Activity description

This activity is about using graphical methods to find a suitable model to connect two sets of data.

There are a variety of ways in which you could use these data, depending on whether you want to use a log graph to find an exponential model (the intended focus of this activity) or whether you prefer to use the data more generally to find and discuss a wider selection of models.

Students can use their model(s) to make predictions. They should also be encouraged to work out percentage errors to test and compare the accuracy of their model(s).

Suitability

Level 3 (Advanced)

Time

1–3 hours

Resources

Student information sheet, slideshow, spreadsheet

Equipment

Graph paper, rulers
Optional: computers with Excel, graphic calculators, internet access

Key mathematical language

Functions, exponential, logarithm, mathematical models

Notes on the activity

The first four slides can be used to introduce the activity. These slides, and page 1 of the student information sheet, give the number of cars with engines of 2 litres and above which were licensed in Great Britain at the end of each year from 1994 to 2007. Students are asked to suggest possible types of functions that could be used to model the data. As well as exponential functions, students might suggest linear, quadratic or other polynomial functions. At this stage you could let students find and check a range of such functions.

If you prefer to concentrate on the exponential model, page 2 of the student information sheet shows how to use natural logarithms to find such a model from a linear graph. Students can work through the 'To answer' section to complete the working, then use percentage errors and/or a graph to check how good the model is. More recent data are given near the end of this section, and students are asked to compare these with predictions before finding a new model.

The method for finding an exponential model is also given on slide 5 of the slideshow. Later slides give the solutions to the 'To answer' section – these can be used as a demonstration or an on-going check of students' work.

During the activity

Students can work individually or in pairs or small groups to share the work. The activity can be completed using graph paper and calculators, spreadsheets or graphic calculators. If possible, use a combination of these methods, so that comparisons can be made and discussed within the group.

Points for discussion

Students should be asked for suggestions of possible functions to use to model the data. They should also be encouraged to discuss how they could check how well their model fits the data.

Students can be asked to think about:

- why the percentage error is a better measure of accuracy than the difference between the actual value and the value predicted by the model;
- what is indicated by a negative percentage error;
- whether their model is valid for all values of t.

Extensions

If you prefer to use the data for a general discussion about models, students could be asked to use a graphic calculator or spreadsheet to find one or more models and compare them. Advise students to use t to represent the number of years after 1994 – this gives more user-friendly coefficients.

The following functions were obtained from a graphic calculator:

$N = 172t + 1362$ ($r = 0.991$), $N = 5.91t^2 + 94.9t + 1516$,
$N = -0.0688t^4 + 1.32t^3 + 0.499t^2 + 87.6t + 1540$, $N = 1507e^{0.0704t}$ ($r = 0.999$)

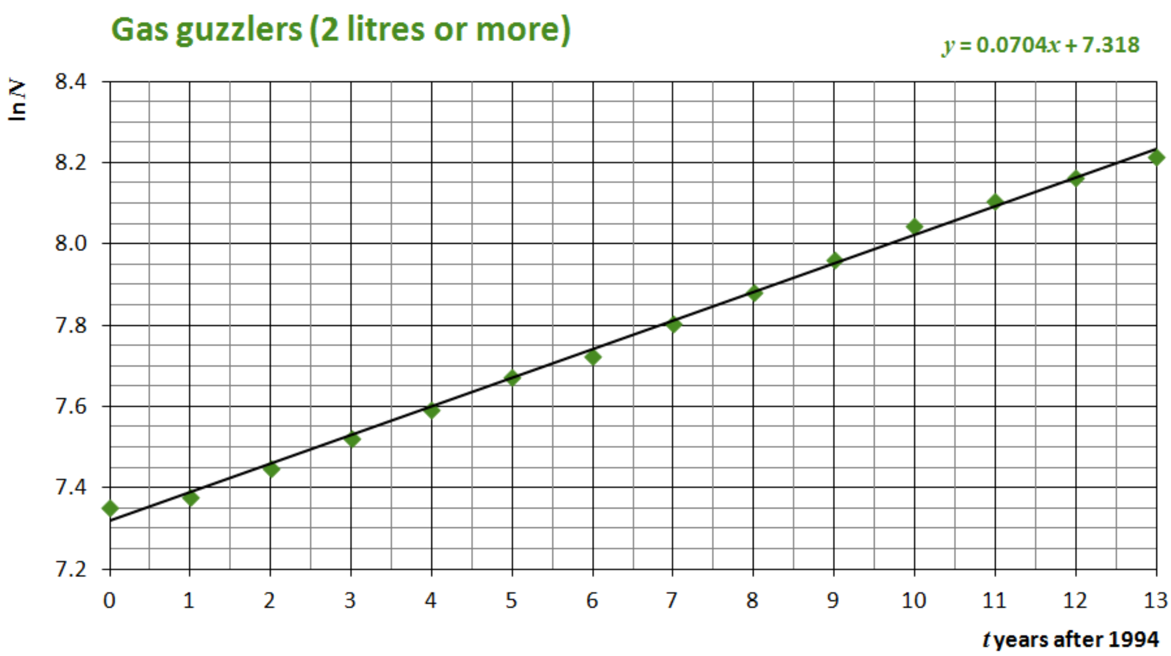
Students could compare the accuracy of the models they find using graphs or percentage errors.

