```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
NOTE: ODS statements in the SAS Studio environment may disable some output features.
69
70         ************************************************************************
71         *Problem 1: PROC UNIVARIATE
72         Read the Data set school 2 final.csv'.Name it as  s2f.  Then do the following analysis:
73         Use PROC UNIVARIATE to analyze the variables f1, f2, f3, f4. Which variable is normal?
74         Which variable is right-skewed? Which variable is left-skewed?;
75         ************************************************************************;
76         proc import datafile='/home/u63743369/week7/school 2 final.csv'  out=s2f replace;
77         run;

NOTE: Unable to open parameter catalog: SASUSER.PARMS.PARMS.SLIST in update mode. Temporary parameter values will be saved to
WORK.PARMS.PARMS.SLIST.
78          /************************************************************************
79          *   PRODUCT:   SAS
80          *   VERSION:   9.4
81          *   CREATOR:   External File Interface
82          *   DATE:      04MAR24
83          *   DESC:      Generated SAS Datastep Code
84          *   TEMPLATE SOURCE:  (None Specified.)
85          ************************************************************************/
86             data WORK.S2F    ;
87             %let _EFIERR_ = 0; /* set the ERROR detection macro variable */
88             infile '/home/u63743369/week7/school 2 final.csv' delimiter = ',' MISSOVER DSD lrecl=32767 firstobs=2 ;
89                informat ClassID best32. ;
90                informat ChildID best32. ;
91                informat Gender $6. ;
92                informat ClassAge $7. ;
93                informat f1 best32. ;
94                informat f2 best32. ;
95                informat f3 best32. ;
96                informat f4 best32. ;
97                format ClassID best12. ;
98                format ChildID best12. ;
99                format Gender $6. ;
100               format ClassAge $7. ;
101               format f1 best12. ;
102               format f2 best12. ;
103               format f3 best12. ;
104               format f4 best12. ;
105            input
106                        ClassID
107                        ChildID
108                        Gender  $
109                        ClassAge  $
110                        f1
111                        f2
112                        f3
113                        f4
114            ;
115            if _ERROR_ then call symputx('_EFIERR_',1);  /* set ERROR detection macro variable */
116            run;

NOTE: The infile '/home/u63743369/week7/school 2 final.csv' is:
      Filename=/home/u63743369/week7/school 2 final.csv,
      Owner Name=u63743369,Group Name=oda,
      Access Permission=-rw-r--r--,
      Last Modified=01Mar2024:00:25:34,
      File Size (bytes)=3348

NOTE: 87 records were read from the infile '/home/u63743369/week7/school 2 final.csv'.
      The minimum record length was 35.
      The maximum record length was 37.
NOTE: The data set WORK.S2F has 87 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              9285.46k
      OS Memory           32796.00k
      Timestamp           03/04/2024 04:23:30 AM
      Step Count                        110  Switch Count  2
      Page Faults                       0
      Page Reclaims                     164
      Page Swaps                        0
      Voluntary Context Switches        12
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           272
```

```
87 rows created in WORK.S2F from /home/u63743369/week7/school 2 final.csv.



NOTE: WORK.S2F data set was successfully created.
NOTE: The data set WORK.S2F has 87 observations and 8 variables.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.06 seconds
      user cpu time       0.04 seconds
      system cpu time     0.01 seconds
      memory              9285.46k
      OS Memory           33056.00k
      Timestamp           03/04/2024 04:23:30 AM
      Step Count                        110  Switch Count  10
      Page Faults                       0
      Page Reclaims                     2478
      Page Swaps                        0
      Voluntary Context Switches        85
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           320


117
118        proc univariate data=s2f;
119           var f1 f2 f3 f4;
120           histogram / normal;
121        run;

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.47 seconds
      user cpu time       0.34 seconds
      system cpu time     0.03 seconds
      memory              14737.75k
      OS Memory           38916.00k
      Timestamp           03/04/2024 04:23:30 AM
      Step Count                        111  Switch Count  0
      Page Faults                       0
      Page Reclaims                     4043
      Page Swaps                        0
      Voluntary Context Switches        2988
      Involuntary Context Switches      2
      Block Input Operations            0
      Block Output Operations           1560


122
123        ************************************************************************
124        Variable f1 appears to follow a normal distribution.
125        Variable f2 is left-skewed.
126        Variables f3 and f4 do not appear to follow a normal distribution;
127        ************************************************************************;
128
129
130
131        ************************************************************************
132        2.Use PROC UNIVARIATE with option plot to graph the histogram, box-plot,
133        and Normal Probability Plot for the variable f3. And only print the graphs.;
134        ************************************************************************;
135
136        ODS SELECT Histogram BoxPlot QQPlot PLOTS;
137        proc univariate data=s2f PLOT;
138            var f3;
139        run;

WARNING: Output 'QQPlot' was not created.  Make sure that the output object name, label, or path is spelled correctly.  Also,
         verify that the appropriate procedure options are used to produce the requested output object.  For example, verify that
         the NOPRINT option is not used.
WARNING: Output 'BoxPlot' was not created.  Make sure that the output object name, label, or path is spelled correctly.  Also,
         verify that the appropriate procedure options are used to produce the requested output object.  For example, verify that
         the NOPRINT option is not used.
WARNING: Output 'Histogram' was not created.  Make sure that the output object name, label, or path is spelled correctly.  Also,
         verify that the appropriate procedure options are used to produce the requested output object.  For example, verify that
         the NOPRINT option is not used.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.08 seconds
      user cpu time       0.03 seconds
      system cpu time     0.01 seconds
      memory              3370.12k
      OS Memory           38748.00k
```

```
      Timestamp           03/04/2024 04:23:30 AM
      Step Count                        112  Switch Count  0
      Page Faults                       0
      Page Reclaims                     425
      Page Swaps                        0
      Voluntary Context Switches        214
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           400


140
141
142       ************************************************************************
143       3.Add a new variable difference defined as difference=f4-f2. Analyze
144       the variable difference and graph its histogram, Boxplot, qqplot.
145        Print only the TestsNormality and TestsForLocation. Does the mean of
146        the variable difference equal to 0 statistically?;
147       ************************************************************************;
148
149       data s2f;
150           set s2f;
151           difference = f4 - f2;
152       run;

NOTE: There were 87 observations read from the data set WORK.S2F.
NOTE: The data set WORK.S2F has 87 observations and 9 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              963.03k
      OS Memory           37804.00k
      Timestamp           03/04/2024 04:23:30 AM
      Step Count                        113  Switch Count  2
      Page Faults                       0
      Page Reclaims                     130
      Page Swaps                        0
      Voluntary Context Switches        13
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


153       ODS SELECT PLOTS;
154
155       proc univariate data=s2f plot;
156           var difference;
157           ODS SELECT TestsNormality TestsForLocation;
158       run;

WARNING: Output 'TestsNormality' was not created.  Make sure that the output object name, label, or path is spelled correctly.
         Also, verify that the appropriate procedure options are used to produce the requested output object.  For example, verify
         that the NOPRINT option is not used.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.08 seconds
      user cpu time       0.04 seconds
      system cpu time     0.01 seconds
      memory              3244.90k
      OS Memory           38748.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        114  Switch Count  1
      Page Faults                       0
      Page Reclaims                     371
      Page Swaps                        0
      Voluntary Context Switches        221
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           400


159
160       ************************************************************
161       *the mean of the variable "difference" is not equal to 0.;
162       ************************************************************
163
164       ************************************************************************
165       4.Use PROC UNIVARIATE to analyze the variable difference by Gender.
166       ************************************************************************;
167
168       ods select BasicMeasures;
169       proc univariate data=s2f;
```

```
170           class Gender;
171       run;

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.10 seconds
      user cpu time       0.10 seconds
      system cpu time     0.00 seconds
      memory              956.28k
      OS Memory           37544.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        115  Switch Count  0
      Page Faults                       0
      Page Reclaims                     57
      Page Swaps                        0
      Voluntary Context Switches        1
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           32


172
173       ************************************************************************
174       5.For the variable difference, calculate custom percentiles from 5 to
175        100 by 5 and export these percentiles to an xlsx file named percentiles.xlsx.
176       ************************************************************************;
177
178       proc univariate data=s2f noprint;
179           var difference;
180           output out=percentiles pctlpre=P_ pctlpts=5 to 100 by 5;
181       run;

NOTE: The data set WORK.PERCENTILES has 1 observations and 20 variables.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              899.09k
      OS Memory           37804.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        116  Switch Count  2
      Page Faults                       0
      Page Reclaims                     116
      Page Swaps                        0
      Voluntary Context Switches        13
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


182       proc print data=percentiles;
183       run;

NOTE: There were 1 observations read from the data set WORK.PERCENTILES.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              750.25k
      OS Memory           37544.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        117  Switch Count  0
      Page Faults                       0
      Page Reclaims                     65
      Page Swaps                        0
      Voluntary Context Switches        1
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           24


184       /* Export percentiles to an xlsx file */
185       proc export data=percentiles
186           dbms=xlsx
187           outfile='/home/u63743369/week7/percentiles.xlsx'
188           replace;
189       run;

NOTE: The export data set has 1 observations and 20 variables.
NOTE: "/home/u63743369/week7/percentiles.xlsx" file was successfully created.
NOTE: PROCEDURE EXPORT used (Total process time):
      real time           0.01 seconds
      user cpu time       0.01 seconds
```

```
      system cpu time     0.00 seconds
      memory              3446.31k
      OS Memory           39216.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        118  Switch Count  0
      Page Faults                       0
      Page Reclaims                     482
      Page Swaps                        0
      Voluntary Context Switches        23
      Involuntary Context Switches      0
      Block Input Operations            16
      Block Output Operations           16


190
191       ************************************************************************
192       *Problem 2: PROC TTEST
193       1.Two sections of a statistics course took a standardized final.
194       Random samples were drawn from each section as follows:
195                     Section A: 65, 68, 75, 78, 70
196                     Section B: 50, 59, 71, 80, 65.;
197       ************************************************************************;
198
199       data scores;
200           input section $ score;
201           datalines;

NOTE: The data set WORK.SCORES has 10 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              675.53k
      OS Memory           37544.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        119  Switch Count  2
      Page Faults                       0
      Page Reclaims                     93
      Page Swaps                        0
      Voluntary Context Switches        14
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


212           ;
213       run;
214
215       proc ttest;
216          class section;
217          var score;
218       run;

NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.23 seconds
      user cpu time       0.10 seconds
      system cpu time     0.06 seconds
      memory              9831.31k
      OS Memory           44244.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        120  Switch Count  48
      Page Faults                       0
      Page Reclaims                     25552
      Page Swaps                        0
      Voluntary Context Switches        1010
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           984


219
220       ************************************************
221       Based on the test I did, I didn't find enough proof
222       to say Section A is better than Section B. So, I can't agree with what the professor said.
223       ************************************************;
224
225
226
227       ************************************************************************
228       3.An experiment is conducted to show that blood pressure levels can be
229       consciously reduced in people trained in this program. The blood pressure
230       measurements (in millimeters of mercury) listed in the table below represent
```

```
231        readings before and after biofeedback training of six subjects."
232        Do the data provide enough evidence to indicate that the mean blood
233        pressure level decreases after training? Use α = 0.05.
234        ************************************************************************;
235
236        data blood_pressure;
237            input Before After;
238            datalines;

NOTE: The data set WORK.BLOOD_PRESSURE has 6 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              673.46k
      OS Memory           38064.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        121  Switch Count  2
      Page Faults                       0
      Page Reclaims                     91
      Page Swaps                        0
      Voluntary Context Switches        18
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


245        ;
246        run;
247
248        proc ttest;
249            paired Before*After;
250        run;

NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.27 seconds
      user cpu time       0.13 seconds
      system cpu time     0.04 seconds
      memory              9816.25k
      OS Memory           44244.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        122  Switch Count  30
      Page Faults                       0
      Page Reclaims                     16626
      Page Swaps                        0
      Voluntary Context Switches        954
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           1160


251
252
253        ************************************************************************
254        Problem 3: RPOC FREQ
255        In a study of the television viewing habits of children, a developmental
256        psychologist selects a random sample of 300 first graders - 100 boys
257        and 200 girls. Each child is asked which of the following TV programs they
258        like best: The Lone Ranger, Sesame Street, or The Simpsons. The results are
259         shown in the contingency table below.
260        ************************************************************************;
261
262        data tv_viewing;
263           input gender $ program $ count @@;
264           datalines;

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.
NOTE: The data set WORK.TV_VIEWING has 6 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              677.09k
      OS Memory           38064.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                        123  Switch Count  2
      Page Faults                       0
      Page Reclaims                     90
      Page Swaps                        0
      Voluntary Context Switches        12
      Involuntary Context Switches      0
      Block Input Operations            0
```

```
      Block Output Operations          264


267       ;
268       run;
269
270       proc print data=tv_viewing;
271          title 'TV Viewing Habits of Children';
272       run;

NOTE: There were 6 observations read from the data set WORK.TV_VIEWING.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              732.06k
      OS Memory           38064.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                       124  Switch Count  1
      Page Faults                      0
      Page Reclaims                    64
      Page Swaps                       0
      Voluntary Context Switches       9
      Involuntary Context Switches     0
      Block Input Operations           0
      Block Output Operations          0


273
274       proc freq data=tv_viewing;
275          tables gender * program / chisq;
276          title 'Chi-Square Test for Independence';
277       run;

NOTE: There were 6 observations read from the data set WORK.TV_VIEWING.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.02 seconds
      user cpu time       0.02 seconds
      system cpu time     0.01 seconds
      memory              1355.25k
      OS Memory           38840.00k
      Timestamp           03/04/2024 04:23:31 AM
      Step Count                       125  Switch Count  5
      Page Faults                      0
      Page Reclaims                    213
      Page Swaps                       0
      Voluntary Context Switches       37
      Involuntary Context Switches     0
      Block Input Operations           0
      Block Output Operations          528


278
279       ********************************************************************************************************
280       it appears that there are no significant differences in TV program preferences between boys and girls;
281       ********************************************************************************************************;
282
283
284
285
286
287
288       OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
298
```