An Introduction to
Python For Data Science

November, 2017

GSA

D2D
DATA TO DECISIONS

What Is Python?

Python is:

- A High Level
- General Purpose Language
- Object Oriented (with full support for other paradigms)
- **Interpreted**
- Created in the early 1990's
- Python 3.6 is the current version

GSA

D2D
DATA TO DECISIONS

1.Interpreter, modules and variables

Simple Hello World Program

1.Interpreter, modules and variables

- Python Interpreter Vs Python modules
  - ➢ Use Interpreter for testing or quick experiments
  - ➢.py files – modules that can be rerun, need a Python interpreter to be executed
- Variables
  - ➢Get assigned values
  - ➢Can be reused, manipulated, reassigned … etc.
- Printing
  - ➢Visual output of your code – is It working as it should be?

GSA

D2D
DATA TO DECISIONS

2.Variables and Operators (and comments)

2.Variables and Operators (and comments)

- Variables
  - ➢ No type declaration necessary (Python figures out the type)
  - ➢ first assignment created the variable
  - ➢Assignment is done using “=“

- Operations
  - ➢Carried out on variables (operands)
  - ➢Operator can behave differently based on the data type
    - + Adds Integers, concatenates Strings
  - ➢Strongly-Typed is the way! (No implicit type conversions)

- Comments Keep code clean and readable
  - ➢Whoever reads your code will thank you!
  - ➢# used to comment a single line

GSA

D2D
DATA TO DECISIONS

3.More on Variables

3.More on Variables

- Multiple Assignment
  - x,y=4,6
- Basic Data Types:
  - Integers (Default for Numbers) 5, 17 , 3000
  - Floats :5.3, 7.324, -34.11, 5/2
  - Strings: “Bob”, ‘John’, “Kevin’s”, “””Mark’s car is “black”””””
- Variable names are :
  - **Case Sensitive!**
  - Can **NOT** start with a number
  - Can contain underscores, letters, numbers
  - Can NOT be a reserved word (if, elif, global, return, pass, import …..etc.)

GSA

D2D
DATA TO DECISIONS

3.More on Variables (Continued)

Additional Notes:

- Python binds variables to **object references**
  - ➢Assigning a variables created a reference to an object, NOT a copy of the object
- A variable name does not imply the object type, the object referenced does
  - ➢X=7 , X="Bob" is completely fine.
- Some datatypes are **mutable**, some are **immutable**

GSA

D2D
DATA TO DECISIONS

4.Mutable Vs Immutable

4.Mutable Vs Immutable

REMEMBER:

- Python binds variables to **object references**

Mutable:

- content of objects of immutable types can be changed after they are created
- More memory is assigned than needed
- Support methods that change the object in place
- Examples: list, set, dict

Immutable:

content of objects of immutable types cannot be changed after they are created

Hashable!

- Examples: tuple, frozenset, float

5.More Data Types

5.More Data Types

- Tuples
  - ➢A collection of “Elements”
  - ➢Can be sliced
    - Elements Accessible individually using [n] or [-n]
    - Ranges [1:2], [:2], [2:], [1:-1], [:]
  - ➢Elements cannot be changed (**immutable**)
  - ➢Check for element presence using “in” clause
- Lists
  - ➢Like a Tuple, but with added functionalities
  - ➢Slower but more useful
  - ➢Elements can be inserted, appended, removed, deleted, “popped” and changed
- Use **len(x)** to find length, **x.index(n)** on lists and tuples to “know your way”
- A string is also a sequence type, closer to a tuple (immutable)!

GSA

D2D
DATA TO DECISIONS

5.More Data Types (Continued)

- Sets (and Frozen Sets)
  - ➢An Unordered collection of “**Unique**” and “**Immutable**” objects
  - ➢Items Cannot be accessed using an index
  - ➢Sets are **mutable** while frozensets are **immutable**

- Dictionaries
  - ➢Unordered Key Value Pairs
  - ➢Keys have to be **immutable**

GSA

D2D
DATA TO DECISIONS

5.More Data Types (Continued)

- You can “**Add**” sequences:
  - [1,2,3]+[4,5,6]
  - (1,2,3)+(4,5,6)
  - "Hello"+" "+"World! " (Look familiar?)

- You can “**Multiply**” a sequence and an integer:
  - [1,2,3]*3
  - (1,2,3)*2
  - "Hello"*3

GSA

D2D
DATA TO DECISIONS

6.Conditionals and loops

6.Conditionals and loops

- Code blocks are identified using **Indentation** **(no { } here!)**
  - ➢Standard is 4 white spaces – tabs not recommended

- Conditions can be evaluated using **if, elif, else** statements
- = used for assignment, == used for comparison
- != is the opposite of ==

- Loops allow you to execute a block of code several times using **while** or **for..in**
- Else condition in loops are executed when condition is false
- Stop a loop using break
- **Watch out for infinite loops!**

GSA

D2D
DATA TO DECISIONS

7.Functions

7.Functions

- Functions are defined using the keyword **def**
  - ➢def addition_function(x,y):
- Values are returned using the **return** keyword (even if not present!)
  - ➢None value
- Functions take arguments
- A function can be an argument to another function
  - ➢addition_function(3,addition_function(3,5))
- No types are defined for arguments or return types
- Functions can call other functions
- Objects have scopes

GSA

D2D
DATA TO DECISIONS

8.Scopes

GSA

D2D
DATA TO DECISIONS

8.Scopes

- Objects have scopes
- Be careful of what you are trying to reference
- Use of **return** to make an object available

GSA

D2D
DATA TO DECISIONS

9.Modules

GSA

D2D
DATA TO DECISIONS

9.Modules

- A module gains access to code in another module by importing it
- Modules provide a way of code reuse
- Python comes with a library of standard modules
  - ➢ Such as the datetime module
  - ➢ …or the statistics module
    - Import statistics
    - print(statistics.mean([1,2,3,4,5,6]))

GSA

D2D
DATA TO DECISIONS

Questions?

GSA

D2D
DATA TO DECISIONS

Resources

- Official Python Docs

https://www.python.org/doc/

- Free eBook : O'REILLY's A WhirlWind Tour of Python

http://www.oreilly.com/programming/free/a-whirlwind-tour-of-python.csp

- Python Beginners Guide Wiki

https://wiki.python.org/moin/BeginnersGuide/Programmers

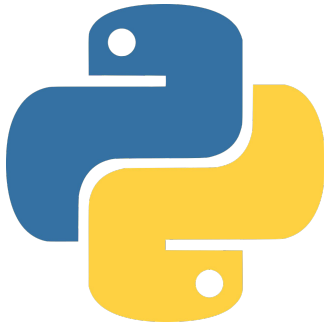
- Non-Programmer's Tutorial for Python 3

https://en.wikibooks.org/wiki/Non-Programmer%27s_Tutorial_for_Python_3

GSA

D2D
DATA TO DECISIONS

Thank You



GSA

D2D
DATA TO DECISIONS