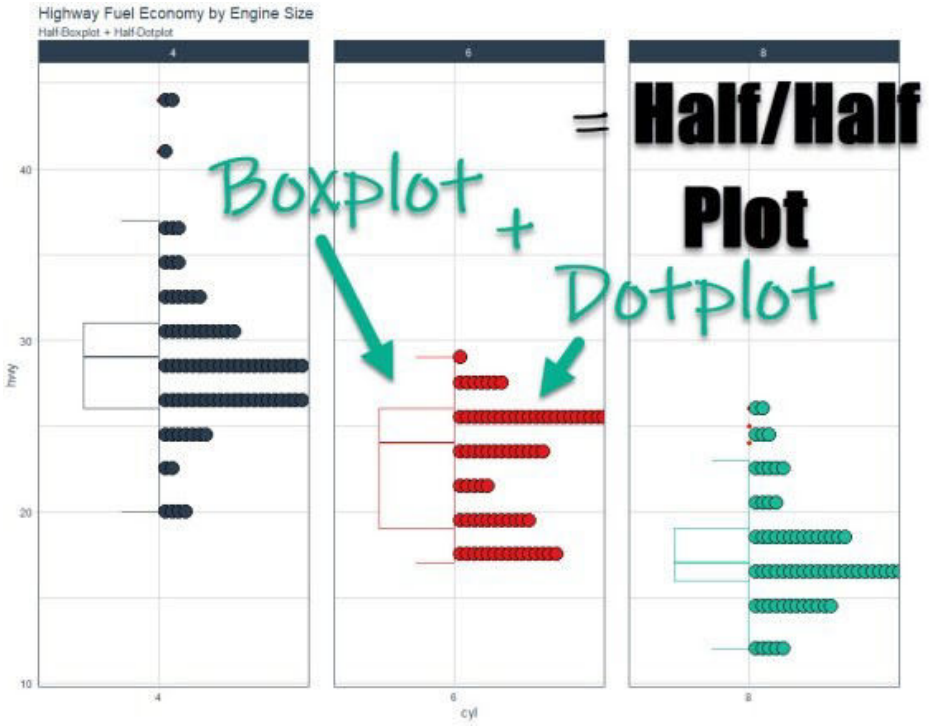
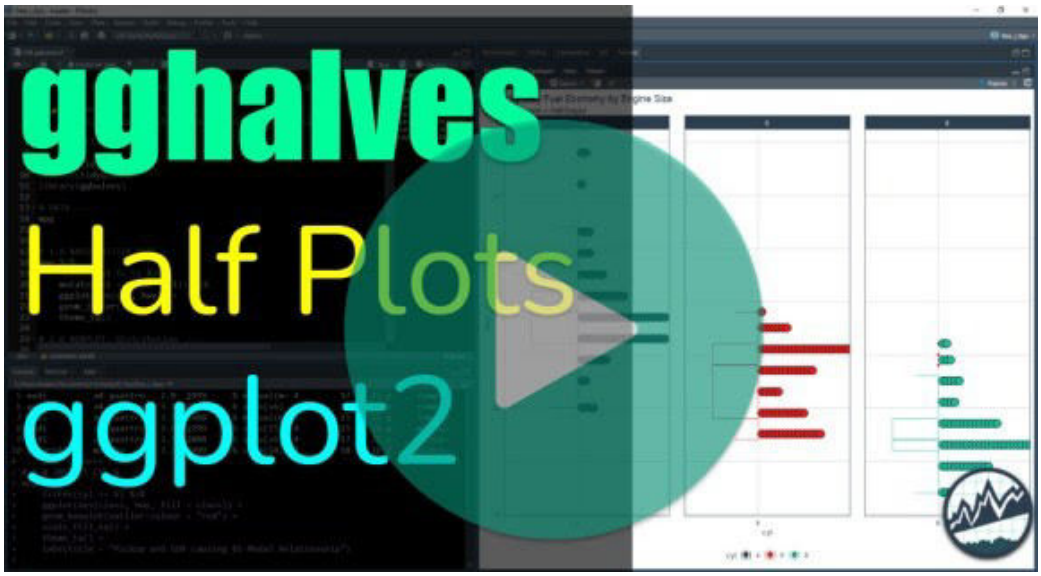
What is gghalves?

`gghalves` is a new R package that makes it easy to compose your own half-plots using `ggplot2`.



gghalves Video Tutorial
For those that prefer Full YouTube Video Tutorials.

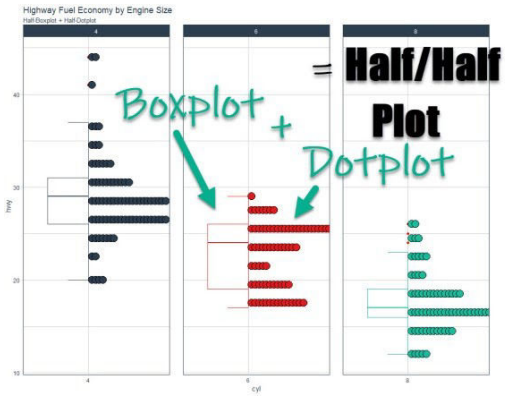
Learn how to use `gghalves` in our free 8-minute YouTube video.



(Click image to play tutorial)

Watch our full YouTube Tutorial

What are Half Plots?
Combining two plots side-by-side.



Half/Half Plots are a way to showcase two plots side-by-side. Here's a common example:

1. Showing a **Boxplot** to identify outliers and quantiles
2. Showing a **Dotplot** to identify distribution

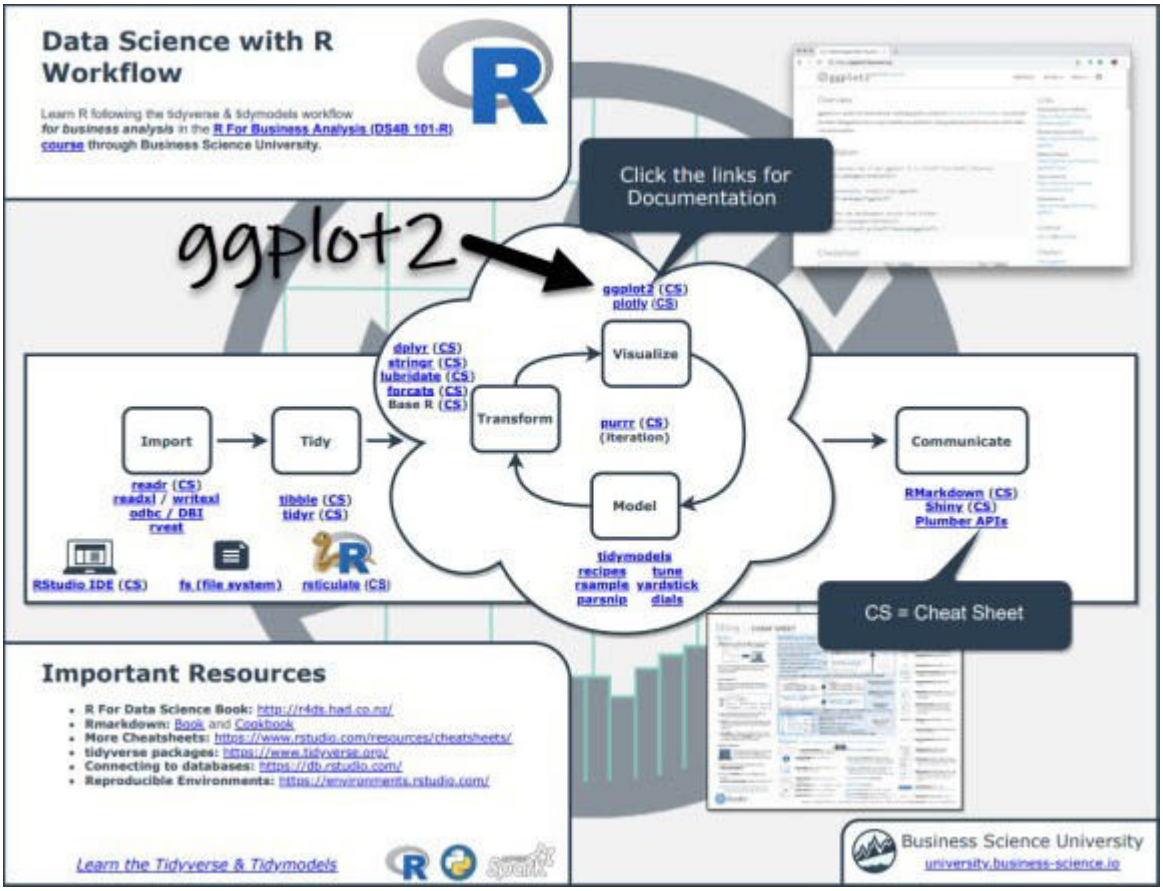
We can easily do this with a half-plot thanks to `gghalves`.

Before we get started, get the R Cheat Sheet

`gghalves` is great for making customized `ggplot2` plots. But, you'll still need to learn how to wrangle data with `dplyr` and visualize data with `ggplot2`. For those topics, I'll use the Ultimate R Cheat Sheet to refer to `dplyr` and `ggplot2` code in my workflow.

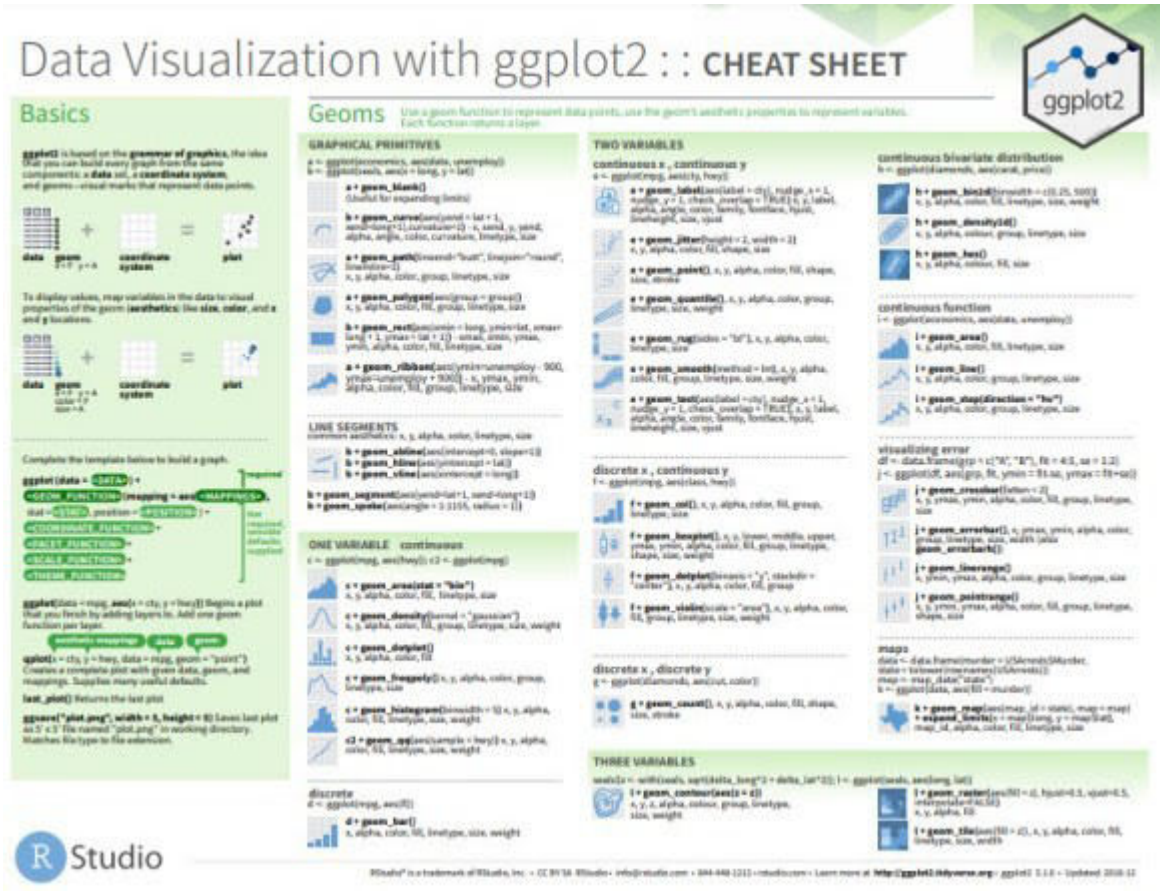
Quick Example:

Download the Ultimate R Cheat Sheet **Then Click the "CS" next to "ggplot2"** opens the Data Visualization with GGplot2 Cheat Sheet.



Now you're ready to quickly reference `ggplot2` functions.

Onto the tutorial.



How gghalves works

The `gghalves` package extends `ggplot2` by adding several new "geoms" (ggplot geometries) that allow us to add half plots. In this tutorial, we'll cover:

- `geom_half_boxplot()`: For creating half-boxplots
- `geom_half_dotplot()`: For creating half-dotplots

Pro Tip:

Simply type "geom_half" in your R console and hit Tab to show all of the half plotting geoms available.

Load the Libraries and Data

First, run this code to:

1. **Load Libraries:** Load `gghalves`, `tidyverse` and `tidyquant`.
2. **Import Data:** We're using the `mpg` dataset that comes with `ggplot2`.

```
# LIBRARIES ----

library(tidyverse)
library(tidyquant)
library(gghalves)

# DATA ----
mpg
```

Get the code.

Make the Half-Boxplot / Half-Dotplot

Next, we can combine a half-boxplot and half-dotplot. This has the advantage of showing:

- **Quantiles and Outliers (Boxplot)**
- **Distribution (Dotplot)**

Business Goal

Suppose we have a question:

> What effect does **Engine Size (number of Cylinders)** have on **Vehicle Highway Fuel Economy (Highway MPG)**?

We can visualize this with `gghalves` by making half-plots of Cylinder vs Highway.

Half-Plot Visualization Code

Using the Ultimate R Cheat Sheet, we can make a `ggplot` from the ggplot2 data visualization cheat sheet. We'll add `geom_half_boxplot()` and `geom_half_dotplot()` to make the half-plots of Cylinder vs Highway.

```r
# 3.0 HALF-BOXPLOT / HALF-DOTPLOT ----
mpg %>%
    filter(cyl != 5) %>%
    mutate(cyl = factor(cyl)) %>%
    ggplot(aes(cyl, hwy, color = cyl)) +

    geom_half_boxplot(outlier.color = "red") +
    geom_half_dotplot(
        aes(fill = cyl),
        dotsize = 0.75,
        stackratio = 0.5,
        color = "black"
    ) +

    facet_grid(cols = vars(cyl), scales = "free_x") +
    scale_color_tq() +
    scale_fill_tq() +
    theme_tq() +
    labs(
        title = "Highway Fuel Economy by Engine Size",
        subtitle = "Half-Boxplot + Half-Dotplot"
    )
```

Get the code.

Half-Plot Visualization

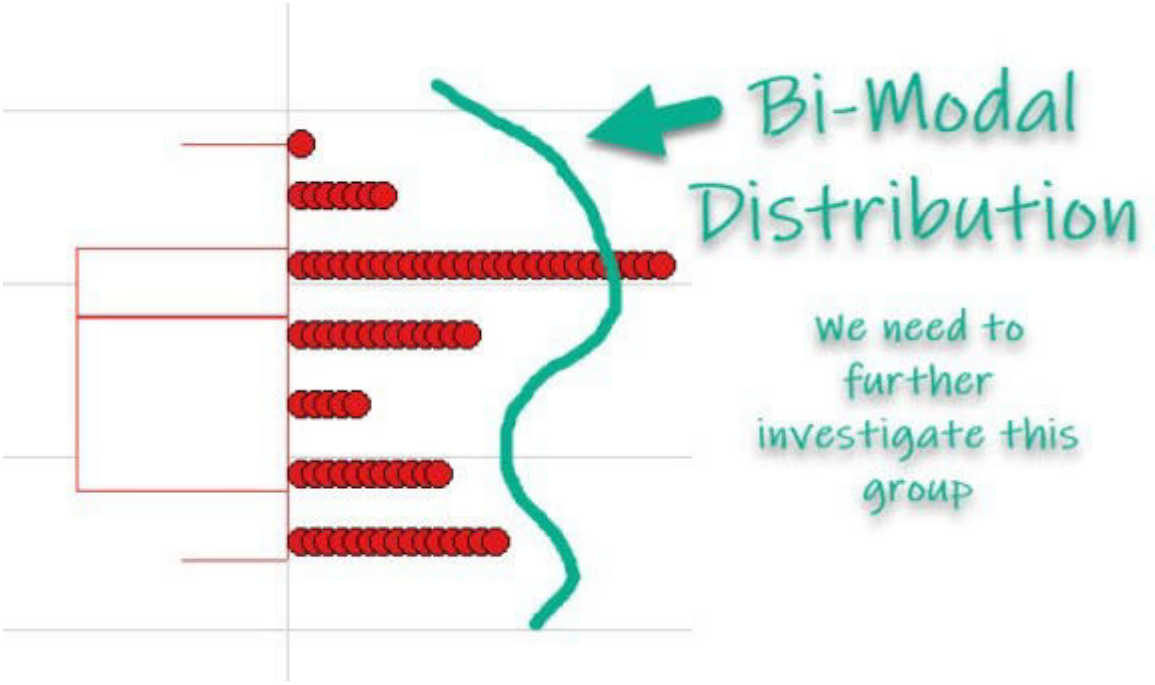
Here is the visualization. We can explore to find an interesting relationship between Engine Size and Fuel Economy.

Get the code.

Insights: Bimodal Distribution of 6-Cylinder Engine Class

Generally speaking, fuel economy goes down as engine size increases. But, the 6-Cylinder engine has something unique going on that has been uncovered by the `gghalves::geom_half_dotplot()`.

The 6-Cylinder Engine class of car has a **bimodal distribution**, which is when there are two peaks. This generally indicates that there are two different populations within the group. We need to investigate with `ggplot2`.

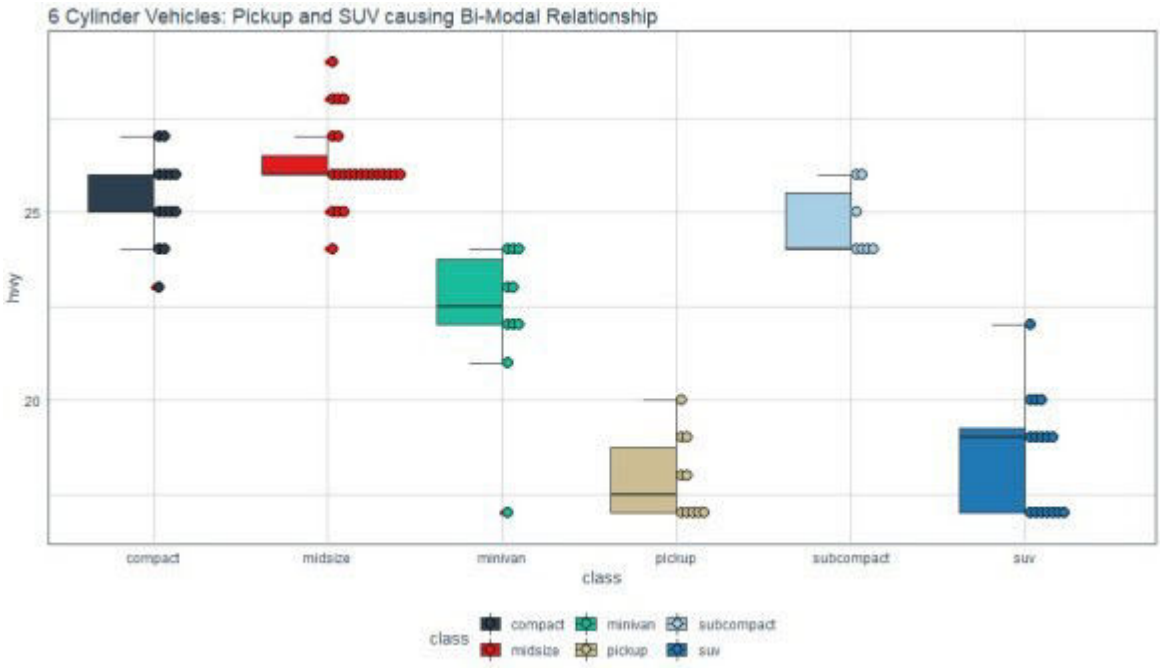


Get the code.

Exploring the Bimodal Relationship

We can explore the 6 Cylinder Vehicle Class a bit further to identify the cause of the Bimodal Distribution. It looks like:

- **SUV and Pickup classes** have much lower fuel economy
- **Compact, Midsize, Minivan, and Subcompact** have much higher fuel economy


6 Cylinder Vehicles: Pickup and SUV causing Bi-Modal Relationship
hwy
25
20
compact
midsize
minivan
pickup
subcompact
suv
class
class
compact
midsize
minivan
pickup
subcompact
suv