

# Using Linear Regression to Explore Environmental Factors Affecting Vector-borne Diseases

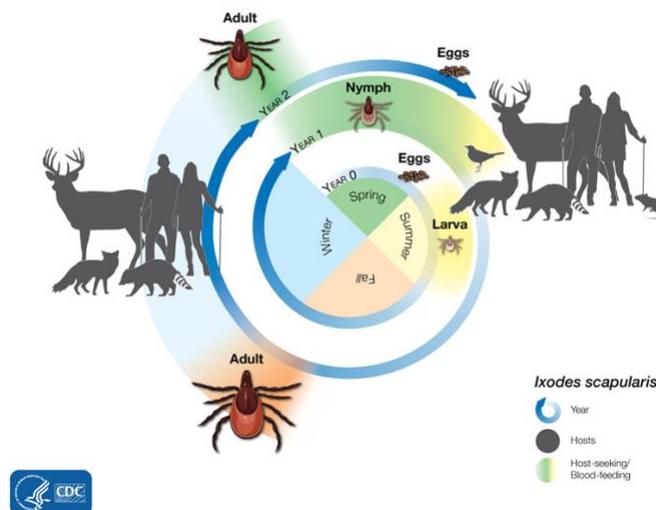
## OPENING ACTIVITY

### Background

Lyme disease is a zoonotic disease caused by bacteria from the genus *Borrelia*. Within the first month of infection, a person may develop initial symptoms of Lyme disease such as fatigue, headache, fever, joint pain, and a bullseye skin rash. Without treatment, Lyme disease may progress with more serious consequences for the cardiovascular, skeletal, and nervous systems (CDC 2019)

*Borrelia* bacteria are commonly found in natural hosts such as small rodents and birds. Ticks can acquire *Borrelia* from these host animals and subsequently transmit the bacteria to humans through a tick bite. The ticks are referred to as the “vector” for Lyme disease since the tick transmits the bacteria from the reservoir host to humans. To understand the epidemiology and spread of Lyme disease, it is important to consider the habitat and range of the vector (CDC 2019).

Ticks belonging to the genus *Ixodes* are the most common vectors of Lyme disease and are found across North America, Europe, and Northern Asia. Ticks thrive in humid environments where they can use vegetation as protection from the sun. *Ixodes* ticks have four life stages: eggs, larvae, nymphs, and adults. After the eggs hatch to larvae, the ticks require a feeding to progress to each subsequent life stage. Ticks spend over 90% of their lives in a fasting state off-host. It is during these off-host times where the ticks are most vulnerable to climate and environmental factors such as temperature and humidity. Therefore, environmental scientists and epidemiologists often study the impact of environmental conditions on Lyme disease transmission (Drummond 2004).



## Introduction Questions

1. Describe the mode of transmission for Lyme disease.
2. Describe the habitat and optimal environmental conditions for the vector that carries Lyme disease.
3. Describe environmental conditions that could contribute to transmission of Lyme disease to humans. Explain your reasoning.
4. Before viewing the data below, come up with a hypothesis regarding the relationship between Lyme disease occurrence and temperature. Explain your reasoning.

## Graphing and Data Analysis

Table 1. Historical data on temperature and Lyme disease cases in Baltimore, MD.

<https://www.cdc.gov/lyme/stats/maps.html>

<b>Month</b>	<b>Average High Temp (F) Baltimore, MD</b>	<b>Lyme Disease Cases</b>
January	41.2	3
February	44.8	2
March	53.9	2
April	64.5	4
May	73.9	5
June	82.7	15
July	87.2	22
August	85.1	13
September	78.2	6

October	67.0	5
November	56.3	4
December	46.0	1

1. You will create a scatter plot for the data above. Before doing so, answer the following questions.
  - a. What is the dependent variable in this data set?
  - b. What is the independent variable?
  - c. What variable will you use on the X-axis?
  - d. What variable will you use on the y-axis?
2. After discussing your responses for the question above, make a scatterplot of the data. Describe the relationship between the data.
3. Calculate the correlation coefficient.
4. What does the value of the correlation coefficient suggest about the relationship between these two variables? Discuss the strength and direction of the correlation.
5. Calculate the line of best fit.
6. Interpret the slope from the model.
7. Is it possible to interpret the y intercept from the model? Why or why not?
8. How many Lyme disease cases should there be if the temperature is 91.2? Is this an example of extrapolation or interpolation?
9. What should the temperature be if there are 10 cases of Lyme disease?
10. What is the residual of the cases of Lyme disease in August?

## Conclusion Questions

1. Does the data appear to support your hypothesis? Explain your reasoning
2. Are there any other environmental factors that could influence Lyme disease occurrence? List them and describe how.
3. How could climate change influence the habitat or location of the ticks (and, thus, the occurrence of Lyme disease)?
4. There are many other human health issues that could be exacerbated by climate change. Provide an example.

## Optional Additional Activity

Use this activity to discuss why correlation does not imply causation.

Ice Cream Sales (Per \$1000)	Swimming Pool Attendance (Per 1000 People)
2.3	0.24
3.4	0.36
3.49	0.36
3.58	0.49
6.9	1.32
9.2	5.87
9.1	6.73
9.3	6.25
9.05	5.83
6.45	1.28
3.69	0.42
2.98	0.12

1. Create a scatter plot to display the data.
2. Based on the scatter plot, what is the relationship between the 2 variables? Weak/Strong? Positive/Negative?
3. Are there any other variables that would account for this relationship?
4. Is this relationship association or causation? Explain.

## Resources:

Brownstein, J.S., Holford, T.R. & Fish, D. 2005. Effect of Climate Change on Lyme Disease Risk in North America. *EcoHealth* 2: 38–46. <https://doi.org/10.1007/s10393-004-0139-x>

Center for Disease Control. Lyme Disease: What you need to know. (15 August 2019). Retrieved July 6, 2020, from <https://www.cdc.gov/lyme/resources/brochure/lymediseasebrochure-P.pdf>

Drummond, R. 2004. Ticks and What You Can Do About Them. Wilderness Press.  
Co-Authors (Alphabetical): Andy Adams, Jessica Adams, John Bray, Tami Imbierowicz, Suzanne Lenhart, Breonna Martin. All co-authors contributed equally. This material is based upon work supported by the



National Science Foundation under Grant No. 1919613. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.