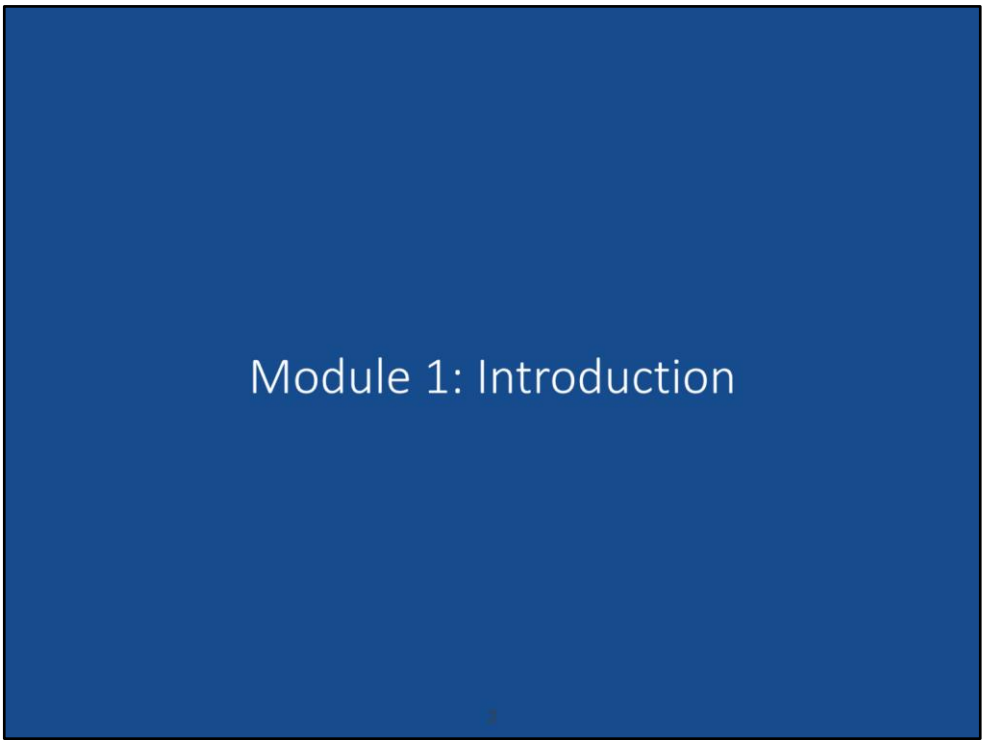
Yale University

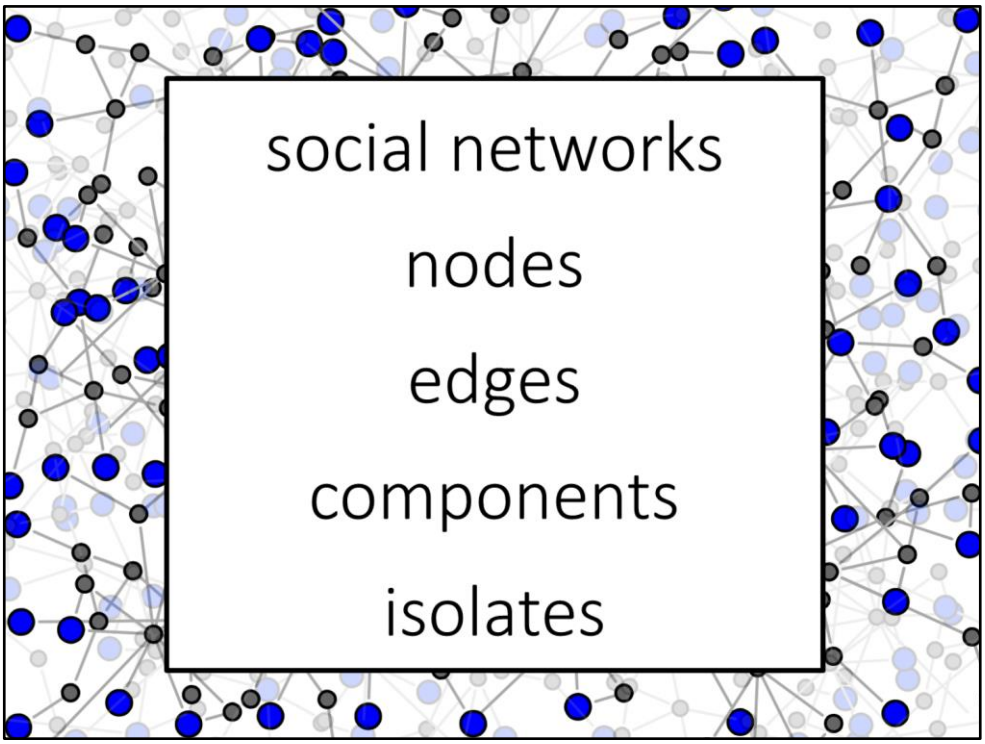
Social Network Analysis for Criminal Justice Practitioners and Analysts

Module 1: Introduction

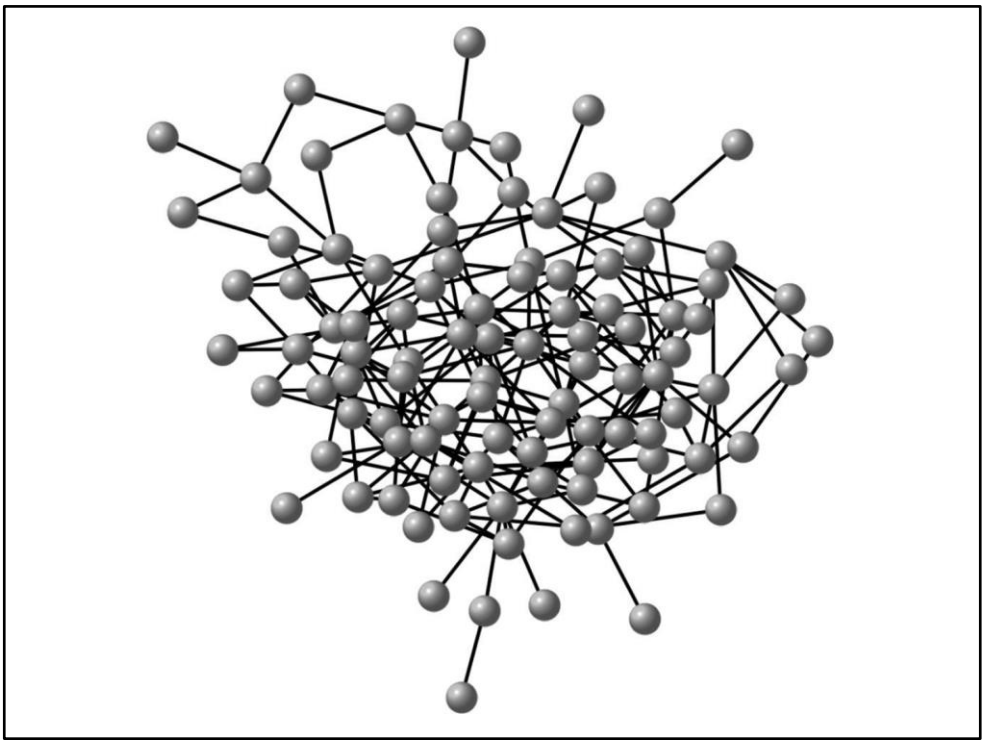
Andrew V. Papachristos



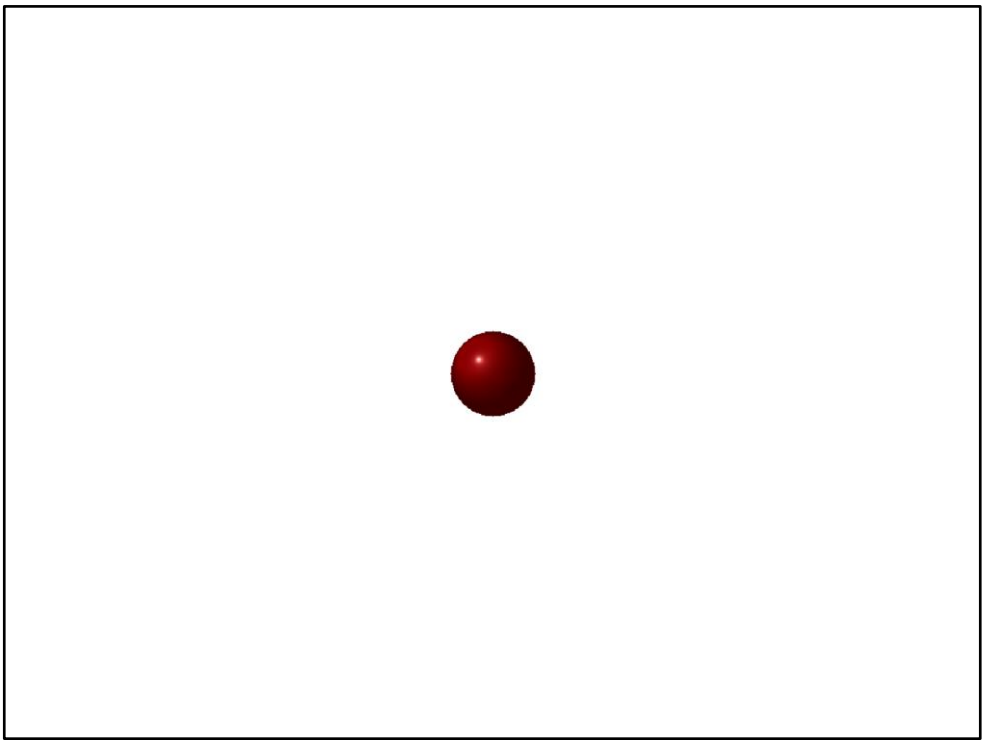
Welcome to Module 1: Introduction. The goal of this module is to familiarize participants with the logic and potential of social network analysis. Social network analysis often goes by the acronym SNA. Module 1: Introduction includes definitions of social network analysis, basic social network concepts, examples of social network analysis research, and criminal justice and law enforcement applications of social network analysis.



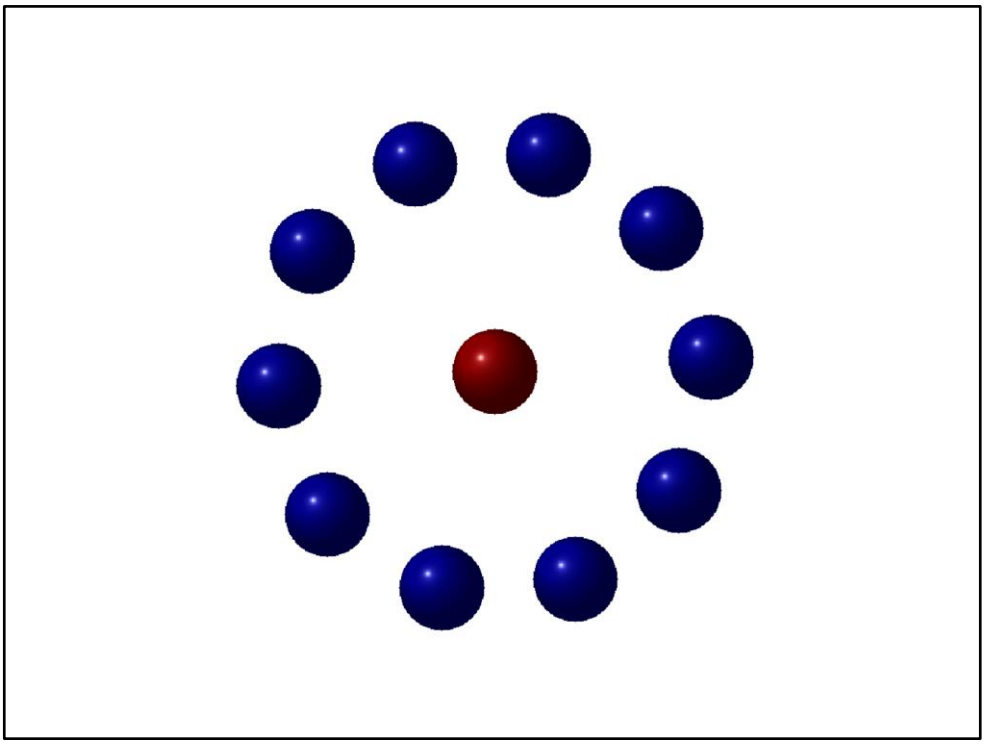
This module defines the following concepts: social networks, nodes, edges, components, and isolates.



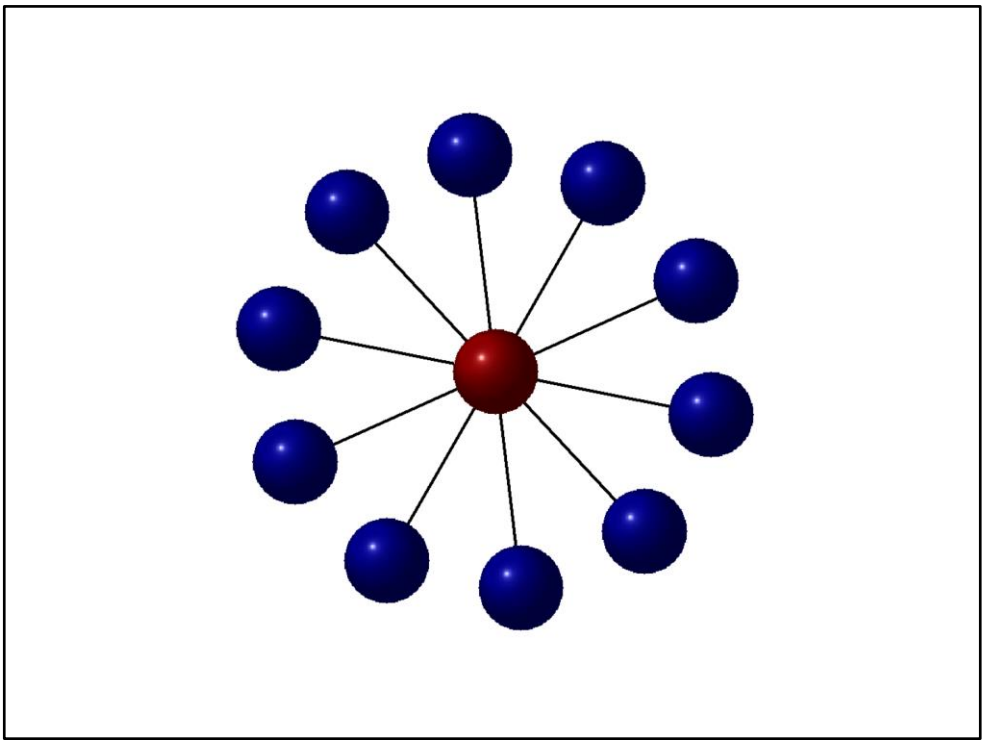
Before we begin, it is important to recognize that social networks are different than networking. In a business class you learn the value of networking. We go to conferences to network. We build online professional profiles to network. Networking is what gets us jobs. That idea of agentic, active-verb networking is not the same as social networks. Social networks are the large systems of social connections and social relationships in which individuals are embedded. Social relationships can be friendships, work relationships, neighbor relationships, associations with classmates—or any of the other various ways in which we are connected with others.



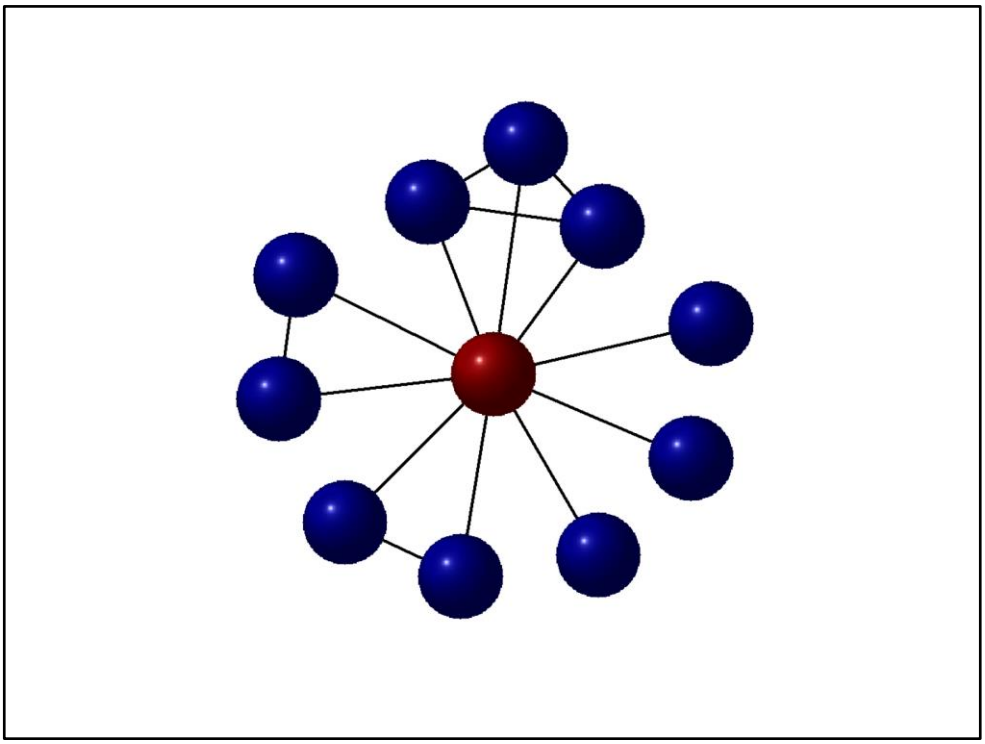
Here’s a thought experiment: Imagine yourself.



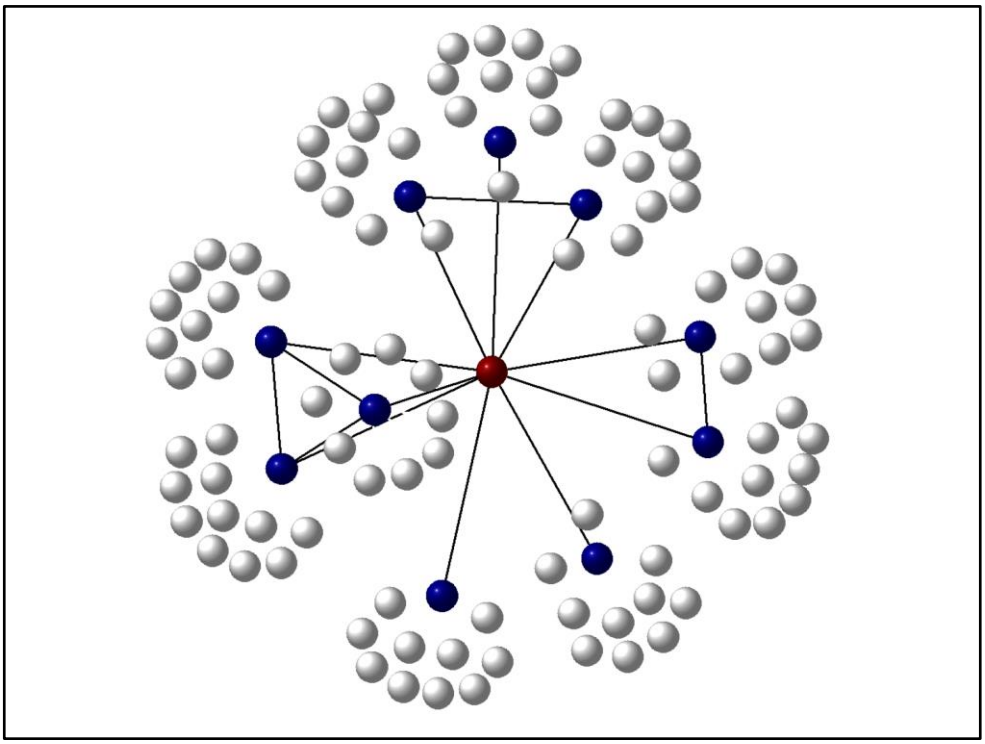
Imagine all of the people whom you know.



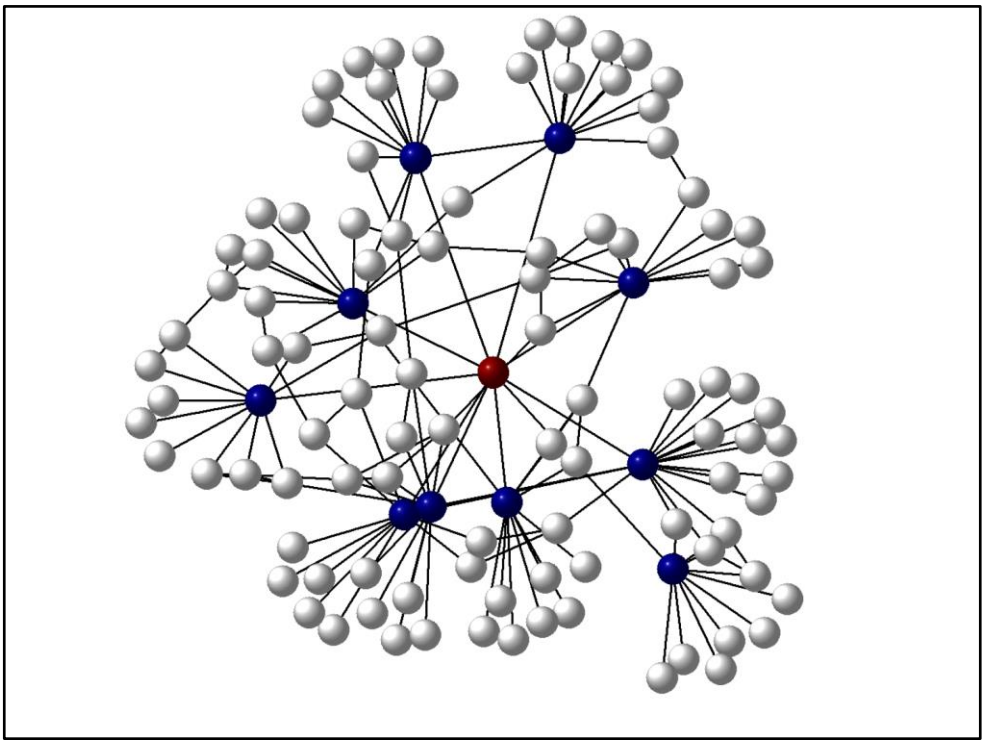
Imagine all of the relationships you have with those people.



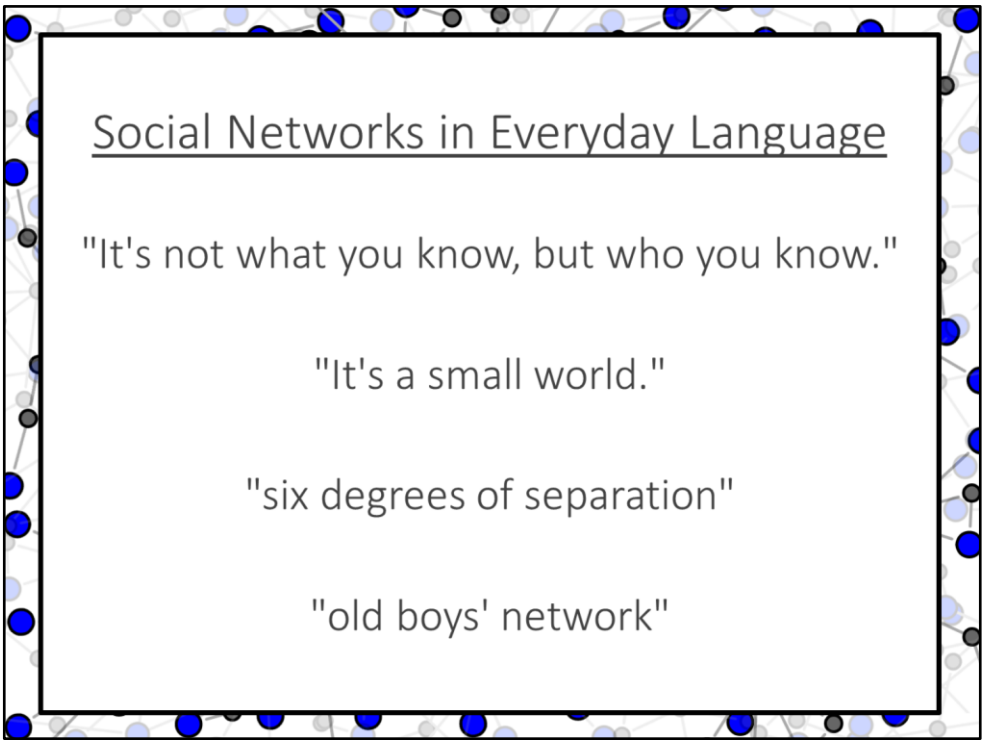
Imagine all of the relationships that exist between the people you know.



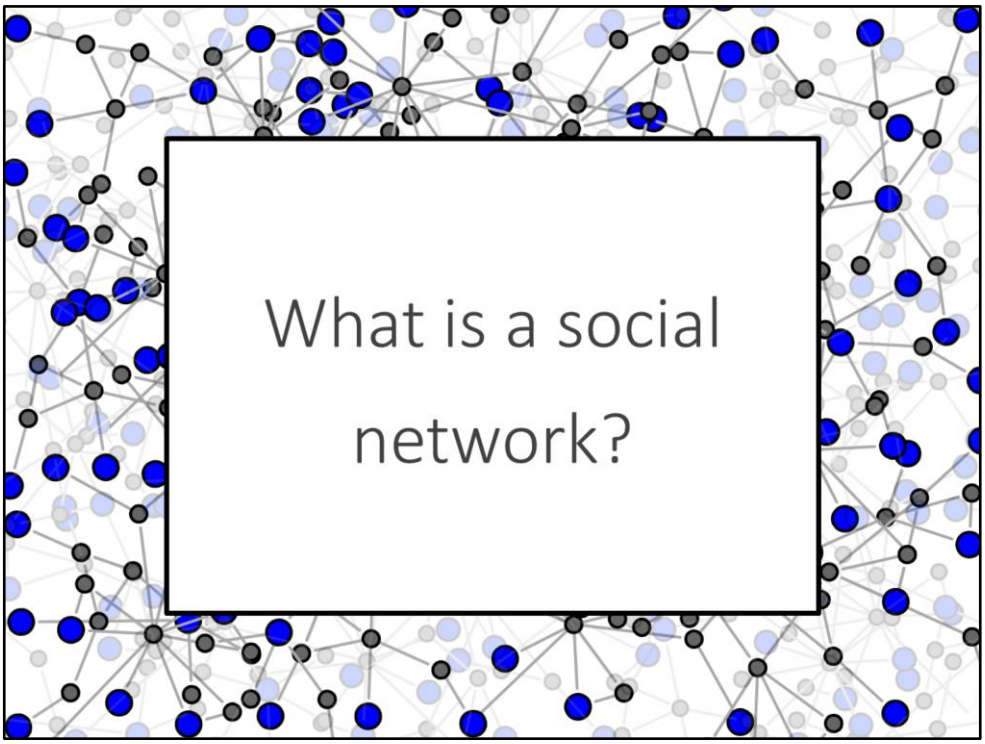
Now imagine all of the people your people know.



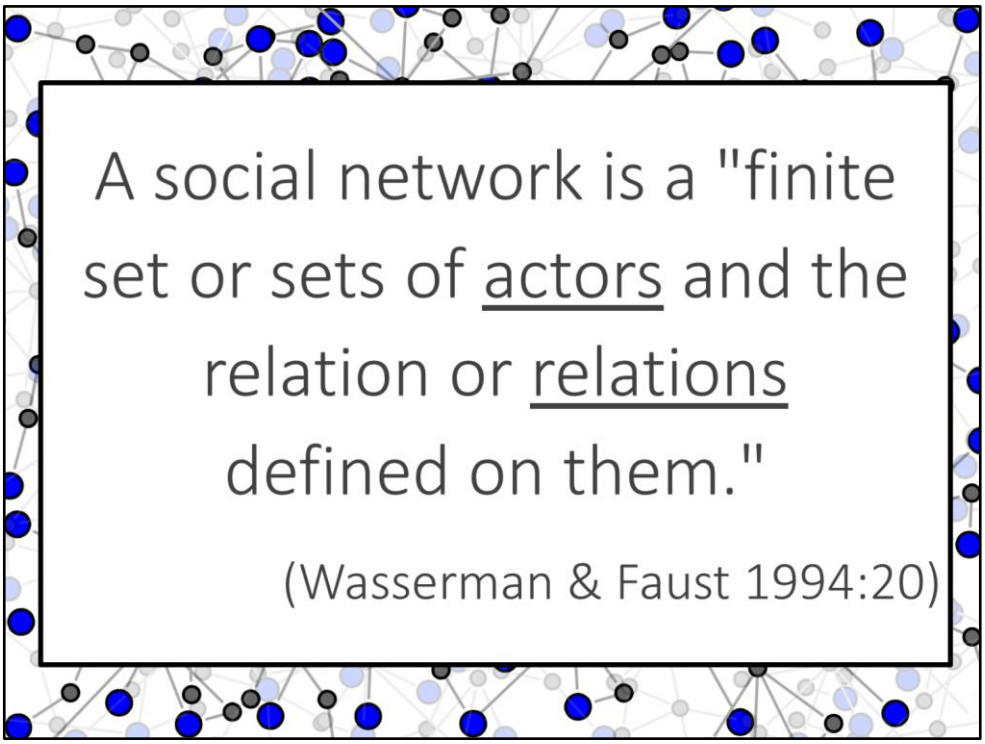
Imagine all of the relationships your people have with their people and their people have with their people. This is likely a dizzying thought experiment, but it is exactly the map of these layers and layers of connections that is your social network. Our social networks are so ubiquitous that we seldom think about them. We take them for granted. But there is an entire science dedicated to the study of social networks. Social network analysis actually measures and maps out the social connections and relationships between individuals.



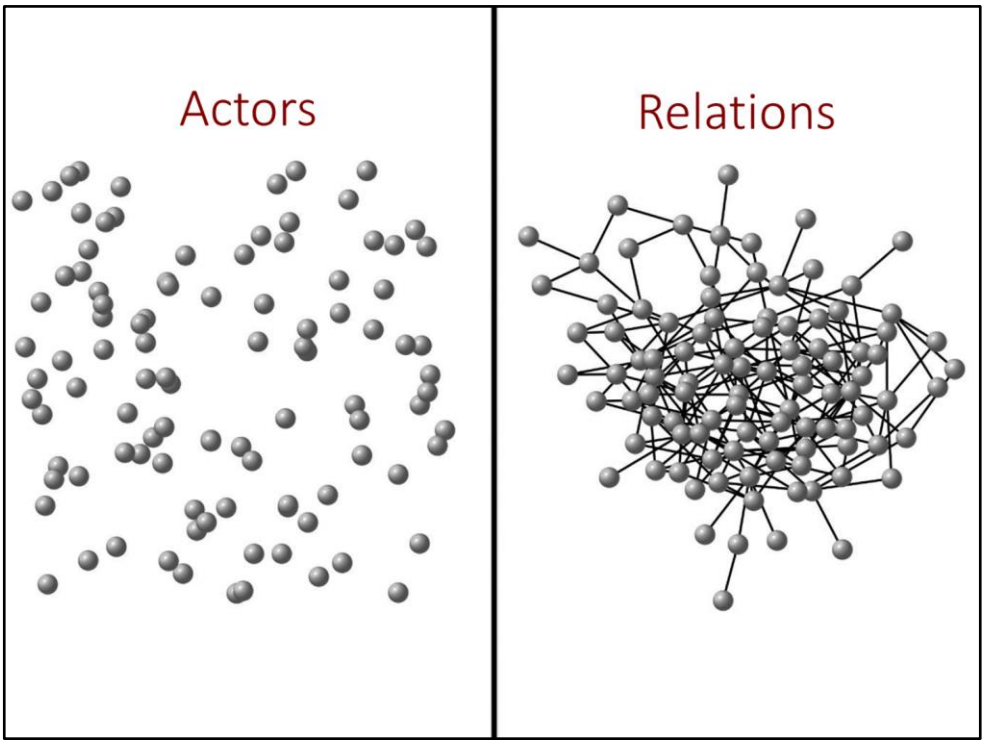
Some social network concepts are prevalent in day-to-day life. Are you familiar with these phrases referring to social networks? “It’s not what you know, but who you know” refers to the idea that skills and talents are often secondary to the social connections one has with others in the business or work world. “It’s a small world” refers to those moments when individuals realize their overlapping connections to others—a friend of a friend is a friend. “Six degrees of separation” refers to the theory that everyone can be connected to everyone else in six steps or less. Six degrees of separation is a synonym for the small world phenomenon. The “old boys’ network” can refer to actual groups of male pupils or a metaphor for cronyism, but the central idea is that personal social ties are the key to individual success. Central to all of these phrases is the idea that social networks are a meaningful and influential part of the social world. Social network analysis begins with this premise by combining information on relationships with powerful analytical tools in order to test empirically the various ways in which social networks form, grow, or influence.



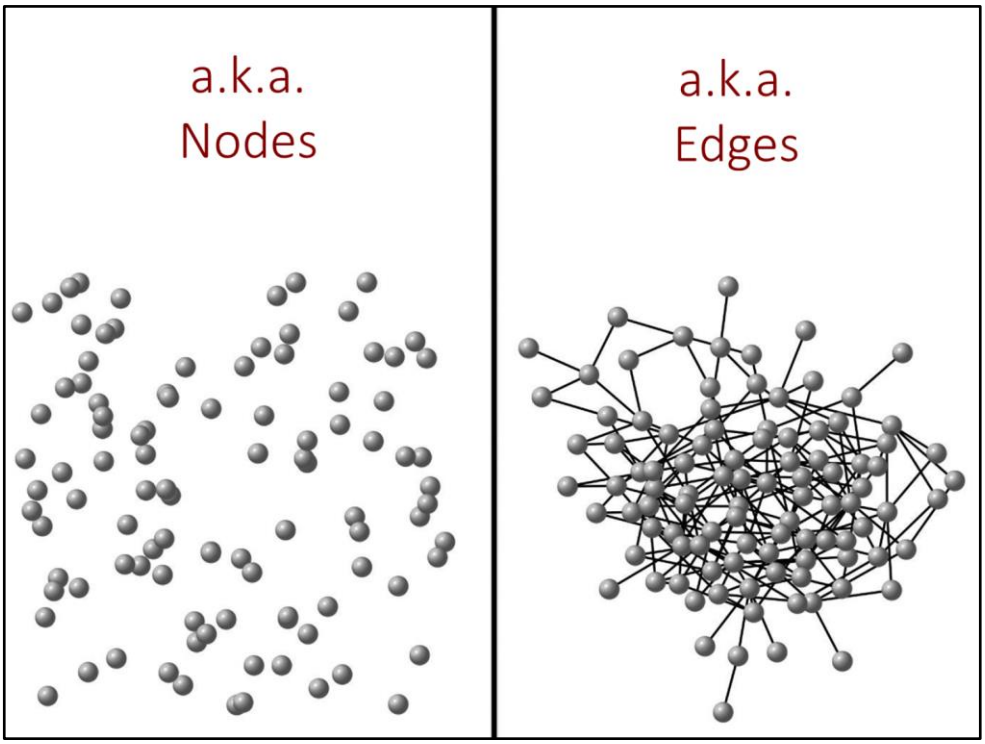
So what is a social network? This will be a central question to this introduction module.



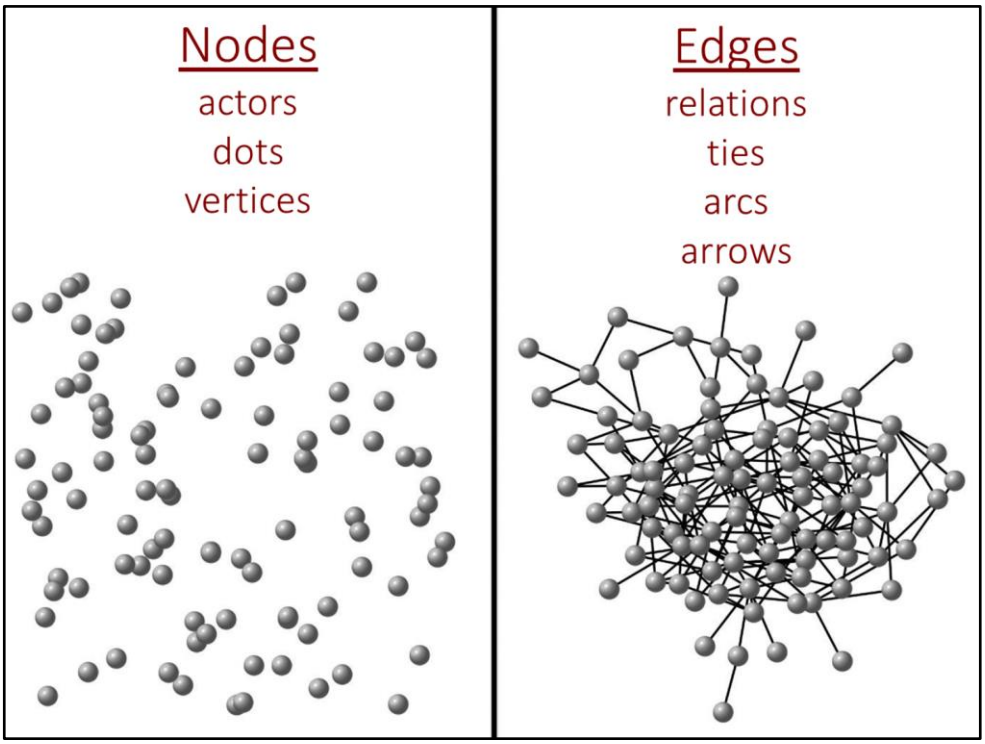
A classic definition of social networks comes from Stanley Wasserman and Katherine Faust's 1994 book titled, *Social Network Analysis: Methods and Applications*. In their book, Wasserman and Faust define social networks as "a finite set or sets of actors and the relation or relations defined on them" (Wasserman & Faust 1994:20). Notice that this definition proposes two essential features of social networks: the actors and their relations.



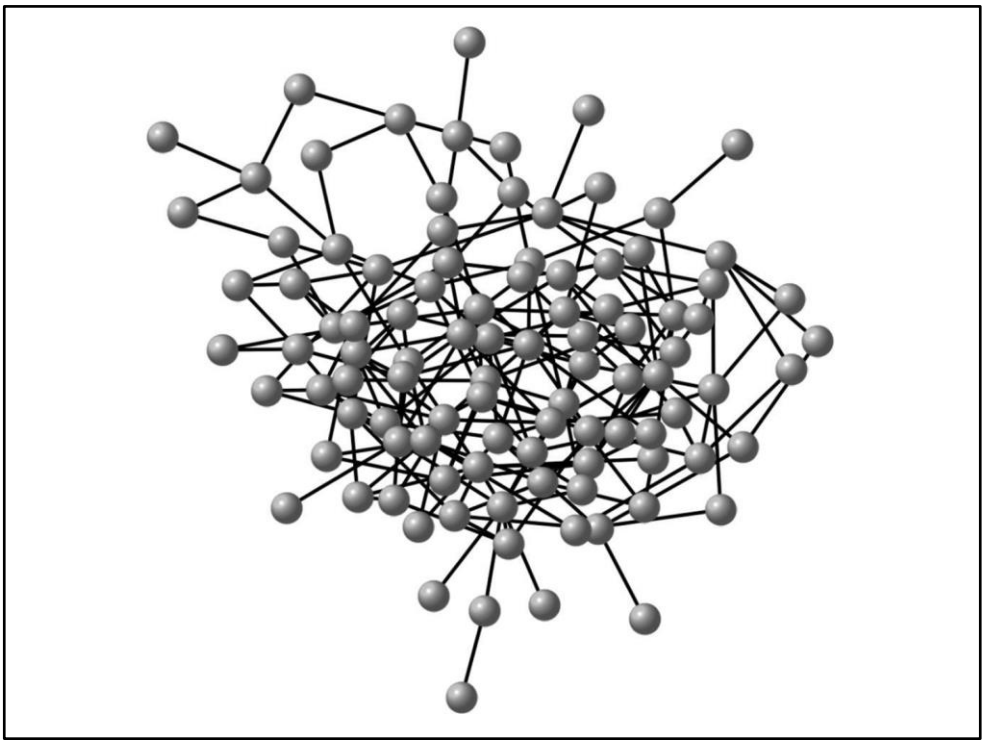
We could study just actors and their actions or beliefs—much research does this. But when we study only actors then we are ignoring the larger social networks surrounding those individuals. We forget that individual actions and beliefs are influenced by systems of relationships. We do not want to forget this larger system of influence when we think about individuals. A social network requires both the actors and their relationships. It is a *both/and* definition: *both* actors *and* relations. The blob of connections on the right can tell us more than just the actors on the left. Without the connections on the right, we are just random dots on a page.



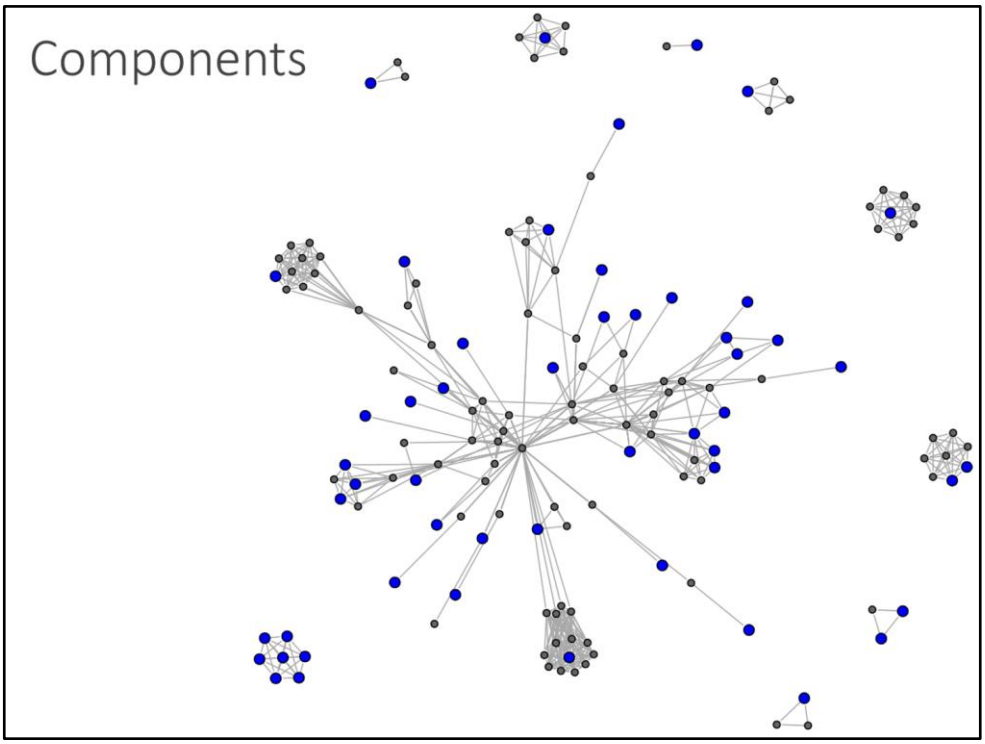
Here are two new SNA concepts: nodes and edges. Nodes and edges are two fundamental terms in social network analysis. They are basic social network analysis jargon. Nodes are the dots or the points in a network. Nodes represent people or actors. Edges are the lines in a network. The edges represent relationships and connections.



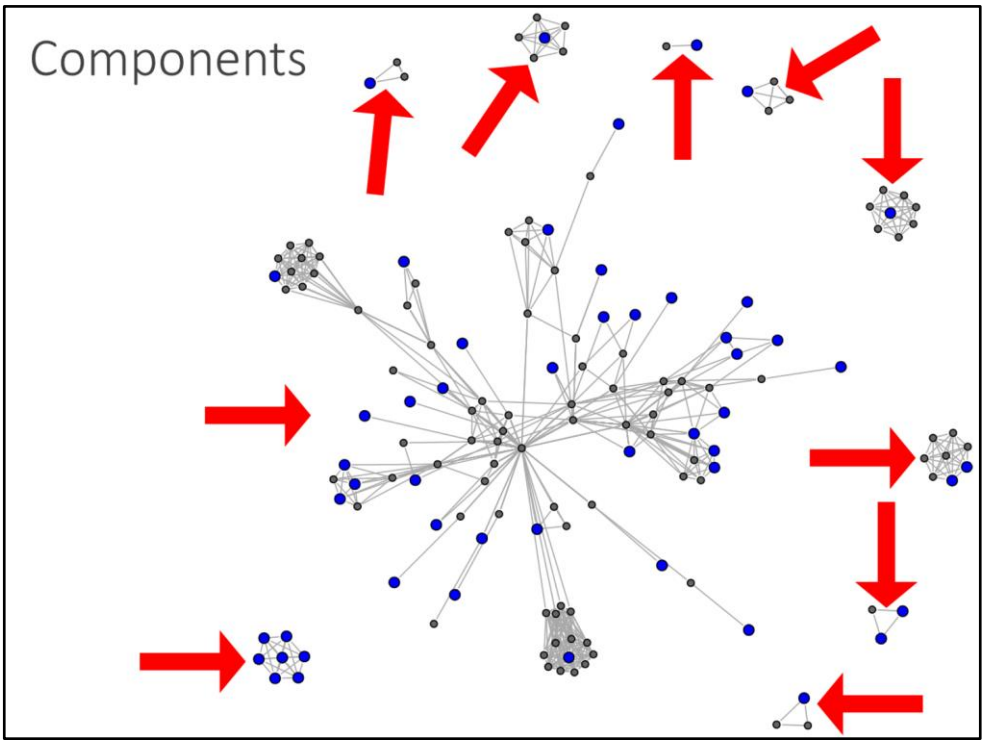
Some other synonyms for nodes include actors, dots, and vertices. Some synonyms for edges include relations, ties, arcs, and arrows.



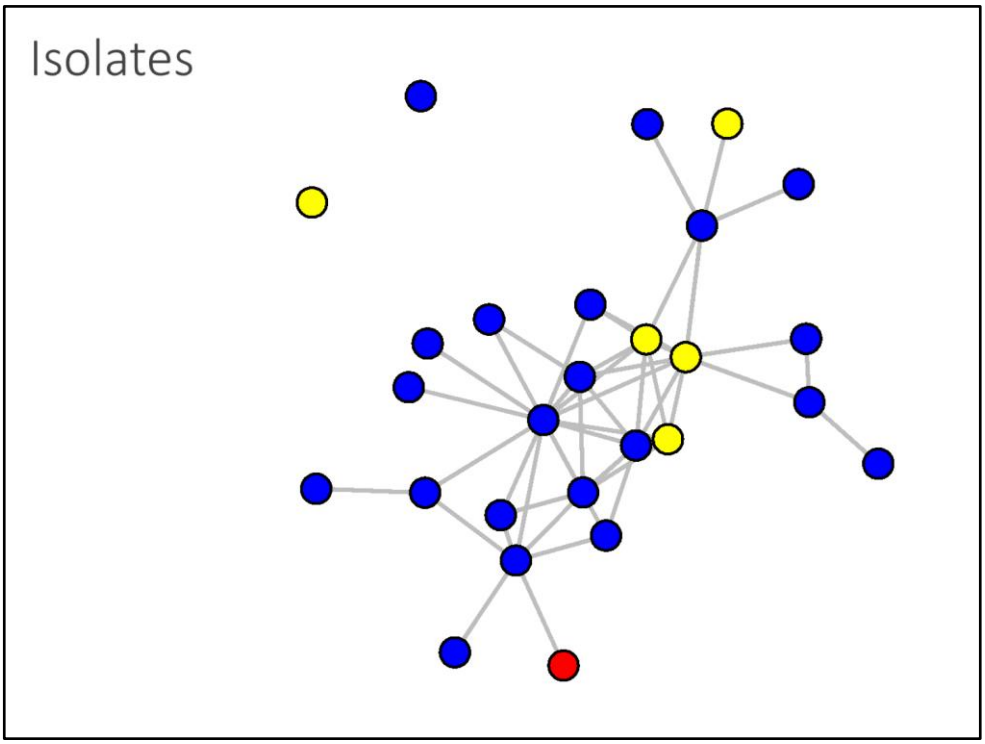
What could this social network possibly represent? What could this be an abstraction of? Perhaps this is a network of Facebook friendships, co-arrests, family ties or rivalries between gang members, criminal associations between Prohibition era bootleggers, friendships among residents in a college dorm, politicians and their committee memberships or co-sponsorship of bills, people who loan money to each other, or people who give advice to each other. Nodes don't only represent individuals, they can represent neighborhoods, schools, words, groups, or nations. So perhaps this is a network of countries and their trading agreements or alliances, organizations and political affiliations, drug trafficking between groups, or the pecking order of chickens. Actually this is just a fake network made using a random number generator. But let's say this is a friendship network in a particular neighborhood. Each node is a resident of the neighborhood and each edge is a mutual friendship. The two people connected by each edge are friends. So given that, what do you see in this network? If you were asked to describe this friendship network, what would you say? One thing you could say about this friendship network, is that all of the residents are connected in a single component. This brings us to another SNA concept: components. Components are the different parts of a network that are separate from other connected parts of a network. In this hypothetical neighborhood friendship network, all of the residents are connected to each other through direct and indirect friendships in one single component. No one is alone or completely outside of the component, and there are no completely separate groups of friends making up different components. In social networks, however, this is not always the case.



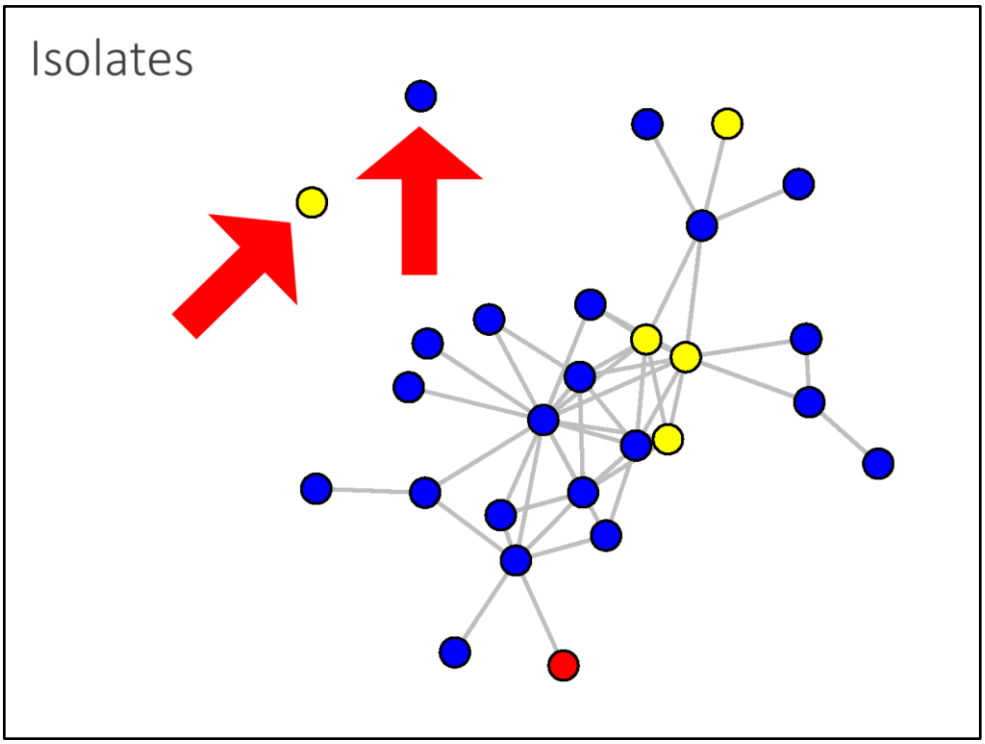
How many networks do you see in this image? This is actually a tricky question. There is one more important piece to the Wasserman and Faust definition of social networks needed to answer this question: “a finite set or sets…” (p. 20). What this means is that a social network has a clearly defined boundary. As analysts, we know who will be in the network and who will be out of the network. We might use a roster, a category, an event, etc. to define this boundary, but whatever the boundary it needs to be finite or have limits. In this network the analyst identified these individuals as the finite set of actors then mapped out the relationships between them. Even though not all of the sections are connected, they still make up a single network. The different sections of the network that do not connect are called components. Notice in this image that there is one large component in the center of the network and several smaller components around the periphery. All of these components make up a one social network.



There are actually 10 components in this network, or 10 separate pieces of this network, as indicated by the 10 red arrows. Components can be useful for identifying subgroups within a larger network. Perhaps you had a social network of all of the gangs in Boston. Different gangs might be connected within the same component, but likely there might be distinct components for the different gangs. Perhaps the social ties represent co-working relationships. We might expect a lot of separate components of co-workers unless individuals frequently change jobs and form new co-working relationships with new people thus bringing together different components over time. If there is a new job in one component, it is unlikely that people in other components will hear about that job. Or if a disease enters one component it will likely only infect people within that component and people in the other components will be safe.



The final SNA concept for this module is also related to components. The last social network term is isolates. Isolates are nodes with zero ties or zero edges. Isolates are individuals who do not have social connections to the rest of the group. Isolates are their own component — isolates are their own sections of a network that are separate from other connected parts of the network. Examples of isolates could be a loner in a high school or a hermit who lives in isolation. But it is quite difficult to think of examples of individuals who are completely isolated from all possible social ties. What is more common, however, is that individuals might not have a certain type of social tie or connection. Imagine a negative or rare social tie like a murder or an arrest. It is likely that you do not belong in a network of murders or arrests. If we mapped out a negative or rare social tie in your community it is likely that you and many people would be isolates because you do not share this negative or rare social tie with each other. You are not murdering each other or being arrested together.



In this example network, there are two isolates. The social tie connecting the largest component did not include the two individuals indicated by the red arrows. There are three components in this network: the largest component plus the two isolates. This is actually a network of criminals and their criminal relationships from 1900 to 1919. In this example, the two isolates represent solo-offenders in crime, and the other criminals are connected to each other through their criminal activities.

Uses of SNA

Spread of disease, fashion, information, technology, and capital

Political and social influence

Attitudes, opinions, and behaviors

Friendships, marriages, and relationships

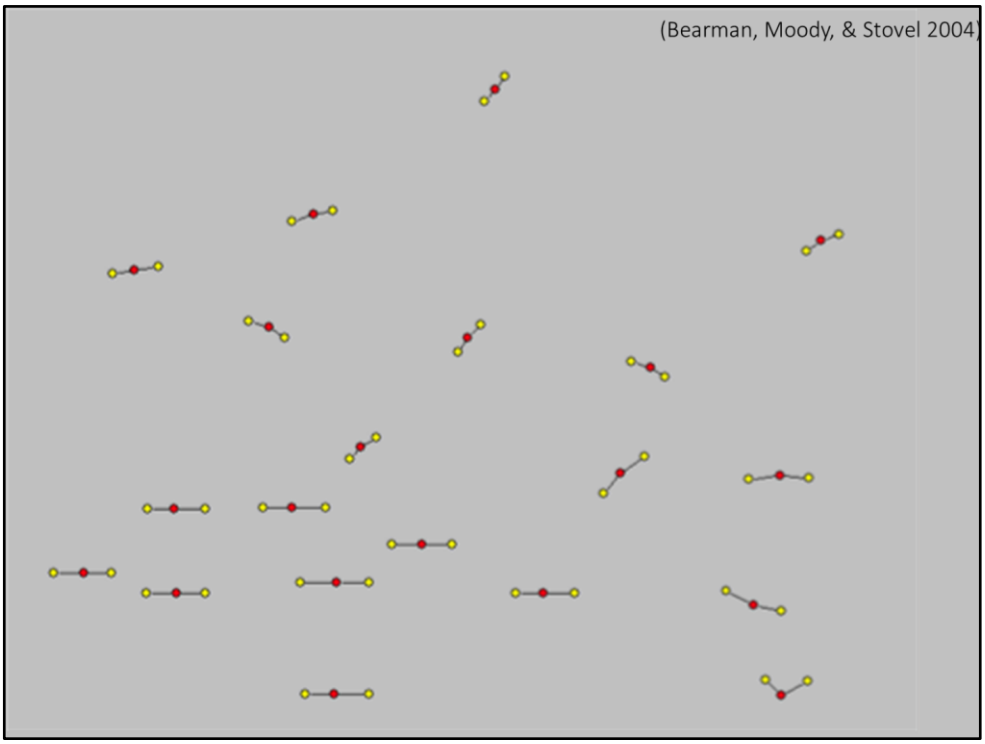
Human societies and animal societies

Social network analysis has been used to research a range of topics. The list here includes examples of a variety of research. Click on the hyperlinks in the slide above to see some example publications. A few of the full links have been included here in the notes: spread of disease (http://journals.lww.com/aidsonline/Abstract/1997/05000/Concurrent_partnerships_and_the_spread_of_HIV.12.aspx), political influence (http://prq.sagepub.com/content/56/4/449.short), opinions (http://www.jstor.org/stable/10.1086/518527), behaviors (http://connectedthebook.com), friendships (http://www.jstor.org/discover/10.1086/653658?uid=3739832&uid=2&uid=4&uid=3739256&sid=21103982393241), marriages (http://books.google.com/books/about/Family_and_Social_Network.html?id=XVwMqHb56TsC), human societies (http://www.nature.com/nature/journal/v481/n7382/full/nature10736.html%3FWT.ec_id%3DNATURE-20120126), animal societies (http://press.princeton.edu/titles/8841.html). Let's look at one of these applied SNA research examples in more depth.

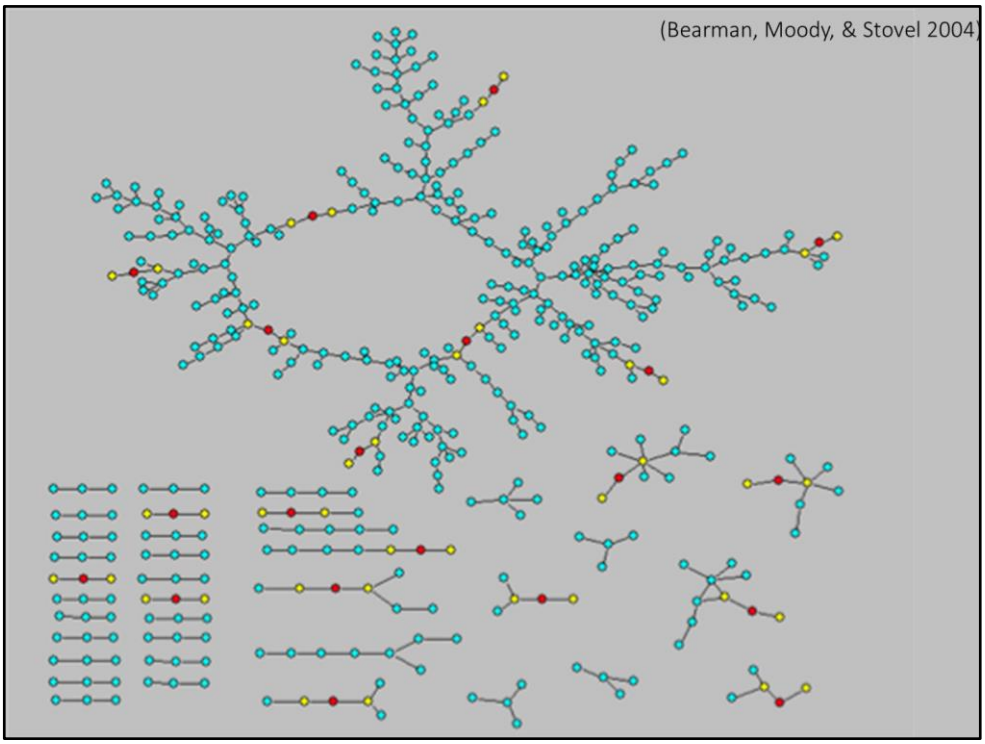
What predicts the risk of getting a sexually transmitted disease?

23

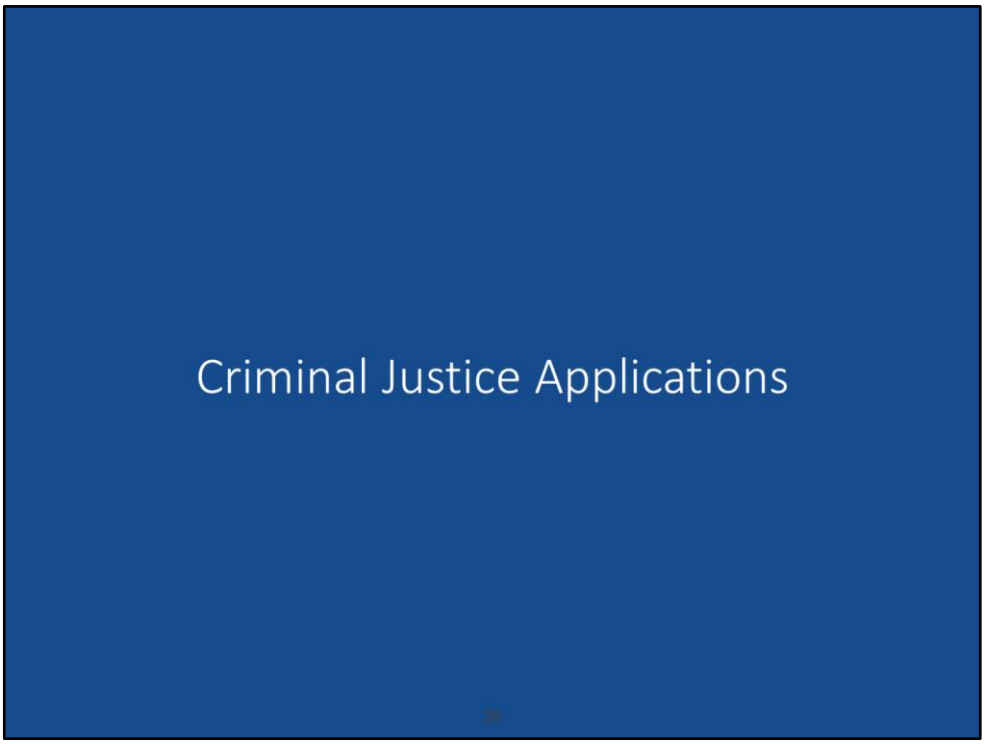
A non-social network way of thinking about what predicts the risk of getting a sexually transmitted disease would likely focus on an individual's behaviors and activities such as unprotected sex or needle sharing.



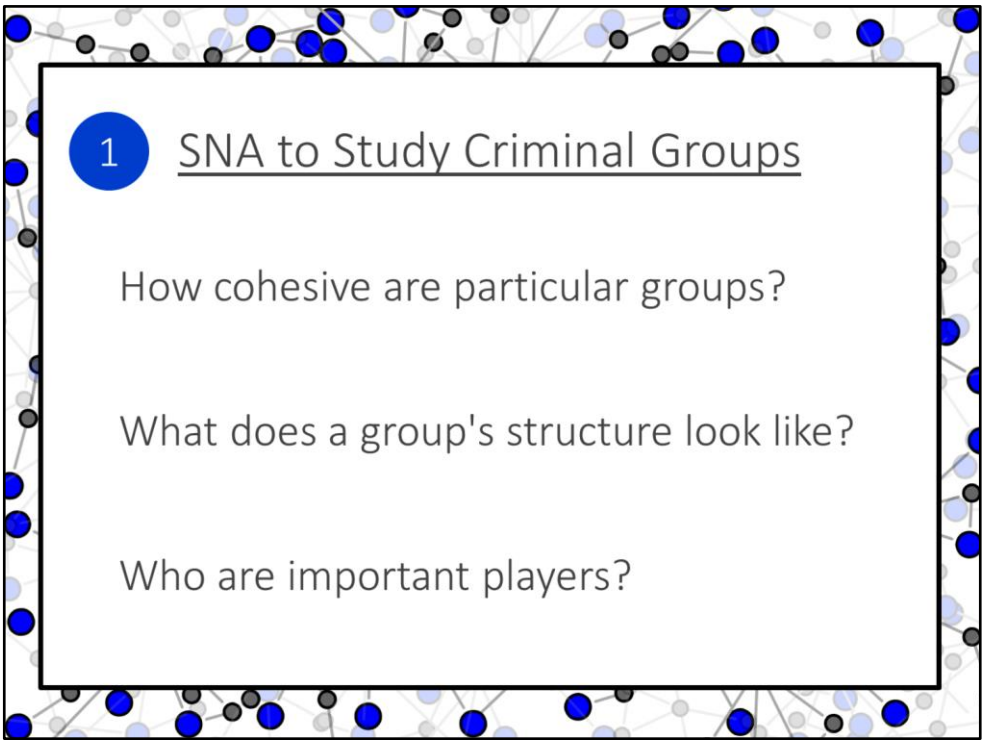
Take individual's behaviors and activities a step further. Unprotected sex and needle sharing require other individuals. That is how disease spreads. The image above considers that these risky behaviors include multiple people. If the red nodes are infected with HIV, then the yellow nodes are at risk of contracting HIV through particular activities with infected nodes. This image focuses on the infected individuals and whom those individuals might put at risk. We can take the logic of social networks a step further to ask: Are these individuals connected in other ways? Whom are the at risk yellow nodes connected to?



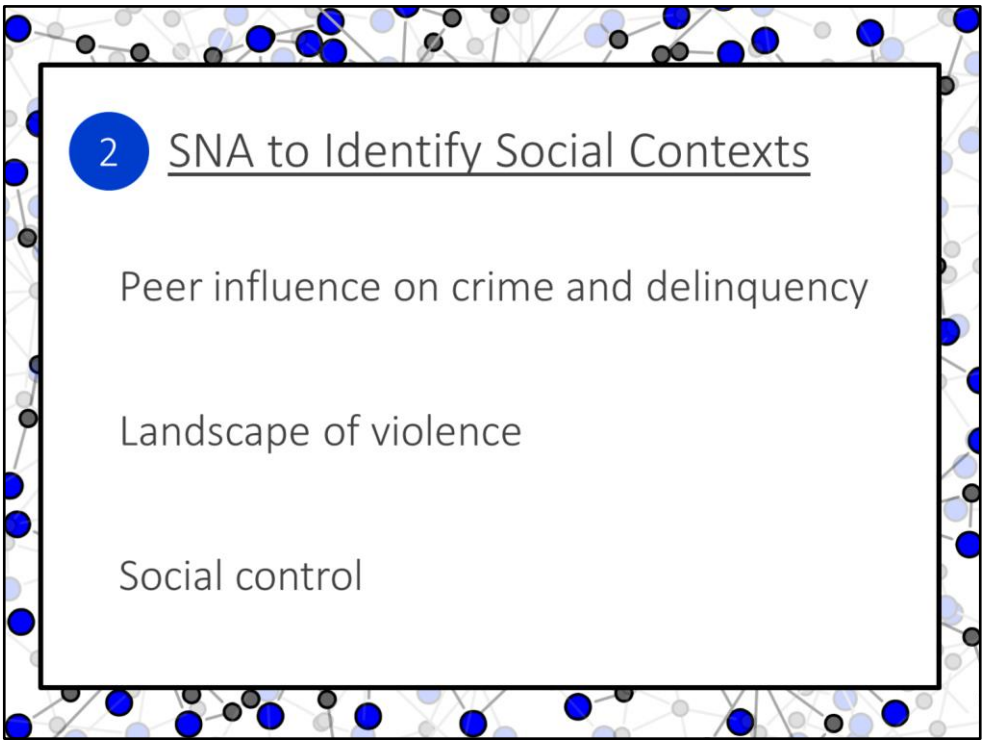
This network image is from Peter Bearman, James Moody, and Katherine Stovel's 2004 article, "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks." This is a bird-eye view of a sex network of a high school in the Midwest during the 1990s. Bearman, Moody, and Stovel's research found that 52 percent of students who were sexually active in this school were located in a single largest component. The two most distant individuals in the largest component were 37 steps apart. The majority of sexually active students in this high school had only one or two partners. Building off of the ideas from the previous slide, the red infected nodes and the yellow at risk nodes are part of various components in this sex network through which disease could spread. In the instances of the smaller components, the disease would remain rather contained. In the instance of the largest component, the disease could put much of the school at risk. If we focus only on the infected persons and their direct at risk partners, then we miss the larger network. We also miss locations in the network that are not at risk of contracting the disease such as the smaller components with no infected or at risk nodes. Consider the implications of this research compared to the implications of the previous network image in terms of public health and policies.



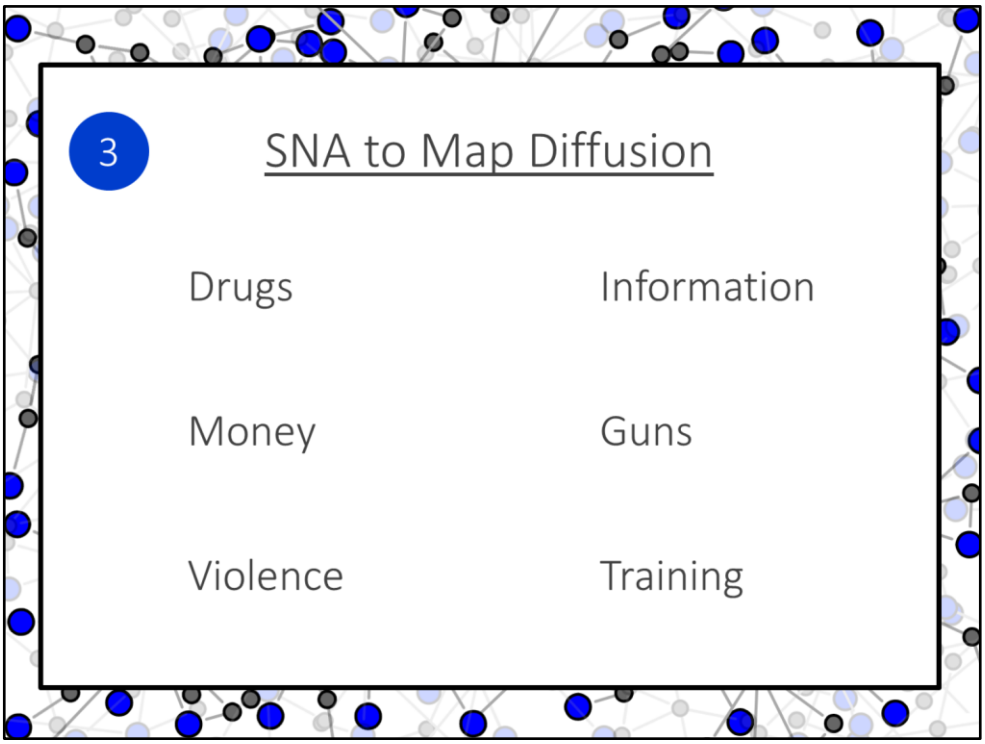
At this point in the introduction module, can you think of some questions that could possibly be answered using social network analysis? What questions are common to criminal justice that might have relational answers or explanations?



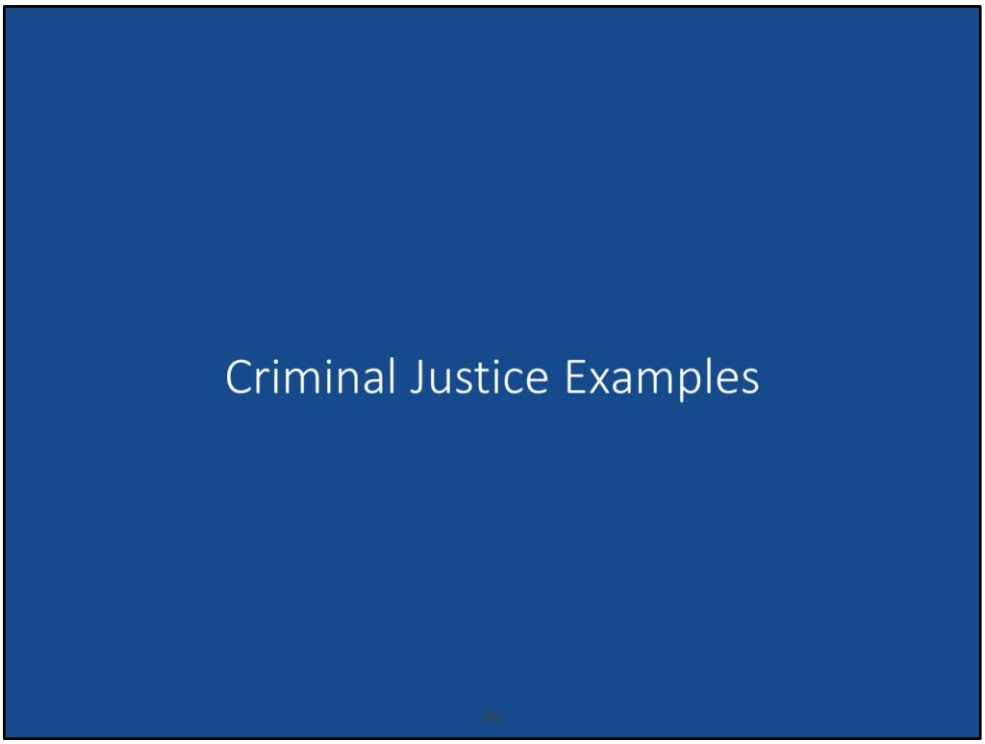
One set of possible questions could be about criminal groups or criminal networks. Mapping out the network of an entire criminal group could tell us how cohesive a criminal group is. We could learn about the structure of the criminal group. Is the group structure hierarchical or flat, dense or sparse, connected or disconnected? Or we could use the network to identify the important, central, or powerful players in the criminal group.



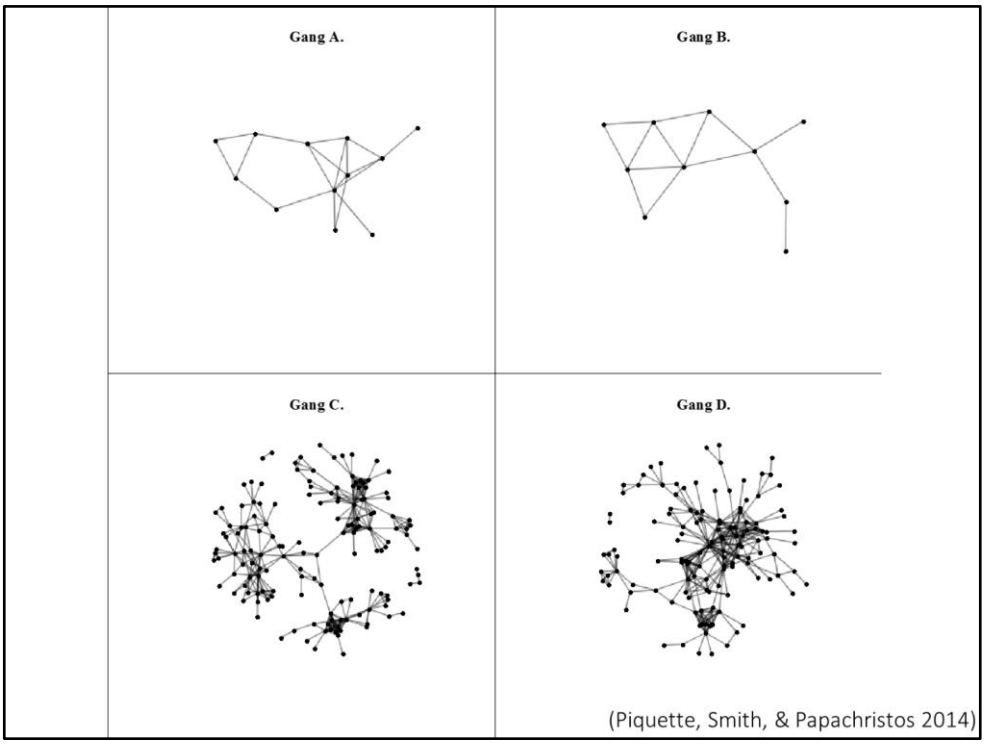
A second approach to social network analysis in criminal justice applications could be to map out broader social relationships to identify the contexts in which crime and deviance occur. For example, social network analysis has been used to look at the influence of peer networks on criminal activities and delinquent behaviors among adolescents, such as peer influence on smoking and petty theft. Crime and delinquency occur in a larger context of social networks. Similarly, shootings or other violent events can be identified in larger contexts such as arrest networks or affiliation networks. Not every social tie in the network is a violent tie, but a larger context of affiliation can be used to analyze landscapes of violence. Mapping out the rivalries and alliances between criminal groups, such as gangs, can help pinpoint where violent exchanges might be more likely to occur. In this example, the nodes of the network would be the different gangs rather than individuals. Social networks can be used to examine social control and instances with less social control. The networks could identify contexts and conditions for positive relationships versus negative relationships.



A third approach considers relationships as a conduit for transferring material items or information. The social network could map out the path of exchanges of drugs, money, or guns. The network could map out violent events, training, or the spread of information. A social network could map out the path of a single weapon exchanged over time among associates.

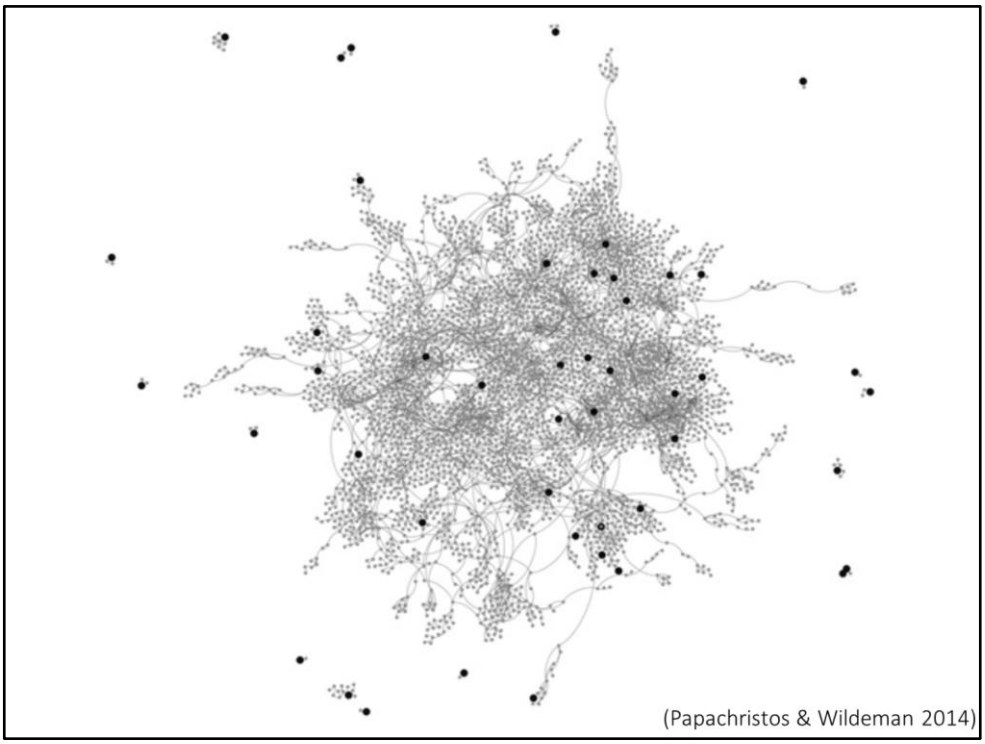


These applications provided some potential questions that utilize the logic of social network analysis. They might seem a bit abstract at this point, so now let's look at some concrete examples of SNA research in the criminal justice context.



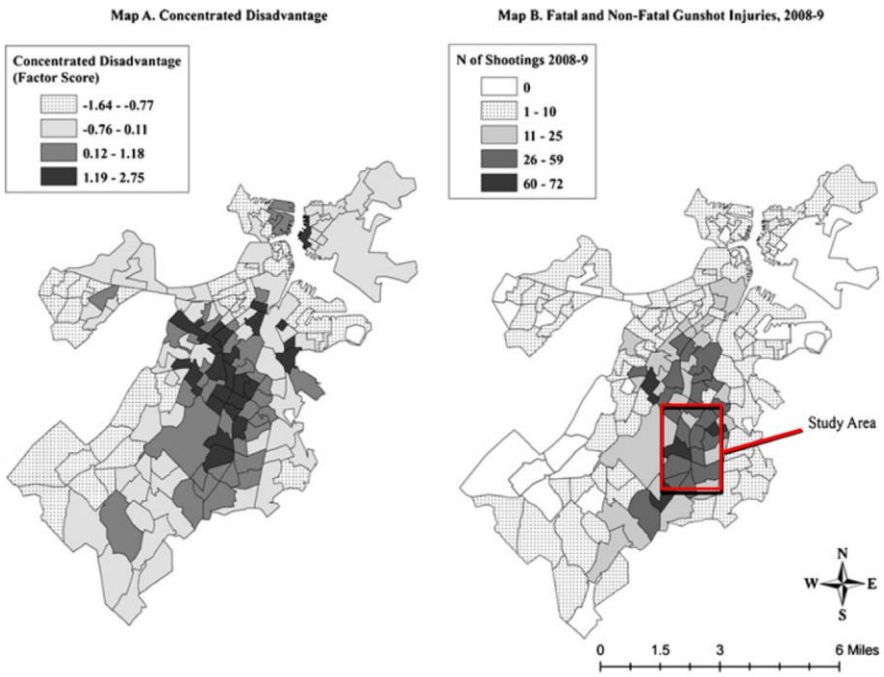
(Piquette, Smith, & Papachristos 2014)

These plots illustrate the networks of four gangs in Chicago using data from the Chicago Police Department. Each node represents a person and each edge represents a non-arrest observation made by the police – sometimes called “field observations” or “field contacts.” By aggregating police field observations over time and across multiple locations, networks emerge representing the patterns of direct and indirect associations among gang members and their non-gang companions. All four of these gangs are from the West Side neighborhood of Chicago and are part of larger gang nations that date back to the 1960s. By virtually any survey or qualitative account, all four of these gangs should have the following: a hierarchical structure, formalized sets of rules and regulations, a division of labor, and intricate drug-dealing portfolios. Yet, the patterns of association shown in this figure reveal that some of the networks deviate from such accounts. Gangs A and B, for example, are smaller in size and less dense than gangs C and D. Moreover, the networks of gangs A and B look more like ordinary friendship networks and are not hierarchical. Gang B contains overlapping triangles of individuals, and this pattern suggests that the network contains small cliques as opposed to some formalized division of labor. In contrast, gangs C and D have a starburst shape with dense pockets of individuals linked together by one or two ties. These gangs are more centralized, a trait that is especially evident in gang C with a single individual joining three different parts of the network. Comparisons across gang networks could reveal how gang structures have changed over time or identify which members are best located to send a message to the other members.

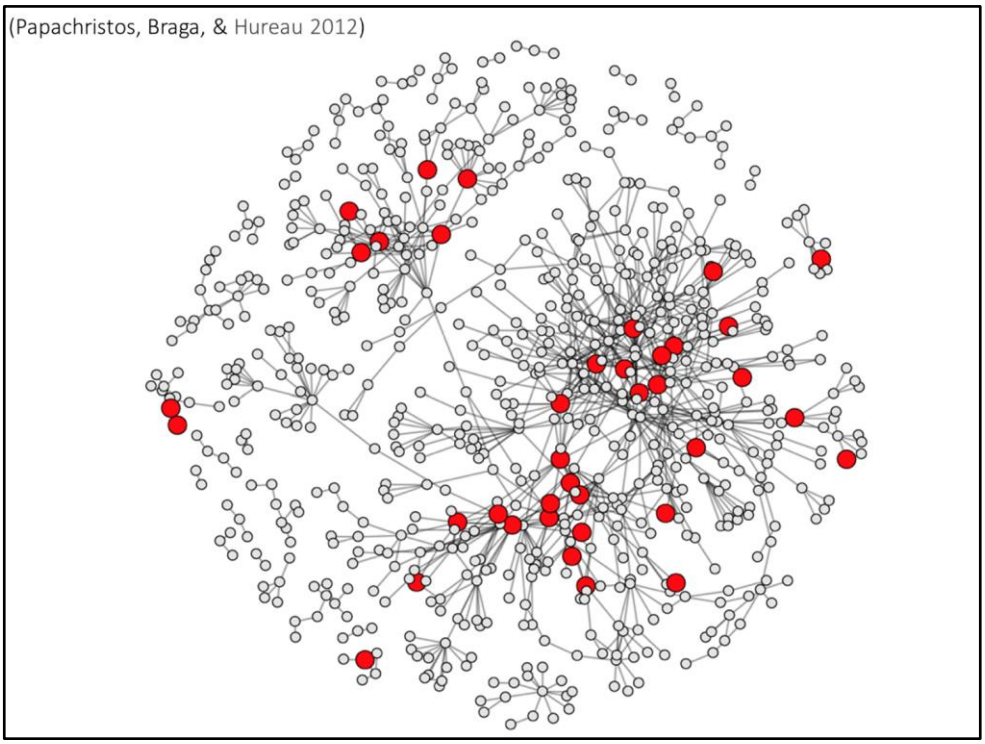


Andrew Papachristos and Christopher Wildeman analyzed a co-arrest network of 3,718 high-risk individuals using homicide and police records. The network is from a predominantly African American and disadvantaged community of 82,000 residents living within a 6-square mile area in Chicago. This community has one of the highest homicide rates in the city, and 41% of the community's homicide victims were located in the co-arrest network. Papachristos and Wildeman's goal was to test the likelihood of being a gunshot homicide victim when controlling for individual characteristics, network position or social distance, and spatial distance. They found that each social tie further away from a homicide victim decreased one's odds of being a homicide victim by 57%. In other words, the further removed an individual from a homicide victim in the co-arrest network, the less likely he or she will be shot. Conversely, those co-arrested with homicide victims were at the greatest risk of gun violence. The takeaway finding from this research is that the risk of homicide in urban areas is even more highly concentrated than previously thought. Homicide spreads through specific types of behaviors and in specific segments of the population. The implications of this finding argue against sweeping policies and practices based on categorical distinctions such as gang membership or race and, instead, focus on intervention and prevention efforts that consider the observable and risky behavior of individuals. Using network techniques to pinpoint groups and individuals at risk for victimization might provide more direct points of intervention and a more efficacious use of limited resources.

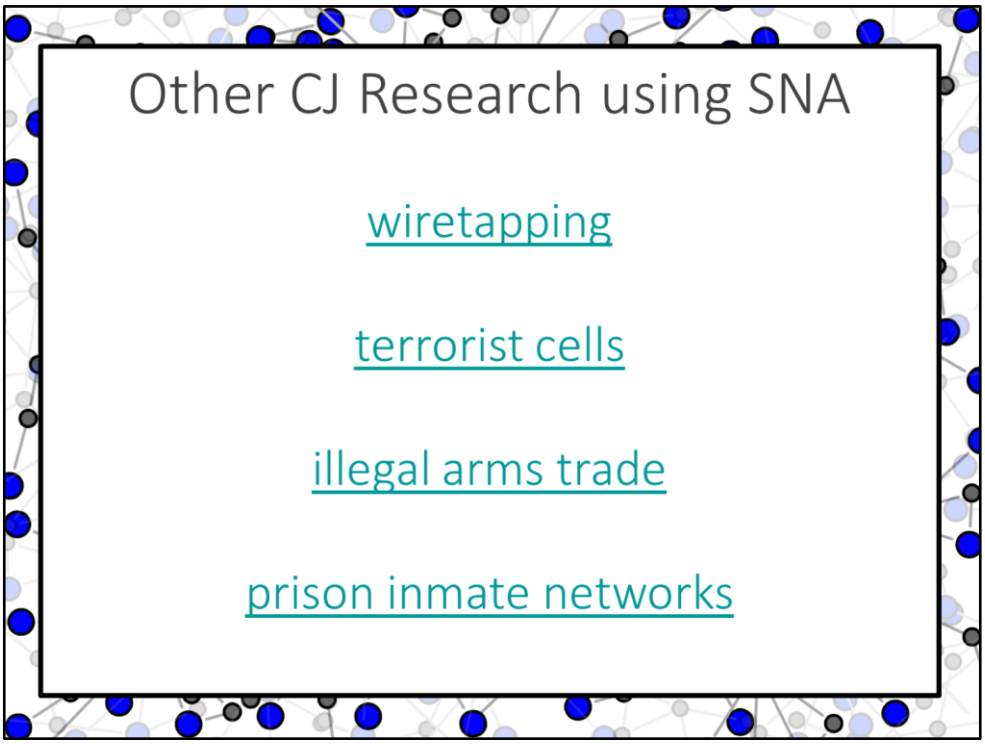
(Papachristos, Braga, & Hureau 2012)



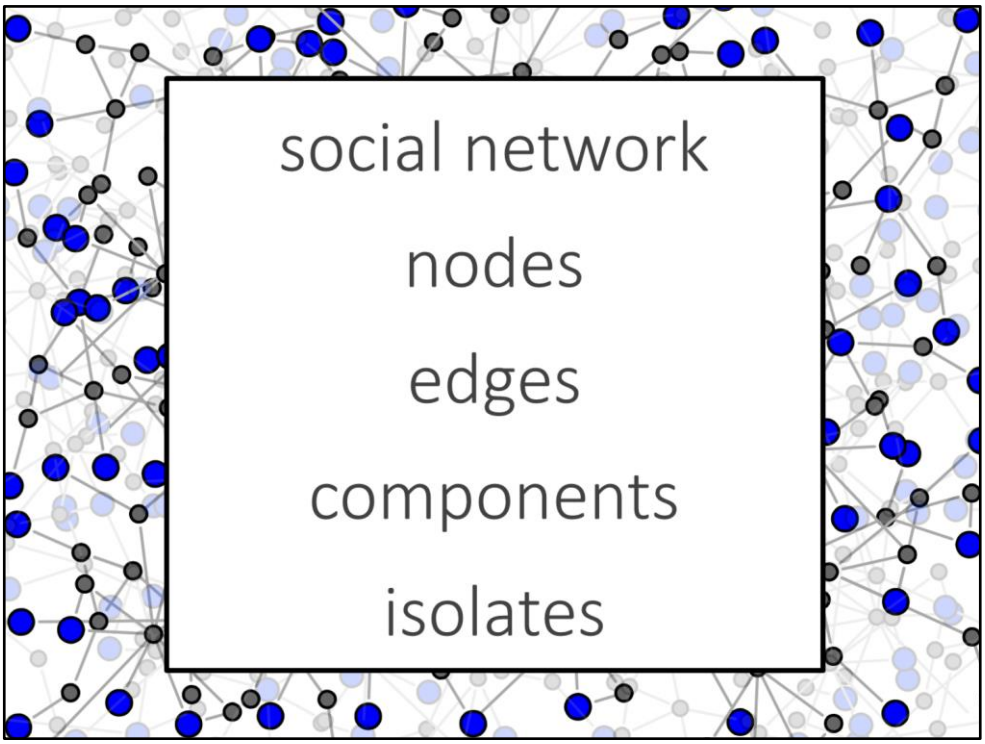
In an article titled, "Social Networks and the Risk of Gunshot Injury," published in 2012 in the *Journal of Urban Health*, Andrew Papachristos, Anthony Braga, and David Hureau, examined a network in Boston. Consistent with prior neighborhood research, Boston neighborhoods with high levels of socio-economic disadvantage tend to also be neighborhoods with high levels of crime. The geographic area of the Papachristos et al. (2012) study includes Boston's two predominantly Cape Verdean neighborhoods that exhibit some of the highest concentration of gunshot injuries. The study area is situated within a larger predominately African American section of Boston, but most of the violence within the study area occurred between Cape Verdean gang-involved youth. Fatal and non-fatal shootings involving Cape Verdean gang-involved youth more than tripled from 12 shootings in 1999 to 47 shootings in 2005. This neighborhood analysis was important context to the arrest network shown in the next slide.



This network was generated from Field Intelligence Observation (FIOs) cards. FIOs are records of non-criminal encounters or observations made by the police and include the reason for the encounter, location, and the names of all individuals involved. There are 763 individuals and 1,869 edges in this network. Edges between individuals represent situations in which two or more individuals were observed in each other's presence by the police and recorded in FIO data. The red nodes indicate fatal and non-fatal gun violence victims during 2008–2009, which also come from official police data. This figure comprises 7 separate components, although 76% of the individuals are connected in the largest component containing 579 individuals. While only 5% of the individuals in this network were victims of gun violence, 85% of gunshot victims from this community were in the largest component. On average, any person is roughly five handshakes removed from a gunshot victim within the largest component. Like the Chicago case, Papachristos et al. found in the Boston study that the closer one was to a gunshot victim, the greater the probability of one's own victimization. In the largest component, every step away further from a shooting victim decreased the odds of getting shot by about 25%. Descriptively and statistically there is much that social network analysts can do with criminal justice data.



Other examples of criminal justice research using social network analysis include: Carlo Morselli, Cynthia Giguere, and Katia Petit's work on drug trafficker communication networks, "The Efficiency/ Security Trade-Off in Criminal Networks" (http://www.sciencedirect.com/science/article/pii/S0378873306000268). Valdis Kreb's work on "Mapping Networks of Terrorist Cells" (http://insna.org/PDF/Connections/v24/2001_I-3-7.pdf). Gisela Bichler and Aili Malm's work on the global arms trade (http://www.tandfonline.com/doi/abs/10.1080/17440572.2013.787928). And Derek Kreager and colleagues' new study researching the social networks of prison inmates (http://www.tandfonline.com/doi/abs/10.1080/07418825.2015.1016090).



To review, a social network is a set of actors and the relationships between those actors. Social networks are the large systems of social connections and social relationships in which individuals are embedded. Nodes are the dots in social networks that represent individuals or actors. Edges are the lines or arrows in social networks that represent different types of relationships, connections, or associations. Components are the different parts of a network that are separate from other connected parts of a network. Isolates are nodes with zero edges. They are separate components that do not have ties to anyone else in the network. SNA has been used to study a variety of different topics. The logic of SNA has broad potential in answering criminal justice and law enforcement related questions. Examples of criminal justice research include analyzing the structure of gang networks or using arrest and affiliation networks to study neighborhood homicide.

REFERENCES

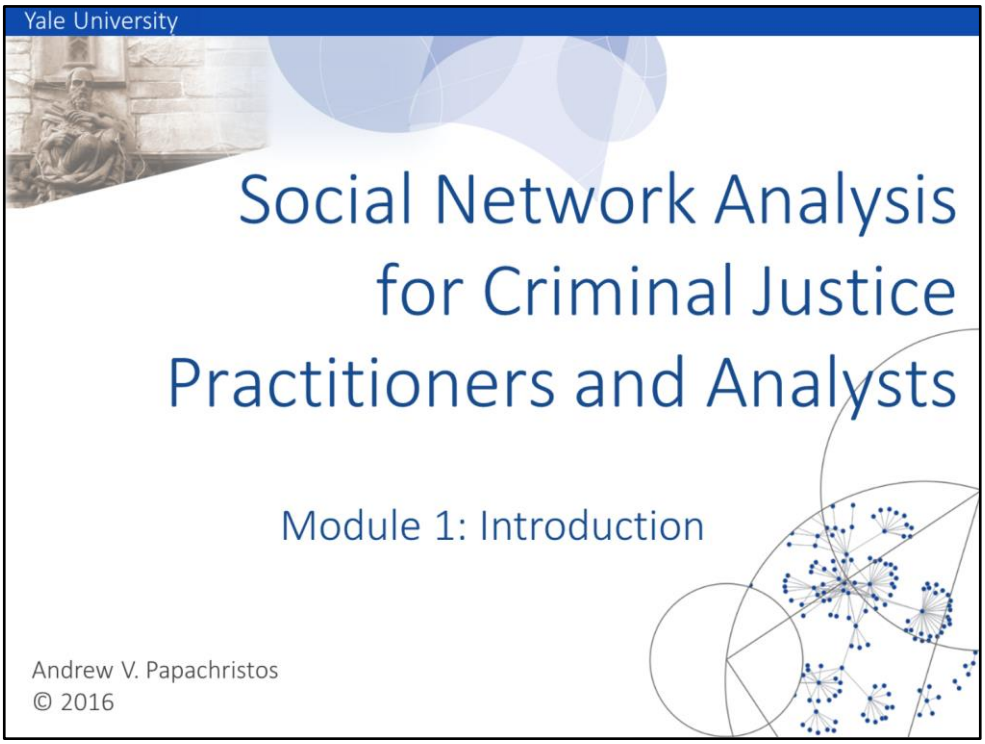
Bearman, Peter S., James Moody, and Katherine Stovel. 2004. "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks." *American Journal of Sociology* 110(1):44-91.

Papachristos, Andrew V., Anthony A. Braga, and David M. Hureau. 2012. "Social Networks and the Risk of Gunshot Injury." *Journal of Urban Health* 89(6):992-1003.

Papachristos, Andrew V. and Christopher Wildeman. 2014. "Network Exposure and Homicide Victimization in an African American Community." *American Journal of Public Health* 104(1):143-50.

Piquette, Jenny C., Chris M. Smith, and Andrew V. Papachristos. 2014. "Social Network Analysis of Urban Street Gangs." Pp. 4981-91 in *Encyclopedia of Criminology and Criminal Justice*, Vol. 62, edited by G. J. N. Bruinsma and D. L. Weisburd. New York: Springer.

Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications.* New York: Cambridge University Press.



This concludes the tutorial section of Module 1: Introduction. Module 1 defined social network analysis, introduced several network terms, presented broad examples of research using social network analysis, and provided applied criminal justice examples of social network analysis. The next part of the Module 1: Introduction is the lab. In the lab section, participants will be guided through some hands on activities that draw from the ideas presented in this tutorial. For participants on the NAVCAP track of this training, there is a second lab called “Introduction NAVCAP that covers installing NAVCAP on to your computer. For participants interested in advanced analysis using RStudio software, there is a lab called “Introduction RStudio" that includes installing and starting RStudio on your computer. The next module is Module 2: Data, which provides the nuts and bolts of building and working with social network data.