In [1]:
```
!pip install pandas
```

```
Requirement already satisfied: pandas in e:\anaconda\lib\site-packages (1.5.3)
Requirement already satisfied: numpy>=1.21.0 in e:\anaconda\lib\site-packages (from pandas) (1.23.5)
Requirement already satisfied: pytz>=2020.1 in e:\anaconda\lib\site-packages (from pandas) (2022.7)
Requirement already satisfied: python-dateutil>=2.8.1 in e:\anaconda\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: six>=1.5 in e:\anaconda\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

In [2]:
```
import pandas as pd
```

In [3]:
```
df=pd.read_csv('psa.csv')
print(df)
```

```
   Roll no. student name  grade  marks
0    190503   khiza hayat      A   72.0
1    190457         usama      c   64.0
2    190451      ehtisham    NaN    NaN
3    190433           ali    NaN   53.0
4    190441         raza      b   66.0
```

In [4]:
```
df.head()
```

Out[4]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	NaN	NaN
3	190433	ali	NaN	53.0
4	190441	raza	b	66.0

In [5]:
```
df.shape
```

Out[5]: (5, 4)

In [6]:
```
df.isnull()
```

Out[6]:

	Roll no.	student name	grade	marks
0	False	False	False	False
1	False	False	False	False
2	False	False	True	True
3	False	False	True	False
4	False	False	False	False

In [7]:
```
df.isnull().sum()
```

Out[7]:
```
Roll no.         0
student name     0
grade            2
marks            1
dtype: int64
```

In [8]:
```
df.isnull().sum().sum()
```

Out[8]: 3

filling the null values

In [9]:
```
df2=df.fillna(value=0)
df2
```

Out[9]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	0	0.0
3	190433	ali	0	53.0
4	190441	raza	b	66.0

filling the null values with a previous value

In [10]:
```
df4=df.fillna(method='pad')
df4
```

Out[10]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	c	64.0
3	190433	ali	c	53.0
4	190441	raza	b	66.0

In [11]:
```
df4.isnull().sum()
```

Out[11]:
```
Roll no.         0
student name     0
grade            0
marks            0
dtype: int64
```

In [12]:
```
# filling the null value with coming value to next row value
```

In [13]:
```
df5=df.fillna(method='bfill')
df5
```

Out[13]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	b	53.0
3	190433	ali	b	53.0
4	190441	raza	b	66.0

In [12]:
```
# filling the null value with coming value to next row value
```

In [13]:
```
df5=df.fillna(method='bfill')
df5
```

Out[13]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	b	53.0
3	190433	ali	b	53.0
4	190441	raza	b	66.0

In [14]:
```
# filling the null value with coming value to next coloumn  value
```

In [15]:
```
df6=df.fillna(method='pad',axis=1)
df6
```

Out[15]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	ehtisham	ehtisham
3	190433	ali	ali	53.0
4	190441	raza	b	66.0

In [16]:
```
df7=df.fillna(method='bfill',axis=1)
df7
```

Out[16]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	NaN	NaN
3	190433	ali	53.0	53.0
4	190441	raza	b	66.0

In [18]:
```
#filling diffferent null values in diffret coulomns
```

In [20]:
```
df8=df.fillna({'grade':'abcd',
               'marks':'defg'})
df8
```

Out[20]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.0
1	190457	usama	c	64.0
2	190451	ehtisham	abcd	defg
3	190433	ali	abcd	53.0
4	190441	raza	b	66.0

In [21]:
```
# filling the value with the mean of a coloumn
df9=df.fillna(value=df['marks'].mean())
df9
```

Out[21]:

	Roll no.	student name	grade	marks
0	190503	khiza hayat	A	72.00
1	190457	usama	c	64.00
2	190451	ehtisham	63.75	63.75
3	190433	ali	63.75	53.00
4	190441	raza	b	66.00

In [2]:
```
!pip install pandas
```

```
Requirement already satisfied: pandas in e:\anaconda\lib\site-packages (1.5.3)
Requirement already satisfied: numpy>=1.21.0 in e:\anaconda\lib\site-packages (from pandas) (1.23.5)
Requirement already satisfied: python-dateutil>=2.8.1 in e:\anaconda\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in e:\anaconda\lib\site-packages (from pandas) (2022.7)
Requirement already satisfied: six>=1.5 in e:\anaconda\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

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

In [6]:
```
data={'Names':['khizar','raza','usama','raza'],
     'Marks':[72,66,64,66],
     'City':['RYK','RWP','FSD','RWp']}
df=pd.DataFrame(data)
df
```

Out[6]:

	Names	Marks	City
0	khizar	72	RYK
1	raza	66	RWP
2	usama	64	FSD
3	raza	66	RWp

In [8]:
```
df.drop_duplicates(subset='Names',keep='first')
```

Out[8]:

	Names	Marks	City
0	khizar	72	RYK
1	raza	66	RWP
2	usama	64	FSD

In [9]:
```
df.drop_duplicates(subset='Names',keep='last')
```

Out[9]:

	Names	Marks	City
0	khizar	72	RYK
2	usama	64	FSD
3	raza	66	RWp

In [12]:
```
df.drop_duplicates(subset='Names',keep=False,inplace=True)
df
```

Out[12]:

	Names	Marks	City
0	khizar	72	RYK
2	usama	64	FSD

In [1]:
```
!pip install pandas
```

```
Requirement already satisfied: pandas in e:\anaconda\lib\site-packages (1.5.3)
Requirement already satisfied: pytz>=2020.1 in e:\anaconda\lib\site-packages (from pandas) (2022.7)
Requirement already satisfied: python-dateutil>=2.8.1 in e:\anaconda\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: numpy>=1.21.0 in e:\anaconda\lib\site-packages (from pandas) (1.23.5)
Requirement already satisfied: six>=1.5 in e:\anaconda\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

In [2]:
```
import pandas as pd
```

In [3]:
```
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [10]:
```
##define our dataset
pf= pd.read_csv('dt.csv')
#pf= [11,10,12,14,12,15,14,13,15,102,12,14,17,19,107,10,13,12,14,12,,108,12,11,14,13,15,10,15,12,10,14,13,15,10]
```

In [11]:
```
plt.hist(pf)
```

Out[11]:
```
(array([29.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  4.]),
 array([  10. ,   20.1,   30.2,   40.3,   50.4,   60.5,   70.6,   80.7,   90.8,
         100.9, 111. ]),
 <BarContainer object of 10 artists>)
```