In [1]:

```
import pandas as pd
import numpy as np
```

In [2]:

```
df=pd.read_csv('student_marks.csv')
```

In [5]:

```
df
```

```
...
```

In [6]:

```
df['Name']
```

```
...
```

In [7]:

```
df[['Name','Gender']]
```

```
...
```

In [8]:

```
# iloc and loc
df.loc[2:5]
```

```
...
```

In [9]:

```
df.loc[2:5,2:3]
```

```
...
```

In [10]:

```
df.loc[2:5,'Name']
```

```
...
```

In [11]:

```
df.loc[2:5,['Name','Gender']]
```

```
...
```

In [12]:

```
names=df.loc[2:5,['Name','Gender']]
```

In [13]:

```
names
```

```
...
```

In [18]:

```
df.loc[:,'Physics':'Biology':2]
```

...

In [20]:

```
df.loc[df['Gender']=='M']
```

...

In [21]:

```
df.loc[df['Gender']=='M','Name']
```

...

In [22]:

```
males=df.loc[df['Gender']=='M','Name']
```

In [23]:

```
males
```

Out[23]:

```
0      John
1    Suresh
2    Ramesh
7    Ritesh
9    Mukesh
Name: Name, dtype: object
```

In [24]:

```
df.loc[df['Gender']=='M',['Gender','Name']]
```

...

In [26]:

```
df.loc[~(df['Gender']=='M')]
```

...

In [27]:

```
#iloc
df.iloc[2:5]
```

Out[27]:

	Name	Gender	DOB	Maths	Physics	Chemistry	English	Biology	Economics	History
2	Ramesh	M	25-05-1989	25	54	89	76	95	87	56
3	Jessica	F	12-08-1990	78	96	86	63	54	89	75
4	Jennifer	F	02-09-1989	58	96	78	46	96	77	83

In [29]:

```
df.iloc[2:5,0]
```

Out[29]:

```
2      Ramesh
3     Jessica
4    Jennifer
Name: Name, dtype: object
```

In [30]:

```
df.iloc[:,2:5]
```

...

In [31]:

```
df.iloc[:,[0,7,10]]
```

...

In [32]:

```
cl=df.iloc[:,[0,7,10]]
```

In [33]:

```
cl
```

Out[33]:

	Name	Biology	Civics
0	John	21	65
1	Suresh	90	2
2	Ramesh	95	74
3	Jessica	54	45
4	Jennifer	96	53
5	Annu	55	52
6	pooja	75	61
7	Ritesh	25	87
8	Farha	78	89
9	Mukesh	58	77

In [34]:

```
type(cl)
```

Out[34]:

```
pandas.core.frame.DataFrame
```

In [5]:

```
df
```

Out[5]:

	Name	Gender	DOB	Maths	Physics	Chemistry	English	Biology	Economics	History
0	John	M	05-04-1988	55	45	56	87	21	52	89
1	Suresh	M	04-05-1987	75	96	78	64	90	61	58
2	Ramesh	M	25-05-1989	25	54	89	76	95	87	56
3	Jessica	F	12-08-1990	78	96	86	63	54	89	75
4	Jennifer	F	02-09-1989	58	96	78	46	96	77	83
5	Annu	F	05-04-1988	45	87	52	89	55	89	87
6	pooja	F	04-05-1987	55	64	61	58	75	58	64
7	Ritesh	M	25-05-1989	54	76	87	56	25	56	76
8	Farha	F	12-08-1990	55	63	89	75	78	75	63
9	Mukesh	M	02-09-1989	96	46	77	83	58	83	46

In [3]:

```
df['sci_total']=df['Maths']+df['Physics']+df['Chemistry']+df['Biology']
```

In [4]:

```
df['sci_total']
```

Out[4]:

```
0    177
1    339
2    263
3    314
4    328
5    239
6    255
7    242
8    285
9    277
Name: sci_total, dtype: int64
```

In [42]:

```
#concat
data1={'Name':['Jay','Raj','Kumar'],
      'Age':[20,25,23],
      'dept':['Hr','Sales','Account']}
df1=pd.DataFrame(data1)
```

In [51]:

```
data2={'deptid':list(range(0,3)),
      'count':list(range(10,40,10))}
df2=pd.DataFrame(data2,dtype=int)
```

In [48]:

```
newdf=pd.concat([df1,df2])
```

In [52]:

```
newdf
```

...

In [53]:

```
newdf=pd.concat([df1,df2],ignore_index=True)
newdf
```

...

In [55]:

```
newdf.isna().sum()
```

...

In [57]:

```
newdf.notna().sum()
```

...

In [59]:

```
newdf['Name'].replace(np.NaN,'abc')
```

...

In [60]:

```
newdf
```

...

In [61]:

```
newdf['Name'].replace(np.NaN,'abc',inplace=True)
```

In [62]:

```
newdf
```

...

