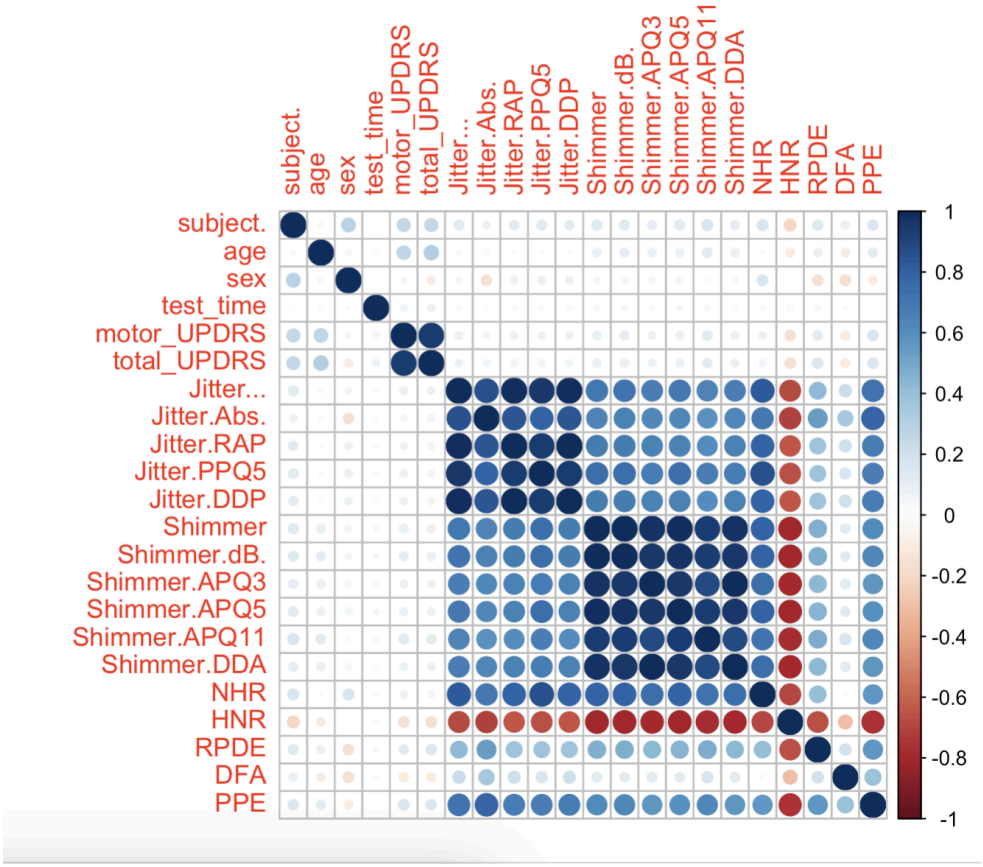
1. Use the appropriate functions to obtain descriptive information about the variables included in the dataset (paste or include a screenshot with the resulting information).

```
  subject. age sex test_time motor_UPDRS total_UPDRS Jitter...
1        1  72   0    5.6431      28.199      34.398   0.00662
2        1  72   0   12.6660      28.447      34.894   0.00300
3        1  72   0   19.6810      28.695      35.389   0.00481
4        1  72   0   25.6470      28.905      35.810   0.00528
5        1  72   0   33.6420      29.187      36.375   0.00335
6        1  72   0   40.6520      29.435      36.870   0.00353
  Jitter.Abs. Jitter.RAP Jitter.PPQ5 Jitter.DDP Shimmer Shimmer.dB.
1   3.380e-05    0.00401     0.00317    0.01204 0.02565       0.230
2   1.680e-05    0.00132     0.00150    0.00395 0.02024       0.179
3   2.462e-05    0.00205     0.00208    0.00616 0.01675       0.181
4   2.657e-05    0.00191     0.00264    0.00573 0.02309       0.327
5   2.014e-05    0.00093     0.00130    0.00278 0.01703       0.176
6   2.290e-05    0.00119     0.00159    0.00357 0.02227       0.214
  Shimmer.APQ3 Shimmer.APQ5 Shimmer.APQ11 Shimmer.DDA      NHR    HNR
1      0.01438      0.01309       0.01662     0.04314 0.014290 21.640
2      0.00994      0.01072       0.01689     0.02982 0.011112 27.183
3      0.00734      0.00844       0.01458     0.02202 0.020220 23.047
4      0.01106      0.01265       0.01963     0.03317 0.027837 24.445
5      0.00679      0.00929       0.01819     0.02036 0.011625 26.126
6      0.01006      0.01337       0.02263     0.03019 0.009438 22.946
     RPDE     DFA     PPE
1 0.41888 0.54842 0.16006
2 0.43493 0.56477 0.10810
3 0.46222 0.54405 0.21014
4 0.48730 0.57794 0.33277
5 0.47188 0.56122 0.19361
6 0.53949 0.57243 0.19500
> names(parkinsons_updrs)
 [1] "subject."      "age"           "sex"           "test_time"
 [5] "motor_UPDRS"   "total_UPDRS"   "Jitter..."     "Jitter.Abs."
 [9] "Jitter.RAP"    "Jitter.PPQ5"   "Jitter.DDP"    "Shimmer"
[13] "Shimmer.dB."   "Shimmer.APQ3"  "Shimmer.APQ5"  "Shimmer.APQ11"
[17] "Shimmer.DDA"   "NHR"           "HNR"           "RPDE"
[21] "DFA"           "PPE"
>
```

```
> summary(parkinsons_updrs)
    subject.          age              sex            test_time        motor_UPDRS      total_UPDRS
 Min.   : 1.00   Min.   :36.0   Min.   :0.0000   Min.   : -4.263   Min.   : 5.038   Min.   : 7.00
 1st Qu.:10.00   1st Qu.:58.0   1st Qu.:0.0000   1st Qu.: 46.847   1st Qu.:15.000   1st Qu.:21.37
 Median :22.00   Median :65.0   Median :0.0000   Median : 91.523   Median :20.871   Median :27.58
 Mean   :21.49   Mean   :64.8   Mean   :0.3178   Mean   : 92.864   Mean   :21.296   Mean   :29.02
 3rd Qu.:33.00   3rd Qu.:72.0   3rd Qu.:1.0000   3rd Qu.:138.445   3rd Qu.:27.596   3rd Qu.:36.40
 Max.   :42.00   Max.   :85.0   Max.   :1.0000   Max.   :215.490   Max.   :39.511   Max.   :54.99
   Jitter...         Jitter.Abs.          Jitter.RAP        Jitter.PPQ5         Jitter.DDP
 Min.   :0.000830   Min.   :2.250e-06   Min.   :0.000330   Min.   :0.000430   Min.   :0.000980
 1st Qu.:0.003580   1st Qu.:2.244e-05   1st Qu.:0.001580   1st Qu.:0.001820   1st Qu.:0.004730
 Median :0.004900   Median :3.453e-05   Median :0.002250   Median :0.002490   Median :0.006750
 Mean   :0.006154   Mean   :4.403e-05   Mean   :0.002987   Mean   :0.003277   Mean   :0.008962
 3rd Qu.:0.006800   3rd Qu.:5.333e-05   3rd Qu.:0.003290   3rd Qu.:0.003460   3rd Qu.:0.009870
 Max.   :0.099990   Max.   :4.456e-04   Max.   :0.057540   Max.   :0.069560   Max.   :0.172630
    Shimmer          Shimmer.dB.      Shimmer.APQ3       Shimmer.APQ5      Shimmer.APQ11
 Min.   :0.00306   Min.   :0.026   Min.   :0.00161   Min.   :0.00194   Min.   :0.00249
 1st Qu.:0.01912   1st Qu.:0.175   1st Qu.:0.00928   1st Qu.:0.01079   1st Qu.:0.01566
 Median :0.02751   Median :0.253   Median :0.01370   Median :0.01594   Median :0.02271
 Mean   :0.03404   Mean   :0.311   Mean   :0.01716   Mean   :0.02014   Mean   :0.02748
 3rd Qu.:0.03975   3rd Qu.:0.365   3rd Qu.:0.02057   3rd Qu.:0.02375   3rd Qu.:0.03272
 Max.   :0.26863   Max.   :2.107   Max.   :0.16267   Max.   :0.16702   Max.   :0.27546
  Shimmer.DDA           NHR                HNR              RPDE             DFA
 Min.   :0.00484   Min.   :0.000286   Min.   : 1.659   Min.   :0.1510   Min.   :0.5140
 1st Qu.:0.02783   1st Qu.:0.010955   1st Qu.:19.406   1st Qu.:0.4698   1st Qu.:0.5962
 Median :0.04111   Median :0.018448   Median :21.920   Median :0.5423   Median :0.6436
 Mean   :0.05147   Mean   :0.032120   Mean   :21.680   Mean   :0.5415   Mean   :0.6532
 3rd Qu.:0.06173   3rd Qu.:0.031463   3rd Qu.:24.444   3rd Qu.:0.6140   3rd Qu.:0.7113
 Max.   :0.48802   Max.   :0.748260   Max.   :37.875   Max.   :0.9661   Max.   :0.8656
      PPE
 Min.   :0.02198
 1st Qu.:0.15634
 Median :0.20550
 Mean   :0.21959
 3rd Qu.:0.26449
 Max.   :0.73173
>
```

2. Calculate the correlation between the different attributes (include the figure produced by R in your answer).



3. Divide the input dataset into training and testing. a. Split the datasets using 80% for training and 20% for testing. b. How many examples will be used for training and how many for testing?

parkinsons_updrs	5875 obs. of 22 variables
test	1175 obs. of 22 variables
train	4700 obs. of 22 variables

4. Build a multiple linear regression model containing all the input variables to predict the output variable. a. Which predictors have a significant impact in the prediction? b. How does the model perform? Provide the R2 and RSE.

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.453e+01  2.583e+00  17.238  < 2e-16 ***
Jitter...      1.657e+02  1.915e+02   0.866 0.386736
Jitter.Abs.   -5.139e+04  8.007e+03  -6.418 1.52e-10 ***
Jitter.RAP    -5.715e+04  4.138e+04  -1.381 0.167383
Jitter.PPQ5    1.516e+01  1.736e+02   0.087 0.930418
Jitter.DDP     1.911e+04  1.380e+04   1.385 0.166037
Shimmer        1.042e+02  5.655e+01   1.843 0.065344 .
Shimmer.dB.   -3.490e+00  4.373e+00  -0.798 0.424873
Shimmer.APQ3  -1.669e+04  4.138e+04  -0.403 0.686746
Shimmer.APQ5  -1.863e+02  5.090e+01  -3.659 0.000256 ***
Shimmer.APQ11  1.287e+02  2.413e+01   5.334 1.01e-07 ***
Shimmer.DDA    5.521e+03  1.379e+04   0.400 0.688994
NHR           -2.289e+01  5.381e+00  -4.254 2.14e-05 ***
HNR           -4.153e-01  6.048e-02  -6.866 7.47e-12 ***
RPDE           2.380e+00  1.606e+00   1.482 0.138319
DFA           -2.858e+01  1.995e+00 -14.328  < 2e-16 ***
PPE            2.014e+01  2.569e+00   7.838 5.63e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.699 on 4683 degrees of freedom
Multiple R-squared:  0.1087,    Adjusted R-squared:  0.1057
F-statistic: 35.71 on 16 and 4683 DF,  p-value: < 2.2e-16
```

a. (Intercept), Jitter.Abs., Shimmer.APQ5, Shimmer.APQ11, NHR, HNR, DFA, and PPE have significant impact in the prediction.
b. R2 = 0.09672578, RSE = 7.679084