are averages of three items, so the standard deviation for these averages will be $\frac{2}{\sqrt{3}} = 1{\cdot}15$. Keeping in the coded form of the data, we should plot warning limits and the grand average plus or minus two standard deviations, i.e. at $1 \pm 2{\cdot}3$. Action limits would go at the grand average plus or minus three standard deviations, i.e. at $1 \pm 3{\cdot}5$. The control chart for salesmen's averages is shown in Fig. 100. It will be seen that they are in perfect control. In a similar way, since the district averages are the averages of four items, the standard deviation for these averages will be $\frac{2}{\sqrt{4}} = 1$, and the warning limits will go at 1 ± 2, with the action limits at 1 ± 3. The control chart for district averages is shown in Fig. 101. It will be seen that there is a marked suggestion of lack of control in this case. This reflects the fact that, although the variance ratio was below the 5% level, it was approaching it (a value 4 as against the 5% level of 5).

For our next illustration of the Analysis of Variance technique we turn to agriculture. Suppose we wished to investigate the effect of five different manurial treatments on the yield of wheat. We should take a block of land and subdivide it into plots of equal area so that it had the appearance of a chess board with five rows and five columns. One of the awkward things about field trials of this kind is that the soil in our experimental plot might show a systematic variation in fertility apart from any treatment applied by us in the course of the experiment. If our block contained a highly fertile strip which coincided with the five plots down one side of our block and we decided to apply one of our five treatments to this fertile strip, then when we found a high yield from this strip we should attribute it to our treatment when in fact it was due to the high fertility of the soil before ever we applied any manurial treatment to it. To get round this kind of difficulty we would be well advised to apply each treatment to five plots out of the total of twenty-five, the plots being chosen at random in such a way that each treatment occurred once and only once in each row and column of our chess board. Denoting the five treatments by the letters A, B, C, D, and E, one possible arrangement would be that shown in the following diagram, where the numbers are to be taken as the yields of wheat measured in bushels per acre.

A 13	*B* 9	*C* 21	*D* 7	*E* 6
D 9	*E* 8	*A* 15	*B* 7	*C* 16
B 11	*C* 17	*D* 8	*E* 10	*A* 17
E 8	*A* 15	*B* 7	*C* 10	*D* 7
C 11	*D* 9	*E* 8	*A* 15	*B* 11

For convenience in computing we shall code the yields by subtracting 10 bushels per acre in every case. The coded data are then as shown:

						Row Totals
	A 3	*B* −1	*C* 11	*D* −3	*E* −4	6
	D −1	*E* −2	*A* 5	*B* −3	*C* 6	5
	B 1	*C* 7	*D* −2	*E* 0	*A* 7	13
	E −2	*A* 5	*B* −3	*C* 0	*D* −3	−3
	C 1	*D* −1	*E* −2	*A* 5	*B* 1	4
Column Totals	2	8	9	−1	7	Grand Total $= T = 25$
Treatment Totals	*A*	*B*	*C*	*D*	*E*	Number of items $= N = 25$
	25	−5	25	−10	−10	Correction Factor $\frac{T^2}{N} = 25$

The Analysis of Variance then proceeds as follows:

Between Column Sum of Squares. Each column total is the sum of five items. Divide the sum of the squares of the column totals by the number of items going to make each total and subtract the correction factor. We get

$$\tfrac{1}{5}[2^2+8^2+9^2+(-1)^2+7^2]-25=14{\cdot}8 \text{ with 4 d.f.}$$

Between Row Sum of Squares. Each row total is the sum of five items. Divide the sum of the squares of the row totals by the number of items going to make up each row total and subtract the correction factor. We get

$$\tfrac{1}{5}[6^2+5^2+13^2+(-3)^2+4^2]-25=26 \text{ with 4 d.f.}$$

Between Treatment Sum of Squares. Each treatment total is the sum of five items. Divide the sum of the squares of the treatment totals by the number of items going to make up each treatment total and subtract the correction factor. We get

$$\tfrac{1}{5}[25^2+(-5)^2+25^2+(-10)^2+(-10)^2]-25=270 \text{ with 4 d.f.}$$

Total Sum of Squares. We find the sum of the squares of all the items in the table and subtract the correction factor. We get

TABLE OF SQUARED VALUES

	9	1	121	9	16
	1	4	25	9	36
	1	49	4	0	49
	4	25	9	0	9
	1	1	4	25	1
Totals	16 +	80 +	163 +	43 +	111 = 413

Hence, subtracting the Correction Factor, we find

$$\text{Total Sum of Squares} = 413 - 25 = 388$$

Since there are altogether 25 items, there is a total of 24 d.f.

