

Computer Science 134C

Introduction to Computer Science, in Python

Lecture #15 (Plotting Data)

October 15, 2018

Keywords csv, matplotlib, virtual environment

We start thinking about data presentation.

1. Office hours tonight, 7-9pm. Exam tomorrow: 6-7:30pm or 7:30-9pm, here.
2. Questions?
3. Simple plotting of data in Python. Many great packages; we'll use matplotlib.

- (a) System must be explicitly installed on your machine. It does not get shipped with Python. Use `pip` in a *virtual environment*. Details in lab.
- (b) Import the package, using this idiomatic expression:

```
import matplotlib.pyplot as plt
```

When we import `x`, all the symbols (e.g. `s`) are imported into the current environment and are referred to using the *name space* syntax: `x.s`. Because some packages have names that hard to type, or conflict with symbols in your scripts, the `as y` suffix allows you to refer to `s` using the symbol `y.s`. After this import, all plotting package symbols start with `plt`.

- (c) Typically, plots are constructed piecemeal as the symbol `plt`. The configuration of the plot is determined by several method calls to `plt`.
 - (d) To set the plot title to string `s`, use `plt.title(s)`.
 - (e) To set the axis labels, use `plt.xlabel(s)` or `plt.ylabel(s)`.
 - (f) To perform the simplest plot, for each data point collect the `x` and `y` values separately. Then use a command like `plt.plot(xl,yl,'r+')`, where `xl` and `yl` are lists and `'r+'` indicates you want to plot with red crosses. To plot with blue lines, use `'b-'`. More help on the format strings can be found in `help(plt.plot)`.
 - (g) To save the plot in a file, use `plt.savefig(filename)`. We'll use pdf files.
4. Application: Plotting the distribution of word lengths in the dictionary.
 5. Reading CSV files. CSV files are a common format for databases and spreadsheets.
 - (a) First you need to import `csv`
 - (b) The `csv` package provides a reader and a writer. We'll focus on the reader, here.
 - (c) To read the rows of a csv file, use

```
for row in csv.reader(open(filename)):  
    ... process row ...
```
 - (d) Each row is returned as a list of *strings*. You must explicitly convert to the type you wish.
 6. Example: Plotting sports data.