Report on Learning Python

Doyeon Kim

1927550

Contents:

1. Introduction to Python
2. Features in Python
   1. Variables, expressions and statements
   2. Conditional execution
   3. Functions
   4. Loops and iterations
   5. Strings
   6. Files
   7. Lists
   8. Dictionaries
   9. Tuples
   10. Regular expressions
   11. Network programming
   12. Using web services
   13. Object-oriented programming
   14. Data visualization
3. Conclusion
4. Introduction to Python

Python is a high-level, general-purpose programming language. Guido van Rossum invented it in 1991, and the Python Software Foundation developed it. Its syntax lets programmers to express concepts in fewer lines of code, and it was primarily built with code readability in mind. Python is a programming language that allows you to operate more quickly and efficiently with systems. Python is a high-level scripting language that is interpreted, interactive, and object-oriented. Python is intended to be a very readable language. It typically uses English terms instead of punctuation, and it has fewer syntactical structures than other languages.

1. Features in Python
   1. Variables, expressions and statements

Constants states the fixed values such as numbers, letters and strings. String constants uses either single quotes or double quotes for expression, where numeric constants are simply expressed with the number itself. However, variable names or identifiers are unable to be expressed with reserved words such as false, none and true.

A variable is a named place in the memory where the users can store values and retrieve using the name assigned. The contents of a variable can be changed when assigned from the original one. When assigning, the name is always located on the left side of the equal sign, and the value to set is on the right side. The equal sign is therefore called an assignment operator. When a new variable is assigned to a location storing a value, it replaces the originally stored variable. Naming the variables must start with a letter or underscore, and it is case sensitive that it differentiate the capital and lower case letters.

There are six basic mathematical operators. They are also called as a numeric expressions, and the majority of them are expressed equivalently to the pure math. As from the table below, addition, subtraction, multiplication and division symbols are the same with the operators used in pure math, where power and remainder are different.

|  |  |
| --- | --- |
| Operator | Operation |
| + | Addition |
| - | Subtraction |
| \* | Multiplication |
| / | Division |
| \*\* | Power |
| % | remainder |

Table 1. Basic Numeric Expressions Provided in Python

As equivalent to the pure math, the operator precedence rules are parenthesis, power, multiplication, addition, and left to right. It is from the most significant rule to the lowest significant rule. Therefore, when writing codes, it is highly recommended to use parentheses to separate the significant variables.

Python variables, literals, and constants have a type that specifies such as whether it is an integer or a string. It is important to separate the variable types since certain operations are prohibited such as adding an integer and a string. Because of such implications, converting types is necessary. It can be controlled with built-in functions int() or float(). The usage converts between strings and integers, but it generates an error if the string contains one or more non-numeric characters. Moreover, Python is available to read data from the user inputs. The inputs are stored as a string, and it also can be converted to integer or float if the condition meets. To be more specific, the string data type is a sequence of characters that can be concatenated with the operator +, and sliced by selecting the ranges. The string type in Python is stored by single character of a string and can be retrieved with square brackets. The index value of the characters in the string begins from 0, such that (the last index of the character) +1 gives the length of a string.

The codes are ignored if they are written after a sharp symbol “#” in Python. The purpose of the comment is to describe the sequence of code, document who wrote and turn off a line of code. It majorly demonstrates the code specifically that let programmers easier to understand other’s codes.

* 1. Conditional execution

The comparison operators are used to compare and generate one result between yes or no. They are therefore used and evaluated for Boolean expressions. The decision makings with the use of Boolean expressions can be separated in three different categories: one-way, two-way and multi-way decisions. They are compared with if-then statements. In one way decision, only “if” is utilized to for conditions, where other decisions use both “if” and “else” or “elif” for true and false logical expressions. The “elif” is additionally used for multi-way decision to indicate and compare with more conditions. It is important to note that “elif” is contrasted when “if” statement is false.

When using if function, it is also possible to compare the variables using “is” and “is not” operators. Python provides an “is” operator for logical expressions, which is similar to equal comparator ==, but stronger.

Aside from if-then or if-else statements, Python also provides a function that tests a block of code for errors, try-except structure. At try state block, Python first execute the code then check if there is an error created. It passes to the except block if any kind of error is detected. From the except block, it catches the error and provide solution as a safety net.

* 1. Functions

Functions in Python have two different kinds. One is built-in functions that are provided by Python such as print(), input(), type(), float() or int(). Another one is functions that can be defined with by the programmer. The functions are stored and named by the use of reserved word, “def”. When defined, the code inside the function is reusable. Some functions take arguments or parameters as an input for computations and returning results. However, functions do not always return a value. Such kind of function is called a void function.

Functions that operate on other functions are called higher-order functions. The built in higher-order function in Python is max function. As obvious from its name, it returns the largest of its inputted arguments. Not only the numerical values, it compares the string and returns the highest order from the alphabet.

There are several objectives of functioning certain lines of code. They are to organize the code into paragraphs and to not to repeat code every time. When the long and complex code is chunked and paragraphed, it is better to comprehend. Furthermore, making a library of common things reduces the effort of repeating.

* 1. Loops and iterations

Loops are repeated steps, that continues until a Boolean of an execution does not meet the loop condition. It can be reached with the loop called “while”. However, if the iteration variable inside the loop is not changing, the loop will never end although the condition is well specified, making an infinite loop. It is thus very important to break the statement if the end of loop is never reached. Another way to finish an iteration is using continue statement. The only difference between continue and break statements is that continue jumps to the top of the loop, and break jumps out of the loop and proceed to execute the next line.

The previously explained loop is indefinite loop. In contrasting to this, definite loops are the loops executed with an exact number of times, and can be ran using the construct called “for”. Definite loops are necessary to have changing explicit iteration variables each time. These iteration variables executes through a sequence or set. The iteration variables can also be replaced with a range function that returns a sequence of numbers. With the usage of range(n), the for loop will continue to execute until when the loop is repeated n times.

* 1. Strings

According to the previous explanations from the variable types, string stores the non-numerical values. In this section, it will demonstrate the library of string. The functions are implemented in every string, and they return a new string without modifying the original string. The library functions contain: capitalize, center, endswith, find, lstrip, replace, lower, rstrip, strip, and upper.

* 1. Files

In Python, it is available to open, read, write and close files. The files are handled by stating specific name of the file and setting the mode. There are different modes of opening files. “r” means that it will be opened in reading mode, where “w” means that it will be opened in writing mode. When the file is opened in reading mode, it is unable to add or overwrite things.

It is significant to note that to read each line of a file, a for loop with “in” is used. As an example, “for line in file:” executes each line in the file in string format.

* 1. Lists

List contains a collection of variables. They are represented by square brackets around the separated elements by commas. In Python, mixture of variables can be stored in one list. Lists are similar to strings since single element can be executed by stating the index in square brackets. The only difference is that strings are unchangeable, but lists are mutable that the element at certain index can be changed.

Because list is similar to string, the functions such as len(), max() and range() operates the same as they are applied to string. It is also equivalent that lists are able to be concatenated with addition operation. Likewise, splitting lists are also available by setting the ranges or using split function. However, there is an additional built-in function that is only applied to lists: append. The append method add elements to empty lists.

* 1. Dictionaries

Dictionaries are data structure for mapping keys to values. They allow users rapid searching operations. In dictionaries, mapped collections are separated by commas. The keys are located on the left of colons, and the corresponding values are on the right side for each collection. Because of such positionings, dictionaries have no order.

Unlike others, dictionary is unable to find the values by referencing a key in the dictionary. To look up properly, it is necessary to use “in” operator. However, it only checks the existence of the key. To retrieve both key and value stored, “get” method is required. The get method, on the other hand, does not produce any error as it needs a default value provision.

* 1. Tuples

Tuples function like lists that arrange elements in another kind of sequence. The only divergence of tuple with list is that it is immutable, and the built-in functions of list do not work with tuples. Nevertheless, since it is unmodifiable, tuples are more efficient than lists in terms of memory usage.

Tuples are available to be put on the left-hand side of an assignment statement. For instance, “(x, y) = (99, 98)”. It is therefore able to put tuples in dictionaries, and comparison operators work with it as well. Not only that, but it also has the ability to sort a list of tuples by the descending order with the key or values in dictionaries. The operation can be done with built-in “sorted” function. It takes a sequence as a parameter and returns a sorted sequence.

* 1. Regular expressions

The use of regular expressions provides particular characters, words, or patters of a string. It is also known as regex. Their major purpose of the usage is matching strings and extracting elements from a string. The functions of regex are “search” that checks the matching string with the regular expression, and “findall” that extracts portions of a string that matches the regex parameters. In Python, to employ regular expressions, it is essential to import the library “re”.

|  |  |
| --- | --- |
| ^ | Matches the beginning of a line |
| $ | Matches the end of the line |
| . | Matches any character |
| \s | Matches whitespace |
| \S | Matches any non-whitespace character |
| \* | Repeats a character zero or more times |
| \*? | Repeats a character zero or more times (non-greedy) |
| + | Repeats a character one or more times |
| +? | Repeats a character one or more times (non-greedy) |
| [aeiou] | Matches a single character in the listed set |
| [^XYZ] | Matches a single character not in the listed set |
| [a-z0-9] | The set of characters can include a range |
| ( | Indicates where string extraction is to start |
| ) | Indicates where string extraction is to end |

Table 2. Regular Expressions

When writing the regular expressions, it is necessary to be careful of greedy and non-greedy matchings. First of all, greedy matching refers to repeat characters \* and + push outward in both directions to match the largest possible string. In opposition, the non-greedy matchings compare the characters as few as possible. Hence, each results will generate as the following figure.

