Seaborn

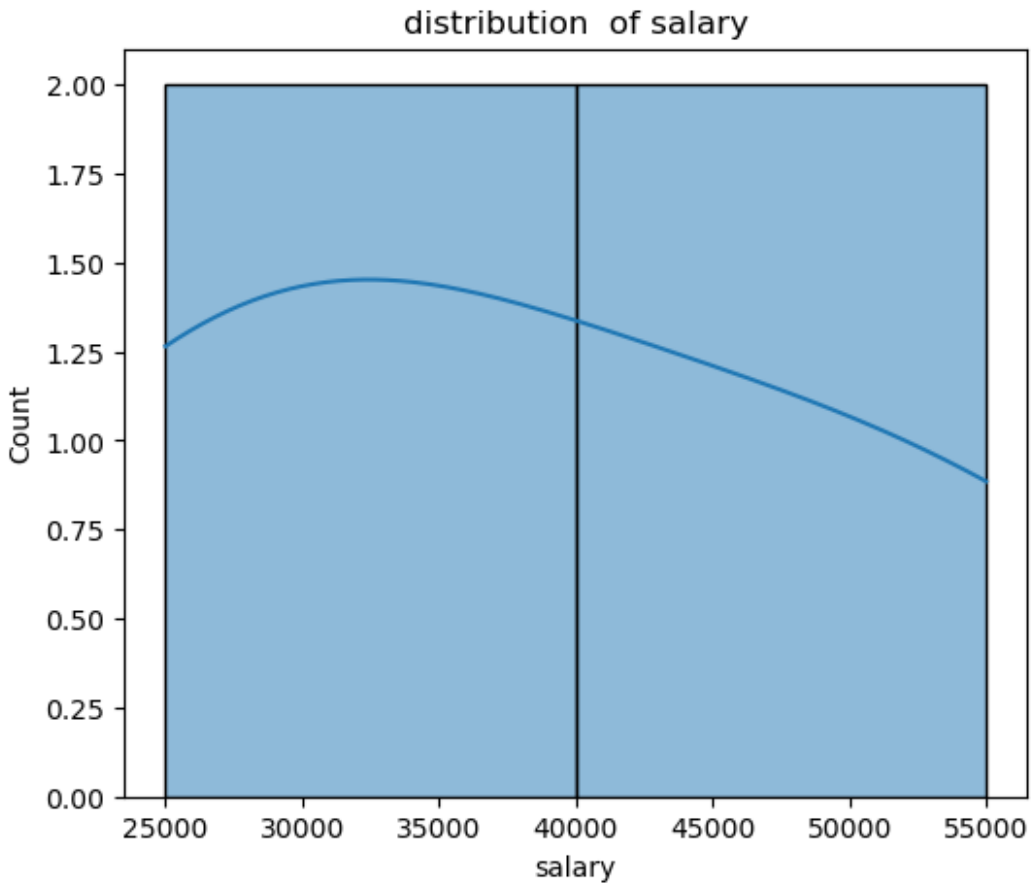
```
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

mydata={'name':['ram','sam','joe','asha'],
        'Age':[12,23,34,52],
        'salary':[30000,55000,42000,25000],
       'exp':[2,1,3,10]}

df=pd.DataFrame(mydata)
df.head()
```

	name	Age	salary	exp
0	ram	12	30000	2
1	sam	23	55000	1
2	joe	34	42000	3
3	asha	52	25000	10

```
plt.figure(figsize =(6,5))
sns.histplot(df['salary'],kde=True,bins=2)
plt.title('distribution  of salary')
plt.show()
```



Conclusion

- negative skew,larger salry value
- No outliner detected
- average salary is around 33000
- majarity salary value b/w 25000 and 55000

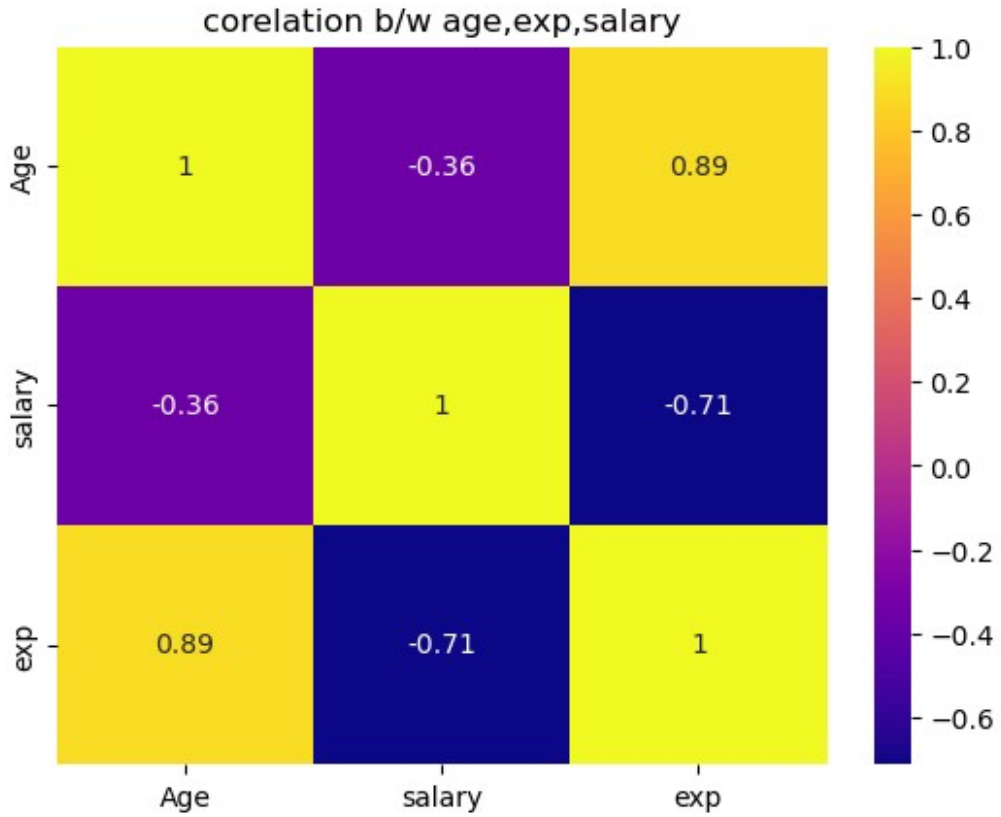
corelation matrix

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

```
   Age  salary  exp
0   12   30000    2
1   23   55000    1
2   34   42000    3
3   52   25000   10
```

```
# heatmap
plt.figure(figure=(6,5))
```

```
sns.heatmap(ndf.corr(),cmap='plasma',annot=True)
plt.title('corelation b/w age,exp,salary')
plt.show()
```

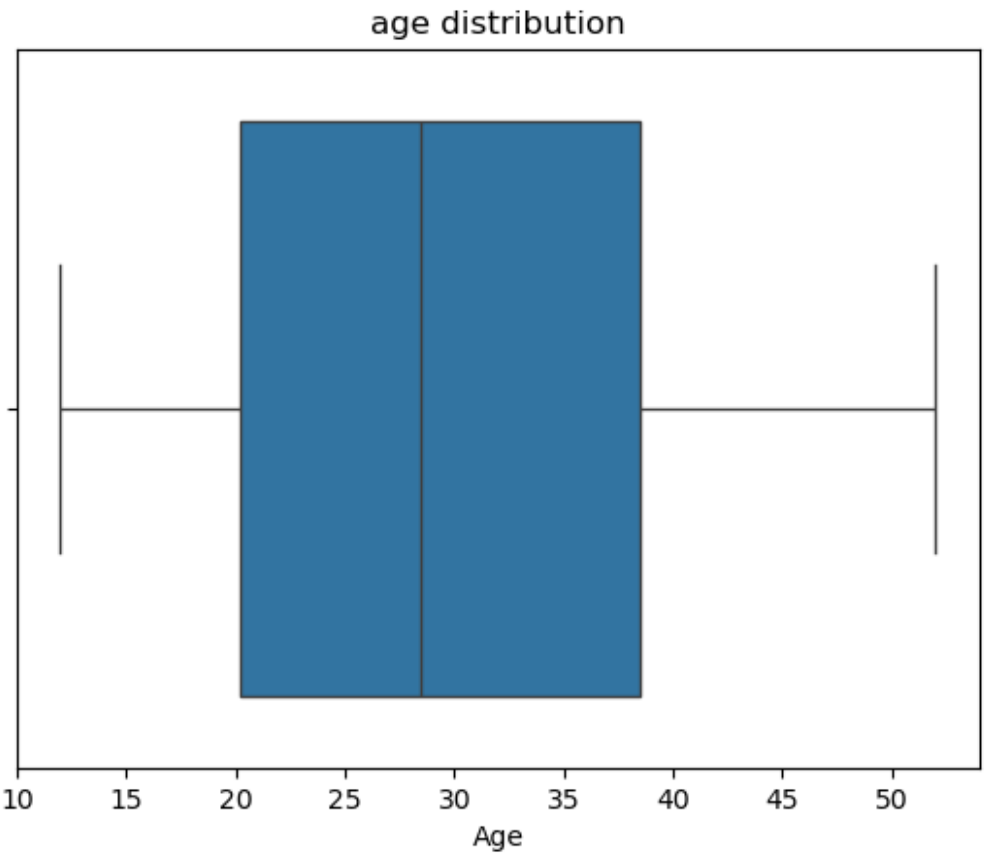


Conclusion:

- age sal exp are the corelations
- age and salary is the lowest corelation

Box plot

```
plt.figure(figure=(6,5))
sns.boxplot(x=df['Age'])
plt.title('age distribution')
plt.show()
```



conclusion:

- the average value is the 28
- no outliner
- the majority of ages b/w 20 to 39

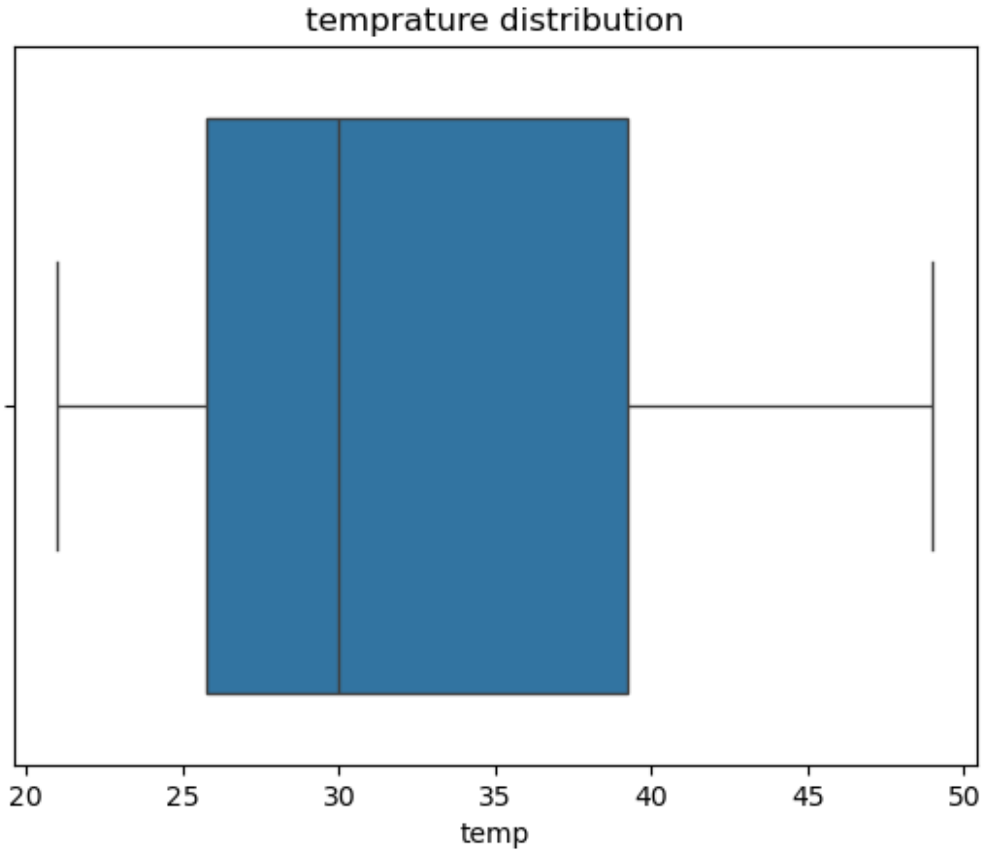
find the outliner in the following data :temp=[21,47,39,22,31,33,29,26,27,25,49,40]

```
mydata={'name':['A','B','c','d','e','f','g','h','i','j','k','l'],
        'temp':[21,47,39,22,31,33,29,26,27,25,49,40],}
df=pd.DataFrame(mydata)
df.head(12)
```

	name	temp
0	A	21
1	B	47
2	c	39
3	d	22
4	e	31
5	f	33
6	g	29
7	h	26
8	i	27
9	j	25

```
10    k    49
11    l    40
```

```
plt.figure(figure=(6,5))
sns.boxplot(x=df['temp'])
plt.title('temprature distribution')
plt.show()
```



Conclusion:

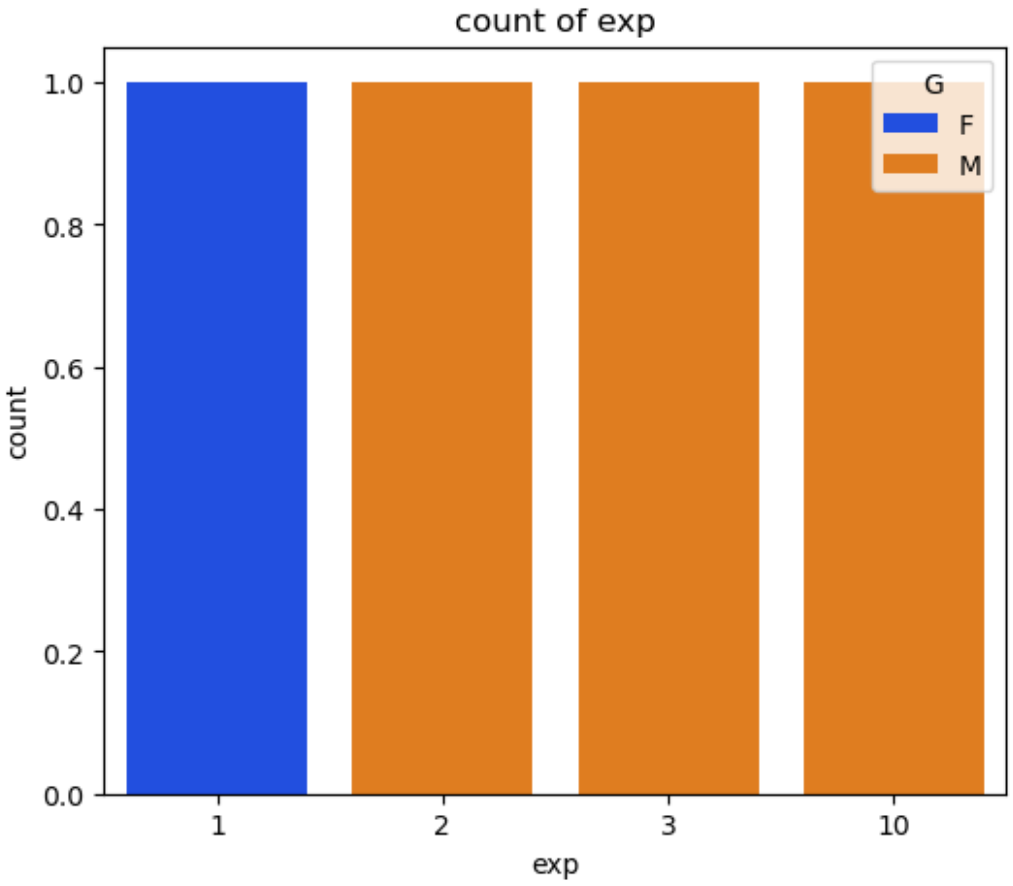
- no outliner
- it is the positive skew
- the majority of temprature is the 26 to 40
- the average of temprature is 30

```
mydata1={'name':['ram','sam','joe','asha'],
         'Age':[12,23,34,52],
         'salary':[30000,55000,42000,25000],
        'exp':[2,1,3,10],
        'G':['M','F','M','M']}
df=pd.DataFrame(mydata1)
df
```

```
   name  Age  salary  exp  G
0   ram   12   30000    2  M
1   sam   23   55000    1  F
2   joe   34   42000    3  M
3  asha   52   25000   10  M
```

```
plt.figure(figsize=(6,5))
sns.countplot(x=df['exp'],palette='bright',hue=df['G'])
plt.title('count of exp')
```

