SIR Worksheet Instructions

1. Open Microsoft Excel® or Google Sheets. These instructions were written for Excel 2016. There may be differences in older versions of Excel and in Google Sheets.
2. Set up your spreadsheet exactly as it appears below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| 1 | Total population (N) | Isolation (q) | Inf. frac. (f) | Rem. frac. (v) | Duration (T) | R0 |
| 2 | 10000 | 0 | 0.001 | 0 | 5 | 2.5 |
| 3 |  |  |  |  |  |  |
| 4 |  | Recovery Coef. (g) | Carryover (p) | Infectivity (b) | B |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 | Time (days) | Susceptible | Infectious | Removed | New Infections |  |

Note that only 3 data values (plus the total population) are needed to define a simulation with the SIR model in the absence of isolation or vaccination. The initial choices for T and R0 are estimates for the Spanish flu, which has been estimated to have killed at least 1% of the world’s population in the last pandemic before COVID-19. We are using yellow highlighting to call attention to the cells that contain the model data. N=10000 is large enough that rounding off the class counts to integers changes the model only slightly without making the numbers unreadably large. The population size is just a little bigger than the number of people on a large cruise ship and roughly that of a small rural city.

1. Enter the time 0 in cell A8.
2. The time corresponding to each increment needs to be entered in column A. Rather than enter the values individually, we can use an Excel formula. Enter the formula “=A8+1” into cell A9, which will set the next time step to be 1 day. Now select cell A9. Hover the mouse pointer over the lower right-hand corner so that the cursor appears as a plus sign. Click and drag the cell down to A128. You should see a column of times from 0 to 120.
3. Calculate the parameters g, p, b, and B from Equations 1 in document 1-2 by entering “=1/E2” in cell B5, “=1-B5” in cell C5, “F2/E2” in cell D5, and “=D5/A2” in cell E5. Use the formulas from Equations 1 to check the results (with a hand calculator).
4. Calculate the initial populations from Equations 8, 9, and 10 by entering “=C2\*A2” in cell C8, “=D2\*A2” in cell D8, and “=A2-C8-D8” in cell B8. The simulation will start with 9990 susceptible, 10 infected, and none removed.
5. Calculate the initial number of new infections from Equation 2 by entering “=MIN($E$5\*B8\*C8, B8)” in cell E8. The first formula calculates the quantity B\*S\*I, and the MIN function uses the smaller of this result and the number of susceptibles. This is necessary to make sure that the number calculated as becoming infected does not exceed the number who could have become infected.
6. Calculate the numbers in each class for time 1 from Equations 4, 5, and 6 by entering “=B8-E8” in cell B9, “=$C$5\*C8+E8” in cell C9, and “=D8+$B$5\*C8” in cell D9. Note that there is no isolation term in the formulas for I and R because nobody has been sick long enough to have symptoms. Check the spreadsheet results with hand calculations. You should get 9985, 13, and 2, rounded to the nearest integer.
7. Complete the calculations for time 1 by dragging the formula from cell E8 down to E9.
8. Spreadsheets are often hard to read because the formulas result in answers with too many decimal places. To keep this from happening, highlight cells B8 through E9. Click on Format in the Cells menu and then Format Cells at the bottom of the drop-down menu. In the dialog box, choose Number and 0 decimal places.
9. Complete the calculations by highlighting cells B9 through E9, hovering the mouse over the lower right corner of cell E9 so that it becomes a plus sign, and then double clicking. This should fill in the table of data down to row 128.
10. The spreadsheet is now set up almost correctly, but the isolation term Iq has not been entered. Edit cell C11 so that it reads “=$C$5\*C10+E10-$B$2\*$C$5^3\*C10”, which matches Equation 6. Then double-click the fill handle of cell C11 so that the new formula is carried down through column C. The numbers you see in any cell should be unchanged by these modifications because q=0.
11. Change cell B2 to 0.1. This should change the results in cells C11 through C128. Notice that the total of B15, C15, and D15 should be the same as A2. It isn’t because we still need to add the isolated individuals to the Removed column.
12. Edit cell D11 by appending “+$B$2\*$C$5^3\*C10” to the formula that is already there. Double-click the fill handle of cell D11 to copy the formula down the column. When finished, the totals of columns B, C, and D for any row should be the number in cell A2 (although the difference could be 1 because of the rounding process).
13. Restore the original scenario by entering 0 into cell B2.

**Make an X – Y Scatter Plot**

.

1. Click on the Insert tab at the top of the screen.
2. Select the Scatter icon and choose the graph with straight lines and no data points on it.



1. ***Do not trust that the program will select the correct data, even if the graph looks correct.*** To select the correct data, right click the chart area and choose “Select Data” from the drop down menu. Remove any data series that is automatically selected. You will be adding a series for each model.
2. To add the data, start by clicking the "Add" icon. Then name the series “Susceptible” and click the icon to the right of the typing field for the X values (shown here).

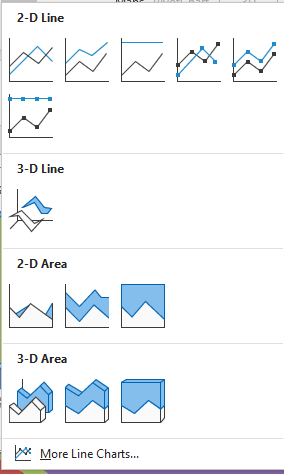


1. Once you do this, the spreadsheet will be active and you can highlight any column of cells you wish to make the X values for the graph. **Do not highlight column headings, just data points.** Highlight cells A8 through A12. *\*The formula for the X values should read* "=Sheet1!$A$8:$A$12". Now edit that formula by changing that final number to 128 and click the arrow icon on the right of the line.
2. Do the same for the corresponding Y values, but using column B to get the susceptible population instead of column A. Note that you must highlight an equal number of X and Y values for a scatter plot, so the final number should again be 128. Hit OK until the dialog boxes go away.
3. Move the plot so that the upper left corner of the plot window is in cell F4. This positions it so that you can easily see how it is affected by changes in the parameters.
4. Right-click on the graph and choose “Select Data” again. Repeat steps 19 through 21 to add three more series to the graph, one for the Infected population, one for the Removed population, and one for the New Infections count. For each of these, the last row for both X and Y should be 128.
5. Note the range of values on the X and Y axes. Software cannot think through the problem of choosing meaningful variable ranges, so we often have to change the default values manually. Population graphs should always start at 0 and go to a reasonable value. Right-click on the numbers on the Y axis and choose Format Axis from the menu. Change the Y axis range so that the maximum is 10000. The minimum value is currently set to 0 using the default algorithm of the program. Type in the number 0 instead so that the program will not be able to choose a different minimum value. Similarly, set the range on the X axis to be 0 to 120.
6. Graphs should always have axis labels, legends, and a title (in professional scientific papers, a caption is used instead of a title). Left-click on the plot area and choose the + sign that appears to the right of the graph. Check the boxes for Axes, Axis Titles, Chart Title, Gridlines, and Legend.
7. If the legend appears to the right of the graph, change it so that it appears on the bottom instead. To do this, put the cursor where the legend is and double-click. Choose Legend Options in the Format Legend dialog box and click the radio button that puts the legend on the bottom. Close the Format Legend box.
8. The generic axis and chart labels need to be changed. Left-click on each of these and type “Days” for the X axis, “Populations” for the Y axis, and “Epidemic Progress” for the chart title.

**Make an Area Plot:**

Another way of visualizing data for an epidemic model is with a stacked area plot.

1. Select the block of cells from B7 to D128. This includes all the data for Susceptible, Infected, and Removed, along with the column headings.
2. Click on the Insert tab at the top of the screen.
3. Select the Line or Area Chart icon and choose the right-most icon in the 2D Area group. This will automatically create the area plot. You do not need to modify the data set. You will need to move it to a position just under the scatter plot.



1. As with any plot, there should be axis labels. Repeat the procedure you used to get the axis labels on the scatter plot.

A stacked area plot shows how the proportion of population groups changes over time. Any vertical line drawn at a particular time will be divided into colors, with the length of each colored segment matching the height of the corresponding population on the line graph. Make sure that the same color codes are being used for the area and line plots.

**Final Setup:**

Graphs give a good indication of the progress of a simulation, but there are also important properties that are useful in characterizing what happens. While these quantities can be found by visual examination of the data, this examination can be automated using formulas. This makes is much easier to quickly identify the effect of any change in the current scenario.

1. The maximum number of individuals in class I is a measure of the stress the epidemic causes on the health care and economic systems. Type “maximum %I” in cell G1 and “=100\*MAX(C8:C128)/A2” in cell G2. Use the Format Cells menu to specify that cell G2 is a number with one decimal place. In our starter scenario, the maximum number of infected individuals at any one time is 24.9%.
2. The number of susceptibles at the end of the scenario is a measure of the risk of a resurgence of the epidemic. Type “final %S” in cell H1 and “=100\*B128/A2” in cell H2. As before, set the format to number with one decimal place. In our starter scenario, 9.4% of the population remains susceptible at the end of the epidemic.
3. Highlight cells G1:H2 in light blue to make it easier to identify key results.