In [1]:

```
import numpy as np
```

In [2]:

```
from matplotlib import pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

```
Matplotlib is building the font cache; this may take a moment.
```

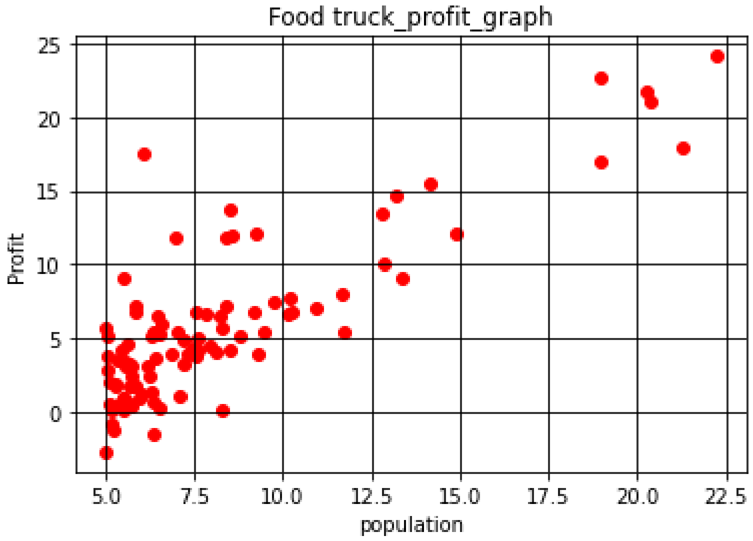
In [9]:

```
data=pd.read_csv("ai dataset.txt")
print(data.shape)
```

```
(97, 2)
```

In [10]:

```
x=data[['population']].values
y=data[['profit']].values
%matplotlib inline
plt.scatter(x,y,c='r',label="scatter_data")
plt.xlabel('population')
plt.ylabel('Profit')
plt.title('Food truck_profit_graph')
plt.grid(True,color='k')
plt.show()
```



In [11]:

```
k=LinearRegression()
k.fit(x,y)
LinearRegression()
print('c value:',k.intercept_)
```

```
c value: [-3.89578088]
```

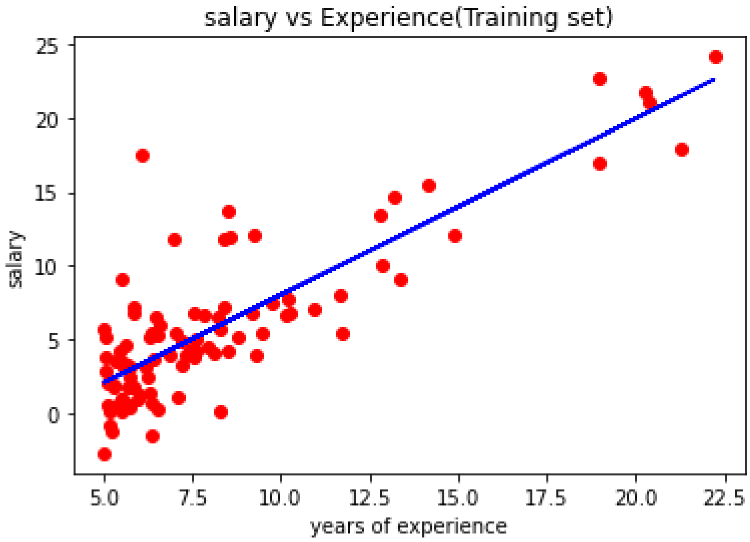
In [12]:

```
print('m value:',k.coef_)
```

```
m value: [[1.19303364]]
```

In [14]:

```
y_pred=k.predict(x)
plt.scatter(x,y,color='red')
plt.plot(x,y_pred,color='blue')
plt.title('salary vs Experience(Training set)')
plt.xlabel('years of experience')
plt.ylabel('salary')
plt.show()
```



In [15]:

```
from sklearn.metrics import r2_score
r_sq=r2_score(y,y_pred)
r_sq
```

Out[15]:

```
0.7020315537841397
```

In [17]:

```
from sklearn.metrics import mean_squared_error
rmse=mean_squared_error(y,y_pred)
rmse
```

Out[17]:

```
8.953942751950358
```

In [18]:

```
n1=4.5
n2=6.5
print("profit from 4500 people city is",k.predict([[n1]])*10000,'$')
print("profit from 6500 people city is",k.predict([[n2]])*10000,'$')
```

```
profit from 4500 people city is [[14728.70520541]] $
profit from 6500 people city is [[38589.37808921]] $
```

In [ ]: