In [6]:

```
pip install numpy
```

Requirement already satisfied: numpy in c:\users\sasidhar royal\appdata\lo
cal\programs\python\python311\lib\site-packages (1.24.3)
Note: you may need to restart the kernel to use updated packages.

In [7]:

```
pip install pandas
```

Requirement already satisfied: pandas in c:\users\sasidhar royal\appdata\l
ocal\programs\python\python311\lib\site-packages (2.0.2)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\sasidhar
royal\appdata\local\programs\python\python311\lib\site-packages (from pand
as) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from pandas) (202
3.3)
Requirement already satisfied: tzdata>=2022.1 in c:\users\sasidhar royal\a
ppdata\local\programs\python\python311\lib\site-packages (from pandas) (20
23.3)
Requirement already satisfied: numpy>=1.21.0 in c:\users\sasidhar royal\ap
pdata\local\programs\python\python311\lib\site-packages (from pandas) (1.2
4.3)
Requirement already satisfied: six>=1.5 in c:\users\sasidhar royal\appdata
\local\programs\python\python311\lib\site-packages (from python-dateutil>=
2.8.2->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.

In [8]:

```
import numpy as np
import pandas as pd
```

In [9]:

```
pip install Seaborn
```

Requirement already satisfied: Seaborn in c:\users\sasidhar royal\appdata
\local\programs\python\python311\lib\site-packages (0.12.2)
Requirement already satisfied: numpy!=1.24.0,>=1.17 in c:\users\sasidhar r
oyal\appdata\local\programs\python\python311\lib\site-packages (from Seabo
rn) (1.24.3)
Requirement already satisfied: pandas>=0.25 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from Seaborn) (2.
0.2)
Requirement already satisfied: matplotlib!=3.6.1,>=3.1 in c:\users\sasidha
r royal\appdata\local\programs\python\python311\lib\site-packages (from Se
aborn) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b!=3.6.1,>=3.1->Seaborn) (1.0.7)
Requirement already satisfied: cycler>=0.10 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from matplotlib!=
3.6.1,>=3.1->Seaborn) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\sasidhar roya
l\appdata\local\programs\python\python311\lib\site-packages (from matplotl
ib!=3.6.1,>=3.1->Seaborn) (4.39.4)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\sasidhar roya
l\appdata\local\programs\python\python311\lib\site-packages (from matplotl
ib!=3.6.1,>=3.1->Seaborn) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b!=3.6.1,>=3.1->Seaborn) (23.1)
Requirement already satisfied: pillow>=6.2.0 in c:\users\sasidhar royal\ap
pdata\local\programs\python\python311\lib\site-packages (from matplotlib!=
3.6.1,>=3.1->Seaborn) (9.5.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b!=3.6.1,>=3.1->Seaborn) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\sasidhar r
oyal\appdata\local\programs\python\python311\lib\site-packages (from matpl
otlib!=3.6.1,>=3.1->Seaborn) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from pandas>=0.25-
>Seaborn) (2023.3)
Requirement already satisfied: tzdata>=2022.1 in c:\users\sasidhar royal\a
ppdata\local\programs\python\python311\lib\site-packages (from pandas>=0.2
5->Seaborn) (2023.3)
Requirement already satisfied: six>=1.5 in c:\users\sasidhar royal\appdata
\local\programs\python\python311\lib\site-packages (from python-dateutil>=
2.7->matplotlib!=3.6.1,>=3.1->Seaborn) (1.16.0)
Note: you may need to restart the kernel to use updated packages.

In [10]:

```
pip install matplotlib
```

Requirement already satisfied: matplotlib in c:\users\sasidhar royal\appda
ta\local\programs\python\python311\lib\site-packages (3.7.1)Note: you may
need to restart the kernel to use updated packages.

