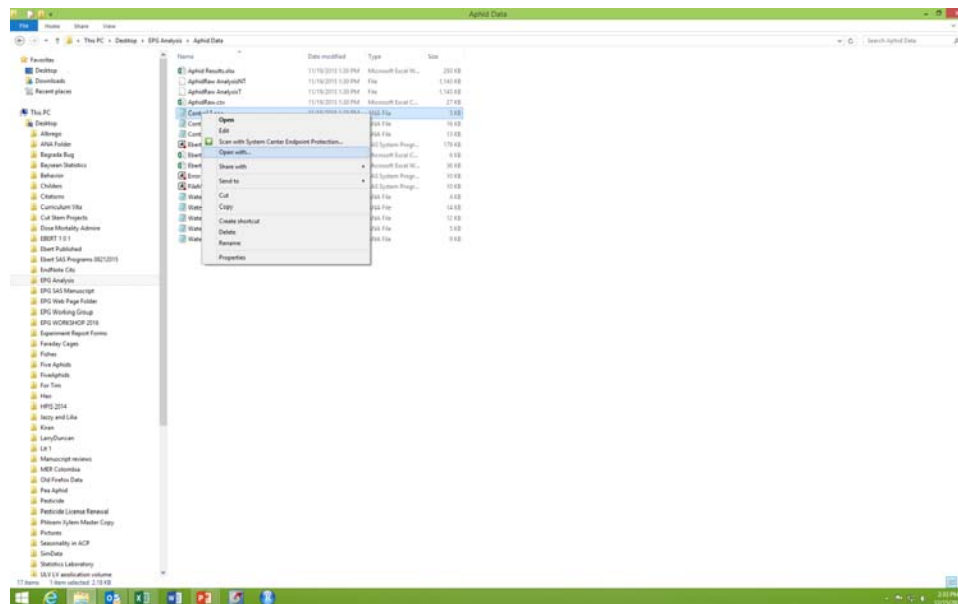# SAS Programs Read the raw data

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

- Open control 1.ana using "open with" by right clicking the file name to bring up the menu.
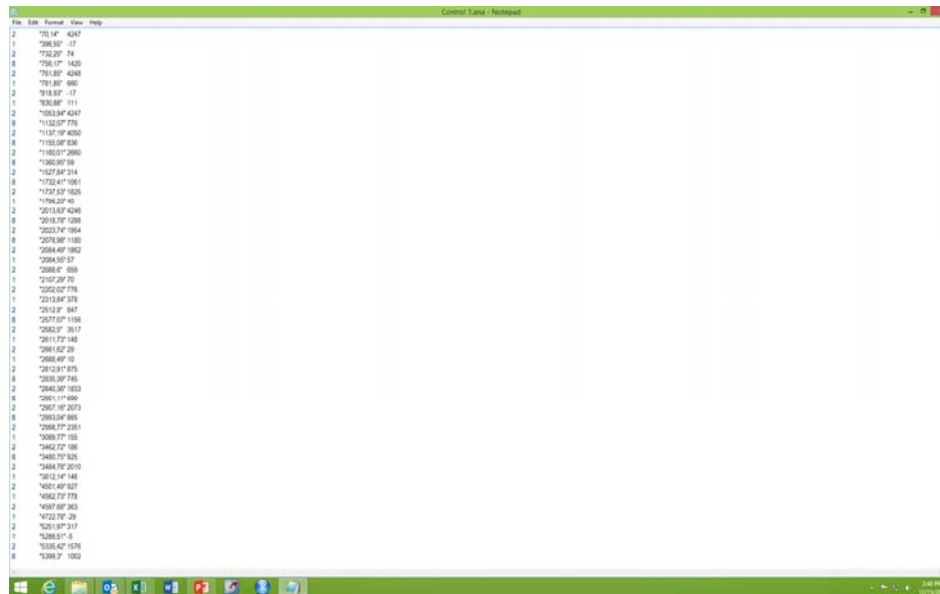
The Raw data

- 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.
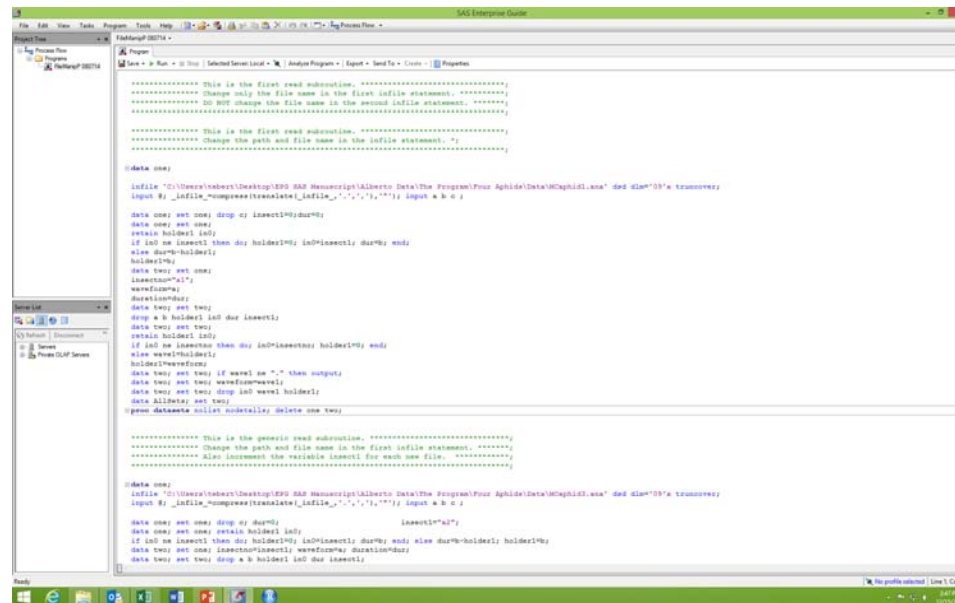

- Note the quotes, and the comma for a decimal

Read the raw data

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

Modifying the Program

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

'C:\Users\tebert\Desktop\EPG SAS Manuscript\Alberto Data\The Program\Four Aphids\Data\MCaphid1.ana'

- Here is what needs to be there:

'C:\Users\tebert\Desktop\EPG Analysis\Aphid Data\Control 1.ana'

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

Modifying File Names

- Single click here and copy the now highlighted path.
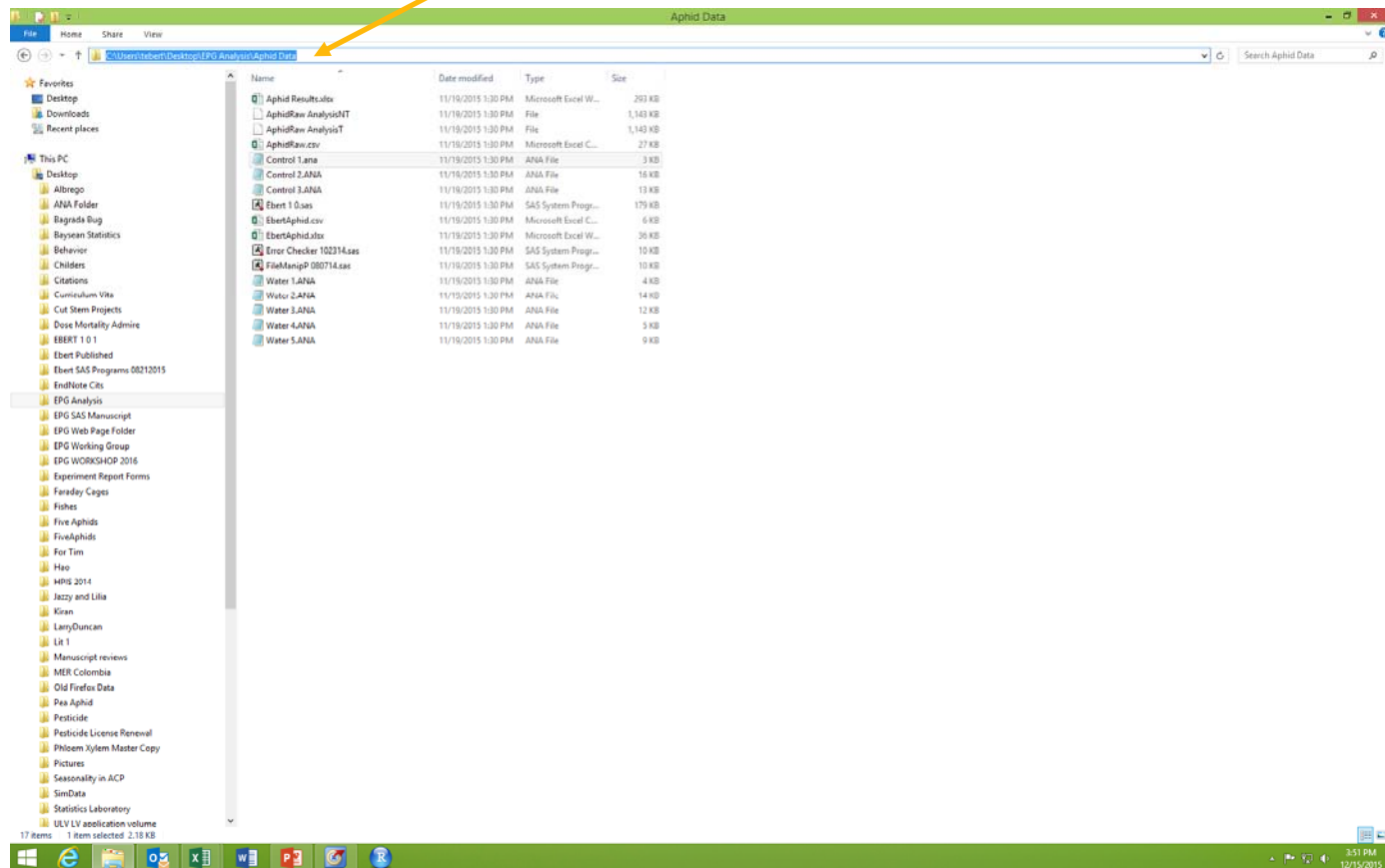
Change the file name

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

    'C:\Users\tebert\Desktop\EPG SAS Manuscript\Alberto Data\The Program\Four Aphids\Data\MCaphid1.ana'

  - New

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

- Then type in a slash and the file name.

  'C:\Users\tebert\Desktop\EPG Analysis\Aphid Data/control 1.ana'

Complete this task

- Copy the file name that you have just corrected, and paste it into the next three places where a file name is used.
- Change the file names to Control 2.ana, Control 3.ana, and Water 1.ana.

More Data?

- We are missing Water 2.ana through Water 5.ana.

- Copy the relevant section, and then change the file name.

  - The relevant section is all of this, and there must be one block like this for every file:

```
data one;
infile 'C:\Users\tebert\Desktop\EPG Analysis\Aphid Data/control 1.ana' dsd dlm='09'x
truncover;
input @; _infile_=compress(translate(_infile_,'.',','),'"'); input a b c ;

data one; set one; drop c; dur=0;                              insect1="a4";
data one; set one; retain holder1 in0;
if in0 ne insect1 then do; holder1=0; in0=insect1; dur=b; end; else dur=b-holder1;
holder1=b;
data two; set one; insectno=insect1; waveform=a; duration=dur;
data two; set two; drop a b holder1 in0 dur insect1;
data two; set two; retain holder1 in0;
if in0 ne insectno then do; in0=insectno; holder1=0; end; else wave1=holder1;
holder1=waveform;
data two; set two; if wave1 ne "." then output; data two; set two; waveform=wave1;
data two; set two; drop in0 wave1 holder1; data two; set two; proc append base=allsets
data=two;
proc datasets nolist nodetails; delete one two;
```

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

  ◦ make changes as needed to this and the other segments.

```
data one;
infile 'C:\Users\tebert\Desktop\EPG Analysis\Aphid Data/control 1.ana' dsd dlm='09'x
truncover;
input @; _infile_=compress(translate(_infile_,'.',','),'""'); input a b c ;

data one; set one; drop c; dur=0;                                    insect1="a4";
data one; set one; retain holder1 in0;
if in0 ne insect1 then do; holder1=0; in0=insect1; dur=b; end; else dur=b-holder1;
holder1=b;
data two; set one; insectno=insect1; waveform=a; duration=dur;
data two; set two; drop a b holder1 in0 dur insect1;
data two; set two; retain holder1 in0;
if in0 ne insectno then do; in0=insectno; holder1=0; end; else wave1=holder1;
holder1=waveform;
data two; set two; if wave1 ne "." then output; data two; set two; waveform=wave1;
data two; set two; drop in0 wave1 holder1; data two; set two; proc append base=allsets
data=two;
proc datasets nolist nodetails; delete one two;
```

Saving output

- Change the path and file name for where to save the output.

- The existing code looks like this:

proc export data=allsets outfile='C:\Users\tebert\Desktop\EPG SAS Manuscript\Alberto Data\The Program\Four Aphids\MCaphidRaw.csv' dbms=csv replace;

- It should now look like this:

proc export data=allsets outfile='C:\Users\tebert\Desktop\EPG Analysis\Aphid Data\AphidData1.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 AphidData1.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.ana using notepad and record the first two entries.

    | 2 | "70,14" | 4247 |
    | 1 | "396,55" | -17 |

  - Open the file Water 5.ana using notepad and record the last two entries at the end of the file.

    | 2 | 21523,28 | -225 |
    | 12 | 21599,61 | -81 |

Confirm the results

- The First entry for the AphidData1 file is:

ins	wave	dur
a1	C	326.41

- And the last entry is:

  b5 C 76.33

- In looking at the FileManip, one can see that if the waveform code is 2 then the waveform is C. The input matches the output.

- Now subtract the first number in quotes from the second.

  - You should get 396.55-70.14 = 326.41 and 21599.61-21523.28=76.33. This matches the data in AphidData1.csv.

  - Our confidence is now much higher that the program works properly.

  - Note that there are a total of 2257 observations in this data set. The Excel file has a header and 2257 rows of data. This number may help us later on.