Estimation of precision

S Yasar

Institute for Reference Materials and Measurements (IRMM)
Geel, Belgium

http://www.irmm.jrc.be
http://www.jrc.ec.europa.eu

Objective

To illustrate a simple-fast way of estimating precision in some analytical cases.

using a statistic program:MINITAB.

while complying ISO rules .

Validasyon

- **Validasyon (CLSI): “Bir sistem veya yöntemin beklendiği şekilde çalıştığını kanıtlama eylemi veya sürecidir.”**

- **• Validasyon (WHO-BS/95.1793): Bir metodun, sürecin, cihazın veya sistemin beklendiği gibi çalıştığını ve istenen sonuçları verdiğini kanıtlama işidir.**

Kriterler

- Doğruluk (Accuracy)
- Kesinlik (Precision)
- Özgünlük (Specifity)
- Doğrusallık (Linearity)
- Gözlenebilme sınırı (LOD; Limit of Detection)
- Tayin sınırı (LOQ; Limit of Quantification)
- Çalışma Aralığı (Range)
- Tutarlılık (Ruggedness)
- Sağlamlık ya da Kararlılık (Robustness)
- Stabilite

Objective

To illustrate a simple-fast way of estimating precision in some analytical cases.

using a statistic program:MINITAB.

while complying ISO rules .

CONTENT

- Definition & precision conditions
- Statistical model of estimating precision
- List of ISOs concerning the estimation of precision.
- Applications in Statistical Program: MINITAB

1, 2, 3, 4 Analytical cases

  - Defining model
  - Data entry (Minitab)
  - Run (how to proceed)
  - Output (results)
  - Summary (interpretative) : Estimated variance components and calculated repeatability and reproducibility StDev.

- Conclusion

*Precision:

*ISO 5725-3:1994

The closeness of agreement between independent test results obtained under the stipulated conditions :

Repeatability condition:
the same method
identical test items
the same operator & equipment
within short intervals of time.

Reproducibility condition:
the same method
different laboratories
different operators
different equipments.

Intermediate precision condition :

the same method in the same laboratory

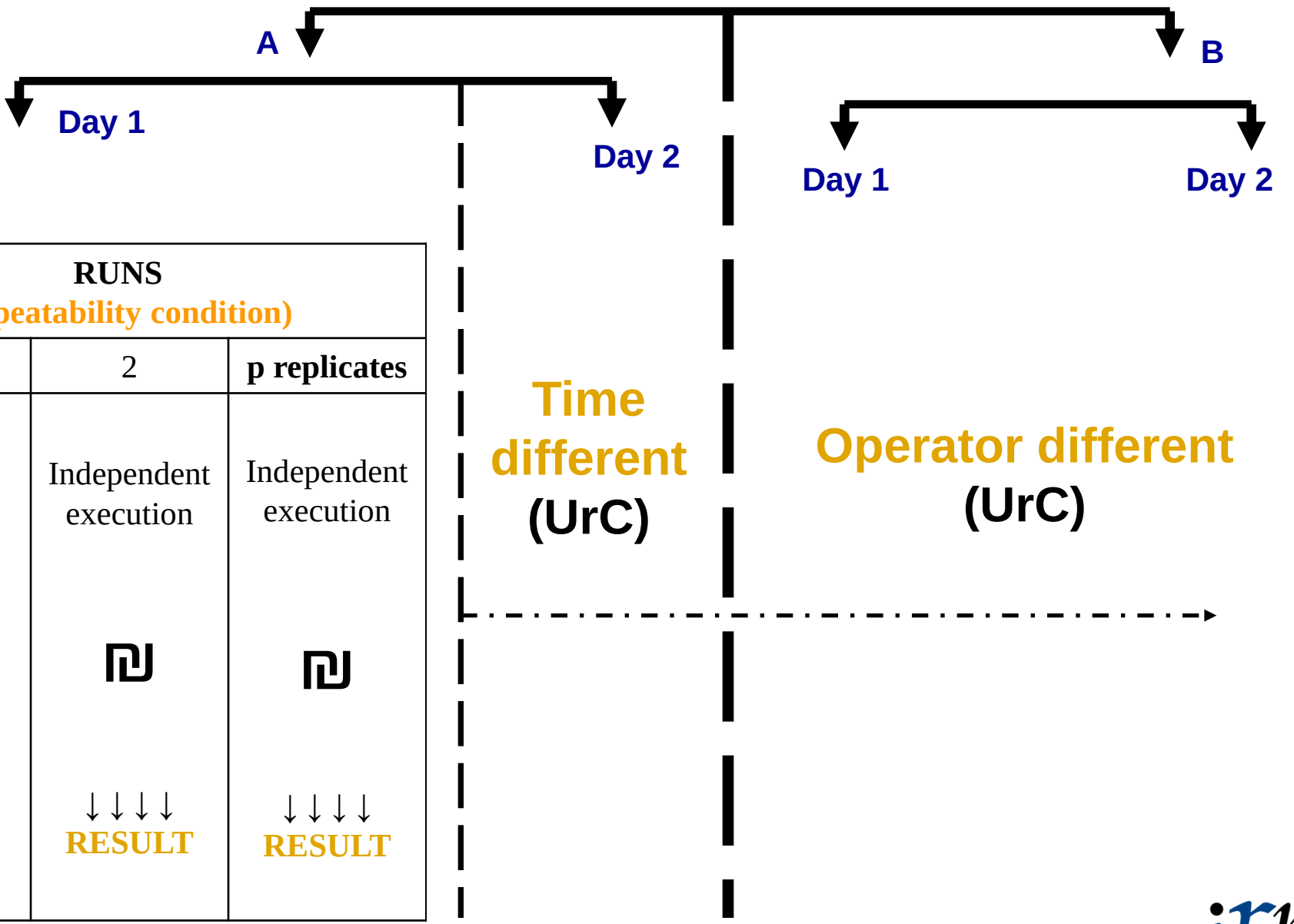
under 1, 2 or 3 factor different conditions

(operator different)

(time different)

(equipment different)

In-house variability conditions



RUNS (Repeatability condition)		
1	2	**p replicates**
Samp.Preap ↓ Extraction ↓ Cleanup ↓ **Injections** 1, 2, 3….**n** ↓↓↓↓ **RESULT**	Independent execution ₪ ↓↓↓↓ **RESULT**	Independent execution ₪ ↓↓↓↓ **RESULT**

Basic Statistical MODEL

$$Y = m + B + e$$

Y, *test result*

- **m,** *general mean (expectation)*
- **B,** *laboratory component of bias under repeatability condition (UrC)*
- **e,** *random error occurring in every measurement UrC*

The discussion of these three components and extentions of the basic model can be found in the respective ISO 5725:1994 series.

TERM B- LABORATORY COMPONENT OF BIAS
INTERMEDIATE PRECISION CONDITION

The basic model is to be extended to cover a number of effects which are due to changes in operator, equipment, time:

Then the model changes to

$$Y = m + B_{(operator)} + B_{(equipment)} + B_{(days)} + \ldots + e$$

The models and their variance components

$Y = m + B + e$ →

Repeatability standard deviation

$S_r = \sqrt{var\ (e)}$

$Y = m + B + e$ →

Reproducibility standard deviation

$S_R = \sqrt{var\ (B) + var\ (e)}$

$Y = m + Bo + B(1) + B(2) + B(3) + \ldots + e$

Intermediate precision

$Sinter = \sqrt{var\ (Bo) + var\ (B1) \ldots . + var\ (e)}$

HOW to estimate?

ISO sets rules of calculating precision for various models

- **ISO 5725-1:** Definitions: Accuracy, trueness and precision.
- ***ISO 5725-2:*** *Basic method estimating r and R stdv.*
- ***ISO 5725-3:*** *Defines intermediate precision (Nested designs).*
- **ISO 5725-4:** Defines basic method determining trueness
- *ISO 5725-5: Estimates precision in split-plot experiments.*
- **ISO 5725-6:** Use of Accuracy values in practice.

APPLICATION
IN STATISTICAL PROGRAM

Statistical Programs

- SAS:
- Advanced,
- easy,
- May be friendly (but specific commands must be known)

- Statistica, SPSS:
- Advanced,
- complicated,
- Previous experience is needed

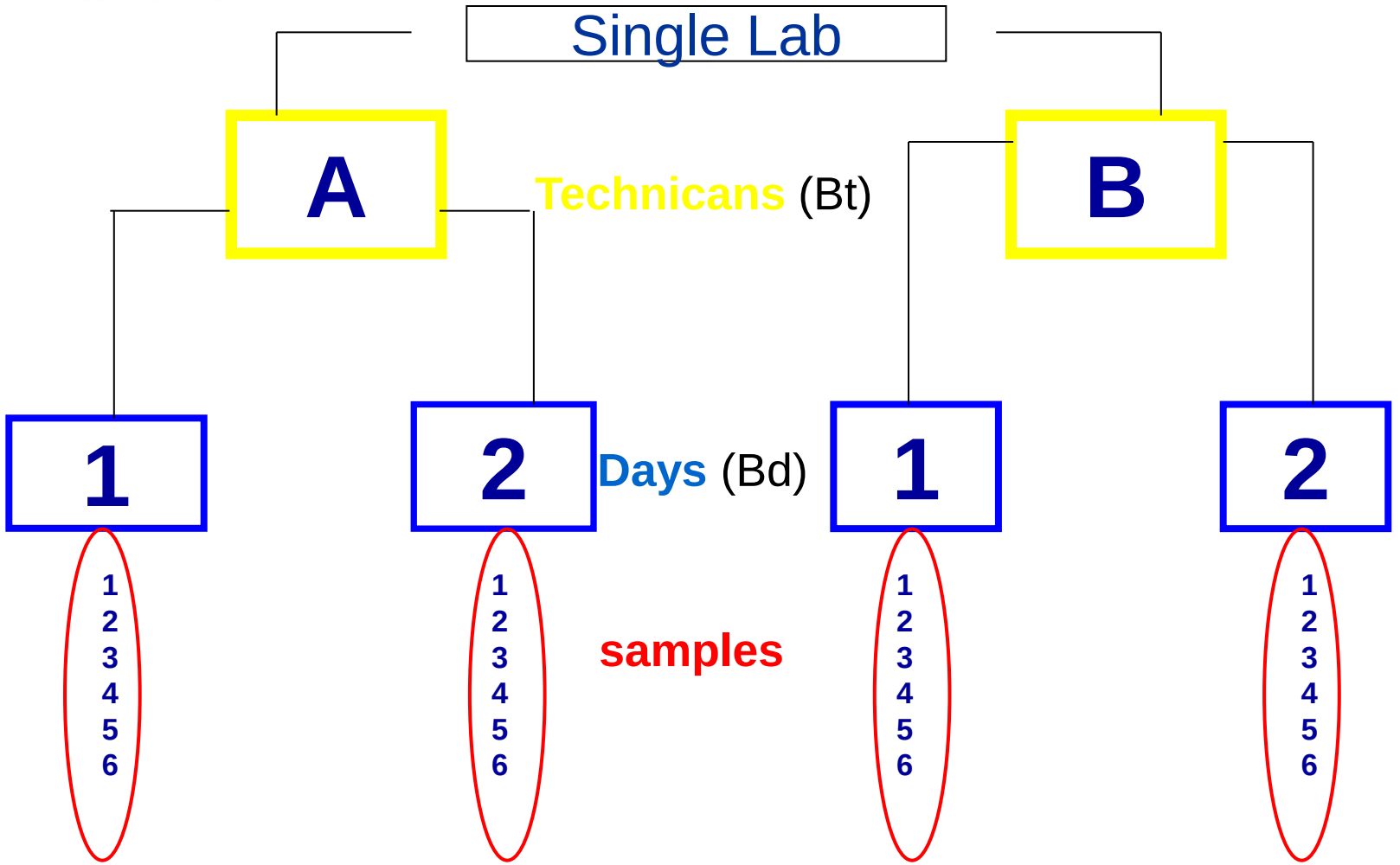
- MINITAB
- Not advanced,
- easy,
- user friendly (no experience)

MINITAB INC™
Making Data Analysis Easier

Case 1: A GC method to determine GTH (glyceroltriheptanoate) in MBM samples

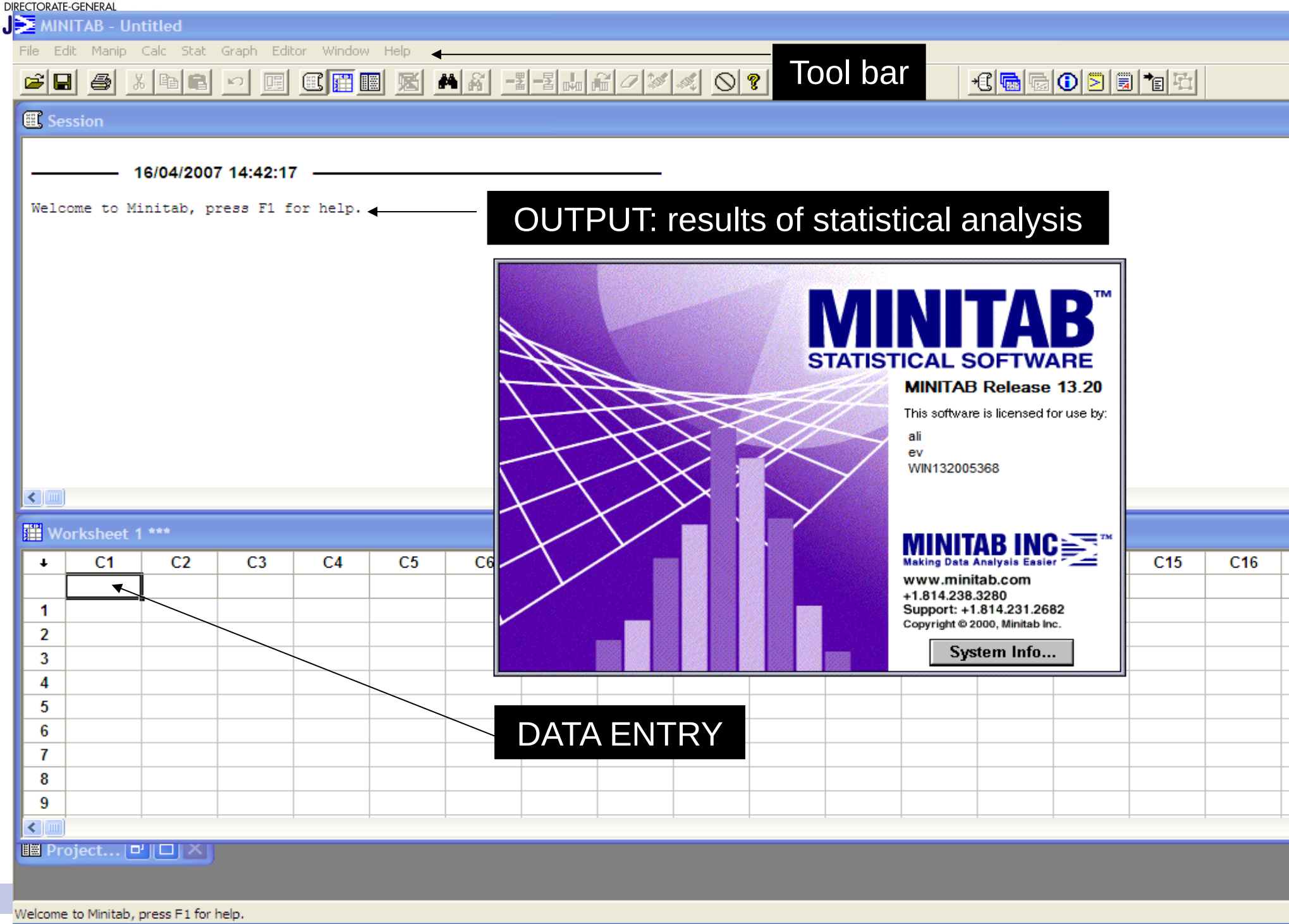
RECOVERY, %		Technician 1		Technician 2	
Conc, ppm	Replicates	Day 1	Day 2	Day 1	Day 2
10ppm	Sample 1	94.8	107.0	102.8	111.3
	Sample 2	105.5	116.3	106.6	127.2
	Sample 3	108.7	119.1	106.4	111.7
	Sample 4	113.4	123.3	110.6	115.2
	Sample 5	108.9	119.3	109.4	115.5
	Sample 6	102.9	114.1	108.9	114.5
100ppm	Sample 1	106.5	95.6	102.7	99.6
	Sample 2	108.5	97.4	101.0	100.3
	Sample 3	108.0	97.0	103.4	100.2
	Sample 4	105.1	94.4	103.9	101.7
	Sample 5	96.2	86.2	99.7	98.4
	Sample 6	104.8	94.2	100.1	103.2

$$Y = m + \underline{Bo}_{technician} + \underline{B(1)}_{days} \ldots + \underline{e}$$



the model: 3 factors fully nested experiment-*ISO 5725-3*

APPLYING INTO MINITAB



DATA ENTRY

MINITAB - RECOVERY RATE.MPJ - [Worksheet 1(DOSAGE = 10) ***]

save the file

File Edit Manip Calc Stat Graph Editor Window Help

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C
	DOSAGE	TECHNICIAN	DAYS	RECOVERY												
1	10	1	1	94.8												
2	10	1	1	105.5												
3	10	1	1	108.7												
4	10	1	1	113.4												
5	10	1	1	108.9												
6	10	1	1	102.6												
7	10	1	2	107.0												
8	10	1	2	116.3												
9	10	1	2	119.1												
10	10	1	2	123.3												
11	10	1	2	119.3												
12	10	1	2	114.1												
13	10	2	1	102.8												
14	10	2	1	106.6												
15	10	2	1	106.4												
16	10	2	1	110.6												
17	10	2	1	109.4												
18	10	2	1	108.9												
19	10	2	2	111.3												
20	10	2	2	127.2												
21	10	2	2	111.7												
22	10	2	2	115.2												
23	10	2	2	115.5												
24	10	2	2	114.5												

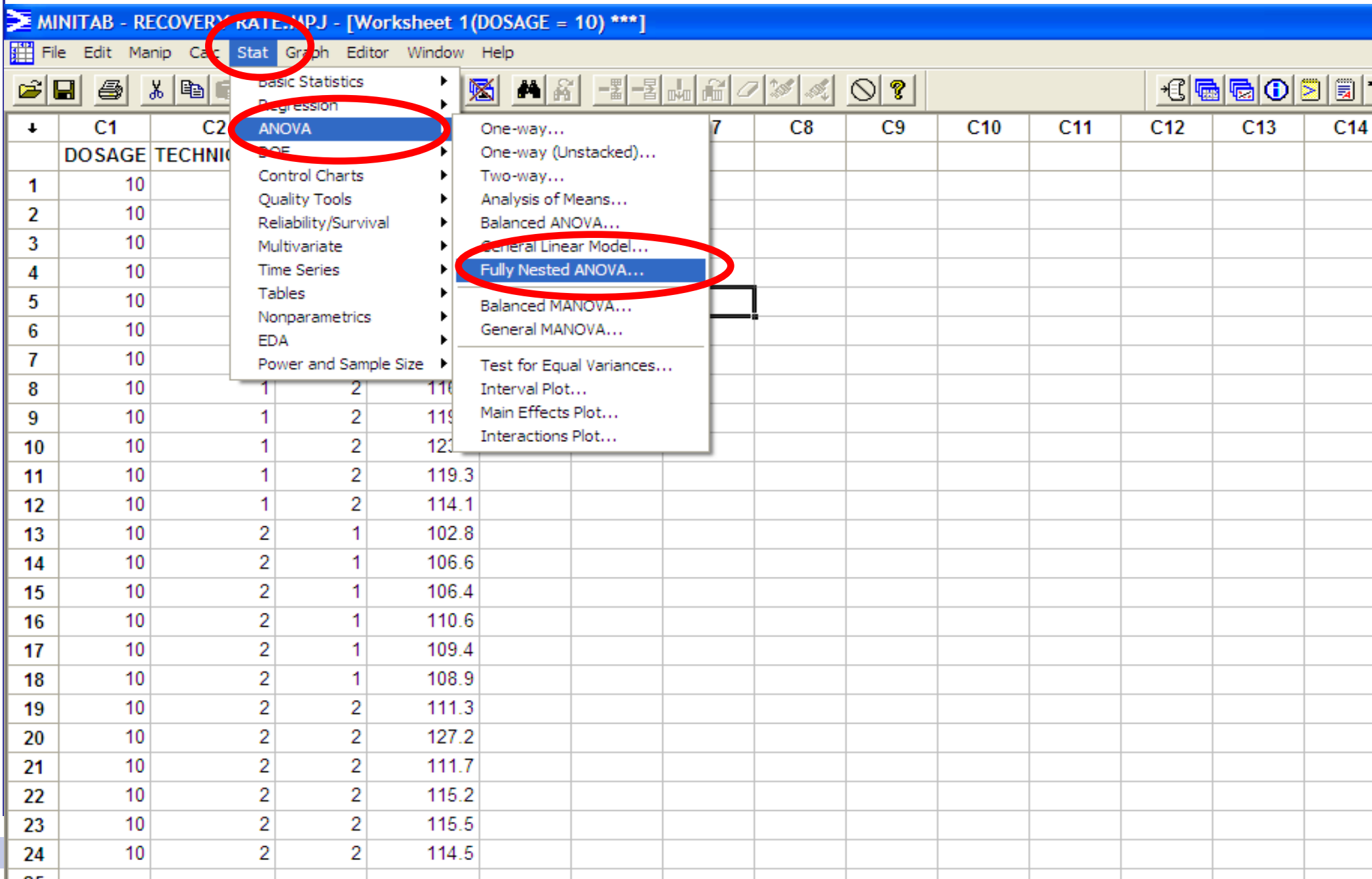
Write down the names of your variables IN EACH CELL by clicking on each cell

Defining each level of variables: TEXT or NUMERICAL

Response: Numeric

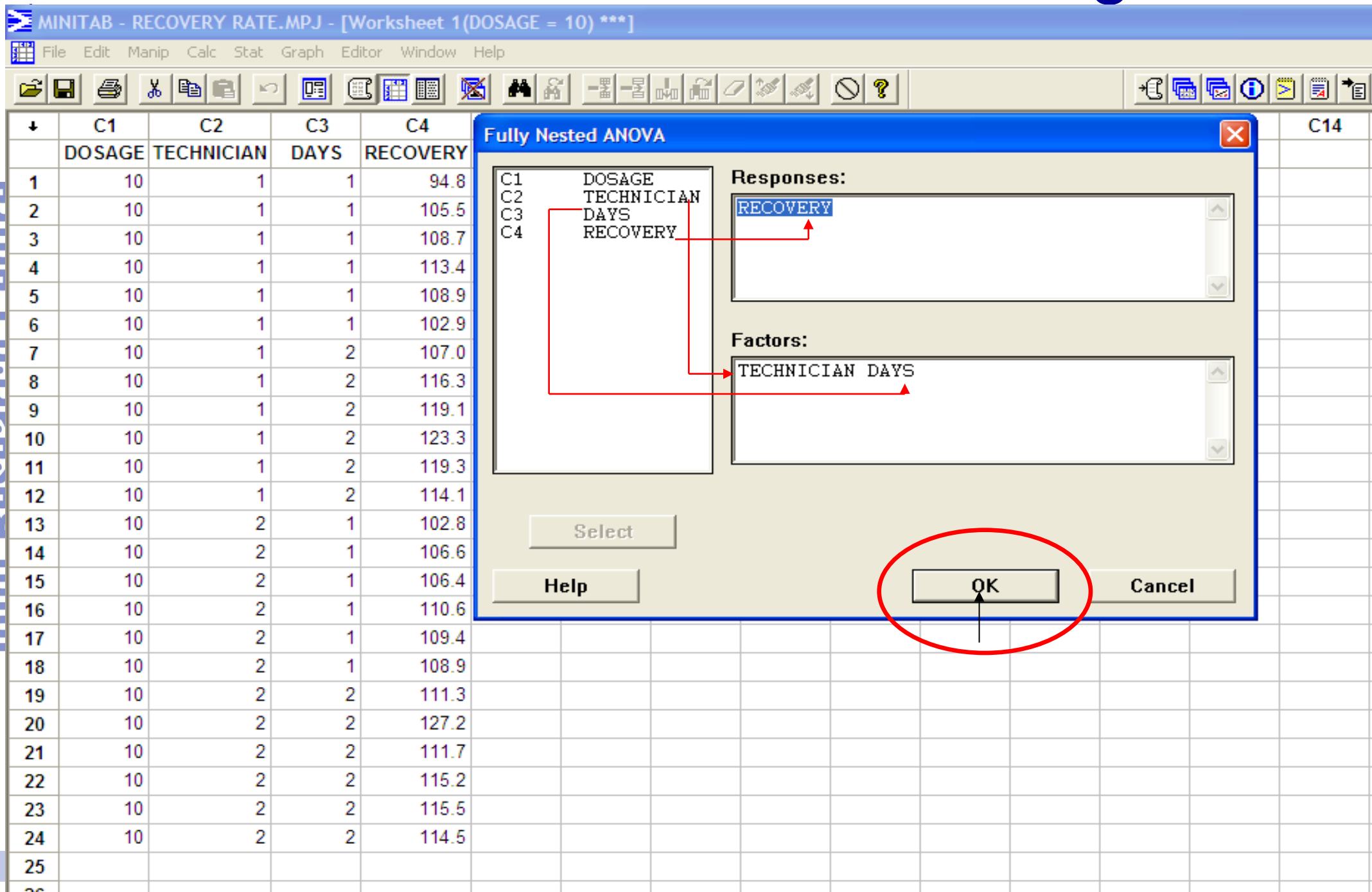
•**Precision after outlier identification and elimination.**

ANALYSIS



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	DOSAGE	TECHNI												
1	10													
2	10													
3	10													
4	10													
5	10													
6	10													
7	10													
8	10		1	2	116									
9	10		1	2	119									
10	10		1	2	123									
11	10		1	2	119.3									
12	10		1	2	114.1									
13	10		2	1	102.8									
14	10		2	1	106.6									
15	10		2	1	106.4									
16	10		2	1	110.6									
17	10		2	1	109.4									
18	10		2	1	108.9									
19	10		2	2	111.3									
20	10		2	2	127.2									
21	10		2	2	111.7									
22	10		2	2	115.2									
23	10		2	2	115.5									
24	10		2	2	114.5									

Model: Nested design



	C1	C2	C3	C4
	DOSAGE	TECHNICIAN	DAYS	RECOVERY
1	10	1	1	94.8
2	10	1	1	105.5
3	10	1	1	108.7
4	10	1	1	113.4
5	10	1	1	108.9
6	10	1	1	102.9
7	10	1	2	107.0
8	10	1	2	116.3
9	10	1	2	119.1
10	10	1	2	123.3
11	10	1	2	119.3
12	10	1	2	114.1
13	10	2	1	102.8
14	10	2	1	106.6
15	10	2	1	106.4
16	10	2	1	110.6
17	10	2	1	109.4
18	10	2	1	108.9
19	10	2	2	111.3
20	10	2	2	127.2
21	10	2	2	111.7
22	10	2	2	115.2
23	10	2	2	115.5
24	10	2	2	114.5
25				

OUTPUT

MINITAB - RECOVERY RATE.MPJ - [Session]

File Edit Manip Calc Stat Graph Editor Window Help

16/04/2007 16:42:40

```
Welcome to Minitab, press F1 for help.
Retrieving project from file: C:\Program Files\MTBWIN\Data\RECOVERY RATE.MPJ
```

Results for: Worksheet 1(DOSAGE = 10)

Nested ANOVA: RECOVERY versus TECHNICIAN, DAYS

Analysis of Variance for RECOVERY

Source	DF	SS	MS	F	P
TECHNICI	1	1.9267	1.9267	0.007	0.942
DAYS	2	565.2083	282.6042	9.908	0.001
Error	20	570.4833	28.5242		
Total	23	1137.6183			

ANALYSIS OF VARIANCE

Variance Components

Source	Var Comp.	% of Total	StDev
TECHNICI	-23.390*	0.00	0.000
DAYS	42.347	59.75	6.507
Error	28.524	40.25	5.341
Total	70.871		8.418

ESTIMATION OF VARIANCE COMPONENTS

* Value is negative, and is estimated by zero.

Expected Mean Squares

```
 1 TECHNICI   1.00(3) +  6.00(2) + 12.00(1)
 2 DAYS       1.00(3) +  6.00(2)
 3 Error      1.00(3)
```

SUMMARY

Nested Design Anova table: Level 1

Sources of variance	SS	df	MS	Estimated Variance Components
Technici (Bo)	**1.927**	**1**	**1.927**	**MS=28.524 + 6(S^2_{B1}) + 12 (S^2_{Bo})**
DAYS (B1)	**565.21**	**2**	**282.604**	**MS=28.524 + 6(S^2_{B1})**
Residual error (*e*)	**570.48**	**20**	**28.524**	**MS= (S^2_e)=28.524**
Total	**1137.62**	**23**		

Estimated Variance components

Variance Source	Component	% of total	StDev	
Technici (S^2_{Bo})	**-23.390***	**0.0000**	**0.000**	
DAYS (S^2_{B1})	**42.347**	**59.75**	6.507	
Error (S^2_e) = S^2_r	**28.524**	**40.25**	**5.341**	**S_r repeatability Stdev**
Total ($S^2_e + S^2_{Bo} + S^2_{B1}$)=$S^2_{int}$	**70.871**		**8.418**	**S_{int} INTERMEDIATE PRECISION**

*Value is negative, and is estimated by zero ***ISO 5725-3:1994***

Case 2*

*ISO 5725-2 Annex B

Interlaboratory Study

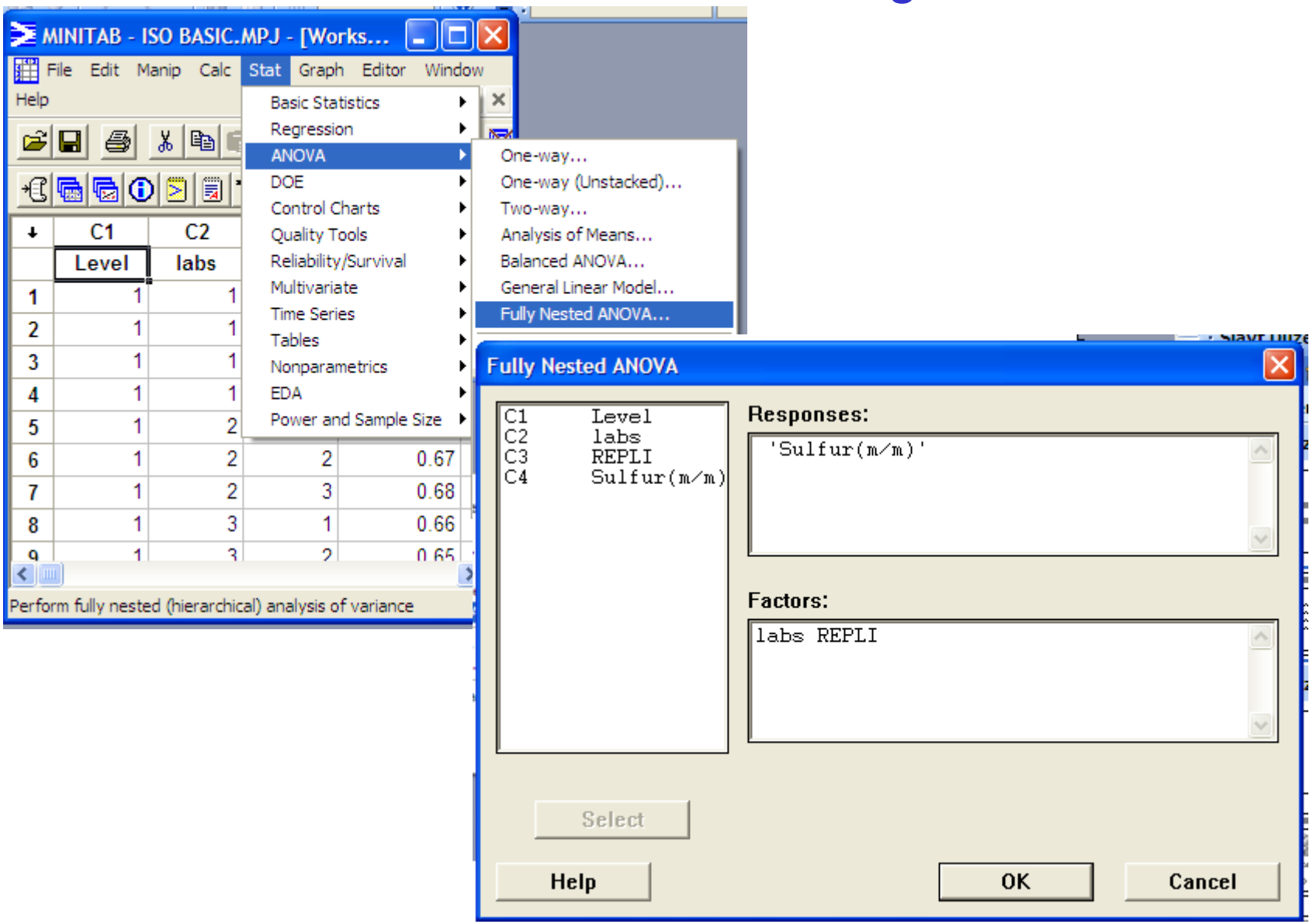
- 8 laboratories
- 4 COAL samples differing in Sulfur content
- The standardised analtyical method determined S content by mass (% m/m)

Labs	Sulfur Levels			
	Level 1	Level 2	Level 3	Level 4
1	0.71 0.71 0.70 0.71	4	4	4
2	3	3	3	3
3	3	3	3	3
4	3	3	3	3
5	5	4	5	5
6	3	3	3	3
7	3	3	3	3
8	3	3	3	3

$Y = m + \underline{B} + \underline{e}$

Precision must be estimated at each level of concentration

Model: Nested design



OUTPUT

MINITAB - ISO BASIC.MPJ - [Session]

File Edit Manip Calc Stat Graph Editor Window Help

Results for: Worksheet 1(Level = 1)

Nested ANOVA: Sulfur(m/m) versus labs, REPLI

Analysis of Variance for Sulfur(m

Source	DF	SS	MS
labs	7	0.0126	0.0018
REPLI	19	0.0043	0.0002
Total	26	0.0169	

Variance Components

Source	Var Comp.	% of Total	StDev
labs	0.000	67.12	0.022
REPLI	0.000	32.88	0.015
Total	0.001		0.026

Expected Mean Squares

1	labs	1.00(2) + 3.35(1)
2	REPLI	1.00(2)

SUMMARY

Nested Design Anova table: Level 1

Sources of variance	SS	df	MS	F	P
Between Lab (B)	0.01255	7	0.00179	7.85	0.000163
Within-Lab (*random error, e)*	0.00434	19	0.00023		
Total	0.0169	26			

Variance components

Variance Source	Component	% of total	StDev	
Lab (S^2B)	0.0004603	67.14	0.022	
Error (S^2_r) =	0.0002322	32.88	0.015	S_r repeatability Stdev
Total ($S^2_r + S^2B = S^2_R$	0.0006925		0.026	S_R reproducibility Stdev

Case 3

Vitamin A in Cattle Feed: Single lab

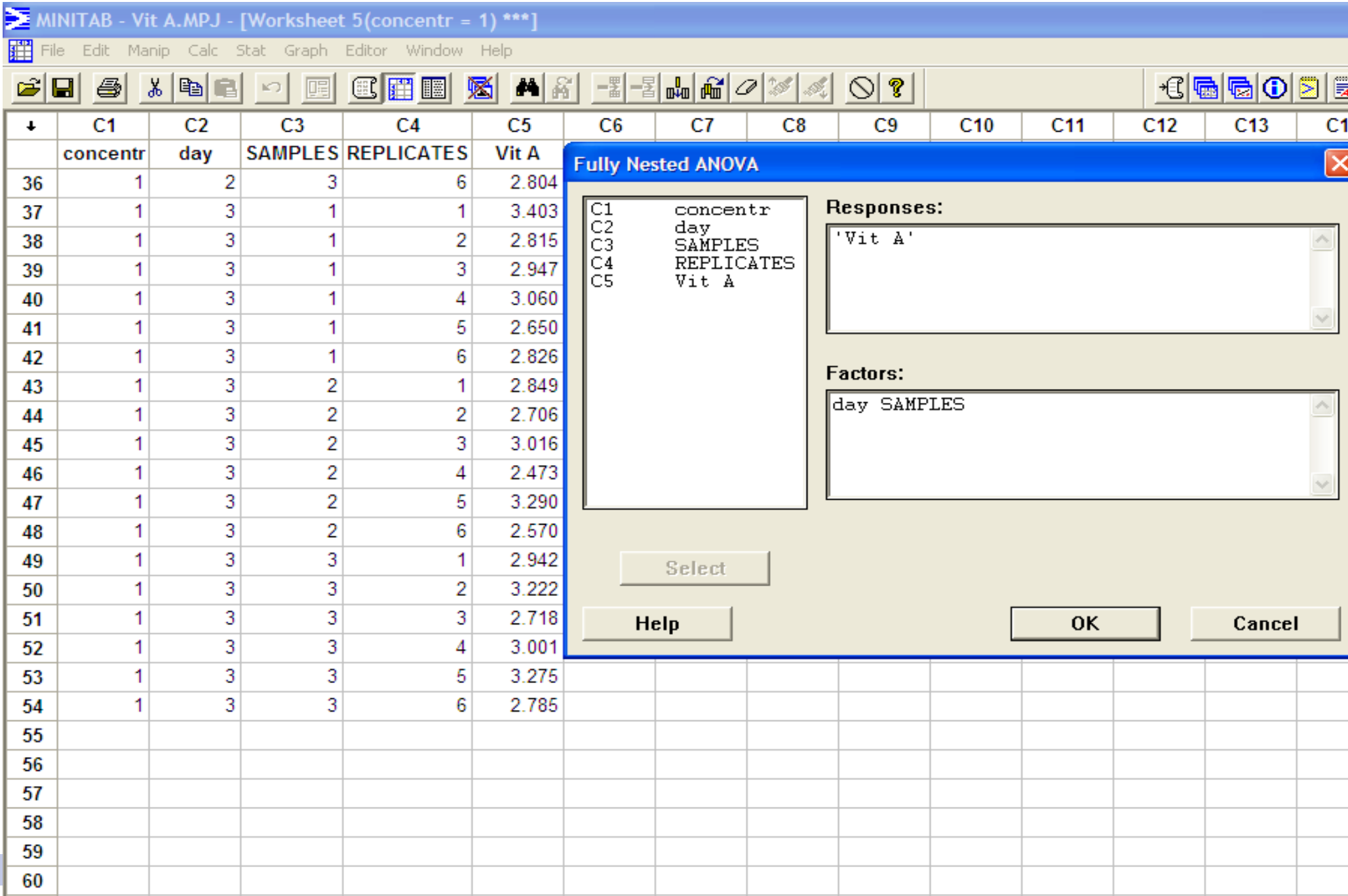
3 days, 3 analyte levels, 3 independent sample preparations/day, 6 injections for each one.

Days	Samples (replicates)	Injections	Vitamin A		
			@Level 1	@Level2	@Level 3
1	1	6	1	2	3
1	2	6	1	2	3
1	3	6	1	2	3
2	1	6	1	2	3
2	2	6	1	2	3
2	3	6	1	2	3
3	1	6	1	2	3
3	2	6	1	2	3
3	3	6	1	2	3

$$Y = m + \underline{Bo}_{days} + \underline{B(1)}_{samples} + \underline{e}$$

Precision must be estimated at each level of concentration

Model: Nested design



↓	C1	C2	C3	C4	C5
	concentr	day	SAMPLES	REPLICATES	Vit A
36	1	2	3	6	2.804
37	1	3	1	1	3.403
38	1	3	1	2	2.815
39	1	3	1	3	2.947
40	1	3	1	4	3.060
41	1	3	1	5	2.650
42	1	3	1	6	2.826
43	1	3	2	1	2.849
44	1	3	2	2	2.706
45	1	3	2	3	3.016
46	1	3	2	4	2.473
47	1	3	2	5	3.290
48	1	3	2	6	2.570
49	1	3	3	1	2.942
50	1	3	3	2	3.222
51	1	3	3	3	2.718
52	1	3	3	4	3.001
53	1	3	3	5	3.275
54	1	3	3	6	2.785
55					
56					
57					
58					
59					
60					

OUTPUT

MINITAB - Vit A.MPJ - [Session]

File Edit Manip Calc Stat Graph Editor Window Help

Nested ANOVA: Vit A versus day, SAMPLES

Analysis of Variance for Vit A

Source	DF	SS	MS	F	P
day	2	0.1426	0.0713	1.255	0.350
SAMPLES	6	0.3409	0.0568	1.719	0.138
Error	45	1.4873	0.0331		
Total	53	1.9708			

Variance Components

Source	Var Comp.	% of Total	StDev
day	0.001	2.13	0.028
SAMPLES	0.004	10.47	0.063
Error	0.033	87.40	0.182
Total	0.038		0.194

Expected Mean Squares

1	day	1.00(3) + 6.00(2) + 18.00(1)
2	SAMPLES	1.00(3) + 6.00(2)
3	Error	1.00(3)

SUMMARY

Nested Design Anova table: Level 1

Sources of variance	SS	df	MS	F	p
Days (Bo)	**0.1426**	**2**	**0.0713**	**1.255**	**0.350**
Sample (B1)	**0.3409**	**6**	**0.0568**	**1.719**	**0.138**
Residual error (*random error, e)*	**1.4873**	**45**	**0.0331**		
Total	**1.9708**	**53**			

Estimated Variance components

Variance Source	Component	% of total	StDev	
Days (S^2_{Bo})	**0.001**	**2.13**	**0.028**	
Sample (S^2_{B1})	**0.004**	**10.45**	0.063	
Error (S^2_e)	**0.033**	**87.40**	0.182	
$S^2_r = (S^2_e) + (S^2_{B1})$	**0.037**		0.192	S_r repeatability Stdev
Total $(S^2_e + S^2_{Bo} + S^2_{B1}) = S^2_{int}$	**0.038**		0.194	S_{int} INTERMEDIATE PRECISION

Case 4 In-house validation study LC-MS/MS method determining 7 coccidiostats

- Cattle Feeds — Poultry
- low, intermediate, high — L, I, H
- Time different (3 subsequent days) — 1, 2, 3

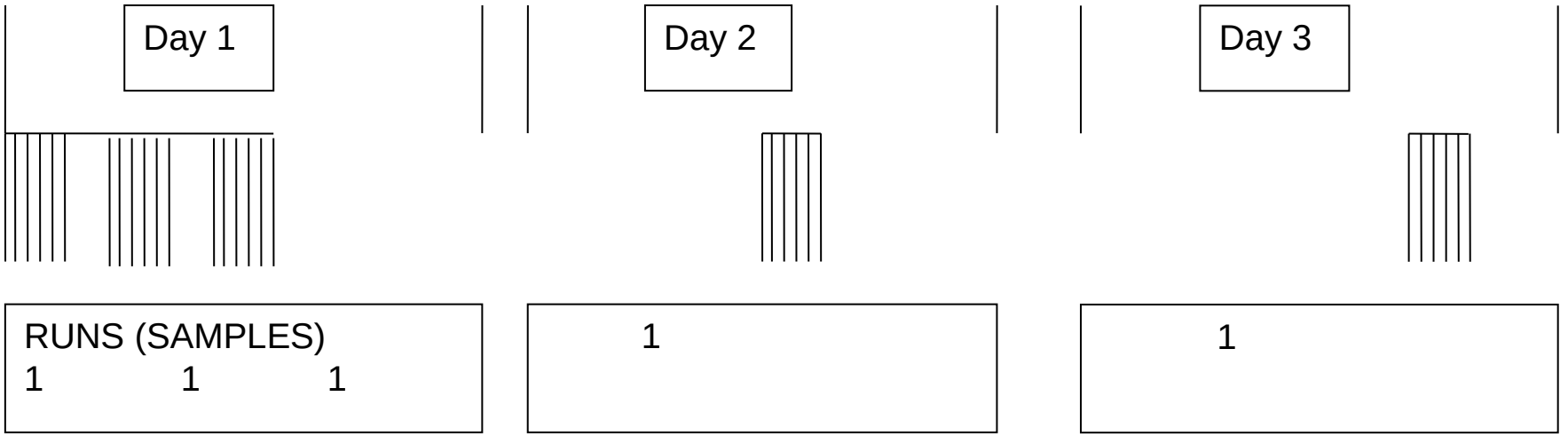
Same Analyst

- 6 injections x 5 runs x 3 days

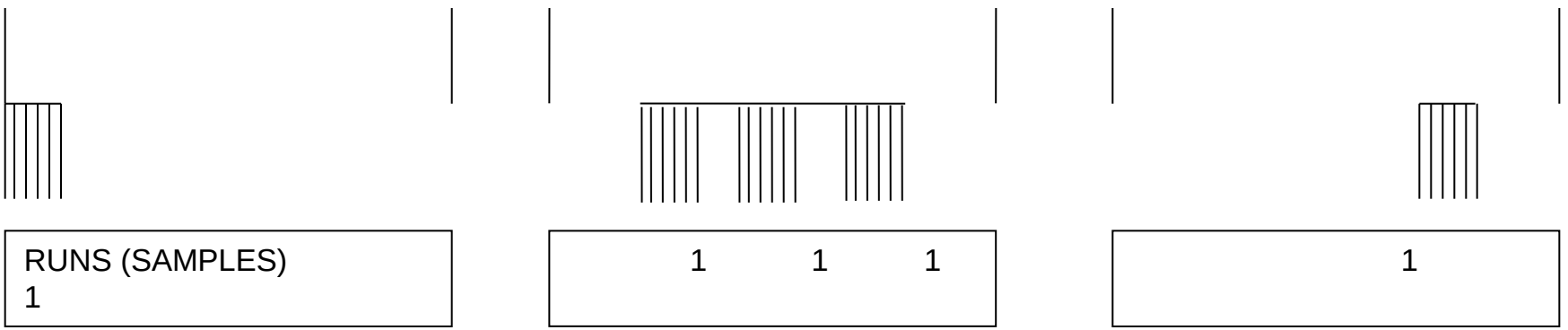
Case 4

Joint Research Centre

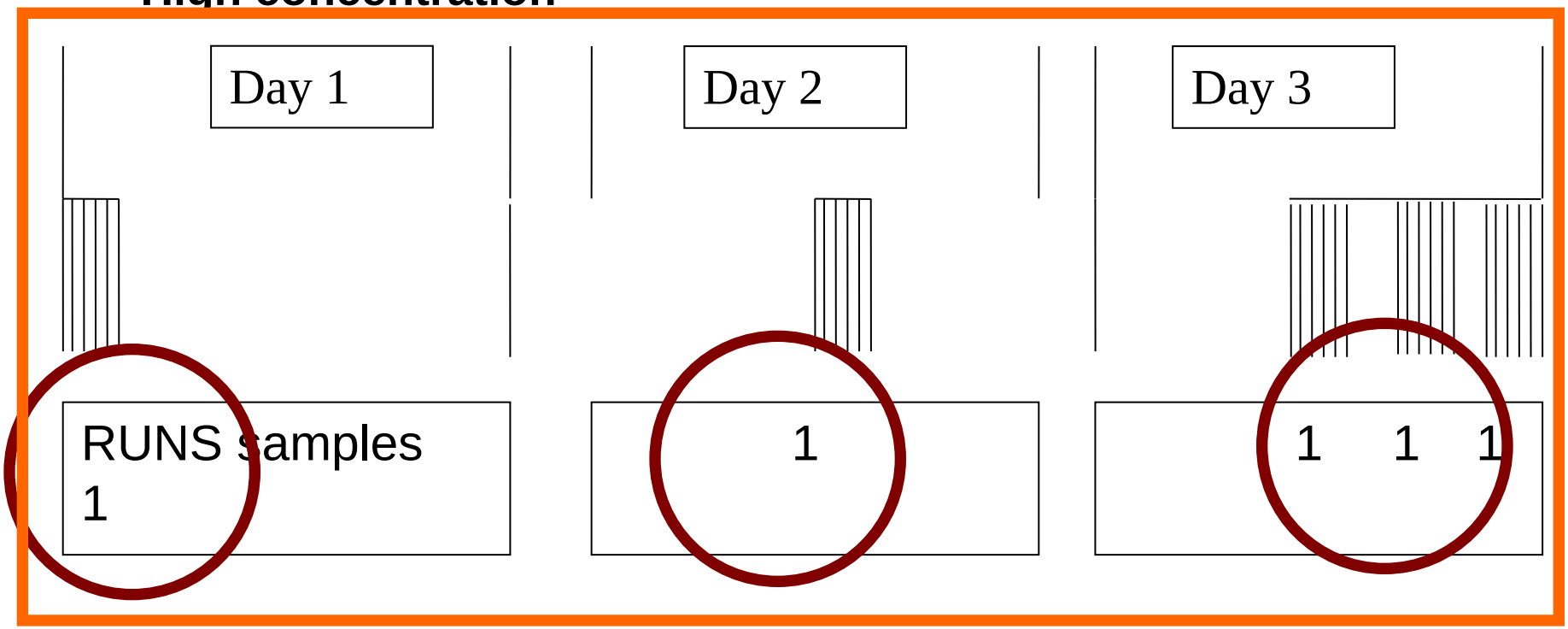
Low concentration



Intermediate concentration



High concentration



$$Y = m + \underline{Bo}_{days} + \underline{B(1)}_{samples} + \underline{e}$$

The statistical model of 3 factors-staggered nested-design:

