UNITED STATES DEPARTMENT OF THE INTERIOR

GEOLOGICAL SURVEY

USER'S MANUAL FOR BOX_DIAGRAM

By Carl A. Carlson

U.S. Geological Survey
Open-File Report 82-624
1982

Table of Contents

List of Figures

INTRODUCTION

This is the user's manual for the program box_diagram. This program plots modified box diagrams from data contained in STATPAC files. See Figure 1 for an example of the box diagrams produced by this program. Box diagrams were devised by Tukey (1977), extended by McGill and others (1978), adapted for geochemical data by Rose and others (1979, p. 32), and further discussed for earth science data by Kleiner and Graedel(1980). They can summarize large amounts of data in an easily understood manner. See Carlson (1982) for an example of their use in summarizing geochemical data. The STATPAC system has been described by VanTrump and Miesch (1977). This section describes the features of rassmp. The next section gives step by step procedures for using box_diagram. The following section contains FORTRAN and BASIC listings of the program. A bibliography completes this manual.

Box_diagram has these features:

- Runs on the U.S. Geological Survey Multics computer in Menlo Park, Calif.
- Plots box diagrams on a Zeta 6300s off-line plotting system (Zeta Research, 1977, 1978).
- Plots a box diagram for every chosen column in a STATPAC file. Columns are specified by including the desired column numbers in a column number file. See Figure 2.
- Gives the user the option of using qualified values (VanTrump and Miesch, 1977) to compute statistics according to the method of Cohen (1959) and as discussed for geological data by Miesch (1967). It will plot a horizontal recurved line on the box diagram at the point of censurship.
- Gives the user the option of using percentiles, arithmetic statistics, or geometric statistics in construction of the box diagrams.
- If the user chooses geometric statistics, gives the user the option of estimating the population mean with the method of Sichel (1952) as discussed for geological data by Miesch (1967).
- Gives the user the option of using either the inner range (the central 65%) or the outer range (the central 95%) for the upper and lower limits of the box.
- Plots a title on the finished diagram.

Figure 1. Box diagrams produced by this program.

Figure 2. Using qedx to produce a column number file.

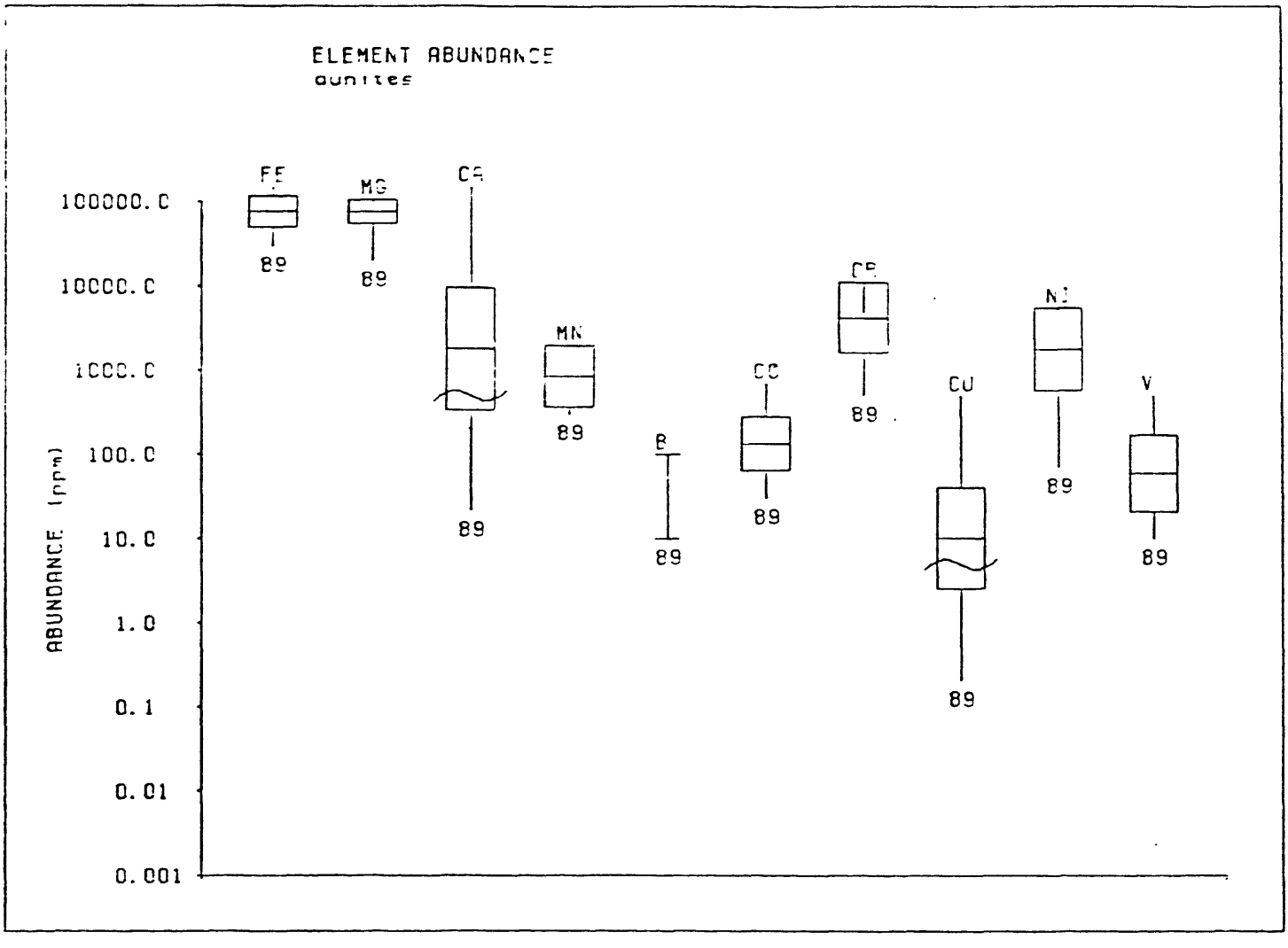

Figure 1. Box diagrams produced by this program.

```
e x
e
1
2
3
25
23
30
\f
w cols <-- cols is the name of the column number file
q
r 1159.0 $0.12 $14.15 0.279 30.530
```

Figure 2. Using qedx to produce a column number file.

OPERATING INSTRUCTIONS

Use the following procedures to use boxdiagram. Follow each typed each typed entry with a carriage return.

Step 1. Prepare the STATPAC File:

If your data is in a RASS file, it must be converted to a STATPAC file using the RASS-STATPAC program b860. The use of b360 is documented in its source code. Figure 3 presents an example of the necessary control file and Figure 4 presents an example of the execution procedure.

If your data is not in the computer, you can enter it from the terminal using the STATPAC program entrys. There is an unpublished document describing entrys avaliable from the STATPAC documentation file on the U.S. Geological Survey Multics computer in Reston, Virginia. Figure 5 illustrates the use of entrys to enter a small data set into a new STATPAC file.

If your data is already in a STATPAC file, you may not need to change the file at all. If you do want to alter the file, you can use the STATPAC programs lookst and edstat to examine the data and then change it. Both lookst and edstat are self documenting. Figures 6 and 7 illustrate the use of the two programs.

The program box_diagram will plot the chemical symbol above the box diagram as was shown in Figure 1, e.g. Cu. This symbol is read from the column header contained in the STATPAC file. To plot the correct chemical symbol, each column header must contain the element symbol preceeded by a dash, e.g. -Cu. The data is assumed to be in parts per million. If the data is in percent, the symbol must be immediately followed by a %-sign, e.g. -Ca%. Many statpac files already follow these conventions. If it is necessary to alter the column headers, use edstat as illustrated in Figure 8.

Step 2. Prepare a Column Number File:

You must tell box_diagram the column numbers of those STATPAC columns which you want plotted. To do this, create a column number file using a Multic editor as illustrated in Figure 2. The program box_diagram will plot box diagrams for each column from left to right in the order given in the column number file. If insuficient information is available in the column to plot a box diagram, a space will be left in the plot where the box diagram would have been plotted.

Step 3. Running box_diagram:

You must run box_diagram while connected to a Zeta 6300s offline plotting system with the A39 asynchronous communication option installed. Use the following login proceedure:

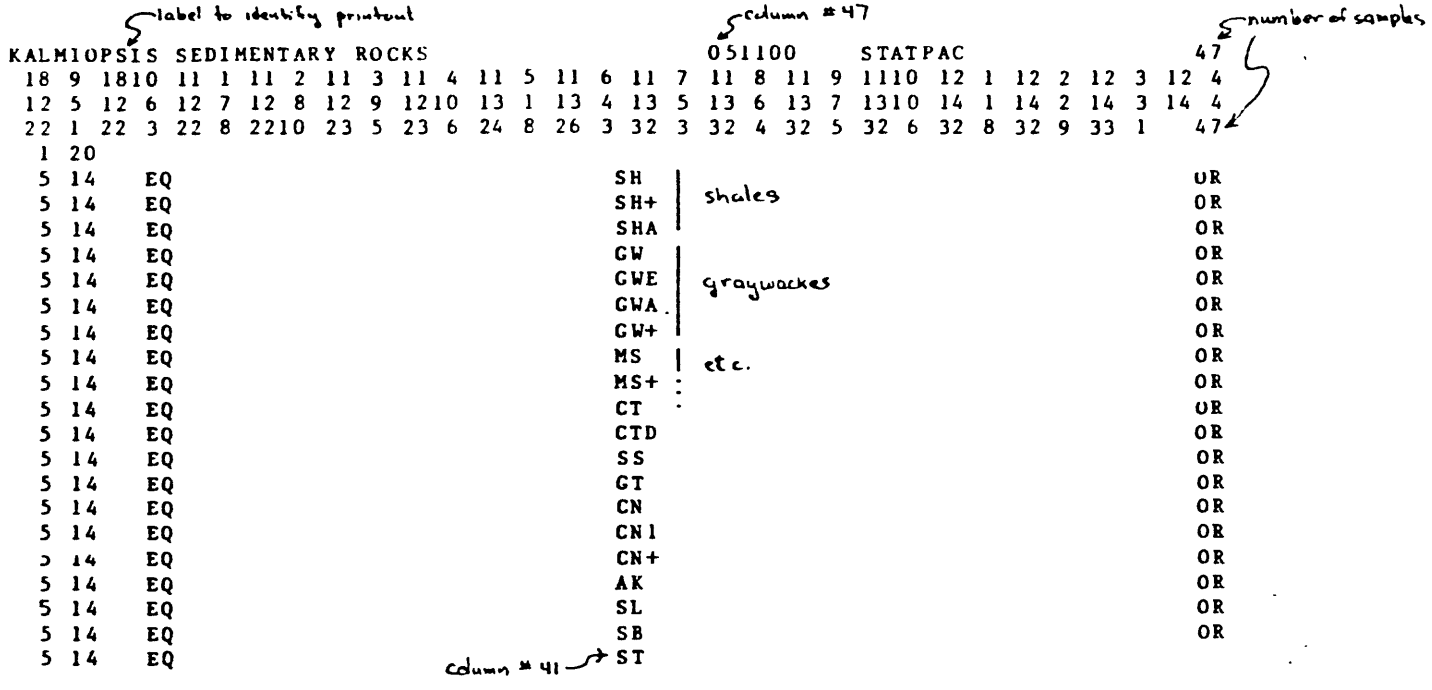

Figure 3. An example of a b860 control file.

```
ec >udd>STATPAC>gvtlib>object>pgmproc.menlo.statpac.ec b860 y n o
Do you want printer output on TTY     ?    no

What is program control file name     ?    <scc>kalm.sedrx.b860

What is input binary file name        ?    kalm.rx.u4 ← RASS file pathname

What is 2nd output binary file name   ?    kalm.sedrx.sp ← new STATPAC pathname
KALMIOPSIS SEDIMENTARY ROCKS
There are   281 samples that satisfy the test criteria.

Finished reading cards.


STOP

Do you want your printer file queued ?    no
r $10.94 $13.70
ls -ft 3
Segments = 42, Lengths = 1314.

r w    28  file11
r w    17  kalm.sedrx.sp
r w     2  b860.!BBBJLggCJmnmjB.list

Multisegment-files = 3, Lengths = 1026.

r w   447  med.sky.kalm
r w   291  medford.u21
r w   288  kalm.sacrosanct

r $0.07 $14.09
dl file11
r $0.03 $14.16
dl b860.!*.list
r $0.06 $14.22
```

Figure 4. An example of the b860 execution procedure.

```
r 1120.4 $0.08 $5.20 0.207 12.648
entrys


ENTER FILE ID -
newfile.sp <--                    multics path name
ENTER DATA SET ID (MAX 8 CHAR) -
STATPAC
ENTER NUMBER OF ROWS -
3
ENTER NUMBER OF COLUMNS -
3
DO YOU WANT TO ENTER COLUMNS IDS ?
yes
ENTER ID FOR COLUMN  1 -
-CU
ENTER ID FOR COLUMN  2 -
-N:
ENTER ID FOR COLUMN  3 -
-FE:
DO YOU WANT TO ENTER ROW IDS ?
yes
WILL YOU HAVE ANY QUAL CODES IN THE DATA SET ?
yes

ROW   1
ENTER ID FOR ROW     1 -
001CAC81
DO YOU HAVE ANY QUAL CODES IN THIS ROW ?
yes
ENTER DATA VALUE CR QUAL CODE CR UNTIL NEXT MESSAGE

1.0
L <-- qualifier code of L
5.0

7.0 <-- blank lines for unqualified data
 <--

ROW   2
ENTER ID FOR ROW     2 -
002CAC81
DO YOU HAVE ANY QUAL CODES IN THIS ROW ?
yes
ENTER DATA VALUE CR QUAL CODE CR UNTIL NEXT MESSAGE

0
B
.4

12.0


ROW   3
ENTER ID FOR ROW     3 -
003CAC81
DO YOU HAVE ANY QUAL CODES IN THIS ROW ?
yes
ENTER DATA VALUE CR QUAL CODE CR UNTIL NEXT MESSAGE

4.0

8.0

15.0

DO YOU WISE TO CHANGE ANY DATA ELEMENT ?
no

ANOTHER SET OF DATA FOR THIS FILE ?
no

ANOTHER SET OF DATA FOR A DIFFERENT FILE ?
no

STOP
r 1124.1 $0.73 $5.94 1.427 14.074
```

Figure 5. Using entrys to input data from the keyboard.

```
asl >xml>statpac>object
r 1127.6 $0.04 $6.74 0.172 15.517
lookst

Enter Input Statpac File Name = newfile.sp

How many data sets in this file do you want to skip ? 0

Data Set ID   = -STATPAC -
No of Rows    =      3
No of Columns =      3

Do you want to print Input Column IDs ? yes
  -CU           -NI           -FE

Do you want to print the data for this data set ? yes

Do you want selected rows ? no

Do you want selected columns ? yes

How many columns pairs(-1 = RowID, 0 = Lat/Long) ? 1

Column selector pairs :
From-To (xxx,xxx) ? 1,2

Do you want to print the selected column IDs ? yes

Selected Column IDs :
  -CU           -NI

DATA:

Row No =    1  Row ID = 001CAC81          Location =   0*  0'  0"    0*  0'  0"
  1.00000E+00   5.00000E+00

Row No =    2  Row ID = 002CAC81          Location =   0*  0'  0"    0*  0'  0"
  0.00000E+00   4.00000E-01

Row No =    3  Row ID = 003CAC81          Location =   0*  0'  0"    0*  0'  0"
  4.00000E+00   8.00000E+00

What do you want next ?
      next  -  read next DS
      skip  -  skip DSs from here & list
      new   -  list DS in a new file
      stop  -  exit from pgm

stop

STOP
r 1129.1 $0.24 $6.98 0.488 16.005
```

Figure 6. Using lookst to review a STATPAC file.

```
asr >uml>statpac>object -after working_dir
r 1144.6 $0.09 $9.92 0.187 23.040
edstat

Enter input filename(return=>newfile): newfile.sp

Enter input dataset name: STATPAC

COMMAND: h
The following commands are entered after the prompt
message COMMAND: (either number or alphabetic codes)
the codes & corresponding operations are:
 0. ("q")  - quit and close files
 1. ("rh") - replace selected column header names
 2. ("rf") - replace selected field numbers (sample ids)
 3. ("rd") - change dataset name and/or column count
 4. ("lh") - print header information
 5. ("l")  - (or "ls")print selected data (by rows/columns)
 6. ("i")  - insert selected data (by rows)
 7. ("d")  - delete selected data (by rows)
 8. ("r")  - replace selected data (by rows)
10. ("nd") - find new dataset in file
11. ("tr") - transformation generator
12. ("s")  - save data on output file (also "wr")
13. ("lr") - list row ids & row numbers
14. ("rv") - replace data with new value if criteria are true
15. ("fr") - find row numbers of specified row ids
16. ("bs") - bastat (basic statistics)
17. ("nf") - read new file
18. ("cp") - copy subset of current file to a new file
 help => print brief help message.
 help long => print help file.(>uml>statpac>doc>helped.info)

COMMAND: r
SELECT ROWS: 2



replace data in row     2   (002CAC81          )
data: 5.0,0B,7.0

COMMAND: s
SELECT OPERATION(a,r,n,q,help): n
Enter new output dataset id:STATPAC
Enter new output filename:  newfile.edited.sp
EDSTAT has written      3 rows &    3 columns
in dataset -STATPAC - on output file newfile.edited.sp

COMMAND: q
 leaving edstat

STOP
r 1146.7 $0.47 $10.39 1.131 24.171
ls -ft 1

Segments = 28, Lengths = 81.

r w     1  newfile.edited.sp

r 1146.9 $0.19 $10.58 0.273 24.444
```

Figure 7. Using edstat to modify a STATPAC file.

```
asr >uml>statpac>object -after working_dir
r 1155.6 $0.17 $13.08 0.207 28.471
edstat

Enter input filename(return=>newfile): newfile.edited.sp

Enter input dataset name: STATPAC

COMMAND: rh
SELECT COLS: 2



Enter line:
-CA%

COMMAND: lh



 for dataset STATPAC , the number of rows & columns =      3     3

do you wish to print the column ids?(y/n)  y
                          column ids
 -CU      -CA%      -FE%

COMMAND: s
SELECT OPERATION(a,r,n,q,help): n
Enter new output dataset id:STATPAC
Enter new output filename:  newfile.edit2.sp
EDSTAT has written      3 rows &    3 columns
in dataset -STATPAC - on output file newfile.edit2.sp

COMMAND: q
 leaving edstat

STOP
r 1157.4 $0.64 $13.72 1.239 29.710
ls -ft 1

Segments = 29, Lengths = 82.

r w     1  newfile.edit2.sp

r 1157.5 $0.21 $13.93 0.286 29.996
```

Figure 8. An example of changing STATPAC column headers.

Figure 3. An example of a b860 control file.

Figure 4. An example of the b860 execution procedure.

Figure 5. Using entrys to input data from the keyboard.

Figure 6. Using lookst to review a STATPAC file.

Figure 7. Using edstat to modify a STATPAC file.

Figure 8. An example of changing STATPAC column headers.

Type: COR ON
SCA X.XX Y.YY (scales the x and y directions of the plot)
ROT OFF or ROT ON
COP COM PLO or COP COM TAP
carriage return, line feed

Multics Prints: the login message.

Type: MAP

Then, continue with the standard login procedure. You must precede all upper case letters (as in your personid) with a shift L. This will produce a \ before every letter you entend to be upper case. This is necessary because the zeta is an upper case only terminal. For example, I login this way: L \C\CARLSON. Multics reads this as l CCarlson. The command MAP used above causes this to happen.

Step 4. Type: ASR >UDD>\W\MIN\RES>\C\CARLSON>CACLIB>OBJECT

Step 5. Type: BOX_DIAGRAM

Multics Prints: ENTER INPUT STATPAC FILE NAME

Type: the path name of your STATPAC file.

Multics Prints: ENTER NAME OF THE FILE CONTAINING SELECTED COL. NUMBERS

Type: the path name of your column number file.

Multics Prints: DO YOU WANT TO INCLUDE QUALIFIED DATA (YES OR NO)?

Type: YES or NO. If you type NO, box_diagram will use only the non-qualified data to compute statistics.

Multics Prints: DO YOU WANT TO USE PERCENTILES (YES OR NO)?

Type: YES or NO. If you type YES, go to Step 6, otherwise

Multics Prints: DO YOU WANT A LOG TRANSFORMATION (YES OR NO)?

Step 5A. Type: YES or NO. If you type YES,

Multics Prints: DO YOU WANT TO USE THE GEOMETRIC MEAN (YES OR NO)?

Type: YES or NO. If you type NO, box_diagram will estimate the population mean with the method of Sichel (1952).

Step 6. Multics Prints: DO YOU WANT TO USE THE INNER RANGE (YES OR NO)?

Type: YES or NO. If you type NO, box_diagram will use the outer range.

Multics will now display all your choices and

Multics Prints: ARE THESE CHOICES ACCEPTABLE TO YOU (YES OR NO)?

Type: YES or NO. If you type NO, Multics will repeat the above procedure. If you type YES, there will be a delay while calculations are performed, then

Multics Prints: ENTER TITLE FOR DIAGRAM.

Type: a name to identify the finished plot.

Multics will now plot the box diagrams which you have requested. If you get an error message refering to ENDPL,

Type: START

You will see the data used by the zeta plotter on the zeta terminal screen. You can ignore it.

```
C     ************************************************************
C     *  PROG:  box_diagram  PLOTS A BOX DIAGRAM FROM DATA IN A  *
C     *  STATPAC FILE.  USES A FILE WHICH CONTAINS THE COLUMN    *
C     *  NUMBERS TO USED (ONE NUMBER PER LINE ).  ASSUMES THE    *
C     *  ELEMENT NAMES ARE INCLUDED IN THE COLUMN HEADINGS AND   *
C     *  ARE PRECEEDED BY A -.  CARL A CARLSON  OCT. 14,1981     *
C     ************************************************************
      EXTERNAL asr(DESCRIPTORS)
      EXTERNAL dsr(DESCRIPTORS)
      EXTERNAL dl(DESCRIPTORS)
      EXTERNAL fo(DESCRIPTORS)
      EXTERNAL ro
      CALL fo ("error_junk")
      CALL dsr(">udd>WMinRes>CCarlson>zeta")
      CALL asr(">udd>WMinRes>CCarlson>zeta","-after","working_dir")
      CALL asr(">uml>statpac>object","-after","working_dir")
      CALL ro
      CALL box_diagram_read
      CALL box_diagram_prep
      CALL dl("box_diagram.temp")
      CALL dl("error_junk")
      WRITE(0,1000)"finished"
1000  FORMAT(V)
      END

C     ********************************************************************
C     *  SUBR box_diagram_read:  READS A STATPAC FILE, COMPUTES MEANS    *
C     *  AND DEVIATIONS FOR EACH COLUMN SELECTED IN A FILE OF COLUMN     *
C     *  NUMBERS, AND PRINTS OUTPUT FILE FOR USE BY box_diagarm_plot     *
C     *  CARL A CARLSON OCT. 12, 1981                                    *
C     ********************************************************************
      SUBROUTINE box_diagram_read
      CHARACTER*32 SEGNAM
      CHARACTER*3 ANS
      CHARACTER*32 SPFILE
      CONM=50
      CONN=500
      CRL=.16
      CRH=.84
      ORL=.025
      ORH=.975
      DIMENSION ID(2),KOLID(50,2),IRID(4),LOC(2),X(50),IA(50)
      DIMENSION XVALU(500),ICOLN(50),IQVALU(500),RMEAN(50),RDEV(50,2)
     +,XMIN(50),XMAX(50)
C
C     ********* INPUT PARAMETERS ***************************************
C
5300  ERROR=0.0
      WRITE(0,1000)"enter input statpac file name"
      READ(0,1000)SPFILE
```

```
1000  FORMAT(V)
      WRITE(0,1000)"enter name of file containing selected col. numbers"
      READ(0,1000)SEGNAM
5000  WRITE(0,1000)"do you want to include qualified data (yes or no)?"
      READ(0,1000)ANS
      IF (ANS.EQ."no")QUAL=0.0
      IF (ANS.EQ."yes") QUAL=1.0
      ERROR=0.0
      IF ((ANS.NE."no").AND.(ANS.NE."yes"))ERROR=1.0
      IF (ERROR.EQ.1.0)WRITE(0,1000)"what's that answer again?"
      IF (ERROR.EQ.1.0)GOTO 5000
5040  WRITE(0,1000)"do you want to use percentiles (yes or no)?"
      READ(0,1000)ANS
      IF (ANS.EQ."yes")STATS=2.0
      IF (ANS.EQ."yes")GOTO 5030
      IF (ANS.EQ."no")GOTO 5020
      WRITE(0,1000)"what's that answer again?"
      GOTO 5040
5020  WRITE(0,1000)"do you want a log transformation (yes or no)?"
      READ(0,1000)ANS
      IF (ANS.EQ."yes" ) STATS=1.0
      ERROR=0.0
      IF (ANS.EQ."no") STATS=0.0
      IF ((ANS.NE."yes").AND.(ANS.NE."no"))ERROR=1.0
      IF (ERROR.EQ.1.0) WRITE(0,1000)"what's that answer again?"
      IF (ERROR.EQ.1.0) GOTO 5020
      IF (STATS.NE.1.0)GOTO 5030
5070  WRITE(0,1000)"do you want to use the geometric mean (yes or no)?"
      READ(0,1000)ANS
      IF (ANS.EQ."yes")GEOM=1.0
      IF (ANS.EQ."no")GEOM=0.0
      ERROR=0.0
      IF ((ANS.NE."yes").AND.(ANS.NE."no")) ERROR=1.0
      IF (ERROR.EQ.1.0) WRITE(0,1000)"what's that answer again?"
      IF (ERROR.EQ.1.0) GOTO 5070
5030  WRITE(0,1000)"do you want to use the inner range (yes or no)?"
      READ(0,1000)ANS
      ERROR=0.0
      IF (ANS.EQ."yes")RANGE=1.0
      IF (ANS.EQ."no")RANGE=2.0
      IF ((ANS.NE."yes").AND.(ANS.NE."no"))ERROR=1.0
      IF (ERROR.EQ.1.0)WRITE(0,1000)"what's that answer again?"
      IF(ERROR.EQ.1.0)GOTO 5030
C
C     ************** WRITE OUT USER CHOICES **********************
C
      WRITE(0,1000)"you have made these choices:"
      WRITE(0,1000)"statpac file is ",SPFILE
      WRITE(0,1000)"column number file is ",SEGNAM
      IF (QUAL.EQ.0.0)WRITE(0,1000)"use only unqualified values."
      IF (QUAL.EQ.1.0)WRITE(0,1000)"use all values."
      IF (STATS.EQ.0.0)WRITE(0,1000)"use arithmetic statistics."
      IF (STATS.EQ.1.0)WRITE(0,1000)"use log transformed data."
      IF ((STATS.EQ.1.0).AND.(GEOM.EQ.1.0))WRITE(0,1000)"use geom. mean"
```

```
      IF ((STATS.EQ.1.0).AND.(GEOM.EQ.0.0))WRITE(0,1000)"estimate popula
     +tion mean by adjusting geom. mean."
      IF (STATS.EQ.2.0)WRITE(0,1000)"use percentiles."
      IF (RANGE.EQ.1.0)WRITE(0,1000)"use the inner range."
      IF (RANGE.EQ.2.0)WRITE(0,1000)"use the outer range."
5400  WRITE(0,1000)"are these choices acceptable (yes or no)?"
      READ(0,1000)ANS
      IF(ANS.EQ."yes")goto 400
      IF (ANS.EQ."no") GOTO 5300
      WRITE(0,1000)"what's that answer again?"
      GOTO 5400
C
C     ************ OPEN AND WRITE HEADER TO MY DATA FILE **********
C
400   OPEN(12,FORM="FORMATTED",FILE="box_diagram.temp")
      OPEN(11,FORM="FORMATTED",FILE=SEGNAM)
      OPEN(10,FORM="UNFORMATTED",FILE=SPFILE)
      RX=18.0
      RY=12.0
      WRITE(12,3000)RX,RY
3000  FORMAT(2F5.1)
C
C     ********* READ STATPAC HEADER ********************************
C
      READ(10,END=100)ID,N,M,(KOLID(I,1),KOLID(I,2),I=1,M)
      OPEN(13,FORM="FORMATTED",FILE="box_diagram.temp2")
C
C     ********* TEST FOR ANY ERRORS ***********************************
C
      IF (M.GT.CONM)WRITE(0,1000)"number of columns exceeds ",CONM
      IF (N.GT.CONN)WRITE(0,1000)"number of rows exceeds ",CONN
      IF ((M.GT.CONM).OR.(N.GT.CONN))GOTO 100
C
C     ********* READ SELECTED COLUMN NUMBERS AND WRITE NAMES **********
C
      DO 210 I=1,CONM
         READ(11,1000,END=410) ICOLN(I)
         WRITE(13,4000)KOLID(ICOLN(I),1),KOLID(ICOLN(I),2)
4000     FORMAT(2A4)
210   CONTINUE
      NCOLMS=I
      GOTO 420
410   NCOLMS=I-1
420   CONTINUE
      CLOSE(13)
      CALL box_diagram_name
C
C     ********* LOOP OVER SELECTED COLS ********************************
C
      DO 220 I=1,NCOLMS
         XMIN(I)=1000000
         XMAX(I)=-1000000
```

```
C
C     ********* LOOP OVER ROWS TO READ ALL X AND QUAL VALUES ***********
C
         DO 230 J=1,N
            CALL GETLST(10,NR,IRID,LOC,X,IA,M,$100)
            XVALU(J)=X(ICOLN(I))
            IQVALU(J)=IA(ICOLN(I))
230      CONTINUE
C
C     ********* COMPUTE RMEAN, DEV, AND NUM ***************************
C
         IF (STATS.EQ.2.0) GOTO 250
C
C     ********* USE ARITHMETIC OR LOG STATISTICS **********************
C
            ICOUNT=0
            NCOUNT=0
            SUMX=0
            SUMX2=0
            RLLIMT=0.001
            IF (STATS.EQ.1.0) RLLIMT=ALOG10(RLLIMT)
            XMIN(I)=999999.9
            XMAX(I)=-999999.9
C
C     ********* LOOP OVER ROWS ****************************************
C
            DO 240 J=1,N
               IF(XVALU(J).LE.0.0)GOTO 240
               IF(IQVALU(J).EQ."B")GOTO 240
C
C     ********* COUNTS ************************************************
C
               IF ((IQVALU(J).NE." ").AND.(IQVALU(J).NE."G"))
     +                NCOUNT=NCOUNT+1
               ICOUNT=ICOUNT+1
C
C     ********* LOG TRANSFORM *****************************************
C
               IF (STATS.EQ.1.0) XVALU(J)=ALOG10(XVALU(J))
C
C     ********* LARGEST LOWER LIMIT OF DETECTION **********************
C
               IF ((QUAL.EQ.1.0).AND.(IQVALU(J).NE." ").AND.
     +            (IQVALU(J).NE."G").AND.(RLLIMT.LT.XVALU(J)))
     +            RLLIMT=XVALU(J)
C
C     ********* MAX AND MIN *******************************************
C
               IF (QUAL.EQ.1.0) GOTO 7020
               IF ((QUAL.EQ.0.0).AND.(IQVALU(J).EQ." ")) GOTO 7020
               GOTO 7010
7020              IF(XVALU(J).LT.XMIN(I))XMIN(I)=XVALU(J)
                  IF(XVALU(J).GT.XMAX(I))XMAX(I)=XVALU(J)
```

```
C
C     ********* UNQUALIFIED SUMS ************************************
C
7010          IF ((IQVALU(J).NE." ").AND.(IQVALU(J).NE."G")) GOTO 240
                  SUMX=SUMX+XVALU(J)
                  SUMX2=SUMX2+XVALU(J)*XVALU(J)
240         CONTINUE
C
C     ********* NO VALID X VALUES: PHONY MEANS AND DEVIATIONS ********
C
            IF (ICOUNT.EQ.0)GOTO 5060
            IF ((ICOUNT-NCOUNT-1).LE.0) GOTO 5060
            GOTO 310
5060           RMEAN(I)= .001
               RDEV(I,1)=.001
               RDEV(I,2)=.001
               IF (STATS.EQ.1.0) XMIN(I)=10**XMIN(I)
               IF (STATS.EQ.1.0) XMAX(I)=10**XMAX(I)
               RLLIMT=0.001
               GOTO 260
C
C     ********* COMPUTE MEANS AND DEVIATIONS *************************
C
310         IF (QUAL.EQ.0.0)ICOUNT=ICOUNT-NCOUNT
            IF (QUAL.EQ.0.0)NCOUNT=0
            RMEAN(I)=SUMX/(ICOUNT-NCOUNT)
            DEV=SQRT(SUMX2/(ICOUNT-NCOUNT-1)-RMEAN(I)**2)
            IF ((NCOUNT.NE.0).AND.(QUAL.EQ.1.0))
     +         CALL LAMDA(RMEAN(I),DEV,ICOUNT,NCOUNT,RLLIMT)
            IF(RMEAN(I).EQ.0.001)GOTO 5060
            RDEV(I,1)=RMEAN(I)-DEV*RANGE
            RDEV(I,2)=RMEAN(I)+DEV*RANGE
            IF ((STATS.EQ.1.0).AND.(GEOM.EQ.0.0))CALL TAU(RMEAN(I),
     +      DEV,ICOUNT)
            IF (NCOUNT.EQ.0)GOTO 6000
            IF (QUAL.EQ.0.0) GOTO 6000
C
C       *************** ESTIMATE MINIMUM VALUE ***********************
C
               RMAX=XMAX(I)
               RMIN=RDEV(I,2)-(RMAX-RDEV(I,1))
               IF (RMIN.LT.XMIN(I))  XMIN(I)=RMIN
C
C       *************** RETRANSFORM LOG DATA *************************
C
6000        IF (STATS.EQ.1.0)RDEV(I,1)=10**RDEV(I,1)
            IF (STATS.EQ.1.0)RDEV(I,2)=10**RDEV(I,2)
            IF ((STATS.EQ.1.0).AND.(GEOM.EQ.1.0))RMEAN(I)=10**RMEAN(I)
            IF (STATS.EQ.1.0) RLLIMT=10**RLLIMT
            IF (STATS.EQ.1.0) XMIN(I)=10**XMIN(I)
            IF (STATS.EQ.1.0) XMAX(I)=10**XMAX(I)
            GOTO 260
```

```
C
C     ********* USE PERCENTILES ***************************************
C
250      ICOUNT=0
C
C     ********* LOOP OVER ROWS, COUNT SAMPLES WITHOUT QUAL= B *********
C
         DO 270 J=1,N
            IF (IQVALU(J).EQ."B")GOTO 270
            IF (((IQVALU(J).NE." ").OR.(IQVALU(J).NE."G")).AND.
     +         (QUAL.EQ.0.0))GOTO 270
             ICOUNT=ICOUNT+1
            IF(XVALU(J).LT.XMIN(I))XMIN(I)=XVALU(J)
            IF(XVALU(J).GT.XMAX(I))XMAX(I)=XVALU(J)
270      CONTINUE
C
C     ********* BUBBLE SORT ALL IGNORING QUALIFIERS ******************
C
20       IFLAG=0
         DO 21 J=1,N-1
           IF (XVALU(J).GE.XVALU(J+1)) GOTO 21
              TEMP=XVALU(J)
              XVALU(J)=XVALU(J+1)
              XVALU(J+1)=TEMP
              IQTEMP=IQVALU(J)
              IQVALU(J)=IQVALU(J+1)
              IQVALU(J+1)=IQTEMP
              IFLAG=1
21       CONTINUE
         IF (IFLAG.EQ.1)GOTO 20
C
C      ***************** FIND LARGEST LOWER LIMIT OF DETECTION ******
C
         RLLIMT=-999999.0
         DO 6010 J=1,N
            IF ((IQVALU(J).EQ." ").OR.(IQVALU(J).EQ."G")) GOTO 6010
            IF (RLLIMT.LT.XVALU(J)) RLLIMT=XVALU(J)
6010     CONTINUE
         IF (ICOUNT.EQ.N) RLLIMT=0.001
C
C     ********* ESTIMATE MEAN WITH MEDIAN ETC ************************
C
         IF (ICOUNT.GT.0)GOTO 320
C
C     ********* NO VALID POINTS **************************************
C
         RLLIMT=.001
         RMEAN(I)=.001
         RDEV(I,1)=.001
         RDEV(I,2)=.001
         GOTO 260
C
C     ********* 1 OR MORE VALID POINT ********************************
C
```

```
320       ICT=(ICOUNT-1)/2.0 + 1.5
          IF (ICT.LT.1)ICT=1
          ICT=ICOUNT-ICT+1
          RMEAN(I)=XVALU(ICT)
          RL=CRL
          RH=CRH
          IF (RANGE.EQ.2.0) RL=ORL
          IF (RANGE.EQ.2.0) RH=ORH
          ICT=(ICOUNT-1)*RL + 1.5
          IF (ICT.LT.1)ICT=1
          ICT=ICOUNT-ICT+1
          RDEV(I,1)=XVALU(ICT)
          ICT=(ICOUNT-1)*RH + 1.5
          IF(ICT.LT.1)ICT=1
          ICT=ICOUNT-ICT +1
          RDEV(I,2)=XVALU(ICT)
C
C     ********* WRITE OUTPUT FILE ************************************
C
260       WRITE(12,2000)RMEAN(I),RDEV(I,1),RDEV(I,2),XMIN(I),XMAX(I),
     +                  RLLIMT,ICOUNT
2000      FORMAT(6(F19.9,1X),I3)
          CLOSE(10)
          OPEN(10,FORM="UNFORMATTED",FILE=SPFILE)
          READ(10,END=100)ID,N,M,(KOLID(IK,1),KOLID(IK,2),IK=1,M)
220    CONTINUE
       CLOSE(10)
       CLOSE(11)
       CLOSE(12)
       RETURN
100    WRITE(0,1000)"ERROR ON 10"
       STOP
       END
5 sub "box_diagram_name"
6 rem ************************************************************
7 rem *  SUB box_diagram_name.  BASIC SUB TO EXTRACT CHEMICAL    *
8 rem *  SYMBOLS FROM COLUMN HEADERS                             *
9 rem ************************************************************
10 file #10:"box_diagram.temp2"
20 file #11:"box_diagram.temp3"
30 if end #10 then 260
40    linput #10:d$
50    let a=pos(d$,"-",1)
60    if a<>0 then 90
70       print #11:using "   0"
80       goto 30
90    let a$=seg$(d$,a+1,a+1)
100   let b$=seg$(d$,a+2,a+2)
110   let c$=seg$(d$,a+3,a+3)
120   if a$<>" " then 150
130      print #11:using "   0"
140      goto 30
150   if b$<>"%" then 180
160      print #11:using "<  1",a$
```

```
170       goto 30
180    if b$<>" " then 210
190       print #11:using "<  0",a$
200       goto 30
210    if c$<>"%" then 240
220       print#11:using "<< 1",a$,b$
230       goto 30
240    print #11:using "<< 0",a$,b$
250    goto 30
260 file #10:"junk"
270 file #11:"junk1"
280 call "dl":"box_diagram.temp2"
290 call "rn":"box_diagram.temp3","box_diagram.temp2"
300 subend
C      ****************************************************************
C      *  TAU:  A SUBROUTINE TO ADJUST THE LOG TRANSFORMED MEAN TO    *
C      *  A BETTER ESTIMATE OF THE POPULATION MEAN USING THE METHOD   *
C      *  OF SICHEL (1952) AS MODIFIED BY MIESCH (1967)               *
C      *  BY CACARLSON DEC. 7, 1981                                   *
C      ****************************************************************
       SUBROUTINE TAU(RMEAN,DEV,ICOUNT)
       DIMENSION TABLE(8,4)
       DATA ((TABLE(I,J),I=1,8),J=1,4) /1.0,1.25,1.692,2.31,2.973,
      + 3.739,4.556,5.0,1.0,1.250,1.744,2.450,3.209,4.151,5.226,10.0,
      + 1.0,1.250,1.817,2.590,3.562,4.858,6.433,100.0,1.0,1.250,1.817,
      + 2.590,3.650,5.005,6.661,200.0/
       N=ICOUNT
       X=10**(DEV)
       IF ((X.GE.1.0).AND.(X.LE.7.0)) GOTO 10
C
C      ************  DEVIATION TOO LARGE OR MEANINGLESS ************
C
          RMEAN=.001
          DEV=.001
          RETURN
10     IF (N.GT.200)N=200
       IF (N.GE.5)GOTO 20
C
C      ************  TOO FEW SAMPLES FOR GOOD RESULTS ***********
C
          RMEAN=.001
          DEV=.001
          RETURN
C
C      ************ FIND THE INDICES OF THE INTERPOLATION AREA ***********
C
20     IXLEFT=X
       IXRITE=IXLEFT+1
       IF (IXRITE.EQ.8)IXRITE=7
       INLOW=1
       IF (N.GE.10)INLOW=2
       IF (N.GE.100)INLOW=3
       IF (N.GE.200)INLOW=4
```

```
      INHIGH=INLOW+1
      IF (INHIGH.EQ.5)INHIGH=4
C
C     ************ INTERPOLATE TAU ************

      TAULOW=TABLE(IXLEFT,INLOW)+(X-IXLEFT)*(TABLE(IXRITE,INLOW)
     +       -TABLE(IXLEFT,INLOW))
      TAUHI =TABLE(IXLEFT,INHIGH)+(X-IXLEFT)*(TABLE(IXRITE,
     +       INHIGH)-TABLE(IXLEFT,INHIGH))
      RTAU=TAULOW+(N-TABLE(8,INLOW))*(TAUHI-TAULOW)/(TABLE(8,INHIGH)-
     +     TABLE(8,INLOW))
C
C     ************ ADJUST RMEAN *****************
C
      RMEAN=RTAU*10**RMEAN
1000  FORMAT(V)
      RETURN
      END
C      ***************************************************************
C      *  SUBROUTINE LAMDA:  ADJUST MEAN AND DEVIATION USING THE   *
C      *  METHOD OF COHEN.  WRITTEN BY CCARLSON.  DEC. 4, 1981     *
C      ***************************************************************
      SUBROUTINE lamda(RMEAN,DEV,ICOUNT,NCOUNT,RLLIMT)
      DIMENSION TABLE(3,15)
      DATA ((TABLE(I,J),I=1,3),J=1,15)/
     + .000,.000,.000,.050,.069,.086,.111,.144,.161,
     + .169,.228,.253,.239,.309,.357,.311,.405,.463,.393,.507,.580,
     + .495,.618,.700,.593,.741,.837,.704,.873,.985,.831,1.019,1.146,
     + .979,1.180,1.314,1.140,1.359,1.526,1.329,1.583,1.756,1.554,1.831,
     + 2.000/
      H=FLOAT(NCOUNT)/FLOAT(ICOUNT)
      IF (H.GT.0.7)RMEAN=.001
      IF (H.GT.0.7)RETURN
      IF (H.EQ.0.0) RETURN
      PAR=(DEV/(RMEAN-RLLIMT))**2
C
C     ********** COMPUTE INDICES FOR TABLE ************************
C
      ILOW=INT(H/.05)+1.0
      IHIGH=ILOW+1
      ILEFT=2
      IRIGHT=3
      IF (PAR.LT.0.65) ILEFT=1
      IF (PAR.LT.0.65) IRIGHT=2
C
C     ********** FIND LOWER AND UPPER LAMDA ***********************
C
      RLAML=TABLE(ILEFT,ILOW)+(PAR-(ILEFT-1)*0.65)*(TABLE(IRIGHT,ILOW)-
     +      TABLE(ILEFT,ILOW))/0.65
      RLAMH=TABLE(ILEFT,IHIGH)+(PAR-(ILEFT-1)*0.65)*(TABLE(IRIGHT,
     +      IHIGH)-TABLE(ILEFT,IHIGH))/0.65
C
C     ********** INTERPOLATE TO FIND RIGHT LAMDA ******************
C
```

```
      RLAMDA=RLAML+(H-(ILOW-1)*.05)*(RLAMH-RLAML)/0.05
C
C     ********** COMPUTE ADJUSTED MEAN AND DEVIATION ***************
C
1000   FORMAT(V)
      RMEAN=RMEAN-RLAMDA*(RMEAN-RLLIMT)
      DEV=SQRT(DEV**2+RLAMDA*(RMEAN-RLLIMT)**2)
      RETURN
      END
C       ***************************************************************
C       *  SUBR  box_diagram_prep:  READS DATA FROM A FILE CONTAINING *
C       *  AVERAGES, DEVIATIONS, MIN, MAX, & NO. OF SAMPLES FOR USE   *
C       *  BY box_diagram_plot.  WRITTEN BY CARL A CARLSON ON         *
C       *  OCT. 13, 1981                                              *
C       ***************************************************************
      SUBROUTINE box_diagram_prep
      DIMENSION RMEAN(50),RDEV(50,2),RMIN(50),RMAX(50),IRNAM(50,3),
     +ITITL(40),INAM(2),ICOUNT(50,4),ICT(4),RLLIMT(50)
C
C     ******** OPEN FILES *********
C
      OPEN(10,FORM="FORMATTED",FILE="box_diagram.temp")
      OPEN(11,FORM="FORMATTED",FILE="box_diagram.temp2")
C
C     ******** INPUT TITLE ********
C
      WRITE(0,1000)"enter title for diagram."
1000  FORMAT(V)
      READ(0,100)(ITITL(I),I=1,40)
100   FORMAT(40A1)
C
C     ******** READ DIMENSIONS OF PLOT *********
C
      READ(10,200)RX,RY
200   FORMAT(2F5.1)
C
C     ******** READ ELEMENT NAMES FROM temp2 ********

     READ(11,2000,END=400)(IRNAM(I,1),IRNAM(I,2),IRNAM(I,3),I=1,50)
2000  FORMAT(2A1,1X,I1)
      I=I+1
400   INUM=I-1
500   CLOSE(11)
C
C     ******** READ STATISTICS *********
C
      READ(10,3000,END=600)(RMEAN(I),RDEV(I,1),RDEV(I,2),RMIN(I),RMAX(I)
     +,RLLIMT(I),ICOUNT(I,1),ICOUNT(I,2),ICOUNT(I,3),I=1,50)
3000  FORMAT(6(F19.9,1X),3A1)
600   CLOSE(10)
C
C     ******** COUNT CHARS IN ICOUNT *********
C
```

```
      DO 700 I=1,INUM
         INCT=1
         IF (ICOUNT(I,2).NE." ") INCT=2
         IF (ICOUNT(I,1).NE." ") INCT=3
         ICOUNT(I,4)=INCT
700   CONTINUE
      DO 800 I=1,INUM
        IF (IRNAM(I,3).EQ.0)GOTO 800
            RMEAN(I)=RMEAN(I)*10000.0
            RDEV(I,1)=RDEV(I,1)*10000.0
            RDEV(I,2)=RDEV(I,2)*10000.0
            RMIN(I)=RMIN(I)*10000.0
            RMAX(I)=RMAX(I)*10000.0
            IF (RLLIMT(I).NE.0.001) RLLIMT(I)=RLLIMT(I)*10000.0
800   CONTINUE

C
C     ******** CALL box_diagram_plot ***********
C
      CALL box_diagram_plot(RMEAN,RDEV,RMIN,RMAX,INUM,IRNAM,ITITL,RX,RY
     +ICOUNT,RLLIMT)
      RETURN
      END
C     ****************************************************************
C     *  SUBR  box_diagram_plot:  A FORTRAN SUBROUTINE WHICH PLOTS    *
C     *  ABUNDANCE DATA IN A BAR CHART LIKE FORM ON THE ZETA PLOTTER. *
C     *  WRITTEN BY C.A. CARLSON OCT, 13,1981                         *
C     ****************************************************************
      SUBROUTINE box_diagram_plot(RMEAN,RDEV,RMIN,RMAX,INUM,IRNAM,
     +ITITL,RX,RY,ICOUNT,RLLIMT)
      DIMENSION RMEAN(50),RDEV(50,2),RMIN(50),RMAX(50),IRNAM(50,3),
     +ITITL(40),INAM(2),ICOUNT(50,4),ICT(3),RLLIMT(50),CUT(7,2)
      DATA (CUT(I,1),CUT(I,2), I=1,7)/
     +  -.125,0.0,.125,.125,.25,.125,.75,-.125,.875,-.125
     +  ,1.125,0.0,1.25,.125/
C*************************SCALE DATA
      SY=RY/8
      SX=RX/(INUM*2+1)
      OFFSET=SX/2.0-.23
       DO 15 I=1,INUM
          IF (RMEAN(I).EQ.0.0)GOTO 15
             RMEAN(I)=ALOG10(RMEAN(I))
             RMIN(I)=ALOG10(RMIN(I))
             RMAX(I)=ALOG10(RMAX(I))
             RDEV(I,1)=ALOG10(RDEV(I,1))
             RDEV(I,2)=ALOG10(RDEV(I,2))
             RLLIMT(I)=ALOG10(RLLIMT(I))
15     CONTINUE
C*********************************PLOT AXIS
      CALL PLOTS(53,0,-1)
      CALL AXIS(0.,0.,ITITL,-40,RX,0.0,999.,RX)
      CALL AXIS(0.,0.,15HABUNDANCE (ppm),15,RY,90.,999.,SY)
      CALL PLOT (-3.5,-1.,3)
      CALL PLOT (-3.5,RY+3.5,2)
```

```
      CALL PLOT (RX+1.,RY+3.5,2)
      CALL PLOT (RX+1.,-1.,2)
      CALL PLOT (-3.5,-1.,2)
      CALL PLOT (0.0,0.0,3)
      CALL SYMBOL(-2.5,-.125,.25,9H    0.001,0.,9)
      CALL SYMBOL(-2.5,SY-.125,.25,  8H    0.01,0.,8)
      CALL SYMBOL(-2.5,2*SY-.125,.25,7H    0.1,0.,7)
      CALL SYMBOL(-2.5,3*SY-.125,.25,7H    1.0,0.,7)
      CALL SYMBOL(-2.5,4*SY-.125,.25,7H   10.0,0.,7)
      CALL SYMBOL(-2.5,5*SY-.125,.25,7H  100.0,0.,7)
      CALL SYMBOL(-2.5,6*SY-.125,.25,7H 1000.0,0.,7)
      CALL SYMBOL(-2.5,7*SY-.125,.25,7H10000.0,0.,7)
      CALL SYMBOL( -2.5,8*SY-.125,.25,8H100000.0,0.,8)
C**************************PLOT TITLES
       CALL SYMBOL(-2.5,(RY-4.0)/2.0,.25,15HABUNDANCE (ppm),90.0,15)
       CALL SYMBOL(2.0,2.5+RY,.25,17HELEMENT ABUNDANCE,0.0,17)
       CALL SYMBOL(2.0,2.1+RY,-.25,ITITL,0.0,40)
C**************************PLOT BARS
       DO 20 I=1,INUM
       INAM(1)=IRNAM(I,1)
       INAM(2)=IRNAM(I,2)
       X=((I-1)*2+1)*SX
       IF (RMEAN(I).NE.-3.0) GOTO 40
C ******************* PLOT JUST A LINE
       Y1=(RMIN(I)+3.0)*SY
       Y2=(RMAX(I)+3.0)*SY
       CALL PLOT (X+SX/4.0,Y1,3)
       CALL PLOT (X+3.0*SX/4.0,Y1,2)
       CALL PLOT (X+SX/2.0,Y1,2)
       CALL PLOT (X+SX/2.0,Y2,2)
       CALL PLOT (X+SX/4.0,Y2,2)
       CALL PLOT (X+3.0*SX/4.0,Y2,2)
       RDEV(I,1)=RMIN(I)
       GOTO 50
C ********************** PLOT THE WHOLE BOX
40     Y=(RMEAN(I)+3)*SY
       CALL PLOT(X,Y,3)
       CALL PLOT(X+SX,Y,2)
       CALL PLOT(X+SX,(RDEV(I,2)+3)*SY,2)
       CALL PLOT(X,(RDEV(I,2)+3)*SY,2)
       CALL PLOT(X,(RDEV(I,1)+3)*SY,2)
       CALL PLOT(X+SX,(RDEV(I,1)+3)*SY,2)
       CALL PLOT(X+SX,Y,2)
       CALL PLOT(X+SX/2,(RDEV(I,1)+3)*SY,3)
       CALL PLOT(X+SX/2,(RMIN(I)+3)*SY,2)
       CALL PLOT(X+SX/2,(RDEV(I,2)+3)*SY,3)
       CALL PLOT(X+SX/2,(RMAX(I)+3)*SY,2)
C**************************PLOT CUT OFF SYMBOL ********
       IF (RLLIMT(I).EQ.-3.0) GOTO 50
       Y1=(RLLIMT(I)+3)*SY
       CALL PLOT (X-SX/4.0,Y1-SY/16.0,3)
       DO 30 J=1,7
          CALL PLOT (X+CUT(J,1)*SX,Y1+CUT(J,2)*SY/2.0,2)
30     CONTINUE
```

```
C**************************PLOT ELEMENT NAME
50      IF (RMAX(I).GT.(RDEV(I,2))) RLOFF=RMAX(I)+3
        IF (RMAX(I).LE.(RDEV(I,2))) RLOFF=RDEV(I,2)+3
        CALL SYMBOL(X+OFFSET+.04,(RLOFF*SY+.1),-.25,INAM,0.0,2)
C**************************PLOT NUMBER OF SAMPLES
        CHAROF=0.1
        RLOFF=RMIN(I)+3
        IF (RMIN(I).GT.(RDEV(I,1))) RLOFF=RDEV(I,1)+3
        ICT(1)=ICOUNT(I,1)
        ICT(2)=ICOUNT(I,2)
        ICT(3)=ICOUNT(I,3)
        INC=ICOUNT(I,4)
        IF(INC.EQ.2)CHAROF=0.2
        IF(INC.EQ.1)CHAROF=0.3
3000    FORMAT(F5.2,I4)
        CALL SYMBOL(X+OFFSET-CHAROF,(RLOFF*SY-.45),-.25,ICT,0.0,3)
20      CONTINUE
C****************************CONCLUDE
        CALL PLOT(0.,0.,999)
        RETURN
        END
```

BIBLIOGRAPHY

Carlson, C. A., Page, N. J, Grimes, D. J., and Leinz, R. W., 1982, Map showing geochemical characteristics of rock samples from the Kalmiopsis Wildernes Area, southwestern Oregon: U.S. Geological Survey Miscellaneous Field Studies Map MF 1240-C.

Cohen, A. C., Jr., 1959, Simplified estimates for the normal distribution when samles are singly censored or truncated: Technometrics, v. 1, no. 3, p. 217-237.

Kleiner, B., and Graedel, T. E., 1980, Exploratory Data Analysis in the Geophysical Sciences: Reviews of Geophysics and Space Physics, v. 18, no. 3, p. 699-717.

McGill, Robert, Tukey, J. W., and Larsen, W. A., 1978, Variations of box plots: The American Statistician, v. 32, no. 1, p. 12-16.

Miesch, A. T., 1967, Methods of computation for estimating geochemical abundance: U.S. Geological Survey Professional Paper 574-B, 15 p.

Rose, A. W., Hawkes, H. E., and Webb, J. S., 1979, Geochemistry in mineral exploration (2d ed.): London, Academic Press, 657 p.

Sichel, H. S., 1952, New methods in the statistical evaluation of mine sampling data: Transactions of the Institute of Mining and Metallurgy v. 61, p. 261-388.

Tukey, J. W., 1977, Exploratory data analysis: Reading, Mass, Addison-Wesley.

VanTrump, George, Jr., and Miesch, A. T., 1977, The U.S. Geological Survey RASS-STATPAC System for management and statistical reduction of geochemical data: Computers and Geosciences, v. 3, no. 3, p. 475-488.

Zeta Research Inc., 1977, Insturmentation manual, 3600 Series Incremental Plotters, Serieals F/158: publication no. 43-1-005, Concord, California.

_____1978, User's manual, Model 6300 Off-line plotter controller: publication no. 43-1-025, Concord, Calif.