TABLE OF ANALYSIS OF VARIANCE

Source	Sum of squares	d.f.	Variance estimate
Columns	14·8	4	3·7
~~Columns~~ Rows	26	4	6·5
Treatments	270	4	67·5
Residual	77·2	12	6·4
Total	388	24	

The residual sum of squares is that portion of the total sum of squares not accounted for by row, column, or treatment effects. Inspection of the table of the Analysis of Variance shows at once

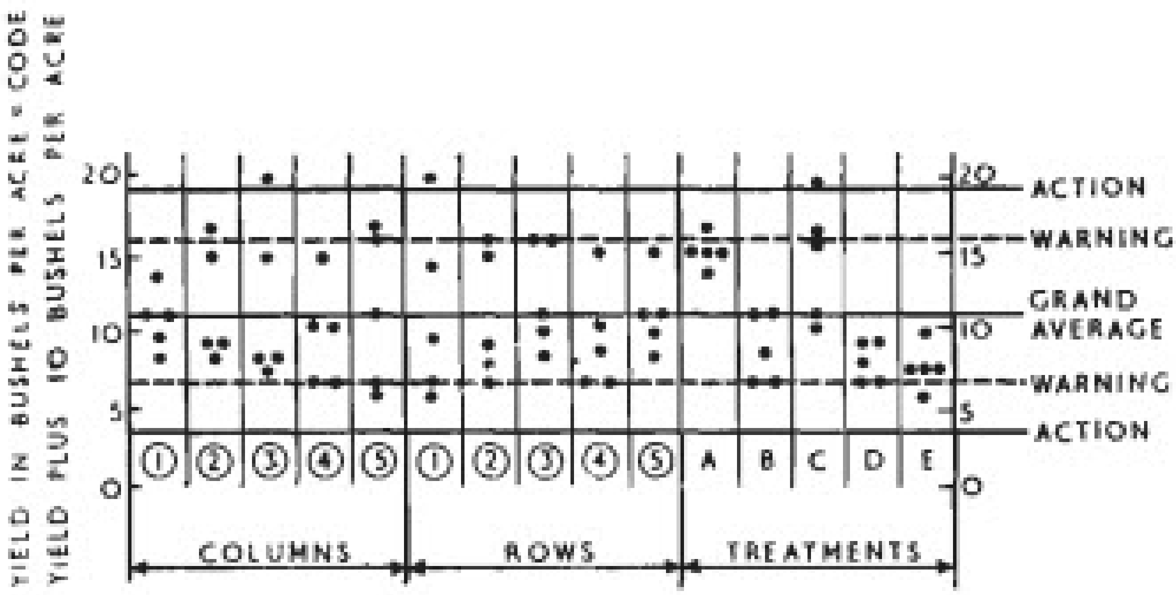

Fig. 102. Control Chart for latin square analysis on manurial treatment for wheat

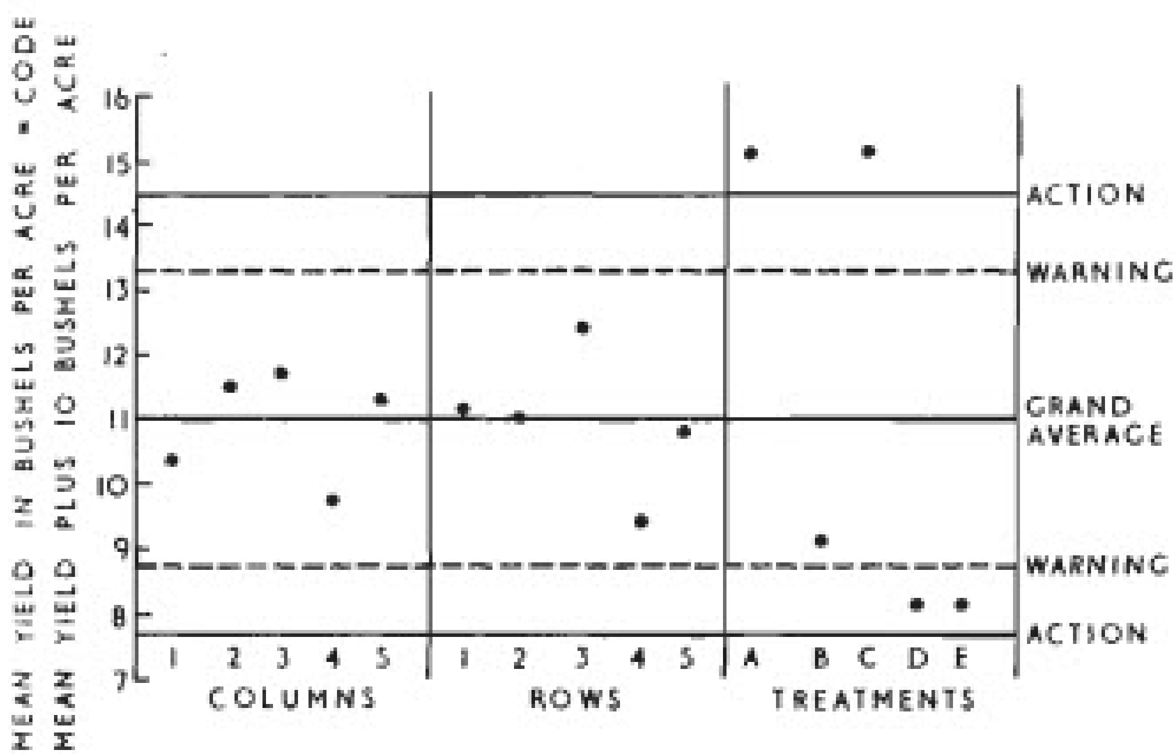

Fig. 103. Control Chart for row, column, and treatment averages in latin square analysis on manurial treatments

that there can be no question of the Estimates of Variance based on row or column degrees of freedom being significantly greater than the residual variance estimate. There is therefore no reason to suppose that there is any significant change in fertility of the soil across the rows or columns. Any apparent variation is simply a reflexion of those other causes of variation which we normally describe as experimental error. The reader may confirm for himself, however, that the treatment effect is highly significant as judged by the F test. We are justified, therefore, in believing that treatments A and C really do give a higher yield than other treatments, and may proceed to calculate confidence limits for their yields in bushels per acre.

This time we shall plot the individual plot yields in a control chart. The residual variance is 6·43 which gives us a standard deviation $\sqrt{6\cdot4} = 2\cdot5$ bushels per acre for individual plot yields. The plot yields are shown in Fig. 102 according to the three ways of looking at them: (a) by rows, (b) by columns, (c) by treatments. Compare this with Fig. 103.

An extremely useful design in Analysis of Variance is the so-called 'Factorial'. In order to illustrate this design type we shall take a fairly complex example so that the reader has a chance to acquire what we may well term the 'routine' of the analytical procedure. The example is typical of the situations in which a factorial design suggests itself as suitable. Silvered mica condensers are manufactured from small mica plates, silvered on each side, and finally impregnated with petroleum jelly to keep out moisture. The silver is applied to the mica plates in the form of a spray or printing ink and the vehicle carrying the silver is then driven off by a firing process. Broadly speaking, there are two variables in the firing process: temperature and time. The impregnation process may well be carried on by immersing the plates in a bath of petroleum jelly heated to a steady temperature somewhere below the 'flash point'. For the impregnation process we again have the variables, temperature and time. Suppose, now, we wished to investigate what combination of firing and impregnation times and temperatures would give our condensers the highest quality as measured by the 'loss angle' of the plates (the lower the loss angle, the better the condensers). To investigate this problem, we might choose three likely firing temperatures, three likely firing times, three likely impregnation temperatures, and three likely