



Independent Research in Data Science

Schedule

Tuesday and Thursday 9:30AM to 10:50AM *Phi Beta Kappa Hall 221*

Instructor

Dr. **Dan** Runfola

danr@wm.edu

ISC 1269

Office Hours: 11-1P TR

757.221.1970

Please let me know if you have any documented disabilities that may impact your performance in this class.

Course Description: This course is designed to permit students with a focus in Data Science to engage in independent research after completing DATA 201 (Data Driven Decisionmaking). Working closely with a program faculty member as an advisor, each student will conduct a substantial research project focusing on synthesis and critical analysis, to solve problems in an applied and/or academic setting, to create original material or original scholarship, and to communicate effectively with a diversity of audiences.

Prerequisite(s): None.

Credit Hours: 3

Materials:

A free SciClone account is needed in order to access the William and Mary High Performance Cluster (HPC). Registration can be started at <https://hpc.wm.edu/acctreq/>.

A computer is required to access the SciClone cluster. Students without access to a reliable computer can use resources in the Data Science and Visualization Lab (ISC 1109).

Course Objectives:

1. Provide students with a critical understanding of the decision making process and the use of data and intuition in applied contexts.
2. Develop students ability to communicate findings, analysis, and visualization skills for future courses (and jobs).

3. Expose students to real-world problems that are being engaged with by contemporary problem solvers and decision makers.
4. Provide an opportunity to earn credit towards the COLL 400 requirement.

Grade Distribution:

Reading Assignments	10%
Projects	60%
Final Report	30%

Letter Grade Distribution:

≥ 93.00	A	73.00 - 76.99	C
90.00 - 92.99	A-	70.00 - 72.99	C-
87.00 - 89.99	B+	67.00 - 69.99	D+
83.00 - 86.99	B	63.00 - 66.99	D
80.00 - 82.99	B-	60.00 - 62.99	D-
77.00 - 79.99	C+	≤ 59.99	F

Time Commitment: Excelling in college level course work typically requires on average three to four hours per credit per week. For each credit you take in this class, in addition to the time we meet as a class each week, you should expect to spend three additional hours on average reading, writing, or otherwise preparing for this class on a weekly basis.

Attendance: This class does not have an attendance policy. However, it will be difficult to learn enough to pass the class without regular participation, as the majority of course content relevant to tests and assignments will be covered in class.

Discussions and Reading Assignments: Most lecture sessions will begin with a discussion of the assigned materials. As such, most weeks students are required to write a short summary (no more than 1 page) reflecting on a given weeks assigned readings - these summaries can represent questions the material raised, commentary, or critiques.

Classroom Behavior: Please remain civil during discussions to promote the open exchange of ideas and foster a culture of open dialogue. Please bear in mind that all students are entitled to their own opinion. You are expected to listen attentively to each person speaking. Please refrain from eating during class (and, if you must, make sure it isn't loud!).

Teacher-student conferences: Students performing at a C level or below are required to schedule a meeting with the instructor to discuss class performance.

Late / Poor Performance Policy: Assignments will not be accepted late, excepting in documented circumstances (i.e., an illness with a doctor's note).

Final Project: The final report will involve the application of the skills taught throughout the semester, and will represent your ability to provide an answer to a real-world question using the data and skills at your disposal. It will include a 5 page written report (inclusive of data visualizations and tables) designed as an executive summary for an external decisionmaker.

Important Dates: The add and drop deadline this semester is September 8th, and withdrawal deadline is October 27th.

Do not cheat!

Academic dishonesty is taken very seriously. Make sure to cite all of your work, and do not turn in work that is not yours! Cases of academic dishonesty will be evaluated and acted upon in accordance with William and Mary policies, which can be found at <http://www.wm.edu/offices/deanofstudents/services/studentconduct/>

Course Outline:

The course outline can be found below. The weekly content might change as it depends on the progress of the class.

Week	Content
Week 1	<ul style="list-style-type: none">• 8/31 - Introduction
Week 2	<ul style="list-style-type: none">• 9/5 - Eric Walter, HPC Account Setup; 9/7 - Problem Definitions
Week 3	<ul style="list-style-type: none">• 9/12 and 9/14 - Defining your Question - Literature Review• Project 1 assigned, due midnight Fri Sept 22 - Annotated Bibliography pt. 1
Week 4	<ul style="list-style-type: none">• 9/18 and 9/21 - Defining your Question - Literature Review• Project 2 assigned, due midnight Fri Sept 29 - Annotated Bibliography pt. 2
Week 5	<ul style="list-style-type: none">• 9/26 and 9/28 - Data Sources• Project 2 due midnight, Friday Sept 29.• Project 3 assigned, due midnight Friday October 13th - Data Source Identificaiton
Week 6	<ul style="list-style-type: none">• 10/3 and 10/5 - Data Sources II
Week 7	<ul style="list-style-type: none">• 10/10 and 10/12 - Dataset Integration pt. 1• Project 3 due by midnight, Friday October 13.
Week 8	<ul style="list-style-type: none">• 10/10 and 10/12 - Dataset Integration pt. 2• Project 4 assigned, due midnight Friday 11/3 - Dataset Integration

Week	Content
Week 9	<ul style="list-style-type: none"> • 10/24 and 10/26 - Model Selection I
Week 10	<ul style="list-style-type: none"> • 10/31 and 11/2 - Model Selection II • Project 4 due by Midnight Friday, November 3rd. • Project 5 assigned, due midnight Friday 12/1 - Contrasting Model Approaches
Week 11	<ul style="list-style-type: none"> • 11/7 and 11/9 - Visualization and Communication
Week 12	<ul style="list-style-type: none"> • 11/14 and 11/16 - Visualization and Communication II
Week 13	<ul style="list-style-type: none"> • 11/21 - Hands on Learning, 11/23 - No class (Thanksgiving Break)
Week 14	<ul style="list-style-type: none"> • 11/27 and 11/30 - Visualization and Communicaiton III • Project 5 due on Friday, December 1st by midnight.
Week 15	<ul style="list-style-type: none"> • 12/5 and 12/7 - Final Project Introductions, Data Validation and Uncertainty I • Final project assigned, due on December 17th by midnight.
Final Project	<ul style="list-style-type: none"> • Due on December 17th by Midnight