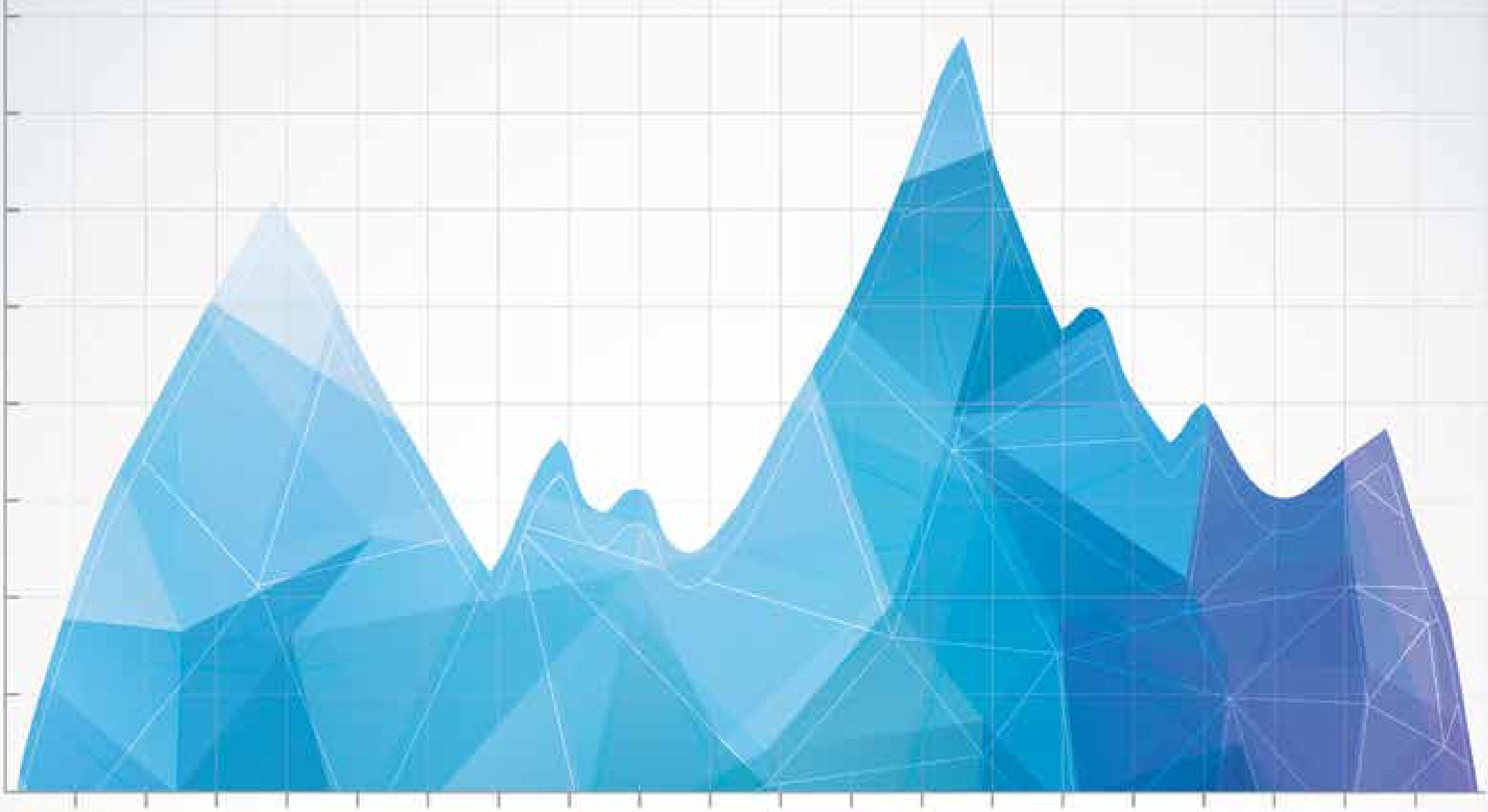


DEMYSTIFYING DATA

Using Google Trends and climate-change data to understand graphs and patterns

Carolyn Dash and Barbara Hug

We constantly encounter data—in the form of graphs—that convey information about weather, medicine, politics, finances, and nutrition. These graphs are intended to help us visualize data for easy interpretation; however, approximately 41% of adults in the United States have low graph literacy (Galesic and Garcia-Retamero 2011).

In this article, we describe an activity that we created to help students

- integrate and evaluate multiple graphs to answer questions,
- choose and interpret units in graphs, and
- analyze and interpret data.

In this activity, students use Google Trends (see "On the web") and climate-change data to understand patterns in basic time-series graphs. The activity aligns with several areas of the *Next Generation Science Standards* (NGSS Lead States 2013) and *Common Core State Standards* (NGAC and CCSSO 2010) (Figure 1, p. 52).

Warm-up discussion and demonstration

Teachers can begin this lesson by demonstrating how Google Trends works. Google Trends, a free tool, allows users to visualize the frequency of a particular Google search term. Teachers can start by plotting search terms *with* strong patterns through time (e.g., each school year drives searches for "The Scarlet Letter") and *without* strong patterns through time (e.g., movie and book releases have historically driven "Harry Potter" searches). Teachers can also show students how to add data to pre-existing graphs to visualize how often multiple search terms relate to each other. For example, "The Scarlet Letter" is positively correlated with "Nathaniel Hawthorne," whereas "day camp" is anticorrelated with both terms—suggesting that few summer campers are researching 19th-century literature.

This demonstration enables students to grasp data patterns and inevitably results in students suggesting search terms. Before searching for their terms, ask students to predict what the resulting graph will look like and hypothesize

FIGURE 1

Connections to the standards.

Next Generation Science Standards (NGSS Lead States 2013)

HS-ESS3: Earth and Human Activity

Performance Expectation

- HS-ESS3-5: Analyze geoscience data and the results from global climate models to make an evidence-based forecast of the current rate of global or regional climate change and associated future impacts to Earth systems.

Disciplinary Core Idea

- ESS3.D: Global Climate Change: Though the magnitudes of human impacts are greater than they have ever been, so too are human abilities to model, predict, and manage current and future impacts.

Science and Engineering Practice

- Analyzing and Interpreting Data: Analyze data using computational models in order to make valid and reliable scientific claims.

Crosscutting Concepts

- Stability and Change: Change and rates of change can be quantified and modeled over very short or very long periods of time. Some system changes are irreversible.
- Cause and Effect: Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.
- Patterns: Empirical evidence is needed to identify patterns.

Connections to the Nature of Science

- Science knowledge is based on empirical evidence.

Common Core State Standards (NGAC and CCSSO 2010)

ELA/Literacy

- RST.11-12.7: Integrate and evaluate multiple sources of information presented in diverse formats and media (e.g., quantitative data, video, multimedia) in order to address a question or solve a problem.

Mathematics

- HSN-Q.A.1 Use units as a way to understand problems and to guide the solution of multi-step problems; choose and interpret units consistently in formulas; choose and interpret the scale and the origin in graphs and data displays.

a potential mechanism. For example, students may suggest that searches for “skeleton” will peak at the end of each October (Halloween) and every fourth January (coinciding with the Winter Olympics, with its skeleton sledding event). To engage advanced learners, ask them what search terms might correlate with their classmates’ suggestions and why. Explain the distinction between correlation and causation: Without a mechanistic explanation, students can’t infer causation between terms because one term doesn’t necessarily cause the trending of another term.

If time allows, teachers can expand upon this basic classroom activity to help students discover how to generate real data and graphs and understand how the data connect to their lives. For example, one surveyed high school teacher commented: “My students loved this activity. I turned it into a competition of who could make the most interesting graph on their own time. Students voted on the graphs.”

Extensions such as this allow students to explore data directly. Students can use these searches to make creative connections to their own interests. Or they might make connections to previous units or other classes by using only science search terms. For example, students could examine the graph for “Charles Darwin” and try to explain why peaks seem to occur annually each semester. Students can also compare graphs for “global warming” and “greenhouse effect” or “coral reefs” and “acidification” and explain possible correlations. The Google Trends database lets students make predictions and test their own hypotheses. It helps them discover how to generate real data and graphs and connect that data to their lives.

Teachers can also use this activity to meaningfully introduce appropriate vocabulary (Figure 2). By developing a connection between the vocabulary and the data patterns, students begin to understand these scientific words. Our colleagues have used this lesson at levels ranging from introductory high school biology to graduate-level biology seminars; the only major difference between the high school– and graduate-level learners is the vocabulary and the depth of the discussion.

Part I: Identifying patterns in Google Trends data

In this part, students work together or individually to read and interpret four Google Trend–generated graphs (Figure 3, p. 54). Students predict the search terms used to create each pattern and write rationales for each prediction. The purpose isn't for students to correctly identify the search terms but to create evidence-based inferences and support these inferences with data. When there are two terms on one graph, teachers can encourage students to hypothesize reasons why the terms would or wouldn't be correlated.

Teachers can project the graphs or provide a worksheet and have students work alone or in groups to answer questions about each graph. A sample worksheet, answer key, lesson plan, and slides are available online (see "On the web"). The worksheet is particularly useful as a means of keeping the class focused and as an assessment. Students use their graph-reading skills and abilities to understand patterns over time to interpret scientific data in Part II of this activity.

Part II: Interpreting trends in climate-change data

Part II asks students to make evidence-based inferences from two iconic data sets often used as evidence of climate change: the Keeling Curve (Figure 4, p. 55) and the Vostok Lake Ice-Core Record (Figure 5, p. 56). Students first explore the Keeling Curve, named for Charles David Keeling who, in 1958, began measuring atmospheric carbon dioxide (CO_2) (in parts per million [ppm]) at the Mauna Loa Observatory in Hawaii (Figure 4). This

FIGURE 2

Useful terminology for describing time-series data.

Basic terms

Oscillation: A repetitive variation around a mean value that can be described using the following terms:

- *Amplitude:* The magnitude or size of an oscillation, typically measured as the difference between the values of the peak and trough for each cycle (e.g., amplitude increases after 2010 in Figure 3B, p. 54).
- *Baseline:* The average value of each cycle through time. For many oscillations the baseline is stable (e.g., Figure 3C); however, the baseline can also shift through time (e.g., Figure 3D).
- *Frequency:* Inverse period or the number of cycles in a given amount of time (e.g., annually in Figure 3A).
- *Period:* Duration of one cycle of the oscillation or the time between successive peaks or troughs (e.g., one year in Figure 3A).
- *Phase:* Describes where the oscillation is at each point in a cycle (e.g., the blue line is highest in summer and lowest in winter in Figure 3C).

Comparative terms

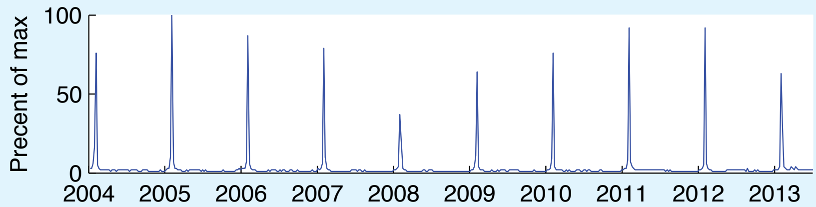
- *Correlation:* When the change in one signal is associated with similar changes in another signal (e.g., Figure 3D, p. 54); antonyms include anticorrelation, inverse correlation, negative correlation (e.g., Figure 3C).
- *Offset:* When two or more oscillations have different baselines (e.g., Figure 3D).
- *Phase shift:* When two or more oscillations have the same frequency but their peaks and troughs don't align in time; synonyms include antiphase and out-of-phase (e.g., Figure 3C).

FIGURE 3

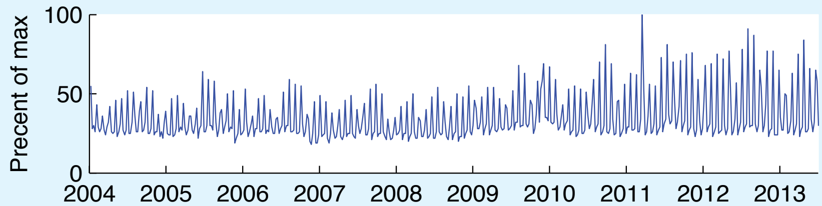
Google Trend outputs.

The below graphs represent the output of Google Trend searches conducted from 2004 to 2013. The search terms in each figure are (A) *groundhog*, (B) *full moon*, (C) *swimming* (blue) and *skiing* (red), and D) *Romeo* (blue) and *Juliet* (red).

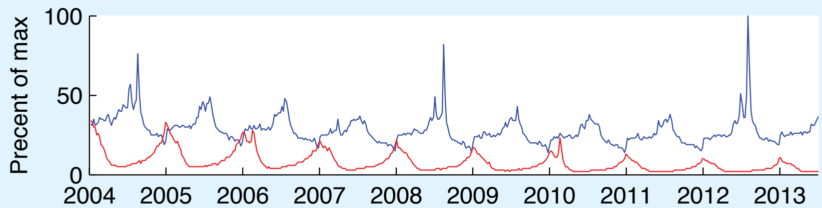
A *Groundhog:* This figure displays a very clear annual period. Support for the choice of *groundhog* could be that the search term frequency peaks once a year in late winter. Good alternative predictions include, but are not limited to: *MLK Day*, *Valentine*, *Super Bowl*, *Oscars*.



B *Full moon:* This example is similar to Figure 1A, but the period lasts approximately one month. Good answers are hard for this pattern because of its unusual frequency, but astute students may answer with references to lunar cycles.



C *Swimming* and *skiing:* Students should notice that the trends are phase shifted or anticorrelated; when the one increases in summer, the other decreases and vice versa. Other good answers would be *ice cream* and *hot cocoa*, *sandcastle* and *snowman*, etc., because these are all driven by searches most likely conducted in summer and winter.



D *Romeo* and *Juliet:* These search trends both have the same frequency and phase, are correlated with each other, and have a baseline shift. Strong answers are duos such as *Bonnie* and *Clyde* or *Bert* and *Ernie*, but students may also suggest similar words such as *army* and *navy*. Perceptive students will notice that both search terms are highest during the school year and thus might make more academic suggestions such as *Lewis* and *Clark*.

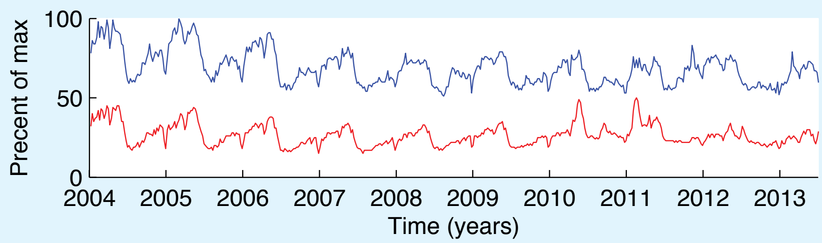


figure is still continuously updated (e.g., in mid-August 2014 the CO_2 concentration was 397 ppm). The full record and current carbon dioxide reading is available online (see "On the web").

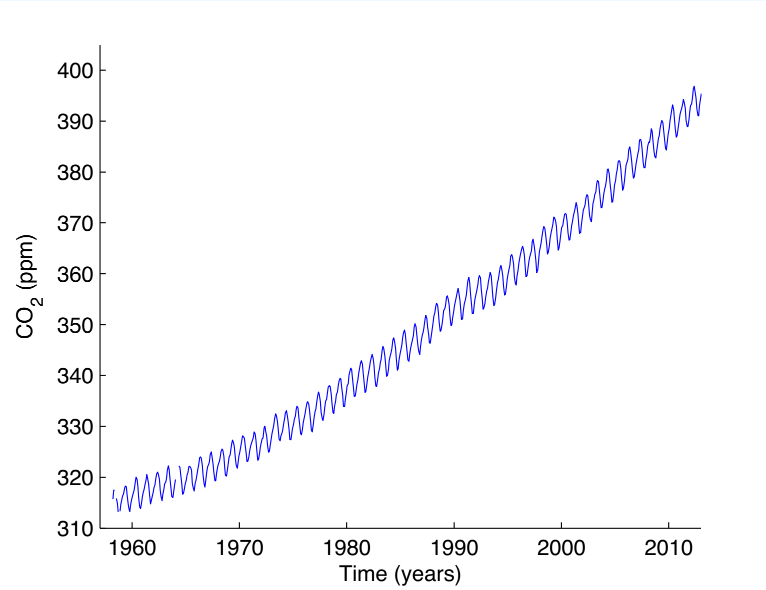
Using the vocabulary and skills gained during Part I, students read and discuss the graphs, then make informed observations of general trends and patterning (Figure 4). Two trends should be immediately apparent: Baseline atmospheric carbon dioxide increases through time, and there is an annual period of an increase and decrease of about 6 ppm. The baseline shift relates to increased burning of fossil fuels, and the annual increase and decrease relate to the cycle of leaf growth and loss each year in the Northern Hemisphere.

Students should consider if the increasing trend is part of a cycle, and teachers can ask how students' answers would change if they only saw a single year of data. This helps students appreciate how the figures' scale influences visualization and interpretation of data. Students often suggest that the increase in carbon dioxide might be part of a cycle but that more data is necessary to test this claim.

FIGURE 4

The Keeling Curve.

The graph below shows changes in atmospheric carbon dioxide (CO_2) concentration (in parts per million [ppm]) over time (Keeling et al. 2001).



That leads to the next data set. Figure 5 (p. 56) is a 420,000-year ice-core record of reconstructed carbon dioxide (ppm) and temperature (°C) from Vostok Lake in Antarctica (Petit et al. 1999). First, students orient themselves to the figures and describe the basic patterns: the approximately 100,000-year phase in both carbon dioxide and temperature and the close correlation between carbon dioxide and temperature over the length of the record. Students should conclude that, given this data set, carbon dioxide and temperature appear correlated and cycle together through time.

At this point, students typically want as much data as possible to draw the most informed conclusion. Some ask how to extend the data to the present day. They can do this by plotting today's concentration of atmospheric carbon dioxide (ppm) from the Keeling Curve (Figure 4) onto the figure of atmospheric carbon dioxide (ppm) from Vostok Lake (Figure 5).

Afterward, they can draw a line to connect the plot to the rest of the carbon dioxide graph. Then they see that the concentration of atmospheric carbon dioxide is far greater today than at any time in the past 420,000 years. This new data point is far outside carbon dioxide's natural cycle. Teachers can then ask students what effect the increase in atmospheric carbon dioxide will likely have on temperature.

Students generally suggest that the unprecedentedly high carbon dioxide correlates with an increase in temperature and that the recent climate change is not part of the 420,000-year cycle. Depending on their prior knowledge, students may recognize the greenhouse effect, or the absorption of infrared radiation by greenhouse gases like carbon dioxide, causing rising temperatures.

To engage more advanced students, ask what conclusions they might deduce about carbon dioxide and temperature if they only had the portion of the graph from about 20,000 years ago to the present. Over that timescale, although students would see a correlation between carbon dioxide and temperature, both would appear to be severely increasing for 20,000 years instead of over the past approximately 150 years. This scale can help students differentiate between the influence of natural glacial or interglacial cycles and the recent anthropogenic alteration of atmospheric carbon dioxide. Follow the activity with a discussion on the importance of scale and evidence-based inferences, correlation versus causation, and other evidence and effects of climate change.

Conclusions

We originally created this activity for a high school teacher workshop. Most participants used the activity the following year and reported that it not only enables students to better grasp basic graphing concepts but also allows them to apply these skills and vocabulary to interpreting real geoscience data.

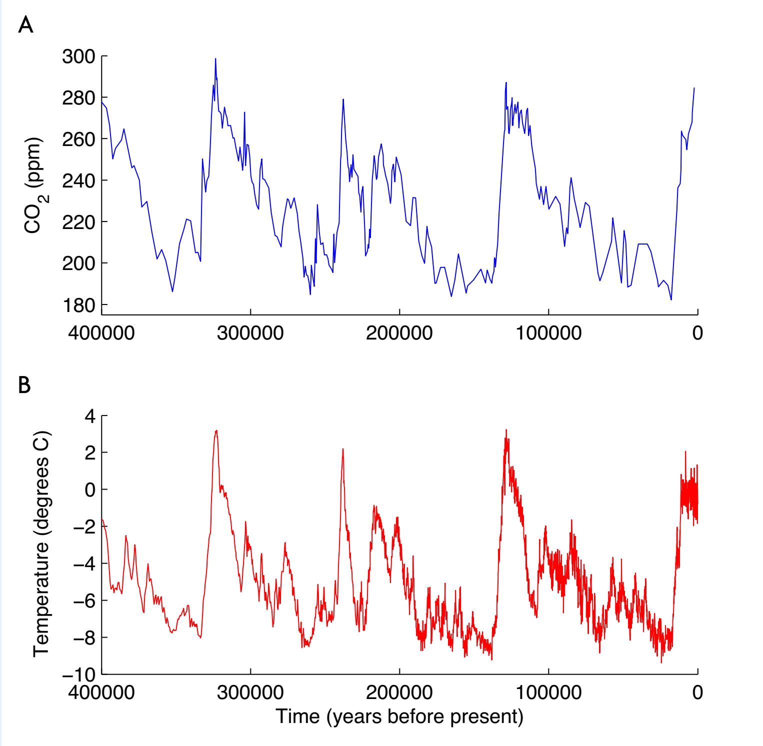
Climate change is too complex for students to fully understand with one or two graphs, but by explaining the strengths and weaknesses of any one piece of evidence, teachers can help students reach their own views about this topic. Teachers can extend the discussion of cutting-edge, climate-change science by referring to the Intergovernmental Panel on Climate Change's Fifth Assessment Report from Working Group I (see "On the web"). ■

Carolyn Dash (cdash@hamilton.edu) *is a visiting professor of Environmental Studies at Hamilton College in Clinton, New York, and developed this activity as a graduate student at the University of Illinois at Urbana–Champaign with Barbara Hug* (bhug@illinois.edu), *a clinical associate professor in the College of Education at the University of Illinois at Urbana–Champaign.*

FIGURE 5

The Vostok ice-core record.

The below graphs show changes in (A) atmospheric carbon dioxide (CO_2) concentration (parts per million [ppm]) and (B) temperature (degrees Celsius) over the course of 420,000 years (until 1950) (Petit et al. 1999). (**Note:** The time periods shown in these graphs refer to years before 1950 [labeled "present"], addressing changes in the ratios of the atmosphere's carbon isotopes due to nuclear weapons testing.)



Acknowledgments

This lesson was funded by the National Science Foundation Grant #0816610 (awarded to Feng Sheng Hu and David Nelson at the University of Illinois at Urbana–Champaign). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation and the University of Illinois. We would like to thank the codevelopers of the workshop where this lesson was first used: Ryan Kelly, Hillary Lauren, James Planey, and Michael Urban. We would also like to thank the teachers that attended our workshop for their participation and feedback and three anonymous reviewers for their comments, which greatly improved this manuscript.

On the web

Carbon dioxide reading: *http://keelingcurve.ucsd.edu*

Google Tends: *www.google.com/trends*

The Intergovernmental Panel on Climate Change's Fifth Assessment Report: *www.ipcc.ch/index.htm*

The Keeling Curve: *http://keelingcurve.ucsd.edu*

More information and additional curriculum materials: *www.nsta.org/highschool/connections.aspx*

References

Galesic, M., and R. Garcia-Retamero. 2011. Graph literacy: A cross-cultural comparison. *Medical Decision Making* 31: 444–457.

Keeling, C.D., S.C. Piper, R.B. Bacastow, M. Wahlen, T.P. Whorf, M. Heimann, and H.A. Meijer. 2001. *Exchanges of atmospheric CO_2 and $^{13}CO_2$ with the terrestrial biosphere and oceans from 1978 to 2000.* San Diego, CA: Scripps Institution of Oceanography.

National Governors Association Center for Best Practices and Council of Chief State School Officers (NGAC and CCSSO). 2010. *Common core state standards.* Washington, DC: NGAC and CCSSO.

NGSS Lead States. 2013. *Next Generation Science Standards: For states, by states.* Washington, DC: National Academies Press.

Petit, J.R., et al. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* 399: 429–436.