Learning Outcomes: By the end of this activity, I will help you to:
1. Generate and test scientific hypotheses based on provided ecological background information and data
2. Analyze provided datasets utilizing descriptive statistics for center and spread
3. Construct at least two figures and at least one table that are clear and readable

Background:

Throughout this semester, we will explore the fundamentals of ecology by analyzing and interpreting data collected either by you or other ecologists. In this class, we will use Microsoft Excel as our primary software to manage, analyze, and visualize data. Do not worry if this is your first-time using Excel. The objective of this tutorial is to ensure that you have the computational skills necessary to begin working with these data. Even if you have used Excel before, you may be surprised about the software's quantitative depth and the new functions you will learn.

In this exercise, we will manage, analyze, and visualize data from a study of white-tailed deer (*Odocoileus virginianus*) harvested from two different counties in Florida – Hillsborough and Pasco which differ in their land-use, land-cover as well as population size (Table 1; Figure 1). In this study, deer were tagged and monitored throughout their lives to collect information on their life histories (Figure 2). Four variables were recorded on deceased individuals: body weight (kg); brain weight (g); age at death (years); and gestation length (days).

Table 1: General demographic information about the two counties surveyed.

County	**Hillsborough County**	**Pasco County**
Human Population	1.4 million	530,000
% Natural	16	30
% Urban/Residential	59	23
% Agriculture	25	47
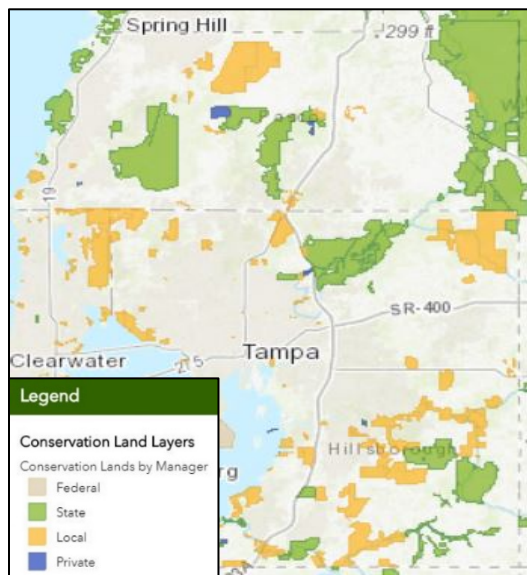
Figure 1 (right): A male white-tailed deer (Odocoileus virginianus); Figure 2 (left): Park areas in Pasco and Hillsborough County. Data provided from Florida Natural Lands Inventory

Thought Questions: (to be completed based ONLY on background information provided, i.e. before any analysis of the actual data)

1. Why might we expect populations of deer within the two counties to differ in the life history data collected?
2. Why might we expect populations of deer within the two counties to be similar in the life history data collected?
3. How might we visualize the data collected and what might each of those plots tell us about the data?
4. From these questions, generate at least two hypotheses about the data that you can later test.

Excel Tutorial Instructions

Open the Excel file titled "PCB3043L_Excel_Tutorial_Dataset_Students" and save it to your local or removable drive. Remember to Save (CTRL + S) often. If you make a mistake or accidentally delete something, use the Undo (CTRL + Z) feature.

Before we get started, we need to make sure you have the "Data Analysis Tool-pack" already downloaded in Excel. You can find instructions for both PCs and Macs here.

Part I: Descriptive Statistics

Remember, descriptive statistics are useful for understanding the **center** (mean, median) and **spread** (standard deviation, standard error, min, max) of the data. In this case, there are two populations for which we need descriptive data for. **The descriptive statistics you decide to calculate should relate to the hypotheses you created during the thought questions exercise.**

Option 1: Use a pivot table
- Highlight all the data
- Select the "Insert" tab and then select "Pivot Table;" the table will open in a new sheet.
- You can then use the PivotTable Fields Tab (Figure 1) to generate descriptive statistics
   - Drag your population variable (e.g., site") from the top field into the "Rows" field and then drag any of the independent variables into the "Values" field.
   - You can change the descriptive statistic from SUM by clicking the statistic and selecting "Value Field Settings." A menu will appear with all the possible options for you to select (Figure 2).
   - You can add as many descriptive statistics as you want to your table by dragging the data variable multiple times into the "Values" field and selecting different statistics using the "Value Field Settings" menu.
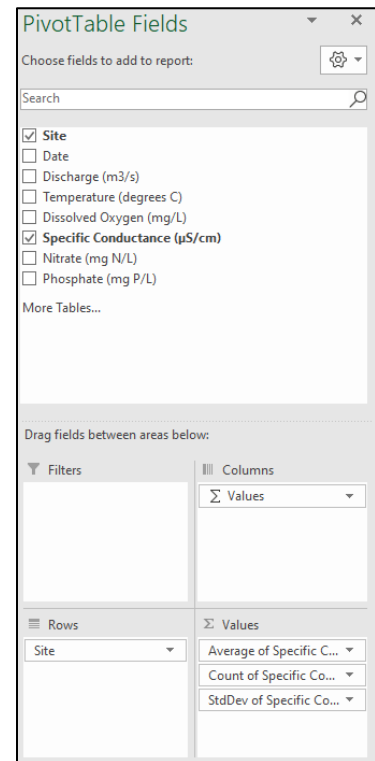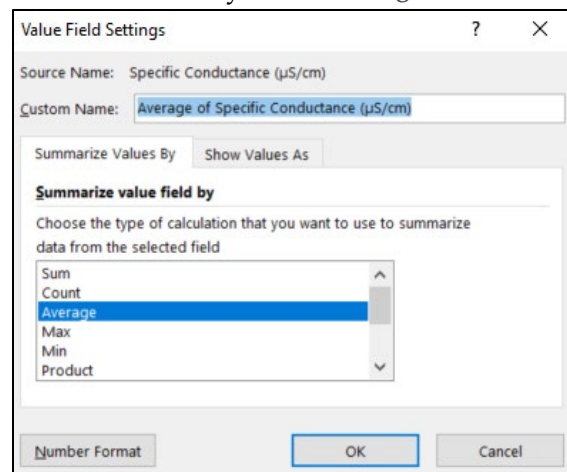

Figure 1

Figure 2

Option 2: Utilize any of the following formulas by typing them into an empty cell. The cell range of data you wish to include in the formula goes in the parentheses.
- Mean =AVERAGE()
- Median =MEDIAN()
- Count =COUNT()
- Standard deviation =STDEV()
- Variance =VAR()
- Minimum =MIN()
- Maximum =MAX()
- Standard Error = STDEV()/SQRT(COUNT())

Once you have all the descriptive statistics you need, **create a table** to highlight values that best support (or reject) the hypotheses you created during the thought questions. Remember, a hypothesis can either be supported or not supported by the data. **Don't worry if your hypothesis is not supported as unsupported hypotheses help further scientific inquiry just as much (if not more) than supported hypotheses!**

Be sure your table is easy to read and includes information about center **and** spread. Feel free to add borders and shading as needed. Don't forget an informative caption that includes both background information (sample size, location of sample collection, sample species, etc.) and highlights your results! This caption should be ABOVE your table.

CHECKPOINT – are you on the right track?
- One table with descriptive statistics describing center and spread that relate to your hypotheses
- Informative table caption above table

Part II: Exploratory Data Analysis

In addition to descriptive statistics, it's important to visualize the data to help see patterns that may not be as clear from just numbers. There are multiple types of graphs including box and whisker plots, scatter plots, bar charts, pie charts, histograms, and more. Each of these graphs answers a different question about your data and the graphs you select to create for this report should relate back to your generated hypotheses. **You should have at least one figure per hypothesis (two or more total).**

To create a figure, just highlight either the ORIGINAL data (e.g., scatterplots, box and whisker plots) or DESCRIPTIVE STATISTICS (e.g., column or bar chart) you would like to graph and select insert → "recommended charts". The recommended chart will then be one generated by Excel. If you would like to create a different chart, select the "All charts" tab and scroll through all the available options.

Be sure your figures are easy to read and include information about center **and** spread. Don't forget an informative figure caption that includes both background information and highlights your results! It should be BELOW your figure. For formatting tips for both the figure and figure caption, refer to the *figure creation guide* on Canvas.

CHECKPOINT – are you on the right track?
- At least two figures that relate to your hypotheses
- Informative figure caption below figure

Part III: Hypothesis Testing

While calculating descriptive statistics and visualizing data will provide evidence to either support or reject your hypotheses, only hypothesis tests (e.g., independent samples t-test, ANOVA, etc.) will give you a p-value so you can say with some amount of confidence that you should reject your hypothesis or not. The most commonly used p-value in scientific research is 0.05 which is what we will also use in this lab.

To conduct a hypothesis test in Excel, go to the "Data" tab and select the "data analysis" button all the way to the right of the screen. A window will pop up (Figure 3) where you can select the test you would like to perform. For this assignment, we will be using a t-test (which is for when you have one or two populations). Since we will be comparing stream measurements from Giles Run and Thompson Creek, we have two populations and will select the "Two-sample Assuming Unequal Variances" option.
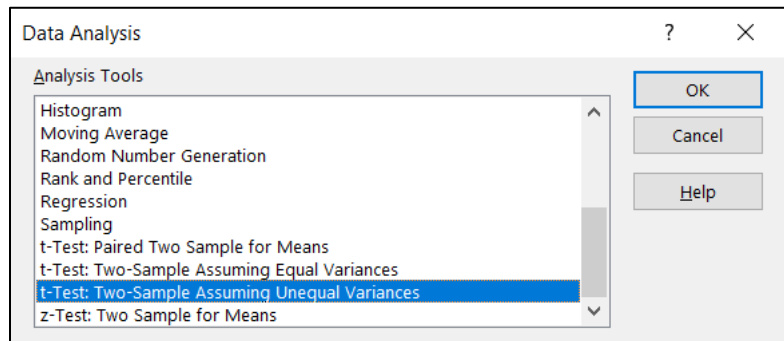

Figure 3

A window will then pop-up (Figure 4) where you can input the data for the variables you would like to test. The alpha (p-value) is already set to 0.05 so you don't need to change that. You then need to select an output range which could be empty cells within the worksheet you are working with or in a
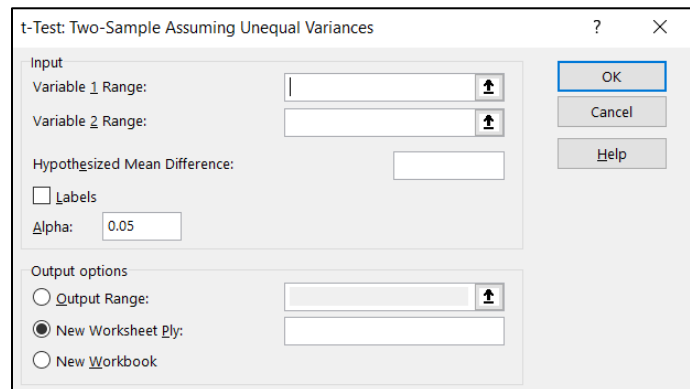

Figure 4

new sheet. You should conduct as many hypothesis tests as necessary to answer the hypotheses you generated at the beginning of the exercise.

When interpreting your results think about your hypothesis. Does it make sense for population one to have a mean greater than OR less than the mean of population two (or vice versa)? If yes, then look at the p-value for the two-tail test. If no, then look at the p-value for the one-tail test. You can also use the results to confirm your calculated mean for each population (Figure 5).

t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	Variable 2
Mean	447.4568	98.59864
Variance	93614.59	1480.458
Observations	44	44
Hypothesized Mean Difference	0	
df	44	
t Stat	7.50406	
P(T<=t) one-tail	1.05E-09	
t Critical one-tail	1.68023	
P(T<=t) two-tail	2.1E-09	
t Critical two-tail	2.015368	

Figure 5

CHECKPOINT – are you on the right track?
- Two p-values – one for each hypothesis you are testing

Summary Questions

1. Were your hypotheses supported by the data or not? When answering this question, please refer to the table and figures you created as well as the results of the hypothesis tests. This should be **no more than 4-6 sentences**.
2. If you were to ask a follow-up research question utilizing the data already collected or new data, what would it be?

Post-Lab Report Submission

Your results (as a PDF), along with your edited Excel workbook, should be uploaded to Canvas for your Methods in Ecology II Post-Lab Report. Results include: full sentence responses to the thought and summary questions as well as any tables and figures you created. Title each file "LastName_Excel_Tutorial" and with the appropriate extension (.xlsx, .PDF). You can review the rubric for your submission at any time on Canvas under the assignment to ensure you have all the components.