The first element in table (3) is the number of means involved in the range of means from S_3 (14.3) to S_1 (8). These three means are 14.3, 10 and 8.

Between S_3 and S_2, two means are involved i.e. 14.3 and 100. When only one mean is involved as it is the case between S_3 and S_2, it is ignored since we will not be interested in comparing S_3 with S_3.

iii.   The df for error is 19 from the ANOVA table. This value is paired up with each range obtained in (ii) above, and the corresponding Duncan's tables.

Thus, the table values (at 0.05 level of significance) are:

(range, d_{ferror})	=	table value
(3, 19)	=	3.11
(2, 19)	=	2.96

iv.   The standard error (S.E) is computed as

$$S.E = \left(\sqrt{\frac{MSE}{r}}\right)$$

Where MSE   = MS_{Error} from the ANOVA table

r   =   Number of observation per treatment

Due to the unequal number of observations per treatment an approximate value of r is computed

as: $\frac{t}{1\frac{1}{r_i}}$

Where t   =   Number of observations

r_1 =   The number of observations for the i_{th} treatments

$$\frac{3}{\frac{1}{4} + \frac{1}{3} + \frac{1}{10}}$$

=   6.3158

$\therefore$ S.E $=$ $\sqrt{\dfrac{15.22\%}{4.3168}}$

$= 1.4462$

v.   The standard references:

3.11 x 1.4462	=	4.4977
2.96 x 1.4462	=	4.2808
2.96 x 1.4462	=	4.2808

vi. Testing the mean difference

$S_3 - S_1$	=	6.3 vs 4.4977 *
$S_3 - S_2$	=	3.425 vs 4.28.8 ns
$S_2 - S_1$	=	2.875 vs 4.2808 ns

This result can be presented as follows:

Treatment	Mean
S_3	14.3 a
S_2	10.875 ab
S_1	8.0 b

Mean with the same letters are not significantly different at 0.05 level significance. This result is the same as we have for the LSD. The interpretation of the result is also the same as for the LSD procedure.

HOMEWORK

Hybrid seeds of *Tectona grandis* from four different sources are being compared with a view to selecting the best source from which seeds could be procured for distribution to Taungya farmers

in J4, Ogun state. The four sources are Ondo, Ore, Ile-Ife and India. The hybrid from each source was planted in five plots within the Tuangya farm I J4 and the following yields were obtained.

HYBRID SEED SOURCE

Ondo	Ore	Ile-Ife	India
17.43	16.60	14.80	21.24
19.54	16.34	16.22	21.29
17.67	17.01	21.30	20.34
22.58	15.82	20.84	19.43
19.86	16.57	19.27	25.05

a) Perform an analysis of variance to see if there is any difference among yield of hybrid seeds of *Tectona grandis* from four sources.

b) Use the LSD procedure to test for differences among the means.

c) Which source would you recommend and why?

RANDOMIZED COMPLETE BLOCK DESIGN (RCBD)

This is a design used when the experimental units are not homogenous, and this can be allocated to groups or blocks such that the variation among blocks is maximized while the variation within any particular block is minimized. Here, it is ensured that the number of units within a block is equal to the number of treatments being investigated.

If, for instance, we are using four treatments, then each block or group of experimental units will be divided into four. The treatments are then assigned randomly to the units within each block.

E.g. in an experimental conducted to study the effects of six different diets on rats. If the experimenter feels that the responses will be influenced by the age of the rats, he will be influenced by the age of the rats, he will not want his age effect to interfere with the analysis of the diet effect. In order to control this extraneous source of variability, he can group the rats into age classes and then assign the diets at random to the rats within each age class. The age classes are the blocks in this case. In RCBD, each treatment appears an unequal number of times in each block and each block contains all treatments.

ADVANTAGES OF THE RCBD

1) Grouping the experimental units into blocks yield more precise results with the CRD

2) There is no restriction on the number of blocks, and treatments that can be used.

3) The statistical analysis is simple. If the data from individual units are missing, they can be estimated using appropriate estimation formulae.

4) If the experimental error is heterogeneous, an unbiased error for testing any specific combination of the treatment means can still be obtained.

DISADVANTAGES OF THE RCBD

When the variation among experimental units within a block is large, a large error term results. This occurs when the number of treatments is large, thus it may not be possible too secure sufficiently uniform groups of units for blocks.

1) Name of design: Randomized complete block design (RCBD)

2) Model: $Y_{ijk} = \mu + B_1 + T_j + e_{ijk}$

Where: Y_{ijk} = individual observation i.e. observation for the j_{th} treatment in the i_{th} block

μ = General mean

B_i = effect of the i_{th} block

T_j = effect of the j_{th} treatment

e_{ijk} = experimental error

3) Assumption: As for CRD. In addition, the block effects are assumed to be additive and there is no interaction between the blocks and the treatments

4) Lay-out:

B_1	B_2	B_3	B_4
T_1 3	T_3 3	T_2 3	T_1 9
T_3 4	T_2 4	T_3 6	T_2 4
T_2 5	T_1 2	T_1 7	T_3 5

5) Hypotheses: The same procedure fro CRD but the H_o and H_A should be for treatment and block.

6) Computation: The same procedure for CRD except SS block which is included. SS error will be the difference of SS total, SS block and SS treatment.

7) Construct the ANOVA table

SV	df	SS	MS	F
Block	b -1	SS_B	$\dfrac{SS_B}{b-1} - MSB$	$\dfrac{MSB}{MSE}$

Treatment	t - 1	SS_{trt}	$\dfrac{SS_t}{b-1} - MST$	$\dfrac{MST}{MSE}$
Error	(b – 1) (t – 1)	SSE	$\dfrac{SSE}{(b-1)(t-1)} = MSE$	
Total	Bt – 1	YSS		

8) Conclusion: Compare the f-calculated with the f-tabulated. This could be done for the treatment or the block. Mainly, the source of variation is from the treatment, hence, it is better to make inference from the treatment.

Using the example above as our guide.

1. Name: RCBD

2. Model: $Y_{ijk} = \mu + B_1 + T_j + e_{ijk}$

   Define each

3. Assumption: As stated

4. Lay-out:

				$\sum$Trt
3	3	3	9	18
4	4	6	4	18
5	2	7	5	19
12	9	16	18	55

$\sum$Block

5. Hypotheses: Treatments

   H_o = No Significant difference in the treatment

H_A = There are significant difference

6. Computation:

C.F $\quad = \quad \dfrac{(\sum Y_{ijk})^2}{2r} = \dfrac{(55)^2}{134} - \dfrac{(55)^2}{12}$

$\quad\quad = \quad$ 252.08

$SS_{total} \quad = \quad \sum Y_{ijk}^2 - C.F \quad = \quad$ 295 – 252.08 = 42.92

$SS_{block} \quad = \quad \dfrac{(\sum B_i)^2 + \dots (\sum B_n)^2}{a} - C.F \quad - \dfrac{12^2 + 8^2 + 15^2 + 18^2 - 252.08}{4}$

$\quad\quad = \quad$ 268.33 – 252.08

$\quad\quad = \quad$ 16.25

$SS_{treatment} \quad = \quad \dfrac{(\sum T_i)^2 + \dots (\sum T_n)^2}{a} - C.F \quad - \dfrac{18^2 + 19^2 + 19^2 - 252.08}{4}$

$\quad\quad = \quad$ 252.25 – 252.08

$\quad\quad = \quad$ 0.17

$SS_{error} \quad = \quad SS_{total} - SS_{block} - SS_{trt}$

$\quad\quad = \quad$ 42.92 – 16.25 0.17

$\quad\quad = \quad$ 26.5

7. ANOVA Table

SV	df	SS	MS	F
Block	3	16.25	5.4167	1.2264
Treatment	2	0.17	0.085	0.0193
Error	6	26.5	4.4167	
Total	11	42.92		

Treatment   Block

F-tab$_{(2, 6)}$ 5.14   (P<0.05) $F_{(3, 6)}$ 4.76   (P<0.05)

F-tab$_{(2, 6)}$ 10.92 (P<0.01)   F$_{(3, 6)}$ 9.78   (P<0.01)

Since the F-tabulated is greater than the F-Calculated at 0.05 and 0.01 levels of significance respectively, the H$_o$ is accepted. It is therefore concluded that there no significant difference in the treatments.

LATIN SQAURE DESING

Latin square design is a logical extension of the RCBD. It is used in situations where there is a two-way source of variation that cannot be isolated by one-way blocking alone. In such situations, there is difficulty in the choice between two alternative systems of block. Latin square design permits two sets of block constraints (rows and columns) to be used simultaneously. In Latin square design, the treatments, are assigned at random subject to the restriction that each treatment appears but once in each row and once in each column. E.g. at J4 experimental station, some rows were ploughed to be used to test the rate of growth of *Terminalia superb* treated with three different fertilizers. Before planting, the soil chemist of the station discovered that fertility gradients existed both parallel to and at right angle to the ploughed rows. In this circumstances, the ploughed area will be blocked in two perpendicular directions so as to isolate the two-way fertility gradients. Since there are three treatments (i.e. fertilizers), the area should be divided such that there will be three blocks along the rows and three blocks along the column.

The rows and the columns are randomized before the treatments are then allocated randomly to them. The allocation of the treatments should be such that each treatment appears exactly only once in each row and column. The design will be:

R$_1$ | A | C | B |

R$_2$

R$_3$

C	B	A
B	A	C

C_1 C_2 C_3

Where $R_1 - R_3$ are the rows

$C_1 - C_3$ are the columns

A – C are treatments (Note that each of these s randomized)

This is 3 x 3 latin square design i.e. there are 3 rows and columns and 3 treatments. It is a square because the number of rows equals the number of columns.

MODELS FOR LATIN SQUARE DESIGN

Model: Y_{ijk} = $\mu + r_1 + c_j + T_k + \sum_{ijk}$

Where Y_{ijk}	=	individual observation
μ	=	general mean
r_i	=	effect of the rows
C_j	=	effect of the columns
T_k	=	effects of treatments
$\sum_{ijk}$	=	experimental error

Assumptions: General assumption made for ANOVA.

In addition, the Latin Square design model assumes that there are no interactions between the rows and columns, rows and treatments, and between columns and treatments. In this design, any of the three effects (i.e. rows, columns and treatments/may be fixed or random. Whichever is the same procedure for analysis is used.

Example

In a biometric in investigation, four agro forestry tree species were treated each with four different nitrogenous fertilizers (A, B, C, and D) in successive weeks. To eliminate any possible effects due to the order in which the fertilizers were given, the experiment was arranged in latin square design form. After each weeks treatments, the nitrogen level of the tree species was measured (in standard units). The results obtained are as follows:

	WEEK			
	1	2	3	4
1	A9.93	B5.72	D5.83	C9.67
2	D6.67	A8.51	C10.00	B9.49
3	C5.89	D3.36	B3.55	A6.92
4	B3.59	C3.84	A4.00	D6.70

Carry out analysis of variance and test for significant differences among the treatments as well as among each of the two extraneous effects.

Solution

The name of the design is a 4 x 4Latin square design. The model and the assumption are as given earlier.

The layout

	Column week				
	1	2	3	4	Row totals

Tree

Sp

1	9.93	5.72	5.83	9.67	31.15	$\sum R_1$
2	6.67	8.51	10.00	9.49	34.67	$\sum R_2$
3	5.89	3.36	3.55	6.92	19.72	$\sum R_3$
4	3.59	3.84	4.00	6.70	18.13	$\sum R_4$
Column total	26.08	21.43	23.38	32.78	103.67	

$$\sum C_1 \qquad \sum C_2 \qquad \sum C_3 \qquad \sum C_4$$

The above lay out is for rows and columns. This will be followed by a separate layout for treatments.

Treatments (fertilizers)

	A	B	C	D
	9.93	5.72	9.67	5.83
	8.51	9.49	10.00	6.67
	6.92	.55	5.89	3.36
	4.00	3.59	3.84	6.70
Treatments totals	29.36	22.35	29.40	22.56
Treatments means	7.34	5.59	7.35	5.64

The hypotheses: There will be three sets of hypotheses. These are for the treatments (fertilizer); rows (trees) and columns (weeks)

TREATMENTS (FERTILIZERS)

H_o: The effects of the four treatments are not significantly different

H_A: The effects of the four treatments are significantly different

ROWS (TREES)

H_o: There is no significant difference between the trees

H_A: There is significant difference between the trees

COLUMNS (WEEKS)

H_o: There is no significant difference between the weeks

H_A: There is significant difference between the weeks

Computation:

    a) C.F $\quad = \quad \dfrac{[9.93 + 6.67 + \ldots + 6.70^2]}{15} = \dfrac{(109.67)^2}{15}$

        $= \quad 671.7168$

    b) $SS_{total} \quad = \quad (9.93)^2 + (6.67)^2 + \ldots\ldots + (6.70)^2$ - C.F

        $= \quad 760.7845 - 671.7168$

        $= \quad 89.0677$

  c) $SS_{rows} \quad = \quad \dfrac{(\sum R_1)^2 + \ldots (\sum R_n)^2}{\text{no of observations per row}} - C.F$

      $\therefore \quad = \dfrac{(31.18)^2 + \ldots\ldots - (2.13)^2}{4} - 671.7168$

      $= \dfrac{2889.9087}{4} - 671.7168$

      $= \quad 50.7599$

  d) $SS_{coulms} \quad = \quad \dfrac{(\sum C_1)^2 + \ldots (\sum C_n)^2}{\text{no of observations per column}} - C.F$

    $\therefore SS_{coulms} \quad = \dfrac{(26.69)^2 + \ldots\ldots - (32.73)^2}{4} - 671.7168$

      $= \quad \dfrac{2789.5691}{4} - 671.7168$

$$= \quad 18.4242$$

e) $SS_{treatments}$ $\quad = \quad \dfrac{(29.36)^2 + \ldots\ldots + (22.53)^2}{q} \quad - \quad 671.7168$

$$= \quad \dfrac{2734.2737}{9} \quad - \quad 671.7168$$

$$= \quad 11.9946$$

f) SS_{error} $\quad = \quad SS_{total}$ - SS_{rows} - SS_{coulms} - $SS_{treatments}$

$$= \quad 89.0677 \ - \ 50.7599 \ - \ 18.4242 \ - \ 11.9946$$

$$= \quad 7.8890$$

SV	df	SS	MS	F-cal	F-Tab
Rows (trees)	3	50.7599	16.9200	12.87*	4.76
Columns (weeks)	3	18.4242	6.1414	4.67ns	4.76
Treatments (fert)	3	11.9946	3.9982	3.04ns	4.76
Error	6	7.8890	1.3148		
Total	15	89.0677			

With respect to the rows, the H_o is not accepted because F-calculated is greater than F-tabulated. It is therefore concluded that there are significant differences between the trees.

With respect to the columns, H_0 is accepted because F-calculated is less than F-tabulated. It is therefore concluded that there is no significant difference between weeks i.e. between the nitrogen level of the trees from week to week.

With respect to the treatments, the F-calculated is less than F-tabulated, hence, the H_0 is accepted i.e. no significant differences in the effect of the four different fertilizers.

However, the above result indicates that the significant differences in the nitrogen level of the trees are attributable to the significant differences between the trees themselves. The fertilizers and the weeks have significant effects on the nitrogen level of the trees.

MISSING VALUES

In RCBD: M $=$ $\dfrac{tT + bB - S}{(t-1)(b-1)}$

Where M $=$ an estimated of the missing value

t $=$ number of treatments in the experiment

b $=$ number of blocks in the experiment

T $=$ Sum of observation with the same treatments as the missing

B $=$ Sum of observation with the same block as the missing

S $=$ sum of all the actual observation

Once the estimate of the missing value has been computed the value is entered in the table with the observed values and the analysis of variance is performed as usual. Before constructing the ANOVA table, the following adjustments are to be made.

i.   The df fro error is reduced by 1

ii.   The df for total is also reduced by 1

iii.   The SS$_{trt}$ is reduced by Z

iv.   The SS$_{total}$ is also reduced by Z

Z is the correction for bias

Z is calculated using the following formula:

Z $=\dfrac{[B-(t-1)M]^2}{t(t-1)}$

After those adjustments, the ANOVA table is then constructed.

SV	df	SS	MS	F
Blocks	b – 1	SSb	MSb	
Treatments	t – 1	(SSt) – Z	MSt	$\frac{MSt}{MSE}$
Error	(b – 1)(t – 1)	SSE	MSE	
Total	(bt – 1) - 1	(SST) – Z		

It should be noted that the df blocks and SS blocks are not adjusted at all. Also, the df treatments and the SSE are not adjusted. The SSE is not adjusted because the estimated missing value makes no contribution to it.

Example: The data below (with one missing observation) were collected from a field experiment involving five treatments in three blocks.

TREATMENTS

		A	B	C	D	E
B l o c k s	I	2.62	2.67	1.04	1.88	1.74
	II	2.58	2.52	M	2.82	2.10
	III	2.65	1.92	0.91	2.65	2.78

M denotes the missing observation