# Exploratory Data Analysis Report: Dengue Fever Outbreaks in Sri Lanka

ASHUTOSH KUMAR

## 1 Introduction

This report summarizes the key findings from an Exploratory Data Analysis (EDA) of dengue fever outbreaks in Sri Lanka from 2007 to 2024. The analysis aims to uncover patterns and relationships between dengue cases and various environmental factors to aid in predicting future outbreaks.

## 2 Temporal Trends

### 2.1 Overall Trend

- A significant spike in dengue cases occurred in 2017, with weekly cases exceeding 10,000 for 3-4 weeks.
- Post-2017, there has been an increase in the baseline number of cases compared to pre-2016 levels.
- Periodic outbreaks occur approximately every 1-2 years, with notable peaks in 2019, early 2020, and 2023.

### 2.2 Seasonality

- Strong seasonal patterns are observed, with peaks typically occurring in June, July, December, and January.
- These peak months correspond with Sri Lanka's rainy seasons, suggesting a strong correlation between increased rainfall and dengue outbreaks.

## 3 Spatial Distribution

### 3.1 High-Risk Districts

- Colombo district consistently reports the highest number of dengue cases, particularly in June and July.
- Gampaha and Ratnapura districts also show high dengue incidence, with patterns similar to Colombo but at slightly lower levels.

### 3.2 Low-Risk Areas

- Northern districts such as Jaffna, Kilinochchi, Mannar, and Mullaitivu generally report lower dengue cases year-round.
- This pattern may be due to lower population density or different environmental conditions in these areas.

## 4 Environmental Factors

### 4.1 Temperature

- Dengue cases peak when average maximum temperatures are between 28°C and 32°C.
- Warmer temperatures, particularly in the range of 25°C to 30°C, create favorable conditions for mosquito breeding and virus transmission.

### 4.2 Precipitation

- A clear increase in dengue cases is observed when rainfall is between 50mm and 200mm.
- Excessive rainfall (beyond 300mm) appears to reduce dengue cases, possibly due to the flushing out of mosquito breeding sites.

### 4.3 Wind Speed

- Fewer dengue cases are reported when average wind speeds exceed 25 km/h.
- Cases tend to be more prevalent in areas with average wind speeds between 15-23 km/h, suggesting that calmer conditions favor mosquito activity.

## 5 Lag Effects and Predictors

- Total precipitation and rainfall from the previous week show the highest positive correlation (0.096) with current dengue cases.
- Average minimum temperature from the previous week also shows a positive correlation (0.061) with dengue cases.
- These lagged correlations suggest that monitoring rainfall and temperature data could help predict and prevent dengue outbreaks.

## 6 Clustering Analysis

K-means clustering revealed three distinct groups based on temperature and dengue case numbers:

1. Low temperature (below 25°C) and low case numbers
2. Moderate to high temperatures (25°C-32°C) with the highest number of cases
3. Extreme temperatures (above 32°C) with relatively fewer cases

This clustering reinforces the importance of temperature as a key variable in predicting dengue outbreaks, with the 25°C-32°C range being particularly conducive to outbreaks.

# 7 Conclusion and Recommendations

The EDA reveals complex interactions between environmental factors and dengue outbreaks in Sri Lanka. Key findings include:

1. The critical role of temperature (25°C-32°C range) and moderate rainfall (50mm-200mm) in dengue transmission.
2. Strong seasonal patterns with peaks in June-July and December-January.
3. Spatial variation in risk, with Colombo, Gampaha, and Ratnapura being high-risk areas.
4. Lagged effects of rainfall and temperature on dengue cases.

Based on these insights, I recommend:

1. Implementing targeted preventive measures in high-risk districts, especially during peak seasons.
2. Developing predictive models that incorporate lagged environmental variables, particularly rainfall and temperature.
3. Focusing mosquito control efforts during periods of moderate rainfall and optimal temperature ranges.
4. Conducting further research to understand the factors contributing to lower dengue incidence in northern districts.

These findings provide a foundation for developing more accurate predictive models and effective strategies for dengue outbreak prevention in Sri Lanka.