

Seeing the Compile and Execution Phases in a SAS DATA Step

Reading a SAS Data Set

The DATA step may be used to read in data from an existing SAS data set, and this is what we will be using here. The general form of such a DATA step is:

```
DATA output-SAS-data-set;  
    SET input-SAS-data-set;  
    additional SAS statements  
RUN;
```

By default, the **SET** statement reads all of the

- **observations** from the input SAS data set
- **variables** from the input SAS data set.

Compile and Execution phases

In the following slides, we will see how the DATA step first

- *compiles* (sets things up, and checks for syntax errors) and then
- *executes* (runs the code, looping through the DATA step line by line for each record that it reads in)

In our example, there will be 4 lines of code in the DATA step, and 10 records, so the DATA step really executes 40 lines of instructions. (In this example, all 4 lines are used for each record—in some cases, based on how the code is written, fewer lines may be used for some records.)

Example

We will use a small input SAS data set of airline records from Los Angeles (LAX) and Dallas Fort Worth (DFW). The values for FirstClass and Economy are the number of passengers in each group.

VIEWTABLE: Ia.Dfwlax					
	Flight	Dest	Date	FirstClass	Economy
1	439	LAX	12/11/2000	20	137
2	921	DFW	12/11/2000	20	131
3	114	LAX	12/12/2000	15	170
4	982	dfw	12/12/2000	5	85
5	439	LAX	12/13/2000	14	196
6	982	DFW	12/13/2000	15	116
7	431	LaX	12/14/2000	17	166
8	982	DFW	12/14/2000	7	88
9	114	LAX	12/15/2000	.	187
10	982	DFW	12/15/2000	14	31

Compiling the DATA Step

```
libname ia 'SAS-data-library';  
data onboard;  
→ set ia.dfwlax;  
    Total=FirstClass+Economy;  
run;
```

PDV

Flight	Date	Dest	FirstClass	Economy
	.		.	.

Compiling the DATA Step

```
libname ia 'SAS-data-library';  
data onboard;  
→ set ia.dfwlax;  
   Total=FirstClass+Economy;  
run;
```

Variable names are created in the PDV (Program Data Vector) in memory and initialized to missing values

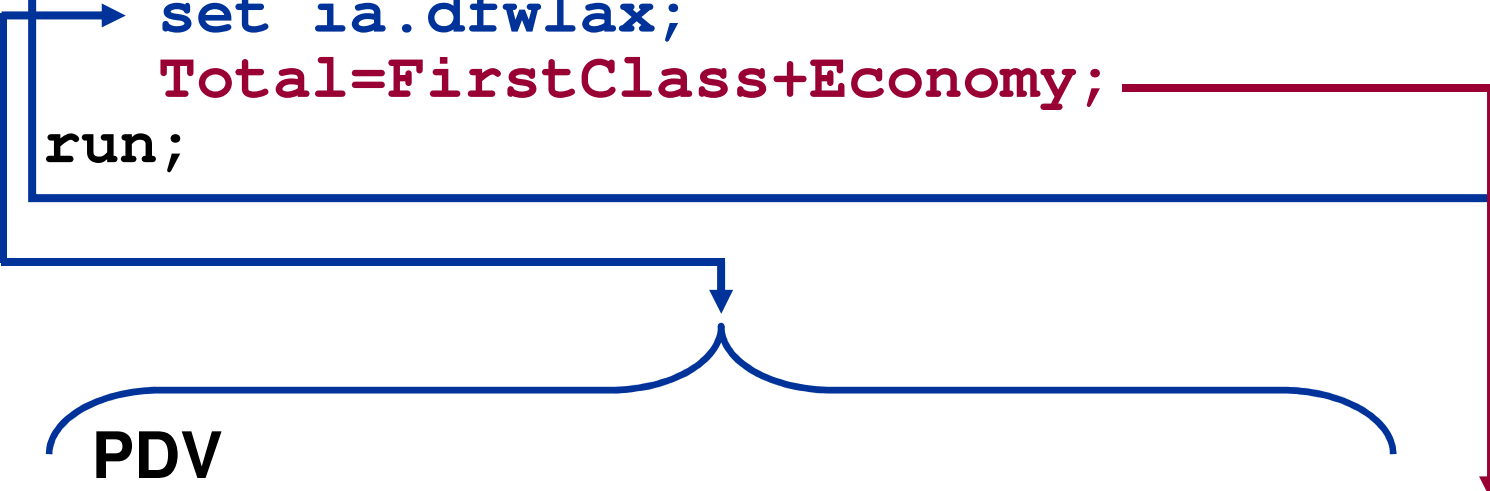
PDV

Flight	Date	Dest	FirstClass	Economy
	.		.	.

Compiling the DATA Step

```
libname ia 'SAS-data-library';  
data onboard;  
→ set ia.dfwlax;  
   Total=FirstClass+Economy;  
run;
```

PDV



Flight	Date	Dest	FirstClass	Economy	Total

Compiling the DATA Step

```
libname ia 'SAS-data-library';  
data onboard;  
→ set ia.dfwlax;  
   Total=FirstClass+Economy;  
run;
```

Each additional
variable name is
created in the PDV
and initialized to a
missing value

PDV

Flight	Date	Dest	FirstClass	Economy	Total

Executing the DATA Step

Compilation finished.
Execution starts.

ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170

```
data onboard;  
  set ia.dfwlax;  
  Total=FirstClass+Economy;  
run;
```

PDV


Flight	Date	Dest	FirstClass	Economy	Total

onboard

Flight	Date	Dest	FirstClass	Economy	Total
--------	------	------	------------	---------	-------

Executing the DATA Step

ia.dfwlax



Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170

```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
run;
```

PDV


Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	.

onboard

Flight	Date	Dest	FirstClass	Economy	Total
--------	------	------	------------	---------	-------

Executing the DATA Step

ia.dfwlax



Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170

```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
run;
```

PDV


Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

onboard

Flight	Date	Dest	FirstClass	Economy	Total
--------	------	------	------------	---------	-------


Executing the DATA Step

ia.dfwlax



Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
Automatic return	00	LAX	15	170

```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
  run;
```



PDV

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

onboard Automatic output



Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

Executing the DATA Step

ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW		
114	12/12/00	LAX		

Reinitialize Total
to missing

data onboard;

```
set ia.dfwlax;
Total=FirstClass+Economy;
run;
```

PDV

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	.

onboard

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

Executing the DATA Step

ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170



```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
run;
```

PDV

Flight	Date	Dest	FirstClass	Economy	Total
921	12/11/00	DFW	20	131	.

onboard

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

Executing the DATA Step

ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170



```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
run;
```

PDV

Flight	Date	Dest	FirstClass	Economy	Total
921	12/11/00	DFW	20	131	151

onboard

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157

Executing the DATA Step

ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
Automatic return	00	LAX	15	170

```
data onboard;
  set ia.dfwlax;
  Total=FirstClass+Economy;
  run;
```

PDV

Flight	Date	Dest	FirstClass	Economy	Total
921	12/11/00	DFW	20	131	151

onboard Automatic output

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157
921	12/11/00	DFW	20	131	151

Executing the DATA Step

And so on, for all the records in ia.dfwlax

Flight	Date	Dest	FirstClass	Economy
439	12/11/00	LAX	20	137
921	12/11/00	DFW	20	131
114	12/12/00	LAX	15	170

```
data onboard;  
  set ia.dfwlax;  
  Total=FirstClass+Economy;  
run;
```

PDV

Flight	Date	Dest	FirstClass	Economy	Total
114	12/12/00	LAX	15	170	185

onboard

Flight	Date	Dest	FirstClass	Economy	Total
439	12/11/00	LAX	20	137	157
921	12/11/00	DFW	20	131	151
114	12/12/00	LAX	15	170	185