Precision must be estimated at each level of concentration for all feeds

MINITAB - Poultry Feed_MADURAMICIN.MPJ - [Worksheet 1 ***]

File Edit Manip Calc Stat Graph Editor Window Help

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	DAYS	SAMPLE	LEVEL	CONCENTRA	RECOVERY									
1	1	1	1	0.86913	86.913									
2	1	1	1	0.81134	81.134									
3	1	1	1	0.84030	84.030									
4	1	1	1	0.85616	85.616									
5	1	1	1	0.88691	88.691									
6	1	1	1	0.86013	86.013									
7	1	2	1	0.91174	91.174									
8	1	2	1	0.87100	87.100									
9	1	2	1	0.90275	90.275									
10	1	2	1	0.87656	87.656									
11	1	2	1	0.94021	94.021									
12	1	2	1	0.89210	89.210									
13	1	3	1	0.86196	86.196									
14	1	3	1	0.89246	89.246									
15	1	3	1	0.87497	87.497									
16	1	3	1	0.90037	90.037									
17	1	3	1	0.88017	88.017									
18	1	3	1	0.90212	90.212									
19	2	4	1	1.02288	102.288									
20	2	4	1	0.94677	94.677									
21	2	4	1	0.96107	96.107									
22	2	4	1	0.96268	96.268									
23	2	4	1	0.94640	94.640									
24	2	4	1	0.86762	86.762									
25	3	5	1	0.85335	85.335									
26	3	5	1	0.86391	86.391									
27	3	5	1	0.84498	84.498									
28	3	5	1	0.87969	87.969									
29	3	5	1	0.88105	88.105									

Staggered nested design

SAMPLES (days)

Joint Research Centre

DATA ENTRY

MINITAB - Poultry Feed_MADURAMICIN.MPJ - [Worksheet 1(LEVEL = 1) ***]

File Edit Manip Calc Stat Graph Editor Window Help

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	DAYS	SAMPLE	LEVEL	CONCENTRA	RECOVERY									
1	1	1	1	0.86913	86.913									
2	1	1	1	0.81134	81.134									
3	1	1	1	0.84030	84.030									
4	1	1	1	0.85616	85.616									
5	1	1	1	0.88691	88.691									
6	1	1	1	0.86013	86.013									
7	1	2	1	0.91174	91.174									
8	1	2	1	0.87100	87.100									
9	1	2	1	0.90275	90.275									
10	1	2	1	0.87656	87.656									
11	1	2	1	0.94021	94.021									
12	1	2	1	0.89210	89.210									
13	1	3	1	0.86196	86.196									
14	1	3	1	0.89246	89.246									
15	1	3	1	0.87497	87.497									
16	1	3	1	0.90037	90.037									
17	1	3	1	0.88017	88.017									
18	1	3	1	0.90212	90.212									
19	2	4	1	1.02288	102.288									
20	2	4	1	0.94677	94.677									
21	2	4	1	0.96107	96.107									
22	2	4	1	0.96268	96.268									
23	2	4	1	0.94640	94.640									
24	2	4	1	0.86762	86.762									
25	3	5	1	0.85335	85.335									

ANALYSIS

MINITAB - Poultry Feed_MADURAMICIN.MPJ - [Worksheet 1(LEVEL = 1) ***]

File Edit Manip Calc Stat Graph Editor Window Help

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	DAYS	SAMPLE	LEVEL	CONCENTRA	RECOVERY									
1	1	1	1	0.86913	86.913									
2	1	1	1	0.81134	81.134									
3	1	1	1	0.84030	84.030									
4	1	1	1	0.85616	85.616									
5	1	1	1	0.88691	88.691									
6	1	1	1	0.86013	86.013									
7	1	2	1	0.91174	91.174									
8	1	2	1	0.87100	87.100									
9	1	2	1	0.90275	90.275									
10	1	2	1	0.87656	87.656									
11	1	2	1	0.94021	94.021									
12	1	2	1	0.89210	89.210									
13	1	3	1	0.86196	86.196									
14	1	3	1	0.89246	89.246									
15	1	3	1	0.87497	87.497									
16	1	3	1	0.90037	90.037									
17	1	3	1	0.88017	88.017									
18	1	3	1	0.90212	90.212									
19	2	4	1	1.02288	102.288									
20	2	4	1	0.94677	94.677									
21	2	4	1	0.96107	96.107									
22	2	4	1	0.96268	96.268									
23	2	4	1	0.94640	94.640									
24	2	4	1	0.86762	86.762									
25	3	5	1	0.85335	85.335									

Fully Nested ANOVA

C1 DAYS
C2 SAMPLE
C3 LEVEL
C4 CONCENTRA
C5 RECOVERY

Responses:

RECOVERY

Factors:

DAYS SAMPLE

Select

Help OK Cancel

OUTPUT

MINITAB - Poultry Feed_MADURAMICIN.MPJ - [Session]

File Edit Manip Calc Stat Graph Editor Window Help

Results for: Worksheet 1(LEVEL = 1)

Nested ANOVA: RECOVERY versus DAYS, SAMPLE

Analysis of Variance for RECOVERY

Source	DF	SS	MS
DAYS	2	310.3019	155.1510
SAMPLE	2	64.0291	32.0146
Error	25	229.6309	9.1852
Total	29	603.9620	

Variance Components

Source	Var Comp.	% of Total	StDev
DAYS	14.659	53.02	3.829
SAMPLE	3.805	13.76	1.951
Error	9.185	33.22	3.031
Total	27.649		5.258

Expected Mean Squares

1	DAYS	1.00(3) + 6.00(2) + 8.40(1)
2	SAMPLE	1.00(3) + 6.00(2)
3	Error	1.00(3)

SUMMARY

Nested Design Anova table: RECOVERY Level 1

Sources of variance	SS	df	MS	F	p
Days (Bo)	**310.3019**	**2**	**155.1510**	**16.89**	**0.0001**
Sample (B1)	**64.0291**	**2**	**32.0146**	**3.49**	**0.0462**
Residual error (*e*)	**229.6309**	**25**	**9.1852**		
Total	**603.9620**	**29**			

Estimated Variance components

Variance Source	Component	% of total	StDev	
Days (S^2_{Bo})	**14.659**	**53.02**	**3.829**	
Sample (S^2_{B1})	**3.805**	**13.16**	1.951	
Error (S^2_e)	**9.185**	**33.22**	3.031	
Total	**27.649**		5.258	
$S^2_r = (S^2_e + S^2_{B1})$	**12.99**		3.604	**S_r repeatability Stdev**
$(S^2_e + S^2_{Bo} + S^2_{B1}) = S^2_{int}$	**27.649**		5.258	**S_{int} INTERMEDIATE PRECISION**

CONCLUSION

MINITAB INC™
Making Data Analysis Easier

- **Practical to estimate precision**
- **Can be used for data storage and exploratory data analysis**
- **Provides control charts**
- **Advanced Statistical Analysis**

THANK YOU FOR YOUR ATTENTION