# Build a Network Graph

***Create a dataset of people's restaurant recommendations. Use Connect the Dots to analyze it as a network.***

## What is Connect the Dots?

Connect the Dots shows you how your data is connecte with a visual network diagram. It helps you answer questions about the interconnectedness of the data points in your dataset. This hands-on activity helps participants learn about network data by creating and analyzing a dataset about their own interests.

## Learning Goals

- Understanding that relationships between things can be a type of data to analyze
- Understanding when something can be analyzed as network data (such as friendships, web page hyperlinks, and even favorite restaurants)
- Awareness of some common algorithmic analyses of network diagrams
- Experience analyzing a network diagram to answer questions

### Total time

30 Minutes

### Audience

3 - 100 people. Ages 12+.
*We designed this for around 30 people, but you can run it with more or less.*

### Space

- a projector connected to a computer
- ability to break out into small groups of 3 clustered around a computer/phone/tablet

### Supplies

Computers/phones/tablets (1 for every 3 people)

### Reminders

- The recommendations for restaurants are of course limited to the experiences of people in the room... this isn't as thorough as Yelp.
- The algorithmic measurements can give you some insight into the graph and how to think about it. Is it very clique-y? Is it very connected?

## RUN THE ACTIVITY

### Solving a Problem

Networks can be useful for finding certain types of answers that aren't apparent in spreadsheet data. Show a picture of "spaghetti chart" network diagram and ask how many people know how to read it. Introduce the ideas of nodes (dots) and edges (connections) on that chart. Explain how network diagrams capture connections between things, and Connect the Dots introduces analyzing those connections.

### Share Inspirational Examples

Show a few network diagrams that reveal interesting insights. One example is theyrule.net, which shows connections between board members of Fortune 500 companies; showing how a small group of powerful people advise the majority of companies.

## Collect the Data

1. Create a public Google Spreadsheet and write the short url on a blackboard. Add a header row with two columns: “Name” and “Restaurant”
2. Break the room into pairs of people. Ask each pair to talk about their favorite restaurants in a popular town or neighborhood close to your workshop location (pick some place lots of people are likely to have been to)
3. Tell each person to add three rows to the shared spreadsheet, each including their name in the first column and then their favorite restaurant in the next (one row for each restaurant). Tell them it is fine if the restaurants are the same as their partner. Ask them to make sure they spell it right.
4. Open Connect the Dots (https://databasic.io/connectthedots) and copy and paste the petrson-restaurant data into the “Paste Data” tab. Click on “Graph”.

## Introduce the Tool

Show everyone the various parts of the output. The diagram of their data is on the left, showing the connections from people to restaurants. Introduce the two types of scores under the chart, which summarize how connected the dots are in various ways. Introduce the concept of Betweenness Centrality, showing the top “connectors” in the group. Explain how this chart can answer questions the spreadsheet can’t, for instance it can be used to figure out what restaurant you should try that you’d be likely to enjoy (based on the restaurants people you are connected to on the chart love).

## Analyze the Results

Break the room into groups of three and have each pick a dataset to analyze; they can choose the one we just created, or one of the sample sets. Ask them to look for interesting insights into the data based on the network chart. Give them 10 minutes, and tell them that each group will have a chance to share-back their top finding at the end. As they share-back focus on these questions and topics:

- What are the benefits and limits of analyzing data as a network? Could their finding have been discovered in the spreadsheet version of this data?
- How is each algorithm useful or not for making statements about the network and the data it shows?

## Glossary

- **Community** - A group of nodes in a network graph that have more connections to each other than to other nodes outside the community.
- **Bipartite graph** - A network graph with two types of nodes. For instance, people connected to restaurants they love but not to other people.
- **Betweeness centrality** - Scores how much of a “connector” each node is. A high score means lots of nodes can go through one to quickly connect to other nodes.
- **Degree** - The immediate number of connections a node has.
- **Edge** - A connection between one node and another node.
- **Node** - The “dots” in the network - the central organizing things that is connected to others.

data BASIC.io

Engagement Lab civic media

DataBasic is a suite of easy-to-use web tools for beginners that introduce concepts of working with data. DataBasic is a project of the Engagement Lab at Emerson College and the MIT Center for Civic Media. Supported by the Knight Foundation.