtype: int64
```

line chart

In [32]:
```
from matplotlib import pyplot as plt
```

In [35]:
```
plt.plot(X,y)
plt.title("line chart")
plt.xlabel("X axis")

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



In [ ]:
```

```