

The Friendship Paradox: Teacher's Guide

This MIT BLOSSOMS lesson serves three purposes:

- A basic introduction to graph theory as well as a review of mean and variance on a discrete distribution. We chose to focus on the Friendship Paradox specifically because it relates well to social networks that virtually all high-school students use on a daily basis.
- An example of a mathematical proof. This BLOSSOMS lesson may also be one of the first times that high school students see a mathematical proof, but the proof itself uses mathematics that should be familiar to high school students.
- An experience using real-world data to analyze and support a hypothesis, via the social network data that we provide.

It is not our goal that students should come away from this lesson being able to perform similar proofs on their own, but rather that they develop an understanding of how seemingly non-mathematical statements can be translated into math, formalized, and proved.

The lesson alternates between short video segments and activities. Most of the six activities should take around 10 minutes, depending on the length of the class discussion, and are meant to be done with partners or in small groups with a class-wide discussion at the end. The lesson can easily be split across two class periods.

Segment 1 introduces the question of whether students have more friends than their friends do. In **Activity 1**, students discuss this among themselves. You may have them start off discussing in pairs or small groups, and then open the discussion up to the whole class. To wrap up the discussion, ask them how we might explore this question in a more scientific way (leading into the use of the fake datasets in Activity 2).

Segment 2 begins to explore the question with more rigor, and introduces the notion of social networks. At this point, students should understand that apps like Facebook, Twitter, etc., are all examples of social networks. In **Activity 2**, students use the provided social network data to calculate the number of friends a person has, as well as the average number of friends their friends have. This activity helps illuminate the intuition behind the Friendship Paradox: students with more friends than friends-of-friends are the more popular students

Segment 3 introduces the notion of a graph and the accompanying vocabulary (vertices, edges, etc.). Students see how a social network naturally translates into a graph, and in **Activity 3** they perform this translation themselves by creating physical graphs using our provided datasets, lengths of string for the edges, and themselves as the vertices. Students should be able to calculate quantities such as the number of friends-of-friends a person has given a social network graph.

In **Segment 4**, we formalize the Friendship Paradox in terms of graph theory, and introduce the two quantities needed to perform the proof (the average number of friends a person has, and the average number of friends their friends have). Students work out these numbers for a particular instance of a social network in **Activity 4**.

Segment 5 reviews the solution to Activity 4, with the goal of getting the students ready to calculate these quantities for a general social network. Doing so is the subject of **Activity 5**.

In **Segment 6** we review the solution to Activity 5, and then show how the variance among the degrees in a social network appears important for the proof. To help this concept sink in, **Activity 6** requires students to study extreme cases of degree-variation in graphs (one where variance is zero, i.e., all vertices have the same degree, and one where the variance among the degrees is high). This helps give students a sense of what a social network “looks like” when variance is low vs. high.

In **Segment 7**, we complete the proof of the Friendship Paradox. The proof relies heavily on the definitions of the mean and variance, and algebraic manipulations of these quantities. Students should be familiar with mean and variance, as well as summation notation/manipulation of summations before this lesson. However, we do not expect them to be able to complete such a proof on their own. The goal of including the proof is to show students how we can use mathematical rigor to prove a seemingly non-mathematical statement (“my friends have more friends than I do”).