1. Scatterplot, Cleveland dotplot, QQ plot and Coplot all show every data point.
2. Histogram and boxplots show aggregated or summarized data
3. A conditional variable in terms of graphical data exploration is a variable plotted against the independent and dependent variables separately which allows for a determination of the relationship between them. Conditional variables are necessary because the relationship between independent and dependent variables can be masked by another variable.
4. Common measures of spread or dispersion include variance, standard deviation, coefficient of variation, median absolute deviation, range, and interquartile range
5. The sample standard deviation(SD) is the root mean squared deviation from the mean and the square root of the variance. Unlike variance which has squared units, SD has the same units as the measured sample variable and is therefore interpretable. In a normal distribution, approximately 68%, 95% and over 99% of the variables are within ±1 SD, ±2 SD and ±3 SD. The range yields the absolute spread of values from the minimum to maximum.
6. Two of important reasons to preform numerical and/ or graphical data exploration is to observe patterns within the data which can generate new questions and to determine if there are outlies within the data that could have disproportionate effects on the analysis. Using the Oregon birds data set I could perform conditional histograms for the elevation of sampling sites conditional on precipitation level. If there is a clear difference in shape or center of the distribution, then an investigation of the effect of precipitation would be warranted. I could preform a box plot on the distribution of slopes at each sampling site and determine where the data skews and outlies are to determine where my focus in explanation needs be placed and what values can be less emphasized.