ISSUES : DATA SET

Effects of bison grazing on plant diversity in a tallgrass prairie (Konza Prairie LTER)

Harmony J. Dalgleish, Division of Biology, Kansas State University, Manhattan, KS 66506. Current address: Utah State University, College of Natural Resources, Logan, UT, 84322, h.dalgleish@usu.edu.

Teresa M. Woods, College of Education, Department of Secondary Education, Kansas State University, Manhattan, KS 66506, twoods@ksu.edu.

THE ECOLOGICAL QUESTION:
How does grazing by large ungulates affect plant community diversity?

ECOLOGICAL CONTENT:
Diversity, herbivory, plant species composition, quantifying diversity, influence of sample size, species accumulation curves, tallgrass prairie ecosystem

WHAT STUDENTS DO:
The exercise is aimed at beginning and intermediate-level ecology students and is divided into three parts. Part 1 introduces biodiversity and includes a short, small group discussion activity to introduce the concepts of richness, evenness, and diversity indices. Part 2 contains three computer-based activities that explore three concepts in detail (species richness, the importance of sample size, and diversity indices) using plant species composition data from the Konza Prairie LTER site. After mastering the analytical techniques presented in Part 2, as well as understanding some of the biases of each technique, students will apply their knowledge to compare plant species diversity in two different tallgrass prairie plant communities: one that has been grazed by bison for >10 years and one that has been protected from bison grazing for >20 years. Part 3 challenges students to consider the assumptions of the methods they have learned and how these assumptions may affect their conclusions. Part 3 ends with a brief summary and conclusions about the effects of bison grazing on plant species diversity.

STUDENT-ACTIVE APPROACHES:
Small group discussion, guided discussion, writing to learn, computer-based projects, calculation

SKILLS:
The exercise develops a conceptual understanding of diversity while employing quantitative reasoning skills and data presentation techniques. Specifically, students will practice the following skills: using Microsoft Excel spread sheets, calculating and interpreting diversity indices, presenting data using figures and tables, analyzing results, and critical thinking.

ASSESSABLE OUTCOMES:
Students will produce tables and figures to present the results of their data analyses. In addition, students will respond to the questions posed in the exercise. Instructors can decide whether to use the questions for small group discussion, for written analysis to hand in for evaluation, or both.

**SOURCE:**

**ACKNOWLEDGEMENTS:**
We would like to thank Dr. Victoria Clegg for the opportunity to develop this exercise in her Principles of College Teaching class at KSU; Dr. Gene Towne for collecting the LTER data; Dr. David Hartnett for the inspiration to develop this exercise and for helpful comments on previous drafts; and Jayne Jonas for helpful comments on previous drafts. In addition, we are indebted to three anonymous reviewers and the *TIEE* editors for significantly improving this exercise.

**OVERVIEW**

The research
The Konza Prairie Biological Station (KPBS) encompasses 3,487 hectares of native tallgrass prairie. Established as a preserve in 1971, it is currently owned by The Nature Conservancy. The Kansas State University Division of Biology manages KPBS as a field research station.

KPBS lies within the larger Flint Hills region of Kansas and northern Oklahoma, which are characterized by steeply sloping hills (Figure 1). Because the shallow soils of the region are not conducive to farming, KPBS and much of the surrounding Flint Hills have never been ploughed. The Flint Hills constitute the largest contiguous area of unplowed tallgrass prairie remaining in North America. The Flint Hills are at the southern and western edge of the historical extent of tallgrass prairie, which extended to southern Minnesota in the north and Indiana in the east.
The tallgrass prairie at KPBS is dominated by the perennial, warm-season grasses *Andropogon gerardii* (big bluestem), *Schizachyrium scoparium* (little bluestem), and *Sorghastrum nutans* (Indiangrass). Other important, though less abundant groups include warm-season and cool-season grasses, composites, legumes, and other forbs (Photo 1). A few woody species such as *Ceanothus cuneatus* (buckbrush) and *Rhus glabra* (smooth sumac) are locally common. The grasses can reach heights of two and a half meters in the wettest years, which is why this vegetation is known as tallgrass prairie.

![Konza Prairie and Flint Hills region map](image)

**Figure 1.** Map of the conterminous United States showing the location of Konza Prairie within the Flint Hills region of Kansas.

**Photo 1.** *Lespedeza stuevei* growing amongst a mixture of big bluestem and Indian grass. Photo by T. M. Woods
KPBS was one of the original eight sites of the NSF-funded Long-Term Ecological Research (LTER) program (www.lternet.edu). A fully replicated, watershed-level experimental design has been in place since 1977 to test the effects of a key driver of tallgrass prairie ecosystems: fire (Knapp and Seastedt 1998). In 1987, bison were re-introduced to KPBS to study the effects of another important ecosystem driver in tallgrass prairie: ungulate grazing (Photo 2). Lastly, meteorological instruments have collected weather data since the 1970’s to examine the importance of the third important driver of tallgrass prairie ecosystems: a variable continental climate.

Over the years, the Konza LTER program has built a long-term database derived from the watershed-level fire and grazing experiments. One of the long-term datasets collected at KPBS includes the plant species composition dataset. Twice a year, a researcher samples permanent transects in several watersheds to estimate the relative abundances of all the plant species that are present. Along each transect, there are five circular plots, each 10 m² in area (Figure 2). Within each plot, the researcher identifies each species and estimates how much of the plot is covered by that species (a measure of abundance, which we will explore more later in the exercise).
Figure 2. Konza Prairie is divided into watershed units, and each watershed has a different treatment combination of fire frequency and grazing. Long-term research watersheds, such as N1B detailed here, have four permanent transects in the uplands and four permanent transects in the lowlands. Each transect contains five permanent circular plots.

We will use plant community data from two of these watersheds in this exercise. One has been grazed by bison since 1992 while the other has been protected from grazing by both bison and cattle since 1981, and both are burned annually (Vinton et al. 1993, Hartnett et al. 1996, Collins et al. 1998, Collins and Steinauer 1998, Knapp et al. 1999). We will use data from 20 plots (four transects) located in the lowland areas of each watershed.

References


**STUDENT INSTRUCTIONS**

Before beginning
Make sure that your version of Microsoft Excel has the diversity plug-in available free from the University of Reading in the United Kingdom. Go to this website:  
http://www.rdg.ac.uk/ssc/software/diversity/diversity.html and follow these directions:

1. Click on **Get the Diversity Add-In now**.
2. In the File Download window, click **Save**
3. In the Save As window, find a place to save the download where you will remember its location, and click **Save**.
4. Click **Open**.
5. Unzip the folder by selecting **Extract All** from the **File** menu.
6. Open a blank file in **Excel**.
7. From the **Tools** menu, select **Add-Ins…**
8. If the name of the **Diversity** module is not in the list of add-ins, click the **Browse** button and locate it in the folder in which you saved it.
9. Once you have found the **Diversity** module, make sure that the check-box against its name is checked, or simply double-click it.
10. The Diversity functions are now available for use. You can check by opening the function wizard by clicking on the $f$ key. At or near the bottom of the Function Category list, you should find a category called **User Defined**. Click on this and a list of the functions should appear in the panel on the right. The add-in should remain available in subsequent Excel sessions until you remove it.

Download the Konza data set from the TIEE website.
Part 1: Introduction: What is biodiversity?

In the broadest sense, biodiversity encompasses all the variety of life forms on Earth. Humans rely upon the earth’s biodiversity resources for food, fiber, fuel, building materials, medicine, and technological advances. For example, a large percentage of medicines are derived from biochemicals found in plants, animals or microbes. The enzyme that allows biologists to replicate DNA quickly and easily was discovered in a bacterium, *Thermus aquaticus*, living in the hot springs of Yellowstone National Park. Furthermore, plants and algae are primarily responsible for maintaining oxygen in the earth’s atmosphere, as well as supporting organisms that rely on their photosynthetic abilities.

We do not know exactly how many species live on Earth with us. One educated guess is about 14 million (of which only about 1.7 million have been described), and that is likely an underestimate (Adams et al. 2000). Of course, this number is changing as new species are discovered and others disappear. While it has been estimated that one new species evolves every three years, it has also been estimated that around 3 species per hour go extinct, many due to human activities (Magurran 2004). The need for conservation in recent years has made biodiversity a household word.

In order to conserve biodiversity, we first need to know how to quantify it. In addition, we need a way to compare how biodiversity changes through time and differs from place to place. Ecologists are often interested in the effects of disturbance, introduced species, land management practices, and other human activities on biodiversity. Biodiversity can refer to diversity at any biological scale from genetic diversity to ecosystem diversity. This exercise focuses on species diversity, specifically on how grazing affects plant species diversity. But how do we measure species diversity? Is it just a count of the total number of species in a given area, or are there other variables important to consider as well?

The objectives of this exercise are 1) to learn different ways that biologists measure species diversity and 2) to apply these new tools to answer the question, “How does grazing, a common land management practice across grasslands worldwide, affect plant diversity?”
Introduction activity: Measuring biodiversity

Question 1. Examine Table 1. This table contains two species lists from similar sites. Given the information in Table 1, which site do you think has higher plant diversity and why?

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andropogon gerardii</td>
<td>Andropogon gerardii</td>
</tr>
<tr>
<td>Schizachyrium scoparium</td>
<td>Schizachyrium scoparium</td>
</tr>
<tr>
<td>Bouteloua curtipendula</td>
<td>Bouteloua curtipendula</td>
</tr>
<tr>
<td>Eragrostis spectabilis</td>
<td>Eragrostis spectabilis</td>
</tr>
<tr>
<td>Koeleria macrantha</td>
<td>Koeleria macrantha</td>
</tr>
<tr>
<td>Dichanthelium oligosanthes</td>
<td>Dichanthelium oligosanthes</td>
</tr>
<tr>
<td>Panicum virgatum</td>
<td>Poa pratensis</td>
</tr>
<tr>
<td>Poa pratensis</td>
<td>Sorghastrum nutans</td>
</tr>
<tr>
<td>Sorghastrum nutans</td>
<td>Carex meadii</td>
</tr>
<tr>
<td>Carex meadii</td>
<td>Amorpha canescens</td>
</tr>
<tr>
<td>Amorpha canescens</td>
<td>Asclepias viridis</td>
</tr>
<tr>
<td>Symphyotrichum ericoides</td>
<td>Baptisia australis</td>
</tr>
<tr>
<td>Baptisia australis</td>
<td>Lespedeza capitata</td>
</tr>
<tr>
<td>Lespedeza capitata</td>
<td>Oxalis stricta</td>
</tr>
<tr>
<td>Oxalis stricta</td>
<td>Solidago canadensis</td>
</tr>
</tbody>
</table>

Question 2. Table 2 describes the same two sites, but includes some further information about the plant community: the abundance of each species. Abundance is the number of individuals of each species in the area. For instance, imagine walking through these plant communities. In which community are you more likely to actually encounter more species given the abundance of each? Using the information presented in Table 2, which community do you now think is more diverse? Does the new information change your assessment and why?
### Table 2. Comparison of plant biodiversity at two sites.

