DATA SCIENCES (DS)

DS 97: SPECIAL TOPICS
1-9 Credits/Maximum of 9
Formal courses given infrequently to explore, in depth, a comparatively narrow subject that may be topical or of special interest.

DS 99: Foreign Studies
1-12 Credits/Maximum of 999
Courses offered in foreign countries by individual or group instruction.

International Cultures (IL)

DS 120: Scripting for Data Sciences
1 Credits
Introductory course in computer-based scripting languages for use in data analyses. DS 120 Scripting for Data Sciences (1) This introductory course aims to teach practical skills in data manipulation and preprocessing scripting, including the fundamentals of an interpreted programming language for use in the data sciences. The goal of the course is to provide an accessible (no pre-requisites) and brief (1 credit) introduction, preparing students for hands-on data analytics assignments in DS 200 Introduction to Data Sciences. This practical course teaches fast manipulation of datasets on the Unix command line, scripting in spreadsheets, and fundamental control structures and data manipulation in a modern interpreted programming language. It is expected that students gain an overview of the available tools and techniques that allows them to acquire basic proficiency in select techniques in the course of applications in most other courses in Data Sciences.

DS 197: Special Topics
1-9 Credits/Maximum of 9
Formal courses given infrequently to explore, in depth, a comparatively narrow subject that may be topical or of special interest.

DS 199: Foreign Studies
1-12 Credits/Maximum of 12
Courses offered in foreign countries by individual or group instruction.

DS 200: Introduction to Data Sciences
4 Credits
The course introduces students to data sciences, an emerging discipline focused on the knowledge and skills needed to harness the power of data to advance science and engineering, address complex national and global challenges, inform public policy, and improve human lives. It demonstrates how the discipline of data science integrates knowledge and skills in computer sciences, statistics, and informatics (with exposure to application domains such as life science, health science, cyber security, astronomy, etc). Through a combination of lectures, hands-on labs, and case studies, students are introduced to the "big picture" of data sciences including elements of understanding data through exploratory data analysis, testing hypotheses against data, building predictive models, all using real-world examples. The course also introduces students to opportunities to specialize in Applied Data Sciences (with an emphasis on data sciences applications in the real world), Computational Data Sciences (with an emphasis on well-engineered data analytics systems), and Statistical Data Sciences (with an emphasis on advanced statistical theory and methods).

DS 220: Data Management for Data Sciences
3 Credits
The course introduces students to the fundamentals of data models: organizing, managing, and using different types of data that arise in real-world applications. The course introduces students to several alternative data models and database solutions, emphasizing their strengths and limitations in the context of real-world applications. Topics covered include the relational databases, key-value stores, column-oriented databases, vector-space databases, graph databases, and distributed file systems together with their applications in solving real-world big data management problems. Upon completion of the course, the students will be able to choose an appropriate data model and database solution for a given application, and use the chosen database to organize, manage, and use data in the context of specific applications.

Enforced Prerequisite at Enrollment: CMPSC 121 or CMPSC 131

DS 294: Research Project
1-12 Credits/Maximum of 12
Supervised student activities on research projects identified on an individual or small-group basis.

DS 296: Independent Studies
1-18 Credits/Maximum of 18
Creative projects, including research and design, that are supervised on an individual basis and that fall outside the scope of formal courses.

DS 297: Special Topics
1-9 Credits/Maximum of 9
Formal courses given infrequently to explore, in depth, a comparatively narrow subject that may be topical or of special interest.

DS 299: Foreign Studies
1-12 Credits/Maximum of 12
Courses offered in foreign countries by individual or group instruction.

DS 300: Privacy and Security for Data Sciences
3 Credits
The course provides students with the knowledge and skills to analyze and implement protection strategies for data privacy and security.

Enforced Prerequisite at Enrollment: DS 220

DS 310: Machine Learning for Data Analytics
3 Credits
The course teaches students the principles of machine learning (and data mining) and their applications in the data sciences. DS 310 Machine
Learning for Data Analytics (3) The course introduces the principles of machine learning (and data mining), representative machine learning algorithms and their applications to real-world problems. Topics to be covered include: principled approaches to clustering, classification, and function approximation from data, feature selection and dimensionality reduction, assessing the performance of alternative models, and relative strengths and weaknesses of alternative approaches. The course will include a laboratory component to provide students with hands-on experience with applications of the algorithms to problems from several domains. Prerequisites for the course include basic proficiency in programming, elementary probability theory and statistics, and discrete mathematics.

**Enforced Prerequisite at Enrollment:** (CMPSC 121 or CMPSC 131) and STAT 318

DS 320: Data Integration

3 Credits

Recommended Preparations: DS 310 Modern data-intensive applications (healthcare, security, public policy, science, commerce, crisis management, education, among others) increasingly call for integration of multiple types of data from disparate sources. This course introduces students to the principles and the practice of data integration, with particular emphasis on relational, knowledge-based, graph-based, and probabilistic methods. Carefully crafted assignments will help enhance the students’ mastery of both the theoretical underpinnings as well as practical aspects of data integration. The students will work in teams to solve representative data integration problems drawn from real-world applications. Upon completion of the course, students should be able design, implement, and evaluate data integration solutions to support data intensive applications.

**Enforced Prerequisite at Enrollment:** DS 220 and STAT 318. Recommended Preparation: DS 310

DS 330: Visual Analytics for Data Sciences

3 Credits

The course introduces visual analytics methods and techniques that are designed to support human analytical reasoning with data. DS 330 Visual Analytics for Data Sciences (3) Visual analytics is the science of combining interactive visual interfaces and information visualization techniques with automatic algorithms to support analytical reasoning through human-computer interaction. People use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data, and to communicate their findings effectively for decision-making. This course will serve as an introduction to the science and technology of visual analytics and will include lectures on both theoretical foundations and application methodologies. The goals of this course are for students to (1) develop a comprehensive understanding of this emerging, multidisciplinary field, and (2) apply that understanding toward a focused research problem in a real-world application or a domain of personal interest.

**Enforced Prerequisite at Enrollment:** DS 220
Networks constitute a useful representational abstraction for data from many real-world applications, such as social networks, biomolecular networks, brain networks, among others. This course aims to cover the conceptual, algorithmic, and applied aspects of network analytics to prepare the students to analyze network data. Specific topics to be covered include different kinds, e.g., undirected, directed, weighted, unweighted, labeled, unlabeled, networks; network properties and network statistics; statistical models of static and dynamic networks; classes of networks, different classes of network structure, e.g., random, small-world, modular hierarchical; methods for community detection, information propagation, node and link classification, network embedding, and their applications in life sciences e.g., predicting protein function, social sciences, e.g., analyzing social ties, e-commerce e.g., recommender systems.

Enforced Prerequisite at Enrollment: Data Science Major or 7th semester standing

Writing Across the Curriculum
DS 442: Artificial Intelligence

3 Credits

This course provides an overview of the foundations, problems, approaches, implementation, and applications of artificial intelligence. Topics covered include problem solving, goal-based and adversarial search, logical, probabilistic, and decision theoretic knowledge representation and inference, decision making, and learning. Through programming assignments that sample these topics, students acquire an understanding of what it means to build rational agents of different sorts as well as applications of AI techniques in language processing, planning, vision.

Enforced Prerequisite at Enrollment: CMPSC 221. Enforced Concurrent at Enrollment: CMPSC 465
Cross-listed with: CMPSC 442

DS 494: Research Project

1-12 Credits/Maximum of 12

Supervised student activities on research projects identified on an individual or small-group basis.

DS 496: Independent Studies

1-18 Credits/Maximum of 18

Creative projects, including research and design, that are supervised on an individual basis and that fall outside the scope of formal courses.

DS 497: Special Topics

1-9 Credits/Maximum of 9

Formal courses given infrequently to explore, in depth, a comparatively narrow subject that may be topical or of special interest.