In [7]:
```
import openpyxl
import pandas as pd

workbook = openpyxl.Workbook()
sheet = workbook.active
data = [
    ['NAME','DOMAIN','AGE','LOCATION','SALARY','EXP'],
    ['ALEX','TESTING',25,'BNG',5000,2],
    ['BARB','JAVA',30,'CHE',10000,3],
    ['CHERRY','C',35,'PUNE',15000,4],
    ['DIPAN','DA',38,'MUMBAI',20000,5],
    ['ESWAR','DS',40,'HYD',50000,6]
]

for row in data:
    sheet.append(row)
workbook.save('data.xslx')
```

In [9]:
```
data
```

Out[9]:
```
[['NAME', 'DOMAIN', 'AGE', 'LOCATION', 'SALARY', 'EXP'],
 ['ALEX', 'TESTING', 25, 'BNG', 5000, 2],
 ['BARB', 'JAVA', 30, 'CHE', 10000, 3],
 ['CHERRY', 'C', 35, 'PUNE', 15000, 4],
 ['DIPAN', 'DA', 38, 'MUMBAI', 20000, 5],
 ['ESWAR', 'DS', 40, 'HYD', 50000, 6]]
```

In [11]:
```
import os
os.getcwd()
```

Out[11]:
```
'C:\\Users\\kavya'
```

In [19]:
```
emp = pd.read_excel(r'C:\\Users\\kavya\\data.xslx')
emp
```

Out[19]:

	NAME	DOMAIN	AGE	LOCATION	SALARY	EXP
0	ALEX	TESTING	25	BNG	5000	2
1	BARB	JAVA	30	CHE	10000	3
2	CHERRY	C	35	PUNE	15000	4
3	DIPAN	DA	38	MUMBAI	20000	5
4	ESWAR	DS	40	HYD	50000	6

In [21]:
```
emp.shape
```

Out[21]:
```
(5, 6)
```

In [23]:
```
emp.columns
```

Out[23]:
```
Index(['NAME', 'DOMAIN', 'AGE', 'LOCATION', 'SALARY', 'EXP'], dtype='object')
```

In [25]:
```
len(emp.columns)
```

Out[25]: 6

In [29]:
```
len(emp)
```

Out[29]: 5

In [31]:
```
emp
```

Out[31]:

	NAME	DOMAIN	AGE	LOCATION	SALARY	EXP
0	ALEX	TESTING	25	BNG	5000	2
1	BARB	JAVA	30	CHE	10000	3
2	CHERRY	C	35	PUNE	15000	4
3	DIPAN	DA	38	MUMBAI	20000	5
4	ESWAR	DS	40	HYD	50000	6

In [33]:
```
emp['SALARY']
```

Out[33]:
```
0     5000
1    10000
2    15000
3    20000
4    50000
Name: SALARY, dtype: int64
```

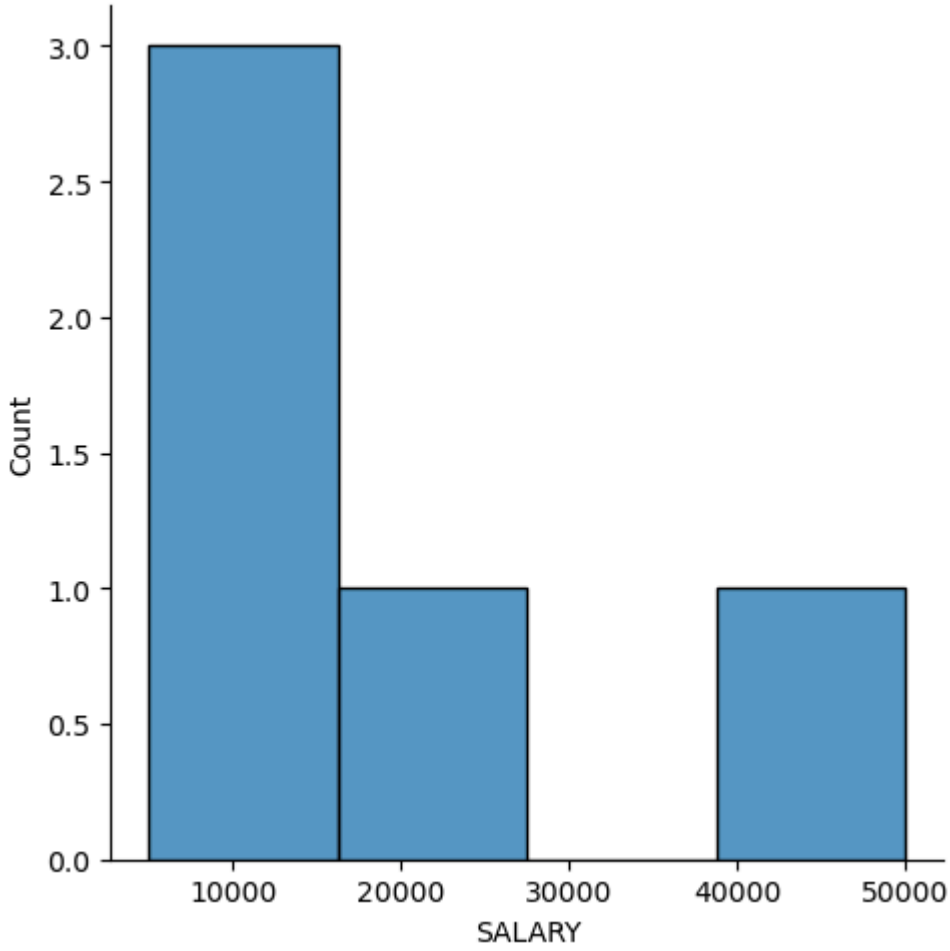
In [35]:
```
emp[['SALARY' , 'EXP']]
```

Out[35]:

	SALARY	EXP
0	5000	2
1	10000	3
2	15000	4
3	20000	5
4	50000	6

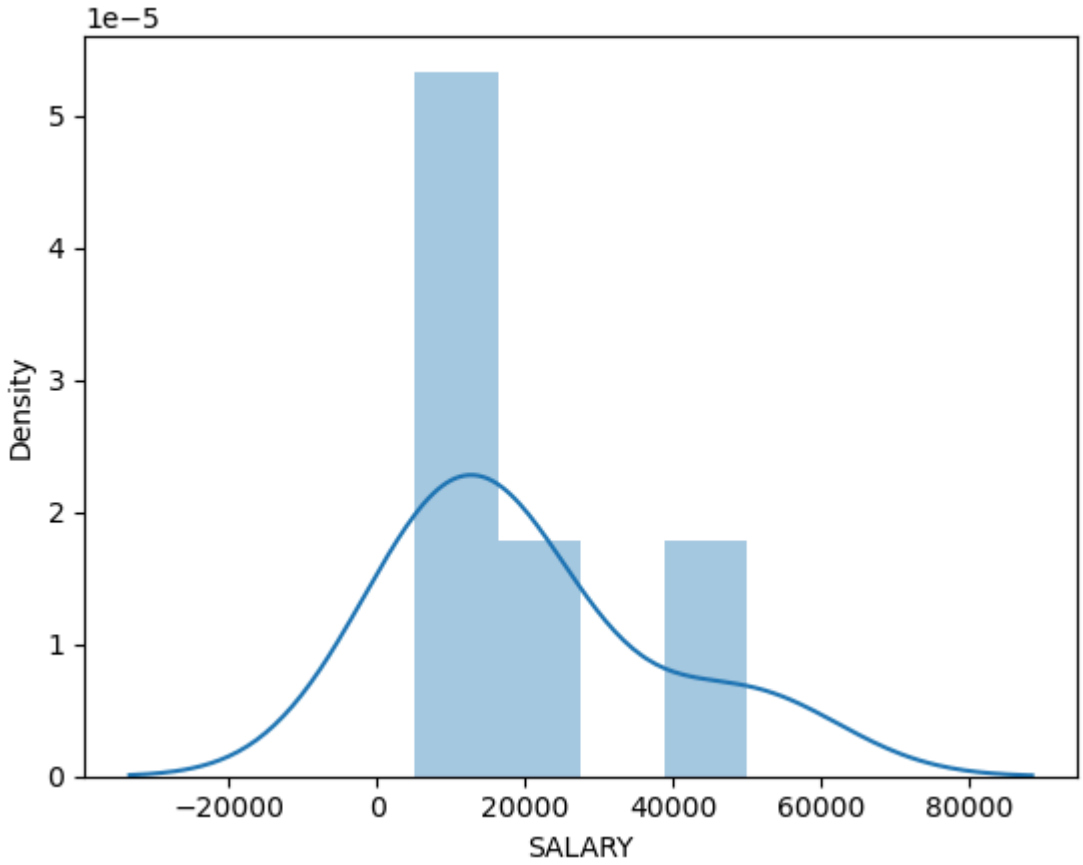
In [37]:
```
import numpy as np # ND array
import matplotlib.pyplot as plt #visualization
import seaborn as sns #statistic visualization
```

In [39]:
```
vis1 = sns.displot(emp['SALARY'])
```

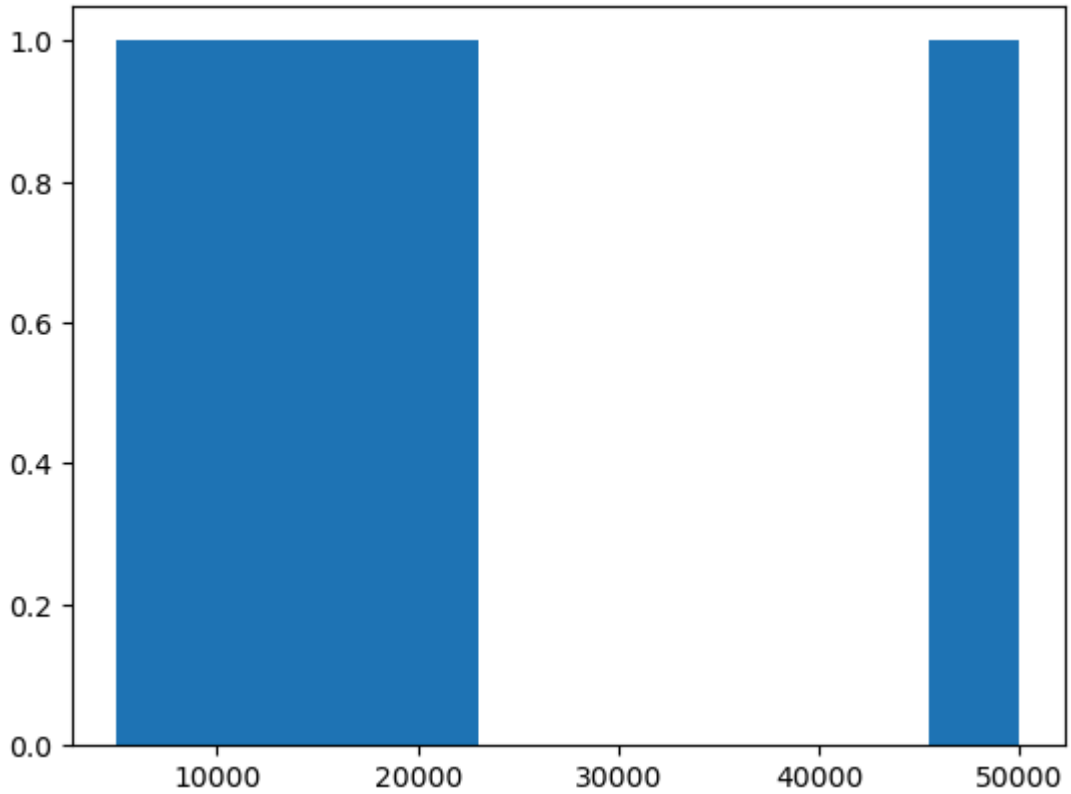


In [43]:
```
vis2 = sns.distplot(emp['SALARY'])
```

```
C:\Users\kavya\AppData\Local\Temp\ipykernel_14856\826855712.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  vis2 = sns.distplot(emp['SALARY'])
```

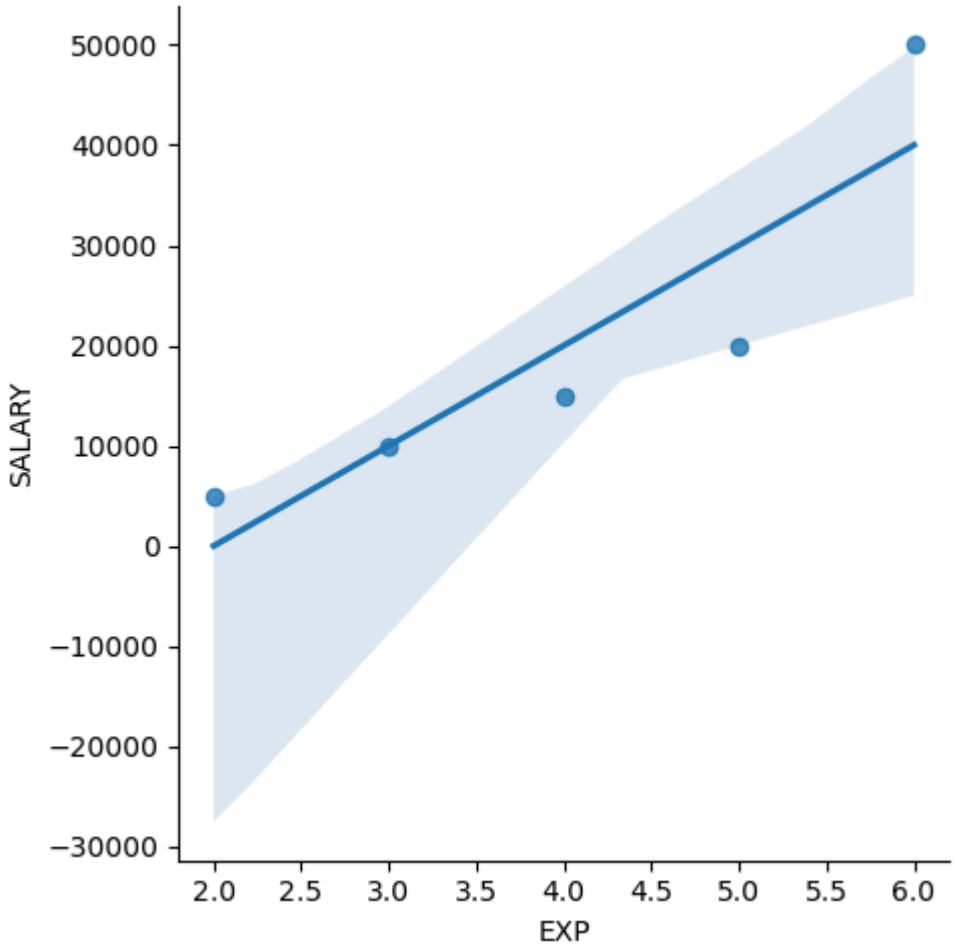


In [45]:
```
vis3 = plt.hist(emp['SALARY'])
```

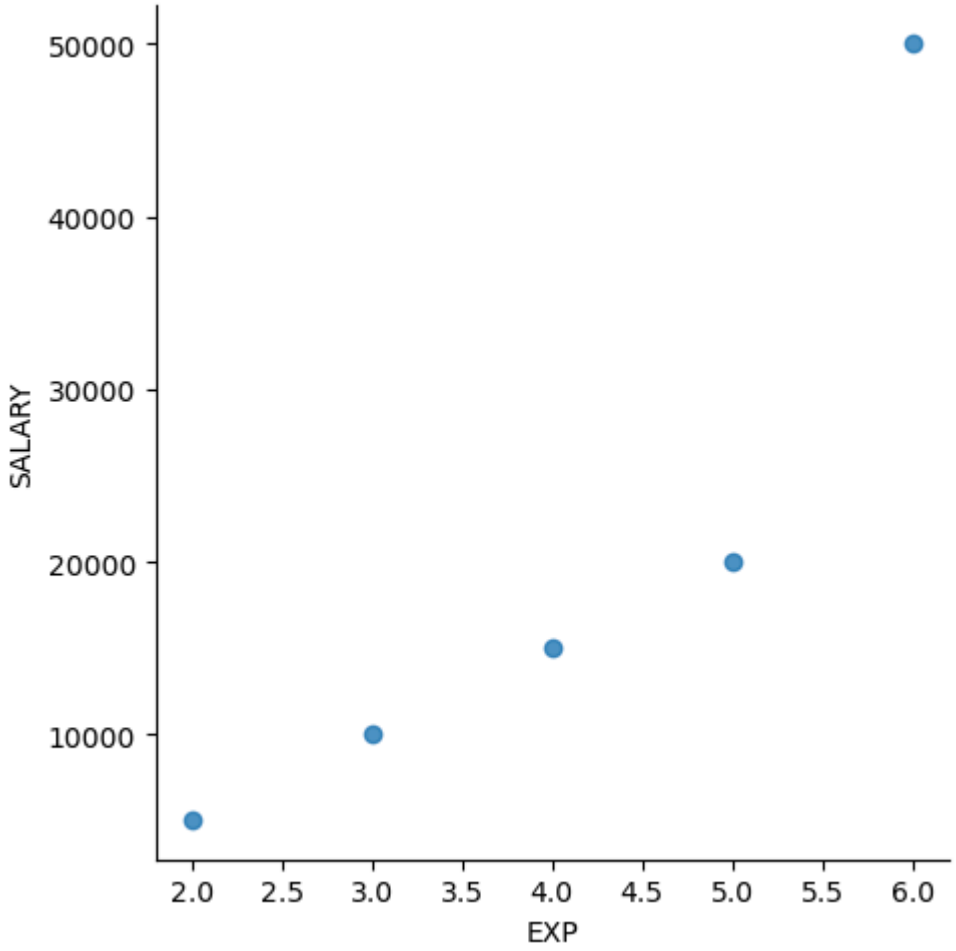


In [49]:
```
plt.rcParams['figure.figsize']=5,1
```

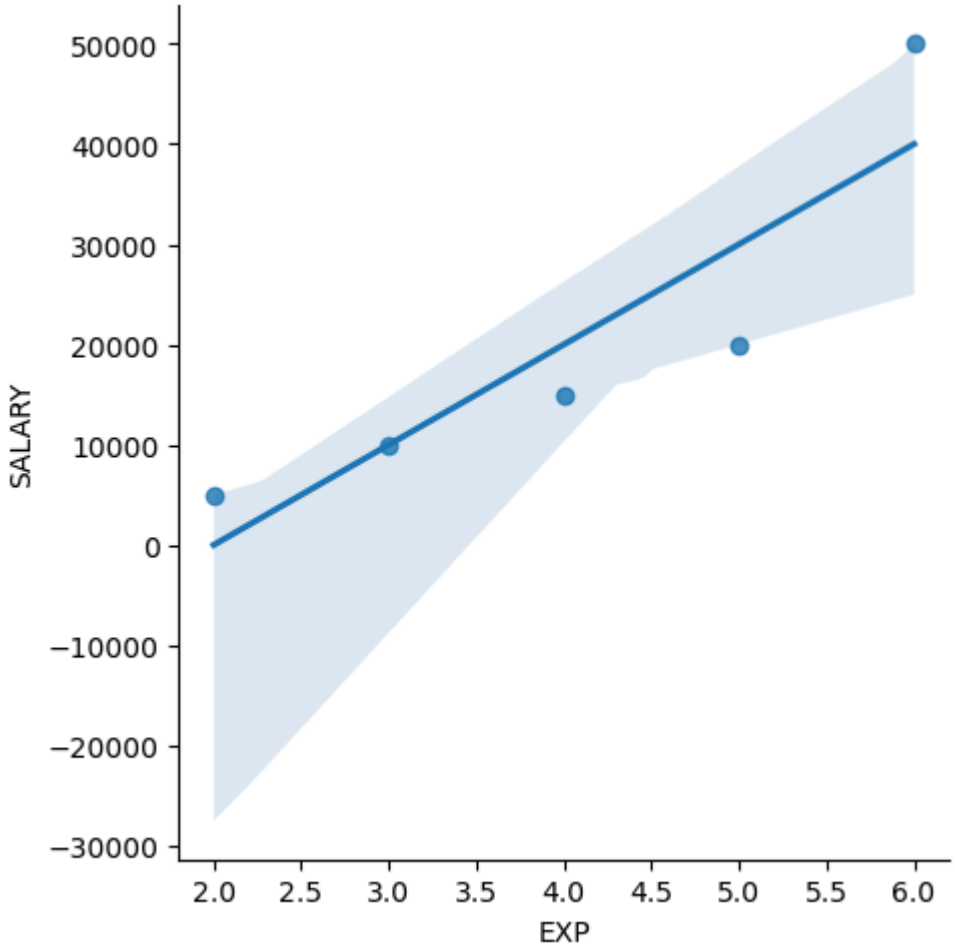
In [51]:
```
vis5 = sns.lmplot(data=emp, x = 'EXP', y = 'SALARY')
```



In [53]:
```
vis5 = sns.lmplot(data=emp, x = 'EXP', y = 'SALARY' , fit_reg = False)
```



In [55]:
```
vis5 = sns.lmplot(data=emp, x = 'EXP', y = 'SALARY' , fit_reg = True)
```


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

In [ ]:

In [ ]:

In [ ]:

In [ ]: