Proudly presented by: Sin Myeong Eun, Loh Yuan Jian Keith, Harvey Kristanto Lauw

# DENGUE PALS

## OVERVIEW

### PROBLEM:

Ever since Singapore first went into circuit breaker period on 7th April, the number of dengue cases spiked up to the extent that it crossed the 10,000 mark in June.
As the majority of the fatalities observed in 2020 occurred in the active dengue clusters, it is essential for us to focus on the identified cluster areas so that better measures can be implemented for these areas.

### MOTIVATION:

Our project goal is to study the dengue clusters across Singapore from February 2020 to July 2020, which is inclusive of the circuit breaker period. Our application will allow users to interactively analyse how dengue spread in Singapore during the time period of February 2020 to July 2020. We hope that the results will be able to assist the relevant authorities in preparing the nation for the more dengue-prone season in the near future.

## APPROACH

### DATA WRANGLING

1. Reformatted "Date" attribute in YYYY-MM-DD and converted to Epidemic-week ("eweek") instead
2. Transform CRS to EPSG 3414 (SVY 21)
3. Check for NA values, invalid geometry % duplicated points
4. Converted coordinates to geometry points

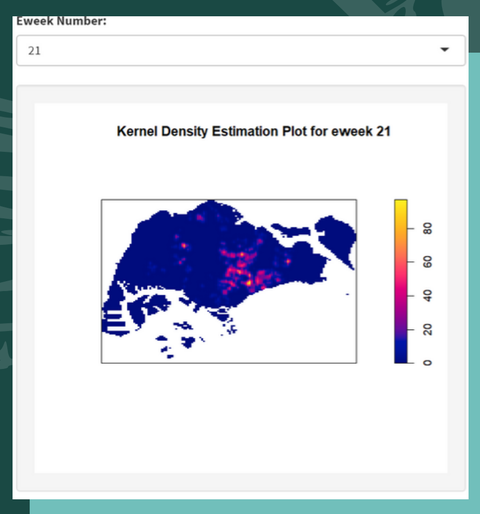
### METHODOLOGIES USED

1. Spatial Point Patterns Analysis: 1st & 2nd Order
2. Geographically weighted regression
3. Spatio-temporal Point Pattern Analysis

## RESULT & KEY FINDINGS

### 1ST ORDER POINT PATTERN ANALYSIS

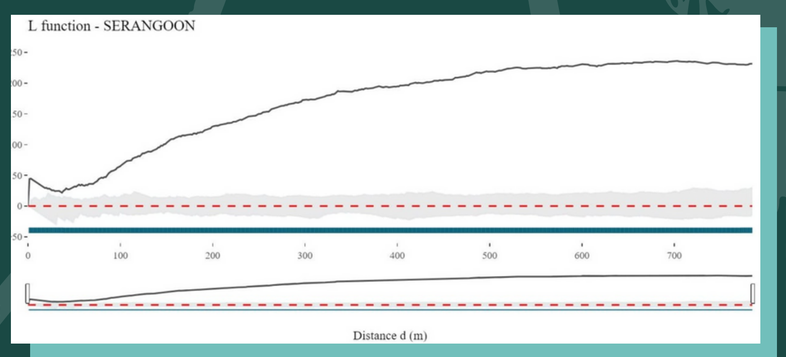
Upon studying the e-weekly and monthly increase in dengue cases across SG from the barcharts, we can delve further into inspect the Kernel Density Estimation plots for specific e-week/month to observe how the "hotspots" of dengue spread change. Moreover, we can deep-dive into the KDE plot of each planning area.



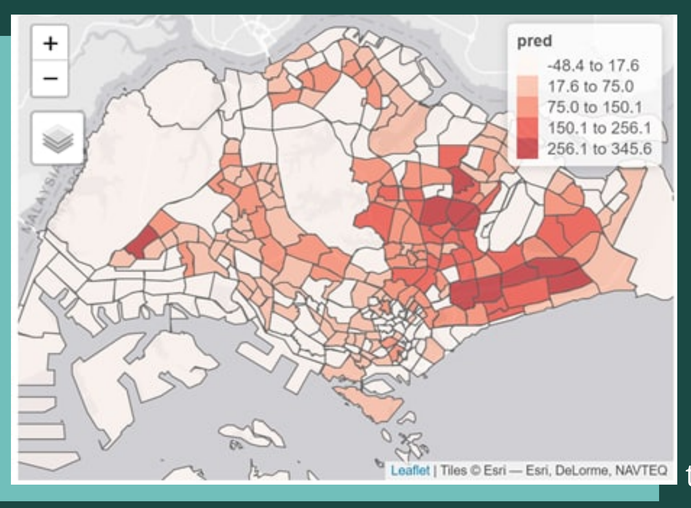

### 2ND ORDER POINT PATTERN ANALYSIS

Using different combinations of filters such as Planning Area, Week of 2020, randomness test of dengue cases can be carried out.
H0 = The distribution of dengue cases at MARINE PARADE in Week 28 are randomly distributed.
H1 = The distribution of dengue cases at MARINE PARADE in Week 28 are not randomly distributed.
The null hypothesis will be rejected if p-value is smaller than alpha value of 0.05 (95% Confidence Interval).

The point pattern of the dengue cases largely lies above the envelope and resembles clustering throughout, hence spatial clustering is statistically significant and we reject the null hypothesis (H0).

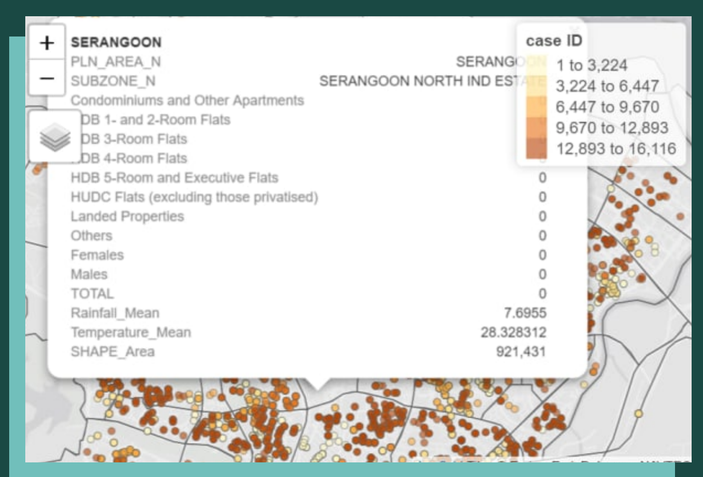



### GEOGRAPHICALLY WEIGHTED REGRESSION (GWR)



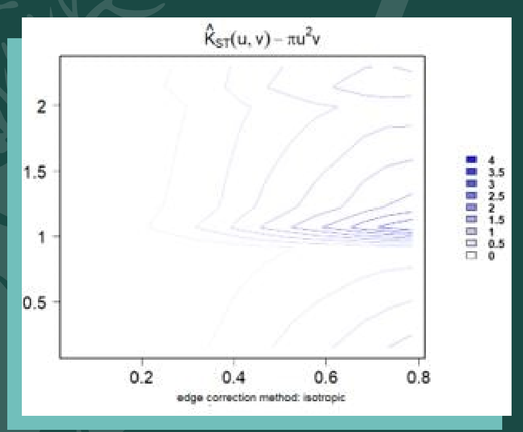

"Predicted Number of Dengue Cases" plot shows predicted dengue outputs based on the historical data from February 2020 to July 2020, which was computed by inputting highly significant variables into the model and training it.

The "Breakdown of Regression outcome by Subzone" plot displays all the numerical values of every variable which was attributed to the training of the model as well as the predicted dengue cases.

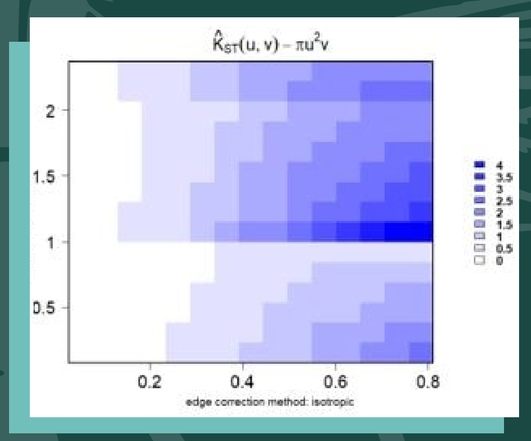


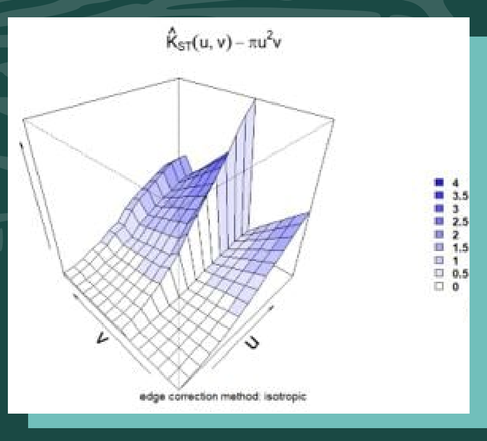

### SPATIO-TEMPORAL POINT PATTERN ANALYSIS

On top of spatial point patterns of the dengue cases in Singapore, we will be looking at the "time" aspect as well. The specific analysis method that we will be applying is the STIKhat function, which computes an estimate of the Space-Time Inhomogeneous K-function.




Looking at the plots, if the STIKhat value that is estimated through u and v variable is positive, the dengue cases to be analyzed show signs of clustering. If the STIKhat value is negative, the dengue cases represent a regular distribution. The intensity of the STIKhat value describes the degree of clustering, where a larger value represents a more concentrated cluster.




## FUTURE WORKS

**Logistic Regression analysis on the relationship between human demographics and dengue patients**

It would be another interesting finding to analyse if there's any specific human demographics that is often observed in dengue patients.

**More in-depth spatio-temporal analysis**

Coupled with a better understanding of stpp package, we can futher delve into this particular geospatial aspect to study the spatio-temporal analysis through the use of Anisotropic Space-Time inhomogenous K-function as well as Space-Time Inhomogenous Pair Corrleation LISA function.

SPECIAL THANKS TO PROF KAM TIN SEONG FOR HIS ENDLESS SUPPORT