DengAI: Predicting Disease Spread

**Group C :** G.A Akila Shashith Ganegoda, M.L.N.Kaushalee, K. G. Inuka Manmitha, R Nilakshan, S.P.A.G.T Rajapakshahorizontal line

# Introduction

Using environmental data collected by U.S. Federal Government agencies, can you predict the number of dengue fever cases reported each week in San Juan, Puerto Rico and Iquitos, Peru?

* Dengue fever is a mosquito-borne disease in tropical and subtropical parts of the world.
* Symptoms of dengue fever can range from mild to severe, and in severe cases can lead to death.
* Climate change will likely produce distributional shifts with significant public health implications worldwide.
* Dengue fever has spread in recent years, with many cases occurring in Latin America.
* Predicting the number of dengue cases weekly in specific locations can help improve research initiatives and resource allocation to help fight life-threatening pandemics.

# Problem Definition

Using environmental data to aid in public health initiatives, predict the number of dengue fever cases reported weekly in San Juan, Puerto Rico, and Iquitos, Peru.

# Description of the Dataset

* The goal is to predict total cases for each (city, year, weekofyear) in the test set.
* The test set has data for two cities, San Juan and Iquitos, spanning 5 and 3 years, respectively.
* The test data are sequential and non-overlapping with any training data.
* The features include climate data, precipitation measurements, dew point temperature, air temperature, relative humidity, specific humidity, and vegetation index.

|  |  |
| --- | --- |
| * city – City abbreviations: sj for San Juan and iq for Iquitos * week\_start\_date – Date given in yyyy-mm-dd format * station\_max\_temp\_c – Maximum temperature * station\_min\_temp\_c – Minimum temperature * station\_avg\_temp\_c – Average temperature * station\_precip\_mm – Total precipitation * station\_diur\_temp\_rng\_c – Diurnal temperature range * precipitation\_amt\_mm – Total precipitation * ndvi\_se – Pixel southeast of city centroid * ndvi\_sw – Pixel southwest of city centroid * ndvi\_ne – Pixel northeast of city centroid * ndvi\_nw – Pixel northwest of city centroid | * reanalysis\_sat\_precip\_amt\_mm – Total precipitation * reanalysis\_dew\_point\_temp\_k – Mean dew point temperature * reanalysis\_air\_temp\_k – Mean air temperature * reanalysis\_relative\_humidity\_percent – Mean relative humidity * reanalysis\_specific\_humidity\_g\_per\_kg – Mean specific humidity * reanalysis\_precip\_amt\_kg\_per\_m2 – Total precipitation * reanalysis\_max\_air\_temp\_k – Maximum air temperature * reanalysis\_min\_air\_temp\_k – Minimum air temperature * reanalysis\_avg\_temp\_k – Average air temperature * reanalysis\_tdtr\_k – Diurnal temperature range |

**Proposed Solution**

We propose to utilise the dataset for a predictive task, such as a regression task, to forecast the total cases of dengue fever predictions (city, year, week of year) and the predicted total cases.

**Mapping the Problem:** The solution involves mapping environmental factors to disease incidence, aiding resource allocation and research efforts to combat pandemics.

horizontal line

(Resource: <https://www.drivendata.org/competitions/44/dengai-predicting-disease-spread/page/80/>)