Leveraging machine learning and time series analysis for early detection and prediction of disease outbreaks in Boston

CSTE

Stories FROM THE Field

CONTRIBUTOR: *Aanchal Gupta, MS, GIS Manager, Boston Public Health Commission*

CATEGORY: **Syndromic Surveillance**

The Boston Public Health Commission used machine learning processes and syndromic surveillance data to identify symptom anomalies and geographic hotspots in their data, allowing them to respond with resources where they were most needed.

The "What"

Symptom-based detection of disease outbreaks is a crucial element of public health surveillance, facilitating early warning and rapid response to emerging health threats. By promptly identifying symptoms, health authorities can allocate resources more effectively, ensuring that the most affected areas receive the necessary medical attention and supplies. This approach enhances public health messaging, guiding individuals on when to seek care and how to prevent transmission. The Boston Public Health Commission's (BPHC) goal is to develop a machine learning (ML) and time series framework that enables early detection of potential outbreaks in Boston. The framework will leverage syndromic surveillance to promote health equity, enhance early warning systems for diseases, and uncover age-specific insights within the data to inform targeted public health interventions and outreach efforts.
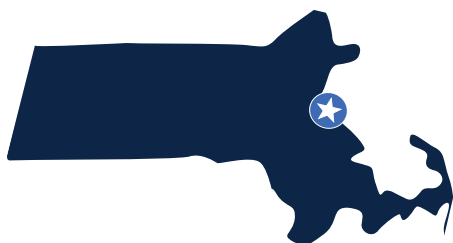
The "So What"

Preliminary analysis indicates that ML methods were more likely to flag or identify potential anomalies than the traditional time series methods. When segmented by demographic categories, reported symptoms were predominantly observed among residents aged 65 and older, likely reflecting the higher frequency of emergency department visits within this age group. Anomalies in reported symptoms were observed among Asian residents, although the findings were not statistically significant. Along with using syndromic surveillance to identify emerging diseases, dashboards were created to track reportable diseases and acute illnesses as part of this project. Heatmaps were created to highlight potential hotspots, making it easier to understand geographic patterns and trends. Such visualizations are invaluable for public health officials, enabling them to allocate resources efficiently and implement targeted interventions where they are required.

The "Now What"

In future iterations of this project, analysis should be conducted at the zip code level to identify and visualize outbreak concentrations more effectively. Based on feedback from public health officials, breaking down the dashboards by type of symptoms to identify disease transmission type can enhance intervention response.

→ More stories are available at **stories.cste.org**

SUBMITTED MARCH 2025