OHSU graduate students desire data analysis and data manipulation skills - a needs assessment for BioData Club

Eric Leung1 Benjamin Cordier1 Eric Earl2 Philip Robinson2 Danielle Robinson3 Asia Mitchell4 Geoffrey Schau5 Kylie McPherson6
Jessica Minnier7 Lilly Winfree8 Letisha Wyatt8 Robin Champieux8 Ted Laderas1
1Department of Medical Informatics & Clinical Epidemiology 2Biomedical Engineering 3Developmental Cognition and Neuroimaging Labs 4Code for Science and Society 5Molecular & Medical Genetics 6Ontology Development Group

Introduction

Along with the experimental wet lab skills, today’s biomedical research requires complementary scientific computing skills. To gain these computing skills, however, students learn these skills in an ad hoc fashion. Furthermore, scientific computing is not easy and individuals will come across obstacles.

BioData Club aims to fill this gap to be a resource and community for students, postdocs, and faculty who want to improve their scientific computing skills to facilitate their own research and gain transferable skills. Here we present preliminary results of a needs assessment survey of scientific computing skills to facilitate their own research and gain transferable skills. We had majority male respondents (N=13), along with six females and one non-binary individual. A majority (N=20) of responders were between the ages of 21 and 39. We had ten graduate students and six research staff, with the rest being faculty, postdocs, and undergraduates.

Objectives

- Assess graduate student needs across graduate programs at OHSU
- Assess gap between utilized scientific computing skills and interested skills
- Assess spectrum of learning styles and barriers to learning computing skills

Methods

We developed an anonymous, 12-question GoogleForms survey and asked the Graduate Student Association (GSO) to help send the survey to all graduate students on November 1st, 2017. There were five question types: information about the student (e.g. which department), skill interests and needs, learning styles, barriers to learning, and skill level.

An earlier, anonymous, 11-question SurveyMonkey survey (in summer 2017) was sent to individuals who had previously attended BioData Club events. Five questions were demographic information (e.g. gender, employment status). The rest were on attendance and interest in previous events and topics.

Results

All survey responses were submitted by November 3rd, 2017. We received N=27 responses which ranged across all departments of OHSU except Molecular Microbiology and Immunology (MMI), Molecular and Medical Genetics (MMG), and Biochemistry and Molecular Biology (Figure 2).

The top three skills that were of interest and in use are manipulating and cleaning data, visualizing data, and machine learning and statistics (Figure 3).

We received N=23 responses from our summer 2017 survey. In addition to shared interest in data analysis skills in our recent survey, programming tools (e.g. specialized libraries) (N=17) and programming practices (e.g. documentation, software design) (N=17) were of interest. Out of our previous events, R (N=16) and Git (N=10) were attended most by responders.

We find that graduate students across most departments at OHSU desire and use various scientific computing skills. There is a strong demand for data manipulation and analysis skills. However, there is a lack of interest in reproducibility, which contrasts the present conversations on reproducibility [I].

Conclusions and Impact

We hope that our preliminary needs assessment of the graduate student population will inform future course and workshop development.

We developed an anonymous, 12-question GoogleForms survey and asked the Graduate Student Association (GSO) to help send the survey to all graduate students on November 1st, 2017. There were five question types: information about the student (e.g. which department), skill interests and needs, learning styles, barriers to learning, and skill level.

An earlier, anonymous, 11-question SurveyMonkey survey (in summer 2017) was sent to individuals who had previously attended BioData Club events. Five questions were demographic information (e.g. gender, employment status). The rest were on attendance and interest in previous events and topics.

Methods

We developed an anonymous, 12-question GoogleForms survey and asked the Graduate Student Association (GSO) to help send the survey to all graduate students on November 1st, 2017. There were five question types: information about the student (e.g. which department), skill interests and needs, learning styles, barriers to learning, and skill level.

An earlier, anonymous, 11-question SurveyMonkey survey (in summer 2017) was sent to individuals who had previously attended BioData Club events. Five questions were demographic information (e.g. gender, employment status). The rest were on attendance and interest in previous events and topics.

Results

All survey responses were submitted by November 3rd, 2017. We received N=27 responses which ranged across all departments of OHSU except Molecular Microbiology and Immunology (MMI), Molecular and Medical Genetics (MMG), and Biochemistry and Molecular Biology (Figure 2).

The top three skills that were of interest and in use are manipulating and cleaning data, visualizing data, and machine learning and statistics (Figure 3).

We received N=23 responses from our summer 2017 survey. In addition to shared interest in data analysis skills in our recent survey, programming tools (e.g. specialized libraries) (N=17) and programming practices (e.g. documentation, software design) (N=17) were of interest. Out of our previous events, R (N=16) and Git (N=10) were attended most by responders.

We had majority male respondents (N=13), along with six females and one non-binary individual. A majority (N=20) of responders were between the ages of 21 and 39. We had ten graduate students and six research staff, with the rest being faculty, postdocs, and undergraduates.

Conclusions and Impact

We find that graduate students across most departments at OHSU desire and use various scientific computing skills. There is a strong demand for data manipulation and analysis skills. However, there is a lack of interest in reproducibility, which contrasts the present conversations on reproducibility [I].

We hope that our preliminary needs assessment of the graduate student population will inform future course and workshop development.

References


Acknowledgements

BioData Club is supported by the OHSU Library and Department of Medical Informatics & Clinical Epidemiology (DMICE). We thank the GSO and graduate students who participated and voices their opinions.