Module 4:

RENAME THE FILE TO INCLUDE YOUR COMPANY, GROUP NUMBER, AND LAST NAMES

E.G. KAMEN1_GROVES_MODULE_4.IPYNB

Team Members:

Andrea and Grace

Project Title:

SIR Modeling of Measles Transmission in Nigeria Using Reported Case Data

Project Goal:

The goal of this project is to convert cumulative measles case data into estimates of the susceptible, infected, and recovered populations, and use these to visualize outbreak dynamics in Nigeria.

Disease Background:

Using your assigned disease, fill in the following bullet points.

MEASLES

- Prevalence & incidence Measles is still very common around the world, especially in places where fewer people get vaccines. The World Health Organization reports about 9 million cases and over 100,000 deaths every year, mostly in young children. In the United States, measles used to infect 3–4 million people each year before the vaccine existed. Today, it is much less common because most people are vaccinated, but small outbreaks still happen every year, usually linked to unvaccinated people or to travelers who bring the virus back from countries where measles is widespread.
- Economic burden Measles is expensive for families, hospitals, and public health systems. Outbreaks in the United States can cost millions of dollars because they require contact tracing, testing, vaccinations, and sometimes school or workplace interventions. For families, hospital stays and treatment for complications (like pneumonia or dehydration) can cost thousands of dollars. In lower-income

countries, the financial impact is even worse because families may lose income while caring for sick children, and healthcare systems often struggle with limited resources.

- Risk factors (genetic, lifestyle) & Societal determinants The biggest risk factor for measles is not being vaccinated. Babies are also at risk because they are too young to receive the vaccine. People with weak immune systems (like those with HIV or cancer) are more likely to have severe disease. Social factors also affect measles risk. Crowded living conditions, limited access to healthcare, poor nutrition (especially vitamin A deficiency), and misinformation or fear about vaccines can all increase the chance of outbreaks. Travel to areas where vaccination rates are low also raises exposure risk.

- Symptoms Measles often starts like a bad cold: high fever, cough, runny nose, and red, watery eyes. A few days later, tiny white spots called Koplik spots may appear inside the mouth. Then, a red, blotchy rash develops, starting on the face and spreading to the rest of the body. The rash usually lasts about a week. Measles can cause serious complications, such as ear infections, pneumonia, diarrhea, or even brain swelling (encephalitis). In rare cases, a long-term brain disease called SSPE can appear years after a person recovers. These complications are most common in young children and people with weak immune systems.

- Diagnosis Doctors diagnose measles by looking at a person's symptoms, especially the combination of fever, rash, and cold-like signs. Koplik spots inside the mouth are a strong clue. To confirm the diagnosis, healthcare providers usually order a blood test to check for measles antibodies (IgM) or use PCR testing to detect the virus in a throat swab or urine sample. Laboratory confirmation is important because measles is rare in many countries due to vaccination, and public health officials need to take quick action when a suspected case appears.

- Biological mechanisms (anatomy, organ physiology, cell & molecular physiology) Measles is caused by a very contagious virus that spreads through the air when an infected person breathes, coughs, or sneezes. The virus enters the body through the nose, mouth, or lungs, infecting immune cells there and then spreading to the bloodstream. Measles weakens the immune system by destroying important immune cells, which can make someone vulnerable to other infections for weeks or even months. The rash happens because the immune system targets infected cells in the small blood vessels of the skin. The virus can affect many organs, leading to lung infections (pneumonia), brain swelling (encephalitis), and severe diarrhea, which explains why the disease can be dangerous.

Sources: https://www.cdc.gov/measles/ https://www.who.int/news-room/fact-sheets/detail/measles

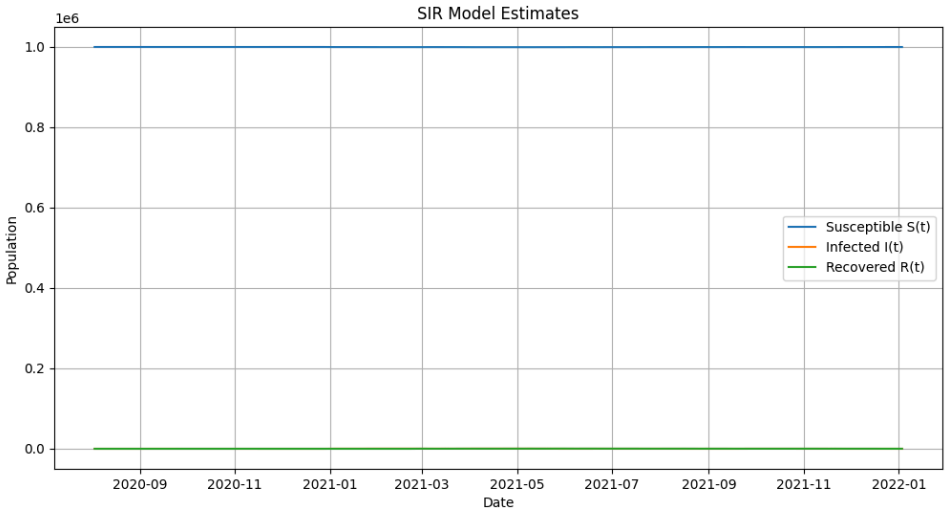
Dataset:

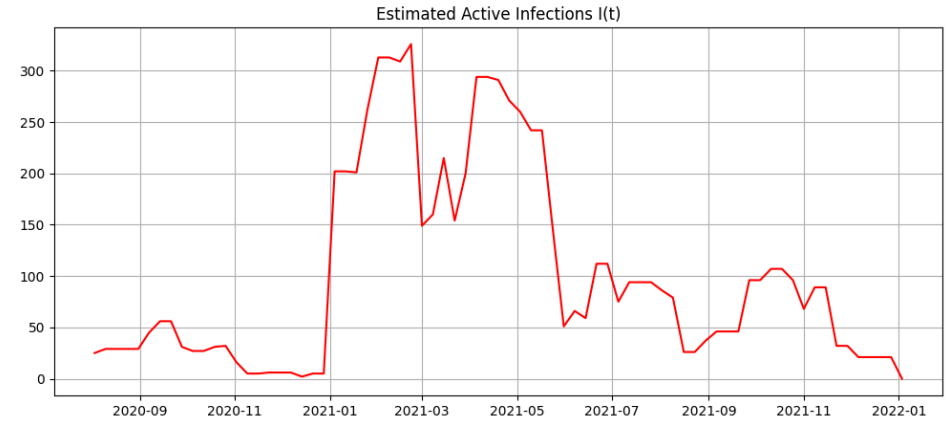
(Describe the data set you will analyze. Cite the source(s) of the data. Describe how the data was collected -- What techniques were used? What units are the data measured in? Etc.)

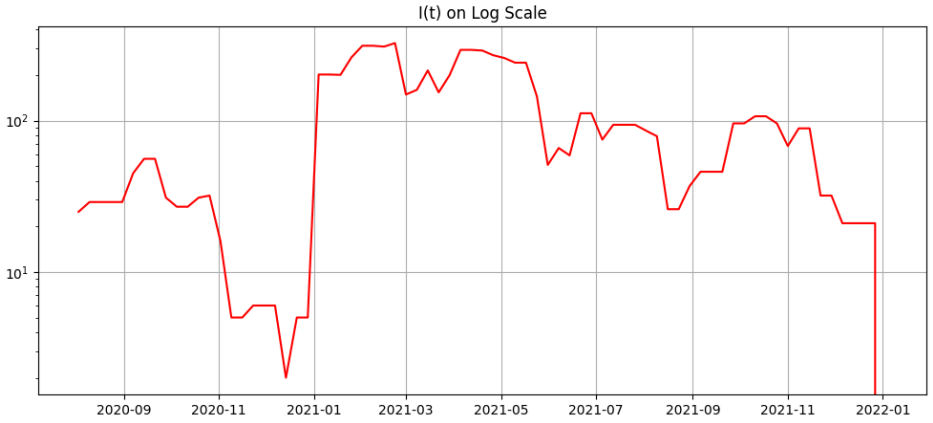
In [ ]:
```
## LOAD YOUR DATASET HERE.

# 1. Read in the csv file of cumulative cases.

# 2. Use the convert_cumulative_to_SIR function to convert cumulative cases

# 3. Plot S, I, R over time.
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


# Your provided conversion function

def convert_cumulative_to_SIR(df, date_col='date', cumulativ
                              population=None, infectious_period=8, recovere
                              new_case_col='new_cases', I_col='I_est', R_col

    df = df.copy()

    # Ensure date column sorted
    if date_col in df.columns:
        df[date_col] = pd.to_datetime(df[date_col])
        df = df.sort_values(date_col).reset_index(drop=True)

    if cumulative_col not in df.columns:
        raise ValueError(f"Column '{cumulative_col}' not found in dataframe.

    # Compute new cases
    df[new_case_col] = df[cumulative_col].diff().fillna(df[cumulative_col].i
    df[new_case_col] = df[new_case_col].clip(lower=0)

    # Estimate I(t)
    df[I_col] = df[new_case_col].rolling(window=infectious_period, min_perio

    # Estimate R(t)
    df[R_col] = df[cumulative_col].shift(infectious_period).fillna(0)

    # Compute S(t)
    if population is not None:
        df[S_col] = population - df[I_col] - df[R_col]
        df[S_col] = df[S_col].clip(lower=0)
    else:
        df[S_col] = np.nan

    return df
```

```
# 1. LOAD YOUR DATASET


df = pd.read_csv("measles_nigeria_data_2020-2021_new_cases.csv")

print("Columns in file:", df.columns.tolist())

# Your columns are:
# date
# confirmed_cases
date_col = "date"
cumulative_col = "confirmed_cases"   # <--- IMPORTANT FIX



# 2. CONVERT CUMULATIVE → S, I, R


population = 1_000_000   # Replace with real value if you know it

df_sir = convert_cumulative_to_SIR(
    df,
    date_col=date_col,
    cumulative_col=cumulative_col,
    population=population,
    infectious_period=8,
    new_case_col="new_cases",
    I_col="I_est",
    R_col="R_est",
    S_col="S_est"
)

print(df_sir.head())



# 3. PLOT S, I, R


plt.figure(figsize=(12,6))
plt.plot(df_sir[date_col], df_sir["S_est"], label="Susceptible S(t)")
plt.plot(df_sir[date_col], df_sir["I_est"], label="Infected I(t)")
plt.plot(df_sir[date_col], df_sir["R_est"], label="Recovered R(t)")

plt.xlabel("Date")
plt.ylabel("Population")
plt.title("SIR Model Estimates")
plt.legend()
plt.grid(True)
plt.show()


# Optional: I(t) only
plt.figure(figsize=(12,5))
plt.plot(df_sir[date_col], df_sir["I_est"], color="red")
```

```
plt.title("Estimated Active Infections I(t)")
plt.grid(True)
plt.show()


# Optional: I(t) log scale
plt.figure(figsize=(12,5))
plt.plot(df_sir[date_col], df_sir["I_est"], color="red")
plt.yscale("log")
plt.title("I(t) on Log Scale")
plt.grid(True)
plt.show()
```







In [ ]:

Data Analyis:

Methods

IN A SUMMARY, DESCRIBE THE METHODS YOU USED TO ANALYZE AND MODEL THE DATA.

Analysis

(Describe how you analyzed the data. This is where you should intersperse your Python code so that anyone reading this can run your code to perform the analysis that you did, generate your figures, etc.)

In [2]:
```
## PYTHON CODE TO BUILD AND FIT AN SIR MODEL GOES HERE. INTERSPERSE COMMENTS
```

Verify and validate your analysis:

(Describe how you checked to see that your analysis gave you an answer that you believe (verify). Describe how your determined if your analysis gave you an answer that is supported by other evidence (e.g., a published paper).

Conclusions and Ethical Implications:

(Think about the answer your analysis generated, draw conclusions related to your overarching question, and discuss the ethical implications of your conclusions.

Limitations and Future Work:

(Think about the answer your analysis generated, draw conclusions related to your overarching question, and discuss the ethical implications of your conclusions.

NOTES FROM YOUR TEAM:

We don't have notes yet.

In [ ]:

In [ ]:

QUESTIONS FOR YOUR TA:

We don't have questions yet for our TA.

In [ ]: