Noah Tigner

GEOG 495

Exercise 2

1. The typical length of the hydrologic unit along Summit Cr. is between 4 and 10. The shortest is 2, the longest is 28, and the range is 26.
2. The strip plot allows us to see the information along a single axis. All information is visible, but it may be difficult to see it all, due to overlapping boxes. By using the stacking method, the boxes are stacked in a way that allows us to get a better sense of the frequency of any given length. Most lengths sit between 3 and 15, and anything above or below that is less common. 7.5 appears the most frequently. This was not as easy to see in the previous chart.
3. The boxplot shows the quartiles of the data, as well as a single outlier (if the range is at the default). Less information is provided, but the overall data is summarized. Not all observations are shown, but the median and distribution are much easier to see compared to the univariate scatter plot and strip plot.
4. Viewing a histogram of the "Yes" variable, we are able to see the distribution of the frequency that communes preferred to join the EU. Values between 45 and 50 were most likely to occur, although values between 25 and 75 were observed.
5. The stripchart shows the "clustering" of data between 40 and 60, but it is difficult to see exactly which values are most frequent, due to overlap. The histogram makes it much easier to compare the frequencies of different 'Yes' value ranges.
6. When setting the histogram's breaks attribute to 20, it becomes apparent that there were no observations between 32 and 34, 58 and 60, 66 and 70. Multiple breaks between 40 and 60 have the same frequency. When set to 40, we see that around 54 is very frequent, but the values lower and higher are surprisingly much less frequent.
7. The density line gives a smoother overview of the data than the histogram, and is less dependent on how the plot is set up (i.e. bin width). However, the histogram loses the least amount of information about the individual values of the variable (more information is lost as bin width increases & number of bins decreases).
8. The composite plot contains the histogram, the density line, and the rug. The rug function creates tick marks for each observation below the x axis. This plot combines the

Noah Tigner

GEOG 495

Exercise 2

advantages of the histogram and the density plot, while still showing each individual value of the “Yes” variable.