In [65]:

```
dmean=np.mean(newdf['deptid'])
```

In [66]:

```
dmean
```

Out[66]:

```
1.0
```

In [67]:

```
newdf['deptid'].replace(np.NaN,dmean,inplace=True)
```

In [68]:

```
newdf
```

Out[68]:

	Name	Age	dept	deptid	count
0	Jay	20.0	Hr	1.0	NaN
1	Raj	25.0	Sales	1.0	NaN
2	Kumar	23.0	Account	1.0	NaN
3	abc	NaN	NaN	0.0	10.0
4	abc	NaN	NaN	1.0	20.0
5	abc	NaN	NaN	2.0	30.0

In [69]:

```
df
```

Out[69]:

	Name	Gender	DOB	Maths	Physics	Chemistry	English	Biology	Economics	History
0	John	M	05-04-1988	55	45	56	87	21	52	89
1	Suresh	M	04-05-1987	75	96	78	64	90	61	58
2	Ramesh	M	25-05-1989	25	54	89	76	95	87	56
3	Jessica	F	12-08-1990	78	96	86	63	54	89	75
4	Jennifer	F	02-09-1989	58	96	78	46	96	77	83
5	Annu	F	05-04-1988	45	87	52	89	55	89	87
6	pooja	F	04-05-1987	55	64	61	58	75	58	64
7	Ritesh	M	25-05-1989	54	76	87	56	25	56	76
8	Farha	F	12-08-1990	55	63	89	75	78	75	63
9	Mukesh	M	02-09-1989	96	46	77	83	58	83	46

In [92]:

```
import datetime
dt=datetime.datetime.now()
import warnings
warnings.filterwarnings("ignore")
```

In [93]:

```
df.info()
```

```
...
```

In [94]:

```
df['DOB']=pd.to_datetime(df['DOB'])
```

In [76]:

```
df['year']=df['DOB'].dt.year
```

In [77]:

```
df['year']
```

...

In [78]:

```
df
```

...

In [79]:

```
df['Month']=df['DOB'].dt.month
df
```

...

In [80]:

```
df['day']=df['DOB'].dt.day
```

In [81]:

```
df
```

...

In [91]:

```
total=df[~(df['sci_total']>=300) & (df['Gender']=='M')]
total
```

...

In [8]:

```
def getgrade(marks):

    if(marks>=300):
        return "A+"
    elif (marks>=200 and marks<300):
        return 'B'
    else:
        return 'C'

df['Grade']=df['sci_total'].apply(getgrade)
```

In [9]:

```
df
```

...

In [10]:

```
df.groupby('Grade')
```

Out[10]:

```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x00000166F15320A0>
```

In [11]:

```
df.groupby('Grade').count()
```

```
...
```

In [14]:

```
df.groupby('Grade')['Maths'].agg(['mean','max','min','count'])
```

Out[14]:

	mean	max	min	count
Grade				
A+	70.333333	78	58	3
B	55.000000	96	25	6
C	55.000000	55	55	1

In [17]:

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

In [20]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]

plt.plot(xdata,ydata)
plt.show()
```

```
...
```

In [39]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]


plt.plot(xdata,ydata)
plt.xlabel("X-axis", color='r')
plt.ylabel("Y-axis",color='g')
plt.title("Example of line chart" ,fontstyle='italic',fontweight='20')
plt.show()
```

```
...
```

In [27]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]


plt.plot(xdata,ydata,'-.',color='#FF0000')
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.title("Example of line chart")
plt.show()
```

...

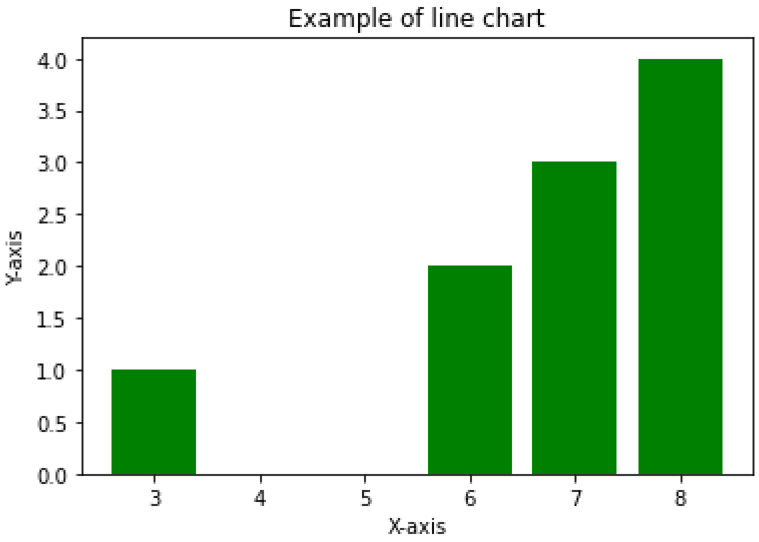
In [31]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]
xdata1=[5,6,9,7]
ydata1=[2,3,4,5]

plt.plot(xdata,ydata,'-.',color='#FF0000',linewidth=5)
plt.plot(xdata1,ydata1,'-.',color='g')
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.title("Example of line chart")
plt.show()
```

...

In [45]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]

plt.bar(xdata,ydata,width=0.8,color='g')
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.title("Example of line chart")
plt.show()
```



In [51]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]
mycolors=['red','hotpink','blue','g']
plt.bar(xdata,ydata,width=0.8,color=mycolors)
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.title("Example of line chart")
plt.show()
```

...

In [57]:

```
xdata=[3,6,7,8]
ydata=[1,2,3,4]
mycolors=['red','pink','blue','g']
plt.bar(xdata,ydata,width=0.8,color=mycolors,alpha=1)
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.title("Example of line chart")
plt.show()
```

...

In [ ]: