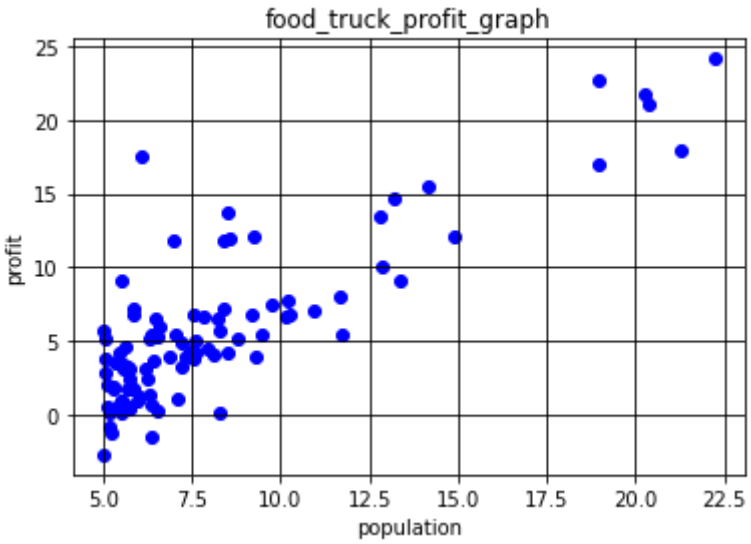
In [6]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
df=pd.read_csv("dataset1.csv")
df
print(df.shape)
```

```
(97, 2)
```

In [7]:
```
print(df.shape)
x=df[['population']].values
y=df[['profit']].values
```

```
(97, 2)
```

In [8]:
```
%matplotlib inline
plt.scatter(x,y,c='b',label='scatter_data')
plt.xlabel("population")
plt.ylabel("profit")
plt.title("food_truck_profit_graph")
plt.grid(True,color="k")
plt.show()
```



In [9]:
```
k=LinearRegression()
k.fit(x,y)
```

Out[9]:
```
LinearRegression()
```

In [11]:
```
print("c value:",k.intercept_)
```

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

In [12]:
```
print("m value:",k.coef_)
```

```
m value: [[1.19303364]]
```

In [14]:
```
y_pred=k.predict(x)
plt.scatter(x,y,color="black")
```

```
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 [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 [19]:
```
n1=4.5
n2=6.5
print("profit from 45000 people city is",k.predict([[n1]])*10000,'$')
print("profit from 6500 people city is ",k.predict([[n2]])*10000,'$')
```

```
profit from 45000 people city is [[14728.70520541]] $
profit from 6500 people city is  [[38589.37808921]] $
```

In [ ]: