Datafest 2020

lemonchocolatecheesecake

Background

Our members mostly come from Singapore, where an outbreak of the dengue virus is currently happening concurrent to the COVID-19 situation.

Local news articles predict that this outbreak would be the most significant in recent years because:

1) More people being around potential mosquito breeding sites due to lockdown measures
2) The serotype of dengue is dominating this outbreak has low community immunity
3) An increasing temperature as "dengue season" from June to October comes

Our investigation is whether a relationship between the dengue outbreak and lockdown measures due to COVID-19 exist.

Tests and Results

Residuals:
```
    Min      1Q  Median      3Q     Max
-3.2953 -0.3051  0.0643  0.4666  1.2883
```

Coefficients:
```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.801e+01  2.256e-01 124.143   <2e-16 ***
data[, 2]   5.412e-05  6.444e-04   0.084    0.933
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.7254 on 69 degrees of freedom
Multiple R-squared: 0.0001022, Adjusted R-squared: -0.01439
F-statistic: 0.007055 on 1 and 69 DF, p-value: 0.9333

Residuals:
```
    Min      1Q  Median      3Q     Max
-6.621 -3.563 -1.353   2.279 20.632
```

Coefficients:
```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.86063    2.18776   4.050 0.000112 ***
data4[, 2]  -0.04653    0.03981  -1.169 0.245730
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 5.331 on 86 degrees of freedom
Multiple R-squared: 0.01564,   Adjusted R-squared: 0.004189
F-statistic: 1.366 on 1 and 86 DF, p-value: 0.2457

Residuals:
```
    Min      1Q  Median      3Q     Max
-3.3508 -0.5252  0.0649  0.5803  1.7957
```

Coefficients:
```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.741e+01  7.107e-02 385.670  < 2e-16 ***
data2[, 2]  9.867e-04  2.555e-04   3.862 0.000134 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.8473 on 340 degrees of freedom
Multiple R-squared: 0.04203,   Adjusted R-squared: 0.03922
F-statistic: 14.92 on 1 and 340 DF, p-value: 0.0001344

Residuals:
```
    Min      1Q  Median      3Q     Max
-6.524 -4.484 -1.299   2.786 29.601
```

Coefficients:
```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.654205   0.459296  14.488   <2e-16 ***
data3[, 2]  -0.002413   0.001576  -1.532    0.126
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 5.688 on 412 degrees of freedom
Multiple R-squared: 0.005662, Adjusted R-squared: 0.003249
F-statistic: 2.346 on 1 and 412 DF, p-value: 0.1264

Conflicting results of 2 regression tests between temperature and no. of cases in different years

Results of 2 regression tests between rainfall and no. of cases in different years, showing no correlation

The dengue virus has four serotypes; a serotype maintaining dominance over many years has the community gaining herd immunity to it. From year to year, changes in the dominant serotype are common. We wanted to test to see if there was a correlation between the current outbreak and outbreaks in other years.

```
        Pearson's Chi-squared test

data:  data6
X-squared = 63.388, df = 9, p-value = 2.967e-10
```

Result of test in 2013 where the dominant serotype changed after a long period of time

```
        Pearson's Chi-squared test

data:  data7
X-squared = 35.498, df = 9, p-value = 4.866e-05
```

```
        Pearson's Chi-squared test

data:  data8
X-squared = 122.16, df = 9, p-value < 2.2e-16
```

Result of tests in 2015 & 2019 where there was no extended period of time in which one serotype was dominant

Conclusion

We were unable to draw any useful results due to our tests being inconclusive