

# Importing libraries

```
In [1]: import pandas as pd  
import numpy as np  
from numpy import cov  
from scipy.stats import spearmanr  
from scipy.stats import pearsonr
```

```
In [2]: df=pd.read_csv(r"C:\Users\user\Desktop\Ash\Datasets\USA_Housing.csv")
df
```

Out[2]:

	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\nDanielstown, 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
...	...	...	...	...	...	...	...
4995	60567.944140	7.830362	6.137356	3.46	22837.361035	1.060194e+06	USNS Williams\nFPO AP 30153- 7653
4996	78491.275435	6.999135	6.576763	4.02	25616.115489	1.482618e+06	PSC 9258, Box 8489\nAPO AA 42991-3352
4997	63390.686886	7.250591	4.805081	2.13	33266.145490	1.030730e+06	4215 Tracy Garden Suite 076\nJoshualand, VA 01...
4998	68001.331235	5.534388	7.130144	5.44	42625.620156	1.198657e+06	USS Wallace\nFPO AE 73316
4999	65510.581804	5.992305	6.792336	4.07	46501.283803	1.298950e+06	37778 George Ridges Apt. 509\nEast Holly, NV 2...

5000 rows × 7 columns

## Mean

```
In [3]: print(df.mean())
```

```
Avg. Area Income          6.858311e+04
Avg. Area House Age       5.977222e+00
Avg. Area Number of Rooms  6.987792e+00
Avg. Area Number of Bedrooms 3.981330e+00
Area Population           3.616352e+04
Price                    1.232073e+06
dtype: float64
```

## Median

```
In [4]: print(df.median())
```

```
Avg. Area Income          6.880429e+04
Avg. Area House Age       5.970429e+00
Avg. Area Number of Rooms  7.002902e+00
Avg. Area Number of Bedrooms 4.050000e+00
Area Population           3.619941e+04
Price                    1.232669e+06
dtype: float64
```

## Mode

```
In [5]: print(df.mode())
```

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	\
0	17796.631190	2.644304	3.236194	
1	35454.714659	2.683043	3.950225	
2	35608.986237	2.797215	3.950973	
3	35797.323122	2.797619	3.969632	
4	35963.330809	2.922736	4.027931	
...	...	...	...	
4995	101599.670580	8.973441	10.024375	
4996	101928.858060	8.991399	10.144988	
4997	102881.120902	9.008900	10.219902	
4998	104702.724257	9.125283	10.280022	
4999	107701.748378	9.519088	10.759588	

	Avg. Area Number of Bedrooms	Area Population	Price	\
0	4.38	172.610686	1.593866e+04	
1	NaN	3285.450538	3.114052e+04	
2	NaN	3883.448164	8.859177e+04	
3	NaN	4114.489353	1.430274e+05	
4	NaN	5727.485885	1.515271e+05	
...	...	...	...	
4995	NaN	68311.695822	2.318286e+06	
4996	NaN	69553.988327	2.330290e+06	
4997	NaN	69575.449464	2.332111e+06	
4998	NaN	69592.040236	2.370231e+06	
4999	NaN	69621.713378	2.469066e+06	

	Address
0	000 Adkins Crescent\nSouth Teresa, AS 49642-1348
1	000 Todd Pines\nAshleyberg, KY 90207-1179
2	001 Steve Plaza\nJessicastad, UT 25190
3	0010 Gregory Loaf\nSouth Ericfort, VA 34651-0718
4	00149 Raymond Knolls\nNew Jason, UT 75026
...	...
4995	Unit 9774 Box 4511\nDPO AE 44963
4996	Unit 9778 Box 2114\nDPO AP 59374
4997	Unit 9785 Box 0790\nDPO AP 60371-0797
4998	Unit 9831 Box 7128\nDPO AA 54705
4999	Unit 9871 Box 9037\nDPO AP 37275-9289

```
[5000 rows x 7 columns]
```

# Statistical data

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

Out[6]:

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

## Sum

In [7]: `print(df.sum())`

```
Avg. Area Income      342915544.919799
Avg. Area House Age    29886.110176
Avg. Area Number of Rooms  34938.959255
Avg. Area Number of Bedrooms  19906.65
Area Population      180817580.192873
Price      6160363270.711784
Address      208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
dtype: object
```

## Cumsum

In [12]: df.cumsum()

Out[12]:

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price	Address
0	7.954546e+04	5.682861	7.009188	4.09	2.308680e+04	1.059034e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
1	1.587941e+05	11.685761	13.740009	7.18	6.325987e+04	2.564924e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
2	2.200812e+05	17.551651	22.252737	12.31	1.001420e+05	3.623912e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
3	2.834264e+05	24.739887	27.839465	15.57	1.344523e+05	4.884529e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
4	3.434086e+05	29.780442	35.678853	19.80	1.608064e+05	5.515473e+06	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
...	...	...	...	...	...	...	...
4995	3.426402e+08	29860.333757	34913.654931	19890.99	1.806696e+08	6.155352e+09	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
4996	3.427186e+08	29867.332892	34920.231694	19895.01	1.806952e+08	6.156835e+09	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
4997	3.427820e+08	29874.583483	34925.036775	19897.14	1.807285e+08	6.157866e+09	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
4998	3.428500e+08	29880.117871	34932.166918	19902.58	1.807711e+08	6.159064e+09	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
4999	3.429155e+08	29886.110176	34938.959255	19906.65	1.808176e+08	6.160363e+09	208 Michael Ferry Apt. 674\nLaurabury, NE 3701...

5000 rows × 7 columns

# Count

```
In [8]: print(df.count())
```

```
Avg. Area Income          5000
Avg. Area House Age        5000
Avg. Area Number of Rooms  5000
Avg. Area Number of Bedrooms 5000
Area Population            5000
Price                     5000
Address                   5000
dtype: int64
```

## Min

```
In [9]: print(df.min())
```

```
Avg. Area Income          17796.63119
Avg. Area House Age         2.644304
Avg. Area Number of Rooms   3.236194
Avg. Area Number of Bedrooms 2.0
Area Population            172.610686
Price                     15938.657923
Address                   000 Adkins Crescent\nSouth Teresa, AS 49642-1348
dtype: object
```

## Max

```
In [10]: print(df.max())
```

```
Avg. Area Income          107701.748378
Avg. Area House Age         9.519088
Avg. Area Number of Rooms  10.759588
Avg. Area Number of Bedrooms 6.5
Area Population            69621.713378
Price                     2469065.594175
Address                   Unit 9871 Box 9037\nDPO AP 37275-9289
dtype: object
```

## Covariance

```
In [13]: cov(df["Avg. Area House Age"],df["Avg. Area Number of Rooms"])
```

```
Out[13]: array([[ 0.98298536, -0.00940227],  
                [-0.00940227,  1.01170049]])
```

## Correaltion

```
In [14]: pearsonr(df['Avg. Area House Age'],df['Avg. Area Number of Rooms'])
```

```
Out[14]: (-0.009428292244463858, 0.5050729943851315)
```

```
In [15]: spearmanr(df['Avg. Area House Age'],df['Avg. Area Number of Rooms'])
```

```
Out[15]: SpearmanrResult(correlation=-0.0051978531999141275, pvalue=0.7132801669661967)
```

```
In [ ]:
```

```
In [ ]:
```