

Paper AD16

PROC TIMEPLOT - An Old Way to Plot New Data

James W. Terry Jr.
Carolina Population Center
University of North Carolina at Chapel Hill

PROC TIMEPLOT was almost a forgotten SAS® procedure at The University of North Carolina's Carolina Population Center until it was recently used to make critical decisions in the analysis of accelerometer data. Accelerometers, a small device usually worn on the body, were used at UNC-CH to measure the physical activity of postpartum women minute-by-minute over a 7-day period. The accelerometers produced over 2 million observations that had wear-time and compliance issues that needed to be analyzed.

The distinctive features of PROC TIMEPLOT were perfect for bringing these issues into focus and it was so easy to use. This presentation will use this and other real examples to demonstrate the ease of use and features of PROC TIMEPLOT and will demonstrate the plotting of time across the horizontal axis and the vertical axis.

An Introduction to PROC TIMEPLOT

Unlike PROC PLOT which plots data on the X-axis and the Y-axis, PROC TIMEPLOT only plots across the horizontal or X-axis. PROC TIMEPLOT plots one or more variables on a horizontal axis that represents the values of the variable(s) you are examining. And a “plot” is one or more horizontal plots or lines that can extend into multiple pages.

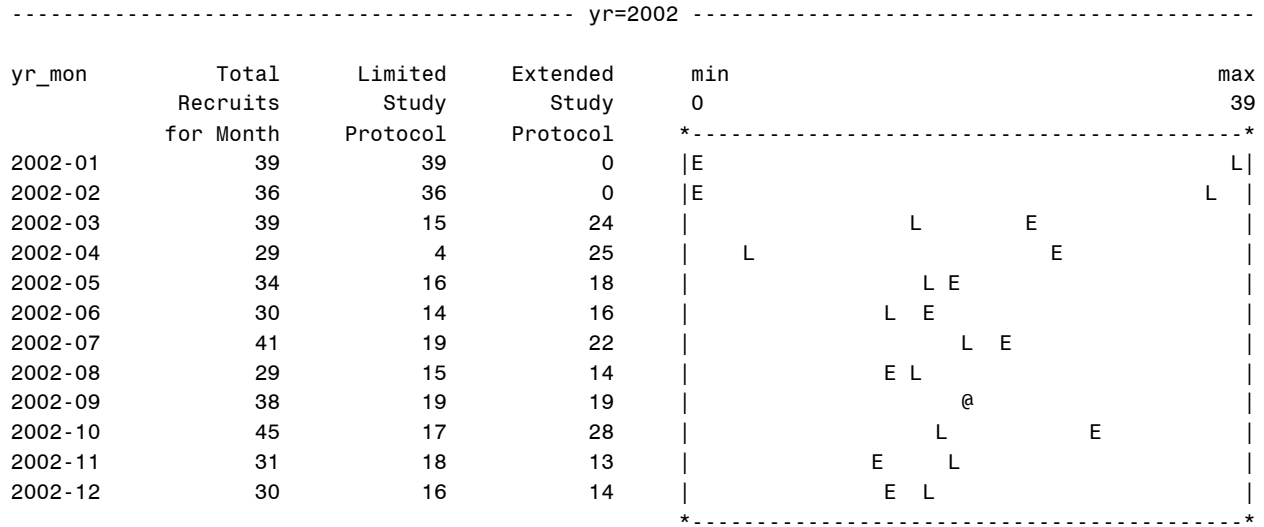
A vertical axis or Y-axis can be created when sequential observations are linked and presented in order as seen in Example 1 where the vertical axis represents a month-by-month representation of the data. However, the vertical axis can also represent separate entities with data about each entity being presented on the horizontal axis as seen in Example 3.

PROC TIMEPLOT – Example 1

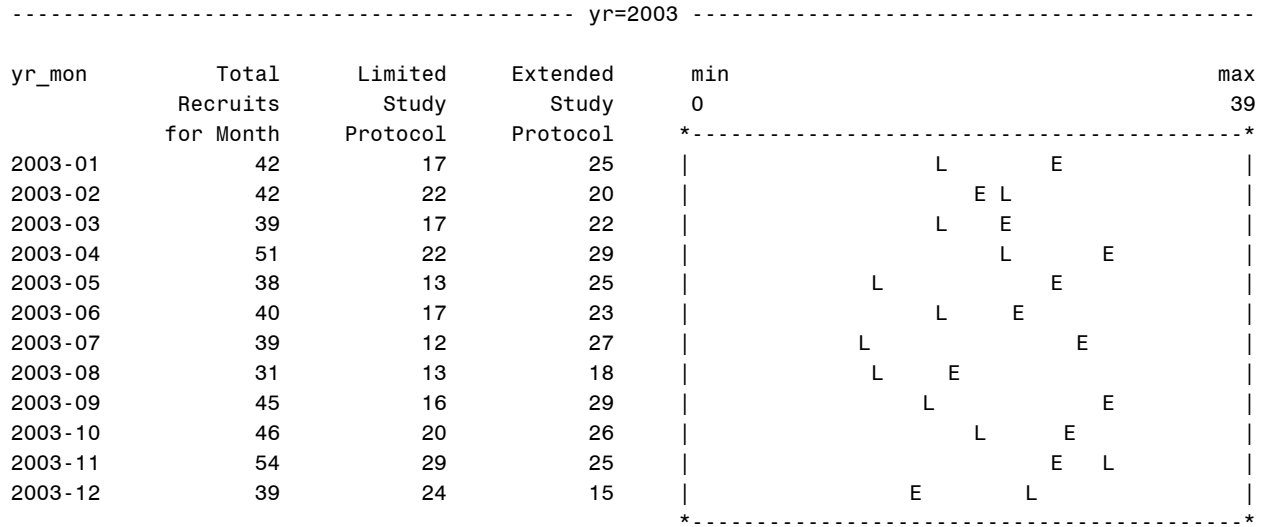
The following example is a simple plot of the number of women recruited for a pregnancy study by research protocol by year over a two year period.

```
proc timeplot data=plot_recruits uniform;  
  plot protocol_1='L' protocol_2='E' / overlay;  
    id yr_mon total;  
    by yr;  
      where yr in (2002,2003);  
title1 Pregnancy Study Recruitment by Month by Study Protocol;  
title2 2002-2003;  
run;
```

Pregnancy Study Recruitment by Month by Study Protocol
2002-2004



Pregnancy Study Recruitment by Month by Study Protocol
2002-2004



This example plots a dataset, DATA=PLOT_RECRUITS, that has one observation for every month of the pregnancy study. The horizontal axis represents the number of recruits by protocol and the vertical axis represents each month of the pregnancy study. The UNIFORM option was used to create a uniform scale across all BY groups

The PLOT statement plots two variables, *protocol_1* and *protocol_2*. Since both variable names begin with the same letter, the '=' option is used to give each one a unique plotting symbol. The OVERLAY option causes both variables to be plotted on the same horizontal axis. The ID statement identifies the variables you want printed in addition to the variables that are plotted. By

default, the values for the plotted variables are printed, so only the *yr_mon* and *total* variables are specified. And the BY statement is used to create plots for each BY group with each plot appearing as a separate page.

PROC TIMEPLOT – Example 2

The following example plots over the course of a 24-hour day each time the accelerometer was worn. It plots the date worn, the beginning and ending times and the average level of physical activity recorded during that time. The horizontal axis represents the time of day based on a 24-hour clock and the vertical axis is in order of the date and time the accelerometer was worn.

```
proc timeplot data=plot;
  plot (_start _end) = '*' / overlay hiloc ref='12:00'T;
  id personid date_stamp avg_act;
  format _start _end time5. date_stamp mmddyy10.;
  title1 3-Month PostPartum Actigraph Active (Wear-Time) Segments;
```

3-Month PostPartum Accelerometer Active (Wear-Time) Segments						
SAS PROC TIMEPLOT						
personid	date_stamp	avg_act	_start	_end	min	max
					3:46	23:09

5582	08/10/2005	426.60	9:10	23:00		*-----*
5582	08/11/2005	379.30	9:18	22:44		*-----*
5582	08/12/2005	265.80	8:30	22:48		*-----*
5582	08/13/2005	262.60	7:40	22:43		*-----*
5582	08/14/2005	257.00	7:54	22:23		*-----*
5582	08/15/2005	259.60	8:06	23:09		*-----*
5582	08/16/2005	2.90	3:46	4:56	*-*	
5582	08/16/2005	264.80	7:29	16:33		*-----*
5582	08/16/2005	158.40	17:39	22:37		*-----*

This example plots a dataset, DATA=PLOT, that has one observation and five variables for every time the women wore the accelerometer. The horizontal axis represents time over the course of a 24-hour day and the vertical axis represents each time the accelerometer was worn.

The PLOT statement plots two variables, the start (*_start*) and ending (*_end*) times the accelerometer was worn and the = '*' causes the asterisk to be used as the plotting symbol. The OVERLAY option causes all variables to be plotted on the same horizontal axis. The HILOC option causes the right-most and left-most plotting symbols to be connected by hyphens. The REF='12:00'T option draws a reference line at 12 o'clock noon.

The ID statement identifies the variables you want printed in addition to the variables that are plotted. By default, the values for the plotted variables are printed, so only the *personid*, *date_stamp* and *avg_act* variables are specified.

When the accelerometer is worn multiple times per day, as on 8/16/2005, each observation is plotted on a separate horizontal axis. This creates a distinctive step-down pattern which becomes very noticeable when this occurs on multiple days for an individual woman.

PROC TIMEPLOT became much more useful in our research when we added BY group processing, see below:

```
proc timeplot data=plot uniform;
  plot (_start _end) = '*' / overlay hiloc ref='12:00'T;
  id date_stamp avg_act;
  by personid;
  format _start _end time5. date_stamp mmddyy10.;
title1 3-Month PostPartum Actigraph Active (Wear-Time) Segments;
```

The addition of the BY statement produced a one-page plot, like what you see above, for each of the 100+ subjects and these could be viewed page-by-page on a PC to get a complete view of the variety of accelerometer wear patterns. And since “wear time” is determined based on different algorithms, it allowed the researchers to graphically see across the whole group of subjects how the different algorithms changed wear-time patterns. The UNIFORM option was added to create a uniform scale across all BY groups and to create a consistent reference line.

PROC TIMEPLOT – Example 3

For a pregnancy study, the following example plots the recruitment, clinic visits and delivery pattern for study individuals across their pregnancy. The four plotted variables represent the gestational age that each event happens.

```
proc timeplot data=plot1;
  plot recruit a_visit b_visit delivery / overlay hiloc axis = 0 to 44
  ovpchar='R';
  id patid;
  where pin3plus = 1;
title Pregnancy Study Timing Patterns;
run;
```

Pregnancy Study Timing Patterns						Monthly Report	
Study ID	recruit	a_visit	b_visit	delivery	min		max
					0		44

84139	15	15	25	34		R-----b-----d	
83296	11	19	27	34		r-----a-----b-----d	
82762	15	16	24	35		ra-----b-----d	
82984	11	16	.	36		r---a-----d	
83766	18	18	26	36		R-----b-----d	
81526	10	15	27	36		r---a-----b-----d	
83301	8	8	.	37		R-----d	
83494	7	16	24	37		r-----a-----b-----d	
83161	12	15	25	37		r--a-----b-----d	
82832	11	15	25	38		r---a-----b-----d	
83779	13	17	26	38		r---a-----b-----d	
81082	17	17	26	38		R-----b-----d	
83463	11	17	28	38		r-----a-----b-----d	
83160	17	17	28	38		R-----b-----d	
84690	14	18	28	38		r---a-----b-----d	

84295	12	17	25	39		r----	a-----	b-----	d	
81924	12	16	26	39		r----	a-----	b-----	d	
81720	16	16	26	39		R-----	b-----	d		
82738	15	18	26	39		r--a-----	b-----	d		
81414	10	15	28	39		r----	a-----	b-----	d	
82697	11	19	28	39		r-----	a-----	b-----	d	
84659	14	20	28	39		r-----	a-----	b-----	d	
82172	12	15	25	40		r--a-----	b-----	d		
85175	17	19	25	40			r-a-----	b-----	d	
83887	10	16	27	40		r----	a-----	b-----	d	
84411	17	17	27	40			R-----	b-----	d	
82018	15	15	28	40			R-----	b-----	d	
85056	15	18	28	40		r--a-----	b-----	d		
85487	16	19	29	40		r--a-----	b-----	d		
83412	18	19	29	40			ra-----	b-----	d	
83527	18	19	29	40			ra-----	b-----	d	
84235	14	18	.	41		r----	a-----	b-----	d	
84938	19	19	25	41			R-----	b-----	d	
85278	17	19	26	41			r-a-----	b-----	d	
84475	13	18	27	41		r----	a-----	b-----	d	
84262	16	19	26	42		r--a-----	b-----	d		

This example plots a dataset, DATA=PLOT1, that has one observation and six variables for every woman in the study. The horizontal axis represents the weeks spent in the pregnancy study from recruitment to delivery.

The PLOT statement plots four variables, the week of recruitment (*recruit*), the first clinic visit (*a_visit*), the second clinic visit, if completed, (*b_visit*) and delivery (*delivery*). Since no plotting symbol was specified, by default, PROC TIMEPLOT uses the first letter of the variable. Again, the OVERLAY option causes all variables to be plotted on the same horizontal axis and the HILOC option causes the right-most and left-most plotting symbols to be connected by hyphens. The AXIS = 0 to 44 option specifies a horizontal axis from zero to 44 to cover a pregnancy up to 44 weeks. The OVPCCHAR='R' option is used to highlight when the week of recruitment and the week of the first clinic visit occur in the same week.

Again the ID statement identifies the variables you want printed in addition to the variables that are plotted. By default, the values for the plotted variables are printed, so only the *patid* variable is specified.

This example is a monthly report for the researchers to review the occurrence of preterm and post-term births, so the observations were sorted on the *delivery* variable which represents the week of pregnancy in which delivery occurred. The report also allows a review of the weeks the clinic visits are occurring to ensure that the study protocol is being followed. There is the flexibility to sort the data on other variables depending on what variable is of most interest.

In Summary

PROC TIMEPLOT from SAS is an excellent tool for presenting and analyzing time series data. I am sure it was based on line printer technology which seems obsolete these days, but its ease of use and flexibility still make it relevant today. With a little forethought and the ability of PROC TIMEPLOT to present time series data on either a horizontal or vertical axis, it can still produce very meaningful plots. And with the ability to list additional variables along side the plots, you can even create what I would call a “plot-enhanced” report.

This paper has tried to cover the major statements and options, but by no means has it covered all of them. If you are working with time series data and need an easy way to view or represent time series data, whether it be by time-of-day, day-of-the-month, month-of-the-year, or some other representation, I encourage you to look at PROC TIMEPLOT.

Trademark Citation

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates UDA registration.

Other brand and product names are trademarks of their respective companies.

Contact Information

Contact the author at:

Jim Terry

Email: jim_terry@unc.edu