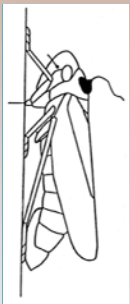


# EPG Data Analysis 101

- Ebert 1.0: Reading data files and checking your work
- by
- T.A. Ebert
- M.E. Rogers

EPG  
Work-  
shop





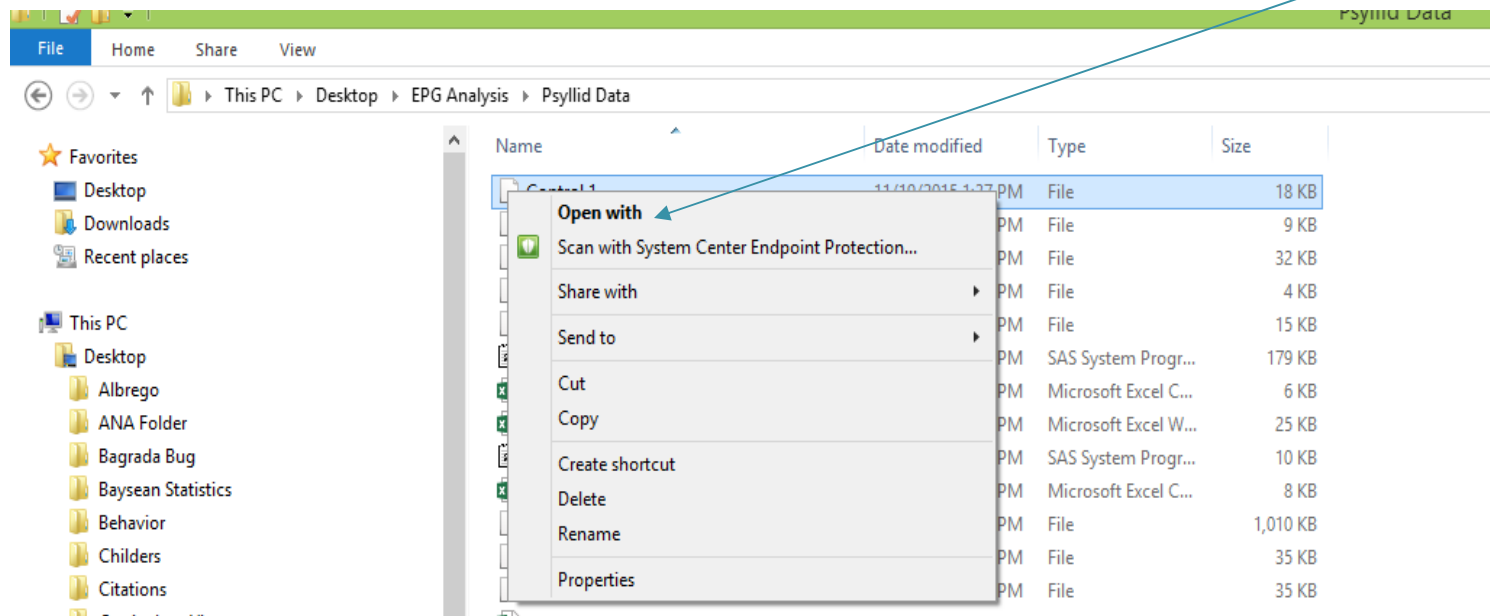
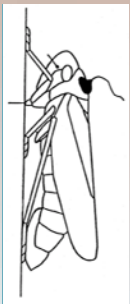
# Formatting Note

- 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

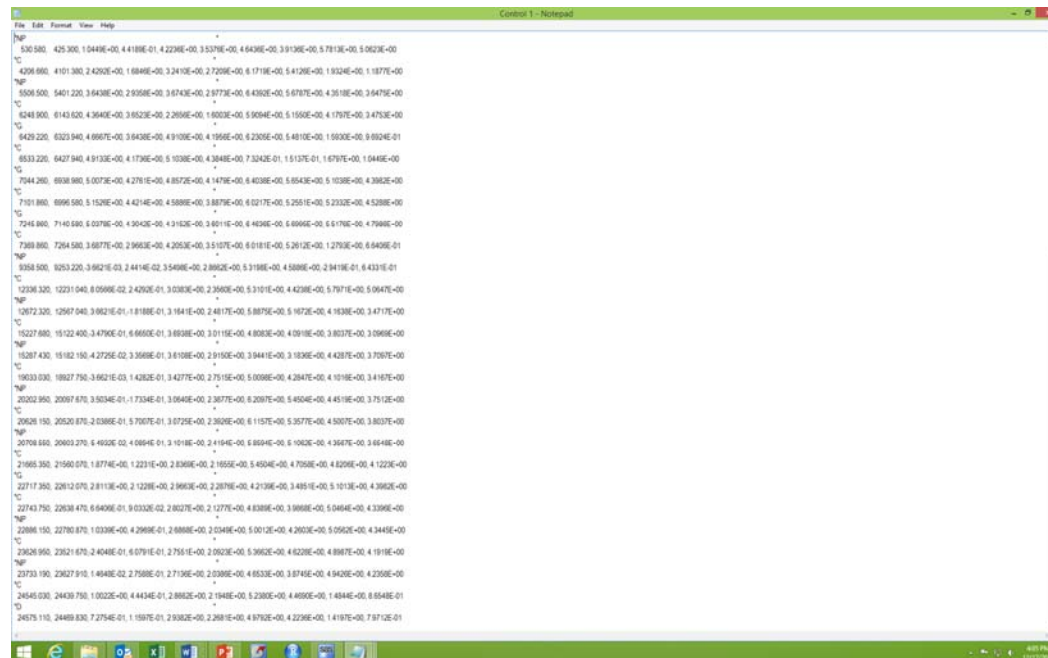
- First look at the contents of one of the raw data files.
- Open “Control 1” in the “Psyllid Data” folder using “open with” by right clicking the file name to bring up the menu.

EPG  
Work-  
shop



# The Raw data

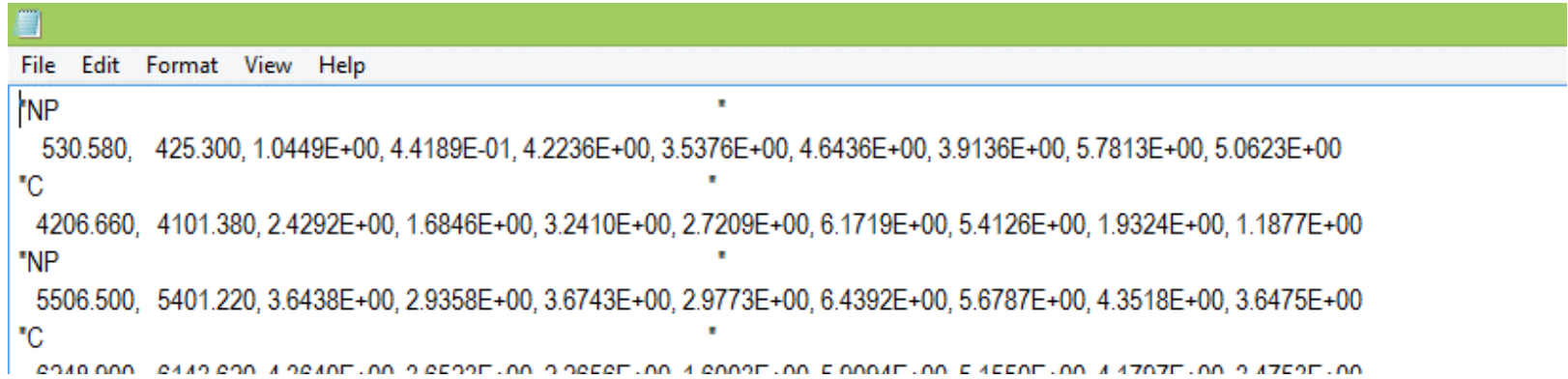
- Open with notepad or wordpad.
- Note: These programs will show you the ASCII contents. No hidden formatting. No codes. What you see is an ASCII representation of your file contents.



```
Control 1 - Notepad
530.580, 425.300, 1.0449E+00, 4.4188E-01, 4.2238E+00, 5.5378E+00, 6.6438E+00, 3.9138E+00, 5.7813E+00, 5.0823E+00
^
4258.880, 4101.300, 2.4292E+00, 1.6888E+00, 3.2419E+00, 2.7209E+00, 6.1719E+00, 5.4139E+00, 1.9324E+00, 1.1877E+00
^
5508.500, 5401.220, 3.6438E+00, 2.9358E+00, 3.6743E+00, 2.9773E+00, 6.4362E+00, 5.6787E+00, 4.3618E+00, 3.6473E+00
^
6243.900, 6143.620, 4.3643E+00, 3.8523E+00, 2.2658E+00, 1.8003E+00, 5.9094E+00, 5.1550E+00, 4.1797E+00, 3.4753E+00
^
8429.220, 8329.940, 4.8867E+00, 3.8438E+00, 4.9109E+00, 4.1998E+00, 6.2265E+00, 5.4819E+00, 1.9808E+00, 9.8934E-01
^
8533.220, 8427.940, 4.9133E+00, 4.1738E+00, 5.1038E+00, 4.3848E+00, 7.3342E-01, 1.5137E-01, 1.6797E+00, 1.0448E+00
^
7044.260, 6938.880, 5.0073E+00, 4.2781E+00, 4.8872E+00, 4.1479E+00, 6.4038E+00, 5.6543E+00, 5.1038E+00, 4.3882E+00
^
7101.880, 6996.580, 5.1529E+00, 4.4214E+00, 4.5889E+00, 3.8879E+00, 6.0217E+00, 5.2551E+00, 5.2332E+00, 4.5208E+00
^
7246.880, 7140.680, 5.0378E+00, 4.3042E+00, 4.3162E+00, 3.8011E+00, 4.4824E+00, 6.4866E+00, 5.6178E+00, 4.7988E+00
^
7389.880, 7284.580, 3.6877E+00, 2.9663E+00, 4.2053E+00, 3.5107E+00, 6.0181E+00, 5.2012E+00, 1.2793E+00, 6.6408E-01
^
9358.500, 9253.220, 3.6821E+00, 2.4414E+00, 3.5489E+00, 2.8862E+00, 5.3198E+00, 4.5888E+00, 2.9419E-01, 6.4318E-01
^
12338.320, 12231.940, 8.0588E+02, 2.4292E-01, 3.0383E+00, 2.3980E+00, 5.3101E+00, 4.4238E+00, 5.7971E+00, 5.0847E+00
^
18172.320, 12967.940, 3.8821E-01, 1.8188E-01, 3.1841E+00, 2.4817E+00, 5.8875E+00, 5.1872E+00, 4.1838E+00, 3.4717E+00
^
15227.880, 15122.400, 3.4790E-01, 6.6660E-01, 3.8838E+00, 3.0115E+00, 4.8803E+00, 4.1818E+00, 3.8837E+00, 3.0588E+00
^
15287.430, 15182.150, 4.2725E+02, 3.3698E-01, 3.8108E+00, 2.9150E+00, 3.9441E+00, 3.1839E+00, 4.4287E+00, 3.7507E+00
^
18033.030, 18867.750, 3.6621E-03, 1.4282E-01, 3.4077E+00, 2.7519E+00, 5.0868E+00, 4.2847E+00, 4.1018E+00, 3.4187E+00
^
20302.880, 20087.670, 3.9334E-01, 1.7334E-01, 3.0648E+00, 2.3877E+00, 6.2087E+00, 5.4504E+00, 4.4518E+00, 3.7512E+00
^
20828.150, 20520.870, 2.8388E-01, 5.7087E-01, 3.9725E+00, 2.3628E+00, 6.1157E+00, 5.3677E+00, 4.5307E+00, 3.8637E+00
^
20708.880, 20683.270, 6.4920E+02, 4.0884E-01, 3.1018E+00, 2.4194E+00, 6.8894E+00, 5.1042E+00, 4.3647E+00, 3.6648E+00
^
21885.300, 21560.070, 1.8748E+00, 1.2231E+00, 2.8388E+00, 2.1855E+00, 5.4504E+00, 4.7598E+00, 4.8288E+00, 4.1222E+00
^
22717.350, 22612.070, 2.8113E+00, 2.1228E+00, 2.8863E+00, 2.2878E+00, 4.2198E+00, 3.4891E+00, 5.1013E+00, 4.3862E+00
^
22743.750, 22638.470, 6.6408E-01, 9.0332E+02, 2.8027E+00, 2.1277E+00, 4.8388E+00, 3.9888E+00, 4.0484E+00, 4.3388E+00
^
22888.100, 22780.870, 1.0338E+00, 4.2988E-01, 2.8888E+00, 2.0348E+00, 5.0012E+00, 4.2002E+00, 5.0592E+00, 4.3448E+00
^
23828.950, 23621.670, 2.4048E-01, 6.0791E-01, 2.7515E+00, 2.0002E+00, 5.3862E+00, 4.6228E+00, 4.8887E+00, 4.1918E+00
^
23733.190, 23627.910, 1.4848E+02, 2.7588E-01, 2.7198E+00, 2.0388E+00, 4.8833E+00, 3.8748E+00, 4.9408E+00, 4.2308E+00
^
24545.030, 24439.750, 1.0022E+00, 4.4434E-01, 2.8932E+00, 2.1948E+00, 5.2008E+00, 4.8895E+00, 1.4848E+00, 8.6648E-01
^
24575.110, 24469.830, 7.2754E-01, 1.1887E-01, 2.9382E+00, 2.2881E+00, 4.8792E+00, 4.2298E+00, 1.4197E+00, 7.8712E-01
```

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

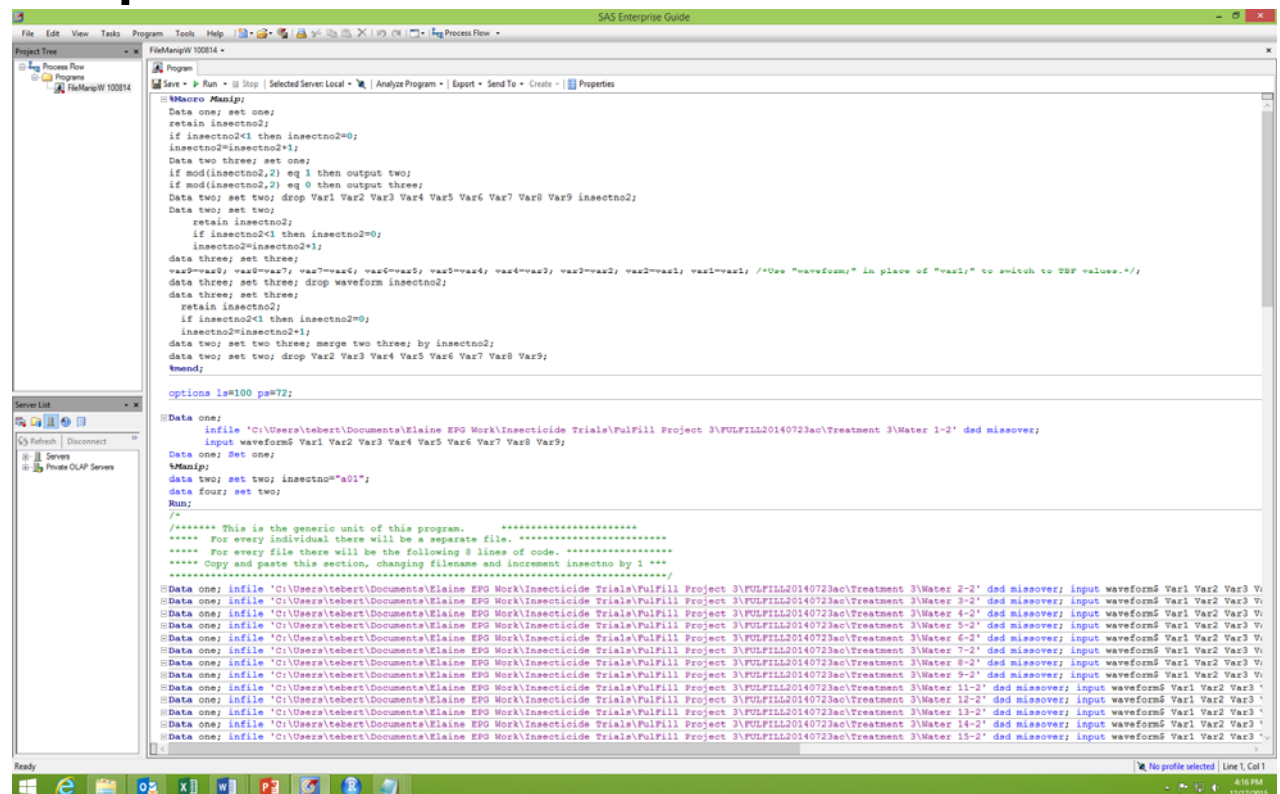


*NP					*					
	530.580,	425.300,	1.0449E+00,	4.4189E-01,	4.2236E+00,	3.5376E+00,	4.6436E+00,	3.9136E+00,	5.7813E+00,	5.0623E+00
*C					*					
	4206.660,	4101.380,	2.4292E+00,	1.6846E+00,	3.2410E+00,	2.7209E+00,	6.1719E+00,	5.4126E+00,	1.9324E+00,	1.1877E+00
*NP					*					
	5506.500,	5401.220,	3.6438E+00,	2.9358E+00,	3.6743E+00,	2.9773E+00,	6.4392E+00,	5.6787E+00,	4.3518E+00,	3.6475E+00
*C					*					
	6240.000,	6142.620,	4.2840E+00,	3.6522E+00,	3.2858E+00,	4.8002E+00,	5.0004E+00,	5.4550E+00,	4.4707E+00,	3.4752E+00

# Read the raw data

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

This is what you should see.



```
File ManipW 100814
Program
Save Run Stop Selected Server: Local Analyze Program Export Send To Create Properties
Macro Manip;
Data one; set one;
retain insectno2;
if insectno<1 then insectno2=0;
insectno2=insectno2+1;
Data two three; set one;
if mod(insectno2,2) eq 1 then output two;
if mod(insectno2,2) eq 0 then output three;
Data two; set two; drop Var1 Var2 Var3 Var4 Var5 Var6 Var7 Var8 Var9 insectno2;
Data two; set two;
retain insectno2;
if insectno<1 then insectno2=0;
insectno2=insectno2+1;
data three; set three;
var2=var5; var3=var7; var4=var6; var5=var4; var6=var3; var7=var2; var8=var1; var9=var1; /*Use "waveform" in place of "var1" to switch to TSF values.*/
data three; set three; drop waveform insectno2;
data three; set three;
retain insectno2;
if insectno<1 then insectno2=0;
insectno2=insectno2+1;
data two; set two three; merge two three; by insectno2;
data two; set two; drop Var2 Var3 Var4 Var5 Var6 Var7 Var8 Var9;
%mend;

options ls=100 ps=72;

Data one;
infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 1-2' dds misover;
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 # lines of code. *****/
**** Copy and paste this section, changing filename and increment insectno by 1 ****/
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 2-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 3-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 4-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 5-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 6-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 7-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 8-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 9-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 10-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 11-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 12-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 13-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 14-2' dds misover; input waveform Var1 Var2 Var3 V
Data one; infile 'C:\Users\tebert\Documents\Eline EPG Work\Insecticide Trials\PulFill Project 3\PULFILL20140723ac\Treatment 3\Water 15-2' dds misover; input waveform Var1 Var2 Var3 V
Ready
No profile selected | Line 1, Col 1
4:16 PM
12/7/2005
```

# Details

Unique Insect number: "a" is the treatment.

This is the first insect

Path and File Name

```
Server List
Refresh Disconnect
Servers
Private OLAP Servers

Data one;
  infile 'C:\Users\tebert\Documents\Elaine EPG Work\Insecticide Trials\FulFill Project 3\FULFILL20140';
  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
```

2<sup>nd</sup>

3<sup>rd</sup>

4<sup>th</sup>

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 I-2'

- Here is what needs to be there:

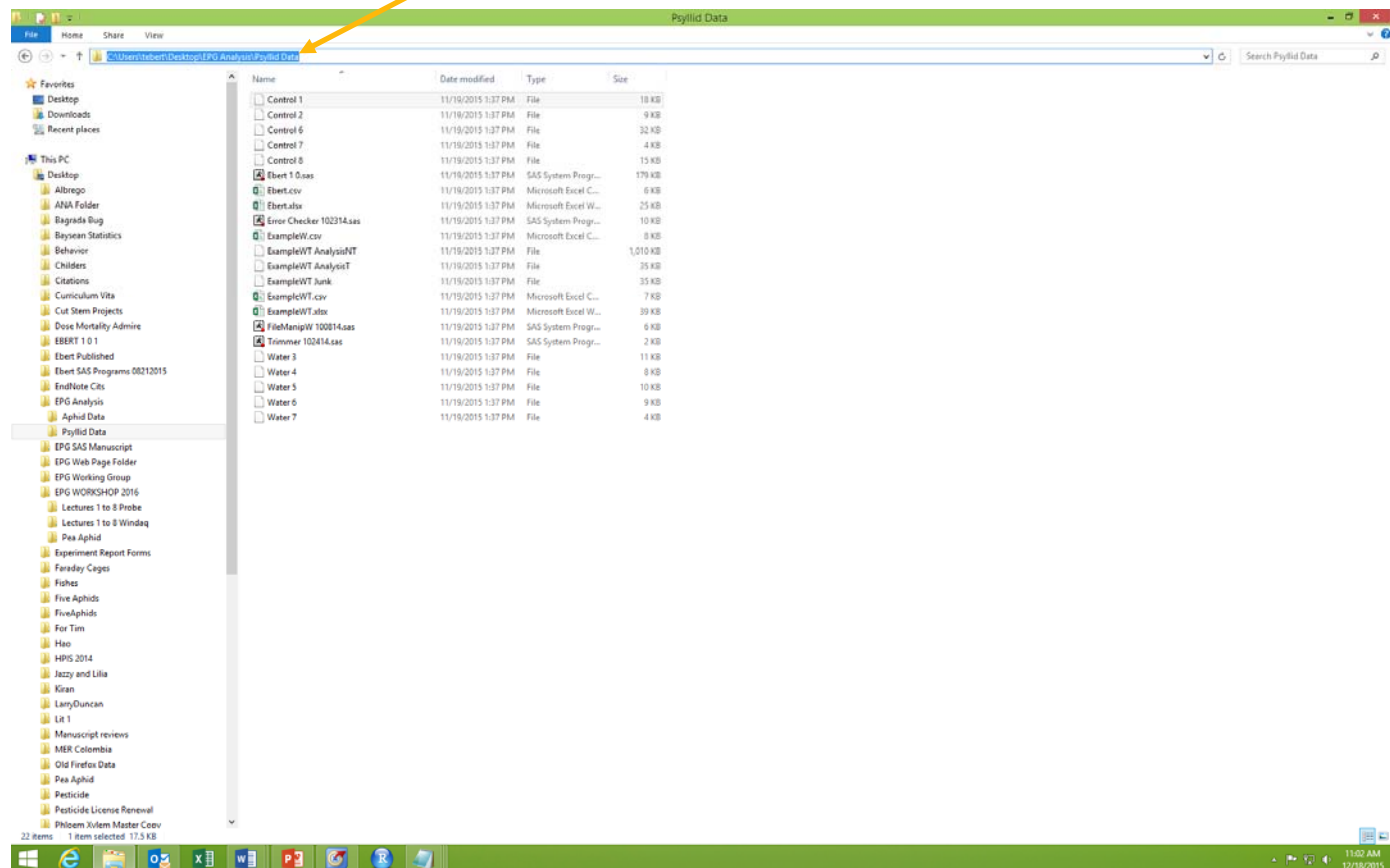
'C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data\Control I'

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

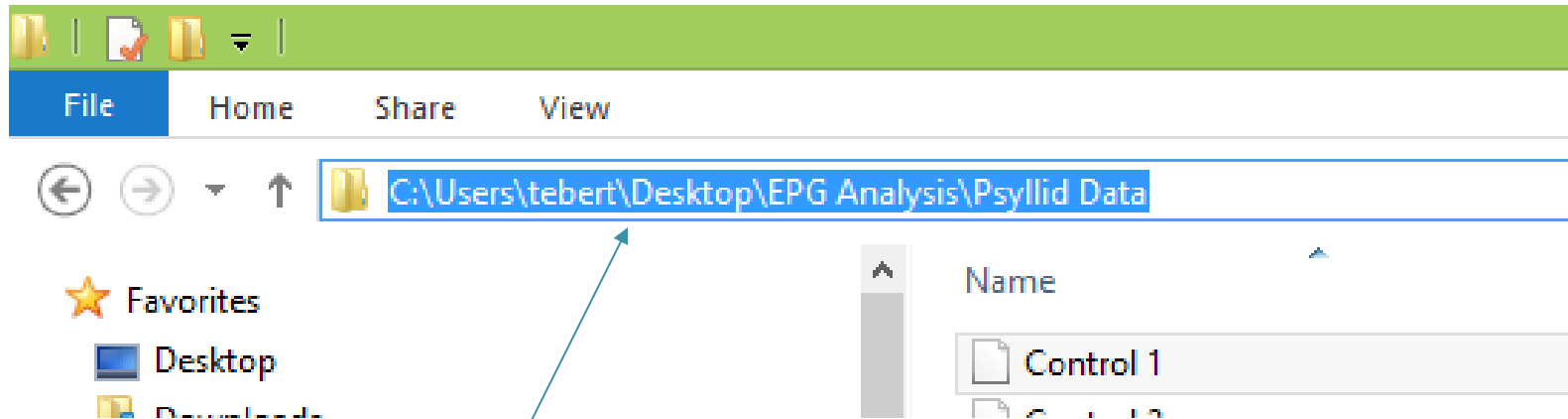


# Modifying File Names

- 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.

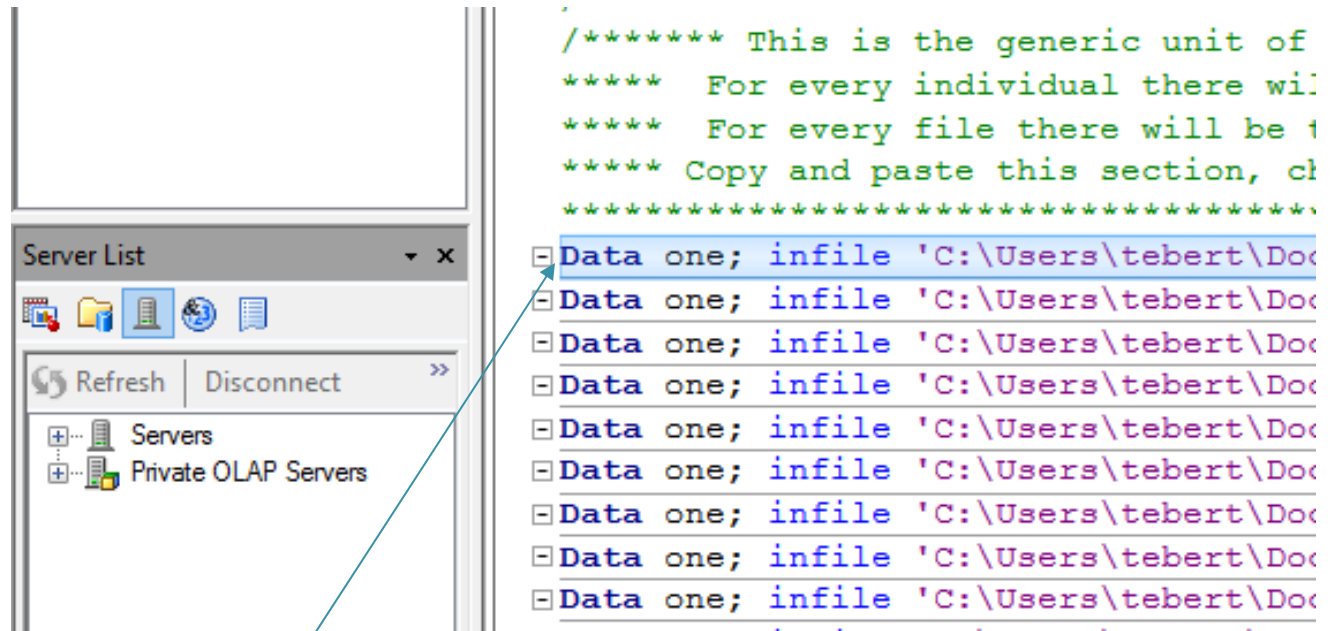
- 'C:\Users\tebert\Desktop\EPG Analysis\Psyllid Data\control 1'



## 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.





```
./***** This is the generic unit of
**** For every individual there wil
**** For every file there will be t
**** Copy and paste this section, ch
*****/
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
Data one; infile 'C:\Users\tebert\Doc
```

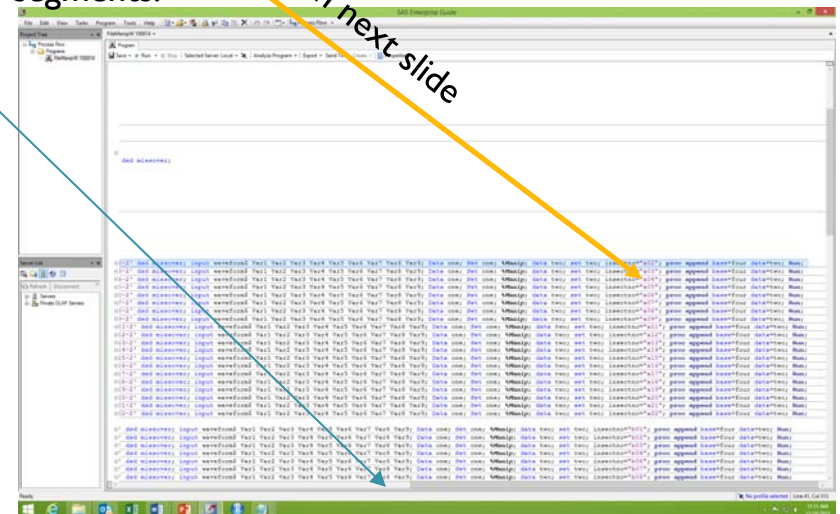
Highlight one line (the entire line)

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 3m/I 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



```
two; insectno="a02"; proc append base=fou  
two; insectno="a03"; proc append base=fou  
two; insectno="a04"; proc append base=fou  
two; insectno="a05"; proc append base=fou  
two; insectno="a06"; proc append base=fou  
two; insectno="a07"; proc append base=fou  
two; insectno="a08"; proc append base=fou  
two; insectno="a09"; proc append base=fou  
t two; insectno="a11"; proc append base=fc  
t two; insectno="a12"; proc append base=fc  
t two; insectno="a13"; proc append base=fc  
t two; insectno="a14"; proc append base=fc  
t two; insectno="a15"; proc append base=fc
```



# What you should have

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

```

***** 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 filenames and incrementing loopup by 1. *****
*****
*****
***** End Fundamental unit *****
***** Paste one copy for each individual and make changes *****
***** Do not copy or paste anything below this line *****
***** Please change the file name in the Proc Export line *****
***** before running the program, by old copy of *****
***** this file will be replaced due to the "replace" statement *****
*****
****data four; set four; waveform=compress(waveform);

```

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\Insecticide 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.



# 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 AphidData I 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 I"





## 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.