

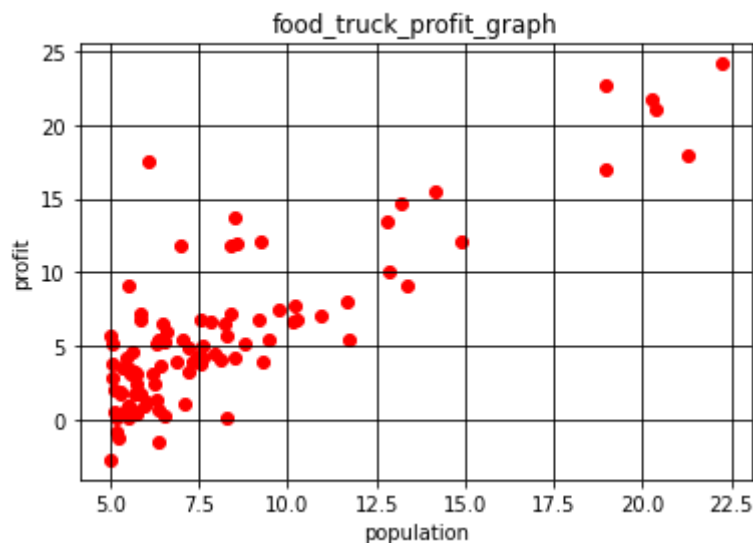
```
In [1]: import numpy
        from matplotlib import pyplot as plt
        import pandas as pd
        from sklearn.model_selection import train_test_split
        from sklearn.linear_model import LinearRegression
```

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

(97, 2)

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

```
In [19]: %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 [5]: k = LinearRegression()
```

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

Out[6]: LinearRegression()

```
In [7]: print('c value:', k.intercept_)
```

c value: [-3.89578088]

```
In [8]: print('m value:',k.coef_)
```

m value: [[1.19303364]]

```
In [9]: y_pred = k.predict(x)
```

```
In [10]: 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 [11]: from sklearn.metrics import r2_score
r_sq = r2_score(y,y_pred)
r_sq
```

```
Out[11]: 0.7020315537841397
```

```
In [12]: from sklearn.metrics import mean_squared_error
rmse = mean_squared_error(y, y_pred)
rmse
```

```
Out[12]: 8.953942751950358
```

```
In [16]: 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 [ ]:
```