



ggplot2 패키지의 qplot() 함수는
이미 지원이 중단(deprecated)되었지만

사용 방법이 간단하다 보니
R 언어 학습에 활용하는 경우가 있어
본 부록 PDF를 제공하게 되었습니다.

내용 구성이 빈약할 수 있습니다.

== ggplot2 citation ==

H. Wickham. ggplot2: Elegant Graphics for Data Analysis.

Springer-Verlag New York, 2016.

qplot 그래프 그리기

[주의] qplot은 더 이상 사용이 권장되지 않습니다. ggplot 함수 사용을 장려하기 위해 qplot 함수는 deprecated 되었습니다. (지원 중단)

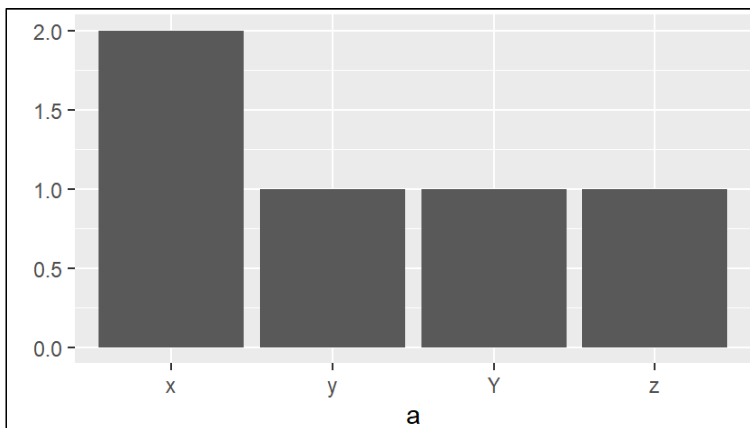
A Quick Plot

- **Quick Plot:** 빈도수 그래프를 그리는 ggplot2 패키지의 함수이다.

ggplot2 패키지를 설치하고 불러왔으면 `qplot()` 함수를 이용해서 빈도 막대 그래프를 그려보도록 하자.

```
a <- c("x", "x", "y", "Y", "z")
qplot(a)
```

위 코드를 실행하면



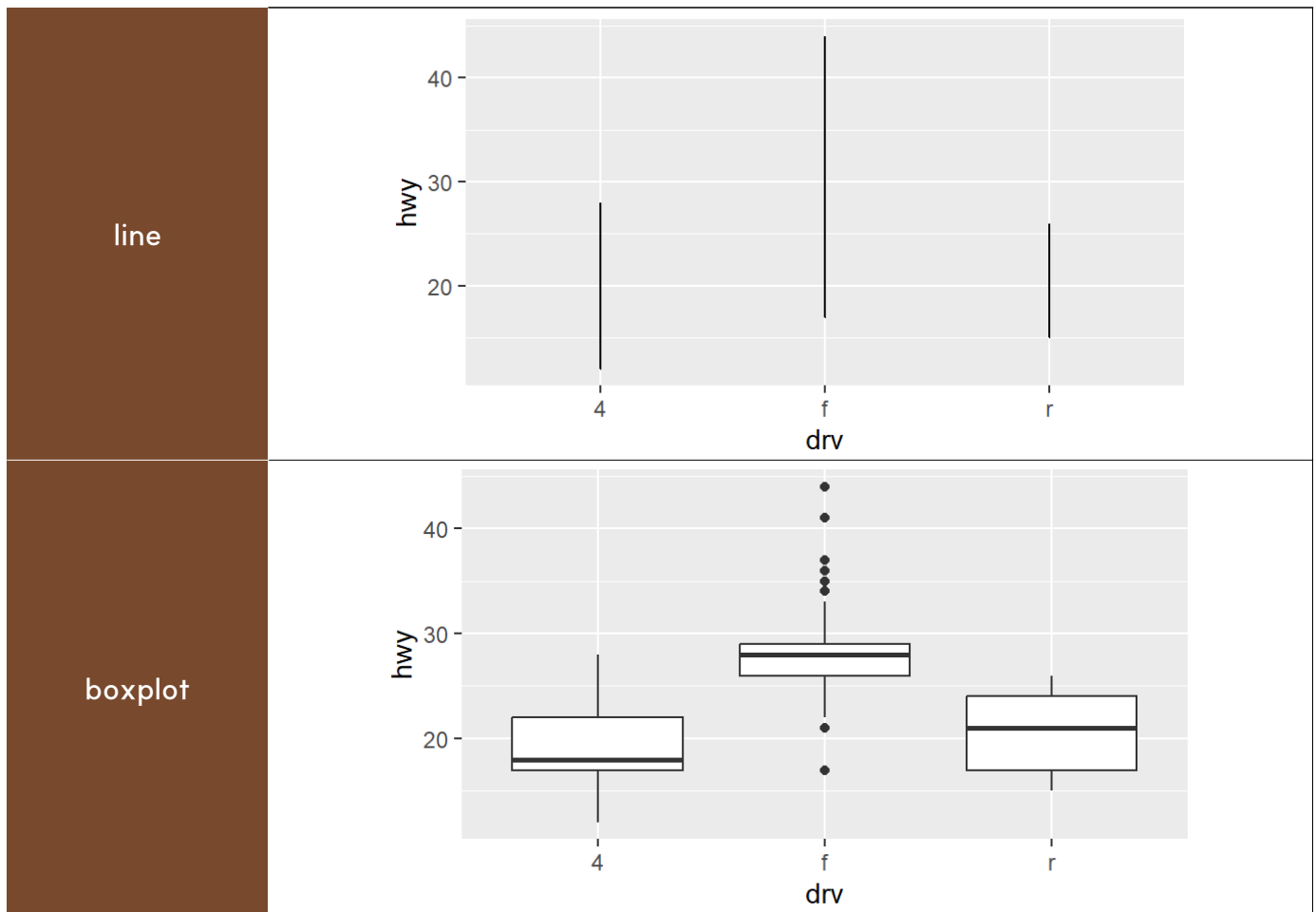
왼쪽 그림과 같이 `a` 변수에서 특정 문자열이 나오는 빈도를 나타낸 빈도수 그래프가 파일 창에 표시된 것을 확인할 수 있다.

B qplot의 인자

- **qplot의 인자:** `qplot()`에도 매개변수들이 있는데, `data`, `x`, `y`, `geom`, `colour`가 있다.

data	x	y	geom	colour
어떤 데이터를 가져올 것인가?	x축 변수는 무엇으로 할 것인가?	y축 변수는 무엇으로 할 것인가?	어떤 그래프 형태로 나타낼 것인가?	어떤 변수를 기준으로 색을 구분할 것인가?

`data`, `x`, `y`, `colour`에는 다음표 없이 해당하는 변수를 지정하면 되지만, `geom`은 다음표로 형태에 대한 것을 나타내주어야 한다. 다음 표는 `geom`에 넣을 수 있는 그래프 형태의 예시이다.

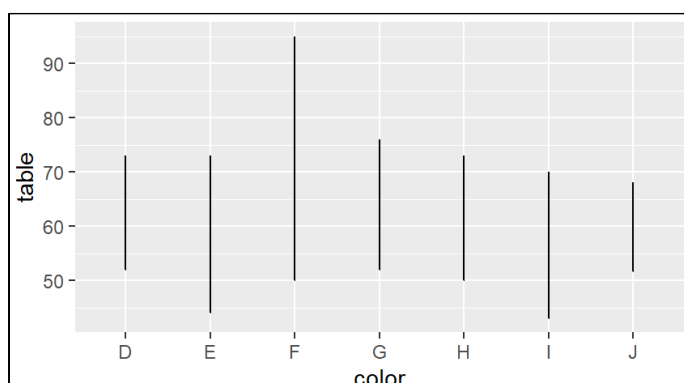


이 부록에서는 line과 boxplot만 알아보도록 하자.

ggplot2에는 기본적으로 diamonds(50,000개 이상의 다이아몬드에 대한 가격, 크기, 품질 등 정보), msleep(다양한 동물의 수면 시간과 관련된 정보), economics(미국의 시간에 따른 실업률, 개인 소득 등의 정보), presidential(미국 대통령 임기 정보), midwest(미국 중서부 지역의 인구 조사 정보), mpg(1999~2008년 사이 미국에서 출시된 자동차들의 연비 관련 정보) 데이터가 탑재되어 있다.

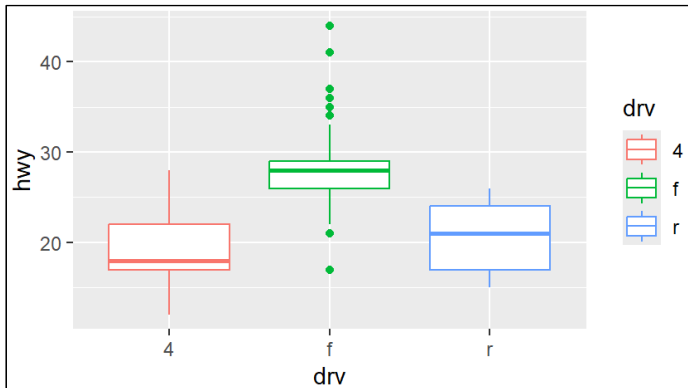
```
qplot(data = diamonds, x = color, y = table)
```

위 코드를 실행하면 diamonds에서 x축에는 color라는 변수의 데이터를, y축에는 table라는 변수의 데이터로 그래프를 그리게 된다.

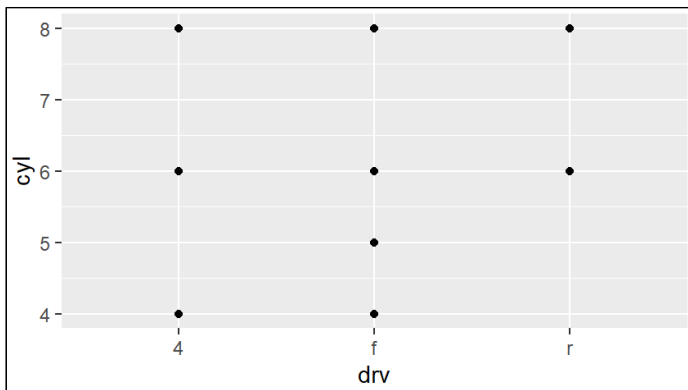


```
qplot(data = mpg, x = drv, y = hwy, geom = "boxplot", colour = drv)
```

위 코드를 실행하면 mpg에서 x축에는 drv라는 변수의 데이터를, y축에는 hwy라는 변수의 데이터를, 데이터 형태는 상자 그림 형태로, 색 구분은 drv 변수 기준으로 그래프를 그리게 된다. 참고로, 색 구분은 x축에 해당하는 변수를 넣어주면 좋다.



```
qplot(data = mpg, x = drv, y = cyl)
```



연습 문제

1. 다음 중 qplot 함수에서 그래프 형태로 지정할 수 없는 것은?

- ① line
- ② histogram
- ③ rail
- ④ boxplot
- ⑤ path

2. qplot 함수에서 색상을 구분하고 싶을 때 사용하는 매개변수의 이름은?

- ① geom
- ② data
- ③ x
- ④ y
- ⑤ colour

3. qplot 함수의 기본 사용법에 대한 설명으로 올바른 것은?

- ① data 인자에는 데이터 프레임을 지정한다.
- ② y축에 사용할 변수는 사용할 수 없다.
- ③ x축에 사용할 변수에는 반드시 따옴표를 사용해야 한다.
- ④ geom 인자에는 변수명을 직접 입력한다.
- ⑤ colour 인자에는 색상 코드를 입력한다.

4. 다음 <조건>을 모두 만족하는 R 코드를 작성하여라.

코드

- ggplot2 패키지의 mpg 데이터를 사용한다.
- x축에는 drv 변수의 데이터를, y축에는 hwy 변수의 데이터를 사용하는 그래프를 그린다.
- 그래프는 상자 그림 형태로 그린다.

5. 다음 <조건>을 모두 만족하는 R 코드를 작성하여라.

코드

- ggplot2 패키지의 diamond 데이터를 사용한다.
- x축에는 cut 변수의 데이터를, y축에는 clarity 변수의 데이터를 사용하는 그래프를 그린다.

<정답 및 해설>

문제 1. 정답: ③ | **해설:** qplot에 rail이라는 그래프 형태는 없다.

문제 2. 정답: ⑤ | **해설:** qplot에서 색상을 구분할 때는 colour를 사용한다.

문제 3. 정답: ① | **해설:** [1] data 파라미터에는 데이터 프레임을 지정한다. (O) [2] y축에 사용할 변수를 정의할 수 있다. (X) [3] 따옴표를 적으면 안 된다. (X) [4] geom은 어떤 그래프 형태로 나타낼 것인지 정하는 것으로 그래프 형태에 대한 것이 들어가야 한다. (X) [5] colour는 어떤 변수를 기준으로 색을 구분할 것인지 정하는 것으로 변수명(열 이름)이 들어가야 한다. (X)

문제 4. 모범 답안:

```
library(ggplot2)
qplot(data = mpg, x = drv, y = hwy, geom = "boxplot")
```

 |
평가 기준: 모범 답안과 같으면 정답으로 인정한다.

문제 5. 모범 답안:

```
library(ggplot2)
qplot(data = diamond, x = cut, y = clarity)
```

 | **평가 기준:** 모범
답안과 같으면 정답으로 인정한다.