

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

mydata={"Name":["Ram","Sam","Joe","Asha"],
        "Age":[22,23,47,26],
        "Salary":[20000,45000,65000,38000],
        "Exp":[2,4,2,1]}
```

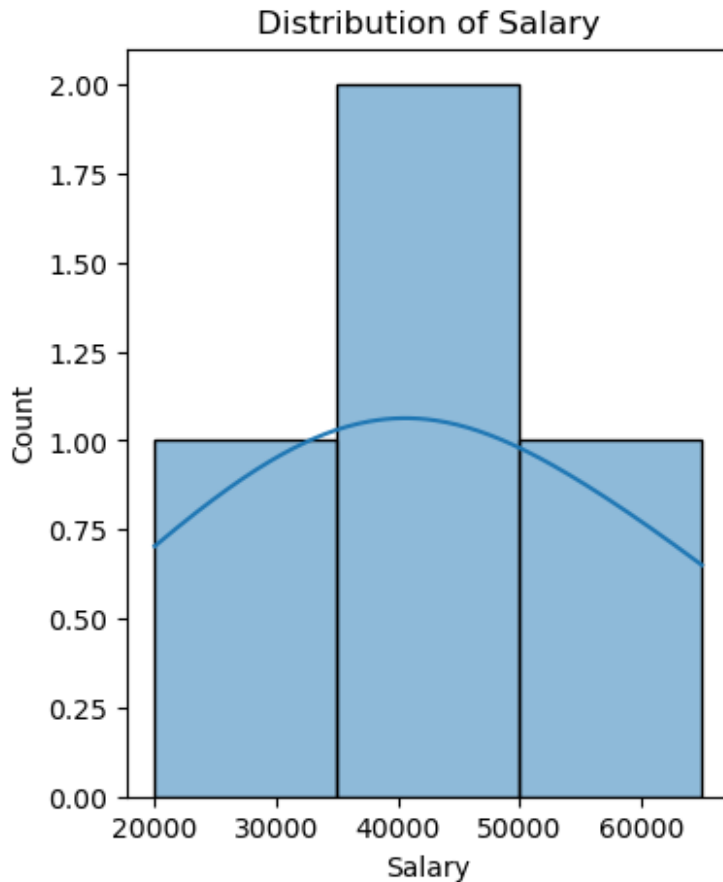
```
df=pd.DataFrame(mydata)
```

```
df
```

	Name	Age	Salary	Exp
0	Ram	22	20000	2
1	Sam	23	45000	4
2	Joe	47	65000	2
3	Asha	26	38000	1

Histogram

```
plt.figure(figsize=(4,5))
sns.histplot(df["Salary"],kde=True,bins=3)
plt.title("Distribution of Salary")
plt.show()
```



1.The salary values are evenly distributed 2.No outlier detected 3.Majority of the salary values are between 40000 and 50000 4.The average of the salary is around 40000

Correlation Matrix

```
ndf=df.select_dtypes(include=["number"])
ndf.head()
```

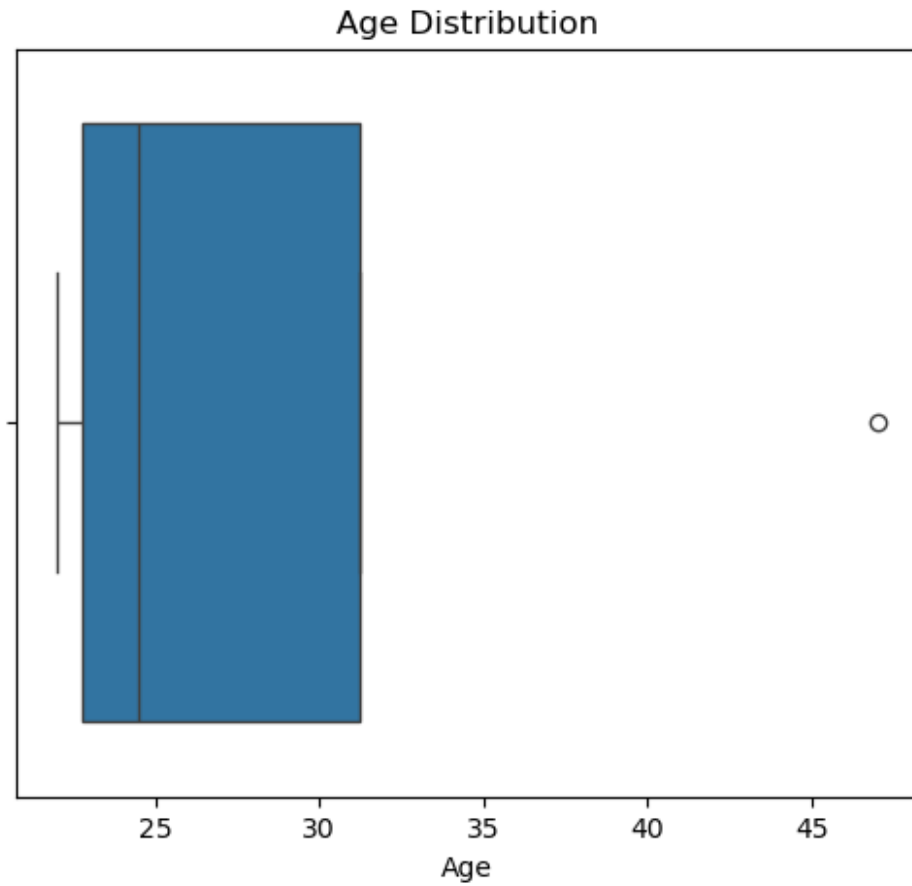
	Age	Salary	Exp
0	22	20000	2
1	23	45000	4
2	47	65000	2
3	26	38000	1

```
plt.figure(figsize=(6,5))
sns.heatmap(ndf.corr(),cmap="viridis",annot=True)
plt.title("Correletion Matrix of Age,Salary and Experiance")
plt.show()
```



1.We can say that Age,Salary and Experience 2.Age affects the experience and Experinece affects the Salary

```
plt.figure(figsize=(6,5))  
sns.boxplot(x=df["Age"])  
plt.title("Age Distribution")  
plt.show()
```

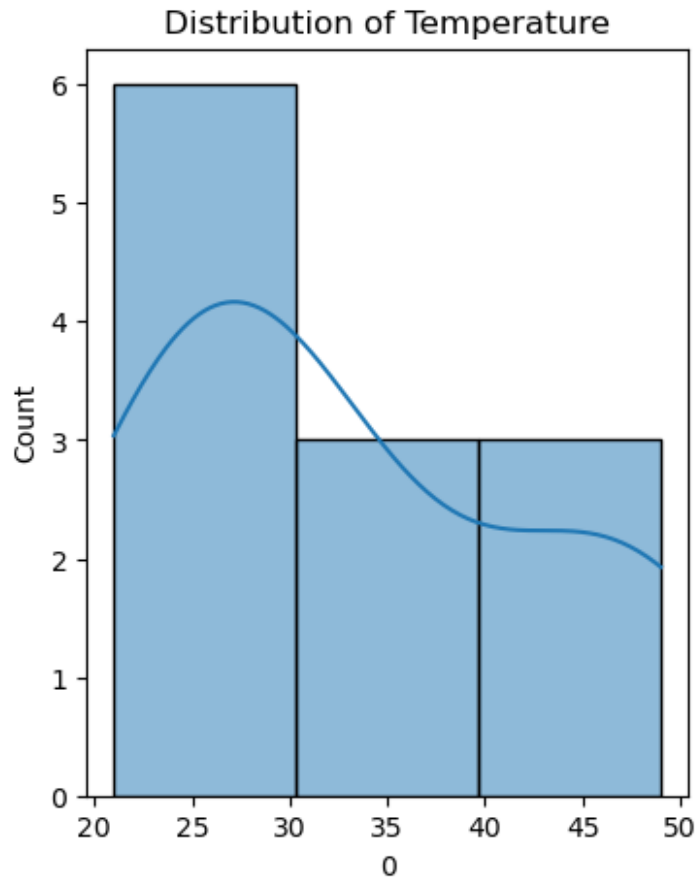


1. Abnormal outlier is around 45
2. The average age is around 25

```
temp = [21,47,39,22,31,33,29,26,27,25,49,46]
ad=pd.DataFrame(temp)
ad
```

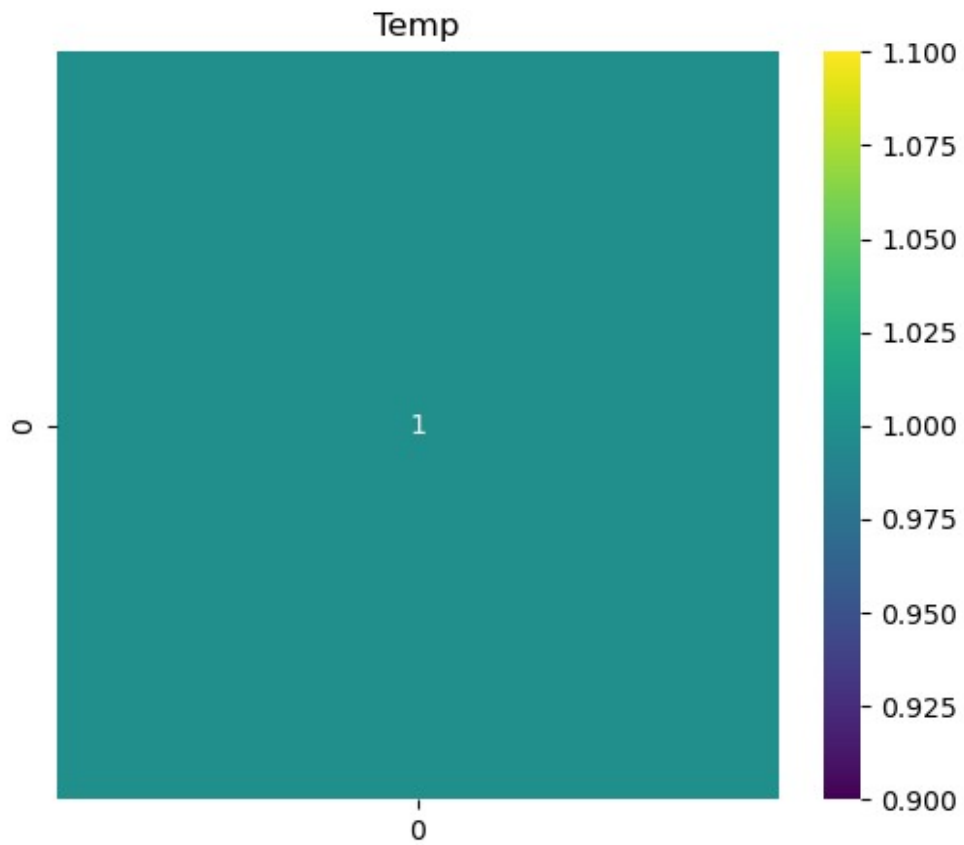
```
0
0  21
1  47
2  39
3  22
4  31
5  33
6  29
7  26
8  27
9  25
10 49
11 46
```

```
plt.figure(figsize=(4,5))
sns.histplot(ad[0],kde=True,bins=3)
plt.title("Distribution of Temperature")
plt.show()
```

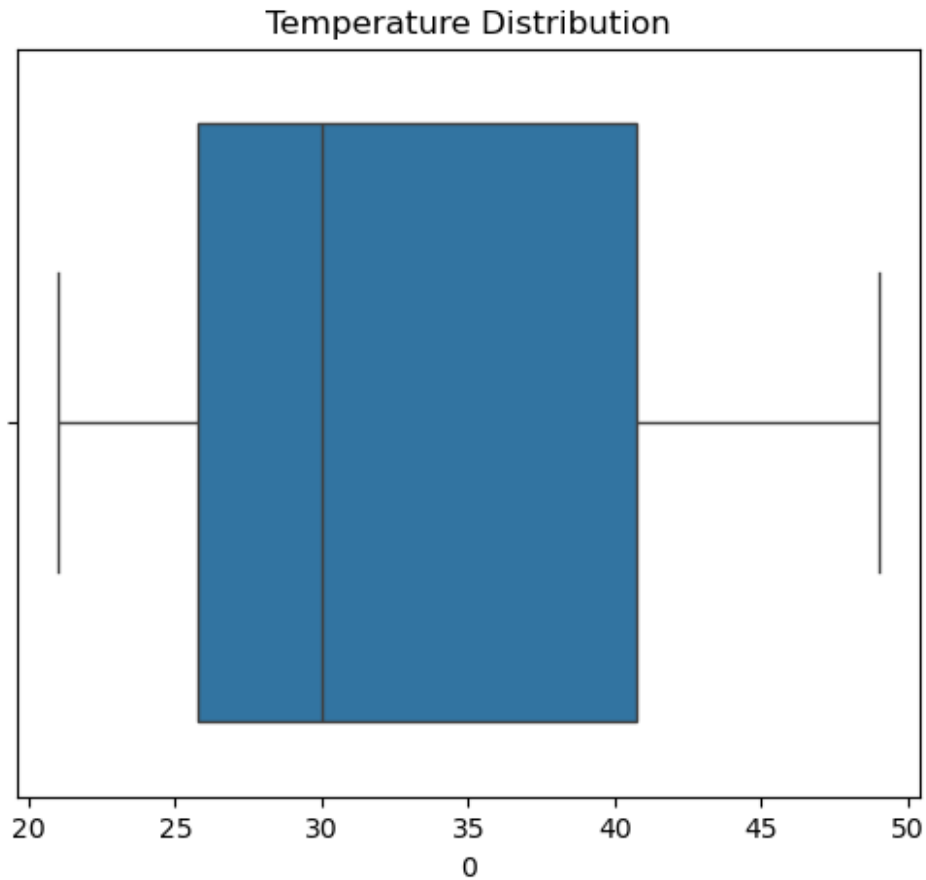


1.positive skew 2.No outlier detected

```
plt.figure(figsize=(6,5))
sns.heatmap(ad.corr(),cmap="viridis",annot=True)
plt.title("Temp")
plt.show()
```



```
plt.figure(figsize=(6,5))
sns.boxplot(x=ad[0])
plt.title("Temperature Distribution")
plt.show()
```



1. Abnormal outlier is not detected 2. The average Temperature is 30

```
mydata1={"Name":["Ram","Sam","Joe","Asha"],
         "Age":[22,23,47,26],
         "Salary":[20000,45000,65000,38000],
         "Exp":[2,4,2,1],
         "G":["Male","Male","Female","Male"]}
```

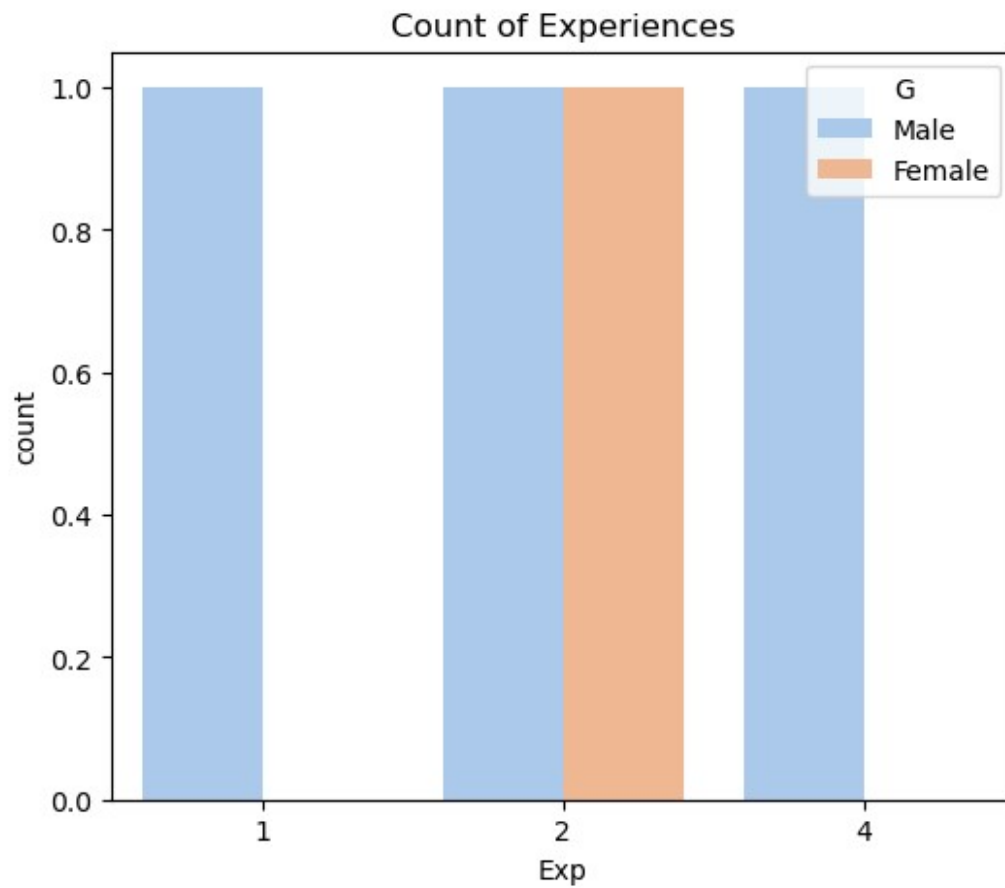
```
df1=pd.DataFrame(mydata1)
df1
```

	Name	Age	Salary	Exp	G
0	Ram	22	20000	2	Male
1	Sam	23	45000	4	Male
2	Joe	47	65000	2	Female
3	Asha	26	38000	1	Male

Count plot

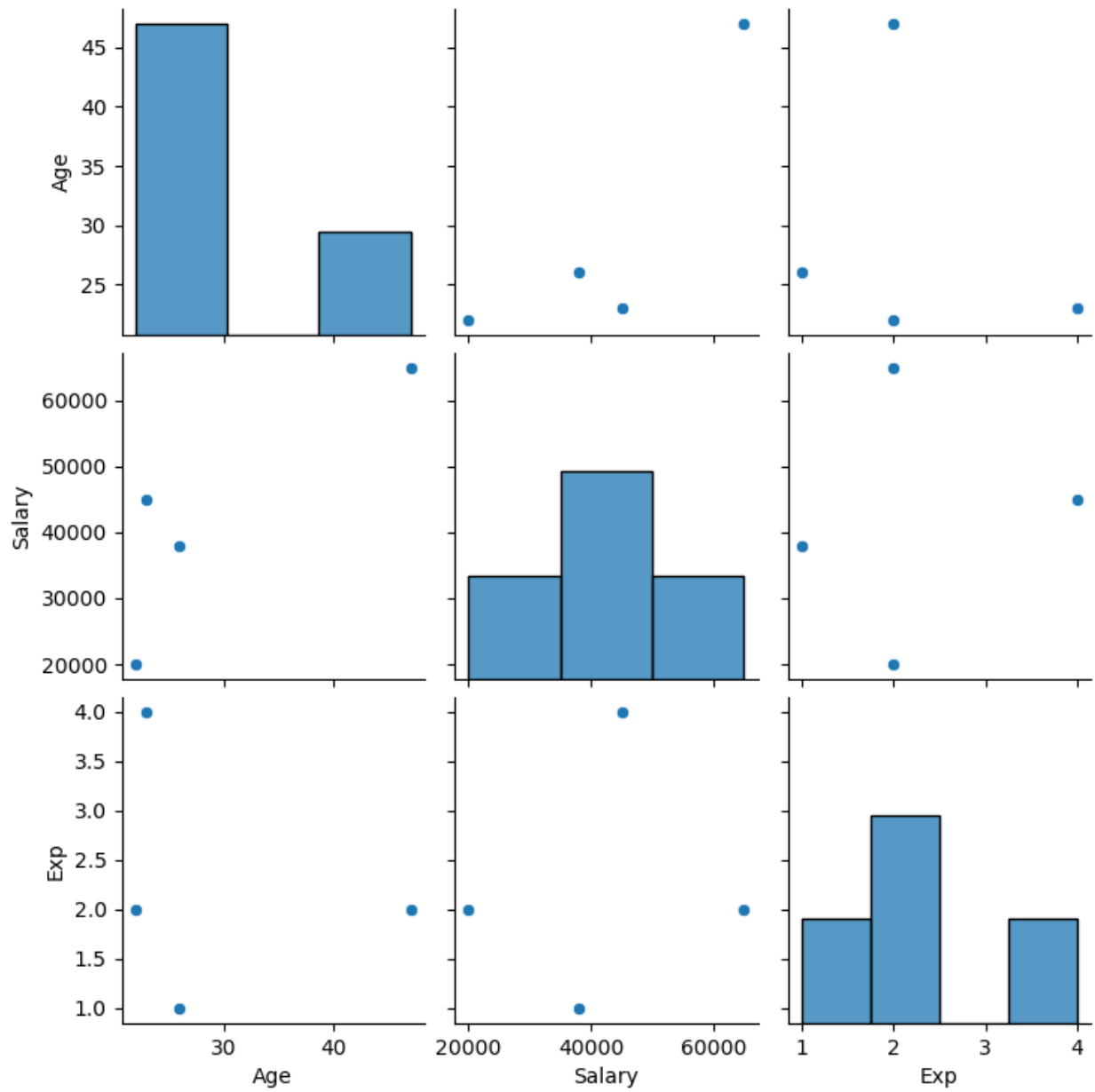
```
plt.figure(figsize=(6,5))
sns.countplot(x=df1["Exp"],palette="pastel",hue=df1["G"])
```

```
plt.title("Count of Experiences")  
plt.show()
```

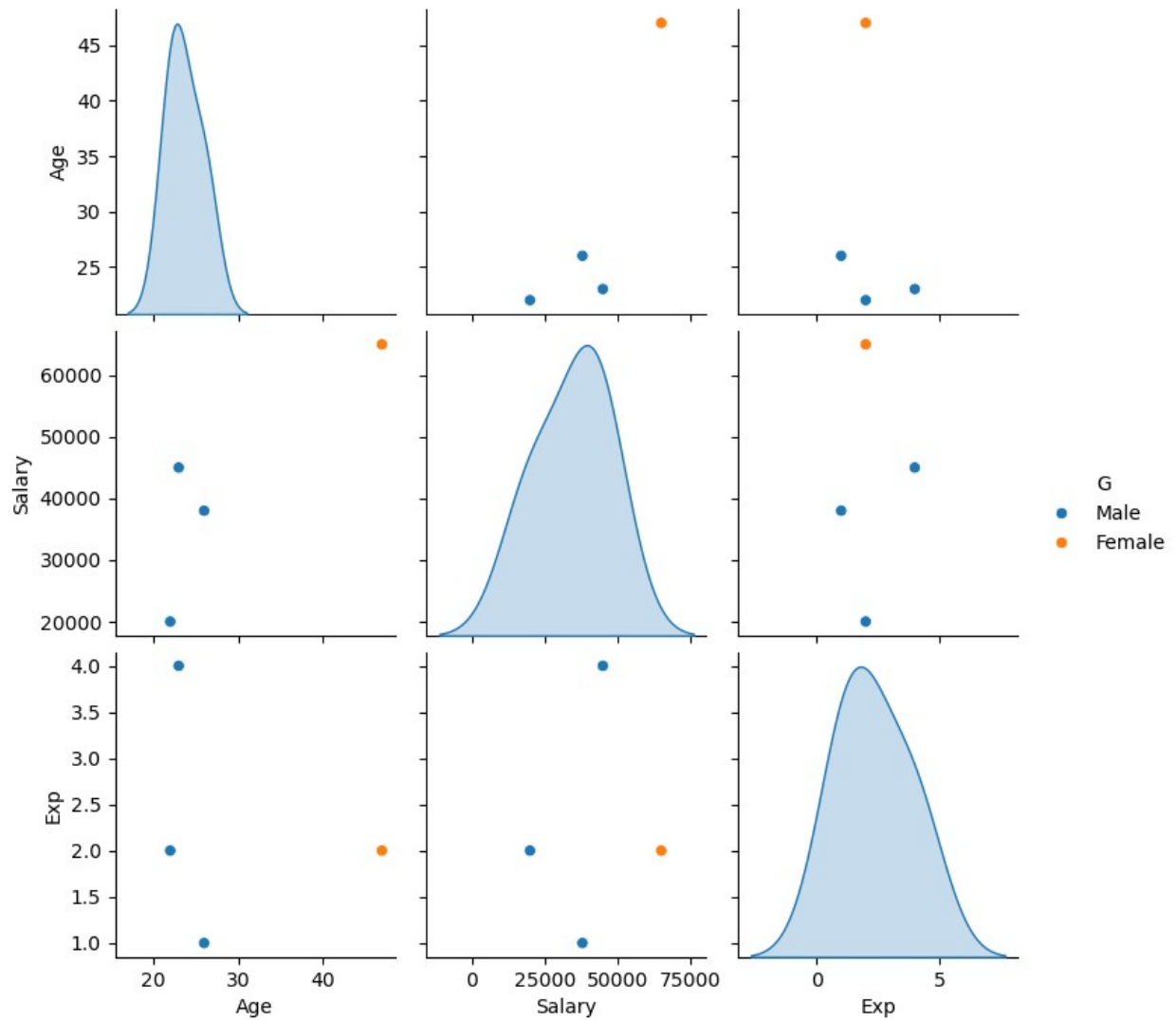


Pair plot

```
sns.pairplot(df1)  
<seaborn.axisgrid.PairGrid at 0x26105c68d40>
```

```
sns.pairplot(df1,hue="G")  
<seaborn.axisgrid.PairGrid at 0x26105d95160>
```



1. It shows the scatter plot matrix visualizing the relationship between Age, Experience and Salary