

# Cross Section Assignment

Jacques Rossouw<sup>a</sup>

<sup>a</sup>*Stellebosch, Western Cape, South Africa*

---

---

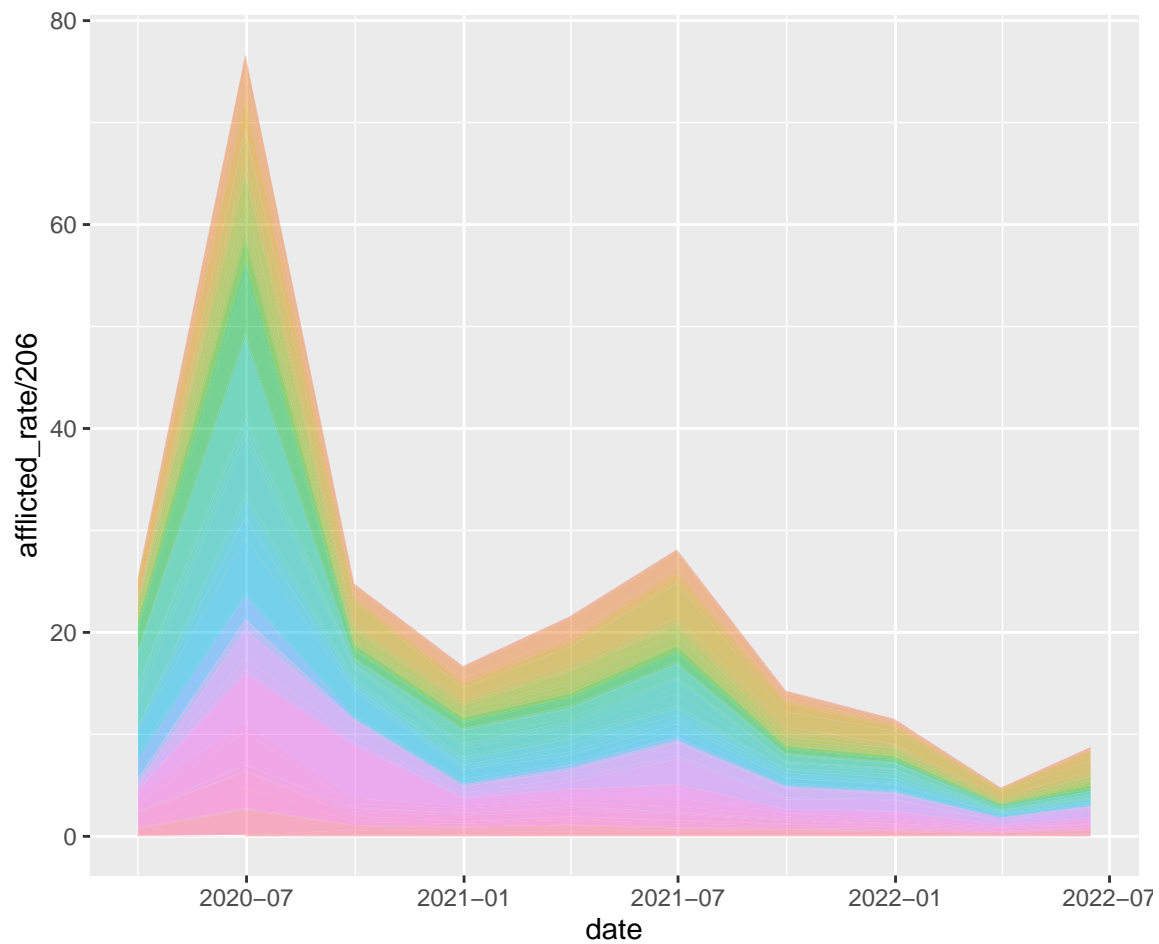
```
## # A tibble: 7 x 1
##   continent
##   <chr>
## 1 Asia
## 2 <NA>
## 3 Europe
## 4 Africa
## 5 North America
## 6 South America
## 7 Oceania
```

