Welcome to the HDSI MS Graduate Program!

The goal of the master’s program is to teach students knowledge and skills required to be successful at performing data driven tasks, and lay the foundation for future researchers who can expand the boundaries of knowledge in Data Science itself. To meet its goals, the Master of Science in Data Science (MS-DS) program consists of two components: formal courses, as well as a terminating thesis or a course-directed comprehensive examination.

**Course Requirements and General Overview**

There are Introductory/Foundational, Core, and Elective and Research requirements (Group A, B, and C courses) for the master’s program. These course requirements are intended to ensure that students are exposed to (1) fundamental concepts and tools (Foundation), (2) advanced, up-to-date views in topics central to Data Science for all students (the Core requirement), and (3) a deep, current view of their research or application are (the Elective requirement). Courses may not fulfill more than one requirement.

The master of science in data science program is structured as a total of twelve (12) 4-unit courses grouped into foundational (Group A), core (Group B), and specialization (Group C) areas as described in detail here.

Successful completion of the program requires completion of a Thesis (Plan I) or a course-based comprehensive examination (Plan II) that tests integrative knowledge across multiple courses. Out of the 48 units, at least 40 units must be using graduate-level courses. In addition, 2 out of 10 graduate courses can be in areas not directly related to data science but a domain specialization such as economics, biology, medicine, etc upon approval of the student’s faculty advisor.

To fulfill the course requirements:

A maximum of 4 courses can be taken from Group A (Foundation courses).
A minimum of 6 courses need to be taken from Group B (Core courses). Among these core courses, all students have to take three required core courses (DSC240, DSC260 and DSC 241).
A minimum of 3 courses need to be taken from Group C (Electives courses/research credits).
The required number of elective courses and research credits varies depending on whether a student is taking MS degree Plan I or Plan II. See details below.

**Thesis or Comprehensive Exam Requirements**

The MS/DS degree can be pursued under either the Thesis option (Plan I) or the Comprehensive Examination option (Plan II). The comprehensive examination option follows a course-based comprehensive examination plan under the supervision of a comprehensive examination committee. For full-time students, all the requirements can be completed within one to two years. Students must register for a minimum of three quarters for residency requirements. To maintain good academic standing, students must be making timely and satisfactory progress toward
completion of degree requirements and must maintain a minimum overall GPA of 3.0 at UC San Diego.

Approved Elective Courses and Research Credits

The number of Elective and Research units required varies by degree (see below). Electives are chosen from graduate courses in DSC, CSE, Cognitive Science, ECE, and Mathematics or from other departments as approved. Please refer to the HDSI website for a list of approved Electives. Courses must be completed for a letter grade, except Research units that are taken on a Satisfactory/Unsatisfactory basis. Seminar and teaching units may not count toward the Electives and Research requirement, although both are encouraged.

Course Requirements – Detailed Overview

There are Introductory, Core, and Elective and Research requirements (Group A, B, and C courses below) for the masters program. These course requirements are intended to ensure that students are exposed to (1) fundamental concepts and tools (Foundation), (2) advanced, up-to-date views in topics central to Data Science for all students (the Core requirement), and (3) a deep, current view of their research or application area (the Elective requirement). Courses may not fulfill more than one requirement.

The master of science in data science program is structured as a total of twelve (12) 4-unit courses grouped into foundational, core and specialization areas as described below. Successful completion of the program requires completion of a Thesis or a course-based comprehensive examination that tests integrative knowledge across multiple courses. Out of the 48 units, at least 40 units must be using graduate-level courses. In addition, 2 out of 10 graduate courses can be in areas not directly related to data science but a domain specialization such as economics, biology, medicine etc upon approval of the student’s faculty advisor.

Group A: Introductory courses: maximum of four course credit

These courses seek to provide five critical foundational knowledge and skills that each student graduating from the master’s program is expected to receive at a graduate level: programming skills, data organization and methods skills, numerical linear algebra, multivariate calculus, probability and statistics.

The program is designed so that students lacking in any (and all) of these foundational knowledge and skills can take credit for a maximum of four courses from the following five courses: DSC 200, DSC 202, DSC 210, DSC 211 and DSC 212.

1. DSC 200: Data Science Programming
2. DSC 202: Data Management for Data Science
3. DSC 210: Numerical Linear Algebra
4. DSC 211: Introduction to Optimization
5. DSC 212: Probability and Statistics for Data Science
**Group B: Core Courses: 3 required courses, minimum of six courses**

These courses build upon foundational courses. All students must take three required core courses: DSC 240, DSC 241 (*), and DSC 260. In addition, students can select at least three out of the following core courses: DSC 203, DSC 204A (*), DSC 204B, DSC 206, DSC 215, DSC 242, DSC 243, DSC 244, DSC 245, DSC 250, DSC 261.

(*) Depending on academic preparation, a M.S. student can take an advanced course on Applied Statistics, such as MATH 282B instead of DSC 241. Similarly, instead of DSC 204A, a student can take a course on Algorithms, such as CSE 202: Design and Analysis of Algorithms.

The list of courses with names are below:

1. DSC 240: Machine Learning
2. DSC 260: Data Ethics and Fairness
3. DSC 203: Data Visualization and Scalable Visual Analytics
4. (*)DSC 204A: Scalable Data Systems (or CSE 202)
5. DSC 204B: Big Data Analytics and Applications
6. DSC 206: Algorithms for Data Science
7. DSC 215: Statistical critical thinking
8. (*)DSC 241: Statistical Models (or MATH 282B)
9. DSC 242: High-dimensional Probability and Statistics
10. DSC 243: Advanced Optimization
11. DSC 244: Large-Scale Statistical Analysis
12. DSC 245: Introduction to Causal Inference
13. DSC 250: Advanced Data Mining
14. DSC 261: Responsible Data Science

**Group C: Elective and Specialization Courses: remaining course credit requirements**

The MS students can take advantage of electives to complete their course of study. These courses can be advanced courses in core Data Science subjects listed under Group B as research topics (DSC 291) courses, or they can be graduate (or upper-division undergraduate) courses in other departments subject to approval by the student’s HDSI faculty advisor.

As a matter of guidance, students can choose from the following elective or specialization tracks to complete course requirements.

**General Elective Courses:**

DSC 205, DSC 231, DSC 251, DSC 252, DSC 253, DSC 254, DSC 213, DSC 214

CSE 234, MATH 181 A-B-C, MATH 284, MATH 285, MATH 287A-B, COGS 243.
**Specialization Areas: minimum of 3 courses required**

Upon prior approval from a graduate advisor, students can sign up for an available specialization area for an “Master of Science in Data Science with specialization in specialization-area” degree. A specialization requires a minimum of three courses in a specialization area.

- **Specialization: Bioengineering**
  - BENG 218, BENG 203, BENG 211, BENG 213, BENG 221, BENG 230A-B, BENG 276, COGS 278, PHYS 278, FMPH 223, FMPH 226

- **Specialization: Business (Marketing)**
  - MGT 475, MGT 477, MGT 489, MGTA 455, MGTA 479

- **Specialization: Business (Supply Chain and Technology)**
  - MGT 450, MGT 451, MGTA 456, MGTA 463

- **Specialization: Business (Finance)**
  - MGT 407, MGTF 402, MGTF 404, MGTF 405, MGTF 406, MGTF 415

- **Specialization: Machine Vision and Interaction Design**
  - COGS 202, COGS 220, COGS 225, COGS 283

- **Specialization: Computational Neuroscience**
  - COGS 260, BGGN 246, BGGN 260, COGS 260 (or NEU 282), COGS 280

- **Specialization: Networks**
  - MATH 261A, MATH 277A, MATH 289A, MATH 289B, DSC 205, BNFO 286, POLI 287, SIOB 276, ECE 227, MAE 247

Availability of all specializations is not guaranteed. Additional specialization areas may be added by student petition.

**MS Plan 1: Thesis Option**

The student must sign up for a minimum of 8 and a maximum of 12 units of DSC 299 (Thesis Research) which can also be used to meet Group C requirements. The student will perform thesis research under the guidance of a thesis advisor and a thesis committee consisting of at least three members. The Master thesis produced by the student must be approved by the committee. It is required that at least two members of the committee are members of the HDSI faculty council; one of the three committee members can be an HDSI industry fellow with an adjunct appointment or a faculty member drawn from another department or division. Chair of the committee shall be appointed by the Graduate program committee. Alternatively, the industry
fellow may be requested to serve as the fourth member of the committee. The committee must be approved by the Graduate Division by the end of the fourth quarter in the MS program.

**MS Plan 2: Comprehensive Exam Option**

Comprehensive examination option is designed to evaluate a student’s ability to apply their Data Science knowledge into practical problem solving in a specific domain. In order to do so, students using this option are tested in three course-hosted comprehensive examinations, each in a different subject area.

This plan conforms to the guidance issued by the Graduate Council dated February 11, 2016 confirming compliance of this plan to Senate Regulation 700 and has been adopted on UC San Diego Campus.

In this format of the Comprehensive Examination, the students must answer comprehensive questions in their chosen domain in each of the three selected courses. The comprehensive examination is integrated into the host courses, and in most cases, the associated work serves dual purposes, contributing independently to the student’s course grade and comprehensive examination score. The comprehensive examination typically consists of a specific class assignment or examination, or a portion thereof, that has been explicitly approved by the MS program committee.

Determination of the outcome on the comprehensive exam is separate from the grade in the host course. The students are required to successfully pass the comprehensive examination in three courses drawn from each of the three areas: **ML/computing, math/statistics, algorithms/systems**. Students are permitted up to five attempts, that is, five different courses. No more than three course-hosted comprehensive examinations can be taken in a single quarter, and no comprehensive examination can be repeated in a single quarter. The courses marked for comprehensive examination can be taken only for a letter grade.

Course-hosted examinations are registered at the beginning of each quarter and students must register in advance by the specified deadline for the examination. The examination is supervised by a faculty committee responsible for the content, evaluation and administration of the examination which is separate from the course instructor who is responsible for the course grade but not success in the comprehensive examination.