

# ENGR-1330-Lesson03

August 29, 2021

Download this page as a jupyter notebook at [Lesson 3](#)

## 1 ENGR 1330 Computational Thinking with Data Science

Copyright © 2021 Theodore G. Cleveland and Farhang Forghanparast

Last GitHub Commit Date: 13 July 2021

### 1.1 Lesson 3 Data Structures:

- Data structures; lists, arrays, tuples, sets, dictionaries
- Name, index, contents; keys

---

```
[18]: # Script block to identify host, user, and kernel
import sys
! hostname; ! whoami; ! pwd;
print(sys.executable)
```

```
atomickitty
sensei
/home/sensei/engr-1330-webroot/1-Lessons/Lesson03
/opt/jupyterhub/bin/python3
```

```
[19]: %%html
<!-- Script Block to set tables to left alignment -->
<style>
  table {margin-left: 0 !important;}
</style>
```

```
<IPython.core.display.HTML object>
```

---

### 1.2 Objectives

1. Awareness of data structures available in Python to store and manipulate data
2. Implement arrays (lists), dictionaries, and tuples
3. Address contents of lists , dictionaries, and tuples

---

### 1.3 Data Structures and Conditional Statements

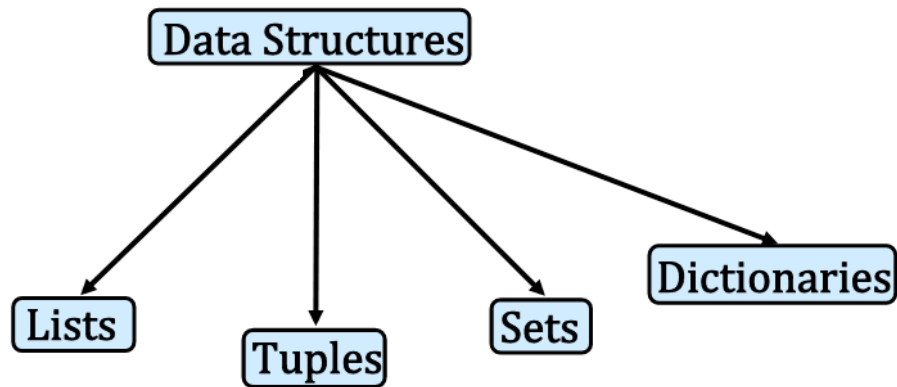
Computational thinking (CT) concepts involved are:

- **Decomposition** : Data interpretation, manipulation, and analysis of NumPy arrays
- **Abstraction** : Data structures; Arrays, lists, tuples, sets, and dictionaries
- **Algorithms** : Conditional statements

### 1.4 What is a data structure?

Data Structures are a specialized means of organizing and storing data in computers in such a way that we can perform operations on the stored data more efficiently.

In our iPython world the structures are illustrated in the figure below



#### 1.4.1 Lists

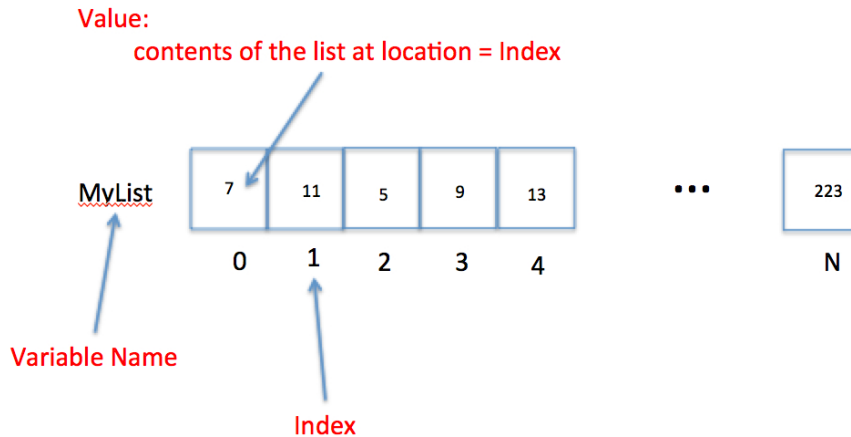
A list is a collection of data that are somehow related. It is a convenient way to refer to a collection of similar things by a single name, and using an index (like a subscript in math) to identify a particular item.

Consider the “math-like” variable  $x$  below:

$$\begin{aligned}x_0 &= 7 && (1) \\x_1 &= 11 && (2) \\x_2 &= 5 && (3) \\x_3 &= 9 && (4) \\x_4 &= 13 && (5) \\&\dots && (6) \\x_N &= 223 && (7) \\&&& (8)\end{aligned}$$

The variable name is  $x$  and the subscripts correspond to different values. Thus the **value** of the variable named  $x$  associated with subscript 3 is the number 9.

The figure below is a visual representation of a the concept that treats a variable as a collection of cells.



In the figure, the variable name is `MyList`, the subscripts are replaced by an index which identifies which cell is being referenced. The value is the cell content at the particular index.

So in the figure the value of `MyList` at `Index = 3` is the number 9.

In engineering and data science we use lists a lot - we often call them vectors, arrays, matrices and such, but they are ultimately just lists.

To declare a list you can write the list name and assign it values. The square brackets are used to identify that the variable is a list. Like:

```
MyList = [7,11,5,9,13,66,99,223]
```

One can also declare a null list and use the `append()` method to fill it as needed.

```
MyOtherList = [ ]
```

Python indices start at **ZERO**. A lot of other languages start at ONE. It's just the convention.

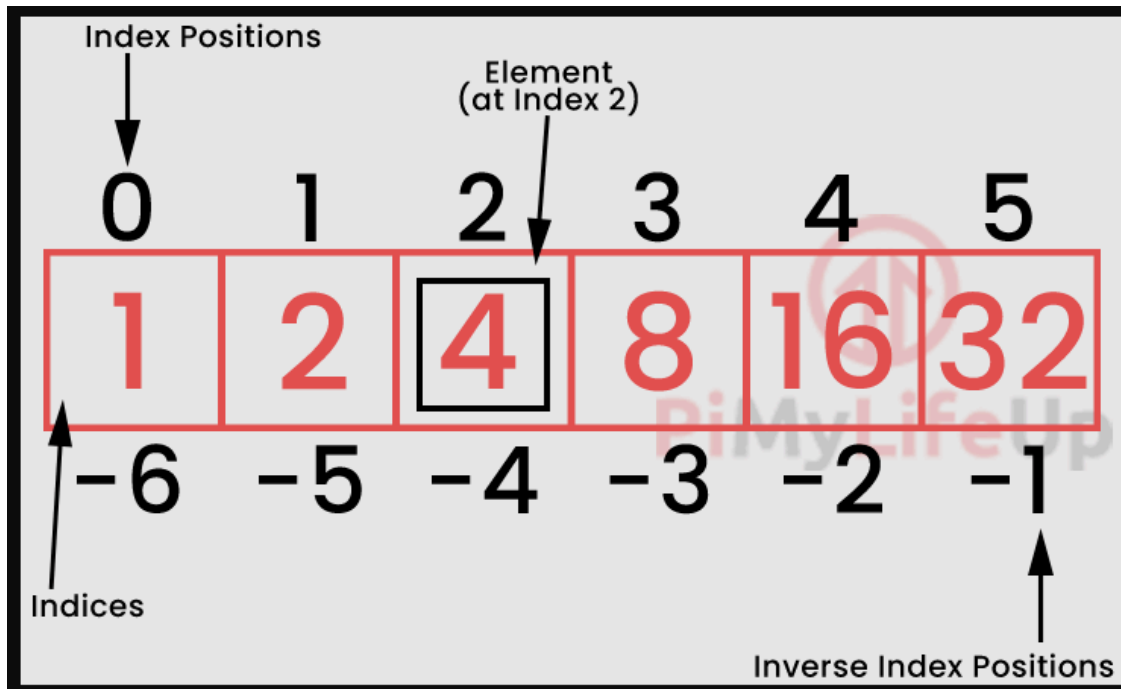
The first element in a list has an index of 0, the second an index of 1, and so on. We access the contents of a list by referring to its name and index. For example

`MyList[3]` has a value of the number 9.

## 1.4.2 Arrays

Arrays are special lists that are used to store only elements of a specific data type, and require use of an external dependency (package) named **array**. The package is installed with core python, so other than importing it into a script nothing else special is needed.

Arrays are: - Ordered: Elements in an array can be indexed - Mutable: Elements in an array can be altered



Data type that an array must hold is specified using the type code when it is created - 'f' for float - 'd' for double - 'i' for signed int - 'I' for unsigned int

More types are listed below

Type Code	C Data Type	Python Data Type	Minimum Size in Bytes
'b'	signed char	int	1
'B'	unsigned char	int	1
'h'	signed short	int	2
'H'	unsigned short	int	2
'i'	signed int	int	2
'I'	unsigned int	int	2
'l'	signed long	int	4
'L'	unsigned long	int	4
'q'	signed long long	int	8
'Q'	unsigned long long	int	8
'f'	float	float	4
'd'	double	float	8

To use arrays, a library named 'array' must be imported

```
[20]: import array
```

Creating an array that contains signed integer numbers

```
[21]: myarray = array.array('i', [1, 2, 4, 8, 16, 32])
```

```
[22]: myarray[0] #1-st element, 0-th position
```

```
[22]: 1
```

```
[23]: import array as arr #import using an alias so the calls don't look so funny
```

```
[24]: myarray = arr.array('i', [1, 2, 4, 8, 16, 32])  
myarray[0] #1-st element, 0-th position
```

```
[24]: 1
```

Lists: Can store elements of different data types; like arrays they are (arrays are lists, but lists are not quite arrays!) - Ordered: Elements in a list can be indexed - Mutable: Elements in a list can be altered - Mathematical operations must be applied to each element of the list

### 1.4.3 Tuple - A special list

A tuple is a special kind of list where the **values cannot be changed** after the list is created. Such a property is called **immutable**. It is useful for list-like things that are static - like days in a week, or months of a year. You declare a tuple like a list, except use round brackets instead of square brackets.

```
MyTupleName = ("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec")
```

Tuples are often created as output from packages and functions.

### 1.4.4 Dictionary - A special list

A dictionary is a special kind of list where the items are related data **PAIRS**. It is a lot like a relational database (it probably is one in fact) where the first item in the pair is called the key, and must be unique in a dictionary, and the second item in the pair is the data. The second item could itself be a list, so a dictionary would be a meaningful way to build a database in Python.

To declare a dictionary using curly brackets

```
MyPetsNamesAndMass = { "Dusty":7.8 , "Aspen":6.3, "Merrimee":0.03}
```

To declare a dictionary using the `dict()` method

```
MyPetsNamesAndMassToo = dict(Dusty = 7.8 , Aspen = 6.3, Merrimee = 0.03)
```

Dictionary properties - Unordered: Elements in a dictionary cannot be - Mutable elements: Elements in a dictionary can be altered - Immutable keys: Keys in a dictionary cannot be altered

### 1.4.5 Sets - A special list

Sets: Are used to store elements of different data types - Unordered: Elements in a set cannot be indexed - Mutable: Elements in a set can be altered - Non-repetition: Elements in a set are unique

Elements of a set are enclosed in curly brackets `{ }` - Creating sets that contains different data types - Sets cannot be nested

**What’s the difference between a set and dictionary?** From <https://stackoverflow.com/questions/34370599/difference-between-dict-and-set-python>

“Well, a set is like a dict with keys but no values, and they’re both implemented using a hash table. But yes, it’s a little annoying that the {} notation denotes an empty dict rather than an empty set, but that’s a historical artifact.”

## 1.5 Readings

1. Computational and Inferential Thinking Ani Adhikari and John DeNero, Computational and Inferential Thinking, The Foundations of Data Science, Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND) Chapter 4 Subpart 3 <https://www.inferentialthinking.com/chapters/04/3/Comparison.html>
2. Computational and Inferential Thinking Ani Adhikari and John DeNero, Computational and Inferential Thinking, The Foundations of Data Science, Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND) Chapter 4 [https://www.inferentialthinking.com/chapters/04/Data\\_Types.html](https://www.inferentialthinking.com/chapters/04/Data_Types.html)
3. Learn Python in One Day and Learn It Well. Python for Beginners with Hands-on Project. (Learn Coding Fast with Hands-On Project Book – Kindle Edition by LCF Publishing (Author), Jamie Chan [https://www.amazon.com/Python-2nd-Beginners-Hands-Project-ebook/dp/B071Z2Q6TQ/ref=sr\\_1\\_3?dchild=1&keywords=learn+python+in+a+day&qid=1611108340&sr=1-3](https://www.amazon.com/Python-2nd-Beginners-Hands-Project-ebook/dp/B071Z2Q6TQ/ref=sr_1_3?dchild=1&keywords=learn+python+in+a+day&qid=1611108340&sr=1-3)
4. Theodore G. Cleveland, Farhang Forghanparast, Dinesh Sundaravadivelu Devarajan, Turgut Batuhan Baturalp (Batu), Tanja Karp, Long Nguyen, and Mona Rizvi. (2021) Computational Thinking and Data Science: A WebBook to Accompany ENGR 1330 at TTU, Whitacre College of Engineering, DOI (pending)

[ ]: