

```
In [1]: import numpy as np
```

```
In [2]: X= np.array([2,4,6,8,10,12])  
np.mean(X)
```

```
Out[2]: 7.0
```

```
In [3]: np.median(X)
```

```
Out[3]: 7.0
```

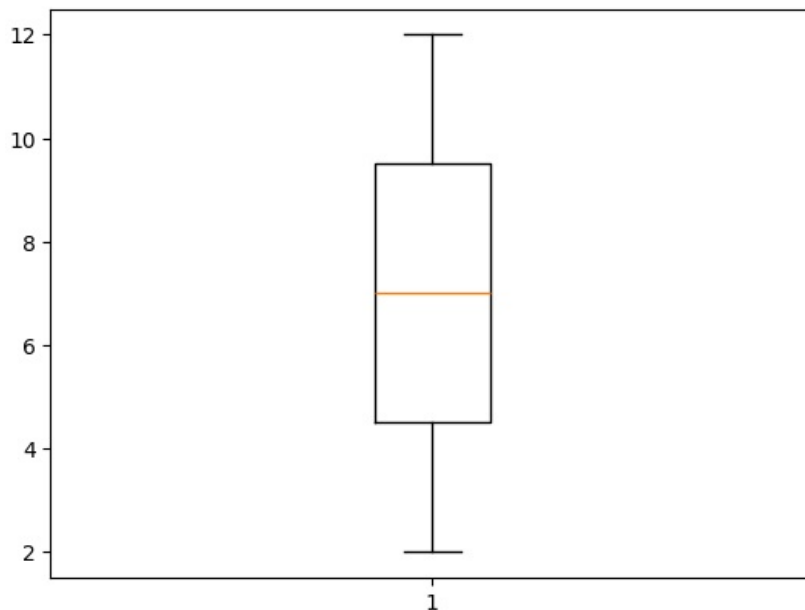
```
In [4]: import matplotlib.pyplot as plt  
import numpy as np  
import pandas as pd
```

```
In [5]: X= np.array([2,4,6,8,10,12])  
df= pd.DataFrame(X)  
print (df)
```

```
   0  
0  2  
1  4  
2  6  
3  8  
4 10  
5 12
```

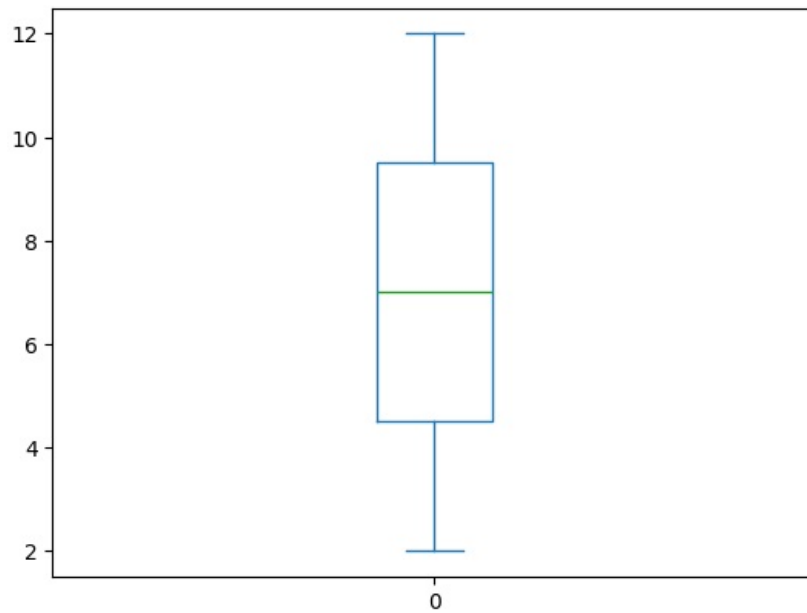
```
In [6]: plt.boxplot(X)
```

```
Out[6]: {'whiskers': [<matplotlib.lines.Line2D at 0x1ded701d100>,  
  <matplotlib.lines.Line2D at 0x1ded701d280>],  
 'caps': [<matplotlib.lines.Line2D at 0x1ded701d5b0>,  
  <matplotlib.lines.Line2D at 0x1ded701d8b0>],  
 'boxes': [<matplotlib.lines.Line2D at 0x1ded6fcb500>],  
 'medians': [<matplotlib.lines.Line2D at 0x1ded701db80>],  
 'fliers': [<matplotlib.lines.Line2D at 0x1ded701de80>],  
 'means': []}
```



