In [1]:

```
import pandas as pd
```

In [2]:

```
import numpy as np
```

In [3]:

```
from matplotlib import pyplot as plt
```

In [13]:

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
data = pd.read_csv("profitdataset(1).txt")
print(data.shape)
```

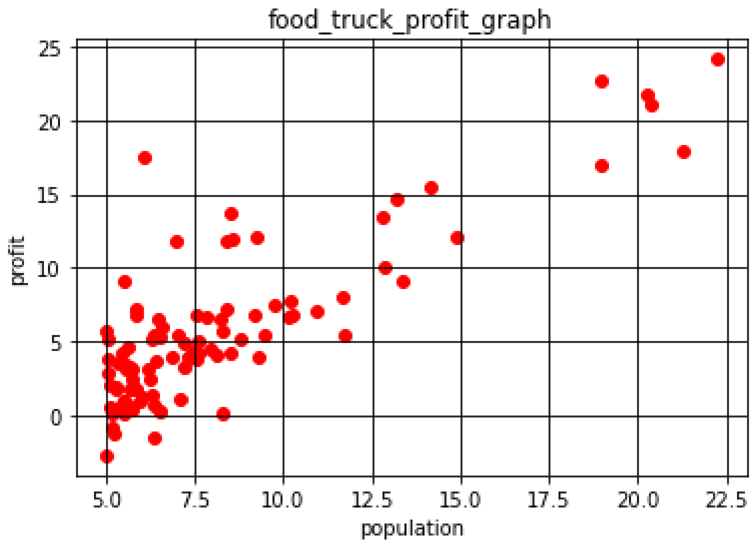
```
(97, 2)
```

In [14]:

```
x = data[["population"]].values
y = data[["profit"]].values
```

In [15]:

```
%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 [16]:

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

Out[16]:

```
LinearRegression()
```

In [17]:

```
print("cvalue:",k.intercept_)
```

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

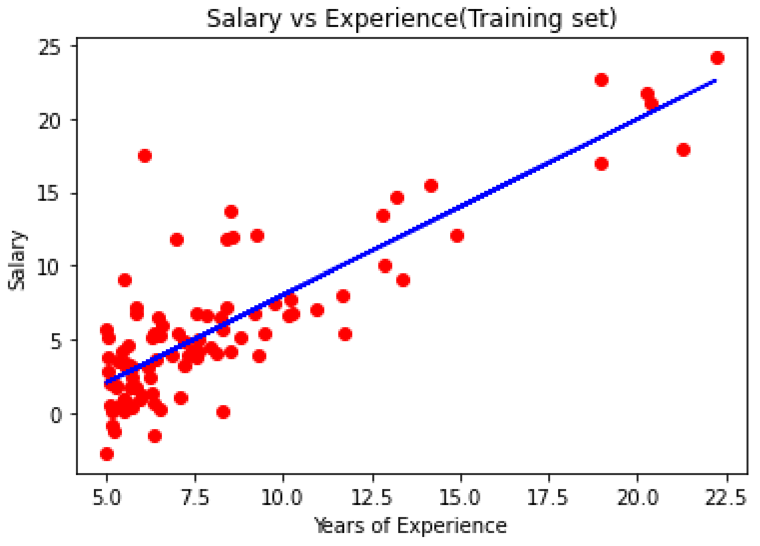
In [18]:

```
print("mvalue:",k.coef_)
```

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

In [19]:

```
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 [20]:

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

Out[20]:

```
0.7020315537841397
```

In [21]:

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

Out[21]:

```
8.953942751950358
```

In [22]:

```
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 [ ]:

```
*
```