MPA 634  
Data Science and R for Administrators  
Homework #2

Aesthetic Mappings and Facets

1. Problems 1 – 6 from Exercises 3.3.1

3.3.1:

1.) *What’s wrong with this code? Why are the points not blue?*

The problem with the code is that the color argument is *inside* the aes function, which is why the points aren’t blue. To correctly change the color of the graph, the color needs to be defined *outside* of the aes function. For example:

ggplot(data = mpg) +

geom\_point(mapping = aes(x = displ, y = hwy), color = "blue")

2.) *Which variables in mpg are categorical? Which variables are continuous? (Hint: type ?mpg to read the documentation for the dataset). How can you see this information when you run mpg?*

Categorical Variables: manufacturer, model, year, cyl, trans, drv, fl, class

Continuous Variables: displ, cty, hwy (displ could be continuous since the engine displacement can theoretically be any volume, and doesn’t necessarily *have* to be certain values)

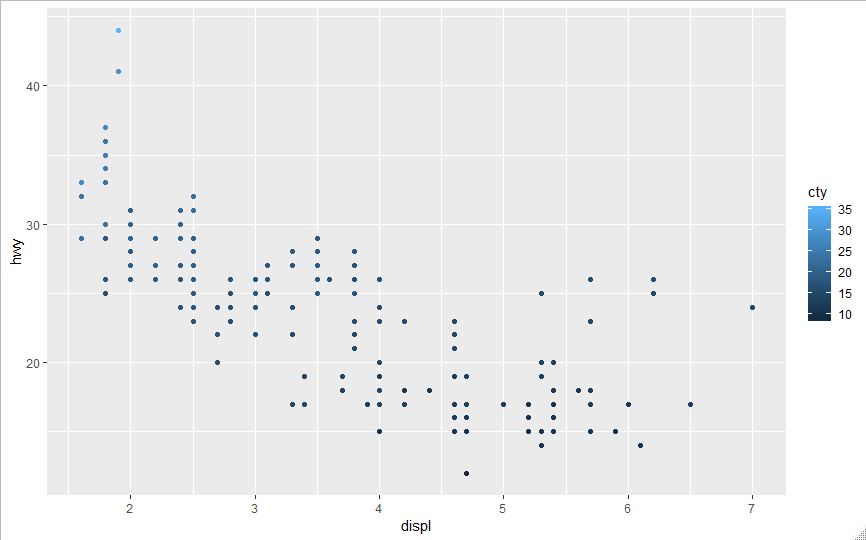
You can see this information by clicking on the data ‘mpg’ in the data window and the table of information will appear. You can also see it in the console when you load the data running the code ‘data(mpg)’.

3. ) Map a continuous variable to color, size, and shape. How do these aesthetics behave differently for categorical vs. continuous variables?

Continuous Variable, cty, to color:

#Map a continuous variable to color

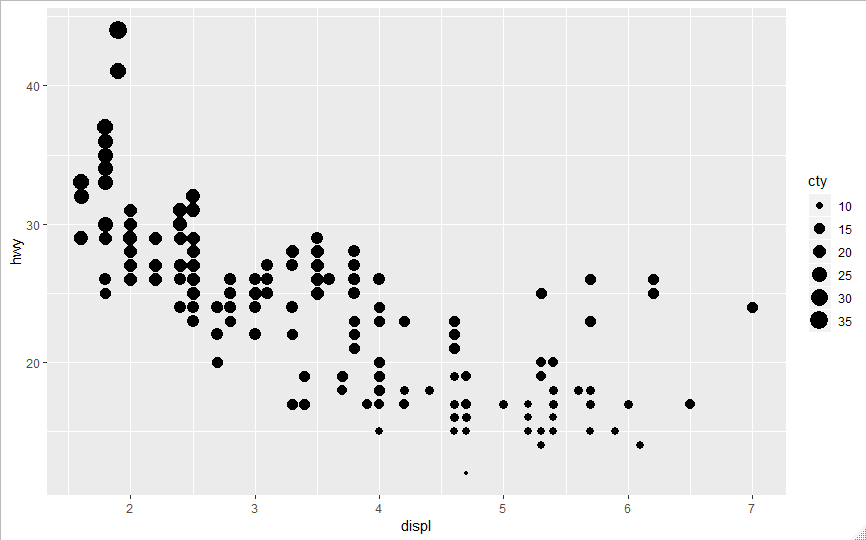
ggplot(data = mpg, mapping = aes(x = displ, y = hwy, color = cty)) +

 geom\_point()

Continuous Variable, cty, to size:

#Map a continuous variable to size

ggplot(data = mpg, mapping = aes(x = displ, y = hwy, size= cty)) +

geom\_point()

Continuous Variable, cty, to shape:

#Map a continuous variable to shape

ggplot(data = mpg, mapping = aes(x = displ, y = hwy, shape = cty)) +

geom\_point()

Error: A continuous variable can not be mapped to shape

The aesthetics of continuous variables behave differently from the aesthetics of categorical variables because they are not discrete variables. In other words, since there are infinite values possible for a continuous variable, e.g., 1.0, 1.001, 1.000001 etc., the aesthetics behave more as a gradient of colors or sizes etc. to reflect the different potential values of a continuous variable. For example, using cty to define the color for the points in the first graph results in different shades of blue to show the difference in cty values, but if cty was a categorical variable, then entirely different and discrete colors would have been used for the different points. Furthermore, for aesthetics that are not continuous, such as shape, continuous variables cannot be mapped to that aesthetic.

4.) What happens if you map the same variable to multiple aesthetics?

5.) What does the stroke aesthetic do? What shapes does it work with? (Hint: use ?geom\_point)

6.) What happens if you map an aesthetic to something other than a variable name, like aes(colour = displ < 5)? Note, you’ll also need to specify x and y.

1. Problem 4 from Exercises 3.5.1

3.5.1:

4.) Take the first faceted plot in this section:

ggplot(data = mpg) +

geom\_point(mapping = aes(x = displ, y = hwy)) +

facet\_wrap(~ class, nrow = 2)

What are the advantages to using faceting instead of the colour aesthetic? What are the disadvantages? How might the balance change if you had a larger dataset?