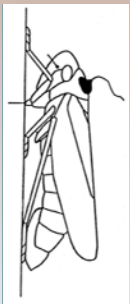


# Congratulations

- You have a clean data file.
- Open Ebert 1.0.
- Change the infile statement to read the data in `AphidData\at.csv`.
- Change the ODS HTML file= statement to give the program a place to dump the output.
- See next slide for a visual.
- Run the program.

EPG  
Work-  
shop



# Change infile and ODS filenames

- Here and here

The screenshot displays the SAS Enterprise Guide interface with a code editor window. Two blue arrows originate from the bullet point 'Here and here' in the slide above. One arrow points to the line: `infile "&libref;&libname.&dataset;" dataset=`. The other arrow points to the line: `ods output="&libref;&libname.&dataset;"`. The code is a SAS program for processing insect data, including comments in English and Spanish, and various SAS statements like `infile`, `length`, `input`, `if`, `then`, `else`, `ods`, and `ods output`.

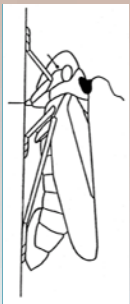
EPG  
Work-  
shop



# Results

- Open the results file in Word. Allow Word some time to process the file. Something like 2 minutes should be good in most cases.
- Select all (Control a) and copy (Control c).
- Go to Excel, and Paste special, then choose “text” from the menu choices.
- This strips off formatting and avoids problems with merged cells in Excel.

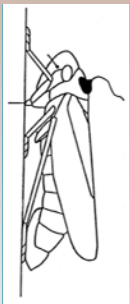
EPG  
Work-  
shop



# Condensing Output

- SAS produces many pages of output. You want a simple table with the results.
- Each Glimmix statement produces results that have all the same format.
- Use this feature to your advantage.

EPG  
Work-  
shop



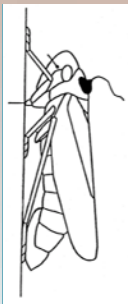
# The Output in Excel

- You should now have this.

Book2 - Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	ANOVA & LSD of TmFrstPrbFrmStrt																						
2																							
3																							
4	The GLMMIX Procedure																						
5	Model Information																						
6	Data Set	WORK.EBERT																					
7	Response Variable	TmFrstPrbFrmStrt																					
8	Response Distribution	Gaussian																					
9	Link Function	Identity																					
10	Variance Function	Default																					
11	Variance Matrix	Diagonal																					
12	Estimation Technique	Restricted Maximum Likelihood																					
13	Degrees of Freedom Method	Residual																					
14																							
15	Class Level Information																						
16	Class	Levels	Values																				
17	trt		2 a b																				
18																							
19	Number of Observations Read		8																				
20	Number of Observations Used		8																				
21																							
22	Dimensions																						
23	Covariance Parameters		1																				
24	Columns in X		3																				
25	Columns in Z		0																				
26	Subjects (Blocks in V)		1																				
27	Max Obs per Subject		8																				
28																							
29	Optimization Information																						
30	Optimization Technique	None																					
31	Parameters		3																				
32	Lower Boundaries		1																				
33	Upper Boundaries		0																				
34	Fixed Effects	Not Profiled																					
35																							
36	Fit Statistics																						
37	-2 Res Log Likelihood		28.93																				
38	AIC (smaller is better)		34.93																				
39	AICC (smaller is better)		46.93																				
40	BIC (smaller is better)		34.3																				
41	CAIC (smaller is better)		37.3																				
42	HQIC (smaller is better)		32.43																				
43	Pearson Chi-Square		27.77																				
44	Pearson Chi-Square / DF		4.03																				
45																							

EPG  
Work-  
shop



# Extracting the useful bits

- Copy the contents in cell E47. This is important. It process only works this way, but I don't know why.

Copy this

The screenshot shows an Excel spreadsheet with the following data:

Effect	Num DF	Den DF	F Value	Pr > F
Intercept	1	6	1.18	0.3185

Parameter	Estimate	Standard Error	DF	T Value	Pr >  T
μ	0.4253	0.2422	6	1.76	0.093
σ <sup>2</sup>	5.1552	0.9622	6	5.36	0.0008

Source	Sum of Squares	DF	Mean Square	F Value	Pr > F
Model	1.18	1	1.18	0.3185	0.5815
Error	1.94	6	0.3233		
Total	3.12	7			

EPG  
Workshop



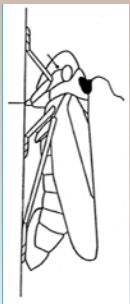
# Extracting

- Type this into cell I48

=IF(\$E47="Pr > F",E48,"")

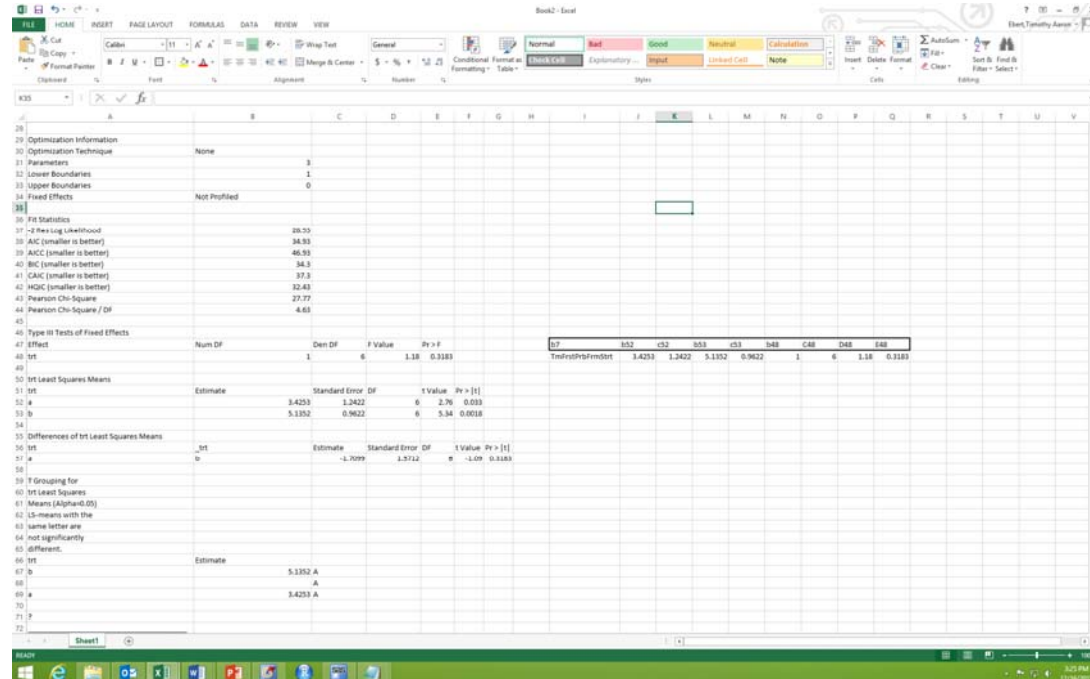
- Make sure that the Pr > F is entered into the formula using the paste command.
- Fill right to cell L48
- In these cells, change the e48, f48, g48, etc... to useful cells.
- I will typically have the first cell (that currently has e48) to cell b7 (the variable name).

EPG  
Work-  
shop



# Extracting

- So I will have something like this, where the contents of the cells in the highlighted box indicate the cell that I have used in the formula.



The screenshot shows an Excel spreadsheet with a table of parameter estimates highlighted in the bottom right corner. The table is as follows:

	b7	b52	b53	b53	b53	b48	b48	b48	b48
TmFirstPmFmDmT	3.4253	1.2422	5.1352	0.9622	1	6	1.18	0.3183	

EPG  
Work-  
shop

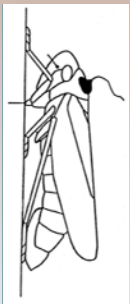




# Extracting

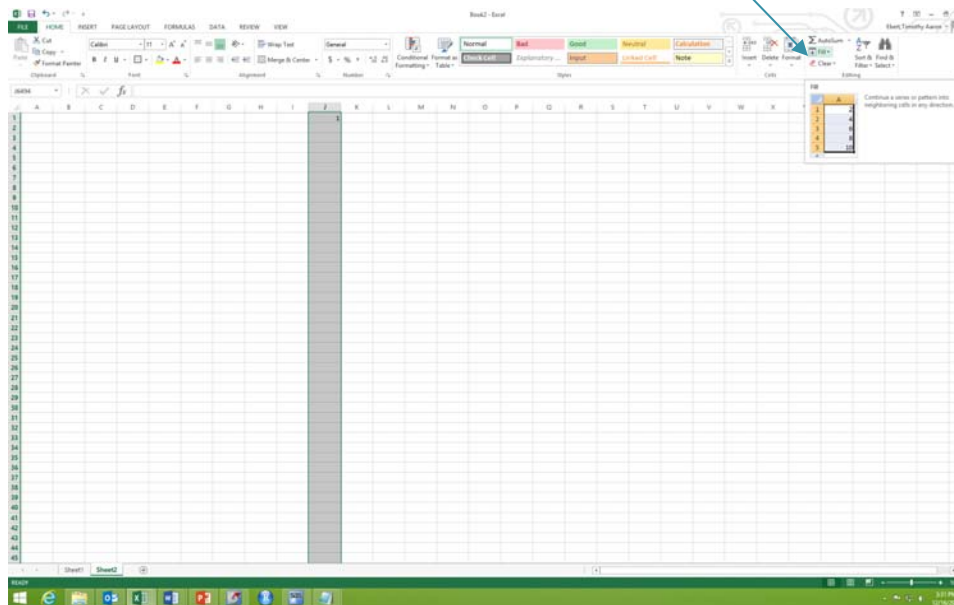
- Fill cells i48 through q48 down to the end of the results (row 649 l).
- Copy columns l through q.
- Open another worksheet, and paste values.
- In cell j1 type the number one.
- Select all cells from j1 through j649 l.

EPG  
Work-  
shop

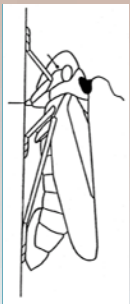


# Extracting

- Right click the fill menu, and select series.



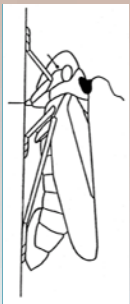
EPG  
Work-  
shop



# Extracting

- The default (type=linear, step=1) is great
- Column J should now have integers from one to 6491
- Select columns A through J
- Sort column A in descending order.
- Select cells A89 through j89 through A6491 through j6491.
- Delete

EPG  
Work-  
shop



# Extracting

- Select columns J through A (not A through J) and sort smallest to largest. You should now have this:

	A	B	C	D	E	F	G	H	I	J
1	b7	0.52	0.53	0.55	0.48	0.48	0.48	0.48	0.48	47
2	TmFstPrt	3.4253	1.2422	5.1352	0.9822	1	6	3.18	0.3183	48
3	CtoFstPrt	1.1114	0.9032	0.8205	0.4216	1	3	0.09	0.7781	119
4	NumF	0.6667	0.322	0.4	0.2894	1	6	0.43	0.537	190
5	DurFstPrt	6.0817	1.3956	7.0742	1.081	1	6	0.32	0.5943	261
6	DurSndPrt	4.767	0.7626	5.264	0.6605	1	5	0.24	0.6432	332
7	ShrtCdnE	7.5971	0.5434	8.1391	0.2717	1	3	0.8	0.438	403
8	DurSndE	5.2824	0.889	5.0289	0.7899	1	5	0.05	0.8966	474
9	TIDurF	7.8382	0.4992	7.8382	0.4992	1	2	0	1	545
10	DurInpD	7.205	0.8027	5.9406	0.6218	1	6	1.7	0.24	616
11	meanpD	2.3457	0.4811	1.8658	0.3417	1	6	1.4	0.2817	687
12	meanPOS	3.8457	0.4811	1.8658	0.3417	1	6	1.4	0.2817	751
13	meanNPD	1.4874	0.8564	2.8253	0.6834	1	6	1.57	0.1566	852
14	meanF	7.8382	0.4992	7.8382	0.4992	1	2	0	1	923
15	TmSndPrt	8.7984	0.3429	8.5446	0.2853	1	8	0.35	0.578	994
16	TmFstPrt	8.77	0.3587	8.4982	0.2778	1	6	0.36	0.5711	1065
17	TmSndPrt	7.5971	0.5434	8.1391	0.2717	1	3	0.8	0.438	1136
18	NumG	0	0.2108	0.2	0.1833	1	6	0.56	0.4816	1207
19	NumInPrt	0.3333	0.3728	0.4828	0.2887	1	6	0.1	0.7619	1316
20	NmbrSrt	0				1	6			1387
21	NumE1	0.5774	0.4731	1.0928	0.3665	1	6	0.74	0.4221	1446
22	NumLngE	0				1	6			1517
23	NumSngE	0.5774	0.5004	0.5464	0.3876	1	6	0	0.9636	1576
24	NumE2	-5.4617	0.3192	0.4828	0.2473	1	6	1.43	0.2769	1647
25	NumLngE	0	0.2108	0.2	0.1833	1	6	0.56	0.4816	1718
26	DurFstE	2.897	1.9361	4.8866	0.9681	1	3	0.83	0.4296	1789
27	CtrE1E10	4.6042	0.781	3.9897	0.9151	1	3	0.5	0.5321	1840
28	DurE1Fst	5.1784	1.1204	0	0	0				1950
29	Port2InD	2.088	0.1293	0	0	0				2004
30	TIDurF	6.2873	1.5534	5.7793	0.7797	1	5	0.05	0.7676	2056
31	TIDurE1	6.2873	1.4852	5.3384	0.7426	1	3	0.46	0.5454	2129
32	TIDurE1F	5.2719	1.2139	0	0	0				2219
33	TIDurSng	6.2873	1.372	4.7277	0.7921	1	2	0.97	0.4287	2273
34	TIDurE1F	6.5344	0.8412	0	0	0				2344
35	TIDurE2	8.1818	0.685	0	0	0				2376
36	MnDurE1	5.189	0.8598	4.6093	0.4299	1	3	0.36	0.589	2452
37	MnDurE2	5.8133	0.3385	0	0	0				2523
38	NumInPrt	1.8455	0.9668	2.0412	0.7024	1	6	2.47	0.1668	2577
39	NmbrC	4.1789	0.7664	2.5121	0.9338	1	6	2.89	0.1402	2644
40	NmbrSrt	2.8595	1.0428	0.8828	0.8077	1	6	2.25	0.1846	2719
41	NumNP	3.8703	0.9198	2.0412	0.7125	1	6	2.47	0.187	2790
42	NmbrPD	7.3297	1.6371	8.0572	1.2681	1	6	0.12	0.7374	2861
43	NmbrPD	0	0	0	0	0				2932
44	NmbrPOS	7.3297	1.6371	8.0572	1.2681	1	6	0.12	0.7374	2991
45	NmbrE1e	0				1	6			3062

EPG Workshop



# Done

- If this was real data from your experiment then you should go and figure out why some variables like NumLngE2 have periods. In this case, treatment A had no E2, so it is easily explained.
- Further down there is a #Name? This is for the variables TtlDurF4 and TtlDurF5. The trimmed recording only goes to 2.8 hours. So this is not really a problem.

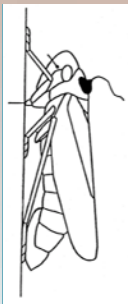
EPG  
Work-  
shop



# Insect Number

- The format of the first insect number is very important.
- If the first insect number is a 1, and you type in “a1” then you will get an error message when you get to insect “a10”.
- To avoid this type in a leading space “ a1”

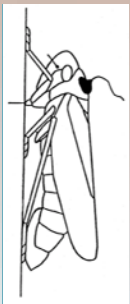
EPG  
Work-  
shop



# Insect Number

- In sorting insect “a11” will appear after a1, and you will have to go all the way to insect “a19” before you get to insect “a2”.
- To avoid this issue type in zeros.
- So insect “a1” is now insect “a01” if you have fewer than 100 insects per treatment, or insect “a001” if you have more that 100 but fewer than 1000.

EPG  
Work-  
shop



# The Next Steps

- Here is a list of steps.
  - Errors: An exercise where you try to find as many errors as possible.
  - Introduction to Backus 1.0
  - Introduction to Ebert X.X
  - Data Analysis
  - How the programs work, customizing the analysis
    - Programming FileManip
    - Programming Error Checker
    - Programming Trimmer
    - Programming Ebert X.X
- From this point forward there is no distinction between Windaq and Probe data. The process is the same in both cases.

EPG  
Work-  
shop

