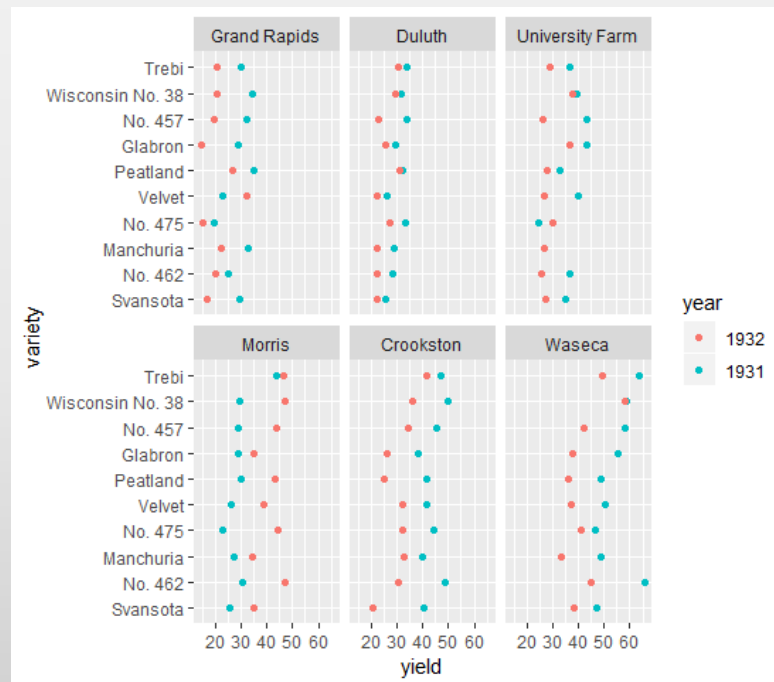


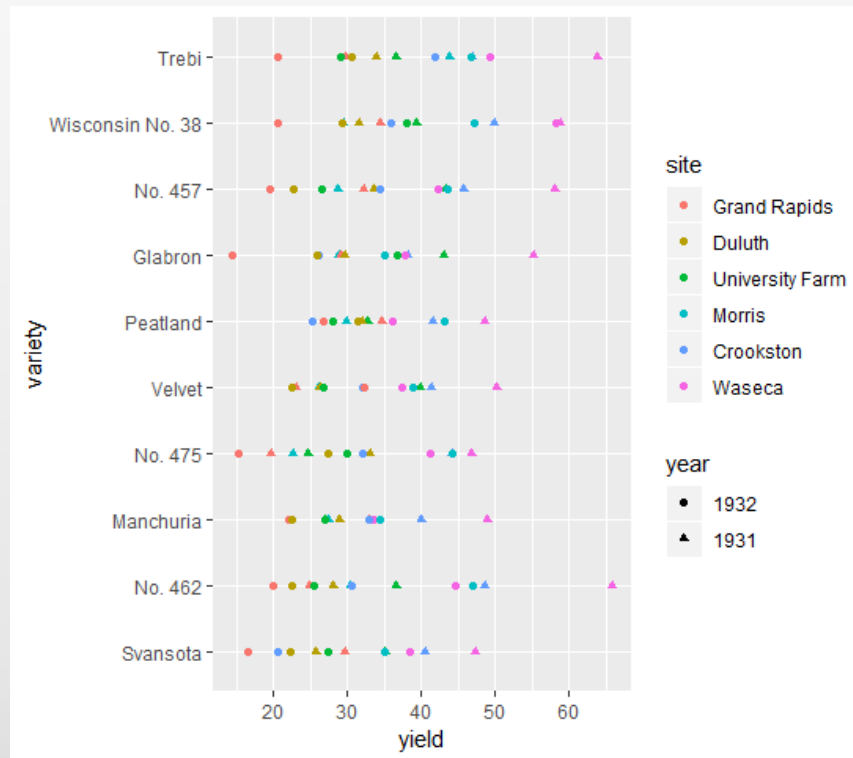
- 연습문제

1. 패키지 `lattice`에 있는 데이터 프레임 `barley`는 미네소타 주 농경학자들이 보리 종류에 따른 수확량의 차이를 비교하기 위해 2년간 경작하여 얻은 자료이다. 설명변수로는 6군데 경작지(site), 10종류의 보리(variety), 경작 년도(year)이고 반응변수는 수확량(yield)이다.

1) 세 설명변수의 조합에 따른 수확량의 분포를 알아보는 다음의 그래프를 작성해 보자.



2) 보리 종류(variety)에 따른 수확량(yield)의 비교분석에서 경작지(site)와 년도(year)의 효과를 단순 반복으로 처리한 다음의 그래프를 작성해보자.



3) 각 보리종류(variety)의 평균 수확량을 계산하여 다음과 같이 크기 순으로 나타내자.

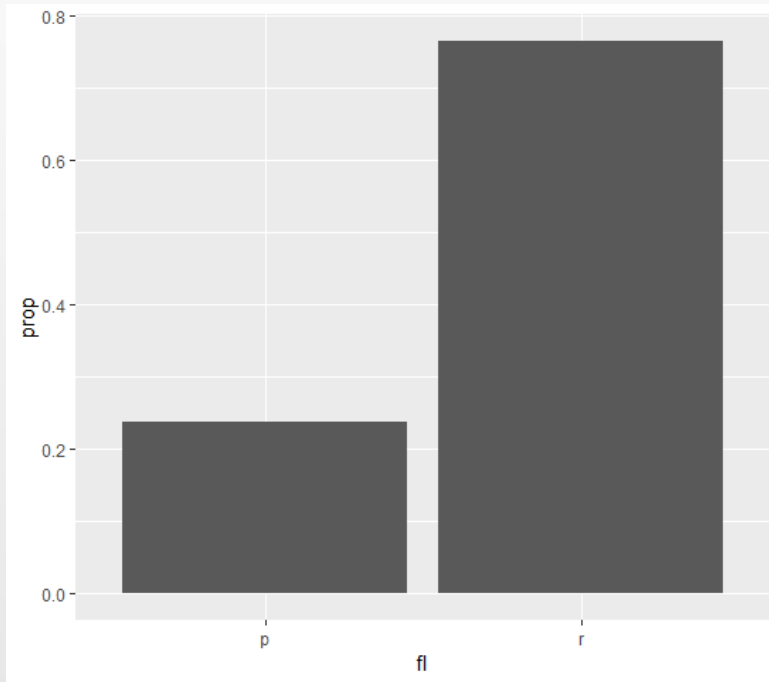
```
# A tibble: 10 x 2
  variety          mean_yield
  <fct>          <dbl>
1 Trebi          39.4
2 Wisconsin No. 38 39.4
3 No. 457         35.8
4 No. 462         35.4
5 Peatland        34.2
6 Glabron         33.3
7 velvet         33.1
8 No. 475         31.8
9 Manchuria       31.5
10 Svansota       30.4
```

2. 데이터 프레임 mpg의 변수 hwy는 자동차의 고속도로 연비를 나타낸다. 범주형 변수인 f1(연료 종류), trans(변속기 종류)에 따른 hwy의 분포를 알아보자.

1) 변수 f1의 종류별 빈도를 다음과 같이 구하라.

```
# A tibble: 5 x 2
  f1      n
  <chr> <int>
1 c         1
2 d         5
3 e         8
4 p        52
5 r       168
```

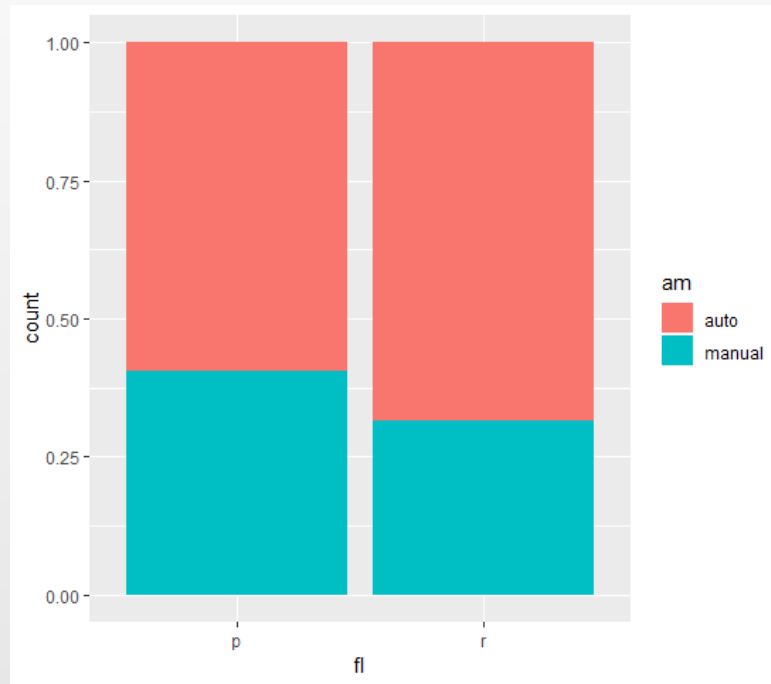
2) 변수 f1에서 c(CNG), d(Diesel), e(Ethanol)는 제외하고 p(Premium)와 r(Regular)만을 대상으로 상대도수에 의한 막대 그래프를 작성해 보자.



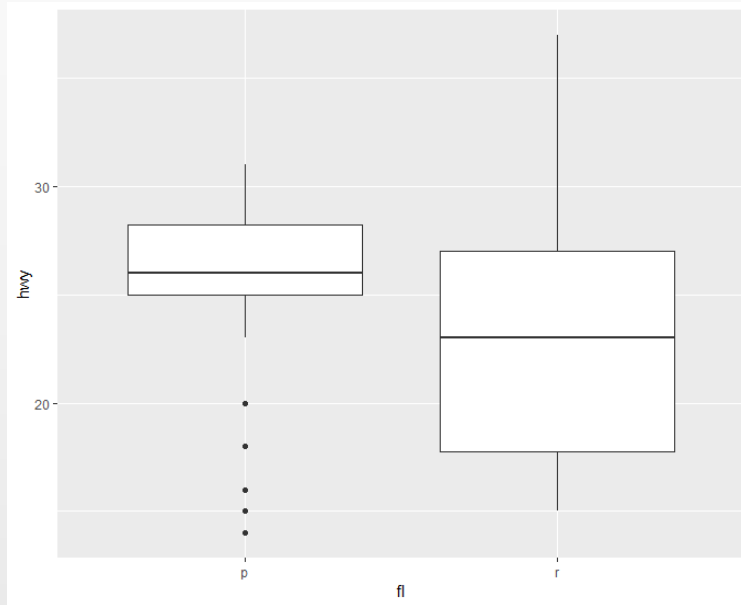
3) 변수 trans의 종류별 빈도를 다음과 같이 구하라.

```
# A tibble: 10 x 2
  trans      n
  <chr>    <int>
1 auto(av)      5
2 auto(l3)      2
3 auto(l4)     78
4 auto(l5)     33
5 auto(l6)      6
6 auto(s4)      3
7 auto(s5)      3
8 auto(s6)     16
9 manual(m5)    56
10 manual(m6)   18
```

- 4) 변수 `trans`의 범주를 'auto'와 'manual'로 통합한 변수 `am`을 생성하고 변수 `f1`과의 관계를 다음과 같은 막대 그래프로 나타내 보자.



5) 변수 `f1`에 따른 `hwy`의 분포를 상자그림으로 나타내 보자.





6) 변수 am과 f1에 따른 hwy의 분포 비교를 위해 상자그림을 작성해 보자.

