pda-401-03

March 19, 2025

#Seaborn

[2]:
```
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

[11]:
```
mydata={"Name":["Ravi","Sam","Joe","Asha"],
        "Age":[23,22,26,47],
        "Salary":[12000,13000,14000,20000],
        "Exp":[2,1,3,4]}
```

[13]:
```
df=pd.DataFrame(mydata)
df.head()
```

[13]:

	Name	Age	Salary	Exp
0	Ravi	23	12000	2
1	Sam	22	13000	1
2	Joe	26	14000	3
3	Asha	47	20000	4

1.Histogram

[35]:
```
plt.figure(figsize=(4,4))
sns.histplot(df['Salary'],kde=True,bins=2)
plt.title('Distribution of salary')
plt.show()
```

1.positive skew,larger salary value

2.no outlines detected

3.Average salary is around 13000

4.Majority salary values are between 12000 to 16000

Corelation matrix

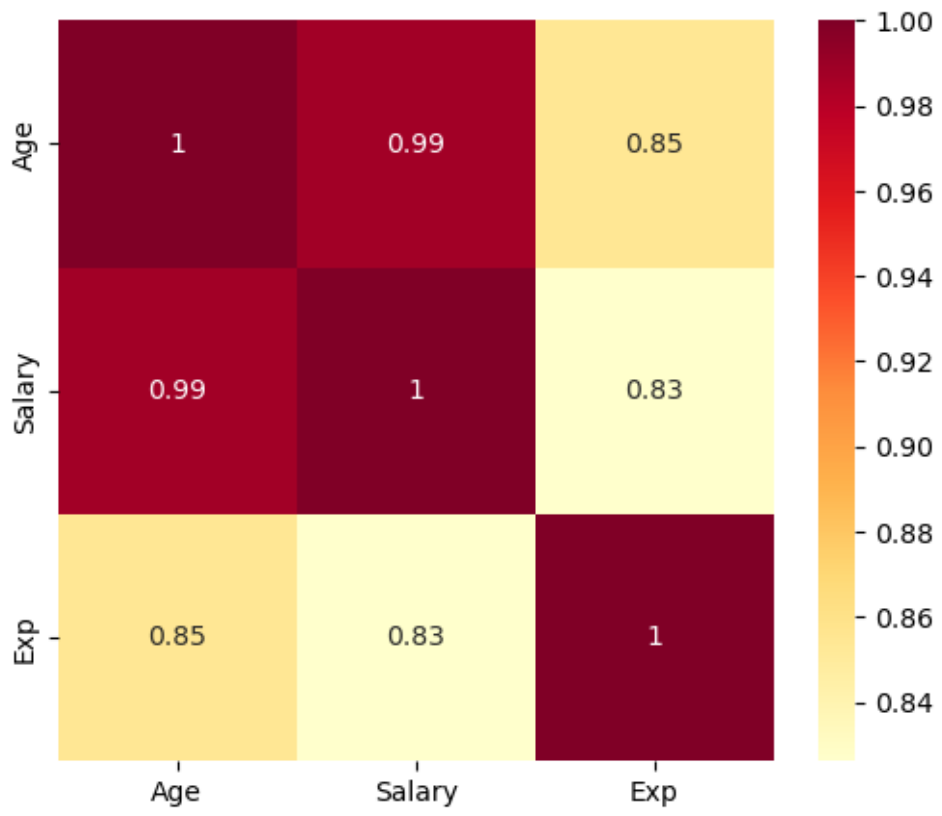
[29]:
```
#step 1 filter numerical data
ndf=df.select_dtypes(include=['number'])
ndf.head()
```

[29]:

	Age	Salary	Exp
0	23	12000	2
1	22	13000	1
2	26	14000	3
3	47	20000	4

[39]:
```
#step 2 use heat map
plt.figure(figsize=(6,5))
sns.heatmap(ndf.corr(),cmap='OrRd',annot=True)
```

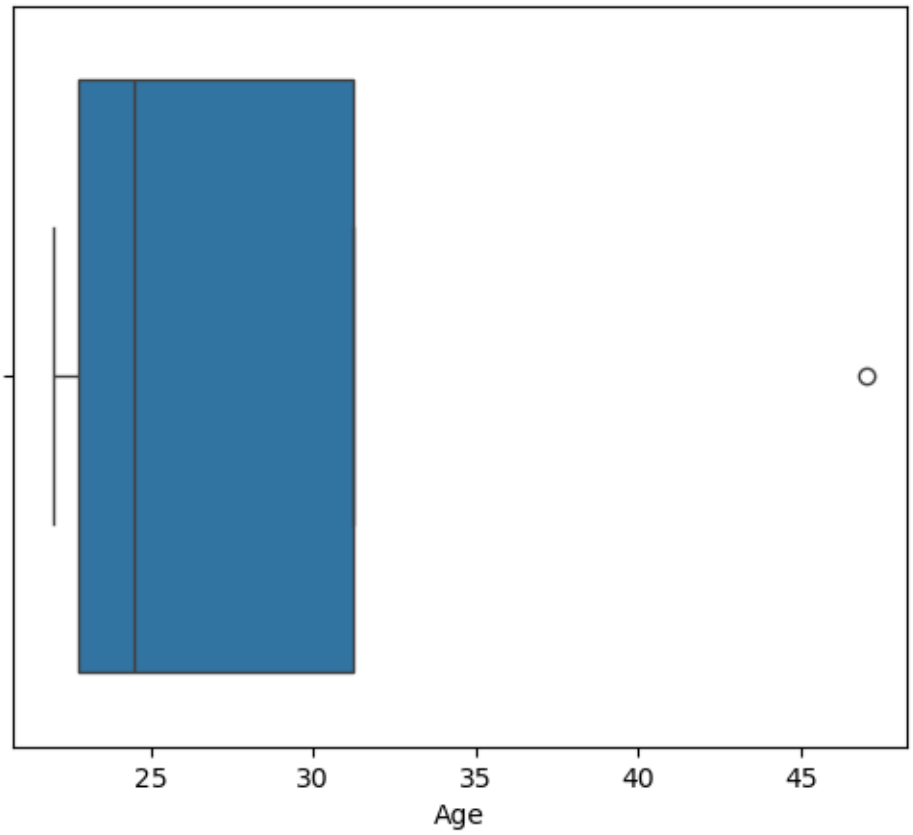
[39]: <Axes: >



[ ]:
```
1.It shows the corelation of age ,salary and exp
2.Here the age salary and exp contains equal corelation
```

[43]:
```
plt.figure(figsize=(6,5))
sns.boxplot(x=df['Age'])
```

[43]:
```
<Axes: xlabel='Age'>
```



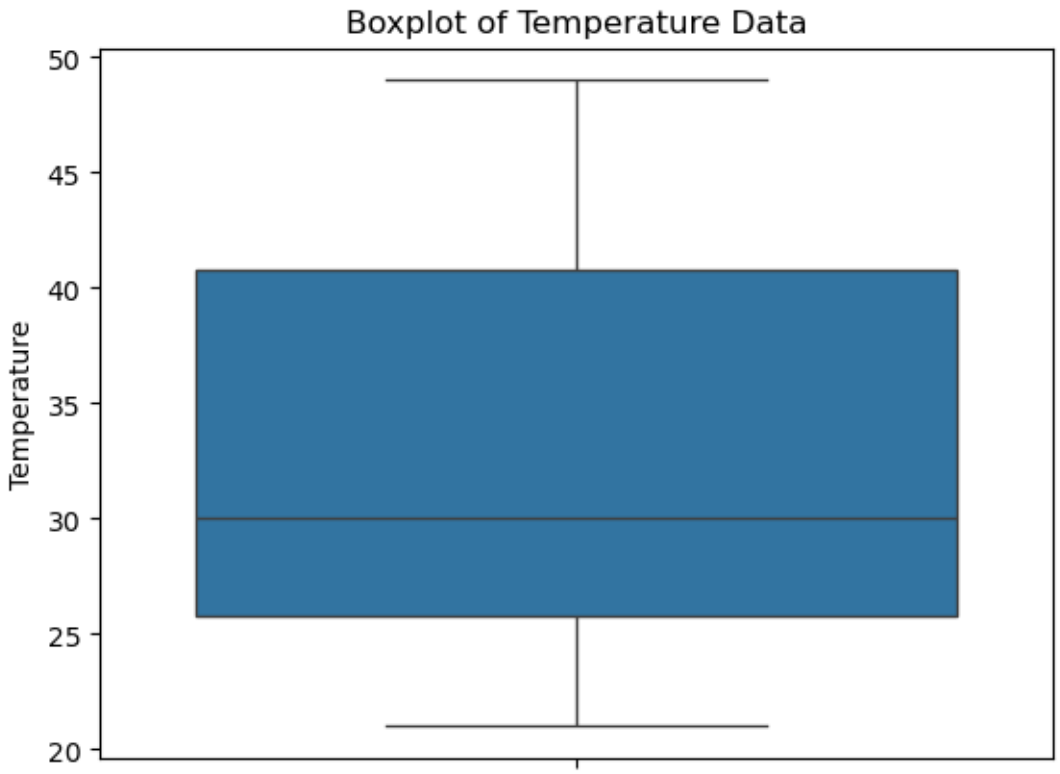
1.The boxplot has interpreted same as histogram

2.the distribution based on age

3.average age value is 25

Find the outline in the following data : temp=[21,47,39,22,31,33,29,26,27,25,49,46]

[51]:
```
temp = [21, 47, 39, 22, 31, 33, 29, 26, 27, 25, 49, 46]
sns.boxplot(temp)
plt.title('Boxplot of Temperature Data')
plt.ylabel('Temperature')
plt.show()
```



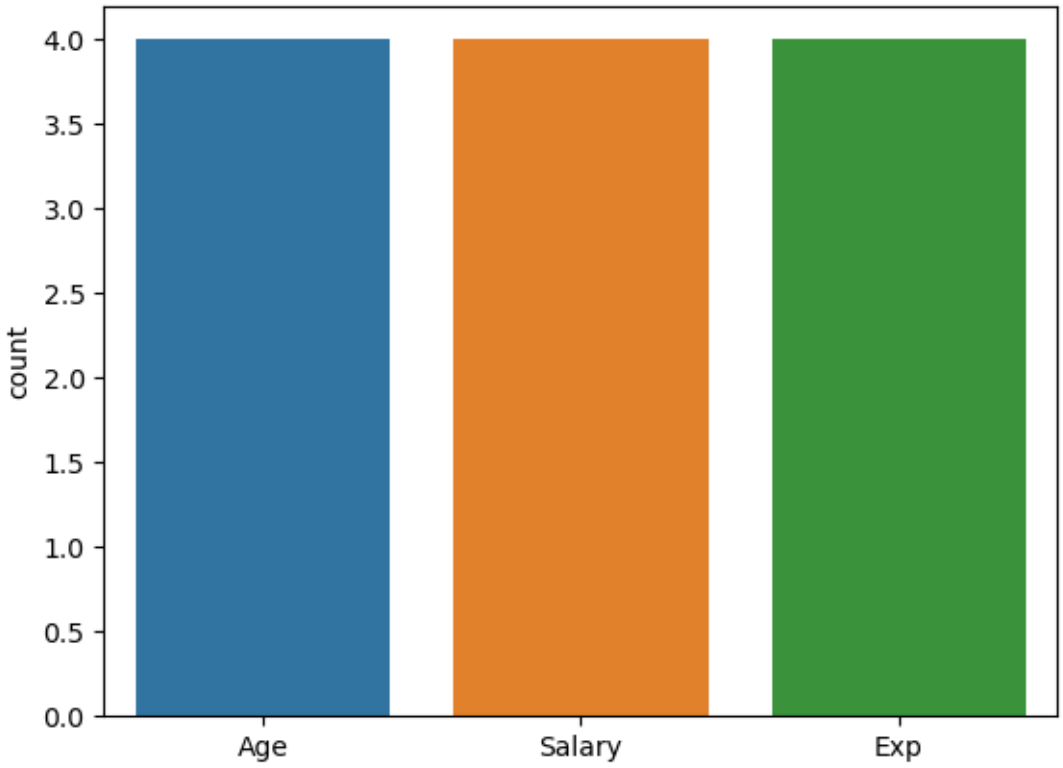
1.the distribution based on temparature

2.average value is 30

3.abnormal outline is not detected

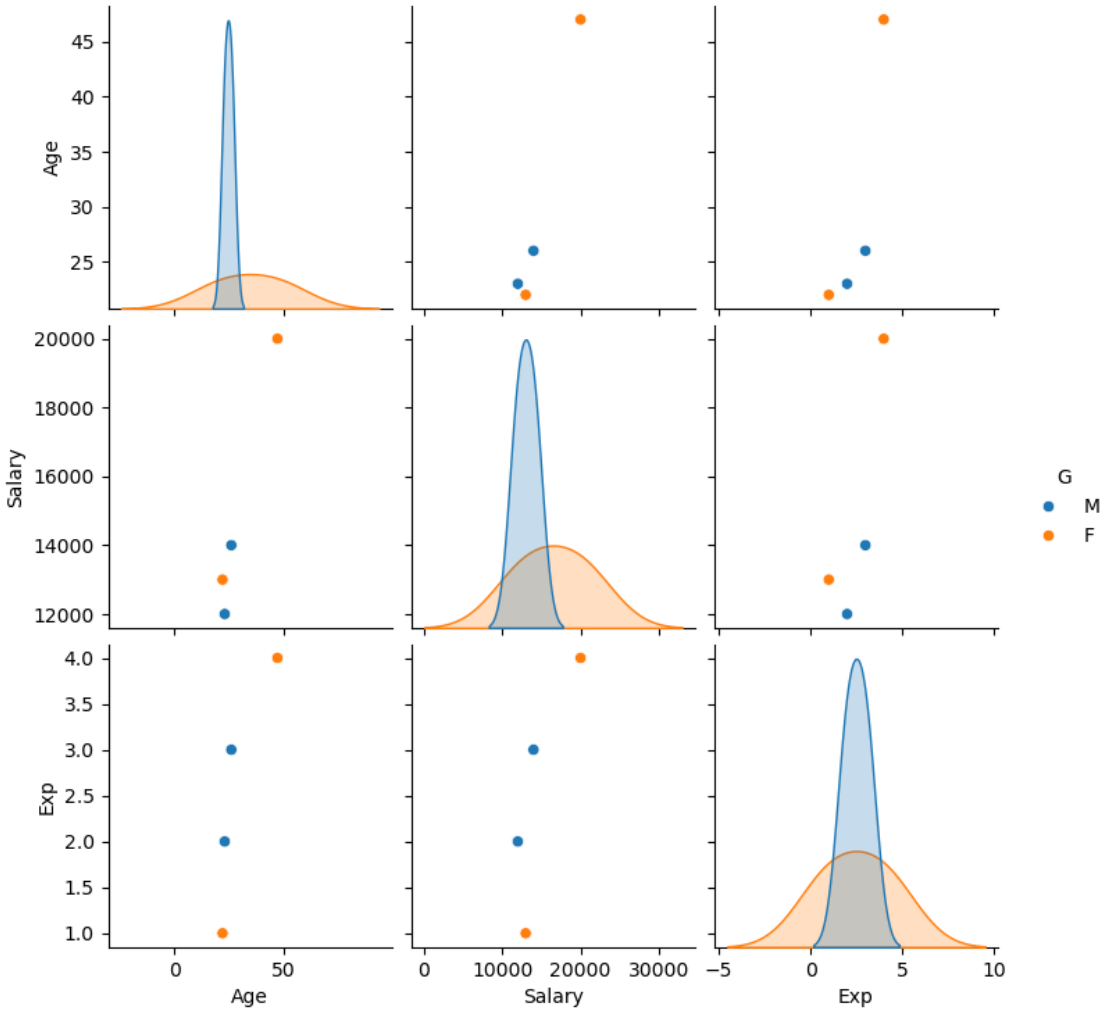
[59]:
```
mydata1={"Name":["Ravi","Sam","Joe","Asha"],
         "Age":[23,22,26,47],
         "Salary":[12000,13000,14000,20000],
         "Exp":[2,1,3,4],
          "G":["M","F","M","F"]}
df1=pd.DataFrame(mydata1)
sns.countplot(df1)
```

[59]:
```
<Axes: ylabel='count'>
```



[63]:
```
sns.pairplot(df1, hue='G')
```

[63]:
```
<seaborn.axisgrid.PairGrid at 0x1ee7e48c8c0>
```



[ ]:
```
1. the orange scatters are showing the genders
```