Mapping Dengue Hotspots and Analyzing the Impact of Weather, Sanitation and Water Supply on Disease Spread in the Philippines: A Multi-Year Analysis

Presented by Hugo, Angelica H.

CPE15L: Cognate/Professional Course 1

Problem Statement

- Dengue fever remains a significant public health concern in the Philippines, with outbreaks occurring annually and widespread transmission across various regions.
- Despite ongoing efforts to control the disease, the persistence of dengue cases, especially in specific hotspot areas, continues to challenge public health management and response strategies

Problem Statement

- The geographical distribution of dengue cases is influenced by various factors, including weather conditions, sanitation, and water supply.
- Comprehensive understanding of how these factors interact in different regions is lacking.

Problem Statement

Objectives

- 1. To identify and map the dengue hotspots across different regions in the Philippines, using multi-year data analysis.
- 2. To analyze the correlation between weather conditions, sanitation, and water supply with the spread of dengue in various regions.
- 3.To provide actionable insights and recommendations for improving public health awareness, preparedness, and prevention strategies in identified high-risk areas.

DENGUE CASES DATA:

- Dengue cases and deaths data for the years 2016–2021 were obtained from the Humanitarian Data Exchange (HDX) platform. This dataset provides detailed reports on dengue cases and fatalities across various regions per down with dates when it was recorded in the Philippines.
 - https://data.humdata.org/dataset/philippine-denguecases-and-deaths

DENGUE CASES DATA:

- Additional, Dengue data by region for 2019–2022 were sourced from the Philippine Statistics Authority's (PSA) Environment Statistics. This dataset includes regional-level statistics on dengue cases.
 - https://psa.gov.ph/statistics/environmentstatistics/node/1684061201

WEATHER DATA:

- Historical weather data was obtained from the Climate Change Knowledge Portal by the World Bank. It provides comprehensive climate data, including historical trends for the Philippines, accessible for various timeframes.
 - https://climateknowledgeportal.worldbank.org/country/philippine
 s/climate-data-historical

SANITATION AND WATER SUPPLY DATA:

- Data on regional sanitation situation, and water supply conditions were from the PSA Environment Statistics. Dataset includes situations that may influence disease transmission.
 - https://psa.gov.ph/statistics/environment-statistics/node/1684061201
 - https://psa.gov.ph/statistics/environment statistics/highlights/component-5-human-settlements-and environmental-health

GEOSPATIAL DATA:

- Regional boundary geospatial data were obtained using GeoJSON files provided at the following link:
 - https://raw.githubusercontent.com/renatomaaliw3/public_files/ma ster/Data%2oSets/GeoData/gadm41_PHL_1.json
- Region-level geospatial data were extracted from a CSV file:
 - https://raw.githubusercontent.com/ixahh/CPE15L/refs/heads/main/DataSets/PH-REGION-GPD.csv

GEOSPATIAL DATA:

- Province-level geospatial data were obtained from a separate CSV file:
 - https://raw.githubusercontent.com/ixahh/CPE15L/refs/heads/main/DataSets/PH-PROVINCE-GPD.csv

INITIAL CLEANUP IN EXCEL:

- Removed unnecessary rows and columns to focus on essential data.
- Standardized column names (e.g., 'Dengue_Cases', 'Sanitation HH Total') for consistency.

DATA CLEANING AND TRANSFORMATION IN PYTHON:

- Handling Missing Values: Retained rows with critical data and dropped irrelevant ones.
- Mapping Region Codes to Names: Used a dictionary to map region codes to region names for readability.

- Filtering Data by Year: Filtered data based on selected year.
- Grouping and Aggregating Data: Summarized data by region (regCode) for total cases and totals for other parameters.
- Re-mapping Region Codes: Applied region name mapping again post-aggregation for easy visualization.

WEATHER DATA PREPROCESSING:

- Grouping by Month and Region: Calculated monthly totals and averages for weather data.
- Sorting Data by Month: Arranged weather data chronologically.

• Color Palette: Used distinct colors for regions when plotting weather data.

TOP CASES DATA PREPROCESSING:

- Aggregating by Region and Town: Summed cases and deaths by town for identifying top cases.
- Identifying High/Low Cases: Used nsmallest() and nlargest() to find towns with the most and least dengue cases.
- Plotting: Created horizontal bar graphs to visually compare the top 10 towns with highest/lowest cases.

1. Initial Cleanup in Excel:

4 A	В	С	D E	F		G l		1 .	J	K L	М		N	0	Р	Q
							Sanit	ation_I								VaterSupply
RegC	RegionNames	Ye T	lengue_ Dengue_I es ths	Sanitation HH Tota		n_HH_Total_li Sanitation_l Total Value			ion_	Sanitation_Unimpo	Sanitation OpenDefeca	tio VaterSupply	HH Total	VaterSupply HH Total Values	₩aterSupply_HH_Tot ved2	al_lm mproved_P ent Service
2 NCR	National Capital Region	2019		13	3385	3206	3206000	94.7	79.8	14.9	4.9	0	3385		3385000	3381
3 CAR	Cordillera Administrative Region	2019	12222	28	431	408	408000	94.6	86.2	8.5	4.5	0.9	431		431000	413
4	llocos Region	2019	19288	86	1232	1207	1207000	98	82.9	15.1	0	0	1232		1232000	1170
5 II	Cagayan Valley	2019		16 66	868	852	852000 2770000	98.1 97.1	90.2	7.9 8.9	0	0	868 2853		868000	848 2819
7 IV-A	Central Luzon CALABARZON	2019 2019		29	2853 3876	2770 3764	3764000	97.1	88.2 88.7	8.4	1.7	1.3	2853 3876		2853000 3876000	3787
8 IV-B	MIMAROPA Region	2019	10621	60	754	694	694000	92	82.9	9.1	1.7	6.4	754		754000	689
9 V 10 VI	Bicol Region	2019 2019		49 40	1300 1870	1104 1664	1104000 1664000	84.9 89	74.2 83.3	10.7 5.7	3.6 2.2	11.5 8.8	1300 1870		1300000 1870000	1203 1704
11 VII	Western Visayas Central Visayas	2019		+0 31	1883	1721	1721000	91.4	81.4	10	3.4	5.2	1883		1883000	1774
12 VIII	Eastern Visayas	2019	26157	85	1083	964	964000	89	77.3	11.7	2.1	8.9	1083		1083000	1044
13 IX	Zamboanga Peninsula	2019		77	851	766	766000	90 94	83.9	6.1 13.3	4.5 3.2	5.5	851		851000	789 1075
14 X 15 XI	Northern Mindanao Davao Region	2019 2019		88 39	1144 1320L	1075 1247	1075000 1247000	94.5	80.7 79.2	13.3 15.3	3.5	2.8 2	1144 1320		1144000 1320000	1281
16 XII	SOCCSKSARGEN	2019	19603 13	39	1156	1072	1072000	92.7	74.2	18.3	4.9	2.4	1156		1156000	1099
17 XIII	Caraga	2019	8892	63	634	597	597000	94.1	84.9	9.2	2.3	3.6	634		634000	596
18 BARMM	Bangsamoro Autonomous Region in	2019	6370	11	670	306	306000	45.6	32.7	12.9	38.2	16.1	670		670000	552
19 NCR	National Capital Region	2020	7338	97	3449	3392	3392000	98.4	79	19.4	1.5	0	3449		3449000	3449
20 CAR	Cordillera Administrative Region	2020	3729	10	439	428	428000	97.4	86.8	10.6	17	0.9	439		439000	424
21 I	llocos Region	2020		60	1252	1239	1239000	99	83	16	0	0.3	1252		1252000	1244
22 II	Cagayan Valley	2020	1340	23	883	863	863000	97.7	87.5	10,3	0	0	883		883000	856
23 III 24 IV-A	Central Luzon CALABARZON	2020 2020		39 06	2923 3970	2870 3904	2870000 3904000	98.2 98.3	87.2 87.4	11.1 10.9	!	0.8	2923 3970		2923000 3970000	2908 3906
25 IV-B	MIMAROPA Region	2020		33	770	708	708000	92	81	11	3.1	5	770		770000	747
26 V	Bicol Region	2020	1565	35	1319	1161	1161000	88	76	12.1	3.5	8.4	1319		1319000	1234
27 VI 28 VII	Western Visayas	2020 2020		44 00	1907 1927	1657 1737	1657000 1737000	86.9 90.1	80.4 75.6	6.5 14.6	2.1	11.1 7.9	1907 1927		1907000 1927000	1797 1798
29 VIII	Central Visayas Eastern Visayas	2020		37	1108	1032	1032000	93.1	82	11	1.3	7.5 5.6	1108		1108000	1067
30 IX	Zamboanga Peninsula	2020	2506	28	863	789	789000	91.4	79.8	11.7	2.9	5.7	863		863000	776
31 X	Northern Mindanao	2020		75 51	1168 1356	1093	1093000 1298000	93.6	82.5 72	11.2 23.7	5.2 3.3	1.2	1168 1356		1168000 1356000	1149 1312
32 XI 33 XII	Davao Region SOCCSKSARGEN	2020 2020	1012	56	1181	1298 1094	1094000	95.7 92.6	72.5	20.1	4.8	2.6	1181		1181000	1140
34 XIII	Caraga	2020		25	649	618	618000	95.2	87.1	8.1	2.1	2.7	649		649000	618
35 BARMM	Bangsamoro Autonomous Region in	2020	1200	7	683	382	382000	55.9	39.3	16.7	24.4	19.6	683		683000	558
36 NCR	National Capital Region	2020		26	3512	3447	3447000	98.2	86.7	11.2	0.3	0.7	3512		3512000	3475
	Cordillera						407000									
37 CAR	Administrative Region Ilocos Region	2021 2021	652 8546	37 96	447 1270	437 1262	437000 1262000	97.6 99.3	86.3 86.4	11.1 12.6	0.2	0.8 0.4	447 1270		447000 1270000	422 1259
39 II	Cagayan Valley	2021		21	898	879	879000	97.9	86.6	10.7	0.6	0.9	898		898000	881
40 111	Central Luzon	2021		71	2991	2958	2958000	98.9	92	6.5	0.2	0.7	2991		2991000	2971
41 IV-A 42 IV-B	CALABARZON MIMAROPA Region	2021 2021	8926 19 1502	59 31	4082 787	4030 723	4030000 723000	98.7 91.9	91.8 80.1	6.6 10.8	0.1	1 62	4082 787		4082000 787000	4011 737
43 V	Bicol Region	2021	798	41	1338	1246	1246000	93.1	84.6	7.4	0.3	5.7	1338		1338000	1262
44 VI	Western Visayas	2021	2467	45	1945	1819	1819000	93.5	86.3	6.7	0.3	5.7	1945		1945000	1853
45 VIII	Central Visayas	2021 2021	2604 (889	21	1970 1134	1830 1059	1830000 1059000	92.9 93.5	78.9 83.8	12.7 7.7	0.3	6 5.5	1970 1134		1970000 1134000	1866 1083
46 VIII	Eastern Visayas	2021	1074	40	074	700	700000	00.4	00.5	7.0	2.0	5.5 F.0	074		074000	000
4 +	Sheet3	FINA	AL DATA SETS	Population	Sheet5	Notes DataSet	Summary	Visualiza	ations	WaterSupply-Sur	(+) : 4					•

2. Handling Missing Values

```
# Merge the region and province dataframes based on the 'regCode' column
x = gdfphRegion.merge(gdfphProvince, on='regCode', how = 'inner')
# Rename 'provDesc' column to 'NAME 1' for consistency across datasets
x = x.rename(columns = {'provDesc' : 'NAME 1'})
# Merge the geographical data with the province-region data
gpdPhFinal = gdfPh.merge(x, on = 'NAME 1', how = 'inner')
# Map 'RegCode' in the dengue data to corresponding region codes
gdfdengData['RegCode'] = gdfdengData['RegCode'].map(regionCode)
# Rename 'RegCode' column to 'regCode' for consistency with other dataframes
gdfdengData = gdfdengData.rename(columns = {'RegCode' : 'regCode'})
# Ensure 'regCode' in the geographical data is of integer type for accurate merging
gpdPhFinal['regCode'] = gpdPhFinal['regCode'].astype(int)
# Merge the dengue data with the geographical data based on 'regCode'
gpdPhFinal = gpdPhFinal.merge(gdfdengData, on = 'regCode', how = 'inner')
# Map region codes in the merged data to region names for easier identification
gpdPhFinal['Region'] = gpdPhFinal['regCode'].map(regionName)
# Drop unnecessary columns from the data to clean it up
phdataSorted = gpdPhFinal.drop(['GID_1', 'GID_0', 'VARNAME_1',
                                    'NL_NAME_1', 'TYPE_1', 'ENGTYPE_1',
```

```
# Further clean the data by removing columns related to sanitation and water supply
phdataCleaned = phdataSorted.drop(['Sanitation_Basic', 'Sanitation_Limited', 'Sanitation_Unimproved',
                                   'Sanitation OpenDefecation', 'WaterSupply Basic',
                                   'WaterSupply Limited', 'WaterSupply Unimproved',
                                   'WaterSupply SurfaceWater', 'Sanitation Improved Percent',
                                   'WaterSupply Improved Percent Service', axis = 1)
# Filter the cleaned data to create separate dataframes for each year
yearDf2019 = phdataCleaned[(phdataCleaned['Year'] == 2019.0)]
yearDf2020 = phdataCleaned[(phdataCleaned['Year'] == 2020.0)]
yearDf2021 = phdataCleaned[(phdataCleaned['Year'] == 2021.0)]
yearDf2022 = phdataCleaned[(phdataCleaned['Year'] == 2022.0)]
# Aggregate (dissolve) the data by region for each year to sum dengue cases by region
regions dissolved2019 = yearDf2019.dissolve(by='Region', as index=False)
regions dissolved2020 = yearDf2020.dissolve(by='Region', as index=False)
regions dissolved2021 = yearDf2021.dissolve(by='Region', as index=False)
regions dissolved2022 = yearDf2022.dissolve(by='Region', as index=False)
# Define custom colormaps for each year
cmap 2019 = mcolors.LinearSegmentedColormap.from list("cmap 2019", ['#F0F8FF', '#0000FF']) # Light blue to dark blue for 2019
cmap 2020 = mcolors.LinearSegmentedColormap.from list("cmap 2020", ['#FFFAF0', '#FF4500']) # Ivory to red for 2020
cmap 2021 = mcolors.LinearSegmentedColormap.from list("cmap 2021", ['#F0FFF0', '#006400']) # Honeydew to dark green for 2021
cmap 2022 = mcolors.LinearSegmentedColormap.from list("cmap 2022", ['#FAF0E6', '#8B008B']) # Linen to dark magenta for 2022
```

3. Weather Data Processing

```
# --- Second Graph: Monthly Climatology ---
# Plotting the second graph (Climate Data)
bars = ax2.bar(tempMonth['Category'], tempMonth['Precipitation'], color='lightblue', alpha=0.6, label='Precipitation')
# Creating the secondary axis for temperature
ax2 twin = ax2.twinx()
# Plotting the line graphs for the different temperature types
line_min = ax2_twin.plot(tempMonth['Category'], tempMonth['Average Minimum Surface Air Temperature'], color='r', marker='o', lak
line mean = ax2 twin.plot(tempMonth['Category'], tempMonth['Average Mean Surface Air Temperature'], color='g', marker='s', label
line max = ax2 twin.plot(tempMonth['Category'], tempMonth['Average Maximum Surface Air Temperature'], color='y', marker='^', lak
# Adding labels and title to the second graph
ax2.set xlabel('Month', fontsize=14, fontweight='bold')
ax2 twin.set ylabel('Temperature (°C)', fontsize=14, fontweight='bold')
ax2.set ylabel('Precipitation (mm)', fontsize=14, fontweight='bold')
ax2.set title('Monthly Climatology of Temperature & Precipitation (1991-2020)', fontsize=16, fontweight='bold')
# Combining the legends into one for the second graph
handles, labels = ax2.get legend handles labels()
handles.extend([line min[0], line mean[0], line max[0]])
labels.extend(['Average Minimum Temperature', 'Average Mean Temperature', 'Average Maximum Temperature'])
ax2.legend(handles, labels, loc='upper left', fontsize=12)
# Displaying the second graph's x-axis with rotated labels
ax2.set xticks(range(1, 13))
ax2.set_xticklabels(['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec'], fontsize=12)
ax2.tick params(axis='x', rotation=45)
```

4. Top Cases Preprocessing

```
] # Aggregate the data to ensure unique combinations of Year and Region
 aggregated_data = phdataCleaned.groupby(['Year', 'Region'], as_index=False).agg({'Dengue_Cases': 'sum'})
 # Convert 'Year' to integer to remove the decimal
 aggregated data['Year'] = aggregated data['Year'].astype(int)
 # Pivot the data for bar graph
 pivot data = aggregated data.pivot(index="Year", columns="Region", values="Dengue Cases")
 # Plot the bar graph
 pivot data.plot(kind="bar", figsize=(14, 8), colormap="tab20")
 # Adding labels and title
 plt.xlabel("Year", fontsize=12)
 plt.ylabel("Dengue Cases", fontsize=12)
 plt.title("Dengue Cases by Region (2019-2022)", fontsize=14)
 plt.grid(axis='y', linestyle='--', alpha=0.7)
 # Customizing ticks
 plt.xticks(fontsize=10, rotation=0)
 plt.yticks(fontsize=10)
 plt.tight_layout()
 plt.show()
```

```
# Drop the columns "Date" and "Year"
townCases = townCases.drop(columns=["Date", "Year"])
# Group by unique Region and Town, summing Cases and Deaths
aggregated townCases df = (
    townCases.groupby(["Region", "Town"], as index=False)
    .agg({"Cases": "sum", "Deaths": "sum"})
# Top 10 towns with the lowest cases across all regions
top 10 lowest = aggregated townCases df.nsmallest(10, "Cases")
# Top 10 towns with the highest cases across all regions
top 10 highest = aggregated townCases df.nlargest(10, "Cases")
# Create a figure with two horizontal bar graphs side by side
fig, axes = plt.subplots(1, 2, figsize=(16, 8), sharey=False)
# Plotting the top 10 towns with the lowest cases
axes[0].barh(top 10 lowest["Town"], top 10 lowest["Cases"], color="lightcoral")
axes[0].set title("Top 10 Towns with the Lowest Dengue Cases (2016-2021)", fontsize=1
axes[0].set xlabel("Cases", fontsize=12)
axes[0].set ylabel("Town", fontsize=12) # Label for the Y-axis on the left graph
axes[0].invert yaxis() # Align with the left graph
axes[0].tick params(axis="y", labelsize=10)
```

Visualizațions, Elnsights

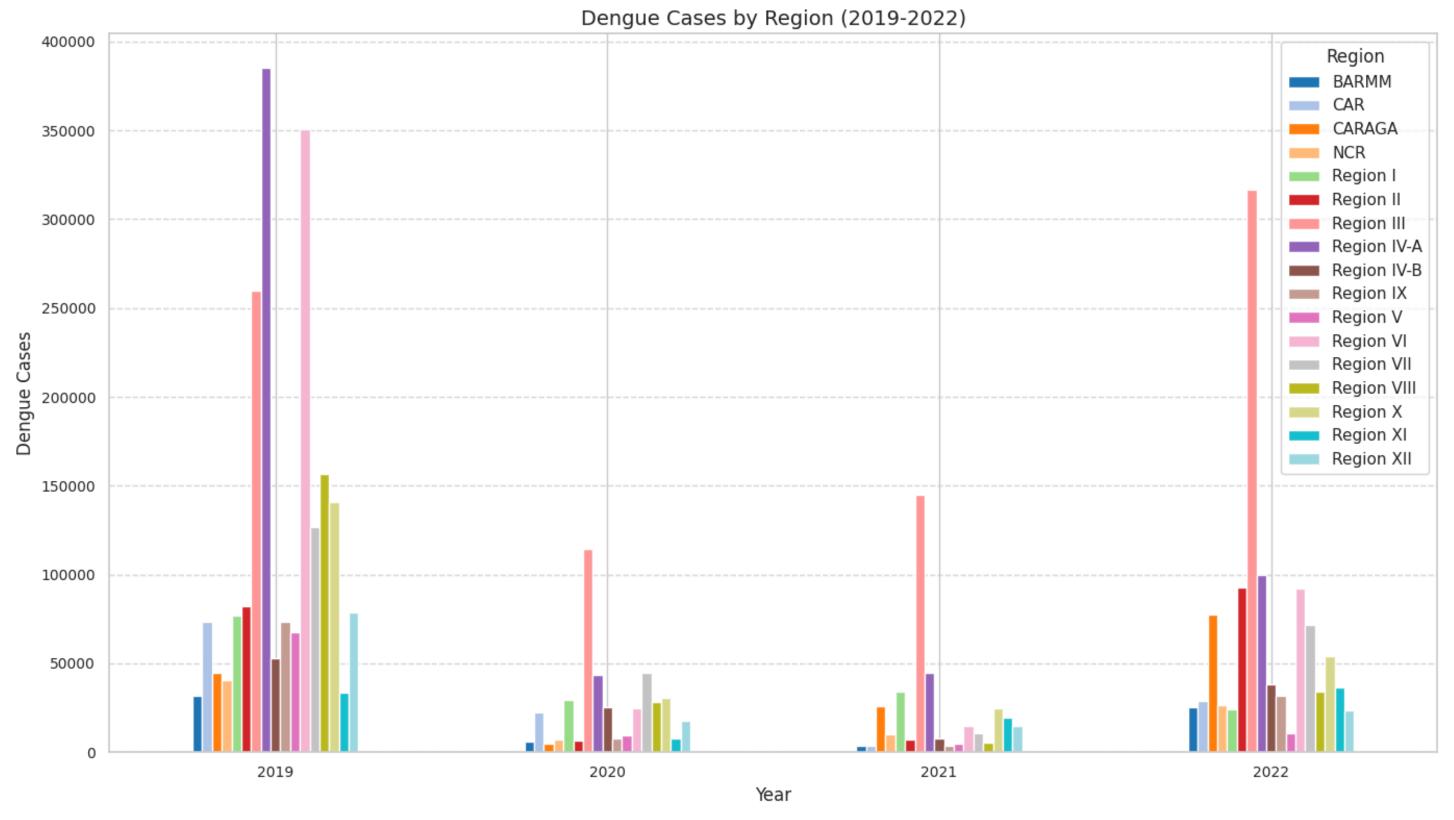


Figure 1. Dengue Cases per Region from 2019-2022

Dengue Deaths by Region and Year

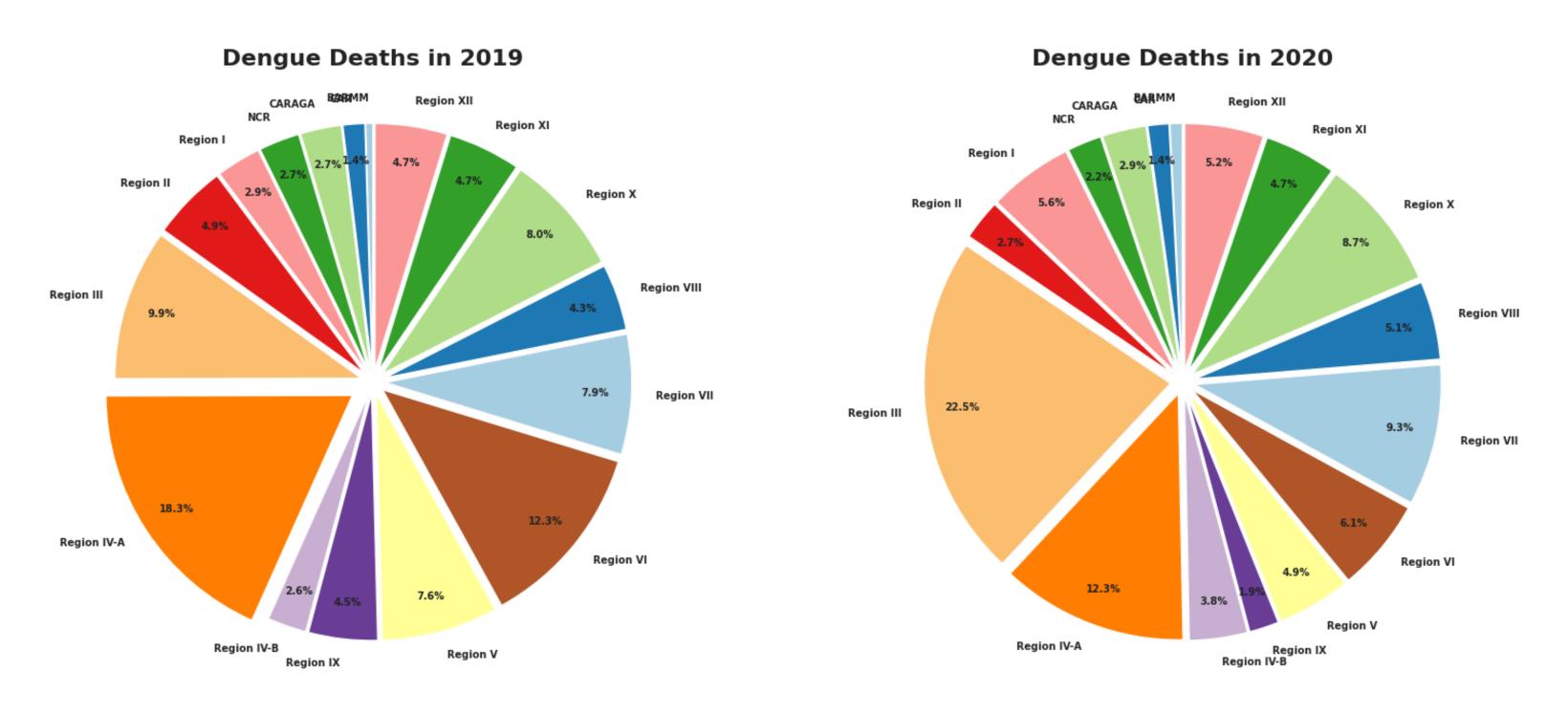


Figure 2. Percentage of Dengue Deaths per Region

Dengue Deaths by Region and Year

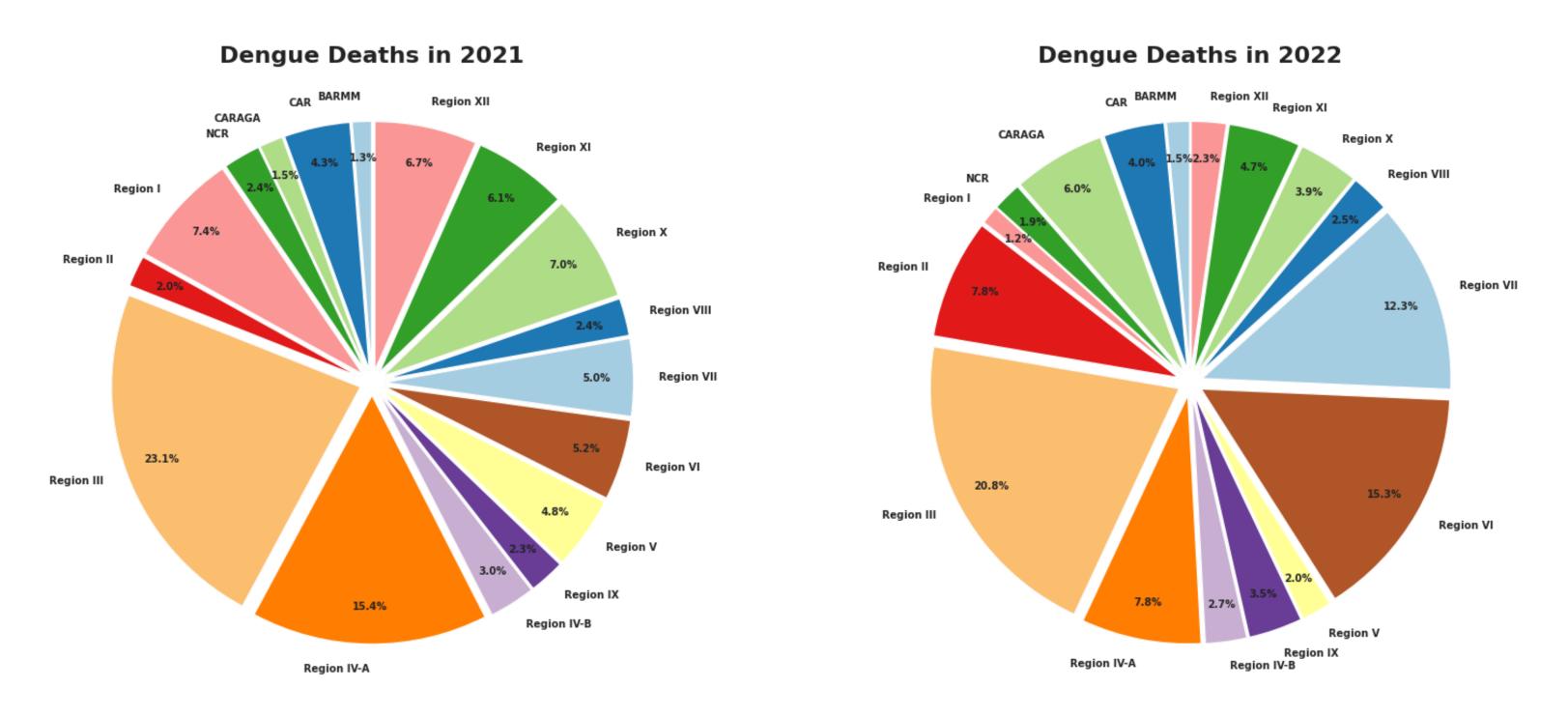


Figure 2. Percentage of Dengue Deaths per Region

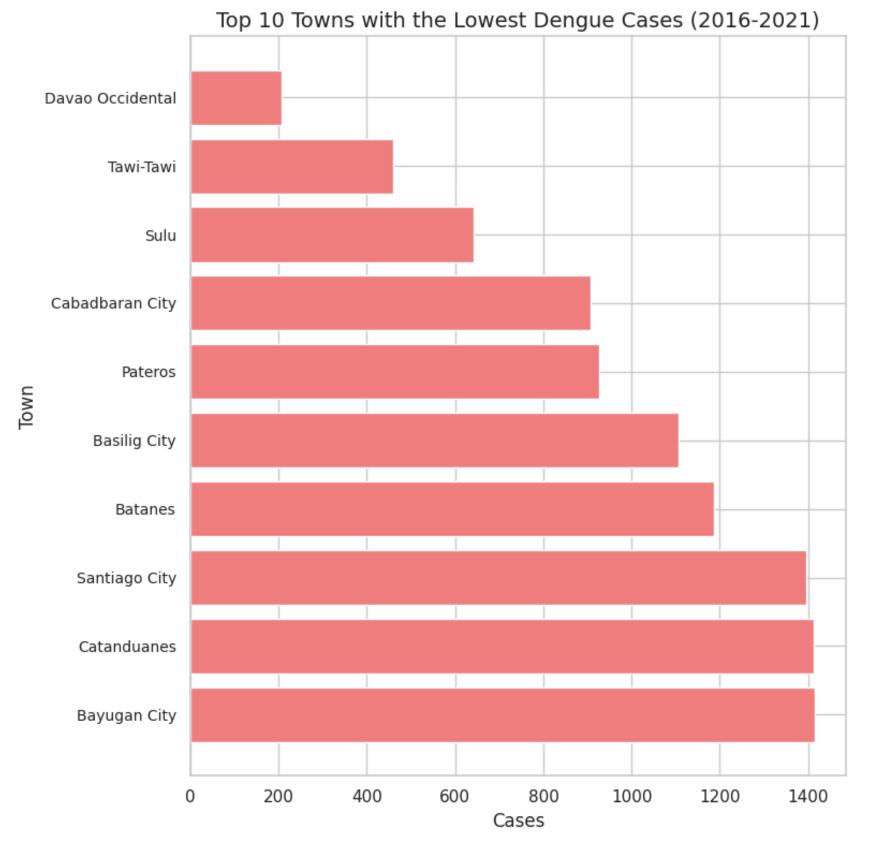


Figure 3. Top 10 Towns/City with the Lowest Total of Dengue Cases from 2016–2021

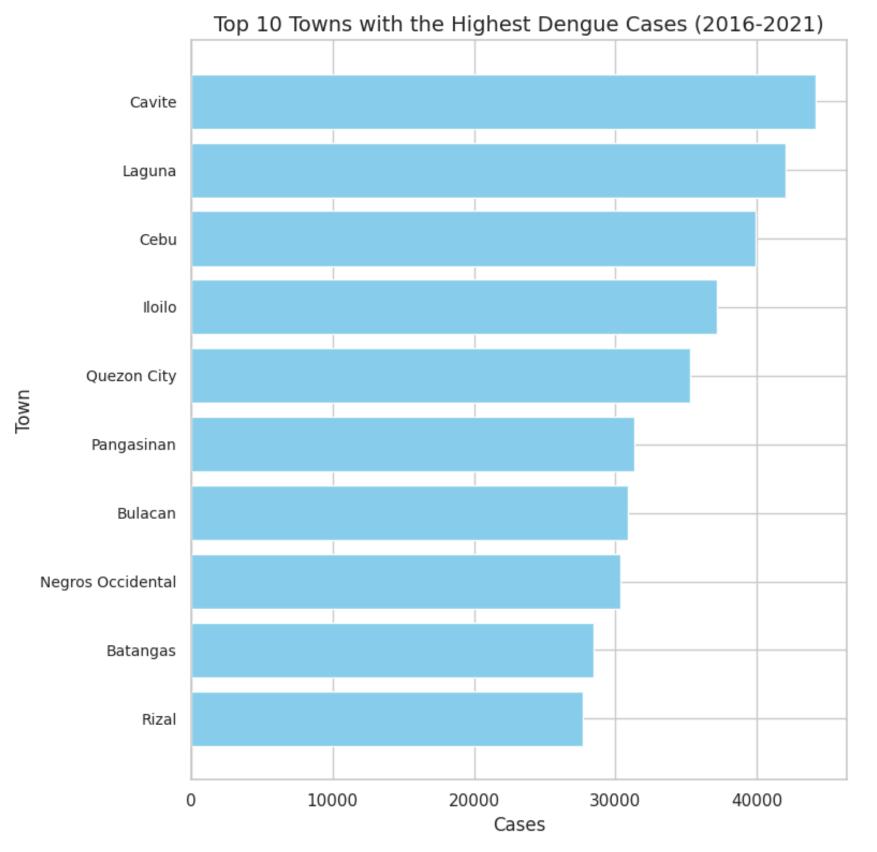


Figure 4. Top 10 Towns/City with the Highest Total of Dengue Cases from 2016–2021

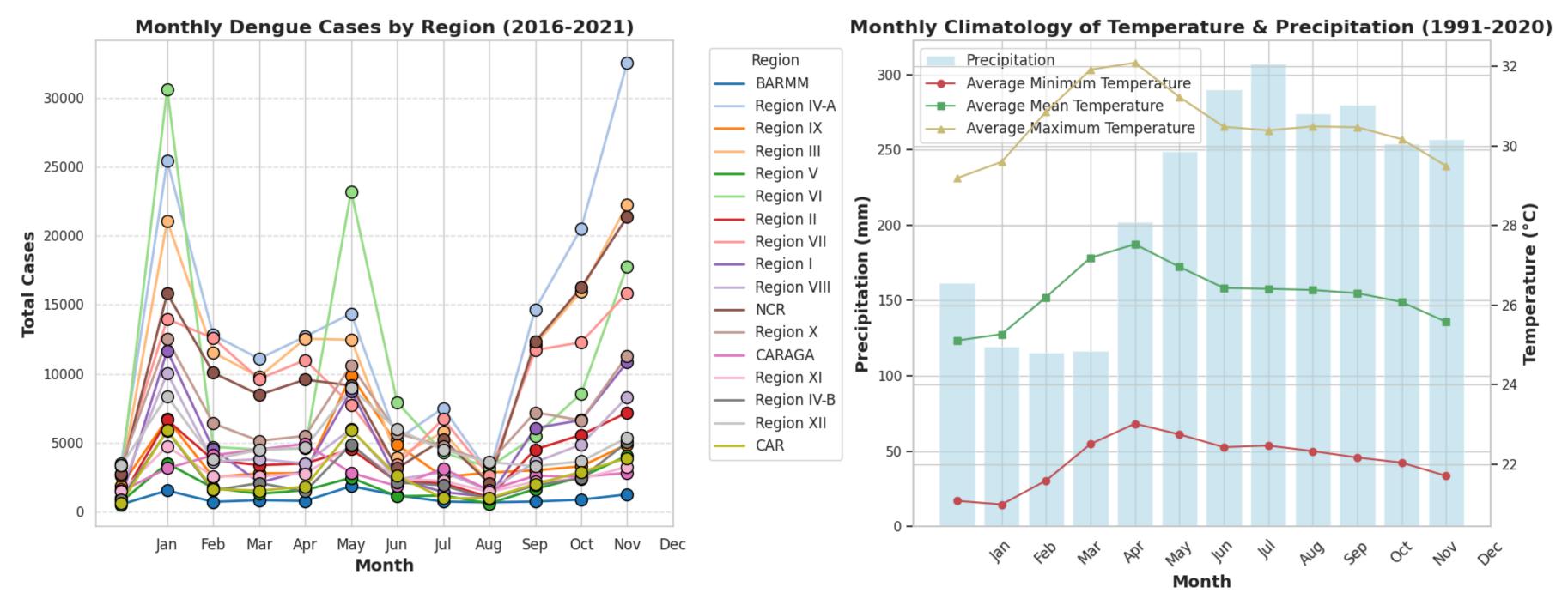


Figure 5. Regional Dengue Cases progression per Month from 2016–2021

Figure 6. Philippines Monthly Climatology of Temperature & Precipitation from 1991–2020

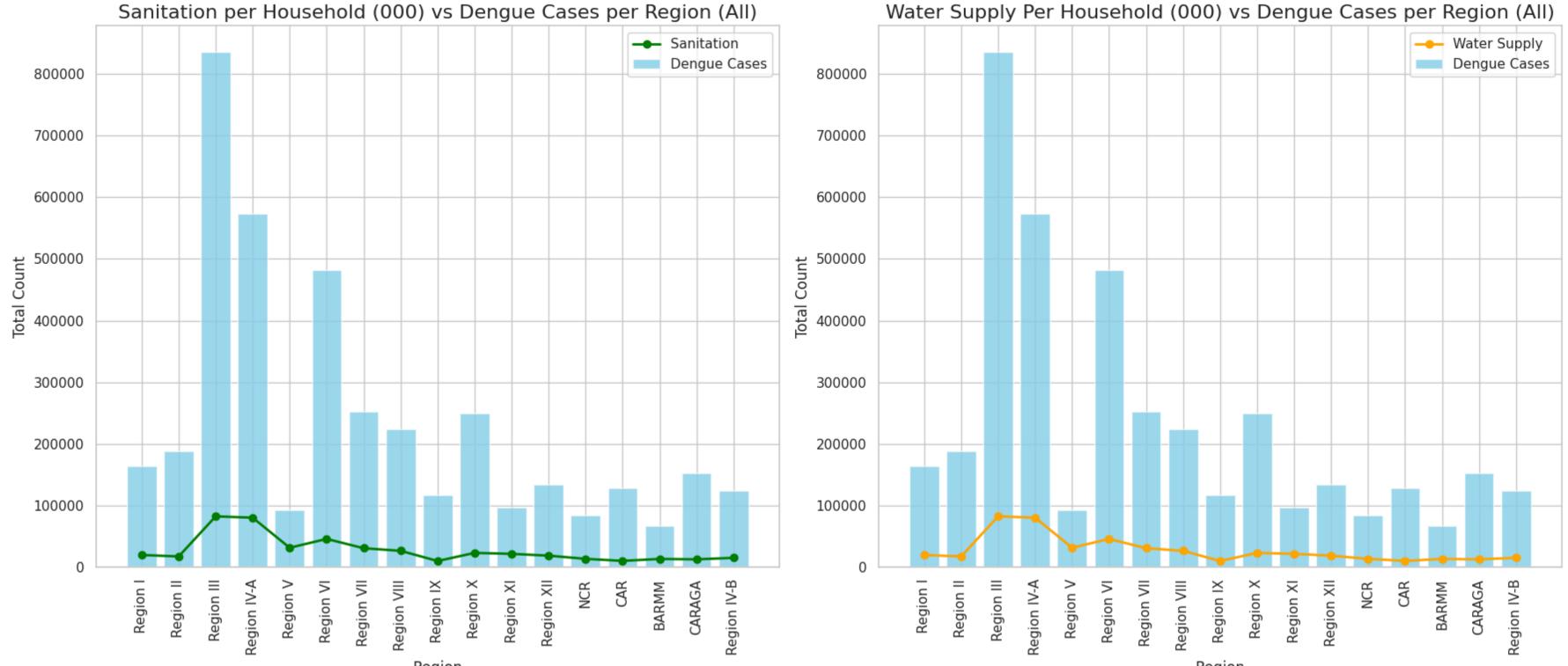
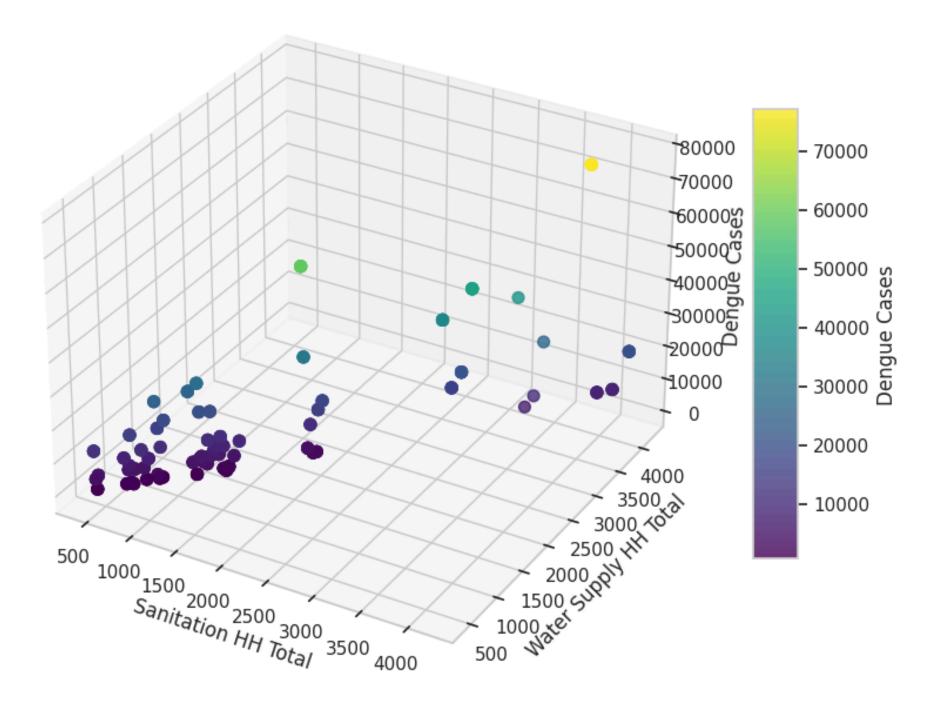


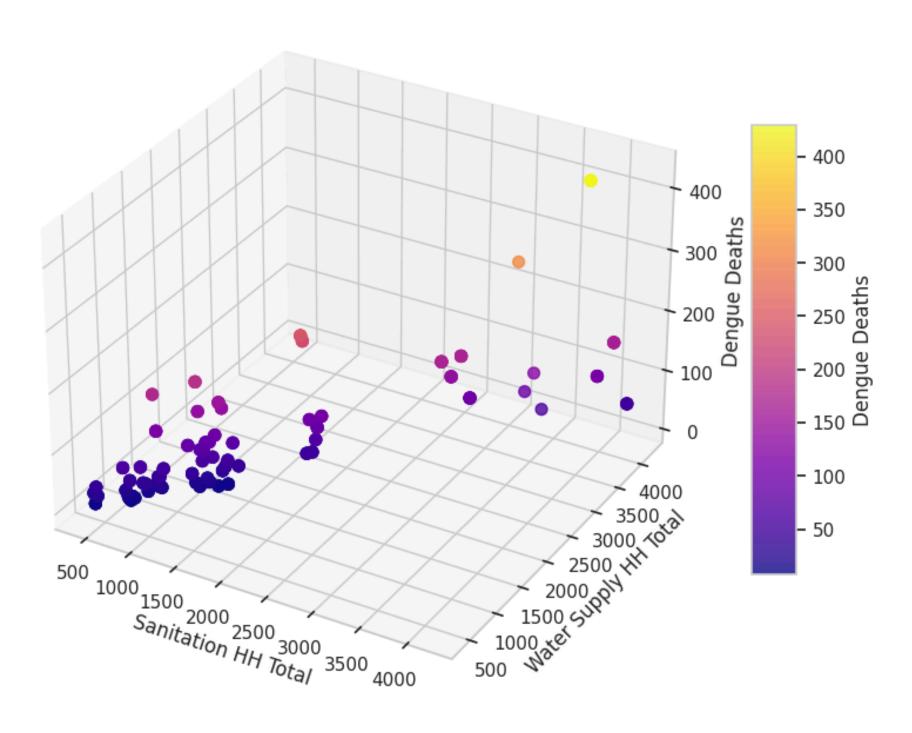
Figure 7. Sanitation per Household vs Dengue Cases per Region

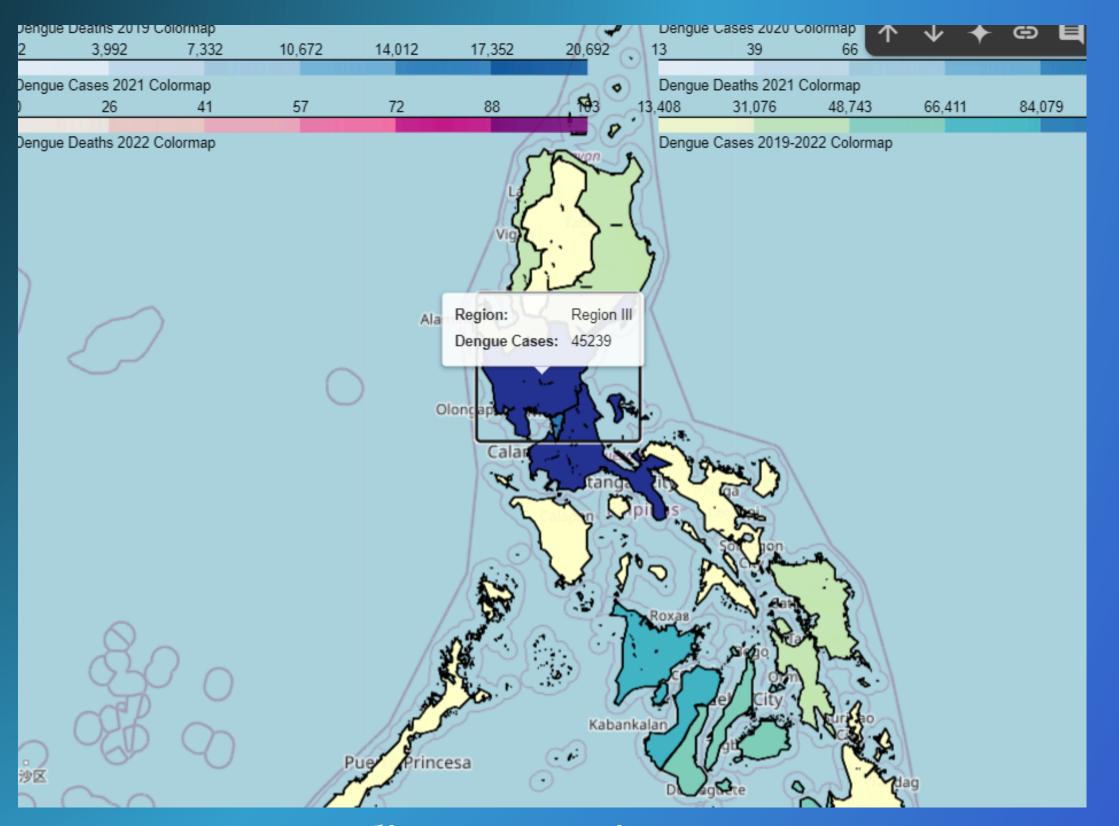
Figure 8. Sanitation per Household vs Dengue Cases per Region

Dengue Cases vs Sanitation and Water Supply

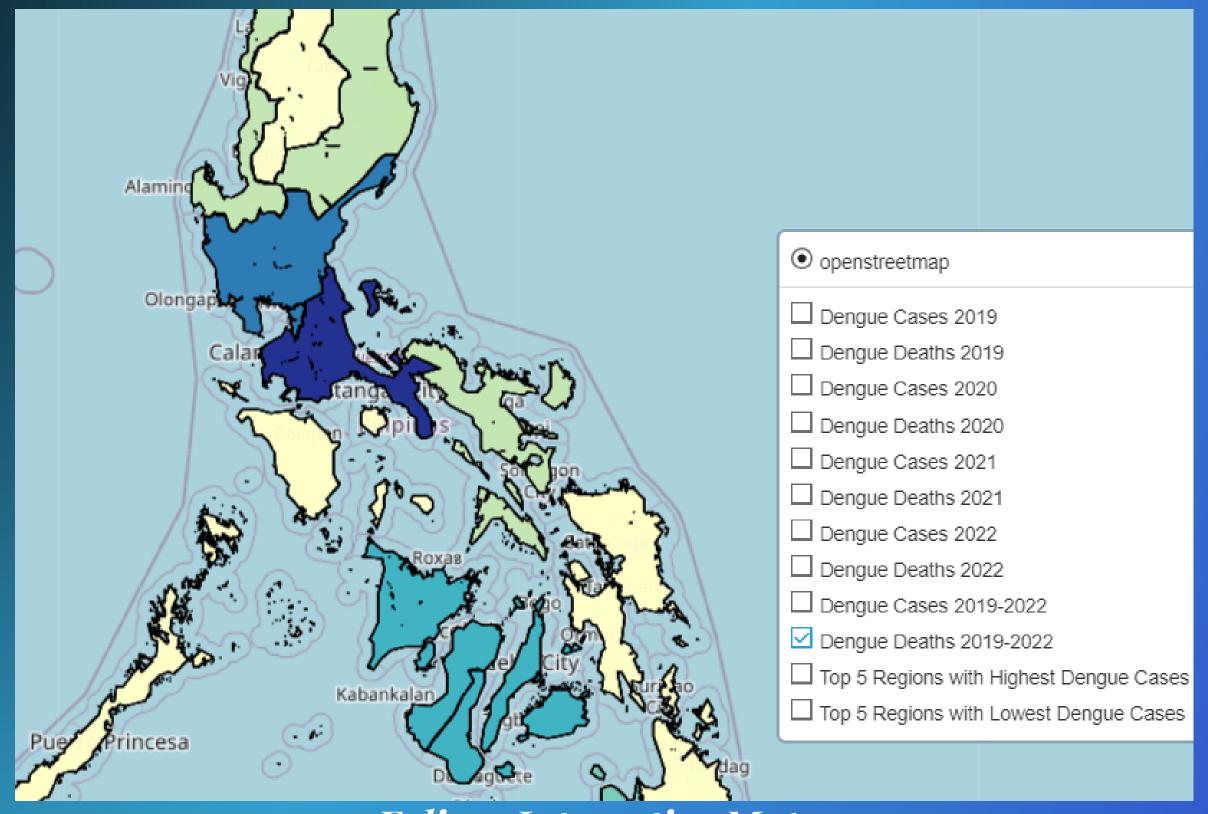


Dengue Deaths vs Sanitation and Water Supply

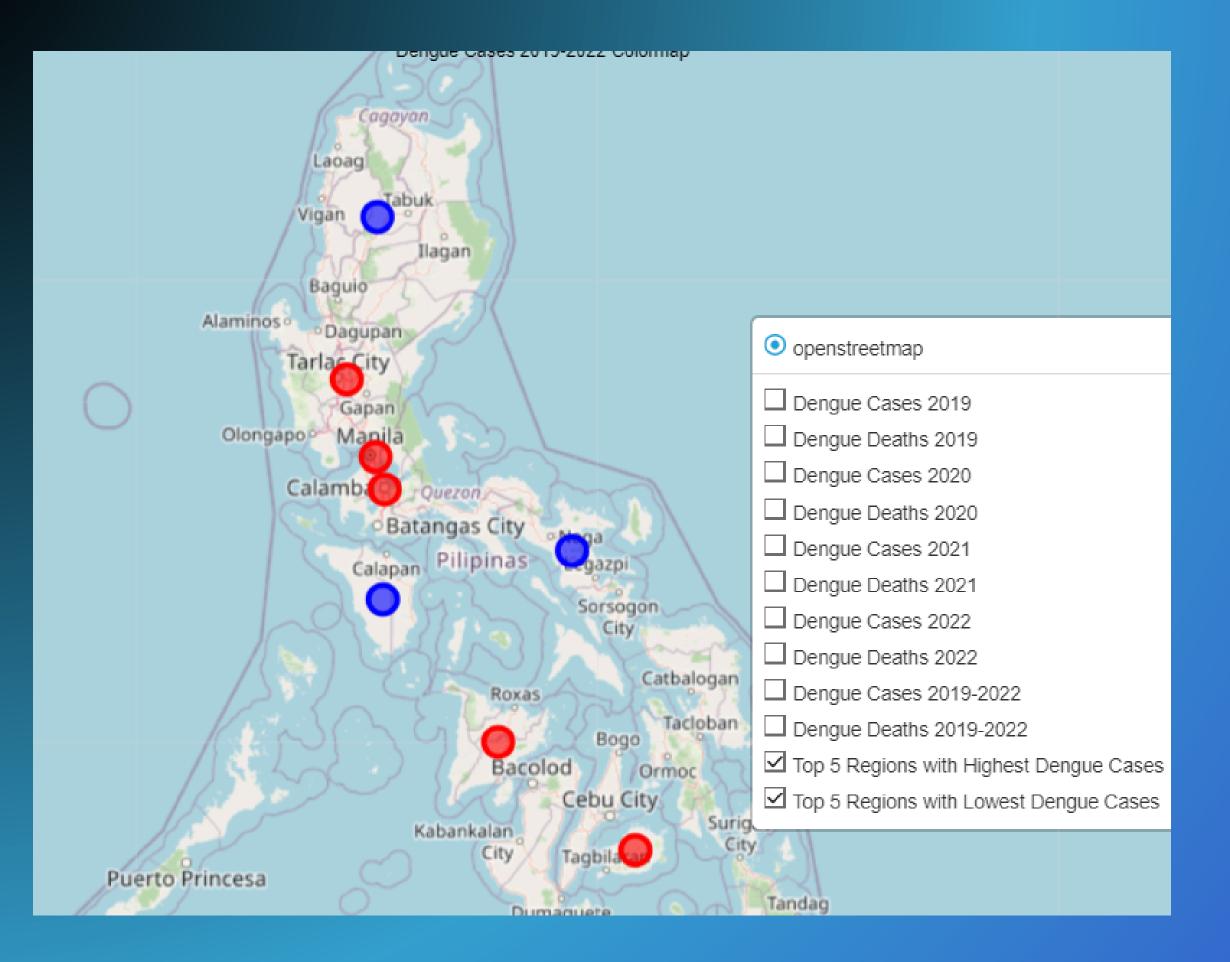




Folium Interactive Maps Mapping of Dengue Cases Per Region and Year
(2019-2022)



Folium Interactive Maps Mapping of Dengue Deaths Per Region and Year
(2019-2022)



Folium Interactive Maps Top 5 Regions with the Highest
and Lowest Dengue Cases

- The Top 3 City/Town of those who got the most Dengue Cases from 2016–2021 comes from Cavite, Laguna, and Cebu
- Dengue Cases and Deaths of 2019–2022 are more widespread in Region III, Region IV-A, and Region VI
- Dengue Disease from Years 2016–2021 usually has influx from the months of February, June, and November

- When the temperature rises on specific months like February, June, and November there seems to be an influx Dengue contraction.
- Although there is a high level Sanitation per Household in Region IV, Region III, and Region VI dengue is still widespread at their location. Which may be due to the highly urbanized city in this Region.

- Having water supply per Household in Region IV, Region III, and Region VI seems to make dengue widespread at their location.
- Due to the highly urbanized cities (in Region IV-A and Region VI) and wide wet farm lands (Region III). Breeds the most mosquitos when the temperature is hot and there is a lot of pollution in the Area.

Conclusions

Dengue cases are most prevalent in Cavite, Laguna, and Cebu, with Regions III, IV-A, and VI showing widespread cases and deaths. Outbreaks peak in February, June, and November, likely due to higher temperatures promoting mosquito breeding. Despite good sanitation levels in these regions, dengue persists, possibly due to urbanization, water supply, and environmental factors. Urbanized cities and wet farmlands create ideal conditions for mosquito breeding, exacerbated by pollution. Addressing these factors is key to controlling the spread of dengue.

Conclusions

References

- Department of Health (DOH). (2021). Dengue surveillance report. Department of Health, Philippines. https://www.doh.gov.ph
- Drakou, K., Nikolaou, T., Vasquez, M., Petric, D., Michaelakis, A., Kapranas, A., Papatheodoulou, A., & Koliou, M. (2020). The effect of weather variables on mosquito activity: A snapshot of the main point of entry of Cyprus. International Journal of Environmental Research and Public Health, 17(4), 1403. https://doi.org/10.3390/ijerph17041403
- Jovi De Leon, & Jovi De Leon. (2023, August 30). DOH: Dengue cases in CL down by 63%. SunStar Publishing Inc. https://www.sunstar.com.ph/pampanga/local-news/doh-dengue-cases-in-cl-down-by-63
- Parsons, D. (2024, February 16). Greening Urban Landscapes: A Climate-Resilient Future with Mosquito Challenges in Mind. https://blogs.biomedcentral.com/bugbitten/2024/02/16/greening-urban-landscapes-a-climate-resilient-future-with-mosquito-challenges-in-mind/#:~:text=However%2C%20due%20to%20the%20avail,and%20lower%20predator%20number s%2C
- Philippine dengue Cases and deaths Humanitarian data exchange. (n.d.). https://data.humdata.org/dataset/philippine-dengue-cases-and-deaths

References

- Statista. (2024, October 18). Number of dengue cases Philippines 2022, by region. https://www.statista.com/statistics/1498116/philippines-number-dengue-cases-by-region/#:~:text=In%202022%2C%2oCentral%2oLuzon%2oregistered,the%2ocountry%2opeaked%2 oin%202019
- Villanueva, R. (2023, January 16). Philippines logs 220,705 dengue cases in 2022. Philstar.com. https://www.philstar.com/nation/2023/01/17/2238093/philippines-logs-220705-dengue-cases-2022
- World Bank Climate Change Knowledge Portal. (n.d.). https://climateknowledgeportal.worldbank.org/country/philippines/climate-data-historical
- World Health Organization: WHO & World Health Organization: WHO. (2024, April 23). Dengue and severe dengue. https://www.who.int/NEWS-ROOM/FACT-SHEETS/DETAIL/DENGUE-AND-SEVERE-DENGUE

THANK YOU FOR LISTENING!

Link for Python (Google Colab File):
https://colab.research.google.com/drive/1RekrSo8QNSAyqJ5KlCPgo
nYTN7hElc7p?usp=sharing

Link for Folium Map & PDF File Download: https://drive.google.com/drive/folders/1UKKSV25IA934gLa9Xx_c5FY
FDKU2n7eH?usp=drive_link

IMPORTANT LINKS: