Computing for Medicine: Phase 3, Seminar 1 Project

Jennifer Campbell
Associate Professor, Teaching Stream
campbell@cs.toronto.edu

Seminar 1 Project

- The project handout, starter code, and data are posted:
  - http://http//c4m.cdf.toronto.edu/cohort2/phase3/

- Two options for doing your work:
  - Use the Computer Science Teaching Labs computing network from your computer (internet connection required).
  - Use only your computer.

WORKING ON CS TEACHING LABS NETWORK

CS Teaching Labs Account

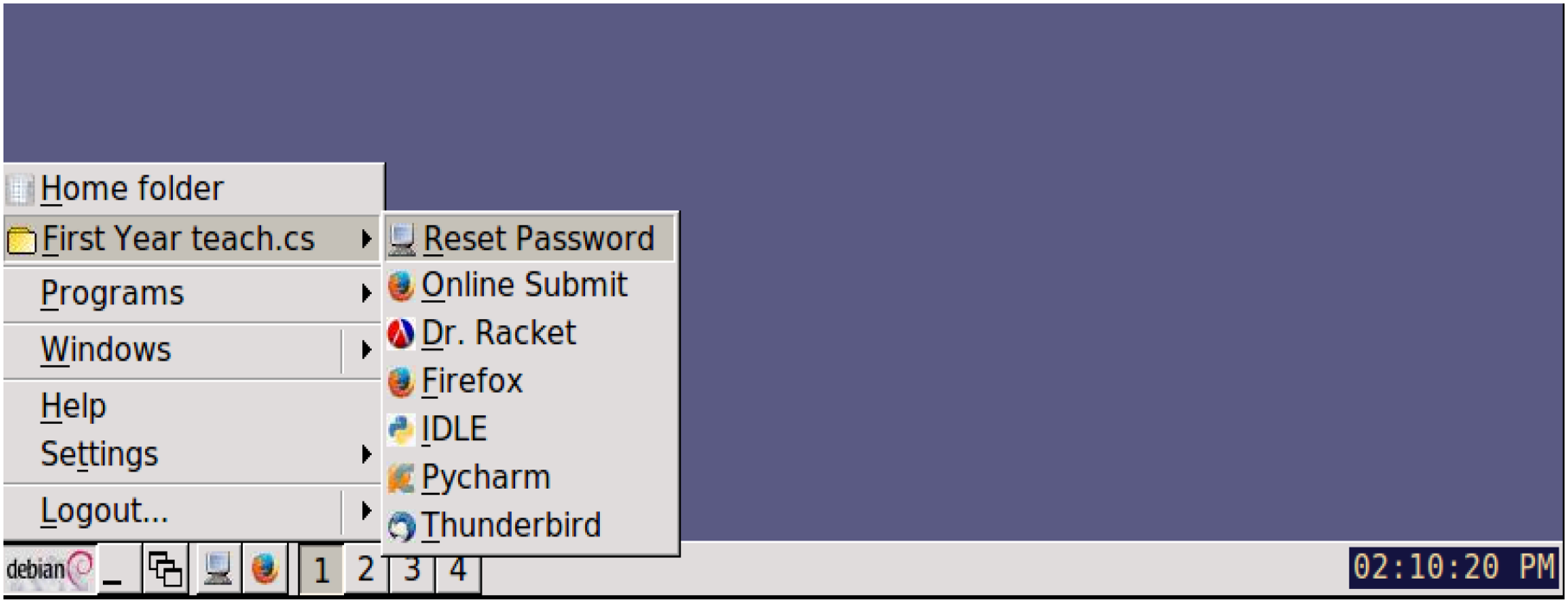
- Information you will need:
  - **Username**: your UTORid
  - **Password**: See email "C4M: Phase 3 account information" for your temporary password.
  - **Key**: For the cdf_nx_key.key file, see:
  - http://www.teach.cs.toronto.edu/using_cdf/remote_access_server.html

Connecting to CS Teaching Labs

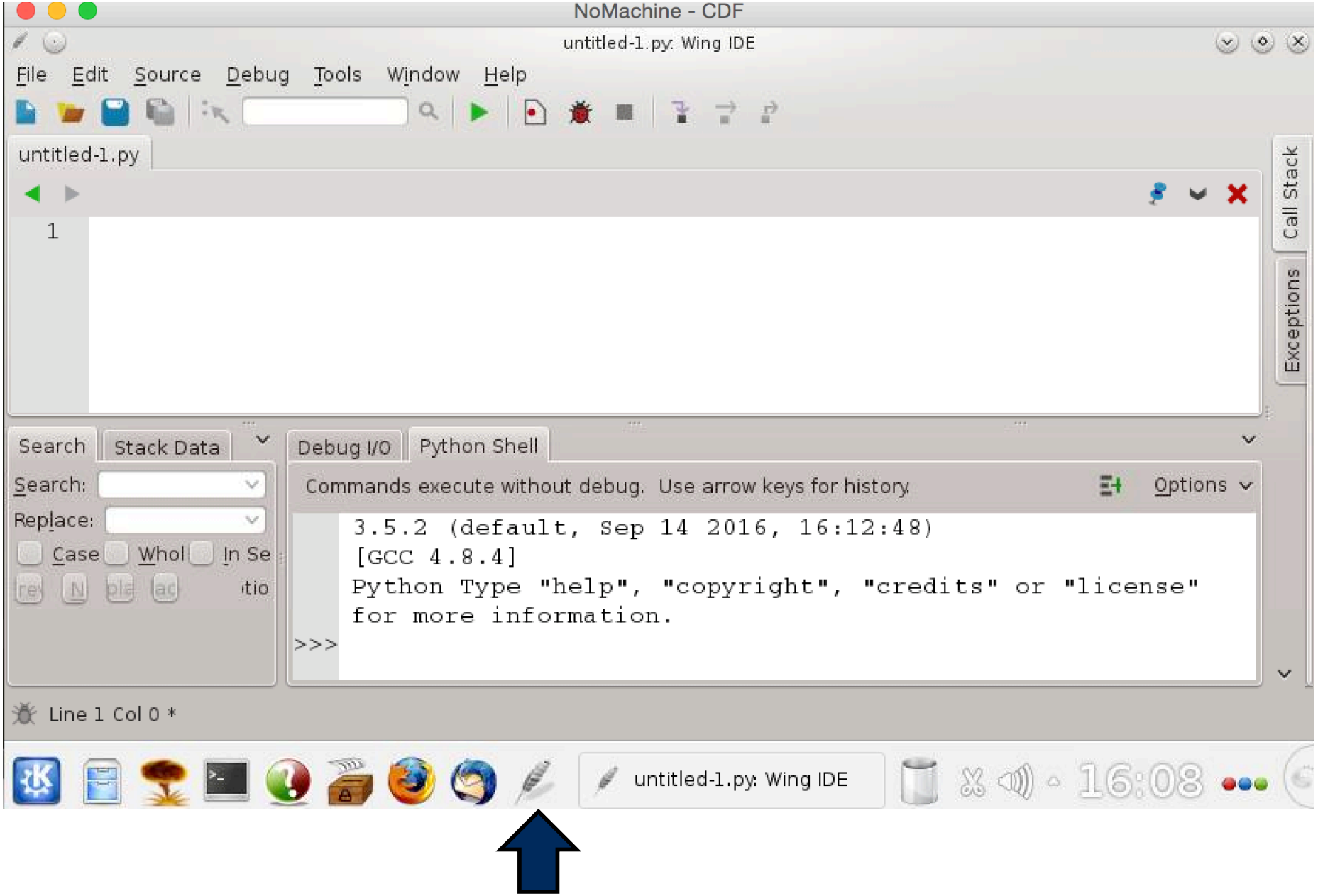
- To connect to the CS Teaching Labs from your computer, please follow the instructions here:
- http://www.teach.cs.toronto.edu/using_cdf/remote_access_server.html

Changing your password

- Once you've logged in to the CS Teaching Labs, please change your password.
- At the bottom of the screen, to change your password use:
  - Application Launch Menu -> First Year teach.cs-> Reset Password



Wing 101 IDE


NoMachine - CDF
untitled-1.py: Wing IDE
File Edit Source Debug Tools Window Help
untitled-1.py
1
Call Stack
Exceptions
Search
Stack Data
Search:
Replace:
Case
Whol
In Se
Debug I/O
Python Shell
Commands execute without debug. Use arrow keys for history.
Options
3.5.2 (default, Sep 14 2016, 16:12:48)
[GCC 4.8.4]
Python Type "help", "copyright", "credits" or "license"
for more information.
>>>
Line 1 Col 0 *
untitled-1.py: Wing IDE
16:08

WORKING ON YOUR OWN LAPTOP

Software installation

- To work on your own laptop, you will need to install the following software:
  - http://www.nltk.org/install.html
  - http://scikit-learn.org/stable/install.html

For Windows users, follow the instructions on the websites above.

For Mac OS X users, open a Terminal window and type these commands (you'll need Python 3.4 or higher installed first):

```
sudo pip3 install –U numpy
sudo pip3 install –U nltk
sudo pip3 install –U scipy
sudo pip3 install –U scikit-learn
```

Download data (and tagger)

- Download the ZIP file that contains the project starter code and data from the C4M website:
  http://http//c4m.cdf.toronto.edu/cohort2/phase3/

- (Optional – not needed unless you want to run the preprocessing code yourself)
  - Download the Stanford Log-linear Part-Of-Speech Tagger from http://nlp.stanford.edu/software/stanford-postagger-2015-12-09.zip and save it in the same directory as the starter code and data.
  - Unzip the file stanford-postagger-2015-12-09.zip

MODULES & PACKAGES

Modules and Packages used in project

- `nltk` (Natural Language Toolkit)
  - http://www.nltk.org/api/nltk.html
- `csv` (Comma Separated Values files)
  - https://docs.python.org/3/library/csv.html
- `math` (Mathematical functions)
  - https://docs.python.org/3/library/math.html
- `numpy` (Numerical Python)
  - https://docs.scipy.org/doc/numpy/reference/routines.html

Click on the links above for documentation for each module/package. You can also use `dir()` and `help()`.

NumPy

- A scientific computing package for Python.
- For this project, you'll use NumPy's N-dimensional array.
  - NumPy's 2D array vs Python's nested lists
    - NumPy's array may contain only elements of the same type, whereas Python's lists may contain different types.
    - NumPy's arrays are more efficient and take less space.
    - NumPy supports a variety of array operations.

NumPy 2D array demo

```
>>> import numpy as np
>>> my_array = np.array([[1, 2, 3], [4, 5, 6]])
>>> my_array.shape
(2, 3)
>>> my_array.size
6
>>> my_array.sum()
21
>>> my_array.min()
1
>>> my_array.max()
6
>>> my_array.mean()
3.5
>>> my_array.var()
 2.9166666666666665
>>> my_array.std()
1.707825127659933
```

UPCOMING SEMINARS

Seminar 2: Dr. Chris McIntosh

- Tuesday October 17 2017 4-6pm
- Location: DCS Innovation Lab
- Topic: Computer Vision
- https://www.researchgate.net/profile/Chris_Mcintosh

SEMINAR FEEDBACK

Phase 3, Seminar 1 Survey

- Tomorrow, you will be invited to complete a short survey.
- We would like your (anonymous) feedback on the seminar.