

Name(s) \_\_\_\_\_ Period \_\_\_\_\_ Date \_\_\_\_\_

# Project Guide - Tell a Data Story






## Background

Computing tools allow people from all different fields to learn from data. By transforming and visualizing data we can find patterns that help us learn about our world.

For this project you will

- Select a dataset from the Data Library
- Use the Data Visualizer to make a visualization that helps us learn something interesting about the data
- Write a response that explains:
  - what your visualizations show
  - new insights or decisions that can be made based on this visualization
  - possible problems with your analysis including potential bias in the data or your interpretation

## Tell a Data Story

Data Analysis Process	Written Response																																			
 <p>Collect or Choose Data</p>	<p>Select a dataset from the Data Library. Read the metadata to understand what information is available in the table</p> <p><b>Dataset Name:</b> Daily Weather</p> <p><b>Short Description:</b> Five day weather forecast for 100 US cities, updated every five days, including a description of the conditions.</p>																																			
 <p>Clean and/or Filter</p>	<p><b>Did you filter or clean the data? Why or why not?</b></p> <p>Yes, I filtered the data to just look at the state of California since that is where all the wildfires have been happening this summer.</p>																																			
 <p>Visualize and Find Patterns</p>	<p><b>Create a visualization. Copy/paste the chart below.</b></p> <p><b>What has the weather been like in California while these summer wildfires are burning?</b></p> <p><b>What has the weather been like in California while these summer wildfires are burning?</b></p> <table border="1"> <thead> <tr> <th></th><th colspan="6">City</th></tr> <tr> <th>Condition Description</th><th>Eureka</th><th>Los Angeles/Oxnard</th><th>Sacramento</th><th>San Diego</th><th>San Francisco Bay Area/Monterey</th><th>San Joaquin Valley/Hanford</th></tr> </thead> <tbody> <tr> <td>few clouds</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> <tr> <td>scattered clouds</td><td>0</td><td>0</td><td>0</td><td>0</td><td>3</td><td>0</td></tr> <tr> <td>sky is clear</td><td>4</td><td>4</td><td>5</td><td>5</td><td>1</td><td>5</td></tr> </tbody> </table> <p>Created using Code.org App Lab on 2020/08/25 from the Daily Weather table in project</p> <p>X Values: Condition Description, Y Values: City, Filtered State to "California"</p>		City						Condition Description	Eureka	Los Angeles/Oxnard	Sacramento	San Diego	San Francisco Bay Area/Monterey	San Joaquin Valley/Hanford	few clouds	1	1	0	0	1	0	scattered clouds	0	0	0	0	3	0	sky is clear	4	4	5	5	1	5
	City																																			
Condition Description	Eureka	Los Angeles/Oxnard	Sacramento	San Diego	San Francisco Bay Area/Monterey	San Joaquin Valley/Hanford																														
few clouds	1	1	0	0	1	0																														
scattered clouds	0	0	0	0	3	0																														
sky is clear	4	4	5	5	1	5																														



New Information

**What information does your visualization show:**

My crosstab visualization shows that California is expected to have mostly clear sunny weather the next five days (8/26 - 8/30) across most of the state. In fact, in all places besides the SF Bay Area, at least four out of the next five days are forecast to be clear days. There is no rain forecast in any part of the state over the next five days.

**What new insights or decisions can be made based on this visualization:**

This visualization shows me that the weather will continue to be unfavorable for slowing the wildfires that are already burning in the state. The fact that there is no rain at all in the forecast means that fire crews will have to continue to fight fires without the help of Mother Nature. If I was in charge of coordinating the response to these fires, I would make sure to ask for more help and reinforcements from neighboring communities and states because judging from this data the fires may continue to get worse and the effort required to contain and fight them will only grow in size.

**What are the possible problems with your analysis including potential bias in the dataset or your interpretation:**

I am making assumptions about other conditions such as temperature and humidity. I am not sure what those conditions are forecast to be and how they will affect the spread of wildfires.

There is also potential bias in what constitutes a few clouds vs scattered clouds vs sky is clear. Who decides what percentage of the sky must be covered in clouds as the threshold for each category? Depending on how these distinctions are made, there could be a bias toward more or less clouds which could have an effect on planning on where to distribute firefighters to fight fires.

## Rubric

Category	Extensive Evidence	Convincing Evidence	Limited Evidence	No Evidence
<b>Collect or Choose Data</b>	Dataset is correctly identified and description is complete.	Dataset is correctly identified and description is mostly complete.	Dataset is correctly identified and description is somewhat complete.	Dataset is not identified or description is missing.
<b>Clean/Filter Data</b>	Explanation for cleaning and/or filtering is complete and reasonable.	Explanation for cleaning and/or filtering is complete and mostly reasonable.	Explanation for cleaning and/or filtering is somewhat complete or somewhat reasonable.	Explanation for cleaning and/or filtering is incomplete.
<b>Visualize and Find Patterns</b>	Visualization is readable and includes a title, and citation.	Visualization is mostly readable and includes a title, and citation.	Visualization is somewhat readable and/or is missing a title or citation. .	Visualization is unreadable or missing.
<b>New Information:</b> Interpreting the visualization	The visualization is thoroughly explained.	The visualization is mostly explained.	The visualization is somewhat explained.	The visualization is not explained.
<b>New Information:</b> Insights or Decisions	Insights or decisions are reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and somewhat effectively linked to information displayed in the visualization	Insights or decisions are missing.
<b>New Information:</b> Bias	Possible problems with analysis or potential bias are reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are mostly reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are somewhat reasonable and addressed.	Possible problems with analysis or potential bias are not addressed.