



## Importing Data

### Importing Data Using a DATA Step

SP4R02d02.sas

1. The **allnames.csv** data set includes the five observations from **Example\_Data** and 195 more observations for a total of 200. Use a DATA step to import the **allnames.csv** data set.

```
data sp4r.all_names;
  length First_Name $ 25 Last_Name $ 25;
  infile "&path\allnames.csv" dlm=' ';
  input First_Name $ Last_Name $ age height;
run;
```

2. Check the log to verify that the data set **sp4r.all\_names** was created with 200 observations and four variables.
3. Navigate to the **SP4R** library and select **All\_names**.

	First_Name	Last_Name	age	height
1	Jordan	Bakeman	27	68
2	Bruce	Wayne	35	70
3	Walter	White	51	70
4	Henry	Hill	65	66
5	JeanClaude	VanDamme	55	69
6	Damion	Goodenow	33	81
7	Shannan	Moehrle	41	74
8	Zena	Seigfried	17	77
9	Cheree	Waldschmidt	41	77

### Importing Data Using PROC IMPORT

SP4R02d02.sas

1. Import the file **baseball.csv**. This file refers to MLB player statistics from the 1986 season.




The first row of **baseball.csv** denotes the variable names.

The variables in the data set are listed below.


<b>Name</b>	player's name
<b>Team</b>	team at the end of the season
<b>Natbat</b>	times at bat
<b>Nhits</b>	number of hits

<b>Nhome</b>	number of home runs
<b>Nruns</b>	number of runs
<b>Nrbi</b>	number of RBIs
<b>League</b>	league (American, National)
<b>Division</b>	division (East, West)
<b>Position</b>	position on field
<b>Nerror</b>	number of errors

```
proc import out=sp4r.baseball
  datafile= "&path\baseball.csv" DBMS=CSV REPLACE;
  getnames=yes;
  datarow=2;
run;
```

 Changing GETNAMES= to NO changes the variable names to **VAR1**, **VAR2**, **VAR3**, and so on.

2. Open the **baseball** data table to ensure that it was read in correctly. Then close the data table before you move to the next step.

 You *must* close the data table *before* you write to it. If you use the subsequent DATA step when the data table is open, SAS throws an error.

3. Submit the DATA step to rename the variables **nAtBat**, **nHits**, **nHome**, **nRuns**, **nRBI**, and **nError**.

```
data sp4r.baseball;
  set sp4r.baseball;
  rename nAtBat = At_Bats
         nHits = Hits
         nHome = Home_Runs
         nRuns = Runs
         nRBI = RBIs
         nError = Errors;
run;
```

Navigate to the **SP4R** library and select **Baseball**.

Partial Table View

VIEWTABLE: Sp4r.Baseball											
	Name	Team	At_Bats	Hits	Home_Runs	Runs	RBIs	League	Division	Position	Errors
1	Allanson, Andy	Cleveland	293	66	1	30	29	American	East	C	20
2	Ashby, Alan	Houston	315	81	7	24	38	National	West	C	10
3	Davis, Alan	Seattle	479	130	18	66	72	American	West	1B	14
4	Dawson, Andre	Montreal	496	141	20	65	78	National	East	RF	3
5	Galaraga, Andres	Montreal	321	87	10	39	42	National	East	1B	4
6	Griffin, Alfredo	Oakland	594	169	4	74	51	American	West	SS	25
7	Newman, Al	Montreal	185	37	1	23	8	National	East	2B	7
8	Salazar, Argenis	Kansas City	298	73	0	24	24	American	West	SS	9
9	Thomas, Andres	Atlanta	323	81	6	26	32	National	West	SS	19

## Creating a SAS Data Set from Delimited Data by Hand

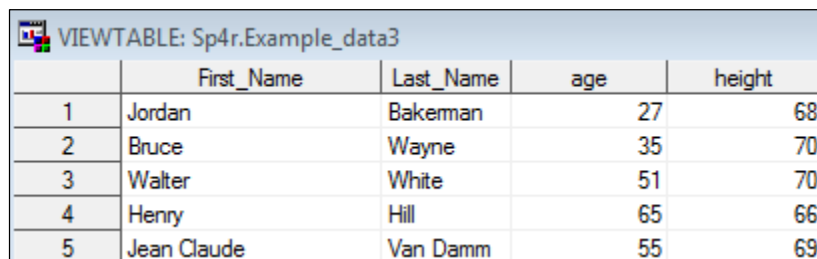
### SP4R02d02.sas

In order to create a SAS data set from delimited data by hand, simply replace the data file path with the word DATALINES in the INFILE statement.

1. Create a data set by hand from \* delimited data.

```
data sp4r.example_data3;  
  length First_Name $ 25;  
  infile datalines dlm='*';  
  input First_Name $ Last_Name $ age height;  
  datalines;  
  Jordan*Bakerman*27*68  
  Bruce*Wayne*35*70  
  Walter*White*51*70  
  Henry*Hill*65*66  
  Jean Claude*Van Damme*55*69  
;run;
```

2. Navigate to the SP4R directory to view the data table.



	First_Name	Last_Name	age	height
1	Jordan	Bakerman	27	68
2	Bruce	Wayne	35	70
3	Walter	White	51	70
4	Henry	Hill	65	66
5	Jean Claude	Van Damm	55	69