**Applied Statistics for Bioinformatics**

**The University of Toledo**

**Bioinformatics, Proteomics and Genomics, COMLS**

**BIPG 6500 001 (27679)**

**Instructor**: Sadik A. Khuder. Ph.D.

**Email**: sadik.khuder@utoledo.edu

**Office Hours**: By arrangement

**Office Location**: Room 12 RHC

**Office Phone**: (419) 383-4089

**Term**: Spring 2022

**Class Location**: 127 HEB

**Class Day/Time**: Wednesday 9:00-12:00

**Lab Location**: 127 HEB

**Lab Day/Time**:  **Credit Hours**: 3

CATALOG/COURSE DESCRIPTION\*

This course will provide students with practical statistical and data analysis skills to perform rigorous analysis of high-throughput biological data. The course assumes familiarity with the statistical methods and with R programming. The course covers the statistical concepts necessary to design experiments and analyze high-dimensional data generated by high-throughput technologies. Also included are stochastic modeling and statistical methods applied to problems such as mapping disease-associated genes, SNP and mutation analysis, transcriptomics, miRNA, DNA methylation and epigenetics, proteomics, metabolomics, and metagenomics.

## COURSE OVERVIEW

Application of statistical methods in bioinformatics. This course is the second in a two semester sequence examining topics in biostatistics for graduate students in bioinformatics or related fields.

STUDENT LEARNING OUTCOMES
After completion of the course, students should be able to:

L1. Identify statistical problems that accompany big omics datasets within the life sciences.

L2. Analyze large-scale data generated from emerging genomic and proteomic techniques.

L3. Apply statistical techniques to solve bioinformatics problems

L4. Apply advanced statistical techniques for biomarker and disease gene identification

L5. Combine data from different platforms

L6. Analyze large-scale genetic data.

L7. Construct [integrative analysis of omics and health record data.](https://www.ncbi.nlm.nih.gov/pubmed/29536440)

TEACHING METHODOLOGY

A variety of teaching methods will be used, including in class lectures, exercises, quizzes, project, and online instructional lectures and videos.

PREREQUISITES AND COREQUISITES
*None*

## TEXTS AND ANCILLARY MATERIALS

There is no required text. All the required materials will be available on Blackboard. Readings will consist of original literature, review articles, and R based books.

TECHNOLOGY REQUIREMENTS
*None*

ACADEMIC POLICIES\*
[Graduate Policies](http://www.utoledo.edu/policies/academic/graduate/): <http://www.utoledo.edu/policies/academic/graduate/>

**UNIVERSITY POLICIES**

### Policy Statement on Non‐Discrimination on the Basis of Disability (ADA)\*

The University is an equal opportunity educational institution. Please read [The University’s Policy Statement on Nondiscrimination on the Basis of Disability Americans with Disability Act Compliance.](http://www.utoledo.edu/policies/administration/diversity/pdfs/3364_50_03_Nondiscrimination_o.pdf)
Students can find this policy along with other university policies listed by audience on the [University Policy webpage](http://www.utoledo.edu/policies/audience.html/#students) (http://www.utoledo.edu/policies/audience.html/#students).

### Academic Accommodations\*

The University of Toledo embraces the inclusion of students with disabilities. We are committed to ensuring equal opportunity and seamless access for full participation in all courses. For students who have an accommodations memo from Student Disability Services, I invite you to correspond with me as soon as possible so that we can communicate confidentially about implementing accommodations in this course. For students who have not established affiliation with Student Disability Services and are experiencing disability access barriers or are interested in a referral to healthcare resources for a potential disability or would like information regarding eligibility for academic accommodations, please contact the [Student Disability Services Office](http://www.utoledo.edu/offices/student-disability-services/index.html) (http://www.utoledo.edu/offices/student-disability-services/) by calling 419.530.4981 or sending an email to StudentDisability@utoledo.edu.

## ACADEMIC AND SUPPORT SERVICES

Please follow this link to view a comprehensive list of [Student Academic and Support Services](http://www.utoledo.edu/studentaffairs/departments.html) (http://www.utoledo.edu/studentaffairs/departments.html) available to you as a student (please refer to the face-to-face syllabus guidelines for more guidance/details).

## SAFETY AND HEALTH SERVICES FOR UT STUDENTS\*

Please use the following link to view a comprehensive list [Campus Health and Safety Services](http://www.utoledo.edu/offices/provost/utc/docs/CampusHealthSafetyContacts.pdf) available to you as a student (please refer to the face-to-face syllabus guidelines for more guidance/details).

COURSE EXPECTATIONS

Students are expected to attend lectures and participate in class discussions, successfully complete all assignments and projects, and pass all quizzes and exams.

OVERVIEW OF COURSE GRADE ASSIGNMENT\*

**Assignments**  30%

There are 11 assignments in this course. These assignments are intended to improve skills in bioinformatics and statistical analysis of genomic and proteomic data. The lowest assignment score will be dropped from the calculation of your final grade.

Students registering for BIPG 720 are required to do an additional assignment.

**Quizzes**  15%

There are 5 quizzes in this course. Each quiz will cover the material presented in that particular week. Few questions from previous quizzes may also be included. Further instructions will be supplied for each particular quiz.

**Project**  20%

A project will explore a topic of the course in greater depth. A written report, along with a power point presentation, is required. A project that addresses a question through data analysis with a written report and summary of conclusions would be sufficient. More detailed descriptions of the project along with suggested topics will be posted to the course webpage at the appropriate times during the term. The due date for project submission is last week of the course.

**Final examination** 30%

The final exam will be a take-home exam. The exam will cover all the materials presented in the course.

Final Grading

≥ 90 A, 87-89 A-, 84-86 B+ , 80-83 B, 77-79 B- , 74-76 C+, 70-73 C, < 70 D, < 50 F.

**COURSE SCHEDULE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **Date** | **Topic** | **LEARNING OUTCOME(S)** | **A/Q\*** |
| 1 | 1/19 | Analysis of gene expression data | L1 , L2 | A1 |
| 2 |  1/26 | RNA-seq I | L1 , L2  |  A2 |
| 3 | 2/2 | RNA-seq II | L1 , L2 | A3, Q1 |
| 4 | 2/9 | Non-Coding RNA | L2 , L3 | A4 |
| 5 | 2/16 | DNA methylation & epigenetics | L2 , L3 |  A5 |
| 6 | 2/23 | Proteomics I | L1 , L2  | Q2 |
| 7 | 3/2 | Proteomics II | L1 , L2  | A6 |
| 8 | 3/16 | Metabolomics | L3 , L4  |  A7 |
| 9 | 3/23 | [Chromatin](https://www.googleadservices.com/pagead/aclk?sa=L&ai=DChcSEwj8ntLNmqzmAhWQhsAKHXYXDvwYABAFGgJpbQ&ohost=www.google.com&cid=CAASE-RobzvlViDzfo55sRYeXtxcH4A&sig=AOD64_1bMeDsuYQg89zaBeDNbFAuPStgpA&q=&ved=2ahUKEwjzpcjNmqzmAhUE-6wKHdr8DJMQ0Qx6BAgQEAE&adurl=) Immunoprecipitation | L3 , L4 | A8,Q3 |
| 10 | 3/30 | Metagenomics  | L3 , L4 | A9 |
| 11 | 4/6 | Combining omics data  | L5  | A10 |
| 12 | 4/13 | Analysis of genetic data | L6 | Q4 |
| 13 | 4/20 | GWAS studies I | L6 | A11 |
| 14 |  4/27 | GWAS studies II  | L6 |  Q5 |
| 15 |  5/4 | **Biomedical informatics** | L7 | Final exam |

*\* Assignments/Quizzes*