CODING FOR BIOLOGISTS (1-DAY WORKSHOP)

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY (KNUST), GHANA

Preamble/Introduction

Computers are increasingly becoming invasive in our daily lives, schools and workplaces. In this post genomic era, characterized by the availability of high-throughput technologies coupled with the advent of powerful computational tools, biomedical science is now seeing the explosion of mass biological data than ever. However, mining of these huge data sets can in fact lead to significant rediscoveries and deeper biological insights such as the molecular mechanisms of certain diseases which remain refractory to current treatment therapies.

More importantly, biological big data is predicted to out-strip astronomical data by several factors in the upcoming years[1] despite recent developments in astronomy, globally such as the establishment of the Square Kilometer Array (SKA) project between Australia, South Africa and other partner countries. For instance, the MeerKAT dishes, the world's largest and sophisticated radio telescopes are expected to generate data equivalent to more than 5 million DVDs per day or about 64 DVDs per second over the life span of the SKA project[2].

On the other hand, astronomers are not the only culprits in this "data-flood" saga and there is now huge demand for biologists to be able to easily manipulate, interpret and bring meaning to current science. Genomics as well as other omics disciplines, medical imaging for treatment and diagnosis, immunological research with the focus of developing high-quality mechanistic and dynamic models of signaling pathways and components, immune cells and their mechanisms of actions, structural derivation and even the evolution of infectious pathogens such as viruses are all key biological avenues that demands high computational power. There is therefore the need for every biologists (luckily to be born in this century and beyond) to expand his/her statistical and computational toolbox in order to push the boundaries of science.

Aim

This workshop was therefore organized to build the capacity of non-computer scientists and non-engineering students and professionals particularly biologist in computer programing and use of Linux environment and terminals.

Objectives

At the end of the workshop, participants were expected to

1. Have a fair idea of the Linux operating system, shells, terminals, kernels etc.
2. Know how to use the command prompt/terminal to change paths/directories and be able to work around with files and folders in the terminal (such as listing, reading, writing and terminating)
3. Know how to access remote/secure connections using the terminal
4. Demonstrate an understanding of the basic syntaxes in python, working with variables, accept input from users, print-out and perform mathematical calculations
5. Be conversant with using the Jupyter notebook, the multipurpose notebook for combining text, codes, results, images and even sound clips.
6. Complete a programming projects such as conversion tables for distance, temperature, time etc.

Workshop Proceedings

The program was held at Oldham Laboratories at the Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Ghana on Wednesday, 26th April, 2017. In all, close to thirty participants including postgraduates, undergraduates and teaching assistants attended the program. The facilitators were Samuel Terkper Ahuno and Dr. Alexander Kwarteng, all from the Biochemistry Dept., KNUST but started by Dr. Alexander Kwarteng giving the opening remarks and addressing the participants of the need to be involved in the data science revolution with focus on biomedical research.

He also stressed the need for students to acquire as much specialized skills whenever there is the opportunity and to be involved in volunteering and community engagement activities such as science communication or outreaches.
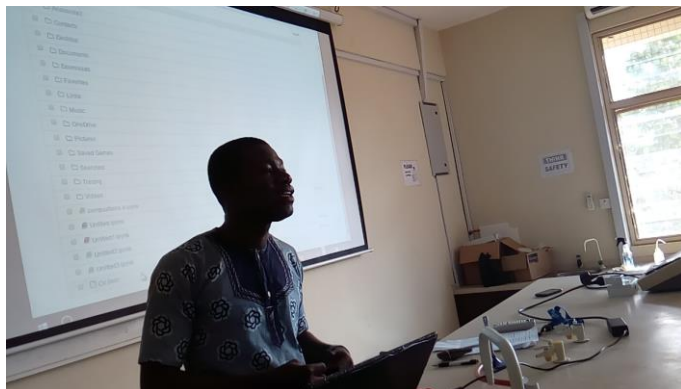
Thereafter, the practical session kicked started beginning with an introduction to the Linux operating system, shells and terminals. However, given that most of the participants were "die-hard fans" of the Microsoft Windows operating system, efforts were also made to provide students with an alternative command prompt terminal "Anaconda command prompt" with comparable capabilities and security to that of Linux.
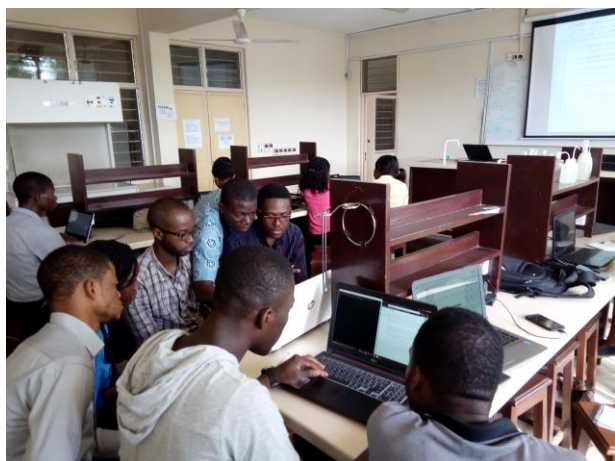
Moreover, navigating through files, folders, changing paths/directories, creating, opening and saving files in command line were as well captured in the demonstrations.

Afterwards, students were taken through some of the syntax and logic behind computer programming. Various keywords such as variables, strings, conditionals, floats, integers

were explained with examples. In addition, a fair idea of how to take input (as characters) from users, processes it and display the output was given. In order to just challenge them, participants were tasked to create codes for scientific computation such as conversion tables for (mm-cm, degrees to Fahrenheit, hrs-seconds etc.).

Challenges and perspective

Although not uncommon, challenges with logistics (computers/laptops) for the training session were met. However, we anticipated this hurdle long ago, and therefore requested that participants come along with their personal computers. But even the available computers were still not enough and we had to pair some students to ensure participants follow through the practical session.

Moreover, similar to the challenges in learning a new language, these participants had an initial tough time grabbing the syntax and the logic behind the python language. Participants were therefore encouraged to practice what they have been introduced to and to look through the now volumes of literature and online courses with focus on the python language. In light of this, a WhatsApp platform was created for the participants to share ideas and challenges as they practice.

Finally, the lack of enough trainers was as well a huge impediment to the smooth running of the programming. Therefore, we believe that using the "Training the Trainers (TTT)" approach where members with passion for this course would be trained to facilitate subsequent workshops should be adopted in subsequent trainings.

Conclusion

Taken together, this workshop was successful despite the short notice and challenges. It is the vision of BSI Ghana that her members learn how to code in at least one computer programming language. Finally, we look forward to future workshops (higher/advanced) in other universities across the country, high schools, communities and even around the globe.

References

[1]Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. (2015) Big Data: Astronomical or Genomical? *PLoS Biol* 13(7): e1002195. https://doi.org/10.1371/journal.pbio.1002195

[2]Thomas Bennett (SKA SA science Data Processing Team) *2016 Newsletter for South Africa's Square Kilometer Array Project*. Pg. 12

Compiled by

Samuel Terkper Ahuno

Revised and Approved by

Alexander Kwarteng (PhD)

   (BSI-Ghana)

Thursday, May 04, 2017