**AI Section 2: Sources of Data and Data Types**

**Intro to Data Sources andTypes**

Objectives

* Students will be able to identify the various data types used in computer programming and understand their innate differences when working with them.
* Students will be able to look at data sets and determine what kinds of Data Types they are

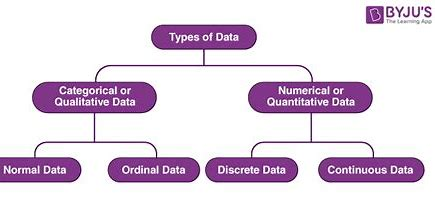
Main Learning Goal

* By the end of this lesson, students should be able to recognize all the major data types used in programming, some various methods associated with them in Matlab (we will touch on more methods in Matlab), and how to use these data types to create variables and perform various operations. Namely, students will be able to work with integers, floating-point numbers, strings and Booleans.

Focus Question

Before beginning, take a moment to think about data that could be stored in a variable and how it could be differentiated based on its type. Why do we separate data into 'types'?

What are Types of Data?



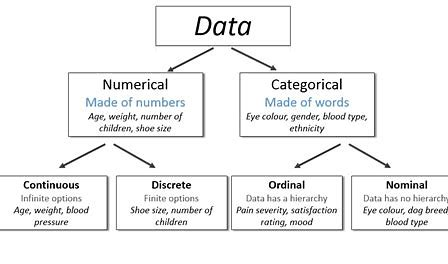


Image from BYU and AskDataScience. com

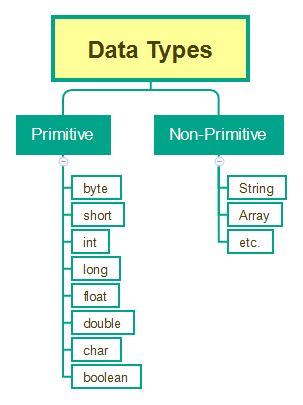
**Data can be categorized into two main types based on if they are composed or words or numbers. There are other data types we will also learn about next.**

The term “data type” in software programming describes the kind of value a variable possesses and the kinds of mathematical, relational, or logical operations that can be performed on it without leading to an error. Numerous programming languages, for instance, utilize the data types string, integer, and floating point to represent text, whole numbers, and values with decimal points, respectively. An interpreter or compiler can determine how a programmer plans to use a given set of data by looking up its data type.

The data comes in different forms. Examples include:

* your name – a string of characters
* your age – usually an integer
* the amount of money in your pocket- usually decimal type
* today’s date – written in date time format

Data is usually categorized as **Primitive** or **non-primitive**based on how it is put together.



**Image from**[**Java for Testers – Using Data Types - QAFoxLinks to an external site.**](https://www.qafox.com/java-for-testers-using-data-types/)

**Primitive Data Types** are predefined data types that include basic values like text or numbers. They are the most fundamental type of data and are used as the foundation for more complex data types. Most computer languages probably employ some variation of these simple data types.

Common types of primitive data from [Data Types in Programming - GeeksforGeeksLinks to an external site.](https://www.geeksforgeeks.org/data-types-in-programming/#common-composite-data-types)

Common Primitive Data Types in Programming:

**Some common primitive datatypes are as follow:**

| **Data Type** | **Definition** | **Examples** |
| --- | --- | --- |
| **Integer (int)** | represent numeric data type for numbers without fractions | 300, 0 , -300 |
| **Floating Point (float)** | represent numeric data type for numbers with fractions | 34.67, 56.99, -78.09 |
| **Character (char)** | represent single letter, digit, punctuation mark, symbol, or blank space | a , 1, ! |
| **Boolean (bool)** | True or false values | true- 1, false- 0 |
| **Date** | Date in the YYYY-MM-DD format (ISO 8601 syntax) | 2024-01-01 |
| **Time** | Time in the hh:mm:ss format for the time of day, time since an event, or time interval between events | 12:34:20 |

There are other data types as well. **Composite types** fall into four main categories: semi-structured (stores data as a set of relationships); **multimedia (stores data as images, music, or videos)**; homogeneous (needs all values to be of the same data type); and tabular (stores data in tabular form).

**Non-primitive data types**contain data that you put together as an "object" There are 4 main types of non-primitive data types, String, Class, Array, and Interface.

String

Strings are a group of characters surrounded by double-quotes. Strings are mostly used to store chunks of text and information. Under the hood, this class is actually storing each character inside of an array, so if you make a string that holds your name “Beyonce”.  what that string will really look like is {‘B’, ‘e’, ‘y’, ‘o’, ‘n’, ‘c', ‘e’}, this makes it possible to access each character individually like an array. The String class has a lot of methods that we can use to manipulate strings.

Array

Arrays are a way of storing information in a “list” format. Arrays are very helpful when we store data that will not have a lot of added elements, for lists that require more manipulation there are Linked Lists. Arrays can’t store integers or anything not wrapped in quotations. Array ={name, "Beyonce", "42"}

What types of data are used in computer programming?

