Cross Section Assignment

Jacques Rossouw^a

^aStellebosch, Western Cape, South Africa

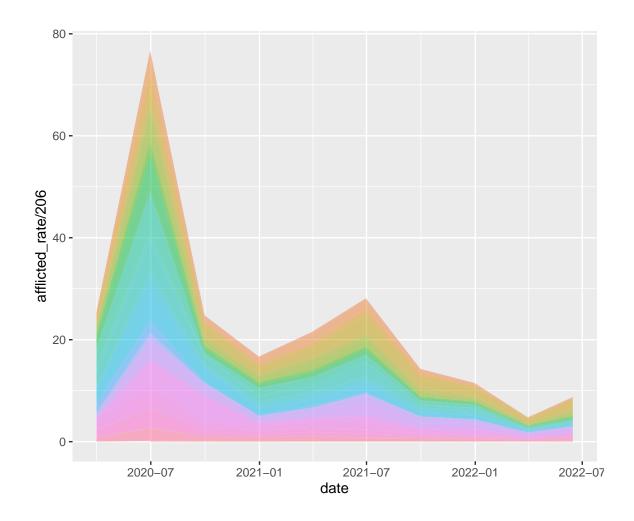
```
## # A tibble: 7 x 1
## continent
## <chr>
## 1 Asia
## 2 <NA>
## 3 Europe
## 4 Africa
## 5 North America
## 6 South America
## 7 Oceania
```

Share of woman vs men in SA about equal

We need to have a look at the correlation between the different variables to see what type of relationships exist. Furthermore, it is important to

1. Cumulative Values

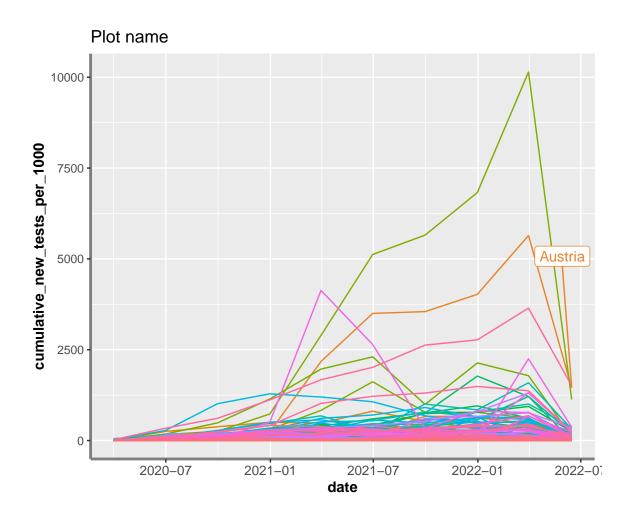
```
##
                       mean
                                 sd
                                      min
                                               max
                                                      range
## afflicted_rate
                                                    1479.96
                      23.37
                             88.83 0.00
                                           1479.96
## reproduction_rate
                       0.77
                              0.44 - 0.01
                                              2.06
                                                       2.08
## new_tests
                     117.76 460.31 0.00 10142.16 10142.16
## new_vaccinations
                      70.79 166.06 0.00
                                           1689.38
                                                    1689.38
## stringency_index
                      44.81 25.15 0.00
                                             99.06
                                                      99.06
```

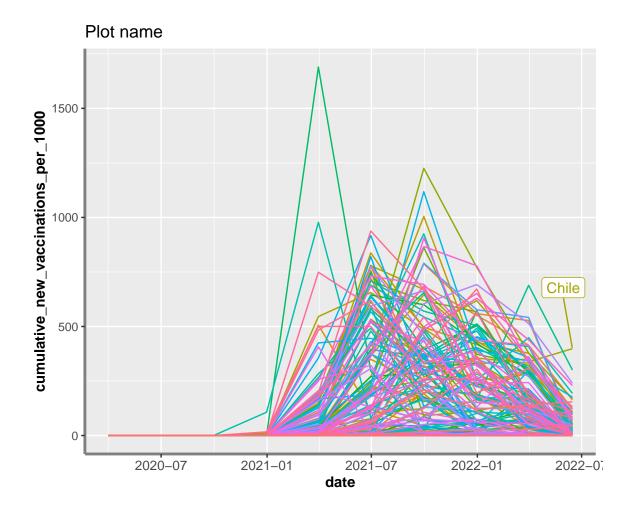


2. Scaling

##		mean	sd	min	max	range
##	afflicted_rate	23.37	88.83	0.00	1479.96	1479.96
##	reproduction_rate	0.77	0.44	-0.01	2.06	2.08
##	new_tests	117.76	460.31	0.00	10142.16	10142.16
##	new_vaccinations	70.79	166.06	0.00	1689.38	1689.38
##	stringency_index	44.81	25.15	0.00	99.06	99.06
##	new_tests_cum_per_1000	487.10	1789.23	0.00	32919.30	32919.30
##	new_vaccinations_cum_per_1000	275.41	558.44	0.00	3041.92	3041.92

Thus, want to scale: new_test, new_vaccinations





2.1. Country specific feature scaling

Now, to check the scales of the features that remain constant per country:

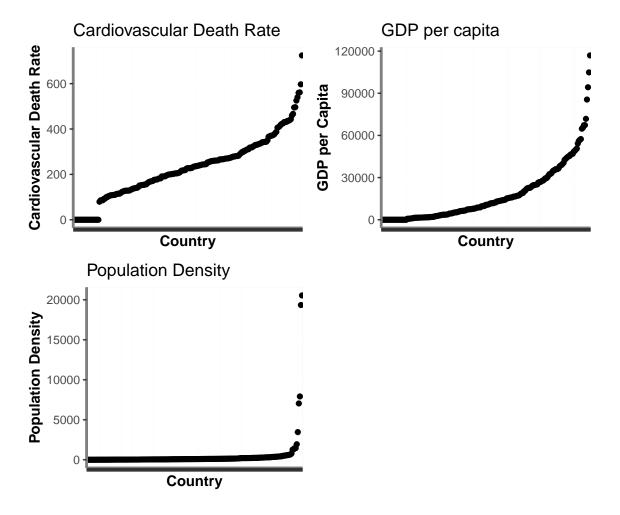
##		mean	sd	min	max	range
##	gdp_per_capita	17697.35	20539.28	0	116935.60	116935.60
##	population_density	444.44	2094.60	0	20546.77	20546.77
##	median_age	27.58	12.79	0	48.20	48.20
##	aged_65_older	7.90	6.48	0	27.05	27.05
##	extreme_poverty	7.83	16.76	0	77.60	77.60
##	cardiovasc_death_rate	226.19	135.19	0	724.42	724.42
##	diabetes_prevalence	7.52	4.59	0	23.36	23.36
##	handwashing_facilities	21.89	32.72	0	99.00	99.00
##	hosp_beds_1k	2.38	2.51	0	13.80	13.80
##	life_expectancy	73.36	9.08	0	86.75	86.75

human_development_index 0.63 0.28 0 0.96 0.96 ## smokers 14.38 12.75 0 45.95 45.95

Additional features that need to be scaled are this

- gdp_per_capita
- population_density
- cardiovasc_death_rate

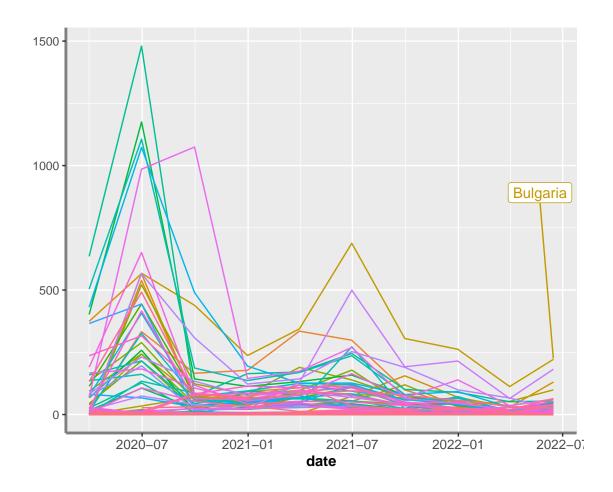
2.1.1. Plot



Now we can check all the descriptive stats for all the columns

mean sd min max range

##	afflicted_rate	23.37	88.83	0.00	1479.96	1479.96
##	reproduction_rate	0.77	0.44	-0.01	2.06	2.08
##	new_tests	0.00	1.00	-0.26	21.78	22.03
##	new_vaccinations	0.00	1.00	-0.43	9.75	10.17
##	stringency_index	44.81	25.15	0.00	99.06	99.06
##	population_density	444.44	2094.60	0.00	20546.77	20546.77
##	median_age	27.58	12.79	0.00	48.20	48.20
##	aged_65_older	7.90	6.48	0.00	27.05	27.05
##	gdp_per_capita	17697.35	20539.28	0.00	116935.60	116935.60
##	extreme_poverty	7.83	16.76	0.00	77.60	77.60
##	cardiovasc_death_rate	226.19	135.19	0.00	724.42	724.42
##	diabetes_prevalence	7.52	4.59	0.00	23.36	23.36
##	handwashing_facilities	21.89	32.72	0.00	99.00	99.00
##	hosp_beds_1k	2.38	2.51	0.00	13.80	13.80
##	life_expectancy	73.36	9.08	0.00	86.75	86.75
##	human_development_index	62.94	28.48	0.00	95.70	95.70
##	smokers	14.38	12.75	0.00	45.95	45.95
##	new_tests_cum_per_1000	487.10	1789.23	0.00	32919.30	32919.30
##	${\tt new_vaccinations_cum_per_1000}$	275.41	558.44	0.00	3041.92	3041.92
##	population_density_norm	0.00	1.00	-0.21	9.60	9.81
##	cardiovasc_death_rate_norm	0.00	1.00	-1.67	3.69	5.36
##	gdp_per_capita_log	8.21	3.22	0.00	11.67	11.67



3. Correlation

```
0 0.2 0.2-0.070.040.0-0.150.080.40.140.3-0.280.290.470.1-0.160.19
                extreme poverty
           handwashing facilities
                                   0
                                          0.3 0.3-0.09.170.130.130.09-0.10.070.1450.08.060.1430.1420.0450.07
           cardiovasc_death_rate 0.2 0.3
                                                <mark>1-0.23</mark>.210.230.240.2<del>2</del>0.010.3-0.05.130.09<mark>0.3-</mark>80.0<del>7</del>0.140.17
    cardiovasc_death_rate_norm 0.2 0.3
                                                 -0.23.210.230.240.2<del>2</del>0.010.3-0.05.130.090.340.970.140.17
        population_density_norm_0.070.090.230.23 -0.010.120.140.040.040.120.00.140.06.190.02.050.06
                stringency_index_0.04.170.210.240.01
                                                           0.40.370.370.120.250.150.140.090.020.00.130.04
               reproduction_rate=0.00.130.230.230.120.4
                                                              0.580.450.110.370.280.30.040.00.130.160.15
      human_development_index=0.16.130.240.240.18.370.58
                                                                   0.730.210.530.550.490.190.240.220.20.23
             gdp_per_capita_log_0.08.090.220.2<del>2</del>0.04.370.450.73
                                                                       0.160.410.440.340.220.040.160.130.15
                    afflicted_rate _0.1_0.40.0-0.0-0.04.120.110.210.16
                                                                           0.260.360.250.03.160.020.030.02
                        smokers_0.14.070.3 0.3-0.12.250.370.590.410.26
                                                                               0.560.510.040.230.220.180.21
                  aged_65_older_<mark>0.3-0.1-50.0-50.0-50.00.1-50.2-80.5-50.4-40.3-60.5-6</mark>
                                                                                   0.580.110.470.240.240.28
                  hosp_beds_1k_0.280.08.130.130.140.140.30.490.340.250.510.58
                                                                                       0.030.330.140.160.18
            diabetes_prevalence_0.29.06.09.09.09.00.00.01.19.220.03.040.110.03
                                                                                           0.220.04 0 0.01
                  life_expectancy_0.470.130.380.38.190.020.00.240.040.160.230.470.330.22
                                                                                                0.180.220.26
       new_tests_cum_per_1000_<mark>0.1-0.1-20.0-70.0-70.0-20.001.130.220.160.020.220.240.140.040.18</mark>
                                                                                                    0.150.32
               new_vaccinations=0.160.050.140.160.050.130.160.20.130.030.180.240.16 0 0.220.15
                                                                                                        0.62
new_vaccinations_cum_per_1000-0.1-0.0-70.1-70.107.0-60.04.1-50.2-30.1-50.02.2-10.2-80.1-80.0-10.2-60.3-20.6-2
```

4. Regressions

```
## Pooling Model
##
## Call:
## plm(formula = world_df$afflicted_rate ~ world_df$stringency_index,
## data = world_df, model = "pooling", index = c("location",
## "date"))
##
## Unbalanced Panel: n = 206, T = 7-10, N = 2050
##
```

0

```
## Residuals:
##
       Min.
              1st Qu.
                         Median
                                  3rd Qu.
                                               Max.
   -44.2658 -27.7049 -18.8655
                                  -4.1618 1446.7652
##
##
## Coefficients:
##
                            Estimate Std. Error t-value Pr(>|t|)
## (Intercept)
                             4.16180
                                        3.98103 1.0454
                                                            0.296
## world_df$stringency_index 0.42865
                                        0.07748 5.5324 3.564e-08 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:
                           16169000
## Residual Sum of Squares: 15931000
## R-Squared:
                  0.014725
## Adj. R-Squared: 0.014244
## F-statistic: 30.6074 on 1 and 2048 DF, p-value: 3.5636e-08
##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek
## % Date and time: Wed, Jun 22, 2022 - 22:12:51
## \begin{table}[!htbp] \centering
##
    \caption{}
##
    \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \[-1.8ex]\
## \hline \\[-1.8ex]
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\
## \cline{2-2}
## \\[-1.8ex] & afflicted\_rate \\
## \hline \\[-1.8ex]
## stringency\_index & 0.429$^{***}$ \\
    & (0.077) \\
##
   & \\
##
## Constant & 4.162 \\
   & (3.981) \\
##
    & \\
## \hline \\[-1.8ex]
## Observations & 2,050 \\
```

```
## R$^{2}$ & 0.015 \\
## Adjusted R$^{2}$ & 0.014 \\
## F Statistic & 30.607$^{***}$ (df = 1; 2048) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:} & \multicolumn{1}{r}{$^{*}}p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
## # A tibble: 5 x 7
                           estimate std.error statistic p.value conf.low conf.high
##
     term
##
     <chr>
                              <dbl>
                                        <dbl>
                                                  <dbl>
                                                           <dbl>
                                                                    <dbl>
                                                                              <dbl>
                          -1.07e+1 4.25
                                                  -2.51 1.21e- 2 -1.90e+1 -2.34
## 1 (Intercept)
                           2.44e-1 0.0785
                                                  3.11 1.91e- 3 9.02e-2 0.398
## 2 stringency_index
## 3 smokers
                           1.57e+0 0.156
                                                  10.0 3.41e-23 1.26e+0 1.87
                                                 -4.77 1.95e- 6 -4.09e-1 -0.171
## 4 handwashing_faciliti~ -2.90e-1 0.0608
                            3.90e-4 0.0000989
                                                   3.95 8.22e- 5 1.96e-4 0.000584
## 5 gdp_per_capita
## Pooling Model
##
## Call:
## plm(formula = afflicted_rate ~ stringency_index + smokers + handwashing_facilities +
       gdp_per_capita, data = world_df, model = "pooling", index = c("location",
##
##
       "date"))
##
## Unbalanced Panel: n = 206, T = 7-10, N = 2050
##
## Residuals:
##
               1st Qu.
       Min.
                         Median
                                   3rd Qu.
                                                Max.
   -85.2003 -27.8280 -8.3143
                                   8.6270 1415.3689
##
##
## Coefficients:
##
                               Estimate
                                           Std. Error t-value
## (Intercept)
                         -10.667536303
                                          4.247631108 -2.5114
## stringency_index
                            0.244094865
                                         0.078534554 3.1081
## smokers
                            1.567425974
                                         0.156093111 10.0416
## handwashing_facilities -0.290028848
                                          0.060775400 -4.7721
## gdp_per_capita
                            0.000390246
                                          0.000098897 3.9460
```

```
##
                                      Pr(>|t|)
## (Intercept)
                                      0.012102 *
## stringency_index
                                      0.001909 **
## smokers
                       < 0.00000000000000022 ***
## handwashing_facilities
                                   0.000001952 ***
## gdp_per_capita
                                   0.000082163 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:
                        16169000
## Residual Sum of Squares: 14649000
## R-Squared:
                  0.093978
## Adj. R-Squared: 0.092206
\#\# F-statistic: 53.0297 on 4 and 2045 DF, p-value: < 0.000000000000000222
```

5. Semester

	mean	sd	min	max	range
afflicted_rate	23.88	90.45	0.00	1379.93	1379.93
reproduction_rate	0.79	0.39	0.00	1.39	1.39
new_tests	234.60	840.38	0.00	12479.18	12479.18
new_vaccinations	141.02	293.40	0.00	1995.94	1995.94
stringency_index	45.81	24.36	0.00	95.61	95.61
population_density	443.11	2091.15	0.00	20546.77	20546.77
median_age	27.51	12.83	0.00	48.20	48.20
aged_65_older	7.88	6.49	0.00	27.05	27.05
gdp_per_capita	17632.21	20530.92	0.00	116935.60	116935.60
extreme_poverty	7.84	16.74	0.00	77.60	77.60
cardiovasc_death_rate	226.04	135.38	0.00	724.42	724.42
diabetes_prevalence	7.51	4.59	0.00	23.36	23.36
handwashing_facilities	21.85	32.69	0.00	99.00	99.00
hosp_beds_1k	2.38	2.51	0.00	13.80	13.80
life_expectancy	73.35	9.07	0.00	86.75	86.75
human_development_index	62.79	28.57	0.00	95.70	95.70
smokers	14.34	12.75	0.00	45.95	45.95
new_tests_cum_per_1000	539.63	1904.57	0.00	32919.30	32919.30
${\tt new_vaccinations_cum_per_1000}$	309.12	590.76	0.00	3041.92	3041.92
population_density_norm	0.00	1.00	-0.21	9.61	9.83
	new_tests new_vaccinations stringency_index population_density median_age aged_65_older gdp_per_capita extreme_poverty cardiovasc_death_rate diabetes_prevalence handwashing_facilities hosp_beds_1k life_expectancy human_development_index	afflicted_rate 23.88 reproduction_rate 0.79 new_tests 234.60 new_vaccinations 141.02 stringency_index 45.81 population_density 443.11 median_age 27.51 aged_65_older 7.88 gdp_per_capita 17632.21 extreme_poverty 7.84 cardiovasc_death_rate 226.04 diabetes_prevalence 7.51 handwashing_facilities 21.85 hosp_beds_1k 2.38 life_expectancy 73.35 human_development_index 62.79 smokers 14.34 new_tests_cum_per_1000 539.63 new_vaccinations_cum_per_1000 309.12	afflicted_rate 23.88 90.45 reproduction_rate 0.79 0.39 new_tests 234.60 840.38 new_vaccinations 141.02 293.40 stringency_index 45.81 24.36 population_density 443.11 2091.15 median_age 27.51 12.83 aged_65_older 7.88 6.49 gdp_per_capita 17632.21 20530.92 extreme_poverty 7.84 16.74 cardiovasc_death_rate 226.04 135.38 diabetes_prevalence 7.51 4.59 handwashing_facilities 21.85 32.69 hosp_beds_1k 2.38 2.51 life_expectancy 73.35 9.07 human_development_index 62.79 28.57 smokers 14.34 12.75 new_tests_cum_per_1000 539.63 1904.57 new_vaccinations_cum_per_1000 309.12 590.76	afflicted_rate 23.88 90.45 0.00 reproduction_rate 0.79 0.39 0.00 new_tests 234.60 840.38 0.00 new_vaccinations 141.02 293.40 0.00 stringency_index 45.81 24.36 0.00 population_density 443.11 2091.15 0.00 median_age 27.51 12.83 0.00 aged_65_older 7.88 6.49 0.00 gdp_per_capita 17632.21 20530.92 0.00 extreme_poverty 7.84 16.74 0.00 cardiovasc_death_rate 226.04 135.38 0.00 diabetes_prevalence 7.51 4.59 0.00 handwashing_facilities 21.85 32.69 0.00 hosp_beds_1k 2.38 2.51 0.00 life_expectancy 73.35 9.07 0.00 numan_development_index 62.79 28.57 0.00 new_tests_cum_per_1000 539.63 1904.57 0.00 new_vaccinations_cum_per_1000 309.12 590.76	afflicted_rate 23.88 90.45 0.00 1379.93 reproduction_rate 0.79 0.39 0.00 1.39 new_tests 234.60 840.38 0.00 12479.18 new_vaccinations 141.02 293.40 0.00 1995.94 stringency_index 45.81 24.36 0.00 95.61 population_density 443.11 2091.15 0.00 20546.77 median_age 27.51 12.83 0.00 48.20 aged_65_older 7.88 6.49 0.00 27.05 gdp_per_capita 17632.21 20530.92 0.00 116935.60 extreme_poverty 7.84 16.74 0.00 77.60 cardiovasc_death_rate 226.04 135.38 0.00 724.42 diabetes_prevalence 7.51 4.59 0.00 23.36 handwashing_facilities 21.85 32.69 0.00 99.00 hosp_beds_1k 2.38 2.51 0.00 36.75 human_development_index 62.79 28.57 0.00 45.95 <

```
## cardiovasc_death_rate_norm
                                     0.00
                                              1.00 -1.67
                                                              3.68
                                                                         5.35
                                              3.24 0.00
## gdp_per_capita_log
                                     8.19
                                                             11.67
                                                                        11.67
## Pooling Model
##
## Call:
## plm(formula = world_semester_df$afflicted_rate ~ world_semester_df$stringency_index +
##
       world_semester_df$aged_65_older + world_semester_df$gdp_per_capita +
       world semester df$extreme_poverty + world semester df$diabetes_prevalence +
##
##
       world semester df$smokers + world semester df$handwashing facilities +
       world semester df$life expectancy + world semester df$human development index,
##
       data = world_semester_df, model = "pooling", index = c("location",
##
           "date"))
##
##
## Unbalanced Panel: n = 206, T = 4-5, N = 1029
##
## Residuals:
##
        Min.
               1st Qu.
                                   3rd Qu.
                          Median
                                                Max.
## -112.7418 -23.1967
                         -5.0964
                                    9.4482 1292.0280
##
## Coefficients:
##
                                                Estimate Std. Error t-value
## (Intercept)
                                             -1.5621e+01 2.7556e+01 -0.5669
## world_semester_df$stringency_index
                                              2.3950e-01 1.1939e-01 2.0061
## world_semester_df$aged_65_older
                                              4.2226e+00 5.8789e-01 7.1826
## world semester df$gdp per capita
                                              5.9982e-05 1.6397e-04 0.3658
                                             -5.7901e-02 1.8417e-01 -0.3144
## world_semester_df$extreme_poverty
## world_semester_df$diabetes_prevalence
                                             -1.4533e+00 6.0902e-01 -2.3863
## world semester df$smokers
                                              6.0482e-01 2.7379e-01 2.2090
## world_semester_df$handwashing_facilities -1.9429e-01 8.8908e-02 -2.1853
## world_semester_df$life_expectancy
                                              4.2146e-02 3.6933e-01 0.1141
## world semester df$human development index -3.0913e-02 1.3425e-01 -0.2303
##
                                              Pr(>|t|)
## (Intercept)
                                               0.57091
## world semester df$stringency index
                                               0.04511 *
## world_semester_df$aged_65_older
                                             1.319e-12 ***
## world_semester_df$gdp_per_capita
                                               0.71459
## world_semester_df$extreme_poverty
                                               0.75330
```

```
## world_semester_df$diabetes_prevalence
                                              0.01720 *
## world_semester_df$smokers
                                               0.02739 *
## world_semester_df$handwashing_facilities
                                              0.02909 *
## world_semester_df$life_expectancy
                                              0.90917
## world_semester_df$human_development_index
                                              0.81793
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:
                           8410500
## Residual Sum of Squares: 7116200
## R-Squared:
                  0.15389
## Adj. R-Squared: 0.14642
## F-statistic: 20.5928 on 9 and 1019 DF, p-value: < 2.22e-16
```