In [1]:

```
import pandas as pd
house = pd.read_csv('house_infos.csv')
```

In [2]:

```
print(house)
```

```
      LotFrontage  LotArea Street  OverallQual  YearBuilt BsmtQual  \
0            65.0     8450   Pave            7       2003       Gd
1            80.0     9600   Pave            6       1976       Gd
2            68.0    11250   Pave            7       2001       Gd
3            60.0     9550   Pave            7       1915       TA
4            84.0    14260   Pave            8       2000       Gd
5            85.0    14115   Pave            5       1993       Gd
6            75.0    10084   Pave            8       2004       Ex
7             NaN    10382   Pave            7       1973       Gd
8            51.0     6120   Pave            7       1931       TA
9            50.0     7420   Pave            5       1939       TA
10           70.0    11200   Pave            5       1965       TA
11           85.0    11924   Pave            9       2005       Ex
12            NaN    12968   Pave            5       1962       TA
13           91.0    10652   Pave            7       2006       Gd
14            NaN    10920   Pave            6       1960       TA
15           51.0     6120   Pave            7       1929       TA
16            NaN    11241   Pave            6       1970       TA
17           72.0    10791   Pave            4       1967      NaN
18           66.0    13695   Pave            5       2004       TA
19           70.0     7560   Pave            5       1958       TA
20          101.0    14215   Pave            8       2005       Ex
21           57.0     7449   Pave            7       1930       TA
22           75.0     9742   Pave            8       2002       Gd
23           44.0     4224   Pave            5       1976       Gd
24            NaN     8246   Pave            5       1968       TA
25          110.0    14230   Pave            8       2007       Gd
26           60.0     7200   Pave            5       1951       TA
27           98.0    11478   Pave            8       2007       Ex
28           47.0    16321   Pave            5       1957       TA
29           60.0     6324   Pave            4       1927       TA
...           ...      ...    ...          ...        ...      ...
1430         60.0    21930   Pave            5       2005       Gd
1431          NaN     4928   Pave            6       1976       Gd
1432         60.0    10800   Pave            4       1927       TA
1433         93.0    10261   Pave            6       2000       Gd
1434         80.0    17400   Pave            5       1977       TA
1435         80.0     8400   Pave            6       1962       TA
1436         60.0     9000   Pave            4       1971       TA
1437         96.0    12444   Pave            8       2008       Ex
1438         90.0     7407   Pave            6       1957       TA
1439         80.0    11584   Pave            7       1979       TA
1440         79.0    11526   Pave            6       1922       Ex
1441          NaN     4426   Pave            6       2004       Gd
1442         85.0    11003   Pave           10       2008       Ex
1443          NaN     8854   Pave            6       1916       TA
1444         63.0     8500   Pave            7       2004       Gd
1445         70.0     8400   Pave            6       1966       TA
1446          NaN    26142   Pave            5       1962       TA
1447         80.0    10000   Pave            8       1995       Gd
1448         70.0    11767   Pave            4       1910       Fa
1449         21.0     1533   Pave            5       1970       Gd
1450         60.0     9000   Pave            5       1974       Gd
1451         78.0     9262   Pave            8       2008       Gd
1452         35.0     3675   Pave            5       2005       Gd
1453         90.0    17217   Pave            5       2006       Gd
1454         62.0     7500   Pave            7       2004       Gd
1455         62.0     7917   Pave            6       1999       Gd
1456         85.0    13175   Pave            6       1978       Gd
1457         66.0     9042   Pave            7       1941       TA
1458         68.0     9717   Pave            5       1950       TA
```

```
1459         75.0     9937   Pave            5       1965       TA

      TotalBsmtSF CentralAir  GrLivArea  PoolArea  GarageArea
0             856          Y       1710         0         548
1            1262          Y       1262         0         460
2             920          Y       1786         0         608
3             756          Y       1717         0         642
4            1145          Y       2198         0         836
5             796          Y       1362         0         480
6            1686          Y       1694         0         636
7            1107          Y       2090         0         484
8             952          Y       1774         0         468
9             991          Y       1077         0         205
10           1040          Y       1040         0         384
11           1175          Y       2324         0         736
12            912          Y        912         0         352
13           1494          Y       1494         0         840
14           1253          Y       1253         0         352
15            832          Y        854         0         576
16           1004          Y       1004         0         480
17              0          Y       1296         0         516
18           1114          Y       1114         0         576
19           1029          Y       1339         0         294
20           1158          Y       2376         0         853
21            637          Y       1108         0         280
22           1777          Y       1795         0         534
23           1040          Y       1060         0         572
24           1060          Y       1060         0         270
25           1566          Y       1600         0         890
26            900          Y        900         0         576
27           1704          Y       1704         0         772
28           1484          Y       1600         0         319
29            520          N        520         0         240
...           ...        ...        ...       ...         ...
1430          732          Y       1838         0         372
1431          958          Y        958         0         440
1432          656          Y        968         0         216
1433          936          Y       1792         0         451
1434         1126          Y       1126         0         484
1435         1319          Y       1537         0         462
1436          864          Y        864         0         528
1437         1932          Y       1932         0         774
1438          912          Y       1236         0         923
1439          539          Y       1725         0         550
1440          588          Y       2555         0         672
1441          848          Y        848         0         420
1442         1017          Y       2007         0         812
1443          952          N        952         0         192
1444         1422          Y       1422         0         626
1445          814          Y        913         0         240
1446         1188          Y       1188         0         312
1447         1220          Y       2090         0         556
1448          560          N       1346         0         384
1449          630          Y        630         0           0
1450          896          Y       1792         0           0
1451         1573          Y       1578         0         840
1452          547          Y       1072         0         525
1453         1140          Y       1140         0           0
1454         1221          Y       1221         0         400
1455          953          Y       1647         0         460
1456         1542          Y       2073         0         500
```

```
1457          1152           Y        2340          0          252
1458          1078           Y        1078          0          240
1459          1256           Y        1256          0          276

[1460 rows x 11 columns]
```

In [3]:

```
cont =0
for i in range (0,1459):
        if ((house.loc[i][1]>5000) and (house.loc[i][7]=='N')):
            cont=cont+1
print('Quantidade de casas que possuem “LotArea” > 5000 e “CentralAir” = {}'.format(
cont))
```

```
Quantidade de casas que possuem “LotArea” > 5000 e “CentralAir” = 87
```

In [4]:

```
copy = house.filter(items=['GrLivArea'])
copy = copy.loc[copy['GrLivArea']>2000]
copy
```

Out[4]:

	GrLivArea
4	2198
7	2090
11	2324
20	2376
35	2452
46	2149
58	2945
64	2034
65	2473
66	2207
69	2287
70	2223
80	2142
85	2417
93	2291
112	2696
113	2259
114	2320
118	3222
130	2157
131	2054
147	2035
159	2462
161	2668
167	2161
175	2158
176	2060
178	2234
183	2082
185	3608
...	...
1284	2447
1292	2372

	GrLivArea
1298	5642
1302	2526
1312	2810
1313	2599
1315	2112
1328	2792
1342	2169
1346	2156
1349	2358
1350	2634
1353	3238
1360	2601
1372	2097
1373	2633
1381	2117
1386	2784
1387	2526
1395	2482
1409	2093
1416	2290
1417	2450
1423	2201
1426	2127
1440	2555
1442	2007
1447	2090
1456	2073
1457	2340

214 rows × 1 columns

In [5]:

```
copy["SalePrice"]=0
copy
```

Out[5]:

	GrLivArea	SalePrice
4	2198	0
7	2090	0
11	2324	0
20	2376	0
35	2452	0
46	2149	0
58	2945	0
64	2034	0
65	2473	0
66	2207	0
69	2287	0
70	2223	0
80	2142	0
85	2417	0
93	2291	0
112	2696	0
113	2259	0
114	2320	0
118	3222	0
130	2157	0
131	2054	0
147	2035	0
159	2462	0
161	2668	0
167	2161	0
175	2158	0
176	2060	0
178	2234	0
183	2082	0
185	3608	0
...	...	...
1284	2447	0
1292	2372	0

	GrLivArea	SalePrice
1298	5642	0
1302	2526	0
1312	2810	0
1313	2599	0
1315	2112	0
1328	2792	0
1342	2169	0
1346	2156	0
1349	2358	0
1350	2634	0
1353	3238	0
1360	2601	0
1372	2097	0
1373	2633	0
1381	2117	0
1386	2784	0
1387	2526	0
1395	2482	0
1409	2093	0
1416	2290	0
1417	2450	0
1423	2201	0
1426	2127	0
1440	2555	0
1442	2007	0
1447	2090	0
1456	2073	0
1457	2340	0

214 rows × 2 columns

In [6]:

```
copy.shape
```

Out[6]:

```
(214, 2)
```

In [7]:

```
cont=0
s=0
e=0
h=0
for e in range (0,214):
    for h in range (0,1459):
        if copy.index[e]==house.index[h]:
            s =  ((house.loc[h][0]+house.loc[h][1]+house.loc[h][6]+house.loc[h][8])*
house.loc[h][3])
            copy['SalePrice'][h]=s
```

In [8]:

```
copy
```

```
Out[8]:
```

	GrLivArea	SalePrice
4	2198	141496.0
7	2090	NaN
11	2324	139572.0
20	2376	142800.0
35	2452	136760.0
46	2149	115171.0
58	2945	181030.0
64	2034	NaN
65	2473	106264.0
66	2207	NaN
69	2287	134855.0
70	2223	127344.0
80	2142	96828.0
85	2417	159080.0
93	2291	64590.0
112	2696	98014.0
113	2259	NaN
114	2320	64008.0
118	3222	121527.0
130	2157	121583.0
131	2054	NaN
147	2035	NaN
159	2462	163562.0
161	2668	162342.0
167	2161	115128.0
175	2158	96354.0
176	2060	80022.0
178	2234	197424.0
183	2082	104531.0
185	3608	277550.0
...	...	...
1284	2447	77634.0
1292	2372	50130.0

	GrLivArea	SalePrice
1298	5642	759520.0
1302	2526	111224.0
1312	2810	NaN
1313	2599	169866.0
1315	2112	85344.0
1328	2792	88812.0
1342	2169	NaN
1346	2156	NaN
1349	2358	66736.0
1350	2634	78080.0
1353	3238	160376.0
1360	2601	65530.0
1372	2097	91210.0
1373	2633	NaN
1381	2117	NaN
1386	2784	146636.0
1387	2526	70920.0
1395	2482	130632.0
1409	2093	164318.0
1416	2290	57868.0
1417	2450	NaN
1423	2201	NaN
1426	2127	100625.0
1440	2555	88488.0
1442	2007	141120.0
1447	2090	107120.0
1456	2073	101250.0
1457	2340	88200.0

214 rows × 2 columns

In [9]:

```
copy["AvgOverallQual"]=0
copy
```

Out[9]:

	GrLivArea	SalePrice	AvgOverallQual
4	2198	141496.0	0
7	2090	NaN	0
11	2324	139572.0	0
20	2376	142800.0	0
35	2452	136760.0	0
46	2149	115171.0	0
58	2945	181030.0	0
64	2034	NaN	0
65	2473	106264.0	0
66	2207	NaN	0
69	2287	134855.0	0
70	2223	127344.0	0
80	2142	96828.0	0
85	2417	159080.0	0
93	2291	64590.0	0
112	2696	98014.0	0
113	2259	NaN	0
114	2320	64008.0	0
118	3222	121527.0	0
130	2157	121583.0	0
131	2054	NaN	0
147	2035	NaN	0
159	2462	163562.0	0
161	2668	162342.0	0
167	2161	115128.0	0
175	2158	96354.0	0
176	2060	80022.0	0
178	2234	197424.0	0
183	2082	104531.0	0
185	3608	277550.0	0
...	...	...	...
1284	2447	77634.0	0
1292	2372	50130.0	0

	GrLivArea	SalePrice	AvgOverallQual
1298	5642	759520.0	0
1302	2526	111224.0	0
1312	2810	NaN	0
1313	2599	169866.0	0
1315	2112	85344.0	0
1328	2792	88812.0	0
1342	2169	NaN	0
1346	2156	NaN	0
1349	2358	66736.0	0
1350	2634	78080.0	0
1353	3238	160376.0	0
1360	2601	65530.0	0
1372	2097	91210.0	0
1373	2633	NaN	0
1381	2117	NaN	0
1386	2784	146636.0	0
1387	2526	70920.0	0
1395	2482	130632.0	0
1409	2093	164318.0	0
1416	2290	57868.0	0
1417	2450	NaN	0
1423	2201	NaN	0
1426	2127	100625.0	0
1440	2555	88488.0	0
1442	2007	141120.0	0
1447	2090	107120.0	0
1456	2073	101250.0	0
1457	2340	88200.0	0

214 rows × 3 columns

In [10]:

```
cont=0
s=0
e=0
h=0
for e in range (0,214):
    for h in range (0,1459):
        if copy.index[e]==house.index[h]:
            s =  (house.loc[h][4]/house.loc[h][3])
            copy['AvgOverallQual'][h]=s
```

```
C:\Users\DANIELWILLIANSIGNACI\Anaconda3\lib\site-packages\ipykernel_laun
cher.py:9: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-do
cs/stable/indexing.html#indexing-view-versus-copy
  if __name__ == '__main__':
```

In [11]:

```
copy
```

```
Out[11]:
```

	GrLivArea	SalePrice	AvgOverallQual
4	2198	141496.0	250
7	2090	NaN	281
11	2324	139572.0	222
20	2376	142800.0	250
35	2452	136760.0	250
46	2149	115171.0	286
58	2945	181030.0	200
64	2034	NaN	285
65	2473	106264.0	250
66	2207	NaN	281
69	2287	134855.0	279
70	2223	127344.0	281
80	2142	96828.0	328
85	2417	159080.0	248
93	2291	64590.0	318
112	2696	98014.0	286
113	2259	NaN	325
114	2320	64008.0	324
118	3222	121527.0	284
130	2157	121583.0	280
131	2054	NaN	333
147	2035	NaN	285
159	2462	163562.0	286
161	2668	162342.0	222
167	2161	115128.0	250
175	2158	96354.0	325
176	2060	80022.0	331
178	2234	197424.0	223
183	2082	104531.0	286
185	3608	277550.0	189
...	...	...	...
1284	2447	77634.0	319
1292	2372	50130.0	378

	GrLivArea	SalePrice	AvgOverallQual
1298	5642	759520.0	200
1302	2526	111224.0	249
1312	2810	NaN	248
1313	2599	169866.0	222
1315	2112	85344.0	328
1328	2792	88812.0	320
1342	2169	NaN	250
1346	2156	NaN	281
1349	2358	66736.0	234
1350	2634	78080.0	393
1353	3238	160376.0	249
1360	2601	65530.0	384
1372	2097	91210.0	285
1373	2633	NaN	200
1381	2117	NaN	328
1386	2784	146636.0	282
1387	2526	70920.0	319
1395	2482	130632.0	250
1409	2093	164318.0	283
1416	2290	57868.0	471
1417	2450	NaN	249
1423	2201	NaN	327
1426	2127	100625.0	284
1440	2555	88488.0	320
1442	2007	141120.0	200
1447	2090	107120.0	249
1456	2073	101250.0	329
1457	2340	88200.0	277

214 rows × 3 columns

In [12]:

```
import pandas as pd
year = pd.read_csv('year_condition.csv')
year
```

Out[12]:

	YearBuilt	YearCondition
0	1800	Very Old
1	1801	Very Old
2	1802	Very Old
3	1803	Very Old
4	1804	Very Old
5	1805	Very Old
6	1806	Very Old
7	1807	Very Old
8	1808	Very Old
9	1809	Very Old
10	1810	Very Old
11	1811	Very Old
12	1812	Very Old
13	1813	Very Old
14	1814	Very Old
15	1815	Very Old
16	1816	Very Old
17	1817	Very Old
18	1818	Very Old
19	1819	Very Old
20	1820	Very Old
21	1821	Very Old
22	1822	Very Old
23	1823	Very Old
24	1824	Very Old
25	1825	Very Old
26	1826	Very Old
27	1827	Very Old
28	1828	Very Old
29	1829	Very Old
...	...	...
189	1989	Old
190	1990	Old

	YearBuilt	YearCondition
191	1991	Old
192	1992	Old
193	1993	Old
194	1994	Old
195	1995	Old
196	1996	Old
197	1997	Old
198	1998	Old
199	1999	Old
200	2000	New
201	2001	New
202	2002	New
203	2003	New
204	2004	New
205	2005	New
206	2006	New
207	2007	New
208	2008	New
209	2009	New
210	2010	New
211	2011	New
212	2012	New
213	2013	New
214	2014	New
215	2015	New
216	2016	New
217	2017	New
218	2018	New

219 rows × 2 columns

In [13]:

```
result = pd.concat([year,copy])
```

```
C:\Users\DANIELWILLIANSIGNACI\Anaconda3\lib\site-packages\ipykernel_laun
cher.py:1: FutureWarning: Sorting because non-concatenation axis is not
aligned. A future version
of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=True'.

To retain the current behavior and silence the warning, pass sort=False

  """Entry point for launching an IPython kernel.
```

In [14]:

```
result
```

Out[14]:

	AvgOverallQual	GrLivArea	SalePrice	YearBuilt	YearCondition
0	NaN	NaN	NaN	1800.0	Very Old
1	NaN	NaN	NaN	1801.0	Very Old
2	NaN	NaN	NaN	1802.0	Very Old
3	NaN	NaN	NaN	1803.0	Very Old
4	NaN	NaN	NaN	1804.0	Very Old
5	NaN	NaN	NaN	1805.0	Very Old
6	NaN	NaN	NaN	1806.0	Very Old
7	NaN	NaN	NaN	1807.0	Very Old
8	NaN	NaN	NaN	1808.0	Very Old
9	NaN	NaN	NaN	1809.0	Very Old
10	NaN	NaN	NaN	1810.0	Very Old
11	NaN	NaN	NaN	1811.0	Very Old
12	NaN	NaN	NaN	1812.0	Very Old
13	NaN	NaN	NaN	1813.0	Very Old
14	NaN	NaN	NaN	1814.0	Very Old
15	NaN	NaN	NaN	1815.0	Very Old
16	NaN	NaN	NaN	1816.0	Very Old
17	NaN	NaN	NaN	1817.0	Very Old
18	NaN	NaN	NaN	1818.0	Very Old
19	NaN	NaN	NaN	1819.0	Very Old
20	NaN	NaN	NaN	1820.0	Very Old
21	NaN	NaN	NaN	1821.0	Very Old
22	NaN	NaN	NaN	1822.0	Very Old
23	NaN	NaN	NaN	1823.0	Very Old
24	NaN	NaN	NaN	1824.0	Very Old
25	NaN	NaN	NaN	1825.0	Very Old
26	NaN	NaN	NaN	1826.0	Very Old
27	NaN	NaN	NaN	1827.0	Very Old
28	NaN	NaN	NaN	1828.0	Very Old
29	NaN	NaN	NaN	1829.0	Very Old
...	...	...	...	...	...
1284	319.0	2447.0	77634.0	NaN	NaN
1292	378.0	2372.0	50130.0	NaN	NaN

	AvgOverallQual	GrLivArea	SalePrice	YearBuilt	YearCondition
1298	200.0	5642.0	759520.0	NaN	NaN
1302	249.0	2526.0	111224.0	NaN	NaN
1312	248.0	2810.0	NaN	NaN	NaN
1313	222.0	2599.0	169866.0	NaN	NaN
1315	328.0	2112.0	85344.0	NaN	NaN
1328	320.0	2792.0	88812.0	NaN	NaN
1342	250.0	2169.0	NaN	NaN	NaN
1346	281.0	2156.0	NaN	NaN	NaN
1349	234.0	2358.0	66736.0	NaN	NaN
1350	393.0	2634.0	78080.0	NaN	NaN
1353	249.0	3238.0	160376.0	NaN	NaN
1360	384.0	2601.0	65530.0	NaN	NaN
1372	285.0	2097.0	91210.0	NaN	NaN
1373	200.0	2633.0	NaN	NaN	NaN
1381	328.0	2117.0	NaN	NaN	NaN
1386	282.0	2784.0	146636.0	NaN	NaN
1387	319.0	2526.0	70920.0	NaN	NaN
1395	250.0	2482.0	130632.0	NaN	NaN
1409	283.0	2093.0	164318.0	NaN	NaN
1416	471.0	2290.0	57868.0	NaN	NaN
1417	249.0	2450.0	NaN	NaN	NaN
1423	327.0	2201.0	NaN	NaN	NaN
1426	284.0	2127.0	100625.0	NaN	NaN
1440	320.0	2555.0	88488.0	NaN	NaN
1442	200.0	2007.0	141120.0	NaN	NaN
1447	249.0	2090.0	107120.0	NaN	NaN
1456	329.0	2073.0	101250.0	NaN	NaN
1457	277.0	2340.0	88200.0	NaN	NaN

433 rows × 5 columns