<table>
<thead>
<tr>
<th></th>
<th>Site 1</th>
<th>Abundance (# of individuals)</th>
<th>Site 1</th>
<th>Abundance (# of individuals)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andropogon gerardii</td>
<td>90</td>
<td>Andropogon gerardii</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Schizachyrium scoparium</td>
<td>15</td>
<td>Schizachyrium scoparium</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Bouteloua curtipendula</td>
<td>2</td>
<td>Bouteloua curtipendula</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Eragrostis spectabilis</td>
<td>1</td>
<td>Eragrostis spectabilis</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Koeleria macrantha</td>
<td>1</td>
<td>Koeleria macrantha</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Dichanthelium oligosanthes</td>
<td>1</td>
<td>Dichanthelium oligosanthes</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Panicum virgatum</td>
<td>1</td>
<td>Poa pratensis</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Poa pratensis</td>
<td>1</td>
<td>Sorghastrum nutans</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Sorghastrum nutans</td>
<td>1</td>
<td>Carex meadii</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Carex meadii</td>
<td>1</td>
<td>Amorpha canescens</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Amorpha canescens</td>
<td>1</td>
<td>Asclepias viridis</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Symphyotrichum ericoidees</td>
<td>1</td>
<td>Baptisia australis</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Baptisia australis</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lespedeza capitata</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxalis stricta</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solidago canadensis</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Question 3.** In your lab groups, discuss your answers to questions 1 and 2. Based upon your discussion, what do you think are the characteristics of a diverse community?

**Conclusions: Introduction**

Measuring species diversity is based upon two ideas: species richness and species evenness. Species richness refers to the total number of species (e.g., data that were presented in Table 1). Species evenness measures how equally represented the species are, in other words, do all of the species have equal abundances (e.g., Site 2 in Table 2) or are they quite skewed with a few being very abundant and others rare (e.g., Site 1 in Table 2)?

Sometimes biologists are simply interested in species richness and use this alone as a measure of biodiversity. Using the species richness data provided in Table 1, it is logical to conclude that Site 1 is more diverse because it has higher species richness (greater number of species). Often, however, biologists use a combination of species richness
and species evenness to calculate what is called a **diversity index**. Using the data from Table 2, it is logical to conclude that Site 2 is more diverse. Even though it contains four fewer species than Site 1, if you were walking through these two communities, you would likely encounter more species in Site 2 because it has higher evenness. Most likely in Site 1, you would only encounter *Andropogon gerardii* (big bluestem) and perhaps also *Schizachyrium scoparium* (little bluestem) because they are relatively high in abundance while the other species are each represented by only one individual. In the activities that follow, we will first explore species richness as a measure of biodiversity, and then we will explore diversity indices as measures of biodiversity. We will use these measures of biodiversity to answer the question of how bison grazing affects plant diversity. Lastly in Part 3, we will consider the biases of each biodiversity measure.

**Part 2: Activities**

Using data from the Konza Prairie LTER, we will explore three concepts: species richness, sample size, and diversity. Within each concept, there will be an activity in which you will learn a new technique for measuring biodiversity, analyze data using this technique, and synthesize the data in a table or graph. You will then be asked to interpret your results and apply them to answer our broader ecological question, “How does grazing affect plant diversity?”

**Concept 1: Exploring species richness**

By counting the number of species (species richness) in a given sample we can address the question, “How diverse is this area,” in a common sense, straightforward way.” Often when people talk about conserving biodiversity, what they mean is maximizing the number of species that live in a preserve or wildlife refuge. Species richness is also easy to interpret, which is not always the case with ecological indices, as we’ll see later.

We will now use species richness to compare the effects of bison grazing on plant diversity in tallgrass prairie using data from the Konza Prairie LTER site in Kansas.

**Activity 1: Species richness**

The Activity 1 spreadsheet in the Excel file contains plant species composition data from two different watersheds on Konza Prairie in 2004. One watershed is grazed by a bison herd that has been present for 14 years (Bison-Present), and one watershed has been protected from grazing for 25 years (Bison-Absent). These are considered two different experimental treatments.

**Directions**

1. Open the Konza Diversity Exercise Excel workbook. Select the Activity 1 worksheet by clicking on the tab in the lower left-hand corner labeled Activity 1. Notice that within this worksheet there are six different data subsets for each watershed. The data for the Bison-Absent watershed are highlighted in yellow, and the data for the Bison-Present watershed are highlighted in green. Each data subset contains the plant species present in that watershed and is based on a number of 10 m² plots that were sampled (refer back to Figure 2 for an illustration of the sample plots). For instance, some subsets were
based on 11 sampled plots, some on 16 plots, and some on other numbers of plots. The number of plots sampled is indicated at the top of each subset as Sample Size, or N. In the plant species lists, each species is identified by a numerical code (unique to each species) in one column and the genus, species and sometimes variety in the next column. Species are listed if an individual appears in any one of the N sampled plots per data subset. Thus, these are presence/absence data sets.

2. Calculate the species richness for each data set for each of the two watersheds.

It may first appear that the quickest way to determine the number of species present is to just count the number of rows by hand; however, learning to use the new “NumSpec” (number of species) function will be a much easier and faster way to do the more complicated calculations later in the exercise.

To use the “NumSpec” function to calculate species richness:
   a. Select cell A59, which should be the first white cell located under the species code column of the first data subset.
   b. Under the Insert menu, select Function.
   c. In the Function category box, select User Defined.
   d. In the Function name box, select Numspec for number of species (i.e., species richness)
   e. Click ok.
   f. Using your cursor, select the cells containing all of the species codes in the first data subset.
   g. Click ok and the species richness value should now be displayed in cell A59.
   h. The correct result is 40.
   h. Repeat steps a-g for all twelve data subsets.

3. Complete Table 3 with your results. Be sure to give the table a descriptive title.

<table>
<thead>
<tr>
<th>Data subset</th>
<th>Bison-Absent</th>
<th>Bison-Present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Species richness</td>
<td>Sample size</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Questions for Activity 1

Question 1. If you refer only to data subset 1 in both treatments, what would you conclude about the effect of bison grazing on species richness? How do your conclusions change if you compare only data subsets 3? Only data subsets 6?

Question 2. Recall that all of the data within a treatment (Bison-Present or Bison-Absent) were collected from the same watershed. How is it possible to reach different conclusions in Question 1 from the different subset comparisons? What could account for the differences in species richness among plots within a treatment?

Question 3. Do you feel that you have enough information to accurately characterize the true species richness for each watershed? Why/why not? State your best estimate of species richness for each treatment, given only the information in Table 3.

Concept 2: Exploring sample size or When is it enough?
In Activity 1, you discovered the biggest problem with measuring and comparing species richness: sample size has a very large effect on the estimate of species richness. Clearly, the more sampling we do, the more accurate our estimate of species richness can be. Yet it’s not practical to examine every single plant in the entire watershed. We must rely on smaller samples to represent the whole. How do we know when we’ve sampled ‘enough’, that is, when our sampling effort is a good representation of the whole?

Figure 3. An example of a species accumulation curve. Sampling is adequate if the curve approaches an asymptote, which is the inflection point or where the line begins to become level. The asymptote is also the estimate of species richness.
To solve this sampling problem, ecologists often use a tool called the **species accumulation curve** to estimate the species richness of a given area of interest (see Figure 3). These curves plot the cumulative number of species encountered (shown on the **y**-axis) against the cumulative number of samples taken, or the **sampling effort** (shown on the **x**-axis). If additional sampling effort does not result in new species being found (where the curve levels off), then it is likely that sampling has been adequate to estimate species richness.

The biologists that collected the data shown in Figure 3 appear to have sampled adequately because the curve approaches an **asymptote**, in other words, the curve levels off with increased sampling effort. The best estimate of species richness of this plant community is 102, shown by the point where the curve approaches the asymptote.

**Activity 2: Species accumulation curves**
Next you will create two species accumulation curves: one for the Bison-Absent watershed (data listed in the Activity 2a spreadsheet) and one for the Bison-Present watershed (data listed in the Activity 2b spreadsheet).

**Directions**
1. The Activity 2a spreadsheet contains plant species composition data for the watershed without bison. The Activity 2b spreadsheet contains the plant species composition data for the watershed with bison. Each sheet has 20 subsets of data. The first subset contains the species present in the first plot, the second subset contains the total species present in the first and second plots combined, the third contains the species present in the first, second, and third plots combined, etc. With the data arranged this way, we can calculate the cumulative species richness as we add an additional plot one at a time. These are the data needed to create a species accumulation curve.

2. Calculate the species richness for each subset in the Activity 2a and 2b spreadsheets. The species richness values should be in row 59 of Activity 2a and in row 99 of Activity 2b in the cells with black borders. Refer to the instructions under “Before Beginning” if you need additional help. Here is where the “NumSpec” function will really speed things up.

3. Plot the cumulative species richness against sampling effort (number of plots). Note that at the right of all of the data subsets there are two columns labeled “Number of plots” and “Species richness.” The first column contains the number of plots (1, 2, 3, ..., 20). The second column will contain the species richness values you calculated in step 2.

   To get these values in column form from row 59, **CTRL click** on each value, and click **Copy** in the Edit menu. Move the cursor an empty column cell, click **Paste Special** in the **Edit** menu, mark **Values** and **Transpose**, then **OK**. Now you...
should have two columns of numbers: number of plots and a corresponding species richness value.

In order to create the graph:

a. **Highlight** the column of species richness values. Do not highlight the column title.
b. Click on the **Chart Wizard** icon in the tool bar.
c. Under chart type, select **Line**. Several Chart sub-types are available. The default option is the one we want to use for this graph. Click **Next**.
d. The next screen shows a Preview of your graph. If everything looks correct, click **Next**. If things are not correct, re-select the data you wish to graph.
e. The next screen allows you to add important information to your graph. Click the **Titles** tab. Add **labels** for both axes (x-axis should be “Number of plots”) and a **title** indicating the treatment.
f. Click the **Legend** tab. Because we have only a single line on our graph, we do not need a legend. Uncheck the **Show legend** box. Click **Next**.
g. The last window asks if you wish to display the chart in a new worksheet or as an object in the current worksheet. Choose **As an object** and click **Finish**.
h. If you wish to move your graph, use click and drag.
i. Repeat steps a-h with Activity 2b. Recall that your species richness values should be in row 99.

**Questions for Activity 2**

**Question 4.** Were an adequate number of samples taken? Provide evidence to support your claim.

**Question 5.** Can we estimate species richness from these species accumulation curves? Why or why not? Support your answer with the data and figures you’ve just examined.

**Question 6.** Using the new information obtained from the species accumulation curves, how would you answer our original question: What is the effect of bison grazing on plant species richness?

**Conclusions: Species richness and sample size**

Species richness is intuitively easy to interpret and is often a central concern of conservationists. While it does not necessarily give the whole picture of diversity (remember, diversity includes both richness and evenness!) it can be very useful in many situations.

Species accumulation curves are only one of the tools available to biologists to estimate species richness. In the mid 1990’s, significant advances occurred in what are called
nonparametric estimators of species richness that use mathematical formulae rather than relying on interpretation of plots to calculate species richness. For instance, Magurran (2004) discusses several of these estimators and their appropriate use for different applications. There are specific computer programs that biologists use to calculate these estimates. Although we will not explore them further in this exercise, such estimators are well described in Magurran (2004).

**Question 7.** When might species richness be the primary diversity measure of interest? What cautions do biologists need to keep in mind when interpreting species richness?

**Concept 3: Exploring diversity**
Recall that measuring species diversity is based upon two ideas: species richness and species evenness. Now that we know how to measure species richness and how to interpret these measures, we will include evenness to our diversity measures. Although many indices have been developed that incorporate both species richness and species evenness into a single measure of diversity, we will explore two of the most common, the Shannon Index and the Simpson Index. But first we need to understand how to measure evenness and relative abundance.

**Evenness and relative abundance**
Refer back to Table 2, and recall that evenness measures how equally abundant the species in a community are. For example, a community with nearly equal abundances of all species, such as in Site 2 of Table 2, would be considered extremely even. Therefore, in order to measure evenness, we need to know the abundance of each species present in the community and how that abundance compares with all the other species in the community. The number of individuals is an obvious and common measure of abundance and should be used when an individual is readily distinguishable, such as with birds, mammals, insects, trees, and annual plants. Several organisms, such as clonal plants, fungi, and bacteria can pose difficulties in discerning where one individual ends and another begins. In such cases, biomass, frequency, cover, and cover class are other measures of abundance. To understand how abundant a species is in a given community, we need to compare it to the rest of the community, and we can do this by using relative abundance. To calculate a relative abundance value for a given species, we take the abundance of the given species and divide it by the total abundance of all species in the community of interest. In our case we will use frequency as our abundance measure to address the question: what is the effect of bison grazing on plant diversity in tallgrass prairie?

Frequency is defined as the proportion of total sample plots in which that species was found. For instance, if *Lespedeza capitata* (round headed bush clover) were found in one-fourth of the plots sampled, its frequency would measure 0.25. The Excel spreadsheet for Activity 3 already contains the frequency values for each species found in each treatment.

To calculate relative frequency for our Konza data, we will take the frequency of each species and divide that by the sum of the frequencies for all of the species in the community. Using the example from above to calculate relative frequency for *L. capitata*, we would take 0.25/(sum of the frequencies calculated for all species in the community). We will calculate the relative frequencies in this way in our Excel spreadsheet.

**Shannon Diversity Index**

The Shannon Index uses relative abundance data (in our case, we will use relative frequency) to incorporate species evenness and species richness into a single measure of diversity, represented by $H'$. The Shannon Diversity Index typically falls between 1.5 and 4.0, with lower values indicating lower diversity, and higher values indicating higher diversity.

Here is the equation for the Shannon Index:

$$H' = - \sum p_i \ln(p_i)$$

We’ll build this equation up, starting with the term on the far right:

- $p_i$ is the relative abundance of each species $i$, as explained above.
- $\ln(p_i)$ is the natural logarithm of the relative abundance.
- $p_i \ln(p_i)$ is the relative abundance of species $i$, multiplied by the natural logarithm of the relative abundance ($p_i$). We calculate this product for every species in the community.
- $\sum p_i \ln(p_i)$ is the sum of the $p_i \ln(p_i)$ products that we calculated for every species in the community.
- $-\sum p_i \ln(p_i)$ is the negative sign of the sum that we calculate. This is necessary because taking the natural log gives us a negative number. Positive numbers are much easier to interpret. This gives us $H'$, the Shannon Diversity Index.

One could easily use a calculator or spreadsheet to figure each of these steps. But we are going to take advantage of a program that has already been written to do these calculations in Excel in one step. Mathematically it is a measure of the average degree of “uncertainty” in predicting the identity (species) of an individual drawn randomly from the community.

**Activity 3: Shannon Diversity Index**

**Directions**

1. Go to the Activity 3 spreadsheet in the Konza Diversity Exercise Excel workbook. These data were taken from the watershed with bison present. Notice that there are four data subsets within this spreadsheet. Each subset contains the species present in that watershed and their abundance measured by frequency.
2. Calculate the species richness for each subset using the `Numspec` function (see Activity 1 for review).

3. Calculate the Shannon Index for the first data subset, N = 4.
   a. Select an empty cell below the frequency data for the first subset.
   b. Under the Insert menu, select `Function`.
   c. In the Function category box, select `User Defined`.
   d. In the Function name box, select `Shannon`.
   e. Click OK.
   f. Select the cells containing all of the frequency data in the first data subset.
   g. Click OK and the Shannon Diversity Index value should now be displayed in the cell. The answer for data subset 1 should be approximately 3.705.
   h. Repeat steps a-g for all four data subsets.

3. Complete the first three columns of Table 5 with your results (the last column, D, will be filled in later). Again, give the table a descriptive title.

### Table 5.

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>N</th>
<th>H'</th>
<th>D</th>
</tr>
</thead>
</table>

S = species richness, N = sample size, H' = Shannon Index, D = Simpson Index

**Simpson Index**

A potential drawback of the Shannon Diversity Index is that it is very sensitive to species richness. That is, if an additional species is added, the Shannon Index rises disproportionately, especially if the new species is not abundant. A different measure, the **Simpson Diversity Index**, is heavily weighted by abundances of the most common species, and produces a measure that is less sensitive to species richness. For instance, we would likely consider a community *truly* more diverse if it had equal abundances of 25 species, than if it had one very dominant species, and 24 extremely rare species. The Simpson Index, represented as D, gives more weight to the relative abundance of species in a community in that it is based on the probability that two individuals drawn from an infinitely large community are the same species. This is the equation:

\[ D = \sum p_i^2 \]

- \( p_i \) is the relative abundance of species \( i \) (and this is then squared).
• $\sum p_i^2$ is the sum of the values that we calculated for every species in the community.

As $D$ increases, diversity decreases, which is rather counter-intuitive. For the value to be more intuitive, the Simpson Index is usually expressed as 1-D or 1/D, so that the value actually rises as the community becomes more diverse. We can use the Excel spreadsheet to calculate the Simpson Index in a few easy steps. Although the diversity plug-in we installed does calculate the Simpson Index, it does so using a different measure of abundance than we are using, so we will program the formula ourselves in Excel to take into account the measure of abundance taken at Konza Prairie, that is, frequency.

**Activity 3: Simpson Diversity Index**

**Directions**

To calculate the Simpson Index for each data set:

1. Open the Activity 3 spreadsheet again to calculate the Simpson Index. Place your cursor in the cell next to the first frequency data point in the empty column of first data set (cell D5). This will become your formula cell.
2. Type = (3. Click on the first frequency data point. You will see that the reference for this cell shows up in the formula cell so that it should now read = (C5
4. Within the formula cell, type / (sum(5. Using your cursor, select the entire column of frequency values.
6. Your formula should now look like this: = (C5 / (SUM($C$5:$C$49)
7. Add $ as shown here: $= (C5/(SUM($C$5:$C$49))
8. Type ^2 and hit enter
9. Your formula should read: =(C5/(SUM($C$5:$C$49)))^2 and the correct result is 4.5E-05.
10. Fill this formula all the way down this column until you reach row 49. One way to do this is to highlight the entire column between row 5 and 49. Under the Edit menu, select Fill and then Down.
11. Place your cursor in the first empty cell under the column of numbers you just created (cell D50).
12. Type the following formula: = 1 – sum(13. Select all of the data just above this cell (cells D5:D49)
14. Within the formula cell, type ) and hit Enter. The entire formula should read =1-sum(D5:D49)
15. The value in this cell is the Simpson Index and should be approximately 0.97 for the first data subset.
16. Repeat steps 1– 15 for the remaining 3 data subsets on the worksheet.

Complete the last column in Table 5 with your results.
Question 8. How does species richness affect the Shannon and Simpson Diversity Indices? Given what you already know about how sample size affects the calculation of species richness, how do you think sample size would affect the Shannon Index? The Simpson Index?

Activity 3: Comparison

Now that we know something about how to calculate diversity indices as well as a little bit about their pitfalls, let us return again to our original question: “How does the presence of bison affect plant species diversity?” Calculate the Simpson Index using data provided in the Comparison spreadsheet. Refer to step 2 above for directions.

Question 9. What is the effect of grazing on tallgrass prairie diversity? Compare your answer with the conclusions you drew based upon species richness alone. Why are your conclusions similar or different?

Question 10. When might a diversity index (instead of just species richness) be the primary diversity measure of interest? Why?

Conclusions: Diversity indices
Despite the fact that several studies have shown that the Simpson Diversity Index performs much better, the Shannon Diversity Index is the most widely used. In fact, Magurran (2004) states, “The Simpson’s index is one of the most meaningful and robust diversity measures available.” Further, she points out the serious shortcomings of the Shannon Index, including its sensitivity to species richness, and thus, sample size. One benefit to using the Shannon index is that it allows a researcher to compare their results to past studies. While the Simpson Diversity Index is often a better choice, it is not a panacea. All diversity indices, including the Simpson Diversity Index and the Shannon Diversity Index, are based on some weighing of relative abundance and species richness, but each index weighs these measures differently. For this reason, scientists often calculate several indices for their data, or choose one index carefully based on what they know about their study system and the assumptions made by a given index.

Part 3: Assumptions and Conclusions
In her classic book on measuring biodiversity, Dr. Anne Magurran (2004) lists three assumptions of biodiversity measurement.

1. All species are equal. In other words, rare or endangered species are no more important than the most common species in the community.
2. All individuals are equal. In other words, the largest redwood in the forest is given equal weight with the tiniest redwood seedling.

3. Species abundances must be measured appropriately. In addition, it is important to use the same abundance measure at sites or times that you wish to compare.

**Question 11.** Why is it important to understand the assumptions behind biodiversity measurements? What information would you need to evaluate whether or not the assumptions had been met or are appropriate?

**Assumptions Specific to Shannon and Simpson**
In addition to the general assumptions listed above, the Shannon and Simpson Indices have a few more assumptions associated with them.

**Shannon Index**
Assumption 1: Individuals are randomly sampled from an infinitely large community.

Assumption 2: All species are represented in your data set. In other words, if a species is present in a community, you sampled it. If you didn’t encounter the species, it doesn’t occur in the community.

**Simpson Index**
Assumption 1: Individuals are randomly sampled from an infinitely large community.

There is a version of the Simpson Index that does not require Assumption 1. If the researcher feels that this assumption cannot be met, the modified Simpson Index can be used. As you can imagine, the modified Simpson Index is a bit more computationally complex, which is why we will not go over it for this exercise. There is no equivalent correction for the Shannon Index.

**Question 12.** Recall that the Shannon Index has been widely used. How often do you think the assumptions for using the Shannon Index have been met? The Simpson Index?

**Conclusions**
By comparing just these two watersheds (Bison-Present and Bison-Absent), our data indicate that bison grazing may increase plant diversity. Long-term research at Konza Prairie substantiates this indication. Statistical analysis conducted on data from many replicate watersheds of these treatments over time has shown that bison have two main effects on the tallgrass prairie plant community that serve to increase plant diversity. First, bison create a more heterogeneous environment by wallowing, by creating grazing lawns, and by excreting wastes, opening more sites for different types of plants to
colonize and grow, increasing species richness (Hartnett et al. 1996). Second, bison prefer to graze on the major dominant grasses, such as *A. gerardii* (big bluestem). By grazing down the dominants, bison create a competitive release of the sub-dominant vegetation allowing plants to flourish that may otherwise be out-competed by big bluestem if the bison were not present, often leading to increased species evenness (Knapp et al. 1999).

We have practiced two main tools that biologists use to quantify species diversity:

1. **species richness** or simply counting the number of species present in an area
2. **species diversity** that incorporate species richness with information about the abundance of each species in the community.

In both cases, we have explored the influence of **sample size** on the measurement of species diversity.

Conserving biodiversity is an important goal, but through the activities presented here, we have seen that it is not always straightforward to measure what we wish to conserve. Applying the appropriate tools to measure biodiversity is an important step toward our goal of conservation as well as understanding patterns of biodiversity loss in communities.

**FACULTY NOTES**

**Objectives and audience**

The objective of these activities is to give students experience calculating species richness and diversity indices as well as to have the students critically evaluate and interpret the calculations. In particular, they should think about what factors may affect the indices, especially sample size and the initial assumptions of each calculation. At the same time, students are given real data to use for their analysis and a real ecological question to address with their new skills. The student instructions are written for introductory to intermediate level students. The instructions assume that students have little or no experience with Excel. Students with more experience with Excel will likely take less time to complete the activities.

**Suggestions for using this dataset**

This dataset is divided into four parts: Overview, Part 1: Introduction, Part 2: Activities and Part 3: Assumptions and Conclusions.

The Overview contains reading material that contains background information on Konza Prairie, the study design and data collection.
Part 1: Introduction contains a short description of biodiversity, species diversity, and the ecological question. It also contains a short Activity that asks students to compare two communities given a list of species in each community, with or without information on abundance. The purpose of the Activity is to get students thinking about the two main components of species diversity (richness and evenness) using a student-active approach of “writing before discussion.” This activity could also be structured using the “Think-Pair-Share” method. More description of these methods can be found on the TIEE website: http://tiee.ecoed.net/teach/teach_glossary.html#studentactive. Time for the introduction activity (if it is all completed within class time, see suggestion below for another option), including some discussion time, is about 20 minutes.

**Suggestion:** The Overview and Introduction can be assigned as pre-class readings. In addition, we suggest that students begin the Introduction Activity before coming to class by writing answers to Questions 1 and 2. Question 3 asks the students to discuss their answers to Questions 1 and 2 in small groups (~5 minutes of discussion should be appropriate) and this discussion may be a great way to start off the class. Structuring the class this way will also allow for maximum computer time in class.

Part 2: Activities contains the computer work that students will complete and is divided into three activities that altogether will require approximately 4 hours, largely depending upon student familiarity with Excel. One to three questions are posed after each activity. Some simply ask the students to compare their results, while other questions require more thought and analysis. The questions should be considered during class (rather than having the students go back to the questions after completing all of the computer work) as they help to focus student attention throughout the exercise.

**Suggestion:** The activities should be done in order, but one option is to spread the activities over two classes, with students completing activities 1 and 2 in one class period and activity 3 in another. Another option is that if students have enough class time to become comfortable with Excel and the format of the dataset, they can finish the activities on their own outside of class.

Part 3: Assumptions and Conclusions does not require any computer-based data calculations. Rather, it challenges students to consider the assumptions of the methods they have learned and how these assumptions may affect the conclusions they have drawn. It should take approximately 20 minutes.

**Suggestion:** Part 3 can be assigned as an outside-of-class activity to reinforce what the students have learned in class.

**Logistics**
In order to allow maximum class time for completing the exercise, we strongly suggest that instructors make sure that the diversity plug-in is installed in the copy of Microsoft Excel the students will be working with. The plug-in is available free from the University
of Reading in the United Kingdom. Instructions for downloading and installing the plug-in are given in the “Before beginning” section of the student instructions.

We decided to use the diversity plug-in to speed up the calculation process for species richness and for the Shannon Index so that students hopefully would focus on interpreting the data rather than getting bogged down with calculating and using Excel. However, the Simpson Index algorithm used by the plug-in is not compatible with the frequency measurement of abundance we use in these activities; therefore, the activity walks students though calculating the Simpson Index by programming formulae in Excel. Having students use an already programmed formula for the Shannon Index and then programming their own formulae for the Simpson Index allows them to experience both methods and saves time as well. Be aware, however, that some students may notice that Simpson is an option with the diversity plug-in, and if they use it instead of following the instructions, they will not calculate the correct answer.

Another issue with the Simpson index is the counter-intuitive nature of its initial calculation for D, and why using the formula 1-D is preferable because it is more intuitive. Faculty may wish to illustrate this point with an example during class.

One important issue this dataset does not consider is replication and statistical comparison of the treatments. While students compare the effect of sample size (i.e. number of plots sampled for species identification) on species diversity measurement, they are still only comparing a single grazed watershed with a single control watershed. We chose to do this for simplicity. As it is, there is a substantial amount of data for students to deal with and many concepts to master without replication. For more advanced students, the instructor may wish to introduce the idea of treatment replication and discuss how it would affect the comparison of the grazed and control treatments. The exercise could be expanded over more class periods by obtaining data for more watersheds and/or more years from the Konza Prairie LTER website. Students would then be able to fully address the ecological question posed in this exercise and see if they obtain results similar to those already published in the cited literature. Deeper analysis of more data may also be an appropriate independent project for more advanced students.

Notes on the Excel spreadsheets
There are seven spreadsheets in the Faculty Excel Workbook for this exercise. The spreadsheets are labeled Activity 1, Activity 2a, Activity 2b, Activity 3, Comparison, Table 3 and Table 5. Descriptions of the data found in each spreadsheet follow below.

*Activity 1:* This worksheet contains a total of 12 data subsets, including six data subsets for the Bison-Absent watershed (highlighted yellow in the worksheet) and six data subsets for the Bison-Present watershed (highlighted green in the worksheet). Each data subset is labeled 1 through 6 and has a sample size that corresponds to it. At the bottom of each data subset, just below the highlighted cells, we’ve outlined a cell in black to
indicate where students should do their calculations. The answers for faculty are summarized in the Table 3 worksheet.

**Activity 2a:** This worksheet contains the data required to create a species accumulation curve for the Bison-Absent watershed. It contains 20 subsets of data, and each is labeled with its sample size (number of plots that were sampled). At the bottom of each data subset, just below the highlighted cells, we’ve outlined a cell in black to indicate where students should do their calculations. To the far right of the 20 data subsets, we’ve created two columns where students will copy and paste their results. The student instructions walk them through the copy and paste techniques. In the faculty workbook, the answers are already available and a sample graph has been created just below the two columns of data (columns BL and BJ).

**Activity 2b:** This worksheet is set up just like the Activity 2a worksheet and contains the data required to create a species accumulation curve for the Bison-Present watershed.

**Activity 3:** This worksheet contains four data subsets all from the Bison-Present watershed. Students use these data to compare the effects of species richness on Shannon and Simpson diversity indices. Calculations are present in the faculty workbook and the answers are also summarized for faculty in the Table 5 worksheet.

**Comparison:** This worksheet contains the entire species lists and frequency values for the Bison-Absent (highlighted yellow) and Bison-Present (highlighted green) watersheds. The activity asks students to use these data to calculate Simpson diversity indices for each watershed and draw conclusions. The correct values are in bold at the bottom of the data lists.

**Table 3:** The student version of the Excel workbook does not contain this spreadsheet. This spreadsheet contains the correct values for Table 3, which students fill out as part of Activity 1.

**Table 5:** The student version of the Excel workbook does not contain this spreadsheet. This spreadsheet contains the correct values for Table 5, which students fill out as part of Activity 3.
Assessment and Evaluation
Various options can be used for evaluation, and they range from assessing basic levels of understanding and comprehension to assessing higher levels of synthesis and application. How you choose to evaluate students depends on your learning goals for your class.

Completion of the Excel activities and calculations, and submission of the two tables and two graphs from Part 2 would document basic understanding of the material and calculations of indices, as well as an ability to use Excel. More complete comprehension of the concepts and their complexity could be assessed by evaluating answers to all or a subset of the 12 questions posed in Parts 2 and 3. Whether you decide to have students answer these questions individually, as a small or large group, or a combination of both will depend on the degree you choose to emphasize independent or collaborative work. Since doing science is dependent on working within a community of scientists and exchanging ideas, we highly recommend incorporating group work or discussion to some extent. If you evaluate written work, be clear about writing expectations (e.g., do you expect complete sentences, or are short lists acceptable?).

Further, a synthesis activity can be assigned. One option is to have students write a short essay whereby they choose the “best” measure of biodiversity and defend their positions. This assignment could be modified by assigning groups of students to a measure of biodiversity and asking that they formulate a defense of why it is “best”. The groups could then have an oral debate in class. The instructor may wish to define a specific situation, either using the one described in the exercise or one or several new situations. One advantage to defining several new situations is the opportunity to highlight that the “best” measure is context-dependent.

Another synthesis activity would be to ask students to write an explanation for a non-scientist friend or relative about how plant ecologists might measure plant species diversity of a given area, and how they would compare that with another area. This is based on the wisdom we all know that if you want to learn something well, try teaching it.

A sample of a basic rubric is included below, which can be modified and adjusted to your own needs. The number of points you wish to assign each category depends on what you choose to emphasize.
Sample Assessment Rubric

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| Basic understanding | • Able to use Excel data sets  
• Able to calculate indices using Excel  
• Turns in completed graphs and tables  
• Graphs and tables labeled appropriately |
| Comprehension       | • Answers all or a subset of the 12 questions  
• Answers in complete sentences (or whatever standards you have indicated), or through active group discussion  
• Answers show understanding and comprehension of the concepts and complexities |
| Group work          | • Asks questions when understanding is not clear  
• Listens to and is respectful of others  
• Shows evidence of applying learning to different situations  
• Is responsible for own tasks within the group in a timely manner |
| Synthesis           | • Is able to synthesize and apply concepts to new situation (according to assignment given)  
• Can articulate in own words the concepts covered |

RESOURCES

- ESA’s biodiversity fact sheet  

- Konza Prairie Biological Station  
  [http://www.konza.ksu.edu/konza](http://www.konza.ksu.edu/konza)

- Tallgrass plant species links:  
  [http://www.lib.ksu.edu/wildflower/](http://www.lib.ksu.edu/wildflower/)  
  [http://www.noble.org/imagegallery/](http://www.noble.org/imagegallery/)

- Other tallgrass prairie sites in the Great Plains:  
  o NPS Tallgrass Prairie National Preserve  
    [http://www.nps.gov/tapr/](http://www.nps.gov/tapr/)
  o Oklahoma Tallgrass Prairie Preserve  

- Tallgrass prairie photos  
  [http://www.naturalplanet.org](http://www.naturalplanet.org)
• Excel tutorials
  http://www.usd.edu/trio/tut/excel/
  http://www.fgcu.edu/support/office2000/excel/
  http://www.baycongroup.com/el0.htm

REFERENCES


COPYRIGHT STATEMENT

The Ecological Society of America (ESA) holds the copyright for TIEE Volume 3, and the authors retain the copyright for the content of individual contributions (although some text, figures, and data sets may bear further copyright notice). No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright owner. Use solely at one’s own institution with no intent for profit is excluded from the preceding copyright restriction, unless otherwise noted. Proper credit to this publication must be included in your lecture or laboratory course materials (print, electronic, or other means of reproduction) for each use.

To reiterate, you are welcome to download some or all of the material posted at this site for your use in your course(s), which does not include commercial uses for profit. Also, please be aware of the legal restrictions on copyright use for published materials posted at this site. We have obtained permission to use all copyrighted materials, data, figures, tables, images, etc. posted at this site solely for the uses described in the TIEE site.
Lastly, we request that you return your students’ and your comments on this activity to the TIEE Managing Editor (tieesubmissions@esa.org) for posting at this site.

GENERIC DISCLAIMER

Adult supervision is recommended when performing this lab activity. We also recommend that common sense and proper safety precautions be followed by all participants. No responsibility is implied or taken by the contributing author, the editors of this Volume, nor anyone associated with maintaining the TIEE web site, nor by their academic employers, nor by the Ecological Society of America for anyone who sustains injuries as a result of using the materials or ideas, or performing the procedures put forth at the TIEE web site, or in any printed materials that derive therefrom.