Requirement already satisfied: contourpy>=1.0.1 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b) (1.0.7)
Requirement already satisfied: cycler>=0.10 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from matplotlib)
(0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\sasidhar roya
l\appdata\local\programs\python\python311\lib\site-packages (from matplotl
ib) (4.39.4)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\sasidhar roya
l\appdata\local\programs\python\python311\lib\site-packages (from matplotl
ib) (1.4.4)
Requirement already satisfied: numpy>=1.20 in c:\users\sasidhar royal\appd
ata\local\programs\python\python311\lib\site-packages (from matplotlib)
(1.24.3)
Requirement already satisfied: packaging>=20.0 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b) (23.1)
Requirement already satisfied: pillow>=6.2.0 in c:\users\sasidhar royal\ap
pdata\local\programs\python\python311\lib\site-packages (from matplotlib)
(9.5.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\sasidhar royal
\appdata\local\programs\python\python311\lib\site-packages (from matplotli
b) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\sasidhar r
oyal\appdata\local\programs\python\python311\lib\site-packages (from matpl
otlib) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\sasidhar royal\appdata
\local\programs\python\python311\lib\site-packages (from python-dateutil>=
2.7->matplotlib) (1.16.0)

In [11]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```
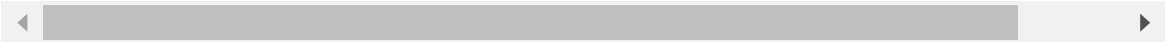
In [12]:

```python
df=pd.read_csv(r"C:\Users\SASIDHAR ROYAL\Downloads\USA_Housing.csv")
df
```

Out[12]:

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price	
0	79545.458574	5.682861	7.009188	4.09	23086.800503	1.059034e+06	208 Michae 674\nLau
1	79248.642455	6.002900	6.730821	3.09	40173.072174	1.505891e+06	188 Johr Suite Kath
2	61287.067179	5.865890	8.512727	5.13	36882.159400	1.058988e+06	912 Stravenue\nD \
3	63345.240046	7.188236	5.586729	3.26	34310.242831	1.260617e+06	USS Barnet
4	59982.197226	5.040555	7.839388	4.23	26354.109472	6.309435e+05	USNS Raym
...	...	...	...	...	...	...	
4995	60567.944140	7.830362	6.137356	3.46	22837.361035	1.060194e+06	USNS Willi AP 3
4996	78491.275435	6.999135	6.576763	4.02	25616.115489	1.482618e+06	PSC 8489\nAPO
4997	63390.686886	7.250591	4.805081	2.13	33266.145490	1.030730e+06	4215 Tra Suite 076\nJ
4998	68001.331235	5.534388	7.130144	5.44	42625.620156	1.198657e+06	USS Wallac
4999	65510.581804	5.992305	6.792336	4.07	46501.283803	1.298950e+06	37778 Geo Apt. 509\n

5000 rows × 7 columns

In [13]:

```
df.head()
```

Out[13]:

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price	A
0	79545.458574	5.682861	7.009188	4.09	23086.800503	1.059034e+06	208 Michael Fe 674\nLaurabt
1	79248.642455	6.002900	6.730821	3.09	40173.072174	1.505891e+06	188 Johnsor Suite 079 Kathleer
2	61287.067179	5.865890	8.512727	5.13	36882.159400	1.058988e+06	9127 Eli Stravenue\nDani WI 0
3	63345.240046	7.188236	5.586729	3.26	34310.242831	1.260617e+06	USS Barnett\nF
4	59982.197226	5.040555	7.839388	4.23	26354.109472	6.309435e+05	USNS Raymonc AE

In [14]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999
Data columns (total 7 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Avg. Area Income              5000 non-null   float64
 1   Avg. Area House Age           5000 non-null   float64
 2   Avg. Area Number of Rooms     5000 non-null   float64
 3   Avg. Area Number of Bedrooms  5000 non-null   float64
 4   Area Population               5000 non-null   float64
 5   Price                         5000 non-null   float64
 6   Address                       5000 non-null   object
dtypes: float64(6), object(1)
memory usage: 273.6+ KB
```

In [15]:

```python
df.describe()
```

Out[15]:

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price
count	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5.000000e+03
mean	68583.108984	5.977222	6.987792	3.981330	36163.516039	1.232073e+06
std	10657.991214	0.991456	1.005833	1.234137	9925.650114	3.531176e+05
min	17796.631190	2.644304	3.236194	2.000000	172.610686	1.593866e+04
25%	61480.562388	5.322283	6.299250	3.140000	29403.928702	9.975771e+05
50%	68804.286404	5.970429	7.002902	4.050000	36199.406689	1.232669e+06
75%	75783.338666	6.650808	7.665871	4.490000	42861.290769	1.471210e+06
max	107701.748378	9.519088	10.759588	6.500000	69621.713378	2.469066e+06

In [16]:

```python
df.columns
```
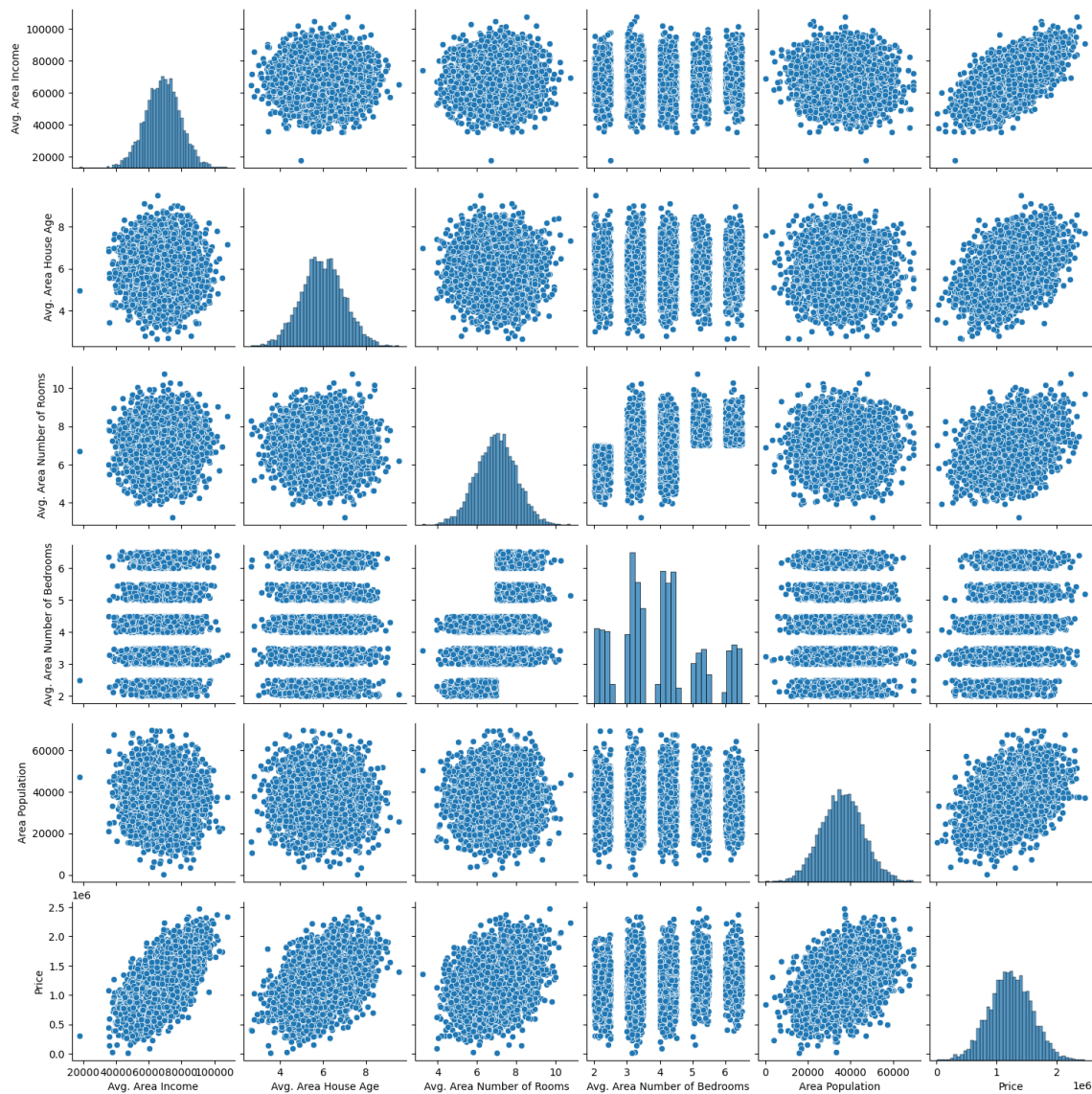
Out[16]:

```
Index(['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Roo
ms',
       'Avg. Area Number of Bedrooms', 'Area Population', 'Price', 'Addres
s'],
      dtype='object')
```
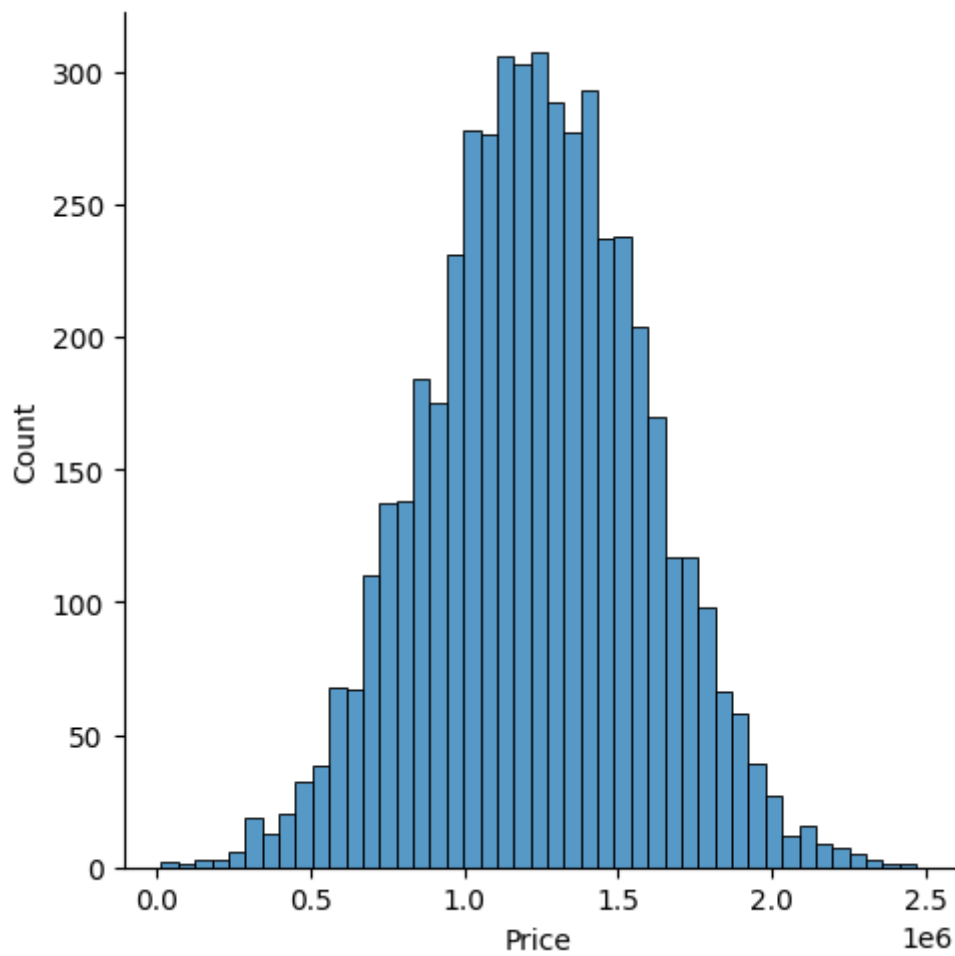
In [17]:

```
sns.pairplot(df)
```

Out[17]:

`<seaborn.axisgrid.PairGrid at 0x20c8f86f010>`
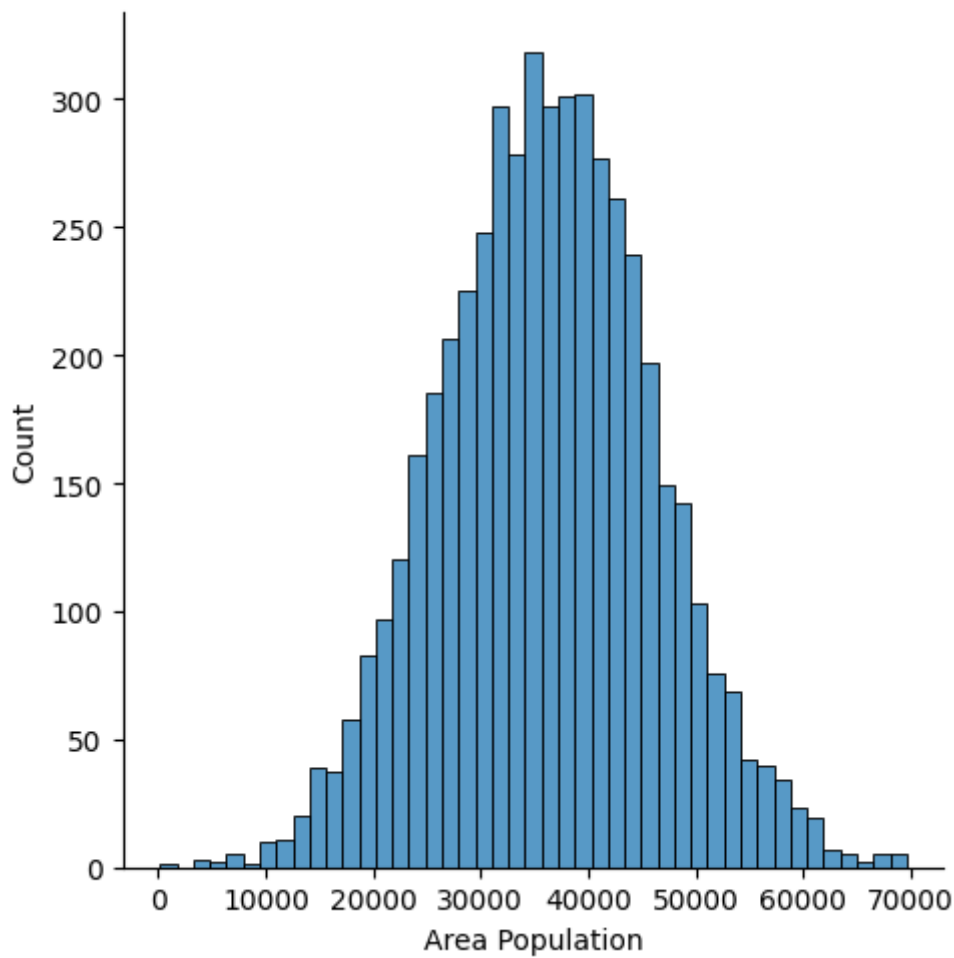
In [18]:

```python
sns.displot(df['Price'])
```

Out[18]:

```
<seaborn.axisgrid.FacetGrid at 0x20c92b26010>
```

In [19]:

```python
sns.displot(df['Area Population'])
```

Out[19]:

```
<seaborn.axisgrid.FacetGrid at 0x20c9413bf10>
```
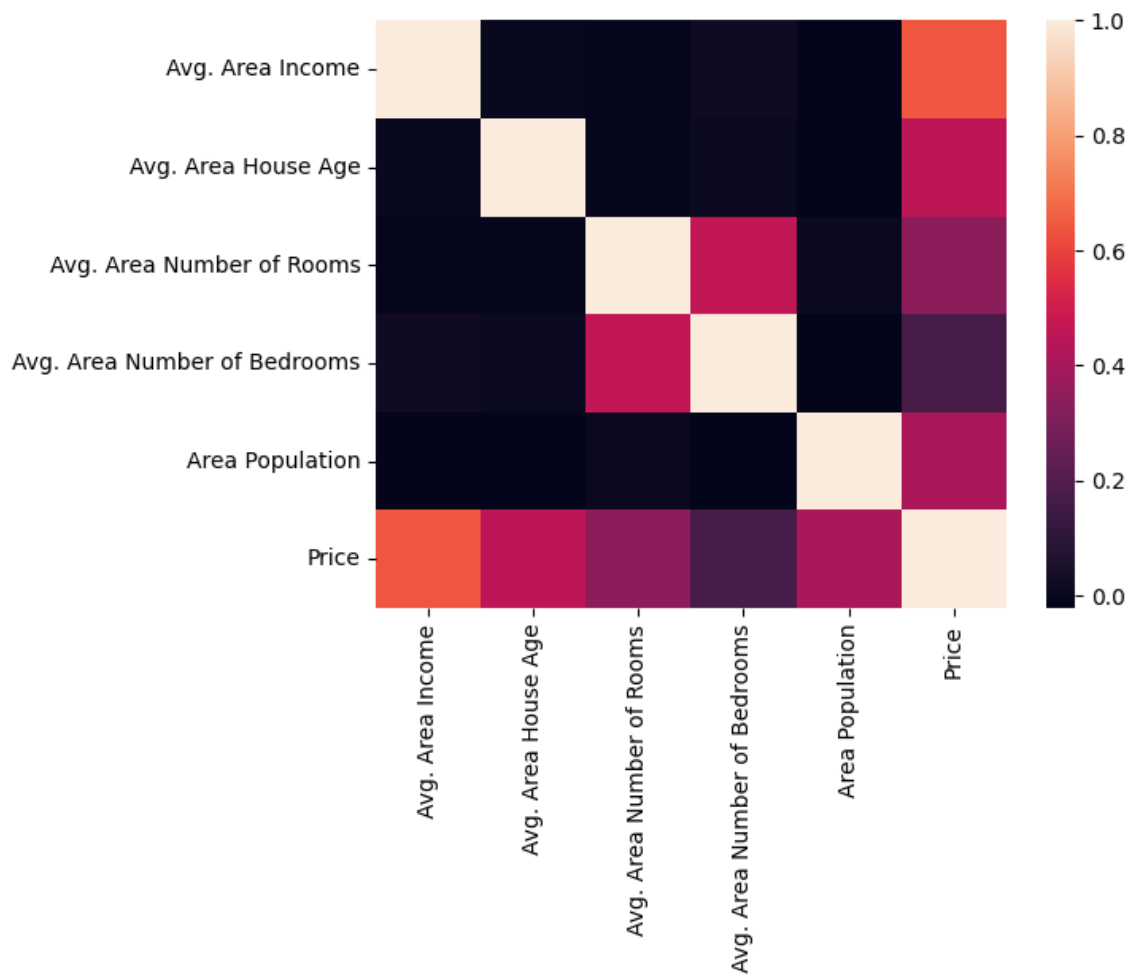

In [20]:

```python
Housedf=df[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms',
        'Avg. Area Number of Bedrooms', 'Area Population', 'Price']]
```

In [21]:

```python
sns.heatmap(Housedf.corr())
```

Out[21]:

```
<Axes: >
```


In [22]:

```python
x=Housedf[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms',
       'Avg. Area Number of Bedrooms', 'Area Population']]
y=df['Price']
```

In [23]:

```
pip install sklearn
```

```
Collecting sklearn
  Using cached sklearn-0.0.post5.tar.gz (3.7 kB)
  Installing build dependencies: started
  Installing build dependencies: finished with status 'done'
  Getting requirements to build wheel: started
  Getting requirements to build wheel: finished with status 'done'
  Preparing metadata (pyproject.toml): started
  Preparing metadata (pyproject.toml): finished with status 'done'
Building wheels for collected packages: sklearn
  Building wheel for sklearn (pyproject.toml): started
  Building wheel for sklearn (pyproject.toml): finished with status 'done'
  Created wheel for sklearn: filename=sklearn-0.0.post5-py3-none-any.whl s
ize=2968 sha256=2c22b3da1dd41fa627c765374757df2ed2b971c32a9ef8a9f0aaf5c34f
ed72e4
  Stored in directory: c:\users\sasidhar royal\appdata\local\pip\cache\whe
els\5f\28\a6\4e4fc2959e4ed9b33bf517703534fd8b19b76a842f74c9ed4c
Successfully built sklearn
Installing collected packages: sklearn
Successfully installed sklearn-0.0.post5
Note: you may need to restart the kernel to use updated packages.
```

In [24]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=101)
```

In [25]:

```
pip install scikitlearn
```

```
Note: you may need to restart the kernel to use updated packages.

ERROR: Could not find a version that satisfies the requirement scikitlearn
(from versions: none)
ERROR: No matching distribution found for scikitlearn
```

In [26]:

```
pip install --upgrade pip
```

```
Requirement already satisfied: pip in c:\users\sasidhar royal\appdata\loca
l\programs\python\python311\lib\site-packages (23.1.2)
Note: you may need to restart the kernel to use updated packages.
```

