Identify a Driving Question:
• What do you wonder about this topic?
• What might someone learn from using the data you collect?

Are you collecting new data or using existing datasets?

Collect data related to the question:
What will you measure to inform your question?
• What tools can you use to collect this data? What are the affordances and limitations of using a computational tool?
• How frequently will you collect data? Are you collecting data on a certain day or during certain times?
• Is the data you are collecting able to be compiled easily (using the same units, avoiding open-ended responses to be able to group)?
• Could there be bias in your data (e.g. human error, missing data or questions that lead to certain responses)?

Locate data related to the question:
• What data that has been collected can help to answer your question?
• What dataset(s) can you compile information from?
• If there is more than one data set being used, is the data comparable?
• Identify cases and attributes in the raw dataset(s). How will you modify or use to answer your question?
• Could there be bias in the instruments used to collect the data (e.g. questions that lead to certain responses)?

Structure the dataset to be analyzed to inform the question
• How might you structure your dataset to help you or someone else to answer a question related to your findings?
• Is there any data missing?
• How can you provide context or cues (e.g. titles, labels, colors) to help someone else understand your dataset?
Part 2: Using Data Moves to Analyze Data

Use data moves (Erickson et al., 2019) to answer the questions about the dataset. Examples of data moves are:

- **Filtering** is selecting only a subset of the data that is available to explore. For example, you may view only a portion of the data because only some of it is relevant to the question you are exploring.

- **Grouping** is used to make comparisons between different subgroups of a data set.

- **Summarizing** is computing a value (e.g. mean, median, mode) to summarize a group.

- **Calculating** is creating a new attribute based on values of one or more existing attributes. For example, if an attribute is distance in kilometers, you may want to convert that value to distance in miles.

- **Merging/joining** is combining more than one dataset together. You may add more cases, or add more attributes to existing cases.

- **Making a hierarchy** is grouping data using multiple levels.

In order to complete data moves, you should become familiar with a program that allows you to manipulate data. Google Sheets, Microsoft Excel, CODAP (Concord Consortium) are all great resources for using data moves to analyze your data.

---

**Question:**

<table>
<thead>
<tr>
<th>Data Move(s)</th>
<th>Sketch or describe how you organized or illustrated the data:</th>
<th>What does this tell you about the relationships/patterns within your dataset?</th>
<th>What new questions do you have?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dive deeper into the data set by combining moves</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
Explore the sources that reported the data. Use the flowchart below to consider error and/or bias that could be present.

1. **Did you collect new data or use existing datasets?**
   - **Yes**
     - **Review the sources where the data was collected**
       - Are the sources that collected this data trustworthy?
       - Is the data relevant, up to date, and comparable?
   - **Not Sure**
     - **Think about the implications of data that is: collected by untrustworthy sources, not updated or relevant, or not comparable across multiple datasets.**
       - What type of organizations can I trust to have collected valid data on this topic?
       - How might the interest or bias of groups change the story or findings we can learn from this dataset?
   - **Self-reported**
     - **Think about the implications of self-reported data.**
       - How might people’s biases or perceptions change the story or findings we can learn from this dataset?
   - **Observed**
     - **Think about the implications of observed data.**
       - How was the data collected and by who?
       - How might the process or tool used for data collection change the story or findings we can learn from this dataset?

2. **Was the data self-reported or observed?**
   - **Self-reported**
     - **Think about the implications of self-reported data.**
       - How might people’s biases or perceptions change the story or findings we can learn from this dataset?
   - **Observed**
     - **Think about the implications of observed data.**
       - How was the data collected and by who?
       - How might the process or tool used for data collection change the story or findings we can learn from this dataset?

3. **Review your data findings with a partner.**
   - Are there any new questions you haven’t explored yet?
   - What data might be missing?
   - How might that data help you better understand the data you have already?