Spring 2018
GNET 742: Introduction to UNIX and Python Programming for biomedical data analysis

An introductory course designed to present the fundamentals of UNIX operating system followed by introduction to Python programming with an emphasis on analysis of bio-medical data. Class will utilize a combination of lectures complemented with practical computer instruction.

Calendar ........................................................................................................................................2
Location and time ..........................................................................................................................3
Instructors .....................................................................................................................................3
    Joel Parker, Ph.D.
    Hemant Kelkar, Ph.D. ................................................................................................................3
Class description .........................................................................................................................3
Class Syllabus ................................................................................................................................3
Class grade (subject to change) ....................................................................................................3
Auditing and class size ................................................................................................................4
Recommended readings ...............................................................................................................4
    General readings ........................................................................................................................4
Homework assignments ...............................................................................................................4
Class Examinations ....................................................................................................................4
Linux server access for class .......................................................................................................4
<table>
<thead>
<tr>
<th>Date (2018)</th>
<th>Instructor</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-Jan</td>
<td>W</td>
<td>Hemant Kelkar</td>
</tr>
<tr>
<td>15-Jan</td>
<td>M</td>
<td>MLK Day</td>
</tr>
<tr>
<td>16-Jan</td>
<td>Tu</td>
<td>Hemant Kelkar</td>
</tr>
<tr>
<td>17-Jan</td>
<td>W</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>22-Jan</td>
<td>M</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>23-Jan</td>
<td>Tu</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>24-Jan</td>
<td>W</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>29-Jan</td>
<td>M</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>30-Jan</td>
<td>Tu</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>31-Jan</td>
<td>W</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>5-Feb</td>
<td>M</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>6-Feb</td>
<td>Tu</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>7-Feb</td>
<td>W</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>12-Feb</td>
<td>M</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>13-Feb</td>
<td>Tu</td>
<td>Joel Parker</td>
</tr>
<tr>
<td>14-Feb</td>
<td>W</td>
<td>Joel Parker</td>
</tr>
</tbody>
</table>
**Location and time**
Time: 11 am – 12 pm, Mon-Tue-Wed
Lectures and computer lab: Biogen Idec Classroom 307, Health Sciences Library.
Materials: All learning materials will be posted on “Sakai” at Sakai https://sakai.unc.edu.

**Instructors**

*Joel Parker, Ph.D.*
Lineberger Cancer Center and Department of Genetics
5214 Marsico Building
919-966-9614, parkerjs@email.unc.edu

*Hemant Kelkar, Ph.D.*
Center for Bioinformatics and Department of Genetics
4159 Genome Science Building,
919-962-4458, hkelkar@unc.edu

**Teaching Assistant for GNET742**
TBD

**Class description**
This class is designed to teach the fundamentals of UNIX operating system and Python programming using practical “hands-on” computer instruction. This module will concentrate on applications of Python programming to biomedical data/analysis. Target audience is biomedical scientists who are interested in getting familiar with computer clusters for manipulating, parsing, analyzing biological text format data.

**Class Syllabus**

(HK) Introduction to UNIX operating system
Accessing UNIX account using the secure shell (SSH) program. Basic UNIX commands followed by navigation of the UNIX file system. Transferring files from local desktop to the UNIX server using secure file transfer protocol.

(JP) Introduction to Python including history, advantages, and disadvantages of the language

(JP) Definition of a variable and basic uses. Understanding of variable types including scalars, strings

(JP) Reading and writing files and processing of command line arguments

(JP) Understanding basic logic and program control with conditional statements and loops

(JP) Complex data structures focused on processing large genomic data

(JP) Subroutines, regular expressions and modules

(JP) Review and knowledge application

**General Class Policies**
Since we are going to be learning about a new operating system and a programming language learning will be in incremental steps. If you miss a class, then you may not be able to understand material being covered next day. That said, we understand that sometimes there are unexpected emergencies. If you miss a class, please email the instructors with an explanation as soon as practical. Unexplained absences may lead to a penalty reduction in final class grade. Programming logic may be intuitive to some, whereas others may find
it hard. If you think you are in the latter category, then please get in touch with an instructor early on. If you wait till the last week of class, then we will not be able to provide additional help.

**Class grade (subject to change)**

90 % Assignments + 10% Class Participation.

We will have at least 2 (or more) assignments. Expect the final assignment to build upon all you will learn in class. It may take 6-8 h of effort to complete. Some assignments may contain questions that provide additional credit. While those questions are not required to complete the assignment, you will not get the maximum possible points, unless those questions are answered.

There will be NO FINAL EXAM for this class.

**Auditing and class size**

Class is limited to twenty (20) students. Auditing will not be allowed due to the hands-on computer instruction.

**Recommended readings**

**Recommended reading**


**Homework assignments**

Homework assignments will be due by 6 pm on Friday of the following the week they were originally assigned in. You can generally expect to have the grades for the assignments available a week after the due date.

We encourage talking with the instructors or teaching assistant (via email may work best) if you have questions about assignments. Searching the web for a specific question/step is acceptable but you should not copy and paste Python code available on the internet as is.

**Class Examinations**

There will be NO FINAL EXAM for this class.

**Linux server access for class**

Compute cluster access will be provisioned for this class with assistance from ITS-Research Computing and will make use of the “killdevil” cluster. If you are a registered student for this module then an account will be automatically added for you on “killdevil” cluster. Logins will be based on your “onyen”. Access to this server will remain available through the duration of the module and will allow you to work outside the class hours for assignments/projects.

Access to “killdevil” from an off-campus location will require use of UNC-VPN software. It is available by visiting [https://vpn.unc.edu](https://vpn.unc.edu) when you are off-campus. Help with VPN (installation/troubleshooting) available from ITS help desk (919-962-HELP) and [http://help.unc.edu/?s=VPN](http://help.unc.edu/?s=VPN).

You could also choose to use your own laptop/desktop. Python distributions are available for most OS.