

# DATA SCIENCE & ENGINEERING (DATA\_ENG)

---

**DATA\_ENG 200-0 Foundations of Data Science (1 Unit)** This course will cover the fundamentals of data science and the context within which this field operates. Students will learn how to design their data analysis by learning to think critically about what questions are answerable with data and they will learn about common pitfalls in data analytics such as algorithmic bias and best practices for handling the sensitive data of others. It will also introduce students to computational thinking, a methodology for problem-solving the technological challenges they will encounter as data scientists. This course will also introduce the steps of the data science lifecycle and common tools and techniques for data science. We will cover data exploration, the principles of data cleaning and integration, version control, and building reproducible data science pipelines. This course is reserved for students pursuing the McCormick Data Science and Engineering Minor. We encourage students to take this early in their studies for the minor. It is the first part of a two-part sequence with DATA\_ENG 300-0. Prerequisite: COMP\_SCI 150-0.

**DATA\_ENG 300-0 Data Engineering Studio (1 Unit)** Data Engineering Studio will teach students how to build a sustainable data science lifecycle. Students will analyze data in multiple contexts (e.g., SQL, building machine learning models), share their findings with peers, and practice iteratively refining their analysis based on feedback from the instructor, course staff, and peers. Students will also hone their practical skills and they will become acquainted with the common pitfalls in applying data analytics to real-world datasets. Moreover, students will learn how to analyze and visualize data from multiple data models, including graph analytics, time series data, and relational data. This course is reserved for students pursuing the McCormick Data Science and Engineering Minor. We encourage students to take this course at the end of their studies in the minor. It is the second part of a two-part sequence with DATA\_ENG 200-0. DATA\_ENG 300-0 has a “flipped classroom” format. Students are responsible for watching a lecture before each class. Then they will work collaboratively with their teams, the instructor, and course staff to learn how to solve various challenges in the data science pipeline. Prerequisite: DATA\_ENG 200-0 and 1 unit from each of the following core areas; Statistics Foundations, Intermediate Programming/Algorithmic Skills, and Applied Machine Learning.