

# Data Variable Descriptions Used in Project

Yao Yao

February 23, 2017

```
#Working directory is linked on GitHub.
```

```
# Saved the file as a csv to use read.csv to import the rolling sales for Queens,  
# skipping 4 header lines from the original file.
```

```
qns <- read.csv("rollingsales_queens.csv",skip=4,header=TRUE)
```

BUILDING CLASS CATEGORY - The type of building it is, used to filter family homes GROSS SQUARE FEET - The total square feet of a building YEAR BUILT - The year the building was built SALE PRICE - The price of the building, used to filter buildings not for sale vs for sale LAND SQUARE FEET - The land square feet of where the building resides

```
#Check the data attributes.
```

```
head(qns)
```

```
## BOROUGH NEIGHBORHOOD BUILDING.CLASS.CATEGORY
## 1 4 AIRPORT LA GUARDIA 01 ONE FAMILY DWELLINGS
## 2 4 AIRPORT LA GUARDIA 01 ONE FAMILY DWELLINGS
## 3 4 AIRPORT LA GUARDIA 02 TWO FAMILY DWELLINGS
## 4 4 AIRPORT LA GUARDIA 03 THREE FAMILY DWELLINGS
## 5 4 AIRPORT LA GUARDIA 03 THREE FAMILY DWELLINGS
## 6 4 AIRPORT LA GUARDIA 03 THREE FAMILY DWELLINGS
## TAX.CLASS.AT.PRESENT BLOCK LOT EASE.MENT BUILDING.CLASS.AT.PRESENT
## 1 1 976 61 NA A5
## 2 1 976 63 NA A5
## 3 1 976 70 NA B1
## 4 1 949 15 NA C0
## 5 1 949 56 NA C0
## 6 1 949 59 NA C0
## ADDRESS APARTMENT.NUMBER ZIP.CODE RESIDENTIAL.UNITS
## 1 21-21 80TH STREET 11370 1
## 2 21-17 80TH STREET 11370 1
## 3 21-03 80TH STREET 11370 2
## 4 19-08 81ST STREET 11370 3
## 5 19-69 80TH STREET 11370 3
## 6 19-63 80TH STREET 11370 3
## COMMERCIAL.UNITS TOTAL.UNITS LAND.SQUARE.FEET GROSS.SQUARE.FEET
## 1 - 1 1,800 1,224
## 2 - 1 1,800 1,224
## 3 - 2 1,800 1,224
## 4 - 3 2,112 4,300
## 5 - 3 2,000 2,835
## 6 - 3 2,000 2,835
## YEAR.BUILT TAX.CLASS.AT.TIME.OF.SALE BUILDING.CLASS.AT.TIME.OF.SALE
## 1 1950 1 A5
## 2 1950 1 A5
## 3 1950 1 B1
## 4 1985 1 C0
## 5 1945 1 C0
```

```
## 6      1945      1      C0
##  SALE.PRICE  SALE.DATE
## 1  $660,000  7/26/2016
## 2  $275,500 11/18/2016
## 3    $-      6/13/2016
## 4  $940,000  4/14/2016
## 5    $-      8/15/2016
## 6  $470,000  4/15/2016
```

# summary(qns)

```
##      BOROUGH      NEIGHBORHOOD
## Min.   :4  FLUSHING-NORTH : 2575
## 1st Qu.:4  ASTORIA       : 1165
## Median :4  BAYSIDE       : 1132
## Mean   :4  FOREST HILLS  : 1052
## 3rd Qu.:4  JACKSON HEIGHTS: 993
## Max.   :4  FLUSHING-SOUTH : 854
##      (Other)      :18549
##
##      BUILDING.CLASS.CATEGORY
## 01 ONE FAMILY DWELLINGS      :8357
## 02 TWO FAMILY DWELLINGS      :5681
## 10 COOPS - ELEVATOR APARTMENTS :3867
## 13 CONDOS - ELEVATOR APARTMENTS :1735
## 03 THREE FAMILY DWELLINGS     :1235
## 09 COOPS - WALKUP APARTMENTS  :1226
## (Other)                      :4219
##
## TAX.CLASS.AT.PRESENT  BLOCK      LOT      EASE.MENT
## 1      :15342      Min.   : 13  Min.   : 1.0  Mode:logical
## 2      : 7213      1st Qu.: 2694  1st Qu.: 16.0  NA's:26320
## 4      : 1797      Median : 5938  Median : 39.0
## 2A     : 629      Mean   : 6614  Mean   : 203.7
## 1B     : 429      3rd Qu.:10076  3rd Qu.: 81.0
##      : 373      Max.   :16322  Max.   :8007.0
## (Other): 537
##
## BUILDING.CLASS.AT.PRESENT      ADDRESS      APARTMENT.NUMBER
## A1      : 3870      120 BEACH 26 STREET : 127      :23536
## D4      : 3867      63-14 QUEENS BOULEVARD: 66  2A      : 48
## A5      : 2034      31-35 31ST STREET : 63  2B      : 48
## B3      : 1954      112-45 39TH AVENUE : 60  3B      : 47
## B2      : 1850      131-05 40TH ROAD : 55  3A      : 45
## A2      : 1593      42-60 CRESCENT STREET : 54  4A      : 35
## (Other):11152      (Other)      :25895  (Other): 2561
##
##      ZIP.CODE      RESIDENTIAL.UNITS  COMMERCIAL.UNITS  TOTAL.UNITS
## Min.   : 0      1      :5673      0      :12815      1      :6080
## 1st Qu.:11360      0      :5154      -      :12104      1      :5597
## Median :11375      1      :4959      1      : 562      0      :4264
## Mean   :11261      2      :3030      1      : 489      2      :2980
## 3rd Qu.:11419      -      :2703      2      : 88      2      :2669
## Max.   :11697      2      :2699      2      : 78      -      :2144
##      (Other):2102      (Other): 184      (Other):2586
##
## LAND.SQUARE.FEET  GROSS.SQUARE.FEET  YEAR.BUILT
## 0      : 5754      0      : 6033      Min.   : 0
## -      : 2877      -      : 3330      1st Qu.:1925
## 4,000 : 1217      1,600 : 109      Median :1940
```

```

##   2,500 : 822      1224 : 103      Mean   :1825
##   2,000 : 708      1,440 : 82      3rd Qu.:1959
##  4000   : 687      1,224 : 76      Max.    :2016
## (Other):14255      (Other):16587
## TAX.CLASS.AT.TIME.OF.SALE BUILDING.CLASS.AT.TIME.OF.SALE
## Min.    :1.000          D4      : 3867
## 1st Qu.:1.000          A1      : 3861
## Median :1.000          A5      : 2032
## Mean    :1.529          B3      : 1972
## 3rd Qu.:2.000          B2      : 1873
## Max.    :4.000          R4      : 1735
##                               (Other):10980
##      SALE.PRICE      SALE.DATE
## $-      : 8226      4/5/2016 : 210
## $10     : 209      11/10/2016: 177
## $450,000 : 156      6/30/2016 : 174
## $650,000 : 150      2/29/2016 : 170
## $250,000 : 137      11/22/2016: 161
## $600,000 : 137      10/28/2016: 158
## (Other)  :17305      (Other)   :25270

```

```
str(qns)
```

```

## 'data.frame': 26320 obs. of 21 variables:
## $ BOROUGH : int 4 4 4 4 4 4 4 4 4 4 ...
## $ NEIGHBORHOOD : Factor w/ 60 levels "AIRPORT LA GUARDIA",...: 1 1 1 1 1 1 1 1 2 2 ...
## $ BUILDING.CLASS.CATEGORY : Factor w/ 44 levels "01 ONE FAMILY DWELLINGS": 1 1 1 1 1 1 1 1 1 1 ...
## $ TAX.CLASS.AT.PRESENT : Factor w/ 11 levels " ", "1", "1A", "1B",...: 2 2 2 2 2 2 6 6 2 2 ...
## $ BLOCK : int 976 976 976 949 949 949 949 949 15828 15829 ...
## $ LOT : int 61 63 70 15 56 59 1012 1025 53 22 ...
## $ EASE.MENT : logi NA NA NA NA NA NA ...
## $ BUILDING.CLASS.AT.PRESENT : Factor w/ 125 levels " ", "A0", "A1",...: 7 7 11 15 15 15 91 91 3 3 ...
## $ ADDRESS : Factor w/ 23093 levels "-00 136TH AVENUE",...: 9305 9287 9243 84 ...
## $ APARTMENT.NUMBER : Factor w/ 1193 levels " ", "0.02", "1",...: 1 1 1 1 1 1 225 3 1 1 ...
## $ ZIP.CODE : int 11370 11370 11370 11370 11370 11370 11370 11370 11691 11691 ...
## $ RESIDENTIAL.UNITS : Factor w/ 111 levels " - ", "1",...: 2 2 14 22 22 22 2 2 2 2 ...
## $ COMMERCIAL.UNITS : Factor w/ 36 levels " - ", "1",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ TOTAL.UNITS : Factor w/ 120 levels " - ", "1",...: 2 2 15 22 22 22 2 2 2 2 ...
## $ LAND.SQUARE.FEET : Factor w/ 4202 levels " - ", "1,000",...: 265 265 265 553 491 4 ...
## $ GROSS.SQUARE.FEET : Factor w/ 4200 levels " - ", "1,000",...: 178 178 178 1781 1348 ...
## $ YEAR.BUILT : int 1950 1950 1950 1985 1945 1945 0 0 2002 2005 ...
## $ TAX.CLASS.AT.TIME.OF.SALE : int 1 1 1 1 1 1 2 2 1 1 ...
## $ BUILDING.CLASS.AT.TIME.OF.SALE: Factor w/ 124 levels "A0", "A1", "A2",...: 6 6 10 14 14 14 90 90 2 2 ...
## $ SALE.PRICE : Factor w/ 3272 levels "$- ", "$1",...: 2590 1188 1 3178 1 2018 ...
## $ SALE.DATE : Factor w/ 359 levels "1/1/2017", "1/10/2017",...: 288 71 245 187 30 ...

```