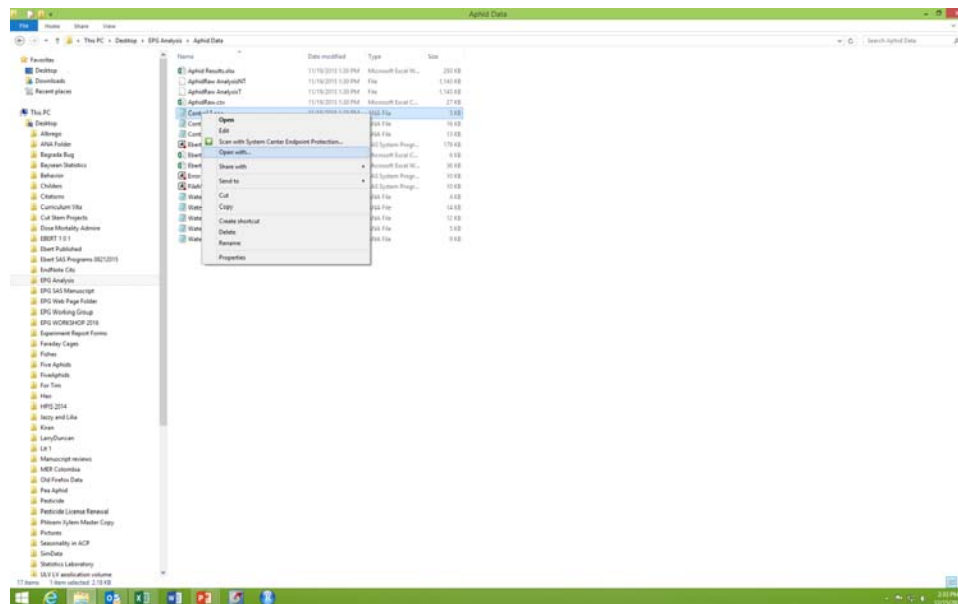


SAS Programs Read the raw data

- 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.

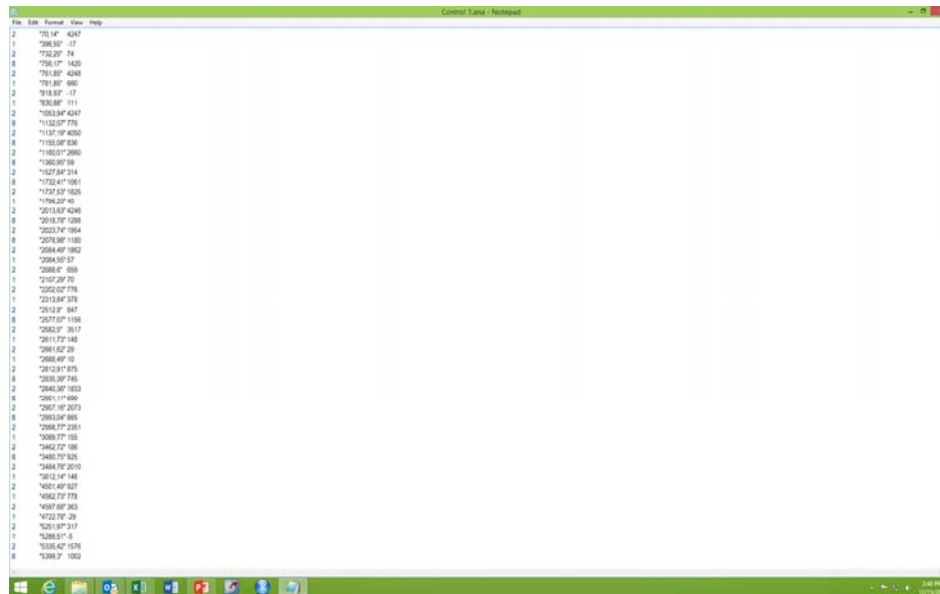


EPG
Work-
shop



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.

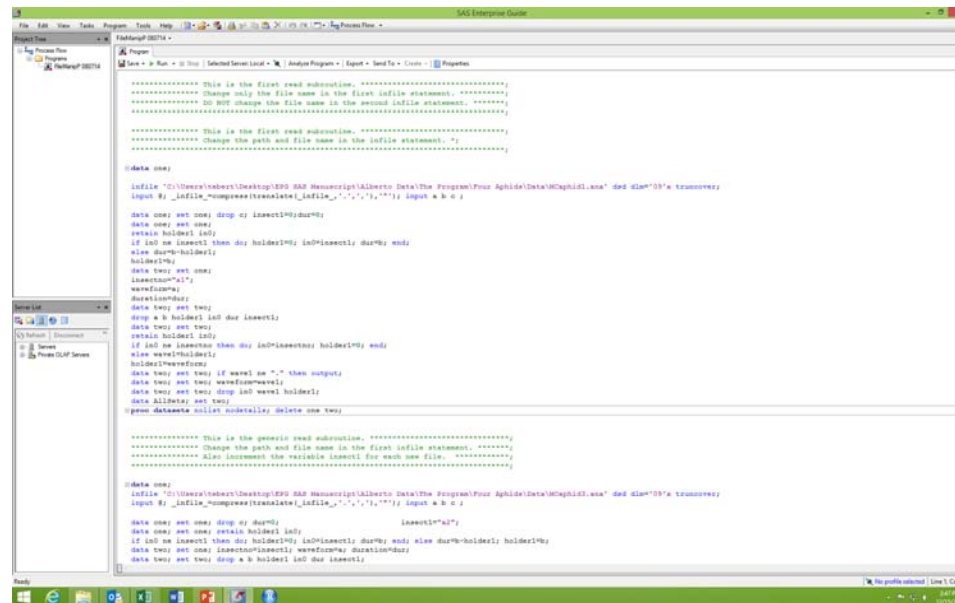


```
File Edit Format View Help
2 *30.14* 4247
1 *306.50* -17
2 *332.20* 74
8 *396.17* 1420
2 *761.80* 4248
1 *781.80* 866
2 *818.80* -17
1 *830.80* 1511
2 *1003.84* 4247
8 *1132.07* 776
2 *1137.10* 4500
8 *1165.00* 536
2 *1180.07* 2060
8 *1262.00* 78
2 *1278.43* 314
8 *1328.47* 1661
2 *1373.57* 1626
1 *1396.27* 45
2 *1513.57* 4248
8 *1518.70* 1288
2 *1523.71* 1664
8 *1578.80* 1180
2 *1584.40* 1802
1 *1684.07* 15
2 *1688.0* 858
1 *1707.00* 70
2 *2202.00* 776
1 *2313.84* 376
2 *2512.0* 847
8 *2573.07* 1196
2 *2582.7* 2617
1 *2611.77* 448
2 *2681.62* 20
1 *2688.00* 10
2 *2812.81* 875
8 *2835.00* 746
2 *2845.00* 1633
8 *2961.11* 499
2 *2967.10* 2073
8 *2993.04* 883
2 *2998.77* 2361
1 *3080.77* 166
2 *3462.72* 186
8 *3483.70* 925
2 *3484.70* 2010
1 *3812.14* 448
2 *4001.00* 167
1 *4062.72* 773
2 *4097.00* 363
1 *4722.00* 20
2 *5051.00* 317
2 *5306.14* 5
2 *5306.40* 1676
8 *5398.7* 1002
```

- 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.



```
***** This is the first read substitution. *****
***** Change only the file name in the first infile statement. *****
***** DO NOT change the file name in the second infile statement. *****
*****

***** This is the first read substitution. *****
***** Change the path and file name in the infile statement. *****
*****

@data one;
infile "C:\Users\alberto\Desktop\SGS SAS Memorign\Alberto Data\The Program\Four Aphids\Data\Omphid.sas" dds dds="09's transover;
input # _infile_ @@;

data one; set one; drop #;
retain holder;
if not # insecnt then do; holder=; insecnt=; dur=; end;
else dur=holder;
holder=;
data two; set one;
insecnt="1";
wareform=;
duration=;
data two; set two;
drop # b holder;
retain holder;
if not # insecnt then do; holder=;
else wareform=;
holder=wareform;
data two; set two; if wareform="." then output;
data two; set two; drop insecnt;
data AllData; set two;
proc datasets nolist nodetail; delete one two;

***** This is the generic read substitution. *****
***** Change the path and file name in the first infile statement. *****
***** Also comment the variable list for each use file. *****
*****

@data one;
infile "C:\Users\alberto\Desktop\SGS SAS Memorign\Alberto Data\The Program\Four Aphids\Data\Omphid.sas" dds dds="09's transover;
input # _infile_ @@;

data one; set one; drop #;
retain holder;
if not # insecnt then do; holder=; insecnt=; dur=; end; else dur=holder; holder=;
data two; set one; insecnt=; wareform=; duration=;
data two; set two; drop # b holder;
proc datasets nolist nodetail; delete one two;
```



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\MCaphid I.ana'

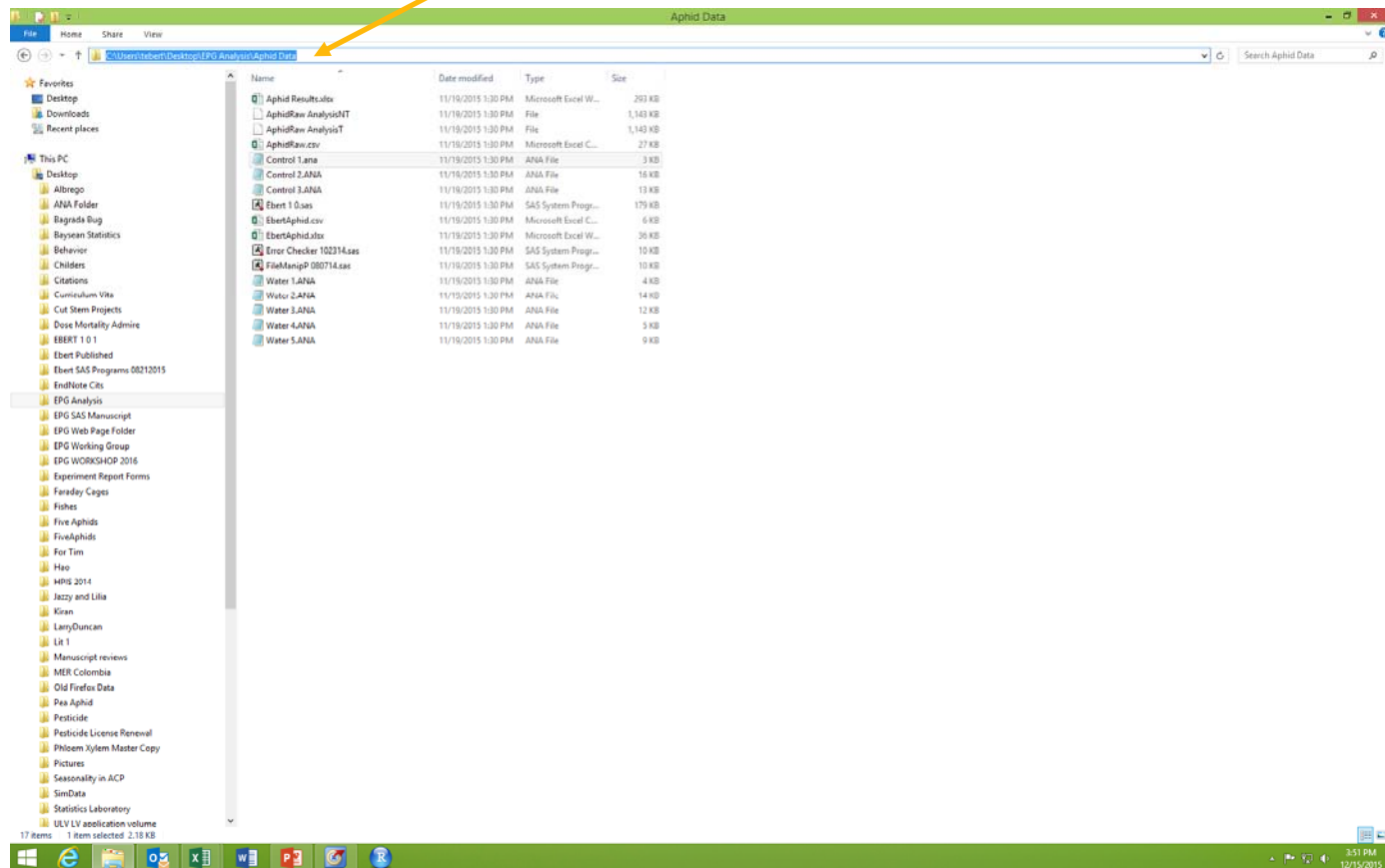
- Here is what needs to be there:

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

- I suggest that you do not type this in.
- Go to “File Explorer”
- Single mouse click on the file “Control I.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\MCaphid 1.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 wavel=holder1;
holder1=waveform;
data two; set two; if wavel ne "." then output; data two; set two; waveform=wavel;
data two; set two; drop in0 wavel 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 wavel=holder1;
holder1=waveform;
data two; set two; if wavel ne "." then output; data two; set two; waveform=wavel;
data two; set two; drop in0 wavel 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\AphidData I.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 AphidData I.csv using Excel.
- You should have this:

ins	wave	dur
a1	C	326.41
a1	NP	335.7
a1	C	23.92
a1	PD	5.68
a1	C	20
a1	NP	37.08
a1	C	11.95
a1	NP	223.06
a1	C	28.13
a1	PD	5.12
a1	C	17.89
a1	PD	4.93
a1	C	200.94
a1	PD	166.89
a1	C	264.57
a1	PD	5.12
a1	C	58.7
a1	NP	217.4
a1	C	5.15
a1	PD	4.96
a1	C	55.24
a1	PD	5.51
a1	C	0.06
a1	NP	4.05
a1	C	18.69
a1	NP	94.73
a1	C	111.82
a1	NP	198.96
a1	C	64.27
a1	PD	5.43
a1	C	29.23
a1	NP	49.89
a1	C	26.87
a1	NP	124.42
a1	C	22.48
a1	PD	4.97
a1	C	60.75
a1	PD	6.05
a1	C	85.88
a1	PD	5.73
a1	C	91
a1	NP	372.95
a1	C	18.03
a1	PD	4.03

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.