

MET MA 603:
Assignment Project Exam Help
SAS Programming and
<https://powcoder.com>
Applications
Add WeChat powcoder

*Importing Data with the Data Step
- Part I*

Entering Data with the Data Step

The **Data Step** can be used to import data from an external file.

This is the most flexible method, giving almost complete control over how the data is read from a file and put into a SAS dataset. However, it can also be the most complicated, and can be very difficult to debug without a full understanding of how the code is being executed.

Use this method when flexibility or control is needed. This is the only method that will work for some types of data structures or formats.

Infile and Input Statements

The **infile** statement tells SAS where to find the external file that is to be imported. Its function is similar to the datafile command in Proc Import.

Like the datalines/cards method, this method of importing data uses an **input** statement. Note that, unlike the datalines/cards method, the **infile** statement comes before the input statement.

The input statement is the key to this method. The example below is a simple case. With more complicated cases, the input statement will require more sophistication.

```
data city_pops1;  
infile "C:\Users\govonlu\Data\city_populations1.txt" ;  
input City $ State $ Population ;  
run;
```

List-Style Input

The first two examples use **List-Style** input. List-Style input relies on delimiters. That is, each item must be separated by a special character. The default delimiter is the space. Use the special symbol '09'x to indicate tab-delimited.

Assignment Project Exam Help
<https://powcoder.com>
Add WeChat powcoder

```
data city_pops7 ;  
infile "C:\Users\govonlu\Data\  
city_populations7.txt" delimiter='-' ;  
input City $ State $ Population ;  
run;
```

With List-Style input, SAS passes over all delimiter characters until it reaches a non-delimiter character. Non-delimiter characters are written to the dataset. Once a delimiter character is encountered SAS recognizes it has reached the end of an item of data.

The default delimiter for importing data with the Data Step is the space character, regardless of the file extension.

Fixed-Width Input

When the external data is structured in columns, **Fixed-Width** input can be used in the Input statement. After each variable, specify the starting and ending columns where the data can be found.

Assignment Project Exam Help

Note that this method does not rely on delimiters. It doesn't matter what is between the columns, as long as the data is structured in columns, there doesn't need to be any separation at all. <https://www.powcoder.com> Add WeChat powcoder

```
data city_pops4;  
infile "C:\Users\govonlu\Data\city_populations4.txt" ;  
input City $ 1-12 State $14-15 Population 17-23;  
run;
```

With fixed-width input, SAS immediately moves to the column number that is specified, and reads in everything within the range.

Length Statement

When SAS creates new variables the default length is 8. The **Length** statement can be used for specifying different lengths. The Length statement should proceed the input statement, otherwise the size will have already been determined before SAS encounters the Length statement.

```
data city_pop$8;  
infile "C:\Users\govonlu\Data\city_populations8.txt" ;  
length City $16. State $2. ;  
input City $ State $ Population ;  
run;
```

Note the distinction between the amount of characters that are read in and the length of the variable. The amount of characters read in is determined by the input statement, while the size of the variable is specified by the Length statement.

Infile Statement Options

The options below can be included in the input statement.

FirstObs specifies which row to start reading data from.

Obs specifies how many rows to read.

Sometimes a row of data doesn't have as many characters as are indicated by the input statement. By default, SAS will continue on to the next row to finish reading in the data.

Trunccover and **Missover** both instruct SAS to stop at the end of the row. With trunccover, the partial amount of data is written to the dataset, while with Missover it is dropped.

```
data city_pops3;  
  infile "C:\Users\govonlu\Data\  
city_populations3.txt"  
          firstobs=2      trunccover ;  
input City $ 1-12 State $14-15 Population 17-23 ;  
run;
```

Practice

Use the List-Style method to import scores1.txt

Use the List-Style method to import scores2.txt

Use the List-Style method to import scores3.txt

Use the Fixed-Width method to import scores3.txt

Use one of the methods to import scores4.txt

Use one of the methods to import scores5.txt

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

Readings

- Textbook sections 2.4, 2.5, 2.6, 2.14, 2.15
- “Understanding '09'x”

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder