







Descriptive Statistics for Numeric Attributes:

finished\_sqft bedrooms bathrooms listing\_price

count 7.345400e+04 79075.000000 76672.000000 8.411200e+04

mean 2.284015e+03 3.074334 2.274377 1.067901e+06

std 1.205212e+04 12.353739 9.383289 6.827188e+06

min 0.000000e+00 0.000000 0.000000 0.000000e+00

25% 1.168000e+03 2.000000 1.000000 3.250000e+05

50% 1.658000e+03 3.000000 2.000000 5.500000e+05

75% 2.336000e+03 4.000000 3.000000 1.000000e+06

max 1.304622e+06 3440.000000 2560.000000 9.390000e+08

Unique Values for Categorical Attributes:

home\_type ['condo' 'single-family' 'cooperative' 'multi-family' 'misc' 'townhouse'

'land' 'unknown' 'mobile' None]

status ['SOLD' 'OTHER' 'UNKNOWN' 'OFF\_MARKET' 'FOR\_SALE' 'PENDING' 'WITHDRAWN'

'COMING\_SOON']

Min listing\_contract\_date: 1900-01-01 00:00:00, Max listing\_contract\_date: 2024-12-27 06:00:00

Min on\_market\_date: NaT, Max on\_market\_date: NaT

Min pending\_date: 1800-01-01 00:00:00, Max pending\_date: 2024-03-27 19:21:42

Min last\_sold\_date: 1800-01-01 00:00:00, Max last\_sold\_date: 2025-08-01 00:00:00

Min off\_market\_date: 1800-01-01 00:00:00, Max off\_market\_date: 2024-03-27 19:20:18

Number of missing values per column:

id 0

listing\_key 78

source\_system 0

address 0

usps\_address 4

status 0

listing\_contract\_date 534

on\_market\_date 84112

pending\_date 57164

last\_sold\_date 67141

off\_market\_date 53019

original\_listing\_price 33808

listing\_price 0

last\_sold\_price 68416

home\_type 12

finished\_sqft 10658

lot\_size\_sqft 16505

bedrooms 5037

bathrooms 7440

year\_built 4835

new\_construction 0

has\_pool 0

state\_market\_id 0

city\_market\_id 0

zipcode\_market\_id 0

neighborhood\_level\_1\_market\_id 12863

neighborhood\_level\_2\_market\_id 14130

neighborhood\_level\_3\_market\_id 69461

long 5

lat 5

crawler 0

dtype: int64

Outlier Homes:

id finished\_sqft listing\_price

27 186061 16000.0 21000000.0

51 201161 4820.0 5995000.0

78 315906 2226.0 5900000.0

84 319861 12200.0 1075000.0

87 453401 0.0 4500000.0

... ... ...

83913 37466528 4220.0 1399900.0

83989 37486866 NaN 6000000.0

84035 37528922 2700.0 518000.0

84036 37529004 5760.0 899900.0

84050 37532082 3550.0 1500000.0

Outliers in 'finished\_sqft' column:

id listing\_key source\_system ... lat crawler outlier

27 186061 358560 SFAR ... 37.795891 SFAR -1

51 201161 342932 SFAR ... 37.792541 SFAR -1

84 319861 360373 SFAR ... 37.748022 SFAR -1

89 547854 363274 SFAR ... 37.802682 SFAR -1

124 1909180 None ATTOM ... 37.783362 ATTOM -1

... ... ... ... ... ... ...

83832 37515342 802832824816 BRIGHT MLS ... 39.938372 BRIGHT MLS -1

83855 37455887 802827853984 BRIGHT MLS ... 40.006420 BRIGHT MLS -1

83913 37466528 802828848464 BRIGHT MLS ... 39.970383 BRIGHT MLS -1

83923 37467931 802828173276 BRIGHT MLS ... 39.945300 BRIGHT MLS 1

84036 37529004 802834245106 BRIGHT MLS ... 39.956350 BRIGHT MLS -1