

Data can be presented in a lot of different ways. As budding STEM students, understanding not only what the data means, but HOW to present the data and demonstrate any key relationship(s) that might exist is a vitally important skill.

Data can be presented using a variety of computer or cloud-based programs. Some are free to use and some require a subscription. Each offers a particular interface and features, so it is good to know what each offers and which one would work best for your needs. Some common programs include Microsoft Excel, Google Sheets, Microsoft Office Excel Online, Apache OpenOffice Calc, and other database software. Online specialty programs include [GeoGebra](#), [R](#), [Tuva](#), and [Tableau](#) among others. Consequently, literacy in one or more of these programs is necessary to present data as well as compute statistical analysis.

VII. Instructions for using Google Sheets to present data

4. When **comparing data** of different information, they need to share an attribute in common so we can analyze any relationships between the two. What attribute of Chinook salmon populations and SRKW calf births do both data sets have in common?

5. Once you determine the attribute that allows you to compare that data, you need to see if all the data is usable for analysis. What value of the attribute you determined in the question above will you have to use as your **starting value** and why?

6. Now create a *data table* of **Chinook salmon populations** and **SRKW births**.
(NOTE: depending on which program you decide to use, you may need to search for tutorials on how to do so. Here are links for instructions on in [Google Sheets](#) ([video tutorial](#)); [Tableau](#); and [R](#)).
 - a. To begin, if you need to refresh and/or learn about data representation, statistics, spreadsheets, etc. proceed with these lessons from *Illustrative Mathematics*:
 - Lesson for [Data Representation](#).
 - Lesson for [Calculating Measures of Center and Variability](#).
 - Lesson for [Spreadsheet Computations](#), including the *mean*, *mode*, and *median* (example [lesson here](#)).

- Lesson for [Using Technology to Represent Data](#).
- b. Don't forget to *name* your Google Sheet and share it with your instructor(s).
Copy and paste a digital link of your Google Sheet data table here:
- _____
- c. Now we will want to **create a chart** in order to compare and visually present the data of Chinook populations and SRKW births.
- d. First we must decide what *type* of chart to create. Follow Google Sheet instructions for [Types of Charts](#) ([video](#) tutorial).
- e. What type of chart(s) did you end up creating?
- f. Next, you will need to decide **how to analyze** some of the data from your data table. Some options include creating box plots, histograms, scatter plots, etc.
- g. Using Google Sheets, follow these instructions on how to make a [Boxplot](#) ([video](#) tutorial); [Histogram](#) ([video](#) tutorial); and [Scatter Plot](#) ([video](#) tutorial) of your data. (**NOTE:** Be sure to keep these in the same Google Sheet of the file that you shared above).
- h. What does each of these tell you about the data?
- i. Next, calculate the mean, median, mode of Chinook salmon and SRKW data:
1. Mean:
 2. Median:
 3. Mode:
- j. What is the meaning of each of these values? (Review the lessons above, and also read [HERE](#) if necessary).

7. Finally, create a **scatter plot** of **Chinook population** vs **SRKW births**.
- a. Which data should be the *independent variable*? And on which axis of the chart is the independent variable plotted?

 - b. Which data should be the *dependent variable*? And on which axis of the chart is the dependent variable plotted?

 - c. Next, calculate and graph the **correlation coefficient** of the data. In Google Sheets, following the instructions on [Calculating the Correlation Coefficient](#) ([video](#) tutorial).

 - d. What does the value of the correlation coefficient mean about the data?

 - e. Next, calculate and graph the **linear regression** of the data. In Google Sheets, following the instructions on [Calculating the Regression](#) ([video](#) tutorial).

 - f. What does the value of the linear regression mean about the data?

VIII. Interpret the data

8. Make a claim based on the evidence that helps answer the question why Puget Sound Chinook salmon populations haven't shown improvements in recent years. Connect the pattern in the data to potential patterns and/or environmental changes. Justify your reasoning using data.

9. Discuss the connections and implications of Puget Sound Chinook salmon and Southern Resident Killer Whale populations. Justify your conclusions using data.

10. Science is an ongoing process. What new questions do you think should be investigated? What future data should be collected to answer them?