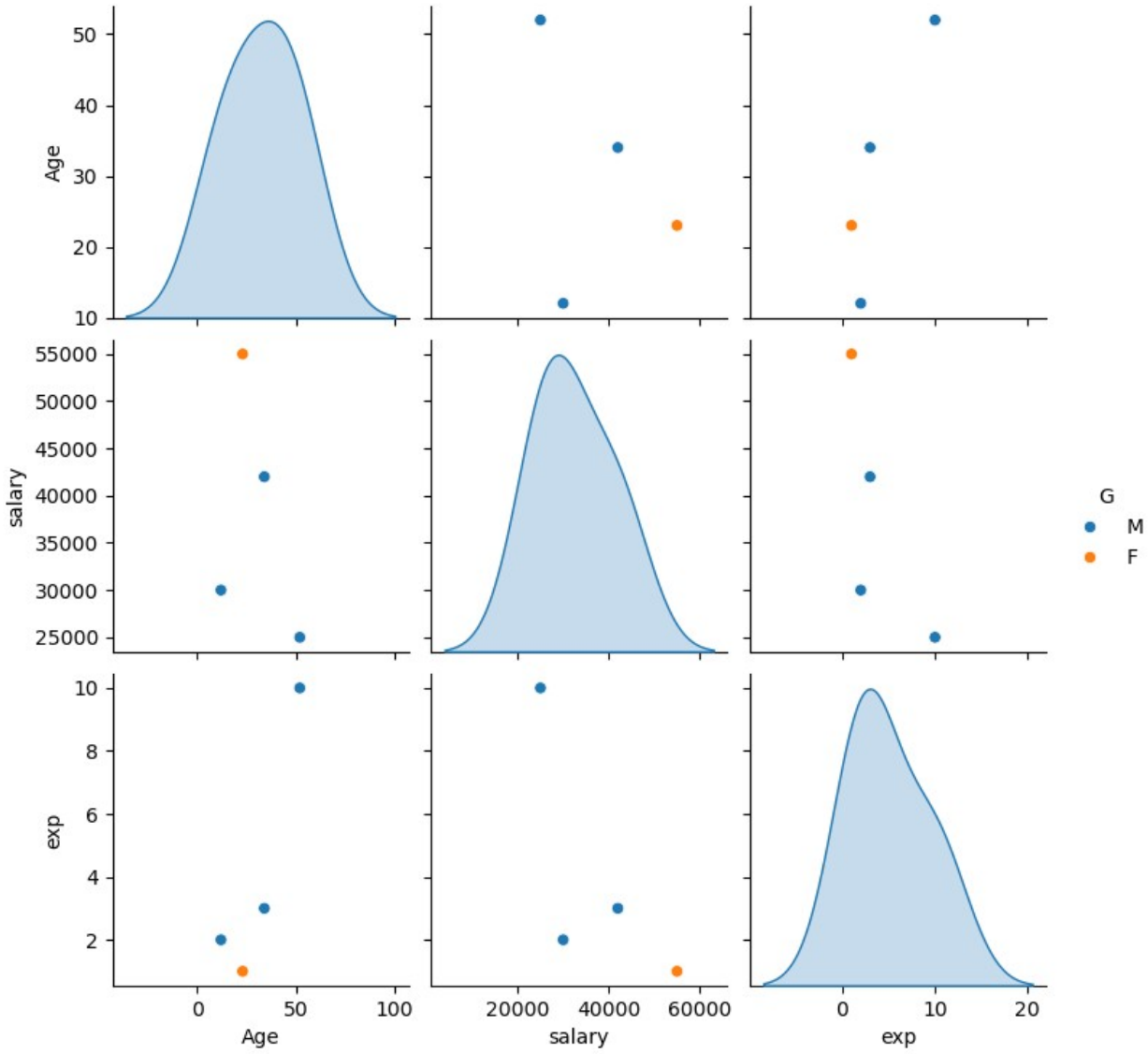
```
Text(0.5, 1.0, 'count of exp')
```



pair plot

```
sns.pairplot(df,hue ='G')
```

```
<seaborn.axisgrid.PairGrid at 0x26925c68740>
```


Age
salary
exp
G
M
F
0
50
100
20000
40000
60000
0
10
20