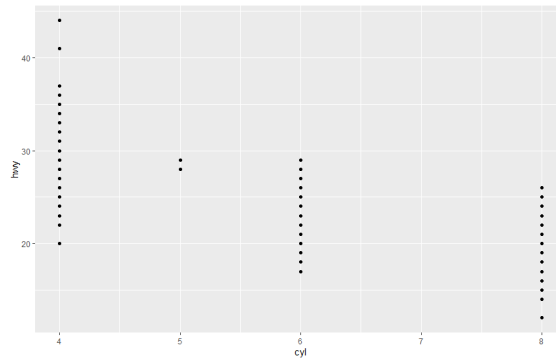


Assignment 6

1. Make a scatter plot of *hwy* vs *cyl*.

```
fig <- ggplot(data=mpg, aes(x=cyl, y=hwy)) \+ geom_point()
```



2. Analyze the vis in terms of the three step framework.

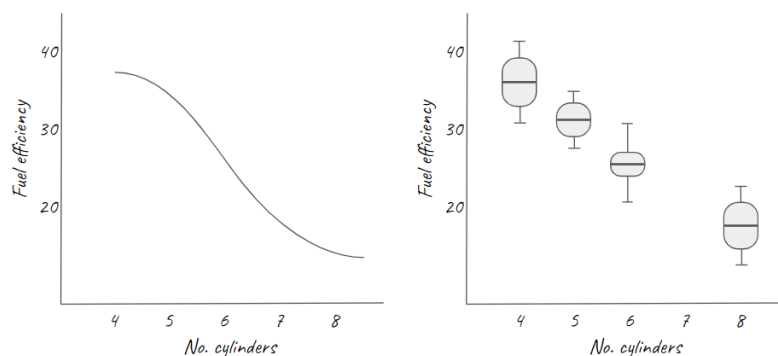
What. A car's fuel efficiency on the highway measured on miles per gallon and the number of cylinders in a car's engine. **Why.** To visualize the relationship between a car's fuel efficiency on the highway and the car's engine number of cylinders. **How.** A dot representing a specific data point with fuel efficiency on the y-axis and number of cylinders on the x-axis.

3. Discuss this vis critically. Why is this vis poor?

The vis does provide some intuition that fuel efficiency decreases as the number of cylinders in the engine increases. However, it lacks several details. First, the dataset has 234 entries and the vis fails to convey that information. Second, the vis does not portray clearly the relationship between *hwy* and *cyl*. For example, we can't safely say the efficiency is different by number of cylinders because we don't know the actual number of data points.

4. Describe the better alternative graphic in your mind by the three step framework.

What. A car's fuel efficiency on the highway measured on miles per gallon and the number of cylinders in a car's engine. **Why.** To visualize the relationship between a car's fuel efficiency on the highway and the car's engine number of cylinders. **How.** A boxplot that shows both the average fuel efficiency per number of cylinders as well as the spread. Or a line that shows the tendency of the average efficiency for number of cylinders.



"Written by Carlos Vintimilla"