

## Part1 - Explore

IMSOHYUN

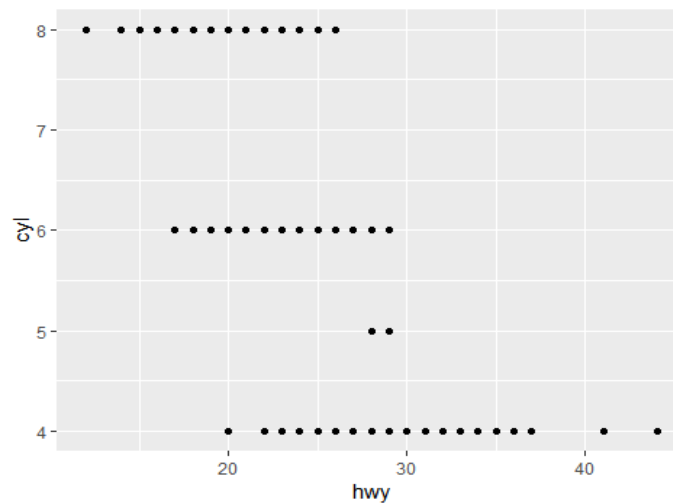
2019 년 3 월 9 일

### Use mpg data

- manufacturer: 자동차 제조사
- model : 자동차 모델명
- displ : 엔진크기 (liter)
- year : 연식
- cyl: 실린더 수
- trans: transmission 타입
- drv: f= front-wheel drive, r = rear wheel drive, 4 = 4wd
- cty : 도심도로 마일리지 (mile/gallon)
- hwy : 고속도로 마일리지 (mile/gallon)
- fl : fuel type
- class : 자동차 타입

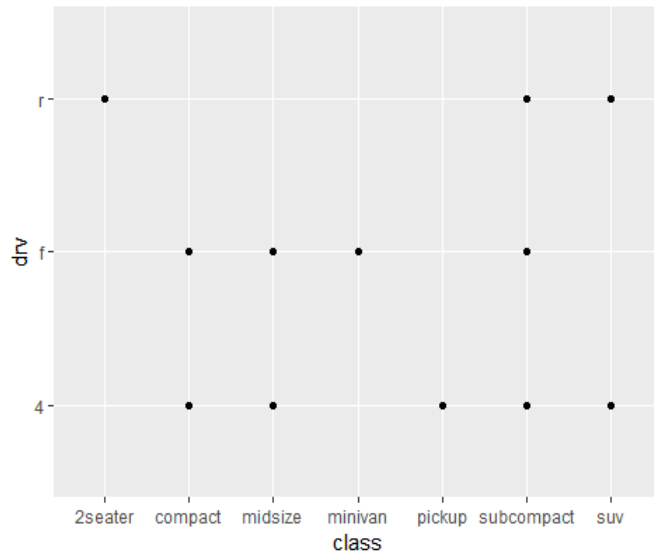
#### 1. Make a scatterplot of hwy vs cyl.

```
ggplot(mpg, aes(hwy, cyl)) + geom_point()
```



## 2. What happens if you make a scatterplot of class vs drv. Why is the plot not useful?

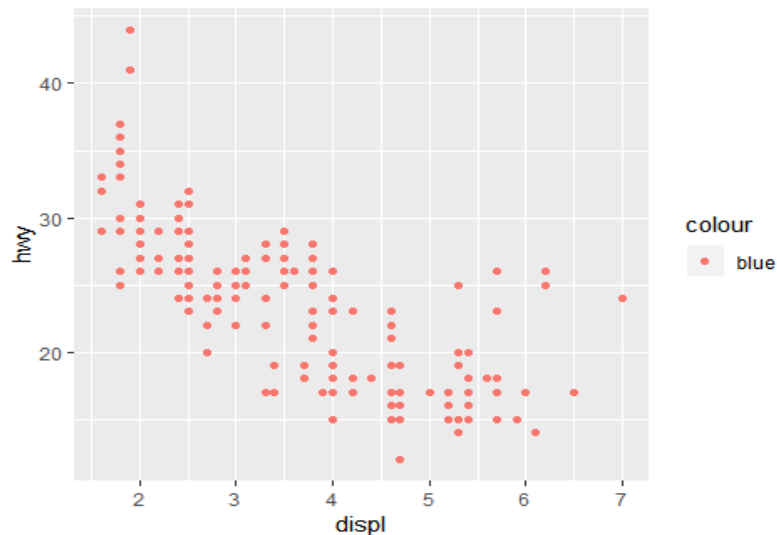
```
ggplot(mpg, aes(class, drv)) + geom_point()
```



Class 와 drv 는 범주형 자료이므로  
왼쪽의 그래프는 x,y 의 관계를  
효과적으로 보여주지 못한다.

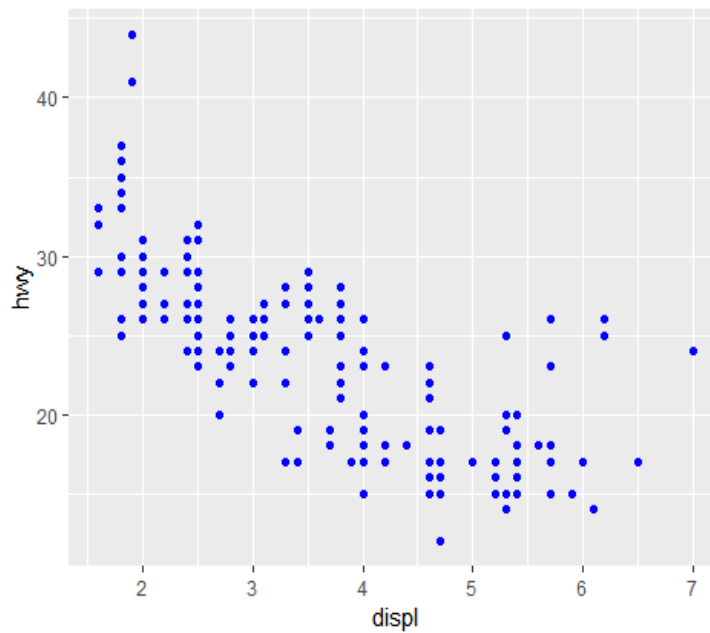
## 3. What's gone wrong with this code? Why are the points not blue?

```
ggplot(mpg) + geom_point(aes(displ, hwy, color = "blue"))
```



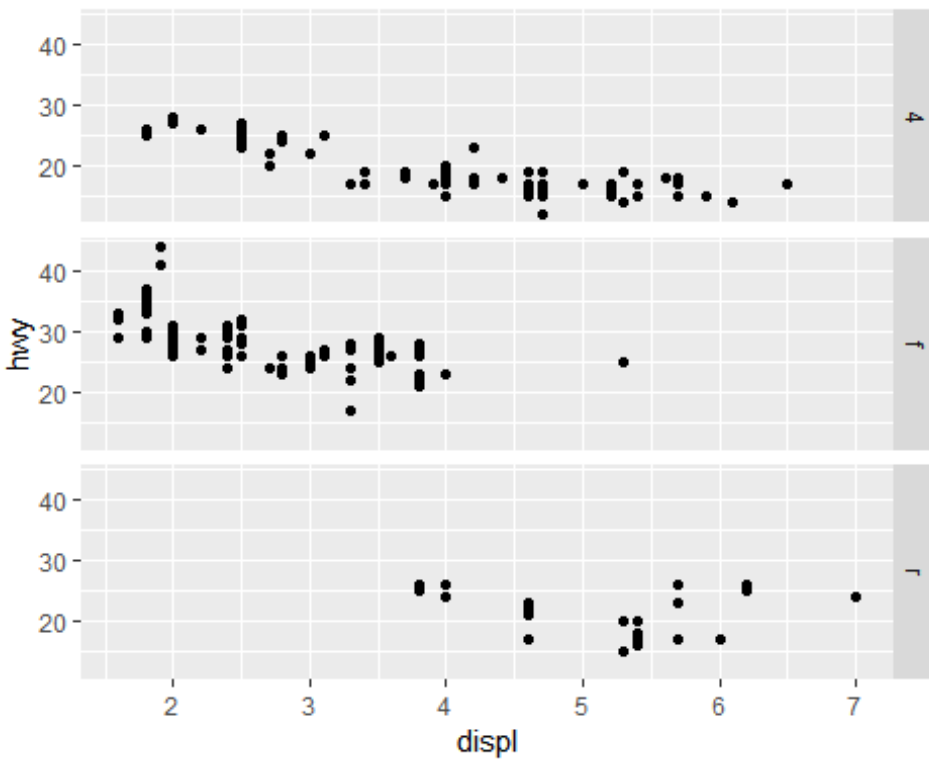
보통 aes 안에서 color 를 지정하는 경우에는 class 별로 색을 다르게 하고 싶을 때 사용한다. 전체 점을 파란색으로 지정하고 싶을 때의 코드는 다음과 같다.

```
ggplot(mpg) + geom_point(aes(displ,hwy), color = "blue")
```

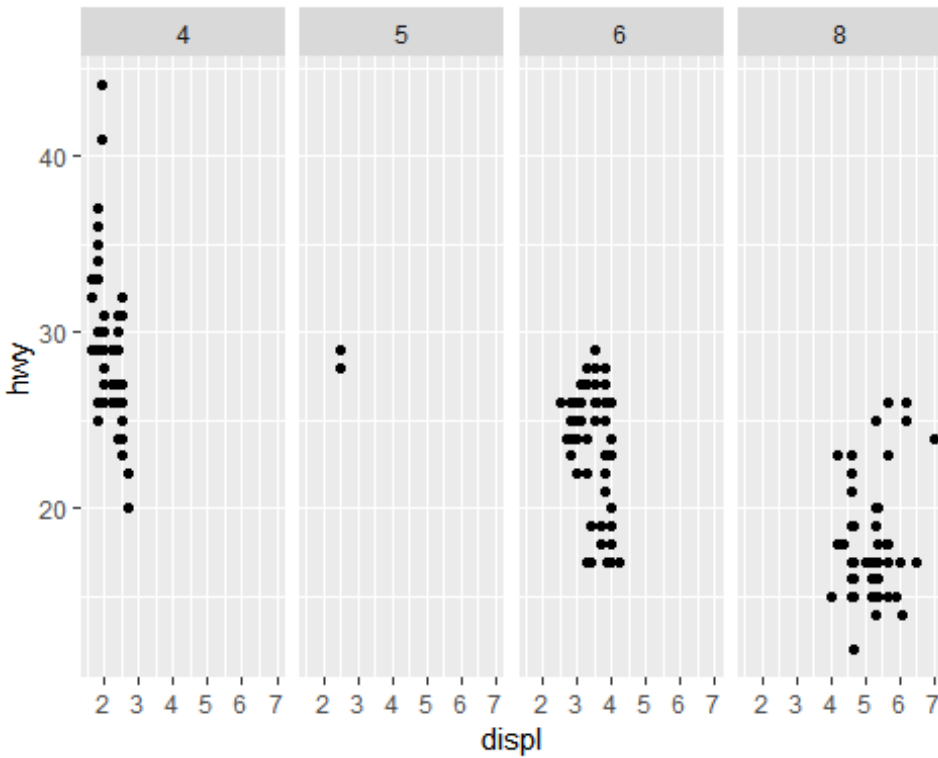


4. What plots does the following code make? What does `'.'` do?

```
ggplot(mpg) +  
geom_point(aes(displ,hwy)) + facet_grid(drv ~ .)
```



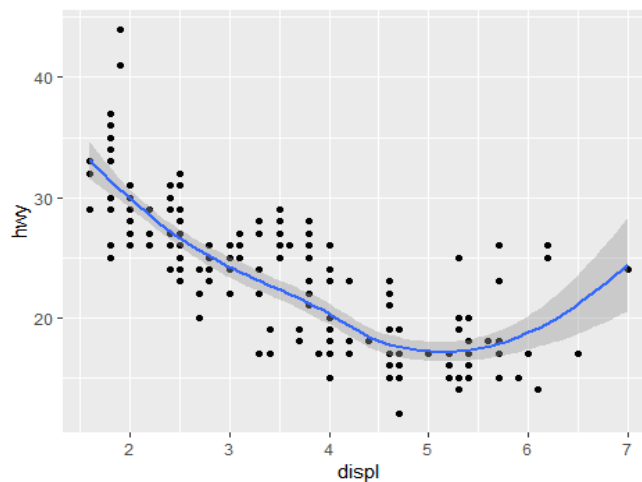
```
ggplot(mpg) +  
geom_point(aes(displ,hwy)) + facet_grid(. ~ cyl)
```



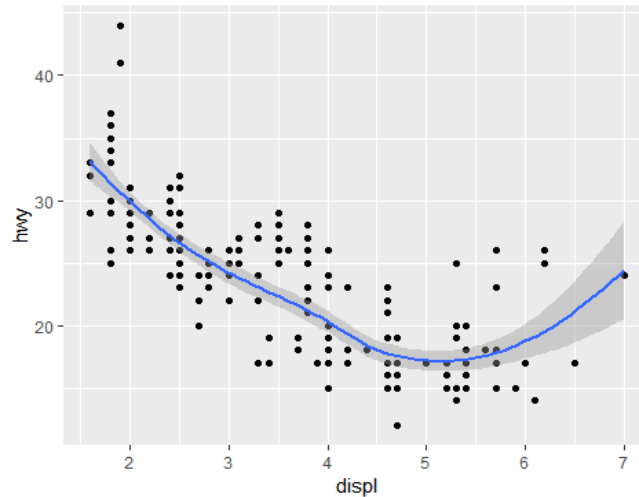
"attribute ~ dot"표기법은 각 속성 값에 대해 여러 개의 행 방향 플롯을 표시하고 "dot ~ attribute"를 사용하면 행 속성이 누락되어 표시가 열 단위로 표시된다.

### 5. Will these two graphs look different? Why/why not?

```
ggplot(mpg, aes(displ,hwy)) + geom_point() + geom_smooth()
```



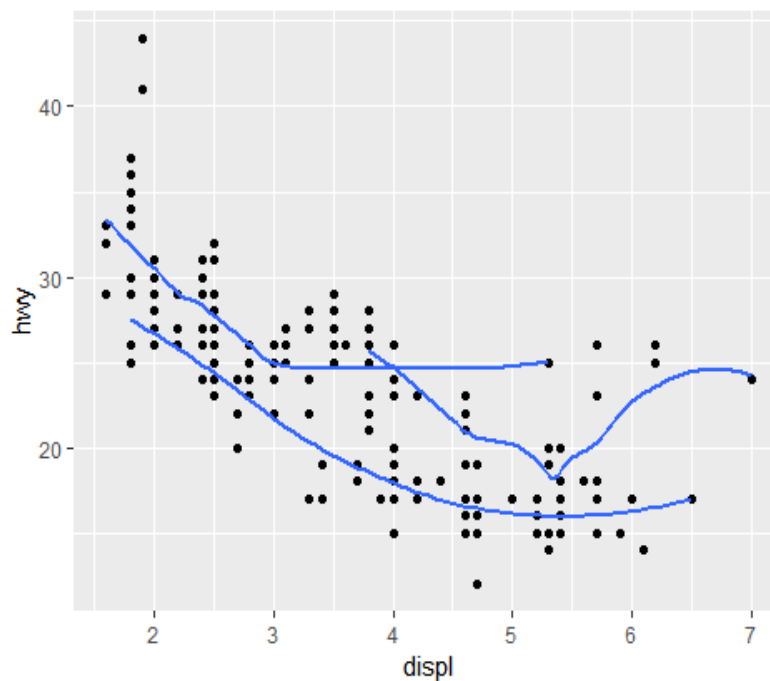
```
ggplot(mpg) + geom_point(aes(displ, hwy)) + geom_smooth(aes(displ, hwy))
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



geom\_point 와 geom\_smooth 에서 지정된 x,y 축이 모두 같기 때문에 두 그래프는 같다.

## 6. Recreate the R code necessary to generate the following graphs

```
ggplot(mpg, aes(displ,hwy)) +
  geom_point() +
  geom_smooth(aes(group=drv),se = FALSE)
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



```
ggplot(mpg, aes(displ, hwy)) +  
  geom_point(aes(color=drv)) +  
  geom_point(shape = 21, color = "white", stroke = 2)
```

