In [3]:
```
import pandas as pd
import numpy as np
```

In [4]:
```
pwd
```

Out[4]:
```
'C:\\Users\\DELL'
```

In [5]:
```
df = pd.read_csv('housingdata.csv')
```

In [6]:
```
df.head()
```

Out[6]:

	Unnamed: 0	price	lot_size	waterfront	age	land_value	construction	air_cond	fuel
0	1	132500	0.09	No	42	50000	No	No	Electric
1	2	181115	0.92	No	0	22300	No	No	Gas
2	3	109000	0.19	No	133	7300	No	No	Gas
3	4	155000	0.41	No	13	18700	No	No	Gas
4	5	86060	0.11	No	0	15000	Yes	Yes	Gas

In [ ]:
```

```

In [21]:
```
#min price of the house

df['price'].min()
```

Out[21]:
```
5000
```

In [13]:
```
#median value of living area

df['living_area'].median()
```

Out[13]:
```
1634.5
```

In [14]:
```
df.columns
```

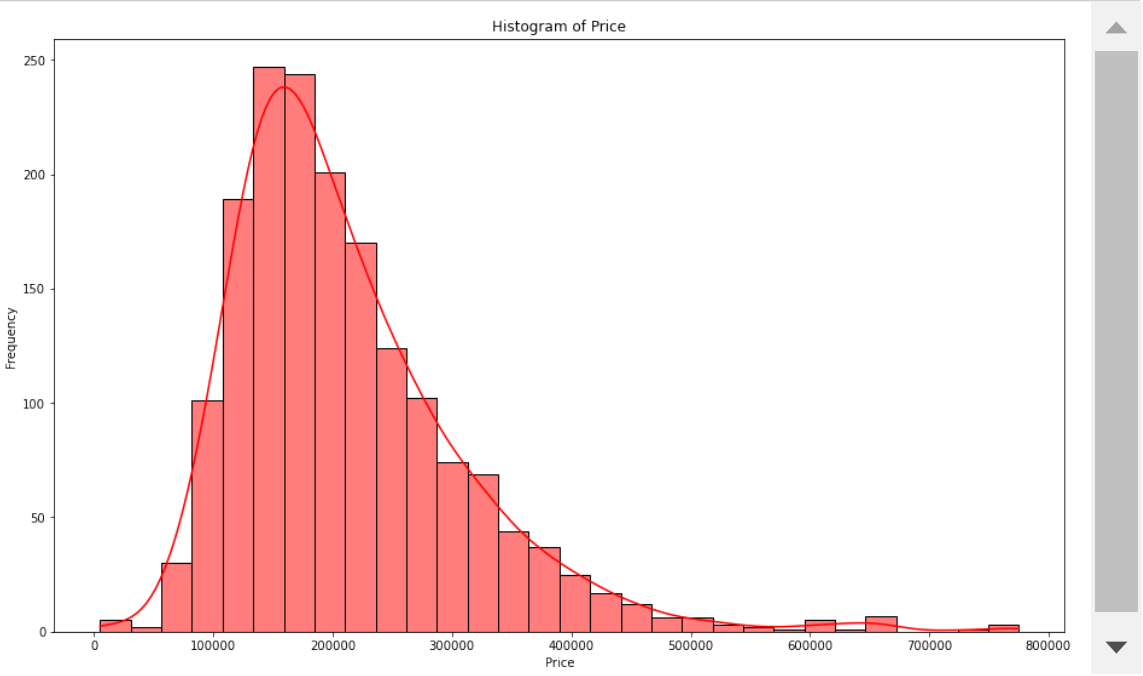
Out[14]:
```
Index(['Unnamed: 0', 'price', 'lot_size', 'waterfront', 'age', 'land_valu
e',
       'construction', 'air_cond', 'fuel', 'heat', 'sewer', 'living_are
a',
       'fireplaces', 'bathrooms', 'rooms'],
      dtype='object')
```

In [18]:
```
#no of rows and columns

num_rows, num_columns = df.shape
print(f"Number of rows: {num_rows}")
print(f"Number of columns: {num_columns}")
```

```
Number of rows: 1728
Number of columns: 15
```

In [23]:
```
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import shapiro
import pandas as pd
```

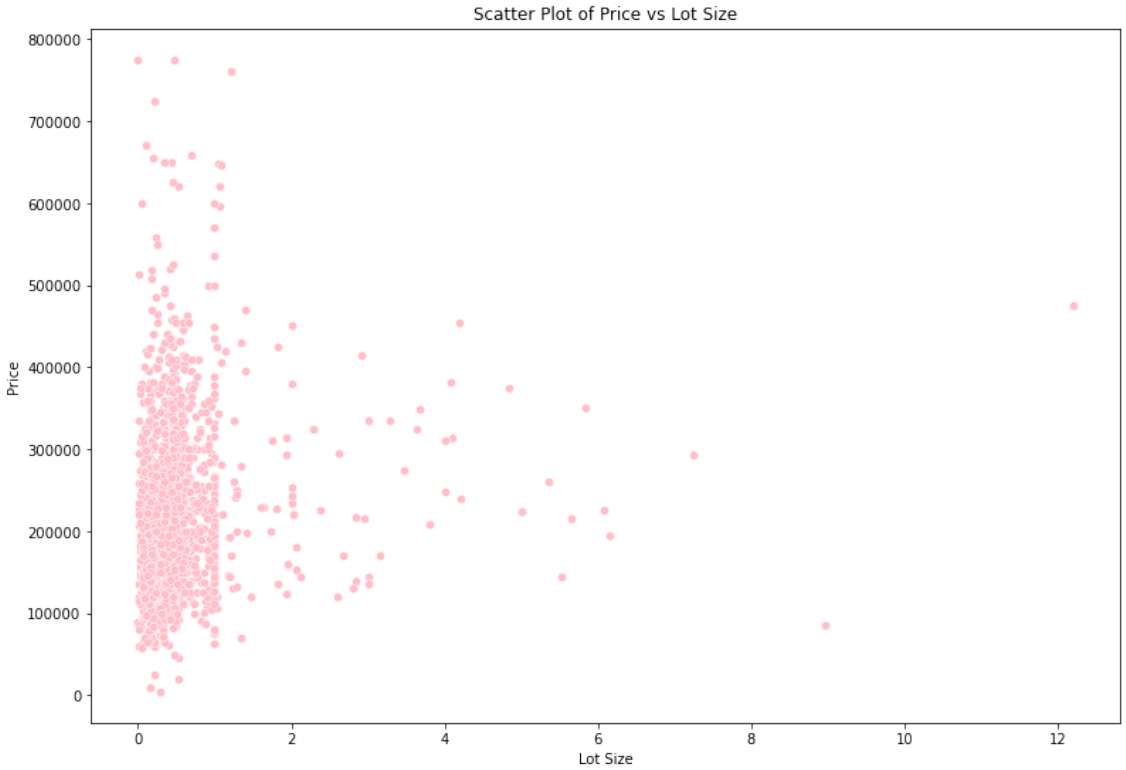
In [27]:
```
#histogram of price

plt.figure(figsize=(15, 9))
sns.histplot(df['price'], kde=True, color='red', bins=30)
plt.title('Histogram of Price')
plt.xlabel('Price')
plt.ylabel('Frequency')
plt.show()
```



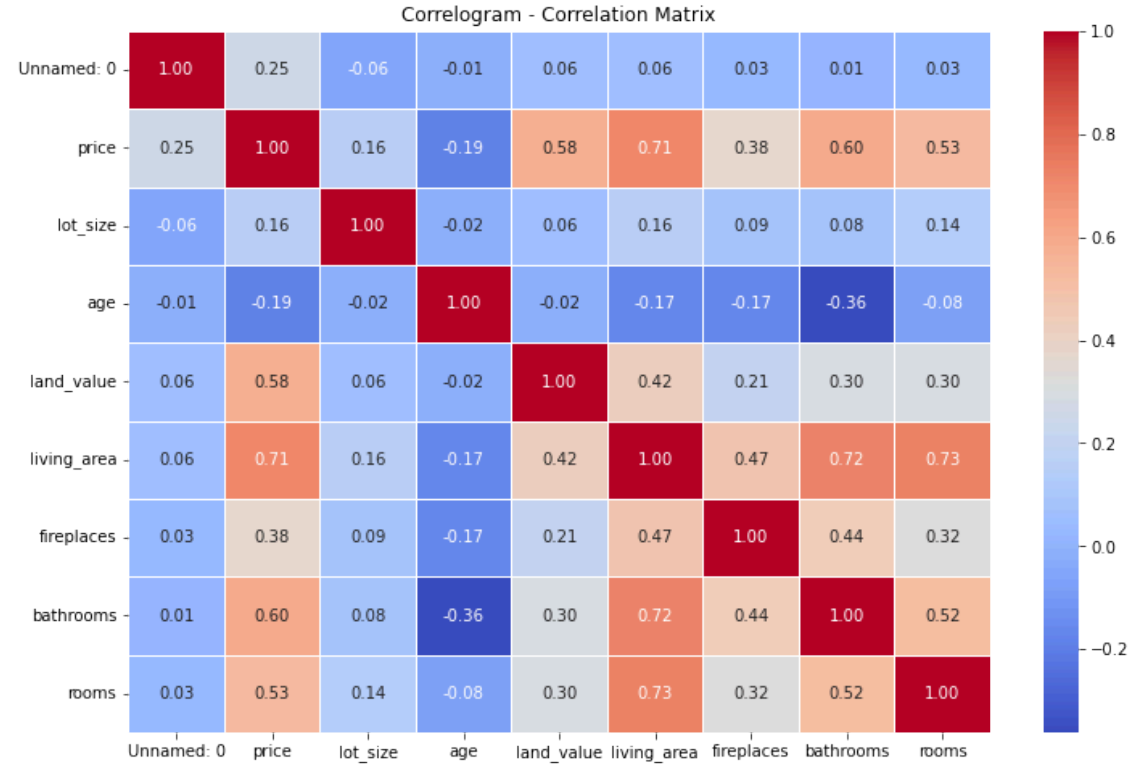
In [34]:
```
#houses having waterfronts

df['waterfront'].value_counts()
```

Out[34]:
```
No     1713
Yes      15
Name: waterfront, dtype: int64
```

In [36]:
```
#plot scatterplot

plt.figure(figsize=(13, 9))
sns.scatterplot(x='lot_size', y='price', data=df, color='pink')
plt.title('Scatter Plot of Price vs Lot Size')
plt.xlabel('Lot Size')
plt.ylabel('Price')
plt.show()
```



In [37]:
```
#plot correleogram

correlation_matrix = df.corr()
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f", li
plt.title('Correlogram - Correlation Matrix')
plt.show()
```



In [ ]:
```
#interpretation

Interpreting the Correlogram:

Values near 1 imply a significant positive association.
Values near -1 imply a strong negative connection.

Values close to zero imply a weak or no association.
Look for patterns and relationships among the variables. For example, if th
```

In [54]:

In [ ]:

In [ ]: