Aesthetic Mappings in ggplot2

The goal of this exercise is to practice using aesthetic mappings in **ggplot2**. You will not need to edit or improve the graphs. The goal here is to simply practice assigning variables to graphical parameters. You will work with editing and improving graphs in the next assignment.

Generate code to produce the graphs described. Note that you will be using data sets included with the **ggplot2** package, so you will not need to download any data. Deliver the results as an HTML webpage generated from an R Markdown file. Use headers to label each graph result. Make sure to include both the code and the resulting graphs on the webpage.

In this first section, you will produce graphs using the data provided in the **midwest** data frame included in the **ggplot2** package. This data frame contains demographic data for counties in the Midwest states of Illinois, Indiana, Michigan, Ohio, and Wisconsin. You should save the data set to a new data frame for use in your code.

Graph 1: Create a histogram of the percent of the population below poverty ("percbelowpoverty") using a binwidth of 5. (2 Points)

Graph 2: Create a kernel density plot of the percent of the population below poverty ("percbelowpoverty") and include a fill color of your choosing. (2 Points)

Graph 3: Create a combined histogram and kernel density plot of the percent of the population below poverty ("percbelowpoverty"). Make sure to define the y-axis as density. For the histogram, use a binwidth of 5, a color of your choosing, and a fill of you choosing. For the kernel density component, set a color, line weight, fill color, and transparency. (2 Points)

Graph 4: Create a kernel density plot of the percent of the population below poverty ("percbelowpoverty") with separate curves for each state. Assign density to the y-axis and use a fill color to differentiate the states. Also provide a transparency so that curves can be more easily visualized where they overlap. (2 Points)

Graph 5: Create a kernel density plot with the same settings as Graph 4; however, only include curves for the states of Indiana and Michigan. (2 Points)

Graph 6: Create a kernel density plot of percent of the population with a college degree ("percollege") with density mapped to the y-axis and states differentiated using a fill color. Also provide a transparency so that curves can be more easily visualized where they overlap. (2 Points)

Graph 7: Create a kernel density plot with the same settings as Graph 6; however, only include curves for the states of Indiana and Michigan. (2 Points)

Graph 8: Create a box plot of percent of the population with a college degree ("precollege") with different plots for each state. Also, differentiate the states using a fill color. (2 Points)

Graph 9: Create a combined violin and box plot of percent of the population with a college degree ("percollege") with different plots for each state. Differentiate the states using a fill applied to the violin plot. Alter the box plots so that they are all the same color (fill not used to differentiate states) and fit within the associated violin plot. (2 Points)

Graph 10: Create a bar graph of mean percent of the population with professional employment ("percprof") by state. You will need to summarize the county-level data by state to obtain the state means. Within geom_bar() you will need to set stat equal to "identity." (2 Points)

Graph 11: Create a scatter plot of percent of the population with a college degree ("percollege") mapped to the x-axis and percent of the population below the poverty line ("percbelowpoverty") mapped to the y-axis. (2 Points)

Graph 12: Add a loess curve to the graph created in 12. (2 Points)

Graph 13: Create a scatter plot with percent of the population with a college education ("precollege") mapped to the x-axis, percent of the population below the poverty line ("percbelowpoverty") mapped to the point color and population density ("popdensity") mapped to the point size. (2 Points)

Now, you will switch to the **economics** data frame included with **ggplot2**. These data represent US economic time series data from http://research.stlouisfed.org/fred2. You should generate a new data frame from these data.

Graph 14: Create a new variable of unemployment rate using the number of unemployed people ("unemploy") and the total population ("pop"). Create a time series line graph with the date ("date") mapped to the x-axis and your new unemployment rate variable mapped to the y-axis. (2 Points)

For the last set of graphs, you will use the **mpg** data also included with **ggplot2**, which contains a subset of the fuel economy data that the EPA makes available on http://fueleconomy.gov and contains only vehicle models which had a new release every year between 1999 and 2008.

Graph 15: Create a box plot with the class of the vehicles ("class") mapped to the x-axis and fill color and the highway fuel economy ("hwy") mapped to the y-axis. (2 Points)

Graph 16: Reproduce the box plot created in Graph 15, but only include data from 2008. (2 Points)

Graph 17: Generate a kernel density plot of city fuel economy ("cty"). Make sure to map density to the y-axis. To differentiate the model year, map the year ("year") to the fill color. Note that you will need to define it as a factor. Also, provide a transparency so that data can be visualized in overlapping areas. (2 Points)

Graph 18: Use just models from 2008, create a scatter plot with city fuel economy ("cty") mapped to the x-axis and highway fuel economy ("hwy") mapped to the y-axis. Also, add a loess curve. (2 Points)

Graph 19: Create a box plot for just models from 2008 to compare highway fuel economy ("hwy") between different drive systems (4-wheel, front-wheel, rear-wheel) using the "drv" field. Assign "drv" to the x-axis and fill color. Map the highway fuel economy ("hwy") to the y-axis. (2 Points)

Graph 20: Add a violin plot to the graph generated in Graph 19 to produce a combined violin and box plot. The fill color for the violin plot should differentiate the drive type ("drv"). The box plots should all have the same fill color and should fit within the associated violin plot. (2 Points)