Correlation Matrix for All Columns

Descriptive statistics

```
df.describe()
```

| | | | number of | | number of | | living area | | lot area | | number of |

```
df['Number of schools nearby'].value_counts()
```

```
3    4973
2    4853
1    4794
Name: Number of schools nearby, dtype: int64
```

```
df['condition of the house'].value_counts()
```

```
3    9350
4    3874
5    1278
2     100
1      18
Name: condition of the house, dtype: int64
```

8 rows × 23 columns

▾ Handling the missing values

```
df.isnull().sum()
```

```
id                                      0
Date                                    0
number of bedrooms                      0
number of bathrooms                     0
living area                             0
lot area                                0
number of floors                        0
waterfront present                      0
number of views                         0
condition of the house                  0
grade of the house                      0
Area of the house(excluding basement)   0
Area of the basement                    0
Built Year                              0
Renovation Year                         0
Postal Code                             0
Lattitude                               0
Longitude                               0
living_area_renov                       0
lot_area_renov                          0
Number of schools nearby                0
Distance from the airport               0
Price                                   0
dtype: int64
```

```
df.dropna()
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	num vi
0	6762810145	42491	5	2.50	3650	9050	2.0	0	
1	6762810635	42491	4	2.50	2920	4000	1.5	0	
2	6762810998	42491	5	2.75	2910	9480	1.5	0	
3	6762812605	42491	4	2.50	3310	42998	2.0	0	
4	6762812919	42491	3	2.00	2710	4500	1.5	0	
...	...	...	...	...	...	...	...	...	
14615	6762830250	42734	2	1.50	1556	20000	1.0	0	
14616	6762830339	42734	3	2.00	1680	7000	1.5	0	
14617	6762830618	42734	2	1.00	1070	6120	1.0	0	
14618	6762830709	42734	4	1.00	1030	6621	1.0	0	
14619	6762831463	42734	3	1.00	900	4770	1.0	0	

14620 rows × 23 columns

✓ 0s   completed at 10:51 AM   ● ✕