```
from google.colab import drive
```

```
# importing all the libraries for the analysis
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
df=pd.DataFrame({'Name':['John','Sam','Helix','Fenna','Mixi','Smith','Mor'],'Salary': [50000, 54000, 50000, 189000, 55000, 44000, 59000]
    'Hours': [41, 40, 36, 30, 35, 39, 40],'Grade': [50, 50, 46, 95, 50, 5, 57]})
```

```
df
```

	Name	Salary	Hours	Grade
0	John	50000	41	50
1	Sam	54000	40	50
2	Helix	50000	36	46
3	Fenna	189000	30	95
4	Mixi	55000	35	50
5	Smith	44000	39	5
6	Mor	59000	40	57

Next steps: Generate code with df | View recommended plots | New interactive sheet

Generate | 10 random numbers using numpy | Close

```
df.shape
(7,4)
```

```
(7, 4)
```

```
df['Salary'].mean()
```

```
np.float64(71571.42857142857)
```

```
df['Salary'].median()
```

```
54000.0
```

```
df['Salary'].mode().values[0]
```

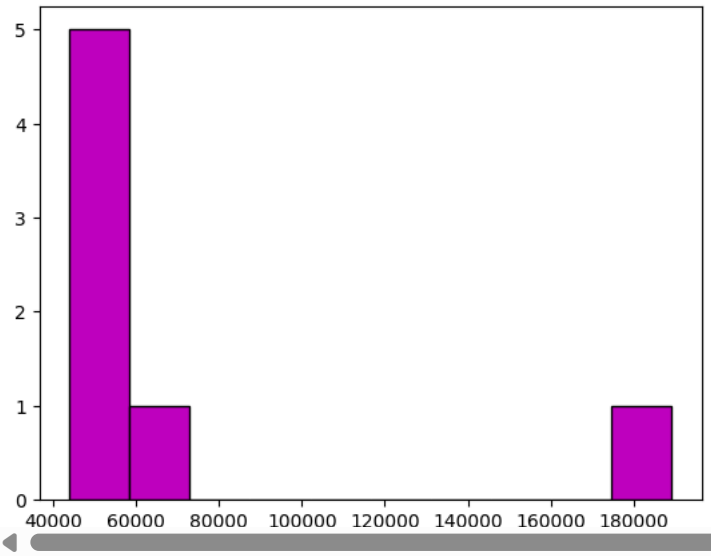
```
np.int64(50000)
```

```
df['Salary'].mode().values[0]
```

```
np.int64(50000)
```

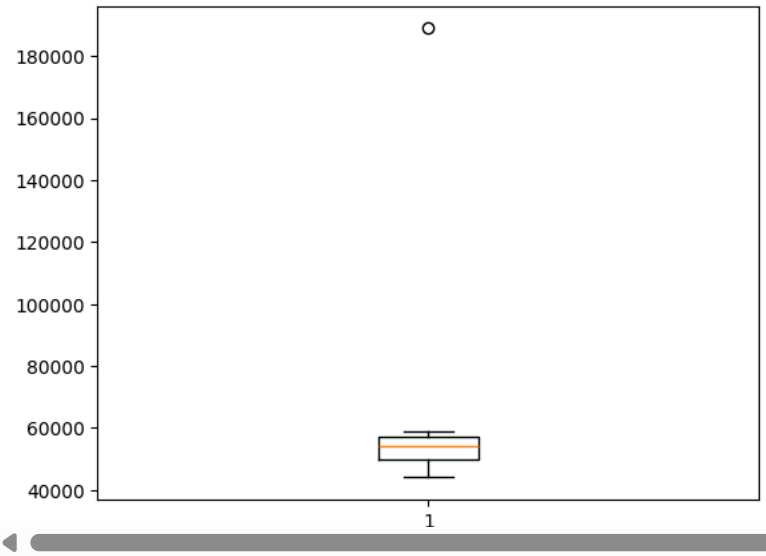
```
plt.hist(df['Salary'],bins=10,color='m',edgecolor='k')
```

```
(array([5., 1., 0., 0., 0., 0., 0., 0., 0., 1.]),
 array([ 44000.,  58500.,  73000.,  87500., 102000., 116500., 131000.,
        145500., 160000., 174500., 189000.]),
 <BarContainer object of 10 artists>)
```




```
plt.boxplot(df['Salary'])
```

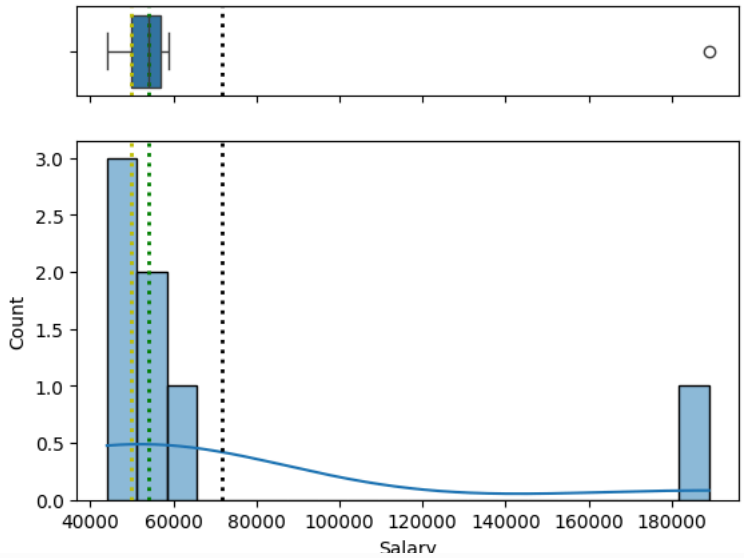
```
{'whiskers': [<matplotlib.lines.Line2D at 0x7ed7d8a6f690>,
  <matplotlib.lines.Line2D at 0x7ed7d8a800d0>],
 'caps': [<matplotlib.lines.Line2D at 0x7ed7d8a80b50>,
  <matplotlib.lines.Line2D at 0x7ed7d8a81450>],
 'boxes': [<matplotlib.lines.Line2D at 0x7ed7d8b72f10>],
 'medians': [<matplotlib.lines.Line2D at 0x7ed7d8a81e90>],
 'fliers': [<matplotlib.lines.Line2D at 0x7ed7d8a82790>],
 'means': []}
```



Double-click (or enter) to edit

```
# to create subplots for showing multiple plots together by specifying the number of rows and#
f,(ax_box,ax_hist)=plt.subplots(2,sharex=True,gridspec_kw={"height_ratios":(0.5,2)})
# creating a boxplot by passig the data and giving the x values and stating the axis
sns.boxplot(data=df,x=df['Salary'],ax=ax_box)
ax_box.axvline(df['Salary'].mean(),color='k',linestyle='dotted',linewidth=2)
ax_box.axvline(df['Salary'].median(),color='g',linestyle='dotted',linewidth=2)
ax_box.axvline(df['Salary'].mode().values[0],color='y',linestyle='dotted',linewidth=2)
sns.histplot(data=df,x=df['Salary'],ax=ax_hist,kde=True)
ax_hist.axvline(df['Salary'].mean(),color='k',linestyle='dotted',linewidth=2)
ax_hist.axvline(df['Salary'].median(),color='g',linestyle='dotted',linewidth=2)
ax_hist.axvline(df['Salary'].mode().values[0],color='y',linestyle='dotted',linewidth=2)
```

```
<matplotlib.lines.Line2D at 0x7ed7db9aca10>
```




```
from google.colab import files
uploaded = files.upload()
```

Choose files tip.csv

- **tip.csv**(text/csv) - 7943 bytes, last modified: 03/06/2025 - 100% done

```
dftips=pd.read_csv('tip.csv')
```

```
dftips
```

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...	...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

Next steps: Generate code with dftips | View recommended plots | New interactive sheet

```
dftips.shape
```

```
(244, 7)
```

```
sns.set_style('dark')
sns.distplot(dftips['tip'],bins=10)
```

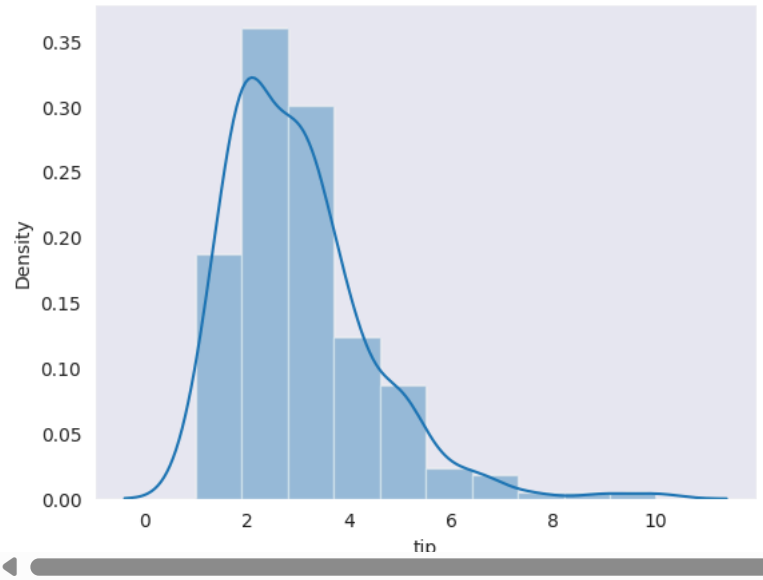
```
<ipython-input-60-77e2db850e86>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('dark')
sns.distplot(dftips['tip'],bins=10)
```

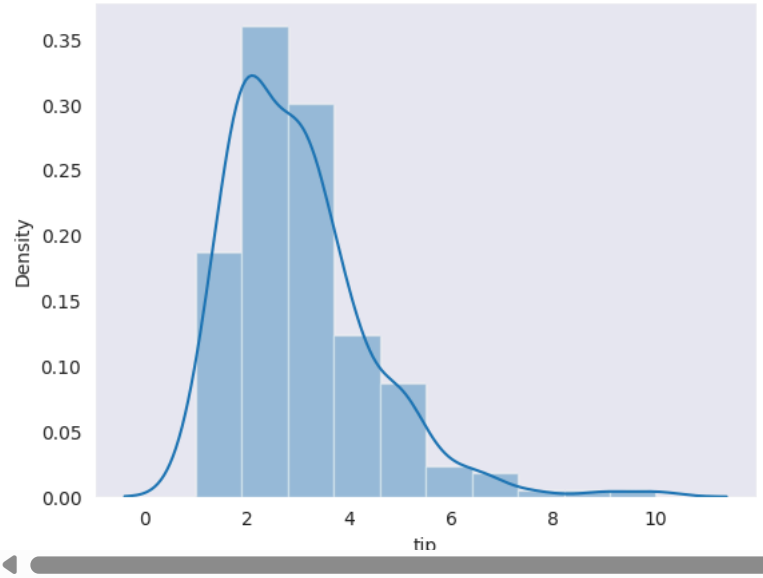
```
<ipython-input-62-77e2db850e86>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('dark')
sns.distplot(dftips['tip'],bins=10)
```

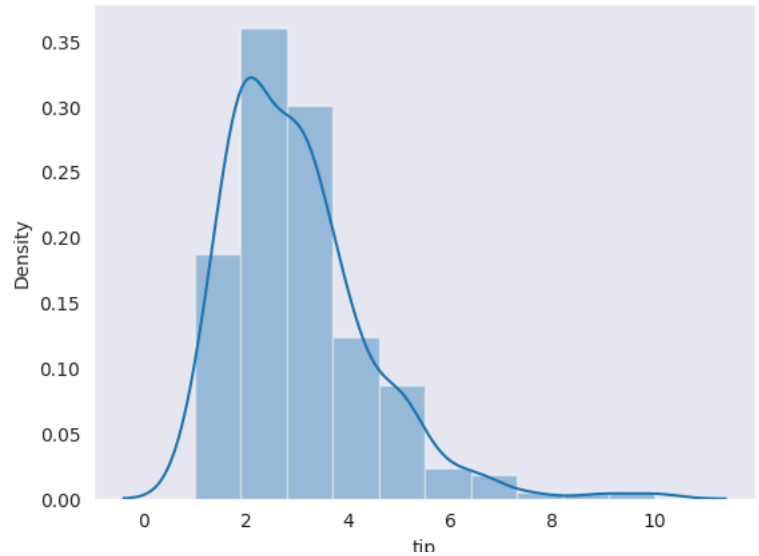
```
<ipython-input-63-77e2db850e86>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('ticks')
sns.distplot(dftips['tip'],bins=10)
```

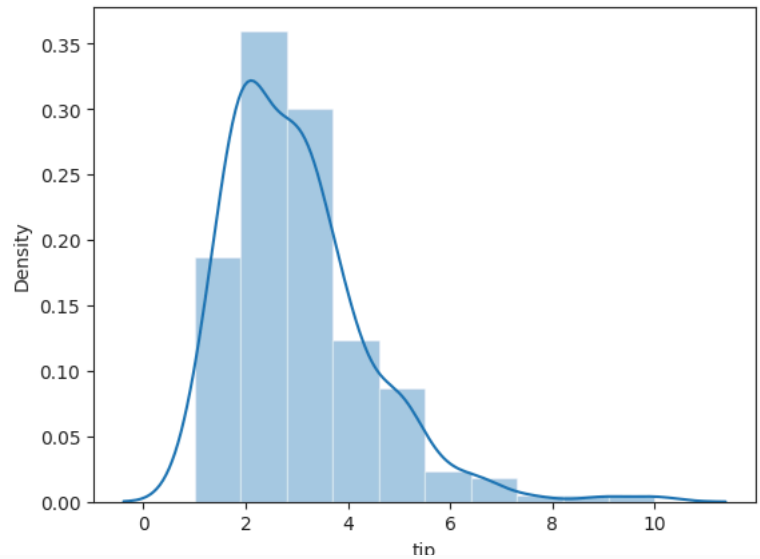
```
<ipython-input-64-2c17eaf9a7ee>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('white')
sns.distplot(dftips['tip'],bins=10)
```

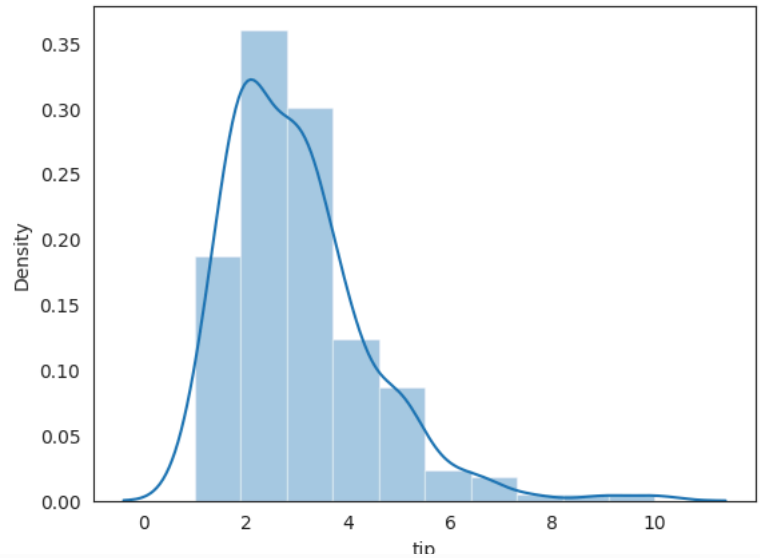
```
<ipython-input-65-59b3eada40d0>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('whitegrid')
sns.distplot(dftips['tip'],bins=10)
```

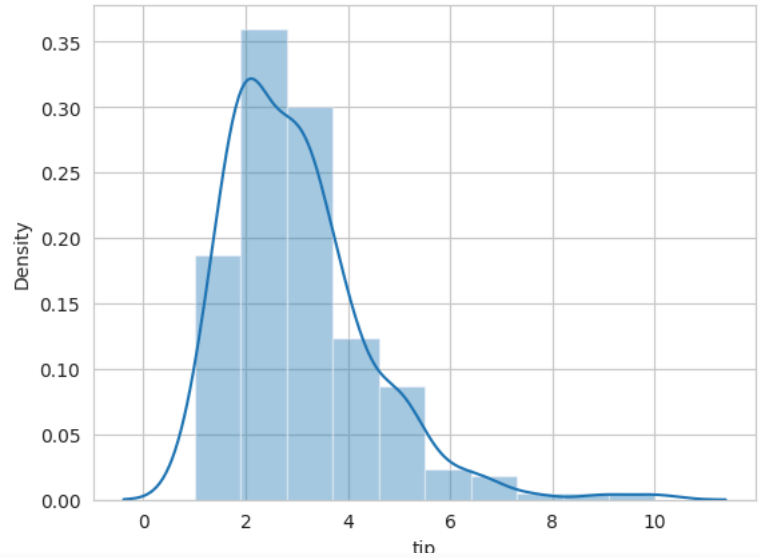
```
<ipython-input-66-9809adb276df>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10)
<Axes: xlabel='tip', ylabel='Density'>
```




```
sns.set_style('dark')
sns.distplot(dftips['tip'],bins=10,kde=False)
```

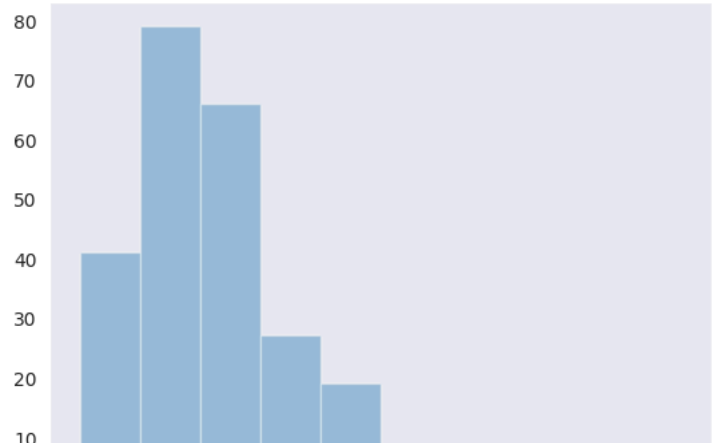
```
<ipython-input-67-8245468656fb>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10,kde=False)
<Axes: xlabel='tip'>
```




```
sns.set_style('dark')
sns.distplot(dftips['tip'],bins=10,rug=True)
```

```
<ipython-input-68-d6e212561aca>:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dftips['tip'],bins=10,rug=True)
<Axes: xlabel='tip', ylabel='Density'>
```