**Ggplot tutorial**

**Haldre Rogers**

**EEOB590**

1. Grammar of graphics
   1. Required
      1. a data set
      2. a set of geoms—visual marks that represent data points,
      3. aesthetic properties of the geom like size, color, and x and y locations. These are referred to as "mappings". Aesthestics are “something you can see”.
   2. additional elements with sensible defaults, so not required
      1. coordinate functions
      2. scale functions
      3. facet functions
      4. theme functions
2. Start with ggplot(data, aes()) + then add a new layer to a plot with a geom\_\*() or stat\_\*() function. Each layer includes a geom, a set of aesthetic mappings, and a default stat and position adjustment.
3. Data set
   1. Use whatever subset you want. Need to have a column for every variable you want to plot.
4. Geoms or stat
   1. Stat vs geom
      1. Some plots visualize a transformation of the original data set. Use a stat to choose a common transformation to visualize.
      2. stat functions and geom functions both combine a stat with a geom to make a layer, i.e. stat\_bin(geom="bar") does the same as geom\_bar(stat="bin")
   2. Lots of geoms to choose from.
      1. Common options: geom\_histogram(), geom\_point(), geom\_boxplot(), geom\_line(), geom\_smooth()
   3. Arguments within geoms or stats
      1. Position ="", choose dodge, jitter, nudge, stack, fill (stacked with all heights equalized)
5. Aesthetics
   1. Tells ggplot what you want to be the x and y coordinates and what you want to be represented by color or size…. This is where you identify the columns in your dataset that you want to be on your graph.
6. Coordinate system
   1. Not usually something you need to adjust unless you want to flip x&y axes, or put into polar (circular) coordinates.
7. Scale functions
   1. Used to change visual representation of the aesthetic functions
      1. Function is always "scale" underscore, then whatever aesthetic you want to adjust, underscore, then the pre-packaged scale to use (e.g. manual, continuous, discrete, identity), then the parentheses .
      2. Inside the parentheses, have
         1. values (which are the names of the values/levels for that particular aesthetic)
         2. limits (range of values to include in mapping)
         3. breaks (breaks to use in the legend/axis)
         4. name (title to use in legend/axis)
         5. labels (labels to use in legend/axis)
8. Facet functions
   1. Facet\_grid and facet\_wrap - used for divide a plot into subplots based on the values of one or more discrete variables.
9. Theme functions
   1. Make the graph look pretty in terms of axes, font etc.