Homework 2

This homework is due on the deadline posted on edX. Please submit a .pdf file of your output and upload a .zip file containing your .Rmd file. Do NOT include your name or EID in your filenames.

Problem 1: We will work with the `mpg` dataset provided by **ggplot2**. See here for details: https://ggplot2.tidyverse.org/reference/mpg.html (https://ggplot2.tidyverse.org/reference/mpg.html)

Make two different strip charts of highway fuel economy versus number of cylinders, the first one without horizontal jitter and second one with horizontal jitter. Explain in 1-2 sentences why the plot without jitter is highly misleading.

Hint: Make sure you do not accidentally apply vertical jitter. This is a common mistake many people make.

```
ggplot(mpg, aes(x=cyl, y=hwy)) + geom_jitter(position=position_jitter(0))
```
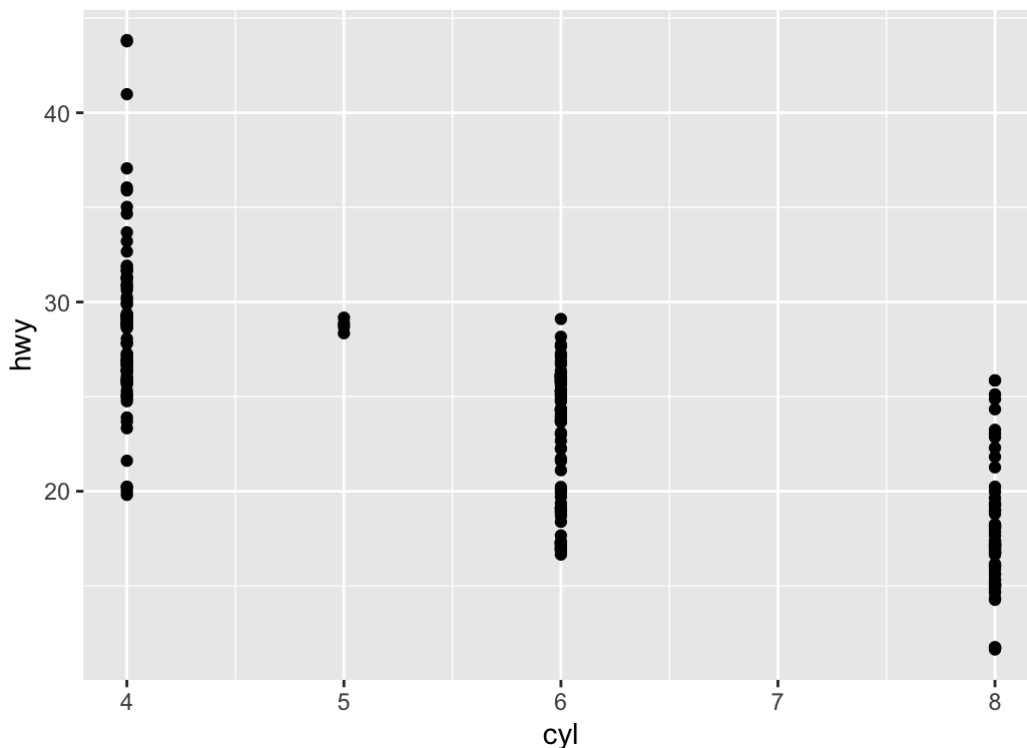


```
ggplot(mpg, aes(x=cyl, y=hwy)) + geom_jitter(position=position_jitter(.2))
```
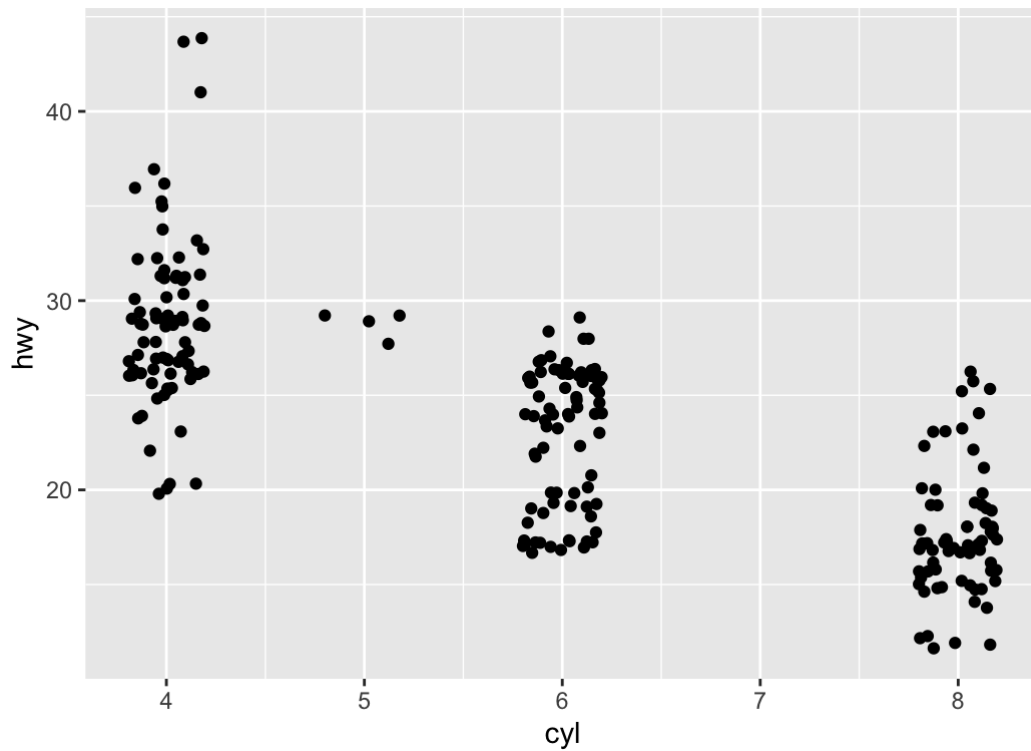
In the first chart, without jitter, it's easy to misinterpret the spread of highway efficiency for cars in each cylinder class. You see that the line in dense/sparse, but it's extremely hard to tell how dense/sparse.

Problem 2: For this problem, we will continue working with the `mpg` dataset. Visualize the distribution of each car's city fuel economy by class and type of drive train with (i) boxplots and (ii) ridgelines. Make one plot per geom and do not use faceting. In both cases, put city mpg on the x axis and class on the y axis. Use color to indicate the car's drive train.
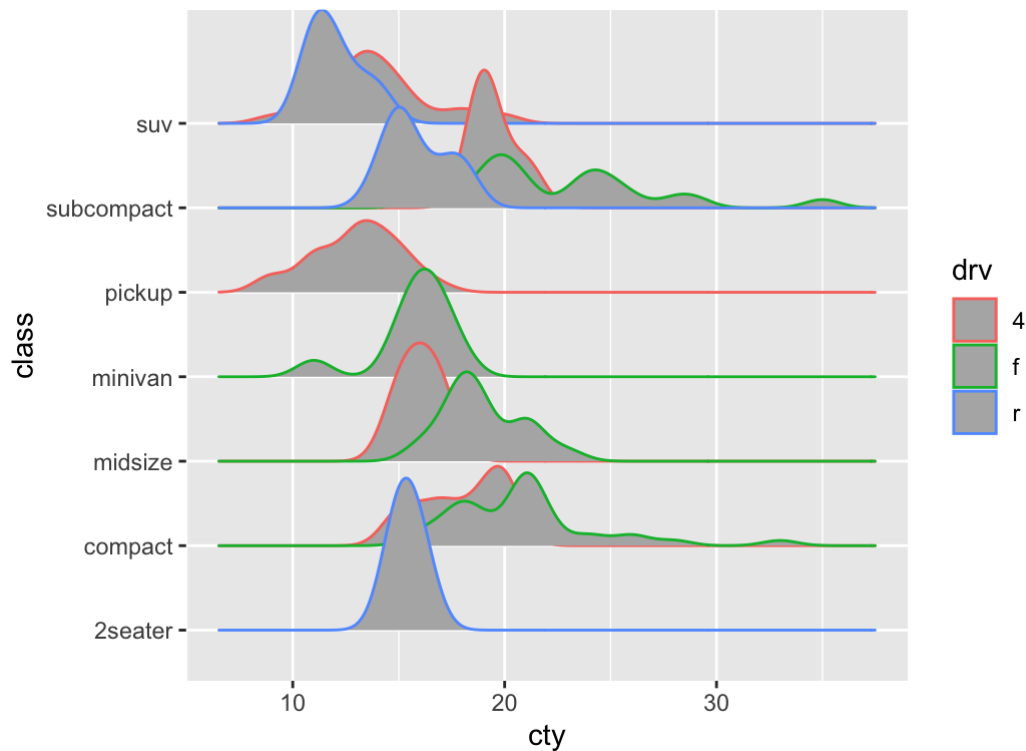
The boxplot ggplot generates will have a problem. Explain what the problem is. (You do not have to solve it.)

```
ggplot(mpg, aes(x=cty, y=class)) + geom_boxplot(aes(color = drv))
```
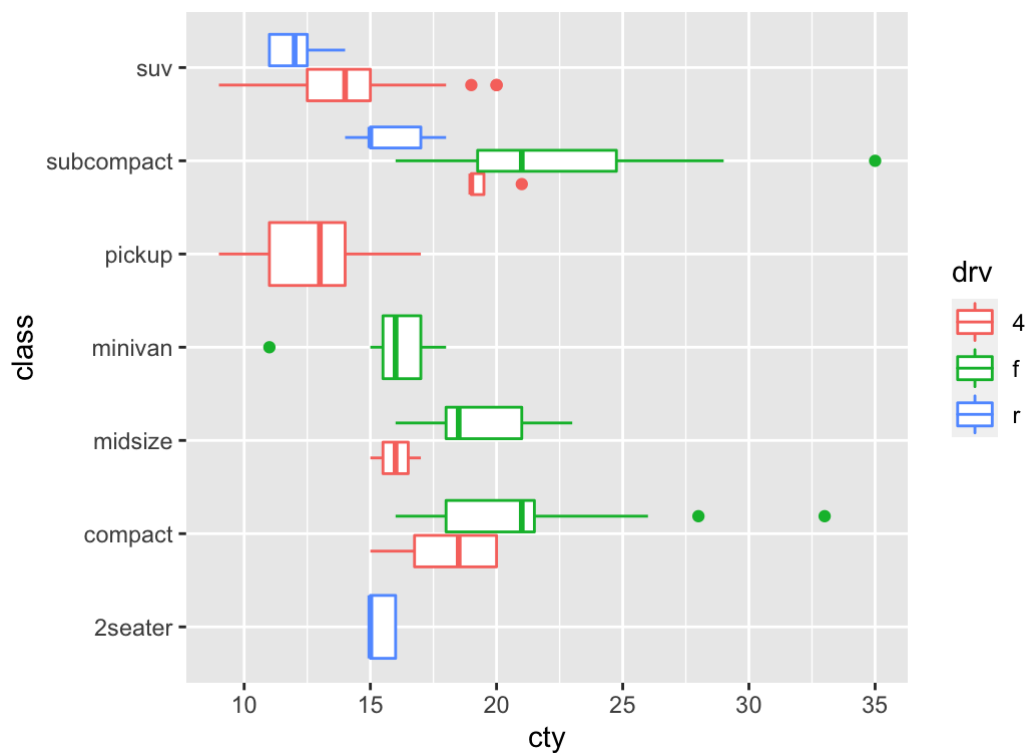
```
ggplot(mpg, aes(x=cty, y=class))+ geom_density_ridges(aes(color = drv))
```

```
## Picking joint bandwidth of 0.828
```


In the first plot, it can be hard to tell which colored boxplot belongs to what vehicle class. You really have to read into the plot and decide "okay this must be the boxplot for subcompact rear-wheels because it is closer to the this boxplot, this must be the boxplot for compact front wheels because it is closer to that boxplot, etc."