Seaborn_Basics

August 19, 2022

```
[1]: import pandas as pd
```

```
[3]: data=pd.read_excel('Sample.xlsx')
     data.head()
```

```
[3]:    Sno   Name  Age Gender    profession
     0    1   XYZ1   34      M   sw engineer
     1    2   XYZ2   45      F        Doctor
     2    3   XYZ3   26      M     Architetc
     3    4   XYZ4   28      M   sw engineer
     4    5   XYZ5   29      M        Doctor
```

```
[4]: data.tail()
```

```
[4]:    Sno    Name  Age Gender profession
     5    6   XYZ6   30      F  Architetc
     6    7   XYZ7   54      M     Farmer
     7    8   XYZ8   36      M     Doctor
     8    9   XYZ9   29      F     Farmer
     9   10  XYZ10   38      F  Architetc
```

```
[5]: data.describe()
```

```
[5]:             Sno        Age
     count  10.00000  10.000000
     mean    5.50000  34.900000
     std     3.02765   8.812239
     min     1.00000  26.000000
     25%     3.25000  29.000000
     50%     5.50000  32.000000
     75%     7.75000  37.500000
     max    10.00000  54.000000
```

```
[6]: import matplotlib.pyplot as plt
     import seaborn as sns
     sns.distplot(data['Age'])
```

1

```
/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed in a
future version. Please adapt your code to use either `displot` (a figure-level
function with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)
```

[6]: <AxesSubplot:xlabel='Age', ylabel='Density'>
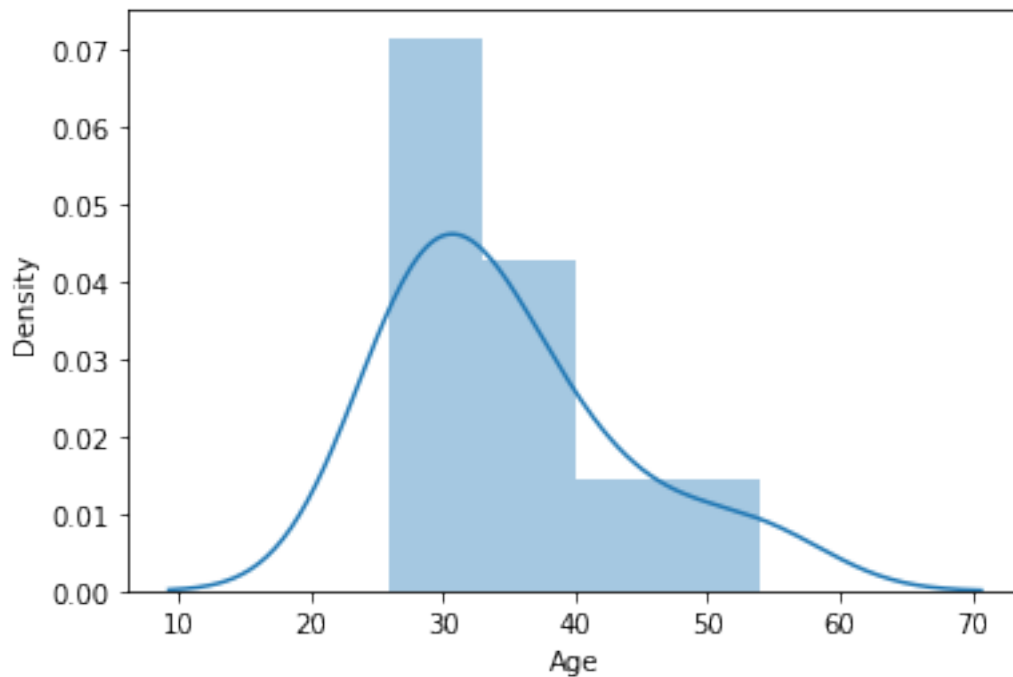
[7]:
```
Salary_df=pd.read_csv('SalaryGender.csv')
Salary_df.head()
```

[7]:
```
   Salary  Gender  Age  PhD
0   140.0       1   47    1
1    30.0       0   65    1
2    35.1       0   56    0
3    30.0       1   23    0
4    80.0       0   53    1
```

[8]:
```
Salary_df.tail()
```

[8]:
```
    Salary  Gender  Age  PhD
95    18.6       1   26    0
96   152.0       1   56    1
97     1.8       1   28    0
```

```
    98    35.0      0    44    0
    99     4.0      0    24    0
```

[9]: `Salary_df.describe()`

[9]:
	Salary	Gender	Age	PhD
count	100.000000	100.000000	100.000000	100.000000
mean	52.524500	0.500000	46.880000	0.390000
std	42.220933	0.502519	15.271469	0.490207
min	0.250000	0.000000	20.000000	0.000000
25%	20.000000	0.000000	31.500000	0.000000
50%	39.300000	0.500000	49.000000	0.000000
75%	75.500000	1.000000	60.000000	1.000000
max	190.000000	1.000000	77.000000	1.000000

[10]:
```python
plt.subplot(2,2,1)
sns.distplot(Salary_df['Salary'])
plt.subplot(2,2,2)
sns.distplot(Salary_df['Gender'])
plt.subplot(2,2,3)
sns.distplot(Salary_df['Age'])
plt.subplot(2,2,4)
sns.distplot(Salary_df['PhD'])
```

/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed in a
future version. Please adapt your code to use either `displot` (a figure-level
function with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)
/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed in a
future version. Please adapt your code to use either `displot` (a figure-level
function with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)
/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed in a
future version. Please adapt your code to use either `displot` (a figure-level
function with similar flexibility) or `histplot` (an axes-level function for
histograms).
  warnings.warn(msg, FutureWarning)
/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed in a
future version. Please adapt your code to use either `displot` (a figure-level
function with similar flexibility) or `histplot` (an axes-level function for
histograms).

```
warnings.warn(msg, FutureWarning)
```

[10]: <AxesSubplot:xlabel='PhD', ylabel='Density'>