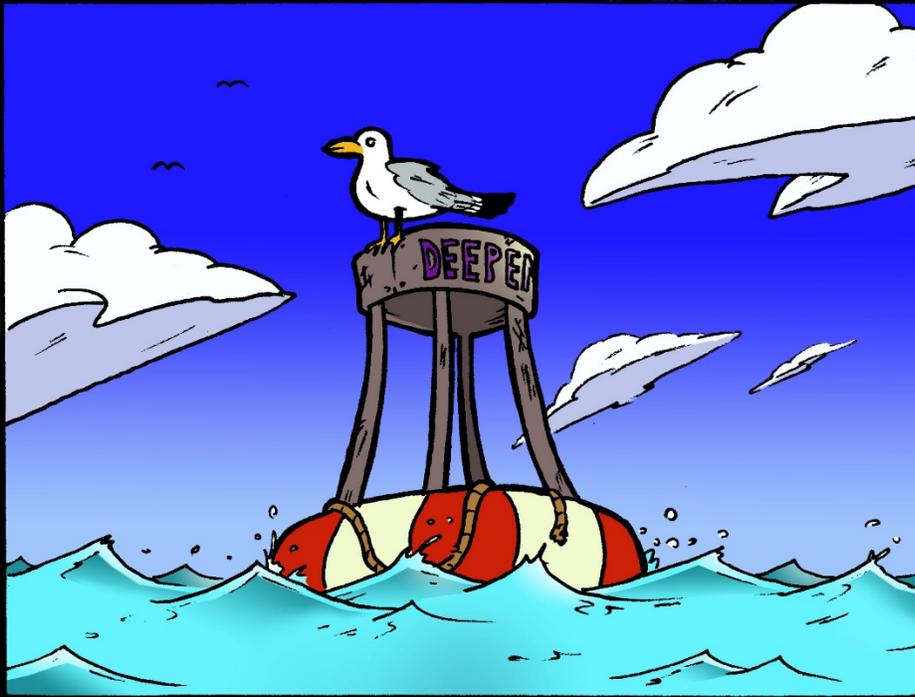


Nature of Science

Teaching and Learning Module Grades 6–12



Dear Educator,

We are pleased to present the 6th in a series of teaching and learning modules developed by the DEEPEND (Deep-Pelagic Nekton Dynamics) Consortium and their consultants. DEEPEND is a research network focusing primarily on the pelagic zone of the Gulf of Mexico, therefore most of the lessons will be based around this zone. Whenever possible, the lessons will focus specifically on events or aspects of the Gulf of Mexico or work from the DEEPEND scientists.

All modules in this series aim to engage students in grades 6 through 12 in STEM disciplines, while promoting student learning of the marine environment. We hope these lessons enable teachers to address student misconceptions and apprehensions regarding the unique organisms and properties of marine ecosystems. We intend for these modules to be a guide for teaching. Teachers are welcome to use the lessons in any order, use just portions of lessons, and may modify the lessons as they wish. Furthermore, educators may share these lessons with other school districts and teachers; however, please do not receive monetary gain for lessons in any of the modules. Moreover, please provide credit to photographers and authors whenever possible.

Given that education reform strives to incorporate authentic science experiences and our five previous teaching modules have focused on incorporating authentic experience, this last module aims to provide a deeper understanding of conducting and communicating science. Therefore, this 6th module focuses on the nature of science, specifically discussing what is science, how is science conducted, and how is science communicated. We have provided several activities and extensions within this module such that lessons can easily be adapted for various grade and proficiency levels.

Additional teaching modules, and materials such as animations, videos, and blog posts (kids and adults), will also be posted on the DEEPEND website as they become available (<http://www.deependconsortium.org/>). We hope you and your students dive into these materials and benefit from the adventure.

Sincerely,

K. Denise Kendall, Ph.D. on behalf of the DEEPEND Education and Outreach Team



Contents

Scientific Investigation	5
Observations, Questions, and Hypotheses	5
Experimental Design, Data Collection, and Data Analyses	6
Disseminating Information	8
What is the DEEPEND Consortium?	10
Background for Students	12
You are a Scientist.....	13
Experimental Design	14
Statistics	15
Visualizing Data	18
Reading Primary Literature	19
Exploring Scientific Investigations.....	21
Objective	21
Next Generation Science Standards	21
Activity Summaries	21
Scientific Investigation	21
Visualizing Data	21
Random Sampling	22
What is Probability?	22
Calculating Descriptive Statistics	22
Reliable Sources	23
Meet the DEEPEND Crew	23
Activities.....	24
Scientific Investigation	25
Visualizing Data	28
Random Sampling	30
What is probability.....	32
Calculating Descriptive Statistics	34
Calculating Descriptive Statistics	36

Reliable Sources	38
Meet some members of the DEEPEND crew!	39
Heather Bracken-Grissom, PhD	40
Chuanmin Hu, PhD	41
Jose V. Lopez, Ph.D.....	42
Laura Timm, PhD Candidate	43
References and resources for teachers and students	44

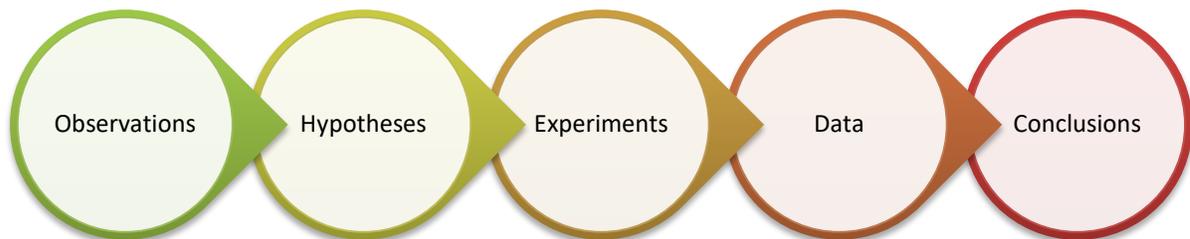
Scientific Investigation

Scientists address questions about the phenomena observed in nature. While there is no one linear process to follow when conducting a scientific investigation, there are fundamental methods that underlie scientific investigations. These fundamental methods are often referred to as the scientific method. In this module, we will explore the foundations of scientific investigations.

Observations, Questions, and Hypotheses

Most scientific investigations begin with an observation. Scientists are constantly observing the world around them, and taking note of aspects, such as behaviors and interactions. From these observations, questions often come to light. Questions are the foundation of scientific investigation, but not all questions can be tested in a scientific framework. For a scientist to pursue a question, the phenomenon observed and variables pursued must be well-defined and testable with elements that can be manipulated and managed. Keep in mind, not all scientists conduct scientific investigations, as some pursue more descriptive avenues of science.

Nevertheless, they follow underlying components of the scientific process to report unbiased information about the natural world.



After defining a question of interest, scientists speculate what may be causing what they are observing. Then, they formulate hypotheses based on their observations and prior knowledge and insights from previous scientific research. A hypothesis is an explanation to the question being investigated; the hypothesis aims to explain the phenomena being investigated. Scientists use their hypotheses to articulate what they think is responsible for a specific phenomenon, and what they expect to see in their data. These expectations, or consequences, when testing a hypothesis are referred to as predictions. Just like the question, the hypothesis must be testable. More importantly, the hypothesis must be falsifiable. Some hypotheses are written in

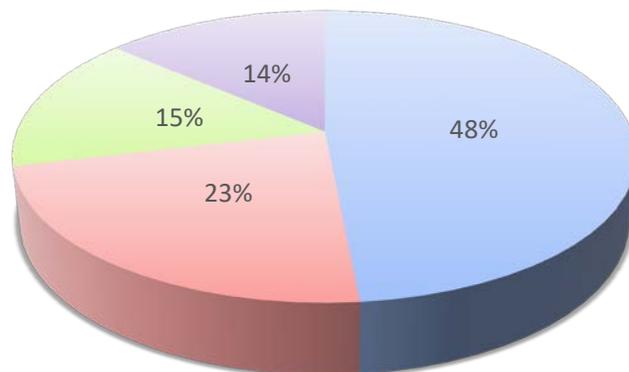
an if... then format — *If* this causes the hypothesis, *then* these consequences will be observed. Other times, hypotheses are merely a simple explanatory statement with predictions stated separately.

An explanatory hypothesis often is referred to as the alternative hypothesis because most data are statistically analyzed against a null hypothesis —a hypothesis that states there is no difference. Therefore, an alternative hypothesis states the pattern that is expected if underlying predictions are true. Thus, an alternative hypothesis cannot be a hypothesis of no effect (e.g., fertilizer runoff does not impact seaweed growth).

Experimental Design, Data Collection, and Data Analyses

Once the research question and hypothesis have been outlined, the scientist is ready to design an experiment to test the question being addressed. Just like in all previous steps of the scientific investigation, the experiment must be well defined. Each experiment must consider multiple variables (independent, dependent, and control) as well as the treatments (experimental and control) so that the scientist can draw the most accurate and unconfounded conclusions from the data. An independent variable is any variable that is changed or manipulated to observe a response; this variable leads to the experimental treatments. A dependent variable is any variable that is affected by the experiment; it is measured, counted, or otherwise observed for a response. A control variable is any condition or factor that is kept the same across all treatments. A control treatment serves as a baseline for scientists to compare experimental treatments to; therefore, this is the group that is not manipulated in the experiment.

Data are collected via an experiment that is designed to test the question and variables. There are two main types of data: quantitative and qualitative. Quantitative data are data consisting of numbers that are the result of counting or measuring (e.g., there are 20 ants on the playground), whereas qualitative



data are data that describe types using words or letters (e.g., ants are present on the playground).

Quantitative data can be further sub-divided into categorical (discrete) or continuous data. Categorical data are based on counting in distinct, mutually exclusive categories. With categorical data, there can be no intermediates. An example of this is that a species is either present or absent from the study site. Meanwhile, continuous data are data that stem from taking measurements on a numerical scale that has an obvious order. For example, an investigator could measure the body length of each ant on the playground. Other examples of continuous data are temperature, height, and pH. In continuous data, there can be intermediate values, such as 3.7 mm in the case of body length of an ant.

After data collection is completed, scientists analyze the data to draw conclusions regarding the phenomena they are investigating using statistics. Statistics is a branch of mathematics involving the analysis and presentation of data.

There are two primary types of statistical analyses employed by scientists: descriptive and inferential statistics. Descriptive statistics are used to summarize and present data in an informative way. Scientists often use descriptive statistics to describe trends within the data. Central tendency statistics are the most commonly used descriptive statistics. Measures of central tendency includes the mean (average value), median (middle value), mode (most common value), and variance (dispersion) of the data set.

It is rare for scientists to census an entire population. There are many reasons including that it may be difficult to define the population, the population size may be too large, or logistically it may not be feasible to account for the entire population. Therefore, scientists take a sample from the population of interest. A sample is simply a subset of a population; a statistic (summary value of the sample data collected) characterizes the sample. A statistic can change with repeated sampling of a population. In rare instances where an entire population is sampled, a parameter summarizes the data. The parameter does not change with repeated sampling of the population because the entire population has been censused. From a sample, scientists can draw conclusions and make decisions about the population by making inferences from the data. Inferential statistics are used to make an inference using the data collected; they allow scientists to draw conclusions and extrapolate from the data collected. Inferential

statistics allow scientists to determine something about a population, based on the sample collected. Inferential statistics are based on the principle of probability. Specifically, inferential statistics are based on the premise of incorrectly rejecting the null hypothesis, also referred to as Type I error. Scientists set a level of significance, or alpha value when evaluating their data. Oftentimes the alpha value is set at 0.05, which indicates a 1 in 20 chance that a true null hypothesis is rejected. Sometimes, scientists will be more stringent and set the alpha value even lower at 0.01 where there is even less chance of rejecting a true null hypothesis. In cases where scientists reject the null hypothesis, they are saying that there is a discrepancy that is unlikely to be the result of chance alone. Some basic statistical tests that are commonly used for inferential statistics include the t-test, correlation, analysis of variance (ANOVA), and chi-squared test. The statistical test used by a scientist is determined by the type of data collected. For example, chi-squared tests are used to determine the relationship between two categorical variables, whereas t-tests are used to make comparisons of the means of two groups of continuous variables. From the results of inferential analyses, scientists either reject their hypothesis or find support for the hypothesis.

Scientists often form new questions from the results obtained by a study. This is because scientists sometimes come to unexpected conclusions after collecting and analyzing data from an experiment, other times it is because scientists observed another phenomenon during their experiment.

Disseminating Information

Scientists communicate their findings to other scientists and the public via oral and written forms. Scientists teach and attend conferences where they share their findings with others. Many scientists publish primary research articles that are written primarily for other researchers in the field to disseminate research findings globally. Some primary research articles are rewritten, or summarized, as popular science articles. Popular science articles are written with the general public as the primary audience; therefore, they are condensed versions of the primary research highlighting the key findings and written for someone who is not an expert in the field.

All scientists must be able to locate, read, and understand scientific literature, both review and primary research articles. Review articles summarize data and conclusions from many studies, whereas primary research articles summarize data and conclusions from an original

experiment. Review articles are an excellent starting point for scientists who want to gauge what is and is not known about a specific phenomenon. Primary research articles can be difficult to read and understand by individuals who are not experts in the field, but with practice and persistence proficiency can be reached. We will focus our attention on primary research articles as they are more prominent in most scientific fields.

Most primary research articles contain six sections, usually in the following order: abstract, introduction, materials and methods, results, discussion, and literature cited. However, each journal has unique guidelines for publications, so some articles may have more sections and others less, and some of these sections may be in a different order.

The abstract is a succinct summary of the entire paper. Typically, this section highlights the research question(s), hypothesis, notable results, and conclusions. Most abstracts are 300 words or less, so scientists must condense their research goals, methods, and major findings into a few key sentences. Minute details are left out of the abstract due to length limitations. From the abstract, the reader can determine what the study is about and if the study addresses topics they are interested in knowing more about. If so, the reader would continue reading the paper.

The introduction provides background information necessary for understanding the study; it sets the stage for the study. Specifically, it describes the phenomenon being investigated. This section will highlight prior research conducted in the field, and citations will be included for these studies (they are great resources for additional background information). The author will also explain how the current study will expand or clarify knowledge in the field. While replication is important in the sciences, we rarely see scientists publishing primary research articles that simply replicate a previous experiment for the same purposes and by the same means. During and after reading the introduction, scientists will look up any unclear terms or concepts before proceeding to the remainder of the paper because a clear understanding of the phenomenon being investigated will make the experiment and findings more comprehensible.

The materials and methods section often follows the introduction. In this section, the authors describe how the research questions were answered. This will include the technical details for experiment methods and/or field work, or they will reference previously used techniques. Statistical methods will be outlined as well as any specific equipment, programs, and software

employed to assist with data acquisition and analyses. This section typically provides enough information to replicate the research study. Some journal articles will move this section to a “supplementary information” section online or to the end of the journal article.

Data are presented in the results section. Data are highlighted by employing figures, graphs, and tables as well as verbal descriptions. The researcher does not interpret the data in this section, instead the data are presented in a manner that the reader can easily understand. General trends of the data are described; it is uncommon for raw data values to be presented. Instead, descriptive statistics often are utilized to provide the reader with data insight. Results of inferential statistics are also reported in this section.

In the discussion section, the authors summarize their conclusions and relate them to the research question(s). This is the section in which the data are interpreted. The findings and conclusions are put into context by comparing them to other research studies (this is another great place to find other articles of interested on the topic). Sometimes the authors will address shortcomings of their study, or questions that remain unanswered.

Most research articles are written by experts in the field with the intention of conveying information to other experts in the field. It is common to take notes and look up terms and methods that might be unfamiliar while reading an article; most scientists do. Sometimes its easiest to write notes and thoughts on the article while reading.

What is the DEEPEND Consortium?

The DEEPEND Consortium is a network of scientists, educators, and ship captains and crewmembers from NOVA Southeastern University, University of South Florida, Florida International University, Texas A&M University, Florida Atlantic University, U.S. Naval Research Laboratory, The National Systematics Laboratory, San Antonio Zoo, and WhaleTimes. The sole mission of this collective team is to characterize the oceanic ecosystem of the Gulf of Mexico, specifically the northern region near the United States mainland. The DEEPEND scientists aim to assess and characterize the overall ecology, genetics, acoustics, oceanography, and biochemistry of the region. The data obtained by the DEEPEND team will provide baseline information for future scientific investigations. The importance of such data was highlighted during the Deepwater Horizon oil spill in 2010 before which extensive baseline data did not

exist. Consequently, the work of the DEEPEND team has been funded by a grant from BP and The Gulf of Mexico Research Initiative. Data and information obtained by the team is not solely for scientific use, but also aims to promote public awareness and involvement in the ecosystem.



Background for Students



You are a Scientist

Who is a scientist, and what does a scientist do? A scientist is a person who uses a systematic process to ask and answer questions about the world that surrounds him/her. This means that you can be a scientist!

We often think of a scientist as someone who wears a lab coat and uses words that we cannot understand, but that stereotype doesn't begin to capture the diversity of scientists around us. This is because being a scientist is based on how individuals think about the world around them; it does not matter what they wear, what language they speak, or how they look.

Scientists engage in scientific investigations to unbiasedly explore and describe the world. Let's briefly explore the underlying principles of scientific investigations:

Observation	<ul style="list-style-type: none">• Documenting what is seen in the natural world.
Research Question	<ul style="list-style-type: none">• Forming a question based on observations.• Must be well defined and scientifically testable.
Hypotheses	<ul style="list-style-type: none">• An explanation to the question being investigated - what is causing the phenomenon?• Has associated predictions.
Experimental Design	<ul style="list-style-type: none">• Outlines what kind of data will be collected.• Explains how data will be collected.
Data Analyses	<ul style="list-style-type: none">• Drawing conclusions from data.• Utilizes statistics.
Dissemination	<ul style="list-style-type: none">• Sharing of knowledge with others.

It is uncommon for scientific investigations to be linear, often scientific investigations require revisiting and repeating components of the process.

Experimental Design

Prior to conducting an experiment, a scientist must carefully design how they will carry out the study. Each experiment must consider multiple variables (independent, dependent, and control) as well as the treatments (experimental and control). It is important to consider these components so that the most accurate conclusions can be drawn about the phenomenon.

An **independent variable** is any variable that is changed or manipulated to observe a response; this variable leads to the experimental treatments. For example, changing or manipulating the number of flowers available for bees to pollinate.



A **dependent variable** is any variable that is affected by the experiment; it is measured, counted, or otherwise observed for a response. Sometimes studies have a single dependent variable; in other studies, there may be multiple dependent variables. If we continue with our example of manipulation of number of flowers, we could measure pollination rates or bee abundance in the different treatments.

A **control variable** is any condition or factor that is kept the same across all treatments. An example is ensuring flowers are the same species.



A **control treatment** serves as a baseline for scientists to compare experimental treatments to; therefore, this is the group that is not manipulated in the experiment.

There are two main types of data collected by scientists: quantitative and qualitative. Quantitative data are data consisting of numbers that are the result of counting or measuring (e.g., there are 20 bees), whereas qualitative data are data that describe types using words or letters (e.g., bees are present).

Statistics

Scientists use statistics to analyze data in order to draw conclusions about the phenomena they are investigating. Statistics is a branch of mathematics involving the analysis and presentation of data.

There are two primary types of statistics used by scientists: descriptive and inferential statistics.

Descriptive statistics are used to summarize and present data in an informative way. Scientists often use

descriptive statistics to describe trends within the data. Measures of central tendency are often used in descriptive statistics, such as the mean (average value), median (middle value), mode (most common value), and variance (dispersion) of the data set. How are these values calculated?


$$19 + 1 = 20$$

The **mean** is the average of the data set and provides us with information about the central point of the data. The mean is calculated by adding all the data values (data points) and then dividing that sum by the total number of data values. For example, consider this example data set: 1, 2, 2, 3, 4. The mean is calculated by adding $1+2+2+3+4$ which equals 12 and then dividing by the number of data values, 5. Thus, $12/5$ which equals 2.4.

The **median** is the middle value of the data set. To calculate the median, start by ordering all your data from the lowest to highest data value. If there is an odd number of data values, then the median is simply the middle number. If there are an even number of data values, then the median is the mean of the two central data values, which can be calculated by adding the two central data values and dividing by two. Using the example data set above, the median is 2.

The **mode** is the most common number in the data set. This value is determined by simply looking for which data value occurs most often. If no data values are repeated, then the mode is equal to zero. Using the example data set above, the data value 2 is found twice, and, therefore, is the mode.

Measures of dispersion describe how different data values occur with respect to one of the central tendency points (typically the mean). The simplest calculation for dispersion is the **range** which represents the span of your data, essentially the difference between the largest and smallest data value. Range is calculated by subtracting the minimum data value from the maximum data value. Using the example data set above, the range is $4 - 1 = 3$.

Another measure of dispersion is sample variance. Sample variance is the degree of variation around the mean. It is calculated by summing the difference of each data value from the mean, squaring that value, and then subsequently dividing by the number of samples minus one.

Another calculation of dispersion is standard deviation which quantifies the scattering of the data values in the sample. Essentially, it tells us the scattering of the data values from each other; how much do they vary from one another. It can be calculated by taking the square root of the sample variance.

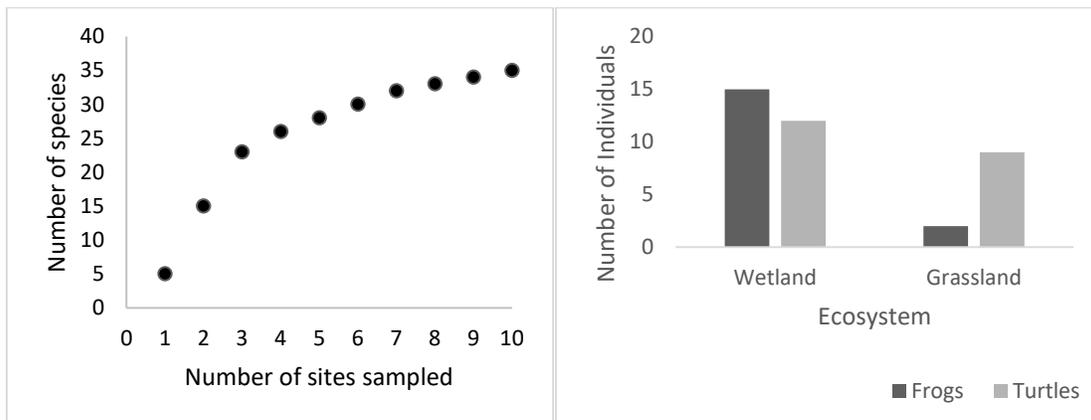
The last common measure of dispersion is standard error. Standard error quantifies the dispersion of the data around the mean. We can calculate this value by dividing the standard deviation by the square root of the number of data values. It is an indicator of confidence, if there is a large standard error then the values are widely spread around the mean. In this case, there is a lot of inconsistency in the data. However, if the standard error is small then the values are clustered around the mean and you can be more certain about the data. Standard error will typically decrease as sample sizes increase, this is because

typically the more samples you quantify the more representative the data is of the population.

Inferential statistics are used to make an inference using the data collected; they allow scientists to draw conclusions and extrapolate for the data collected. Specifically, inferential statistics allow scientists to determine something about a population based on the sample being studied. Inferential statistics are based on the principle of probability. Scientists set a level of significance, or alpha (α) when evaluating their data. Oftentimes alpha is set at 0.05, which indicates a 1 in 20 chance that a true null hypothesis is accidentally rejected. In cases where scientists reject the null hypothesis, they are saying that there is a discrepancy that unlikely to be by chance. This means that there is a difference in what they are seeing. From the results of inferential analyses scientists either reject their hypothesis or find support for the hypothesis. Some common inferential statistical tests use by scientists include: Student's t-test, one-way analysis of variance (ANOVA), correlation, chi-squared test, diversity indices, Kruskal-Wallis test, and Mann-Whitney U test. Many of these analyses can be completed using online calculators or Microsoft Excel, but scientists oftentimes use more specialized programs dedicated for statistical analyses, such as R, SPSS, and SAS.

Visualizing Data

After data collection, scientists often graph their data. This allows them to visualize, or see, potential trends in the data set. They may use line graphs, bar graphs, histograms, pie charts, among other graph formats, depending on the types of data collected.

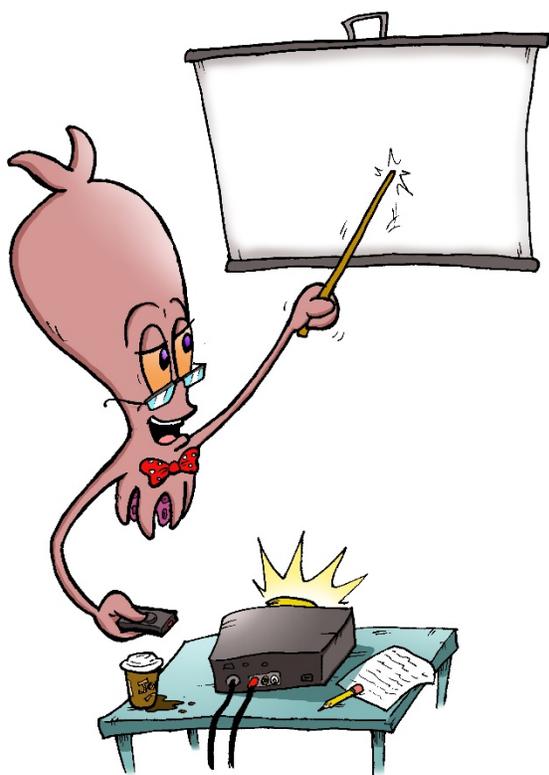


Visualizations also enable a scientist to see the distribution of data values (e.g., normal, skewed, bimodal). A **normal distribution** occurs when there is a unimodal, bell-shaped curve or pattern for a continuous variable. Data values are equally likely to occur on one side of the mean as the other. A skewed distribution occurs when data values are more likely to occur on one side than the other side of the mean. This results in an asymmetrical pattern with many data points on one side and a tapering tail of fewer data points on the other. A bimodal distribution occurs when there are two peaks in the data set, on both side of the mean. In this distribution, there are less data points at the mean than at values above and below the mean.

Visualizations of the data allow scientists to not only see the overall trends, but also individual data points. This is essential when scientists are looking at variation, or for outliers in their data. Sometimes there can be error in transcribing data, while other times this natural variation exists. Accurate data is the foundation to draw conclusions later, therefore this step in the scientific process is essential.

These visualizations allow scientist to begin drawing conclusions about their data, but it's still essential for them to conduct statistical analyses.

Reading Primary Literature



After scientists conduct experiments and draw conclusions they share their findings, usually by giving a presentation or writing a paper. All scientists must be able to locate, read, and understand scientific literature, both review and primary research articles. Review articles summarize data and conclusions from many studies, whereas primary research articles summarize data and conclusions from an original experiment. Review articles are an excellent starting point for scientists who want to gauge what is and is not known about a specific phenomenon thus far. Primary research

articles can be difficult to read and understand by individuals who are not experts in the field, but with practice and persistence proficiency can be reached. Let's outline the main sections of a primary research article.

Most primary research articles contain six sections, usually in this order: abstract, introduction, materials and methods, results, discussion, and literature cited. However, each journal has unique guidelines for publications, so some articles may have more sections and others less and some of these sections may be in a different order.

The **abstract** is a succinct summary of the paper. This section highlights the research question(s), notable results, and conclusions, all in just a few hundred words. When reading the abstract ask yourself if the paper addresses the phenomenon you are interested in learning more about.

The **introduction** provides background information necessary for understanding the study; it sets the stage for the study. It provides information on what is already known about the phenomenon and how this study will contribute/expand on this knowledge. While reading this section, look up concepts that are unclear to you or what words are unfamiliar. A solid foundation of the background information will make the remainder of the paper easier to understand.

The **materials and methods** section often follows the introduction. This section describes how the study was conducted. The scientist will discuss methods used to collect data, where the data were collected, and how the data were analyzed. Enough information is provided so others can replicate the study, this is because replication is important in the sciences.

Data are presented in the **results** section. In this section, general trends of the data are described, it is uncommon for raw data values to be presented. The data will be described both visually and in sentence form.

In the **discussion** section, the authors summarize their conclusions and associate them to the research question(s). Sometimes the authors will address shortcomings of their study, or questions that remain unanswered.

Exploring Scientific Investigations

Objective

Students will take on the role of a scientist. The primary objective of all of these activities is to make students aware that they are all capable of being scientists.

Next Generation Science Standards

The practices, lessons, and concepts outlined in this module underpin all areas of the Next Generation Science Standards for middle and high school students. From our core understanding of science, historical and current practices, and the cross disciplinary applications of these ideas we see the applications of the nature of science.

Activity Summaries

Scientific Investigation

This activity helps students understand the primary stages of scientific investigations by having them ask a question, design an experiment, and show their expected results. It is important to clarify to students that scientific investigations are not linear; however, oftentimes scientists revisit previous steps in order to gain a better understanding of the phenomenon they are exploring. An activity I enjoy doing with my students prior to this activity is to have them go outside, even if it's just on school property, and observe the world around them. I ask students to spread out around the school and make a list of ten questions that come to mind while observing the natural world around them. Then, I have students form groups in the classroom to discuss their questions and determine if they are scientifically testable.

A great extension to this activity is to have students design investigations they can actually conduct. Let them be scientists, including having to revise their hypotheses and experimental designs. A great underlying component of this extension is that some students will be faced with reconciling misconceptions.

Visualizing Data

This activity teaches students about the basics of interpreting graphs presented in textbooks, primary literature, and other publications. Students first learn about graphs, then they are asked to apply what they learned to a graph presented in a paper published by members of the DEEPEND crew. An extension to this activity is to have students read the paper. This extension is most appropriate for advanced or upper level high school students. Another extension would be to have students graph their own data. By conducting simple experiments or surveys in the

classroom students can obtain data. They can then choose how to best represent their data and share it with peers.

Random Sampling

Students are often perplexed by the idea of random sampling. This activity starts off by having student sample from a population of Skittles. An added bonus is they can eat their population when finished! This activity just asks them to take one random subsample and compare it to the entire population. A very brief activity that would provide students with exposure to random sample. If you are not time pressed, you could easily have students take multiple samples (with and without replacement) and compare results. Students could be asked to summarize their data both visually and using words, thereby reiterating data representation. Then, students could be asked to present their findings to other groups for comparison.

The second portion of this activity invites students to be members of the DEEPEND crew. They are out on a cruise and must interpret data and determine if it representative of the population in the area. Students are tasked to think like a scientist about their data. An extension to this activity would be to have students research diversity data within the Gulf of Mexico.

What is Probability?

This activity revisits basic principles of probability. While probability is addressed many times throughout K-12 education, we find students are still confused by it during their undergraduate experiences. Therefore, this activity revisits the basic principles of probability using items students are familiar with: coins and dice. There are many online simulations that can be used to shorten the length of this activity, but the tangible experience of flipping the coin and collecting data provides a unique experience regarding data collection. Students are asked to not only consider the probability of an event occurring, but also to visualize it using graphs.

For students who understand these basic principles, this activity could be expanded to address AND and OR rules. For instance, students can be asked to calculate the probability of rolling a 3 or 6 when rolling two dice. Then they can carry out a simulation and analyze the data. Then, students can be asked to calculate the probability of rolling a 3 and a 6 when rolling two dice. Once again, they can complete a simulation and analyze the data.

Calculating Descriptive Statistics

Students are asked to calculate basic descriptive statistics for two data sets in this activity. Students can be asked to complete the first data set in pairs and then work interpedently on the second data set to determine proficiency. In addition to calculating these basic statistics,

students are asked to graph their data. This is important as scientists often visualize their data prior to completing any statistical analyses as this provides an overview of the data.

More advanced students can collect data (or be given data) and conduct inferential statistical analyses on the data. There are numerous free online statistical packages, and Microsoft Excel, that can be used to analyze data.

Reliable Sources

Students are bombarded by information on the internet. Some of this information is reliable, whereas others are not. This activity asks students to evaluate the reliability of three articles of their choice. Students read the articles while concurrently answering questions about the articles. Then, students are asked to determine how many of their initial three articles they would consider reliable sources. This activity is a great brainstorming activity prior to discussing finding resources with students as it asks students to first draw their own conclusions. An extension of this activity would be to provide students with a primary research article and a corresponding popular science article. Then have students compare the two publications and determine if there are any misconceptions in either article.

Meet the DEEPEND Crew

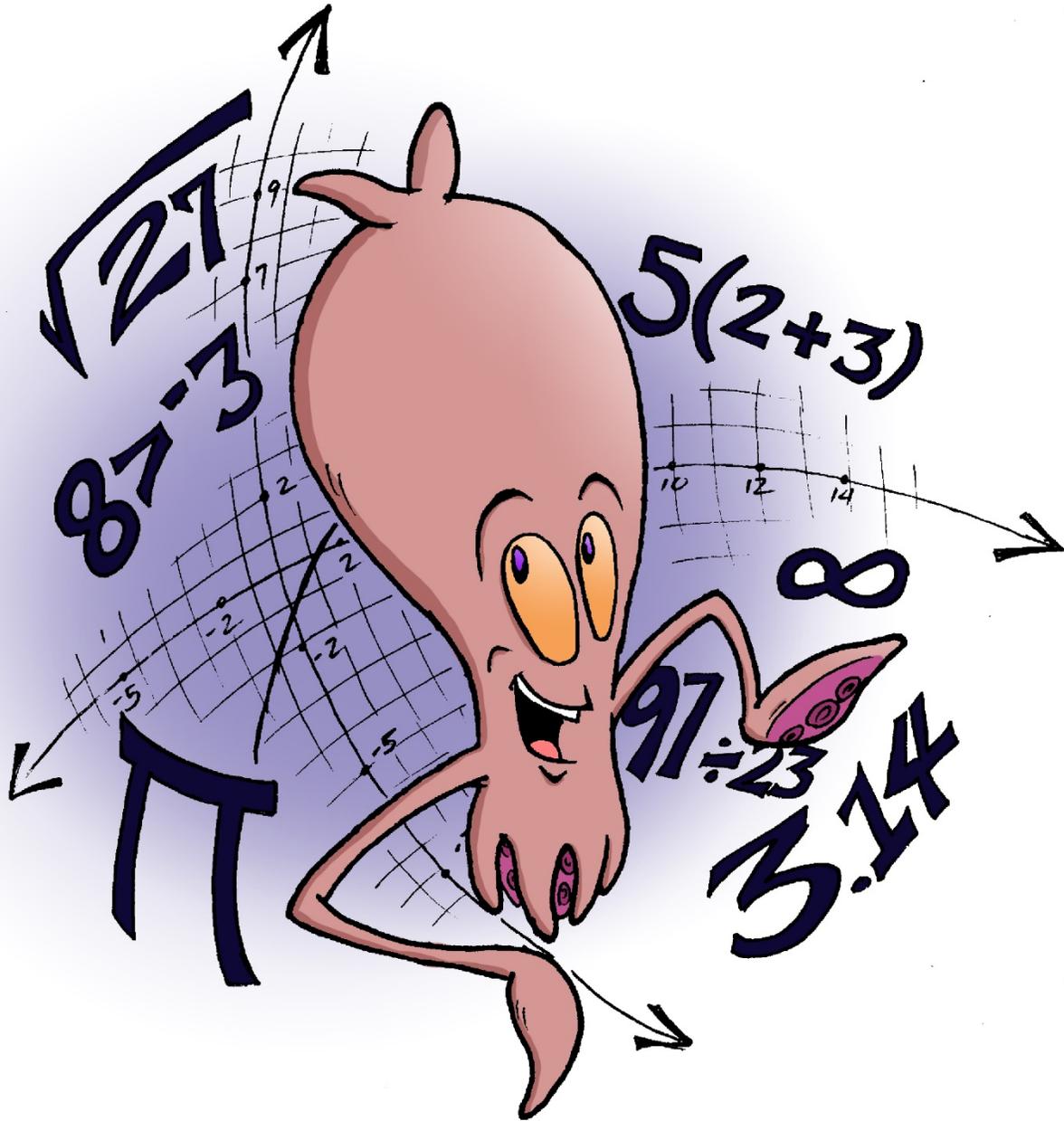
This section introduces students to some members of the DEEPEND crew. Information about other members of the DEEPEND crew can be found on the DEEPEND consortium website:

<http://www.deependconsortium.org/>

The DEEPEND crew consists of individuals from a diversity of backgrounds and professions. The objective of introducing students to members of the DEEPEND crew are to: 1) get them interested and curious about the STEM fields; and 2) make students aware that anyone can pursue a career in the STEM fields. In addition to introducing members of the DEEPEND crew, students can be encouraged to visit websites of the affiliated institutions and research programs. Students can also be encouraged to read publications that come from the DEEPEND consortium and affiliated institutions.

A great extension to introducing students to members of the DEEPEND crew is to make connections with local higher education institutions. Oftentimes, graduate students, faculty members, and outreach coordinators will be happy to speak with your classes about their work. Sometimes these discussions will be in person, and other times via Skype. These interactions allow students to learn more about the STEM fields and to ask questions. All it takes is an email or phone call to build a relationship with local scientists.

Activities



Scientific Investigation

1. What phenomenon would you like to explore? Why?
2. Describe any observations you have about this phenomenon.
3. What questions do you have about your observations?
4. Write a hypothesis for one of your questions.

5. Describe how you would collect your data to test your hypothesis. Explain your variables and treatments.

6. What would your data look like if your hypothesis is supported?

SCIENTIFIC INVESTIGATION

QUESTION:

Empty rounded rectangular box for the question.

HYPOTHESIS:

Empty rounded rectangular box for the hypothesis.

MATERIALS:

Empty rounded rectangular box for materials.

METHODS:

Empty rounded rectangular box for methods.

OBSERVATIONS:

Large empty rounded rectangular box for observations.

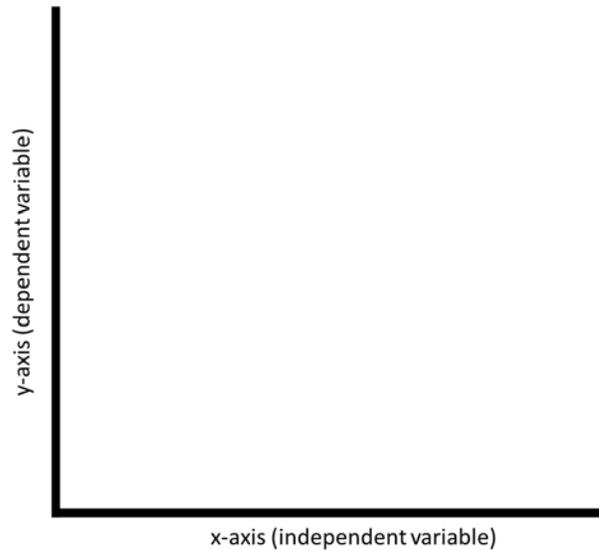
CONCLUSIONS:

Large empty rounded rectangular box for conclusions.

Visualizing Data

It is important that scientists can both read, and generate, visual representations of their data. These visualizations can include graphs, photographs, and models. Let's practice some of these skills now, focusing specifically on graphs.

First and foremost, it is important to determine what the author is showing in the graph. Read the caption, take note of any legends, and identify the axes. Almost every graph has axes, specifically an x-axis and a y-axis. The x-axis runs horizontally and typically represents the independent variable, while the y-axis runs vertically and typically represents the dependent variable. Sometimes, the x- and y-axes are swapped depending on the data being represented and how it is visualized. With just this information you should be able to simply state what the graph is showing.



Now take a closer look at the axes. Each axis will have units associated with it and will have a specific scale. How were the data measured and what range do the data points span?

Not only do graphs provide insightful information, but they also allow scientists to portray a lot of information in limited space. Now that you know what the graph is showing, take a look at the overall patterns present. Are outcomes for each treatment equal, or is one more or less than others? How much variation is there within each treatment relative to any differences seen? Are there any patterns in relationships – positive, negative, u-shaped, others? Are any data points outliers? How are the data points distributed – normal, skewed, bimodal?

Lastly, does the data represented in the graph provide support for the hypothesis being studied? Does it align with the predictions? Make your own summation and inferences from the graph prior to taking a closer look at data analyses and conclusions.

Apply what you've learned about reading graphs by interpreting the graph below. This graph comes from a manuscript written by DEEPEND consortium members and colleagues. If you're interested in learning more about the data, please refer to the original article:

Judkins, H., Vecchione, M., Cook, A., and Sutton, T. 2016. Diversity of midwater cephalopods in the northern Gulf of Mexico: comparison of two collecting methods. *Marine Biodiversity*.

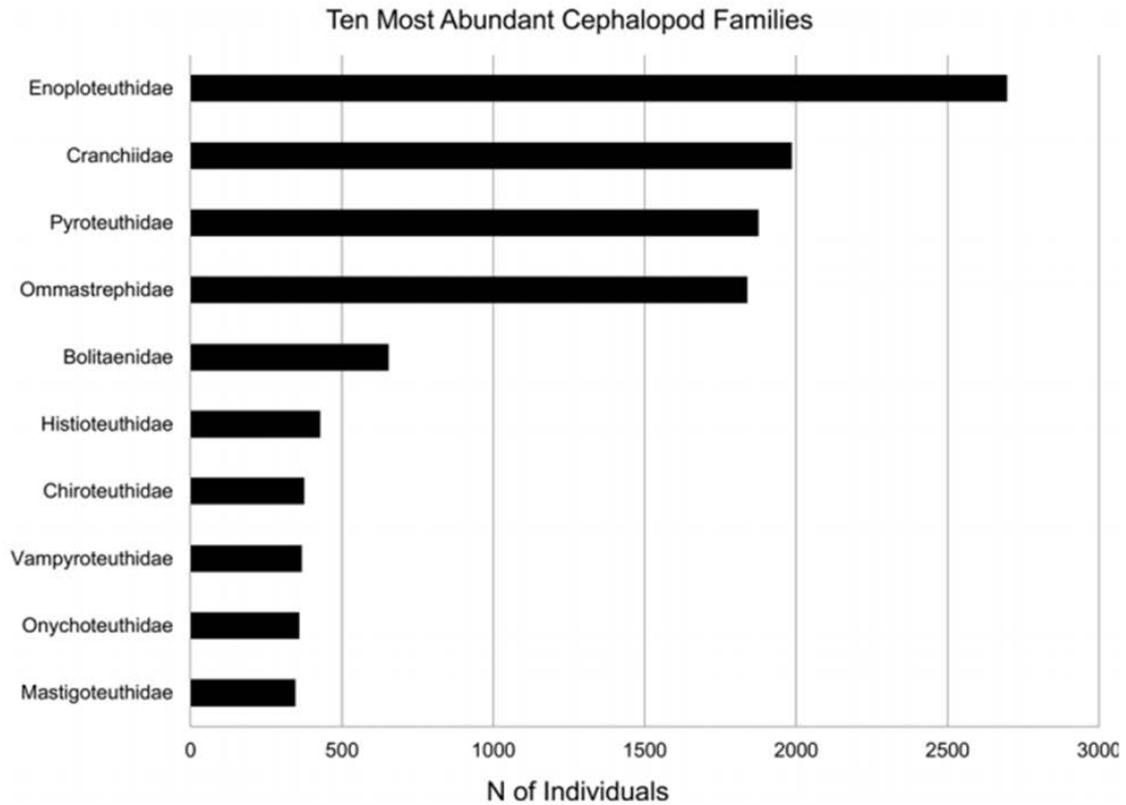


Fig. 2 Number of individuals for the ten most abundant cephalopod families from ONSAP

What is represented in this graph?

Do you notice any trends in the data?

What conclusions can you draw from this graph?

Random Sampling

It is impossible for scientists to collect data on every individual in a population or for every component of every system. Therefore, scientists take information on a random subset of the population or areas and extrapolate information from these data. Ideally when selecting a random sample, each individual or area of a larger population are equally likely to be selected for sampling. Scientists must be unbiased in their selections, this means they are not looking to include only data points that support their hypothesis. This is because they want the sample to be representative of the larger population or area.

Some random sampling techniques include:

- Simple random sampling – each member has the same chance of being selected.
- Systematic sampling – random sampling in a systematic manner, such as selecting every tenth individual or measuring at every three meters.
- Stratified sampling – subdivision of the population and randomly sampling within each.
- Cluster sampling – subdivision of the population or system into geographic sections and sampling within each.

One of the quickest simple random sampling methods used is to generate a random number list that represents the individuals or areas that are then sampled.

Let's practice systematic random sampling. Fill a bag with Skittles. This is your population. Randomly draw out each individual and report data (the color) for every 5th individual you draw.

Color	Tally	Skittles (%)
Green		
Yellow		
Orange		
Red		
Purple		

Now count every individual in the population (the entire bag) and record the data below.

Color	Tally	Skittles (%)
Green		
Yellow		
Orange		
Red		
Purple		

Was your smaller random sample representative of the entire population of Skittles?

You've been invited out onto a cruise with the DEEPEND crew! You are assigned the task of determining shrimp populations within the water column. You need to sample from the surface to 100m deep in the water column. After much discussion with colleagues, you decide that you will sample every 10m in the water column. In order to obtain a representative count, you have also decided to sample each region for 60min while traveling a steady rate of 5mph. Here is the data you obtain:



Depth (meters)	Number of individuals
0 (surface)	0
10	5
20	12
30	33
40	40
50	25
60	12
70	21
80	8
90	7
100	5

Use this data to draw a graph of the vertical distribution of shrimp in the water column.

How many total shrimp were caught in this study?

Do you think that your data include every individual of shrimp living in the area? Yes or no, explain your answer choice.

Do you think your data could be representative of the shrimp population in the area? Yes or no, explain your answer choice.

Propose other ways you could have randomly sampled for shrimp distribution.

What is probability

At its core, probability is the likelihood that something will (or will not) occur. Consider a fair coin with two sides — heads and tails. What is the likelihood if the coin were flipped that it would land heads up? $\frac{1}{2}$ or 50%. We obtain this value by taking the number of ways to get heads and dividing it by the total number of different outcomes. Typically, the probability an event will occur is denoted as $P(\text{event})$, where $P(\text{event})$ is the probability of the event occurring and; for example, $P(\text{heads}) = 0.5$ or $P(H) = 0.5$. The probability for any event will always lie between 0 and 1; moreover, the sum of all possible outcomes will always equal 1.

Let's make things a bit more complex. What is the probability of rolling a 3 on a fair die with six sides numbered 1, 2, 3, 4, 5, and 6? You should determine the probability is $\frac{1}{6}$. What is the probability of obtaining an even number? Show your work.

Just because you expect an outcome a certain proportion of the time, does not mean that you will acquire exactly that outcome each time you sample. This is due to variation in sampling and in the natural population. Let's go back to our coin. You expect to have it land on heads 50% of the time. Flip a coin ten times, how many times does it land on heads?



Now flip the coin another ten times (for a total of 20 flips), how many times does it land on heads this time?

Repeat the coin flipping again; you should now have a total of 30 flips. How many heads this time?

Just because we expect a specific outcome to occur so often, does not mean that will always be the case, as natural variation will occur. You are more likely to get closer to the expected probability the larger your sample size. Flip the coin 100 times, how many times does it land on heads?

Now 200 times, how many times does it land on heads?

Draw a graph showing the frequency of heads obtained for 10, 100, and 200 flips. What do you notice?

Compile your coin flip data with that of your classmates. How many times does the coin land on heads?

Calculating Descriptive Statistics

As a member of the DEEPEND crew you are interested in the surface temperature of the water off the shore of Florida. You record surface temperature over the course of a year and obtain the following data:

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Temperature (°C)	13	21	24	25	29	30	30	31	28	26	24	22

Calculate the mean:

Calculate the median:

Calculate the range:

Calculate the mode:

Calculate the sample variance:

Calculate the standard deviation:

Calculate the standard error:

Calculating Descriptive Statistics

As a member of the DEEPEND crew you are interested in the number of copepods captured during DEEPEND cruises. You record the number of individual copepods captured over the course of a year and obtain the following data:

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Number of copepods	9	22	30	17	12	15	7	52	23	17	10	12

Calculate the mean:

Calculate the median:

Calculate the range:

Calculate the mode:

Calculate the sample variance:

Calculate the standard deviation:

Calculate the standard error:

Create a graph:

Reliable Sources

Scientists communicate their findings not only to other scientists, but also to policy makers, educators, and the general public. Dissemination of information is vital to the enhancement of our overall understanding of the natural world. An essential process of scientific written communication is peer review. Through peer review, the work of scientists is evaluated by other experts in the field. These experts aim to be impartial in their evaluation and strive to determine that the work meets scientific standards and that the conclusions are supported by the data obtained.

Even though most scientific communications are evaluated prior to publication, it is still essential that you evaluate all sources you are obtaining information from. Using Google, find three articles about a topic of your choice. Write their citations here:

- 1.
- 2.
- 3.

Read each of the articles. While reading the articles answer the following questions for each:

1. Does this article present current information? When was it published?
2. Who published the article?
 - a. What is the education level of the author(s)?
 - b. What other articles has the author(s) published?
3. Is the article peer reviewed?
4. Is the article a primary source?
5. Are references cited? What kind of references are these?
6. Is the article biased?
7. Is the source credible?

Of the three articles you selected, how many would you deem reliable?

What factors determined which articles you deemed reliable versus unreliable?

Meet some members of the DEEPEND crew!



Heather Bracken-Grissom, PhD

Dr. Bracken-Grissom's interest in marine biology began in the 8th grade. Now, Dr. Bracken-Grissom studies the genetic diversity and connectivity of midwater crustaceans. In fact, her favorite species in the Gulf of Mexico is a deep-sea shrimp called *Systellaspis debilis*, which has the ability to produce bioluminescence.

The most exciting component of her work is discovering new species and exploring novel habitats. Dr. Bracken-Grissom's work provides insights regarding the sustainability of our planet, as it is reliant on the ocean and their inhabitants. By understanding the organisms that live in the ocean, how they are related, and how they persist is needed for a healthy ocean and thus a healthy planet.

While Dr. Bracken-Grissom's interest in science started in middle school, she persisted primarily due to great support and mentorship. Her advice to you is never stop exploring because begin a scientist takes passion and hearth. She also thinks being a scientist is a great career choice if you really want to make a difference in the world.



Chuanmin Hu, PhD

Dr. Hu is a trained physicist specializing in measuring and modeling light. Dr. Hu finds light in nature extremely fascinating; have you ever noticed that the sky isn't always the same shade of blue and it isn't always homogenous in color? These details are natural phenomenon that Dr. Hu investigates! His training and expertise regarding light allows him to study algae blooms, oil spills, and water transparency in oceans, estuaries, and lakes.

Dr. Hu's expertise has applications for monitoring oil spills. How does he do this? Well, Dr. Hu can use the glint of the sun over the ocean (direct reflection of the sun) to search for oil. With this information, he can also estimate the severity of an oil spill. He can also apply this technique to assess human environmental impacts from on/in water uses. Such a novel technique!

Not only does Dr. Hu contribute to oil spill research, but he also applies these techniques to monitoring and detecting algal blooms. Algal blooms are a rapid increase in growth rate and accumulation of photosynthetic species. This information is vital in order to provide the general public and scientists with information about potential marine hazards.

Dr. Hu's research is not only applicable for the general public and scientific community, but also to anyone navigating vessels on water ways. The information Dr. Hu gleans from water temperature and transparency can be used to safely guide ships by providing information regarding water depth.

It's amazing how many applications there are from scientific work. Dr. Hu is always impressed and surprised by the diversity of life in the Gulf of Mexico. He says to always keep curious and be logical in everyday life if you also want to pursue interested in the sciences.

Jose V. Lopez, Ph.D.

Dr. Lopez studies marine microbial communities in the Gulf of Mexico including the bacterial symbionts of angler fish. A symbiosis is a long-term, close physical interaction between individuals of two different species in which both benefit. Dr. Lopez's passion in the field of genetics allows him to determine and understand DNA and protein sequences obtained from living organisms. Not only does he analyze these sequences, but he is able to use these analyses to infer evolutionary relationships. Dr. Lopez infer evolutionary relationships of symbionts, but he also identifies and locates pathogenic organisms and assesses seawater quality to determine ecosystem health.

The broader impacts of Dr. Lopez's work can provide information about how human recreational use of the Gulf of Mexico alters the ecosystem. Humans enjoy lying on the beach, but they also enjoy swimming in and recreating on the waters.

Dr. Lopez has been passionate about evolutionary genetics for years and is excited that he has been able to pursue his passion as a career. His advice to others who are interested in the sciences is to follow your passion whether it is focusing on a specific organism or the broad applications; if you do this then the world is your oyster.

Dr. Lopez has had the opportunity to venture out on the Gulf of Mexico many times with the DEEPEND crew. During these ventures, he has enjoyed seeing the diversity of organisms hauled up by the MOCNESS net. He enjoys seeing the organisms that come from the depths of the Gulf of Mexico, but also fishing near patches of sargassum seaweed (check out the neat images to the right! © J. Lopez). By fishing or netting near these patches, he can see what other life these patches support. Another memorable moment on the Gulf of Mexico for Dr. Lopez is watching flying fish with large pectoral fins swimming next to the boat during the night. He said they appear eerily luminescent.

Doesn't Dr. Lopez's work sound interesting? What else would you be interested in learning from Dr. Lopez?

Check out more here:

<http://www.deependconsortium.org/index.php/research/ecology/microbial-community-ecology>



Laura Timm, PhD Candidate

Ms. Timm is a PhD candidate at Florida International University. This means that she is a graduate student pursuing a Doctor of Philosophy. Since she was a child, Ms. Timm has been interested and curious about the world that surrounds her. She particularly remembers watching nature documentaries and taking trips to the beach inspiring her curiosity. As Ms. Timm progressed through her education and chose a career, she became particularly interested in DNA. This is because DNA can provide insights into an individual animal, a population, and even a whole species.

Now, as a PhD candidate, Ms. Timm studies the genetics of shrimp living in the Gulf of Mexico. She uses genetic analyses to diagnose species and community health and resilience. This research is valuable because it allows us to assess the health and diversity of the Gulf of Mexico and consequently work to conserve the many economies and habitats it provides.

One of the biggest challenges Ms. Timm has faced in her scientific career is refining her own skills and learning how to use constructive criticism. Whether you are pursuing a career in the sciences or any other career, being about the take constructive criticism and apply it to your work and experiences is a valuable skill.

Whether or not you are interested in pursuing a scientific career, Ms. Timm encourages you to become a scientifically literate member of society. If you're interested in something, then learn as much as you can about it. Also, look for opportunities to interact with or contribute to science or conservation efforts.

Ms. Timm has had the opportunity to participate in DEEPEND cruises in the Gulf of Mexico. Her favorite experience so far was seeing a waterspout! She said it was memorable because she had never seen one before; it was huge, lasted for a long time, and reminded her that weather doesn't only happen on land!



References and resources for teachers and students

Booth, W.C., Colomb, G. G., and Williams, J. M. 2008. The Craft of Research. The University of Chicago Press.

DEEPEND Consortium. <http://www.deependconsortium.org/>

Evolution and the Nature of Science Institutes. <http://www.indiana.edu/~ensiweb/home.html>

Flick, L.B., and Lederman, N.G. (Eds). 2006. Scientific Inquiry and Nature of Science. Implications for Teaching, Learning, and Teacher Education. Springer.

Handlesman, J., Miller, S., and Pfund, C. 2007. Scientific Teaching. W. H. Freeman and Company.

Saul, W., Reardon, J., Pearce, C., Dieckman, D., and Neutze, D. 2002. Science Workshop. Heinemann.

© 2017 K. Denise Kendall and DEEPEND Consortium

Photo Credits: Dante Fenolio, San Antonio Zoo and K. Denise Kendall, The University of Alabama in Huntsville.

Squirt Stills: Stanton Broadway, San Antonio Zoo.

Clip Art Credits: Stanton Broadway, San Antonio Zoo and K. Denise Kendall, University of Alabama in Huntsville.

About the Author:

Dr. K. Denise Kendall is a graduate of the Department of Ecology and Evolutionary Biology at the University of Tennessee-Knoxville (2013). She holds a Florida Professional Teaching Certificate (Biology 6–12) from the University of West Florida. Dr. Kendall is committed to the advancement of science education in K-12 and higher education through the integration of authentic scientific experiences into course curricula.