**Homework 4** ***Biostatistics 203A***

***Due date: Week5 on Friday 6PM***

In this homework, we will use XPT file data set. The goals of this homework is

1. Read XTP files to SAS files
2. Merge XPT files and compare the file attributes, and Data QC
3. Use the ods function to generate tables and reports.

.

**Exercise 1. Read XPT files to SAS data set [20 points; 5 points x4]**

Using the pseudo code below, read XPT files and list the attribute of data set using Proc Contents

Make a folder for XPT and XPTread in your computer.

libname xptfile0 XPORT ".. \LabHomework3\XPT\TR.xpt";

libname ino ".. \LabHomework3\XPTread";

**proc** **copy** in=xptfile0 out=ino memtype=data;

**run**;

**proc** **contents** data=ino.IC order=varnum;

**run**;

**proc** **print** data=ino.IC (firstobs=**1** obs=**5**);

**run**;

1. Find out the number of data file label, number of row and columns

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | File label | # of row | # of columns |
| IC |  |  |  |
| TR |  |  |  |
| TU |  |  |  |
| RS |  |  |  |

1. The file set contains the CDISC STDM style. The elements of data are the record of RECIST reading by two readers. Merge two files of TR and RS files after sorting by SUBJIDN**,** date,VISIT (VISITNUM), and reader ID. The variable name of date is TRDTC in TR file and RSDTC in RS file. The variable name of reader ID is TRVALID in TR file and RSVALID in RS file. Find the number of subject and number of visit that appears in both files after merging to TR and RS file.

When the rows that were available in TR file only, what was the visit (VISITNUM)?

When the rows that were available in RS file only, what was the mRECIST1.1 response (RSORRES)?

**Here are the pseudo code:**

**proc** **sort** data=TR ;

by SUBJIDN date VISIT readerID ;

**run**;

**proc** **sort** data=RS ;

by SUBJIDN date VISIT readerID;

**run**;

**data** ino.TRRS;

merge ino.TR (rename = (old\_name =new\_name oldname2=new\_name2))

ino.RS (rename = (old\_name =new\_name oldname2=new\_name2));

by SUBJIDN date VISIT readerID;

**run**;

**data** TRRS2;

set ino.TRRS;

if TRSEQ~=**.** & RSSEQ~=**.** then merge=**3**;

if TRSEQ=**.** & RSSEQ~=**.** then merge=**2**;

if TRSEQ~=**.** & RSSEQ=**.** then merge=**1**;

**run**;

**proc** **freq** data=TRRS\_first;

table merge\*VISITNUM;

**run**;

**proc** **freq** data=TRRS2;

where merge=;

table ;

**run**;

|  |  |
| --- | --- |
| Merge |  |
| # of row only appears in TR file |  |
| # of row only appears in RS file |  |
| # of row only appears in both files |  |

|  |  |
| --- | --- |
| Merge | VISITNUM |
| rows only appears in TR file |  |

|  |  |  |
| --- | --- | --- |
| Merge | Subject ID | mRECIST1.1 response |
| rows only appears in RS file |  |  |

1. Find the number of subject and number of visit that appears in files after merging to TR and RS file.

**Data** ;

Set ;

by SUBJIDN VISITNUM;

if (first.SUBJIDN=**1** or first.VISITNUM=**1**) then output ;

**run**;

**proc** **freq** data= ;

table ;

**run**;

|  |  |  |
| --- | --- | --- |
| Merge | # of Subject ID | # of visit |
| Merged File : Total |  |  |

|  |  |
| --- | --- |
| Merge | # of visit |
| Merged File : Total |  |
| Only from TR |  |
| Only from RS |  |
| Both available |  |

1. Use TR file and summarize the sum of target lesion at baseline using mean and standard deviation by the accepted reader ()

[**TRTEST=**Sum of Diameter, **TRACPTFL=”Y”, VISIT=”Screening”]**

ods listing close;

ods rtf body="example.rtf";

title1 "Baseline Target Disease Burden";

**proc**;

**run**;

ods rtf close;

ods listing;