CSSS 567 Homework 1

Jess Kunke, 10/15/2021

Instructions

Please submit either

1. a knitted R Markdown file (html or pdf) that shows your final code and results, or
2. a pdf of your report and a separate R file with your code.

Please title all files you submit with "HW1_" and your last name. For example, I would submit my homework as HW1_Kunke.pdf (and HW1_Kunke.R if I had a separate R file with my code).

Your code should be neatly organized and commented, and I should be able to run it from start to finish to reproduce your results.

For the more specific questions below, there is often just one right answer (e.g. the graph is either directed or undirected), but other decisions are more open-ended, such as how you plot the graph, exactly what details you decide to include, etc. As long as you hit all the questions outlined here and provide reasonable results, you are on track.

If you have questions, feel free to reach out!

Data

For this homework, we will use the `UKfaculty` data set from the `igraphdata` package:

```
library(igraphdata)
data(UKfaculty)
```

1. Data description and visualization

Take a look at the help documentation using `?UKfaculty` as we've done in class, and answer the following:

- In a sentence, what does this data set represent?
- What paper, book, or other resource is the source of this data set?

Use skills from Lab 1 to answer the following questions. Summarize your answers in paragraphs and also provide the code you used to obtain these answers.

- How many vertices and edges does the data set have?
- Is this graph directed or undirected? What other properties can you comment on?
- Describe each of the attributes included in this data set. What are they called, what format/type do they have, what are they attributes of (vertices, edges, etc), what do they mean?
- Compute and comment on two different centrality measures for this data set.
- Compute and briefly comment on some measure of assortativity for this data set.
- Compute/explore at least one other aspect of the data (see labs for ideas) and comment on it.

Plot the network. Try a few layouts, then choose one you think is informative or appropriate (there is no particular right answer), and comment on the graph structure.

2. Graph partitioning

Pick two graph partitioning methods (including at least one spectral method), state which two methods you chose, and apply them to this data set. Do the results look similar? Would you draw different conclusions from one partitioning versus the other? Bonus: why might you prefer one over the other?