In [27]:

```
pip install scikit-learn
```

Requirement already satisfied: scikit-learn in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (1.2.2)
Requirement already satisfied: numpy>=1.17.3 in c:\users\sasidhar royal\ap
pdata\local\programs\python\python311\lib\site-packages (from scikit-lear
n) (1.24.3)
Requirement already satisfied: scipy>=1.3.2 in c:\users\sasidhar royal\app
data\local\programs\python\python311\lib\site-packages (from scikit-learn)
(1.10.1)
Requirement already satisfied: joblib>=1.1.1 in c:\users\sasidhar royal\ap
pdata\local\programs\python\python311\lib\site-packages (from scikit-lear
n) (1.2.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\sasidhar r
oyal\appdata\local\programs\python\python311\lib\site-packages (from sciki
t-learn) (3.1.0)
Note: you may need to restart the kernel to use updated packages.

In [28]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=101)
```

In [29]:

```
from sklearn.model_selection import train_test_split
```

In [30]:

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=101)
```

In [32]:

```
from sklearn.linear_model import LinearRegression
lm=LinearRegression()
lm.fit(x_train,y_train)
print(lm.intercept_)
```

-2641372.6673013885

In [33]:

```
coeff_df=pd.DataFrame(lm.coef_,x.columns,columns=['coefficient'])
coeff_df
```
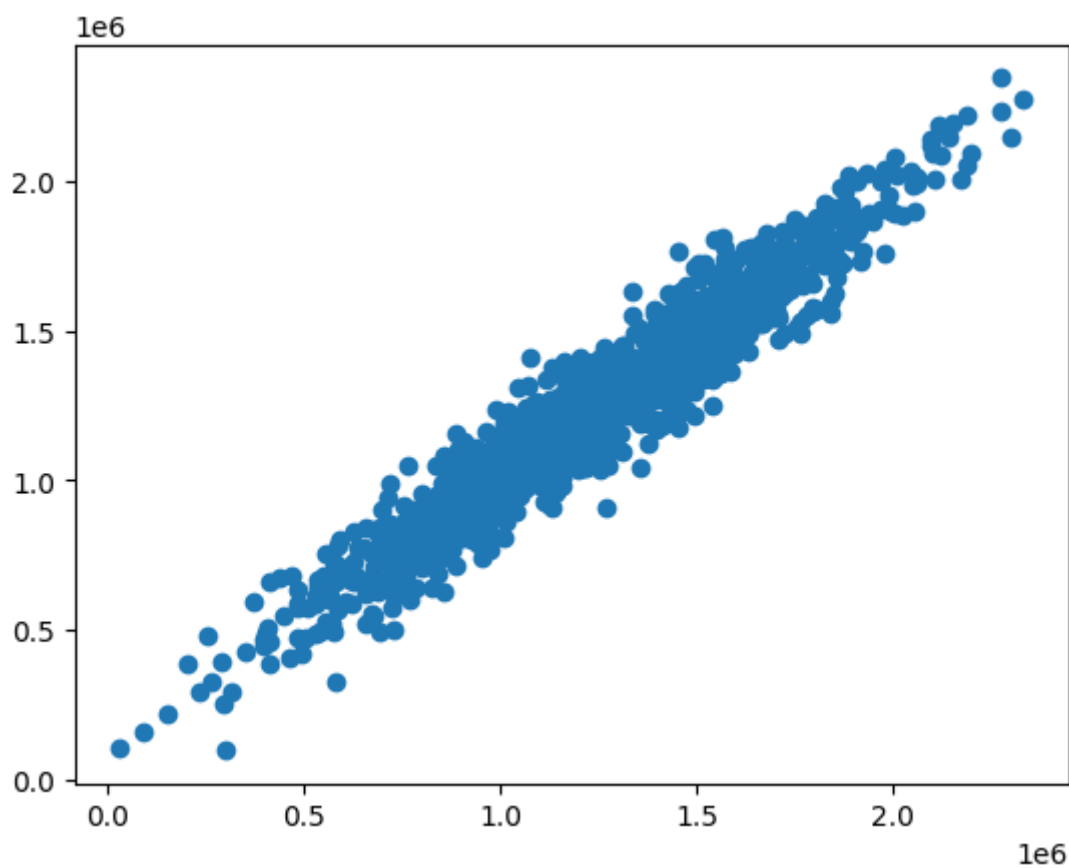
Out[33]:

	coefficient
Avg. Area Income	21.617635
Avg. Area House Age	165221.119872
Avg. Area Number of Rooms	121405.376596
Avg. Area Number of Bedrooms	1318.718783
Area Population	15.225196

