An Introduction to Two-Way ANOVA

ANOVA : What is it?

- An ANOVA (Analysis of Variance), sometimes called an F test, is closely related to the t test.
- The major difference:

t - test	ANOVA tests
measures the difference between the means of two groups	measures the difference between the means of more than two groups

Two-Way ANOVA (Factorial ANOVA)

- Extension of one way ANOVA

- There are two independent variables (Hence the name two way)

- *Two-way ANOVA* is an extension of the paired *t* test to more than two treatments

What is two way independent ANOVA?

- Two Independent VariableS (IVs)
  - ✓ Two-way = 2 IVs
  - ✓ Three-way = 3 IVs
- Different participants in all conditions
  - ➢ Independent = "Different Participants"
- Several independent variables is known as a factroial design

Two-Way ANOVA

- **"Two-Way" means groups are defined by 2 independent variables (IVs)**
- **These IVs are typically called *factors***
- **With 2-Way ANOVA, there are two main effects and 1 interaction, so there are 3 *F* tests**
- **All, some, or none may be significant**

- there are three types of ANOVA analysis available:

  1) Single Factor ANOVA

  2) Two-Factor ANOVA Without Replication

  3) Two-Factor ANOVA with Replication

Each ANOVA test type is explained below:

Single Factor ANOVA

Single Factor ANOVA tests the effect of just one factor.

Example: the teaching method, on the measured outputs. The measured outputs are the mean test scores for the groups that had the different teaching methods applied to them. The Null Hypothesis for this one factor states that varying that factor has no effect on the outcome.

Two-Factor ANOVA Without Replication

Two-Factor ANOVA Without Replication - Allows testing of the original factor plus one other factor. For example, in addition to testing teaching methods, you could also test an additional factor, such as whether differences in teaching ability caused additional variation in the outcome of test average scores. Each factor has a Null Hypothesis which states that varying that factor had no effect on the outcome.

Two-Factor ANOVA With Replication

Two-Factor ANOVA With Replication allows for testing both factors as above. This method also allows us to test the effect of interaction between the factors upon the measured outcome. The test is replicated in two places. This allows for analysis of whether the interaction between the two factors has an effect on the measured outcome. The Null Hypothesis for this interaction test states that varying the interaction between the two factors has no effect on the measured outcome. Each of the other two factors being tested also has its own Null Hypothesis.

