

CH5

2023-07-31

```
library(tidyverse)
dat <- read.csv("C:\\Users\\phl02\\Desktop\\P\\bio\\ch5\\Ch5_chb.csv")
dat1<-dat
```

```
dim(dat1)
```

```
## [1] 30 21
```

```
colnames(dat1)
```

```
## [1] "id"      "gender"  "age"     "treat_gr" "lc"      "hcc"
## [7] "b_alt"   "b_plt"   "b_alb"   "m6_alt"   "m6_plt"   "m6_alb"
## [13] "m12_alt" "m12_plt" "m12_alb" "m18_alt"  "m18_plt"  "m18_alb"
## [19] "m24_alt" "m24_plt" "m24_alb"
```

1. R base 그래프

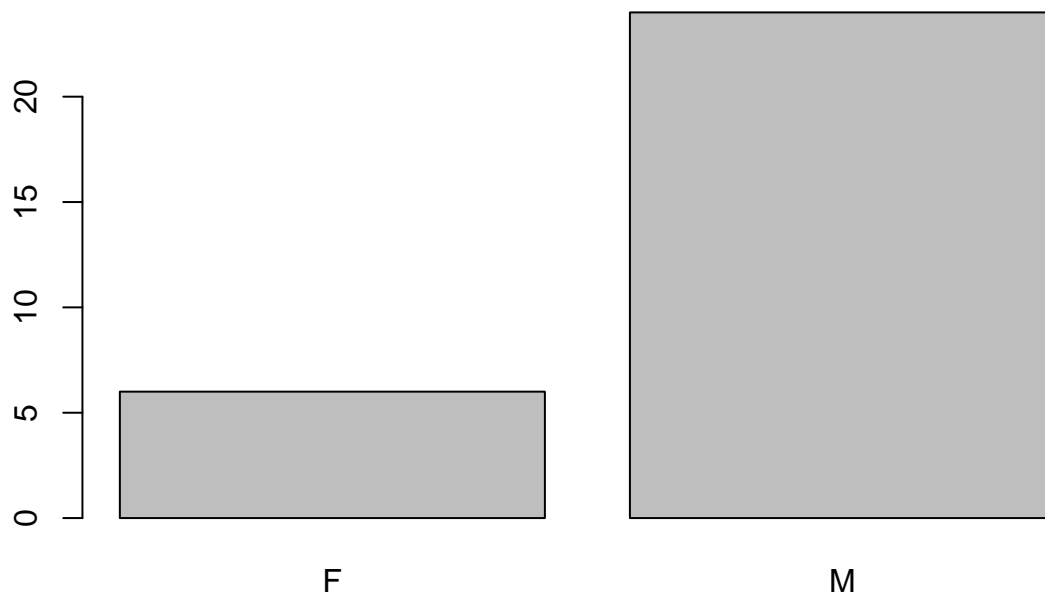
1.1 기본 그래프 그려보기

```
table(dat1$gender)
```

```
##  
##  F  M  
##  6 24
```

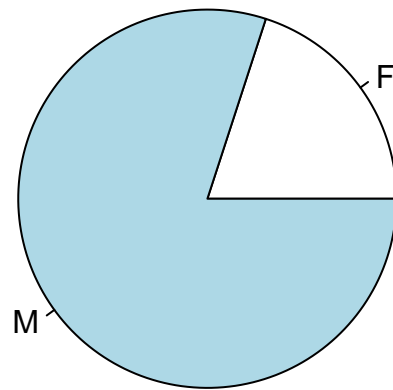
(1) 막대 그래프

```
barplot(table(dat1$gender))
```



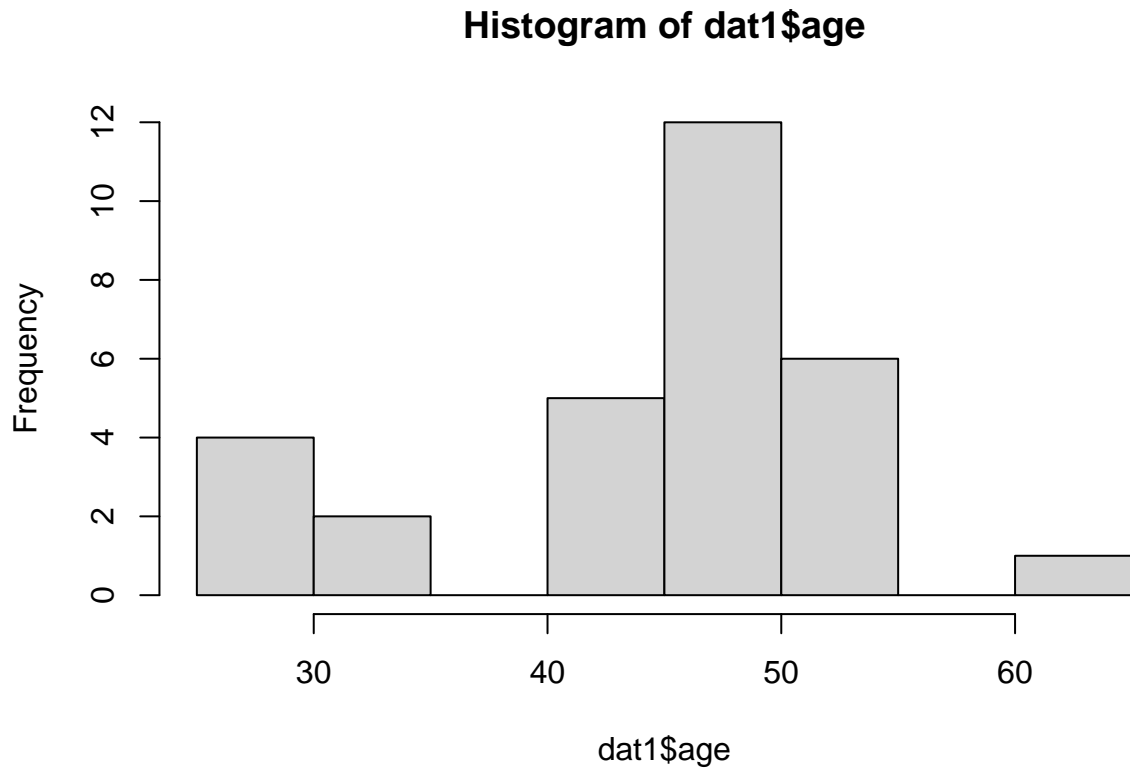
(2) 파이 그래프

```
pie(table(dat1$gender))
```



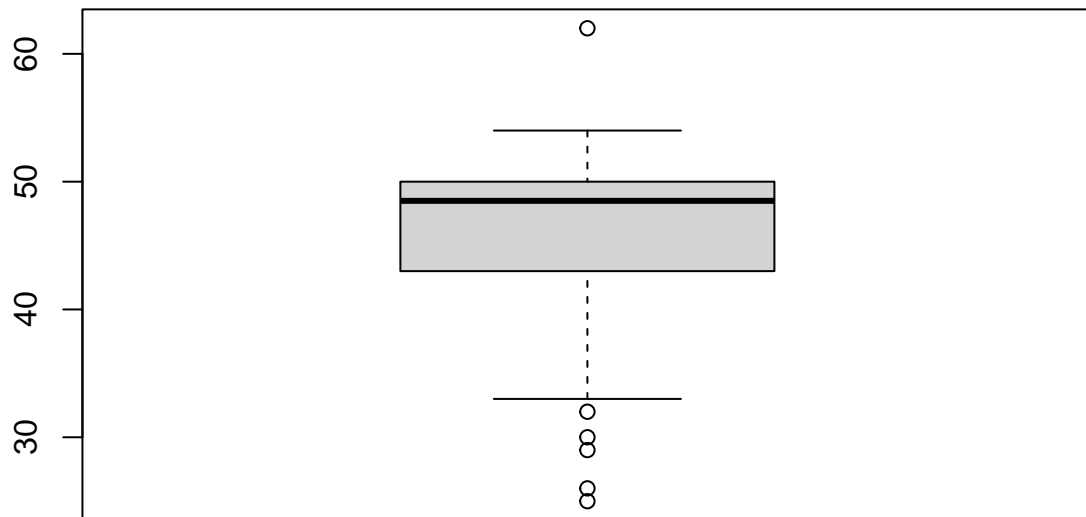
(3) 히스토그램

```
hist(dat1$age)
```



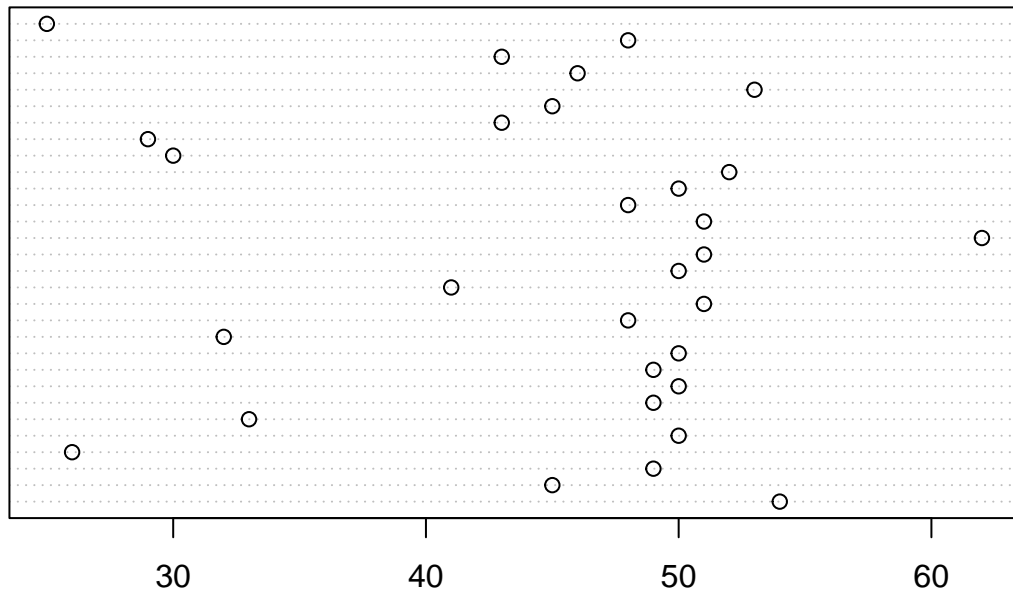
(4) 박스 그래프

```
boxplot(dat1$age)
```



(5) 점 그래프

```
dotchart(dat1$age)
```



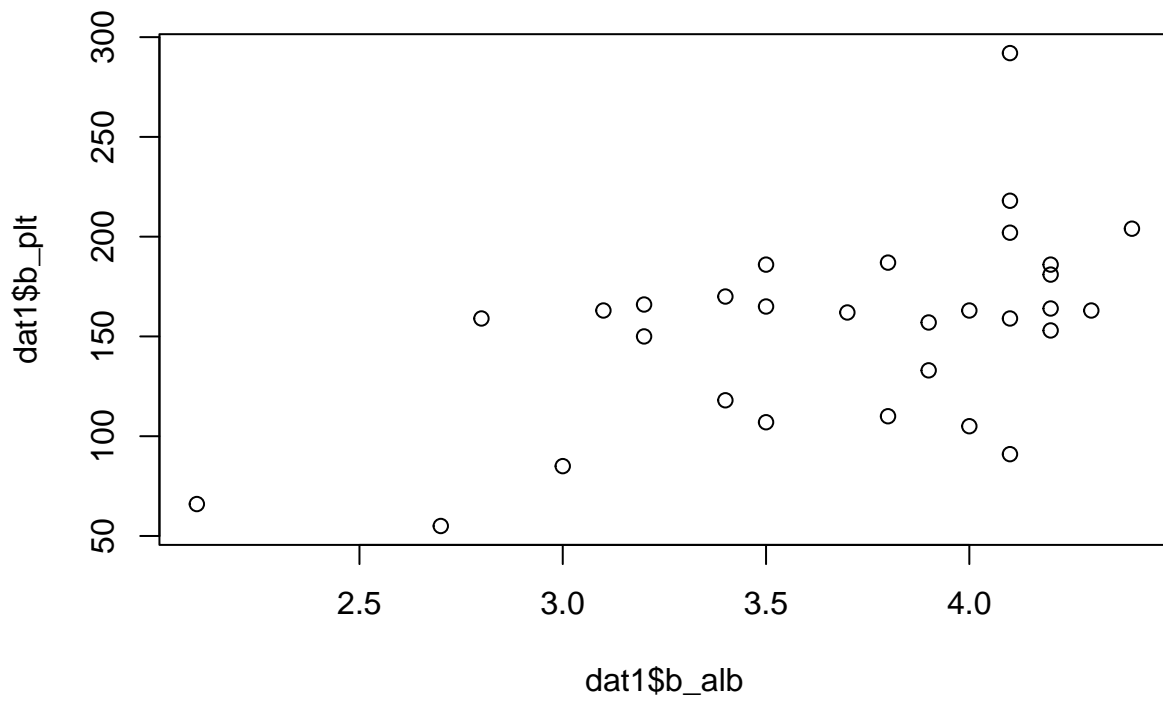
(6) 줄기 앞 그래프

```
stem(dat1$age)
```

```
##  
## The decimal point is 1 digit(s) to the right of the |  
##  
## 2 | 569  
## 3 | 023  
## 3 |  
## 4 | 133  
## 4 | 556888999  
## 5 | 00000111234  
## 5 |  
## 6 | 2
```

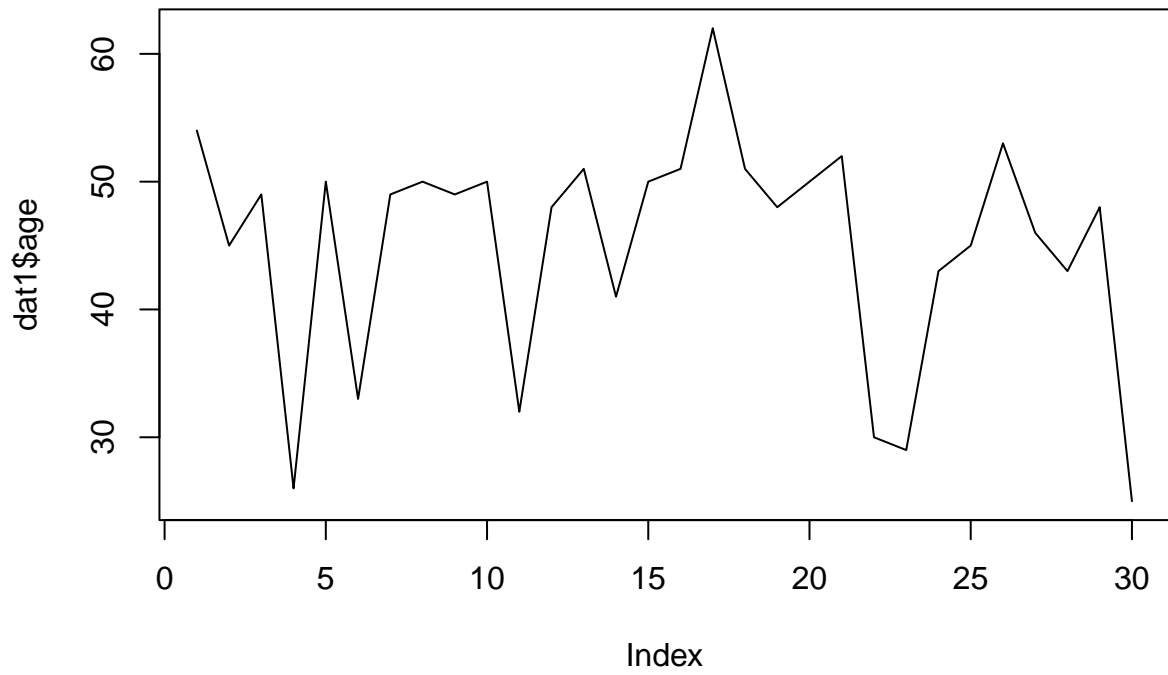
(7) 산점도

```
plot(dat1$b_alb, dat1$b_plt)
```



(8) 선 그래프

```
plot(dat1$age, type='l')
```

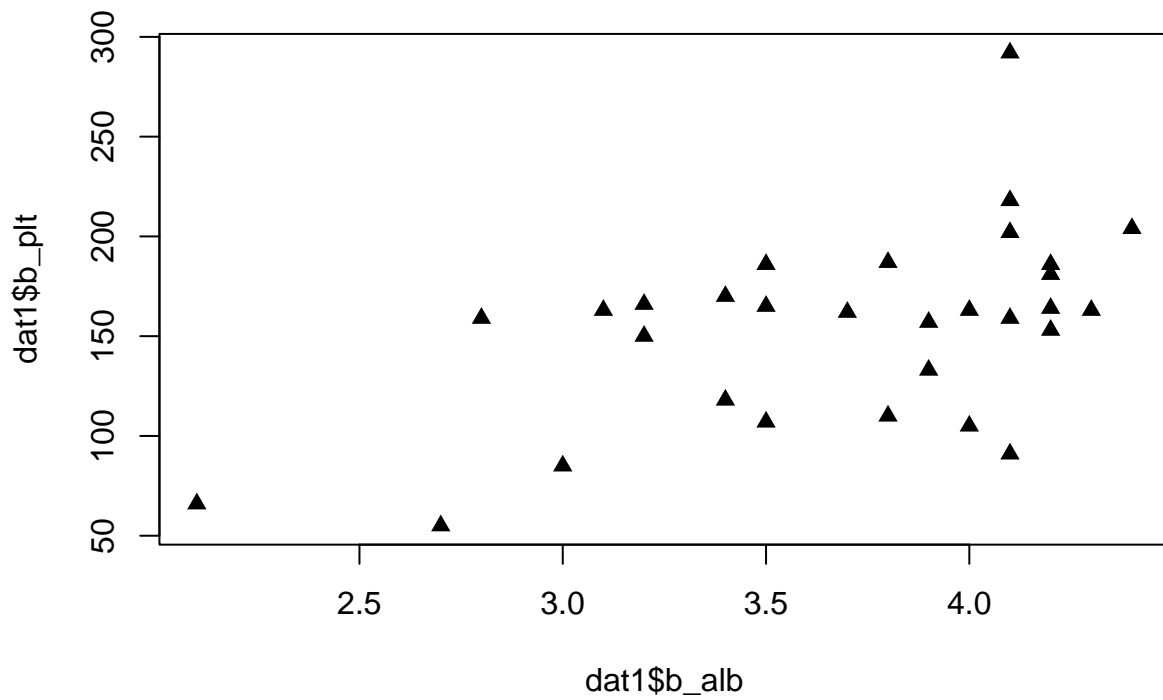


1.2 기본 그래프의 옵션

점, 선, 축 등의 모양, 색상, 굵기, 음영, 투명도, 추가선

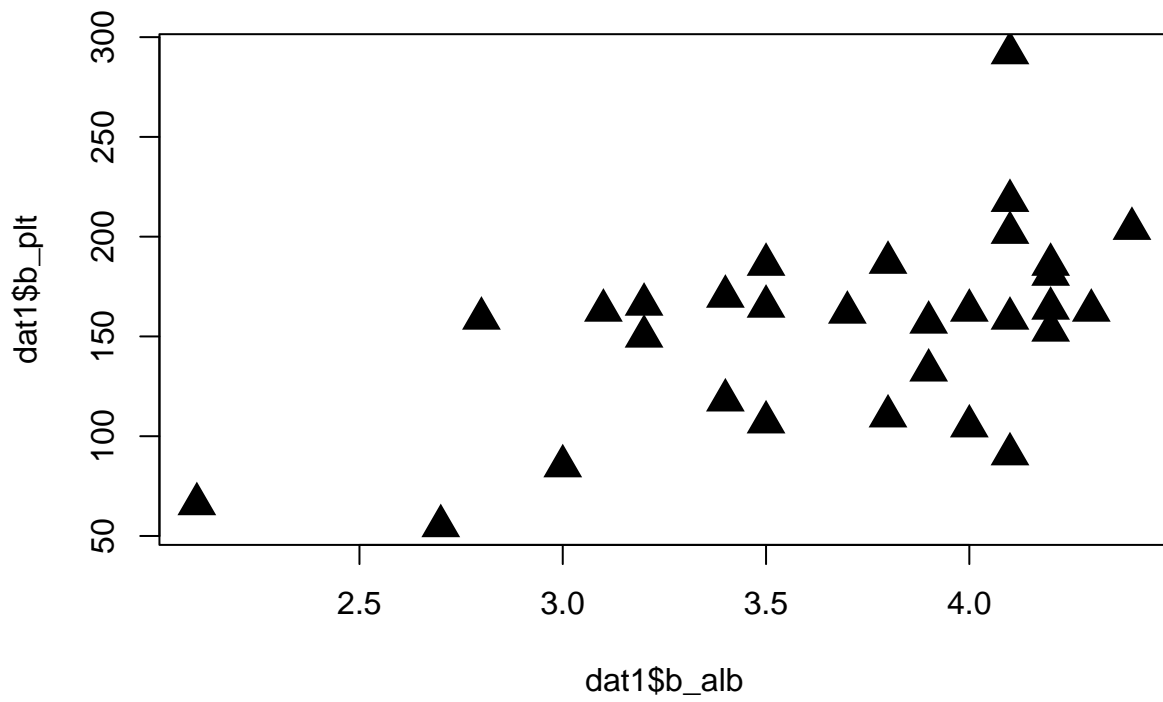
(1) 점 모양

```
plot(dat1$b_alb, dat1$b_plt, pch=17)
```



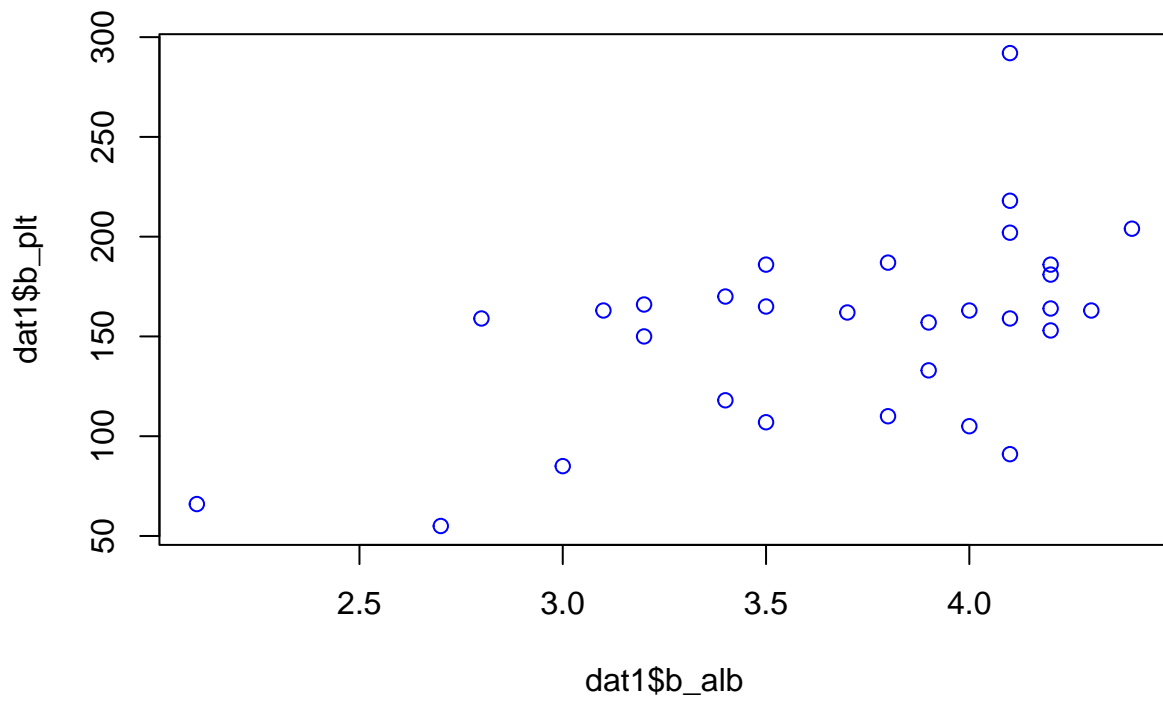
(2) 점 크기

```
plot(dat1$b_alb, dat1$b_plt, pch=17, cex=2)
```



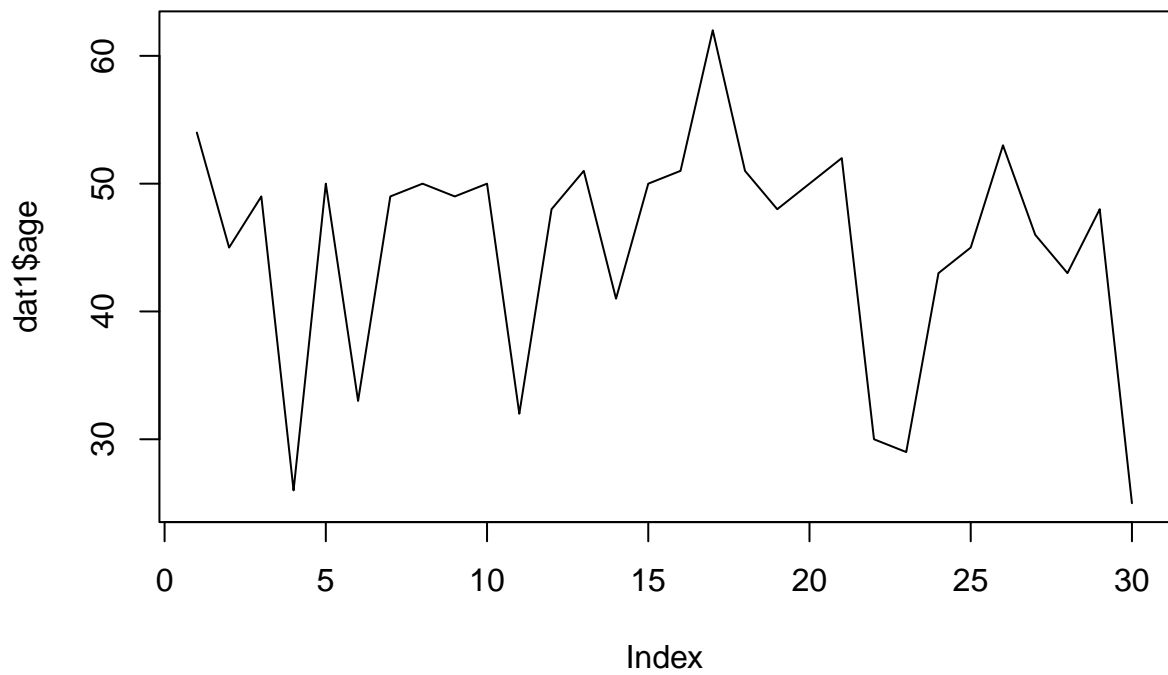
(3) 색깔 변경

```
plot(dat1$b_alb, dat1$b_plt, col='blue')
```

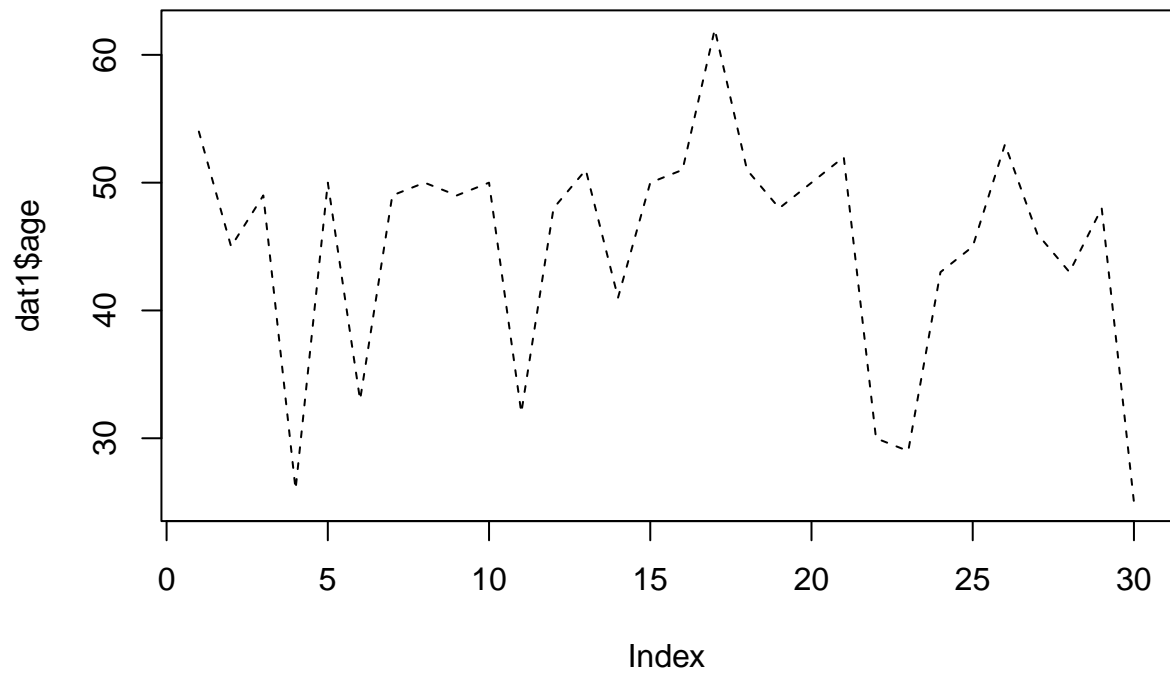


(4) 선 종류

```
plot(dat1$age, type='l')
```

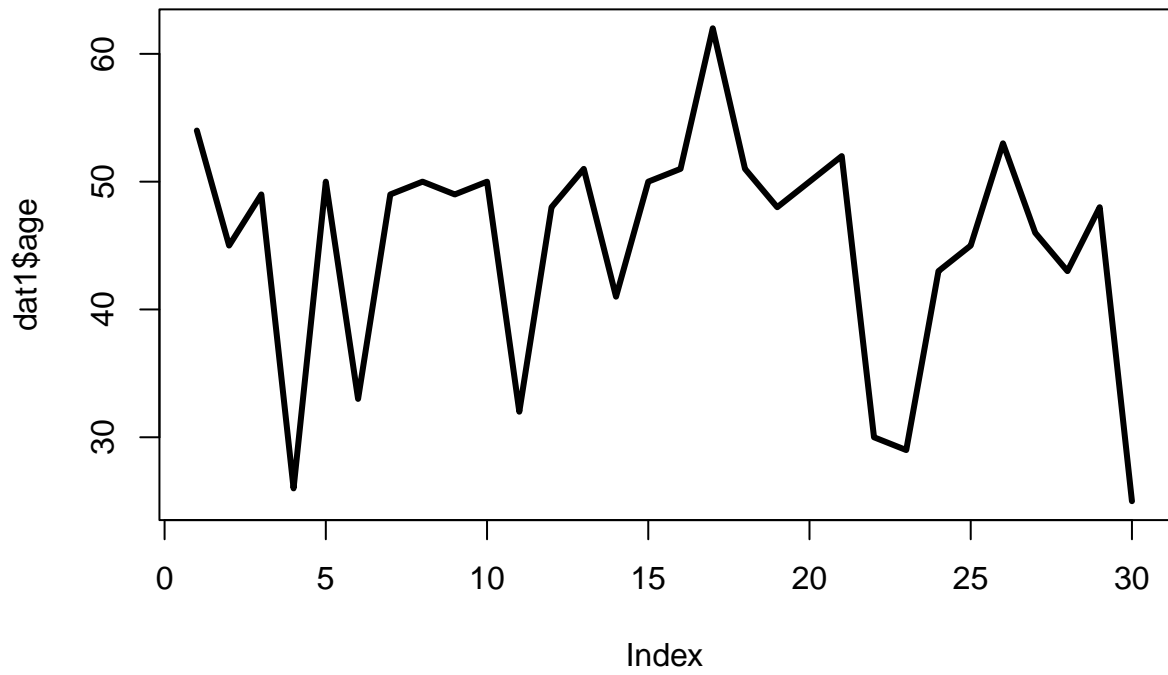


```
plot(dat1$age, type='l', lty=2)
```



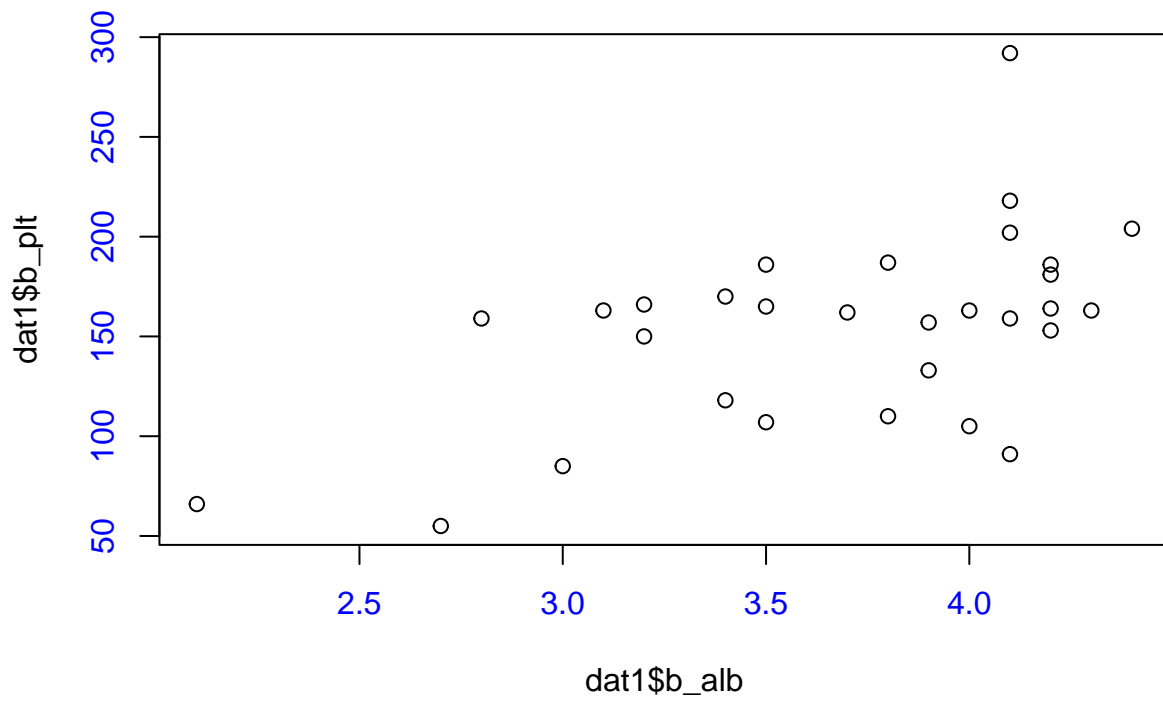
(5) 선 두께

```
plot(dat1$age, type='l', lwd=3)
```



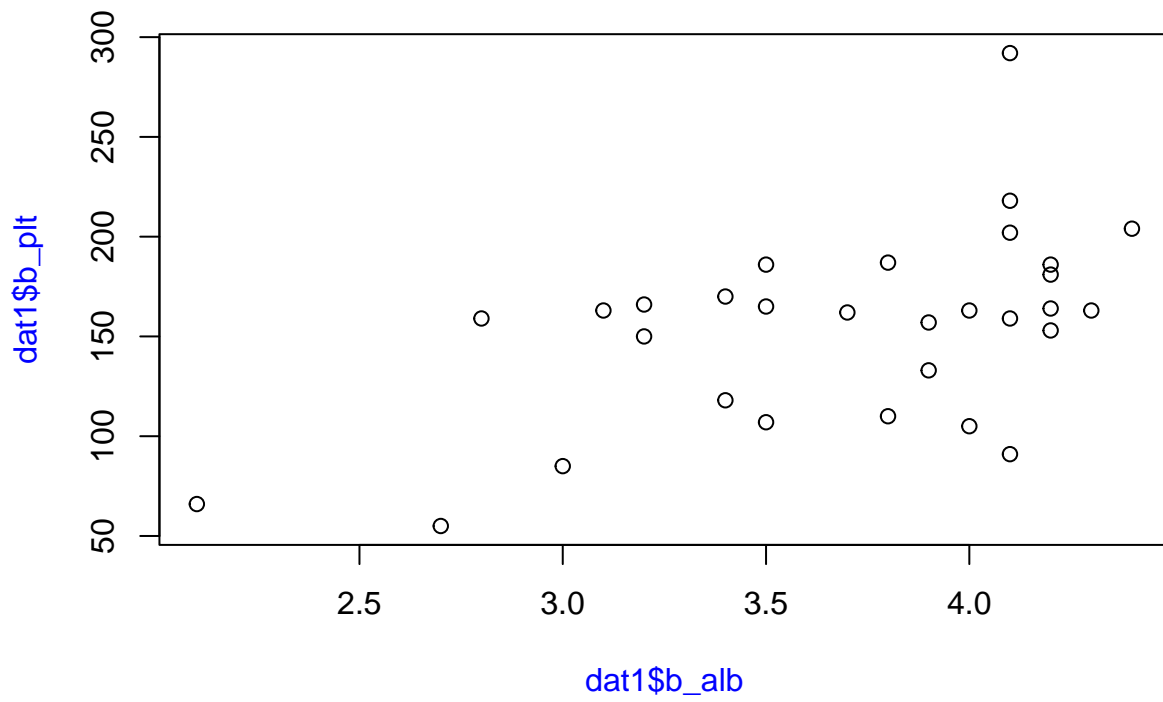
(6) 축 색상

```
plot(dat1$b_alb, dat1$b_plt, col.axis='blue')
```



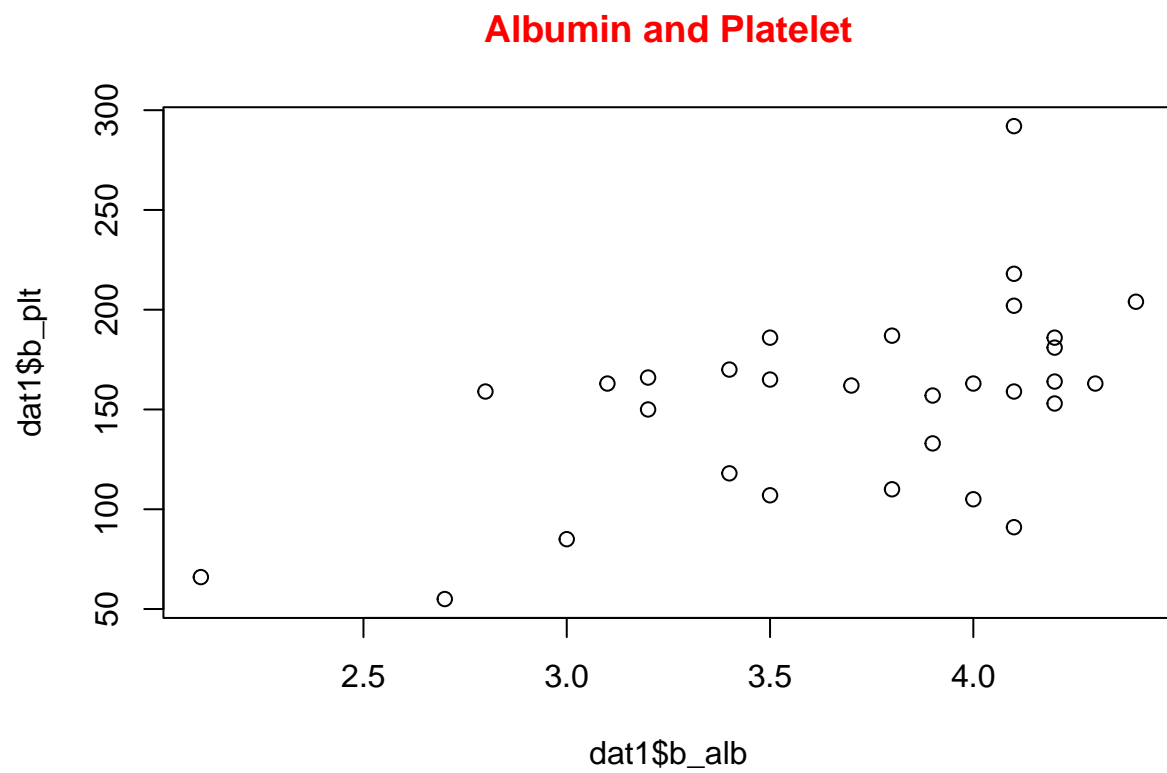
(7) 축 이름 색상 변경

```
plot(dat1$b_alb, dat1$b_plt, col.lab='blue')
```



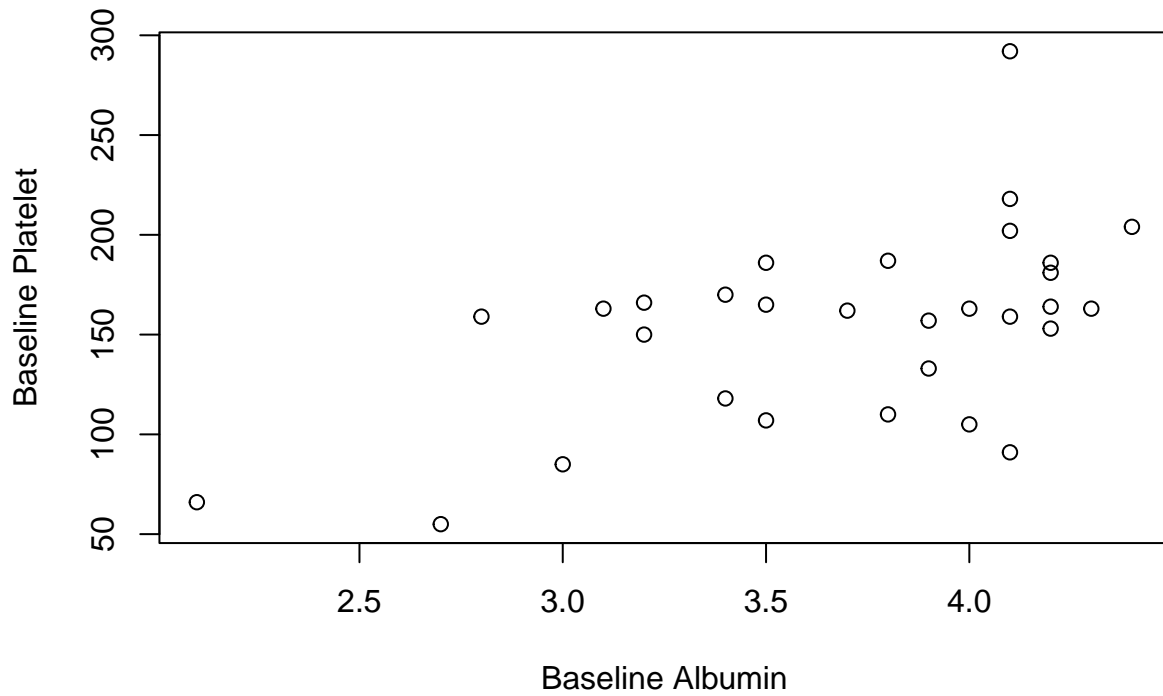
(8) 그래프 제목 붙이고 색상 변경

```
plot(dat1$b_alb, dat1$b_plt,  
     main='Albumin and Platelet',  
     col.main='red')
```



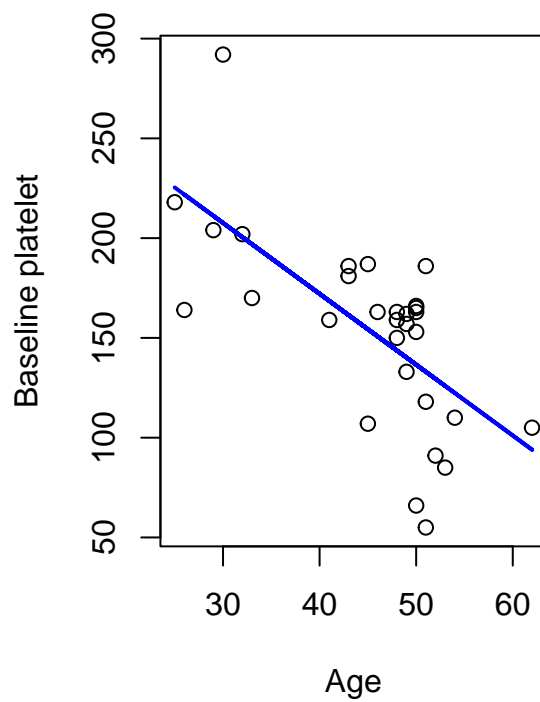
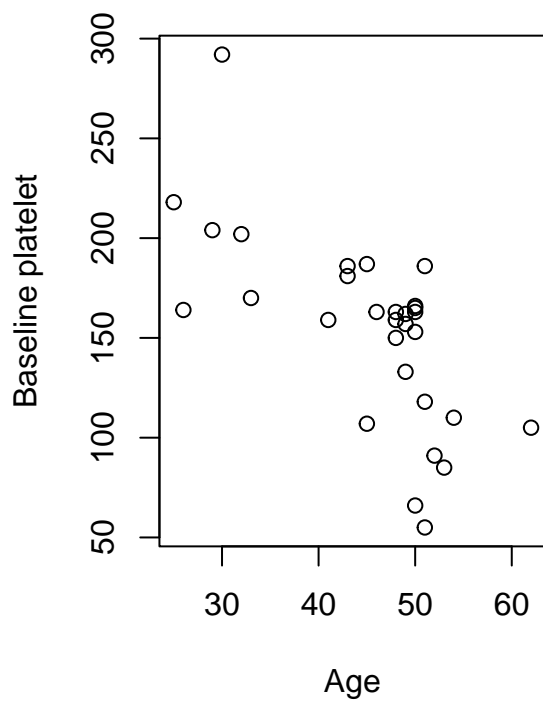
(9) 축 이름 변경

```
plot(dat1$b_alb, dat1$b_plt, xlab='Baseline Albumin',ylab='Baseline Platelet')
```



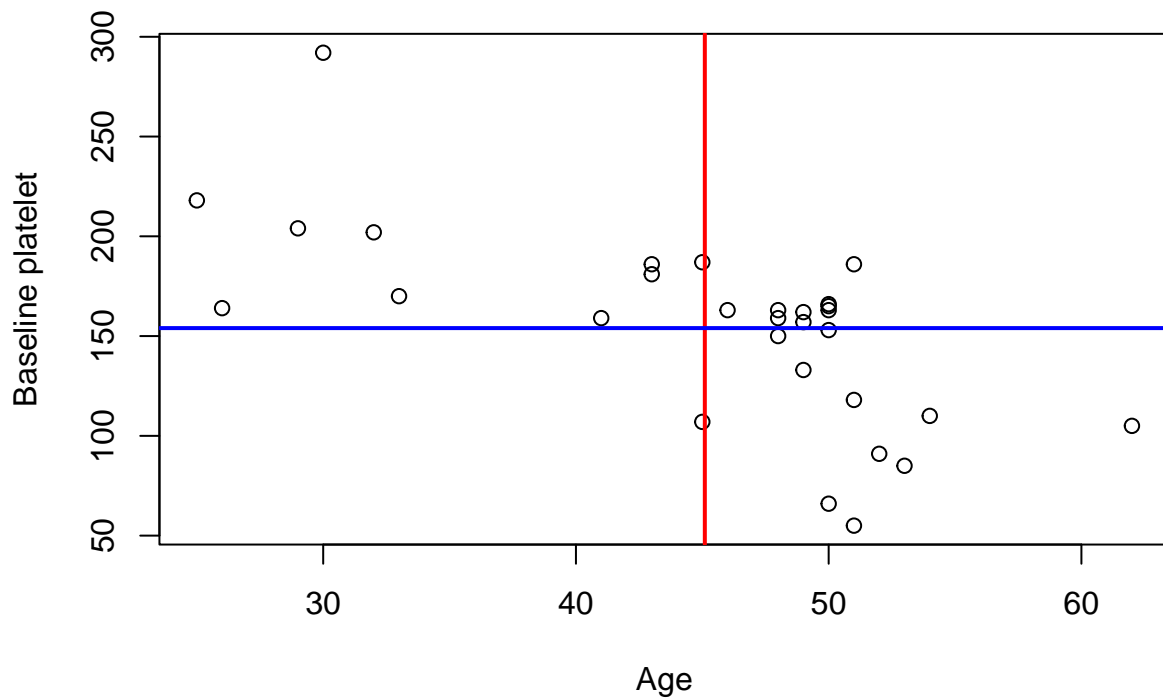
(10) 그래프 선 추가하기

```
par(mfrow=c(1,2))
plot(dat1$age, dat1$b_plt,
     xlab='Age', ylab='Baseline platelet')
fit<-lm(b_plt~age, data=dat1)
plot(dat1$age, dat1$b_plt,
     xlab='Age', ylab='Baseline platelet')
lines(dat1$age, fit$fitted.values, col='blue', lwd=2)
```



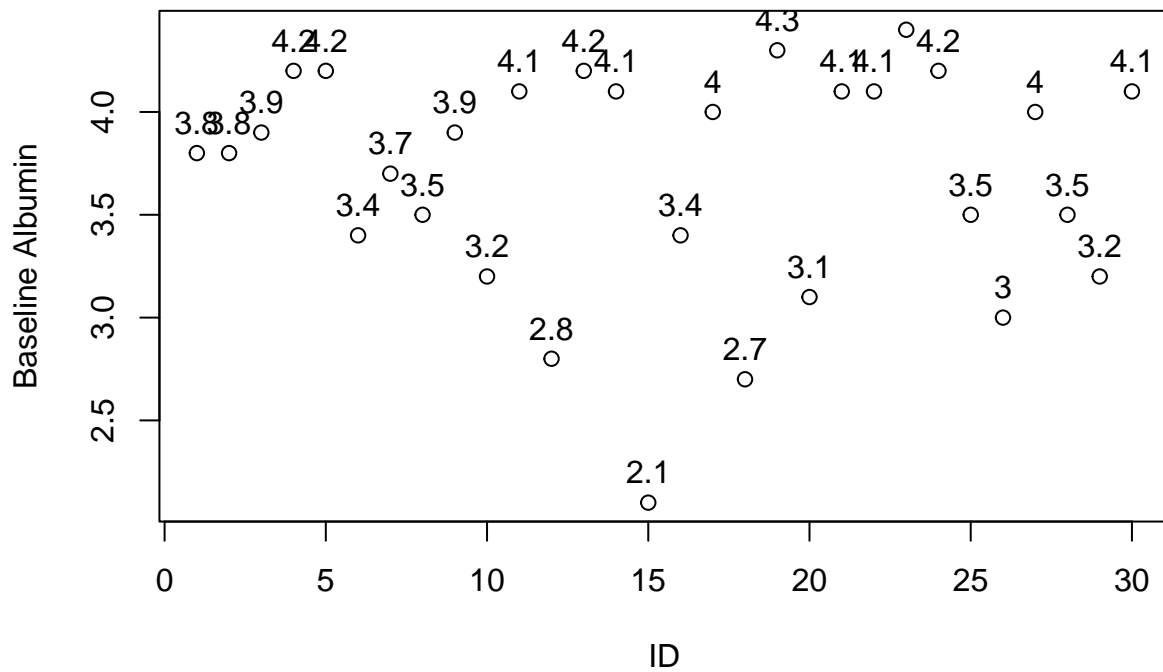
(11) 그래프에 수평 혹은 수직 라인 추가

```
plot(dat1$age, dat1$b_plt,  
      xlab='Age', ylab='Baseline platelet')  
abline(v=mean(dat1$age), col='red', lwd=2)  
abline(h=mean(dat1$b_plt), col='blue', lwd=2)
```



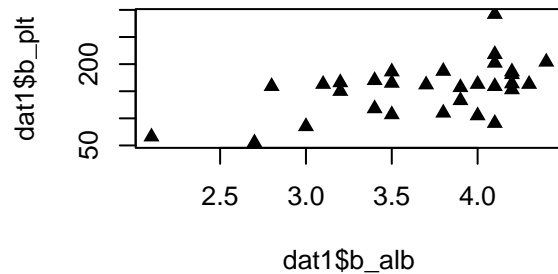
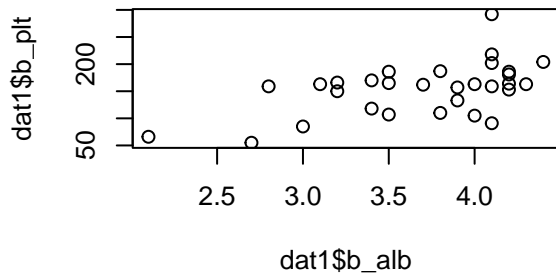
(12) 그래프에 text추가

```
plot(dat1$id, dat1$b_alb,  
      xlab='ID', ylab='Baseline Albumin')  
text(dat1$id, dat1$b_alb,  
      labels=dat1$b_alb, pos=3)
```

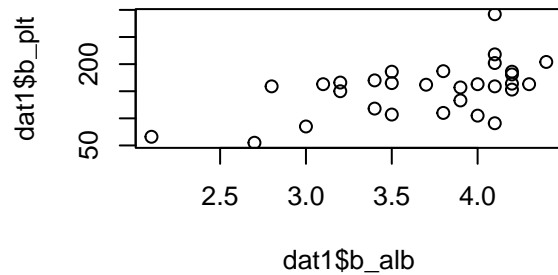
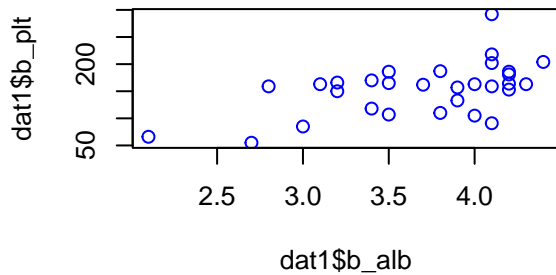


1.3 다중 그래프 그리기

```
par(mfrow=c(2,2))
plot(dat1$b_alb, dat1$b_plt)
plot(dat1$b_alb, dat1$b_plt, pch=17)
plot(dat1$b_alb, dat1$b_plt, col='blue')
plot(dat1$b_alb, dat1$b_plt,
      main='Association between Albumin and Platelet')
```



Association between Albumin and Plate



2. ggplot2

2.1 ggplot2 기본 문법

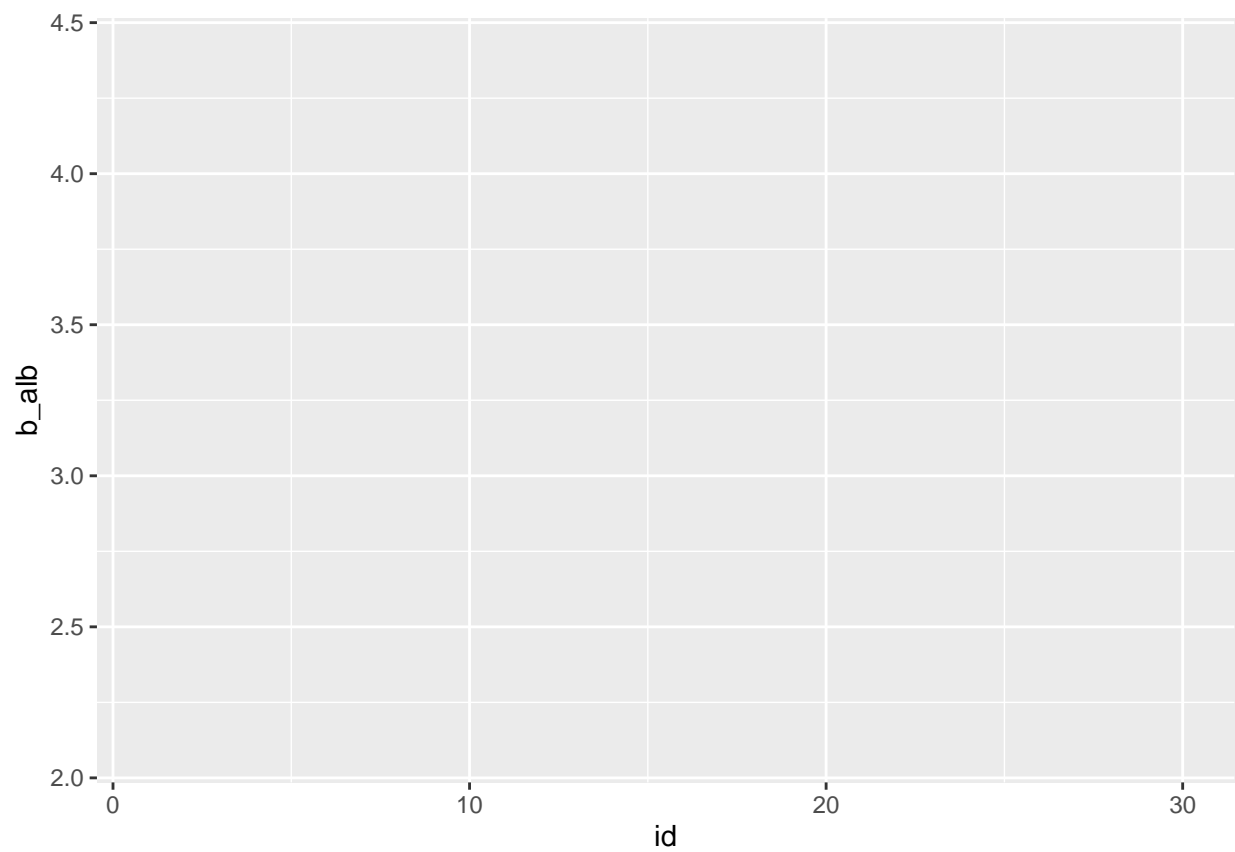
(1) 레이어

여러 개의 레이어가 차곡차곡 쌓여서 만들어지는 그래픽 문법이 ggplot2

(2) 시작

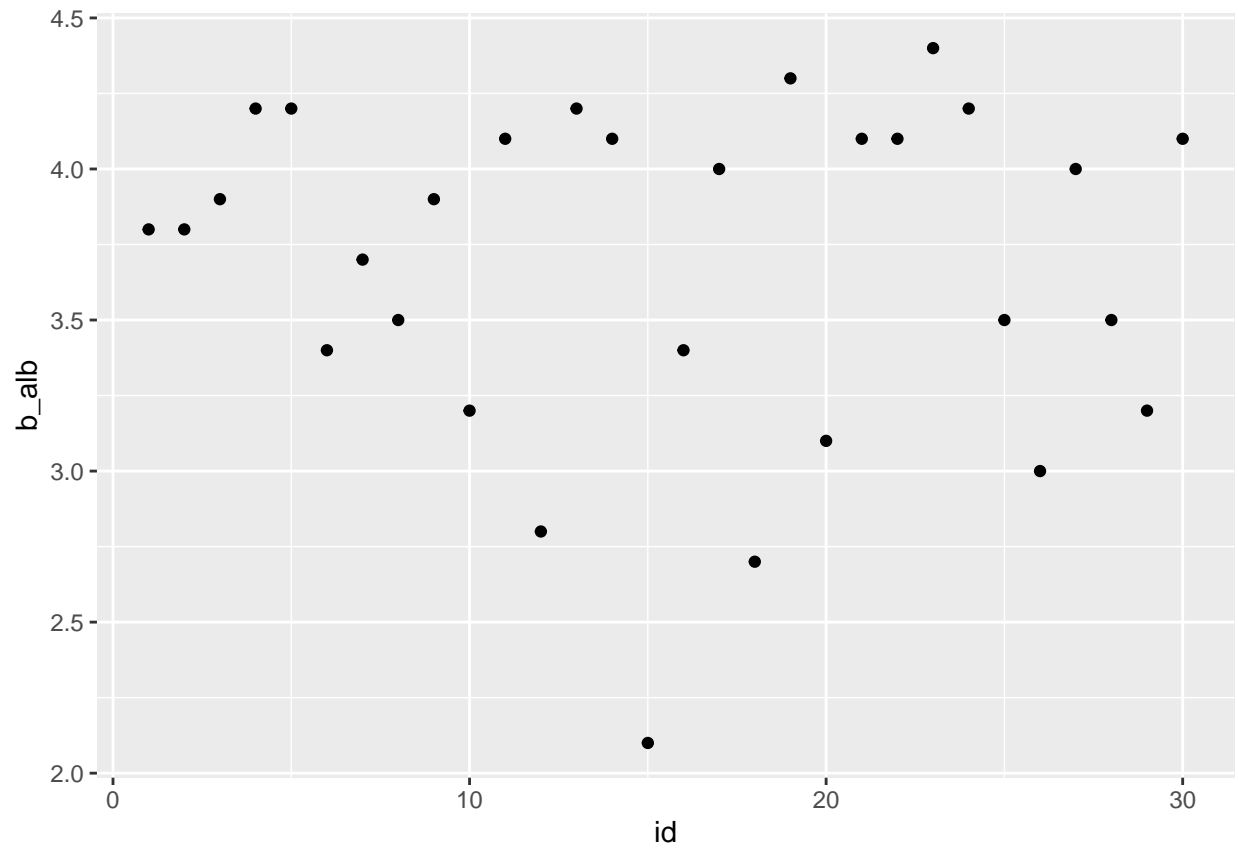
그래프 도면

```
ggplot(dat1, aes(x=id, y=b_alb))
```



geometrices 추가하기

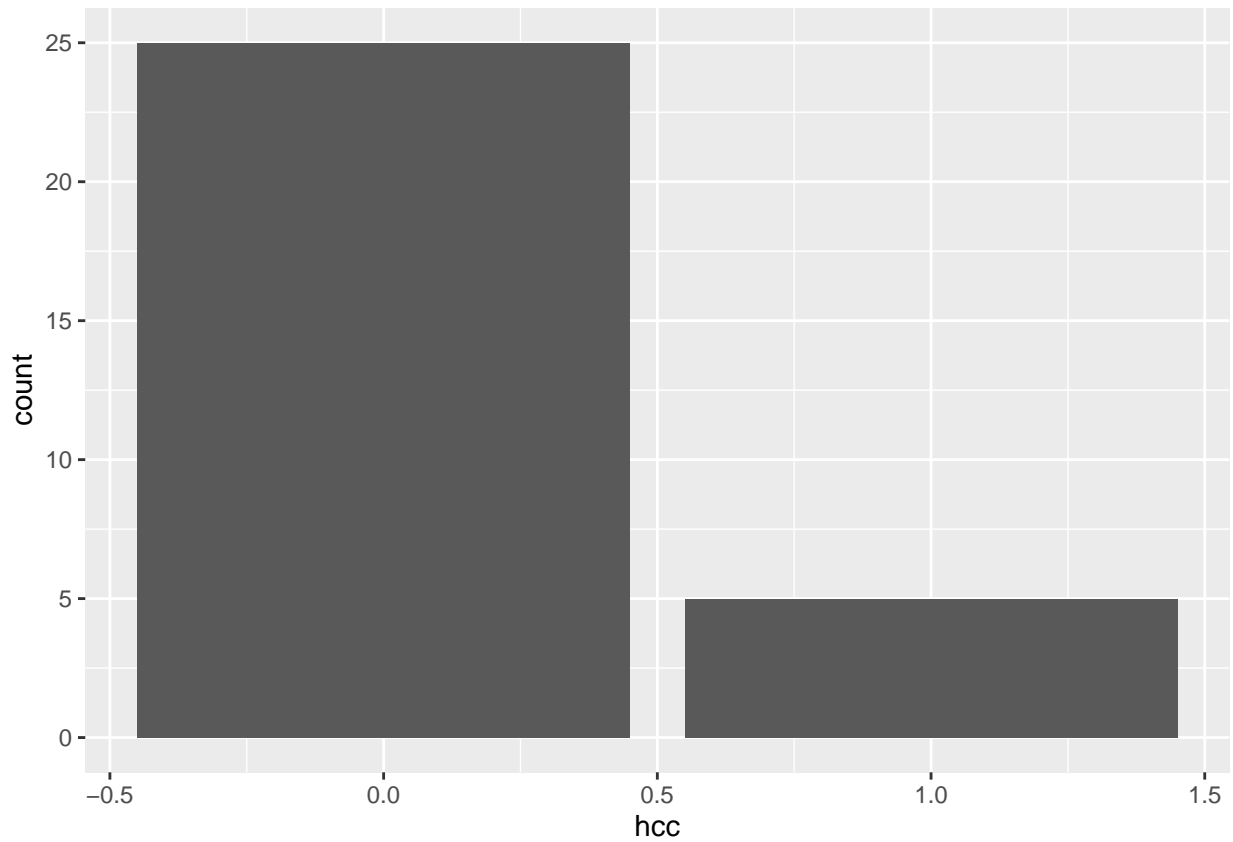
```
ggplot(dat1, aes(x=id, y=b_alb))+  
  geom_point()
```



2.2 막대 그래프

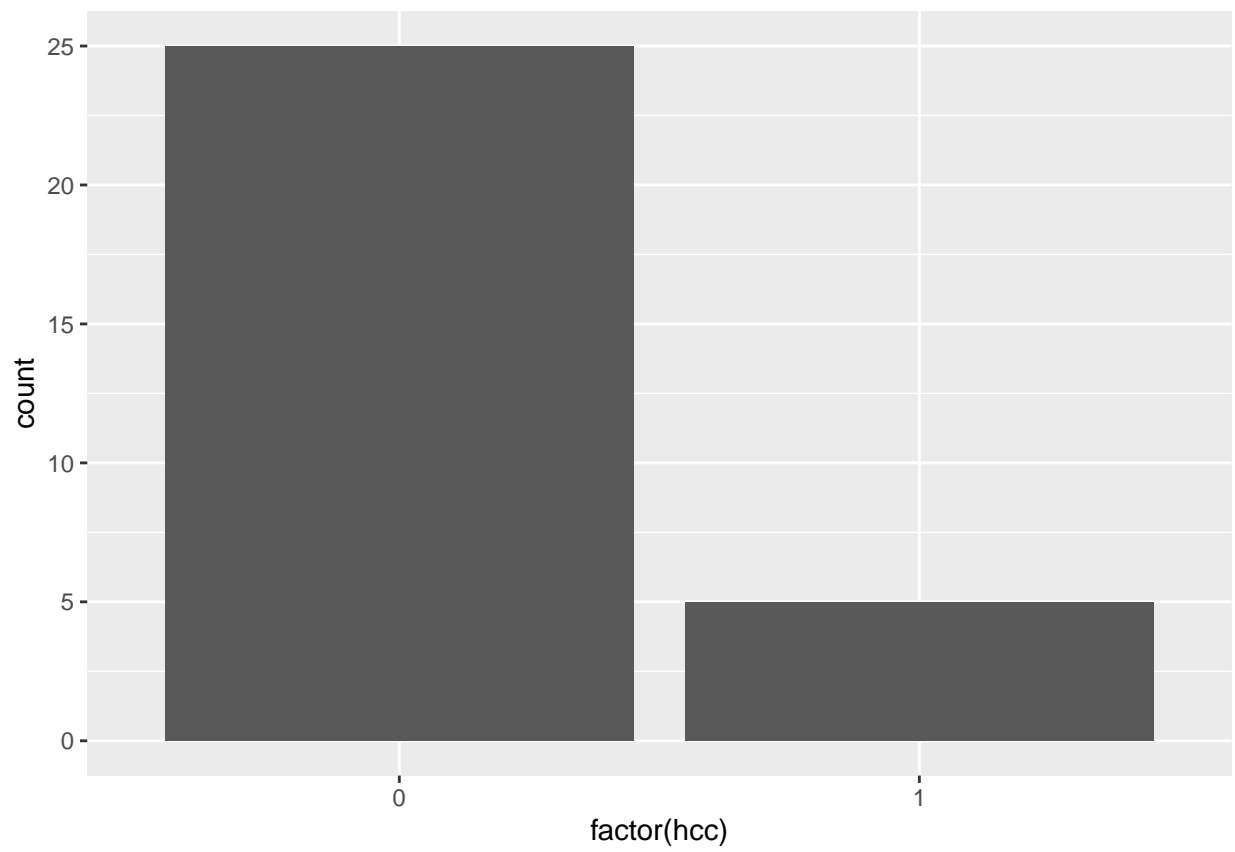
(1) 빈도

```
ggplot(dat1, aes(x=hcc))+  
  geom_bar()
```



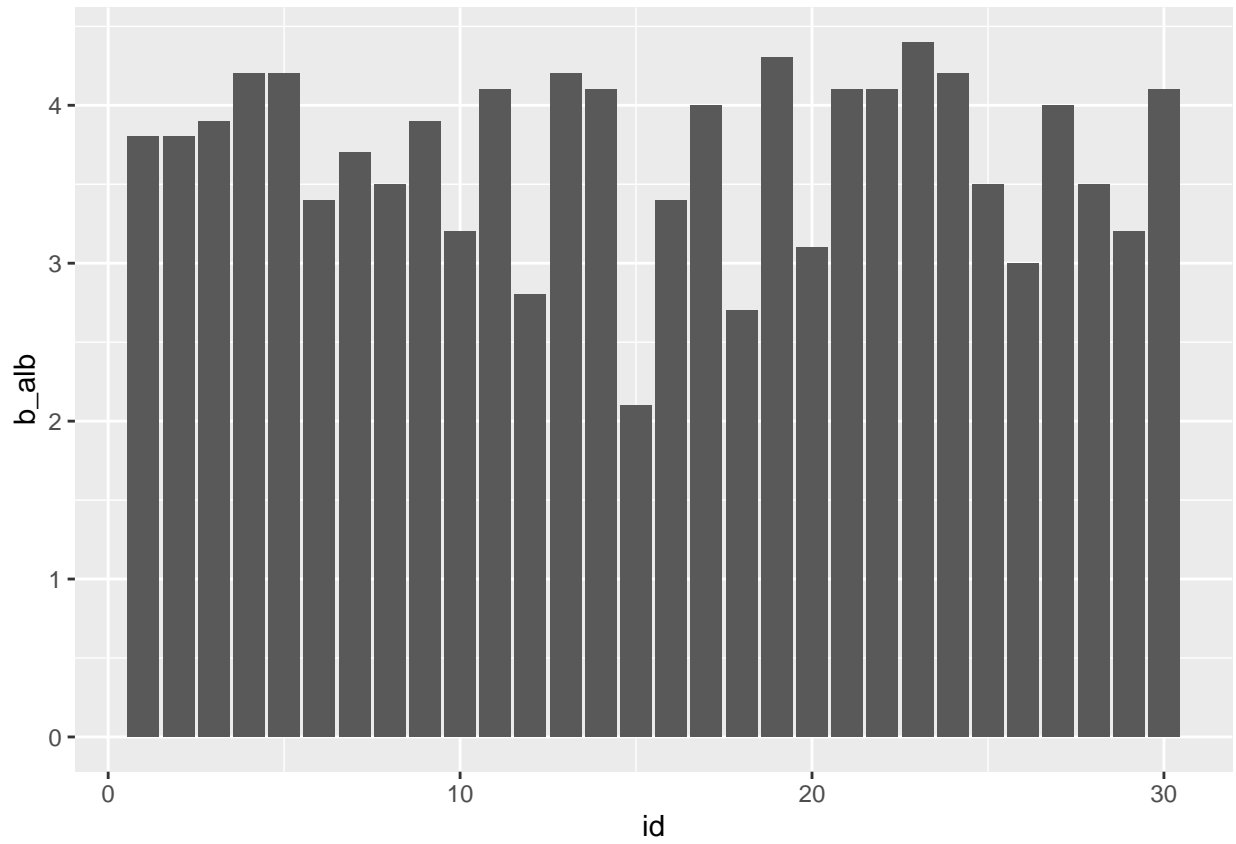
hcc가 numeric형태여서 x축이 연속형이 나옴 범주형이므로 factor로 변경

```
ggplot(dat1, aes( x=factor(hcc) ))+  
  geom_bar()
```



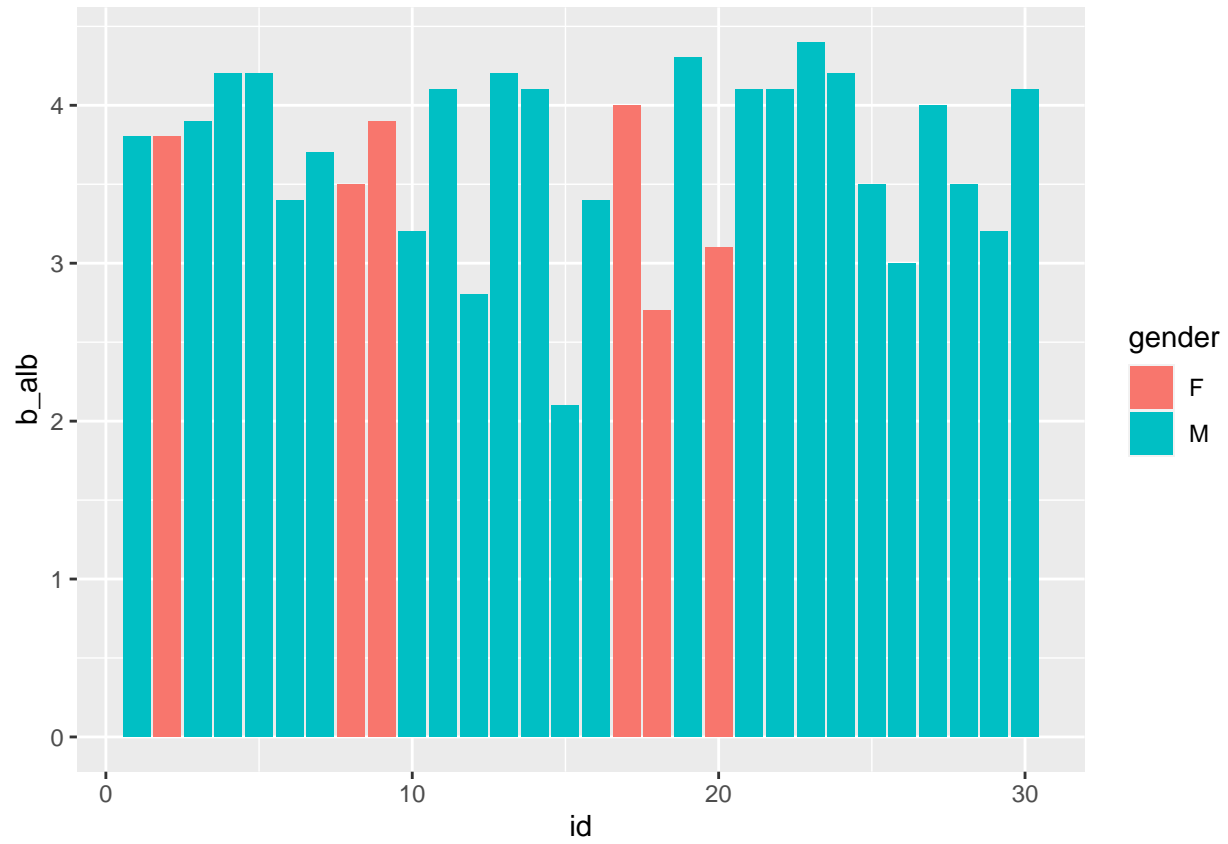
(2) 고유값

```
ggplot(dat1, aes(x=id, y=b_alb))+  
  geom_bar(stat='identity')
```



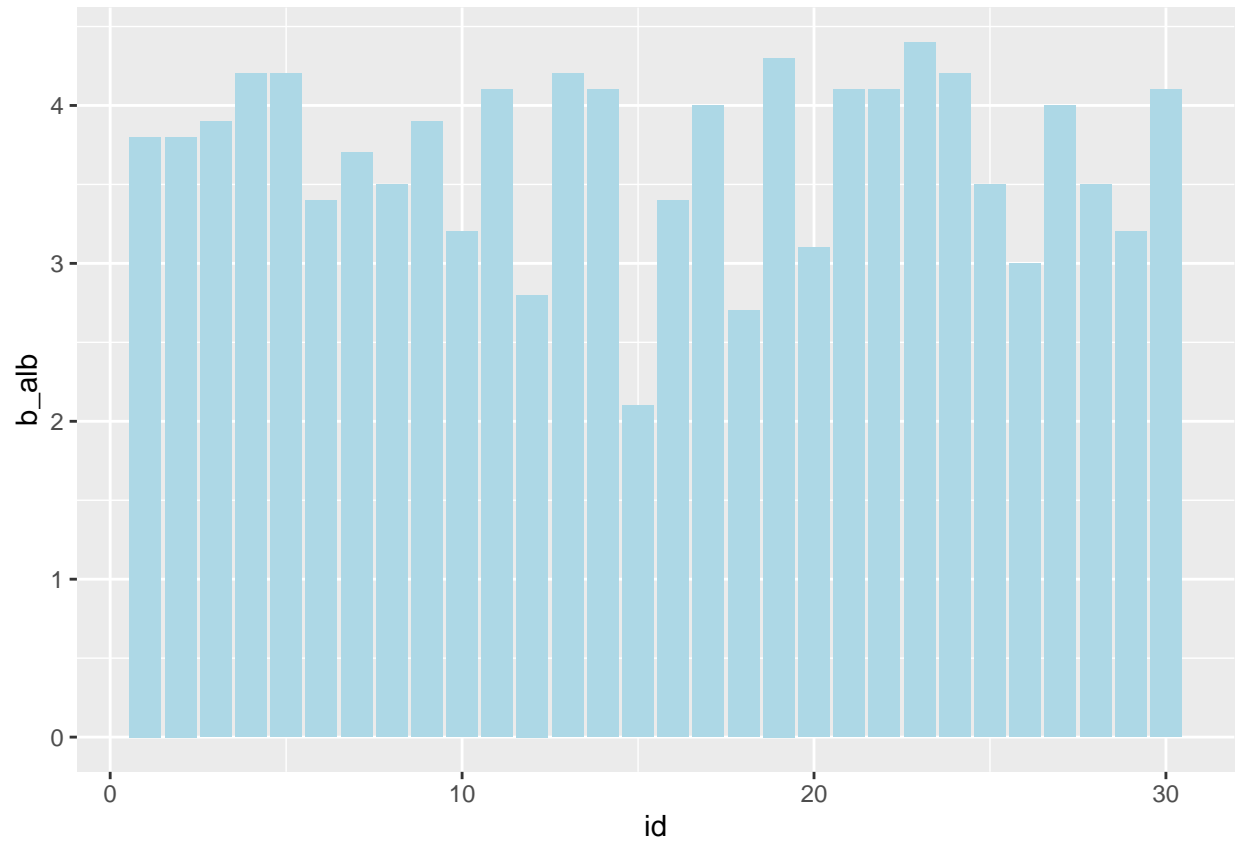
(3) 막대 그래프 색상 변경하기

```
ggplot(dat1, aes(x=id, y=b_alb, fill=gender))+  
  geom_bar(stat='identity')
```



ggplot2는 코드 순서대로 실행되기 때문에 색상이 덮이게 됨

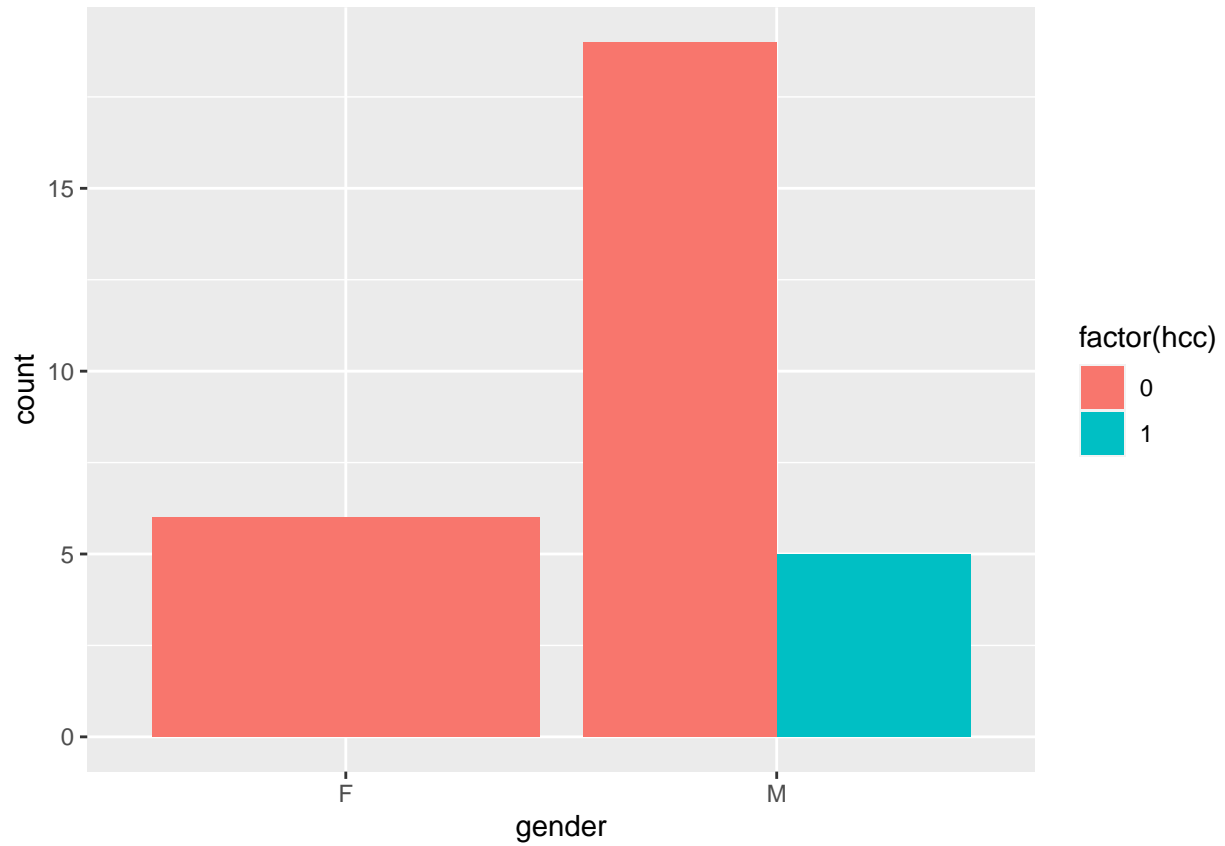
```
ggplot(dat1, aes(x=id, y=b_alb, fill=gender))+  
  geom_bar(stat='identity', fill='lightblue')
```



(4) 옆으로 나란한 막대 그리기

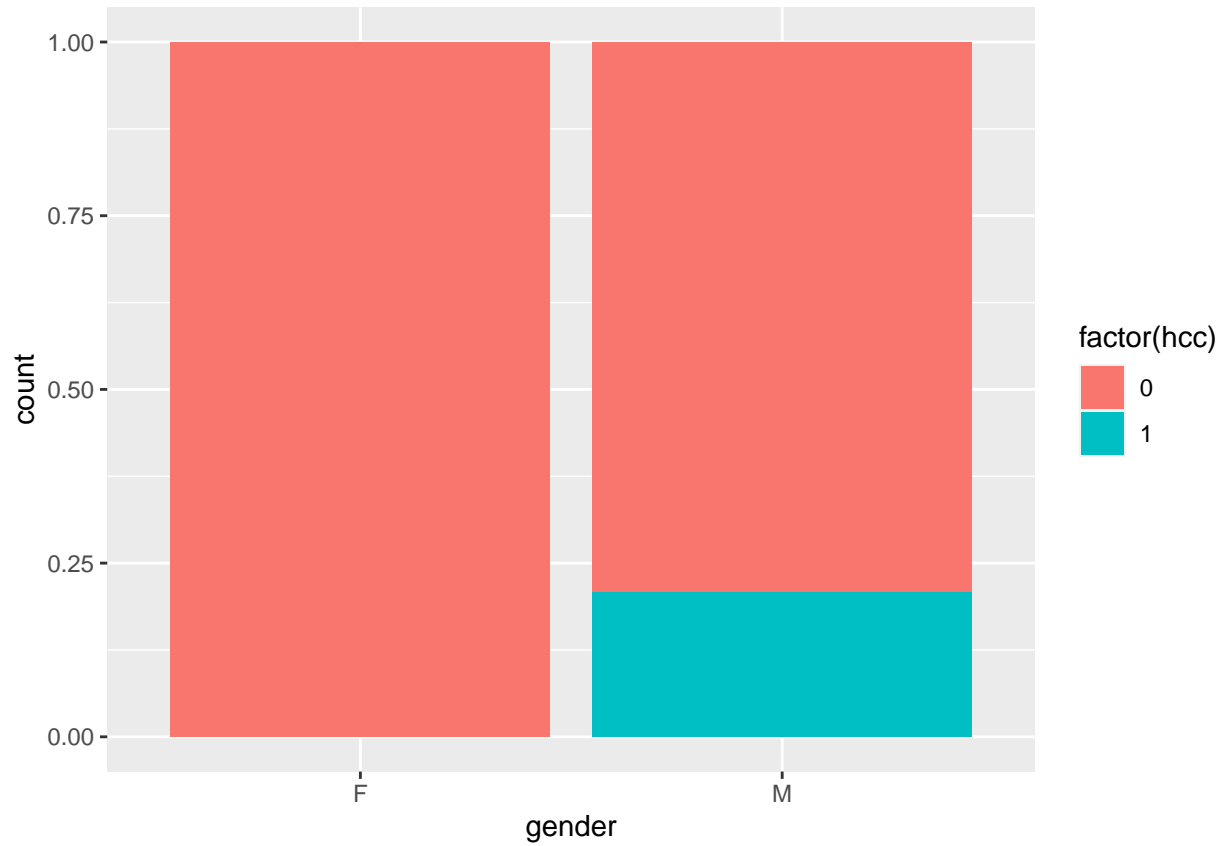
성별에 따른 간암 환자수와 간암 유무에 따른 숫자를 각각 표시

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge')
```



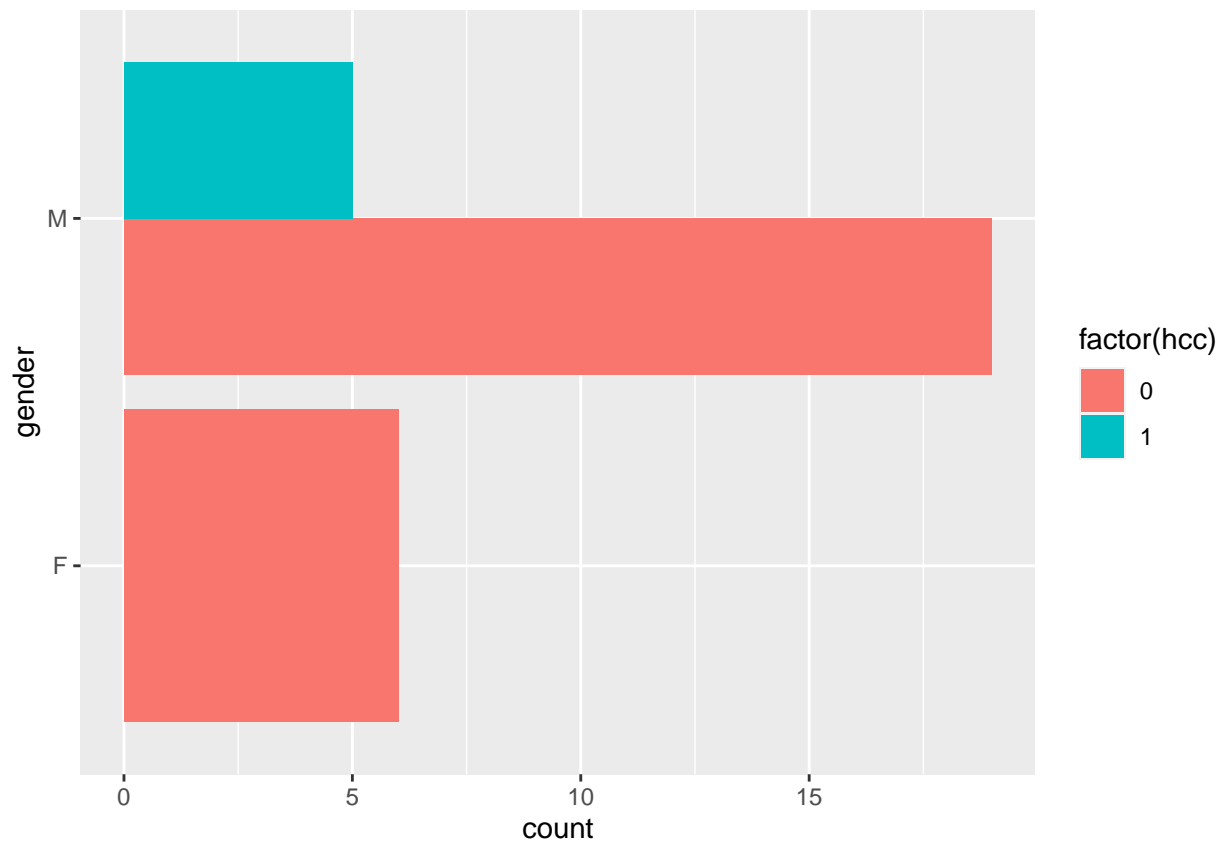
(5) 누적 비율 막대 그래프 그리기

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(position='fill')
```



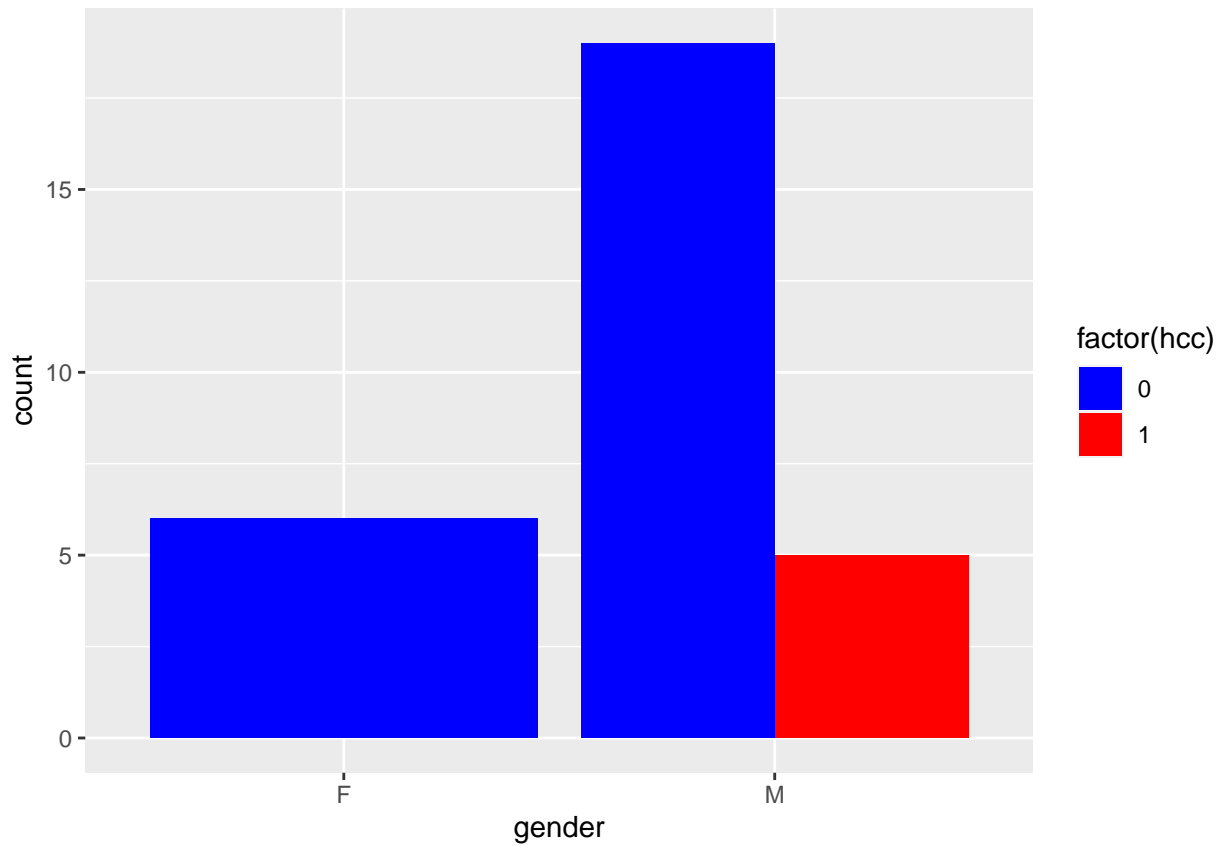
(6) 막대 그래프에서 x축,y축 변경

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge')+  
  coord_flip()
```



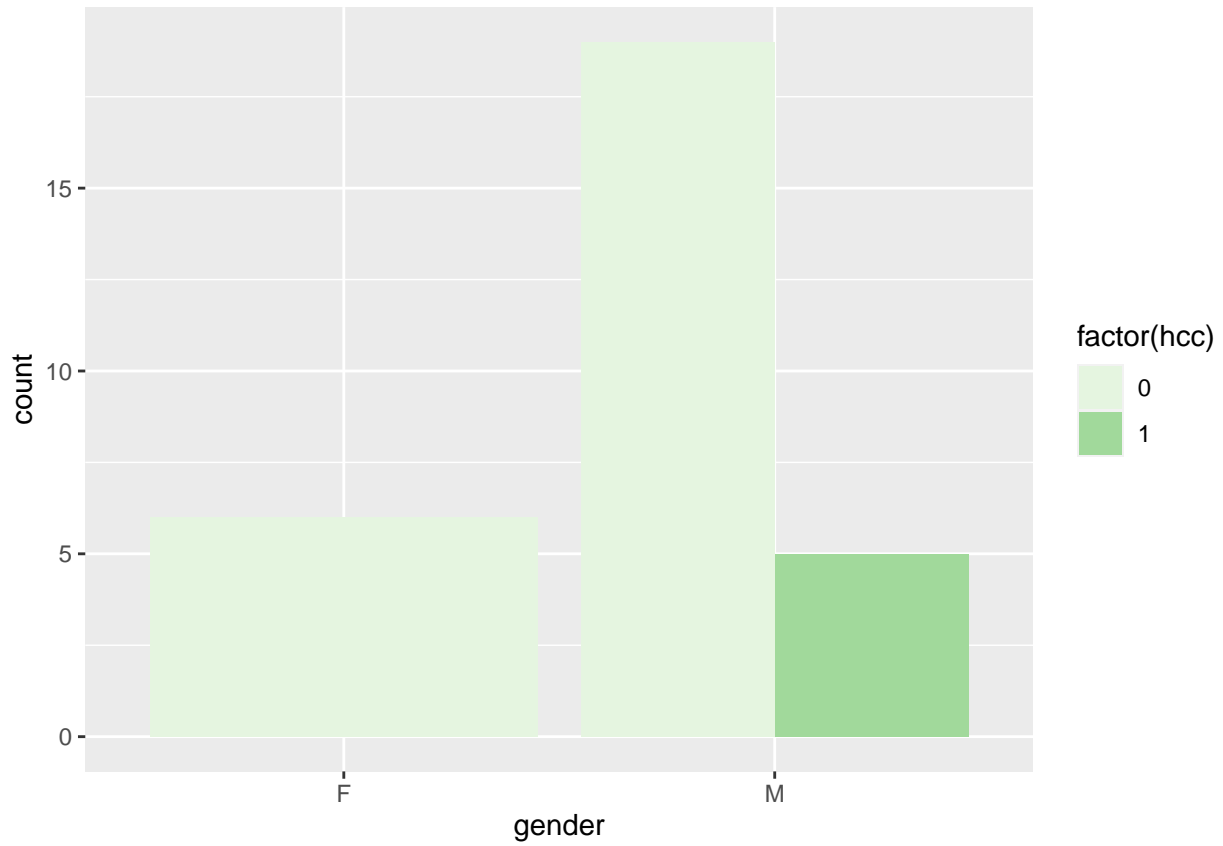
(7) 색상 변경

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge')+  
  scale_fill_manual(values=c('blue','red'))
```



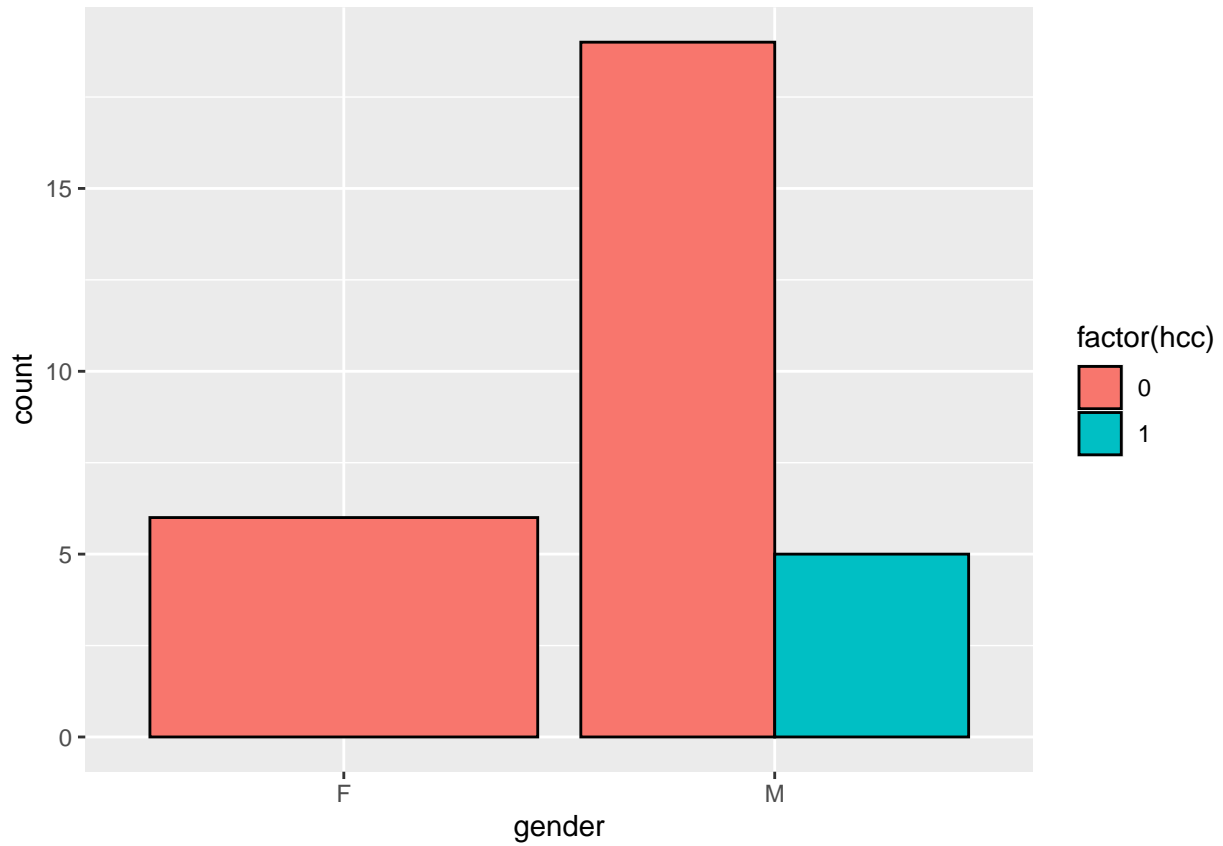
(8) 자동 색상 선택

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge')+  
  scale_fill_brewer(palette='Greens')
```



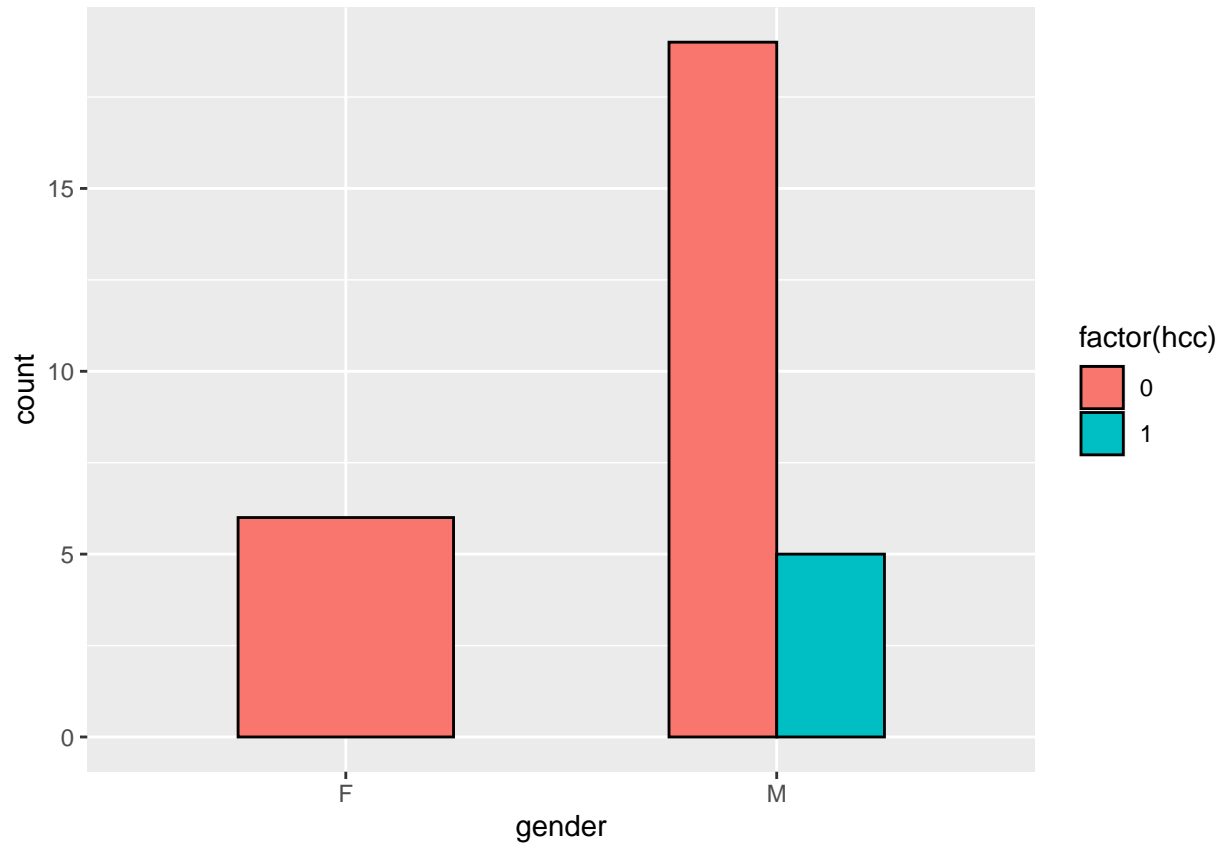
(9) 테두리 색 입히기

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge', color='black')
```



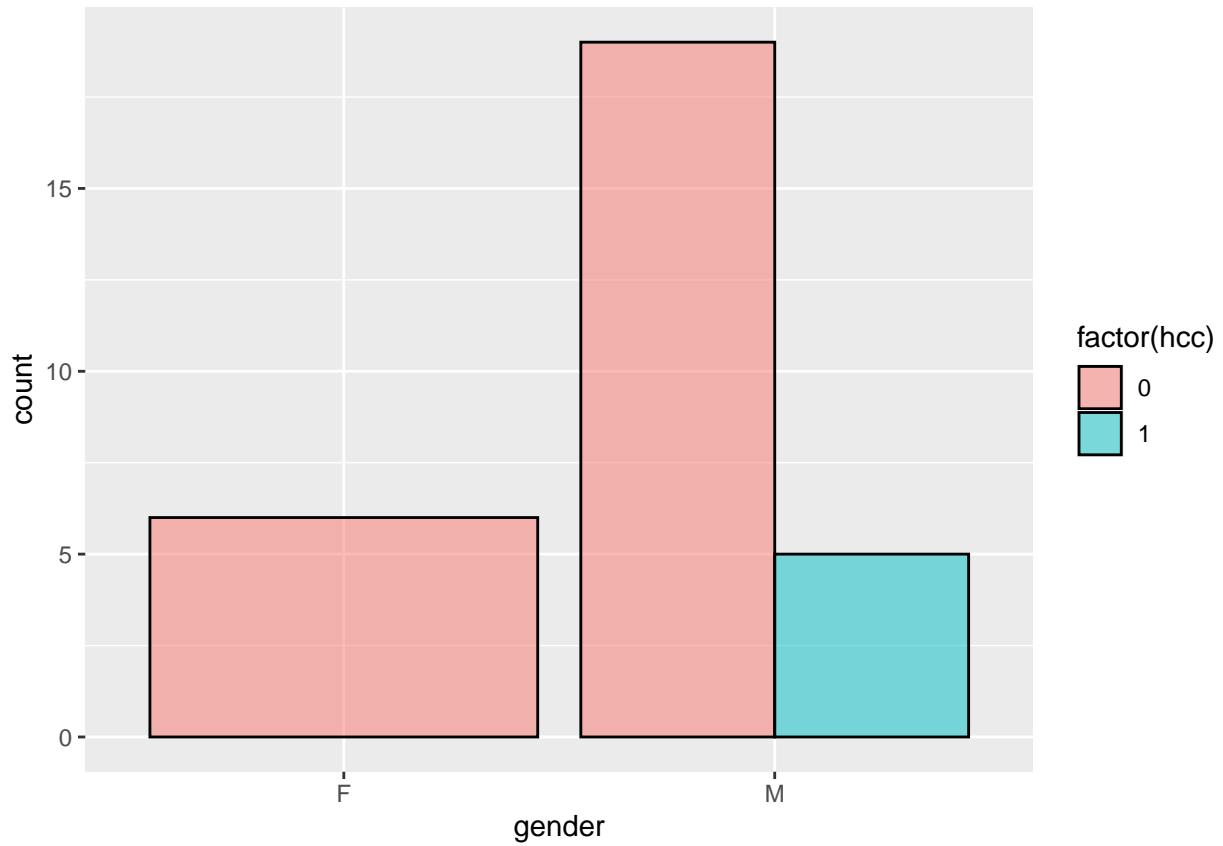
(10) 너비 조절

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge', color='black', width=0.5)
```



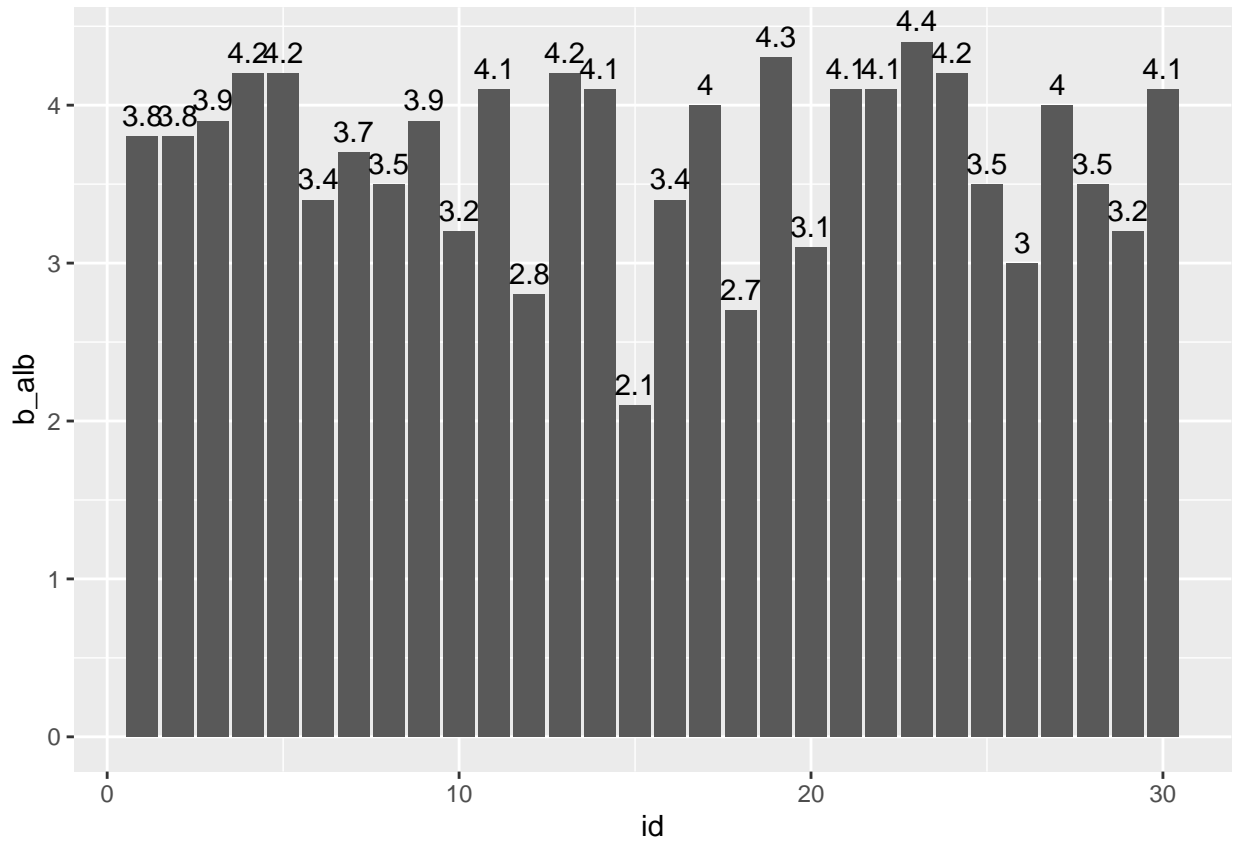
(11) 투명도 조절

```
ggplot(dat1, aes(x= gender, fill=factor(hcc)))+  
  geom_bar(stat='count', position='dodge', color='black', alpha=0.5)
```



(12) 막대 그래프에 데이터 값 추가하기

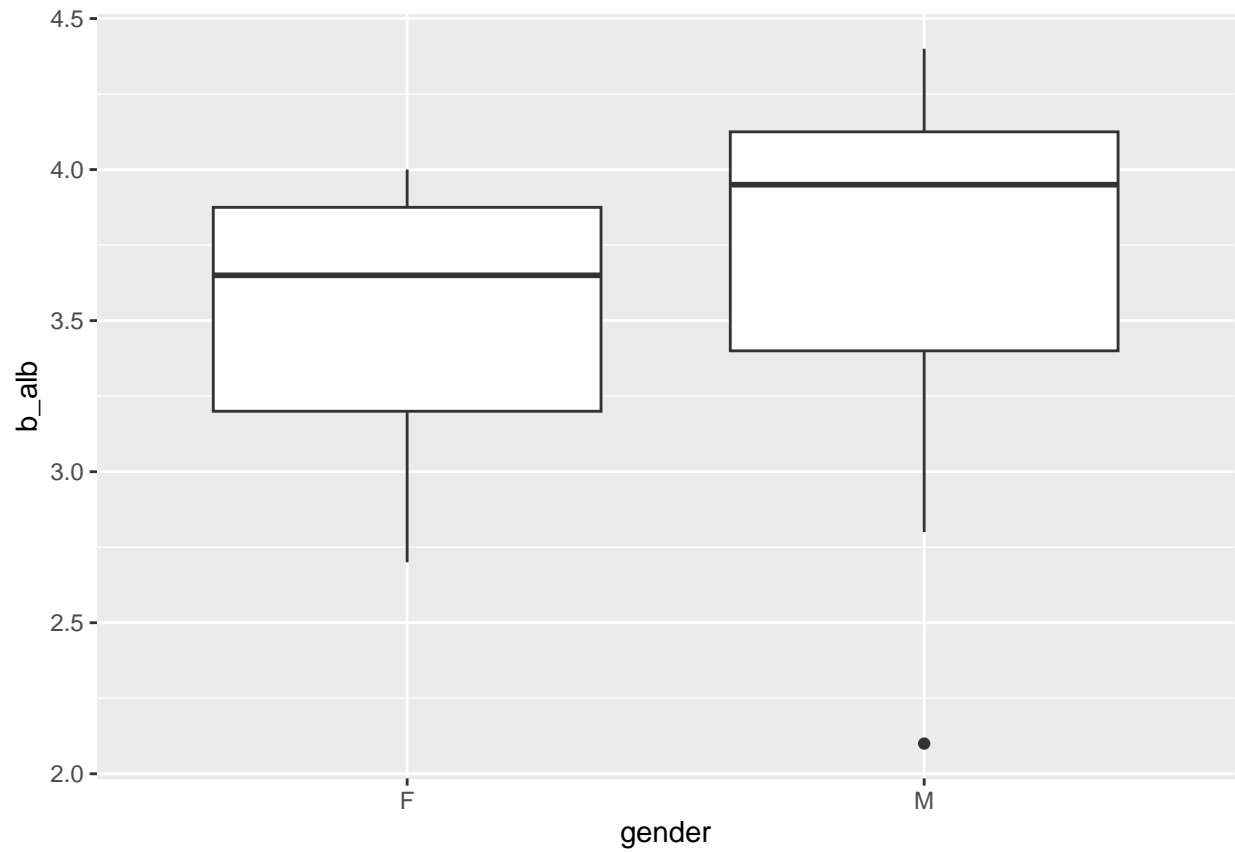
```
ggplot(dat1, aes(x=id, y=b_alb))+  
  geom_bar(stat='identity')+  
  geom_text(aes(label=b_alb), vjust=-0.5)
```



2.3 박스 그래프

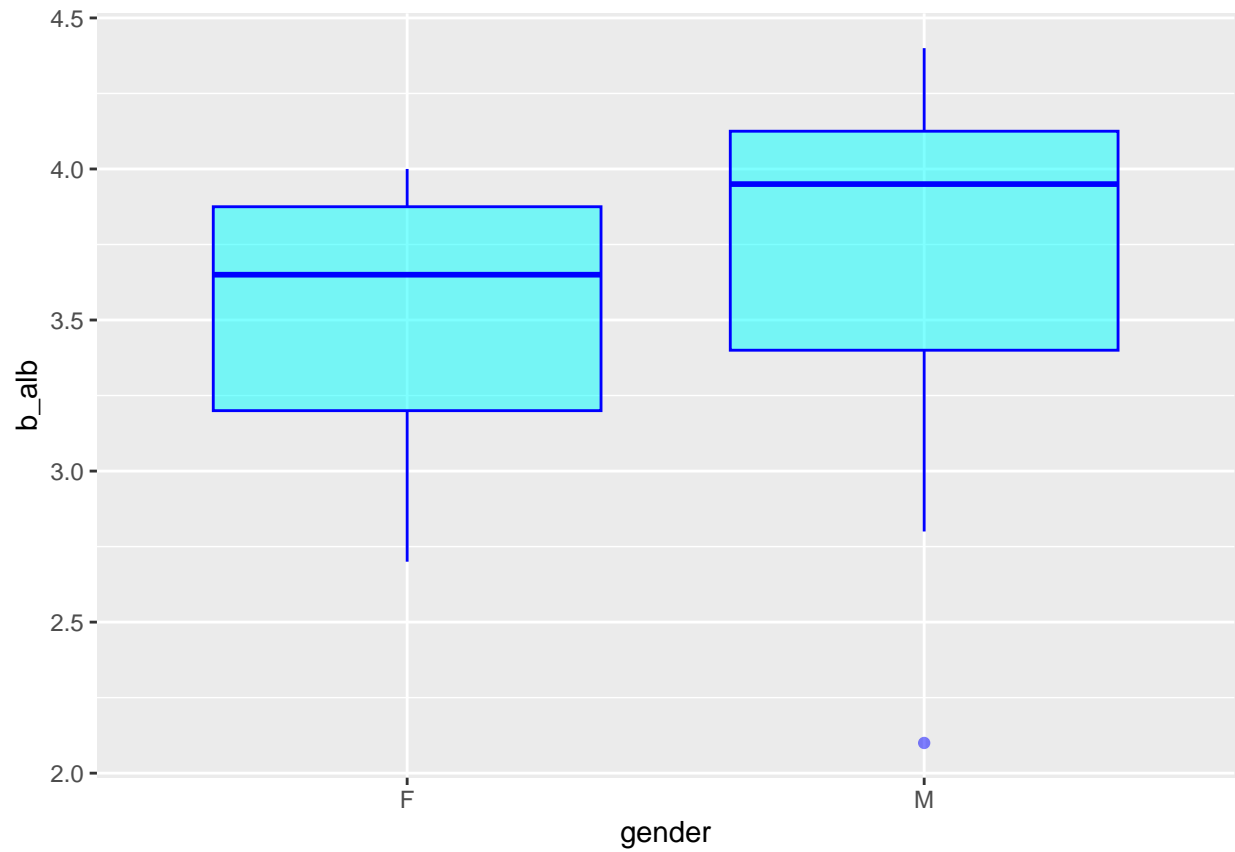
(1) 박스 그래프 그리기

```
ggplot(dat1, aes(x=gender, y=b_alb))+  
  geom_boxplot()
```



(2) 색상, 폭, 투명도 변경

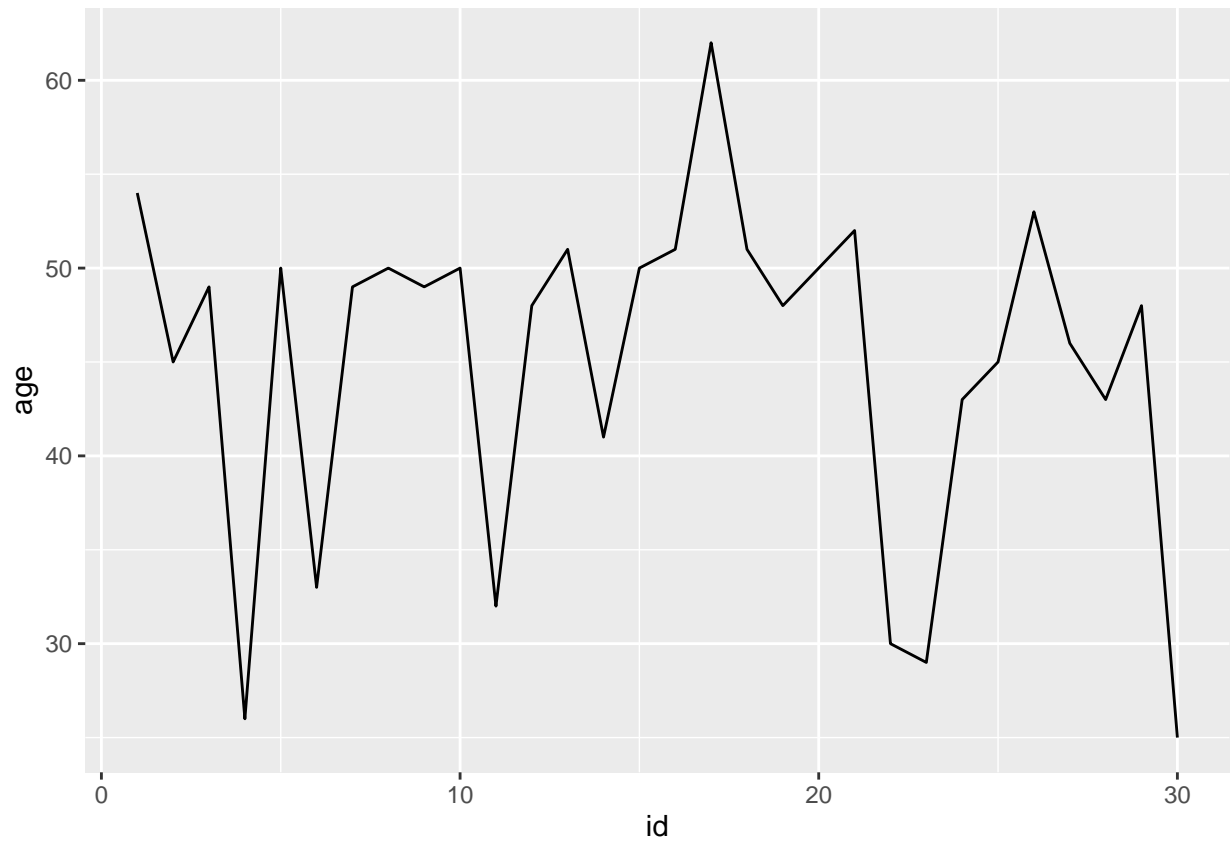
```
ggplot(dat1, aes(x=gender, y=b_alb))+  
  geom_boxplot(fill='cyan', color='blue', alpha=0.5)
```



2.4 선 그래프

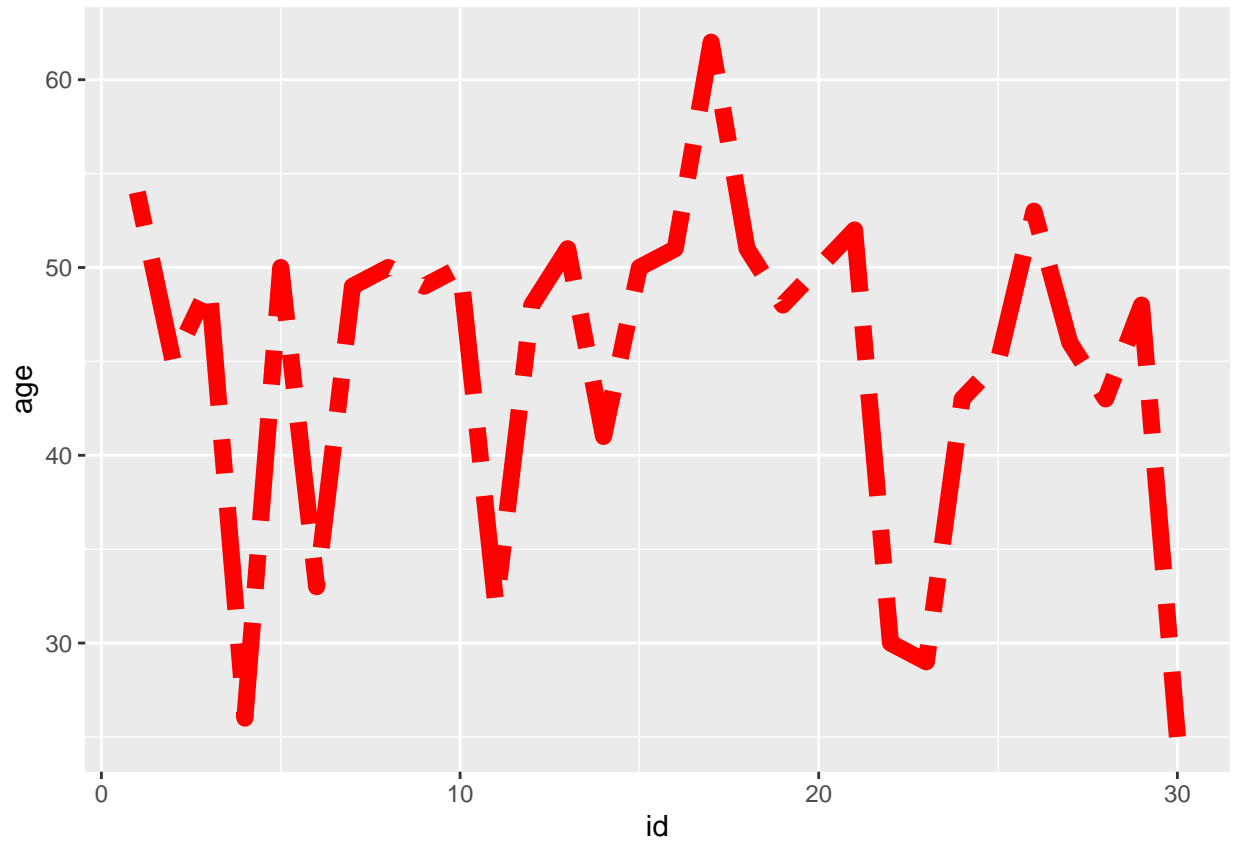
(1) 기본

```
ggplot(dat1, aes(x=id, y=age))+  
  geom_line()
```



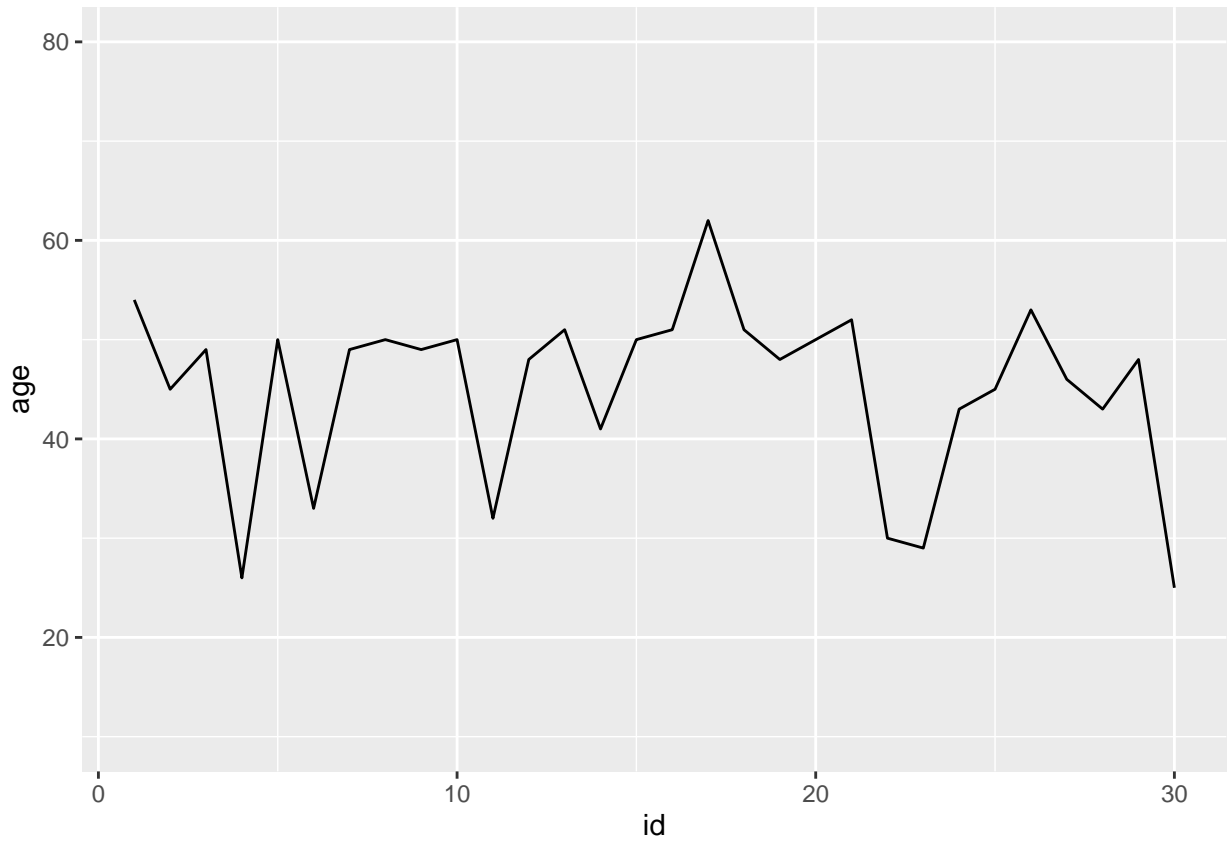
(2) 선 종류 변경 및 두께

```
ggplot(dat1, aes(x=id, y=age))+  
  geom_line(color='red', linetype=6,linewidth=3)
```



(3) 축 최소, 최대 변경

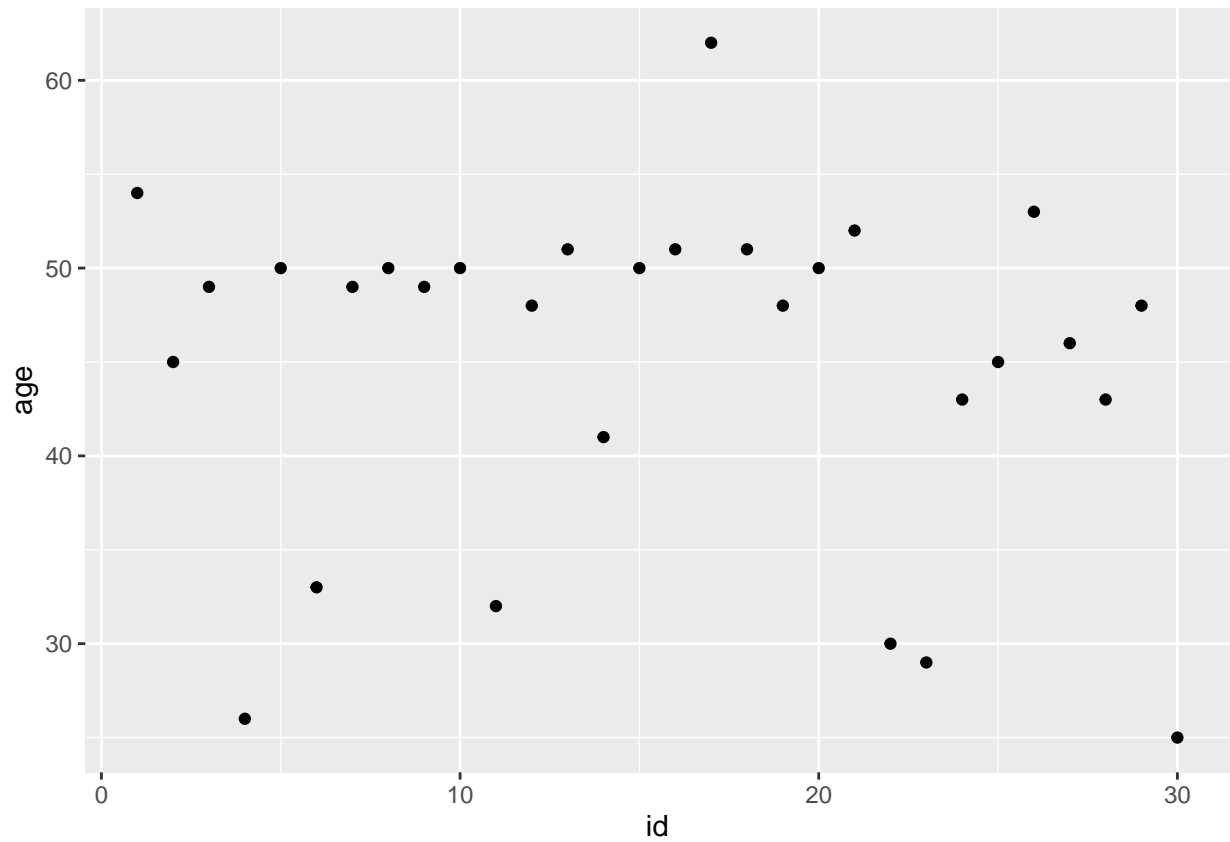
```
ggplot(dat1, aes(x=id, y=age))+  
  geom_line()+  
  ylim(c(10,80))
```



2.5 산점도

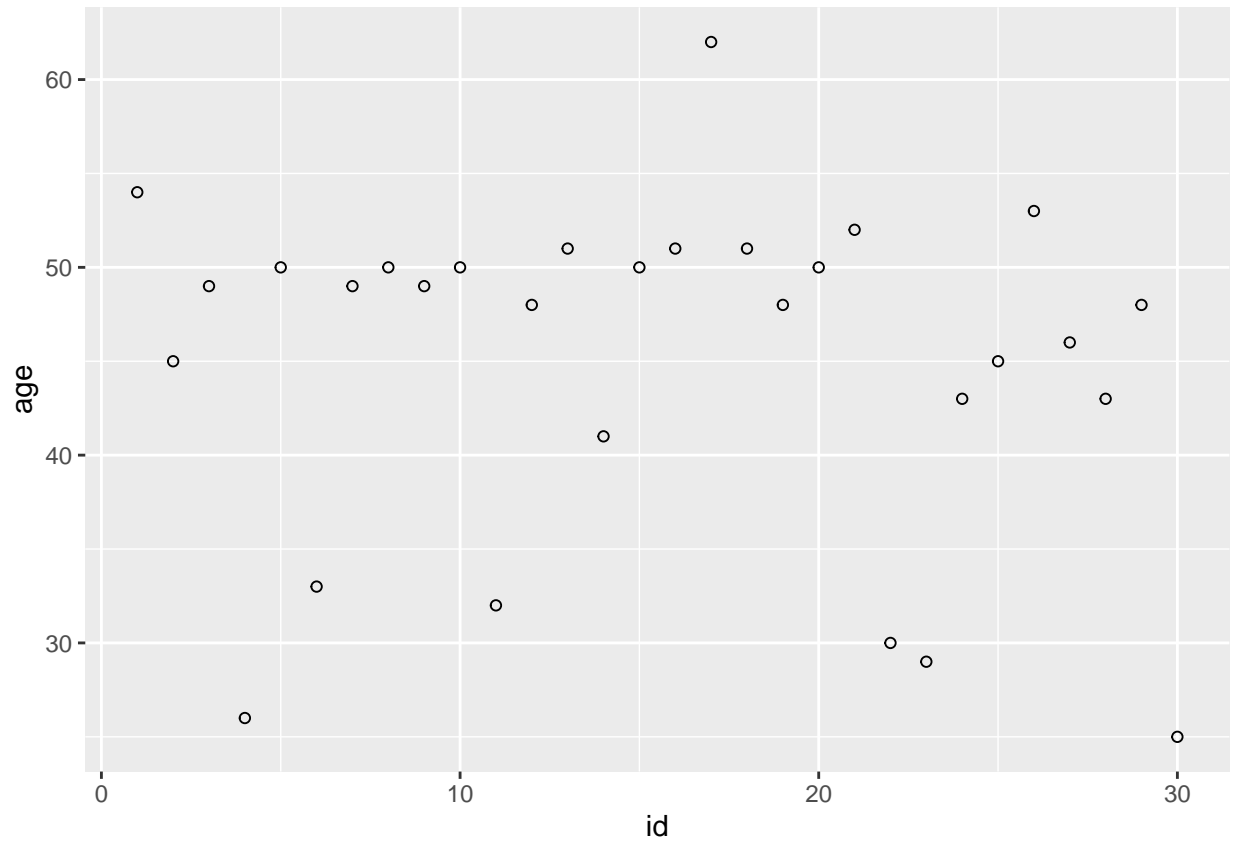
(1) 기본

```
ggplot(dat1, aes(x=id, y=age))+  
  geom_point()
```



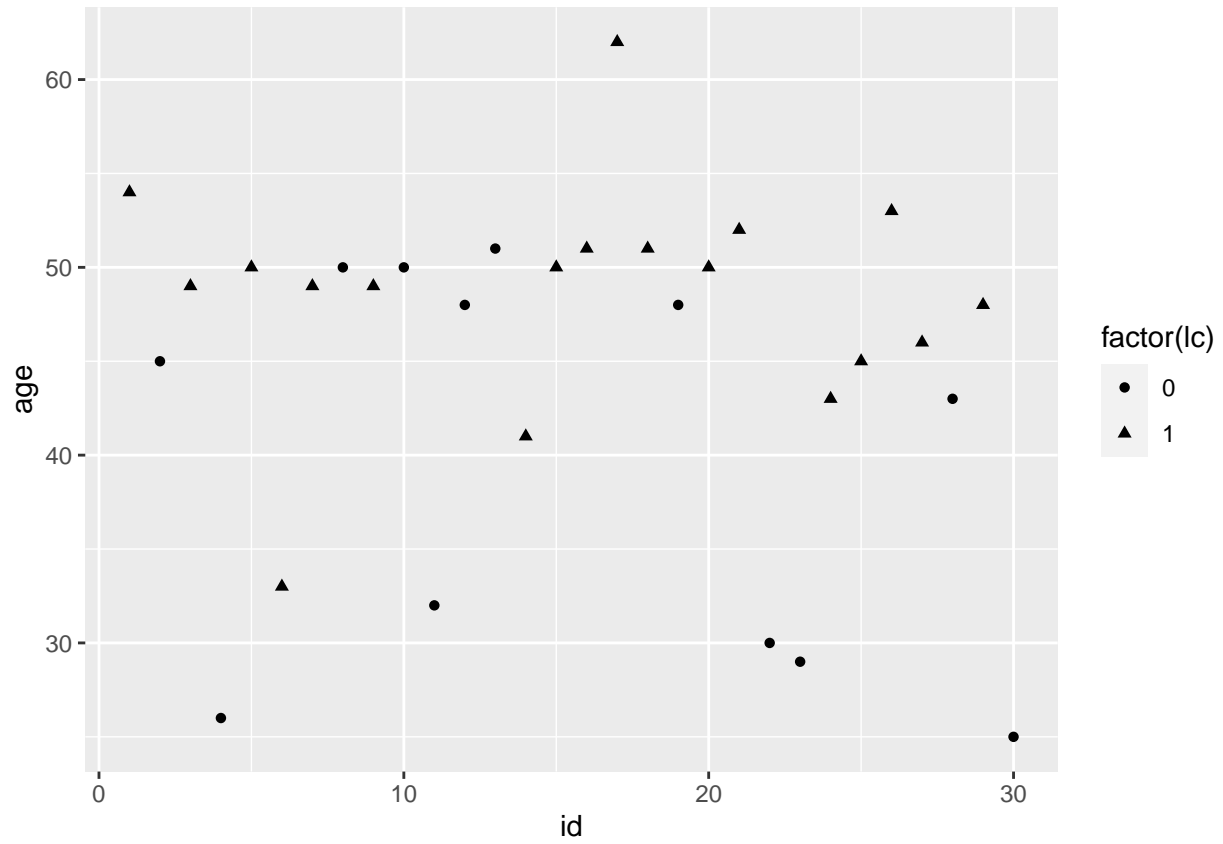
(2) 점 모양 변경

```
ggplot(dat1, aes(x=id, y=age))+  
  geom_point(shape=1)
```



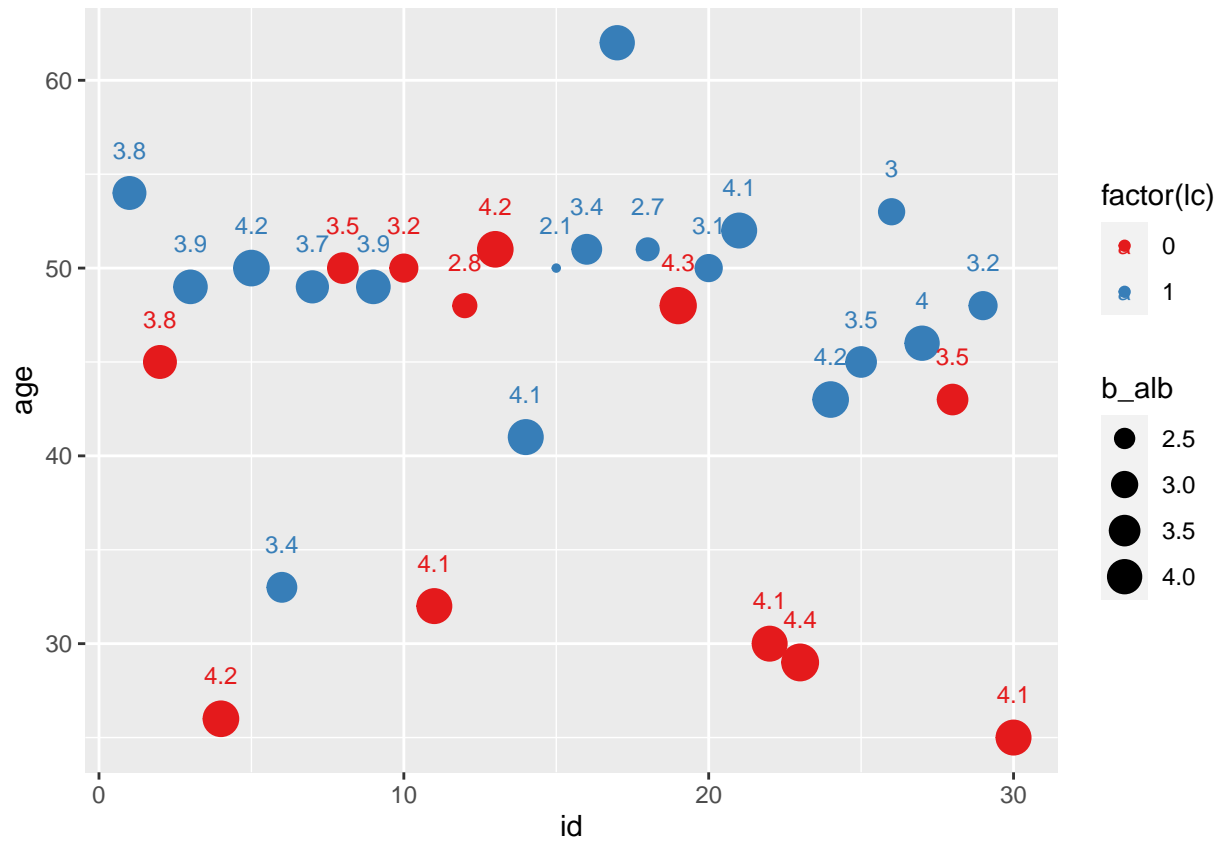
(3) 그룹에 따라 다르게 표시

```
ggplot(dat1, aes(x=id, y=age, shape=factor(lc)))+  
  geom_point()
```



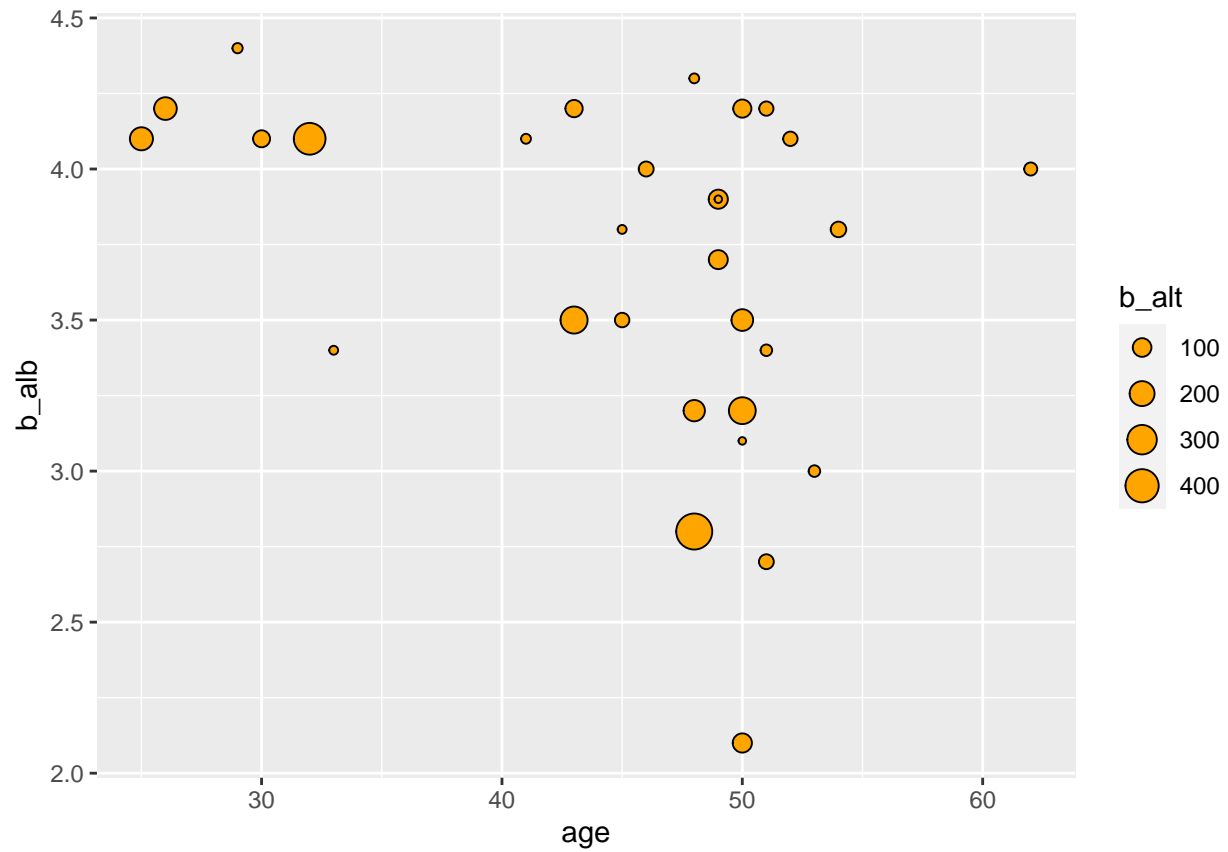
(4) 복합조건

```
ggplot(dat1, aes(x=id, y=age, color=factor(lc), size=b_alb))+  
  geom_point()+  
  geom_text(aes(label=b_alb), vjust=-2, size=3)+  
  scale_color_brewer(palette='Set1')
```



2.6 버블 그래프

```
ggplot(dat1, aes(x=age, y=b_alb))+  
  geom_point(aes(size=b_alt),shape=21, color='black', fill='orange')
```

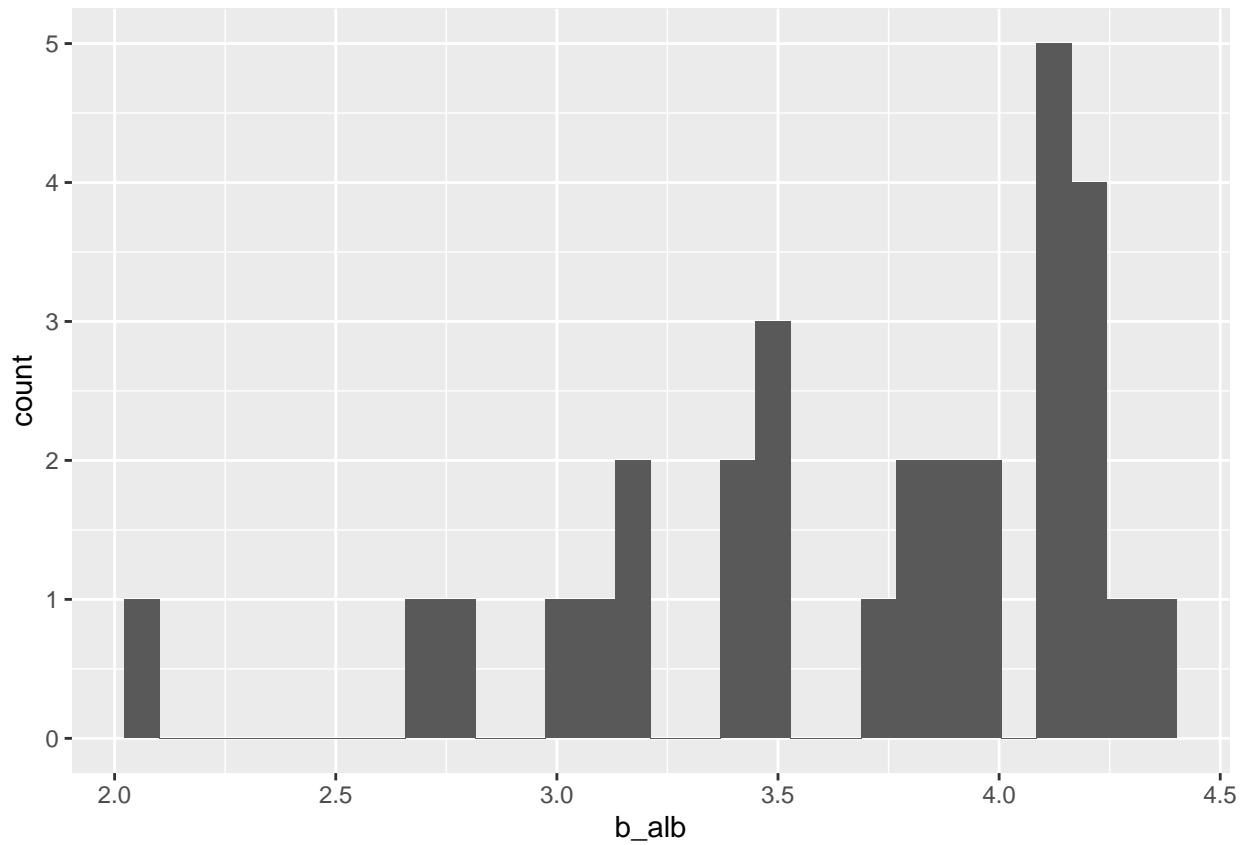


2.7 히스토그램

(1) 히스토그램 그리기

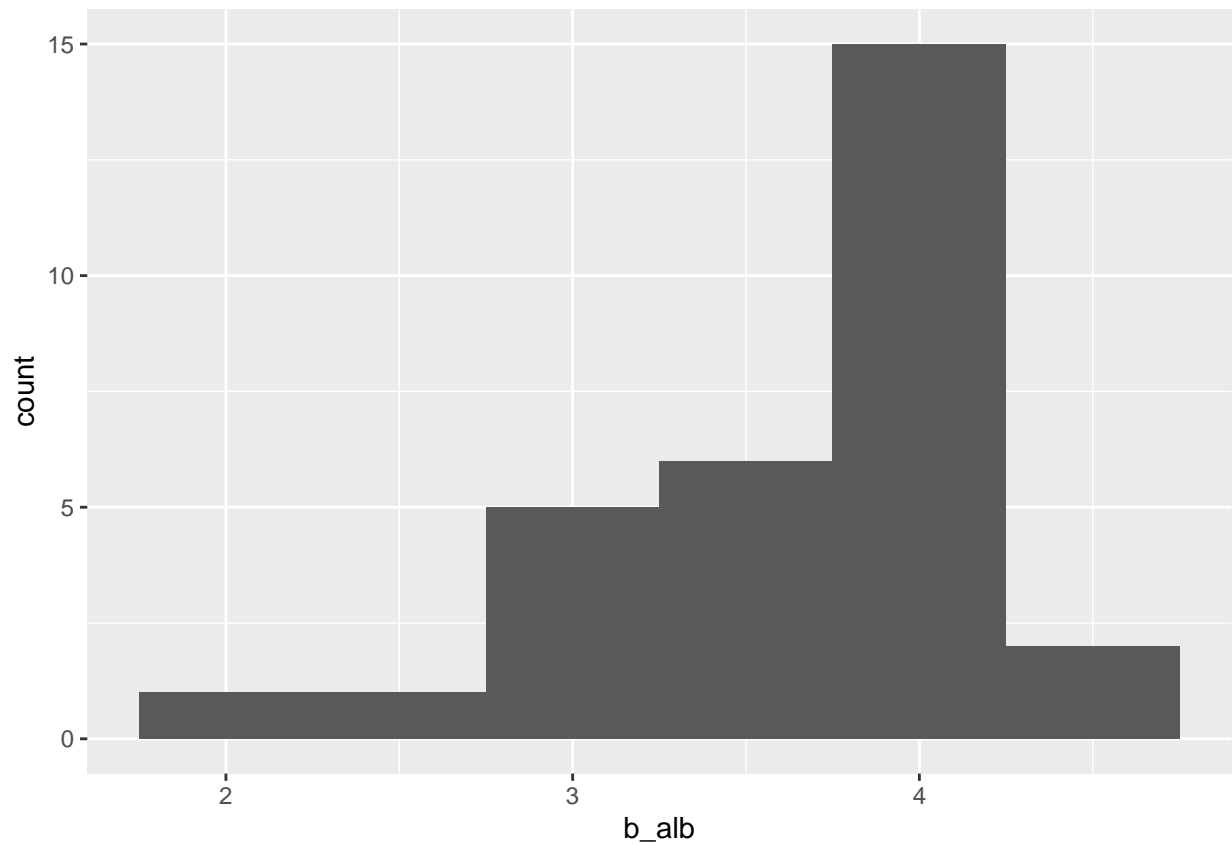
```
ggplot(dat1, aes(x=b_alb))+  
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



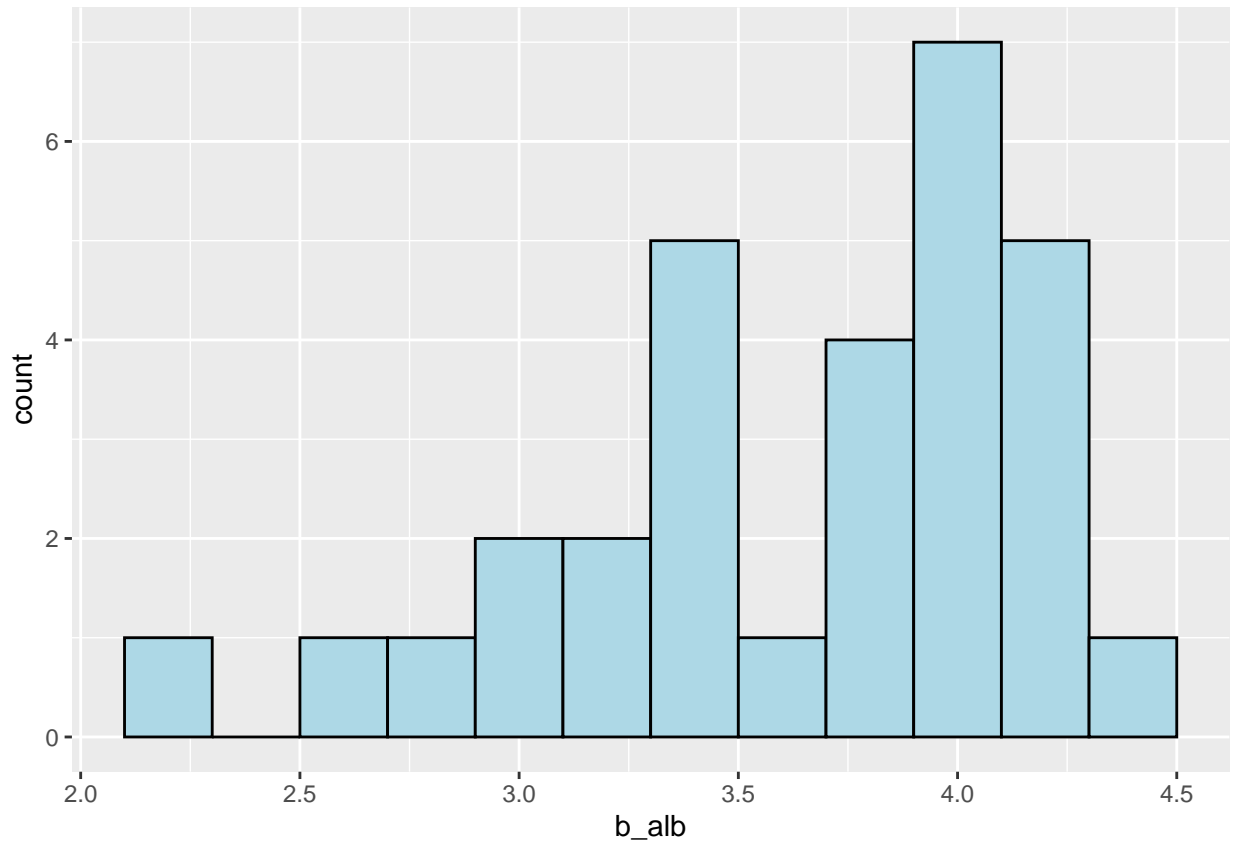
(2) 너비 조절

```
ggplot(dat1, aes(x=b_alb))+  
  geom_histogram(binwidth = 0.5)
```



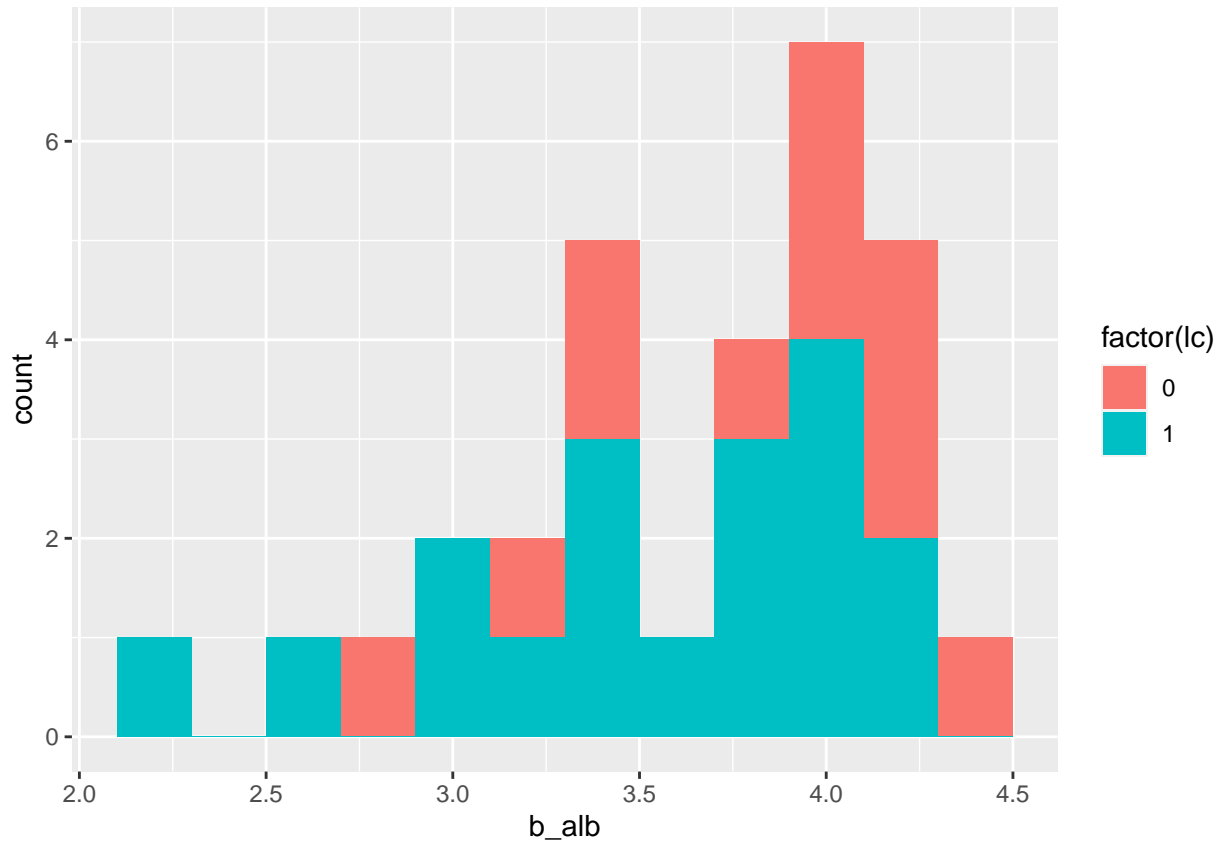
(3) 색상과 테두리

```
ggplot(dat1, aes(x=b_alb))+  
  geom_histogram(binwidth=0.2, fill='lightblue',color='black')
```



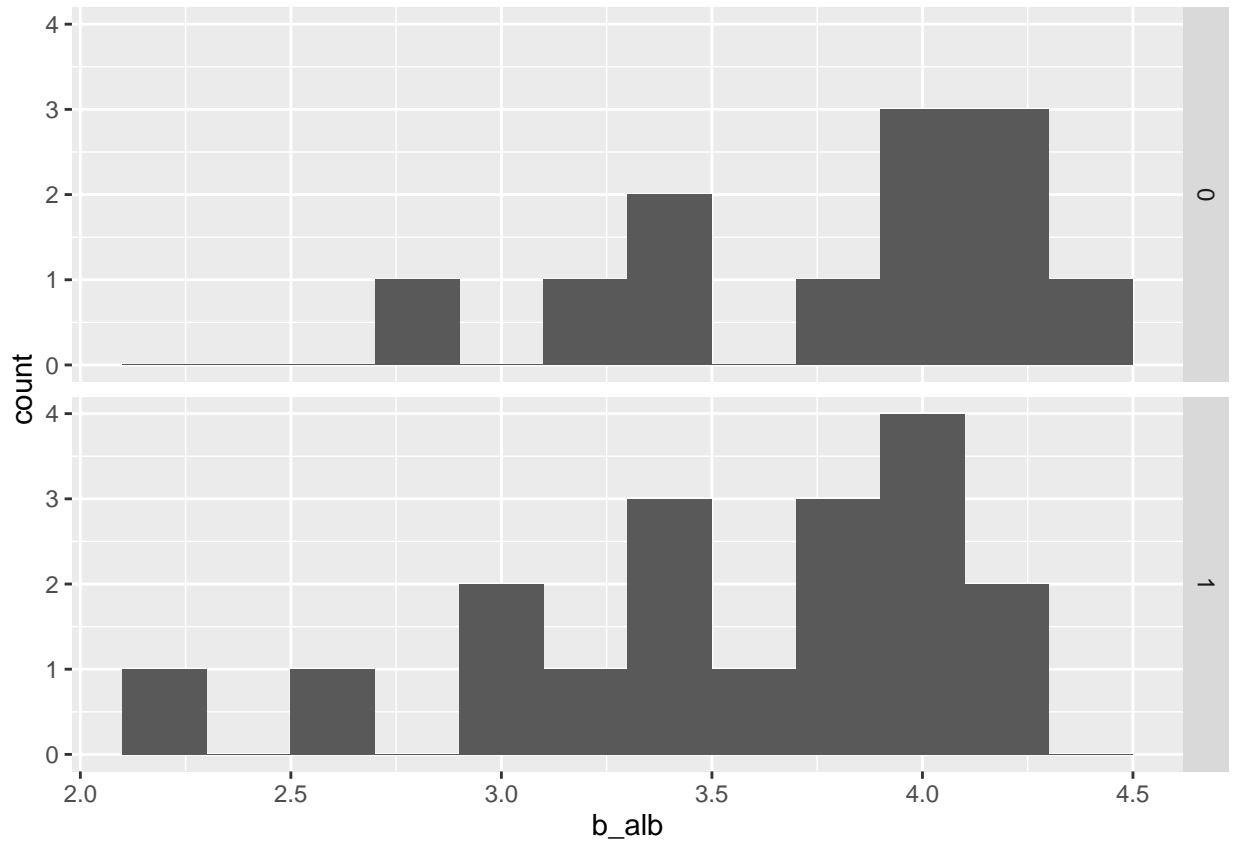
(4) 그룹별 히스토그램

```
ggplot(dat1, aes(x=b_alb, fill=factor(lc)))+  
  geom_histogram(binwidth = 0.2)
```



(5) 그룹별로 나누어서 히스토그램

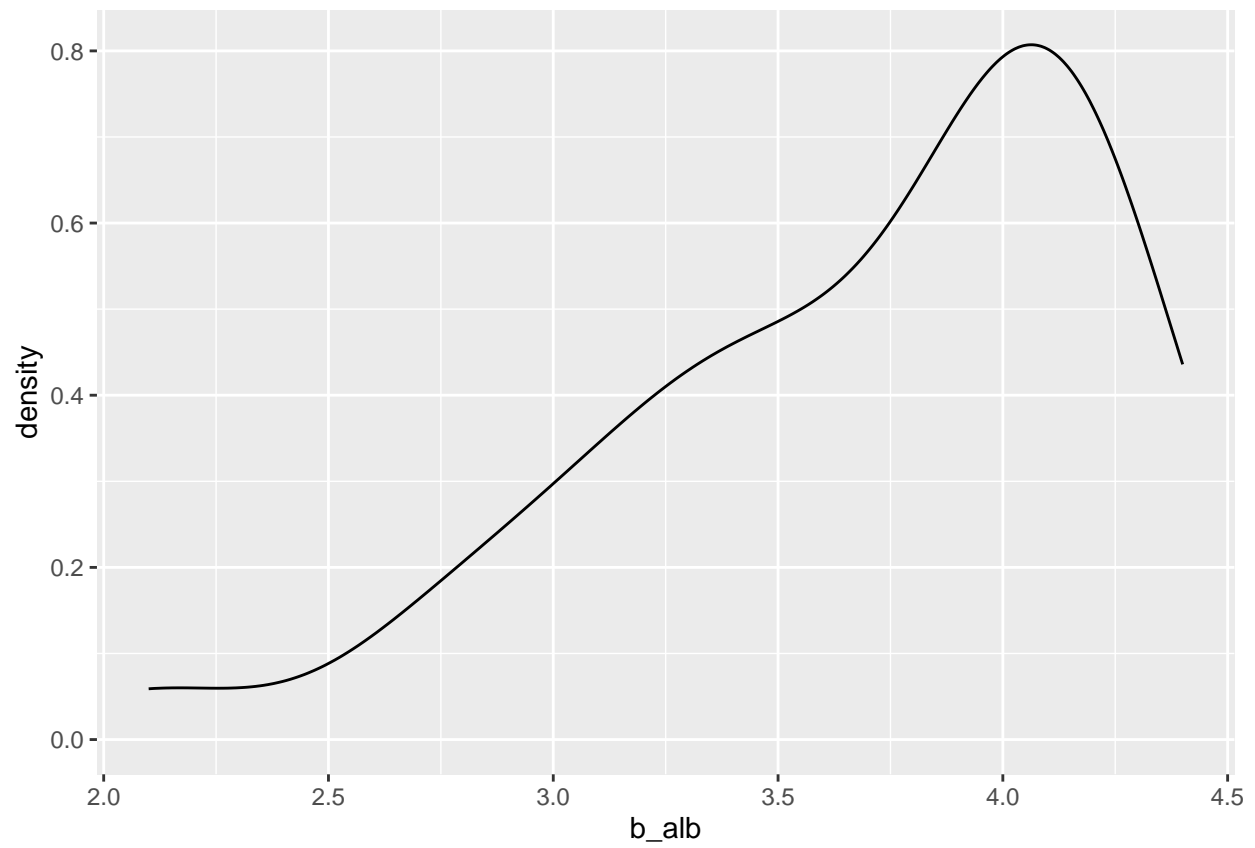
```
ggplot(dat1, aes(x=b_alb))+  
  geom_histogram(binwidth = 0.2)+  
  facet_grid(lc~.)
```



2.8 밀도 그래프

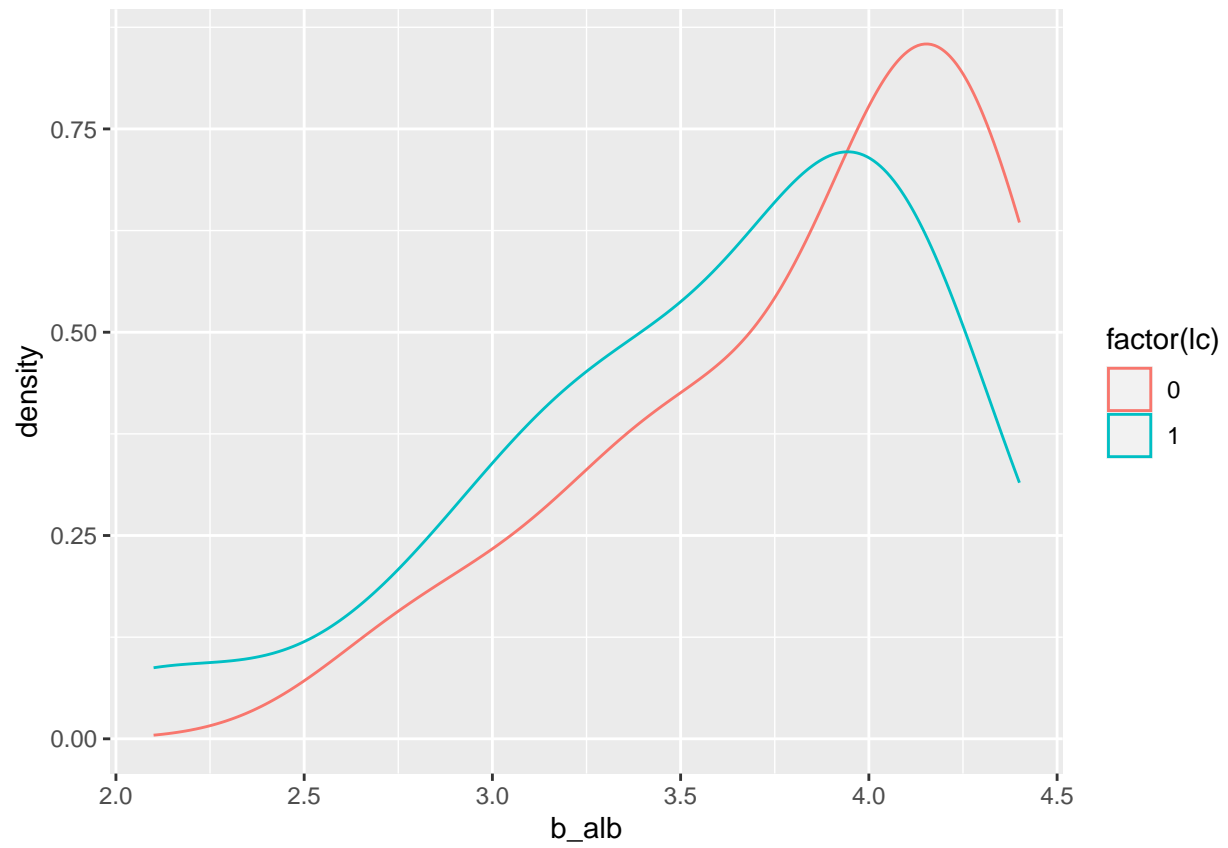
(1) 밀도 곡선 그리기

```
ggplot(dat1, aes(x=b_alb))+  
  geom_density()
```



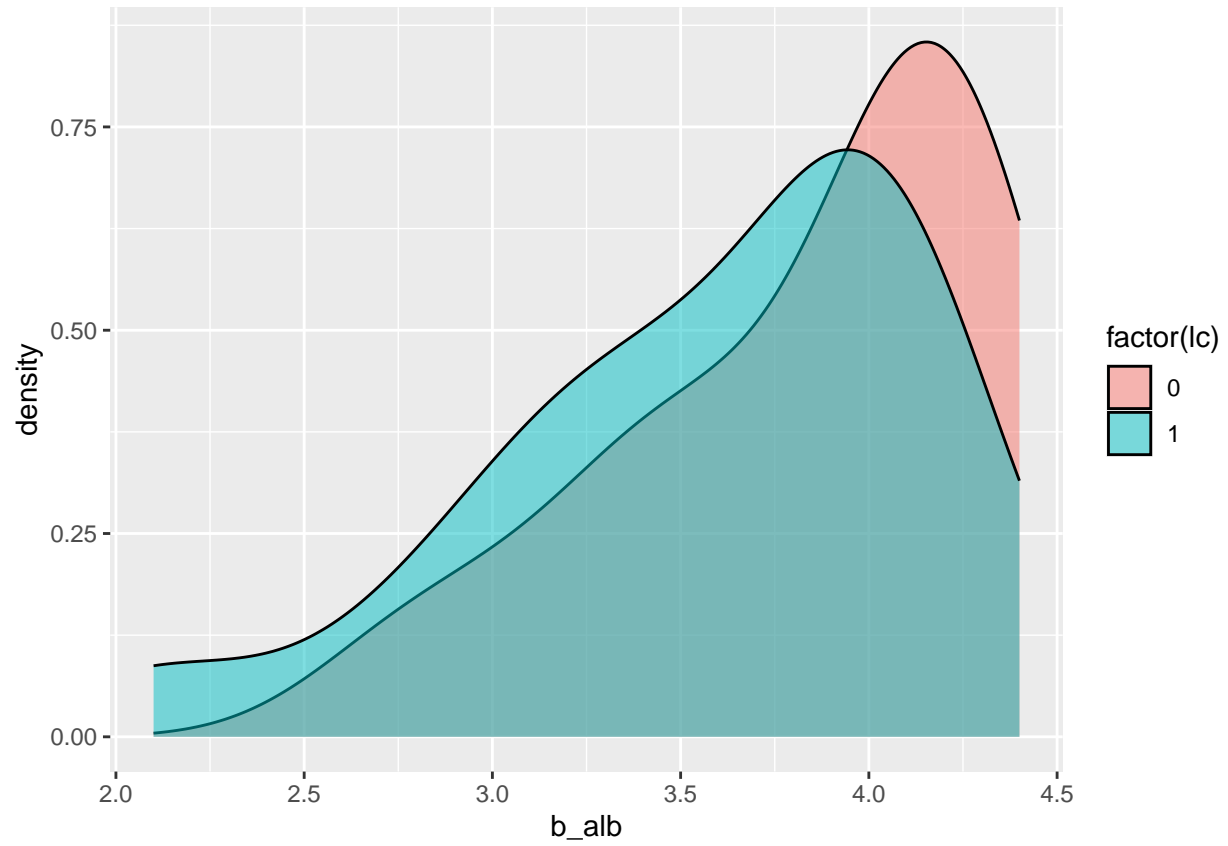
(2) 그룹별 밀도 곡선 겹쳐 그리기

```
ggplot(dat1, aes(x=b_alb, color=factor(lc)))+  
  geom_density()
```



(3) 그룹별 밀도 곡선 겹쳐 그리기/영역 다르게

```
ggplot(dat1, aes(x=b_alb, fill=factor(lc)))+  
  geom_density(alpha=0.5)
```



3. ggplot2의 다양한 옵션

알부민과 관련된 변수들을 중심으로 선택해서 데이터를 새로 만들었음

```
albu<-dat1 %>%  
  select(id, age, gender, treat_gr, lc, contains('alb')) %>%  
  gather(6:10, key='observation', value='albumin')
```

변수들 확인

```
str(albu)
```

```
## 'data.frame':   150 obs. of  7 variables:  
## $ id          : int  1 2 3 4 5 6 7 8 9 10 ...  
## $ age         : int  54 45 49 26 50 33 49 50 49 50 ...  
## $ gender      : chr  "M" "F" "M" "M" ...  
## $ treat_gr    : chr  "ETV" "ETV" "ETV" "ETV" ...  
## $ lc         : int  1 0 1 0 1 1 1 0 1 0 ...  
## $ observation: chr  "b_alb" "b_alb" "b_alb" "b_alb" ...  
## $ albumin     : num  3.8 3.8 3.9 4.2 4.2 3.4 3.7 3.5 3.9 3.2 ...
```

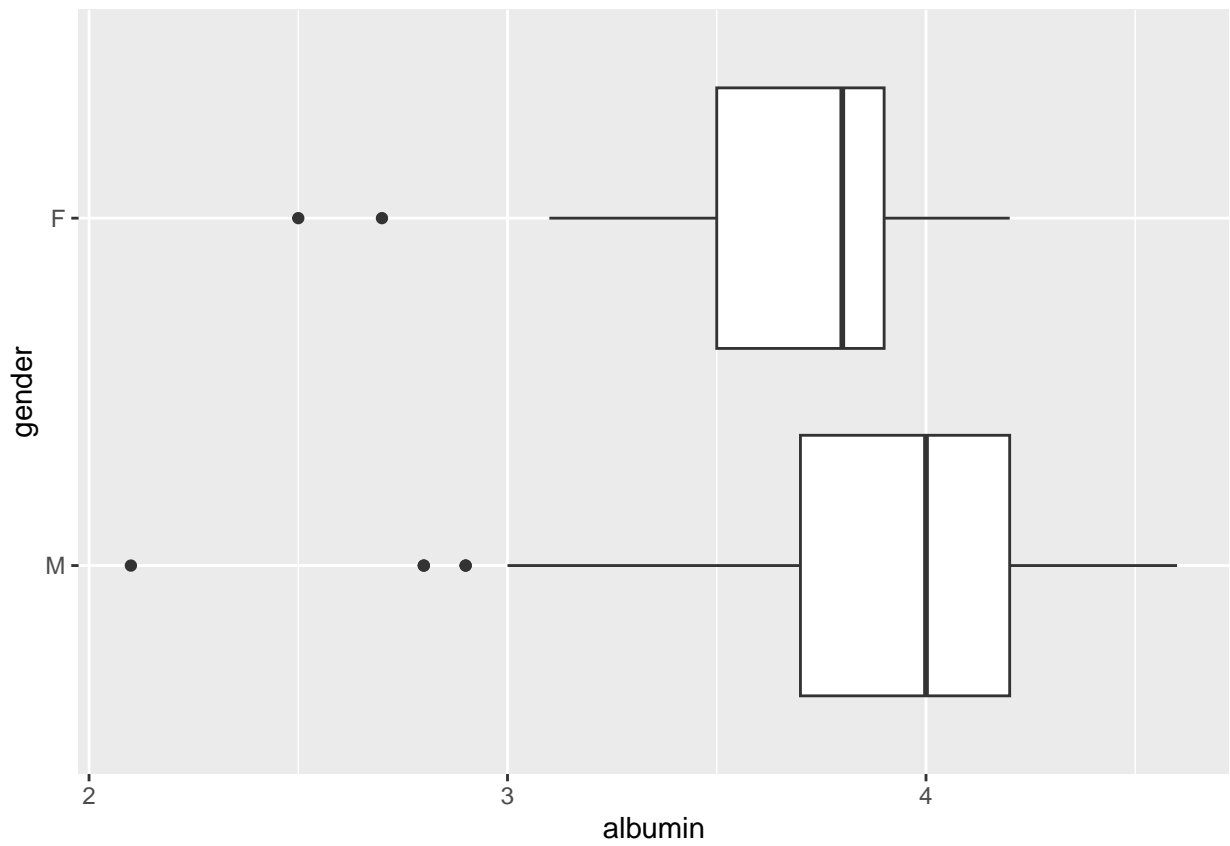
factor형으로 변경

```
albu$gender<-factor(albu$gender, levels = c('M', 'F'))  
albu$treat_gr<-factor(albu$treat_gr)  
albu$lc<-factor(albu$lc)
```

3.1 축

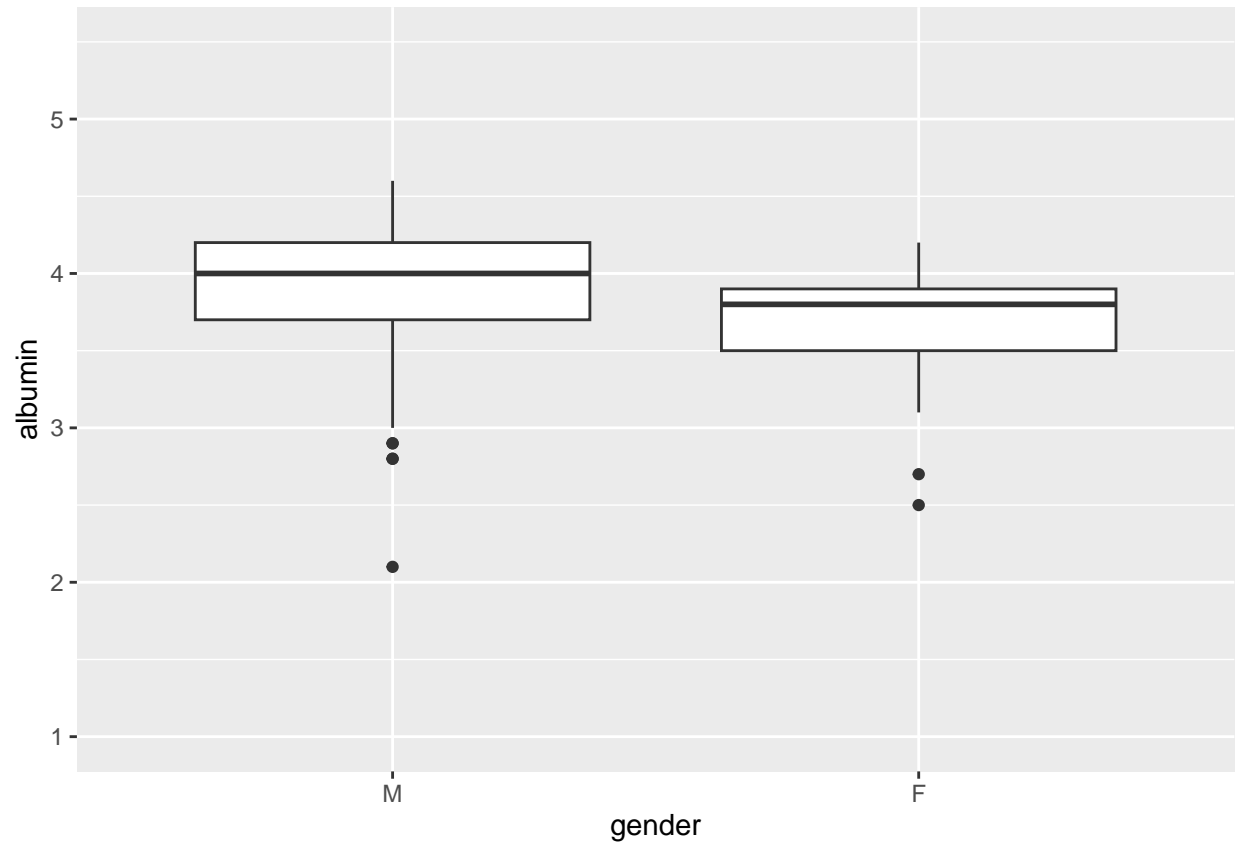
(1) x축y축 서로 바꾸기

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  coord_flip()
```



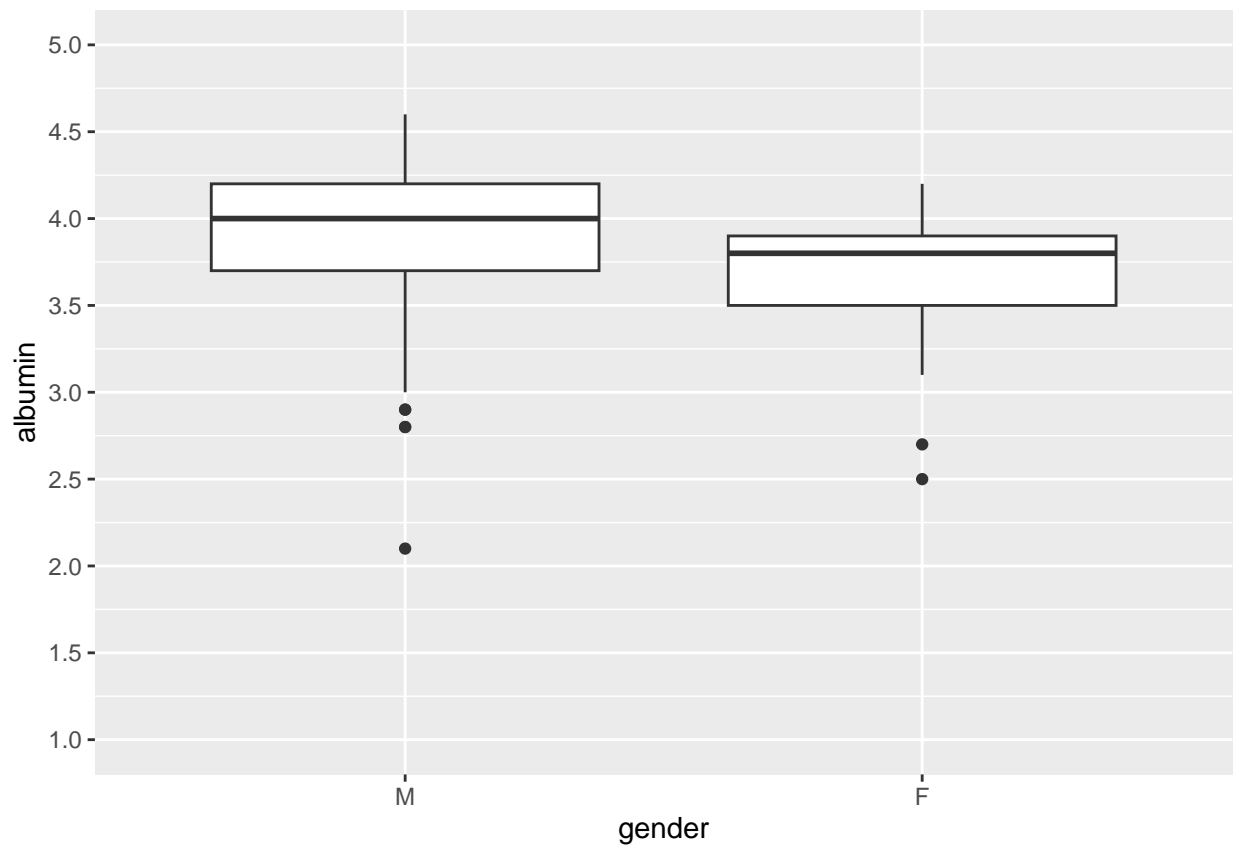
(2) 축의 범위 설정

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  ylim(1,5.5)
```



(3) 축의 작은 눈금 설정

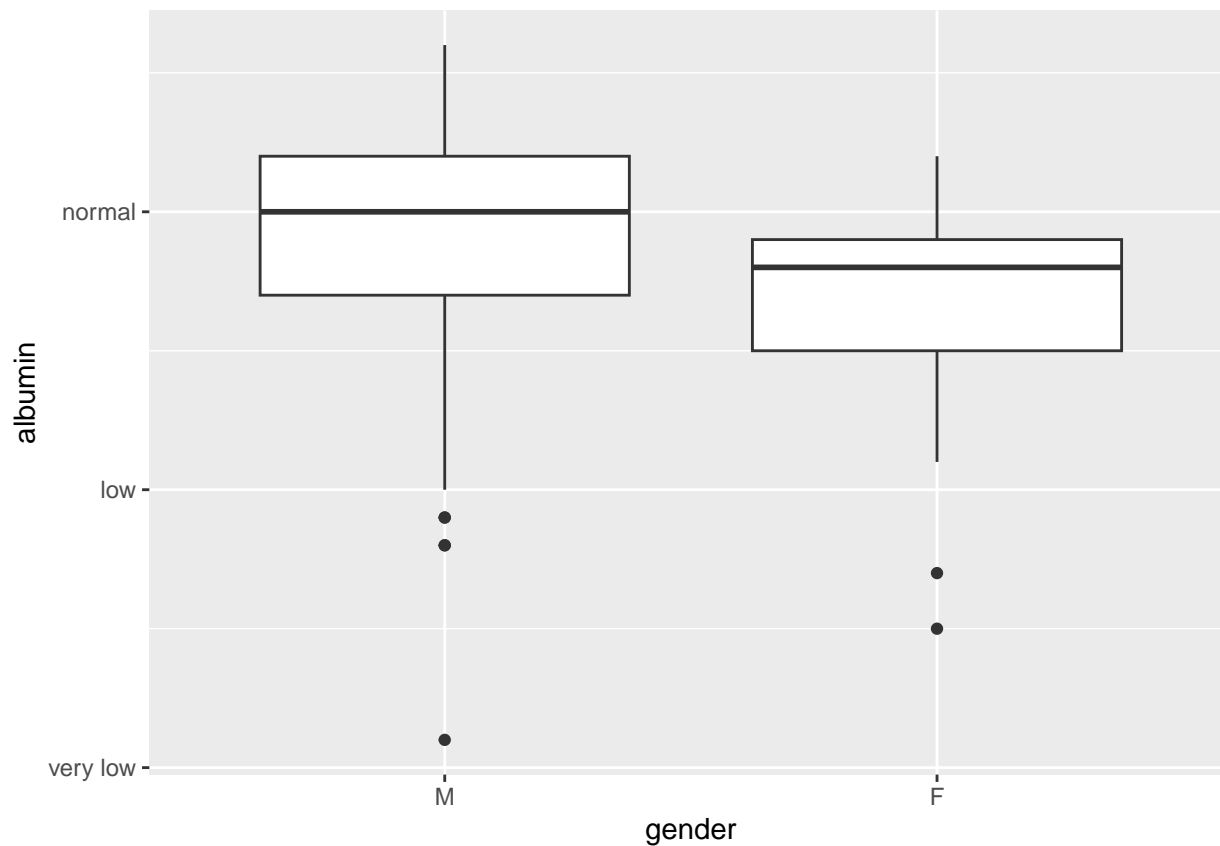
```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  scale_y_continuous(limits=c(1,5), breaks=c(seq(1,5,0.5)))
```



ylim랑 scale_y_continuous 같이 사용할 수 없음

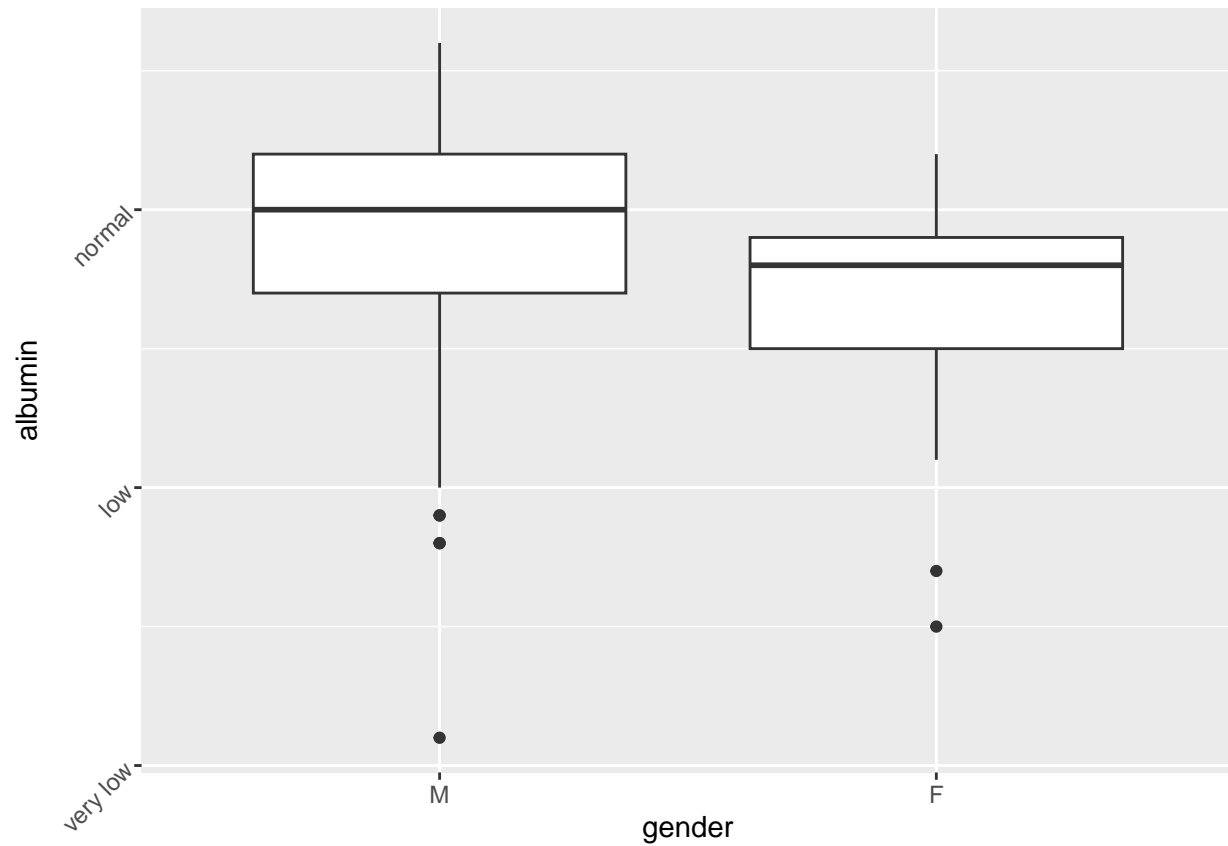
(4) 축 눈금 임의로 설정

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  scale_y_continuous(labels=c('very low', 'low', 'normal', 'high'))
```



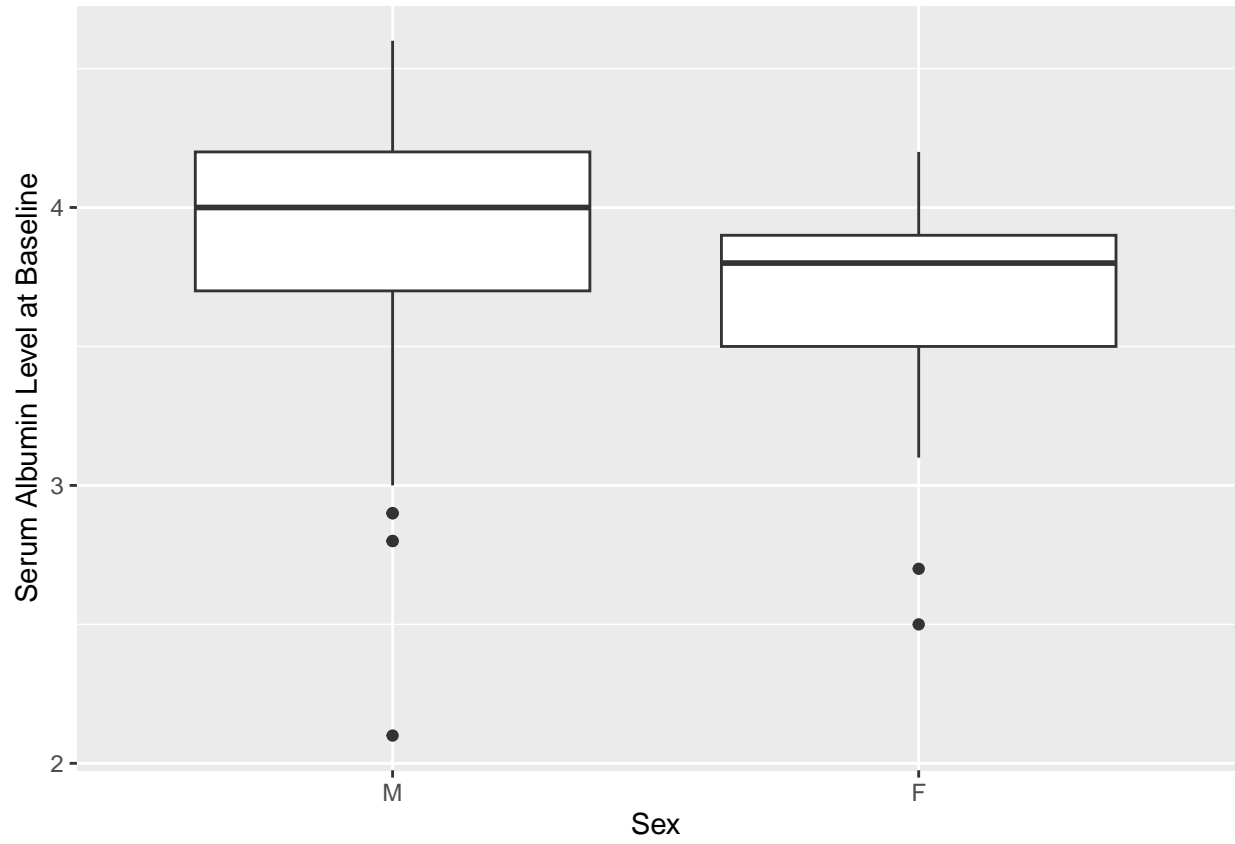
(5) 축의 이름의 위치,형태 변경

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  scale_y_continuous(labels=c('very low','low','normal','high'))+  
  theme(axis.text.y=element_text(angle=45))
```



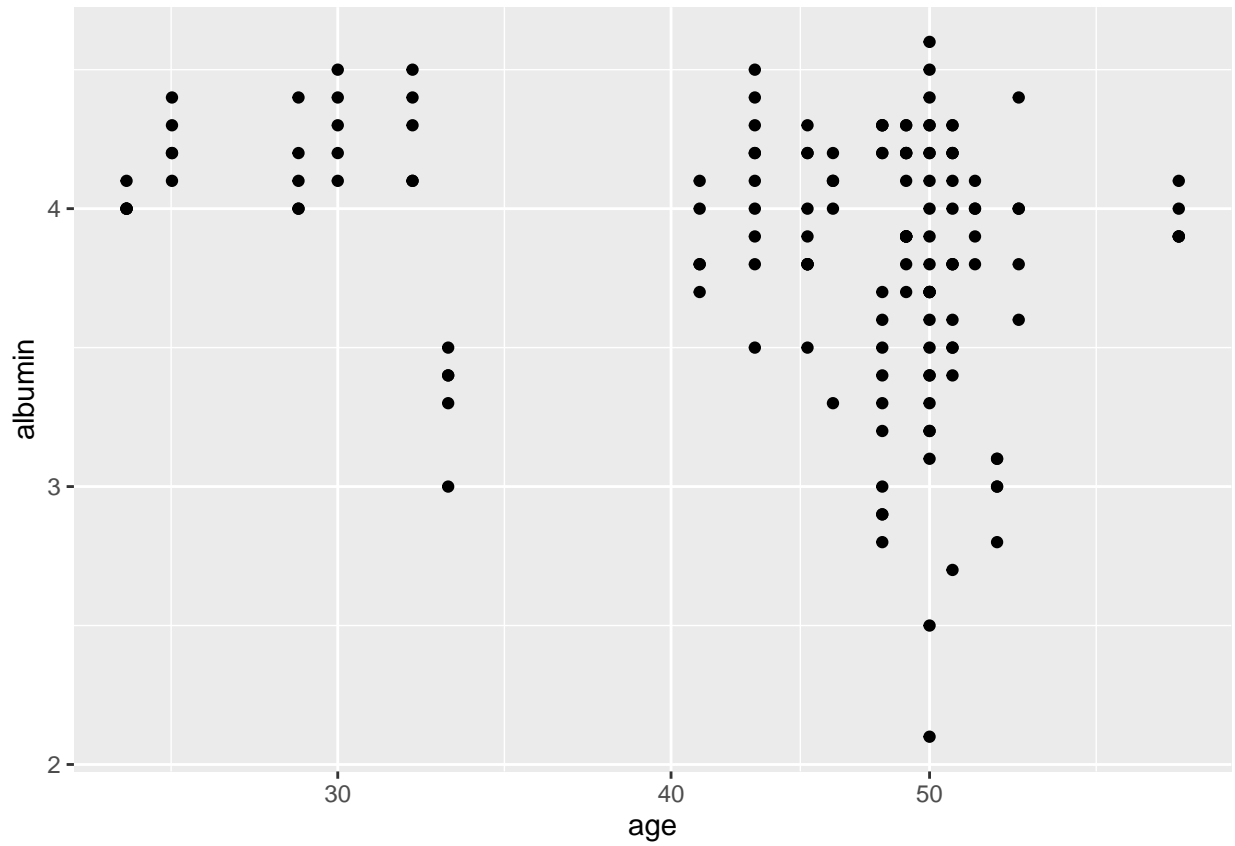
(6) 축 이름 변경

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  xlab('Sex')+  
  ylab('Serum Albumin Level at Baseline')
```



(7) 로그 변환 축 사용

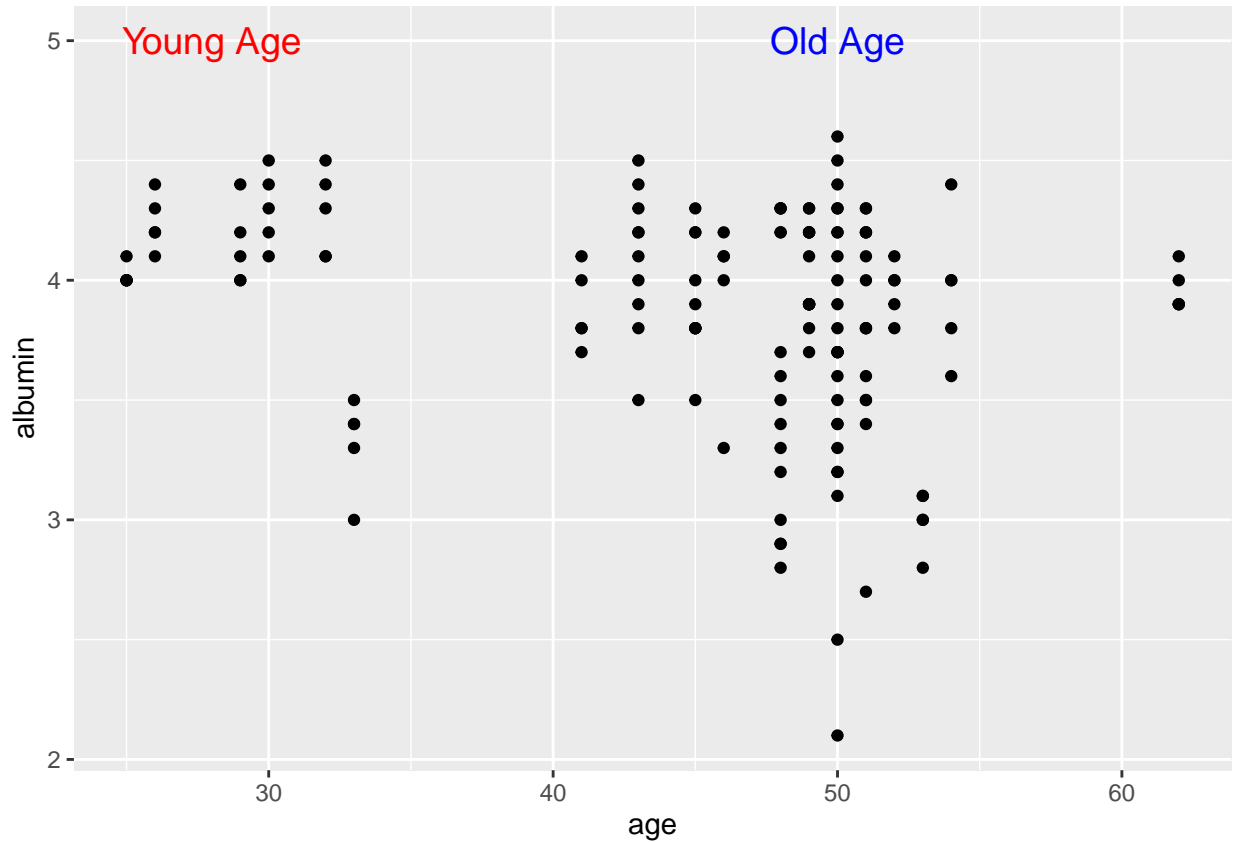
```
ggplot(albu, aes(x=age, y=albumin))+  
  geom_point()+  
  scale_x_log10()
```



3.2 주석

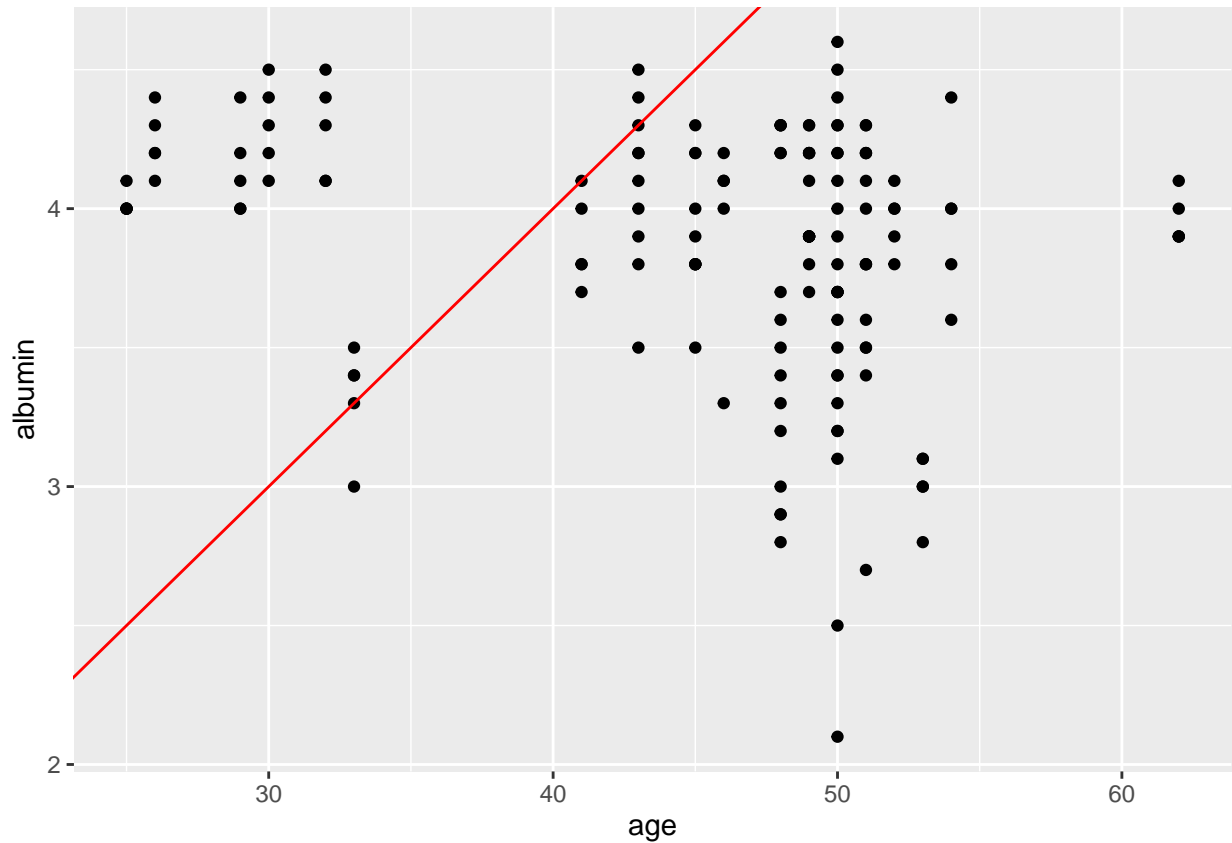
(1) 텍스트 주석 넣기

```
ggplot(albu, aes(x=age, y=albumin))+  
  geom_point()+  
  annotate('text',x=28, y=5, label='Young Age', color='red', size=5)+  
  annotate('text',x=50, y=5, label='Old Age', color='blue', size=5)
```



(2) 선 추가하기

```
ggplot(albu, aes(x=age, y=albumin))+  
  geom_point()+  
  geom_abline(intercept=0, slope=0.1, color='red')
```



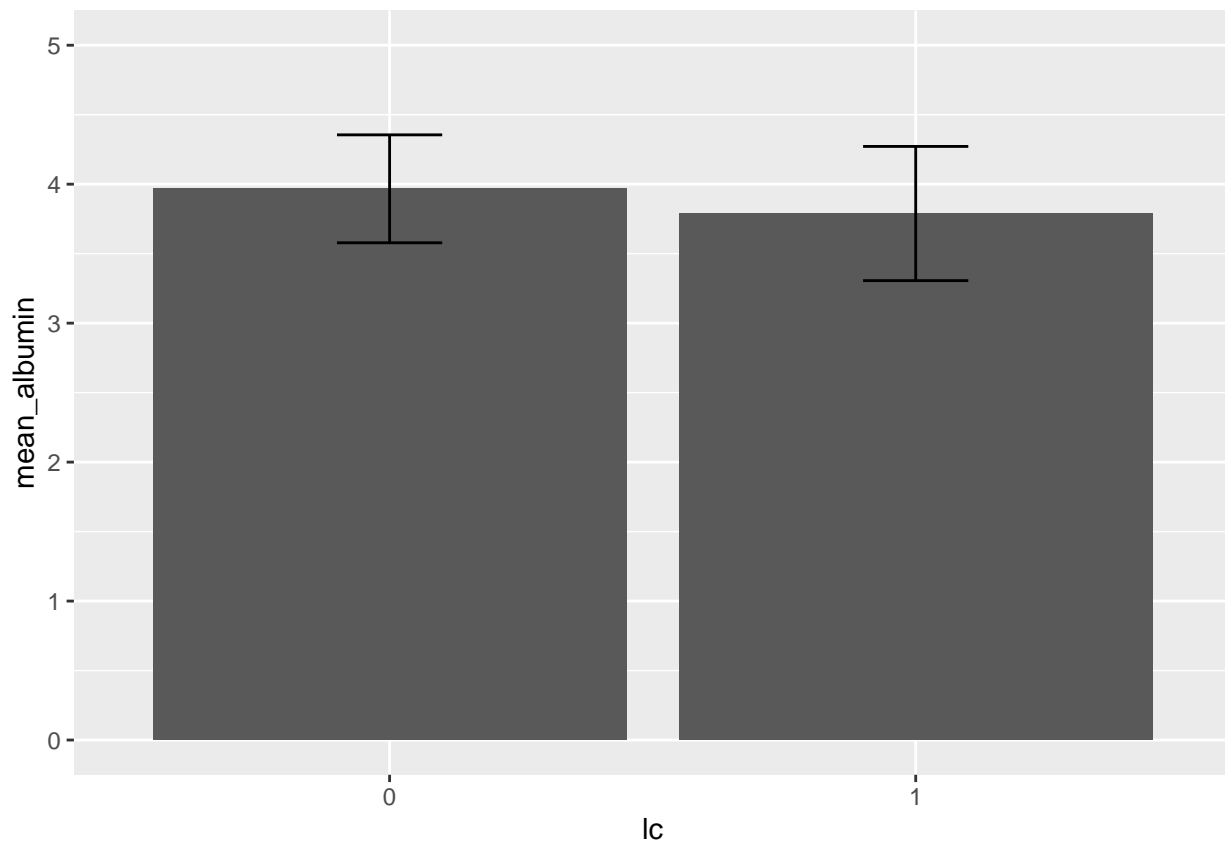
hline:수평선 vline:수직선 abline:기울기와 절편이 있는 직선

(3) 오차 막대 추가

```
albu1<-albu %>%  
  group_by(lc) %>%  
  summarize( mean_albumin = mean(albumin, na.rm=T),  
             sd_albumin = sd(albumin, na.rm=T))  
head(albu1)
```

```
## # A tibble: 2 x 3  
##   lc   mean_albumin sd_albumin  
##   <fct>         <dbl>     <dbl>  
## 1 0             3.97       0.388  
## 2 1             3.79       0.483
```

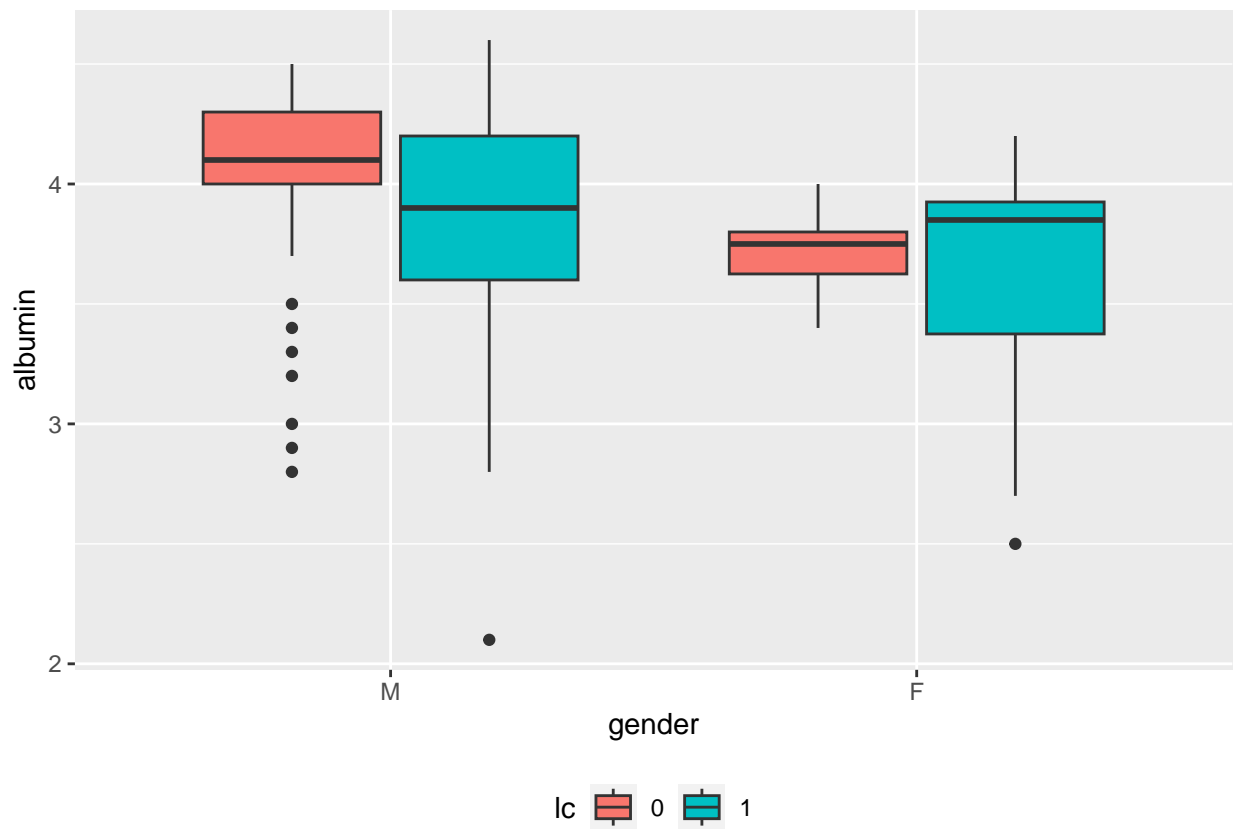
```
ggplot(albu1, aes(x=lc, y=mean_albumin))+  
  geom_bar(stat='identity')+  
  ylim(c(0,5))+  
  geom_errorbar(aes(ymin=mean_albumin-sd_albumin,  
                   ymax=mean_albumin+sd_albumin), width=0.2)
```



3.3 legend

(1) 위치 변경

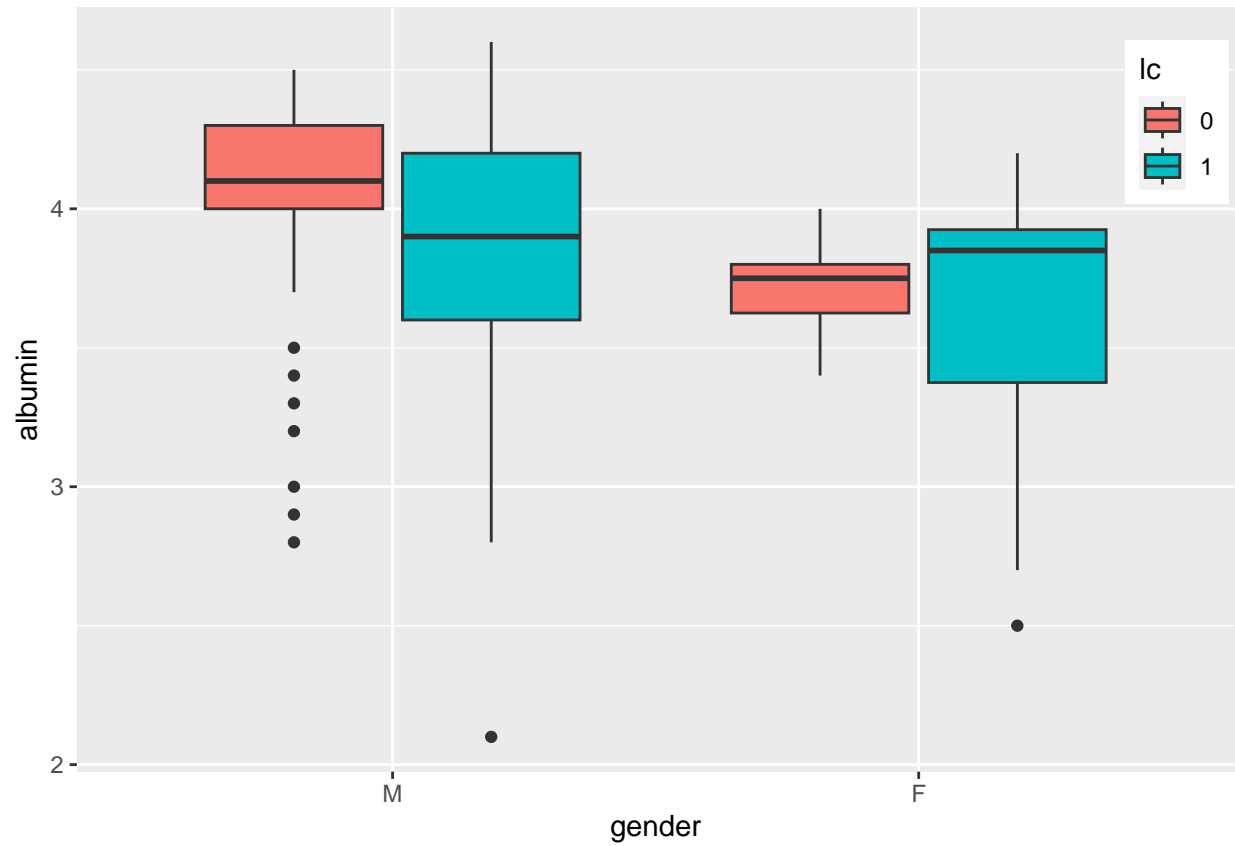
```
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+  
  geom_boxplot()+  
  theme(legend.position = 'bottom')
```



position에는 'top', 'bottom', 'left', 'right'가 있음

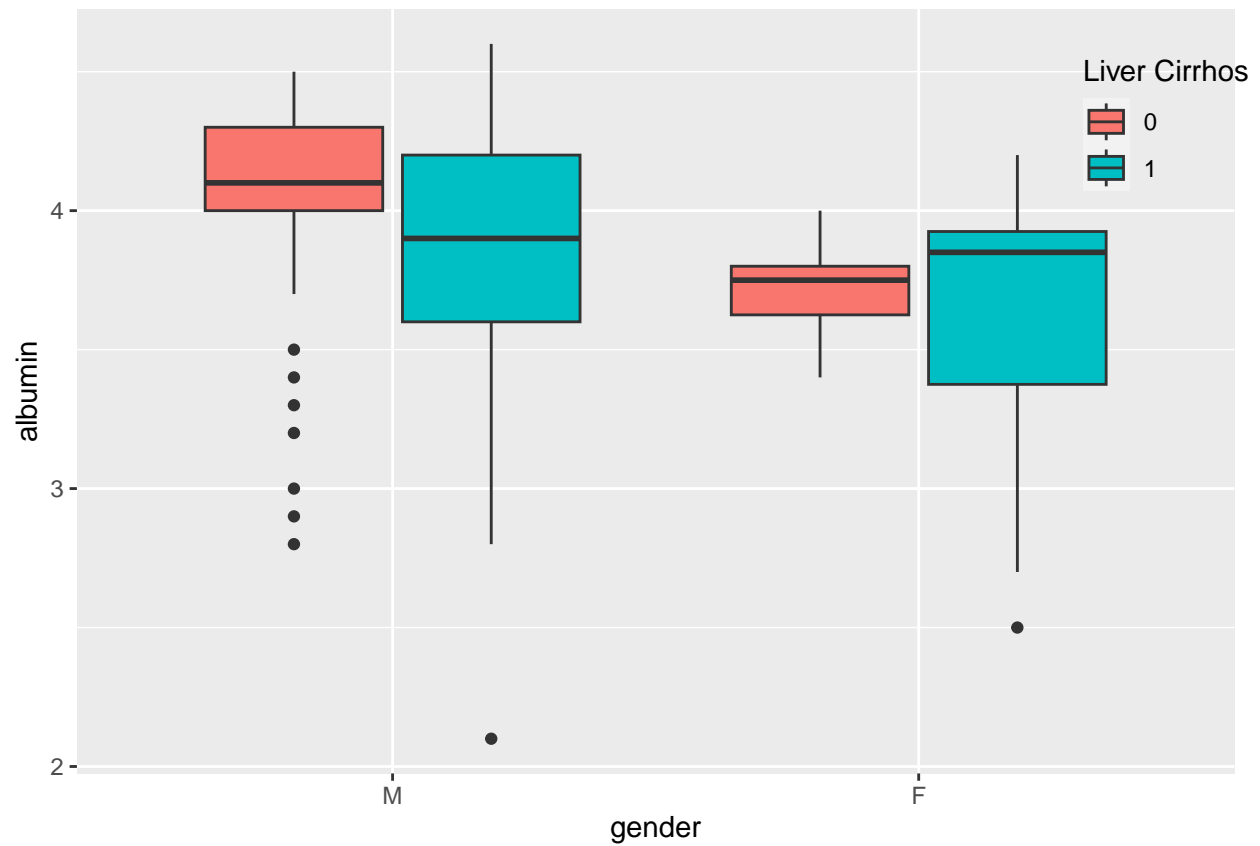
(2) 그래프 안에 포함

```
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+  
  geom_boxplot()+  
  theme(legend.position = c(0.95,0.85))
```



(3) 제목 바꾸기

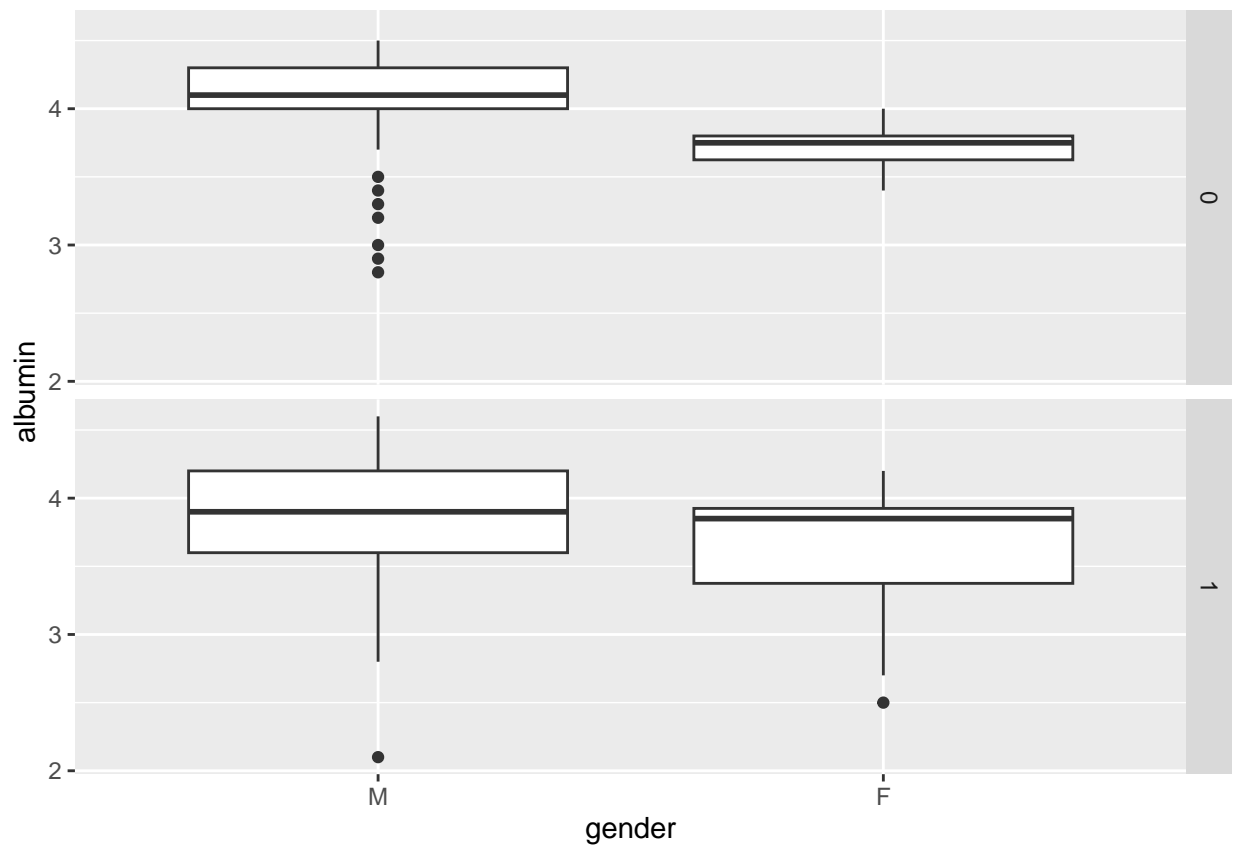
```
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+  
  geom_boxplot()+  
  theme(legend.position = c(0.95,0.85),  
        legend.background = element_blank())+  
  labs(fill='Liver Cirrhosis')
```



3.4 분할

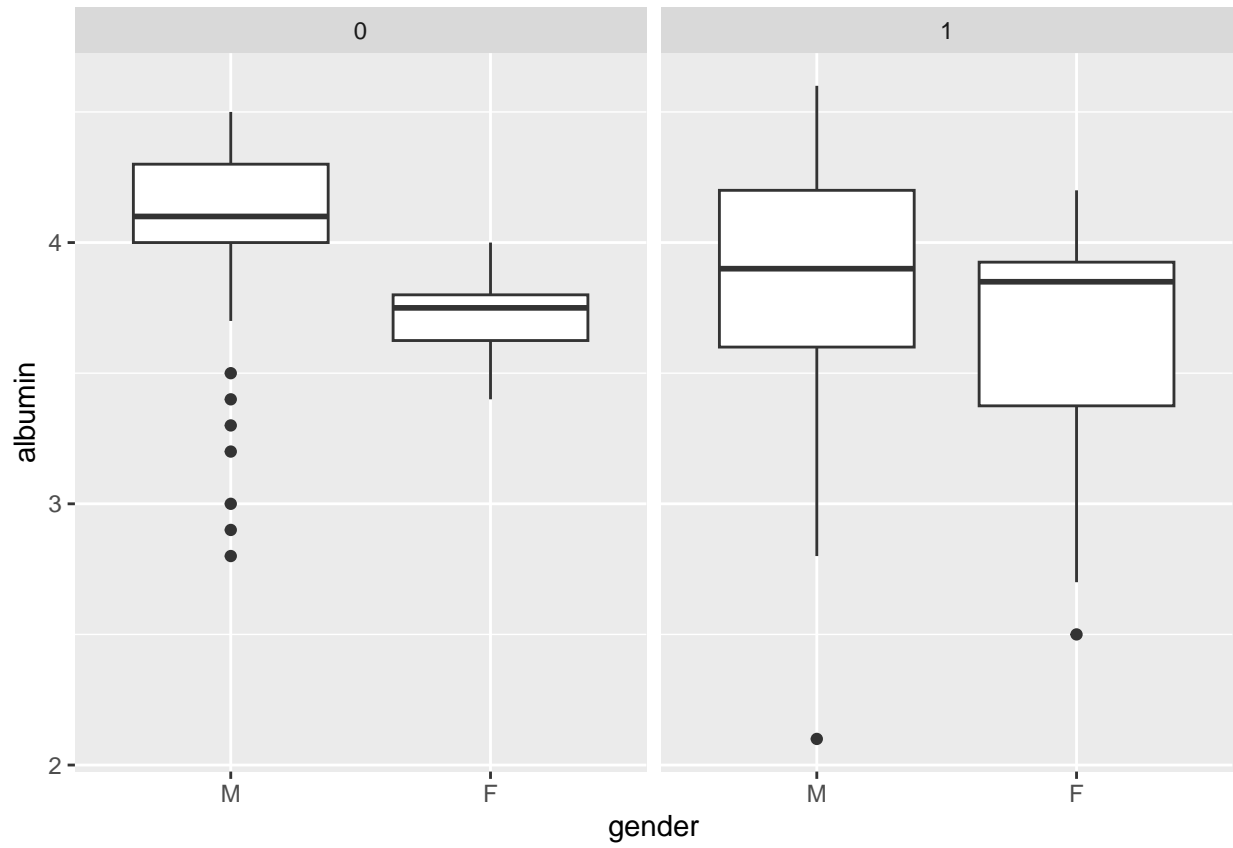
(1) 변수에 따라 화면 자동 분할

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  facet_grid(lc~.)
```



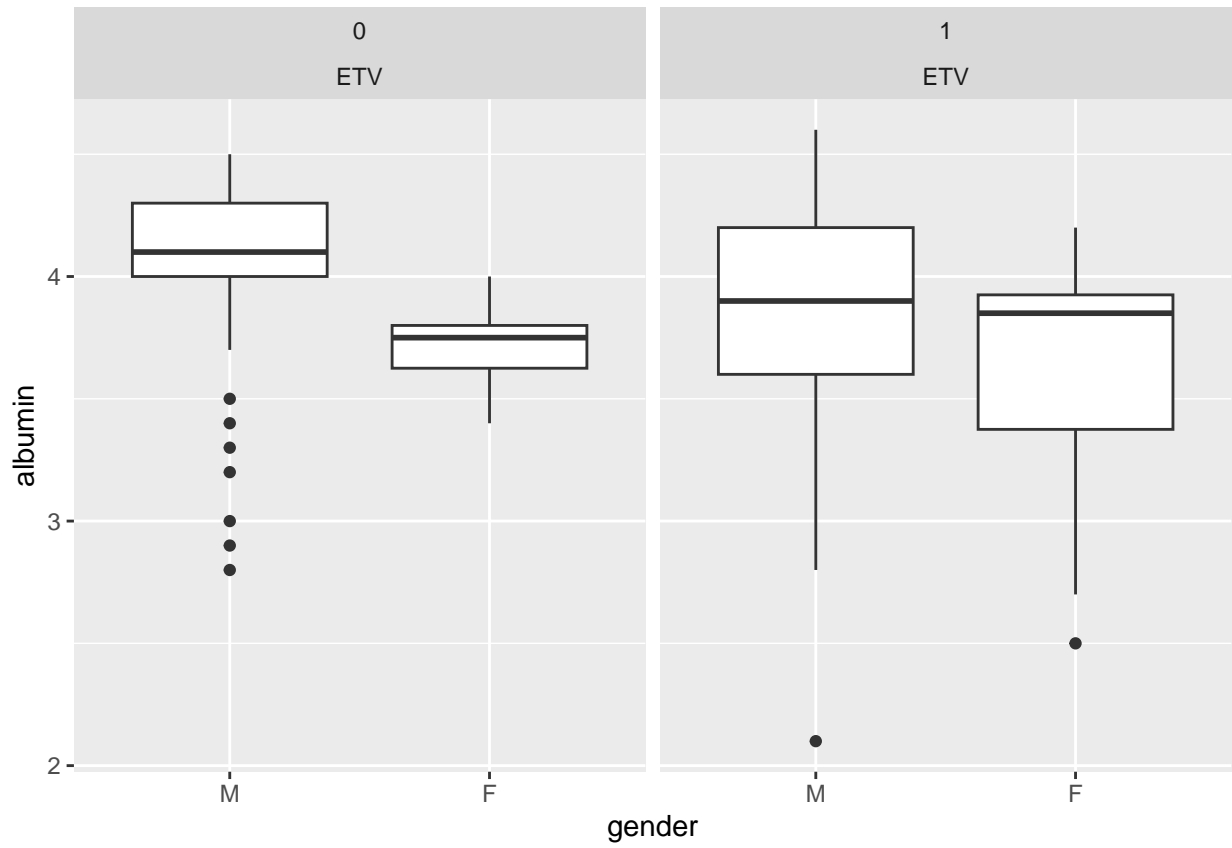
세로로 나누어보기

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  facet_grid(~lc)
```



동시 분할

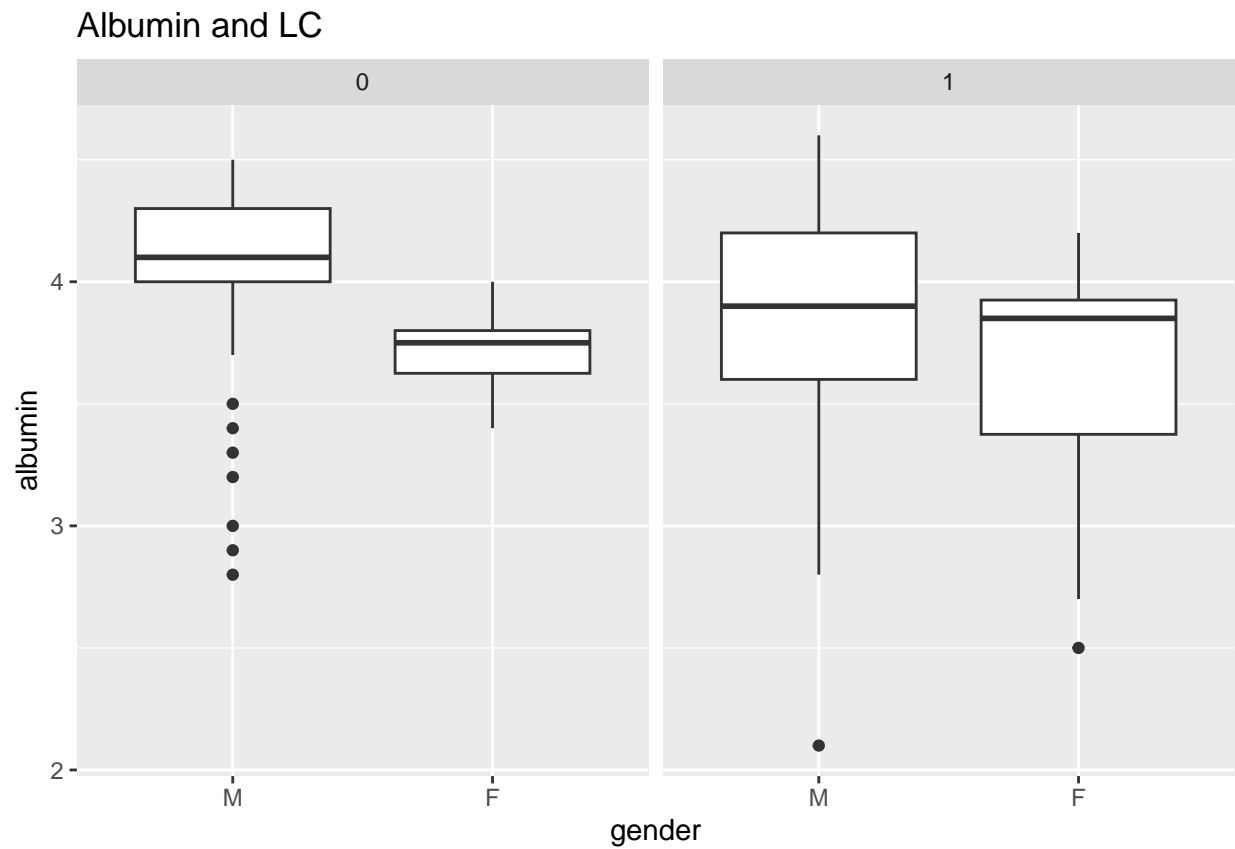
```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  facet_grid(~lc+treat_gr)
```



3.5 테마

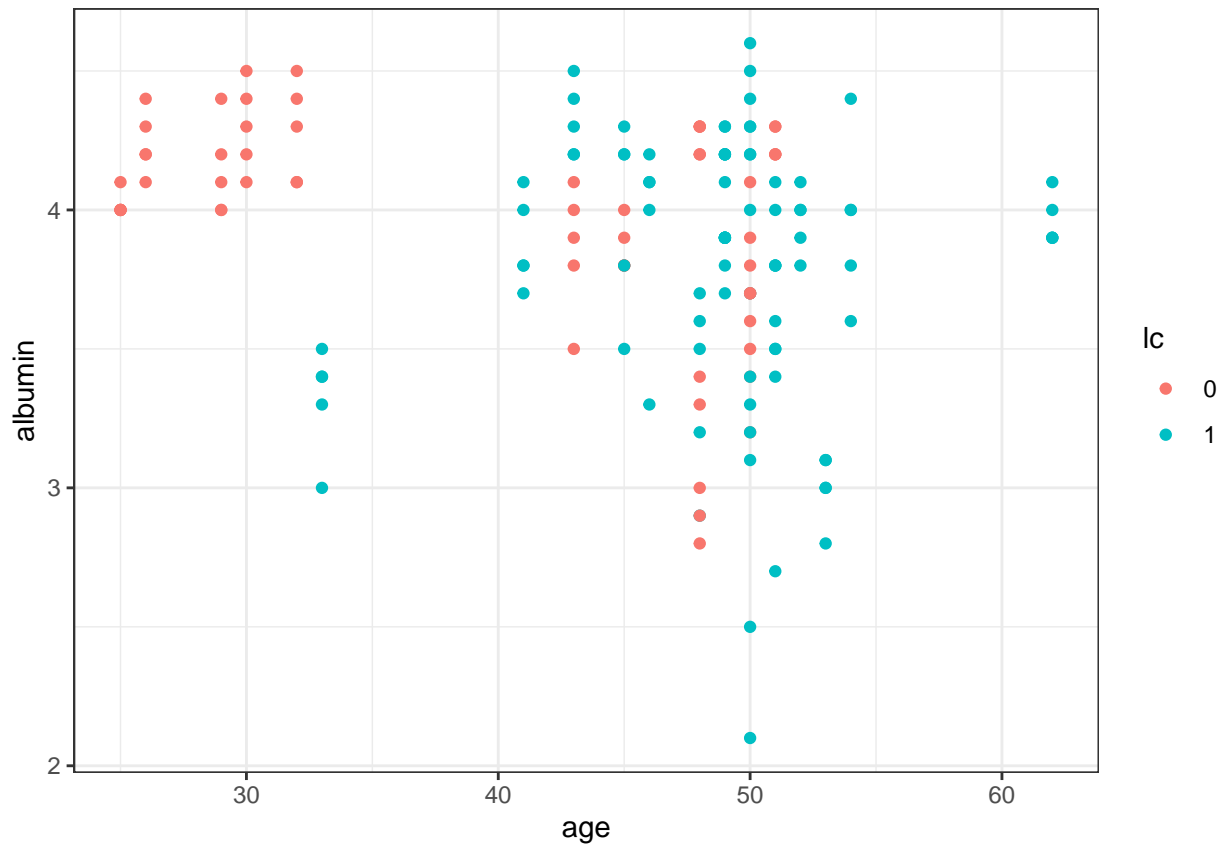
(1) 그래프 제목 붙이기

```
ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  facet_grid(~lc)+  
  ggtitle('Albumin and LC')
```



(2) 기본으로 제공하는 테마들

```
ggplot(albu, aes(x=age, y=albumin, color=lc))+  
  geom_point()+  
  theme_bw()
```



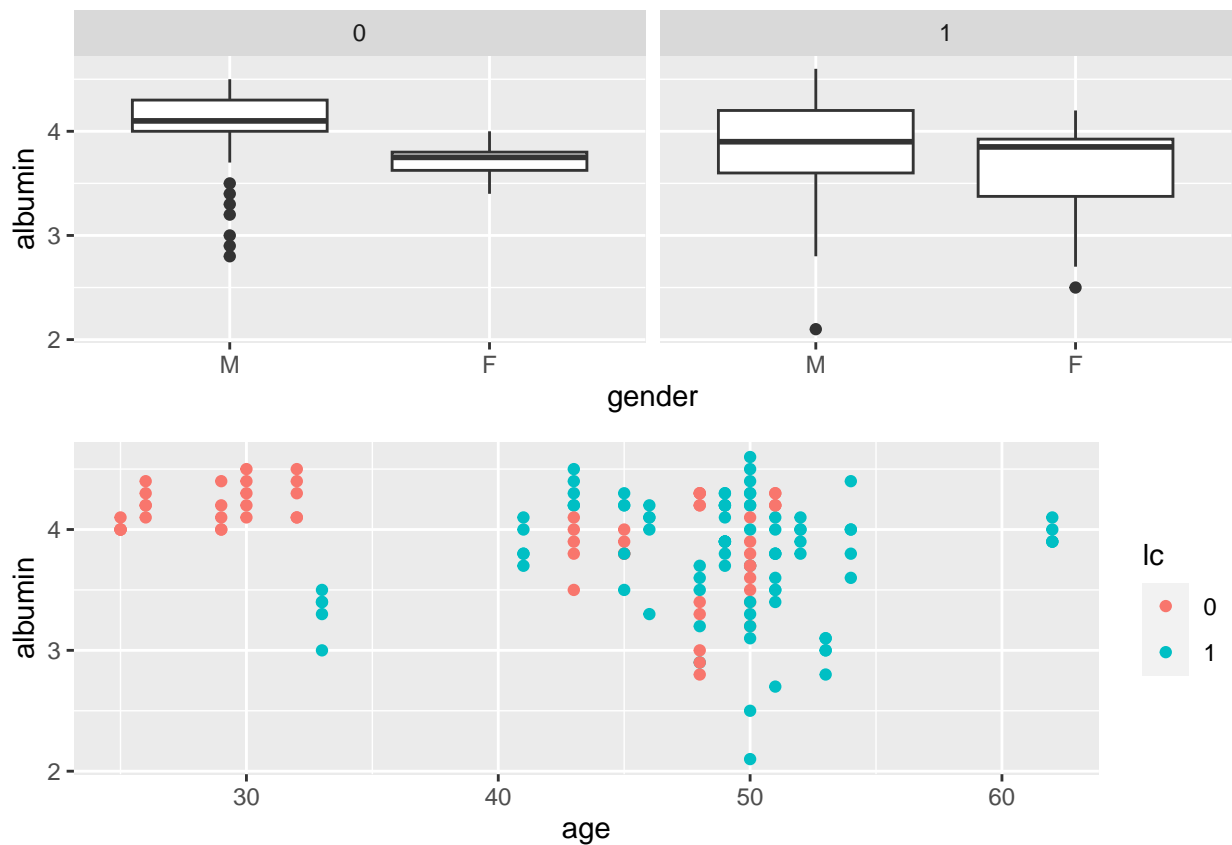
theme_bw(): 그래프의 배경색을 하얗게
theme_dark(): 그래프의 배경색을 어둡게
theme_minimal(): 그래프의 표시가 최소화
theme_classic(): x,y축 축이 진하게 표시됨

3.6 한 화면에 그래프 여러 개 그리기

(1) 그래프 여러 개 그리기

```
alb1<-ggplot(albu, aes(x=gender, y=albumin))+  
  geom_boxplot()+  
  facet_grid(~lc)  
  
alb2<-ggplot(albu, aes(x=age, y=albumin, color=lc))+  
  geom_point()
```

```
library(gridExtra)  
grid.arrange(alb1, alb2)
```



3.7 ggplot2 클릭만으로 하기

필요한 패키지들

```
library(esquisse)  
library(officer)  
library(rvg)
```

클릭만 하면 ggplot2의 기능들을 실행할 수 있는 패키지임

3.8 출판을 위한 출력

(1) pdf로 저장

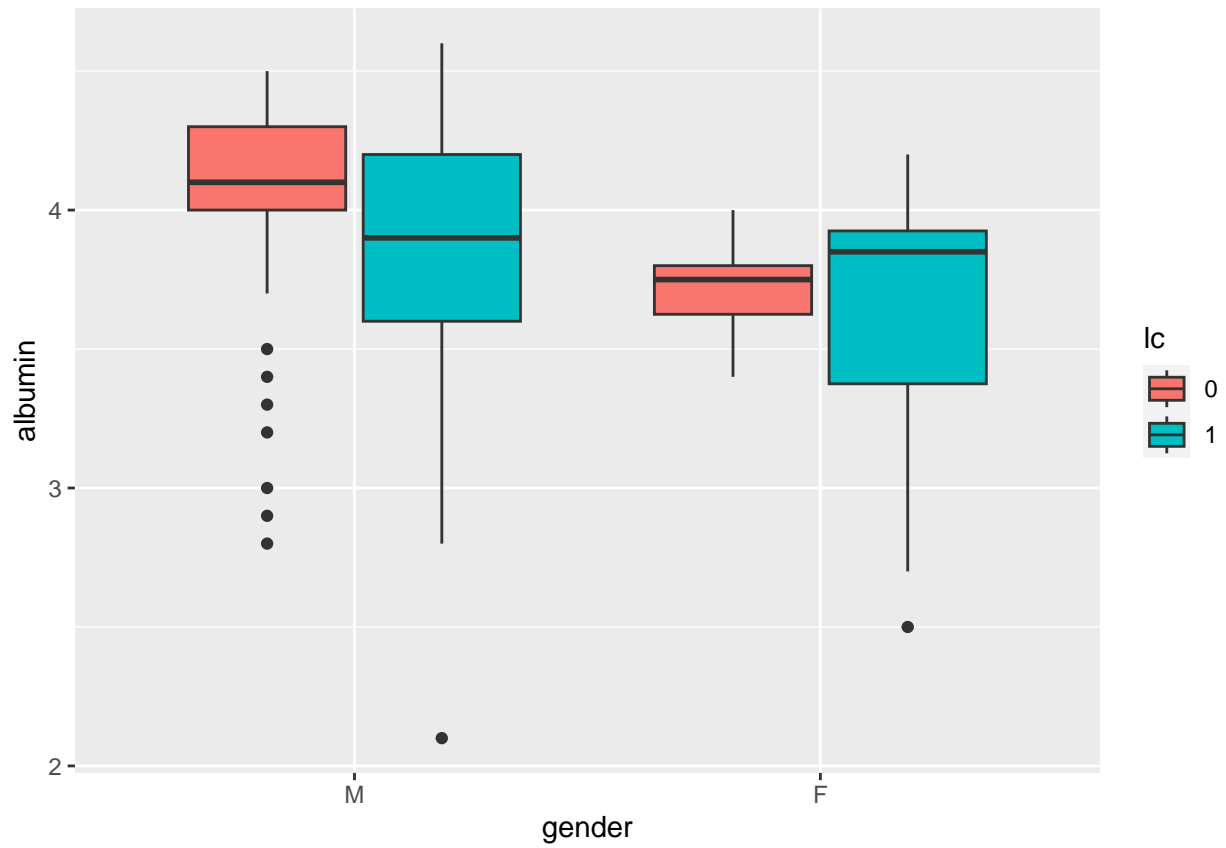
```
pdf('plot1.pdf',width=8, height=8)
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+
  geom_boxplot()
```

(2) tiff로 저장

```
tiff('plot1.tiff',width=1200, height=1800, res=300)
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+
  geom_boxplot()
```

(3) ggsave 이용

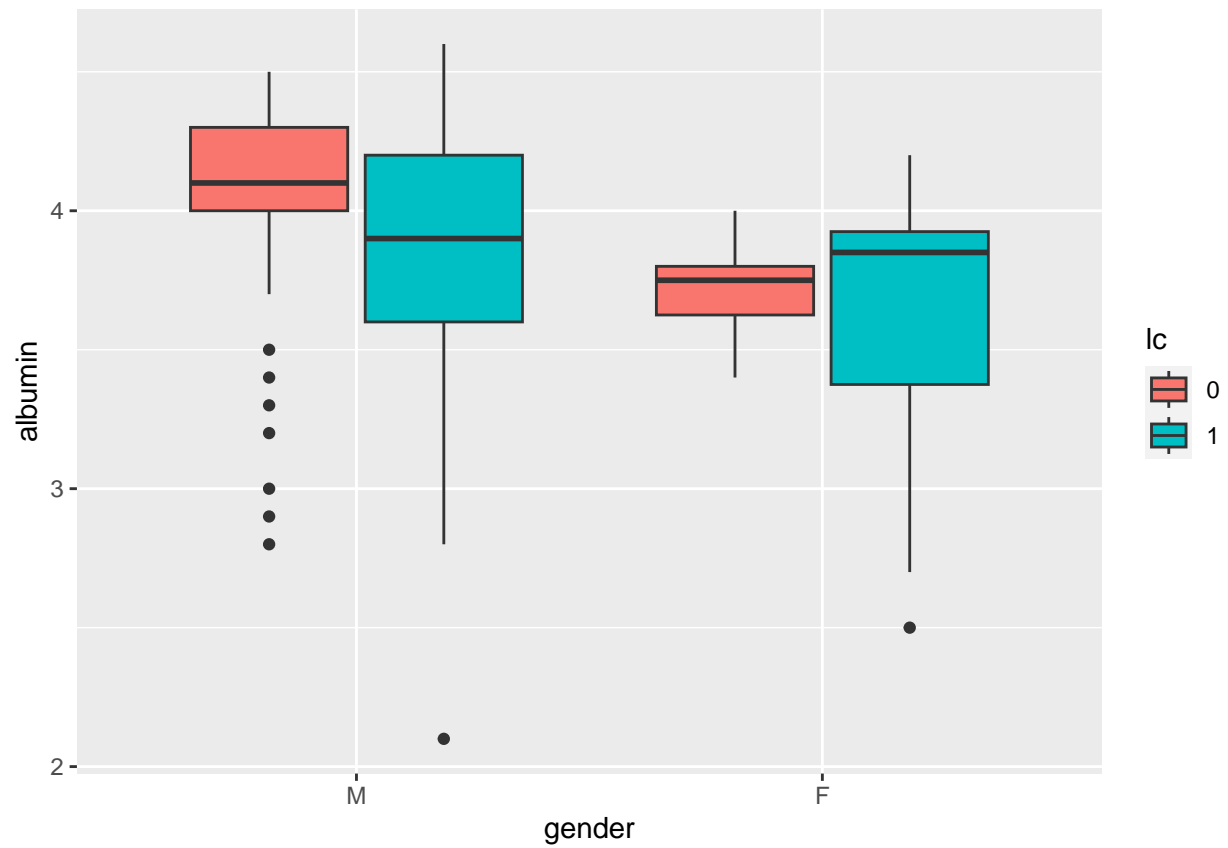
```
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+  
  geom_boxplot()
```



```
ggsave('albumin_graph.pdf', width = 10, height=12)
```



```
ggplot(albu, aes(x=gender, y=albumin, fill=lc))+
  geom_boxplot()
```



```
ggsave('albumin_graph.tiff', width = 6, height=8, dpi=300)
```

4. ggpubr 패키지

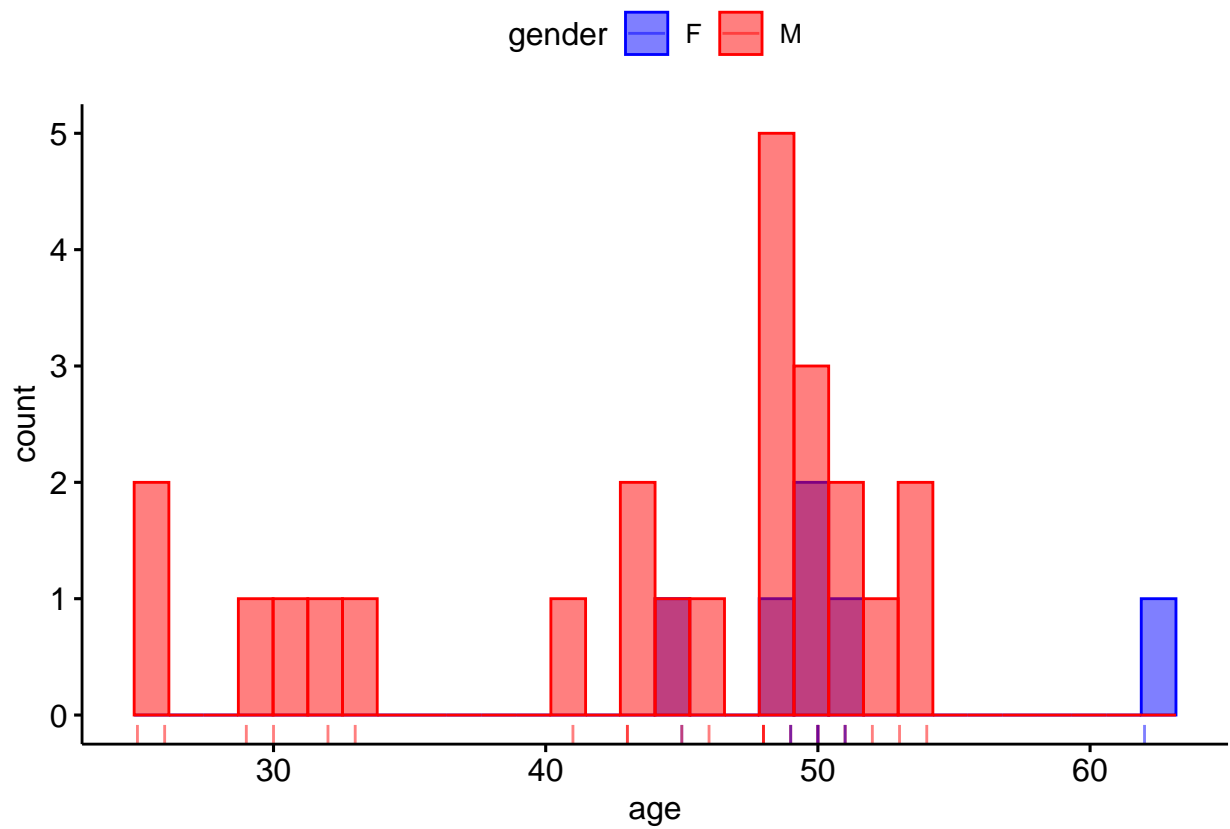
```
library(ggpubr)
```

논문 출판을 위한 그래프를 만들어주는 패키지

4.1 히스토그램

(1) 히스토그램

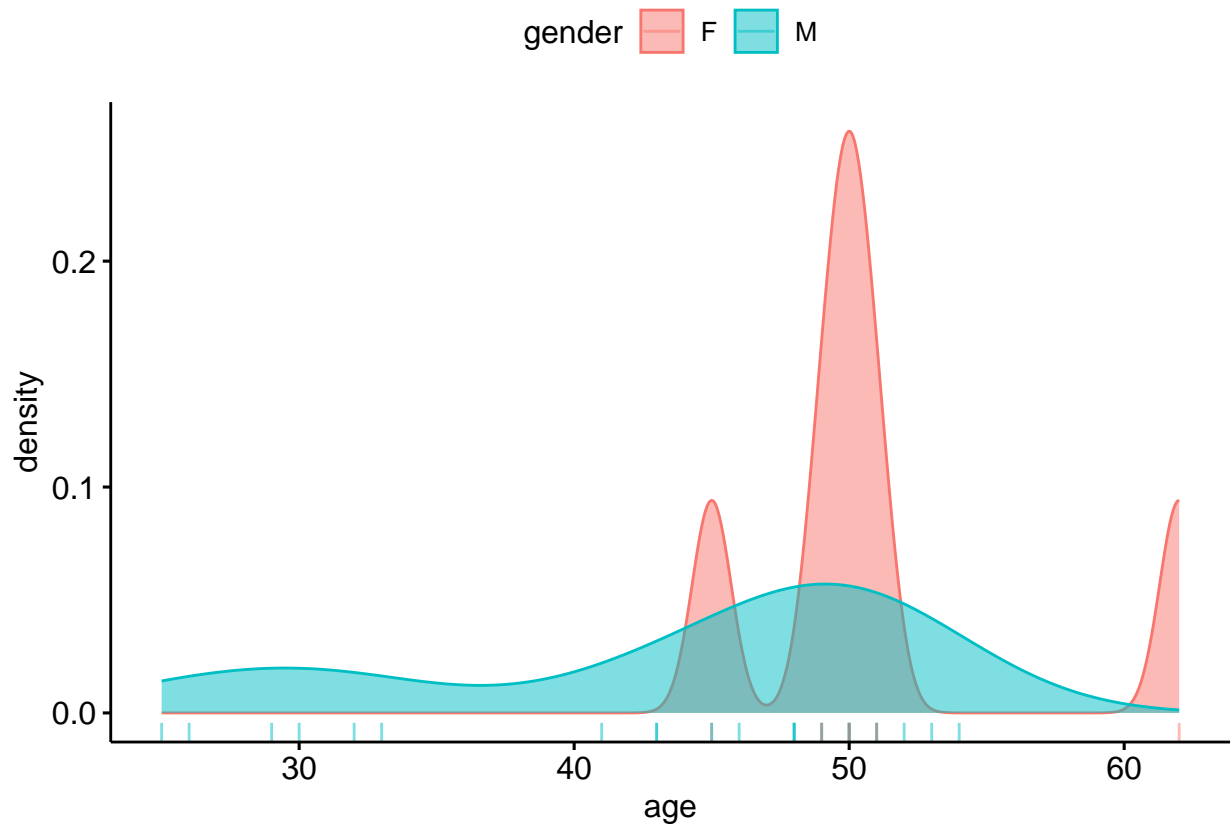
```
gghistogram(dat1,  
             x='age',  
             color='gender', fill='gender',  
             palette=c('blue','red'),  
             rug=TRUE)
```



4.2 밀도 그래프

(1) 밀도 곡선

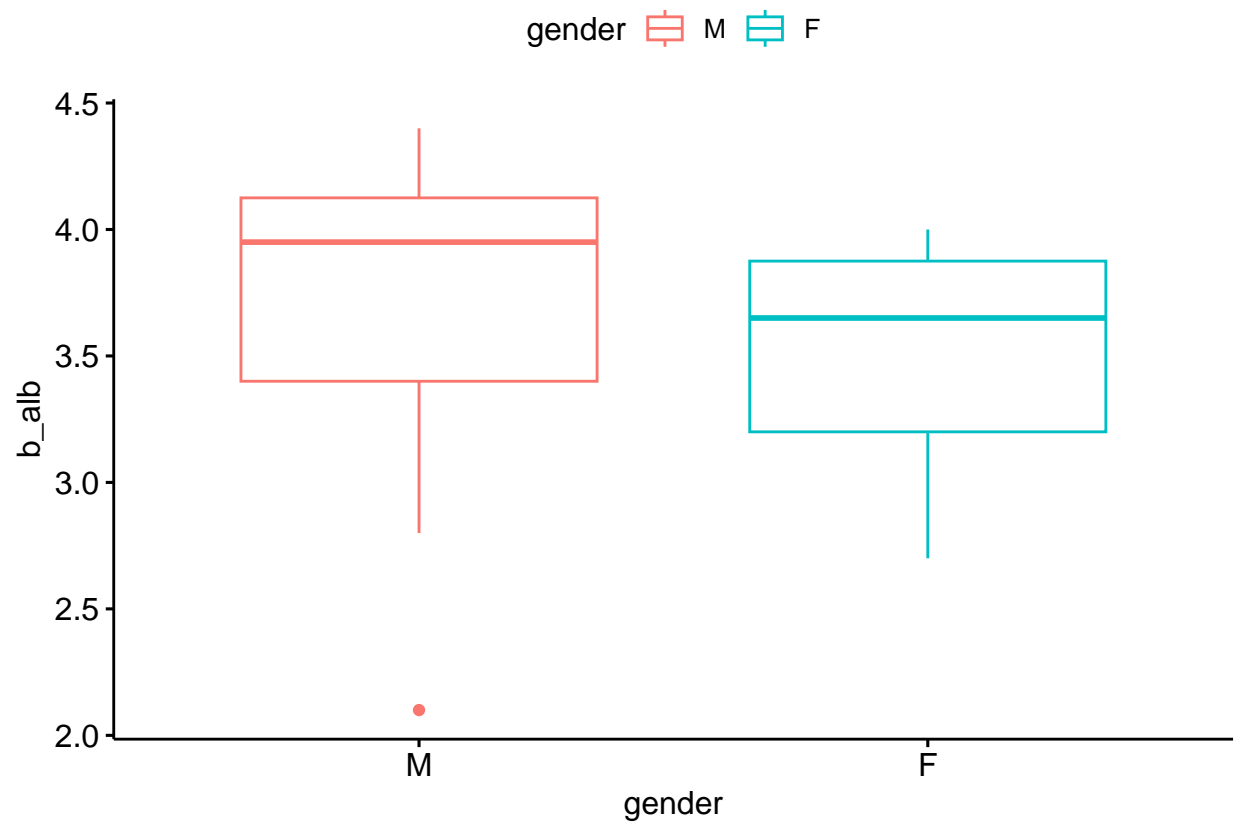
```
ggdensity(dat1,  
  x='age',  
  color='gender', fill='gender',  
  rug=TRUE)
```



4.3 박스 그래프

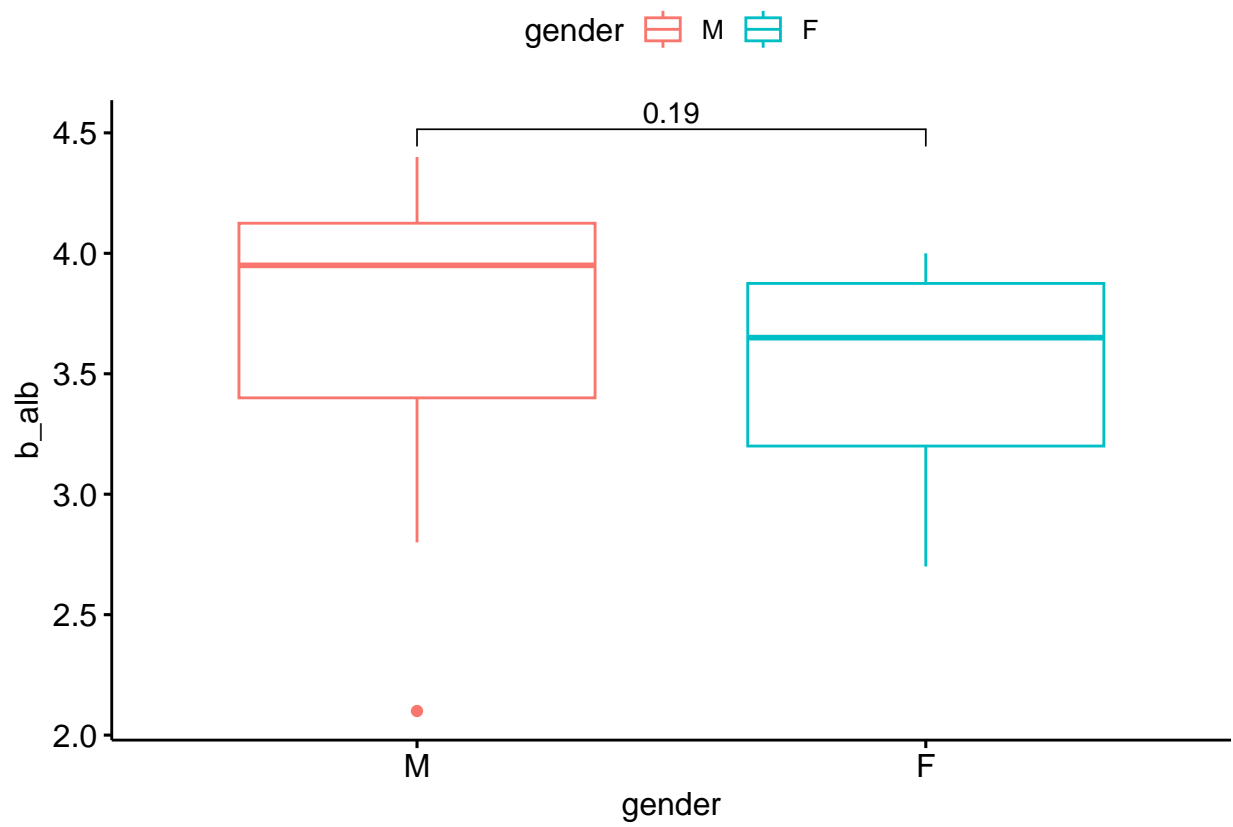
(1) 박스 그래프

```
ggboxplot(dat1,  
          x='gender', y='b_alb',  
          color='gender')
```



(2) 평균비교를 통한 p값 표시하기

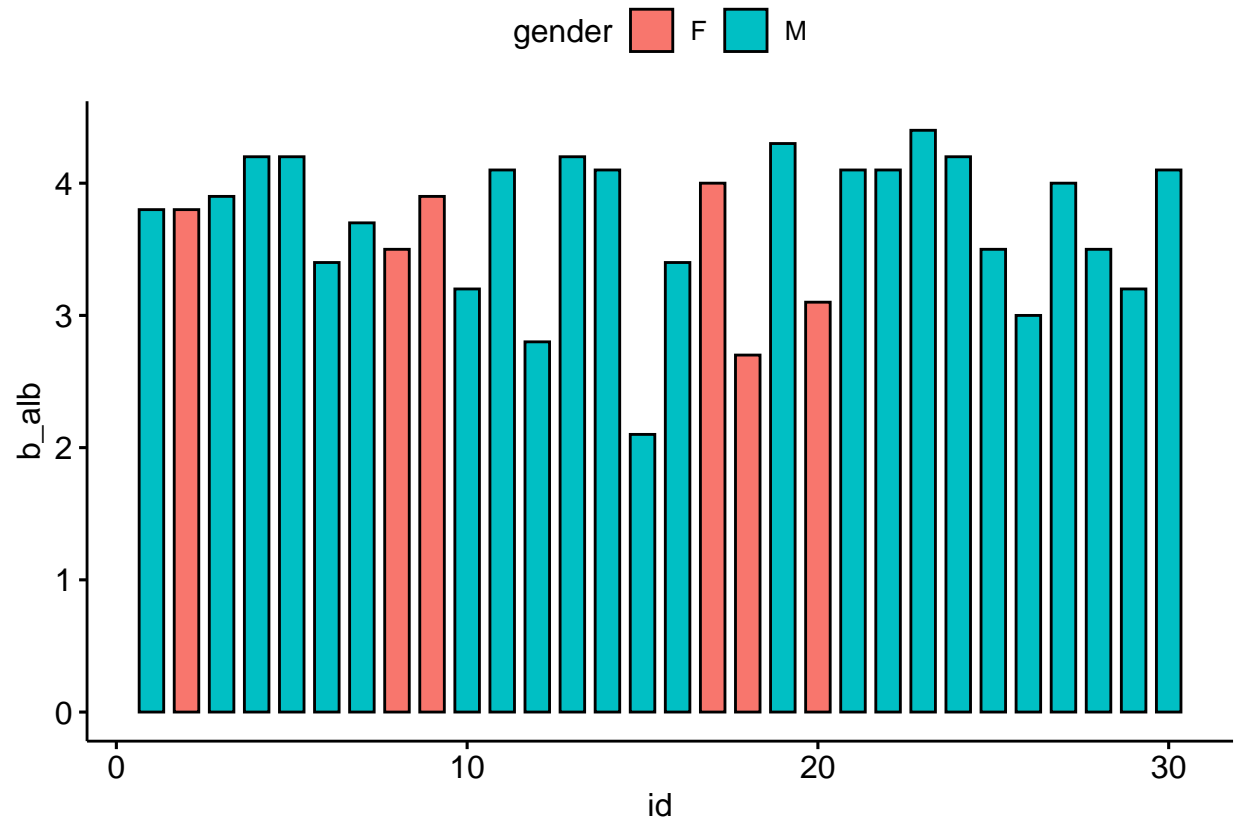
```
ggboxplot(dat1,  
  x='gender', y='b_alb',  
  color='gender')+  
  stat_compare_means(comparisons= list(c('M','F')))
```



4.4 막대 그래프

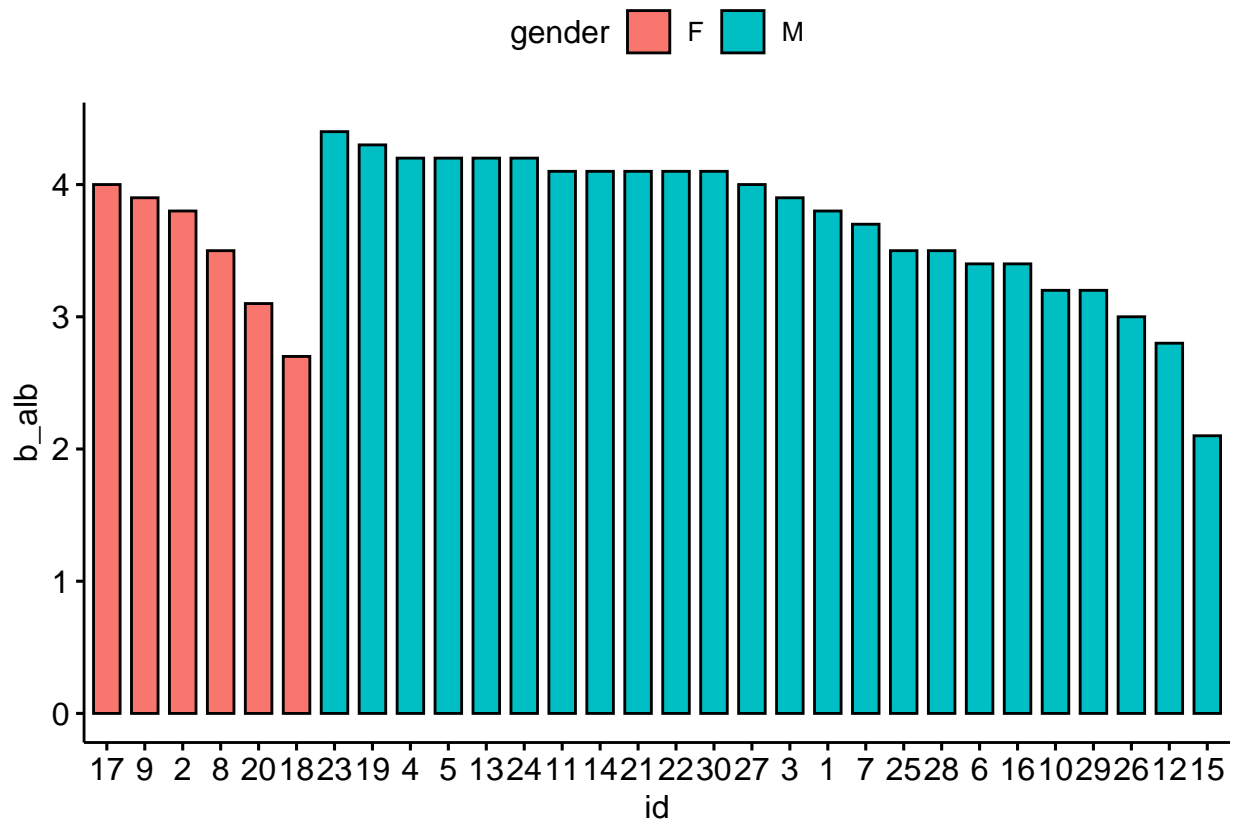
(1) 막대 그래프

```
ggbarplot(dat1,  
          x='id', y='b_alb',  
          fill='gender')
```



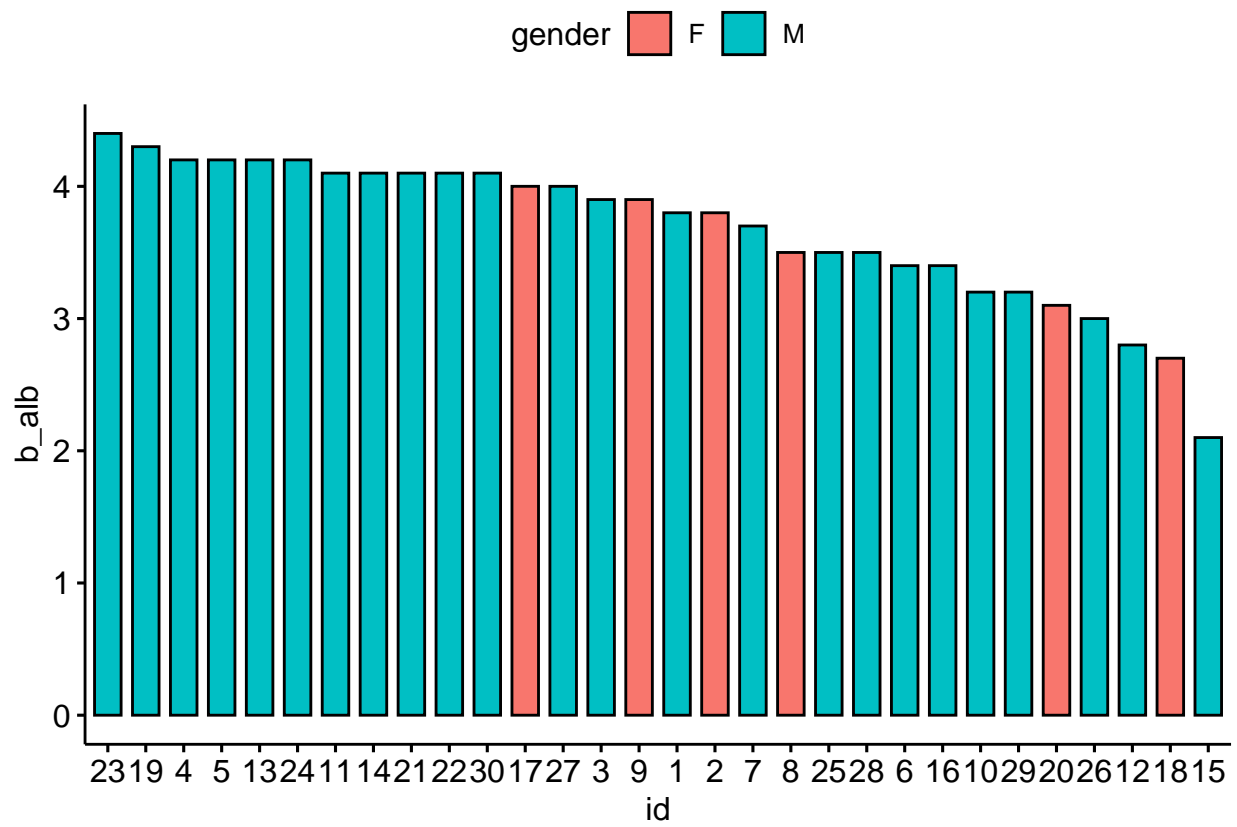
내림차순으로 막대 배열

```
ggbarplot(dat1,  
  x='id', y='b_alb',  
  fill='gender',  
  sort.val='desc')
```



성별 구분 없애기

```
ggbarplot(dat1,  
  x='id', y='b_alb',  
  fill='gender',  
  sort.val='desc',  
  sort.by.group=FALSE)
```



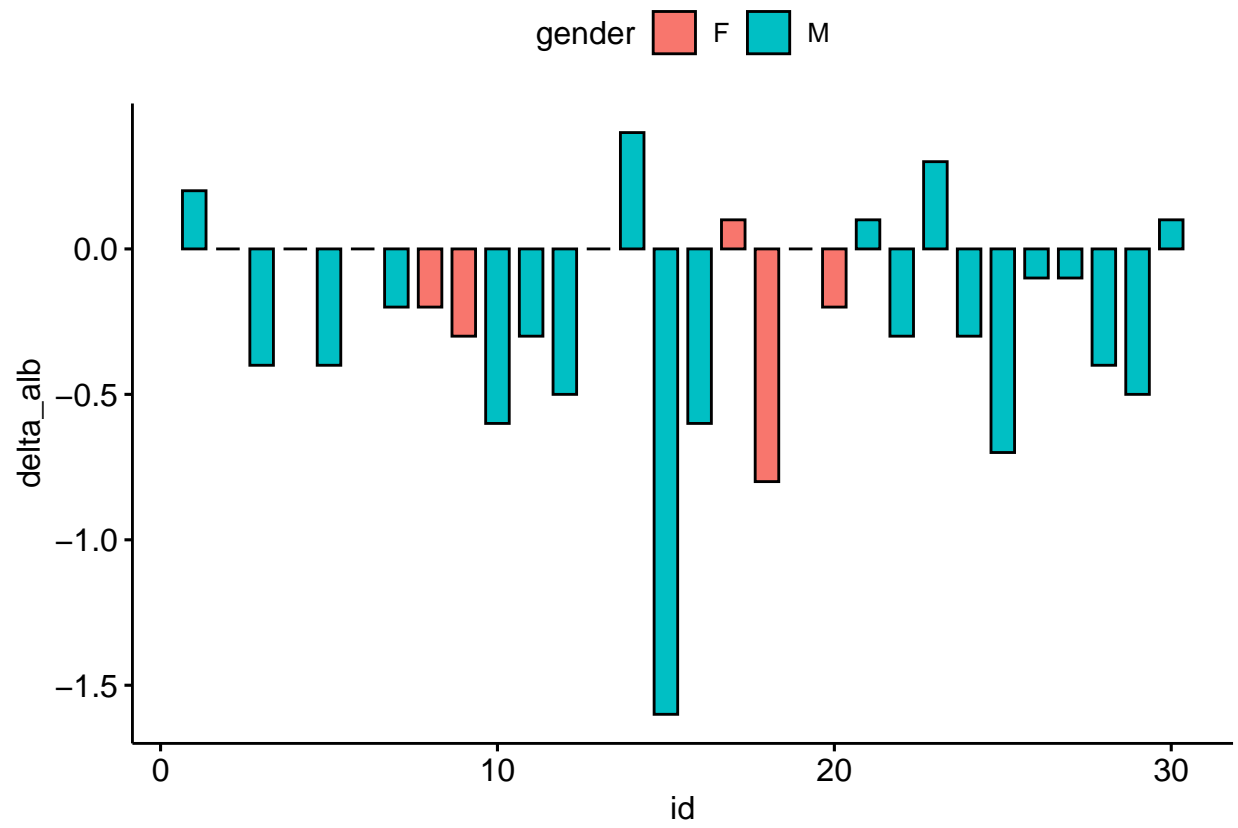
4.5 워터폴 (waterfall) 그래프

증양의 크기 변화나 혈청학적 마커의 증감을 보여주는데 탁월한 그래프

```
water.dt<-dat1 %>%  
  select(id,gender,b_alb,m6_alb) %>%  
  mutate(delta_alb=b_alb-m6_alb)  
water.dt
```

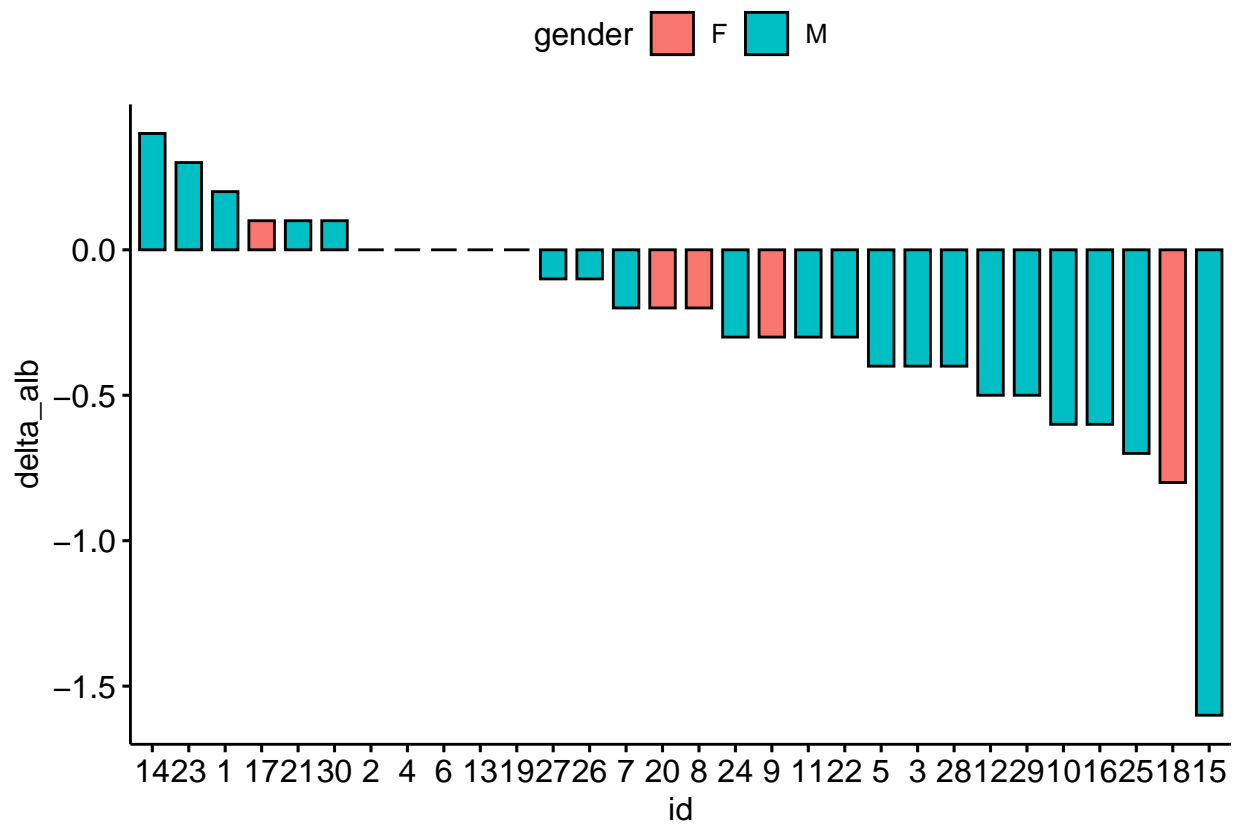
##	id	gender	b_alb	m6_alb	delta_alb
## 1	1	M	3.8	3.6	0.2
## 2	2	F	3.8	3.8	0.0
## 3	3	M	3.9	4.3	-0.4
## 4	4	M	4.2	4.2	0.0
## 5	5	M	4.2	4.6	-0.4
## 6	6	M	3.4	3.4	0.0
## 7	7	M	3.7	3.9	-0.2
## 8	8	F	3.5	3.7	-0.2
## 9	9	F	3.9	4.2	-0.3
## 10	10	M	3.2	3.8	-0.6
## 11	11	M	4.1	4.4	-0.3
## 12	12	M	2.8	3.3	-0.5
## 13	13	M	4.2	4.2	0.0
## 14	14	M	4.1	3.7	0.4
## 15	15	M	2.1	3.7	-1.6
## 16	16	M	3.4	4.0	-0.6
## 17	17	F	4.0	3.9	0.1
## 18	18	F	2.7	3.5	-0.8
## 19	19	M	4.3	4.3	0.0
## 20	20	F	3.1	3.3	-0.2
## 21	21	M	4.1	4.0	0.1
## 22	22	M	4.1	4.4	-0.3
## 23	23	M	4.4	4.1	0.3
## 24	24	M	4.2	4.5	-0.3
## 25	25	M	3.5	4.2	-0.7
## 26	26	M	3.0	3.1	-0.1
## 27	27	M	4.0	4.1	-0.1
## 28	28	M	3.5	3.9	-0.4
## 29	29	M	3.2	3.7	-0.5
## 30	30	M	4.1	4.0	0.1

```
ggbarplot(water.dt,  
          x='id', y='delta_alb',  
          fill='gender')
```



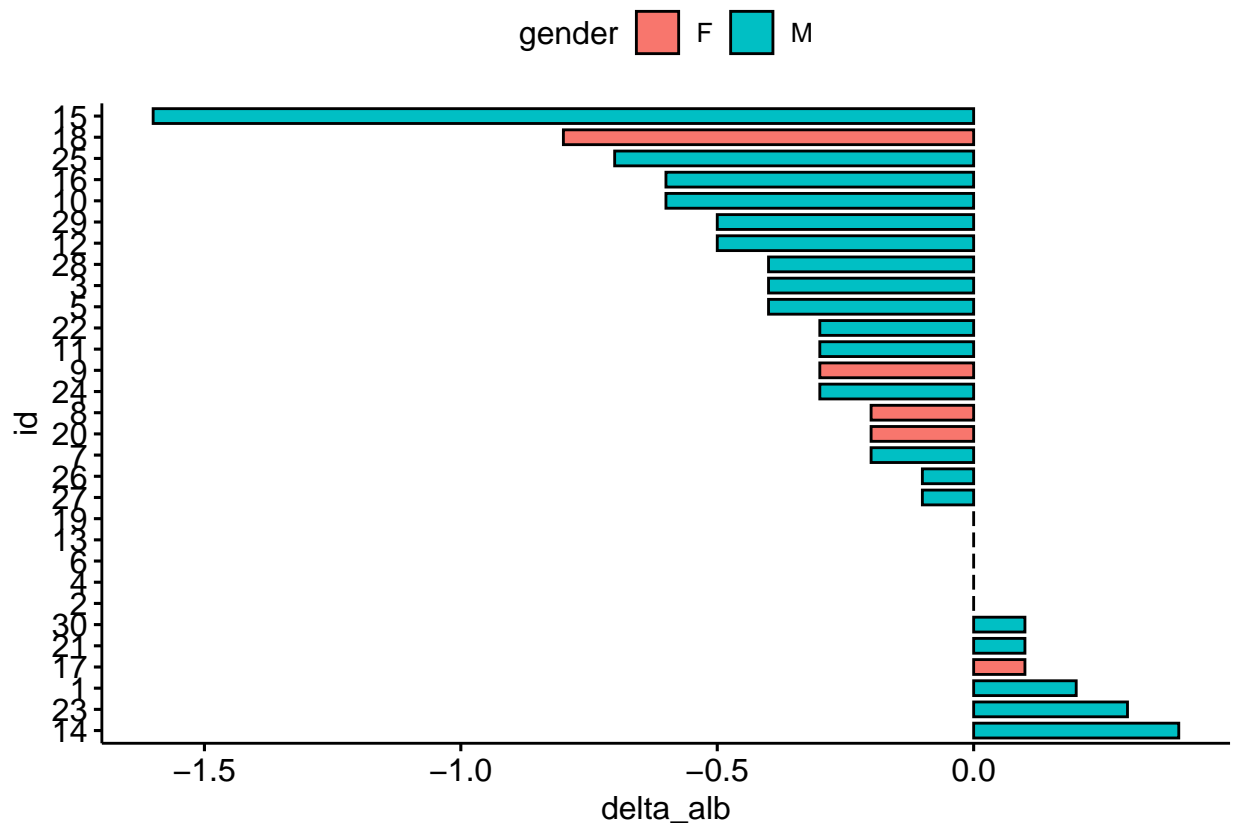
내림차순으로 막대 배열

```
ggbarplot(water.dt,  
          x='id', y='delta_alb',  
          fill='gender',  
          sort.val='desc',  
          sort.by.groups = FALSE)
```



가로로 나타내기

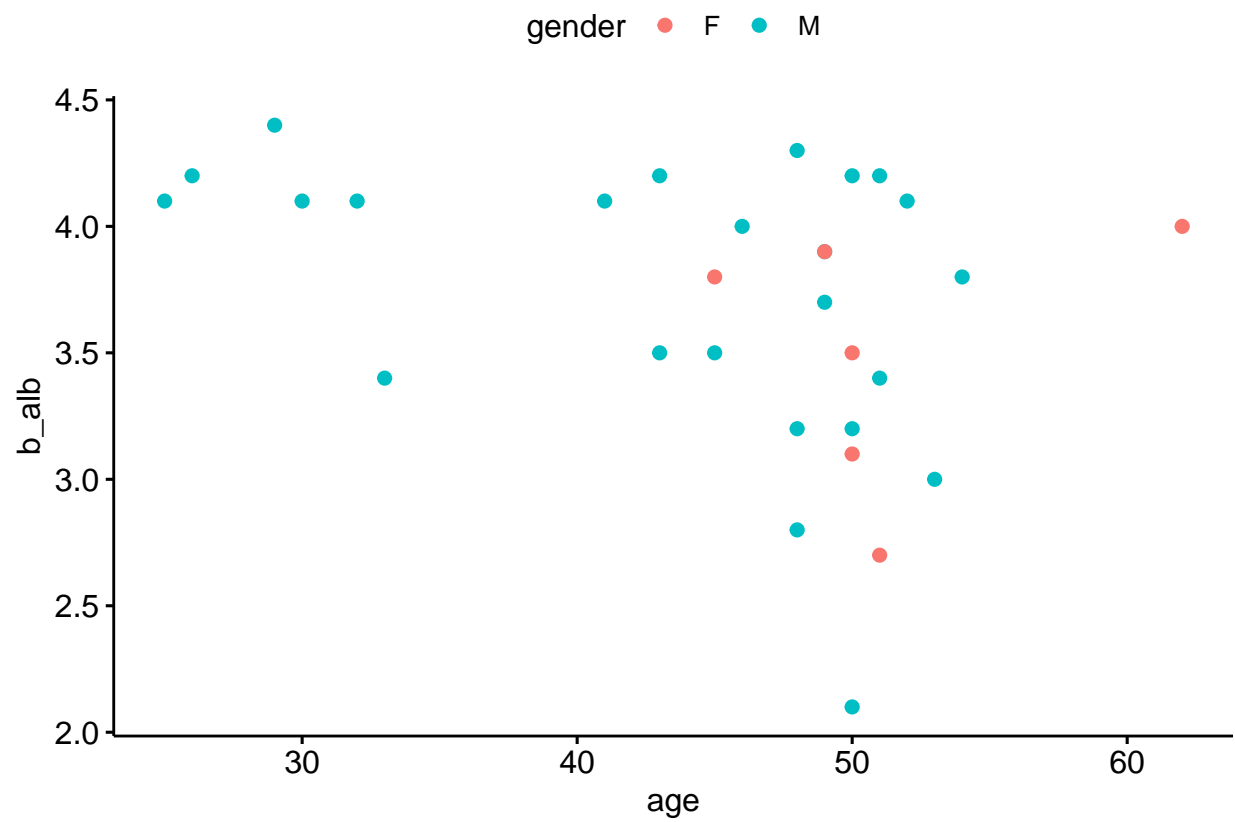
```
ggbarplot(water.dt,  
          x='id', y='delta_alb',  
          fill='gender',  
          sort.val='desc',  
          sort.by.groups = FALSE,  
          rotate=TRUE)
```



4.6 산점도

(1) 산점도

```
ggscatter(dat1,  
          x='age',  
          y='b_alb',  
          color='gender')
```

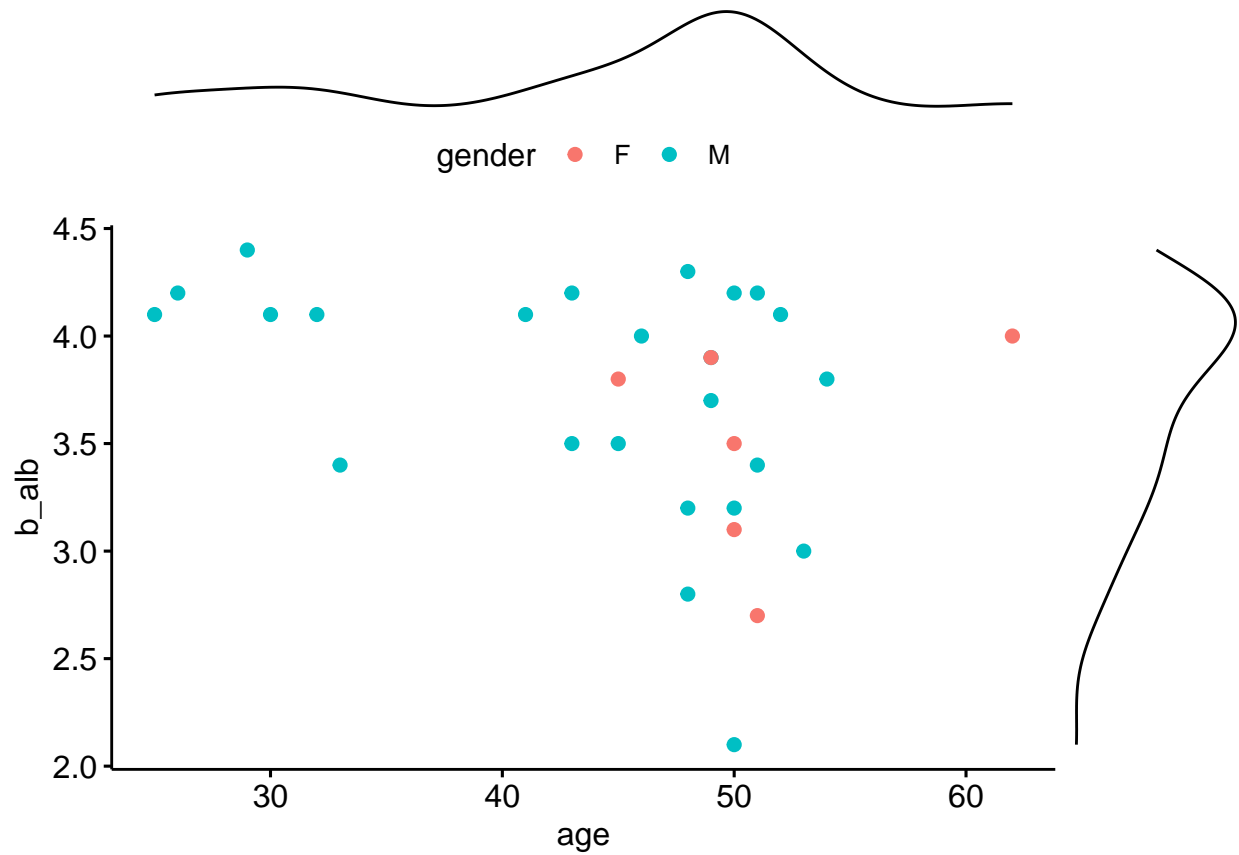


(2) 산점도에 추가 그래프 넣기

- 밀도

```
library(ggExtra)

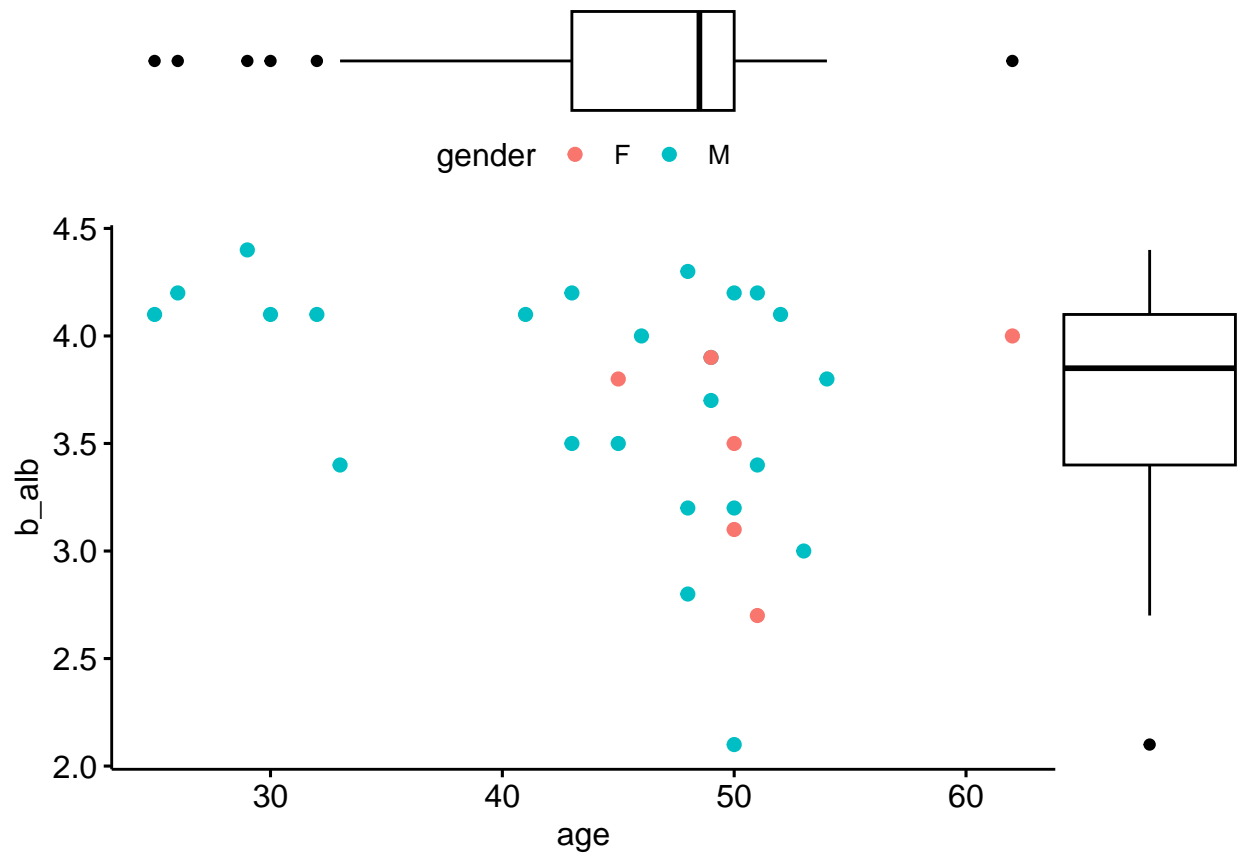
ggscatter(dat1, x='age', y='b_alb',
          color='gender') %>%
  ggMarginal(type='density')
```



- 박스

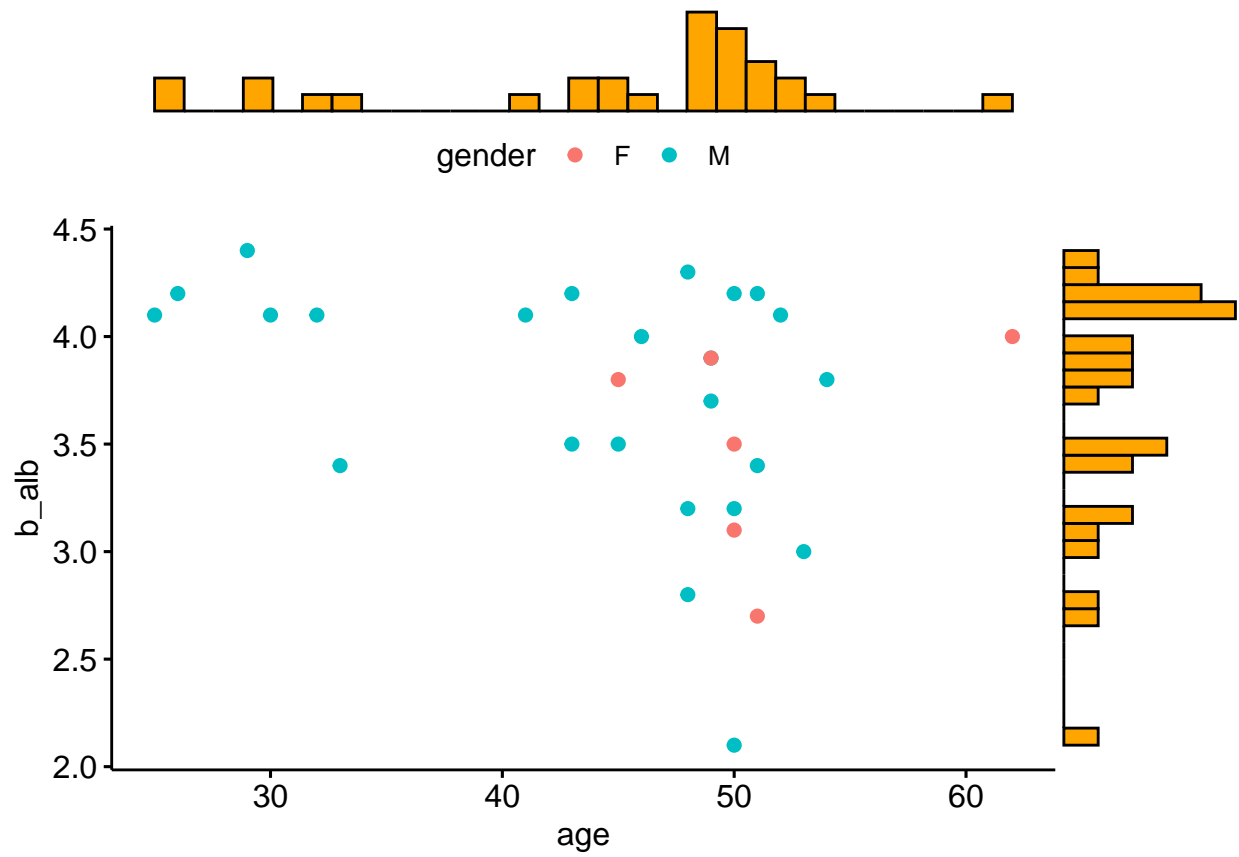
```
ggscatter(dati, x='age', y='b_alb',
          color='gender') %>%
  ggMarginal(type='boxplot')
```

```
## Warning: Continuous x aesthetic
## i did you forget `aes(group = ...)`?
## Continuous x aesthetic
## i did you forget `aes(group = ...)`?
```



- 히스토그램

```
ggscatter(dat1, x='age', y='b_alb',
          color='gender') %>%
  ggMarginal(type='histogram', fill='orange')
```

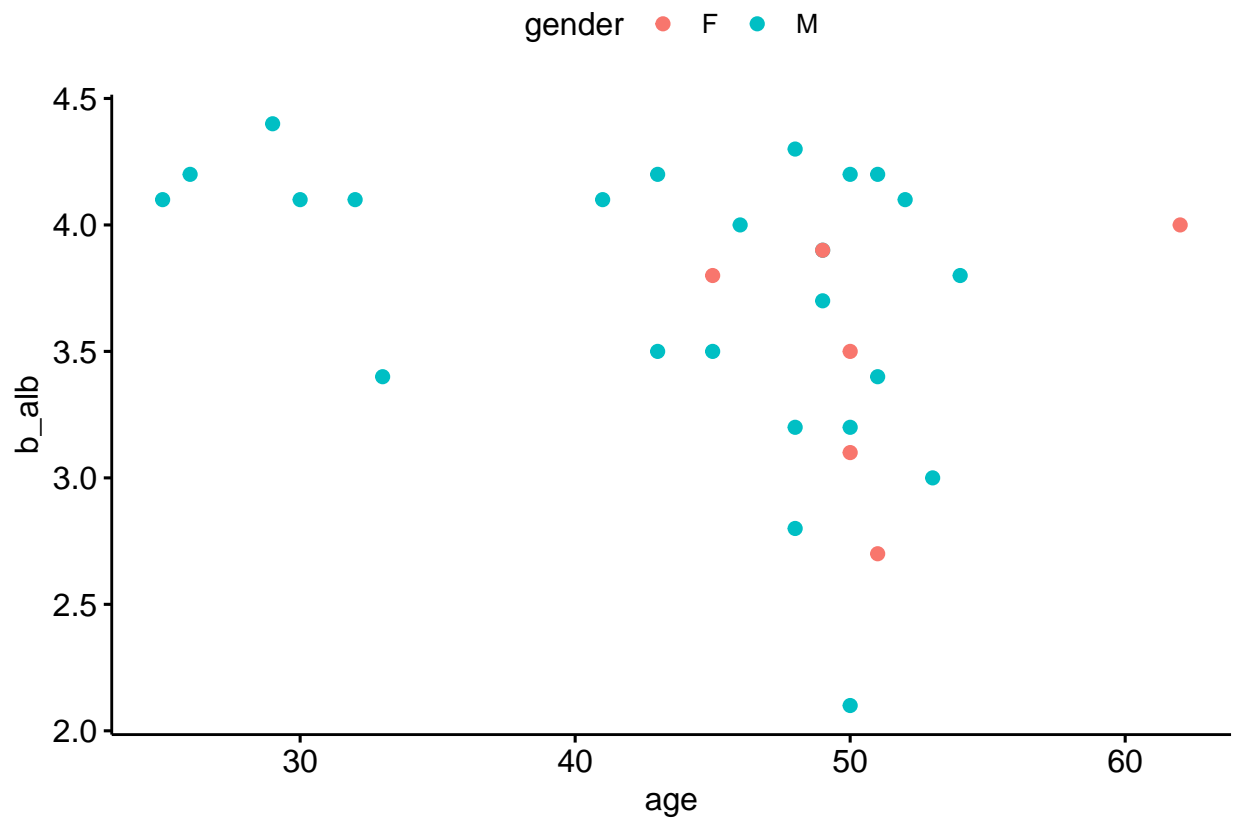


5. 상관관계를 그려주는 패키지

5.1 GGally 패키지

(1) 변수들 간의 상관관계 한눈에 보기

```
ggscatter(dat1,  
          x='age',  
          y='b_alb',  
          color='gender')
```

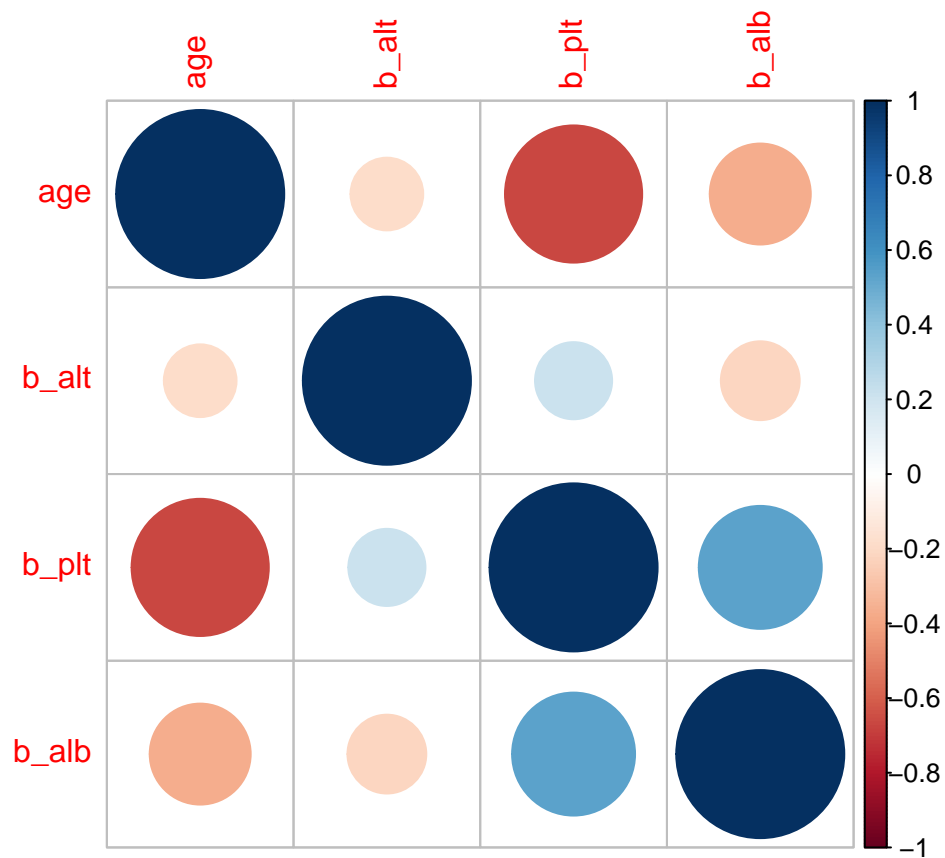


5.2 corrplot 패키지

(1) 한눈에 변수들의 상관관계 보기

```
library(corrplot)

cor.dt<-dat1 %>%
  select(age, b_alt, b_plt, b_alb) %>%
  na.omit()
corrplot(cor(cor.dt))
```



(2) 다양한 method 옵션 변화 주기

```
par(mfrow=c(2,2))
corrplot(cor(cor.dt), method='square')
corrplot(cor(cor.dt), method='ellipse')
corrplot(cor(cor.dt), method='number')
corrplot(cor(cor.dt), method='color')
```

