Research

Tim Chartier

Depart. of Math & CS

U Turn
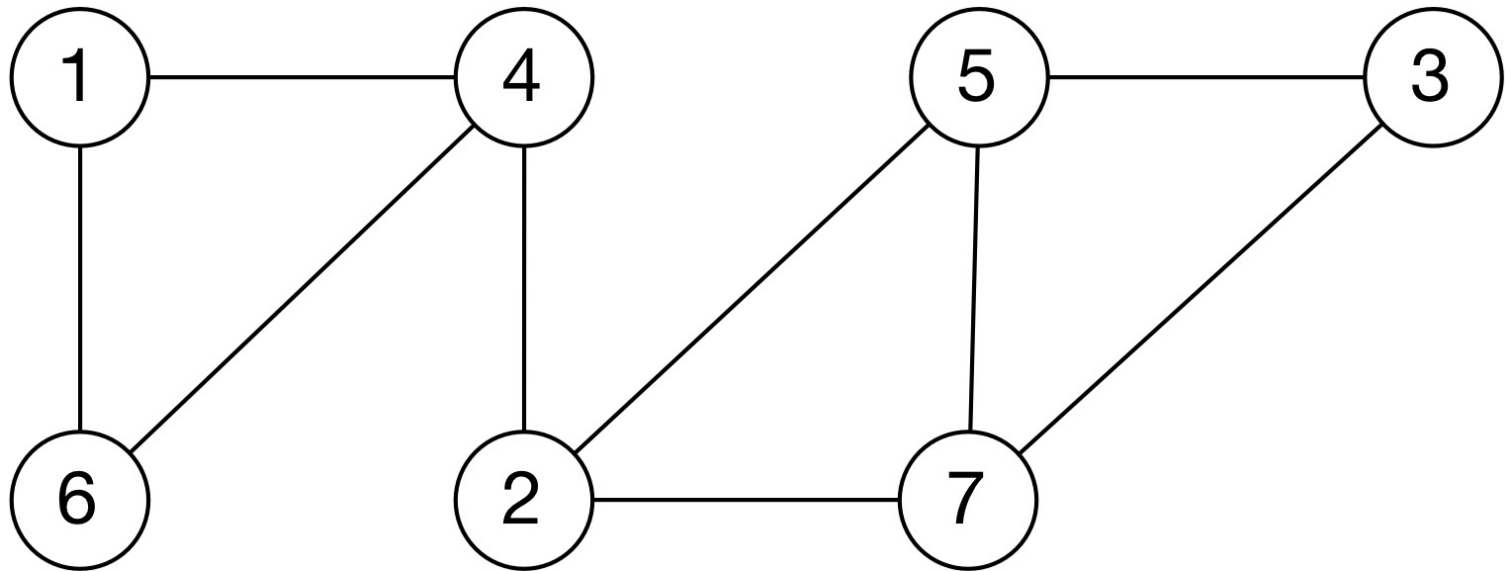
Let's share our research ideas.

MORE TOPIC

grouping

graphic groupie

- Let's learn to cluster vertices in a graph.
- Note, our graph is undirected.

graph 2 matrix

First, we from the adjacency matrix.

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

diagonally

Form diagonal matrix *D*, where d_{ii} = row sum of row i.

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Laplacian

Form Laplacian matrix *L = D − A*, so

$$L = \begin{pmatrix} 2 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 3 & 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & 2 & 0 & -1 & 0 & -1 \\ -1 & -1 & 0 & 3 & 0 & -1 & 0 \\ 0 & -1 & -1 & 0 & 3 & 0 & -1 \\ -1 & 0 & 0 & -1 & 0 & 2 & 0 \\ 0 & -1 & -1 & 0 & -1 & 0 & 3 \end{pmatrix}.$$

find a clique

- Miroslav Fiedler proved that the e-vector associated with second smallest e-value (called the Fiedler vector) will partition a graph into maximally intraconnected components and minimally interconnected components.
- Think of a close group of friends. They have strong friendships in their clique but fewer connections outside the group.

e-vectors

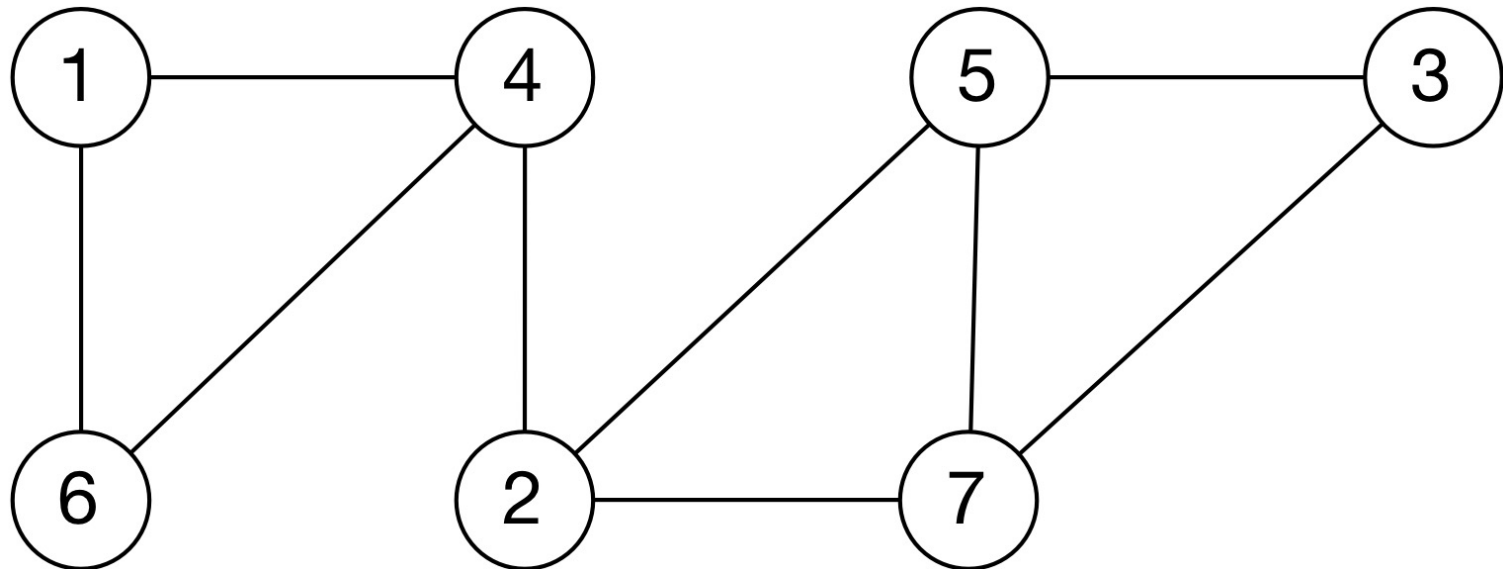- From linear algebra, **v** is an eigenvector of A with associated eigenvalue λ if
$$A\mathbf{v} = \lambda\mathbf{v}$$
- We can compute these numerically.
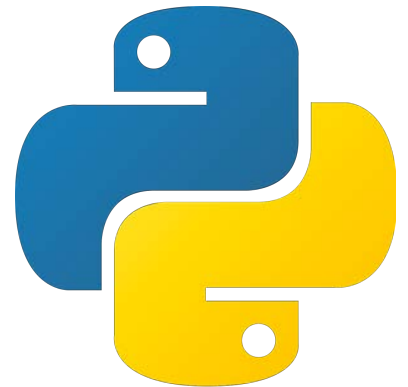
find the sign

- To cluster, use the sign of each element of the Fielder eigenvector.
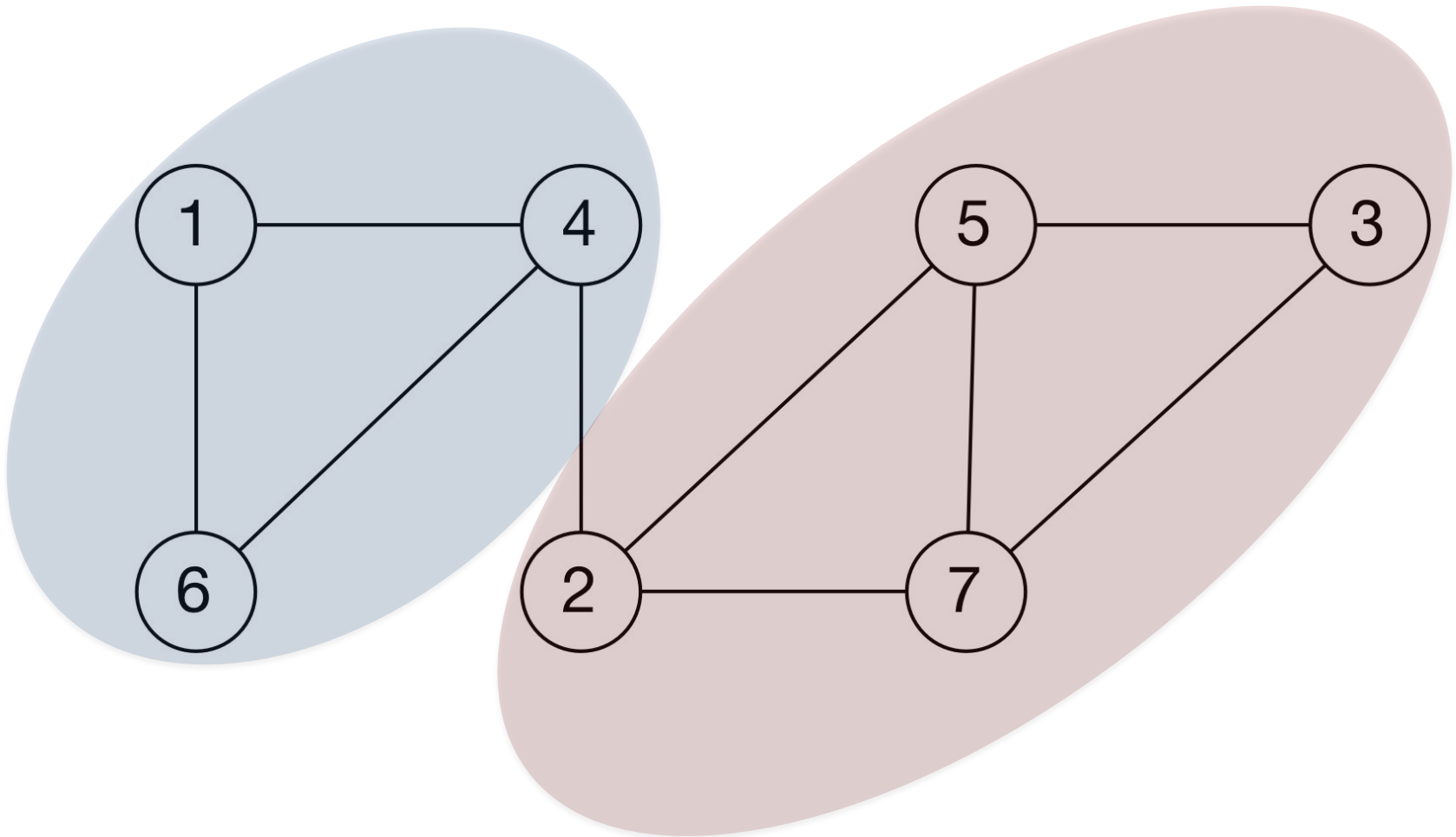- The rows with the same sign are placed in the same cluster.

cluster

So, we place nodes 1, 4, and 6 into one cluster and nodes 2, 3, 5, and 7 into the other.

$$
\begin{pmatrix}
0.4801 \\
-0.1471 \\
-0.4244 \\
0.3078 \\
-0.3482 \\
0.4801 \\
-0.3482
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
+ \\
- \\
- \\
+ \\
- \\
+ \\
-
\end{pmatrix}
$$

# clustered

Here are the clusters.

more clusters

Returning to our first example, if we want more clusters, add the next eigenvector (associated with the 3rd smallest e-value).

$$
\begin{pmatrix}
0.4801 & -0.2143 \\
-0.1471 & 0.6266 \\
-0.4244 & -0.6515 \\
0.3078 & 0.2735 \\
-0.3482 & 0.0900 \\
0.4801 & -0.2143 \\
-0.3482 & 0.0900
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
+ & - \\
- & + \\
- & - \\
+ & + \\
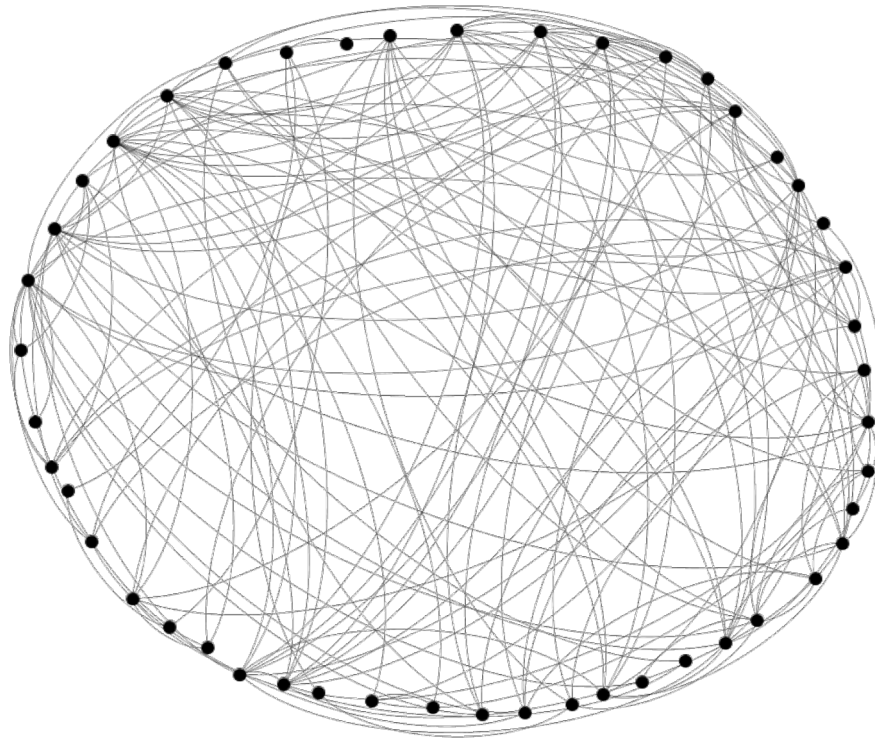- & + \\
+ & - \\
- & +
\end{pmatrix}
$$

clustered

Here are the clusters. Note the number of singletons, which can indicate over-clustering.
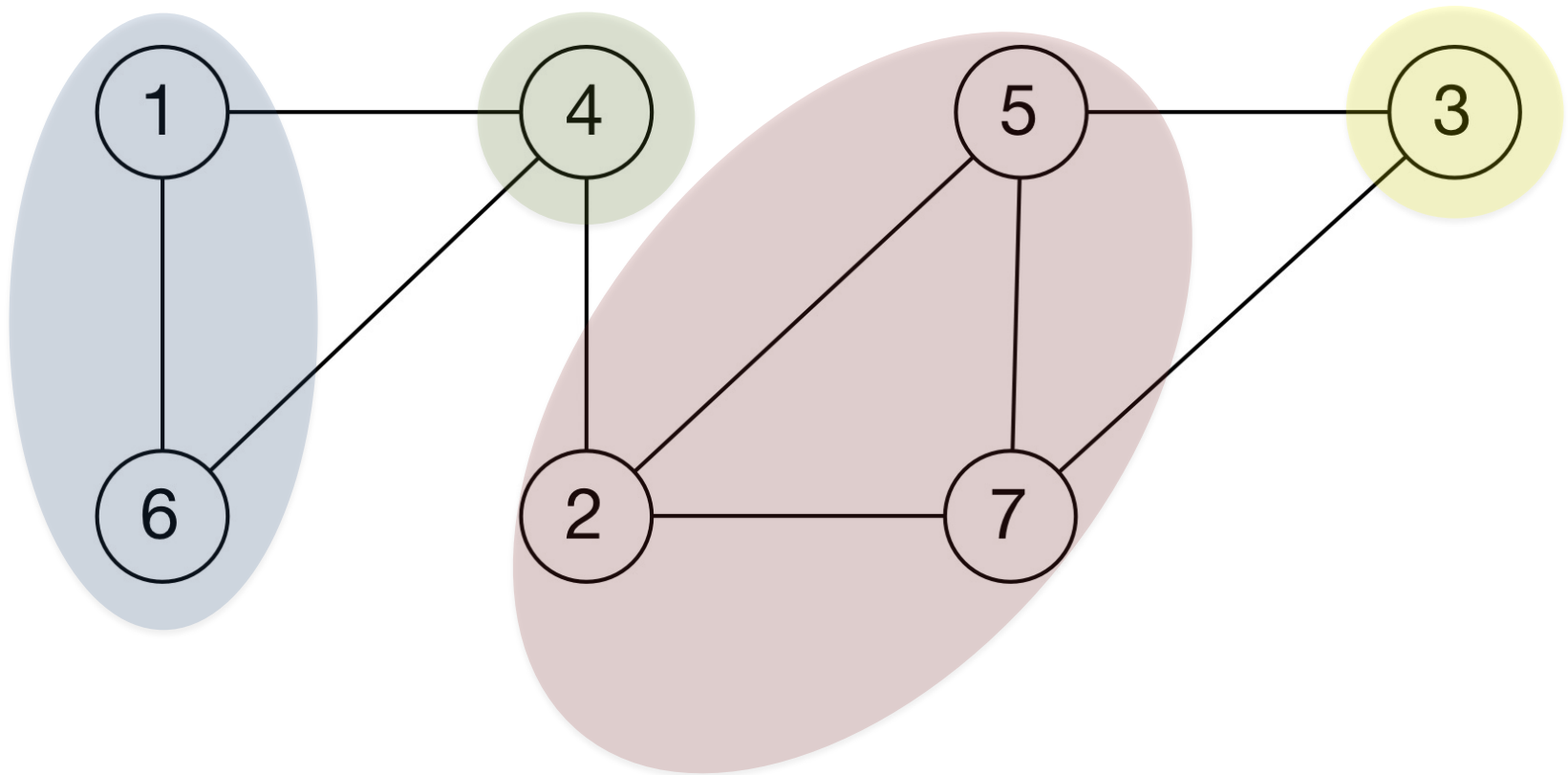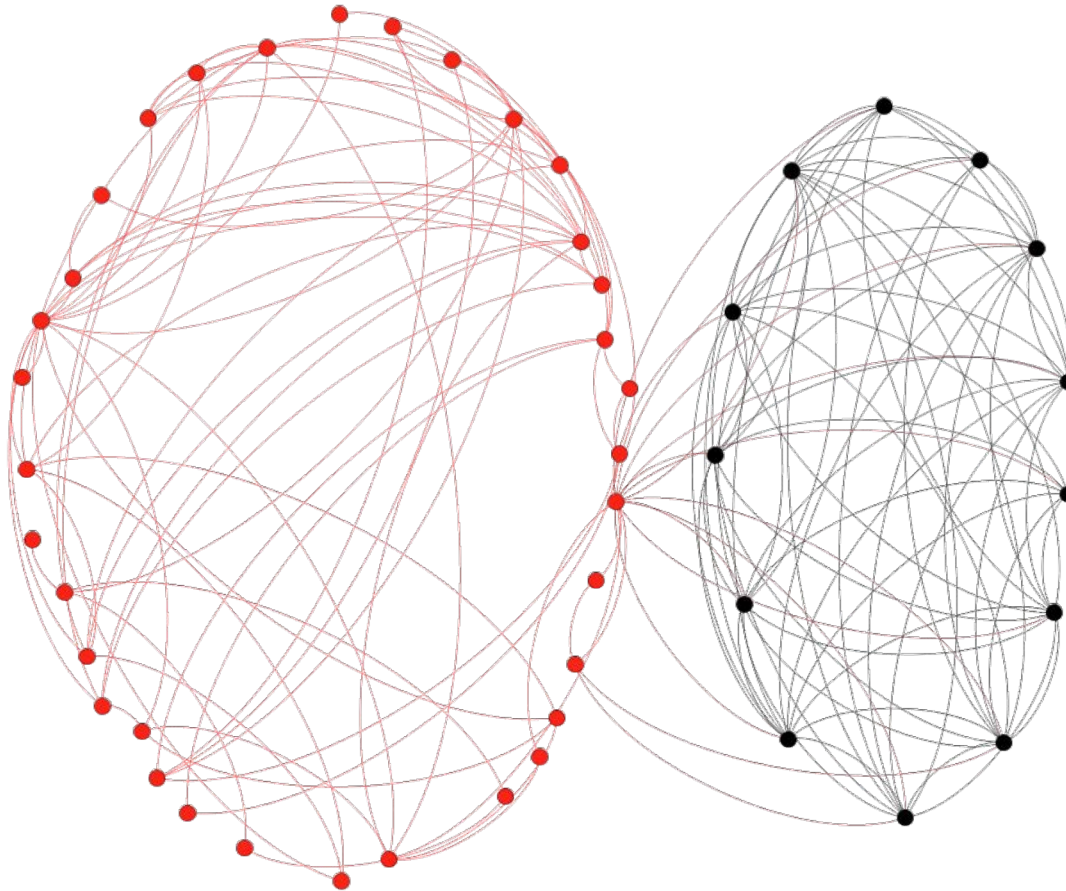
friend-ly graph

We can also create a Facebook network where two nodes (people) are connected by an edge if they are Facebook friends.
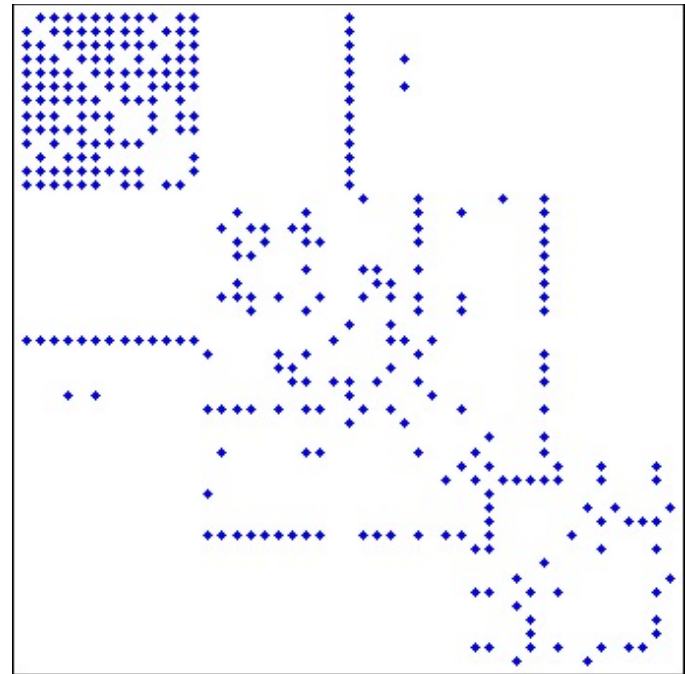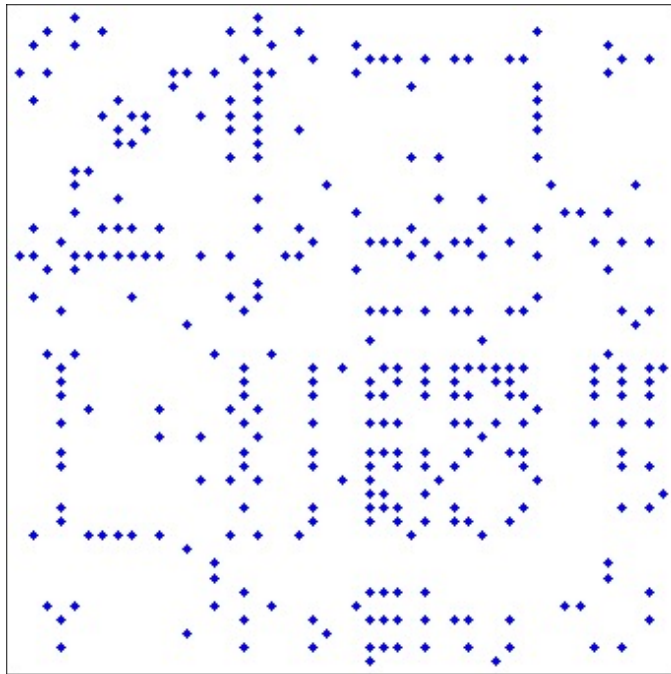
groups of friends

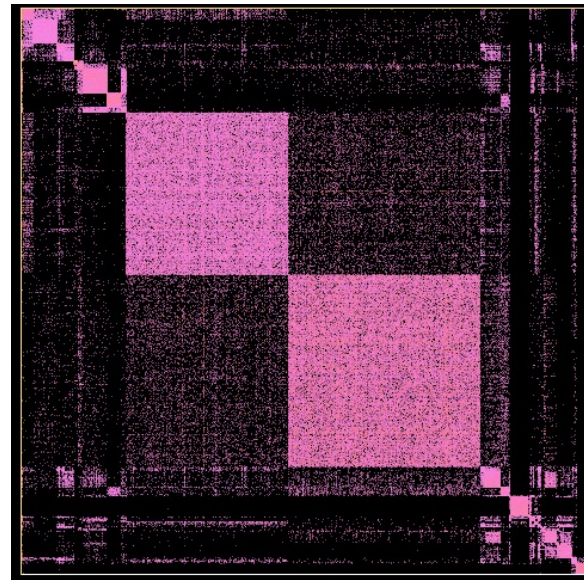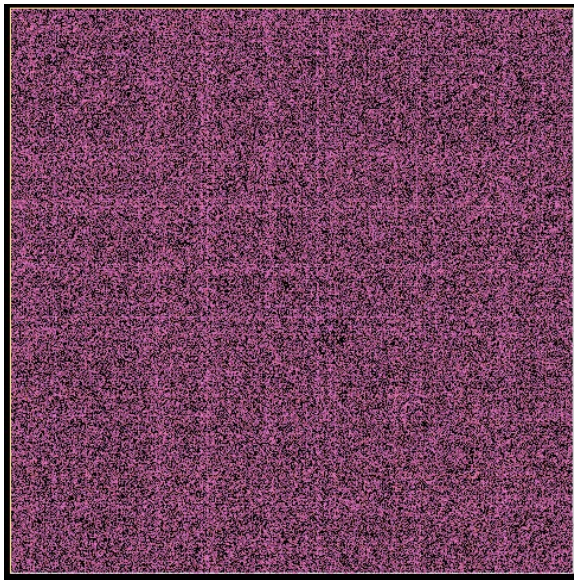If we use the Fiedler method, we find:

matrix friends

Let's create a spy plot of the Facebook adjacency matrix. Then, reorder according to clusters.

movie app

- Two movies connected if either is within the 20 most similar movies.

- Clustering finds mathematical movie genres.

Pictures thanks to David Gleich.

cluster & rank

CHOOSE YOUR SIDE

Clustering and ranking can go hand in hand.

Next steps

Let's talk about what lies ahead and what to keep in mind as you progress through research.
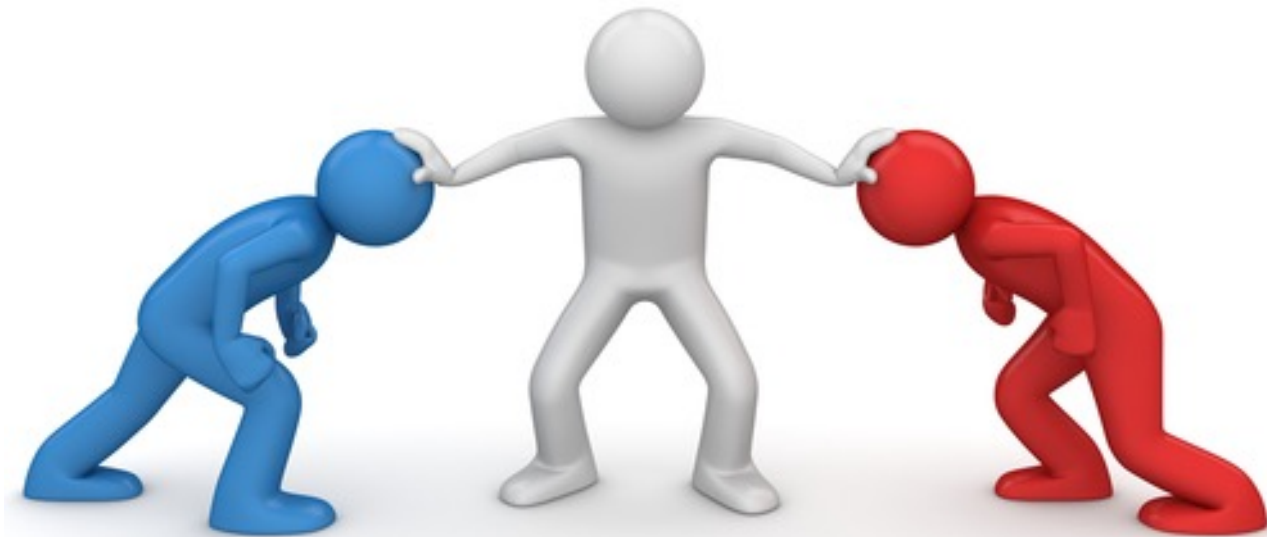
Big Goal

- Find your data!

- Got all your data or do you need to augment/change it?

- Is it ready to be ranked?

Game?

- In many ranking methods, items compete over something. I often ask, "What's the game?"

- Also, is it one-versus-one or one-versus-many?

Who's number 1?

- Which ranking method(s) will you use?

- Do you need to change the algorithm?

- Will you use more than one method and compare them?

- What do you expect find?

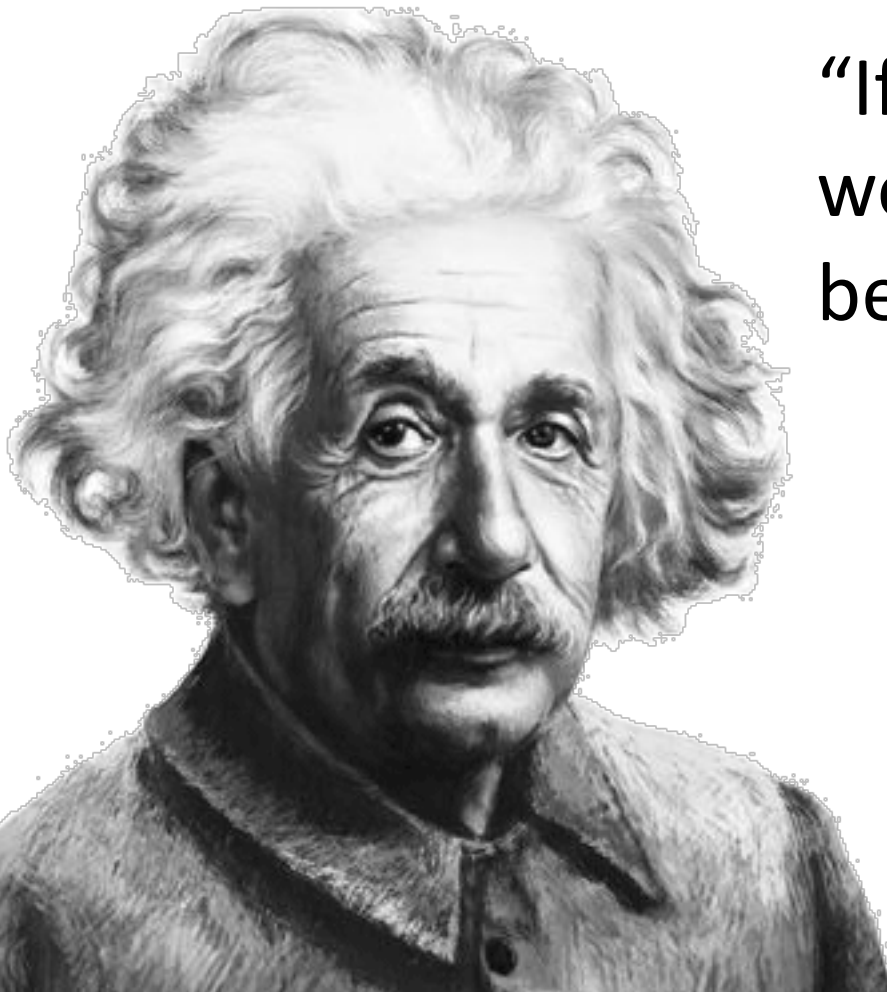- Be careful of "making" the method state what you expect.

Success?

- How will you show your work's success?

- Ranking methods will rank.

- Why should someone care?

- Why should someone believe in your work?
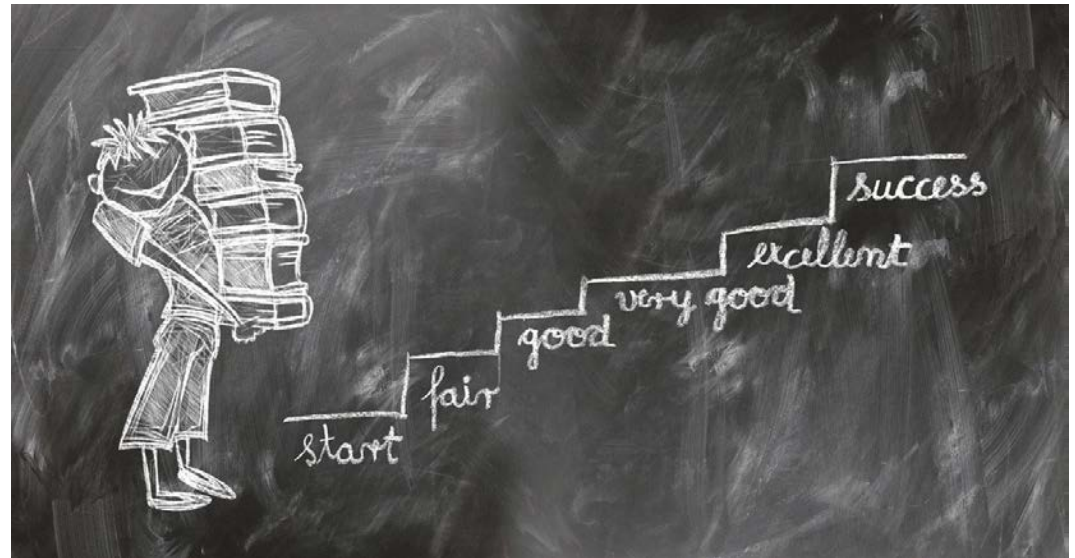
Beware
confusion coming

"If we knew what we were doing, it wouldn't be called research."

– Albert Einstein

Stuck?

Stuck? Come to our meetings or connect on Schoology ready to explain where you are stuck and any ideas you have.

Remember

- Since our one-on-one sessions are on different days, there are no longer due dates on Schoology.

- So, you will need to remind yourself when components are due. Note that the requirements are stated in the homework.

Homework

- In our first individual session, we will discuss and refine project proposal/plan.

- Start preliminary list of research articles/resources/data sources.

- Due: Research proposal and plan in which you propose a topic for research, including what application you intend to rank and which ranking methods you will use.

- This paper should be 2-3 pages long.