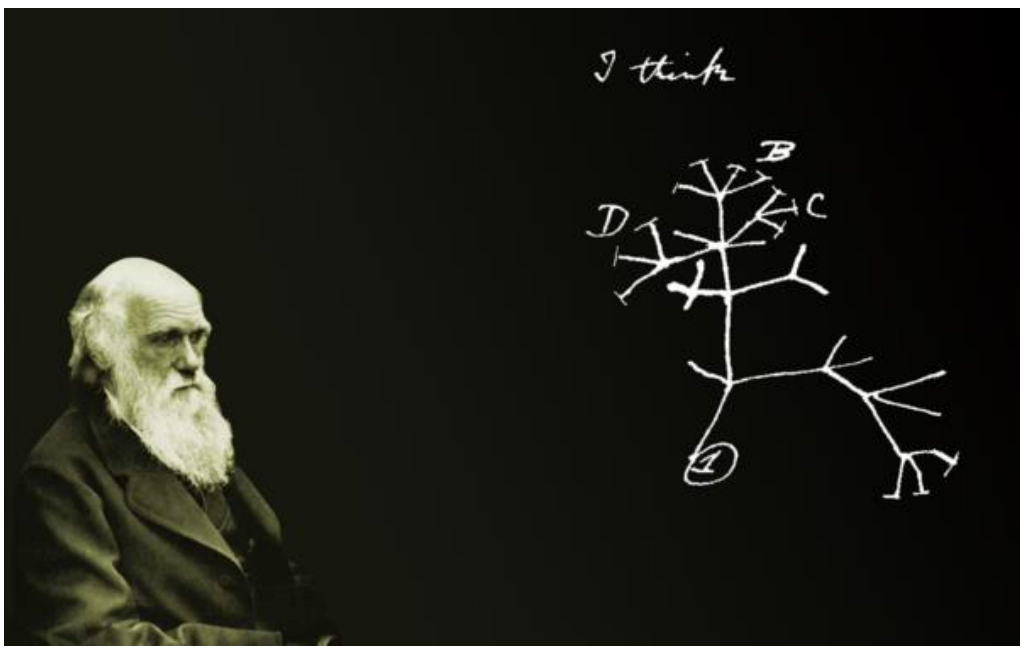
Keys to Success

AP Biology



The reason a lot of people do not recognize opportunity is because it usually goes around wearing overalls looking like hard work.

Thomas Edison

Our class goal is to understand how the human body is organized and regulated, and how behavior and physiology influence each other.

What you need to do to achieve our aim:

Trust your teacher; reflect on what motivates and demotivates you; celebrate the peaks and reflect on the valleys in your progress; and make your best effort to do better today than you did yesterday.

What your teacher will do:

Listen carefully to what demotivates you and stop doing it. Listen to your hypotheses regarding improving your enthusiasm and learning and test them.

Contents

Calendar

Calendar

How to Successfully Study

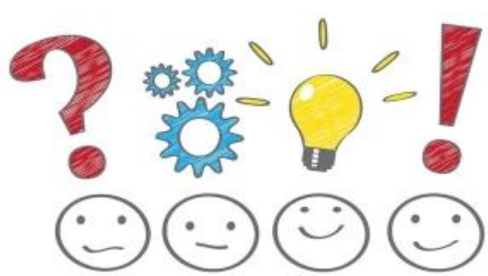
It does not matter whether you are writing an English literature paper, reviewing algebra problems, or finishing up a chemistry lab report. There are a few key elements every successful student needs to include in a study plan.

- ***Time-Management*** – It is not the amount of time you spend studying that matters. It's what you can accomplish during that time. Spending 40 hours to prepare for an exam and only earning a C clearly was a waste of your time. Develop a study plan and learn how to manage your time effectively to maximize your results.

- ***Motivation*** – If you are not motivated and have a poor attitude, your study session will not be very productive. You have just one opportunity to pass that Geometry exam or ace the term paper. Pick a time of day where you can get motivated to prepare for tests, write essays, and solve problems.

- ***Concentration*** – The ability to concentrate is one of the more important study skills you need to develop. You won't always be able to study in absolute silence or be able to spend as much time as you would like on a particular project. Learn how to overcome distractions so you can focus all your attention on your studies.

- ***When in doubt, ask*** – If you aren't sure about a particular topic, don't be shy. Ask your instructor, family, or friends for help. It is important to address the problem area as soon as possible. Otherwise, you will end up having to spend even more time studying to catch up.



The 5 Academic Keys to Learning

1) Repetition

Repetition must come from memory. Reading a bit of information over and over is NOT an effective way to learn.

2) Elaboration

Our brains like complexity and they especially like a good story. Information incorporated into a story is much easier to recall. Mnemonics are a simple why to elaborate.

3) Spacing (Multiple exposures)

Your brain is designed to forget. You need to reexpose yourself to learned information before you forget it. Each time you successfully recall a bit of information, you can extend the interval between exposures. For example, if you test yourself one day after learning a bit of information and still remember it, you can be certain that you will remember the information for 2 or more days.

4) Making connections between key ideas and the supporting details

Every bit of information should be learned within the context of big ideas. Rote memorization of random facts is an ineffective way to learn.

5) Making it interesting

Your brain doesn't pay attention to or remember boring things. You need to find a way to trick yourself into thinking boring information is interesting. Humor is a good tool for this.

The 3 Physiological Keys to Learning

1) No bad stress

Characteristics of good stress: 1) temporary; 2) empowering; 3) results in something predictably good.

Characteristics of bad stress: 1) long term; 2) intense; 3) caused by things we cannot control

Good stress is good for learning and your brain. Bad stress inhibits learning and damages the bits of your brain necessary for learning.

2) Adequate sleep

Sleep consolidates memories and is essential for all kinds of learning. Without adequate sleep you will be significantly WORSE at the following: 1) paying attention; 2) remembering; 3) thinking logically; 4) learning math.

3) Aerobic exercise everyday

Aerobic exercise increases your heartrate to about 120 to 140 beats per minute for at least 20 continuous minutes.

The following are some of the benefits of aerobic exercise: 1) protects against the negative effects of bad stress; 2) improves quality of sleep; 3) reduces depression; 4) improves reasoning; 5) improves long term memory; 6) helps you solve problems.

Name____Period_____

MANDATORY Study Hours Log First Quarter

30 pts weekly

Weekly Study Hours:

1. Every week you will complete 3 MANDATORY Study times: 1/2 hour will be completed in Mr. Burke's classroom, 1 hour will be completed with a classmate from any class period, 1 hour will be completed at home.
2. All 3 MANDATORY times require a supervisor's initials: Mr. Burke, your Study Buddy (may be different peers weekly), and a parent/guardian.
3. This MANDATORY Study Hours Log is due every Friday by 3:30 p.m. (or earlier)

	Week 1	Week 2	Week 3
Mr. Burke			
Study Buddy			
Home			
			/30
	Week 4	**Week 5**	**Week 6**
Mr. Burke			
Study Buddy			
Home			
	/30	/30	/30
	Week 7	**Week 8**	**Week 9**
Mr. Burke			
Study Buddy			
Home			
	/30	/30	/30

Quarterly Reflection:

20 pts

1. In two complete paragraphs:
    a. Describe your study routine and habits. Specifically explain how you used at least one of the academic keys to learning and at least one of the physiological keys to learning.
    b. Explain how you will improve your study routine and habits in the coming weeks.
2. Type the reflection double spaced in size 12 font with your name and period typed into the upper left corner.
3. Take time to proof read your reflection to ensure that it is insightful, purposeful, and clear.

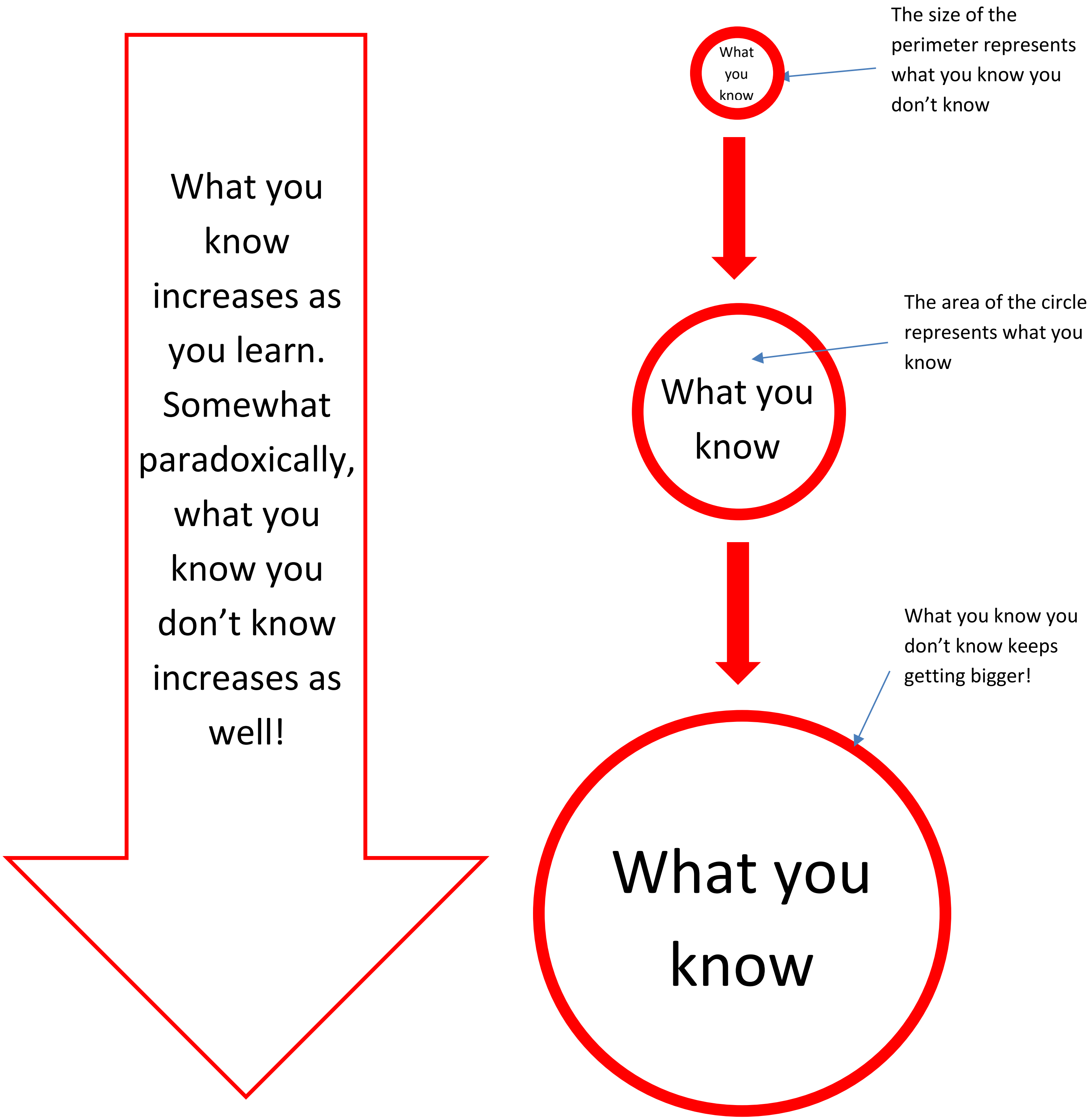
Question Log: Questioners Questioning Questions

Our imagination is stretched to the utmost, not as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.

Richard Feynman

Question Types:

1) **Basic science knowledge** (these are questions that can be answered by looking up facts in your notes or on Wikipedia *(They should never be on your list).*
   Examples: What part of the brain is most responsible for attention? What are the plantar flexor and dorsi flexor muscles that move the foot?
2) **Application**
   Examples: What would happen if the level of dopamine decreased in the striatum? How does the stride and cadence of a runner affect musculoskeletal health?
3) **Societal**
   Examples: Is technology use affecting the way people think? Why are epidemiologists concerned about decreases in vaccination rates?



If you don't have more questions after you engage in a learning activity than when you started, you haven't learned anything.

“What I don’t know” and “What science doesn’t know”

Number each question. Answer, and **put a check mark** next to each question with an answer. **Circle** the questions that science doesn’t know.

Questions and answers	Question Type

Number each question. Answer, and **put a check mark** next to each question with an answer. **Circle** the questions that science doesn't know.

Questions and answers	Question Type

What is WOOP?

WOOP is a method for building self-control

What should I use WOOP for?

Use WOOP for help with any kind of wish, whether it's something hard ("I want to start a school newspaper") or comparatively small ("I want to get an A in Science this quarter"). WOOP works equally well for academic, athletic, or personal wishes.

WOOP is especially helpful for anyone who procrastinates, who feels anxious about taking the first step—and the next step—toward their wish.

What makes WOOP effective?

Often, people fixate on how great it would feel to achieve their wish—but overlook the obstacles to that wish. Moreover, it is possible that imagining a wish gives you the illusion of achievement. The positive feelings from imagining success could reduce your motivation to do the hard work needed for real success. WOOP works because it prepares your brain for action and guides you through those in-between and oft-forgotten steps needed to achieve your dreams.

How do I become an expert WOOPer?

Like any skill, WOOP takes practice and patience to master. It's common to struggle at first. When practicing WOOP, be willing to experiment with different approaches. Make sure you are pursuing a goal that is meaningful to you. Be realistic, and think deeply without interruption when you WOOP.

WOOP Checklist:

<u>WISH</u>

☐ Specific and **important to YOU**

☐ Can be accomplished in time frame

☐ Challenging but **feasible**

<u>OUTCOME</u>

☐ Fulfilling and **motivating**

☐ **Clearly visualized** (Close your eyes and really imagine it)

☐ Reduced to most crucial aspects

<u>OBSTACLE</u>

☐ Inner obstacle, not outside barrier (it must be **something you have control over**)

☐ Clearly visualized

☐ Reduced to most crucial aspects

<u>PLAN</u>

☐ Stated as "when ... then..." (What is **the exact thing you will do**?)

☐ Observable action rather than internal decision (**what does it look like?**)

☐ You have all resources or skills needed to implement plan

☐ Plan should be easy to remember

Wish:

What is an important wish that you want to accomplish? Your wish should be challenging but feasible.

Outcome:

What will be the best result from accomplishing your wish? How will you feel? How will your wish improve your life? Close your eyes and really imagine it.

Obstacle:

What is the main obstacle inside of you that might prevent you from accomplishing your wish? Take time to really imagine an obstacle that you have control over.

Plan:

What's an effective action you could do to tackle the obstacle? Make a when-then plan.

When…**obstacle**…then…**I will**.

Learning Objective Self-Assessment

Background
The Essential Knowledge statements provided in the AP Biology Curriculum Framework are scientific claims describing phenomenon occurring in the physical world. These statements represent conclusions drawn from myriad scientific experimentations and observations. This is in contrast with conjecture, an opinion formed with insufficient supporting information. The Essential Knowledge statements are supported by multiple lines of evidence with examples found in widely published scientific research.

Scientific claims are fully understood when evaluated against supporting evidence found in scientific literature. Students of AP Biology will not only learn statements of Essential Knowledge but be able to evaluate and describe the supporting evidence.

Scientific claims can be incorporated into a broader concept, the scientific explanation. There is research to suggest that a scientific explanation will incorporate three features. These features include the claim, the supporting evidence, and the reasoning for how the evidence supports the claim (McNeill & Krajcik, 2008).

Assignment

1) Read each Learning Objective from your 'Objective Self-Assessment' document.

2) Enter one of the words below in the "Post-comprehension level" column, next to each of the Learning Objectives.

    Can – I can demonstrate I've learned the objective.
    Think– I think I can demonstrate I've learned the objective.
    Cannot – I cannot demonstrate I've learned the objective.

3) Complete the following for each Learning Objective you entered "Think" or "Cannot."
    a. State the Learning Objective.
    b. Use your notes and lecture materials to research the concept. Describe the evidence that supports this Learning Objective.
    c. Provide the reasoning behind your selection of the evidence.

Learning Objective
Evidence that supports this claim
How does the evidence support the claim?

***Do not include two Subobjective analyses on one piece of paper. Each will be turned in separately.**

Cornell Note taking Directions

When taking notes it is important to be studying while you take them rather than after you take them. This will improve the quality of notes you take and reduce your study time over the long run. In the short run, however, you need to be willing to spend more time taking notes than you normally do.

1) Before reading each section, write "initial guess" in the cue column and write the answer in the details column. If you are taking notes on a reading from a book or article, turn the section heading or title into a question.
    - this primes your brain for the information that follows
    - guessing wrong has been shown to improve comprehension and recall
2) Now read the section and take notes as you see fit.
    - cues such as questions, vocabulary words, and unlabeled diagrams go in the "cue column", which is approximately a third of the page on the left side
    - details such as answers to questions, definitions, and labeled diagrams go in the "details column" to the right of the cue column
        - details begin in line with each cue they explain
        - leave space between the end of details for one cue and details for another cue so you can revise your notes without making them messy
    - we are all visual learners so arrange your notes in a visually rich way
        - you must include all of the provided diagrams in the order they appear in the text
    - You must include at least one higher order question in the cue column that you cannot answer using Wikipedia
        - Highlight the question
        - Leave space in the details section to answer the question
3) In the summary section of your notes answer the title in 3 to 5 sentences.
    - cross out the initial guess you made (do not erase your initial guess)

Make-up assignments

1) **Vocabulary Index Card (cannot be used for bones or terms without definitions)**

Front side has a cartoon or diagram

1. **Write your name, period, and M#1 on the front of a 4 X 6 note card in the upper right hand corner**
2. The diagram or cartoon must have at least 4 different colors
3. There should be a minimum of writing and explanations on this side of the vocabulary card

Back side explains the word

4. Link-word
   - A word useful for remembering the word
   - The meaning of the link-word should be known
   - Good link-words are related to the vocabulary word or rhyme with it
5. Definition written 3 times from memory
   - Memorize the definition first
   - Each time you write the definition, cover up what you have written to ensure you are writing the definition from memory rather than copying the definition over and over
     a. The definition for a muscle is the origin, insertion, and action of the muscle
6. Write 2 sentences using the vocabulary word correctly

2) **Riddle Index Card**

Front side has the riddle and artwork

1. **Write your name, period, and M#2 on the front of a 3 X 5 note card in the upper right hand corner**
2. Riddle must contain one or more clues
   - The clues do not have to rhyme, but must refer to aspects of the answer
3. The riddle card must have artwork on the front
   - The artwork must reflect the topic
   - The artwork must have a clue that might help solve the riddle
   - The artwork must have at least 4 different colors

Back side

4. Has the answer to the riddle
   - The answer must be clearly and cleanly written across the top of the card
5. Below the answer an explanation of how the clues and artwork lead to the answer must be given
   - Explanation must be at least 2 sentences

3) Single Frame Cartoon

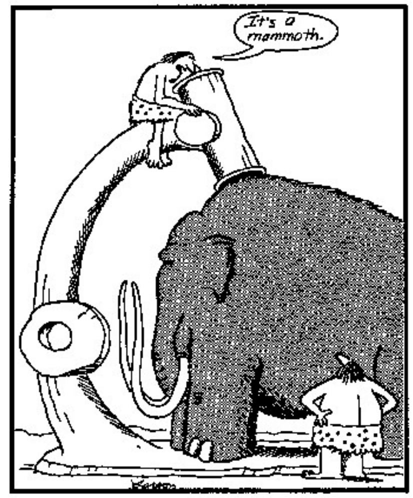
The cartoon does NOT have to be funny

Front side

1. **Write your name, period, and M#3 on the front of an 8.5 X 11" piece of paper in the upper right hand corner**
2. Draw a single frame cartoon like the Farside cartoon to the right
   - Use a minimum of 4 colors
3. Write a caption that is no more than 2 sentences (speaking bubbles are OK but not encouraged)

Back side

1. The concept the cartoon illustrates is stated
2. Write a paragraph that is at least 4 sentences explaining how the cartoon shows the concept stated



Early microscope

4) Explain One Term (cannot be used for bones or terms without definitions)

1. Cut an 8.5 X 11" piece of paper in half so you have an 8.5 X 5.5" piece of paper
2. Fold the paper in half so you have a card that is 4.25 X 5.5"

Front of card

3. **Write your name, period, and M#4 in the upper right hand corner on the front of the card**
4. Write the word and illustrate the word on the front of the card

Inside top half of card

5. Write the type of word it is (noun, verb, adverb, adjective) and how you know
6. Identify and define any morphemes

Inside bottom half of card

7. Define the word
8. Write a sentence using the vocabulary word correctly

Back of card (not shown)

9. Write the definition of the word 5 times
   - Memorize the definition first
   - Each time you write the definition, cover up what you have written to ensure you are writing the definition from memory rather than copying the definition over and over
   - The definition for a muscle is the origin, insertion, and action of the muscle

Carnivore

Front of Card

<u>Carni</u>**vore**

Flesh Feeding

noun because it is a type of thing

An organism that only eats other animals

The carnivore refused to eat salad.

5) Compare and Contrast Two Terms (counts for 2 vocabulary words and cannot be used for bones or terms without definitions)

1. Start with an 8.5 X 11” piece of paper in the landscape position
2. Fold the paper by bringing the two ends to the center of the paper
3. The result should be a 5.5 X 8.5” paper with two flaps that are like shudders covering a window

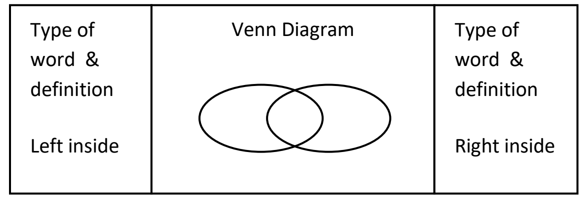
Word and diagram Left front	Word and diagram Right front

Left side

4. One term and an illustration on the front
5. The type of word (noun, verb, adjective, or adverb) and how you know on the inside left flap

Right side

6. **Write your name, period, and M#5 on the front in the upper right hand corner**
7. One term and an illustration on the front
8. The type of word (noun, verb, adjective, or adverb) and how you know on the inside left flap



Inside center

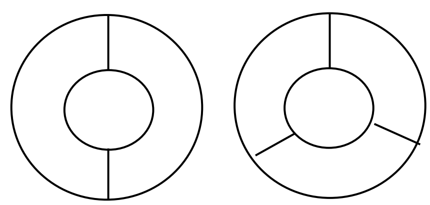
9. A Venn diagram to compare and contrast the two terms

Back of card

10. Write the definition of the words 5 times
    - Memorize the definition first
    - Each time you write the definition, cover up what you have written to ensure you are writing the definition from memory rather than copying the definition over and over
    - The definition for a muscle is the origin, insertion, and action of the muscle

6) Bull’s Eye comparison 2 or 3 words (counts for 2 to 3 words)

1. **Write your name, period, and M#6 in the upper right hand corner of an 8.5 X 11” piece of paper**
2. Compare and contrast two or three topics
    - differences between the topics belong in the outside sections
    - Similarities between the topics belong in the center
    - Words, pictures, graphs, equations, etc. should be used
3. The diagram must include a minimum of 4 colors
    - the colors must be used to emphasize similarities and differences
4. Write a 4-6 sentence paragraph that explains how the diagram shows similarities and differences and how the colors help clarify the similarities and differences



2 words 3 words

7) Four Window Concept Riddle (counts for 4 vocabulary words)

1. Start with an 8.5 X 11” piece of paper in the landscape position
2. Fold the paper by bringing the two ends to the center of the paper
3. The result should be a 5.5 X 8.5” paper with two flaps that are like shudders covering a window
4. Cut the two flaps (shudders) in half so you have 4 flaps (windows)
5. **Write your name, period, and M#7 in the upper right hand corner of the upper right front flap**

Riddle 1 Illustration	Riddle 2 Illustration
Riddle 3 Illustration	Riddle 4 Illustration

Front Four Flaps

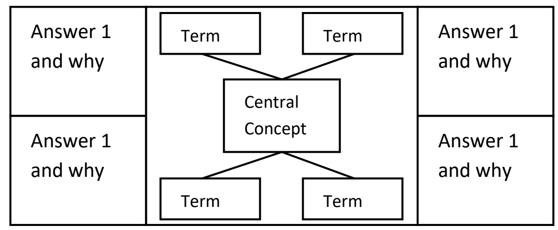
6. Select 4 terms (vocabulary words, processes, phrases, etc. that are related to a single fact)
7. Place a riddle on the front of each flap
    - Illustrate the riddle

Inside flaps (windows)

8. Write the answer to each riddle
    - Explain the answer to the riddle

Inside center

9. Make a concept map with the terms as spokes connecting to the central concept
    - Write the reason the term is connected to the central concept on the line connecting the term to the concept



8) Fables (counts for up to 3 vocabulary words)

A fable is a short story that teaches a moral or lesson.

1. **Write your name, period, and M#8 on the front of an 8.5 X 11” piece of paper in the upper right hand corner**

The fable

2. 3 paragraphs
3. The vocabulary words used in the fable must be highlighted

The moral or lesson

4. Write the moral or lesson after the last paragraph
5. Illustrate the moral or lesson

Explanation

6. Below the illustration, explain how the fable and illustration teach the vocabulary words
7. Your explanation must be at least 3 sentences per vocabulary word

9) Limerick

Limerick format

1. A limerick has exactly 5 lines
2. The last words of the first, second, and fifth lines rhyme with each other
3. The first, second, and fifth lines are longer than the third and fourth lines
4. The last words of the third and fourth lines rhyme with each other
5. The pattern of sounds follows the following: Da DUM da da DUM da da DUM

Limerick

1. **Write your name, period, and M#9 in the upper right hand corner of an 8.5 X 11" piece of paper**
2. The form and pattern of limerick writing must be followed
3. The limerick must be G or PG rated
4. The vocabulary words must be addressed in the limerick

Illustration

5. An illustration about the vocabulary words must follow the limerick
6. Use at least 4 colors

Explanation

7. A 3 to 5 sentence explanation of how the limerick and the illustration are related to the vocabulary words

10) Haiku

Haiku is a minimalist, contemplative poetry form from Japan that emphasizes nature, color, season, contrasts, and surprises. It usually has 3 lines and 17 syllables distributed in a 5, 7, and 5 syllable pattern. It should show a sensation, impression, or drama of a specific fact or vocabulary word.

Haiku

1. **Write your name, period, and M#10 in the upper right hand column of an 8.5 X 11" piece of paper**
2. The form and pattern of a haiku must be followed

Illustration

3. An illustration about the fact or vocabulary word must follow the haiku
4. Use at least 4 colors

Explanation

5. A 3 to 5 sentence explanation of how the haiku and the illustration are related to the vocabulary word

11) Cinquain

A cinquain is a five-line poem written about a single concept, object, or idea. An American poet developed cinquains after examining the Japanese haiku format. The format is a short, unrhymed poem of twenty-two syllables and five lines. The five lines contain 2, 4, 6, 8, then 2 syllables. Each line is supposed to deal with a specific aspect of the cinquian's topic.

Raindrop
Moisture, Falling
Sustain, Nourish, Cleansing
Teardrop, Diamond, Dropping, Earthward
Dewdrop

1. The first line consists of two syllables (a one word title)
2. The second line consists of four syllables (2 words describing the title)
3. The third line consists of six syllables (3 words stating action)
4. The fourth line consists of eight syllables (4 words expressing a feeling)
5. The last line consists of two syllables (1 word that is another word for the title)

Cinquain

1. **Write your name, period, and M#11 in the upper right hand column of an 8.5 X 11" piece of paper**
2. The form and pattern of a cinquain must be followed

Illustration

3. An illustration about the fact or vocabulary word must follow the cinquain
4. Use at least 4 colors

Explanation

5. A 3 to 5 sentence explanation of how the cinquain and the illustration are related to the vocabulary word

12) Tee Shirt Design

Design artwork for a tee shirt that represents a vocabulary word

1. **Write your name, period, and M#12 in the upper right hand corner of an 8.5 X 11" piece of paper**

Front of shirt

2. Artwork that represents the vocabulary word or fact
3. Must use at least 4 colors

Back of shirt

4. A 1 or 2 line cute or clever saying using the vocabulary word or fact
    - must be G or PG appropriate

The explanation

5. A 2 paragraph explanation of how the tee shirt represents the vocabulary word or fact
    - 1 paragraph for explaining how the artwork represents the vocabulary word or fact
- 1 paragraph for explaining how the saying represents the vocabulary word or fact

13) Acrostic Poem

An acrostic poem, sometimes called a name poem, uses a word for its subject. Then each line of the poem begins with a letter from the subject word. This type of poetry doesn't have to rhyme.

Acrostic Poem

1. **Write your name, period, and M#13 in the upper right hand corner of an 8.5 X 11" piece of paper**
2. The form and pattern of an acrostic poem must be followed

Illustration

3. An illustration about the fact or vocabulary word must follow the poem
4. Use at least 4 colors

Explanation

5. A 3 to 5 sentence explanation of how the poem and the illustration are related to the vocabulary word

14) Song, Rap, or Rhyming Poem

1. **Write your name, period, and M#14 in the upper right hand corner of an 8.5 X 11” piece of paper**

Illustration

2. An illustration about the fact or vocabulary word must follow the song, rap, or poem
3. Use at least 4 colors

Explanation

4. A 2 to 3 paragraph explanation of how the song, rap, or poem and the illustration are related to the vocabulary word

15) Vehicle Name

As part of a design team for a new model vehicle, you must select a name for the model. The name must reflect the vehicle’s abilities, which must be related to a vocabulary word

The Vehicle

1. **Write your name, period, and M#15 in the upper right hand corner of an 8.5 X 11” piece of paper**
2. Come up with a model name for the vehicle that is inspired by a vocabulary word or fact
3. Explain how the model name of the vehicle fits its abilities

Magazine Advertisement

4. Write an advertising slogan to be used to represent and show the vehicle’s features
5. Create an advertisement showing the vehicle emphasizing its features and name

Explanation

6. Explain in 2 to 3 paragraphs how the slogan and the advertisement represent the vocabulary word

16) Letter to the Editor

1. **Write your name, period, and M#16 in the upper right hand corner of an 8.5 X 11” piece of paper**

The letter

2. The letter must be 2 to 3 paragraphs long
3. Each use of the vocabulary word must be highlighted
4. State your opinion in the first paragraph
5. Use at least 5 specific facts to support your opinion

Illustration

6. An illustration that captures your opinion stated in the first paragraph
7. You must use at least 4 colors

Explanation

8. Explain how the letter and illustration are related to the vocabulary word in 3 to 5 sentences

17) People in Your neighborhood Flip Book (minimum of 4 vocabulary words)

Select a word and imagine it as if it represents a person in a neighborhood. Describe the people (words) that live in the neighborhood. Each person (a minimum of 4) is described on one of the pages of the flip book. **Write your name, period, and M#17 on the front page of the book.**

Making the Flip Book

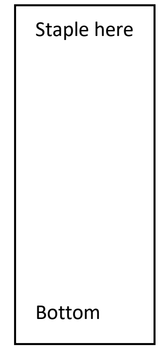
1. Your book must include at least 4 pages
2. Each page must be exactly 6cm by 12cm
3. Staple the pages together at the top

Front of each page

4. Draw and color a picture of the person in appropriate work or leisure clothing
5. Write the name of the person across the bottom of the flip page

Back of each page

6. Write the name of the person across the top of the flip page
7. Describe the job or workplace of the person
8. Explain how the job or workplace fits the person's name
9. Explain how the person's job helps the neighborhood function



18) Newspaper Article

1. **Write your name, period, and M#18 in the upper right hand corner of an 8.5 X 11" piece of paper**

The Article

2. Must be 2 to 3 paragraphs
3. Must contain the How, Who, What, When, Where, and Why about the vocabulary word
4. Must contain at least two interesting facts people could use in common conversations

The Illustration

5. Must represent the vocabulary word
6. Must have at least 4 colors
7. Must have a caption of 2 to 3 sentences explaining the graphic

19) Tattoo or Body Art

You are in charge of developing a tattoo to allow the world to know about one of the vocabulary words or facts

1. **Write your name, period, and M#19 in the upper right hand corner of an 8.5 X 11" piece of paper**

The Tattoo

2. The centerpiece of the tattoo must have a slogan or phrase as part of the tattoo
3. The surrounding artwork must demonstrate the vocabulary word in a real life situation
4. You must use at least 4 colors
5. The artwork must be suitable for all ages and appropriate for viewing in all social situations
6. The best location of the tattoo on the body must be written underneath the tattoo

The Explanation

7. 2 to 3 paragraphs
8. Explain how the artwork represents the vocabulary word
9. Explain why the tattoo belongs on a particular part of the body

20) Design a Clothing Line

You are a fashion designer and owner of a clothing company. Your next line of clothing will be named after a vocabulary word

1. **Write your name, period, and M#20 in the upper right hand corner of an 8.5 X 11" piece of paper.**

The logo

2. Design a logo that has the name of the new clothing line on it somewhere
3. The logo must use or apply the vocabulary word

Illustration

4. Illustrate one article of clothing from the new clothing line
5. The logo must appear somewhere on the article of clothing
6. You must use at least 4 colors

The Explanation

7. 3 paragraphs
    - 1 paragraph describing how the article of clothing represents the vocabulary word
    - 1 paragraph that explains how the name of the clothing line will help sell it
    - 1 paragraph that explains how the illustration represents the vocabulary word and how wearing the clothing would help a student learn the word

21) Design a Toy

Apply your knowledge of fun and science to design the hottest and best-selling toy of the season. Use a vocabulary word or fact for inspiration. The toy cannot cause serious bodily injury as part of its normal use.

1. **Write your name, period, and M#21 in the upper right hand corner of an 8.5 X 11" piece of paper**

The toy

2. Name the toy
3. What is the vocabulary word used when playing with the toy?
4. What are the most fun features of the toy?
5. What age group is the toy designed to reach?
6. How will playing with the toy help teach the vocabulary word

The illustration

7. Consider #'s 3-6 above when you illustrate the toy being used by a happy customer
8. Write a slogan for advertising the toy above the illustration
9. Use a minimum of 4 colors

The Explanation

10. 1 paragraph explaining how the slogan will help market the toy
11. 1 paragraph explaining how playing with the toy will help the user learn the vocabulary word

22) Public Service Announcement

You are charged with writing a public service announcement for the radio. The topic is one of the vocabulary words.

1. **Write your name, period, and M#22 in the upper right hand corner of an 8.5 X 11" piece of paper**

The announcement

2. Must be no less than 20 seconds and no more than 30 seconds when read aloud
3. The word must be mentioned at least twice during the announcement
4. A description of any sound effects or music that would accompany the commercial should be listed in parentheses and highlighted inside the body of the announcement
5. After the announcement, there must be a one sentence declaration of the organization responsible for developing the public service announcement

The Explanation

6. Explain how and why the public service announcement would help people understand the word

23) Pet Name

You are the proud owner of a new and unique pet. You decide to name the pet after one of your vocabulary words.

1. **Write your name, period, and M#23 in the upper right hand corner of an 8.5 X 11" piece of paper**
2. Name the pet after one of your vocabulary words

The illustration

3. Draw your pet doing a trick that represents its name
4. You must use at least 4 colors

The Explanation

5. 1 paragraph explaining how the behavior of the pet fits its name
6. 1 paragraph that describes the trick you illustrated and explains how the trick represents your pets name (the vocabulary word)

24) Design a Magazine Advertisement

Design a magazine advertisement about a vocabulary word

1. **Write your name, period, and M#24 in the upper right hand corner of an 8.5 X 11" piece of paper**

The Advertisement

2. Identify the magazine the ad will be placed in
    - The magazine must be appropriate of teenagers
3. The standard header and footer used by the magazine must be placed above and below the ad
4. The ad must be no more than half a page
    - Include claims, selling points, guarantees, etc. about the vocabulary word
    - Illustrations, diagrams, graphs, etc. must occupy at least half of the ad
5. Use at least 4 colors

The Explanation

6. A 2 to 3 paragraph explanation of how the ad explains the vocabulary word
    - explain why the magazine was selected to be advertised in
    - explain how the artwork helps explain the vocabulary word
    - explain how the claims, selling points, etc. help develop understanding of the vocabulary word

Conversation Starters (for group work and argumentation sessions)

1) What other things did you try?
2) I noticed _____.
3) For these reasons _____ I think _____.
4) I don't know if I agree with _____ because _____.
5) I disagree with _____ because _____.
6) Would _____ be better if _____?
7) One way to modify _____ is _____.
8) I have a question about _____.
9) How do you know _____?
10) Why did you decide to do _____?
11) _____ could be improved if _____.
12) Is _____ always that way?
13) What evidence do you have for _____?
14) Is there another possible solution?

The 4 Ground Rules of Critique in Science

(1) Be Respectful

Critique is how we identify errors or flaws in our ideas. In science, we always critique ideas, not people. When we are critiquing ideas, we never say hurtful things.

(2) Be Specific

Even if you are being respectful, you are not doing anybody any favors if you are vague. In science, we always make specific and detailed comments about what needs to be improved.

(3) Be Helpful

Critique is more than identifying a flaw or error in an idea; it is also about offering suggestions for ways to improve it. In science, we always offer specific and detailed suggestions for how to make things better.

(4) Use Scientific Criteria

Scientists use empirical and theoretical criteria to determine if an idea is valid or acceptable. In science, we always use scientific criteria to critique arguments and reports.

ADI Investigation Proposal TGB Version

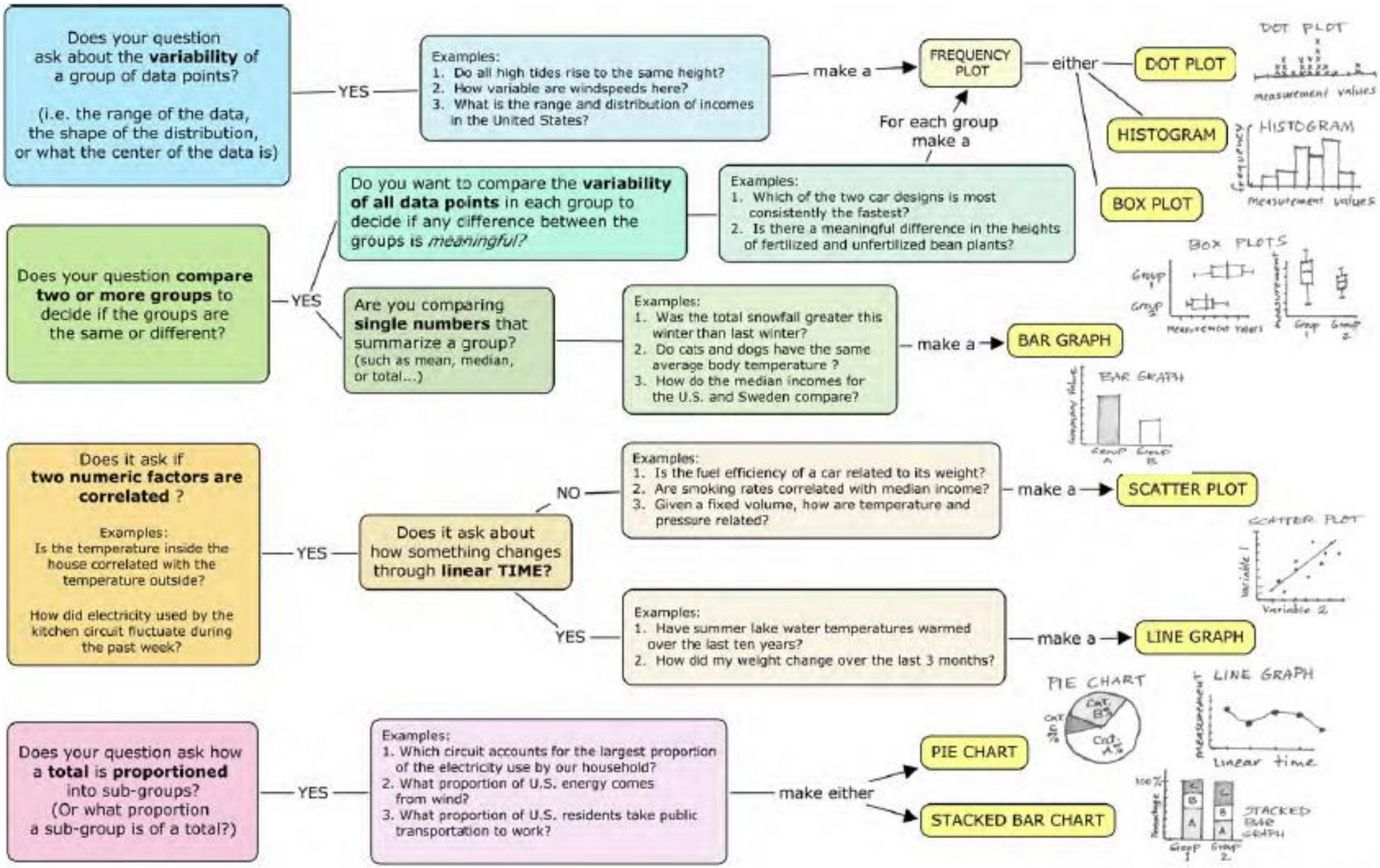
Guiding Question:

Claim:

Alternative claims:

Procedure:	What data will you collect? How will this data help you answer the guiding question? What safety precautions will you follow?

Graph Choice Chart



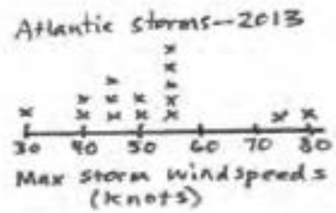
Variability questions: **Frequency plot** (3 kinds)

Kind of data:

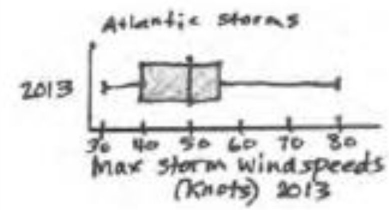
One categorical group and
One numeric variable (one axis)

Frequency plots show how variable the group is. Describe variability by range, measure of center (mean, median, or mode), and the shape of the distribution.

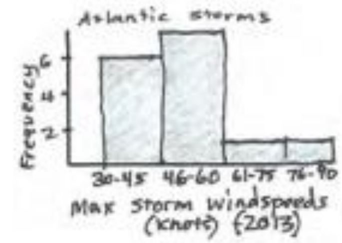
Dot plot



Box & whisker plot



Histogram

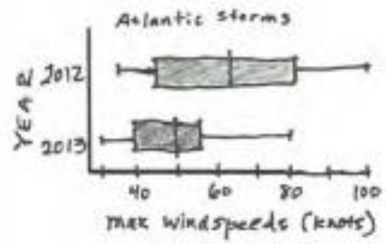


Comparing groups questions: **Frequency plots** OR **Bar graph**

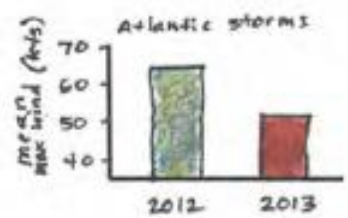
Kind of data:

Two or more categorical groups &
One numeric variable

Frequency plots *allow you to compare how variable the groups are.* ***Bar graphs*** *only show a single number (ie. sum, average, percent or count) for each group.*



(To compare two groups of values)



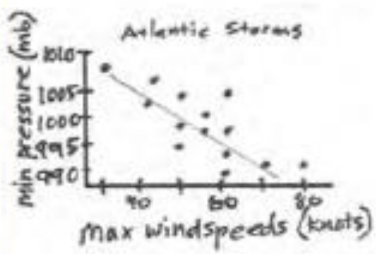
(To compare two summary values)

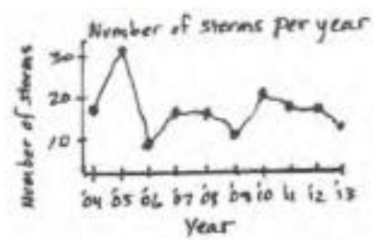
Correlation questions: **Scatter plot** OR **Line graph** (for time series)

Kind of data:

Two numeric variables

Both variables must be continuously numeric. Connect dots only if one variable is linear time (i.e. days, years...) Put time on the X-axis.
Show correlation with a 'line of best fit'.



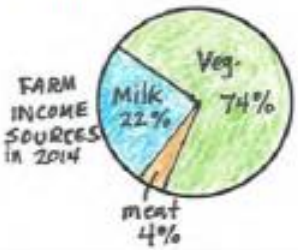


Proportion (percentage) questions: **Pie chart** OR **Stacked bar graph**

Kind of data:

Size of a subgroup as a **percentage** of the whole group (Total of sub-groups must = 100%)

In pie charts and stacked bar graphs, all sub-group percentages must total 100%.



Criteria for an informative graph:

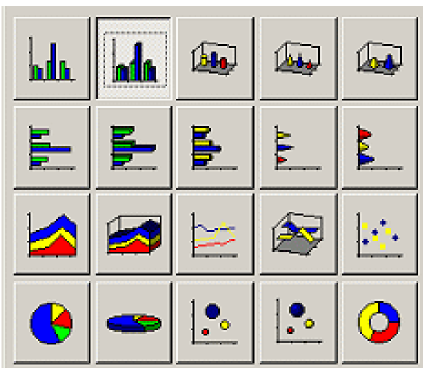
___ Graph type fits the question
___ Axes are drawn & scaled correctly
___ Axes are labeled clearly, correctly
___ Units are given
___ Data are plotted accurately
___ Legend is present, if needed
___ Graph is overall neat & legible
___ Title and/or caption present
___ Trend line shown (scatter plot or line graph only)
___ Graph helps answer the question

(There are other kinds of questions and other kinds of graphs, and often more than one graph type is useful for a given question. Learn to graph data for these basic kinds of questions first.)

Guiding Question:

Our Claim:

Our Evidence:



ANALYSIS

ILLUSTRATE A TREND, DIFFERENCE, OR A RELATIONSHIP AND DESCRIBE IT

- This graph indicates…
- This graph shows…
- This graph suggests…

INTERPRETATION

EXPLAIN WHAT THE ANALYSIS MEANS

Our Justification of the Evidence:

USE YOUR SCIENTIFIC KNOWLEDGE AND ANALYSIS TO SUPPORT YOUR INTERPRETATION

The 3 elements of a valid justification:

1) Theory or known scientific knowledge
2) Specific data from experiment related to the theory or scientific knowledge
3) Explanation of HOW the data relate to the theory or scientific knowledge

REASON

EXPLAIN WHY THE EVIDENCE MATTERS AND HOW IT RELATES TO ESTABLISHED SCIENTIFIC KNOWLEDGE

The 3 Elements of a Valid Justification

WHAT IT IS

WHAT YOU ARE DOING

Background knowledge usually given in pre-lab reading

Theory or known scientific knowledge

Paraphrasing established fact and/or theory

Data from your analysis of the experiment

Specific data from analysis of the experiment related to the theory or scientific knowledge

Connecting data to theory or knowledge

Your understanding of how the data APPLY to the theory or knowledge

Explanation of HOW the data relate to the theory or scientific knowledge

Explaining the relationship between the data and theory or knowledge

Identifying weaknesses in experimental design and/or execution **IF** the data are not in line with theory or knowledge

Argumentation Sample Questions

1) Why did you do _____? (reference a specific step in their procedures)
2) Why didn't you _____ (suggest a specific method, technique, etc. that could have been part of the procedure)
3) Why does ______ (reference specific evidence used) matter?
4) Why is ______ (reference specific evidence used) important?
5) What variables did you control for?
6) Why did you control for _____? (reference a specific variable controlled for)
7) What variables were NOT controlled for?
8) How could you have controlled for _____? (reference a specific variable that was not controlled for)
9) What are the strengths of your experimental design?
10) What could have made your experimental design better?
11) Why did you choose to collect _____? (reference specific data collected)
12) Why didn't you collect data on _____?
13) What is the relationship between the independent variable and the dependent variable?
14) How does your data support your claim?
15) How confident are you in your claim and why?
16) Why did you choose to illustrate your data with a _____? (reference the type of chart they used (scatter plot with best fit line, histogram, bar chart, etc.)
17) Why did you organize your data table the way you did?
18) What could you have done to make your data easier to understand?
19) What data would have strengthened your claim?
20) What statistics did you use to support your claim?
21) What statistics could you have used to strengthen your claim?

Argumentation Session Rubric

10 points	8 points	6 points	4 points	0 points
Asked several challenging and original questions at each station	Asked several challenging and original questions	Asked a few questions from the sample list at each station	Asked a few questions from the sample question list	Did not ask questions

Argumentation session feedback

We want to be helpful and supportive so everyone can get the most out of each scientific experience. Leave one positive comment and one respectful suggestion for improvement for each group you visit during an argumentation session.

We like this because…

We suggest that you change this to…

We suggest you add…

Will you clarify…?

Guide for writing an investigative report

Section 1: The introduction

What are you trying to do and why?

Be sure to:

- Provide a context for the investigation
- Explain the task
- Make the guiding question clear

You can find this information in the introduction to the lab handout

Section 2: The method

What did you do and why did you do it that way?

Be sure to explain…

- How you collected your data and why you decided to do it that way
- What type of data you collected and why you collected that data
- How you analyzed your data and why you decided to analyze it that way

You can find this information in the introduction to the lab handout

Section 3: The argument

What is your argument?

Be sure to…

- Provide your claim
- Support your claim with evidence
- Use a figure to present your evidence and reference it
- Provide a justification for your evidence

You can find all this information on your white board

*must include a data table and graph

**no more than 2 pages typed, double spaced, 12 point font, 1 inch margins (all around)

ADI Investigation Report Peer Review Guide-TGB version

Report by________________________

Reviewed by______________________ Reviewed by______________________ Reviewed by______________________

YOU MUST ADDRESS THE FOLLOWING WORDS AND CONCEPTS IN YOUR INTRODUCTION:	**Section 1: Introduction and Guiding Question**	**Reviewer Rating**		
	Is there enough background information and is it accurate?	☐ No	☐ Partially	☐ Yes
	Is the guiding question explicit and did the author **EXPLAIN HOW** the question is related to the background information?	☐ No	☐ Partially	☐ Yes
	Reviewers: **EXPLAIN HOW** the author could improve this part of his or her report.	Author: What revisions did you make in your report?		

Section 2: Method	**Reviewer Rating**		
Did the author describe the procedures and **EXPLAIN WHY** the procedures were used?	☐ No	☐ Partially	☐ Yes
Did the author describe what data were collected and **EXPLAIN WHY** the data were collected?	☐ No	☐ Partially	☐ Yes
Reviewers: **EXPLAIN HOW** the author could improve this part of his or her report.	Author: What revisions did you make in your report?		

<table>
<tr><th colspan="2">Section 3: The Argument</th><th colspan="3">Reviewer Rating</th></tr>
<tr><td colspan="2">Did the author make a claim consistent with the evidence that answers the guiding question?</td><td>☐No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td rowspan="4">Quality of Evidence</td><td>Did the author describe how he/she analyzed the data and EXPLAIN WHY the analysis helped him/her answer the guiding question?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Is the analysis of the data appropriate and free from errors?</td><td>☐No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Is the author’s interpretation of the analysis valid?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>MANDATORY DATA</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td rowspan="2">Presentation of Evidence</td><td>Are tables and graphs correctly formatted, labeled, and are metric units used?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Are specific values from tables and graphs referenced in the body of the text?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td rowspan="3">Justification of Evidence</td><td>Did the author state WHAT scientific concept is being defended?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Did the author DESCRIBE the data from their analysis used as evidence in support of the scientific concept?</td><td>☐No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Did the author EXPLAIN HOW the evidence supports the underlying scientific concept?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td colspan="2">Did the author EXPLAIN HOW his or her claim agrees with the claims made by other groups and PROVIDE REASONS for disagreements?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td colspan="2">Reviewers: EXPLAIN HOW the author could improve this part of his or her report.</td><td colspan="3">Author: What revisions did you make in your report?</td></tr>
</table>

<table>
<tr><th>Mechanics</th><th colspan="3">Reviewer Rating</th></tr>
<tr><td>Organization: Is each section easy to follow? Do paragraphs include multiple sentences: Do paragraphs begin with a topic sentence?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Grammar: Are the sentences complete? Is there proper subject-verb agreement in each sentence? Are there run-on sentences?</td><td>☐No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Conventions: Did the author use appropriate spelling, punctuation, paragraphing and capitalization:</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Word Choice: Did the author use the appropriate words (e.g., there vs. their, to vs. too, than vs. then, etc.)?</td><td>☐ No</td><td>☐ Partially</td><td>☐ Yes</td></tr>
<tr><td>Fluency: Does the paper as a whole flow well? Do the sections follow each other to give a complete idea of the lab?</td><td>☐ No</td><td>☐ Partially</td><td>☒ Yes</td></tr>
</table>

Peer Review Feedback

If you think the author…	Weak Feedback	Strong Feedback
…wrote something that was inaccurate.	That is not right.	We disagree with _____. We think you should change it to _____.
…needs to make a change to a table, graph or figure.	Fix this.	We think you need to reorganize your _____. Here is how we would change it: _____.
… forgot to include something important	Write more.	We suggest adding _____ and _____.
… included an important piece of information, but did not provide enough details about it.	Add more detail.	We think you need to be more specific about _____. We suggest making the following changes: _____

Annotating Text

Annotating Text
☐ **UNDERLINE** concepts you think might be useful for understanding or solving the problem
☐ **Box** information you think might be helpful for designing your investigation
☐ ← Write **notes** in the left margin
☐ → Write **questions and answers** in the right margin
Each paragraph must have something underlined or boxed, **AND** have something written in the margins (a question and/or note).

AP BIOLOGY EQUATIONS AND FORMULAS

STATISTICAL ANALYSIS AND PROBABILITY

Standard Error

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

***NOTE:** For the purposes of the AP Exam, students will not be asked to manipulate or derive this equation; however, they must know the underlying concepts and applications.*

Mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Standard Deviation

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

***NOTE:** For the purposes of the AP Exam, students will not be asked to manipulate or derive this equation; however, they must know the underlying concepts and applications.*

Chi-Square

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

s = sample standard deviation (i.e., the sample based estimate of the standard deviation of the population)

$\bar{x}$ = mean

n = size of the sample

o = observed individuals with observed genotype

e = expected individuals with observed genotype

Degrees of freedom equals the number of distinct possible outcomes minus one.

CHI-SQUARE TABLE

	Degrees of Freedom							
p	1	2	3	4	5	6	7	8
0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51
0.01	6.64	9.21	11.34	13.28	15.09	16.81	18.48	20.09

LAWS OF PROBABILITY

If A and B are mutually exclusive, then P (A or B) = P(A) + P(B)

If A and B are independent, then P (A and B) = P(A) x P(B)

HARDY-WEINBERG EQUATIONS

$p^2 + 2pq + q^2 = 1$ — p = frequency of the dominant allele in a population

$p + q = 1$ — q = frequency of the recessive allele in a population

METRIC PREFIXES

Factor	Prefix	Symbol
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	µ
10^{-9}	nano	n
10^{-12}	pico	p

Mode = value that occurs most frequently in a data set

Median = middle value that separates the greater and lesser halves of a data set

Mean = sum of all data points divided by number of data points

Range = value obtained by subtracting the smallest observation (sample minimum) from the greatest (sample maximum)

RATE AND GROWTH

Rate

dY/dt

Population Growth

$dN/dt = B - D$

Exponential Growth

$$\frac{dN}{dt} = r_{\max} N$$

Logistic Growth

$$\frac{dN}{dt} = r_{\max} N \left(\frac{K - N}{K} \right)$$

dY = amount of change

t = time

B = birth rate

D = death rate

N = population size

K = carrying capacity

r_{max} = maximum per capita growth rate of population

Temperature Coefficient Q_{10}

NOTE: *For use with labs only (optional)*

$$Q_{10} = \left(\frac{k_2}{k_1} \right)^{\frac{10}{t_2 - t_1}}$$

Primary Productivity Calculation

mg O_2/L x 0.698 = mL O_2/L

mL O_2/L x 0.536 = mg carbon fixed/L

t_2 = higher temperature

t_1 = lower temperature

k_2 = metabolic rate at t_2

k_1 = metabolic rate at t_1

Q_{10} = the *factor* by which the reaction rate increases when the temperature is raised by ten degrees

SURFACE AREA AND VOLUME

Volume of a Sphere

$V = 4/3 \pi r^3$

Volume of a Cube (or Square Column)

$V = l w h$

Volume of a Column

$V = \pi r^2 h$

Surface Area of a Sphere

$A = 4 \pi r^2$

Surface Area of a Cube

$A = 6 a$

Surface Area of a Rectangular Solid

$A = \Sigma$ (surface area of each side)

r = radius

l = length

h = height

w = width

A = surface area

V = volume

Σ = Sum of all

a = surface area of one side of the cube

Water Potential (Ψ)

$\Psi = \Psi p + \Psi s$

Ψp = pressure potential

Ψs = solute potential

The water potential will be equal to the solute potential of a solution in an open container, since the pressure potential of the solution in an open container is zero.

The Solute Potential of the Solution

$\Psi s = -iCRT$

i = ionization constant (For sucrose this is 1.0 because sucrose does not ionize in water).

C = molar concentration

R = pressure constant (R = 0.0831 liter bars/mole K)

T = temperature in Kelvin (273 + °C)

Dilution – used to create a dilute solution from a concentrated stock solution

$C_i V_i = C_f V_f$

i = initial (starting)

C = concentration of solute

f = final (desired)

V = volume of solution

Gibbs Free Energy

$\Delta G = \Delta H - T\Delta S$

ΔG = change in Gibbs free energy

ΔS = change in entropy

ΔH = change in enthalpy

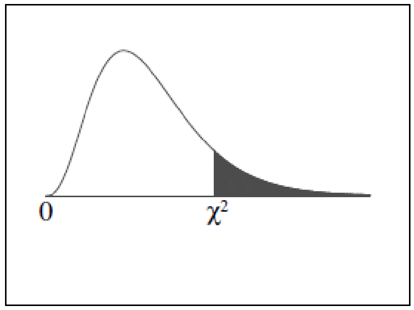
T = absolute temperature (in Kelvin)

pH = – log [H+]

NOTE: *For the purposes of the AP Exam, students will not be asked to manipulate or derive this equation; however, they must know the underlying concepts and applications.*

Chi-Square Distribution Table

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

Additional help.

https://www.khanacademy.org/math/probability/statistics-inferential

http://onlinestatbook.com/2/index.html

What statistics test should I use?
http://www.socscistatistics.com/tests/what_stats_test_wizard.aspx

How do I know what kind of graph to make?

1) Does the graph help you visualize a pattern in the data?
2) Does the graph provide insight into the data?
3) Does the graph help you visualize the range and distribution of the data?
4) Does the graph help you answer the question you are investigating?
5) Does the graph facilitate making predictions about the data?

Representing quantitative data.

Scientists collect large quantities of data. They use tables, graphs, and charts to organize, interpret, and illustrate the data they collect. Different data types are best illustrated with a particular type of chart. For example, histograms are better for illustrating quantitative data and bar charts are best for illustrating qualitative data. Additionally, some charts are better for illustrating different aspects of and answering different questions about the same data set.

Methane is an important greenhouse gas. Many countries around the world are interested in decreasing greenhouse gas emissions. Table 1 shows the sources of methane emissions in the U.S. by category over several years. Methane emissions are measured in teragrams of Carbon dioxide (CO_2) equivalents. In other words, 1 $TgCO_2$ of methane is the amount of heat a thousand tons of carbon dioxide absorbs. Data that is continuously variable, like methane emissions, is called quantitative data.

Table 1 U.S. Methane Emissions by Source ($TgCO_2$ Equivalents)

Source Category	1990	2000	2005	2006	2007	2008	2009
Natural Gas Systems	189.8	209.3	190.4	217.7	205.2	211.8	221.2
Enteric Fermentation	132.1	136.5	136.5	138.8	141	140.6	139.8
Landfills	147.4	111.7	112.5	111.7	111.3	115.9	117.5
Coal Mining	84.1	60.4	56.9	58.2	57.9	67.1	71
Manure Management	31.7	42.4	46.6	46.7	50.7	49.4	49.5
Petroleum Systems	35.4	31.5	29.4	29.4	30	30.2	30.9
Wastewater Treatment	23.5	25.2	24.3	24.5	24.4	24.5	24.5
Forest Land Remaining Forest Land	3.2	14.3	9.8	21.6	20	11.9	7.8
Rice Cultivation	7.1	7.5	6.8	5.9	6.2	7.2	7.3
Stationary Combustion	7.4	6.6	6.6	6.2	6.5	6.5	6.2
Abandoned Underground Coal Mines	6	7.4	5.5	5.5	5.6	5.9	5.5
Mobile Combustion	4.7	3.4	2.5	2.3	2.2	2	2
Composting	0.3	1.3	1.6	1.6	1.7	1.7	1.7
Petrochemical Production	0.9	1.2	1.1	1	1	0.9	0.8
Iron and Steel Production & Metallurgical Coke Production	1	0.9	0.7	0.7	0.7	0.6	0.4
Field Burning of Agricultural Residue	0.3	0.3	0.2	0.2	0.2	0.3	0.2
Ferroalloy Production	+	+	+	+	+	+	+
Silicon Carbide Production and Consumption	+	+	+	+	+	+	+
Incineration of Waste	+	+	+	+	+	+	+
International Bunker Fuels	0.2	0.1	0.1	0.2	0.2	0.2	0.1
Total for U.S.	674.9	659.9	631.4	672.1	664.6	676.7	686.3

Source: Source: Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2009

The amount of data in Table 1 above makes interpreting the significance of the data and finding patterns in the data difficult. A pie chart like the one in Figure 1 is an effective way to illustrate the major sources of methane emissions. However, a line graph like the one in Figure 2 would be a better choice for illustrating how emissions from the 4 largest sources have changed over time.

The graph below clearly shows that emissions from natural gas systems have been increasing while emissions from the next 3 largest sources have either decreased or increased slightly. Enteric fermentation results in methane emissions from the digestion of plant matter by organisms such as cows.

The pie chart and graph together give the public and policy makers a clear picture of where methane emissions are coming from and how emissions are changing over time. This information is important for making good public policy decisions.

Highly variable quantitative data is usually best illustrated using a scatter plot with a best fit line. For example, Figure 3 shows the relationship between brain mass and body mass in mammals. A best fit line generally results in half of the points above the line and half of the points below the line. Points above the line indicate an above average brain size and points below the line indicate a below average brain size.

Notice that an elephant has a larger brain than a human. However, the elephant's brain lies on the best fit line meaning its brain is the size we would expect for a mammal of its size. Although the human brain is significantly smaller than the elephant's brain, our brains are much larger than we would expect for a mammal of our size.

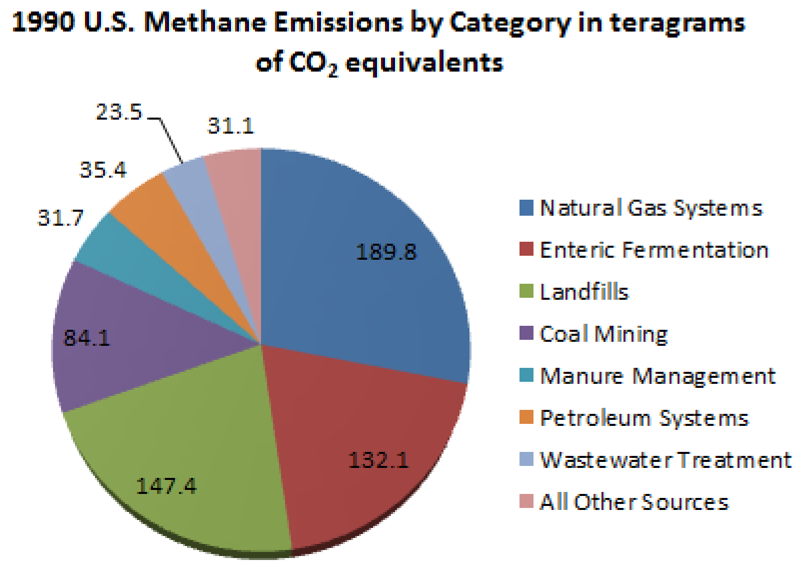


Figure 1

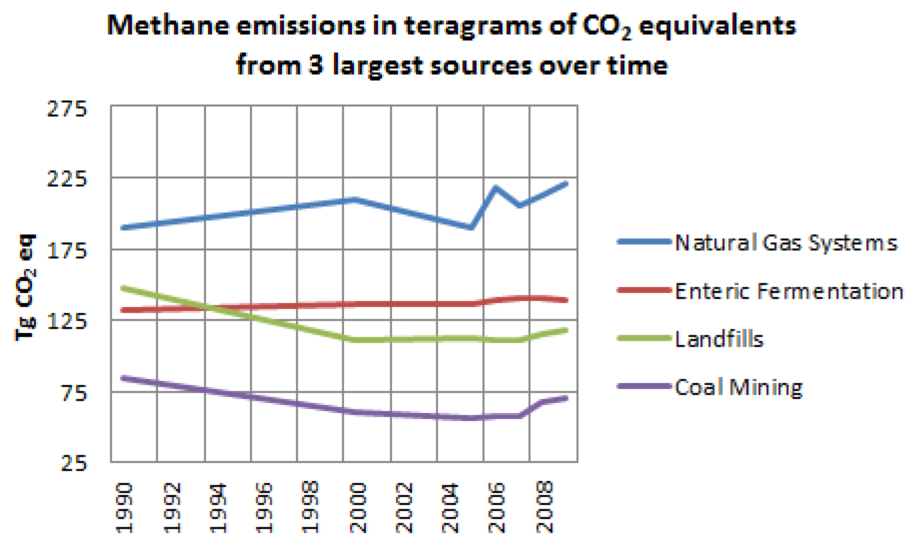


Figure 2

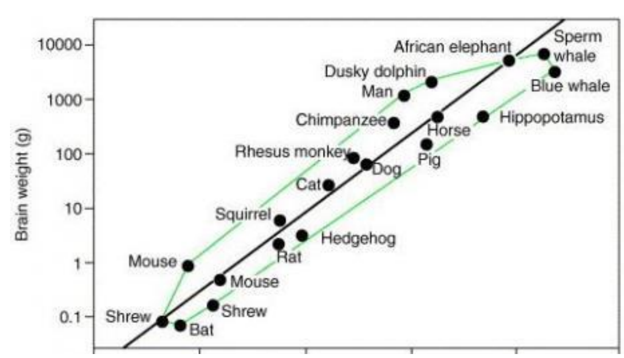
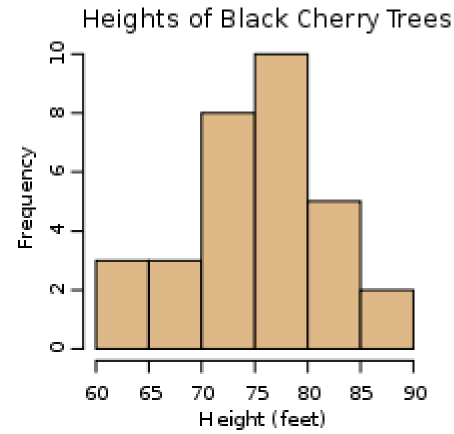


Figure 3 The dependence of brain mass on body mass in select mammals

Sometimes illustrating quantitative data with a histogram provides useful information. For example, the histogram to the right shows the frequency of cherry trees that fall within various ranges of heights. The 2 most frequently found ranges of tree heights are 70-75 feet and 75-80 feet. A histogram rather than a bar chart is the preferred method for representing continuously varying data such as height.



Representing qualitative data.

Data that is categorical is qualitative data. Examples of qualitative data are gender, color, nationality and taxonomic ranking. The table to the right shows the 35 phyla of kingdom Animalia and the approximate number of species described in each. Given that a species is either in phylum Chordata, or Arthropoda or some other phylum, a bar chart rather than a histogram is the appropriate way to represent species number per phylum. A pie chart would be another way of representing this information.

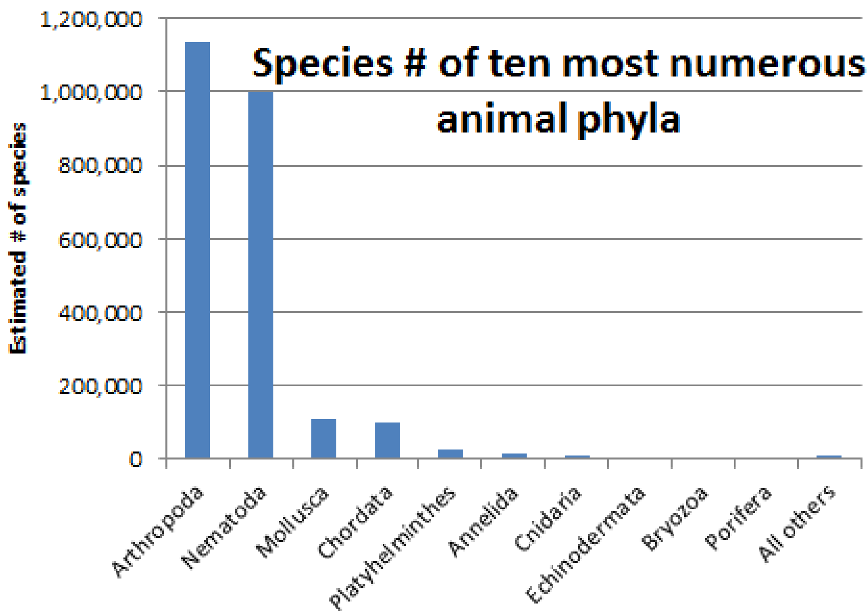


Figure 4

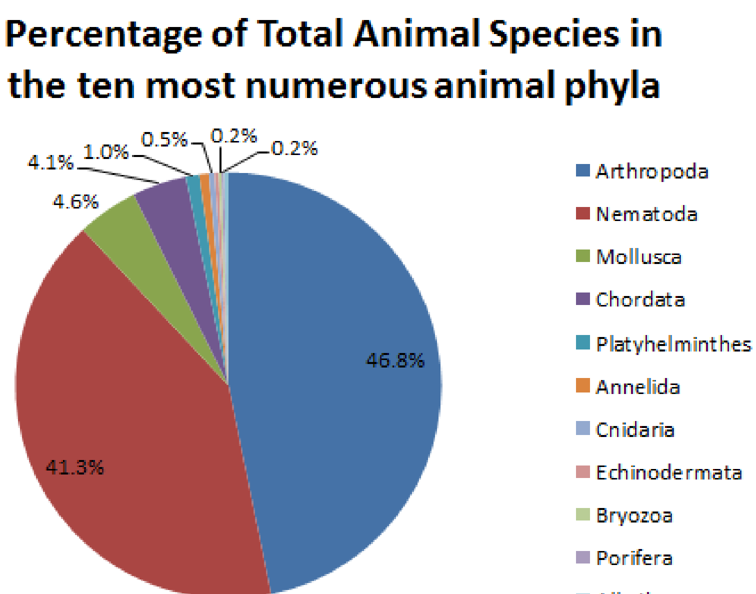


Figure 5

Animal Phyla	
Phylum	**Estimated # of species**
Acanthocephala	756
Acoelomorpha	
Annelida	17,000
Arthropoda	1,134,000
Brachiopoda	500
Bryozoa	5,000
Chaetognatha	100
Chordata	100,000
Cnidaria	11,000
Ctenophora	100
Cycliophora	3
Echinodermata	7,000
Entoprocta	150
Gastrotricha	690
Gnathostomulida	100
Hemichordata	100
Kinorhyncha	150
Loricifera	122
Micrognathozoa	1
Mollusca	112,000
Nematoda	1,000,000
Nematomorpha	320
Nemertea	1,200
Onychophora	200
Orthonectida	20
Phoronida	20
Placozoa	1
Platyhelminthes	25,000
Porifera	5,000
Priapulida	16
Rhombozoa	75
Rotifera	2,000
Sipuncula	320
Tardigrada	1,000
Xenoturbellida	2
Total: 35	2,423,946

Notice how the bar chart (Figure 4) and the pie chart (Figure 5) above facilitate interpreting the animal phyla data table. You and all other organisms with a back bone are in phylum Chordata. This phylum includes all sharks, fish, amphibians, reptiles, birds, and mammals (it also includes some other species without jaws, and some without back bones). Although biology classes typically focus on our phylum, just about all animals on the planet are arthropods or nematodes.

Figure 6 A platypus (member of phylum Chordata)

Arthropods include crustaceans such as crabs, spiders, and insects. Nematodes are mostly less than 0.1 inches. They are the most numerous animals on the planet and should not be confused with earth worms, which belong to phylum Annelida.

Figure 7 Giraffe Weevil (member of phylum Arthropoda)

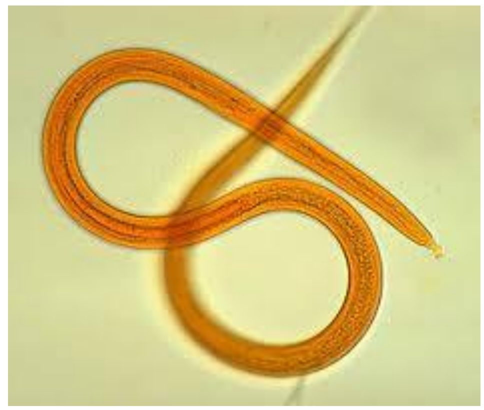


Figure 8 The common name for most members of phylum Nematoda is round worm

What are statistics?

Statistics are numbers generated through random sampling that allow us to make inferences about a population. For example, I could infer the mean height of everyone on the planet (symbolized by μ) by measuring the height of a random sample of people and determining the mean of the sample (symbolized by M or $\bar{X}$).

1) Take a random **sample**

2) Calculate mean ($\bar{X}$) of sample (this is a **statistic**)

3) Make **inference** about the mean (μ) of the population

Population (populations are often too large to measure)

Mean of the sample ($\bar{X}$) is a statistic that is an estimate of the mean of the population (μ).

What is a box and whiskers plot

Box and whiskers or box plots are a way of visualizing data and comparing distributions of data sets from different treatment groups.

1) **The first step is to rank the data from the smallest value to the largest value. The table below is time to react to a stimulus in hundredths of a second.**

Each sample has a rank number to make analysis easier

Samples are ranked from smallest to largest

1	14
2	15
3	16
4	16
5	17
6	17
7	17
$\boxed{8}$	$\boxed{17}$
9	17
10	18
11	18
12	18
13	18
14	18
15	18
$\boxed{16}$	$\boxed{19}$
17	19
18	19
19	20
20	20
21	20
22	20
23	20
$\boxed{24}$	$\boxed{20}$
25	21
26	21
27	22
28	23
29	24
30	24
31	29

There are 31 samples in this data set

29 hundredths of a second is the longest reaction time

2) **Calculate the bottom 25% or 25th percentile (the lower hinge). 25% of samples are below this point.**

R (n + 1) = the rank number or the percentile of interest

Fraction or percentile (25%=0.25, 50%=0.5, and 75%=0.75)

Number of samples

0.25 (31 + 1) = 8

17 hundredths of a second is the 8th ranked sample, so 25% of samples are at or below this point. This is the lower hinge.

3) **Calculate the median (50% of samples are above and below this point). The median is equal to the 50th percentile.**

0.5 (31 + 1) = 16

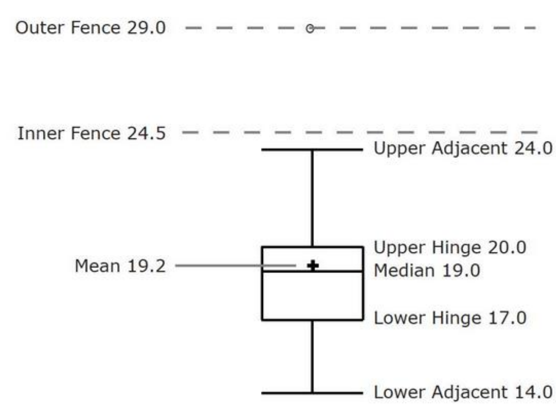
19 hundredths of a second is the 16th ranked sample, so this is the median. 50% of samples are above and 50% of samples are below this point.

4) **Calculate the bottom 75% (the upper hinge). 75% of samples are below this point.**

0.75 (31 + 1) = 24

20 hundredths of a second is the 24th ranked sample, so 75% of the samples are below this. This is the upper hinge.

5) Plot the "box". ***The middle 50% of your data set is contained within the box.***
6) There are different ways of representing the whiskers. ***The whiskers show the spread of your data***. The inner fences equal 1.5 times the difference between the 75th percentile and the 25th percentile. In this case that is 1.5 * (20 - 17) = 4.5. So add 4.5 to the upper hinge (75th percentile). This equals 24.5; and subtract 4.5 from the lower hinge (25th percentile). The lower inner fence is not shown in the box plot to the right.



7) The upper whisker is the sample data point closest to and below the upper inner fence. The upper inner fence is 17 + 4.5 = 24.5 and our closest sample data point is 24 hundredths of a second (see data table above).
8) The bottom whisker is the sample data point closest to and above the lower inner fence. The lower inner fence is 17 - 4.5 = 13.5 and our closest sample data point above that is 14 hundredths of a second (see data table above).
9) The outer fence is 3 times the difference between the 75th percentile and the 25th percentile.
10) Data points beyond the outer fence are not to be ignored, but they may indicate an irregularity in the sampling technique.

What is the mean?

The average of a sample is the mean. Below is the equation you will get on the reference sheet for the AP exam. Suppose I want to know the height of a population of rabbits. I catch, measure, and release 5 rabbits. Each rabbit in my sample is represented in the equation below by X_i.

X_1 = 21cm, X_2 = 23cm, X_3 = 27cm, X_4 = 14cm, and X_5 = 24cm

To the last sample

Sum

mean

Sample i

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Number of samples

From sample 1

The equation above says the following: the mean ($\bar{X}$) equals the sum of the samples (X_i) from the first sample (X_1) to the last sample (in this case X_5) divided by the number of samples (n).

$$\bar{X} = \frac{X_1 + X_2 + X_3 \ldots + X_n}{n}$$ is a simpler way of stating $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

For my rabbit height sample

$$\bar{X} = \frac{21cm + 23cm + 27cm + 14cm + 24cm}{5}$$

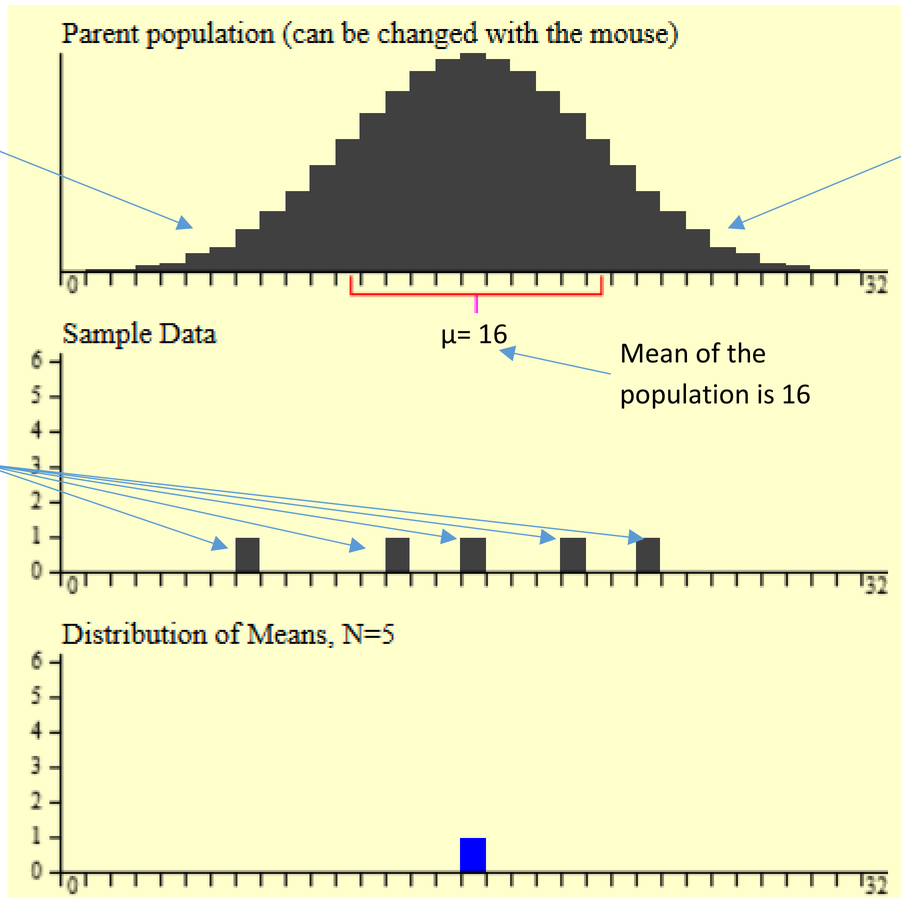
$$\bar{X} = \frac{109cm}{5}$$

$$\bar{X} = 21.8cm$$

We are almost never able to measure an entire population, so we use statistics to estimate parameters of a population. If you could measure the height of every rabbit in a population, it would look something like the graph labeled parent population below. Go to http://onlinestatbook.com/stat_sim/sampling_dist/index.html to experiment with this statistic.

A few really short rabbits

A few really tall rabbits



μ= 16

Mean of the population is 16

This is a random sample of the heights of 5 rabbits

$\bar{X}$= 15.8

This is my statistic. The **mean of my sample** is approximately equal to the mean of my population. In symbols this means $\bar{X} \approx \mu$

What is the standard deviation?

The standard deviation of a population (σ) is a measure of the spread of a parameter of a population. You can also think about it as a way of quantifying how far from normal a sample is. The standard deviation of a sample (S) is an estimate of standard deviation of the population (σ).

1) Take a random **sample**

2) Calculate standard deviation (S) of sample (this is a **statistic**)

3) Make **inference** about the standard deviation (σ) of the population

Population (populations are often too large to measure)

Below is the equation you will get on the reference sheet for the AP exam. It is assumed that you know $\sum_{i=1}^{n}$ is the same thing as $\sum$

The equation below reads as follows: the standard deviation of the sample is equal to the square root of the sum of each sample minus the sample mean squared divided by the number of samples minus 1.

$$S = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}}$$

$$S = \sqrt{\frac{\boxed{\sum (x_i - \overline{x})^2}}{n-1}}$$

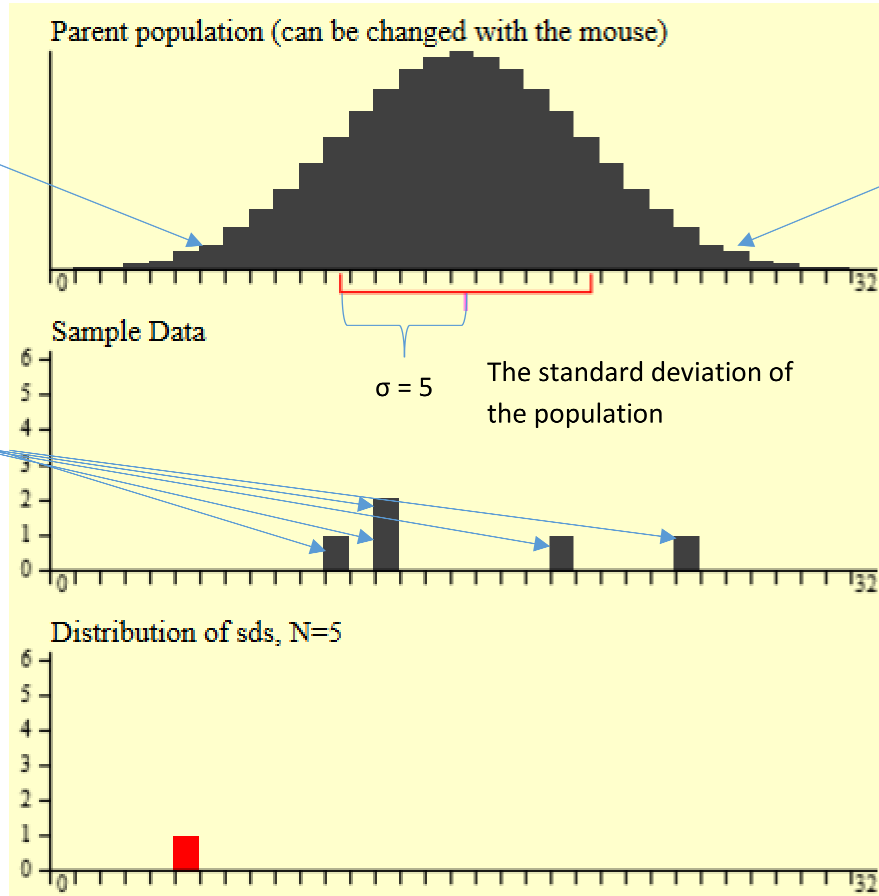
This is a measure of the variance of the sample. Some samples will be above the mean and some below. As a result, the positive differences would cancel the negative differences and on average, we would calculate a variance of 0. Squaring the difference of each allows us to get rid of negative values.

We are almost never able to measure an entire population, so we use statistics to estimate parameters of a population. If you could measure the height of every rabbit in a population, it would look something like the graph labeled parent population below. Go to http://onlinestatbook.com/stat_sim/sampling_dist/index.html to experiment with this statistic.

A few really short rabbits

A few really tall rabbits

This is a random sample of the heights of 5 rabbits



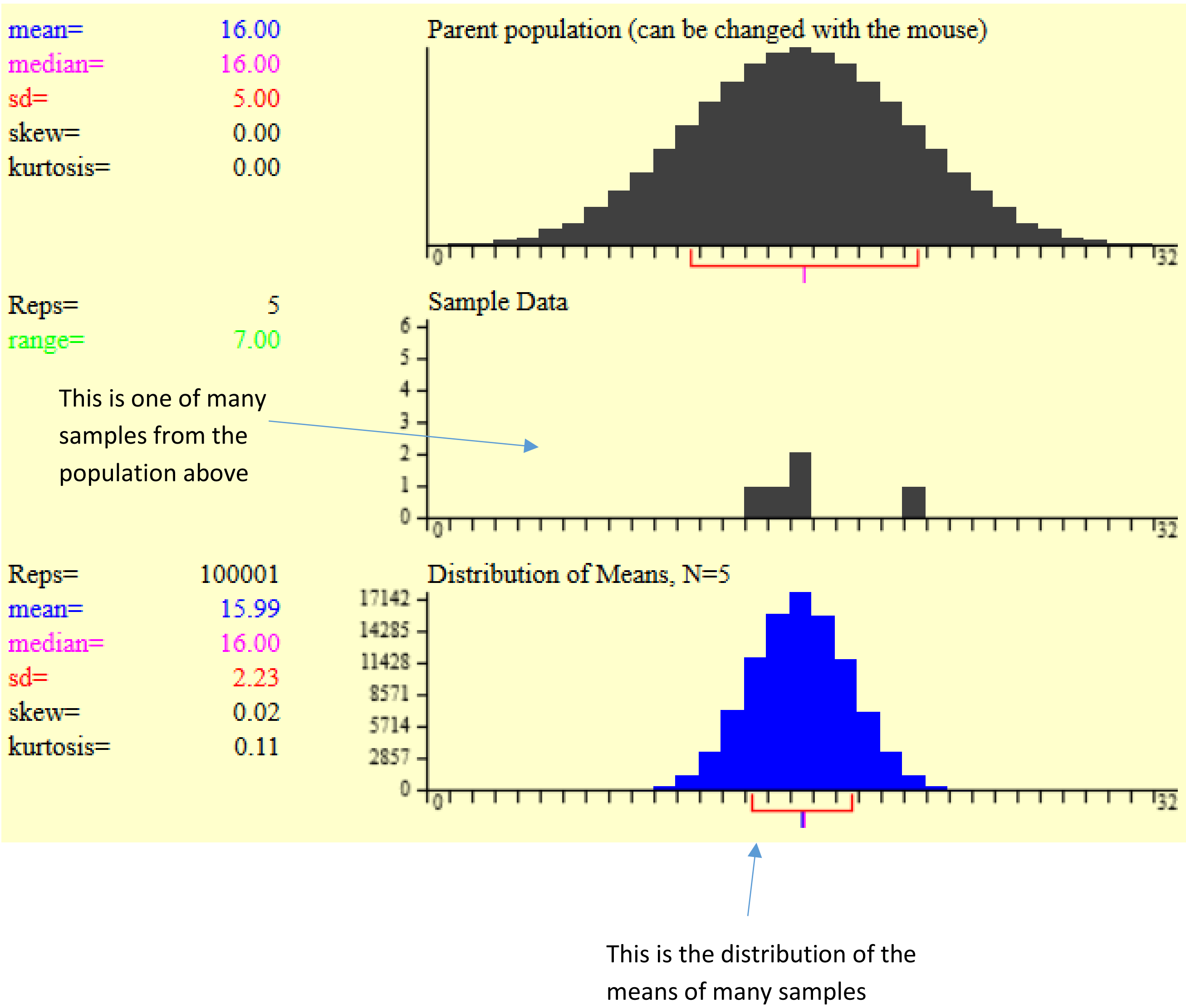
S = 5.89

This is my statistic. The **standard deviation (S) of my sample** is approximately equal to the standard deviation of my population. In symbols this means, $S \approx \sigma$

What is the central limit theorem?

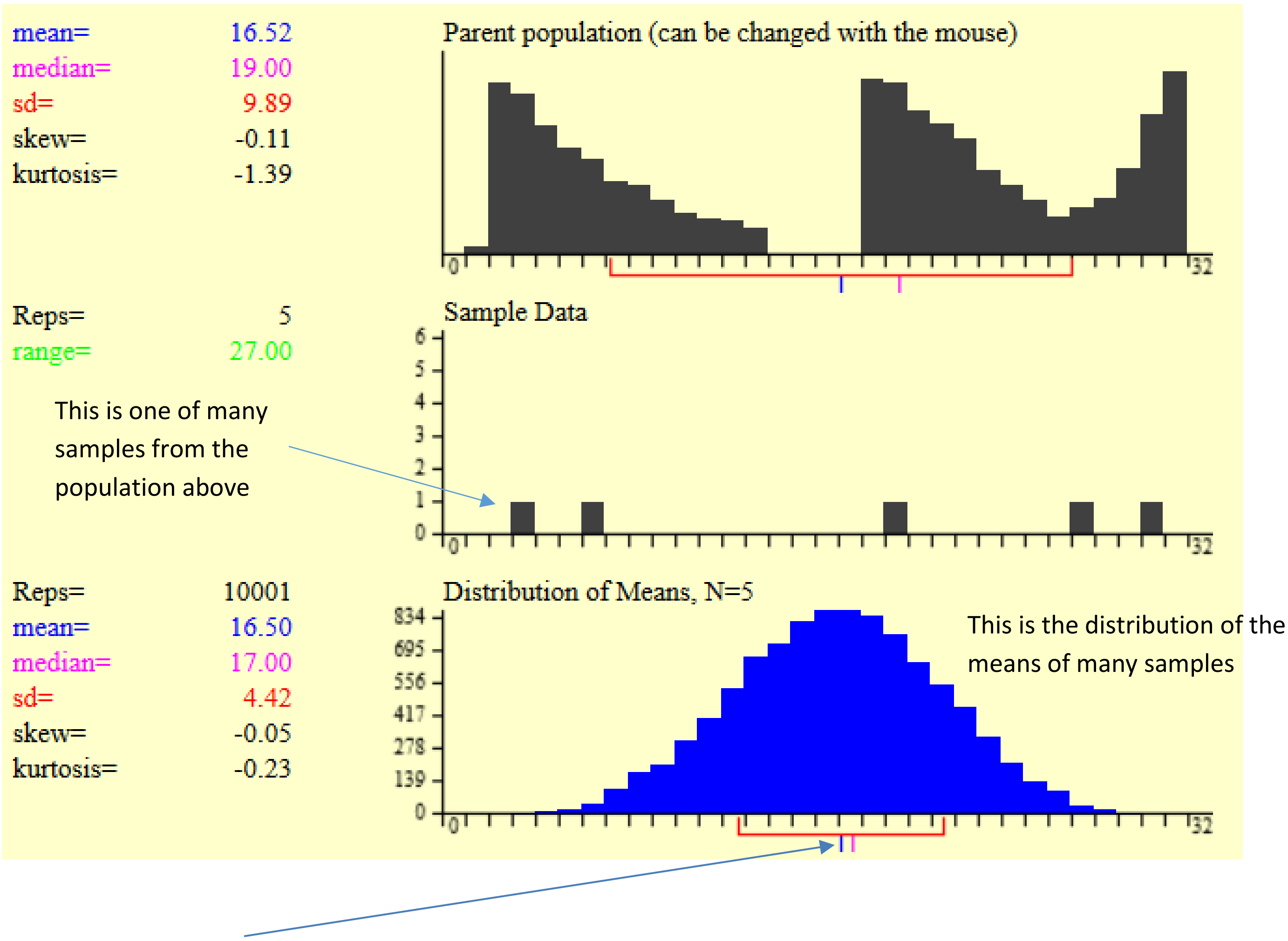
The central limit theorem says that the mean of a large random sample of means taken from a population will be normally distributed regardless of the true distribution of the population.

Back to our rabbit example. The height of our rabbit population is normally distributed. You know this because the skew = 0 and the kurtosis = 0. Skew means there are more large or small individuals than what are found in a normal distribution. Kurtosis means there are more or less individuals around the mean than are found in a normal distribution. Go to http://onlinestatbook.com/stat_sim/sampling_dist/index.html to experiment with this statistic.



Notice the mean of the sample means is nearly identical to the population mean and the distribution of the means is normally distributed (the skew and kurtosis are nearly 0). This isn't surprising, but the same holds even if my population is not normally distributed.

Below is a rabbit population whose height is not normally distributed.



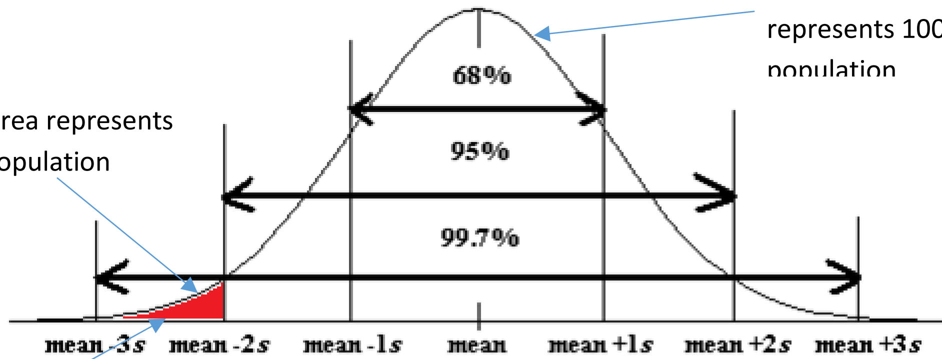
Notice that the <u>mean of the sample</u> means (16.50) is nearly identical to the **true mean** (16.52) of the population and the distribution of the means approaches a normal distribution.

What is the empirical rule?

The empirical rule states that the majority of data clusters around the mean in a normal distribution. More specifically, 68% of all samples collected will be within 1 standard deviation from the mean, 95% of all samples will be within 2 standard deviations from the mean, and 99.7% will fall within 3 standard deviations from the mean.

The area under this curve represents 100% of the population

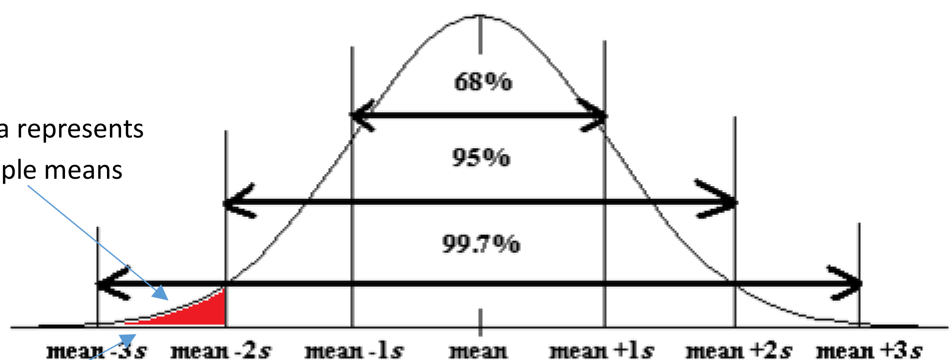
This shaded area represents 2.5% of the population



To continue with our rabbit example, if height is normally distributed in a population, it would be possible, but unlikely to randomly sample a rabbit this short. 97.5% of all the rabbits in this population are taller than this.

What if rabbit height is not normally distributed? It doesn't matter as long as we are using the mean of a sample of something. Remember the central limit theorem? The means of a set of samples will be normally distributed regardless of the distribution of what is being measured.

This shaded area represents 2.5% of the sample means



It would be possible, but unlikely to randomly sample a population and get a mean this small. 97.5% of the time the mean will be larger than this

The fact that means of samples are normally distributed regardless of the distribution of what is being sampled, makes statistical analysis possible. We call this the empirical rule because the mathematics behind it has been experimentally confirmed. As a result, we can put specific numbers to the probability that an experimental result happened by chance. All we need to know to calculate the probability of sampling a particular mean from a population is the standard deviation of the means from repeated sampling of a population. This is called the standard error of the mean.

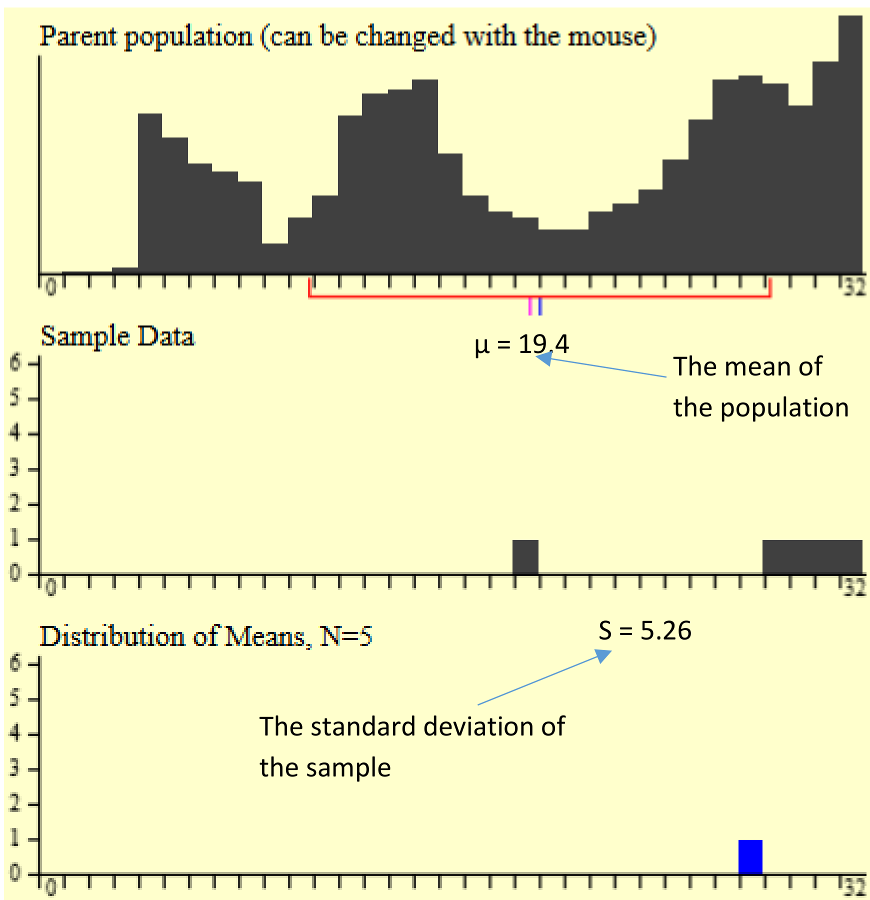
What is standard error of the mean?

It is the standard deviation of the means in a sampling distribution.

The equation below says the standard error of the mean of a sample is equal to the standard deviation of the sample divided by the square root of the sample size. This statistic gives us the expected spread of a sample of means taken from a population based on a single sample. It is used to tell us how far from normal the sample mean is, and it is the basis for the probability statistics used in hypothesis testing.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Below is a rabbit population whose height is not normally distributed. It is the same one used in the discussion about the central limit theorem above.



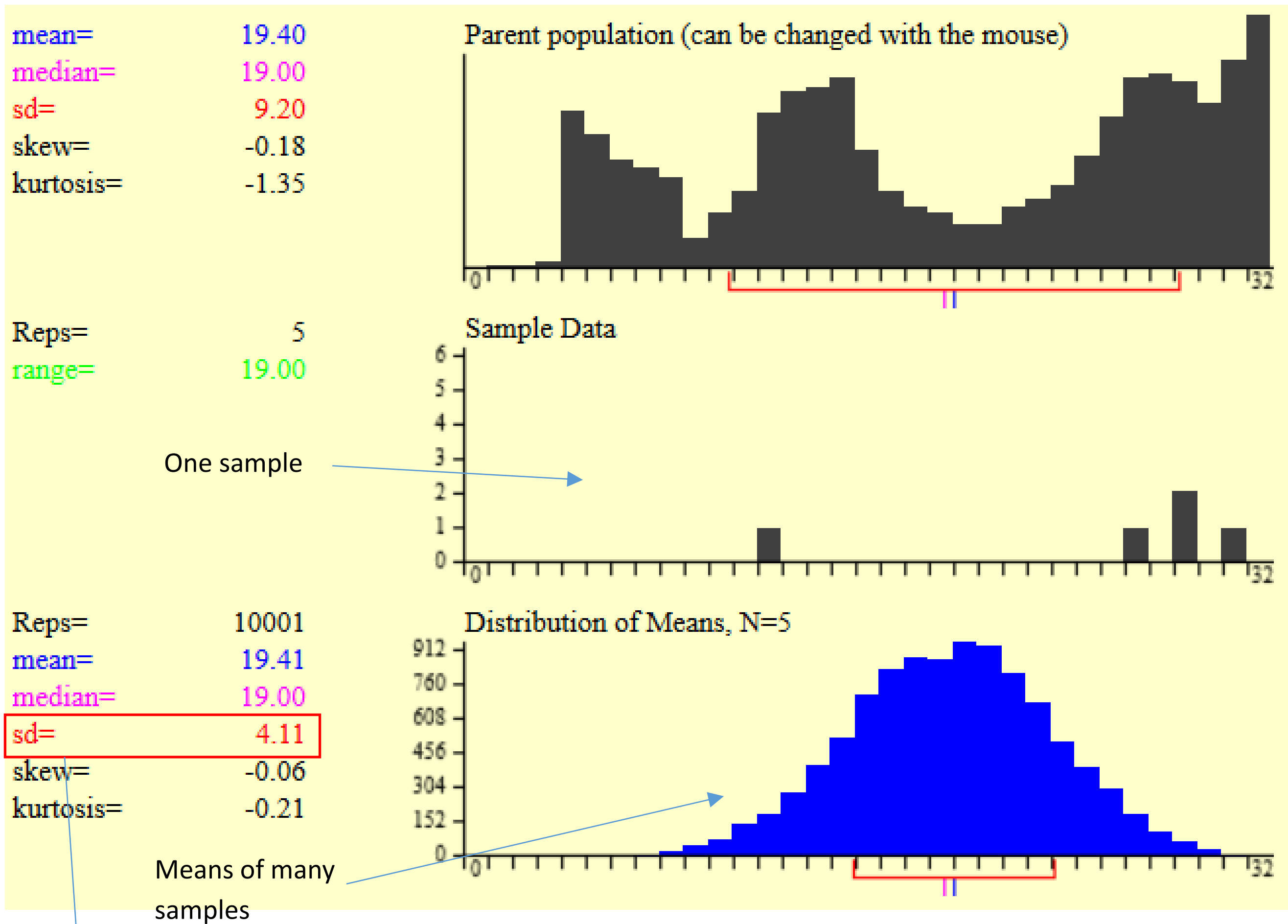
$\bar{X}$ = 28.2

$SE_{\bar{X}}$ = 2.35 (standard error of mean)

The mean of the sample data

Remember that we normally can't see a population. We are using statistics to make inferences about the population. My sample size was small (n=5). My sample mean was not very close to the population mean. If I repeatedly sampled this population with an n=5, I would expect to find a standard deviation of the means of my sampling distribution, which is called the standard error, to be 2.35.

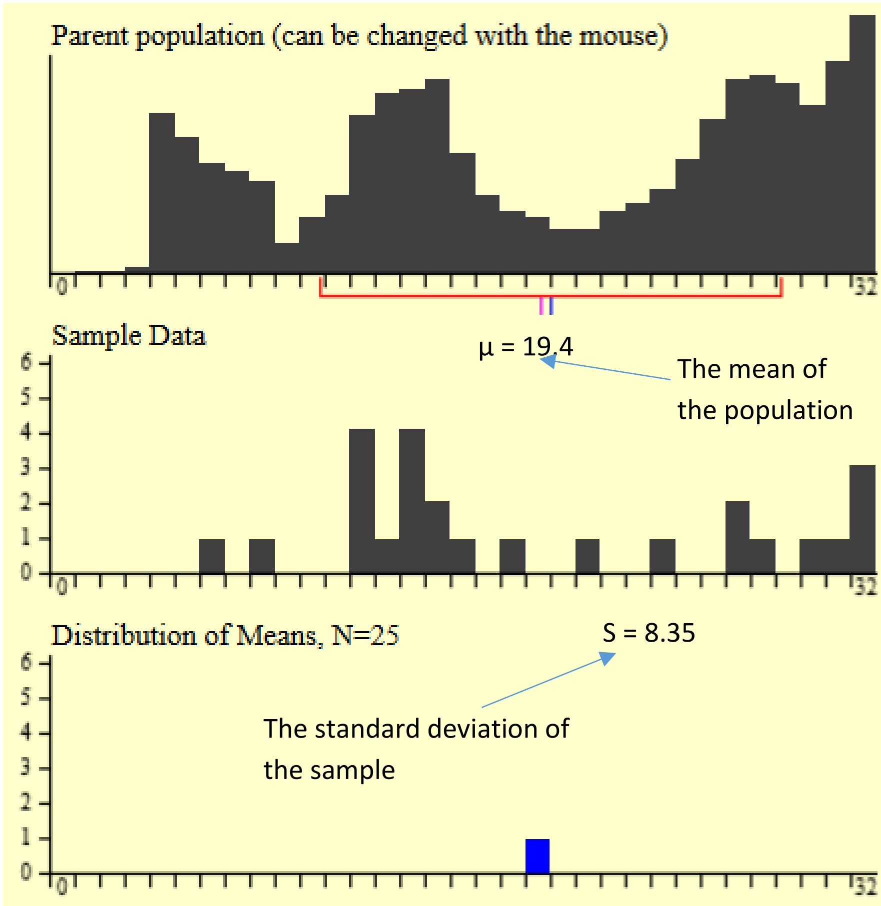
This is the same population used above.



The standard deviation of the many means is 4.11, which is significantly greater than what was predicted by the standard error of the mean of one sample calculated above ($SE_{\bar{X}}$ = 2.35)

The take home lesson is small sample sizes are poor estimates of population parameters. Small sample sizes also result in more spread in the data. As a result, using small sample sizes makes it difficult to determine whether two samples from two different treatments are different. The reason this is true should be evident by looking at the distribution of means above. Large differences between sample means are to be expected. Remember the empirical rule? 68% of our samples will be 19.41 ± 4.11.

This is the same population that was used above. The difference is we are taking a sample of n = 25 rather than n = 5.

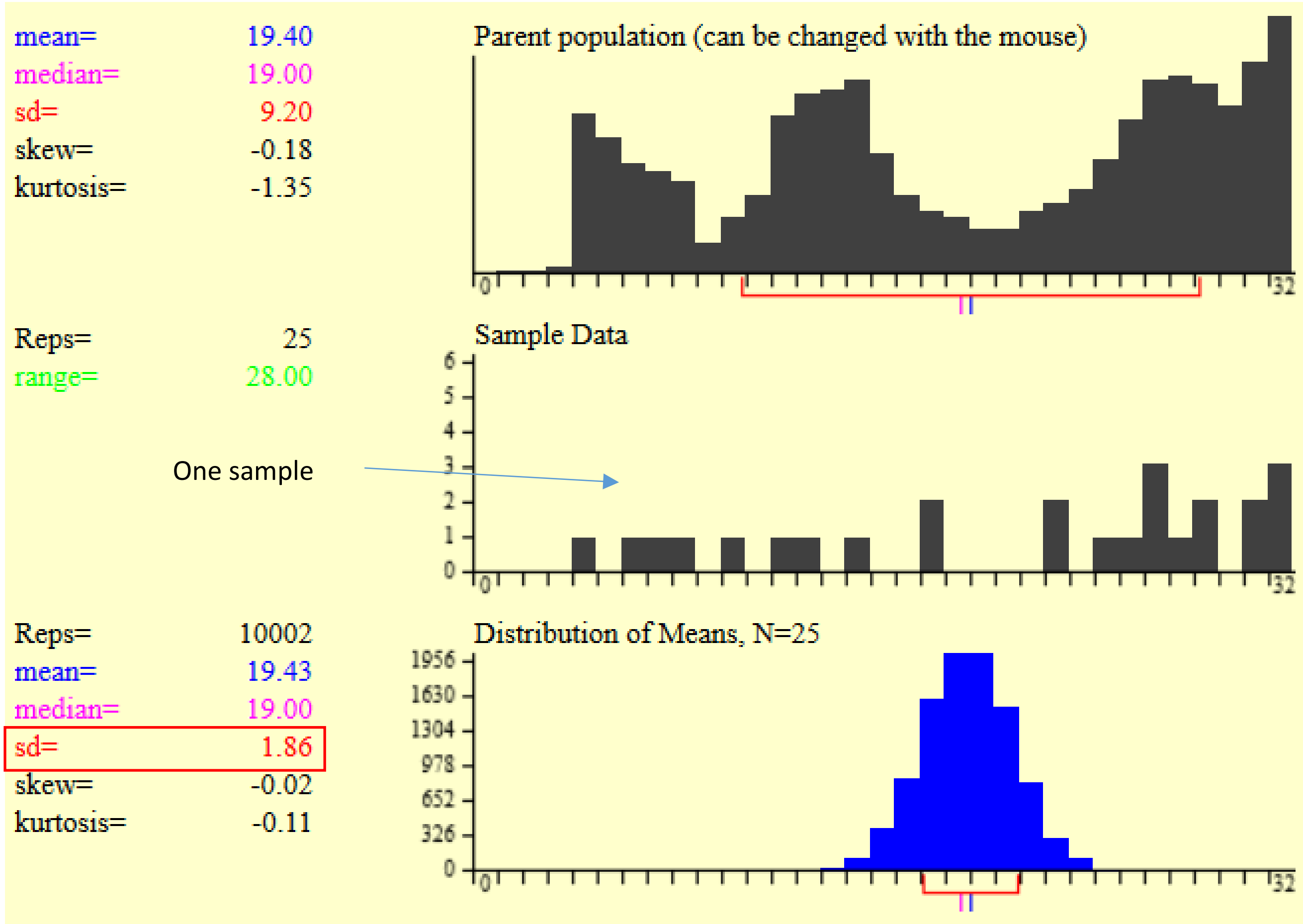


$\bar{X}$ = 19.16 The mean of the sample data

$SE_{\bar{X}}$ = 1.67 (standard error of mean)

Review the results from taking a sample of n = 5 above. With a sample of 25, we have a sample mean that is close to the population mean and a standard error that is significantly smaller. Since my standard error is small, if I repeatedly took n = 25 samples from this population, I would expect the means to all be pretty close to 19.16. Remember the empirical rule? 68% will be 19.16 ± 1.67.

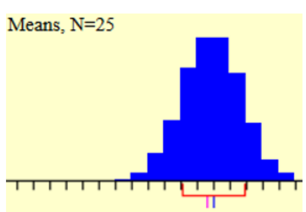
This is the same population used above.

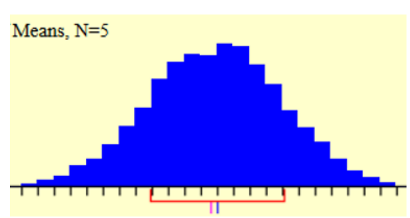


A smaller standard error means it will be easier to determine if differences between two samples taken from two different treatments are different because a small standard error tells us that any two sample means should not be very different. In other words, the probability that 2 samples are different due to random sampling differences decreases as standard error decreases.

You need to be able to explain why large sample sizes are necessary in scientific research using the equation for standard error to the right. Notice that the sample size is in the denominator. This means standard error is inversely proportional to the sample size (n). In other words, as sample size increases standard error decrease.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$



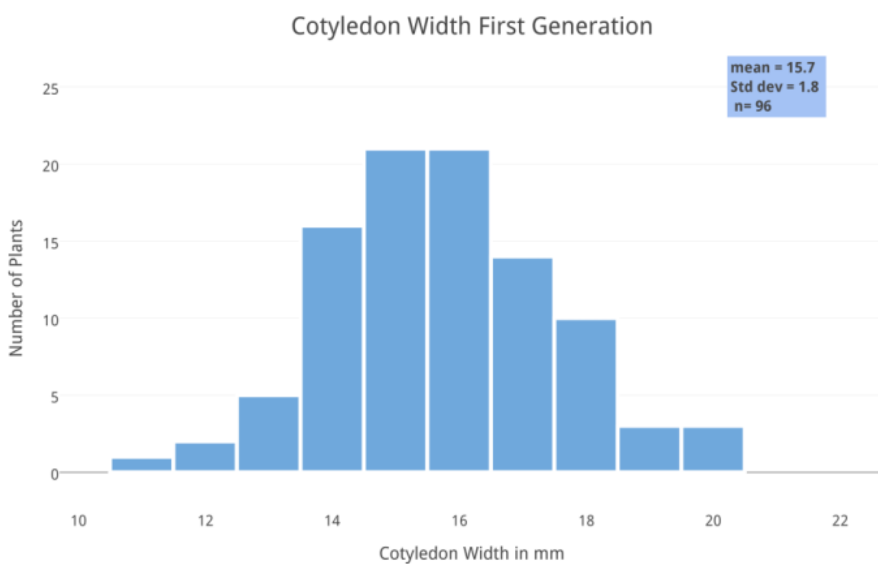


The 2 distributions to the right illustrate the results of the equation for standard error. There is much less variation in means as sample size increases. As a result, large differences between groups are unlikely with a large sample. Conversely, large differences between groups is likely with a small sample size. As the complexity of the phenomenon under investigation increases, so does the need for larger sample sizes.

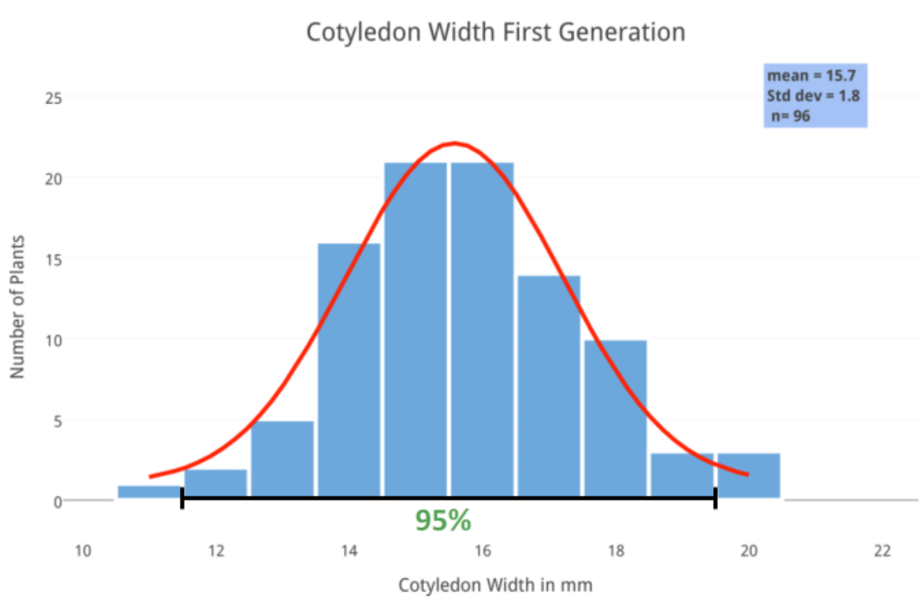
How are graphs with standard deviation bars interpreted?

Cotyledons are the first leaf or leaves that grow on a seed plant. Sometimes called embryonic leaves.

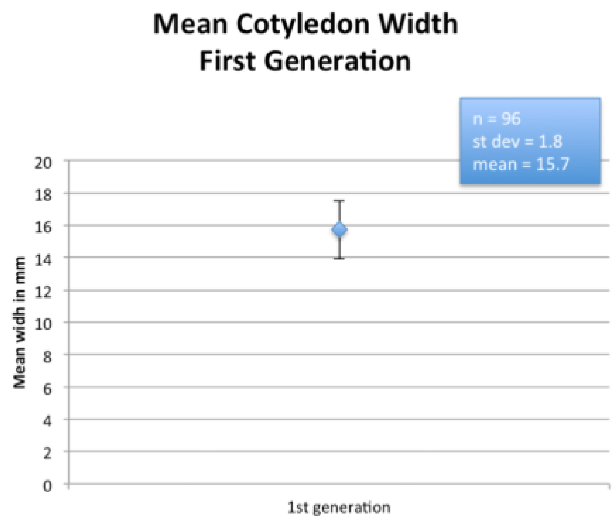
Below is a histogram of a fairly large sample of cotyledon widths measured after a set number of days post germination.

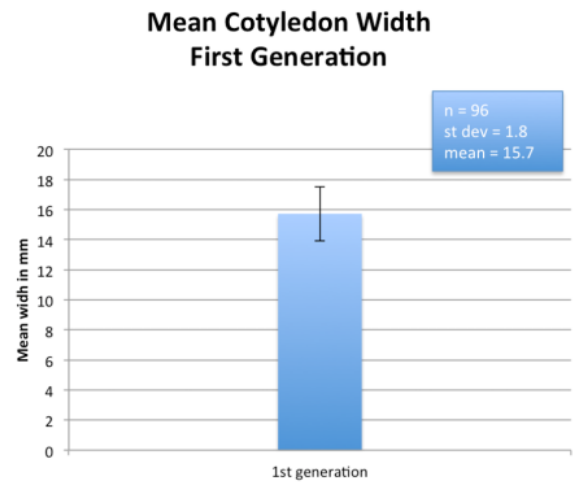


The line marks a normal distribution. Cotyledon width is the type of characteristic that is typically normally distributed. The empirical rule tells us that 68% of random samples will fall within 1 standard deviation from the mean, 95% will fall within 2, and 99.7 will fall within 3. The importance of the graph below is it tells us that it is unlikely to randomly sample a cotyledon that is smaller than 12.1mm or larger than 19.3mm. We know this because the mean is 15.7mm the standard deviation is 1.8mm, and cotyledon width is normally distributed.



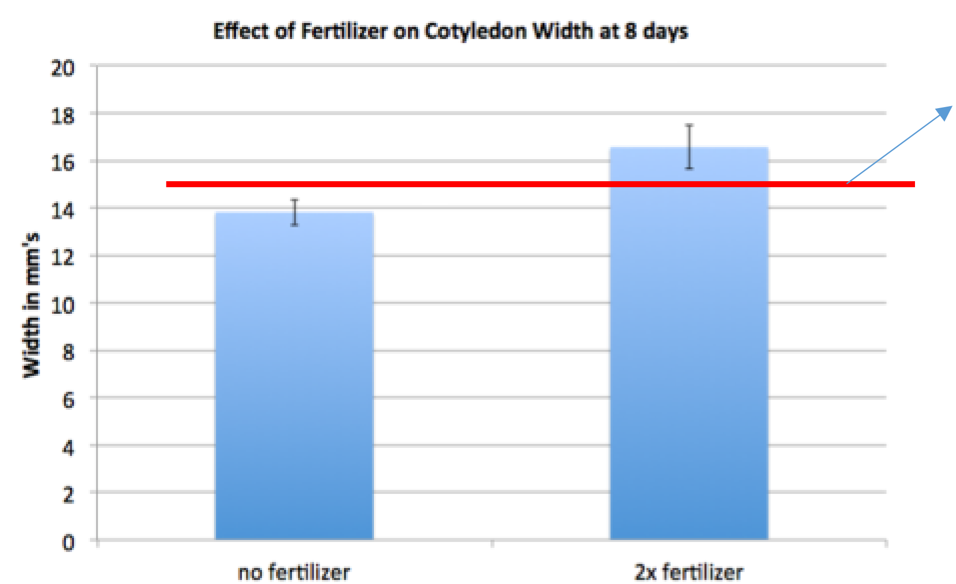
The two graphs below provide identical information and summarize the graph above. The mean is shown as a single point in the graph on the right and as a single bar in the graph on the left. The bars show ± 1 standard deviation. The advantage of a graph like this is it captures the nature of a large data set with a single point and error bars.





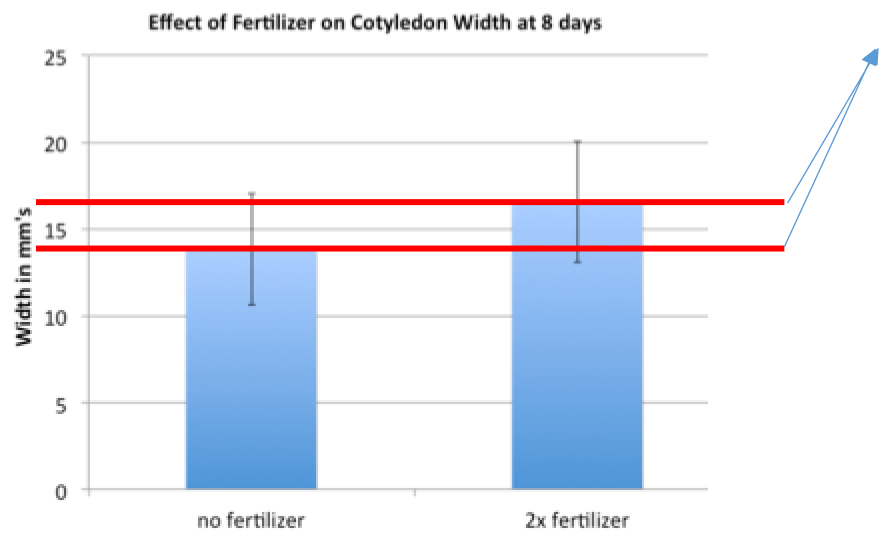
How are graphs with standard error bars interpreted?

The graphs above describe a data set. We often want to test hypotheses, which often involves sampling from 2 different populations that are subjected to 2 different treatments. The graph below summarizes the nature of 2 data sets. The means and ± standard errors of the means are plotted. The question we want to answer is, "Was there a treatment effect?" Remember that means are normally distributed and the standard error of the mean is an estimate of the standard deviation of the distribution of sample means. The empirical rule tells us that 95% of the means for each treatment will fall within the error bars (the bars indicate ±2 standard errors). The means are different, but are they different enough to conclude that the difference is due to the treatment effect and not due to chance?

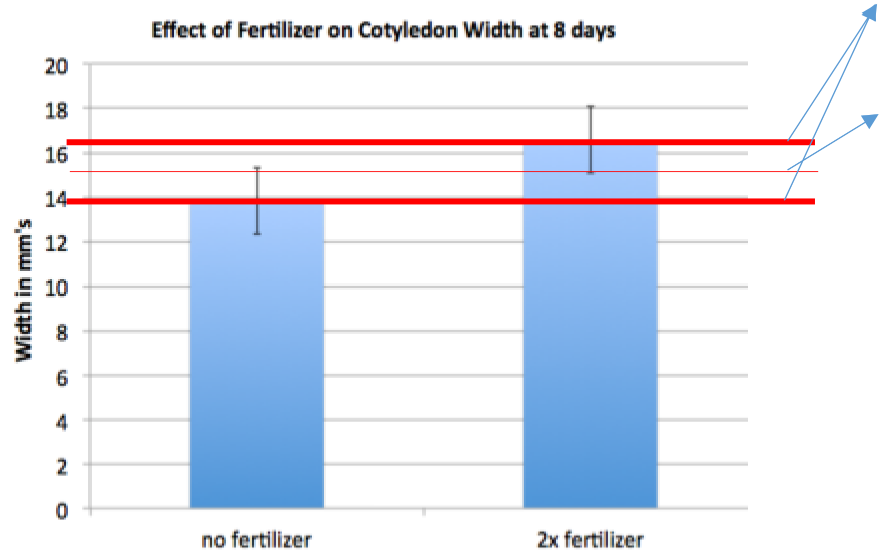


Notice that the standard error bars do not overlap. This means that it would be highly unlikely to randomly sample a mean from either treatment that would be found in the other. Therefore we can conclude that there is a reasonable probability that there was a treatment effect.

Once again we can see that the means of the 2 treatments are different. Are they different enough to conclude that there is a treatment effect?

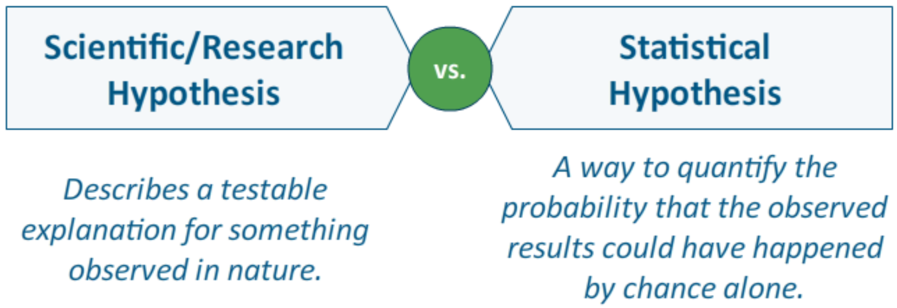


The variation between groups must be larger than the variation within groups to claim with any certainty that there is a treatment effect. There is significant overlap between the expected means of the 2 populations. Additionally, the means of each sample are within 2 standard errors of the mean of the other sample. As a result, despite the fact that the means are different, we cannot say with significant certainty that the difference is due to the treatment. In other words, we cannot rule out the possibility that the difference is due to chance.



The means of the 2 samples do not fall within 2 standard errors of the mean of the other treatment group. However, there is still overlap in the expected means of the 2 groups. We need a more precise measure of the probability that the difference is due to chance. This graph doesn't tell us the there is no treatment effect, nor does it tell us that there is one.

What is the difference between a scientific hypothesis and a statistical hypothesis?



A scientific/research hypothesis could be aspirin relieves headache pain. Hypotheses are explanations. You can think of a hypothesis as a special kind of explanation, one that is testable and falsifiable. A statistical hypothesis is a bit more complicated and counterintuitive. In statistics, we test a null hypothesis (H_O), which always explains the difference by saying it is due to chance. If we reject our null hypothesis, we can say that the alternative hypothesis (H_A) cannot be ruled out at this time.

For example, I want to test whether or not aspirin relieves headache pain. I get a bunch of people with headaches and randomly assign them into 2 groups. One group, the experimental group, gets aspirin. The other group, the control group, gets a sham treatment called a placebo. The experimenter should not know who is in what group and the subjects shouldn't either. This is called a double blind study and it's more difficult to accomplish than it sounds.

My null hypothesis (H_O) would be differences in control and treatment groups are due to chance.

My alternative hypothesis (H_A) would be differences in control and treatment groups are caused by the treatment (aspirin).

There are many different and confusing ways of stating a H_0 and H_A. Keep in mind that statistical hypothesis testing is a way of assigning a specific probability that the difference in treatment groups was due to chance.

There are 2 possible outcomes to my experiment: 1) differences in control group and experimental group are unlikely due to chance, or 2) differences in control group and experimental group are likely due to chance. If the results are unlikely to be due to chance we reject the null hypothesis, and if the results are likely to be due to chance we accept the null hypothesis.

Notice that we never prove our alternative hypothesis. In our aspirin example, we didn't prove that the aspirin relieved headaches. All we can say is the reduction in pain in our aspirin receiving group was not likely due to chance.

For more on the complexity of designing quality studies go to https://www.youtube.com/watch?v=MT9QPKz2a-8&index=9&list=PL8MfjLNsf_miVcNu6eJMNigAMNwQkk_B9. For a discussion on statistical significance, start at 19:03.

How do I know whether or not a difference between 2 or more treatment groups is significant?

Review the central limit theorem and the empirical rule. The central limit theorem tells us that if we take large random samples of a population, the distribution of the sample means will be normally distributed regardless of the distribution of the characteristic we are measuring. The empirical rule tells us that 68% of values from a normally distributed parameter will fall within 1 standard deviation from the mean, 95% will fall within 2 standard deviations from the mean, and 99.7% will fall within 3 standard deviations from the mean. Review standard error of the mean. Remarkably, we don't need a distribution of sample means to estimate the standard deviation of a distribution of sample means. The standard deviation of the distribution of sample means is estimated by the standard error statistic. The equation is given below.

$$SE_{\overline{x}} = \frac{s}{\sqrt{n}}$$

The equation above and the empirical rule tell us that if we have large random samples from two or more populations, it is unlikely that the means will be very different unless the means really are different.

The statement above begs the question, "how different is very different?"

There are 2 kinds of errors we can make. One is thinking 2 things are different when they aren't (type I error), and the other is not thinking 2 things aren't different when they are (type II error). Type I and type II errors are defined by whether we accept or reject the null hypothesis and by whether the null hypothesis is true or false. A type I error is called a false positive, which is rejecting a true null hypothesis. A type II error is called a false negative, which means accepting a false null hypothesis. This is confusing because false null is a double negative. For example, most spiders are harmless to people. In fact the vast majority of spiders are beneficial to people (they eat stuff that annoy us and other things that spread disease and eat our food). If I find a random spider it might be dangerous or harmless and I might kill it or let it live. If I kill a harmless spider, I've made a type I error. I failed to see that there was no danger. If I let a dangerous spider live, I've made a type II error. I failed to recognize danger.

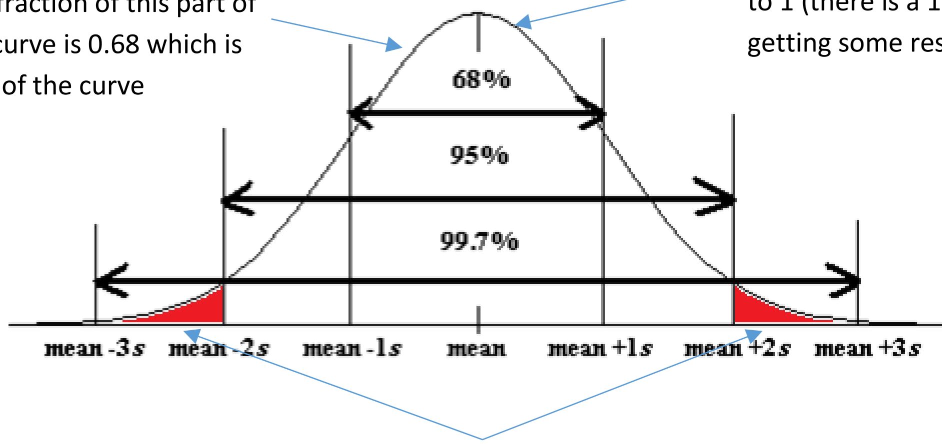
This was a long winded way of saying we need to strike a balance between finding a difference when there isn't (type I error) and not finding a difference when there is (type II error). We say a difference is significant when we think the probability of the difference strikes a balance between making these 2 types of errors.

What does statistically significant mean?

When we think there is a 5% or less probability that the difference found in an experimental outcome is due to chance, we say the difference is statistically significant. It is important to understand what statistically significant means. It doesn't mean the null hypothesis is false. It means the difference is unlikely to be due to chance, so the null hypothesis might be false. Additionally, it doesn't mean the alternative hypothesis is true. It means we can't rule out the possibility that it is true. Statistical tests return p-values, which are fractions of the area under a probability distribution curve like the one below.

The area under this curve adds up to 1 (there is a 100% probability of getting some result)

The fraction of this part of the curve is 0.68 which is 68% of the curve



Let's say we want to know if a population of rabbits at one golf course is taller than a population of rabbits at another golf course. I take a large random sample from each population and determine the mean of each. Then I run a statistical test, like a Student's t-test, to see if the results are statistically different. If I get a p-value of 0.05 or less, it means the probability of randomly sampling a difference of that magnitude is 5%. In other words the difference is more than 2 standard deviations from a mean difference of 0.

This also means that if I were to run the exact same protocol for sampling rabbits from the 2 populations 20 times and the populations were NOT different, I would get the result I got 1 time. This is important to understand so I'll repeat it. If there is no difference in 2 or more populations or treatment groups or whatever else we are studying, I will randomly sample a difference 1 out of 20 times due to chance alone. A p-value of 0.05 tells us the probability of making a type I error (getting a false positive). The reason we don't use a lower p-value for determining significance is because as we decrease the probability of making a type I error (false positive), we increase the probability of making a type II error (false negative).

All of this assumes my protocol was flawless. There are more than a million medical papers published each year. The majority show statistically significant effects. This means 10s of thousands of false positive results are published every year. It's worse than that because medical research is expensive, so research often starts with small sample sizes. It's really hard to randomly sample a human population and it is often immoral to use humans in medical trials. As a result, animal models are often used that can only approximate what might happen in a human. Worse still is the unavoidable bias and financial incentive to "prove" something has an effect. What this all means is most published medical research is probably wrong!

What statistical tests do I need know?

You need to be able to interpret any statistical test. The good news is they all give p-values so once you learn how to interpret one test, you know how to interpret all tests!

What is the chi-squared test?

This is the only statistical test you are required to know how to calculate by hand. The chi-squared test (χ^2) is useful for determining if an outcome is different than expected. It could be used to determine experimentally whether or not a coin is weighted. If I flipped a coin 100 times and got heads 60 times, can I say the coin is weighted? The chi-squared statistic tells us the probability that a difference in outcomes from expectation is due to chance alone. Below is the equation for calculating the chi-squared statistic and the abbreviated chi-squared probability distribution table that you will be given on the AP exam. Take a look at the larger distribution table above. It may help you wrap you mind around this idea.

This equation reads the chi-squared statistic equals the sum of the observed minus the expected outcomes squared divided by the expected number of outcomes

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

Let's return to our coin problem. You flipped a coin 100 times and got heads 60 times. You suspect that the coin is weighted because you know there is an equal chance of getting heads and tails. That is your alternative hypothesis. The null hypothesis is the difference is due to chance.

Outcome	Observed	Expected	$\frac{(o-e)^2}{e}$
Heads	60	50	2
Tails	40	50	+ 2
		χ^2 =	4

$\chi^2 = 4$. Is this a statistically significant result? In other words, can we reject the null hypothesis? We have 2 possible outcomes, so our degrees of freedom is 1. Degrees of freedom is possible outcomes minus 1. See the discussion below.

4 is a bit larger than 3.84. This means our p-value is little less than 0.05. This means we can reject the null hypothesis and entertain the possibility that the coin is weighted because you would expect to get 60 heads out of 100 a little less than 1 out of every 20 times you flipped a coin 100 times.

$\chi^2 = 4$

CHI-SQUARE TABLE

	Degrees of Freedom							
p	1	2	3	4	5	6	7	8
0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51
0.01	6.64	9.21	11.34	13.28	15.09	16.81	18.48	20.09

What are degrees of freedom?

Degrees of freedom is the number of ways something can vary. It is calculated by subtracting 1 from the number of possible outcomes of an experiment. The equation is $df = n-1$ were n is the number of possible outcomes.

For example, you could be asked to figure out if an inheritance pattern follows Mendelian inheritance laws. Do blue and brown eyes follow a Mendelian inheritance pattern? Let's assume we have a population that only has blue and brown eyed individuals. Since individuals can either be blue eyed or brown eyed, we have 2 possible outcomes, so our degrees of freedom is 2-1=1.

Imagine we have a population that either has curly or straight hair and blue or brown eyes. Do these 2 characteristics follow a Mendelian inheritance pattern? You have 4 possible outcomes: 1) blue/straight, 2) blue/curly, 3) brown/straight, 4) brown/curly. Degrees of freedom is 4-1 = 3.

Glossary

Alternative Hypothesis
In hypothesis testing, the null hypothesis and an alternative hypothesis are put forward. If the data are sufficiently strong to reject the null hypothesis, then the null hypothesis is rejected in favor of an alternative hypothesis. For instance, if the null hypothesis were that $\mu_1 = \mu_2$ then the alternative hypothesis (for a two-tailed test) would be $\mu_1 \neq \mu_2$.

Analysis of Variance
Analysis of variance is a method for testing hypotheses about means. It is the most widely-used method of statistical inference for the analysis of experimental data.

Average
(i) The (arithmetic) mean

(ii) Any measure of central tendency

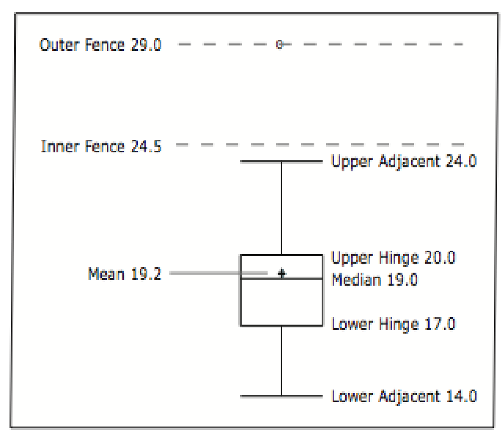
Bar Chart
A graphical method of presenting data. A bar is drawn for each level of a variable. The height of each bar contains the value of the variable. Bar charts are useful for displaying things such as frequency counts and percent increases. They are not recommended for displaying means (despite the widespread practice) since box plots present more information in the same amount of space.

Bias
1. A sampling method is biased if each element does not have an equal chance of being selected. A sample of internet users found reading an online statistics book would be a biased sample of all internet users. A random sample is unbiased. Note that possible bias refers to the sampling method, not the result. An unbiased method could, by chance, lead to a very non-representative sample.

2. An estimator is biased if it systematically overestimates or underestimates the parameter it is estimating. In other words, it is biased if the mean of the sampling distribution of the statistic is not the parameter it is estimating, The sample mean is an unbiased estimate of the population mean. The mean squared deviation of sample scores from their mean is a biased estimate of the variance since it tends to underestimate the population variance.

Bimodal Distribution
A distribution with two distinct peaks.

Box Plot
One of the more effective graphical summaries of a data set, the box plot generally shows mean, median, 25th and 75th percentiles, and outliers. A standard box plot is composed of the median, upper hinge, lower hinge, higher adjacent value, lower adjacent value, outside values, and far out values. An example is shown below. Parallel box plots are very useful for comparing distributions.



Central Tendency
There are many measures of the center of a distribution. These are called measures of central tendency. The most common are the mean, median, and, mode. Others include the trimean, trimmed mean, and geometric mean.)

Confidence Interval
A confidence interval is a range of scores likely to contain the parameter being estimated. Intervals can be constructed to be more or less likely to contain the parameter: 95% of 95% confidence intervals contain the estimated parameter whereas 99% of 99% confidence intervals contain the estimated parameter. The wider the confidence interval, the more uncertainty there is about the value of the parameter.

Confounding
Two or more variables are confounded if their effects cannot be separated because they vary together. For example, if a study on the effect of light inadvertently manipulated heat along with light, then light and heat would be confounded.

Constant
A value that does not change. Values such as π, or the mass of the Earth are constants.

Continuous Variables
Variables that can take on any value in a certain range. Time and distance are continuous; gender, SAT score and "time rounded to the nearest second" are not. Variables that are not continuous are known as discrete variables. No measured variable is truly continuous; however, discrete variables measured with enough precision can often be considered continuous for practical purposes.

Dependent Variable
A variable that measures the experimental outcome. In most experiments, the effects of the independent variable on the dependent variables are observed. For example, if a study investigated the effectiveness of an experimental treatment for depression, then the measure of depression would be the dependent variable.

Descriptive Statistics

1. The branch of statistics concerned with describing and summarizing data.
2. A set of statistics such as the mean, standard deviation, and skew that describe a distribution.

Deviation Scores
Scores that are expressed as differences (deviations) from some value, usually the mean. To convert data to deviation scores typically means to subtract the mean score from each other score. Thus, the values 1, 2, and 3 in deviation-score form would be computed by subtracting the mean of 2 from each value and would be -1, 0, 1.

Degrees of Freedom
The degrees of freedom of an estimate is the number of independent pieces of information that go into the estimate. In general, the degrees of freedom for an estimate is equal to the number of values minus the number of parameters estimated en route to the estimate in question. For example, to estimate the population variance, one must first estimate the population mean. Therefore, if the estimate of variance is based on N observations, there are N-1 degrees of freedom.

Discrete Variables
Variables that can only take on a finite number of values are called "discrete variables." All qualitative variables are discrete. Some quantitative variables are discrete, such as performance rated as 1,2,3,4, or 5, or temperature rounded to the nearest degree. Sometimes, a variable that takes on enough discrete values can be considered to be continuous for practical purposes. One example is time to the nearest millisecond.

Distribution
The distribution of empirical data is called a frequency distribution and consists of a count of the number of occurrences of each value. If the data are continuous, then a grouped frequency distribution is used. Typically, a distribution is portrayed using a frequency polygon or a histogram.

Mathematical equations are often used to define distributions. The normal distribution is, perhaps, the best known example. Many empirical distributions are approximated well by mathematical distributions such as the normal distribution.

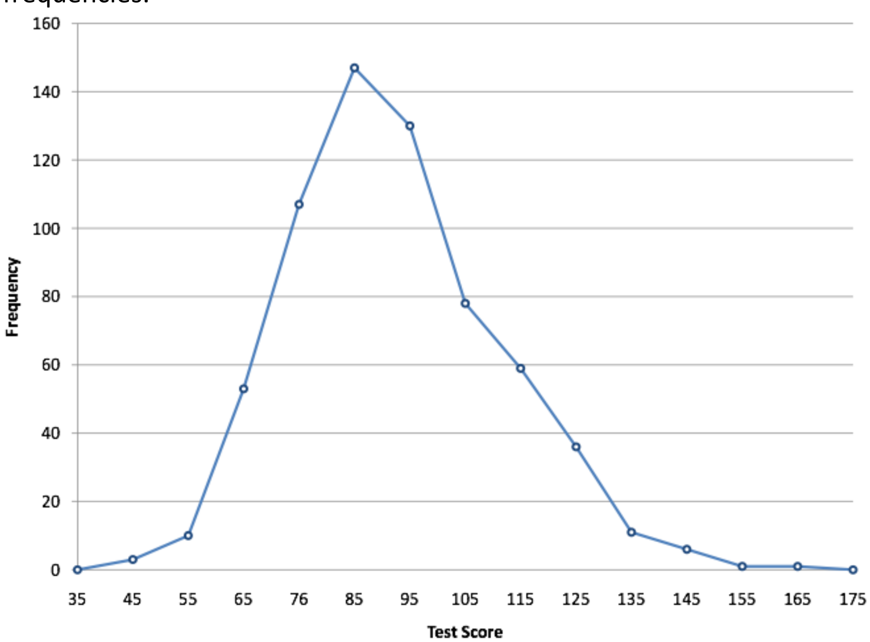
Expected Value
The expected value of a statistic is the mean of the sampling distribution of the statistic. It can be loosely thought of as the long-run average value of the statistic.

False Positive
A false positive occurs when a diagnostic procedure returns a positive result while the true state of the subject is negative. For example, if a test for strep says the patient has strep when in fact he or she does not, then the error in diagnosis would be called a false positive. In some contexts, a false positive is called a false alarm. The concept is similar to a Type I error in significance testing.

Far Out Value
One of the components of a box plot, far out values are those that are more than 2 steps beyond the nearest hinge. They are beyond an outer fence.

Frequency Distribution
For a discrete variable, a frequency distribution consists of the distribution of the number of occurrences for each value of the variable. For a continuous variable, it is the number of occurrences for a variety of ranges of variables.

Frequency Polygon
A frequency polygon is a graphical representation of a distribution. It partitions the variable on the x-axis into various contiguous class intervals of (usually) equal widths. The heights of the polygon's points represent the class frequencies.



Frequency Table
A table containing the number of occurrences in each class of data; for example, the number of each color of M&Ms in a bag. Frequency tables often used to create histograms and frequency polygons. When a frequency table is created for a quantitative variable, a grouped frequency table is generally used.

Grouped Frequency Table
A grouped frequency table shows the number of values for various ranges of scores. Below is shown a grouped frequency table for response times (in milliseconds) for a simple motor task.

Range	Frequency
500-600	3
600-700	6
700-800	5
800-900	5
900-1000	0
1000-1100	1

Grouped Frequency Distribution
A grouped frequency distribution is a frequency distribution in which frequencies are displayed for ranges of data rather than for individual values. For example, the distribution of heights might be calculated by defining one-inch ranges. The frequency of individuals with various heights rounded off to the nearest inch would then be tabulated.

H-Spread
One of the components of a box plot, the H-spread is the difference between the upper hinge and the lower hinge.

Independence
Two variables are said to be independent if the value of one variable provides no information about the value of the other variable. These two variables would be uncorrelated so that Pearson's r would be 0.

Two events are independent if the probability the second event occurring is the same regardless of whether or not the first event occurred.

Independent Events
Events A and B are independent events if the probability of Event B occurring is the same whether or not Event A occurs. For example, if you throw two dice, the probability that the second die comes up 1 is independent of whether the first die came up 1. Formally, this can be stated in terms of conditional probabilities: $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

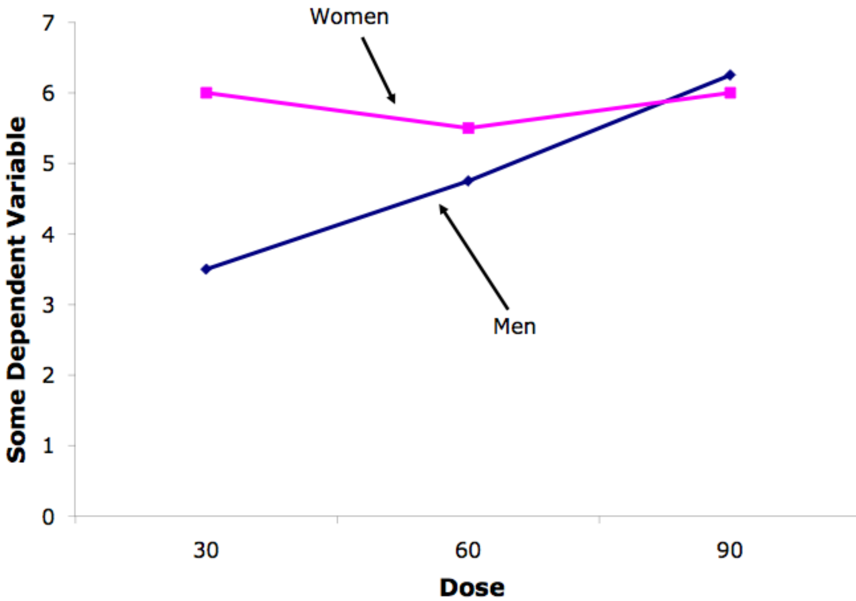
Independent Variable (Factor)
Variables that are manipulated by the experimenter, as opposed to dependent variables. Most experiments consist of observing the effect of the independent variable(s) on the dependent variable(s).

Inferential Statistics
The branch of statistics concerned with drawing conclusions about a population from a sample. This is generally done through random sampling, followed by inferences made about central tendency, or any of a number of other aspects of a distribution.

Inner Fence
In a box plot, the lower inner fence is one step below the lower hinge while the upper inner fence is one step above the upper hinge.

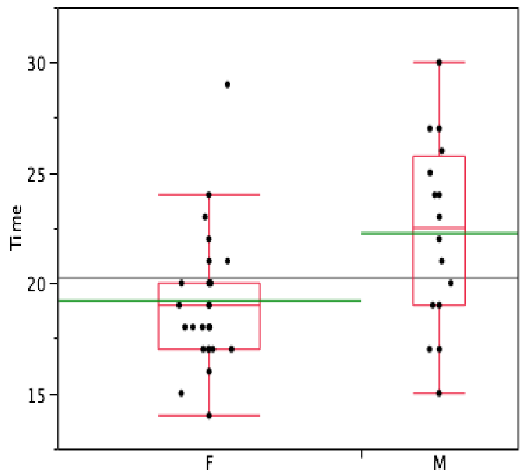
Interaction
Two independent variables interact if the effect of one of the variables differs depending on the level of the other variable.

Interaction Plot
An interaction plot displays the levels of one variable on the X axis and has a separate line for the means of each level of the other variable. The Y axis is the dependent variable. A look at this graph shows that the effect of dosage is different for males than it is for females.



Interquartile Range
The Interquartile Range (IQR) is the 75th percentile minus the 25th percentile. It is a robust measure of variability.

Jitter
When points in a graph are jittered, they are moved horizontally so that all the points can be seen and none are hidden due to overlapping values. An example is shown below:



Kurtosis
Kurtosis measures how fat or thin the tails of a distribution are relative to a normal distribution. It is commonly defined as:

$$\sum \frac{(X-\mu)^4}{N\sigma^4} - 3$$

Distributions with long tails are called leptokurtic; distributions with short tails are called platykurtic. Normal distributions have zero kurtosis.

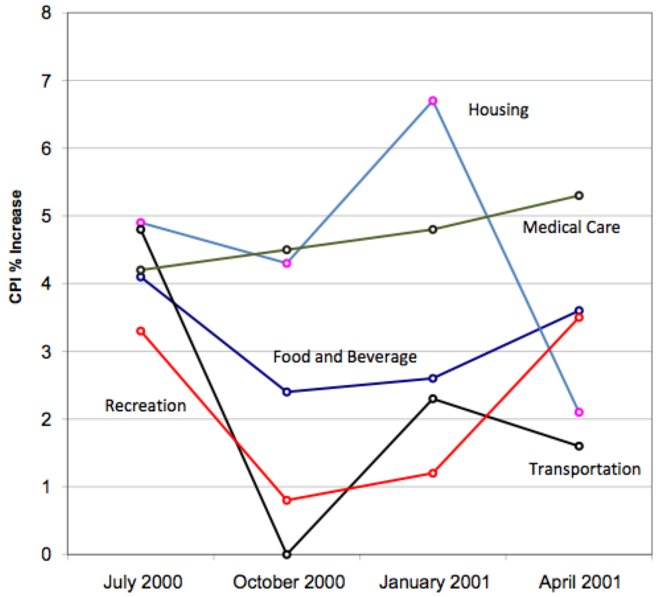
Level
When a factor consists of various treatment conditions, each treatment condition is considered a level of that factor. For example, if the factor were drug dosage, and three doses were tested, then each dosage would be one level of the factor and the factor would have three levels.

Levels of Measurement
Measurement scales differ in their level of measurement. There are four common levels of measurement:

1. Nominal scales are only labels.
2. Ordinal Scales are ordered but are not truly quantitative. Equal intervals on the ordinal scale do not imply equal intervals on the underlying trait.
3. Interval scales are are ordered and equal intervals equal intervals on the underlying trait. However, interval scales do not have a true zero point.
4. Ratio scales are interval scales that do have a true zero point. With ratio scales, it is sensible to talk about one value being twice as large as another, for example.

Line Graph
Essentially a bar graph in which the height of each par is represented by a single point, with each of these points connected by a line. Line graphs are best used to show change over time, and should not be used if your X-axis is not

an ordered variable. An example is shown below.



Linear Regression
Linear regression is a method for predicting a criterion variable from one or more predictor variable. In simple regression, the criterion is predicted from a single predictor variable and the best-fitting straight line is of the form

$$Y' = bX + A$$

where Y' is the predicted score, X is the predictor variable, b is the slope, and A is the Y intercept. Typically, the criterion for the "best fitting" line is the line for which the sum of the squared errors of prediction is minimized. In multiple regression, the criterion is predicted from two or more predictor variables.

Linear Relationship
There is a perfect linear relationship between two variables if a scatterplot of the points falls on a straight line. The relationship is linear even if the points diverge from the line as long as the divergence is random rather than being systematic.

Logarithm
The logarithm of a number is the power the base of the logarithm has to be raised to in order to equal the number. If the base of the logarithm is 10 and the number is 1,000, then the log is 3 since 10 has to be raised to the 3rd power to equal 1,000.

Lower Adjacent Value
A component of a box plot, the lower adjacent value is smallest value in the data above the inner lower fence.

Lower Hinge
A component of a box plot, the lower hinge is the 25th percentile. The upper hinge is the 75th percentile.

Margin of Error
When a statistic is used to estimate a parameter, it is common to compute a confidence interval. The margin of error is the difference between the statistic and the endpoints of the interval. For example, if the statistic were 0.6 and the confidence interval ranged from 0.4 to 0.8, then the margin of error would be 0.20. Unless otherwise specified, the 95% confidence interval is used.

Mean
Also known as the arithmetic mean, the mean is typically what is meant by the word "average." The mean is perhaps the most common measure of central tendency. The mean of a variable is given by (the sum of all its values)/(the number of values). For example, the mean of 4, 8, and 9 is 7. The sample mean is written as M, and the population mean as the Greek letter mu (μ). Despite its popularity, the mean may not be an appropriate measure of central tendency for skewed distributions, or in situations with outliers. Other than the arithmetic mean, there is the geometric mean and the harmonic mean.

Median
The median is a popular measure of central tendency. It is the 50th percentile of a distribution. To find the median of a number of values, first order them, then find the observation in the middle: the median of 5, 2, 7, 9, and 4 is 5. (Note that if there is an even number of values, one takes the average of the middle two: the median of 4, 6, 8, and 10 is 7.) The median is often more appropriate than the mean in skewed distributions and in situations with outliers.

Misses
Misses occur when a diagnostic test returns a negative result, but the true state of the subject is positive. For example, if a person has strep throat and the diagnostic test fails to indicate it, then a miss has occurred. The concept is similar to a Type II error in significance testing.

Mode
The mode is a measure of central tendency. It is the most frequent value in a distribution: the mode of 3, 4, 4, 5, 5, 5, 8 is 5. Note that the mode may be very different from the mean and the median.

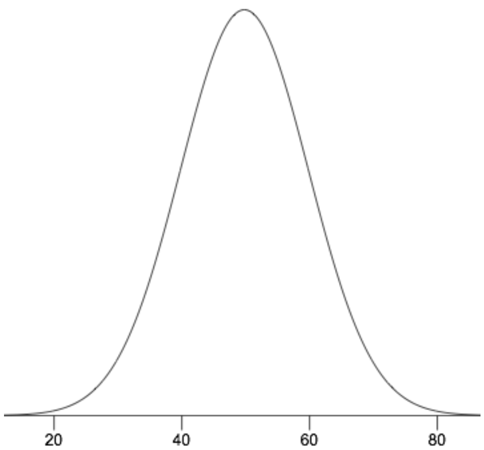
Nominal Scales
A nominal scale is one of four commonly-used levels of measurement. No ordering is implied, and addition/subtraction and multiplication/division would be inappropriate for a variable on a nominal scale. {Female, Male} and {Buddhist, Christian, Hindu, Muslim} have no natural ordering (except alphabetic). Occasionally, numeric values are nominal: for instance, if a variable were coded as Female = 1, Male =2, the set {1,2} is still nominal.

Non-representative
A non-representative sample is a sample that does not accurately reflect the population.

Normal Distribution
One of the most common continuous distributions, a normal distribution is sometimes referred to as a "bell-shaped distribution." If μ is the distribution mean, and σ the standard deviation, then the height (ordinate) of the normal distribution is given by

$$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

A graph of a normal distribution with a mean of 50 and a standard deviation of 10 is shown below.



If the mean is 0 and the standard deviation is 1, the distribution is referred to as the "standard normal distribution."

Null Hypothesis
A null hypothesis is a hypothesis tested in significance testing. It is typically the hypothesis that a parameter is zero or that a difference between parameters is zero. For example, the null hypothesis might be that the difference between population means is zero. Experimenters typically design experiments to allow the null hypothesis to be rejected.

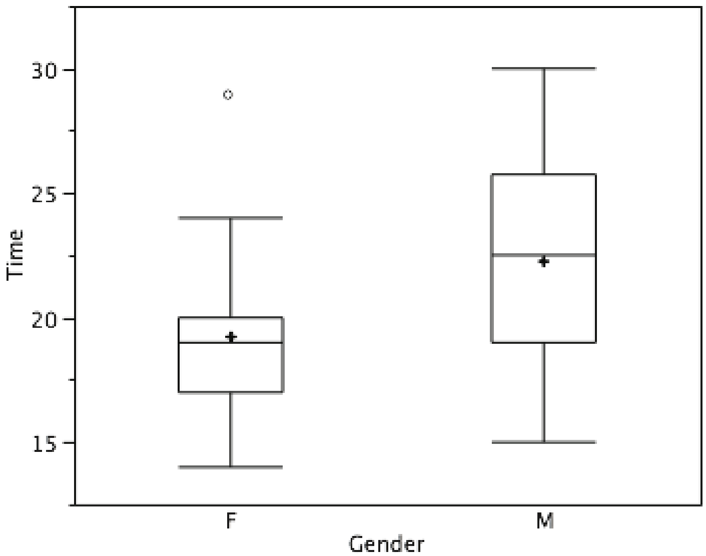
Ordinal Scales
One of four commonly-used levels of measurement, an ordinal scale is a set of ordered values. However, there is no set distance between scale values. For instance, for the scale: (Very Poor, Poor, Average, Good, Very Good) is an ordinal scale. You can assign numerical values to an ordinal scale: rating performance such as 1 for "Very Poor," 2 for "Poor," etc, but there is no assurance that the difference between a score of 1 and 2 means the same thing as the difference between a score of and 2 and 3.

Outer Fence
In a box plot, the lower outer fence is two steps below the lower hinge whereas the upper inner fence is two steps above the upper hinge.

Outlier
Outliers are atypical, infrequent observations; values that have an extreme deviation from the center of the distribution. There is no universally-agreed on criterion for defining an outlier, and outliers should only be discarded with extreme caution. However, one should always assess the effects of outliers on the statistical conclusions.

Outside Values
A component of a box plot, outside values are more than one step beyond the nearest hinge but not more than two steps. They are beyond an inner fence but not beyond an outer fence.

Parallel Box Plots
Two or more box plots drawn on the same Y-axis. These are often useful in comparing features of distributions. An example portraying the times it took samples of women and men to do a task is shown below.



Parameter
A value calculated in a population. For example, the mean of the numbers in a population is a parameter. Compare with a statistic, which is a value computed in a sample to estimate a parameter.

Pearson's r
Pearson's correlation is a measure of the strength of the linear relationship between two variables. It ranges from -1 for a perfect negative relationship to +1 for a perfect positive relationship. A correlation of 0 means that there is no linear relationship.

Percentiles
There is no universally accepted definition of a percentile. Using the 65th percentile as an example, some statisticians define the 65th percentile as the lowest score that is *greater* than 65% of the scores. Others have defined the 65th percentile as the lowest score that is *greater than or equal* to 65% of the scores. A more sophisticated definition is given below.

The first step is to compute the rank (R) of the percentile in question. This is done using the following formula:

$$R = P/100 \times (N + 1)$$

where P is the desired percentile and N is the number of numbers. If R is an integer, then the Pth percentile is the number with rank R. When R is not an integer, we compute the Pth percentile by interpolation as follows:

1. Define IR as the integer portion of R (the number to the left of the decimal point).
2. Define FR as the fractional portion or R.
3. Find the scores with Rank IR and with Rank IR + 1.

4. Interpolate by multiplying the difference between the scores by FR and add the result to the lower score.

Pie Chart
A graphical representation of data, the pie chart shows relative frequencies of classes of data. It is a circle cut into a number of wedges, one for each class, with the area of each wedge proportional to its relative frequency. Pie charts are only effective for a small number of classes, and are one of the less effective graphical representations.

Placebo
A device used in clinical trials, the placebo is visually indistinguishable from the study medication, but in reality has no medical effect (often, a sugar pill). A group of subjects chosen randomly takes the placebo, the others take one or another type of medication. This is done to prevent confounding the medical and psychological effects of the drug.

Population
A population is the complete set of observations a researcher is interested in. Contrast this with a sample which is a subset of a population. A population can be defined in a manner convenient for a researcher. For example, one could define a population as all girls in fourth grade in Houston, Texas. Or, a different population is the set of all girls in fourth grade in the United States. Inferential statistics are computed from sample data in order to make inferences about the population.

Positive Association
There is a positive association between variables X and Y if smaller values of X are associated with smaller values of Y and larger values of X are associated with larger values of Y.

Predictor
A predictor variable is a variable used in regression to predict another variable. It is sometimes referred to as an independent variable if it is manipulated rather than just measured.

Probability Distribution
For a discrete random variable, a probability distribution contains the probability of each possible outcome. The sum of all probabilities is always 1.0. See binomial distribution for an example.

Probability Value
In significance testing, the probability value (sometimes called the p value) is the probability of obtaining a statistic as different or more different from the parameter specified in the null hypothesis as the statistic obtained in the experiment. The probability value is computed assuming the null hypothesis is true. The lower the probability value, the stronger the evidence that the null hypothesis is false. Traditionally, the null hypothesis is rejected if the probability value is below 0.05.

Qualitative Variable
Also known as categorical variables, qualitative variables are variables with no natural sense of ordering. They are therefore measured on a nominal scale. For instance, hair color (Black, Brown, Gray, Red, Yellow) is a qualitative variable, as is name (Adam, Becky, Christina, Dave . . .). Qualitative variables can be coded to appear numeric but their numbers are meaningless, as in male=1, female=2. Variables that are not qualitative are known as quantitative variables.

Quantitative Variable
Variables that are measured on a numeric or quantitative scale. Ordinal, interval and ratio scales are quantitative. A country’s population, a person’s shoe size, or a car’s speed are all quantitative variables. Variables that are not quantitative are known as qualitative variables.

Random Assignment
Random assignment occurs when the subjects in an experiment are randomly assigned to conditions. Random assignment prevents systematic confounding of treatment effects with other variables.

Random Sampling
The process of selecting a subset of a population for the purposes of statistical inference. Random sampling means that every member of the population is equally likely to be chosen.

Range
The difference between the maximum and minimum values of a variable or distribution. The range is the simplest measure of variability.

Ratio Scale
One of the four basic levels of measurement, a ratio scale is a numerical scale with a true zero point and in which a given size interval has the same interpretation for the entire scale. Weight is a ratio scale, Therefore, it is meaningful to say that a 200 pound person weighs twice as much as a 100 pound person.

Regression
Regression means “prediction.” The regression of Y on X means the prediction of Y by X.

Regression Coefficient
A regression coefficient is the slope of the regression line in simple regression or the partial slope in multiple regression.

Regression Line
In linear regression, the line of best fit is called the regression line.

Relative Frequency
The proportion of observations falling into a given class. For example, if a bag of 55 M & M's has 11 green M&M's, then the frequency of green M&M's is 11 and the relative frequency is 11/55 = 0.20. Relative frequencies are often used in histograms, pie charts, and bar graphs.

Relative Frequency Distribution
A relative frequency distribution is just like a frequency distribution except that it consists of the proportions of occurrences instead of the numbers of occurrences for each value (or range of values) of a variable.

Representative Sample
A representative sample is a sample chosen to match the qualities of the population from which it is drawn. With a large sample size, random sampling will approximate a representative sample; stratified random sampling can be used to make a small sample more representative.

Robust
Something is robust if it holds up well in the face of adversity. A measure of central tendency or variability is considered robust if it is not greatly affected by a few extreme scores. A statistical test is considered robust if it works well in spite of moderate violations of the assumptions on which it is based.

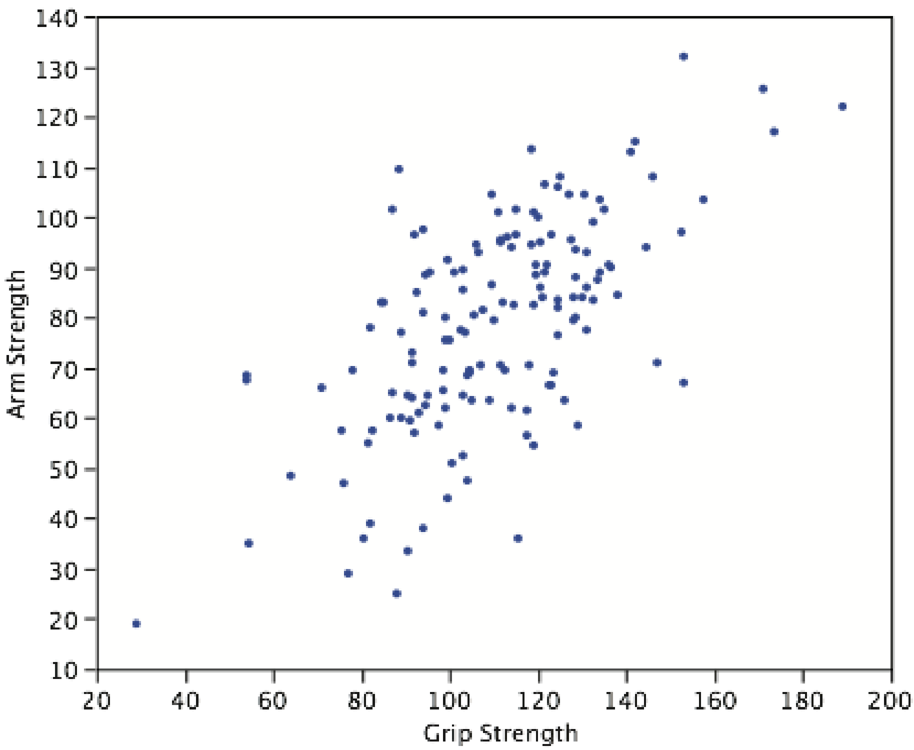
Sample
A sample is a subset of a population, often taken for the purpose of statistical inference. Generally, one uses a random sample.

Sampling Distribution

A sampling distribution can be thought of as a relative frequency distribution with a very large number of samples. More precisely, a relative frequency distribution approaches the sampling distribution as the number of samples approaches infinity. When a variable is discrete, the heights of the distribution are probabilities. When a variable is continuous, the class intervals have no width and and the heights of the distribution are probability densities.

Scatter Plot

A scatter plot of two variables shows the values of one variable on the Y axis and the values of the other variable on the X axis. Scatter plots are well suited for revealing the relationship between two variables. The scatter plot shown below illustrates the relationship between grip strength and arm strength in a sample of workers.



Significance Level

In significance testing, the significance level is the highest value of a probability value for which the null hypothesis is rejected. Common significance levels are 0.05 and 0.01. If the 0.05 level is used, then the null hypothesis is rejected if the probability value is less than or equal to 0.05.

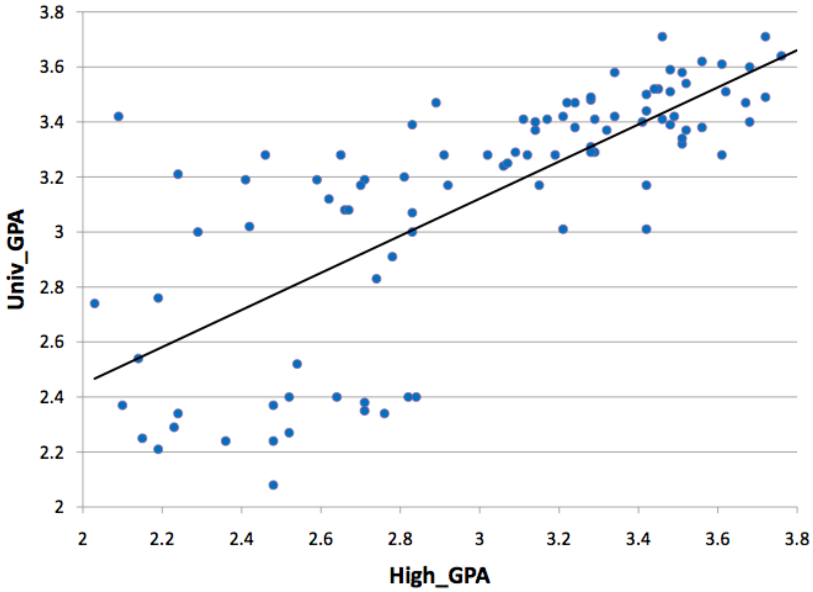
Significance Testing

A statistical procedure that tests the viability of the null hypothesis. If data (or more extreme data) are very unlikely given that the null hypothesis is true, then the null hypothesis is rejected. If the data or more extreme data are not unlikely, then the null hypothesis is not rejected. If the null hypothesis is rejected, then the result of the test is said to be significant. A statistically significant effect does not mean the effect is important.

Skew

A distribution is skewed if one tail extends out further than the other. A distribution has a positive skew (is skewed to the right) if the tail to the right is longer. It has a negative skew (skewed to the left) if the tail to the left is longer.

Slope
The slope of a line is the change in Y for each change of one unit of X. It is sometimes defined as "rise over run" which is the same thing. The slope of the black line in the graph is 0.675 because the line increases by 0.675 each time X increases by 1.0.



Standard Deviation
The standard deviation is a widely used measure of variability. It is computed by taking the square root of the variance. An important attribute of the standard deviation as a measure of variability is that if the mean and standard deviation of a normal distribution are known, it is possible to compute the percentile rank associated with any given score.

Standard Error
The standard error of a statistic is the standard deviation of the sampling distribution of that statistic. For example, the standard error of the mean is the standard deviation of the sampling distribution of the mean. Standard errors play a critical role in constructing confidence intervals and in significance testing.

Standard Error of the Mean

the standard error of the mean is the standard deviation of the sampling distribution of the mean. The formula for the standard error of the mean in a population is:

$$\sigma_m = \frac{\sigma}{\sqrt{N}}$$

where σ is the standard deviation and N is the sample size. When computed in a sample, the estimate of the standard error of the mean is:

$$s_M = \frac{s}{\sqrt{N}}$$

Standard Normal Distribution
The standard normal distribution is a normal distribution with a mean of 0 and a standard deviation of 1.

Step
One of the components of a box plot, the step is 1.5 times the difference between the upper hinge and the lower hinge. See also: H-spread.

Symmetric Distribution
In a symmetric distribution, the upper and lower halves of the distribution are mirror images of each other. In a symmetric distribution, the mean is equal to the median.

t distribution
The t distribution is the distribution of a value sampled from a normal distribution divided by an estimate of the distribution's standard deviation. In practice, the value is typically a statistic such as the mean or the difference between means and the standard deviation is an estimate of the standard error of the statistic. The t distribution in leptokurtic.

t test
Most commonly, a significance test of the difference between means based on the t distribution. Other applications include (a) testing the significance of the difference between a sample mean and a hypothesized value of the mean and (b) testing a specific contrast among means.

Third Variable Problem
A type of confounding in which a third variable leads to a mistaken causal relationship between two others. For instance, cities with a greater number of churches have a higher crime rate. However, more churches do not lead to more crime, but instead the third variable, population, leads to both more churches and more crime.

Two Tailed
The last step in significance testing involves calculating the probability that a statistic would differ as much or more from the parameter specified in the null hypothesis as does the statistics obtained in the experiment.

A probability computed considering differences in both direction (statistic either larger or smaller than the parameter) is called two-tailed probability. For example, if a parameter is 0 and the statistic is 12, a two-tailed probability would be the he probability of being either ≤ -12 or ≥12. Compare with the one-tailed probability which would be the probability of a statistic being ≥ to 12 if that were the direction specified in advance.

Type I Error
In significance testing, the error of rejecting a true null hypothesis.

Type II Error
In significance testing, the failure to reject a false null hypothesis.

Unbiased
A sample is said to be unbiased when every individual has an equal chance of being chosen from the population.

An estimator is unbiased if it does not systematically overestimate or underestimate the parameter it is estimating. In other words, it is unbiased if the mean of the sampling distribution of the statistic is the parameter it is estimating, The sample mean is an unbiased estimate of the population mean.

Upper Hinge
The upper hinge is one of the components of a box plot; it is the 75th percentile.

Upper Adjacent Value
One of the components of a box plot, the higher adjacent value is the largest value in the data below the upper inner fence.

Variability
Variability refers to the extent to which values differ from one another. That is, how much they vary. Variability can also be thought of as how spread out a distribution is. The standard deviation and the semi-interquartile range are measures of variability.

Variable
Something that can take on different values. For example, different subjects in an experiment weigh different amounts. Therefore “weight” is a variable in the experiment. Or, subjects may be given different doses of a drug. This would make “dosage” a variable. Variables can be dependent or independent, qualitative or quantitative, and continuous or discrete.

Variance
The variance is a widely used measure of variability. It is defined as the mean squared deviation of scores from the mean. The formula for variance computed in an entire population is:

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

where σ^2 represents the variance, μ is the mean, and N is the number of scores.

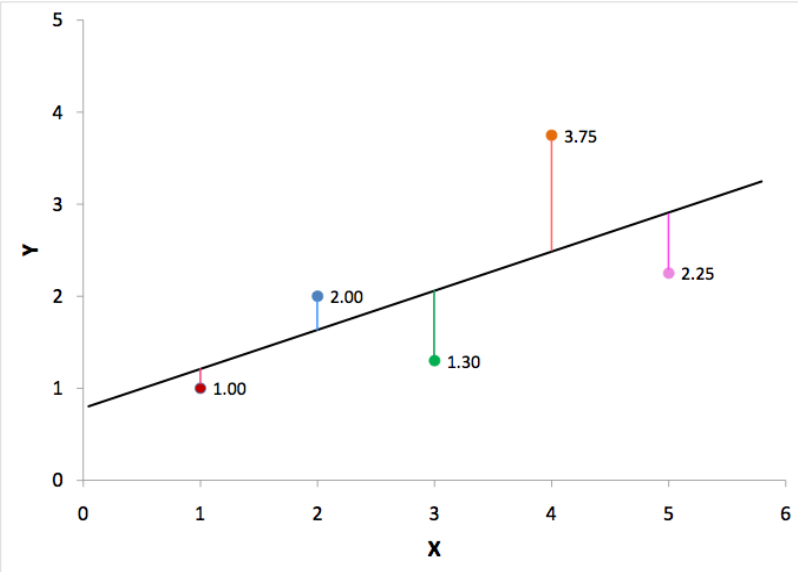
When computed in a sample in order to estimate the variance in the population, the formula is:

$$s^2 = \frac{\sum(X-M)^2}{N-1}$$

where s^2 is the estimate of variance, M is the sample mean, and N is the number of scores in the sample.

Y Intercept

The Y-intercept of a line is the value of Y at the point that the line intercepts the Y axis. It is the value of Y when X equals 0. The Y intercept of the black line shown in the graph is 0.785.



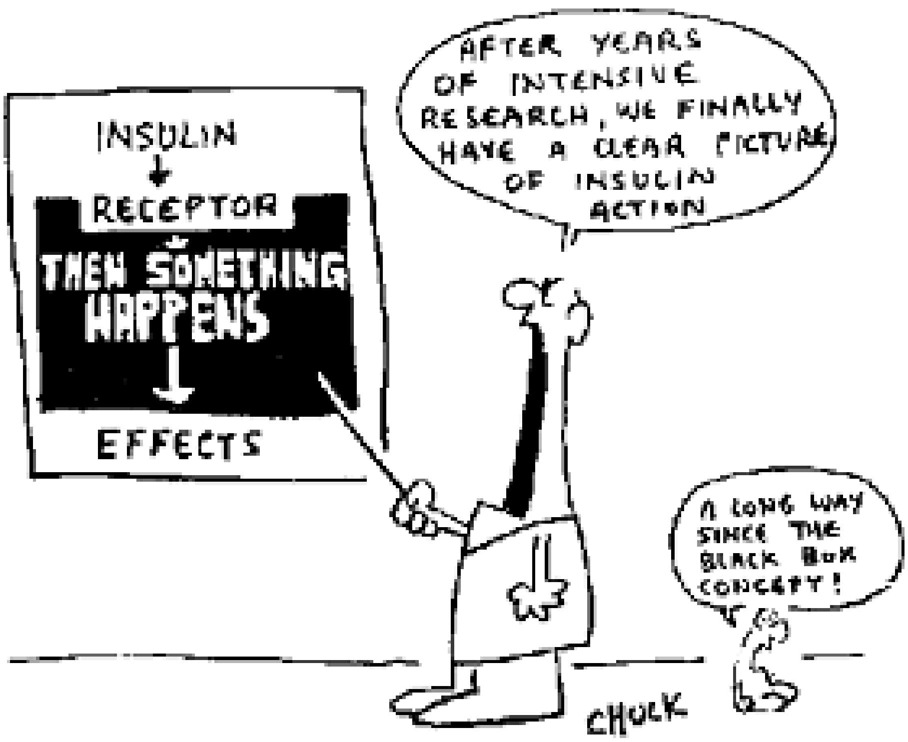
z score

The number of standard deviations a score is from the mean of its population. The term "standard score" is usually used for normal populations; the terms "z score" and "normal deviate" should only be used in reference to normal distributions. The transformation from a raw score X to a z score can be done using the following formula:

$$z = (X - \mu)/\sigma$$

Transforming a variable in this way is called "standardizing" the variable. It should be kept in mind that if X is not normally distributed then the transformed variable will not be normally distributed either.

Big Idea 2: Biological systems utilize free energy and molecular building blocks to grow, to reproduce and to maintain dynamic homeostasis.



Enduring understanding 2.A: Growth, reproduction and maintenance of the organization of living systems require free energy and matter.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 2.A.1: All living systems require constant input of free energy.	2.1 I can explain how biological systems use free energy based on empirical data that all organisms require constant energy input to maintain organization, to grow and to reproduce.		
	2.2 I can justify a scientific claim that free energy is required for living systems to maintain organization, to grow, or to reproduce, but that multiple strategies for obtaining and using energy exist in different living systems.		
	2.3 I can predict how changes in free energy availability affect organisms, populations, and/or ecosystems.		
Essential knowledge 2.A.2: Organisms capture and store free energy for use in biological processes.	2.4 I can use representations to pose scientific questions about what mechanisms and structural features allow organisms to capture, store and use free energy.		
	2.5 I can construct explanations of the mechanisms and structural features of cells that allow organisms to capture, store or use free energy.		
	2.41 I can evaluate data to show the relationship between photosynthesis and respiration in the flow of free energy through a system.		
Essential knowledge 2.A.3: Organisms must exchange matter with the environment to grow, reproduce and maintain organization.	2.6 I can use calculated surface area-to-volume ratios to predict which cell(s) might eliminate wastes or procure nutrients faster by diffusion.		
	2.7 I can explain how cell size and shape affect the overall rate of nutrient intake and the rate of waste elimination.		
	2.8 I can justify the selection of data regarding the types of molecules that an animal, plant or bacterium will take up as necessary building blocks and excrete as waste products.		
	2.9 I can represent graphically or model quantitatively the exchange of molecules between an organism and its environment, and the subsequent use of these molecules to build new molecules that facilitate dynamic homeostasis, growth and reproduction.		

Enduring understanding 2.B: Growth, reproduction and dynamic homeostasis require that cells create and maintain internal environments that are different from their external environments.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	**Learning Objective**	**Pre-comprehension level**	**Post-comprehension level**
Essential knowledge 2.B.1: Cell membranes are selectively permeable due to their structure.	2.10. I can use representations and models to pose scientific questions about the properties of cell membranes and selective permeability based on molecular structure.		
	2.11 I can construct models that connect the movement of molecules across membranes with membrane structure and function.		
Essential knowledge 2.B.2: Growth and dynamic homeostasis are maintained by the constant movement of molecules across membranes.	2.12 I can use representations and models to analyze situations or solve problems qualitatively and quantitatively to investigate whether dynamic homeostasis is maintained by the active movement of molecules across membranes.		
Essential knowledge 2.B.3: Eukaryotic cells maintain internal membranes that partition the cell into specialized regions.	2.13 I can explain how internal membranes and organelles contribute to cell functions.		
	2.14 I can use representations and models to describe differences in prokaryotic and eukaryotic cells.		

Enduring understanding 2.C: Organisms use feedback mechanisms to regulate growth and reproduction, and to maintain homeostasis.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 2.C.1: Organisms use feedback mechanisms to maintain their internal environments and respond to external environmental changes.	2.15 I can justify a claim made about the effect(s) on a biological system at the molecular, physiological or organismal level when given a scenario in which one or more components within a negative regulatory system is altered.		
	2.16 I can connect how organisms use negative feedback to maintain their internal environments.		
	2.17 I can evaluate data that show the effect(s) of changes in concentrations of key molecules on negative feedback mechanisms.		
	2.18 I can make predictions about how organisms use negative feedback mechanisms to maintain their internal environments.		
	2.19 I can make predictions about how positive feedback mechanisms amplify activities and processes in organisms based on scientific theories and models.		
	2.20 I can justify that positive feedback mechanisms amplify responses in organisms.		
Essential knowledge 2.C.2: Organisms respond to changes in their external environments.	2.21 I can justify the selection of the kind of data needed to answer scientific questions about the relevant mechanism that organisms use to respond to changes in their external environment.		
	2.42 I can pose a scientific question concerning the behavioral or physiological response of an organism to a change in its environment.		

Enduring understanding 2.D: Growth and dynamic homeostasis of a biological system are influenced by changes in the system's environment.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 2.D:1: All biological systems from cells and organisms to populations, communities and ecosystems are affected by complex biotic and abiotic interactions involving exchange of matter and free energy.	2.22 I can refine scientific models and questions about the effect of complex biotic and abiotic interactions on all biological systems, from cells and organisms to populations, communities and ecosystems.		
	2.23 I can design a plan for collecting data to show that all biological systems (cells, organisms, populations, communities and ecosystems) are affected by complex biotic and abiotic interactions.		
	2.24 I can analyze data to identify possible patterns and relationships between a biotic or abiotic factor and a biological system (cells, organisms, populations, communities or ecosystems).		
Essential knowledge 2.D:2: Homeostatic mechanisms reflect both common ancestry and divergence due to adaptation in different environments.	2.25 I can construct explanations based on scientific evidence that homeostatic mechanisms reflect continuity due to common ancestry and/or divergence due to adaptation in different environments.		
	2.26 I can analyze data to identify phylogenetic patterns or relationships, showing that homeostatic mechanisms reflect both continuity due to common ancestry and change due to evolution in different environments.		
	2.27 I can connect differences in the environment with the evolution of homeostatic mechanisms.		
Essential knowledge 2.D:3: Biological systems are affected by disruptions to their dynamic homeostasis.	2.28 I can use representations of models to analyze quantitatively and qualitatively the effects of disruptions to dynamic homeostasis in biological systems.		
Essential knowledge 2.D:4: Plants and animals have a variety of chemical defenses against infections that affect dynamic homeostasis.	2.29 I can create representations and models to describe immune responses.		
	2.30 I can create representations of models to describe nonspecific immune defenses in plants and animals.		
	2.43 I can connect the concept of cell communication to the functioning of the immune system.		

Enduring understanding 2.E: Many biological processes involved in growth, reproduction and dynamic homeostasis include temporal regulation and coordination.

Essential knowledge 2.E.1: Timing and coordination of specific events are necessary for the normal development of an organism, and these events are regulated by a variety of mechanisms.	2.31 I can connect concepts in and across domains to show that timing and coordination of specific events are necessary for normal development in an organism and that these events are regulated by multiple mechanisms.		
	2.32 I can use a graph or diagram to analyze situations or solve problems (quantitatively or qualitatively) that involve timing and coordination of events necessary for normal development in an organism.		
	2.33 I can justify scientific claims with scientific evidence to show that timing and coordination of several events are necessary for normal development in an organism and that these events are regulated by multiple mechanisms.		
	2.34 I can describe the role of programmed cell death in development and differentiation, the reuse of molecules, and the maintenance of dynamic homeostasis.		
Essential knowledge 2.E.2: Timing and coordination of physiological events are regulated by multiple mechanisms.	2.35 I can design a plan for collecting data to support the scientific claim that the timing and coordination of physiological events involve regulation.		
	2.36 I can justify scientific claims with evidence to show how timing and coordination of physiological events involve regulation.		
	2.37 I can connect concepts that describe mechanisms that regulate the timing and coordination of physiological events.		
Essential knowledge 2.E.3: Timing and coordination of behavior are regulated by various mechanisms and are important in natural selection.	2.38 I can analyze data to support the claim that responses to information and communication of information affect natural selection.		
	2.39 I can justify scientific claims, using evidence, to describe how timing and coordination of behavioral events in organisms are regulated by several mechanisms.		
	2.40 I can connect concepts in and across domain(s) to predict how environmental factors affect responses to information and change behavior.		

Big Idea 3: Living systems store, retrieve, transmit and respond to information essential to life processes.

Enduring understanding 3.A: Heritable information provides for continuity of life.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 3.A.1: DNA, and in some cases RNA, is the primary source of heritable information.	3.1 I can construct scientific explanations that use the structures and mechanisms of DNA and RNA to support the claim that DNA and, in some cases, RNA are the primary sources of heritable information.		
	3.2 I can justify the selection of data from historical investigations that support the claim that DNA is the source of heritable information.		
	3.3 I can describe representations and models that illustrate how genetic information is copied for transmission between generations.		
	3.4 I can describe representations and models illustrating how genetic information is translated into polypeptides.		
	3.5 I can explain how heritable information can be manipulated using common technologies.		
	3.6 I can predict how a change in a specific DNA or RNA sequence can result in changes in gene expression.		
Essential knowledge 3.A.2: In eukaryotes, heritable information is passed to the next generation via processes that include the cell cycle and mitosis or meiosis plus fertilization.	3.7 I can make predictions about natural phenomena occurring during the cell cycle.		
	3.8 I can describe the events that occur in the cell cycle.		
	3.9 I can explain, using visual representations or narratives, as to how DNA in chromosomes is transmitted to the next generation via mitosis, or meiosis followed by fertilization.		
	3.10 I can represent the connection between meiosis and increased genetic diversity necessary for evolution.		
	3.11 I can evaluate evidence provided by data sets to support the claim that heritable information is passed from one generation to another generation through mitosis, or meiosis followed by fertilization.		
Essential knowledge 3.A.3: The chromosomal basis of inheritance provides an understanding of the pattern of passage (transmission) of genes from parent to offspring.	3.12 I can construct a representation that connects the process of meiosis to the passage of traits from parent to offspring.		
	3.13 I can pose questions about ethical, social or medical issues surrounding human genetic disorders.		
	3.14 I can apply mathematical routines to determine Mendelian patterns of inheritance provided by data sets.		

Essential knowledge 3.A.4: The inheritance pattern of many traits cannot be explained by simple Mendelian genetics.	3.15 I can explain deviations from Mendel's model of the inheritance of traits.		
	3.16 I can explain how the inheritance patterns of many traits cannot be accounted for by Mendelian genetics.		
	3.17 I can describe representations of an appropriate example of inheritance patterns that cannot be explained by Mendel's model of the inheritance of traits.		

Enduring understanding 3.B: Expression of genetic information involves cellular and molecular mechanisms.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 3.B.1: Gene regulation results in differential gene expression, leading to cell specialization.	3.18 The student is able to describe the connection between the regulation of gene expression and observed differences between kinds of organisms.		
	3.19 The student is able to describe the connection between the regulation of gene expression and observed differences between individuals in a population.		
	3.20 The student is able to explain how the regulation of gene expression is essential for the processes and structures that support efficient cell function.		
	3.21 The student can use representations to describe how gene regulation influences cell products and function.		
Essential knowledge 3.B.2: A variety of intercellular and intracellular signal transmissions mediate gene expression.	3.22 The student is able to explain how signal pathways mediate gene expression, including how this process can affect protein production.		
	3.23 The student can use representations to describe mechanisms of the regulation of gene expression.		

Enduring understanding 3.C: The processing of genetic information is imperfect and is a source of genetic variation.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Essential Knowledge	Essential Knowledge
Essential knowledge 3.C.1: Changes in genotype can result in changes in phenotype.	3.24 The student is able to predict how a change in genotype, when expressed as a phenotype, provides a variation that can be subject to natural selection.		
	3.25 The student can create a visual representation to illustrate how changes in a DNA nucleotide sequence can result in a change in the polypeptide produced.		
	3.26 The student is able to explain the connection between genetic variation in organisms and phenotypic variation in populations.		
Essential knowledge 3.C.2: Biological systems have multiple processes that increase genetic variation.	3.27 The student is able to compare and contrast processes by which genetic variation is produced and maintained in organisms from multiple domains.		
	3.28 The student is able to construct an explanation of the multiple processes that increase variation within a population.		
Essential knowledge 3.C.3: Viral replication results in genetic variation and viral infection can introduce genetic variation into the hosts.	3.29 The student is able to construct an explanation of how viruses introduce genetic variation in host organisms.		
	3.30 The student is able to use representations and appropriate models to describe how viral replication introduces genetic variation in the viral population.		

Enduring understanding 3.D: Cells communicate by generating, transmitting and receiving chemical signals.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	Learning Objective	Pre-comprehension level	Post-comprehension level
Essential knowledge 3.D.1: Cell communication processes share common features that reflect a shared evolutionary history.	3.31 I can describe basic chemical processes for cell communication shared across evolutionary lines of descent.		
	3.32 I can generate scientific questions involving cell communication as it relates to the process of evolution.		
	3.33 I can use representation(s) and appropriate models to describe features of a cell signaling pathway.		
Essential knowledge 3.D.2: Cells communicate with each other through direct contact with other cells or from distance via chemical signaling.	3.34 I can construct explanations of cell communication through cell-to-cell contact or through chemical signaling.		
	3.35 I can create representation(s) that depict how cell-to-cell communication occurs by direct contact or from a distance through chemical signaling.		
Essential knowledge 3.D.3: Signal transduction pathways link signal reception with cellular response.	3.36 I can describe a model that expresses the key elements of signal transduction pathways by which a signal is converted to a cellular response.		
Essential knowledge 3.D.4: Changes in signal transduction pathways can alter cellular response.	3.37 I can justify claims based on scientific evidence that changes in signal transduction pathways can alter cellular response.		
	3.38 I can describe a model that expresses key elements to show how change in signal transduction can alter cellular response.		
	3.39 I can construct an explanation of how certain drugs affect signal reception and, consequently, signal transduction pathways.		

Enduring understanding 3.E: Transmission of information results in changes within and between biological systems.

For each of the **Learning Objectives**, indicate your comprehension level by filling in the appropriate code from the choices below.

Can – I can demonstrate I've learned the objective.
Think– I think I can demonstrate I've learned the objective.
Cannot – I cannot demonstrate I've learned the objective.

Essential Knowledge	**Learning Objective**	**Pre-comprehension level**	**Post-comprehension level**
Essential knowledge 3.E.1: Individuals can act on information and communicate it to others.	3.40 I can analyze data that indicate how organisms exchange information in response to internal changes and external cues, and which can change behavior.		
	3.41 I can create a representation that describes how organisms exchange information in response to internal changes and external cues, and which can result in changes in behavior.		
	3.42 I can describe how organisms exchange information in response to internal changes or environmental cues.		
Essential knowledge 3.E.2: Animals have nervous systems that detect external and internal signals, transmit and integrate information, and produce responses.	3.43 I can construct an explanation, based on scientific theories and models, about how nervous systems detect external and internal signals, transmit and integrate information, and produce responses.		
	3.44 I can describe how nervous systems detect external and internal signals.		
	3.45 I can describe how nervous systems transmit information.		
	3.46 I can describe how the vertebrate brain integrates information to produce a response.		
	3.47 I can create a visual representation of complex nervous systems to describe/explain how these systems detect external and internal signals, transmit and integrate information, and produce responses.		
	3.48 I can create a visual representation to describe how nervous systems detect external and internal signals.		
	3.49 I can create a visual representation to describe how nervous systems transmit information.		
	3.50 The student is able to create a visual representation to describe how the vertebrate brain integrates information to produce a response.		

Labs

Virtual Labs