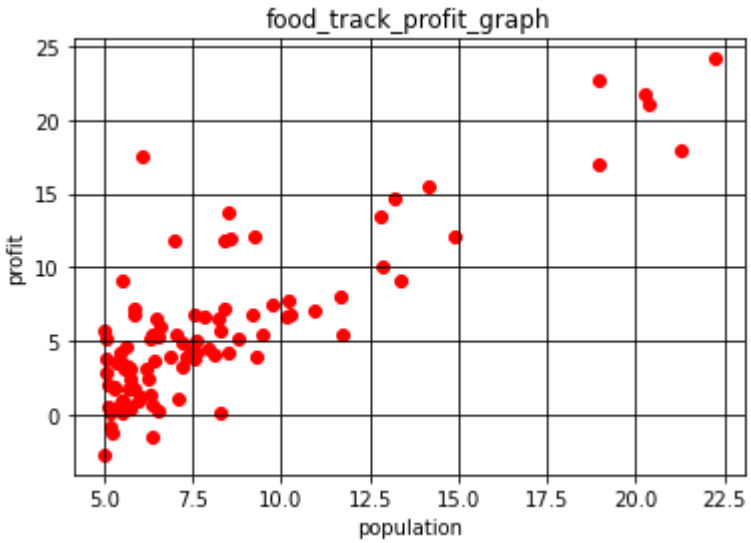
In [11]:

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
data=pd.read_csv("dataset.csv")
print (data.head())
print(data.shape)
```

```
   population   profit
0      6.1101  17.5920
1      5.5277   9.1302
2      8.5186  13.6620
3      7.0032  11.8540
4      5.8598   6.8233
(97, 2)
```

In [12]:

```
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_track_profit_graph')
plt.grid(True,color='k')
plt.show()
```




In [14]:

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

```
c values: [-3.89578088]
```

In [15]:

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

```
m values: [[1.19303364]]
```

In [16]:

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


In [17]:

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

Out[17]:

```
0.7020315537841397
```

In [18]:

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

Out[18]:

```
8.953942751950358
```

In [21]:

```
n1=4.5
n2=6.5
print('profit from 45000 people city is',k.predict([[n1]])*10000,'$')
print('profit from 65000 people city is',k.predict([[n2]])*10000,'$')
```

```
profit from 45000 people city is [[14728.70520541]] $
profit from 65000 people city is [[38589.37808921]] $
```

In [ ]:

```

```