In [2]:

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

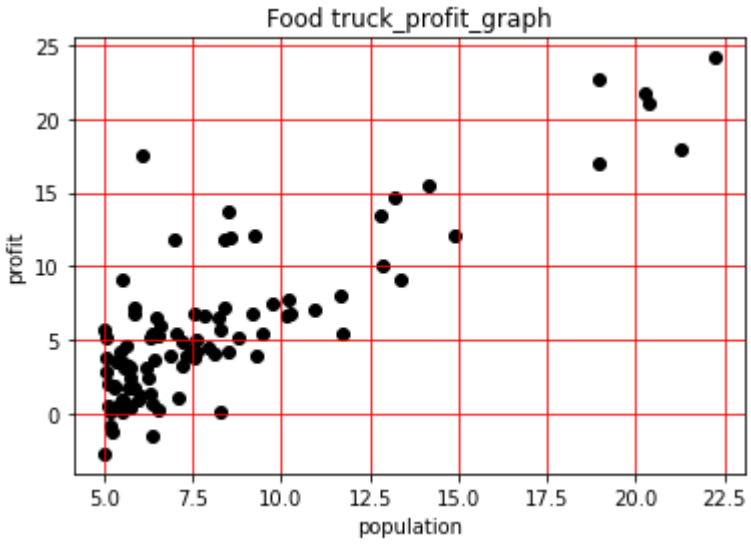
In [4]:

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

```
(97, 2)
```

In [24]:

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



In [6]:

```
k = LinearRegression()
k.fit(x,y)
```

Out[6]:

```
LinearRegression()
```

In [8]:

```
print('c value',k.intercept_)
```

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

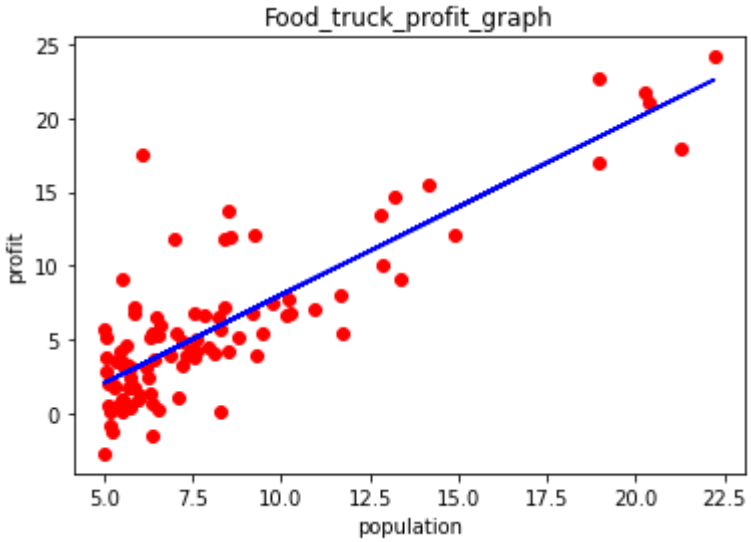
In [10]:

```
print('mvalue',k.coef_)
```

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

In [11]:

```
y_pred = k.predict(x)
plt.scatter(x,y,color = "red")
plt.plot(x,y_pred,color = 'blue')
plt.title("Food_truck_profit_graph")
plt.xlabel('population')
plt.ylabel('profit')
plt.show()
```



In [14]:

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

Out[14]:

```
0.7020315537841397
```

In [16]:

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

Out[16]:

```
8.953942751950358
```

n1 = 4.5 n2 = 6.5

In [17]:

```
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 [18]:

```
print(data)
```

```
    population    profit
0       6.1101  17.59200
1       5.5277   9.13020
2       8.5186  13.66200
3       7.0032  11.85400
4       5.8598   6.82330
..         ...       ...
92      5.8707   7.20290
93      5.3054   1.98690
94      8.2934   0.14454
95     13.3940   9.05510
96      5.4369   0.61705

[97 rows x 2 columns]
```

In [ ]: