

DATA SCIENCE (DATASCI)

DATASCI 217 Introduction to Python and Data Science Tools (1-2 Units) Fall

Instructor(s): John Kornak

Prerequisite(s): BIostat 213 or equivalent (knowledge of probability/statistics and familiarity with programming concepts, e.g., from using R)

Restrictions: This course is part of the Health Data Science Masters and Certificate Program and may have space limitations.

Activities: Lecture, Workshop

This course provides an introduction to essential tools and skills for data science, focusing on Python programming and industry-relevant tools. Students will learn command line basics, version control with Git, documentation with Markdown, remote execution, and high-performance computing (HPC). Integrated throughout the course, the Python component covers syntax, flow control, data management, visualization, libraries for data science, and algorithms and data structures common in interviews.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: Letter Grade, P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No

DATASCI 220 Data Science Program Seminar I (1 Units) Fall, Winter, Spring

Instructor(s): John Kornak

Prerequisite(s): BIostat 202 and BIostat 213

Restrictions: This course is restricted to students enrolled in the Certificate in Health Data Science and the Master's degree in Health Data Science (first year students).

Activities: Seminar, Independent Study

This seminar series covers topics in data science algorithms, ethics, biases, and applications. Students will be exposed to current topics on Data Science and Machine Learning/Biostatistics and Health Data applications, discuss issues in data science, present their work, and learn how to critically evaluate research literature. External speakers will be invited to give presentations on potential careers in health data science across the biotech industry, government and academia.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? Yes

DATASCI 221 Data Science Program Seminar II (1 Units) Fall, Winter, Spring

Instructor(s): John Kornak

Prerequisite(s): DATASCI 220

Restrictions: This course is restricted to students enrolled in year 2 of the Master's in Health Data Science program.

Activities: Seminar, Independent Study

This course covers advanced topics of data science methods, ethics and biases. The focus in this second year of the seminar program will be on students presenting their research work progress from their Capstone projects. Additionally, students will also learn how to critically evaluate research literature.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? Yes

DATASCI 222 Data Science Capstone Project (8 Units) Fall, Winter, Spring

Instructor(s): John Kornak

Prerequisite(s): BIostat 202, BIostat 213, BIostat 214, BIostat 216, DATASCI 220, DATASCI 225

Restrictions: This course is restricted to 2nd year students in the Master's in Health Data Science program.

Activities: Project

Capstone project requirement for students in the Masters in Health Data Science program. Students will write a first author paper researching a problem in health data science and analyzing data using appropriate data science methodology; present their work at a scientific conference; generate a portfolio of code, analyses and data products; and write a detailed report on the background methodology and technical issues that were considered as well as implemented for the submitted publication.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? Yes

DATASCI 223 Applied Data Science with Python (2 Units) Spring

Instructor(s): John Kornak

Prerequisite(s): Familiarity with programming concepts, including loops, variables, and functions. Ideally, hands-on experience writing and running scripts such as in: Python, R, Bash, or other programming languages.

Restrictions: This course is part of the Health Data Science Masters and Certificate Program and may have space limitations. Auditing is not permitted.

Activities: Lecture, Project, Workshop

Survey of Data Science methods in Python, starting with common data science tools and processes and spending one week per topics learning to build common ML/AI solutions.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: Letter Grade, P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No

DATASCI 224 Understanding Machine Learning: From Theory to Applications (3 Units) Spring

Instructor(s): Jean Feng

Prerequisite(s): BIostat 216

Restrictions: This course is part of the Health Data Science Masters and Certificate Program and may have space limitations. Auditing is not permitted.

Activities: Lecture, Project

This course teaches the mathematical foundations of machine learning (ML). Each week, the course surveys a different algorithm to examine its underlying machinery, covering topics such as linear algebra, calculus, and optimization. ML algorithms range from linear models to gradient boosting and deep learning. The course also discusses newer concepts such as model fairness and ML for causal inference. Upon course completion, students should be able to learn new ML algorithms independently.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: Letter Grade, P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No

DATASCI 225 Advanced Machine Learning for the Biomedical Sciences II (3 Units) Spring

Instructor(s): Gilmer Valdes

Prerequisite(s): BIostat 213, BIostat 216 and BIostat 208. Exceptions to these prerequisites may be made with the consent of the Course Director, space permitting.

Restrictions: This course is part of the Health Data Science Masters and Certificate Program and may have space limitations. Auditing is not permitted.

Activities: Lecture, Project

This course covers the underlying formulation of machine learning algorithms. Its focus is on providing deep understanding of machine learning methodology. This is an advanced course in machine learning and its objective is to provide students with a strong foundation so that they can properly manipulate and customize black box machine learning library packages. Students will implement popular machine learning algorithms and customize them to best satisfy specific needs in medicine.

School: Graduate Division

Department: Clinical and Epidemiological Research Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: Letter Grade, P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No

DATASCI 226 Bayesian Methods and Gaussian Processes (2-3 Units) Fall

Instructor(s): John Kornak

Prerequisite(s): Basic knowledge of probability and statistics (BIostat 200 and BIostat 208 equivalent); programming skills in R (BIostat 213 and BIostat 214 equivalent); some familiarity with calculus and linear algebra (especially for the extra Gaussian processes unit).

Restrictions: This course is part of the Health Data Science Masters and Certificate Program and may have space limitations. Auditing is not permitted.

Activities: Lecture, Project

This course provides an introduction to Bayesian statistics, Markov Chain Monte Carlo (MCMC) sampling, and Gaussian Processes. The first two units cover the fundamentals of Bayesian methods and MCMC, and the final optional unit explores Gaussian processes. Students will gain practical skills in applying these techniques to real-world problems using R, STAN, and JAGS.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? No

Does enrollment in this course require instructor approval? No

Course Grading Convention: Letter Grade, P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No

DATASCI 300 Data Science Educational Practice (2 Units) Fall, Winter, Spring, Summer

Instructor(s): Staff

Prerequisite(s): Students must have previously taken the course they EA for.

Restrictions: This course is restricted to 2nd year students in the Master's in Health Data Science program.

Activities: Lab science, Discussion

Masters in Health Data Science students are expected to act as an educational apprentice (EA). This experience involves leading a weekly small-group discussion section of 10-15 students, holding office hours and grading homework assignments and projects. This requirement will provide students with valuable teaching experience without having a significant time impact on their Capstone project work. In all cases, students will have taken the course they are asked to EA during their first year.

School: Graduate Division

Department: Health Data Science Program

May the student choose the instructor for this course? Yes

Does enrollment in this course require instructor approval? No

Course Grading Convention: P/NP (Pass/Not Pass) or S/U (Satisfactory/Unsatisfactory)

Graduate Division course: Yes

Is this a web-based online course? No

Is this an Interprofessional Education (IPE) course? No

May students in the Graduate Division (i.e. pursuing Master or PhD) enroll in this course? Yes

Repeat course for credit? No