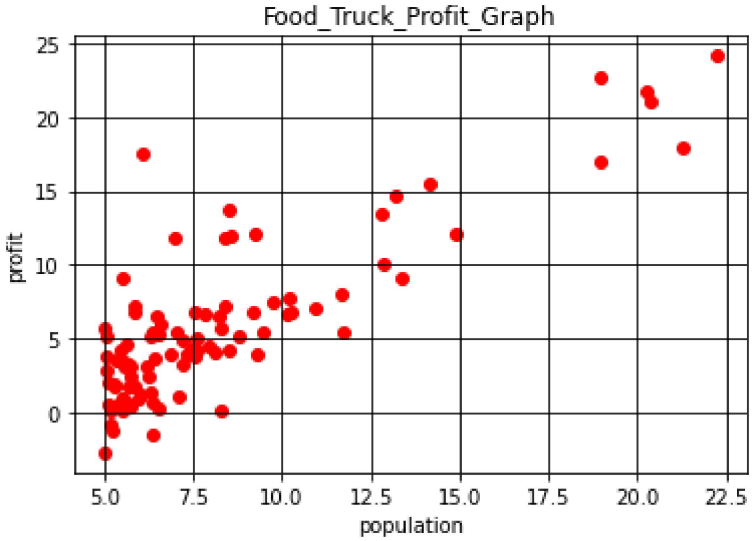
In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

In [2]:
```
data=pd.read_csv("ProfitDataset.txt")
print(data.shape)
```

```
(97, 2)
```

In [3]:
```
x=data[["population"]].values
y=data[["profit"]].values
```

In [5]:
```
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 [6]:
```
k=LinearRegression()
k.fit(x,y)
```

Out[6]:
```
LinearRegression()
```

In [7]:
```
print("C value:",k.intercept_)
```

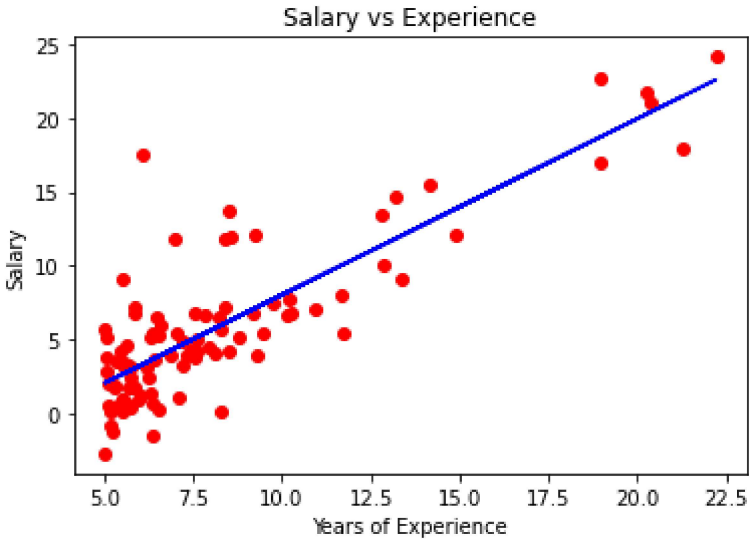
```
C value: [-3.89578088]
```

In [9]:
```
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')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.show()
```



In [16]:
```
from sklearn.metrics import r2_score
rsq=r2_score(y,y_pred)
print(rsq)
```

```
0.7020315537841397
```

In [15]:
```
from sklearn.metrics import mean_squared_error
rmse=mean_squared_error(y,y_pred)
print(rmse)
```

```
8.953942751950358
```

In [18]:
```
n1=4.5
n2=6.5
print("Profit from 45000 people city is:",k.predict([[n1]])*1000,'$')
print("Profit from 65000 people city is:",k.predict([[n2]])*1000,'$')
```

```
Profit from 45000 people city is: [[1472.87052054]] $
Profit from 65000 people city is: [[3858.93780892]] $
```

In [ ]: