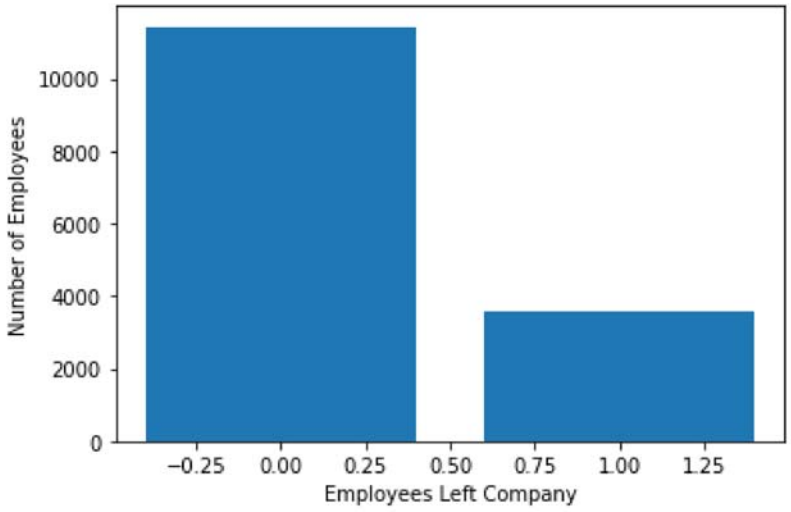
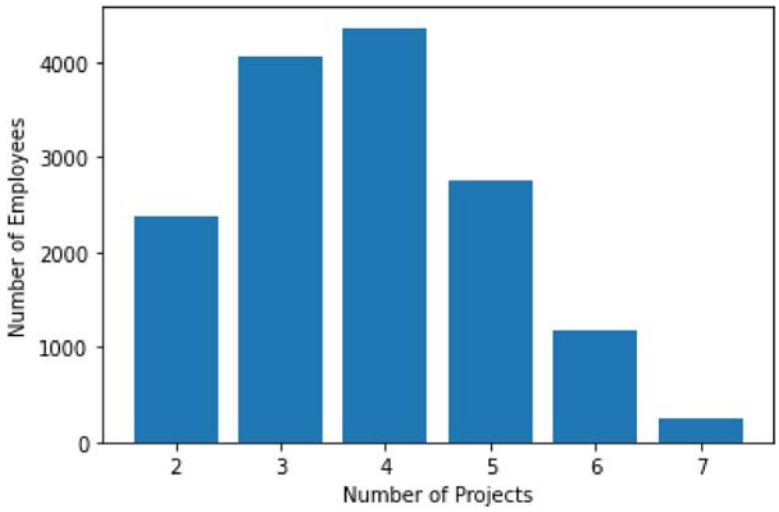
Project Title: Employment Prediction

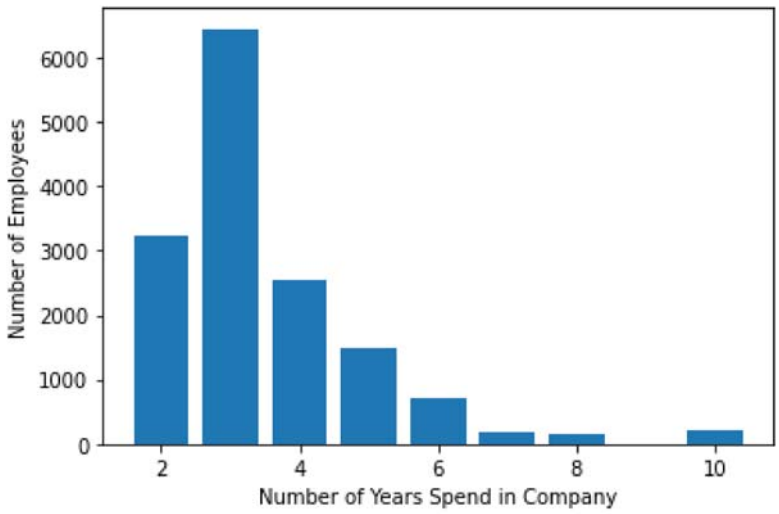
Roll No: 21BCE528, 21BCE526

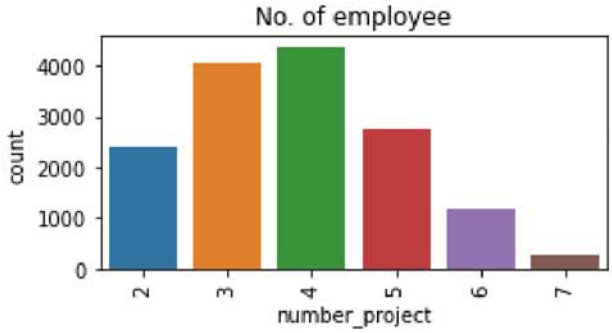
In [ ]:

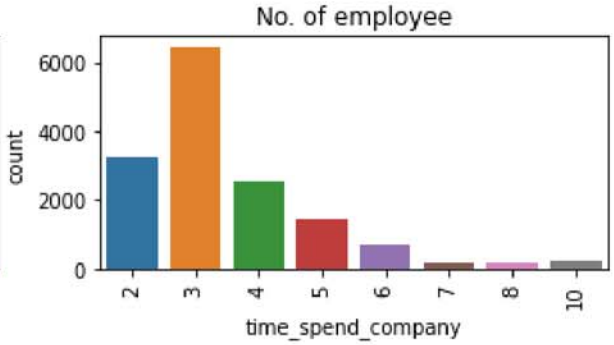
```
import pandas
import matplotlib.pyplot as plt
import seaborn as sns

data=pandas.read_csv('HR_comma_sep.csv')


left = data.groupby('left')
left.mean()

data.describe()

left_count=data.groupby('left').count()
plt.bar(left_count.index.values, left_count['satisfaction_level'])
plt.xlabel('Employees Left Company')
plt.ylabel('Number of Employees')
plt.show()

data.left.value_counts()
num_projects=data.groupby('number_project').count()
plt.bar(num_projects.index.values, num_projects['satisfaction_level'])
plt.xlabel('Number of Projects')
plt.ylabel('Number of Employees')
plt.show()

time_spent=data.groupby('time_spend_company').count()
plt.bar(time_spent.index.values, time_spent['satisfaction_level'])
plt.xlabel('Number of Years Spend in Company')
plt.ylabel('Number of Employees')
plt.show()
features=['number_project','time_spend_company','Work_accident','left', 'promotion_last
fig=plt.subplots(figsize=(10,15))
for i, j in enumerate(features):
    plt.subplot(4, 2, i+1)
    plt.subplots_adjust(hspace = 1.0)
    sns.countplot(x=j,data = data)
    plt.xticks(rotation=90)
    plt.title("No. of employee")
```

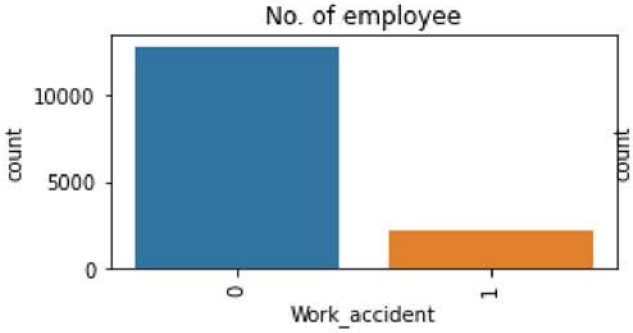

Number of Employees
10000
8000
6000
4000
2000
0
−0.25
0.00
0.25
0.50
0.75
1.00
1.25
Employees Left Company

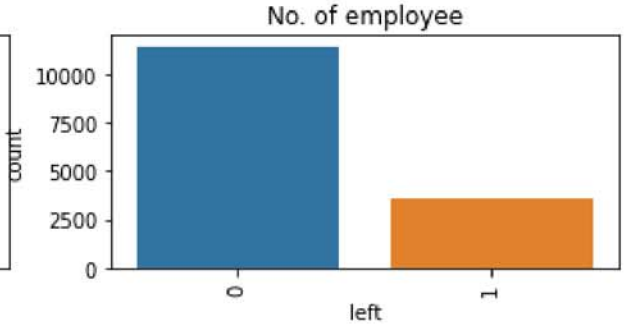

Number of Employees
4000
3000
2000
1000
0
2
3
4
5
6
7
Number of Projects

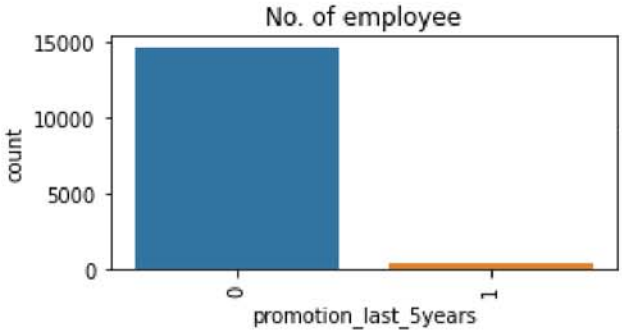

Number of Employees
6000
5000
4000
3000
2000
1000
0
2
4
6
8
10
Number of Years Spend in Company

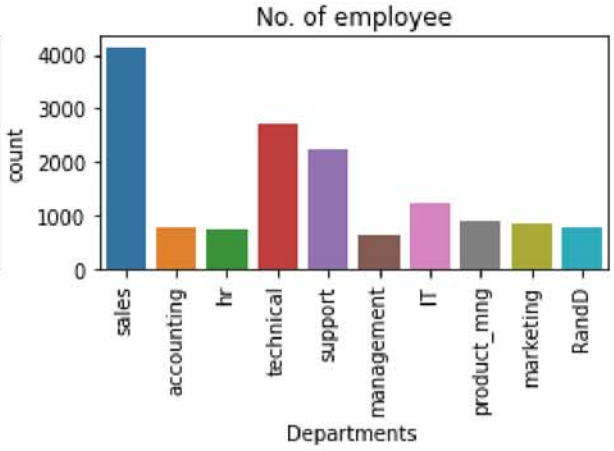

No. of employee
count
number_project
2
3
4
5
6
7

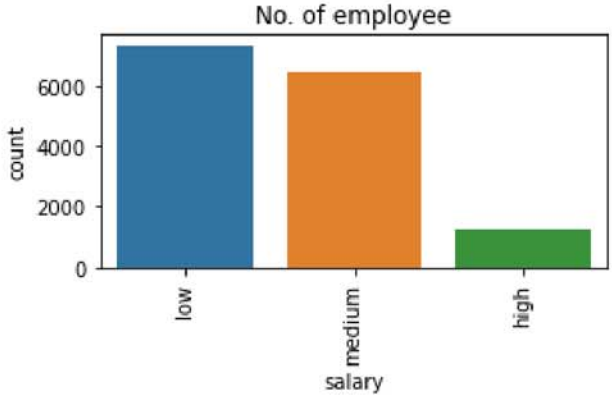

No. of employee
count
time_spend_company
2
3
4
5
6
7
8
10


No. of employee
count
Work_accident
0
1


No. of employee
count
left
0
1


No. of employee
count
promotion_last_5years
0
1


No. of employee
count
Departments
sales
accounting
hr
technical
support
management
IT
product_mng
marketing
RandD


No. of employee
count
salary
low
medium
high