Hazel Ortiz

ECO 602

September 18, 2021

Week 3 Questions

I worked alone on these questions (to be indicated per instructions)

Q1: Which of the plot types show every data point?

A scatterplot matrix (pair plot) shows every combination of variables.

Q2: Which of the plot types show aggregated or summarized data?

The box plot shows summarized distribution data.

Q3: Explain what a conditional variable means in the context of graphical data exploration.

Based on some of the word usage in the McGarigal reading, specifically the Coplot description, I believe that a conditional variable is equivalent to a dependent variable.

Q4: List at least three of the common measures of spread or dispersion that were mentioned in the readings.

Three common measures of spread of dispersion are range, coefficient of variation, and interquartile range.

Q5: Choose two of the measures in your list and explain how they capture different aspects of the concept of spread.

Measures of spread measure how spread out the data is. Range measures the spread by measuring the minimum and maximum values of a data, which can be found in most graphical displays. There is also interquartile range, as shown in a boxplot, shows the range of values between the 25th and 75th quantiles of data.

Q6: List two of the important reasons to perform data exploration (numerical and/or graphical). For each of the two reasons you identify, describe the quantities, or plots you would use and the insight you would gain.

Data exploration is necessary to address required questions for analyzing data. Screening my data for any extreme values or outliers in my species data is another example of why data exploration is important. I would use a boxplot to screen my species data for any species that have a low or no prevalence in the area that I am looking at so that I can remove the species from my data. If the species has low to no prevalence there will not be enough or any data for the questions that I would be asking of the data. Exploratory data that reveals the relationship between variables and their patterns is also important in creating and appropriate environmental model. I haven’t explored this aspect of my dataset yet, but I would maybe use a scatterplot to explore the relationship of invasive species and climate affected habitats to identify any patterns or any missing information to create a better model.