Data in Action: Making Predictions

LESSON 2: Making Climate Predictions

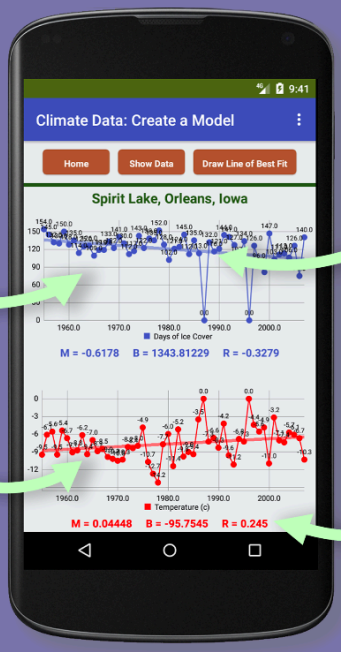
Ice skater and high school student Selena has collected lake ice data for a few years. Using her own data and some online historical data, she thinks that the number of annual ice days each winter may be declining. But the data on the graph goes up and down over the years, and it's hard to say for sure.

To learn more about the future of ice on her local lake, Selena creates a mathematical model in her app, a representation of something happening in the real world. Scientists use models to understand problems and make predictions.

Once you pick up some of the skills in this guide, you can modify your app to show any data for better understanding and making future predictions.

The data in this app is from a scientific study of the number of days a lake was frozen each winter.

This chart gives the annual average winter temperature around the lake.



The dark line shows a trend in the data since 1955 and helps you make predictions.

These numbers provide useful information about the trend.

Prerequisites:

- Before starting the lesson, students should complete **Lesson 1: Visualizing Climate Data**.
- Comfort navigating App Inventor Designer and Blocks screens.
- Up-to-date App Inventor Companion (version 2.70 or higher) for Android devices (**iOS not available yet**; iOS chart capability is currently in development)
- Some prior knowledge of algebraic and statistical terms such as slope, x and y intercepts, and correlation coefficient could be helpful.

What Can You Do?

PART 1: Get to Know the App Template

1. **Get the app template**
2. **Check out the design**
3. **Review the code**
4. **Look over the data**

PART 2: Add a Mathematical Model

5. **Add a Trendline component***
6. **Draw a line of best fit***

PART 3: What Can You Do with a Model?

7. **Get information about the line of best fit***
8. **Use the information to make predictions**

***New App Inventor Skills**

APPENDICES

- **Appendix A:** Link your app to a spreadsheet of your choice
- **Appendix B:** Create a Google Sheets credential for free to read and write to private spreadsheets

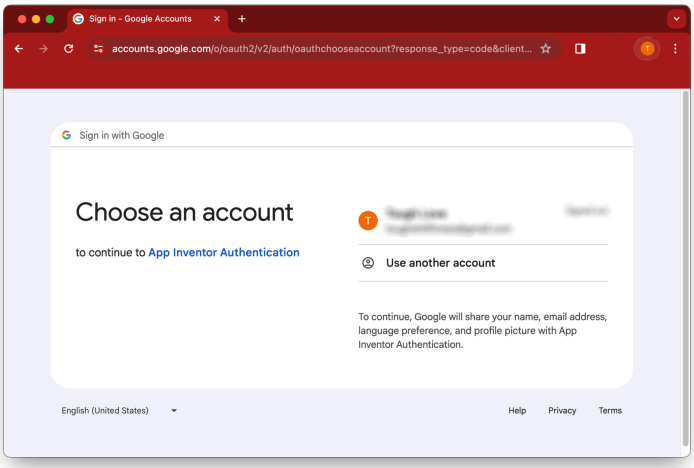
A note on the data used in this module:

- The lesson relies on historical data about one lake from a larger study of over 3,000 Northern Hemisphere lakes and the impacts of climate change: Sharma, Sapna et al. "Loss of Ice Cover, Shifting Phenology, and More Extreme Events in Northern Hemisphere Lakes." Journal of Geophysical Research: Biogeosciences 126, no. 10 (2021): e2021JG006348. https://doi.org/10.1029/2021JG006348.
- Additional lake data can be found in the paper above and at the **National Snow and Ice Data Center**.

PART 1: Get to Know the App Template

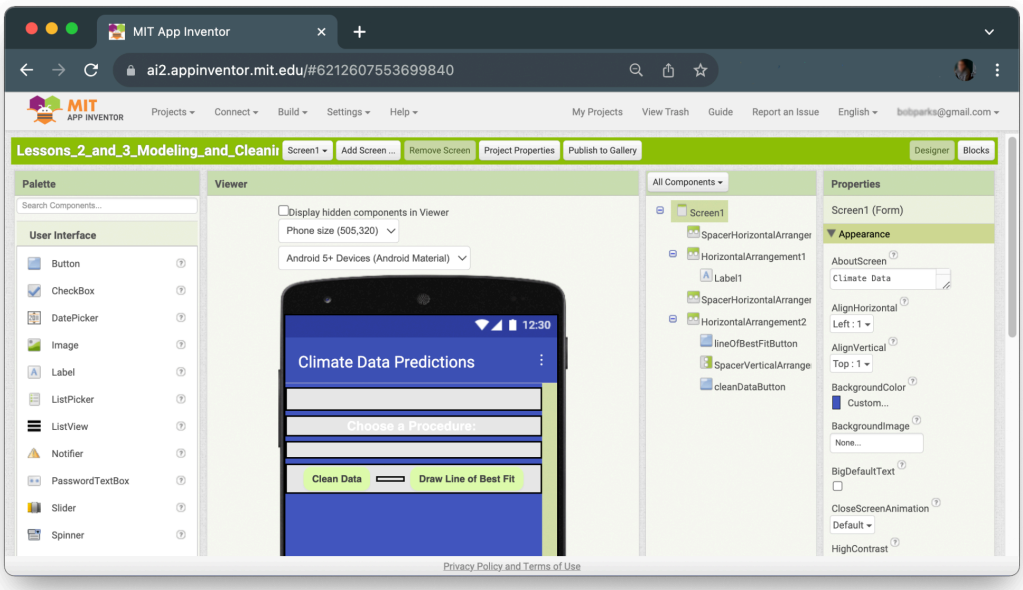
1. Get the app template

- **For Gmail users, [click this link to load the app automatically](#).**
  - Using a Gmail account is preferred because you can save apps to the account over time.



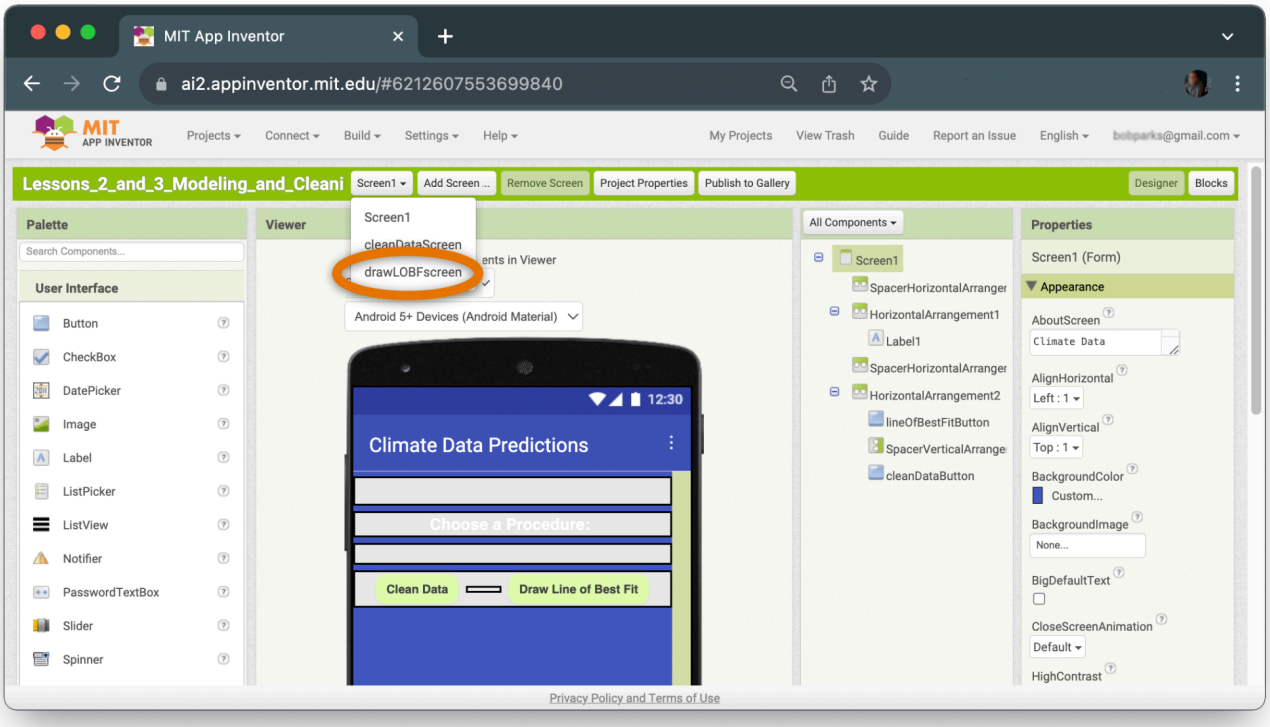
- **No Gmail account? [Log in anonymously using the *code.appinventor.mit.edu* site](#).**
  - The code.appinventor.mit.edu method also allows you to save apps, but you do have to keep track of an anonymous code and type it in to retrieve your apps each time.

- After logging in, the template should appear on the screen:
  - Please wait up to a minute for the app to reach your browser.

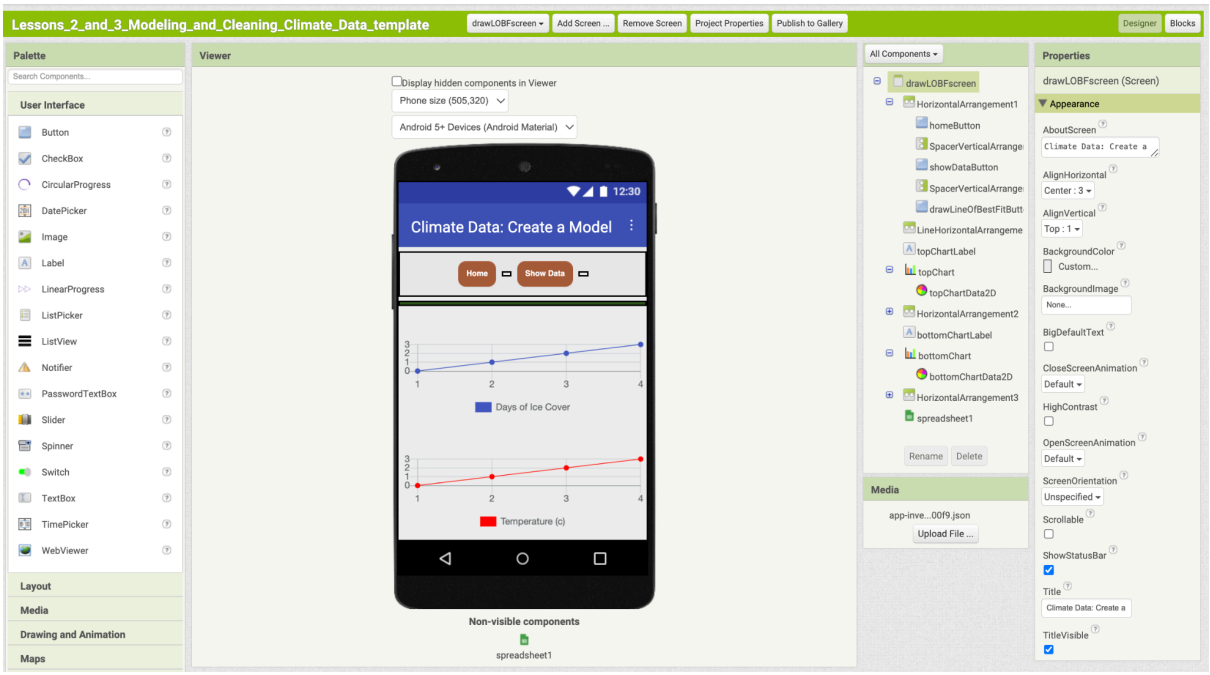


2. Check out the design

- The template will open on Screen1, a welcome screen. Click the screen selection button to switch to **drawLOBFScreen**. (LOBF stands for “line of best fit” — we’ll define this math term shortly!)



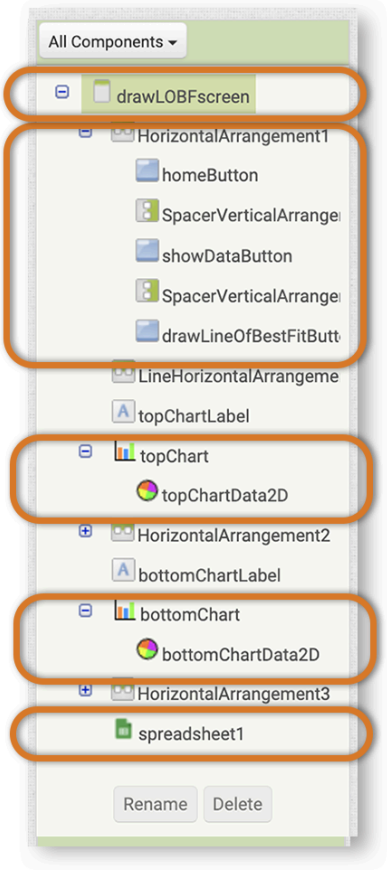
- You should see the **drawLOBFScreen:**



What Does It Do?

Look over the components in **Designer**. Can you guess what they do?

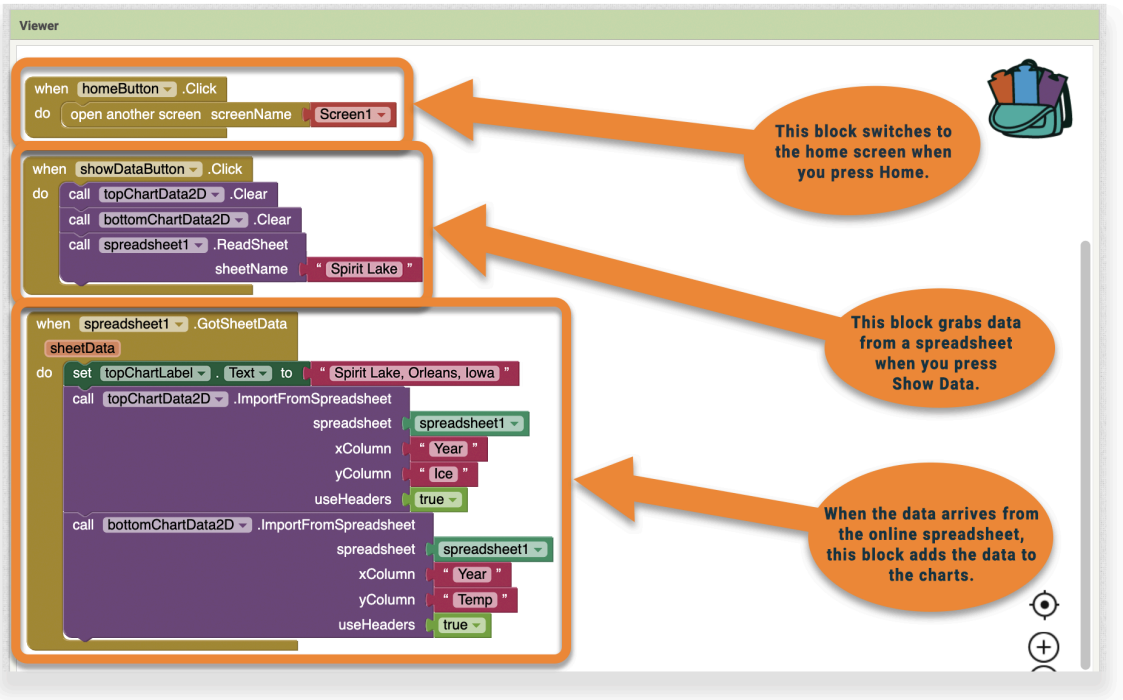
- The Screen component represents one screen inside the app.
- The HorizontalArrangement component contains the three Buttons, with spacer VerticalArrangements between them for appearance. (The last button is currently set to be invisible).



- The topChart and topChartData2D components show the lake ice duration data.
- The bottomChart and bottomChartData2D components show the average winter temperature around the lake.
- The Spreadsheet component gets the data from a Google Sheets spreadsheet

3. Review the code

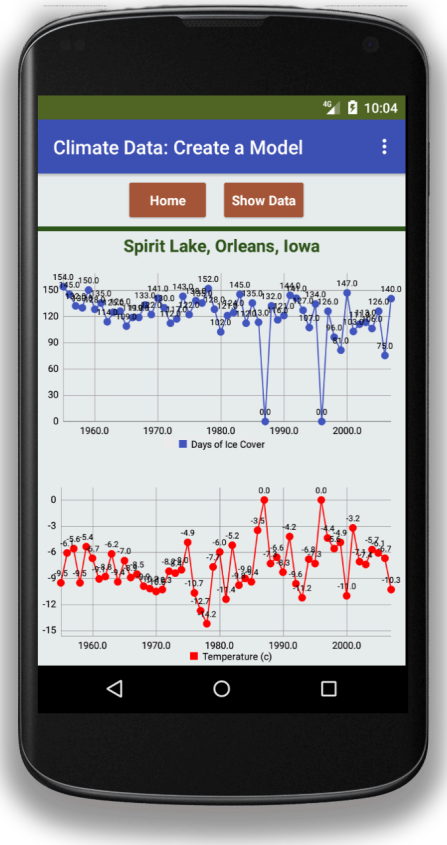
Click on **Blocks** in App Inventor to see the template code. You will add code to add a mathematical model in this lesson.



4. Look over the data

The template uses [data similar to what we used in Lesson 1](). Start by testing the app:

- **Test your code** with App Inventor Companion (**version 2.70 or higher**) on your phone or tablet.
- Click the **Show Data** button. A graph should appear.
  - The top graph in blue shows the **"ice duration"** — the number of days a lake in Iowa froze each year from 1955 to 2007.
  - The bottom graph in red shows the average winter temperatures around the lake.
  - You might notice two extra data points if you are familiar with Lesson 1. Can you spot them? We will explain more in Lesson 3 — Data Cleaning!
  - Can you spot a general trend in the number of days of ice? Can you spot a trend in the average winter temperature?



Trend-spotting by eye is hard! Next, you will create a mathematical model called a *line of best fit*. The line appears over the data and helps make sense of it. For Selena, adding a line of best fit:

- Clarifies trends to understand the data better
- Reports how confident her model is in representing the data accurately
- Predicts the future behavior of the research subject (in this case, Spirit Lake!)

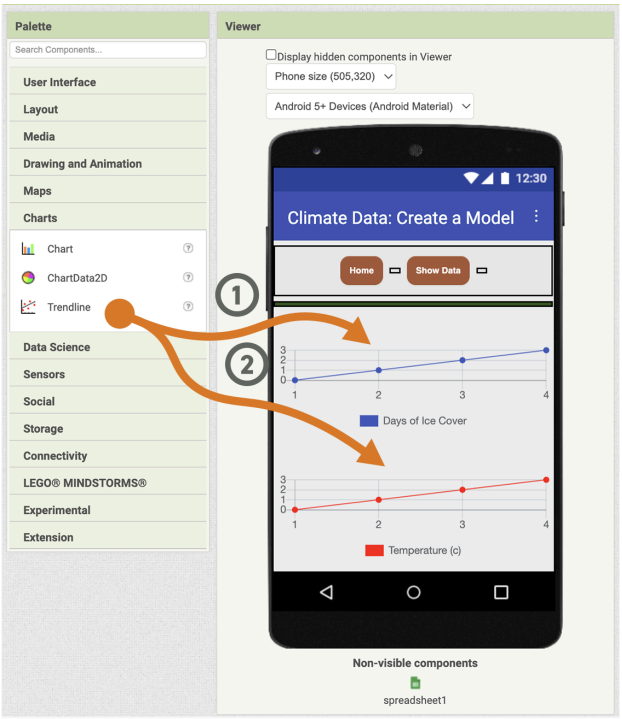
A few notes before you start:

- Terminology discussed in this lesson will be more familiar to those who have studied some high school algebra and statistics. Stick with it! This project can give you the basic idea of a "line of best fit" without knowing the nitty-gritty math to construct it.
- In the app, you will use a regression model. A regression is a statistical method for relating a dependent variable (in this case, *time*) to an independent variable (in this case, *the number of days of ice cover*). You will use linear regression, which uses a "line of best fit," but many other regression models (including quadratic, cubic, exponential, and logarithmic) are possible.

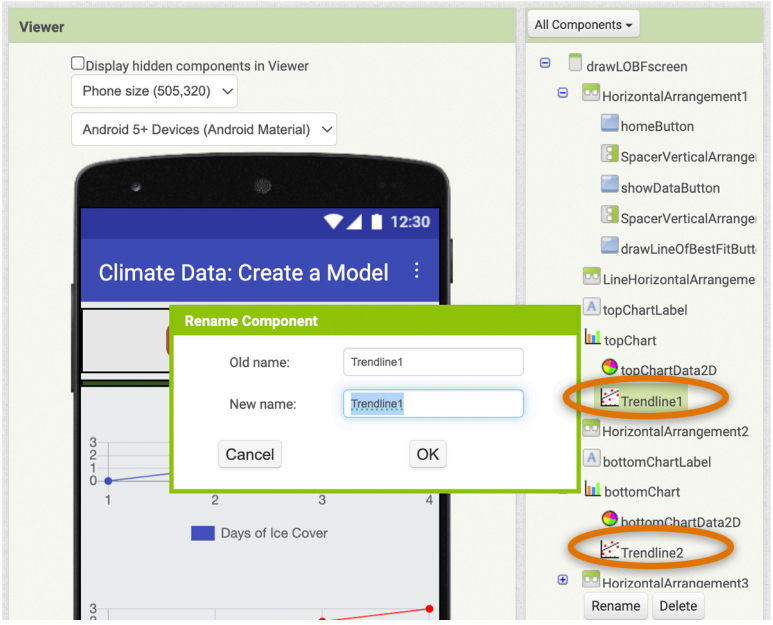
PART 2: Add a Mathematical Model

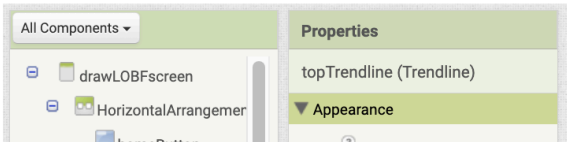
5. Add a Trendline component (NEW SKILL)

- A trend line will help you summarize patterns in the data and make predictions about future events.
- From the **Charts** drawer, drag out **Trendline** components for the top and bottom charts.



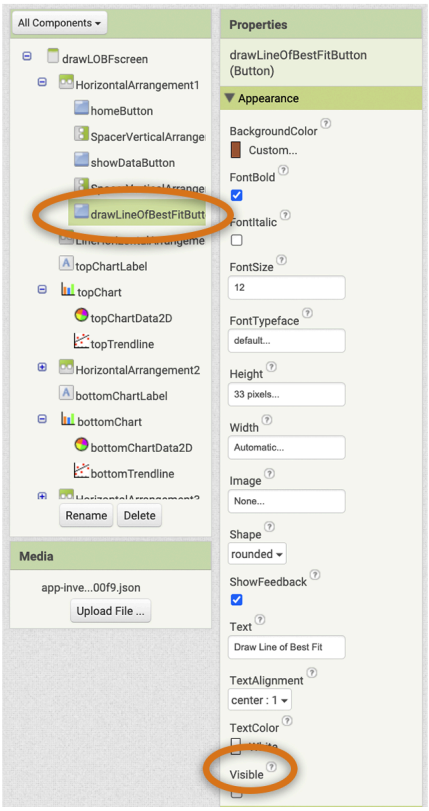
- **Rename them:**
  - "topTrendline" and
  - "bottomTrendline"





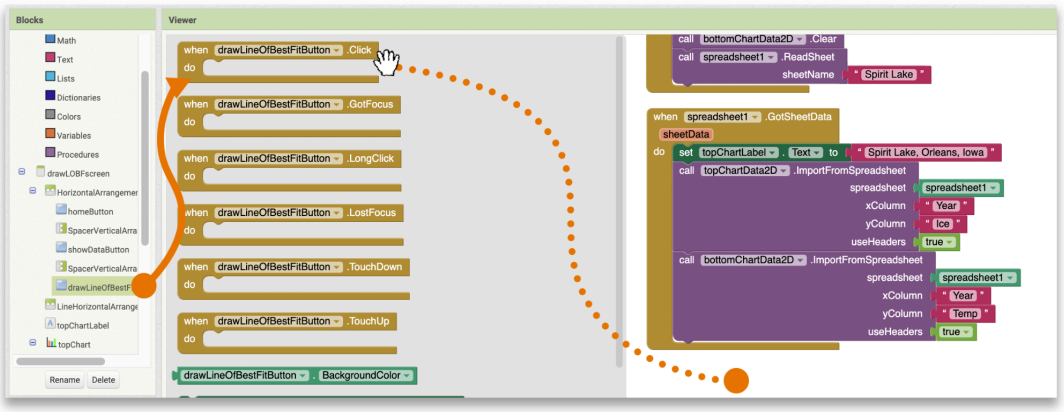
- **Change the line thickness:**
    - Highlight the topTrendline and adjust the StrokeWidth. (We increased the line to 4.)

- Make the **drawLineOfBestFitButton** visible.

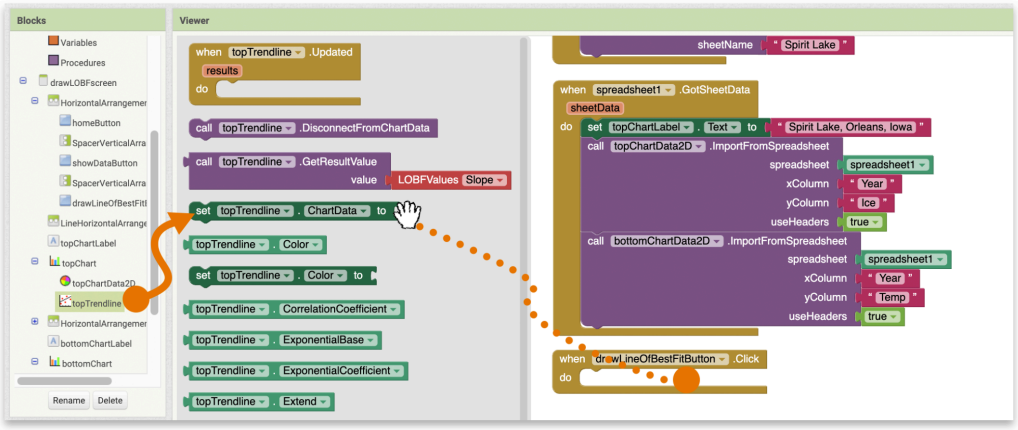


6. Draw a line of best fit (NEW SKILL)

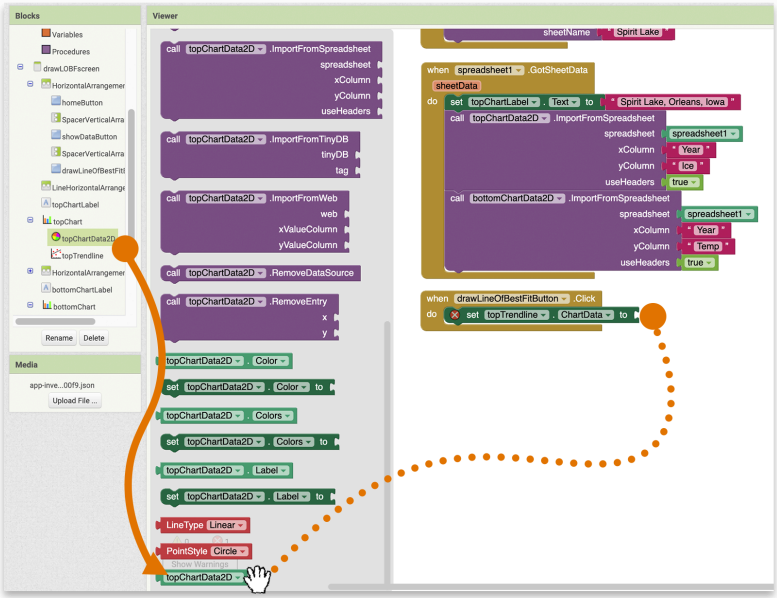
- You don't have to know how to calculate a line of best fit to see that it is a good way of showing a general trend for your data. Use this trend line to help understand what's going on with the lake ice data over time.
- To start, drag out the button block **when.drawLineOfBestFitButton.Click** —



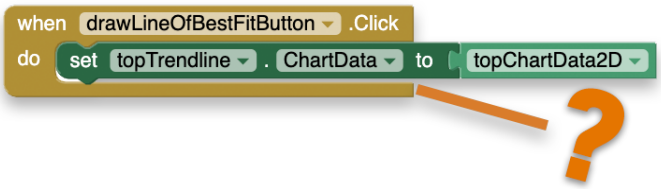
- You need to tell the trendline component which data to use. From the **topTrendline** drawer, drag out the **set topTrendline.ChartData to** block:



- From the **topChartData2D** drawer, drag out the **topChartData2D** block (look all the way to the bottom!):



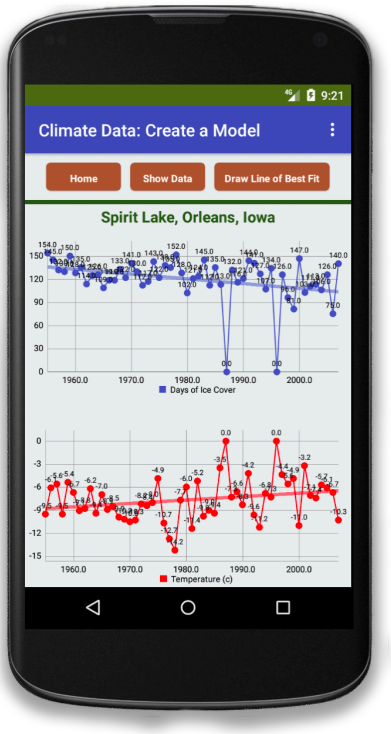
- What code will you add to draw a trend line in the bottom chart? (Try this step independently, then check your code in the testing box on the next page.)



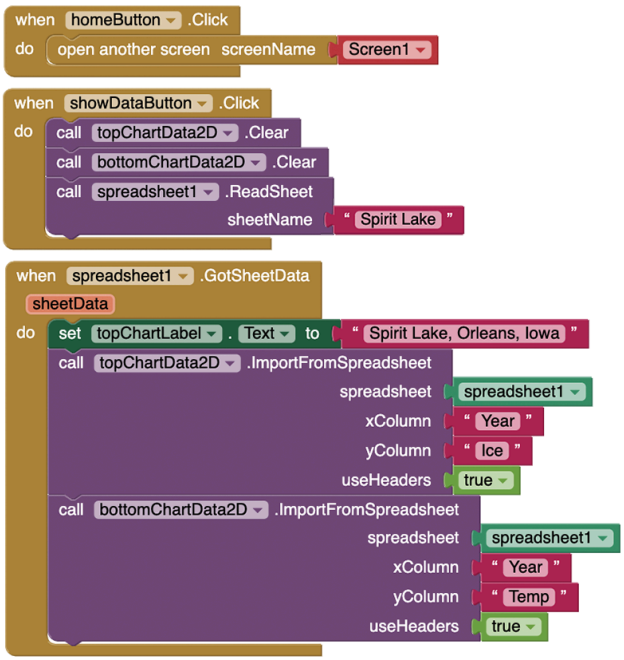
. . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . .
. . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . .
. . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . .
. . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . .
. . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . . . testing . .

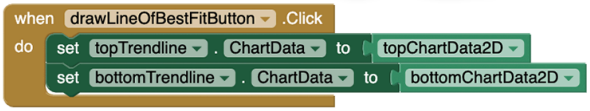
Test your code.

- Click the **Show Data** button.
- Once the graph appears (it may take a few seconds), click on **Draw Line of Best Fit**. Your app should look like the one to the right.



- If the buttons are not working, check your code:



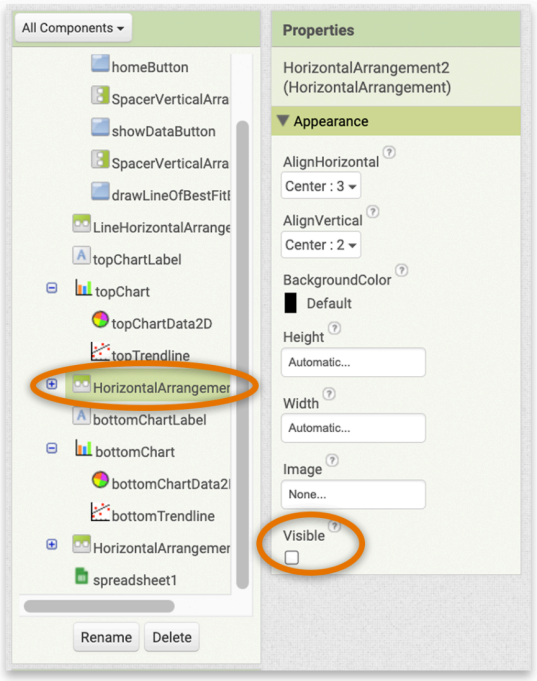


PART 3: What Can You Do with a Mathematical Model?

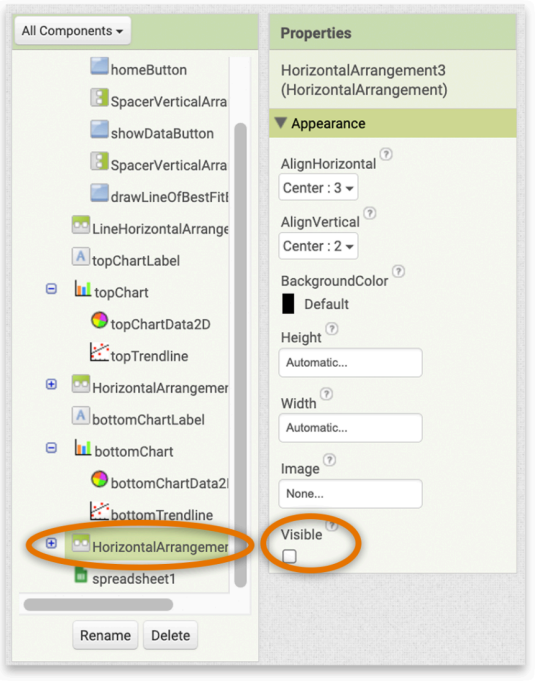
7. Get information about the line of best fit (NEW SKILL)

- Start by seeing what information is available about the line of best fit. In Designer, make **HorizonalArrangment2** and **HorizonalArrangment3** visible. These will show important labels for the trend lines.

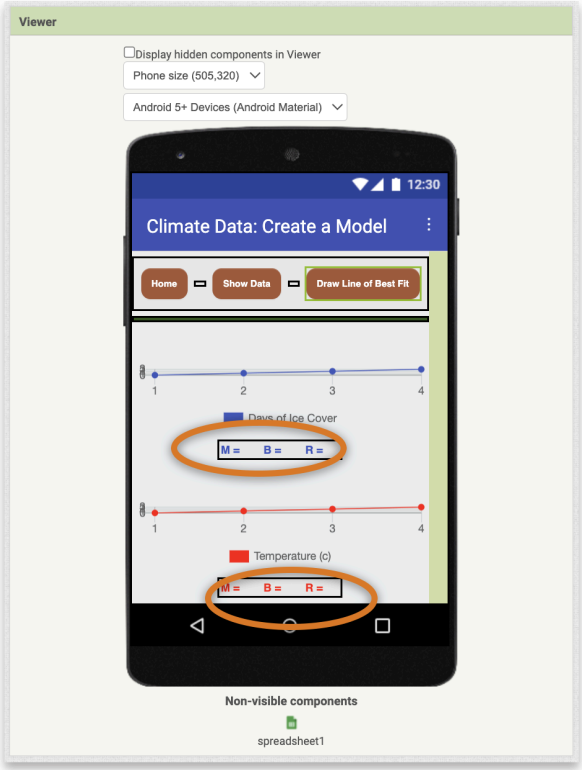
1.)



2.)



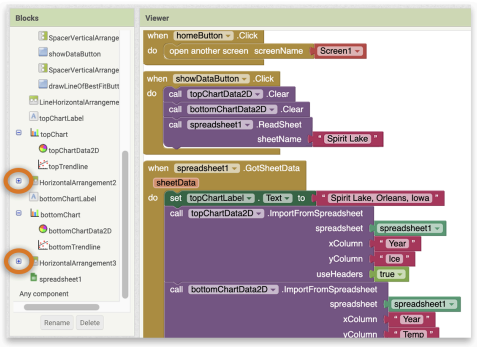
- Check out the new labels you just made visible:



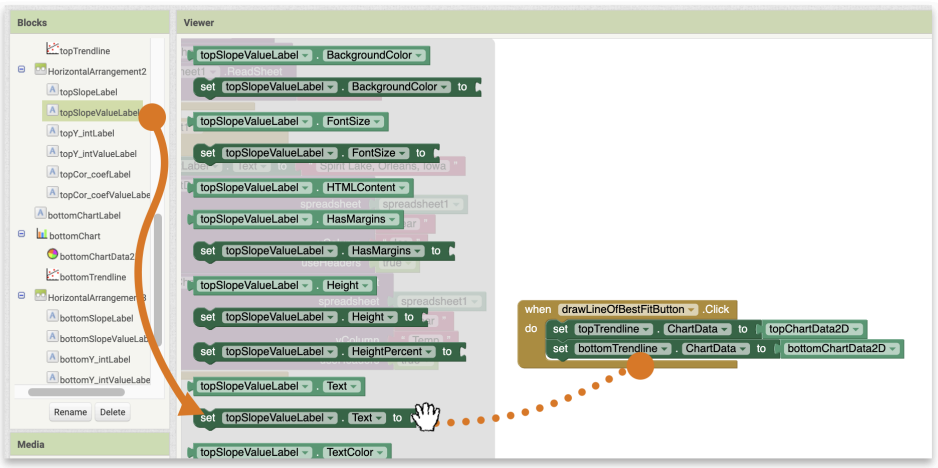
Discuss: What are M, B, and R?

- While live testing your app, you should see values appear for M, B, and R. What are these? (You might recall “M” and “B” from math class, but maybe not “R”.)
- Do you remember the classic formula y=mx + b from algebra class? (Good memories? Chills?)
  - M = Slope of a line
  - B = Where the line crosses the y-axis (the “y-intercept”)
  - R = Correlation coefficient (This one may be new unless you’ve taken some precalculus or statistics!)
- For our ice skater, Selena, the values provide important information about the line of best fit. How can she use the values to make predictions about the future of Spirit Lake?

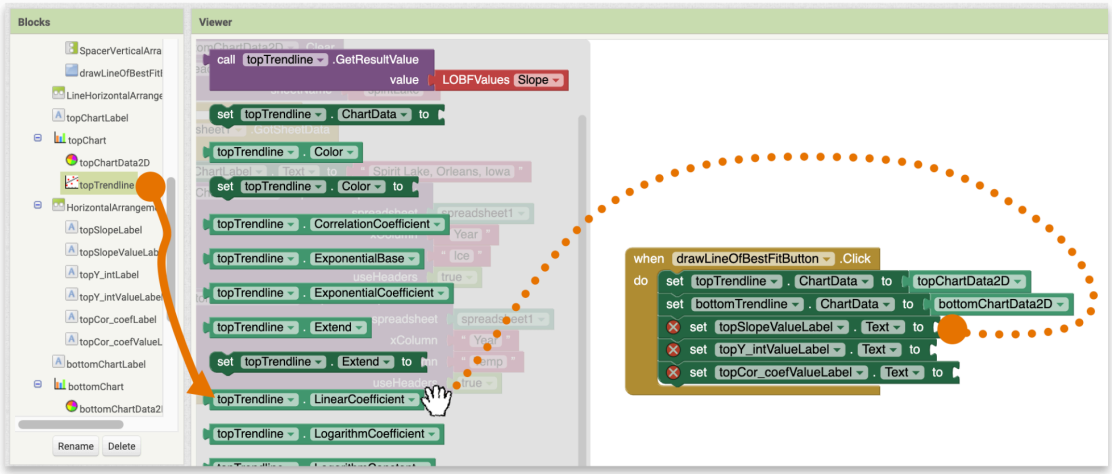
- In **Blocks**, code the labels to show information about the line of best fit.
    - First, click on the little **+** signs next to HorizonalArrangment2 and HorizonalArrangment3 to show all the labels:



    - **For each "value" label for the top chart**, drag out all the set text blocks for **topSlopeValueLabel**, **topY_intValueLabel**, and **topCor_coefValueLabel**. Drop them in **drawLineOfBestFitButton.Click**

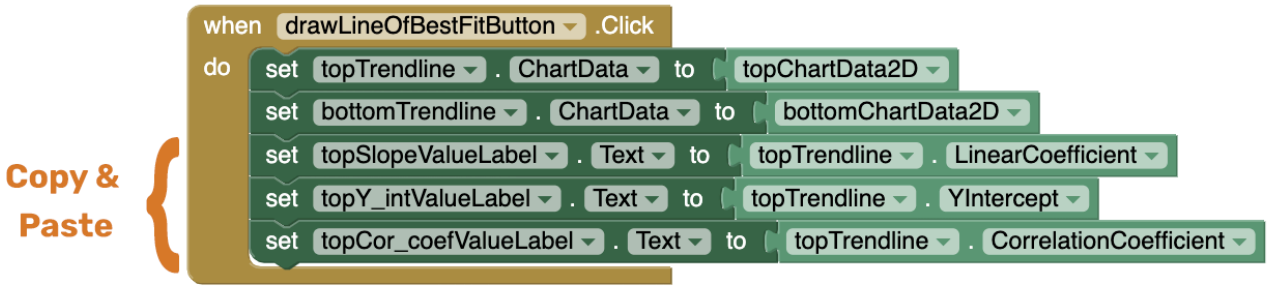


    - **Add the slope value to the label.**
        - When considering many different mathematical models, ***slope* can also be described as "linearCoefficient"**
        - From the **topTrendline** drawer, drag out a **topTrendline.LinearCoefficient** block and snap it to the topSlopeValueLabel block:

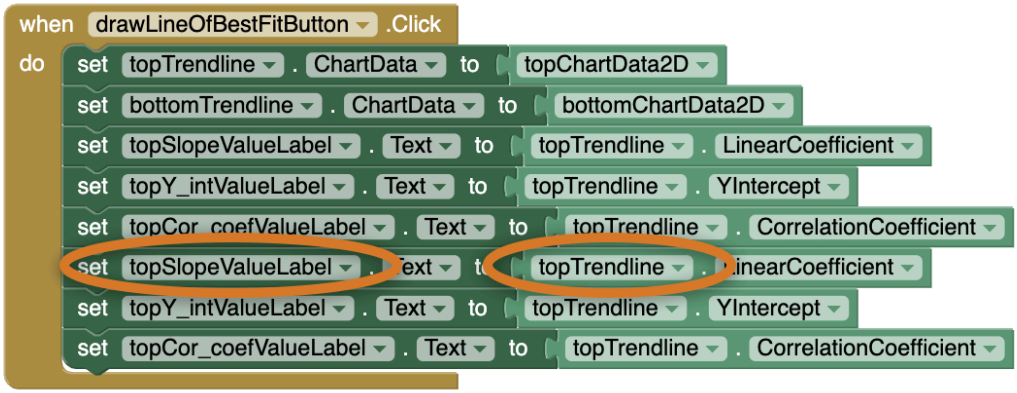


- From the same **topTrendline** drawer, drag out the **topTrendline.Yintercept** **topTrendline.CorrelationCoefficient** blocks and snap them into their labels.

- The values for the bottom trendline still need to appear. To save time, **copy & paste the three top labels** and place the copies below.

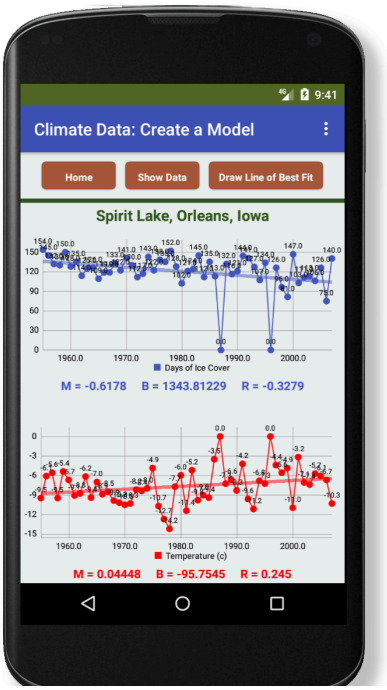


- **Change the labels and values to reflect the bottom chart and trendline.** (Try these steps independently, then check your code in the testing box on the next page.)

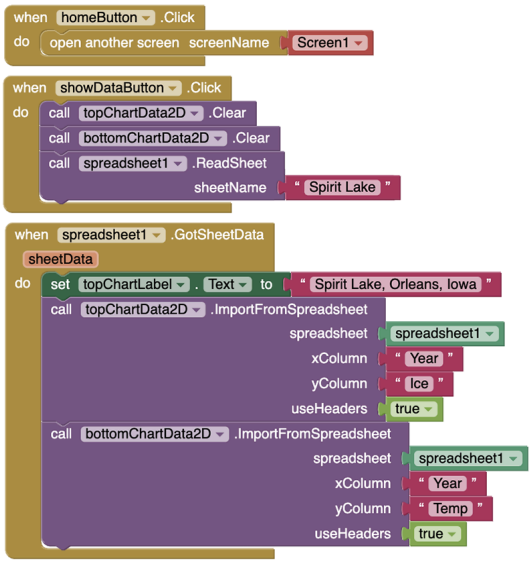


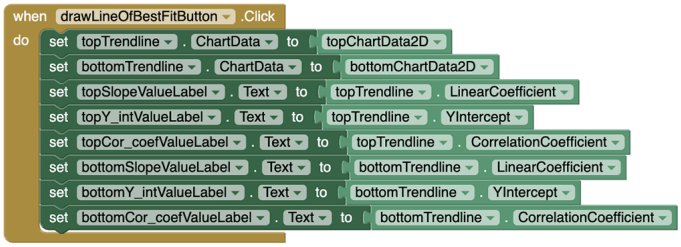
Test your code.

- Click the **Show Data** button.
- Once the graph appears (it may take a few seconds), click on **Draw Line of Best Fit**. Your app should look like the one to the right:



- If the buttons are not working, check your code:





8. Use the values for M, B, and R to make predictions

Discuss: Data science projects are all different!

Data science often requires us to use our judgment and research skills to determine what information is most important. Understanding what the numbers mean depends on the subject and context of the study (part of a study's *domain*). Learning about the domain (in this case, lake ice data!) helps data scientists focus on the specific information they need to use to answer questions or solve problems.

Examine the values for the **lake ice model** (the top chart):

- What does **M** (the slope of a line) mean in the real context of lake ice? How could you use M to describe the rate and direction of change for the annual days of ice cover on Spirit Lake? For example, for every 10 years, how many fewer days of ice will appear on the lake, according to the model?
- Now, consider the meaning of **B** (the y-intercept). The y-intercept is the number of ice duration days when x is 0 — in other words, a long, long time ago when the calendar year was 0 (the start of the Common Era). Can you explain why the y-intercept is a meaningless number in this case?
- What does **R** (the correlation coefficient) mean in the context of the lake ice model? If you have studied statistics, you will know that the correlation coefficient measures how well a linear model represents the data. An R of 0 means no correlation, and an R of +1 or -1 means a perfect correlation for either a positive or negative trend. In this case, R is far from -1, so the relationship between the negative trend line and the limited ice data is not very good. To be sure of a strong correlation, we need to compare data from many lakes. (Note: this is done with thousands of lakes in the study linked on page 2.)
- Finally, let's look at a value that does not appear but is easy to calculate with simple algebra — **the x-intercept**.
    - What does the x-intercept represent in this context? (Keep in mind that this data comes from only one lake. But if the data were representative of thousands of lakes, what would be the implications for climate change?)
    - Using your model for the value for M and B — and your trusty slope-intercept form y = mx +b — calculate the x-intercept of the line. Use a piece of scratch paper and a pencil to do your calculations.

Next, look at the values for the **temperature model** (the bottom chart). The values are winter average temperatures for the region around Spirit Lake.

- What does **M** imply about temperature rise every 10 years? How about every 100 years? (Again, the data has too few data points to be generalized with much confidence.)

    - The United Nations predicts a 3°C increase by 2100. Based on your own limited data, is your prediction higher or lower?
- Is **B** meaningful for the temperature model? Why or why not?
- According to **R**, how well does the temperature model represent the temperature data? How would adding more temperature data points provide more confidence in what the model is telling us?
- As with the lake ice, what does the **x-intercept** tell us? Calculate this value and explain the implications of your results.

EXTENSIONS

When you are done coding this part of the app, consider adding custom features or try one of the ideas below:

- **Fix the Bug.** There is a minor bug in the app. You get an error if you press the Draw Line of Best Fit button before the Show Data button.
    - Consider using these blocks in the **drawLineofBestFitButton.Click** event handler to fix the issue —

APPENDICES

- **Appendix A:** Link your app to a spreadsheet of your choice
- **Appendix B:** Create a Google Sheets credential for free to read and write to private spreadsheets