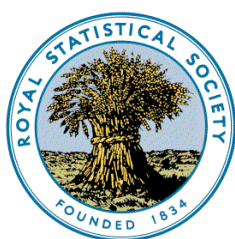


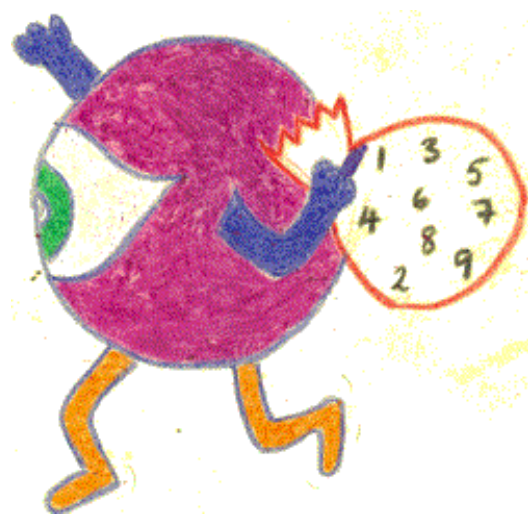
Relevant & Engaging

Statistics & Data Handling

**Using Autograph with
CensusAtSchool Data**



Centre for
Statistical Education



Chapter 4

Second Edition: published in the UK in 2009 by

The Royal Statistical Society Centre for Statistical Education
The University of Plymouth
Plymouth
PL4 8AA
UK

© 2009 The Royal Statistical Society Centre for Statistical Education

All rights reserved. No part of this booklet may be reprinted or reproduced or utilised in any form or by any electronic, mechanical or other means, now known or hereafter invented, including photocopying and recording, or in any information storage and retrieval system without permission in writing from the copyright holders, or a licence permitting restricted copying.

The advice and information in this booklet are believed to be true and accurate at the date of printing, but neither the authors, nor the publisher can accept any legal responsibility or liability for errors or omissions.

We would like to thank the many contributors to this book. These include: Doreen Connor, Neville Davies, John Marriott, Alan Catley (Chapter 4) and Mark Crowley (Chapter 6) and those who reviewed and helped in it's production including Alison Davies, Claire Webster, Peter Holmes & Elizabeth Gibson.

This booklet is aimed at all secondary level teachers: there are hints and tips that we hope will be useful to support the teaching and learning of statistics and data handling. We hope that you find the material useful. Please email or write to us with suggestions for improvements. We will try to respond to all communications.

Doreen Connor

The Royal Statistical Society Centre for Statistical Education
The University of Plymouth
2009

email info@censusatschool.org.uk

www.censusatschool.org.uk

www.rsscse.org.uk

Chapter 4

Using Autograph with CensusAtSchool Data

This chapter presents personal experiences from Alan Catley a teacher and ICT expert.

Teaching and learning data handling and statistics makes for a far more meaningful experience if learners are involved in working with real data. Incorporating resources such as *CensusAtSchool* into the teaching and learning of statistical topics at Key Stages 2, 3, GCSE (both tiers) and also in S1 at A-level can motivate learning and enhance success whilst having fun!

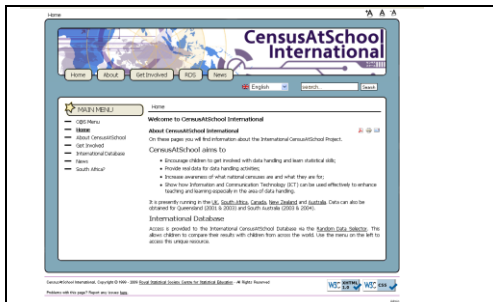


Fig 1. *CensusAtSchool* is 'multi-national'.



Fig 2. Access to all documents and the online questionnaire is very easy.

The starting point for all my classes in recent years has been getting learners to complete the current phase *CensusAtSchool* questionnaire (see figures 1 and 2). With Level 1 learners, completing a paper copy of this was the focus of a complete lesson, which involved accuracy of measuring and double-checking recorded values in an attempt to eliminate nonsense data. With a group of S1 learners, completing the questionnaire was a 10-minute job at the end of a pure maths lesson prior to commencing the statistics module.

Once learners had completed the paper copy they completed the online version of the questionnaire. There was no need for me to use lesson time to do this, as all learners have internet access (if not at home or via school/college then they could use the public library). Once the whole group had completed the online questionnaire, it only required a quick email to *CensusAtSchool* to retrieve the spreadsheet containing all the data responses for an individual class. This data can then be used to cover a large amount of the syllabus by discussing appropriate hypotheses with the class. This has proved to be a lot better than teaching, for example, scatter graphs and correlation using fictitious textbook examples. To add further interest, we looked at random samples, requested and retrieved from the *CensusAtSchool* database, drawn from different areas of the country. This provoked much interest by comparing these to their own responses, using statistical techniques relevant to the course being followed.

Phase 7 *CensusAtSchool* Questionnaire For Pupils Aged 11 plus

1 State the first part of your postcode (e.g. LL1 or LL24)

2 Gender
 Male Female

3 a) Please state your age in completed years
 _____ Years

4 Complete the following measurements.
 HEIGHT _____ cm
 FOOT LENGTH _____ cm
 Belly button to floor _____ cm

8 How do you usually travel to school?
 Walk Car
 Cycle Other
 Bicyc Other

9 In the last week, approximately how much time did you spend, to the nearest hour, on each of the following activities?

Activity	Hours
Playing computer games	
Reading (or newspapers)	
Watching TV (films)	
Playing/listening to music	
Playing board/card games	
Doing homework	
Doing jobs at home	
Working for pay	

12 Over the next few years, do you expect the following to get better, worse or stay the same?

Crime
 Better Worse Same

NHS
 Better Worse Same

Quality of education
 Better Worse Same

The way your local area is policed
 Better Worse Same

Public transport
 Better Worse Same

Environmental quality
 Better Worse Same

Politics/news
 Better Worse Same

13 Think about someone you most look up to. This could be someone you know personally, or

Fig 3. Part of the questionnaire

	A	B	C	D	E	F	G	H	I
1	yeargroup	region	gender	age	height	footLength	wristCirc	thumbCirc	bellyButton
2	Year 9	West Midlands	F	13	153	23	160	50	97
3	Year 9	South	F	13	161	22.5	150	53	98
4	Year 8	South	M	12	160	32	150	60	80
5	Year 9	South East	F	13	160	24	155	62	92
6	Year 9	North West	F	14	171	23	150	60	103
7	Year 9	South East	F	14	160	23	160	80	99
8	Year 9	East Midlands	F	13	90	32	120	20	50
9	Year 9	North West	F	13	148	22	135	60	100
10	Year 10	West Midlands	M	14	151	24	150	50	97
11	Year 9	East Midlands	F	13	165	27	180	65	135
12	Year 8	South East	M	13	200	23	190	23	50
13	Year 9	East Midlands	F	14	159	23	136	53	93
14	Year 9	West Midlands	F	13	120	24	129	47	70
15	Year 9	North West	F	13	160	22	150	50	102
16	Year 7	South	M	11	152	22	160	40.5	90

Fig 4. A section of a random data sample

Once you have your class's responses to the questionnaire you have a wealth of possibilities for the teaching of topics such as scatter graphs and correlation, box plots, cumulative frequency, mean and median calculation, etc. Using real data gives a great deal more meaning for the learners than ploughing through textbook examples and exercises. I will attempt to give a few examples of how this can be done. Using such data enabled the teacher to cover a lot of the S1 syllabus with great efficiency as well as helping level 1 learners to appreciate what handling data is all about.

The interactive use of Autograph through the classroom 'data projector'

Example 1 – working with single-variable data


Start by opening all 3 applications – Autograph/Excel/Word


Collect your data in Excel and then highlight the column required

(eg the height column in Fig 4)

Copy this column (use the top cell in the column to name the data set)

Switch to the Autograph page and proceed as follows:

 Open a new **Statistics** Page

 Click 'Enter Grouped Data' – see Fig 5. for the 'window' that opens.

Choose 'Use Raw data' and then 'Edit' as highlighted in Fig 5.

Paste the data (use 'Ctrl V' or a 'Right Click') and, if required, tick the 'Column Header' boxes as shown.

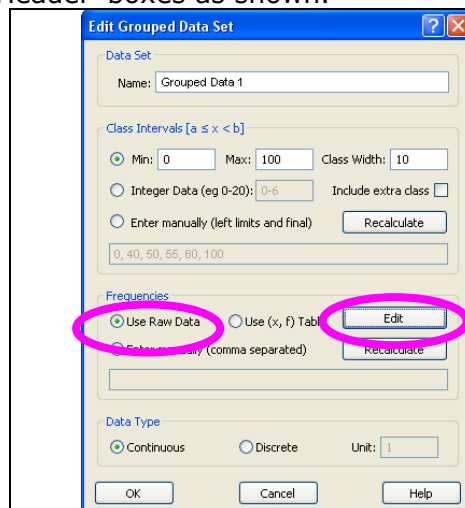


Fig 5. Hint: Choose your own 'Class Intervals' rather than accept the 'defaults'

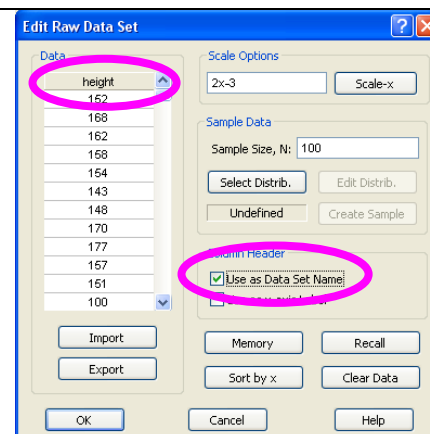




Fig 6. Hint: use the top cell in the spreadsheet column to name the Data Set

Nothing appears to happen – but lots of options light up on the 'Tool Bar'

 Histogram – options to choose from are shown in Fig 7a.

 Box Plot – based on Raw or Grouped Data (Fig 7b.)

 Cumulative Frequency options also available

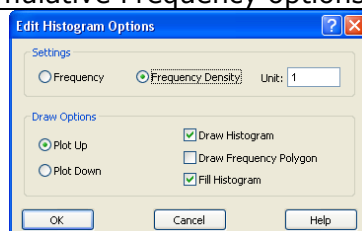


Fig 7a. Histogram Options

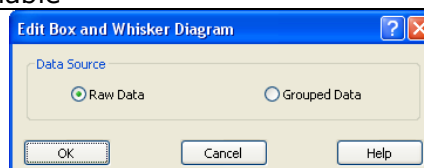


Fig 7b. Box and Whisker Options. See also options for Cumulative Frequency

The diagrams produced (e.g. Fig 8a and 8b.) can have **'dynamic'** text boxes added to enhance learning.

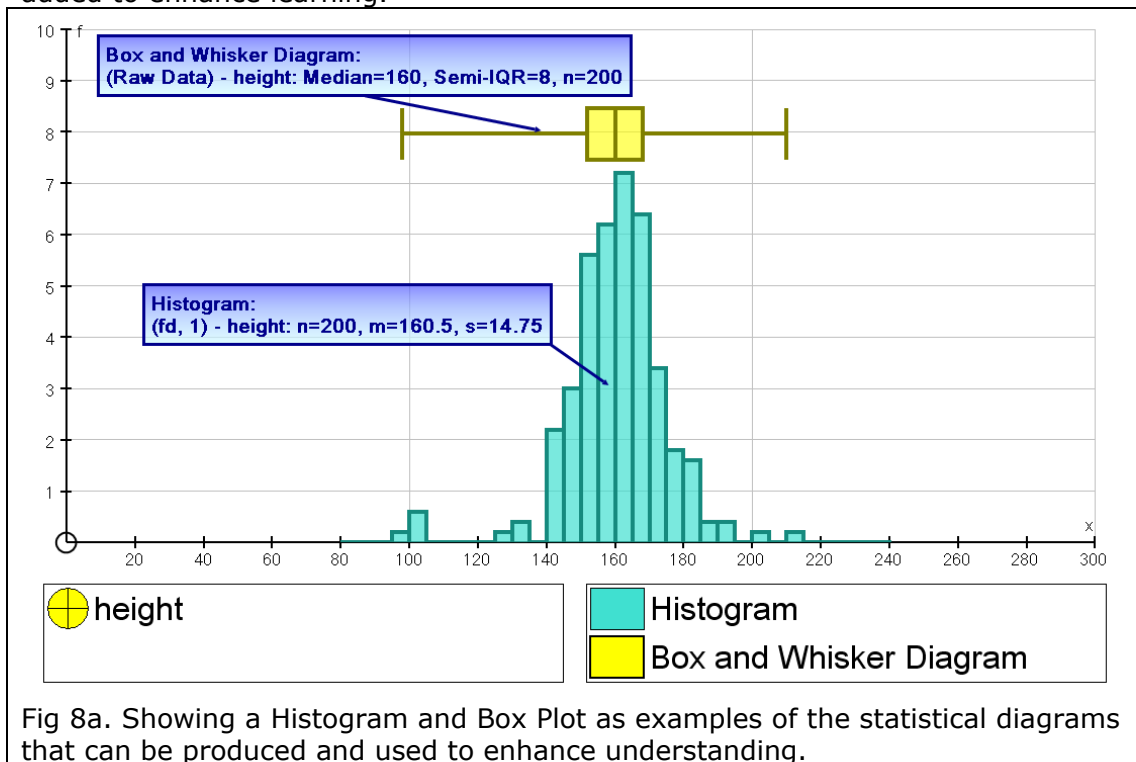


Fig 8a. Showing a Histogram and Box Plot as examples of the statistical diagrams that can be produced and used to enhance understanding.



The 'Drag' and various 'Zoom' options can be used as appropriate to get a clearer picture.



Selecting the 'Probability by Area' option introduces the idea of areas under a Histogram (dynamically) as shown in Fig 8b.

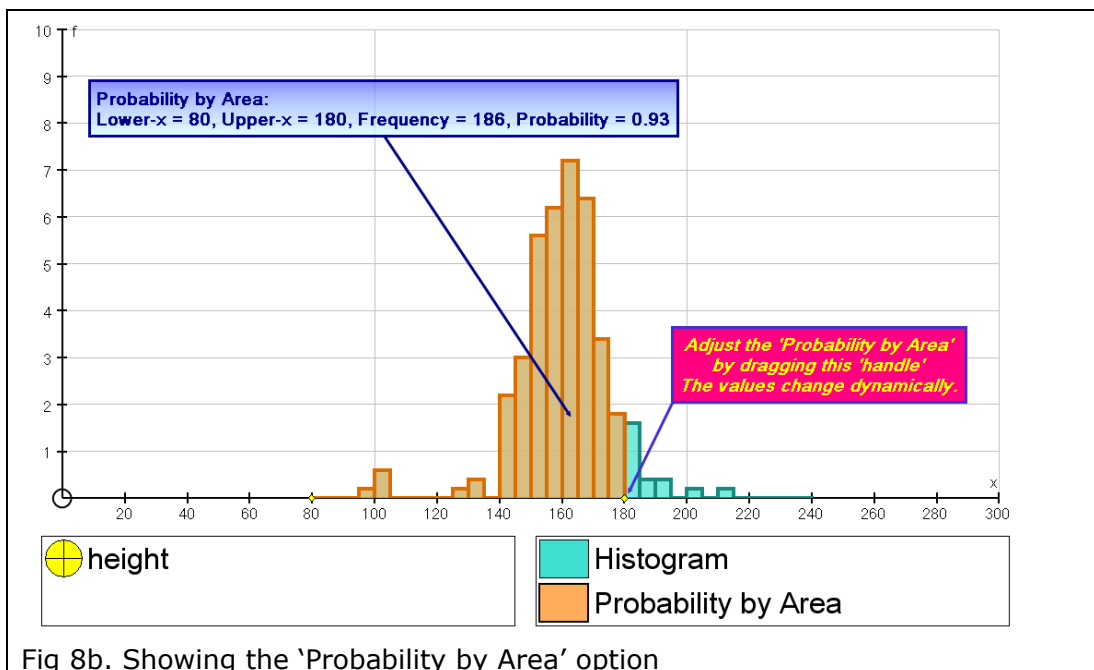


Fig 8b. Showing the 'Probability by Area' option

Copying results from Autograph to Word

This can be used creatively by the teacher during a lesson delivered through a Data Projector in many and varied ways. Here is just one idea:

Example – Data Table

Start by highlighting the data you require, e.g. the data table shown in Fig 9. Copy and paste this data as shown in Fig 12.