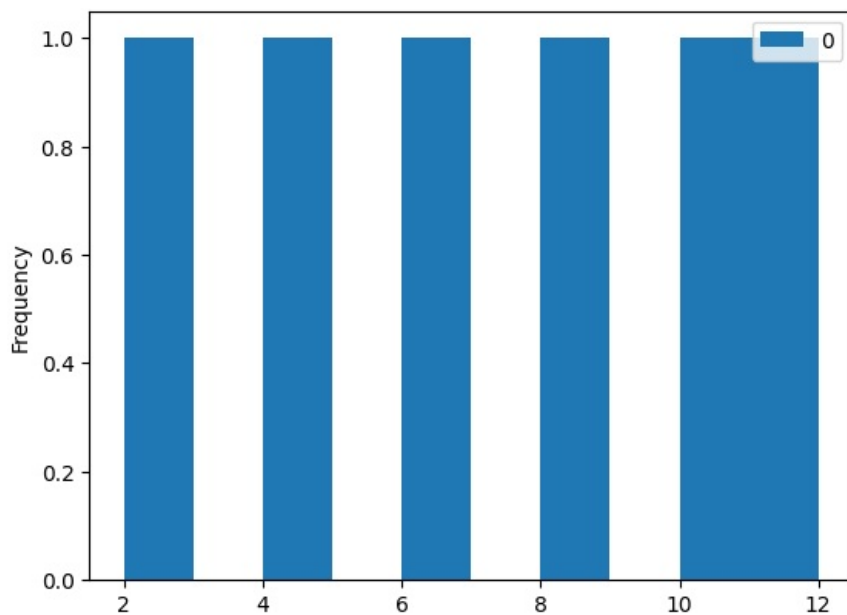
```
In [7]: df.plot.box()
```

```
Out[7]: <Axes: >
```



```
In [8]: df.plot.hist()
```

```
Out[8]: <Axes: ylabel='Frequency'>
```



```
In [9]: data = {
    "Name": ["Amit", "Renuka", "Raj", "Shital", "Vikram", "Ananya", "Rohan"],
    "Gender": ["Male", "Female", "Male", "Female", "Male", "Female", "Male"],
    "Marks": [85, 80, 78, np.nan, 76, 82, np.nan],
    "Age": [np.nan, 21, 22, np.nan, 24, np.nan, 26]
}
df = pd.DataFrame(data)
print(df)
```

	Name	Gender	Marks	Age
0	Amit	Male	85.0	NaN
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
3	Shital	Female	NaN	NaN
4	Vikram	Male	76.0	24.0
5	Ananya	Female	82.0	NaN
6	Rohan	Male	NaN	26.0

In [10]: `df.head()`

Out[10]:

	Name	Gender	Marks	Age
0	Amit	Male	85.0	NaN
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
3	Shital	Female	NaN	NaN
4	Vikram	Male	76.0	24.0

In [11]: `df.tail()`

Out[11]:

	Name	Gender	Marks	Age
2	Raj	Male	78.0	22.0
3	Shital	Female	NaN	NaN
4	Vikram	Male	76.0	24.0
5	Ananya	Female	82.0	NaN
6	Rohan	Male	NaN	26.0

In [12]: `df.count()`

Out[12]:

```
Name      7
Gender     7
Marks      5
Age        4
dtype: int64
```

In [13]: `df.isnull()`

Out[13]:

	Name	Gender	Marks	Age
0	False	False	False	True
1	False	False	False	False
2	False	False	False	False
3	False	False	True	True
4	False	False	False	False
5	False	False	False	True
6	False	False	True	False

In [14]: `df.isnull().sum()`

Out[14]:

```
Name      0
Gender     0
Marks      2
Age        3
dtype: int64
```

In [15]: `df.dropna()`

Out[15]:

	Name	Gender	Marks	Age
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
4	Vikram	Male	76.0	24.0

In [16]: `df.fillna(0)`

Out[16]:

	Name	Gender	Marks	Age
0	Amit	Male	85.0	0.0
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
3	Shital	Female	0.0	0.0
4	Vikram	Male	76.0	24.0
5	Ananya	Female	82.0	0.0
6	Rohan	Male	0.0	26.0

In [17]: df['Marks'].fillna(df['Marks'].mean())

Out[17]:

0	85.0
1	80.0
2	78.0
3	80.2
4	76.0
5	82.0
6	80.2

Name: Marks, dtype: float64

In [18]: df['Age'].fillna(df['Age'].median())

Out[18]:

0	23.0
1	21.0
2	22.0
3	23.0
4	24.0
5	23.0
6	26.0

Name: Age, dtype: float64

In [19]: df.fillna(method='bfill')

C:\Users\Apurva\AppData\Local\Temp\ipykernel\_10028\2831856154.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.  
df.fillna(method='bfill')

Out[19]:

	Name	Gender	Marks	Age
0	Amit	Male	85.0	21.0
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
3	Shital	Female	76.0	24.0
4	Vikram	Male	76.0	24.0
5	Ananya	Female	82.0	26.0
6	Rohan	Male	NaN	26.0

In [20]: df.fillna(method='pad')

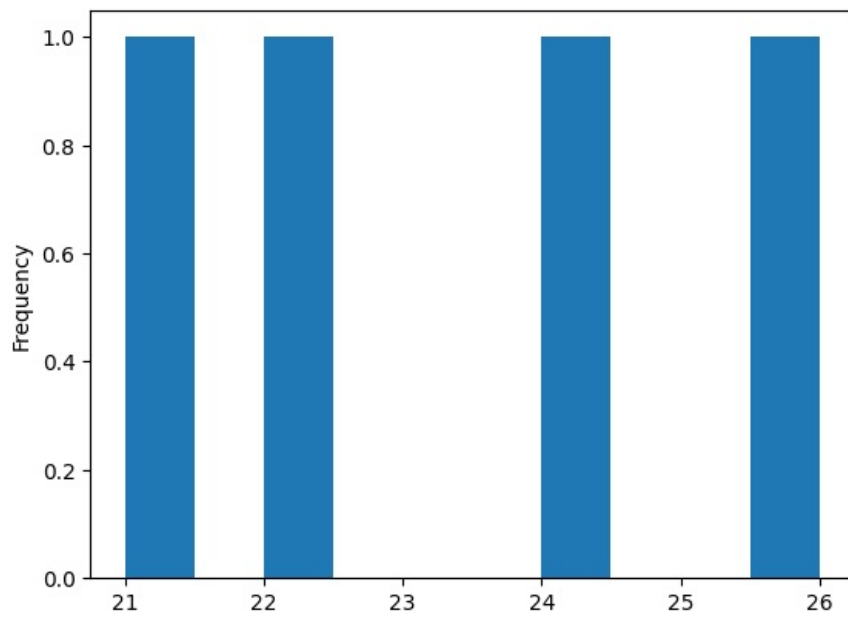
C:\Users\Apurva\AppData\Local\Temp\ipykernel\_10028\196241153.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.  
df.fillna(method='pad')

Out[20]:

	Name	Gender	Marks	Age
0	Amit	Male	85.0	NaN
1	Renuka	Female	80.0	21.0
2	Raj	Male	78.0	22.0
3	Shital	Female	78.0	22.0
4	Vikram	Male	76.0	24.0
5	Ananya	Female	82.0	24.0
6	Rohan	Male	82.0	26.0

In [21]: df['Age'].plot.hist()

Out[21]: <Axes: ylabel='Frequency'>



In [ ]:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js