

```
In [2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
import pickle

#Load the dataset
dataset=pd.read_csv(r"E:\Naresh IT\data science class notes\machine learning\22nd,

# split the data into independent and dependent variable
x=dataset['sqft_living'].values.reshape(-1,1)
y=dataset['price'].values

# splitting the data for training and testing
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=1/3,random_state=0)

# Train the model
regressor=LinearRegression()
regressor.fit(x_train,y_train)

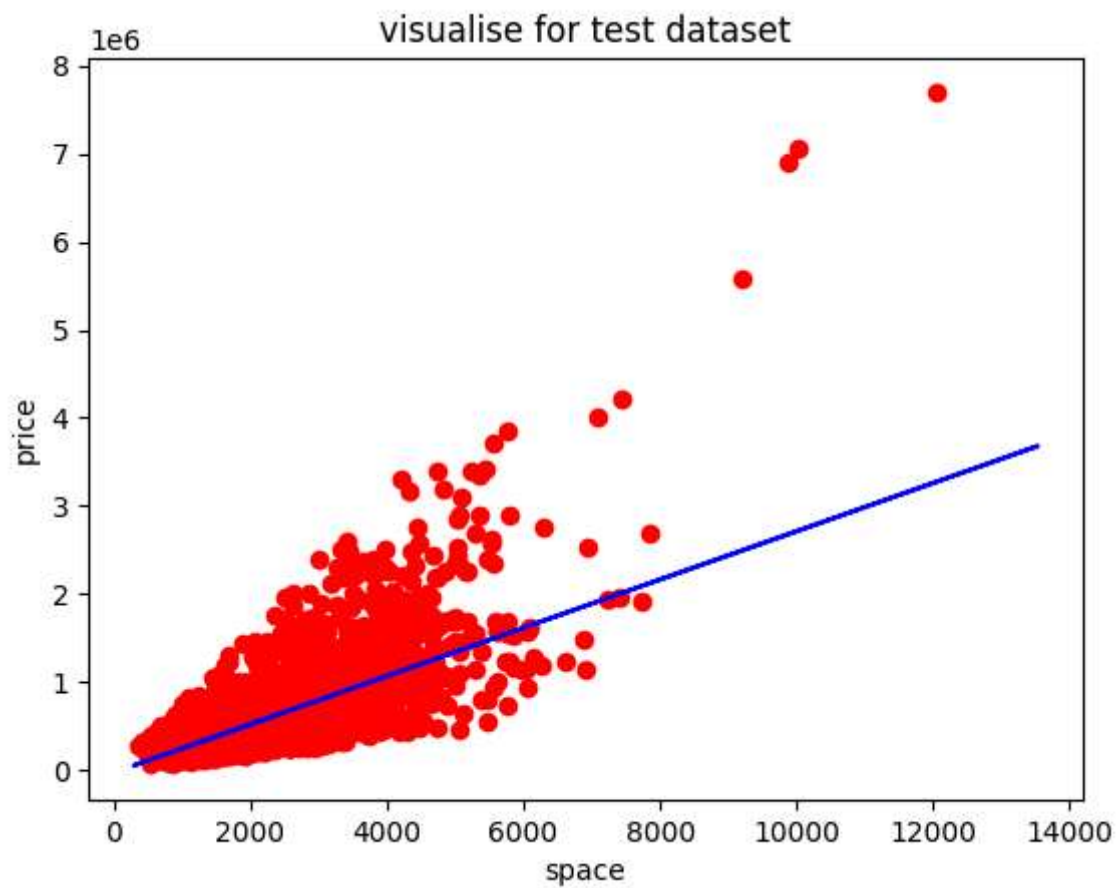
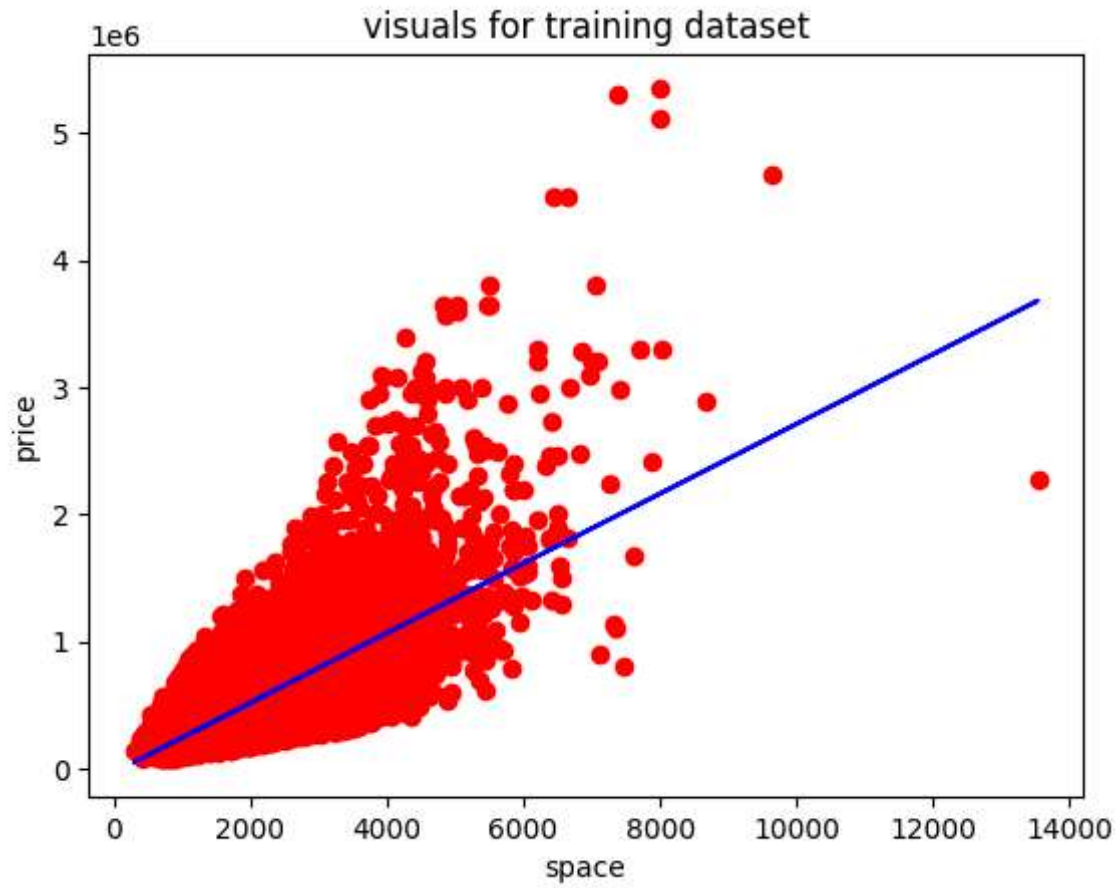
# predicting the prices
y_pred=regressor.predict(x_test)

# visualize the training results
plt.scatter(x_train,y_train,color='red')
plt.plot(x_train,regressor.predict(x_train),color='blue')
plt.title("visuals for training dataset")
plt.xlabel("space")
plt.ylabel("price")
plt.show()

# visualise the test set
plt.scatter(x_test,y_test,color="red")
plt.plot(x_train,regressor.predict(x_train),color='blue')
plt.title("visualise for test dataset")
plt.xlabel("space")
plt.ylabel("price")
plt.show()

# check the model performance
bias=regressor.score(x_train,y_train)
variance=regressor.score(x_test,y_test)
print(f"Training score (R^2): {bias:.2f}")
print(f"Testing score (R^2): {variance:.2f}")

# save the model to the disk
filename="house_pred_regression_model.pkl"
with open(filename,'wb') as file:
    pickle.dump(regressor,file)
print("Model has been pickled and saved as house_pred_regression_model.pkl")
```



Training score (R^2): 0.49

Testing score (R^2): 0.50

Model has been pickled and saved as house_pred_regression_model.pkl