Course Syllabus for DS 710: Programming for Data Science

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

Course Description

Computer programming is an essential part of data science. When working with large data sets, it's especially important to be able to write effective, efficient code to help you organize and understand the data. In this course, we'll introduce you to two of the most widely-used programming languages for data science: Python and R. You'll gain experience working with real-world data, and leave the course with skills you can apply in other courses in the MS Data Science Program as well as on the job!

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

- Use GitHub, R, and Python for basic data analysis and version control.
- Use control flow.
- Perform exploratory data analysis in R.
- Use strings, lists, and tuples in python.
- Test statistical hypotheses using R functions.
- Create readable functions for efficient task management.
- Manipulate real-world data to enable further analysis.
- Choose and implement efficient methods for analyzing data sets.
- Gather real-world data that can be analyzed to address a modeling question.
- Synthesize analysis of gathered data for stakeholders.

Grading Policy

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

Activity	Percentage of Grade
Homework	60%
Participation in required learning activities	5%
Quizzes	15%
Final Project	20%

Total	100%

Final grades are assigned using the following scale.

Note: A final grade within 0.5% of the next highest letter grade will be rounded up <u>only</u> if the average homework score is greater than the cutoff for that letter.

─────────────────

A   90-100%

─────────────────

B   80-89%

─────────────────

C   70-79%

─────────────────

D   60-69%

─────────────────

F   At or below 59%