

빅콘 EDA part1

Kim Na Yoon 1402013

8/12/2020

#맥북으로 한글폰트 깨짐현상 해결

```
library(extrafont)

## Registering fonts with R

library(ggplot2)
ggplot2::theme()

## Named list()
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi FALSE
## - attr(*, "validate")= logi TRUE

theme_set(theme_grey(base_family='NanumGothic'))
```

김수현

```
rm(list=ls())
shop_data<-read.csv("/Users/gimnayun/Desktop/final_data.csv")
```

#자료:37372 개

```
library(ggplot2)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

#shop_data<-shop_data%>%filter(total_sales!=0)
#dim(shop_data)
#total_sales 가 0 인걸 제외하니 자료 35379 개
#####
```

#total_sales 총 판매액이 0 인 데이터는 아예 예측시 제외해야 하는지 메일로 물어봄

#기온

#알아보기 힘드니깐... #주차별해가지고 계절별로 나눠 보기

#계절별 판매량을 본다

```
max(shop_data$temp)
```

```
## [1] 36.7
```

```
max(shop_data$total_sales)
```

```
## [1] 322009000
```

```
min(shop_data$total_sales)
```

```
## [1] 0
```

```
min(shop_data$temp)
```

```
## [1] -10.8
```

```
max(shop_data$week)
```

```
## [1] 53
```

#53 주까지 있음 #봄 3₅ 월/1022 주차 #여름 6₈ 월/2335 주차 #가을 9₁₁ 월/3648 주차 #겨울 12₂ 월/19 주차+49~53 주차

```
shop_data$season<-ifelse(shop_data$week>=10&shop_data$week<=22,"spring",ifelse(shop_data$week>=23&shop_data$week<=35,"summer",ifelse(shop_data$week>=36&shop_data$week<=48,"fall","winter")))
```

```
a<-shop_data%>%filter(season=="spring")%>%summarise(mean(Sales_quantity))
```

```
b<-shop_data%>%filter(season=="summer")%>%summarise(mean(Sales_quantity))
```

```
c<-shop_data%>%filter(season=="fall")%>%summarise(mean(Sales_quantity))
```

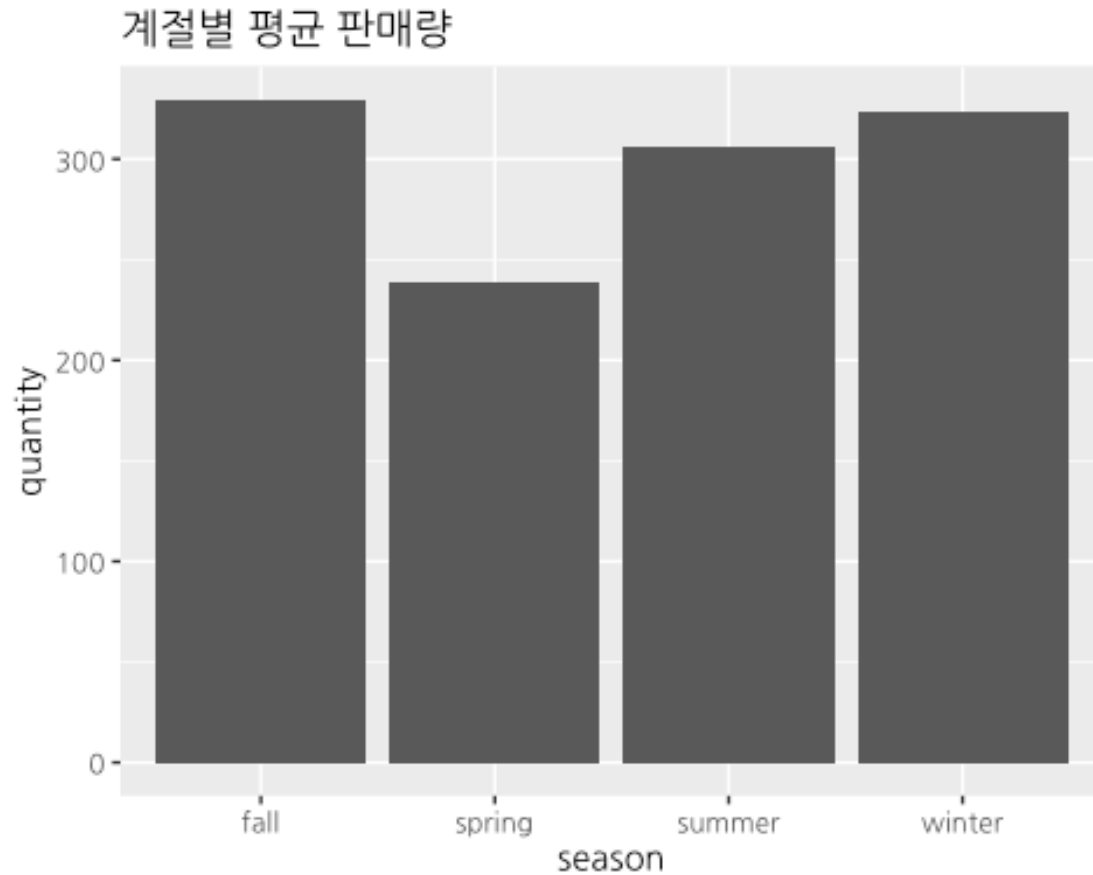
```
d<-shop_data%>%filter(season=="winter")%>%summarise(mean(Sales_quantity))
```

```
season<-c("spring","summer","fall","winter")
```

```
quantity<-c(a,b,c,d)
```

```
season_quantity<-data.frame(season,quantity)
```

```
ggplot(data=season_quantity)+geom_col(aes(x=season,y=quantity))+ggtitle("계절별 평균 판매량")
```



#봄에

판매량이 살짝 줄어든거처럼 보인다

#강수량

```
min(shop_data$precipitation)
```

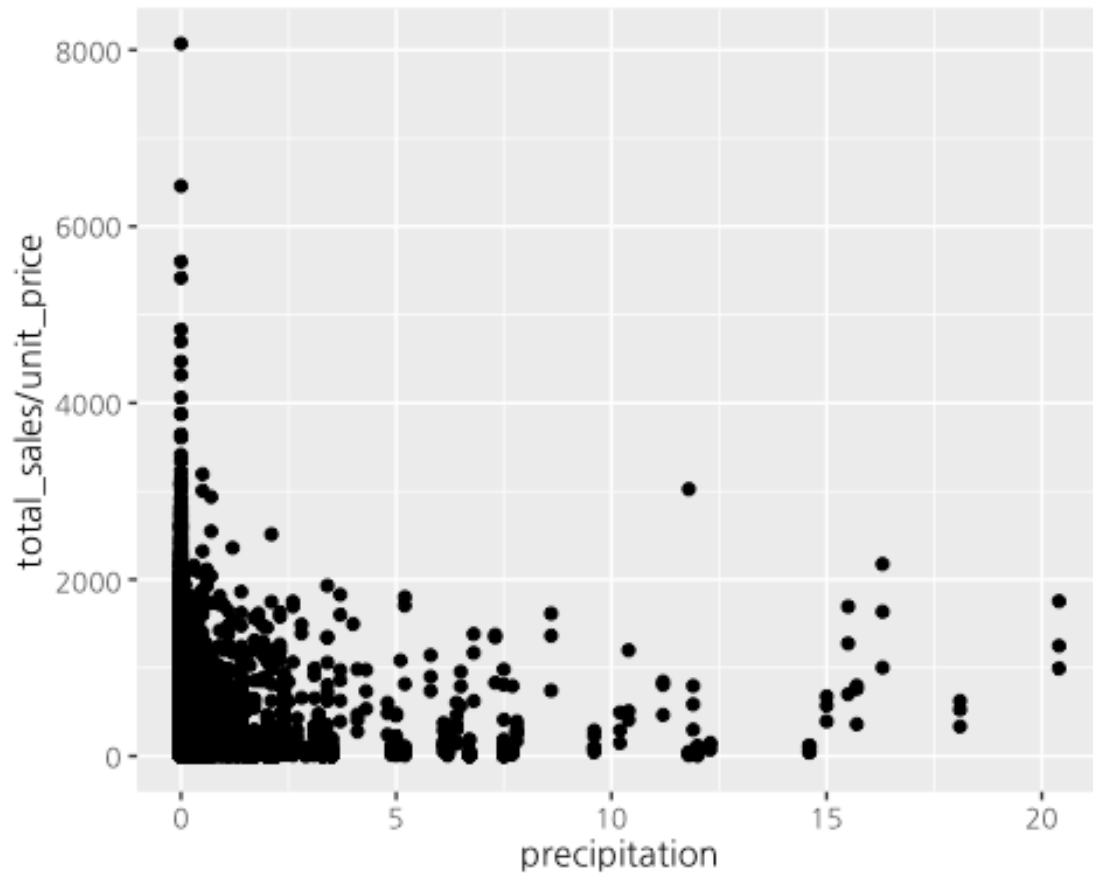
```
## [1] 0
```

```
max(shop_data$precipitation)
```

```
## [1] 20.4
```

일단 y 축 갯수로 해가지고 산점도

```
ggplot(data=shop_data)+geom_point(aes(x=precipitation,y=total_sales/unit_price))
```

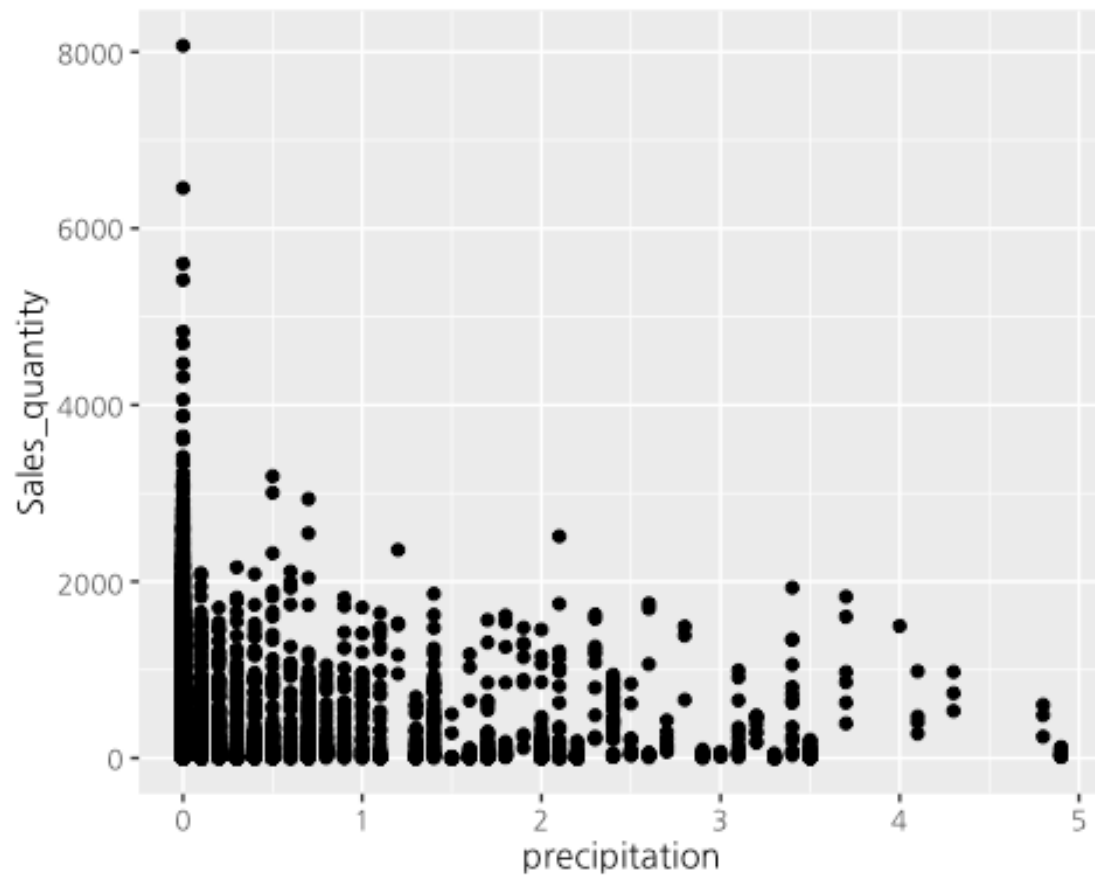


#0~5 사이에 자료들이 몰려 있으므로 #0~5 를 한그룹으로 보고 5~21 을 또다른 한 그룹으로 보아 서로의 평균 매출액 비교

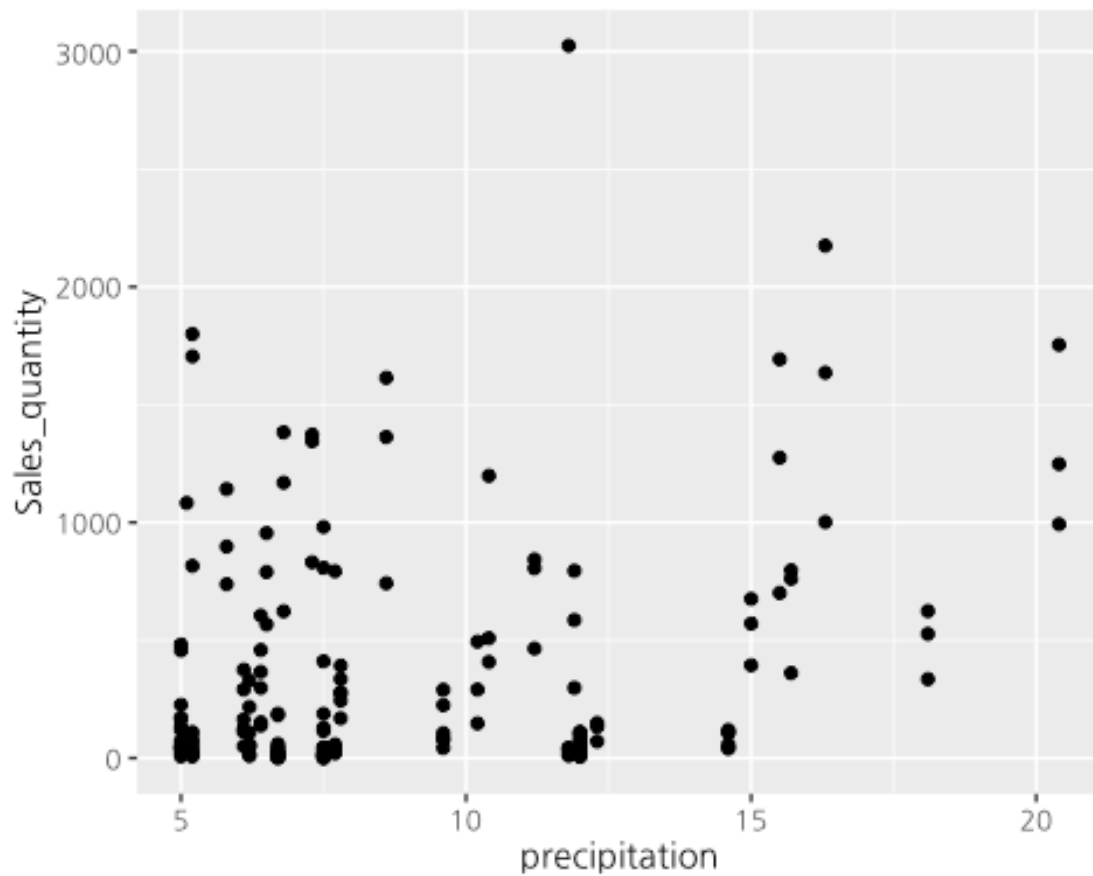
```
rain_small<-shop_data%>% filter(precipitation>=0&precipitation<5)
rain_big<-shop_data%>%filter(precipitation>=5)
```

#각각나눠거 그려보기...의미없는듯...

```
ggplot(data=rain_small)+geom_point(aes(x=precipitation,y=Sales_quantity))
```



```
ggplot(data=rain_big)+geom_point(aes(x=precipitation,y=Sales_quantity))
```



#small 그룹과 big 그룹 평균 비교

```
mean(rain_small$precipitation)
```

```
## [1] 0.05022588
```

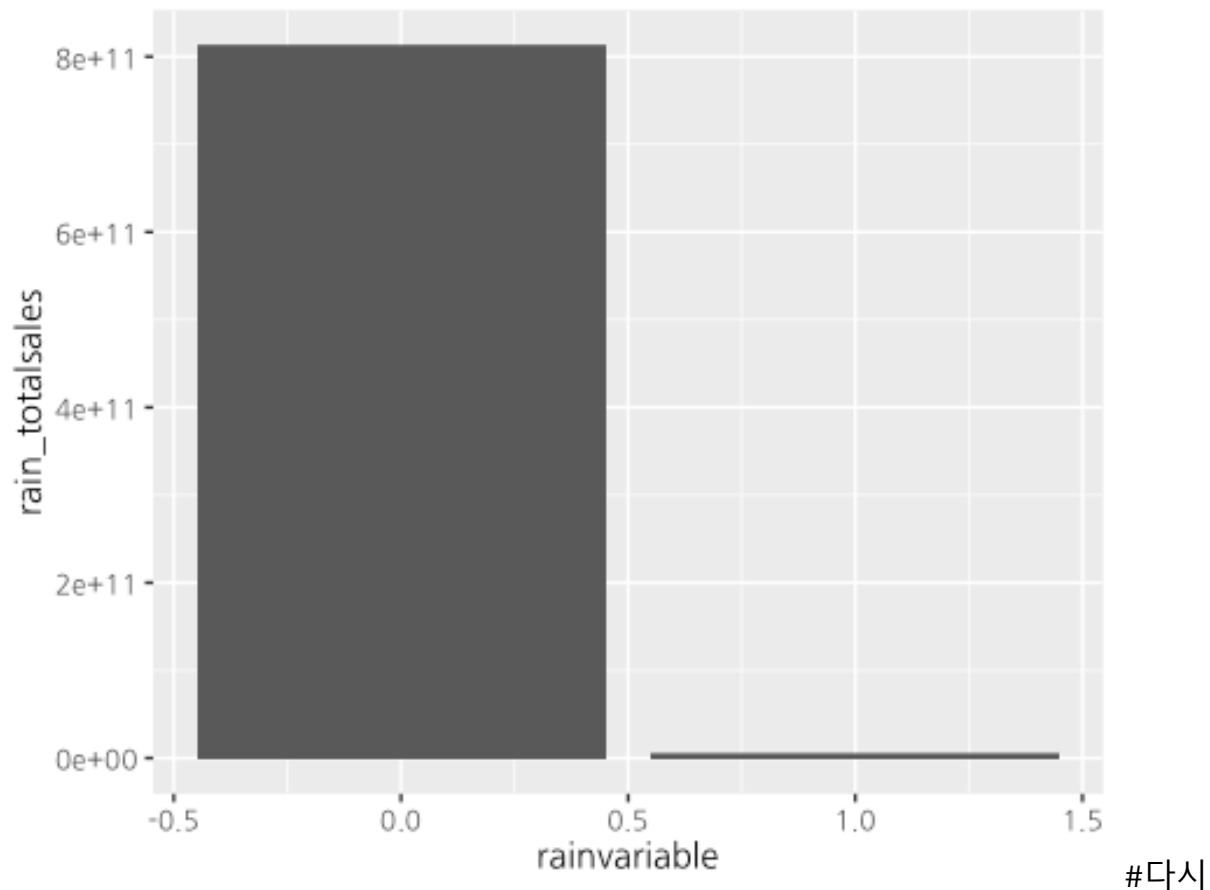
```
mean(rain_big$precipitation)
```

```
## [1] 8.966304
```

```
shop_data$rainvariable<-ifelse(shop_data$precipitation>=0& shop_data$precipitation<5,0,1)
```

```
shop_data$rain_totalsales<-ifelse(shop_data$rainvariable==0,mean(rain_small$total_sales),mean(rain_big$total_sales))
```

```
ggplot(data=shop_data)+geom_col(aes(x=rainvariable,y=rain_totalsales))
```



나눠봄 4 개그룹으로

```
shop_data$rainvariable<-ifelse(shop_data$precipitation>=0& shop_data$precipitation<3,1,ifelse(shop_data$precipitation>=3&shop_data$precipitation<5,2,ifelse(shop_data$precipitation>=5&shop_data$precipitation<10,3,4)))
mean1<-shop_data %>% filter(rainvariable==1) %>% summarise(mean(Sales_quantity))
mean1

##   mean(Sales_quantity)
## 1                297.3657

mean2<-shop_data %>% filter(rainvariable==2) %>% summarise(mean(Sales_quantity))
mean2

##   mean(Sales_quantity)
## 1                275.5667

mean3<-shop_data %>% filter(rainvariable==3) %>% summarise(mean(Sales_quantity))
mean3
```

```
##    mean(Sales_quantity)
## 1          280.8862

mean4<-shop_data %>% filter(rainvariable==4) %>% summarise(mean(Sales_quantit
y))
mean4

##    mean(Sales_quantity)
## 1          479.7869

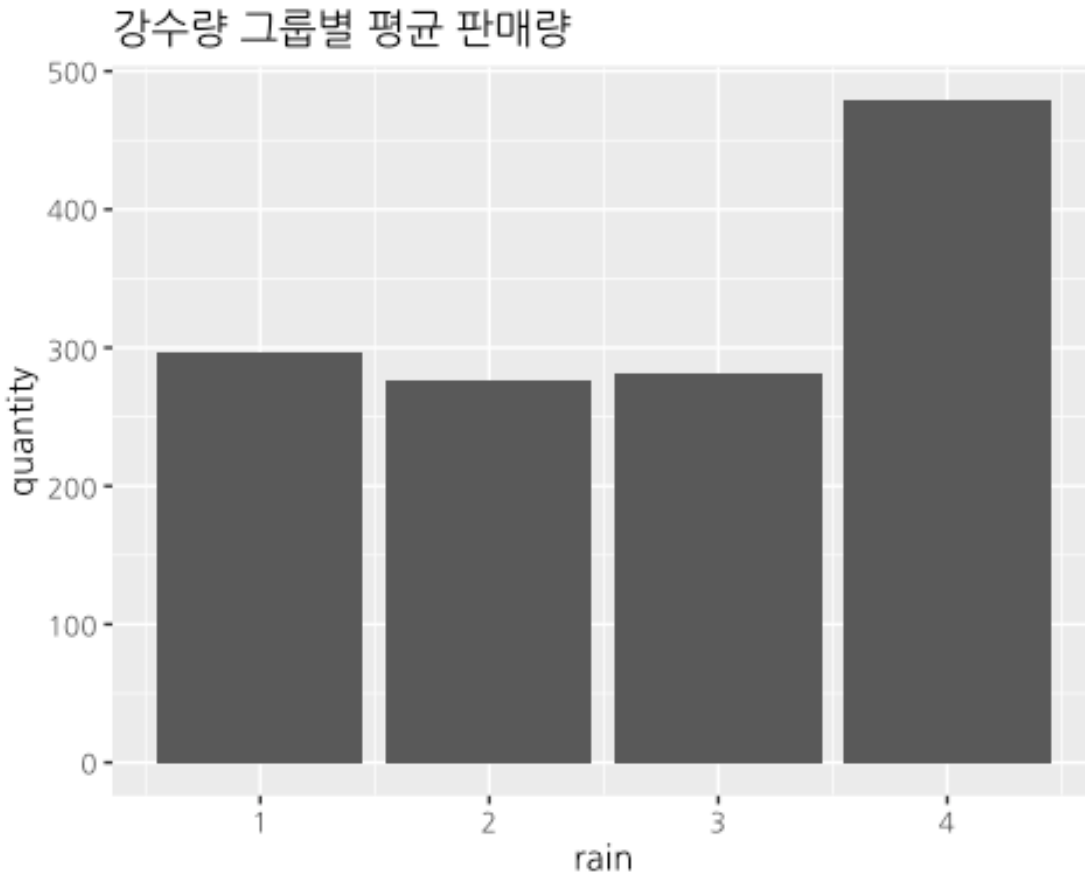
shop_data$rain_quantity<-ifelse(shop_data$rainvariable==1,mean1,ifelse(shop_d
ata$rainvariable==2,mean2,ifelse(shop_data$rainvariable==3,mean3,mean4)))
```

#ggplot(data=shop_data)+geom_col(aes(x=rainvariable,y=rain_quantity) #안그려져서
따로 데이터프레임만들

```
rainframe<-data.frame(rain=c(1,2,3,4),
                      quantity=c(297.3657,275.5667,280.8862,479.7869))
rainframe

##    rain quantity
## 1     1 297.3657
## 2     2 275.5667
## 3     3 280.8862
## 4     4 479.7869

ggplot(data=rainframe)+geom_col(aes(x=rain,y=quantity))+ggtitle("강수량 그룹별
평균 판매량")
```

#강수량이 높은 4 그룹에서 판매량이 높아지는 걸 확인할 수 있음

#최저기온,최고기온,평균기온은 하루매출총액으로 따져야 할듯함 #일별 총판매량
구하기/일별 평균 판매량

```
#head(shop_data$date,10)
#library(tidyr)
#tempdata<-separate(shop_data,date,into=c("year","mon","day","t"))
#head(tempdata)
#str(tempdata)
#tempdata$mon<-as.numeric(tempdata$mon)
#tempdata$day<-as.numeric(tempdata$day)
#str(tempdata)
#tempdata$day_quantity<-0

#for (i in c(1,3,5,7,8,10,12)){
#  for (j in 1:31){
#    if (tempdata$mon==i&tempdata$day==j){
#      tempdata[tempdata$mon==i&tempdata$day==j,"day_quantity"]<-mean(tempdata[tempdata$mon==i&tempdata$day==j,"Sales_quantity"])
#    }
#  }
#}
```

```

#}

#tempdata$day_quantity
#최저기온
#str(shop_data)
#ggplot(data=shop_data)+geom_point(aes(y=Low_temp,x=total_sales/unit_price))

#최고기온
#ggplot(data=shop_data)+geom_point(aes(x=high_temp,y=total_sales/unit_price))

#평균기온

```

#미세먼지 #미세없는날 0, 주의보 1, 경보 2 로 데이터 변경 #3 개의 그룹별 평균 판매량을 비교

```

str(shop_data)

## 'data.frame': 37372 obs. of 33 variables:
## $ date : Factor w/ 20588 levels "2019-01-01 06:00:00 UTC",...: 1
1 2 2 3 3 4 5 6 7 ...
## $ Date : Factor w/ 366 levels "2019-01-01 00:00:00 UTC",...: 1 1
1 1 1 1 1 1 1 1 1 ...
## $ month : int 1 1 1 1 1 1 1 1 1 1 ...
## $ time : int 6 6 6 6 6 6 7 7 7 8 ...
## $ min : int 0 0 20 20 40 40 0 20 40 0 ...
## $ week : int 1 1 1 1 1 1 1 1 1 1 ...
## $ day_week : Factor w/ 7 levels "금","목","수",...: 7 7 7 7 7 7 7 7
7 7 ...
## $ holiday : int 1 1 1 1 1 1 1 1 1 1 ...
## $ exposure_time : int 20 20 20 20 20 20 20 20 20 20 ...
## $ M_code : int 100346 100346 100346 100346 100346 100346 100305
100305 100305 100808 ...
## $ Product_code : int 201072 201079 201072 201079 201072 201079 200974
200974 200974 202377 ...
## $ Product_name : Factor w/ 1687 levels "\" 100%리얼 착즙 석류 \", 총 4 박
스",...: 1517 1523 1517 1523 1517 1523 1099 1099 1099 267 ...
## $ Product_group : Factor w/ 11 levels "가구","가전",...: 7 7 7 7 7 7 6 6 6
7 ...
## $ unit_price : int 39900 39900 39900 39900 39900 39900 59000 59000 5
9000 59900 ...
## $ total_sales : int 2099000 4371000 3262000 6955000 6672000 9337000 6
819000 15689000 25370000 16133000 ...
## $ Sales_quantity : int 52 109 81 174 167 234 115 265 430 269 ...
## $ temp : num -7.9 -7.9 -7.9 -7.9 -7.9 -7.9 -7.7 -7.7 -7.7 -7.7
...
## $ cold_wave : logi NA NA NA NA NA NA ...

```

[illegible]

[illegible]

```

## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## ..$ : num 297
## .. [list output truncated]

b[is.na(b)]<-0
b<-ifelse(shop_data$fine_dust=="주의보",1,
          ifelse(shop_data$fine_dust=="경보",2,0))

b[is.na(b)]<-0
table(b)

## b
##      0
## 37372

head(b)

## [1] 0 0 0 0 0 0

shop_data$dustvariable<-b
mean1<-shop_data %>% filter(b==0) %>% summarise(mean(Sales_quantity))
mean1

##      mean(Sales_quantity)
## 1                297.5393

mean2<-shop_data %>% filter(b==1) %>% summarise(mean(Sales_quantity))
mean2

##      mean(Sales_quantity)
## 1                      NaN

mean3<-shop_data %>% filter(b==2) %>% summarise(mean(Sales_quantity))
mean3

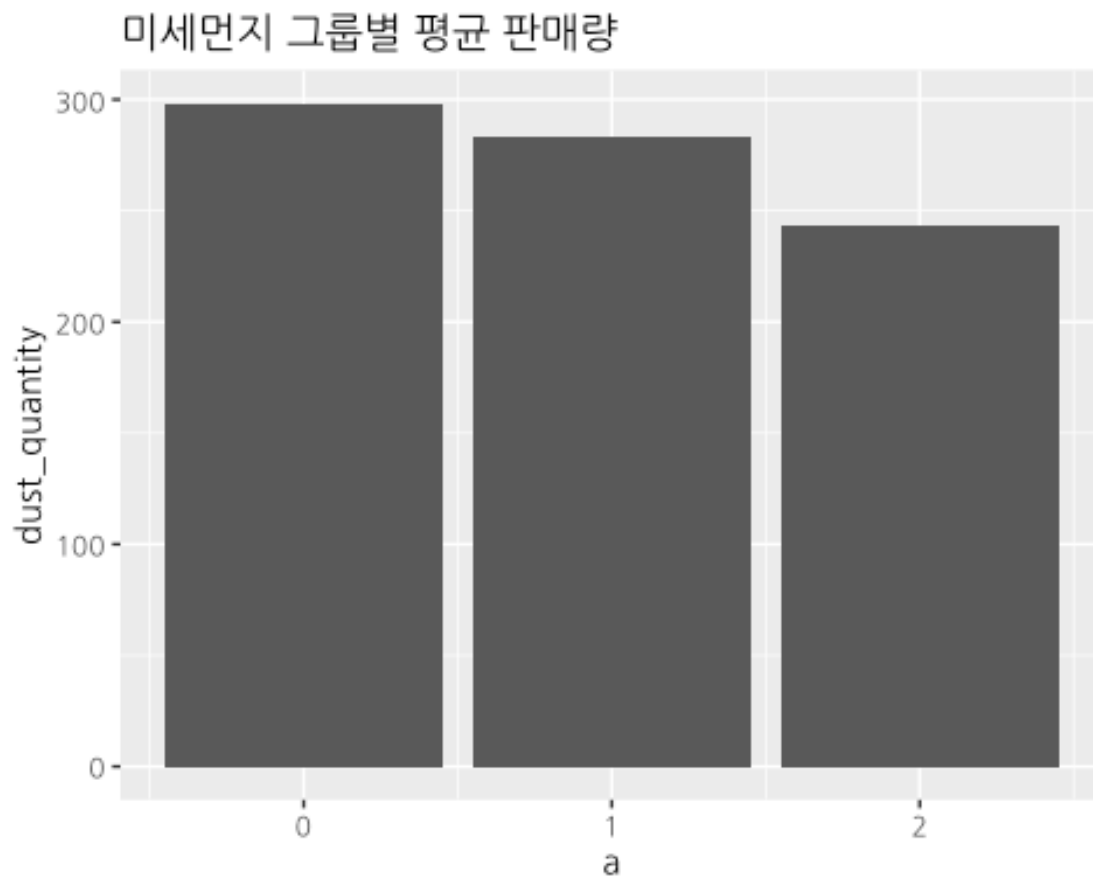
```

```
## mean(Sales_quantity)
## 1 NaN

dust_quantity<-c(298.3561,283.3523,242.8)
dustframe<-data.frame(a=c(0,1,2),dust_quantity)
dustframe

## a dust_quantity
## 1 0 298.3561
## 2 1 283.3523
## 3 2 242.8000

ggplot(data=dustframe)+geom_col(aes(x=a,y=dust_quantity))+ggtitle("미세먼지 그룹별 평균 판매량")
```



#미세먼지에 따른 차이는 거의 없다고 보임

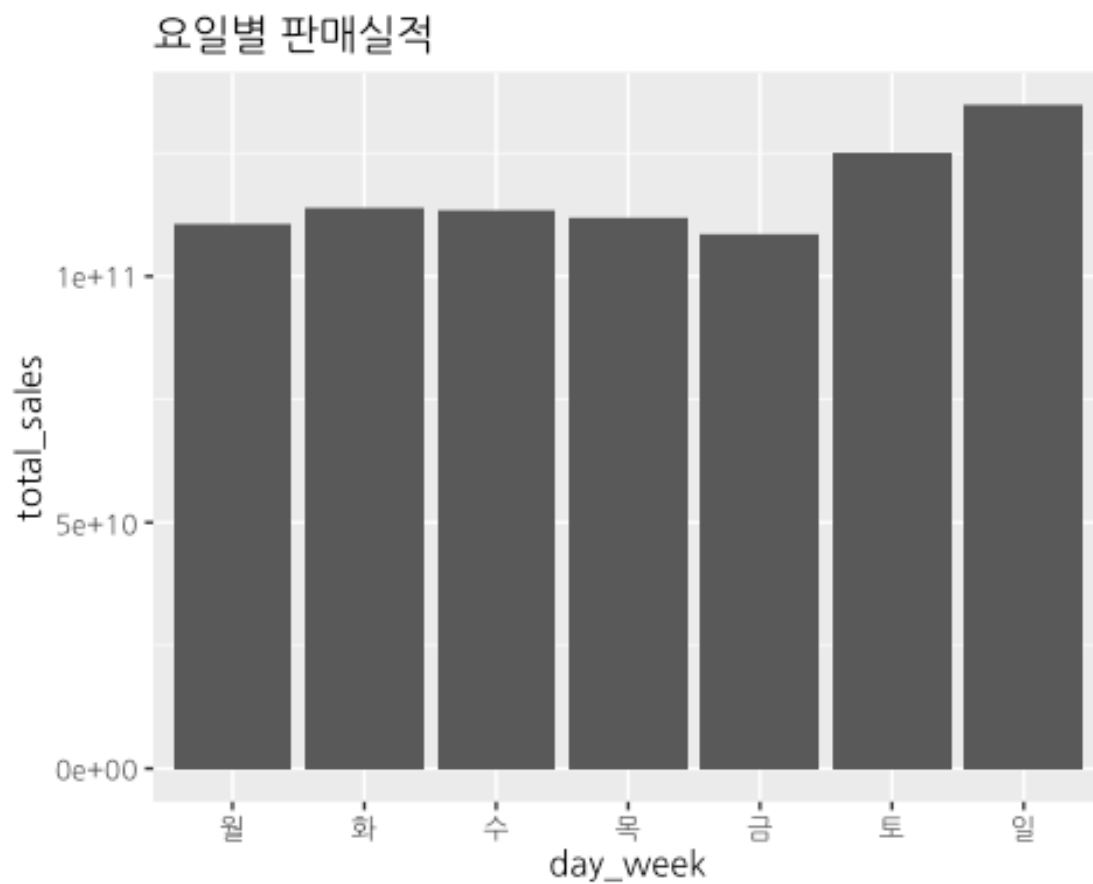
####이상 기상데이터 담당 김수현이었습니다

#- 상품군 편성 횟수는 count 문제 생각해 봐야할 것 같아 질문을 넣지 않았습니다. #- 휴일/비휴일의 경우 각자 해당하는 일수로 나눠야 할지(각 휴일/비휴일별 평균)?

##요일별## # 총 매출액이 요일에 따라 편차가 있는가?

```
week<- ggplot(data=shop_data, aes(x=day_week, y=total_sales))+  
  geom_col()+  
  scale_x_discrete(limits=c("월", "화", "수", "목", "금", "토", "일"))+  
  ggtitle("요일별 판매실적")
```

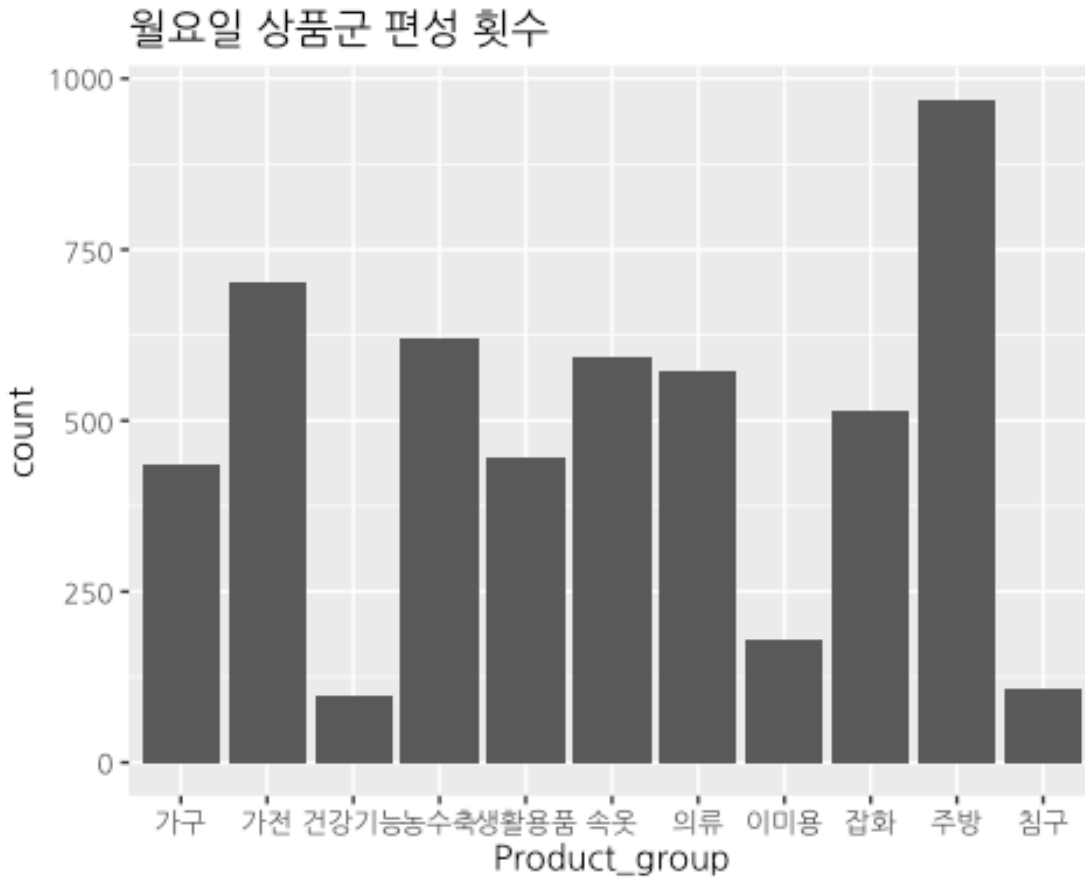
week



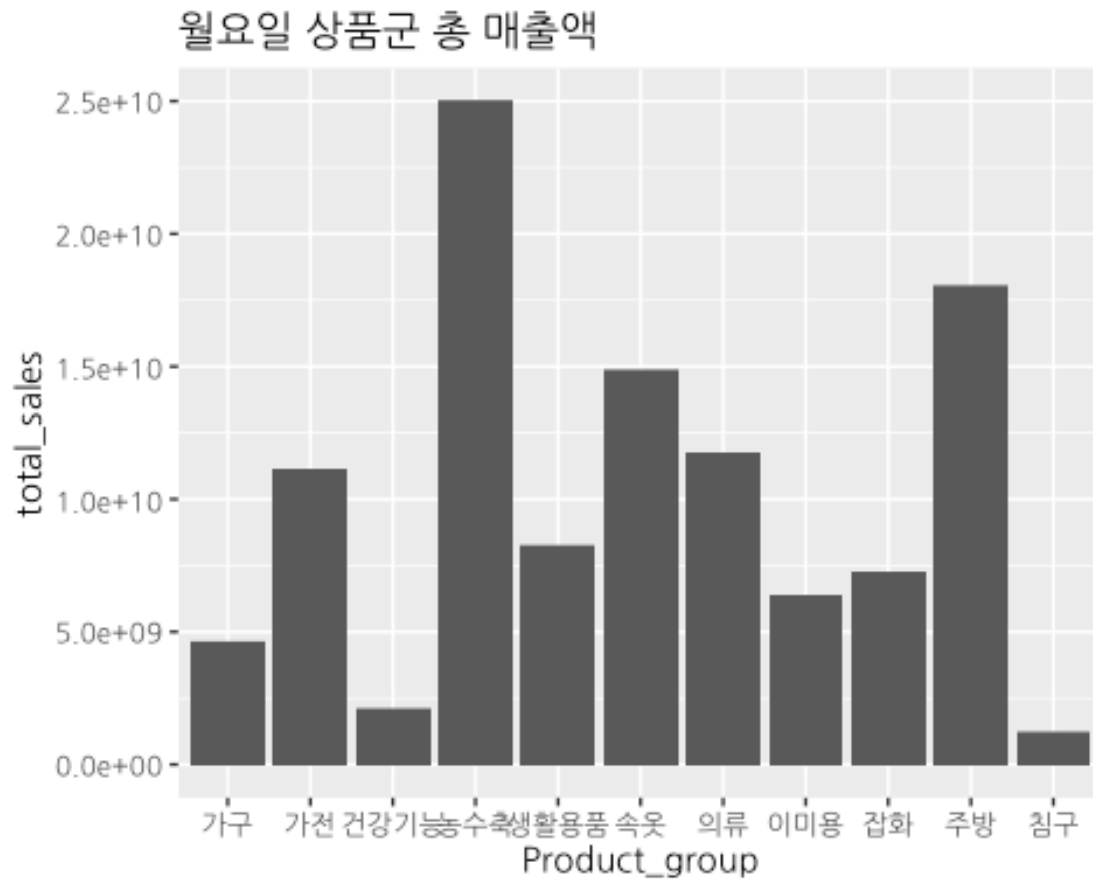
#월요일

상품군에 따른 총 매출액

```
Monday<-shop_data %>% filter(day_week=="월")  
ggplot(data=Monday)+  
  geom_bar(mapping=aes(x=Product_group))+  
  ggtitle("월요일 상품군 편성 횟수")
```



```
ggplot(data=Monday, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("월요일 상품군 총 매출액")
```

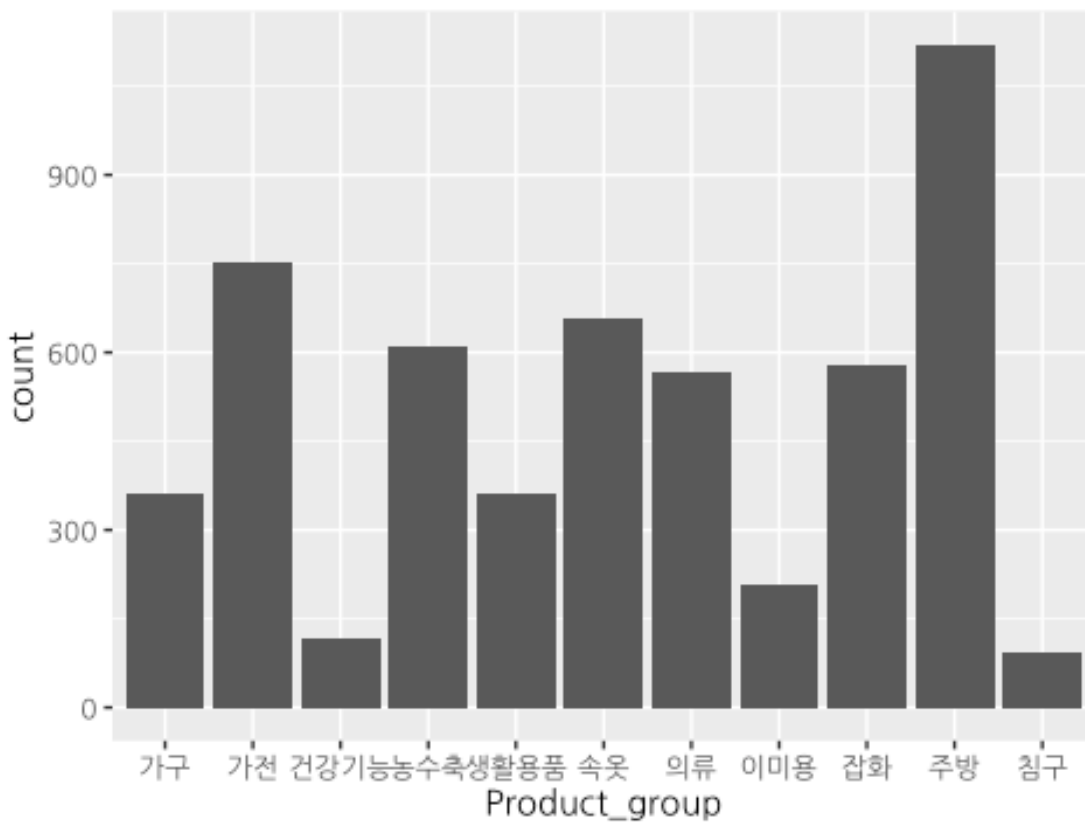



#화요일

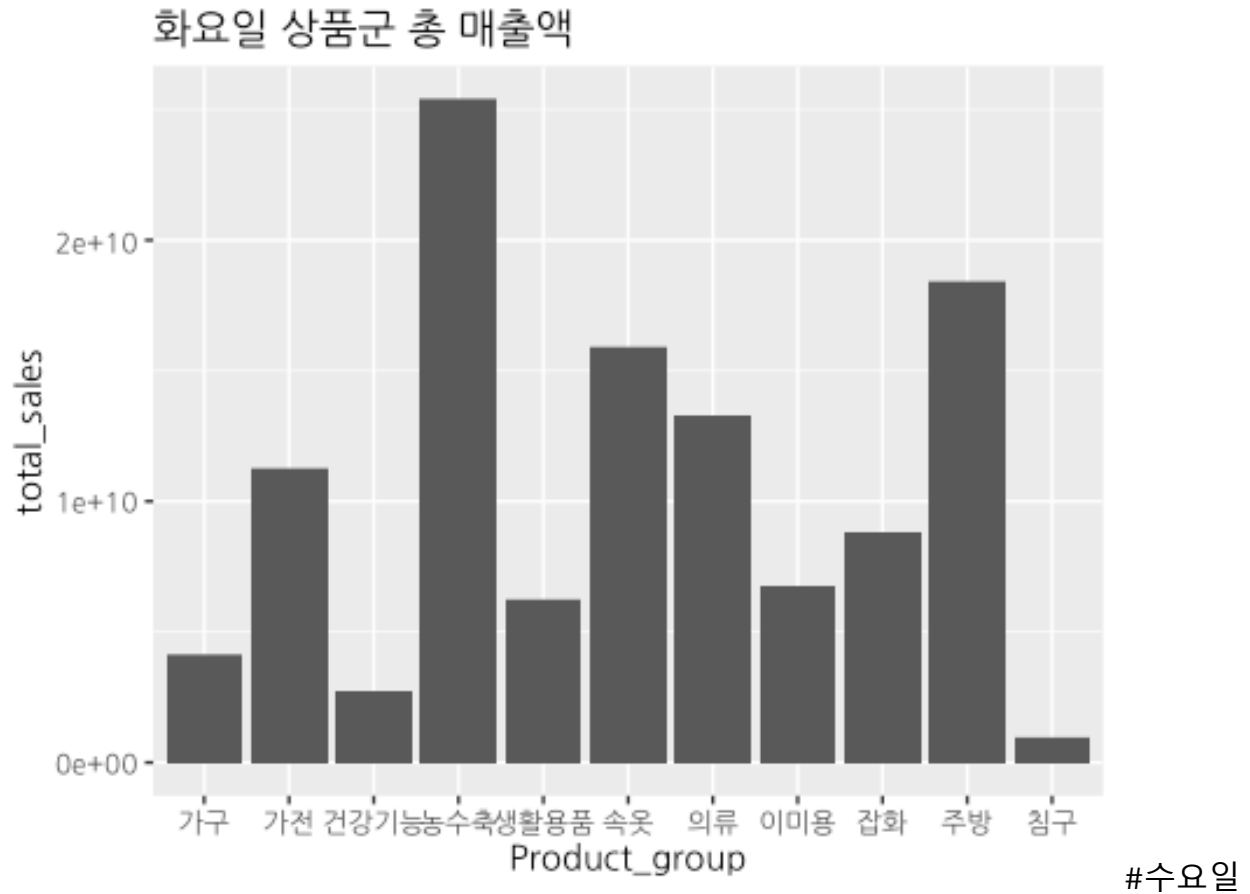
상품군에 따른 총 매출액

```
Tuesday<-shop_data %>% filter(day_week=="화")
ggplot(data=Tuesday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("화요일 상품군 편성 횟수")
```

화요일 상품군 편성 횟수



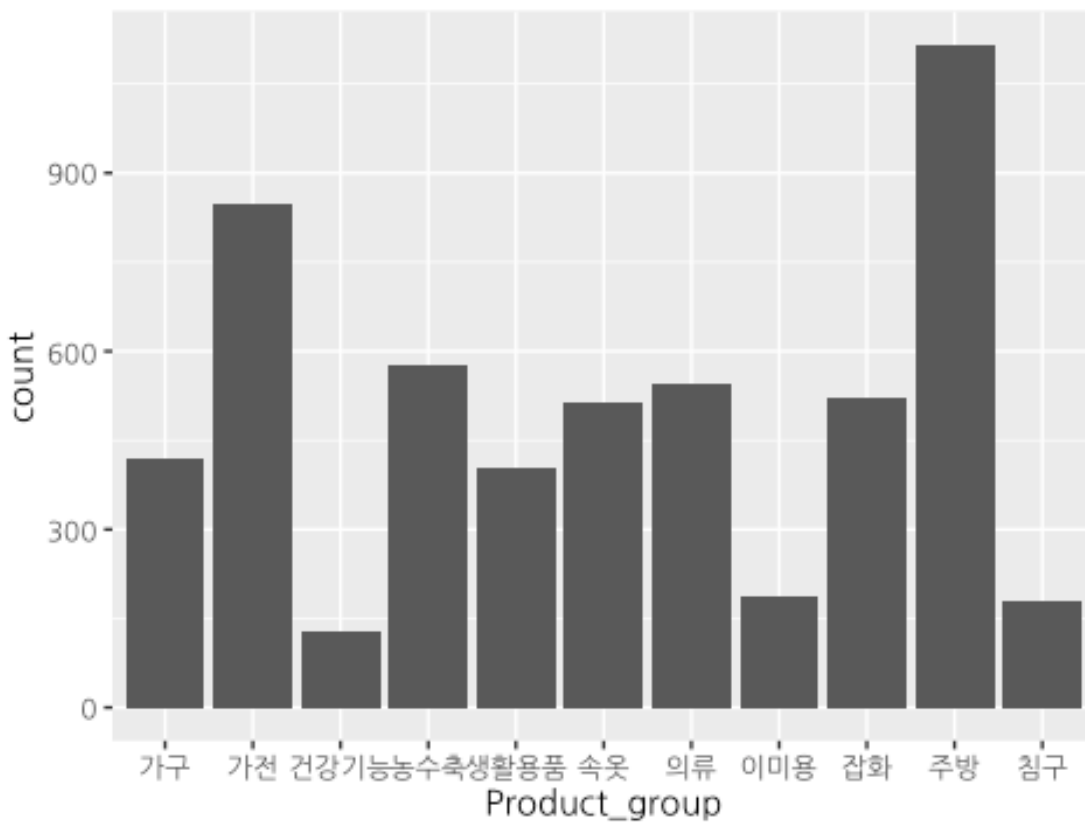
```
ggplot(data=Tuesday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("화요일 상품군 총 매출액")
```



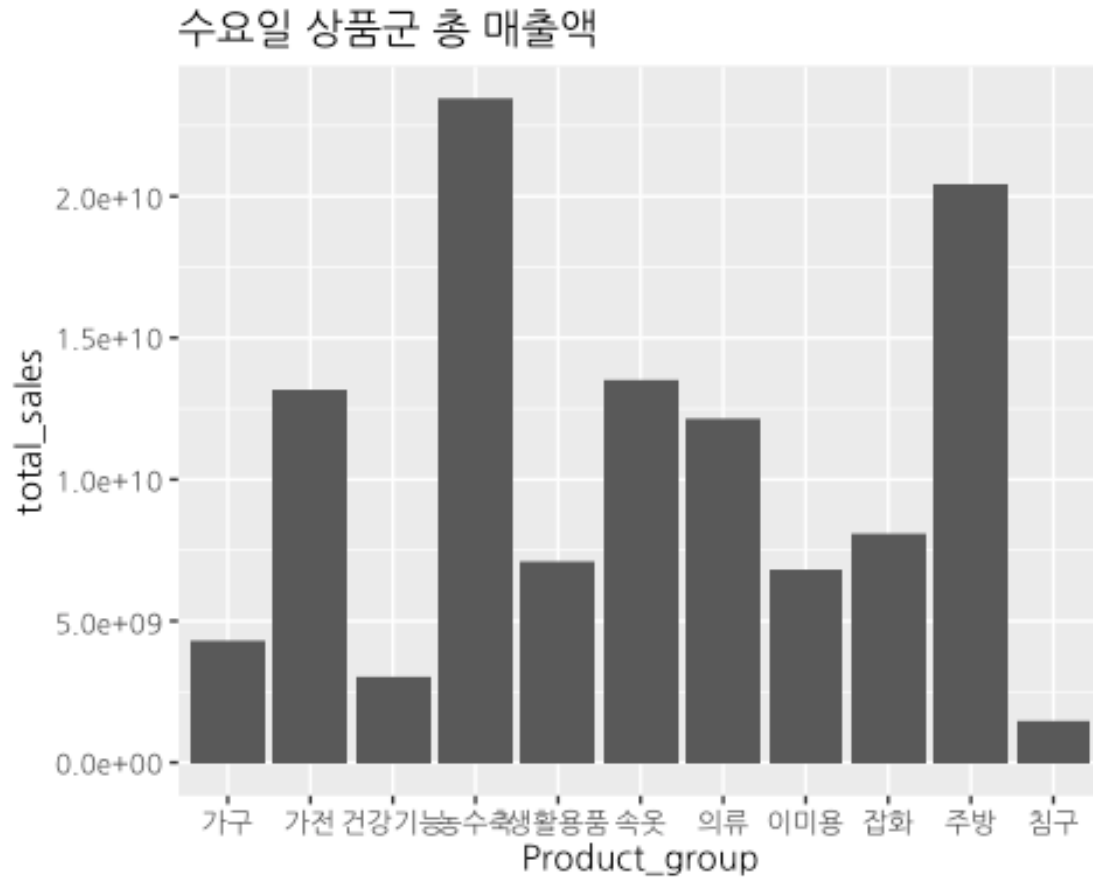
상품군에 따른 총 매출액

```
Wednesday<-shop_data %>% filter(day_week=="수")
ggplot(data=Wednesday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("수요일 상품군 편성 횟수")
```

수요일 상품군 편성 횟수



```
ggplot(data=wednesday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("수요일 상품군 총 매출액")
```

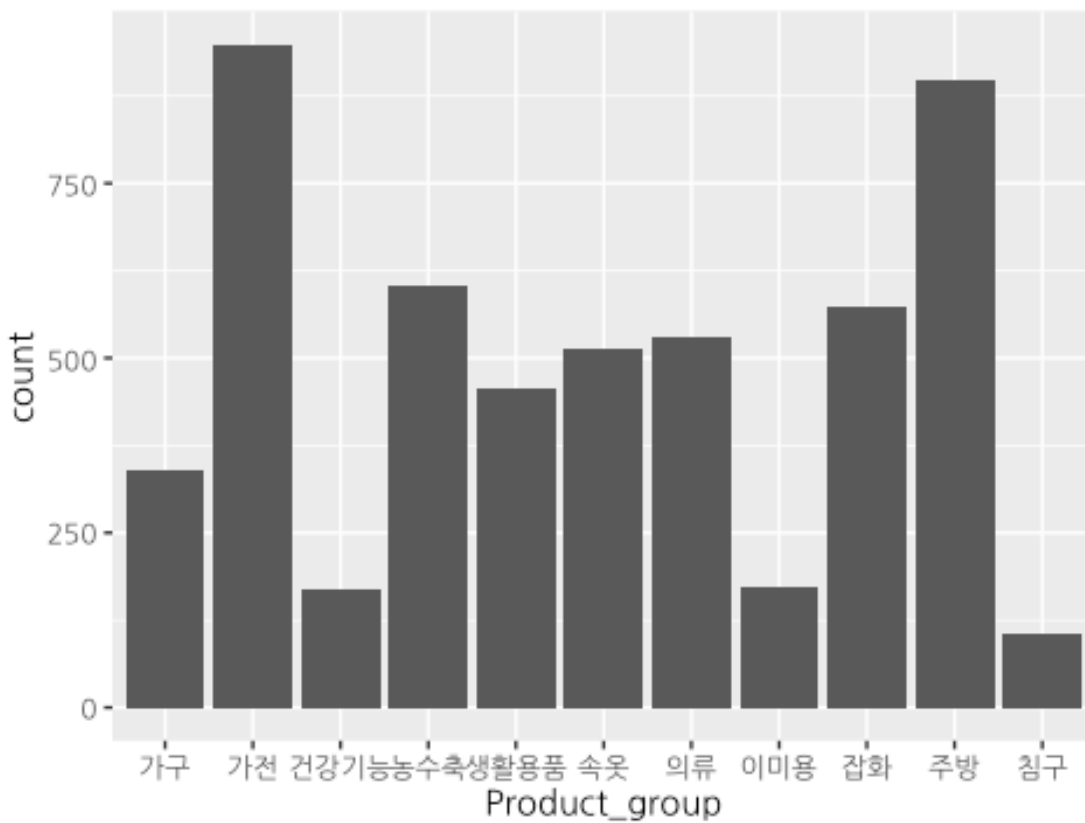


#목요일

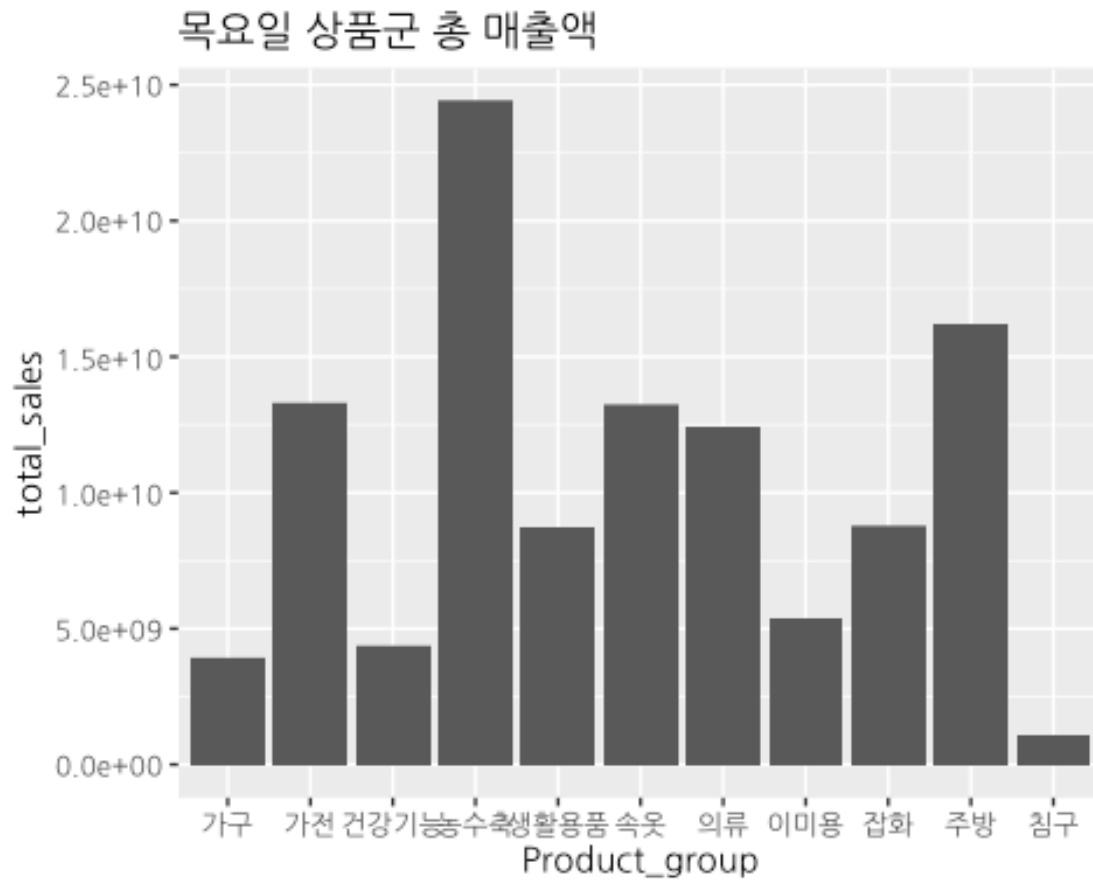
상품군에 따른 총 매출액

```
Thursday<-shop_data %>% filter(day_week=="목")
ggplot(data=Thursday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("목요일 상품군 편성 횟수")
```

목요일 상품군 편성 횟수



```
ggplot(data=Thursday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("목요일 상품군 총 매출액")
```

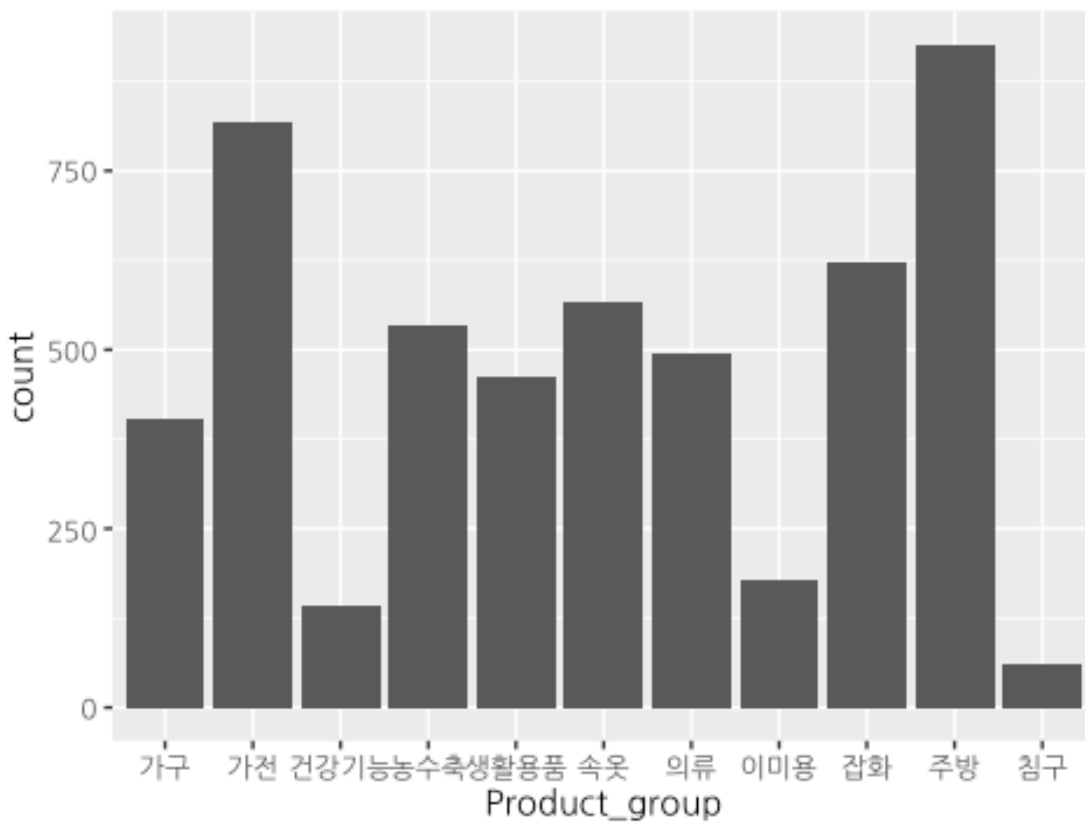


#금요일

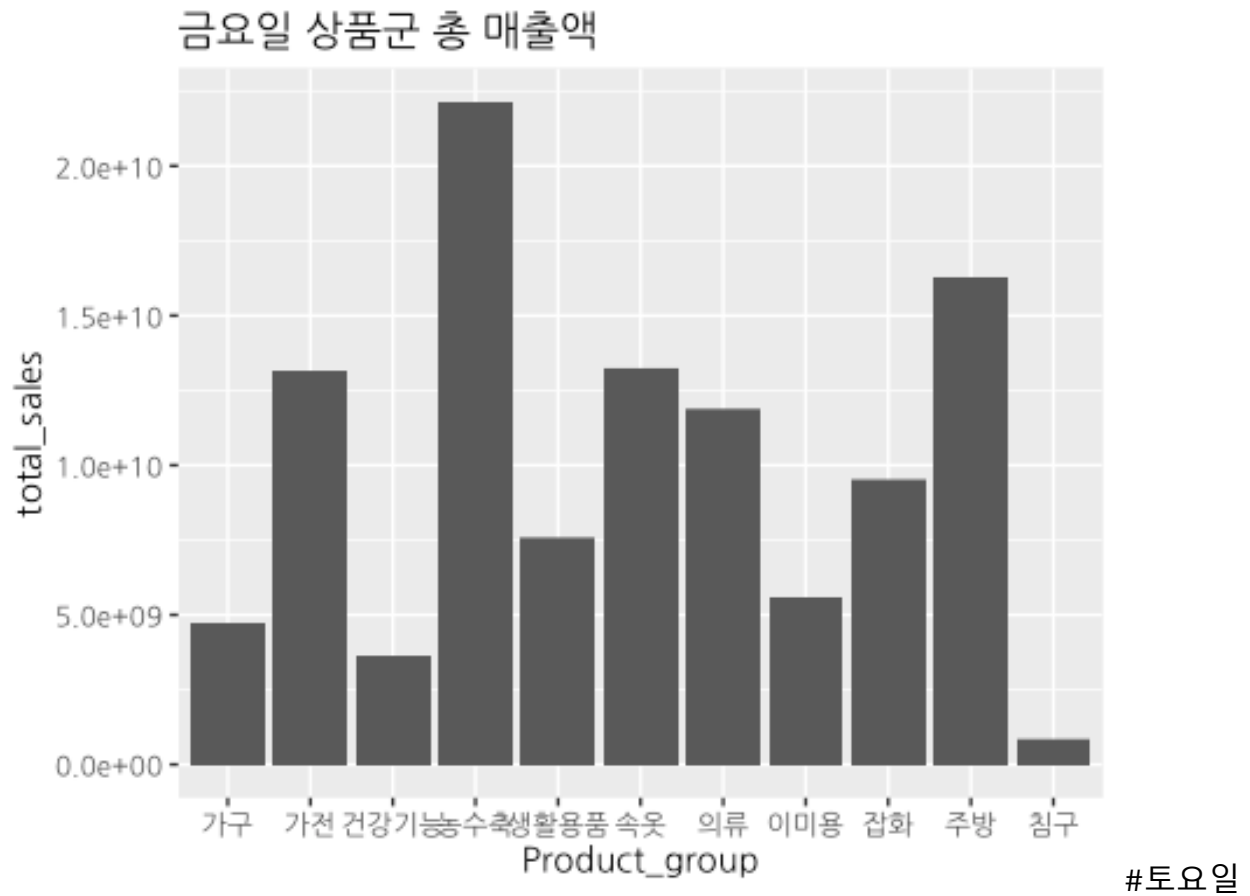
상품군에 따른 총 매출액

```
Friday<-shop_data %>% filter(day_week=="금")
ggplot(data=Friday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("금요일 상품군 편성 횟수")
```

금요일 상품군 편성 횟수

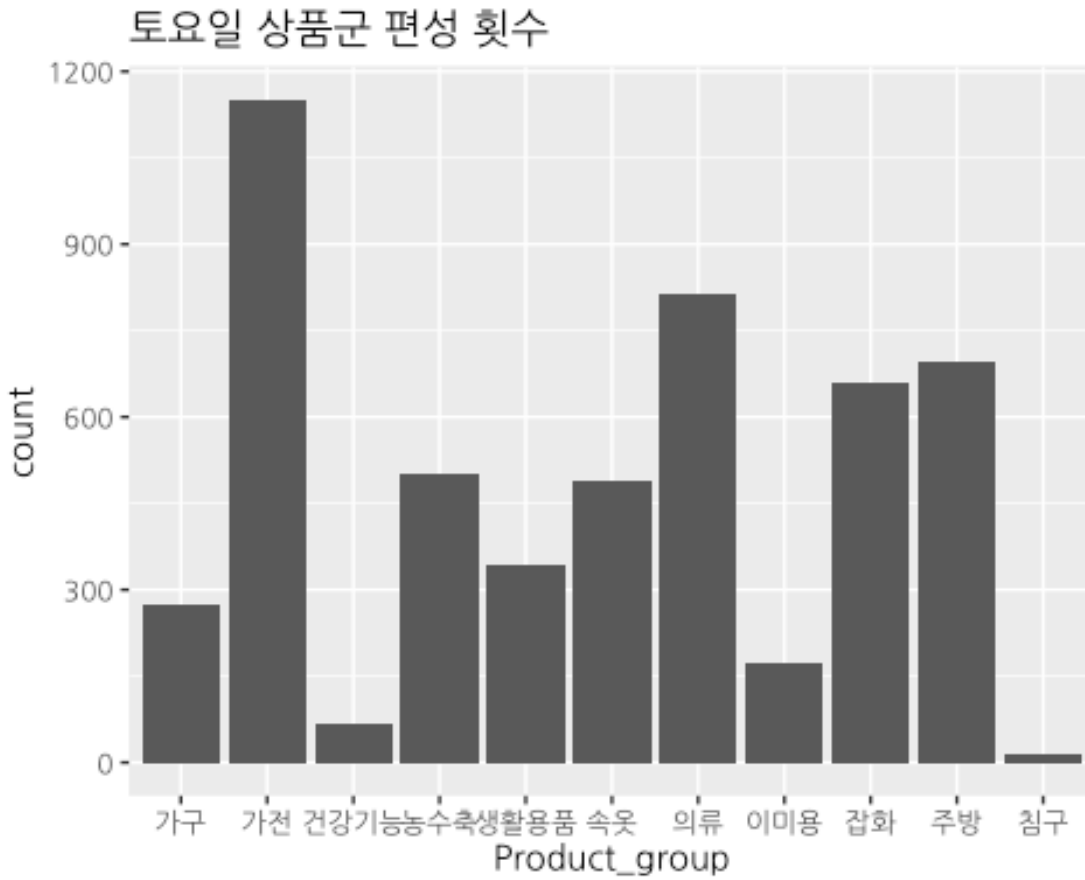


```
ggplot(data=Friday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("금요일 상품군 총 매출액")
```

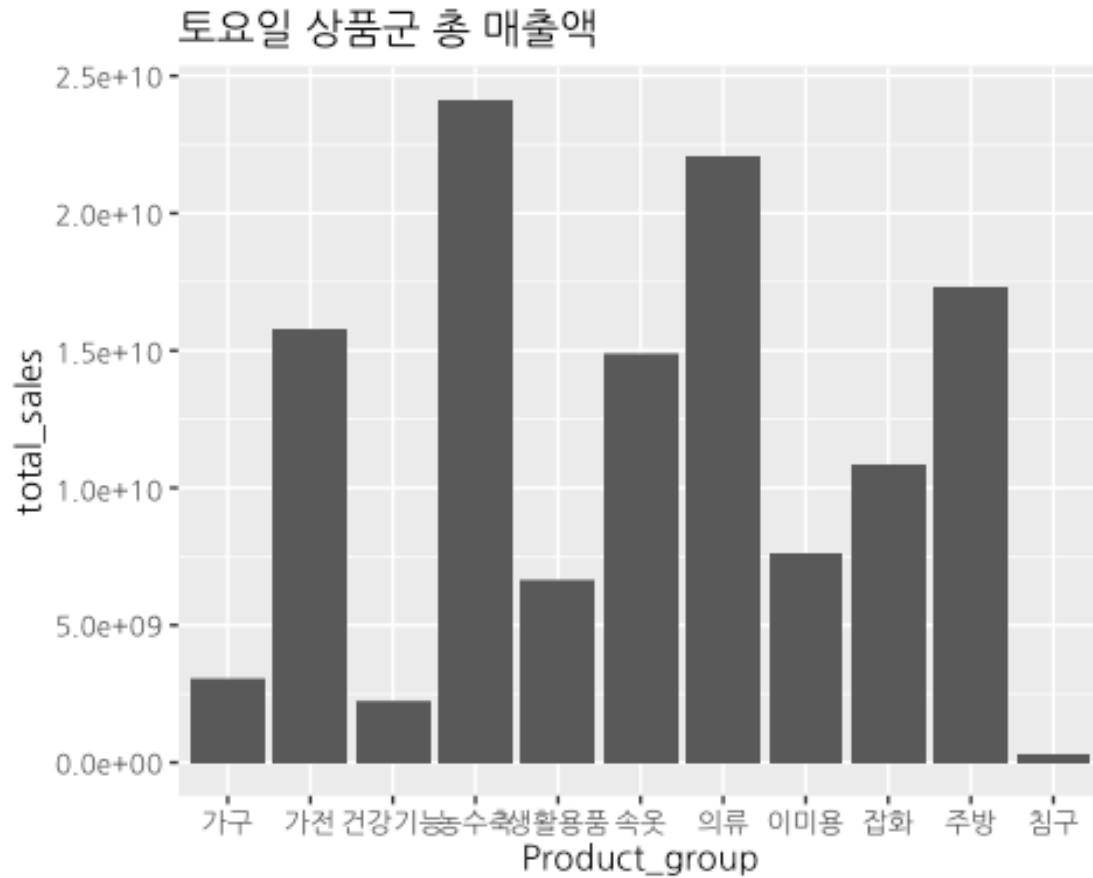



상품군에 따른 총 매출액

```
Saturday<-shop_data %>% filter(day_week=="토")
ggplot(data=Saturday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("토요일 상품군 편성 횟수")
```

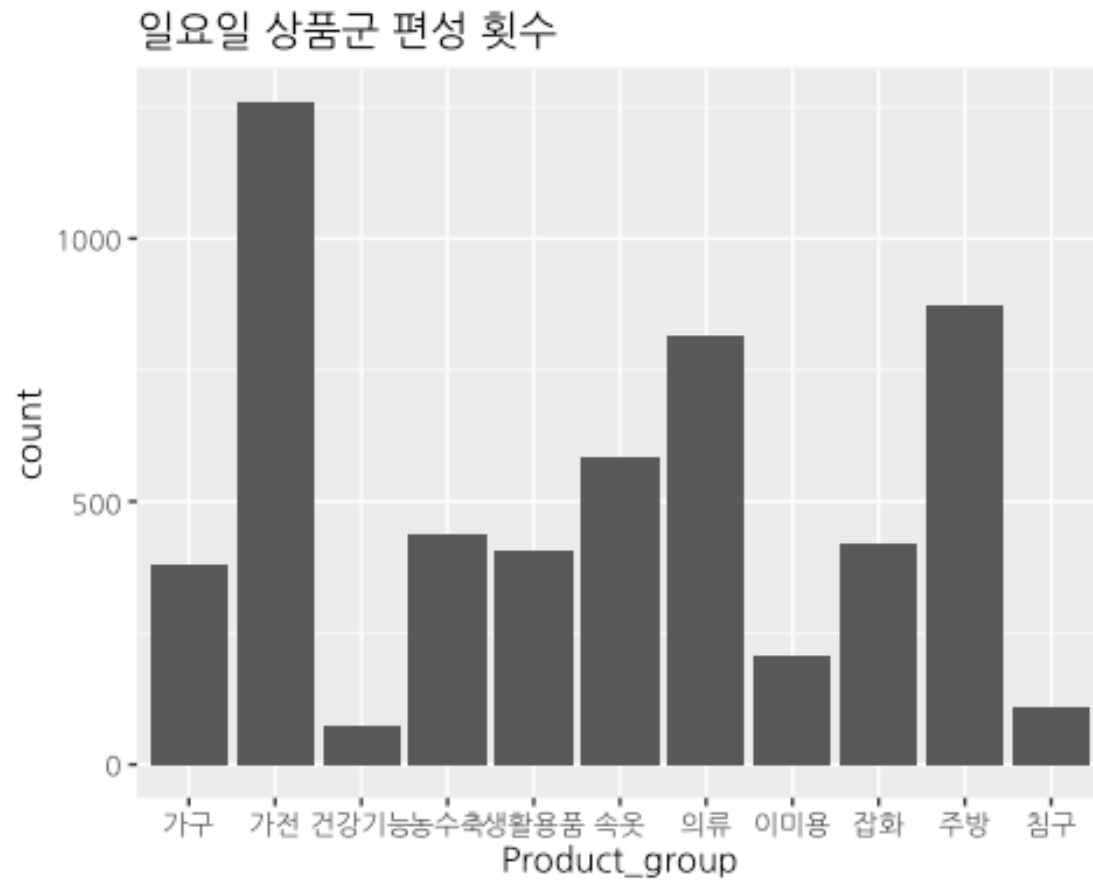


```
ggplot(data=Saturday, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("토요일 상품군 총 매출액")
```

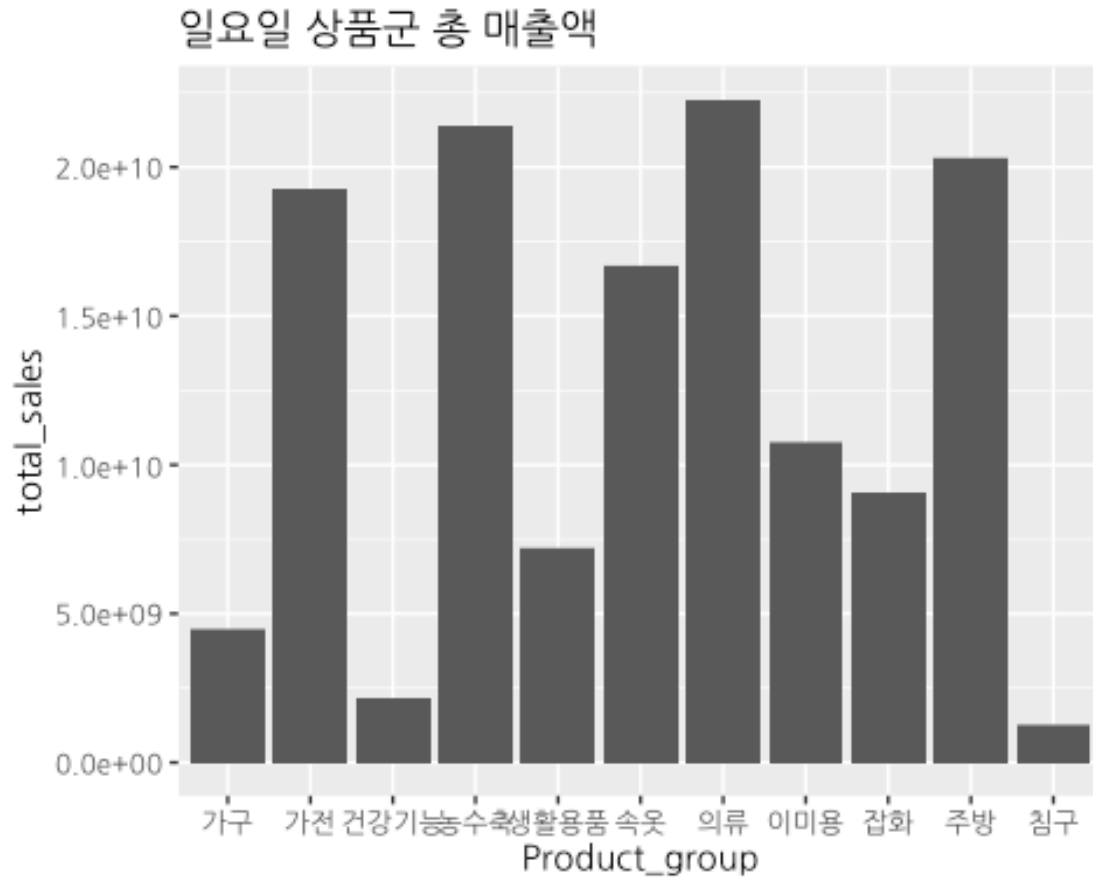


#일요일 상품군에 따른 총 매출액

```
Sunday<-shop_data %>% filter(day_week=="일")
ggplot(data=Sunday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("일요일 상품군 편성 횟수")
```



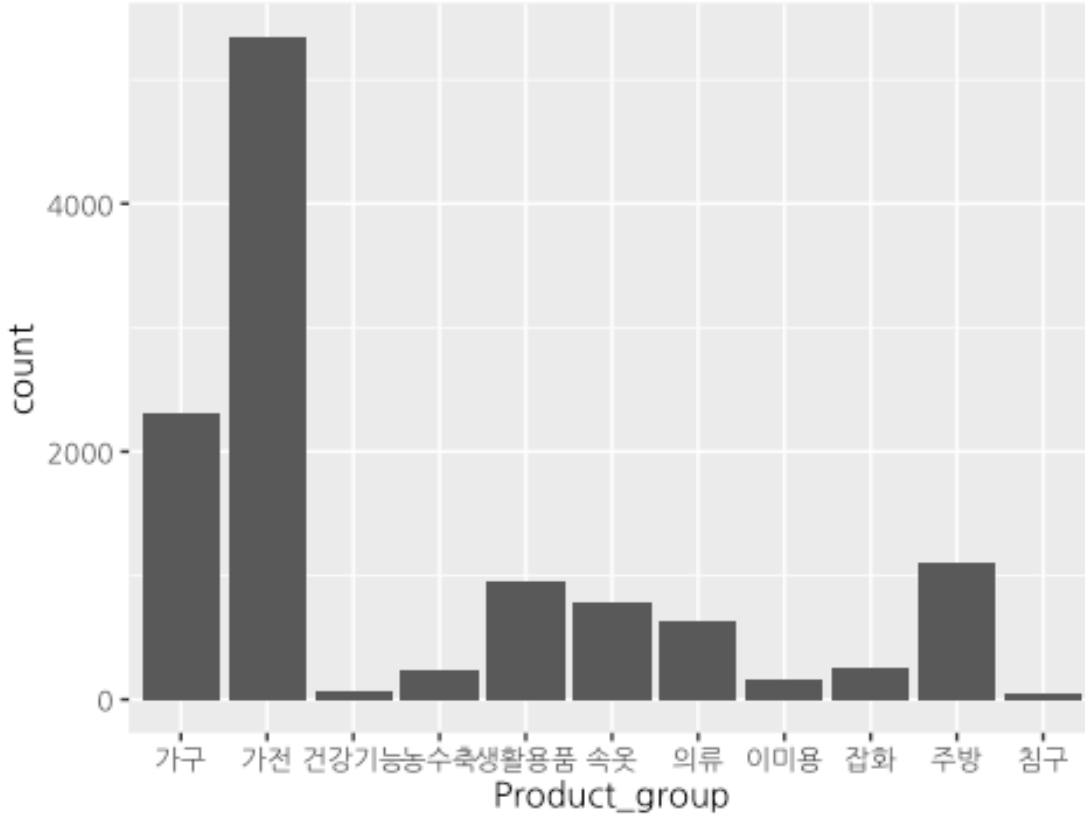
```
ggplot(data=Sunday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("일요일 상품군 총 매출액")
```



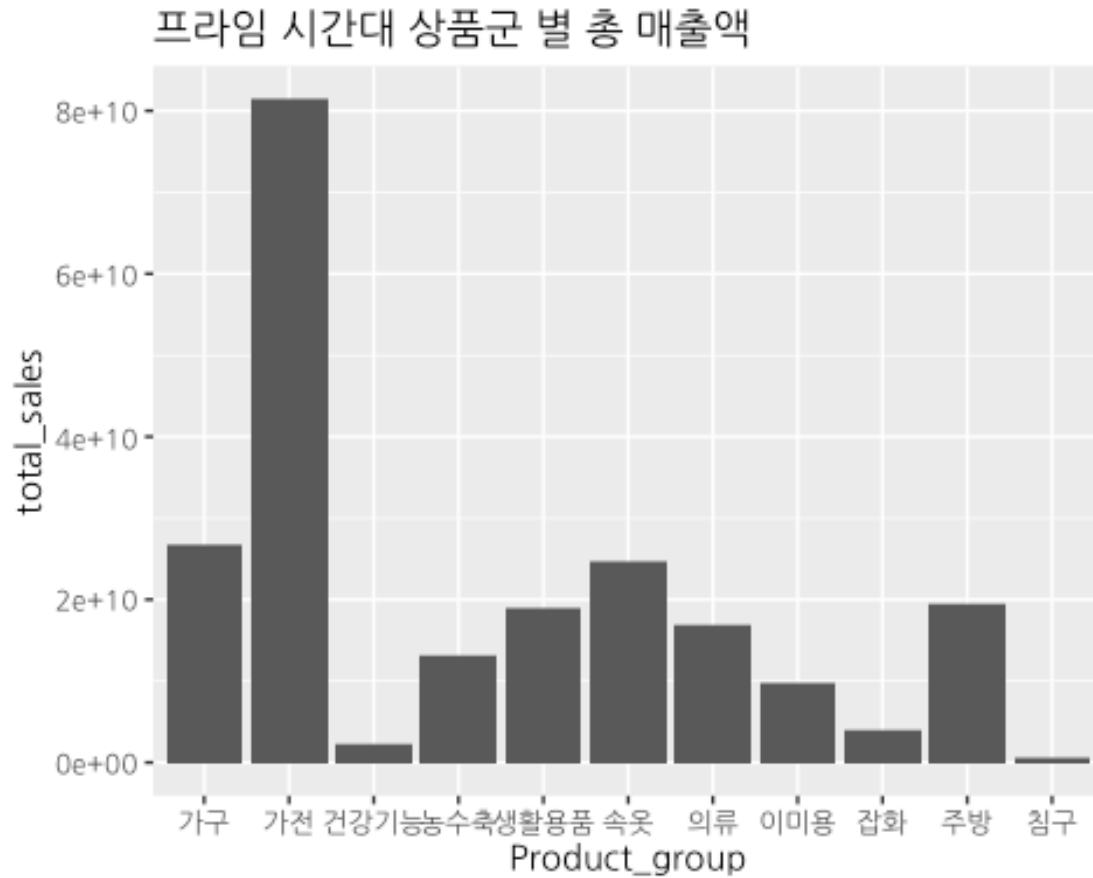
##시간대별## #주요 시간대의 상품군별 매출액 차지하는 정도? #프라임 시간대의 상품군별 매출액은 어떻게 되는가?

```
Primetime<-shop_data %>% filter(time>=20 & time<=23)
ggplot(data=Primetime)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("프라임 시간대 상품군 별 편성 횟수")
```

프라임 시간대 상품군 별 편성 횟수

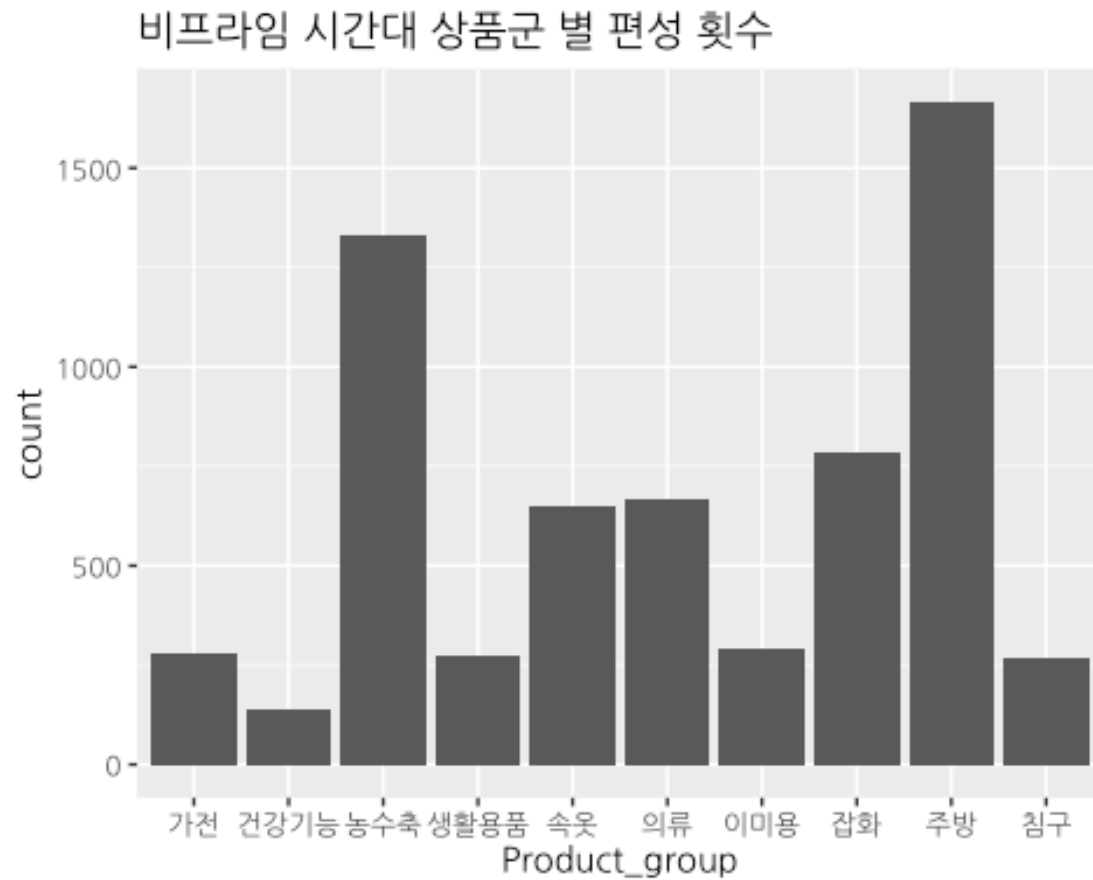


```
ggplot(data=Primetime, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("프라임 시간대 상품군 별 총 매출액")
```

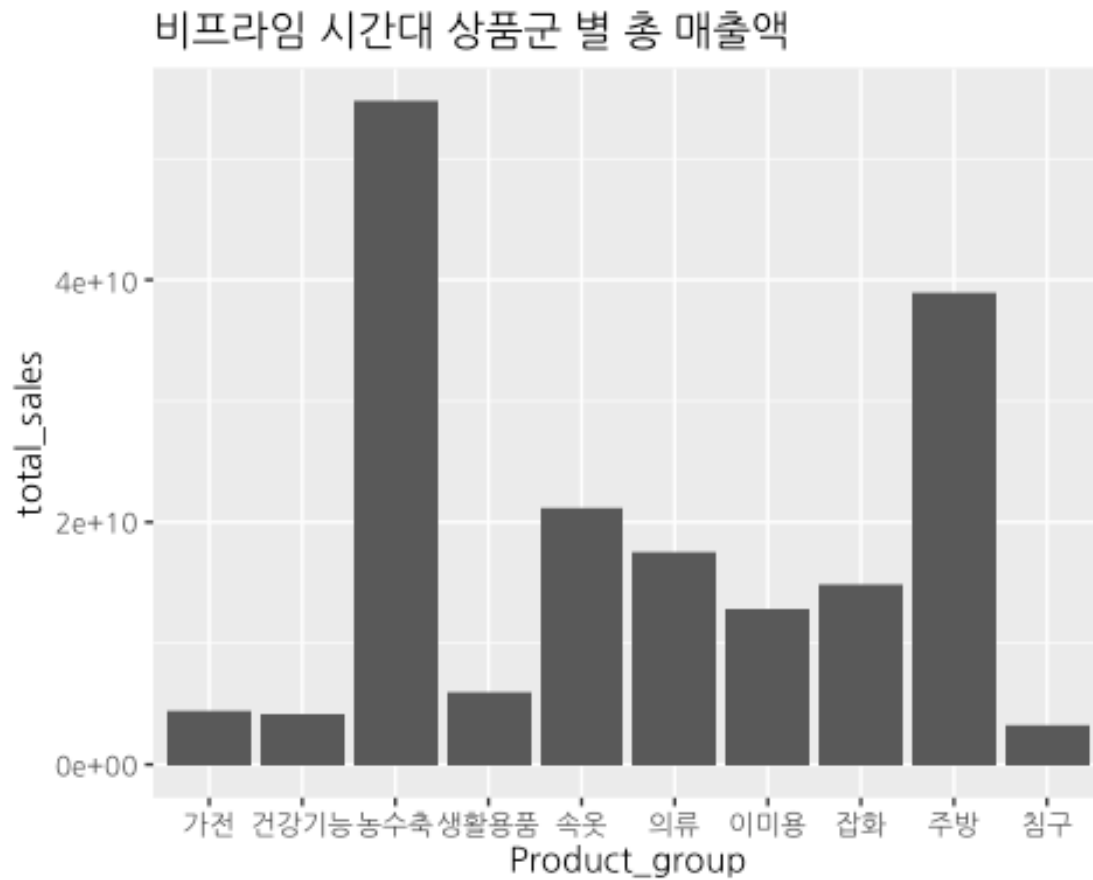


#비프라임 #비프라임 시간대의 상품군별 매출액은 어떻게 되는가?

```
NPrimetype<-shop_data %>% filter(time>=13 & time<=16)
ggplot(data=NPrimetype)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("비프라임 시간대 상품군 별 편성 횟수")
```

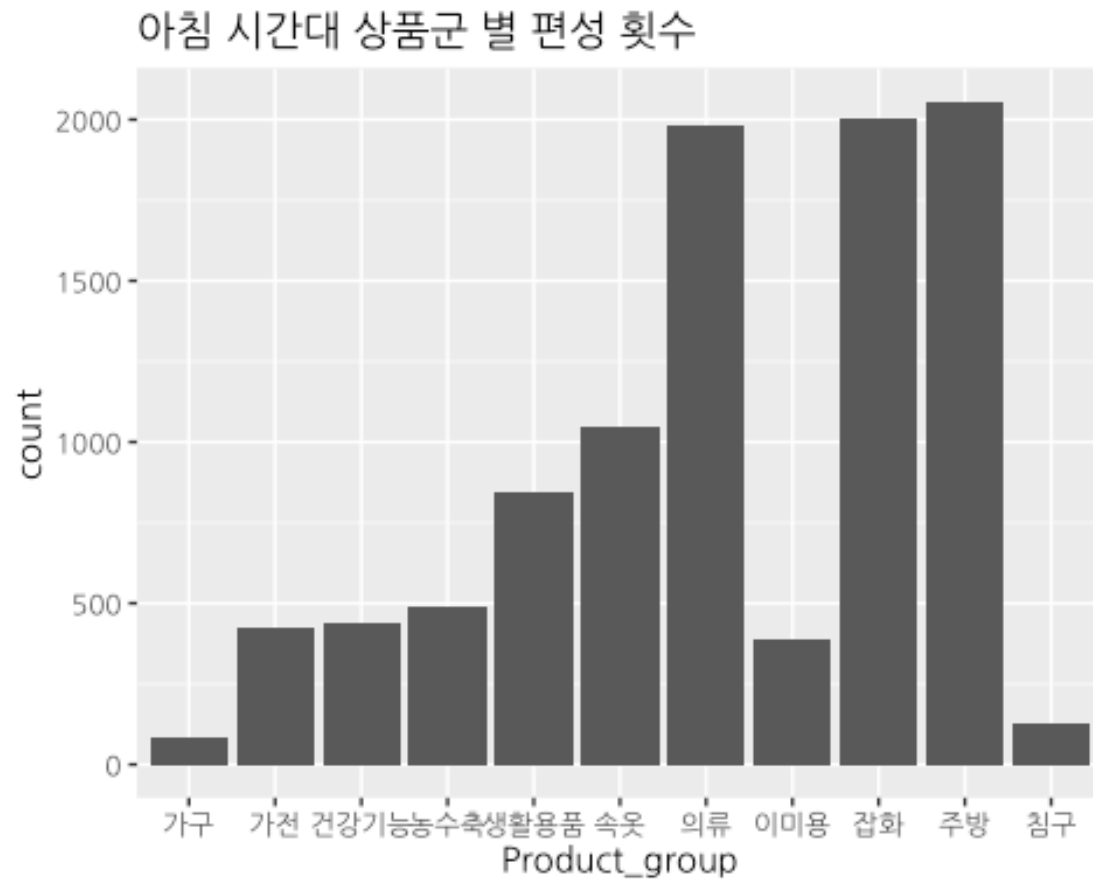


```
ggplot(data=NPrimetype, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("비프라임 시간대 상품군 별 총 매출액")
```

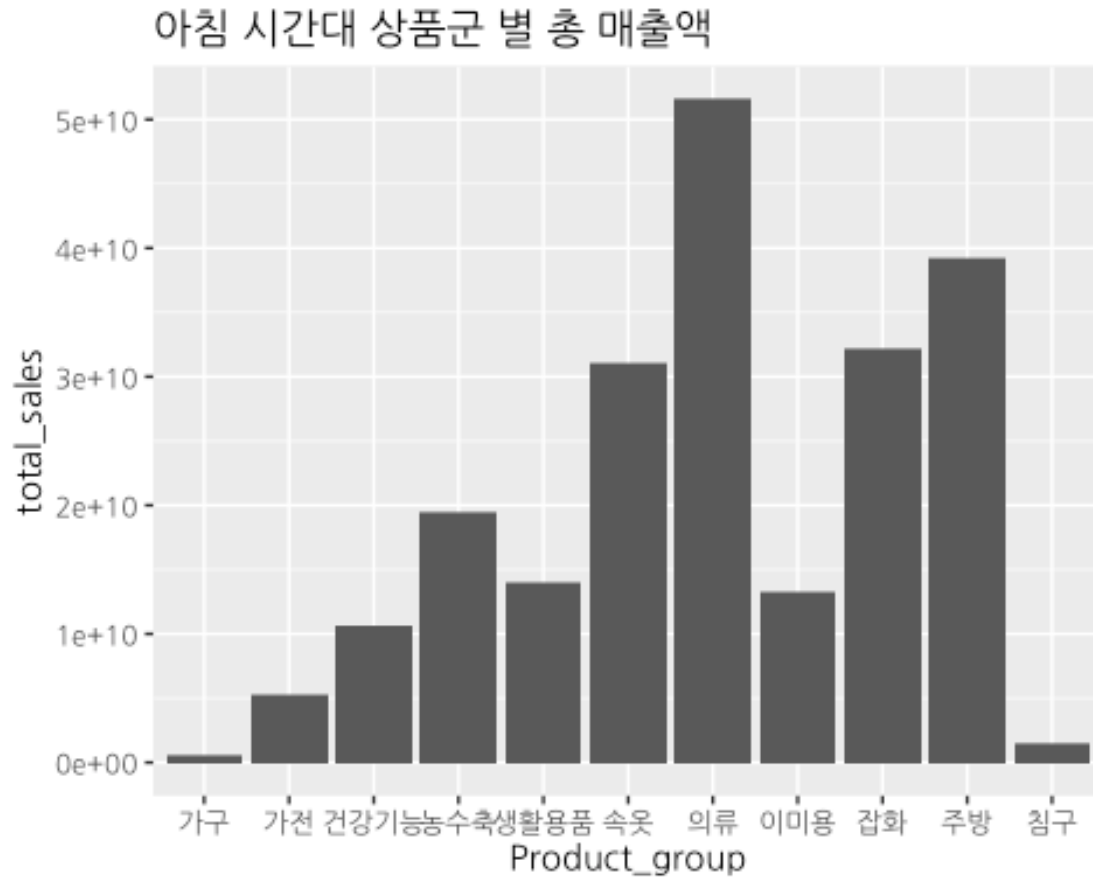



#아침시간 #아침 시간대의 상품군별 매출액은 어떻게 되는가?

```
Morning<-shop_data %>% filter(time>=6 & time<=11)
ggplot(data=Morning)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("아침 시간대 상품군 별 편성 횟수")
```

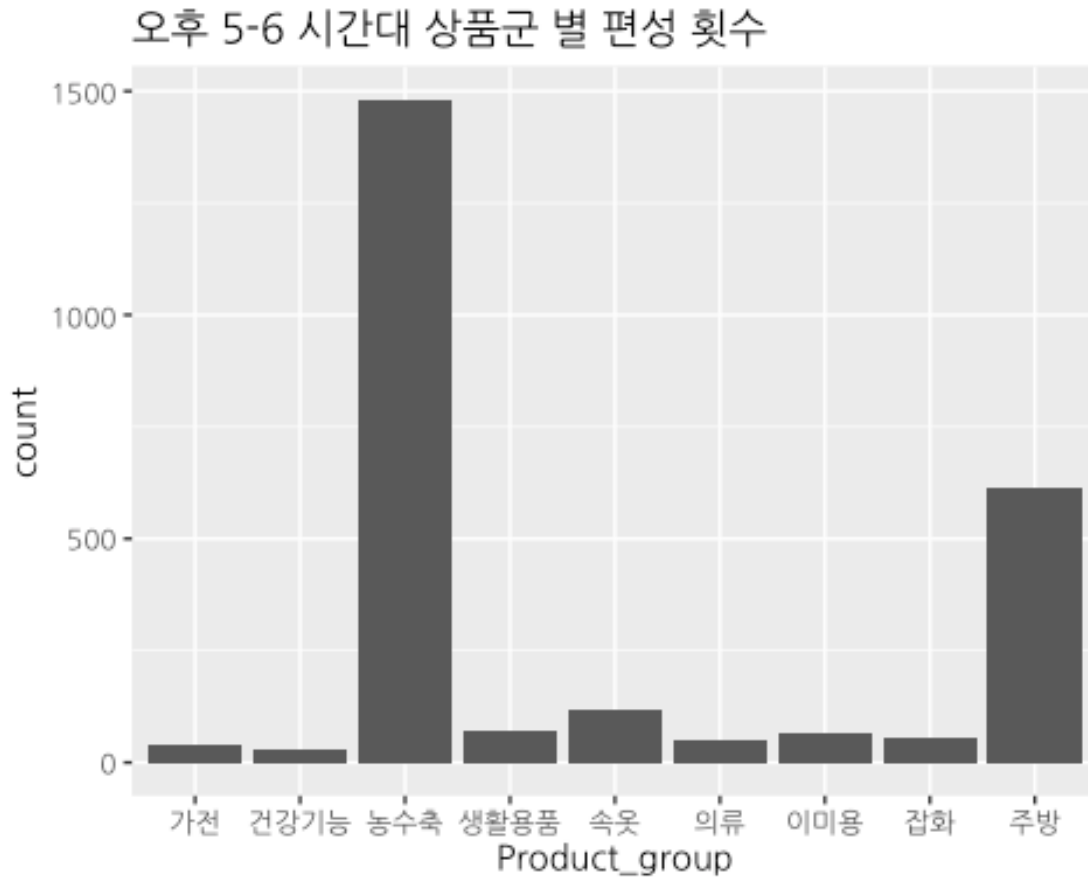


```
ggplot(data=Morning, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("아침 시간대 상품군 별 총 매출액")
```

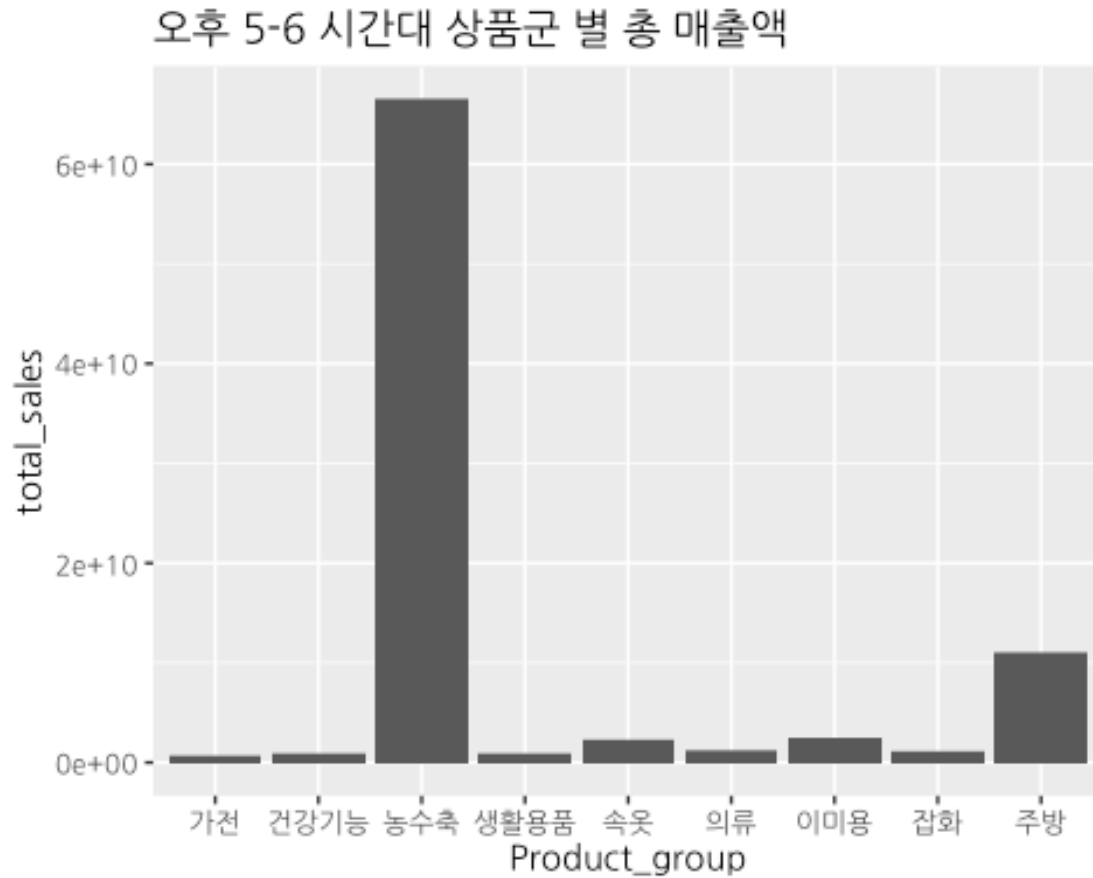


#오후시간 #오후 시간대의 상품군별 매출액은 어떻게 되는가?

```
Afternoon<-shop_data %>% filter(time>=17 & time<=18)
ggplot(data=Afternoon)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("오후 5-6 시간대 상품군 별 편성 횟수")
```

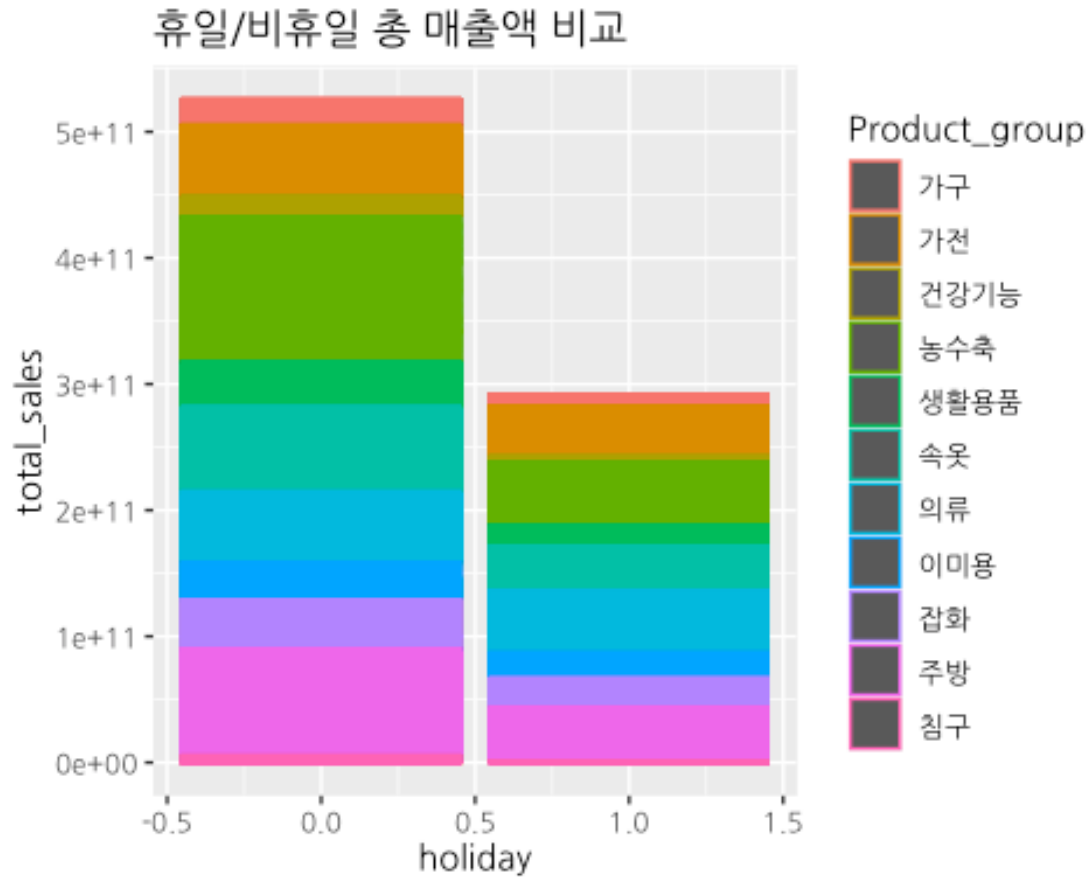


```
ggplot(data=Afternoon, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("오후 5-6 시간대 상품군 별 총 매출액")
```



#휴일/비휴일 #휴일/비휴일의 매출액 차이는 얼마나 나는가?

```
ggplot(data=shop_data, aes(x=holiday, y=total_sales, group=Product_group, color=Product_group))+
  geom_col(position="stack")+
  ggtitle("휴일/비휴일 총 매출액 비교")
```

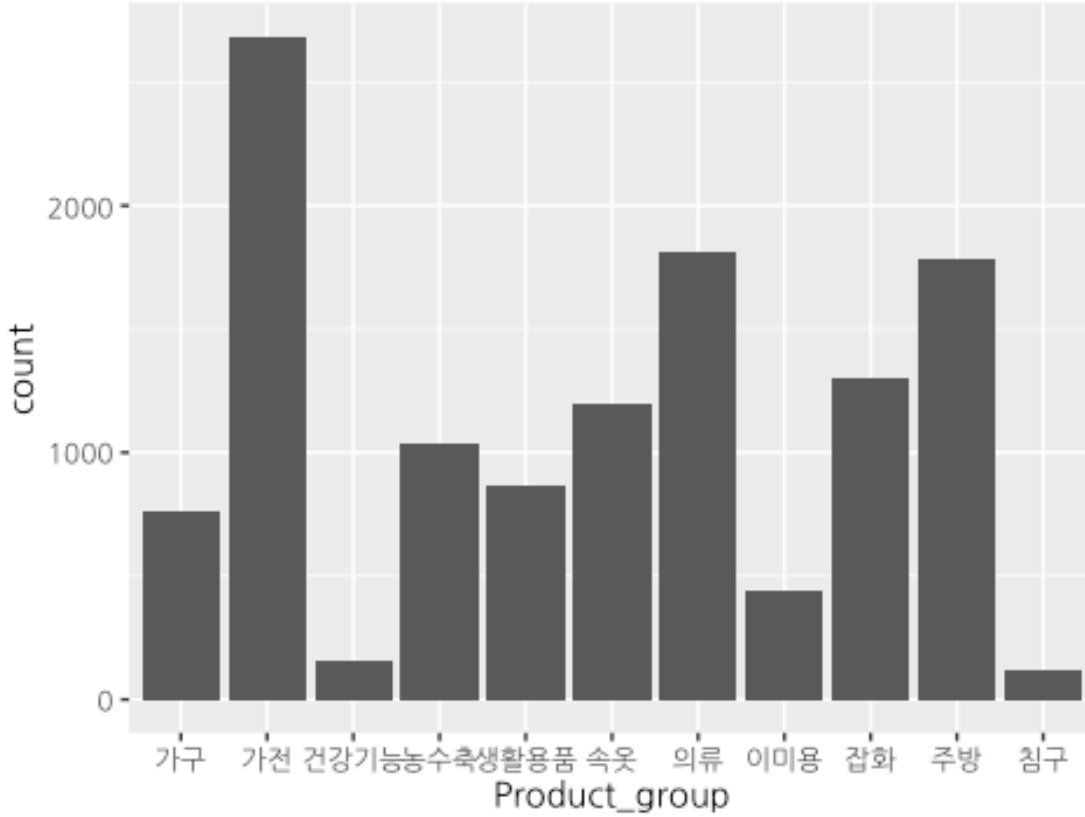


#휴일의

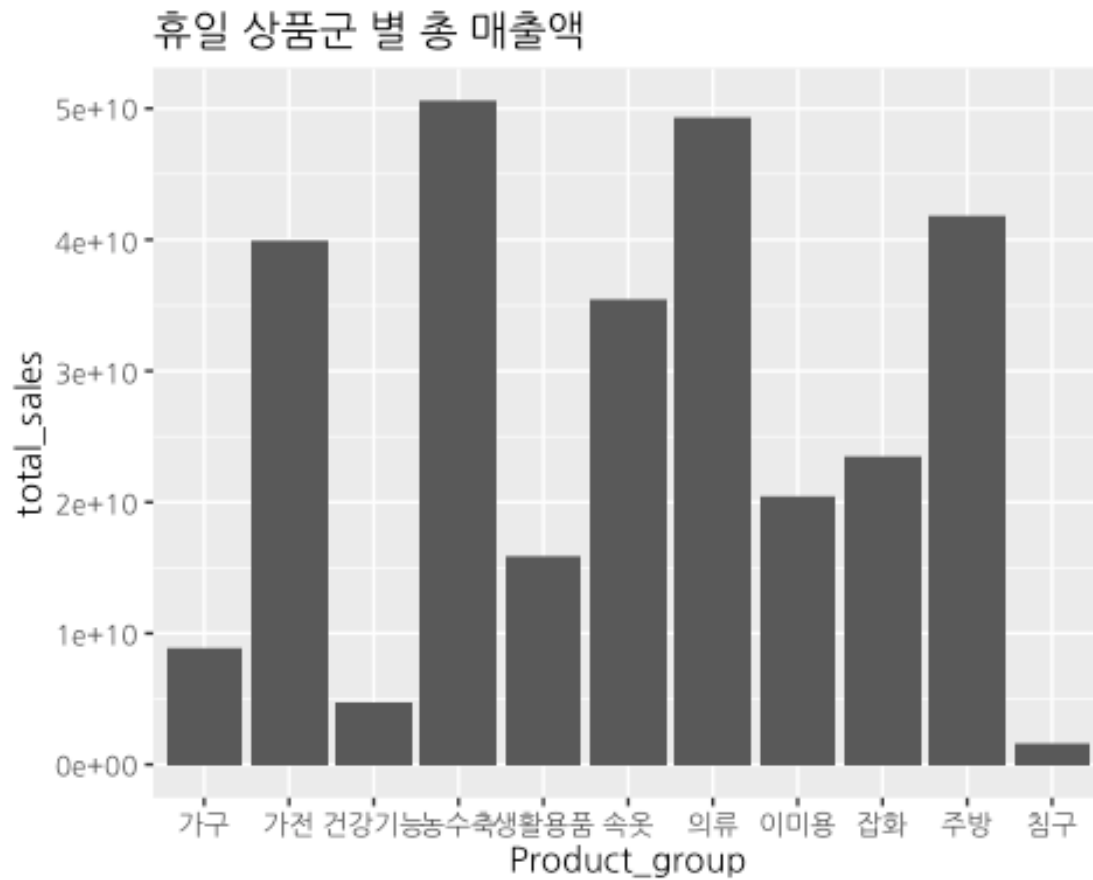
제품군별 매출액 차이는 얼마나 나는가? #휴일의 시간대별 매출액 추이?

```
holiday<-shop_data %>% filter(holiday==1)
ggplot(data=holiday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("휴일 상품군 별 편성 횟수")
```

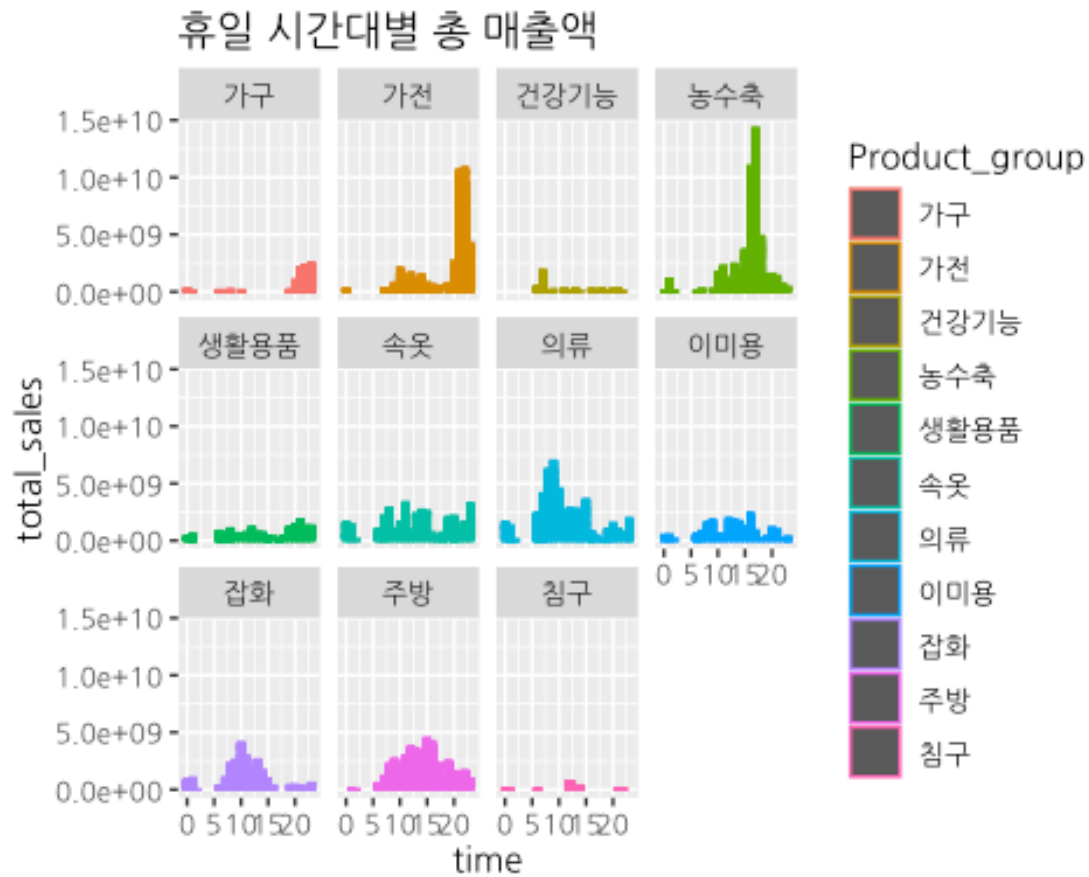
휴일 상품군 별 편성 횟수



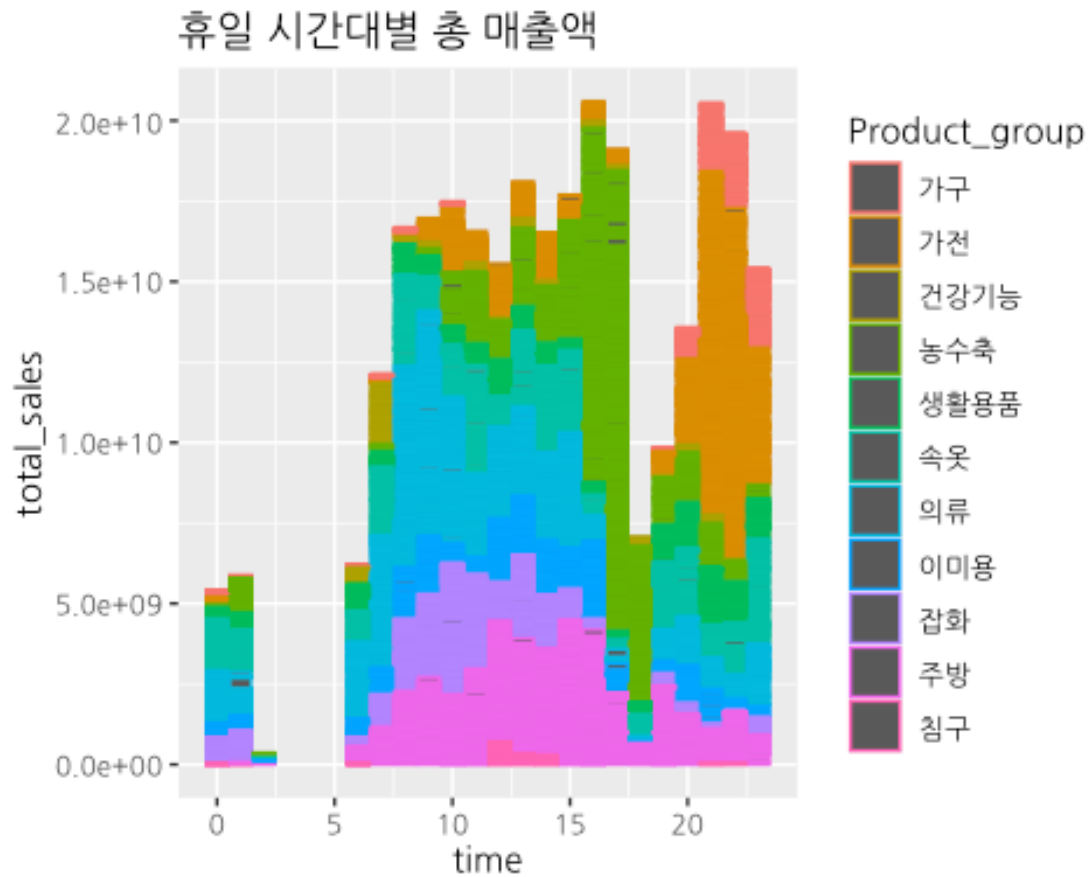
```
ggplot(data=holiday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("휴일 상품군 별 총 매출액")
```



```
ggplot(data=holiday,aes(x=time, y=total_sales,group=Product_group,colour = Product_group))+
  geom_col()+
  ggtitle("휴일 시간대별 총 매출액")+
  facet_wrap(~ Product_group)
```

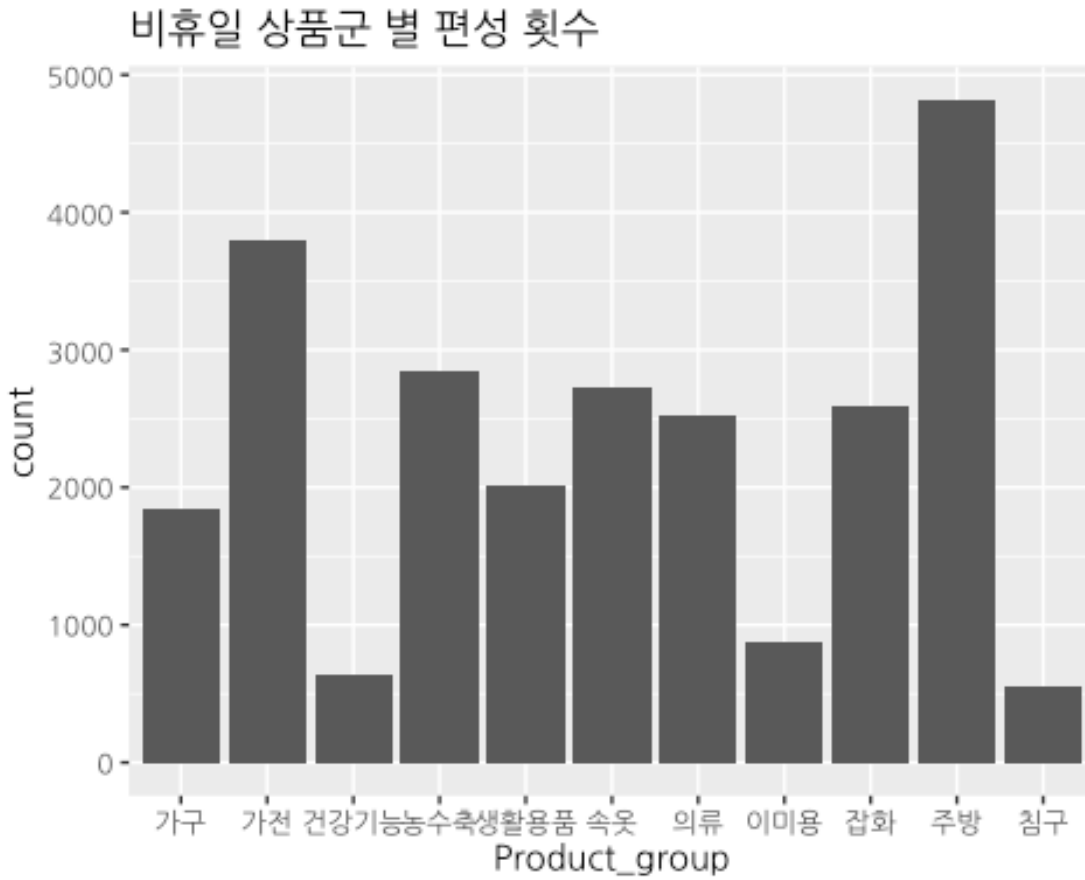



```
ggplot(data=holiday,aes(x=time, y=total_sales,group=Product_group,colour = Product_group))+
  geom_col()+
  ggtitle("휴일 시간대별 총 매출액")
```

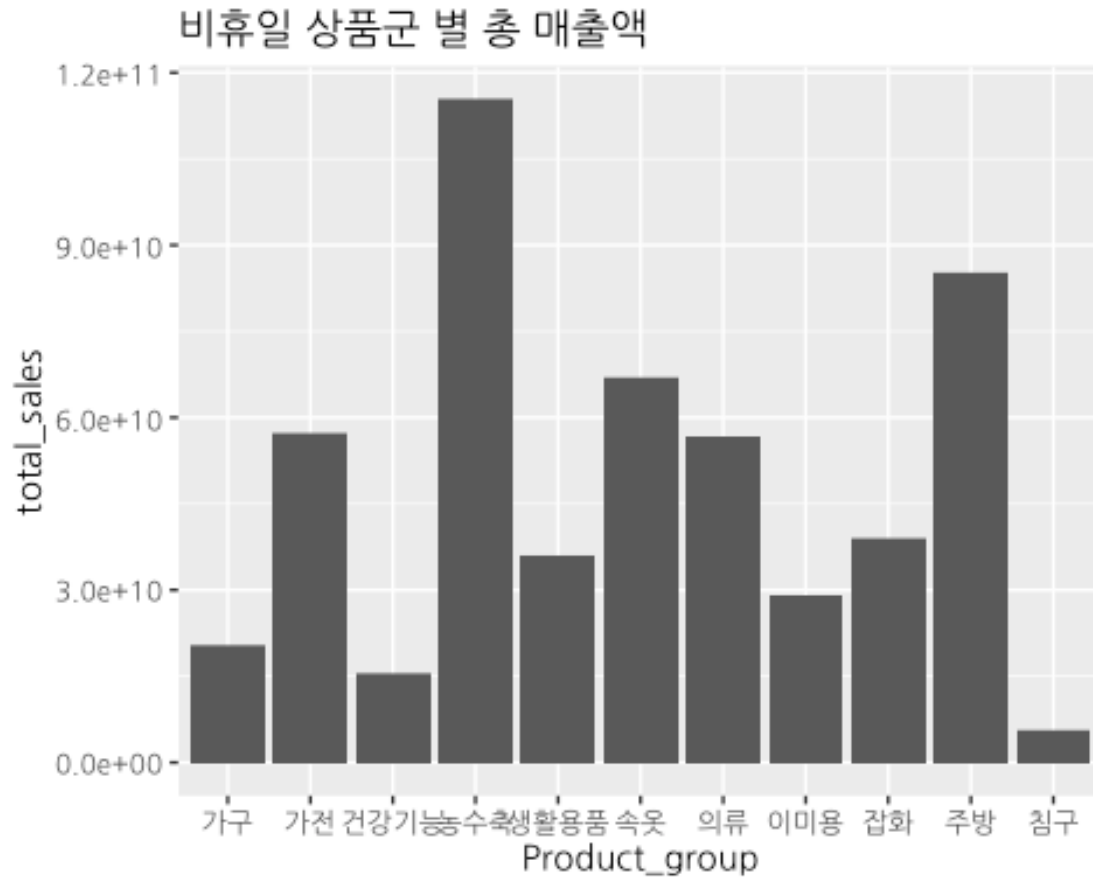


#비휴일의 제품군별 매출액 차이는 얼마나 나는가? #비휴일의 시간대별 매출액 추이?

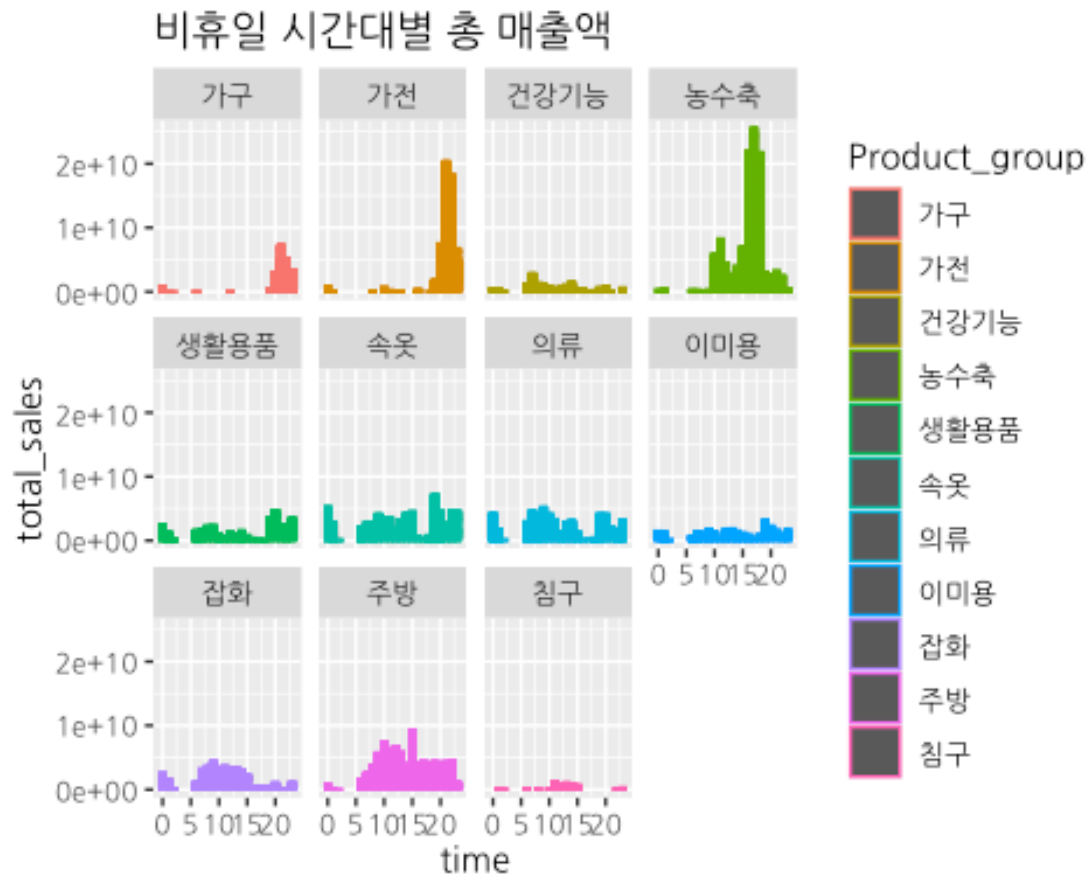
```
Nholiday<-shop_data%>% filter(holiday==0)
ggplot(data=Nholiday)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("비휴일 상품군 별 편성 횟수")
```



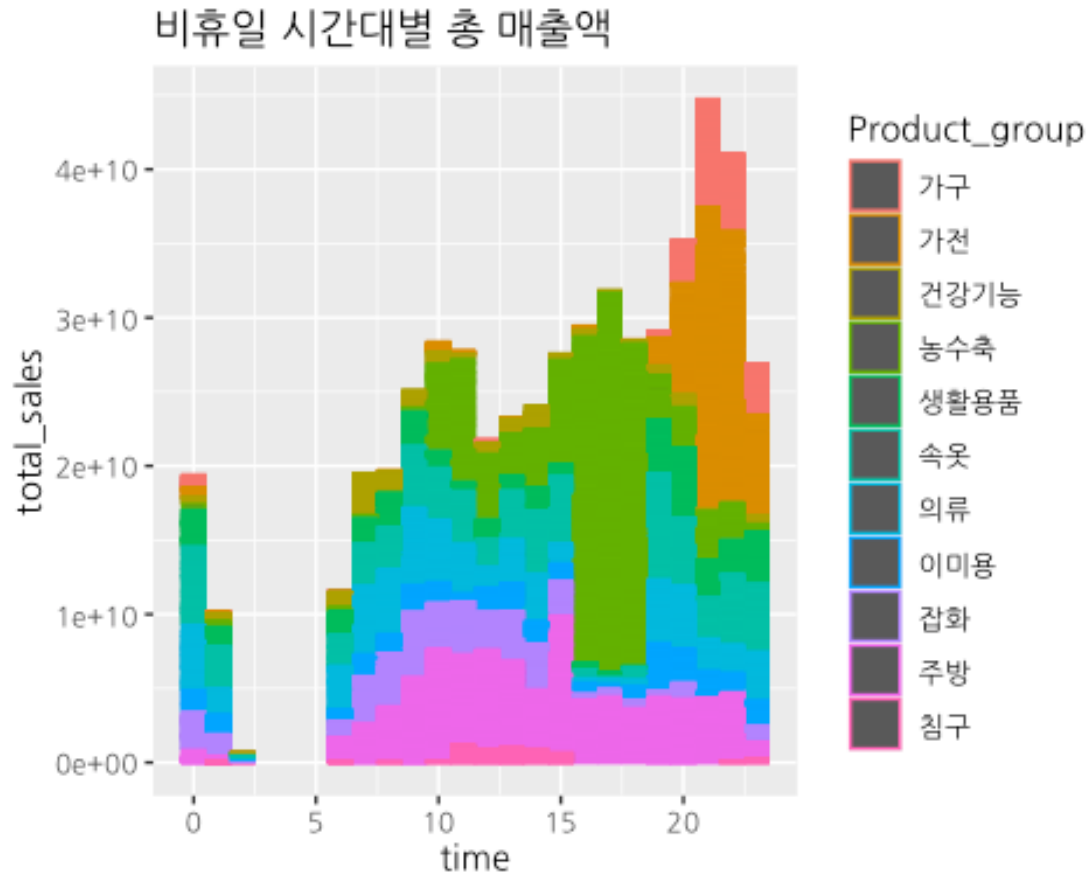
```
ggplot(data=Nholiday, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("비휴일 상품군 별 총 매출액")
```



```
ggplot(data=Nholiday,aes(x=time, y=total_sales,group=Product_group,colour = Product_group))+
  geom_col()+
  ggtitle("비휴일 시간대별 총 매출액")+
  facet_wrap(~ Product_group)
```



```
ggplot(data=Nholiday,aes(x=time, y=total_sales,group=Product_group,colour = Product_group))+
  geom_col()+
  ggtitle("비휴일 시간대별 총 매출액")
```



```
time_holiday_counts <- shop_data %>%
  group_by(time, Product_group, holiday) %>%
  tally
ggplot(data = time_holiday_counts, aes(x = time, y = n, color = holiday, group = holiday)) +
  geom_line() +
  facet_wrap(~ Product_group) +
  ggtitle("휴일/비휴일 상품군 별 편성 횟수 비교")
```

휴일/비휴일 상품군 별 편성 횟수 비교

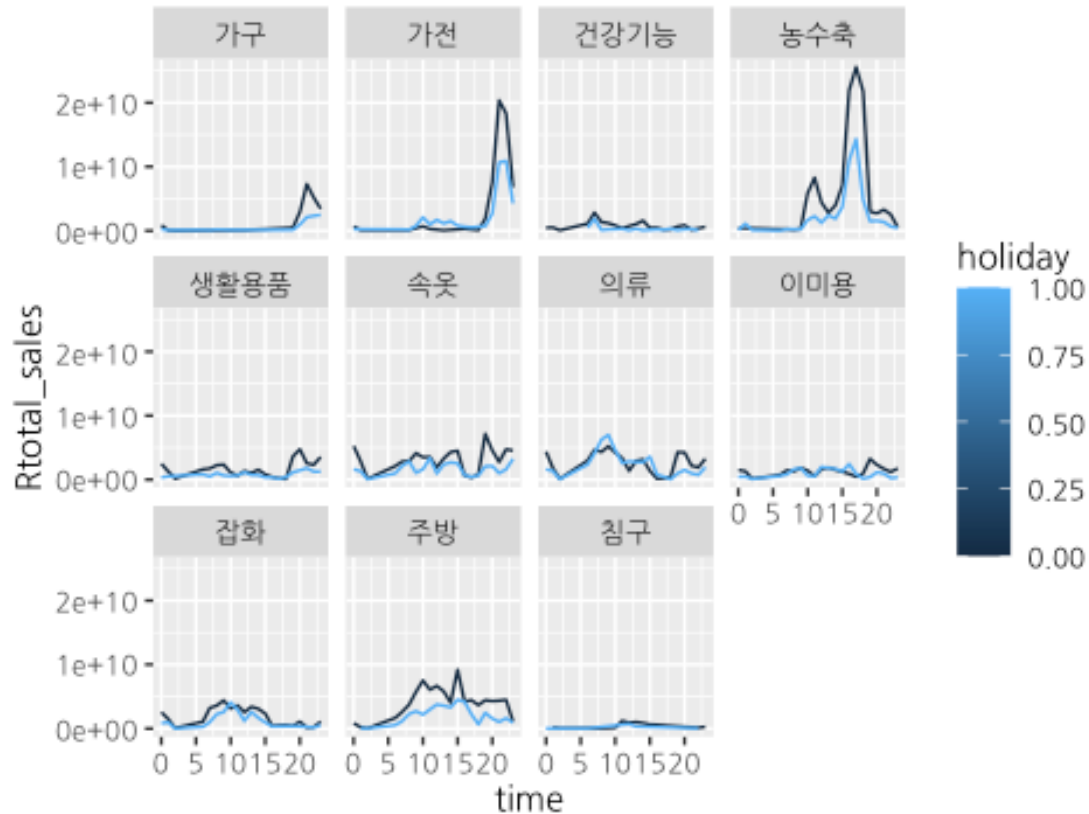


```
time_holiday_total <- shop_data %>%
  group_by(time, Product_group, holiday) %>%
  summarize(Rtotal_sales = sum(total_sales))

## `summarise()` regrouping output by 'time', 'Product_group' (override with
## `.groups` argument)

ggplot(data = time_holiday_total, aes(x = time, y = Rtotal_sales, color = holiday, group = holiday)) +
  geom_line() +
  facet_wrap(~ Product_group) +
  ggtitle("휴일/비휴일 상품군 별 매출액 비교")
```

휴일/비휴일 상품군 별 매출액 비교

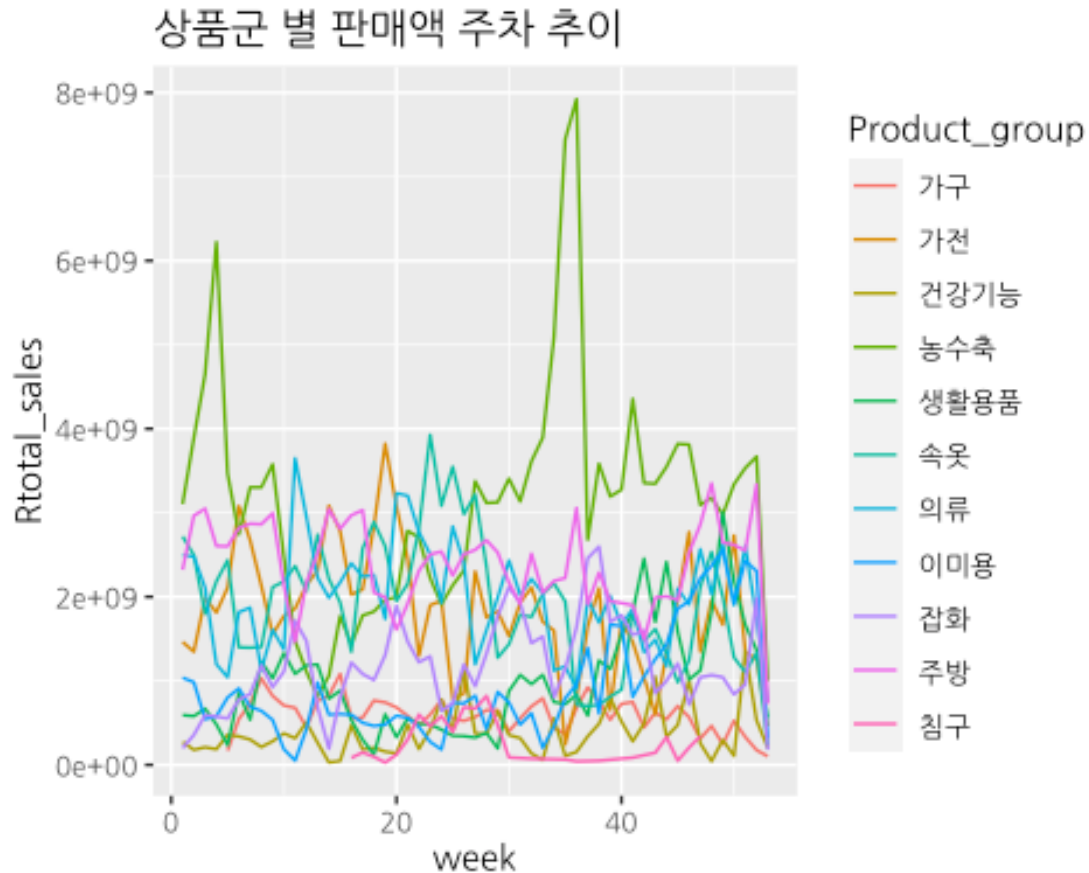


##주차별

```
byweek<-shop_data %>% group_by(week, Product_group) %>%
  summarize(Rtotal_sales=sum(total_sales))

## `summarise()` regrouping output by 'week' (override with `.groups` argumen
t)

ggplot(data=byweek, aes(x=week, y=Rtotal_sales, group=Product_group, color=Pr
oduct_group))+
  geom_line()+
  ggtitle("상품군 별 판매액 주차 추이")
```

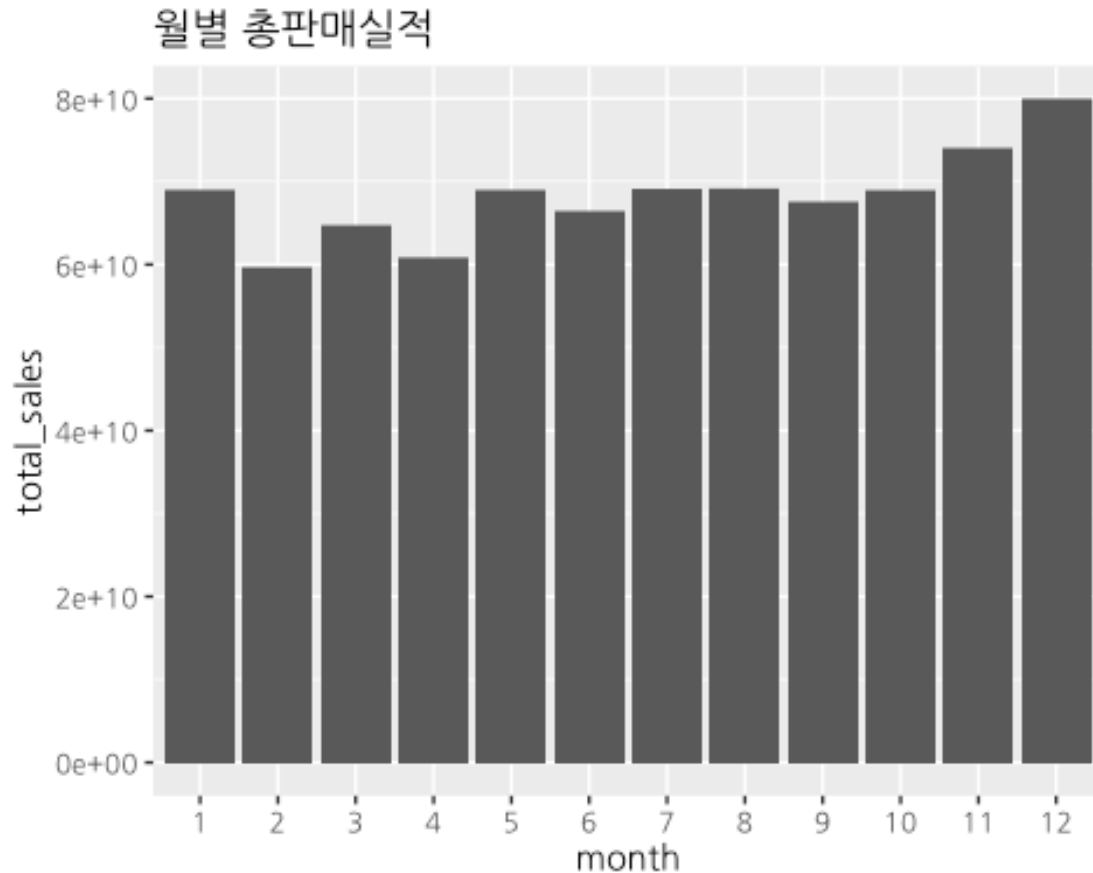



나윤

#월별로 그래프를 그려보려고 한다. #그런데 월별을 일단 character 으로 바꿔줘보자
#월별로 총 판매 실적을 보고자 한다.

```
shop_data$month<-as.character(shop_data$month)

ggplot(data=shop_data, aes(x=month, y=total_sales))+
  geom_col()+
  scale_x_discrete(limits=c("1","2","3","4","5","6","7","8","9","10","11","12"))+
  ggtitle("월별 총판매실적")
```



#12 월의 총판매액이 가장 많고 #2 월과 4 월의 판매액이 가장 적다 #월별로 상품군에 따라 편성 횟수를 각각 보고자 한다.

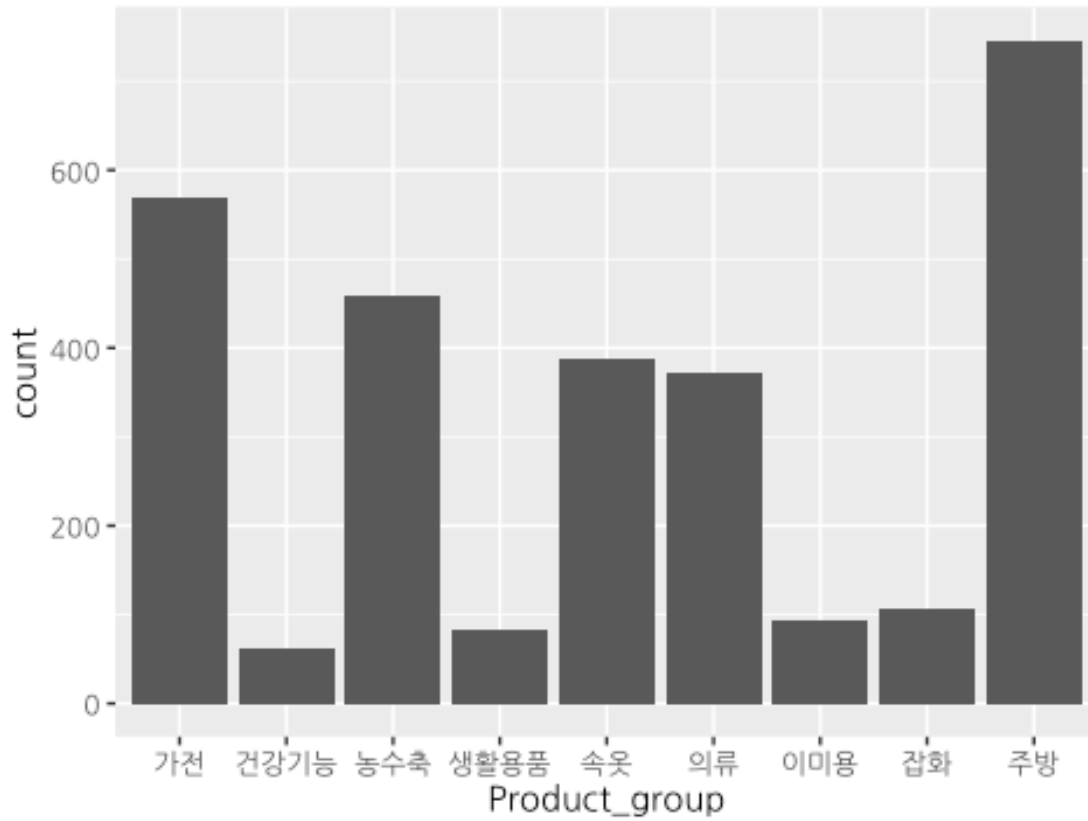
그 전에 먼저 상품군의 숫자가 몇개인지 보고자 한다.

```
summary(shop_data$Product_group)
```

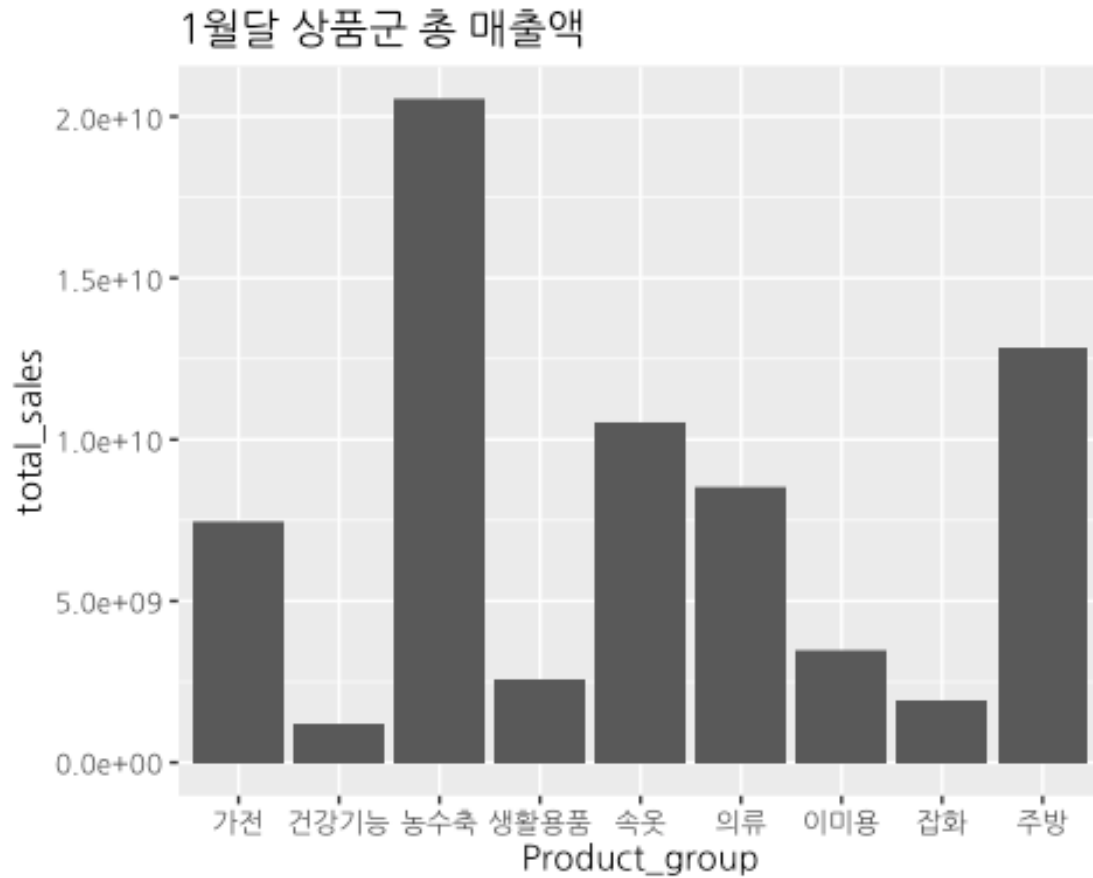
```
##      가구      가전 건강기능  농수축 생활용품  속옷  의류  이미용
##      2614      6479      790      3884      2877      3919      4340      1305
##      잡화      주방      침구
##      3891      6604      669
```

```
month1<-shop_data %>% filter(month=="1")
ggplot(data=month1)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("1 월달 상품군 편성 횟수")
```

1월달 상품군 편성 횟수



```
ggplot(data=month1, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("1 월달 상품군 총 매출액")
```



```
summary(month1$Product_group)
```

```
##      가구      가전 건강기능      농수축 생활용품      속옷      의류      이미용
##      0      570      63      458      83      389      372      93
##      잡화      주방      침구
##      107      746      0
```

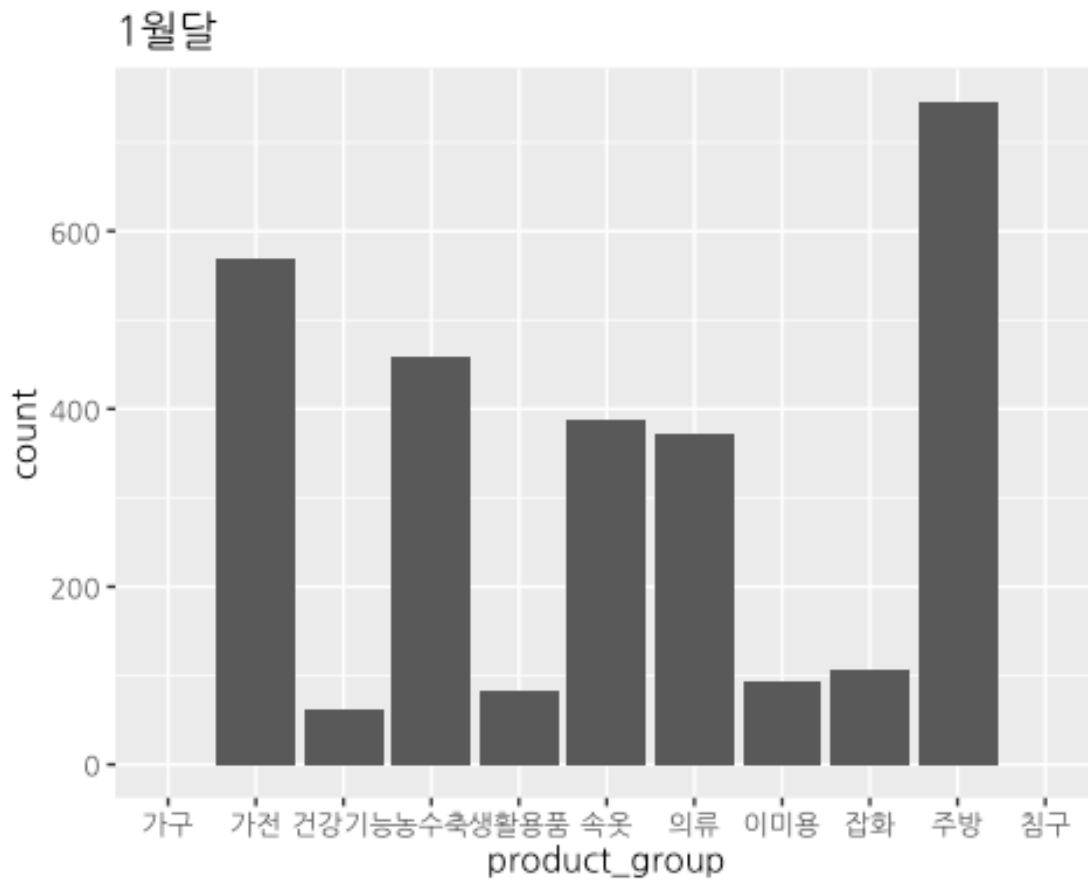
```
count<-as.data.frame(table(month1$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

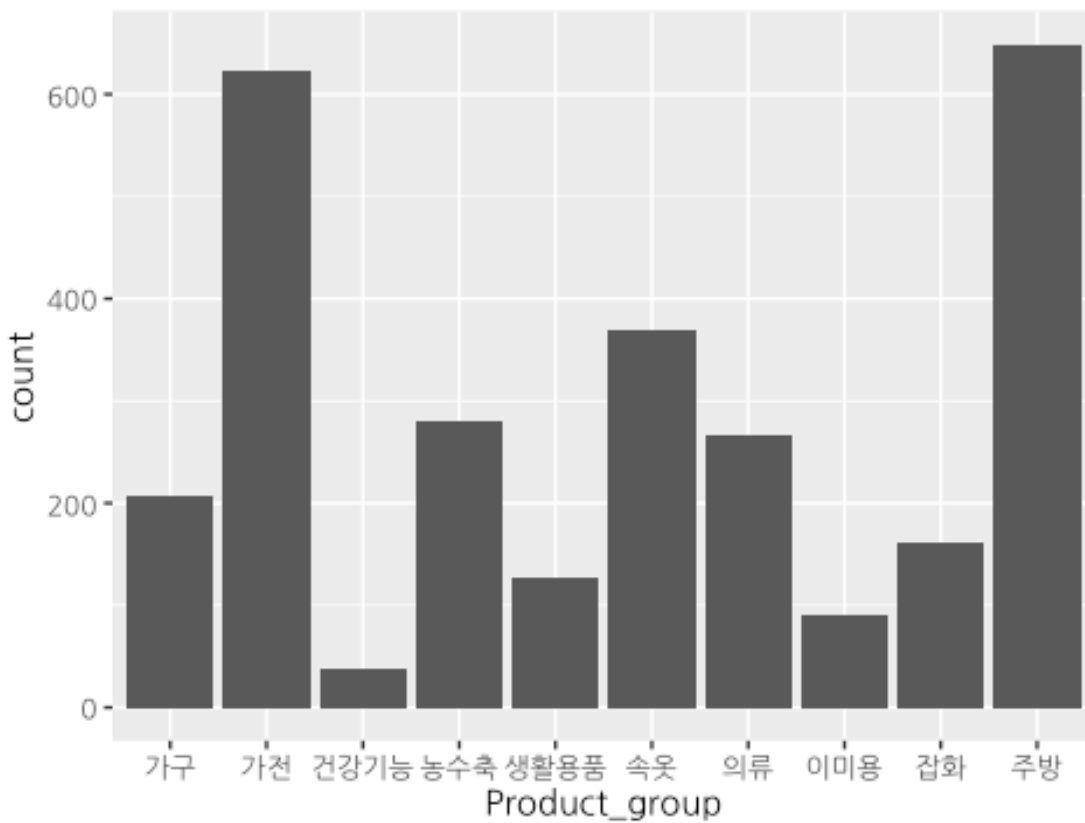
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("1 월
달")+geom_col(position = "stack")
```

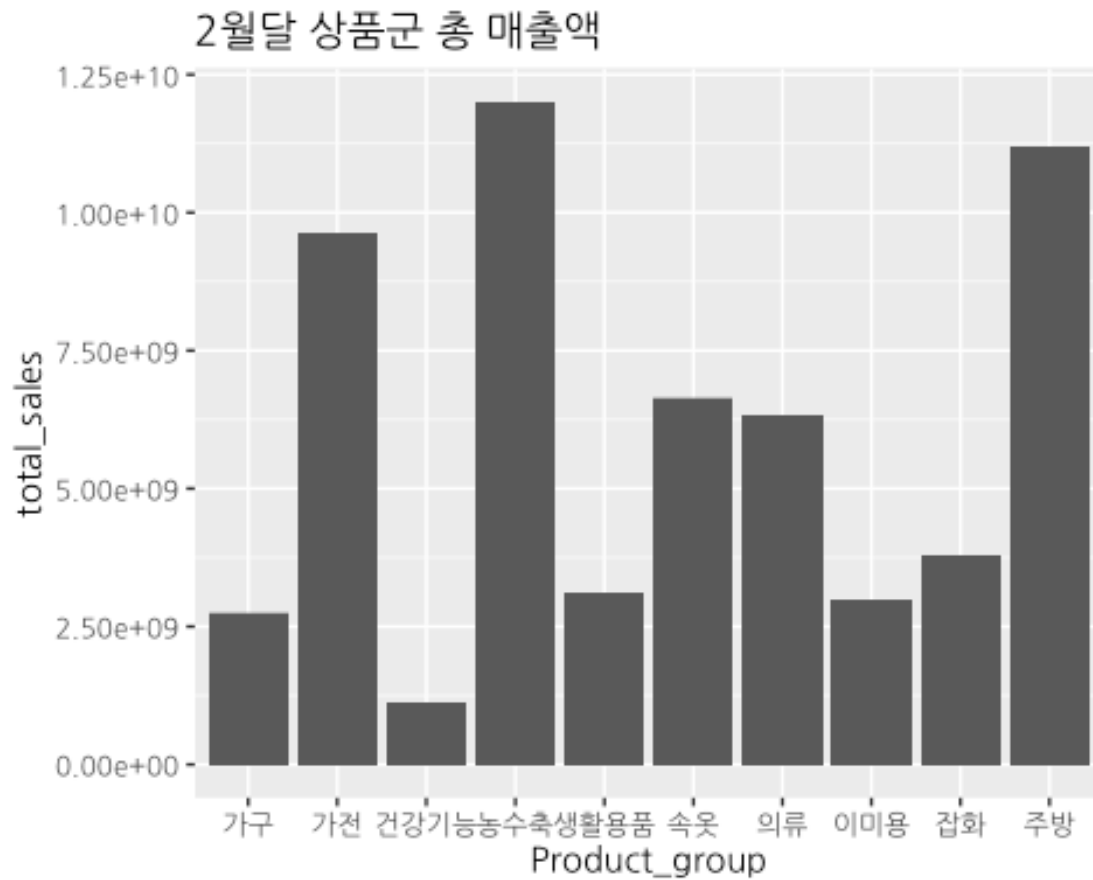


```
month2<-shop_data %>% filter(month=="2")
ggplot(data=month2)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("2 월달 상품군 편성 횟수")
```

2월달 상품군 편성 횟수



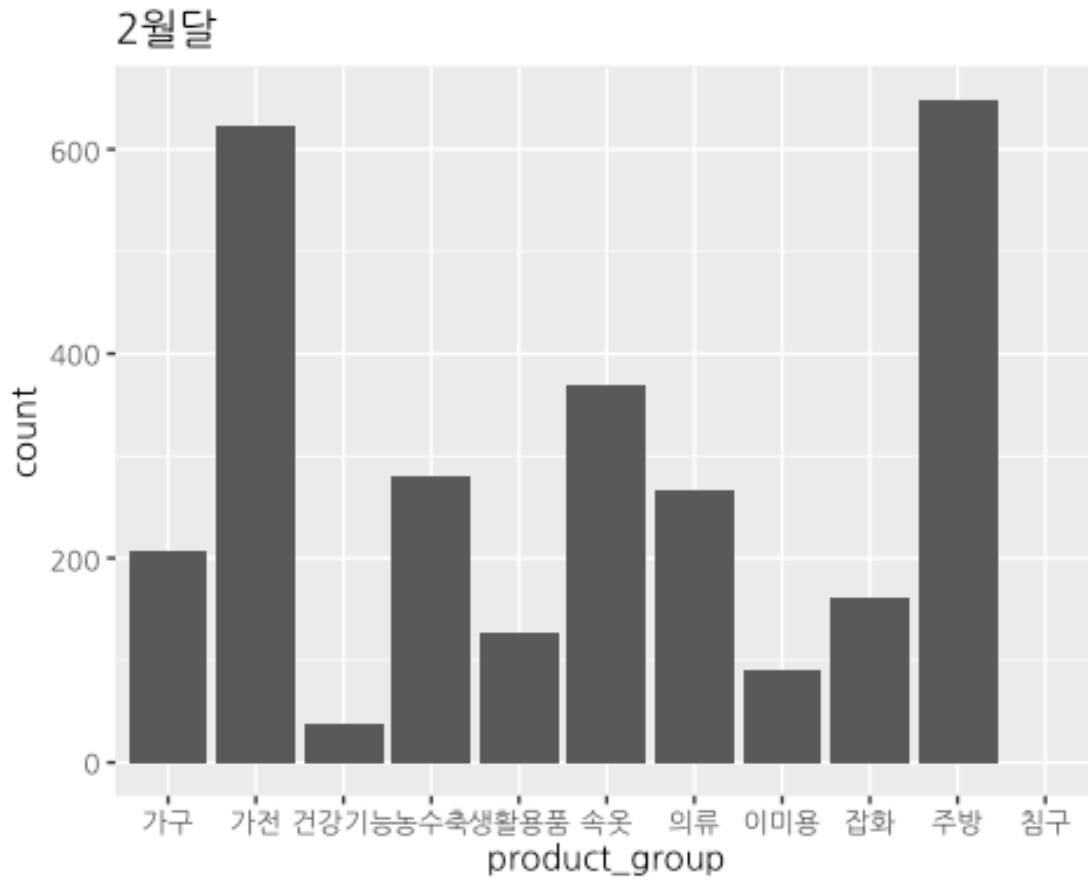
```
ggplot(data=month2, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("2 월달 상품군 총 매출액")
```



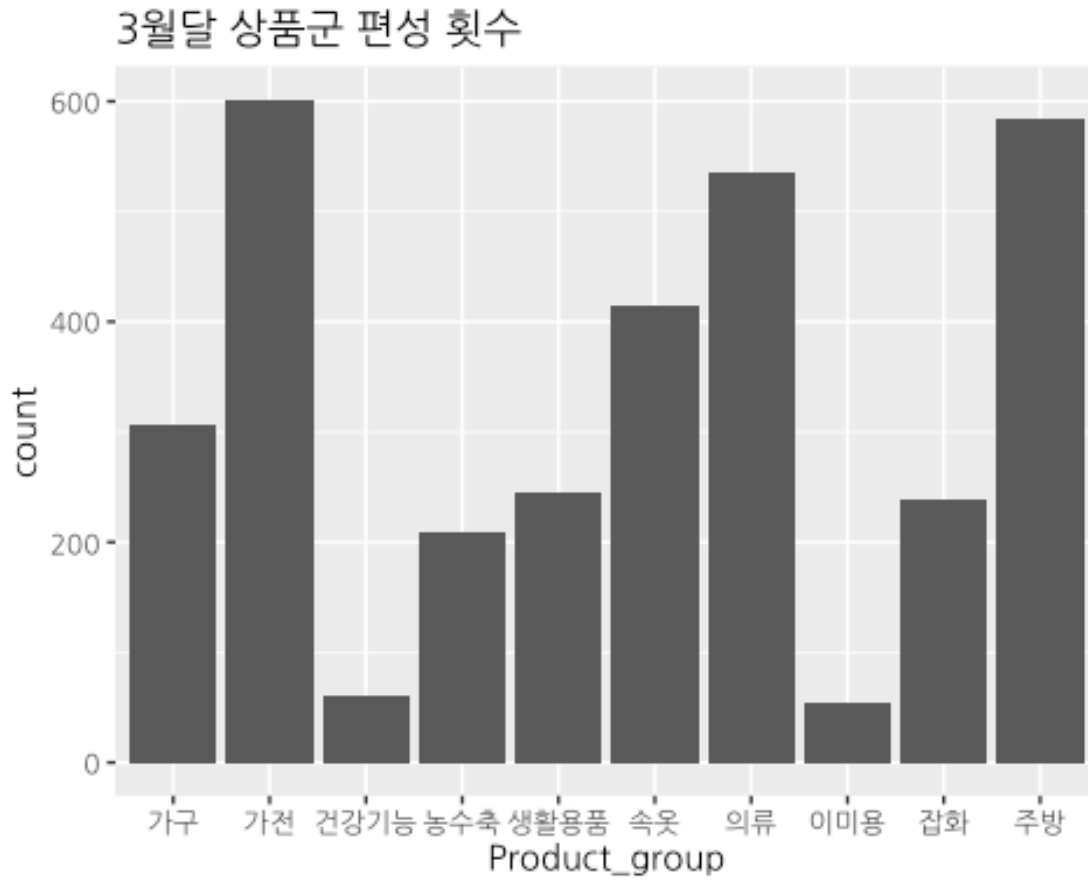
```
count<-as.data.frame(table(month2$Product_group))

library(ggplot2)
count$Freq<-as.numeric(count$Freq)
count$Var1<-as.character(count$Var1)

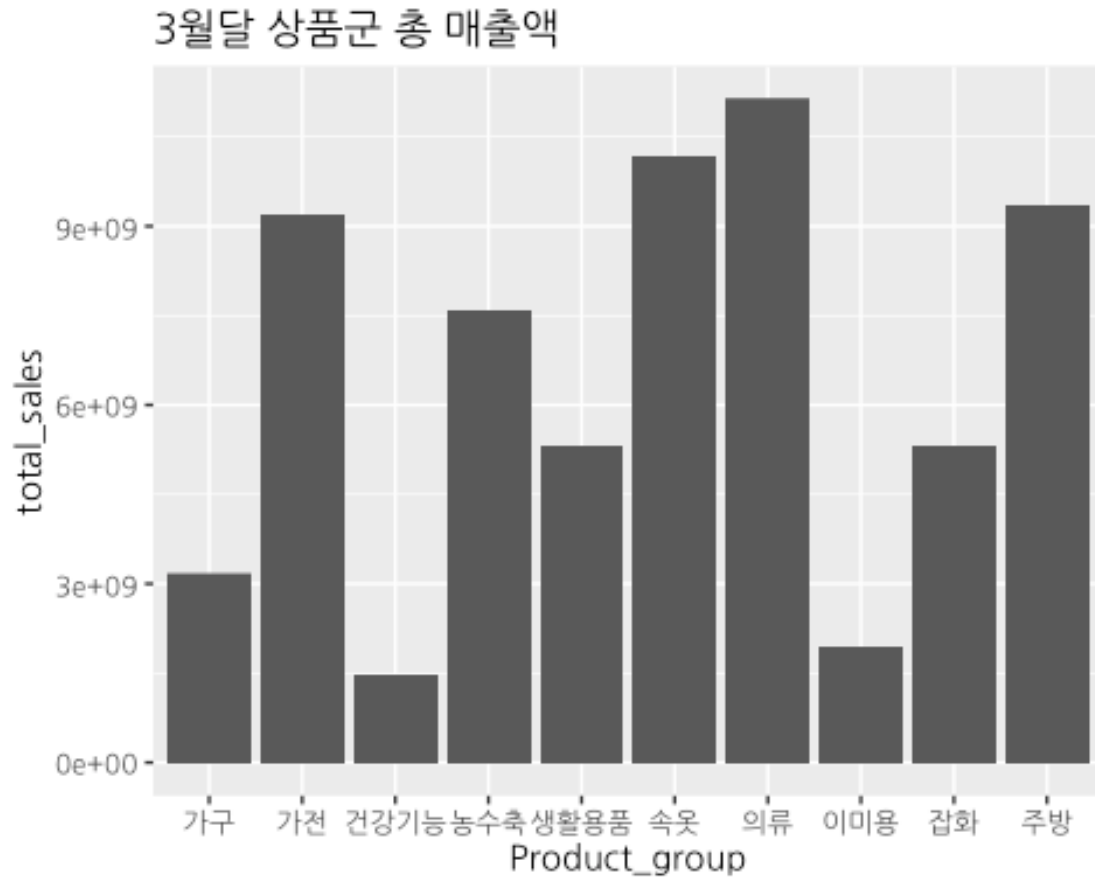
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("2 월
달")+geom_col(position = "stack")
```



```
month3<-shop_data %>% filter(month=="3")
ggplot(data=month3)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("3 월달 상품군 편성 횟수")
```

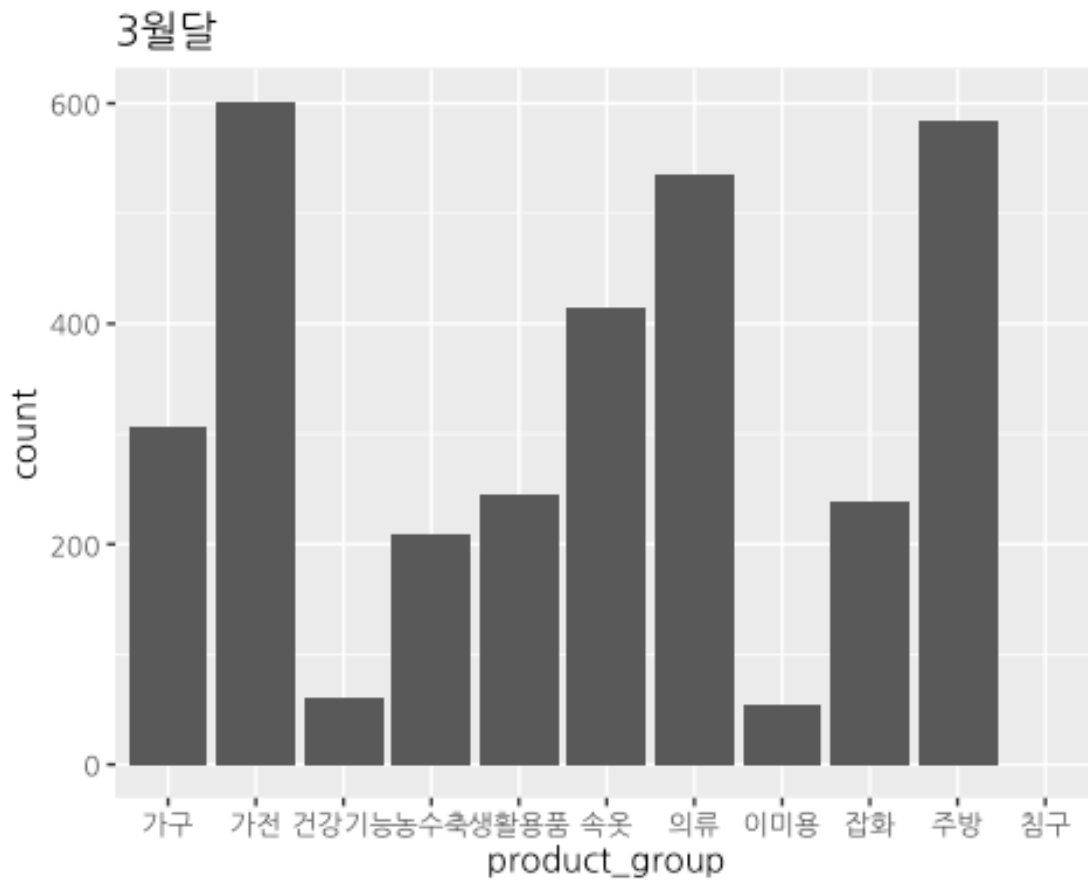
```
ggplot(data=month3, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("3 월달 상품군 총 매출액")
```



```
count<-as.data.frame(table(month3$Product_group))

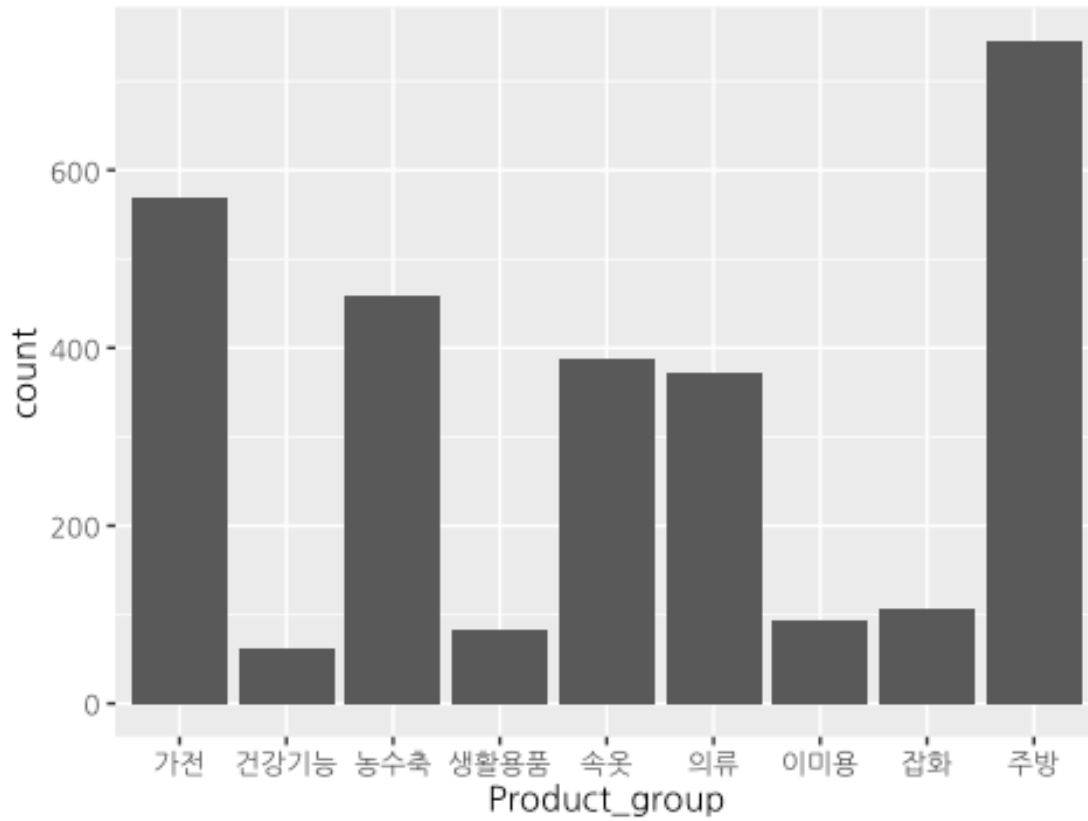
library(ggplot2)
count$Freq<-as.numeric(count$Freq)
count$Var1<-as.character(count$Var1)

ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("3 월
달")+geom_col(position = "stack")
```

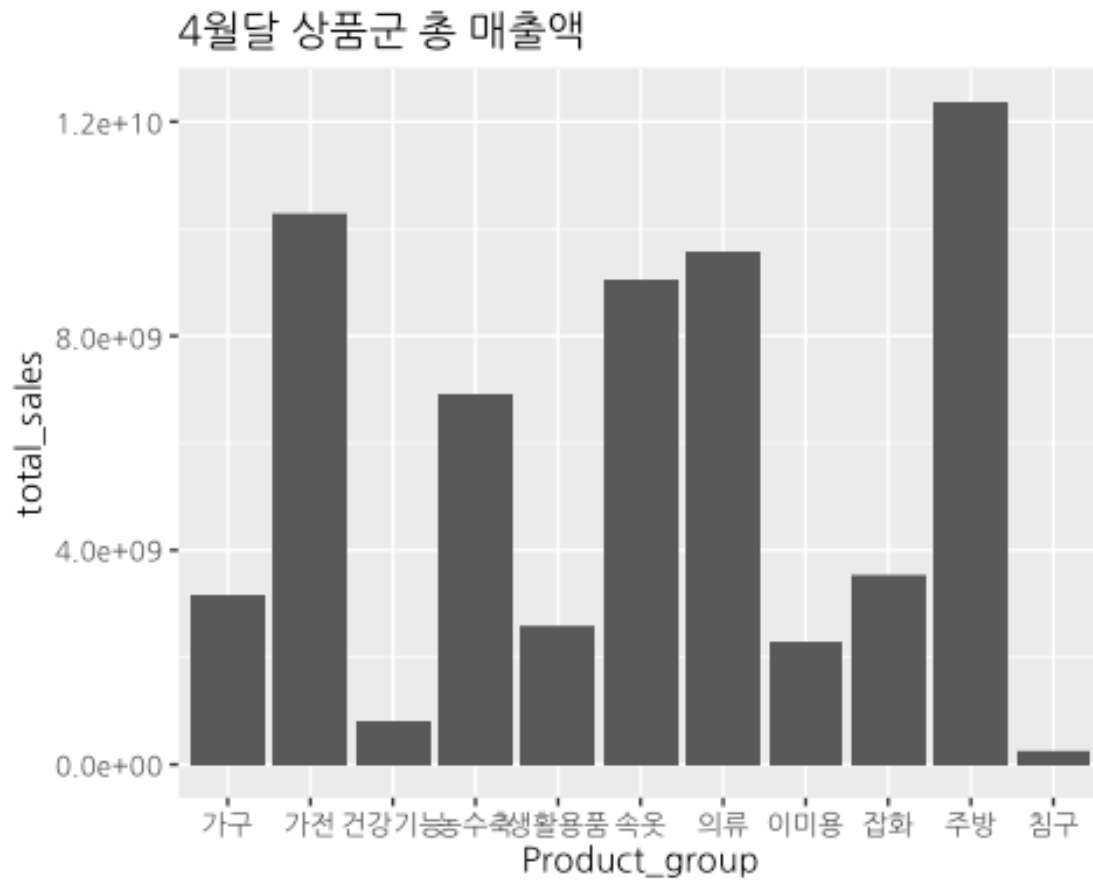


```
month4<-shop_data %>% filter(month=="4")
ggplot(data=month1)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("1 월달 상품군 편성 횟수")
```

1월달 상품군 편성 횟수



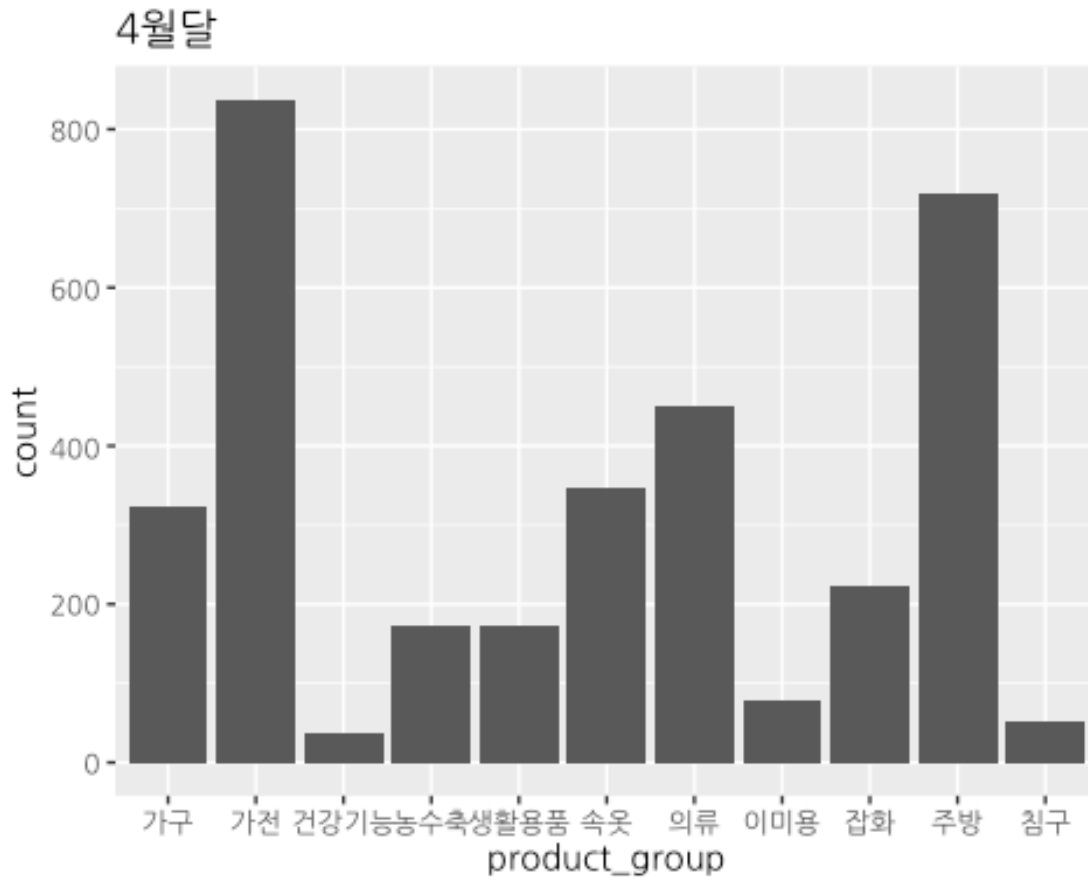
```
ggplot(data=month4, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("4 월달 상품군 총 매출액")
```



```
count<-as.data.frame(table(month4$Product_group))

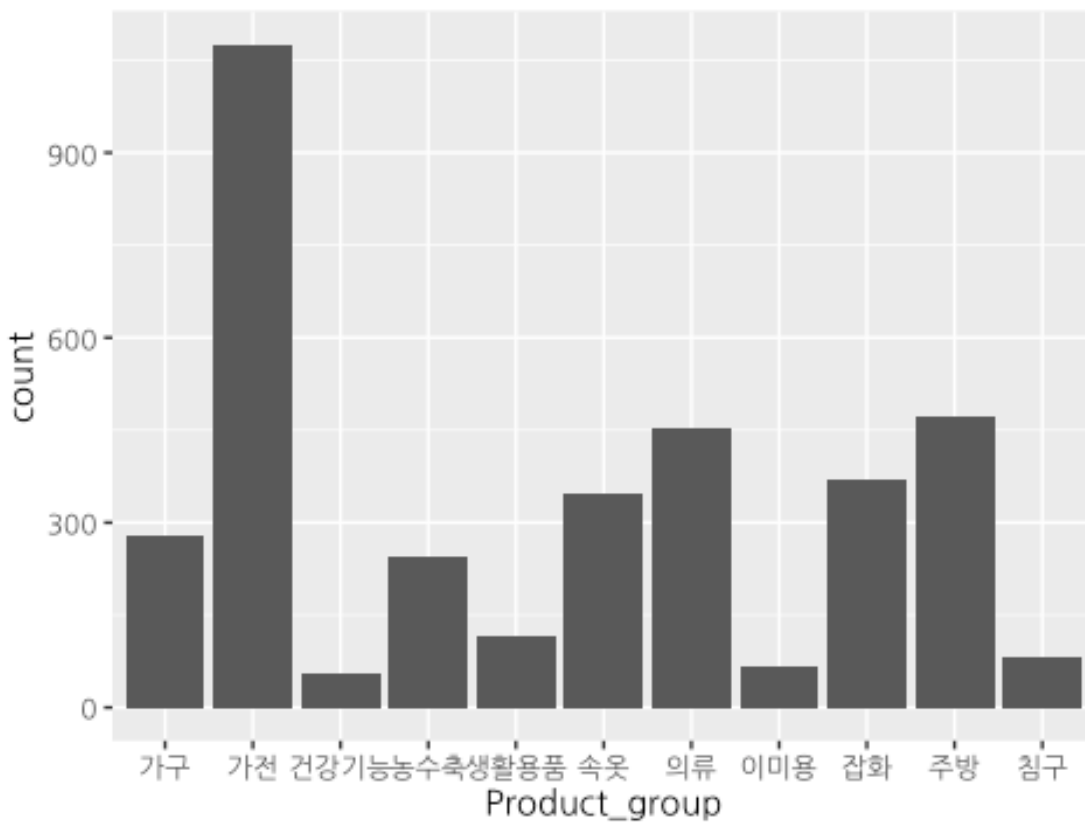
library(ggplot2)
count$Freq<-as.numeric(count$Freq)
count$Var1<-as.character(count$Var1)

ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("4 월
달")+geom_col(position = "stack")
```

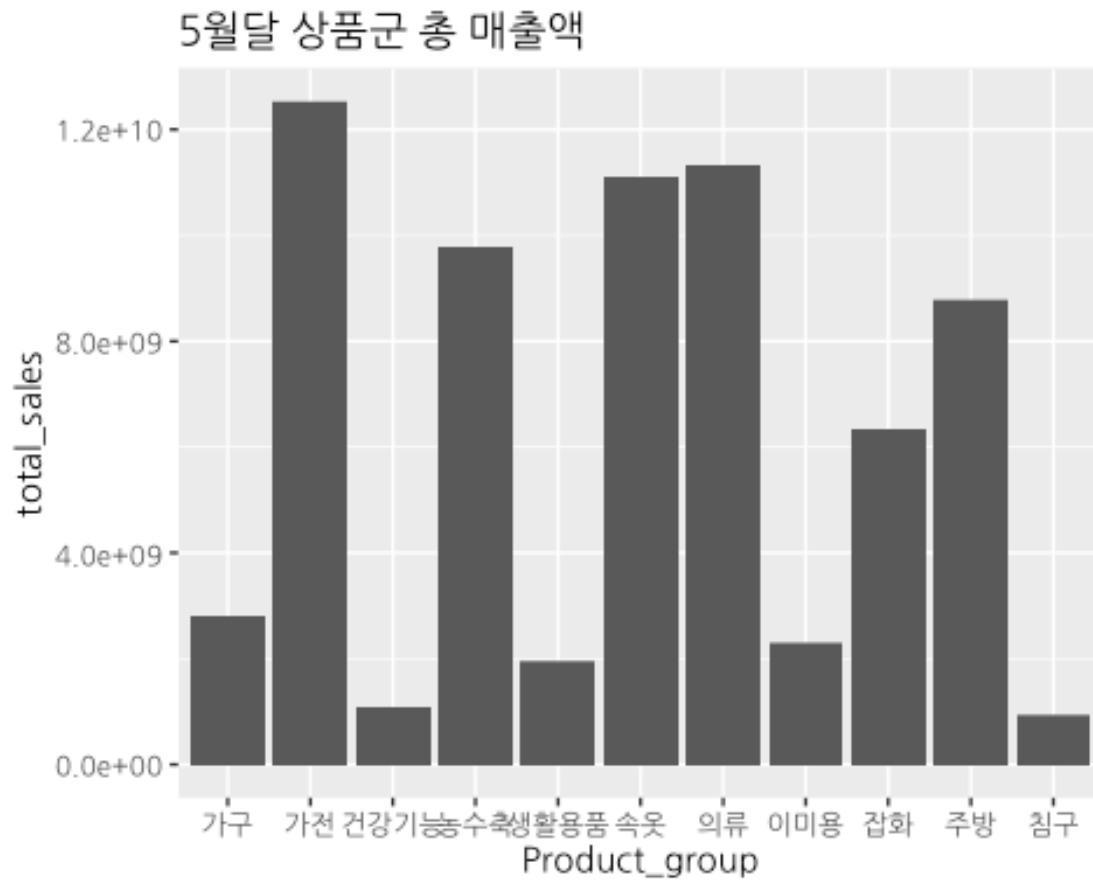


```
month5<-shop_data %>% filter(month=="5")
ggplot(data=month5)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("5 월달 상품군 편성 횟수")
```

5월달 상품군 편성 횟수



```
ggplot(data=month5, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("5 월달 상품군 총 매출액")
```



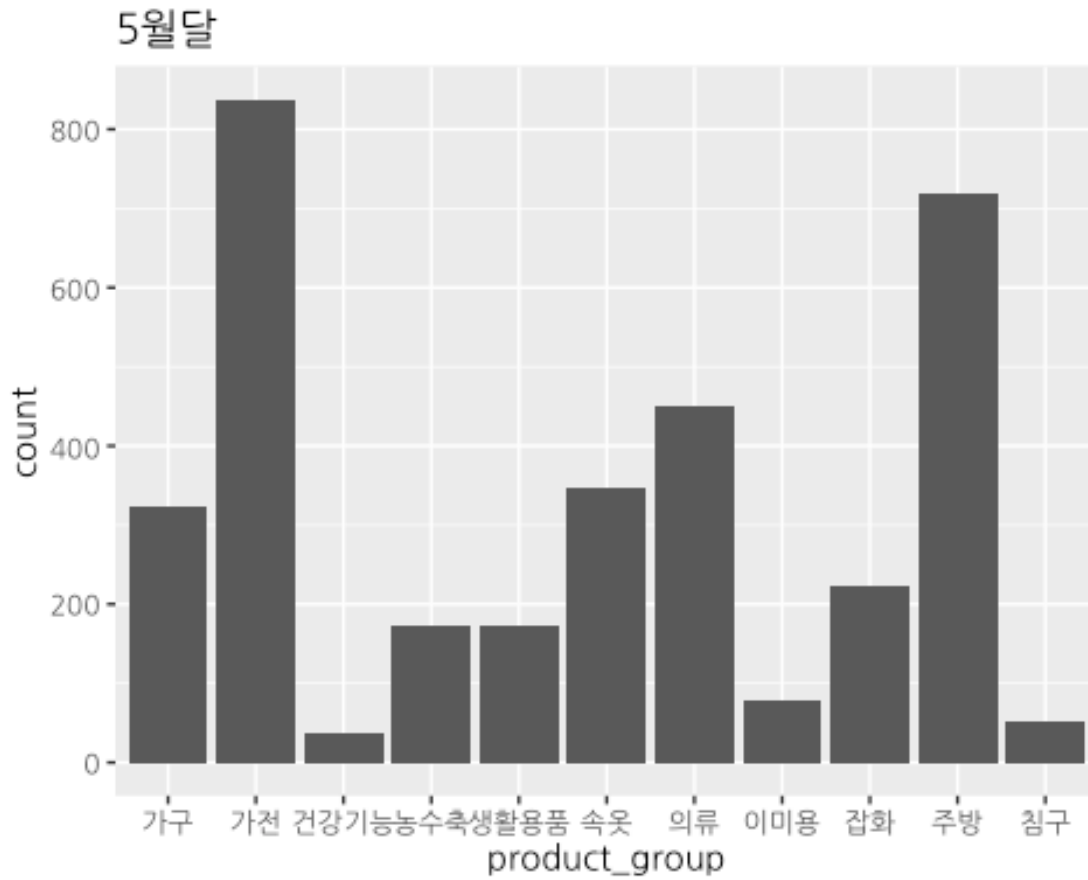
```
count<-as.data.frame(table(month5$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

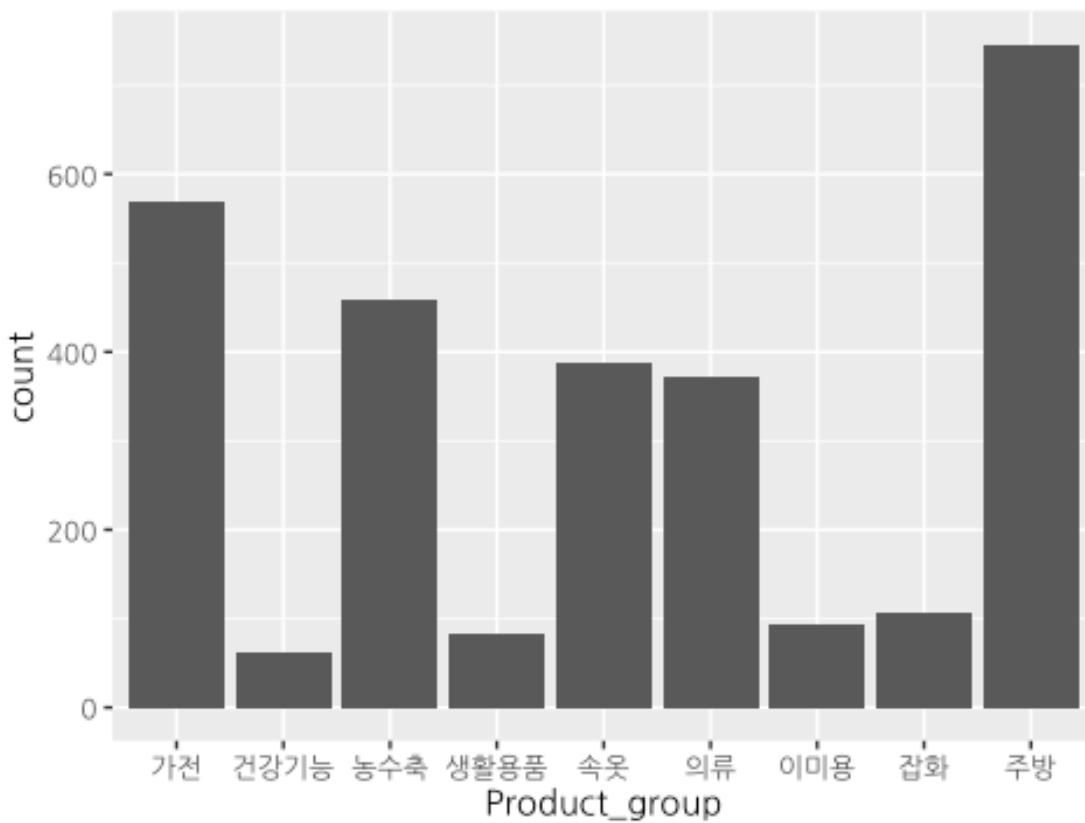
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("5 월  
달")+geom_col(position = "stack")
```

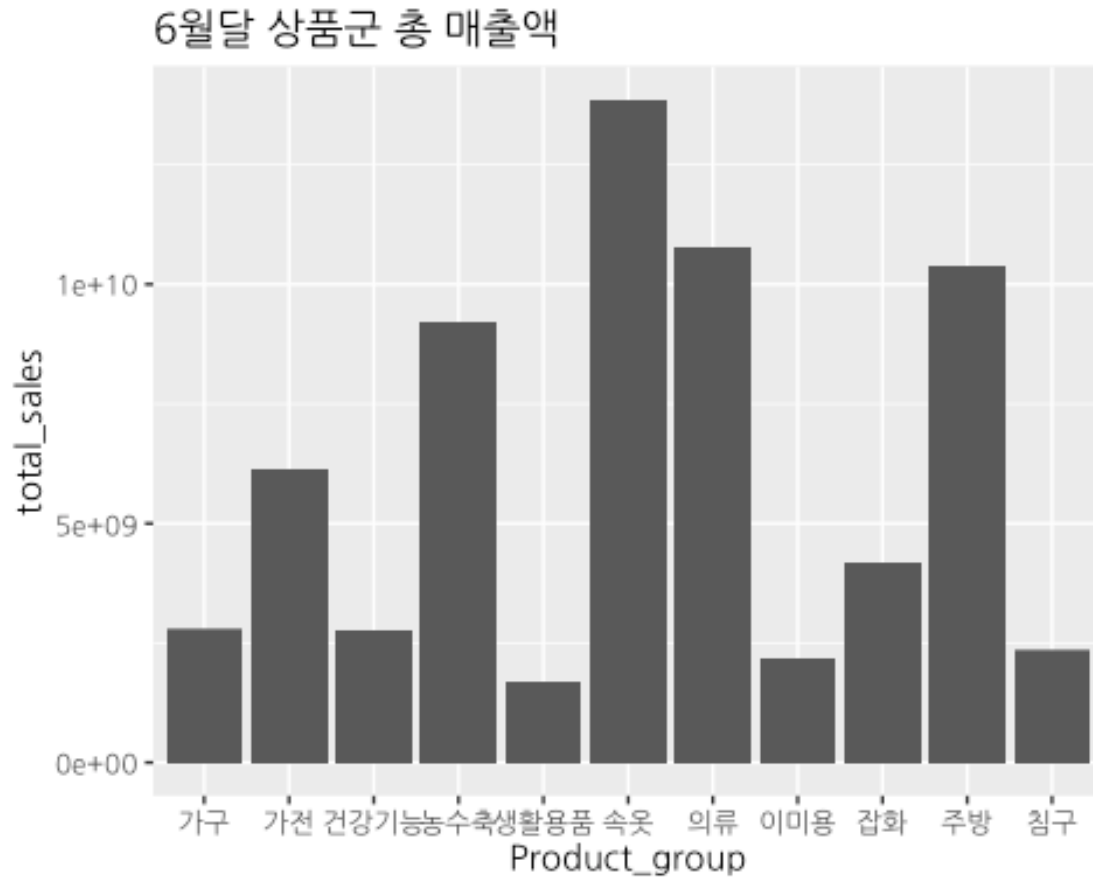



```
month6<-shop_data %>% filter(month=="6")
ggplot(data=month1)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("6 월달 상품군 편성 횟수")
```

6월달 상품군 편성 횟수



```
ggplot(data=month6, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("6 월달 상품군 총 매출액")
```



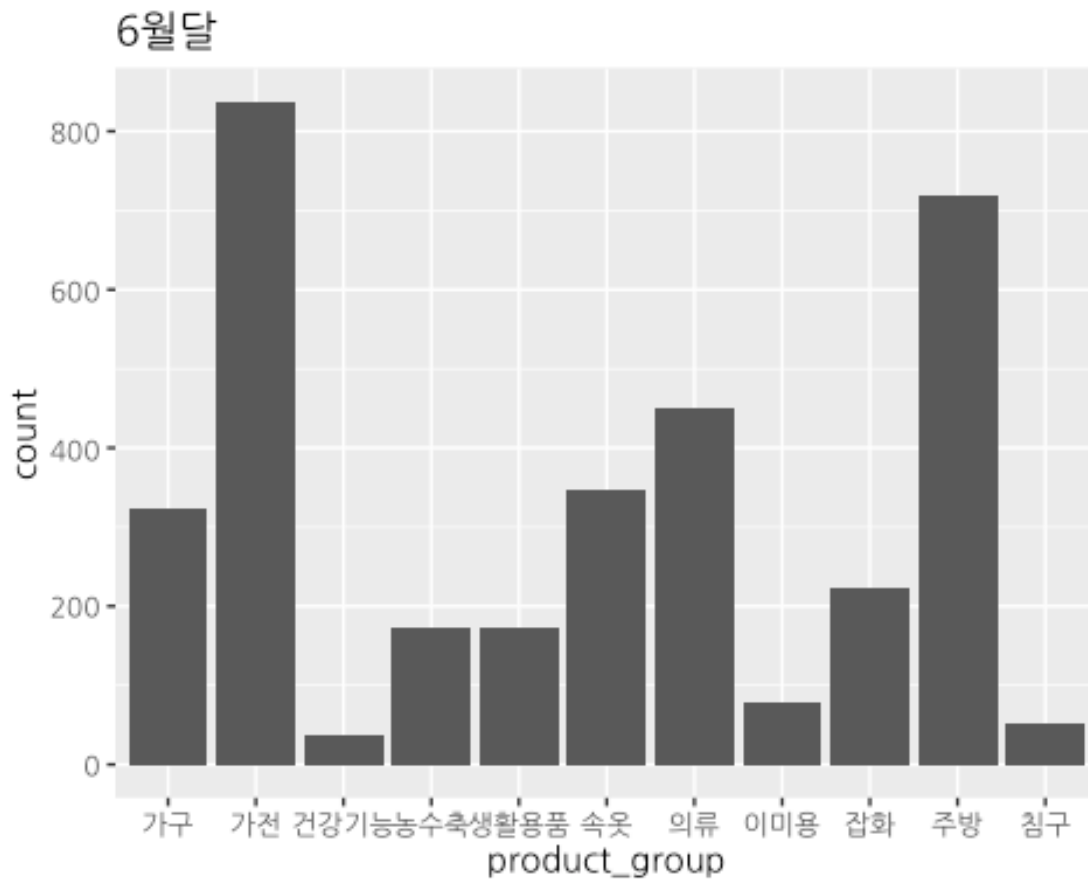
```
count<-as.data.frame(table(month6$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

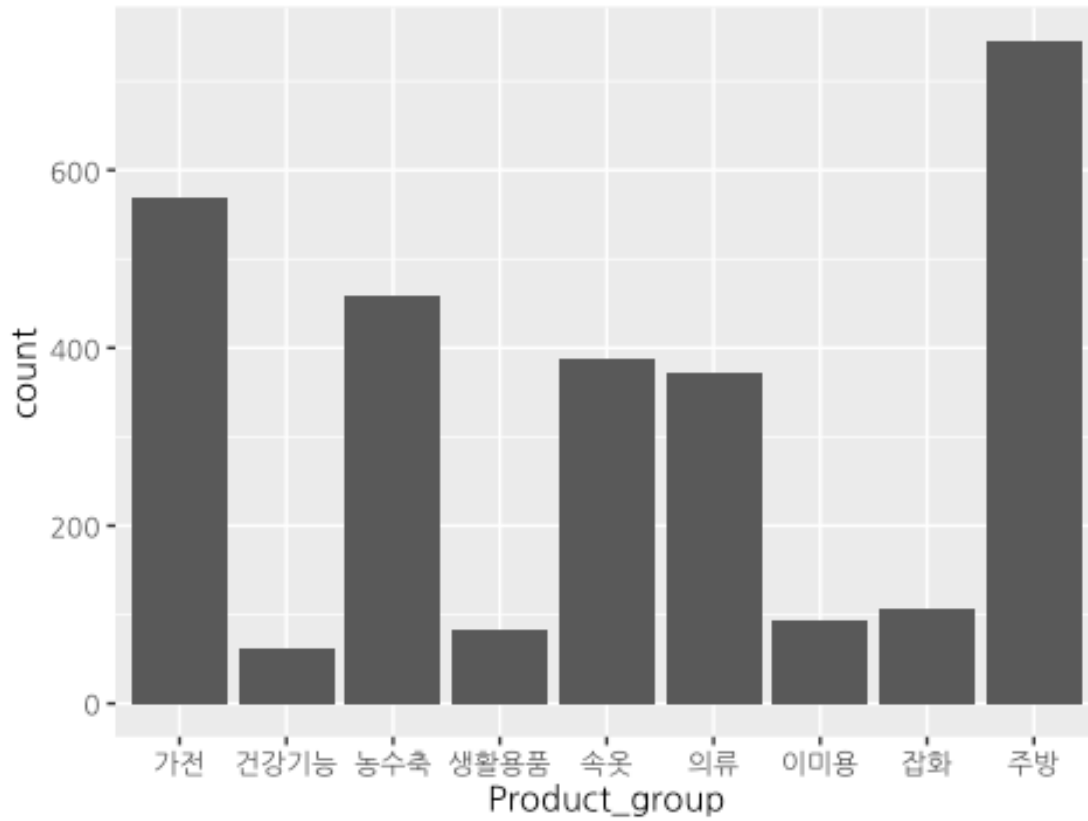
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("6 월  
달")+geom_col(position = "stack")
```

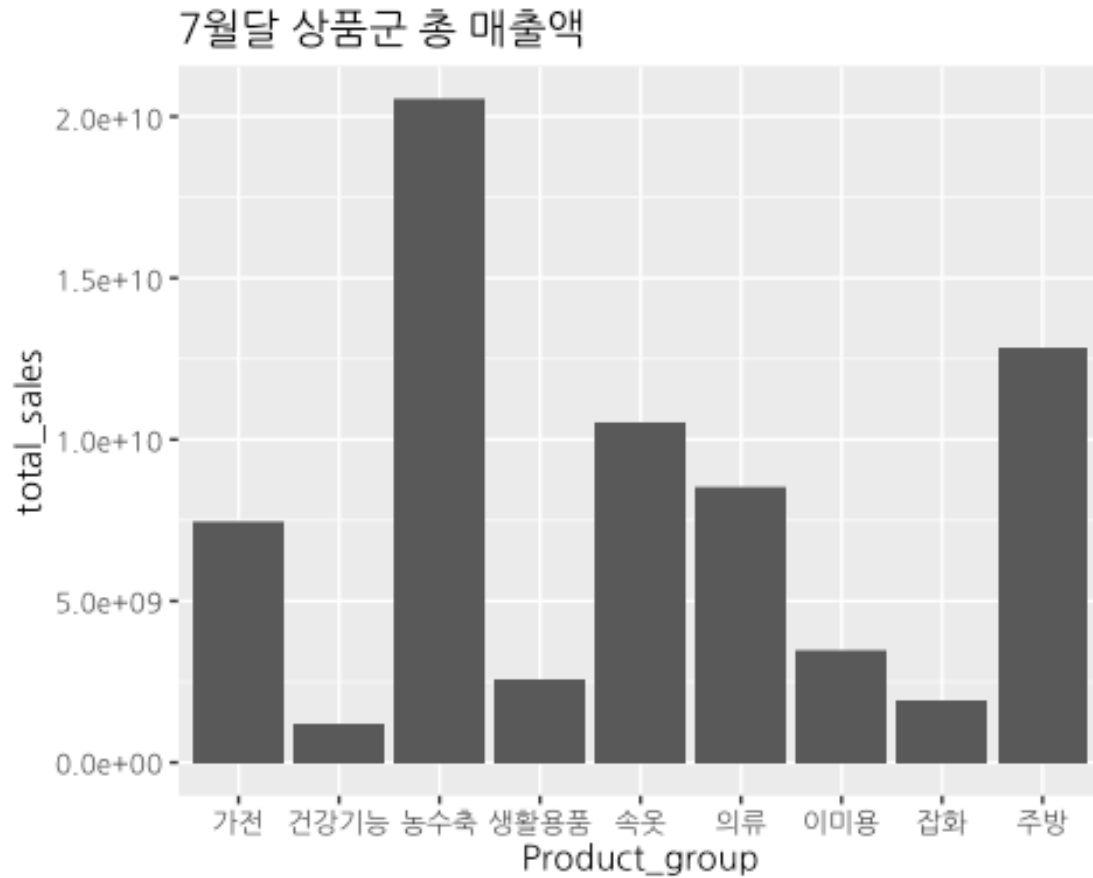


```
month7<-shop_data %>% filter(month=="1")
ggplot(data=month1)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("7 월달 상품군 편성 횟수")
```

7월달 상품군 편성 횟수



```
ggplot(data=month7, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("7 월달 상품군 총 매출액")
```



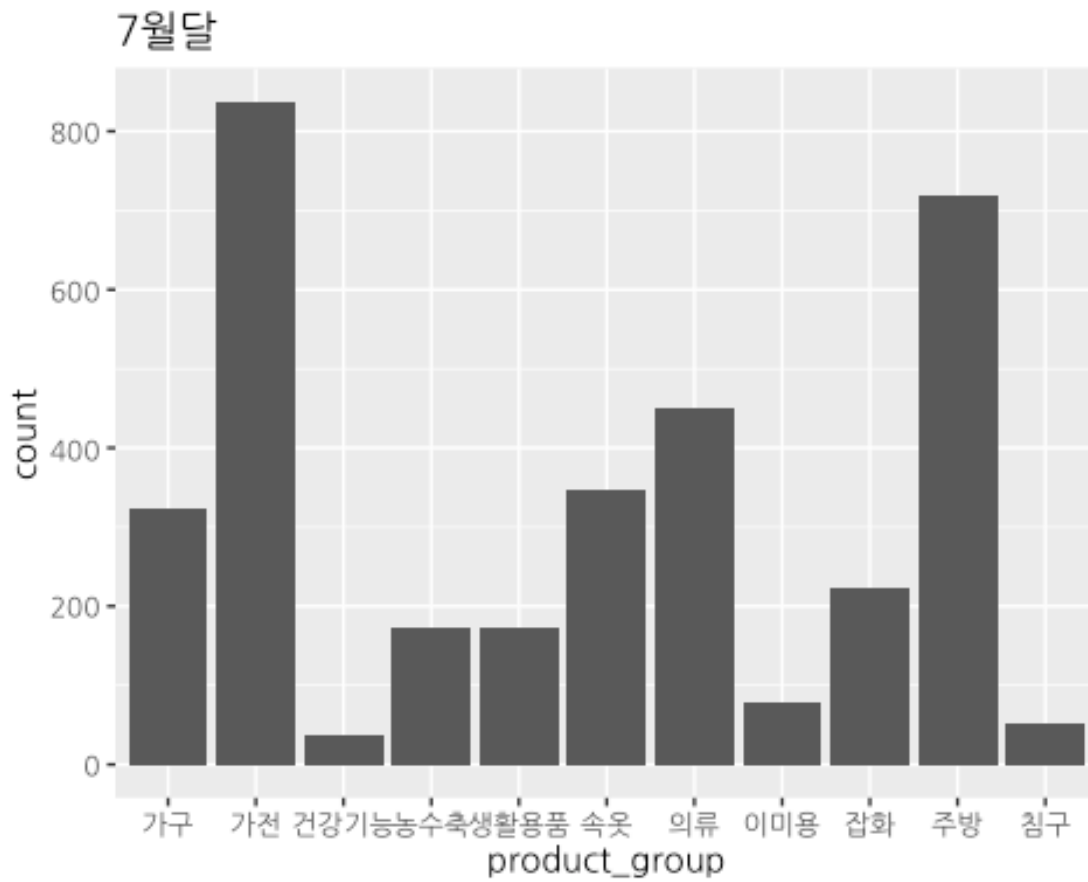
```
count<-as.data.frame(table(month7$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

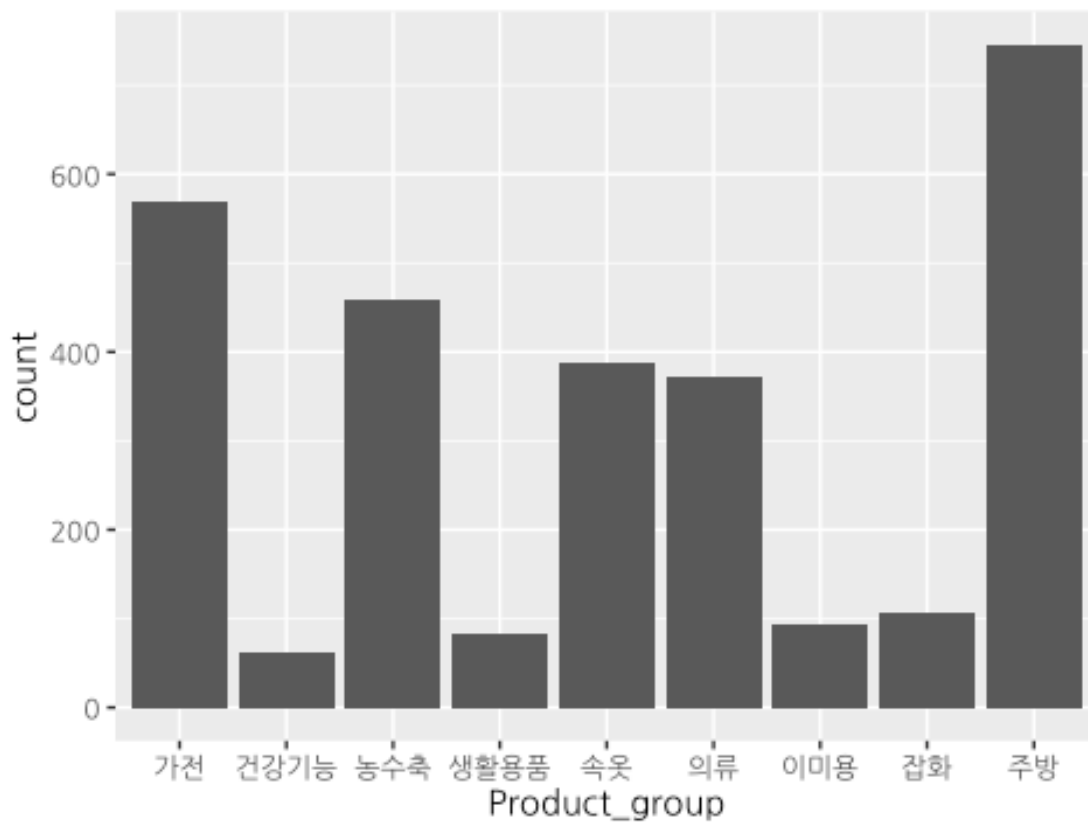
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("7 월  
달")+geom_col(position = "stack")
```

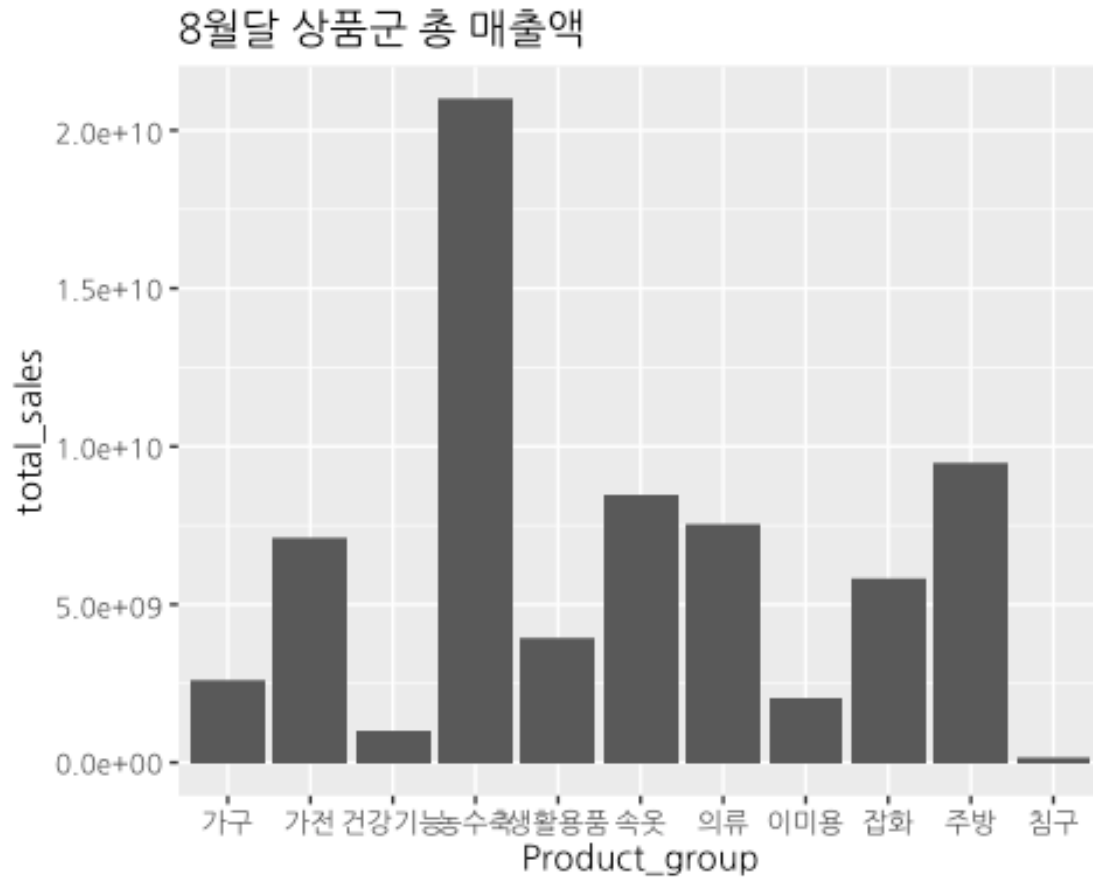


```
month8<-shop_data %>% filter(month=="8")
ggplot(data=month1)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("8 월달 상품군 편성 횟수")
```

8월달 상품군 편성 횟수



```
ggplot(data=month8, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("8 월달 상품군 총 매출액")
```

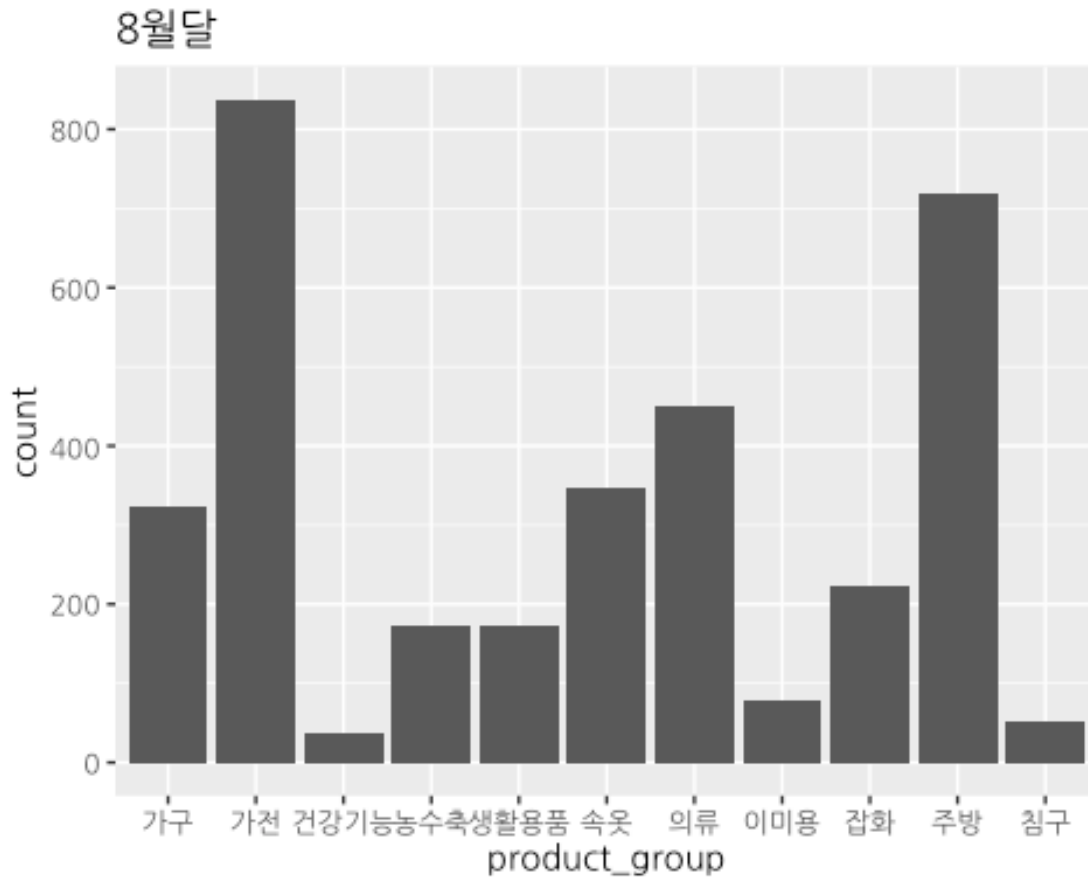
```
count<-as.data.frame(table(month8$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

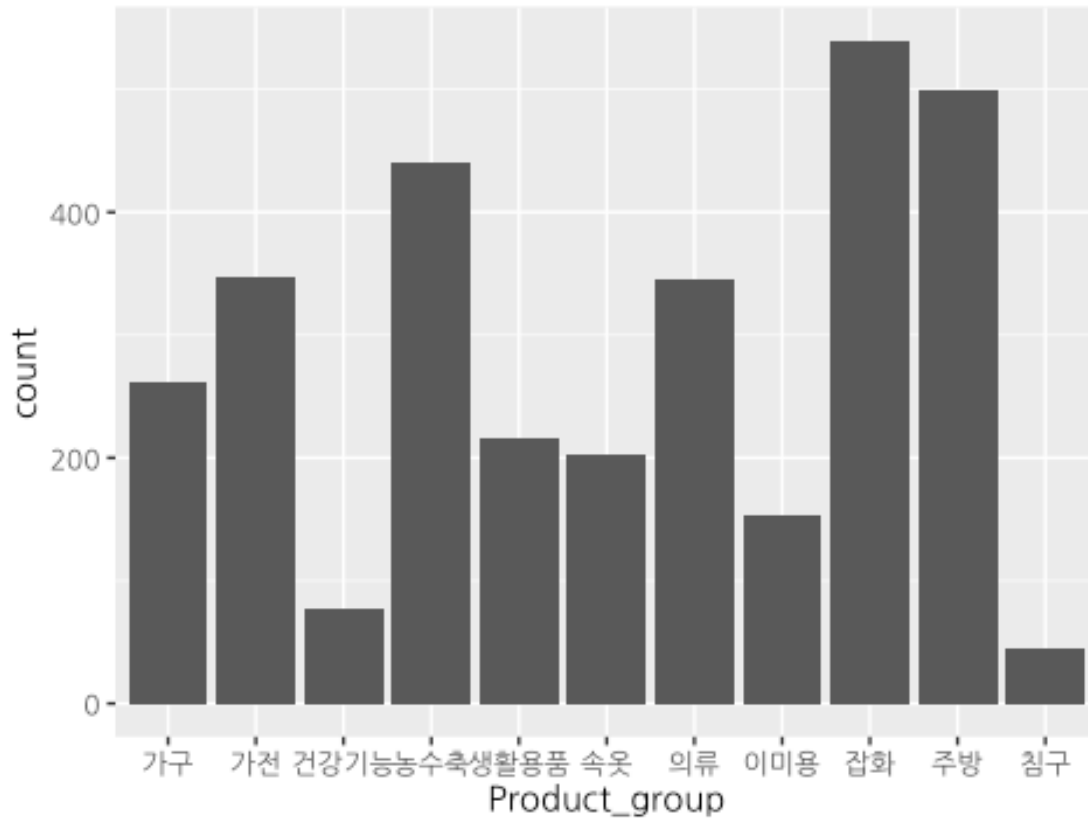
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("8 월  
달")+geom_col(position = "stack")
```

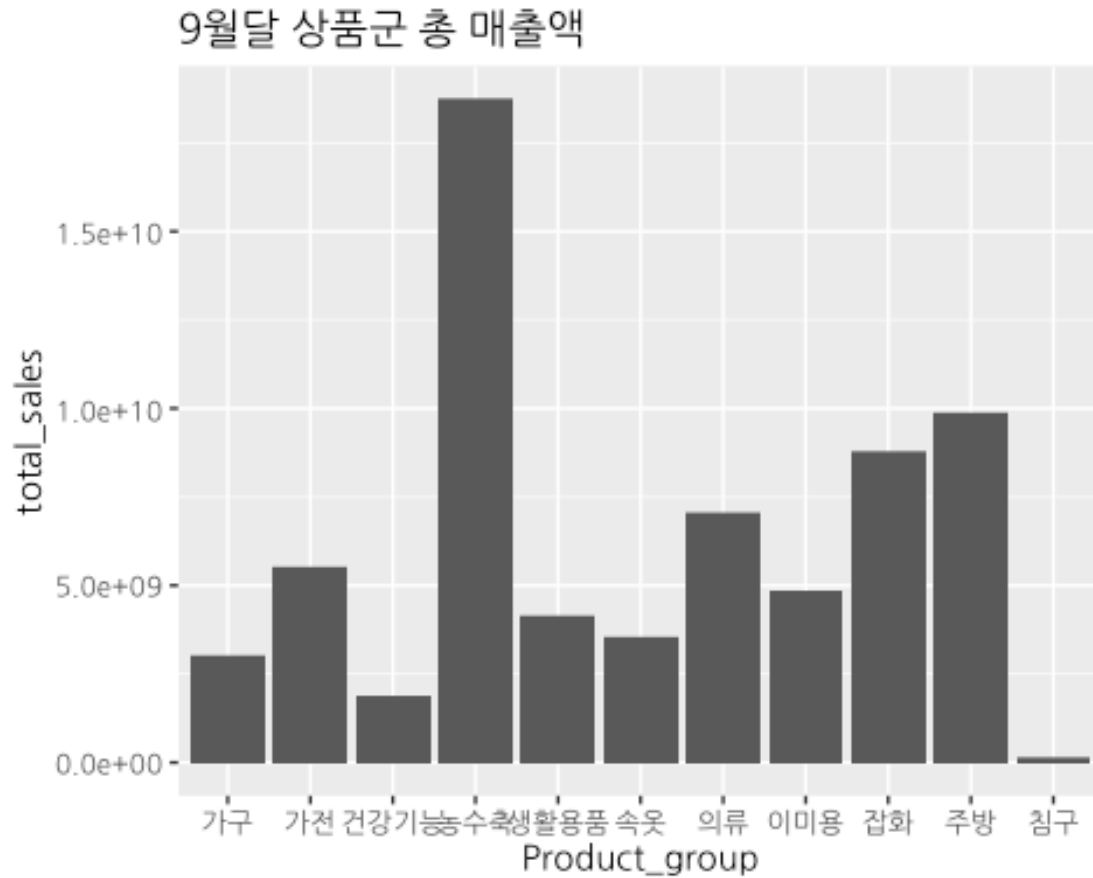


```
month9<-shop_data %>% filter(month=="9")
ggplot(data=month9)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("1 월달 상품군 편성 횟수")
```

1월달 상품군 편성 횟수



```
ggplot(data=month9, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("9 월달 상품군 총 매출액")
```



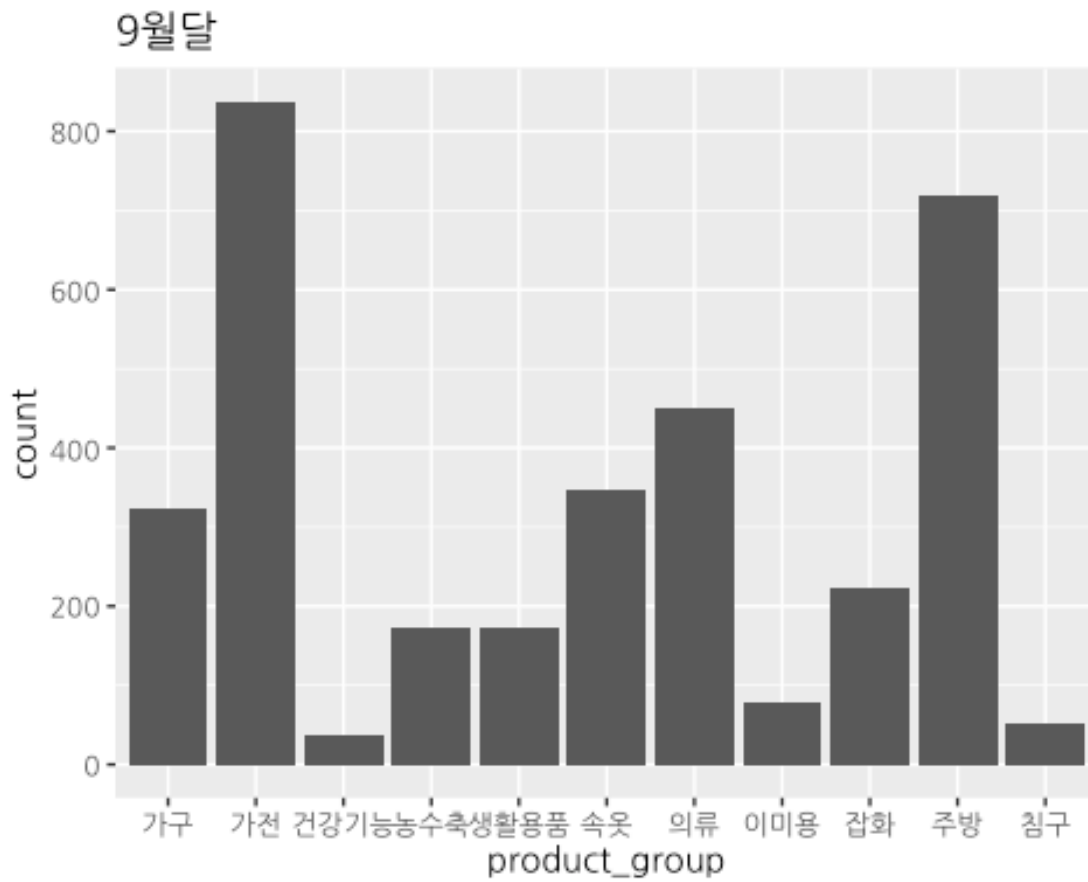
```
count<-as.data.frame(table(month9$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

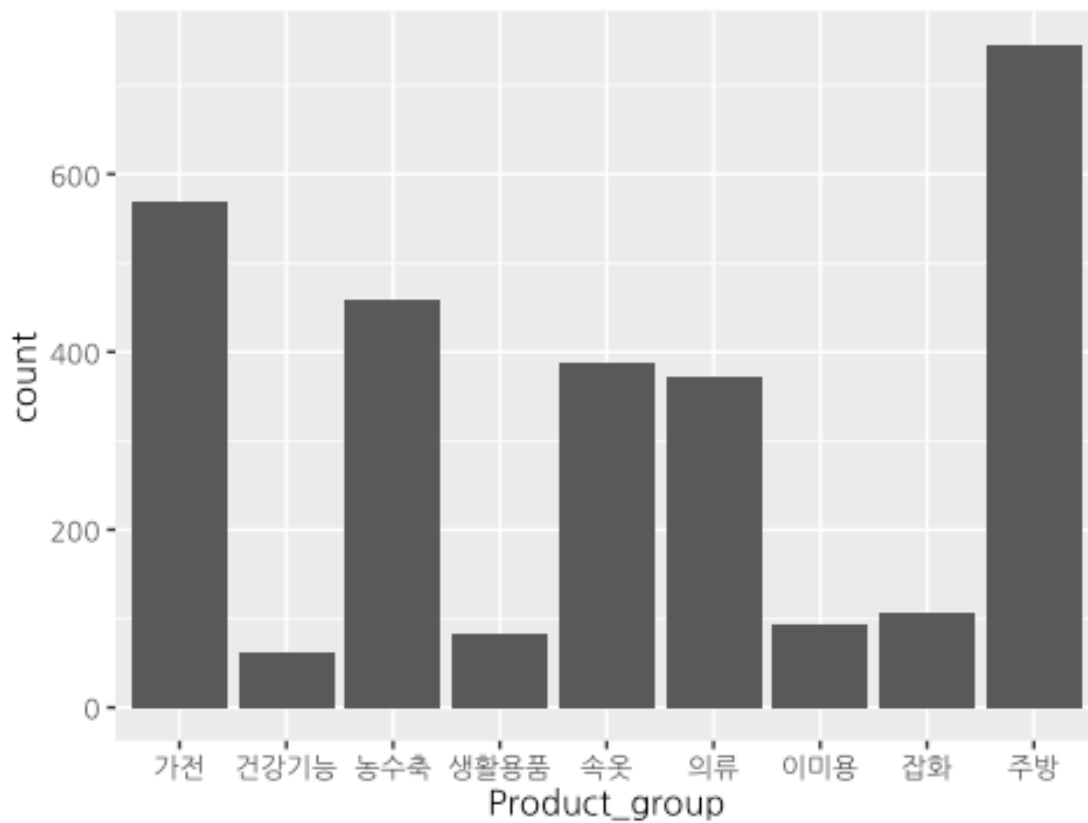
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("9 월  
달")+geom_col(position = "stack")
```

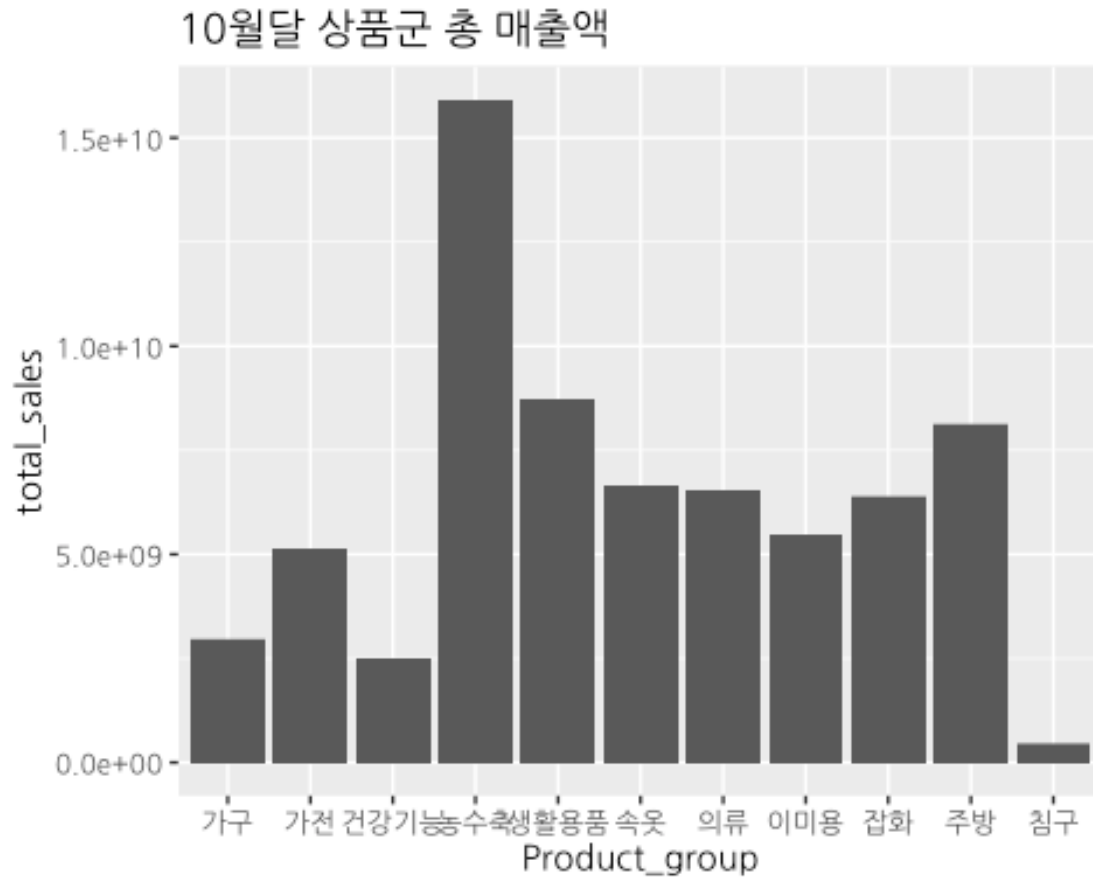


```
month10<-shop_data %>% filter(month=="10")
ggplot(data=month10)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("10 월달 상품군 편성 횟수")
```

10월달 상품군 편성 횟수



```
ggplot(data=month10, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("10 월달 상품군 총 매출액")
```



```
count<-as.data.frame(table(month10$Product_group))

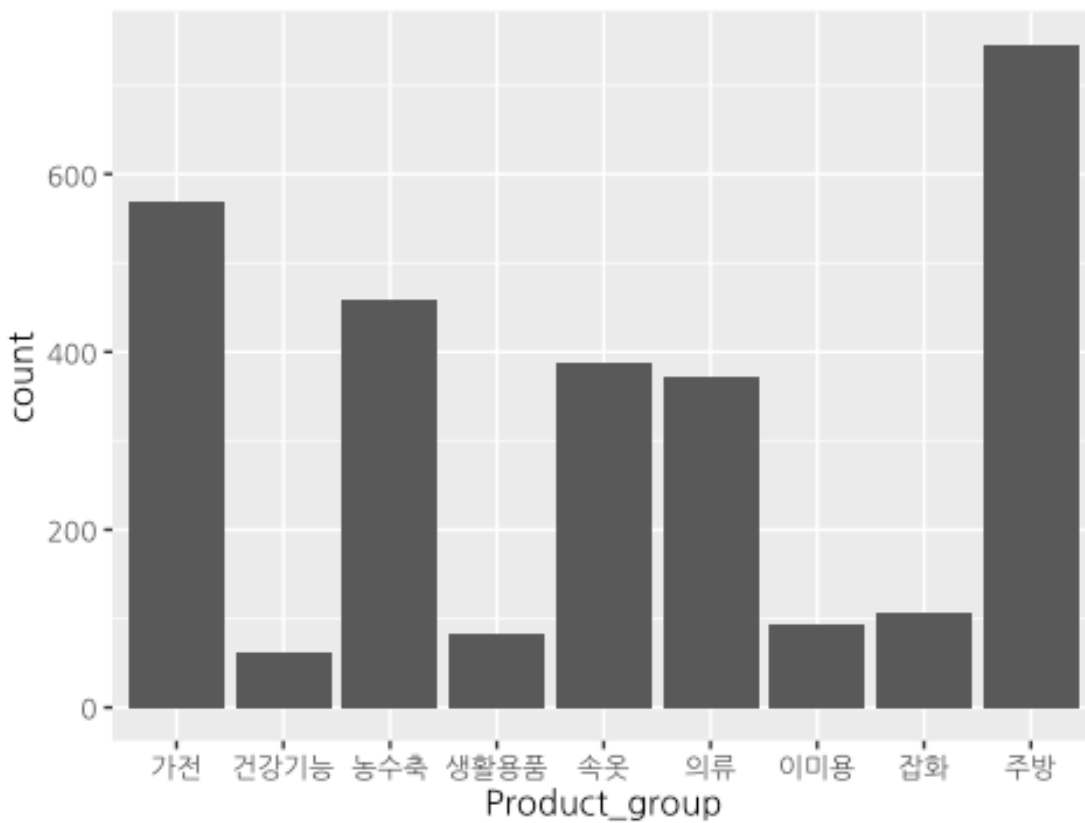
library(ggplot2)
count$Freq<-as.numeric(count$Freq)
count$Var1<-as.character(count$Var1)

ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("10월달")+geom_col(position = "stack")
```

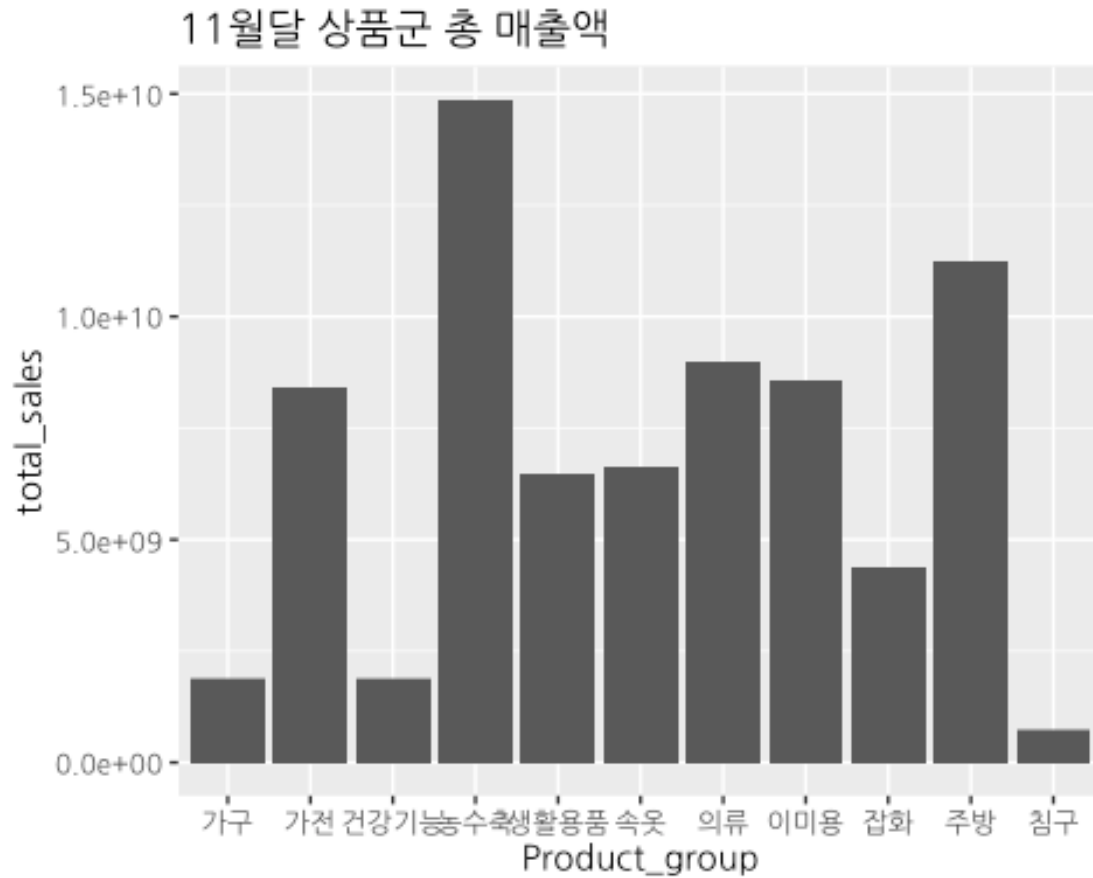


```
month11<-shop_data %>% filter(month=="11")
ggplot(data=month11)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("11 월달 상품군 편성 횟수")
```


11월달 상품군 편성 횟수



```
ggplot(data=month11, aes(x=Product_group, y=total_sales))+  
  geom_col(position="stack")+  
  ggtitle("11 월달 상품군 총 매출액")
```



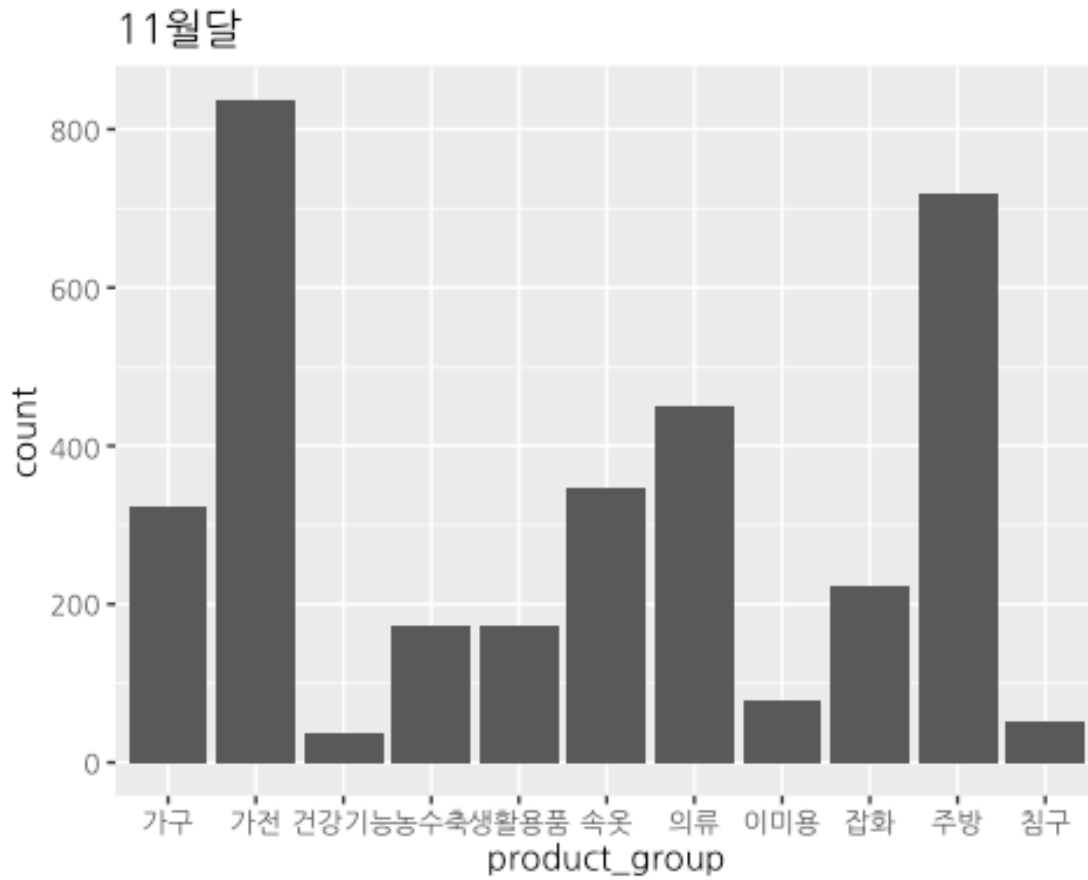
```
count<-as.data.frame(table(month11$Product_group))
```

```
library(ggplot2)
```

```
count$Freq<-as.numeric(count$Freq)
```

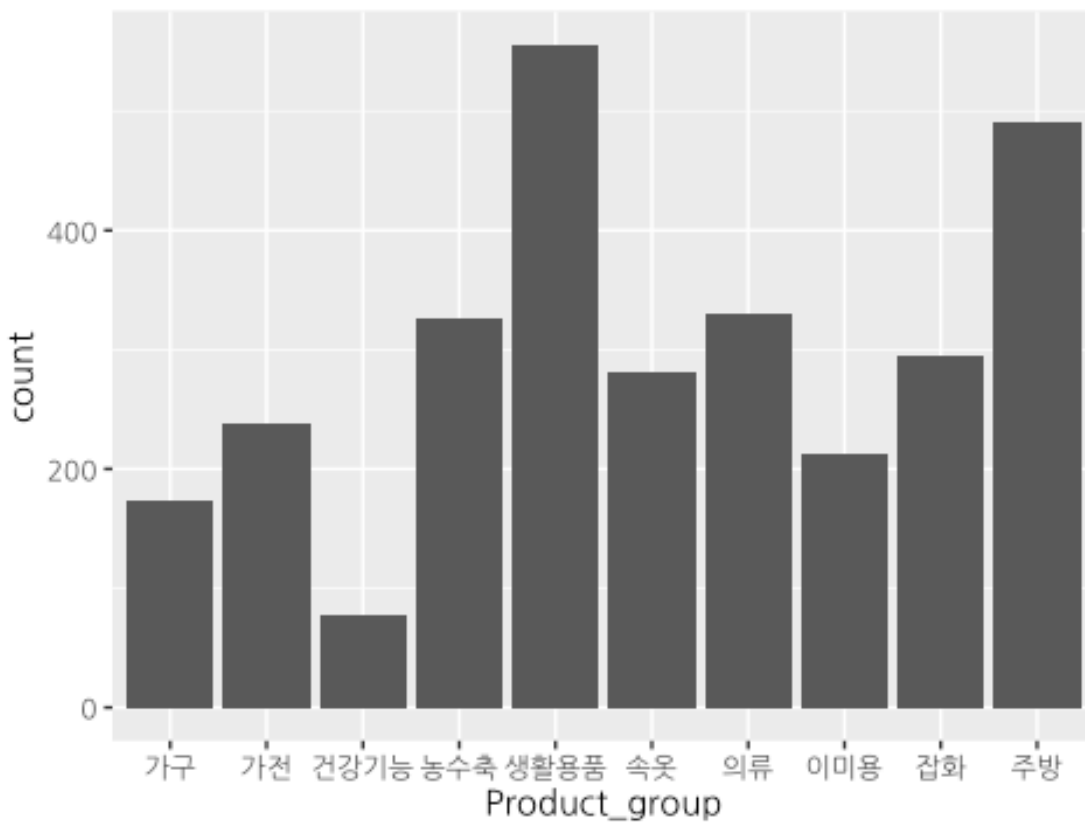
```
count$Var1<-as.character(count$Var1)
```

```
ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("11월달")+geom_col(position = "stack")
```

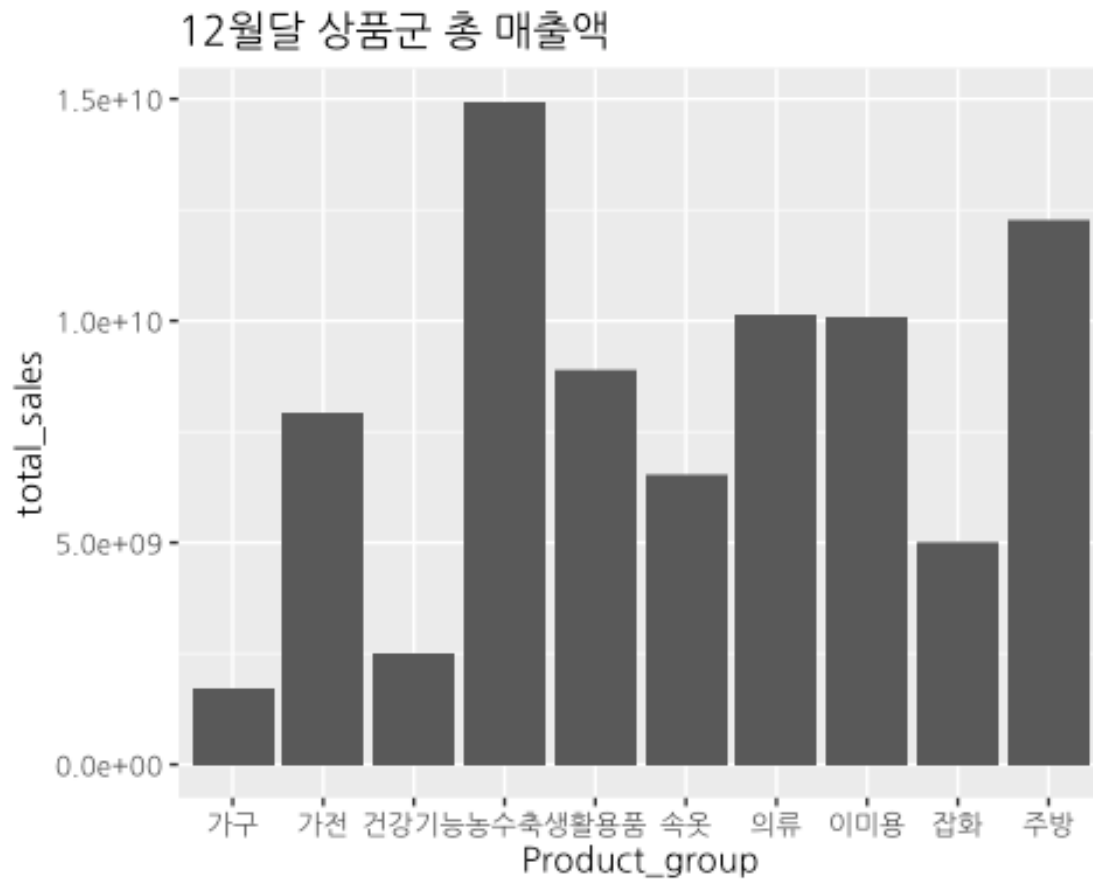


```
month12<-shop_data %>% filter(month=="12")
ggplot(data=month12)+
  geom_bar(mapping=aes(x=Product_group))+
  ggtitle("12 월달 상품군 편성 횟수")
```

12월달 상품군 편성 횟수



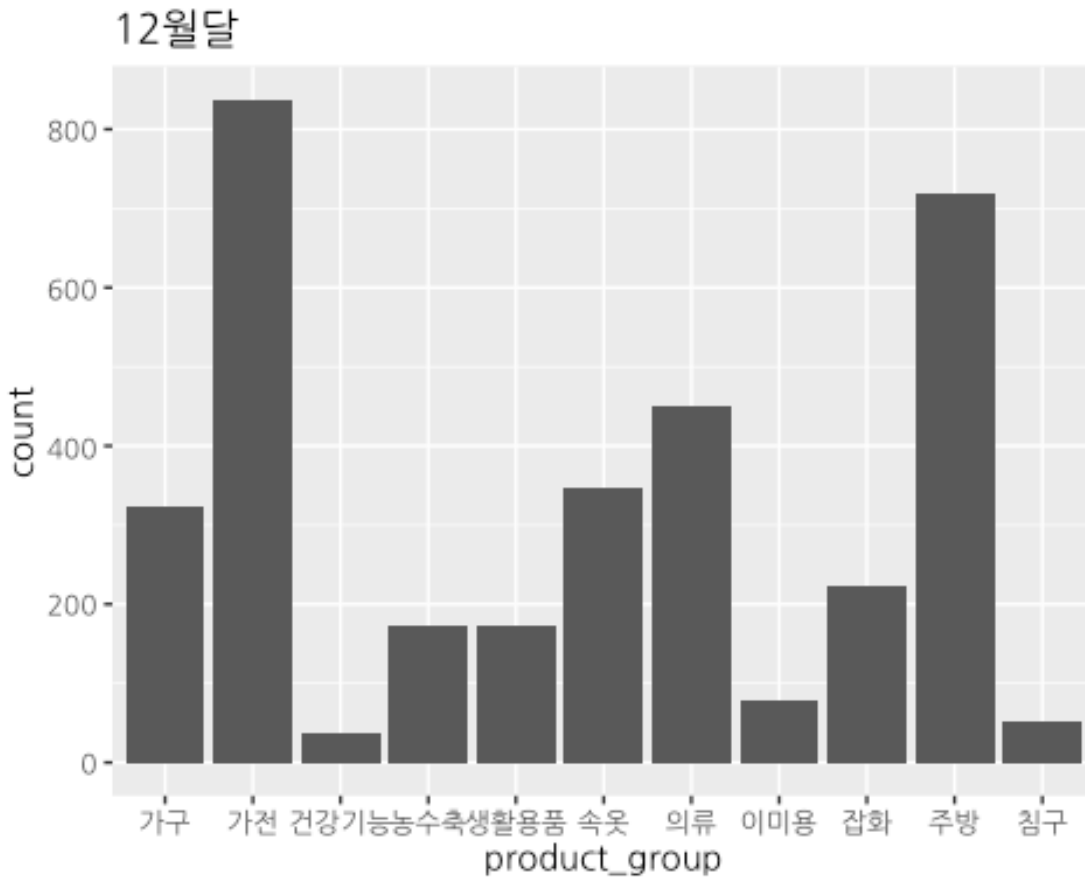
```
ggplot(data=month12, aes(x=Product_group, y=total_sales))+
  geom_col(position="stack")+
  ggtitle("12 월달 상품군 총 매출액")
```



```
count<-as.data.frame(table(month12$Product_group))

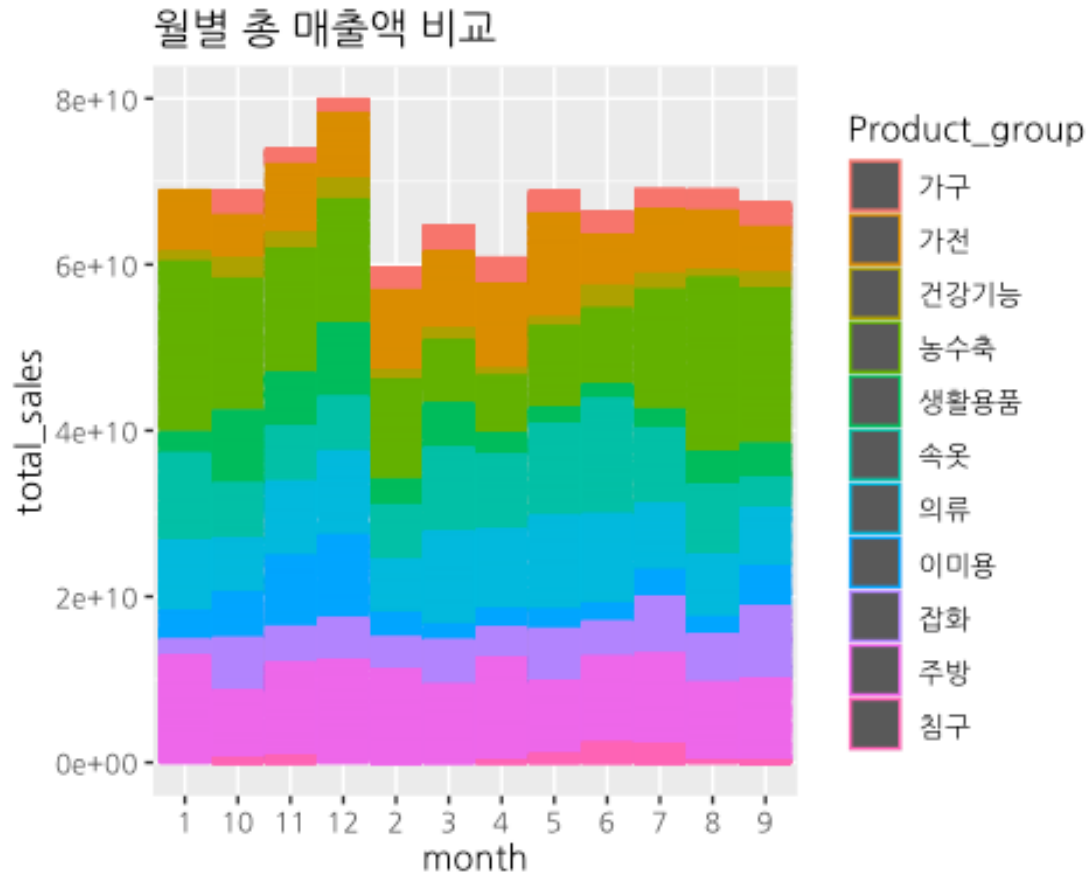
library(ggplot2)
count$Freq<-as.numeric(count$Freq)
count$Var1<-as.character(count$Var1)

ggplot(count,aes(Var1,Freq))+xlab("product_group")+ylab("count")+ggtitle("12
월달")+geom_col(position = "stack")
```



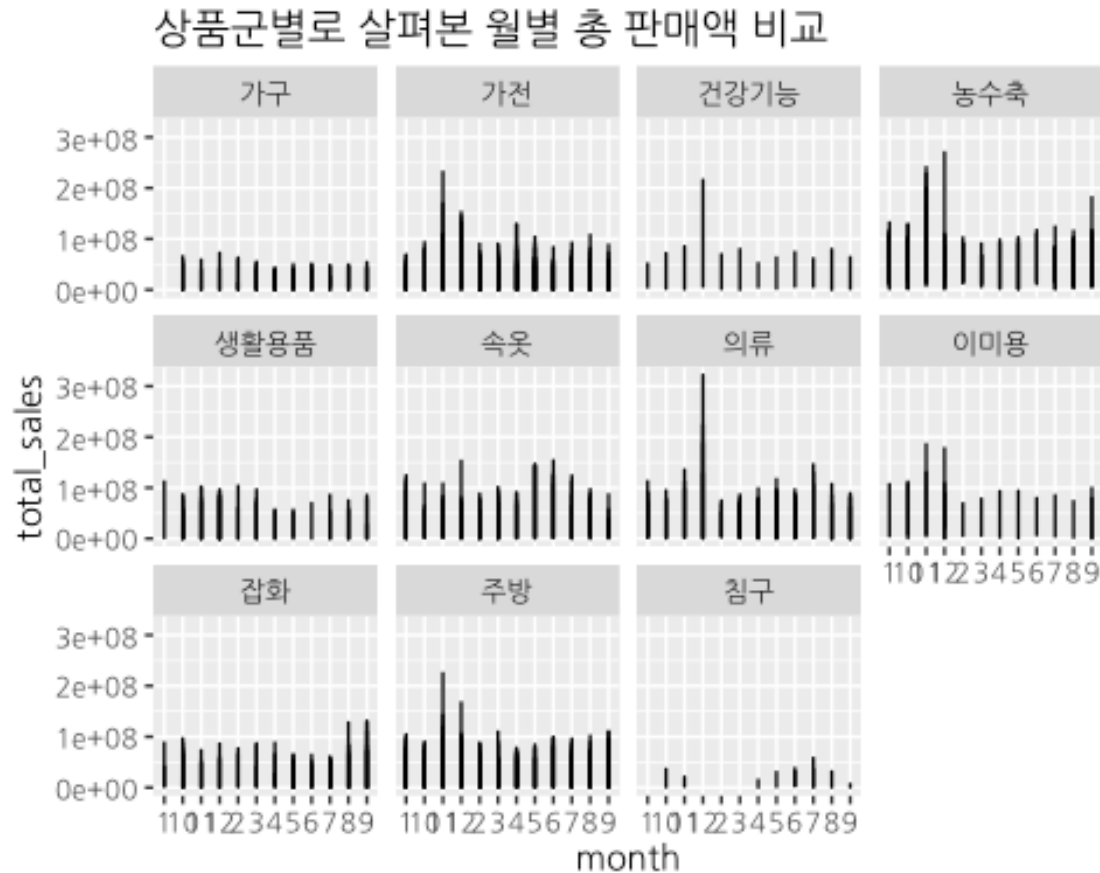
#한번에 그래프 그리기 #월별 총 매출액을 하나의 그래프로 합쳐 그렸다.

```
ggplot(data=shop_data, aes(x=month, y=total_sales, group=Product_group, color=Product_group))+
  geom_col(position="stack")+
  ggtitle("월별 총 매출액 비교")
```



#상품군별로 살펴본 월별 총 판매액 비교

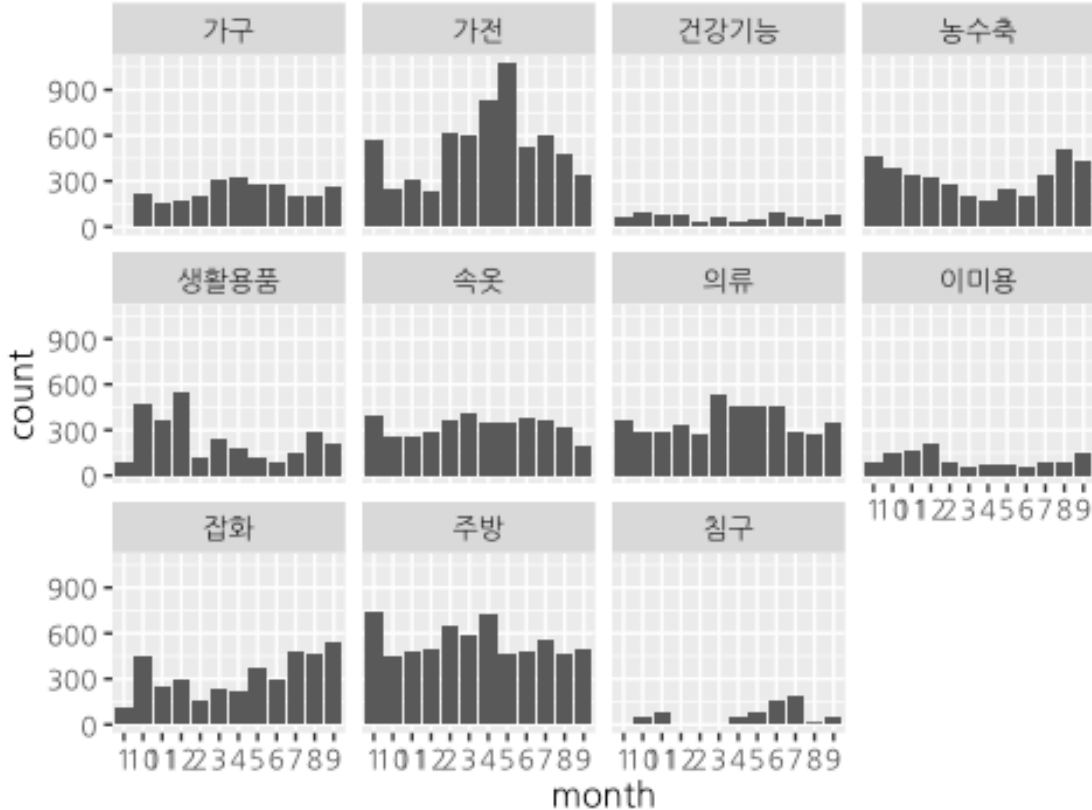
```
ggplot(data = shop_data, aes(x = month, y = total_sales)) +
  geom_line() +
  facet_wrap(~ Product_group) +
  ggtitle("상품군별로 살펴본 월별 총 판매액 비교")
```



#상품군별로 살펴본 월별 총 판매 횟수의 비교

```
ggplot(data = shop_data, aes(x = month)) +
  geom_bar() +
  facet_wrap(~ Product_group) +
  ggtitle("상품군별로 살펴본 월별 총 판매 횟수의 비교")
```


상품군별로 살펴본 월별 총 판매 횟수의 비교



#상품군

별 월별 총 판매액 주차 추이

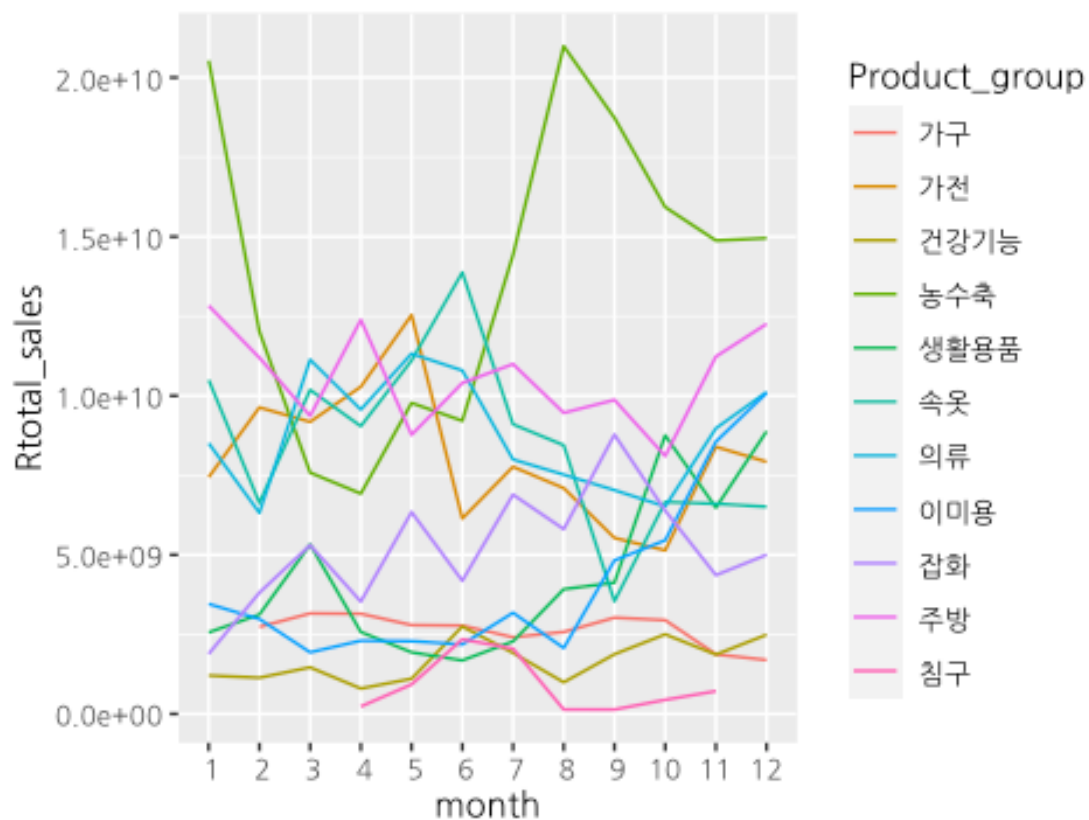
```
bymonth<-shop_data %>% group_by(month, Product_group) %>%
  summarize(Rtotal_sales=sum(total_sales))

## `summarise()` regrouping output by 'month' (override with `.groups` argument)

ggplot(data=bymonth, aes(x=month, y=Rtotal_sales, group=Product_group, color=
Product_group))+
  geom_line()+
  ggtitle("상품군 별 판매액 주차 추이")+scale_x_discrete(limits=c(1:12))

## Warning: Continuous limits supplied to discrete scale.
## Did you mean `limits = factor(...)` or `scale*_continuous()`?
```

상품군 별 판매액 주차 추이



신승은

#환경 세팅

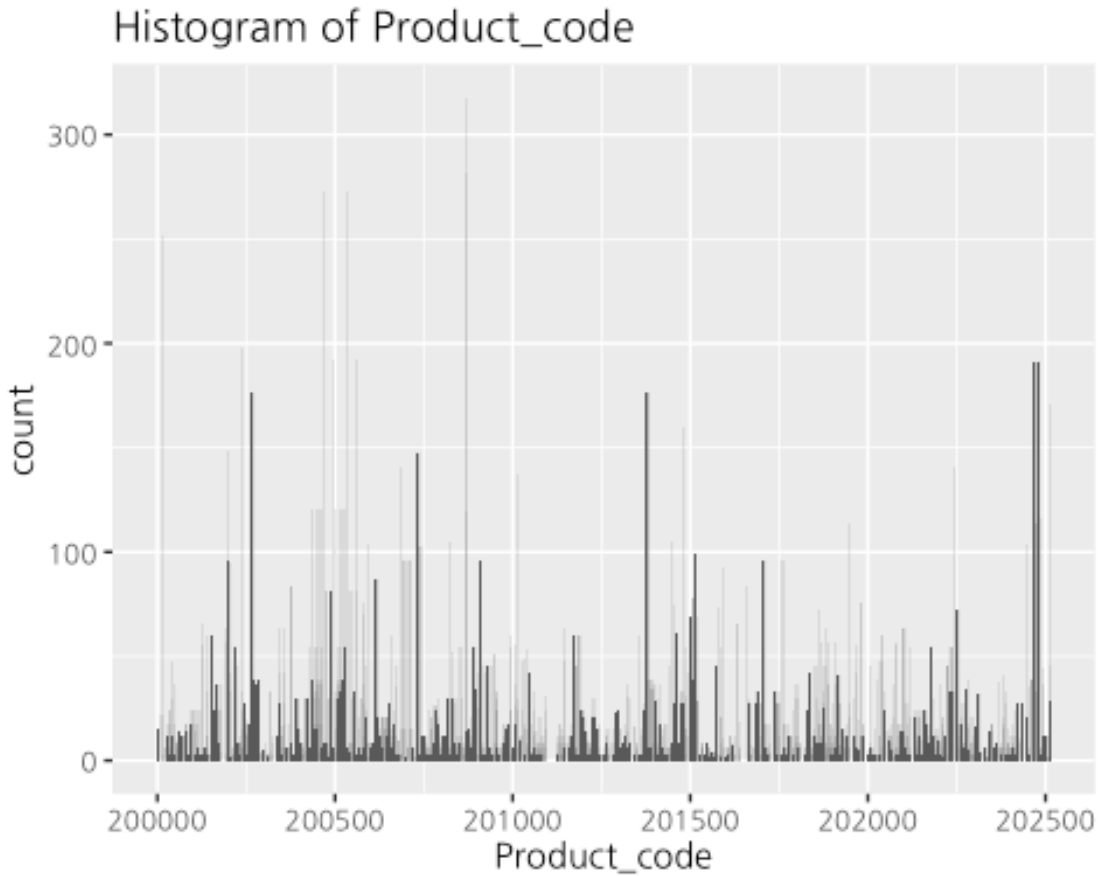
#데이터 불러오기

```
shop_data <- as.data.frame(shop_data)
```

Product_code

상품코드별 histogram

```
ggplot(data=shop_data)+geom_bar(mapping=aes(x=Product_code))+ggtitle("Histogram of Product_code")
```

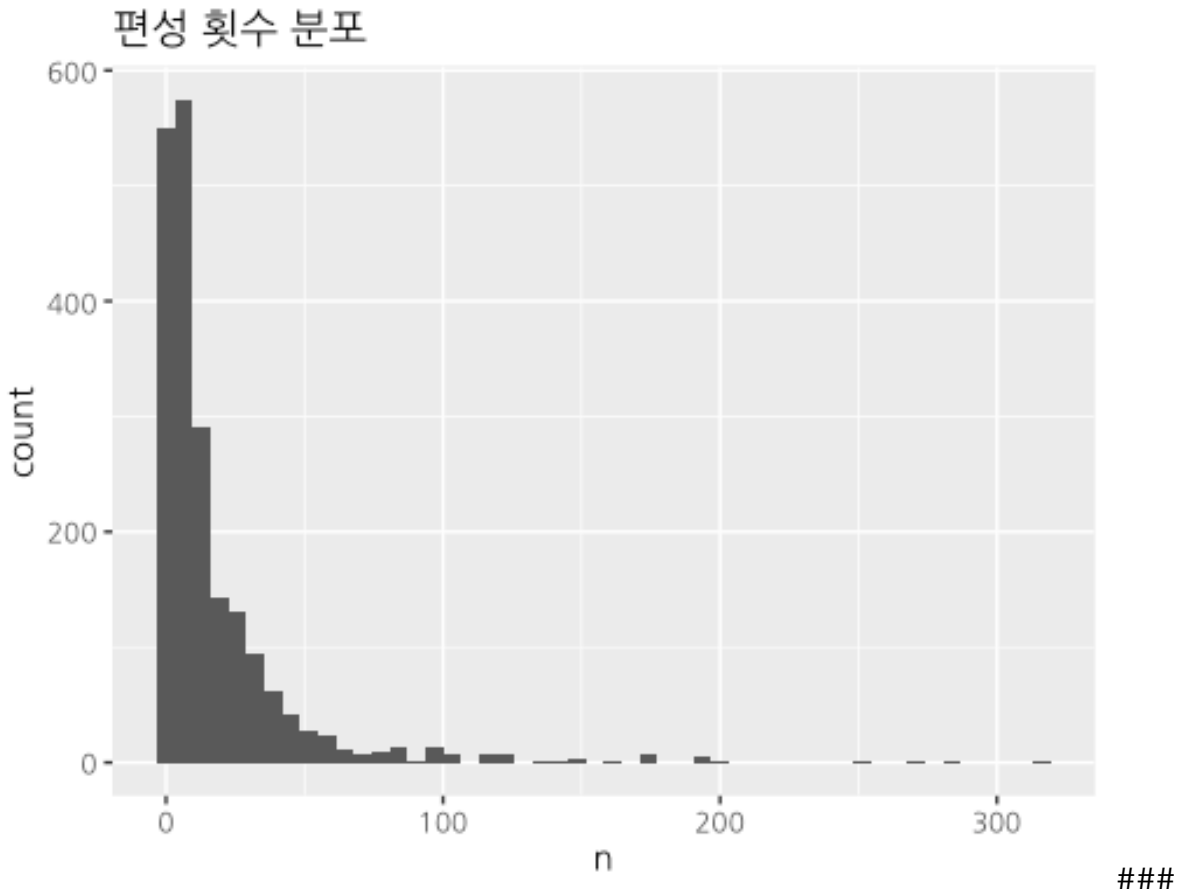


```
freq1<- shop_data %>% count(Product_code) #table 과 같음
```

설명: 제일 많이 편성된 횟수는 318 번이며, 100 번 이상 편성된 상품은 총 50 개이다.

상품코드별 편성횟수의 분포

```
ggplot(data=freq1)+geom_histogram(aes(x=n),bins=50)+ggtitle("편성 횟수 분포")
```



설명: 편성 횟수가 0 번~100 번 사이에 많이 분포하고 있으며, 같은 상품에 대해 3 번, 6 번 정도를 제일 많이 편성하는 것으로 보인다.

상품코드별 평균 매출액 계산, 데이터프레임으로 묶기

```
pc_sales<- aggregate(shop_data$total_sales,by=list(shop_data$Product_code),FU
N=sum)
pc_info<- data.frame(pc_sales,freq1[,2])
names(pc_info)<- c("Product_code","total_sales","freq")
mean_sales<- pc_info$total_sales/pc_info$freq #평균 매출액
pc_info<- data.frame(pc_info,mean_sales)
p<- subset(shop_data,select=c('Product_code','Product_name','Product_group','
unit_price'))
product<- unique(merge(pc_info,p,by="Product_code"))
product<- data.frame(product, quantity=product$total_sales/product$unit_price)
head(product)
```

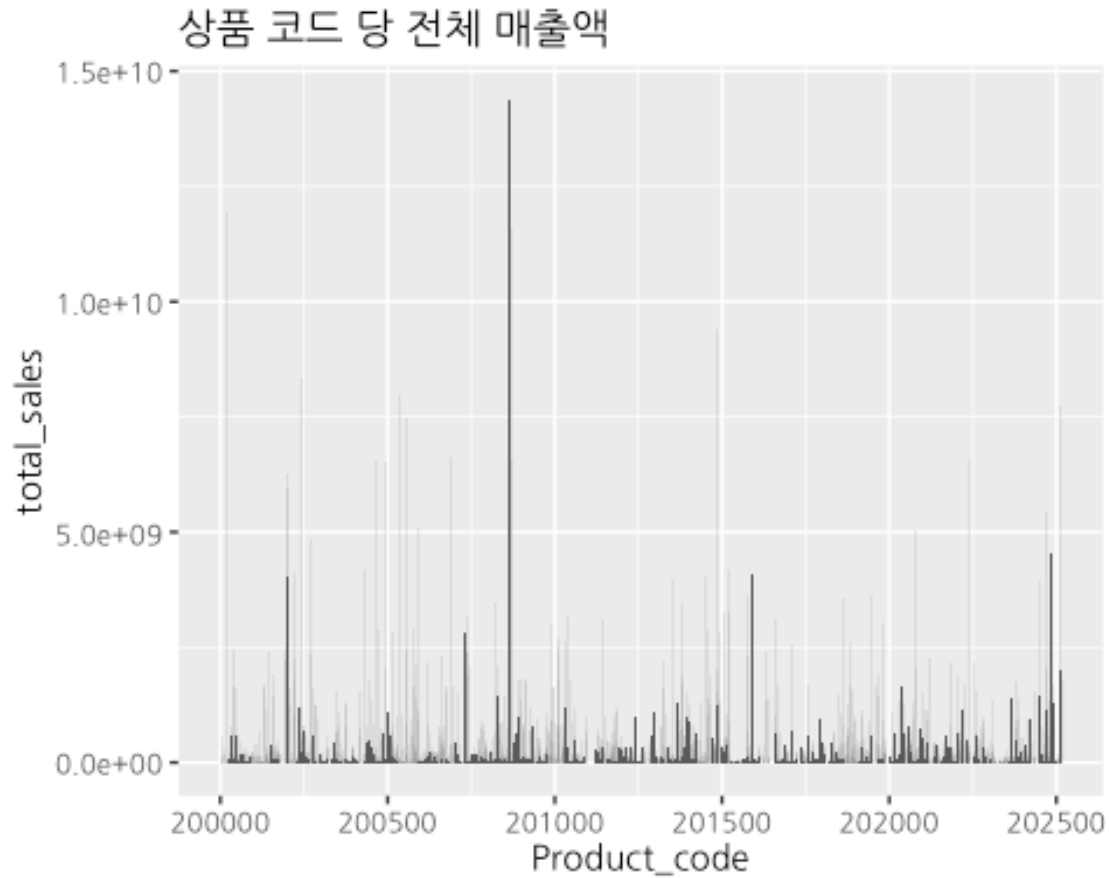
```
##      Product_code total_sales freq mean_sales
## 1          200000    76562000     6   12760333
## 7          200001    151336000     4    37834000
## 11         200002    571247000    15    38083133
```

```
## 26      200008      273214000      22      12418818
## 48      200009      171727000      22      7805773
## 70      200013      92222000       3      30740667
##
##                               Product_name Product_group unit_price  quant
ity
## 1      엘로엘 아쿠아클린 마스크      이미용      79900      958.222
8
## 7  국내생산 스텐락 심플 스텐밀폐용기 17 종      주방      99000      1528.6465
## 11      이보은의 우삼겹 12 팩세트      농수축      49900      11447.835
7
## 26      AAA 19 F/W 남성 트레킹화 3 종      잡화      49800      5486.22
49
## 48      AAA 19 F/W 여성 트레킹화 3 종      잡화      49800      3448.33
33
## 70      AAB 의 든든한 콩국수      농수축      39900      2311.32
83
```

상품코드별 그래프

설명: 상품코드별 총 매출액이 가장 높았던 상품은 농수축 상품군의 안동간고등어 20 팩이다. 그 다음은 농수축 상품군의 AAB 의 소곱창전골 800g x 8 팩이다. 이 두가지 상품은 편성 빈도도 매우 높게 나타났다. (각각 318, 252 번)

```
ggplot(data=product)+geom_col(aes(x=Product_code,y=total_sales))+ggtitle("상품 코드 당 전체 매출액")
```



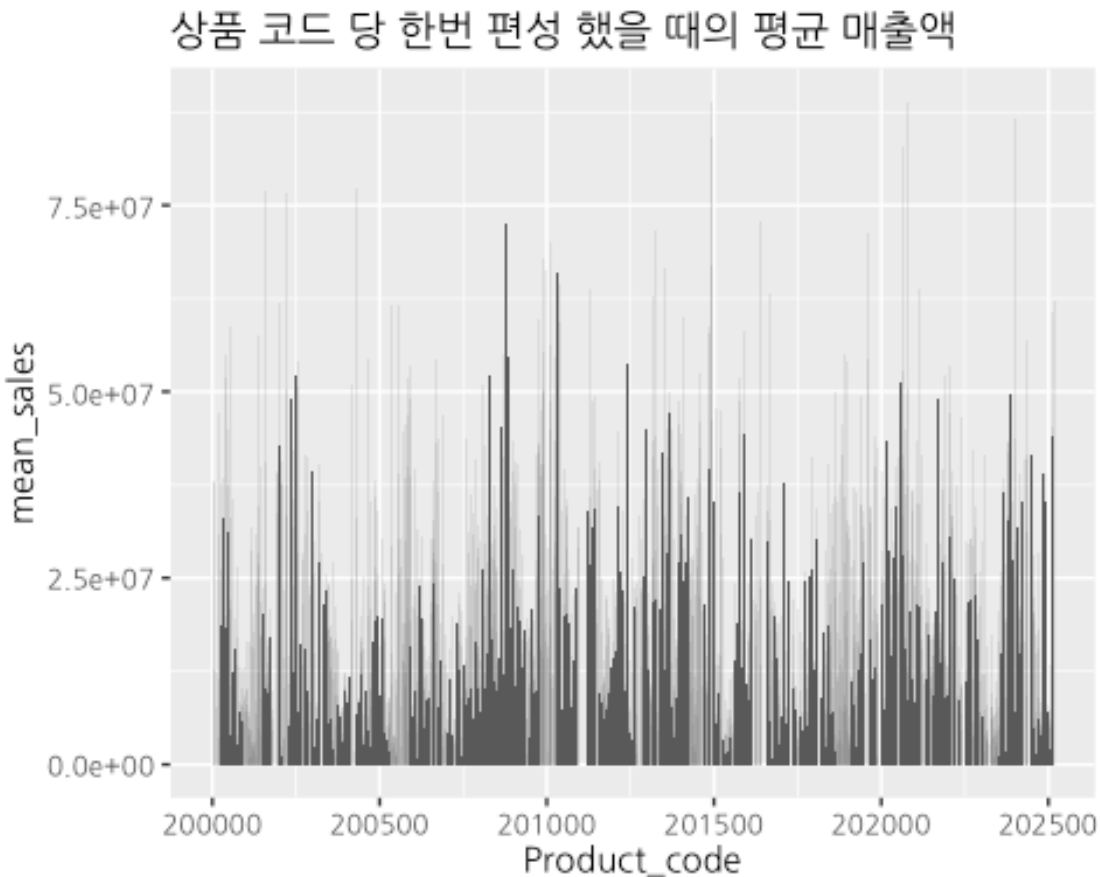
```
product_order1<- product[order(-product$total_sales),]; head(product_order1)
```

##	Product_code	total_sales	freq	mean_sales	Product_na
## 14809	200864	14391949000	318	45257701	안동간고등어 20팩
## 73	200014	11923972000	252	47317349	AAB 의 소곱창전골 800g x 8팩
## 15133	200868	11607152000	282	41160113	한일 대용량 스텐 분쇄믹서기
## 23150	201479	9417609000	160	58860056	국내산 손질갑오징어 8팩
## 2920	200236	8327167000	198	42056399	에코라믹 통주물 스톤 냄비세트
## 9001	200533	7996652000	273	29291766	일시불 LG 통돌이 세탁기
##	Product_group	unit_price	quantity		
## 14809	농수축	30900	465758.87		
## 73	농수축	55900	213308.98		
## 15133	주방	109800	105711.77		

```
## 23150      농수축      49900 188729.64
## 2920       주방      60900 136735.09
## 9001       가전      439000 18215.61
```

설명: 상품코드 당 평균 매출액이 가장 높았던 상품은 농수축 상품군의 피시원 국내산 절단 핫 꽃게 2.4kg 이다. 그 다음은 CERINI

```
ggplot(data=product)+geom_col(aes(x=Product_code,y=mean_sales))+ggtitle("상품
코드 당 한번 편성 했을 때의 평균 매출액 ")
```



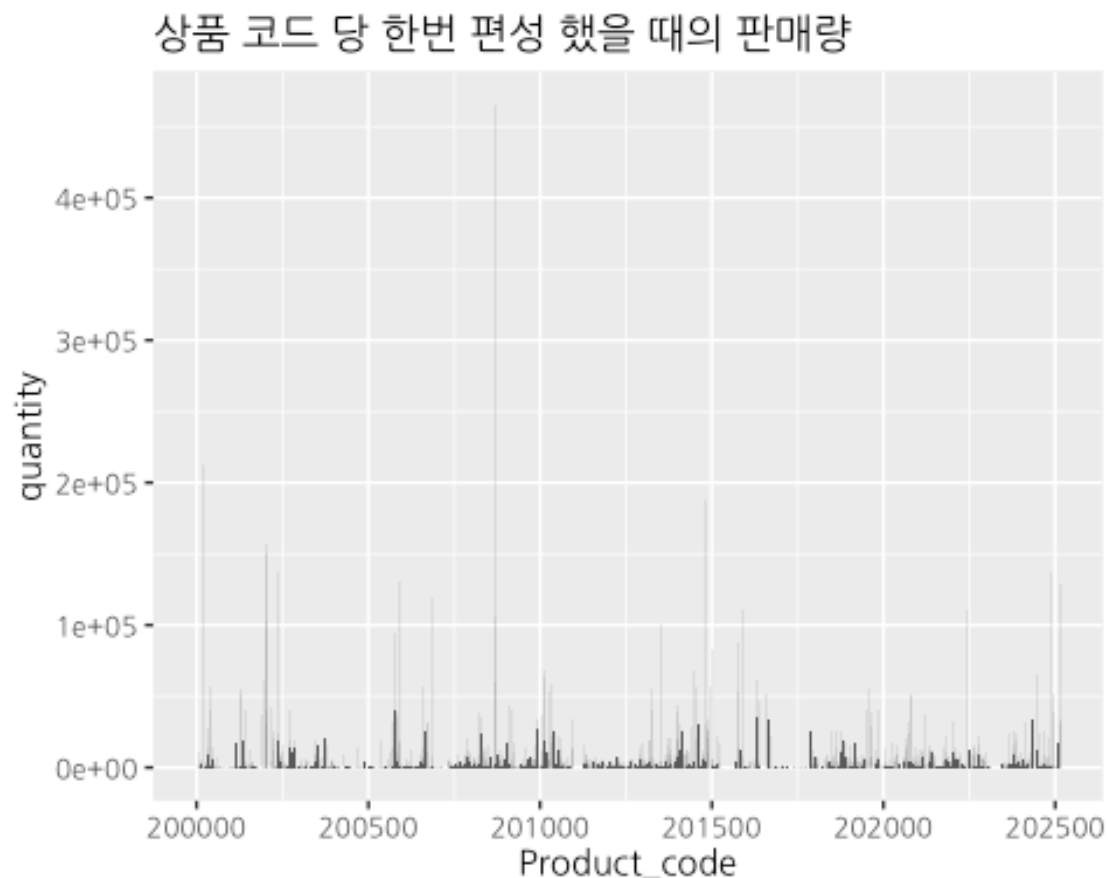
```
product_order2<- product[order(-product$mean_sales),]; head(product_order2)
```

```
##      Product_code total_sales freq mean_sales
## 23449      201486 1331715000   15  88781000
## 34689      202395  259678000    3  86559333
## 23464      201487 1264077000   15  84271800
## 30546      202064  496645000    6  82774167
## 5460       200427 4162757000   54  77088093
## 1591       200152 1924038000   25  76961520
##                                     Product_name Product_group unit_price
## 23449      피시원 국내산 절단 핫 꽃게 2.4kg      농수축      49900
```

## 34689	CERINI by PAT 남성 풀패키지 기모니트 3 종	의류	79900
## 23464	피시원 국내산 절단 햇 꽃게 7 팩	농수축	49900
## 30546	고창 꿀 고구마 10kg	농수축	30900
## 5460	LG 전자 통돌이 세탁기 TR14WK1(화이트)	가전	489000
## 1591	대동모피 19 F/W 폭스 베스트	의류	149000
##	quantity		
## 23449	26687.675		
## 34689	3250.038		
## 23464	25332.204		
## 30546	16072.654		
## 5460	8512.796		
## 1591	12913.007		

설명: 상품코드 당 판매량이 가장 높았던 상품은 농수축 상품군의 안동간고등어 20 팩이다. 그 다음은 농수축 상품군의 AAB 의 소금창전골 800g x 8 팩이다. 이 트렌트는 총 매출액을 비교했을 때와 비슷하다.

```
ggplot(data=product)+geom_col(aes(x=Product_code,y=quantity))+ggtitle("상품  
코드 당 한번 편성 했을 때의 판매량")
```




```
product_order3<- product[order(-product$quantity),]; head(product_order3)
```

	Product_code	total_sales	freq	mean_sales	
## 14809	200864	14391949000	318	45257701	
## 73	200014	11923972000	252	47317349	
## 23150	201479	9417609000	160	58860056	
## 2167	200197	6239718000	149	41877302	
## 2316	200199	5946118000	96	61938729	
## 36816	202484	4528056000	116	39034966	

	Product_name	Product_group	unit_price
## 14809	안동간고등어 20 팩	농수축	30900
## 73	AAB 의 소금창전골 800g x 8 팩	농수축	55900
## 23150	국내산 손질갑오징어 8 팩	농수축	49900
## 2167	비버리힐스폴로클럽 남성기초세트(2018FW 골드)	이미용	39800
## 2316	비버리힐스폴로클럽 남성기초세트(골드+콜라겐)	이미용	39800
## 36816	옛날 그 쥐포 110 장 (11 팩)	농수축	32900

	quantity
## 14809	465758.9
## 73	213309.0
## 23150	188729.6
## 2167	156776.8
## 2316	149399.9
## 36816	137630.9

느낀점: 코드별로 일단 매출액, 수량, 평균 매출액 등으로 그려보긴 했지만, 특별한 유의성이 있는지 알아보는 건 쉽지는 않음. 그냥 어떤 상품이 매출이 높은지, 판매수량이 높은지 확인하는 정도. 하지만 왜 높은지 생각해볼 필요성은 있다고 봄. 나중에 편성할 때 필요할 수도 있다고 생각함.

시간변수를 함께 고려하면서 추이를 살펴보는게 더 의미있다.

Product_group

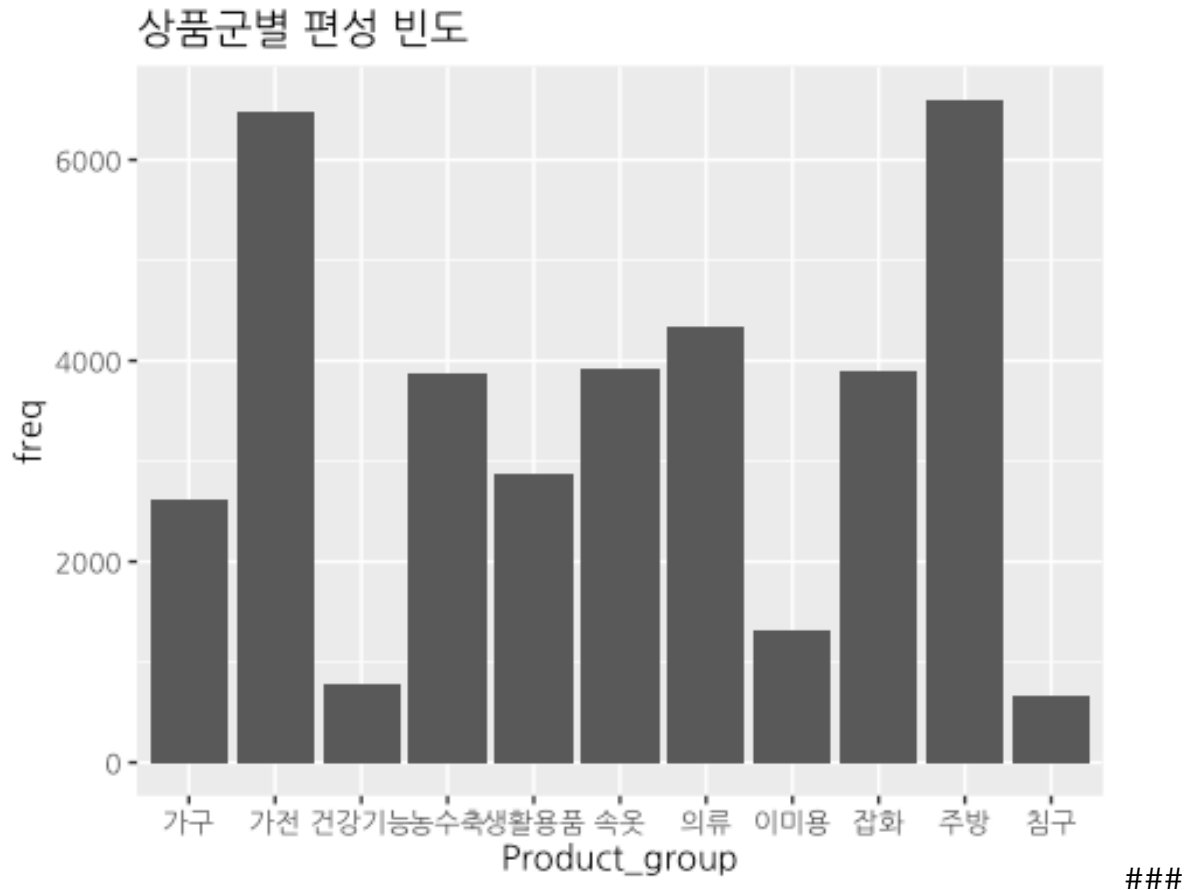
데이터 정리

```
freq2<- shop_data %>% count(Product_group)
pg_sales<- aggregate(shop_data$total_sales,by=list(shop_data$Product_group),FUN=sum)
pg_mean_sales<- aggregate(shop_data$total_sales,by=list(shop_data$Product_group),FUN=mean)
pg_info<- data.frame(pg_sales,pg_mean_sales,freq2[,2])
pg_info<- pg_info[,-3]
names(pg_info)<- c("Product_group","total_sales","mean_sales","freq")
```

그래프

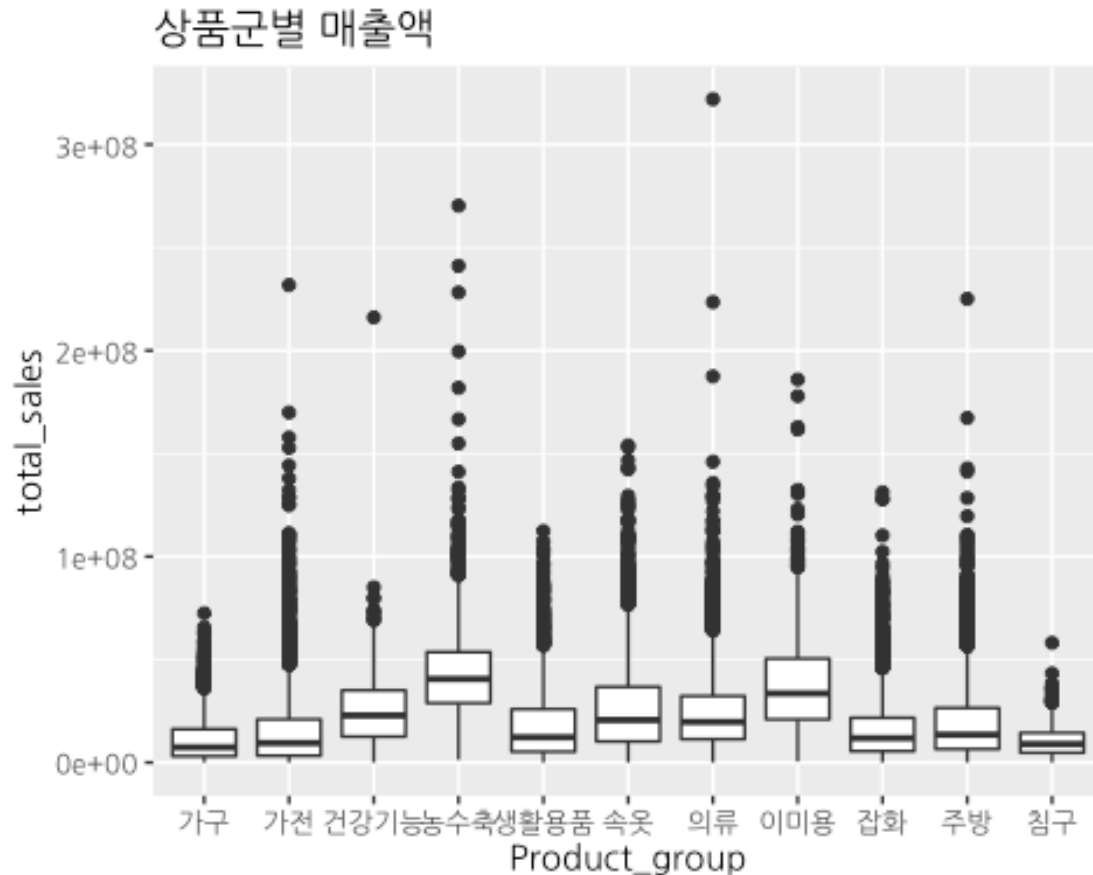
가전, 주방 상품군의 편성 빈도가 매우 높다.

```
ggplot(data=pg_info)+geom_col(mapping=aes(x=Product_group,y=freq))+ggtitle("상품군별 편성 빈도")
```



의류, 농수축, 가전, 주방의 상품군에서 이상치가 나타난다. 이는 어떤 특정한 상품에서 아주 높은 매출액을 기록한 것이다. ### 상품군들중 가장 높은 평균 매출액을 가지는 상품군은 농수축이다.

```
ggplot(data=shop_data,aes(x=Product_group,y=total_sales))+geom_boxplot()+ggtitle("상품군별 매출액")
```



##고유정 # 데이터 불러오기

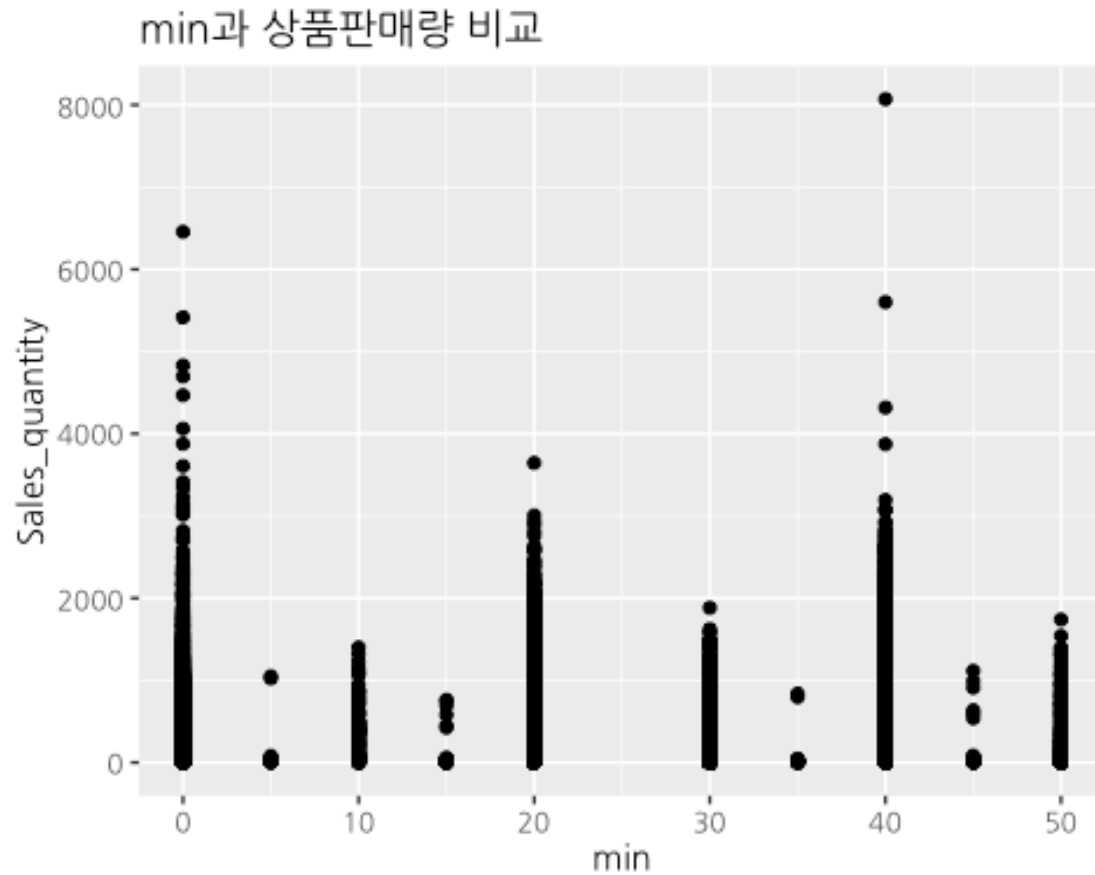
```
shop_data <- as.data.frame(shop_data)
min <- as.data.frame(table(shop_data$min))
```

총판매량은 농수축이 제일 많지만 방영횟수는 가전, 주방이 제일 많음

방영횟수 자체도 0-20-40 이 압도적으로 많아서 전반:0, 중반:20, 후반:40 으로 봄.

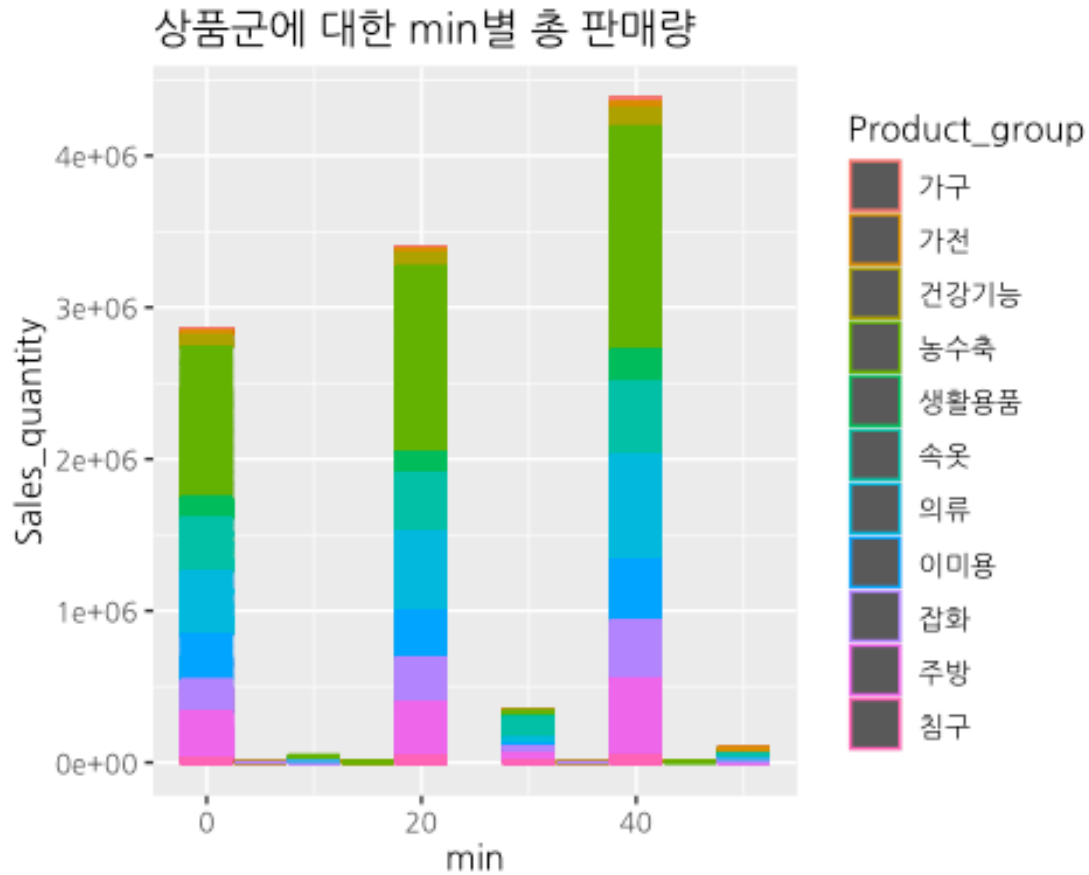
min 과 상품판매량 비교

```
library(ggplot2)
ggplot(data = shop_data, mapping = aes(x = min, y = Sales_quantity)) +
  geom_point() +
  ggtitle("min 과 상품판매량 비교")
```



상품군별 min 별 총판매량

```
ggplot(data=shop_data,aes(x=min, y=Sales_quantity,group=Product_group,colour  
= Product_group))+  
  geom_col()+  
  ggtitle("상품군에 대한 min 별 총 판매량")
```

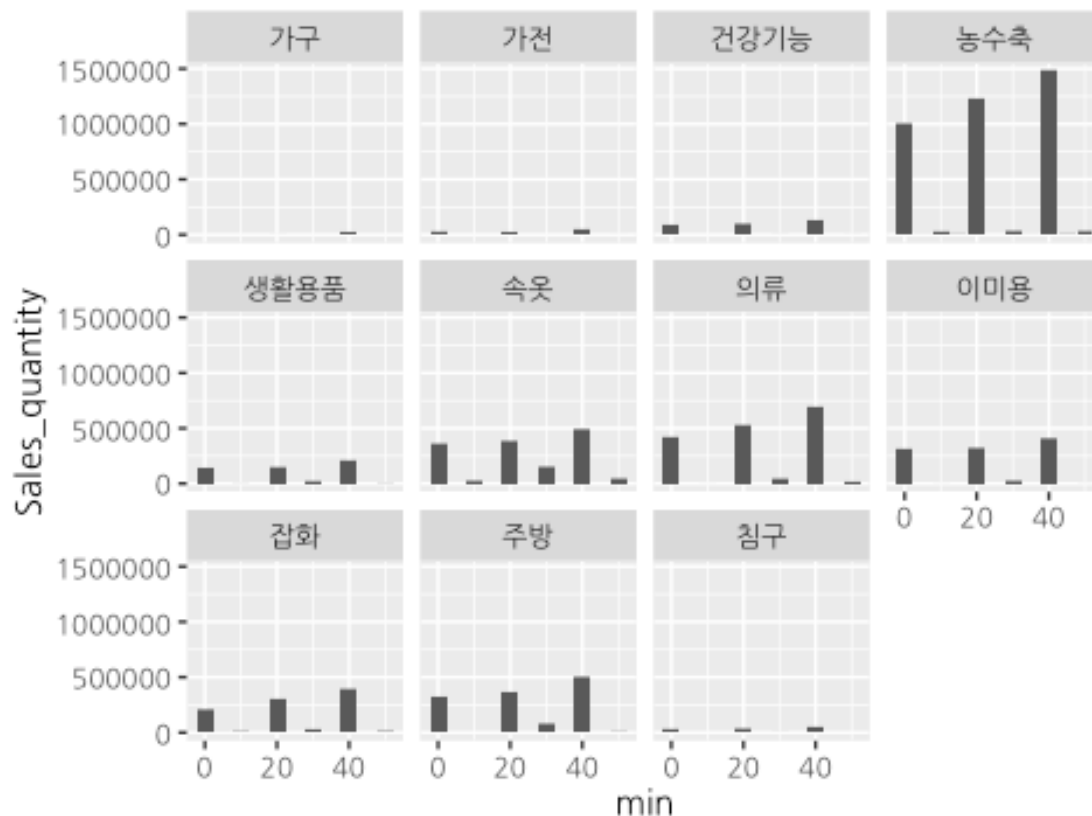


농수축 판매량 압도적으로 많음

상품군별로 min 과 상품판매량 비교

```
ggplot(data = shop_data, aes(x = min, y = Sales_quantity)) +
  geom_col() +
  facet_wrap(~ Product_group) +
  ggtitle("상품군별로 살펴본 시간대별 총 판매량 비교")
```

상품군별로 살펴본 시간대별 총 판매량 비교



min 에 대한 상품군별 방영횟수

```
library(tidyverse)

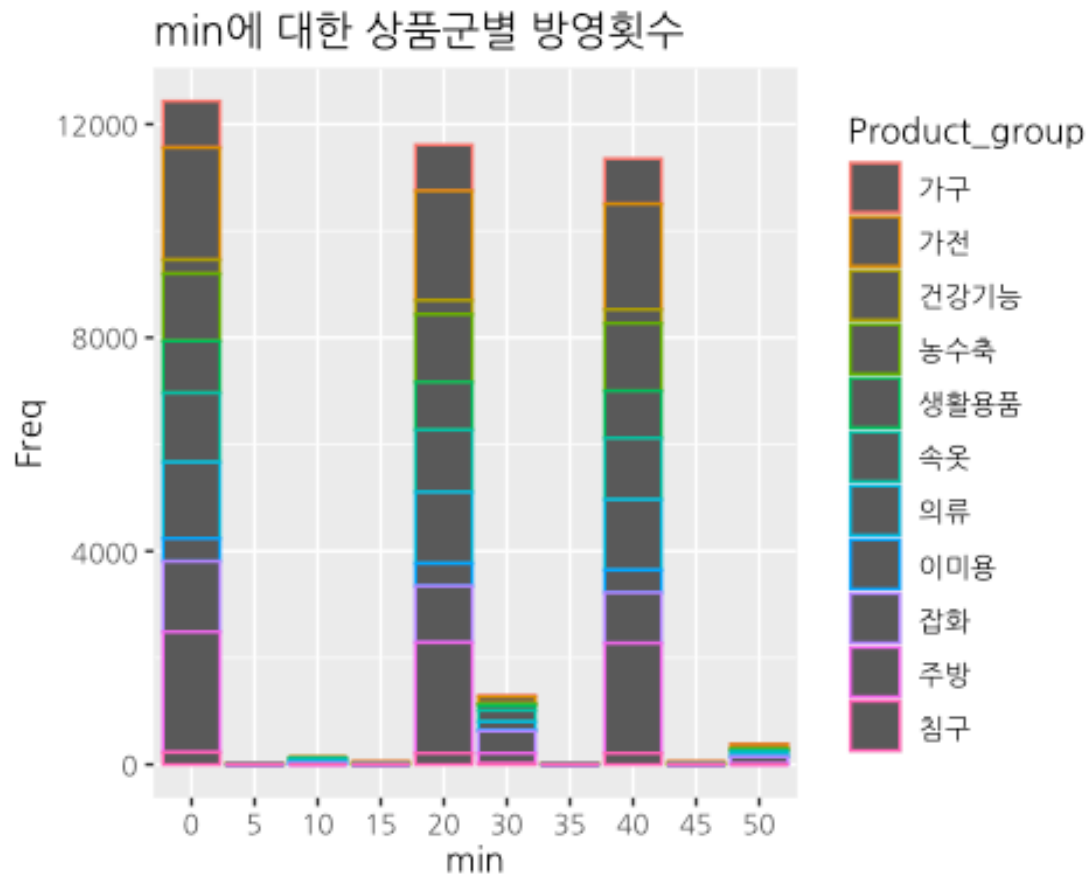
## — Attaching packages — tidyverse 1.3.0 —

## ✓ tibble 2.1.3      ✓ purrr 0.3.3
## ✓ tidyr 1.0.2       ✓ stringr 1.4.0
## ✓ readr 1.3.1       ✓ forcats 0.5.0

## — Conflicts — tidyverse_conflicts() —
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

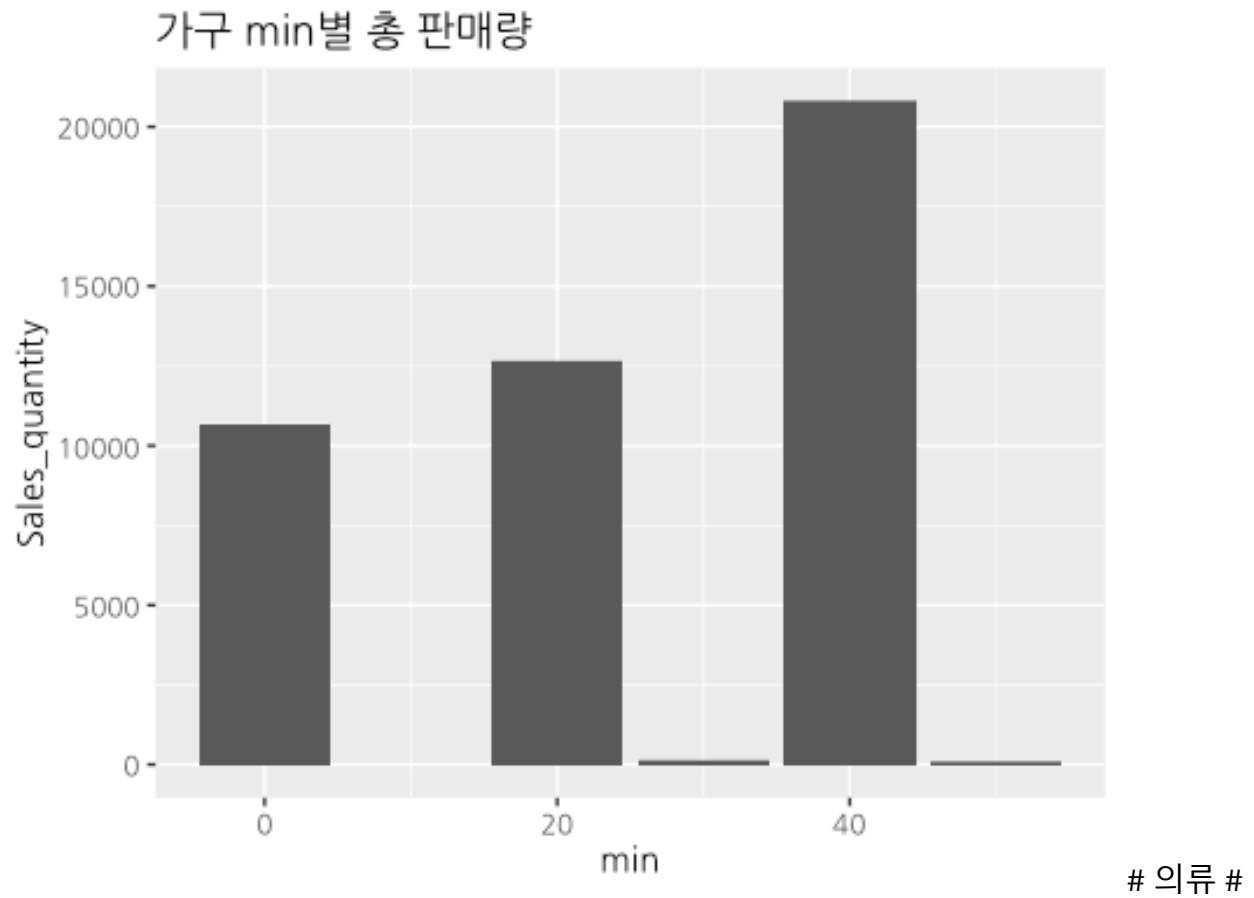
min_count <- as.data.frame(table(select(shop_data, Product_group, min)))

ggplot(data=min_count, aes(x=min, y=Freq, group=Product_group, colour = Product_group)) +
  geom_col() +
  ggtitle("min 에 대한 상품군별 방영횟수")
```

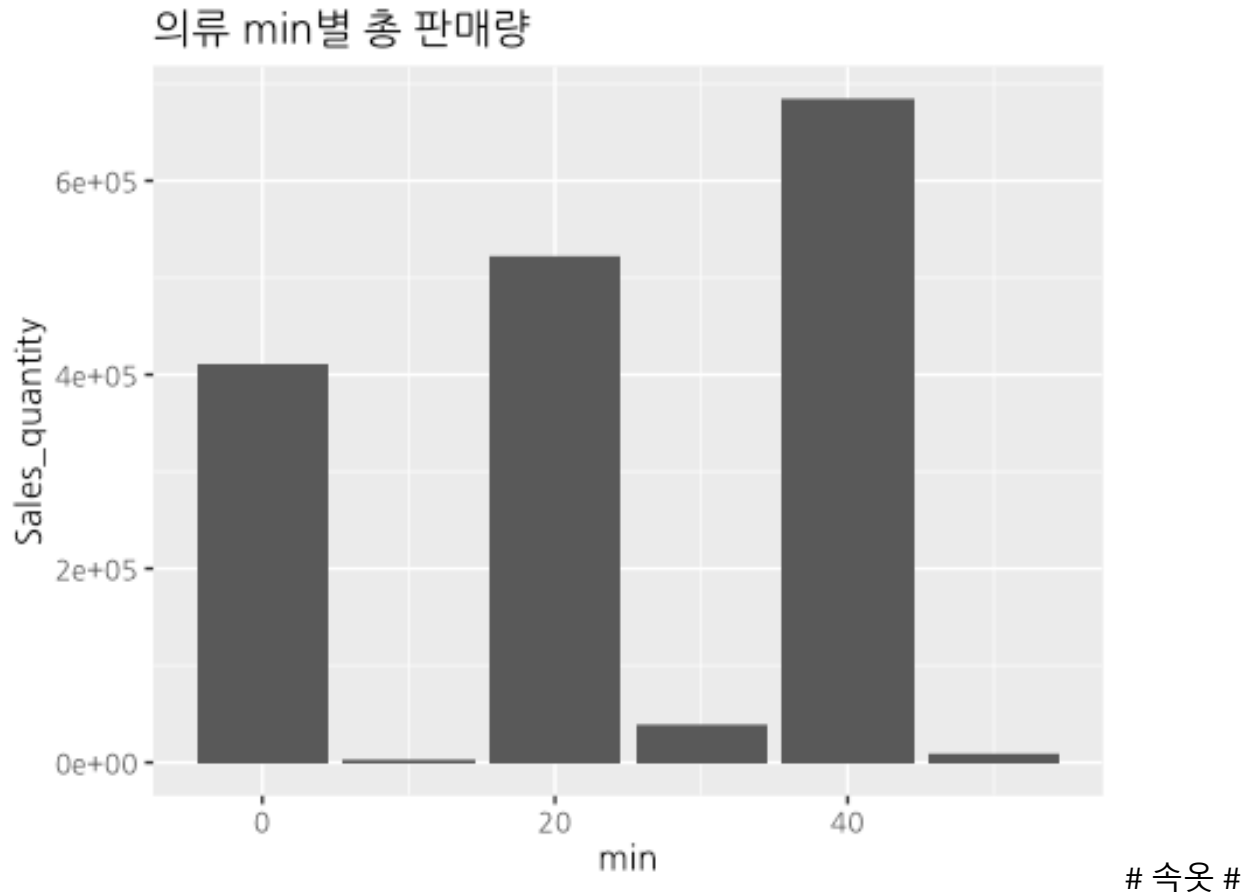


```
furniture <-shop_data %>% filter(Product_group=="가구")

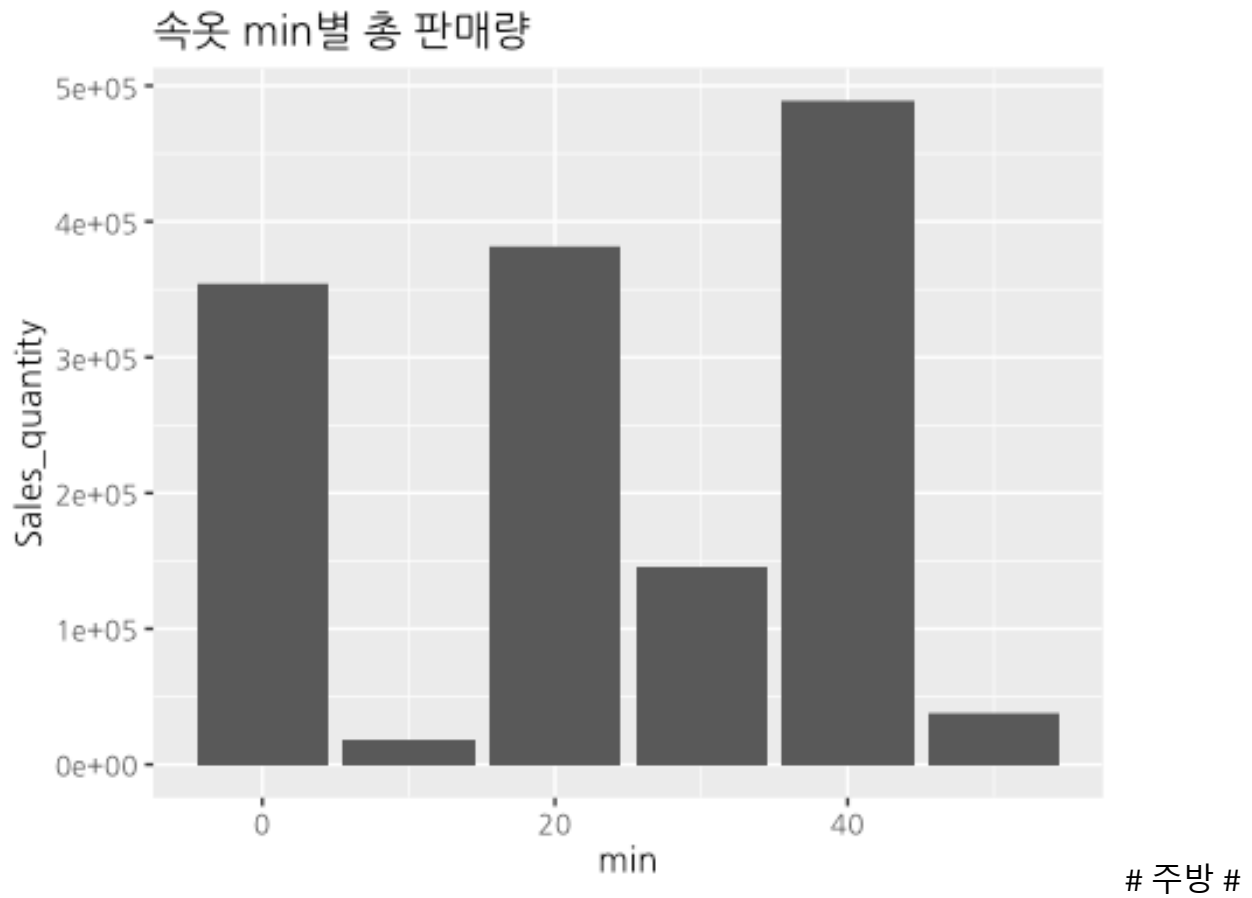
ggplot(data=furniture, aes(x=min, y=Sales_quantity))+
  geom_col(position="stack")+
  ggtitle("가구 min 별 총 판매량")
```



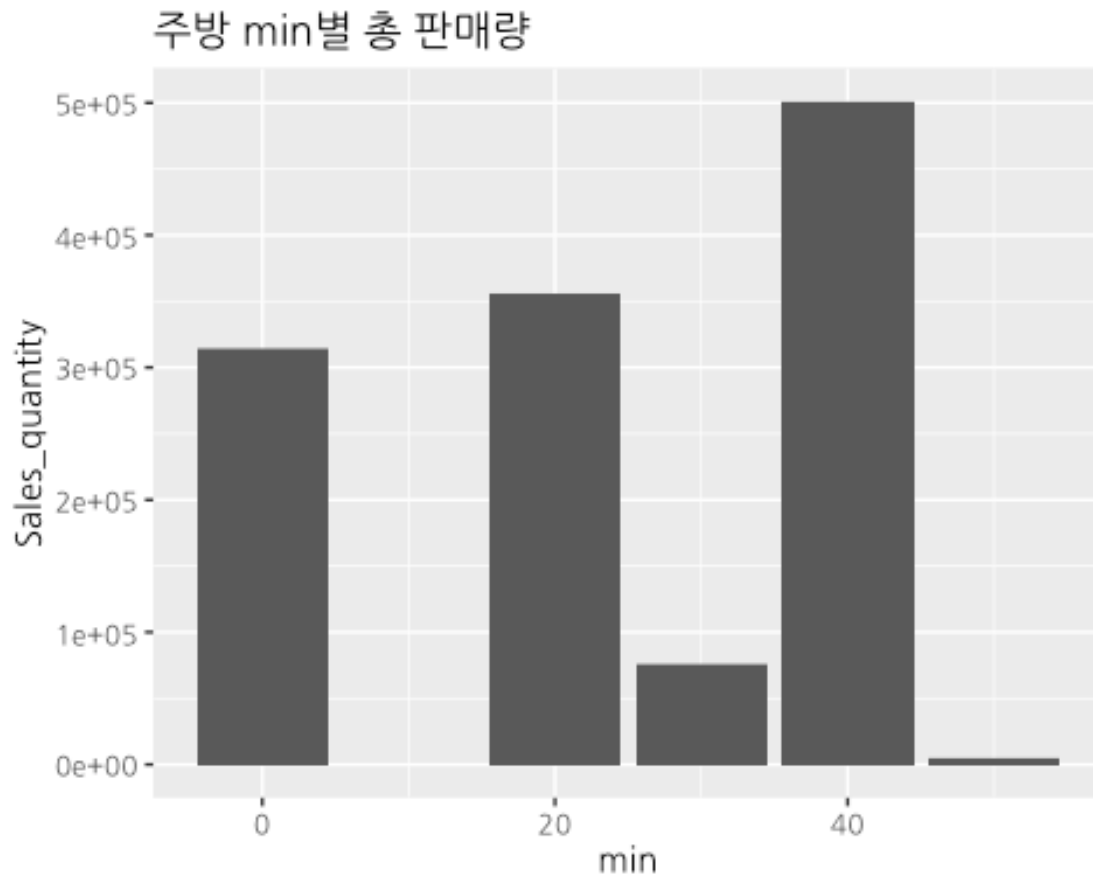
```
clothes <- shop_data %>% filter(Product_group=="의류")  
  
ggplot(data=clothes, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("의류 min 별 총 판매량")
```

```
underwear <- shop_data %>% filter(Product_group=="속옷")  
  
ggplot(data=underwear, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("속옷 min 별 총 판매량")
```



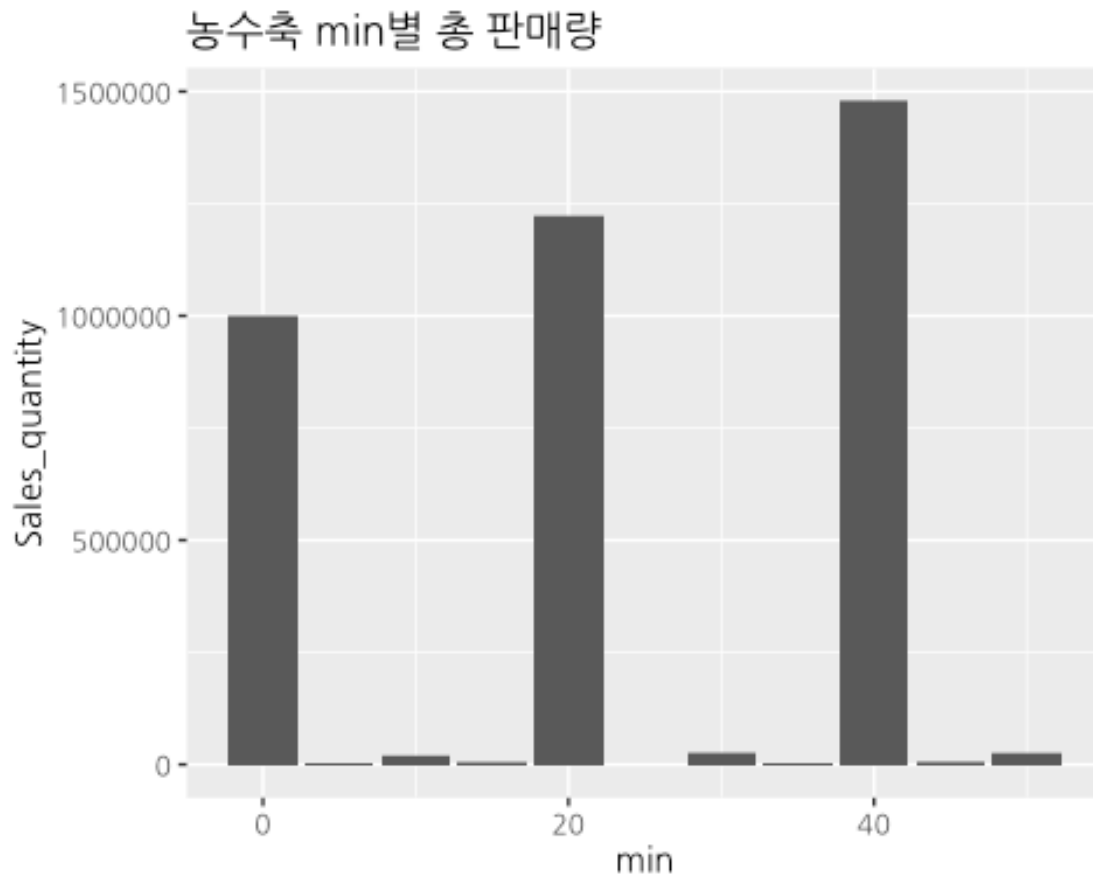
```
kitchen <-shop_data %>% filter(Product_group=="주방")  
  
ggplot(data=kitchen, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("주방 min 별 총 판매량")
```



농수축

```
agriculture <- shop_data %>% filter(Product_group=="농수축")

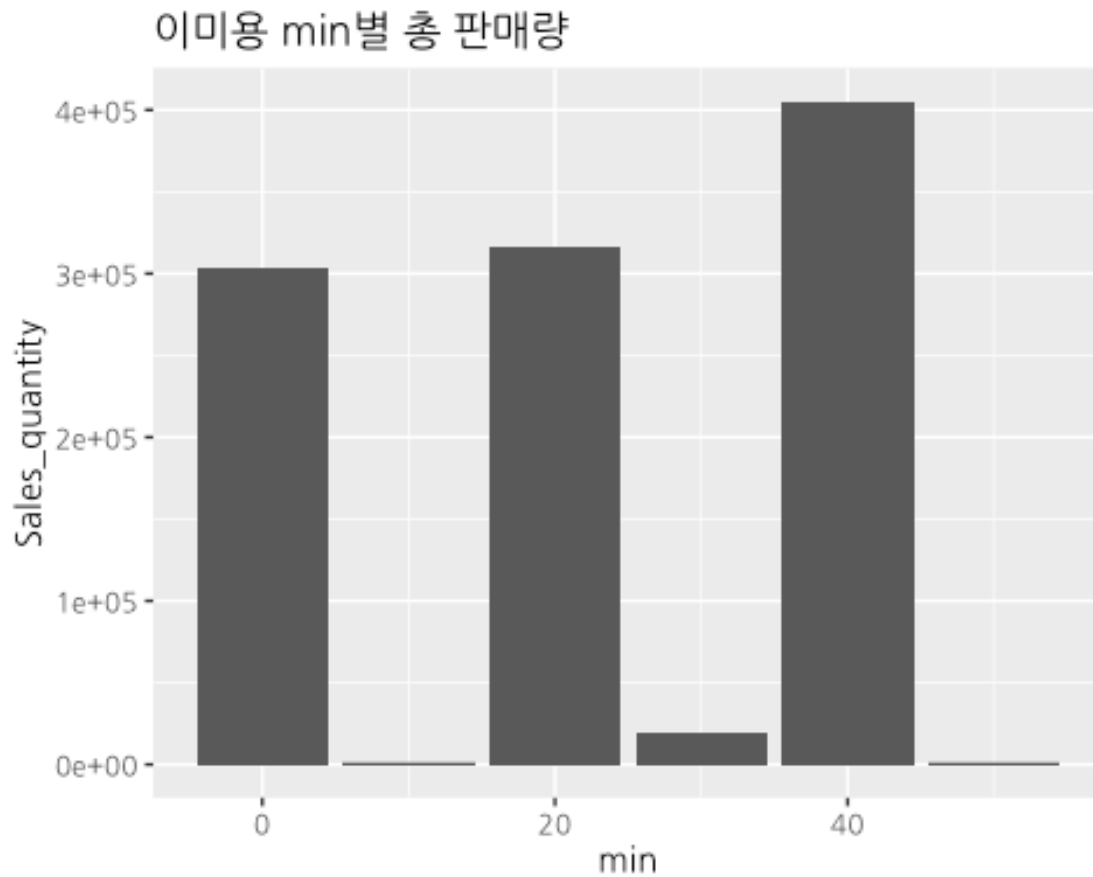
ggplot(data=agriculture, aes(x=min, y=Sales_quantity))+
  geom_col(position="stack")+
  ggtitle("농수축 min 별 총 판매량")
```



이미용

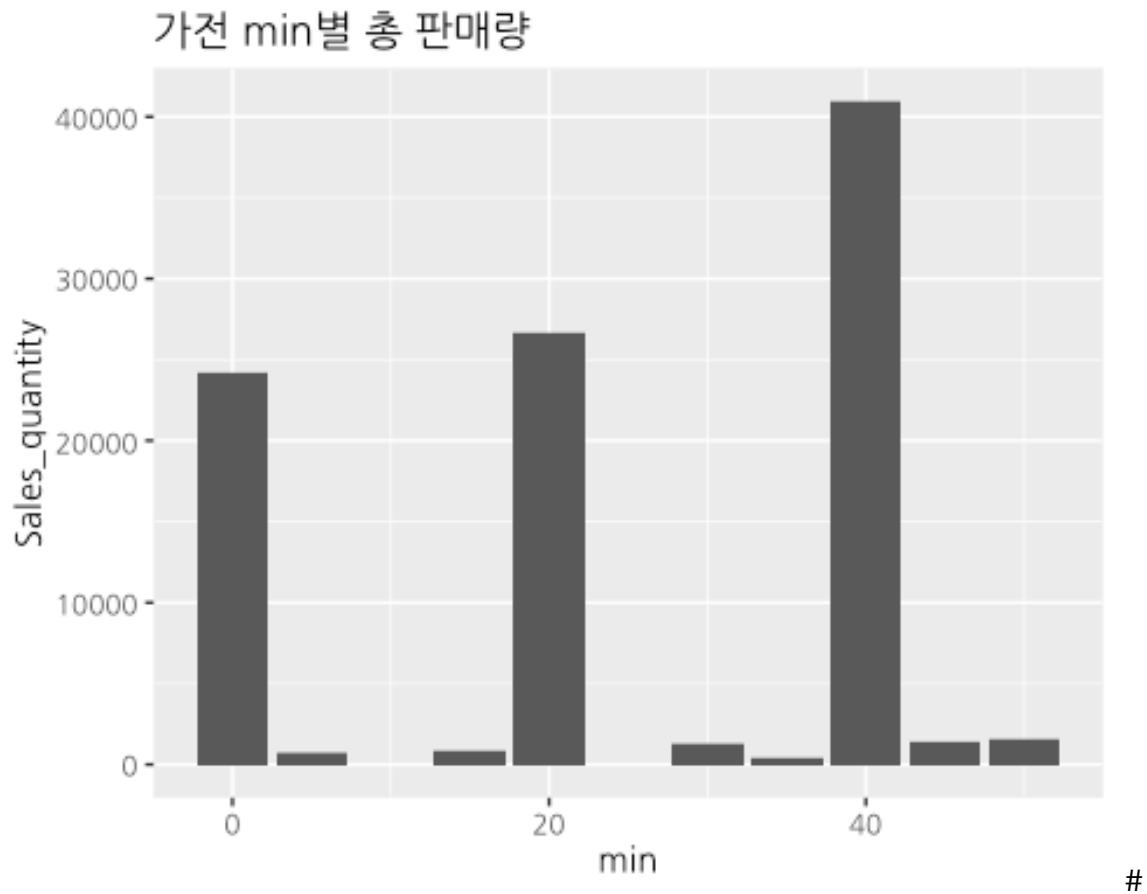
```
beauty <- shop_data %>% filter(Product_group=="이미용")
```

```
ggplot(data=beauty, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("이미용 min 별 총 판매량")
```



가전

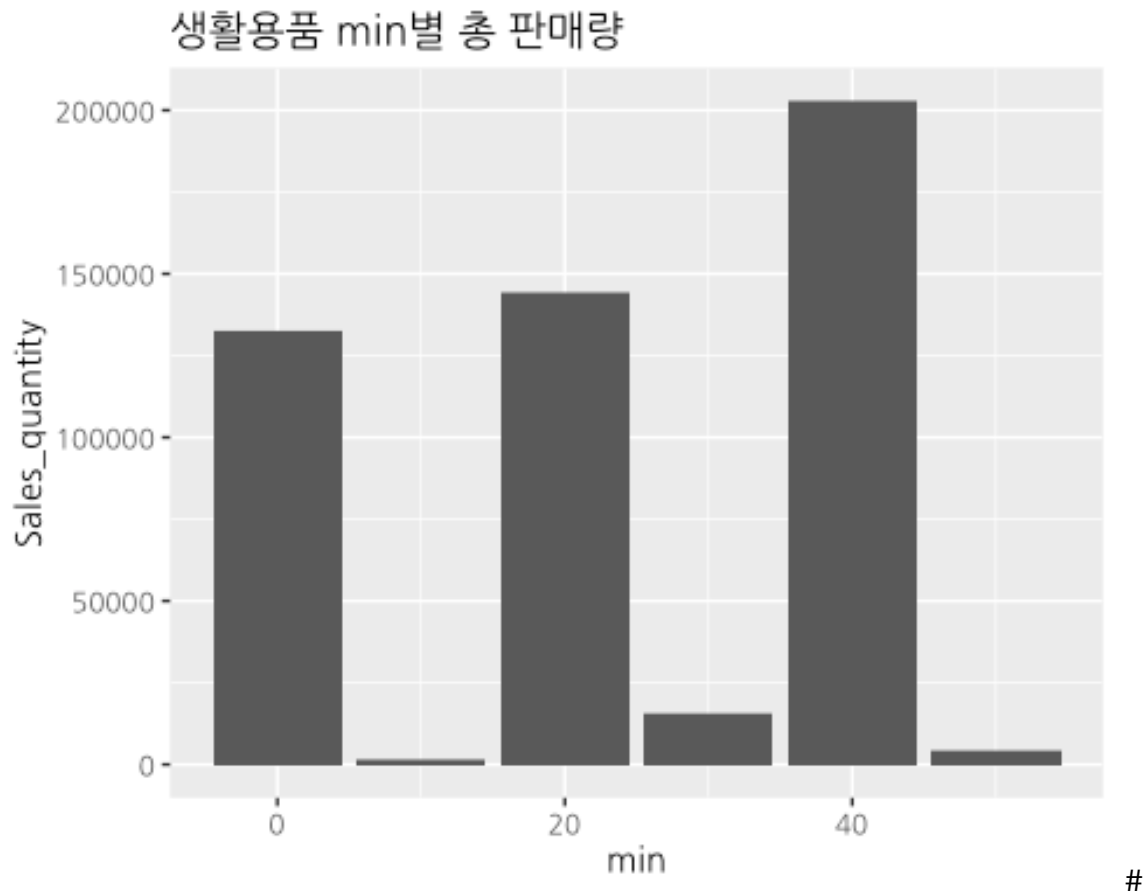
```
electronics <- shop_data %>% filter(Product_group=="가전")  
  
ggplot(data=electronics, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("가전 min 별 총 판매량")
```



생활용품 #

```
living <- shop_data %>% filter(Product_group=="생활용품")

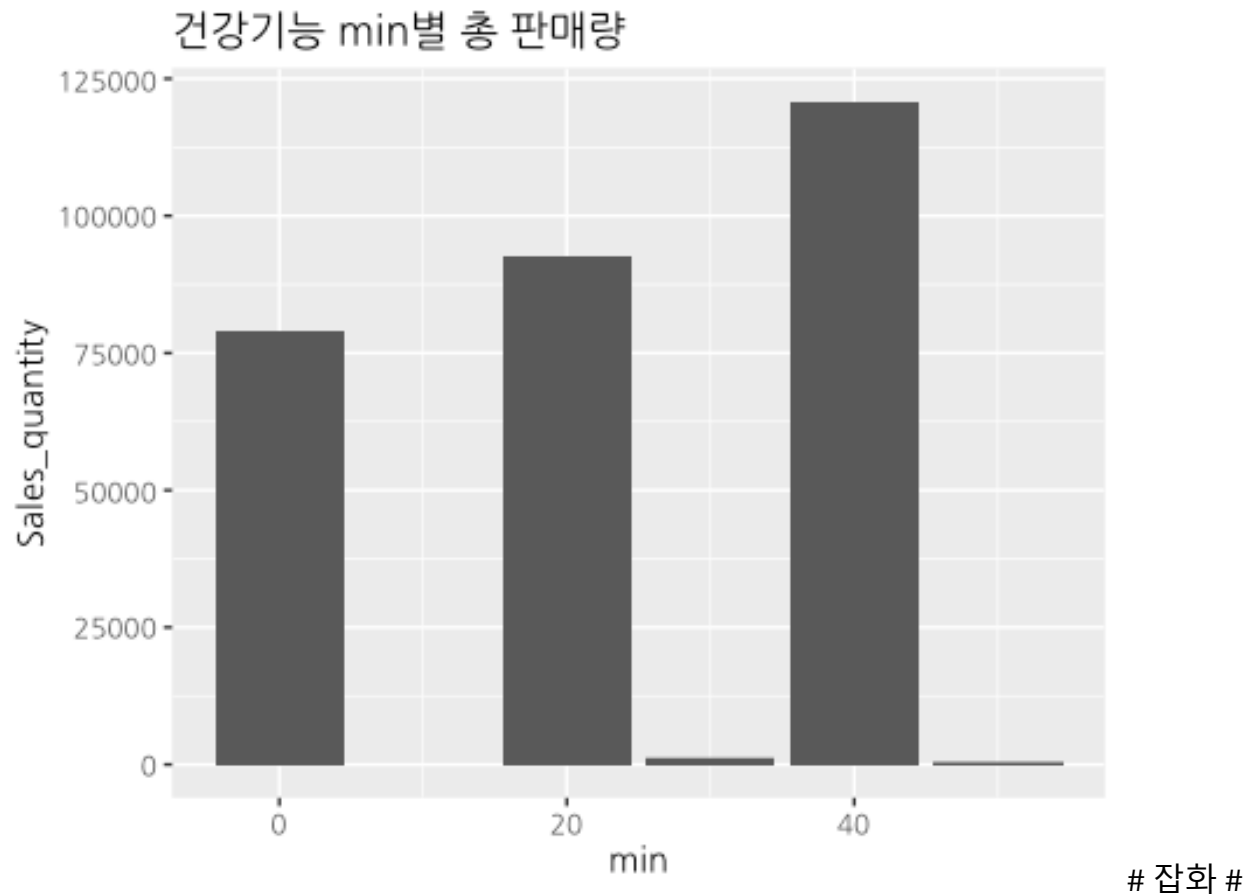
ggplot(data=living, aes(x=min, y=Sales_quantity))+
  geom_col(position="stack")+
  ggtitle("생활용품 min 별 총 판매량")
```



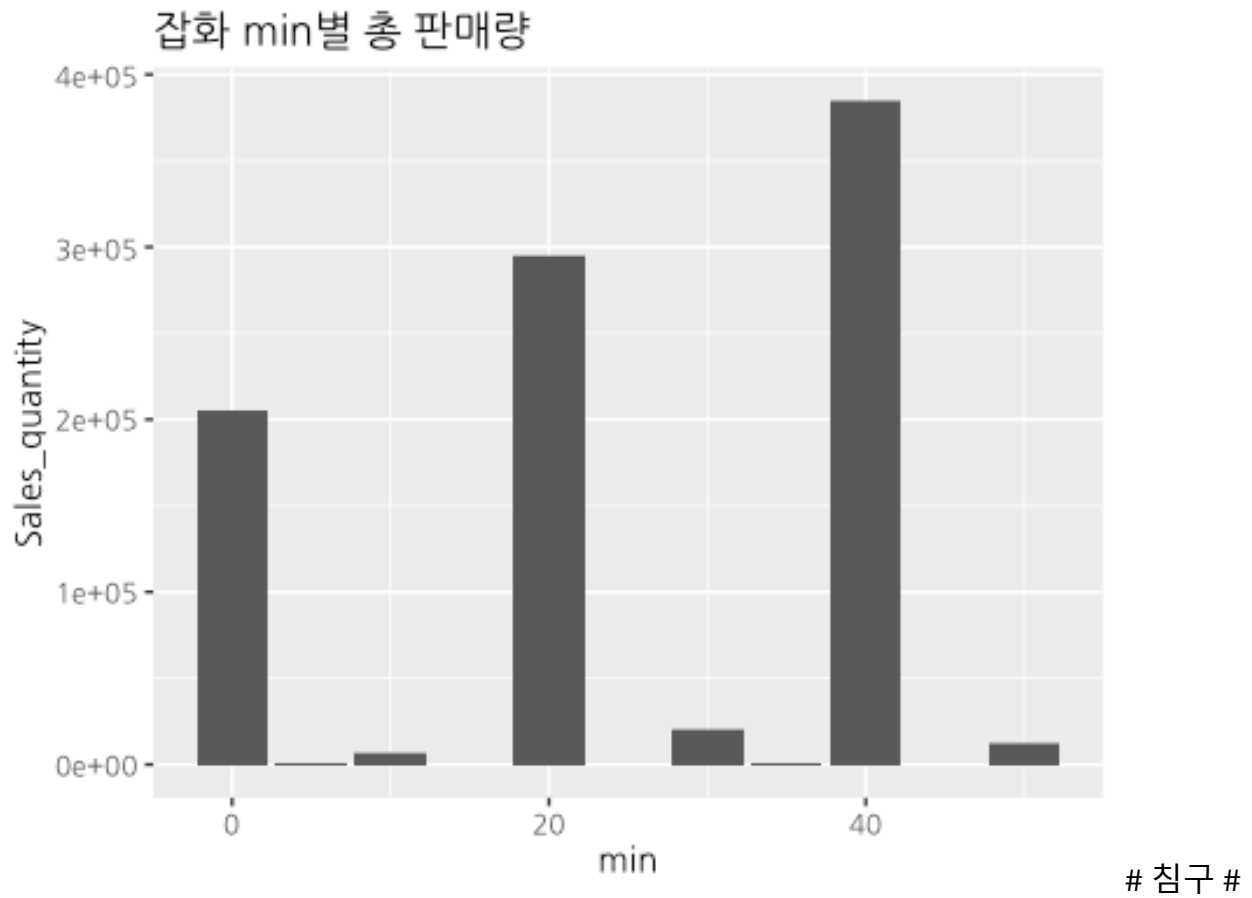
건강기능 #

```
health <- shop_data %>% filter(Product_group=="건강기능")
```

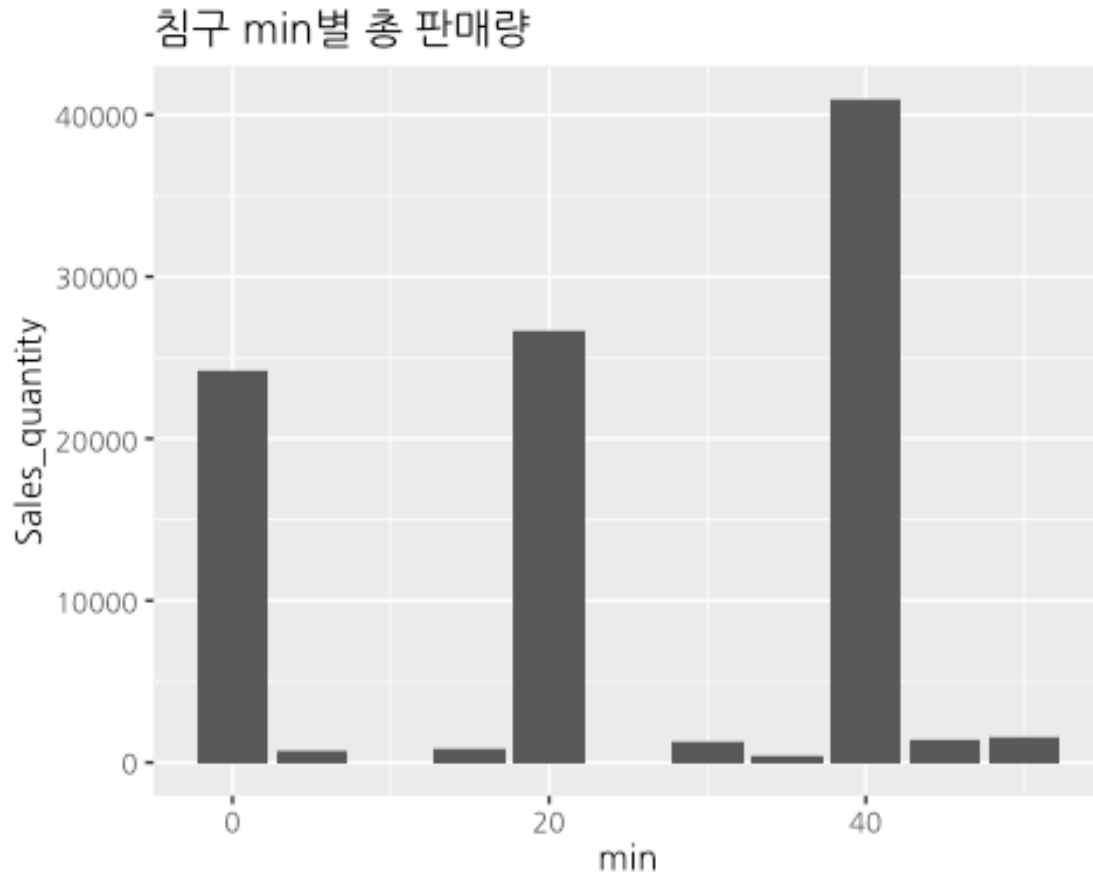
```
ggplot(data=health, aes(x=min, y=Sales_quantity))+
  geom_col(position="stack")+
  ggtitle("건강기능 min 별 총 판매량")
```



```
goods <- shop_data %>% filter(Product_group=="잡화")  
  
ggplot(data=goods, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("잡화 min 별 총 판매량")
```

```
bed <- shop_data %>% filter(Product_group=="침구")  
  
ggplot(data=electronics, aes(x=min, y=Sales_quantity))+  
  geom_col(position="stack")+  
  ggtitle("침구 min 별 총 판매량")
```



상품들의 min 별 판매량 추이는 비슷하다. 전반 < 중반 < 후반 의 형태를 보임.

정각에 방송시작인 경우가 많아 방영횟수는 0 > 20 > 40 순

전반과 중반은 차이가 크지 않은데 후반(40)에 판매량이 급증하는 경우가 많이 보임.