### Project

#### Amritha

### 24/11/2022

housing=read.csv("C:\\Users\\Amritha\\Desktop\\Project\_files\\dataset\\housing.csv")

#### head(housing)

```
##
             id
## 1 7049658318
## 2 7046392406
## 3 7037810876
## 4 7035234026
## 5 7036718653
## 6 7048921428
                                                                                                    url
## 1
          https://denver.craigslist.org/apa/d/arvada-washer-dryer-hookups-full-size/7049658318.html
## 2
          https://lansing.craigslist.org/apa/d/okemos-pet-friendly-community-online/7046392406.html
       https://annarbor.craigslist.org/apa/d/ann-arbor-2-bedrooms-townhome-for-rent/7037810876.html
## 3
## 4
               https://saginaw.craigslist.org/apa/d/fitness-center-community-garden/7035234026.html
## 5 https://gainesville.craigslist.org/apa/d/gainesville-3133-sw-42nd-place-3-br-2/7036718653.html
##
        https://treasure.craigslist.org/apa/d/vero-beach-family-friendly-spacious-3/7048921428.html
##
                      region
                                                       region_url price
## 1
                       denver
                                   https://denver.craigslist.org 1739 townhouse
## 2
                      lansing
                                  https://lansing.craigslist.org
                                                                   1055 apartment
## 3
                                 https://annarbor.craigslist.org
                   ann arbor
                                                                   1945
                                                                             condo
## 4 saginaw-midland-baycity
                                  https://saginaw.craigslist.org
                                                                    830 apartment
## 5
                 gainesville https://gainesville.craigslist.org
                                                                   1100
                                                                            duplex
                                 https://treasure.craigslist.org
              treasure coast
                                                                    883 apartment
##
     sqfeet beds baths cats_allowed dogs_allowed smoking_allowed wheelchair_access
## 1
       1166
               3
                   1.5
                                   1
                                                 1
                                                                 1
## 2
        972
               2
                   1.5
                                   1
                                                 1
                                                                                    0
## 3
       1310
               2
                   2.5
                                   1
                                                 1
                                                                 0
                                                                                    0
               2
## 4
       1140
                   1.5
                                                                                    0
                                   1
                                                 1
                                                                  1
## 5
       1270
                   2.0
                                   0
                                                                                    0
##
       1200
                   2.0
                                   1
##
     electric_vehicle_charge comes_furnished laundry_options
                                                                  parking_options
## 1
                            0
                                                   w/d in unit
## 2
                            0
                                            0
                                                   w/d hookups
## 3
                            0
                                            0
                                                   w/d in unit
                                                                   attached garage
## 4
                            0
                                            0 laundry in bldg
                                                                           carport
## 5
                            0
                                                   w/d hookups off-street parking
## 6
                                                   w/d hookups
##
                                                         image_url
```

## 1 https://images.craigslist.org/00Q0Q\_5m9ea2vDyiJ\_600x450.jpg

```
## 2 https://images.craigslist.org/00i0i_fg5z2eefE9V_600x450.jpg
## 3 https://images.craigslist.org/01010_b2dicCKeR4m_600x450.jpg
## 4 https://images.craigslist.org/00c0c_95vUgcdIXQm_600x450.jpg
## 5 https://images.craigslist.org/00F0F_c4K9sSHJDNa_600x450.jpg
## 6 https://images.craigslist.org/00P0P_1FE6EvPnc1B_600x450.jpg
##
## 2 OKEMOS APARTMENTS & TOWNHOMES Meridian Meadows in Okemos, Michigan, offers convenient living in a
## 4
## 5
## 6
##
         lat
                  long state
## 1 39.8471 -105.0530
              -84.3994
## 2 42.7137
                          mi
## 3 42.2794
              -83.7840
                          mi
## 4 43.4141
             -84.0206
                          mi
## 5 29.6132
             -82.3873
                          fl
## 6 27.6165 -80.3831
                          f1
summary(housing)
##
          id
                                                                 region_url
                             url
                                               region
           :7.005e+09
##
  Min.
                        Length: 50000
                                            Length: 50000
                                                                Length: 50000
   1st Qu.:7.036e+09
                         Class : character
                                            Class : character
                                                                Class : character
## Median :7.043e+09
                         Mode :character
                                            Mode :character
                                                                Mode : character
  Mean
           :7.041e+09
##
    3rd Qu.:7.048e+09
##
    Max.
           :7.051e+09
##
##
        price
                                                sqfeet
                                                                   beds
                             type
           :0.000e+00
                        Length: 50000
                                                          0
                                                                     :0.000
##
                                            Min.
                                                              Min.
                                            1st Qu.:
##
    1st Qu.:8.150e+02
                         Class : character
                                                        750
                                                              1st Qu.:1.000
   Median :1.060e+03
                         Mode : character
                                            Median:
                                                        950
                                                              Median :2.000
           :5.674e+04
##
  Mean
                                                      1025
                                                              Mean
                                                                     :1.895
                                            Mean
    3rd Qu.:1.450e+03
                                            3rd Qu.:
                                                      1155
                                                              3rd Qu.:2.000
##
    Max.
           :2.768e+09
                                            Max.
                                                   :999999
                                                              Max.
                                                                     :8.000
##
##
        baths
                     cats_allowed
                                       dogs_allowed
                                                        smoking_allowed
##
    Min.
           :0.000
                    Min.
                           :0.0000
                                      Min.
                                             :0.0000
                                                       Min.
                                                              :0.000
                    1st Qu.:0.0000
    1st Qu.:1.000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.000
  Median :1.000
                    Median :1.0000
                                      Median :1.0000
                                                       Median :1.000
    Mean
           :1.482
                            :0.7132
##
                    Mean
                                      Mean
                                             :0.6928
                                                       Mean
                                                              :0.734
##
    3rd Qu.:2.000
                    3rd Qu.:1.0000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.000
                           :1.0000
##
    Max.
          :6.500
                    Max.
                                      Max.
                                             :1.0000
                                                       Max.
                                                              :1.000
##
```

Min.

Mean

:0.00000

:0.05066

:1.00000

1st Qu.:0.00000

Median :0.00000

3rd Qu.:0.00000

laundry\_options

Class : character

Mode :character

Length: 50000

wheelchair\_access electric\_vehicle\_charge comes\_furnished

1st Qu.:0.00000

Median :0.00000

3rd Qu.:0.00000

:0.00000

:0.01392

:1.00000

##

##

##

##

Min.

## Mean

:0.00000

:0.07844

:1.00000

1st Qu.:0.00000

3rd Qu.:0.00000

## Median :0.00000

Min.

Mean

Max.

```
parking_options
                       image_url
                                          description
                                                                  lat
## Length:50000
                       Length:50000
                                          Length:50000
                                                             Min. : 2.202
  Class :character
                                          Class : character
                       Class : character
                                                             1st Qu.: 33.514
  Mode :character
                      Mode :character
                                          Mode :character
                                                             Median: 37.976
##
##
                                                             Mean : 37.200
##
                                                             3rd Qu.: 41.180
##
                                                             Max.
                                                                   :102.036
                                                             NA's
##
                                                                    :249
##
         long
                         state
           :-163.89
                      Length:50000
##
   1st Qu.:-104.70
                      Class :character
## Median : -86.47
                      Mode :character
## Mean
          : -92.43
## 3rd Qu.: -81.25
## Max.
          : 94.16
## NA's
           :249
nrow(housing)
## [1] 50000
ncol(housing)
## [1] 22
dim(housing)
## [1] 50000
                22
length(housing)
## [1] 22
colnames(housing)
   [1] "id"
                                  "url"
##
    [3] "region"
                                  "region_url"
  [5] "price"
##
                                  "type"
  [7] "sqfeet"
                                  "beds"
  [9] "baths"
                                  "cats allowed"
##
## [11] "dogs_allowed"
                                  "smoking_allowed"
## [13] "wheelchair_access"
                                  "electric_vehicle_charge"
## [15] "comes_furnished"
                                  "laundry_options"
## [17] "parking_options"
                                  "image_url"
## [19] "description"
                                  "lat"
## [21] "long"
                                  "state"
\#\#unimportant column
```

```
summary(housing["id"])
##
          id
           :7.005e+09
## Min.
## 1st Qu.:7.036e+09
## Median :7.043e+09
         :7.041e+09
## Mean
## 3rd Qu.:7.048e+09
## Max. :7.051e+09
nrow(unique(housing["id"]))
## [1] 50000
\#\#unimportant column
summary(housing["url"])
##
        url
## Length:50000
## Class :character
## Mode :character
sum(is.na(housing["url"]))
## [1] 0
nrow(unique(housing["id"]))
## [1] 50000
\#\#important column–categorical data
summary(housing["region"])
##
       region
## Length:50000
## Class :character
## Mode :character
sum(is.na(housing["region"]))
## [1] 0
nrow(unique(housing["region"]))
## [1] 291
```

```
summary(housing["region_url"])
##
     region_url
##
   Length: 50000
   Class : character
   Mode :character
sum(is.na(housing["region_url"]))
## [1] 0
nrow(unique(housing["region_url"]))
## [1] 299
regi=unique(housing[3:4])
nrow(regi)
## [1] 300
d <- function(x) table(x)[x] > 1
onetomany=regi[d(regi$region) & !d(regi$region_url),]
nrow(onetomany)
## [1] 18
d1=onetomany[order(onetomany$region),]
##
              region
                                                region_url
## 34
                          https://albanyga.craigslist.org
              albany
## 243
                            https://albany.craigslist.org
              albany
                          https://athensga.craigslist.org
## 133
              athens
## 4110
              athens
                        https://athensohio.craigslist.org
                        https://columbusga.craigslist.org
## 44
            columbus
## 421
            columbus
                          https://columbus.craigslist.org
## 36
        fayetteville
                             https://fayar.craigslist.org
## 62
        fayetteville
                      https://fayetteville.craigslist.org
## 106
             jackson
                           https://jackson.craigslist.org
## 1581
                               https://jxn.craigslist.org
             jackson
## 93
        jacksonville
                            https://onslow.craigslist.org
## 276
       jacksonville
                      https://jacksonville.craigslist.org
## 4116
              monroe
                          https://monroemi.craigslist.org
## 4718
                            https://monroe.craigslist.org
              monroe
## 160
                         https://rochester.craigslist.org
           rochester
## 383
           rochester
                               https://rmn.craigslist.org
## 957
         springfield
                       https://springfield.craigslist.org
         springfield https://springfieldil.craigslist.org
## 1856
```

```
h1=housing
h1$region[h1$region_url=="https://albanyga.craigslist.org"]="albany, GA"
h1$region[h1$region url=="https://albany.craigslist.org"]="albany, NY"
h1$region[h1$region url=="https://athensga.craigslist.org"]="athens, GA"
h1$region[h1$region url=="https://athensohio.craigslist.org"]="athens, OH"
h1$region[h1$region_url=="https://columbusga.craigslist.org"]="columbus, GA"
h1$region[h1$region url=="https://columbus.craigslist.org"]="columbus, OH"
h1$region[h1$region_url=="https://fayar.craigslist.org"]="fayetteville, AR"
h1$region[h1$region_url=="https://fayetteville.craigslist.org"]="fayetteville, NC"
h1$region[h1$region_url=="https://jxn.craigslist.org"]="jackson, MI"
h1$region[h1$region_url=="https://jackson.craigslist.org"]="jackson, MS"
h1$region[h1$region_url=="https://jacksonville.craigslist.org"]="jacksonville, FL"
h1$region[h1$region url=="https://onslow.craigslist.org"]="jacksonville, NC"
h1$region[h1$region_url=="https://monroe.craigslist.org"]="monroe, LA"
h1$region[h1$region url=="https://onslow.craigslist.org"]="monroe, MI"
h1$region[h1$region_url=="https://rmn.craigslist.org"]="rochester, MN"
h1$region[h1$region_url=="https://onslow.craigslist.org"]="rochester, NY"
h1$region[h1$region url=="https://springfieldil.craigslist.org"]="springfield, IL"
h1$region[h1$region_url=="https://onslow.craigslist.org"]="springfield, MO"
housing=h1
h1=NULL
regi=unique(housing[3:4])
d <- function(x) table(x)[x] > 1
manytoone=regi[!d(regi$region) & d(regi$region_url),]
summary(manytoone)
##
      region
                       region_url
## Length:2
                       Length:2
## Class :character
                       Class : character
                      Mode :character
## Mode :character
d2=manytoone[order(manytoone$region),]
d2
                                         region_url
            st louis https://stlouis.craigslist.org
## 6595
## 9
       st louis, MO https://stlouis.craigslist.org
h1=housing
h1$region[h1$region_url=="https://kansascity.craigslist.org"]="kansas city, MO"
h1$region[h1$region_url=="https://stlouis.craigslist.org"]=" st louis, MO"
```

```
library(stringr)
r2=h1[str_detect(h1$region, "kansas"),]
r2[sample(nrow(r2), 10,replace=TRUE), ]
##
## 46720 7037842864
## 10538 7035189185
## 44166 7051215838
## 48059 7043082823
         7049762991
## 331
## 19431 7051037580
## 10687 7049607326
## 6148 7049722080
## 31942 7050857783
## 5923 7049716427
##
## 46720
             https://kansascity.craigslist.org/apa/d/kansas-city-new-homes-have-arrived-3/7037842864.htm
## 10538 https://kansascity.craigslist.org/apa/d/kansas-city-as-low-as-650-month-500-down/7035189185.ht
## 44166
                https://kansascity.craigslist.org/apa/d/overland-park-the-perfect-pair-79/7051215838.htm
## 48059
               https://kansascity.craigslist.org/apa/d/kansas-city-3-bedroom-possible-4th/7043082823.htm
## 331
             https://kansascity.craigslist.org/apa/d/kansas-city-welcome-to-your-new-home/7049762991.html
## 19431
             https://kansascity.craigslist.org/apa/d/leawood-proper-design-smartly-priced/7051037580.htm
## 10687
           https://kansascity.craigslist.org/apa/d/overland-park-49-deposit-st-month-move/7049607326.html
## 6148
           https://kansascity.craigslist.org/apa/d/kansas-city-just-like-that-youve-found/7049722080.htm
            https://kansascity.craigslist.org/apa/d/kansas-city-updated-studios-ready-for/7050857783.html
## 31942
## 5923
                 https://kansascity.craigslist.org/apa/d/kansas-city-stunning-two-bedroom/7049716427.htm
##
                                                  region url price
                  region
## 46720 kansas city, MO https://kansascity.craigslist.org
                                                               924 manufactured
## 10538 kansas city, MO https://kansascity.craigslist.org
                                                               500
                                                                           house
## 44166 kansas city, MO https://kansascity.craigslist.org
                                                               739
                                                                       apartment
## 48059 kansas city, MO https://kansascity.craigslist.org
                                                               750
                                                                           house
         kansas city, MO https://kansascity.craigslist.org
                                                               600
                                                                       apartment
## 19431 kansas city, MO https://kansascity.craigslist.org
                                                               739
                                                                       apartment
## 10687 kansas city, MO https://kansascity.craigslist.org
                                                               739
                                                                       apartment
## 6148 kansas city, MO https://kansascity.craigslist.org
                                                               600
                                                                       apartment
## 31942 kansas city, MO https://kansascity.craigslist.org
                                                               600
                                                                       apartment
         kansas city, MO https://kansascity.craigslist.org
                                                                       apartment
##
         sqfeet beds baths cats_allowed dogs_allowed smoking_allowed
## 46720
           1216
                   3
                          2
                                       1
                                                     1
                                                                      1
## 10538
           1000
                   3
                          1
                                       1
                                                     1
                                                                      1
## 44166
            701
                   1
                          1
                                       1
                                                     1
                                                                      1
## 48059
           1232
                   3
                                       0
                                                     0
                          1
                                                                      1
## 331
            480
                   0
                          1
                                                     1
                                                                      1
                                       1
            701
## 19431
                   1
                          1
                                       1
                                                     1
                                                                      1
## 10687
            641
                          1
                                       1
                                                     1
                                                                      1
## 6148
            480
                   0
                          1
                                       1
                                                     1
                                                                      1
## 31942
            480
                   0
                          1
                                       1
                                                     1
                                                                      1
                   2
                          2
## 5923
           1087
                                       1
                                                     1
##
         wheelchair_access electric_vehicle_charge comes_furnished laundry_options
## 46720
                                                                          w/d hookups
                          0
## 10538
                                                   0
                                                                          w/d hookups
## 44166
                          0
                                                   0
                                                                   O laundry in bldg
## 48059
                          0
                                                   0
```

```
## 331
                                                  0
                                                                  0 laundry on site
## 19431
                         0
                                                                  0 laundry in bldg
                                                  0
## 10687
                         0
                                                  0
                                                                  0 laundry in bldg
                         0
## 6148
                                                  0
                                                                  0 laundry on site
## 31942
                         0
                                                  0
                                                                  0 laundry on site
## 5923
                         0
                                                                        w/d in unit
                                                  0
            parking_options
## 46720 off-street parking
## 10538
            attached garage
## 44166 off-street parking
## 48059 off-street parking
## 331
       off-street parking
## 19431 off-street parking
## 10687 off-street parking
## 6148 off-street parking
## 31942 off-street parking
## 5923
##
                                                            image_url
## 46720 https://images.craigslist.org/00X0X_bosZgEqin9P_600x450.jpg
## 10538 https://images.craigslist.org/00J0J_4mxSVBaS3hM_600x450.jpg
## 44166 https://images.craigslist.org/00i0i_euNShZGzvWd_600x450.jpg
## 48059 https://images.craigslist.org/00Y0Y_4cg6SWVTVsv_600x450.jpg
         https://images.craigslist.org/00000_jBWCGojmrRx_600x450.jpg
## 19431 https://images.craigslist.org/00v0v_i9aIP1TVJ1k_600x450.jpg
## 10687 https://images.craigslist.org/00I0I_3Iiiaxx3dnC_600x450.jpg
## 6148 https://images.craigslist.org/00x0x_7U9K1oy0H2q_600x450.jpg
## 31942 https://images.craigslist.org/00B0B_idpZxZJV7jI_600x450.jpg
## 5923 https://images.craigslist.org/01616_e0sBgM3B3DL_600x450.jpg
##
## 46720
## 10538
## 44166
## 48059 two story 3 bedroom house possible 4th bedroom bonus room.1 bathroom up stairs plenty of park
## 19431
## 10687
## 6148
## 31942
## 5923
##
                     long state
             lat
## 46720 39.0803 -94.7806
## 10538 38.9249 -94.4983
## 44166 38.9928 -94.6771
## 48059 39.1042 -94.4923
         39.2133 -94.5743
                             ks
## 19431 38.9667 -94.6169
                             ks
## 10687 38.9928 -94.6771
                             ks
## 6148 39.2133 -94.5743
## 31942 39.2133 -94.5743
                             ks
## 5923 39.2627 -94.6970
housing=h1
h1=NULL
summary(housing)
```

```
##
          id
                             url
                                                region
                                                                 region_url
##
           :7.005e+09
                         Length:50000
                                            Length:50000
                                                                 Length:50000
    Min.
    1st Qu.:7.036e+09
                                             Class : character
                         Class : character
                                                                 Class : character
   Median :7.043e+09
                                                                Mode : character
##
                         Mode : character
                                            Mode :character
    Mean
           :7.041e+09
##
    3rd Qu.:7.048e+09
##
    Max.
           :7.051e+09
##
        price
##
                             type
                                                 sqfeet
                                                                    beds
##
           :0.000e+00
                         Length: 50000
                                                                      :0.000
    Min.
                                            Min.
                                                   :
                                                          0
                                                              Min.
    1st Qu.:8.150e+02
                         Class : character
                                             1st Qu.:
                                                        750
                                                              1st Qu.:1.000
    Median :1.060e+03
                                                        950
##
                         Mode :character
                                            Median :
                                                              Median :2.000
           :5.674e+04
                                                      1025
##
    Mean
                                             Mean
                                                              Mean
                                                                      :1.895
                                             3rd Qu.:
##
    3rd Qu.:1.450e+03
                                                       1155
                                                              3rd Qu.:2.000
##
    Max.
           :2.768e+09
                                            Max.
                                                    :999999
                                                              Max.
                                                                      :8.000
##
##
        baths
                      cats_allowed
                                       dogs_allowed
                                                        smoking_allowed
##
    Min.
           :0.000
                    Min.
                           :0.0000
                                      Min.
                                             :0.0000
                                                        Min.
                                                              :0.000
    1st Qu.:1.000
                    1st Qu.:0.0000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.000
##
                                                        Median :1.000
##
    Median :1.000
                    Median :1.0000
                                      Median :1.0000
##
    Mean
           :1.482
                    Mean
                            :0.7132
                                      Mean
                                              :0.6928
                                                        Mean
                                                               :0.734
##
    3rd Qu.:2.000
                    3rd Qu.:1.0000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.000
##
    Max.
           :6.500
                            :1.0000
                                      Max.
                                              :1.0000
                                                        Max.
                                                               :1.000
                    Max.
##
##
    wheelchair_access electric_vehicle_charge comes_furnished
                                                                  laundry_options
           :0.00000
                      Min.
                              :0.00000
                                                Min.
                                                       :0.00000
                                                                  Length: 50000
##
    1st Qu.:0.00000
                      1st Qu.:0.00000
                                                1st Qu.:0.00000
                                                                  Class :character
    Median :0.00000
                      Median :0.00000
                                                Median :0.00000
                                                                  Mode :character
##
    Mean
           :0.07844
                      Mean
                              :0.01392
                                                Mean
                                                       :0.05066
                      3rd Qu.:0.00000
    3rd Qu.:0.00000
                                                3rd Qu.:0.00000
##
    Max.
           :1.00000
                      Max.
                              :1.00000
                                                Max.
                                                       :1.00000
##
##
    parking_options
                         image_url
                                           description
                                                                     lat
   Length:50000
                        Length: 50000
                                           Length:50000
                                                                      : 2.202
##
                                                               Min.
                                                               1st Qu.: 33.514
##
    Class :character
                        Class :character
                                           Class :character
##
    Mode :character
                       Mode :character
                                           Mode :character
                                                               Median: 37.976
##
                                                               Mean
                                                                     : 37.200
##
                                                               3rd Qu.: 41.180
##
                                                               Max.
                                                                       :102.036
                                                               NA's
##
                                                                       :249
##
         long
                          state
##
           :-163.89
                      Length: 50000
   Min.
    1st Qu.:-104.70
                      Class : character
##
   Median : -86.47
                      Mode :character
           : -92.43
    Mean
    3rd Qu.: -81.25
##
           : 94.16
##
    Max.
##
    NA's
           :249
sum(is.na(housing$lat))
```

## [1] 249

```
h1=na.omit(housing)
dim(h1)
```

#### ## [1] 49751 22

housing=h1
summary(housing)

```
##
          id
                             url
                                                region
                                                                 region_url
           :7.005e+09
##
    Min.
                         Length: 49751
                                            Length: 49751
                                                                Length: 49751
    1st Qu.:7.036e+09
                         Class : character
                                             Class : character
                                                                Class : character
##
    Median :7.043e+09
                         Mode : character
                                            Mode :character
                                                                Mode : character
    Mean
           :7.041e+09
    3rd Qu.:7.048e+09
##
    Max.
           :7.051e+09
##
##
        price
                                                 sqfeet
                                                                    beds
                             type
           :0.000e+00
                         Length: 49751
                                                                      :0.000
##
    Min.
                                            Min.
                                                  :
                                                          0
                                                              Min.
                                                              1st Qu.:1.000
    1st Qu.:8.150e+02
                         Class : character
                                             1st Qu.:
                                                        750
##
   Median :1.060e+03
##
                         Mode :character
                                             Median:
                                                        950
                                                              Median :2.000
    Mean
           :5.701e+04
                                                       1024
                                                              Mean
##
                                            Mean
                                                                     :1.893
    3rd Qu.:1.450e+03
                                                              3rd Qu.:2.000
##
                                             3rd Qu.:
                                                       1154
    Max.
           :2.768e+09
                                            Max.
                                                    :999999
                                                              Max.
                                                                      :8.000
##
##
        baths
                     cats allowed
                                       dogs allowed
                                                        smoking allowed
                           :0.0000
##
    Min.
           :0.000
                    Min.
                                      Min.
                                             :0.0000
                                                        Min.
                                                               :0.0000
    1st Qu.:1.000
                    1st Qu.:0.0000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.0000
##
                    Median :1.0000
    Median :1.000
                                      Median :1.0000
                                                        Median :1.0000
##
    Mean
          :1.481
                    Mean
                            :0.7135
                                      Mean
                                              :0.6928
                                                        Mean
                                                               :0.7344
    3rd Qu.:2.000
                    3rd Qu.:1.0000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
   Max.
           :6.500
                    Max.
                            :1.0000
                                      Max.
                                              :1.0000
                                                        Max.
                                                               :1.0000
##
##
    wheelchair_access electric_vehicle_charge comes_furnished
                                                                  laundry_options
##
    Min.
           :0.00000
                      Min.
                              :0.00000
                                                Min.
                                                       :0.00000
                                                                  Length: 49751
    1st Qu.:0.00000
                      1st Qu.:0.00000
                                                1st Qu.:0.00000
                                                                  Class : character
   Median :0.00000
                      Median :0.00000
                                                Median :0.00000
                                                                  Mode :character
##
           :0.07867
                                                       :0.05011
    Mean
                      Mean
                              :0.01393
                                                Mean
    3rd Qu.:0.00000
##
                      3rd Qu.:0.00000
                                                3rd Qu.:0.00000
##
    Max.
           :1.00000
                      Max.
                              :1.00000
                                                Max.
                                                       :1.00000
##
    parking_options
                         image_url
                                           description
                                                                     lat
##
    Length: 49751
                        Length: 49751
                                           Length: 49751
                                                               Min.
                                                                       : 2.202
##
    Class : character
                        Class : character
                                           Class : character
                                                               1st Qu.: 33.514
##
    Mode :character
                        Mode :character
                                           Mode :character
                                                               Median: 37.976
                                                                     : 37.200
##
                                                               Mean
##
                                                               3rd Qu.: 41.180
##
                                                               Max.
                                                                      :102.036
##
                          state
         long
           :-163.89
                      Length: 49751
##
    Min.
##
    1st Qu.:-104.70
                      Class : character
    Median : -86.47
                      Mode :character
##
    Mean
          : -92.43
    3rd Qu.: -81.25
    Max.
          : 94.16
```

#### h1=NULL

```
housing$type=as.factor(housing$type)
housing$region=as.factor(housing$region)
housing$laundry_options=as.factor(housing$laundry_options)
housing$parking_options=as.factor(housing$parking_options)
housing$state=as.factor(housing$state)
summary(housing)
```

```
##
          iд
                             url
                                                                   region
##
    Min.
           :7.005e+09
                         Length: 49751
                                              omaha / council bluffs :
##
    1st Qu.:7.036e+09
                         Class : character
                                              inland empire
                                                                         508
    Median :7.043e+09
                         Mode :character
                                              denver
                                                                         503
                                                                         496
##
    Mean
           :7.041e+09
                                              reno / tahoe
    3rd Qu.:7.048e+09
                                              fort collins / north CO:
                                                                         493
##
    Max.
           :7.051e+09
                                              minneapolis / st paul :
                                                                         488
##
                                              (Other)
                                                                      :46734
##
     region_url
                            price
                                                                        sqfeet
                                                        type
##
    Length: 49751
                                :0.000e+00
                                              apartment
                                                          :41020
                        Min.
                                                                    Min.
##
    Class : character
                        1st Qu.:8.150e+02
                                              house
                                                           : 4370
                                                                    1st Qu.:
                                                                                750
    Mode : character
                        Median :1.060e+03
                                              townhouse
                                                          : 1849
                                                                    Median:
                                                                                950
##
                        Mean
                                :5.701e+04
                                              condo
                                                              921
                                                                    Mean
                                                                               1024
##
                        3rd Qu.:1.450e+03
                                              duplex
                                                              683
                                                                    3rd Qu.:
                                                                               1154
##
                                :2.768e+09
                        Max.
                                              manufactured:
                                                             577
                                                                    Max.
                                                                            :999999
##
                                              (Other)
                                                             331
                                                          :
##
         beds
                         baths
                                       cats allowed
                                                         dogs allowed
##
    Min.
           :0.000
                     Min.
                            :0.000
                                      Min.
                                              :0.0000
                                                        Min.
                                                                :0.0000
##
    1st Qu.:1.000
                     1st Qu.:1.000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.0000
    Median :2.000
                     Median :1.000
                                      Median :1.0000
                                                        Median :1.0000
##
##
    Mean
           :1.893
                     Mean
                            :1.481
                                      Mean
                                              :0.7135
                                                        Mean
                                                                :0.6928
                     3rd Qu.:2.000
    3rd Qu.:2.000
                                                        3rd Qu.:1.0000
##
                                      3rd Qu.:1.0000
##
    Max.
           :8.000
                            :6.500
                                      Max.
                                              :1.0000
                                                        Max.
                                                                :1.0000
                     Max.
##
    smoking_allowed
                      wheelchair_access electric_vehicle_charge comes_furnished
##
    Min.
           :0.0000
                      Min.
                              :0.00000
                                         Min.
                                                 :0.00000
                                                                   Min.
                                                                           :0.00000
                      1st Qu.:0.00000
##
    1st Qu.:0.0000
                                         1st Qu.:0.00000
                                                                   1st Qu.:0.00000
##
    Median :1.0000
                      Median :0.00000
                                         Median :0.00000
                                                                   Median :0.00000
    Mean
          :0.7344
                      Mean
                              :0.07867
                                         Mean
                                                 :0.01393
                                                                   Mean
                                                                           :0.05011
##
    3rd Qu.:1.0000
                      3rd Qu.:0.00000
                                         3rd Qu.:0.00000
                                                                   3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                              :1.00000
                                         Max.
                                                 :1.00000
                                                                   Max.
                                                                           :1.00000
##
##
               laundry options
                                           parking_options
                                                               image_url
##
                       :10316
                                                    :17805
                                                              Length: 49751
                                                              Class : character
##
    laundry in bldg
                       : 5166
                                 off-street parking:16669
    laundry on site
                       : 7410
                                                    : 5413
                                                              Mode : character
                                 carport
    no laundry on site:
                          462
                                 attached garage
                                                    : 5071
##
    w/d hookups
                       : 9397
                                 detached garage
                                                    : 2420
                       :17000
##
    w/d in unit
                                 street parking
                                                    : 1987
##
                                 (Other)
                                                       386
##
    description
                             lat
                                                 long
                                                                   state
##
    Length: 49751
                        Min.
                                 2.202
                                           Min.
                                                   :-163.89
                                                                      : 6324
                                                               ca
                                                                      : 5975
    Class : character
                        1st Qu.: 33.514
                                           1st Qu.:-104.70
                                                               fl
    Mode :character
                        Median: 37.976
                                           Median : -86.47
                                                                      : 3486
                                                               nc
```

```
##
                                : 37.200
                                           Mean
                                                   : -92.43
                                                                      : 2705
                                                              mi
                        3rd Qu.: 41.180
##
                                           3rd Qu.: -81.25
                                                                      : 2594
                                                              ga
##
                        Max.
                               :102.036
                                           Max.
                                                     94.16
                                                              СО
                                                                      : 2088
##
                                                               (Other):26579
housing=subset(housing,beds!=0)
housing=subset(housing,baths!=0)
housing=subset(housing,price!=0)
housing=subset(housing,sqfeet!=0)
housing$baths=as.integer(housing$baths)
housing=subset(housing,type!='assisted living')
summary(housing)
##
          id
                             url
                                                                   region
           :7.005e+09
                         Length: 47812
##
    Min.
                                              omaha / council bluffs :
                                                                         509
    1st Qu.:7.036e+09
##
                         Class : character
                                              inland empire
                                                                         496
    Median :7.043e+09
                         Mode :character
                                                                         484
##
                                             denver
##
    Mean
           :7.041e+09
                                                                         479
                                             ventura county
    3rd Qu.:7.048e+09
                                             orlando
##
    Max.
           :7.051e+09
                                             fort collins / north CO:
                                                                         471
                                                                      :44897
##
                                              (Other)
##
                            price
     region_url
                                                        type
                                                                        sqfeet
                                                          :39244
##
   Length: 47812
                        Min.
                               :1.000e+00
                                             apartment
                                                                    Min.
                                                                                  1
                        1st Qu.:8.250e+02
                                                                    1st Qu.:
                                                                               770
##
    Class : character
                                             house
                                                          : 4307
##
    Mode :character
                        Median :1.072e+03
                                             townhouse
                                                          : 1836
                                                                    Median:
                                                                               960
##
                        Mean
                               :5.928e+04
                                              condo
                                                             895
                                                                    Mean
                                                                              1040
                                                                              1163
##
                        3rd Qu.:1.455e+03
                                             duplex
                                                          :
                                                             667
                                                                    3rd Qu.:
##
                        Max.
                                :2.768e+09
                                             manufactured:
                                                             571
                                                                    Max.
                                                                           :999999
##
                                              (Other)
                                                             292
                                                         dogs_allowed
##
         beds
                         baths
                                       cats_allowed
                                                                :0.0000
##
           :1.000
                            :1.000
                                             :0.0000
    Min.
                     Min.
                                      Min.
                                                        Min.
    1st Qu.:1.000
                     1st Qu.:1.000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.0000
##
##
    Median :2.000
                     Median :1.000
                                      Median :1.0000
                                                        Median :1.0000
##
    Mean
          :1.949
                     Mean
                            :1.454
                                      Mean
                                             :0.7125
                                                        Mean
                                                                :0.6928
##
    3rd Qu.:2.000
                     3rd Qu.:2.000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
##
    Max.
           :8.000
                     Max.
                            :6.000
                                      Max.
                                              :1.0000
                                                        Max.
                                                                :1.0000
##
##
    smoking allowed
                      wheelchair_access electric_vehicle_charge comes_furnished
    Min.
                             :0.00000
                                         Min.
                                                 :0.00000
                                                                   Min.
##
           :0.0000
                      Min.
                                                                          :0.00000
                                         1st Qu.:0.00000
##
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                                                   1st Qu.:0.00000
##
   Median :1.0000
                      Median :0.00000
                                         Median : 0.00000
                                                                   Median :0.00000
##
    Mean
          :0.7358
                      Mean
                             :0.07787
                                         Mean
                                                 :0.01387
                                                                   Mean
                                                                          :0.04813
    3rd Qu.:1.0000
                                         3rd Qu.:0.00000
##
                      3rd Qu.:0.00000
                                                                   3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.00000
                                         Max.
                                                 :1.00000
                                                                   Max.
                                                                          :1.00000
##
##
              laundry_options
                                           parking_options
                                                               image_url
##
                       : 9918
                                                             Length: 47812
                                                    :17144
##
                       : 4825
                                 off-street parking:15931
                                                             Class : character
    laundry in bldg
    laundry on site
                       : 6796
                                                             Mode : character
                                 carport
                                                    : 5228
##
    no laundry on site:
                          407
                                 attached garage
                                                    : 4939
##
    w/d hookups
                       : 9289
                                 detached garage
                                                    : 2344
##
    w/d in unit
                       :16577
                                 street parking
                                                    : 1866
##
                                 (Other)
                                                       360
##
    description
```

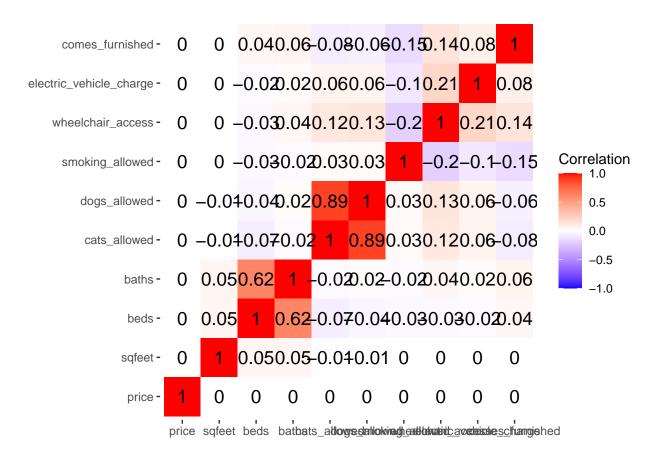
long

state

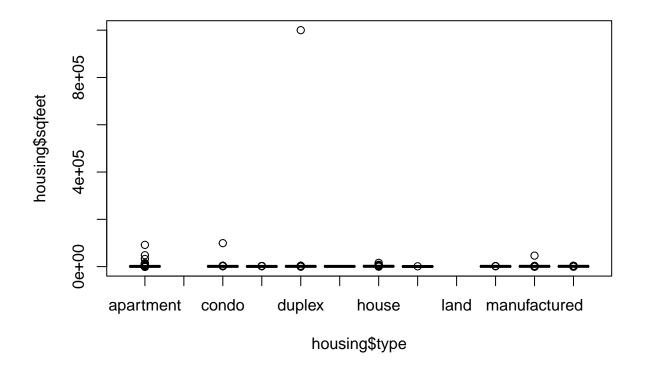
lat

```
Min. : 2.202
                                                                    : 5977
## Length:47812
                                          Min.
                                                  :-163.89
                                                             ca
## Class :character
                        1st Qu.: 33.500
                                         1st Qu.:-101.33
                                                             fl
                                                                    : 5867
  Mode :character
                       Median: 37.974
                                          Median : -86.31
                                                             nc
                                                                    : 3410
##
                               : 37.188
                                                 : -92.25
                                                                    : 2616
                        Mean
                                          Mean
                                                             mi
##
                        3rd Qu.: 41.183
                                          3rd Qu.: -81.25
                                                             ga
                                                                    : 2549
##
                        Max.
                               :102.036
                                          Max.
                                                 : 94.16
                                                                    : 2013
                                                             СО
##
                                                             (Other):25380
library(dplyr)
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##
       filter, lag
## The following objects are masked from 'package:base':
##
       intersect, setdiff, setequal, union
h1=housing
summary(h1$state)
##
          al
               ar
                    az
                          ca
                               СО
                                    ct
                                         dc
                                              de
                                                    fl
                                                         ga
                                                              hi
                                                                   ia
                                                                        id
                                                                              il
##
    364 1112 578 1158 5977 2013
                                                                       817 1711 1200
                                   681
                                        415
                                             371 5867 2549
                                                             319 1385
    ks
          ky
               la
                    ma
                         md
                               me
                                    mi
                                         mn
                                              mo
                                                    ms
                                                         mt
                                                              nc
                                                                   nd
                                                                        ne
                                                                              nh
                                                                                   пj
## 1280
        979 1300
                   895 1364
                               83 2616 1415
                                             401
                                                  903
                                                        229 3410
                                                                  610
                                                                       480
                                                                            305 1028
     nm
          nv
               ny
                    oh
                          ok
                               or
##
   457
        531 1799 1193
                          12
library(inops)
##
## Attaching package: 'inops'
## The following object is masked from 'package:base':
##
##
       <<-
h1=h1[h1$state %out#% 1:100,]
summary(h1$state)
##
                                    ct
                                                    fl
                                                                                   in
     ak
          al
               ar
                    az
                          ca
                               СО
                                         dc
                                              de
                                                         ga
                                                              hi
                                                                   ia
                                                                        id
                                                                              il
    364 1112
              578 1158 5977 2013
                                   681
                                        415
                                             371 5867
                                                       2549
                                                             319 1385
                                                                       817 1711 1200
##
     ks
          ky
               la
                          md
                               me
                                    mi
                                                              nc
                                                                   nd
                    ma
                                         mn
                                              mo
                                                    ms
                                                         \mathtt{mt}
                                                                        ne
## 1280
        979 1300
                   895 1364
                                0 2616 1415 401 903
                                                        229 3410
                                                                  610
                                                                       480
                                                                            305 1028
##
     nm
          nv
               ny
                    oh
                          ok
                               or
    457
        531 1799 1193
                                0
```

```
housing=h1
h1=NULL
library(reshape2)
x=housing[,c(5,7:15)]
summary(x)
                                                       baths
##
       price
                         sqfeet
                                          beds
## Min. :1.000e+00
                     Min. :
                                1 Min. :1.000 Min.
                                                          :1.000
## 1st Qu.:8.250e+02
                     1st Qu.:
                               770
                                    1st Qu.:1.000
                                                   1st Qu.:1.000
## Median :1.070e+03
                     Median :
                                   Median :2.000 Median :1.000
                               960
## Mean :5.940e+04
                     Mean : 1041
                                   Mean :1.949 Mean :1.454
## 3rd Qu.:1.455e+03
                      3rd Qu.: 1163
                                   3rd Qu.:2.000
                                                    3rd Qu.:2.000
## Max.
        :2.768e+09 Max.
                          :999999
                                    Max. :8.000 Max. :6.000
                 dogs_allowed smoking_allowed wheelchair_access
##
   cats allowed
## Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median: 1.000 Median: 1.0000 Median: 0.00000
## Mean
        :0.713 Mean
                       :0.6935
                                Mean
                                       :0.7364
                                                 Mean :0.07799
## 3rd Qu.:1.000
                  3rd Qu.:1.0000
                                 3rd Qu.:1.0000
                                                 3rd Qu.:0.00000
## Max. :1.000 Max.
                        :1.0000 Max.
                                       :1.0000
                                                 Max. :1.00000
## electric_vehicle_charge comes_furnished
## Min. :0.00000
                      Min. :0.00000
## 1st Qu.:0.00000
                         1st Qu.:0.00000
## Median :0.00000
                         Median :0.00000
## Mean :0.01385
                         Mean :0.04795
## 3rd Qu.:0.00000
                         3rd Qu.:0.00000
## Max. :1.00000
                         Max. :1.00000
#calculate correlation coefficients, rounded to 2 decimal places
cor_df <- round(cor(x), 2)</pre>
#melt the data frame
melted_cor <- melt(cor_df)</pre>
#view head of melted data frame
head(melted_cor)
##
           Var1 Var2 value
## 1
          price price
## 2
          sqfeet price
## 3
           beds price
                         0
## 4
          baths price
                         0
## 5 cats allowed price
## 6 dogs_allowed price
library(ggplot2)
#create correlation heatmap
ggplot(data = melted_cor, aes(x=Var1, y=Var2, fill=value)) +
 geom tile() +
 geom_text(aes(Var2, Var1, label = value), size = 5) +
```



boxplot(housing\$sqfeet ~ housing\$type)



```
h1=housing
max(h1$sqfeet)
```

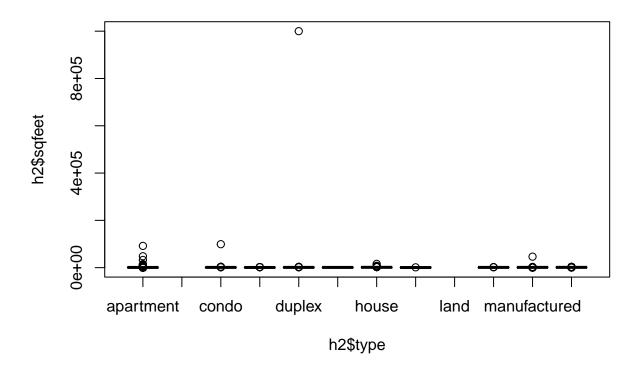
## [1] 999999

```
h2=h1[h1$sqfeet>150 ,] summary(h2)
```

```
##
          id
                              url
                                                                   region
##
            :7.005e+09
                         Length: 47601
                                              omaha / council bluffs :
    Min.
    1st Qu.:7.036e+09
                         Class : character
                                                                          496
##
                                              inland empire
    Median :7.043e+09
                         Mode :character
                                              denver
                                                                          484
                                              ventura county
    Mean
           :7.041e+09
                                                                          479
    3rd Qu.:7.048e+09
                                                                          476
##
                                              orlando
           :7.051e+09
                                              fort collins / north CO:
                                                                         471
##
    Max.
##
                                              (Other)
                                                                       :44687
##
     region_url
                            price
                                                                         sqfeet
                                                        type
##
    Length: 47601
                                :1.000e+00
                                              apartment
                                                           :39088
                                                                    Min.
                                                                                155
    Class :character
                        1st Qu.:8.250e+02
                                              house
                                                                    1st Qu.:
                                                                                770
##
                                                           : 4273
##
    Mode :character
                        Median :1.072e+03
                                              townhouse
                                                           : 1830
                                                                    Median:
                                                                                960
##
                        Mean
                                :5.953e+04
                                              condo
                                                                    Mean
                                                                               1043
                                                              891
                        3rd Qu.:1.455e+03
##
                                              duplex
                                                              663
                                                                    3rd Qu.:
                                                                               1165
##
                        Max.
                                :2.768e+09
                                                              565
                                                                    Max.
                                                                            :999999
                                              manufactured:
##
                                              (Other)
                                                             291
##
                                       cats_allowed
                                                         dogs_allowed
         beds
                         baths
```

```
## Min. :1.000
                   Min. :1.000
                                   Min.
                                         :0.0000
                                                          :0.0000
                                                   Min.
  1st Qu.:1.000
                   1st Qu.:1.000
                                   1st Qu.:0.0000
                                                   1st Qu.:0.0000
## Median :2.000
                   Median :1.000
                                   Median :1.0000
                                                   Median :1.0000
## Mean :1.949
                   Mean :1.455
                                   Mean
                                        :0.7133
                                                   Mean
                                                          :0.6938
##
   3rd Qu.:2.000
                   3rd Qu.:2.000
                                   3rd Qu.:1.0000
                                                   3rd Qu.:1.0000
##
  Max. :8.000
                   Max.
                          :6.000
                                   Max. :1.0000
                                                   Max. :1.0000
##
##
   smoking_allowed wheelchair_access electric_vehicle_charge comes_furnished
##
   Min.
          :0.0000
                    Min.
                           :0.00000
                                     Min.
                                            :0.00000
                                                             Min.
                                                                   :0.00000
                                                             1st Qu.:0.00000
##
   1st Qu.:0.0000
                    1st Qu.:0.00000
                                      1st Qu.:0.00000
  Median :1.0000
                    Median :0.00000
                                     Median :0.00000
                                                             Median :0.00000
## Mean
         :0.7373
                    Mean
                           :0.07807
                                      Mean
                                           :0.01382
                                                             Mean
                                                                    :0.04777
   3rd Qu.:1.0000
                    3rd Qu.:0.00000
                                      3rd Qu.:0.00000
                                                             3rd Qu.:0.00000
##
   Max. :1.0000
                          :1.00000
                    Max.
                                     Max.
                                            :1.00000
                                                             Max.
                                                                    :1.00000
##
##
             laundry_options
                                        parking_options
                                                         image_url
##
                     : 9901
                                               :17114
                                                        Length: 47601
                    : 4770
                                                        Class : character
##
   laundry in bldg
                              off-street parking:15801
##
   laundry on site
                    : 6750
                                               : 5212
                                                        Mode :character
                              carport
   no laundry on site: 394
                              attached garage
##
                                               : 4917
##
   w/d hookups
                    : 9257
                              detached garage
                                              : 2338
##
   w/d in unit
                     :16529
                              street parking
                                               : 1860
##
                              (Other)
                                               : 359
##
   description
                           lat
                                                             state
                                            long
## Length:47601
                            : 2.202
                      Min.
                                       Min.
                                              :-163.89
                                                         ca
                                                                : 5960
  Class : character
                      1st Qu.: 33.496
                                       1st Qu.:-102.82
                                                         fl
                                                                : 5861
## Mode :character
                      Median : 37.973
                                       Median : -86.38
                                                         nc
                                                                : 3404
##
                      Mean : 37.176
                                       Mean
                                              : -92.28
                                                         mi
                                                                : 2614
##
                      3rd Qu.: 41.158
                                       3rd Qu.: -81.27
                                                         ga
                                                                : 2540
##
                      Max.
                             :102.036
                                       Max.
                                              : 94.16
                                                                : 2012
                                                         СО
##
                                                         (Other):25210
```

boxplot(h2\$sqfeet~h2\$type)



```
housing=h2
h1=h2=NULL
summary(housing$id)
                           Median
##
        Min.
               1st Qu.
                                       Mean
                                               3rd Qu.
## 7.005e+09 7.036e+09 7.043e+09 7.041e+09 7.048e+09 7.051e+09
dim(housing)
## [1] 47601
                22
sum(duplicated(housing))
## [1] 0
#Q1 <- quantile(housing$sqfeet, .25)
Q3 <- quantile(housing$sqfeet, .99)
IQR <- IQR(housing$sqfeet)</pre>
#Q1
QЗ
## 99%
## 2432
```

```
IQR
```

```
## [1] 395
```

```
no_outliers <- subset(housing,housing$sqfeet < (Q3 + 1.5*IQR))
dim(no_outliers)</pre>
```

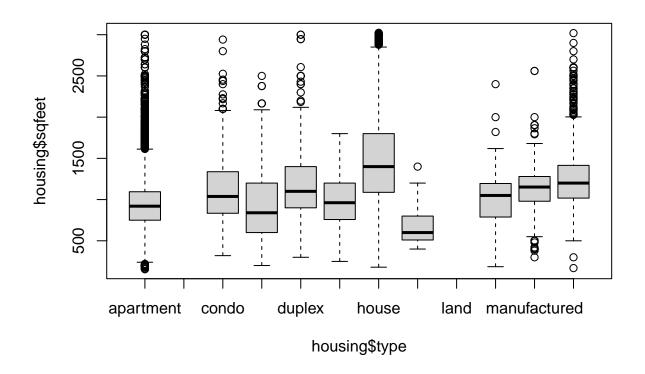
**##** [1] 47438 22

summary(no\_outliers\$sqfeet)

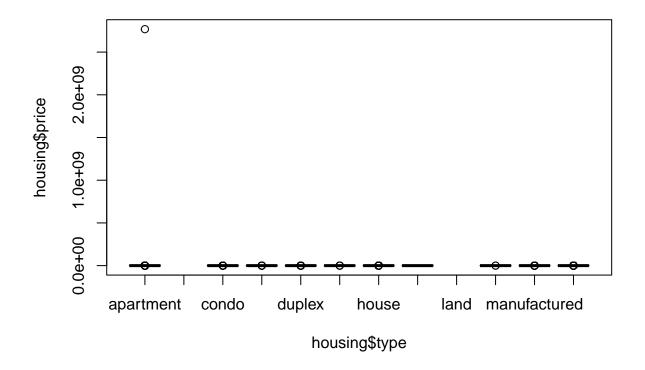
```
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 155 770 960 1005 1160 3023
```

housing=no\_outliers no\_outliers=NULL

boxplot(housing\$sqfeet~housing\$type)



boxplot(housing\$price ~ housing\$type)



```
h1=housing
max(h1$price)
```

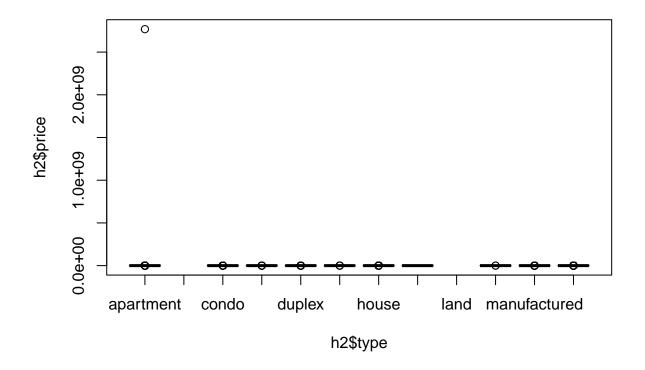
## [1] 2768307249

```
h2=h1[h1$price>150 ,]
summary(h2)
```

```
##
          id
                             url
                                                                  region
##
           :7.005e+09
                         Length: 47308
                                             omaha / council bluffs :
    Min.
    1st Qu.:7.036e+09
                         Class : character
                                             inland empire
                                                                         492
##
    Median :7.043e+09
                         Mode :character
                                             denver
                                                                         478
    Mean
           :7.041e+09
                                             orlando
                                                                         474
    3rd Qu.:7.048e+09
                                                                         472
##
                                             ventura county
           :7.051e+09
                                             fort collins / north CO:
##
    Max.
                                                                         469
##
                                             (Other)
                                                                      :44416
##
     region_url
                            price
                                                                        sqfeet
                                                        type
##
    Length: 47308
                               :1.510e+02
                                             apartment
                                                          :38974
                                                                    Min.
                                                                           : 155
    Class :character
                        1st Qu.:8.250e+02
                                             house
                                                                    1st Qu.: 770
##
                                                          : 4136
##
    Mode :character
                        Median :1.072e+03
                                             townhouse
                                                          : 1828
                                                                    Median: 960
##
                        Mean
                                :5.987e+04
                                             condo
                                                                    Mean
                                                                           :1005
                                                             880
                        3rd Qu.:1.450e+03
##
                                             duplex
                                                             660
                                                                    3rd Qu.:1160
##
                        Max.
                                :2.768e+09
                                                             562
                                                                   Max.
                                                                           :3023
                                             manufactured:
##
                                             (Other)
                                                             268
##
                                       cats_allowed
                                                         dogs_allowed
         beds
                         baths
```

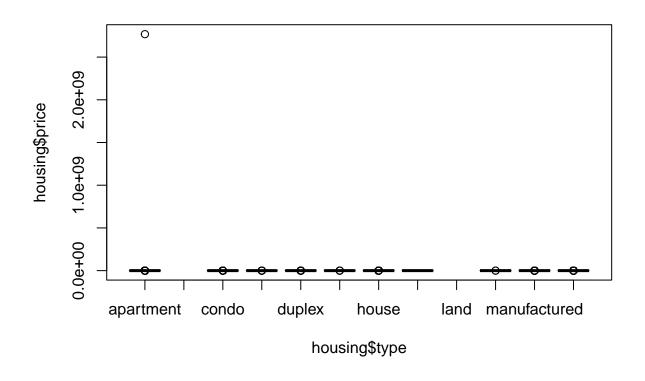
```
## Min. :1.000
                   Min. :1.000
                                   Min.
                                          :0.0000
                                                           :0.0000
                                                    Min.
  1st Qu.:1.000
                   1st Qu.:1.000
                                   1st Qu.:0.0000
                                                   1st Qu.:0.0000
                                                   Median :1.0000
## Median :2.000
                   Median :1.000
                                   Median :1.0000
## Mean :1.941
                   Mean :1.449
                                   Mean
                                         :0.7146
                                                   Mean
                                                           :0.6946
##
   3rd Qu.:2.000
                   3rd Qu.:2.000
                                   3rd Qu.:1.0000
                                                    3rd Qu.:1.0000
##
   Max. :8.000
                   Max.
                          :6.000
                                   Max.
                                         :1.0000
                                                    Max.
                                                           :1.0000
##
##
   smoking_allowed wheelchair_access electric_vehicle_charge comes_furnished
##
   Min.
          :0.0000
                    Min.
                           :0.00000
                                      Min.
                                             :0.00000
                                                              Min.
                                                                    :0.00000
                                                              1st Qu.:0.00000
##
   1st Qu.:0.0000
                    1st Qu.:0.00000
                                      1st Qu.:0.00000
  Median :1.0000
                    Median :0.00000
                                      Median :0.00000
                                                              Median :0.00000
##
  Mean
         :0.7377
                    Mean
                           :0.07821
                                      Mean
                                            :0.01385
                                                              Mean
                                                                     :0.04712
   3rd Qu.:1.0000
                    3rd Qu.:0.00000
                                      3rd Qu.:0.00000
                                                              3rd Qu.:0.00000
##
         :1.0000
                           :1.00000
   Max.
                    Max.
                                      Max.
                                             :1.00000
                                                              Max.
                                                                    :1.00000
##
##
             laundry_options
                                        parking_options
                                                          image_url
##
                     : 9865
                                                :17020
                                                         Length: 47308
                    : 4741
                                                         Class : character
##
   laundry in bldg
                              off-street parking:15740
##
   laundry on site
                    : 6729
                                                : 5183
                                                         Mode :character
                              carport
   no laundry on site: 391
##
                              attached garage
                                               : 4813
##
   w/d hookups
                     : 9186
                              detached garage
                                               : 2333
##
   w/d in unit
                     :16396
                              street parking
                                                : 1860
##
                              (Other)
                                                : 359
##
   description
                           lat
                                                              state
                                             long
## Length:47308
                             : 2.202
                      Min.
                                        Min.
                                               :-163.89
                                                          ca
                                                                 : 5909
  Class : character
                      1st Qu.: 33.495
                                        1st Qu.:-101.32
                                                          fl
                                                                 : 5839
  Mode :character
##
                      Median : 37.974
                                        Median : -86.28
                                                          nc
                                                                 : 3388
##
                      Mean
                             : 37.178
                                        Mean
                                               : -92.24
                                                          mi
                                                                 : 2601
##
                      3rd Qu.: 41.159
                                        3rd Qu.: -81.26
                                                                 : 2534
                                                          ga
##
                      Max.
                             :102.036
                                        Max.
                                               : 94.16
                                                          СО
                                                                 : 1997
##
                                                          (Other):25040
```

boxplot(h2\$price~h2\$type)

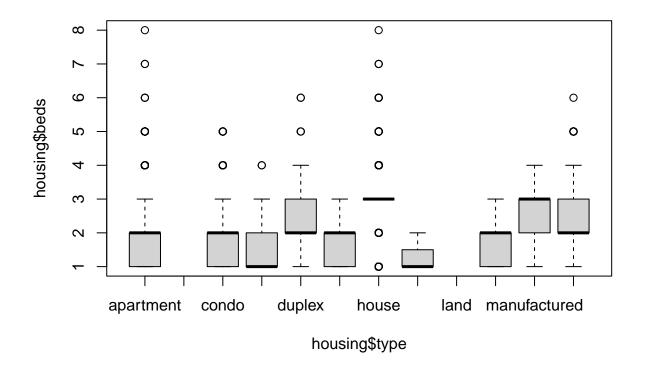


housing=h2 h1=h2=NULL

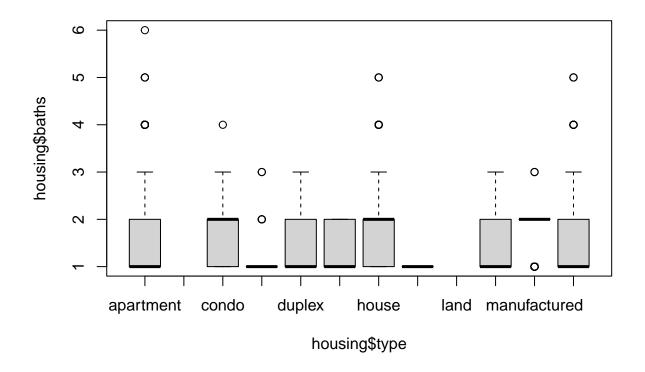
boxplot(housing\$price ~ housing\$type)



```
\#Q1 \leftarrow quantile(housing\$price, .25)
Q3 <- quantile(housing$price, .99)
IQR <- IQR(housing$price)</pre>
QЗ
##
    99%
## 3600
IQR
## [1] 625
no_outliers <- subset(housing, housingprice < (Q3 + 1.5*IQR))
dim(no_outliers)
## [1] 47121
                 22
housing=no_outliers
no_outliers=NULL
boxplot(housing$beds ~ housing$type)
```



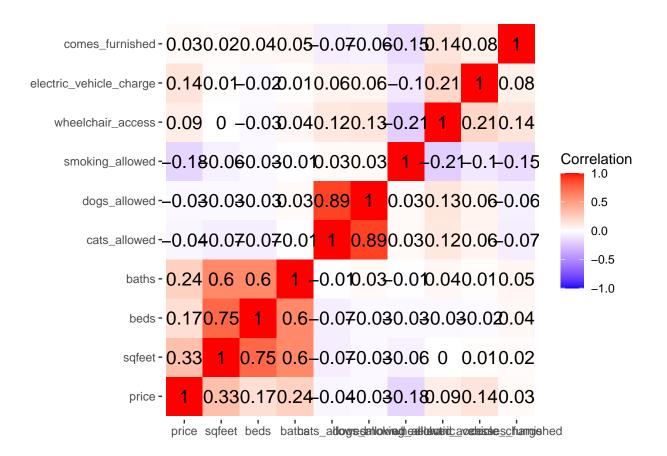
boxplot(housing\$baths ~ housing\$type)



```
library(reshape2)
x=housing[,c(5,7:15)]
summary(x)
```

```
cats allowed
##
                        sqfeet
                                         beds
                                                         baths
        price
##
    Min.
           : 151
                    Min.
                            : 155
                                    Min.
                                            :1.000
                                                     Min.
                                                             :1.000
                                                                      Min.
                                                                              :0.0000
    1st Qu.: 825
                                                                      1st Qu.:0.0000
##
                    1st Qu.: 770
                                    1st Qu.:1.000
                                                     1st Qu.:1.000
##
    Median:1066
                    Median: 959
                                    Median :2.000
                                                     Median :1.000
                                                                      Median :1.0000
##
    Mean
            :1214
                    Mean
                            :1002
                                            :1.938
                                                             :1.446
                                                                      Mean
                                                                              :0.7153
                                    Mean
                                                     Mean
##
    3rd Qu.:1450
                    3rd Qu.:1157
                                    3rd Qu.:2.000
                                                     3rd Qu.:2.000
                                                                      3rd Qu.:1.0000
    Max.
            :4519
                            :3023
                                            :8.000
                                                     Max.
                                                             :6.000
                                                                      Max.
                                                                              :1.0000
##
                    Max.
                                    Max.
     dogs_allowed
                                        wheelchair_access electric_vehicle_charge
##
                      smoking_allowed
    Min.
           :0.0000
                              :0.0000
                                        Min.
                                                :0.00000
                                                            Min.
                                                                   :0.00000
##
                      Min.
    1st Qu.:0.0000
                      1st Qu.:0.0000
                                        1st Qu.:0.00000
                                                            1st Qu.:0.00000
##
    Median :1.0000
                                        Median :0.00000
##
                      Median :1.0000
                                                            Median :0.00000
            :0.6953
##
    Mean
                      Mean
                              :0.7385
                                        Mean
                                                :0.07786
                                                            Mean
                                                                   :0.01367
    3rd Qu.:1.0000
                      3rd Qu.:1.0000
                                        3rd Qu.:0.00000
                                                            3rd Qu.:0.00000
##
##
    Max.
            :1.0000
                      Max.
                              :1.0000
                                        Max.
                                                :1.00000
                                                            Max.
                                                                   :1.00000
##
    comes furnished
##
    Min.
            :0.0000
##
    1st Qu.:0.00000
    Median :0.00000
##
##
    Mean
            :0.04626
##
    3rd Qu.:0.00000
    Max.
           :1.00000
```

```
{\it \#calculate \ correlation \ coefficients, \ rounded \ to \ 2 \ decimal \ places}
cor_df <- round(cor(x), 2)</pre>
#melt the data frame
melted_cor <- melt(cor_df)</pre>
#view head of melted data frame
head(melted_cor)
            Var1 Var2 value
##
## 1
           price price 1.00
         sqfeet price 0.33
## 2
            beds price 0.17
## 3
## 4
            baths price 0.24
## 5 cats_allowed price -0.04
## 6 dogs_allowed price -0.03
library(ggplot2)
#create correlation heatmap
ggplot(data = melted_cor, aes(x=Var1, y=Var2, fill=value)) +
 geom_tile() +
  geom_text(aes(Var2, Var1, label = value), size = 5) +
  scale_fill_gradient2(low = "blue", high = "red",
                       limit = c(-1,1), name="Correlation") +
 theme(axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        panel.background = element_blank())
```



# housing\$ppa=housing\$price/housing\$sqfeet summary(housing)

```
##
          id
                             url
                                                                   region
           :7.005e+09
##
                         Length: 47121
    Min.
                                              omaha / council bluffs :
    1st Qu.:7.036e+09
                         Class : character
                                              inland empire
                                                                         492
    Median :7.043e+09
                         Mode :character
                                              denver
                                                                         478
    Mean
           :7.041e+09
                                              orlando
                                                                         474
    3rd Qu.:7.048e+09
                                              fort collins / north CO:
                                                                         469
##
           :7.051e+09
##
                                             daytona beach
                                                                         468
##
                                              (Other)
                                                                       :44233
                                                                   sqfeet
                            price
##
     region_url
                                                   type
    Length: 47121
##
                        Min.
                               : 151
                                        apartment
                                                     :38876
                                                               Min.
                                                                      : 155
    Class : character
                        1st Qu.: 825
                                        house
                                                     : 4088
                                                               1st Qu.: 770
    Mode :character
                        Median:1066
                                                               Median: 959
##
                                        townhouse
                                                     : 1819
##
                        Mean
                               :1214
                                        condo
                                                        860
                                                               Mean
                                                                      :1002
##
                        3rd Qu.:1450
                                        duplex
                                                        655
                                                               3rd Qu.:1157
##
                        Max.
                                        manufactured:
                                                        558
                                                               Max.
                                                                      :3023
                                                        265
##
                                        (Other)
##
         heds
                         baths
                                       cats_allowed
                                                         dogs_allowed
    Min.
           :1.000
                             :1.000
                                             :0.0000
                                                        Min.
                                                                :0.0000
                     Min.
                                      1st Qu.:0.0000
    1st Qu.:1.000
                     1st Qu.:1.000
                                                        1st Qu.:0.0000
    Median :2.000
                     Median :1.000
                                      Median :1.0000
                                                        Median :1.0000
##
    Mean
           :1.938
                     Mean
                            :1.446
                                      Mean
                                             :0.7153
                                                        Mean
                                                                :0.6953
    3rd Qu.:2.000
                     3rd Qu.:2.000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
```

```
##
    Max.
           :8.000
                            :6.000
                                     Max.
                                             :1.0000
                                                       Max.
                                                               :1.0000
                     Max.
##
##
    smoking allowed
                      wheelchair access electric vehicle charge comes furnished
                             :0.00000
                                                :0.00000
    Min.
          :0.0000
                      Min.
                                        Min.
                                                                  Min.
                                                                         :0.00000
##
##
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                         1st Qu.:0.00000
                                                                  1st Qu.:0.00000
##
    Median :1.0000
                      Median :0.00000
                                        Median :0.00000
                                                                  Median :0.00000
    Mean :0.7385
                             :0.07786
                                                :0.01367
                      Mean
                                        Mean
                                                                  Mean
                                                                         :0.04626
    3rd Qu.:1.0000
##
                      3rd Qu.:0.00000
                                         3rd Qu.:0.00000
                                                                  3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.00000
                                        Max.
                                                :1.00000
                                                                  Max.
                                                                         :1.00000
##
##
              laundry_options
                                           parking_options
                                                              image_url
##
                       : 9835
                                                             Length: 47121
                                                   :16972
##
    laundry in bldg
                       : 4732
                                off-street parking:15709
                                                             Class : character
                       : 6718
                                                   : 5167
##
    laundry on site
                                carport
                                                             Mode :character
    no laundry on site:
                                                   : 4748
##
                          391
                                attached garage
##
    w/d hookups
                       : 9171
                                detached garage
                                                   : 2316
##
    w/d in unit
                       :16274
                                street parking
                                                   : 1853
##
                                (Other)
                                                      356
##
    description
                                                                  state
                             lat
                                                long
##
    Length: 47121
                        Min.
                                  2.202
                                           Min.
                                                  :-163.89
                                                              fl
                                                                     : 5816
##
    Class : character
                        1st Qu.: 33.492
                                           1st Qu.:-101.30
                                                                     : 5806
                                                              ca
    Mode :character
                        Median: 37.974
                                          Median : -86.26
                                                                     : 3384
                                                              nc
                               : 37.186
                                                  : -92.19
##
                        Mean
                                          Mean
                                                                     : 2597
                                                              mi
##
                        3rd Qu.: 41.176
                                           3rd Qu.: -81.26
                                                              ga
                                                                     : 2532
##
                        Max.
                               :102.036
                                                                     : 1989
                                          Max.
                                                  : 94.16
                                                              СО
##
                                                              (Other):24997
##
         ppa
          : 0.1250
##
    Min.
##
    1st Qu.: 0.8824
##
    Median: 1.1192
##
    Mean : 1.2805
##
    3rd Qu.: 1.5113
##
    Max. :10.9388
##
```

### housing\$bba=housing\$beds/housing\$sqfeet summary(housing)

```
##
          id
                             url
                                                                  region
##
   Min.
           :7.005e+09
                         Length: 47121
                                             omaha / council bluffs :
                                                                        507
##
   1st Qu.:7.036e+09
                         Class : character
                                             inland empire
                                                                        492
   Median :7.043e+09
                         Mode :character
                                             denver
                                                                        478
    Mean
           :7.041e+09
                                                                        474
##
                                             orlando
##
    3rd Qu.:7.048e+09
                                             fort collins / north CO:
                                                                        469
##
    Max. :7.051e+09
                                             daytona beach
                                                                        468
##
                                             (Other)
                                                                     :44233
##
    region_url
                            price
                                                                  sqfeet
                                                  type
##
   Length: 47121
                        Min.
                              : 151
                                        apartment
                                                    :38876
                                                              Min.
                                                                     : 155
    Class : character
                        1st Qu.: 825
                                                    : 4088
                                                              1st Qu.: 770
                                        house
                                                              Median: 959
##
    Mode :character
                        Median:1066
                                        townhouse
                                                    : 1819
##
                        Mean
                               :1214
                                        condo
                                                       860
                                                              Mean
                                                                     :1002
##
                        3rd Qu.:1450
                                        duplex
                                                    :
                                                       655
                                                              3rd Qu.:1157
##
                               :4519
                                       manufactured:
                        Max.
                                                       558
                                                              Max.
                                                                     :3023
##
                                        (Other)
                                                       265
                                                    :
```

```
##
         beds
                         baths
                                       cats allowed
                                                         dogs_allowed
           :1.000
                            :1.000
                                            :0.0000
                                                               :0.0000
##
    Min.
                     Min.
                                      Min.
                                                        Min.
                     1st Qu.:1.000
                                                        1st Qu.:0.0000
##
    1st Qu.:1.000
                                      1st Qu.:0.0000
                                      Median :1.0000
    Median :2.000
                     Median :1.000
                                                        Median :1.0000
##
##
    Mean
           :1.938
                     Mean
                            :1.446
                                      Mean
                                             :0.7153
                                                        Mean
                                                               :0.6953
    3rd Qu.:2.000
                     3rd Qu.:2.000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
##
    Max.
           :8.000
                            :6.000
                                             :1.0000
##
                     Max.
                                      Max.
                                                        Max.
                                                               :1.0000
##
##
    smoking allowed
                      wheelchair_access electric_vehicle_charge comes_furnished
           :0.0000
                             :0.00000
                                         Min.
                                                :0.00000
##
    Min.
                      Min.
                                                                  Min.
                                                                          :0.00000
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                         1st Qu.:0.00000
                                                                  1st Qu.:0.00000
    Median :1.0000
##
                      Median :0.00000
                                         Median :0.00000
                                                                  Median :0.00000
##
    Mean
           :0.7385
                      Mean
                             :0.07786
                                         Mean
                                                :0.01367
                                                                  Mean
                                                                          :0.04626
    3rd Qu.:1.0000
##
                      3rd Qu.:0.00000
                                         3rd Qu.:0.00000
                                                                  3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.00000
                                                                          :1.00000
                                         Max.
                                                 :1.00000
                                                                  Max.
##
##
              laundry_options
                                                              image_url
                                           parking_options
##
                       : 9835
                                                             Length: 47121
                                                    :16972
                       : 4732
                                                             Class : character
##
    laundry in bldg
                                off-street parking:15709
##
    laundry on site
                       : 6718
                                 carport
                                                   : 5167
                                                             Mode : character
##
    no laundry on site:
                          391
                                attached garage
                                                   : 4748
    w/d hookups
                       : 9171
                                 detached garage
                                                    : 2316
##
    w/d in unit
                       :16274
                                 street parking
                                                    : 1853
                                 (Other)
                                                      356
##
##
    description
                             lat
                                                long
                                                                   state
##
   Length: 47121
                        Min.
                               : 2.202
                                           Min.
                                                   :-163.89
                                                              fl
                                                                      : 5816
##
    Class : character
                        1st Qu.: 33.492
                                           1st Qu.:-101.30
                                                                      : 5806
                                                              ca
                        Median: 37.974
##
    Mode :character
                                           Median: -86.26
                                                                       3384
                                                              nc
##
                        Mean
                               : 37.186
                                                   : -92.19
                                                                      : 2597
                                           Mean
                                                              mi
                                           3rd Qu.: -81.26
##
                        3rd Qu.: 41.176
                                                                      : 2532
                                                              ga
##
                        Max.
                               :102.036
                                           Max.
                                                   : 94.16
                                                                      : 1989
##
                                                              (Other):24997
##
                            bba
         ppa
                              :0.0004167
##
          : 0.1250
    Min.
                       Min.
##
    1st Qu.: 0.8824
                       1st Qu.:0.0015385
    Median : 1.1192
                       Median: 0.0019048
##
##
    Mean
          : 1.2805
                       Mean
                             :0.0019355
##
    3rd Qu.: 1.5113
                       3rd Qu.:0.0022472
##
    Max.
           :10.9388
                       Max.
                              :0.0228571
##
```

housing\$apb=housing\$sqfeet/housing\$beds
housing=subset(housing,apb>=200)
summary(housing)

```
##
          id
                              url
                                                                   region
##
    Min.
           :7.005e+09
                         Length: 47087
                                              omaha / council bluffs:
##
    1st Qu.:7.036e+09
                          Class : character
                                                                         491
                                              inland empire
    Median :7.043e+09
                          Mode :character
                                              denver
                                                                         478
            :7.041e+09
                                                                         474
##
    Mean
                                              orlando
##
    3rd Qu.:7.048e+09
                                              daytona beach
                                                                         468
##
    Max.
           :7.051e+09
                                              fayetteville, NC
                                                                         468
##
                                              (Other)
                                                                      :44201
##
     region_url
                             price
                                                   type
                                                                    sqfeet
```

```
Length: 47087
                        Min. : 151
                                       apartment
                                                    :38849
                                                             Min. : 200
                                                             1st Qu.: 770
##
                                       house
    Class : character
                        1st Qu.: 825
                                                    : 4087
                                       townhouse
##
    Mode :character
                        Median:1068
                                                    : 1818
                                                             Median: 960
##
                        Mean
                               :1214
                                       condo
                                                       858
                                                             Mean
                                                                     :1003
##
                        3rd Qu.:1450
                                       duplex
                                                    :
                                                       654
                                                             3rd Qu.:1158
##
                        Max.
                               :4519
                                       manufactured:
                                                       557
                                                                     :3023
                                                             Max.
##
                                        (Other)
                                                       264
##
         beds
                         baths
                                      cats allowed
                                                        dogs allowed
##
    Min.
           :1.000
                    Min.
                            :1.000
                                     Min.
                                            :0.0000
                                                       Min.
                                                               :0.0000
##
    1st Qu.:1.000
                    1st Qu.:1.000
                                     1st Qu.:0.0000
                                                       1st Qu.:0.0000
    Median :2.000
                    Median :1.000
                                     Median :1.0000
                                                       Median :1.0000
    Mean
          :1.938
                          :1.446
                                            :0.7156
##
                    Mean
                                     Mean
                                                       Mean
                                                               :0.6955
##
    3rd Qu.:2.000
                    3rd Qu.:2.000
                                     3rd Qu.:1.0000
                                                       3rd Qu.:1.0000
##
                            :6.000
    Max.
          :8.000
                    Max.
                                     Max.
                                            :1.0000
                                                       Max.
                                                              :1.0000
##
##
    smoking_allowed
                     wheelchair_access electric_vehicle_charge comes_furnished
##
           :0.0000
                                        Min.
                                                :0.00000
    Min.
                     Min.
                             :0.00000
                                                                  Min.
                                                                         :0.00000
##
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                        1st Qu.:0.00000
                                                                  1st Qu.:0.00000
    Median :1.0000
                     Median :0.00000
                                        Median : 0.00000
                                                                 Median :0.00000
##
##
    Mean
          :0.7386
                     Mean
                             :0.07788
                                        Mean
                                                :0.01363
                                                                  Mean
                                                                         :0.04606
##
    3rd Qu.:1.0000
                     3rd Qu.:0.00000
                                        3rd Qu.:0.00000
                                                                  3rd Qu.:0.00000
##
    Max.
          :1.0000
                     Max.
                             :1.00000
                                        Max.
                                                                  Max.
                                                                         :1.00000
                                                :1.00000
##
##
              laundry_options
                                           parking_options
                                                             image url
##
                                                            Length: 47087
                       : 9828
                                                   :16964
##
    laundry in bldg
                       : 4728
                                off-street parking:15692
                                                            Class : character
##
    laundry on site
                       : 6709
                                                   : 5166
                                                            Mode : character
                                carport
##
    no laundry on site:
                          387
                                attached garage
                                                  : 4747
##
    w/d hookups
                                detached garage
                       : 9168
                                                   : 2313
##
    w/d in unit
                       :16267
                                street parking
                                                   : 1850
##
                                (Other)
                                                   :
                                                      355
##
    description
                             lat
                                                                  state
                                                long
##
    Length: 47087
                        Min.
                               : 2.202
                                          Min.
                                                  :-163.89
                                                             fl
                                                                     : 5815
                        1st Qu.: 33.491
                                                                     : 5796
##
    Class :character
                                          1st Qu.:-101.29
                                                             ca
##
    Mode :character
                        Median: 37.974
                                          Median: -86.26
                                                                     : 3383
                                                             nc
##
                        Mean
                               : 37.184
                                          Mean
                                                  : -92.18
                                                             mi
                                                                     : 2593
##
                        3rd Qu.: 41.176
                                           3rd Qu.: -81.26
                                                             ga
                                                                     : 2532
##
                        Max.
                               :102.036
                                          Max.
                                                  : 94.16
                                                                     : 1988
                                                             CO
##
                                                              (Other):24980
##
                            bba
                                                 apb
         ppa
                              :0.0004167
                                                  : 200.0
##
    Min.
          : 0.1250
                      Min.
                                           Min.
    1st Qu.: 0.8821
                       1st Qu.:0.0015373
                                            1st Qu.: 445.8
##
    Median: 1.1187
                      Median: 0.0019048
                                           Median: 525.0
##
    Mean
          : 1.2794
                       Mean
                              :0.0019315
                                           Mean
                                                  : 559.2
    3rd Qu.: 1.5105
                       3rd Qu.:0.0022434
                                            3rd Qu.: 650.5
##
    Max.
           :10.3713
                      Max.
                              :0.0050000
                                            Max.
                                                   :2400.0
##
```

# housing\$b\_a=housing\$beds\*housing\$sqfeet summary(housing)

```
## id url region
## Min. :7.005e+09 Length:47087 omaha / council bluffs: 507
## 1st Qu.:7.036e+09 Class :character inland empire : 491
```

```
478
    Median :7.043e+09
                         Mode :character
                                             denver
##
    Mean
           :7.041e+09
                                             orlando
                                                                       474
##
    3rd Qu.:7.048e+09
                                             daytona beach
                                                                       468
           :7.051e+09
                                                                       468
##
    Max.
                                             fayetteville, NC
##
                                             (Other)
                                                                    :44201
##
     region url
                            price
                                                  type
                                                                  sqfeet
##
    Length: 47087
                        Min.
                               : 151
                                        apartment
                                                     :38849
                                                              Min.
                                                                     : 200
    Class : character
                        1st Qu.: 825
                                                              1st Qu.: 770
##
                                        house
                                                     : 4087
    Mode :character
##
                        Median:1068
                                        townhouse
                                                     : 1818
                                                              Median: 960
##
                                                        858
                                                                     :1003
                        Mean
                               :1214
                                        condo
                                                              Mean
##
                        3rd Qu.:1450
                                        duplex
                                                        654
                                                              3rd Qu.:1158
##
                               :4519
                                                        557
                                                                     :3023
                        Max.
                                        manufactured:
                                                              Max.
##
                                        (Other)
                                                        264
##
                         baths
                                       cats_allowed
         beds
                                                         dogs_allowed
                                                               :0.0000
##
                            :1.000
                                                        Min.
    Min.
           :1.000
                     Min.
                                      Min.
                                             :0.0000
##
    1st Qu.:1.000
                     1st Qu.:1.000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.0000
##
    Median :2.000
                     Median :1.000
                                                        Median :1.0000
                                     Median :1.0000
##
    Mean
          :1.938
                     Mean :1.446
                                      Mean
                                            :0.7156
                                                        Mean
                                                               :0.6955
                                                        3rd Qu.:1.0000
##
    3rd Qu.:2.000
                     3rd Qu.:2.000
                                      3rd Qu.:1.0000
##
    Max.
           :8.000
                     Max.
                            :6.000
                                     Max.
                                             :1.0000
                                                        Max.
                                                               :1.0000
##
##
    smoking_allowed
                      wheelchair_access electric_vehicle_charge comes_furnished
##
    Min.
           :0.0000
                      Min.
                             :0.00000
                                         Min.
                                                :0.00000
                                                                  Min.
                                                                          :0.00000
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                         1st Qu.:0.00000
                                                                  1st Qu.:0.00000
##
##
    Median :1.0000
                      Median :0.00000
                                         Median :0.00000
                                                                  Median :0.00000
    Mean
           :0.7386
                      Mean
                             :0.07788
                                         Mean
                                                :0.01363
                                                                  Mean
                                                                          :0.04606
##
    3rd Qu.:1.0000
                      3rd Qu.:0.00000
                                         3rd Qu.:0.00000
                                                                  3rd Qu.:0.00000
##
           :1.0000
                             :1.00000
    Max.
                      Max.
                                         Max.
                                                :1.00000
                                                                  Max.
                                                                          :1.00000
##
##
              laundry_options
                                           parking_options
                                                              image_url
##
                       : 9828
                                                    :16964
                                                             Length: 47087
##
    laundry in bldg
                       : 4728
                                off-street parking:15692
                                                             Class : character
##
    laundry on site
                       : 6709
                                carport
                                                   : 5166
                                                             Mode : character
                                                   : 4747
##
    no laundry on site:
                          387
                                attached garage
##
    w/d hookups
                       : 9168
                                detached garage
                                                   : 2313
##
    w/d in unit
                       :16267
                                street parking
                                                    : 1850
##
                                 (Other)
                                                      355
##
    description
                                                                  state
                             lat
                                                long
##
    Length: 47087
                        Min.
                                 2.202
                                                  :-163.89
                                                                     : 5815
                                           Min.
                                                              fl
##
    Class : character
                        1st Qu.: 33.491
                                           1st Qu.:-101.29
                                                                      : 5796
                                                              ca
##
    Mode :character
                        Median: 37.974
                                           Median : -86.26
                                                                      : 3383
                                                              nc
                               : 37.184
##
                        Mean
                                           Mean
                                                  : -92.18
                                                                     : 2593
                                                              тi
##
                        3rd Qu.: 41.176
                                           3rd Qu.: -81.26
                                                              ga
                                                                     : 2532
##
                        Max.
                               :102.036
                                                  : 94.16
                                           Max.
                                                                      : 1988
                                                              СО
##
                                                              (Other):24980
##
         ppa
                            bba
                                                 apb
                                                                   b_a
    Min.
##
         : 0.1250
                       Min.
                              :0.0004167
                                            Min.
                                                   : 200.0
                                                              Min.
                                                                         200
                                                                     :
                                            1st Qu.: 445.8
    1st Qu.: 0.8821
                       1st Qu.:0.0015373
                                                              1st Qu.:
                                                                        807
##
    Median: 1.1187
                       Median :0.0019048
                                            Median: 525.0
                                                              Median: 1904
                                                  : 559.2
                                                                    : 2152
##
    Mean
          : 1.2794
                       Mean
                              :0.0019315
                                            Mean
                                                              Mean
##
    3rd Qu.: 1.5105
                       3rd Qu.:0.0022434
                                            3rd Qu.: 650.5
                                                              3rd Qu.: 2500
           :10.3713
                                                   :2400.0
##
    Max.
                       Max.
                             :0.0050000
                                            Max.
                                                              Max.
                                                                     :21272
##
```

### housing\$b\_b=housing\$beds\*housing\$baths summary(housing)

```
##
                             url
                                                                  region
          id
##
    Min.
           :7.005e+09
                         Length: 47087
                                              omaha / council bluffs:
                                                                        507
##
    1st Qu.:7.036e+09
                                              inland empire
                                                                        491
                         Class :character
    Median :7.043e+09
                                                                        478
##
                         Mode :character
                                              denver
##
    Mean
           :7.041e+09
                                              orlando
                                                                        474
    3rd Qu.:7.048e+09
                                              daytona beach
                                                                        468
           :7.051e+09
##
    Max.
                                              fayetteville, NC
                                                                        468
                                              (Other)
##
                                                                     :44201
##
                            price
     region_url
                                                   type
                                                                   sqfeet
   Length: 47087
##
                        Min.
                                : 151
                                        apartment
                                                     :38849
                                                               Min.
                                                                      : 200
    Class : character
                                                               1st Qu.: 770
##
                        1st Qu.: 825
                                        house
                                                     : 4087
##
    Mode :character
                        Median:1068
                                        townhouse
                                                     : 1818
                                                               Median: 960
##
                        Mean
                                :1214
                                        condo
                                                        858
                                                               Mean
                                                                      :1003
##
                        3rd Qu.:1450
                                                        654
                                        duplex
                                                               3rd Qu.:1158
##
                        Max.
                                :4519
                                        manufactured:
                                                        557
                                                               Max.
                                                                      :3023
##
                                        (Other)
                                                        264
##
         beds
                         baths
                                       cats_allowed
                                                         dogs_allowed
##
           :1.000
                            :1.000
                                      Min.
                                             :0.0000
                                                        Min.
                                                                :0.0000
    Min.
                     Min.
    1st Qu.:1.000
                     1st Qu.:1.000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.0000
##
##
    Median :2.000
                     Median :1.000
                                      Median :1.0000
                                                        Median :1.0000
    Mean
          :1.938
                     Mean
                            :1.446
                                      Mean
                                             :0.7156
                                                        Mean
                                                                :0.6955
    3rd Qu.:2.000
                     3rd Qu.:2.000
##
                                      3rd Qu.:1.0000
                                                        3rd Qu.:1.0000
##
    Max.
           :8.000
                     Max.
                            :6.000
                                      Max.
                                             :1.0000
                                                        Max.
                                                                :1.0000
##
##
    smoking allowed
                      wheelchair_access electric_vehicle_charge comes_furnished
##
    Min.
           :0.0000
                      Min.
                              :0.00000
                                         Min.
                                                 :0.00000
                                                                   Min.
                                                                           :0.00000
##
    1st Qu.:0.0000
                      1st Qu.:0.00000
                                         1st Qu.:0.00000
                                                                   1st Qu.:0.00000
##
    Median :1.0000
                      Median :0.00000
                                         Median :0.00000
                                                                   Median :0.00000
##
    Mean
           :0.7386
                      Mean
                              :0.07788
                                         Mean
                                                 :0.01363
                                                                   Mean
                                                                           :0.04606
##
    3rd Qu.:1.0000
                      3rd Qu.:0.00000
                                         3rd Qu.:0.00000
                                                                   3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.00000
                                         Max.
                                                 :1.00000
                                                                   Max.
                                                                          :1.00000
##
##
               laundry_options
                                           parking_options
                                                               image_url
##
                       : 9828
                                                    :16964
                                                              Length: 47087
                       : 4728
##
    laundry in bldg
                                 off-street parking:15692
                                                              Class : character
##
    laundry on site
                       : 6709
                                 carport
                                                    : 5166
                                                              Mode :character
    no laundry on site:
                                                    : 4747
##
                          387
                                 attached garage
##
    w/d hookups
                       : 9168
                                 detached garage
                                                    : 2313
##
    w/d in unit
                       :16267
                                 street parking
                                                    : 1850
##
                                 (Other)
                                                       355
    description
##
                             lat
                                                 long
                                                                   state
##
    Length: 47087
                        Min.
                                :
                                 2.202
                                           Min.
                                                   :-163.89
                                                               fl
                                                                      : 5815
##
    Class : character
                        1st Qu.: 33.491
                                           1st Qu.:-101.29
                                                                      : 5796
                                                               ca
    Mode : character
                        Median: 37.974
                                           Median: -86.26
                                                                      : 3383
                                                               nc
##
                                                                        2593
                        Mean
                                : 37.184
                                           Mean
                                                   : -92.18
                                                               mi
##
                        3rd Qu.: 41.176
                                           3rd Qu.: -81.26
                                                                      : 2532
                                                               ga
##
                        Max.
                                :102.036
                                           Max.
                                                      94.16
                                                                      : 1988
                                                   :
##
                                                               (Other):24980
##
         ppa
                            bba
                                                  apb
                                                                    b_a
          : 0.1250
                               :0.0004167
    Min.
                       Min.
                                            Min.
                                                    : 200.0
                                                               Min.
                                                                         200
```

```
1st Qu.: 0.8821
                      1st Qu.:0.0015373
                                          1st Qu.: 445.8
                                                            1st Qu.: 807
##
   Median: 1.1187
                      Median :0.0019048
                                          Median: 525.0
                                                           Median: 1904
   Mean
         : 1.2794
                      Mean
                             :0.0019315
                                          Mean : 559.2
                                                           Mean
                                                                 : 2152
   3rd Qu.: 1.5105
                      3rd Qu.:0.0022434
                                          3rd Qu.: 650.5
                                                           3rd Qu.: 2500
##
##
   Max.
          :10.3713
                      Max.
                             :0.0050000
                                          Max.
                                                 :2400.0
                                                           Max.
                                                                  :21272
##
##
         b_b
##
   Min.
          : 1.000
##
   1st Qu.: 1.000
##
   Median : 2.000
   Mean
          : 3.074
   3rd Qu.: 4.000
##
##
   Max.
           :36.000
##
```