텍스트, 스크린샷, 모니터, 실내이(가) 표시된 사진

자동 생성된 설명

Fig. 1, Result Generation of Greedy and Non-greedy Matching

* 1. Network programming

In network programming, there are two main ways to connect the computer to a server. They are Transport Control Protocol, simply known as TCP and User Datagram Protocol, UDP. In this report, it will focus on TCP connection. TCP connection is a connection that is built on top of Internet Protocol, asks the server every time for accessing permission. To connect to the internet, an internet socket or network socket is used at the endpoint of the communication flow. When it successfully obtains the permission, the network provides the port number to the devices. Under TCP/IP control, there is an application layer protocol. It defines the system that runs on separate end systems which passes a message between application processes (clients and servers). One of the application type is World Wide Web, and the layer protocol is Hypertext Transfer Protocol, known as HTTP. It is dominant on the internet and invented only for the Web.

With such network programming method, there are procedures to follow to retrieve data from the server. Each time the user clicks on the link of a new page, the browser makes a connection to the web server and issues a “GET” request. If the request is approved, then the server responds by returning the HTML document to the browser, letting the user to access the web page. When accessing to the web, every images or documents are sent or downloaded by bytes. Because of the small number, it is required to convert the strings into the networking standard that the server can understand. This is called encoding and decoding, which is changing the Unicode strings into UTF-8 bytes or vice versa. The HTTP request in Python can be done by the following diagram.

텍스트, 모니터, 스크린샷, 실내이(가) 표시된 사진

자동 생성된 설명

Fig. 2, HTTP Request in Python

However, since HTTP is so common, Python has a library that does all the socket work automatically and makes web pages look like a file, which can be written as figure n.

텍스트, 모니터, 스크린샷, 실내이(가) 표시된 사진

자동 생성된 설명

Fig. 3, Built-in HTTP Request Library in Python

* 1. Using web services

In this section, the report will demonstrate the commonly used formats in HTTP request and response. They are extensible markup language, XML, and JavaScript object notation, JSON. The formats are used since an agreed way to represent data is needed between applications and across networks. The primary purpose of extensible markup language XML is to help information systems share structured data. The basic XML contains five structures: start tag, end tag, text content, attribute and self-closing tag. From the structures, the tags indicate the beginning and ending of elements, and attribute defines keyword or value pairs on the opening tag on XML. JSON, however, is more readable and writable data-interchange format compared to XML for humans. It thus represents data as nested lists and dictionaries.

In some web services provide application program interface (API) implemented interfaces, which is a programming language used to build an application. These APIs are not free in such websites, thus data providers issue limited number of API keys per day, or even payment is needed for usage. The web services provide infrastructure for networked apps to communicate with one another, and SOAP and REST are two approaches of the web services.

* 1. Object-oriented programming

Object oriented means a program is made up of many cooperating objects. It is made up of one or more objects working together to make use of each other’s capabilities. The type of programming has: class, method or message, field or attribute, and object or instance. First of all, class defines the abstract characteristics of an object. It classifies objects such as the class animal containing dogs, cats, or rabbits. Instance in OOP refers to the actual object created at runtime. Lastly, methods are verbs. As an instance, Lassie, a dog, has ability to bark. From this, bark() is one of Lassie’s methods. The methods and messages are often having multiples, like sit() or eat().

In addition to such methods, Python is also available for dir() and type() methods. They are built-in functions unlike others. From the methods, dir() command lists capabilities. When executed, there are two types of methods that are accessible. The ones with underscores are used by Python itself. Therefore, the other methods are the real operations that can be performed. It is similar to type() method, which informs about the variables.

In object oriented programming, it is typical to use constructors. The constructor is a method to set up some instance variables during the object creation. It sometimes renew the instances or set initial values.

Not only them, but OOP also has inheritance. Inheritance is subclasses that inherits the attributes and behaviors to child classes. For example, animal, dog and cat classes in a Python program can make inheritance from animal classes to dog and cat classes.

* 1. Data visualization

In Python, it provides the availability of visualizing the data such as using Google Map or ranking the web pages. Page rank with networks is possible to be visualized with a simple web page crawler, which is a computer software that performs systematic and automated search of the World Wide Web. Additionally, the geographical data can be obtained by using the Google Geodata API. The visualization throw a browser using Google Maps API can represent as the picture below.

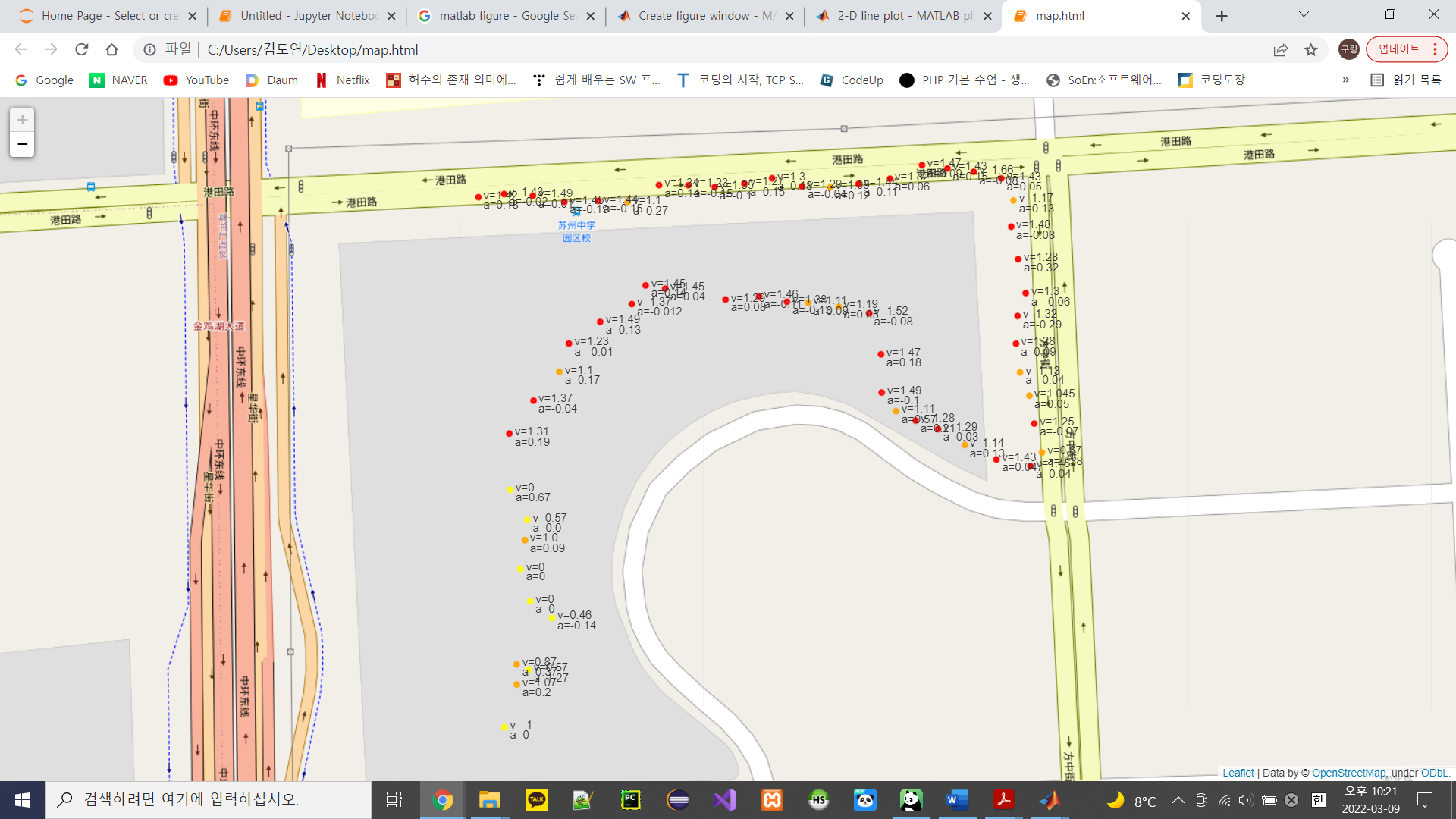


Fig. 4, Geodata of Locations with Velocity and Acceleration Data at Wenhua Apartment in Suzhou

1. Conclusion

Python provides adequately comprehensive and broad functions. It is evident that the programming language Python is a perfect instrument for Machine Learning. Furthermore, it is equally important that Python provides simplicity and consistency. The ease of access to libraries and flexibility of modifications allow building great frameworks for machine learning on Artificial Intelligence.