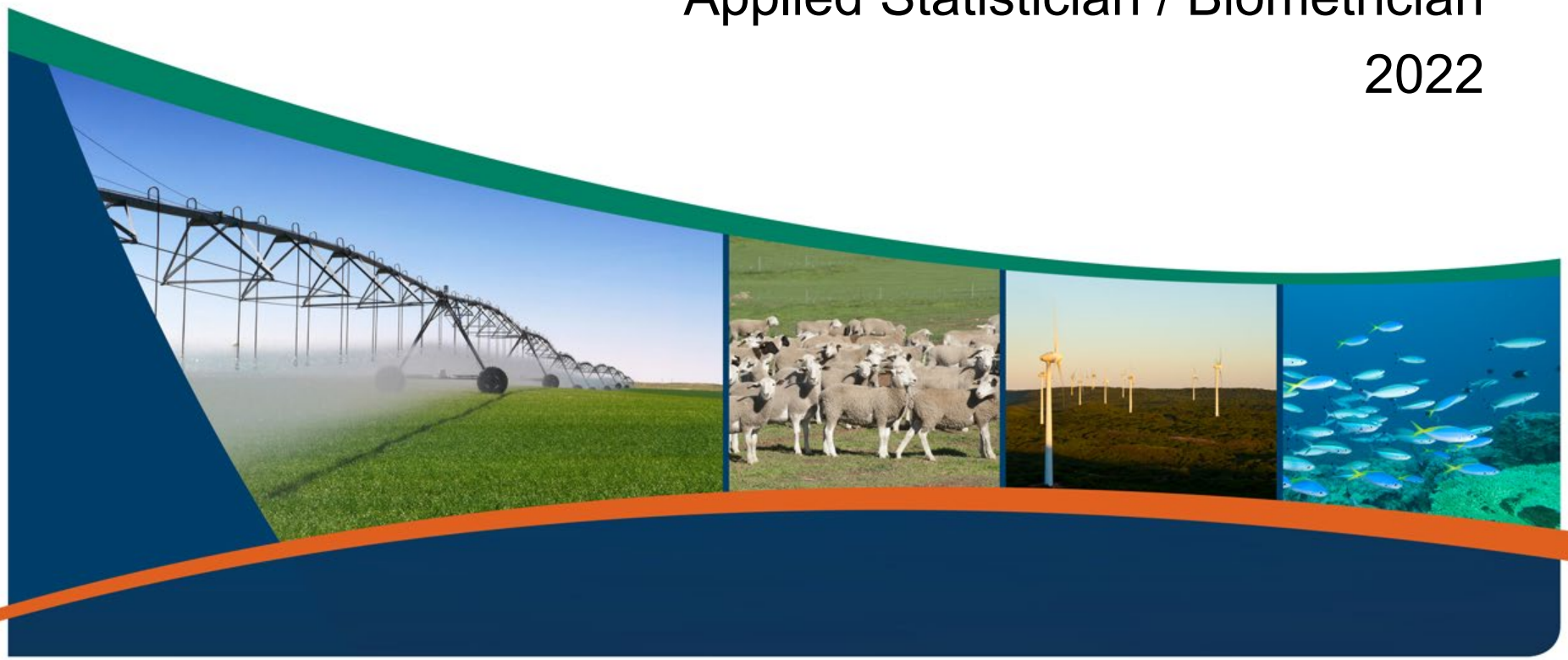
Department of
Primary Industries and Regional Development

GOVERNMENT OF WESTERN AUSTRALIA

SAGI
Statistics for the Australian Grains Industry
National | North | West | South

Presenting Trial Results

Andrew van Burgel
Applied Statistician / Biometrician
2022



Example

- **Consider this trial:**
  - 5 crops were evaluated as a green manure (GM) with 3 GM methods.
  - It was a split plot design with crop type as whole plot.
  - Response variable is yield from wheat sown the following year.

		Yield	Yield	Yield
Crop!	GM!	Rep 1	Rep 2	Rep 3
CANOLA	PLOUGH	1.95	1.69	1.72
CANOLA	SLASH	1.81	1.99	1.54
CANOLA	SPRAY	1.67	1.39	1.45
FIELD PEAS	PLOUGH	2.18	2.00	2.31
FIELD PEAS	SLASH	2.30	2.18	2.07
FIELD PEAS	SPRAY	1.99	2.01	2.07
MUSTARD	PLOUGH	1.59	1.69	2.07
MUSTARD	SLASH	1.86	1.74	1.95
MUSTARD	SPRAY	1.58	1.63	1.83
OATS	PLOUGH	1.60	1.60	1.52
OATS	SLASH	1.68	1.75	1.27
OATS	SPRAY	1.66	1.53	1.37
VETCH	PLOUGH	1.90	1.97	1.76
VETCH	SLASH	1.80	1.85	2.01
VETCH	SPRAY	1.67	1.71	1.94

Analysis results

- **Output from Genstat:**
  - How would you summarise the results graphically for a paper or presentation?

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Rep.Crop stratum					
Crop	4	1.62141	0.40535	10.45	0.001
Residual	10	0.38801	0.03880	2.36	
Rep.Crop.GM stratum					
GM	2	0.21264	0.10632	6.47	0.007
Crop.GM	8	0.06375	0.00797	0.48	0.853
Residual	20	0.32880	0.01644		
Total	44	2.61461			

Tables of means

Crop	CANOLA	FIELD PEAS	MUSTARD	OATS	VETCH
	1.689	2.123	1.771	1.554	1.846

GM	PLOUGH	SLASH	SPRAY
	1.837	1.853	1.700

Crop	GM	PLOUGH	SLASH	SPRAY
CANOLA		1.785	1.780	1.501
FIELD PEAS		2.165	2.182	2.021
MUSTARD		1.783	1.849	1.681
OATS		1.575	1.566	1.521
VETCH		1.877	1.886	1.775

Standard errors of means

Table	Crop	GM	Crop GM
e.s.e.	0.0657	0.0331	0.0892
Except when comparing means with the same level(s) of Crop			0.0740

Standard errors of differences of means

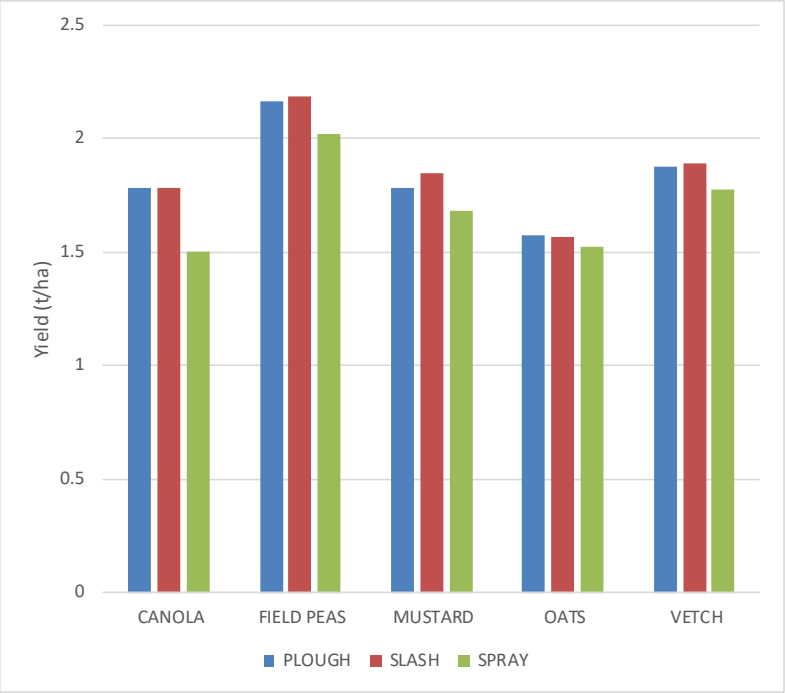
Table	Crop	GM	Crop GM
s.e.d.	0.0929	0.0468	0.1262
Except when comparing means with the same level(s) of Crop			0.1047

Least significant differences of means (5% level)

Table	Crop	GM	Crop GM
l.s.d.	0.2069	0.0977	0.2599
Except when comparing means with the same level(s) of Crop			0.2184

Interaction

- **Given that the interaction is not significant (and not even close at p=0.853), I would not present all treatment means as below:**
  - The risk is that readers/ viewers will make conclusions from the non-significant results
    - eg. based on the LSD (~0.22) yields are significantly lower with spray for Canola only.
  - Instead I would focus on the significant main effects



Different approach for different situations

In general…

- **If the interaction is statistically significant**
  - Focus on the interaction effect
- **If the interaction is not statistically significant**
  - Focus on the main effects

But… there are exceptions…

- **Interaction is close to significant and is meaningful**
  - Focus on the interaction effect
- **Interaction is significant but relatively small**
  - Focus on the main effects

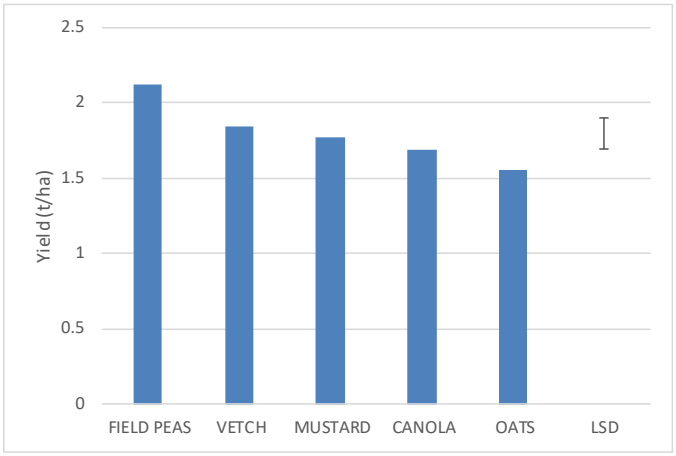
Graphing the main effect

- **Consider the crop type main effect:**
  - CANOLA 1.689
  - FIELD PEAS 2.123
  - MUSTARD 1.771
  - OATS 1.554
  - VETCH 1.846
  - Statistical results: SE = 0.0657; SED = 0.0929; LSD = 0.2069; p = 0.001

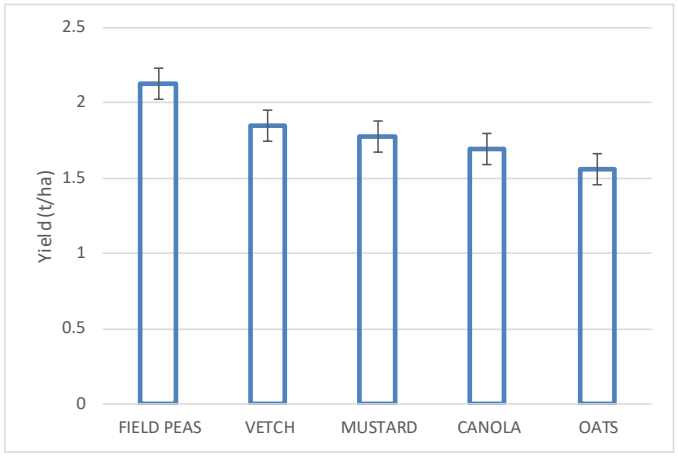
- **How would you graph these crop type means?**

Some options

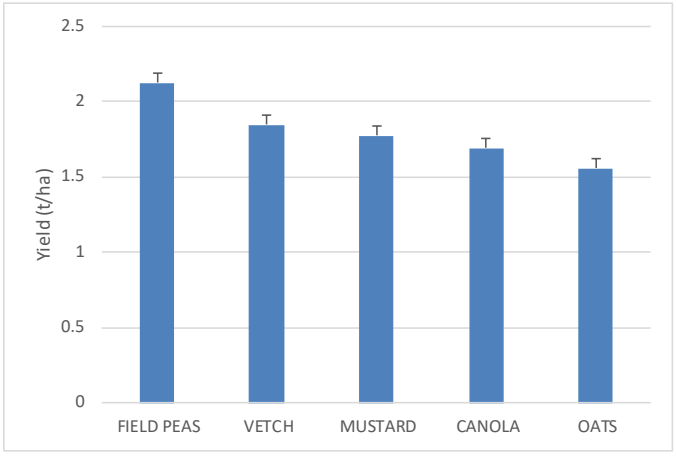
- **Single LSD bar**



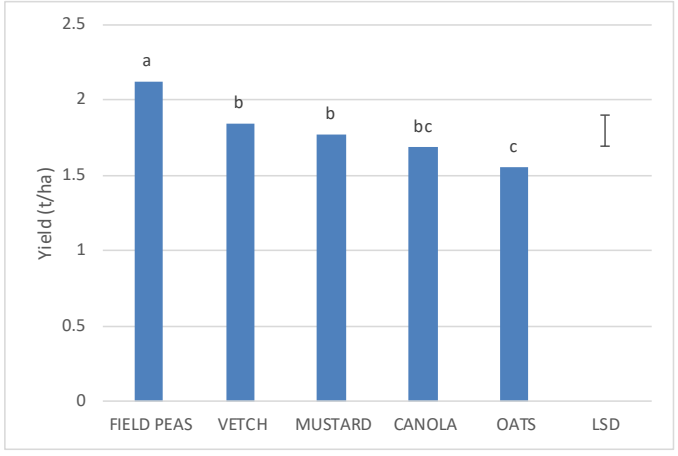
- **+/- SED or +/- half-LSD bars**



- **SE bars**



- **Significance lettering**



Treatments sharing no common letters are significantly different ($p<0.05$)

Demonstration

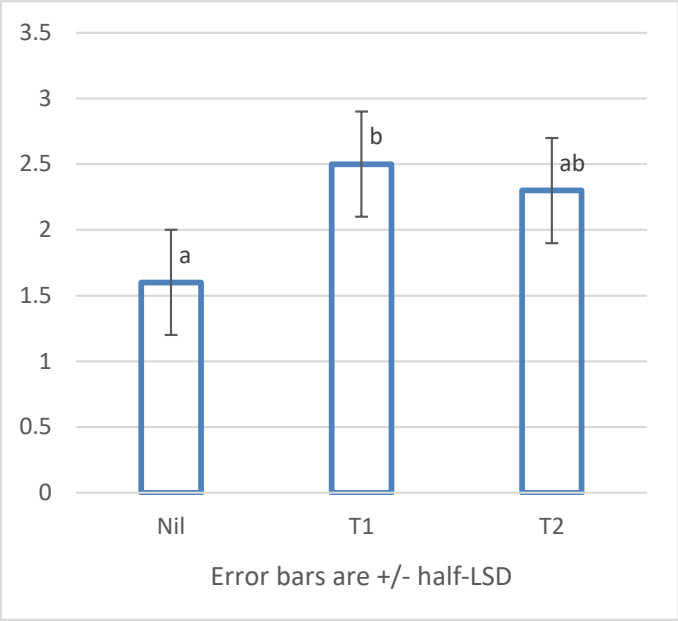
- **How to produce such a graph in Excel…**
  - Note: more powerful tools (eg. ggplot in R) can produce more elegant looking graphs, but additional learning and effort is required.

Significance traps

- p<0.05 should not be applied strictly.
  - p=0.08 is not the same as p=0.80.

Suggestion: quote exact p-values, rather than p<0.05, p>0.05

- If the effect of treatment 1 (T1) is significantly different to the Nil, but the effect of treatment 2 (T2) is not significantly different to the Nil, this does not imply T1 is significantly better than T2.

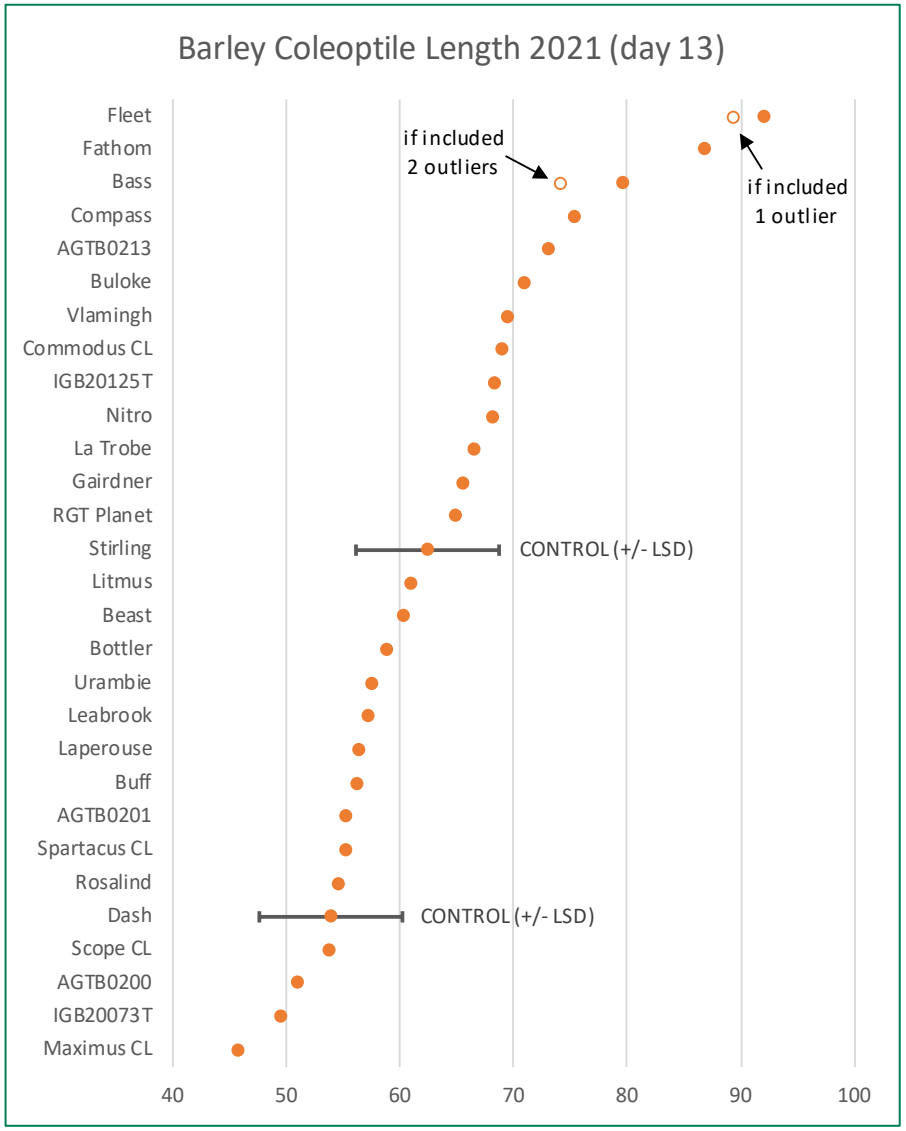


- "Statistical significance" is different to the more common meaning of significance (importance)
  1. An effect can be statistically significant but not significant
     - if the effect is too small to be important
  2. An effect can be significant but not statistically significant
     - if the effect is large, but there is high variability

Suggestion: use "statistically significant" instead of just "significant" when describing statistical significance.

Other graphing tips

- Transpose the usual x and y axes if many groups and/or long group names
- Gridlines helpful visual aid
- Text on the graph is quicker to interpret compared to legends or captions.



Key messages

- **Avoid presenting any results that are not even close to statistically significant.**
- **In most cases include some statistical information (e.g. LSD) to enable valid comparison of treatment means.**
- **If including error bars, make sure you define what they represent. Are they SE, SED, LSD, half-LSD, etc.**

*** Questions**

Thank you

Visit dpird.wa.gov.au