In [57]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import PolynomialFeatures
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error
```

In [58]:
```
# Load the dataset
data = pd.read_csv("/Users/user/Downloads/Housing_Data.csv")
```

In [59]:
```
# Display the first few rows of the dataset
print(data.head())
```

```
   Avg. Area Income  Avg. Area House Age  Avg. Area Number of Rooms  \
0      79545.458574             5.682861                   7.009188   
1      79248.642455             6.002900                   6.730821   
2      61287.067179             5.865890                   8.512727   
3      63345.240046             7.188236                   5.586729   
4      59982.197226             5.040555                   7.839388   

   Avg. Area Number of Bedrooms  Area Population         Price  \
0                          4.09     23086.800503  1.059034e+06   
1                          3.09     40173.072174  1.505891e+06   
2                          5.13     36882.159400  1.058988e+06   
3                          3.26     34310.242831  1.260617e+06   
4                          4.23     26354.109472  6.309435e+05   

                                             Address  
0  208 Michael Ferry Apt. 674\nLaurabury, NE 3701...  
1  188 Johnson Views Suite 079\nLake Kathleen, CA...  
2  9127 Elizabeth Stravenue\nDanieltown, WI 06482...  
3                          USS Barnett\nFPO AP 44820  
4                         USNS Raymond\nFPO AE 09386  
```

In [ ]:

In [ ]:

In [60]:

```
# Check for any missing values
print(data.isnull().sum())
```

```
Avg. Area Income                0
Avg. Area House Age             0
Avg. Area Number of Rooms       0
Avg. Area Number of Bedrooms    0
Area Population                 0
Price                           0
Address                         0
dtype: int64
```

In [86]:

```
ep 1: Preprocess the data
 data[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms', 'Avg. Area Number of Bedrooms'
 data['Price']           # Target variable
```

In [99]:

```
# Step 2: Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state=0)
```

In [100]:

```
# Step 3: Create polynomial features
poly = PolynomialFeatures(degree=3)  # Adjust the degree as needed
X_poly_train = poly.fit_transform(X_train)
X_poly_test = poly.transform(X_test)
```

In [101]:

```
# Step 4: Fit the model
model = LinearRegression()
model.fit(X_poly_train, y_train)
```

Out[101]:

```
LinearRegression()
```

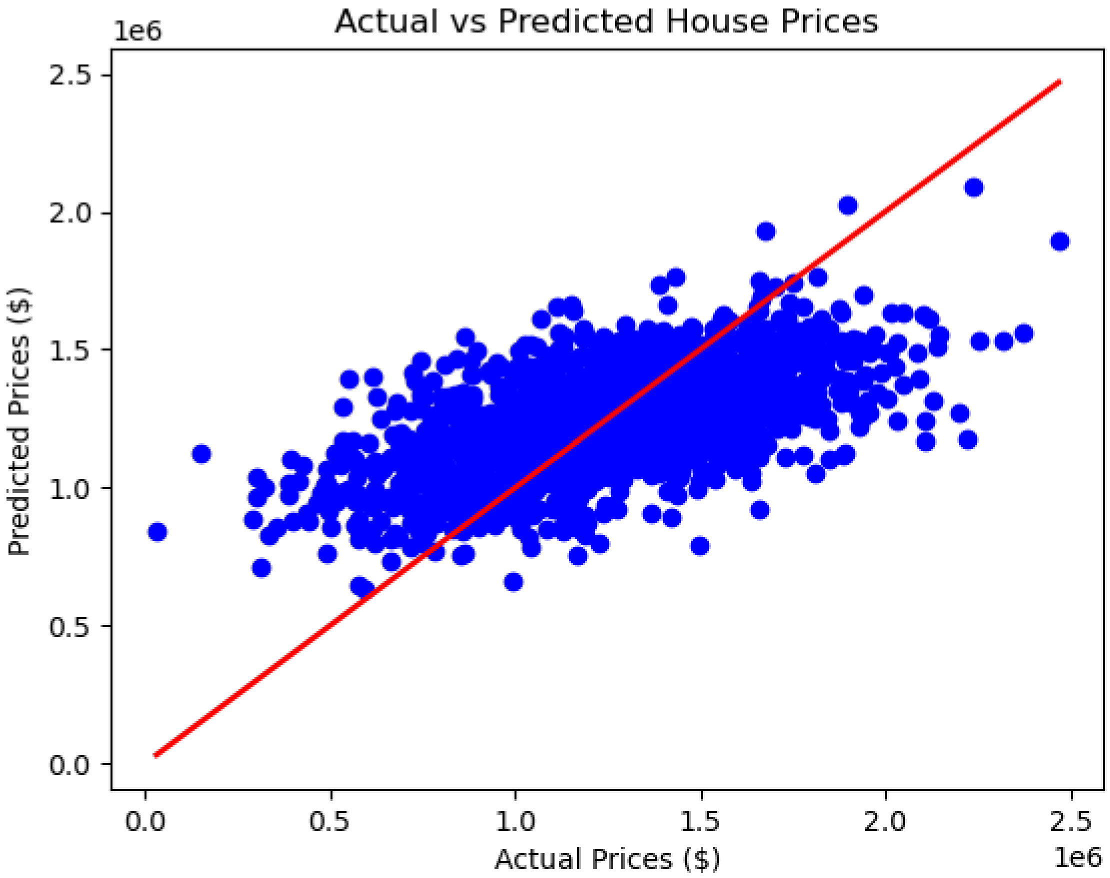
In [102]:

```
# Step 5: Make predictions
y_pred = model.predict(X_poly_test)
```

In [103]:

```
# Step 6: Evaluate the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

In [104]:
```
# Step 7: Visualize the results
plt.scatter(y_test, y_pred, color='blue')
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], color='red', linewidth=2)
plt.xlabel('Actual Prices ($)')
plt.ylabel('Predicted Prices ($)')
plt.title('Actual vs Predicted House Prices')
plt.show()
```



```
In [105]: # Calculate Mean Absolute Error
          mae = mean_absolute_error(y_test, y_pred)
          print(f'Mean Absolute Error (MAE): {mae:.2f}')

          Mean Absolute Error (MAE): 233581.09

In [106]: # Calculate Mean Squared Error
          mse = mean_squared_error(y_test, y_pred)
          print(f'Mean Squared Error (MSE): {mse:.2f}')

          Mean Squared Error (MSE): 85121794210.41

In [107]: # Calculate Root Mean Squared Error
          rmse = np.sqrt(mse)
          print(f'Root Mean Squared Error (RMSE): {rmse:.2f}')

          Root Mean Squared Error (RMSE): 291756.40

In [ ]:
```