Faculty Member Contact Information

Name	Dr. Carolyn Butts-Wilmsmeyer
Contact Info	
SIUE Email	cbuttsw@siue.edu
Campus Box	1651
Department	Biological Sciences

1 Funded, 3-5 Unfunded URCA Assistants

	This position is **ONLY** open to students who have declared a major in this discipline.	
	This project deals with social justice issues.	
X	This project deals with sustainability (green) issues.	
X	This project deals with human health and wellness issues.	
X	This project deals with community outreach.	
X	This mentor's project is interdisciplinary in nature.	

Are you willing to work with students from outside of your discipline? If yes, which other disciplines?

- Yes, BIOL, CHEM, ENSC, CS, CMIS, MATH & STAT

How many hours per week will your student(s) be required to work in this position? (Minimum is 6 hours per week; typical is 9)

- 9 hours

Will it be possible for your student(s) to earn course credit?

- Yes-- BIOL 493, BIOL 492, BIOL 492M, ENSC 498, ENSC 499, CS 495 (0-3 credit hours)

Location of research/creative activities:

- SIUE Campus (Science West or Biotechnology Laboratory Incubator)

Brief description of the nature of the research/creative activity?

Students participating in this research project will gain experience in data science as it relates to biologically, environmentally, and/or economically important phenomena. Although there is considerable flexibility in the nature of the work students may take on, meaning that projects can be designed with student interests in mind, all students can expect to emerge with training in how to mine large datasets so as to make informed decisions about their research and the world around them. Work may be in the area of bioinformatics, sustainable crop production, development of high-throughput phenotyping procedures (using either digital means or efficient phytochemistry protocols), epidemiology, statistical programming, or community and economic development. While seemingly quite variable, the shared overall objective of each of these project areas is to develop new ways of handling large datasets and use this information in order to either (a) utilize our natural resources in a sustainable manner or (b) promote the health and vitality of our communities. Although not exhaustive, below is a list of the types of projects which are available for students to work on:

1. Projects with genetic and environmental components that cross over into the human health fields. The biostatistical tools that we use are transferrable between disciplines. With numerous datasets being posted electronically, meta-analysis of multiple datasets or even reanalysis of publicly available datasets to address complex questions are now possible. However, this means that novel biostatistical techniques and analysis pipelines that are tailored for complex datasets are also needed. A student interested in this type of research would gain experience in data querying and cleaning, the integration of multiple datasets, the use of various biostatistical tools, and the different techniques used to visualize complex data problems. My interest is in learning how we can best build an applied analysis pipeline to analyze publicly available data to answer complex questions. The specific question can be flexible. Examples of projects that a student could undertake in this area of research include:

a. The development of machine learning models to predict long-term patient outcomes in aging-related diseases OR other public health concerns, and to identify key time points for intervention.

b. The use of data science techniques to understand how various environmental, genetic, demographic, or behavioral factors may influence disease incidence and severity.

2. Developing new phytochemical assessment tools using analytical chemistry techniques in order to monitor crop loss. Billions of dollars are lost each year due to postharvest loss and

mycotoxins each year in the US. In developing countries, mycotoxins may go undetected, posing a significant health hazard to people who consume these toxins in their diets; even when detected, these toxins pose a significant threat to a sustainable food supply. However, the protocols that are currently used for the detection of these compounds are time consuming and are unrealistic for widescale use. The student working on this project would modify standard laboratory protocols to work in a high-throughput setting, thereby enabling the rapid detection of these toxins. This information will be used to help create new crop varieties that are naturally resistant to the development of mycotoxins and other negative outcomes associated with plant diseases.

3. Using spectral imagery to study morphological and phytochemical changes in plants due to abiotic stress. The goal of this project is to combine remote sensing and artificial intelligence (AI) techniques to capture plant responses to common abiotic stress factors, especially those expected to become more problematic due to global climate change. Students working on this project can expect to gain experience working with image processing, spectral imagery, vegetative indices and other metrics of plant health, greenhouse experimentation, and statistical programming (e.g. Python and/or R).

Brief description of student responsibilities?

Depending on the exact research project, the research may be solely computer based or may be a combination of computer-based and laboratory or community based. The student will be responsible for formatting and cleaning the research data, summarizing that data, running statistical models, and compiling research results. Students in community projects may be responsible for tasks such as compilation of surveys and meeting with community stakeholders (remotely). Students engaged in biochemistry research will be working in a laboratory setting for the majority of their project. Although students should feel comfortable with a computer, no prior coding experience is required, as this is something that can be learned during the research experience. Basic statistical experience is required (e.g. calculation of means and summary statistics), but advanced predictive modeling can also be learned during the research experience.

URCA Assistant positions are designed to provide students with *research or creative activities* experience. As such, there should be measurable, appropriate outcome goals. What exactly should your student(s) have learned by the end of this experience?

1. Which software programs are commonly used for statistical, bioinformatics, and/or general data analysis.
2. How to write basic and intermediate commands in those software programs.
3. For epidemiology and statistical genetics projects, that the genotype-by-environment interaction influences complex traits (e.g. disease response) more than it does simple traits (e.g. blood type), and this can have an impact on which bioinformatics tools are best suited

for a given circumstance.

4. How to recognize outliers and other troublesome data points, and how to perform data collection in a manner that reduces the error (i.e. background noise).

5. In the case of biostatistics students, how to perform common statistical analyses. In the case of biochemical students, how to perform common laboratory protocols while operating under standard safety procedures.

6. How to interpret data and results.

7. How to summarize those analyses for presentation to a scientific audience

Requirements of Students

If the position(s) require students to be available at certain times each week (as opposed to them being able to set their own hours) please indicate all required days and times:

- Schedules can be flexible

If the location of the research/creative activities involves off campus work, must students provide their own transportation?

- If working on the phytochemical project, the student must provide their own transportation to the Biotechnology Laboratory Incubator (BLI). There is a bus stop directly across from the BLI, if needed.

Must students have taken any prerequisite classes? Please list classes and preferred grades:

- No prerequisite classes are required, although it might be helpful if students have taken introductory courses in mathematics and statistics.

Other requirements or notes to applicants:

- N/A