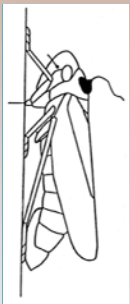# EPG Data Analysis 101

- Ebert 1.0: Reading data files and checking your work

- by
- T.A. Ebert
- M.E. Rogers

EPG Work-shop

Formatting Note

- Starting with this lecture, there will be screen shots to help you find things. Sometimes critical details are hard to see.
- The first screen shot will be an overview.
- Where needed, the following slide contains detailed portions of the overview.
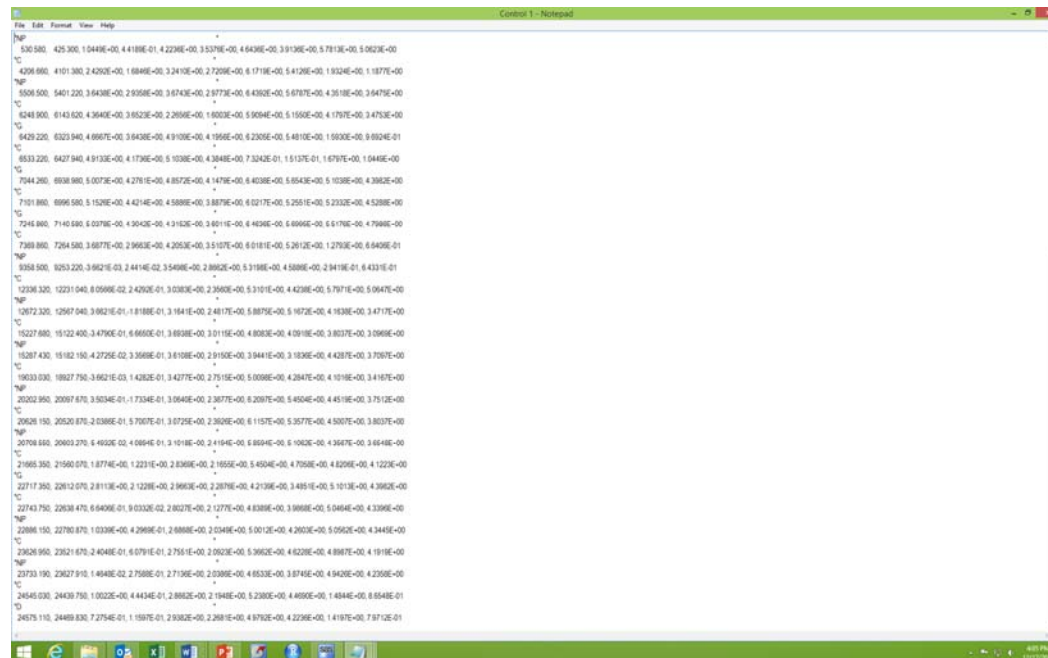
SAS Programs Read the raw data

- First look at the contents of one of the raw data files.

- Open "Control 1" in the "Psyllid Data" folder using "open with" by right clicking the file name to bring up the menu.
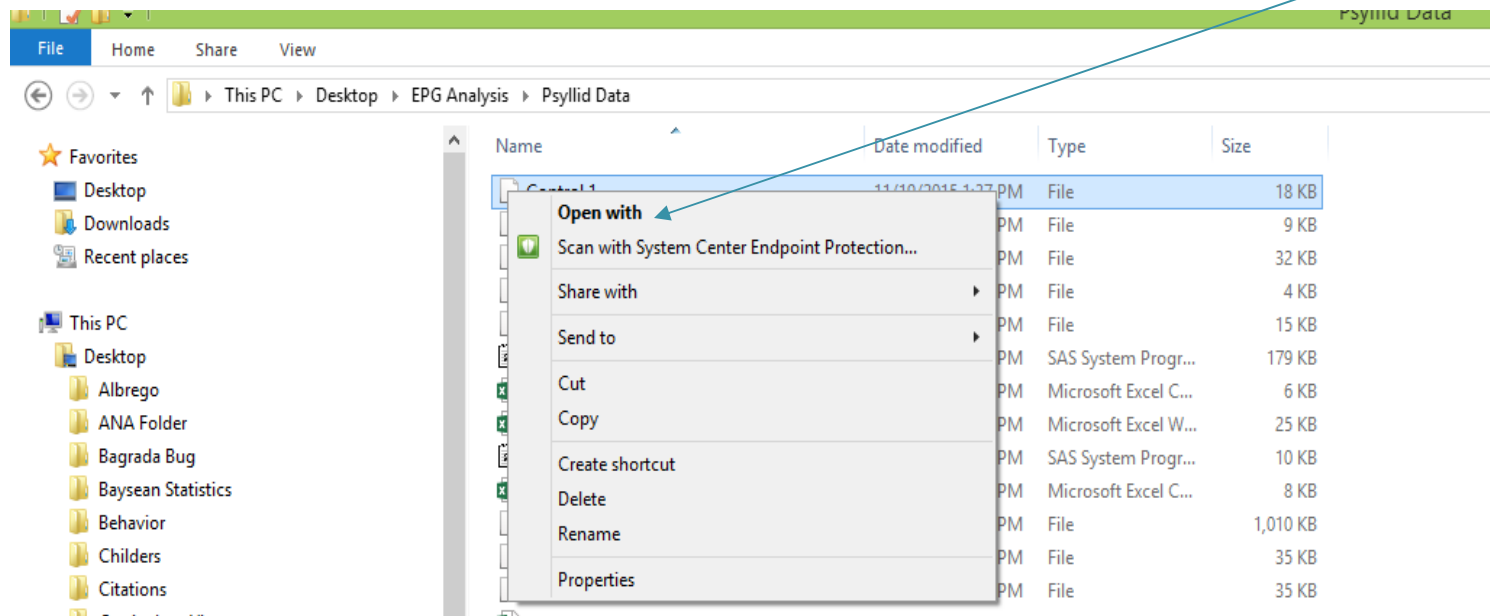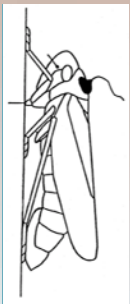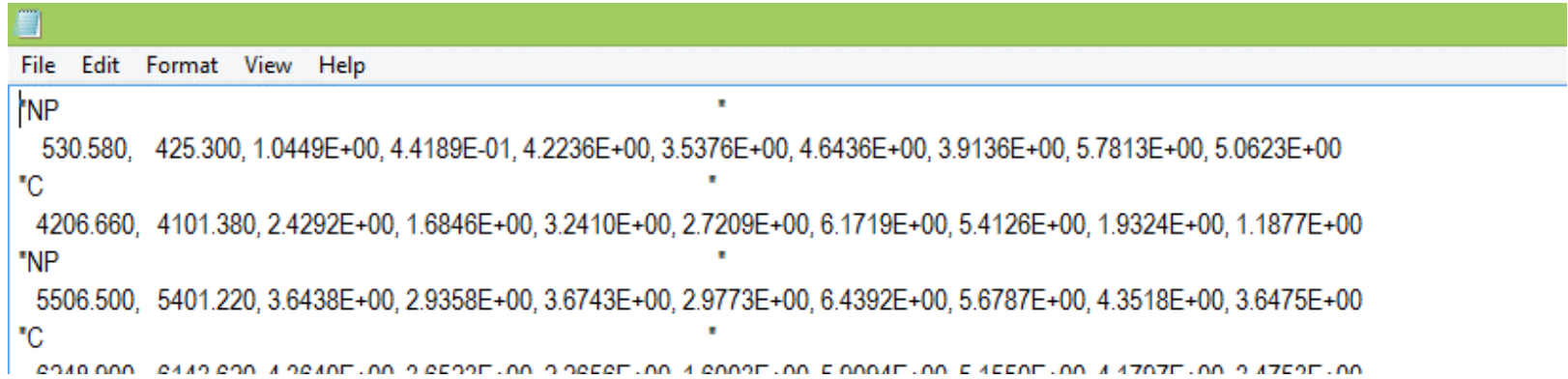
The Raw data

- Open with notepad or wordpad.
  - Note: These programs will show you the ASCII contents. No hidden formatting. No codes. What you see is an ASCII representation of your file contents.

Note the quotes and a long text string beginning with the waveform code (clearer on next slide). Numerical data is on the following line. This is very different from the ANA file format.
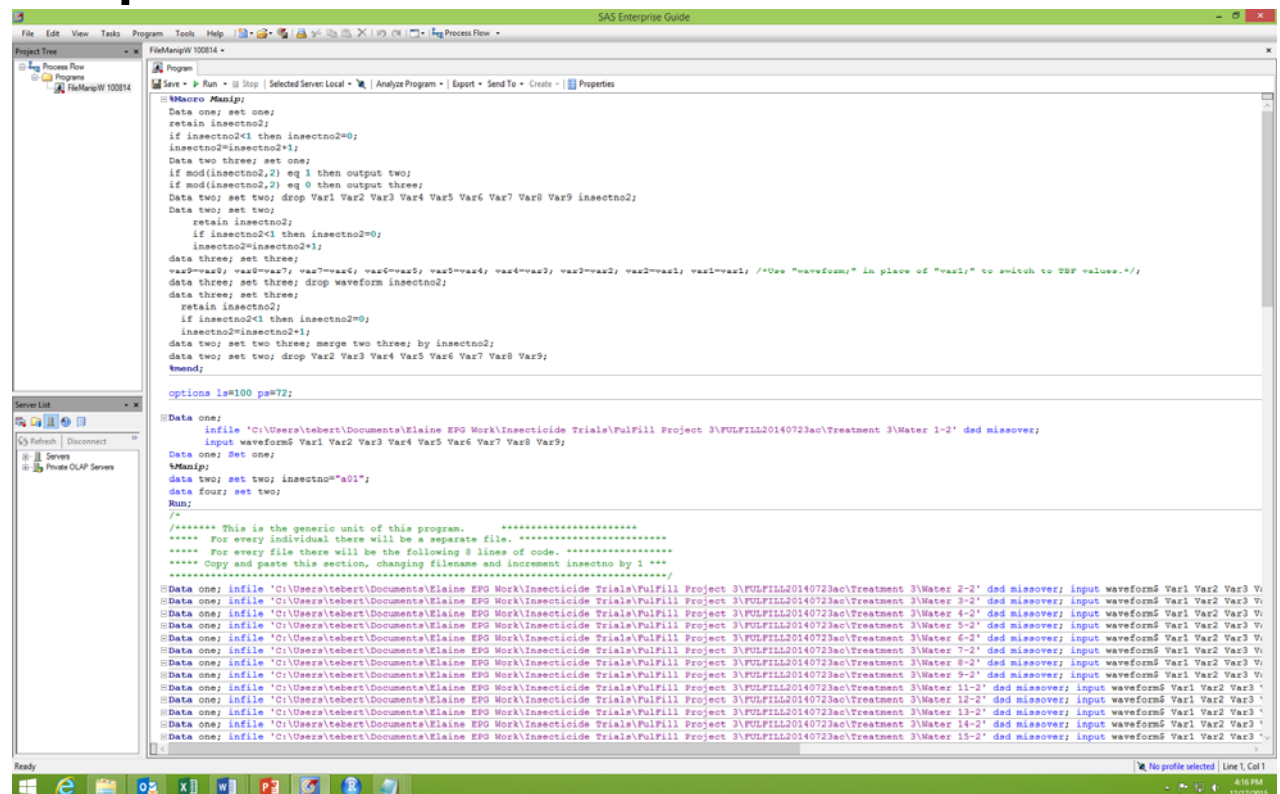
Details

Read the raw data

- Close notepad.
- Right click "FileManipW.sas" and open in SAS Enterprise Guide.

This is what you should see.

Details

Unique Insect number: "a" is the treatment.

This is the first insect

Path and File Name

```
Data one;
     infile 'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL2014C
     input waveform$ Var1 Var2 Var3 Var4 Var5 Var6 Var7 Var8 Var9;
Data one; Set one;
%Manip;
data two; set two; insectno="a01";
data four; set two;
Run;
/*
/******* This is the generic unit of this program.     ***********************
*****   For every individual there will be a separate file. **************************
*****   For every file there will be the following 8 lines of code. ******************
***** Copy and paste this section, changing filename and increment insectno by 1 ***
************************************************************************************/
Data one; infile 'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL2
Data one; infile 'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL2
Data one; infile C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL2
```

Server List

Refresh | Disconnect

- Servers
- Private OLAP Servers

2ⁿᵈ
3ʳᵈ
4ᵗʰ

Here are the second,
third, and fourth insects

Path and File Name

Note: there is text that
does not fit on screen.

Be aware that SAS wants a \ in the path name not a /

Modifying the Program

- The file names are wrong.
- Here is what is there now:

'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL20140723ac\Treatment 3\Water 1-2'

- Here is what needs to be there:

'C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data\Control 1'

- I suggest that you do not type this in.
- Go to "File Explorer"
- Single mouse click on the file "Control 1"

Modifying File Names

- Single click here and copy the now highlighted path.
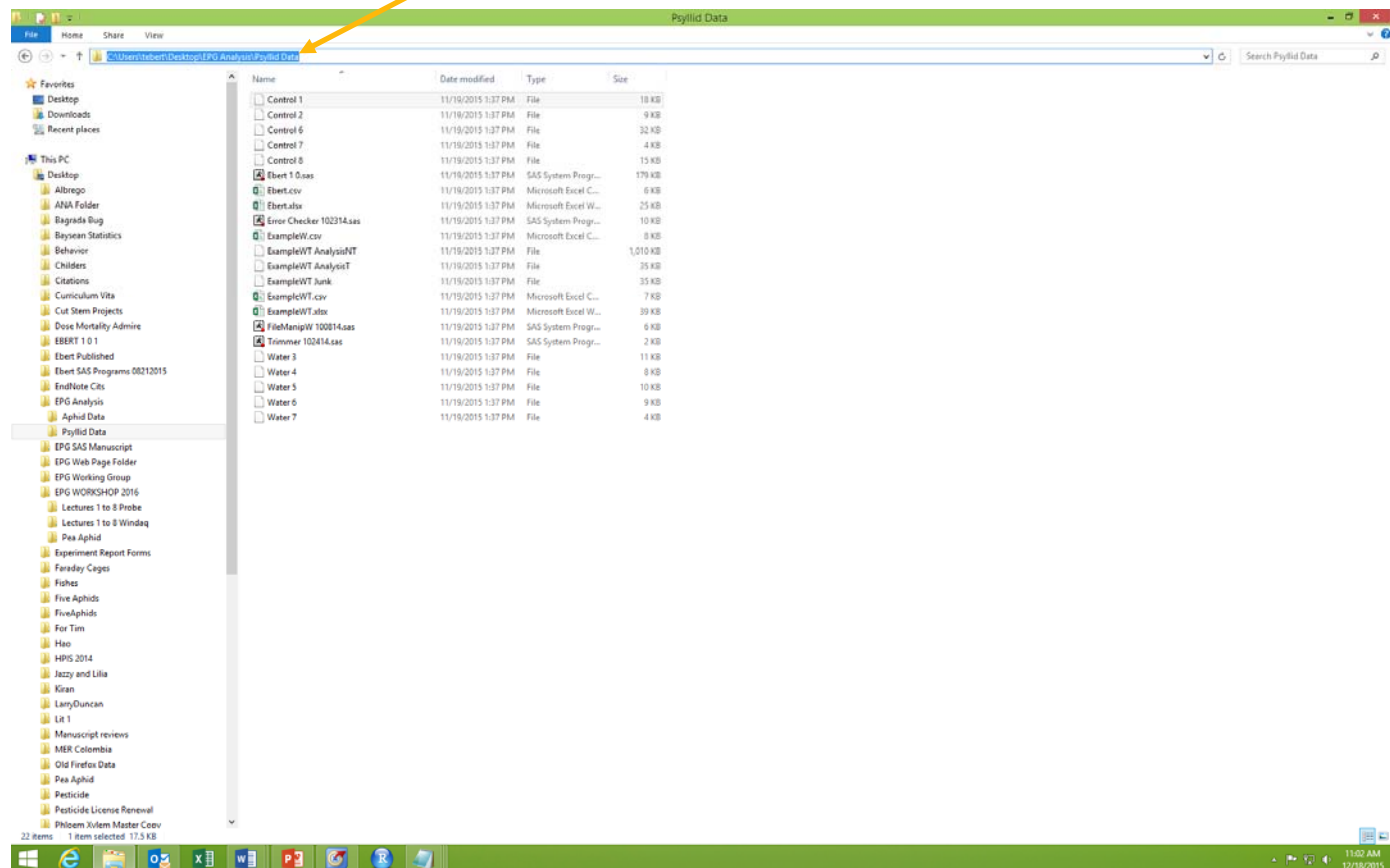
Details
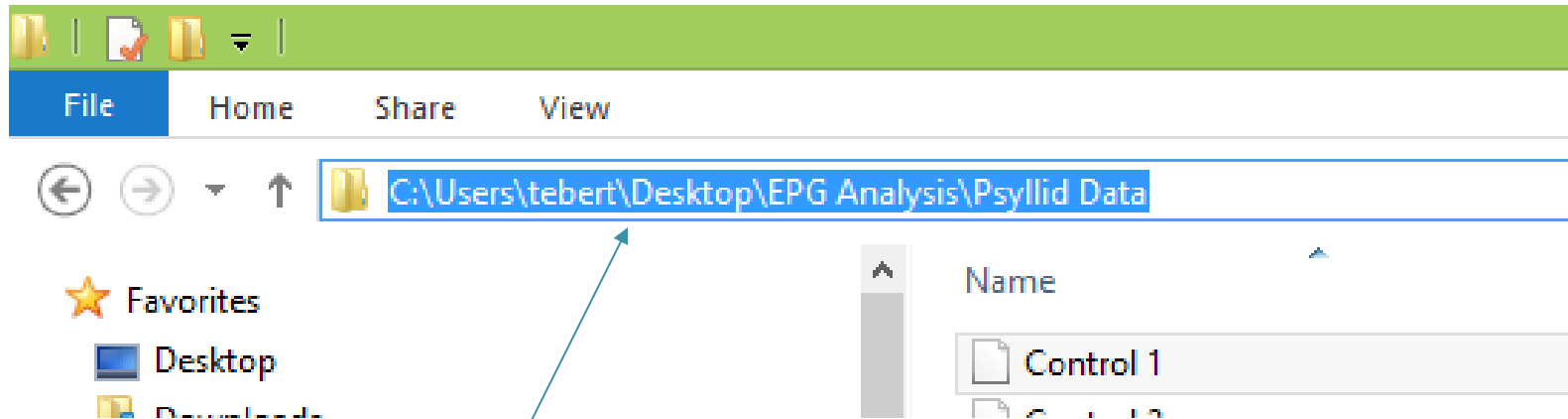
If you do not see this path, then you should change your settings for Explorer. See the small lecture on Path Names.

Change the file name

- Return to SAS, and highlight the old file name, and paste the new path here.
  - Original

    'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL20140723ac\Treatment 3\Water 1-2 '

  - New

    'C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data'

- Then type in a slash \ and the file name.
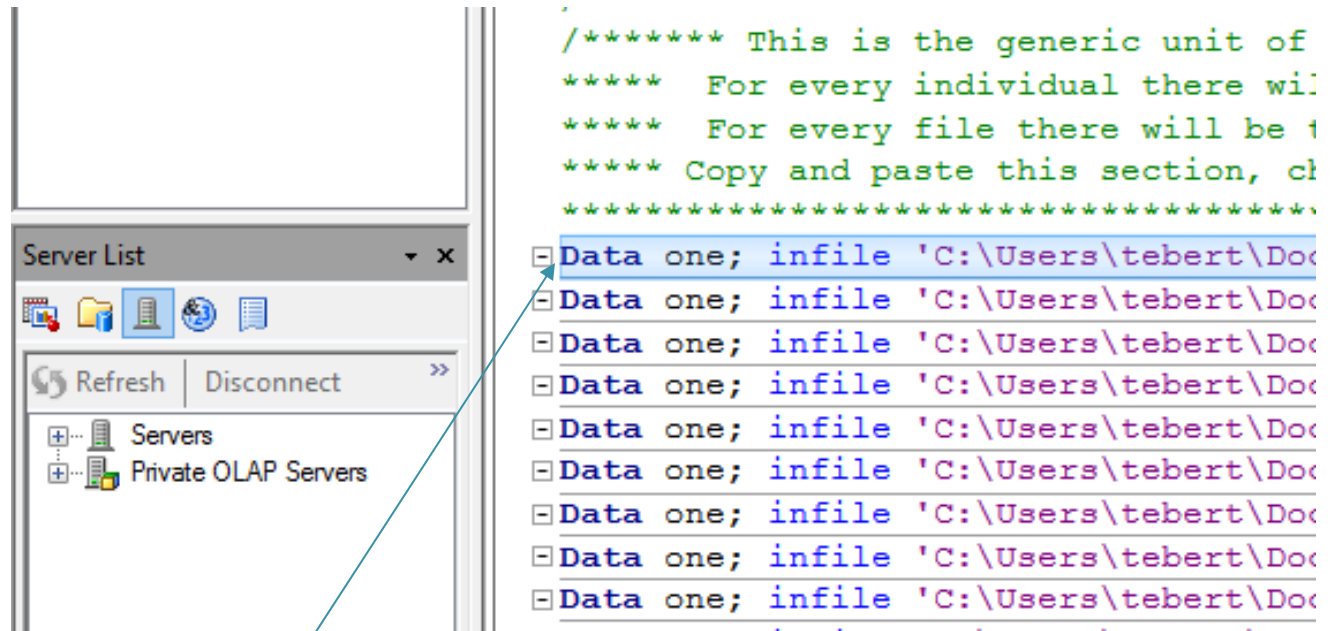
  'C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data\control 1'

Complete this task

- Copy the file name that you have just corrected, and paste it into the next five places where a file name is used.
- Change the file names to Control 2, Control 6, Control 7, Control 8, and Water 3.

More Data?

- We are missing Water 4 through Water 7.
- Use the previous approach to get these, so that there is one of these lines for all remaining files.

Highlight one line (the entire line)
Copy
And paste as many new lines as needed (3 in this case).

Correct Treatments

- Each insect must have a unique designation.

- Use treatment codes A, B, C.
  - Keep track of what A means in a separate location. It is easier for you to do this than for me to program the computer to deal with "Aphid Imidacloprid 3ml/l 25 mV 10Jan2015."
  - In this case control is A, and Water is B.

- Here is insect 4 in treatment A
  - You will need to scroll to the right to see this.
  - make changes as needed to this and the other segments.
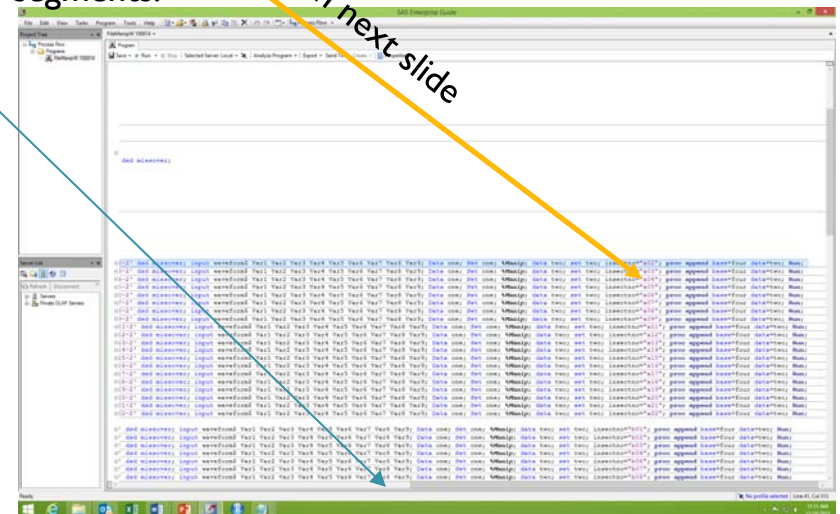
Detail on next slide

Detail

What you should have

- Delete all the extra lines that have old file names. The final results should look like this.

Note: a graphics problem produces a small error in this view. This should be Manip.

Saving output

- Change the path and file name for where to save the output (scroll down).
- The existing code looks like this:

proc export data=five outfile='C:\Users\tebert\Documents\Elaine EPG Work\Inecticide Trials\FulFill Project 3\Raw Data\Raw Data TRT3.csv' dbms=csv replace;

- It should now look like this:

proc export data=five outfile='C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data\PsyllidData1.csv' dbms=csv replace;

Run the program, Errors?

- Run the program. F3 will do this, or use menu.
- If SAS tells you that there is an error, then check the following:
  - Check the file names if SAS cannot find a file. Be aware of spaces in file names and differences between a 1 and an l (a one and a lower case L)
  - The quotes about file names where SAS says there is an error.
  - The quotes about the insect number.

Check the result

- Open file PsyllidData1.csv using Excel.
- You should have this:

Confirm the results

- In running a program that you have no experience in using, it is good to make a few observations to check the results.

  ○ Open the file "Control 1" using notepad and record the first two entries.

  "NP   "
    530.580,   425.300, 1.0449E+00, 4.4189E-01, 4.2236E+00, 3.5376E+00, 4.6436E+00, 3.9136E+00, 5.7813E+00, 5.0623E+00

  "C   "
    4206.660, 4101.380, 2.4292E+00, 1.6846E+00, 3.2410E+00, 2.7209E+00, 6.1719E+00, 5.4126E+00, 1.9324E+00, 1.1877E+00

  ○ Open the file "Water 7" using notepad and record the last two entries at the end of the file.

  "NP   "
    888.640,   840.000,-4.6997E-01,-1.1841E-01, 3.4912E-01, 2.2473E+00,-4.5776E-01, 6.8481E-01

  "C   "
    1397.440, 1348.800,-4.8340E-01,-1.2695E-01,-4.6387E-01, 1.4844E+00,-4.5532E-01, 6.7993E-01

Confirm the Results

- Take a look at the numbers in Control 1
  - It would be a great idea to enter them by hand into Excel.
- For Control 1, you will note that the difference between the first column and the second column is 105.28.
  - The first column is time from the start of recording.
  - The second column is time from when you told Windaq that you had placed the insect on the plant.
  - Use the second column of numbers.

Confirm the Results

- Note that the second column is a constantly increasing set of numbers. These values are all time from start of recording to the end of that behavior.

- If these were time from start of recording to the beginning of that behavior then the first behavior would be at time 0, and you would need a dummy waveform for the end of the final behavior (Code 12 is used in Probe).

- These are the two ways to calculate TBF (Time from Beginning of File).

Confirm the results

- The First two entries for the AphidData1 file are:

insectno	waveform	Dur
a01	NP	425.3
a01	C	3676.08

- And the last two entries are:

b07	E1	162.08
b07	E2	25279.5

- Subtract the first number from the second number from the second column: 4101.38-425.3 to get 3676.08.

Data from file "Control 1"

Confirm the Results

- Do the same thing for the last entry.
- The numbers match.
- We have greater confidence that the process worked as advertised.
- Note: There are 528 data entries and one line containing column headings in the Excel file.