Thursday July 17 - Student Workshop Day 1

** Get "Charlie Ticket" passes activated for 1 week at Kenmore ($19)

7:00 am – Breakfast (700 Commonwealth Avenue, Warren Towers, BU)

9:00 am – SCI 352 – 590 Commonwealth Avenue

- Welcome & Overview

9:30 am – SCI 352

- "Informal Chat" #1 with Sandy Pentland, MIT

10:30 am – Science Center Lounge

- Get to Know You Activity

11:30 am – Lunch (700 Commonwealth Avenue, Warren Towers, BU)

1:00 pm – SCI Courtyard

- Activity: Columbian Hypnosis

1:30 pm – SCI 130

- Gephi: Visual Representation of Networks

2:30 pm – SCI 130

- Graphr: Hands-on Network Visualization Tool

3:15 pm – SCI 130

- Install NetLogo & Network Interactives

4:00 pm – SCI 352

- Wrap Up
- Introduce "Talk Show"

5:00 pm – Dinner (700 Commonwealth Avenue, Warren Towers, BU)

"The Big Picture of Network Science"

- Networks are everywhere. Network science is the study of how stuff is *connected* to other stuff.

- Network science has at its roots in interdisciplinary science. Network science has grown from crossing the long-held traditional boundaries between disciplines – *making new connections where they did not exist before.*

- This interdisciplinary nature can be massively intimidating for learners, teachers, and science researchers. Even though science is, at its heart, about making new discoveries, most human beings are comfortable with certainty – with knowing things. But making new connections, thinking outside the box, tapping into creativity, looking at problems in many different ways, being able to dive in deep or step way back, asking what may seem like outrageous questions and taking risks is where you find the thrill of "doing science".

- Data are everywhere and science is about making sense of data. "Small data" we can plot some points and fit a line and draw some conclusions. As we accumulate more and more data, we develop more techniques for making sense of data. In an information society, how scientists make sense of "big data" is critical and has given rise to innovative ways to visualize data through networks. Science and art overlap.

- Computational advances have helped create exciting opportunities in network science. And vice versa: we build bigger and faster computers and more sophisticated tools because we push at the limits of what is possible.

- Since you are spending a full school year on science research, your job is not to pre-define tasks and tackle them. Rather, the challenge is to begin to feel comfortable looking at problems with a creative mind and start by asking questions. Creativity is not just in traditional works of art or music. We can be creative in business, in life, in art, in music, and in science. This is not encouraged so much in schools and, unfortunately, in the human lifespan we are most creative as young children and tend to lose it as we become more and more enmeshed in societal norms.

- Over the next many days, our job is to give you tools. To provide you with a "backpack" you will take back home with you, encourage you to think outside the box and try new things, and help you grow confidence as young scientists. Your job is simple: to be curious.

- You will work in teams because science is rarely a solitary pursuit. It's not important that you have the best GPA in the group. Everyone can have something to contribute. We all have our special talents. Discovering what they are and using them to your group's advantage is how you can grow beyond what you can do as individuals (i.e., the whole is greater than the sum of its parts).

"Informal Chats"

We have several "informal chats" from scientists who study all things networks. We would like you to approach these chats as opportunities. As high school students, it is not all that common to hear from prominent scientists in such an intimate setting.

While some of what they talk about you may not quite understand, try to listen with an open mind.

Try not to get bogged down in what you don't understand, but by what is being said that may remind you of what you already do know.

Also, as you listen to our speakers, have your notebook out and jot down questions you have. The important thing to keep in mind is that this is not a class. You are not being tested on the content.

We are expecting that each student jot down at least a couple questions per speaker. Make sure you do, because if there are no questions asked from the audience, one of us may call on you! Our speakers will leave here with very positive opinions of all of us if we ask questions!!

Remember, these are human beings who do this research for their living. So in addition to questions you may have about the science concepts or methods or whatever, there is also a story behind the research.

So questions like the "Talk Show" would be very relevant.

Kinds of questions that may be interesting are:

- How did you get that data?
- Why did you decide to study THIS particular question?
- Who else helps you solve these problems? Do you work alone? Do you supervise others?
- Who funds this study? And how did you convince them to fund you?
- How long have you been working on this problem?

These are just samples of questions you may want to ask. You get the idea!!

Good luck!

Alex ‘Sandy’ Pentland

http://web.media.mit.edu/~sandy/

Alex `Sandy’ Pentland has helped create and direct MIT’s Media Lab, the Media Lab Asia, and the Center for Future Health. He chairs the World Economic Forum's Data Driven Development council, is Academic Director of the MIT-Harvard-ODI Big Data and People Project, and is a member of the Advisory Boards for Google, Nissan, Telefonica, Monument Capital, and the Minerva Schools.

In 2012 Forbes named Sandy one of the 'seven most powerful data scientists in the world’, along with Google founders and the CTO of the United States, and in 2013 he won the McKinsey Award from Harvard Business Review. He is among the most-cited computational scientists in the world, and a pioneer in computational social science, organizational engineering, wearable computing (Google Glass), image understanding, and modern biometrics. His research has been featured in Nature, Science, and Harvard Business Review, as well as being the focus of TV features on BBC World, Discover and Science channels. His most recent book is `Social Physics,' published by Penguin Press.

Over the years Sandy has advised more than 50 PhD students. Almost half are now tenured faculty at leading institutions, with another one-quarter leading industry research groups and a final quarter founders of their own companies. Sandy's research group and entrepreneurship program have spun off more than 30 companies to date, three of which are publicly listed and several that serve millions of poor in Africa and South Asia. Recent spin-offs have been featured in publications such as the Economist and the New York Times, as well as winning a variety of prizes from international development organizations.

Interesting experiences include winning the DARPA 40th Anniversary of the Internet Grand Challenge, dining with British Royalty and the President of India, staging fashion shows in Paris, Tokyo, and New York, and developing a method for counting beavers from space.

"Get to Know You" Activity

Part 1: Gathering Data

Each student is handed a sheet of 24 yellow stickers and writes their initials on all

Each adult is handed a sheet of 24 blue stickers and writes their initials on all

Everyone is handed a white paper that they put their initials on top.

For 20 minutes, if you talk some with another person and get to know them a bit, then you each place one of your stickers on the other person's white paper. Do not place a sticker on the other person if you just know him or her. You have to talk for a bit in order to exchange stickers.

After, everyone will have a number of stickers on their page.

Part 2: Making Sense of Data

Count up how many total dots are on your paper. Write it as "Number of Connections".

Group Activity 1: Plot the histogram of number of connections using sticky notes.

Question: How would you describe this histogram and does it make sense?

Group Activity 2: Using the blackboard, how would you go about making a visual representation of the social connections in the data you collected? Make sure every "node" of the network (person) is identified.

Group Activity 3 (Optional): What if I asked each of you to count the number of triangles you have or connected triplets. So, if I am connected to X and X is connected to Y and Y is connected back to me, that's a triangle. On your paper write the number of triangles you have. Can you tell that from only your paper? This is a type of clustering and helps us to understand real-world networks. In a social network for example, nodes create tight groups with a large number of connections between them.

Group Activity 4 (Optional): You see that some of you have blue and some of you have yellow stickers on your paper. Can you learn anything about the picture on the blackboard if we make use of that information? How can you do that? First, label all the blues and yellow nodes. Does that help learn anything? Second, label all the blue-blue connections and all the yellow-yellow connections. Does that help? If not, then redraw the network with blue nodes on the left side of board, and yellow nodes on right side. Does this show clustering? Weak links?

Part 3: Visualizing Data

The instructor will create a Google Drive spreadsheet that can be edited by anyone with a link. A shortened URL will be created using e.g. tinyurl.com and then shared with all students.

Each student will enter the result of the activity into the shared spreadsheet (in a single row) in the following format:

Initials-of-student Initials-on-sticker-1 Initials-on-sticker-2 …

For example:

HS PT SU CC LS …

where "HS" is the student's own initials, and the rest are the initials they have on their white paper.

This data entry will take some time (20~30 minutes). Students may need to take turns as *Google Drive* may not allow all the students to have access to the spreadsheet simultaneously.

Once the data entry is over, the instructor will guide the students to install *Gephi* to their computers. The installer is available from their website and from the thumb drives distributed in the classroom.

In the meantime, the instructor will download the spreadsheet in a text format and correct any problems in it due to mistakes in data entry.

The whole class will play with *Gephi* following the tutorial slides. Then the data will be imported into *Gephi* and visualized. Several network analysis techniques are tested, e.g., degree, centrality measures, community detection, followed by discussions and Q&A's.

"Columbian Hypnosis" Activity

Count participants off 1 and 2. Have all the ones pair up with a two.

Start by having the pairs face each other about arms length apart.

The "ones" will hypnotize the "twos." To do so they will put their arm straight out, 6" from the face of the person in front. Fingers pointed up.

When the leader says start, the number "twos" will be hypnotized and follow whatever the hand in front of them does. There face my always remain six inches from the hand.

If hand twists to the right, they must twist their head with it.

If the hand moved towards their face they must back up to keep the six inch space.

If the hand lowers they must lower.

After about 30 seconds, announce that they are switching roles and all the number ones will be hypnotized.

Note: Not mentioning they are switching roles at the beginning usually leads to a good conversation about what the number ones though about when they were told about their initial roles and how they felt when roles were reversed.

After both sides have time to play the role of hypnotist and hypnotized, debrief.

Part 2 (Optional)

Ask for a volunteer. They will be the center.

Ask for two more volunteers. They will both be hypnotized by the first volunteer.

Ask for four more volunteers. Two will be assigned to each person who volunteered the second time.

After everyone has joined you will have a huge tree of people being hypnotized and hypnotizing.

Once the leader says start everyone will have to follow the same directions to the original activity. It may get out of hand, which is ok, because you can ask about the different roles (center vs outside).

Debrief about what they noticed how it felt. Talk about how they influenced each other.

Graphr: Visualizing Congressional Collaborations

Visit the interactive website at: **http://polymer.bu.edu/graphr**

Graphr is a simple software tool we wrote to explore collaborations within the US Congress. Google "congressional statements" to see many examples of statements. As you can see, statements tend to come from groups, indicating collaboration. The more "signers" to the statement indicates a level of support and therefore seeks to influence opinions on legislation, appointments, and all government affairs. Of course, there is a long history here as there were 56 signers of the Declaration of Independence.

Suggestions for exploration:

1. Select 2011 colored by party. How many clusters do you see? What are the primary attributes of each cluster? *A cluster can be described as a collection of highly interconnected nodes.*

2. Select 2010 colored by party. Notice the small cluster of republicans. Try dragging them around to uncover their structure (click a node after dragging it to unpin it). Who is the senator represented by the most connected node in this cluster (hover over the node to see)? Look him up on Wikipedia. Is there anything special you can identify about him?

3. Select 2008. Notice the pairs of isolated connected nodes. Hover over them. What property do these pairs share that is causing them to be connected?

4. Select 2011 colored by party. Are there any nodes that stand out as having unusual connections? Look them up on Wikipedia. Can you explain what's happening?

5. Select 2009 colored by gender. Does this suggest that women are more likely to work with each other in congress? Try hovering over each woman and/or coloring them by different properties. What can you learn from this?

Network Interactives - http://www.bu.edu/networks/activities/

Network Optimizer

Using the parameters in the software, you can design your own network which will have certain characteristics. For example, very vulnerable to attacks, not very vulnerable, etc. It's up to you! Some of the parameters you may not know what they mean, and that's okay. In science, computers and elsewhere we often are in a situation where we don't know everything. So, don't stress too much about it. You can explore, and maybe you will start to figure it out.

If you want to save a network here and import it into Network Attack, simply click on Export and save the network as *somefilename*.txt

The .txt extension is needed!

Then you can load it to Network Attack by clicking File > Load

Network Attack

An example on how to use:

File > Load > network.txt (or the others)

Maybe the best way to start is to look only at one network on one side. So on right side select Network > None

Maybe the best way to start is to look only at one network on one side. So on right side select Network > None

Attack Size: 1
Attack Type: Random

You will see a large random network. In this simulation, when you click on "Step", the computer selects a node at random, kills it, and all its links (or connections).

Now try another attack type: "Targeted"

In this simulation, the computer selects the most connected node, kills it, and all its links (connections). At each step, it selects the next most connected node.

Now put the network on both the left and right sides.

Select Random for the left. Select Targeted for the right. Select "Step".

Question: Which attack mode is more efficient in destroying the network?

NetLogo 5.1RC1 Install - http://ccl.northwestern.edu/devel/5.1-RC1/

Once you to the install, open the "models" folder and create an empty folder titled "User Community Models".

Copy the .nlogo file in extensions > nw > models into the "User Community Models" folder

Copy the 3 .nlogo files in extensions > nw > demo into the "User Community Models" folder

Explore a selection of these models. When you load a new model, you will see some buttons at the top: Interface, Info, Code. "Info" will give you some tips on what's going on in that particular model.

In particular, explore:

- Files > Model Library > User Community Models

There you will find the network models you copied.

Other models we thought may be interesting are here:

- Sample Models > Art > Follower
- Sample Models > Biology > AIDS
- Sample Models > Biology > Virus
- Sample Models > Computer Science > PageRank
- Sample Models > Earth Science > Fire
- Sample Models > Social Science > (all of them)
- Sample Models > Networks > (all of them)
- Curricular Models > epiDEM > (all of them)

In addition, in the "Model Library" you will see a link to "Go to User Community Models web page".

This web page has models others have uploaded. We will come back to this later in our workshop. In the meantime, take a few minutes and download any one that is interesting.

The procedure is to download it, and move it into the models > User Community Models folder as you did earlier.

"Talk Show"

You will be hosting 9 "talk shows" throughout our workshop, usually in the mornings and right after lunch. The goal is to find different ways for you to learn about scientists. There will be 2 hosts per show and each show will last about 15 minutes and interview one scientist.

While you may do any kind of format you want, and ask any question you want (within reason of course), we have some starter questions that may help you formulate your own talk show.

Suggested questions:

- When you were growing up, what did you think of science?
- And what did you think of scientists – did you have an idea of what scientists did, or certain kinds of scientists?
- What first got you interested in science – was there one experience that got you interested?
- When you were young did you know anyone who was a scientist?
- Did you have science teachers who influenced you? If so, how?
- When did you first realize that YOU could be a scientist?
- How did you choose your field? Did you know right away what you wanted to do or did you try a few different things first?
- What do you love about working in science?
- Is there anything you DON'T like about being a scientist?
- Why is being a scientist important?
- When did you first hear about network science?
- How did you decide to concentrate on network science?