Logic of One Way ANOVA

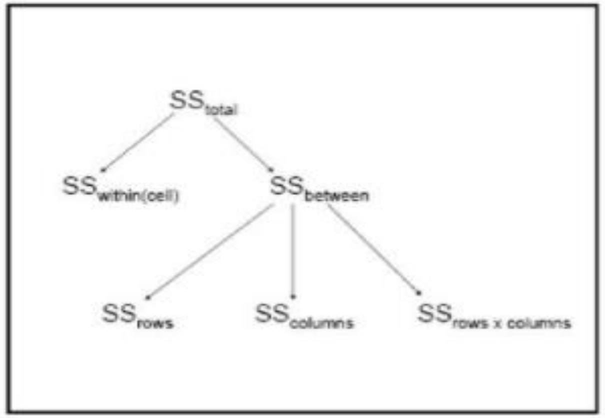
One Way ANOVA

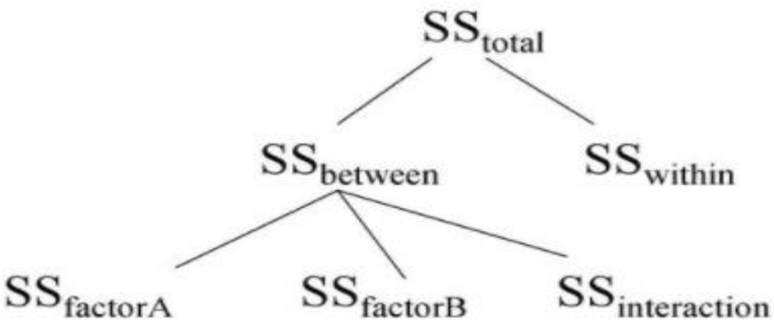
$$SS_{total}$$

$$SS_{between} \qquad\qquad SS_{within}$$

Logic of Two Way ANOVA

Two Way ANOVA





Main Effect 1

$$F = \frac{s^2_{factorA}}{s^2_w}$$

Main Effect 2

$$F = \frac{s^2_{factorB}}{s^2_w}$$

Interaction

$$F = \frac{s^2_{interaction}}{s^2_w}$$

Assumptions for the Two Factor ANOVA

1. Observations within each sample are independent
2. Populations are normally or approximately normally distributed
3. Populations from which the samples are selected must have equal variances (homogeneity of variance)
4. The groups must have the same sample size

Factors

- The two independent variables in a two-way ANOVA are called factors
- The idea is that there are two variables, factors, which affect the dependent variable
- Each factor will have two or more levels within it
- The degrees of freedom for each factor is one less than the number of levels

Hypotheses

- There are three sets of hypothesis with the two-way ANOVA
- The null hypotheses for each of the sets are given below

1. The population means of the first factor are equal. This is like the one-way ANOVA for the row factor

2. The population means of the second factor are equal. This is like the one-way ANOVA for the column factor

3. There is no interaction between the two factors

Treatment Groups

- Treatment Groups are formed by making all possible combinations of the two factors

- All treatment groups must have the same sample size for a two-way ANOVA

- For example,

  if the first factor has 3 levels and
  the second factor has 2 levels,
  Then, 3x2=6 different treatment groups

Main Effect

- The main effect involves the independent variables one at a time
- The interaction is ignored for this part
- Just the rows or just the columns are used, not mixed
- This is the part which is similar to the one-way analysis of variance

Interaction Effect

- The interaction effect is the effect that one factor has on the other factor

- The degrees of freedom here is the product of the two degrees of freedom for each factor

Within Variation

- The Within variation is the sum of squares within each treatment group

- The within variance = $\dfrac{\text{within variation}}{\text{its degrees of freedom}}$

- The within group is also called the error

F-Tests

- There is an F-test for each of the hypotheses
- The F-test is the mean square for each main effect and the interaction effect divided by the within variance

$$\text{F-test} = \frac{\text{mean square}}{\text{within variance}}$$

- The numerator df come from each effect, and the denominator df is the df for the within variance in each case

Advantages of two-way ANOVA

1. More efficient to study two factors (A and B) simultaneously, rather than separately

2. We can investigate interactions between factors (can investigate complex associations)

3. In a two way anova (A×B design) there are four sources of variations

1. Variation due to factorA
2. Variation due to factor B
3. Variation due to the interactive effect of A & B
4. Within cell (error) variation

BASIC TWO -WAY ANOVA TABLE

Source of variation	SS	df	MS [SS/df]	F	Pvalue	Fcrit
Main effect A						
Main effect B						
Interactive effect						
Within						
Total						

TWO – WAY ANOVA WITHOUT REPLICATION WITH EXAMPLE

A Comparison of Dissolution of Various Tablet Formulations:

- **Eight laboratories were requested to participate in an experiment**
- **Objective :**

  To compare the dissolution rates of two generic products and a standard drug product
- **Purpose:**

