MODELING RAINFALL WITH CODAP

Simulating Vacations

In this activity, you are going to be introduced to CODAP (Common Online Data Analysis Platform). We will use CODAP to model the rainfall in different cities and simulate the number of hours it might rain during hypothetical vacations.

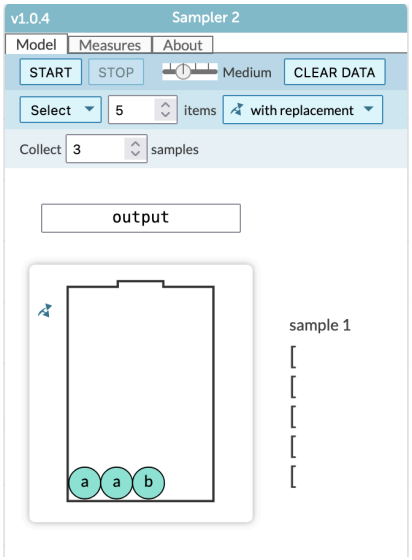
- To open CODAP navigate your web browser to https://codap.concord.org/
- Click the orange "Launch CODAP" button in the webpage's navigation bar.
- Finally, click the "Create New Document" button.

This will open up a new CODAP workspace.

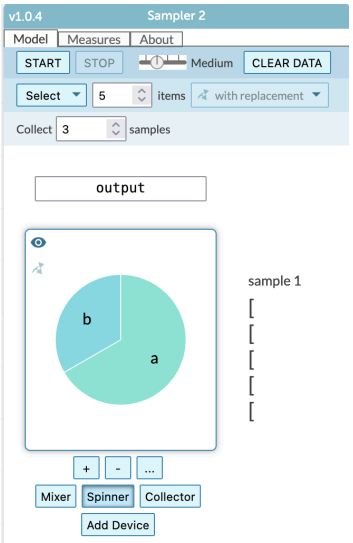
Modeling Rainfall in Seattle

In the CODAP toolbar, click "Plugins" and then select Simulation > Sampler. This will create a sampler in your workspace. The default sampling device is called a mixer. (It looks like a bin with three balls in it.) You can change the sampling device by clicking on the mixer.

- Change the sampling device to a spinner.

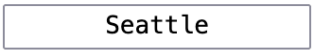


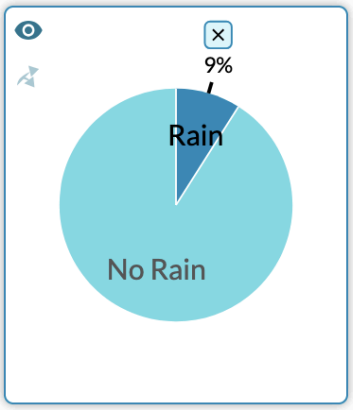
Default sampling device is a mixer.



After changing the sampling device to a spinner

To model the rainfall in Seattle, we need to set up our spinner to have two sections "Rain" and "No Rain". These sections also need to mimic the probabilities of rain and no rain in the city of Seattle, namely 9% and 91%, respectively. To do this:



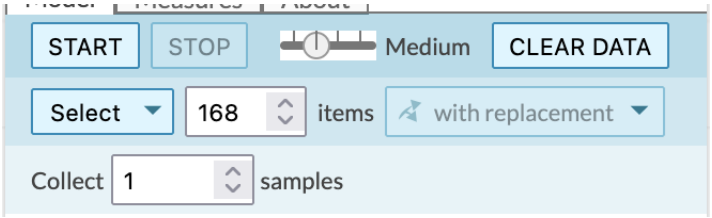


- Click on "output" and change the text label to "Seattle".
- Click on the "a" spinner section label and change it to "Rain".
- Click on the "b" spinner section label and change it to "No Rain".
- Click on one of the percentages associated with Rain or No Rain and change it to the appropriate percentage. (This should automatically set the other section percentage since we only have two sections.)

Simulating Data For a Hypothetical Vacation

Above the sampling device, we are going to change the number of items and samples we will be collecting.

- Change the number of items from "5" to "168". This reflects the number of hours in Albert Hoffman's vacation.
- Change the number of samples from "3" to "1". This will simulate one hypothetical vacation.



Vocabulary

In a simulation, we generate multiple samples of data from a defined model. Each sample generated is referred to as a **trial** of the simulation. In our example each trial represents a hypothetical vacation of 168 hours.

- To simulate data for the first trial, click the "Start" button. (After you watch the spinner select a couple values, move the speed slider to "Fastest" to increase the speed of the data generation.)

Plotting the Results of the Trial

Recall that we want to summarize the number of hours it rained on Albert Hoffman's vacation. The easiest way to do this is to plot the 168 simulated values from the hypothetical vacation that we just simulated. The data generated from the sampler is collected in a window that has three different tables in it. The first table is called "experiments", the second table is called "samples", and the third table is called "items". The results for our trial are stored in the "items" table. To create a plot of these values:

- Click on the Graph icon in the CODAP toolbar. This will open a blank graph in your workspace.
- Click and hold on the "Seattle" attribute name in the "items" table and drag it to the *x*-axis of the graph in your workspace.

This should create a binned dotplot. Each bin includes either "Rain" or "No Rain".

Questions

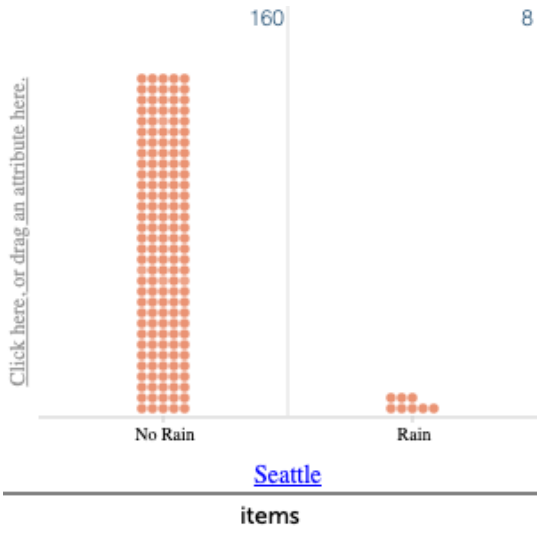
1. Identify the cases in this plot using the context of the problem. E.g., what does each dot in this plot represent?

Counting the Hours of Rain

We want to count the number of dots in the "Rain" bin. To do this:

- On the right-side of the graph, select the ruler icon. (This is a way to add measures to your plot.)
- Check the box next to "Count".

This should add the counts for each bin to your plot.



Questions

2. Record the number of hours of rain for the first hypothetical vacation you generated in the table below.

Hypothetical Vacation	Hours of Rain
1	
2	
3	
4	
5	

Run Additional Trials

To run an additional trial, that is generate data for another hypothetical vacation:

- Click the "Clear Data" button in the sampler.
- Click the "Start" button.

This will generate data for another hypothetical vacation. It will also plot the data and count the number of cases in each bin in the plot.

Questions

3. Record the number of hours that it rained from this hypothetical vacation in the table above.

4. Carry out 3 more trials. Don't forget to clear the data from each trial before running the new trial. Record the number of hours of rain from each hypothetical vacation to the table above.

5. Add the five values from your table into the instructor's computer.

After all the students have entered their simulated values into the instructor's computer, the instructor will create a plot of everyone's data. (While you wait, you can move on to modeling Olympia's rainfall.)

1. Sketch the plot of the number of hours of rain.

2. Identify the cases in this plot using the context of the problem. E.g., what does each dot in this plot represent?

3. Fill in this sentence using your plot as evidence:

   If Albert Hoffman was vacationing in Seattle, it would not be unexpected that it would rain between _____ and _____ hours of his vacation.

Modeling Rainfall in Olympia

4. Open up a new CODAP workspace and set up a sampler to model a hypothetical vacation in Olympia, which has measurable rain in 13% of recorded hours. Sketch a picture of your sampler below. Include enough detail so that someone could replicate your sampler in CODAP.

Automating the Counting and Recording of the Number of Hours of Rain

Rather than manually recording and building the dotplot of the number of hours of rain from many hypothetical vacations, we will use CODAP to do this.

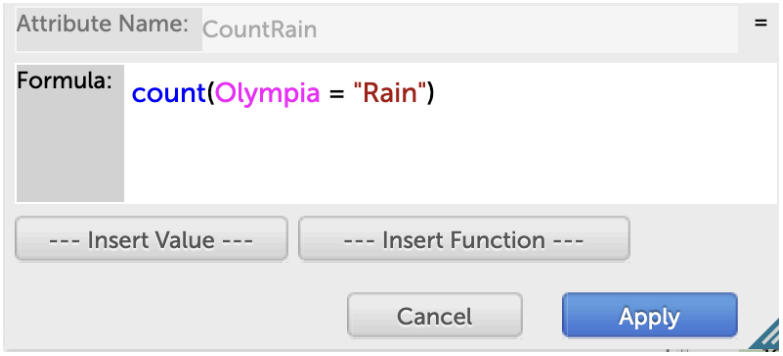
- Change the Collect value above the sampler from "1" to "5".
- Click "Start"

Now in the Sampler Data window, you should see five rows in the samples table. Each row corresponds to a different trial of the simulation. If you click on one of the rows, it highlights the 168 simulated hours for that trial in the items table.

We want to have CODAP compute the number of hours of rain for each trial. Note that this is a summary measure at the trial level. To compute a summary measure at the trial level, we will be working in the samples table.

- Click the "+" in the **samples** table. This will add another column (i.e., attribute) to that table.

- Name this attribute CountRain. (Note that attribute names cannot include spaces!)
- Click on the CountRain name (it should look like a hyperlink) and select Edit Formula… This will open the formula editor.
- We are going to use the count() function to count the number of values in the items table that are "Rain" for each trial. In my sampler, I named the sampler "Olympia" and the label on the spinner section was "Rain". (I use upper-case "O" on Olympia and upper-case "R" on Rain.) So for my formula I would type the following:



IMPORTANT

If you used different names, spelling, or capitalization, you need to change the formula to use exactly the name of the attribute in the items column and how you labelled rain in your spinner.

After clicking "Apply", the count for the number of rainy hours in each trial should appear in the CountRain attribute column.

Questions

5. Create a graph of the data in the CountRain attribute.
6. Identify the cases in this plot using the context of the problem. E.g., what does each dot in this plot represent?

Run More Trials

In simulation studies, we need to run several trials of the simulation to get a sense for the patterns and variation in the summary measures that we are interested in.

- Above the sampler, click the "Clear Data" button to remove the data from your initial five trials.
- Change the collect value to "500".
- Click "Start"

Questions

7. Sketch the graph of the number of hours of rain for the 500 simulated vacations.

8. Fill in this sentence using your plot as evidence:

   If Albert Hoffman was vacationing in Olympia, it would not be unexpected that it would rain between _____ and _____ hours of his vacation.

Modeling Rainfall in Astoria

Open up a new CODAP workspace and set up a sampler to model a hypothetical vacation in Astoria, which has measurable rain in 16% of recorded hours. Run 500 trials of the simulation. Use CODAP to count the number of hours of rain generated in each trial.

9. Sketch a picture of your sampler below. Include enough detail so that someone could replicate your sampler in CODAP.

10. Create a graph of the data in the attribute you created to count the number of hours of rain in each trial.

11. Sketch the graph of the number of hours of rain for the 500 simulated vacations.

12. Fill in this sentence using your plot as evidence:

    If Albert Hoffman was vacationing in Astoria, it would not be unexpected that it would rain between _____ and _____ hours of his vacation.

Evaluating Where Albert Hoffman Vacationed

Suppose Albert Hoffman visited a city in the Pacific Northwest for one week. He used the methodology that the Western Regional Climate Center uses to collect hourly data from his vacation. Based on this, on his vacation it rained only 15 out of the 168 hours.

13. If Albert Hoffman claimed his vacation was in **Seattle, WA**, would you believe his claim? Use the results of your simulation and the range of values you gave in Question #3 to justify your answer.

14. If Albert Hoffman claimed his vacation was in **Olympia, WA**, would you believe his claim? Use the results of your simulation and the range of values you gave in Question #8 to justify your answer.

15. If Albert Hoffman claimed his vacation was in **Astoria, OR**, would you believe his claim? Use the results of your simulation and the range of values you gave in Question #12 to justify your answer.