\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*QUESTION ONE\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname customer '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066'; /\*question 1a\*/

title "Data set with female customers"; /\*question 1b(i)\*/

**data** youth;

set customer.customer\_dim;

where Customer\_Gender = 'F';

**run**;

**proc** **print** data=youth;

var customer\_lastname customer\_firstname customer\_gender;

**run**;

title;

title "Data set where customer\_Age is between 18 and 36"; /\*question 1b(ii)\*/

**data** youth;

set customer.customer\_dim;

where customer\_age between **18** and **36**;

**run**;

**proc** **print** data=youth;

var customer\_lastname customer\_firstname customer\_age;

**run**;

title;

title "Data set that have the word Gold in their Customer\_Group"; /\*question 1b(iii)\*/

**data** youth;

set customer.customer\_dim;

where customer\_group contains 'Gold';

**run**;

**proc** **print** data=youth;

var customer\_lastname customer\_firstname customer\_group;

**run**;

title;

title "Data set containing only selected variables";

**data** youth;

set customer.customer\_dim;

keep customer\_name customer\_age customer\_birthdate customer\_gender customer\_group;

**run**;

**proc** **contents** data=youth;

**proc** **print** data=youth;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*QUESTION TWO\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname prg1 '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066'; /\*2a\*/

title "Data from products from Great Britain

(GB), Spain (ES), or Netherlands (NL) and Product\_Category values

that end in the word Sports."; /\*2b\*/

**data** sports;

set prg1.product\_dim;

where (supplier\_country in ('GB', 'ES', 'NL')) and (product\_category like '%Sports');

**run**;

**proc** **print** data=work.sports;

var product\_name supplier\_country product\_category;

**run**;

title;

title "Data set not including variables Product\_ID, Product\_Line, Product\_Group, and Supplier\_ID"; /\*2c\*/

**data** sports;

set prg1.product\_dim;

drop Product\_ID Product\_Line Product\_Group Supplier\_ID;

**run**;

**proc** **contents** data=work.sports;

**run**;

title;

title "Data set with labels added"; /\*2d\*/

**data** sports;

set prg1.product\_dim;

label Product\_Category = 'Sports Category'

Product\_Name = 'Product Name (Abbrev)'

Supplier\_Name = 'Supplier Name (Abbrev)';

**run**;

**proc** **contents** data=work.sports;

**run**;

**proc** **print** data=work.sports(obs=**5**) label;

**run**;

title;

title "product name and supplier name formatted to 15 characters"; /\*2e\*/

**data** sports;

set prg1.product\_dim;

format product\_name $15. supplier\_name $15.;

**run**;

**proc** **print** data=work.sports label;

**run**;

title;

title "contents to show updated formats and labels"; /\*2f\*/

**data** sports;

set prg1.product\_dim;

label Product\_Category = 'Sports Category'

Product\_Name = 'Product Name (Abbrev)'

Supplier\_Name = 'Supplier Name (Abbrev)';

format product\_name $15. supplier\_name $15.; **run**;

**proc** **contents** data=work.sports;

**run**;

**proc** **print** data=work.sports(obs=**14**) label; /\*2g\*/

**run**;

title;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*QUESTION THREE\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname Nhanes3 '/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066';

**data** examsub1;/\*3a\*/

set Nhanes3.exam;

keep hsageir hssex dmaracer bmpwt bmpht pep6g1 pep6h1 pep6i1 pep6g3 pep6h3 pep6i3 sppfvc sppfev1;

label hsageir = 'age'

hssex = 'gender'

dmaracer = 'race'

bmpwt = 'wt\_kg'

bmpht = 'ht\_cm'

pep6g1 = 'sbp1'

pep6h1 = 'sbp2'

pep6i1 = 'sbp3'

pep6g3 = 'dbp1'

pep6h3 = 'dbp2'

pep6i3 = 'dbp3'

sppfvc = 'fvc'

sppfev1 = 'fvc1';

format pep6g1 **12.** pep6h1 **12.** pep6i1 **12.** pep6g3 **12.** pep6h3 **12.** pep6i3 **12.**; /\*3b\*/

**run**;

**proc** **print** data=examsub1 (obs=**7**);

**run**;

**proc** **contents** data=examsub1;

**run**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*QUESTION FOUR\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname NH "/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066"; /\*4a\*/

**data** labsub1; /\*4b\*/

set NH.lab;

keep seqn hgp htp tcp tgp lcp hdp fbpsi crp sgp urp; /\*4c\*/

label seqn ='sequence number'

hgp ='hemoglobin (g/dl)'

htp ='hematocrit (%)'

tcp ='cholesterol (mg/dl)'

tgp ='triglycerides (mg/dl)'

lcp ='low density lipoprotein (mg/dl)'

hdp ='high density lipoprotein (mg/dl)'

fbpsi ='fibrinogen (mg/dl)'

crp ='C-reactive protein (mg/dl)'

sgp ='plasma glucose (mg/dl)'

urp ="urinary creatinine (mg/dl)";

**run**;

**proc** **contents** data=labsub1; /\*4d\*/

title "proc contents step to verify that the data set labsub1 has the correct variables.";

**run**;

**proc** **print** data=labsub1(obs=**5**) label;

title "procedure print to verify the first 5 observations on the data set labsub1.";

**run**;

title;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*QUESTION FIVE\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

libname nh3 "/home/u62109636/my\_shared\_file\_links/jhshows0/STA5066"; /\*5a\*/

**data** mortsub1; /\*5b\*/

set nh3.mortality;

where eligstat = **1**;

keep SEQN MORTSTAT;/\*5c\*/

label MORTSTAT = "Mortality Status";

**run**;

**proc** **contents** data=mortsub1; /\*5d\*/

**run**;

**proc** **print** data=mortsub1(obs=**100**) label;/\*5e\*/

**run**;