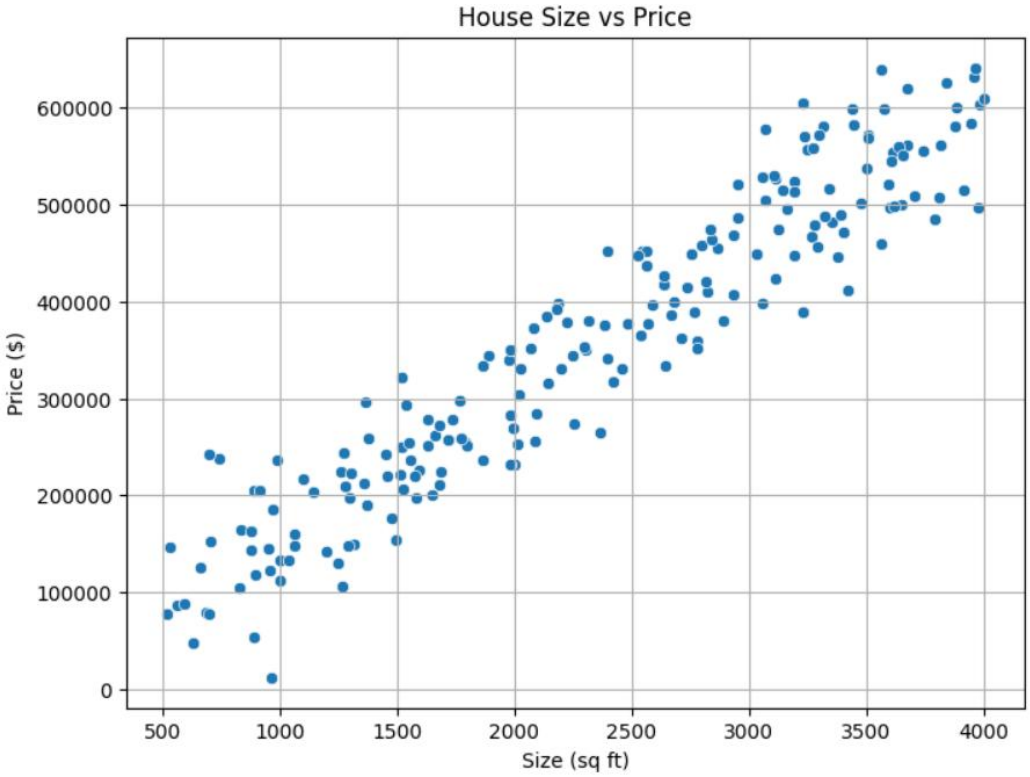
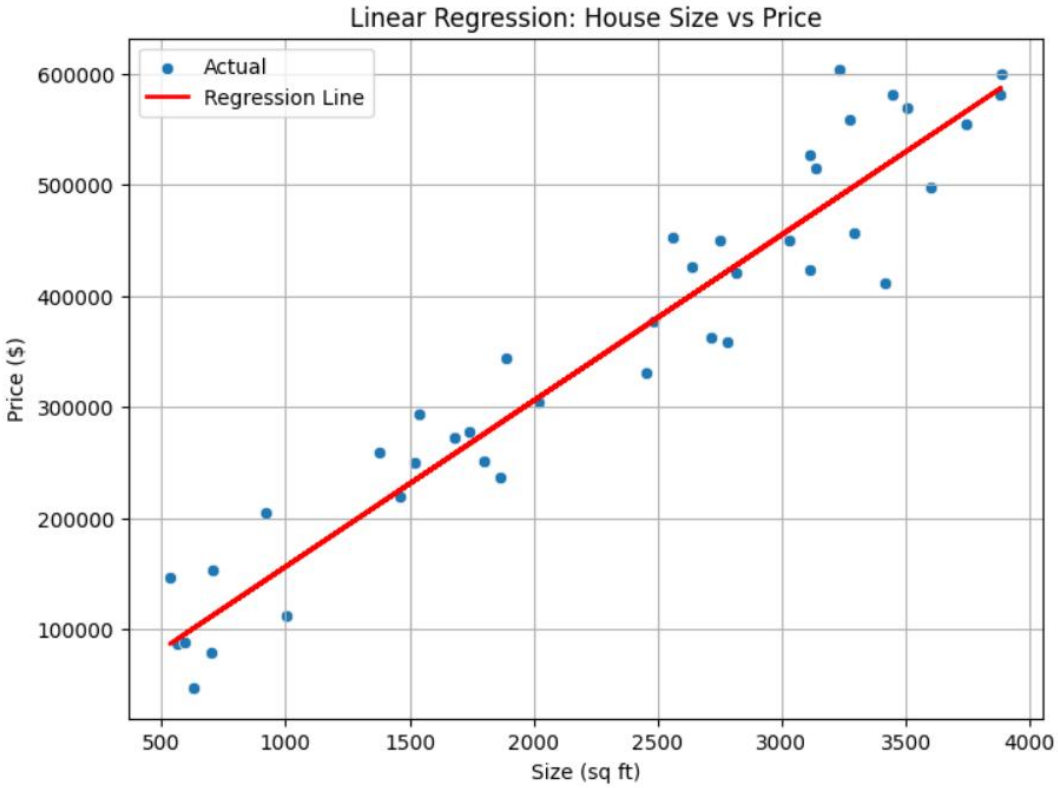
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.model_selection import train_test_split

np.random.seed(42)
house_size = np.random.randint(500, 4000, 200)  # in square feet
house_price = house_size * 150 + np.random.normal(0, 50000, 200)  # price with noise

df = pd.DataFrame({
    'HouseSize': house_size,
    'Price': house_price
})

plt.figure(figsize=(8, 6))
sns.scatterplot(x='HouseSize', y='Price', data=df)
plt.title('House Size vs Price')
plt.xlabel('Size (sq ft)')
plt.ylabel('Price ($)')
plt.grid(True)
plt.show()

X = df[['HouseSize']]
y = df['Price']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
model = LinearRegression()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)

plt.figure(figsize=(8, 6))
sns.scatterplot(x=X_test['HouseSize'], y=y_test, label='Actual')
plt.plot(X_test, y_pred, color='red', linewidth=2, label='Regression Line')
plt.title('Linear Regression: House Size vs Price')
plt.xlabel('Size (sq ft)')
plt.ylabel('Price ($)')
plt.legend()
plt.grid(True)
plt.show()

mse = mean_squared_error(y_test, y_pred)
rmse = np.sqrt(mse)
r2 = r2_score(y_test, y_pred)

print(f"Model Coefficient (Slope): {model.coef_[0]:.2f}")
print(f"Model Intercept: {model.intercept_:.2f}")
print(f"Mean Squared Error (MSE): {mse:.2f}")
print(f"Root Mean Squared Error (RMSE): {rmse:.2f}")
print(f"R² Score: {r2:.4f}")
```

OUTPUT





Model Coefficient (Slope): 149.41
Model Intercept: 6721.43
Mean Squared Error (MSE): 2139805804.92
Root Mean Squared Error (RMSE): 46258.04
R^2 Score: 0.9182