Strictly speaking, a data type is a classification that determines what type of value a variable can have, which in turn determines what type of mathematical and logical operators can be applied to the variable.

You may look at the above definition with great confusion. A simpler way is to think about a data type as something akin to a label that tells you what information you are holding and how you can interact or manipulate it.

Variable Categories

Below is an overview of all common variable categories by type.

*Numeric Variables*

**int** -> Used to store integer values (whole numbers).

**float** -> Used to store floating-point values (decimal numbers).

*String Variables*

**str** -> Used to store text or string values. Strings are enclosed in single or double quotes.

*Boolean Variables*

**bool** -> Used to store boolean values (True or False), which represent binary conditions.

*Sequence Variables*

**list**- Modifiable collection of values / elements.

**tuple** -> Non-modifiable collection of values / elements.

*Mapping Variables*

**dict** -> Known to humans as dictionaries. These are used to store key-value pairs, where a *key* is a unique sequence of characters that corresponds to a *value* that can be of any type. These key-value pairs are important when we dive deeper into this variable type.

Data Types? What Data Types?

Matlab has a several data types, but we will be covering the four core data types that serve as the backbone of this programming language and on which other data types are built. Let us first give a simple overview of each of these data types before we examine them in greater detail.

The links will take you to the MatLab page

* [Numeric TypesLinks to an external site.](https://www.mathworks.com/help/matlab/numeric-types.html)  
  Integer and floating-point data
* [Characters and StringsLinks to an external site.](https://www.mathworks.com/help/matlab/characters-and-strings.html)  
  Text in character arrays and string arrays
* [Dates and TimeLinks to an external site.](https://www.mathworks.com/help/matlab/date-and-time-operations.html)  
  Arrays of date and time values that can be displayed in different formats
* [Categorical ArraysLinks to an external site.](https://www.mathworks.com/help/matlab/categorical-arrays.html)  
  Arrays of qualitative data with values from a finite set of discrete, nonnumeric data
* [TablesLinks to an external site.](https://www.mathworks.com/help/matlab/tables.html)  
  Arrays in tabular form whose named columns can have different types
* [TimetablesLinks to an external site.](https://www.mathworks.com/help/matlab/timetables.html)  
  Time-stamped data in tabular form
* [StructuresLinks to an external site.](https://www.mathworks.com/help/matlab/structures.html)  
  Arrays with named fields that can contain data of varying types and sizes
* [Cell ArraysLinks to an external site.](https://www.mathworks.com/help/matlab/cell-arrays.html)  
  Arrays that can contain data of varying types and sizes
* [Function HandlesLinks to an external site.](https://www.mathworks.com/help/matlab/function-handles.html)  
  Variables that allow you to invoke a function indirectly
* [DictionariesLinks to an external site.](https://www.mathworks.com/help/matlab/dictionary.html)  
  Map data with keys that index values
* [Time SeriesLinks to an external site.](https://www.mathworks.com/help/matlab/time-series.html)  
  Data vectors sampled over time
* [Data Type IdentificationLinks to an external site.](https://www.mathworks.com/help/matlab/data-type-identification.html)  
  Determine data type of a variable
* [Data Type ConversionLinks to an external site.](https://www.mathworks.com/help/matlab/data-type-conversion.html)  
  Convert between numeric arrays, strings and character arrays, dates and times, cell arrays, structures, or tables

Type Casting in Programming:

* Converting a single data type value—such as an integer int, float, or double—into another data type is known as typecasting. You have the option of doing this conversion manually or automatically. The conversion is done in two ways, automatically by the compiler and manually by a programmer.
* Type casting is sometimes known as type conversion. For example, a programmer can type cast a long variable value into an int if they wish to store it in the program as a simple integer. Thus, type casting is a technique that allows users to utilize the cast operator to change values from one data type to another.
* Type casting is used when imagine you have an age value, let’s say 30, stored in a program. You want to display a message on a website or application that says “Your age is: 30 years.” To display it as part of the message (a string), you would need to convert the age (an integer) to a string.
* Simple explanation of type casting can be done by this example:
* Imagine you have two types of containers: one for numbers and one for words. Now, let’s say you have a number written on a piece of paper, like “42,” and you want to put it in the container meant for words. Type casting is like taking that number, converting it into words, and then putting it in the container for words. Similarly, in programming, you might have a number (like 42) stored as one type, and you want to use it as if it were another type (like a word or text). Type casting helps you make that conversion.



To better understand this concept before delving into code examples, let us think about variables as containers. Imagine that you are working at a restaurant and have two containers, where one container is labeled “Chum” and the other “Beef.” Logically, you would neither want to nor be able to combine both containers due to their different labels. Now, suppose that you wanted to be sneaky and still combine both despite their different contents. How would you do that? The most logical approach would be to temporarily change the label of one of the containers to match the label of the other, which would then not draw any suspicion from those around you. In the end, you can interact with both containers while keeping the same contents for both. This process is exactly what casting is doing: you can convert the value of a variable and change it to a different data type without modifying the value itself.

TO DO:

1. Assignment: Identifying Data Types in Sports and Music Data