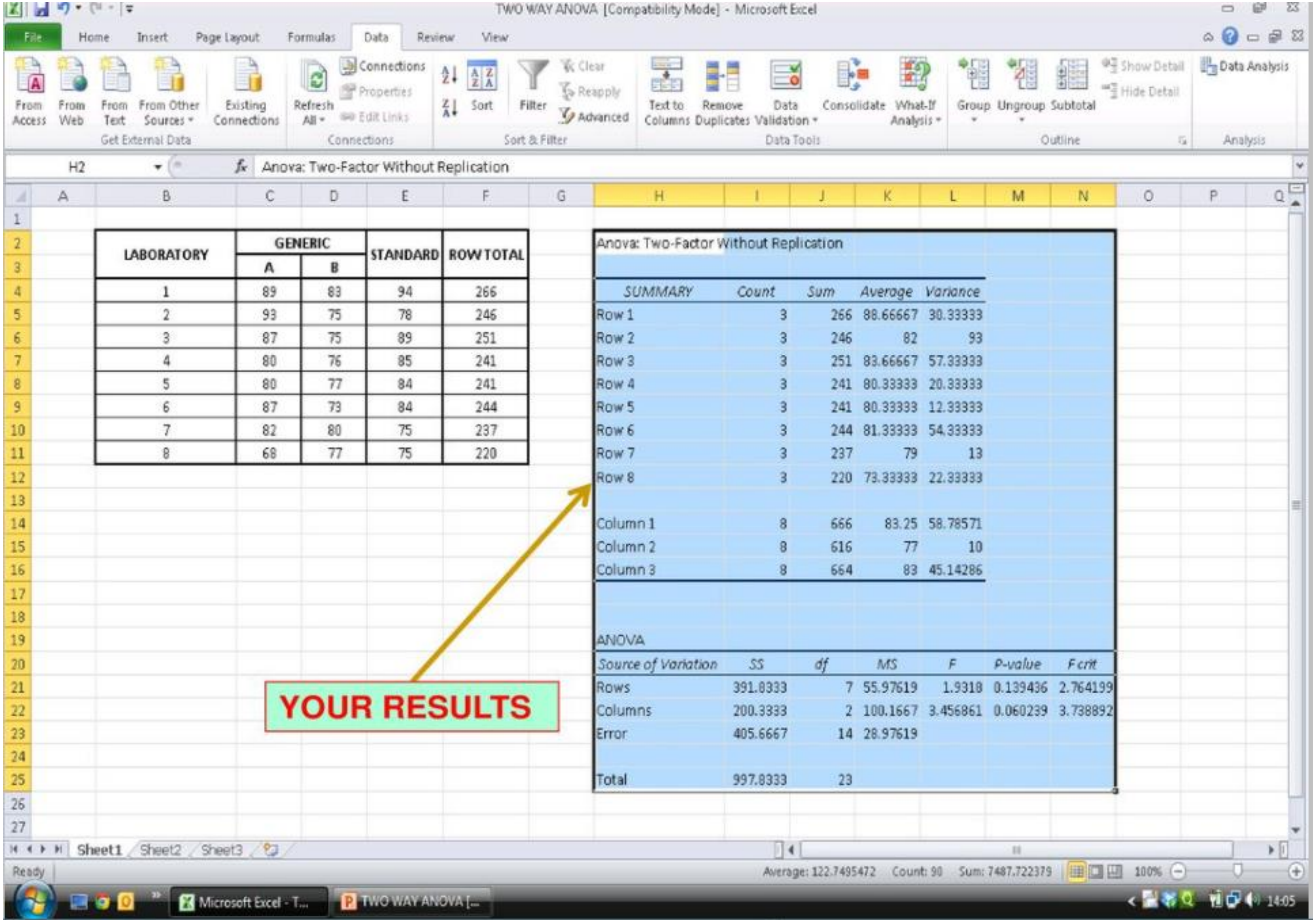
(a) To determine if the products had different rates of dissolution, and

(b) To estimate the laboratory variability (differences)

c) Interaction effect between lab & product

Tablet Dissolution After 30 Min for Three Products (% Dissolution)

LABORATORY	GENERIC		STANDARD	ROW TOTAL
	A	B		
1	89	83	94	266
2	93	75	78	246
3	87	75	89	251
4	80	76	85	241
5	80	77	84	241
6	87	73	84	244
7	82	80	75	237
8	68	77	75	220



