Infectious Disease Modeling in the Time of COVID

October 8, 2020

Presenters:
Flavia Camponovo, PhD - Postdoctoral Research Fellow
Caroline Buckee, PhD - Associate Professor of Epidemiology

Learning objectives:

- Understand the concepts of mathematical modelling of infectious diseases, including importance and source of parameters
- Utilize epidemiological data to inform infectious disease modeling parameters
- Understand how dynamics of COVID-19 in the US relate to the biological and behavioral processes that spread the virus

Recap of Prof. Buckee's Presentation: Compartmental models in epidemiology

S-I-R models are compartmental models used to define simplified infectious disease dynamics. In these models, the population is divided between **S**usceptibles, **I**nfected, and **R**ecovered, and transition from one compartment to the other based on parameters which define the transmission dynamics. There are many variations of this model. One variation of the compartment model we will discuss in the exercise is the S-E-I-R model, where a susceptible individual is assumed to be infected but not yet infectious for some time (***E**xposed),* before becoming infectious. This period between exposed and infectious is called the **latent period**.

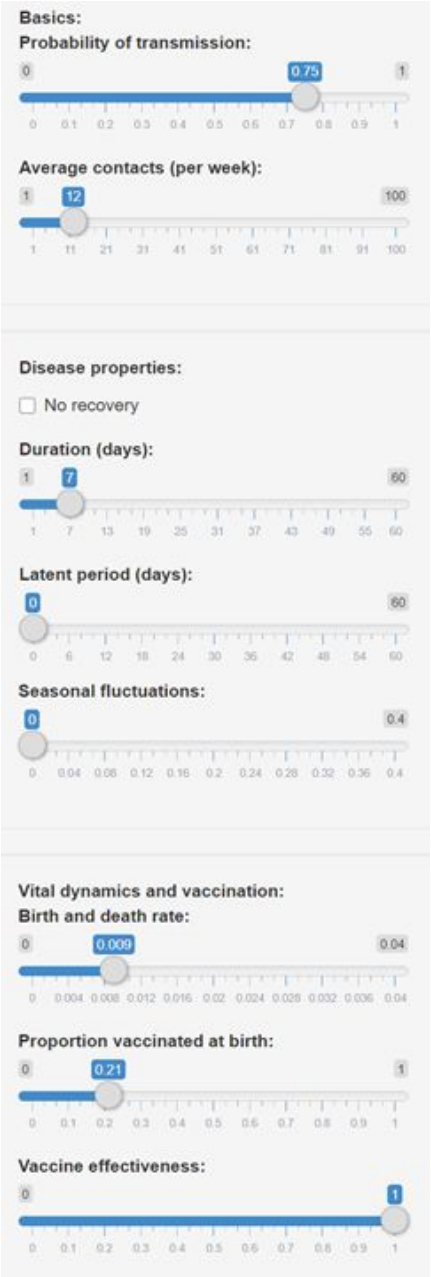
Summary of the compartments:

S: Number of **S**usceptible individuals. If in contact with an infectious person, a susceptible becomes either *exposed* or *infected* depending on the model, and transitions into the corresponding compartment.
E: Number of **E**xposed individuals. Exposed individuals have been infected but are not yet infectious. Once infectious, they transition into the *Infectious* compartment.
I: Number of **I**nfectious individuals. Infectious individuals can infect the susceptible population.
R: Number of **R**ecovered individuals. These individuals have recovered from the infection, and cannot be infected again (they are immune).

Note: The resource shared in today's activity has many parameters and much more information than we will be able to address with you in the short time we have. Questions in *italics* will not be covered today, but have been included in case you would like to further your understanding and learn more.

Use the SIR module developed by Mathew Kiang, available at https://mkiang.shinyapps.io/DiseaseDynamics/ to model infectious diseases.

The parameters on the page are described below—note that certain parameters are only shown when they are applicable (Eg. “Proportion vaccinated at birth” will only be shown if the “Birth and death rate” are adjusted to over 0).



Basics

- **Probability of transmission:** the probability that an infectious person will infect a susceptible person at any one contact.
- **Average contacts (per week):** the number of people an infected person will run into.

Disease properties

- **No recovery:** when checked, if somebody becomes infected, they will be infectious forever—never recovering.
- **Duration (days):** how long an infected person remains infected. This determines 1) how many days a person can infect susceptible hosts and 2) how long it takes before recovering.
- **Latent period (days):** the time between being infected and being infectious (at which point they are neither susceptible nor infected, but are “exposed”).
- **Seasonal fluctuations:** a function that emulates seasonal fluctuations in contact rates.

Vital dynamics and vaccination

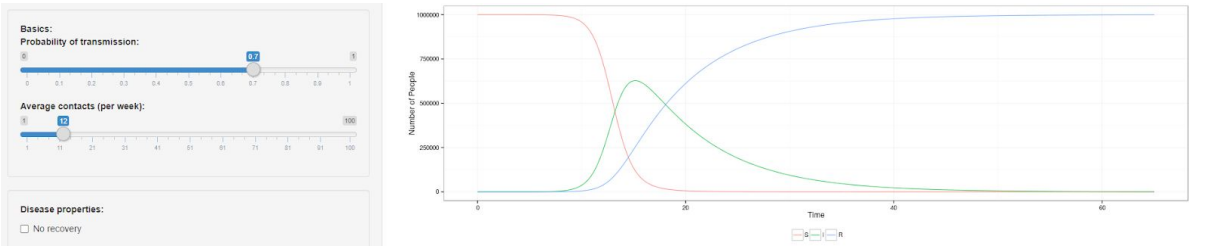
- **Birth and death rate:** in this model, kept equal to each other at 0. Adjusting it will make below options appear
- **Proportion vaccinated at birth**: assumes vaccination occurs immediately at birth (therefore, this is the proportion of new births who never enter a susceptible stage).
- **Vaccine effectiveness:** probability of a vaccine actually working.

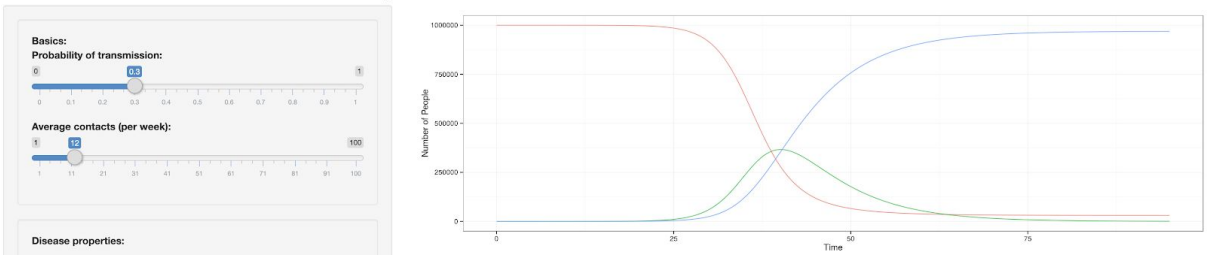
Copied from: https://mathewkiang.com/2013/12/20/shiny-desolve-interactive-ode-models/

When in breakout rooms, pick one person to share their screen and manipulate the website parameters. Other group members will read the instructions and provide guidance. You have 15 minutes to explore **Part I A and B** in your breakout room. Please be mindful of the time. **After Part I A and B, we will return to the main room for discussion before proceeding to Part II. Select a group representative to share your results.**

<u>Part I - A</u> Understanding the general behavior of the model:

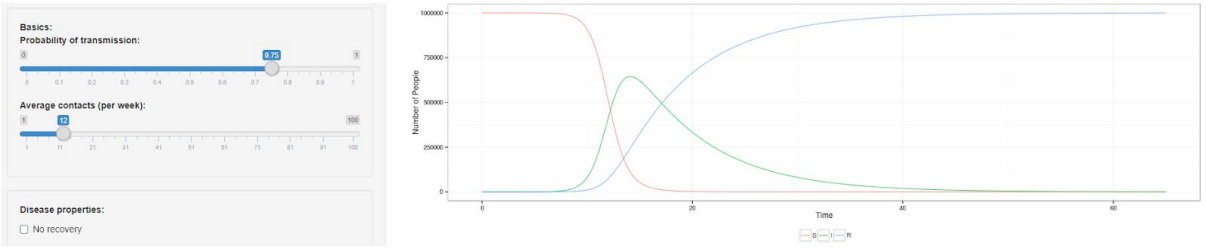
1. Start the activity without adjusting any of the parameters. This is a S*usceptible-Infected-Recovered (S-I-R)* model, where infected people recover after some time and become immune.
    a. What is the impact of the probability of transmission on the disease dynamics outcome?
    The probability of transmission indicates how likely an infected person will infect a susceptible person upon contact. A disease with a lower probability of transmission will take more time to spread across a population. This can be observed in the plots by checking how long it takes for the number of infected individuals (green curve) to peak, as illustrated by the two plots below where it takes approximately 15 days for the peak when the probability of transmission is 0.75 (upper plot) versus 40 days if the probability of transmission is reduced to 0.3 (lower plot). In addition, with a lower probability of transmission, the peak is also lower, meaning that there are less people infected simultaneously in a given population. This has the additional advantage of avoiding overwhelming health care centers and workers.
    You can also notice that for a low transmission probability, the outbreak ends (no more infected people) before all the susceptible population has been exposed (the susceptible curve never reaches 0).

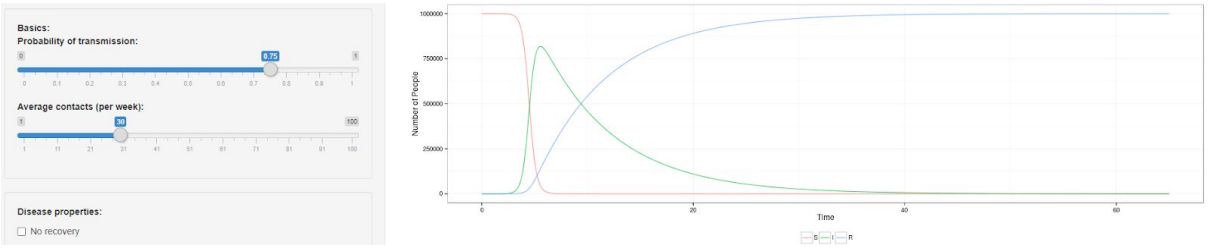




b. What is the impact of the average number of contacts on the disease dynamics outcome?

The average contacts per week determines the opportunities for the disease to spread from an infected person to a susceptible host. The higher the average contact rate, the faster the disease can spread across the population. This can be observed in the plot by checking how long it takes for the number of infected (green curve) to peak, as illustrated by the two plots below where it takes approximately 15 days for the number of infected to peak when considering 12 contacts per week (upper plot) versus 5 days if considering 30 contacts per week (lower plot). Similarly to decreasing the transmission probability, if the average contacts are decreased sufficiently, the infection can end before the entire population becomes infected (in this example, you can try, for example, probability of transmission = 0.75 and average contact = 6 or lower).

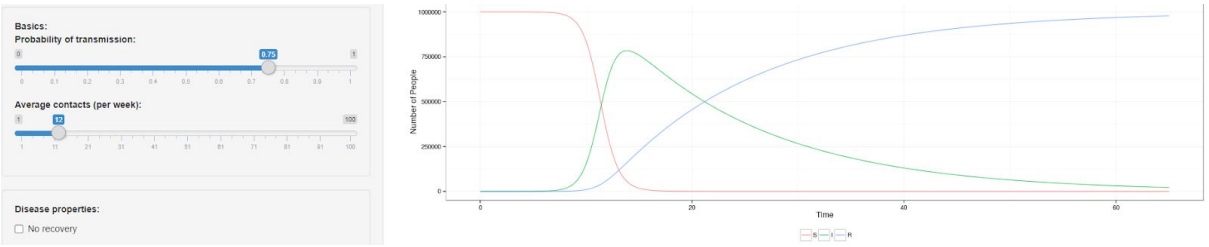




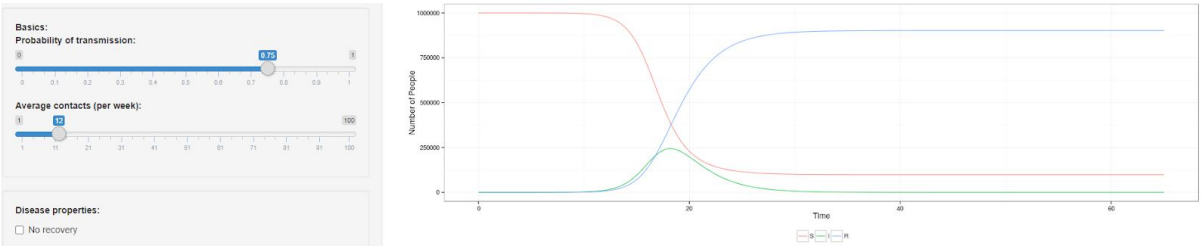
c. What is the impact of the duration of the infection on the disease dynamics outcome?

The duration of the infection determines *1)* how many days a person can infect susceptible hosts and *2)* how long it takes before recovering. Thus, a disease with a long infection period can result in a high proportion of the population staying infected at the same time, and a longer time to obtain a fully recovered population compared to a disease with shorter infection period. This is illustrated with the plots below, where the *Infected* peak is higher and wider when duration is set to 14 days (upper plot) compared to 2 days (lower plot), and the time for the *Recovered* population to reach all people is more than 60 days versus 25 days. Here again, a disease with a short infectious period is more likely to disappear before the entire population becomes infected (ie Susceptibles people remain after the end of the outbreak)

Probability of transmission 0.75; Average contact 12; Duration 14 days:

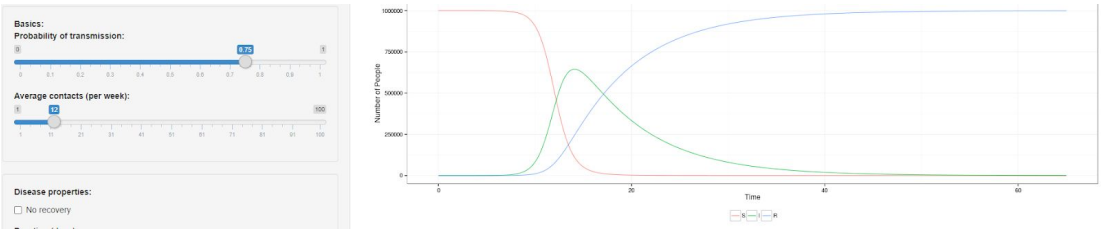


Probability of transmission 0.75; Average contact 12; Duration 2 days:

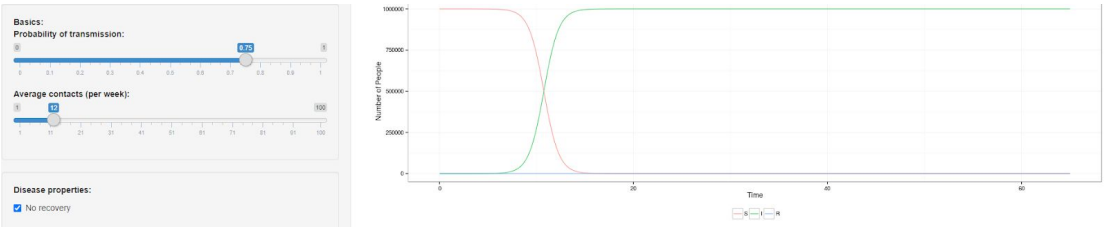


1. ***Go to Part I - B on the next page first. If time allows, feel free to come back and dig deeper with these additional models:***
    a. *What changes when assuming no recovery? Check the "No recovery" box in the Disease Properties section. This is a Susceptible-Infected (SI) model.*
    This model assumes that once infected, a person remains infected, thus there is no "Recovered" and this becomes an SI model. Because there is no recovery and once infected a person remains infectious for ever, the number of Infected people (green line) rises until the entire population is infected and decreasing the transmission probability or contact rate will only slow down the spread but will never avoid the entire susceptible population to become infected.

Unchecked



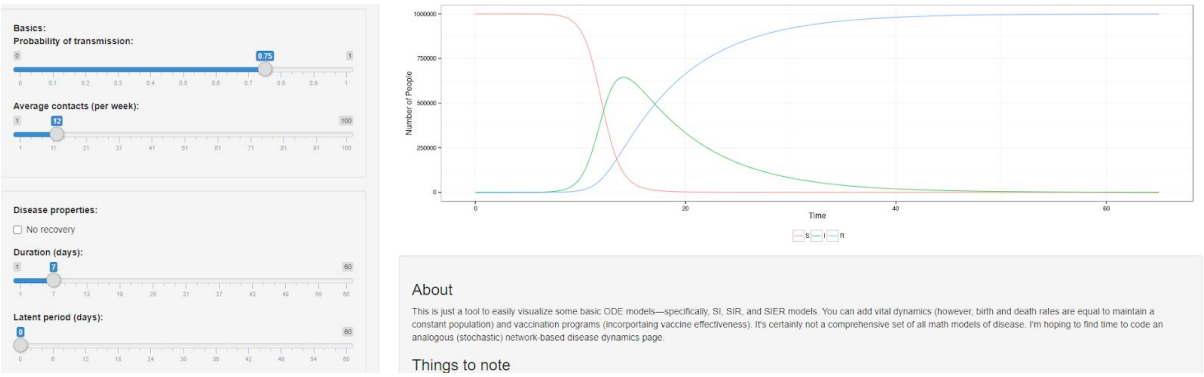
Checked



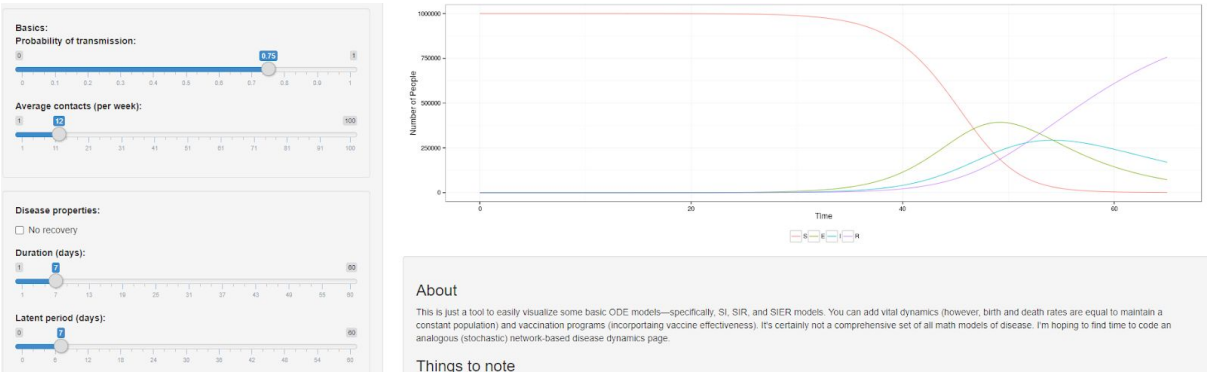
b. *What changes by adding a Latent period? (Make sure "No recovery" is unchecked). This is a Susceptible-Exposed-Infected-Recovered (S-E-I-R) model, with the **E**xposed population defined as the fraction of people who got infected, but are not yet infectious.*

A disease with a latent period implies that people who are infected remain non infectious for some time (latent period) before they can transmit the disease. This results in a delay of the peak outbreak, and the outbreak (the *Infected* curve) becomes delayed and "flattened". This is illustrated by the plots below where both models assume a 7 days infectious period but the second model includes a 7 days latent period.

Probability of transmission 0.75; Average contact 12; Duration 7 days; no Latent period:



Probability of transmission 0.75; Average contact 12; Duration 7 days; Latent period 7 days:



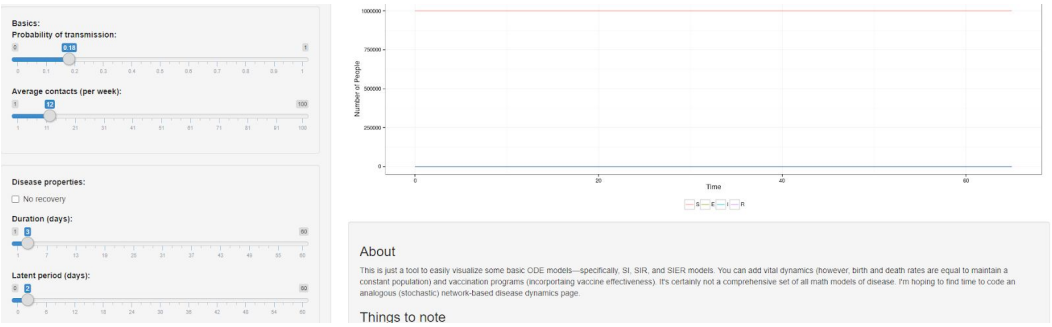
<u>Part I - B</u> Example with specific disease parameters:

1. Simulate two disease-like outbreaks. What parameters do you have to change? Describe the impact of changing each of these parameters on the outbreak trajectory.

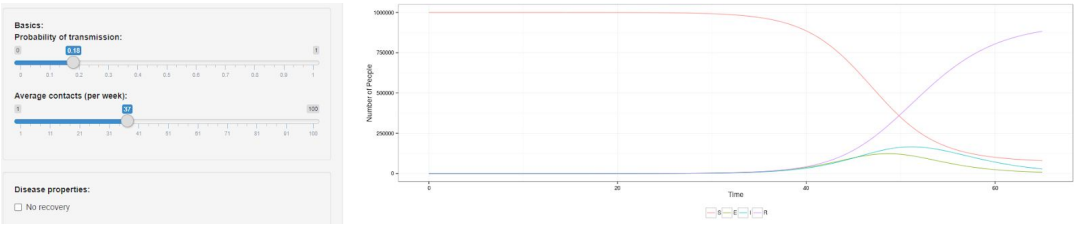
	Probability of transmission	Duration	Latent period
Influenza	0.18	3	2
COVID-19	0.5	14	2

Generally, increasing the following parameters results in:
Probability of transmission - Faster spread within the population. All peaks shift to the left
Duration - Leveling effect on number of infected after it peaks. The infected remain infected for a longer period and cannot transition to recovered. Thus, infected curve flattens out while the recovered curve decreases. For influenza parameters, shifts the curves to the left.
Latent period - Shifts all curves to the right into the future. A longer latent period means a longer time period when people are infected but not infectious, so they cannot spread the disease and the effects are all delayed from the original model.

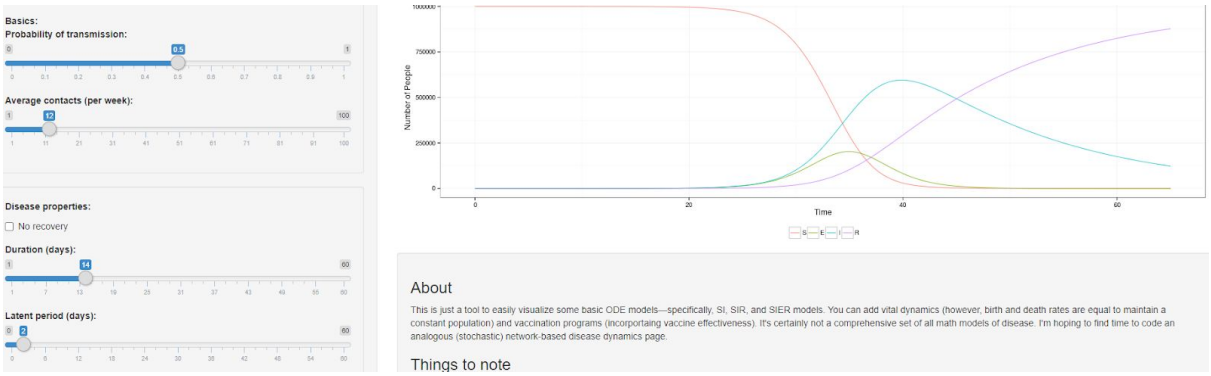
Influenza



Even after 65 days, the number of infected is stable and has not noticeably increased. Adjusting the other parameters such as increasing Average contacts to above 30 will change the output to increase the number of infected.

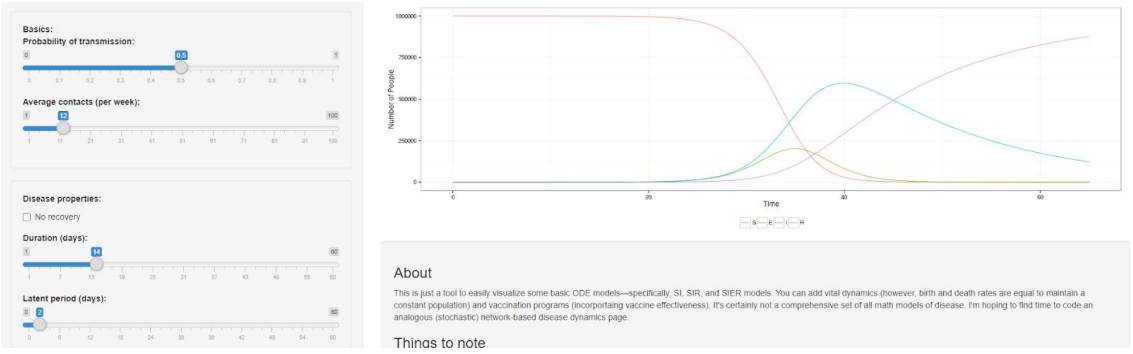


COVID-19

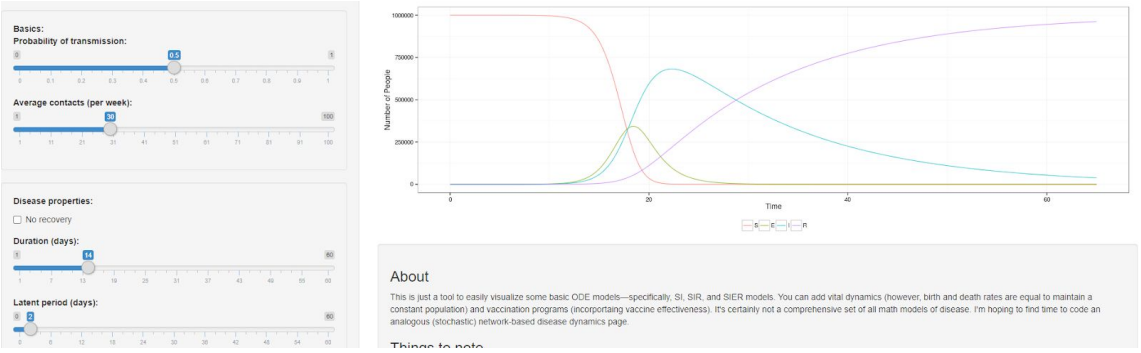


2. One measure taken during the Covid-19 pandemic is physical distancing, or lockdown, both of which aim to reduce the average number of contacts. Let's take the COVID-19-like parameters, and change the average contact rate to lower or higher values. What happens if you change the average contacts? By reducing the average contact rate, the peak of the outbreak is delayed, and the curve is flattened. From a health system's perspective, this has the advantage of gaining some time to prepare and to limit the overwhelming of the hospital by not having too many cases at once.

   Probability of transmission 0.5; Average contact 12; Duration 14 days; Latent period 2 days:



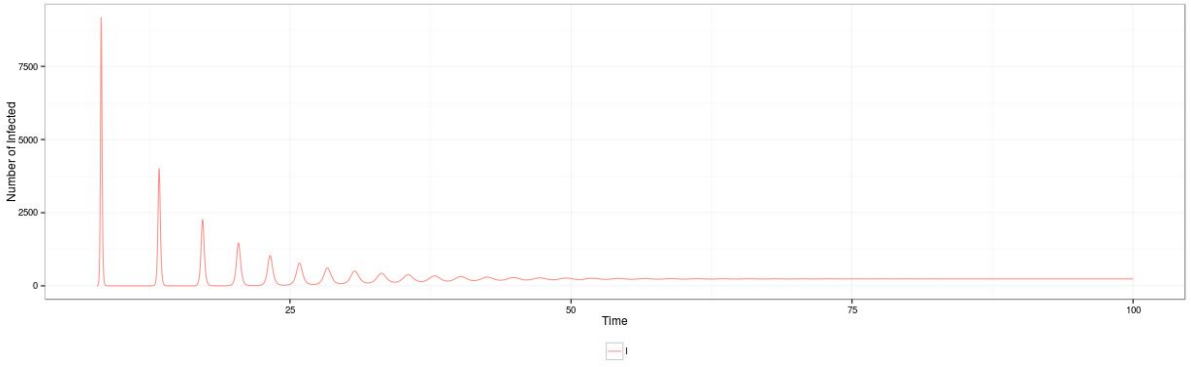
Probability of transmission 0.5; Average contact 30; Duration 14 days; Latent period 2 days:



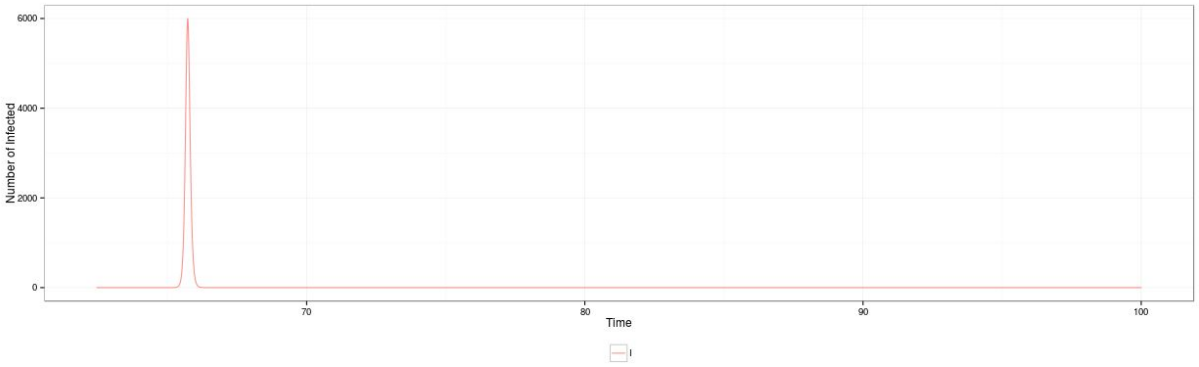
3. ***If time allows, feel free to consider this case study on measles:***
In the case of measles, an effective prevention measure is vaccination of young children. Let's take the measles-like parameters and an average contact rate of 25, and look at the effect of vaccination. Set the Birth and death rate to 0.013. Adding birth and death rates result in a dynamic population, with newborns entering the population as Susceptibles. The proportion of vaccinated children then determines how many of the newborns are immunized and thus protected, or remain susceptible to the infection. Change the Time scale to Years and increase the Time max to 100. Check both boxes in the Plot settings to show the infectious curve only, and the infections past initial outbreak.You will see that without any vaccination (Proportion vaccinated at birth = 0) over the years there are recurrent outbreaks and the population is never disease-free (ie the curve for the infected population never reaches 0). Given the parameters we defined in this model, can you tell what should be the minimum proportion of vaccinated children at birth to avoid any future outbreaks after the initial peak?

	Probability of transmission	Duration	Latent period
Measles	0.9	7	7

Probability of transmission: 0.9
Average contacts: 25
Duration: 7
Latent period 7
Birth and death rate 0.013
Proportion vaccinated at birth: 0
Time max: 100 years



By adjusting the Proportion vaccinated at birth parameter, the model indicates that in this setting, you need to have more than 85% of the children vaccinated at birth in order to eradicate the disease. This shows the importance of a high coverage of immunized children, especially for a disease like measles with a very high transmission probability.



After Part I A and B, we will return to the main room for discussion before proceeding to Part II. Select a group representative to share your results.

When in breakout rooms, pick one person to share their screen and generate the graphs. You have 10 minutes to explore **Part II A and B** in your breakout room. Please be mindful of the time. **After Part II A and B, we will return to the main room for discussion. Select a group representative to share your results.**

Part II - A Graphing real Covid-19 data

Download files for Part II from the Life Sciences Outreach Program website at https://lifesciencesoutreach.fas.harvard.edu/lecture-1-100820

Covid-19 cases and deaths data in files:

covid19-statistics_MAcounties.xlsx	Massachusetts counties[1]
covid19-statistics_USstates.xlsx	US states and territories[1]
covid19-statistics_USA.xlsx	US as a whole[1]
covid19-statistics_China.xlsx	China[2]
covid19-statistics_Italy.xlsx	Italy[2]
covid19-statistics_Southkorea.xlsx	South Korea[2]

Feel free to graph the variables in whatever way is most convenient for you. If you would like graphing assistance, we have provided instructions:

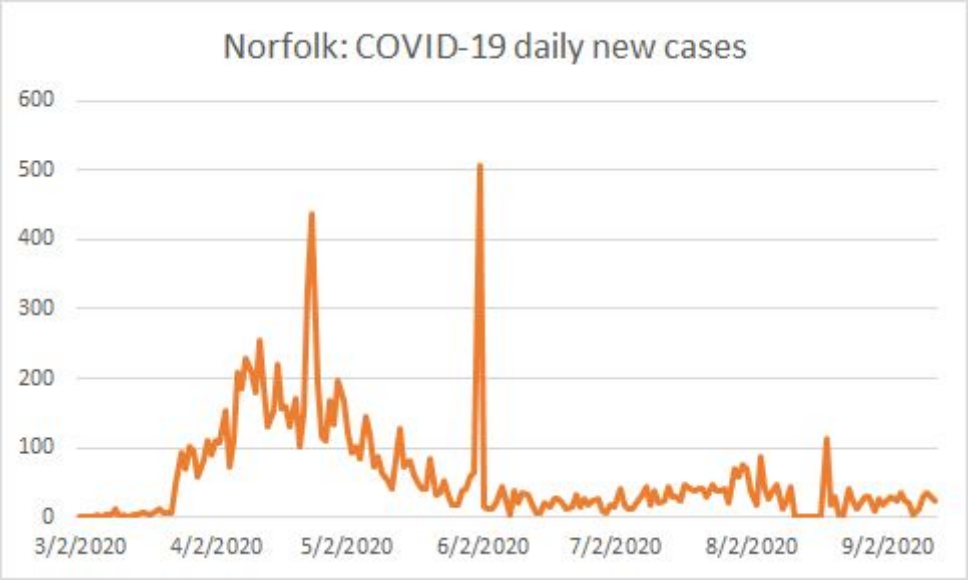
- How to make graphs using Microsoft Excel
- How to make graphs using Google Sheets

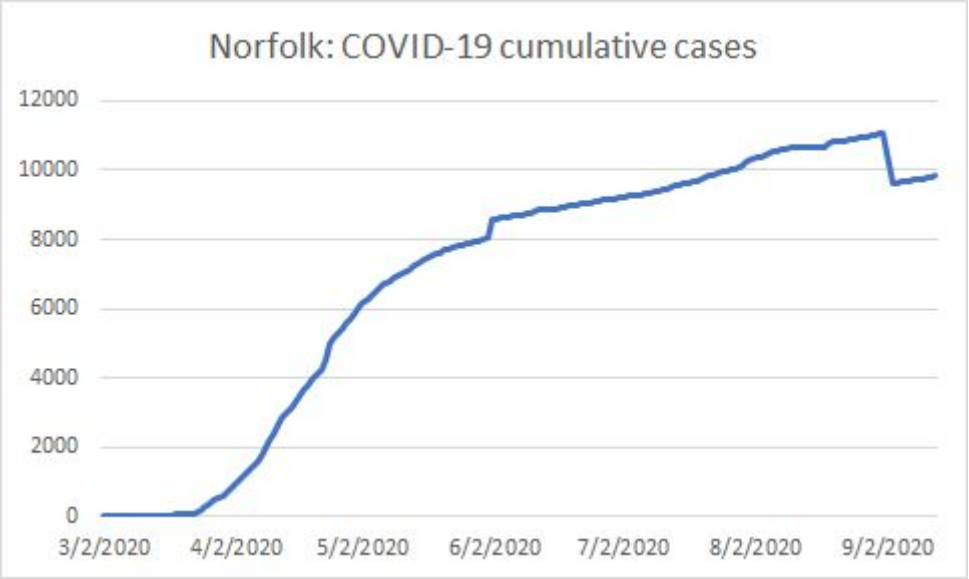
You may take screenshots of the graphs to share and paste in your worksheet below.

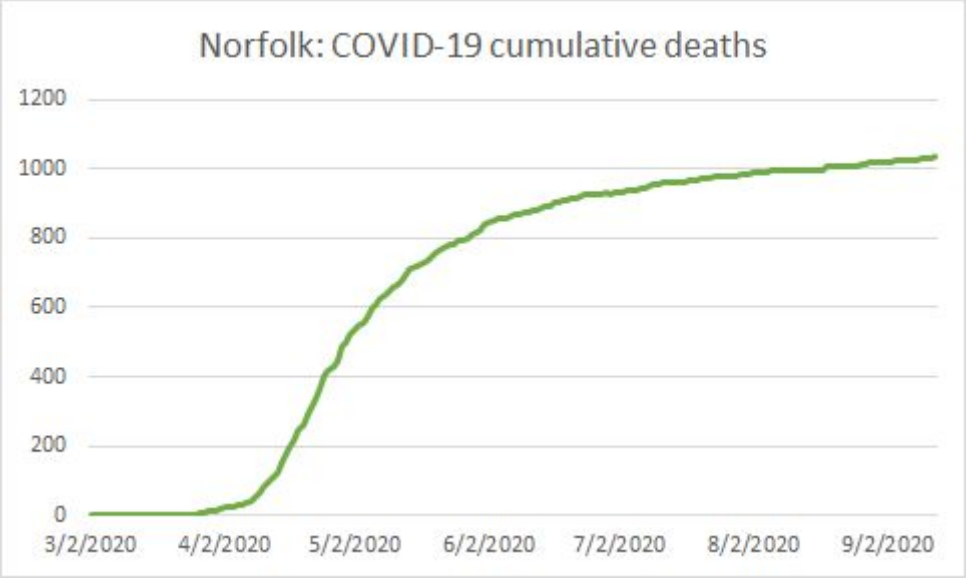
[1] downloaded from https://github.com/nytimes/covid-19-data
[2] downloaded form ECDC https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

Plot COVID-19 new cases, cumulative cases, and deaths counts on three separate graphs for two geographic regions as follows:
Breakout room 1: Norfolk county and Suffolk county
Breakout room 2: Massachusetts and Florida
Breakout room 3: USA and China
Breakout room 4: USA and Italy
Breakout room 5: USA and South Korea
Breakout room 6: Massachusetts and USA

1. Norfolk county, Massachusetts

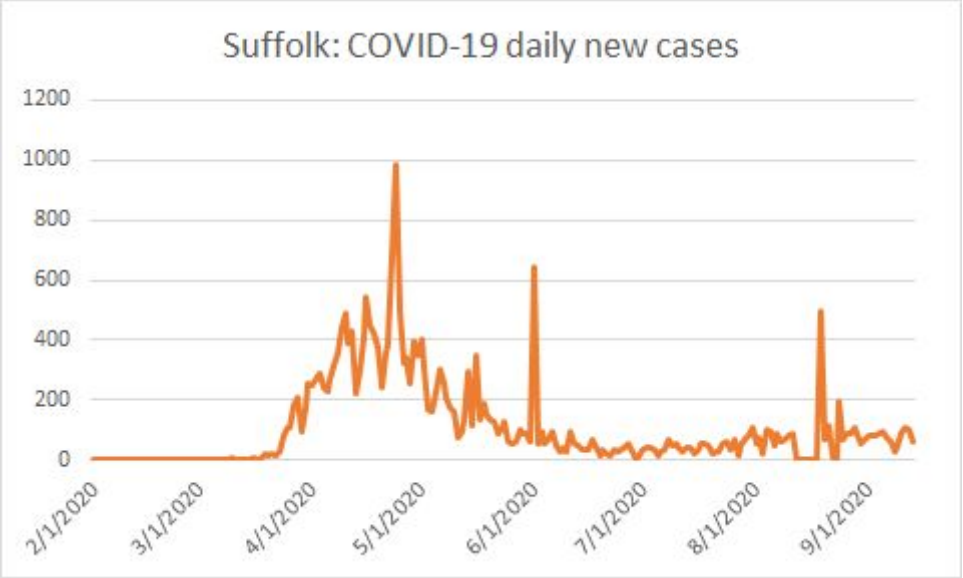


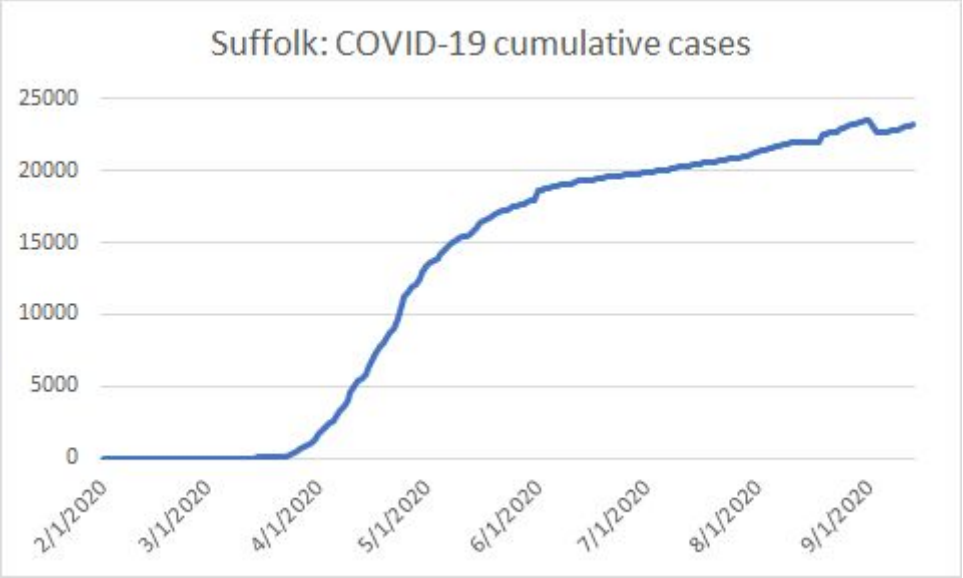


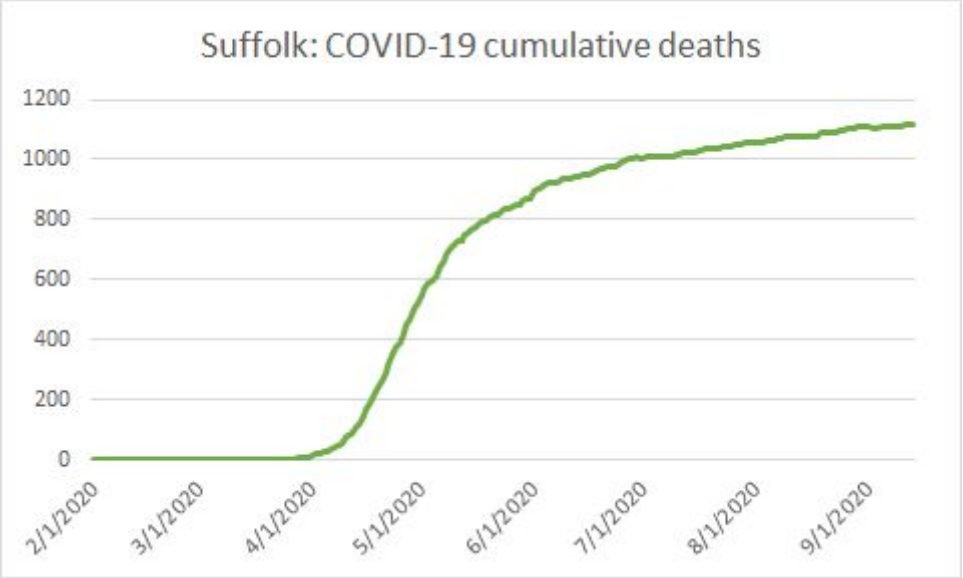


Cities in Suffolk county: Boston, Chelsea, Revere, Winthrop

2. Suffolk county, Massachusetts





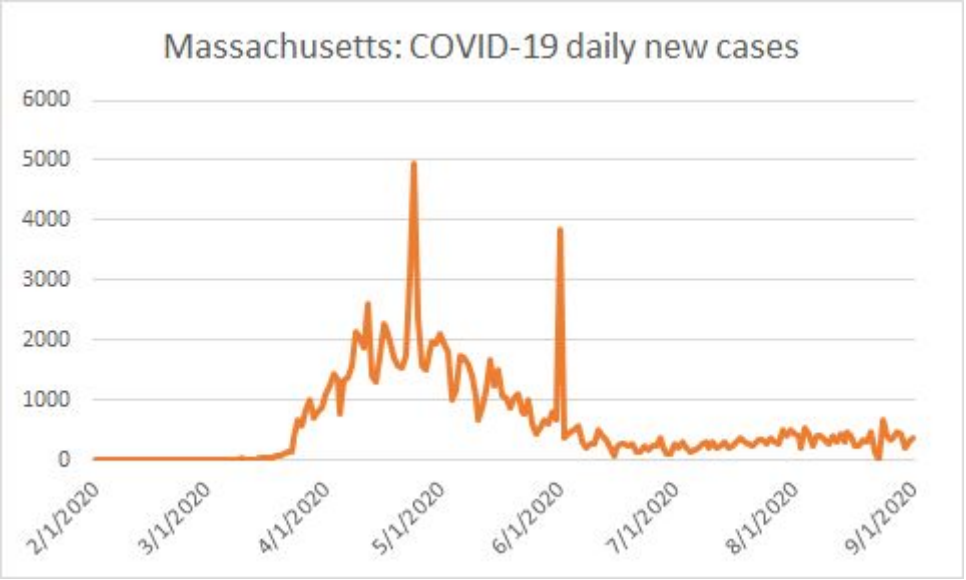


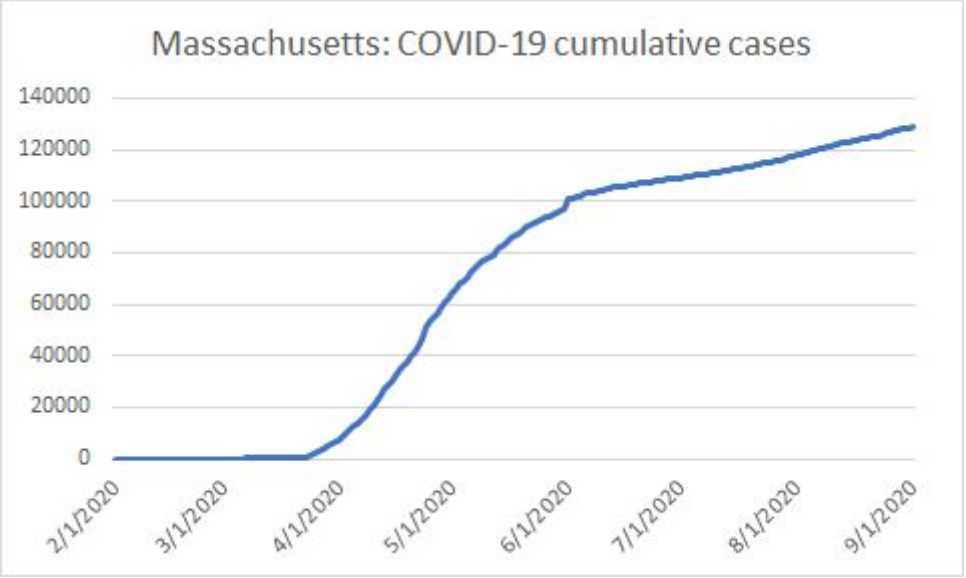
Cities in Norfolk country: Braintree, Franklin, Quincy, Randolph, Weymouth

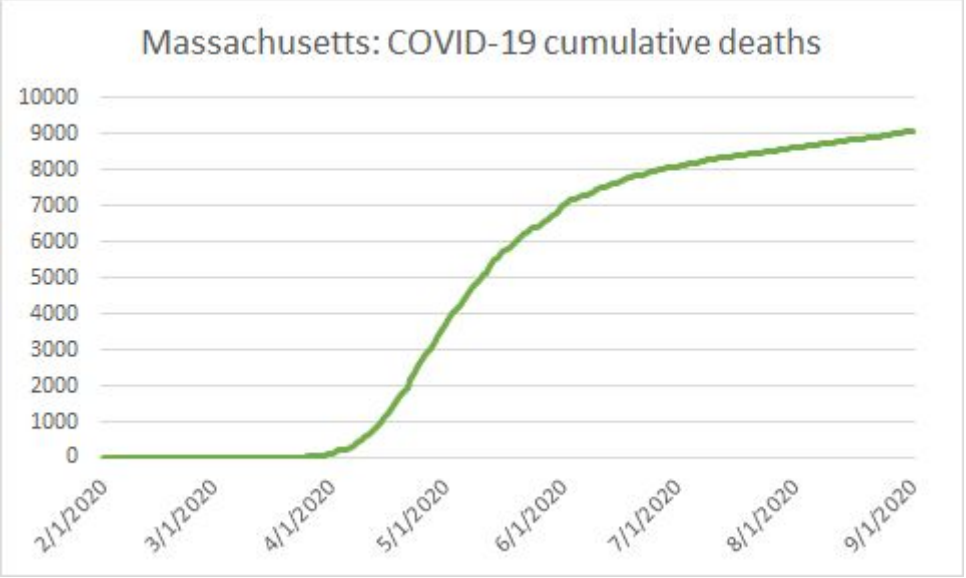
Interesting table

State Rank (per capita income)	US Rank (per capita income)	County	Per capita income	Median family income	Population	Number of households	Density of people per square mile
1	25	Norfolk	$46,920	$108,943	677,296	257,451	401
7	193	Suffolk	$32,835	$61,449	735,701	288,240	1637

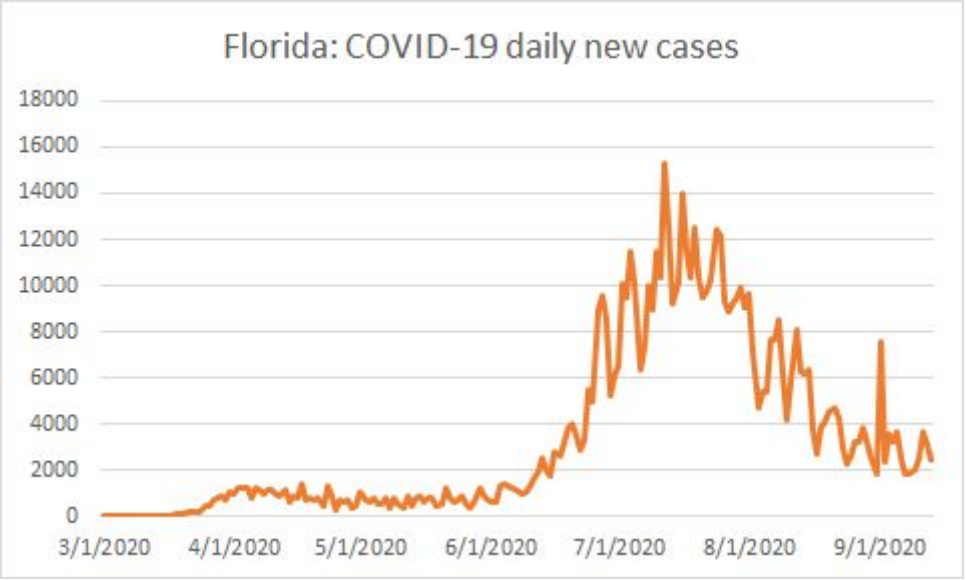
3. Massachusetts

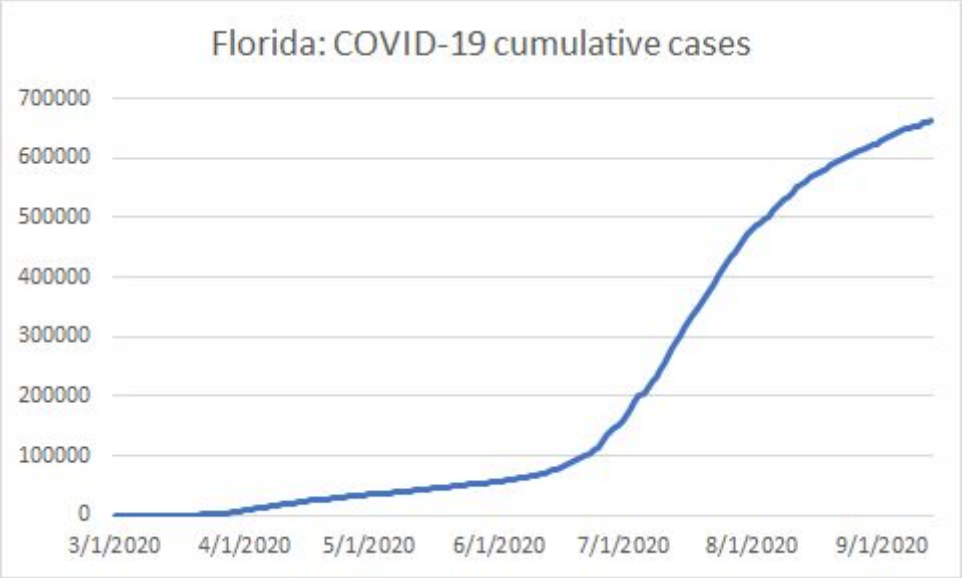


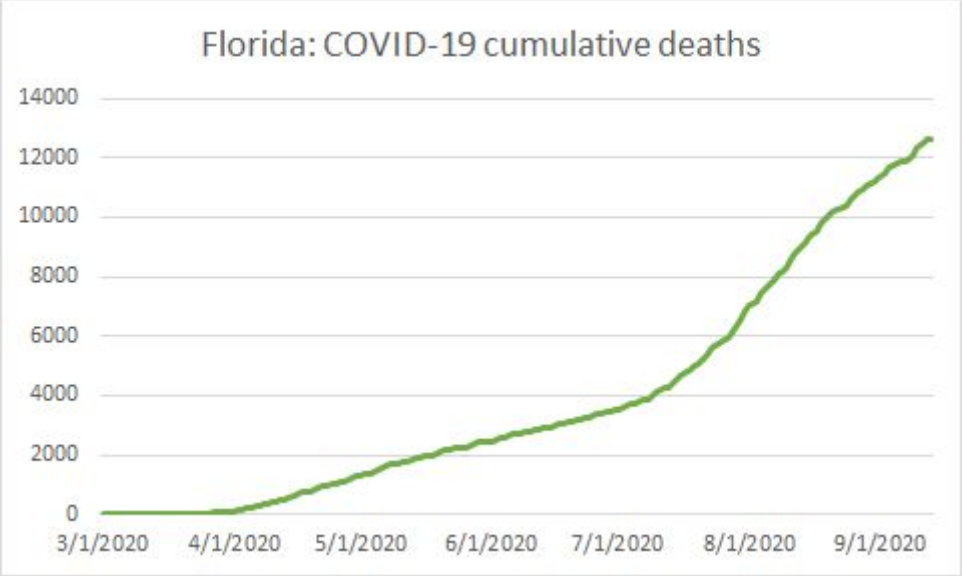




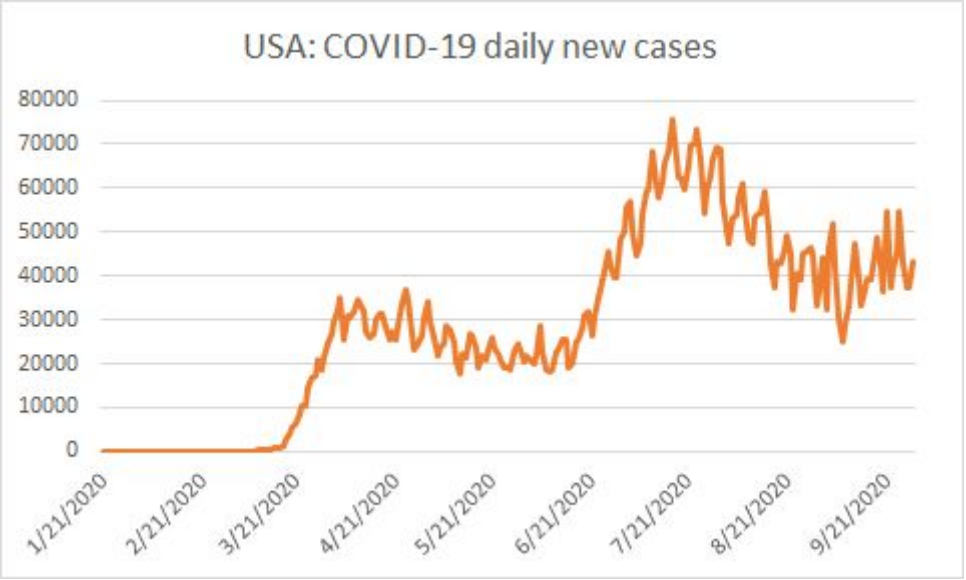
4. Florida

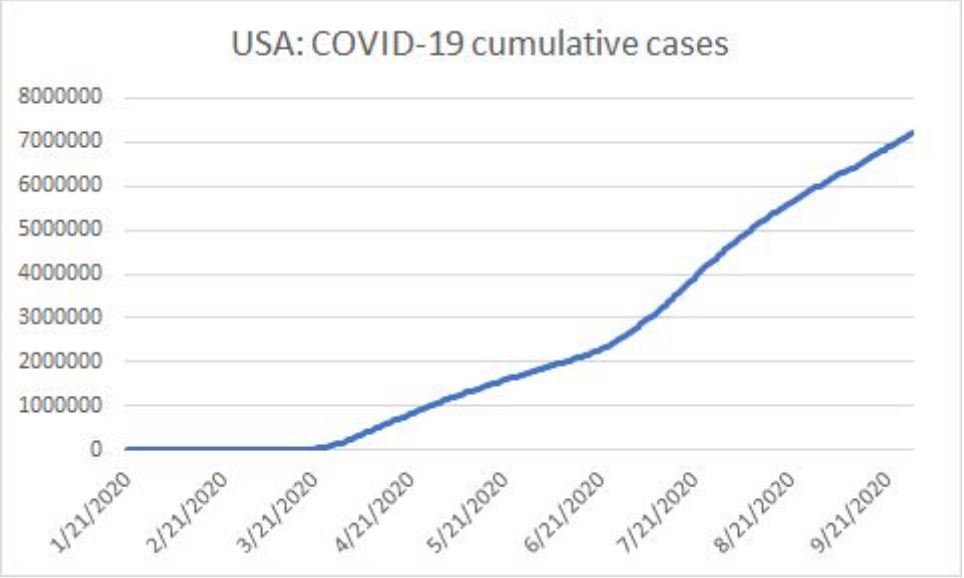

Florida: COVID-19 daily new cases
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
3/1/2020
4/1/2020
5/1/2020
6/1/2020
7/1/2020
8/1/2020
9/1/2020

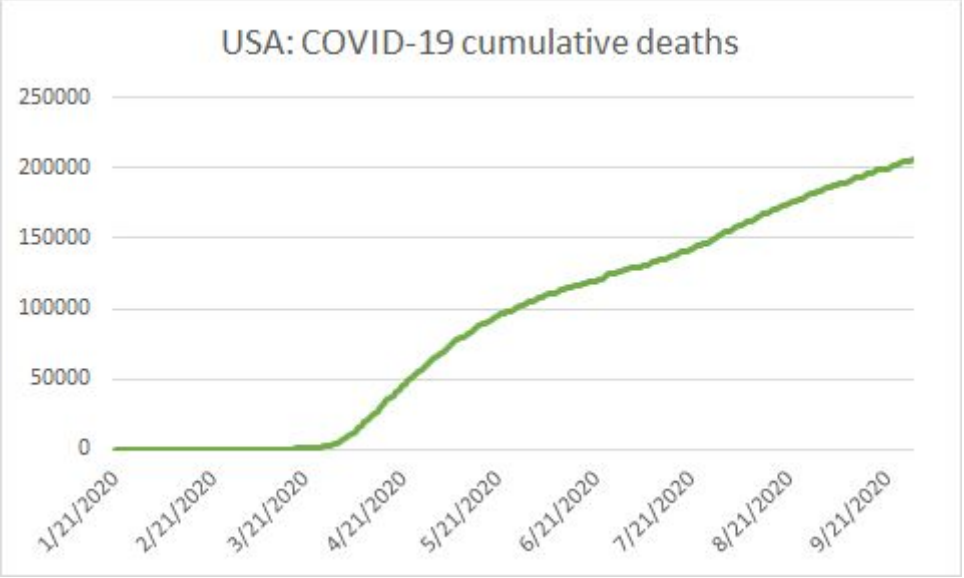

Florida: COVID-19 cumulative cases
700000
600000
500000
400000
300000
200000
100000
0
3/1/2020
4/1/2020
5/1/2020
6/1/2020
7/1/2020
8/1/2020
9/1/2020


Florida: COVID-19 cumulative deaths
14000
12000
10000
8000
6000
4000
2000
0
3/1/2020
4/1/2020
5/1/2020
6/1/2020
7/1/2020
8/1/2020
9/1/2020

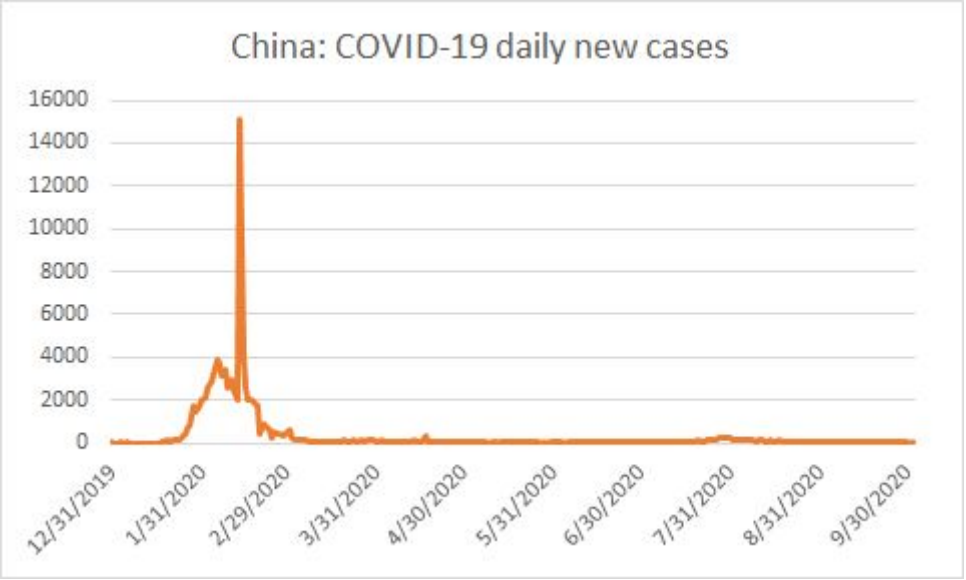
5. USA

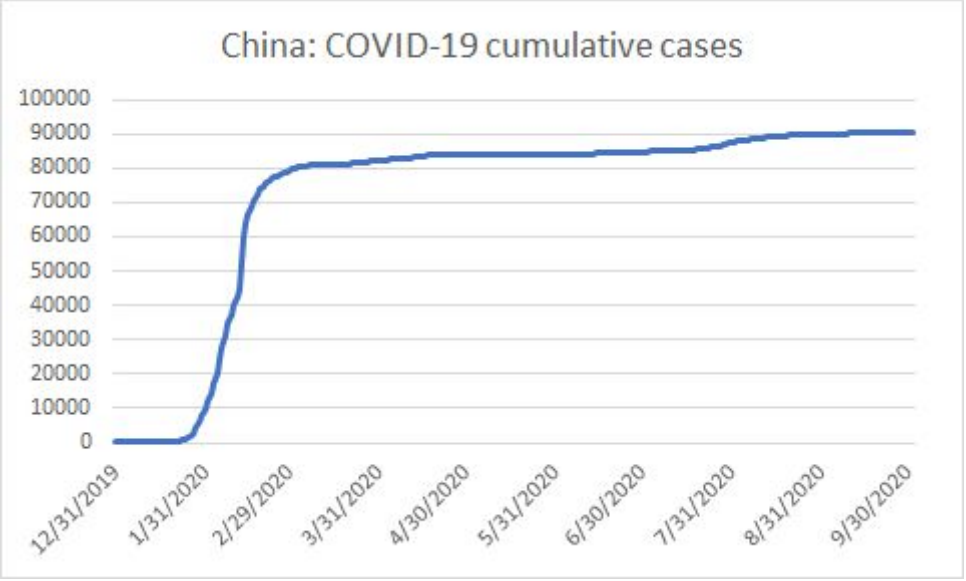

USA: COVID-19 daily new cases
80000
70000
60000
50000
40000
30000
20000
10000
0
1/21/2020
2/21/2020
3/21/2020
4/21/2020
5/21/2020
6/21/2020
7/21/2020
8/21/2020
9/21/2020

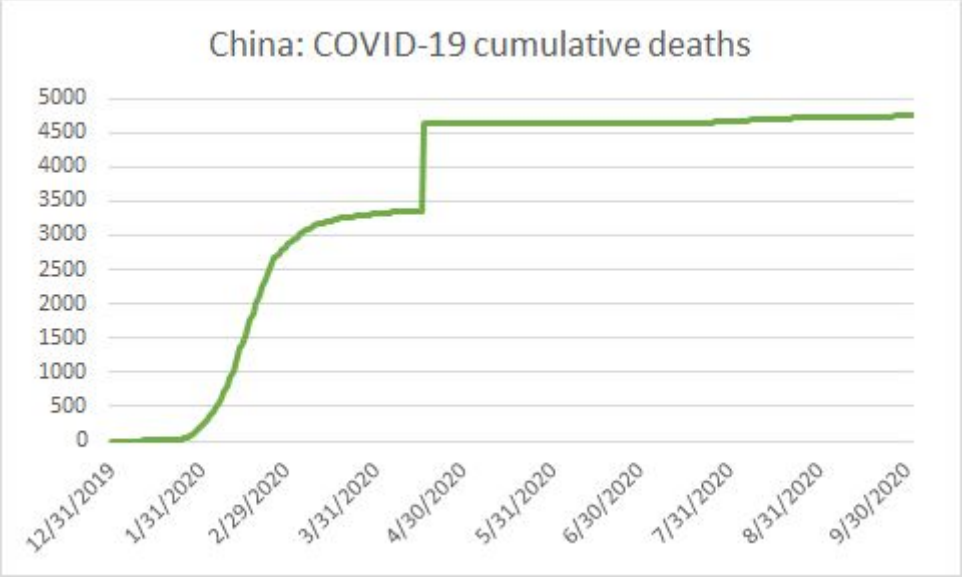

USA: COVID-19 cumulative cases
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
1/21/2020
2/21/2020
3/21/2020
4/21/2020
5/21/2020
6/21/2020
7/21/2020
8/21/2020
9/21/2020


USA: COVID-19 cumulative deaths
250000
200000
150000
100000
50000
0
1/21/2020
2/21/2020
3/21/2020
4/21/2020
5/21/2020
6/21/2020
7/21/2020
8/21/2020
9/21/2020

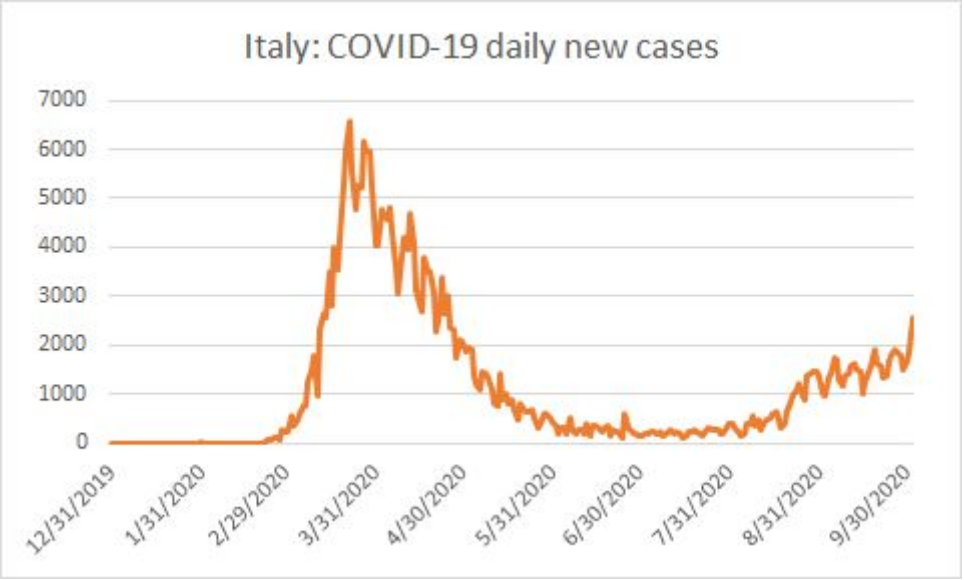
6. China

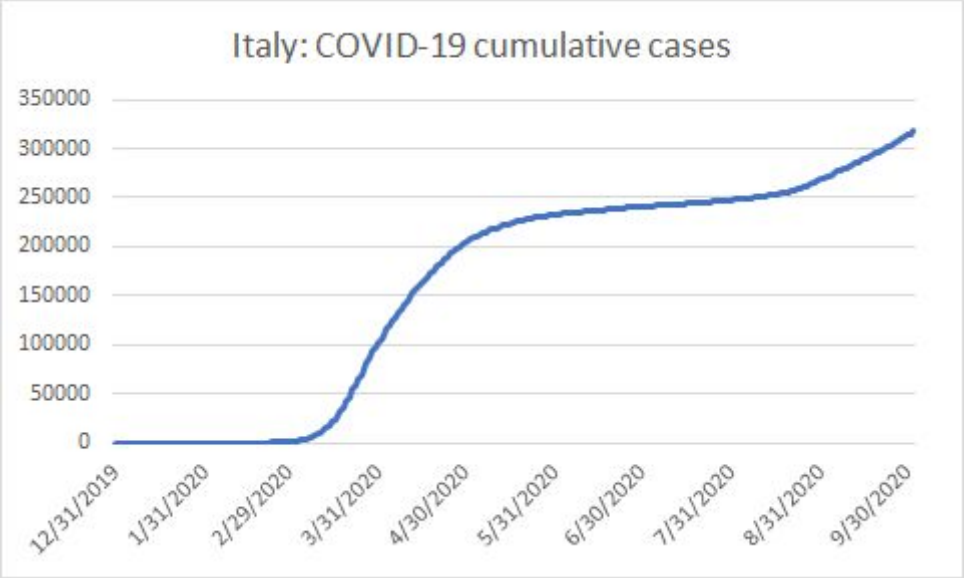

China: COVID-19 daily new cases
16000
14000
12000
10000
8000
6000
4000
2000
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020

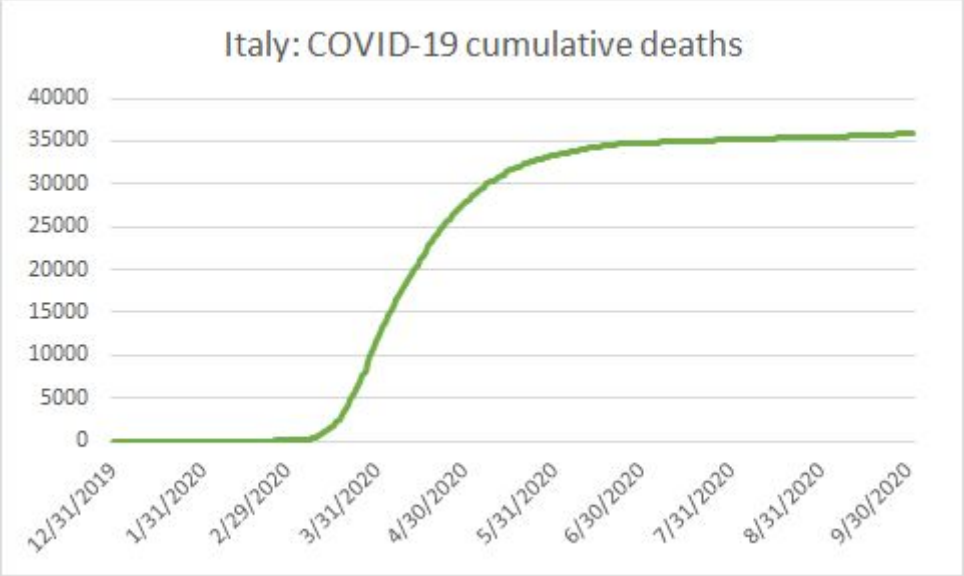

China: COVID-19 cumulative cases
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020


China: COVID-19 cumulative deaths
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020

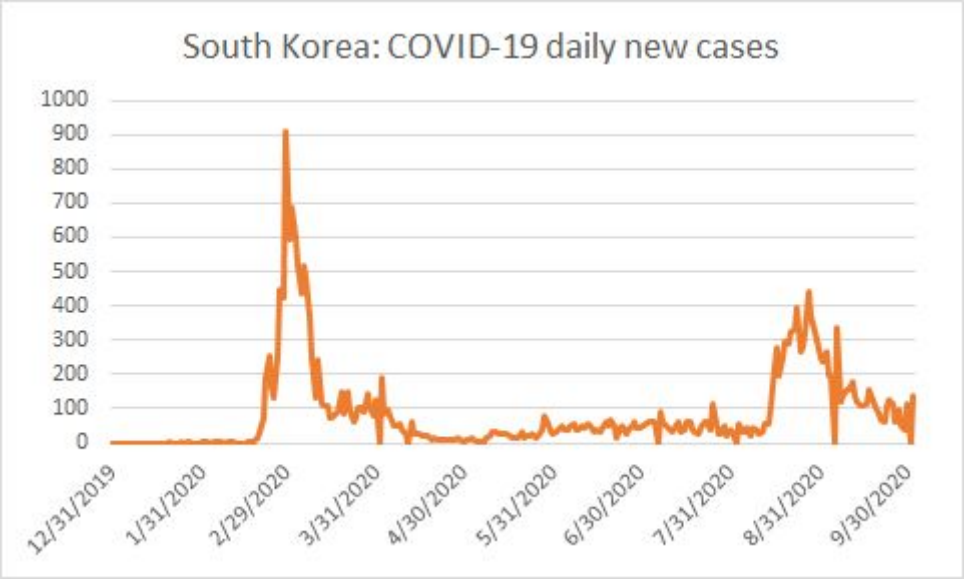
7. Italy

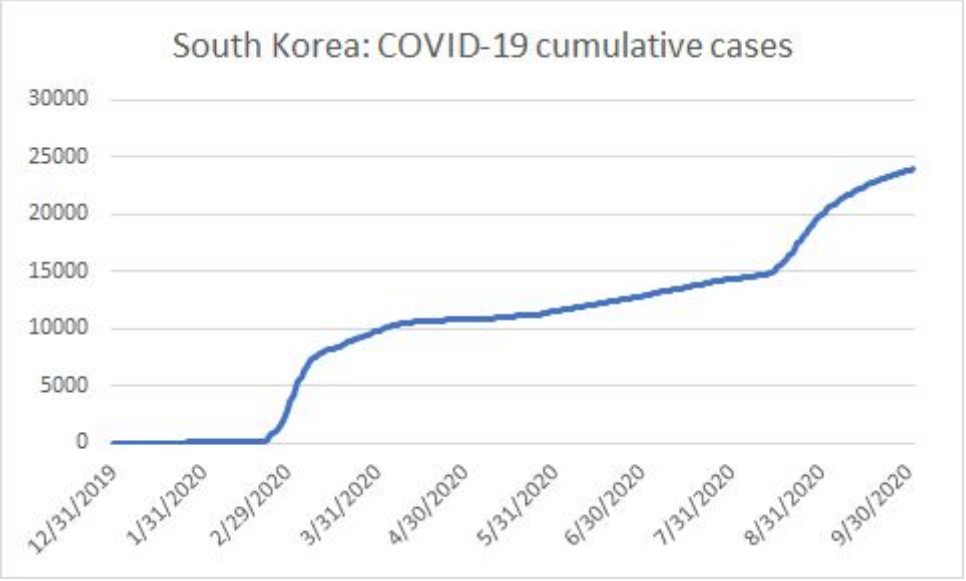

Italy: COVID-19 daily new cases
7000
6000
5000
4000
3000
2000
1000
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020

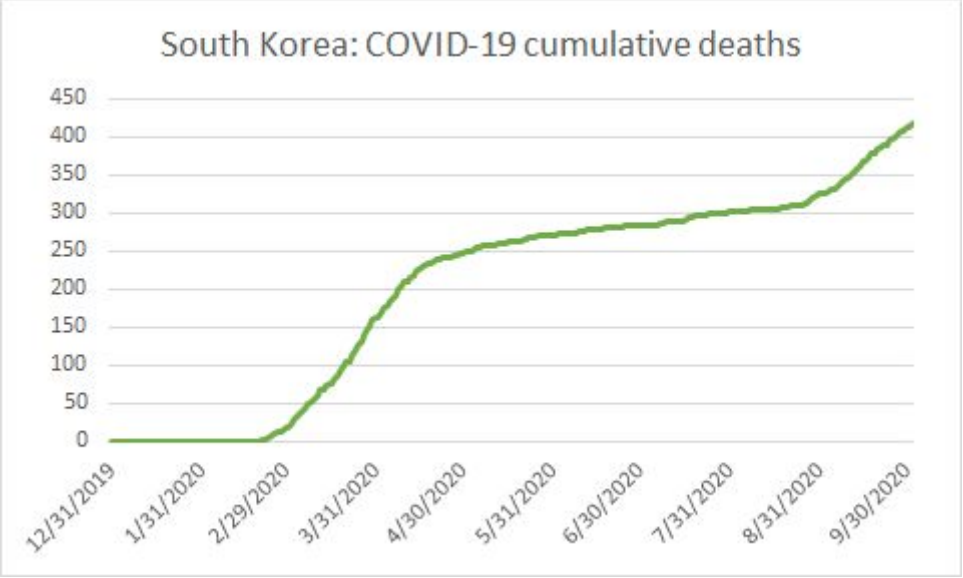

Italy: COVID-19 cumulative cases
350000
300000
250000
200000
150000
100000
50000
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020


Italy: COVID-19 cumulative deaths
40000
35000
30000
25000
20000
15000
10000
5000
0
12/31/2019
1/31/2020
2/29/2020
3/31/2020
4/30/2020
5/31/2020
6/30/2020
7/31/2020
8/31/2020
9/30/2020

8. South Korea







<u>Part II - B</u> Analyzing the differences between models and real data

1. Why does the "real world" data look different to the SIR models? In addition to the uncertainties around parameter values, what factors can influence the dynamics, which were not accounted for in the models?
   In the real world, humans change their behavior according to new information such as staying at home to prevent spread. Note that in the SEIR modeling exercise in Part I, the parameters are set at the beginning of the simulations, and remain constant over the course of the outbreak.

   Other factors that can influence the dynamics include access to healthcare, health literacy, and population density.

2. What do you think are the reasons for the observed decrease in the COVID-19 case counts?
   - Strict stay at home policies and social distancing
   - More thorough understanding of COVID-19 transmission that was then communicated to the public
   - Widespread use of PPE

   In part I in the SIS models, the decrease in the infected population was due to a decrease of the susceptible population, with the outbreak ending when there is nobody left to infect. In the case of the COVID-19 outbreak, the data shows a decrease in cases due to behavioral changes, but the important difference with the SEIR models is that, as soon as the lockdown or other measures are relaxed, the infection will spread again as there remains a large pool of susceptible population. This can be observed in the dataset, for example, in Italy.

3. Discuss how changing the scale from national to county level, changes the disease dynamics observed
   - Trends more reflective of a specific population are seen when the scope is narrowed. At a national level, the dynamics are averaged over a larger population so you will lose nuances particular to a certain group of people. Similarly, SIR and SEIR models illustrated in Part I also show what the disease dynamics are expected to be on average, and the heterogeneities among smaller populations might get lost.
   - This illustrates that even for the same disease and outbreak, the parameters defining the disease such as the probability of transmission and average contact rates, can be highly context and setting specific.

More information on case counts and deaths in the US and by state can be found here:
https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html

More information on case counts and deaths in the the world by country can be found here:
https://qap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html#global-overview-tab

Updated datasets for your classroom or personal use:
If you are interested in downloading and reformatting publicly available COVID-19 datasets like the ones used today, utilize our tutorial:
https://docs.google.com/document/d/1de5mFTujUf9graP3uSMYOcK8in5_1vv6WAx_mV4ZlCo/edit?usp=sharing

Data is updated daily.