Zika, a “disease of the poor”?

A project mapping confirmed instances of the Zika virus in 2016 in countries in South and Central America

A project by Daria Locher

Research summary

This project maps confirmed instances of the Zika virus in 2016 in countries in South and Central America alongside several factors to determine which has the highest correlation.

The Zika virus is a mosquito-borne disease that can cause neurological effects including Guillain-Barré syndrome, as well as serious birth defects in pregnant women, including microcephaly. The first cases were identified in Africa in the 1940s, but there was an extreme outbreak across the globe in 2016, leading the World Health Organization to declare a Public Health Emergency of International Concern (PHEIC) for Zika-related microcephaly and neurological disorders.[1] Background research indicates that poverty and inequality had large effects on which populations were affected by the virus. A July 2016 article by the United Nations Population Fund points to Brazil as an example of where the virus hit hardest, and identifies marginalization, poverty, and poor healthcare systems as important factors.[2] Meanwhile, mosquito populations are found across the world, and converge around wet areas, including rainforests.[3]

The aim of this research was to answer the questions:

- Is there a notable correlation between occurrences of the disease across countries with the country's per capita GDP?
- Do other factors have a larger impact on Zika cases, for example amount of forest area within the country?

In an attempt to answer these questions, this study looks at reported Zika cases across Central and South America in 2016 in the first seven months of the WHO's Public Health Emergency. This research shows that Brazil, while singled out as a location with a especially large epidemic, was not the location in this region with the worst outbreak — rather, other, smaller countries had case loads that affected significantly larger portions of their populations. Additionally, this research reveals a much stronger correlation between Zika cases and forest coverage within countries than the cases with countries' income levels.

Context for this study

This data used in this report was initially collected for a university class project in the course "GIS in Area Studies" at Washington University in St. Louis, but this report was written in 2024 as an exercise of how to represent GIS data better using the design and analysis expertise I developed as a human rights researcher and graphic designer after college.

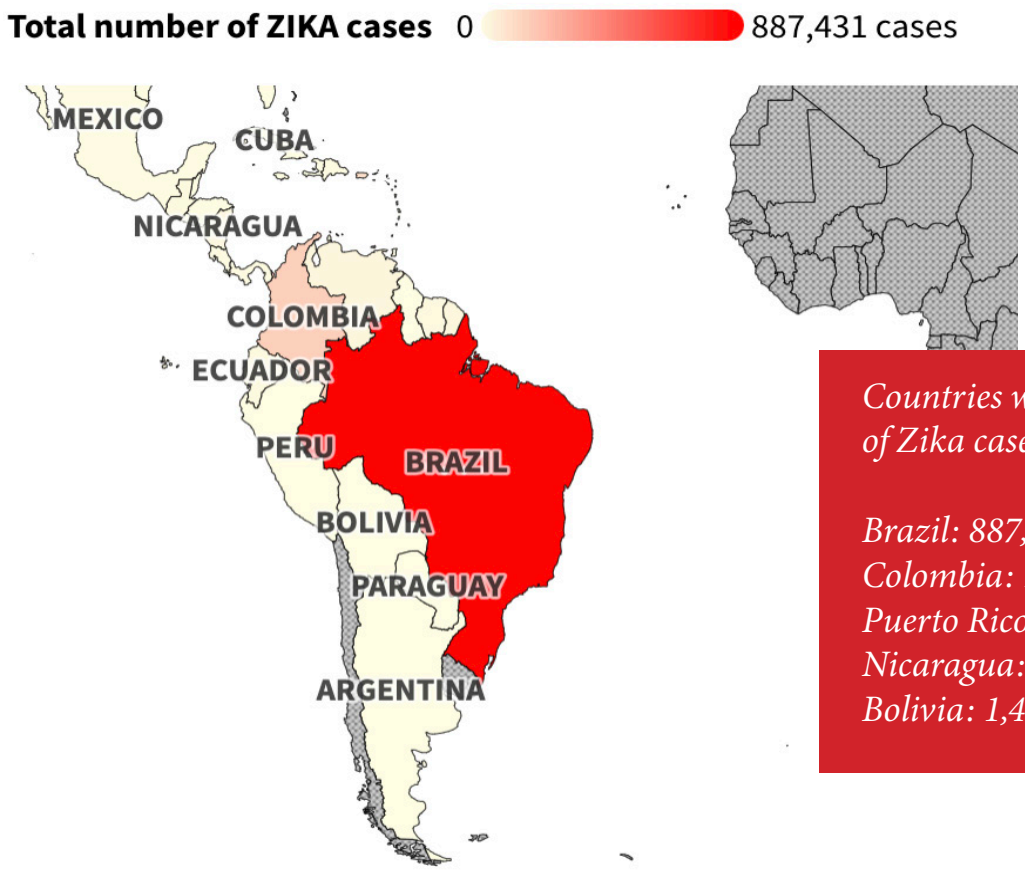
[1] Zika virus. World Health Organization, December 8, 2022.
[2] Poverty, inequality at the heart of the Zika outbreak. United Nations Population Fund. July 21, 2016.
[3] Where Mosquitoes Live. Center for Disease Control. April 20, 2024.

The data for the 2016 reported cases of Zika in Central and South America were sourced from the Humanitarian Data Exchange, which covers the reporting period from February 12, 2016 until September 15, 2016. I cleaned non-confirmed cases from the dataset to have more standardization between countries' data. The 2016 GDP per capita of the countries came from the United Nations' National Accounts Main Aggregates database and World Bank Group national accounts data. Data for the French overseas regions came from Statistica, and data for Saint Barthelemy from the Institut d'Émission des Départements d'Outre-Mer. Please note that the GDP data for Saint Martin was only available from 2014; and for Saint Barthelemy from 2010, while the rest of the data is sourced from the more-relevant year 2016. In all charts unless otherwise indicated, the French overseas regions of Martinique, Guadeloupe, and French Guinea are merged due to data issues.

Analysis

First, I mapped total confirmed cases of the disease by the country level, and immediately realized that looking at Zika cases purely by country-level totals is not useful, because high-population countries are overrepresented. In this first map, Brazil has by far the highest total confirmed cases, which tracks with the country having the largest population of these countries — double any other country, with Mexico (half the population) and Colombia (a quarter of the population) being the next highest-populated countries in the region.

ZIKA CASES IN 2016 - OVERALL CASES PER COUNTRY

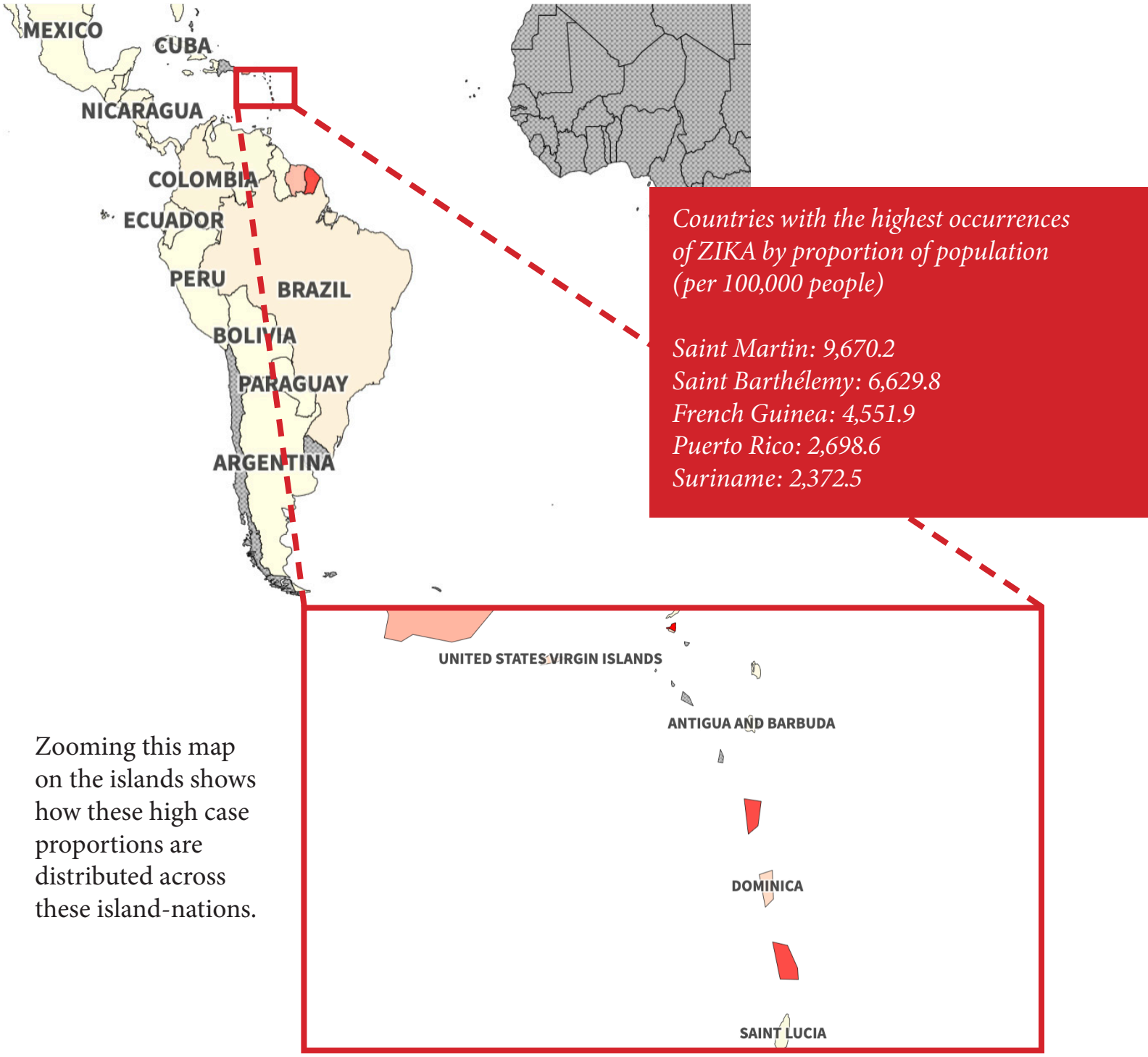


Countries with the highest occurrences of Zika cases in the region:

Brazil: 887,431
Colombia: 154,510
Puerto Rico: 95,756
Nicaragua: 12,908
Bolivia: 1,438

After dismissing this map due to the population size being too overwhelmingly strong a factor, I remapped the data by looking at the number of Zika cases per 100,000 people. This new mapping revealed that the countries with the highest proportion of Zika cases to be Saint Martin, Saint Barthélemy, and the French Guinea, respectively.

ZIKA CASES IN 2016 - CASES PER 100,000 PEOPLE

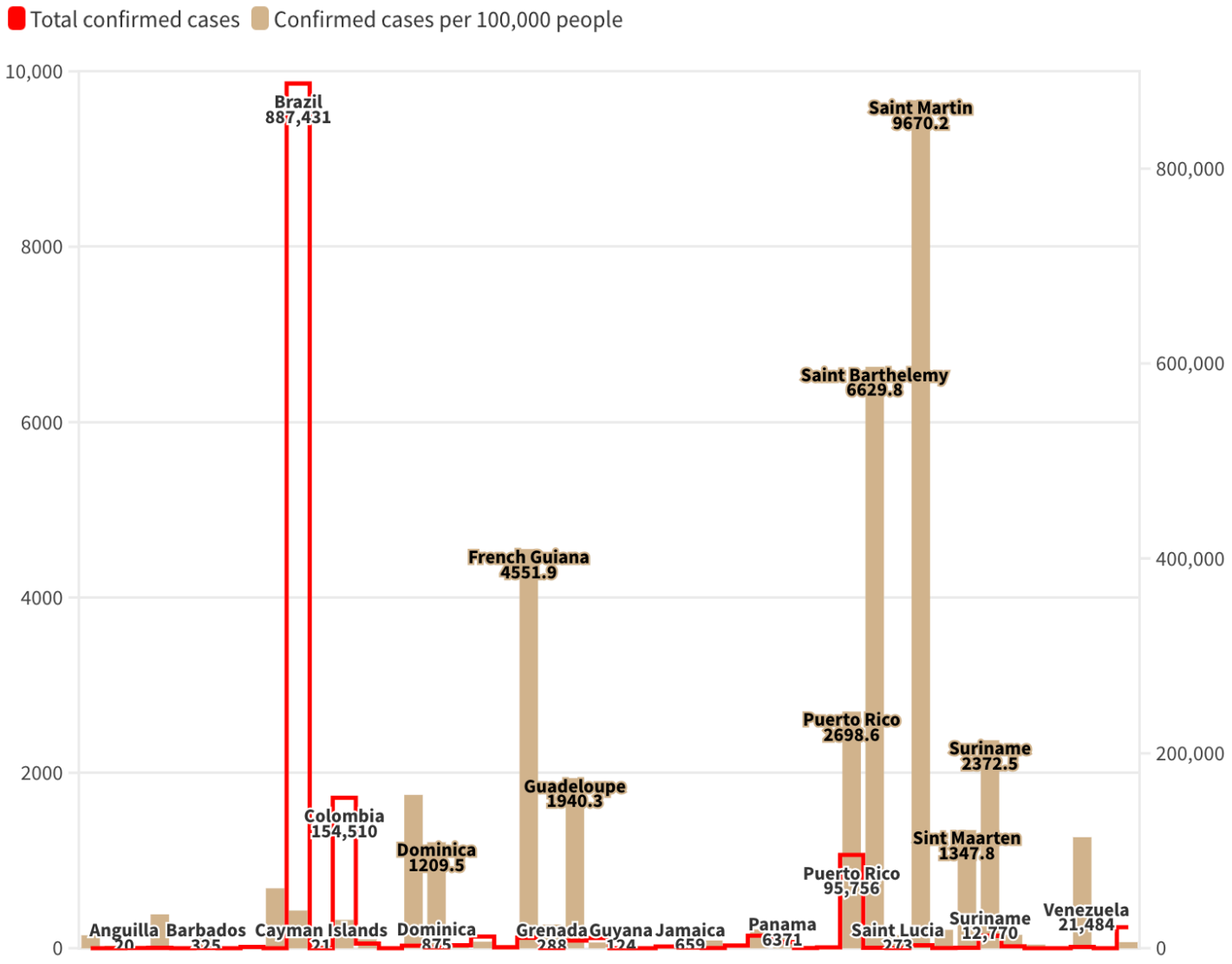


Zooming this map on the islands shows how these high case proportions are distributed across these island-nations.

The issue with looking at country-level total cases becomes even more clear when total cases and proportional cases are compared in a bar chart.

Brazil in particular has such a high number of total confirmed cases that a large number of other countries are flattened in the y-axis. When compared with proportional case numbers, not only do more countries become visible in the chart, but also Brazil is revealed as having a much less severe epidemic than other countries.

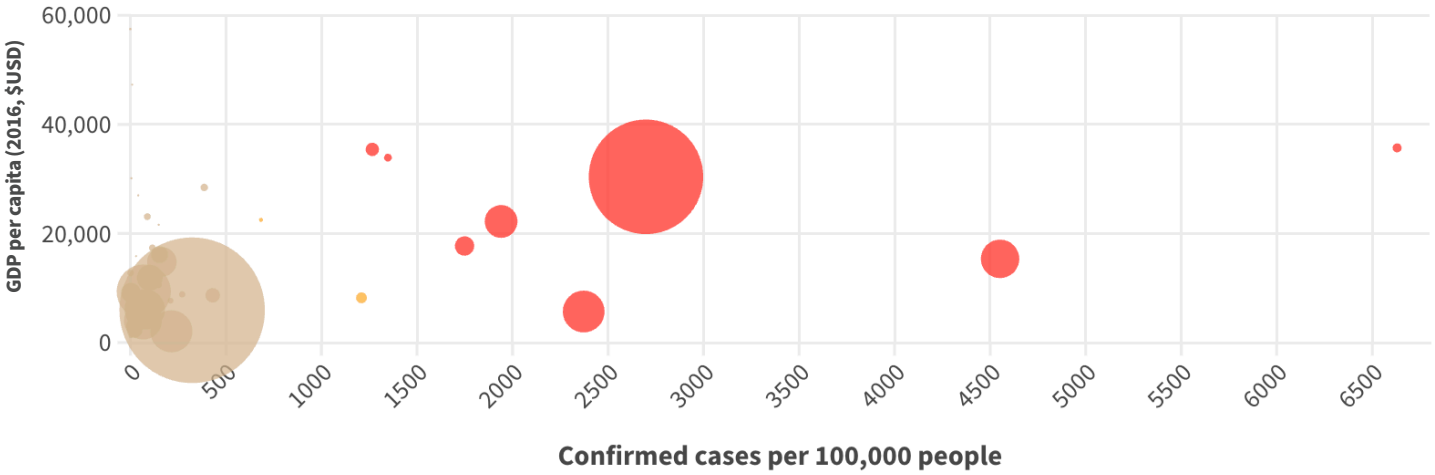
COMPARING COUNTRIES' 2016 TOTAL CONFIRMED ZIKA CASES WITH CASES PER 100,000 PEOPLE



Source: Humanitarian Data Exchange • Please note that the French overseas territories of Martinique, Guadeloupe, and French Guinea are separated into their own columns in this chart.

Now that the fallacy of looking at total case numbers per country has been dismissed, I can consider how two factors are correlated with Zika cases: the country's GDP or the amount of land area covered in forest. These were mapped on a chart comparing proportional case numbers, GDP per capita in countries, and percent of country covered in forest. The chart is color-coded to indicate countries with proportional numbers of cases that were especially high (70% or more above the region's average), medium, and low (30% or less than the average). The

COMPARING 2016 ZIKA CASES WITH GDP PER CAPITA



High/low numbers of cases Low Medium High

The size of the dots is determined by the total number of cases in each country.

Please note that a few extreme outliers were removed:
- The Cayman Islands from the GDP per capita chart (very high GDP; very low cases)
- Saint Martin from the GDP per capita chart (very high cases)

COMPARING 2016 ZIKA CASES WITH COUNTRIES' FOREST COVERAGE

High/low numbers of cases Low Medium High

bubbles in the chart are also sized with the total number of cases in different countries — the two largest bubbles are Brazil (tan, bottom left of chart) and Colombia (red, in the middle of the chart).

Looking at the chart only with your eyes, the strengths of correlations between these two factors is not clear. In fact, the data that stands out is the disparity between total case numbers in different countries. Nonetheless, forest coverage ends up being the stronger predictor of Zika cases. Looking at the strengths of the trend lines of these two relationships, the forest to case correlation has an r-squared value of 0.2809, and the GDP to case correlation — even after removing outliers — is only 0.0723. Given these relatively low values, it is evident that neither of these factors are entirely correlated with case numbers in these countries. Further study is required.

Issues with this analysis to be addressed in a later study

- The largest issue with this study is ignoring geopolitical and cultural considerations in these breakdowns, which surely had large impacts on how effectively the countries were able to minimize transmission of the disease.
- It would be interesting to look at why islands and small countries did so poorly using a qualitative analysis, comparing infrastructure and healthcare systems. The analysis by forest area should even out the impact of islands' wetlands on mosquito populations, but further analysis is needed to confirm this assumption.
- The case number data was based on country-reported stats, and there was a significant level of inconsistency between and within the datasets. It would have been useful to compare how the disease spread over time, but the datasets had significant outlier values when looking at reported cases by week. These were probably due to the countries' health systems not prioritizing consistent reporting while they were responding to the outbreak.
- The data visualization tool used, Flourish, did not specify between the French overseas regions of Martinique, Guadeloupe, and French Guinea. In a future study, these should ideally be separated out, since they have notable differences in income levels, forest area, and other factors that could have been expected to influence Zika cases in 2016.

Links to source material and interactive versions of the charts used in this paper can be found at this link: darlocher.wixsite.com/website/post/zika-a-disease-of-the-poor