[2부] 그래프 디자인의 기본 원칙

목차

17장 잉크양 비례의 법칙
- 17.1 선형 축 도표 그리기
- 17.2 로그 축이 있는 도표 그리기
- 17.3 데이터 값을 면적으로 나타내기

18장 오버플로팅: 겹치는 점 처리 방법
- 18.1 반투명 값과 지터링
- 18.2 이차원 히스토그램
- 18.3 등고선 그래프

19장 효과적인 색 사용을 위한 조언
- 19.1 목적 없는 무분별한 색 사용은 금물
- 19.2 비단조적 색상 스케일 구성
- 19.3 색각 이상자를 배려하자

20장 불필요한 기호화
- 20.1 불필요한 기호화는 범례 디자인을 망친다
- 20.2 범례가 없어도 좋은 도표

21장 여러 개의 패널로 구성된 도표
- 21.1 소형 다중패널
- 21.2 복합형 도표

22장 제목, 캡션, 표의 효과적 활용
- 22.1 도표 제목과 캡션
- 22.2 축과 범례 제목
- 22.3 표, 제대로 만들기

23장 데이터 이해를 돕는 시각화의 주변 요소
- 23.1 적절한 수준의 요소 활용
- 23.2 배경 격자
- 23.3 쌍 데이터

24장 축 레이블의 글자는 큼직하게

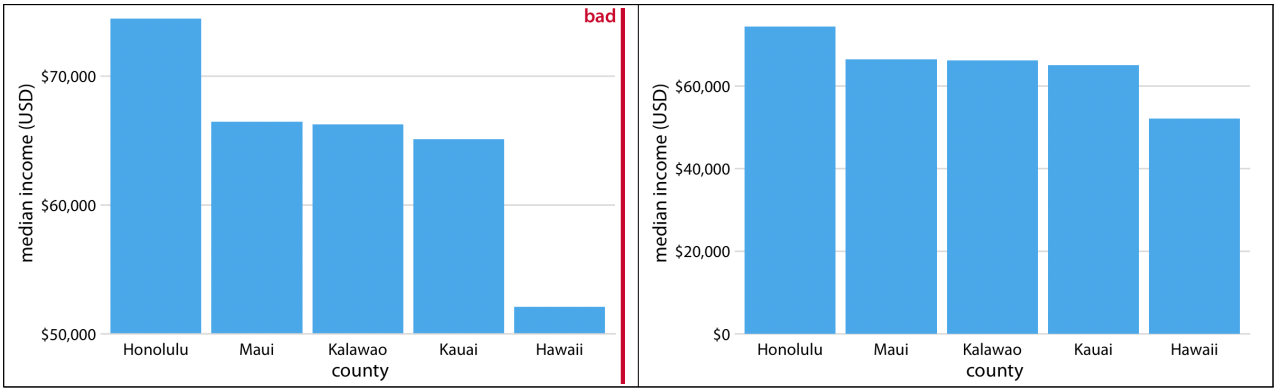
25장 선 그림은 피하자

26장 3차원 그래픽과 차트는 그만
- 26.1 불필요한 3차원 그래픽은 지양하자
- 26.2 3차원 위치 스케일도 이제 그만
- 26.3 3차원 시각화가 필요한 경우

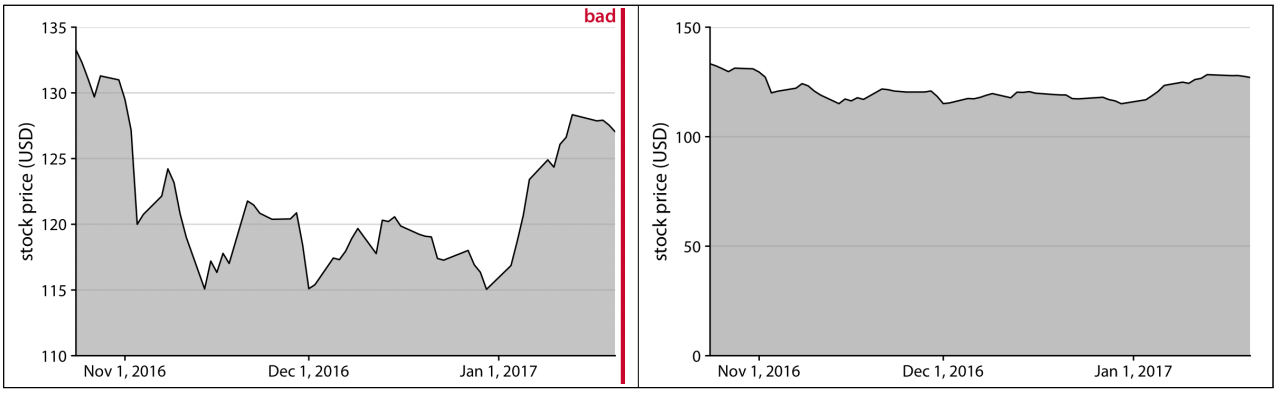
17장 잉크양 비례의 법칙

색조를 넣은 영역이 수치 값을 나타낼 때, 색조가 들어간 영역의 면적은 해당 값과 정비례해야 한다.

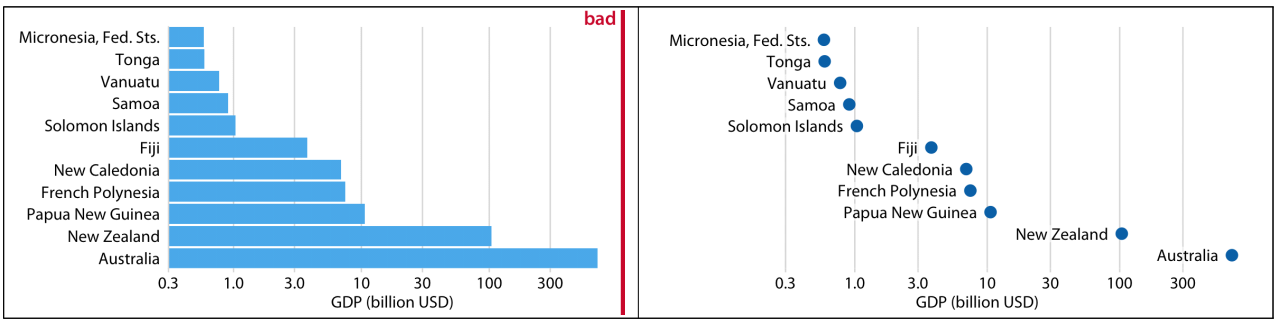
17.1 선형 축 도표 그리기



선형 스케일에 막대 그래프를 그릴 땐 항상 0에서 시작해야 한다.

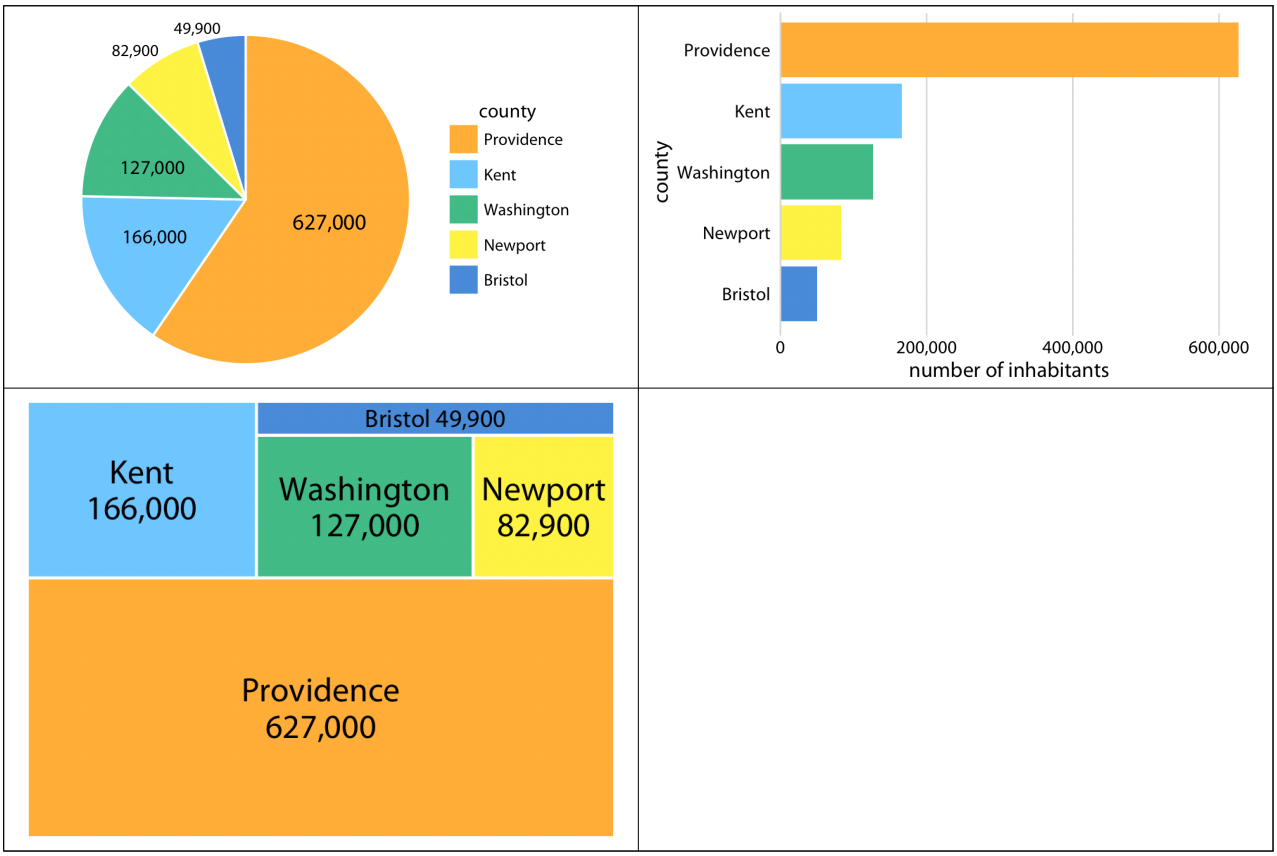


17.2 로그 축이 있는 도표 그리기



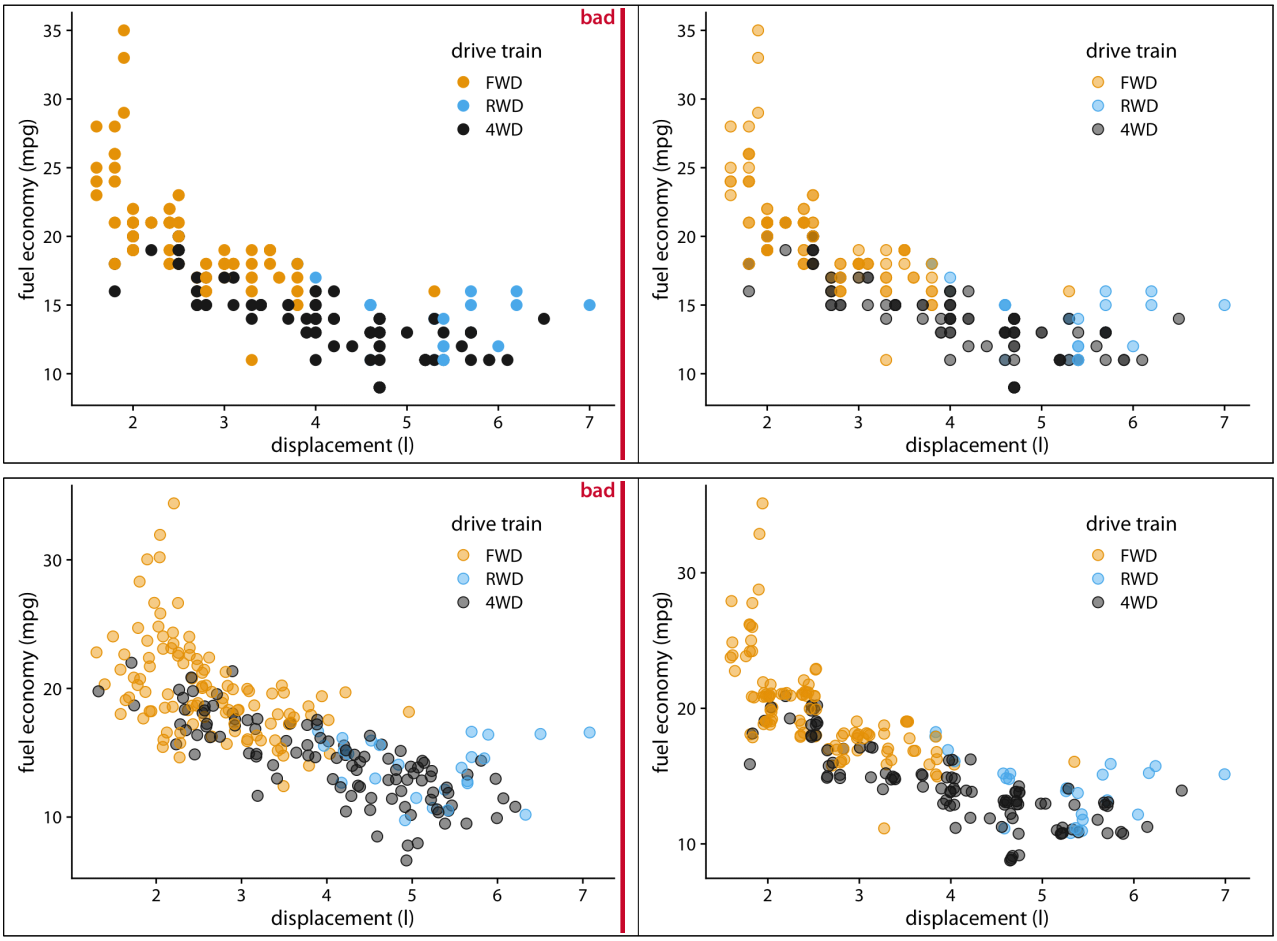
로그 스케일에 그리는 막대는 비율을 비교할 때 쓰이며, 0이 아닌 1에서 출발해야 한다.

17.3 데이터 값을 면적으로 나타내기

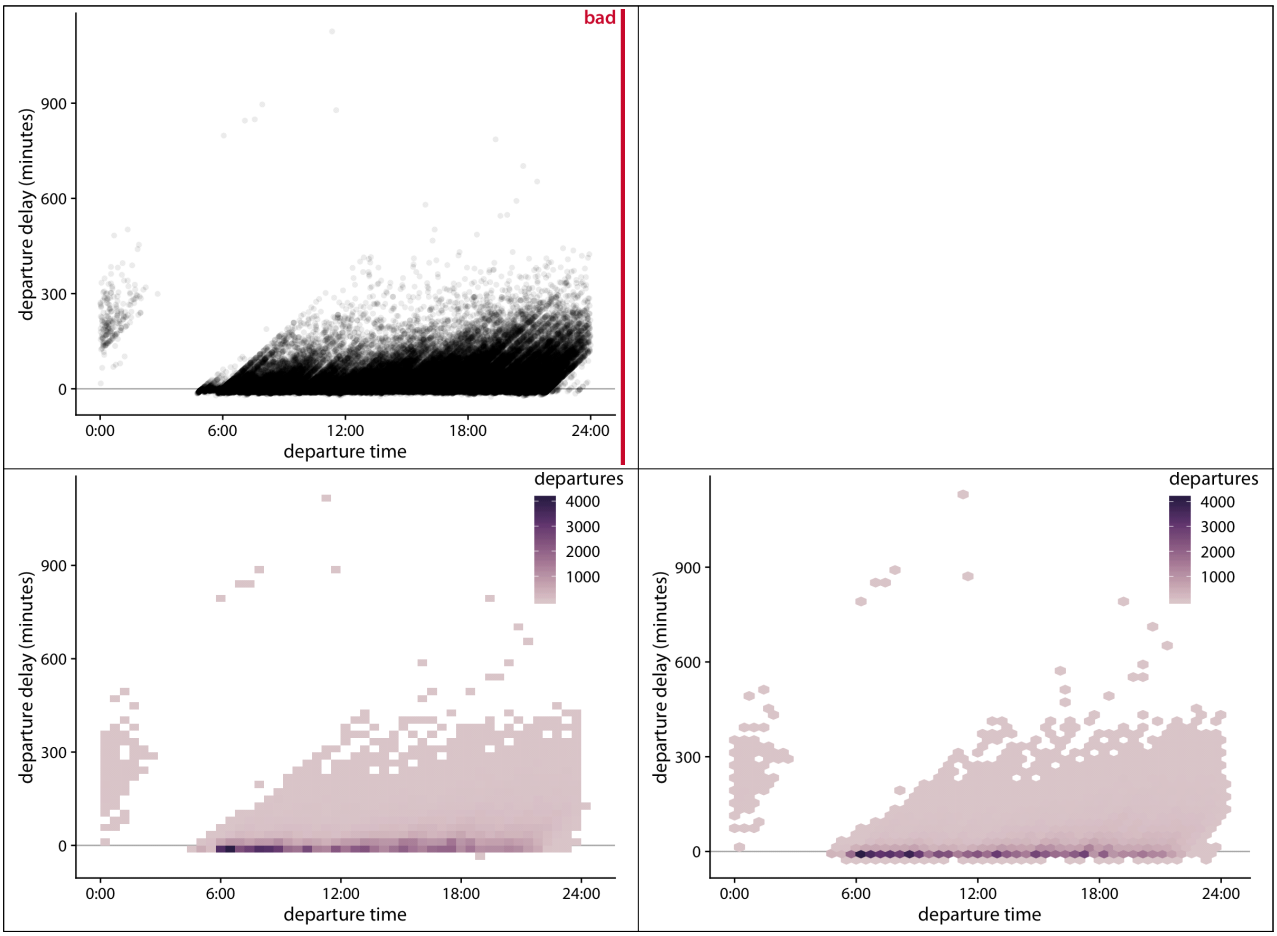


18장 오버플로팅: 겹치는 점 처리 방법

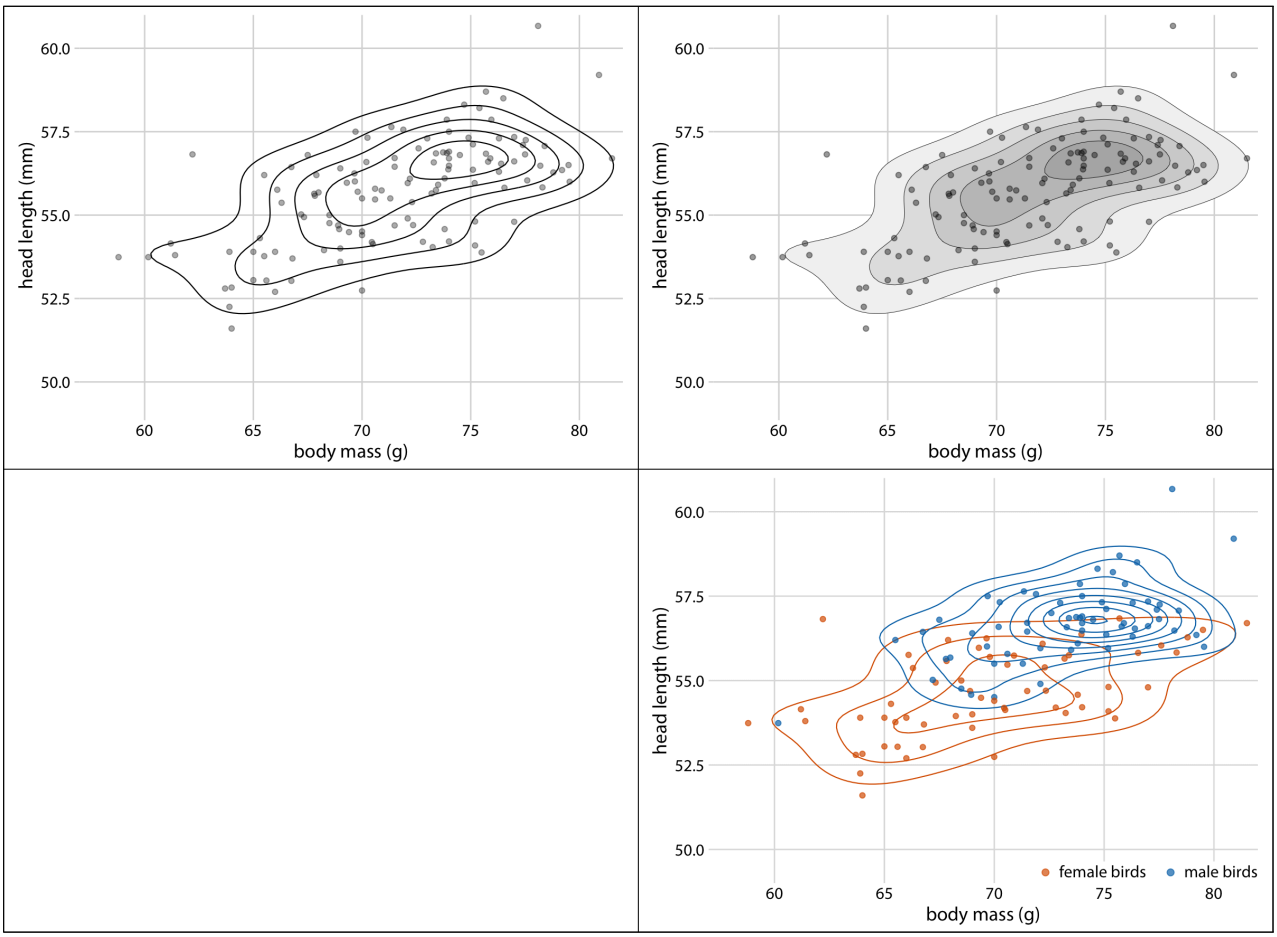
18.1 반투명 값과 지터링

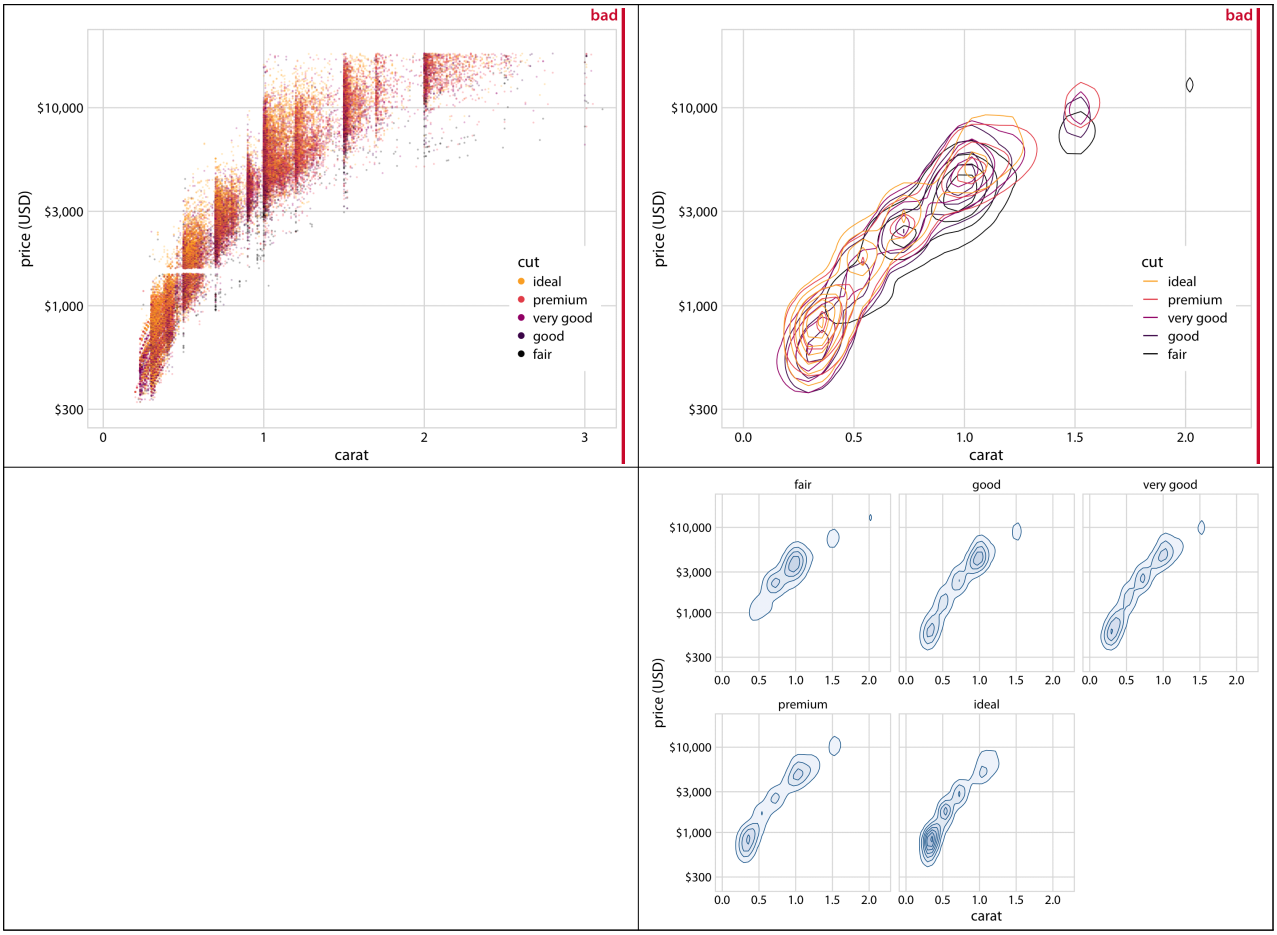


18.2 이차원 히스토그램



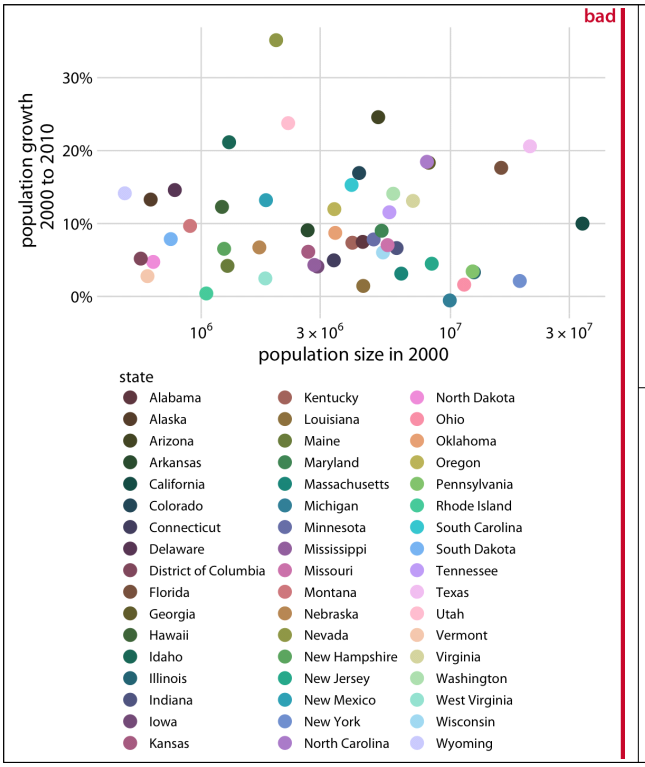
18.3 등고선 그래프

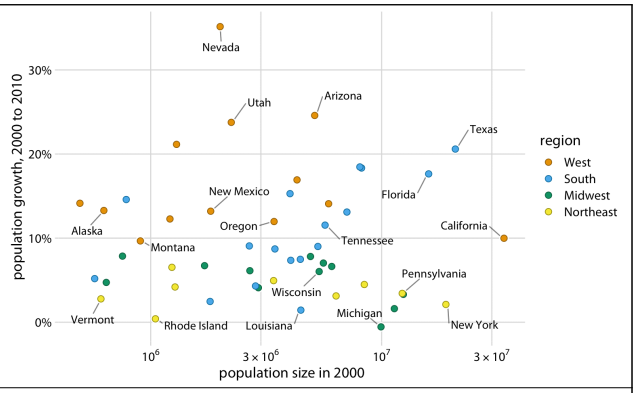

60.0
57.5
55.0
52.5
50.0
head length (mm)
60
65
70
75
80
body mass (g)
female birds
male birds


bad
price (USD)
$10,000
$3,000
$1,000
$300
0
1
2
3
carat
cut
ideal
premium
very good
good
fair
bad
price (USD)
$10,000
$3,000
$1,000
$300
0.0
0.5
1.0
1.5
2.0
carat
cut
ideal
premium
very good
good
fair
fair
good
very good
premium
ideal
price (USD)
$10,000
$3,000
$1,000
$300
0.0
0.5
1.0
1.5
2.0
carat

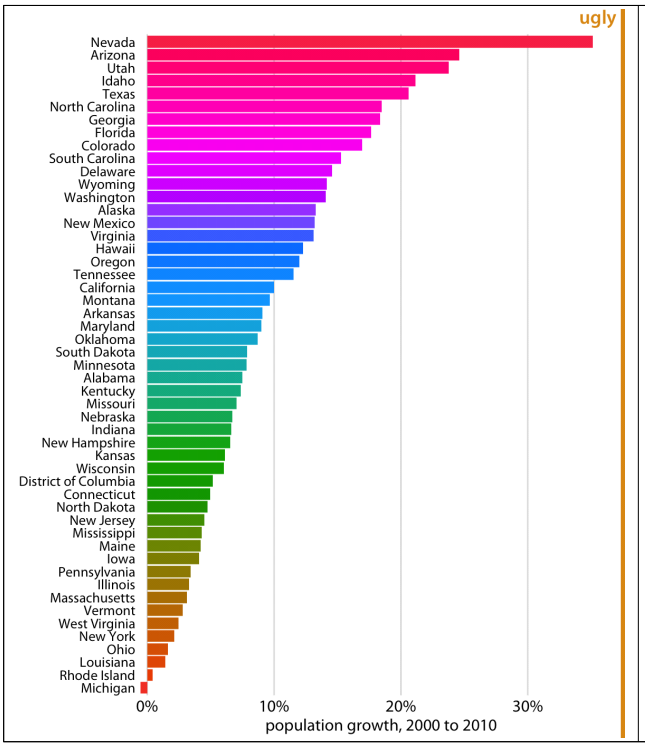
19장 효과적인 색 사용을 위한 조언

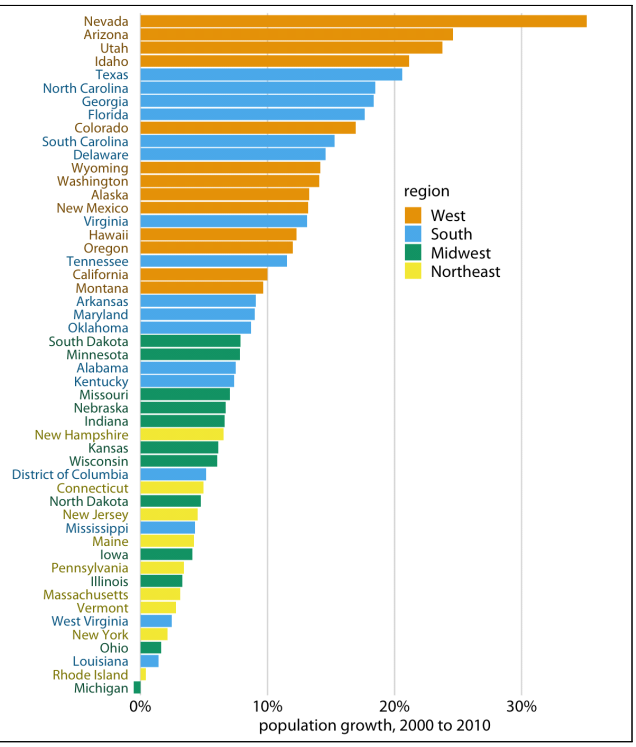
19.1 목적 없는 무분별한 색 사용은 금물





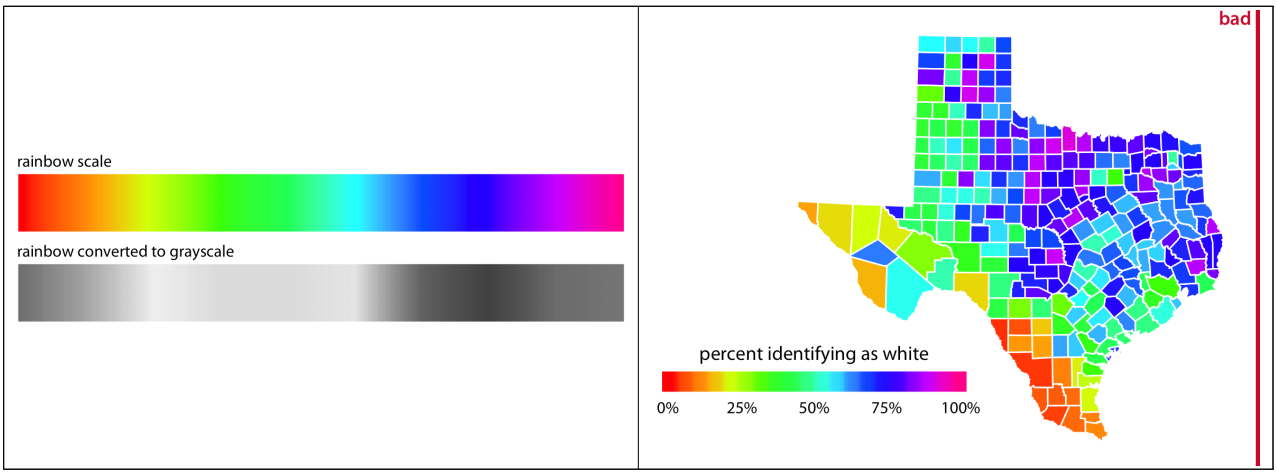
구분해야 할 범주가 8개 이상이라면 데이터 포인트에 색을 입히는 대신 바로 옆에 레이블을 달자.





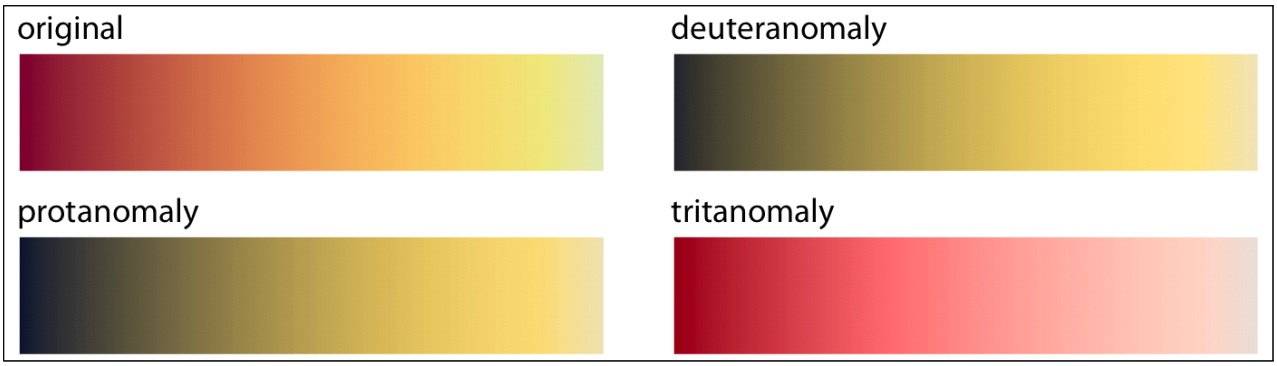
넓은 면적에는 너무 채도가 높은 색을 입히지 않는다. 색이 튀면 독자가 도표를 꼼꼼하게 살피기 힘들기 때문이다.

19.2 비단조적 색상 스케일 구성

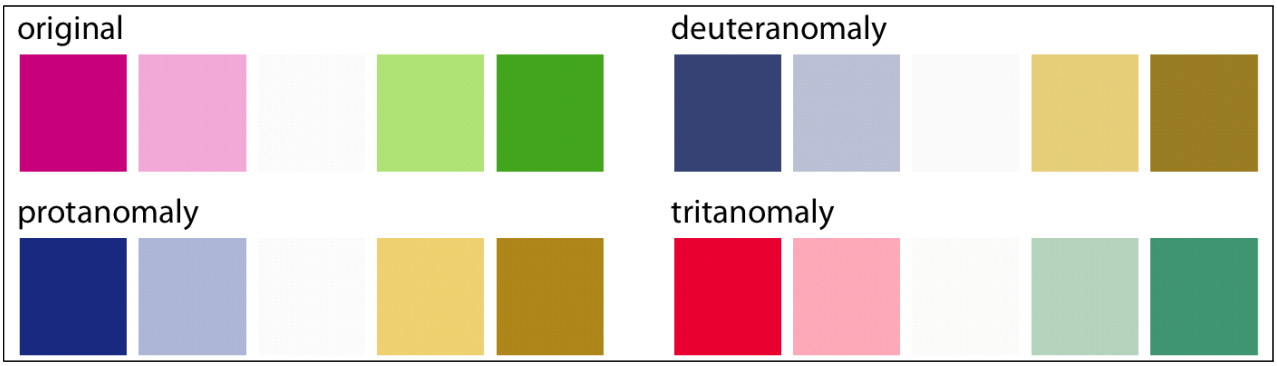


19.3 색각 이상자를 배려하자

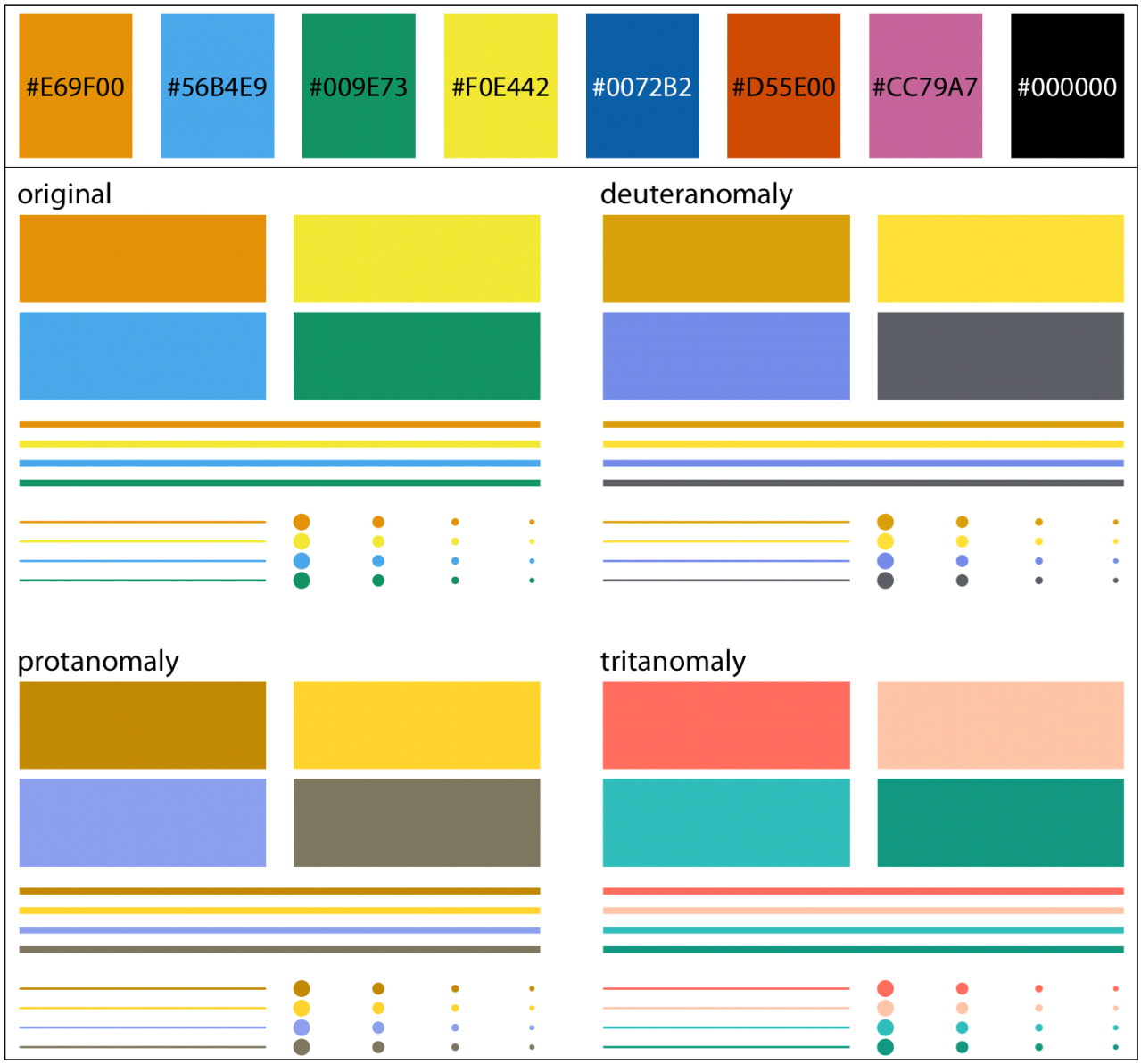
- 짙은 빨간색에서 밝은 노란색으로 변하는 순차적 색상 스케일



- ColorBrewer PiYG (분홍-연두) 스케일



- 모든 색각 이상자가 구별할 수 있는 정성적 색상 모음(Okabe and Ito 2008)



색각 이상자에게 도표를 문제없이 보여주려면 특정 색상 스케일에만 의존하지 말고, 직접 색각 이상 시뮬레이션을 돌려서 자신이 만든 도표를 시험해보자.

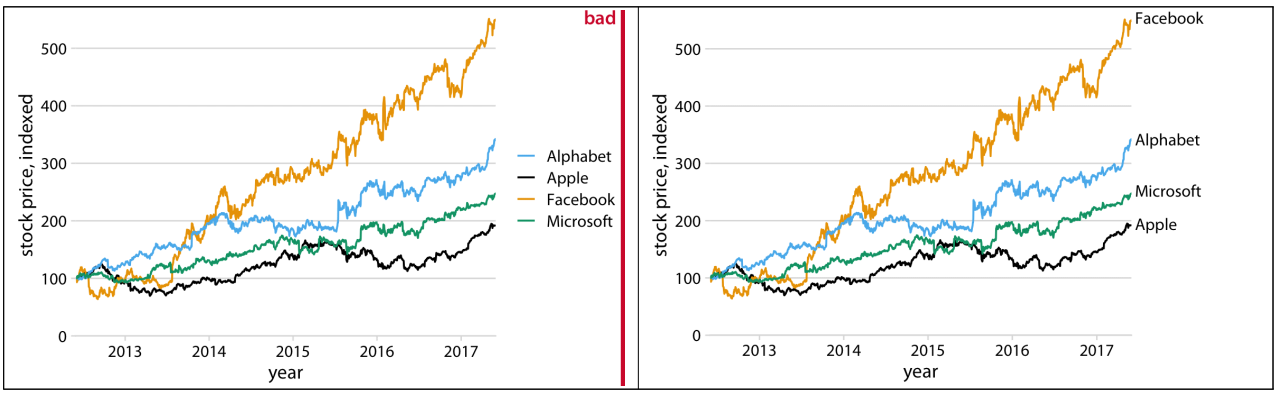
20장 불필요한 기호화

20.1 불필요한 기호화는 범례 디자인을 망친다

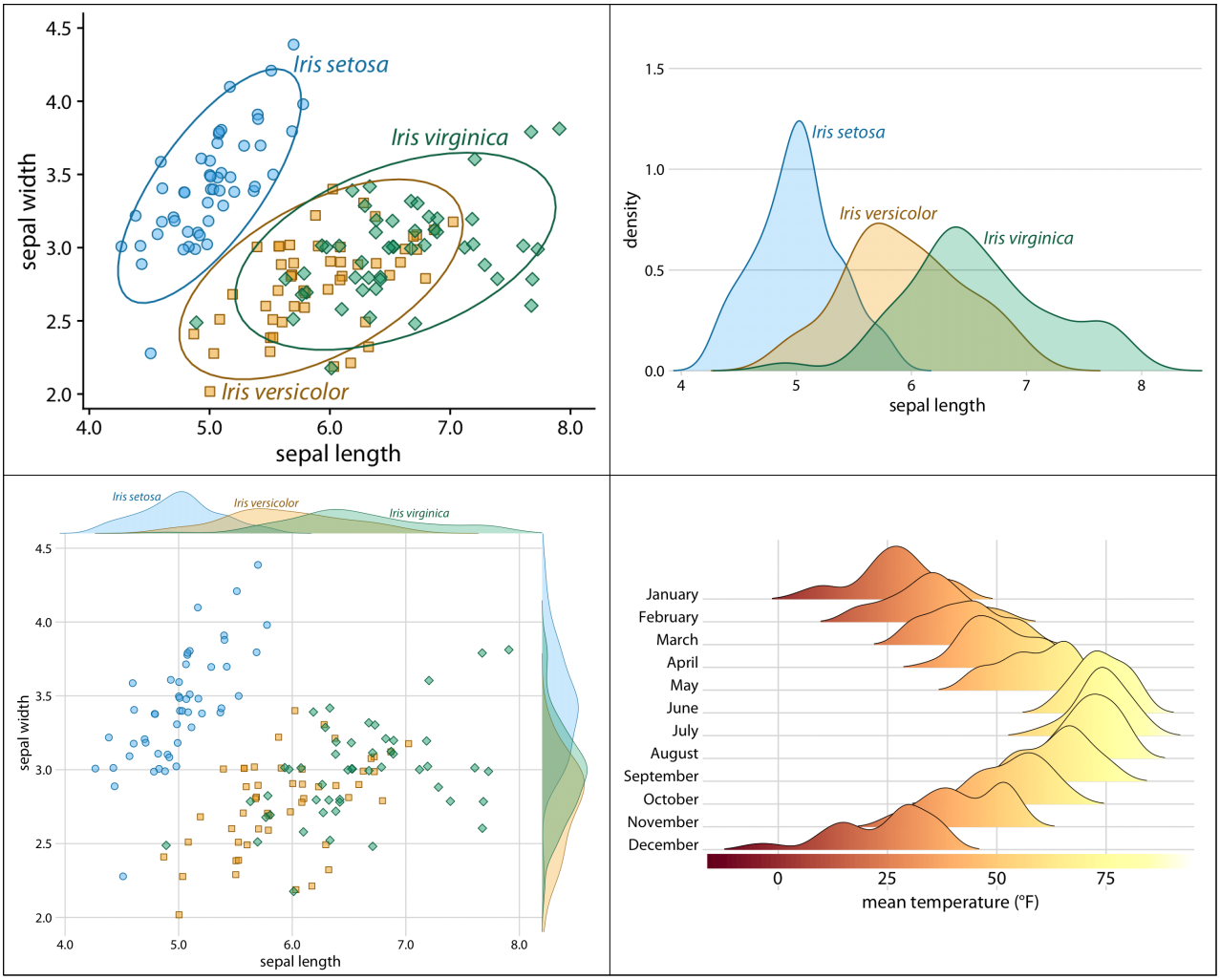


데이터에 시각적 순서가 있으면 범례도 같은 순서로 배열하자.

20.2 범례가 없어도 좋은 도표



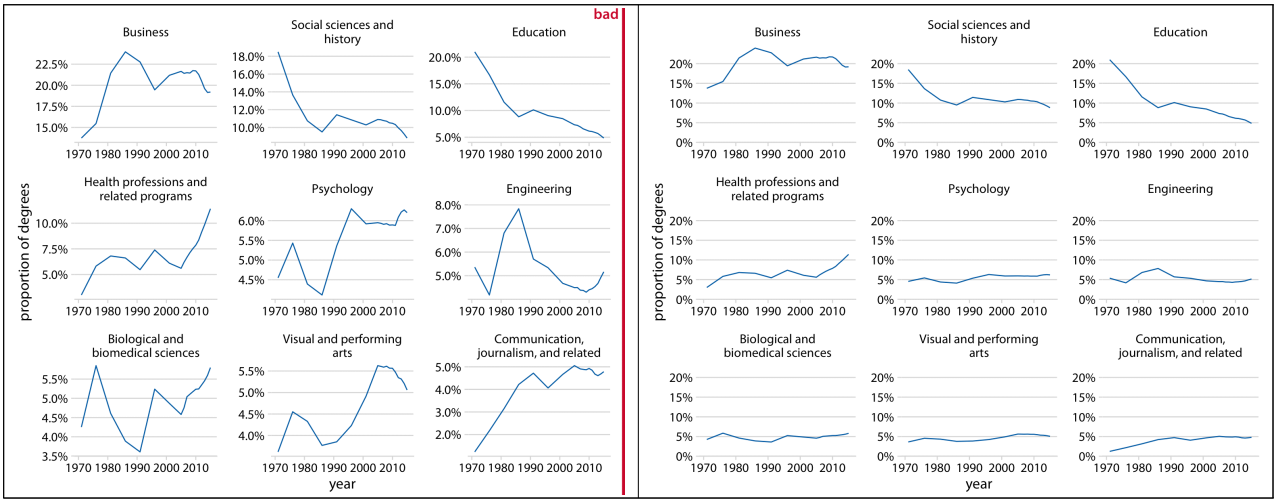
가능하면 범례를 따로 넣을 필요가 없도록 도표를 디자인하자.


Iris setosa
Iris virginica
Iris versicolor
sepal width
sepal length
density
January
February
March
April
May
June
July
August
September
October
November
December
mean temperature (°F)

21장 여러 개의 패널로 구성된 도표

21.1 소형 다중패널

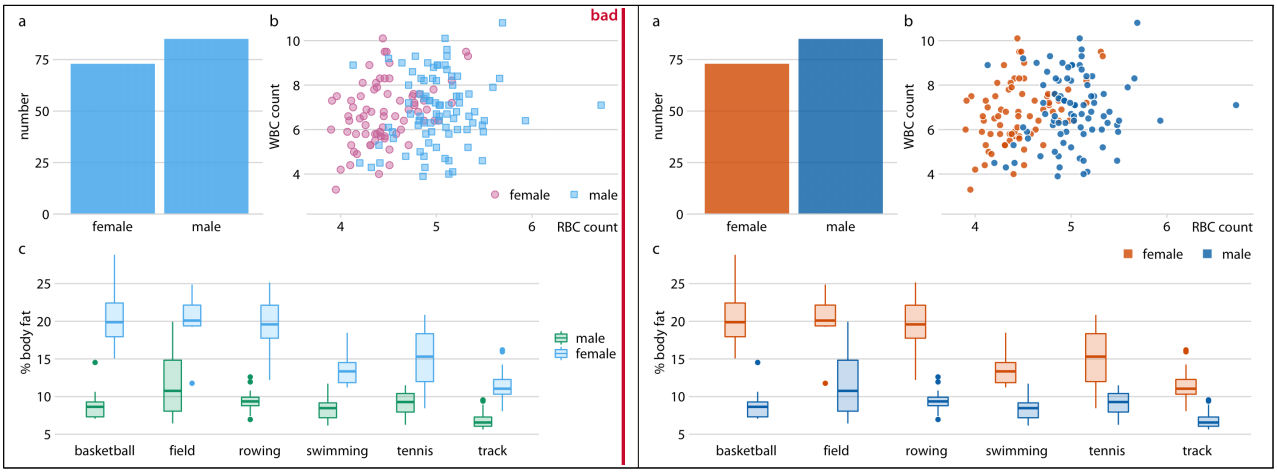
- 모든 패널의 y축을 동일하게 설정



소형 다중패널 도표에 패널들을 배열할 땐 항상 의미 있고 논리적은 순서를 적용하자.

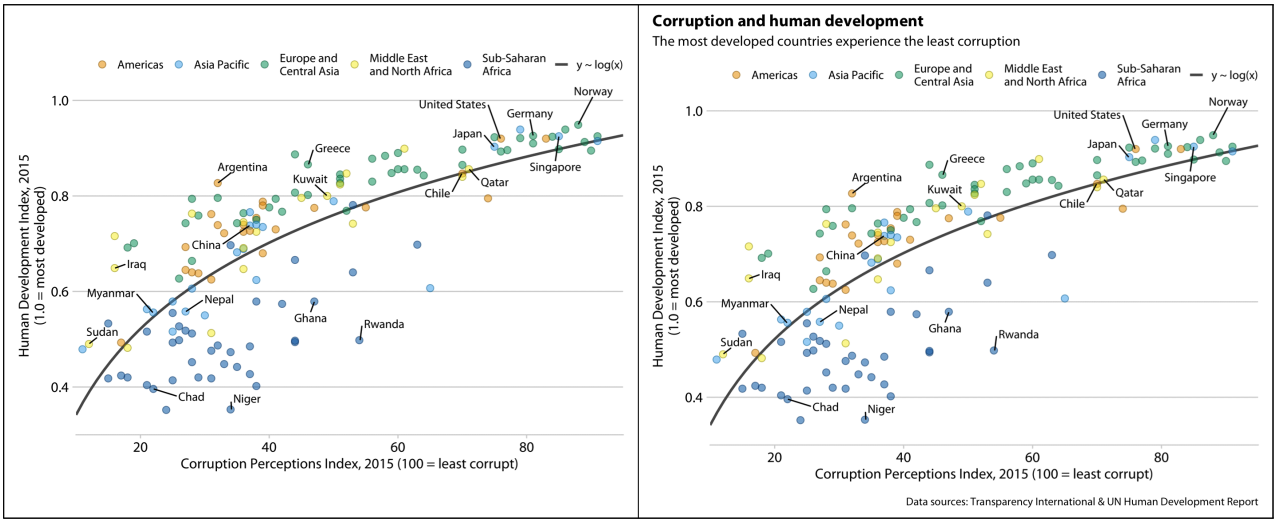
21.2 복합형 도표

- 모든 패널에서 여성 데이터를 왼쪽에 두고, 성별을 같은 색으로 표시



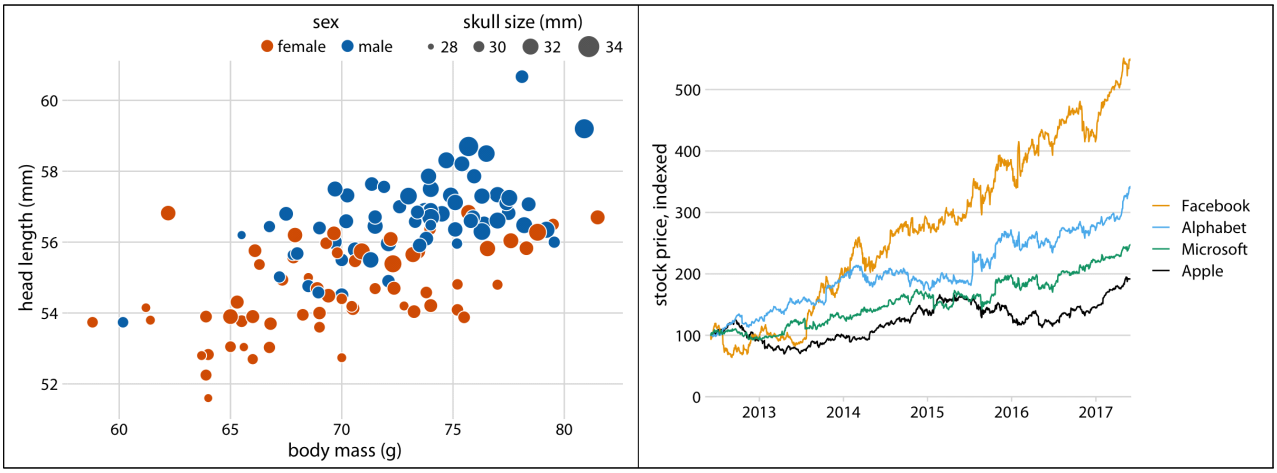
22장 제목, 캡션, 표의 효과적 활용

22.1 도표 제목과 캡션



도표마다 아래쪽에 캡션 상자를 넣는 레이아웃을 사용한다면, 도표 제목은 도표 상단에 넣지 말고 캡션 맨 첫 문장으로 쓰자.

22.2 축과 범례 제목



22.3 표, 제대로 만들기

- 세로 선을 넣지 않는다.
- 데이터의 행과 행 사이에 사로 선을 넣지 않는다(가로 선은 헤더 행과 첫 데이터 행을 구분하는 용도, 또는 표 전체의 테두리 용도로 쓸 수 있다).
- 글자 열은 왼쪽을 기준으로 정렬해야 한다.
- 숫자 열은 오른쪽을 기준으로 정렬하고, 모든 숫자의 소수점 자릿수를 통일한다.
- 열에 숫자나 문자가 1글자만 들어갈 경우에는 가운데로 정렬한다.
- 헤더 칸은 아래쪽 데이터 열과 정렬을 맞춰야 한다. 즉 문자 열의 헤더 칸은 왼쪽 정렬, 숫자 열의 헤더 칸은 오른쪽 정렬이다.

a

ugly

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

b

ugly

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

c

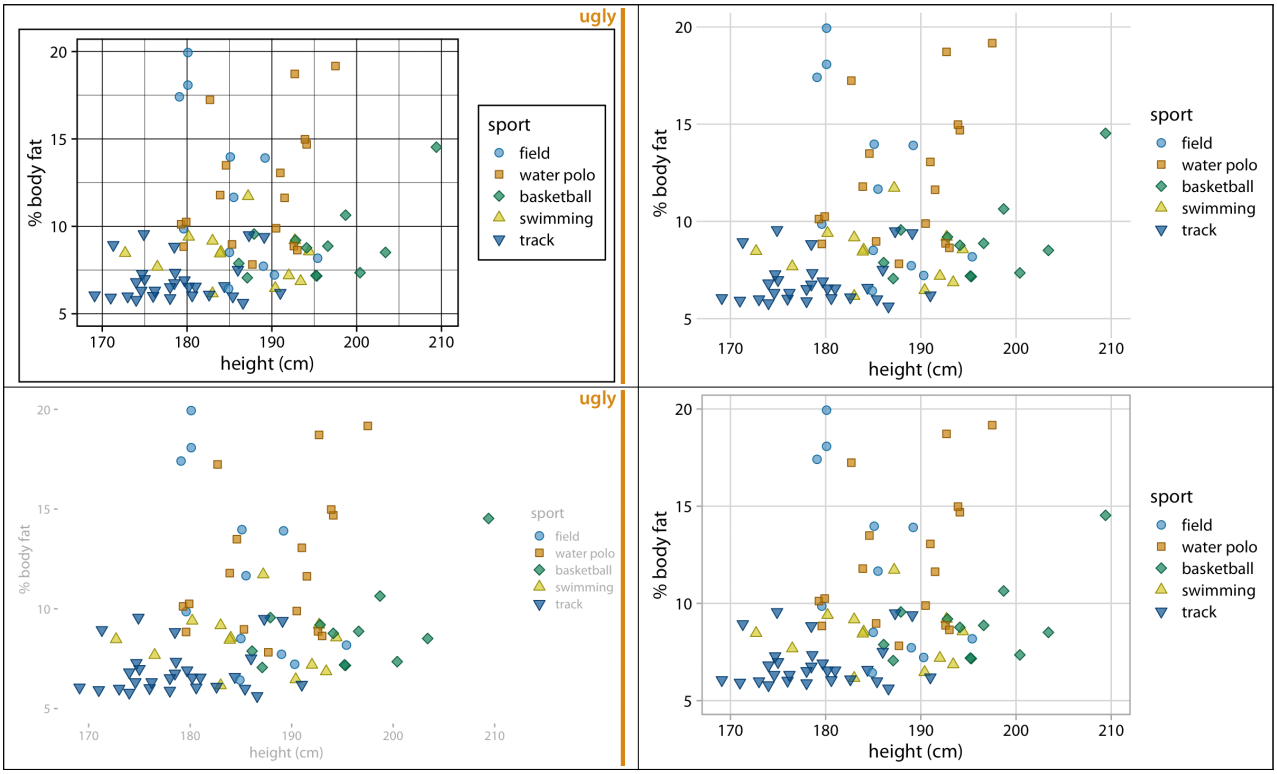
Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

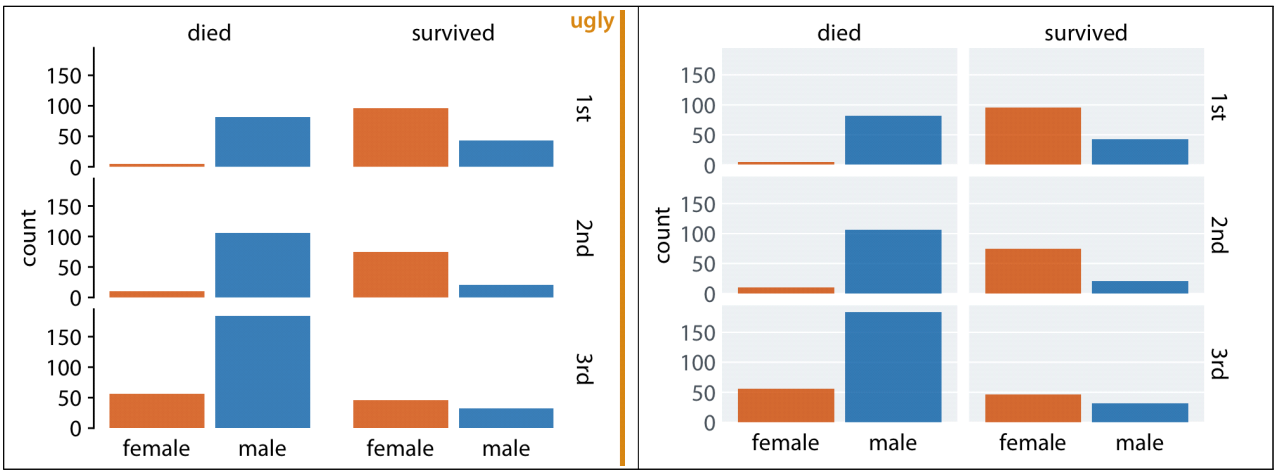
d

Rank	Title	Amount
1	*Star Wars: The Last Jedi*	$71,565,498
2	*Jumanji: Welcome to the Jungle*	$36,169,328
3	*Pitch Perfect 3*	$19,928,525
4	*The Greatest Showman*	$8,805,843
5	*Ferdinand*	$7,316,746

23장 데이터 이해를 돕는 시각화의 주변 요소

23.1 적절한 수준의 요소 활용



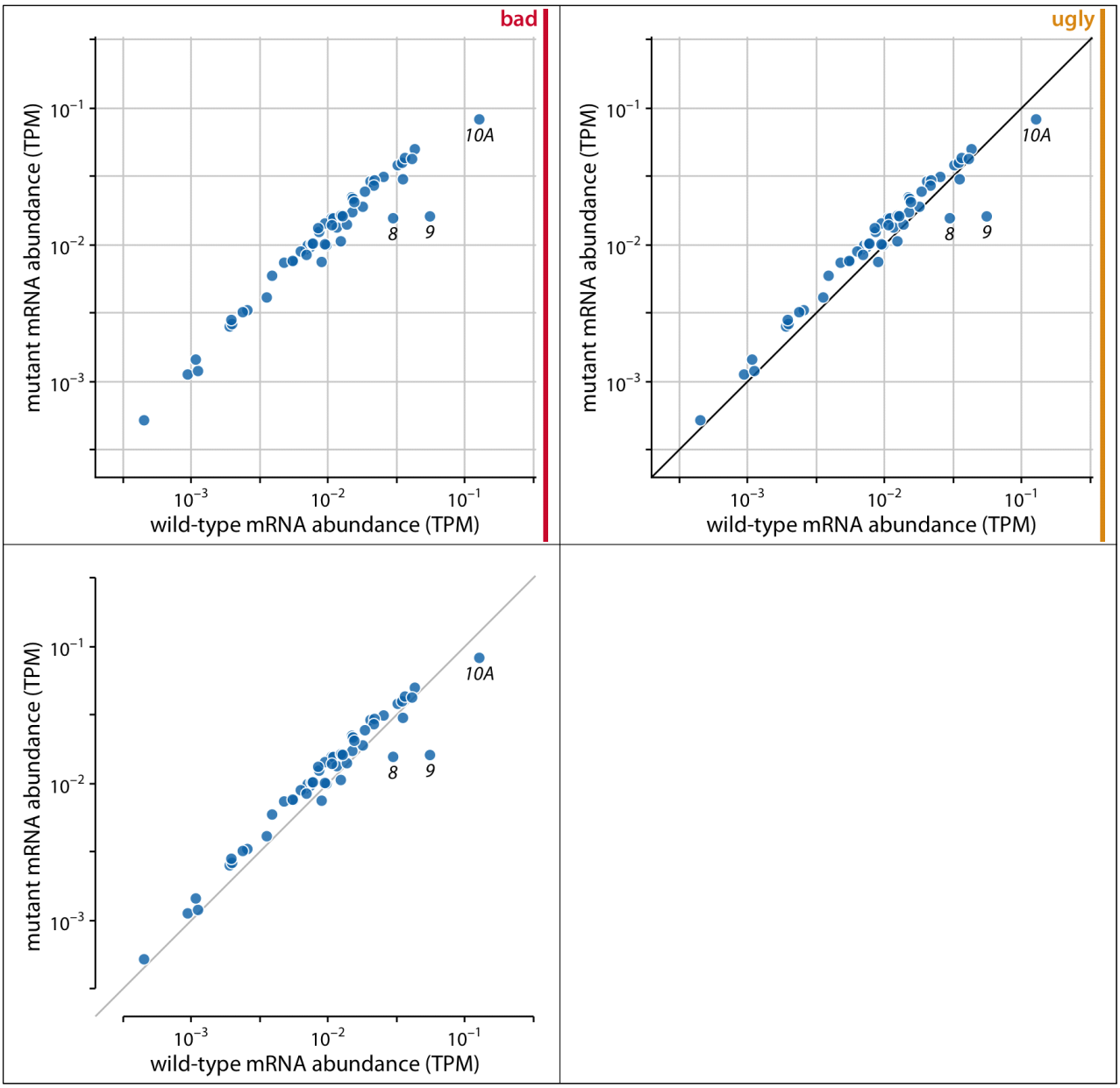


23.2 배경 격자

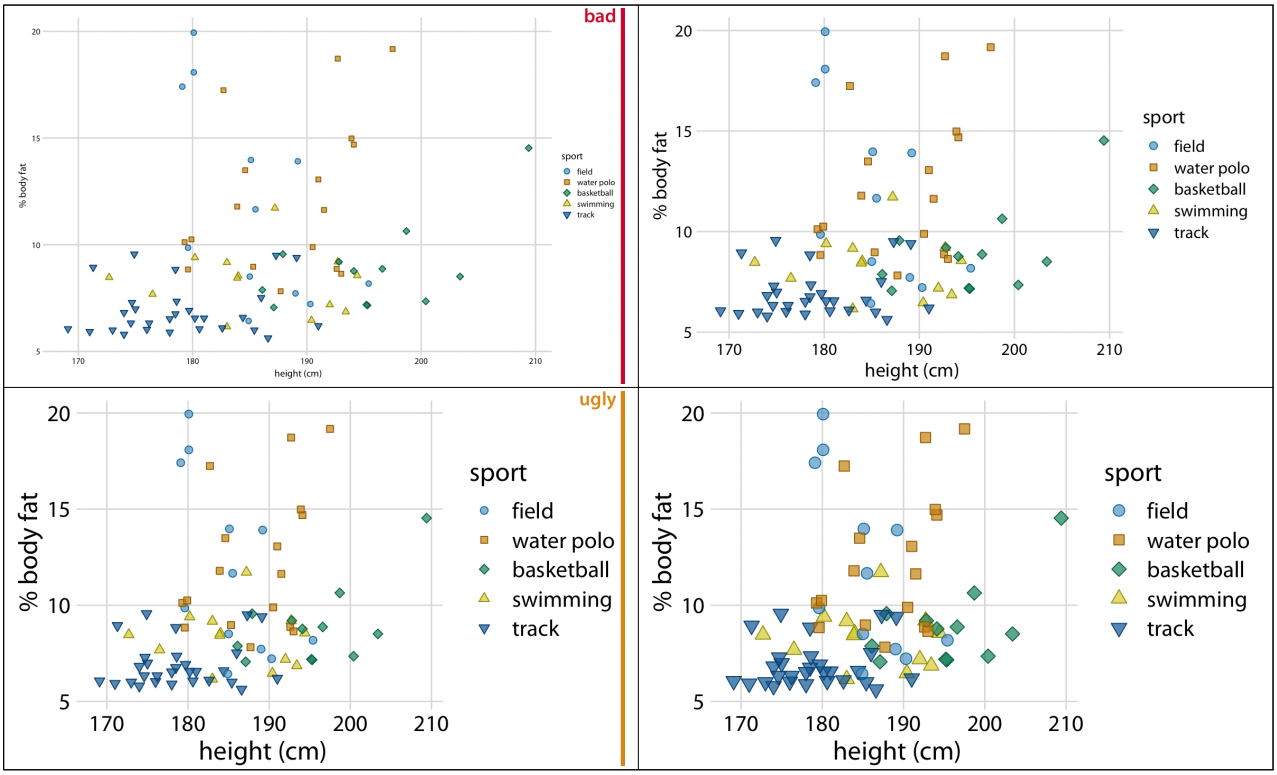


중요한 정보를 전달하는 변수의 진행 방향과 직각을 이루도록 격자선을 긋는 것이 가장 효과적이다.

23.3 쌍 데이터

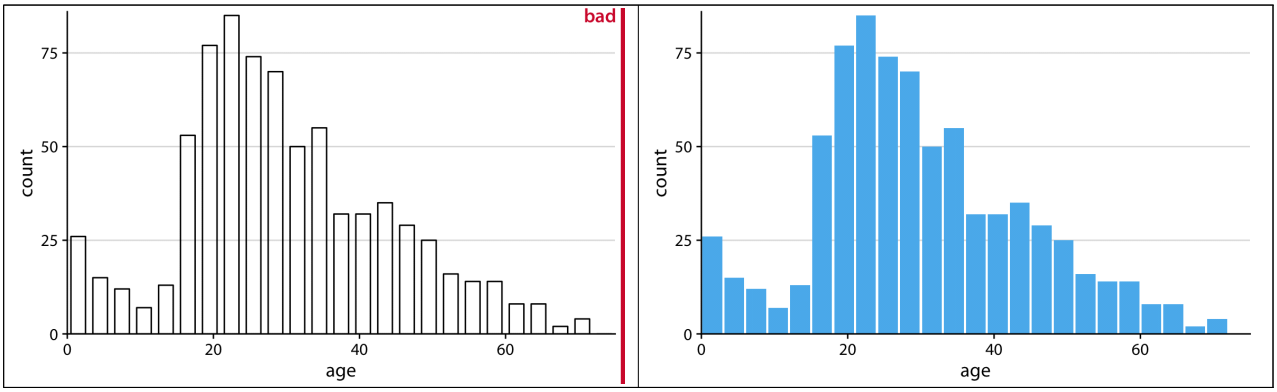

bad
ugly
mutant mRNA abundance (TPM)
wild-type mRNA abundance (TPM)
10A
8
9
10^{-1}
10^{-2}
10^{-3}

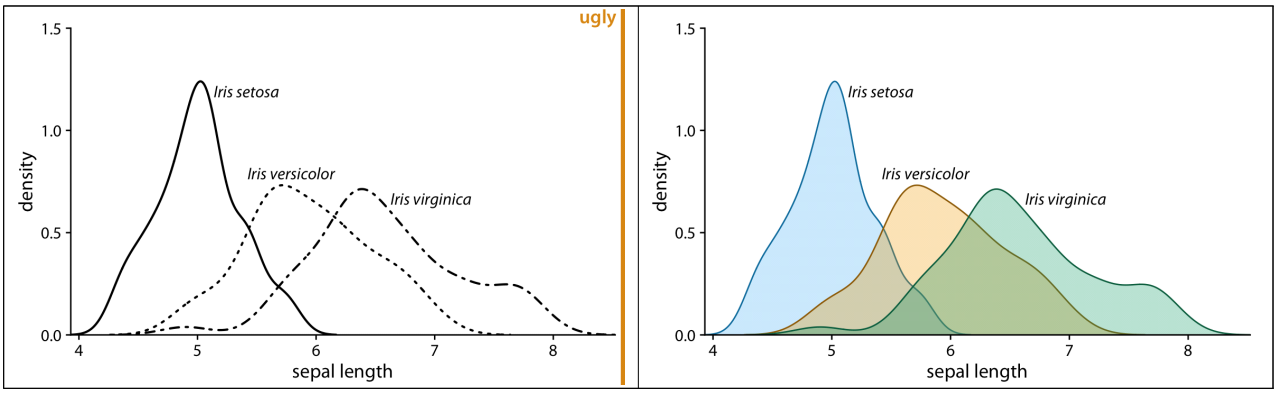
24장 축 레이블의 글자는 큼직하게

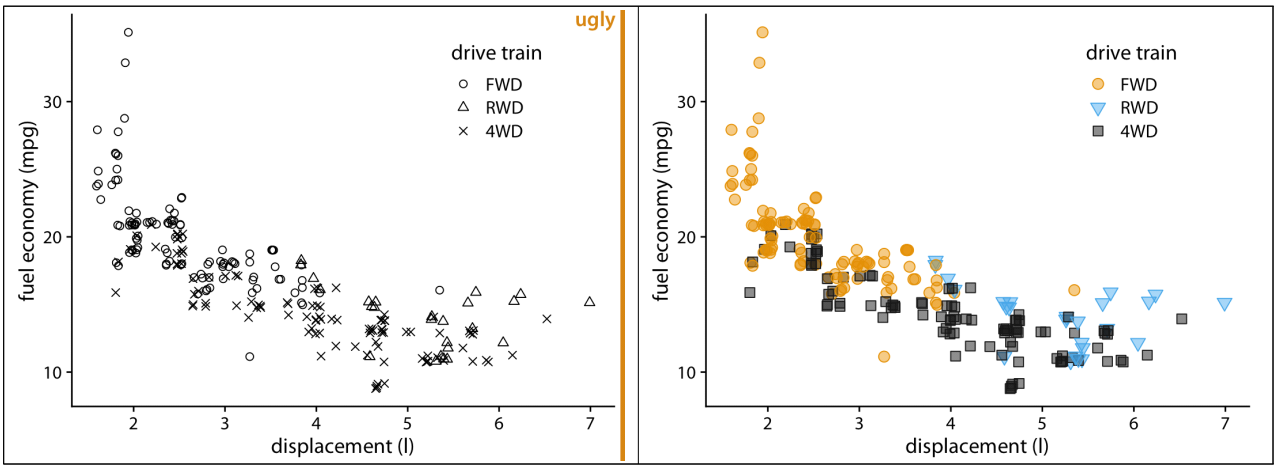


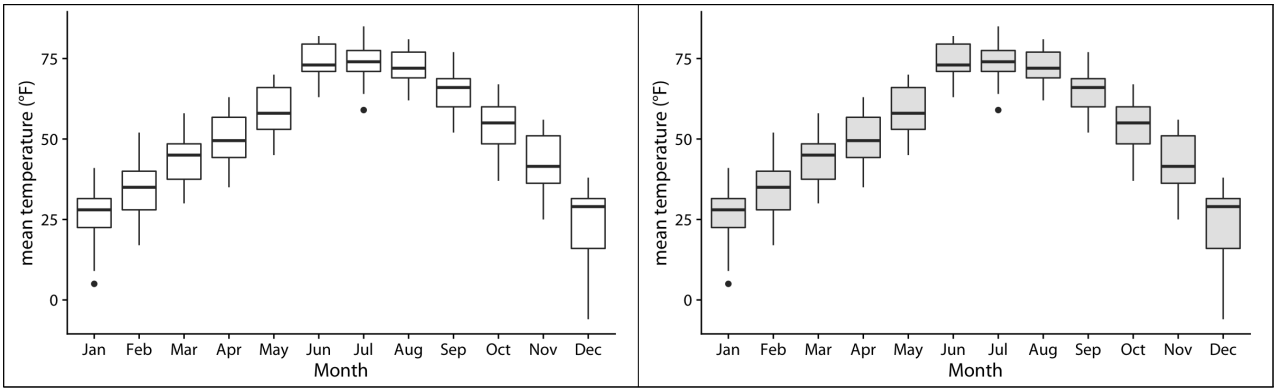
항상 도표를 축소해보고 축 레이블의 크기가 적당한지 확인하자.

25장 선 그림은 피하자


bad
count
age
0
25
50
75
0
20
40
60

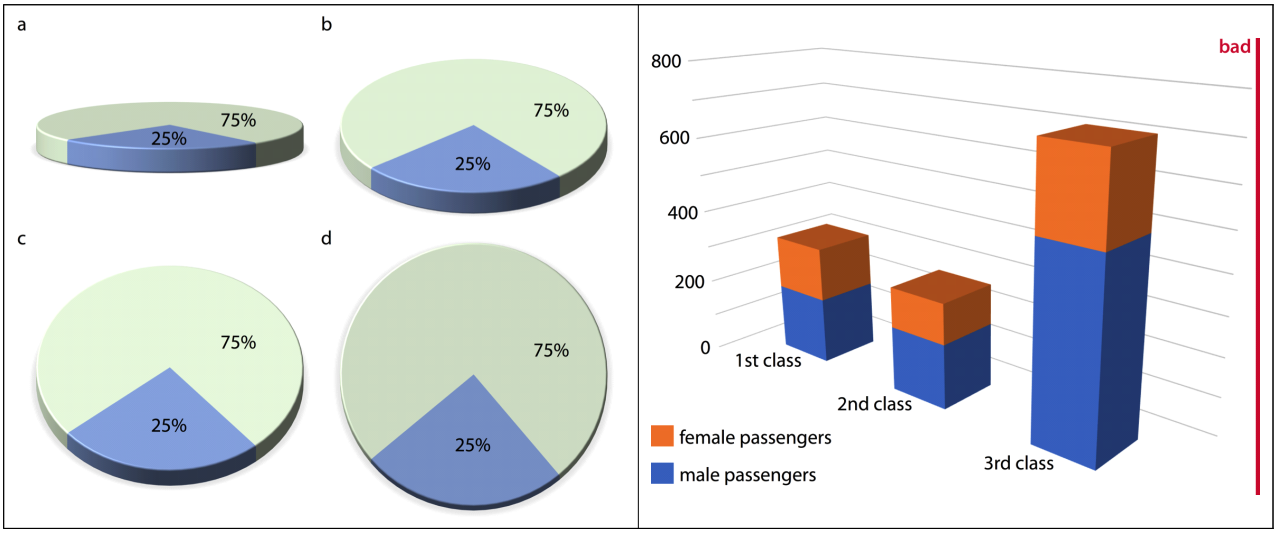

ugly
density
sepal length
Iris setosa
Iris versicolor
Iris virginica
0.0
0.5
1.0
1.5
4
5
6
7
8


ugly
fuel economy (mpg)
displacement (l)
drive train
FWD
RWD
4WD
10
20
30
2
3
4
5
6
7

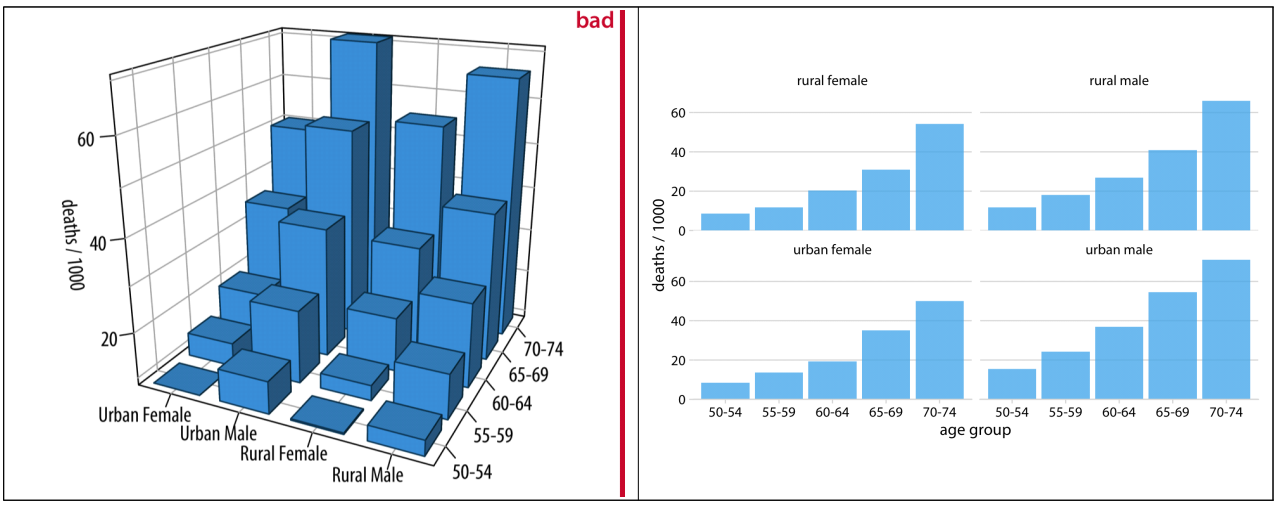

mean temperature (°F)
0
25
50
75
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
mean temperature (°F)
0
25
50
75
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

26장 3차원 그래픽과 차트는 그만

26.1 불필요한 3차원 그래픽은 지양하자



26.2 3차원 위치 스케일도 이제 그만



26.3 3차원 시각화가 필요한 경우

