```python
 # Import libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
import matplotlib.pyplot as plt

# Step 1: Load dataset
df = pd.read_csv('house_prices_large.csv')# Ensure this CSV is in your working directory
print(df.head())

# Step 2: Define features and target
X = df[['Area', 'Bedrooms', 'Bathrooms', 'Parking', 'Age']]
y = df['Price']

# Step 3: Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=42)

# Step 4: Train Random Forest model
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Step 5: Predict
predictions = model.predict(X_test)

# Step 6: Display results
print("Predicted vs Actual Prices:")
for pred, actual in zip(predictions[:10], y_test.values[:10]):
    print(f"Predicted: ${pred:,.2f} | Actual: ${actual:,.2f}")

# Step 7: Evaluate
print("\nModel Performance:")
print("Mean Squared Error:", mean_squared_error(y_test, predictions))
print("R² Score:", r2_score(y_test, predictions))

# Step 8: Plot
plt.scatter(y_test, predictions, color='blue', alpha=0.6)
plt.plot([y.min(), y.max()], [y.min(), y.max()], '--r')
plt.xlabel('Actual Price')
plt.ylabel('Predicted Price')
plt.title('Actual vs Predicted House Prices')
plt.grid(True)
plt.show()
```

```
      Area  Bedrooms  Bathrooms  Parking  Age    Price
  0   1660         2          1        1   20   251796
  1   2094         2          2        1   10   342409
  2   1930         4          3        1   27   346390
  3   1895         5          1        2    4   357028
  4   2438         1          1        0    3   376363
Predicted vs Actual Prices:
Predicted: $312,459.98 | Actual: $292,836.00
Predicted: $216,687.97 | Actual: $211,036.00
Predicted: $378,958.07 | Actual: $384,058.00
Predicted: $221,187.14 | Actual: $228,294.00
Predicted: $417,917.04 | Actual: $420,443.00
Predicted: $242,996.17 | Actual: $250,706.00
Predicted: $210,489.53 | Actual: $192,980.00
Predicted: $362,878.01 | Actual: $358,059.00
Predicted: $177,887.74 | Actual: $181,867.00
Predicted: $187,449.67 | Actual: $180,957.00

Model Performance:
Mean Squared Error: 245925990.92489204
R² Score: 0.9676761439580798
```


Actual vs Predicted House Prices