Activity #3
Connect the Dots

Purpose

Connect the Dots demonstrates how data can be analyzed using a visual network diagram. This helps participants answer questions about the interconnectedness of the data points in a 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 are also data
- Understanding when something can be analyzed as networked data (such as friendships, web page hyperlinks, and even favorite restaurants)
- Awareness of common algorithmic analyses of network diagrams
- Ability to analyze a network diagram to answer questions

Time

30 Minutes

Supplies

- Ability to break out into groups of three
- Projector
- Computers, phones, tablets *1 for every 3 participants*

Introductory Activity Questions

- **HOW CAN LARGE DATA SETS BE VISUALIZED IN A WAY THAT MAKES SENSE?**

- **HOW ARE DATASETS CONNECTED AND WHY DO THE CONNECTIONS MATTER?**

Instructions

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 "Place".

2 Break the room into pairs of participants. Give the groups a choice of various public community spaces to choose from and tell them to select a type of space as their topic (for example: public parks or restaurants in the community) Ask each pair to talk about their chosen topic.

3 Tell each person to add three rows to the shared spreadsheet, each including their name in the first column and then their topic choice in the next (one row for each choice). Tell them it is fine if they choose the same topic.

4 Open Connect the Dots (https://databasic.io/connectthedots) and copy and paste the person-topic data into the "Paste Data" tab. Click on "Graph".

Introduce the Tool

Show participants 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 the connection between nodes. Introduce the concept of Betweenness Centrality, showing the top "connectors" in the group. Explain how this chart can answer questions the spreadsheet can't.

Debrief

Break the room into groups of three and have each group pick a dataset to analyze; they can choose the one they 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 findings 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?