

Course Syllabus for DS 730: Big Data: High-Performance Computing

NOTE: This syllabus document contains the basic information of this course. The most current syllabus is available in the full course.

Course Description

This course teaches you how to process large datasets efficiently. You will be introduced to non-relational databases. You will learn algorithms that allow for the distributed processing of large data sets across clusters.

Prerequisite: DS 710

Course Objectives

By the end of this course, you will be able to:

- Implement algorithms that allow for the distributed processing of large data sets across computing clusters.
- Create parallel algorithms that can process large datasets.
- Process large datasets efficiently.

Course Components

Activities: This course's activities serve the same purpose as regular assignments do in other courses by asking you to demonstrate your understanding of a particular lesson's content. Activities are smaller than projects, and you generally have a lesson to complete them.

Projects: The projects are larger than activities and ask you to synthesize content from several lessons. You usually have two lessons to complete projects.

Grading

Your mastery of course content is assessed using a variety of methods:

Activity	Percentage
Activities	30%
Projects	70%
Total	100%

Final grades are assigned using the following scale:

90–100%	A
80–89%	B

60–79%	C
0–59%	F