

How can I import data into SAS?

How can I import data into SAS?

If you have an ascii file containing your data, you can easily read it into a SAS dataset. SAS will read your data no matter how it looks in the file, but in the simplest case, data either be in columns or space separated. For example, it could look like this:

```
Daffy      Duck      5
Porky      Pig       3
Road       Runner    4
Hewey      .         3
```

This type of data format is called "column input format"; you tell SAS how to read each variable by specifying the columns it occupies in the file. Missing values are notated by no value in the specified column.

Your data file could also look like this:

```
Daffy Duck 5
Porky Pig 3
Road Runner 4
Hewey . 3
```

This format is called "list input format"; SAS recognizes that a space separates your variables. In list input format:

- Fields must be separated by at least one blank
- Fields must be in a specified order
- Missing values must be represented by a place holder such as a period (.)
- Character values longer than 8 characters need to be treated specially
- Data cannot be in special formats (eg dates)

If your data is in one of these two simple forms, it is very easy to write a SAS program to read it in. The SAS statement 'input' allows you to describe your data format to SAS. To read variables in list input format, simply name the variables on the input line, adding a \$ for variables which are made up of characters instead of numbers. For example, this input statement would read our sample data:

```
input first $ last $ age;
```

To read the same data in column input format, you need to specify the columns each variable occupies:

```
input first $ 1-10 last $ 12-20 age 24;
```

To put it all together into a complete SAS step (A step is a series of SAS statements), add lines resembling the following:

```
data <name>;
  infile '<pathname>';
  input <input specification>;
run;
```

Where <pathname> is the UNIX pathname to your data file, <name> is an 8 character (or less) name for your dataset internal to SAS. By default it will be a temporary dataset, and <input specification> is the input line as above.

A specific example would look like this:

```
data tdata;  
  infile '/mit/joeuser/sasdata';  
  input first $ 1-10 last $ 11-20 age 22;  
run;
```

If your data is not this simple, or you want further information, consult the SAS documentation concerning the "input" and "infile" statements.