NOTE: Copyright (c) 2016 by SAS Institute Inc., Cary, NC, USA.

NOTE: SAS (r) Proprietary Software 9.4 (TS1M6)

Licensed to STATISTICS CANADA - UNLIMITED SAS 9, Site 70186445.

NOTE: This session is executing on the X64\_10PRO platform.

NOTE: Analytical products:

SAS/STAT 15.1

SAS/ETS 15.1

SAS/OR 15.1

SAS/IML 15.1

SAS/QC 15.1

NOTE: Additional host information:

X64\_10PRO WIN 10.0.18362 Workstation

NOTE: SAS initialization used:

real time 5.49 seconds

cpu time 2.43 seconds

1 libname tutorial "U:\BT\Banff 2.07 Tutorial\SAS\_English\Data";

NOTE: Libref TUTORIAL was successfully assigned as follows:

Engine: V9

Physical Name: U:\BT\Banff 2.07 Tutorial\SAS\_English\Data

2

3

4 /\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

5 \* Lesson 3 - PROC Outlier \*

6 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

7

8 /\* Run query to select only data with variable STRAT>1. Proc Outlier will not be run

9 for units in STRAT=1 according to the user's survey specifications. \*/

10 proc sql;

11 create table tutorial.outlier\_indata as

12 select IDENT,

13 AREA,

14 STRAT,

15 HEN\_LT20,

16 HEN\_GE20,

17 HEN\_OTH,

18 HEN\_TOT,

19 EGG\_LAID,

20 EGG\_SOLD,

21 EGG\_VALU,

22 QR\_REV,

23 QR\_EXP,

24 QR\_PROF

25 from tutorial.current

26 where STRAT > 1;

NOTE: Table TUTORIAL.OUTLIER\_INDATA created, with 117 rows and 13 columns.

27 quit;

NOTE: PROCEDURE SQL used (Total process time):

real time 0.15 seconds

cpu time 0.06 seconds

28 data work.indata;

29 set tutorial.outlier\_indata;

30 run;

NOTE: There were 117 observations read from the data set TUTORIAL.OUTLIER\_INDATA.

NOTE: The data set WORK.INDATA has 117 observations and 13 variables.

NOTE: DATA statement used (Total process time):

real time 0.04 seconds

cpu time 0.01 seconds

31 data work.hist;

32 set tutorial.historical;

33 if IDENT='R005' then HEN\_TOT=-HEN\_TOT;

34 if IDENT='R007' then DELETE;

35 if IDENT='R042' then HEN\_TOT=.;

36

37 run;

NOTE: There were 162 observations read from the data set TUTORIAL.HISTORICAL.

NOTE: The data set WORK.HIST has 161 observations and 10 variables.

NOTE: DATA statement used (Total process time):

real time 0.02 seconds

cpu time 0.01 seconds

38

39 /\* First execution of Proc Outlier for variables with only positive values. \*/

40 proc outlier

41 data = work.indata

42 hist = work.hist

43 outstatus = tutorial.outlier\_outstatus1

44 outsummary=outsum1

45 method = historic

46 mei = 4

47 mii = 7

48 mdm = 0.05

49 exponent = 0

50 minobs = 10

51 outlierstat

52 ;

53 id IDENT;

54 var HEN\_TOT HEN\_LT20;

55 by AREA;

56 run;

NOTE: --- Banff System 2.08.002b2DEV developed by Statistics Canada ---

NOTE: PROCEDURE OUTLIER Version 2.06.002b2DEV

NOTE: Created on Oct 27 2021 at 11:04:02

NOTE: Email: banff@statcan.gc.ca

NOTE:

NOTE: DATA = WORK.INDATA

NOTE: HIST = WORK.HIST

NOTE: OUTSTATUS = TUTORIAL.OUTLIER\_OUTSTATUS1

NOTE: OUTSUMMARY = WORK.OUTSUM1

NOTE: METHOD = HISTORIC

NOTE: SIDE = BOTH (default)

NOTE: MII = 7

NOTE: MEI = 4

NOTE: MDM = 0.05

NOTE: EXPONENT = 0

NOTE: MINOBS = 10

NOTE: OUTLIERSTAT

NOTE: REJECTNEGATIVE (default)

NOTE: REJECTZERO (default)

NOTE: ID = IDENT

NOTE: VAR = HEN\_TOT HEN\_LT20

NOTE: WITH not specified.

NOTE: WEIGHT not specified.

NOTE: BY = AREA

NOTE: The following messages are for the BY group:

AREA=91

NOTE: Number of observations in DATA data set .....................: 11

NOTE: dropped because ID only in DATA ........................: 1

NOTE: Number of valid observations in DATA data set ...............: 10

NOTE: Number of observations in HIST data set .....................: 18

NOTE: dropped because ID only in HIST ........................: 8

NOTE: Number of valid observations in HIST data set ...............: 10

NOTE: The following messages are for the variable:

HEN\_TOT

NOTE: Number of records used in Outlier Detection : 7.

WARNING: EI\_Outlier: Number of records is less than MINOBS 10. Outlier not processed.

WARNING: There were 1 observations dropped from the HIST data set because the variable HEN\_TOT is negative.

WARNING: There were 2 observations dropped from the HIST data set because the variable HEN\_TOT is 0.

NOTE: The following messages are for the variable:

HEN\_LT20

NOTE: Number of records used in Outlier Detection : 0.

WARNING: EI\_Outlier: Number of records is less than MINOBS 10. Outlier not processed.

WARNING: There were 6 observations dropped from the DATA data set because the variable HEN\_LT20 is 0.

WARNING: There were 4 observations dropped from the HIST data set because the variable HEN\_LT20 is 0.

NOTE: The above message was for the following BY group:

AREA=91

NOTE: The following messages are for the BY group:

AREA=92

NOTE: Number of observations in DATA data set .....................: 6

NOTE: Number of valid observations in DATA data set ...............: 6

NOTE: Number of observations in HIST data set .....................: 18

NOTE: dropped because ID only in HIST ........................: 12

NOTE: Number of valid observations in HIST data set ...............: 6

WARNING: Number of valid observations in DATA data set less than MINOBS 10. Outlier not processed.

NOTE: The above message was for the following BY group:

AREA=92

NOTE: The following messages are for the BY group:

AREA=93

NOTE: Number of observations in DATA data set .....................: 25

NOTE: Number of valid observations in DATA data set ...............: 25

NOTE: Number of observations in HIST data set .....................: 26

NOTE: dropped because ID only in HIST ........................: 1

NOTE: Number of valid observations in HIST data set ...............: 25

NOTE: The following messages are for the variable:

HEN\_TOT

NOTE: Number of records used in Outlier Detection : 20.

NOTE: The median of the ratios : 1.35503

NOTE: Value of the 25th percentile : -0.35503

Value of the 50th percentile : -0.00227

Value of the 75th percentile : 0.64259

NOTE: Acceptance interval : [ -1.41330 , 2.57720 ]

Exclusion intervals : [ -2.47157 , -1.41330 [ U ] 2.57720 , 4.51181 ]

Imputation intervals : ] -infinity , -2.47157 [ U ] 4.51181 , +infinity [

WARNING: There were 1 observations dropped from the DATA data set because the variable HEN\_TOT is negative.

WARNING: There were 1 observations dropped from the HIST data set because the variable HEN\_TOT is missing.

WARNING: There were 3 observations dropped from the HIST data set because the variable HEN\_TOT is 0.

NOTE: The following messages are for the variable:

HEN\_LT20

NOTE: Number of records used in Outlier Detection : 0.

WARNING: EI\_Outlier: Number of records is less than MINOBS 10. Outlier not processed.

WARNING: There were 16 observations dropped from the DATA data set because the variable HEN\_LT20 is 0.

WARNING: There were 9 observations dropped from the HIST data set because the variable HEN\_LT20 is 0.

NOTE: The above message was for the following BY group:

AREA=93

NOTE: The following messages are for the BY group:

AREA=94

NOTE: Number of observations in DATA data set .....................: 75

NOTE: Number of valid observations in DATA data set ...............: 75

NOTE: Number of observations in HIST data set .....................: 95

NOTE: dropped because ID only in HIST ........................: 20

NOTE: Number of valid observations in HIST data set ...............: 75

NOTE: The following messages are for the variable:

HEN\_TOT

NOTE: Number of records used in Outlier Detection : 67.

NOTE: The median of the ratios : 1.00000

NOTE: Value of the 25th percentile : -0.18182

Value of the 50th percentile : 0.00000

Value of the 75th percentile : 0.12158

NOTE: MDM specified will be ignored because the value of the median of effects is 0.

NOTE: Acceptance interval : [ -0.72727 , 0.48630 ]

Exclusion intervals : [ -1.27273 , -0.72727 [ U ] 0.48630 , 0.85103 ]

Imputation intervals : ] -infinity , -1.27273 [ U ] 0.85103 , +infinity [

WARNING: There were 2 observations dropped from the DATA data set because the variable HEN\_TOT is missing.

WARNING: There were 2 observations dropped from the DATA data set because the variable HEN\_TOT is negative.

WARNING: There were 2 observations dropped from the DATA data set because the variable HEN\_TOT is 0.

WARNING: There were 2 observations dropped from the HIST data set because the variable HEN\_TOT is 0.

NOTE: The following messages are for the variable:

HEN\_LT20

NOTE: Number of records used in Outlier Detection : 11.

NOTE: The median of the ratios : 0.92784

NOTE: Value of the 25th percentile : -0.78430

Value of the 50th percentile : 0.00000

Value of the 75th percentile : 0.26527

NOTE: MDM specified will be ignored because the value of the median of effects is 0.

NOTE: Acceptance interval : [ -3.13719 , 1.06106 ]

Exclusion intervals : [ -5.49009 , -3.13719 [ U ] 1.06106 , 1.85686 ]

Imputation intervals : ] -infinity , -5.49009 [ U ] 1.85686 , +infinity [

WARNING: There were 1 observations dropped from the DATA data set because the variable HEN\_LT20 is missing.

WARNING: There were 1 observations dropped from the DATA data set because the variable HEN\_LT20 is negative.

WARNING: There were 49 observations dropped from the DATA data set because the variable HEN\_LT20 is 0.

WARNING: There were 13 observations dropped from the HIST data set because the variable HEN\_LT20 is 0.

NOTE: The above message was for the following BY group:

AREA=94

NOTE: There were 117 observations read from the data set WORK.INDATA.

NOTE: There were 161 observations read from the data set WORK.HIST.

NOTE: The data set TUTORIAL.OUTLIER\_OUTSTATUS1 has 15 observations and 14 variables.

NOTE: The data set WORK.OUTSUM1 has 7 observations and 24 variables.

NOTE: PROCEDURE OUTLIER used (Total process time):

real time 0.29 seconds

cpu time 0.07 seconds

57 proc print data=outsum1;

NOTE: Writing HTML Body file: sashtml.htm

58 title "outsum1";

59 run;

NOTE: There were 7 observations read from the data set WORK.OUTSUM1.

NOTE: PROCEDURE PRINT used (Total process time):

real time 0.27 seconds

cpu time 0.18 seconds