pls

2022-03-30

```
setwd("/cloud/project/housing crisis/pls_new")
rent.size.kde.df <- read.csv("to team.csv")
names(rent.size.kde.df)
```

```
##  [1] "X"                         "Effective_Rent_SF"
##  [3] "Latitude"                  "Longitude"
##  [5] "City"                      "Style"
##  [7] "Number_Of_Units"           "Transit_Time"
##  [9] "Drive_Time"                "Parking_Ratio"
## [11] "Bed_Count"                 "Size_"
## [13] "Distance_Beach"            "Distance_Bike"
## [15] "FB_All"                    "FB_Youth"
## [17] "Noise"                     "Flood"
## [19] "Distance_Beach_Inv"        "Distance_Bike_Inv"
## [21] "Transit_Time_Inv"          "Drive_Time_Inv"
## [23] "PC1"                       "PC2"
## [25] "PC3"                       "PC4"
## [27] "BIKE_800"                  "BUS_800"
## [29] "BUS_1600"                  "STORE_800"
## [31] "STORE_1600"                "STORE_4800"
## [33] "REST_800"                  "REST_1600"
## [35] "REST_4800"                 "GYM_800"
## [37] "GYM_1600"                  "GYM_4800"
## [39] "PARK_800"                  "PARK_1600"
## [41] "PARK_4800"                 "Distance_HW"
## [43] "ROAD_800"                  "Distance_HW_Inv"
## [45] "Amenities_AC"              "Amenities_Safty"
## [47] "Amenities_Pool"            "Amenities_Indoor_Gather"
## [49] "Amenities_Entertainment"   "Amenities_Outdoor_Gather"
## [51] "Amenities_Gym"             "Amenities_EV"
## [53] "Amenities_Service"
```

```
library(dplyr)
#no size
rent.size.kde.short.stand.pdf = rent.size.kde.df %>% select (Effective_Rent_SF, Style, Number_Of_Units,
                                                             GYM_4800, PARK_800, PARK_1600, PARK_4800, Dis
names(rent.size.kde.short.stand.pdf)
```

```
##  [1] "Effective_Rent_SF"         "Style"
##  [3] "Number_Of_Units"           "Transit_Time"
##  [5] "Drive_Time"                "Parking_Ratio"
##  [7] "Bed_Count"                 "Size_"
##  [9] "Amenities_AC"              "Amenities_Safty"
## [11] "Amenities_Pool"            "Amenities_Indoor_Gather"
## [13] "Amenities_Entertainment"   "Amenities_Outdoor_Gather"
## [15] "Amenities_Gym"             "Amenities_EV"
```

```
## [17] "Amenities_Service"        "Distance_Beach"
## [19] "Distance_Bike"            "FB_All"
## [21] "FB_Youth"                 "Noise"
## [23] "Flood"                    "Distance_Beach_Inv"
## [25] "Distance_Bike_Inv"        "Transit_Time_Inv"
## [27] "Drive_Time_Inv"           "BIKE_800"
## [29] "BUS_800"                  "BUS_1600"
## [31] "STORE_800"                "STORE_1600"
## [33] "STORE_4800"               "REST_800"
## [35] "REST_1600"                "REST_4800"
## [37] "GYM_800"                  "GYM_1600"
## [39] "GYM_4800"                 "PARK_800"
## [41] "PARK_1600"                "PARK_4800"
## [43] "Distance_HW"              "ROAD_800"
## [45] "Distance_HW_Inv"
```

```
install.packages("pls")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
install.packages("micOmics")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
## Warning: package 'micOmics' is not available for this version of R
##
## A version of this package for your version of R might be available elsewhere,
## see the ideas at
## https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages
```

```
library(pls)
```

```
##
## Attaching package: 'pls'
```

```
## The following object is masked from 'package:stats':
##
##     loadings
```

```
rent.size.kde.short.stand.pdf.pls <- plsr(rent.size.kde.short.stand.pdf$Effective_Rent_SF~.,ncomp = 10,
```

```
summary(rent.size.kde.short.stand.pdf.pls)
```

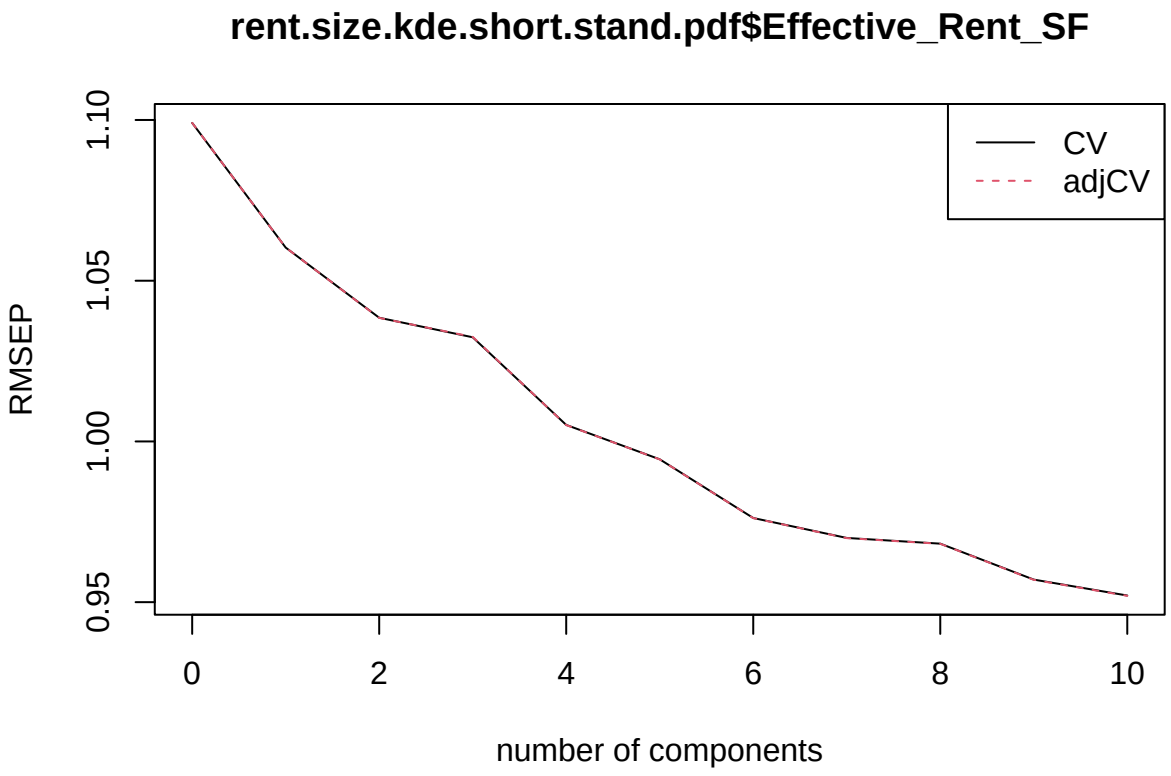
```
## Data:    X dimension: 281 46
##  Y dimension: 281 1
## Fit method: kernelpls
## Number of components considered: 10
##
## VALIDATION: RMSEP
## Cross-validated using 281 leave-one-out segments.
##        (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
## CV           1.099     1.06    1.038    1.032    1.005   0.9944   0.9762
## adjCV        1.099     1.06    1.038    1.033    1.005   0.9944   0.9761
##        7 comps  8 comps  9 comps  10 comps
## CV      0.9700   0.9682   0.9570    0.9521
## adjCV   0.9699   0.9682   0.9569    0.9520
```

```
##
## TRAINING: % variance explained
##                                                 1 comps  2 comps  3 comps
## X                                                 56.82    68.27    90.04
## rent.size.kde.short.stand.pdf$Effective_Rent_SF    8.13    14.01    16.47
##                                                 4 comps  5 comps  6 comps
## X                                                 98.49    99.67    99.87
## rent.size.kde.short.stand.pdf$Effective_Rent_SF   20.70    21.87    25.05
##                                                 7 comps  8 comps  9 comps
## X                                                100.00   100.00   100.00
## rent.size.kde.short.stand.pdf$Effective_Rent_SF   26.03    28.11    31.36
##                                                 10 comps
## X                                                 100.00
## rent.size.kde.short.stand.pdf$Effective_Rent_SF    33.22
```

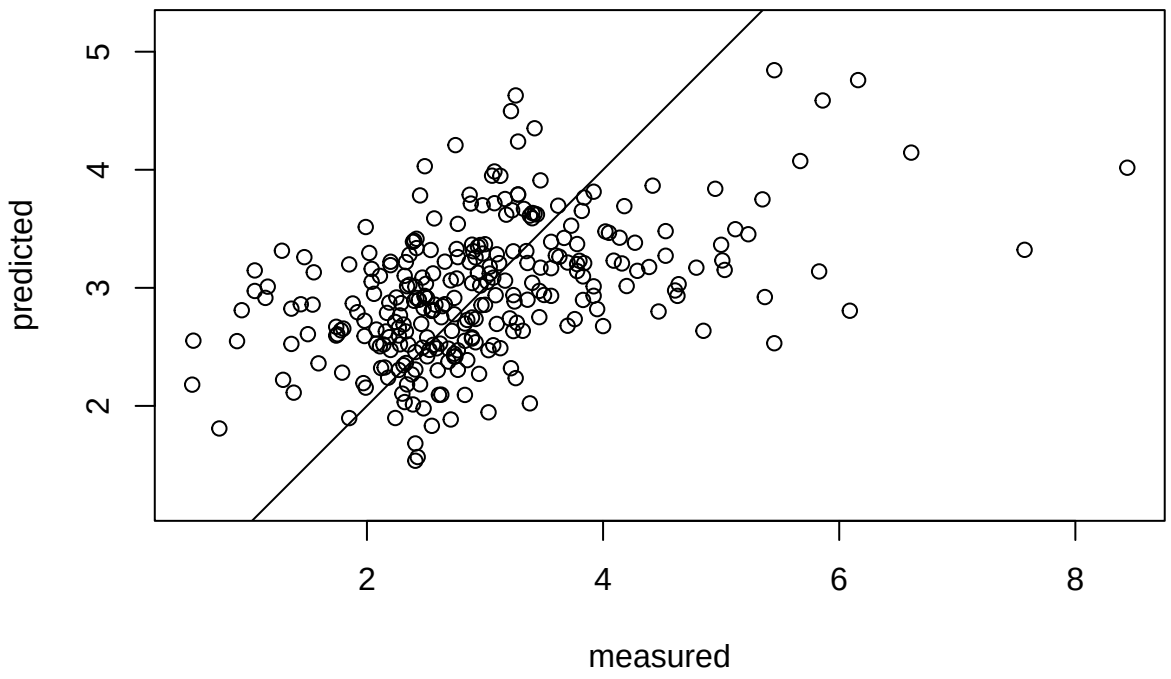
```
rent.size.kde.short.stand.pdf.pls$ncomp
```

```
## [1] 10
```

```
plot(RMSEP(rent.size.kde.short.stand.pdf.pls), legendpos = "topright")
```




```
plot(rent.size.kde.short.stand.pdf.pls, ncomp = 8, asp = 1, line = TRUE)
```




```
explvar(rent.size.kde.short.stand.pdf.pls)
```

```
##       Comp 1       Comp 2       Comp 3       Comp 4       Comp 5       Comp 6
## 5.681782e+01 1.145076e+01 2.176823e+01 8.449232e+00 1.186503e+00 1.939668e-01
##       Comp 7       Comp 8       Comp 9      Comp 10
## 1.312654e-01 9.953166e-04 5.021606e-04 4.027283e-04
```

```
head(predict(rent.size.kde.short.stand.pdf.pls, ncomp = 8))
```

```
## , , 8 comps
##
##   rent.size.kde.short.stand.pdf$Effective_Rent_SF
## 1                                        3.917520
## 2                                        4.235519
## 3                                        3.622745
## 4                                        4.442155
## 5                                        3.899333
## 6                                        4.892105
```

```
RMSEP(rent.size.kde.short.stand.pdf.pls, data = rent.bed.kde.short.stand.pdf.pls)
```

```
##        (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
## CV           1.099     1.06    1.038    1.032    1.005   0.9944   0.9762
## adjCV        1.099     1.06    1.038    1.033    1.005   0.9944   0.9761
##        7 comps  8 comps  9 comps  10 comps
## CV      0.9700   0.9682   0.9570    0.9521
## adjCV   0.9699   0.9682   0.9569    0.9520
```

```
predictions <- predict(rent.size.kde.short.stand.pdf.pls, type="response")
```

```
install.packages("caret")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
```

```
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:pls':
##
##     R2
```

```
varImp(rent.size.kde.short.stand.pdf.pls, scale=FALSE)
```

```
##                              Overall
## StyleGarden              7.387071e-04
## StyleLow-Rise            2.841939e-04
## StyleMid-Rise            4.063700e-04
## Number_Of_Units          5.304685e-04
## Transit_Time             1.281838e-04
## Drive_Time               2.224698e-04
## Parking_Ratio            1.334031e-03
## Bed_Count                2.841255e-03
## Size_                    2.242051e-04
## Amenities_AC             1.773222e-04
## Amenities_Safty          1.823976e-03
## Amenities_Pool           3.692434e-04
## Amenities_Indoor_Gather  6.897951e-04
## Amenities_Entertainment  4.113120e-05
## Amenities_Outdoor_Gather 6.529190e-04
## Amenities_Gym            1.132796e-03
## Amenities_EV             4.099693e-04
## Amenities_Service        1.572944e-03
## Distance_Beach           4.098457e-06
## Distance_Bike            8.500263e-07
## FB_All                   3.481648e-03
## FB_Youth                 3.378364e-03
## Noise                    1.979668e-03
## Flood                    1.466320e-04
## Distance_Beach_Inv       1.591646e-04
## Distance_Bike_Inv        4.016661e-05
## Transit_Time_Inv         1.679277e-07
## Drive_Time_Inv           2.163282e-08
## BIKE_800                 5.974511e-04
## BUS_800                  1.135720e-03
## BUS_1600                 1.198635e-03
## STORE_800                1.953891e-03
## STORE_1600               7.115046e-04
## STORE_4800               1.394647e-03
## REST_800                 1.184228e-03
## REST_1600                9.093112e-04
## REST_4800                1.107074e-03
## GYM_800                  2.513504e-03
```

```
## GYM_1600                9.246790e-04
## GYM_4800                9.245335e-04
## PARK_800                1.368008e-03
## PARK_1600               6.359258e-04
## PARK_4800               7.105378e-04
## Distance_HW             1.625163e-04
## ROAD_800                7.610137e-04
## Distance_HW_Inv         4.450875e-06
```

```
scores <- as.data.frame(rent.size.kde.short.stand.pdf.pls$scores[,1:10])
size.scores <- cbind(scores, predictions, rent.size.kde.short.stand.pdf)
colnames(size.scores)[1:20] <- c("scores ncomp=1", "scores ncomp=2", "scores ncomp=3", "scores ncomp=4"
write.csv(size.scores,"/cloud/project/housing crisis/pls_new/size_scores.csv")
```

```
ssr <- sum((rent.size.kde.short.stand.pdf.pls$residuals)^2)
mean<-mean(rent.size.kde.short.stand.pdf$Effective_Rent_SF)
sst<- sum((rent.size.kde.short.stand.pdf$Effective_Rent_SF-mean)^2)
1-ssr/sst
```

```
## [1] -6.750465
```

```
residual <- predictions - rent.size.kde.short.stand.pdf$Effective_Rent_SF
ssr1 <- sum((residual)^2)
mean<-mean(rent.size.kde.short.stand.pdf$Effective_Rent_SF)
sst<- sum((rent.size.kde.short.stand.pdf$Effective_Rent_SF-mean)^2)
1-ssr1/sst
```

```
## [1] -6.750465
```