Mini Project 1

Firstly, we load our data.

```
## set options
options(scipen = 100)

## load packages
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
```

```
library(psych)
library(pracma)
```

```
##
## Attaching package: 'pracma'
```

```
## The following objects are masked from 'package:psych':
##
##     logit, polar
```

```
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following objects are masked from 'package:psych':
##
##     %+%, alpha
```

```
# load data
dat <- readRDS("house_data.RDS")

# check data
ls(dat)
```

```
##  [1] "AdjSalePrice"    "Bathrooms"       "Bedrooms"        "BldgGrade"
##  [5] "DocumentDate"    "ID"              "ImpsVal"         "LandVal"
##  [9] "NbrLivingUnits"  "NewConstruction" "PropertyID"      "PropertyType"
## [13] "SalePrice"       "SqFtFinBasement" "SqFtLot"         "SqFtTotLiving"
## [17] "TrafficNoise"    "ym"              "YrBuilt"         "YrRenovated"
## [21] "zhvi_idx"        "zhvi_px"         "ZipCode"
```

We found there are 23 variables in this data set.

And now we are going to do some data tidy, including these steps:

- drop all the non-numeric variables
- drop ID and Party ID
- drop the variable with low variance

```
dat <- dat[,sapply(dat, is.numeric)] # drop non-numeric
dat <- dat %>% # drop ID and party ID
  select(!c(ID, PropertyID))
dat %>% # drop low variance variable
  summarise(across(everything(), ~var(scale(., center = FALSE), na.rm = TRUE)))
```

```
##   SalePrice     zhvi_px    zhvi_idx AdjSalePrice NbrLivingUnits   SqFtLot
## 1 0.3177451 0.008517194 0.008517194    0.3173574     0.02399549 0.8591889
##   SqFtTotLiving SqFtFinBasement Bathrooms  Bedrooms  BldgGrade      YrBuilt
## 1      0.161738       0.6919151 0.1107266 0.0672611 0.02307366 0.0002364283
##   YrRenovated TrafficNoise  LandVal   ImpsVal         ZipCode
## 1   0.9488276    0.8788141 0.408023 0.3624847 0.0000002848018
```

Then we create a correlate coefficient matrix to check the relationship between different variables.

```
lowerCor(dat)
```

```
##                SlPrc zhv_p zhv_d AdjSP NbrLU SqFtL SqFTL SqFFB Bthrm Bdrms
## SalePrice       1.00
## zhvi_px         0.10  1.00
## zhvi_idx        0.10  1.00  1.00
## AdjSalePrice    0.99 -0.04 -0.04  1.00
## NbrLivingUnits  0.02  0.02  0.02  0.02  1.00
## SqFtLot         0.13 -0.01 -0.01  0.14 -0.01  1.00
## SqFtTotLiving   0.69 -0.02 -0.02  0.70  0.06  0.20  1.00
## SqFtFinBasement 0.30 -0.01 -0.01  0.30  0.07  0.04  0.41  1.00
## Bathrooms       0.52 -0.03 -0.03  0.53  0.11  0.11  0.76  0.27  1.00
## Bedrooms        0.31  0.00  0.00  0.31  0.17  0.07  0.60  0.31  0.54  1.00
## BldgGrade       0.67 -0.05 -0.05  0.68 -0.05  0.15  0.77  0.14  0.66  0.37
## YrBuilt         0.08 -0.02 -0.02  0.08 -0.11  0.07  0.31 -0.21  0.47  0.15
## YrRenovated     0.11  0.00  0.00  0.11  0.01  0.01  0.07  0.11  0.07  0.05
## TrafficNoise   -0.01  0.01  0.01 -0.01  0.08 -0.01 -0.06  0.05 -0.06 -0.04
## LandVal         0.80 -0.02 -0.02  0.81  0.03  0.07  0.47  0.28  0.31  0.20
## ImpsVal         0.83 -0.03 -0.03  0.83  0.00  0.11  0.76  0.28  0.63  0.35
## ZipCode        -0.07  0.00  0.00 -0.07  0.06 -0.12 -0.20  0.14 -0.20 -0.15
##                BldgG
## SalePrice
## zhvi_px
## zhvi_idx
## AdjSalePrice
## NbrLivingUnits
## SqFtLot
## SqFtTotLiving
## SqFtFinBasement
## Bathrooms
## Bedrooms
## BldgGrade       1.00
## YrBuilt         0.43
## YrRenovated     0.03
## TrafficNoise   -0.06
## LandVal         0.46
## ImpsVal         0.75
## ZipCode        -0.19
##              YrBlt YrRnv TrffN LndVl ImpsV ZipCd
## YrBuilt       1.00
## YrRenovated  -0.25  1.00
## TrafficNoise -0.15  0.02  1.00
## LandVal      -0.11  0.13  0.00  1.00
## ImpsVal       0.18  0.12 -0.01  0.57  1.00
## ZipCode      -0.38  0.08  0.10 -0.07 -0.03  1.00
```

Certificate if the variance and corr variance is equal. This step aims to make sure if the data is correct.

```
var(dat, na.rm = TRUE) == cov(dat, use = "complete.obs")
```

```
##                SalePrice zhvi_px zhvi_idx AdjSalePrice NbrLivingUnits SqFtLot
## SalePrice           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## zhvi_px             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## zhvi_idx            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## AdjSalePrice        TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## NbrLivingUnits      TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtLot             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtTotLiving       TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtFinBasement     TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## Bathrooms           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## Bedrooms            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## BldgGrade           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrBuilt             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrRenovated         TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## TrafficNoise        TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## LandVal             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## ImpsVal             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## ZipCode             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
##                SqFtTotLiving SqFtFinBasement Bathrooms Bedrooms BldgGrade
## SalePrice               TRUE            TRUE      TRUE     TRUE      TRUE
## zhvi_px                 TRUE            TRUE      TRUE     TRUE      TRUE
## zhvi_idx                TRUE            TRUE      TRUE     TRUE      TRUE
## AdjSalePrice            TRUE            TRUE      TRUE     TRUE      TRUE
## NbrLivingUnits          TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtLot                 TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtTotLiving           TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtFinBasement         TRUE            TRUE      TRUE     TRUE      TRUE
## Bathrooms               TRUE            TRUE      TRUE     TRUE      TRUE
## Bedrooms                TRUE            TRUE      TRUE     TRUE      TRUE
## BldgGrade               TRUE            TRUE      TRUE     TRUE      TRUE
## YrBuilt                 TRUE            TRUE      TRUE     TRUE      TRUE
## YrRenovated             TRUE            TRUE      TRUE     TRUE      TRUE
## TrafficNoise            TRUE            TRUE      TRUE     TRUE      TRUE
## LandVal                 TRUE            TRUE      TRUE     TRUE      TRUE
## ImpsVal                 TRUE            TRUE      TRUE     TRUE      TRUE
## ZipCode                 TRUE            TRUE      TRUE     TRUE      TRUE
##                YrBuilt YrRenovated TrafficNoise LandVal ImpsVal ZipCode
## SalePrice         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## zhvi_px           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## zhvi_idx          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## AdjSalePrice      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## NbrLivingUnits    TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtLot           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtTotLiving     TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtFinBasement   TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## Bathrooms         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## Bedrooms          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## BldgGrade         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrBuilt           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrRenovated       TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## TrafficNoise      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## LandVal           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## ImpsVal           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## ZipCode           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
```

```
Mcov <- var(dat, na.rm = TRUE)

# if we centre the data, we get the same result
c_dat <- scale(dat, center = TRUE, scale = FALSE)
Mcov_c <- var(c_dat, na.rm = TRUE)

near(Mcov_c, Mcov) # covariance matrix is same
```

```
##                SalePrice zhvi_px zhvi_idx AdjSalePrice NbrLivingUnits SqFtLot
## SalePrice          FALSE   FALSE     TRUE        FALSE           TRUE   FALSE
## zhvi_px            FALSE   FALSE     TRUE        FALSE           TRUE   FALSE
## zhvi_idx            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## AdjSalePrice       FALSE   FALSE     TRUE        FALSE           TRUE   FALSE
## NbrLivingUnits      TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtLot            FALSE   FALSE     TRUE        FALSE           TRUE   FALSE
## SqFtTotLiving       TRUE    TRUE     TRUE        FALSE           TRUE    TRUE
## SqFtFinBasement    FALSE    TRUE     TRUE        FALSE           TRUE    TRUE
## Bathrooms           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## Bedrooms            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## BldgGrade           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrBuilt             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrRenovated         TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## TrafficNoise        TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## LandVal             TRUE   FALSE     TRUE         TRUE           TRUE   FALSE
## ImpsVal             TRUE   FALSE     TRUE         TRUE           TRUE   FALSE
## ZipCode             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
##                SqFtTotLiving SqFtFinBasement Bathrooms Bedrooms BldgGrade
## SalePrice               TRUE           FALSE      TRUE     TRUE      TRUE
## zhvi_px                 TRUE            TRUE      TRUE     TRUE      TRUE
## zhvi_idx                TRUE            TRUE      TRUE     TRUE      TRUE
## AdjSalePrice           FALSE           FALSE      TRUE     TRUE      TRUE
## NbrLivingUnits          TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtLot                 TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtTotLiving           TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtFinBasement         TRUE            TRUE      TRUE     TRUE      TRUE
## Bathrooms               TRUE            TRUE      TRUE     TRUE      TRUE
## Bedrooms                TRUE            TRUE      TRUE     TRUE      TRUE
## BldgGrade               TRUE            TRUE      TRUE     TRUE      TRUE
## YrBuilt                 TRUE            TRUE      TRUE     TRUE      TRUE
## YrRenovated             TRUE            TRUE      TRUE     TRUE      TRUE
## TrafficNoise            TRUE            TRUE      TRUE     TRUE      TRUE
## LandVal                FALSE           FALSE      TRUE     TRUE      TRUE
## ImpsVal                FALSE           FALSE      TRUE     TRUE      TRUE
## ZipCode                 TRUE            TRUE      TRUE     TRUE      TRUE
##                YrBuilt YrRenovated TrafficNoise LandVal ImpsVal ZipCode
## SalePrice         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## zhvi_px           TRUE        TRUE         TRUE   FALSE   FALSE    TRUE
## zhvi_idx          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## AdjSalePrice      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## NbrLivingUnits    TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtLot           TRUE        TRUE         TRUE   FALSE   FALSE    TRUE
## SqFtTotLiving     TRUE        TRUE         TRUE   FALSE   FALSE    TRUE
## SqFtFinBasement   TRUE        TRUE         TRUE   FALSE   FALSE    TRUE
## Bathrooms         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## Bedrooms          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## BldgGrade         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrBuilt           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrRenovated       TRUE        TRUE         TRUE   FALSE    TRUE    TRUE
## TrafficNoise      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## LandVal           TRUE       FALSE         TRUE   FALSE   FALSE    TRUE
## ImpsVal           TRUE        TRUE         TRUE   FALSE   FALSE    TRUE
## ZipCode           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
```

Then we perform the eigen-decomposition of the covariance matrix and check the results. This step can divided a matrix into a eigen values and eigen vectors.

```
eigMcov <- eigen(Mcov)
eigMcov
```

```
## eigen() decomposition
## $values
##  [1] 326208688074.6403808593750  18235914108.6289024353027
##  [3]   7213699330.8674087524414   2907594354.4751238822937
##  [5]    819768275.0459656715393    265647813.2387284040451
##  [7]       358827.5509719259571       193944.6326406990411
##  [9]       139701.7964463515382         2380.4086192770487
## [11]          540.1920212516691            0.5116549668237
## [13]            0.3719691774605            0.2991274858044
## [15]            0.1831829428816            0.0231452300706
## [17]           -0.0000002413329
##
## $vectors
##                      [,1]              [,2]               [,3]              [,4]
##  [1,] -0.602048190758632 -0.08632353123778 -0.348452357792817 -0.54053693604472
##  [2,] -0.000564338517070 -0.00950159350215 -0.089000806375808 -0.61488790468656
##  [3,] -0.000000001296734 -0.00000002183270 -0.000000204505529 -0.00000141288581
##  [4,] -0.671371692965639 -0.08286364811084 -0.239609377807863  0.56099741735375
##  [5,] -0.000000006144448 -0.00000003859377 -0.000000008546372 -0.00000002763985
##  [6,] -0.006723293476451  0.00466043018045 -0.030830105151754  0.02224672094794
##  [7,] -0.001135948383779  0.00202124256211  0.000533277664980  0.00007708117987
##  [8,] -0.000234642301026  0.00000292508448  0.000491919926047 -0.00007472283429
##  [9,] -0.000000729351419  0.00000185276710  0.000000243518647  0.00000013819191
## [10,] -0.000000506493967  0.00000099188393  0.000000216660954 -0.00000014412608
## [11,] -0.000001427115810  0.00000261541785  0.000001065153434  0.00000047833112
## [12,] -0.000004410954214  0.00006509561867 -0.000045394841596  0.00001309590428
## [13,] -0.000088015559774 -0.00001345352307  0.000496762583082 -0.00011173299429
## [14,]  0.000000010813782 -0.00000005661127  0.000000127913354 -0.00000009564167
## [15,] -0.263765414693926 -0.60932551094283  0.741708017808869 -0.09122188614116
## [16,] -0.342314121099202  0.78376539565179  0.512020607188603 -0.07872871062598
## [17,]  0.000006183686681  0.00001792069680  0.000020563556390  0.00000010802241
##                      [,5]              [,6]               [,7]               [,8]
##  [1,] -0.00092500311640  0.4652803404063  0.000476541187148 -0.0003074456185496
##  [2,] -0.00544316693649 -0.7834992657177 -0.000437591208266  0.0004156446958949
##  [3,] -0.00000001250728 -0.0000018003200 -0.000000001005486  0.0000000009550612
##  [4,]  0.02692770685904 -0.4117496317030  0.000209517529386  0.0000428493239681
##  [5,]  0.00000004177443 -0.0000001245475 -0.000025062140123 -0.0000050602271672
##  [6,] -0.99921330114249 -0.0070690202978  0.003018729316630 -0.0002906679569835
##  [7,] -0.00323299670836  0.0005212248736 -0.928208062774202  0.0930290019685839
##  [8,] -0.00006246921838 -0.0001182740005 -0.371451710927692 -0.2612639639947784
##  [9,] -0.00000083597251  0.0000009084183 -0.000539875515137  0.0000409608569991
## [10,] -0.00000082617401  0.0000002357646 -0.000783085527838  0.0000000857460508
## [11,] -0.00000224710303  0.0000012329735 -0.000464248881369  0.0003040838349096
## [12,] -0.00004169032543  0.0000640169553 -0.009076776876378  0.0224489384445046
## [13,] -0.00001629049033 -0.0004226816314  0.010628450812498 -0.9603369999865978
## [14,]  0.00000017565267 -0.0000003538937  0.000050240745243 -0.0000588135497904
## [15,] -0.02594990125873 -0.0049027311426 -0.000272758510803  0.0005541544397628
## [16,] -0.01154539312026 -0.0055557473969  0.002235045160704  0.0001513859595040
## [17,]  0.00020510603568 -0.0000605486210  0.015565937524174 -0.0182990788949705
##                      [,9]               [,10]               [,11]
##  [1,]  0.0001605859903364  0.00002491936784812  0.00001601220645247
##  [2,] -0.0001696089785111 -0.00002970998039459 -0.00003207715237611
##  [3,] -0.0000000003897174 -0.00000000006829674 -0.00000000007396302
##  [4,]  0.0000486769090081 -0.00002657789099886 -0.00000902005944176
##  [5,]  0.0000054733425377  0.00027964584086260  0.00076541167174310
##  [6,]  0.0011243911186266  0.00010271048128154 -0.00003856132743092
##  [7,] -0.3584009689324924  0.03358636342809108  0.01336252181281341
##  [8,]  0.8899264399971795 -0.03850425258139365 -0.01768203388937795
##  [9,] -0.0002798272589561 -0.00132515557864803 -0.00791193281907092
## [10,] -0.0001600175428835 -0.00040427370362079  0.00204064618307306
## [11,] -0.0007107079906152 -0.00126841722462626 -0.00898079237252579
## [12,] -0.0248098862026032 -0.18747336017040794 -0.98157829401043206
## [13,] -0.2781666267929094 -0.00934492582032079 -0.01324433835270366
## [14,]  0.0001373777623987  0.00079790638674286  0.00203825629196005
## [15,] -0.0004387949233892  0.00002476411992577 -0.00005239734037122
## [16,]  0.0005902904363719 -0.00007526540226400  0.00001118365229877
## [17,]  0.0398118583643448  0.98089322969517778 -0.18890544010198079
##                      [,12]             [,13]              [,14]              [,15]
##  [1,]  0.00000002330388  0.00000002840513 -0.00000003785715 -0.00000007845501
##  [2,] -0.00000026998355  0.00000055065828  0.00000018431264  0.00000025009820
##  [3,]  0.00000073314919 -0.00000598048136 -0.00000753841920  0.00000251087786
##  [4,]  0.00000021655590 -0.00000013826861  0.00000003037894  0.00000002783848
##  [5,]  0.04173835214971 -0.01002457380511 -0.02528383441481  0.05748794849676
##  [6,]  0.00000199247595  0.00000029588147  0.00000013085065  0.00000052940552
##  [7,] -0.00073640461927 -0.00073874171442 -0.00007193205065 -0.00024158555124
##  [8,] -0.00022095224077  0.00028154519593  0.00007289430839 -0.00011132696663
##  [9,]  0.20730867888905  0.09676779395832 -0.04483561203595  0.97025166859262
## [10,]  0.95283982252208  0.20180149168765  0.02075959895647 -0.22436878008565
## [11,] -0.21763777294613  0.97312239856574  0.05370214442059 -0.04661422223617
## [12,]  0.00216721372513 -0.00912406113448 -0.00225772977126 -0.00786282170433
```

```
## [13,] -0.00003609868059 -0.00004708056327 -0.00001108996731 -0.00014139401553
## [14,] -0.00226846417160  0.05254051439851 -0.99701070142537 -0.05225500501467
## [15,]  0.00000025739419 -0.00000057809468 -0.00000009616703  0.00000007654168
## [16,]  0.00000094342987 -0.00000151677087  0.00000011618654 -0.00000052007679
## [17,]  0.00081169875269 -0.00027554945826  0.00040936332018 -0.00031608284324
##                   [,16]                   [,17]
##  [1,]  0.00000002626414  0.0000000000000000000
##  [2,] -0.00000008373433  0.0000029778966408858
##  [3,]  0.00000235974679 -0.9999999994485777588
##  [4,] -0.00000003810206  0.0000000000073452434
##  [5,]  0.99710210942838  0.0000027840628666864
##  [6,]  0.00000000286031  0.0000000000004087554
##  [7,] -0.00000506471354  0.0000000380183176336
##  [8,]  0.00002917137085 -0.0000000260595466882
##  [9,] -0.06477535813131  0.0000021945892905915 4
## [10,] -0.02439577087212 -0.0000012857208647739 5
## [11,]  0.02295026131719 -0.0000064470177690767 3
## [12,]  0.00101972727436  0.0000000558388704588 8
## [13,]  0.00001861667946  0.0000000000276130108 2
## [14,] -0.02164735694476  0.0000070177160875543 9
## [15,] -0.00000001270250  0.0000000000045330387 2
## [16,]  0.00000007464302  0.0000000000077571683 1
## [17,] -0.00013815304227 -0.0000000019625681616 8
```

Then we do another certificate step to make sure the stability of the raw data and process data. If these two are very similar, then we can conclude that there is no bias between our dealing process.

```
s_dat <- scale(dat)
Mcor <- var(s_dat, na.rm = TRUE)
near(cor(dat), Mcor)
```

```
##                SalePrice zhvi_px zhvi_idx AdjSalePrice NbrLivingUnits SqFtLot
## SalePrice           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## zhvi_px             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## zhvi_idx            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## AdjSalePrice        TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## NbrLivingUnits      TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtLot             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtTotLiving       TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## SqFtFinBasement     TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## Bathrooms           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## Bedrooms            TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## BldgGrade           TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrBuilt             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## YrRenovated         TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## TrafficNoise        TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## LandVal             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## ImpsVal             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
## ZipCode             TRUE    TRUE     TRUE         TRUE           TRUE    TRUE
##                SqFtTotLiving SqFtFinBasement Bathrooms Bedrooms BldgGrade
## SalePrice               TRUE            TRUE      TRUE     TRUE      TRUE
## zhvi_px                 TRUE            TRUE      TRUE     TRUE      TRUE
## zhvi_idx                TRUE            TRUE      TRUE     TRUE      TRUE
## AdjSalePrice            TRUE            TRUE      TRUE     TRUE      TRUE
## NbrLivingUnits          TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtLot                 TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtTotLiving           TRUE            TRUE      TRUE     TRUE      TRUE
## SqFtFinBasement         TRUE            TRUE      TRUE     TRUE      TRUE
## Bathrooms               TRUE            TRUE      TRUE     TRUE      TRUE
## Bedrooms                TRUE            TRUE      TRUE     TRUE      TRUE
## BldgGrade               TRUE            TRUE      TRUE     TRUE      TRUE
## YrBuilt                 TRUE            TRUE      TRUE     TRUE      TRUE
## YrRenovated             TRUE            TRUE      TRUE     TRUE      TRUE
## TrafficNoise            TRUE            TRUE      TRUE     TRUE      TRUE
## LandVal                 TRUE            TRUE      TRUE     TRUE      TRUE
## ImpsVal                 TRUE            TRUE      TRUE     TRUE      TRUE
## ZipCode                 TRUE            TRUE      TRUE     TRUE      TRUE
##                YrBuilt YrRenovated TrafficNoise LandVal ImpsVal ZipCode
## SalePrice         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## zhvi_px           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## zhvi_idx          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## AdjSalePrice      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## NbrLivingUnits    TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtLot           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtTotLiving     TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## SqFtFinBasement   TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## Bathrooms         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## Bedrooms          TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## BldgGrade         TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrBuilt           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## YrRenovated       TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## TrafficNoise      TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## LandVal           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## ImpsVal           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
## ZipCode           TRUE        TRUE         TRUE    TRUE    TRUE    TRUE
```

EigMcov and eigMcor are the results obtained after performing eigenvalue decomposition, corresponding to the covariance matrix Mcov and the correlation matrix Mcor, respectively.

```
eigMcor <- eigen(Mcor)
eigMcov
```

```
## eigen() decomposition
## $values
##  [1] 326208688074.6403808593750  18235914108.6289024353027
##  [3]   7213699330.8674087524414   2907594354.4751238822937
##  [5]    819768275.0459656715393    265647813.2387284040451
##  [7]       358827.5509719259571       193944.6326406990411
##  [9]       139701.7964463515382         2380.4086192770487
## [11]          540.1920212516691            0.5116549668237
## [13]            0.3719691774605            0.2991274858044
## [15]            0.1831829428816            0.0231452300706
## [17]           -0.0000002413329
##
## $vectors
##                      [,1]              [,2]               [,3]              [,4]
##  [1,] -0.602048190758632 -0.08632353123778 -0.348452357792817 -0.54053693604472
##  [2,] -0.000564338517070 -0.00950159350215 -0.089000806375808 -0.61488790468656
##  [3,] -0.000000001296734 -0.00000002183270 -0.000000204505529 -0.00000141288581
##  [4,] -0.671371692965639 -0.08286364811084 -0.239609377807863  0.56099741735375
##  [5,] -0.000000006144448 -0.00000003859377 -0.000000008546372 -0.00000002763985
##  [6,] -0.006723293476451  0.00466043018045 -0.030830105151754  0.02224672094794
##  [7,] -0.001135948383779  0.00202124256211  0.000533277664980  0.00007708117987
##  [8,] -0.000234642301026  0.00000292508448  0.000491919926047 -0.00007472283429
##  [9,] -0.000000729351419  0.00000185276710  0.000000243518647  0.00000013819191
## [10,] -0.000000506493967  0.00000099188393  0.000000216660954 -0.00000014412608
## [11,] -0.000001427115810  0.00000261541785  0.000001065153434  0.00000047833112
## [12,] -0.000004410954214  0.00006509561867 -0.000045394841596  0.00001309590428
## [13,] -0.000088015559774 -0.00001345352307  0.000496762583082 -0.00011173299429
## [14,]  0.000000010813782 -0.00000005661127  0.000000127913354 -0.00000009564167
## [15,] -0.263765414693926 -0.60932551094283  0.741708017808869 -0.09122188614116
## [16,] -0.342314121099202  0.78376539565179  0.512020607188603 -0.07872871062598
## [17,]  0.000006183686681  0.00001792069680  0.000020563556390  0.00000010802241
##                      [,5]              [,6]               [,7]               [,8]
##  [1,] -0.00092500311640  0.4652803404063  0.000476541187148 -0.0003074456185496
##  [2,] -0.00544316693649 -0.7834992657177 -0.000437591208266  0.0004156446958949
##  [3,] -0.00000001250728 -0.0000018003200 -0.000000001005486  0.0000000009550612
##  [4,]  0.02692770685904 -0.4117496317030  0.000209517529386  0.0000428493239681
##  [5,]  0.00000004177443 -0.0000001245475 -0.000025062140123 -0.0000050602271672
##  [6,] -0.99921330114249 -0.0070690202978  0.003018729316630 -0.0002906679569835
##  [7,] -0.00323299670836  0.0005212248736 -0.928208062774202  0.0930290019685839
##  [8,] -0.00006246921838 -0.0001182740005 -0.371451710927692 -0.2612639639947784
##  [9,] -0.00000083597251  0.0000009084183 -0.000539875515137  0.0000409608569991
## [10,] -0.00000082617401  0.0000002357646 -0.000783085527838  0.0000000857460508
## [11,] -0.00000224710303  0.0000012329735 -0.000464248881369  0.0003040838349096
## [12,] -0.00004169032543  0.0000640169553 -0.009076776876378  0.0224489384445046
## [13,] -0.00001629049033 -0.0004226816314  0.010628450812498 -0.9603369999865978
## [14,]  0.00000017565267 -0.0000003538937  0.000050240745243 -0.0000588135497904
## [15,] -0.02594990125873 -0.0049027311426 -0.000272758510803  0.0005541544397628
## [16,] -0.01154539312026 -0.0055557473969  0.002235045160704  0.0001513859595040
## [17,]  0.00020510603568 -0.0000605486210  0.015565937524174 -0.0182990788949705
##                      [,9]              [,10]               [,11]
##  [1,]  0.0001605859903364  0.00002491936784812  0.00001601220645247
##  [2,] -0.0001696089785111 -0.00002970998039459 -0.00003207715237611
##  [3,] -0.0000000003897174 -0.00000000006829674 -0.00000000007396302
##  [4,]  0.0000486769090081 -0.00002657789099886 -0.00000902005944176
##  [5,]  0.0000054733425377  0.00027964584086260  0.00076541167174310
##  [6,]  0.0011243911186266  0.00010271048128154 -0.00003856132743092
##  [7,] -0.3584009689324924  0.03358636342809108  0.01336252181281341
##  [8,]  0.8899264399971795 -0.03850425258139365 -0.01768203388937795
##  [9,] -0.0002798272589561 -0.00132515557864803 -0.00791193281907092
## [10,] -0.0001600175428835 -0.00040427370362079  0.00204064618307306
## [11,] -0.0007107079906152 -0.00126841722462626 -0.00898079237252579
## [12,] -0.0248098862026032 -0.18747336017040794 -0.98157829401043206
## [13,] -0.2781666267929094 -0.00934492582032079 -0.01324433835270366
## [14,]  0.0001373777623987  0.00079790638674286  0.00203825629196005
## [15,] -0.0004387949233892  0.00002476411992577 -0.00005239734037122
## [16,]  0.0005902904363719 -0.00007526540226400  0.00001118365229877
## [17,]  0.0398118583643448  0.98089322969517778 -0.18890544010198079
##                      [,12]             [,13]              [,14]              [,15]
##  [1,]  0.00000002330388  0.00000002840513 -0.00000003785715 -0.00000007845501
##  [2,] -0.00000026998355  0.00000055065828  0.00000018431264  0.00000025009820
##  [3,]  0.00000073314919 -0.00000598048136 -0.00000753841920  0.00000251087786
##  [4,]  0.00000021655590 -0.00000013826861  0.00000003037894  0.00000002783848
##  [5,]  0.04173835214971 -0.01002457380511 -0.02528383441481  0.05748794849676
##  [6,]  0.00000199247595  0.00000029588147  0.00000013085065  0.00000052940552
##  [7,] -0.00073640461927 -0.00073874171442 -0.00007193205065 -0.00024158555124
##  [8,] -0.00022095224077  0.00028154519593  0.00007289430839 -0.00011132696663
##  [9,]  0.20730867888905  0.09676779395832 -0.04483561203595  0.97025166859262
## [10,]  0.95283982252208  0.20180149168765  0.02075959895647 -0.22436878008565
## [11,] -0.21763777294613  0.97312239856574  0.05370214442059 -0.04661422223617
## [12,]  0.00216721372513 -0.00912406113448 -0.00225772977126 -0.00786282170433
```

```
## [13,] -0.00003609868059 -0.00004708056327 -0.00001108996731 -0.00014139401553
## [14,] -0.00226846417160  0.05254051439851 -0.99701070142537 -0.05225500501467
## [15,]  0.00000025739419 -0.00000057809468 -0.00000009616703  0.00000007654168
## [16,]  0.00000094342987 -0.00000151677087  0.00000011618654 -0.00000052007679
## [17,]  0.00081169875269 -0.00027554945826  0.00040936332018 -0.00031608284324
##                   [,16]                  [,17]
##  [1,]  0.00000002626414  0.0000000000000000000
##  [2,] -0.00000008373433  0.0000029778966408858
##  [3,]  0.00000235974679 -0.9999999994485777588
##  [4,] -0.00000003810206  0.0000000000073452434
##  [5,]  0.99710210942838  0.0000027840628666864
##  [6,]  0.00000000286031  0.0000000000004087554
##  [7,] -0.00000506471354  0.0000000380183176336
##  [8,]  0.00002917137085 -0.0000000260595466882
##  [9,] -0.06477535813131  0.0000021945892905915 4
## [10,] -0.02439577087212 -0.0000012857208647739 5
## [11,]  0.02295026131719 -0.0000064470177690767 3
## [12,]  0.00101972727436  0.0000000558388704588 8
## [13,]  0.00001861667946  0.0000000000276130108 2
## [14,] -0.02164735694476  0.0000070177160875543 9
## [15,] -0.00000001270250  0.0000000000045330387 2
## [16,]  0.00000007464302  0.0000000000077571683 1
## [17,] -0.00013815304227 -0.0000000019625681616 8
```

```
eigMcor
```

```
## eigen() decomposition
## $values
##  [1] 5.6018600693110149890686 2.0298722169467926512709 1.9191278147356185534989
##  [4] 1.3001771610606518425612 1.0007811293439694555474 0.9948462230660555238870
##  [7] 0.8986962933955856547286 0.8947928084106385737684 0.7771801729175958906026
## [10] 0.5203899109238445674919 0.3377168188234225820743 0.2545110391917222991687
## [13] 0.2232965768879534851354 0.1301758177684894657311 0.1120751402842977556551
## [16] 0.0045008069321234907972 0.0000000000000004451719
##
## $vectors
##              [,1]        [,2]        [,3]         [,4]        [,5]         [,6]
##  [1,] -0.37819612  0.13941611  0.12456037 -0.228851824  0.09289270 -0.012398619
##  [2,]  0.01014419  0.65682061 -0.25287326  0.002279683 -0.01328419  0.003540483
##  [3,]  0.01014419  0.65682061 -0.25287326  0.002279683 -0.01328419  0.003540483
##  [4,] -0.38092191  0.04728309  0.16287703 -0.229931578  0.09498655 -0.013345328
##  [5,] -0.02174748  0.05936637  0.10301880  0.517283126  0.32963208 -0.155170320
##  [6,] -0.07938282 -0.03836911 -0.07885330 -0.034137136 -0.22476482 -0.870968346
##  [7,] -0.37997752 -0.03382650 -0.08412441  0.177644281 -0.06523639  0.003972539
##  [8,] -0.16308421  0.10008297  0.28726574  0.365129387 -0.19697382  0.079455965
##  [9,] -0.32655611 -0.08159000 -0.19417930  0.261320820 -0.02163761  0.089528459
## [10,] -0.22902257 -0.02217832 -0.09000857  0.518929762 -0.13802451  0.050139929
## [11,] -0.35243459 -0.08052204 -0.15345851 -0.096070162  0.05876722  0.050793366
## [12,] -0.11913813 -0.20989614 -0.55401151 -0.029674582  0.11848495  0.093552656
## [13,] -0.04603084  0.09465634  0.27324059  0.050292090 -0.56560876 -0.093136061
## [14,]  0.02204315  0.07711263  0.19242687  0.134623207  0.63799612 -0.322880153
## [15,] -0.29659364  0.09562150  0.26191903 -0.276892786  0.07606004 -0.042102626
## [16,] -0.37391093  0.01393657  0.06494528 -0.107409838  0.06270892  0.075146715
## [17,]  0.07765466  0.13194914  0.40353856  0.067659928  0.07799660  0.261616866
##              [,7]         [,8]        [,9]        [,10]        [,11]
##  [1,]  0.11498710  0.005427198 -0.02473347  0.027903435 -0.079249906
##  [2,] -0.01986510  0.005500685  0.01533887 -0.006182737  0.005665891
##  [3,] -0.01986510  0.005500685  0.01533887 -0.006182737  0.005665891
##  [4,]  0.11850310  0.004427628 -0.02613710  0.029502274 -0.078364617
##  [5,]  0.65598015 -0.200945770  0.21141837 -0.227808805  0.052117951
##  [6,] -0.01586907  0.320435556  0.24848200  0.036137503 -0.069023266
##  [7,] -0.10043582  0.056945568  0.01475014  0.000849951  0.245786849
##  [8,] -0.22907385  0.419445952 -0.30993603 -0.555549357 -0.045327651
##  [9,] -0.11209166 -0.090829304  0.15174944 -0.141507917 -0.229039426
## [10,] -0.02500777  0.033159519 -0.20807072  0.706195099 -0.143919784
## [11,] -0.10582503 -0.067336687  0.19934242 -0.016279291  0.338313298
## [12,] -0.11905987 -0.071580116  0.17840893 -0.265933988 -0.508949857
## [13,] -0.14803522 -0.703142601  0.15019220 -0.107460570 -0.108783540
## [14,] -0.57060399 -0.281040599 -0.13977768  0.026968800 -0.040713508
## [15,]  0.24166176 -0.010198751 -0.26884447  0.056349186 -0.474681402
## [16,] -0.06330338 -0.023070086  0.19209903 -0.013745084  0.407549039
## [17,] -0.17420269  0.290890935  0.71158432  0.176280522 -0.256904774
##              [,12]         [,13]        [,14]         [,15]         [,16]
##  [1,]  0.104675320 -0.2460435550  0.003183568 -0.4137202039  0.7051725594
##  [2,] -0.009988429  0.0209079808  0.004828592  0.0236768063 -0.0499338203
##  [3,] -0.009988429  0.0209079808  0.004828592  0.0236768063 -0.0499338203
##  [4,]  0.105292018 -0.2479946234  0.002555553 -0.4147035811 -0.7055028973
##  [5,] -0.109091291 -0.0547072122  0.007601236  0.0109988910  0.0005188216
##  [6,]  0.013918051 -0.0008035467  0.059269846  0.0358659885  0.0002341412
##  [7,] -0.055871663  0.1067125236 -0.849449746  0.0391533519  0.0004517259
##  [8,] -0.162828987 -0.1166546522  0.163719696 -0.0186650433 -0.0002314837
##  [9,]  0.622914644  0.4780554325  0.184385823 -0.0730907954 -0.0006969019
## [10,] -0.129598350 -0.1869077876  0.177998363  0.0156794345 -0.0002465395
## [11,] -0.593428012  0.3993386340  0.342016274 -0.1780394994  0.0008380123
## [12,] -0.254591352 -0.3871896626 -0.060373911  0.1090401086 -0.0016427036
## [13,] -0.104864165 -0.0966451283 -0.021036602 -0.0100636592  0.0004070034
## [14,] -0.013144158 -0.0074923718 -0.020012310 -0.0008543371  0.0001435488
## [15,] -0.192931956  0.3562773383 -0.052981143  0.4730905271  0.0015549659
## [16,]  0.239210629 -0.3730105123  0.226325107  0.6159147257  0.0016991545
## [17,] -0.092854255  0.0277858868 -0.082596752  0.0017651955  0.0007282313
##                            [,17]
##  [1,] -0.0000000000000027410768339
##  [2,]  0.7071067811865562324769O
##  [3,] -0.7071067811865385799308l
##  [4,]  0.0000000000000026766244976
##  [5,]  0.0000000000000000042411278
##  [6,]  0.0000000000000000033711381
##  [7,]  0.0000000000000000038458335
##  [8,]  0.0000000000000000016009621
##  [9,]  0.0000000000000000032825389
## [10,] -0.0000000000000000028914300
## [11,]  0.0000000000000000007075445
## [12,]  0.0000000000000000123258007
## [13,] -0.0000000000000000018174601
## [14,] -0.0000000000000000052479437
## [15,]  0.0000000000000000327976823
```

```
## [16,]  0.00000000000000198348722
## [17,]  0.00000000000000012802749
```

Then we can calculate the loading matrix to find the relationship between origin variable and PCs. Use Sigma to confirm we need six PCs, and use loading^2 to check which variable is important in each PC.

By using loading ^ 2, we can decide to keep what variables in PCs.

```
loadings <- t(t(eigMcor$vectors)*sqrt(eigMcor$values))
row.names(loadings)<-colnames(dat)
colnames(loadings)<-paste0("PC",seq(1:ncol(loadings)))
loadings <- as.data.frame(loadings)
loadings ^ 2
```

```
##                         PC1          PC2        PC3            PC4
## SalePrice       0.8012469783 0.0394543236 0.02977582 0.068094382933
## zhvi_px         0.0005764576 0.8757138976 0.12271841 0.000006756959
## zhvi_idx        0.0005764576 0.8757138976 0.12271841 0.000006756959
## AdjSalePrice    0.8128383038 0.0045381662 0.05091240 0.068738456243
## NbrLivingUnits  0.0026494171 0.0071540116 0.02036746 0.347903786591
## SqFtLot         0.0353008645 0.0029883540 0.01193284 0.001515153749
## SqFtTotLiving   0.8088128713 0.0023226446 0.01358151 0.041030328561
## SqFtFinBasement 0.1489896533 0.0203324179 0.15836951 0.173338929462
## Bathrooms       0.5973761603 0.0135127135 0.07236186 0.088787240364
## Bedrooms        0.2938250534 0.0009984496 0.01554790 0.350122234749
## BldgGrade       0.6958078330 0.0131612822 0.04519453 0.011999953982
## YrBuilt         0.0795122125 0.0894288381 0.58903551 0.001144911066
## YrRenovated     0.0118694345 0.0181872949 0.14328289 0.003288530729
## TrafficNoise    0.0027219459 0.0120703460 0.07106166 0.023563641006
## LandVal         0.4927832397 0.0185600793 0.13165520 0.099684081961
## ImpsVal         0.7831925870 0.0003942581 0.00809467 0.014999979131
## ZipCode         0.0337805995 0.0353412423 0.31251724 0.005952036615
##                         PC5           PC6          PC7           PC8
## SalePrice       0.0086357950 0.00015293349 0.0118825930 0.00002635565
## zhvi_px         0.0001766076 0.00001247041 0.0003546457 0.00002707422
## zhvi_idx        0.0001766076 0.00001247041 0.0003546457 0.00002707422
## AdjSalePrice    0.0090294932 0.00017717990 0.0126203790 0.00001754142
## NbrLivingUnits  0.1087421861 0.02395373650 0.3867179663 0.03613101994
## SqFtLot         0.0505586872 0.75467627771 0.0002263165 0.09187638192
## SqFtTotLiving   0.0042591106 0.00001569973 0.0090654673 0.00290163211
## SqFtFinBasement 0.0388289918 0.00628071329 0.0471589359 0.15742528907
## Bathrooms       0.0004685517 0.00797403573 0.0112917052 0.00738200710
## Bedrooms        0.0190656462 0.00250105584 0.0005620346 0.00098387276
## BldgGrade       0.0034562838 0.00256666947 0.0100644433 0.00405719586
## YrBuilt         0.0140496499 0.00870699303 0.0127392469 0.00458466160
## YrRenovated     0.3201631676 0.00862962032 0.0196944127 0.44239408109
## TrafficNoise    0.4073569997 0.10371430380 0.2926055529 0.07067415259
## LandVal         0.0057896488 0.00176349535 0.0524842306 0.00009307144
## ImpsVal         0.0039354804 0.00561792531 0.0036013617 0.00047623457
## ZipCode         0.0060882222 0.06809064277 0.0272723562 0.07571516284
##                         PC9            PC10          PC11           PC12
## SalePrice       0.0004754357 0.0004051764641 0.00212104655 0.00278865777
## zhvi_px         0.0001828557 0.0000198925493 0.00001084149 0.00002539224
## zhvi_idx        0.0001828557 0.0000198925493 0.00001084149 0.00002539224
## AdjSalePrice    0.0005309290 0.0004529391518 0.00207392345 0.00282161348
## NbrLivingUnits  0.0347381886 0.0270065980411 0.00091733372 0.00302891289
## SqFtLot         0.0479856737 0.0006795871268 0.00160895426 0.00004930188
## SqFtTotLiving   0.0001690886 0.0000003759384 0.02040186986 0.00079449252
## SqFtFinBasement 0.0746561944 0.1606105857511 0.00069387159 0.00674792221
## Bathrooms       0.0178968225 0.0104205429135 0.01771630636 0.09875604867
## Bedrooms        0.0336467916 0.2595244423645 0.00699509611 0.00427469932
## BldgGrade       0.0308831191 0.0001379112922 0.03865367835 0.08962779461
## YrBuilt         0.0247374479 0.0368024355504 0.08747877305 0.01649658004
## YrRenovated     0.0175313938 0.0060093451534 0.00399649307 0.00279872889
## TrafficNoise    0.0151843915 0.0003784879882 0.00055979593 0.00004397159
## LandVal         0.0561725230 0.0016523580825 0.07609517541 0.00947359816
## ImpsVal         0.0286795332 0.0000983158761 0.05609348684 0.01456356075
## ZipCode         0.3935269288 0.0161710241312 0.02228933128 0.00219437193
##                          PC13           PC14            PC15
## SalePrice       0.0135178011039 0.0000013193455 0.01918327493805
## zhvi_px         0.0000976126830 0.0000030350881 0.00006282833268
## zhvi_idx        0.0000976126830 0.0000030350881 0.00006282833268
## AdjSalePrice    0.0137330371907 0.0000008501587 0.01927457729169
## NbrLivingUnits  0.0006682996503 0.0000075214008 0.00001355835766
## SqFtLot         0.0000001441798 0.0004572965300 0.00014417000114
## SqFtTotLiving   0.0025428037699 0.0939302971727 0.00017180950461
## SqFtFinBasement 0.0030386885643 0.0034892506769 0.00003904516778
## Bathrooms       0.0510315290205 0.0044257346107 0.00059873502956
## Bedrooms        0.0078007589720 0.0041244147532 0.00002755307536
## BldgGrade       0.0356094053702 0.0152273334500 0.00355256489449
## YrBuilt         0.0334756927436 0.0004744920427 0.00133254487098
## YrRenovated     0.0020856527354 0.0000576078297 0.00001135066040
## TrafficNoise    0.0000125348951 0.0000521344468 0.00000008180274
## LandVal         0.0283438253714 0.0003654037231 0.02508405793950
## ImpsVal         0.0310687806026 0.0066680029375 0.04251581086832
## ZipCode         0.0001723973521 0.0008880885140 0.00000034921665
##                          PC16
## SalePrice       0.0022381087853359
## zhvi_px         0.0000112222508495
## zhvi_idx        0.0000112222508497
## AdjSalePrice    0.0022402061595203
## NbrLivingUnits  0.0000000012115086
## SqFtLot         0.0000000002467436
## SqFtTotLiving   0.0000000009184182
```

```
## SqFtFinBasement 0.0000000002411744
## Bathrooms       0.0000000021859170
## Bedrooms        0.0000000002735668
## BldgGrade       0.0000000031607571
## YrBuilt         0.0000000121453159
## YrRenovated     0.0000000007455667
## TrafficNoise    0.0000000000927448
## LandVal         0.0000000108825865
## ImpsVal         0.0000000129943971
## ZipCode         0.0000000023868714
##                                                                                  PC17
## SalePrice       0.00000000000000000000000000000000000000000033448001325323
## zhvi_px         0.00000000000000222585955207743336895328058248284820877
## zhvi_idx        0.00000000000000222585955207732194235041811456533380211
## AdjSalePrice    0.00000000000000000000000000000000000000000031893534430761
## NbrLivingUnits  0.00000000000000000000000000000000000000000000080073805
## SqFtLot         0.00000000000000000000000000000000000000000000050591884
## SqFtTotLiving   0.00000000000000000000000000000000000000000000065842864
## SqFtFinBasement 0.00000000000000000000000000000000000000000000011410110
## Bathrooms       0.00000000000000000000000000000000000000000000047967547
## Bedrooms        0.00000000000000000000000000000000000000000000037218008
## BldgGrade       0.00000000000000000000000000000000000000000000002228616
## YrBuilt         0.00000000000000000000000000000000000000000000676329038
## YrRenovated     0.00000000000000000000000000000000000000000000014704746
## TrafficNoise    0.00000000000000000000000000000000000000000000122604409
## LandVal         0.00000000000000000000000000000000000000000004788660663
## ImpsVal         0.00000000000000000000000000000000000000000001751404916
## ZipCode         0.00000000000000000000000000000000000000000000007296830
```

```
(sigma <- round(eigMcor$values/ncol(dat),2))
```

```
##  [1] 0.33 0.12 0.11 0.08 0.06 0.06 0.05 0.05 0.05 0.03 0.02 0.01 0.01 0.01 0.01
## [16] 0.00 0.00
```

```
(sigmaSum <- cumsum(sigma))
```

```
##  [1] 0.33 0.45 0.56 0.64 0.70 0.76 0.81 0.86 0.91 0.94 0.96 0.97 0.98 0.99 1.00
## [16] 1.00 1.00
```

So we use these 9 variables to do the second PCA.

PC1: Sale Price, Adj Sale Price, SqFtTotLiving, ImpsVal

PC2: Bedrooms

PC3: TrafficNoise

PC4: YrRenovated

PC5: SqFtLot

PC6: NbrLibingUnits

```
selected_variables <- list(
  PC1 = c("SalePrice", "AdjSalePrice", "SqFtTotLiving", "ImpsVal"),
  PC2 = c("Bedrooms"),
  PC3 = c("TrafficNoise"),
  PC4 = c("YrRenovated"),
  PC5 = c("SqFtLot"),
  PC6 = c("NbrLivingUnits")
)
selected_data <- dat[, unlist(selected_variables)]
pca_mod_second <- pca(selected_data, nfactors = ncol(selected_data), covar = FALSE, rotate = "none")

pca_mod_second$loadings
```

```
##
## Loadings:
##                PC1    PC2    PC3    PC4    PC5    PC6    PC7    PC8    PC9
## SalePrice       0.926 -0.126  0.142        -0.125 -0.190  0.207
## AdjSalePrice    0.930 -0.127  0.142        -0.122 -0.189  0.195
## SqFtTotLiving   0.879        -0.171                0.194 -0.313 -0.236
## ImpsVal         0.908         0.105                      -0.232  0.293
## Bedrooms        0.543  0.418 -0.371 -0.178  0.138  0.542  0.203
## TrafficNoise           0.437  0.623  0.468 -0.334  0.298
## YrRenovated     0.145         0.584 -0.236  0.761
## SqFtLot         0.209        -0.284  0.813  0.449
## NbrLivingUnits         0.839 -0.129               -0.517
##
##                  PC1   PC2   PC3   PC4   PC5   PC6   PC7   PC8   PC9
## SS loadings    3.685 1.126 1.044 0.972 0.955 0.769 0.276 0.158 0.014
## Proportion Var 0.409 0.125 0.116 0.108 0.106 0.085 0.031 0.018 0.002
## Cumulative Var 0.409 0.535 0.651 0.759 0.865 0.950 0.981 0.998 1.000
```

So we can get a new PCA results, which can suggest the variables in each PC more clearly. And the cumulative sigma value suggests that we should choose PC1 - PC5 as our results.

PC1: SalePrice, AdjSalePrice, SqFtTotLiving, ImpsVal

PC2: NbrLibingUnits

PC3: TrafficNoise

PC4: SqFtLot

PC5: YrRenovated

So with the second PCA, we finally made 6 PC into 5 PC, and make each PC’s variable become more clear.