

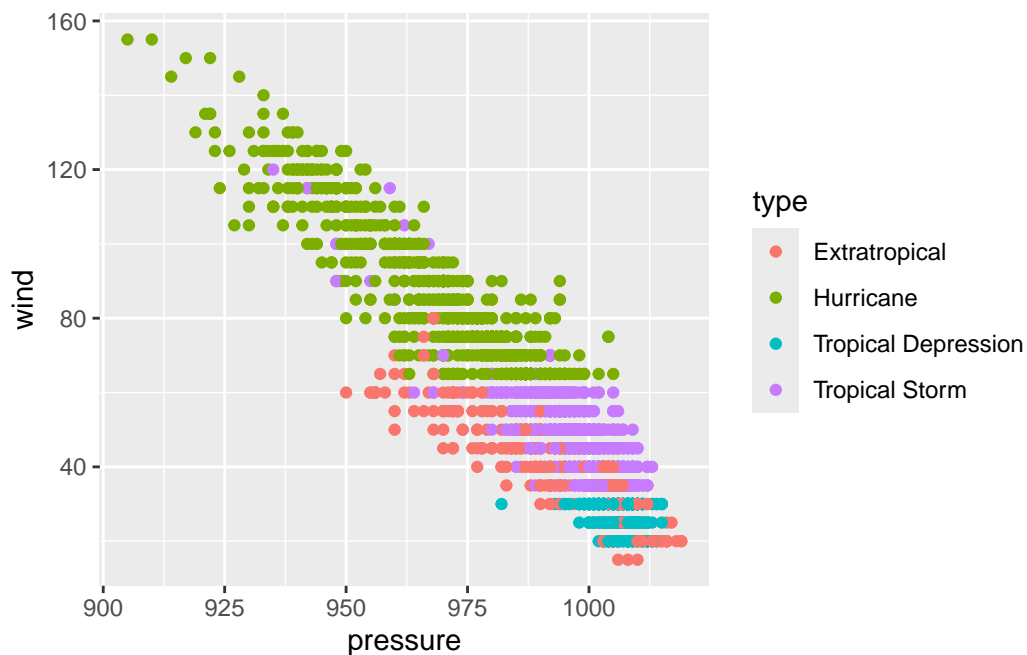
ggformula Worksheet Answers

Casey Crary and Chelsea Wang

Aesthetics (color, size, shape, alpha)

Using the `storms` data frame from the `nasaweather` package, create a scatterplot between `wind` and `pressure`, with color being used to distinguish the `type` of storm. Try jittering the points, making them transparent, and adding another aesthetic too.

```
gf_point(wind ~ pressure, color = ~ type, data = storms)
```

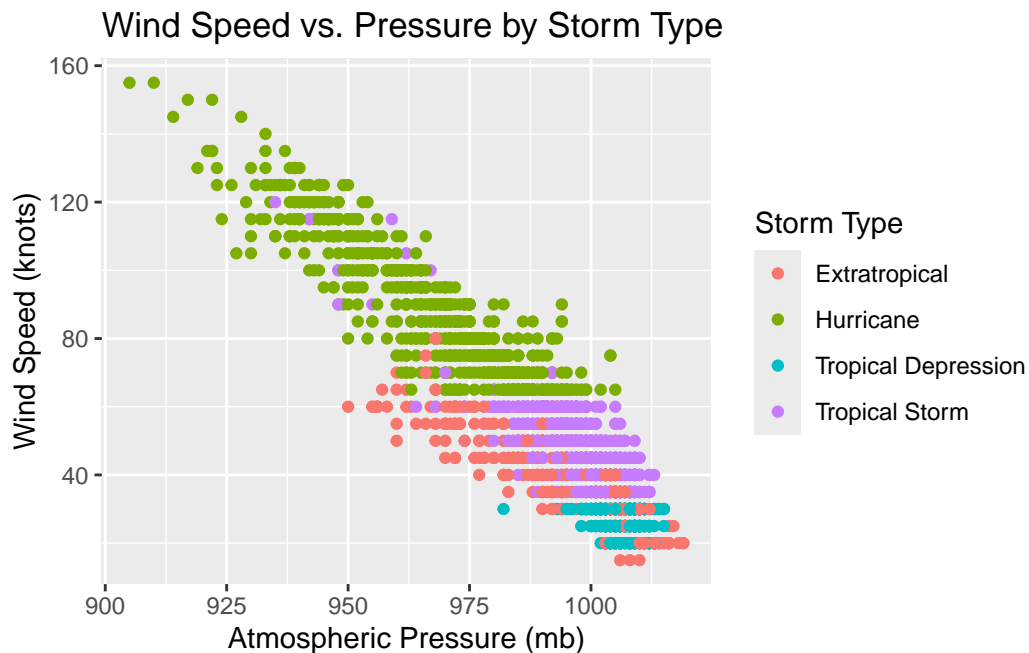


Axis Labels

Save your above plot, then add appropriate axis labels

```
storms_plot <- gf_point(wind ~ pressure, color = ~ type, data = storms)

storms_plot |>
  gf_labs(
    x = "Atmospheric Pressure (mb)",
    y = "Wind Speed (knots)",
    title = "Wind Speed vs. Pressure by Storm Type",
    color = "Storm Type"
  )
```

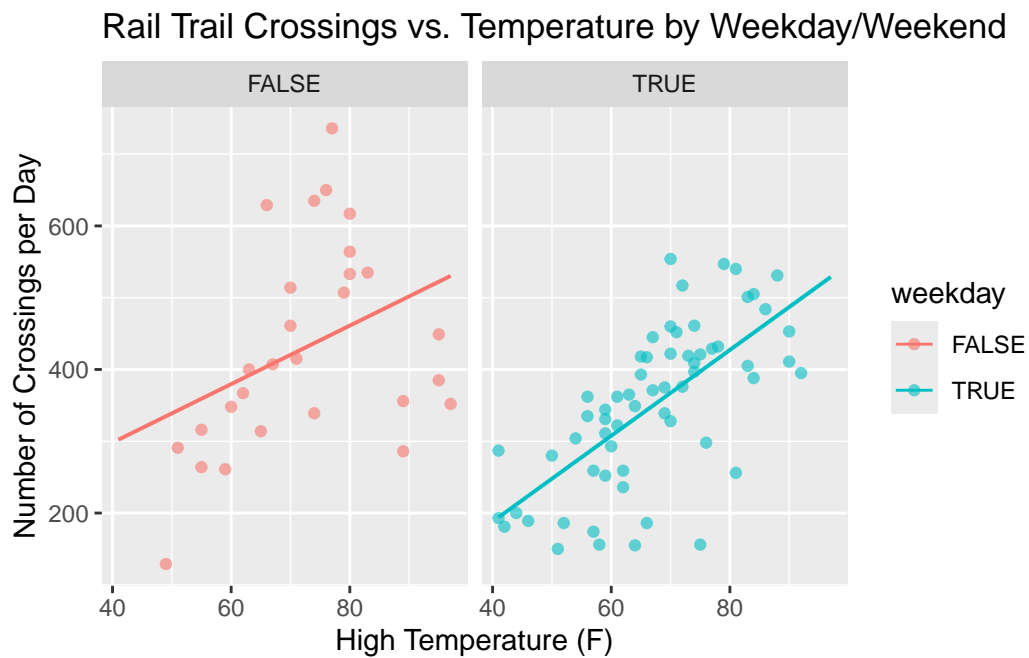


Facets

The RailTrail data set from the mosaicData package describes the usage of a rail trail in Western Massachusetts. Use these data to answer the following questions. (Hint: For information on how to use faceting and add regression lines, see MDSR, Ch 3.2.2)

1. Create a scatterplot of the number of crossings per day volume against the high temperature that day
2. Separate your plot into facets by weekday (an indicator of weekend/holiday vs. weekday)
3. Add regression lines to the two facets

```
gf_point(volume ~ hightemp | weekday, data = RailTrail, color = ~weekday, alpha = 0.6) |>
  gf_lm() %>%
  gf_labs(
    x = "High Temperature (F)",
    y = "Number of Crossings per Day",
    title = "Rail Trail Crossings vs. Temperature by Weekday/Weekend"
  )
```



Geometries (boxplot)

Using the `iris` dataset make a boxplot of `Sepal.Length` across the different species. What can you conclude?

```
#Specified binwidth can be altered
gf_boxplot(Sepal.Length ~ Species, data = iris, fill = ~Species) |>
  gf_labs(
    x = "Species",
    y = "Sepal Length (cm)",
    title = "Distribution of Sepal Length by Iris Species"
  )
```

