

```

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)
n_samples = 200

data = {
    'EngineSize': np.random.uniform(1.0, 5.0, n_samples),
    'Horsepower': np.random.randint(70, 400, n_samples),
    'FuelEfficiency': np.random.uniform(10, 40, n_samples),
}

data['Price'] = (
    data['EngineSize'] * 5000 +
    data['Horsepower'] * 100 +
    data['FuelEfficiency'] * -200 +
    np.random.normal(0, 3000, n_samples)
)

df = pd.DataFrame(data)

sns.pairplot(df)
plt.suptitle("Car Attributes vs Price", y=1.02)
plt.show()

X = df[['EngineSize', 'Horsepower', 'FuelEfficiency']]
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)

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

print("\nModel Evaluation:")
print(f"Mean Squared Error (MSE): {mse:.2f}")
print(f"Root Mean Squared Error (RMSE): {rmse:.2f}")
print(f"R2 Score: {r2:.4f}")

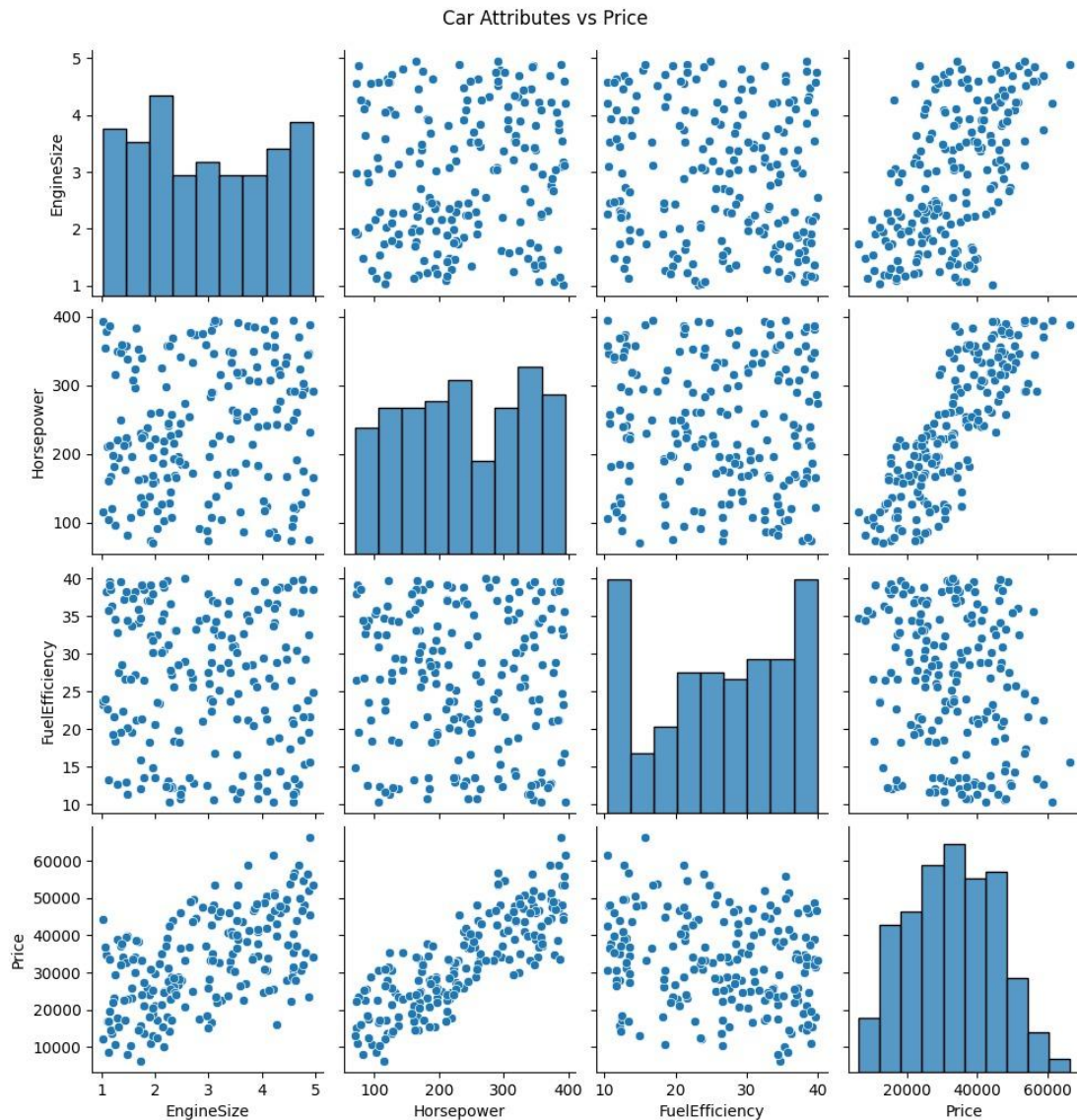
coef_df = pd.DataFrame({
    'Feature': X.columns,
    'Coefficient': model.coef_
}).sort_values(by='Coefficient', ascending=False)

print("\nFeature Influence on Car Price:")
print(coef_df)

plt.figure(figsize=(8, 6))
sns.scatterplot(x=y_test, y=y_pred)
plt.xlabel("Actual Price")
plt.ylabel("Predicted Price")
plt.title("Actual vs Predicted Car Prices")
plt.plot([y.min(), y.max()], [y.min(), y.max()], 'r--')
plt.grid(True)
plt.show()

```

## OUTPUT



### Model Evaluation:

Mean Squared Error (MSE): 10400491.13

Root Mean Squared Error (RMSE): 3224.98

$R^2$  Score: 0.9238

### Feature Influence on Car Price:

Feature Coefficient

0 EngineSize 5325.443312

1 Horsepower 102.296598

2 FuelEfficiency -207.111115

