Introduction to Programming

Practical Class #3

João Fonseca
joao.fonseca@novasbe.pt

Agenda

- What is the purpose of programming?
- Recap from last week
- What is a library?
- Exercises

What is the purpose of programming?

- Programming is used for immensely broad applications
- Some examples of useful applications for business and/or economics include:
  - Data Science (e.g., Machine Learning and Data Mining)
  - Statistic modelling (as well as Algebra operations ranging from very basic to very advanced)
  - Web Scraping (retrieving data from websites for further analysis)
  - Software development (mobile apps, games, general purpose software)
  - Front end development (build websites)
  - Back end development (development that doesn't involve a GUI)
  - …

What is the purpose of programming?

- **Examples of projects:**
  - Project developed last summer at Data Science for Social Good using Python and R:
    - http://dssg-eu.org/tuscany/
  - Machine Learning and Artificial Intelligence:
    - http://www.hansonrobotics.com/robot/sophia/
    - https://www.youtube.com/watch?v=LSHZ_b05W7o
  - (Ethical!) Hacking:
    - https://github.com/Manisso/fsociety
  - Some of my own projects…

Last week's recap

- Terminal commands:
  - sudo – run as administrator (stands for **su**peruser **do**)
  - pwd – **p**rint **w**orking **d**irectory
  - ls – **l**i**s**t contents
  - rm – **r**e**m**ove file/directory
  - cd – **c**hange **d**irectory
  - touch – create file
  - mkdir – **m**a**k**e **dir**ectory

Last week's recap

- Some built-in Python functions:
  - print()
  - input()
  - Math operations: 1+1, 2-1, 3*3, 3/4
  - Logic operations: 2==8/4, 3>1, 6<=10 …
  - int(), str(), float()
  - list(), len(), range()
  - min(), max(), sum()
  - type()

What is a library?

- A library is a collection of functions and methods that allows you to perform many actions without writing your own code
- These functions and methods are optimized to perform given actions and will save you time
- This means you can efficiently perform complex tasks with a few lines of code, instead of writing them yourself

What is a library?

- The main libraries we will be using in this course are:
  - Pandas – Data structures and data analysis tools
  - Matplotlib – Data visualization library
  - Math – Access to mathematical functions
- We'll be using a few others for some utilities, such as:
  - Numpy – Scientific computing
  - Datetime – Date formatting
  - Pandas_profiling – Profiling reports from pandas DataFrame objects
  - And others on an as needed basis

What is a library?

- In order to use a library you will need to import it into your script.
- There are 4 ways to do this in Python:
  - **import <library>**
  - **import <library> as <new name>** - imports library and renames it
  - **from <library> import <function>** - imports only the functions you will need from the library
  - **from <library> import <function> as <new name>** - imports function and renames it. Not as commonly used.

Let’s write our first Python script!

- Exercise:
  - Open Spyder
  - Create a script: 01_my_first_script.py
  - Program it to print “Hello World!” and save the file
  - Run it in the IDE (reminder: IDE=Integrated Development Environment)

Another exercise!

- Exercise:
  - Open Spyder
  - Create a script: 02_using_length.py
  - Request user input: ‘What are you thinking about?’ (**tip:** use “\n” by the of the string to insert a paragraph)
  - Print the number of characters in the passed input as well as the number of words:
    - “Your thought has … characters and ... words!”
  - Once finished, close the IDE and run the script in the terminal

Intro to Pandas

- Pandas is an open source library
- Provides **high-performance and easy to use** data structures and data analysis tools
  - Allows to user to efficiently perform complex operations on data structures (such as a table or column)
- The data structures used by pandas are **Series** (known as a column - 1 dimensional array) and **DataFrames** (known as a table – generally a 2 dimensional, size mutable, tabular structure)
- Capable of **reading and writing plenty of data formats**, such as CSV (Comma Separated Values) and Microsoft Excel files

Intro to Pandas

- Sample dataset:
  - https://www.kaggle.com/caganseval/earthquake
- A 10 minute live demo from the creator of Pandas:
  - https://vimeo.com/59324550
- Let's open a Jupyter notebook and see how it works!