ggplot Scatter Plot Three Continuous Variables and Multiple Categorical Variables

Fan Wang

2021-03-23

Contents

1	ggp	lot Scatter Plot	1
	1.1	Three Continuous Variables and Two Categorical Variables	1

1 ggplot Scatter Plot

Go to the RMD, R, PDF, or HTML version of this file. Go back to fan's REconTools Package, R Code Examples Repository (bookdown site), or Intro Stats with R Repository (bookdown site).

1.1 Three Continuous Variables and Two Categorical Variables

We will generate a graph that is very similar to the graph shown for fs_tib_factors, with the addition that scatter color and shape will be for two separate variables, and with the addition that scatter size will be for an additional continuous variable.

We have three continuous variables:

- 1. y-axis: time for 1/4 Miles (QSEC)
- 2. x-axis: horsepower
- 3. scatter-size: miles per gallon (mpg)

We have two categorical ariables:

- 1. color: vs engine shape (vshaped or straight)
- 2. shape: am shift type (auto or manual)

First, Load in the mtcars dataset and convert to categorical variables to factor with labels.

Second, generate the core graph, a scatterplot with a nonlinear trendline. Note that in the example below color and shape only apply to the jitter scatter, but not the trendline graph.

```
# Graphing
plt_mtcars_scatter <-
    ggplot(tb_mtcars, aes(x=hp, y=qsec)) +
    geom_jitter(aes(size=mpg, colour=vs, shape=am), width = 0.15) +
    geom_smooth(span = 0.50, se=FALSE) +
    theme_bw()</pre>
```

Third, control Color and Shape Information. There will be two colors and two shapes. See all shape listing.

```
# Color controls
ar_st_colors <- c("#33cc33", "#F8766D")
ar_st_colors_label <- c("v-shaped", "straight")
fl_legend_color_symbol_size <- 5
st_leg_color_lab <- "Engine-Shape"
# Shape controls
ar_it_shapes <- c(9, 15)
ar_st_shapes_label <- c("auto", "manuel")
fl_legend_shape_symbol_size <- 5
st_leg_shape_lab <- "Transmission"</pre>
```

Fourth, control the size of the scatter, which will be the MPG variable.

```
# Control scatter point size
fl_min_size <- 3
fl_max_size <- 6
ar_size_range <- c(fl_min_size, fl_max_size)
st_leg_size_lab <- "MPG"</pre>
```

Fifth, control graph strings.

Sixth, combine graphical components.

Eight, replace default legends titles for color, shape and size.

```
# replace the default labels for each legend segment
plt_mtcars_scatter <- plt_mtcars_scatter +
  labs(colour = st_leg_color_lab,</pre>
```

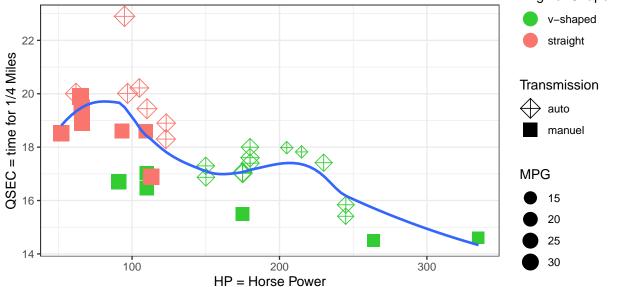
```
shape = st_leg_shape_lab,
size = st_leg_size_lab)
```

Ninth, additional controls for the graph.

```
# Control the order of legend display
# Show color, show shape, then show size.
plt_mtcars_scatter <- plt_mtcars_scatter + guides(
    colour = guide_legend(order = 1, override.aes = list(size = fl_legend_color_symbol_size)),
    shape = guide_legend(order = 2, override.aes = list(size = fl_legend_shape_symbol_size)),
    size = guide_legend(order = 3))
# show
print(plt_mtcars_scatter)</pre>
```

Distribution of HP and QSEC from mtcars

https://fanwangecon.github.io/R4Econ/tabgraph/ggscatter/htmlpdfr/fs_ggscatter_acts_mdiag.html



mtcars dataset, https://fanwangecon.github.io/R4Econ/