|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| V  A  R  I  A  B  L  E | T  Y  P  E | N  T  Y  P  E | R  A  N  G  E | N  A  N  s | O  U  T  L  I  E  R | I  N  S  I  D  E  S |
| id | int64 |  | 1000102 :  9900000190 | - | - | is there some logic?  there are some cutouts,  e.g. 4500000000 : 5000000000 |
| bedrooms | float64 | int64 | 01:33 | - | (yes) | most have 3 or 4 |
| bathrooms | float64 |  | 0.5 : 8 | - | - | most have 2.5  measured in steps of 0.25 |
| sqft\_living | float64 | int64 | 370 : 13.540 | - | - | what is the unit?  median is 2000 |
| sqft\_lot | float64 | int64 | 520: 1.651.359 | - | yes | most are < 30.000  5.000 : 10.000 |
| floors | float64 |  | 1 : 3.5 | - | - | majority hat 1 (10.000) or 2 (8.000) floors measured in steps of 0.5 |
| waterfront | float64 | object | 0 , 1 | 2.391 (0) | - | only 146 with waterfront |
| view | float64 | int64 | 0 : 4 | 64 (-1) | - | 0=19422 |
| condition | int64 | cat | 1 : 5 | - | - | 3=14.000 |
| grade | int64 | int64 | 3 : 13 | - | - | 7=9.000, 8=6.000 |
| sqft\_above | float64 | int64 | 370 : 9.410 | - | - | most have 1.500,  over 6.000 is rare |
| sqft\_basement | float64 | int64 | 0 : 4.820 | 452  (-1) | - | lot of zeros(12.827),  which are also NAs?  or 1 floor houses? |
| yr\_built | int64 | date | 1900 : 2015 | - | - | not up to date data, stops in Beginning 2015,  increasing number 2014=559,  2015=38 |
| yr\_renovated | float64 | date | 0 : 20150 | 3848 | - | need to earse a 0, 0=17.005 can be NAs or new, calculate 2 new column:  How many years are between  built and renovation,  how long is renovation from today |
| zipcode | int64 | cat | 98001 : 98199 | - | - | for some only 50=98039 houses listed for others 600 =98103 |
| lat | float64 |  | 47.1559 : 47.7776 | - | - | most between 47.5 and 47.77  with two main peaks |
| long | float64 |  | -122.519 : -121.315 | - | - | under -122.4 and over -121.7 is rare |
| sqft\_living15 | float64 |  | 399 : 6.210 | - | - | peak at 1700, very few over 5000 |
| sqft\_lot15 | float64 |  | 651 : 871.200 | - | yes | what is the unit? maybe round it,  most under 20.000, thousands values more frequent 5.000 (quite some on 220.000) |
| date | object | date | 2014-05-02 :  2015-05-27 | - | - | we only have sales data from 1 year |
| price | float64 |  | 78.000 : 7.700.000 | - | yes | most between 200.000 : 800.000,  over 1.000.000 is rare |