

BIOF 475: Introduction to Data Science

Learning from data in order to make useful predictions or obtain insights is a cornerstone of modern science. The goal of this course is to introduce students to the basic tools and workflows for doing this, with a focus on biological- and health-related data. In this course, students will learn how to use Python-based tools, particularly Numpy, SciKit-learn, Pandas, and Matplotlib.

Learning Objectives

- Load and clean data
- Choose what type of model (e.g. supervised or unsupervised) to use based on the questions being asked of the data
- Build and validate the chosen model
- Visualize and explain what that model learned from the data

Credits: 2

Class Type: Graduate Course

Prerequisites:

Previous programming experience is not required, but is recommended.

Program: Bioinformatics and Data Science

Availability Summer 2022