

Interpolation and Extrapolation

There are times when data are taken that misses a region of interest. For example, if you know the population of a city at the start of each decade, but wish to guess the population at September, 2002. Graphing the population at 1980, 1990, 2000, and 2010 vs. year and sketching in a smooth curve that connects the points will allow you to guess the population at September, 2002. That process is called interpolation and is usually reliable if the gap is not large compared with the extent of data variability on either side.

On the other hand, if you wish to guess the population at September 2020, you must take the points up to 2010 and try to continue the curve up through 2020. That process is called extrapolation and is much less reliable. Unexpected events such as an economic collapse or a disease epidemic can cause your estimate to be quite wrong.

Just for fun, let's try to extrapolate the Arctic ice extend data in the following graph. The graph obtained from <https://nsidc.org/arcticseaicenews/charctic-interactive-sea-ice-graph/> is able to show data on the arctic ice coverage for past decades. This version shows 2012 as a dashed line, 2018 from January through July as a pale line. The solid line and its surrounding gray band show the average for 2000 through 2009. Try to guess where the 2018 curve will go by this September. When October arrives, you can see what really happened. This extrapolation is very unreliable because so many variables affect this ice coverage, but knowing how past years trended allows some idea of what we might expect.

