

## STAT 541: Test 2

1. List the command needed to write the name of the current data set to the SAS LOG.
2. Write OPTIONS statements for the following output options for data set Traffic (assume it is stored in WORK).
  - (a) Input the first 5 observations and list only the first 4 observations
  - (b) Restore the listing of all observations and add row numbers.
  - (c) What would be a reasonable option for FLOW to control the output display of the variable Location?
3. Refer to the data set Traffic again.
  - (a) Write SQL code to save the name of each County, and the average AADT for each county.
  - (b) Assume you know that ncnty=2, where ncnty is the number of counties; write SQL code to save each County and average AADT for each county as macro variables.
4. For the YearTwo data, suppose we created a macro variable &month with a LET statement:

```
libname STAT541 "c:\STAT 541";  
%let month=12;
```

- (a) Provide commands so that the value of the macro variable could be viewed in the LOG. Graduate students should list two ways (the methods should be distinct), while undergraduate students can list only one.
  - (b) Show how you would subset December records from YearTwo and save them in the permanent SAS directory STAT541.
  - (c) The following code should extract December records assuming &month is assigned the character value DEC:

```
data dec;  
set YearTwo;  
cdate=put(Date,date9.);  
cmonth=substr(cdate3,3);  
if cmonth="&month";  
run;
```

Suppose &month is assigned the value **dec** instead. How could the code be modified?
5. The following question uses the Traffic data set. What will be printed by the following commands?

```

proc sort data=traffic;
by descending aadt;
run;

data traffic; set traffic; rank=_n_;
run;

data _null_;
set traffic;
call symput('busiest' || put(_n_,1.),Location);
run;

%put _user_;

options symbolgen;
%let busiestname=busiest2;
proc print;
var AADT;
where rank=input(substr("&busiestname",8,1),1.);
title "AADT for &&&busiestname";
run;

```

6. Suppose we wanted to write a macro that plotted a histogram when one variable was specified, and plotted a scatterplot when two variables were specified. The code for generic variables would like like this:

```

proc sgplot data=gradebook;
title "Histogram of &y";
histogram &y;
run;

proc sgplot data=gradebook;
title "Scatterplot of &y and &x";
scatter x=&x y=&y;
run;

```

- (a) Undergraduate students should write a macro that plots a histogram for a generic variable from the Gradebook data set.
- (b) Graduate students should write a macro that plots either a histogram or a scatterplot (if a second argument is provided) using %IF, %THEN, %DO %END statements for generic variables from the Gradebook data set.

Data sets

County	Intersection	Road	Location	AADT
Spartanburg	2274	I-85	SC 85 to I-26	59800
Spartanburg	103	S-42-9	S-1081 to S-316	1550
Spartanburg	205	S-42-9	S-644 to Victor Hill Road	250
Greenville	2165	I-26	S-492 to SC-14	89400
Greenville	107	S-23-189	Healthstone Lane to US 29	1400
Greenville	111	S-23-347	Barker Rd to Nealy Ferry Road	2600
Greenville	191	S-23-418	S-435 to end	1200

Table 1: Data Set Traffic: County, Intersection ID, Road, Intersection Location and Average Annual Daily Traffic Count for monitored intersections in South Carolina

ID	Date	TL
380	12/15/11	1340
673	1/23/12	1460
110	1/23/12	1260
122	12/15/11	1300
673	12/15/11	1465
045	1/23/12	1250

Table 2: Data Set YearTwo: ID Implant Number, Date of Recapture, and Total Length of 2-year old striped bass

ID	Q1	Q2	HW1	T1
1	7	8	17	84
2	10	10	19	92
3	.	9	18	94
4	9	10	20	85
5	9	10	20	90
6	9	10	20	96

Table 3: Data Set Gradebook: ID, Quiz 1, Quiz 2, Homework 1, Test 1