All the above models agree reasonably well with the data.

Answers to questions

1 to 3 The completed table and log graph are shown below, with the line of best fit and its equation.

t	N	ln N
0	1558	7.351158
1	1600	7.377759
2	1715	7.447168
3	1844	7.519692
4	1980	7.590852
5	2145	7.670895
6	2262	7.724005
7	2451	7.804251
8	2647	7.881182
9	2869	7.961719
10	3118	8.044947
11	3314	8.105911
12	3512	8.163941
13	3687	8.212568



4 The gradient of the line of best fit gives $k = 0.0704$

The intercept gives ln $N_0 = 7.318$ and so $N_0 = e^{7.318} = 1507$

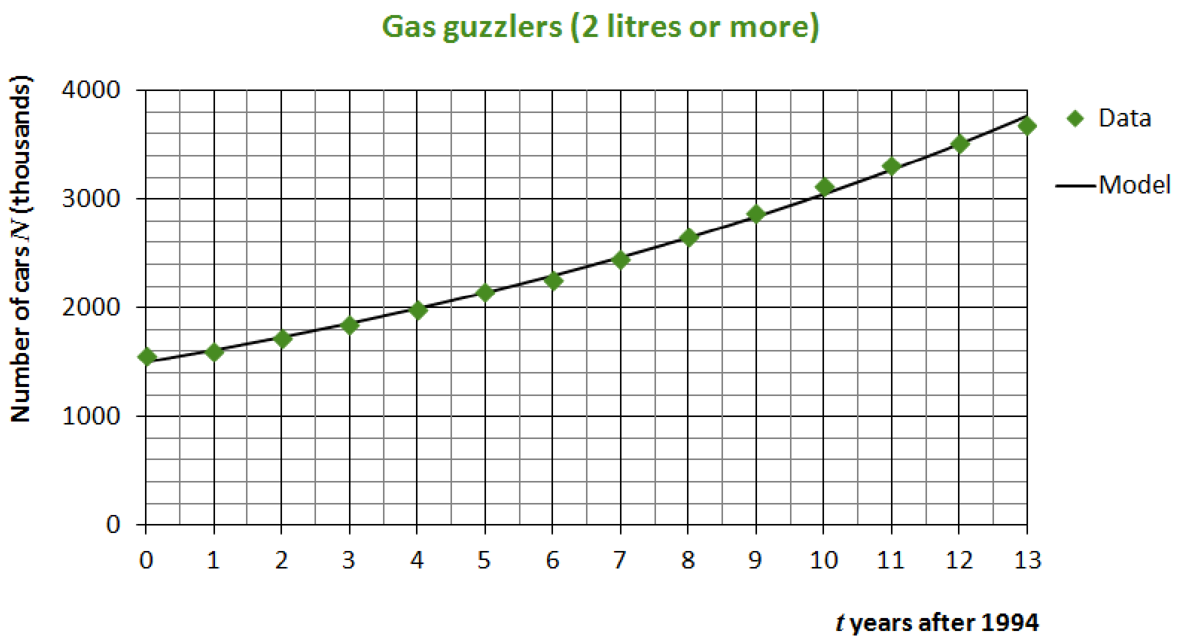
The exponential model is $N = 1507e^{0.0704t}$

5 Comparison of data and values suggested by the model

a The table gives the percentage error in the value suggested by the model for each data value.

Year	Data	Model	% Error
0	1558	1507	–3.2%
1	1600	1617	1.1%
2	1715	1735	1.2%
3	1844	1861	0.9%
4	1980	1997	0.9%
5	2145	2143	–0.1%
6	2262	2299	1.6%
7	2451	2467	0.7%
8	2647	2647	0.0%
9	2869	2840	–1.0%
10	3118	3047	–2.3%
11	3314	3269	–1.4%
12	3512	3508	–0.1%
13	3687	3763	–2.1%

b The graph below also shows how well the model fits the data.



6 Predictions for 2008 and 2009

Substituting $t = 14$ in $N = 1507e^{0.0704t}$ gives
$N = 1507e^{0.0704 \times 14} = 1507e^{0.9856} = 4037$ thousand

Substituting $t = 15$ in $N = 1507e^{0.0704t}$ gives
$N = 1507e^{0.0704 \times 15} = 1507e^{1.056} = 4332$ thousand

7 These predictions are much greater than the actual values:

3731 thousand and 3768 thousand.

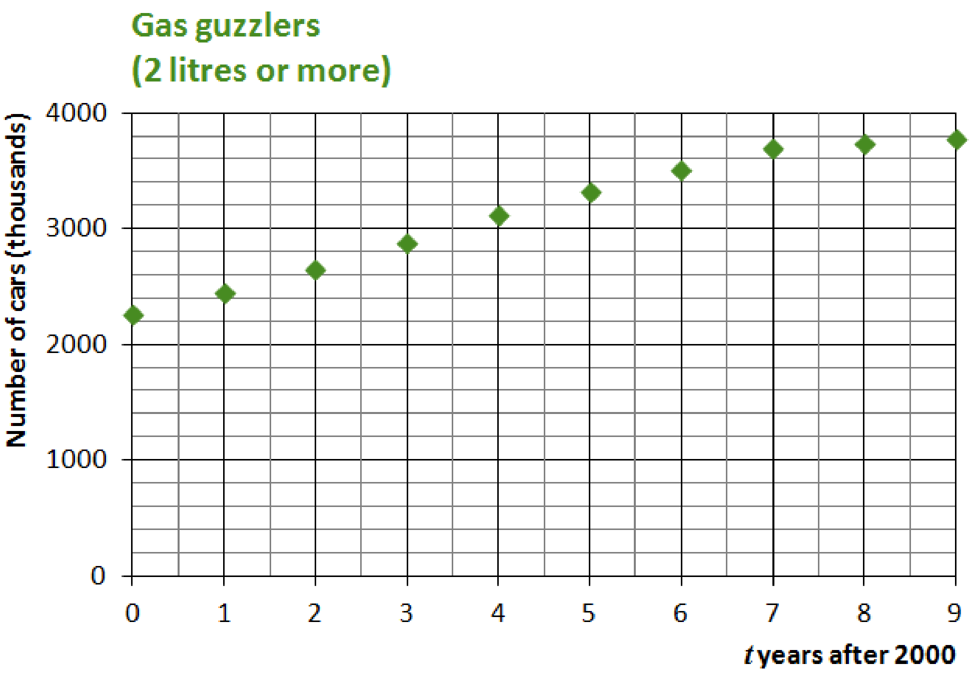
Possible reasons for this are:

- the level of demand for cars with large engines did not increase as much as predicted, because they consume a lot of petrol and are very costly to run
- measures taken by the government to control global warming may discourage their purchase.

In general, this shows that extrapolating beyond the data used to find the model may give a poor prediction, and that it might be advisable to find a new model.

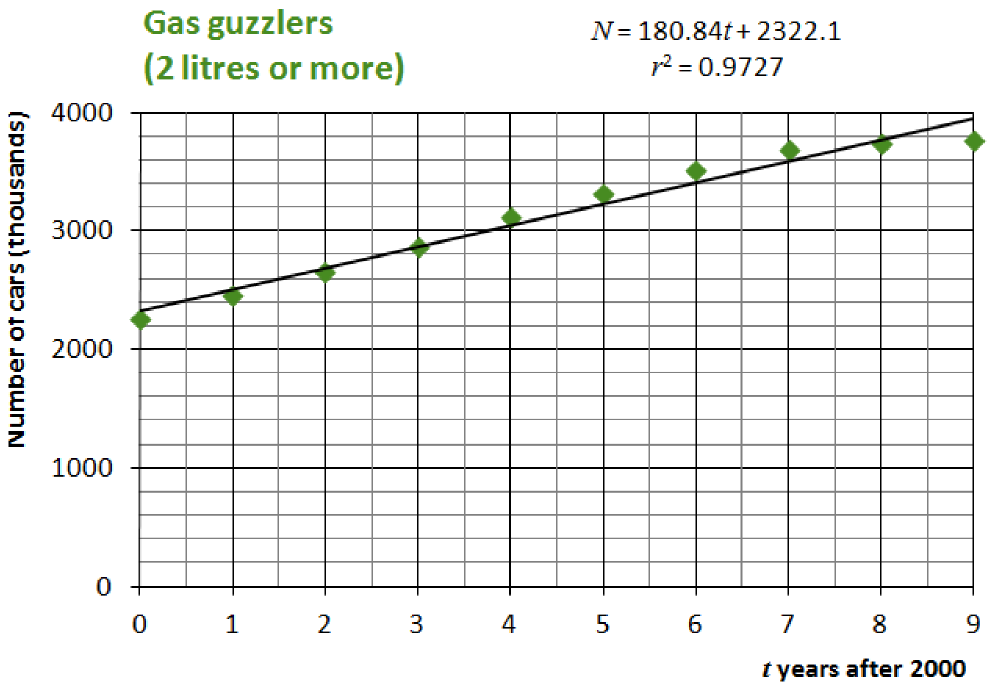
The final request on the worksheet is to use the data for 2000–2009 to find a new model. The graph for 2000–2009 and possible models are given below.

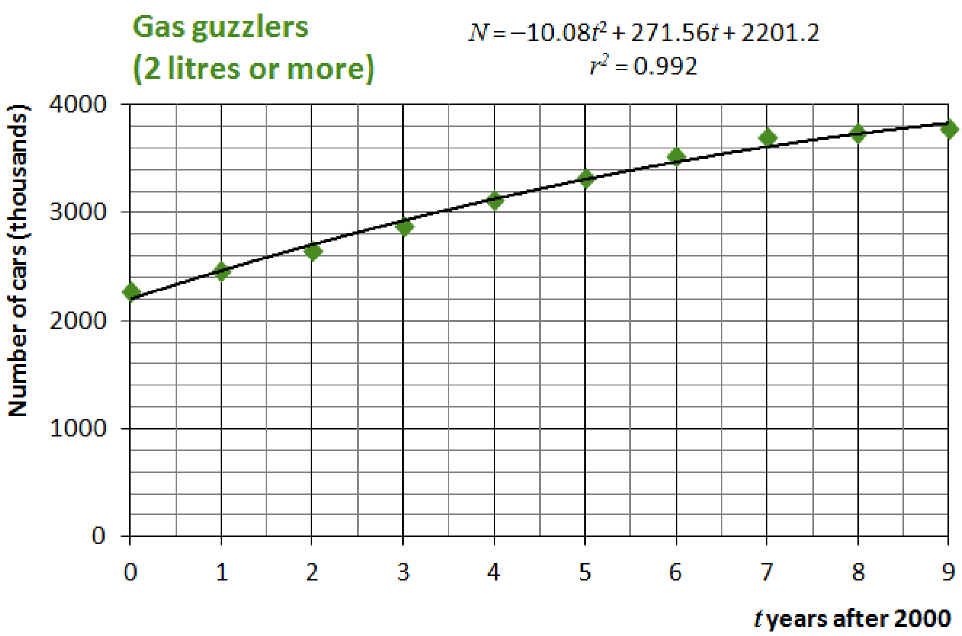
8 Graph for 2000 to 2009 data

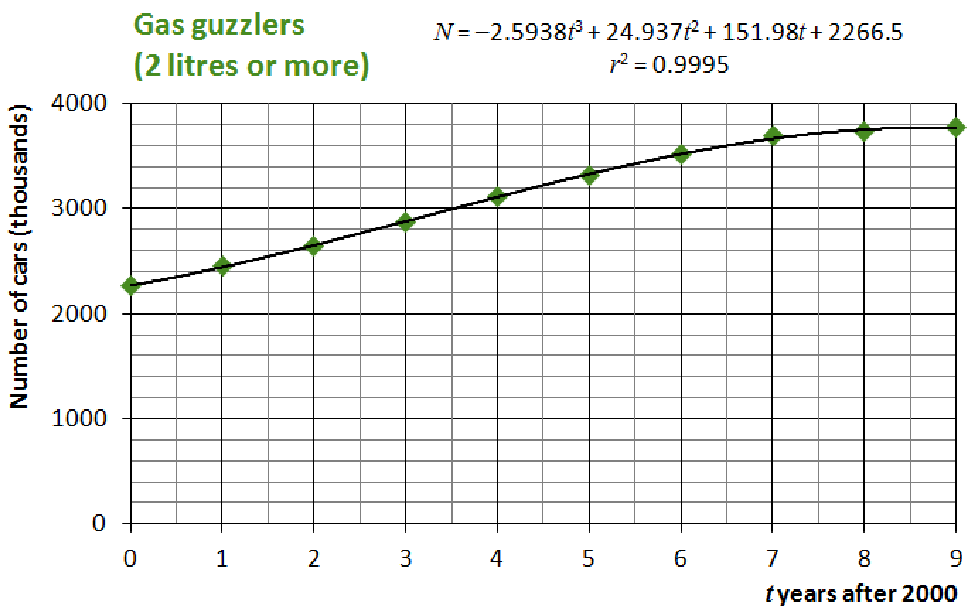


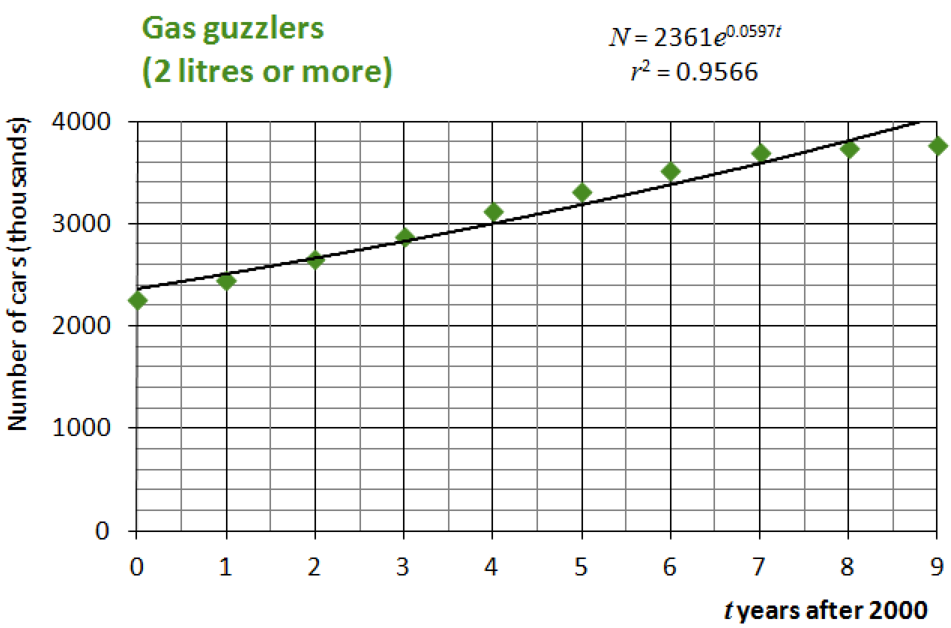
Possible models

The graphs below show models suggested by Excel. The cubic function is the best of these.









The following functions, which are very similar to those above, were obtained from a graphic calculator.

$N = 183t + 2310$ ($r = 0.986$)

$N = -10.1t^2 + 272t + 2201$

$N = -2.59t^3 + 24.9t^2 + 152t + 2267$

$N = 2361e^{0.05975t}$ ($r = 0.978$)