

# BIG DATA VISUALIZATION

PROCESS

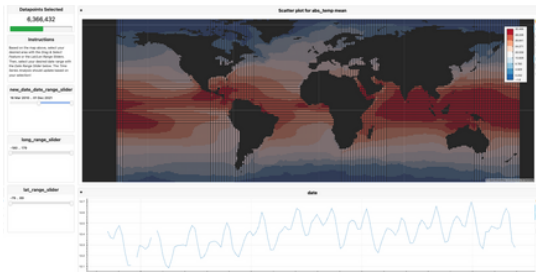
• PICTURE •

PREDICT

## PROBLEM AND PROJECT GOALS

Many researchers need interactive, large-scale visualizations to help observe trends to make and/or support hypotheses. Not having the right amount of computing power can make it difficult to load, transform, and visualize their data. This is where we come in! We have created two dashboards, one that finds a correlation between snowfall and sea surface temperatures and another that uses climate variables to predict malaria outbreaks.

## SNOWFALL

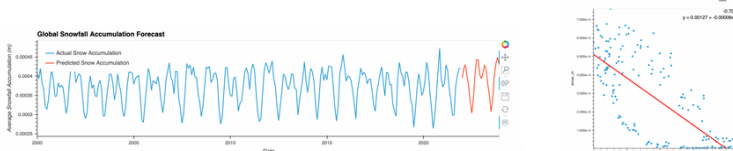


### KEY FEATURES

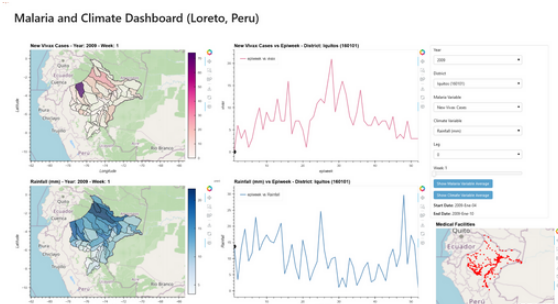
- Global snowfall and sea surface temperature time series analysis for 2000-2021
- Interactive dashboards allowing users to select time ranges and regions, updating time series analysis and scatterplot with lag
  - Users can investigate which part of the oceans will correlate most to increases in snowfall on certain parts of land.

### RESULTS

- Observed lag of 5-6 months typically results in highest correlation
- Noticed a phenomenon where the scatterplot went from a circular to a linear shape as the lag changes
- Machine learning: attempted ARIMA, SARIMA, and Holt Winters to forecast snowfall



## MALARIA



### KEY FEATURES

- Select tools for Malaria case data from Loreto, Peru spanning 2009-2021
- Select tools for various types of climate data from Loreto
- A year select and week slider
- A "lag" feature that displays climate data from the current week - given lag (ranging from 0-24 weeks) so you can visualize the effect on current case counts
- The ability to compute averages of case and climate data with standard deviations
- A mini-map featuring health facilities across Loreto



SCAN THE QR CODE FOR DEMO VIDEOS