In [1]:
```
import openpyxl
import pandas as pd
workbook=openpyxl.Workbook()
sheet=workbook.active
data=[
     ['NAME','DOMAIN','AGE','LOCATION','SALARY','EXP'],
       ['ALEX','TESTING',25,'BANGLORE',5000,2],
       ['BARBIE','JAVA',24,'HYDERABAD',10000,3],
       ['CHERRY','PYTHON',26,'CHENNAI',20000,4],
       ['ESWAR','C',25,'PUNE',25000,5],
       ['SIEMEN','DATA ANALYST',27,'BANGLORE',35000]
]
for row in data:
    sheet.append(row)

workbook.save('data.xlsx')
```

In [2]:
```
data
```

Out[2]:
```
[['NAME', 'DOMAIN', 'AGE', 'LOCATION', 'SALARY', 'EXP'],
 ['ALEX', 'TESTING', 25, 'BANGLORE', 5000, 2],
 ['BARBIE', 'JAVA', 24, 'HYDERABAD', 10000, 3],
 ['CHERRY', 'PYTHON', 26, 'CHENNAI', 20000, 4],
 ['ESWAR', 'C', 25, 'PUNE', 25000, 5],
 ['SIEMEN', 'DATA ANALYST', 27, 'BANGLORE', 35000]]
```

In [7]:
```
import os
os.getcwd()
```

Out[7]:
```
'C:\\Users\\manga'
```

In [9]:
```
emp=pd.read_excel(r'C:\Users\manga\Desktop\NIT Class Notes\Assignments\data.xlsx')
emp
```

Out[9]:

	S.NO	NAME	DOMAIN	AGE	LOCATION	SALARY	EXPERIENCE
0	1	ALEX	TESTING	25	BANGLORE	10000	2
1	2	BARBIE	JAVA	24	HYDERABAD	15000	3
2	3	CHERRY	PYTHON	26	CHENNAI	20000	4
3	4	ESWAR	C	25	PUNE	25000	5
4	5	SIEMSON	DATA ANALYST	27	BANGLORE	50000	6

In [10]:
```
emp.shape
```

Out[10]:
```
(5, 7)
```

In [11]:
```
emp.columns
```

Out[11]:
```
Index(['S.NO', 'NAME', 'DOMAIN', 'AGE', 'LOCATION', 'SALARY', 'EXPERIENCE'], dtype
='object')
```

In [12]:
```
len(emp.columns)
```

Out[12]: 7

In [13]:
```
len(emp)
```

Out[13]: 5

In [14]:
```
emp
```

Out[14]:

	S.NO	NAME	DOMAIN	AGE	LOCATION	SALARY	EXPERIENCE
0	1	ALEX	TESTING	25	BANGLORE	10000	2
1	2	BARBIE	JAVA	24	HYDERABAD	15000	3
2	3	CHERRY	PYTHON	26	CHENNAI	20000	4
3	4	ESWAR	C	25	PUNE	25000	5
4	5	SIEMSON	DATA ANALYST	27	BANGLORE	50000	6

In [15]:
```
emp['SALARY']
```

Out[15]:
```
0    10000
1    15000
2    20000
3    25000
4    50000
Name: SALARY, dtype: int64
```

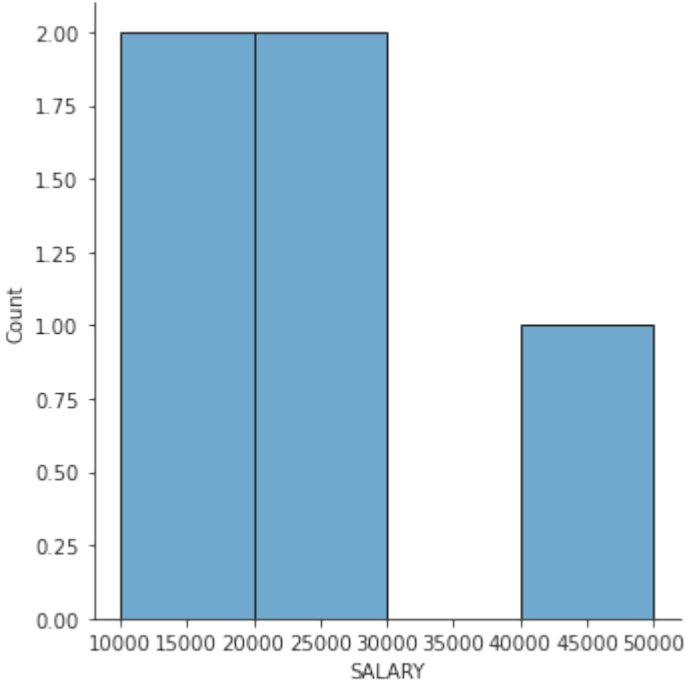
In [19]:
```
emp[['SALARY','EXPERIENCE']]
```

Out[19]:

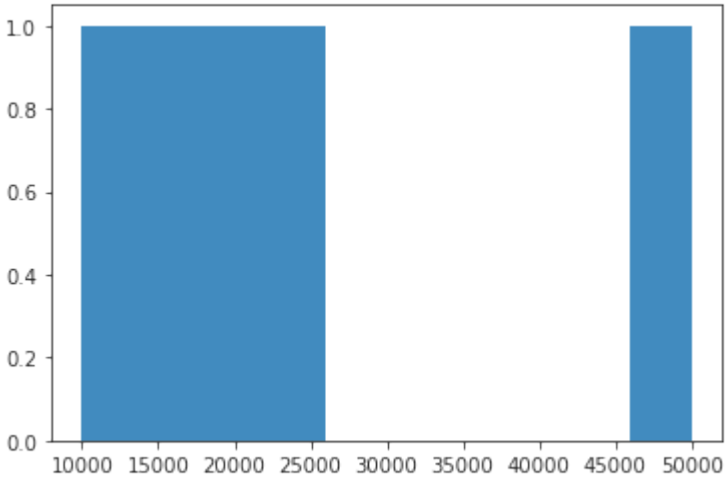
	SALARY	EXPERIENCE
0	10000	2
1	15000	3
2	20000	4
3	25000	5
4	50000	6

In [22]:
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [23]:
```
vis1=sns.displot(emp['SALARY'])
```



In [24]:
```
vis2=plt.hist(emp['SALARY'])
```

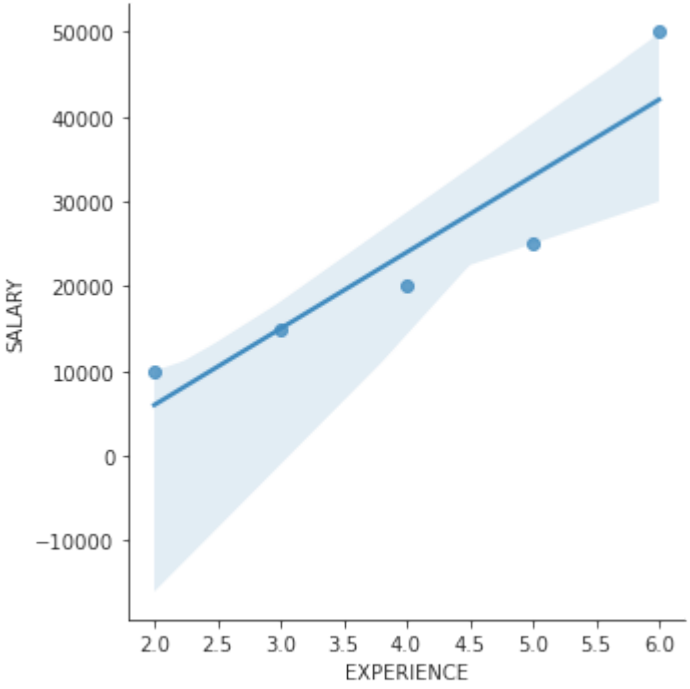


In [27]:
```
plt.rcParams['figure.figsize']=5,1
```

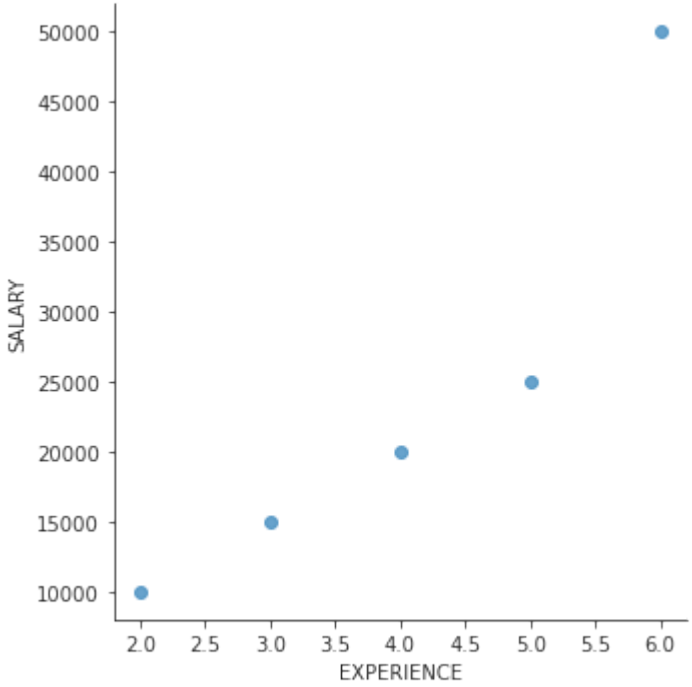
In [28]:
```
vis3=sns.lmplot(data=emp,x='EXPERIENCE',y='SALARY')
emp
```

Out[28]:

	S.NO	NAME	DOMAIN	AGE	LOCATION	SALARY	EXPERIENCE
0	1	ALEX	TESTING	25	BANGLORE	10000	2
1	2	BARBIE	JAVA	24	HYDERABAD	15000	3
2	3	CHERRY	PYTHON	26	CHENNAI	20000	4
3	4	ESWAR	C	25	PUNE	25000	5
4	5	SIEMSON	DATA ANALYST	27	BANGLORE	50000	6


50000
40000
30000
20000
10000
0
−10000
SALARY
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
EXPERIENCE

In [29]:
```
vis4=sns.lmplot(data=emp,x='EXPERIENCE',y='SALARY',fit_reg=False)
```



In [30]:
```
vis4=sns.lmplot(data=emp,x='EXPERIENCE',y='SALARY',fit_reg=True)
```