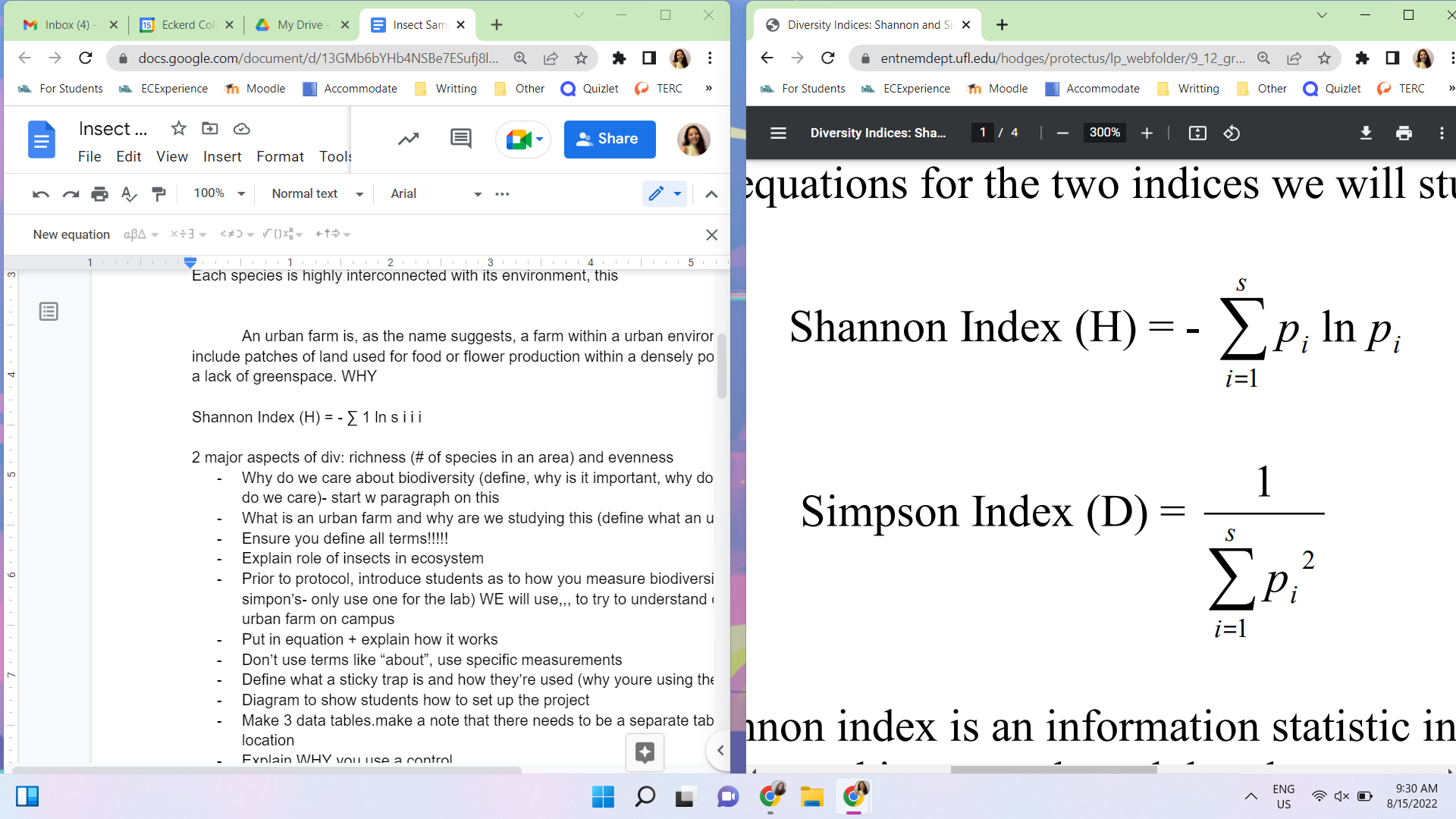
**A brief introduction (what is a diversity index and why do we use it in ecology)**

Eckerd’s urban farm provides us with a unique opportunity to better understand how urban agriculture affects insect populations’ presence. In this lab, we will be examining the differences in flying insect diversity and abundance between the orchard and farm section of Eckerd’s farm. Insects are important for an ecosystem’s success and have been called “biological foundation for all terrestrial ecosystems” (Scudder). This is because in their large numbers, they are able to very effectively maintain soil structure, polinate plants, and serve as a food source for other organisms. Their importance in a+n ecosystem’s biodiversity and health cannot be overstated. In it’s simplest terms, biodiversity is the variety of life in a particular ecosystem or habitat. In each of these ecosystems, all living things are highly dependent on each other- be it directly or indirectly. A healthy ecosystem is one that is diverse; one that has the ability to adapt and recover if a species has a population decline.

Greenspaces, such as urban farms, provide a habitat for many flying insects whose habitats have been otherwise destroyed by urbanization. This will, in turn, lead to a higher population of such species in these greenspaces. Pollinators, which are an integral part of agriculture, are expected to seek refuge in such areas (Sivakoff). Pollinators come in many different forms and sizes, from ants and bees to hummingbirds and bats, their contributions cannot be overstated. Pollinators are an integral part of all flowering plants’ lives, which includes many of the plants being grown on Eckerd’s urban farm (Celley). An urban farm is, as the name suggests, a farm within an urban environment. This would include patches of land used for food or flower production within a densely populated area with a lack of greenspace. The purpose of this study is to understand the diversity of insects on Eckerd’s campus and how the urban farm impacts that.

When conducting experiments in science, there is always an independent variable, dependent variable, and a control group. The independent variable is the condition being manipulated in the experiment. In our case, it is the presence of vegetation. The dependent variable is the factor that is being influenced by the change in the independent variable and is often what is being used to measure the change. For this lab, we are measuring insect biodiversity. Lastly, the control is the unchanged or unaffected variable. This lab uses a nearby rocky area to demonstrate an uncultivated landscape and will serve as a control for the study. A control is important because it allows us to confirm that the results from the study are due to the manipulated variable and not extraneous factors that we did not measure.

The **Shannon Diversity Index** **(H)** is a way to measure the diversity of species within a community. More specifically, it measures species evenness, which describes how similar species abundance is across the collected data. The index has a range of 0-1, with higher values indicating greater diversity. This equation assumes that all species are represented in a sample *and* that they are randomly sampled. In science we often stress the importance of accurate and thorough data, but this should exemplify why it is important. Incomplete or inaccurate data has the potential to give false results.



This is an equation with elements you’re likely familiar with. As a refresher, the **Σ** (sigma) indicates that you will sum the calculations of each of the species (denoted as **s** on top of the sigma). The **p** represents the proportion of individuals of one particular species found divided by the total number of individuals found. This is generally written as n/N when not in an equation. Lastly the natural log (**ln**) is a numerical algorithm whose properties, while interesting, are not integral to your understanding of this lab.

**Importing Data**

Install package:

>istall.packages(“vegan”)

Load the packages and data into your document. Remember to set your working directory before loading your data, otherwise it will raise an error.

>library(vegan)

>library(ggplot2)

>library(tidyr)

>data <- read.csv(“filename.csv”)

To ensure the file was loaded correctly, print it:

>print(data)

The library tidyverse has a function called pivot\_wider that increases the number of columns while decreasing the number of rows. We will use this to consolidate the data and make it look nice in one line of code.

>family\_data <- pivot\_wider(data, Family, names\_from = Collection, values\_from = Collection, values\_fn = length, values\_fill = 0)

>print(data)

The data should now be sorted into a *much* more readable chart.

**Shannon’s Diversity index and an evenness Index**

Diversity indices require numeric inputs, so we will be converting the current string data into numerics in the next few steps.

Starting this process, we will need to isolate the data from each location. For this tutorial, we still start with the rocky location:

>Rocky\_data <- data %>%

filter(Collection.Location %in% 'ROCKY')

>print(Rocky\_data)

Using the Rocky\_data, you will need to find the number of families present. It should look like this:

>Rocky\_fam <- Rocky\_data %>% count(Family)

>print(Rocky\_fam)

Now that we have numeric data, we can use that to populate the diversity indices.

The **shannon diversity index** **(H)** is a way to measure the diversity of species within a community. A higher value indicates a greater diversity of species. Using vegan, the diversity function defaults to the shannon index.

Calculating the Shannon Index:

>R\_Shannon <- diversity(data$ROCKY)

>print(R\_Shannon)

**What is the value of H in the rocky location?**

**Pielou’s evenness index** measures diversity along with species eveness. Eveness of species, or their equilbilty, refers to how similar the abundance of each species is. The function specnumber finds the number of species.

>R\_Pielou <- R\_Shannon/log(specnumber(data$ROCKY))

>print(R\_Pielou)

**What is the value of J?**

**Using the past few steps as a template, what are the H and J values of the other two locations?**

Now that we have all of the values for each location, we will put them into a matrix for easy viewing.

>SP\_Values <- data.frame(Index = c("Shannon's", "Pielou's"),

Rocky = c(R\_Shannon, R\_Pielou),

Orchard = c(O\_Shannon, O\_Pielou),

Farm = c(F\_Shannon, F\_Pielou))

>print(SP\_Values)

**Jaccard’s Index**

Another index commonly used by ecologists is Jaccard’s index. This compares two sets of data to see which members are shared and distinct.

We will start by using the same piovt\_wider function as before, this time using te insect orders instead of the familys.

>order\_data <- pivot\_wider(data, Order, names\_from = Collection, values\_from = Collection, values\_fn = length, values\_fill = 0)

>print(order\_data)

Now we will create a data frame from this data. We need to change the format because the function used only takes a specific type of input.

>order\_data <- data.frame(Rocky = order\_data$ROCKY,

Orchard = order\_data$ORCHARD,

Farm = order\_data$FARM)

>print(order\_data)

Using this new format, we can use the Vigdist function, which is used to calculate a number of indices, to calculate Jaccard’s index.

>vegdist(vals, method = "jaccard", binary = T)

**Graphing Insect Heights**

Using ggplot2, you will create a violin plot of insect heights. This should be familiar, so feel free to try this without the guide code.

>ggplot(data, aes(x=Order, y=Height)) +

geom\_boxplot()+

geom\_dotplot(binaxis='y',

stackdir='center',

dotsize = .5,

fill="blue")+

labs(title="Height at which Insects were Caught",

x = "Order",

y = "Hieght (cm)") +

theme\_bw()

**Post-Lab Questions**

1. Describe what an urban farm is and what benefits having one in our area may present. You should use outside sources for this.
2. In your own words, describe the difference between Shannon’s diversity insex and Pielou’s evenness index
3. What are some ways you could apply the Vegan package to other datasets?
4. What are some benefits to using a boxplot with this type of data?
5. Are there any functions or concepts you struggled with? How did you work through them?