

# SAS/SQL Test

NAME: \_\_\_\_\_ DATE: \_\_\_\_\_

- 1) Given a dataset "work.claims" containing variables:

Variable	Numeric/ Character	Length	Format
Recip_id	C	9	\$9.
Prov_id	C	9	\$9.
Begin_dos	N	8	mmddyy10.
End_dos	N	8	mmddyy10.
Dx1	C	6	\$6.
Dx2	C	6	\$6.
Dx3	C	6	\$6.
Dx4	C	6	\$6.
Dx5	C	6	\$6.

- a. Write the code used to calculate the number of claims each provider had in 2000 (begin\_dos in 2000).

```
proc sql;
create table clm_cnt as
select prov_id, distinct(recip_id) as clm_cnt
from work.claims
where begin_dos between '01JAN2000'd a '31DEC2000'd;
quit;
```

SAS Test

- b. Write the code used to calculate the number of claims that had a diagnosis code between 290.xx and 319.xx coded for each provider in 2000 (begin\_dos in 2000).

```
proc sql;
create table clm_cnt as
select prov_id, distinct( recip_id) as clm_cnt
from work.claims
where begin_dos between '01JAN2000'd a '31DEC2000'd
and (
substr(left(trim(dx1))1,1,3) ge '290' and and substr(left(trim(dx1))1,1,3) le '319'
or substr(left(trim(dx2))1,1,3) ge '290' and and substr(left(trim(dx2))1,1,3) le '319'
or substr(left(trim(dx3))1,1,3) ge '290' and and substr(left(trim(dx3))1,1,3) le '319'
or substr(left(trim(dx4))1,1,3) ge '290' and and substr(left(trim(dx4))1,1,3) le '319'
or substr(left(trim(dx5))1,1,3) ge '290' and and substr(left(trim(dx5))1,1,3) le '319'
)
; quit;
```

SAS Test

c. Write the code used to calculate the number of times a diagnosis code between 290.xx and 319.xx was coded for each provider in 2000 (begin\_dos in 2000).  
(if you can, please use a macro or an array)

```
%macro looper(val);
proc sql;
create table dx_cnt.&val as
select
  prov_id,
  sum(case
    when substr(left(trim(&&dx.&val))1,1,3) ge '290' and and substr(left(trim(&&dx.&val))1,1,3) le '319' then 1
    else 0
  end as diag_ind) as diag_cnt.&val
from clm_cnt

where begin_dos between '01JAN2000'd a '31DEC2000'd
and substr(left(trim(&&dx.&val))1,1,3) ge '290' and and substr(left(trim(&&dx.&val))1,1,3) le '319'

group by prov_id
; quit;

%if &val ="1" %then %do;
data dx_sum;
  set dx_cnt.&val;
run;
%else do;
proc append base = dx_sum data=dx_cnt.&val; run;
%end;

%mend looper;
%looper(val=1);
%looper(val=2);
%looper(val=3);
%looper(val=4);
%looper(val=5);
*** Final output data set =dx_sum;
```

## SAS Test

```
**--- With arrays ---;
data dx_cnt;
  set work.claims (where = (begin_dos between '01JAN2000'd a '31DEC2000'd));
  array diag{5} dx1--dx5;
  array diag_cnt{5} dx_cnt1--dx_cnt5;
  for i=1 to 5;
    if substr(left(trim(diag1{i})),1,3) between 290 and 319 then diag_cnt{i}=1;
    else diag_cnt{i}=0;
  end;

** Then summarize diag_cnt1 - diag_cnt5;
Proc sort data=dx_cnt;
by prov_id;

Proc summary data=dx_cnt;
by prov_id;
var diag_cnt1 – diag_cnt5;
output out =dx_sum sum=;
run;

*** Final output = dx_sum;
```

2) What SAS statement, within a data step, would be used to change the appearance of a date variable for a proc print?

format

3) Name some SAS procedures used to do data summarization.

```
proc summary
proc freq
proc sql
proc univariate
proc report
```

4) What SAS procedure is used to order SAS data sets? How can you remove duplicates?

proc sort, use nodupkey or nodup option

SAS Test

5) What SAS statement(s), can be used within a data step, to prevent the output of certain variables into the output data set?

keep/drop

6) What function do you use to grab a portion of a character variable?

substr

7) What SAS statement is used to allow a variable to not change its value from one data step iteration to the next?

retain

8) What function is used to convert a character variable to numeric?

input

9) What SAS system option(s) within a data step would be used to set the 9th and only the 9th observation from a data set?

if \_N\_ = 9

10) What SQL statement is necessary to build a SAS data set via proc sql?

```
proc sql;  
create table newData as  
select *  
from oldData  
where (constraints)  
; quit;
```

11) What SQL statement would be used to remove records from a SAS data set or dbms table?

delete

SAS Test

12) What SQL statement would be used to add records to a SAS data set or DBMS table?

union/union all

13) If I wanted to do a many to many join, can this be done with: a) only Proc SQL, b) only the merge statement or c) either a) or b)? Give an example of one to one match merging using the data step or proc sql.

Ans. c, both

```
proc sql;
create table myData as
select a.*, b.*
from data1 a
left join data2 b
on a.criteria = b.criteria
; quit;
```

```
proc sort data=data1; by criteria;
proc sort data=data2; by criteria;
data myData;
merge data1(in=a) data2(in=b);
by criteria;
if a or b;
run;
```

14) What SQL statement is used to change values in current records only?

case

15) What are the acceptable values for the first character of a macro variable definition? Please give an example.

letter (ex. %let m=1; ), underscore (%let \_m =1;)

16) What is the first character for all macro variables (being resolved)?

&

SAS Test

17) Name date formats you know in SAS and give examples.

- format date mmddyy10.; 01012001 for 01/01/2001
- format date date9.; 01JAN2001 for 01/01/2001
- format date mmddyy8.; 010101 for 01/01/01
- format date ddmmyy10.; 01012001 for 01/01/2001

18) What subroutine is used to assign a data set variable to a macro variable in the data step? Can I use this macro variable within the data step I created it?

```
call symput('var1',var1);
```

19) If I were to resolve a macro variable within quotes, a) the quotes have to be single quotes, b) the quotes have to be double quotes, or c) either a) or b)?

B) Double

20) What macro statement writes text to the SAS log?

```
%put Some text here;
```