## housing\$apbth=housing\$sqfeet\*housing\$baths summary(housing)

```
##
          id
                             url
                                                                 region
##
   Min.
           :7.005e+09
                         Length: 47087
                                             omaha / council bluffs:
                                                                       507
                                                                       491
    1st Qu.:7.036e+09
                         Class : character
                                             inland empire
   Median :7.043e+09
                         Mode :character
                                             denver
                                                                       478
  Mean
          :7.041e+09
                                             orlando
                                                                       474
    3rd Qu.:7.048e+09
                                                                       468
##
                                             daytona beach
##
    Max.
           :7.051e+09
                                             fayetteville, NC
                                                                       468
##
                                             (Other)
                                                                    :44201
##
    region_url
                            price
                                                                  sqfeet
                                                  type
##
    Length: 47087
                        Min.
                               : 151
                                        apartment
                                                    :38849
                                                             Min.
                                                                     : 200
##
    Class :character
                        1st Qu.: 825
                                                             1st Qu.: 770
                                       house
                                                    : 4087
##
    Mode :character
                        Median:1068
                                       townhouse
                                                    : 1818
                                                             Median: 960
##
                        Mean
                               :1214
                                                       858
                                                             Mean
                                                                   :1003
                                        condo
##
                        3rd Qu.:1450
                                       duplex
                                                       654
                                                             3rd Qu.:1158
                                                    :
                                       manufactured:
##
                        Max.
                               :4519
                                                       557
                                                             Max.
                                                                     :3023
##
                                        (Other)
                                                       264
##
                         baths
                                      cats_allowed
         beds
                                                        dogs_allowed
##
    Min.
          :1.000
                    Min.
                          :1.000
                                     Min.
                                            :0.0000
                                                       Min.
                                                               :0.0000
##
    1st Qu.:1.000
                    1st Qu.:1.000
                                     1st Qu.:0.0000
                                                       1st Qu.:0.0000
    Median :2.000
                    Median :1.000
                                     Median :1.0000
                                                       Median :1.0000
    Mean
          :1.938
                    Mean
                          :1.446
                                             :0.7156
                                                               :0.6955
##
                                     Mean
                                                       Mean
##
    3rd Qu.:2.000
                    3rd Qu.:2.000
                                     3rd Qu.:1.0000
                                                       3rd Qu.:1.0000
##
    Max.
          :8.000
                    Max.
                            :6.000
                                     Max.
                                             :1.0000
                                                       Max.
                                                               :1.0000
##
##
    smoking_allowed
                     wheelchair_access electric_vehicle_charge comes_furnished
##
    Min.
           :0.0000
                      Min.
                             :0.00000
                                        Min.
                                                :0.00000
                                                                  Min.
                                                                         :0.00000
    1st Qu.:0.0000
##
                      1st Qu.:0.00000
                                        1st Qu.:0.00000
                                                                  1st Qu.:0.00000
##
    Median :1.0000
                     Median :0.00000
                                        Median :0.00000
                                                                  Median :0.00000
##
    Mean
          :0.7386
                      Mean
                             :0.07788
                                        Mean
                                                :0.01363
                                                                  Mean
                                                                         :0.04606
##
    3rd Qu.:1.0000
                      3rd Qu.:0.00000
                                        3rd Qu.:0.00000
                                                                  3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.00000
                                        Max.
                                                :1.00000
                                                                  Max.
                                                                         :1.00000
##
##
              laundry_options
                                                             image url
                                           parking_options
##
                       : 9828
                                                   :16964
                                                            Length: 47087
                       : 4728
    laundry in bldg
                                off-street parking:15692
                                                            Class : character
                                carport
    laundry on site
                       : 6709
                                                   : 5166
                                                            Mode :character
```

```
no laundry on site: 387
                                attached garage
                                                   : 4747
##
                       : 9168
    w/d hookups
                                detached garage
                                                   : 2313
##
    w/d in unit
                       :16267
                                street parking
                                                    : 1850
##
                                 (Other)
                                                       355
##
    description
                             lat
                                                long
                                                                  state
    Length: 47087
##
                                 2.202
                                                  :-163.89
                                                                      : 5815
                        Min.
                               :
                                           Min.
                                                              fl
                                           1st Qu.:-101.29
                        1st Qu.: 33.491
    Class : character
                                                                      : 5796
                                                              ca
##
    Mode :character
                        Median: 37.974
                                           Median : -86.26
                                                              nc
                                                                      : 3383
##
                        Mean
                               : 37.184
                                           Mean
                                                  : -92.18
                                                              mi
                                                                      : 2593
##
                        3rd Qu.: 41.176
                                           3rd Qu.: -81.26
                                                              ga
                                                                      : 2532
##
                        Max.
                               :102.036
                                           Max.
                                                  : 94.16
                                                                      : 1988
                                                              CO
##
                                                              (Other):24980
##
                            bba
                                                 apb
                                                                   b_a
         ppa
                              :0.0004167
##
          : 0.1250
                       Min.
                                            Min.
                                                   : 200.0
                                                              Min.
                                                                         200
    1st Qu.: 0.8821
                       1st Qu.:0.0015373
                                            1st Qu.: 445.8
                                                              1st Qu.:
                                                                         807
##
    Median: 1.1187
                       Median: 0.0019048
                                            Median: 525.0
                                                              Median: 1904
          : 1.2794
                                                   : 559.2
##
    Mean
                       Mean
                              :0.0019315
                                            Mean
                                                              Mean
                                                                    : 2152
    3rd Qu.: 1.5105
                       3rd Qu.:0.0022434
                                            3rd Qu.: 650.5
                                                              3rd Qu.: 2500
##
    Max.
          :10.3713
                              :0.0050000
                                                   :2400.0
                       Max.
                                            Max.
                                                              Max.
                                                                     :21272
##
##
         b_b
                          apbth
##
           : 1.000
                      Min.
                             :
                                200
                      1st Qu.: 775
    1st Qu.: 1.000
##
    Median : 2.000
                      Median: 1037
##
##
    Mean
          : 3.074
                      Mean
                            : 1563
    3rd Qu.: 4.000
                      3rd Qu.: 2200
##
    Max.
           :36.000
                            :15000
                      Max.
##
library(reshape2)
x=housing[,c(5,7:15,20,21,23:28)]
summary(x)
```

```
sqfeet
                                        beds
                                                        baths
                                                                      cats allowed
        price
          : 151
                           : 200
                                                                            :0.0000
##
    Min.
                   Min.
                                   Min.
                                          :1.000
                                                    Min.
                                                           :1.000
                                                                    Min.
##
    1st Qu.: 825
                   1st Qu.: 770
                                   1st Qu.:1.000
                                                    1st Qu.:1.000
                                                                     1st Qu.:0.0000
##
    Median:1068
                   Median: 960
                                   Median :2.000
                                                                     Median :1.0000
                                                    Median :1.000
    Mean :1214
                   Mean :1003
                                   Mean
                                         :1.938
                                                    Mean
                                                           :1.446
                                                                     Mean
                                                                            :0.7156
    3rd Qu.:1450
                   3rd Qu.:1158
                                   3rd Qu.:2.000
                                                    3rd Qu.:2.000
                                                                     3rd Qu.:1.0000
##
##
    Max.
           :4519
                   Max.
                           :3023
                                   Max.
                                           :8.000
                                                    Max.
                                                           :6.000
                                                                     Max.
                                                                            :1.0000
##
     dogs_allowed
                      smoking_allowed wheelchair_access electric_vehicle_charge
    Min.
           :0.0000
                     Min.
                            :0.0000
                                       Min.
                                               :0.00000
                                                          Min.
                                                                  :0.00000
##
    1st Qu.:0.0000
                      1st Qu.:0.0000
                                       1st Qu.:0.00000
                                                          1st Qu.:0.00000
##
    Median :1.0000
                      Median :1.0000
                                       Median :0.00000
                                                          Median :0.00000
##
    Mean
           :0.6955
                      Mean
                             :0.7386
                                       Mean
                                               :0.07788
                                                          Mean
                                                                  :0.01363
##
    3rd Qu.:1.0000
                      3rd Qu.:1.0000
                                       3rd Qu.:0.00000
                                                          3rd Qu.:0.00000
##
    Max.
           :1.0000
                      Max.
                             :1.0000
                                       Max.
                                               :1.00000
                                                          Max.
                                                                  :1.00000
                                              long
##
    comes_furnished
                            lat
                                                                 ppa
    Min.
##
           :0.00000
                                                 :-163.89
                      Min.
                              : 2.202
                                         Min.
                                                            Min.
                                                                    : 0.1250
                       1st Qu.: 33.491
                                                            1st Qu.: 0.8821
##
    1st Qu.:0.00000
                                          1st Qu.:-101.29
##
    Median :0.00000
                      Median: 37.974
                                          Median : -86.26
                                                            Median: 1.1187
                            : 37.184
##
    Mean
           :0.04606
                      Mean
                                          Mean
                                               : -92.18
                                                            Mean
                                                                    : 1.2794
    3rd Qu.:0.00000
                       3rd Qu.: 41.176
                                          3rd Qu.: -81.26
                                                            3rd Qu.: 1.5105
   Max.
                                               : 94.16
                                         Max.
##
          :1.00000
                      Max. :102.036
                                                            Max.
                                                                    :10.3713
```

```
##
        bba
                                                            b_b
                            apb
                                            b_a
                                      Min. : 200 Min. : 1.000
          :0.0004167
                       Min. : 200.0
## Min.
   1st Qu.:0.0015373
                       1st Qu.: 445.8
                                      1st Qu.: 807
                                                       1st Qu.: 1.000
## Median :0.0019048
                       Median : 525.0
                                                       Median : 2.000
                                      Median : 1904
##
   Mean :0.0019315
                       Mean : 559.2
                                       Mean : 2152
                                                       Mean : 3.074
##
   3rd Qu.:0.0022434
                       3rd Qu.: 650.5
                                       3rd Qu.: 2500
                                                       3rd Qu.: 4.000
  Max.
         :0.0050000
                       Max. :2400.0
                                      Max. :21272
                                                       Max.
                                                              :36.000
##
##
       apbth
## Min.
         : 200
##
  1st Qu.: 775
## Median: 1037
## Mean : 1563
## 3rd Qu.: 2200
## Max. :15000
#calculate correlation coefficients, rounded to 2 decimal places
cor_df <- round(cor(x), 2)</pre>
#melt the data frame
melted_cor <- melt(cor_df)</pre>
#view head of melted data frame
head(melted_cor)
##
            Var1 Var2 value
## 1
           price price 1.00
## 2
          sqfeet price 0.33
            beds price 0.17
## 3
## 4
           baths price 0.24
## 5 cats_allowed price -0.04
## 6 dogs_allowed price -0.03
library(ggplot2)
#create correlation heatmap
ggplot(data = melted_cor, aes(x=Var1, y=Var2, fill=value)) +
 geom_tile() +
 geom_text(aes(Var2, Var1, label = value), size = 3) +
 scale_fill_gradient2(low = "blue", high = "red",
                      limit = c(-1,1), name="Correlation") +
 theme(axis.title.x = element_blank(),
       axis.title.y = element_blank(),
       panel.background = element_blank())
```

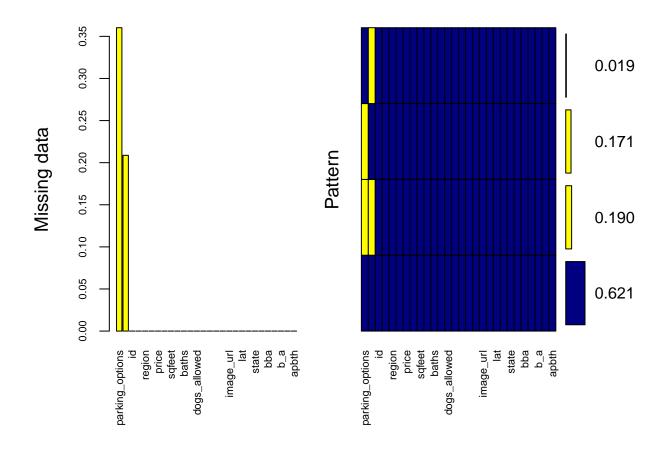
```
apbth -0.310.860.720.880.040.040.05.030.010.060.060.040.25.090.14.840.92 1
                  b_b-<mark>0.210.730.850.89</mark>0.050.040.02.02 0 0.070.060-0.270.4-0.42.84 1 0.92
                  b_a = 0.250.920.910.6+0.040.040.040.040.030.020.32.290.311 0.840.84
                  apb -0.11+0.020.640.26.020.020.060.050.050.020.040.050.130.921 -0.340.420.14
                  bba --0.130 0.620.240.050.06.020.050.040.070.050.050.131 --0.920.290.40.09
                  ppa -<mark>0.72</mark>0.330.340.18 0 -0.020.100.090.130.050.040.271 -0.10.130.320.270.25
                                                                                             Correlation
                 long --0.28.050.02 0 0.050.040.110.040.060.020.081 --0.270.06.050.02 0 0.01
                                                                                                  1.0
                   lat -0.08.010.040.120.040.050.130.02.02 0 1-0.080.04.050.04.030.060.06
                                                                                                  0.5
     comes_furnished -0.030.020.030.050.070.060.15.140.08 1 0 -0.02.050.070.020.050.070.06
electric_vehicle_charge -0.140.040.020.010.060.06-0.10.21 1 0.080.020.06.130.040.050.01 0 0.01
                                                                                                  0.0
    wheelchair_access -0.09 0 -0.08.040.120.13-0.2 1 0.210.140.020.00.090.060.060.000.020.03
                                                                                                  -0.5
     smoking_allowed -0.180.060.030.00.030.03.11 -0.2-0.40.150.10.110.020.030.050.050.020.05
                                                                                                  -1.0
         cats_allowed --0.040.070.070.011 0.890.030.120.060.070.00.05 0 -0.05.020.080.050.05
                baths -0.240.6 0.6 1 -0.00.030.00.040.010.050.12 0 -0.18.240.260.60.890.88
                 beds -0.170.76 1 0.6-0.070.030.030.030.040.020.340.620.640.910.850.72
                sqfeet -0.33 1 0.760.6-0.070.030.060 0.010.020.010.050.330-0.020.730.86
                        1 0.330.170.240.040.030.18.090.140.030.030.28.720.18.110.250.210.31
                      pricsqfeetedsatacsglsoutdaglanthoveducses.chiartosdgppabbaapbb_ab_lapbth
```

```
h1 <- replace(housing, housing=='', NA)
#x=apply(h1,MARGIN = 2,is.na)
#summary(x)
#x=NULL
housing=h1
h1=NULL
summary(housing)</pre>
```

```
##
          id
                              url
                                                                   region
            :7.005e+09
                         Length: 47087
                                              omaha / council bluffs:
##
    Min.
                                                                         507
    1st Qu.:7.036e+09
                         Class : character
                                              inland empire
                                                                         491
                                                                         478
    Median :7.043e+09
                         Mode : character
                                              denver
           :7.041e+09
                                                                         474
##
    Mean
                                              orlando
##
    3rd Qu.:7.048e+09
                                              daytona beach
                                                                         468
    Max.
           :7.051e+09
                                              fayetteville, NC
                                                                         468
##
##
                                              (Other)
                                                                      :44201
##
     region url
                             price
                                                   type
                                                                    sqfeet
##
    Length: 47087
                              : 151
                                                      :38849
                                                                       : 200
                        Min.
                                         apartment
                                                               Min.
    Class : character
                        1st Qu.: 825
                                        house
                                                      : 4087
                                                               1st Qu.: 770
##
    Mode : character
                        Median:1068
                                         townhouse
                                                      : 1818
                                                               Median: 960
##
                        Mean
                                :1214
                                         condo
                                                         858
                                                               Mean
                                                                       :1003
##
                        3rd Qu.:1450
                                         duplex
                                                         654
                                                               3rd Qu.:1158
##
                        Max.
                                :4519
                                        manufactured:
                                                         557
                                                               Max.
                                                                       :3023
##
                                         (Other)
                                                         264
##
                         baths
                                       cats_allowed
                                                          dogs_allowed
         beds
```

```
Min.
          :1.000
                   Min.
                          :1.000
                                    Min.
                                           :0.0000
                                                     Min.
                                                            :0.0000
   1st Qu.:1.000
                   1st Qu.:1.000
                                    1st Qu.:0.0000
                                                    1st Qu.:0.0000
  Median :2.000
                   Median :1.000
                                    Median :1.0000
                                                     Median :1.0000
##
  Mean
         :1.938
                   Mean :1.446
                                    Mean
                                          :0.7156
                                                     Mean
                                                            :0.6955
   3rd Qu.:2.000
                    3rd Qu.:2.000
                                    3rd Qu.:1.0000
                                                     3rd Qu.:1.0000
##
   Max. :8.000
                          :6.000
                                          :1.0000
                                                     Max.
                                                            :1.0000
                   Max.
                                    Max.
##
##
   smoking_allowed wheelchair_access electric_vehicle_charge comes_furnished
##
   Min.
          :0.0000
                     Min.
                            :0.00000
                                      Min.
                                              :0.00000
                                                              Min.
                                                                      :0.00000
##
   1st Qu.:0.0000
                     1st Qu.:0.00000
                                       1st Qu.:0.00000
                                                               1st Qu.:0.00000
   Median :1.0000
                     Median :0.00000
                                       Median :0.00000
                                                               Median :0.00000
##
   Mean
         :0.7386
                     Mean
                            :0.07788
                                       Mean
                                              :0.01363
                                                               Mean
                                                                      :0.04606
##
   3rd Qu.:1.0000
                     3rd Qu.:0.00000
                                       3rd Qu.:0.00000
                                                               3rd Qu.:0.00000
##
   Max. :1.0000
                           :1.00000
                                                               Max.
                     Max.
                                       Max.
                                              :1.00000
                                                                      :1.00000
##
##
              laundry_options
                                         parking_options
                                                           image_url
##
                           0
                              off-street parking:15692
                                                          Length: 47087
                     :
##
   laundry in bldg
                     : 4728
                               carport
                                                : 5166
                                                          Class : character
                                                          Mode : character
   laundry on site
                     : 6709
                              attached garage
                                               : 4747
   no laundry on site: 387
                               detached garage : 2313
##
  w/d hookups
                     : 9168
                               street parking
                                                 : 1850
  w/d in unit
                     :16267
                               (Other)
                                                 : 355
## NA's
                              NA's
                      : 9828
                                                 :16964
##
   description
                            lat
                                              long
                                                               state
## Length:47087
                      \mathtt{Min}.
                             : 2.202
                                        \mathtt{Min}.
                                                :-163.89
                                                           fl
                                                                  : 5815
  Class : character
                      1st Qu.: 33.491
                                         1st Qu.:-101.29
                                                           ca
                                                                  : 5796
##
  Mode :character
                       Median : 37.974
                                        Median : -86.26
                                                                  : 3383
                                                           nc
##
                             : 37.184
                       Mean
                                         Mean
                                                : -92.18
                                                           mi
                                                                  : 2593
##
                       3rd Qu.: 41.176
                                         3rd Qu.: -81.26
                                                                  : 2532
                              :102.036
##
                       Max.
                                         Max.
                                               : 94.16
                                                                 : 1988
                                                           СО
##
                                                           (Other):24980
##
                           bba
                                               apb
                                                                b_a
        ppa
   Min. : 0.1250
                             :0.0004167
                                          Min. : 200.0
                                                           Min.
                                                                  : 200
   1st Qu.: 0.8821
                      1st Qu.:0.0015373
                                          1st Qu.: 445.8
                                                           1st Qu.: 807
##
   Median : 1.1187
                     Median :0.0019048
                                          Median : 525.0
                                                           Median: 1904
##
   Mean
          : 1.2794
                     Mean
                            :0.0019315
                                         Mean : 559.2
                                                           Mean : 2152
                                          3rd Qu.: 650.5
   3rd Qu.: 1.5105
                     3rd Qu.:0.0022434
                                                           3rd Qu.: 2500
##
   Max.
          :10.3713
                     Max.
                             :0.0050000
                                          Max.
                                                :2400.0
                                                           Max. :21272
##
##
        b_b
                         apbth
   Min. : 1.000
                     Min. : 200
##
   1st Qu.: 1.000
                     1st Qu.: 775
  Median : 2.000
                     Median: 1037
##
  Mean
         : 3.074
                     Mean : 1563
   3rd Qu.: 4.000
                     3rd Qu.: 2200
## Max.
          :36.000
                     Max.
                           :15000
##
library(mice)
##
## Attaching package: 'mice'
## The following object is masked from 'package:stats':
```

```
##
##
       filter
## The following objects are masked from 'package:base':
##
##
       cbind, rbind
library(VIM)
## Loading required package: colorspace
## Loading required package: grid
## VIM is ready to use.
## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues
##
## Attaching package: 'VIM'
## The following object is masked from 'package:datasets':
##
##
       sleep
mice_plot <- aggr(housing, col=c('navyblue', 'yellow'),</pre>
                    numbers=TRUE, sortVars=TRUE,
                    labels=names(housing), cex.axis=.7,
                    gap=3, ylab=c("Missing data", "Pattern"))
```



```
Variables sorted by number of missings:
##
##
                   Variable
                                 Count
##
            parking_options 0.3602693
##
            laundry_options 0.2087200
##
                          id 0.0000000
                         url 0.0000000
##
##
                      region 0.0000000
                 region_url 0.0000000
##
##
                       price 0.0000000
##
                        type 0.0000000
                      sqfeet 0.0000000
##
##
                       beds 0.0000000
##
                       baths 0.0000000
               cats_allowed 0.0000000
##
               dogs_allowed 0.0000000
##
            smoking_allowed 0.0000000
##
##
          wheelchair_access 0.0000000
##
    electric_vehicle_charge 0.0000000
            comes_furnished 0.0000000
##
##
                   image_url 0.0000000
                description 0.0000000
##
##
                         lat 0.0000000
##
                        long 0.0000000
##
                       state 0.0000000
##
                         ppa 0.0000000
```

##

```
bba 0.0000000
##
##
                          apb 0.0000000
##
                          b a 0.0000000
##
                          b_b 0.0000000
##
                        apbth 0.0000000
imputed_Data <- mice(housing,m=1, maxit=1, method = "cart")</pre>
##
##
    iter imp variable
##
         1 laundry_options parking_options
## Warning: Number of logged events: 6
summary(imputed_Data)
## Class: mids
## Number of multiple imputations: 1
## Imputation methods:
##
                          id
                                                   url
                                                                          region
                                                     11 11
                                                                               11 11
##
##
                 region_url
                                                 price
                                                                             type
##
                          11 11
##
                      sqfeet
                                                  beds
                                                                            baths
##
##
               cats_allowed
                                         dogs_allowed
                                                                 smoking_allowed
##
##
         wheelchair_access electric_vehicle_charge
                                                                 comes_furnished
##
##
            laundry_options
                                      parking_options
                                                                       image_url
                      "cart"
##
                                                "cart"
##
                description
                                                   lat
                                                                             long
##
##
                                                                              bba
                       state
                                                   ppa
                                                                               11 11
##
##
                                                   b_a
                                                                              b_b
                         apb
                                                    11 11
                                                                               11 11
                          11 11
##
##
                       apbth
##
## PredictorMatrix:
##
               id url region region_url price type sqfeet beds baths cats_allowed
## id
                0
                     0
                            1
                                        0
                                               1
                                                     1
                                                            1
                                                                  1
                                                                        1
                                                                                       1
## url
                1
                     0
                            1
                                        0
                                                    1
                                                                  1
                                                                                       1
                                               1
                                                            1
                            0
                                        0
                    0
                                               1
                                                    1
                                                                                       1
## region
                1
                                                            1
## region_url 1
                    0
                            1
                                        0
                                               1
                                                    1
                                                            1
                                                                                      1
                                        0
                    0
                                               0
## price
                1
                            1
                                                    1
                                                                  1
                                                                        1
                                                                                       1
                                                            1
                            1
                                        0
                                               1
                                                                                       1
## type
                1
##
               dogs_allowed smoking_allowed wheelchair_access
## id
                           1
                                             1
## url
                           1
                                             1
                                                                 1
                           1
                                             1
                                                                 1
## region
                           1
                                             1
## region_url
```

```
## price
                          1
                                           1
                                                               1
## type
                          1
                                           1
               electric_vehicle_charge comes_furnished laundry_options
##
## id
                                      1
## url
                                      1
                                                       1
                                                                        1
## region
                                      1
                                                       1
                                                                        1
## region_url
                                      1
                                                       1
                                                                        1
## price
                                      1
                                                       1
                                                                         1
## type
                                      1
                                                                         1
##
               parking_options image_url description lat long state ppa bba apb b_a
## id
                             1
                                        0
                                                     0
                                                         1
                                                               1
                                                                     1
                              1
                                        0
                                                     0
                                                         1
                                                               1
                                                                     1
                                                                                      1
## url
                                                     0
                                                               1
## region
                             1
                                        0
                                                         1
                                                                     1
                                                                         1
                                                     0
                                                         1
                                                               1
## region_url
                              1
                                        0
                                                                     1
                                                                                      1
## price
                             1
                                        0
                                                     0
                                                         1
                                                               1
                                                                              1
                                                                                      1
                                                                     1
                                                                         1
                                                                                  1
## type
                                        0
                                                     0
                                                         1
                                                               1
                                                                     1
                                                                         1
                                                                              1
                                                                                  1
##
               b_b apbth
## id
                 1
## url
                       1
                 1
## region
                       1
## region_url
                 1
                       1
## price
                 1
## type
                 1
                       1
## Number of logged events:
     it im
                        dep
                                 meth
## 1 0
                            constant
## 2
     0
        0
                            constant
## 3
     0
                            constant
## 4
     0
        0
                             constant
     1
        1 laundry_options
                                 cart
     1 1 parking_options
                                 cart
##
## 1
## 2
## 3
## 4
## 5
                       regionashtabula, regionbend, regiondayton / springfield, regioneast oregon, region
## 6 regionashtabula, regionbend, regioncentral louisiana, regiondayton / springfield, regiondelaware,
completedData <- complete(imputed_Data,1)</pre>
housing=completedData
completedData=imputed_Data=NULL
summary(housing)
##
          id
                             url
                                                                  region
##
           :7.005e+09
                         Length: 47087
   Min.
                                              omaha / council bluffs:
                                                                        507
    1st Qu.:7.036e+09
                         Class : character
                                              inland empire
                                                                        491
   Median :7.043e+09
##
                         Mode :character
                                              denver
                                                                         478
    Mean
           :7.041e+09
                                              orlando
                                                                         474
    3rd Qu.:7.048e+09
##
                                             daytona beach
                                                                        468
##
    Max.
           :7.051e+09
                                             fayetteville, NC
                                                                        468
```

apartment

price

: 151

Min.

(Other)

type

:38849

:44201

: 200

sqfeet

 $\mathtt{Min}.$ 

##

##

##

region\_url

Length: 47087

```
Class :character
                       1st Qu.: 825
                                      house
                                                   : 4087
                                                            1st Qu.: 770
   Mode :character
                       Median:1068
                                      townhouse
                                                   : 1818
                                                            Median: 960
##
                             :1214
                                       condo
                                                      858
                                                            Mean :1003
                       Mean
##
                       3rd Qu.:1450
                                                      654
                                                            3rd Qu.:1158
                                       duplex
##
                       Max.
                              :4519
                                      manufactured:
                                                      557
                                                            Max.
                                                                   :3023
##
                                       (Other)
                                                      264
##
                        baths
                                      cats allowed
         beds
                                                       dogs allowed
                         :1.000
                                                             :0.0000
   Min.
                                    Min.
                                           :0.0000
##
           :1.000
                    Min.
                                                      Min.
##
    1st Qu.:1.000
                    1st Qu.:1.000
                                     1st Qu.:0.0000
                                                      1st Qu.:0.0000
##
   Median :2.000
                    Median :1.000
                                    Median :1.0000
                                                      Median :1.0000
   Mean
         :1.938
                    Mean :1.446
                                     Mean
                                           :0.7156
                                                      Mean
                                                             :0.6955
   3rd Qu.:2.000
                    3rd Qu.:2.000
                                     3rd Qu.:1.0000
##
                                                      3rd Qu.:1.0000
                                            :1.0000
          :8.000
                           :6.000
##
   Max.
                    Max.
                                     Max.
                                                      Max.
                                                             :1.0000
##
##
   smoking_allowed
                     wheelchair_access electric_vehicle_charge comes_furnished
##
   Min.
           :0.0000
                     Min.
                            :0.00000
                                       Min.
                                               :0.00000
                                                                Min.
                                                                       :0.00000
##
   1st Qu.:0.0000
                     1st Qu.:0.00000
                                        1st Qu.:0.00000
                                                                1st Qu.:0.00000
   Median :1.0000
                     Median :0.00000
##
                                       Median :0.00000
                                                                Median :0.00000
##
   Mean
         :0.7386
                     Mean
                            :0.07788
                                       Mean
                                               :0.01363
                                                                Mean
                                                                       :0.04606
##
   3rd Qu.:1.0000
                     3rd Qu.:0.00000
                                        3rd Qu.:0.00000
                                                                3rd Qu.:0.00000
##
   Max.
          :1.0000
                     Max.
                            :1.00000
                                       Max.
                                               :1.00000
                                                                Max.
                                                                       :1.00000
##
##
              laundry_options
                                          parking_options
                                                            image_url
##
                               off-street parking:26096
                                                           Length: 47087
                      :
##
                                                           Class : character
   laundry in bldg
                     : 5869
                                carport
                                                  : 7141
   laundry on site
                      : 8213
                               attached garage
                                                 : 6886
                                                           Mode : character
##
   no laundry on site: 492
                               detached garage
                                                 : 3351
##
   w/d hookups
                      :11956
                                street parking
                                                  : 3010
##
   w/d in unit
                      :20557
                               no parking
                                                     571
##
                                (Other)
                                                  :
                                                      32
##
   description
                            lat
                                               long
                                                                 state
##
   Length: 47087
                       Min.
                              : 2.202
                                          Min.
                                                 :-163.89
                                                            fl
                                                                    : 5815
##
                       1st Qu.: 33.491
                                          1st Qu.:-101.29
                                                                    : 5796
   Class : character
   Mode :character
##
                       Median: 37.974
                                          Median : -86.26
                                                                    : 3383
                                                            nc
                             : 37.184
##
                       Mean
                                          Mean
                                                 : -92.18
                                                            mi
                                                                   : 2593
##
                       3rd Qu.: 41.176
                                          3rd Qu.: -81.26
                                                                   : 2532
                                                            ga
##
                       Max.
                              :102.036
                                          Max. : 94.16
                                                                    : 1988
##
                                                            (Other):24980
##
                           bba
                                                                 b a
         ppa
                                                apb
##
   Min. : 0.1250
                             :0.0004167
                                           Min. : 200.0
                                                                  : 200
                      Min.
                                                            Min.
   1st Qu.: 0.8821
                      1st Qu.:0.0015373
                                           1st Qu.: 445.8
                                                            1st Qu.: 807
   Median : 1.1187
##
                      Median :0.0019048
                                           Median : 525.0
                                                            Median: 1904
   Mean : 1.2794
                      Mean
                             :0.0019315
                                           Mean : 559.2
                                                            Mean : 2152
##
   3rd Qu.: 1.5105
                      3rd Qu.:0.0022434
                                           3rd Qu.: 650.5
                                                            3rd Qu.: 2500
##
   Max.
           :10.3713
                      Max.
                             :0.0050000
                                           Max.
                                                  :2400.0
                                                            Max.
                                                                   :21272
##
##
         b_b
                         apbth
##
          : 1.000
                            : 200
   Min.
                     Min.
   1st Qu.: 1.000
                     1st Qu.: 775
   Median : 2.000
                     Median: 1037
##
##
   Mean
          : 3.074
                     Mean
                            : 1563
   3rd Qu.: 4.000
##
                     3rd Qu.: 2200
##
   Max.
           :36.000
                     Max.
                            :15000
```

##

```
x=write.csv(housing,file="housing_cleaned_l.csv", row.names=FALSE)
h1=housing
h1=h1[h1$lat<68.13,]
h1=h1[h1$lat>19.47,]
h1=h1[h1$long> (-171.69),]
h1=h1[h1$long< (-66.57),]
h1=h1[!(h1$id == 7026190479 | h1$id == 7022766169 | h1$id == 7030114325 | h1$id == 7028422754 | h1$id =
dim(housing)
## [1] 47087
                28
dim(h1)
## [1] 47038
                28
housing=h1
h1=NULL
summary(housing)
```

```
##
          id
                            url
                                                              region
                       Length: 47038
##
  Min.
           :7.005e+09
                                           omaha / council bluffs:
                                                                    507
   1st Qu.:7.036e+09
                        Class : character
                                           inland empire
                                                                    491
                                                                    478
##
   Median :7.043e+09
                       Mode :character
                                           denver
                                                                   474
##
   Mean
          :7.041e+09
                                           orlando
   3rd Qu.:7.048e+09
                                           daytona beach
                                                                    468
##
  Max.
          :7.051e+09
                                          fayetteville, NC
                                                                   468
##
                                           (Other)
                                                                 :44152
##
    region_url
                          price
                                                               sqfeet
                                                type
##
  Length: 47038
                      Min. : 151
                                      apartment
                                                 :38805
                                                           Min.
                                                                : 200
                      1st Qu.: 825
                                                           1st Qu.: 770
   Class : character
                                                  : 4083
##
                                     house
##
   Mode :character
                      Median:1070
                                     townhouse
                                                 : 1818
                                                           Median: 959
##
                      Mean :1214
                                     condo
                                                    857
                                                           Mean :1003
                                                  :
##
                      3rd Qu.:1450
                                     duplex
                                                 :
                                                    654
                                                           3rd Qu.:1158
##
                       Max.
                              :4519
                                     manufactured:
                                                    557
                                                                 :3023
                                                           Max.
                                                    264
##
                                      (Other)
##
         beds
                       baths
                                     cats_allowed
                                                     dogs_allowed
##
          :1.000
                   Min.
                          :1.000
                                   Min.
                                          :0.0000
                                                    Min.
                                                            :0.0000
   Min.
                   1st Qu.:1.000
##
   1st Qu.:1.000
                                    1st Qu.:0.0000
                                                    1st Qu.:0.0000
##
   Median :2.000
                   Median :1.000
                                   Median :1.0000
                                                    Median :1.0000
##
   Mean
         :1.938
                   Mean :1.445
                                   Mean
                                         :0.7154
                                                    Mean
                                                            :0.6952
   3rd Qu.:2.000
                   3rd Qu.:2.000
                                    3rd Qu.:1.0000
                                                    3rd Qu.:1.0000
##
##
   Max.
         :8.000
                   Max.
                          :6.000
                                   Max.
                                         :1.0000
                                                    Max.
                                                            :1.0000
##
##
  smoking_allowed wheelchair_access electric_vehicle_charge comes_furnished
          :0.0000
                                             :0.00000
## Min.
                    Min.
                            :0.00000
                                      Min.
                                                              Min.
                                                                     :0.00000
   1st Qu.:0.0000
                    1st Qu.:0.00000
                                      1st Qu.:0.00000
                                                               1st Qu.:0.00000
##
## Median :1.0000
                    Median :0.00000
                                      Median :0.00000
                                                              Median :0.00000
## Mean :0.7384
                    Mean :0.07796
                                      Mean :0.01365
                                                              Mean :0.04605
## 3rd Qu.:1.0000
                    3rd Qu.:0.00000
                                      3rd Qu.:0.00000
                                                              3rd Qu.:0.00000
```

```
:1.0000
                     Max.
                             :1.00000
                                               :1.00000
                                                                        :1.00000
##
    Max.
                                        Max.
                                                                 Max.
##
              laundry_options
                                          parking_options
##
                                                            image url
##
                                off-street parking:26072
                                                            Length: 47038
                      :
                           0
##
    laundry in bldg
                      : 5866
                                carport
                                                  : 7139
                                                            Class : character
                      : 8210
                                                            Mode : character
##
    laundry on site
                                attached garage
                                                  : 6874
    no laundry on site: 492
                                detached garage
                                                  : 3341
##
    w/d hookups
                      :11950
                                street parking
                                                  : 3009
##
    w/d in unit
                      :20520
                                no parking
                                                     571
##
                                (Other)
                                                      32
##
   description
                            lat
                                             long
                                                               state
                                                                  : 5815
##
   Length: 47038
                                               :-159.50
                       Min.
                              :19.59
                                        Min.
                                                          fl
##
    Class :character
                       1st Qu.:33.49
                                        1st Qu.:-101.30
                                                                  : 5796
                                                          ca
                                                                  : 3383
##
   Mode :character
                       Median :37.97
                                        Median : -86.27
                                                          nc
                                                                  : 2593
##
                       Mean
                              :37.19
                                               : -92.34
                                        Mean
                                                          mi
##
                       3rd Qu.:41.18
                                        3rd Qu.: -81.27
                                                                  : 2532
                                                          ga
##
                              :64.95
                                               : -70.05
                                                                  : 1988
                       Max.
                                        Max.
                                                          CO
##
                                                           (Other):24931
##
         ppa
                           bba
                                                apb
                                                                 b_a
##
    Min.
          : 0.1250
                      Min.
                              :0.0004167
                                           Min.
                                                  : 200.0
                                                            Min.
                                                                    :
                                                                       200.0
                                           1st Qu.: 445.5
##
    1st Qu.: 0.8824
                      1st Qu.:0.0015373
                                                            1st Qu.: 806.2
    Median : 1.1190
                      Median :0.0019048
                                           Median : 525.0
                                                            Median: 1904.0
                                           Mean : 559.3
    Mean
         : 1.2796
                                                            Mean : 2152.4
##
                      Mean
                              :0.0019315
    3rd Qu.: 1.5111
                      3rd Qu.:0.0022447
                                           3rd Qu.: 650.5
                                                            3rd Qu.: 2500.0
##
         :10.3713
##
    Max.
                      Max.
                              :0.0050000
                                           Max.
                                                :2400.0
                                                            Max.
                                                                   :21272.0
##
##
         b_b
                         apbth
          : 1.000
                               200
##
    Min.
                     Min.
                            :
##
    1st Qu.: 1.000
                     1st Qu.:
                               775
##
  Median : 2.000
                     Median: 1036
##
    Mean
         : 3.073
                     Mean : 1563
##
    3rd Qu.: 4.000
                     3rd Qu.: 2200
##
    Max.
         :36.000
                     Max. :15000
##
```

#### #unique(housing\$state)

```
x=write.csv(housing, file="housing_cleaned_1.csv", row.names=FALSE)
```

```
housing$type=as.numeric(housing$type)
housing$state=as.numeric(housing$state)
housing$laundry_options=as.numeric(housing$laundry_options)
housing$parking_options=as.numeric(housing$parking_options)
housing$region=as.numeric(housing$region)
summary(housing)
```

```
##
          iд
                            url
                                                region
                                                             region_url
           :7.005e+09
                        Length: 47038
                                                            Length: 47038
  \mathtt{Min}.
                                            Min.
                                                  : 1.0
  1st Qu.:7.036e+09
                                            1st Qu.: 62.0
##
                        Class : character
                                                            Class : character
## Median :7.043e+09
                        Mode :character
                                            Median :143.0
                                                            Mode :character
## Mean
           :7.041e+09
                                            Mean
                                                  :143.3
  3rd Qu.:7.048e+09
                                            3rd Qu.:220.0
## Max. :7.051e+09
                                                  :299.0
                                            Max.
```

```
##
        price
                                         safeet
                                                         beds
                                                                        baths
                        type
          : 151
##
   Min.
                   Min.
                         : 1.000
                                           : 200
                                                    Min.
                                                           :1.000
                                                                    Min.
                                                                           :1.000
                                    Min.
                   1st Qu.: 1.000
                                     1st Qu.: 770
                                                                    1st Qu.:1.000
    1st Qu.: 825
                                                    1st Qu.:1.000
   Median:1070
                                    Median: 959
##
                   Median : 1.000
                                                    Median :2.000
                                                                    Median :1.000
##
   Mean :1214
                   Mean
                          : 2.189
                                    Mean
                                          :1003
                                                    Mean
                                                           :1.938
                                                                    Mean
                                                                           :1.445
##
    3rd Qu.:1450
                   3rd Qu.: 1.000
                                     3rd Qu.:1158
                                                    3rd Qu.:2.000
                                                                    3rd Qu.:2.000
##
   Max.
           :4519
                   Max. :12.000
                                    Max.
                                           :3023
                                                    Max.
                                                           :8.000
                                                                    Max.
                                                                           :6.000
                      dogs_allowed
                                       smoking_allowed wheelchair_access
##
     cats allowed
##
   Min.
           :0.0000
                     Min.
                            :0.0000
                                      Min.
                                             :0.0000
                                                        Min.
                                                               :0.00000
##
    1st Qu.:0.0000
                     1st Qu.:0.0000
                                      1st Qu.:0.0000
                                                        1st Qu.:0.00000
   Median :1.0000
                     Median :1.0000
                                      Median :1.0000
                                                        Median :0.00000
##
   Mean
          :0.7154
                     Mean
                            :0.6952
                                      Mean
                                              :0.7384
                                                        Mean
                                                               :0.07796
##
    3rd Qu.:1.0000
                     3rd Qu.:1.0000
                                      3rd Qu.:1.0000
                                                        3rd Qu.:0.00000
##
   Max.
          :1.0000
                     Max.
                            :1.0000
                                      Max.
                                              :1.0000
                                                        Max.
                                                               :1.00000
##
    electric_vehicle_charge comes_furnished
                                               laundry_options parking_options
##
   Min.
           :0.00000
                            Min.
                                   :0.00000
                                               Min.
                                                      :2.000
                                                               Min.
                                                                      :2.000
##
    1st Qu.:0.00000
                            1st Qu.:0.00000
                                               1st Qu.:3.000
                                                               1st Qu.:3.000
##
   Median :0.00000
                            Median :0.00000
                                               Median :5.000
                                                               Median :6.000
##
   Mean
          :0.01365
                            Mean
                                   :0.04605
                                               Mean
                                                     :4.703
                                                               Mean
                                                                      :4.871
##
    3rd Qu.:0.00000
                            3rd Qu.:0.00000
                                               3rd Qu.:6.000
                                                               3rd Qu.:6.000
           :1.00000
##
   Max.
                            Max.
                                   :1.00000
                                               Max.
                                                      :6.000
                                                               Max.
                                                                      :8.000
##
                       description
                                                lat
     image_url
                                                                long
##
   Length: 47038
                       Length: 47038
                                                  :19.59
                                                                  :-159.50
                                          Min.
                                                           Min.
   Class : character
                       Class : character
##
                                           1st Qu.:33.49
                                                           1st Qu.:-101.30
##
                       Mode :character
   Mode :character
                                          Median :37.97
                                                           Median: -86.27
##
                                           Mean
                                                 :37.19
                                                           Mean : -92.34
##
                                           3rd Qu.:41.18
                                                           3rd Qu.: -81.27
##
                                           Max.
                                                :64.95
                                                           Max.
                                                                  : -70.05
##
                                           bba
        state
                                                                apb
                         ppa
   Min. : 1.00
                    Min. : 0.1250
                                              :0.0004167
                                                                 : 200.0
                                      Min.
                                                           Min.
##
    1st Qu.: 8.00
                    1st Qu.: 0.8824
                                       1st Qu.:0.0015373
                                                           1st Qu.: 445.5
##
   Median :15.00
                    Median : 1.1190
                                      Median :0.0019048
                                                           Median: 525.0
##
   Mean
         :16.39
                    Mean
                         : 1.2796
                                      Mean
                                            :0.0019315
                                                           Mean
                                                                  : 559.3
##
    3rd Qu.:24.00
                    3rd Qu.: 1.5111
                                       3rd Qu.:0.0022447
                                                           3rd Qu.: 650.5
##
   Max.
         :36.00
                           :10.3713
                                             :0.0050000
                                                           Max.
                                                                  :2400.0
                    Max.
                                      Max.
##
        b a
                           b b
                                           apbth
##
          : 200.0
                      Min.
                           : 1.000
                                       Min. : 200
##
   1st Qu.: 806.2
                      1st Qu.: 1.000
                                       1st Qu.: 775
##
   Median: 1904.0
                      Median : 2.000
                                       Median: 1036
##
   Mean : 2152.4
                      Mean : 3.073
                                       Mean : 1563
                      3rd Qu.: 4.000
   3rd Qu.: 2500.0
                                       3rd Qu.: 2200
   Max. :21272.0
##
                      Max. :36.000
                                       Max.
                                              :15000
library(reshape2)
x=housing[,c(3,5:17,20:28)]
summary(x)
##
                                                         sqfeet
        region
                        price
                                                                         beds
                                         type
                                                                           :1.000
   Min. : 1.0
                    Min. : 151
                                   Min.
                                          : 1.000
                                                     Min. : 200
                                                                    Min.
   1st Qu.: 62.0
                    1st Qu.: 825
##
                                   1st Qu.: 1.000
                                                     1st Qu.: 770
                                                                    1st Qu.:1.000
##
   Median :143.0
                    Median:1070
                                   Median : 1.000
                                                     Median: 959
                                                                    Median :2.000
##
   Mean
         :143.3
                    Mean
                           :1214
                                   Mean : 2.189
                                                     Mean
                                                          :1003
                                                                    Mean
                                                                           :1.938
```

:12.000

3rd Qu.:1158

:3023

Max.

3rd Qu.:2.000

:8.000

Max.

3rd Qu.: 1.000

Max.

3rd Qu.:220.0

Max. :299.0

##

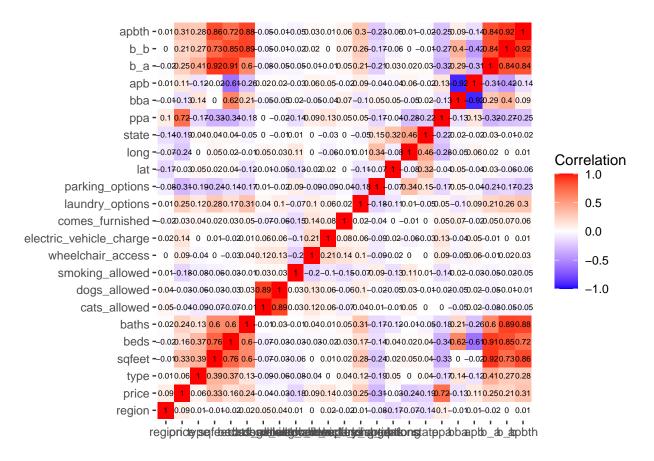
3rd Qu.:1450

:4519

Max.

```
##
       baths
                    cats allowed
                                    dogs_allowed
                                                    smoking allowed
          :1.000
                   Min.
                         :0.0000
                                   Min. :0.0000
                                                    Min. :0.0000
##
   Min.
                                   1st Qu.:0.0000
   1st Qu.:1.000
                   1st Qu.:0.0000
                                                    1st Qu.:0.0000
  Median :1.000
                   Median :1.0000
                                   Median :1.0000
                                                    Median :1.0000
   Mean :1.445
                   Mean :0.7154
                                   Mean
                                          :0.6952
                                                    Mean :0.7384
##
   3rd Qu.:2.000
                   3rd Qu.:1.0000
                                   3rd Qu.:1.0000
                                                    3rd Qu.:1.0000
   Max.
         :6.000
                   Max. :1.0000
                                   Max.
                                          :1.0000
                                                    Max.
                                                          :1.0000
##
   wheelchair_access electric_vehicle_charge comes_furnished
                                                              laundry_options
##
   Min.
          :0.00000
                    Min.
                            :0.00000
                                            Min.
                                                   :0.00000
                                                              Min.
                                                                    :2.000
##
   1st Qu.:0.00000
                    1st Qu.:0.00000
                                            1st Qu.:0.00000
                                                              1st Qu.:3.000
  Median :0.00000 Median :0.00000
                                            Median :0.00000
                                                              Median :5.000
## Mean :0.07796
                     Mean
                          :0.01365
                                            Mean
                                                   :0.04605
                                                              Mean
                                                                    :4.703
   3rd Qu.:0.00000
                     3rd Qu.:0.00000
                                            3rd Qu.:0.00000
                                                              3rd Qu.:6.000
##
  Max.
                                                   :1.00000
                                                                    :6.000
         :1.00000
                     Max. :1.00000
                                            Max.
                                                              Max.
   parking_options
                        lat
                                       long
                                                        state
##
   Min.
          :2.000
                   Min. :19.59
                                  Min.
                                         :-159.50
                                                    Min. : 1.00
##
   1st Qu.:3.000
                   1st Qu.:33.49
                                                    1st Qu.: 8.00
                                  1st Qu.:-101.30
##
  Median :6.000
                   Median :37.97
                                  Median : -86.27
                                                    Median :15.00
                                        : -92.34
  Mean :4.871
                   Mean :37.19
                                  Mean
                                                    Mean
                                                          :16.39
##
   3rd Qu.:6.000
                   3rd Qu.:41.18
                                  3rd Qu.: -81.27
                                                    3rd Qu.:24.00
##
   Max. :8.000
                   Max.
                         :64.95
                                  Max. : -70.05
                                                    Max.
                                                         :36.00
##
                          bba
        ppa
                                             apb
                                                              b_a
                     Min.
##
   Min. : 0.1250
                                                         Min. : 200.0
                            :0.0004167
                                        Min. : 200.0
   1st Qu.: 0.8824
                     1st Qu.:0.0015373
                                        1st Qu.: 445.5
                                                         1st Qu.: 806.2
##
                                        Median : 525.0
   Median : 1.1190
                    Median :0.0019048
                                                         Median: 1904.0
   Mean : 1.2796
                     Mean :0.0019315
                                        Mean : 559.3
                                                         Mean : 2152.4
##
   3rd Qu.: 1.5111
                     3rd Qu.:0.0022447
                                        3rd Qu.: 650.5
                                                         3rd Qu.: 2500.0
                     Max. :0.0050000
##
   Max. :10.3713
                                        Max. :2400.0
                                                         Max. :21272.0
##
        b_b
                        apbth
  Min. : 1.000
                    Min. : 200
                    1st Qu.: 775
##
  1st Qu.: 1.000
## Median : 2.000
                    Median: 1036
  Mean : 3.073
                    Mean : 1563
   3rd Qu.: 4.000
                    3rd Qu.: 2200
##
   Max. :36.000
                    Max. :15000
#calculate correlation coefficients, rounded to 2 decimal places
cor_df <- round(cor(x), 2)</pre>
#melt the data frame
melted_cor <- melt(cor_df)</pre>
#view head of melted data frame
head(melted cor)
##
      Var1 Var2 value
## 1 region region 1.00
## 2 price region 0.09
```

## 3 type region 0.01
## 4 sqfeet region -0.01
## 5 beds region -0.02
## 6 baths region 0.02



```
x=write.csv(housing,file="housing_cleaned_pred_1.csv", row.names=FALSE)
```