TITLE PAGE

COMPUTERIZATION OF MEASURES OF DISPERSION

By

ABDULHAKEEM MOHAMMED

PGD/MCS/2003/2004/1124

A PROJECT SUBMITED TO THE DEPARTMENT OF MATHEMATICS/ COMPUTER SCIENCE.

IN PARTIAL FUL FILMENT OF THE REQIUEMENTS

FOR THE AWARD OF A POSTGRADUATE DIPLOMA

IN COMPUTER SCEINCE.

FEDERAL UNIVERSITY OF TECHNOLOGY,

MINNA, NIGER STATE

November 2004

DECLARATION

I **ABDULHAKEEM MOHAMMED** with registration number *PGD/MCS/2003/2004/1124* hereby declare that this project work is carried out by me under the guidance of my supervisor Professor **K.R ADEBOYE**

ABDULHAKEEM MOHAMMED

SIGNATURE

DATE

CERTIFICATION

I hereby certify that this project is the original work of **MR. MOHAMMED ABDULHAKEEM**. PGD/MCS/2003/2004/1124. Submitted for the award of POSTGRADUATE DIPLOMA IN COMPUTER SCIENCE OF FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA.

———————————————   ———————————————

PROFESSOR   K.R ADEBOYE   DATE
PROJECT SUPERVISOR

———————————————   ———————————————

MR L.N EZEAKO   DATE
HEAD OF DEPARTMENT

———————————————   ———————————————

EXTERNAL EXAMINER   DATE

iii

DEDICATION

This project is dedicated to my beloved late father ALH MOHAMMED

SANI AKANBI IBRAHIM, whom up to this moment I am still enjoying the fruit of what he had labour, may Almighty Allah reward him with the best paradise (Amin) .

ACKNOWLEDGEMENT

All praises and thanks are due to Allah (S.W. A) for given me health & strength to carry out this project work

My sincere appreciation goes to my project supervisor **Professor *K.R* ADEBOYE** who despite his tight schedule was patient to read through my project in which necessary correction and advise are given which increase the reliability quality of this project work.

A profound gratitude goes to the entire *IBAS FAMILY* for giving me both financial and moral support in my PGD programmed may Allah (S.W.T) reward and blessed them all (Amin).

Sincere gratitude is due to the H.O.D Maths/Computer science department MR L.N EZEAKO, DR AKINWANDE, DR YOMI AIYESIMI, PGD CORDINATOR DR U.Y ABUBAKAR, PRINCE BADMUS, MR JIYA MOHAMMED, MR HAKIMI and other lecturer of the department.

A special thanks is due to my friend, college, MUSA MOHAMMADU ISA to him I which him breather feature. Also special appreciation is due my friends UMAR IDRIS GANA, MUKHTAR ISA, ABDULLAHI NDAGANA and my course mates SANUSI ABUBAKAR, ALH LAWAL JIBRIN, IBRAHIM AL-AMIN, UMAR NDAGI and my *little pet*

May ALLAH (S.W.T) in His infinite mercy continue to be with us all (Amin)

ABSTRACT

An introduction to statistical data analysis was explained in chapter one

Measures of dispersion, types of dispersion, area of application, advantages
and disadvantages of each types of dispersion are explained in chapter two

Manual computation of each method of dispersion with a life data was use to explained each of the computational processes.

QBASIC programming language was use in written the programmed because of its debugging aids, it can be learnt virtually on every computer and also because it require little memory space

Steps on how to use the software are explained in chapter four of this project work

Conclusion and recommendation are made base on. the result obtained from research work..

TABLE OF CONTENT

CHAPTER ONE

INTRODUCTION

1.1   STATISTICAL DATA ANALYSIS

Statistics can be defined as a branch of science that deal with scientific method of Collecting, organizing, summarizing, presenting and analyzing data as well as drawing valid conclusion and making reasonable decision on the basis of such data.

The word data is virtually used in every day activities. Although the word carries something of an aura of scientific mystique, it's meaning is quite simple. Data can therefore be thought of as the result of observation. Data are collected in virtually every aspect of life; statements given to a police officer or physician during an interview are data. Data are obtained in the course of scientific inquiry; the position of fossils in an archaeological site, the number of interactions between two members of an animal colony during a period of observation or the spectral composition of light emitted by a star.

1. 2 <u>FORMS OF DATA</u>

Data can be divided into quantitative and qualitative data.

1. 2. 1 <u>QUANTITATIVE</u>

Data are said to be quantitative when they can be expressed numerically. Quantitative data can be classified as measurement or ordinal data.

1. 2.1.1 <u>MEASUREMENT (INTERVAL DATA)</u>

When result of scientific observations are expressed as or translated into numerical value, these data are called measurement or scores. Measurement expressed characteristics such as height, time taken for individual athletes to a race, examination score of an individual students e.t.c.

1.2. 1.2 <u>ORDINAL DATA</u>

These usually have to do with personal opinion. It consists of rank, assignment to ordered categories or of sequential information. For instance two people can be asked to rank ten different brands of cars from best to lowest. Also one hundred participant can be asked to rate objects by placing them in one of five ordered categories, where one is very bad and five is very good, the result will resemble a hundred set of ranks with ties, that is, with several objects sharing the same rank.

1.2.2   <u>QUALITATIVE DATA</u>

These are data that cannot be expressed numerically. Statistical techniques cannot be used to analyze such data are called nominal data.

1.2.2.1   <u>NORMINAL DATA</u>

These are phenomena that cannot be translated into numbers. For instance attribute such as gender, blood group, political party affiliation, e.t.c.

1.3   <u>CLASSIFICATION OF DATA (SOURCE OF DATA)</u>

There are two main sources of data collection and these are published and published sources.

1.3.1   <u>PUBLISHED SOURCES</u>

These are data published as information to the public wher summarized. They are data in which general public are allow access to. These forms of data are also known as published national data because they have produce been published by indigenous media organization within the country.

There are also published international sources such as EPIDEMIOLOGICAI AND STATISTICS BOOKS published yearly by World Health Organization.

1.3.2 UNPUBLISHED SOURCES

These kinds of data are usually out of reach of the public unless on official requisition by the body in-charge or in which the organization is answerable to. These involve files of government and non-governmental agencies and departments. Collection of these types of data is usually difficult and time consuming due to its confidential nature.

1.4 AIMS AND OBJECTIVES

1) To improve speed and efficiency in this aspect of data analysis

2) To aid student who want to make further research in this aspect of analysis

3) Since manual computation is prone to error, this will reduce the error, if not eliminate the error completely.

1.5 SCOPE AND LIMITATION

This project is base on computerization of some measure of dispersion.

The process and program for drawing bar chart, pie chart, histogram, ogive, etc by this programming language is complex and advanced. It is therefore beyond the scope of this program.

The program cannot compute range if the data is grouped

CHAPTER TWO

LITERATURE REVIEW

2.1 MASURES OF DISPERSION

Strategic development is the main objective of every society. And for the fact that we are advancing every day so we are also becoming aware of the need to keep good and reliable records for consultation. These data may look arbitrary and can only be meaningful when processed and analyzed. For the analysis of the data to be made, some characteristics and traits of the data must be known such as the central values of the data, its consistency and variability. If these characteristics are inevitable then some statistical measures are designed to give a mathematical and logical aid to the entire process. These are the measures of Central Tendency and Dispersion

Measures of Dispersion may be defined as a statistics signifying the extent of the scatterness of items around a measure of central tendency. It could be interpreted as measuring how far each data value is from the mean (center) of the data. The knowledge of the variability along with its center can help us visualize the shape of a data set as well as its extremes values. It indicates the variation of observation from a finite population or a sample.

2.2 TYPES OF DISPERSION

The following are types of dispersion;

Range

Mean deviation

Variance

Standard deviation

Coefficient of variation

2.3 Range

The range of a quantitative data set is equal to the largest measurement minus the smallest measurement. Range is easy to compute and easy to understand, but it is insensitive measure of data variation when the data set are large. This is because two data sets can have the same range and be vastly different with respect to data variation.

2.3.1 Advantages Of Range

1) It is easy to compute and understand

2) It is easy to interpret

3) It provides single value

2.3.2 Disadvantages Of Range

1) It only considered the two extreme values

2) It does not encourage equal representation

3) The frequencies are not used at all

6

2.3.3 Area Of Application

1) Industries uses the range in plotting *RANGE CHART* for

2) Statistical quality control technique

2.4 Mean Deviation

The mean deviation also called absolute deviation; make use of an absolute value obtained from a given result. It captures variability in a way that is easy to interpret. Possibilities of capturing dispersion by looking at the deviation of every observation from some constant immediately beg the question of selecting an appropriate constant. It makes intuitive sense that the value of our constant should be typical of the value observed in the data. We might therefore select the most frequently observed value, the mode, but if we want our index of variability to reflect the average difference among all observation, our constant should itself reflect all of the observation. On this bases an appropriate statistic is the arithmetic mean.

2.4.1 Advantages Of Mean Deviation

1) It is easy to compute

2) Taken the absolute value reduces the error term, thereby increases the variability

3) It makes use of all the frequencies; thereby giving room for equal representation

4) Using the absolute values of the deviation reduces the error

terms

2.4.2 Disadvantages Of Mean Deviation

1) Interpreting the result can some time be obscure.

2) Its always difficult to interpret the result

3) The result is not always single value.

2.5 VARIANCE

This is the squared deviation that forms the bases of the most

widely use measure of dispersion Summing all square deviation   from the

mean and dividing by one less than the sample size gives the variance.

2.5.1 Advantages Of Variance

(1) It is singled value

(2) Easy to compute

(3) Each squared deviation enter the sum with the same frequency

(4) It satisfies all the four criteria of good statistic

2.5.2   Disadvantages Of Variance

(1) It is not expressed in the same unit of measurement of the observation

(2) Its computation is usually tedious

2.5.3   Area Of Application Of Variance

Manufacturing industries make use of variance to compare variation of a particular batch produced with a standard.

2.6   Standard Deviation

The standard deviation is a number representing the "*spread*" of a set of numbers around their average value. It is calculated from all the individual deviation from the average. It is an important concept in most statistical calculation because it is a precise indicator of the degree of variability within the set of numbers.

2.6.1   Advantages Of Standard Deviation

(1) The standard deviation is more informative measure of dispersion than the variance because the squaring effect as been reduced

(2) The units of standard deviation are the same as those of the individual Values.

(3) The average absolute deviation actually measures the average distance. .

9

2.6.2 Disadvantages Of Standard Deviation

(1) For Large values, computation of variance is time consuming and tedious.

(2) Variables with deferent units of measurement cannot be compared.

2.6.3 AREA OF APPLICATION OF STANDARD DEVIATION

(1) School makes use of it to determine the performance between sexes

(2) The extend to which different batches of product produce by manufacturing industries varies from a stated standard

2.7 COEFFICIENT OF VARIATION

Otherwise known as coefficient of determination, it is used to compare the variability of two or more sets of data. It is denoted by: it is expressed as a percentage.

2.7.1 ADVANTAGES OF COEFFICINT OF VARIATION

(1) It has all the advantages of standard deviation

(2) It can be used to compare data in different units of measurement

3) it can easily be interpreted since it is in percentage.

2.7.2 DISADVANTAGE OF COEFFICINT OF VARIATION

When you have a voluminous data it is tedious.

2.7.3 AREA OF APPLICATON

Industries use it to compare raw material in different unit of measurement

CHAPTER THREE

System Analyses And Design

3.1 Problem Definition

The purpose of this chapter is to find out whether there is need to

change

the existing system. This is done by defining the problem to be solved.

3.4 <u>Analysis Of Existing System</u>

In analyzing the existing,(i.e. measure of dispersion) with manual

computation, an artificial data is used for each computation.

3.4.1 <u>RANGE</u>

Range for ungrouped data

Range $=X_{max} - X_{min}$

Example of range for ungrouped data   3,5,6,7,23,33,2,20,13,34,10

Range $=34 - 2 = 32$

Example of Range for grouped data

0-9,10-19,20-29,30-39,40-49,50-59 has a range of

Range $= 59-0 = 59$

3.4.2 MEAN DEVIATION

The formula\computation for mean deviation of ungrouped and

grouped

data with there examples are given below;

Deviation for ungrouped data $= \Sigma_i \dfrac{|(x-\bar{x})|}{N}$

3.4.2.1 Mean Deviation For Ungrouped Data

TABLE 1.0

x	$x-\bar{x}$	$|(x-\bar{x})|$
5	0	0
7	2	2
8	3	3
3	-2	2
2	-3	3
TOTAL=25	0	10

Mean, $x = \Sigma_i x_i / \Sigma N$ = 25/5 =5

Mean deviation $= \Sigma |(x-\bar{x})|/N$ =10/5 = 2

the formula for calculating mean deviation for grouped data is given

below

13

Mean deviation$=\{\Sigma f | (x-\bar{x}) | \} / \Sigma f$

3.4.2.2 Mean Deviation For Grouped Data

TABLE 2.0

Scores (x_i)	Frequency (f)	fx	$(x-\bar{x})$	$|(x-\bar{x})|$	$f|(x-\bar{x})|$
0	3	0	-2	2	9
1	6	6	-1	1	6
2	2	4	0	0	0
3	3	6	1	1	9
4	3	12	2	2	6
5	1	5	3	3	3
TOTAL	18	36			33

Mean, $x = \Sigma f_i x_i / \Sigma f_i = 36/18 = 2$

Mean deviation $= = \{\Sigma f | (x-\bar{x}) | \} / \Sigma f = 33/18 = 1.83$

14

3.4.3 VARIANCE

The variance (O'^2) is the sum of the squared deviation from the mean

divided by the size of the population, n.

Mathematically as:

$$O'^2 = \Sigma(x-x)^2/n, \text{ for ungrouped data}$$

3.4.3.1 Variance For Ungrouped Data

TABLE 3.0

x_i	$x_i - \bar{x}$	$\Sigma(x_i - \bar{x})^2$
1	-5	25
3	-3	9
5	-1	1
6	0	0
10	4	16
11	5	25
36		76

$$X = \Sigma x_i /n = 3/6 = 6$$

$$\grave{o}^2 = \Sigma(x_i - \bar{x})^2 / n, \qquad \grave{o}^2 = 76/6 = 12.67$$

3.4.3.2 Variance For Grouped Data

$$O'^2 = \Sigma f_i(x-\bar{x})^2 / N$$

TABLE 4.0

Scores (x_i)	Frequency (f)	fx	$(x-\bar{x})$	$(x-\bar{x})^2$	$f(x-\bar{x})$
0	3	0	-2	4	12
1	6	6	-1	1	6
2	2	4	0	0	0
3	3	9	1	1	3
4	3	12	2	4	12
5	1	5	3	9	9
TOTAL	18	36			42

Solution

Mean, $x = \Sigma f_i x_i / N = 36/18 = 2$

16

$$O'^2 = \Sigma f_i(x-\bar{x})^2 / N = 42 / 18 = 2.33$$

3.4.4 <u>STANDARD DEVIATION</u>

The standard deviation (O') of a distribution is the square root of the variance of that distribution.

<u>UNGROUPED DATA</u>

$$O' = \sqrt{\Sigma (x-\bar{x})^2 / n}$$

With reference to example of variance for ungrouped data in table 3.0,

Standard deviation, $O' = \sqrt{variance}$

$$O' = \sqrt{12.67} = 3.56$$

<u>GROUPED DATA</u>

$$O' = \sqrt{\Sigma f_i(x-\bar{x})^2 / f}$$

With reference to example of variance for grouped data in table 4,

$$O' = \sqrt{\Sigma f_i(x-x)^2 / f}$$

$$= \sqrt{2.33}$$

$$= 1.53$$

3.4.5 <u>COEFFICIENT OF VARITION</u>

C. v = $\dfrac{\text{Standard deviation}}{\text{Mean}}$ * 100

FOR UNGROUPED DATA

With reference from table 3.0 and standard deviation for ungrouped

data,   we have the following result

$x = \Sigma x_i / n = 3/6 = 6$

$O' = \sqrt{\Sigma(x-\tilde{x})^2 / n}$ $= \sqrt{12.67} = 3.56$

C. v $= \dfrac{3.56}{6}$ * 100 $= 59.33$

With reference to table 4.0, the mean and standard deviation for

grouped data we have the following result.

Mean, $x = \Sigma f_i x_i / N = 36/18 = 2$

$O'^2 = \Sigma f_i(x-\bar{x})^2 / N = 42/18 = 2.33$

$C v = \dfrac{O'^2}{X}{}^{*100} = \dfrac{2.33}{2} * \dfrac{100}{1}$ $= 116.5$

CHAPTER FOUR

SYSTEM DEVELOPMENT

4.1 Programming Language Used

Programming language used is Q-BASIC (Quick-Beginners All Purpose

Symbolic Instruction Code)

Q-Basic was is because of the following reasons

It is easy to use and understand

Debugging process is easy

With Q-BASIC interpreter, this program can be learnt on virtually every

computer, because it requires little memory space.

.2 Syntax Of The System Developed

The interpreter has to be installed in the hard disk before it can run the

program. If the interpreter is installed, then one can run the program.

After the installation, then go to the background of the interpreter as in

fig 1 below, although the background is usually blue
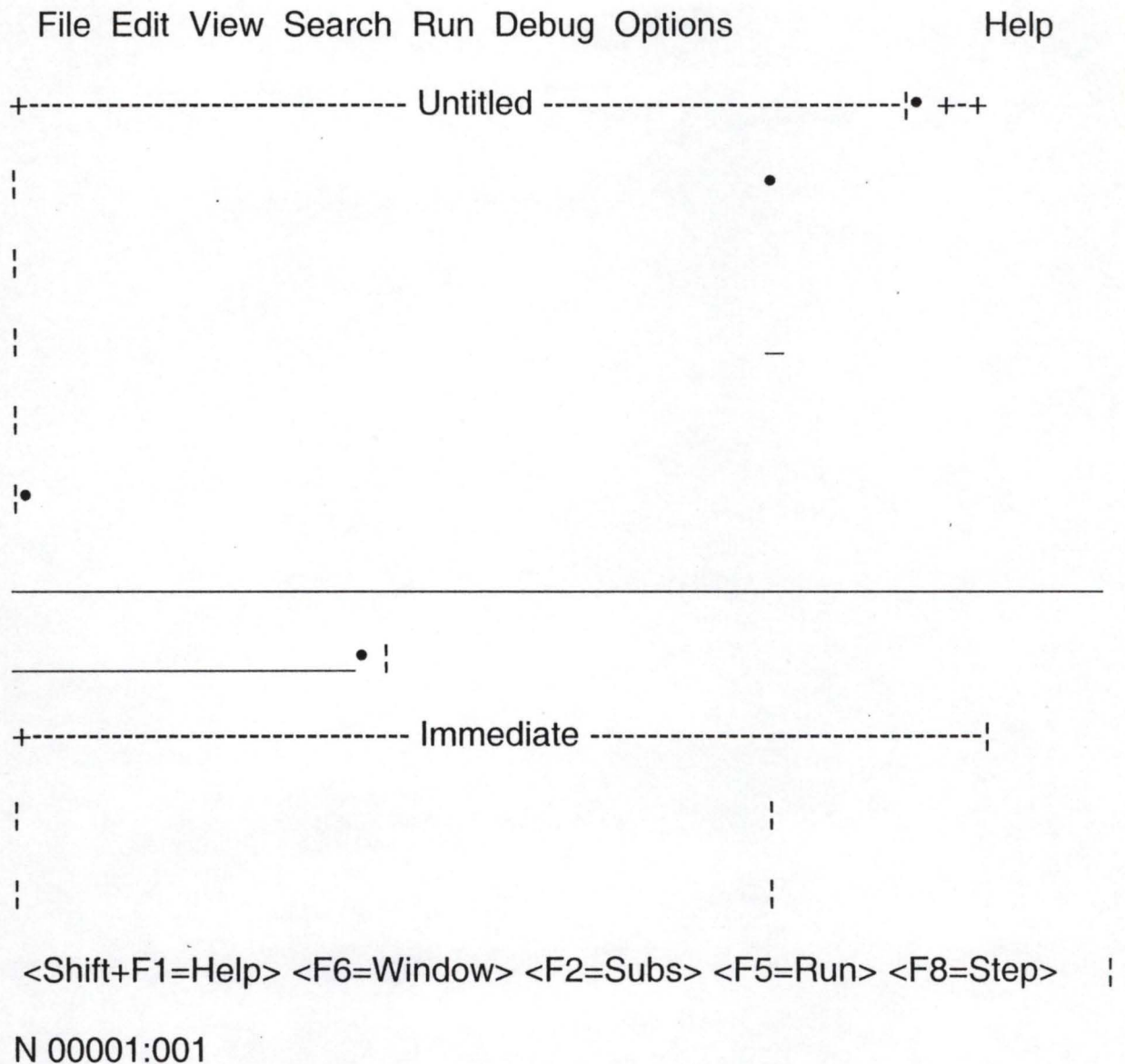
19

Fig1

```
  File  Edit  View  Search  Run  Debug  Options                    Help

+----------------------------- Untitled ----------------------------¦•  +-+

¦                                                          •

¦

¦                                                          _

¦

¦•


_____


_____•  ¦

+----------------------------- Immediate ---------------------------¦

¦                                           ¦

¦                                           ¦

 <Shift+F1=Help> <F6=Window> <F2=Subs> <F5=Run> <F8=Step>      ¦
```

N 00001:001

Click on file menu to bring a sub or pull down menu

Select and double click on " Open…

This takes you to the next background as below

  Subsequently, select [-A-] since the program will be on a floppy

diskette

    Double click [-A-] to take you to the next stage

20

Fig1

File Edit View Search Run Debug Options   Help

+------------------------------- Untitled -----------------------------¦• +-+

¦   •

¦

¦   ─

¦

¦•

_____

_____• ¦

+------------------------------- Immediate -------------------------------¦

¦   ¦

¦   ¦

<Shift+F1=Help> <F6=Window> <F2=Subs> <F5=Run> <F8=Step>   ¦

N 00001:001
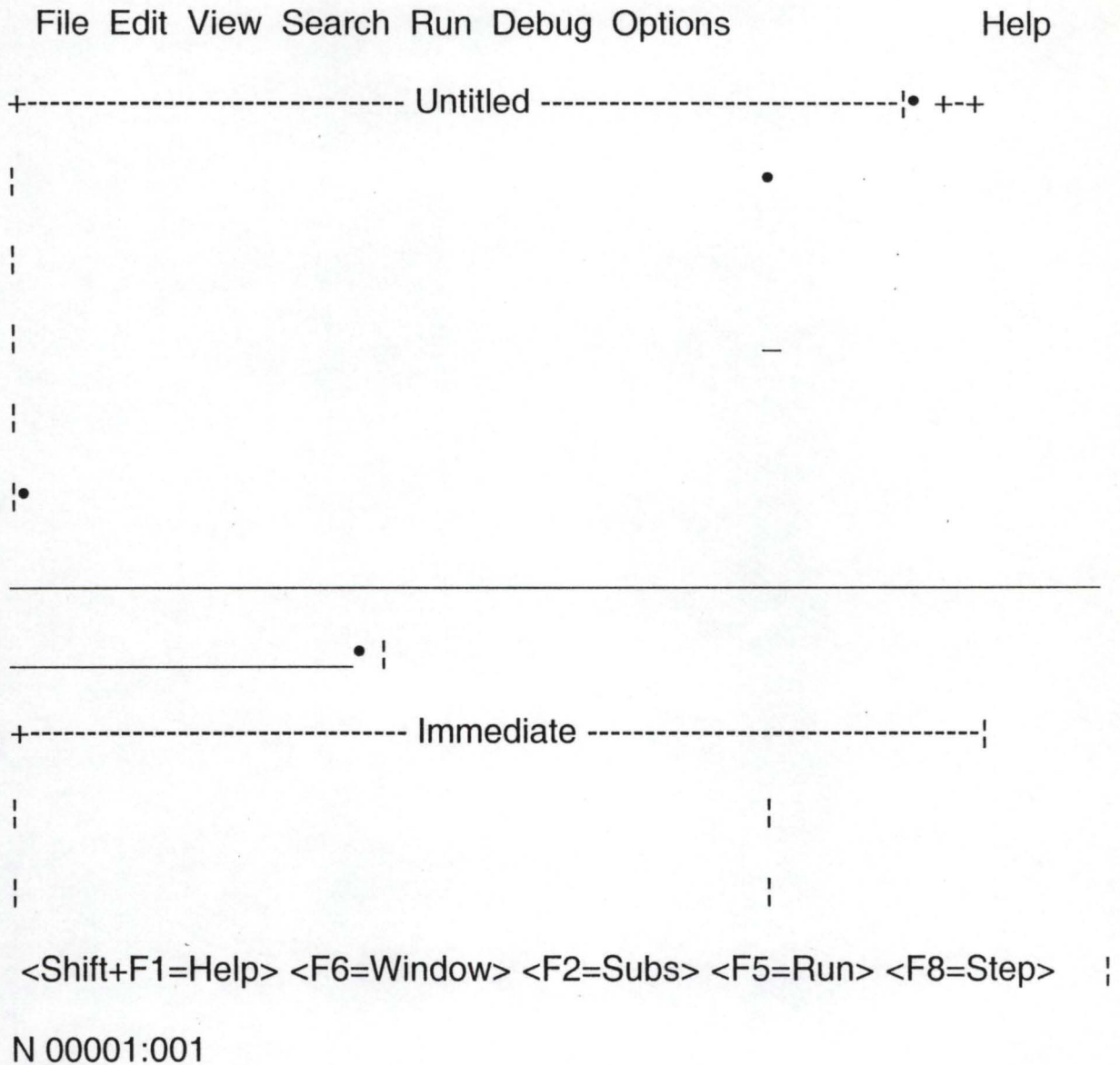
Click on file menu to bring a sub or pull down menu

Select and double click on " Open…

This takes you to the next background as below

  Subsequently, select [-A-] since the program will be on a floppy

diskette

    Double click [-A-] to take you to the next stage

Fig2

```
 File Edit View Search Run Debug Options                    Help
+----------------------------- Untitled ----------------------------¦• +-+
¦    +--------------------------- Open ---------------------------+
¦    ¦ File Name: ¦*.BAS                              ¦ ¦
¦    ¦ C:\MYDOCU~1                                    ¦      _
¦    ¦              Files              Dirs/Drives    ¦
¦    ¦¦ PROF.BAS                       ¦ ¦ ..      • ¦     _
¦    ¦¦                                ¦ ¦ LOGICF~1  ¦     _
¦    ¦¦                                ¦ ¦ MPEGAV   _¦     _
¦    ¦¦                                ¦ ¦ MYMUSI~1 _¦     _
¦    ¦¦                                ¦ ¦ MYPICT~1 _¦     _
¦    ¦¦                                ¦ ¦ QBASIC   _¦     _
¦    ¦¦                                ¦ ¦ [-A-]    _¦     _
¦    ¦¦                                ¦ ¦ [-B-]    • ¦
¦• __¦         < OK >        < Cancel >       < Help >      ¦____• ¦
+----+-----------------------------------------------------------+-----¦
¦.                                              ¦
¦ F1=Help  Enter=Execute  Esc=Cancel  Tab=Next Field  Arrow=Next
Item
```

Fig3

File Edit View Search Run Debug Options   Help

+------------------------------ Untitled ----------------------------¦• +-+

¦   +--------------------------- Open ----------------------------+   •

¦   ¦ File Name: ¦*.BAS   ¦ ¦   _

¦   ¦ A:\   ¦   _

¦   ¦   Files   Dirs/Drives ¦   _

¦   ¦ ¦ PROF.BAS   ¦ ¦ VARIAN~1 • ¦   _

¦   ¦ ¦   ¦ ¦ [-A-]   ¦   _

¦   ¦ ¦   ¦ ¦ [-B-]   _ ¦   _

¦   ¦ ¦   ¦ ¦ [-C-]   _ ¦   _

¦   ¦ ¦   ¦ ¦ [-D-]   _ ¦   _

¦   ¦ ¦   ¦ ¦ [-E-]   _ ¦   _

¦   ¦ ¦   ¦ ¦   • ¦   _

¦   ¦ +• _____ • + +---

¦   +---¦   •

¦• ___¦   < OK >   < Cancel >   < Help >   ¦____• ¦

+-----+---+-----¦

¦   ¦

22

Then double click on 'PROF.BAS' to open the program. Using our mouse click 'Run' and from the pull down menu click Start and thereafter follow subsequent instruction issued by the program

Fig4

File Edit View Search Run Debug Options   Help

+-------------------------------- PROF.BAS --------------------------------¦● +-+

¦REM THIS PROGRAM CALCULATES THE MEAN,VARIANCE

,STANDARD DEVIATION   ●

¦REM COEFFICIENT OF VARIATION

¦REM written by Muhammad Abdulhakeem   _

¦COLOR 7, 4: CLS : COLOR 1, 2: LOCATE 2, 32: PRINT "WELCOME"

¦LOCATE 4, 40: PRINT "TO": LOCATE 6, 35: PRINT "project": LOCATE 8, 40: PRINT "_

¦LOCATE 10, 20: PRINT "computerization of some measures of dispersion"

```
¦LOCATE 12, 35: PRINT "presented by Muhammad Abdulhakeem

(PGD/MCS/2003/2004/1124_

¦LOCATE 24, 26: COLOR 12, 3: PRINT "press any key to continue...":

ans$ = INPUT_

¦MENU:                                                    _

¦COLOR 7, 0: CLS : COLOR 0, 2                             _

¦LOCATE 5, 2: PRINT "DATE:"; DATE$: LOCATE 5, 64: PRINT "TIME:";

TIME$      _

¦COLOR 7, 9: LOCATE 5, 24: PRINT CHR$(201): LOCATE 7, 24:

PRINT CHR$(201)    _

¦FOR i = 25 TO 55                                         _

¦LOCATE 5, i: PRINT CHR$(205): LOCATE 7, i: PRINT CHR$(205):

LOCATE 19, i: PRIN_

¦NEXT i                                                   _

¦LOCATE 5, 56: PRINT CHR$(187): LOCATE 7, 56: PRINT CHR$(187):

LOCATE 19, 24: P_

¦FOR j = 6 TO 18                                          _

¦LOCATE j, 24: PRINT CHR$(186): LOCATE j, 56: PRINT CHR$(186)

+------------------------------- Immediate --------------------------------¦

 <Shift+F1=Help> <F6=Window> <F2=Subs> <F5=Run> <F8=Step>    ¦

N 00001:001
```

Example: using OPTION 1 which is to compute the Mean, Mean deviation, variance, standard deviation, and covariance of data that have already been grouped

From fig4, click: Run

Start

The main menu appears as below, after you type 1 by the side of

Select Option (1-5)? 1

After typing the 1, a command at the bottom of the main menu is there asking you to press any key to continue

```
DATE:10-21-2004        +------------------------------+        TIME:06:22:18

                       ¦     PRINT MAIN MENU          ¦

                       ¦------------------------------¦

                       ¦ 1.Grouped Data               ¦

                       ¦ 2.Grouped Continuous Data    ¦

                       ¦ 3.Grouped Discrete Data

                       ¦ Ungrouped Data

                       ¦ 5.Quit

                       +------------------------------+
```

Select Option (1-5)? 1

Press Any Key To Continue

After pressing any to continue, example press 'Enter' from the key board

and a message appears as thus:

THIS MODULE CALCULATES MEAN, VARIANCE, S.D, C.V, AND

M.D

FOR DATA THAT HAVE ALREADY BEEN GROUPED

At the bottom, an instruction asking you to continue appears. After

pressing any key, the instruction is as thus,

For a data with a class of five groups as below:

Class interval	Frequency
1-9	3
10-19	2
20-29	5
30-39	7
40-49	1

STEP1

Enter the number of classes---→ , type 5

Press Enter

Enter the lower class limit---→, type 1

Press Enter

Enter the upper class limit---→, type 9

Press enter

This completes imputing the data for the first class

Press enter and continue in this manner, and finally the result

is printed on the screen as:

FOR GROUPED DATA

****

S/N	Class Int.	f	x	fx	fx^2	f|x-m|
1	0 - 0	3	0	0	0	72.91666
2	10 - 19	2	14.5	29	420.5	19.61111
3	20 - 29	5	24.5	122.5	3001.25	.9722233
4	30 - 39	7	34.5	241.5	8331.75	71.36111
5	40 - 49	1	44.5	44.5	1980.25	20.19444

18   0   13733.75   185.0555

****

MEAN   = 24.30556   VARIANCE   = 172.2261

STANDARD DEV. =   13.12349   MEAN DEV.   = 10.28086

COEF. OF VAR.   = 53.9938 %

Press Any Key to Continue...

CHAPTER FIVE

SUMMARY/ CONCLUTION AND RECOMMENDATION

5.1   <u>SUMMARY</u>

Computerization can be defined as planned and articulated changes from a manual system to automation using computer system. This project topic; computerization of measure of dispersion was carried out base on the afro mentioned reasons in chapter 1 of this project, some of which are manual computation is prone to error, computation is tedious etc. with a good and reliable program, these computations are carried out yielding reliable and accurate results.

Q BASIC programming language was used due to the fact that one can copy a QBasic compiler on a floppy diskette, install it in his/her computer system with little or no problem. Its compatibility on virtually every computer system has increased its acceptability and usage by programmers. QBasic is easy to learn, understand and it provide debugging aids. lastly, QBasic compiler requires very little space of the hard disk which is as added advantage to its portability.

5.2   CONCLUSION

It is understandable that measure of dispersion indicates spread of

distribution and as such, proper consideration should be taken in

statistical data analysis, this is what necessitate for the project topic:

computerization of some measures of dispersion which were

explained in chapter four of this project work.

Accuracy of every computation is highly stressed and to achieve this,

a   reliable technique should be ensured. A well written program,

tested a

documented program will certainly perform analysis on voluminous
data

for the user.

5.3   RECOMMENDATION

Government and non –government organization should be

encouraged to use computerized method of dispersion to solve any

problem on measure of dispersion.

An individual who to go on research work in which measure of

dispersion will be used should use this software to solve it.

If all the aforementioned recommendation are taken into

consideration, the resulting effect will be speed in computation

process, reduce error   reliability efficiency and easy documentation.

Reference:

Battely I.G Brokes C.J, (1979):   FUNDAMENTAL OF MATHEMATICS
Loxton .S.M   AND STATISTICS
                                   Published By Bro
                                   (Norwich) Ltd.

Benjamin Fruchter (1987):   FUNDAMENTAL STATISTICS IN
Guilford J.P   PSYCHOLOGY AND EDUCATION
                                   Third Edition
                                   Printed By B& Jo Enterprises Pte Ltd

Emuoyibofarhe O . Justice   FUNDAMENTALS OF COMPUTER
Fasanya A. Kola   SCIENCE AND TECHNOLOGY
Sunday A. Reju (2003)   Published by KAY- EM Publishers

Frank J.Kohout (1974)   STATISTICS FOR SOCIAL
SCIENTISTS

                                   Published By John Wiley & Sons, Inc

JIYA MOHAMMED (2004)   INTRODUCTION TO PROGRAMMING
                                   LANGUAGE(QBASIC) UNPUBLISHED
                                   FEDERAL UNIVERSTY OF

    TECHNOLOGY

Ronald J, Guilford (1986):   INTRODUCTORY STATISTICS
Thomas H.Wonnacott   Third Edition

```
 THIS PROGRAM CALCULATES THE MEAN,VARIANCE ,STANDARD DEVIATION
 COEFFICIENT OF VARIATION
 written by Muhammad Abdulhakeem
OR 7, 4: CLS : COLOR 1, 2: LOCATE 2, 32: PRINT "WELCOME"
ATE 4, 40: PRINT "TO": LOCATE 6, 35: PRINT "project": LOCATE 8, 40: PR    "
ATE 10, 20: PRINT "computerization of some measures of dispersion"
ATE 12, 35: PRINT "presented by Muhammad Abdulhakeem (PGD/MCS/2003/200   112
ATE 24, 26: COLOR 12, 3: PRINT "press any key to continue...": ans$ =    UT
:
R 7, 0: CLS : COLOR 0, 2
ATE 5, 2: PRINT "DATE:"; DATE$: LOCATE 5, 64: PRINT "TIME:"; TIME$
R 7, 9: LOCATE 5, 24: PRINT CHR$(201): LOCATE 7, 24: PRINT CHR$(201)
i = 25 TO 55
TE 5, i: PRINT CHR$(205): LOCATE 7, i: PRINT CHR$(205): LOCATE 19, i: PRIN
i
TE 5, 56: PRINT CHR$(187): LOCATE 7, 56: PRINT CHR$(187): LOCATE 19, 24: P
j = 6 TO 18
TE j, 24: PRINT CHR$(186): LOCATE j, 56: PRINT CHR$(186)
j: LOCATE 19, 56: PRINT CHR$(188)
 3, 4: LOCATE 6, 31: PRINT "PRINT MAIN MENU": 'COLOR 7,4:PRINT
E 8, 27: PRINT "1.Grouped Data": LOCATE 10, 27: PRINT "2.Grouped Continou
 12, 27: PRINT "3.Grouped Discrete Data": LOCATE 14, 27: PRINT "Ungroupe
 17, 27: PRINT "5.Quit"
5, 3: LOCATE 21, 27: INPUT "Select Option (1-5)"; opt
 < 1 OR opt > 5 THEN
COLOR 6, 2: LOCATE 12, 25: PRINT "INVALID OPTION,TRY AGAIN": BEEP
 (2): GOTO MENU

 = 1 THEN GOTO GROUP1
 = 2 THEN GOTO group2
 = 3 THEN GOTO GROUP3
 = 4 THEN GOTO UNGROUPED ELSE GOTO QUIT
:
COLOR 1, 2: CLS
E 12, 13: PRINT "THIS MODULE CALCULATES MEAN,VARIANCE,S.D,C.V,&M.D"
E 13, 18: PRINT "FOR DATA THAT HAVE ALREADY BEEN GROUPED"
 12, 3: LOCATE 25, 26: PRINT "Press any key to continue...": ans$ = INPUT

 = 0: s2 = 0: s3 = 0: c = 0: d = 0
2, 1: LOCATE 12, 23: INPUT "Enter the Number of classes--->", n
 l(n), u(n), f(n), x(n)

 = 1 TO n
7, 0: CLS : COLOR 10, 4
 9, 23: INPUT "Enter the Lower class Limit  ---->", l(i)
 11, 23: INPUT "Eenter the Upper class Limit ---->", u(i)
 13, 23: INPUT "Enter the frequency of the class --->", f(i)
(i) = (l(i) + u(i)) / 2: LET s1 = s1 + f(i) * x(i)
 = s2 + f(i) * x(i) ^ 2: LET c = c + f(i)

 = s1 / c: LET var = (s2 / c - m ^ 2): LET sd = SQR(var): LET cvar =  sd
 = 1 TO n
 = d + f(i) * ABS(x(i) - m)
i
dev = d / c: COLOR 7, 1: CLS
TAB(16); "COMPUTED VALUES OF SOME MEASURES OF DISPERSION"
TAB(27); "FOR GROUPED DATA"
TAB(1); STRING$(73, "*")
"S/N"; TAB(8); "Class Int."; TAB(24); "f"; TAB(34); "x"; TAB(44); "  ";
TAB(1); STRING$(73, "-")
 = 1 TO n
i; TAB(7); l(i); "-"; u(i); TAB(23); f(i); TAB(33); x(i); TAB(43);     )

 STRING$(73, "-")
```

```
R 1, 3: LOCATE 12, 17: INPUT "Enter the Number of Data to be compute        "
M x3(n3), f3(n3): COLOR 7, 0: CLS
i = 1 TO n3: COLOR 4, 3: CLS
TE 12, 25: PRINT "Enter the values of x"; i; "--->": LOCATE 12, 55: INPUT
TE 14, 25: PRINT "Enter the values of f"; i; "--->": LOCATE 14, 55: INPUT
k1 = k1 + f3(i) * x3(i): LET k2 = k2 + f3(i) * x3(i) ^ 2: LET z = z + f3(i
 i: max = 0:
i = 1 TO n3
3(i) >= max THEN max = x3(i)
3(i) <= min THEN min = x3(i)
 i
 max - min: LET m = k1 / z: LET var = k2 / z - m ^ 2: LET sd = SQR(var): L
i = 1 TO n3
3 = k3 + f3(i) * ABS(x3(i) - m)
 i
 dev = k3 / z: COLOR 7, 1: CLS : PRINT TAB(15); "COMPUTED VALUES OF SOME M
 TAB(26); "FOR GROUPED DESCRETE DATA": PRINT
 STRING$(76, "*")
 "S/N."; TAB(10); "X"; TAB(25); "f"; TAB(40); "fx"; TAB(55); "fx^2";
 STRING$(76, "-")
 = 1 TO n3
 i; TAB(9); x3(i); TAB(24); f3(i); TAB(39); f3(i) * x3(i); TAB(54);
 i
 STRING$(76, "-")
 TAB(24); z; TAB(39); k1; TAB(54); k2; TAB(68); k3
 STRING$(76, "*"): PRINT
 "MEAN"; TAB(18); "="; m; TAB(48); "VARIANCE"; TAB(62); "="; var: PRINT
 "STANDARD DEV."; TAB(18); "="; sd; TAB(48); "MEAN DEV."; TAB(62);      m
 "COEF. OF VAR."; TAB(18); "="; cvar; "%"; TAB(48); "RANGE"; TAB(62);  "="
 12, 3: LOCATE 25, 26: PRINT "Press Any key To Continue...": ans$ = INPUT

UPED:
 COLOR 1, 2: CLS
E 12, 12: PRINT "THIS MODULE COMPUTES MEAN,VARIANCE,S.D,C.V,M.D & RANGE"
E 13, 30: PRINT "FOR UNGROUPED DATA"
R 12, 3: LOCATE 25, 26: PRINT "Press Any key to Continue..."; ans$ = INPUT
: LET t1 = 0: t2 = 0: t3 = 0
R 1, 3: LOCATE 12, 13: INPUT "Enter the Number of Variables to be Com
 x2(n2): COLOR 7, 0: CLS
i = 1 TO n2
R 4, 3: LOCATE 12, 25: PRINT "Enter the Value of x"; i; "--->"
E 12, 55: INPUT x2(i): LET t1 = t1 + x2(i)
t2 = t2 + x2(i) ^ 2: min = x2(1)
 i: max = 0
i = 1 TO n2
2(i) >= max THEN max = x2(i)
2(i) <= min THEN min = x2(i)
 i
max - min: LET m = t1 / n2: LET var = t2 / n2 - (m) ^ 2
sd = SQR(var): LET cvar = (sd / m) * 100
i = 1 TO n2: LET t3 = t3 + ABS(x2(i) - m): NEXT i: LET mdev = t3 / n2
R 7, 1: CLS : PRINT TAB(13); "COMPUTED VALUES OF SOME MEASURES OF DISPERSI
T TAB(26); "FOR UNGROUPED DATA"
T STRING$(73, "*")
T "S/N."; TAB(15); "Variable(x)"; TAB(45); "x^2"; TAB(65); "|x-m|"
T STRING$(73, "-")
 i = 1 TO n2
NT i; TAB(15); x2(i); TAB(45); x2(i) ^ 2; TAB(65); ABS(x2(i) - m)
T i
NT STRING$(73, "-")
NT TAB(15); t1; TAB(45); t2; TAB(65); t3
NT TAB(1); STRING$(73, "*"): PRINT
NT "MEAN"; TAB(18); "="; m; TAB(48); "VARIANCE"; TAB(62); "="; var: PRINT
NT "STANDARD DEV."; TAB(18); "="; sd; TAB(48); "MEAN DEV."; TAB(62);       m
```

```
NT TAB(23); c; TAB(43); sl; TAB(53); s2; TAB(65); d
NT TAB(1); STRING$(73, "*"): PRINT
NT "MEAN"; TAB(18); "="; m; TAB(48); "VARIANCE"; TAB(62); "="; var: PRINT
JT "STANDARD DEV."; TAB(18); "=", sd; TAB(48); "MEAN DEV."; TAB(62); "="; m
JT "COEF. OF VAR."; TAB(18); "="; cvar; "%"
DR 12, 3: LOCATE 25, 26: PRINT "Press Any Key to Continue...": ans$ = INPUT

p2:
: COLOR 1, 2: CLS
TE 12, 12: PRINT "THIS MODULE IS FOR A FREQUENCY DISTRIBUTION AND COMPUTES
TE 13, 11: PRINT "MEAN,VARIANCE,S.D,C.V,M.D & RANGE FOR THE INPUTED DATA"
R 12, 3: LOCATE 25, 26: PRINT "Press Any Key to Continue...": ans$ = INPUT
R 3, 1: CLS : s1 = 0: s2 = 0: s3 = 0
TE 12, 23: INPUT "Enter Number of Variables--->", n1
M x1(n1), midp(n1), k(n1)
= 1 TO n1
 COLOR 1, 7: LOCATE 12, 25: PRINT "Enter the value of x"; i; "--->": LOCA
 x1(1): COLOR 3, 1
i: max = 0
 = 1 TO n1
(i) >= max THEN max = x1(i)
(i) <= min THEN min = x1(i)

x - min: k = 1 + 3.32 * 1 / LOG(10) * LOG(n1): cw = CINT(r / k)
 <= 0 THEN
 CLS : COLOR 12, 3: LOCATE 12, 10: PRINT "THE DATA INPUTED CAN NOT BE ZRO
 12, 3: LOCATE 25, 26: PRINT "Press Any key to Continue...": ans$ = INPUT

CINT(k): sp = min: 'CINT(min-cw/2)
 7, 1: CLS : PRINT TAB(16); "COMPUTED VALUES OF SOME MEASURES OF DISPERSI
 TAB(27); "FOR GROUPED DATA"
 TAB(1); STRING$(73, "*")
 "S/N."; TAB(8); "Class Int."; TAB(24); "f"; TAB(34); "x"; TAB(44); "   ";
 STRING$(73, "-")
 = 1 TO nc
sp + cw: k = 0
 = 1 TO n1
(j) >= sp AND x1(j) <= uc THEN k(i) = k(i) + 1
i: midp(i) = (sp + uc) / 2
(i) * midp(i): fx2 = k(i) * midp(i) ^ 2
s1 + fx: s2 = s2 + fx2
i; TAB(7); sp; "-"; uc; TAB(23); k(i); TAB(33); midp(i); TAB(43); fx: TA
sp + cw + 1

 / n1: var = s2 / n1 - m ^ 2
SQR(var): cvar = (sd / m) * 100
 = 1 TO nc
(i) * ABS(midp(i) - m): s3 = s3 + d: LOCATE 5 + i, 65: PRINT d

= s3 / n1
 STRING$(73, "-")
 TAB(23); n1; TAB(43); s1; TAB(53); s2; TAB(65); s3
 TAB(1); STRING$(73, "*"): PRINT
 "MEAN"; TAB(18); "="; m; TAB(48); "VARIANCE"; TAB(62); "="; var: PRINT
 "STANDARD DEV."; TAB(18); "="; sd; TAB(48); "MEAN DEV."; TAB(62);
 "COEF.OF VAR."; TAB(18); "="; cvar; "%"; TAB(48); "RANGE"; TAB(62); "=";
 12, 3: LOCATE 25, 26: PRINT "Press Any key To Continue...": ans$ = INPUT

3:
 COLOR 1, 2: CLS
E 12, 12: PRINT "THIS MODULE COMPUTES MEAN,VARIANCE,S.D,C.V,M.D & RANGE"
E 13, 24: PRINT "FOR GROUPED DISCRETE DATA"
 12, 3: LOCATE 25, 26: PRINT "Press Any key To Continue": ans$ = INPUT(1
 LET k1 = 0: k2 = 0: k3 = 0
```

```
INT "COEF.OF VAR."; TAB(18); "="; cvar; "%"; TAB(48); "RANGE"; TAB(62); "--";
LOR 11, 3: LOCATE 24, 24: PRINT "Press Any key to Continue...": ans$ = INPUT

T:
LOR 1, 7: CLS : LOCATE 12, 15: INPUT "ARE YOU SURE YOU WANT TO QUIT THIS PRO
UCASE$(ans$) = "N" THEN GOTO MENU ELSE IF UCASE$(ans$) = "Y" THEN GOTO CONT
UCASE$(ans$) <> "Y" OR ans$ <> "N" THEN
 IF
  : COLOR 6, 2:
ATE 12, 22: PRINT "INVALID ANSWER,PLEASE PRESS Y OR N": BEEP
EP (2): GOTO QUIT
T:
EEN 1: COLOR 4, 5: FOR i% = 2 TO 5: COLOR i%: LOCATE 12, 16: PRINT "GOOD B
EP (1)
T i%
P
```

4

THE FLOW CHART FOR MEASURES OF DISPERSION

Start

M

Display

| 1 Grouped Data | 2 Group continues | 3 Grouped discrete data | 4 Ungrouped data | 5 Exit |

Input OPT

Is 5 < OPT < 1

NO → M

YES

Is OPT = 1 → A

Is OPT = 2 → B

Is OPT = 3 → C

Is OPT = 4 → D

E

Column 1

```
        ┌─────────┐
        │    A    │
        └────┬────┘
             │
    ┌────────▼────────┐
    │  Let S1 = o     │
    │      S2 = o     │
    │      S3 = o     │
    │      C= o       │
    │      d= o       │
    └────────┬────────┘
             │
     ╱───────▼───────╲
    ╱     Input       ╲
    ╲       N         ╱
     ╲───────┬───────╱
             │
  ╱──────────▼──────────╲
 ╱   Dim L (N) U (N)     ╲
│    F (N), X (N)         │
 ╲   FOR I = I TO N      ╱
  ╲──────────┬──────────╱
             │
     ╱───────▼───────╲
    ╱   Input L (1)   ╲
    ╲   U (1), F (1)  ╱
     ╲───────┬───────╱
             │
    ┌────────▼────────┐
    │ Let x (1) = (L (1))
    │ u(11)|2          │
    │ S1= S1 + f(1)* x(1)
    │ S2 = S2 +fou) xx (1) **²
    └────────┬────────┘
             │
  ╱──────────▼──────────╲
 ╱      Me x t I         ╲
  ╲──────────┬──────────╱
             │
    ┌────────▼────────┐
    │ Let m = S₁/ c    │
    │ V2 = S₂/c-m**2   │
    │ Sd = SQR (VA-c)  │
    │ CVa - =          │
    └────────┬────────┘
             │
  ╱──────────▼──────────╲
 ╱       For             ╲
  ╲     I = 1 to N       ╱
  ╲──────────┬──────────╱
             │
        ┌────▼────┐
        │    f    │
        └─────────┘
```

$Let\ m = S_1/ c$

$V2 = S_2/c-m**2$

Column 2

```
        ┌─────────┐
        │    f    │
        └────┬────┘
             │
    ┌────────▼────────┐
    │ Let d = d + f(i)*
    │       Abs        │
    │   Abs (x (1) –n) │
    └────────┬────────┘
             │
  ╱──────────▼──────────╲
 ╱       Next I          ╲
  ╲──────────┬──────────╱
             │
    ┌────────▼────────┐
    │  Let             │
    │  M dev = d/c     │
    └────────┬────────┘
             │
  ╱──────────▼──────────╲
 ╱       For             ╲
  ╲      1 = 1 to N      ╱
  ╲──────────┬──────────╱
             │
     ╱───────▼───────────╲
    ╱ Print L(1), V(1), X(1)
   │  F(1)* X(1), f(1)x X (1)²
    ╲ F(1)* ABS (x(1) – m)
     ╲───────┬───────────╱
             │
  ╱──────────▼──────────╲
 ╱       Me x + I        ╲
  ╲──────────┬──────────╱
             │
     ╱───────▼───────────╲
    ╱  PRINT C, S1, S2    ╲
   │   d, m, ur, Sd        │
    ╲  mdev, cuar         ╱
     ╲───────┬───────────╱
             │
        ┌────▼────┐
        │    m    │
        └─────────┘
```

B

Let S1 = o
S2 = o
S3 = o

Input
N

Dim X_1 (N_1)
Mid p (N_1)
I = 1 To N_1

For
I = 1 To N_1

Input
X_1 (1)

Let
Min = X(1)

NEXT I

Let
Mx = o

For I = 1 To N1

G

G

IS
X1(1)>
Max

MAX =X1(1)

Is
X_1(i)<
min

Min =x1 (1)

Next I

R=max – min
K=1+3.32x log (1)
Cw=cint (r |k)

Is
Cw< 0 —— YES —— M

NO

NC=CINT (K)
SP=MIN

FOR I =1 TOrnl

H

Left column

H

VC = sp + cw

For
S = l to NI

Sp ≥ x1 (1) ≤uc

K (1) =k (1) + 1

Next j

Mid p0) = (sp tvc)/2
FX =k (1) 2 midp (1)
S2=s2+f x2

Print I, sp
Uc, k (1), mid pi)
FX, fx2

Sp =sp + cw+¡

1

Right column

1

Next i

M = s, |N,
Var=s2|ni-mxx2
Sn=s opr (var)
Cvar = (SD)

For
I =1 TO NC

d=
K (1)x AB,
(midp (1) -m
S3 = s3 + d

Next i

M d e v =S3 | Ni

Print N, S1, S2, S3
MEAN VARIAN e
Std. dev, MEAN
DEV
CVAR. RANGE

M

Left column

C (connector)

$K_1 = 0$
$K_2 = 0$
$K_3 = 0$

Input
N_3

REDIM
$X_3(n_3)$

Input $X_3 (i)$

$f_3 (i)$

$K_1 = k_1 + f_3(i) * X_3(i)$
$K_2 = k_2 + f_3(i) * X_3(i)x^2$
$Z = Z + f_3(i)$
$Min = X_3(i)$

Next I

Max = 0

For I = 1
to N_3

J (connector)

Right column

J (connector)

Is
$X_3 (i) \geq$
max — No / YES

Max = $X_3 (i)$

Yes

Is
$X_3 (i) \leq$
min — No

MIN = $X_3 (i)$

Next I

$rg = max - min$
$m = N_1 / N_3$

SD = SQR (Var)
< Var = (Sd/m)
XNS

For
I = 1 To N_3

K (connector)

```
                    ┌──────┐
                    │  K   │
                    └──┬───┘
                       │
                       ▼
        ┌──────────────────────────────┐
        │ K = k₃ + f₃ (i)              │
        │ XABS (X₃h) - m               │
        └──────────────┬───────────────┘
                       │
                       ▼
        ╱──────────────────────────────╲
        ╲          Next I               ╱
         ╲─────────────┬───────────────╱
                       │
                       ▼
        ┌──────────────────────────────┐
        │     Mdev = k₃+ n₃            │
        └──────────────┬───────────────┘
                       │
                       ▼
       ╱────────────────────────────────╲
      ╱        For I = 1 to N₃           ╲
      ╲────────────────┬─────────────────╱
                       │
                       ▼
        ┌──────────────────────────────┐
        │ Print x₃(i), f₃(i)           │
        │ x₃(i), f₃(i), x₃(i)          │
        │ *BS (x-m)                    │
        └──────────────┬───────────────┘
                       │
                       ▼
       ╱────────────────────────────────╲
      ╱          Next I                  ╲
      ╲────────────────┬─────────────────╱
                       │
                       ▼
      ╱──────────────────────────────╲
     ╱ PRINT S₂, S₂, S3              ╱
    ╱  Mean, variance,              ╱
   ╱   Sd, cvar,                   ╱
  ╱    Mdev, range                ╱
  ╲──────────────┬───────────────╱
                 │
                 ▼
            ┌──────┐
            │  M   │
            └──────┘
```

Left column (flowchart):

D

Let $T_1= 0$
$T_2 = 0$
$T_3 = 0$

Input N_2

Rsdim $x_2 (n)$
For $j = 1$ to n_2

Input $x_2 (1)$

Let T1 $=T1x_2 (1)$
$T_2=T_2 + x_2 (2)^2$
Min $= x2 (1)$

Next I

Max = 0

For I = 1 to N_2

K

Right column (flowchart):

K

Is x_2 (1)> mix

YES

Max = $X_2 (1)$

15
$x_2 \le min$

NO

Min= $X_2(1)$

Next I

R=maa-min
$M=T_2 |N_2$
Var= $T_2|N_2$
SD =Sqr (VAR) 0
Che = so |mx 10

For I = 1TO N_2

Mdev= Ts/n_2

Let $T_2= T_2+$ ABS(x (1) -m)

Next I

For I = 1to O_2

RINT I
X (1), $X_2 (1)_2$

L

L

Next I

Print T_1, T_2, T_3
Mean, var, SD dev,
Mder, cvar

M