Julia Vineyard

Week 3 Reading Questions

Worked with: Jessica Bonin

1. Which of the plot types show every data point?
   Cleveland dotplot, scatterplot, pairplot, coplot

2. Which of the plot types show aggregated or summarized data?
   Boxplots, histograms

3. A conditional variable is a third variable upon which the 2 variables being analyzed are dependent upon. Based off of the Zuur coplot example, in my research a conditional variable of vegetation composition and turtle location within the wetland would be the week of year. Both the vegetation structure and location within the available habitat are dependent upon what time of the year the data is collected within.

4. List at least three of the common measures of spread or dispersion that were mentioned in the readings.
   Range, standard deviation, coefficient of variation

5. The range is simply the entire range of values in the sample. This captures all values, including all outliers, as it is from the minimum to the maximum. The standard deviation tells you how tightly the data is centered around the mean. In a normal distribution the standard deviation can tell you approximately what percent of the values are within a certain range. ± 1 standard deviation contains about 95% of the values represented within the dataset and the percent captured increases as more standard deviations are considered.

6. Data exploration can allow you to find errors in your datasets or identify outliers that may not be descriptive of the sample. One way to visualize these errors or outliers is to graph them on a scatterplot or boxplot. A second reason to perform data exploration is that it allows you to confirm that you data are collected in such a way that you can analyze through your desired methods. After exploring your data you may find that the way your data was collected is actually more suitable for analysis through a coplot rather than a scatterplot as you had planned. If you explore your data by plotting it on different graphs you may also find unexpected relationships such as a linear regression or normal distribution.