3E8

Datasheet: Statistical Tables

1 (a)

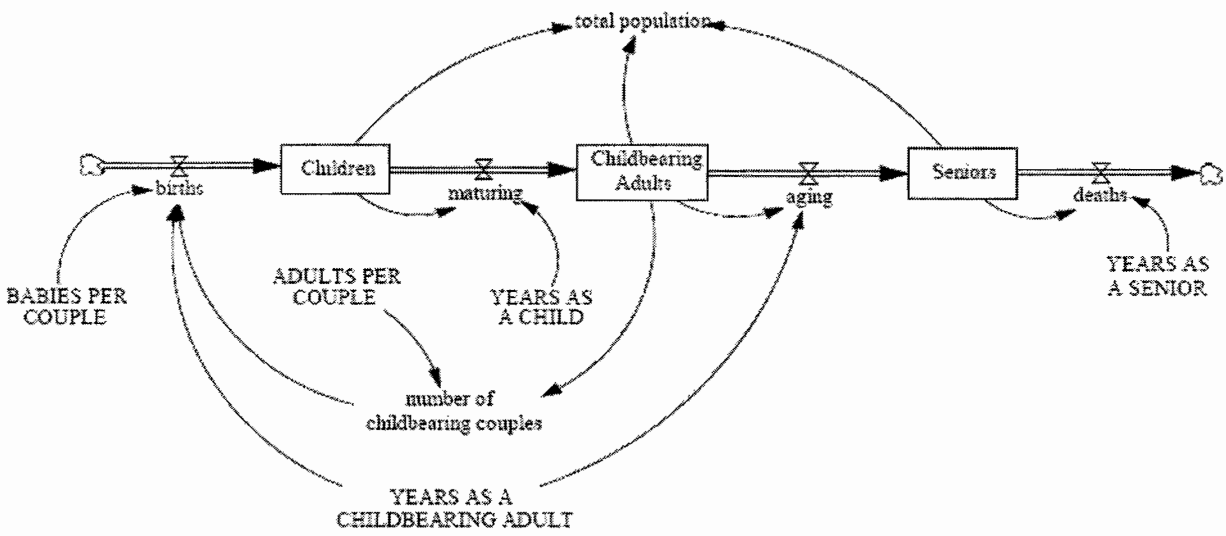
(i) The list includes auxiliary variables, which are used to disaggregate rate equations.

Name	Type	Units
maturing	flow out of children into childbearing adults	people/year
years as a childbearing adult	constant	years
children	stock	people
number of childbearing couples	auxiliary variable	couples
deaths	flow out of seniors	people
years as a child	constant	years
aging	flow out of childbearing adults into seniors	people/year
births	flow into children	people/year
seniors	stock	people
babies per couple	constant	people/couple
childbearing adults	stock	people
adults per couple	constant	people
years as a senior	constant	years
total population	auxiliary variable	people

("BABIES PER COUPLE" has units of people/couple. One could also define "BABIES PER COUPLE" to have units of people/people, making the units effectively dimensionless. Using more specific units, however, often makes it easier to keep track of different variables while formulating equations. The rate of "births" is then the "NUMBER OF CHILDBEARING COUPLES" multiplied by the number of "BABIES PER COUPLE," divided by the number of years over which the couple may have children.

The "total population" should not be modeled as a stock with "births" as inflow and "deaths" as outflow. Rather, the "total population" should be modeled as an auxiliary variable that is the sum of the other three stocks. Although in reality, the total population is a stock, modeling the "total population" as a stock would be duplicating other stocks in the model. In addition, if one modified the model to have outflows of deaths from each stock, then modeling the "total population" as a stock would be wrong, but modeling it as an auxiliary sum would be correct.

(ii) Model Diagram



(iii) Model equations:

ADULTS PER COUPLE = 2
Units: person/couple
Number of adults in a childbearing couple.

aging = Childbearing Adults / YEARS AS A CHILDBEARING ADULT
Units: person/year

If people are childbearing adults of n years, then on average 1/n of all childbearing adults become seniors each year.

BABIES PER COUPLE = 2
Units: person/couple

The number of children born by each couple during the couple's childbearing years.
births = number of childbearing couples * BABIES PER COUPLE / YEARS AS A

CHILDBEARING ADULT
Units: person/year
The number of children born by all childbearing couples per year.

Childbearing Adults = INTEG (maturing-aging, 1000)
Units: person

Number of people between 20 and 45 years of age.
Children = INTEG (births-maturing, 1000)
Units: person

Number of people 20 years of age and younger.
deaths = Seniors / YEARS AS A SENIOR
Units: person/year
If people live n years as a senior, then on average 1/n of all seniors die each year.

maturing = Children / YEARS AS A CHILD
Units: person/year
If people live n years as a child, then on average 1/n of all children will mature to adulthood each year.

number of childbearing couples = Childbearing Adults / ADULTS PER COUPLE
Units: couple

Number of couples who are able to have children.
Seniors = INTEG (aging - deaths, 1000)
Units: person

Number of people between the ages of 45 and 75.
total population = Children + Childbearing Adults + Seniors
Units: person

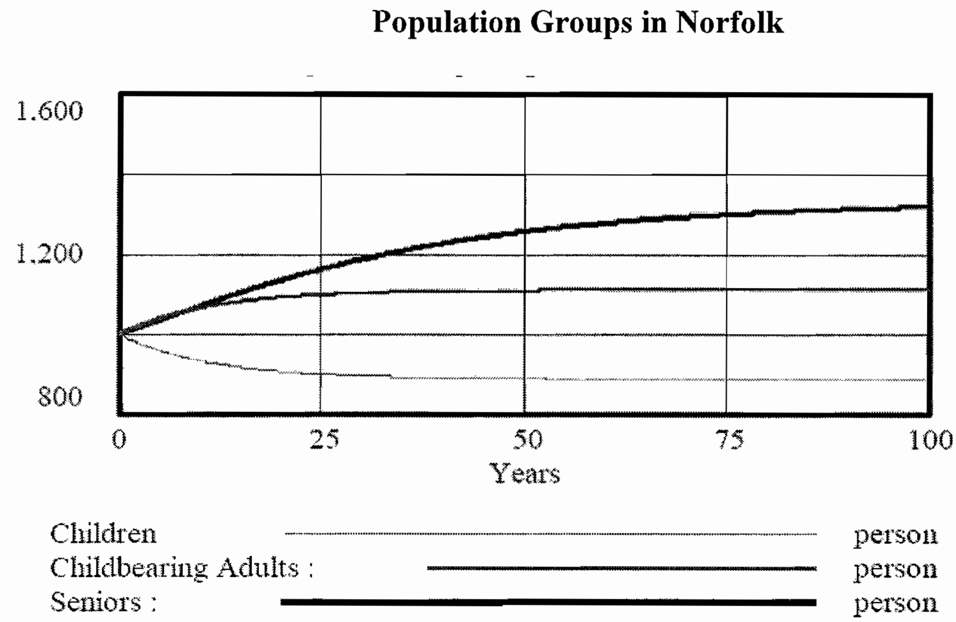
The total population is the sum of all three population groups.
YEARS AS A CHILD = 20
Units: year

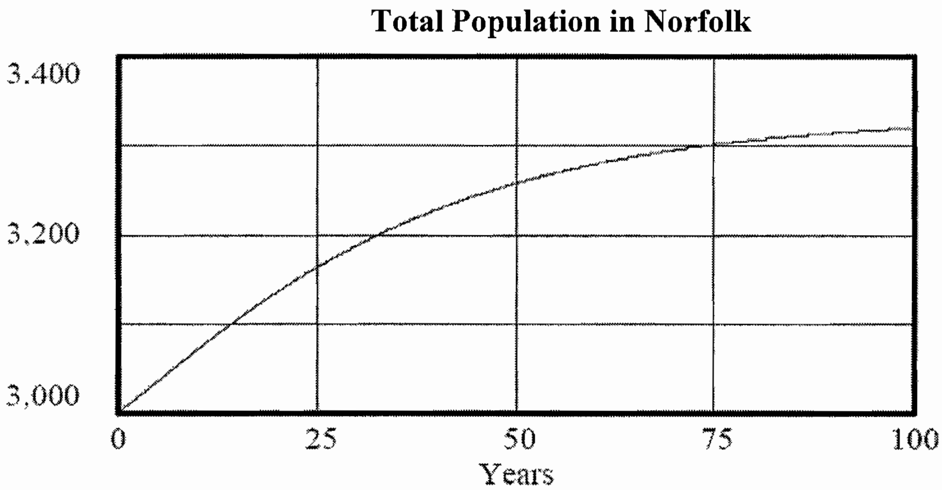
The number of years a person is a child. Assuming that a person is a child from birth until the age of 20, a person lives 20 years as a child.
YEARS AS A CHILDBEARING ADULT = 25
Units: year

The number of years during which people can bear children. Assume that people can bear children from when they are 20 years old until when they are 45 years old, over a period of 25 years.
YEARS AS A SENIOR = 30
Units: year

The number of years during which a person lives as a senior. Assume that on average, a person lives as a senior from age 45 until age 75, over a period of 30 years.

(iv) Model Behaviour





The behaviour of the model makes sense. Each stock starts with 1000 people, and according to the rate equations, during the first time step (DT=0.0625) the stocks behave as follows:

Children: Each person lives 20 years as a child, so 1/20 of the children (50 people initially) leave the stock per year. The stock is, however, being replenished by "births." Initially, there are 1000 childbearing adults, thus 500 childbearing couples. If each couple has 2 children over the course of 25 years, then each year a couple has 2/25 children. The total number of children born per year is 500 * 2 / 25 = 40 children. Because the inflow to the "Children" stock is 40 people while the outflow is 50 people, the net flow into the stock is –10 people per year, so the number of "Children" decreases by 0.0625 * 10 = 0.625 people during the first time step. The stock will keep decreasing until the inflow "births" equals the outflow "maturing."

Childbearing Adults: Each person lives 25 years as a childbearing adult, so 1/25 of the adults (40 people initially) leave the stock each year through the aging process. The stock is, however, replenished by 50 maturing children. The net flow into the stock is 10 people/year, so the number of "Childbearing Adults" increases by 0.0625 * 10 = 0.625 people during the first time step. The stock will continue to grow until the inflow "maturing" equals the outflow "aging."

Seniors: Each person lives 30 years as a senior, so 1/30 of the seniors (1000/30=33 people initially) leave the stock through "deaths". The stock of "Seniors," however, also gains 40 new people through the "aging" inflow. Therefore, the net flow into the stock is 6.67 people per year, so the number of "Seniors" increases by 0.0625 * 6.67 = 0.42 people during the first time step. The stock will continue to grow until the inflow "aging" equals the outflow "deaths."

total population: Add up the net increase/decrease into each stock to ensure total population increases by 0.42 people during the first time step. To find out whether the system will reach equilibrium or exhibit exponential behaviour, one must analyse the dynamics of the birth loop.

Notice that each couple, in its lifetime, gives birth to two children who essentially replace the couple in the population. The exact substitution gives the impression that the total population will remain unchanged. The population will only be unchanged, however, if the time delays for all stocks are the same. Because it takes 25 years for each couple's two children to enter the "Children" stock but only 20 years for children to mature, the net effect is that the number of maturing children will always be larger than the number of children born. The faster children become "Childbearing Adults," however, the more "Childbearing Adults" in the population, hence the greater the birth rate. Equilibrium is reached when the sum of inflows into each stock equals the sum of outflows from the stock:

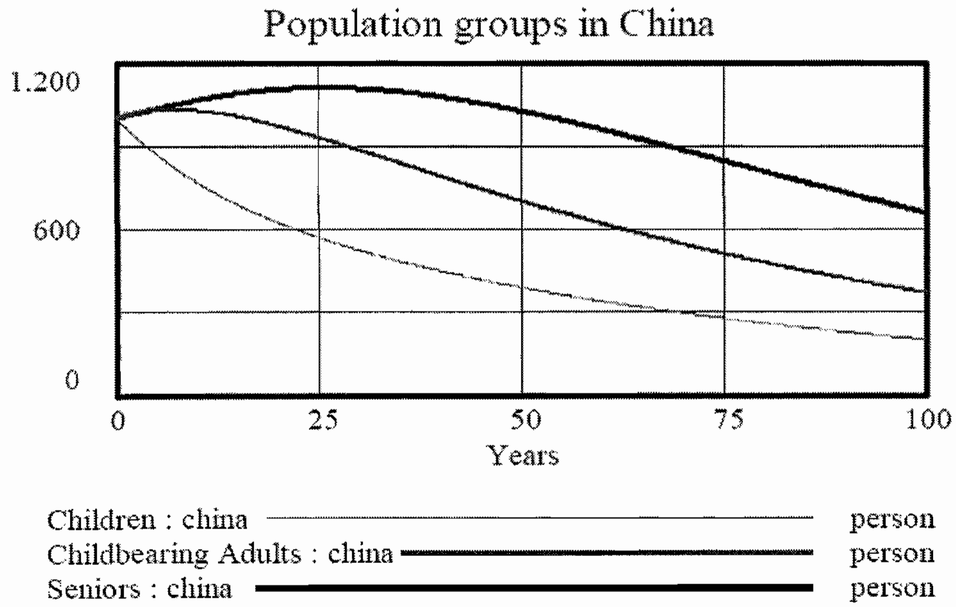
For the stock of "Children" to be in equilibrium:
births = maturing
number of childbearing couples * BABIES PER COUPLE / YEARS AS A CHILDBEARING ADULT = Children / YEARS AS A CHILD
(Childbearing Adults / 2)*2 / 25 = Children / 20
Childbearing Adults / 25 = Children / 20

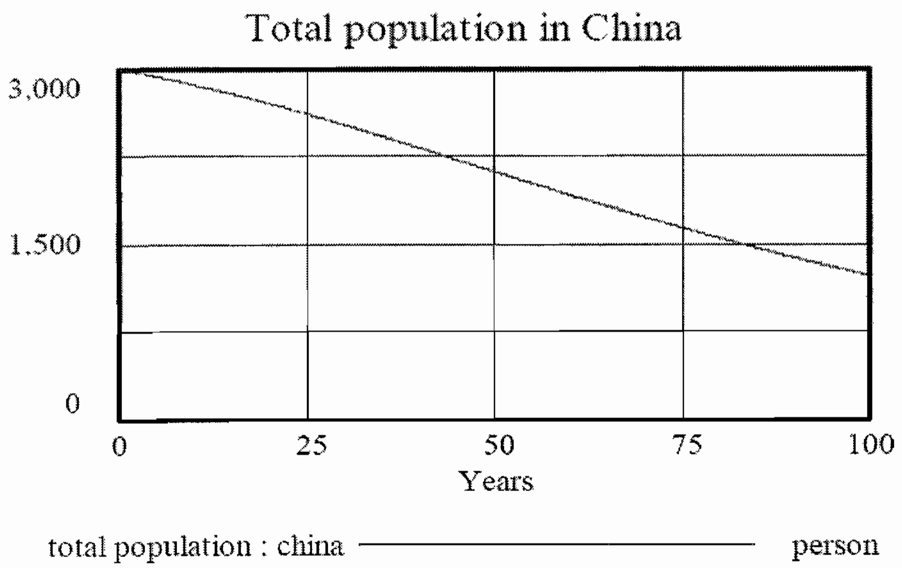
For the stock of "Childbearing Adults" to be in equilibrium:
maturing = aging
Children / YEARS AS A CHILD = Childbearing Adults / YEARS AS A CHILDBEARING ADULT
Children / 20 = Childbearing Adults / 25

For the stock of "Seniors" to be in equilibrium:
aging = deaths
Childbearing Adults / YEARS AS A CHILDBEARING ADULT = Seniors / YEARS AS A SENIOR
Childbearing Adults / 25 = Seniors / 30

The first two equilibrium equations give the same equality. The third equation is different but does not conflict with the first two. Equilibrium is reached when the system reaches the following proportions:
Children / 20 = Childbearing Adults / 25 = Seniors / 30

1 (b) Chinese Demographics





Let us examine the initial behaviour (during the first time step of 0.0625) of each stock in this scenario to find the equilibrium values of the system:

Children: The initial outflow is, as before, 50 people per year. The population contains

1000 “Childbearing Adults” forming 500 childbearing couples. Each couple has only one child in 25 years, or 1/25 children per year. The inflow of “births” is therefore 500 * 1 / 25 = 20 children/year. Thus the net flow in the stock is –30 people per year, so the number of “Children” decreases by 0.0625 * 30 = 1.875 people during the first time step.

The stocks “Childbearing Adults” and “Seniors” behave as before, changing by 0.625 and 0.42 people during the first time step, respectively. The change in the “total population” is –1.875 + 0.625 +0.42 = –0.83 people during the first time step.

Intuitively, one would think that because each couple is replaced by only one person, the stocks should be decreasing throughout the simulation, until the population of each stock reaches zero. The initial increase in the stocks “Childbearing Adults” and “Seniors” is caused by the delays in the system. That is, although the number of “Children” is decreasing quickly, the decrease does not affect the other stocks right away because of the delays. Only after the number of “Childbearing Adults” starts decreasing because of the decrease in “Children,” does the number of “Seniors” start decreasing. Also, as the delay for a stock becomes longer, the effect is felt later. Therefore, the “Seniors” stock peaks after the peak of the “Childbearing Adults” stock.

Using the stock equilibrium equations from before, one can find the relationships of the stocks at equilibrium and show that equilibrium can occur only when the all stocks equal zero:

For the stock of “Children” to be in equilibrium:
births = maturing
number of childbearing couples * BABIES PER COUPLE / YEARS AS A CHILDBEARING ADULT = Children / YEARS AS A CHILD
(Childbearing Adults / 2)*1 / 25 = Children / 20
Childbearing Adults / 50 = Children / 20

For the stock of “Childbearing Adults” to be in equilibrium:
maturing = aging
Children / (YEARS AS A CHILD)
= Childbearing Adults / (YEARS AS A CHILDBEARING ADULT)
Children/20=Childbearing Adults / 25

For the stock of “Seniors” to be in equilibrium:
aging = deaths
Childbearing Adults / (YEARS AS A CHILDBEARING ADULT)
= Seniors / (YEARS AS A SENIOR)
Childbearing Adults / 25 = Seniors / 30

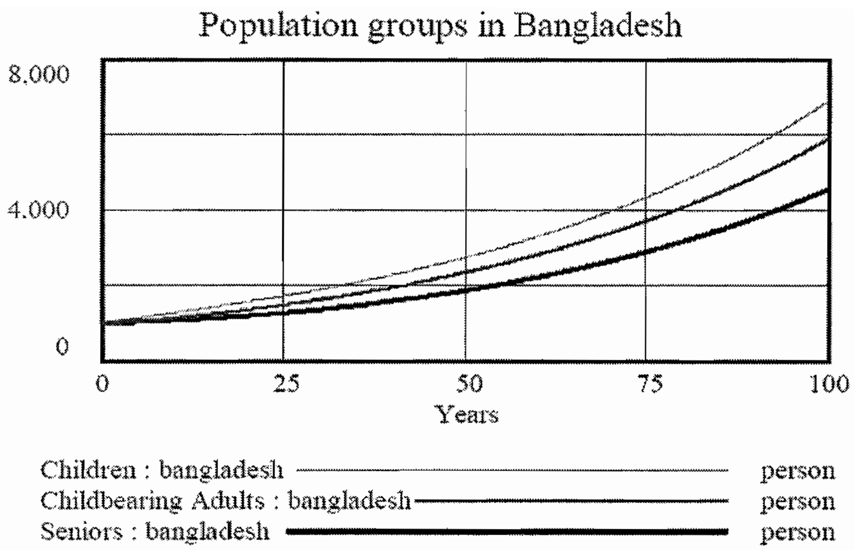
One can see that the equilibrium equations for “Children” and for “Childbearing Adults” cannot hold true at the same time. The equilibrium equation for “Children” states that for every 20 “Childbearing Adults” there are 50 “Children,” while the equilibrium equation for “Childbearing Adults” states that for every 20 “Childbearing Adults” there are 25 Children. Both equations only hold true when the value of both stocks is zero, which implies that the equilibrium value of the “Seniors” stock must also be zero.

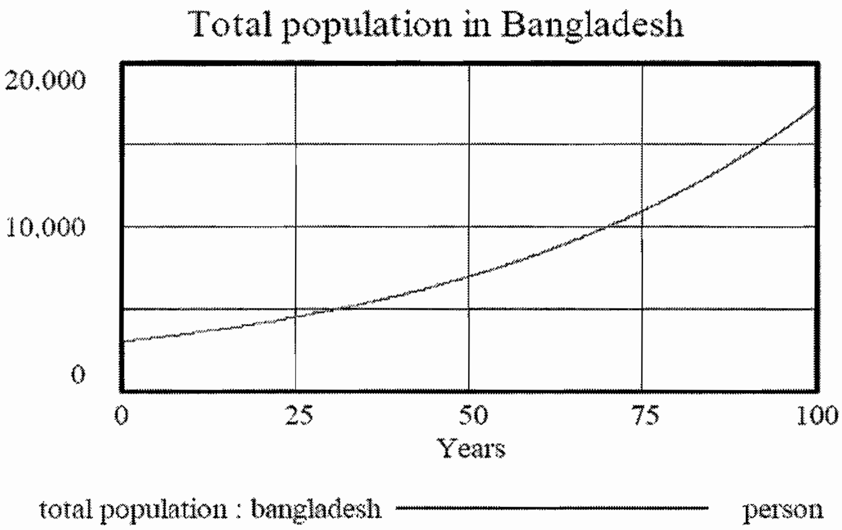
1 (c) <u>Model Behaviour in Bangladesh</u>

(i) Because the number of "BABIES PER COUPLE" is 4, it is easy to see that the system exhibits exponential growth. It is interesting to note, however, that the "Children" stock has the sharpest growth, followed by the "Childbearing Adults" stock, and then the "Seniors" stock. Again, this effect is produced by the delays in the system. (see graphs)

(ii) Many problems could be caused by such explosive population growth. Depletion of resources, overcrowding, and pollution are just some of them. These factors will slow down the exponential growth and cause the population to stabilise at some equilibrium point by increasing the death rate. Social pressure may cause the birth rate to decrease before these natural factors increase the death rate as a population control measure.

(iii) System dynamics tools can help determine important factors such as how many years it will take before the population halves, or what needs to be the number of "BABIES PER COUPLE" allowed in order to control the population by a certain date. Even the awareness of many of these counter-intuitive behaviours of relatively simple systems helps us to make better decisions because we know that such unexpected things can happen and we are more careful and consider non-obvious results. Implementing these policies, however, is another story.


Population groups in Bangladesh
8,000
4,000
0
0
25
50
75
100
Years
Children : bangladesh
person
Childbearing Adults : bangladesh
person
Seniors : bangladesh
person


Total population in Bangladesh
20,000
10,000
0
0
25
50
75
100
Years
total population : bangladesh
person

2 (a)

(i) The doubling time for a first-order positive-feedback system is 0.7 / time constant

The drug-user system fits nicely into a first-order positive-feedback generic structure. Therefore, one can calculate that:

doubling time for Adams = 0.7 / 0.2 per year = 3.5 years and
doubling time for Barton = 0.7 / 0.1 per year = 7 years

Notice, however, that the two schools do not start out with the same initial value of drug users. In fact, Barton's initial number of drug users is double that of Adams. The leader in 10 years will be based upon both the initial value and the doubling time.

Adams:

Years	0	3.5	7	10.5
Drug Users	200	400	800	1600

Barton:

Years	0	7	14
Drug Users	400	800	1600

In 7 years, both school districts will have 800 drug users. During the 3 years remaining before the tenth year, Barton can expect to gain fewer drug users than Adams. By year 10.5, Adams will see 1600 drug users, while Barton will see the same level only after 14 years. The doubling time of the number of drug users at Adams is shorter than the doubling time of the number of drug users at Barton, so the number of drug users at Adams is growing faster than the number of drug users at Barton. Hence, by the 10^{th} year, Adams will have many more drug users than Barton.

This can also be modeled using a first order linear rate equation:

$$dS/dt = rS$$

where S is the stock of drug users in Adams or Barton and f is the fractional influence rate. We solve for S,

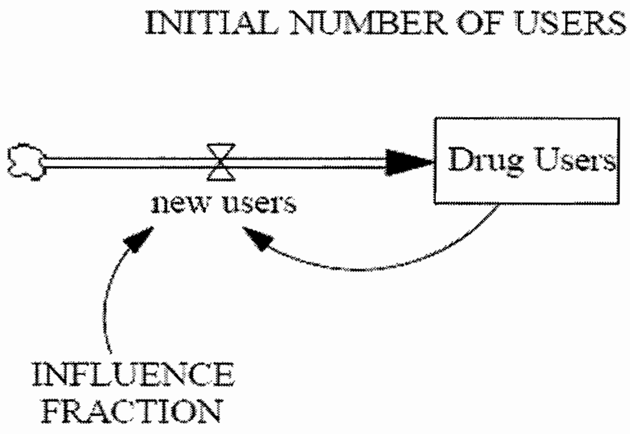
$$S = \exp(rt + C)$$

For Adams, r = 0.2 and at t = 0, S = 200, so after solving for C, we find

$$S_A = 200 * \exp(rt) = 200 * \exp(0.2t)$$
$$S_B = 400 * \exp(rt) = 400 * \exp(0.1t).$$

Therefore, $S_A(t{=}10) = 200 * \exp(2) = 1478$ and $S_B(t{=}10) = 400 * \exp(1) = 1087$

(ii) Model diagram:

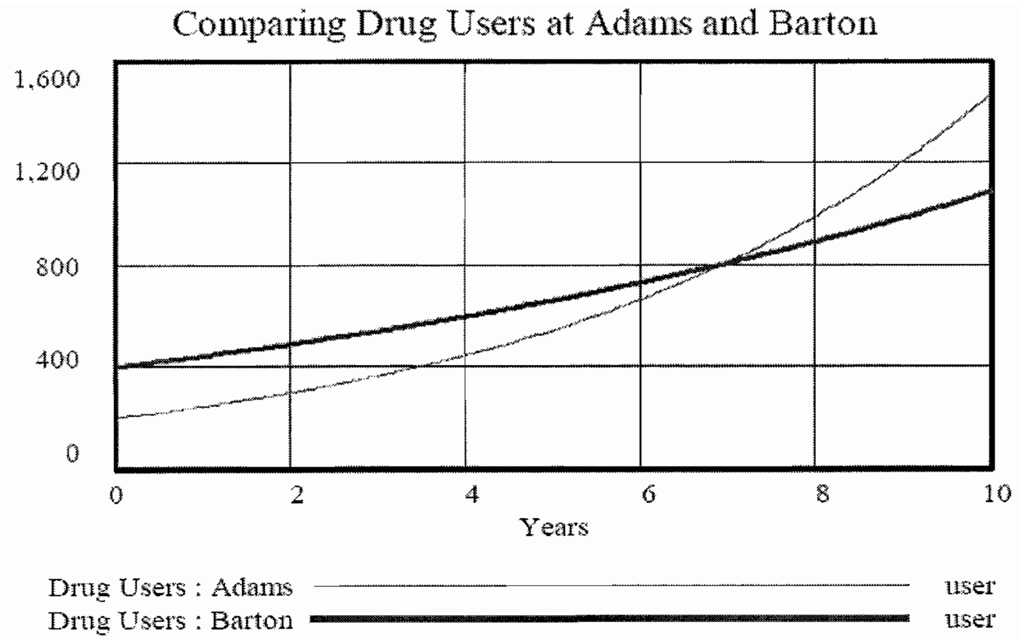


Model equations:
For Adams:
Drug Users = INTEG (new users, INITIAL NUMBER OF USERS)
Units: user
The number of drug users in the school district.

INFLUENCE FRACTION = 0.2
Units: 1/year
The amount of influence that current drug users have on converting others to drug use.
INITIAL NUMBER OF USERS = 200
Units: user

The initial number of drug users in the school district.
new users = Drug Users * INFLUENCE FRACTION
Units: user/year
The number of students who become drug users each year.

For Barton, the "INITIAL NUMBER OF USERS" is 400 users and the "INFLUENCE FRACTION" is 0.1 per year.
Simulated over ten years, the model generates the following behaviours for the two school districts:



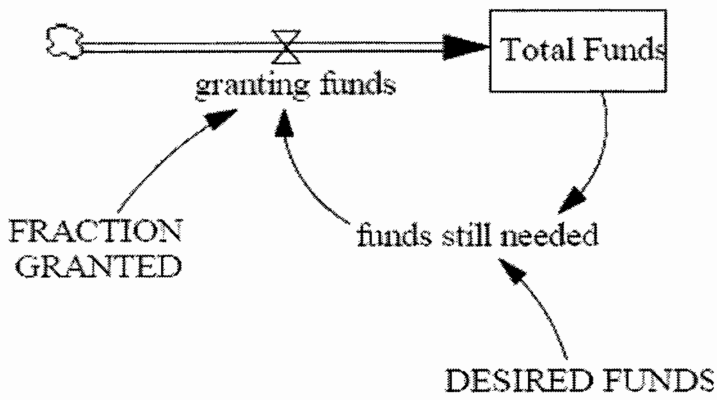
The graph of stock behaviour confirms the calculations from part (i): the number of “Drug Users” at Adams surpasses the number of Barton “Drug Users” around year 7, so Adams has more “Drug Users” in ten years.

2 (b)

(i) Please note that we are again assuming that the events in the system are continuous and not discrete. The project does not receive a bundle of money at the beginning of each quarter, but receives a continuous flow of money, as the Budget Office pays, for example, salaries to the construction workers after every instant of work.

The system fits the generic structure of a goal-seeking negative-feedback system. One can model the system in two different ways. The first model uses an inflow goal-gap structure:

Model diagram:



(ii) Model equations:

DESIRED FUNDS = 1e+006
Units: pound
The total amount of money needed for the project.

FRACTION GRANTED = 0.5
Units: 1/Quarter
The fraction of the needed funds that is granted by the Office of Management and Budget per quarter.

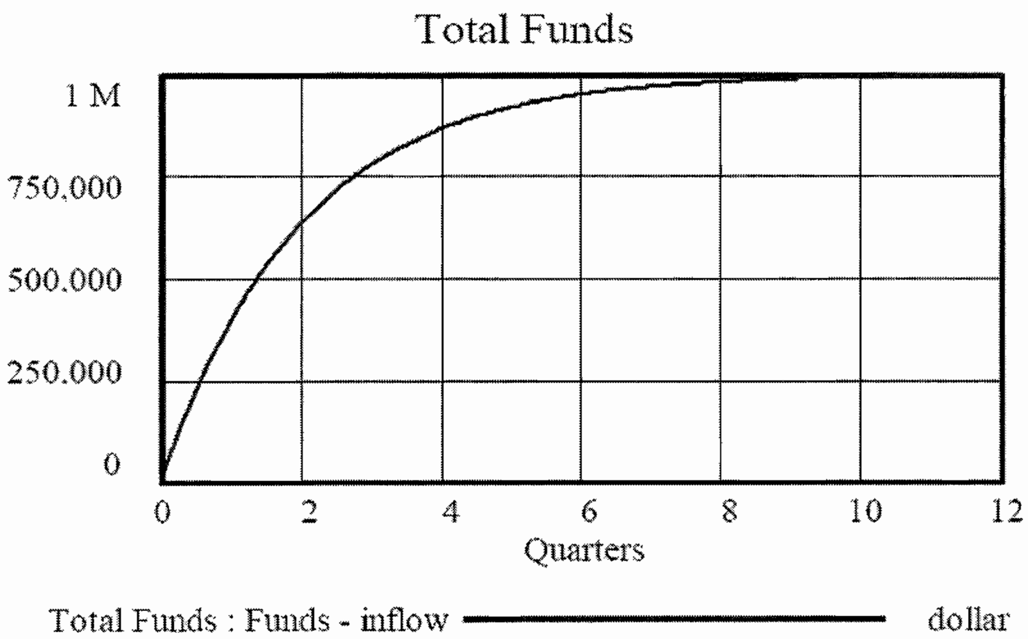
funds still needed = DESIRED FUNDS - Total Funds
Units: pound
The difference between desired and current funds for the project is the amount that still remains to be collected.

granting funds = funds still needed * FRACTION GRANTED
Units: pound/Quarter
The amount of money the Office of Management and Budget grants per quarter.

Total Funds = INTEG granting funds, 0)
Units: pound

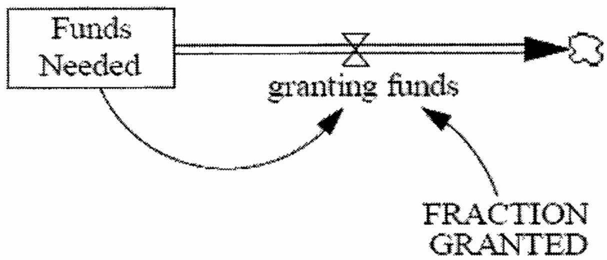
The amount of money that has been granted for the project at any given time.

Model behaviour:



2 (b) Alternatively, one could model the system with an outflow decay structure:

(i) Model diagram:



(ii) Model equations:

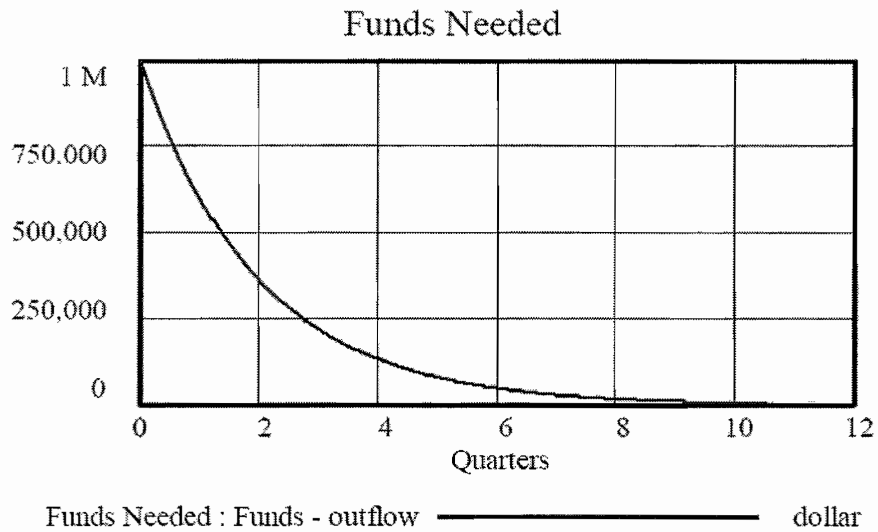
FRACTION GRANTED = 0.5
Units: 1/Quarter
The fraction of the needed funds that is granted by the Office of Management and Budget per quarter.

Funds Needed = INTEG (-granting funds, 1e+006)
Units: pound

The amount of funds that still needs to be collected.

granting funds = Funds Needed * FRACTION GRANTED
Units: pound/Quarter
The amount of money the Office of Management and Budget grants per quarter.

Model behaviour:



(iii) The time constant of the system is 1 / 0.5 = 2 quarters, so the half-life is 1.4 quarters. Ninety-five percent of the goal of "DESIRED FUNDS" is met when "Total Funds" reach £950 000 and "funds still needed" fall to £50 000. This point occurs after 3 time constants, or after 6 quarters. You can either obtain this value exactly by using exponentials arithmetic or by approximating from the behaviour graph. If you make a table, pinpoint the time at which "Total Funds" passes the critical £950 000 mark to find out how many fiscal quarters have passed.

In terms of exponentials,

$$0.95 = 1 - \exp(-0.5t)$$
$$\exp(-0.5t) = 0.05$$
$$-0.5t = -2.996$$
$$T=5.99$$

(iv) With the anonymous gift, the time constant is still equal to 2 quarters, so the half-life does not change either. If a gift of £500 000 is received at time = 0, then the initial value of "Total Funds," as well as the initial value of "Funds Needed," is £500 000. When there was no gift, it took the system one half-life to reach £500 000. The anonymous gift thus cuts one half-life from the total amount of time that will pass before 95% of the funds have been granted. The number of quarters it takes to reach 95% of the goal is then 6 quarters – 1.4 quarters = 4.6 quarters.

Alternatively, can write: 0.95 = 1 – 0.5*exp(-0.5t) and solve for t=4.61.

3. a i) $s^2 = \frac{\left((N_A - 1)s_A^2 + (N_B - 1)s_B^2\right)}{N_A + N_B - 2}$

	variance	df	Sum of squares
Sample A	8.7	49	426.3
Sample B	9.1	49	445.9
Pooled sample		98	872.2

Pooled variance $s^2 = 872.2/98 = 8.9$

ii) H0: $\mu_A - \mu_B = 0$
H1: $\mu_A - \mu_B \neq 0$

test statistic is $t = (\mu_A - \mu_B)/\sqrt{s^2(1/N_A + 1/N_B)}$

under H0 this statistic has a t-distribution with $(N_A + N_B - 2)$ degrees of freedom

iii) Type I error: When the null hypothesis is true and we decide to reject it.
Type II error: When the null hypothesis is false and we decide accept it.

iv) $\mu_A - \mu_B = 62.6 - 64.2 = -1.6$

$\sqrt{s^2(1/N_A + 1/N_B)} = 0.5967$

$t = -1.6/0.5967 = -2.682$

$N_A + N_B - 2 = 98$: the 5% critical value for a t-test with 98 degrees of freedom is 1.984

Because $|-2.682| > 1.984$

Hence we can reject H0 and conclude that $\mu_A - \mu_B \neq 0$

v) The computed probability value using a two-tailed t-distribution is 0.009

b i) $\Sigma d_i^2 = 54$

$$r = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)} = 1 - \frac{6(54)}{10(10^2-1)} = .67$$

ii) $H_0 : p_r \leq 0$

$H_a : p_r > 0$

$$\sigma_{r_s} = \sqrt{\frac{1}{n-1}} = \sqrt{\frac{1}{10-1}} = .33$$

$$z = \frac{r_s - \mu_{r_s}}{\sigma_{r_s}} = \frac{.67 - 0}{.33} = 2.02$$

p-value = 2(1.0000 - .9783) = .0434

iii) p-value $\leq$ .05, reject H_0. Conclude a significant positive rank correlation.

iv) An often used approach to test the impact of one variable on another variable, eg. the impact of training on labour productivity, is regression analysis. For example, we can build up a model that labour productivity is a function of training and other explanatory variables. Labour productivity is the dependent variable in the model, and training will be one of the explanatory variables. The other explanatory variables in the model will include other major factors that theory suggests to have a significant effect on productivity, eg., firms size, innovation, competition, etc.

Regression analysis and correlation test have some fundamental differences. The primary objective of correlation analysis is to measure the strength or degree of linear association between two variables. In regression analysis, we try to estimate or predict the average value of one variable on the basis of the fixed values of other variables.
In regression analysis, there is an asymmetry in the way the dependent and explanatory variables are treated. The dependent variable is assumed to be statistical, random; while explanatory variable are assumed to have fixed values. In correlation analysis, we treat any two variables symmetrically; there is no distinction between the dependent and the independent variables. Both variables are assumed to be random.

To use regression analysis in current problem setting will suffer from problems such as small sample size, omitted variable problems.

4. (i) (a) $b_1 = t \times \text{se}(b_1) = 1.257 \times 0.2174 = 0.2732$

(b) $se(b_2) = b_2 / t = 0.1801/5.754 = 0.0313$

(ii) The estimated coefficient of log(L) is 0.18 indicates that a 1% increase in number of employees increases average turnover of the firm by 0.18%. The positive sign is as expected; more labour inputs should lead to greater output.

The estimated coefficient of log(K) is 0.81 indicates that a 1% increase in capital stock increases average turnover of the firm by 0.81%. The positive sign is as expected; more capital inputs should lead to greater output.

The estimated coefficient of RD is 0.11 indicates that the turnover of firms that invest in R&D is 1.11 units (EXP(0.1069 = 1.1128) higher than those who do not. The positive sign is as expected; investment in R&D is expected to raise productivity of firms and therefore lead to greater output.

(iii) A 99% confidence interval for the slope is given by

$$b_2 \pm t_c \text{se}(b_2) = 0.1801 \pm 2.68 \times 0.0313 = (0.096, 0.264)$$

(iv) $H_0 : \beta_2 = 0.2$

$H_1 : \beta_2 \neq 0.2$,

For testing the above hypothesis,

we calculate $t = (0.1801 - 0.2)/0.0313 = -0.634$.

The critical values for a two-tailed test with a 5% significance level and 49 degrees of freedom are $\pm t_c = \pm 2.01$.

Since $t = -0.634$ lies in the interval (–2.01, 2.01), we do not reject H_0.

The null hypothesis suggests that a 1% increase in number of employees leads to an increase in average turnover of 0.2%. Non-rejection of H_0 means that this claim is compatible with the sample of data.

(v) The assumptions about the error term in the regression model:

- the error term ε is a random variable with a mean of expected value of zero; $E(\varepsilon) = 0$.
- the variance of ε is the same for all values of x.
- the values of ε are independent.
- The error term ε is a normally distributed random variable.

(vi)

Given SSR = $\Sigma(e_t)^2$, so SSR here refers to residual sum of squares, or sum of squares due to residual (error).

SST = SSE + SSR

SSE = SST – SSR = 0.52876 – 0.14632 = 0.38244

SSE here refers to explained sum of squares.

R^2 = SSE/SST = 0.38244 / 0.52876 = 0.7233

R^2 equals 0.7233 means that 72.33% of the variation in the dependent variable is explained by the variables log(L), log(K) and RD.

(vii) The firm AAA residual is

$\hat{e}_0$ = log(310) – 0.2732 – 0.1801 × log(10)-0.8086 x log(200)-0.1069 x 1 = 0.0705

(viii) The prediction is

Log(SALES) = 0.2732 + 0.1801 × log(75)+0.8086 x log(125)+0.1069 x 0 = 2.3065

SALES = 10^(2.306) = 202.5163

(ix) The current model is based on the widely used Cobb-Douglas production function while takes into account the impact of innovation on productivity.
The potential problems of the current model are that:
First, according to economic theory, in addition to labour and capital inputs and innovation, there are other factors may influence the output of firms, for example, labour force skills and management practices of the firms. Moreover, firms in different industries may use different technology. Some treatment needs to be done to control for the industry specific effect, eg. including some industries dummies into the regression model. Missing of all these variables may lead to the omitted variable problem.
Second, while invest in R&D may lead to greater output, firms with greater turnover may more likely to invest in R&D. So there may be endogeneity between the output and R&D variable.
Finally, one limitation of using cross-section datasets is that it fails to capture the dynamics in the input-output, R&D and output relationship.