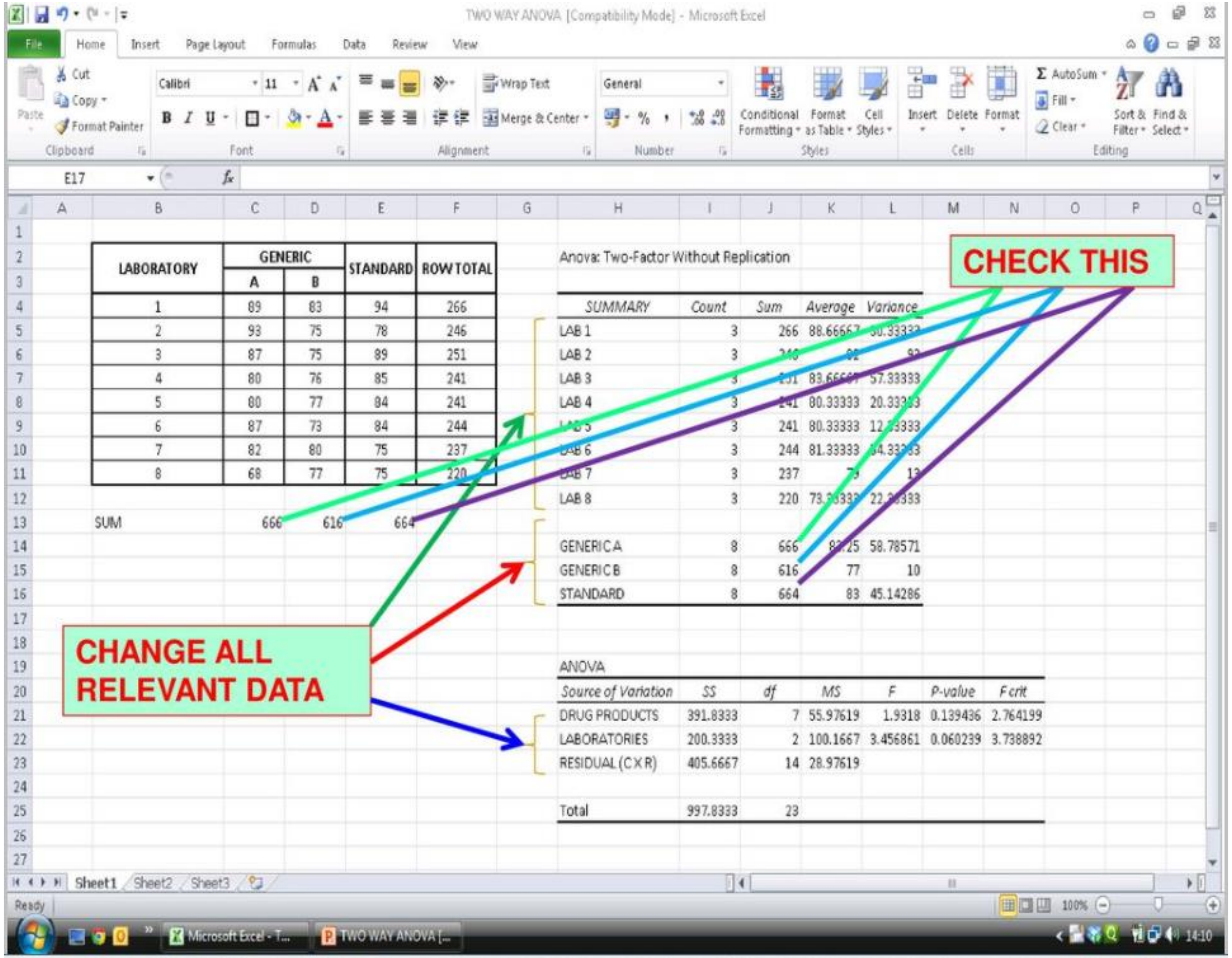
LABORATORY	GENERIC		STANDARD	ROW TOTAL
	A	B		
1	89	83	94	266
2	93	75	78	246
3	87	75	89	251
4	80	76	85	241
5	80	77	84	241
6	87	73	84	244
7	82	80	75	237
8	68	77	75	220

Anova: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance
Row 1	3	266	88.66667	30.33333
Row 2	3	246	82	93
Row 3	3	251	83.66667	57.33333
Row 4	3	241	80.33333	20.33333
Row 5	3	241	80.33333	12.33333
Row 6	3	244	81.33333	54.33333
Row 7	3	237	79	13
Row 8	3	220	73.33333	22.33333
Column 1	8	666	83.25	58.78571
Column 2	8	616	77	10
Column 3	8	664	83	45.14286

