PART 3: The Law of the Few and Power Law Distributions

Instructions

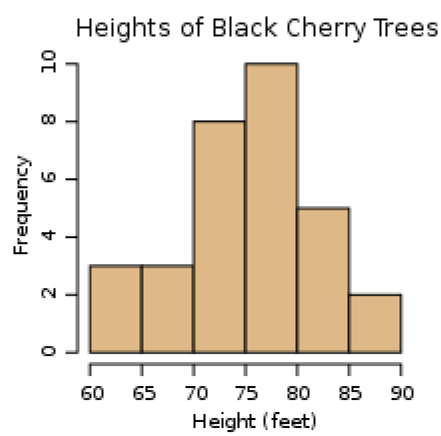
This activity will walk you through loading and analyzing data from our social network (extracted from Facebook). While you work through the activity, there are some questions you need to answer. It should be obvious what you need to answer, but we've made them bold to help you find them.

You need to answer questions in *full sentences* and with some thought or you will only get partial credit. One word answers and/or answers copied and pasted from online sources will get zero credit.

When you are done, email the completed document to: caleb.phillips@colorado.edu with the subject: "BHS Activity 1". If your email address doesn't make your name obvious, then put it somewhere in the document or email. If you don't have an email account, you can print it out and hand it in.

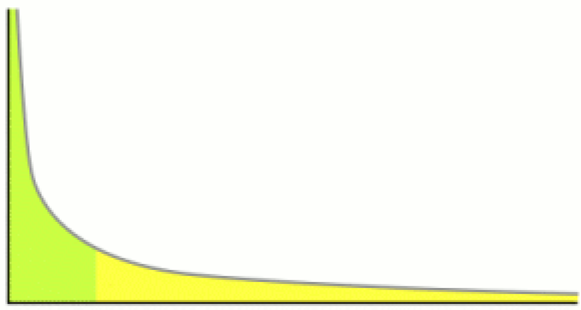
Background

A histogram is a plot that shows the number of times something occurs (the frequency). For instance, here's a plot that shows the number of Black Cherry trees of a certain height.



It shows that the most popular height for Black Cherry trees is approximately 78 feet. And that most cherry trees are between 70 and 80 feet tall. We might call this the "distribution" of Black Cherry tree heights. This is a "normal" distribution because it is shaped like a bell, with a clear peak at the center and sides that slope down.

Another type of distribution, more interesting for us, is a "power law", which looks like so.



Let's say that this plots the frequency of the population of cities. In this distribution, all of the weight is to one side, which means that **most** cities are small, but since there is a long tail off to the right, it also means that a few cities are very large. This is also known as the 80-20 rule. For instance, 20 percent of the people, have 80 percent of the wealth. If we looked at this plot as a plot of the number of people having some amount of wealth, you can see that most people have little wealth (they are on the left), but a few people have lots of wealth (they are on the right). As we discuss the "law of the few" below and how it relates to social networks, this concept will be important...

The Law of the Few in Our Own Social Network

Today we're going to see if we can find evidence of the law of the few in our own social network...

1. Go to this link to see the data we collected from Facebook last week: https://www.google.com/fusiontables/DataSource?snapid=S436487PZui
2. **This is a big "table", where each line is a person in the extended social network of this class. How many people are in the network? Is that more or less than you would expect?**

3. Now, let's see if we can find evidence of the law of the few. We don't have good data showing how many friends each person has since Facebook privacy settings block this, but we do know how many people in *this class* that each person in the list is friends with. Let's look at that data.
4. Click View -> Aggregate. Then select "nconn" next to "Aggregated by:" and finally click Apply.
5. This shows the number of the people in the network having each number of friends within the class. Let's plot this data.
6. Click Visualize -> Line. Then choose "Count" for the Y-axis.
7. **Does this look like a power law? Why or why not?**

8. **How many people in our network are friends with only one person in the class? How many people are friends with 15 people in the class?**

9. **Can you explain how (or why) someone might be "friends" with more than half the people in our class?**

10. **Does this graph support the law of the few? Why or why not?**

11. Now, let's try to find Mavens in the data. Since Mavens collect information, it seems likely they would have many interests relative to most people. Let's look at that.
12. Click Aggregate, then uncheck "nconn" and check "ninterests". Then, click "Apply". Select "Count" for the y-axis as before.
13. **Does this look like a powerlaw? Why or why not?**

14. **How many people have only 1 interest? How many people have more than 400 interests?**

15. **Due to changes in facebook privacy settings, it's difficult to get counts of people's comments or "likes". However, if we could get this data, how would you expect the distribution to be similar or different from that of "interests" given above?**

16. **Do you think the people that appear to be Mavens, Connectors, and Salespeople in facebook, would also be Mavens, Connectors, and Salespeople in "real life" (i.e., as Gladwell describes these roles)? Why or why not?**

17. **With the rest of the time, search around on Google for other examples of Power Law distributions. List three things here that also have been shown to follow a Power law:**