Models

Group 4

2024-05-21

Model regressing livability on distance without considering feature of neighborhood

This model regress the liveability score on the distance of commercial spaces and does not consider characteristics such as age, income.

```
# Model with distance only
## Call:
## lm(formula = livability_score ~ . - Standard_Deviation_Age -
##     Weighted_Average_Age - age_cluster - social_cluster -
demographic_cluster -
##     id - postal_code - prop_0.14 - prop_15.24 - prop_25.44 -
##     prop_45.64 - prop_65. - prop_university - worker - prop_women -
##     pop_density - Aid - Income - single_family_housing, data = df_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05066 -0.15723  0.01341  0.16369  0.66992
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    8.1954773  0.0645748 126.915  < 2e-16 ***
## Distance.GP                    0.0061508  0.0356125   0.173  0.86292
## Within.1.km.GP                -0.0098559  0.0074427  -1.324  0.18584
## Within.3.km.GP                -0.0051672  0.0036349  -1.422  0.15558
## Within.5.km.GP                 0.0011128  0.0024885   0.447  0.65488
## Distance.Hospital             -0.0176822  0.0088458  -1.999  0.04598 *
## Within.5.km.Hospital           0.0174529  0.0124452   1.402  0.16122
## Within.10.km.Hospital         -0.0284147  0.0092826  -3.061  0.00229 **
## Within.20.km.Hospital          0.0165597  0.0066047   2.507  0.01238 *
## Distance.Supermarket           0.1380826  0.0316202   4.367 1.44e-05 ***
## Within.1.km.Supermarket       -0.0328055  0.0066770  -4.913 1.10e-06 ***
## Within.3.km.Supermarket        0.0033401  0.0028916   1.155  0.24841
## Within.5.km.Supermarket        0.0006951  0.0020758   0.335  0.73784
## Distance.Department.Store      0.0100464  0.0120690   0.832  0.40544
## Within.5.km.Department.Store   0.0647023  0.0089826   7.203 1.45e-12 ***
## Within.10.km.Department.Store  0.0201135  0.0045509   4.420 1.14e-05 ***
## Within.20.km.Department.Store -0.0090779  0.0035462  -2.560  0.01067 *
## Distance.Restaurant           -0.0995167  0.0347165  -2.867  0.00427 **
## Within.1.km.Restaurant         0.0023029  0.0003552   6.483 1.64e-10 ***
```

```
## Within.3.km.Restaurant          0.0002244  0.0001277   1.757  0.07926 .
## Within.5.km.Restaurant         -0.0002842  0.0001247  -2.278  0.02300 *
## Distance.Day.Care              -0.0506514  0.0445685  -1.136  0.25612
## Within.1.km.Day.Care            0.0069666  0.0030117   2.313  0.02099 *
## Within.3.km.Day.Care            0.0075513  0.0013114   5.758 1.25e-08 ***
## Within.5.km.Day.Care            0.0015816  0.0010869   1.455  0.14605
## Distance.School                 0.0502357  0.0389988   1.288  0.19810
## Within.1.km.School             -0.0090292  0.0094198  -0.959  0.33811
## Within.3.km.School             -0.0203988  0.0043803  -4.657 3.80e-06 ***
## Within.5.km.School             -0.0113338  0.0027777  -4.080 4.99e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2488 on 741 degrees of freedom
## Multiple R-squared:  0.5906, Adjusted R-squared:  0.5751
## F-statistic: 38.17 on 28 and 741 DF,  p-value: < 2.2e-16
```

Models considering features of neighborhood

This model regress liveability score on distance of commercial space and the feature of neighborhood. Therefore, it allows to control the characteristics of the neighborhood such as age, income. Notice that the age variable is recorded into 2 methods. The first method uses the average and SD value for age of citizens living in the neighborhood. The second method records age of different groups and transfer to proportion. There is 5 bins of age values `prop_0.14`, `prop_15.24`, `prop_25.44`, `prop_45.64`, and `prop_65.`.

In order to choose the best performance model, we run the model in both method recoding age.

```
# Linear model with proportion of age
## Call:
## lm(formula = livability_score ~ . - id - postal_code -
Standard_Deviation_Age -
##     Weighted_Average_Age - age_cluster - social_cluster -
demographic_cluster,
##     data = df_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53256 -0.09279  0.00760  0.09427  0.55218
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    8.050e+00  1.214e+00   6.628 6.61e-11 ***
## prop_0.14                     -1.596e+00  1.199e+00  -1.331 0.183569
## prop_15.24                    -1.042e+00  1.203e+00  -0.866 0.386837
## prop_25.44                    -1.430e+00  1.210e+00  -1.182 0.237500
## prop_45.64                    -6.563e-01  1.222e+00  -0.537 0.591372
## prop_65.                      -5.101e-01  1.217e+00  -0.419 0.675309
```

```
## prop_university                 8.464e-01  7.181e-02  11.788  < 2e-16 ***
## worker                          5.492e-01  1.396e-01   3.933 9.17e-05 ***
## prop_women                      9.827e-01  2.521e-01   3.899 0.000106 ***
## pop_density                    -4.103e-02  5.859e-02  -0.700 0.483983
## Aid                            -1.598e-01  8.796e-02  -1.817 0.069703 .
## Income                          9.310e-01  9.398e-02   9.906  < 2e-16 ***
## single_family_housing           2.245e-01  3.267e-02   6.872 1.36e-11 ***
## Distance.GP                     1.940e-02  2.253e-02   0.861 0.389429
## Within.1.km.GP                 -1.164e-03  4.726e-03  -0.246 0.805465
## Within.3.km.GP                  2.102e-03  2.324e-03   0.904 0.366098
## Within.5.km.GP                 -1.758e-03  1.591e-03  -1.104 0.269756
## Distance.Hospital              -9.382e-03  5.738e-03  -1.635 0.102452
## Within.5.km.Hospital            2.799e-03  8.080e-03   0.346 0.729159
## Within.10.km.Hospital          -1.601e-02  5.966e-03  -2.684 0.007440 **
## Within.20.km.Hospital           1.595e-02  4.388e-03   3.635 0.000298 ***
## Distance.Supermarket            5.279e-02  2.022e-02   2.610 0.009229 **
## Within.1.km.Supermarket        -5.506e-03  4.300e-03  -1.280 0.200779
## Within.3.km.Supermarket        -6.755e-04  1.879e-03  -0.360 0.719315
## Within.5.km.Supermarket         2.180e-03  1.376e-03   1.585 0.113511
## Distance.Department.Store       3.276e-03  7.729e-03   0.424 0.671794
## Within.5.km.Department.Store    2.370e-02  6.021e-03   3.937 9.06e-05 ***
## Within.10.km.Department.Store   1.179e-02  2.927e-03   4.029 6.18e-05 ***
## Within.20.km.Department.Store -7.112e-03  2.420e-03  -2.940 0.003390 **
## Distance.Restaurant            -8.371e-02  2.244e-02  -3.730 0.000206 ***
## Within.1.km.Restaurant          1.583e-03  2.322e-04   6.819 1.92e-11 ***
## Within.3.km.Restaurant          1.480e-05  8.227e-05   0.180 0.857241
## Within.5.km.Restaurant         -7.879e-06  8.651e-05  -0.091 0.927460
## Distance.Day.Care              -2.650e-02  2.883e-02  -0.919 0.358321
## Within.1.km.Day.Care            3.604e-03  2.011e-03   1.792 0.073544 .
## Within.3.km.Day.Care            3.361e-03  8.578e-04   3.918 9.78e-05 ***
## Within.5.km.Day.Care           -2.895e-04  6.979e-04  -0.415 0.678384
## Distance.School                 9.163e-03  2.475e-02   0.370 0.711374
## Within.1.km.School             -1.443e-03  6.152e-03  -0.235 0.814623
## Within.3.km.School             -1.072e-02  2.829e-03  -3.791 0.000163 ***
## Within.5.km.School             -2.296e-03  1.860e-03  -1.235 0.217287
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1555 on 729 degrees of freedom
## Multiple R-squared:  0.8427, Adjusted R-squared:  0.8341
## F-statistic: 97.64 on 40 and 729 DF,  p-value: < 2.2e-16

# Linear model with average and SD of Age

##
## Call:
## lm(formula = livability_score ~ . - id - postal_code - prop_0.14 -
##     prop_15.24 - prop_25.44 - prop_45.64 - prop_65. - age_cluster -
##     social_cluster - demographic_cluster, data = df_train)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55221 -0.09263  0.01135  0.09815  0.51582
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.108e+00  1.510e-01  40.451  < 2e-16 ***
## Weighted_Average_Age           1.732e-02  2.539e-03   6.822 1.89e-11 ***
## Standard_Deviation_Age         4.495e-03  5.125e-03   0.877 0.380672
## prop_university                8.481e-01  7.168e-02  11.831  < 2e-16 ***
## worker                         6.454e-01  1.326e-01   4.867 1.39e-06 ***
## prop_women                     1.230e+00  2.395e-01   5.138 3.57e-07 ***
## pop_density                   -4.956e-02  5.785e-02  -0.857 0.391838
## Aid                           -2.132e-01  8.807e-02  -2.421 0.015720 *
## Income                         8.464e-01  8.965e-02   9.442  < 2e-16 ***
## single_family_housing          2.443e-01  3.088e-02   7.913 9.32e-15 ***
## Distance.GP                    1.082e-02  2.296e-02   0.471 0.637538
## Within.1.km.GP                -8.636e-04  4.822e-03  -0.179 0.857900
## Within.3.km.GP                 2.380e-03  2.368e-03   1.005 0.315286
## Within.5.km.GP                -1.361e-03  1.617e-03  -0.842 0.400169
## Distance.Hospital             -1.090e-02  5.771e-03  -1.888 0.059380 .
## Within.5.km.Hospital           9.947e-03  8.121e-03   1.225 0.221002
## Within.10.km.Hospital         -1.499e-02  6.074e-03  -2.469 0.013795 *
## Within.20.km.Hospital          1.616e-02  4.468e-03   3.617 0.000318 ***
## Distance.Supermarket           5.283e-02  2.058e-02   2.567 0.010466 *
## Within.1.km.Supermarket       -6.672e-03  4.375e-03  -1.525 0.127692
## Within.3.km.Supermarket        3.719e-04  1.909e-03   0.195 0.845600
## Within.5.km.Supermarket        1.579e-03  1.394e-03   1.132 0.257926
## Distance.Department.Store      4.676e-04  7.781e-03   0.060 0.952091
## Within.5.km.Department.Store   2.087e-02  6.068e-03   3.439 0.000616 ***
## Within.10.km.Department.Store  1.168e-02  2.961e-03   3.943 8.80e-05 ***
## Within.20.km.Department.Store -7.071e-03  2.465e-03  -2.869 0.004235 **
## Distance.Restaurant           -7.924e-02  2.261e-02  -3.504 0.000486 ***
## Within.1.km.Restaurant         1.669e-03  2.343e-04   7.122 2.54e-12 ***
## Within.3.km.Restaurant        -4.941e-06  8.389e-05  -0.059 0.953045
## Within.5.km.Restaurant        -1.310e-05  8.684e-05  -0.151 0.880173
## Distance.Day.Care             -1.585e-02  2.918e-02  -0.543 0.587181
## Within.1.km.Day.Care           3.297e-03  2.047e-03   1.611 0.107568
## Within.3.km.Day.Care           3.447e-03  8.719e-04   3.953 8.46e-05 ***
## Within.5.km.Day.Care          -5.997e-05  7.099e-04  -0.084 0.932696
## Distance.School                1.684e-02  2.531e-02   0.665 0.505950
## Within.1.km.School            -3.720e-04  6.251e-03  -0.060 0.952564
## Within.3.km.School            -1.265e-02  2.863e-03  -4.420 1.14e-05 ***
## Within.5.km.School            -2.918e-03  1.871e-03  -1.559 0.119383
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1587 on 732 degrees of freedom
## Multiple R-squared:  0.8354, Adjusted R-squared:  0.8271
## F-statistic: 100.4 on 37 and 732 DF,  p-value: < 2.2e-16
```

Model regress liveability on different measurement of commercial space distance

Model regress liveability on `distance`

```
##
## Call:
## lm(formula = livability_score ~ prop_0.14 + prop_15.24 + prop_25.44 +
##     prop_45.64 + prop_65. + prop_university + worker + prop_women +
##     pop_density + Aid + Income + single_family_housing + Distance.GP +
##     Distance.Hospital + Distance.Supermarket + Distance.Department.Store +
##     Distance.Restaurant + Distance.Day.Care + Distance.School,
##     data = df_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64720 -0.11691  0.00151  0.10587  0.71436
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                7.952939   1.505527   5.282 1.67e-07 ***
## prop_0.14                 -2.599557   1.488512  -1.746 0.081149 .
## prop_15.24                -1.119537   1.494460  -0.749 0.454017
## prop_25.44                -1.129438   1.504249  -0.751 0.452989
## prop_45.64                -0.078364   1.516084  -0.052 0.958791
## prop_65.                  -0.467391   1.511508  -0.309 0.757239
## prop_university            1.000860   0.084668  11.821  < 2e-16 ***
## worker                     0.771837   0.170129   4.537 6.65e-06 ***
## prop_women                 0.865389   0.279384   3.097 0.002024 **
## pop_density                0.371279   0.056253   6.600 7.77e-11 ***
## Aid                       -0.182904   0.091364  -2.002 0.045653 *
## Income                     1.458116   0.103623  14.071  < 2e-16 ***
## single_family_housing      0.191764   0.035738   5.366 1.08e-07 ***
## Distance.GP                0.018008   0.025700   0.701 0.483700
## Distance.Hospital         -0.026486   0.006324  -4.188 3.14e-05 ***
## Distance.Supermarket       0.063535   0.024169   2.629 0.008744 **
## Distance.Department.Store -0.015799   0.008988  -1.758 0.079193 .
## Distance.Restaurant       -0.096510   0.027226  -3.545 0.000417 ***
## Distance.Day.Care         -0.038659   0.034571  -1.118 0.263813
## Distance.School            0.019428   0.027584   0.704 0.481442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1949 on 750 degrees of freedom
## Multiple R-squared:  0.7458, Adjusted R-squared:  0.7394
## F-statistic: 115.8 on 19 and 750 DF,  p-value: < 2.2e-16
```

Model regress liveability on `Within.5.km.CommercialSpace`

```
## Call:
## lm(formula = livability_score ~ prop_0.14 + prop_15.24 + prop_25.44 +
##     prop_45.64 + prop_65. + prop_university + worker + prop_women +
##     pop_density + Aid + Income + single_family_housing + Within.5.km.GP +
##     Within.5.km.Hospital + Within.5.km.Supermarket +
Within.5.km.Department.Store +
##     Within.5.km.Restaurant + Within.5.km.Day.Care + Within.5.km.School,
##     data = df_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.57836 -0.09962 -0.00350  0.11389  0.66409
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   7.375e+00  1.348e+00   5.469 6.15e-08 ***
## prop_0.14                    -1.469e+00  1.336e+00  -1.100  0.27178
## prop_15.24                   -2.608e-01  1.337e+00  -0.195  0.84539
## prop_25.44                   -6.987e-01  1.343e+00  -0.520  0.60304
## prop_45.64                    2.542e-01  1.352e+00   0.188  0.85089
## prop_65.                      2.234e-01  1.352e+00   0.165  0.86879
## prop_university               9.629e-01  7.678e-02  12.541  < 2e-16 ***
## worker                        6.361e-01  1.547e-01   4.112 4.35e-05 ***
## prop_women                    6.218e-01  2.633e-01   2.362  0.01845 *
## pop_density                   5.136e-02  5.556e-02   0.924  0.35558
## Aid                          -1.899e-01  8.625e-02  -2.202  0.02796 *
## Income                        1.002e+00  1.007e-01   9.955  < 2e-16 ***
## single_family_housing         2.806e-01  3.393e-02   8.271 6.05e-16 ***
## Within.5.km.GP                8.881e-04  1.189e-03   0.747  0.45528
## Within.5.km.Hospital          1.170e-03  7.318e-03   0.160  0.87306
## Within.5.km.Supermarket       5.148e-04  1.027e-03   0.501  0.61643
## Within.5.km.Department.Store  1.759e-02  6.213e-03   2.831  0.00477 **
## Within.5.km.Restaurant        3.769e-05  8.193e-05   0.460  0.64559
## Within.5.km.Day.Care          1.188e-03  6.447e-04   1.843  0.06567 .
## Within.5.km.School           -3.274e-03  1.853e-03  -1.767  0.07762 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1755 on 750 degrees of freedom
## Multiple R-squared:  0.7939, Adjusted R-squared:  0.7886
## F-statistic:   152 on 19 and 750 DF,  p-value: < 2.2e-16
```

Model regress liveability on clusters

It is hypothesized that neighborhoods can be categorized into distinct groups based on specific characteristic. Therefore, based on the defined clusters (`demographic_cluster`, `social_cluster`, and `age_cluster`), the model regresses liveability on distance and

clusters is created. Two models uses `distance` and `cluster`; and `distance`, `Within.X.Km.CommercialSpace` and `cluster` are as follows:

```
# Model - `distance` and `cluster`
lm_cluster_distance <- lm(livability_score ~ demographic_cluster +
social_cluster + age_cluster + Distance.GP + Distance.Hospital +
Distance.Supermarket+ Distance.Department.Store + Distance.Restaurant +
Distance.Day.Care + Distance.School, data = df_train)
summary(lm_cluster_distance)

##
## Call:
## lm(formula = livability_score ~ demographic_cluster + social_cluster +
##     age_cluster + Distance.GP + Distance.Hospital + Distance.Supermarket +
##     Distance.Department.Store + Distance.Restaurant + Distance.Day.Care +
##     Distance.School, data = df_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01605 -0.23124  0.00462  0.25425  1.06562
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                8.4436190  0.0390985 215.958  < 2e-16 ***
## demographic_cluster       -0.0424504  0.0312374  -1.359  0.17456
## social_cluster            -0.0006159  0.0154801  -0.040  0.96827
## age_cluster                0.0102354  0.0185439   0.552  0.58114
## Distance.GP               -0.0076723  0.0466371  -0.165  0.86937
## Distance.Hospital         -0.0524825  0.0111543  -4.705 3.01e-06 ***
## Distance.Supermarket       0.1959732  0.0431096   4.546 6.36e-06 ***
## Distance.Department.Store -0.0427317  0.0160587  -2.661  0.00796 **
## Distance.Restaurant       -0.2916043  0.0463485  -6.292 5.31e-10 ***
## Distance.Day.Care         -0.0717676  0.0607002  -1.182  0.23745
## Distance.School            0.1559936  0.0494148   3.157  0.00166 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3582 on 759 degrees of freedom
## Multiple R-squared:  0.1307, Adjusted R-squared:  0.1193
## F-statistic: 11.41 on 10 and 759 DF,  p-value: < 2.2e-16

# Model - `distance`, `Within.X.Km.CommercialSpace` and `cluster`

## Call:
## lm(formula = livability_score ~ . - id - postal_code -
Weighted_Average_Age -
##     Standard_Deviation_Age - prop_0.14 - prop_15.24 - prop_25.44 -
##     prop_45.64 - prop_65. - prop_university - worker - prop_women -
##     pop_density - Aid - Income - single_family_housing, data = df_train)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05067 -0.15588  0.01544  0.16429  0.64657
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   8.1853212  0.0669693 122.225  < 2e-16 ***
## Distance.GP                   0.0056952  0.0356655   0.160  0.87317
## Within.1.km.GP               -0.0096094  0.0074539  -1.289  0.19774
## Within.3.km.GP               -0.0052816  0.0036398  -1.451  0.14718
## Within.5.km.GP                0.0007984  0.0024980   0.320  0.74935
## Distance.Hospital            -0.0180552  0.0088559  -2.039  0.04183 *
## Within.5.km.Hospital          0.0170170  0.0124584   1.366  0.17239
## Within.10.km.Hospital        -0.0289838  0.0092941  -3.119  0.00189 **
## Within.20.km.Hospital         0.0145460  0.0067310   2.161  0.03101 *
## Distance.Supermarket          0.1382454  0.0316226   4.372 1.41e-05 ***
## Within.1.km.Supermarket      -0.0336330  0.0067128  -5.010 6.81e-07 ***
## Within.3.km.Supermarket       0.0034660  0.0028973   1.196  0.23198
## Within.5.km.Supermarket       0.0007185  0.0020852   0.345  0.73050
## Distance.Department.Store     0.0100567  0.0120737   0.833  0.40515
## Within.5.km.Department.Store  0.0639030  0.0090021   7.099 2.96e-12 ***
## Within.10.km.Department.Store 0.0208098  0.0046136   4.511 7.52e-06 ***
## Within.20.km.Department.Store -0.0085103  0.0035632  -2.388  0.01717 *
## Distance.Restaurant           -0.0987487  0.0348172  -2.836  0.00469 **
## Within.1.km.Restaurant         0.0023318  0.0003563   6.545 1.12e-10 ***
## Within.3.km.Restaurant         0.0002212  0.0001278   1.731  0.08383 .
## Within.5.km.Restaurant        -0.0002782  0.0001256  -2.216  0.02700 *
## Distance.Day.Care             -0.0486844  0.0446189  -1.091  0.27558
## Within.1.km.Day.Care           0.0075008  0.0030291   2.476  0.01350 *
## Within.3.km.Day.Care           0.0075188  0.0013117   5.732 1.45e-08 ***
## Within.5.km.Day.Care           0.0017049  0.0010985   1.552  0.12107
## Distance.School                0.0483897  0.0390672   1.239  0.21588
## Within.1.km.School            -0.0098984  0.0094424  -1.048  0.29485
## Within.3.km.School            -0.0203808  0.0043808  -4.652 3.89e-06 ***
## Within.5.km.School            -0.0114415  0.0027956  -4.093 4.73e-05 ***
## age_cluster                   -0.0015947  0.0130945  -0.122  0.90310
## demographic_cluster            0.0387912  0.0247795   1.565  0.11790
## social_cluster                 0.0091519  0.0110123   0.831  0.40621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2488 on 738 degrees of freedom
## Multiple R-squared:  0.5922, Adjusted R-squared:  0.5751
## F-statistic: 34.57 on 31 and 738 DF,  p-value: < 2.2e-16
```

Comparing models

To find the best performance models, we compare 7 models in term of adjusted R squares and residual mean square error (RMSE)

```
##                               Model Adjusted_R_squared      RMSE
## 1        Distance without features          0.5750989 0.2570767
## 2 Features-distances-age propotion          0.8340705 0.1648036
## 3 Features-distances-AgeMean-AgeSD          0.8271036 0.1687580
## 4                Featuess-Distance          0.7393824 0.2030783
## 5                  Features-Within          0.7886404 0.1854019
## 6                Cluster-distances          0.1192619 0.3623502
## 7         Cluster-distances-Within          0.5750621 0.2565696
```

Conclusion

The two best performance models are the models that regress liveability on features, distance, and the number of commercial space within a specific radius (1km, 3km, 5km, 10km, and 20km). The different methods of age measurement do not significantly impact the Adjusted_R_square and RMSE. Therefore, these two measurement can be used to measure age of the neighborhood without affecting the prediction performance of the models.