Visualizing Data

Data visualization is presented in the results and appendix sections using a combination of:

- Written commentary
- Drawings or photographs for illustrating qualitative data *(rare)*
- Graphs to illustrate patterns, trends or relationships with at least one quantitative variable.

Written Commentary

When you include a graph in the results section of your paper, you need to include written commentary that has:

- Reference to Figure that illustrates data
- A description of the results
- A statement of how you elected to visualize data
- An explanation of why you this visualization is appropriate

Example 1:

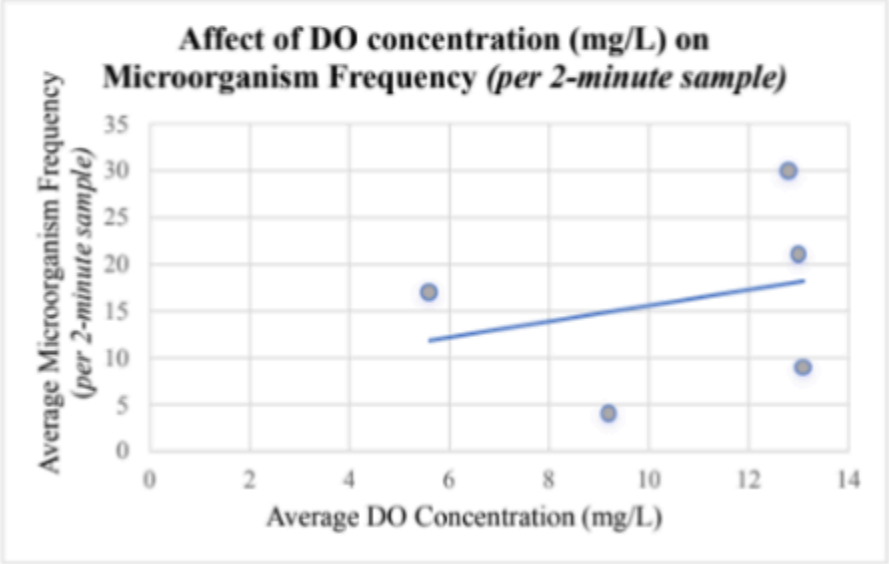
Figure 16: 2004-2012 Change in Toucan Barbet Density of Various Areas ("Santa Lucia Reserve")



Looking at Figure 16, I saw that though the bird population in the primary and secondary forest was declining at a steady rate over the years, the silvapasture was the only area that actually increased in Toucan Barbet bird density. Wanting to know if this was consistent with the trend for average bird density, I graphed the average bird density for the last 8 years in comparison to Figure 17 above. The red represents the Toucan Barbet data and the blue represents the average bird data.

Example 2:

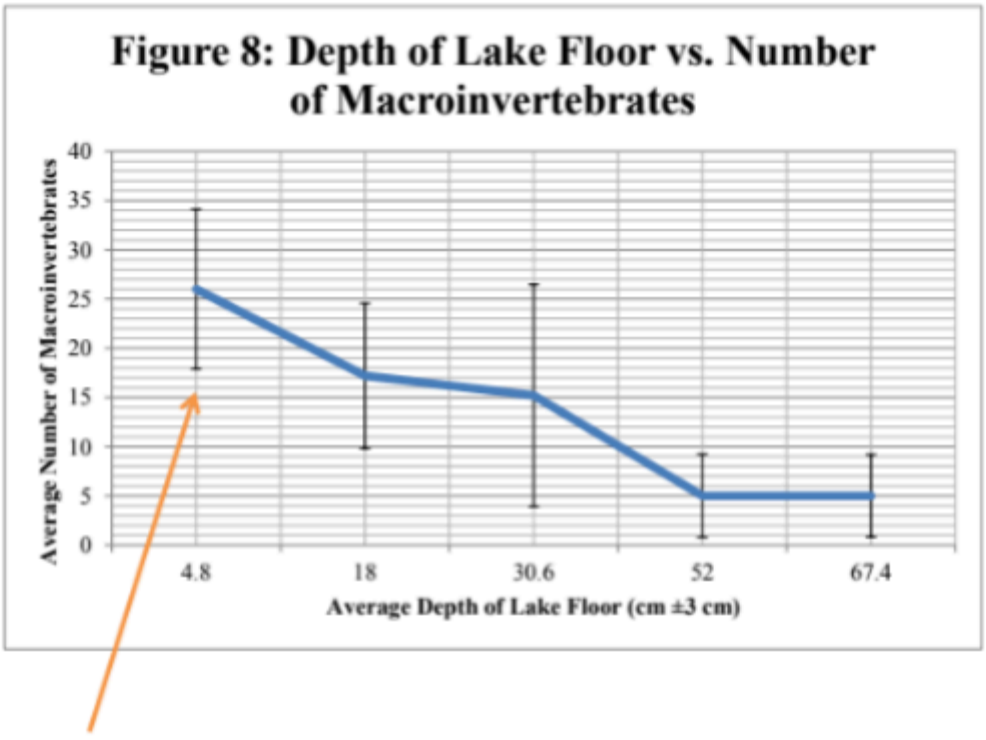
Figure 7b: Graph Comparing Average DO Concentration with Average Microorganism Frequency



Figures 7a and 7b compare the relationship between DO concentration and average microorganism frequency. The scatter-plot graph visually depicts the relationship. Aligning the data shows that a correlating relationship is not present. The line-of-best fit also fails to go through any points because the data is dispersed.

Example 3:

In Figure 8 below, I graphed the mean values I calculated for both the depths and the abundance of macroinvertebrates at each of the five data points of my study. The decrease in abundance of macroinvertebrates is quite visible with this form of presentation. I also graphed the standard deviations of each data point in the form of error bars (*Appendix E*). I took the standard deviation in order to see how far from the mean each average varies, and this number allows me to analyze how accurately my values represent the whole data set.



(Error bars indicate standard deviation)

Graphs

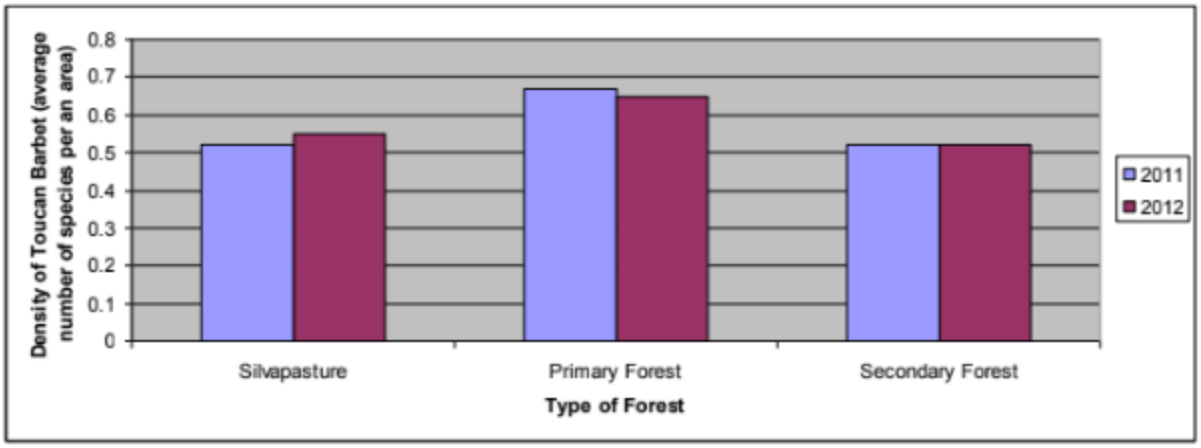
- Be sure you select the correct type of graph for the data you want to visualize!
- For comparative studies, do NOT include information from multiple measures (multiple trials). Graph only the central tendency and spread.
- For correlative studies, add a line of best fit. A line of best fit is only appropriate in comparative studies if both variables are quantitative.
- The figure caption is the graph title. Do not clutter the graph by including both a caption and a title!

- Be sure the caption clearly indicates the type of data being graphed (i.e. mean and SD) on both the X and Y axis.
- Figure captions are placed below the illustration, numbered in sequential order in the sequence in which you refer to the figure within the text of the paper.
- For quantitative data, be sure the axis is labeled with units. Measurement uncertainty indicated (+/-). Measurement precision (how many digits you include in the number) must match the uncertainty of the measurement. So, if your raw data was measured to the tenth of a unit, then your graph axis labeled must also be reported to the tenth of a unit.
- Mark the quantities on both axes and number them at regular intervals. Your axis intervals do not have to be the same on the x and y axis and they do not have to always start at the origin with a value of 0.
- Include a key only if presenting two or more types of data on a single graph (otherwise delete the key).
- The area enclosed by the axes should be roughly square and not disproportionately exaggerated.
- The graph should take up at least ¼ of a page for legibility.
- In general, avoid labeling each point on the graph with a value. That clutters the graph.

Here are some example graphs, with the good and not-so-good indicated:

Example 4:

Figure 15: Comparing 2011 and 2012 Toucan Barbet Data in Various Areas ("Santa Lucia Reserve")



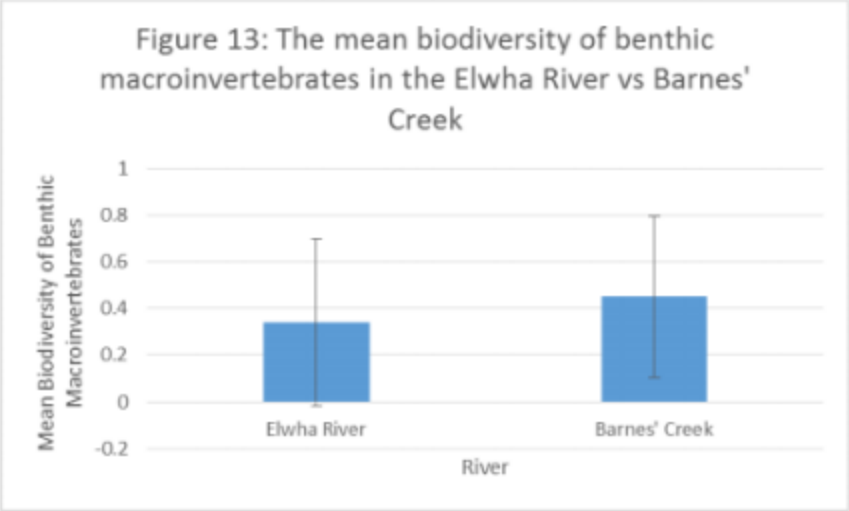
Good

- This is the correct type of graph for the data.
- Only the central tendency is graphed (not each individual data point).
- No line of best fit is included.
- The figure caption is the graph title.
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- Y axis measurement precision matches the uncertainty of the measurement and is consistent (i.e. all values are to the tenths place).
- Axis intervals spaced appropriately.
- There is a key for the two years of data.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Bars are not labeled with individual values.

Not-so-good

- Whenever a mean is graphed, the associated standard deviation should also be graphed (as "error bars").
- The figure caption is a bit vague ("Toucan Barbet *Data*").
- Figure caption is not below the illustration.
- The Y axis unit is unclear. What area? (i.e. per meter? Per hectare?)
- Measurement uncertainty is not indicated (+/-) for the Y axis unit.

Example 5:



Good

- This is the correct type of graph for the data.
- Only the central tendency is graphed (not each individual data point).
- For each mean, the associated standard deviation is also graphed (as “error bars”).
- No line of best fit is included.
- The figure caption is the graph title.
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- Axis intervals spaced appropriately.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Bars are not labeled with individual values.

Not-so-good

- The SD bars should be labeled as such (i.e. in the caption).
- The figure caption is not below the illustration.
- The Y axis unit is unclear.
- Measurement uncertainty is not indicated (+/-) for the Y axis unit.
- Y axis measurement precision is not consistent (i.e. all values are not to the tenths place).

Example 6:

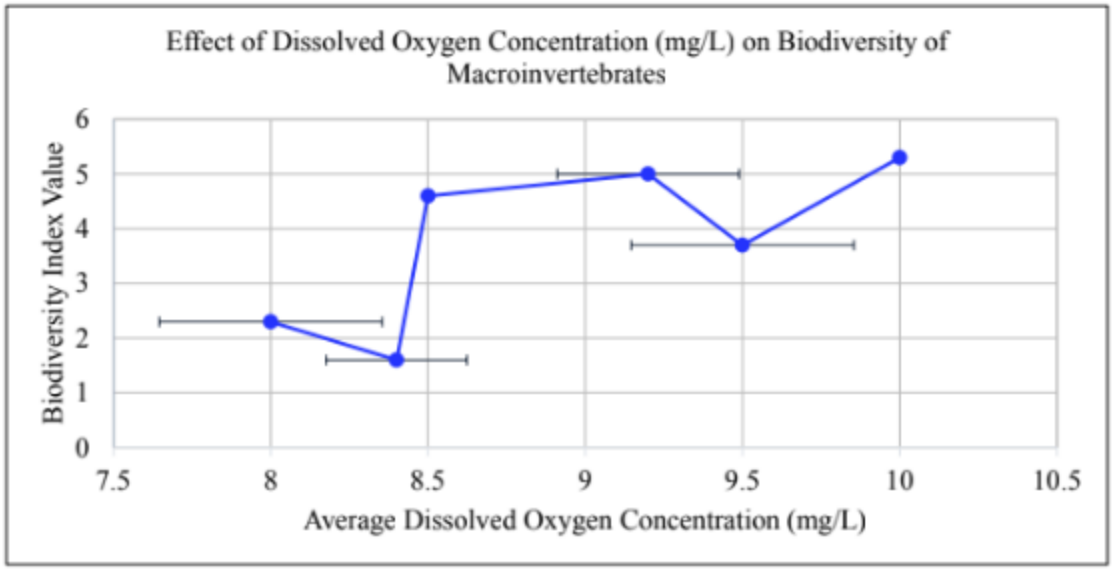


Figure 16: Relationship Between Dissolved Oxygen Content and Biodiversity of Macroinvertebrates (Excel)

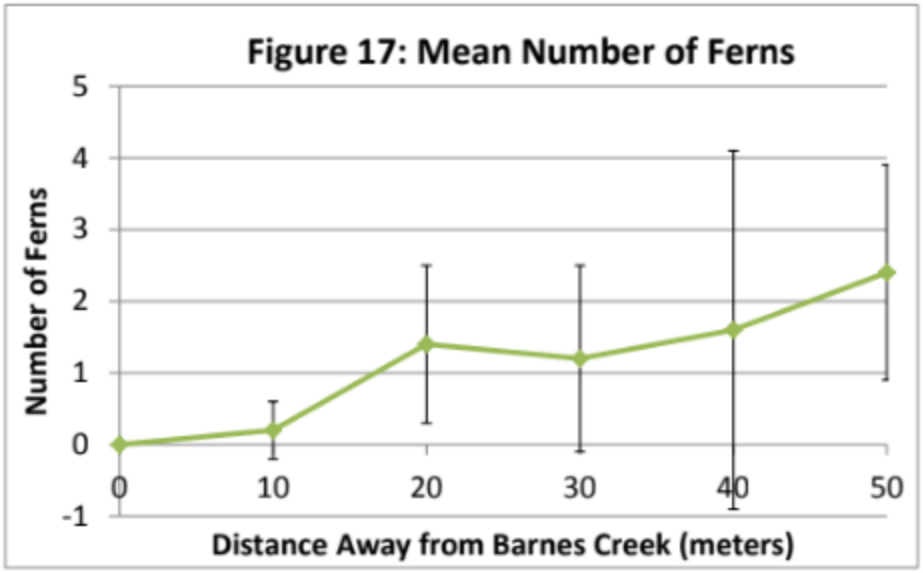
Good

- Only the central tendency is graphed (not each individual data point).
- For each mean, the associated standard deviation is also graphed (as “error bars”).
- The figure caption is below the illustration.
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- Axis intervals spaced appropriately.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Points are not labeled with individual values.

Not-so-good

- This is the correct type of graph for the data, except the points should not be connected dot-to-dot. Rather, there should be a best fit line to illustrate the trend.
- There is a figure caption AND a graph title.
- The SD bars should be labeled as such (i.e. in the caption).
- The Y axis unit is unclear.
- Measurement uncertainty is not indicated (+/-) for the X axis unit.
- X axis measurement precision is not consistent (i.e. all values are not to the tenths place).

Example 7:



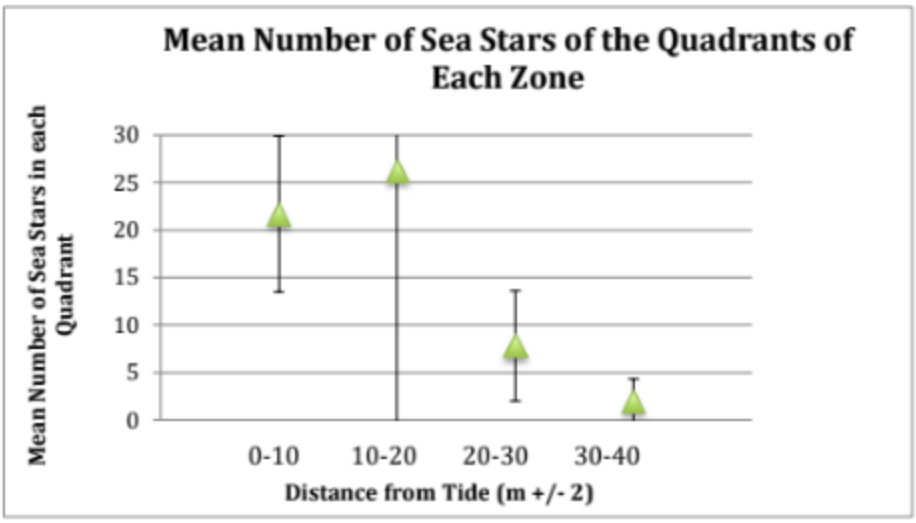
Good

- Only the central tendency is graphed (not each individual data point).
- For each mean, the associated standard deviation is also graphed (as “error bars”).
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- Axis intervals spaced appropriately.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Points are not labeled with individual values.
- Axis measurement precision is consistent (i.e. all values are whole numbers).

Not-so-good

- This is the correct type of graph for the data, except the points should not be connected dot-to-dot. Rather, there should be a best fit line to illustrate the trend.
- The figure caption is not below the illustration.
- The caption is vague, and does not reference the data on both the X and Y axis.
- The SD bars should be labeled as such (i.e. in the caption).
- The Y axis unit is unclear. Ferns per what area?
- Measurement uncertainty is not indicated (+/-) for the X or Y axis unit.

Example 8:



Good

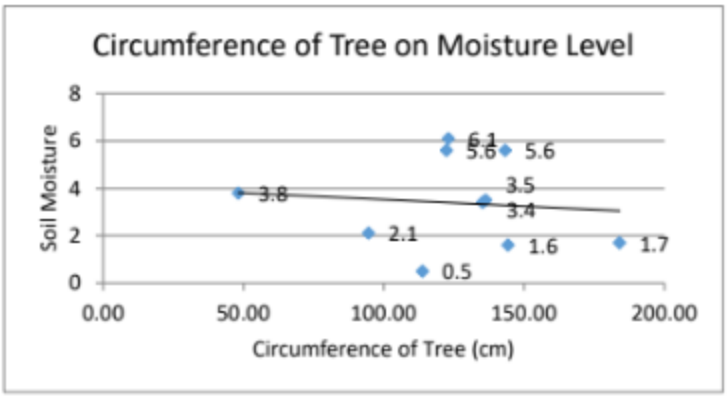
- Only the central tendency is graphed (not each individual data point).
- Axis intervals spaced appropriately.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Points are not labeled with individual values.
- Axis measurement precision is consistent (i.e. all values are whole numbers).

Not-so-good

- For each mean, the associated standard deviation is also graphed (as “error bars"). However, these are not the correct SD values and the 10-20 m value is not fully shown in the graph.
- This is not the correct type of graph for the data. If the X axis values are grouped into ranges (i.e. 0-10), then a histogram is the correct graph to use.
- There is no Figure caption.
- The SD bars should be labeled as such (i.e. in the caption).
- The Y axis unit is unclear. I.e. What is the per area size?
- Measurement uncertainty is not indicated (+/-) for the Y axis unit.

Example 9:

Graph 2: Circumference of Tree on Moisture Level



Good

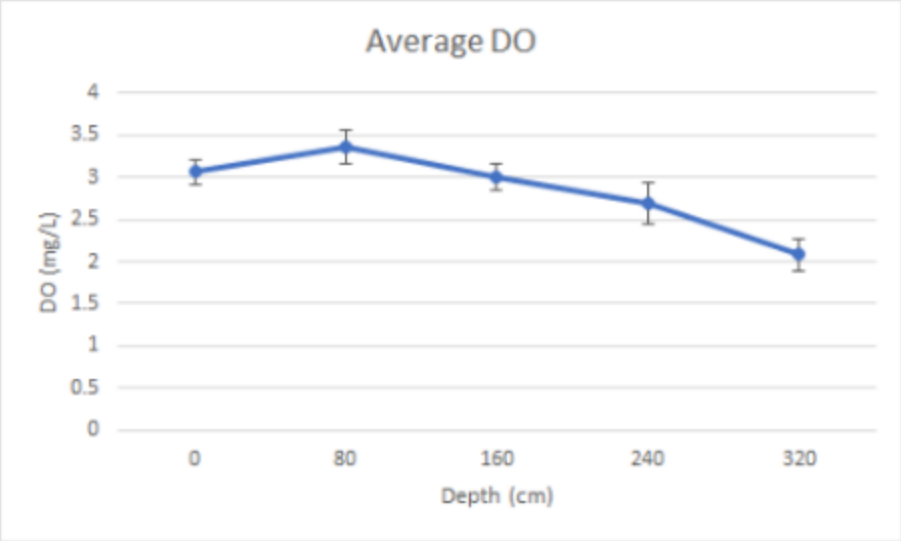
- Axis intervals spaced appropriately.
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.

Not-so-good

- This is the correct type of graph for the data, except the variables are switched. Moisture should be on the X axis and circumference on the Y axis.
- The graph is too small.
- Points are labeled with individual values.
- The figure caption is not below the illustration. Additionally, graphs should be labeled as Figures (not “Graphs.”)
- There is a figure caption AND a graph title.
- The Y axis unit is unclear.
- Measurement uncertainty is not indicated (+/-) for the X or Y axis unit.
- X axis measurement precision is not consistent with the actual precision of the raw data (raw data was collected to the tenths place, graph axis is labeled to the hundredths place).

Example 10:

Figure 1.10: Graph of Average Measurements of Dissolved Oxygen Levels at Different Water Depths



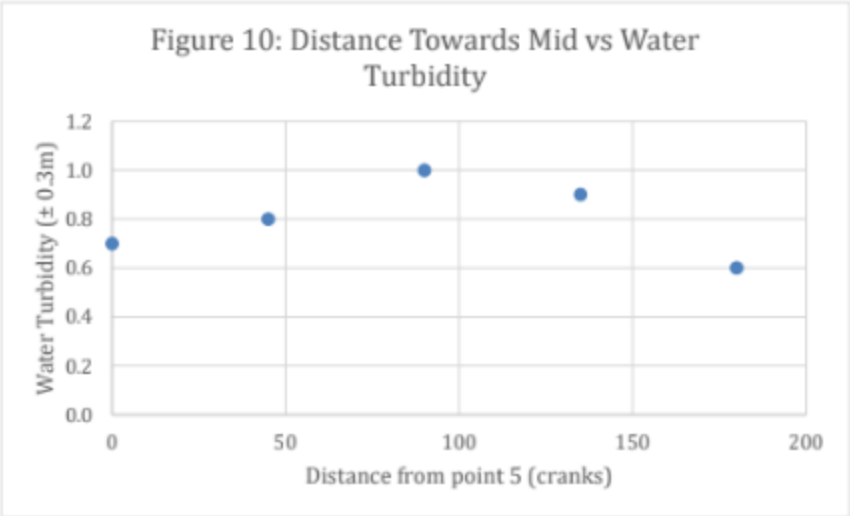
Good

- Only the central tendency is graphed (not each individual data point).
- For each mean, the associated standard deviation is also graphed (as “error bars”).
- Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper.
- Caption references the data on both the X and Y axis.
- The axis units are unclear.
- Axis intervals spaced appropriately.
- There is no key, as one is not needed.
- Axes are not disproportionately exaggerated.
- The graph takes up at least ¼ of a page.
- Points are not labeled with individual values.

Not-so-good

- This is the correct type of graph for the data, except the points should not be connected dot-to-dot. Rather, there should be a best fit line to illustrate the trend.
- The figure caption is not below the illustration.
- There is a figure caption AND a graph title.
- The SD bars should be labeled as such (i.e. in the caption).
- Measurement uncertainty is not indicated (+/-) for the X or Y axis unit.
- Y axis measurement precision is not consistent (i.e. all values are not to the tenths place).

Example 11:



Good	**Not-so-good**
• Figure caption is numbered in sequential order in the sequence in which it is referred to within the text of the paper. • There is not a figure caption AND a graph title. • Axis intervals spaced appropriately. • There is no key, as one is not needed. • Axes are not disproportionately exaggerated. • The graph takes up at least ¼ of a page. • Points are not labeled with individual values. • Axis measurement precision is consistent (i.e. all values are to the tenths place).	• Only the central tendency is graphed (not each individual data point). Except, nowhere on the graph does it say that these are MEAN values! • For each mean, the associated standard deviation should also be graphed (as “error bars”). • This is the correct type of graph for the data, except there should be a best fit line to illustrate the trend. • The figure caption is not below the illustration. • The caption is vague. What does “distance towards mid” mean?. • Measurement uncertainty is not indicated (+/-) for the X or Y axis unit.

Gotelli, Nicholas J., and Aaron M. Ellison. *A Primer of Ecological Statistics*. 2nd ed., Sinauer Associates, Inc., Publishers, 2018.