DATA SCIENCE - GRADUATE CERTIFICATE (ONLINE)

The Master of Science in Data Science program hosted online through the Coursera platform offers stackable graduate-level courses, a graduate certificate, and a fully accredited master's degree in data science. Students earn the same credentials as on-campus students. There are no online or Coursera designations on official CU transcripts or diplomas.

Program Policies
This specialized program does not align with standard campus policies. Please refer to the Special Online Programs (https://catalog.colorado.edu/specialized-programs/) section of the catalog for more information.

Program Requirements
Admission Requirements
There are no admission processes for learners seeking to earn the Data Science Graduate Certificate. All courses attempted and/or completed for credit will appear on an official CU Boulder transcript (unless dropped by the drop deadline) and will count toward the cumulative GPA.

Prerequisite Knowledge
There are no formal prerequisites for the Data Science Graduate Certificate, but we recommend that prospective students be knowledgeable in Python, R Programming, Calculus including derivatives and integrals, and Linear Algebra including matrix multiplication, matrix inversion, and solving linear systems using matrices. If they are not, we encourage them to try out non-credit coursework before attempting for-credit courses.

Required Courses and Credits
To earn the Data Science Graduate Certificate (12 credits), students must complete the following required specializations:

- Data Mining Foundations and Practice (3 credits)
- Data Science Foundations: Statistical Inference (3 credits)

As well as two specializations from the following:

- Introduction to Statistical Learning for Data Science (3 credits)
- Machine Learning (3 credits)
- Statistical Modeling for Data Science (3 credits)

In order to earn a certificate, students must receive a minimum grade of a C or higher in each course. The cumulative GPA for certificate courses must be 3.0 or higher.

The Data Science Graduate Certificate requires 12 credit hours of coursework. Students must complete the follow required courses:

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credit Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTSA 5504</td>
<td>Data Mining Pipeline</td>
<td>1</td>
</tr>
<tr>
<td>DTSA 5505</td>
<td>Data Mining Methods</td>
<td>1</td>
</tr>
<tr>
<td>DTSA 5506</td>
<td>Data Mining Project</td>
<td>1</td>
</tr>
</tbody>
</table>

Additional Requirements
Students must complete an additional 6 credits and can choose to complete any 2 of the following 3 specializations:

Introduction to Statistical Learning for Data Science
- DTSA 5501 Probability Theory: Foundation for Data Science
- DTSA 5502 Statistical Inference for Estimation in Data Science
- DTSA 5503 Hypothesis Testing for Data Science

Machine Learning
- DTSA 5509 Introduction to Machine Learning - Supervised Learning
- DTSA 5510 Unsupervised Algorithms in Machine Learning
- DTSA 5511 Introduction to Deep Learning

Statistical Modeling for Data Science
- DTSA 5512 ANOVA and Experimental Design
- DTSA 5513 Generalized Linear Models and Nonparametric Regression

Learning Outcomes
Upon completing the program, students will be able to:

- Acquire, clean, wrangle and manage data.
- Correctly perform exploratory data analysis in order to assist with the generation of scientific hypotheses.
- Apply principles and methods of probability theory and statistics to draw rational conclusions from data.
- Construct an appropriate statistical model in order to answer important scientific or business-related questions.
- Assess the validity of a statistical model when applied to a particular dataset.
- Use statistical techniques to design an experiment.
- Understand and be able to apply the main computational techniques used to analyze large data sets, including a variety of data mining and machine learning approaches.
- Understand the principles of computer representation, storage and access of large data sets and be able to determine the appropriate approaches for specific problem.
- Clearly communicate the results of a data science analysis to a non-technical audience.