DATA Step

Can I read in multiple raw data files in a single pass without hard-coding all the names?

None of us enjoys having to code for multiple references - even with the ease of copy / paste there's still the possibility of mis-keying.

If you have a number of files in a location, simply reference them with an appropriate wildcard on the INFILE statement. Be careful however that the files are essentially treated as continuous, so options such a FIRSTOBS= will only apply to the first file.

A folder containing the following:

```
names_1.csv
names_2.csv
names_3.csv
other_names.csv
```

Can be read in using the code:

```
data names ;
  infile '\\path\to\folder\names_*.csv dsd ;
  input first_name :$20.
        surname    :$30.
  ;
run ;
```

to select only the similarly named files, although if the layout is comparable, all files can be read with the more general wildcard:

```
data names ;
  infile '\\path\to\folder\*.csv dsd ;
  input first_name :$20.
        surname    :$30.
  ;
run ;
```

Unique solution ID: #1069
Author: Alan D Rudland
Last update: 2023-04-18 15:12