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	391.8333	7	55.97619	1.9318	0.139436	2.764199
Columns	200.3333	2	100.1667	3.456861	0.060239	3.738892
Error	405.6667	14	28.97619			
Total	997.8333	23				


TWO WAY ANOVA [Compatibility Mode] - Microsoft Excel
LABORATORY
GENERIC
A
B
STANDARD
ROW TOTAL
1 89 83 94 266
2 93 75 78 246
3 87 75 89 251
4 80 76 85 241
5 80 77 84 241
6 87 73 84 244
7 82 80 75 237
8 68 77 75 220
SUM 666 616 664
Anova: Two-Factor Without Replication
SUMMARY Count Sum Average Variance
LAB 1 3 266 88.66667 30.33333
LAB 2 3 246 82 93
LAB 3 3 251 83.66667 57.33333
LAB 4 3 241 80.33333 20.33333
LAB 5 3 241 80.33333 12.33333
LAB 6 3 244 81.33333 34.33333
LAB 7 3 237 79 13
LAB 8 3 220 73.33333 22.33333
GENERIC A 8 666 83.25 58.78571
GENERIC B 8 616 77 10
STANDARD 8 664 83 45.14286
ANOVA
Source of Variation SS df MS F P-value F crit
DRUG PRODUCTS 391.8333 7 55.97619 1.9318 0.139436 2.764199
LABORATORIES 200.3333 2 100.1667 3.456861 0.060239 3.738892
RESIDUAL (C X R) 405.6667 14 28.97619
Total 997.8333 23
CHECK THIS
CHANGE ALL RELEVANT DATA
Sheet1 Sheet2 Sheet3

Total sum of squares (TSS)

$$= \sum X^2 - \text{C.T.} = 89^2 + 93^2 + \ldots + 75^2 + 75^2 - \frac{(1946)^2}{24}$$
$$= 158,786 - 157,788.2 = 997.8$$

Column sum of squares (CSS) or product SS

$$= \frac{\sum C_j^2}{R} - \text{C.T.} = \frac{(666^2 + 616^2 + 664^2)}{8} - 157,788.2$$
$$= 200.3\ (C_j \text{ is the total of column } j,\ R \text{ is the number of rows})$$

Row sum of squares (RSS) or laboratory SS

$$= \frac{\sum R_i^2}{C} - \text{C.T.} = \frac{(266^2 + 246^2 + \ldots + 220^2)}{3} - 157,788.2$$
$$= 391.8 (R_i \text{ is the total of row } i,\ C \text{ is the number of columns})$$

Residual ($C \times R$) sum of squares (ESS) = TSS − CSS − RSS

$$= 997.8 - 200.3 - 391.8 = 405.7$$

The ANOVA table is shown in Table 8.10. The *degrees of freedom* are calculated as follows:

$$\text{Total} = N_t - 1 \qquad N_t = \text{total number of observations}$$

$$\text{Column} = C - 1 \qquad C = \text{number of columns}$$

$$\text{Row} = R - 1 \qquad R = \text{number of rows}$$

$$\text{Residual } (C \times R) = (C-1)(R-1)$$

Tests of Significance

To test for differences among *products* (H_0: $\mu_A = \mu_B = \mu_C$), an F ratio is formed:

$$\frac{\text{drug product MS}}{\text{residual MS}} = \frac{100.2}{29} = 3.5$$

The F distribution has 2 and 14 d.f. According to Table IV.6, an F of 3.74 is needed for significance at the 5% level. Therefore, the products are not significantly different at the 5% level.

Table 8.10 Analysis of Variance Table for the Data (Dissolution) from Table 8.8

Source	d.f.	SS	MS	F^a
Drug products	2	200.3	100.2	$F_{2,14} = 3.5$
Laboratories	7	391.8	56.0	$F_{7,14} = 1.9$
Residual ($C \times R$)	14	405.7	29.0	
Total	23	997.8		

TWO – WAY ANOVA WITH REPLICATION

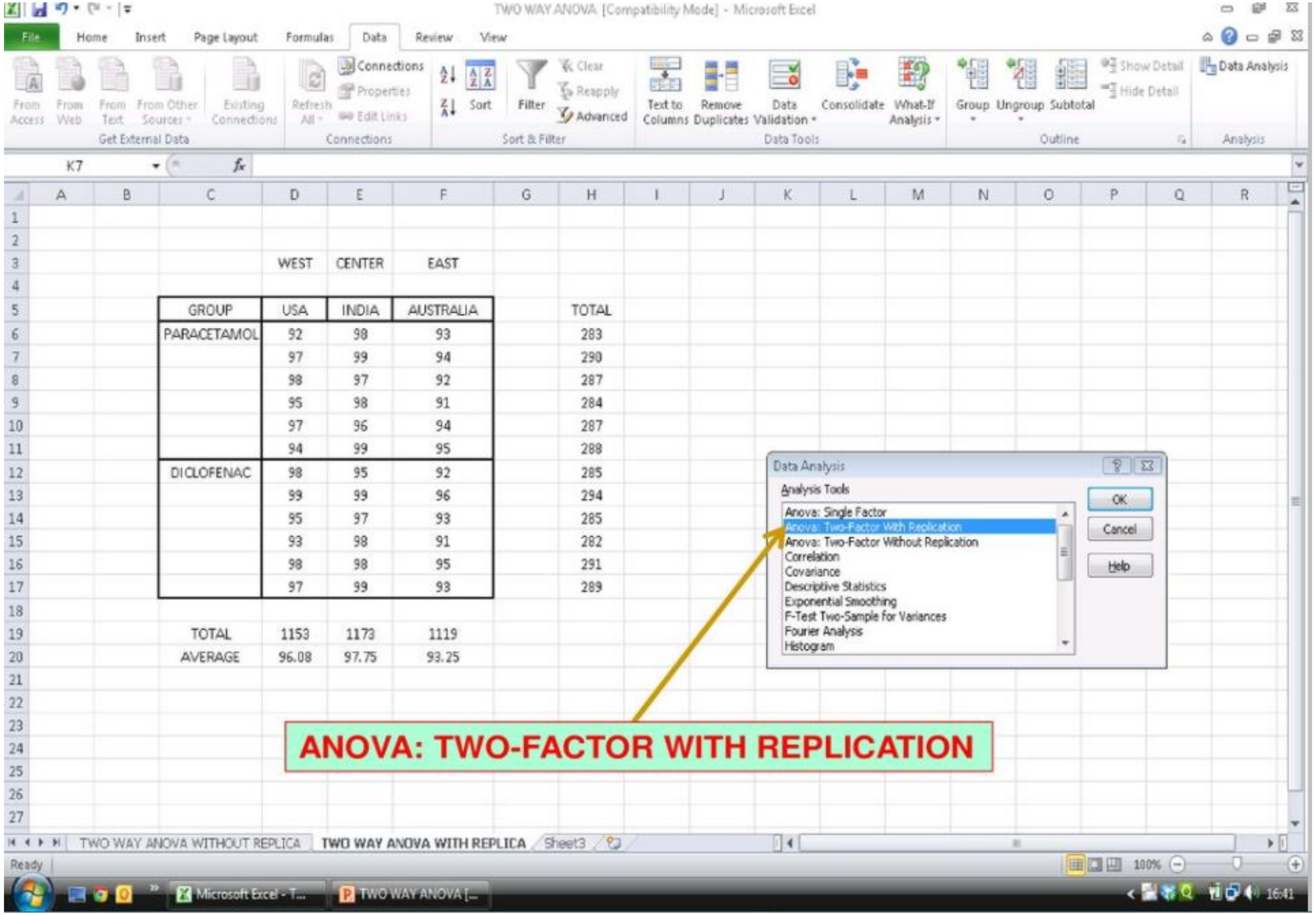
Before discussing an example of the analysis of two-way designs with replications, two points should be addressed regarding the implementation of such experiments.

1. It is best to have equal number of replications for each cell of the two-way design. In the dissolution example, this means that each lab replicates each formulation an equal number of times. If the number of replicates is very different for each cell, the analysis and interpretation of the experimental results can be very complicated and difficult.
2. The experimenter should be sure that the experiment is *properly* replicated. As noted above, merely replicating assays on the same tablet is not proper replication in the dissolution example. Replication is an independently run sample in most cases. Each particular experiment has its own problems and definitions regarding replication. If there is any doubt about what constitutes a proper replicate, a statistician should be consulted.

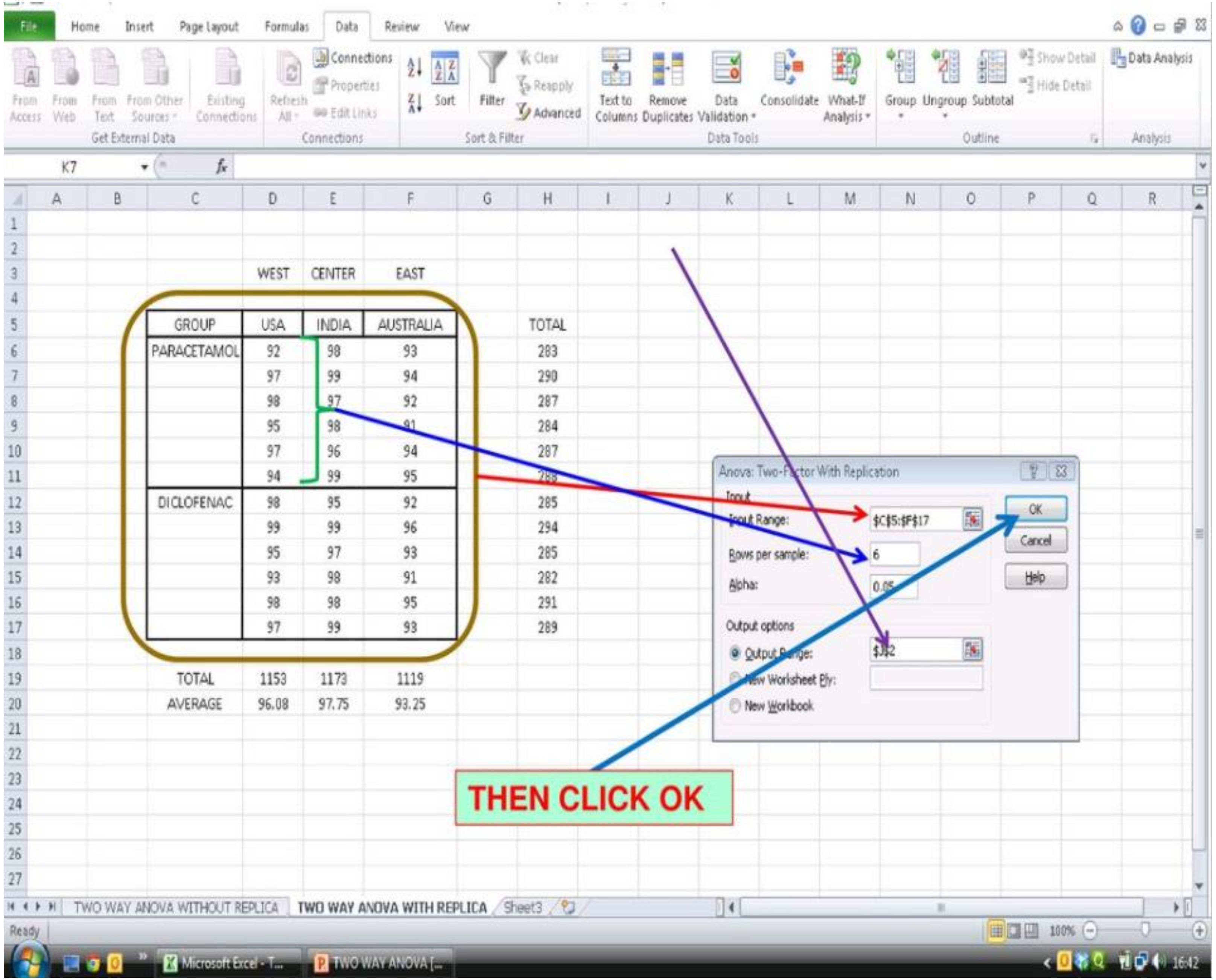
Replicate tablet dissolution data for eight laboratories testing three products (Percent distribution)

- If we take an example that we had test content uniformity for Paracetamol and Diclofenac from USA, INDIA and AUSTRALIA.
- We will take 6 samples from each contry from 6 different places.
- The table can be summarized as follows:

GROUP	WEST	CENTER	EAST
	USA	INDIA	AUSTRALIA
PARACETAMOL	92	98	93
	97	99	94
	98	97	92
	95	98	91
	97	96	94
	94	99	95
DICLOFENAC	98	95	92
	99	99	96
	95	97	93
	93	98	91
	98	98	95
	97	99	93



	WEST	CENTER	EAST	
GROUP	USA	INDIA	AUSTRALIA	TOTAL
PARACETAMOL	92	98	93	283
	97	99	94	290
	98	97	92	287
	95	98	91	284
	97	96	94	287
	94	99	95	288
DICLOFENAC	98	95	92	285
	99	99	96	294
	95	97	93	285
	93	98	91	282
	98	98	95	291
	97	99	93	289
TOTAL	1153	1173	1119	
AVERAGE	96.08	97.75	93.25	



File Home Insert Page Layout Formulas Data Review View
WEST CENTER EAST
GROUP USA INDIA AUSTRALIA TOTAL
PARACETAMOL 92 98 93 283
97 99 94 290
98 97 92 287
95 98 91 284
97 96 94 287
94 99 95 288
DICLOFENAC 98 95 92 285
99 99 96 294
95 97 93 285
93 98 91 282
98 98 95 291
97 99 93 289
TOTAL 1153 1173 1119
AVERAGE 96.08 97.75 93.25
Anova: Two-Factor With Replication
Input
Input Range: $C$5:$F$17
Rows per sample: 6
Alpha: 0.05
Output options
Output Range:
New Worksheet Ply:
New Workbook
OK
Cancel
Help
THEN CLICK OK
TWO WAY ANOVA WITHOUT REPLICA
TWO WAY ANOVA WITH REPLICA
Sheet3

	WEST	CENTER	EAST	
GROUP	USA	INDIA	AUSTRALIA	TOTAL
PARACETAMOL	92	98	93	283
	97	99	94	290
	98	97	92	287
	95	98	91	284
	97	96	94	287
	94	99	95	288
DICLOFENAC	98	95	92	285
	99	99	96	294
	95	97	93	285
	93	98	91	282
	98	98	95	291
	97	99	93	289
TOTAL	1153	1173	1119	
AVERAGE	96.08	97.75	93.25	

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	1.361111111	1	1.361111	0.420240137	0.521748525	4.170876786
Columns	124.2222222	2	62.11111	19.17667238	0.00000432	3.315829501
Interaction	2.888888889	2	1.444444	0.445969125	0.644379106	3.315829501
Within	97.16666667	30	3.238889			
Total	225.6388889	35				

Anova: Two-Factor With Replication

SUMMARY	**USA**	**INDIA**	**AUSTRALIA**	**Total**
PARACETAMOL				
Count	6	6	6	18
Sum	573	587	559	1719
Average	95.5	97.83333	93.16666667	95.5
Variance	5.1	1.366667	2.166666667	6.382353
DICLOFENAC				
Count	6	6	6	18
Sum	580	586	560	1726
Average	96.66667	97.66667	93.33333333	95.88889
Variance	5.066667	2.266667	3.466666667	6.810458
Total				
Count	12	12	12	
Sum	1153	1173	1119	
Average	96.08333	97.75	93.25	
Variance	4.992424	1.659091	2.568181818	

TWO WAY ANOVA WITHOUT REPLICA | TWO WAY ANOVA WITH REPLICA | Sheet3