

F-IF A-SSE Modeling London's Population

Alignments to Content Standards: A-SSE.A.1 F-IF.B.4 F-IF.C.7

Task

The table below shows historical estimates for the population of London.

Year	1801	1821	1841	1861	1881	1901	1921	1939	1961
London population	1,100,000	1,600,000	2,200,000	3,200,000	4,700,000	6,500,000	7,400,000	8,600,000	8,000,000

No data was available in 1941 because of the war.

a. Can the London population data be accurately modeled by a linear, quadratic, or exponential function? Explain.

b. A *logistic growth* equation can be written in the form

$$P(t) = \frac{a}{1 + e^{-b(t-c)}}$$

where a , b , and c are positive numbers and t represents time measured in years. Using the application supplied, determine if the London population data can be accurately modeled by a logistic equation.

c. Explain the shape of the graph of P in terms of the structure of the equation $P(t) = \frac{a}{1 + e^{-b(t-c)}}$. What impact do the values of a , b , and c have on the graph of P ?

IM Commentary

The purpose of this task is to model the population data for London with a variety of different functions. In addition to the linear, quadratic, and exponential models, this task introduces an additional model, namely the logistic model. Logistic growth is characterized by having an stage of exponential growth which then slows down and eventually levels off. This model is used in many biological contexts where initial growth, fueled by general prosperity and availability of resources, can be very rapid. Eventually, however, the growth slows down as the population approaches the "carrying capacity" which is the size which the environment can support in the long term.

Many variables influence the reported population of a large metropolitan city like London: for example, political and economic health and which smaller outlying towns are considered part of London. Thus an exact fit with a relatively simple function can not be expected over a long period of time. As with many modeling tasks built on real data, this task will require patience and the technology supplied (by desmos) is an invaluable tool for analyzing the data and visually checking the fit of the different models to the given data.

The logistic growth model introduced in part (b) of this task is developed in an abstract version in <http://www.illustrativemathematics.org/illustrations/800> and in a concrete population setting in <http://www.illustrativemathematics.org/illustrations/804>. An alternative expression for the logistic growth function, used in these two tasks, is

$$P(t) = \frac{KP_0e^{rt}}{K + P_0(e^{rt} - 1)}.$$

where K , P_0 and r are positive numbers. This reduces to the given form when a , b , and c are chosen appropriately. In this expression, P_0 is the initial population, K is the carrying capacity, and r determines the growth rate.

The historical estimates of the population of London are from <http://www.demographia.com/dm-lon31.htm>. More modern data for the population of London can be found at <http://worldpopulationreview.com/population-of-london/>. For 1981 the population was about 6.6 million while in 2001 it was about 7.2 million. Finally, in 2010 it was about 7.8 million.

There are many questions the teacher might want to pose as the students work on this task including:

- What outside influences would cause a population to grow quickly then level off or decline? Do any of these influences apply to London after 1939?
- What happens to the predictions of the different models when compared with the historical data for 1981, 2001, and 2010?

This task has been designed to support MP4, Modeling with Mathematics. Students are given data and 4 different models which they will use to fit the data. There are at least three aspects to analyzing each model:

- Visually or numerically, how accurate does the model predict the known values of the population?
- What predictions does the model make for future years and are these predictions reasonable?
- Does the model capture the overall trend or shape of the population data?

Ideally, a good model should provide insight into all three of these aspects of the population. The task also supports MP3, Construct Viable Arguments and Critique the Reasoning of Others, as students may well disagree about which models are the best and they will need to articulate their reasoning and try to convince others.

[Edit this solution](#)

Solution

a. Below is a table showing the London population over this period with the differences, second differences, and quotients needed to test for linear, quadratic, and exponential growth.

Year	Population (in thousands)	Change in Population (in thousands)	Change in Population Changes (in thousands)	Successive Population Quotients
1801	1100	---	---	---
1821	1600	500	---	1.45
1841	2200	600	100	1.38
1861	3200	1000	400	1.45
1881	4700	1500	500	1.47

1901	6500	2200	700	1.38
1921	7400	900	-1300	1.14
1939	8600	*1200*	*300*	*1.16*
1961	8000	*-600*	*-1800*	*0.93*

Note the asterisks in the table: this is because we need to calculate differences, second differences, and quotients over *equal* intervals. With no population data to put in for 1941 we use the 1939 population as if it were the 1941 population.

A linear model can show the general trend of growth in the population but will not account for the changing rate of growth or for the decrease in the final decades. A good linear approximation will overestimate for the first few data points and then underestimate when the growth becomes faster in the middle of the period covered by the table. For accuracy a linear model can provide a reasonable fit to the data but it misses other important aspects of the growth.

There are two choices for a quadratic model, one which is concave up (with the rate of growth increasing over time) and one which is concave down (with the rate of growth decreasing over time). Another way of saying this is that a concave up quadratic model will have positive second differences while a concave down quadratic model will have negative second differences. Looking at the data, up through 1901 this can be very well approximated by a quadratic model with positive leading coefficient (concave up). But after 1901 two of the second differences are markedly negative which means that the rate of growth is now decreasing and a concave down quadratic is called for. With separate quadratics (one for 1801-1901 and one for 1901-1961) we could get a very good model. A single quadratic, however, will either miss the rapid growth over the first 100 years or the slowing growth over the last 60 years.

An exponential model fits the data extremely well for the first 120 years of the data as we can see that the successive quotients are all lie within a small range. Over the next 40 year the growth rate slows down dramatically and then for the 20 final years it begins to decrease. No single exponential model will be able to capture all three of these trends.

b. Below is an application which allows you to vary the three parameters a , b , and c . Preliminary values have been set showing that the general shape of the London population can be captured by this logistic equation. (Note: At the moment, the graph displays a very

zoomed out set of axes when viewed in Firefox. The graph appears correctly in chrome.)



We can visually see the influence each parameter has on the shape of the graph (by clicking on the button in the left hand column next to the given parameters) and this is described mathematically in the next part of the question.

c. To understand the shape of the graph of $P(t)$ we begin with the exponential function in the denominator. Since b is a positive number $e^{-b(t-c)}$ steadily decreases as t increases. The parameter b controls how fast this decay is. The larger b is, the more rapidly the exponential function will decay: since the exponential function is the denominator, as it decays the value of the function P increases and so the graph is steeper when b is larger. When $t = c$ we have

$$P(c) = \frac{a}{1 + e^0} = \frac{a}{1 + 1}.$$

When t is large on the other hand, $e^{-b(t-c)}$ will be very small and so the approximate value of P will be a . So a represents the value that the population approaches (according to the model determined by P) as time increases and c is the time when the population $P(c)$ is half of this limiting value a .

Summing this up, the parameter a determines how large the values of the function P

become as time grows. The parameter b determines how rapidly the population $P(t)$ increases: the larger the value of b , the shorter the time span of rapid growth and the more rapid that growth is. The parameter c moves the graph to the left or right without influencing the overall shape so it determines *when* the rapid growth period happens.



F-IF A-SSE Modeling London's Population
Typeset May 4, 2016 at 21:12:53. Licensed by Illustrative Mathematics under a
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License .