Demography and population projection

# Objectives

* Understand the use of projection matrices and the importance of age structure
* Relate population pyramids to past and future growth of populations
* Understand the main determinants of growth rates
* Familiarity with the causes of recent human population growth patterns

# Introduction

A **deme** is a population, so demography is the study of populations, but with an emphasis on their changes through time (**dynamics**) and the factors affecting those changes. Although it is easy to come up with simple models that project future population sizes based only on the current number of individuals and a reproductive rate (which may depend on the density), we will focus here on models that include **age structure**. In these models, the survival and reproductive rates vary with the age or life stage of the organism. This structure can have important consequences for the pace of population growth. For example, if most of the population is in the juvenile stage, future growth may be rapid, but current growth will be slower than for a population that has most individuals in a mature stage.

Data on age-specific survival and growth rates are typically presented as a **life table** (which can be turned into a matrix, as we shall see). For each age group *x*, the life table lists the fraction surviving (lx), the average number of offspring born to a female in that age group (mx) and usually the product of these: lxmx. The lx is usually calculated from the observed numbers of individuals of each age class (Sx), so lx =Sx/S0.

Table 1 A life table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| x | Sx | lx | mx | lx mx |
| 0 | 500 | 1.0 | 0 | 0 |
| 1 | 400 | 0.8 | 2 | 1.6 |
| 2 | 200 | 0.4 | 3 | 1.2 |
| 3 | 50 | 0.1 | 1 | 0.1 |
| 4 | 0 | 0 | 0 | 0 |
| R0 =Σ lx mx | | | | 2.9 offspring | |

The sum of the lx mx values is the net reproductive rate of the population, R0. If R0 > 1.0 the population will increase, and if R0 < 1.0 the population will decrease.

# Calculating the R0 in R: Squirrels

To do this in R, the lx and mx values are read into separate vectors of the same length. You will multiply the elements of these two vectors to yield a vector of lxmx values. The sum function in R will sum the elements of a vector, so you will use that to calculate the R0 value of a squirrel population.

* Start R Studio. Find the script “squirrels.R” in the folder on the desktop of the lab computer or download it from Canvas (right click the file name on a PC, control-click on a Mac, and save it to a place where you can find it). Load the script. Do not run it yet – you need to change a couple of things – read on!

Let’s calculate R0 for this squirrel population using R. The table (11.1) is from the reading.

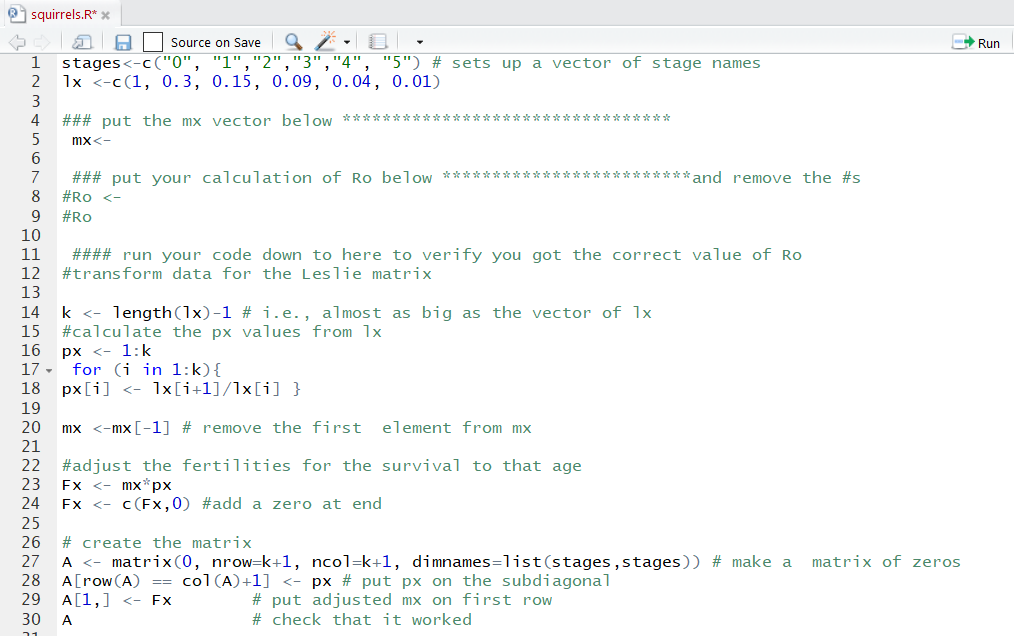


Figure 1 First part of Squirrels.R code

We need to make a vector for lx and one for mx. The lx is shown below. It is in the file “squirrels.r”.

Lx <-c(1, 0.3, 0.15, 0.09, 0.04, 0.01)

* **Write the vector for mx using the data in Table 11.1. Add that to the code in line 5**.

**Question 1.1.** Figure out the expression for R0 =Σ lx mx in R \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. **Insert this expression on line 8 of the code, and as the answer to Question 1.1 on Gradescope.**

* Run your code **down to line 10** to verify that you got the value shown above in Table 11.1.

# Projection matrices

**Projection or transition matrices** are used to predict the future population size when a species has variable survival and fecundity over its lifespan (which is true of many species, including humans). We divide the life span into age classes (e.g., 0-5 years, 6-10 years, etc.) or stages (egg, juvenile, sub-adult, adult) that have reasonably constant survival and fecundity within the stage.

For a species with 5 age classes (stages) we have a 5x5 transition matrix[[1]](#footnote-1), and a 5-stage initial vector:

***M***= 0 =

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *P1* | *F2* | *F3* | *F4* | *F5* |  |  |  | *N1* |
| *G1* | *P2* | 0 | 0 | 0 |  |  |  | *N2* |
| 0 | *G2* | *P3* | 0 | 0 |  |  |  | *N3* |
| 0 | 0 | *G3* | *P4* | 0 |  |  |  | *N4* |
| 0 | 0 | 0 | *G4* | *P5* |  |  |  | *N5* |

Gi = proportion surviving and moving from stage *i* to stage *i+1* in a time interval

Pi = proportion remaining in stage *i* in a time interval

Fi = fecundity (offspring) produced in stage *i* and surviving to enter stage 1

Ni = Number in stage *i* (so ΣNi = N total)

If the time interval and the stage classes are the same duration, then all of the Pi values will be zero, since all individuals will age out of one stage class and into the next. For short-lived species the stage classes may be annual and the time interval one year. For humans, the age-classes are often five years and the projections may also be at five-year intervals. For some species there are stages of varying durations and types.

Given the projection matrix and a starting population vector, we can then predict the population size in the next time interval as: 

For example, in a species with only 3 age classes, and using letters in place of numbers:



The total population size is of course N1 + N2 + N3 .

Here’s what happens over time, starting from 50 newborns:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***M***= |  |  |  |  | Initial *N* | Next year | Second year | Third year |
| 0 | 1 | 5 |  | N1 | 50 | 0 | 10 | 10 |
| 0.2 | 0.5 | 0 |  | N2 | 0 | 10 | 5 | 4.5 |
| 0 | 0.1 | 0.5 |  | N3 | 0 | 0 | 1 | 1 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  | Total N | 50 | 10 | 16 | 15.5 |
|  |  |  |  | R\* | 0.2 | 1.6 | 0.97 | . |

\*  , the rate of change, a value we will examine in more detail below.

Of the 50 initial individuals 50\*0.2=10 survive to the next year and grow into the second age class (N2).

These ten give birth to a total of ten offspring that enter class N1 in the second year. Five of these ten (10 \* 0.5) survive and stay in N2; and one of the ten (10\*0.1) ages into N3. In the next time interval, the five individuals in N2 and the single individual in N3 produce a total of ten offspring that enter N1 in the third year. Two of the previous year’s offspring survive into N2 (10\*0.2), and two and a half remain in the stage (5 \* 0.5). Half an individual (5 \* 0.1) grows into N3, and half an individual (1\*0.5) remains in N3. You might be troubled by the idea of fractional individuals, but the elements of a population matrix often represent densities (e.g., individuals per unit area) rather than numbers of individuals, so it’s not as ridiculous as it sounds.

#### Squirrels projection matrix

The projections can be easily done in R once the matrix is set up. For this exercise, we provide you with most of the code, the goal of which is to convert the information in the life table into the elements of the transition matrix.

Here is what the code does in lines 14-30: First, we calculate the px values (the fourth column of the life table 11.1) from the values in the lx vector: px = lx+1 / lx. These px values correspond to the Gi in the transition matrix above, since this is an annual matrix in which all surviving squirrels move to the next age class the next year. We then multiply the elements of this px vector by the elements of the mx vector you entered earlier. This allows us to calculate the reproductive output for each class. Finally, we compile these values into the matrix A, which has the fecundities (birth rates) in the first row and the px values on the subdiagonal.

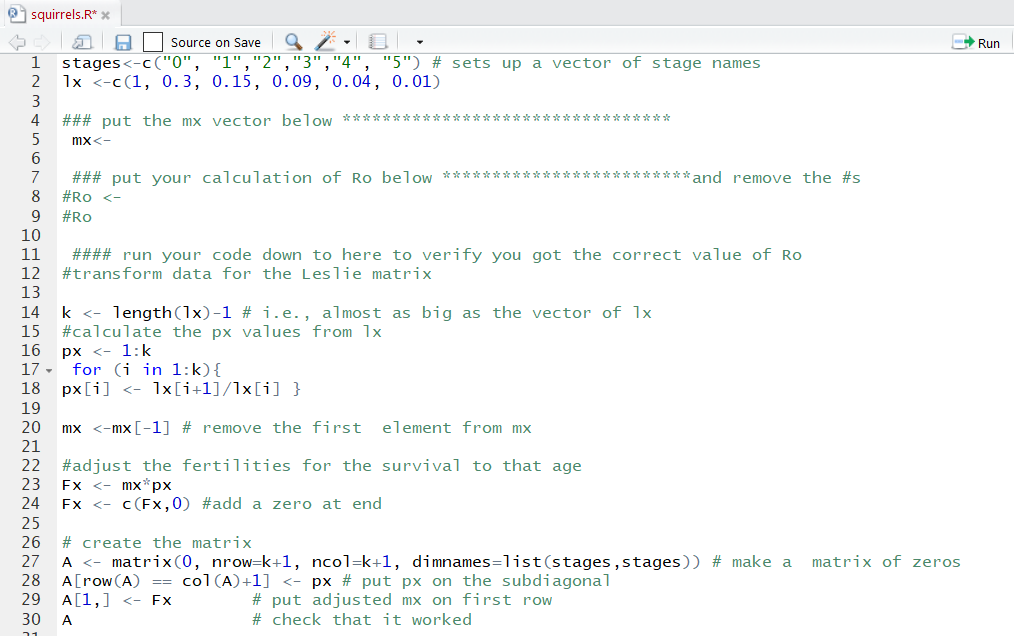


Figure 2. Second Part of Squirrels.R code.

* Your next task is to enter the initial population vector on line 41 of the script file. In the reading, it was 20 newborns and 10 one-year old squirrels. The other 4 age classes are 0s, but they need to be entered too.

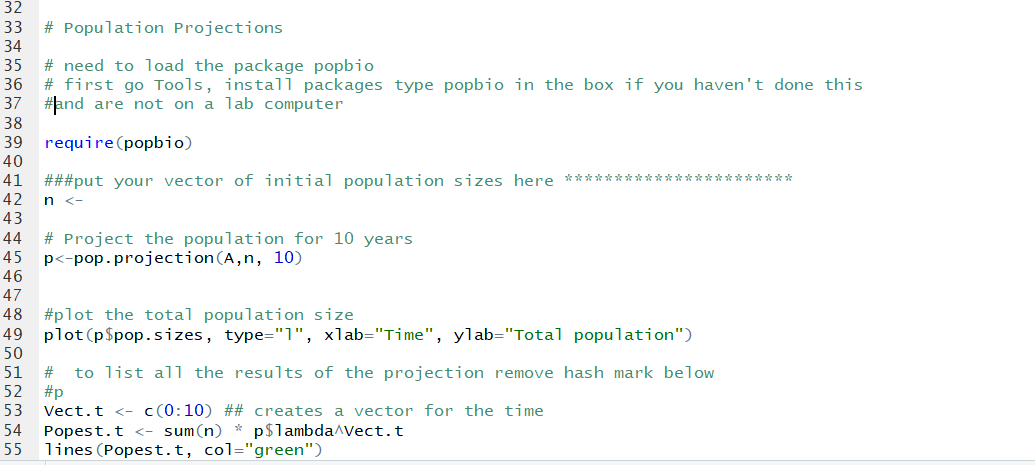


Figure 3. Third part of Squirrels.R code

We are going to use an R package that is designed to handle projection matrices. Packages are sets of R functions, data, and compiled code in a well-defined format, written by users. Often they are based on books, as is the one we will use. To use a package, it must be downloaded, installed once, and then loaded into R for a particular session any time you wish to use the package.

* To install a package: in R Studio – find the Tools menu and pick “Install Packages.” You then just type the name of the package in the empty box and click the install button.

The package we want is called “popbio” and is based on two books, *Matrix Population Models* (Caswell, 2002) and *Quantitative Conservation Biology* (Morris and Doak, 2002).

To load the package for a session, you can either use library(popbio) or require(popbio). We have used the latter because it is a little clearer.

The line of code

p<-pop.projection(A, n, 10)

will multiply the matrix **A** by the population vector (starting with **n**), for **10** years to project the population over that time interval. The data frame of results contains many items: the stable age distribution, the size of each age class at each time, the total population size at each time, and the changes in population size from year to year (Nt+1/Nt).

* Run the remaining lines of code (11-end) in squirrels. Note the correspondence to the values in your reading.
* Note that there are two plots: the total population size vs. time, and the age structure (proportion of the population in each age class) over time.

From here, it is easy to project the population under new assumptions about the fecundity and mortality schedules. For example, imagine that you want to investigate the effects of a habitat change that will lead to new, mostly higher survival probabilities but poorer offspring production. If the life table changes as described below, you can predict what will happen to the population after the change in habitat:

* Replace your lx vector with the following values: lx = c(1, 0.3, 0.3, 0.18, 0.08, 0.02). (Add this vector on a new line. Do not delete the original vector—just comment it out).
* Adjust the mx values so they are 1/3 of their original values by adding the line mx <- mx/3 on the line below your original *mx* vector.
* Re-run everything.

**Question 1.2.** What happens to the population over time with this new scenario?

# The rate of change, R, and λ

Like R0 from the life table, R is useful because it tells us if the population is growing (R>1), shrinking (R<1), or stable (R=1). R is calculated for each time step t as Nt+1 / Nt. In a population with age structure and constant birth and death rates, R fluctuates initially but eventually stabilizes such that the population continues on a trajectory determined by **λ , which is** t**he stable value of R**. We will see that λ can be calculated directly from the matrix, rather than running the projection until the value of R stabilizes.

In the models that we will use, there is no change in the birth and death schedule, there is no density-dependence, and there is no variability due to environment or chance, and thus it would appear that the growth rate should be constant over time. Without considering age structure, a population with a fixed rate of change λ, will grow (or shrink) according to:

(eqn. 1) Nt = N0λt

The initial population size is the sum of the elements of vector n,and λ was calculated by the projection function (and is called p$lambda). The following lines of the code do the above calculation.

Vect.t <- c(0:10) ## creates a vector for the time

Popest.t <- sum(n) \* p$lambda^Vect.t

By also running the command below (still using the modified vectors from Question 1.2), you can compare the growth of the age-structured squirrel population to its growth with no age structure starting from the same total number of squirrels. This command will plot the growth trajectory on the same graph you generated above:

lines(Popest.t, col="green")

So, the black line has age structure, and the green line does not have age structure.

**Question 1.3** Add your names somewhere on the graph of total population vs. time, and upload this (as a PNG, BMP, GIF file) to Gradescope

**Question 1.4** How do the total population size lines (green and black) compare? Are they identical? If not, describe how they differ.

# Age Structure

* **Return to the original values for the lx and mx** **& re-run all the code.**

The code also plots the proportion in each age class or stage over time, using the function stage.vector.plot. (lines 58-61)

**Questions 2.1 & 2.2.** Be sure you are using the original vectors for lx and mx. What are the relative sizes of the age classes? What happens to them over time?

Now we can explore why we should care about age structure.

What if we had a plague that killed all of the young squirrels, so we are left with just older squirrels? Let’s model this scenario.

* Make a new initial vector with no newborns or 1-year-olds, 4 of age 2, 10 of age 3, 12 of age 4, and 4 of age 5. Note that the initial number of squirrels has not changed.
* Re-run the code with this new age structure, adding titles to the graphs that reflect this change.

**Question 2.3.** Explain why the population’s trajectory (black line) is different from the trajectory it would have if there were no age structure (the green line).

Note that in both of the graphs showing the age structures, the lines eventually settle down to a constant proportion in each stage. This is in fact a general property of these projections; each results in a stable age distribution. What will happen if we start the population near the stable age distribution?

We can find the stable age distribution because it is calculated when you do the projection, and is part of the data frame of the results (here called p). It is a vector called stable.stage.

* Type p$stable.stage in the console to see the stable age distribution vector.
* Turn this into a vector of integers, multiplying by 100 and rounding up (you can do this by hand if you like).
* Use this new vector as the starting age structure. Alter the code and re-run it.

**Question 2.4.** What happens to the population size of the projections with and without age structure when you start at the stable age distribution?

**Question 2.5.** What happens to the number in each age class in the structured squirrel population over time?

**Question 2.6** Put your names on the graph of population size vs. time and upload it to Gradescope.

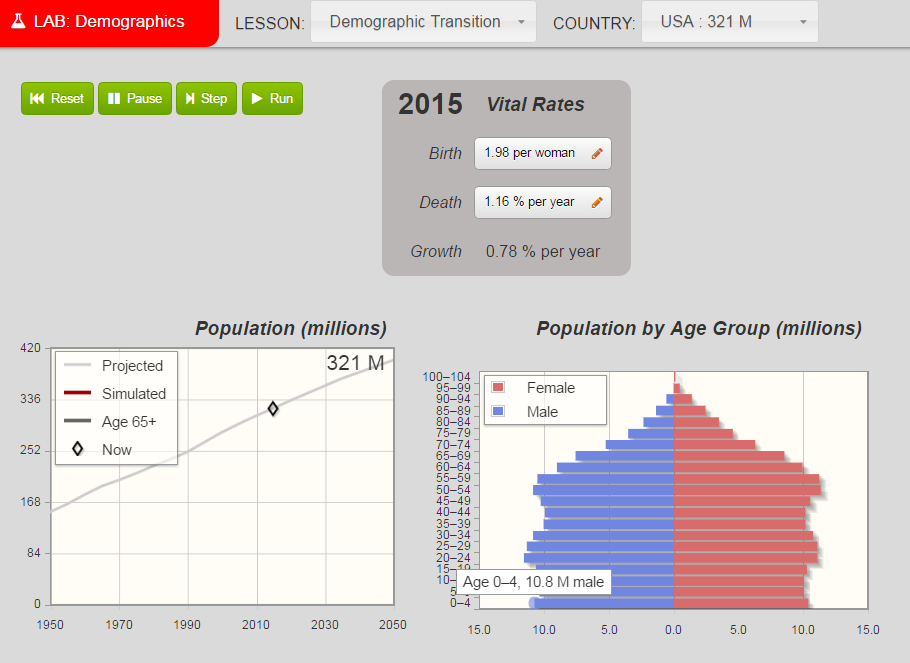
# Population pyramids

A population pyramid divides a population into age classes for males and females. The population pyramid provides an intuitive visual representation of the age structure of a population. The age structure can transiently affect the population growth rate, as you saw above. The shape of the pyramid is also affected by past growth and mortality events (wars, famines, disasters). Here we will use a simulator that plots both population size and the pyramids. The simulator is at:

<https://www.learner.org/wp-content/interactive/envsci/demographics/demog.html>

The simulator allows you to view pyramids and growth trajectories for nine countries at different stages of demographic transition and with different histories of growth.

We will look at Indonesia, Italy, and Nigeria, which have highly contrasting histories and patterns of growth. Select the country you want to view from the pull-down menu at the top of the simulator page.



**Question 3.1 -3.3**. Match each country in the list with its population trend by running the projection to 2050.

**Question 3.4**. The pyramid for Italy is widest in the middle. Why is this?

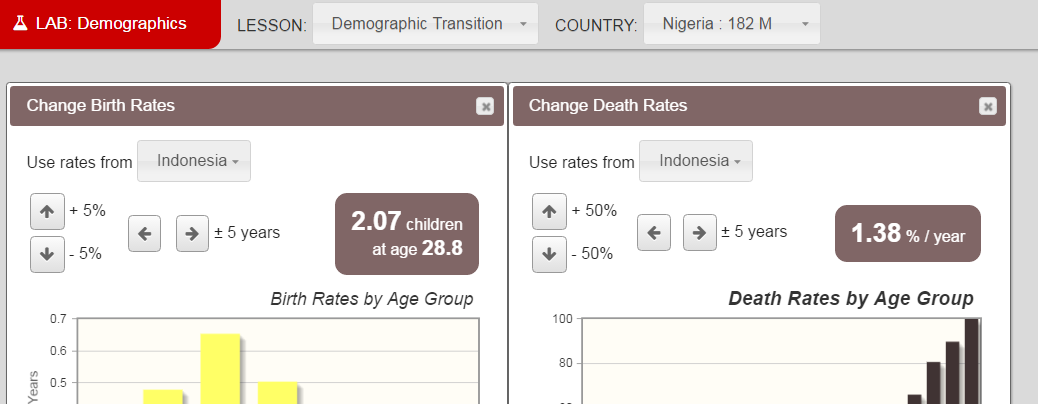
* To see how the population pyramids change in the future, you can run the simulation forward in time. Click the green run button. Click twice to project out to the year 2100. Do this for each of the three countries.

**Question 3.5-3.7**. What happens to the pyramids of these countries?

# Population momentum

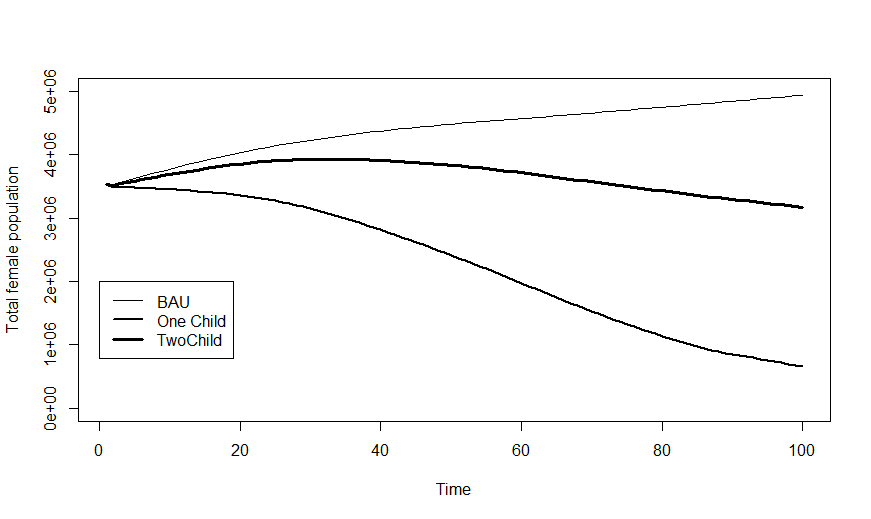
You should note that Nigeria has a very high growth rate. What would happen if we put the brakes on the growth by immediately reducing the births per female to replacement rate fertility (about 2.1 offspring per female)? We can do this by following the instructions below to give Nigeria the birth and death rates of Indonesia.

* Reset Nigeria. Under 2015 *Vital Rates* first click on the box next to “Birth.” Where it says “Use rates from” click the box and choose “Indonesia.” Be sure to click “Apply.” Repeat for “Death” and apply. Verify that the Birth value is now 2.07 per woman. Run the projection to the year 2100.



**Question 4.1.** What happens to the population size of Nigeria when you set its growth to the stable value?

As we saw with Nigeria above, and in the Age Structure part of the lab, age structure can lead to temporary states that affect the growth rate. Below is another illustration of the same principle, using the current world population size.

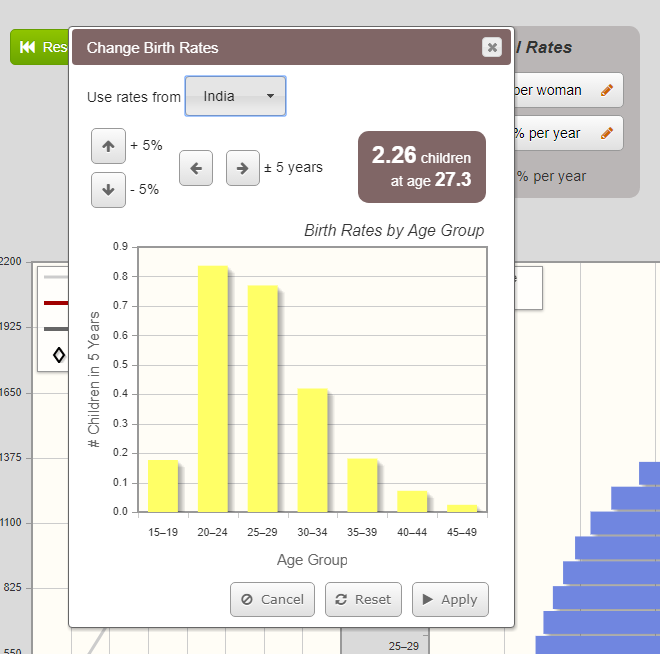


The graph compares “Business as Usual” (BAU), i.e., the growth of the population if it continues at 2015 rates, to the growth if either every couple had two children, starting in 2016, or every couple had only one child.

**Question 4.2.** What is your expectation for the growth of the population if every couple has 2 offspring? How and why does the graph deviate from this expectation?

# Timing of reproduction

In long-lived organisms with multiple potential reproductive episodes like humans, the timing of reproduction can also affect growth rates. We can use the web-based demography simulator to investigate these effects.