In [34]:

```
predictions=lm.predict(x_test)
plt.scatter(y_test,predictions)
```
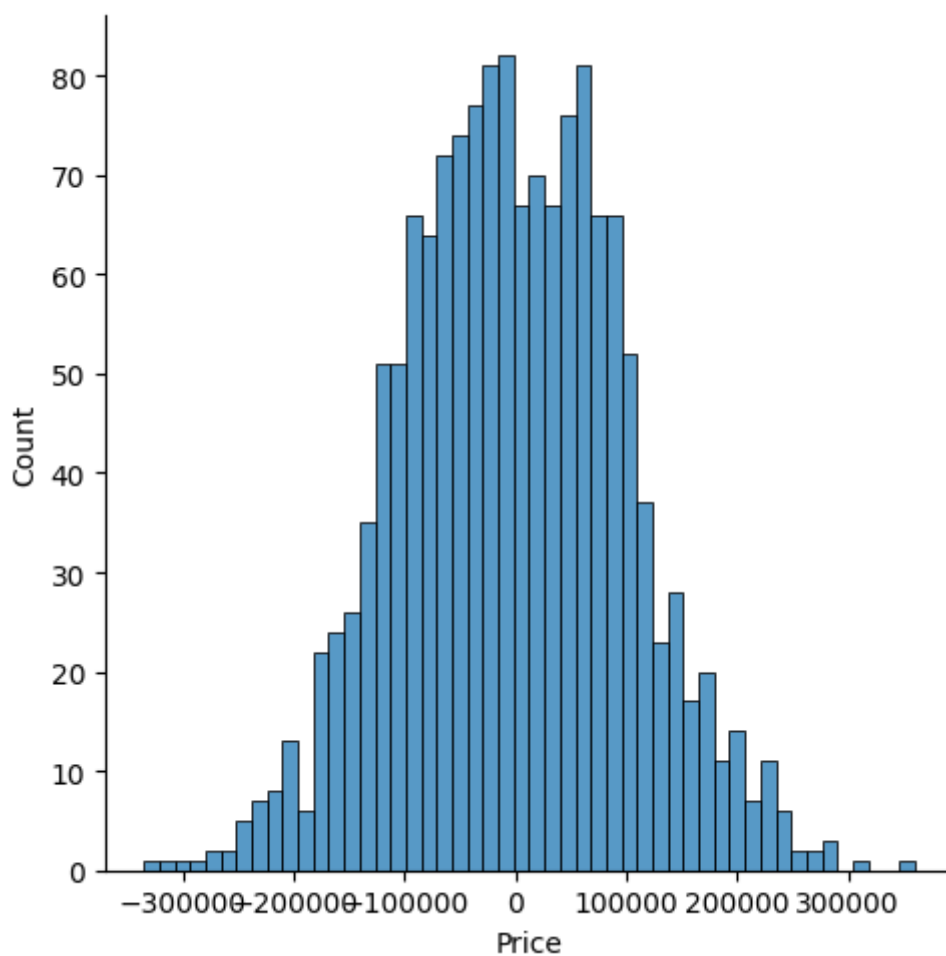
Out[34]:

```
<matplotlib.collections.PathCollection at 0x20c94a65a50>
```

In [36]:

```python
sns.displot((y_test-predictions),bins=50);
```


In [37]:

```python
from sklearn import metrics
```

In [40]:

```python
print('MAE:',metrics.mean_absolute_error(y_test,predictions))
print('MSE:',metrics.mean_squared_error(y_test,predictions))
print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,predictions)))
```

```
MAE: 81257.55795855928
MSE: 10169125565.897568
RMSE: 100842.0823163503
```

In [ ]: