Introductory Laboratory
Lab 1

• Roll call
• Completion of Notice of Risks Forms
• Review of Course Schedule
• Structure of the Course and Grading
• Introduction to Scientific Approach and Data Exploration
• Excel Exercises
• Introduction to Statistics

I. INTRODUCTION TO BIOLOGY 265L

Overall purpose and student learning outcomes

The purpose of BIOL 265L is to introduce you to the fundamentals of field ecology as well as the principles of evolution. The overall goal of the laboratory is to equip you with the methodological, technical, collaborative, and critical thinking skills necessary to address basic ecological and evolutionary questions. During parts of the laboratory, the entire class will function as a large, coordinated research team, during other parts you will work independently or with a partner. Throughout the semester, you will gain hands-on experience in experimental design, field research, data analysis, critical thinking, as well as scientific written and oral communication. If you plan on pursuing a career in science, even if it is not in ecology or evolutionary biology, the ‘hands on’ experience with statistical analysis of this lab should be relevant and ultimately rewarding as ALL scientists rely on statistical analysis to interpret results.

Unlike the lecture, which is equally divided between ecology and evolution, the majority of the laboratory is ecology. Hands-on experiments in evolution are difficult to conduct during a laboratory period or even a semester. Even so, evolution is inescapably intertwined with ecology. The evolutionary history of a species’ ultimately helps explain its ecological characteristics, so most if not all ecology experiments are related to evolutionary biology.

The ecology labs in this course are designed to give you exposure to a variety of key ecological concepts and a diversity of study subjects. Over the course of the semester, the labs will deal with a variety of organisms to address concepts ranging from ecological sampling and experimental design to more theoretical concepts such as optimal foraging and island biogeography. An integral part of the semester will be Wa‘ahila Ridge which is located just adjacent to UH Manoa. It will be your field site where you will collect data for your semester long research project and formal report.

At the end of this course, you will be able to:

1. demonstrate broad-based knowledge of the fundamentals in ecological and evolutionary principles acquired through coursework and readings, including general knowledge as well as developing skills in experimental design, hypothesis testing, critical thinking, data analysis and synthesis.
2. apply appropriate statistics to explore and analyze data using Excel.
3. present data in appropriate figures using Excel.
4. read and understand scientific literature.
5. use knowledge gained in the classroom and outdoors, as well as during discussions to conceive your own project using data collected by the whole class on Wa‘ahila Ridge.
6. demonstrate the ability to find research literature appropriate to your research project.
7. communicate your research project to your peers and teaching assistants in oral and written form.
8. be capable of working productively with others as part of a team.
II. IMPORTANT POLICIES

A. Academic Conduct
Any student caught cheating or plagiarizing will, at a minimum, receive a grade of “zero” on the falsified work and have a written warning filed with the Dean of Students. According to the Student Conduct Code of the University of Hawaii at Manoa available online at http://studentaffairs.manoa.hawaii.edu/policies/conduct_code/proscribed_conduct.php:

The term "cheating" includes, but is not limited to: (1) use of any unauthorized assistance in taking quizzes, tests, or examinations; (2) use of sources beyond those authorized by the instructor in writing papers, preparing reports, solving problems, or carrying out other assignments; (3) the acquisition, without permission, of tests or other academic material belonging to a member of the UH faculty, staff or student (4) engaging in any behavior specifically prohibited by a faculty member in the course syllabus or class discussion.

The term "plagiarism" includes, but is not limited to, the use, by paraphrase or direct quotation, of the published or unpublished work of another person without full and clear acknowledgement. It also includes the unacknowledged use of materials prepared by another person or agency engaged in the selling of term papers or other academic materials.

The Biology Department takes plagiarism very seriously. Several students were given failing grades or have been expelled in the past due to academic dishonesty.

B. Laboratory Attendance
Switching lab sections is not allowed. Regular attendance is expected. Because other lab sections will be using the classroom, there will be little opportunity for you to come into the laboratory at times other than your scheduled period. Therefore, it is important that you attend each laboratory and make full use of the lab resources during your assigned lab period. Do not plan on leaving your lab early.

If extraordinary circumstances require that you attend another laboratory section, you MUST notify your TA (possibly ahead of time), who will attempt to make arrangements for you to attend another lab during the same week.

C. Laboratory Safety Regulations
- No eating, drinking, or smoking in the lab.
- Wash your hands after the exercises, and during, if necessary.
- Tie back long hair.
- Close-toed shoes are required.

These regulations will help protect you from accidents. During some labs you and your clothes may get dirty, so you will want to dress accordingly for the day’s activities. Any injuries should be reported to your TA immediately. Also, make yourself aware of the locations of the first-aid kit, and the fire extinguisher.

D. Clean Up
You must be responsible for your own cleanliness or your grade will suffer. The students in the lab following yours will report a “dirty lab” to their TA, who in turn, will see that 2% is deducted from your final grade for each occurrence. At the end of each period, you should:

1. Properly put away any instruments or equipment you have used.
2. Clean and return all other supplies to their proper places.
3. Clean your workspace of debris from field samples.

E. Laulima, Computers and Email
Computers and the Internet will have important roles in our laboratory. Your TA will communicate with you via email or Laulima. Computers will be used for data storage, analysis, and synthesis. Most data will be entered into Microsoft Excel spreadsheets and shared on your section’s web page on Laulima or as a Google document.

You must check your UH email account regularly (at least daily) as email is the major avenue of communication with your TA outside the laboratory.

F. Lab Manual & Textbooks
Your lab manual will be accessible on your 265L section’s web page on Laulima. **You will have to print the manual from the website each week before coming to class.**

Also, bring your textbooks from the 265 lecture to the lab. The textbooks will be referenced throughout the lab manual for further illustrations, background and examples.

**G. Equipment**

In this laboratory you will use equipment such as digital cameras, GPS personal navigators, wind speed meters, light lux meters, dissecting microscopes and iButtons thermal monitors. Some of these instruments are delicate and expensive. You will be instructed on their proper use and care. Please help us keep the equipment in good condition by taking proper care of it. If a piece of equipment is not working as expected, please let your TA know so that it can be repaired or replaced.

**H. Office Hours**

The TAs for BIOL 265L will hold office hours in Sinclair Library in the Commons Area. A complete schedule will be posted in the lab, outside the room on the front door, and on the website. Feel free to consult ANY of the TAs regarding the factual information covered in the lab. **For questions regarding matters pertaining to grades, you are advised to consult with your own TA only.**

The TAs make every effort to be available during their posted office hours. Remember that they are also students and can sometimes run into unexpected problems that interfere with keeping office hours. Just as they try to accommodate your special needs, we ask that you be patient in accommodating theirs. Come to Sinclair Library at another time if no one is present, or contact your TA via email to set up a special appointment.

**I. Evaluation of Your TA**

To maintain teaching quality and to enable us to continually improve our laboratories, we need your suggestions and constructive criticism. Your evaluation of your TA will help him/her do a better job in teaching you. Please note that your evaluations will be anonymous, as you will complete them online.

You will have two opportunities to evaluate your TA. The first one will be carried out mid-semester and the second one during the last week of the course. Remember that your input counts! Your suggestions and constructive comments will help us continually improve the quality of our teaching.

**J. Saving and backing up data**

For many of the exercises in this lab, you will be working on your lab work in class. In order to save your work before going home, you may want to bring a memory stick or thumb drive of some sort. The alternative is to e-mail your work to yourself as an attachment, but the more backups the better.

**K. General Conduct**

Support an effective and positive learning environment for your classmates and yourself by being on time, silencing your cell phones, and refraining from extraneous talking or activities that may interfere with other students’ opportunities to learn. Your full attention is requested especially during the introductory part of each lab in order to understand and carry out the lab exercises. Therefore also refrain from surfing the web or emailing on your own laptop or the lab computers, using your cell phone in any way other than taking notes, or listening to music. Participation points may be deducted for not complying with these general conduct rules in class.
III. HIKING AND FIELD TRIP REGULATIONS

Students will make three TA-supervised trips to Wa‘ahila Ridge, which is on the Diamond Head side (east) of campus. Hiking in Hawaii is an enjoyable activity. However, hiking has its inherent dangers and we advise students to plan ahead, think clearly, and be safe! The following are requirements and recommendations for hiking in the field.

If you are unable to hike up Wa‘ahila Ridge for health reasons, please talk to your TA.

ABSOLUTELY NO SOLO HIKES! ALWAYS HIKE WITH A BUDDY!

A. Appropriate Footwear
- Hiking Boots or Hiking Shoes (closed-toe and sturdy): Shoes that you do not mind getting scratched up and dirty.
- DO NOT WEAR slippers, high heels, or anything too valuable or inappropriate for walking on rocky and rugged trails.

B. Appropriate Clothing
Wa'ahila Ridge is very exposed to the elements; the sun and heat on the ridge are intense, so do not be fooled. An information sheet on preventing, detecting, and treating heat stress and heat exhaustion can be found in Appendix B.

In general, wear clothes that you do not mind getting sweaty, dirty, or torn.
- Hat for sun protection.
- Lightweight and loose long pants are recommended. Wearing shorts may not protect you from scratches by the vegetation.

- If you have especially delicate or allergic skin, you may wish to bring a lightweight, long-sleeved shirt to cover your arms and/or light gloves for handling plants.
- Rain gear if necessary.
- Backpack to carry your supplies.

C. Miscellaneous Necessities
- Water, 1 liter minimum
- Sunscreen
- Mosquito repellent
- First Aid Kit (TA will carry one)
- Field Notebook and Pencil
- Plastic bag for rubbish

D. Garbage In, Garbage Out
NEVER leave rubbish in the field. ALWAYS take rubbish out of the field. If you see trash that is not your own, PICK IT UP! We have a responsibility to maintain our research site!

E. No Horse-Play
Do not horse-around out in the field. Wa‘ahila Ridge has many obscured drops. Do not cause yourself or anyone else to fall down the ridge.

F. Do not Tread On Other Peoples’ Projects
Stay on the trail. There are over two hundred students conducting projects on the Ridge. Be aware that walking off the trail increases erosion and the likelihood that pitfall traps and other student projects may be trampled. Watch your step in general!

G. Report Possible Dangers And All Injuries
It is everyone’s responsibility to maintain as safe an environment as possible. Report and mark hazards to prevent injury to others!

Report ALL injuries to your TA!
IV. ASSIGNMENTS & GRADING

A. Final Grade Breakdown

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Percentage of Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 Lab Assignments</td>
<td>60</td>
</tr>
<tr>
<td>1 Final Formal Report</td>
<td>20</td>
</tr>
<tr>
<td>10 In-lab Quizzes</td>
<td>10</td>
</tr>
<tr>
<td>1 Final Oral Presentation</td>
<td>5</td>
</tr>
<tr>
<td>In-class Engagement/Team work</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100</td>
</tr>
</tbody>
</table>

The bulk of your grade will be determined by your scores on lab assignments and a formal report. There will be 11 in-lab quizzes, but your worst score will be dropped. Most students do well on the quizzes and in-class engagement, so your lab assignments and formal report will be even more important in determining your grade. This also means it is important to do well on the quizzes, final oral presentation, and in-class engagement. You will not have a chance to make up for missing grades on a final or with extra credit.

B. Lab Assignments

The weekly lab assignments are to be turned in promptly at the beginning of the next lab period (1 week after they were assigned), unless otherwise noted in the laboratory schedule. Do not plan on spending any of your laboratory time finishing an assignment that was due. This includes printing!

This is a writing intensive class and therefore you are expected to be proficient in written English, as it will be a big part of your lab assignments. If you have a hard time with your writing please take advantage of the resources and tutors at the UH Manoa Student Success Center and the Manoa Writing Center located at Sinclair Library. For more information ask your TA and check out the following websites:

www.english.hawaii.edu/writingcenter/

C. Late Lab Assignments

Late lab assignments will NOT be accepted. Ask your TA about what to do if extenuating circumstances prevent you from being able to turn in your work on time.

There is no final exam. You need to turn in all your homework on time in order to succeed in this class.

D. Final Formal Report

You will write a final formal report in the format of a scientific publication, which will be due at the end of the semester. This formal report will include assignments from throughout the semester that deal with your data collection and analyses on Wa‘ahila Ridge. You will have the chance to submit drafts of parts of your formal report to your TA throughout the semester. Make sure to include comments and suggestions from your TA and a peer in your final completed version.

E. In-Lab Quizzes

Prior to attending each lab, you will be required to read the lab manual in order to be able to understand what we will be doing during the lab. Some weeks you will also be required to read some extra reading material, mostly primary scientific literature, to enhance your learning of the lab content and your critical thinking and reading skills. To make sure you are reading the lab manual (and possible supplemental readings) there will be a short quiz or write-up at the beginning of each lab. The quiz will be based on the lab being conducted that day so read over the lab and supplemental reading material carefully before coming to lab and taking the quiz. The quiz could also be about a former lab used as a review and refresher of what you have learned so far during the semester. The quiz will take up the first five minutes of each lab, so make sure you are on time. There will be no make-ups for missed quizzes. However, your worst quiz score will be dropped at the end of the semester.
**F. Oral Presentation**
You will present the analyses and conclusions of your final report to your classmates during your last lab in a group presentation.

**G. In-class Engagement & Team work**
Active participation in class activities and discussions, and kokua towards your classmates and TAs will earn you points. Please help contribute to a positive learning environment. Good class engagement and teamwork will earn you 15 points, showing initiative and going the extra mile will earn you up to 20 pts. Disruptive behavior, consistent tardiness, unexcused absences, etc. will reduce your score. Your TA has the right to remove all in-class engagement points for unacceptable or dangerous behavior.

**H. Class Attendance**
Though you do not receive a grade for attendance per se you will not be able to do well in the course if you do not go to class.

An unexcused absence from a lab period will result in a 0 for that lab assignment AND you will get a 0 for the quiz AND points off your final class participation grade. This will drop your grade by at least one letter.

**I. No Extra Credit**
There is no extra credit available for this course.

**J. Lab Notebook**
You may want to keep a lab notebook or folder throughout the semester (optional). Your lab notebook can be used to record data and take notes during each lab period. In the field, it can be used for notes on environmental conditions & notable observations. It is surprisingly easy to forget when and where you collected data, so, it is important to enter the date, your location and the lab assignment number, along with the entries in your notebook. This information will help you with your lab write-ups and ridge project.

**V. LABORATORY EXERCISES**
Answering ecological and evolutionary questions is a multi-step process, following the scientific method. Usually researchers start with some interesting idea or question, which they then transform into a testable hypothesis. Once a question or hypothesis is formulated the data collecting begins, which can consist of a lot of field/lab work or computer modeling/simulations. This process can be very tedious. After all the necessary data is collected, it usually needs to be entered into the computer, analyzed, and visualized using graphs and figures in order to find some patterns or answers to the original question. This last part, the actual interpretation of the data includes reading a lot of references and studies dealing with similar problems and summarizing it all in a written form that is understandable to other researchers. Ideally, your interpretations should lead to new and exciting inferences about the scientific world.

During the 265L course you will do a little bit of everything, but especially learn how to explore, analyze and interpret data collected by others or as part of the lab exercise. The first few laboratories (including today’s lab) will also introduce you to the basics of statistical analysis. The purpose of this week’s laboratory is to introduce you to the basics of descriptive statistical analysis. In Lab 2 you will learn some basic inferential statistical analysis. Why do you have to do math in an ecology class? Statistics are used in every Science, as they are critical to understanding whether any patterns we see in experimental results are due to real effects or just normal random variation. Before you learn about how to calculate variation in a dataset and test if it is due to some true effect you need to be familiar with some terms in descriptive statistics. You will also learn how to calculate them in Excel.

Before we introduce you to some basic statistics we want to make sure that all of you have some basic understanding of digital cameras, computers and the software Excel. It is important that you learn how to use these two pieces of equipment so that you can complete your laboratory assignments in a timely manner.
A. Laulima
Throughout the semester, you will need to retrieve data & other files for use in lab and for homework assignments. These files will be made available by your TA via the course’s web site on Laulima.

Tell your TA if you do not get their emails or if you are not able to access your sections’ site on Laulima.

B. Using the Digital Cameras
During the semester, you will take pictures of the places you visit and the organisms you observe in the field. The digital cameras you will use in BIOL 265L are very easy to operate.

You should become familiar with the following:
- adjusting the flash
- activating macro mode
- adjusting photo quality
- viewing the photographs you have taken
- deleting photos from the camera
- downloading photos.

Taking Pictures
To take a picture, you must first turn the camera on (“on/off”). There are several picture taking modes. You will mostly use automatic mode (camera symbol). You can frame your picture only on the view screen. On the back of the camera you will notice several buttons. At the top are two rectangular buttons, which control the zoom. “T” is zoom in and “W” is zoom out. When you want to take a picture, hold the trigger button down halfway until you see a yellow frame in the viewfinder (to focus the object to obtain a high quality image) and then press the trigger all the way. When taking close-up pictures (e.g. of plants) use the macro mode. Press the left of the cross button and select the MACRO mode (flower symbol). You can choose other shooting modes (e.g., “portrait”, or “landscape”) by pressing the “menu” button. If you want to activate or deactivate the flash (flash symbol), proceed the same way as for the MACRO mode.

Viewing Pictures
To view the pictures you have taken, press the third button from the top on the back of the camera (arrow symbol). You can use the left and right side of the cross button to scroll through the pictures. Delete a picture by pressing the top part of the same cross button (trash can symbol) and select “ERASE” pressing the menu/ok button, then press the menu/ok button again.

When taking images:
- Hold the camera as steady as possible so you don’t get blurry images.
- Be as close to your subject as possible. Do not include a lot of unnecessary space around your subject. However, for the camera to focus the object, it must be at least 10 inches from the subject.
- Unless you are a professional, leave the camera in the Auto-mode.

1. Photographing Your Labmate
   a) Choose a partner to take photographs of each other
   b) Familiarize yourself with the controls on the digital camera.
   c) Try different angles and distances. You can also try the “Portrait” scene in the scene mode (SCN). Then, delete all but the best picture of your partner.
   d) When you are finished, pass the camera to your partner so that s/he can take a picture of you.
   e) Fill in the picture log and pass the camera to the next pair of students.

2. Preparing the images to be placed on the web
   a) Your TA will download and rename your pictures.
   b) These pictures will be used to create a class web page for your lab section. You can use this web page to help you get to know your classmates or to communicate with them via email. They will not be available to the public.
C. Data Exploration

There are many different programs to explore, analyze and present data. We will use Microsoft Excel during this semester. Microsoft Excel is a simple and commonly used spreadsheet program that is capable of storing, analyzing and graphically presenting data. You must become familiar with the fundamental features of Excel because you will be using Excel to store the data you have collected, create graphs, and conduct statistical analyses. Excel is a very easy program to learn, and everything you need to know for this semester will be described in this manual. Your TAs will assist you. Work through the following exercises to become familiar with the basic Excel tasks. Also, if you would like to learn more about Excel, we have some books to help you and a pdf file describing basic functions of MS Excel 2013 on your section’s Laulima site under “Resources”.

1. THE BASICS

Opening Excel

- Excel can be opened by double clicking the Excel icon inside the “Microsoft Office 2013” folder in the “All Programs” menu,

The Excel environment is composed of a large table known as the “worksheet”, a title bar, a ribbon with several tabs, and a formula bar (Figure 1). Each block in the table is called a cell. You can type anything into a cell – words, numbers, dates, times, formulæ, etc. To move between cells you can point and click with your mouse or use the arrow keys, the “tab” key, and the “enter” key.

Figure 1. The MS Excel environment is organized into a large table with letter labeled columns and number labeled rows.
Selecting Multiple Cells
You may need to select cells when you want to format, cut, copy, delete, or sort data or create graphs. When a cell is selected it has a heavy black outline or is entirely black.

• To select more than one cell you must drag the mouse (hold the left mouse button and move the mouse) over the block of cells you wish to select.
• Noncontiguous cells can be selected by holding the CTRL key between mouse clicks or drags.
• Entire columns or rows can be selected by clicking the letters at the top of the table or the numbers at the left of the table, respectively.

Opening & Saving Files
When you first open Excel, you are presented with a new workbook. Each Excel file is called a workbook. When you open a file or save a file, you are opening or saving an entire workbook. A workbook can contain many worksheets, which are analogous to pages. A worksheet can contain data and/or charts.

• To open a file, click File, Open on the “File” Tab in the Ribbon and browse through the folders until you find the file you want.
• To save a file, click File, Save As and browse to the folder where you would like to save the file (most likely the E: Removable Disk).

2. MANAGING DATA

Sorting
Sorting your spreadsheet data is an important skill to learn. You may need to sort your data to complete your homework assignments. Sorting refers to the alphabetical or numerical ordering of a column of data. The key to successfully sorting your data is selecting the proper cells. Anything that is not selected, will not be sorted. You will need to select all of the data in a table before sorting it. Your data can become irreconcilably jumbled if you sort the data incorrectly by selecting fewer columns than intended.

• Select the data you want to sort. It is OK to select one “header row”, which is a row of names or labels for each column.
the selected row) or column (to the left of the selected column). Then you can insert or delete the columns or rows. On the **Home** tab, in the **Cells** group, click the arrow on the **Insert** button and select the respective command (see figure below). You can also right-click the row header and select **Insert from the shortcut menu.**

### 3. USING FORMULAS

What makes Excel so valuable is its ability to quickly process large amounts of data. Using the formulas in Excel is one way to speed up your calculations substantially. If you learn how to use these formulas, you will almost never use a standard calculator again. Although it is not imperative that you use formulas for this class, it will make your life easier for some assignments.

**Cell Contents & The Formula Bar**

Each cell can contain a different formula. When a formula has been used properly, you will not see it in the cell. You will only see the calculated result of the formula, which is usually a number. The way to find out if a cell has a formula or just a number, is to select the cell in question and look at the formula bar (Figure 1). The formula bar displays the true contents of a cell.

**Basic Math**

To let Excel know that you are using a formula, the first part of every formula is an `=` sign. Try experimenting with the following formulas:

- `=1+1`, `=4-9`, `=15/5`, `=4*6`

You can also perform basic math on the values of specific cells. Note that the columns are labeled with letters, and the rows are labeled with numbers. The cell in the upper left corner of the sheet is called cell A1. Instead of adding 1 and 1 you could add cell A2 to cell B2 with the following formula (of course, there must be values in cells A2 and B2 for this to work):

```
=A2+B2
```

Subtraction, multiplication, and division also work in a similar fashion.

**Sum, Mean**

What if you want to add the values of many cells? Instead of typing out the addition formula you can use the SUM function.

```
=sum(A1:A2) or =sum(A1:A1000)
```

A colon (:) between the two cell references tells Excel to add all the values between cells A1 and A2 (or A1000).

Similarly, you can quickly calculate the mean of a large set of numbers using the AVERAGE function. An example calculation is below. It is important to note that although there are many measures of average, Excel interprets the word average as the mean exclusively and not the mode or median.

```
=average(A1:A1000)
```

You do not have to type out the cell references. If you start a formula, such as `=average(...)` you can then use your mouse to select the cells you want to include in your formula. In the Statistics part of this lab you will be introduced to more functions, such as MEDIAN, STDEV (Standard Deviation), and VAR (Variance).

**Mass Production**

Often you will want to calculate the sums or means of many columns. When you copy and paste a cell with a formula, the formula is automatically adjusted for the new cell. If you copy cell A10 with the formula `=sum(A1:A9)` and paste it into cell B10, then the formula will change in cell B10 to `=sum(B1:B9)`. This feature allows you to make numerous calculations in a fraction of the time it takes with a calculator.
4. **CREATING GRAPHS**
Creating graphs is a major skill required by this course. There are numerous types of graphs in excel. At the very least, you will need to know how to make two basic graphs, *column graphs* and *XY scatter plots*. When the data on your x-axis are categorical (spring, summer, fall, winter or ocean, lake, stream, swamp or species A, species B, species C, etc.), you should make a column or line graph. When the data on your x-axis are numerical and continuous (1, 2, 3 or 0.3, 1.2, 3.7), you should make a XY scatter plot.

**Column & Line Graphs**

1. Before creating a graph, your data should be properly formatted. The categories for the x-axis of the graph should be the first (left-most) column. The y-axis data should be in the second column.

2. Select the data for your graph. If you select column headings, they will be automatically incorporated into the graph.

3. To make a column graph, click the **Column** button on the **Insert** Tab and the following window will appear.

4. Select e.g. **Clustered Column** under the **2-D Column** graphs and the graph will appear on your worksheet.

5. In the **Chart Tools** Tab you can choose among one of two Tabs to edit or adjust your graph. You can chose a layout by clicking on the **Quick Layout** button on the **Design** Tab and change your Axis, Axis Titles, Data Labels, etc. by clicking directly on the respective element in your graph.

6. If you use your graph in a presentation (e.g. in powerpoint) you would add a title. If you use a graph in a text, then it is considered a figure and needs an appropriate caption at the bottom of the figure (make sure that you are not adding a title as well). You also need to give your figure a number so that you can refer to it in the text (see Figure 2).
**XY Scatter Plots**

Scatter plots are made in much the same way as column graphs. The major difference is that a scatter plot has continuous numerical data on the x-axis.

1. Setup the data in your spreadsheet so that the first column has the x-axis data. Select the data.
2. On the **Insert** Tab, click on the **Scatter** plot button and select the **Scatter with only Markers** chart button (see below).
3. Add axis titles etc. the same way you did in the bar graph example. You can delete the Legend to the series if you have only one series (i.e., 1 set of data points as compared to Figure 2 where you have 5 series/data sets).

4. Once you added an appropriate caption (or title) your graph is complete (see Figure 3). Again, you can still edit the graph by double clicking or right-clicking on various parts of the graph.

**D. Introduction to Basic Statistics**

Statistical calculations are often used in ecology to show the relationships between organisms and their environment, and to test hypotheses, or make predictions. Statistics can be divided into two general categories: descriptive and inferential.

**Descriptive statistics** are used to summarize and simplify a large collection of data (measurements) into a few numbers that convey something informative (statistics) about the set of data. You have all used descriptive statistics without even knowing it. For example, if you have ever compared your own test score to the class average (or mean), you have made use of a descriptive statistic, namely the mean.

**Inferential statistics** are used to test the Null-Hypothesis ($H_0$). These statistics make use of the theory of probability to draw conclusions about a large group based on results obtained from a small sample of that group. In the language of statistics, this large group is often referred to as the **population**. This is a general term, which may refer not only to a group of living organisms, but also to any group of objects.

In the context of ecology, **raw data** are data that you collect directly from the organisms or the environments you are studying. For example, most
if not all of the data you will collect in the field are raw data. Raw data sets are difficult to interpret and compare because they are usually comprised of many individual numbers. However, they can be summarized by calculating descriptive statistics such as the mean, median, range, variance and standard deviation. These descriptive statistics are sometimes referred to as **summary statistics**, because they summarize properties of a data set.

In order to illustrate how to calculate these summary statistics, we will use some simple data sets: 1) tree fern height and 2) student height.

**Tree fern height:** five measurements made on the height of *Cybotium glaucum* (hapu‘u or tree fern). Each tree fern was chosen randomly by walking 20 steps and measuring the closest tree fern and repeating the procedure until five measurements were made. The heights (raw data) are as follows:

- 1.2 m
- 3.0 m
- 0.5 m
- 2.3 m
- 1.5 m

**Student height:**
With a lab partner, measure each other’s height with a measuring tape. Get used to metric units, the units mostly used in science, and measure in centimeters (cm) and not feet or inches. Report your height to your TA. You will use the heights of your whole class to explore some basic statistics this week and in the following weeks.

**Frequency Distribution**
The frequency distribution of a group of numbers (measurements) is most easily visualized by plotting the measurements on the x-axis (measurements need to be divided into bins) versus the frequency of occurrence on the y-axis. This type of graph is called a **histogram**. As an example, a histogram from a past year’s final exam scores is shown in Figure 4.

Most frequency distributions are ‘mound-shaped’. One specific type of frequency distribution is called the **normal distribution**. The normal distribution is often referred to as a 'bell-shaped curve'. A perfect example is shown in Figure 5B.

Because some statistical tests that we will be using assume a normal frequency distribution, it is important to understand how to visualize a frequency distribution (using a histogram), so that you can confirm that the frequency distribution of your data is similar to a bell-shaped curve. In comparing the frequency distribution shown in Figure 4 to those shown in Figure 5, it is apparent that Figure 4 is most similar to Figure 5B, which has a normal distribution. This concept will be revisited later when specific statistical tests are discussed in Lab 2.

**Population & Sample Size**
The **sample size** is denoted by a lower case ‘n’ and indicates the number of observations, measurements or individuals measured. In the case of the tree ferns, n = 5 because 5 tree ferns were measured. In contrast, the **population size** is denoted by an upper case ‘N’ and represents the total number of individuals in the population of interest. In our case, if we want to use our sample to make inferences about the heights of tree ferns within the whole forest, then the population size (N) would be the number of tree ferns in the entire forest. It is usually impractical or impossible to
measure or even count an entire population. Nevertheless, we are interested in estimating the properties or characteristics of the entire population \(N\) based on the sample size \(n\). A sample size of five individuals is probably not large enough to get a precise estimate of the average height of the population \(N\). If the sample size is sufficiently large and measurements are made on a representative sample of the population, then the summary statistics for the data set (i.e. the mean, median, and variance of \(n\)) will approximate the summary statistics for the entire population \(N\).

**Sample Mean**

The **sample mean** is simply the average of all measurements taken. It is calculated using the following equation:

\[
\bar{X} = \frac{\sum X}{n}
\]

Where the \(\bar{X}\) represents the sample mean, \(\Sigma\) symbolizes the sum, \(\Sigma x\) represents the sum of all observations, and \(n\) represents the sample size. As indicated by the equation, the mean is calculated by adding up all the observations and dividing by the sample size. Remember that the purpose of calculating the sample mean (\(\bar{X}\)) is to get an estimate for the true population mean (often denoted \(\mu\)). In our example, \(\mu\) represents the mean tree fern height for the entire forest. We usually cannot measure \(\mu\) directly because \(N\) is too large.

Returning to the data for the heights of the tree fern, the sum of the five measurements is 8.5 m, and the sample size is 5, so **the sample mean = 8.5m / 5 = 1.7 meters**. This is our best estimate of the population mean (\(\mu\)), but it is not usually exactly equal to \(\mu\) because our estimate is based only on a sample from the population, rather than measurement of the entire population.

**Sample Median**

The **sample median** is the middle datum or measurement in a ranked list of data. First, you must arrange the data from smallest to largest. Using the tree fern data, the list should appear as follows:

\[
0.5 \ 1.2 \ 1.5 \ 2.3 \ 3.0
\]

The number that is in the center (1.5 m) is the median tree fern height. If the sample size is an even number, then there will be two “middle points”. In that case, the median is the mean of those two “middle points”.

**Mean Versus Median**

When there are a few unusually large or unusually small numbers in your data set, the median often gives a better indication of a "typical" measurement than does the mean. Unusually large or unusually small numbers will skew the mean away from the “typical” value for a population (see Figure 5). For example, household incomes are almost always compared using the median rather than the mean. This is because there are always a few individuals like Bill Gates who makes a huge sum of money (e.g. $1 billion), thus **inflating the mean** household income so that the mean is higher than what most people earn (Figure 5C). The median is much less affected by extreme values, and the median household income is more representative of a “typical” household.

**Range**

Variability is the difference among individuals in a sample. The most basic measure of variability is the **range** of the data, which is the difference between the largest and the smallest data points in a data set. For the tree fern example, the range is:

\[
3.0 - 0.5 = 2.5 \text{ meters}
\]

The range indicates the extremes in your measurements. Notice that data sets B and C in Figure 6 have the same ranges, but they obviously differ in variability.

**Sample Variance**

A much better (but less intuitive) measure of the variability of your data is the **sample variance**. The sample variance is calculated as follows:

\[
s^2 = \frac{\sum (X - \bar{X})^2}{n - 1}
\]
Where \( s^2 \) represents the sample variance and the other letters and symbols are the same as those used to calculate the mean.

Using the mean of 1.7 for the tree fern data, the numerator of the equation is calculated as follows:

\[
(0.5 - 1.7)^2 + (1.2 - 1.7)^2 + (1.5 - 1.7)^2 + (2.3 - 1.7)^2 + (3.0 - 1.7)^2 = 3.78
\]

The denominator of the equation is \( 5 - 1 = 4 \)

\[
s^2 = \frac{3.78}{4} = 0.945 \text{ m}^2
\]

When the sample variance \( (s^2) \) is large, there is a lot of variability in your data. Likewise, if the variance is small, then it tells you that your data are less variable. What is "large" and what is "small"? It is all relative. Consider the samples in Figure 6. Sample A has the lowest variance and sample C has the highest variance.

Confidence Intervals
It is important to remember that the motive behind calculating the sample mean is usually to get an estimate of the population mean. As an estimator of the population mean, it is useful to determine how precise a sample mean is. The standard error of the mean, also called the standard error \( (SE) \) is a measure of how precisely the sample mean estimates the true population mean \( (\mu) \). Notice that \( SE \) is in capital letters to avoid confusion with "s", the standard deviation.

The SE should decrease with increasing sample size; thus, an estimate of the true population mean becomes more precise as the sample size is increased.

Having determined the sample mean and SE, one can calculate a confidence interval \( (CI) \), a range around the sample mean that is likely to contain the true population mean. A CI always has a probability (confidence level) associated with it and the confidence level is chosen arbitrarily. It usually ranges from 90% to 99%. The formula to calculate a 95% confidence interval is as follows:

\[
CI_{95} = \bar{X} \pm (t_{(df)} \cdot SE)
\]

Where \( CI_{95} \) is the 95% confidence interval, \( \bar{X} \) is the sample mean, \( SE \) is the standard error, and \( t_{(df)} \) is the t-value taken from the t-table located in the Appendix on page D-1.

In order to locate the correct t-value on the t-table (Appendix D-1), you need to know two numbers. The first number is the degrees of freedom \( (df) \)
which equals \( n-1 \) when calculating a confidence interval (CI). The second number is the confidence level. The confidence level in most studies is set at 95%, which gives you a 5% chance for error or a 5% chance that the confidence interval does not contain the true mean. In Appendix D-1, confidence levels are listed as \( \alpha \)'s. \( \alpha \) is the probability of making a Type I error and equals one minus the confidence level in proportions. For example, for a confidence level of 95%, \( \alpha = 0.05 \). To find the correct \textbf{t-value} on Appendix D-1, scroll down the \( \alpha \)(2-tail) 0.05 column until you reach the row where \( df = 4 \). Your \textbf{t-value} equals 2.776.

Continuing with our tree fern example, the 95% confidence interval for the sample mean will be:

\[
\text{CI}_{95} = 1.7 \pm (2.776 * 0.435)
\]

\[
= 1.7 \pm 1.21 \text{ m}
\]

You can say with 95% confidence that the true mean height for the entire population of tree ferns in this forest lies between 0.49 and 2.91 m. Conversely, we can say that there is a 5% chance that the true population mean is outside the range of 0.49 to 2.91 m. This is a huge range and you could easily improve this estimate by increasing the sample size, \( n \).

The accuracy of our calculated confidence interval depends on several assumptions that are listed under "Student’s t-test" in Lab 2.

**IV. STATISTICS WITH EXCEL**

It is much easier to use Microsoft Excel to compute your statistics than to calculate them by hand. Under the “Data” Tab, select “Data Analysis”, and choose the appropriate statistical test.

**Statistical Functions for Formulae**

To find a statistical (or any other) function go to the Formulas Tab and click on any of the symbols within the Function Library grouping. Listed below are the formulae you may use in this course. For all of the functions below you must insert a data set (cell range from your spreadsheet) in the parentheses, such as A1:A56.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>=min()</td>
<td>Minimum Value In Sample</td>
</tr>
<tr>
<td>=max()</td>
<td>Maximum Value In Sample</td>
</tr>
<tr>
<td>=sum()</td>
<td>Sum</td>
</tr>
<tr>
<td>=average()</td>
<td>Sample Mean</td>
</tr>
<tr>
<td>=median()</td>
<td>Sample Median</td>
</tr>
<tr>
<td>=stdev()</td>
<td>Sample Standard Deviation</td>
</tr>
<tr>
<td>=vara()</td>
<td>Sample Variance</td>
</tr>
</tbody>
</table>

**Data, Data Analysis...**

**Descriptive Statistics**

If your data analysis requires the use of multiple formulae, it is easier to use the \textbf{Descriptive Statistics} function rather than inputting each individual Excel formula into a column. To locate the Descriptive Statistics function click on Data Analysis in your Data tab. In the Data Analysis window select \textbf{Descriptive Statistics}.

You must enter the \textbf{Input Range} (the data to be analyzed) and the \textbf{Output Range} (the area on the worksheet where you want the results to be placed), and check off the \textbf{Summary Statistics} checkbox and the \textbf{Confidence Interval} checkbox. Once the parameters are set, click \textbf{OK}. The following stats will be calculated for you: Mean, Standard Error, Median, Mode, Standard Deviation, Sample Variance, Range, Minimum, Maximum, Sum, Count, Confidence Interval (95.0%).
*Ignore the $ signs.

**Data, Data Analysis...**

**Histogram**

If you want to look at the distribution of your data it is best to first plot a **histogram**. You first need to add a column on your datasheet with possible **bin categories**, i.e. possible categories of all values you want to plot in the histogram. You can find the range of your bins by sorting the column of the data you want to plot. Then go to **Data Analysis** in the **Data** tab and select **Histogram**. Enter the **Input Range** of the data and the **Bin Range**, as well as the **Output Range**. Then click **OK**.

You now will have two columns, one with the bin categories and one with the frequency of your data in each bin and a **column chart** of these two columns, where the bins are on the x-axis and frequency is on your y-axis.

---

**E. Excel Exercises/ Homework Assignment** *(Total 35 Pts.)*

Before leaving lab, be sure to show at least all the graphs to your TAs. The graphs and your written discussion of them are due next week.

**Be sure to label the units on the axes correctly in your graphs, copy and paste the graphs into a MS Word file, and add an appropriate caption below each figure (do not use titles in the graphs).**

1. **Yield of vegetables:**
   a) Type the following data into a new Excel workbook (units for rainfall are cm/month, units for vegetables are yield in g/m²).

<table>
<thead>
<tr>
<th>Month</th>
<th>Month#</th>
<th>Rainfall</th>
<th>Tomato</th>
<th>Peppers</th>
<th>Bean</th>
<th>Peas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun</td>
<td>6</td>
<td>5.2</td>
<td>132</td>
<td>106</td>
<td>106</td>
<td>129</td>
</tr>
<tr>
<td>Jul</td>
<td>7</td>
<td>7.9</td>
<td>198</td>
<td>115</td>
<td>157</td>
<td>170</td>
</tr>
<tr>
<td>Aug</td>
<td>8</td>
<td>3.2</td>
<td>115</td>
<td>103</td>
<td>93</td>
<td>131</td>
</tr>
<tr>
<td>Sept</td>
<td>9</td>
<td>10.6</td>
<td>226</td>
<td>118</td>
<td>178</td>
<td>162</td>
</tr>
<tr>
<td>Oct</td>
<td>10</td>
<td>6.9</td>
<td>169</td>
<td>111</td>
<td>135</td>
<td>149</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>600</td>
<td>639</td>
<td>556</td>
<td>730</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td>600/12</td>
<td>639/12</td>
<td>556/12</td>
<td>730/12</td>
</tr>
</tbody>
</table>

   b) Sort the data by Rainfall.

   c) Use formulas to calculate the total and average values for each column.
d) Plot the following graphs *(they should look like Figures 1 and 2 below):*
1. Bean yield versus rainfall using an X-Y scatterplot (Figure 1). *(2 pts.)*
2. Tomato and pepper yield versus month using a column graph (Figure 2). *(2 pts.)*

*(Note regarding order of the months: Sort by the “Month#” column to put the months back into chronological order.)*

![Bean Yield vs Rainfall Graph](image1.png)

**Figure 1.** Monthly yield of beans (g/m²) versus rainfall (cm/month).

![Tomato and Pepper Yield Graph](image2.png)

**Figure 2.** Monthly yield (g/m²) of tomatoes and peppers from Summer to Fall.

---

2. **Age of the Hawaiian Islands:**

   Go to the website [seattlecentral.edu/qelp/sets/073/073.html](http://seattlecentral.edu/qelp/sets/073/073.html). Get some information on the dataset under “About the Data”. Then click on “ExcelFile” to download the data (“Save file”).

   a) Open it in Excel and plot age (x-axis) versus distance (y-axis) using an x-y scatterplot. Make sure you label axes appropriately. *(2 pts.)*

   *Note: You can move your pointer onto a data point and see which volcano/island it belongs to. Did you find Oahu?*

   b) Make a column chart of the age (y-axis) of all the Main Hawaiian Islands. Average age of volcanoes where needed (i.e. where there are multiple volcanoes per island). If you don’t know which volcanoes belong to which island, as your fellow students or try to find some information on the web. *(3 pts.)*

3. **Coral Reefs on Coconut Island:**

   Get the file “ReefTransect_2006.xls” from your sections page on Laulima (ask your TA for the exact location) and make a column chart of % cover of species at different locations on the reef. Change the format of the data as needed. *(3 pts.)*

4. **Sulfur emissions of Kilauea:**

   Get the file “Kilauea eruption sulfur.xls” from your sections page on Laulima (ask your TA for the exact location). The data is from the following NASA website: [gcmd.nasa.gov/records/GCMD_USGS_OFR_2007_1114.html](http://gcmd.nasa.gov/records/GCMD_USGS_OFR_2007_1114.html). Open the file in Excel and plot sulfur emissions as a function of date using an x-y scatterplot. *(3 pts.)*

   Calculate the mean *(1 pt.), range (min/max) *(1 pt.), and sample variance *(1 pt.)* using excel formulas and functions AND get the descriptive statistics *(1 pt.)*.
5. Student Height:

Plot student height of your section in a graph. (2 pts.)

What would be the best way to plot this data in order to get the most information out of it? What is the class average? Are there different ways to plot the data? Discuss with your lab partner and then ask your TA for the best way. (2 pts.)

Calculate the mean (1 pt.), range (min/max) (1 pt.), and sample variance (1 pt.) using excel formulas and functions, get the descriptive statistics (1 pt.) and make a histogram (2 pts.). Look at mean, median and histogram: Is the data skewed, normally distributed? (1 pt.)

6. Discussion/Summary:

Look at all your graphs with your lab partner and discuss what these graphs tell you about the data. Think about other ways you could plot the datasets and what questions they would answer. Write up a few sentences for each graph after your discussion. Even though you discussed the graphs with your partner make sure that the summary of your discussion is written in your own words (i.e. don’t just copy your partners sentences, as this would be considered plagiarism). (6 pts.)