Paper 4167-2020

Using Lag and Other Function in SAS ® to Create Final Datasets Temperature Study

Abbas S. Tavakoli, DrPH, MPH, ME, University of South Carolina Thomas Best, BS Student, University of South Carolina Robin B. Dail, PhD, RN, FAAN, University of South Carolina

ABSTRACT

Temperature control is important for the health of premature babies. One way to monitor body temperature in premature infants is through continuous measurement of central (abdominal) and peripheral (foot) skin temperature. Fourteen premature infants, born ≤ 32 weeks gestational age and birthweights < 1500 grams were enrolled for study after Institutional Review Board approval and parental consent for participation in this study. Each infant had one skin temperature probe (thermistor) attached to their abdomen and one skin temperature probe to the sole of one foot. The data downloaded from incubator had several problems such as extra rows, missing value for several minutes, and so forth. The data for each days included about 86,400 rows. The final datasets for each infant included averaged data for every minute of all variables for 28 days. Lag and several functions in SAS were used to prepare data for analyses. Several programs were used to delete unnecessary rows, create minutes from the time each infant was born to the time infants completed data collection, to replace missing minutes, combine and merge different datasets. Several procedures in SAS were used to analyze data such as Means, Freq, Univariate, Gplot, and Sqplot. All data analyses was performed using SAS/STAT® statistical software, version 9.4.

Keywords: SAS, body temperature, infant

INTRODUCTION

Temperature control is important for the health of premature babies. One way to monitor body temperature in premature infants is through continuous measurement of central (abdominal) and peripheral (foot) skin temperature. 1 Abnormal patterns of body temperature with increased central-peripheral temperature difference (CPtd) or a CPtd that is negative (foot higher then abdominal temperature) have been associated with the onset of infection and stressful events in premature infants. 2 Engineers at G. E. Healthcare developed a research software allowing for display of central and peripheral body temperature and the CPtd on their Giraffe Omnibed Carestation™ hybrid incubators and funded the study of 14 infants using these special incubators in a level four neonatal intensive care unit in the southeastern United States.

PROPUSE

This paper describes the use of lag and other function in SAS to combine and merge each dataset to create 14 final dataset for statistical analysis using a within subject, case study design.

METHODOLOGY

Fourteen premature infants, born at 32 weeks gestational age or less and having birthweights of less than 1500 grams were enrolled for study after Institutional Review Board approval and parental consent for participation in this study. Physiological data were measured and stored for the first 28 days of life using a laptop attached to the research incubator in a RS232 port. Each infant had one skin temperature probe (thermistor) attached to their abdomen and one skin temperature probe to the sole of one foot. Infants were maintained in incubator servo control (36.5□-37.0□C) from the feedback of abdominal temperature. Temperatures were measured every minute for 14 days. A laptop computer on the incubator shelf recorded and stored all data. These data included both skin temperatures, inside incubator temperature, and humidity levels. In the data sets we have date, time, abdominal temperature (ABD), foot temperature (FT), inside incubator air temperature(EXT), air temperature(AS), Abdominal set temperature(ISC), Incubator Power output(IP), Humidity set point(HSP), Humidity measure(HUM), and Heat Respiratory. Minutes since birth (MSB) were calculated from birthdate and time and used to anchor all data longitudinally. Data were cleaned to eliminate low and high temperatures due to temperature probes being off the infant and data were coded for missing data. Each variable had approximately 40,320 measures for each infant. The data for each days included about 86,400 rows. The final datasets for each infant included average of all measurement by minutes for all 28 days of data. The data were not ready for data analysis. Thus, lag and several function in SAS used to prepare data for data analysis. Several programs were used to delete unnecessary rows, create minutes from the time the infant was born to the time the infant completed data collection, replace missing minutes, combine and merge different datasets. Several procedures in SAS were used to analyze data such as Means, Freg, Univariate, Gplot, and Sgplot. All data analyses were performed using SAS/STAT® statistical software, version 9.4.3

DATA STEPS

Table 1 shows an example of data converted from text file to excel file. We use infant one and two as an example to present the steps that we went through to create the final data set. For this example, the infant was born on 6/6/2018 9:34 am and data collection was complete on 6/26/2018 at 9:41. Data included all measurements for every second in each row. In addition, there is extra row for each data point.

Table1 Example of data in Excel

Date	Time	MSB	ABD	FT	EXT	AS	ISC	IP	HSP	HUM	Heat Res
6/6/2018	9:34	0									
6/6/2018	12:27:29.147]		31.2	34.3	29.1	36	37	100	0	26	3554
6/6/2018	12:27:29.158]		0								
6/6/2018	12:27:29.158]		31.3	34.3	29.1	36	37	100	0	27	3560
6/6/2018	12:27:29.158]		0								
6/6/2018	12:27:29.158]		31.4	34.3	29.2	36	37	100	0	27	3560
6/6/2018	12:27:29.158]		0								
6/6/2018	12:27:29.158]		31.5	34.3	29.2	36	37	100	0	27	3560
6/6/2018	12:27:29.158]		0								
6/6/2018	12:27:29.158]		31.6	34.3	29.3	36	37	100	0	27	3548
6/6/2018	12:27:29.158]		0								
6/6/2018	12:27:29.158]		31.7	34.3	29.4	36	37	100	0	27	3536

6/6/2018	12:27:29.158]	0				

Table 2 indicates the note for infant one. For example, infant one was born on 6/6/18 at 9:34am. Data collection instruments were attached to the infant in the incubator on 6/6/2018 at 12:27 and data collection was completed on 6/26/2018 at 9:41am. There are several times the incubator was not collecting the data from 6/8/18 at 17:45 to 6/12/18 at 10:44, 6/23/18 at 14:19 to 6/23/18 to 14:14, and 6/24/18 at 6:26 to 6/25/18 at 9:56am. Thus, when we created the data set we replaced those missing minutes.

Table 2. Data note for infant 1

Note
DOB 6/6/18 TOB 0934
ECU001 6-6-2018_draft 2_RD Wed 6/6/2018 1227 -6/7/2018 1420
ECU001 6-7-2018_draft 2_RD Thurs 6/7/2018 1427-6/8/2018 1107
ECU001 6-8-2018_draft 2_RD Fri 6/8/2019 1113-6/8/2018 1744
Missing data from 6/8/18-6/12/18 am.
ECU001 6-12-2018_draft 2_RD 6/12/2018 1044-6/12/2018 2312
ECU001 6-13-2018_draft 2_RD 6/13/2018 0013-6/13/2018 1226
ECU001 6-13-2018-2_draft 2_RD 6/13/2018 1245-6/14/2018 1027
ECU001 6-14-2018_draft 2_RD 6/14/2018 1228-6/15/2018 0924
ECU001 6-15-2018_draft 2_RD 6/15/2018 0931-6/16/2018 0651
ECU001 6-16-2018_draft 2_RD 6/16/2018 0655-06/18/2018 0956
ECU001 6-18-2018_draft 2_RD 6/18/2018 0959-06/19/2018 0942
ECU001 6-19-2018_draft 2_RD 6/19/2018 0945-6/20/2018 0959
ECU001 6-20-2018_draft 2_RD 6/20/2018 1002-6/21/2018 0831
ECU001 6-21-2018_draft 2_RD 6/21/2018 0835-6/22/2018 0940
ECU001 6-22-2018_draft 2_RD 6/22/2018 0945-6/24/2018 0625
Missing 6/23/2018 14:10-6/23/2018 14:14
6/23/2018 1415-6/24/2018 0625
Missing data 6/24/18 0626-6/25/18 0956
ECU001 6-22-3018_2_draft 2_RD 6/25/18 0957-6/25/2018 1004
ECU001 6-25-3018_2_draft 2_RD 6/25/18 1009-6/26/2018 941

Table 3 shows the SAS program to create the master file including id for each infant, date, time (hours and minutes), and Minutes since birth (MSB) from the time the infant was born to the time the infant completed data collection.

Table 3. SAS Program to create master file

SAS Program libname temp 'd:\ABBAST\\temp\test\'; *** add times from date of birth to time recoded ***; timest='06jun2018:09:34:00'dt; timeed='26jun2018:09:41:00'dt; msb=.; interval=intck('minute',timest,timeed); do i = 1 to interval by 1; timest=intnx('minute',timest,1,'same'); date = datepart(timest); timea = timepart(timest); output; end; drop timeed timest i interval; format timea time8. date mmddyy10.; run; data one; set one; $id = _n_;$ timeb=timea; tmp=put(timea,time5.); format timea time8. timeb time5.; drop tmp; run; data two; set one; msb=id -1; run; data temp.ECUM(keep= id DATE TIMEb MSB) ; retain id DATE TIMEb MSB; set two; format date mmddyy10. ; run;

Table 4 shows an example of first 20 observations for the master file, which creates MSB from the time the infant was born to the time the infant completed data collection. The data includes id, date, time (hours and minutes), and MSB.

Table 4 Sample of data master for first 20 observation

Obs	id	date	timeb	Msb
1	1	06/06/2018	9:34	0
2	2	06/06/2018	9:35	1
3	3	06/06/2018	9:36	2
4	4	06/06/2018	9:37	3
5	5	06/06/2018	9:38	4
6	6	06/06/2018	9:39	5
7	7	06/06/2018	9:40	6
8	8	06/06/2018	9:41	7
9	9	06/06/2018	9:42	8

Obs	id	date	timeb	Msb
10	10	06/06/2018	9:43	9
11	11	06/06/2018	9:44	10
12	12	06/06/2018	9:45	11
13	13	06/06/2018	9:46	12
14	14	06/06/2018	9:47	13
15	15	06/06/2018	9:48	14
16	16	06/06/2018	9:49	15
17	17	06/06/2018	9:50	16
18	18	06/06/2018	9:51	17
19	19	06/06/2018	9:52	18
20	20	06/06/2018	9:53	19

Table 5 shows the SAS program to import data from excel to the SAS dataset. We used this procedure for all datasets for all infants. For example for infant1, we imported 15 datasets from excel to SAS datasets.

Table 5 SAS program to import excel file to SAS dataset

```
SAS Program

libname temp 'd:\ABBAST\temp\test\';

PROC IMPORT OUT= temp.inf1 DATAFILE= libname temp "d:\ABBAST\temp\test\"

DBMS=xlsx REPLACE;

SHEET="auto";

GETNAMES=YES;RUN;
```

Table 6 indicates the SAS program to combine all datasets for each infant. In addition, we created the average for all measurement by minutes.

Table6 SAS Program to combine all datasets and create average by minutes

```
SAS Program
libname temp 'd:\ABBAST\temp\test\';
data one;
set temp.INF1 - temp.INF15; *** List all of the data sets here for each
infant ***;
stime=time;
timeb=substr(stime,1,8);
timea = input(substr(timeb,1,2) !! ':' !! substr(timeb,4,2) !! ':' !!
substr(timeb,7,2),time8.);
drop stime time timeb;
format timea
              time8. ;run;
data two (rename= ( timea=time));;
set one ;
cptd = abd - ft; *** Creat new variable Central-peripheral temperature
difference (ABd-FT);
ID=1;
retain abd ft cptd ext aS isc ip HSP HUM HEATRES;;run;
```

```
data temp.INFall(keep= abd ft cptd ext aS isc ip HSP HUM HEATRES
retain abd ft cptd ext aS isc ip HSP HUM HEATRES;
set two ;
format date mmddyy10.;run;
data three;
set temp.infall;
timeb=time;
tmp=put(time,time5.);
l=lag(tmp);
if n_= 1 then msb=msb+0;
else if 1 ne tmp then msb+1;
format time time8. timeb time5.;
drop tmp l;run;
data four;
set three ;
by id;
if first.id then idn=1;
else idn+1;run;
proc means data=four n mean std min max noprint maxdec=2 ;
     class MSB;
           var abd ft cptd ext aS isc ip HSP HUM HEATRES;
       id abd ft cptd ext aS isc ip HSP HUM HEATRES;;
       output out=avgtwob mean= mabd mft mcptd mext maS misc mip mHSP mHUM
mHEATRES;;
     title ' means ';
       title2 ' Temperature Study '; run;
data five;
set avgtwob;
RUN;
data temp.avgeculall
(keep= IDN ID SITE DATE TIMEb mabd mft mcptd mext maS misc mip mHSP mHUM
mHEATRES ;
retain IDN ID SITE DATE TIMEb mabd mft mcptd mext mas misc mip mHSP mHUM
mHEATRES;
set five;
format MABD MFOOT 8.2 MSB 8.;
 If type =1;
 drop _type_ _freq_; run;
```

Table 7 indicates SAS datasets for 20 observations for all measurement. This datasets created average of all measurements from seconds to minutes. We combined all data sets for each infant from the time the infant was born to the time data collection was completed on each infant. In this dataset, MSB is only correct when there is no missing minutes. For this reason, we merged this dataset with first dataset that we created as master file.

Table7 SAS dataset to combine data and create average measurement by MSB

Obs	DATE	TIMEB	MSB	MABD	MFT	MCPTD	MEXT	MAS	MISC	MIP	MHSP	MHUM	MHEATRES	i
1	06/07/20 18	14:21	1727	•	•	•	•		•	•				1

Obs	DATE	TIMEB	MSB	MABD	MFT	MCPTD	MEXT	MAS	MISC	MIP	MHSP	мним	MHEATRES	i d
2	06/07/20 18	14:22	1728							٠				1
3	06/07/20 18	14:23	1729				•	•	•	•	•	•		1
4	06/07/20 18	14:24	1730				•	•		•	•			1
5	06/07/20 18	14:25	1731					•		•	•			1
6	06/07/20 18	14:26	1732					•						1
7	06/07/20 18	14:27	1733	36.40	36.2	0.20	37.35	37.3	36.50	81.60	0.00	24.00	5616.42	1
8	06/07/20 18	14:28	1734	36.40	36.2	0.20	37.20	37.3	36.50	100.0	0.00	24.00	2833.33	1
9	06/07/20 18	14:29	1735	36.40	36.2	0.20	37.13	37.3	36.50	97.20	0.00	24.13	1779.10	1
10	06/07/20 18	14:30	1736	36.40	36.2	0.20	37.08	37.3	36.50	92.80	0.00	24.47	1310.10	1
11	06/07/20 18	14:31	1737	36.40	36.1	0.29	37.00	37.3	36.50	99.87	0.00	23.40	1065.43	1
12	06/07/20 18	14:32	1738	36.43	36.1	0.33	37.08	37.3	36.50	56.47	0.00	23.77	889.30	1
13	06/07/20 18	14:33	1739	36.50	36.1	0.32	37.24	37.3	36.50	27.00	0.00	24.00	959.43	1
14	06/07/20 18	14:34	1740	36.50	36.2	0.30	37.36	37.3	36.50	32.00	0.00	24.00	1218.07	1
15	06/07/20 18	14:35	1741	36.50	36.2	0.30	37.48	37.3	36.50	19.50	0.00	24.20	1492.17	1
16	06/07/20 18	14:36	1742	36.50	36.2	0.30	37.54	37.3	36.50	5.40	0.00	24.67	1831.97	1
17	06/07/20 18	14:37	1743	36.50	36.2	0.21	37.60	37.3	36.50	0.60	0.00	24.67	2338.47	1
18	06/07/20 18	14:38	1744	36.50	36.3	0.20	37.64	37.3	36.50	0.00	0.00	24.50	2972.23	1
19	06/07/20 18	14:39	1745	36.57	36.3	0.27	37.66	37.3	36.50	0.00	0.00	24.40	3641.50	1
20	06/07/20 18	14:40	1746	36.60	36.3	0.30	37.60	37.3	36.50	0.00	0.00	24.03	4297.57	1

Table 8 indicates SAS program to merge the combined dataset with the master file. This program correctly inserted the missing MSB if there is any in dataset. The final dataset created in this program was used for data analysis.

Table 8 SAS program dataset to merged combine data and master dataset

	SAS Program	
libname temp	o 'd:\ABBAST\temp\test\';	

```
data one;
set temp.ecum;
timec=put(timeb,tod5.); run;
proc sort data=one;
by date timec; run;
data two;
set temp.avgecullcombine;
timec=put(timeb,tod5.);
proc sort data=two;
by date timec; run;
data all;
merge one (in=a) two (in=b);
by date timec;
aa=a;
bb=b;
drop aa bb;run;
data temp.avgcullcombaall;
set all;
run;
```

Day and week created from MSB. Data analyses included frequency tables, measure of central and dispersion, and different types of graphs for overall and each infant. Macro used for each procedure to reduce programing (see table 9 and table 10 for example).

Table 9 Example of Macro for means

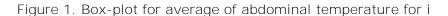
```
SAS Program
Ods rtf;ods listing close;
** means for all infant by days and week ***;
%macro avg (q,t);
proc means data=one n mean std min max maxdec=2;
  class &q;
     var MABT MFT MCPTD MEXT MAS MISC MIP ;
     title ' means /Final data/' &t ;
       title2 ' Temperature Study ';
                                         run;
%mend avg;
%avg(day, by day);
%avg(week, by week);
ods rtf close; ods listing; quit;run;
** Run means by each infant ***;
Ods rtf;ods listing close;
proc sort data =one; by id;
proc means data=one n mean std min max maxdec=2;
       var MABT MFT MCPTD MEXT MAS MISC MIP ;
     title ' means /Final data/by infant';
       title2 ' Temperature Study ';
by id;run;
ods rtf close;ods listing;quit;run;
```

Table 10 Example of Macro for graph

```
SAS Program
```

```
*** Box plot***;
Ods rtf;
ods listing close;
ods select ssplots;
%macro qp (q,t);
proc sort data=one; by week;
proc univariate data = one plot ;
where id= &a;
var MABT MFT MCPTD MEXT MAS MISC MIP ;
by week ;
title "boxplot by week/ each infnat "&t;
title2 ' Temperature Study ';run;
%mend gp;run;
%gp (1,infant1);
%gp (2,infnat2);run;
ods rtf close;ods listing;quit;run;
*** Histogram **:
Ods rtf; ods listing close;
ods graphics /height=1000px width=1000px;
proc freq data=one order=freq noprint;
   tables week* mcptdq/ out=FreqOut2(where=(percent^=.)); by id; run;
PROC SGPLOT DATA = one;
VBAR day / GROUP = mcptdg;
TITLE 'Counts of day by CPTD';
title2 ' Temperature Study '; by id; run;
proc sgplot data=FreqOut2;
 hbarparm category=week response=count / group=mcptdg
      seglabel seglabelfitpolicy=none seglabelattrs=(weight=bold);
 keylegend / opaque across=1 position=bottomright location=inside;
grid;
 yaxis labelpos=top;
                      by id; run; ods rtf close; ods listing; quit; run;
***temperature changes
                       ***;
goptions device = jpeg xpixels = 1500 xmax = 10.5in ypixels = 900 ymax =
6.5in
ftext = 'Swiss' htext = 10pt cback = white ;
%let name=powerpoint;
ods _all_ close;
ods powerpoint file="&name..ppt" style=htmlblue;
options nodate nonumber;
%macro plotinfant;
%DO rate = 0 %TO 43200 %by 1440;
%let daycount=%eval(&daycount + 1);
proc gplot data=one;
        symbol1 i = join color = blue line = 1 w=2 v=point;
        symbol2 i = join color = red line = 1 w = 2 v = point;
        symbol3 i = join color = green line = 1 w = 2 v = point;
        symbol4 i = join color = orange line = 1 w = 2 v = point;
        axis1 order=25 to 40 by 1 label=none;
        axis2 label=none;
        label msb = 'Minutes since birth';
            legend1 label=none
        position=(top center inside)
        mode=share;
        plot mabt*msb = 1 mft*msb = 2 mext*msb = 3 misc*msb
             /overlay legend=legend1 vaxis=axis1 name = "S &rate"
```

Figure 1-3 shows several types of plots used for each measurement to report the data. These graphs created for each infant by days, and weeks. In addition, means, Standard Deviation (SD), and range of all measurement reported by days and weeks for each infant.



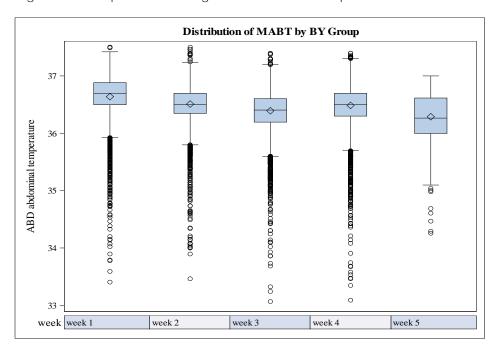


Figure 2. Histogram plot frequency of days by CPTD for infant 2.

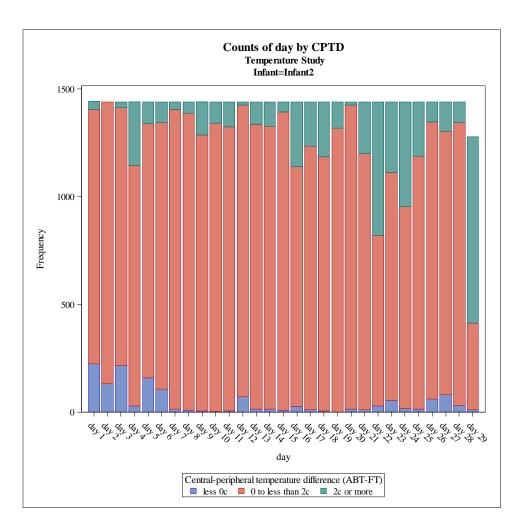
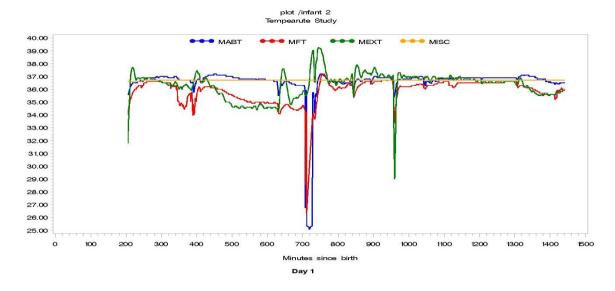


Figure 3. Temperature changes for minutes for day 1 (infant 2).



CONCLUSION

Many times data were not ready in the form needed to be analyze in statistical software. Data collected for this Infant Temperature study went through several

steps in order to be prepared for data analyses. Lag and several functions in SAS were used to prepare data for data analyses. Several programs were used to delete unnecessary rows, create minutes from the time the infant was born to the time the infant completed data collection, replace missing minutes, combine and merge different datasets. Several procedures in SAS were used to analyze data such as Means, Freq, Univariate, Gplot, and Sgplot. All data analyses was performed using SAS/STAT® statistical software, version 9.4. SAS is the most powerful software to handle any type of data.

REFERENCES

- 1. Lyon A, Freer Y. Goals and options in keeping preterm babies warm. Archives in Diseases of Childhood. 2011; 96: F71-F74.
- 2. Knobel-Dail RB, Sloane R, Holditch-Davis D, Tanaka DT. Negative Temperature Differential in Preterm Infants Less Than 29 Weeks Gestational Age: Associations With Infection and Maternal Smoking. Nurs Res. 2017; 66(6): 442-453.
- 3. SAS Institute Incorporated. (2013). SAS for Windows 9.4. Cary, NC: SAS Institute Inc.

CONTACT INFORMATION < HEADING 1>

Your comments and questions are valued and encouraged. Contact the author at:

Abbas S. Tavakoli, DrPH, MPH, ME Clinical Associate Professor College of Nursing University of South Carolina 1601 Greene Street Columbia, SC 29208-4001

Fax: (803) 777-5561

E-mail: abbas.tavakoli@sc.edu

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.

BASIC INSTRUCTIONS

WRITING GUIDELINES

Trademarks and product names

To find correct SAS product names (including use of trademark symbols), if you are a SAS employee, see the $\underline{\text{Master Name List}}$. Otherwise, see $\underline{\text{SAS Trademarks}}$.

- Use superscripted trademark symbols in the first use in title, first use in abstract, and in graphics, charts, figures, and slides.
- Do not abbreviate product names. For example, you cannot use "EM" for SAS® Enterprise Miner™. After having introduced a SAS product name, you can occasionally omit "SAS" for certain products, provided that your editor agrees. For example, after you have introduced SAS® Simulation Studio, you can occasionally use "Simulation Studio."

Writing style

• Use active voice. (Use passive voice only if the recipient of the action needs to be emphasized.) For example:

The product creates reports. (active) Reports are created by the product. (passive)

• Use second person and present tense as much as possible. For example:

You get accurate results from this product. (second person, present tense) The user will get accurate results from this product. (future tense)

• Run spellcheck, and fix errors in grammar and punctuation.

Citing references

All published work that is cited in your paper must be listed in the REFERENCES section.

If you include text or visuals that were written or developed by someone other than yourself, you must use the following guidelines to cite the sources:

- If you use material that is copyrighted, you must mention that you have permission from the copyright holder or the publisher, who might also require you to include a copyright notice. For example: "Reprinted with permission of SAS Institute Inc. from SAS® Risk Dimensions®: Examples and Exercises. Copyright 2004. SAS Institute Inc."
- If you use information from a previously printed source from which you haven't requested copyright permission, you must cite the source in parentheses after the paraphrased text. For example: "The minimum variance defines the distance between cluster (Ward 1984, p. 23)

TIPS FOR USING WORD

To select a paragraph style

- Click the HOME tab. The most common styles in your document are displayed in the top right area of the Microsoft ribbon. If you don't see a style that you want, click the slanted down arrow at the bottom right corner of the Styles area, and scroll through the list. The main styles for this template are headings 1 through 4, PaperBody, and Caption. Avoid using other styles.
- To change a paragraph style, click the paragraph to which you want to apply a style, and then click the style that you want in the ribbon.
- 3. PaperBody (used for most text) is automatically applied when you press Enter at the end of any heading style or the Caption style.

To insert a caption

- 1. Click **REFERENCES** on the main Word menu.
- Click Insert Caption.
- 3. Select the Label type that you want.
- 4. Click OK.

To insert a graphic from a file

- 1. Click **INSERT** on the main Word menu.
- 2. Click Picture.
- 3. In the Insert Picture dialog box, navigate to the file that you want to insert.
- 4. When the name of the file that you want to insert is displayed in the File name box, click Insert.