ABSTRACT
A project management tool helps manage time and resource for the duration of a project. This paper explains how to use PROC CPM to manage studies.

INTRODUCTION
Project management and planning are necessary things to consider before a project begins. The work flow will be smooth when an effort is put forth to manage resources and time-lines and to plan the allocation and management of resources and the management time-lines proactively. The SAS CPM procedure can be used for planning, controlling, and monitoring a project. A typical project consists of several activities that may have precedence and time constraints. Some of those activities may already be in progress and some of them may follow different work schedules. All of the activities may compete for scarce resources. PROC CPM enables you to schedule activities subject to all the constraints.

SYNTAX
The following statements are used in PROC CPM:

PROC CPM options;
  ACTIVITY variable;
  ACTUAL / actual options;
  ALIGNDATE variable;
  ALIGNTYPE variable;
  BASELINE / baseline options;
  CALID variable;
  DURATION / duration options;
  HEADNODE variable;
  HOLIDAY variable / holiday options;
  ID variables;
  PROJECT variable / project options;
  RESOURCE variables / resource options;
  SUCCESSOR variables / lag options;
  TAILNODE variable;

USAGE
In clinical studies, as is the case for most complicated projects, certain discrete tasks depend on other tasks and certain activities cannot start until other activities have finished. For instance, the project team will not be able to review the “Statistical Analysis Plan” (some companies call “Reporting Analysis Plan”) until the Statistician releases it for first draft review. Such constraints among the activities (namely, activity B can start after activity A has finished) are referred to as precedence constraints. Given the precedence constraints and estimated durations of the activities, you can use the critical path method to determine the shortest completion time for the project.

Now we are going to discuss “Activity-On-Arc (AOA) or Activity-On-Edge (AOE)” method. The diagram (figure 1) below shows the AOA representation for the project. This method of representing a project is known also as the arrow diagramming method (ADM). For projects represented in the AOA format, PROC CPM requires the use of the following statements:
PROC CPM options;
  TAILNODE variable;
  HEADNODE variable;
  DURATION variable;

Figure 1

The data set should contain the minimum amount of information required by PROC CPM for an activity network in AOA format, namely, the TAILNODE and HEADNODE variables, which indicate the direction of each arc in the network and the DURATION variable which gives the length of each task. In addition, the data set should contain a variable identifying the name of the task associated with each arc. Note that PROC CPM treats each observation in the data set as a new task, thus enabling you to specify multiple arcs between a pair of nodes.

In addition to the variables specified TAILNODE, HEADNODE and DURATION statements, the output data set (PROC CPM Output dataset) will contain the following new variables.

**E_START** specifies the earliest time an activity can begin, subject to any time constraints and the completion time of the preceding activity.

**E_FINISH** specifies the earliest time an activity can be finished, assuming it starts at E_START.

**L_START** specifies the latest time the activity can be started. This is computed from the activity's latest finish time.

**L_FINISH** specifies the latest completion time of the activity. This is the minimum of the minimum late start time of all successor activities and any upper bound placed on the finish time of the activity by the alignment constraints.

**T_FLOAT** specifies the amount of flexibility in the starting of a specific activity without delaying the project: T_FLOAT = L_START - E_START = L_FINISH - E_FINISH.

**F_FLOAT** specifies the difference between the early finish time of the activity and the early start time of the activity's immediate successors.
Let's discuss CPM procedure with an example:

```sas
options missing = " ";

/* Activity-on-Arc representation - Tasks information dataset */
data taskinfo;
  format task $14.;
  input task & days tail head;
  datalines;
  SAP CREATION 15 1 2
  SAP QC 10 2 3
  RELEASE SAP 1 3 4
  REVIEW SAP 1 4 5
  GET RAWDATA 3 6 7
  EDITCHECK PRGS 3 7 8
  WRITE SPECS 3 7 9
  RELEASE SPECS 1 9 10
  CREATE ANALDS 15 10 11
  QC ANALDS 10 11 12
  CREATE DISPS 15 12 13
  QC DISPS 10 13 14
  FINAL REVIEW 10 14 15
  SUBMISSION 10 15 16
;
run;

/* Tasks description dataset */
data details;
  format task $14. dept $6. descrpt $50.;
  input task & dept & descrpt &;
  label dept = "Department"
   descrpt = "Activity Description";
  datalines;
  SAP CREATION Stat SAP Creation process
  SAP QC Stat QC the completed SAP
  RELEASE SAP Stat Release the finalised SAP
  REVIEW SAP Stat Review the finalised SAP
  GET RAWDATA PROG Receive the raw datasets from Data Management
  EDITCHECK PRGS PROG Write EDIT Check programs to make sure raw datasets were error free
  WRITE SPECS PROG Write the Analysis Datasets specification
  RELEASE SPECS PROD Release the final specification
  CREATE ANALDS PROG Write programs to Create Analysis Datasets
  QC ANALDS PROG Validate the completed Analysis Datasets
  CREATE DISPS PROG Create Displays
  QC DISPS PROG QC the Displays
  FINAL REVIEW PROG Perform the final review
  SUBMISSION MGMT Submit the project
;
run;
```
/* Holidays information dataset */
data holidays;
   format holiday holifin date9.;
   input holiday date9. holifin date9. holidur;
datalines;
01JAN200901JAN2009 1
31MAY200931MAY2009 1
05JUL200905JUL2009 1
06SEP200906SEP2009 1
25NOV200926NOV2009 2
24DEC200901JAN2010 9
31MAY201031MAY2010 1
05JUL201005JUL2010 1
06SEP201006SEP2010 1
25NOV201026NOV2010 2
24DEC201031DEC2010 6
;
run;

proc sort data = taskinfo;
   by task;
run;

proc sort data = details;
   by task;
run;

data resource;
   merge taskinfo(in = a) details;
      by task;
   if a;
run;

/* The project is scheduled using PROC CPM - The network information is conveyed using the TAILNODE and HEADNODE statements. The ID statement is used to transfer project information to the output data set */
proc cpm data = resource holidata=holidays date='14dec09'd out=outtime interval=weekday;* fbdate='30sep10'd;
   tailnode tail;
   headnode head;
   duration days;
   activity task;
   holiday  holiday / holifin=(holifin);
   id dept descrpt;
run;

proc sort data = outtime;
   by tail head;
run;

options ls=90;
title 'Activity-On-Arc Format - Project Schedule';
proc print data = outtime;
   id descrpt;
   var dept e_ : l_ : t_float f_float;
run;
The above **PROC CPM** will process the given datasets per requirement as follows,

- “holidata=holidays” this option will refer the ‘holidays’ dataset.
- “interval=weekday” this option will consider only weekdays in the calculation.
- “date='14dec09'd” refers to the start date of a study.
- “tailnode tail” and “headnode head” will connect the network arcs.
- “duration days” refers to the total duration between two given date values.
- “activity task” and “id dept descipt” refers to the task information.
- “holiday holiday / holifin=(holifin)” refers to the holidays variable.
- “fbdate='30sep10'd” refers to the pre-defined study finish date (or to specify a deadline on the study). If we use this option then E_START and E_FINISH will retain the same output values but the other output variables values will change accordingly.

**OUTPUT**

**Activity-On-Arc Format - Project Schedule**

<table>
<thead>
<tr>
<th>descrpt</th>
<th>dept</th>
<th>E_START</th>
<th>E_FINISH</th>
<th>L_START</th>
<th>L_FINISH</th>
<th>T_FLOAT</th>
<th>F_FLOAT</th>
<th>descrpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP Creation process</td>
<td>Stat</td>
<td>14DEC09</td>
<td>12JAN10</td>
<td>03MAR10</td>
<td>23MAR10</td>
<td>50</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>QC the completed SAP</td>
<td>Stat</td>
<td>13JAN10</td>
<td>26JAN10</td>
<td>24MAR10</td>
<td>06APR10</td>
<td>50</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Release the finalised SAP</td>
<td>Stat</td>
<td>27JAN10</td>
<td>27JAN10</td>
<td>07APR10</td>
<td>07APR10</td>
<td>50</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Review the finalised SAP</td>
<td>Stat</td>
<td>28JAN10</td>
<td>28JAN10</td>
<td>08APR10</td>
<td>08APR10</td>
<td>50</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Receive the raw datasets</td>
<td>PROG</td>
<td>14DEC09</td>
<td>16DEC09</td>
<td>14DEC09</td>
<td>16DEC09</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>from Data Management</td>
<td>PROG</td>
<td>17DEC09</td>
<td>21DEC09</td>
<td>06APR10</td>
<td>08APR10</td>
<td>71</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Write EDIT Check programs</td>
<td>PROG</td>
<td>17DEC09</td>
<td>21DEC09</td>
<td>17DEC09</td>
<td>21DEC09</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>to make sure raw dataset</td>
<td>PROG</td>
<td>17DEC09</td>
<td>21DEC09</td>
<td>17DEC09</td>
<td>21DEC09</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Write the Analysis</td>
<td>PROD</td>
<td>22DEC09</td>
<td>22DEC09</td>
<td>22DEC09</td>
<td>22DEC09</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Datasets specification</td>
<td>PROG</td>
<td>23DEC09</td>
<td>21JAN10</td>
<td>23DEC09</td>
<td>21JAN10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Write programs to Create</td>
<td>PROG</td>
<td>22JAN10</td>
<td>04FEB10</td>
<td>22JAN10</td>
<td>04FEB10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Analysis Datasets</td>
<td>PROG</td>
<td>05FEB10</td>
<td>25FEB10</td>
<td>05FEB10</td>
<td>25FEB10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Validate the completed</td>
<td>PROG</td>
<td>26FEB10</td>
<td>11MAR10</td>
<td>26FEB10</td>
<td>11MAR10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Analysis Datasets</td>
<td>PROG</td>
<td>12MAR10</td>
<td>25MAR10</td>
<td>12MAR10</td>
<td>25MAR10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Create Displays</td>
<td>PROG</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>QC the Displays</td>
<td>MGMT</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Perform the final review</td>
<td>PROG</td>
<td>22JAN10</td>
<td>04FEB10</td>
<td>22JAN10</td>
<td>04FEB10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Submit the project</td>
<td>MGMT</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>26MAR10</td>
<td>08APR10</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**GANTT procedure** produces a Gantt chart of the schedule obtained from PROC CPM.

It shows the early and late start schedules as well as the precedence relationships between the activities. The precedence information is conveyed to PROC GANTT through the TAILNODE= and HEADNODE= options.

```plaintext
/* specify the device on which you want the chart printed */
goptions vpos=50 hpos=80 border;
title f=swiss 'Precedence Gantt Chart';
title2 f=swiss 'Early and Late Start Schedule';

proc gantt graphics data = outtime;
   chart / compress tailnode=tail headnode=head
      font=swiss height=1.5 nojobnum skip=2
      cprec=cyan cmile=magenta
      caxis=black cframe=ligr
      dur=days increment=7;
      id descrpt;
run;
```
The code below will create an output with Holidays option as well,

```
title 'Scheduling Around Holidays';
title2 'Project Schedule';
goptions vpos=50 hpos=80 border;
goptions ftext=swiss;
proc gantt graphics data = outtime holidata=holidays;
    chart / compress
        font=swiss height=1.5 nojobnum skip=2
        dur=days increment=7
        holiday=(holiday) holifin=(holifin)
        cframe=ligr;
    id descrit;
run;
```
**CALANDAR PROCEDURE**

This example shows how you can use the output from CPM to display calendars containing the critical path schedule and the early start schedule. The following program invokes PROC CALENDAR to produce calendar with displays only the critical activities in the project.

```plaintext
proc sort data = outtime out=crit;
   where t_float=0;
   by e_start;
run;

title 'Printing the Schedule on a Calendar';
title2 'Critical Activities in December';

/* print the critical act. calendar */
proc calendar schedule
   data=crit;
   id e_start;
   var task;
   dur days;
run;
```

The output will consists of five pages starting from December 2009 to April 2010 until the end of the project.
CONCLUSION
By using the available smart SAS procedures, all kind of tasks can be done in an efficient and effective manner. Like wise CPM procedure helps you manage a project easily and quickly, also you can track the date at any time. I hope this paper will give you a better idea of scheduling study time using PROC CPM.

REFERENCES
http://support.sas.com

ACKNOWLEDGEMENTS
I would like to thank my managers at Quartesian LLC for given me this wonderful opportunity to explore the ideas.

CONTACT INFORMATION
Your comments and questions are valued and encouraged, contact the author at:

Venkateshwaran Lajapathirajan
Quartesian LLC
5 Independence Way, Suite 300
Princeton, NJ 08540.
Phone: 551-208-8815
E-mail: venkateshwaran.lajapathirajan@quartesian.com
Website: www.quartesian.com

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 USA registration.

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