[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

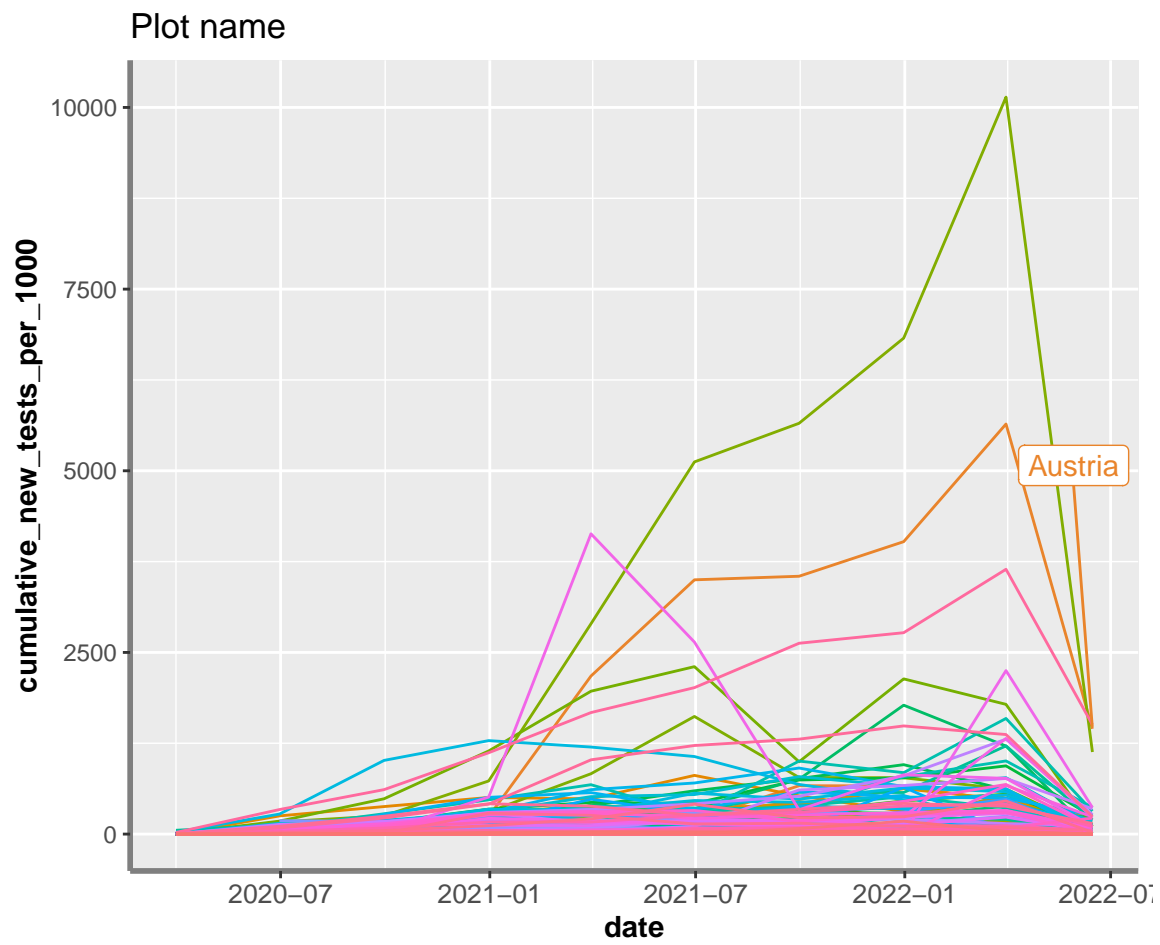
##	mean	sd	min	max	range
## afflicted_rate	23.37	88.83	0.00	1479.96	1479.96
## reproduction_rate	0.77	0.44	-0.01	2.06	2.08
## new_tests	117.76	460.31	0.00	10142.16	10142.16
## new_vaccinations	70.79	166.06	0.00	1689.38	1689.38
## stringency_index	44.81	25.15	0.00	99.06	99.06

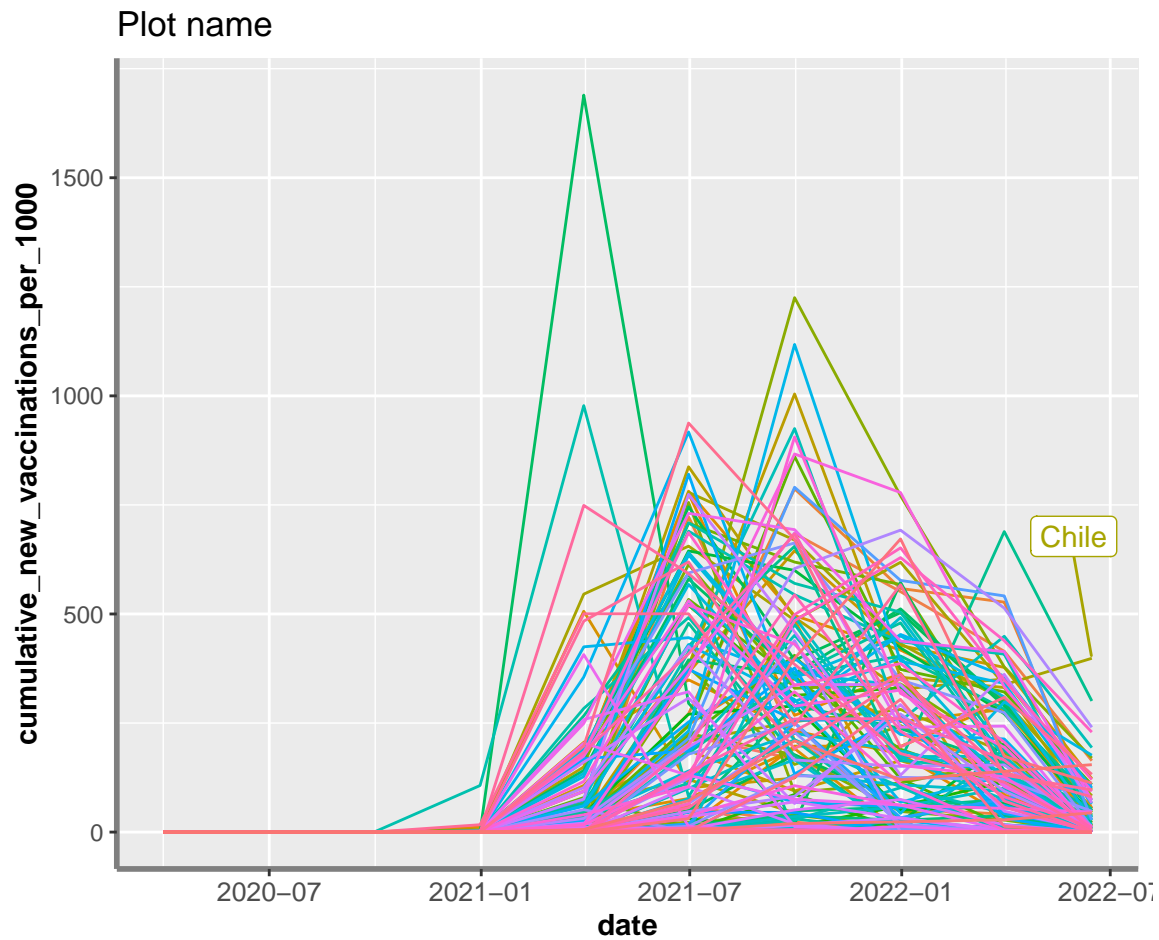


## 2. Scaling

##	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

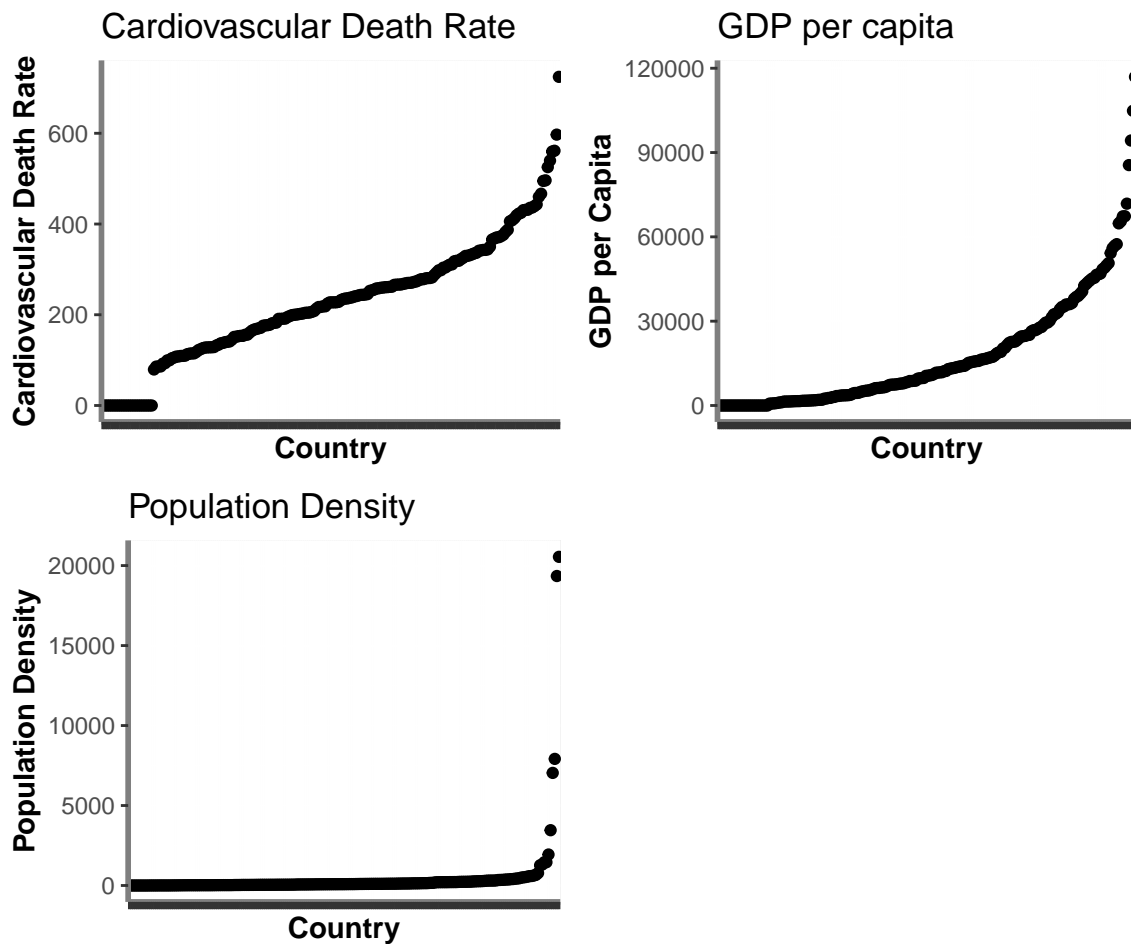
---

```
## 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



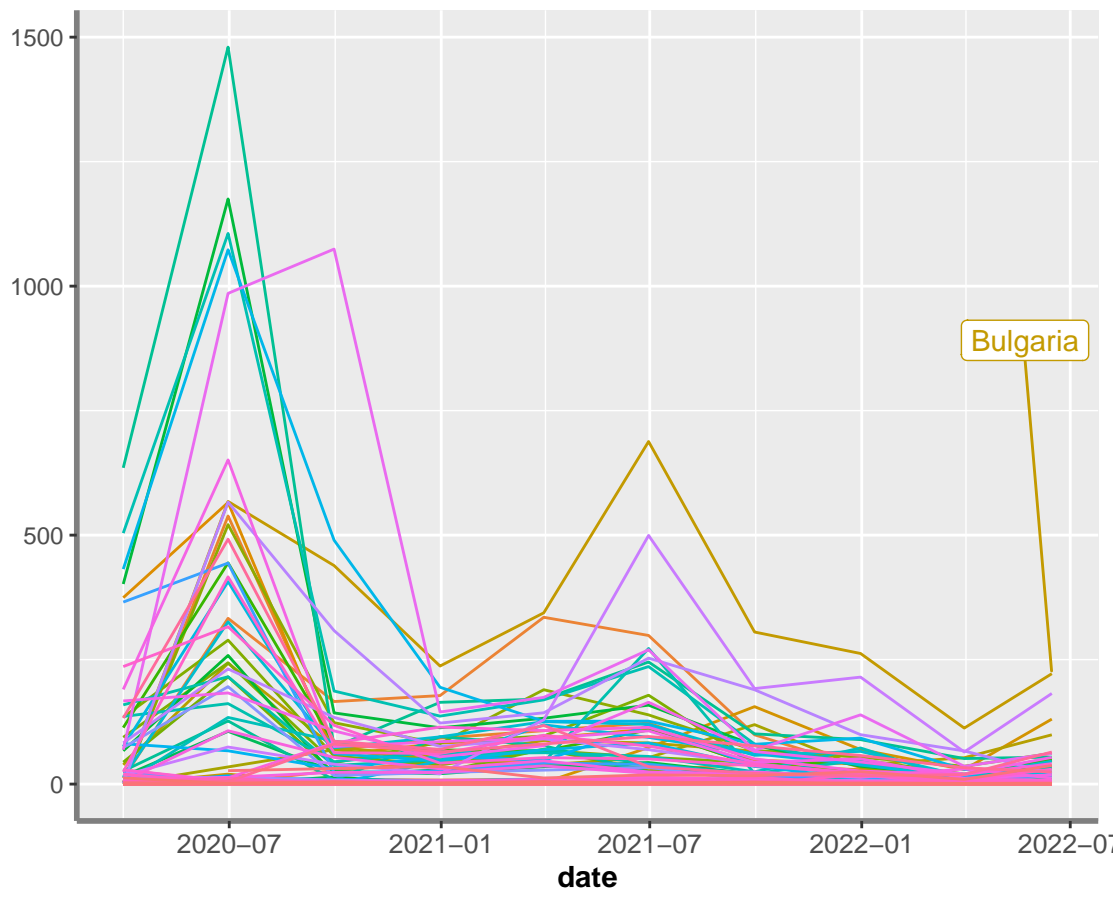
Now we can check all the descriptive stats for all the columns

```
##                                mean      sd    min      max      range
```

---

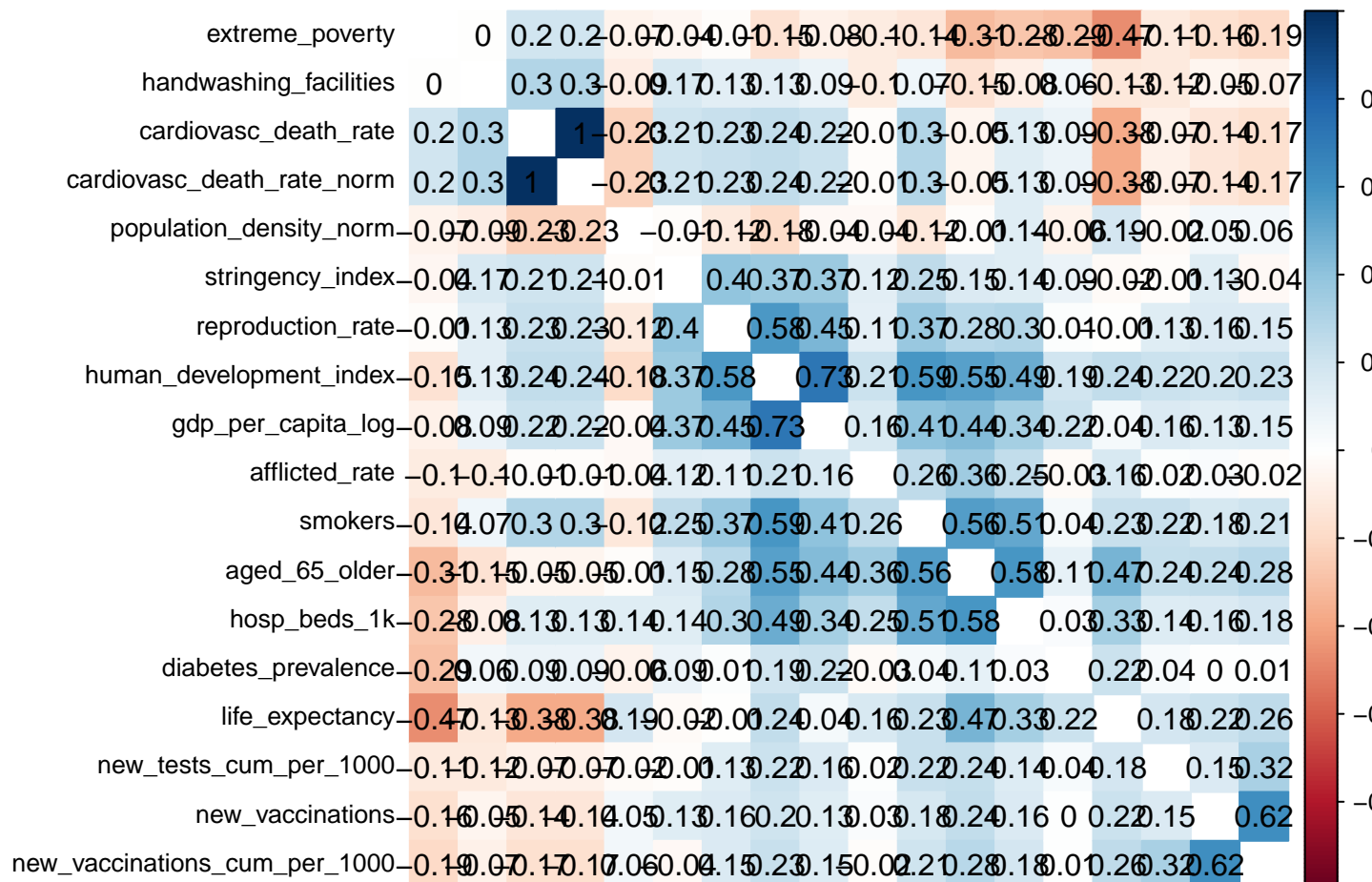
---

## 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
## 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



### 4. Regressions

#### 4.1. OLS

```
## 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
##
```



---

```

## 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 '**' 0.01 '*' 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}
## \hline \hline
##   & \multicolumn{1}{c}{\textit{Dependent variable:}} & \\
## \cline{2-2}
## \hline \hline
##   stringency\_index & 0.429$^{***}$ & \\
##   & (0.077) & \\
##   & & \\
##   Constant & 4.162 & \\
##   & (3.981) & \\
##   & & \\
## \hline \hline
## Observations & 2,050 &

```

---

---

```
## R2 & 0.015 \\
## Adjusted R2 & 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
##   term                estimate std.error statistic  p.value conf.low conf.high
##   <chr>              <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)      -1.07e+1  4.25      -2.51  1.21e- 2 -1.90e+1 -2.34
## 2 stringency_index  2.44e-1  0.0785     3.11  1.91e- 3  9.02e-2  0.398
## 3 smokers          1.57e+0  0.156     10.0  3.41e-23  1.26e+0  1.87
## 4 handwashing_faciliti~ -2.90e-1  0.0608    -4.77  1.95e- 6 -4.09e-1 -0.171
## 5 gdp_per_capita    3.90e-4  0.0000989  3.95  8.22e- 5  1.96e-4  0.000584
```

```
## 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:
##      Min.    1st Qu.    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 '**' 0.01 '*' 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
## 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

---

```

## cardiovasc_death_rate_norm      0.00      1.00 -1.67      3.68      5.35
## gdp_per_capita_log              8.19      3.24  0.00     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.      Median      3rd Qu.      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$gdp_per_capita    5.9982e-05  1.6397e-04  0.3658
## world_semester_df$extreme_poverty   -5.7901e-02  1.8417e-01 -0.3144
## 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 '**' 0.01 '*' 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
```