U.S Housing Price data analysis

In [1]:
```
import numpy as np
import pandas as pd
```

In [3]:
```
usa = pd.read_csv("usa_housing.csv")
usa.head()
```

Out[3]:

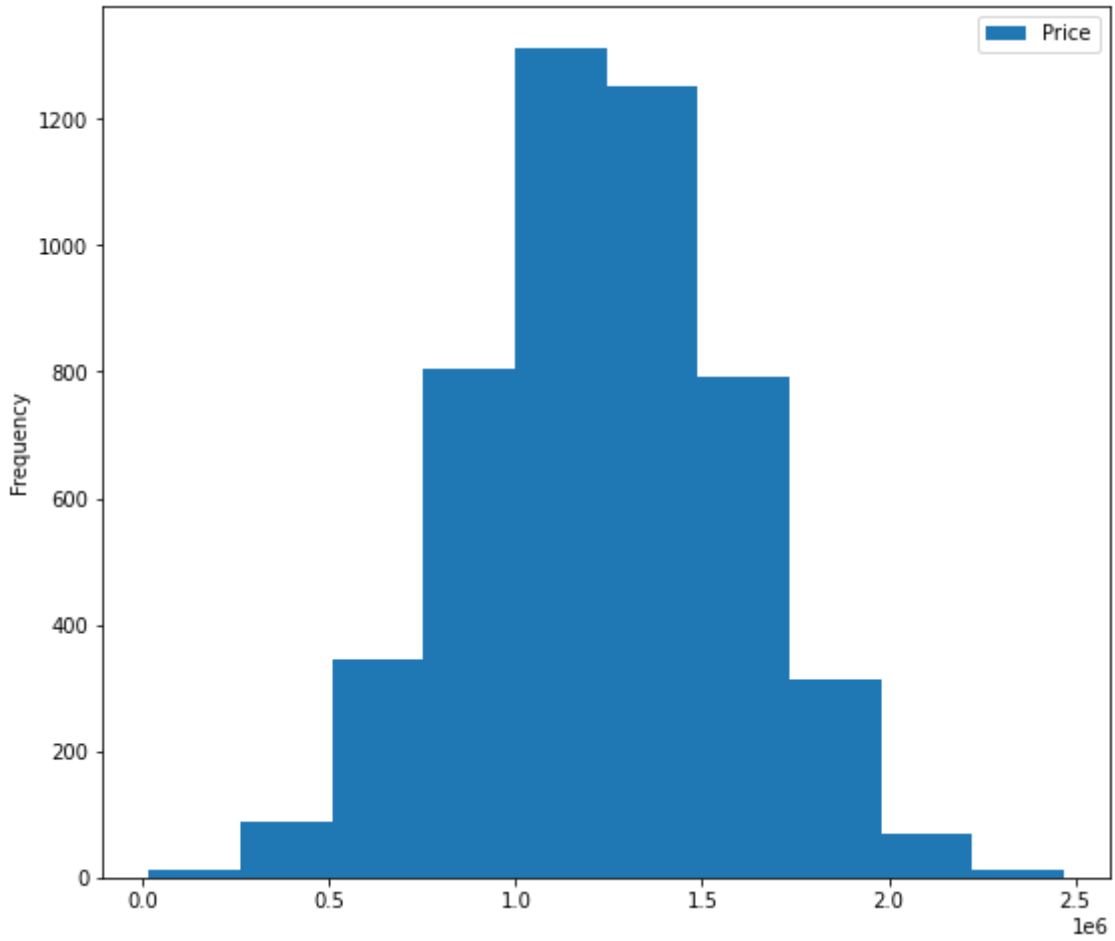
	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price	Address
0	79545.458574	5.682861	7.009188	4.09	23086.800503	1.059034e+06	208 Michael Ferry Apt 674\nLaurabury, NE 3701..
1	79248.642455	6.002900	6.730821	3.09	40173.072174	1.505891e+06	188 Johnson Views Suite 079\nLake Kathleen, CA..
2	61287.067179	5.865890	8.512727	5.13	36882.159400	1.058988e+06	9127 Elizabeth Stravenue\nDanieltown WI 06482..
3	63345.240046	7.188236	5.586729	3.26	34310.242831	1.260617e+06	USS Barnett\nFPO AP 44820
4	59982.197226	5.040555	7.839388	4.23	26354.109472	6.309435e+05	USNS Raymond\nFPO AE 09386

In [16]:
```
y = usa["Price"]
x = usa["Avg. Area Income"]
cor_ = y.corr(x)
print(cor_)
```

```
0.6397337782498941
```

1. Histogram plot for Avg. Area Income

In [29]:
```
import matplotlib.pyplot as plt
df = pd.DataFrame(usa, columns = ["Avg. Area Income","Price"])
df.plot(x = "Avg. Area Income",y = "Price", kind  = "hist", figsize = (9,8))
plt.show()
```

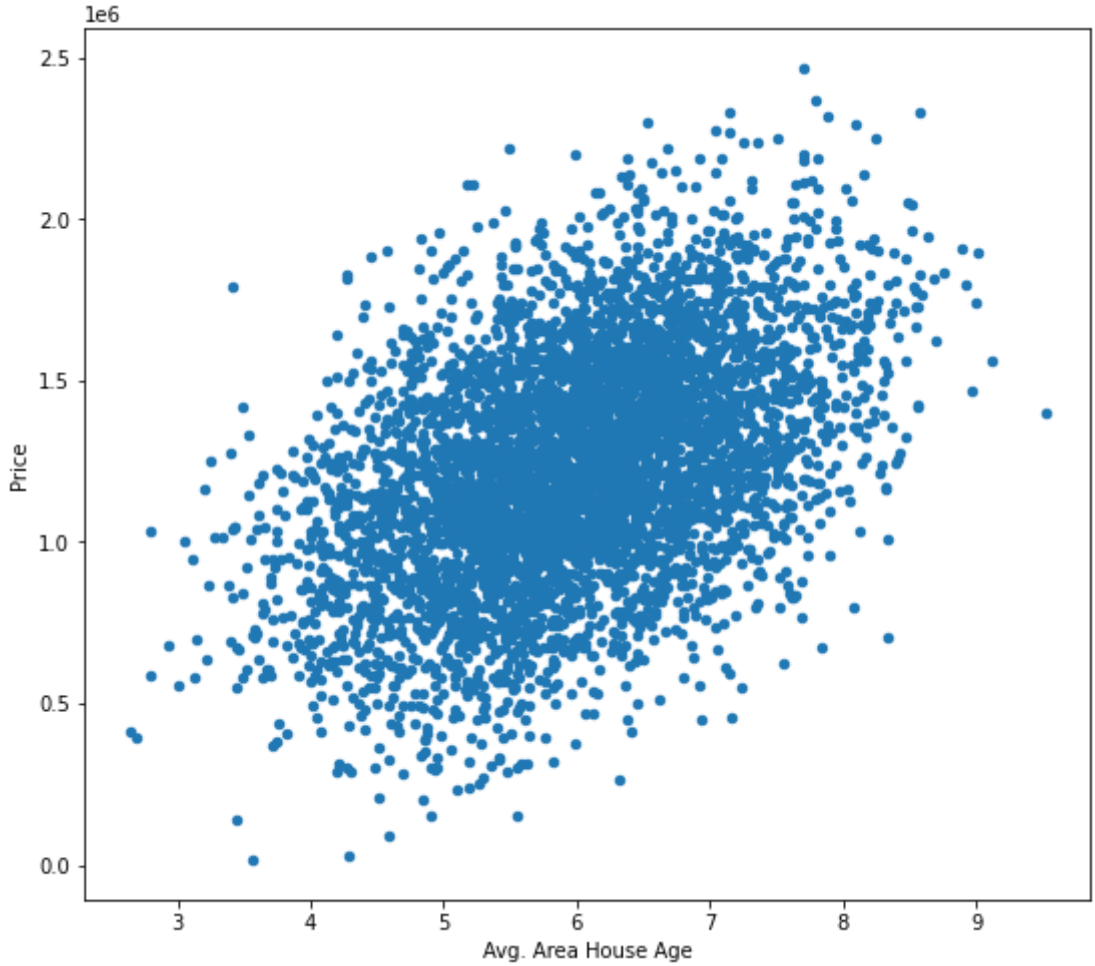


In [19]:
```
x = usa["Avg. Area House Age"]
cor_ = y.corr(x)
print(cor_)
```

```
0.45254253717875587
```

