Welcome to Programming with data

Lab 1

Wednesday, 7 Nov. 2018

Stelios Sotiriadis

Prof. Alessandro Provetti

Portable environments

- Atom
- Sublime
- PyCharm Community Edition
- PySpark
- **To do:**
  - Download your favourite editor
- **To do:**
  - Enable scripts, enable Python execution
  - Run simple Python examples

About Python 2.7

- We will stick to Python 3, standard for Data Science modules

```
print(1/3)
```

- Remember: Advanced architectures implement $\frac{1}{3}$ directly, alongside π and e.

```
from future import division
print(1/3)
```

Exercise 1: Python crash course

- Go through the Crash Course in Python by Joel Grush:
  https://www.oreilly.com/learning/data-science-from-scratch

Exercise 2: Complexity

- Develop four functions as example cases for the following time complexities:
  - O(1)
  - O(n)
  - $O(n^2)$
  - $O(n^3)$ -> Multiply two matrices
    - (See an example: https://www.mathsisfun.com/algebra/matrix-multiplying.html)
  - Hint: Examples will be at GitHub by tonight!

Exercise 2 with Hints: Complexity

- Develop four functions as example cases for the following time complexities:
  - O(1) -> e.g. to access an element of a list
  - O(n) -> e.g. to search for an item (linear search of a list)
  - $O(n^2)$ -> Search for the max values for each row “n” of a “n x m” matrix, return a list of max values for each row “n”.
  - $O(n^3)$ -> Multiply two matrices
    - (See an example: https://www.mathsisfun.com/algebra/matrix-multiplying.html)
  - Hint: Examples will be at GitHub by tonight!

Exercise 3: Linear search and bubble sort

- Develop the following Python functions:
  - LinearSearch:
    - Given a two dimensional list of “n x m” elements, write a function to count occurrences of a given key. The function should also return a matrix of the position of a key in the 2 dimensional list.

e.g.

```
Mylist = [  0 1 2
        0  [1,4,7],
        1  [4,5,2],
        2  [6,4,8]
        ]
key = 4
```

Output:
3, [[0, 1], [1, 0], [2, 1]]
3 occurrences of number 4 in positions: 0,1, 1,0 and 2,1

    - What is the complexity of the linear search in the 2 dimensional array?
  - Develop BubbleSort (hint)
    - It is the simplest sorting algorithm that works by repeatedly swapping the adjacent elements if they are in wrong order.
    - Implement a simple bubble sort algorithm!
    - What is the complexity of Bubble sort?

- Hint: Examples will be at GitHub by tonight!

6 5 3 1 8 7 2 4