

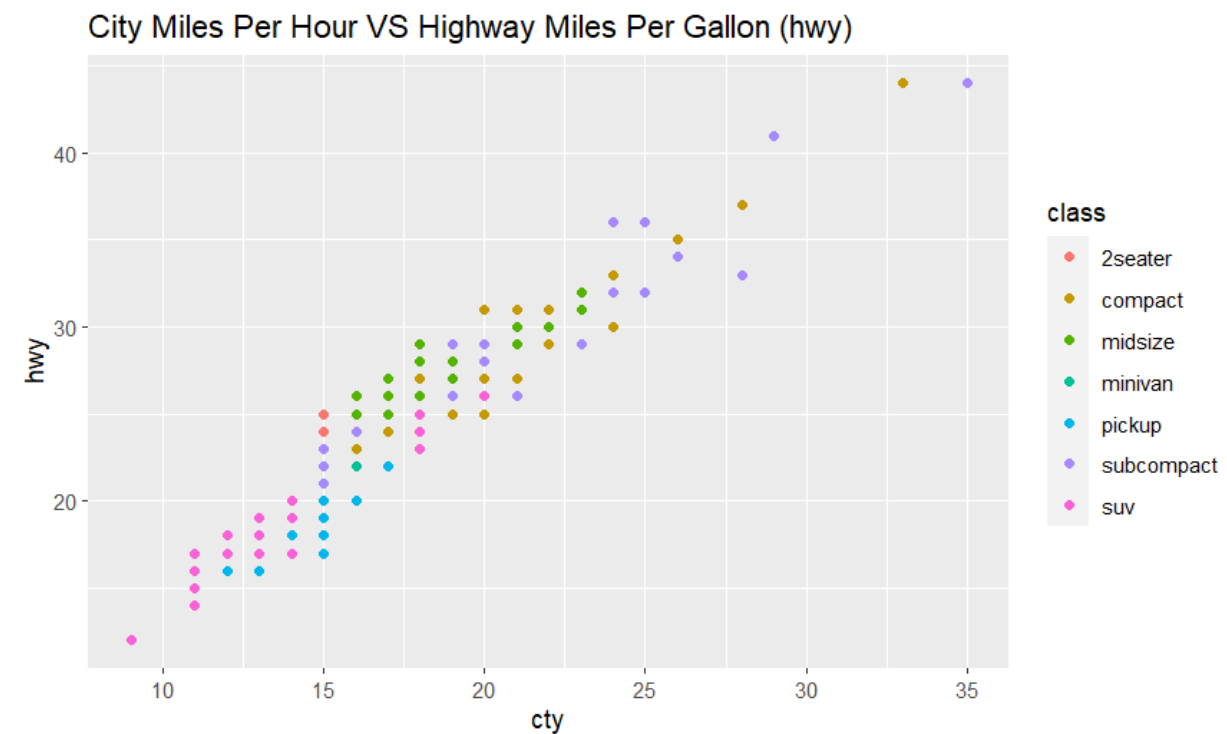
Ex1:

```
data <- mpg
glimpse(data)
```

```
> data <- mpg
> glimpse(data)
Rows: 234
Columns: 11
$ manufacturer <chr> "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", "audi", ...
$ model <chr> "a4", "a4", "a4", "a4", "a4", "a4", "a4", "a4 quattro", "a4 quattro", "a...
$ displ <dbl> 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 3...
$ year <int> 1999, 1999, 2008, 2008, 1999, 1999, 2008, 1999, 1999, 2008, 2008, 1999, ...
$ cyl <int> 4, 4, 4, 4, 6, 6, 6, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 8, 8, 8, 8, 8, 8, ...
$ trans <chr> "auto(15)", "manual(m5)", "manual(m6)", "auto(av)", "auto(15)", "manual(...
$ drv <chr> "f", "f", "f", "f", "f", "f", "f", "f", "4", "4", "4", "4", "4", "4", "4", "4...
$ cty <int> 18, 21, 20, 21, 16, 18, 18, 18, 16, 20, 19, 15, 17, 17, 15, 15, 17, 16, ...
$ hwy <int> 29, 29, 31, 30, 26, 26, 27, 26, 25, 28, 27, 25, 25, 25, 25, 24, 25, 23, ...
$ fl <chr> "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p...
$ class <chr> "compact", "compact", "compact", "compact", "compact", "compact", "compact", "compa...
```

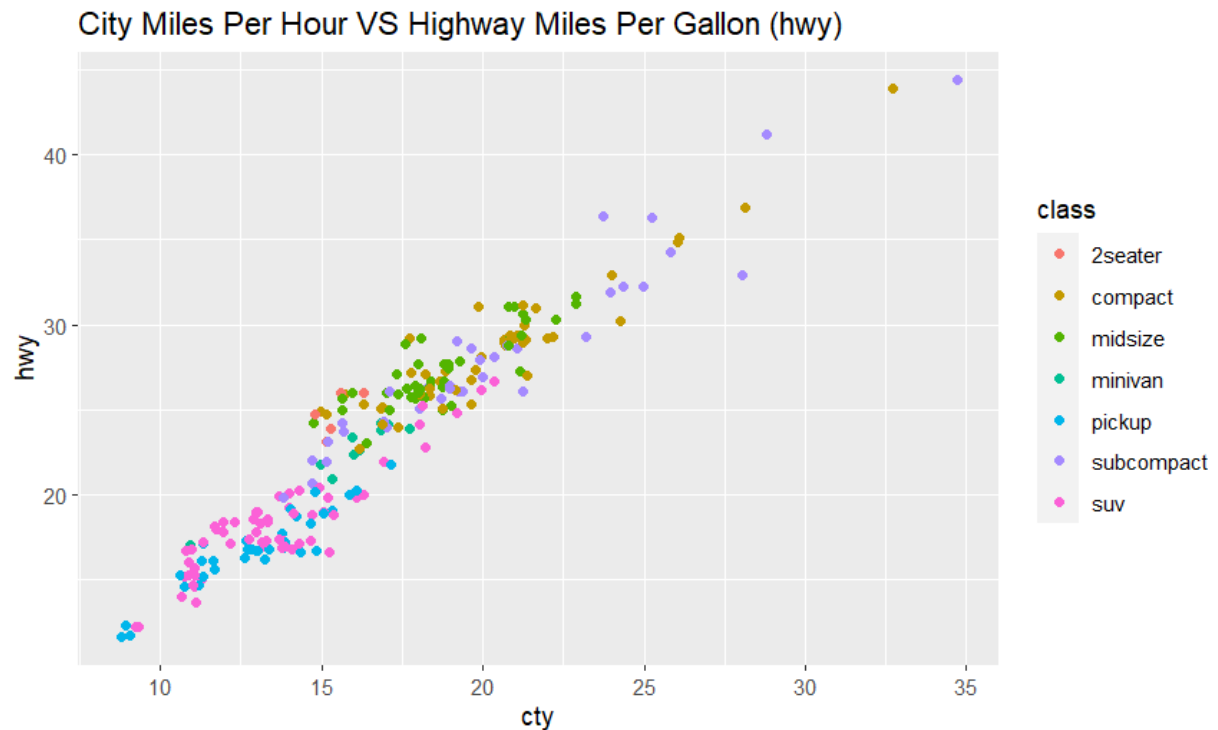
Ex2:

```
ggplot(data=mpg, mapping=aes(x=cty, y=hwy, color=class))+geom_point()+labs(title="City Miles Per
Hour VS Highway Miles Per Gallon (hwy)")
```



Ex3:

```
ggplot(data=mpg, mapping=aes(x=cty, y=hwy, color=class))+geom_jitter()+labs(title="City Miles Per Hour VS Highway Miles Per Gallon (hwy)")
```



- **geom_point:**
 - Advantages:
 - Simple and clear way to visualize points on a plot
 - Easy to control point aesthetics like color, size, shape
 - Disadvantages:
 - Points can overlap if there are many observations in the same location
 - Overlapping points make it hard to see density or count of observations
- **geom_jitter:**
 - Advantages:
 - Prevents overlapping of points by adding random noise/jitter to the points
 - Better for dense datasets as it shows underlying patterns and density clearly
 - Can still see individual observations unlike geom_density()



Disadvantages:

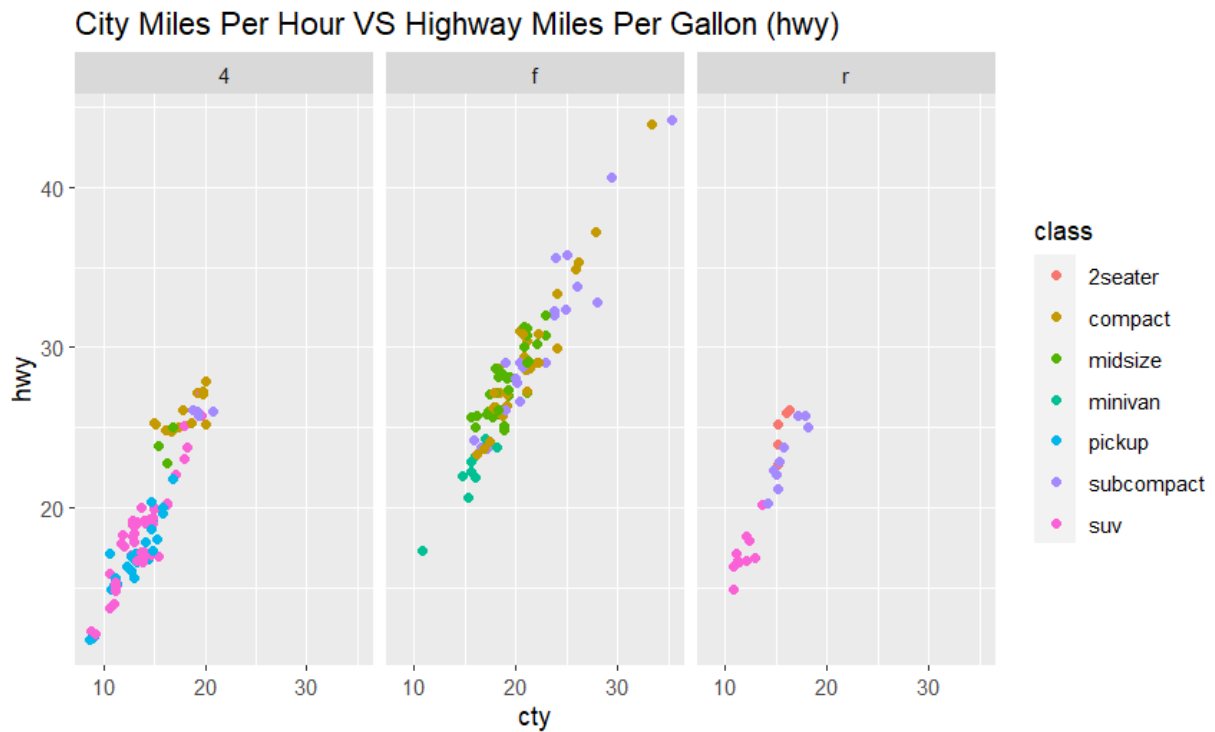
Adds random noise so exact positions are approximations rather than true values

Less precise than `geom_point` if seeing exact positions is important

More cluttered looking than `geom_point` if there are many points

Ex4:

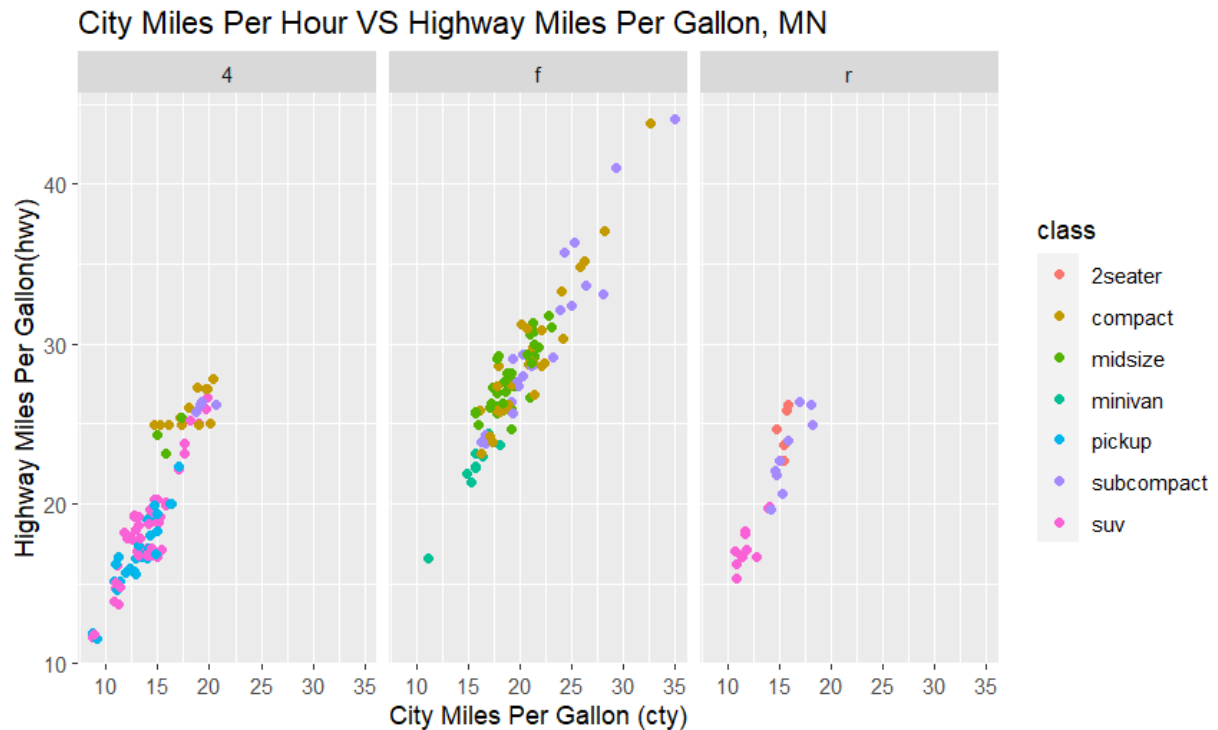
```
ggplot(data=mpg, mapping=aes(x=cty, y=hwy, color=class))+geom_jitter()+labs(title="City Miles Per Hour VS Highway Miles Per Gallon (hwy)")
+facet_wrap(drv~.)
```



A car's class clearly impacts its city MPG (cty). Larger vehicles like SUVs, trucks and vans get fewer cty MPG than smaller cars. The connection between class and cty MPG seems non-linear, with SUVs lowest and tiny two-seaters highest.

Ex5:

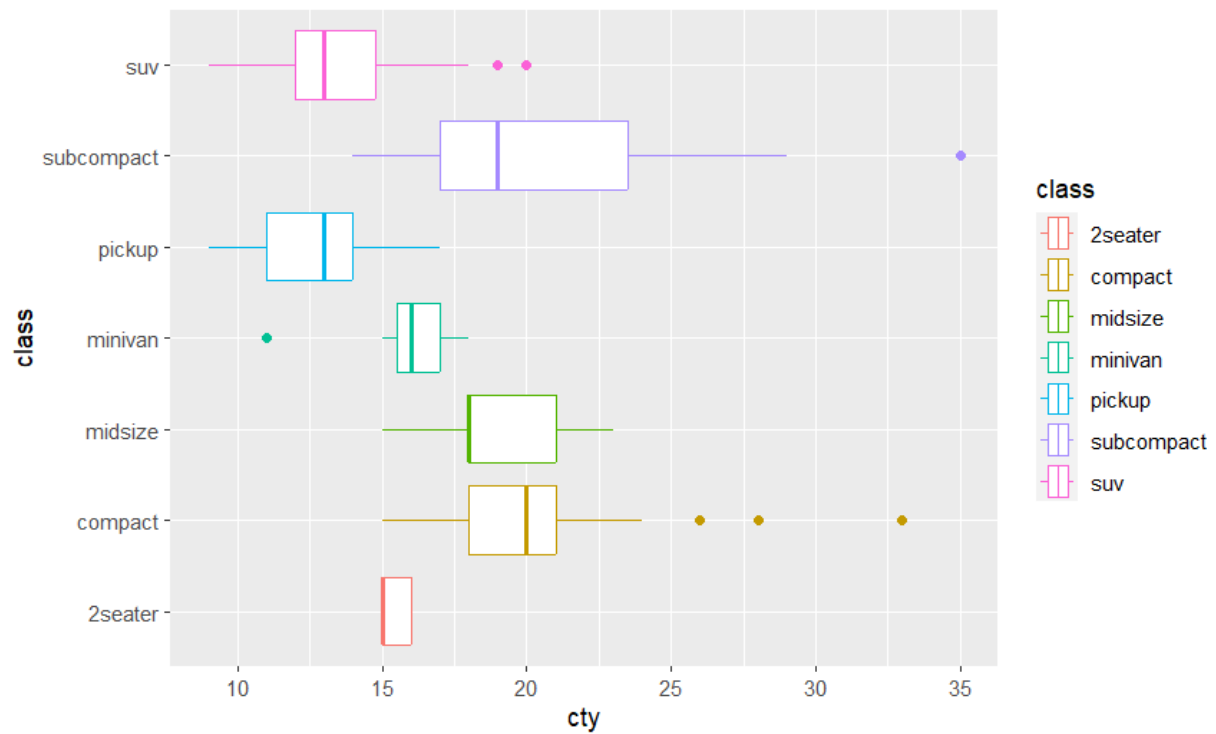
```
ggplot(data=mpg, mapping=aes(x=cty, y=hwy, color=class))+geom_jitter()+labs(title="City Miles Per Hour VS Highway Miles Per Gallon, MN", x="City Miles Per Gallon (cty)", y="Highway Miles Per Gallon(hwy)") +facet_wrap(drv~.)
```



The relationship between city and highway MPG for different vehicle classes. Larger vehicles have lower city MPG than smaller ones. There is a non-linear link between class and city MPG. City MPG also varies significantly within each class.

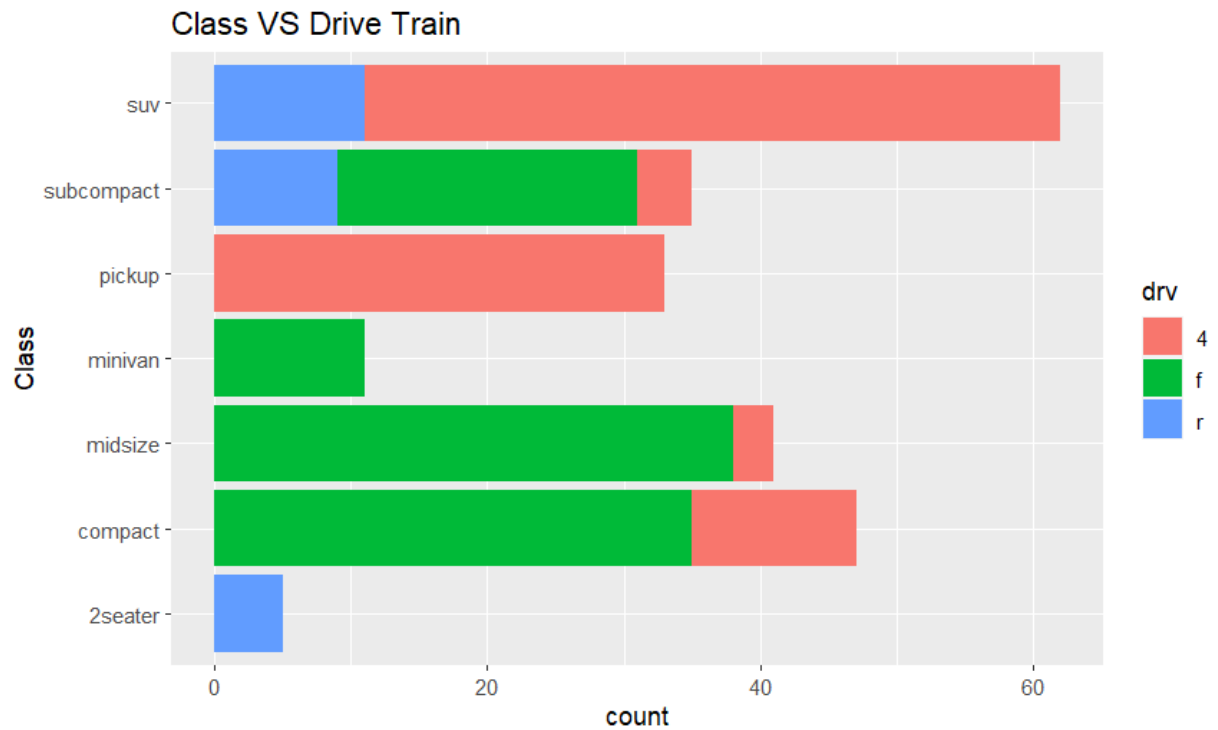
Ex6:

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



Ex7:

```
ggplot(data=mpg)+geom_bar(mapping=aes(x=class, fill=drv))+coord_flip()+labs(title="Class VS Drive Train", x="Class")
```



Ex8:

```
mpg <- mpg %>%
  rename(vehicle_class = class)
ggplot(data=mpg, mapping=aes(x=displ, y=hwy))+geom_point(size=0.5)+labs(title="Highway mpg VS
Displacement",x="Displacement (Litres)", y="Highway mpg")+theme_bw()facet_wrap(~vehicle_class,
nrow=2)
```

