Seaborn

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

mydata={'Name' :['Ram','Sam','Joe','Asha'],
        'Age' :[23,22,26,47],
        'Salary':[1200,50000,56000,54000],
        'Exp':[2,1,3,10]
       }

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

```
    Name  Age  Salary  Exp
0    Ram   23    1200    2
1    Sam   22   50000    1
2    Joe   26   56000    3
3   Asha   47   54000   10
```

Histgram

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

```
C:\Users\DELL\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed
in a future version. Convert inf values to NaN before operating
instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

Distibution of Salary
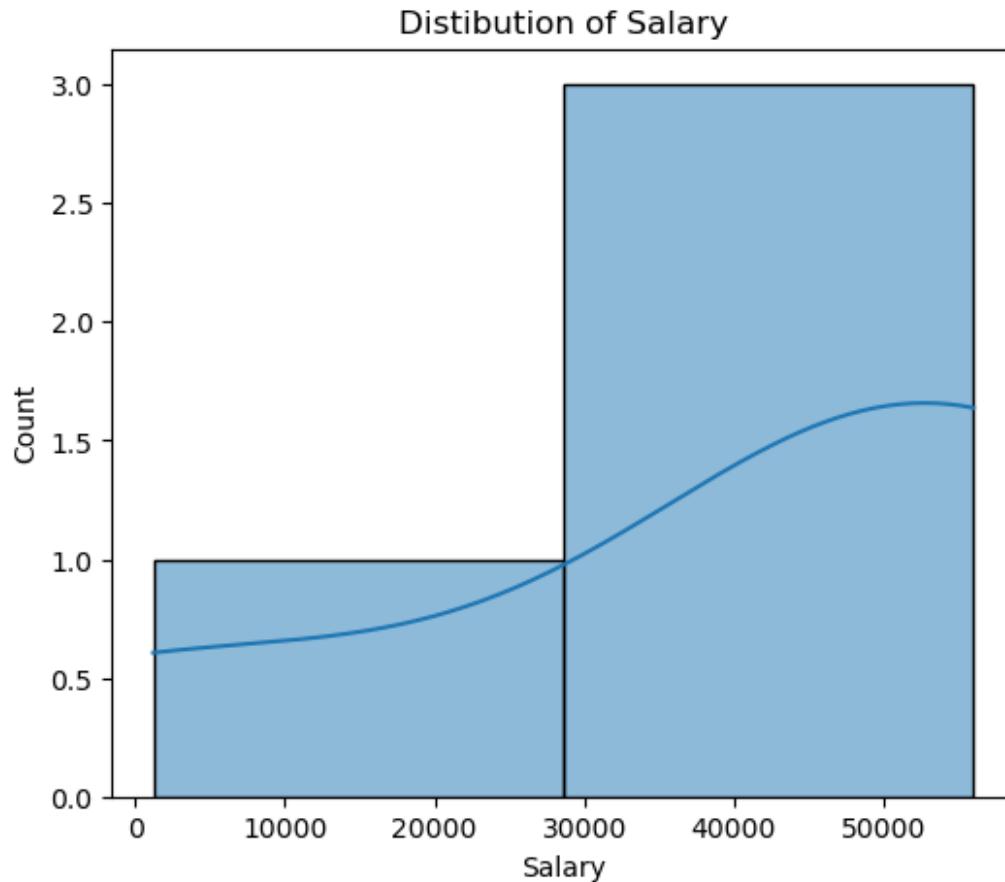
1.There is an negative skew, lower salary,below average 2.The average Salary is 30000 3.There is no outerliers 4.Majority salary is blw 28000 to 56000

Correlation Matrix(heatmap)
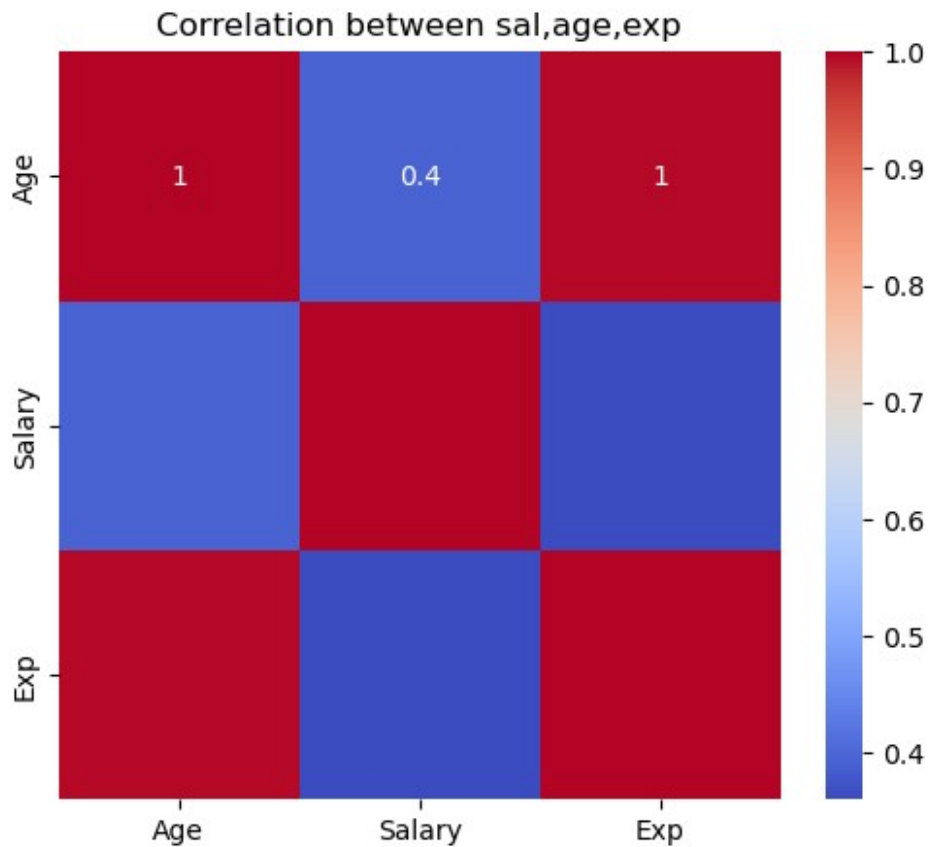
```python
#step1: filter the numerical value
ndf=df.select_dtypes(include =['number'])
ndf.head()
```

```
    Age  Salary  Exp
0    23    1200    2
1    22   50000    1
2    26   56000    3
3    47   54000   10
```
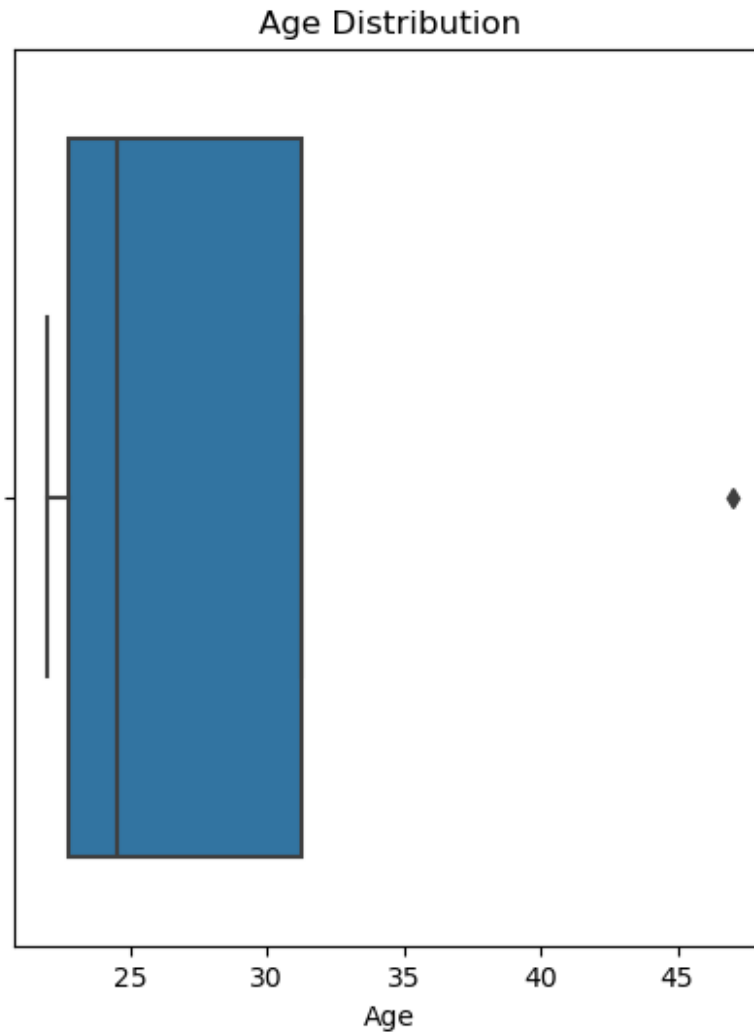
```python
#step2: heatmap
plt.figure(figsize=(6,5))
sns.heatmap(ndf.corr(), cmap='coolwarm', annot=True)
plt.title('Correlation between sal,age,exp')
plt.show()
```

Correlation between sal,age,exp

1.Dark area more correlated 2.light coloured area are less correlated

Box plot

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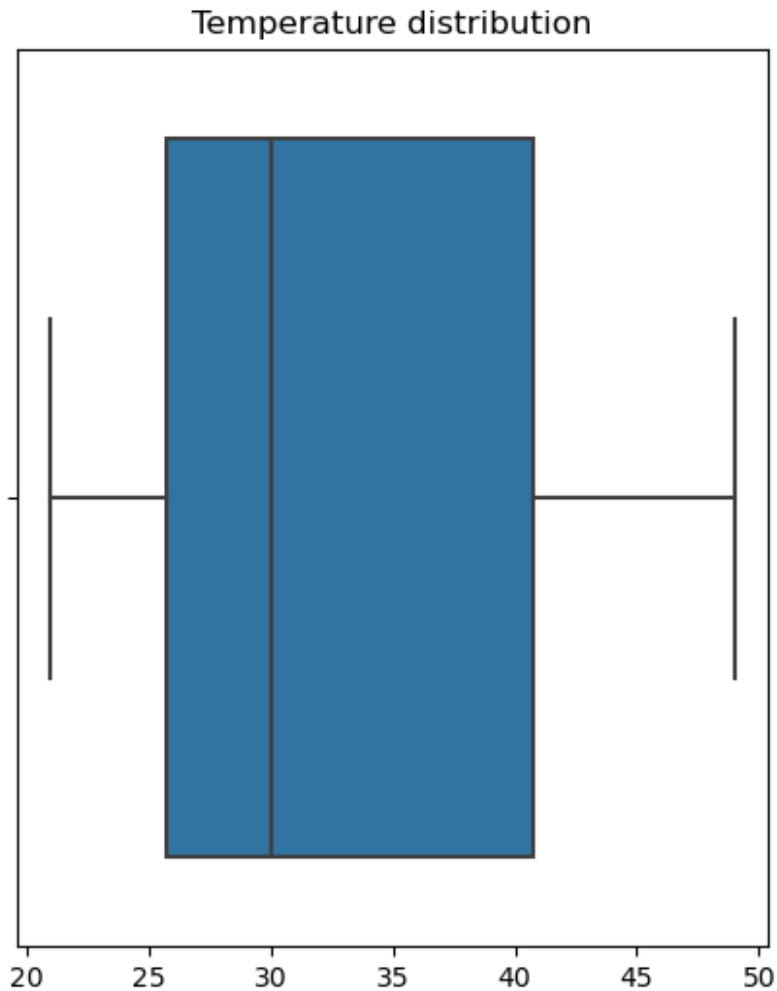
Age Distribution



1.the average age is 24 2.47 is outlier in the data

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

```python
temp =[21, 47, 39,22 , 31, 33, 29, 26, 27, 25, 49, 46]

plt.figure(figsize=(5,6))
sns.boxplot(x = temp)
plt.title("Temperature distribution")
plt.show()
```

Temperature distribution

```
mydata1={'Name' :['Ram','Sam','Joe','Asha'],
         'Age' :[23,22,26,47],
         'Salary':[1200,50000,56000,54000],
         'Exp':[2,1,3,10],
         'G':['M','F','M','M']
        }
df1=pd.DataFrame(mydata1)

plt.figure(figsize=(6,5))
sns.countplot(x=df1['Exp'],palette='pastel',hue=df1['G'])
plt.title('Counts of exp')
plt.show()
```

Counts of exp
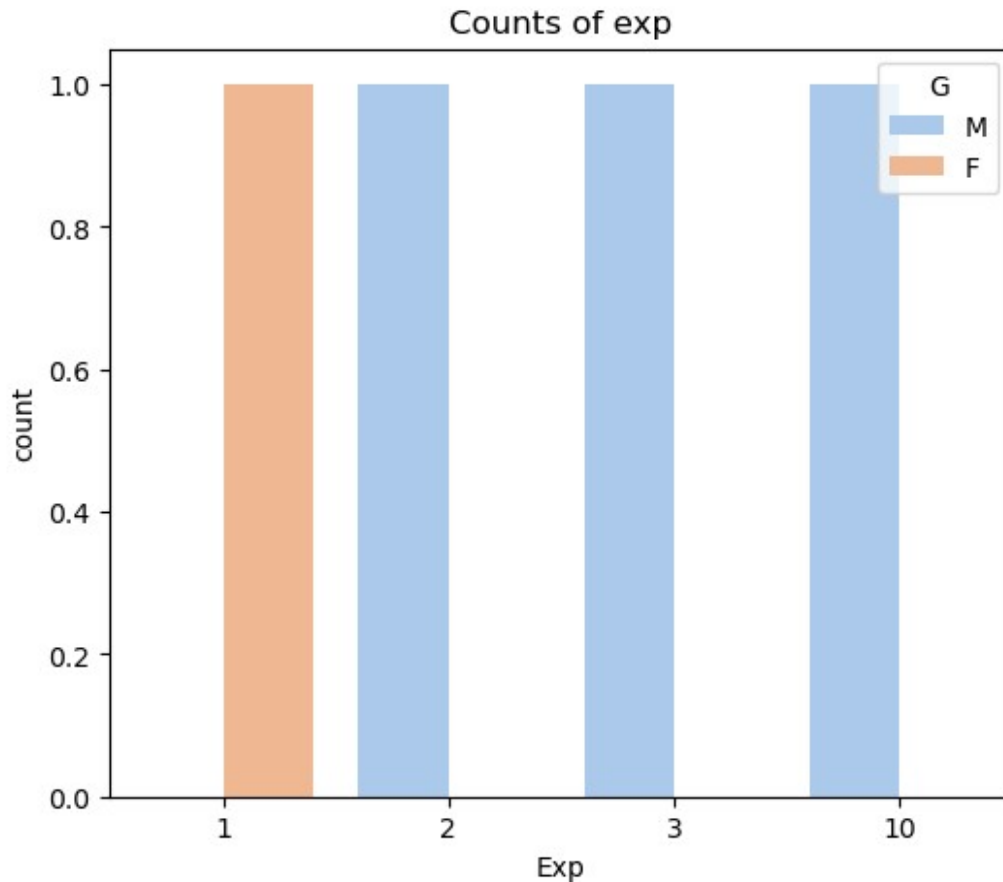
Pair plot

```
sns.pairplot(df1, hue='G')

C:\Users\DELL\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed
in a future version. Convert inf values to NaN before operating
instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\DELL\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed
in a future version. Convert inf values to NaN before operating
instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\DELL\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed
in a future version. Convert inf values to NaN before operating
instead.
  with pd.option_context('mode.use_inf_as_na', True):

<seaborn.axisgrid.PairGrid at 0x2754ab85f50>
```