2. Scatter plot for House Area

In [30]:
```
import matplotlib.pyplot as plt
df = pd.DataFrame(usa, columns = ["Avg. Area House Age","Price"])
df.plot(x = "Avg. Area House Age",y = "Price", kind  = "scatter", figsize = (9,8))
plt.show()
```

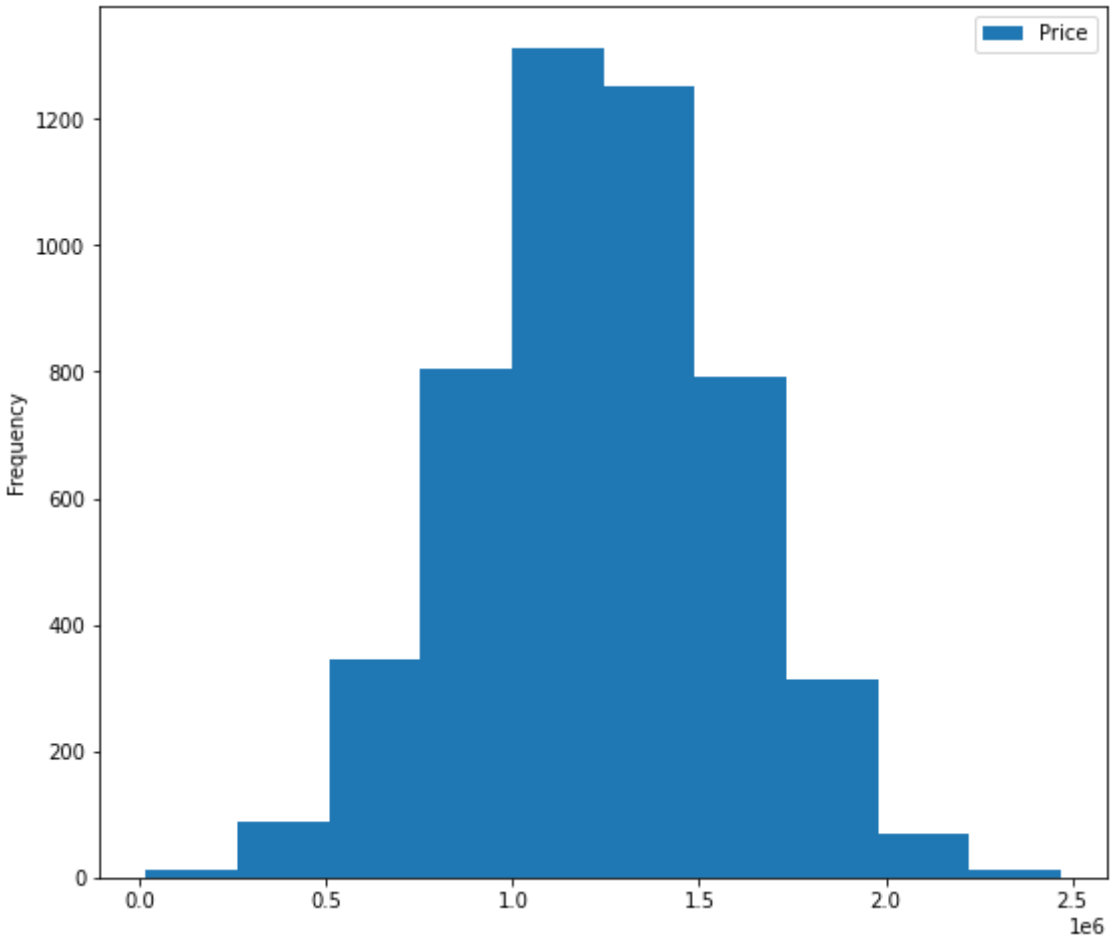


In [25]:
```
x = usa["Avg. Area Number of Rooms"]
cor_ = y.corr(x)
print(cor_)
```

```
0.3356644533667596
```

3. Histogram plot for Avg. Area Number of Rooms

In [27]:
```
import matplotlib.pyplot as plt
df = pd.DataFrame(usa, columns = ["Avg. Area Number of Rooms","Price"])
df.plot(x = "Avg. Area Number of Rooms",y = "Price", kind  = "hist", figsize = (9,
plt.show()
```



From the above data analysis, it is clear that the Avg. Area Income, House Area, number of bedrooms and other factors are affecting the price of the Houses in U.S.

Note:- [In the data analysis, if correlation is greater than 0.2 then it means that there exists a relation between 2 parameters.]

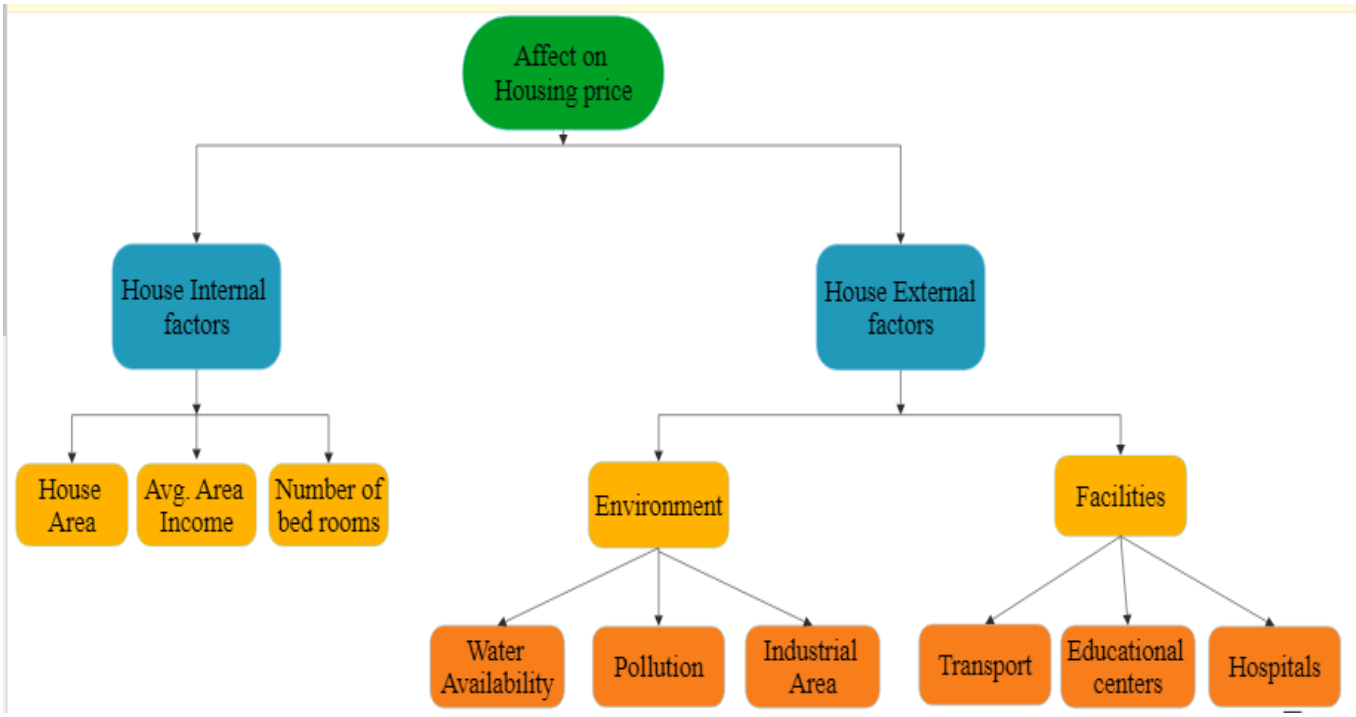


Fig. Flowchart for the factors affecting the Housing prices

Not only the parameters which are mentioned in the flow chart are going to affect the house prices, there are many other factors which affect the cost of living such as

- Population growth
- Unemployment
- Inflation
- Changes in the political party and many others are going to affect the house price in U.S for the next decade.

For example: Nowadays, because of sudden war between Ukraine and Russia has affected the GDP of the United States which causes recession which will also affect the prices of houses.

In conclusion to that, Housing prices are not always same that is because of the changing world affecting the U.S economy.