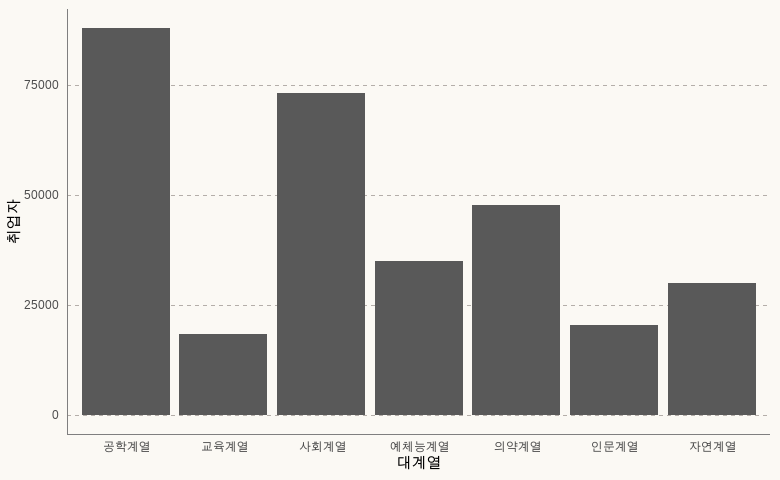
1. 비교(Compare)의 시각화

# 막대 그래프

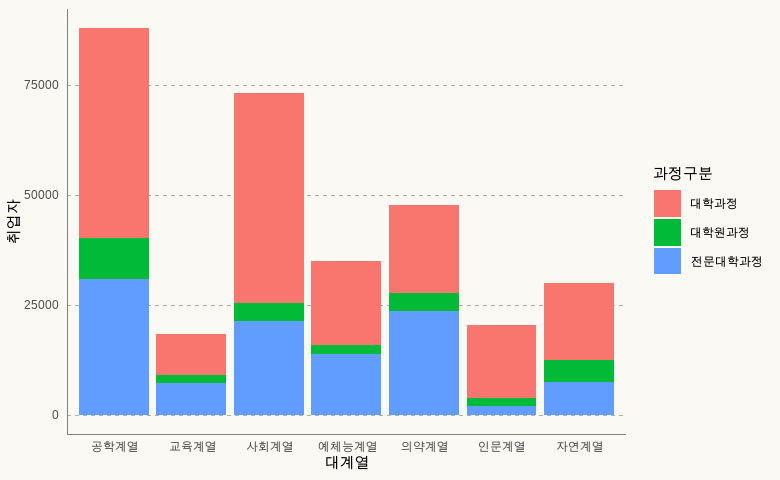
df\_취업통계\_계열별 <- df\_취업통계 |> group\_by(과정구분, 대계열, 중계열) |>   
 summarise(졸업자 = sum(졸업자\_계),   
 취업자 = sum(취업자\_합계\_계),   
 교외취업자 = sum(취업자\_교외취업자\_계),   
 교내취업자 = sum(취업자\_교내취업자\_계),   
 해외취업자 = sum(취업자\_해외취업자\_계),   
 농림어업종사자 = sum(취업자\_농림어업종사자\_계),   
 개인창작활동종사자 = sum(취업자\_개인창작활동종사자\_계),   
 일인창사업자 = sum(`취업자\_1인창(사)업자\_계`),   
 프리랜서 = sum(취업자\_프리랜서\_계),   
 진학자 = sum(진학자\_계),   
 입대자 = sum(입대자),  
 취업불가능자 = sum(취업불가능자\_계),   
 외국인유학생 = sum(외국인유학생\_계),   
 제외인정자 = sum(제외인정자\_계),   
 기타 = sum(기타\_계),   
 미상 = sum(미상\_계),   
 ## 백분률인 취업률은 그 자체로 합계나 평균을 낼 수 없으니 각 그룹별로 재계산  
 취업률 = 취업자 / (졸업자 - (진학자+입대자+취업불가능자+외국인유학생+제외인정자)))

## stack, dodge, fill

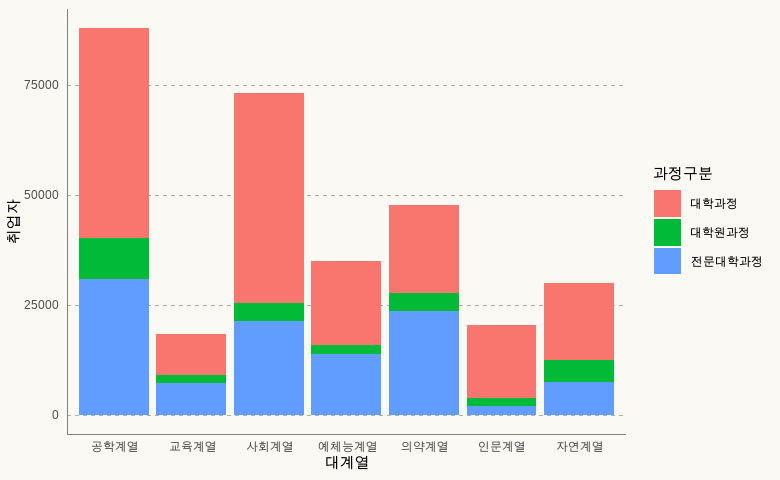
p\_col\_type <- df\_취업통계\_계열별 |>  
 mutate(대계열 = fct\_relevel(대계열, '인문계열', '사회계열', '교육계열', '자연계열', '공학계열', '의약계열', '예체능계열')) |>  
 ggplot()  
  
p\_col\_type +  
 geom\_col(aes(x = 대계열, y = 취업자))



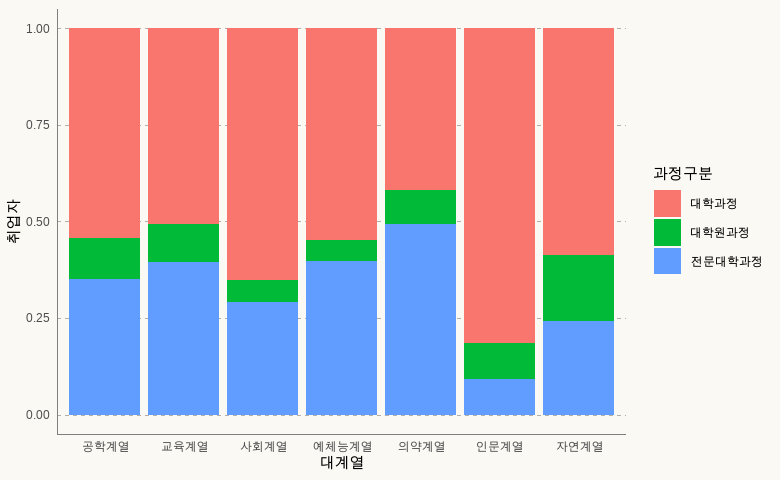
p\_col\_type +  
 geom\_col(aes(x = 대계열, y = 취업자, fill = 과정구분))



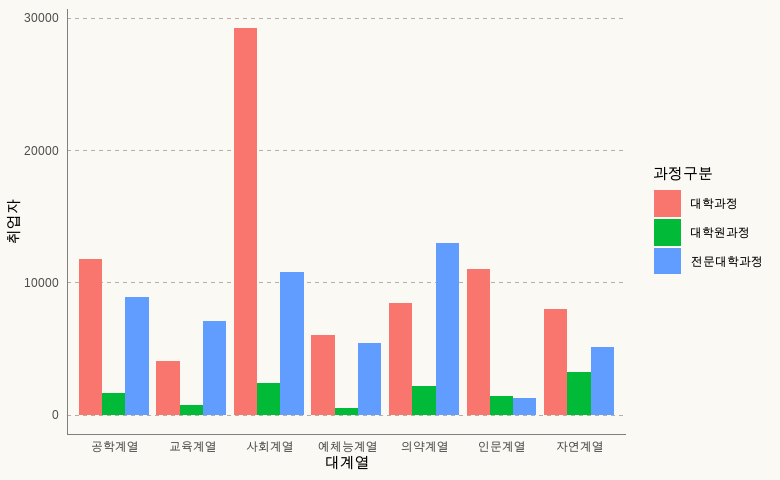
p\_col\_type +  
 geom\_col(aes(x = 대계열, y = 취업자, fill = 과정구분), position = 'stack')



p\_col\_type +  
 geom\_col(aes(x = 대계열, y = 취업자, fill = 과정구분), position = 'fill')

