**Importing an Excel spreadsheet into SAS**

References will be made to the text: **Data Mining using SAS Enterprise Miner: A Case Study Approach (SAS Publishing).**

**Part 1: Importing income spreadsheet**

*(time scale in lab: 35mins)*

**Objectives:**

The aim of this practical is to give is to give you experience of importing external data into the SAS system using the File Import node. This will allow you to immediately start working on your assignment.

**By the end of this practical you will:**

* Be familiar with the purpose of the File import node in SAS Enterprise Miner.

(Note It is advisable not to import directly from Blackboard.)

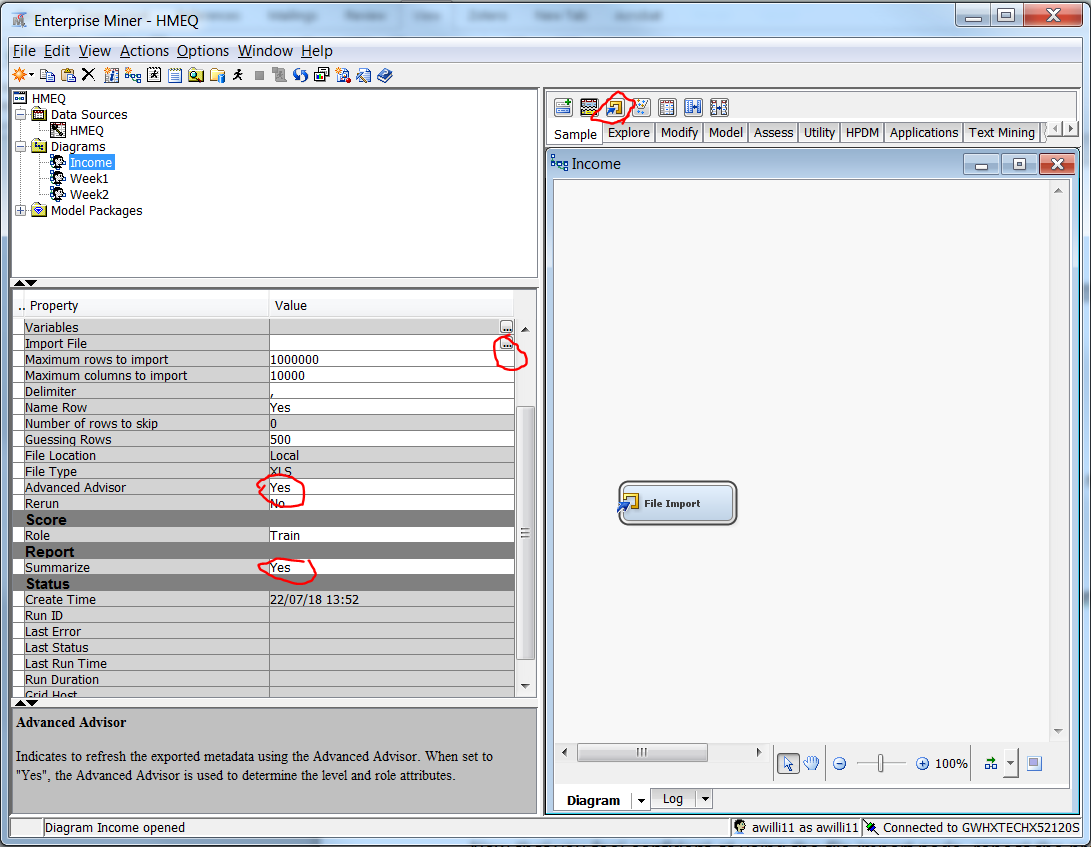
Follow each step:

1) Open the Excel spreadsheet income.xls, which can be found in blackboard.

2) Using the file selection type dropdown box, save the spreadsheet on your H: in Microsoft Excel 5.0 or 95 workbook format. SAS use to be particular about Excel formats.

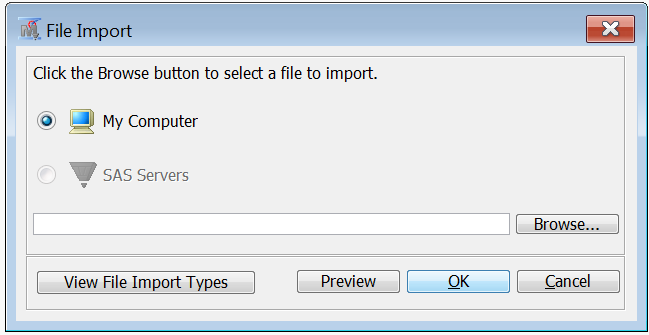
3) Close Excel.  (This is important as SAS can’t import data when the file is open as it is locked for editing).

4) In EM 14.1, create a new diagram in your chosen project, go to the “Sample” ribbon bar and select the third button from the left, file import.



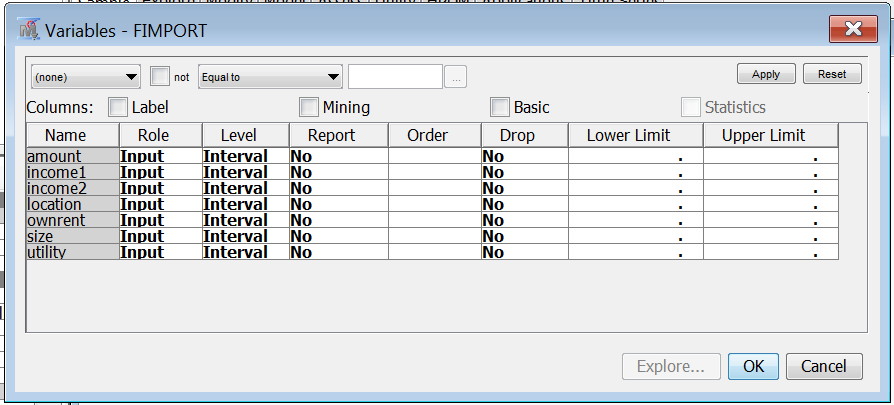
5) In the properties dialog box under Report change setting for summarise from **No to yes**. Under Advanced Adviser settings change from **No to Yes**. Under the train heading select file import ellipse and navigate to the spreadsheet you want to import.

6) A file location dialog box opens, you have the choice to select a file form the server or a local file. Select a local file, browse my computer.



Select the file income.xls from your previously save location, the select “OK”.

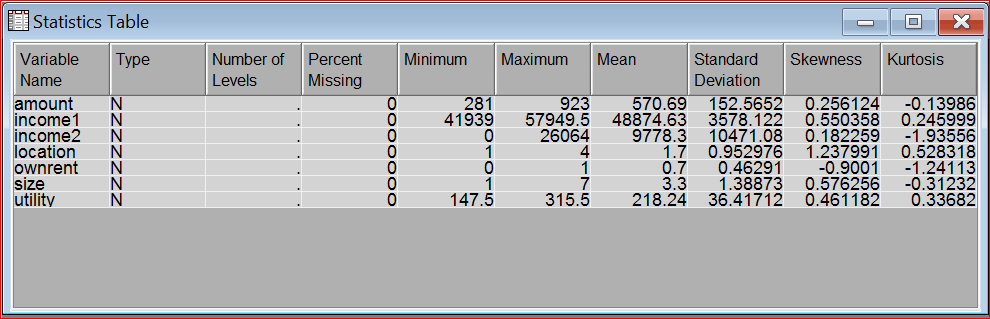
7) To browse your data, select the run command by either clicking on the man button on the main toolbar or right click on the file import node and select run. If you do not select the run command you will not be able to view your data as it has yet to be imported into the SAS format. Then go to the properties tab, under the train section select the variable ellipse, A list of variables with their roles and levels is provided.



8) With the mouse select all the variables and click on the explore button. Browse the tables and histograms and answer the following questions.

9) to view statistical data, click on the file import node right click select the running man, run the node and browse the results. Alternatively click on the black running man in the tool bar.

**Examine the Statistic Table**



Questions

How much are the utility costs for the 50 th observation? Hint you will need to use the explore variables function in the properties bar

234.5

What is the difference between the mean for the first and second income ?

48875 - 9778

What percentage of people own their own property?

70%

**Part 2**

Now that you feel confident at using the file import node, repeat the process on the **organics data set of the assignment**. ***After you have completed each lab using the HMEQ dataset you might wish to gain further practice and a head start on your assignment by repeating the same steps on the Organics data set.*** You are expected to use Enterprise Miner to answer the group activity exercises which can be found on Blackboard.