

## **Graphing Exercises: Introduction to Visualizing your Data through Graphs**

Betsy A. Colburn

Schoolyard Ecology Looking at Data Workshop for Teachers-2019

Harvard Forest

Graphs of your data can provide valuable insights into patterns of change over time; differences between individual trees, species, or sites in growth, phenology, or distribution; and they can help with comparisons between different study trees or sites sampled by your students or by other schools.

Often, sketching out rough graphs by hand can help you identify the kinds of graphs you would like to prepare and the kinds of results you can expect. (This is not usually feasible for an entire, original data set, but it can be useful once you have extracted additional data from your primary field data, by combining information, making calculations based on the field measurements, or by other means.) Examples of extracted additional data include: mean biomass accumulation by tree species in a plot; growing season length for different tree species or different sites; percent of study trees that are infested with hemlock woolly adelgid.

The graphing exercises presented here provide you with opportunities for creating different kinds of graphs with which to visualize results from long-term ecological studies of trees or other organisms, especially those involving the Harvard Forest LTER Schoolyard Science protocols.

For example, you and your students may have studied a number of trees and shrubs representing a number of species in a Changing Forests study or in sampling for Buds, Leaves, and Global Warming. Each species represents a proportion of the total trees studied, and it might be interesting to examine the distribution of species in the study. One way to do this is shown in Exercise 1, which prepares a simple pie graph showing the relative numbers of different species of trees sampled in Athol-Royalston Middle School's study of Buds, Leaves, and Global Warming, and in Applewild School survey for Changing Forests. The individual graphs provide information about each school's study tree population, and comparing the graphs shows some interesting differences between the study populations.

You can make one or both of the same graphs, or create a similar graph with your own students' data on trees studied for bud-break and leaf fall or from Changing Forest plots studied for diameter change and biomass accumulation. Looking at

different pie charts to compare schools' study tree populations can be interesting. Would you expect study populations to be similar for both studies, or would you expect differences? What might be the basis of expected differences, if any?

The same question about distribution of tree species in a study can be examined with different kinds of graphs. Exercises 2 and 3 look at the same data as Exercise 1 but present the story through stacked and individual bar graphs. Bar graphs are also useful for looking at many other kinds of data, such as length of growing season or biomass accumulation in individual trees, or across study sites.

Later, you may want to look at measures such as growing season length or annual biomass accumulation by within a species, or to compare species' behaviors across study sites. Schools looking at hemlock woolly adelgid and how it affects growth in hemlock trees may want to graph growth of trees and compare results for trees or stands with different levels of adelgid infestation. The final graphing exercises look at creating scatter plots and line graphs, or alternately bar graphs, to examine other questions about data, such as changes in measured variables over time.

The overall purpose of the exercises is to help you become more comfortable with creating graphs from schoolyard science data, with looking at the visual presentation of the data in different graphs, and to come up with ideas of interpreting results or sharing ideas with your students.

NOTE: We realize that many teachers work with GoogleSheets in the classroom, and that many of you wish to use GoogleSheets to create graphs of your schoolyard data. This is not a tool regularly used by most scientists, and we do not have a lot of familiarity with GoogleSheets at Harvard Forest. The basic concepts involved in selecting graphs, creating them, looking at the results, and interpreting the story that is told, are the same regardless of the graphing tools that are used to create graphs. Our goal is to help each of you toward meeting your goals for graphing your schoolyard data and for including graphic visualization tools as part of your classroom teaching.

The basic concepts behind data visualization are the same, whether the data are from Changing Forests, Buds, Leaves, and Global Warming, Woolly Bully, or Vernal Pools in the Landscape. We have prepared a separate set of exercise summaries that specify steps for working in GoogleSheets to prepare certain kinds of graphs, and we hope that those will be helpful to teachers who want to graph their data with this tool. These GoogleSheets summaries use Changing Forests data as examples, but the basic steps in creating graphs will work with data from the other protocols as well. Similarly, effective graphs can be created using Excel, other graphics programs that come with different computers, and through hand drawing of pie charts, bar graphs, and scatter plots, and line graphs.

## Graphing Exercise 1. Create a Pie Graph by Selecting a Data Table

This exercise selects the data in an existing data table and creates a pie graph of the data (Figure G1). This exercise suggests using data from trees studied for Buds, Leaves, and Global Warming or from Changing Forests plots.

Figure G1a. Tree species sampled in schoolyard phenology study at Athol-Royalston Middle School, 2004-2007.

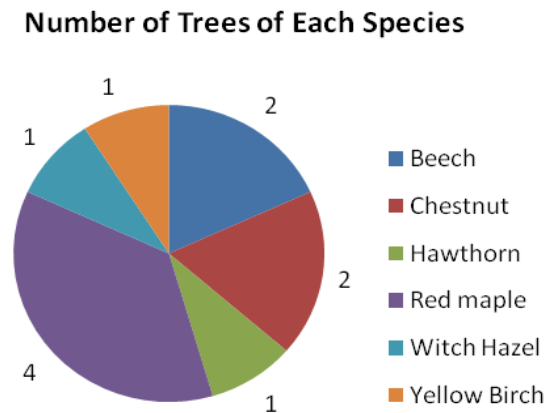
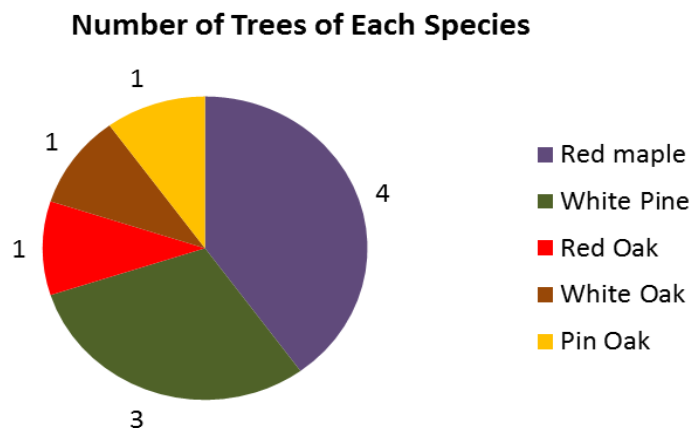


Figure G1b. Tree species sampled in schoolyard Changing Forests study at Applewild School, 2016.



Pie graphs, also called pie charts, can provide useful visual comparisons of numbers of kinds of things in a study. Other examples might be: animals found in vernal pools, number of plots being sampled by different schools working on a

protocol, or any set of information about different numbers of certain kinds of things.

### **Educational Goals of Exercise 1**

- Before graphing, become comfortable with working with data within a spreadsheet
- Select data you want to graph
- Choose a graph type for your data
- Create a graph
- Format the graph
- Consider the mathematics involved in creating a pie chart

### **Summary of Steps for Making a Pie Chart/Pie Graph**

1. Input a new data table of tree-species data into your worksheet (use the data in Table G1a, Table G1b, your own data, or another set of data you want to work with). NOTE: You may want to include only data for trees species found in your study site, or you can include all species if you want to create multiple graphs for comparison.
2. Select both columns of data in the table you created, including the column headings.
3. Instruct the computer to insert a pie graph.
4. Format and edit the completed graph as you choose.<sup>1</sup>
5. Consider what mathematical changes the program made to the data in order to convert the column of tree numbers into a pie with different-sized slices.

<sup>1</sup>NOTE: Formatting and editing a pie chart can include specifying a title, adjusting the legend, changing colors for individual pieces of the pie, deciding whether to include data values on the graph and where to put them, choosing whether to have a solid pie or one with the pieces exploded, and various other options. Take the time to explore ways to change the appearance of the pie chart you have created, and consider how the changes enhance or fail to enhance the story you want to tell.

### **Graph Analysis – Questions**

1. What patterns do the study trees have in common between the two schools?
2. How do the study trees differ between the schools?
3. Is this a useful way to compare study populations? Are there other kinds of graph that might make the comparison clearer?

Table G1. Breakdown of sampled trees by species.

(a) ARM Schoolyard Phenology Study, 2004-2007

<u>Tree Species</u>	<u>Number of Trees</u>
Beech	2
Chestnut	2
Hawthorn	1
Red maple	4
Witch Hazel	1
Yellow Birch	1

(b) AWH Changing Forests Study, 2016.

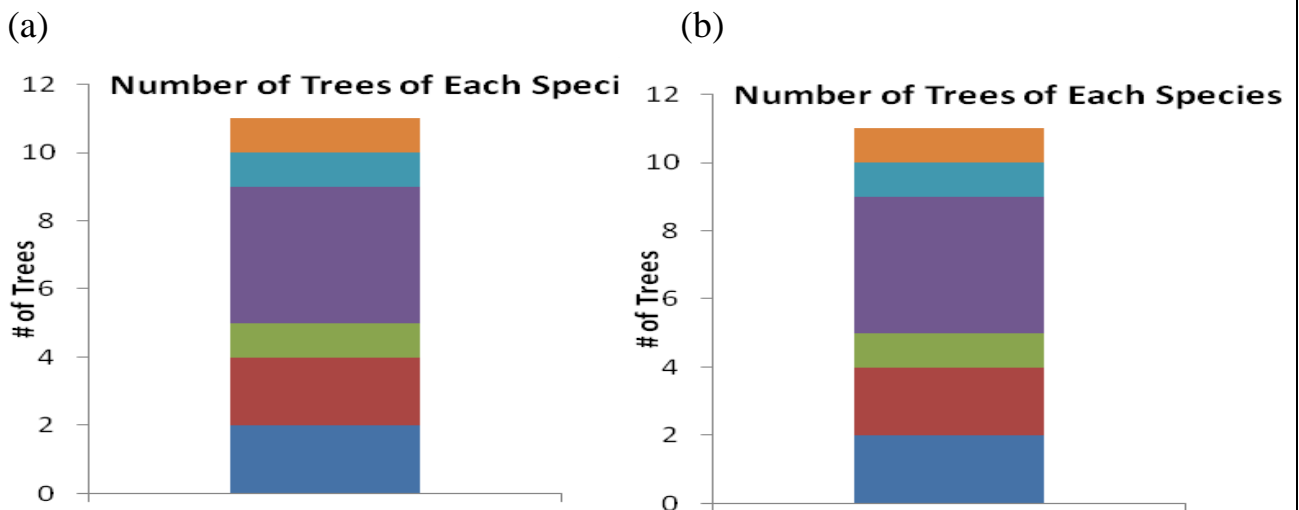
<u>Tree Species</u>	<u>Number of Trees</u>
Red Maple	4
White Pine	3
Red Oak	1
White Oak	1
Pin Oak	1



## Graphing Exercise 2. Create a Stacked Bar Graph

The data you graphed in the pie chart exercise can also be shown as a stacked bar graph. You can create this kind of graph the same way you made the pie chart, by selecting your data and instructing the computer to create a stacked bar graph, and formatting appropriately. Use the data table presented for Exercise 1.

Figure G2a. Trees sampled by schools studying forest ecology as part of Harvard Forest Schoolyard Science Program. (a) Fall phenology study by students at Athol-Royalston Regional Middle School, 2004-2007. (b) Changing Forests study at Applewild School, 2016



## **Educational Goals of Exercise**

- Continue to become comfortable with working with data in a spreadsheet and organizing the data for graphing
- If working in Excel or a comparable program, learn how to select specific data points for graphing from a larger data set
- Practice creating a stacked bar graph from data on tree species data from Buds, Leaves, and Global Warming or Changing Forests studies
- Review the graph and the story it tells, and compare the results with those from the pie chart. Which do you find more informative?

## **Summary of Steps for Making a Simple Stacked Bar Graph**

1. Use the data table you input into your worksheet for Exercise 1, or input a new data table of tree-species data into your worksheet (use the data in Table G1, your own data, or another set of data you want to work with).
2. As in Exercise 1, select both columns of data in the data table, including column headings.
3. Instruct the computer to insert a stacked bar graph. *Note: If the graph does not appear as expected, try the Switch Row/Column option.*
4. Format and edit the completed graph as you choose. See note on editing in Exercise 1.
5. Consider what mathematical changes the program made to the data in order to convert the column of tree numbers into a pie with different-sized slices.
6. Look at the graphs. Is it easier to compare the study tree populations in the bar graphs than in the pie graphs? Is there a general question that emerges from comparing these two study populations? If so, what is it, and how might you answer it?



## **Optional Exercise G2a: Use Source Data Dialog Boxes to Select Data for Graphing**

Some graphing programs, such as Excel, provide “Source Data” dialog boxes that allow you to specify the data that you are going to graph very precisely. This can be very helpful when you are working with a large or complex data set and want to choose specific kinds of data for graphing. With simple data such as those shown in Exercise 1, it is easy to input the data into a worksheet and create a simple bar graph using the methods employed above for the pie graph, but if the data set is large or more complex, inputting a new data table can be time-consuming and subject to errors. If you would like to explore how to select data points individually from a larger data set, the simple nature of the data table G1 lends itself to use as an example. If you create graphs with Excel instead of GoogleDocs or another program, once you are familiar with the Source Data dialog boxes, it will save you time in creating more complex graphs of data on biomass accumulation, leaf fall, bud burst, water-level changes, hemlock woolly adelgid infestation, and other field variables you and your students measure over time for schoolyard ecology studies.

For this exercise you will graph the same tree-species data as in the pie graph in Graphing Exercise 1 by using the Source Data dialog boxes in Excel to specify the data we will be using to create a stacked bar graph as in Figure G2. As before, you may choose to graph data on trees in a Changing Forests plot, or other data you wish to work with.

### **Educational Goals of Exercise if Graphing with Data Source Dialog Boxes in Excel**

- Practice, and become comfortable with, specifying data to be graphed using Data Source dialog boxes in Excel, or comparable graph-creation methods in other graphing programs.
- Become familiar with ways of selecting data in a complex graphing program

## Summary of Steps

1. Click on a blank cell in your worksheet.
2. Tell the computer to insert a graph, and select the type of graph you want to create.
3. Use the Data Source dialog box to select a data range for graphing from a data table.
4. Use the Data Source dialog box to specify individual data to be graphed, creating a series.
5. Format/edit the graph, as appropriate – format axis titles, change font size, adjust the range, alter fonts and colors, and make other changes to make your graph attractive and informative.
6. Take a few minutes to consider how different formatting options enhance or otherwise contribute to (or detract from) the message conveyed by the graph. Also, compare the graph you just made with the pie chart made in Exercise 1. Does one do a better job, in your opinion, of showing the information you want a viewer to see?

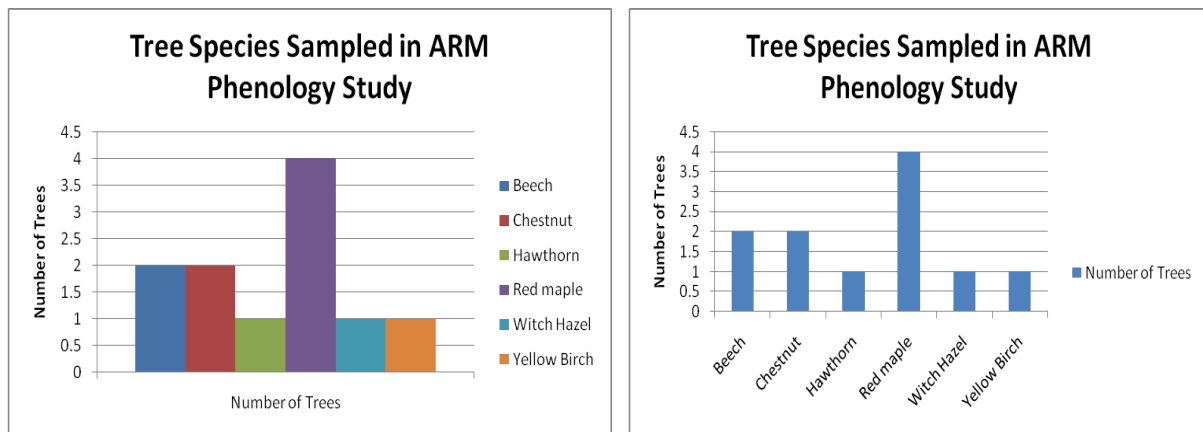
### Graphing Exercise 3. Creating a Simple Bar Graph from an Existing Graph

Most graphing programs allow you to convert one kind of graph into another kind without having to go through all the steps involved in creating the graph from a data table. In this exercise, you will use the Chart Type dialog box in Excel, or the comparable tools in other graphing programs, to convert an existing graph into another kind of graph, specifically the simple bar graphs shown in Figure G3.

Figure G3. Two versions of simple bar graphs produced from the pie graph and stacked bar graph in Figures G1 and G2.

Left, graph produced by specifying that data are in rows.

Right, graph produced by specifying that data are in columns.



### Educational Goals of Exercise

Develop facility with tools provided by your graphing program to change an existing graph into another kind of graph, and consider information from each.

- Practice using graphing tools that allow you to switch between different types of graph for a given set of data.
- Compare the graphs of the data and consider the different information and possible interpretations you can find in the different graphs.

## Summary of Steps

1. Select an existing graph, such as the pie chart you created in Exercise 1 for trees in a Phenology or Changing Forests study.
2. Use the Change Chart Type dialog box and/or the Source Data dialog box in Excel, or the comparable tools in other graphing programs, to convert the graph into a simple bar graph with the number of individuals sampled for each species shown as a separate vertical bar.
3. Edit the graph as needed.
4. Once you are happy with the appearance of the graph, use the Switch Row/Column command in Excel, or the appropriate tool in your graphing program, to switch between the two versions of bar graphs shown in Figure G3.
5. Spend a few minutes thinking about the different ways you have depicted the same set of data. Do the different graphs tell slightly different stories? Do they suggest different questions? Do you find one more compelling than another?

## Note on Other Uses of Bar Graphs

Bar graphs are useful for illustrating many kinds of data. The examples here have a quantitative Y axis but a categorical X axis, relating to individual tree species. The X axis can also be quantitative, reflecting time, size, number, or other values.

As shown in slide show, examples of bar graphs depicting different aspects of Schoolyard Science data might include:

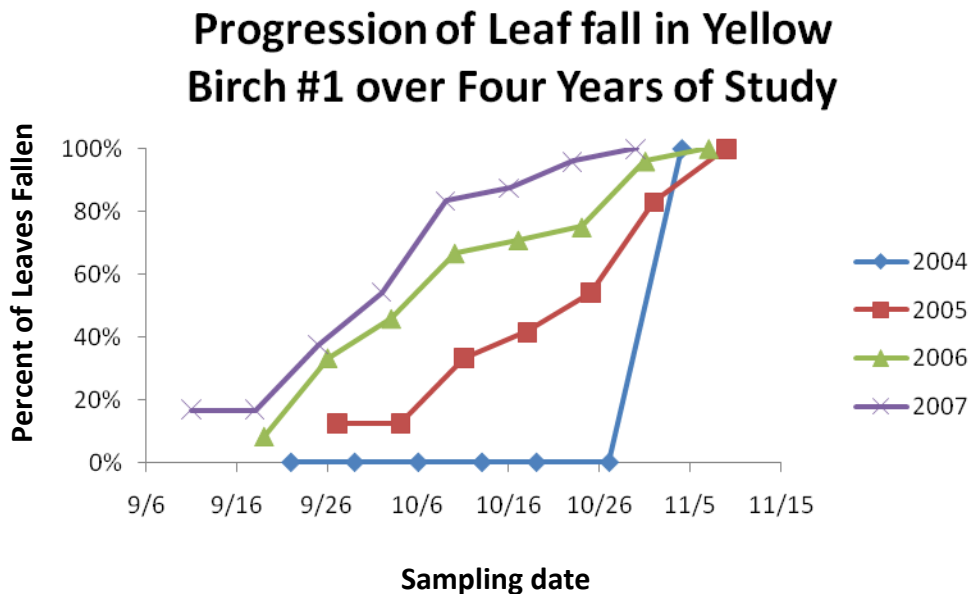
- Biomass increment or branch growth for individual trees at a site (X axis is tree or branch ID, Y is biomass or growth increment);
- Biomass increment by each species at a site, with the biomass increase of all individuals of each species combined for the species values;
- Biomass accumulation in a plot in a given year, with all species and individual trees included in the biomass data point – this might be interesting for comparing across sites;
- Comparison of biomass increment for species or plots between sites;
- Biomass accumulation in a plot in a given year, with all species and individual trees included in the biomass data point – this might be interesting for comparing across sites;
- Dates of 50% budburst or leaf-fall over time;
- Length of growing season over time; and
- Many others.

## Graphing Exercise 4. Graph Leaf-Fall Data from One Tree

This exercise will graph leaf fall from a single tree over four years of sampling. It involves creating a scatter plot with time on the X axis and percentage of leaves fallen on the Y axis, and using study data from multiple years as the data for each sampling period. The graph will be similar to Figure G4.

Similar graphs can be created with data from all of the schoolyard Science study protocols. For example, graphs can be made with Year in the X axis and with the Y axis representing length of growing season, date of bud burst or final leaf fall, growth increment of a tree or a branch, number or percent of trees in a study that were infested with hemlock woolly adelgid, and the different lines representing individual trees, different plots, or different school studies. This example just illustrates the process for making such a graph. Feel free to explore with different data, including your own.

Figure G4. Progression of leaf fall in a yellow birch sampled in 2004-2007 by students at Athol-Royalston Middle School



## Educational Objectives of Exercise

Build on what you learned in the earlier exercises to create a complex graph from multiple years of data.

- Review and practice sorting data in a spreadsheet
- Learn how to convert count data in a spreadsheet to percentages
- If working in Excel or a similar program, become comfortable with using the Source Data dialog or other data-selection features to create a complex graph.
- Increase confidence in formatting the graph to produce a clear presentation of the results
- Further develop skills in interpreting the message the graph provides

## Summary of Steps

(This assumes you are using Buds, Leaves, and Global Warming data on Leaf-fall progression; if using other data, the data manipulation and axis details will be different)

1. Go online to the Harvard Forest LTER Schoolyard Database. Download fall data from Athol-Royalston Middle School (ARM), your own students' data, or another school with multiple years of data, or work with your own Changing Forests data (if you don't have multiple years of data, you can still create a plot of Y vs X to answer questions about your data, it will just have different axes than a leaf-fall graph).

NOTE: When you are downloading data from the Schoolyard LTER database, you have a number of choices – including the school(s) and project(s) whose data you want to work with. *After you click on the general download data option for the Schoolyard LTER database, you will first need to choose a project and then select a school. Don't forget to click submit after selecting a school. After you see a list of appropriate data, click download data. A temp file will appear at the bottom left of the page.*

2. Sort the data to organize the data by tree, so each tree's data are presented chronologically, together.

*NOTE: When sorting for this exercise, you must select/highlight your entire data set. Click Sort by TreeID. Add a sorting level, to sort each TreeID by date. Remember to click the My Data Has Headers option at the top right for easier sorting.*

3. Add a new column to the data table, and in this column insert the percent of leaves fallen for each tree on each date. (The data set shows number of leaves fallen, so you will need to convert the number fallen on each date into a percent of all leaves observed on that date.)
4. Insert a new graph that will show percent of leaves fallen vs. date. Specify that you want a scatter plot with lines connecting the points.
5. Use the Source Data dialog boxes in Excel, or the comparable tool in a different graphing program, to add the leaf-fall data for Yellow Birch Tree YB1 from 2004, 2005, 2006, and 2007 (you may add data from later years, if you wish).

*NOTE: The data from each year will be a separate series or legend entry in your graph. Each year is added individually. Note: The data for 2007 contains NA. These missing/not applicable values give spreadsheet programs fits. They must be removed from the data set.*

6. Edit the graph, as appropriate. You may want to adjust the symbols used for the different years, their colors, and their sizes; edit the legend; format the axes to make them more readable; or make other changes to make the graph more effective at showing the results of the study.
7. Spend a few minutes looking at the graph and thinking about the story it tells and questions that it generates.
  - Is there a pattern over time?
  - In what year was leaf-fall finished earliest? When was it finished latest?
  - What might explain the observed pattern?
  - How does the pattern fit with the basic hypothesis behind your schoolyard science study?

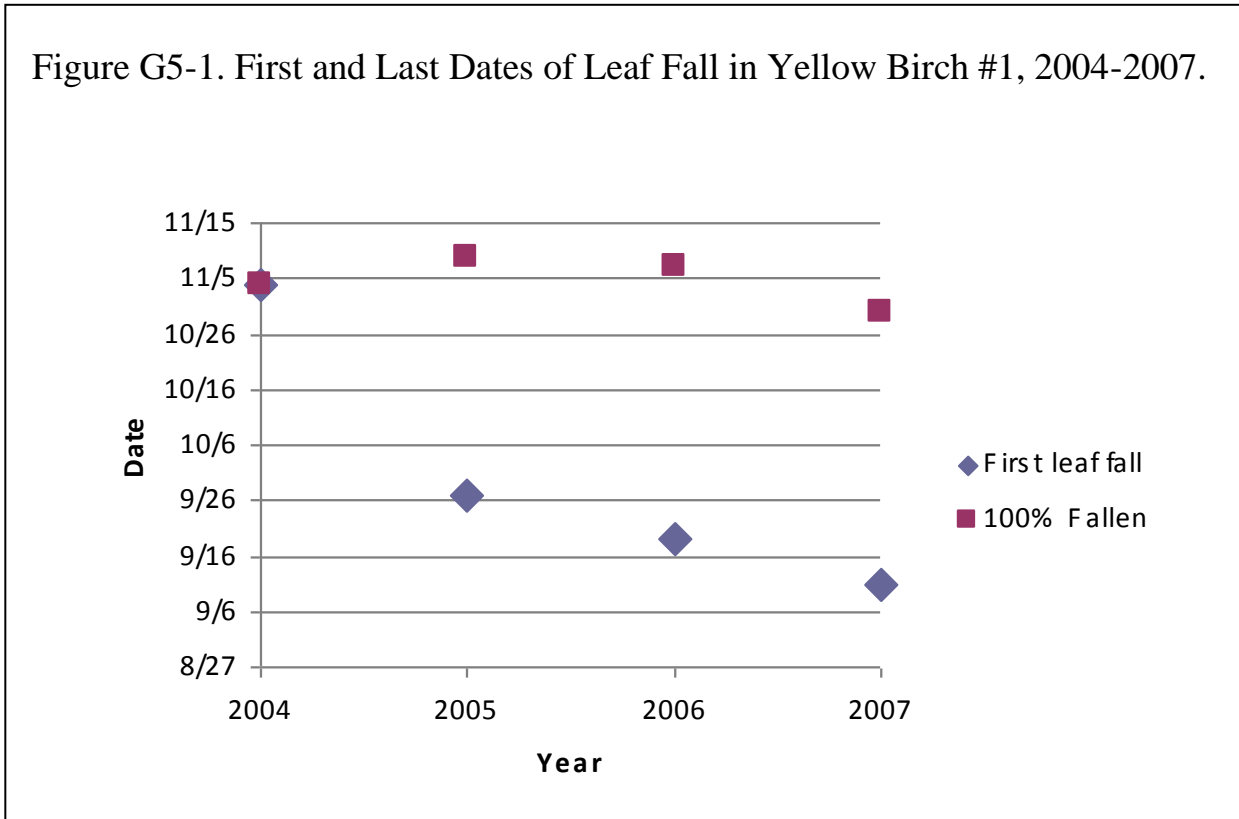




**Graphing Exercise 5. Create a Scatter Plot and a Side-by-Side Bar Graph to Compare First and Last Leaf-fall Dates, Growth Increment in Different Tree Species or Individual Trees, or Other Data**

In this exercise, you will use a scatter plot to compare two sets of data. These data might be the first and last dates when leaves were observed to fall from Yellow Birch #1 during the four years of the ARM phenology study shown below in Figure G5-1. You could also plot the annual biomass increment by species of tree or for two individual trees for different years or sites in a Changing Forests study.

To create this graph, it is necessary to create a new data table. The data table will have two columns of data. For first and last leaf-fall it will contain the first and last leaf-fall dates for tree YB1, organized by year – this data table is provided below in Table G5. For biomass increment, the data table will contain the biomass data for the study population you are interested in. The same is true for any other set of data you might want to compare in this way.



## Educational Goals for Exercise

- Become comfortable with manipulating schoolyard data to extract additional information from the original field results
- Explore multiple ways of graphing data on first and last dates of observed leaf fall over multiple years, or similar data
- Practice reformatting axes to present time in different ways
- Think further about how different ways of presenting the same data set can provide different information to the viewer
- Think about the data and the story told by the graphs here, and how they differ from the stories told by earlier graphs.
- How can these “stories” contribute to understanding of what is going on in the study trees, and our forests at large?

## Summary of Steps

1. Create a new Data Table showing dates of first and last leaf fall, the growth increment for different tree species or individual trees, or other data that you would like to present in a scatter plot looking at a variable over time
2. Select the data table, including column headings
3. Insert a graph, and format as necessary.
4. Look at your graph. What is the story it tells? Does it help address the basic question behind the long-term schoolyard study? Why or why not?

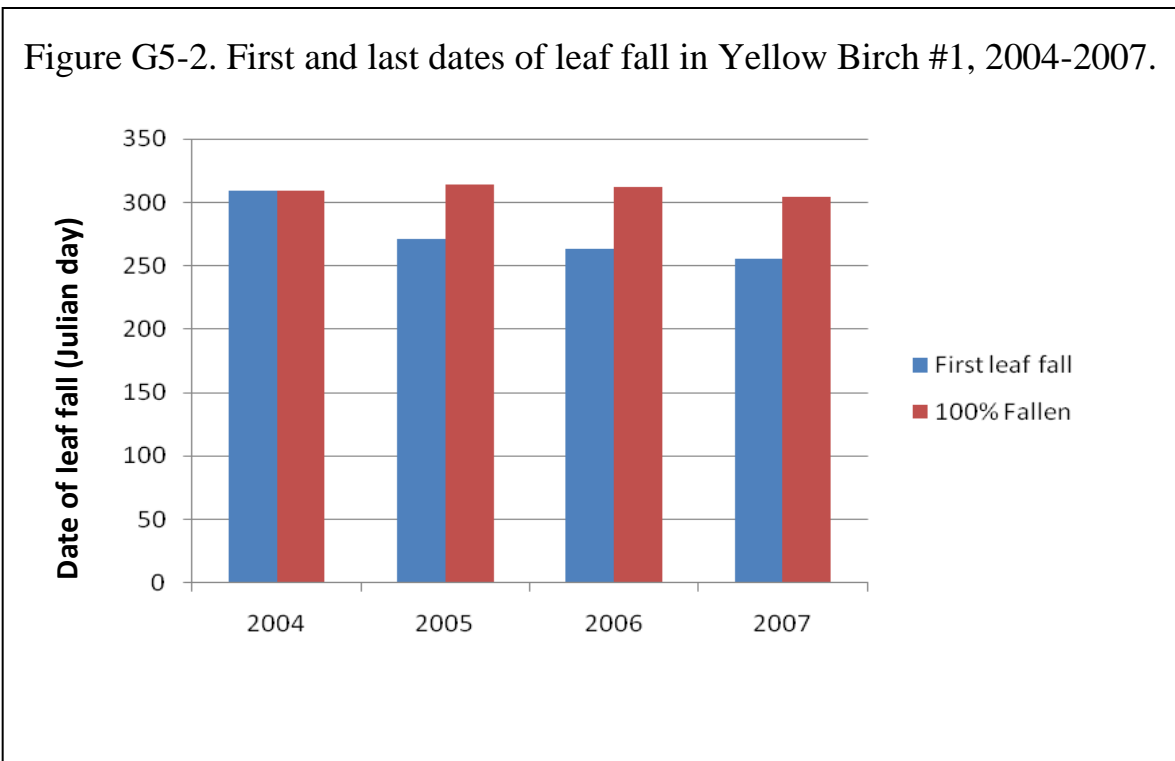
Table G5. Data Table for first and last leaf-fall dates for tree YB1, 2004-2007

Year	First leaf fall	100% Fallen
2004	309	309
2005	271	314
2006	263	312
2007	255	304

### Optional Variation: Graph the Data as a Bar Graph.

In Figure G5-2, the data in Figure G5-1 are graphed as a bar or column graph comparing matched pairs of data. As with the initial scatter chart that simply presented the data as points, it is easy to see how the dates when the students observed the first fallen leaves became earlier each year over the four years for which data are available. The similar but less consistent trend in the date of last leaf fall is also visible.

You can either create a bar graph from scratch, using the same process you used to make your scatter plot, or you can use your graphing program's option for changing an existing graph into another kind of graph.



When you have completed your graphs, once again, take a few minutes to look at them carefully. Think about the “story” the graphs tell and identify questions that are raised by the graphs.

- Do you find the bar graph or the individual points more informative?
- Is there further formatting that would make either graph more informative?

- Are there some anomalies that might need to be explained?
- Are there questions raised by the patterns shown in the graphs that suggest additional research that is needed?