**MATH517 Final Research Project Summer 2022**

The final project is the most rewarding part of the class, offering the students the opportunity to combine and utilize all the knowledge they acquired. Students are asked to select a network of interest to them, map it out and analyze it. Some procedural details enrich this assignment:

1. The project is carried out in pairs. The instructor does not do the pairing, but students are encouraged to find their partners.
2. A few weeks into the course one class is devoted to *preliminary project presentations*. Each group is asked to offer a five-minute presentation with no more than five slides, offering a preview of the dataset they selected. Students are advised to collect their own data - simply downloading a dataset already prepared for network analysis is not acceptable. Indeed, one of the goals of the project is to experience the choices and compromises one must make in network mapping. Manual mapping is allowed, like looking up the ingredients of recipes in a cookbook or the interaction of characters in a novel or a historical text. Digital mapping is encouraged, like scrapping data from a website or a database that is not explicitly organized as a network map, but the students must reinterpret and clean the data to make it amenable for network analysis. For example, one can systematically scrap data from Wikipedia to identify relationships between writers, scientists, or concepts.
3. It is important to always emphasize that the purpose of the final project is to test a student's ability to analyze a network. Consequently, students must stay focused on exploring the network aspect of the data and avoid being carried away by other tempting questions their dataset poses that would take them away from this goal.
4. The course ends with the final project presentations

**Final project guidelines**

Measure:

* N(t), L(t) [t- time if you have a time dependent system);
* P(k) (degree distribution).
* <l> average path length.
* C (clustering coefficient).
* Crand.
* C(k).
* The weighted distribution P(w) if you have a weighted network.

Visualize communities.

Discuss network robustness and spreading, degree correlations, whichever is appropriate.

It is not sufficient to simply measure things– you need to discuss the insights they offer:

* What was your expectation?
* What is the proper random reference?
* How do the results compare to your expectations?
* What did you learn from each quantity you measured?

Time frame will be strictly enforced.

Approximately 12 min + 3 min questions.

No need to write a report—you will hand in the presentation.

Send us an email with names/titles/program.

Come earlier and try out your slides with the projector. Show an entry of the data source—just to have a sense of how the source looks like. On the slide, give your program/name.

Grading criteria:

Use of network tools (completeness/correctness).

Ability to extract information/insights from your data using the network tools.

Overall quality of the project/presentation.

**Final Preliminary Project Guidelines**

Present 5 slides in no more than five minutes.

* Introduce your network, discussion its nodes and links.
* Tell us how you will collect the data and estimate size of the network(N, L). Make sure that N > 100.
* Tel us what questions you are planning to ask. We understand that they may change as you advance with your project and the class.
* Tel us why you care about your network