DS 542 - Python in Data Science

Fall 2021, Week 2

Recap of Last Week

- Recording in Blackboard “Week 1” folder under “Resources”
- Last week we learned about:
  - The History of python
  - How to install and setup python (Poll)
  - The basics rules/standards of python
  - What are third party libraries? (Poll)
  - What are pythons capabilities?
    - What is python good for?

Week 2 Agenda:

1. What are Data Types in python?
2. Data Type: Integer
3. Data Type: Float
4. Python built-in function: Range
5. Data Type: String
6. Equivalence, changing data type, and input()

Week 2 Agenda:

1. **What are Data Types in python?**
2. Data Type: Integer
3. Data Type: Float
4. Python built-in function: Range
5. Data Type: String
6. Equivalence, changing data type, and input()

What are Data Types in python?

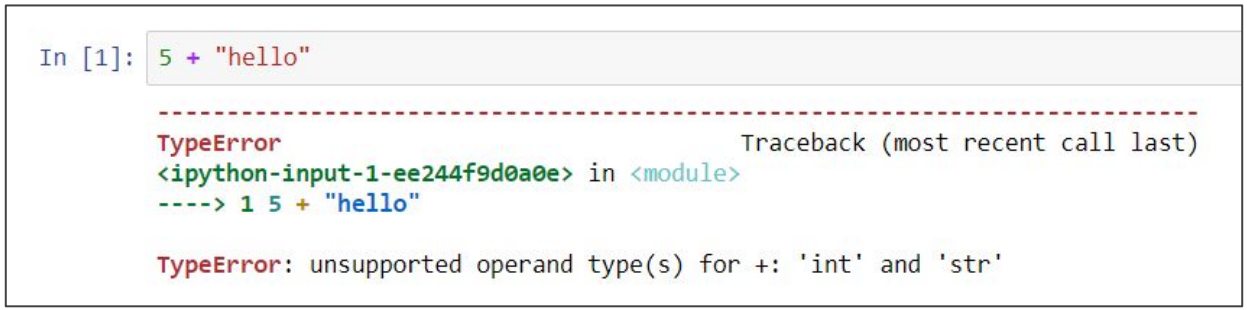
- Data types are a way to categorize of classify data in python
  - For example, the data type of the number “5” is called an “integer”
  - The data type of the word “hello” is a called a “string”
  - The data type of the number 3.14 is called a “float”
- There are many data types in python:
  - integer, float, string, boolean, list, tuple, range, dictionary
    - For all data types including less common ones, see this link: https://www.w3schools.com/python/python_datatypes.asp

All Data Types in python

- There are many data types in python, the most common are:
  - integer
  - float
  - string
  - boolean
  - list
  - tuple
  - range
  - dictionary
- To see all data types including less common ones, see this link: https://www.w3schools.com/python/python_datatypes.asp

Why are Data Types important?

- Data types determine the operations you can perform on data
  - For example, if you have two integers you may add them:
    - 5 + 7 = 12
  - If you have one integer and one float, you still may add them:
    - 5 + 7.58 = 12.58
  - But, if you have one integer and one string you cannot add them:
    - 5 + “hello” = error



Week 2 Agenda:

1. What are Data Types in python?
2. **Data Type: Integer**
3. Data Type: Float
4. Python built-in function: Range
5. Data Type: String
6. Equivalence, changing data type, and input()

Data Type: Integer

- In python, integers are positive or negative whole numbers:
  - 1, 6, -7, 549, -943
- The data type integer in python is labelled “int”
- To best show the properties of integers, let’s move to a Jupyter Notebook
  - Please follow along with me to practice in python

Week 2 Agenda:

1. What are Data Types in python?
2. Data Type: Integer
3. **Data Type: Float**
4. Python built-in function: Range
5. Data Type: String
6. Equivalence, changing data type, and input()

Data Type: Float

- In python, floats are positive or negative real numbers with a fractional part denoted by the decimal symbol
  - 2.5, 3.1415, -9.4, 1.0, -209.111
- It is possible to see a float appear with the scientific notation “e”

```
num = 2/1250000
print(num)
```

```
1.6e-06
```

- The data type float in python is labelled “float”

```
In [8]: type(3.14159)
Out[8]: float
```

- Let’s move back to the Jupyter Notebook to explore the properties of floats

Week 2 Agenda:

1. What are Data Types in python?
2. Data Type: Integer
3. Data Type: Float
4. **Python built-in function: Range**
5. Data Type: String
6. Equivalence, changing data type, and input()

Python built-in function: Range

- In python, ***range()*** is a built-in function used to generate a sequence of numbers in a given range
  - range() only needs 1 argument to run, but you may pass up to 3 arguments:
    - start: integer starting from which the sequence of integers is to be returned
    - stop: integer before which the sequence of integers is to be returned. The range of integers end at stop – 1.
    - step: integer value which determines the increment between each integer in the sequence

```
In [11]: for i in range(1, 10, 2):
             print(i)

1
3
5
7
9
```

Let's try using **range()** in a Jupyter notebook

Week 2 Agenda:

1. What are Data Types in python?
2. Data Type: Integer
3. Data Type: Float
4. Python built-in function: Range
5. **Data Type: String**
6. Equivalence, changing data type, and input()

Data Type: String

- In python, strings are sequences of unicode characters enclosed in single quotes or double quotes.
  - ‘Hello’, “goodbye”, “1-800-phone-number”, ‘test.csv’
- The data type string in python is labelled “str”
  - “str” is actually a built-in python class, so it has many methods
- Strings have many more interesting and complicated properties than integers and floats
- Let’s continue in the Jupyter Notebook

Week 2 Agenda:

1. What are Data Types in python?
2. Data Type: Integer
3. Data Type: Float
4. Python built-in function: Range
5. Data Type: String
6. **Equivalence, changing data type, and input()**

Equivalence, changing data type, and input()

- **Equivalence**: are two variables equal?
- **Changing data type**: force a variable to a specified data type
- **input()** - let the user input the assignment of a variable
- **Variable names best practice**: underscores not camelCase

Let's try all the above in a Jupyter notebook

PEP 8: Style Guide for python Code

- To write code according to python best practice (or “pythonic” code), follow the official PEP 8 guide:
  - https://www.python.org/dev/peps/pep-0008/#introduction
- You will also find many other great resources for new python users at https://www.python.org

Class Follow-ups and Homework:

1. Week 2 Discussion
2. Week 2 Assignment
3. Reading
    3.1. Please read chapters 1 and 6 from your textbook: Automate the Boring Stuff with Python by Al Sweigart