*  Select India. Record the birth rate, run the projection to 2100, and record the population size as the baseline run in the table below.
* Now, we will shift the reproduction 5 years earlier. Click the birth rates box, and then click the left-pointing arrow, which will shift the births 5 years earlier. Click “Apply.”
* Record the new birth rate, reset the population and run the simulation through 2100 again.
* Now, we will shift 5 years later with a similar series of clicks. **Be sure to reset the timing of reproduction to the original distribution before you click the right pointing arrow to shift it 5 years later.** Record the birth rate and population size in 2100 below.

|  |  |  |
| --- | --- | --- |
| Run | Birth rate (offspring per female) | Population size in 2100 |
| Baseline |  |  |
| 5 years earlier |  |  |
| 5 years later |  |  |

**Question 5.1.** Was there a large shift in the number of offspring per female (greater than 0.2 offspring per female different)?

**Question 5.2** What accounts for the changed growth rate when the births are earlier?

# Global patterns

How has the global population growth rate changed recently? The figure below compares population projections based on the survival and fertility rates in the early 1950s to those based on more recent rates (with the same initial vector of population sizes). Why do they differ – what has changed in the fertility and survival rates and why? We explore this issue below.

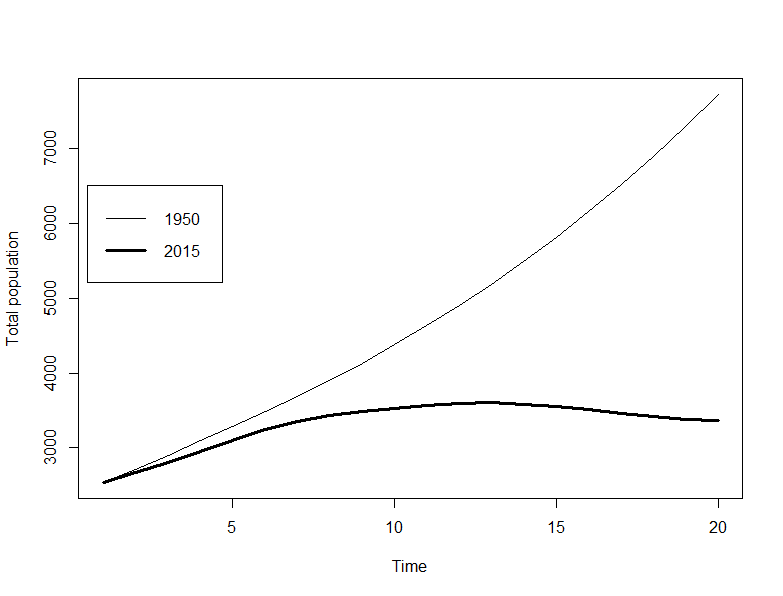


Figure 4 Global population size based on survival and fertility at two different times

Economic development is one possible explanation for these changes. Around the world, most countries that have high fertility rates are less developed, and those with slower growth or shrinking populations are more developed, with higher average incomes per person. Is this a direct effect of income, or are there other variables that have changed that have better explanatory power?

We can get all kinds of data, brilliantly presented, from GapMinder.org. They have data on development and population and economics, back to 1800 for some countries, and they plot them as bubble graphs.

* Go to: <https://www.gapminder.org/tools/> (pick bubbles). There is a link to this page on Canvas.

In a typical graph there are bubbles for each country that are proportional in size to the population (by default, but you can change this) and are coloured according to region (also adjustable). You can mouse over a point to see the country name. On the bottom is a time line. You can drag the slider to move through time, or press play.

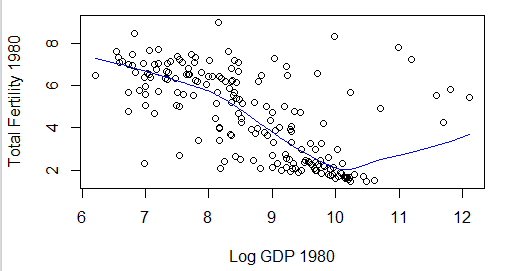
To change which variable is plotted, click on the name of the variable below the plot. You will get a list of variables and categories of variables (with the triangle pointing right). Click to pick the category and variable you want. You may also change the display to a logarithmic axis. You can see fascinating trends and changes here. We will focus on correlates of total fertility, because total fertility is the largest driver of population growth. But what affects total fertility?

* Plot the total fertility (babies per woman) on the y-axis and change the x-axis to “age at first marriage (women),” which you will find in the population category. Total fertility is correlated to the timing of reproduction (which we manipulated earlier in the simulator).

**Question 6.1**. What happens to age at first marriage over time? (Post 1900; data are sparse before then.)

**Question 6.2**. What is the relationship between total fertility and age at first marriage?

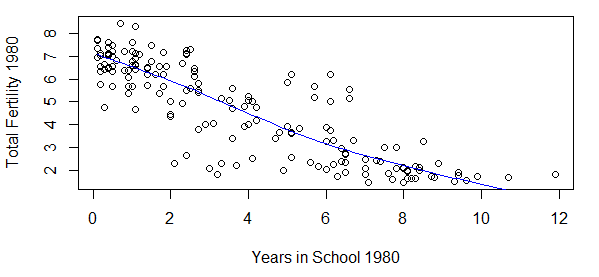
Returning now to the question of economic development, we can plot total fertility against per capita gross domestic product, a measure of wealth. In fact, one can download the data, and we did so in order to calculate the correlation between fertility and GDP. We picked a random year, 1980, although extensive analyses (Lutz & Qiang, 2002) show the pattern holds in most years.



r = -0.60

Figure 5 Relationship between Fertility and Gross Domestic Product in 1980. Each point is a different country.

So there is a correlation to GDP, but are there better explanatory variables? Below we have plotted total fertility vs. the years in school for women older than 25, also for 1980.

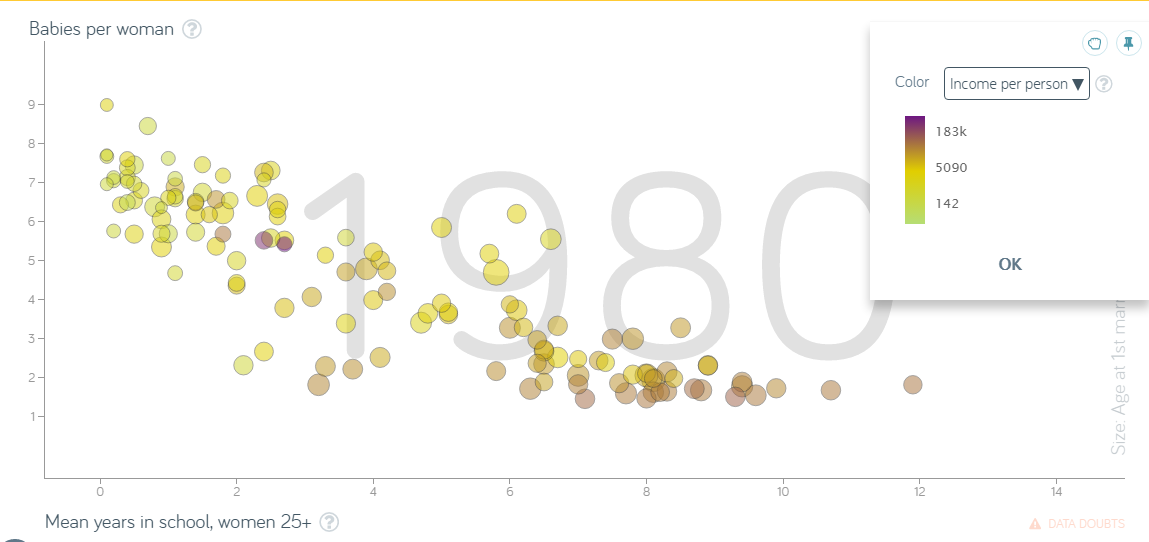


r = -0.85

Figure 6 Relationship between fertility and average years of school attendance in 1980

**Question 6.3**. Which variable explains more of the variation in fertility in 1980?

You may object that there is a relationship between the two explanatory variables: countries with more per capita income have longer education. In the graph below, the points are coloured according to income (colour version on Canvas). You can see that some of the deviations around the line are due to income – i.e., for a given number of years of education, poorer countries have higher fertility than richer countries (with some exceptions). In a much more extensive analysis, Lutz & Qiang (2002) found that female literacy was the single most important factor influencing the fertility rate.



On Gradescope, do not forget to add your partner’s name and submit the assignment.

References Cited

Caswell, Hal.2001. *Matrix population models*. Sunderland, MA; Sinauer Associates.

Lutz, W. & Qiang, R. 2002. Determinants of human population growth. Philosophical Transactions of the Royal Society B: Biological Sciences, 357(1425), 1197–1210. http://doi.org/10.1098/rstb.2002.1121

Morris, W. F. & Doak, DF. 2002. *Quantitative conservation biology: theory and practice of population viability analysis*. Sunderland MA; Sinauer Associates.



1. In this age-based transition matrix the sub- & super-diagonals are zero because organisms do not skip over stages, nor do they get younger over time. If it is stage-based, individuals may transition in and out of a reproductive stage. [↑](#footnote-ref-1)