Table of Values of Histogram [height]:

Class Int.	Mid. Int. (x)	Class Width	Freq.	Cum. Freq.
$80 \leq x < 90$	85	10	0	0
$90 \leq x < 100$	95	10	1	1
$100 \leq x < 110$	105	10	3	4
$110 \leq x < 120$	115	10	0	4
$120 \leq x < 130$	125	10	1	5
$130 \leq x < 140$	135	10	2	7
$140 \leq x < 150$	145	10	26	33
$150 \leq x < 160$	155	10	59	92
$160 \leq x < 170$	165	10	68	160
$170 \leq x < 180$	175	10	26	186
$180 \leq x < 190$	185	10	10	196
$190 \leq x < 200$	195	10	2	198
$200 \leq x < 210$	205	10	1	199
$210 \leq x < 220$	215	10	1	200
$220 \leq x < 230$	225	10	0	200
$230 \leq x < 240$	235	10	0	200
$240 \leq x < 250$	245	10	0	200

Fig 12. Data copied from Fig 9. – the Results Box

Table of Values of Histogram [height]:

Class Int.	Mid. Int. (x)	Class Width	Freq.	Cum. Freq.
$90 \leq x < 100$	95	10	1	1
$100 \leq x < 110$	105	10	3	4
$110 \leq x < 120$	115	10	0	4
$120 \leq x < 130$	125	10	1	5
$130 \leq x < 140$	135	10	2	7
$140 \leq x < 150$	145	10	26	33
$150 \leq x < 160$	155	10	59	92
$160 \leq x < 170$	165	10	68	160
$170 \leq x < 180$	175	10	26	186
$180 \leq x < 190$	185	10	10	196
$190 \leq x < 200$	195	10	2	198
$200 \leq x < 210$	205	10	1	199
$210 \leq x < 220$	215	10	1	200

Fig 13. Here the data table has been 'tidied up' using the 'Convert Text to Table' option in Word (could also use 'Table Autoformat in Word')

The three middle columns can be copied back to Excel from the Word table shown in Fig 13. In Excel the 'Class Width' column can be deleted and the mid-range and frequency values used to e.g. estimate the mean of the frequency distribution. Of course the class do this 'long hand' and the teacher can use the projected images (Fig 14, Fig 10, etc.) to check working!

	A	B	C
1	x	freq	x*freq
2	95	1	95
3	105	3	315
4	115	0	0
5	125	1	125
6	135	2	270
7	145	26	3770
8	155	59	9145
9	165	68	11220
10	175	26	4550
11	185	10	1850
12	195	2	390
13	205	1	205
14	215	1	215
15	Σfreq= 200		
16	Σx*freq= 32150		
17			
18	Mean= 160.75		

Fig 14a. Tabulated data organised in Autograph then brought to Excel via Word

	A	B	C
1	x	freq	x*freq
2	95	1	=x*freq
3	105	3	=x*freq
4	115	0	=x*freq
5	125	1	=x*freq
6	135	2	=x*freq
7	145	26	=x*freq
8	155	59	=x*freq
9	165	68	=x*freq
10	175	26	=x*freq
11	185	10	=x*freq
12	195	2	=x*freq
13	205	1	=x*freq
14	215	1	=x*freq
15	Σfreq= =SUM(B2:B14)		
16	Σx*freq= =SUM(C2:C14)		
17			
18	Mean= =C16/B15		
19			

Fig 14b. Shows the formulae used in the spreadsheet opposite. Note the use of 'labels in formulae' – an often overlooked 'tool' in Excel that is so useful in teaching!

Example 2 – working with 2-variable data

Fig 15. shows part of the responses from one class to a *CensusAtSchool* questionnaire. Questions are discussed with the class about the nature of the scatter graph for e.g. "Height of belly button from the ground" onto "Height"

	B	E	F	G	H	J	K	L	M	N	O	P	
1	schoolcode	postcode	region	gender	age	Height	footLength	wristCirc	thumbCirc	BellyButton	liveln	fashionAcc	P
2	6789	ne25	North East	F	16	160	20	155	45	100	Town	Dont know	A
3	6789	NE25	North East	M	19	175	27	185	67	105	Town	Dyed Hair	C
4	6789	Ne28	North East	M	16	176	27	176	65	110	Town	Rings/Jewellery	C
5	6789	ne29	North East	F	18	157	25.5	156	65	100	Town	Mobile Phone	N
6	6789	NE29	North East	F	17	157	20	152	57	90	Village	Body Piercing(s)	N
7	6789	ne30	North East	M	17	178	26	165	65	105	Town	Mobile Phone	C
8	6789	NE25	North East	F	16	160	23	145	58	100	Town	Body Piercing(s)	A
9	6789	NE29	North East	F	16	170	24	139	49	104	Village	Other	N
10	6789	Ne33	North East	M	17	182	25	133	80	112	Town	Trainers	C
11	6789	NE29	North East	F	16	165	26	164	50	100	Town	Other	N
12	6789	NE29	North East	F	17	159	23	145	55	97	Town	Rings/Jewellery	F
13	6789	NE6	North East	M	16	179	23	160	64	110	City	Sunglasses	C
14	6789	ne26	North East	M	16	194	31	180	70	118	Town	Trainers	C
15	6789	NE30	North East	F	17	165	23	161	56	99	Town	Rings/Jewellery	N
16	6789	NE29	North East	F	16	150	23	160	60	86	Town	Shoes	A
17	6789	NE29	North East	F	16	165	26	164	50	100	Town	Other	N

Fig 15. Class Data from a CensusAtSchool questionnaire

To graph the relevant data on Autograph proceed as follows:

Use the 'Hide Column' option in Excel to ensure the two columns of data are next to each other as shown in Fig 16a.

	G	H	J	N	O
	gender	age	Height	BellyButton	liveln
					fashi
	F	16	160	100	Town
	M	19	175	105	Town
	M	16	176	110	Town
	F	18	157	100	Town
	F	17	157	90	Village
	M	17	178	105	Town
	F	16	160	100	Town
	F	16	170	104	Village
	M	17	182	112	Town
	F	16	165	100	Town
	F	17	159	97	Town
	M	16	179	110	City
	M	16	194	118	Town
	F	17	165	99	Town
	F	16	150	86	Town
	F	16	165	100	Town

Fig 16a. The two columns to be copied from Excel

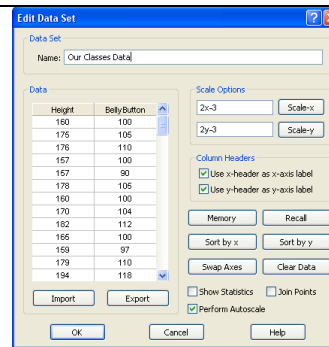




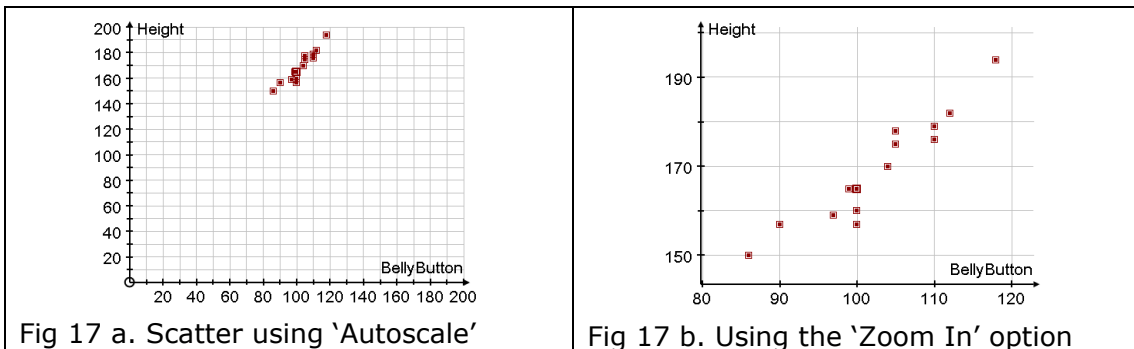
Fig 16b. Pasting the data into a GRAPH page in Autograph

 Open a new **Graph** Page and  Select 'Add Data'
 Paste the data as shown in Fig 16b. and 'Swap Axes' if required
 Use the 'column headers' as axes labels if appropriate (ticked here!)
 Select or de-select the 'Show Statistics' and 'Perform Autoscale' as required.

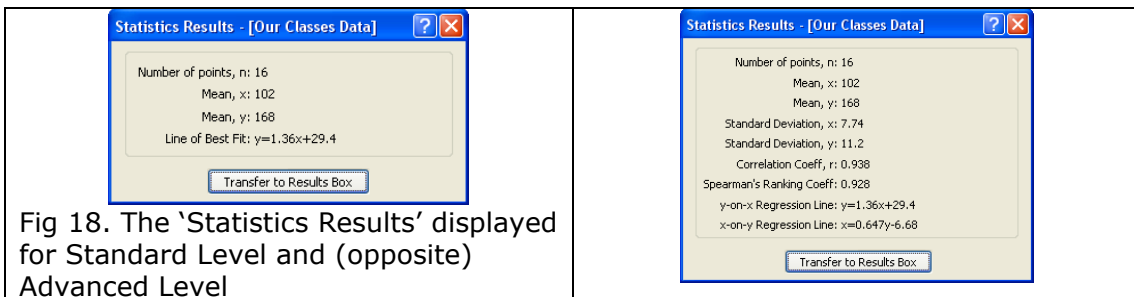
Notes

1. The Autoscale option will always show the data in relation to the origin. The zoom options can then be used as appropriate.
2. Before selecting the 'Show Statistics' option ensure that the appropriate 'level' (Advanced or Standard) is chosen initially – the information given in Advanced level is rather daunting for Key Stage 3 learners!

Fig 17. shows the scatter graph (without selecting the 'Show Statistics' option)



If the option to 'Show Statistics' is chosen from the dialogue box then the window (Fig 18.) opens depending on the level of Autograph chosen at the outset.



The graph shown in Figs 19 and 20 can be used as a creative teaching tool to discuss issues around lines of best fit etc. The use of dynamic text boxes really does bring the straight line fitting to life.

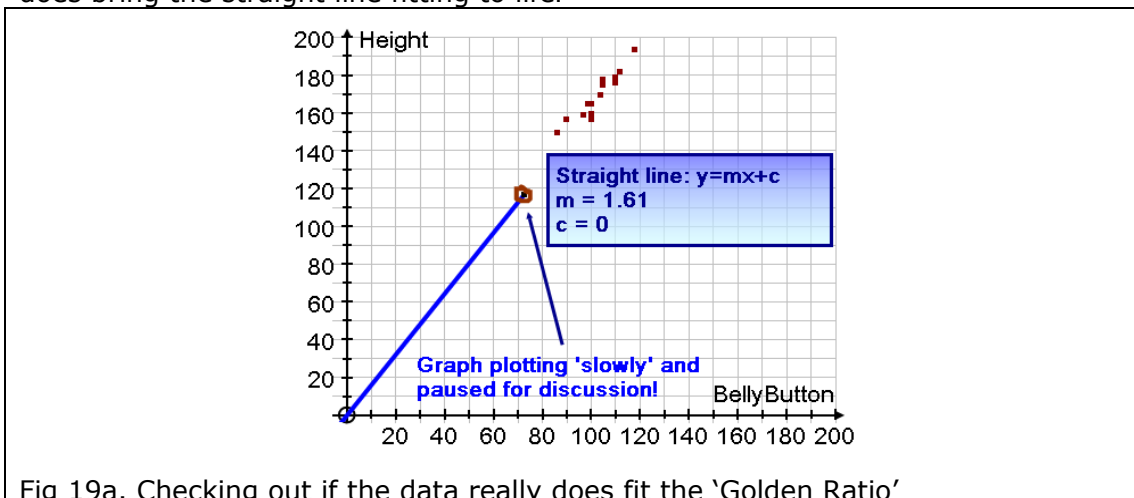


Fig 19a. Checking out if the data really does fit the 'Golden Ratio'

Using $y = mx + c$ is not a good way to search for a line of **best** fit as this should really pass through the (mean \bar{x} , mean \bar{y}) point. The object menu in Autograph (once the data set is selected!) can be used to locate this point (sometimes called the centroid). It is then possible to add a random point away from the data set and use the Autograph 'object menu' (also available from a 'right click' once the **two** points are **selected**) to draw a straight line through these two points – as shown in Fig 19b.

A superb animation can be added by selecting **two** 'objects' on the screen; the data set and this straight line. The option to show the 'residuals' as squares and then trying to minimise the total area of these squares dynamically by moving the 'random point' can certainly enhance understanding of the concept (see Fig 20a/b)

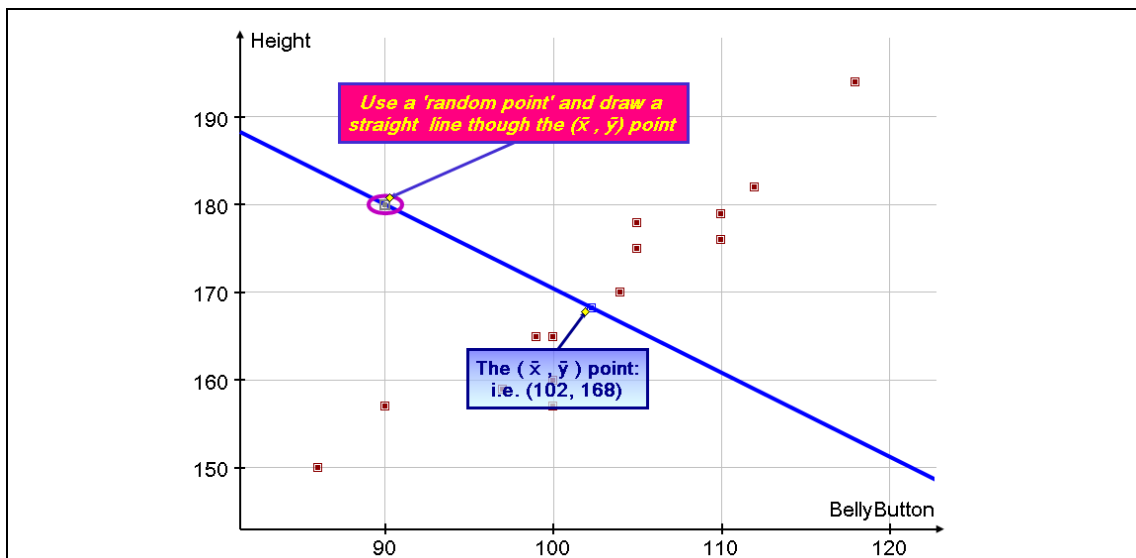


Fig 19b. Any learner who thinks this is a line of best fit (well let's face it there are seven points above and seven below the line!) is in for a shock when the squares move as in Fig 20a. Yes, such students do exist.

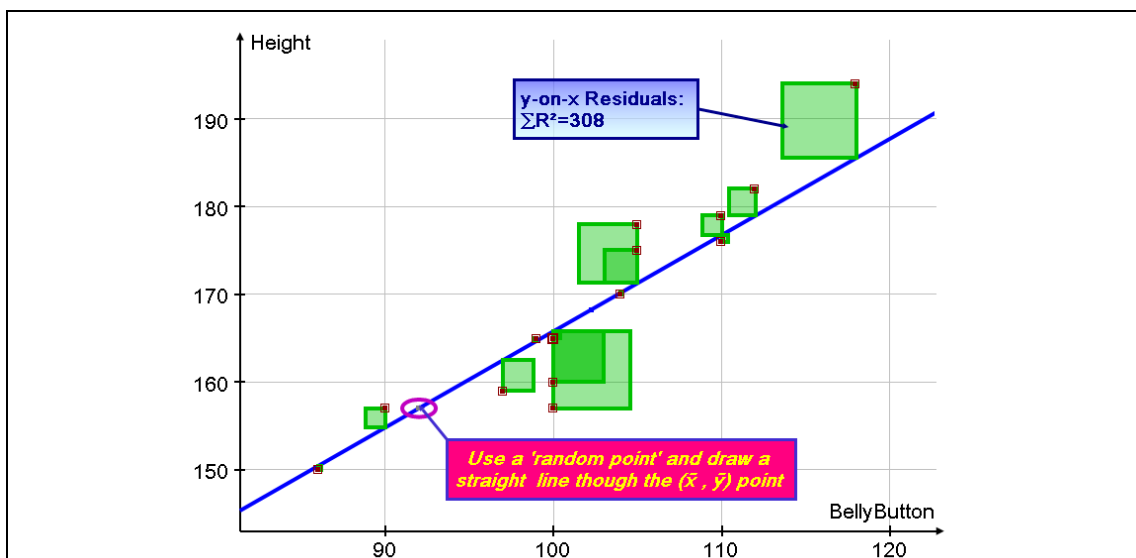


Fig 20a. Zooming in on the data and showing the 'residuals' which can be changed dynamically until the 'least squares area' is found

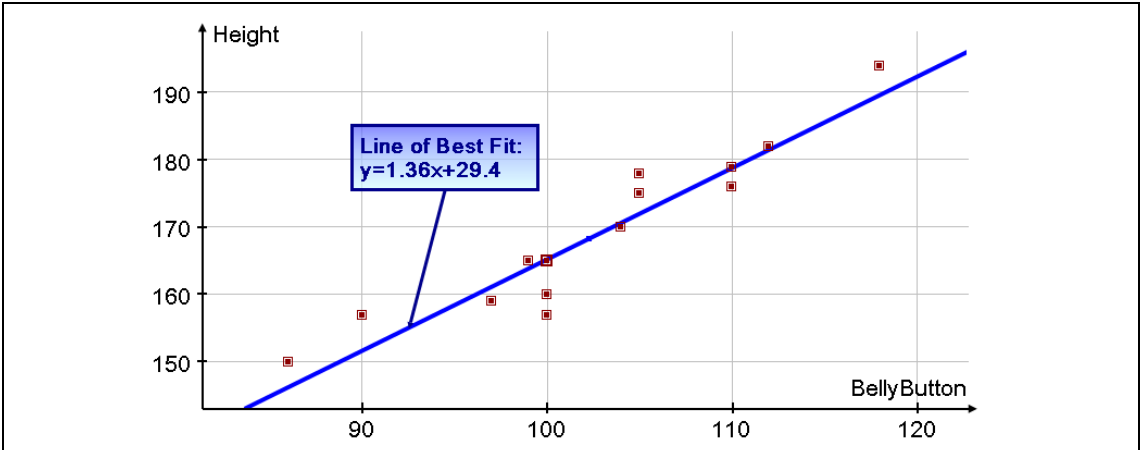


Fig 20b. Of course Autograph can show **directly** the exact line of best fit – but the previous diagrams illustrate the ‘interactive’ approach in the classroom that can support learning

Note – it is possible to dynamically observe the effect on the line of best fit by moving a single point if you hold down the ‘Ctrl’ key!

Example 3 – a ‘theory’ lesson on scatter graphs and lines of best fit

In order that learners might understand the theory of scatter diagrams, it is a good idea to simplify to, say, just five points. (I’ve actually done the following using just 3 points.) Starting with a blank graph page on Autograph, I get the students to choose the points (see Figure21). We discuss possible correlation, and then think what might be a best fit line and/or correlation coefficient. Figure 21 shows how a movable line can easily be added that passes through the centroid (mean of x, mean of y). Rotating this line and then adding the residuals, as squares, provides a lovely animation to help learners understand the concept (see Figure 22). With S1 learners, I like to get them to calculate values such as the correlation coefficient and regression line equation for their chosen points before using the ‘results box’ to confirm (or otherwise) their answers (see Figure 23). I then ask for suggestions on how we might achieve a correlation coefficient of zero (or 1).

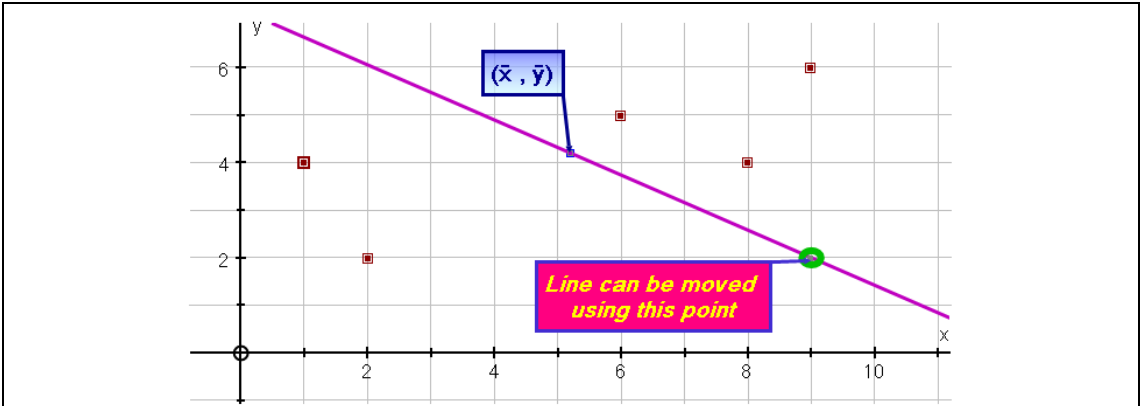


Fig 21. A data set consisting of five points (chosen by five learners) and a movable line that can be animated to promote discussion.

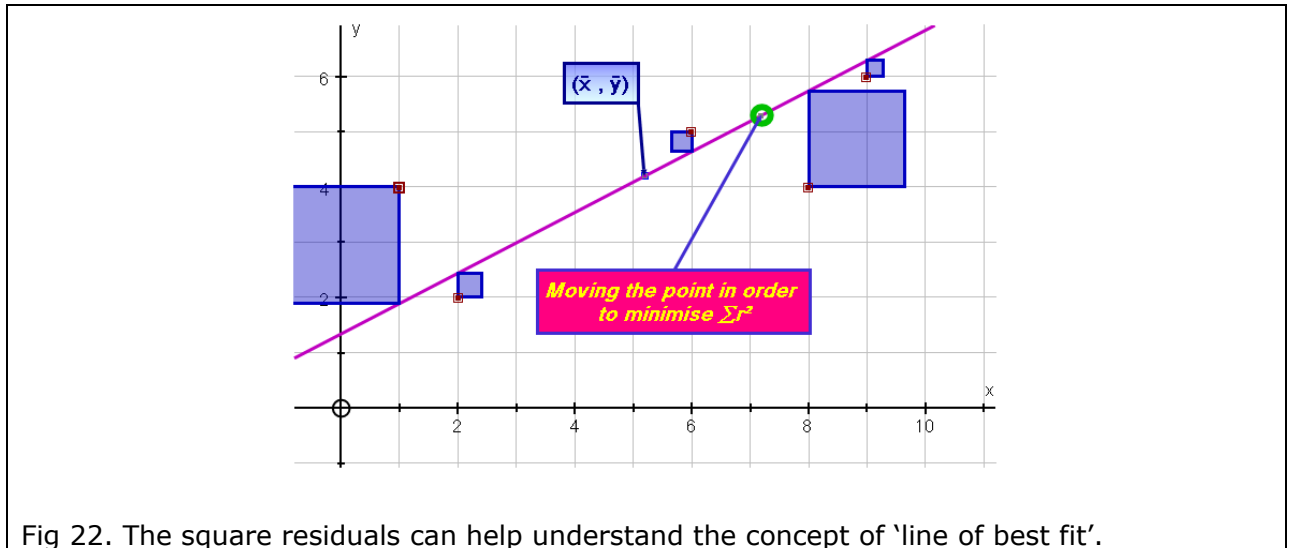


Fig 22. The square residuals can help understand the concept of 'line of best fit'.

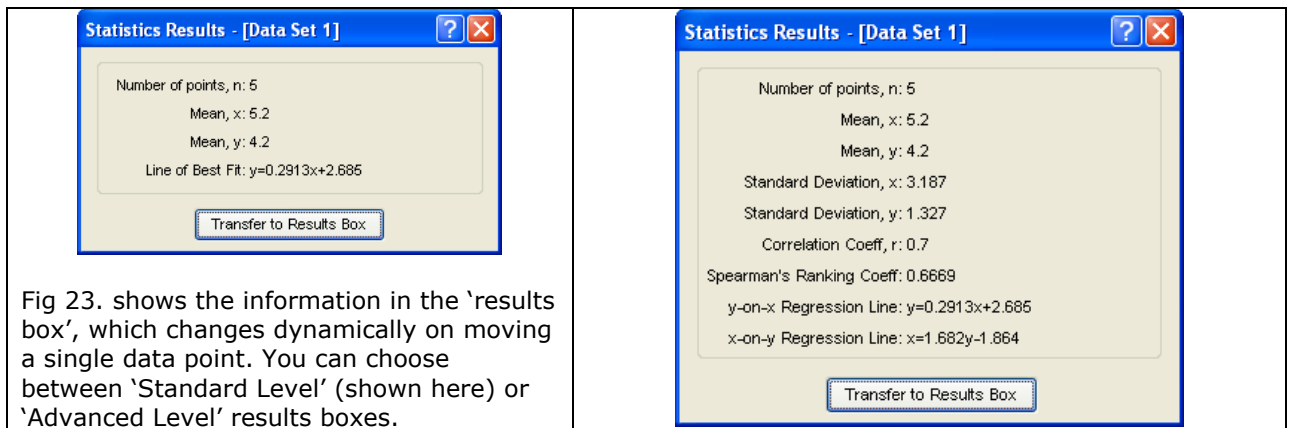


Fig 23. shows the information in the 'results box', which changes dynamically on moving a single data point. You can choose between 'Standard Level' (shown here) or 'Advanced Level' results boxes.

The above example was put to good use with my re-sit GCSE group, some of whom were under the impression that a line of best fit merely had to have the same number of points on each side. The dynamic visualisation certainly put misconceptions to rest.

There are many other ways in which I have used real data to make the teaching and learning of statistics a far more meaningful and fun experience for both teacher and learners.

Previously, I used to find teaching statistics a little tedious, but not any more!

And finally – if you are wondering how the mathematical symbols have appeared in text boxes of e.g. Fig 19 then it is all down the wonderful 'on screen' keyboard available through the 'View' menu. It is fantastic that this also works in Word and Excel etc which makes adding such symbols to worksheets a 'breeze'. Here it is in all its glory (no need to open it fully usually!)



Fig 24. The on screen keyboard