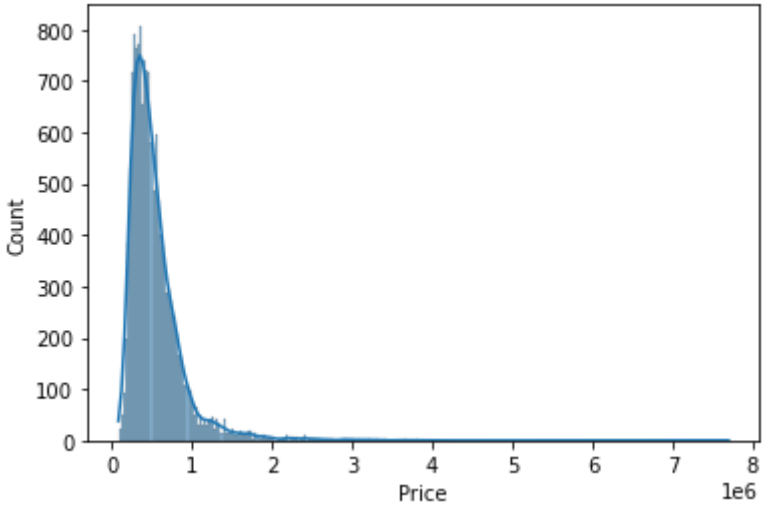
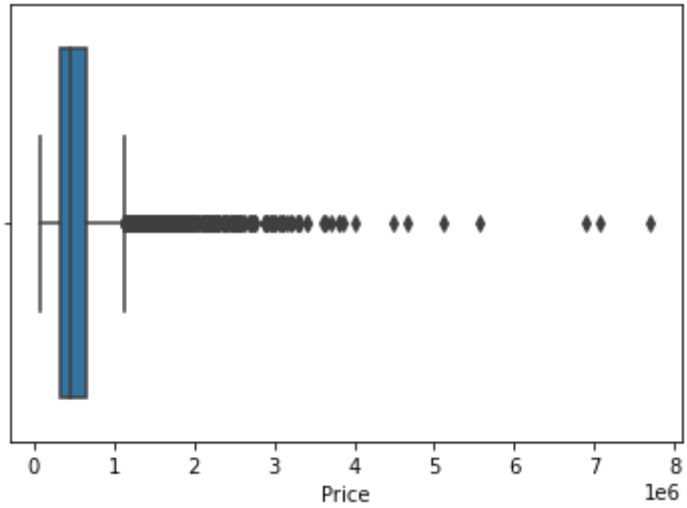
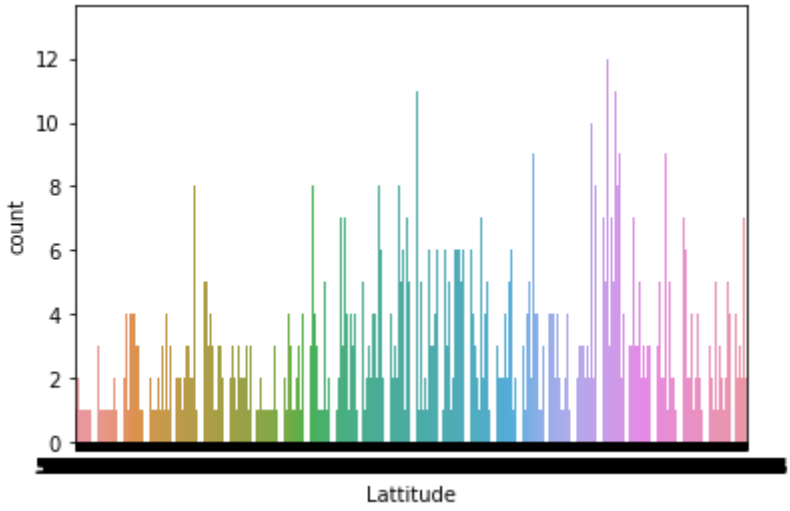
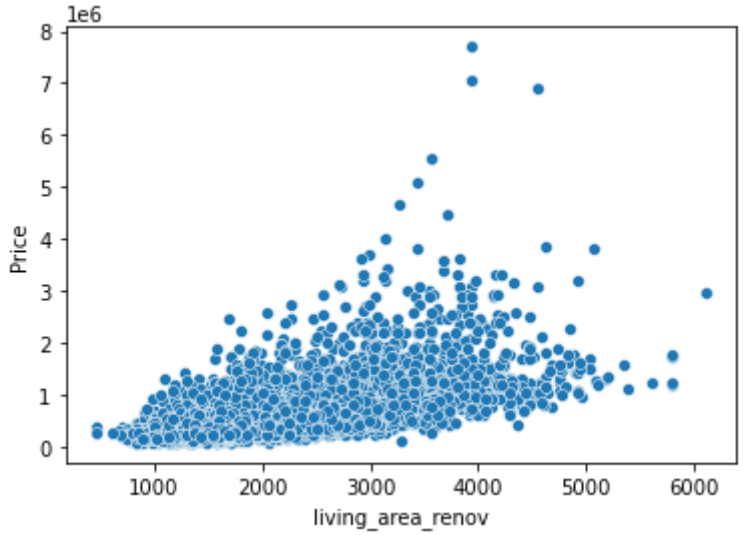
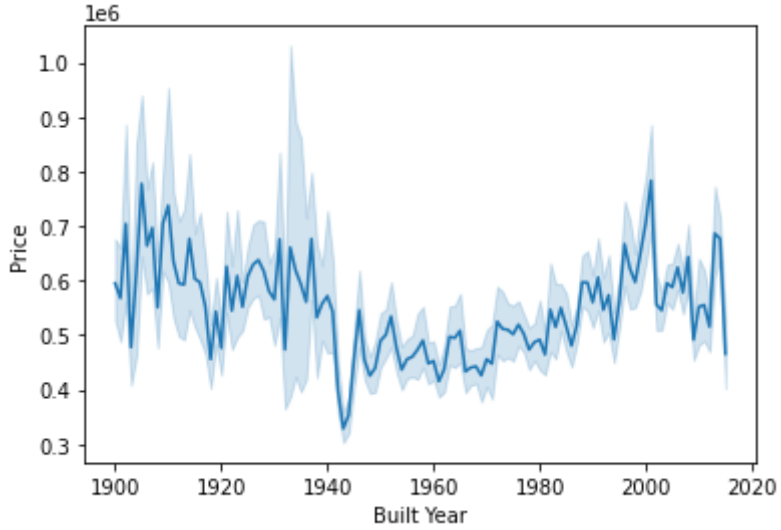
In [2]:
```
# Import necessary libraries
from sklearn.decomposition import PCA
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
df = pd.read_csv('House Price India.csv')

# Univariate Analysis
# Histogram
sns.histplot(df['Price'], kde=True)
plt.show()

# Box plot
sns.boxplot(x=df['Price'])
plt.show()

# Bar chart
sns.countplot(x='Lattitude', data=df)
plt.show()

# Bivariate Analysis
# Scatter plot
sns.scatterplot(x='living_area_renov', y='Price', data=df)
plt.show()

# Line plot
sns.lineplot(x='Built Year', y='Price', data=df)
plt.show()

# Heatmap
sns.heatmap(df.corr(), annot=True)
plt.show()

# Multivariate Analysis
# Pair plot
sns.pairplot(df, hue='Price')
plt.show()

# PCA
pca = PCA(n_components=2)
X = df.drop('Price', axis=1)
y = df['Price']
X_pca = pca.fit_transform(X)
sns.scatterplot(x=X_pca[:, 0], y=X_pca[:, 1], hue=y)
plt.show()
```

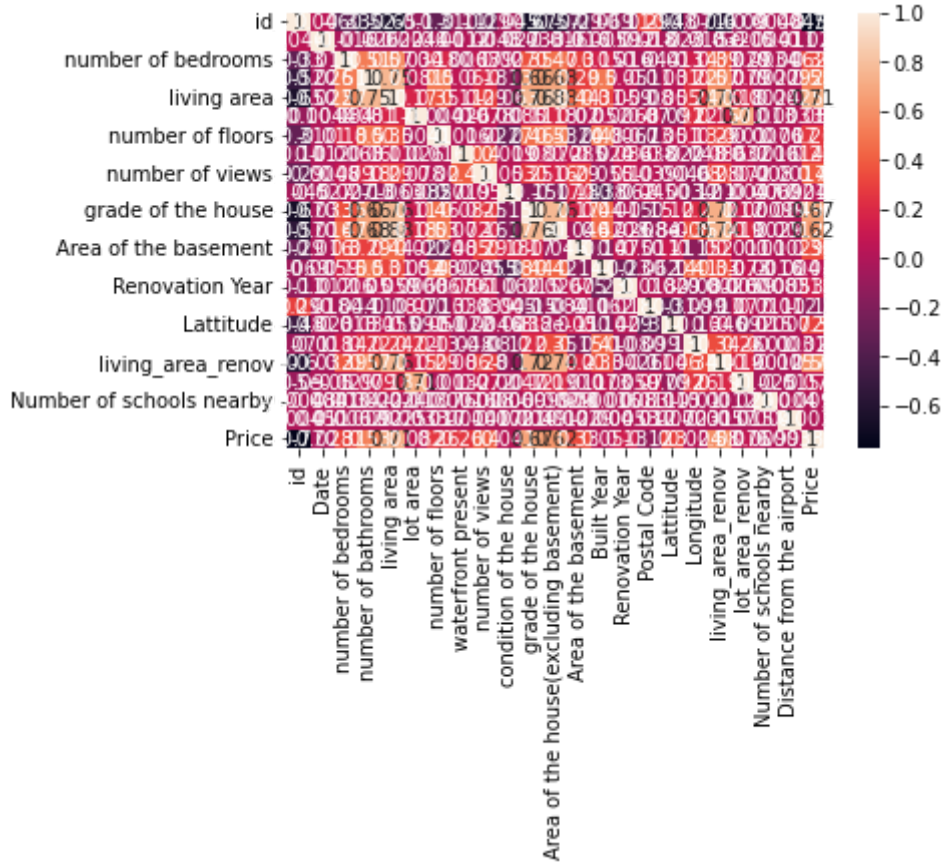

Count
Price
1e6

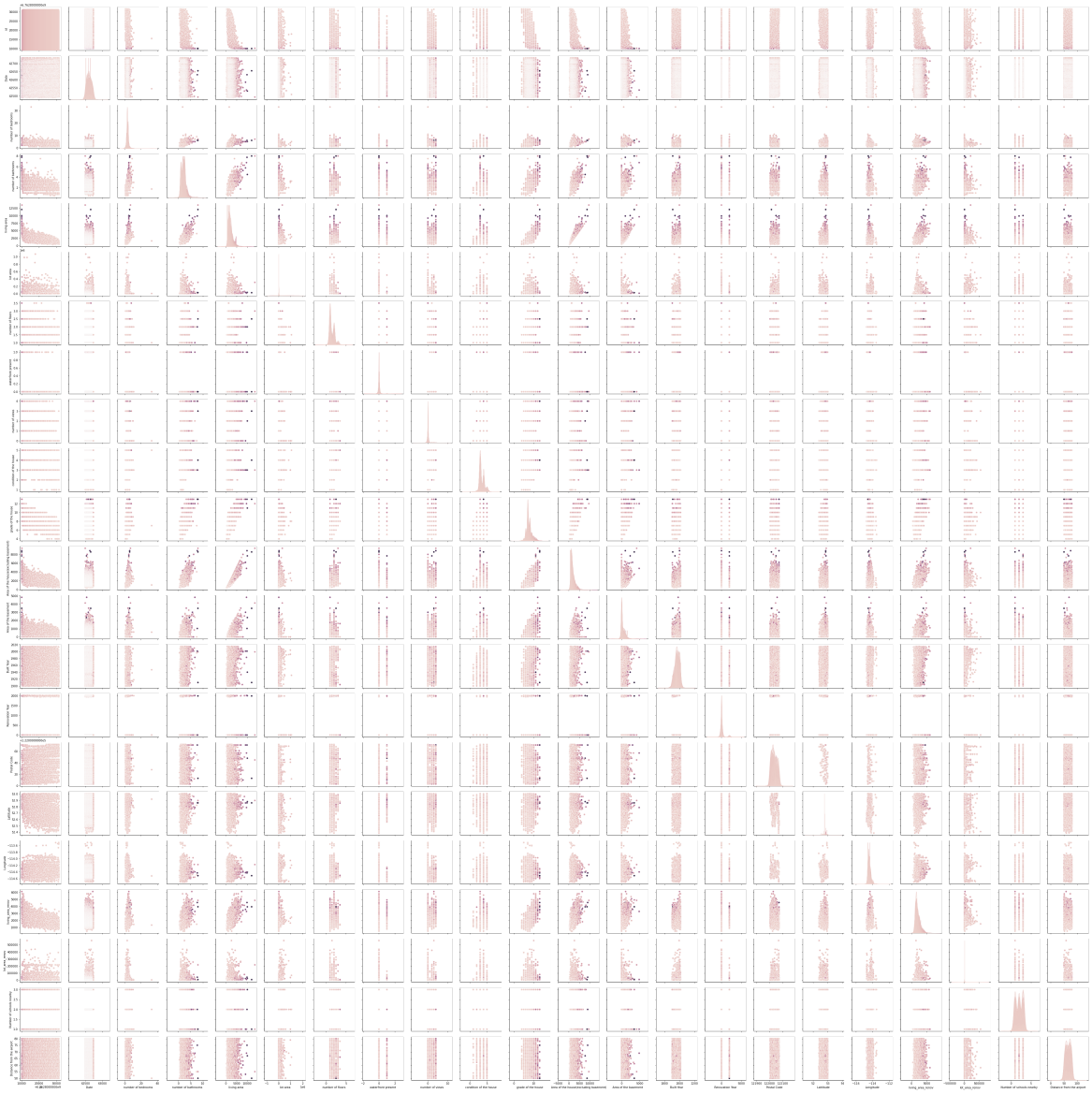

Price
1e6

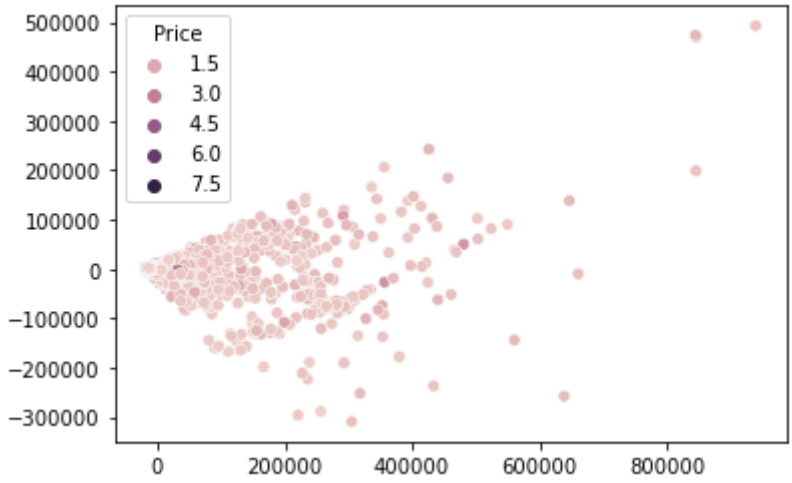

count
Lattitude


1e6
Price
living_area_renov


1e6
Price
Built Year


id
number of bedrooms
living area
number of floors
number of views
grade of the house
Area of the basement
Renovation Year
Lattitude
living_area_renov
Number of schools nearby
Price
Date
number of bathrooms
lot area
waterfront present
condition of the house
Area of the house(excluding basement)
Built Year
Postal Code
Longitude
lot_area_renov
Distance from the airport




Price
1.5
3.0
4.5
6.0
7.5
500000
400000
300000
200000
100000
0
−100000
−200000
−300000
0
200000
400000
600000
800000

In [ ]: