

**FACULTY OF ENGINEERING & TECHNOLOGY**

**DEPARTMENT OF AGRICULTURAL MECHANIZATION & IRRIGATION ENGINEERING**

**ASSIGNMENT REPORT ON APPLICATION OF KNOWLEDGE ACQUIRED FROM MODULES ONE THROUGH THREE USING MATLAB**

**PRESENTED TO   
THE COMPUTER PROGRAMMING COURSE LECTURER  
MR. MASERUKA BENDICTO**

**By GROUP J**

**MEMBERS OF GROUP J**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **NAME** | **REG NUMBER** | **PROGRAM** |
| **1** |  |  |  |
| **2** |  |  |  |
| **3** |  |  |  |
| **4** |  |  |  |
| **5** |  |  |  |
| **6** |  |  |  |
| **7** |  |  |  |
| **8** |  |  |  |
| **9** |  |  |  |
| **10** |  |  |  |
| **11** |  |  |  |
| **12** |  |  |  |
| **13** |  |  |  |
| **14** |  |  |  |
| **15** |  |  |  |
| **16** |  |  |  |
| **17** |  |  |  |
| **18** |  |  |  |
| **19** |  |  |  |
| **20** |  |  |  |

# ABSTRACT

This report details a two-part MATLAB assignment focused on applying data handling and organization skills. The first exercise involved importing a dataset of Olympic athletes from Kaggle, cleaning it, and then using a MATLAB script to process the data.

The script organized the data by year and exported it into a single Excel workbook with separate sheets for each year. The second exercise concentrated on structuring diverse data types by using a MATLAB script to collect various personal attributes such as name, age, and interests from each gruop member and storing them in a single, cohesive structural array variable. The project demonstrated fundamental skills in data handling, organization, and problem-solving within the MATLAB environment, providing practical experience in a complete data workflow.

**Table of Contents**

[ABSTRACT 1](#_Toc208383288)

[ACKNOWLEGEMENT 5](#_Toc208383289)

[DECLARATION 6](#_Toc208383290)

[APPROVAL 7](#_Toc208383291)

[Chapter 1: Introduction 8](#_Toc208383292)

[Chapter 2: Exercise solutions 9](#_Toc208383293)

[Exercise 1: Kaggle Dataset categorizing 9](#_Toc208383294)

[Exercise 2: Group Member Data Storage 11](#_Toc208383295)

[Chapter 3: Conclusion and Learning Experience 13](#_Toc208383296)

[Chapter 4: References and Resources 14](#_Toc208383297)

# ACKNOWLEGEMENT

First, and foremost: we thank God Almighty for His unending love and protection upon our lives (Psalms 100: 5) and for coming through for us when we didn’t think we would make it out of this alive.

Secondly, special regards to our course lecturer, Mr. Maseruka Ben for guiding us in this course which is a vital aspect for our engineering profession.

Appreciation goes to our group members for the commitment and team spirit which simplified work and made it easy for us to complete the task and come up with this report.

# DECLARATION

We, Group J members hereby declare to the best of our knowledge, that this assignment report is a true record of our unending efforts in applying the knowledge we acquired from modules one through three. It is truly an original creation of our own and it has never been used by any other individual for any academic award in any learning institution.

|  |  |
| --- | --- |
| **Signature** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
| **Date** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
|  | for Group J |

# APPROVAL

This is to confirm that this report has been written and presented by Group J, giving details of the assignment carried out.

|  |  |
| --- | --- |
| **Signature** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
| **Date** | **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
|  | Course Lecturer |

## Chapter 1: Introduction

This report provides a detailed overview of the two-part MATLAB assignment. The first exercise focused on data manipulation, specifically the process of importing a dataset from Kaggle, structuring it for analysis, and exporting it to an organized Excel workbook. The second exercise involved a more focused task of structuring heterogeneous data by storing a group of members' personal attributes into a single, cohesive variable. In this assignment we were required to demonstrate fundamental skills in data handling, organization, using code within the MATLAB environment.

**Exercise 1**

Retrieve a unique data set from kaggle.com in Excel format.  
Write a single MATLAB script to:

1. Read the dataset into MATLAB
2. Copy variables for each year
3. Display data for each year in separate tables
4. Convert the tables from (3) into structural arrays
5. Output the variables from the structural arrays into one Excel workbook and separate sheets for data of each year with clear column headings

**Exercise 2**

Each group has different members from different backgrounds, home districts, religions, tribes, villages, courses, interests, ages, names and facial representations.

Write a MATLAB code that can store each members attributes into a single variable. Ensure the code saves the value.

## Chapter 2: Exercise solutions

### Exercise 1: Kaggle Dataset categorizing

The primary objective of this exercise was to create a MATLAB script suitable for handling an external dataset. Our approach was a three-step process:

#### Methodology

**Data Acquisition and Import:** We sourced a unique dataset from Kaggle in Excel format. The MATLAB script was designed to use a function to read the entire .csv file, ensuring all data was available for processing. Our exercise required that the data be of a variety containing a field for the year. We chose a dataset of Olympic athletes over the span of some 100+ years.

We went ahead to clean the data in Excel in order to remove records with missing values and, remove records with uncoherent data as well.

Once the data was clean and usable, our next step was to import it into our MATLAB environment. For that we added the dataset file into our exercise folder for convenience purposes, imported it to MATLAB and proceeded with further processing.

**Data Processing:** Our end task was to structure the data into separate years maintaining the structure of the individual years’ data while keeping each individual year separate. To do this we used the MATLAB script to access the field labelled “Year” and extract unique values. These are the different years that our data has, repeated across multiple records. For our next step, we utilized the loop feature of MATLAB to iterate over our dataset while extracting the records of data that match each of the unique years; for every unique year!

Once we had extracted records of the year of interest, we converted the resulting filtered table into a structural array; and added it to a worksheet in our target Excel workbook.   
At the end of the iteration, we had an Excel workbook with different sheets corresponding to different years of Olympic athlete data.

This final step required the script to output the structured data back into a single Excel workbook. Using MATLAB's built-in functions, we created an Excel workbook and to achieve the separation of the worksheets we used a property called Sheet, and assigned the respective years as sheet names. This ensures the output is easy for a human to interpret.

Below is the code we used to write a script that would accomplish our intended task, we saved the .m script into our exercise folder and executed the script using the Run command in our MATLAB IDE.

#### Code for our MATLAB script for exercise 1

*01:* *% Input and output file paths*

*02:* kaggle\_data\_file = "\athlete\_events.csv";

*03:* output\_file = "\output.xlsx";

*04:*

*05:* *% Load data and initialize structure*

*06:* kaggle\_data = readtable(kaggle\_data\_file);

*07:* struct\_arrays = struct();

*08:* unique\_years = unique(kaggle\_data.Year);

*09:*

*10:* *% Process data for each year*

*11:* *for* i = 1:length(unique\_years)

*12:* year = unique\_years(i);

*13:* year\_data = kaggle\_data(kaggle\_data.Year == year, :);

*14:* currentYear = sprintf('Year\_%d', year);

*15:*

*16:* *% Store data in struct and Excel*

*17:* S = table2struct(year\_data);

*18:* struct\_arrays(i).(currentYear) = S;

*19:* writetable(year\_data, output\_file, 'Sheet', currentYear);

*20:* *end*

*21:*

*22:* *% Display results*

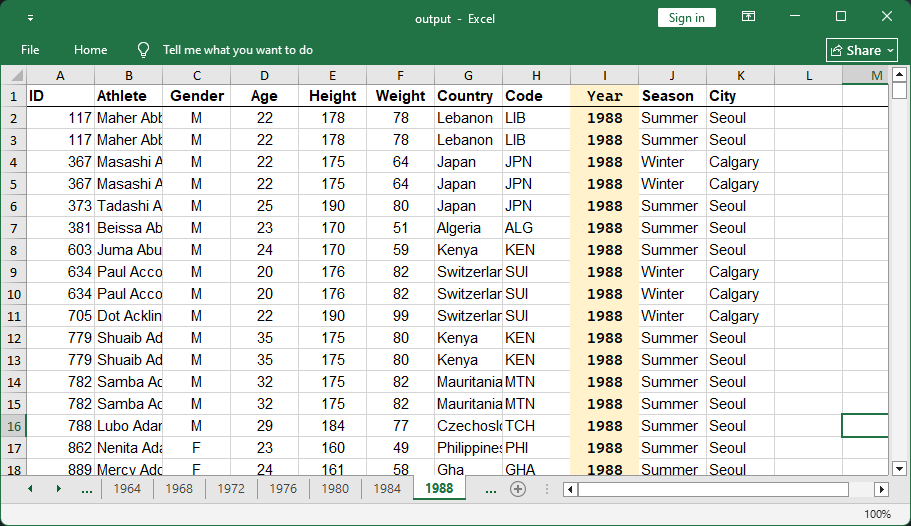
*23:* disp('Loop finished. Displaying the final struct:');

*24:* disp(struct\_arrays);

*25:* disp('All field names in the struct:');

*26:* disp(fieldnames(struct\_arrays));

*27:* disp('Data has been written to Excel file.');



**Figure 1** A screenshot of the output Excel workbook showing the result of running the MATLAB script on the Kaggle dataset, the 1988 worksheet is selected displaying filtered records for 1988

### Exercise 2: Group Member Data Storage

This exercise focused on a different aspect of data management: structuring varied data types into a single variable. The task was to take a set of attributes for each group member including home district, religion, tribe, interests, age, name, and a description for facial representation and store them collectively.

To accomplish this, we chose to utilize a structural array. Each element of the array represents a single group member, and each field within that element (e.g., .name, .age, .interests) holds the corresponding attribute. This approach allows for logical grouping of related data and simplifies future access and manipulation of the information.

Our main task here was to get the data into the MATLAB environment, for that we utilized the MATLAB input feature. Using a script, we prompted the user to enter values for the attributes of the different group members; again, we utilized a loop for this. First, we prompted the user to enter the group size, in terms of number of members; and iterated the attribute prompts for each of the members

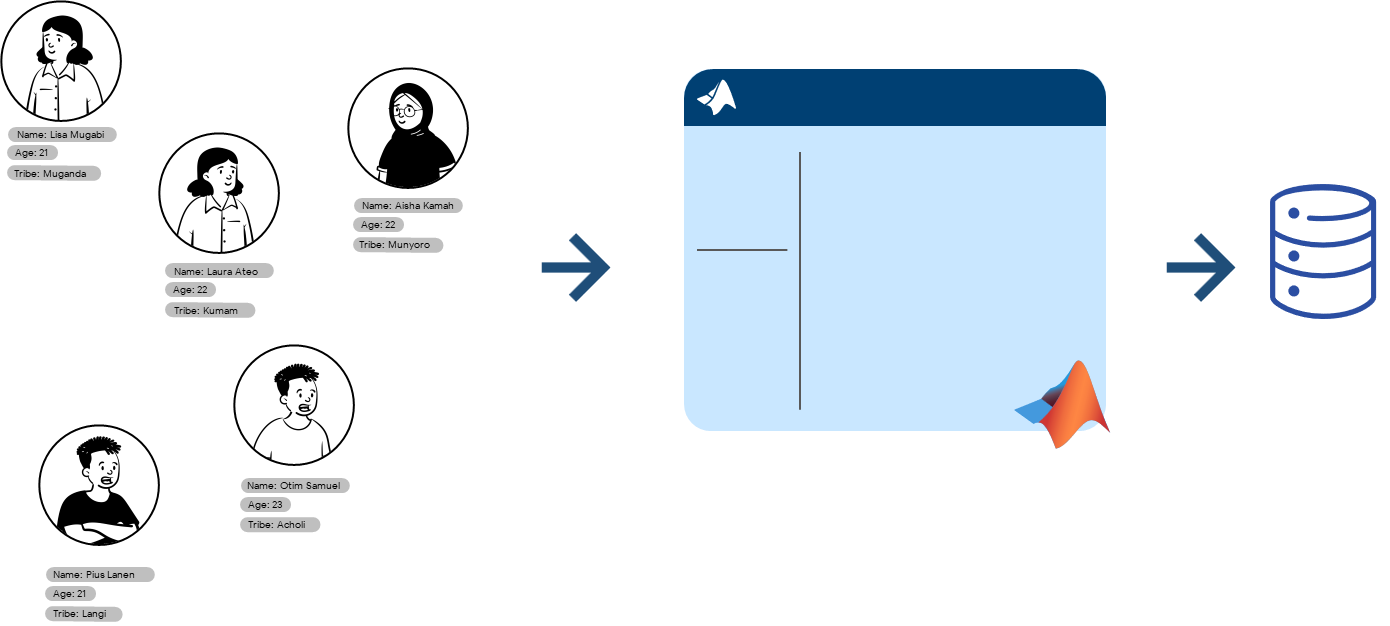


Figure An illustration of the intended task to be accomplished in Exercise 2

Below is the code for the MATLAB script we used to achieve the intended outcome of the exercise.

#### Code for our MATLAB script for exercise 2

*01:* % Greet the user and get initial group information

*02:* disp('Hello! Welcome to the Group Demographics Collector!');

*03:* group\_name = input('Please enter the name of your group: ', 's');

*04:* num\_members = input('Please enter the number of members in your group: ');

*05:*

*06:* % Initialize a struct to store data for all group members

*07:* group\_data = struct();

*08:*

*09:* % Loop to collect details for each member

*10:* *for* i = 1:num\_members

*11:* member\_info = struct();

*12:*

*13:* % Get member details via user input

*14:* member\_info.name = input(sprintf('Please enter the name of member %d: ', i), 's');

*15:* member\_info.home\_district = input(sprintf('Please enter the home district of %s: ', member\_info.name), 's');

*16:* member\_info.religion = input(sprintf('Please enter the religion of %s: ', member\_info.name), 's');

*17:* member\_info.tribe = input(sprintf('Please enter the tribe of %s: ', member\_info.name), 's');

*18:* member\_info.village = input(sprintf('Please enter the village of %s: ', member\_info.name), 's');

*19:* member\_info.courses = input(sprintf('Please enter the courses of %s (comma-separated): ', member\_info.name), 's');

*20:* member\_info.interests = input(sprintf('Please enter the interests of %s (comma-separated): ', member\_info.name), 's');

*21:* member\_info.age = input(sprintf('Please enter the age of %s: ', member\_info.name));

*22:* member\_info.facial\_representation = input(sprintf('Please enter the facial representation of %s: ', member\_info.name), 's');

*23:*

*24:* % Create a valid field name from the member's name (e.g., 'JohnDoe')

*25:* group\_member\_variable = strrep(member\_info.name, ' ', '');

*26:*

*27:* % Add the current member's data to the main group structure

*28:* group\_data.(group\_member\_variable) = member\_info;

*29:* *end*

*30:*

*31:* % Display a final thank you message

*32:* fprintf('Thank you for providing the information about your group: %s\n', group\_name);

*33:*

The MATLAB code ensures that all attributes are correctly assigned and stored within this single variable.

## Chapter 3: Conclusion and Learning Experience

Completing this assignment was a valuable learning experience that reinforced key MATLAB programming concepts and gave us hands-on experience with the foundations we had acquired from Modules 1 through 3. We gained a deeper understanding of how to handle real-world data, which we realised is often messy and requires careful structuring. The process of importing from a third-party source like Kaggle, transforming the data into a more usable format like a structural array, and then exporting it cleanly highlighted a complete data workflow. The second exercise also re-enforced our knowledge of data types and the use of structural arrays in organizing heterogeneous information. The assignment as a whole provided a practical foundation in data management and problem-solving within a technical computing environment.

## Chapter 4: References and Resources

* [kaggle.com](https://www.kaggle.com/) - The primary source for the dataset used in this assignment.
* MATLAB Documentation - Used for syntax and function guidance on readtable(), struct(), and writetable().
* Microsoft Excel Cleaning techniques from Microsoft community forums