

Graph Editing with ggplot2

In the last exercise, you produced a wide variety of graphs and practiced using aesthetic mappings in **ggplot2**. Expanding upon this last assignment, you will now edit graphs to obtain a final product suitable for inclusion in a document, presentation, or paper.

Generate code to produce the graphs described. Note that you will be using a data set included with the **ggplot2** package, a data set from the **gapminder** package, and two provided data sets. 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.

Graph 1: Unemployment Rate Time Series (8 Points)

For this graph you will use 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. Create an edited version of Graph 14 from the prior assignment. You will need to create a new variable of percent unemployed 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 percent unemployment variable mapped to the y-axis. Edit the graph to meet the following additional requirements:

- Y-axis shows percent unemployment as opposed to unemployment rate. Breaks should be defined for 1%, 2%, 3%, 4%, 5%, and 6% with percent signs included with the labels (for example, “1%”). The y-axis should have limits of 1% to 6%. Use the `expand` argument to remove the gaps between the data and y-axis.
- On the x-axis, the limits should be 1965 to 2015 with breaks and labels every five years (1965, 1970, 1975.....) Only the year, not the month and day, should be included as part of the label. You will need to use `scale_x_date()`.
- Change the line color and line width for the time series.
- Make sure to include appropriate axis labels and a main title.
- Change the default theme to a different theme of your choosing.
- Change the axes titles, axes labels, and main title font size.

Graph 2: Monthly Runoff Box Plot (8 Points)

Make a box plot showing the distribution of monthly runoff in a stream using the provided **runoff_data_by_month.csv** data set. These data are for a specific stream in the Fernow Experimental Forest near Parsons, WV. The runoff measurements have been aggregated by month and are in mm units. Edit the graph to meet the following additional requirements:

- Month plotted to the x-axis and fill color for the box plots.
- Runoff (“runoff”) mapped to the y-axis.
- The months should be re-ordered by calendar progression (January, February, March.....) as opposed to alphabetically.
- The y-axis title should be “Runoff [mm]” and the x-axis title should be “Month”.
- Provide a descriptive main title.
- Remove the legend (since the x-axis labels provide the month names, you don’t need an additional legend for the fill color.)
- Change the theme to a theme of your choosing.
- Change the axes titles, axes labels, and main title font size.

- Change the y-axis scale to include the following breaks (0, 50, 100, 150, 200, 250, 300), limits of 0 to 300, and labels including the unit of measurement (for example, “50 mm”).

Graph 3: Gapminder Scatter Plot (8 Points)

You will now create a scatter plot from the Gapminder data available in R. You can obtain this dataset by loading in the **gapminder** package then calling the **gapminder** data object or copying it to a new data frame. Here is a link to the source of these data: <https://www.gapminder.org/>. This data set includes socioeconomic data aggregated to countries. The scatter plot should only include data from 2007. The GDP per capita (“gdpPercap”) should be mapped to the x-axis, life expectancy (“lifeExp”) to the y-axis, the continent (“continent”) to color, and population (“pop”) to size. The graph should meet the following additional requirements:

- The x-axis title should be “GDP per Capita” and the y-axis title should be “Life Expectancy in Years”. In the legend, the continent data should be titled “Continent” and the population data should be titled “Population.”
- Add a descriptive main title.
- Change the theme to a theme of your choosing.
- Change the axes titles, axes labels, and main title font size.
- Define the following breaks for the x-axis: 0, 10000, 20000, 30000, 40000, 50000. Set the limits to 0 to 500000. Include comma separators in the labels.

Graph 4: Faceted Gapminder Scatter Plot (8 Points)

You will now expand upon Graph 3 to include faceting. Extract data for just Africa and Europe and also for just 1977 and 2007. In the facet grid, the columns should be defined using the continents and the rows should be defined using the years. GDP per capita (“gdpPercap”) should be mapped to the x-axis, life expectancy (“lifeExp”) to the y-axis, and population (“pop”) to size. Do not map the continent to color, since you will use continents to define facets. The graph should meet the following requirements:

- The x-axis label should be “GDP per Capita” and the y-axis label should be “Life Expectancy in Years”. In the legend, the population data should be titled “Population.”
- Add a descriptive main title.
- Change the theme to a theme of your choosing.
- Change the axes titles, axes labels, and main title font size.
- Define the following breaks for the x-axis: 0, 10000, 20000, 30000, 40000, 50000. Set the limits to 0 to 500000. Include comma separators in the labels.
- The size aesthetic should include the following breaks: 5e+04, 1e+05, 1e+06, 1e+07, 1e+08, 1.5e+08. The limits should be 5e+04 to 1.5e+08. Use the following labels: "50 Thousand", "100 Thousand", "1 Million", "10 Million", "100 Million", "150 Million".

Graph 5: NDVI Box Plot (8 Points)

You will now make a new box plot of NDVI (normalized difference vegetation index) by land cover type from the provided **mine_classification_with_lidar.csv** data set.

- The land cover type (“class”) should be mapped to the x-axis and fill color. NDVI (“ndvi”) should be mapped to the y-axis.

- Remove the legend (since the x-axis labels provide the land cover type names, you don't need an additional legend for the fill color.)
- Define colors to use to represent each class. Use the following labels: "Barren", "Forest", "Herbaceous", "Woodlands", "Water".
- Apply the following labels to the x-axis: "Barren", "Forest", "Herbaceous", "Woodlands", "Water".
- Provide a main title and appropriate x-axis and y-axis titles.
- Change the theme to a theme of your choosing.
- Change the axes titles, axes labels, and main title font size.