Managing Research Data Effectively

Justin du Toit Grootfontein Agricultural Development Institute

- From field data-sheets to electronic files
- Data entry and storage
- Managing data in Excel
- Managing visual data (photographs)

Entry format (base format)

Date	1 10 2008			
Site	Short grass			
	Disc Height	Number of grass plants	Number of flowering culms	Wattle seedlings
Quadrat 1	3.7 cm	6	0	12
Quadrat 2	6 cm		1	3 (1 dead)
Quadrat 3	5.3 cm	2	1	5
Quadrat 4	4.4 cm	2	2	8
Quadrat 5	6.9 cm		3	0
Date	1 10 2008			
Site	Tall grass			
	Disc Height	Number of grass plants	Number of flowering culms	Wattle seedlings
Quadrat 1	9 cm	2	2	0
Quadrat 2	11 cm	2	4	0
Quadrat 3	12.3 cm	Not record	ed	
Quadrat 4	8 cm	0	0	3
Quadrat 5	8 cm	4	4	1 (dead)

Computers view data differently from the way we do

- Computers can analyse data *incredibly* fast, but it must be in the correct format
- <u>Data sheets</u> are laid out in a way that makes sense to us (see left)
- <u>Data files</u> must be laid out in a way that makes sense to a computer



Dual entry



- To enter a list of 100 values takes between 0.01 and 0.1% of the time that it takes to run an experiment (person working-hours, not duration of the experiment)
- However, vested within these numbers is the *entire cost* of the experiment (usually hundreds of thousands of rands)
- Therefore, it is *vital* to ensure that the data are captured perfectly
- By far the most reliable way of doing this is for two different people to enter the data (having the same person enter it twice leads to unhealthy temptation with the Copy | Paste function!)

Spreadsheet or database?

Spreadsheets

- Easy to use
 Suitable for most
- applications

 Lack of query capacity
- (hence less room for errors...)
- illustration capabilities
- Limited to 65536 rows x 256 columns

Databases

- Require expertise to run correctly
- VERY EASY to ask for the wrong thing, and hence end up with junk data
- Vital for large data sets, especially ongoing ones, and where >1 people enter data
- Unlimited data storage

Backing up data

- Why backup? To have a copy of the data that is free from the hazards your original copy might face, e.g.:
- Theft, a computer crash, losing the storage device, office burns down, etc. ■ Therefore, most commonly backups face exactly the same hazards as the original!
- set up a few internet-based email accounts, and Search & Send it to these regularly, and using e.g. Google Fusion
- (Similarly, photograph or scan lab- or field-data sheets)



Example – Google Fusion

2 K to be applied and the property of the second	ercentarilectri (shies No pixee - Windows Internet I	aplease a			
h GR (he facts (n) 80) - Stageture Macin			2 B 4 8 2 Loop 24			COTO - K Hauthan profile and inclusion and conta			 B) (+) (x) 30 toops 		14	
Parenter 🏤 🗆 anzi 🐮 danga Panar 1 - 🖉 manga Panar Table (m							the test year to	ons (or an				
i - Munaga Passar Table (19. x 440)	and when the second of the second sec						Afreste A D	ALE & Google Factor Tables VARIABLE	ej whatter			
	eproservedh toop		Q -	2 · C @ · ten · bie	er føk	- 8-	🙁 - 💋 tenga Tum	e heles (m. x 📲 desp-care red) b	ioge	5	-0.0.0.0	s- Select Spin-
aver heoritable			justinduto	n@gnail.com <u>Tox</u> <u>this</u> <u>Cee</u>	088315	91.0xf	Coope Neontable			justine	knoiteligenail.com Tox	this I Constant 1 Sig
to place ware a tecover				Classesene (0)	det bit.	Sec.	No place	I du Tot DAFF			Chackee	ana Di desas da
Ve Vew Edit Vesseler Mege							File Vew Edit	Vessier Mege				
Corrent view: All - Show aptions				1.0	900 yf 195	1005.2	Correct view: All - S	how actions				1 - 900 of 195 (
Onter *	Transact 7	Levelade *	Latitude *	Abbuda *			Onto -	Full species 7	Gene *	Section *	Veleter *	
19209		-21.25784	25.690108	1408		8	5/32049	Arietida congenta	Anna	congesta		
/13/2010	1	-01.267984	26.600106	1406	9	8	6/13/2010	Engenie klonawiere	Engunin	Mananiana .		0
/13:2610		-31.257984	25.630106	1406	Ø	8	6/13/2610	Avietida dilloca	Aidda	diluca		0
/32(1)	1	-31.257884	25.530106	5406	Ø		5132510	Ericceptalus spirescens	Disceptake	spinescens		0
/0200		-31.257804	25.630106	1400	9		5/13/2010	Walahola spotilio	Visiekde	motio		P
/13:2010	1	-01.257904	26.600106	1400	9	8	6/13/2010	Tropie koelensidee	Traper	hotesides		9
213(2610		-21.257984	25.620706	1406	9	13	6/13/2610	Pertica phone	Petzia	glubese		9
213(2510	1	-31.257684	25.530106	1406	0	- 8	5/13/2610	Cogrania curvia	Disposite	curate		0
/13/0640		31,257884	25.690106	5400	œ.	- 13	6/13/0010	Ovyeschone cilots	Chrysochome	cilata		9
/130510	1	-21,257984	25-520736	1400	9	8	6/13/2010	Encorpholus proceles	Encentelle	excedes		9
2130610	2	-31.960075	2549685	1305	9	8	6/13/2610	Telichrysum perdoades	Heisbyran	pertaules		9
2430510	2	-31,269075	25-196845	100	Ø	- 8	5/13/2510	2/paphylium incruitatum	2ygophyllum	inouestum		0
213.0244	2	-31,263675	25.496846	1967	8		5/02(4)	Cognitie survie	Disposis	curela		
/13/2610	2	-21,260075	21-24648	1967	9		6/19/2010	Areida éños	Arston	enue		9
2130610	2	-21,262075	21-29682	1988	~		6/13(2610	Cyndopagen planodia	Cyntrapagos	phonode		
-			A	Olteral 6		× 4%	1010010	NGR DOUBLE	Sacto.	ACCORD OF	(a) (stand	4 · 5.18
and the second se	-	-	And in case of the local division of the loc	Contract on the second	-	10020-04	Concession of the local division of the loca		-	100000	-	1000 A



Make your life easier in Excel! Learn the basic functions!

- Basic arithmetic functions (=,-,x,/, sum, Conditionally manipulate data (=if(CONDITION,TRUE,FALSE)
- Simplify data e.g. convert continuous data to binomial (=if(VALUE>x,1,0)
- value or term (=vlookup(VALUE,TABLE,COLUMN))
- Remove unwanted spaces (=trim(TEXT))

- Split data into separate records e.g.
 "Eragrostis curvula" into two columns titled "Genus" and "Species"
- Truncate text e.g. (=left("Eragrostis",3) gives "Era" combined with the same for "curvula" (cur) can be joined (="Era"&"cur") to give a code Eracur
- You NEVER have to re-enter values, and NEVER have to use a calculator in Excel

	A	В		С	D
1	SpeciesFull	Genus		Species	Code
2	Gnidia polycephala	Gnidia		polycephala	Gniala
3	Helichrysum cerastoides	Helichrysum		cerastoides	Heldes
4	Helichrysum zeyheri	Helichrys	um	zeyheri	Heleri
5	Hermannia coccocarpa	Hermannia		coccocarpa	Herrpa
6	Hertia pallens	Hertia		pallens	Herens
7	Heteropogon contortus	Heteropo	gon	contortus	Hettus
	Original names Split using Data Text to Columns		6-let letters tters o left(A	ter code (1 st 3 s of genus, 1 st 3 of species) usir 2,3)&right(B2,	3 ag: .3)

	A	В	С	D	
1	SpeciesFull	RelAb	>10%?	log(RelAb)	
2	Helichrysum zeyheri	16.2	1	1.21	
3	Hermannia coccocarpa	7.7	0	0.89	
4	Heteropogon contortus	31.2	1	1.49	
5	Hyparrhenia hirta	22.6	1	1.35	
6	Lepidium africanum	2.3	0	0.36	
7	Limeum aethiopicum	13.9	1	1.14	
8	Oropetium capense	7.6	0	0.88	
0					
	Data				
ſ	Conditional manipulation:	Arith	nmetic conv	ersion	
	1 if greater than 10,	Log	of original	data	
	otherwise 0	=log	g(B2)		
	=if(B2>10,1,0)				

	Α	B	С	D	E	F	G	Н
					Cumulative	Cumulative		
		Rain	Cumulative		sequential wet	sequential dry	Cumulative	Cumulative
1	Date	(mm)	rainfall	Rain?	days	days	wet days	dry days
2	01-Jan	0	0	0	0	1		1
3	02-Jan	4	4	1	1	0		
4	03-Jan	33	37	1	2	0		
5	04-Jan	14	51	1	3	0		
6	05-Jan	4	55	1	4	0	4	
7	06-Jan	0	55	0	0	1		
8	07-Jan	0	55	0	0	2		
9	08-Jan	0	55	0	0	3		3
10	09-Jan	2	57	1	1	0		
11	10-Jan	8	65	1	2	0	2	
12	11-Jan	0	65	0	0	1		
13	12-Jan	0	65	0	0	2		
14	13-Jan	0	65	0	0	3		
15	14-Jan	0	65	0	0	4		4
10								
					6	ଳ	\sim	
				$\widehat{}$	<u>-</u>	-1.		
			3	-1	124	5+	E	F3,
			÷.	<u>o</u>	<u> </u>),F	2	4,
			B3	6	Ξ) i	7	<u> </u>
			÷+	E.	D3	D3	E3	F3:
				Ē	FG	E	F(F(
					T	Π	Π	Π











Folder: Animals, Honey Badgers, Suicidal tendencies?

