\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Question One \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname shoes '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066';

**proc** **print** data=shoes.shoes\_tracker; /\*1a\*/

var Product\_Category Product\_Name Supplier\_Country Supplier\_Name;

where Product\_Category ne 'Shoes' or /\*since the question asked for observations that don't meet both requirements no observation is met if you use a 'and' qualifier but on the other hand if the question meant observation that didn’t meet both the first and second requirement but not together an 'or' is used and results for both are printed.\*/

Supplier\_Country not in ('GB','US');

**run**;

**proc** **freq** data=shoes.shoes\_tracker nlevels; /\*1b\*/

tables Supplier\_Name Supplier\_ID;

where Supplier\_Name in ('3Top Sports', 'Greenline Sports Ltd') or

Supplier\_ID in (**2963**,**14682**);

**run**;

**proc** **print** data=shoes.shoes\_tracker; /\*1c\*/

var Product\_ID Product\_Name;

where Product\_Name ne propcase(Product\_Name);

**run**;

**data** shoes\_tracker; /\*1d\*/

set shoes.shoes\_tracker;

Supplier\_Country = Upcase(Supplier\_Country); /\*obs 10 uppercase country\*/

if Supplier\_Country='UT' then Supplier\_Country = 'US'; /\*obs 5 is US for country\*/

Product\_Category = 'Shoes'; /\*obs 2 is shoes for category\*/

if \_N\_=**1** then Supplier\_ID = **2963**; /\*obs 1 is 2963 for supplierID\*/

if Supplier\_Name = '3op Sports' then Supplier\_Name = '3Top Sports'; /\*obs 3,7 is 3TopSports for name\*/

if \_N\_=**4** then Product\_ID=**220200300079**; /\*obs 4 is 220200300079\*/

else if \_N\_=**8** then Product\_ID=**220200300129**; /\*obs 8 is 220200300129\*/

Product\_Name = propcase(Product\_Name); /\*obs 3should be proper case\*/

if Supplier\_ID=**14682** and Supplier\_Name = '3Top Sports' then Supplier\_Name = 'Greenline Sports Ltd'; /\*obs 9 should be Greenline...\*/

**proc** **print**;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Question Two \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname qtr '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066';

**data** qtr2; /\*2a\*/

set qtr.qtr2\_2007;

**run**;

**proc** **print**; /\*2b\*/

title 'observations on work.qtr2 that do not meet both of the requirements';

where Delivery\_Date lt Order\_Date or

Order\_Date not between **'01APR2007'd** and **'30JUN2007'd**;

**run**;

**proc** **freq** nlevels; /\*2c\*/

tables Order\_ID Order\_Type;

**run**;

**proc** **print**;

title 'observations for which order\_type does not have a value of 1, 2, or 3'; /\*2d\*/

var Order\_ID Order\_Type;

where Order\_Type not in (**1**,**2**,**3**);

**run**;

**data** qtr; /\*2e\*/

set qtr2;

if Order\_ID=**1242012259** then Delivery\_Date = **'12MAY2007'd**;

if Order\_ID=**1242449327** then Order\_Date = **'26JUN2007'd**;

if \_n\_=**18** then Order\_ID = **1241895587**;

if \_n\_=**19** then Order\_ID = **1241895564**;

if Order\_Type not in (**1**,**2**,**3**) then

Order\_Type =**3**;

**run**;

**proc** **print**; /\*2f\*/

title "Checking that work.qtr meets all the datastep requirements";

where Delivery\_Date lt Order\_Date or /\*this where statement should have no observations\*/

Order\_Date not between **'01APR2007'd** and **'30JUN2007'd** or

Order\_ID = . or

Order\_Type not in (**1**,**2**,**3**);

**run**;

**proc** **freq** nlevels;

title "procedure to show that there are 36 unique observations";

tables Order\_ID;

**run**;

**proc** **print**;

title"procedure to show that the observations have the correct Order\_Type and ID.";

**run**;

title;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Question Three \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname pc '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066';

**proc** **contents** data=pc.price\_current; **run**; /\*3a\*/

**proc** **print** data=pc.price\_current; /\*3b\*/

where Unit\_Cost\_Price = . or

Unit\_Sales\_Price = . or

Factor = . ;

**run**;

**proc** **means** data=pc.price\_current; /\*3c\*/

var Unit\_Cost\_Price Unit\_Sales\_Price Factor; **run**;

where Unit\_Cost\_Price in(**1**, **400**) or

Unit\_Sales\_Price in (**3**, **800**) or

Factor in (**1**, **1.05**);

**run**; /\*From the results Unit\_Cost\_Price is in the correct range

Unit\_Sales\_Price has an observation '5730.00' that is gt 800;

Factor has both observations out of the range of 1-1.05\*/

ods select extremeobs; /\*3d\*/

**proc** **univariate** data=pc.price\_current;

var Unit\_Sales\_Price Factor;

**run**;

**data** price\_current; /\*3e\*/

set pc.price\_current;

if Product\_ID=**220200200022** then Unit\_Sales\_Price = **57.30**;

if Product\_ID=**240200100056** then Unit\_Sales\_Price = **41.20**;

if \_N\_=**14** then Factor = **1.0**;

if \_N\_=**170** then Factor = **1.02**;

if \_N\_=**134** then Factor = **1.0**;

**run**;

**proc** **print**;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Question Four \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname lab '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066';

**proc** **contents** data=lab.labsub1 position; **run**; /\*4a\*/

**proc** **means** data=lab.labsub1; /\*4b\*/

var hgp htp tcp tgp lcp hdp fbpsi crp sgp urp;

**run**;

**proc** **means** data=lab.labsub1 max ; /\*4c\*/

**run**;

**data** labsub2; /\*4d\*/

set lab.labsub1;

if hgp = **88888** then hgp = .;

if htp =**88888** then htp = . ;

if tcp = **888** then tcp = . ;

if tgp = **8888** then tgp = . ;

if lcp = **888** then lcp = . ;

if hdp = **888** then hdp = . ;

if fbpsi = **8888** then fbpsi = . ;

if crp = **88888** then crp = . ;

if sgp = **888** then sgp = . ;

if urp = **88888** then urp = . ;

**run**;

**proc** **means** data=labsub2; /\*4e\*/

**run**; /\* none of the maximum observations contain fill values of 8's\*/