CDS 292: Network Analysis

Syllabus Part 1

Spring, 2022

This syllabus is split into three parts:

1. Part 1: General Course description
2. Part 2: Instructors
3. Part 3: Calendar

Course Introduction.

The purpose of CDS 292 is to introduce the novice data analyst to the basic concepts of networks, as well as the mathematical, statistical, and computational tools that support the application of those concepts. Networks have become an essential tool in the world of data analysis, and thus in disciplines as disparate as sociology, biology, arts, and physics, networks show up now with abundant regularity.

The course uses the programming language `Python`, which is one of the most versatile and extensively used computational tools at present. The versatility of `Python` in data analysis is well-established and currently it is a tool no data scientists can do without.

CDS 292 is a true amalgamated course, with mathematical, computational, statistical, and data analytic material. To have a sense about what the course tries to achieve, by the end of the course the successful student should 1) be able to perform data analysis using networks and its various concepts, 2) understand the fundamental properties of each concept learned to apply them correctly, 3) be able to error check code and libraries constructed by others (lots of shared code on the Internet has errors) and finally, 4) know enough to understand how to build a networks algorithm from scratch, error check it, and use it.

This course is a Quantitative Reasoning Mason Core course, designed to fulfill the relevant learning outcomes. It does this by focusing on the teaching, use, and interpretation of quantitative analysis to study concrete theoretical and real-world problems. **Quantitative reasoning will be expected of students!**

Disclaimer: It is common for students to join this class thinking it is about social media and/or the use of packages to study social networks. This is incorrect.

This class is primarily about an introduction to network concepts and network analysis broadly defined, not focused specifically on social media, or social networks although some do appear. Also, although we do learn how to code routines using `Python` in the network setting, this class is not designed to focus on a computational tool, library, or language.

Advice to succeed in this class

Read the book, slides, notebooks and solutions a few times. There are many little concepts that keep popping up over and over, and they are scattered so unless you read, they will escape you.

Practice a lot of code. Programming is like playing a musical instrument, which can always be done better if you like it, practice, and have some natural talent. I am sure you have all three, but the one you can control is practice. Make variations on the code, try new things, etc.

Dust off your old high-school math. The course does rely on the use of some algebra, matrices, and arithmetic. This goes with the Quantitative Reasoning Core designation. Please, please do not underestimate this. Many students join the class proficient in coding but rusty in math. Data Science relies on people understanding the meaning of quantities and numbers. This class fully embraces that notion and there is an expectation that you will use the class to improve those skills.

Once solutions are provided, study them. It cannot be stressed enough how this is essential prep for exams.

Take full advantage of collaborative work. There are activities designed specifically for you to engage with other students. Make use of them. Peer learning is an important part of learning. But remember your problem set submissions are only your own!

Don't delay in talking to your instructor if there's a problem. If you wait, the problem may get worse and your instructors may have less options/time to help you through it.

Textbook

It is made available to you on Blackboard.

Computer IT Support

If you are having difficulties with your GMU network account, or with software on the GMU network, please contact IT support: support@gmu.edu. The ITS Support Center is in Innovation Hall, Room 226 or can be reached at: 703-993-8870.

Learning Outcomes and Quantitative Reasoning Mason Core

General

1. Students are able to interpret quantitative information (i.e., formulas, graphs, tables, models, and schematics) and draw inferences from them. Students are able to decant information to determine if it has network structure.

2. Given a quantitative problem, students are able to formulate the problem quantitatively and use appropriate arithmetic, algebraic, and/or statistical methods to solve the problem. Matrix algebra is part of the methodology of this course.

3. Students are able to evaluate logical arguments using quantitative network theory reasoning.

4. Students are able to communicate and present quantitative results effectively. This is achieved by a set of continuing in-class activities where students are required to apply network techniques to a data problem, analyze the results, and defend them to the class and instructor.

Specific Learning Lesson Outcomes

For detailed information on each lesson, please see **Syllabus Part 3**

Lesson assignments

For due dates of assignments, please consult **Syllabus Part 3**.

Course Material Availability

Lesson slides, videos, `Python` notebooks, problem set assignments, colaboratory assignments, solutions, other class materials, and assignments and exam scores will be posted to the class Blackboard site.

Policies and Procedures

Numerical/Letter grade correspondence

Grade scored g	Will equal
$97\% \leq g \leq 100\%$	A+
$94\% \leq g < 97\%$	A
$90\% \leq g < 94\%$	A-
$87\% \leq g < 90\%$	B+
$84\% \leq g < 87\%$	B
$80\% \leq g < 84\%$	A-
$77\% \leq g < 80\%$	C+
$74\% \leq g < 77\%$	C
$70\% \leq g < 74\%$	C-
$60\% \leq g < 70\%$	D
$0\% \leq g < 60\%$	F

Grading

- When grading assignments, while correct answers are important, it is more important how you arrived at those answers, including derivations, logic used, algorithms, and good problem solving strategies. Some (minimal) credit is given for the answer and most credit is given for the correct work leading to the answer.
- Full credit is given for correct answers provided ALL supporting work is shown in an organized manner showing how the answer was arrived at.
- Semester grades will be computed from the following percentages:
  - 45% Assignment Grade:
    - 15% for colaboratory assignments

    - 30% for problem sets
  - 25% for midterm exam
  - 30% for final exam
  - Participation has a value of 10% as extra credit (see below)

Assignments

All `Python` Problem Set submissions MUST ORIGINATE ON PAPER, then be scanned, and then be sent. For scanning, the gold standard is a standalone scanner. You may use phone apps to scan if you do not have access to a standalone scanner. Some options:

- Adobe Scan app
- Microsoft Office Lens

Some recommendations:

- **Do not send problem sets pictures**. They are hard to read. Think: if you have trouble reading it from a picture, we have even more trouble as we have to read many such submissions.
- **Send a SINGLE pdf document** containing your solutions.
- **CORRECT INDENTATION OF `Python` CODE IS A CONDITION TO CORRECTNESS**. If you are not sure about your handwriting, use a symbol to mark an indented block. Just the correct number of right pointing arrows, each signifying an indentation level, should be enough. You can use other options too, but make sure they are easy for your teaching assistant to read and decipher.
- **Time extensions may ONLY be granted by the instructor**, must be in written (email) from, and will give a specific time. **Late assignments sent directly to your TA will NOT be graded**.
- The total available points in each Problem Set can vary week to week.
- The semester Problem Set grade will consist of the averaged sum of the individual Problem Set grades. Specifically,

$$\text{Problem Set Grade} = \frac{\text{(Sum of points you obtained)}}{\text{(Sum of points over all sets)}}$$

- Problem Sets are generally released a **week before the due date**.

Colaboratory Assignments

- Every week, there will be a set of coding problems intended for you to solve with your **programming partner(s)**.
- One of you will import the provided notebook to Colaboratory, invite partner(s), and start working.
- An important part of the assignment is that you organize the way you work. This is to avoid overwriting each other's code. There is a video demonstrating how to do this (see Lesson 0 in Blackboard).

- Once you have finished with the notebook. **MAKE SURE THAT AT THE TOP OF THE NOTEBOOK THERE IS A TEXT CELL WITH YOUR NAMES**. Then, invite your teaching assistant (emails specified in Part 2 of the syllabus) to the notebook and he/she will be able to grade it.
- To submit your work, **one of you should go to the submission page in Blackboard, and upload a downloaded version of your colaboratory notebook**. This is just so we keep track of the date you sent your work.
- The file uploaded to Blackboard must have the following structure: `ColaboratoryXX-GroupYY.ipynb`; where `XX` is the number of the colaboratory assignment and `YY` your group number (you can see this in Blackboard).
- All Notebooks have equal weight.
- The semester Notebook grade will consist of the averaged sum of the individual Notebook grades. Specifically,

$$\text{Notebook Grade} = \frac{\text{(Sum of points you obtained)}}{\text{(Sum of points over all Notebooks)}}$$

- Notebooks are generally released **a week before the due date**.

Exams

- The exams will be available on Blackboard with the **Respondus Lockdown Brower**.
- Once you open the exam you will have a set amount of time to complete it (90 minutes for the midterm and 2 hours and 45 minutes for the final) unless you open with less time than the exam duration before the exam closes. In that case, you will only have the remaining time between when you open the exam and the closing time.
- Dates for the midterm and final are stated in Part 3 of the syllabus.
- NOTE ON RESPONDUS: If you have installed RESPONDUS at NVCC or any other college/school/..., it MUST be UNINSTALLED and REINSTALLED for use at MASON. This is due to the internal handshaking that occurs between MASON's Blackboard and RESPONDUS.
- Students will be REQUIRED to show their GMU ID card PRIOR to beginning either the Midterm or Final exams. (Remember: "No GMU ID - No exam grade - No exception").

Participation

This grade is only extra credit. If a student does not feel the need to tap into the participation extra credit grade, they will still have access to 100% of the points. To receive credit for participation, students must:

- attend office hours with both the instructor and the teaching assistants, and
- contribute to any discussions, or pose questions.

Over time, we have found that students find themselves in trouble by misunderstanding certain basic things about the class. In addition, office hours are grossly underused but students that do participate gain a great deal. Therefore, we have introduced an incentive structure to have students actively participate in this activity.

More comments on Grading

- While correct answers matter, it is more important how you arrived at those answers. Completeness in answers (complete thoughts, complete codes, complete plots) is about $\frac{1}{2}$ credit for the problem (or more credit if the problem is particularly challenging).
- Semester grades will be computed as a percentage of points earned divided by total points available in each category. Grades are mathematically determined and are not typically curved.
- There are no extra credit assignments/projects available for CDS-292.

Technology Needed and GMU netlogin

- Each student will require access to Blackboard from GMU. Make sure you have access to your GMU netlogin.
- Each student will require checking email often and therefore also requires have up to date
- Software access for the class:
  - PDF reader with the capacity to display pages in single page view (for slides), and
  - a computer capable of executing `Python` code, and the `networkx` and `matplotlib` libraries,
  - a computer with a browser to use Blackboard and Respondus browser (needs to be Mason Installed or it doesn't work) with a webcam or a phone with a webcam (consult Lesson 0 for details).
  - Access to a Google account is highly recommended (Colaboratory).

Collaboration and Plagiarism

- All CDS-292 activities are subject to GMU's Honor Code and IT policies.
- Collaboration: Students are encouraged to discuss homework problems with each other. Discussion means each student working the problem may talk with someone else (aka: conversation on the logic or software needed to compete the assignment) but will fully work the problem on their own.
- Plagiarism will not be tolerated at any time. Students will be given a zero for any assignment, quiz, or exam where plagiarism is suspected. If plagiarism is suspected a second time for any student, an automatic grade of "F" will be assigned for the course with a report sent to the Dean for further action.
- Collaboration becomes plagiarism when: **A STUDENT COPIES THE WORK OF SOMEONE ELSE, EITHER FROM STUDENTS CURENTLY TAKING CDS-292, OR WHO HAVE TAKEN CDS-292 IN THE PAST OR BY COPYING THE WORK FROM ANY PUBLISHED OR UNPUBLISHED SOURCE**.

- All problem sets, notebook assignments, and exams, including computer programs and associated outputs, turned in for grading represent the student's own work.
- Students may not discuss exam problems with anyone other than your instructor or the CDS-292 TA.
- COLLABORATION DURING AN EXAM OR QUIZ IS NOT AUTHORIZED AT ANY TIME AND THUS PROHIBITED AND IS CONSIDERED PLAGIARISM!
- All plagiarism violations will be reported in writing to the Dean's office.

Prohibited Equipment use during exams and quizzes

Exams are closed book/closed notes. The only allowed equipment is the blackboard site under respondus, paper and pencil. Therefore, other computer equipment, smart phone, calculator or other electronic device use during an exam is prohibited. Individual instructors may announce modifications to this policy which will be specific to their class section only.

Disabilities

If you have a documented learning disability or other condition that may affect academic performance students MUST:

1. Have the need for accommodation on file with Office of Disability Services (SUB I, Rm. 4205; 703-993-2474; http://ods.gmu.edu);
2. Provide the instructor with a copy of the Office of Disability Services accommodation determination prior to receiving any accommodations. We will protect this information as private and will not share the information with anyone other than the class assistants unless authorized in writing by the student or the Office of Disability Services.
3. In case you need it, we can arrange for an alternate testing location for the Midterm and Final Exams with PRIOR notice, which is needed AT LEAST 1 week in advance although it is better if even more time is allowed.
4. PLEASE NOTE: If you are having ANY difficulties with CDS-292 due to personal limitations, PLEASE let the instructor know. We want to help you succeed in CDS-292 and in your GMU career.

CDS 292: Network Analysis

Syllabus Part 2

Spring, 2022

This syllabus is split into three parts:

1. Part 1: General Course description
2. Part 2: Instructors
3. Part 3: Calendar

Section 1: 12.00 - 1.15 PM

Instructor Valentin Vergara Hidd

Email vvergara@gmu.edu

Class Meetings EXPL 2312

Office Hours Every Tuesday from 2PM to 3PM I will hold *walk-in* office hours in this zoom meeting room. If you need additional office hours, send me an email and we can schedule it. Please, allow some time for me to schedule our meeting (i.e. email me at least 1 day before the time you want to meet).

Time to respond emails We created this class to minimize the amount of work you should do over weekends. With that in mind, I will respond to emails within a day during the week (Monday to Friday).

Section DL: Asynchronous class

Instructor Eduardo López

Email elopez22@gmu.edu

Class Meetings This is an asynchornous section. There are no scheduled class meetings.

Office Hours Thursday 1.30PM to 2.45PM in this zoom meeting room. Password is **912471**.

Time to respond emails I will respond to email messages within 24 to 48 hrs, Monday through Friday. Emails sent during weekends and holidays will be responded to on the following Monday.

STARs (both sections)

Formally, the STARs for this class are assigned to a section. However, feel free to reach to both of them for office hours or content questions. If you need help with topics such as grading reviews, contact the STAR assigned to your section.

Radhika Laddha: Section 1

1. **Email** rladdha@gmu.edu
2. **Colaboratory Assignment email** radhikalre@gmail.com
3. **Office Hours** Monday 3.00 - 4.00 PM in this zoom meeting room. Password is **8fTM3T**
4. **Time to respond emails** Weekdays.

Sofia Escoto: Section DL

1. **Email** sescoto@gmu.edu
2. **Colaboratory Assignment email** seescoto2000@gmail.com
3. **Office Hours** Wednesday 4.30PM - 5.30PM in this zoom meeting room. Password is **044711**. Sofia plans to hold in person office hours after the first few weeks of the semester. Time and place will be announced in Blackboard.
4. **Time to respond emails** Usually less than 24 hours.

CDS 292: Network Analysis

Syllabus Part 3

Spring, 2022

This syllabus is split into three parts:

1. Part 1: General Course description
2. Part 2: Instructors
3. Part 3: Calendar

Definition of a week: This semester, each week is defined from the beginning of Saturday (12:00 AM) to the end of the following Friday (11:59 PM)

Snow Day Policy

If Mason is closed due to inclement weather on a class meeting day (only section 1), the class is **canceled**. Section DL will have no changes due to inclement weather.

Calendar

Calendar week	Lesson	Due dates
Jan 22 - Jan 28	Lesson 0: Navigating CDS 292; Lesson 1: Tools	• FERPA form related to Colaboratory use: Wednesday Jan 26th • Signed Statement that Lesson 0 has been read and understood: Wednesday Jan 26th • Colaboratory notebook: Friday Jan 28th • Problem Set: Friday Jan 28th

Jan 29 - Feb 4	Lesson 2: Building blocks of networks	• Colaboratory notebook: Wednesday Feb 2nd, by the end of the day (11:59 pm) • Problem Set: Friday Feb 4th
Feb 5 - Feb 11	Lesson 3 (part 1): Node Degrees, Link Indicators, Network Formulas, and Adjacency Matrices	• Colaboratory notebook: Wednesday Feb 9th • Problem Set: Friday Feb 11th
Feb 12 - Feb 18	Lesson 3 (part 2): Node Degrees, Link Indicators, Network Formulas, and Adjacency Matrices	• Colaboratory notebook: Wednesday Feb 16th • Problem Set: Friday Feb 18th
Feb 19 - Feb 25	Lesson 3 (part 3): Node Degrees, Link Indicators, Network Formulas, and Adjacency Matrices	• Colaboratory notebook: Wednesday Feb 23rd • Problem Set: Friday Feb 25th
Feb 26 - Mar 4	Lesson 4 (part 1): Histograms of Node Degrees	• Colaboratory notebook: Wednesday Mar 2nd • Problem Set: Friday Mar 4th
Mar 5 - Mar 11	Lesson 4 (part 2): Histograms of Node Degrees	• Colaboratory notebook: Wednesday Mar 9th • Problem Set: Friday Mar 11th
Mar 19 - Mar 25 Mar 26 - Apr 1	Review and Midterm Lesson 5 (part 1): Paths in networks	Midterm Available Mar 23rd • Colaboratory notebook: Wednesday Mar 30th • Problem Set: Friday Apr 1st

Apr 2 - Apr 8	Lesson 5 (part 2): Paths in networks	• Colaboratory notebook: Wednesday Apr 6th • Problem Set: Friday Apr 8th
Apr 9 - Apr 15	Lesson 6: Shortest paths in networks	• Colaboratory notebook: Wednesday Apr 13th • Problem Set: Friday Apr 15th
Apr 16 - Apr 22	Lesson 7 (part 1): Triangles, v-shapes and clustering	• Colaboratory notebook: Wednesday Apr 20th • Problem Set: Friday Apr 22nd
Apr 23 - Apr 29	Lesson 7 (part 2): Triangles, v-shapes and clustering	• Colaboratory notebook: Wednesday Apr 27th • Problem Set: Friday Apr 29th
Apr 30 - May 6	Lesson 7 (part 3): Triangles, v-shapes and clustering	• Colaboratory notebook: Wednesday May 4th • Problem Set: Friday May 6th
May 9th - May 11th	Review Sessions	
May 12th	Final Exam	