# SAS ASSIGNMENT 10

# (Editing Data)

/\* Question 1 \*/

libname dataLib "/courses/d649d56dba27fe300/STA5066";

proc print data=dataLib.shoes\_tracker;

where missing(Product\_Category) or Supplier\_Country not in ('GB', 'US');

run;

proc freq data=dataLib.shoes\_tracker nlevels;

tables Supplier\_Name Supplier\_ID;

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

and Supplier\_ID in (2963, 14682);

run;

proc print data=dataLib.shoes\_tracker;

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

run;

data work.shoes\_tracker;

set dataLib.shoes\_tracker;

if n = 10 and Supplier\_Country =: 'gb' then Supplier\_Country = 'GB';

if n = 5 and Supplier\_Country = 'UT' then Supplier\_Country = 'US';

if n = 2 and Product\_Category = 'Shoes' then Product\_Category = '';

if n = 1 and missing(Supplier\_ID) then Supplier\_ID = 2963;

if n in (3, 7) and Supplier\_Name = '3op Sports' then Supplier\_Name = '3Top Sports';

if n = 4 and Product\_ID = '22020030007' then Product\_ID = '220200300079';

if n = 8 and Product\_ID = '22020030012' then Product\_ID = '220200300129';

if n = 3 then Product\_Name = propcase(Product\_Name);

if n = 9 and Supplier\_Name = '3Top Sports' and Supplier\_ID = 14682 then Supplier\_Name = 'Greenline Sports Ltd';

run;

/\* Question 2 \*/

libname dataLib "/courses/d649d56dba27fe300/STA5066";

data work.qtr2;

set dataLib.qtr2\_2007;

run;

proc print data=work.qtr2;

where Delivery\_Date < Order\_Date or not ('01APR2007'd <= Order\_Date <= '30JUN2007'd);

run;

proc freq data=work.qtr2 nlevels;

tables Order\_ID Order\_Type;

run;

proc freq data=work.qtr2;

tables Order\_ID / missing;

run;

proc print data=work.qtr2;

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

run;

data work.qtr2;

set work.qtr2;

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

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

if n = 18 and missing(Order\_ID) then Order\_ID = 1241895587;

if n = 19 and missing(Order\_ID) then Order\_ID = 1241895564;

if n = 2 and Order\_Type = 0 then Order\_Type = 3;

if n = 10 and Order\_Type = 4 then Order\_Type = 3;

run;

/\* Question 3 \*/

libname dataLib "/courses/d649d56dba27fe300/STA5066";

proc contents data=dataLib.price\_current;

run;

proc print data=dataLib.price\_current;

where missing(Unit\_Cost\_Price) or missing(Unit\_Sales\_Price) or missing(Factor);

run;

proc means data=dataLib.price\_current n mean min max;

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

where Unit\_Cost\_Price < 1 or Unit\_Cost\_Price > 400 or

Unit\_Sales\_Price < 3 or Unit\_Sales\_Price > 800 or

Factor < 1 or Factor > 1.05;

run;

ods select ExtremeObs;

proc univariate data=dataLib.price\_current;

var Unit\_Sales\_Price Factor;

run;

data work.price\_current;

set dataLib.price\_current;

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

if Product\_ID = '240200100056' and missing(Unit\_Sales\_Price) then Unit\_Sales\_Price = 41.20;

if n = 14 and Factor = 100.0 then Factor = 1.0;

if n = 170 and Factor = 10.20 then Factor = 1.02;

if n = 134 and Factor = 0.01 then Factor = 1.0;

run;

/\* Question 4 \*/

libname dataLib "/courses/d649d56dba27fe300/STA5066";

proc contents data=dataLib.labsub1 position;

run;

proc means data=dataLib.labsub1;

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

run;

data work.labsub2;

set dataLib.labsub1;

array varList hgp htp tcp tgp lcp hdp fbpsi crp sgp urp;

do i = 1 to dim(varList);

if varList[i] in (8, 88, 888, 8888, 88888) then varList[i] = .;

end;

drop i;

run;

proc means data=work.labsub2;

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

run;