

WORKSHOP DATA SETS

DAVID J. HAND

Data set 1 is a very simple three group continuous variable set.

Data set 2 involves a categorical ordinal scale with missing values.

Data set 3 is a more complicated continuous variable set with two values recorded at each time.

Data set 4 is a still more complicated continuous variable data set with missing values and three variables measured at each time. It has a structure of conditions over time. Interest is in relationships between the variables.

Faculty of Mathematics
The Open University
Milton Keynes
MK7 6AA
England

Editor's Note: As mentioned often during the Workshop, the most informative approach, and one which is strongly recommended when the data are first analysed, is to produce separate graphs of the responses against time for each subject. To aid the reader considering the various approaches, we have done this here for Data sets 1 and 3 using the Macintosh software StatViewII.

The data can also be obtained by sending an email message to srw308@csc2.anu.edu.au.

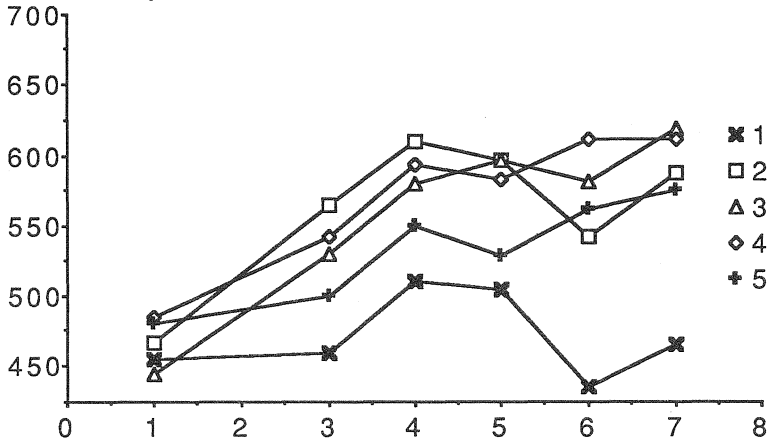
DATA SET 1

These data arose during a study of the effect of a vitamin E diet supplement on the growth of guinea pigs. The body weight (grams) was measured at the end of weeks 1, 3, 4, 5, 6, and 7. All animals were given a growth inhibiting substance during week 1 and the vitamin E therapy was started at the beginning of week 5. The animals were divided into three groups: group 1 received no vitamin E, group 2 a low dose, and group 3 a high dose. The data below give first the animal number, then its group, and then its six weight measures.

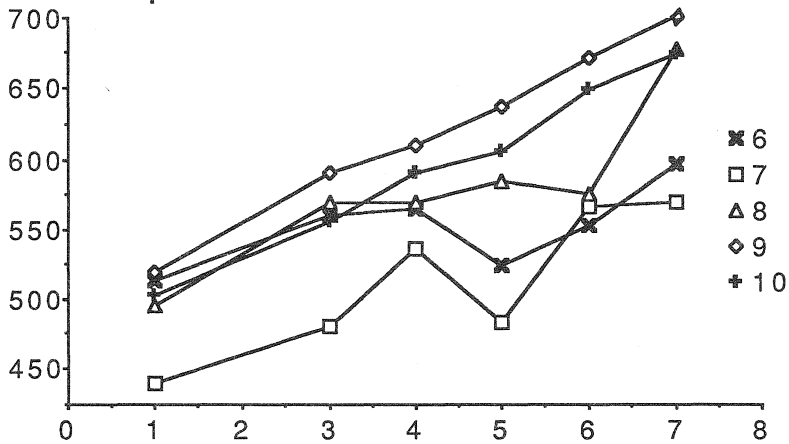
1	1	455 460 510 504 436 466
2	1	467 565 610 596 542 587
3	1	445 530 580 597 582 619
4	1	485 542 594 583 611 612
5	1	480 500 550 528 562 576
6	2	514 560 565 524 552 597
7	2	440 480 536 484 567 569
8	2	495 570 569 585 576 677
9	2	520 590 610 637 671 702
10	2	503 555 591 605 649 675
11	3	496 560 622 622 632 670
12	3	498 540 589 557 568 609
13	3	478 510 568 555 576 605
14	3	545 565 580 601 633 649
15	3	472 498 540 524 532 583

Editor's Note: These data are also considered in Crowder MJ & Hand DJ (1990), Analysis of Repeated Measures. Chapman & Hall: London.

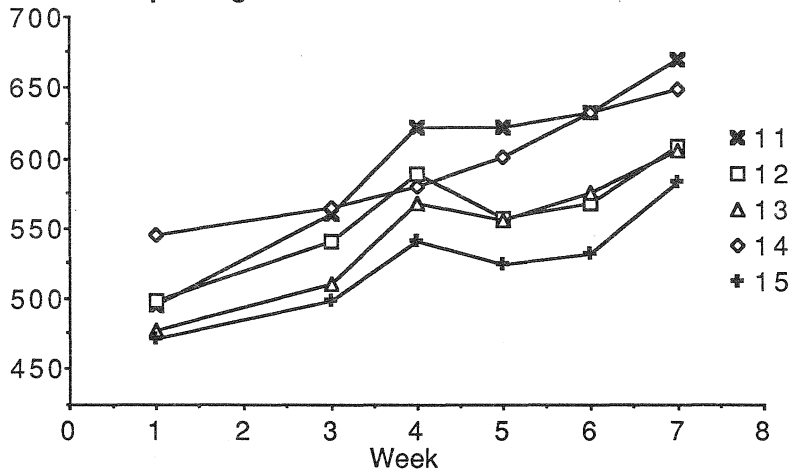
Group 1: No vitamin E



Group 2: Low dose vitamin E



Group 3: High dose vitamin E



DATA SET 2

These data arose during a clinical trial of a treatment for sprains. The first value for each patient is the patient number, the second is the group (1 = treatment, 2 = control) and the next four are tenderness scores on a scale 0 = none, 1 = slight, 2 = moderate, 3 = severe. A BLANK represents a missing value. Primary interest centred on differences between the two groups.

1	1	211	17	2	3000
2	1	21 0	18	2	11 0
3	1	1100	19	2	2220
4	1	110	20	2	1100
5	1	2100	21	2	2210
6	1	21	21	2	2210
7	1	210	22	2	110
8	1	2100	23	2	2222
9	1	200	24	2	210
10	1	2110	25	2	2100
11	1	3221	26	2	3221
12	1	3221	27	2	22 0
13	1	3210	28	2	222
14	1	3233	29	2	2211
15	1	332	30	2	2211
16	1	3211			

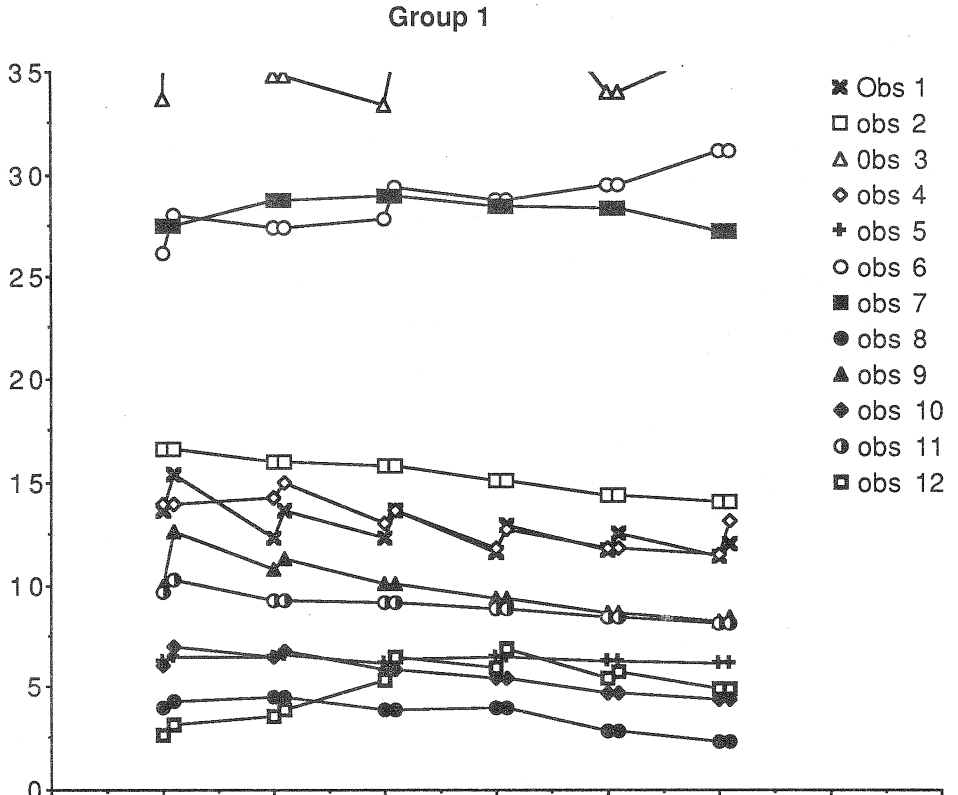
DATA SET 3

These data are part of the data collected during a larger study. I am grateful to Dr Gudrun Sartory and Dr Irene Daum of the Institute of Psychiatry for permission to use the data. The aim of the study was to explore the effects of predictability and control on the anxiety experienced by phobics. In the data extracted below, 24 snake and spider phobics were exposed to 6 repeated presentations of a slide depicting the object of their phobia. Subjects are divided into two groups, indicated by the first column (after the subject number). In level 1 a tone sounded to indicate that the slide was about to switch off, while in level 2 the tone did not presage the disappearance of the image.

At each of the six presentations two measures were taken: first the onset level of skin conductance and second the peak level of skin conductance obtained for this presentation of the slide. The values below are presented in the order onset, peak for presentation 1; onset, peak for presentation 2, etc.

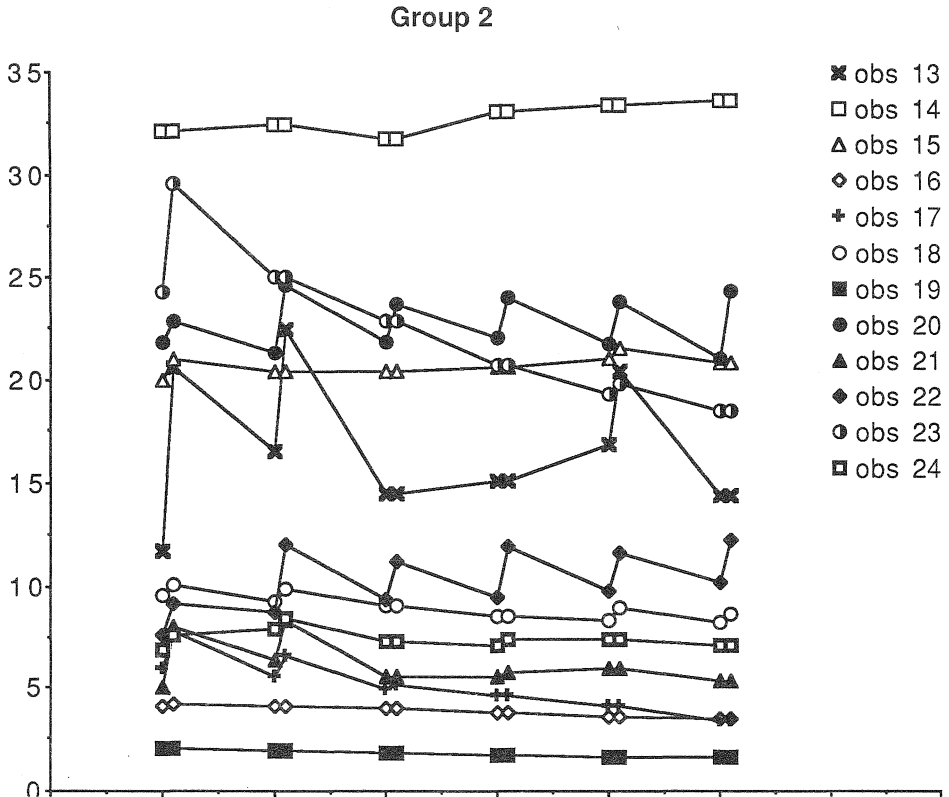
Questions of interest are such things as: Is there acclimatisation over exposures? Are there differences between groups and if so what are the differences? And so on.

