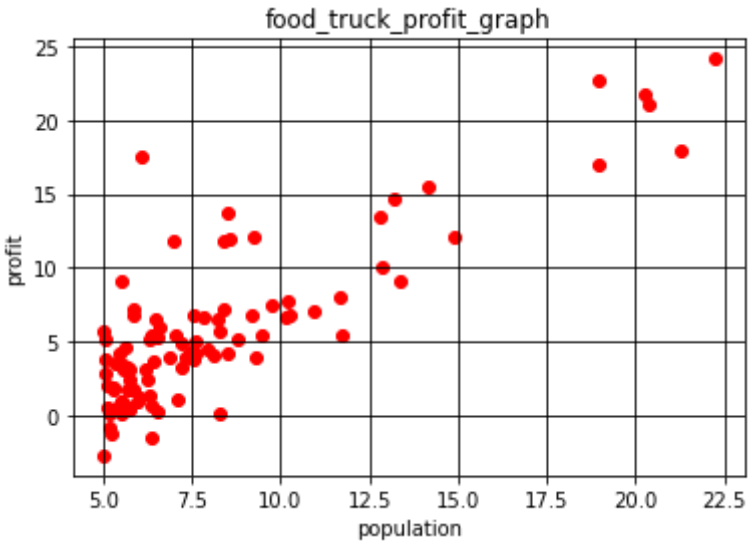
In [12]:
```
import pandas as pd
from matplotlib import pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
data=pd.read_csv("profitdataset.txt")
print(data.shape)
```

```
(97, 2)
```

In [13]:
```
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 [14]:
```
k=LinearRegression()
k.fit(x,y)
```

Out[14]:
```
LinearRegression()
```

In [15]:
```
print('cvalue:',k.intercept_)
```

```
cvalue: [-3.89578088]
```

In [16]:
```
print('mvalue:',k.coef_)
```

```
mvalue: [[1.19303364]]
```

In [18]:
```
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 [19]:
```
from sklearn.metrics import r2_score
r_sq=r2_score(y,y_pred)
r_sq
```

Out[19]:
```
0.7020315537841397
```

In [20]:
```
from sklearn.metrics import mean_squared_error
rmse=mean_squared_error(y,y_pred)
rmse
```

Out[20]:
```
8.953942751950358
```

In [26]:
```
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 [ ]: