Assignment_1

Chase Holland

2023-09-11

Assignment 1

This assignment was complete using data showing all home sales from 2001 - 2018 in the state of Connecticut. This data file can be accessed at the following website: https://catalog.data.gov/dataset/real-estate-sales-2001-2018.

```
#First, I have a summary of all variables.

summary(Real_Estate_Sales_2001_2020_GL_1_)
```

```
##  Serial Number        List Year    Date Recorded
##  Min.   :0.000e+00   Min.   :2001   Min.   :1999-04-05 00:00:00.00
##  1st Qu.:3.044e+04   1st Qu.:2004   1st Qu.:2005-07-25 00:00:00.00
##  Median :7.030e+04   Median :2010   Median :2011-01-19 00:00:00.00
##  Mean   :4.312e+05   Mean   :2010   Mean   :2011-06-19 11:41:32.72
##  3rd Qu.:1.519e+05   3rd Qu.:2016   3rd Qu.:2017-04-05 00:00:00.00
##  Max.   :2.001e+09   Max.   :2020   Max.   :2021-09-30 00:00:00.00
##                                     NA's   :2
##      Town             Address           Assessed Value       Sale Amount
##  Length:997212      Length:997212      Min.   :        0   Min.   :0.000e+00
##  Class :character   Class :character   1st Qu.:    87600   1st Qu.:1.400e+05
##  Mode  :character   Mode  :character   Median :   138390   Median :2.250e+05
##                                        Mean   :   278260   Mean   :3.911e+05
##                                        3rd Qu.:   225560   3rd Qu.:3.650e+05
##                                        Max.   :142858700   Max.   :5.000e+09
##
##   Sales Ratio         Property Type      Residential Type
##  Min.   :      0.0   Length:997212      Length:997212
##  1st Qu.:      0.5   Class :character   Class :character
##  Median :      0.6   Mode  :character   Mode  :character
##  Mean   :     10.4
##  3rd Qu.:      0.8
##  Max.   :1226420.0
##
```

```
#Next, I did the same for just the Assessed Value variable.

summary(Real_Estate_Sales_2001_2020_GL_1_$`Assessed Value`)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##         0     87600    138390    278260    225560 142858700
```

```r
#I then did the same for Sale Amount.

summary(Real_Estate_Sales_2001_2020_GL_1_$`Sale Amount`)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.000e+00 1.400e+05 2.250e+05 3.911e+05 3.650e+05 5.000e+09
```

```r
#Next, I got the standard deviations for both variables.

sd(Real_Estate_Sales_2001_2020_GL_1_$`Assessed Value`)
```

```
## [1] 1418520
```

```r
sd(Real_Estate_Sales_2001_2020_GL_1_$`Sale Amount`)
```

```
## [1] 5347271
```

```r
#I then did a summary of the categorical variables of Town and Residence type, showing how many entries

cbind(table(Real_Estate_Sales_2001_2020_GL_1_$'Town'))
```

```
##                  [,1]
## ***Unknown***       1
## Andover           889
## Ansonia          4445
## Ashford          1385
## Avon             7029
## Barkhamsted      1030
## Beacon Falls     1890
## Berlin           5643
## Bethany          1616
## Bethel           5598
## Bethlehem         888
## Bloomfield       5632
## Bolton           1387
## Bozrah            547
## Branford         8714
## Bridgeport      34201
## Bridgewater       555
## Bristol          16915
## Brookfield       5727
## Brooklyn         2455
## Burlington       3154
## Canaan            429
## Canterbury       1369
## Canton           3730
## Chaplin           594
## Cheshire         7707
## Chester          1053
## Clinton          4589
## Colchester       4439
```

```
## Colebrook          519
## Columbia          1636
## Cornwall           614
## Coventry          4107
## Cromwell          5056
## Danbury          20350
## Darien            6545
## Deep River        1254
## Derby             3282
## Durham            1949
## East Granby       1899
## East Haddam       3269
## East Hampton      4870
## East Hartford    13732
## East Haven        9085
## East Lyme         6555
## East Windsor      3772
## Eastford           548
## Easton            2327
## Ellington         4434
## Enfield          10896
## Essex             2633
## Fairfield        15898
## Farmington        8768
## Franklin           567
## Glastonbury      11828
## Goshen            1568
## Granby            3520
## Greenwich        17390
## Griswold          3682
## Groton            9683
## Guilford          6992
## Haddam            2288
## Hamden           16191
## Hampton            591
## Hartford         18810
## Hartland           448
## Harwinton         1475
## Hebron            2708
## Kent              1263
## Killingly         6322
## Killingworth      1936
## Lebanon           2325
## Ledyard           4452
## Lisbon            1160
## Litchfield        2660
## Lyme               882
## Madison           5964
## Manchester       16380
## Mansfield         3998
## Marlborough       1656
## Meriden          17502
## Middlebury        2268
## Middlefield       1096
```

```
## Middletown       12403
## Milford          17749
## Monroe            5984
## Montville         5526
## Morris             835
## Naugatuck         9191
## New Britain      16405
## New Canaan        6275
## New Fairfield     4142
## New Hartford      2283
## New Haven        21346
## New London        6356
## New Milford       8599
## Newington         9091
## Newtown           8188
## Norfolk            658
## North Branford    2873
## North Canaan       950
## North Haven       6561
## North Stonington  1609
## Norwalk          23960
## Norwich          11908
## Old Lyme          3203
## Old Saybrook      4339
## Orange            3455
## Oxford            4008
## Plainfield        4641
## Plainville        5754
## Plymouth          3679
## Pomfret           1163
## Portland          3257
## Preston           1593
## Prospect          2540
## Putnam            3171
## Redding           2537
## Ridgefield        8024
## Rocky Hill        5608
## Roxbury            912
## Salem             1551
## Salisbury         1536
## Scotland           430
## Seymour           4279
## Sharon            1340
## Shelton          11869
## Sherman           1374
## Simsbury          7423
## Somers            2473
## South Windsor     8389
## Southbury         8220
## Southington      11979
## Sprague            763
## Stafford          4053
## Stamford         32529
## Sterling          1421
```

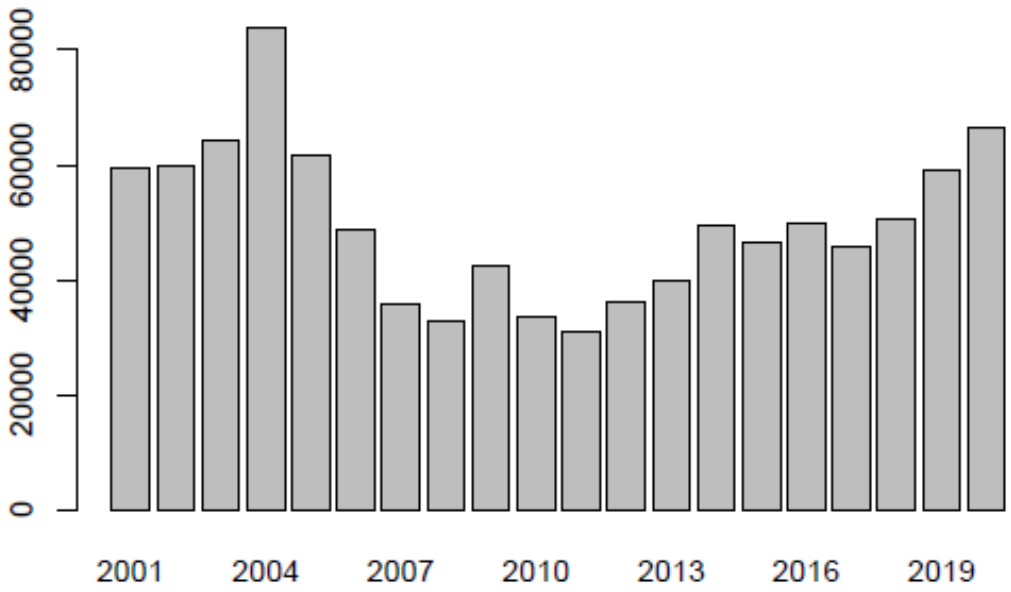
```
## Stonington        6439
## Stratford        16688
## Suffield          4504
## Thomaston         2416
## Thompson          2969
## Tolland           4101
## Torrington       13172
## Trumbull          9785
## Union              261
## Vernon            8662
## Voluntown          734
## Wallingford      11915
## Warren             623
## Washington        1648
## Waterbury        28506
## Waterford         6715
## Watertown         5797
## West Hartford    19854
## West Haven       12308
## Westbrook         2646
## Weston            3268
## Westport          9198
## Wethersfield      7557
## Willington        1262
## Wilton            5387
## Winchester        3723
## Windham           5150
## Windsor           8421
## Windsor Locks     4147
## Wolcott           4635
## Woodbridge        2501
## Woodbury          3407
## Woodstock         3194
```

```
cbind(table(Real_Estate_Sales_2001_2020_GL_1_$`Residential Type`))
```

```
##                  [,1]
## Condo          117779
## Four Family      2418
## Single Family 445016
## Three Family   14081
## Two Family     29609
```

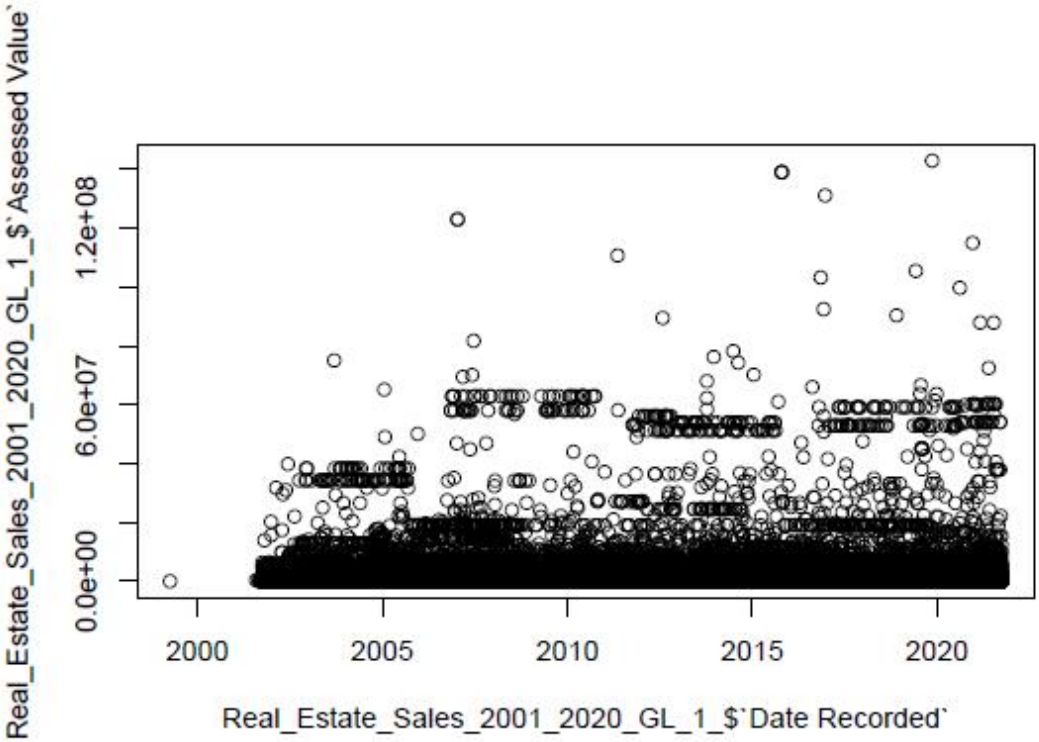
```
#The Following is a bar chart showing the number of listings for each year.

barplot(table(Real_Estate_Sales_2001_2020_GL_1_$`List Year`))
```




```
#Finally, I show a scatterplot of the assessed value of the homes and the dates they were recorded.

plot(Real_Estate_Sales_2001_2020_GL_1_$`Date Recorded`, Real_Estate_Sales_2001_2020_GL_1_$`Assessed Val
```


Real_Estate_Sales_2001_2020_GL_1_$`Assessed Value`
0.0e+00
6.0e+07
1.2e+08
2000
2005
2010
2015
2020
Real_Estate_Sales_2001_2020_GL_1_$`Date Recorded`