1	1	13.6	15.4	12.3	13.7	12.3	13.6	11.6	12.9	11.7	12.5	11.4	12.0
2	1	16.6	16.6	16.0	16.0	15.8	15.8	15.1	15.1	14.4	14.4	14.1	14.1
3	1	33.7	47.8	34.8	34.8	33.4	35.8	40.3	40.3	34.0	34.0	36.1	36.1
4	1	14.0	14.0	14.3	15.0	13.0	13.7	11.8	12.7	11.8	11.8	11.5	13.1
5	1	6.3	6.5	6.5	6.6	6.2	6.4	6.5	6.5	6.3	6.3	6.2	6.2
6	1	26.2	28.0	27.4	27.4	27.8	29.4	28.7	28.7	29.5	29.5	31.1	31.1
7	1	27.5	27.5	28.7	28.7	28.9	28.9	28.4	28.4	28.3	28.3	27.2	27.2
8	1	4.0	4.3	4.5	4.5	3.9	3.9	4.0	4.0	2.9	2.9	2.4	2.4
9	1	10.0	12.6	10.8	11.3	10.1	10.1	9.3	9.3	8.6	8.6	8.2	8.4
10	1	6.1	7.0	6.5	6.8	5.9	5.9	5.4	5.4	4.7	4.7	4.4	4.4
11	1	9.6	10.3	9.2	9.2	9.1	9.1	8.8	8.8	8.4	8.4	8.1	8.1
12	1	2.7	3.2	3.6	3.9	5.3	6.5	6.0	6.9	5.4	5.7	4.9	4.9



(NB For easier visual comparisons between the two Groups, Observation 3's high values have been truncated.)

13	2	11.7	20.6	16.5	22.5	14.5	14.5	15.1	15.1	16.8	20.4	14.4	14.4	
14	2	32.1	32.1	32.4	32.4	31.7	31.7	33.1	33.1	33.4	33.4	33.6	33.6	
15	2	20.0	21.0	20.4	20.4	20.4	20.4	20.6	20.6	21.0	21.6	20.8	20.8	
16	2	4.1	4.2	4.1	4.1	4.0	4.0	3.8	3.8	3.6	3.6	3.5	3.5	
17	2	6.0	7.8	5.5	6.6	4.9	5.1	4.6	4.6	4.1	4.1	3.4	3.4	
18	2	9.5	10.1	9.2	9.9	9.0	9.0	8.5	8.5	8.3	8.9	8.2	8.6	
19	2	2.1	2.1	1.9	1.9	1.8	1.8	1.7	1.7	1.6	1.6	1.6	1.6	
20	2	21.9	22.9	21.4	24.6	21.9	23.7	22.1	22.1	24.0	21.8	23.8	21.0	24.3
21	2	5.0	8.0	6.4	8.3	5.5	5.5	5.5	5.7	6.0	6.0	5.3	5.3	
22	2	7.6	9.1	8.7	12.0	9.3	11.2	9.4	11.9	9.8	11.6	10.2	12.2	
23	2	24.3	29.6	25.0	25.0	22.9	22.9	20.7	20.7	19.3	19.8	18.5	18.5	
24	2	6.9	7.6	7.9	8.4	7.3	7.3	7.1	7.4	7.4	7.4	7.1	7.1	



DATA SET 4

I am grateful to Dr Christopher Bass of King's College Hospital, London for permission to use these data. They describe patients who suffer from panic attacks and control subjects.

The first variable is the subject number.

The second has values 1, meaning that this is a patient who suffers panic attacks, and 2, meaning that this is a control subject.

The next 11 variables are scores on an anxiety scale measured at the following times in minutes: 4, 6, 8, 10, 11, 14, 16, 17, 18, 19, 23.

The next 11 are scores (on the next line in the listing) on a carbon dioxide expiration scale measured at the same times.

The final 11 (on the third line on the listing) are pulse rates measured at the same times.

Times 4, 11, 14, 19, and 23 are rest times.

Times 6, 8, and 10 are times at which the subject is spoken to about the topic which they are anxious about.

Times 16, 17, and 18 are times at which the subject is asked to hyperventilate.

* 's signify missing values.

The researchers were interested in a number of hypotheses, including:

- (i) Do panic patients have lower CO₂ at rest than controls, and a greater fall during panic attacks?
- (ii) Do falls in CO₂ during phobic talk correlate inversely with levels of anxiety?
- (iii) Are ratings of anxiety during overbreathing correlated with speed of recovery of CO₂ level after overbreathing?

11 4 5 8 6 4 4 5 6 6 7 4
 33 30 21 26 27 28 18 15 14 19 22
 108 108 128 116 112 108 124 124 120 108 100
 21 5 6 6 5 3 3 7 8 8 8 4
 40 39 40 40 39 40 23 23 20 25 31
 94 104 106 104 92 * * * * 108 84
 31 5 5 5 5 5 5 6 6 7 6 4
 30 23 25 25 30 30 18 15 13 18 24
 96 96 96 84 96 96 120 108 108 96 84
 41 2 3 3 3 1 1 7 8 7 7 2
 29 29 28 28 28 29 19 17 15 19 22
 90 90 90 90 84 84 96 96 90 84 84
 51 4 5 7 8 4 2 2 3 4 2 1
 35 31 31 30 32 34 19 17 16 27 33
 78 78 84 78 78 78 78 78 84 84 78
 61 3 5 6 7 5 3 7 7 7 4 2
 38 34 28 22 27 33 21 17 18 24 31
 84 108 114 108 90 84 96 108 96 80 78
 71 1 5 6 6 2 1 5 7 8 5 1
 36 38 36 35 35 35 21 18 17 25 30
 86 94 88 82 84 80 104 104 104 96 74
 81 1 4 6 6 4 2 4 6 7 4 1
 32 32 30 32 32 31 19 13 10 10 31
 70 74 72 72 70 68 86 84 82 68 70
 91 1 1 5 2 1 2 2 2 2 1 0

37 36 35 36 37 37 19 17 16 23 29
 80 80 80 78 78 72 96 96 90 80 76
 101 4 8 8 8 3 4 7 7 8 8 5
 22 21 19 19 21 19 14 11 11 14 17
 78 78 78 78 72 78 90 90 90 84 74
 111 2 2 4 6 7 4 7 7 7 7 7
 31 31 31 29 29 30 21 18 17 22 24
 90 90 96 96 96 90 90 90 90 84 78
 121 8 8 8 8 7 6 4 3 2 4 2
 30 29 29 29 30 34 17 16 15 23 29
 124 124 132 130 118 124 140 152 148 130 102
 131 3 4 8 8 7 4 6 8 8 7 8
 26 25 23 23 23 24 19 17 16 19 20
 80 82 86 82 80 86 92 90 84 80 80
 141 2 2 2 3 2 1 4 5 6 3 1
 28 29 28 26 26 27 18 15 14 18 20
 98 110 110 110 108 102 132 134 140 114 78
 151 3 6 8 8 6 8 * * * * *
 35 37 34 34 33 28 * * * * *
 96 98 104 104 98 122 * * * * *
 161 4 7 8 8 3 2 7 8 * 7 3
 34 33 32 32 34 34 19 15 * 23 30
 66 76 74 70 62 64 94 92 * 72 62
 171 4 6 6 6 4 6 8 8 * 7 7
 32 32 30 30 28 26 19 18 * 19 19
 86 86 86 82 78 78 120 122 * 112 94
 181 5 5 8 6 2 3 6 8 * 5 3
 34 34 32 33 33 34 23 20 * 25 32
 96 108 118 116 100 100 124 130 * 102 90
 191 4 6 6 6 4 3 4 6 6 5 2
 32 30 28 27 31 32 16 15 15 19 27
 86 88 90 88 82 92 100 98 92 84 82
 201 2 3 4 4 2 2 7 8 * 4 2
 36 36 37 36 35 35 20 18 * 27 36
 80 86 86 82 78 78 114 122 * 86 80
 211 7 7 6 7 6 5 7 * * 6 6
 31 31 32 31 31 32 15 * * 28 30
 104 100 98 98 98 96 * * * 98 96

221 6 6 8 8 5 3 6 6 6 4 1
 30 31 31 33 30 30 25 21 21 23 28
 114 108 108 114 108 112 120 120 122 108 108
 232 1 6 7 7 6 3 3 4 5 6 4
 40 39 37 36 39 39 18 14 14 17 21
 92 104 102 102 92 84 146 166 166 132 102
 242 2 4 6 7 2 1 5 5 7 3 1
 34 34 34 33 34 33 22 19 18 25 31
 88 100 102 96 90 84 96 96 102 84 84
 252 1 3 3 4 1 0 3 3 3 2 0
 37 38 37 37 38 36 20 16 13 22 32
 84 86 84 88 74 72 108 108 108 84 74
 262 0 0 0 0 0 0 0 1 1 0 0
 34 34 34 33 33 33 19 18 17 27 34
 84 78 88 84 84 84 90 90 90 84 82
 272 1 4 8 8 2 1 2 3 4 2 1
 37 36 36 36 37 37 20 17 15 24 30
 80 100 98 96 92 84 120 126 132 90 80
 282 0 3 6 6 0 0 0 1 1 0 0
 37 37 37 36 36 36 21 20 18 25 31
 64 68 66 66 66 66 72 72 74 66 62
 292 0 0 0 0 0 0 1 4 4 0 0
 35 34 35 35 35 35 25 22 20 28 35
 98 104 106 106 94 94 110 112 118 94 90
 302 1 3 6 7 3 1 6 8 8 3 0
 40 38 37 39 40 41 18 13 10 18 36
 92 104 108 108 92 90 134 160 172 114 90
 312 1 1 4 8 3 0 6 8 8 6 2
 34 34 25 20 23 28 15 14 13 17 20
 74 80 84 88 70 68 122 132 126 92 68
 322 1 1 1 1 1 0 1 2 2 1 0
 32 31 31 32 32 32 21 19 18 29 32
 84 84 84 82 82 82 82 84 84 82 80
 332 0 0 2 3 0 0 3 4 5 0 0
 35 33 33 34 34 35 19 18 17 26 32
 70 78 76 78 74 72 80 80 80 74 70
 342 1 2 2 2 1 1 2 3 3 4 2
 40 40 40 40 40 40 18 16 14 27 38

64 66 64 62 64 64 98 84 80 64 66
 352 0 2 2 2 0 0 1 2 2 1 0
 36 36 36 35 35 35 18 15 13 20 28
 90 98 98 98 90 84 108 112 114 98 84
 362 1 1 2 4 2 1 2 5 * 3 1
 31 31 30 27 28 31 15 13 * 23 27
 90 94 88 86 82 78 104 116 * 82 82
 372 1 4 6 7 0 0 5 6 7 5 0
 39 38 37 36 37 40 21 16 14 20 28
 78 84 80 80 70 68 108 110 114 84 68
 382 2 4 4 5 3 1 5 5 6 3 3
 37 35 35 35 35 34 20 15 13 25 32
 90 96 92 94 94 92 102 122 130 86 82
 392 2 4 6 7 2 2 3 5 5 4 2
 40 40 40 40 40 40 17 15 13 23 33
 78 86 80 80 76 80 100 106 106 78 86
 402 1 4 8 8 2 1 2 6 6 4 1
 38 38 38 38 39 40 23 19 18 26 38
 98 96 98 94 98 94 108 120 122 98 98
 412 3 3 3 3 3 3 3 4 * * 4 6
 27 26 26 25 25 25 22 * * 23 24
 134 130 126 124 124 118 * * * 118 140