TCP1101 PROGRAMMING FUNDAMENTALS ASSIGNMENT

Trimester 1, 2021/2022

by Group 2

LECTURE SECTION: TC2V
TUTORIAL SECTION: TT5V

TEAM MEMBERS

NAME	PHONE NUMBER	EMAIL
Anis Hazirah binti Mohamad Sabry	0182740855	1211300373@student.mmu.edu.my
Muhammad Afif Afham Bin Mohd Asri	0188709636	1201303394@student.mmu.edu.my
Zaim, Sabahaddin	01116391300	1211300368@student.mmu.edu.my
Nuha Awadah Binti Mohd Yusof	0143380062	1211303209@student.mmu.edu.my

TABLE OF CONTENTS

Introduction

The task of analysing large amounts of data is considered to be almost essential for certain areas of employment, such as business. Data analysis refers to the act of cleaning, processing, transforming, and modelling large amounts of data in order to extract information from large pools of data. From analysing data, we are able to extract key decision-making information and are also given an insight into solving problems.

Reporting is also an essential part of data analysis. The data collected for data analysis is often through the form of reports. The same goes for the other way wherein analysed data is compiled together in the form of a report. These reports consist of accurate information and analysis of any given set of data by an individual or a group. Larger pools of data would provide a more accurate report compared to reports from a smaller size collection.

The world of technology is continuously changing, evolving at a rapid pace. Alongside that, the data being collected continuously grows larger and larger, to the point where people have found difficulties in manually analysing such large amounts of data.

Thus, our group has created the Data Analysis Program (BDAP). BDAP is a software program that enables users to analyze any given data file automatically.

Access to this system requires an account created by an administrator. The program allows users to load any given data file and perform operations such as performing statistical analysis, generating a report, and generating an HTML report.

Such statistical operations that the program is capable of are finding the minimum, the maximum, the median, the mean, the variance, the standard deviation, the correlation between any two selected columns, calculating the number of distinct data members, and plotting a histogram of the given data. The program will record a log of all the activities performed by the user while they are still logged in.

Workload

NAME	FUNCTION
Anis Hazirah binti Mohamad Sabry	F2.1, F2.2, F2.3, F2.4, F2.5
Muhammad Afif Afham Bin Mohd Asri	F1.1, F1.2, F1.3, F1.4, F1.5
Zaim, Sabahaddin	F4.1, F4.2, F4.3, F4.4, F4.5
Nuha Awadah Binti Mohd Yusof	F3.1, F3.2, F3.3, F3.4, F3.5, F3.6, F3.7, F3.8, F3.9

Functional Requirements

Below are the main functional requirements that this assignment should deliver.

F1. User Registration and Authentication

This requirement focuses on operations related to user accounts. Users must have registered a user account before using the system. The sub-requirements are shown below.

No	Sub-Requirement	User	Description
F1.1	Create user account	Admin	鐒 Administrator is the only user who can create users and deletes them 鐒 The user profile should contain at least the following data items: 鐒 User name 鐒 Password (must start with letter, must contain a digit and a capital letter) 鐒 User type (Administrator or Buyer) 鐒 status (active, deleted)
F1.2	Login to a user account	Admin User	鐒 Registered user login to the system by typing the username and password 鐒 The login is successful when the typed username and password matches the registered username and password stored in its persistent storage. Otherwise, the login is unsuccessful and the user can login again 鐒 The login page must contain the user name placed at the top of the page (centered).
F1.3	Logout from a user account	Admin User	鐒 Performed only by users who have logged in to the system 鐒 When the user logout, it is directed to the login screen
F1.4	Delete user account	Admin	鐒 Administrator can delete accounts 鐒 By deleting the account, the user can no longer use the deleted account's username and password to enter the system 鐒 The profile of the deleted user and his logs will not be deleted from the system. Just set to inactive or deleted.

F1.5	Change user password	Admin User	鋍 The registered user can change his or her password

F2. Data File Reader

This requirement focuses on operations related to the administrator.

The data file have the following format:

5 Id,Age,math,science,malay 0, 1, 1, 1, 1 100 11920100,19,93,92,87 11920111,20,76,87,75 :	---> number of columns ---> column titles ---> 1 : Can compute, 0 : Cannot ---> number of rows ---> row 0 with data separated by , ---> row 1 :

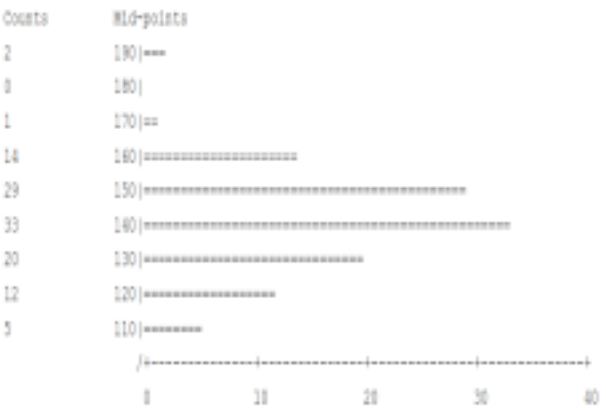
No	Sub-Requirement	User	Description
F2.1	Load data file	User	鐒 The data is loaded to a proper data structure (array, vector, list, …) 鐒 The loader should detect errors: 鐒 File exist error 鐒 Data format error
F2.2	Save as	User	鐒 Save the file with another name
F2.3	Save report	User	鐒 Saves reports into a text file based on the option selected by the user:
F2.4	Save report HTML	User	鐒 Saves a fully generated report of the statistical computations and tables in a single HTML report to be displayed using a web browser.
F2.5	Log	User	鐒 Save all the transactions of the user from log in to logout. 鐒 Your transactions should be stored each on a single line 鐒 Each line will only name the operation done and column titles selected for the operation

F3. Statistical Computation

This requirement focuses on operations related to the User.

No	Sub-Requirement	User	Description
F3.1	Find Minimum	User	⸗ Minimum of a specific column ⸗ Minimum of a complete row (minimum of all the columns)
F3.2	Find Maximum	User	⸗ Maximum of a specific column ⸗ Maximum of a complete row
F3.3	Median	User	⸗ Median of a specific column ⸗ Median of a complete row
F3.4	Mean	User	⸗ Mean of a specific column ⸗ Mean of a complete row
F3.5	Variance	User	⸗ Variance of a specific column ⸗ Variance of a complete row Sample Variance = $s^2 = \frac{\Sigma(X - \overline{X})^2}{n-1}$
F3.6	Standard Deviation	User	⸗ STDV of a specific column ⸗ STDV of a complete row $s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$
F3.7	Correlation between any selected 2 columns	User	⸗ Compute the Pearson's correlation between any two selected columns showing the table of computation of the correlation (correlation measures the strength of association between two quantitative variables). $r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\sum x_i^2 - n\bar{x}^2}\sqrt{\sum y_i^2 - n\bar{y}^2}}.$
F3.8	Distinct data members	User	⸗ Display the distinct numbers (no repetition) and a count of the occurrences of each number (frequency) in a tabular form. ⸗ Your tabular form must include header titles and must be displayed and formatted nicely.

F3.9	Plot A histogram (text mode)	User	鍔 For any selected column, plot a histogram for the data column

			鍔 Example Histogram 

F4. Menu System

This requirement focuses on operations related to the User.

No	Sub-Requirement	User	Description
F4.1	Main page	Admin, User	鏳 Program title, user name, and possible functions of a specific user. 鏳 Design your own style of menu to cover all the functional requirements
F4.2	Statistical Analysis Menu	User	鏳 Contains all the statistical functionalities and their required parameters input and menus.
F4.3	Reports menu	User	鏳 Generate reports for all the statistical computations individually. 鏳 It should contain the column and the statistical values computed.
F4.4	HTML generated Report	User	鏳 Same as F4.3 but stored as an HTML document.
F4.5	Error Checking	User	鏳 Your menu should respond to wrong input of selection with appropriate error messages

Menu system: Design and Motivation

The Menu Design:

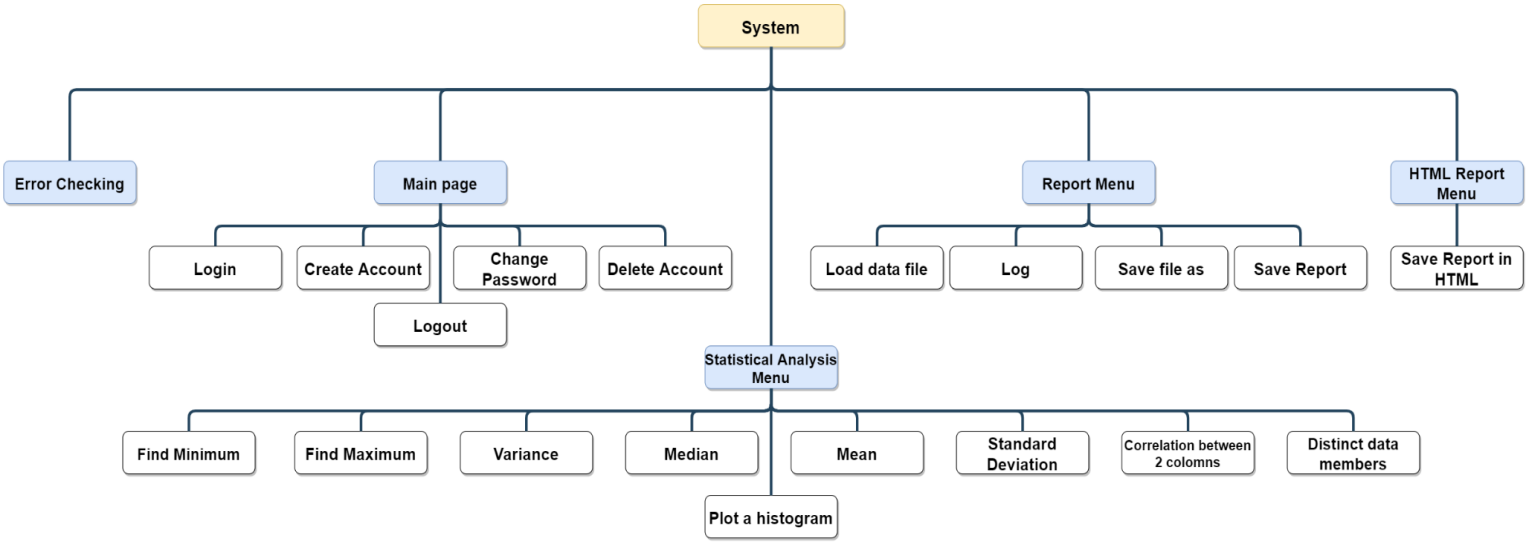
The menu of our program is designed to be simple and user friendly while still being able to efficiently perform its main functions.

Upon starting up the program, the user is prompted to log in to the system. When the user information has been inputted, the program will go through a verification process. Once verified, the user will be shown a menu listing all the commands the program can do. If the user is an admin, you have extra access and commands that are special—for example, deleting users. The menu will display report options, file saving options, and a statistical analysis menu. Despite the number of commands available, it is still easy to read and follow through. For example, when the user wishes to create a report or save it as an HTML, it is within a group of similar commands, making it easier to find it. If the file given to the system contains no errors, the user is able to perform operations through the statistical analysis menu. The design of it is similar to the main menu itself. Although all the functions are compacted into a single program, it is still easy to navigate through.

The Motivation Behind it:

The motivation behind this program's design is that if it is easy to follow and easy to understand, then the program will be satisfactory for the users. Our motivation is to make an easy to use program that maximizes user usage efficiency while still being compact and simple.

Structured Chart



Algorithms

F1. User Registration and Authentication

```
System login page
while in login page
    if user correctly inputs username and password
        allow access to system options
            display option to change password
                allow the current user to change their own password
            if user is admin
                allow access to administrative options
                    display option to create an account
                        input username, password, account type, and status of account
                    display option to delete an account
                        status of deleted user is now inactive or deleted
            logout of account
                return to login page
    else
        return to login page
```

F2. Data File Reader

```
if file exists and no data format error occurred
    load data file
        if save as
            save as another file name set by user
        else if save report
            save as report in text file based on option selected by user
        else if save report HTML
            save as generate report of statistical computations and tables in a HTML report
        else if log
            save all transactions of the user from login and out
                store each transaction in a single line
                    each transaction only name operation done and column titles for operation
else
    return to main menu
```

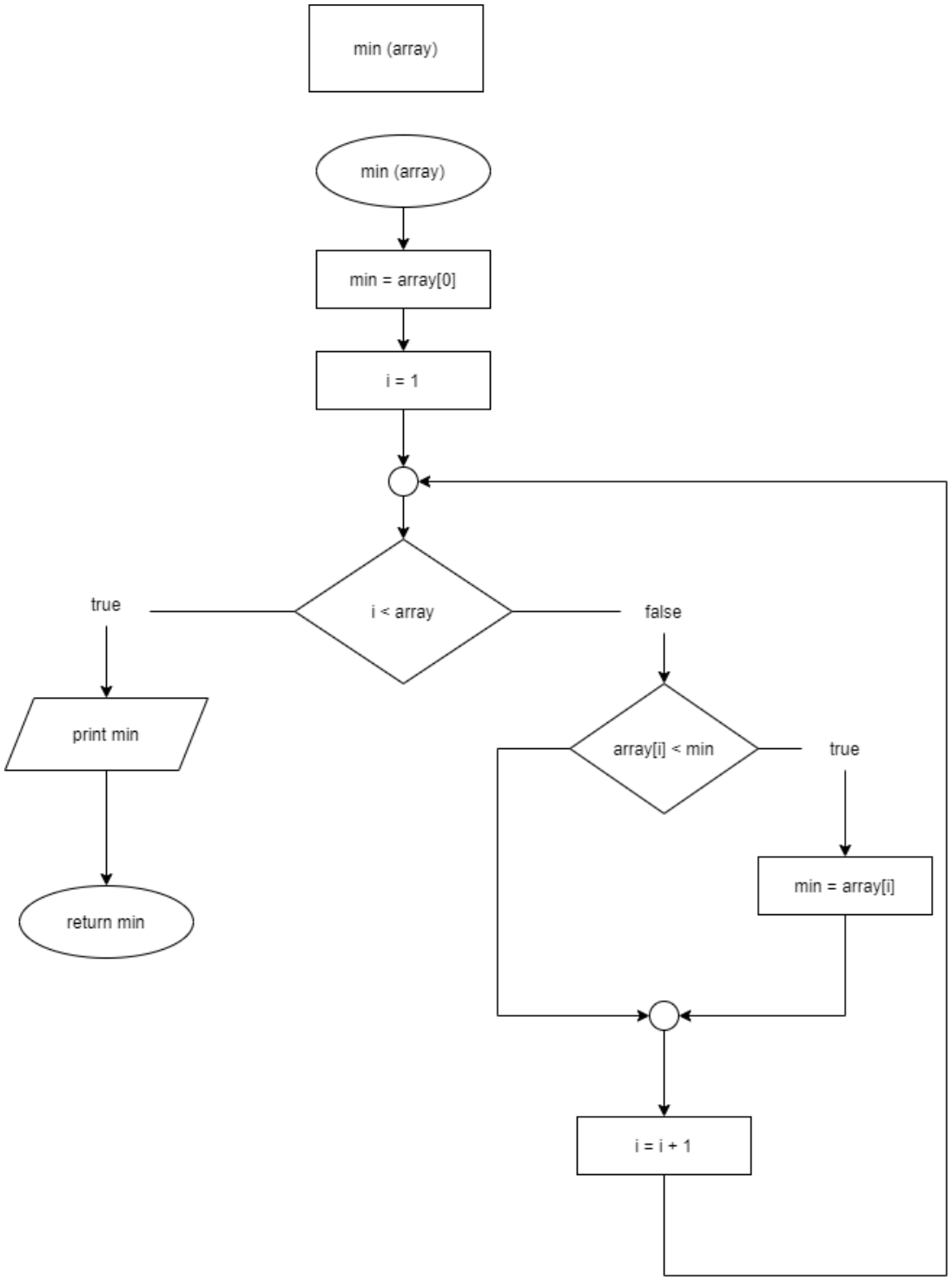
F3. Statistical Computation

User chooses statistical operations

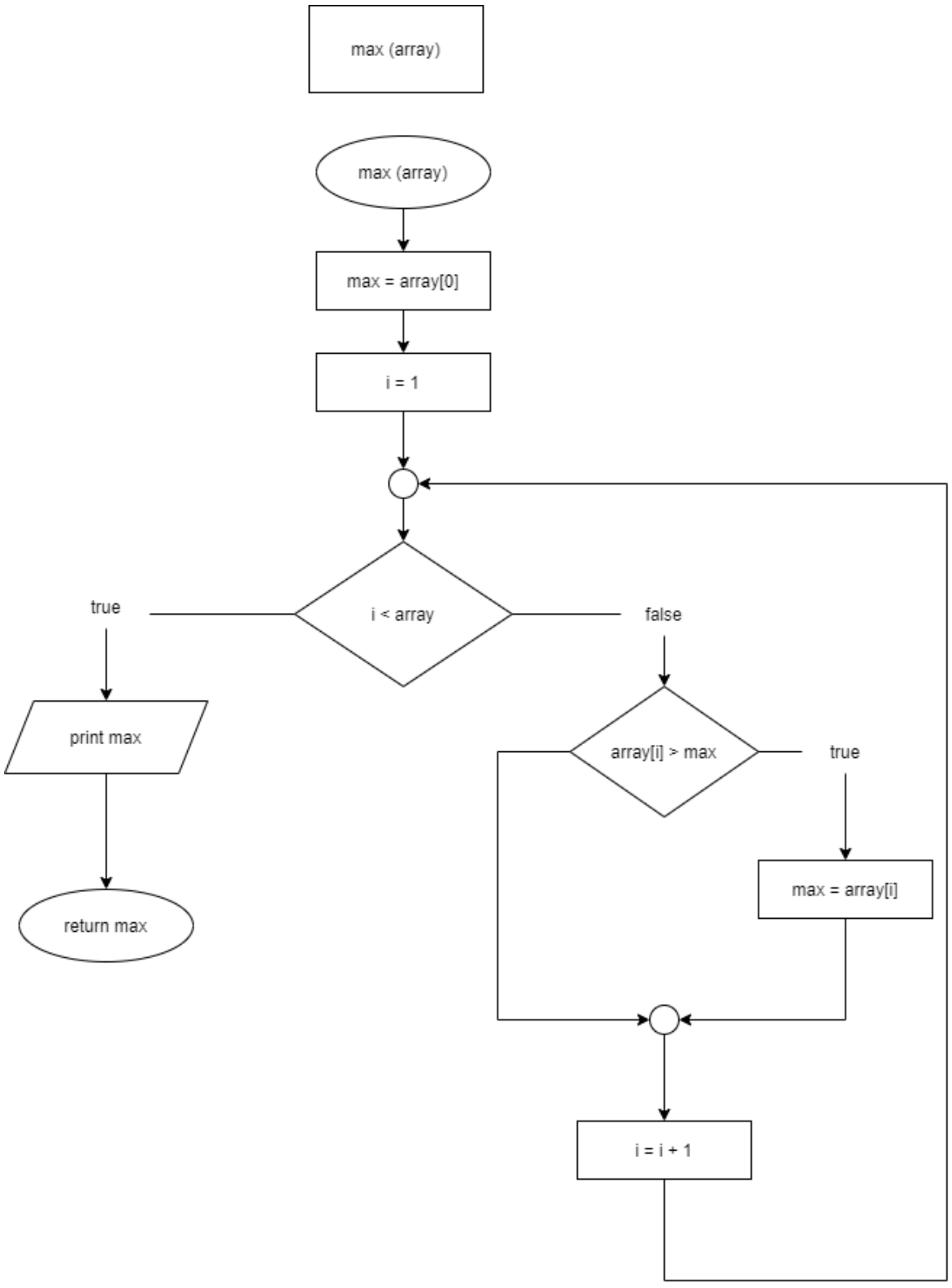
Min and max, median and mean, variance and standard deviation, correlation between two selected columns and number of distinct data members, plot a histogram

If user pick

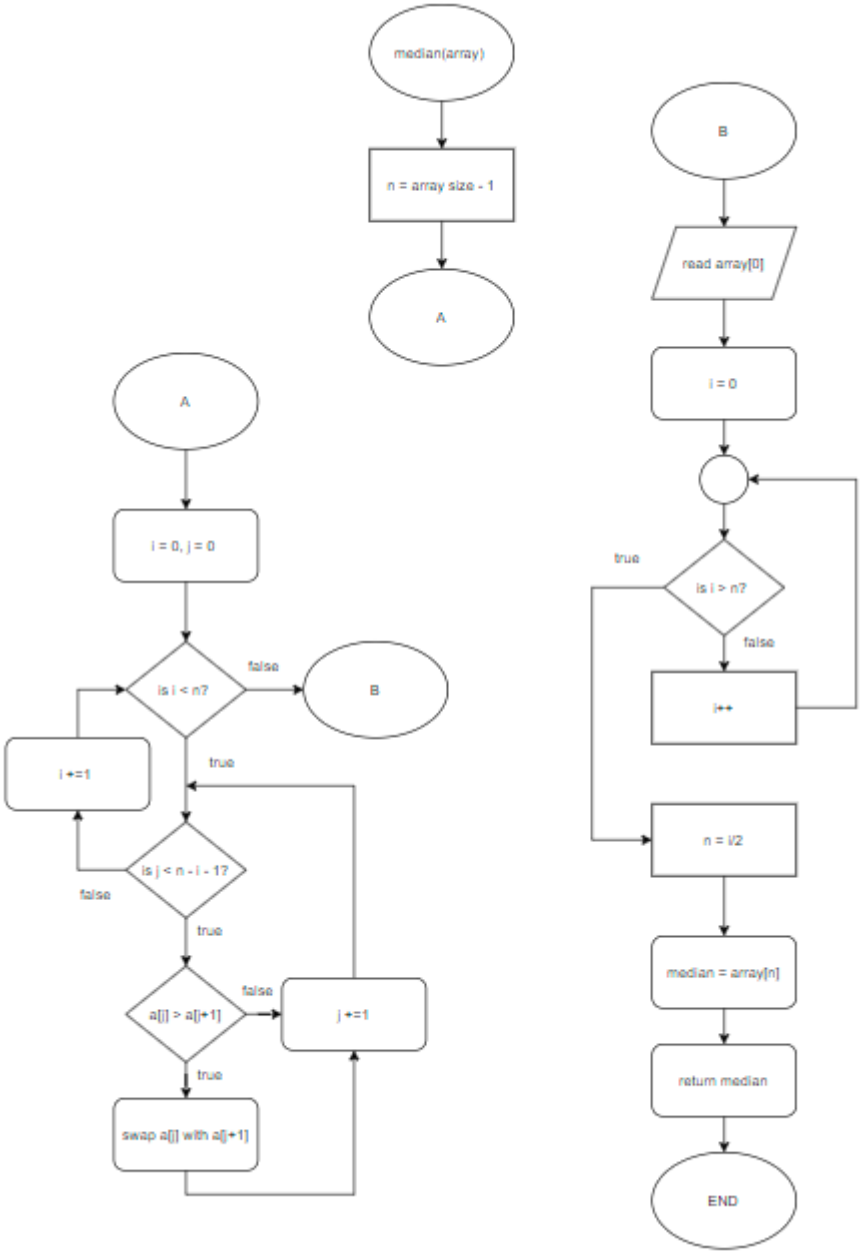
minimum,



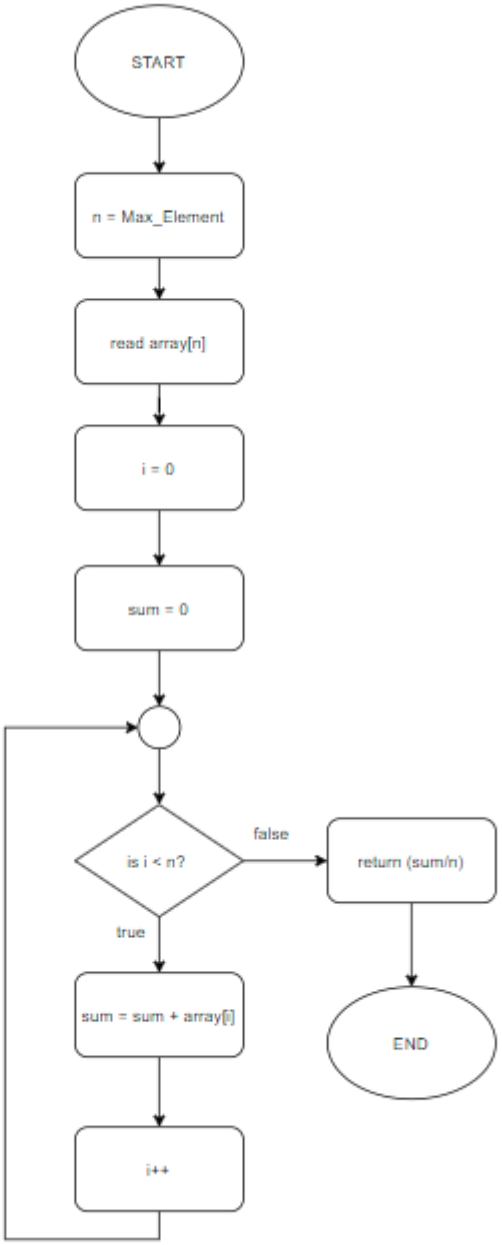
maximum,


max (array)
max (array)
max = array[0]
i = 1
i < array
true
false
print max
return max
array[i] > max
true
max = array[i]
i = i + 1

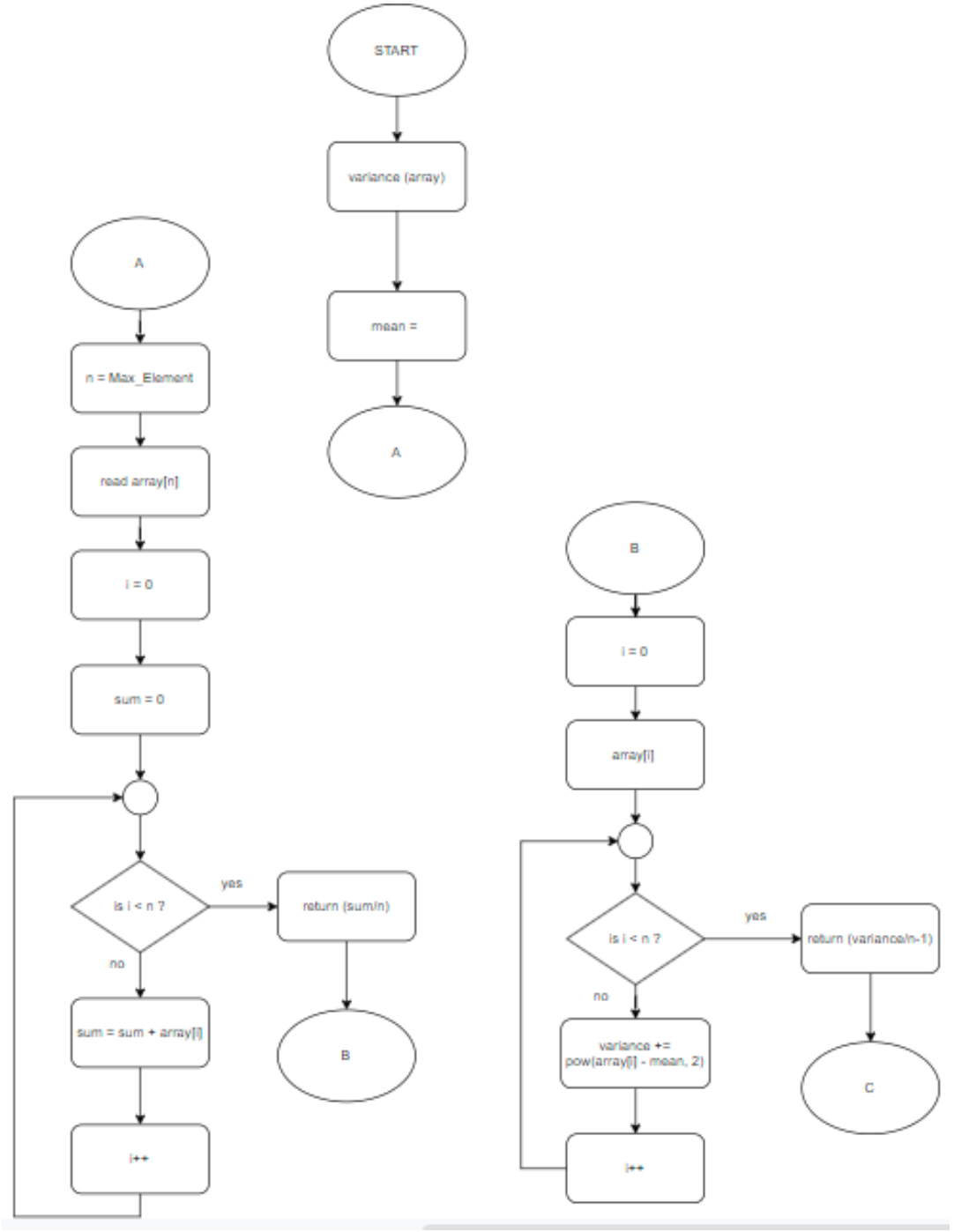
median,


median(array)
n = array size - 1
A
A
i = 0, j = 0
is i < n?
false
B
true
i +=1
is j < n - i - 1?
false
true
a[j] > a[j+1]
false
j +=1
true
swap a[j] with a[j+1]
B
read array[0]
i = 0
true
is i > n?
false
i++
n = i/2
median = array[n]
return median
END

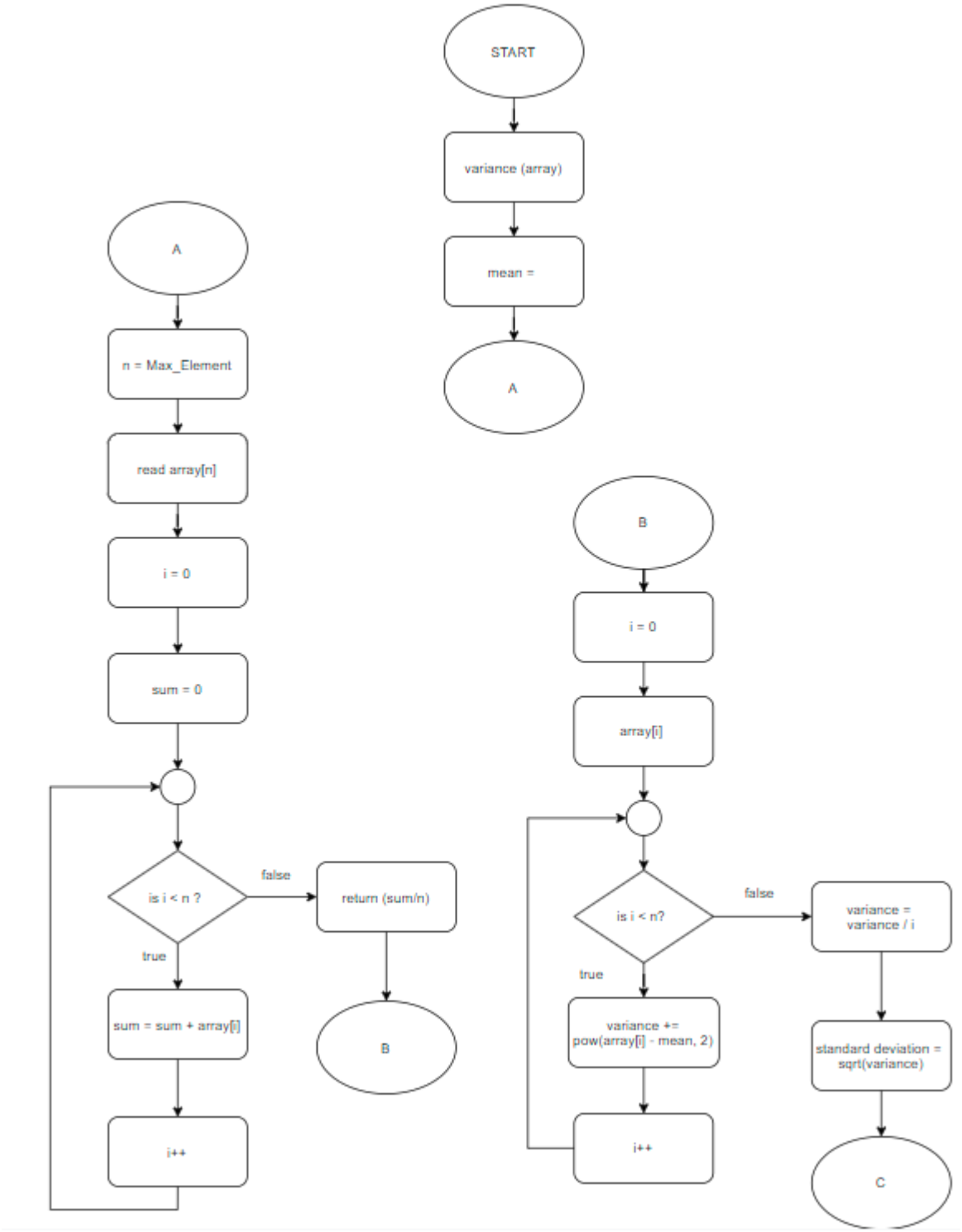
mean,


START
n = Max_Element
read array[n]
i = 0
sum = 0
is i < n?
false
return (sum/n)
true
sum = sum + array[i]
END
i++

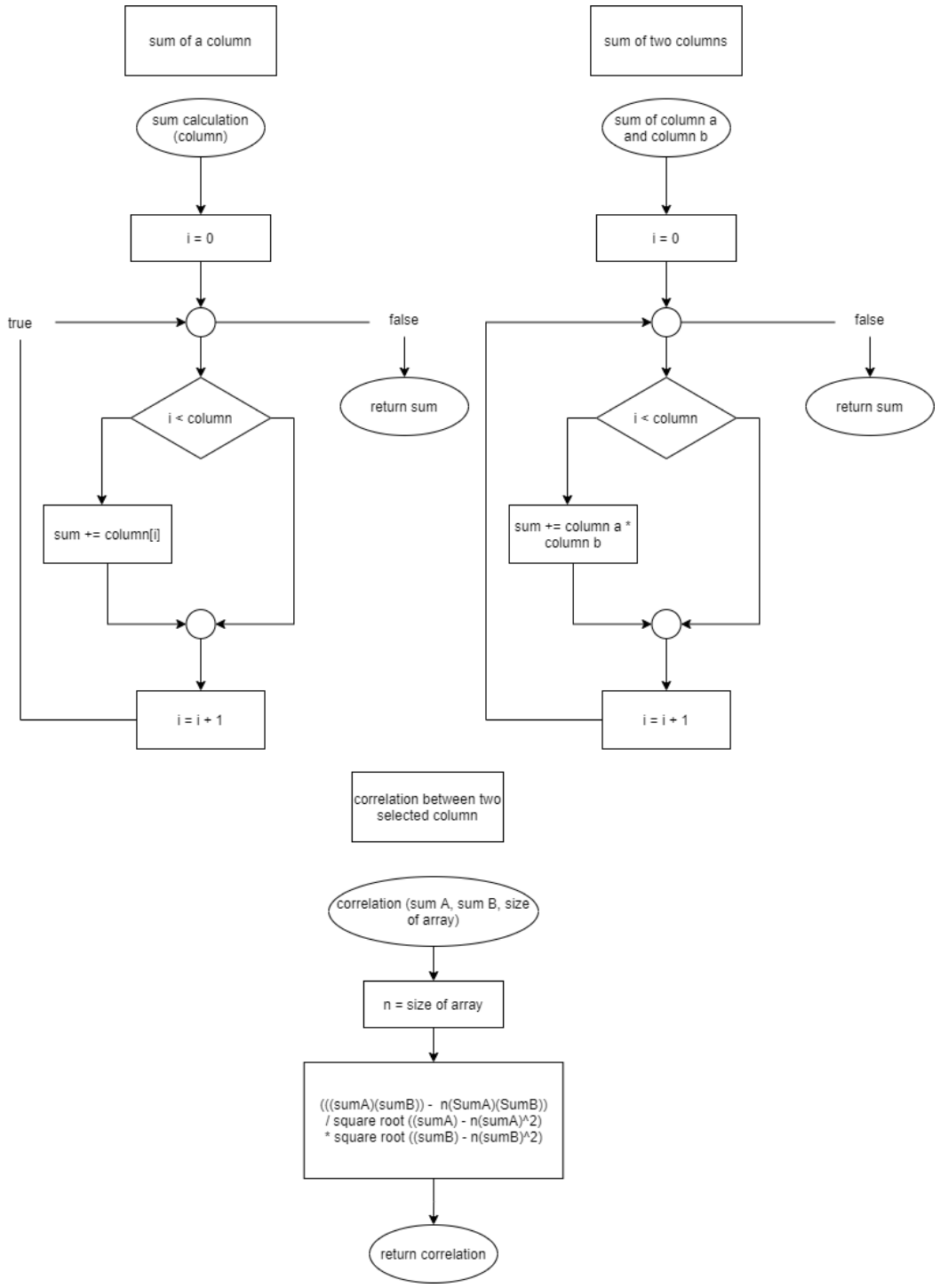
variance,


START
variance (array)
mean =
A
A
n = Max_Element
read array[n]
i = 0
sum = 0
is i < n ?
yes
return (sum/n)
no
sum = sum + array[i]
i++
B
B
i = 0
array[i]
is i < n ?
yes
return (variance/n-1)
no
variance +=
pow(array[i] - mean, 2)
i++
C

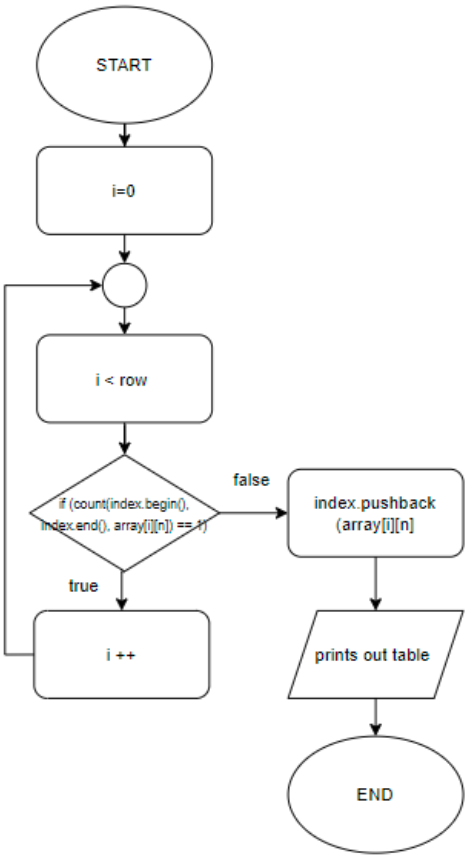
standard deviation,



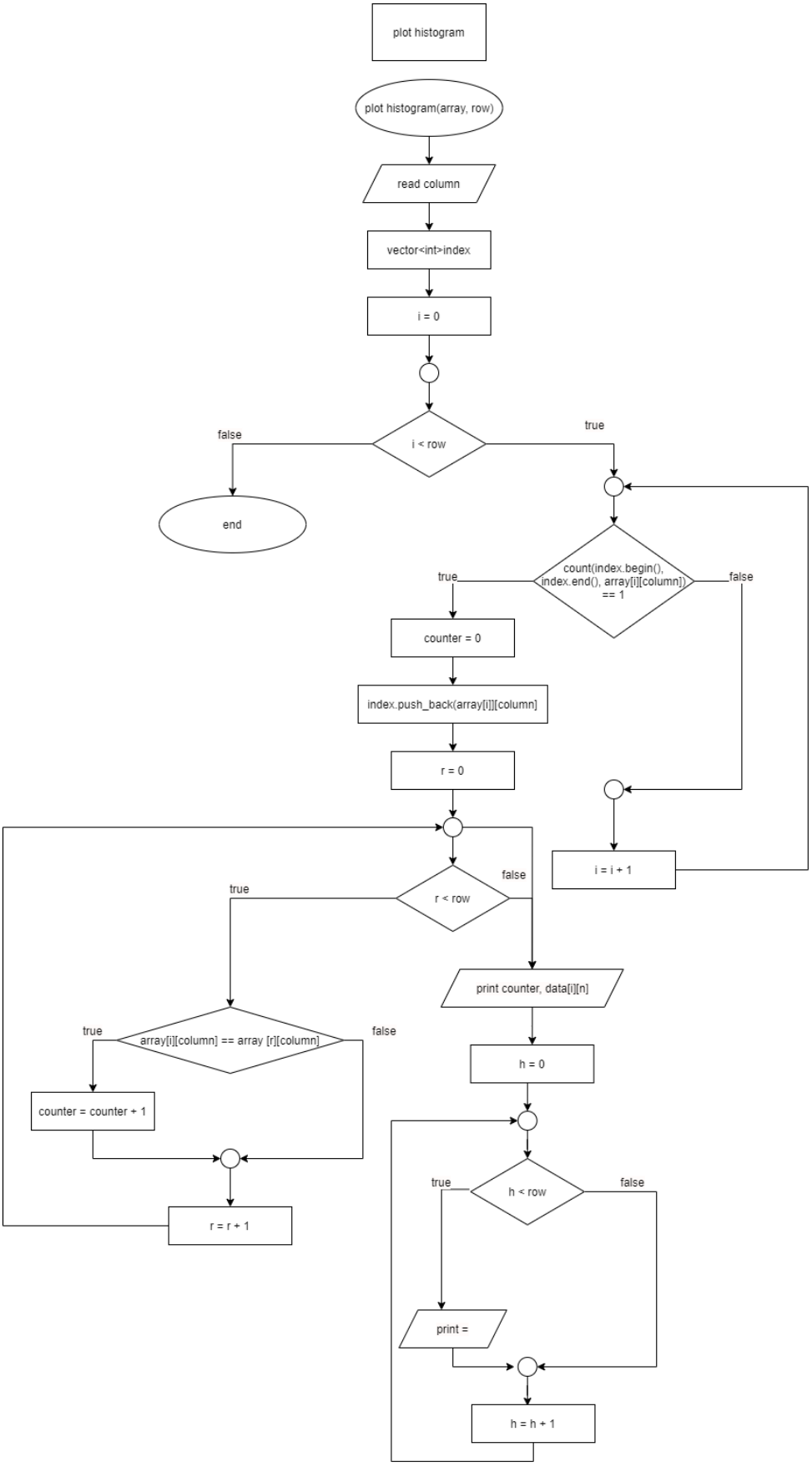
correlation between any two selected columns,



calculate data members



print histogram



exit

return to main page

F4. Menu System

Main page

while in main page

print program title, username, and menu options of a specific username

statistical analysis menu

print all statistical functionalities and their required parameters inputs and menus

reports menu

generate report for all the statistical computations individually

HTML generated report

HTML report of all statistical computations

if wrong input of selection

output an error message

End if user exits the program

Conclusion

The Data Analysis Program is a program that's able to perform statistical analysis on any given file and output the results in the form of a report of a HTML report.

The statistical analysis includes finding the minimum, the maximum, the median, the mean, the variance, the standard deviation, the correlation between any two selected columns, the distinct data members, and plot a histogram.

The program is built upon 4 main functions which are the user registration and authentication, the data file reader, the statistical computation, and the menu system. By combining these 4 functions, the program is built upon a solid foundation that can operate in the most efficient and smoothest way possible. Added with its user-friendly interactive interface, this program is able to guide the user through its system with as little chance of failure as possible. If a failure were to occur, the program can detect them and deal with them accordingly.

The reports and HTML reports produced by the program enlightens the user with the statistical analysis in an easy to comprehend manner. These reports are designed to display the statistical computation information in its utmost clarity.