

Climatic Shifts and Mosquito-Borne Threats: Analyzing the Impact of Climate Change on Dengue Incidence in Singapore

1 Introduction

In 2016, The Straits Times reported it the hottest year on record for Singapore and the world. In the same year, they published another article with the headline 'Dengue cases hit 11,343, surpass last year's figure'. Both articles highlight the correlation between rising temperatures and the escalating number of dengue cases.

Against this backdrop, climate change has become a global predicament over the last few decades, impacting ecosystems, economies, and public health. Notably, the World Health Organization (WHO) has reported a significant global surge in dengue incidence by tenfold over the past two decades, presenting a considerable public health challenge. Given this global context, understanding the local implications, in Singapore, becomes crucial for effective public health management and preventive measures.

It is known and proven that increasing temperatures increases the available habitat for the dengue fever vector, the *Aedes aegypti* mosquito, while concurrently increasing both the longevity of the virus and the mosquito. They also breed more rapidly in stagnant water. With Singapore's tropical humid and wet climate, a pressing concern emerges as we witness the influence of rising temperatures and precipitation on the prevalence of this widespread vector-borne disease.

Therefore, this research aims to explore the relationship between both increasing surface air temperature and rainfall, caused by climate change, and the levels of dengue incidence in Singapore. To achieve this, a thorough approach will be used, including clearly defined aims and objectives, methodical dataset acquisition from different sources and insightful exploratory data analysis.

1.1 Aims and objectives

My primary objective is to discern the influence of climate change variables, specifically temperature and precipitation, on the frequency and intensity of dengue outbreaks in Singapore. Explicitly, I seek to:

1. Showcase Temperature Influence: Demonstrate a correlation between increasing temperatures, attributed to climate change, and the rise in dengue cases over the years.
2. Consider Rainfall as a Factor: Explore the role of heavier rainfall, another consequence of climate change, and its correlation with dengue incidence, recognizing the breeding habits of *Aedes* mosquitoes in stagnant water.
3. Accounting for Pandemic Impact: Investigate the unique case of 2020, where the COVID-19 pandemic influenced daily routines, potentially affecting water storage practices and waste management, and consequently, dengue incidence.

1.2 Dataset acquisition

To realize these aims, I have procured data from reputable sources, primarily tapping into the World Bank's Climate Knowledge Portal for Climate Data, SingStat for Dengue Incidence Records and the National Environment Agency (NEA) for dengue breeding habits in Singapore. The climate and dengue dataset encompasses a time span sufficient for discerning long-term trends, ensuring the reliability of our analyses.

1.3 Data utilisation

With the use of Jupyter Notebook, the dataset will be meticulously explored through an extensive exploratory data analysis and I will employ Python's data analysis libraries to:

- Clean and preprocess the dataset for accuracy and completeness.
- Visualize temporal trends to uncover patterns.
- Conduct statistical analyses to establish correlations between climatic variables and dengue incidence.
- Develop visualizations that effectively communicate insights derived from the dataset.

1.4 Writing style

Throughout this project, findings will be articulated in a concise manner, ensuring that both technical and non-technical audiences can grasp the nuances of the intricate interplay between climate change and dengue levels. Visualizations, code documentation, and written explanations will be presented with clarity and coherence.

1.5 Summary of area of research

In summary, this research focuses on the correlation between climate change and public health, specifically on the escalating dengue incidence in Singapore. This study aims to contribute valuable insights into the evolving dynamics of vector-borne diseases in a changing climate.

2 Relevance of data

2.1 Origin of data

The primary source of dengue incidence data is acquired from Singstat, the Department of Statistics Singapore, which provides comprehensive historical data on dengue cases (from 1966-2021), allowing for a detailed temporal analysis. Additionally, Climate data, specifically temperature and precipitation, is sourced from the World Bank's Climate Knowledge Portal. This data is derived from meteorological stations, satellites, and other reliable sources, providing detailed information on temperature and rainfall variations over time, aligning with the research focus. I have also obtained the NEA Quarterly Dengue Surveillance Data (2020) from the National Environment Agency (NEA) in Singapore to analyse the statistics on dengue-related deaths.

2.2 Justification for data source:

The choice of Singstat as the primary source for dengue incidence data is well-justified due to several key reasons:

1. Comprehensive Historical Coverage: Singstat provides an extensive dataset spanning from 1966 to 2021, enabling a comprehensive temporal analysis. This historical depth is crucial for understanding long-term trends and patterns in dengue incidence.
2. Reliable and Official Source: Singstat, being the Department of Statistics Singapore, is a reliable and official source, ensuring the credibility and accuracy of the data. This is imperative for a trustworthy research outcome.

I have also chosen to use Climate data from the World Bank's Climate Knowledge Portal as the Climate Knowledge Portal compiles data from meteorological stations, satellites, and other reliable sources. This diversity enhances the reliability and accuracy of the temperature and precipitation data, providing a holistic view of climate variations.

The inclusion of NEA's quarterly dengue surveillance data for 2020 is particularly pertinent as it allows for a focused analysis of a specific period, offering insights into the impact of unique circumstances, such as the COVID-19 pandemic, on dengue incidence. The dataset provides essential statistics on breeding habits which have led to a favourable environment for Aedes mosquito breeding.

This aligns seamlessly with the research question, allowing for an in-depth exploration of how different factors may correlate with fluctuations in dengue incidence.

2.3 Suitability for Research Question

The dengue incidence dataset from Singstat includes an extensive coverage of years spanning from 1966 to 2021, which aligns with the research, enabling an investigation into the temporal patterns of dengue incidence.

The Climate Knowledge Portal data covers temperature and precipitation data, directly addressing the climate-related aspects of the research question.

The NEA dengue dataset also includes relevant information on the breeding habits in 2020, addressing the human behaviour aspects of the research.

2.4 Format of data

All climate and dengue incidence data is available in CSV format. This format facilitates seamless integration into data analysis tools like Jupyter Notebooks using Pandas, ensuring ease of data manipulation and analysis. For my data on breeding habits, I have chosen to extract information from the Quarterly Dengue Surveillance data PDF instead. While PDFs can be more challenging to process, this format preserves the integrity of the original data source and provides essential insights for the research.

2.5 Consideration of Alternative Datasets

1. Health Promotion Board (HPB) Data:

- Strengths: Comprehensive health-related data, potentially including additional variables related to public health interventions and awareness campaigns.
- Weaknesses: May lack the spatial and temporal granularity of the Singstat dataset.

1. Singapore Meteorological Service (MSS) Data:

- Strengths: Local meteorological data, potentially with more detailed information on climate variables.
- Weaknesses: May lack the specific temporal alignment needed to correlate climate data with reported dengue cases.

3 Project background

3.1 Interest and relevance of project

The intersection of climate change and the escalating incidence of dengue presents a compelling field of interest due to its profound implications on public health. As the most widespread and rapidly increasing vector-borne disease globally, dengue not only poses an immediate threat to individuals but also serves as an impactful indicator of the broader consequences of climate change. With increasing temperatures and heavier rainfall attributed to climate change, understanding the intricate relationship between these variables and dengue transmission becomes paramount. Hence, this research is particularly relevant as it sheds light on the urgency of climate action, showcasing how environmental shifts contribute to the escalation of dengue, and also emphasizes the urgent need for proactive measures to mitigate the impact of dengue.

3.2 Novelty of Research

While previous studies have explored the link between climate change and vector-borne diseases, the specific nuances of this relationship in the context of Singapore, especially with a focus on dengue incidence, remain relatively unexplored. The focus on the specific impact of temperature and rainfall variations on dengue adds a unique dimension while the integration of historical dengue data from Singstat, covering a substantial period, delves into the historical context of dengue incidence in Singapore. Furthermore, the inclusion of NEA's quarterly data for 2020, a year marked by the COVID-19 pandemic, introduces an unprecedented element, exploring how exceptional circumstances may influence dengue transmission. This nuanced approach contributes to the novelty of the research.

3.3 Scope of work

The scope of this research involves a comprehensive analysis of the relationship between climate variables and dengue incidence in Singapore. The focus will be on analyzing the temporal patterns using data from Singstat, and the World Bank's Climate Knowledge Portal. While the primary emphasis is on temperature, secondary climate factors such as rainfall/precipitation will also be considered. The scope extends to exploring correlations, trends, and potential influence by the COVID19 pandemic. The analysis will not delve into broader socioeconomic factors influencing dengue incidence.

3.4 Analytical data processing pipeline

1. Data Collection: Obtain and preprocess dengue and climate data from Singstat, the World Bank's Climate Knowledge Portal and NEA.
2. Data Exploration: Conduct exploratory data analysis (EDA) to understand the structure and patterns in the datasets.
3. Data cleaning and merging: Extract only the data that is needed.
4. Correlation Analysis: Examine the correlation between climate variables and dengue incidence.
5. Temporal Trends: Visualize spatial and temporal trends in dengue cases and climate variables.

3.5 Evaluation of aims and objectives

The success of this research will be evaluated based on the achievement of the defined aims and objectives. This includes the identification of clear temporal trends and correlations between climate variables and dengue incidence. At the end of this project, I should be able to address the research hypotheses posed throughout the project.

4 Analysis of the Impact of Climate Change on Dengue Incidence

4.1 Annual Surface Air Temperature

Importing data from CSV file

Using the pandas library, the data is extracted and the contents of the CSV file is read and stored into a dataframe. This format offers versatility, supporting efficient data manipulation, integration with NumPy for advanced mathematical operations, and compatibility with statistical analyses. The dataframe's support for qualitative and quantitative data also ensures seamless exploration and integration into various analytical workflows, laying a foundation for comprehensive investigations.

```
In [98]: # import requirements
import requests
import base64
import datetime
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import json
import tabula
from tabula.io import read_pdf
from tabulate import tabulate
from IPython.display import display, HTML
from wordcloud import WordCloud
```

```
In [92]: # Use pandas to read the CSV file into a DataFrame
file_path = 'annual-mean-surface-air-temp-of-sg-1901-2022.csv'
```

```
airtemp_df = pd.read_csv(file_path)
airtemp_df.head()
```

Out[92]:

	Category	Annual Mean	5-yr smooth
0	1901	26.68	26.77
1	1902	26.63	26.68
2	1903	26.63	26.61
3	1904	26.65	26.55
4	1905	26.69	26.53

Understanding and cleaning dataset

As seen above, the dataset encompasses the years 1901 to 2022. However, due to the absence of historical dengue data before 1966, I specifically extracted and retained data spanning 1966 to 2022 to focus the pertinent timeframe.

In [40]:

```
# cleaning data to show only year 1966 onwards to 2022
airtemp_df = airtemp_df.rename(columns={'Category': 'Year'})

# convert the 'Year' column to numeric, handling errors by coercing invalid values to NaN
airtemp_df['Year'] = pd.to_numeric(airtemp_df['Year'], errors='coerce')

# filter data for the desired range of years (1966 to 2021)
airtemp_df = airtemp_df[airtemp_df['Year'].between(1966, 2022)]

# removing last column
airtemp_df = airtemp_df.drop(columns=['5-yr smooth'], errors='ignore')
airtemp_df.head()
```

Out[40]:

	Year	Annual Mean
65	1966	26.78
66	1967	26.54
67	1968	26.63
68	1969	26.89
69	1970	26.78

Processing to remove illegal values

This step is to check for any illegal values within my dataset to ensure accuracy, reliability, and integrity. By cleaning and eliminating such values, the data becomes more suitable for analysis, preventing errors throughout the project.

Hence, I will be conducting a thorough examination of the dataset to identify and address any missing values, ensuring that all relevant information is accounted for. Additionally, I'll be scrutinizing the dataset for duplicates, aiming to eliminate any redundant entries that could skew the integrity of the data.

In [41]:

```
# Check for missing values
if (airtemp_df.isna().sum() == 0).all():
    print("No missing values")
else:
    print("There are {} missing values in the dataframe".format(missing_values.sum()))

# Fill missing values with a specific value or method
# df_filtered.fillna(value=0, inplace=True)

# Drop rows with missing values
# df_filtered.dropna(inplace=True)
```

No missing values

In [42]:

```
# Identify and count the number of duplicates
airtemp_df.duplicated().sum()
print("Number of duplicates:", airtemp_df.duplicated().sum())

# Remove duplicates if have
airtemp_df.drop_duplicates(inplace=True)
```

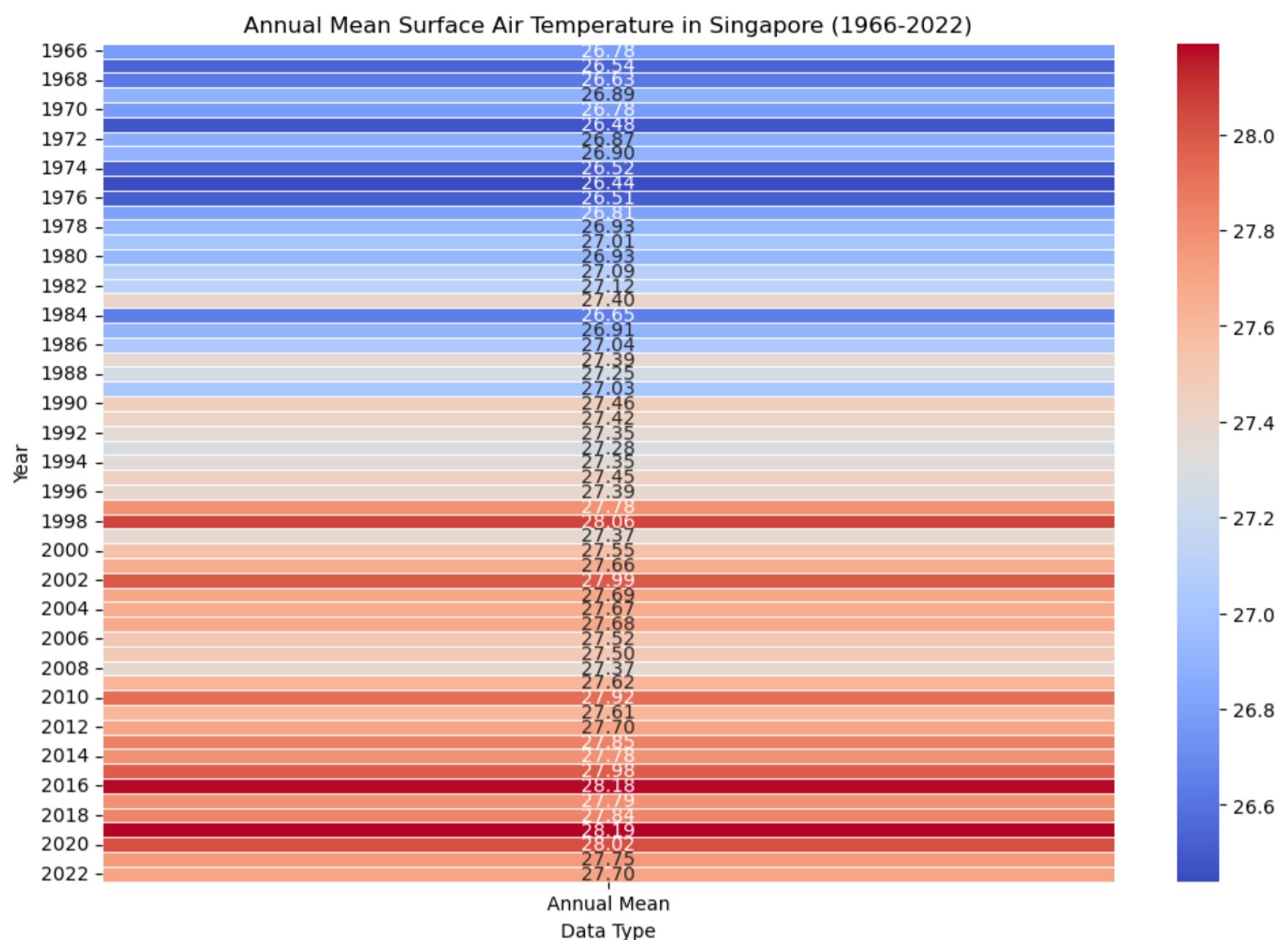
Number of duplicates: 0

Graph of Annual mean surface air temperature in Singapore from 1966 to 2022

In [43]:

```
# Create a heatmap using seaborn
plt.figure(figsize=(12, 8))
sns.heatmap(airtemp_df.pivot_table(index='Year', values='Annual Mean', aggfunc='mean'), cmap='coolwarm', annot=True, fmt=".2f")
plt.title('Annual Mean Surface Air Temperature in Singapore (1966-2022)')
plt.xlabel('Data Type')
```

```
plt.ylabel('Year')
plt.show()
```



Understanding the graph

I have chosen to use a heatmap as a visualization for the Annual Temperature as it is easy to see how the temperature has increased over time from the blue to red gradient. The intensity of the color in each cell of the heatmap reflects the magnitude of the temperature for the corresponding year. Darker and more intense colors (closer to red) indicate higher temperatures, while lighter colors (closer to blue) represent lower temperatures. For example, the annual mean temperature in 1966 is 26.78 degrees celcius and it is a light blue color whereas in 2020 the annual mean temperature is 28.02 degrees and it is a bright red color. As can be seen above, the shift from blue to red from 1966 to 2022 suggests an increase in temperature over the years. In 2019 the temperture is exceptionally high at 28.19 degrees, and as noted is also the time where the COVID19 pandemic happened. Overall, this observation aligns with the broader context of global climate change and its impact on regional temperatures.

4.2 Annual Dengue Cases

Reading CSV files

```
In [44]: annual_dengue_df = pd.read_csv("annual-dengue-cases-1966-2021.csv", header=None, skiprows=10)
annual_dengue_df.head()
```

```
Out[44]:
```

0	1	2	3	4	5	6	7	8	9	...	47	48	49	50	51	52	53	
0	Data Series	2021.0	2020.0	2019.0	2018.0	2017.0	2016.0	2015.0	2014.0	2013.0	...	1975.0	1974.0	1973.0	1972.0	1971.0	1970.0	1969.0
1	Dengue/DHF	5261.0	35266.0	15979.0	3283.0	2767.0	13085.0	11294.0	18326.0	22170.0	...	59.0	229.0	1255.0	12.0	19.0	32.0	60.0
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN						
3	Footnotes:	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN						
4	No cases of plague, smallpox and yellow fever ...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN						

5 rows × 57 columns

Cleaning dataset and removing illegal values

```
In [45]: # Display only rows 9 and 10
annual_dengue_df = annual_dengue_df.iloc[0:2, :]

# convert floating point numbers to integer
numeric_cols = annual_dengue_df.select_dtypes(include='number').columns
annual_dengue_df[numeric_cols] = annual_dengue_df[numeric_cols].applymap(lambda x: int(x) if x == x else x)
annual_dengue_df.head()
```

```
Out[45]:
```

	0	1	2	3	4	5	6	7	8	9	...	47	48	49	50	51	52	53	54	55	56
0	Data Series	2021	2020	2019	2018	2017	2016	2015	2014	2013	...	1975	1974	1973	1972	1971	1970	1969	1968	1967	1966
1	Dengue/DHF	5261	35266	15979	3283	2767	13085	11294	18326	22170	...	59	229	1255	12	19	32	60	158	355	162

2 rows × 57 columns

```
In [46]: # transpose
annual_dengue_df = annual_dengue_df.transpose()
# change column name
annual_dengue_df.columns = annual_dengue_df.iloc[0]
# remove the first row
annual_dengue_df = annual_dengue_df.drop(0)
# rename headers
annual_dengue_df = annual_dengue_df.rename(columns={'Data Series': 'Year'})
annual_dengue_df = annual_dengue_df.rename(columns={'Dengue/DHF': 'Annual Cases'})
annual_dengue_df.head()
```

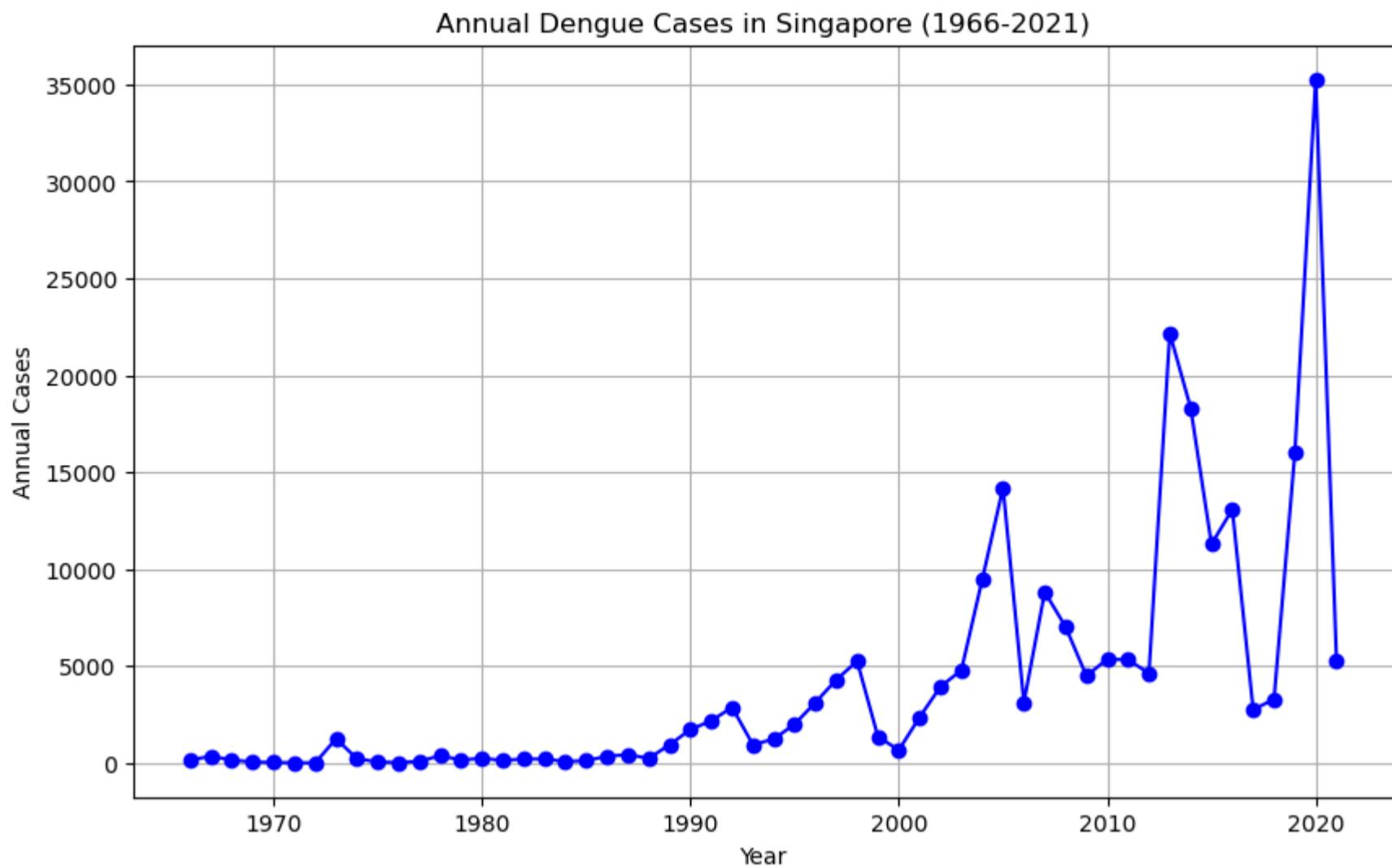
```
Out[46]:
```

	Year	Annual Cases
1	2021	5261
2	2020	35266
3	2019	15979
4	2018	3283
5	2017	2767

Graph of Annual dengue cases in Singapore from 1966 to 2021

```
In [47]: x_axis = annual_dengue_df['Year']
y_axis = annual_dengue_df['Annual Cases']

# Plotting the Line graph
plt.figure(figsize=(10, 6))
plt.plot(x_axis, y_axis, marker='o', linestyle='-', color='b')
plt.title('Annual Dengue Cases in Singapore (1966-2021)')
plt.xlabel('Year')
plt.ylabel('Annual Cases')
plt.grid(True)
plt.show()
```



Understanding the graph

As can be seen in the line graph above, the line slopes upward, indicating a consistent increase in dengue incidence over the years. This suggests a long-term trend of rising dengue cases in Singapore. In 2020, there is an observable spike of dengue cases (about 35000 cases). This spike aligns with the onset of the COVID-19 pandemic, and the correlation could be indicative of the pandemic's influence on dengue transmission which I will be analysing later on.

4.3 Analysis on Dengue Cases and Temperature from 1966 to 2021

```
In [48]: # assuming 'Year' is the column in both DataFrames
annual_dengue_df['Year'] = annual_dengue_df['Year'].astype(str)
airtemp_df['Year'] = airtemp_df['Year'].astype(str)

# merge the two DataFrames based on the 'Year' column
merged_df = pd.merge(annual_dengue_df, airtemp_df, on='Year', how='inner')

# sort the DataFrame by 'Year' 1966-2021
merged_df.sort_values(by='Year', inplace=True)

# create a figure and axis
fig, ax1 = plt.subplots(figsize=(12, 8))

# bar plot for Dengue Cases with rotated x-axis labels
sns.barplot(x='Year', y='Annual Cases', data=merged_df, label='Dengue Cases', ax=ax1, alpha=0.7, palette="rocket_r")

# set y-axis label for Dengue Cases
ax1.set_ylabel('Number of Dengue Cases (Bar)', color='blue')

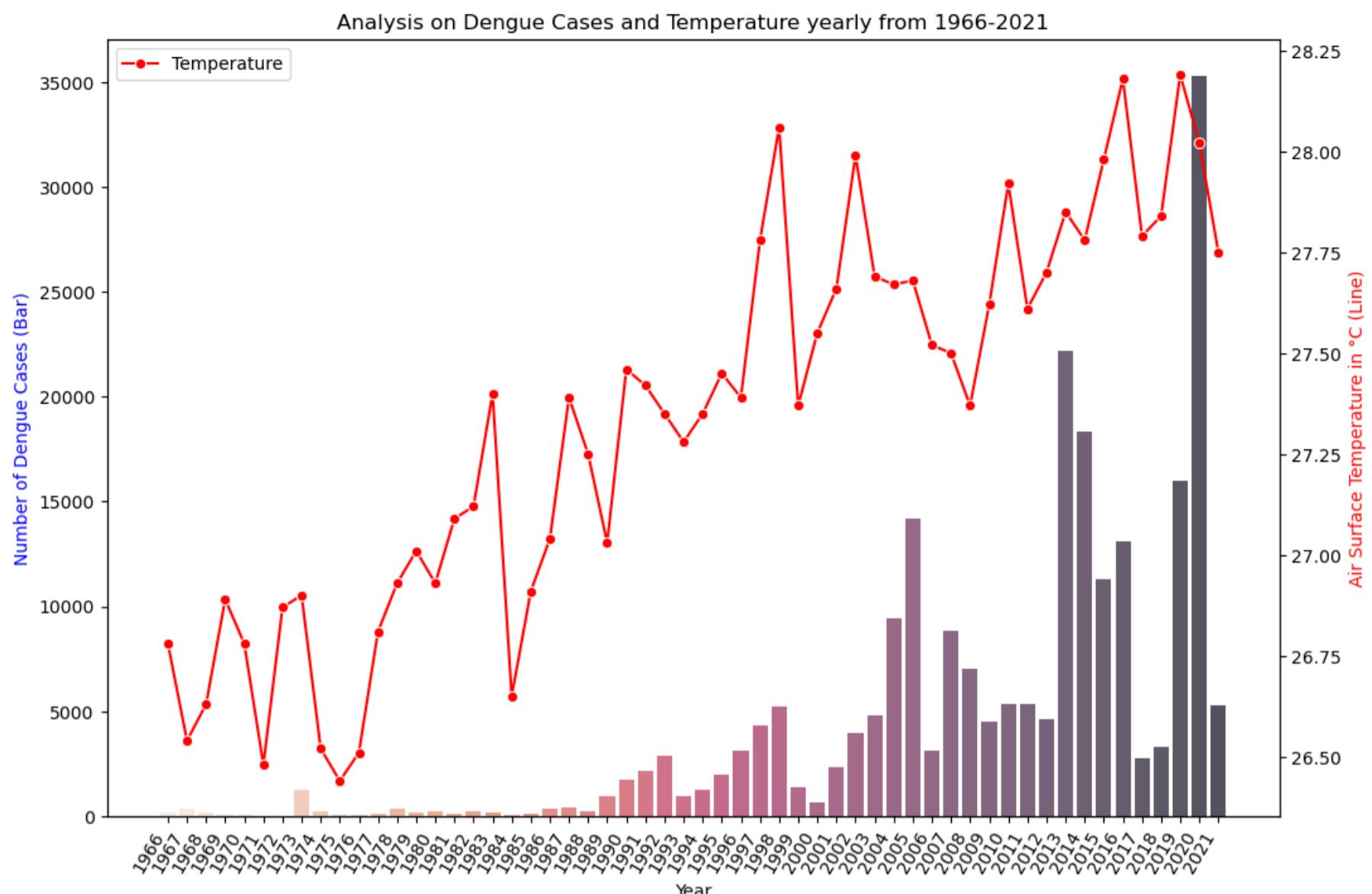
# rotate x-axis labels for Dengue Cases
ax1.set_xticklabels(ax1.get_xticklabels(), rotation=60, ha='right')

# create a second y-axis to plot the temperature line
ax2 = ax1.twinx()
sns.lineplot(x='Year', y='Annual Mean', data=merged_df, marker='o', linestyle='-', color='red', label='Temperature', ax=ax2)

# set y-axis label for Temperature
ax2.set_ylabel('Air Surface Temperature in °C (Line)', color='red')

# set title and x-axis label
plt.title('Analysis on Dengue Cases and Temperature yearly from 1966-2021')
plt.xlabel('Year')

# display the plot
plt.show()
```



Understanding the graph

Correlation between Air Surface Temperature and Dengue Incidence:

- Temperature: The line plot representing the annual mean of air surface temperature shows a general increase from around 26°C in 1966 to about 28°C in 2021. This indicates a trend of rising temperatures over the years.
- Dengue Cases: The bar plot representing the number of dengue cases also shows a general increase over the years, from 162 cases in 1966 to 5261 cases in 2021. This suggests a trend of increasing dengue incidence.

The overall upward trend in both temperature and dengue cases suggests a positive correlation. As temperatures rise, there is an observable increase in the number of dengue cases.

However, while the graph shows a correlation between temperature and dengue cases, it's important to note that correlation does not imply causation. Other factors and interactions could be influencing the observed trends.

Consideration of other factors

It is known that the Aedes mosquito breed more rapidly in stagnant water. Considering Singapore's tropical humid and wet climate, rainfall could also be a factor affecting the number of dengue cases.

Looking at the above graph, we can also see that there was a spike in dengue cases in 2020, at a staggering number of 35266 cases, which is much higher than in 2021, even though the annual mean temperature dropped from the previous year. Hence, I will also be analysing a special case: the spike in Dengue cases in 2020, when the COVID-19 pandemic happened. This might be due to changes in human behaviour during the pandemic, which might impact mosquito breeding and dengue transmission.

This further analysis and consideration of additional factors can provide a more comprehensive understanding of the factors influencing dengue incidence.

4.4 Annual Rainfall

Reading CSV files

```
In [49]: annual_rainfall_df = pd.read_csv("annual-precipitation-of-singapore-1901-2022.csv")
annual_rainfall_df.head()
```

	Category	Annual Mean	5-yr smooth
0	1901	2106.77	2277.17
1	1902	2334.47	2307.09
2	1903	2541.25	2337.37
3	1904	2306.62	2368.47
4	1905	2459.92	2400.59

Cleaning dataset and removing illegal values

```
In [50]: # cleaning data to show only year 1966 onwards to 2022
annual_rainfall_df = annual_rainfall_df.rename(columns={'Category': 'Year'})

# convert the 'Year' column to numeric, handling errors by coercing invalid values to NaN
annual_rainfall_df['Year'] = pd.to_numeric(annual_rainfall_df['Year'], errors='coerce')

# filter data for the desired range of years (1966 to 2021)
annual_rainfall_df = annual_rainfall_df[annual_rainfall_df['Year'].between(1966, 2022)]

annual_rainfall_df.head()
```

	Year	Annual Mean	5-yr smooth
65	1966	2563.65	2494.74
66	1967	3019.72	2444.10
67	1968	2253.60	2384.74
68	1969	2407.16	2328.48
69	1970	2318.44	2286.03

```
In [51]: # Check for NaN values
print(merged_df['Annual Mean'].isnull().sum())
# Check for non-numeric values
non_numeric_values = merged_df[~merged_df['Annual Mean'].apply(np.isreal)]['Annual Mean']
print(non_numeric_values)
```

0
Series([], Name: Annual Mean, dtype: float64)

Graph of Annual Rainfall in Singapore from 1966 to 2021

```
In [52]: # Assuming 'Year' is the column in your DataFrame
annual_rainfall_df['Year'] = annual_rainfall_df['Year'].astype(str)

# Create a figure and axis
fig, ax = plt.subplots(figsize=(12, 8))

# Line plot for Annual Mean
sns.lineplot(x='Year', y='Annual Mean', data=annual_rainfall_df, marker='o', linestyle='-', color='orange', label='Annual Mean')

# Line plot for 5-yr Smooth using backticks
sns.lineplot(x='Year', y='5-yr smooth', data=annual_rainfall_df, marker=None, linestyle='dashdot', color='purple', label='5-yr smooth')

# Set y-axis label for Rainfall
ax.set_ylabel('Rainfall (mm)', color='blue')

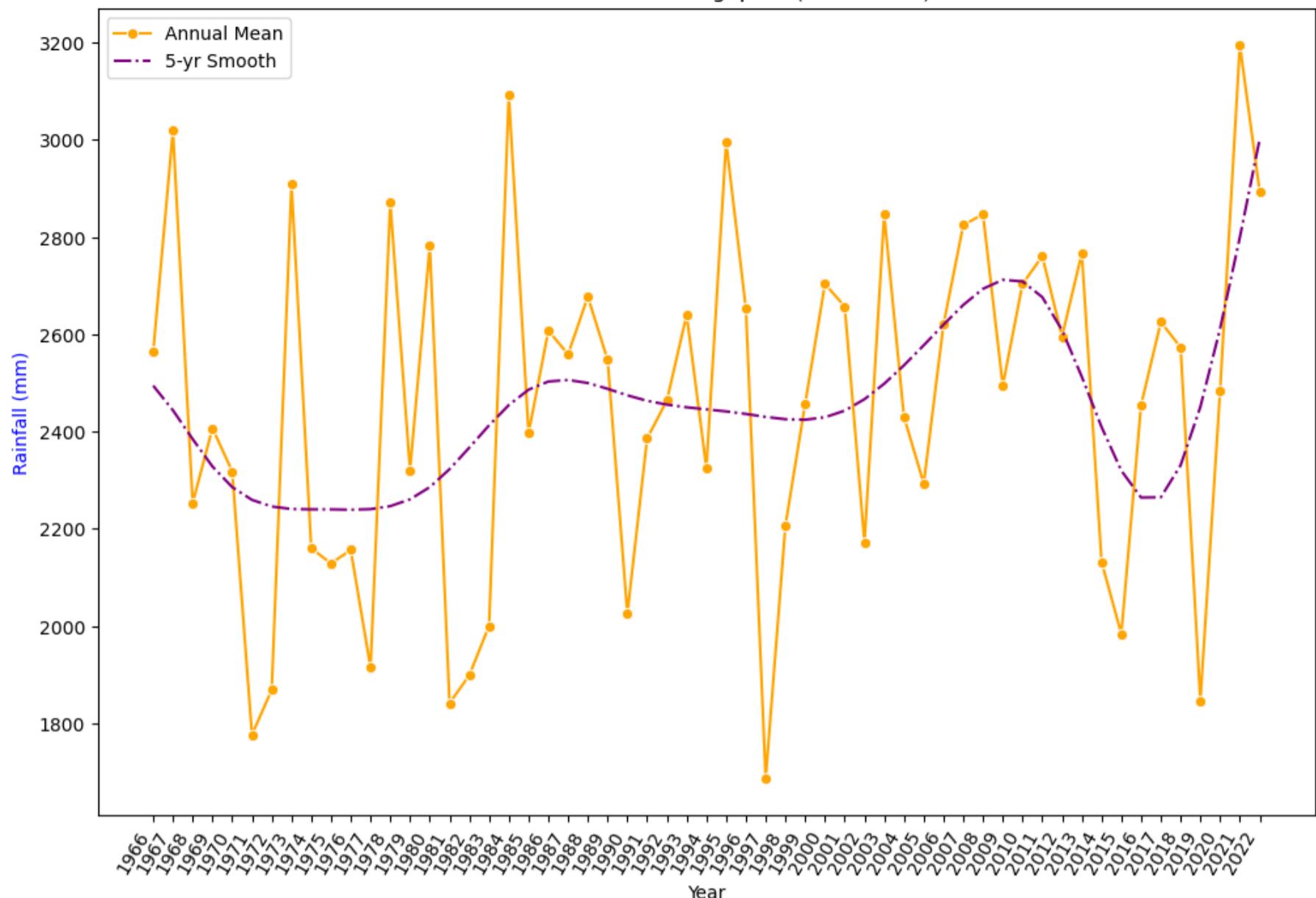
# Rotate x-axis labels
ax.set_xticks(ax.get_xticks()[::1]) # Adjust the interval based on data
ax.set_xticklabels(annual_rainfall_df['Year'][::1], rotation=60, ha='right')

# Set title and x-axis label
plt.title('Annual Rainfall in Singapore (1966-2021)')
plt.xlabel('Year')

# Display the legend
plt.legend()

# Show the plot
plt.show()
```

Annual Rainfall in Singapore (1966-2021)

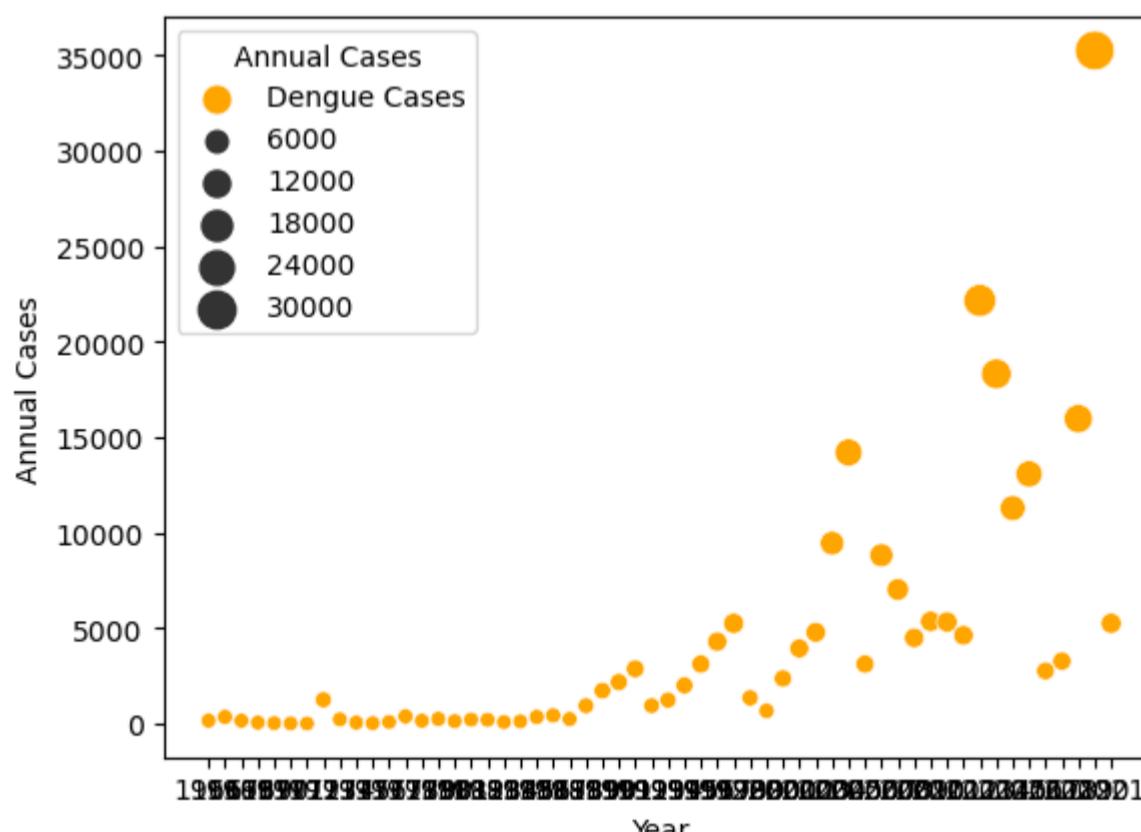


```
In [59]: # removing last column
annual_rainfall_df = annual_rainfall_df.drop(columns=['5-yr smooth'], errors='ignore')
```

4.5 Analysis on Dengue Cases and Rainfall from 1966 to 2021

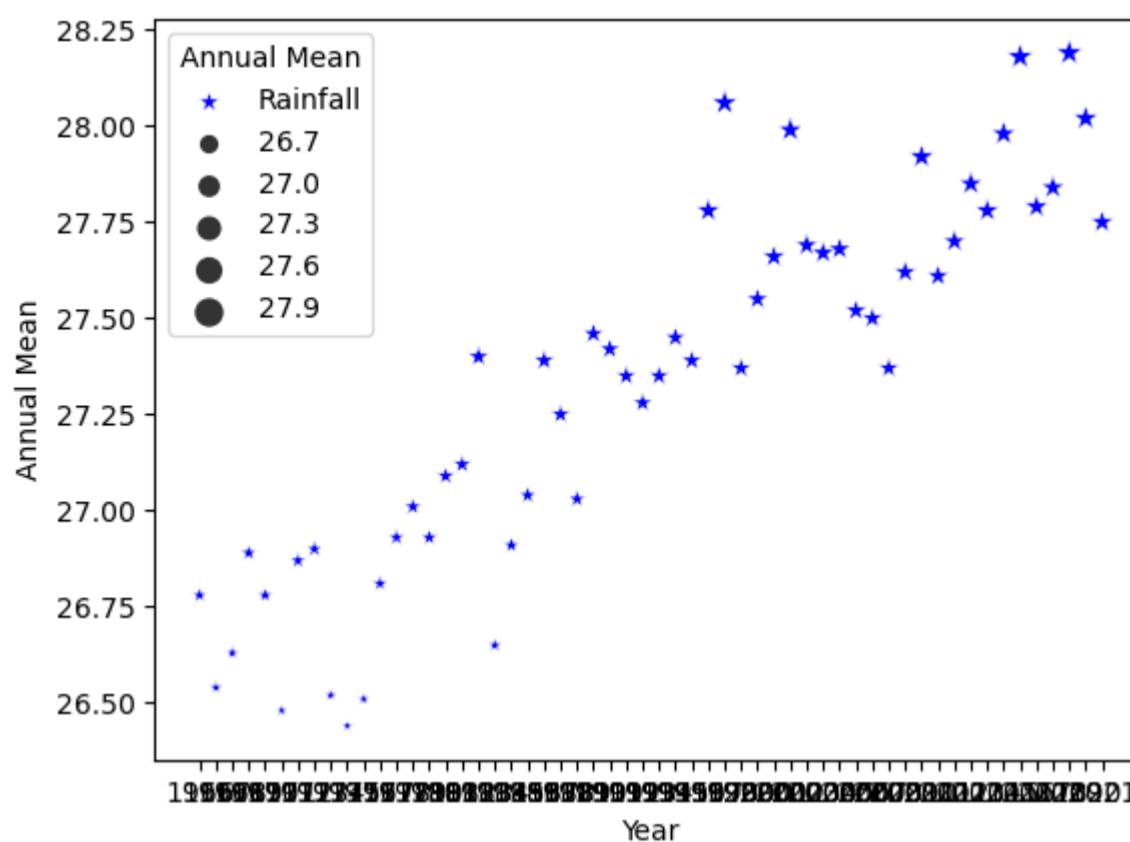
Scatterplot Graph of Dengue Cases

```
In [60]: sns.scatterplot(x='Year', y='Annual Cases', size='Annual Cases', sizes=(30, 200), data=merged_df, color='orange', label='Dengue Cases')
# set y-axis
plt.show()
```



Scatterplot graph of Annual Rainfall

```
In [61]: sns.scatterplot(x='Year', y='Annual Mean', size='Annual Mean', sizes=(20, 100), data=merged_df, color='blue', label='Rainfall')
<Axes: xlabel='Year', ylabel='Annual Mean'>
```



```
In [62]: print(merged_df.dtypes)
```

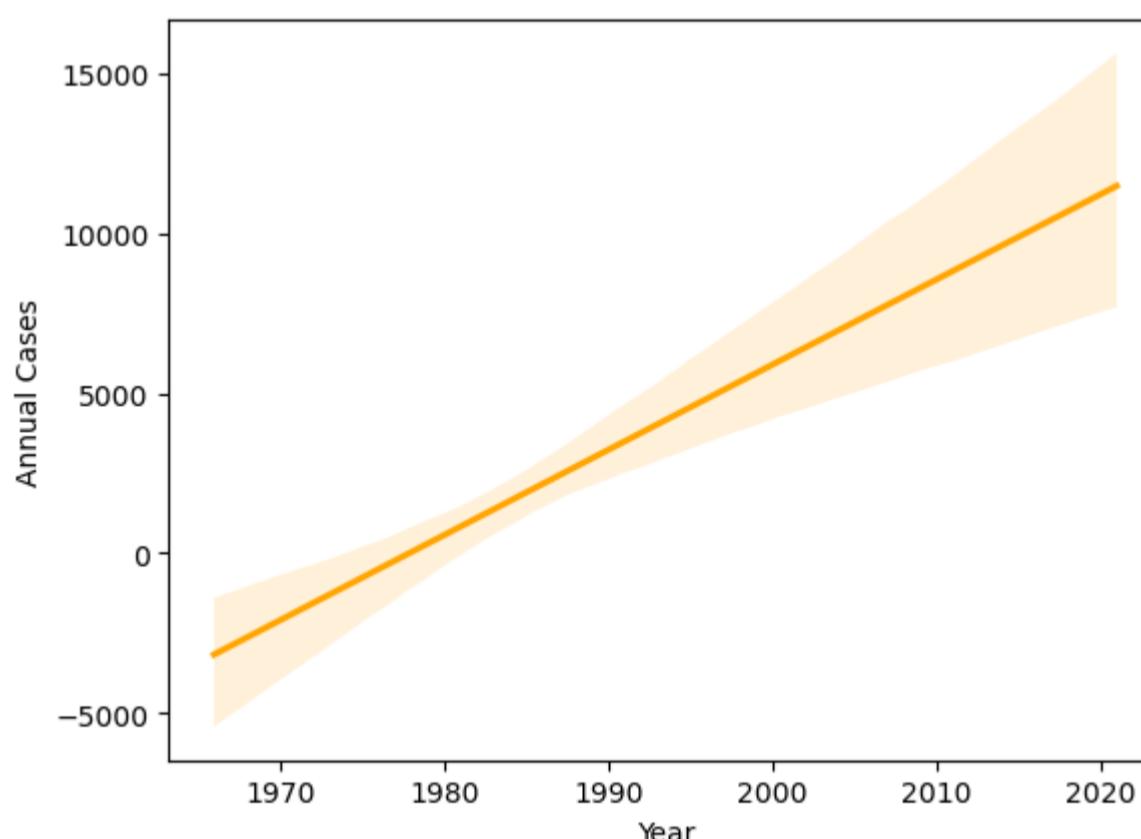
Year	object
Annual Cases	int64
Annual Mean	float64
dtype:	object

```
In [63]: # convert 'Annual Cases' obj to int
merged_df['Annual Cases'] = pd.to_numeric(merged_df['Annual Cases'], errors='coerce')
merged_df['Annual Cases'].dtypes
```

```
Out[63]: dtype('int64')
```

```
In [66]: #draw line of best fit for annual dengue cases
sns.regplot(x='Year', y='Annual Cases', data=merged_df, scatter=False, color='orange')
```

```
Out[66]: <Axes: xlabel='Year', ylabel='Annual Cases'>
```



```
In [67]: print(merged_df.dtypes)
```

Year	int64
Annual Cases	int64
Annual Mean	float64
dtype:	object

Scatterplot graph of Annual Dengue Cases and Annual Rainfall

```
In [65]: annual_dengue_df['Year'] = annual_dengue_df['Year'].astype(str)
annual_rainfall_df['Year'] = annual_rainfall_df['Year'].astype(str)

# merge the two DataFrames based on the 'Year' column
merged_df = pd.merge(annual_dengue_df, annual_rainfall_df, on='Year', how='inner')
merged_df['Year'] = pd.to_numeric(merged_df['Year'], errors='coerce')
```

```

merged_df['Annual Cases'] = pd.to_numeric(merged_df['Annual Cases'], errors='coerce')

# sort the DataFrame by 'Year' 1966-2021
merged_df.sort_values(by='Year', inplace=True)

# Create a figure and axis
fig, ax1 = plt.subplots(figsize=(12, 8))

# Scatter plot for Dengue Cases
sns.scatterplot(x='Year', y='Annual Cases', size='Annual Cases', sizes=(30, 200), data=merged_df, color='orange', label='Dengue Cases')
# set y-axis label for Dengue Cases
ax1.set_ylabel('Number of Dengue Cases', color='orange')

# rotate x-axis labels
ax1.set_xticks(range(len(merged_df['Year'])))
ax1.set_xticklabels(ax1.get_xticklabels(), rotation=60, ha='right')

# create a second y-axis to plot the rainfall scatterplot
ax2 = ax1.twinx()
sns.scatterplot(x='Year', y='Annual Mean', size='Annual Mean', sizes=(20, 100), data=merged_df, color='blue', label='Rainfall')
# set y-axis label for rainfall
ax2.set_ylabel('Annual Rainfall (mm)', color='blue')

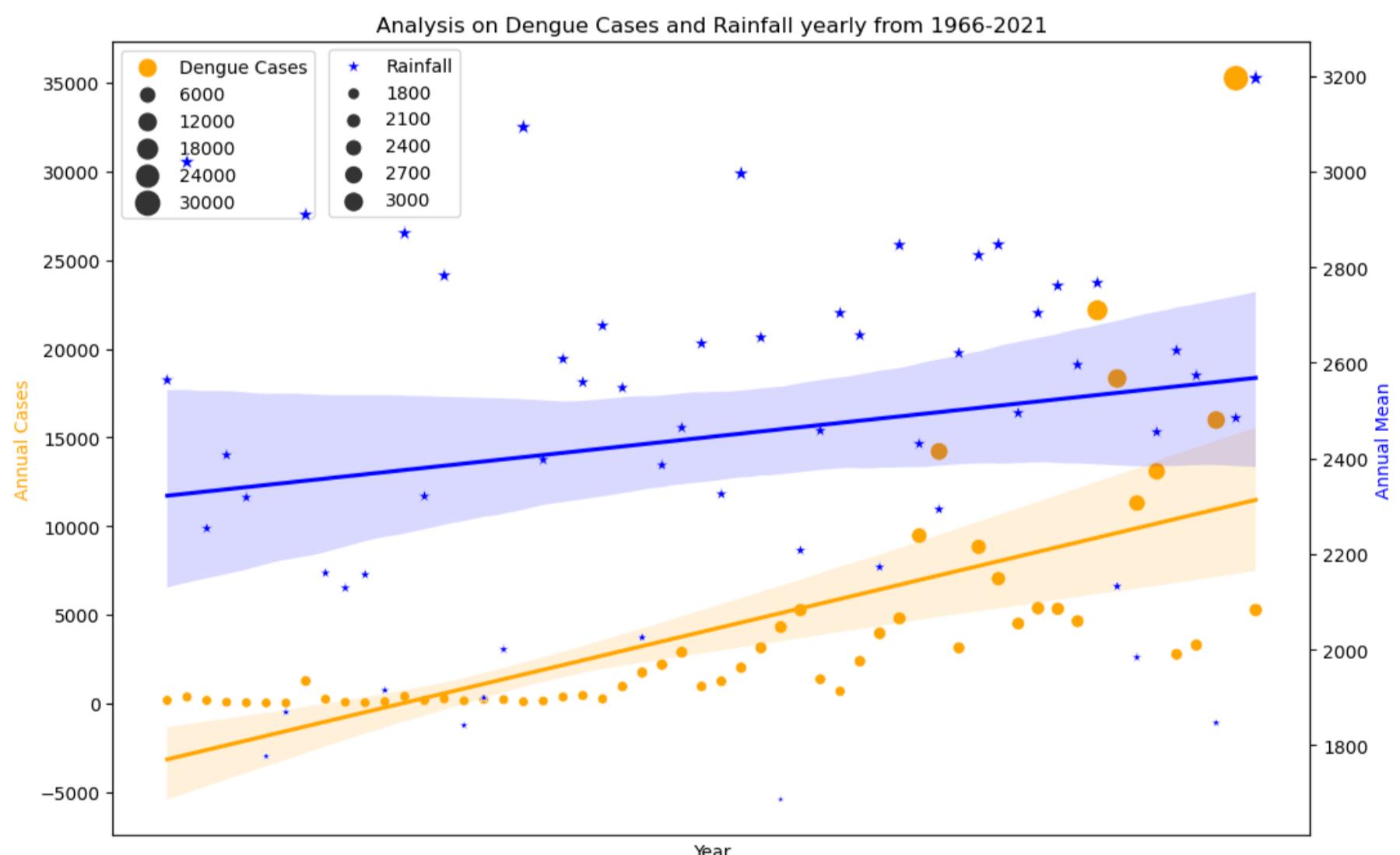
# Line of best fit for dengue cases
sns.regplot(x='Year', y='Annual Cases', data=merged_df, scatter=False, color='orange', ax=ax1)
# Line of best fit for rainfall
sns.regplot(x='Year', y='Annual Mean', data=merged_df, scatter=False, color='blue', ax=ax2)

# set title and x-axis label
plt.title('Analysis on Dengue Cases and Rainfall yearly from 1966-2021')
plt.xlabel('Year')

# Adjust the position of Legends to avoid overlap
ax1.legend(loc='upper left')
ax2.legend(loc=(0.18, 0.78))

# Show the plot
plt.show()

```



Understanding the graph

In the presented graph, I have opted for a scatterplot representation to illustrate the association between annual rainfall and the incidence of dengue cases.

Upon examination of the scatterplot, it is evident that there is no straightforward correlation between annual rainfall and the frequency of dengue cases. Despite an overall upward trend in both variables from 1966 to 2021, with dengue cases rising from 162 cases to 5261 cases and rainfall increasing from 2563.65mm to 3195.10mm, establishing a positive relationship becomes challenging. This is due to the consistent pattern of rainfall over the years, displaying both high and low levels without a discernible trend.

Notably, the year 2020 stands out as a peak in both rainfall and dengue cases. This anomaly reinforces the notion that factors beyond precipitation contribute to the incidence of dengue. While it is well-established that Aedes mosquitoes thrive in stagnant water, the singular emphasis on rainfall may not be sufficient.

Considering the broader context, it is apparent that other variables, such as human behavior, play a significant role. This aligns with the earlier observation of the relationship between temperature and dengue cases. Therefore, the multifaceted nature of these contributing factors underscores the need for a comprehensive understanding of the dynamics influencing the rise in dengue cases.

4.6 Special case: Spike in Dengue Cases due to COVID-19

Analyzing the spike in dengue cases during the COVID-19 pandemic in 2020 is crucial due to the potential impact of human behavior on the breeding and transmission of the dengue virus. The COVID-19 pandemic lockdown led to significant changes in daily routines like water storage practices and waste management, which could have induced favorable conditions for mosquito breeding.

I will be extracting relevant data from the National Environment Agency (NEA), specifically their quarterly dengue surveillance reports. These reports highlight the top 5 breeding habits for each quarter, providing insights into the most common habits contributing to dengue transmission. By leveraging this data, I aim to create a word cloud visualization to emphasize the prevalence of specific breeding habits, aiding in targeted interventions for dengue prevention and control.

Extracting quarterly surveillance data from PDF

```
In [106...]:  
import tabula.io  
# Specify the path to your PDF file  
pdf_path1 = 'q1-2020-dengue-surveillance-data.pdf'  
  
# Read PDF into a DataFrame  
df1 = tabula.io.read_pdf(pdf_path1, pages='all', multiple_tables=True)  
  
# Check the structure of the DataFrame  
# for i, table in enumerate(df1):  
#     print(f"Table {i + 1}: \n{table}\n{'=' * 40}\n")  
  
# Display table  
for i, df in enumerate(df1):  
    # Extract the first row as the header  
    header = ["Top 5 Breeding Habitats in Homes", "Top 5 Breeding Habitats in Public Areas"]  
  
    # Exclude the first row and first column  
    df_subset = df.iloc[1:, 1:]  
  
    # Convert DataFrame to HTML table string  
    df_subset.columns = header  
    table_str1 = df_subset.to_html(index=False)  
  
    # Display the HTML table  
    display(HTML(f"<h4>Quarter 1</h4>{table_str1}"))
```

Quarter 1

Top 5 Breeding Habitats in Homes		Top 5 Breeding Habitats in Public Areas	
Domestic Containers		Covered Perimeter Drains	
Ornamental Containers (vases)		Discarded Receptacles	
Flower Pot Plates/Trays		Gully Traps	
Water Fountains		Covered Carpark Drains	
Toilet Bowls/Cisterns		Inspection Chambers	

```
In [108...]:  
# Specify the path to your PDF file  
pdf_path2 = 'q2-2020-dengue-surveillance-data.pdf'  
  
# Read PDF into a DataFrame  
df2 = tabula.io.read_pdf(pdf_path2, pages='all', multiple_tables=True)  
  
# # Filter out unwanted tables  
desired_num_columns = 3  
df2 = [df for df in df2 if len(df.columns) == desired_num_columns]  
  
# Display table  
for i, df in enumerate(df2):  
    # remove the first column  
    df_subset = df.iloc[:, 1:]  
  
    # Convert the DataFrame to an HTML table string  
    table_str2 = df_subset.to_html(index=False)  
  
    # Display the HTML table in the notebook  
    display(HTML(f"<h4>Quarter 2</h4>{table_str2}"))
```

Quarter 2

Top 5 Breeding Habitats in Homes Top 5 Breeding Habitats in Public Areas

Domestic Containers	Covered Perimeter Drains
Flower Pot Plates/Trays	Discarded Receptacles
Ornamental Containers (vases)	Plants (e.g. Hardened soil and leaf axils)
Toilet Bowls/Cisterns	Domestic containers
Plants (e.g. Hardened soil and leaf axils)	Gully Traps

In [109...]

```
# Specify the path to your PDF file
pdf_path3 = 'q3-2020-dengue-surveillance-data.pdf'

# Read PDF into a DataFrame
df3 = tabula.io.read_pdf(pdf_path3, pages='all', multiple_tables=True)

# Display table
for i, df in enumerate(df3):
    # remove the first column
    df_subset = df.iloc[:, 1:]

    # Convert the DataFrame to an HTML table string
    table_str3 = df_subset.to_html(index=False)

    # Display the HTML table in the notebook
    display(HTML(f"<h4>Quarter 3</h4>{table_str3}"))
```

Quarter 3

Top 5 Breeding Habitats in Homes Top 5 Breeding Habitats in Public Areas

Domestic Containers	Discarded Receptacles
Flower Pot Plates/Trays	Covered Perimeter Drains
Ornamental Containers (vases)	Gully Traps
Toilet Bowls/Cisterns	Plants (e.g. Hardened soil and leaf axils)
Plants (e.g. Hardened soil and leaf axils)	Domestic containers

In [110...]

```
# Specify the path to your PDF file
pdf_path4 = 'q4-2020-dengue-surveillance-data.pdf'

# Read PDF into a DataFrame
df4 = tabula.io.read_pdf(pdf_path4, pages='all', multiple_tables=True)

# Display table
for i, df in enumerate(df4):
    # remove the first column
    df_subset = df.iloc[:, 1:]

    # Convert the DataFrame to an HTML table string
    table_str4 = df_subset.to_html(index=False)

    # Display the HTML table in the notebook
    display(HTML(f"<h4>Quarter 4</h4>{table_str4}"))
```

Quarter 4

Top 5 Breeding Habitats in Homes Top 5 Breeding Habitats in Public Areas

Domestic Containers	Discarded Receptacles
Flower Pot Plates/Trays	Covered Perimeter Drains
Ornamental Containers (vases)	Plants (e.g. Hardened soil and leaf axils)
Plants (e.g. Hardened soil and leaf axils)	Gully Traps
Planter Boxes	Domestic Containers

Common mosquito breeding places

While there are many different types of breeding places where breeding habits take place, (such as in rooftops and covered linkways, corridors and staircases, toilets, ground level (pump room/switch room), apron of building, and in bin rooms), certain breeding habits have surfaced as notably prevalent, both in homes and public areas.

From the tables shown above, it can be seen that the breeding habits in homes are:

- Domestic Containers
- Flower Pot Plates/Trays

- Ornamental Containers (vases)
- Toilet Bowls/Cisterns
- Plants (e.g. Hardened soil and leaf axils)
- Planter Boxes
- Water Fountains

The breeding habits in public areas are:

- Discarded Receptacles
- Covered Perimeter Drains
- Plants (e.g. Hardened soil and leaf axils)
- Gully Traps
- Domestic Containers
- Covered Carpark Drains
- Inspection Chambers

WordCloud

To discern overarching trends, a comparative wordcloud encompassing all breeding habits across the dataset will be created. This broader perspective allows for the identification of the most frequently occurring "habit" across all entries. The frequency of each term is visually emphasized, providing insights into the characteristics that contribute significantly to dengue transmission.

```
In [111...]
# quarter 1 of 2020
homes1 = [
    "Domestic Containers",
    "Ornamental Containers (vases)",
    "Flower Pot Plates/Trays",
    "Water Fountains",
    "Toilet Bowls/Cisterns"
]

public1 = [
    "Covered Perimeter Drains",
    "Discarded Receptacles",
    "Gully Traps",
    "Covered Carpark Drains",
    "Inspection Chambers"
]
```

```
In [112...]
# quarter 2 of 2020
homes2 = [
    "Domestic Containers",
    "Flower Pot Plates/Trays",
    "Ornamental Containers (vases)",
    "Toilet Bowls/Cisterns",
    "Plants (e.g. Hardened soil and leaf axils)"
]

public2 = [
    "Covered Perimeter Drains",
    "Discarded Receptacles",
    "Plants (e.g. Hardened soil and leaf axils)",
    "Domestic containers",
    "Gully Traps"
]
```

```
In [113...]
# quarter 3 of 2020
homes3 = [
    "Domestic Containers",
    "Flower Pot Plates/Trays",
    "Ornamental Containers (vases)",
    "Toilet Bowls/Cisterns",
    "Plants (e.g. Hardened soil and leaf axils)"
]

public3 = [
    "Discarded Receptacles",
    "Covered Perimeter Drains",
    "Gully Traps",
    "Plants (e.g. Hardened soil and leaf axils)",
    "Domestic containers"
]
```

```
In [114...]
# quarter 4 of 2020
homes4 = [
    "Domestic Containers",
    "Flower Pot Plates/Trays",
    "Ornamental Containers (vases)",
    "Plants (e.g. Hardened soil and leaf axils)",
    "Planter Boxes"
]
```

```
[1]
public4 = [
    "Discarded Receptacles",
    "Covered Perimeter Drains",
    "Plants (e.g. Hardened soil and leaf axils)",
    "Gully Traps",
    "Domestic containers"
]
```

```
In [115...]: from collections import Counter

# Compile all breeding habits for each category
all_homes = homes1 + homes2 + homes3 + homes4
all_public = public1 + public2 + public3 + public4

# Count the frequency of each breeding habit
frequency_homes = Counter(all_homes)
frequency_public = Counter(all_public)

# Print the frequency for each breeding habit
print("Frequency of breeding habits in homes:")
print(frequency_homes)

print("\nFrequency of breeding habits in public areas:")
print(frequency_public)
```

Frequency of breeding habits in homes:
Counter({'Domestic Containers': 4, 'Ornamental Containers (vases)': 4, 'Flower Pot Plates/Trays': 4, 'Toilet Bowls/Cisterns': 3, 'Plants (e.g. Hardened soil and leaf axils)': 3, 'Water Fountains': 1, 'Planter Boxes': 1})

Frequency of breeding habits in public areas:
Counter({'Covered Perimeter Drains': 4, 'Discarded Receptacles': 4, 'Gully Traps': 4, 'Plants (e.g. Hardened soil and leaf axils)': 3, 'Domestic containers': 3, 'Covered Carpark Drains': 1, 'Inspection Chambers': 1})

```
In [167...]: # Frequencies for each quarter - Homes
frequency_homes1 = Counter(["DomesticContainers", "OrnamentalContainers (vases)", "FlowerPot Plates/Trays", "WaterFountains",
frequency_homes2 = Counter(["DomesticContainers", "FlowerPot Plates/Trays", "OrnamentalContainers (vases)", "ToiletBowls/Cist
frequency_homes3 = Counter(["DomesticContainers", "FlowerPot Plates/Trays", "OrnamentalContainers (vases)", "ToiletBowls/Cist
frequency_homes4 = Counter(["DomesticContainers", "FlowerPot Plates/Trays", "OrnamentalContainers (vases)", "Plants (e.g. Har

# Frequencies for each quarter - Public Areas
frequency_public1 = Counter(["Covered PerimeterDrains", "DiscardedReceptacles", "GullyTraps", "Covered CarparkDrains", "Inspe
frequency_public2 = Counter(["Covered PerimeterDrains", "DiscardedReceptacles", "Plants (e.g. Hardened soil and leaf axils"),
frequency_public3 = Counter(["DiscardedReceptacles", "Covered PerimeterDrains", "GullyTraps", "Plants (e.g. Hardened soil and
frequency_public4 = Counter(["DiscardedReceptacles", "Covered PerimeterDrains", "Plants (e.g. Hardened soil and leaf axils"),
```

Breeding habits in homes:

```
In [183...]: # Combine the breeding habits and their frequencies for each quarter - Homes
all_data_homes = frequency_homes1 + frequency_homes2 + frequency_homes3 + frequency_homes4

# Create a string with breeding habits
text_homes = " ".join(all_data_homes)

# Define additional stopwords
stopwords = set(["e", "g", "Hardened", "soil", "and", "leaf", "axils", "vases", "Perimeter", "Covered", "Plates", "Trays", "C

# Generate the WordCloud
wordcloud = WordCloud(stopwords=stopwords, background_color="white", width=700, height=400, colormap = "magma").generate(text_h

# Display the WordCloud
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis("off")
plt.title('Breeding Habits in Homes')
plt.show()
```



Although the size of the words are relatively similar due to a tie in the frequencies, the term "containers" in the wordcloud dedicated to homes underscores the significance of household items that can harbor stagnant water. This signals a critical area for targeted intervention in dengue prevention strategies within residential spaces.

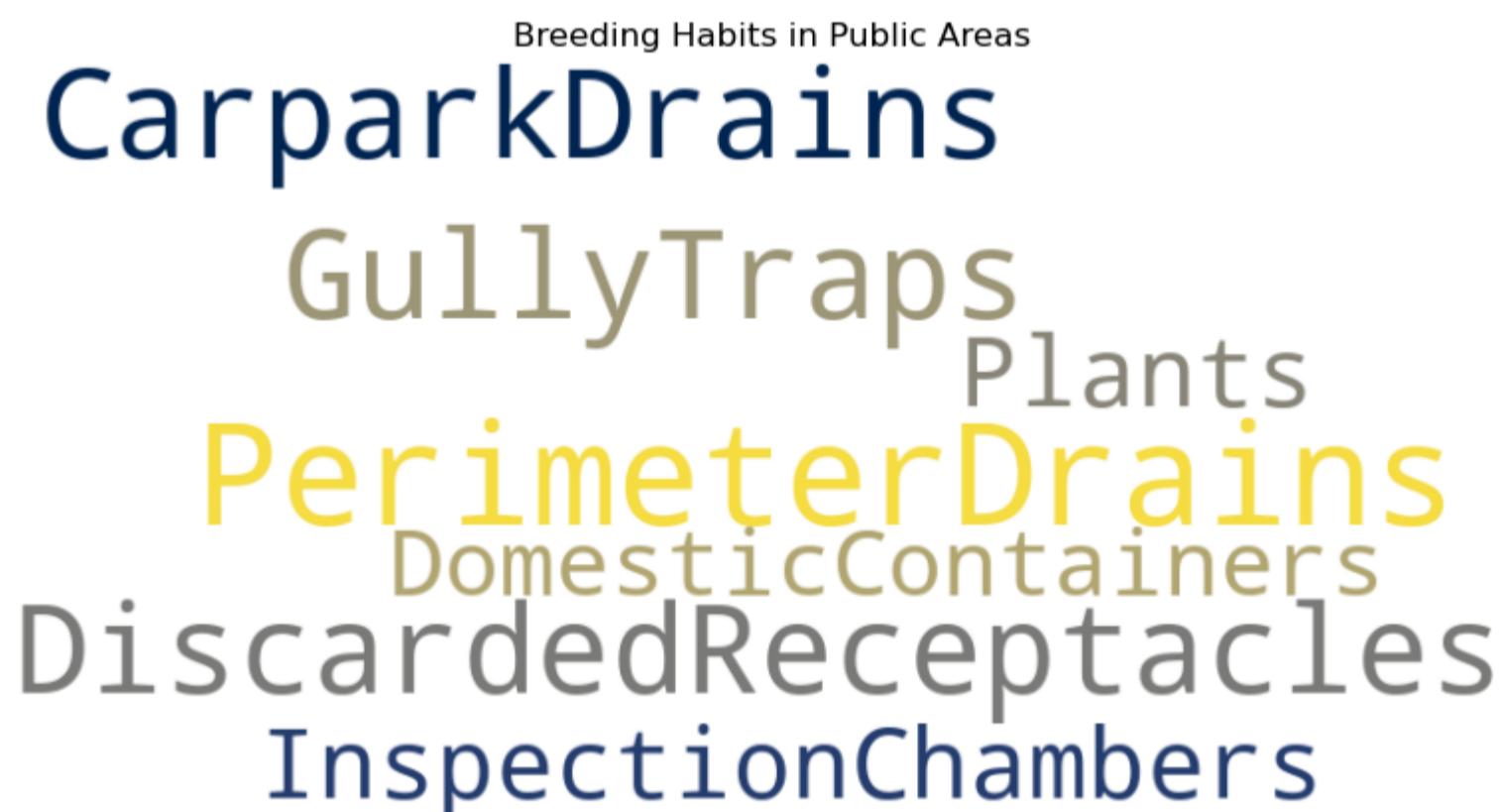
Breeding habits in public areas:

```
In [182...]
# Combine the breeding habits and their frequencies for each quarter - public
all_data_public = frequency_public1 + frequency_public2 + frequency_public3 + frequency_public4

# Create a string with breeding habits
text_homes = " ".join(all_data_public)

# Generate the WordCloud
wordcloud = WordCloud(stopwords=stopwords, background_color="white", width=800, height=400, colormap = "cividis").generate(text_homes)

# Display the WordCloud
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis("off")
plt.title('Breeding Habits in Public Areas')
plt.show()
```



In public areas, the discernible dominance of "drainages" in the respective wordcloud underscores the importance of addressing water accumulation in these spaces. Drainages emerge as key breeding grounds, necessitating strategic measures for public health.

WordCloud of keywords

```
In [177...]
# Assuming 'breeding_habits' is a list of strings containing extracted breeding habits
breeding_habits = ["Domestic Containers", "Covered Perimeter Drains", "Ornamental Containers (vases)", "Discarded Receptacles"]

# Join the breeding habits into a single string
```

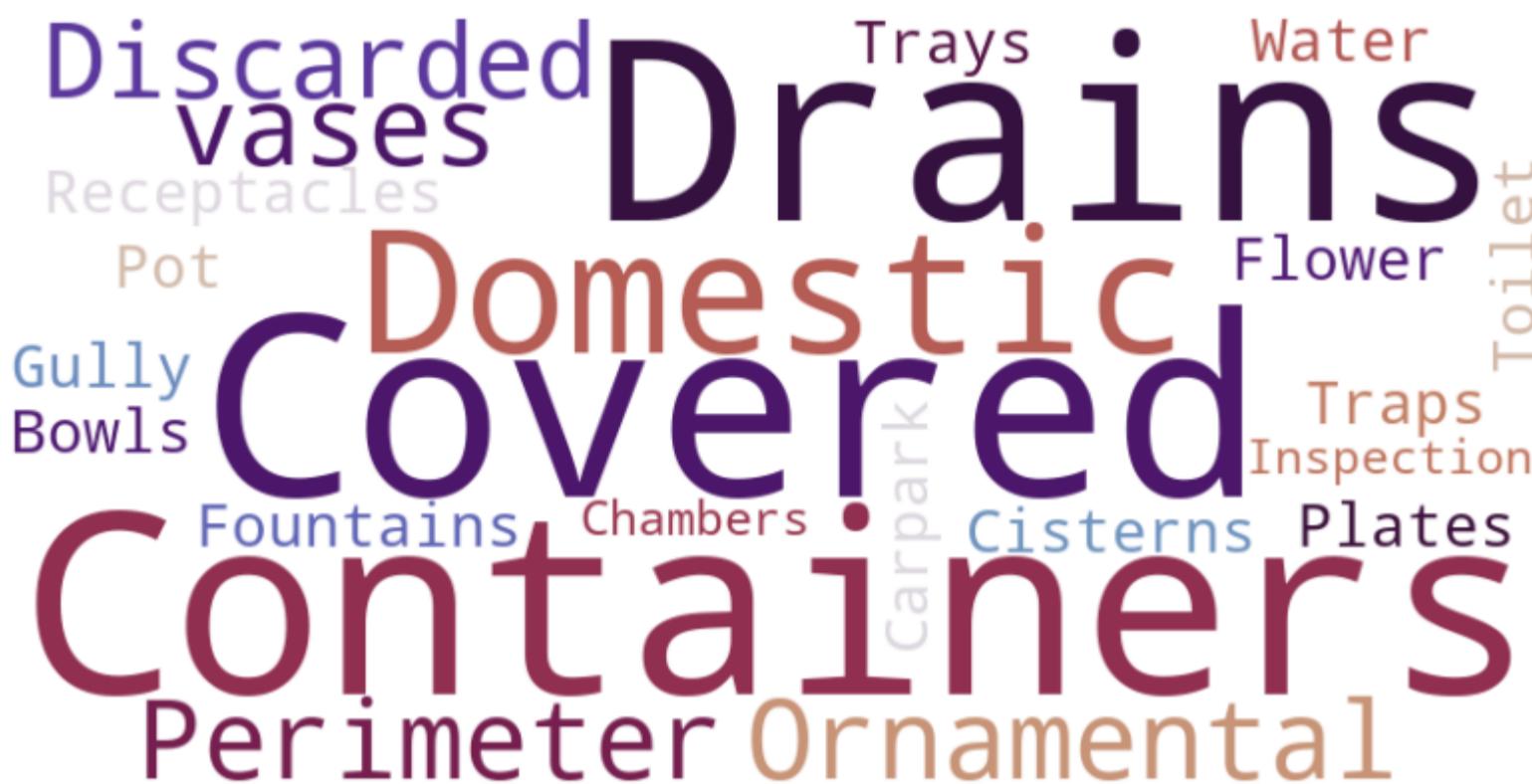
```

text = " ".join(breeding_habits)

# Generate the word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white', colormap = "twilight").generate(text)

# Display the word cloud using matplotlib
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()

```



Understanding the WordClouds

Acknowledging the complexity of prominence determination, where the size of wordclouds and frequency create ties, the emphasis shifts to the wordcloud of keywords. Here, a shared emphasis on "drains", "covered" and "containers" emerge as the most prominent terms, suggesting their heightened importance in the dataset. From this we can also see that there is an overarching characteristic - these findings underscore that places capable of holding and storing water consistently contribute to dengue transmission.

This highlights the importance of individuals heightening their awareness of behaviors that can inadvertently create favorable conditions for breeding grounds. By targeting these specific areas — be it containers in homes or drainages in public spaces, community-driven initiatives can play a pivotal role in mitigating the risk of dengue transmission.

4.7 Conclusion

In the course of this research, I delved into three pivotal factors — air surface temperature, rainfall, and human behavior, specifically breeding habits — to unravel the intricate dynamics influencing the surge in dengue incidence. Contrary to a simplistic understanding, my findings underscore that dengue's escalation is not solely attributed to a single factor but rather emerges from the interplay of various contributors.

The impact of climate, exemplified by rising air surface temperatures and variations in rainfall patterns, emerged as a substantial influence on dengue transmission. However, I also probed into the realm of human behavior, shedding light on breeding habits that significantly contribute to the proliferation of dengue cases.

Understanding that climate plays a significant role in the increase in dengue incidence should not diminish the urgency of human intervention. This study hopes to remind that, while climate conditions set the stage, human choices can modulate the course of the dengue narrative. Beyond addressing breeding habits, individuals and societies possess the means to enact broader changes, such as reducing carbon emissions and adopting sustainable practices. By doing so, we not only mitigate the immediate impact of dengue but also contribute to the larger global effort to curb the acceleration of climate change.

In this context, it's commendable to note that Singapore has proactively implemented enforcement regimes like the Mozzie Wipeout, a concerted effort to eliminate mosquito breeding grounds. However, more individuals and communities should actively participate in these initiatives as overcomplacency can increase the spread of dengue. While government-led programs lay the groundwork, individual actions, such as eliminating stagnant water and practicing responsible waste management, become crucial in fortifying the defenses against dengue transmission.

In conclusion, the fight against dengue demands a holistic approach that recognizes the correlation between climate and human behavior. Understanding the importance will give us the ability and responsibility to make choices that safeguard both our health and the health of our planet.

5 Ethical considerations

5.1 Data provenance

The datasets used in this research are derived from multiple sources, mainly from Singstat.gov.sg, climateknowledgeportal.worldbank.org and nea.gov.sg.

The department of statistics singapore (singstat.gov.sg) provides demographic and socioeconomic data for Singapore, and it is publicly accessible. The data is considered open and is governed by the terms specified in the terms of use on the Singstat website:
<https://www.singstat.gov.sg/terms-of-use>

Climate-related data from the World Bank's Climate Knowledge Portal is sourced from various sources, including meteorological stations and is also open and can be freely accessed. The terms of use for datasets can be found on their website:
<https://www.worldbank.org/en/about/legal/terms-of-use-for-datasets>

The National Environment Agency (NEA) in Singapore provides crucial data for this research, and access is conducted ethically and in compliance with any licensing agreements. The specific terms of use for datasets from NEA.gov.sg can be found on their website:
<https://www.nea.gov.sg/corporate-functions/terms-of-use>

5.2 Usage/reusage of data considerations

As all datasets are acquired from open sources, the analysis does not aim to create new forms of intellectual property. This research aims to contribute insights to the public domain, promoting awareness of the impacts of climate change, specifically the increase in temperature and rainfall, on dengue incidence. Proper attribution to the data sources, including Singstat, the World Bank's Climate Knowledge Portal, and NEA, will be given in any publication or presentation resulting from this analysis, adhering to the specified terms of use requirements outlined by each data source.

5.3 Ethical implications

Careful consideration is given when identifying correlations between climate variables and dengue incidence, steering clear of making discriminatory assumptions. My research aims to disseminate information responsibly and conclusions drawn are grounded in statistical rigor. Acknowledging the potential societal impact of the results, the analysis will refrain from the dissemination of harmful assumptions by prioritizing accuracy.

5.4 Data processing pipeline considerations

Detailed documentation throughout the research will elucidate the steps taken during data processing, providing transparency of the analysis. Anonymization procedures will also be applied if needed, ensuring that individual identities are protected. All data obtained from the various sources will adhere strictly to its stipulated terms and conditions.

5.5 Potential biases of the dataset

Climate and dengue data from the various sources will be scrutinized for imbalances to ensure that there are no potential bises of the dataset in the context of the research question. The research will be conducted with sensitivity to potential demographic disparities within the datasets, ensuring that conclusions are not skewed by demographic or other disparities.

6 References

Dengue cases hit 11,343, surpass last year's figure. The Straits Times. (2016, September 5).

<https://www.straitstimes.com/singapore/environment/dengue-cases-hit-11343-surpass-last-years-figure#main-content>

Tan, A. (2016, December 31). 2016 is hottest year on record for Singapore and the world. The Straits Times.

<https://www.straitstimes.com/singapore/environment/2016-is-hottest-year-on-record-for-singapore-and-the-world>

World Health Organization. (n.d.). Dengue- global situation. World Health Organization. <https://www.who.int/emergencies/diseases-outbreak-news/item/2023-DON498>

Regan, H. (2019, September 19). How climate crisis is accelerating the global spread of Deadly Dengue Fever. CNN.

<https://edition.cnn.com/2019/09/18/health/dengue-asia-outbreak-climate-intl-hnk/index.html>

Gan, N. (2020, July 3). Singapore is on track to face its worst dengue outbreak in history. CNN.

<https://edition.cnn.com/2020/07/03/asia/singapore-dengue-intl-hnk/index.html>

Wikimedia Foundation. (2023, November 30). 2020 dengue outbreak in Singapore. Wikipedia.

https://en.wikipedia.org/wiki/2020_dengue_outbreak_in_Singapore

Impact of climate change in Singapore. National Climate Change Secretariat. (n.d.-a). <https://www.nccs.gov.sg/singapores-climate-action/impact-of-climate-change-in-singapore/>

World Bank Climate Change Knowledge Portal. Climatology | Climate Change Knowledge Portal. (n.d.).
<https://climateknowledgeportal.worldbank.org/country/singapore/climate-data-historical>

Notifications Of Specific Notifiable Diseases. Tablebuilder.singstat.gov.sg. (n.d.). <https://tablebuilder.singstat.gov.sg/table/TS/M870081>

Common Mosquito Breeding Places: other premises. The National Environment Agency. (n.d.). <https://www.nea.gov.sg/docs/default-source/resource/section-b-chapter-6---other-premises-pdf-3-23-mb-.pdf>

Dengue surveillance data, Jan - Mar 2020 - The National Environment Agency. National Environment Agency. (2020, April 27).
[https://www.nea.gov.sg/docs/default-source/default-document-library/q1-2020-dengue-surveillance-data-\(74kb\).pdf](https://www.nea.gov.sg/docs/default-source/default-document-library/q1-2020-dengue-surveillance-data-(74kb).pdf)

Surveillance of dengue virus serotypes 2020 - the national environment ... National Environment Agency. (2020, July 29).
[https://www.nea.gov.sg/docs/default-source/default-document-library/q2-2020-dengue-surveillance-data-\(163kb\)d440a13ca7424ccaa873e93d5c95ba2.pdf](https://www.nea.gov.sg/docs/default-source/default-document-library/q2-2020-dengue-surveillance-data-(163kb)d440a13ca7424ccaa873e93d5c95ba2.pdf)

Surveillance of dengue virus serotypes 2020 - the national environment ... National Environment Agency. (2020, November 2).
[https://www.nea.gov.sg/docs/default-source/default-document-library/q3-2020-dengue-surveillance-data-\(162kb\).pdf](https://www.nea.gov.sg/docs/default-source/default-document-library/q3-2020-dengue-surveillance-data-(162kb).pdf)

Dengue surveillance data - The National Environment Agency. National Environment Agency. (2021, January 28).
[https://www.nea.gov.sg/docs/default-source/default-document-library/q4-2020-dengue-surveillance-data-\(112kb\).pdf](https://www.nea.gov.sg/docs/default-source/default-document-library/q4-2020-dengue-surveillance-data-(112kb).pdf)

User guide and tutorial#. User guide and tutorial - seaborn 0.13.1 documentation. (n.d.). <https://seaborn.pydata.org/tutorial.html>