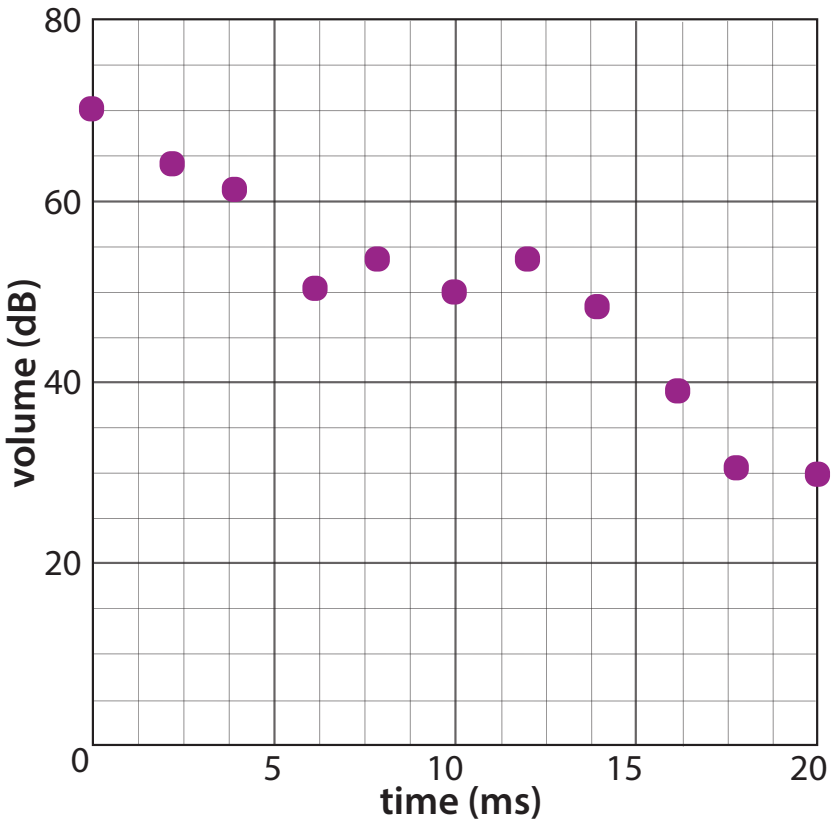
Unit 6: Statistics

Activity 6.1 - Reverberation Time

Recall that sound waves spread out and bounce around. So sometimes we can still hear a sound a bit after it is no longer being made. If we hear it clearly after a short pause, then it is usually called an *echo*. If we hear it fade away–often without a pause–then it is usually a *reverberation*. As the sound waves spread, it weakens and becomes quieter. The loudness of the sound is called volume and is measured in decibels (dB). This definition of volume should not be confused with the volume that describes the amount of space in an object, which is measured in liters or cubic meters.

Suppose that you're setting up a large room for a Hawaiian music concert. To test if it is set up the way you want, you hit kālaʻau sticks together once (at time 0 ms) and use a machine to record the volume of the sound. You can watch a video about kālaʻau sticks on the book's website (`https://www.stemd2.com`).

1. Here is a graph of the volume in the recording from the first setup.



a. Does this scatter plot show a positive association, a negative association, or no association?

The sound volume typically decreases as time increases. So there is a negative association.

b. Describe the real-world reason why your answer to Part 1a makes sense.

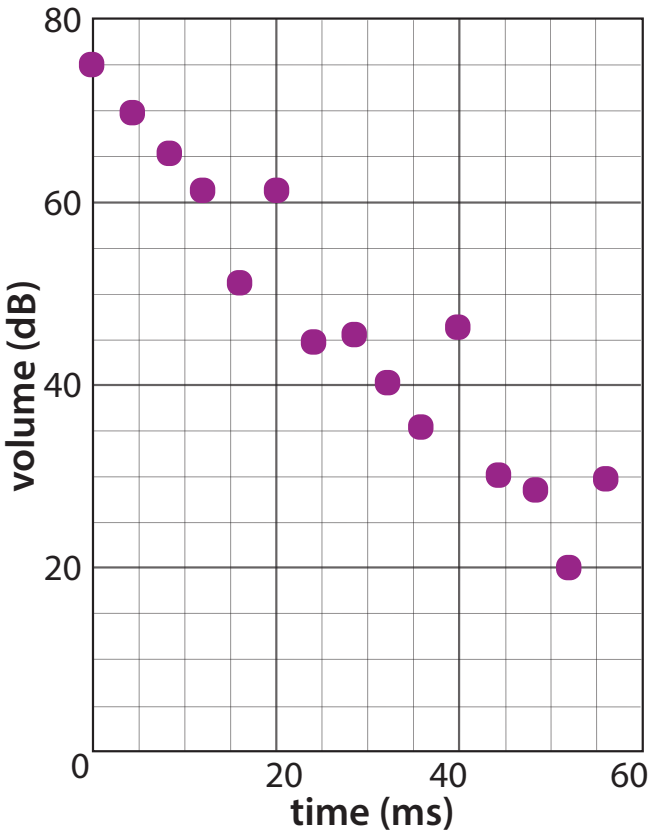
The sound waves spread out and weaken as the air and environment absorbs the energy. The sounds that you hear later have traveled and bounced around more than the sounds that you hear right away so they're weaker.

c. Assume that there is a linear relationship shown in the graph. What is the equation for this line? Let y be loudness (volume) and t be time.

We can choose two points that look like they follow the general trend then write the line that goes through these points. The first and last point happens to look like good choices. They're at (0, 70) and (20, 30). So the equation is $y = -2x + 70$.

$$\begin{aligned} y &= mx + b \\ &= \frac{(30-70)}{(20-0)}x + (70) \\ &= -\frac{40}{20}x + 70 \\ &= -2x + 70 \end{aligned}$$

2. Here is the recording from the second setup.



a. Does this scatter plot show a positive association, a negative association, or no association?

This also shows a negative association for the same reasons as before.

b. Describe any clusters that you see on the plot.

All but three or four points seem to fit nicely on a line. We can consider these points to be clusters. However, since the points are spread out pretty evenly, some people might not consider them to be clusters.

c. Describe some of the outliers on the plot.

The outliers are around (20, 62), (40, 47), and (57, 30). They are fairly evenly spaced out in time and a bit louder than the rest of the sounds.

d. Take a look at the outliers you found in part 2c. Why might there be outliers? Are these readings higher or lower than expected? Why might they be going higher/lower?

These outliers might be echoes. The readings are higher than expected meaning that some of the sound is bouncing around the environment strongly while most of the other sounds fade away.

e. Assume that there is a linear relationship shown in the graph. What is the equation for this line? Let y be loudness and t be time.

The points at (0, 75) and (48, 28) seem to follow the general trend of the data. We can choose to put a line through these or other similar points. This gives us the equation $y = -x + 75$.

$$\begin{aligned} y &= mx + b \\ &= \frac{(28-75)}{(48-0)}x + (75) \\ &= -\frac{47}{48}x + 75 \\ &\approx -(1)x + 75 \end{aligned}$$

3. *Reverberation time* is the time it takes for a sound's volume to drop 60 dB. For example, if a sound starts at 100 dB, the reverberation time is how long it takes for the sound to dampen to 40 dB.

a. What is the reverberation time of the first setup?

The equation for the first setup is $y = -2x + 70$. The y intercept tells us that at $x = 0$, $y = 70$. To find the reverberation time, we need to figure out when the volume y drops 60 dB to $y = 10$. The reverberation time is 30 milliseconds.

$$\begin{aligned} y &= -2x + 70 \\ 70 - 60 &= -2x + 70 \\ 10 &= -2x + 70 \\ -70 &\quad -70 \\ -60 &= -2x \\ \div(-2) &\quad \div(-2) \\ 30 &= x \end{aligned}$$

b. What is the reverberation time of the second setup?

The reverberation time is 60 milliseconds.

$$\begin{aligned} y &= -x + 75 \\ 75 - 60 &= -x + 75 \\ 15 &= -x + 75 \\ -75 &\quad -75 \\ -60 &= -x \\ \div(-1) &\quad \div(-1) \\ 60 &= x \end{aligned}$$

c. One of these setups results in more echoes. Which one do you think it is and why did you think it was that one?

The second setup has more echoes. This is shown in the presence of outliers. The reverberation time is also twice as long. This tells us that the setup is not very good at stopping or scattering sound waves.

d. Which of these setups would you choose for a concert and why? Discuss and share your ideas with other groups and on the on the online comment section.

Most students would agree that the first setup is better for listening to music because there is less reverberation and echo. The students are free to disagree though, as long as they articulate their reasoning.

Activity 6.2 - Music Preferences I

Now that we have the room set up for the concert, let's choose the music. We can do this by asking other students what kind of music they want to hear. Here's the data we've collected by asking 210 students about their music preferences.

Music preference

Age	Hawaiian music	Jawaiian music	TOTAL
10-12	52	78	130
13-15	52	28	80
TOTAL	104	106	210

1. Let's take a look at each age group separately.
Next to letters A-G, write the relative frequency compared to the total of that **row**. Here, two boxes in the first row are filled as examples. Give your answers as a percentage and round to the nearest whole number. You may use a calculator for this part.

Music preference

Age	Hawaiian music		Jawaiian music		TOTAL	
10-12	52	40%	78	A	130	100%
13-15	52	B	28	C	80	D
TOTAL	104	E	106	F	210	G

Music preference

Age	Hawaiian music		Jawaiian music		TOTAL	
10-12	52	40%	78	A **60%**	130	100%
13-15	52	B **65%**	28	C **35%**	80	D **100%**
TOTAL	104	E **50%**	106	F **50%**	210	G **100%**

2. Let's take a look at each music genre separately.
Next to letters A-G, write the relative frequency compared to the total of that **column**. Here, two boxes in the Hawaiian music column are filled as an example. Give your answers as a percentage and round to the nearest whole number. You may use a calculator for this part.

Music preference

Age	Hawaiian music		Jawaiian music		TOTAL	
10-12	52	50	78	A	130	B
13-15	52	C	28	D	80	E
TOTAL	104	100%	106	F	210	G

Music preference

Age	Hawaiian music		Jawaiian music		TOTAL	
10-12	52	50	78	A **74%**	130	B **62%**
13-15	52	C **50%**	28	D **26%**	80	E **38%**
TOTAL	104	100%	106	F **100%**	210	G **100%**

3.a. Explain the meaning of the percentages you wrote in Part 1. What do they tell you?

This table of percentages look at the ages separately. For a certain age, it asks: *what percentage of students in this age group prefer Hawaiian or Jawaiian music?*

b. Explain the meaning of the percentages you wrote in Part 2. What do they tell you?

This table of percentages look at the music genres separately. For a certain music type, it asks: *what percentage of students who love this type of music are ages 10-12 or ages 13-15?*

4. Of the types of songs that you're thinking of playing for the concert, what percent of them should be Hawaiian songs? What percent of them should be Hip Hop? How did you use the data from the table to help you make your decisions?

These answers are very open-ended. For example, students can look at the first table and argue that most students aged 10-12 prefer Jawaiian music, but most students aged 13-15 prefer Hawaiian music. So we should play an equal amount of each type of music. (50% Hawaiian, 50% Jawaiian). Students can also look at the second table and argue that students who like Hawaiian music come from both age groups, but the students who like Jawaiian music are mostly aged 10-12. So it might be more enjoyable to play mostly Hawaiian music since both age groups enjoy it. (70% Hawaiian, 30% Jawaiian).

Activity 6.3 - Music Preferences II

Your kumu noticed that not everyone voted. Maybe some students wanted to hear other kinds of music. To prepare for the next concert, we let students vote for "other music." Here are the results.

Music Preference

Age	Hawaiian	Jawaiian	Other	TOTAL
Age 10-12	55	104	21	180
Age 13-15	57	15	8	80
TOTAL	112	119	29	260

1. Let's take a look at all the votes.
Next to the letters A-I, write the relative frequency compared to the total number of votes (260). Give your answers as a percentage and round to the nearest whole number. Three boxes are already filled. For example, we can see that 21% of all votes were from 10-12 year old students who wanted to hear Hawaiian music. You may use a calculator for this part.

Music Preference

Age	Hawaiian		Jawaiian		Other		TOTAL	
Age 10-12	55	21%	104	A	21	B	180	C
Age 13-15	57	D	15	E	8	F	80	31%
TOTAL	112	G	119	H	29	I	260	100%

Music Preference

Age	Hawaiian		Jawaiian		Other		TOTAL	
Age 10-12	55	21%	104	A **40%**	21	B **8%**	180	C **69%**
Age 13-15	57	D **22%**	15	E **6%**	8	F **3%**	80	31%
TOTAL	112	G **43%**	119	H **46%**	29	I **11%**	260	100%

2. What are some conclusions or statements that you can make about the data? For example, can you say that certain groups have a stronger preference for certain kinds of music? If you cannot make any conclusions, why not?

There are many different answers. We can make precise conclusions like *46% of the students liked Jawaiian music* and *22% of the students surveyed like Hawaiian music and are aged 13-15*. We cannot make big general conclusions like *students like Jawaiian music more than Hawaiian music*. We also have to be careful not to assume too much. For example, just because we can say that 46% of students like Jawaiian music, we cannot say that 46% of students hate Hawaiian music. The might love both types of music, but prefer to listen to Jawaiian. We only ask about what kind of music do they prefer; we do not ask about what they dislike.

3. Share your ideas with your classmates and online. Did you see any conclusions from other groups that you disagree with? How is it possible for two people to look at the same data and come up with very different conclusions? This actually happens a lot. Talk with your other classmates about how or why this is happens.

One of the reasons that we disagree about the interpretation of the data is because we did not clearly describe where the data comes from. This is why it is very important to write a clear Methods section in a lab/science report and cite all the sources of your data. It is not a big issue here, but another major source of confusion happens when we don't choose our independent variables very well. In one example from a textbook, students are separated by whether they preferred math or science. What about students who prefer both or neither? If, for example, 10% of students prefer science, can we say that 90% do not prefer science?