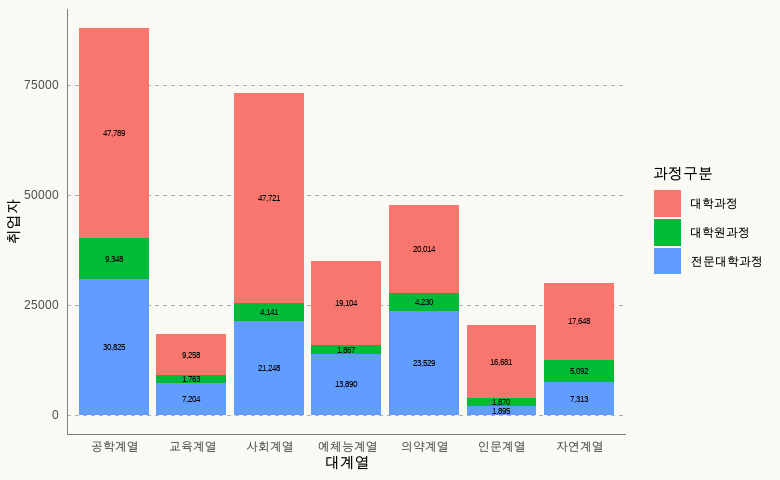


p\_col\_type +  
 geom\_col(aes(x = 대계열, y = 취업자, fill = 과정구분), position = 'dodge')

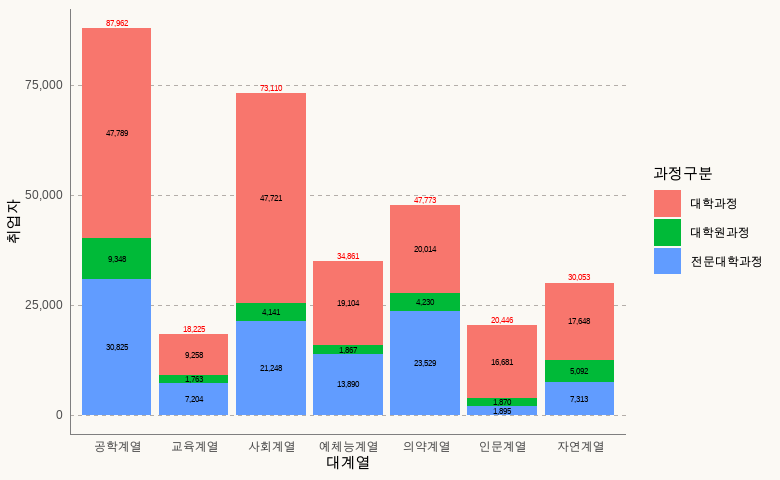


## 데이터 값 넣기

p\_col\_data <- df\_취업통계\_계열별 |> group\_by(과정구분, 대계열) |>  
 summarise(취업자 = sum(취업자)) |>  
 ggplot() +  
 geom\_col(aes(x = 대계열, y = 취업자, fill = 과정구분)) +  
 geom\_text(aes(x = 대계열, y = 취업자, label = scales::comma(취업자, accuracy = 1), fill =과정구분), position = position\_stack(vjust = 0.5), size = 2)  
  
p\_col\_data



p\_col\_data1 <- p\_col\_data +   
 stat\_summary(aes(x = 대계열, y = 취업자, label = scales::comma(..y.., accuracy = 1)), fun = 'sum', geom = 'text', color = 'red', vjust = -0.5, size = 2, inherit.aes = FALSE) +   
 scale\_y\_continuous(labels = scales::comma)  
  
p\_col\_data1



## 축 라벨에 이미지 넣기

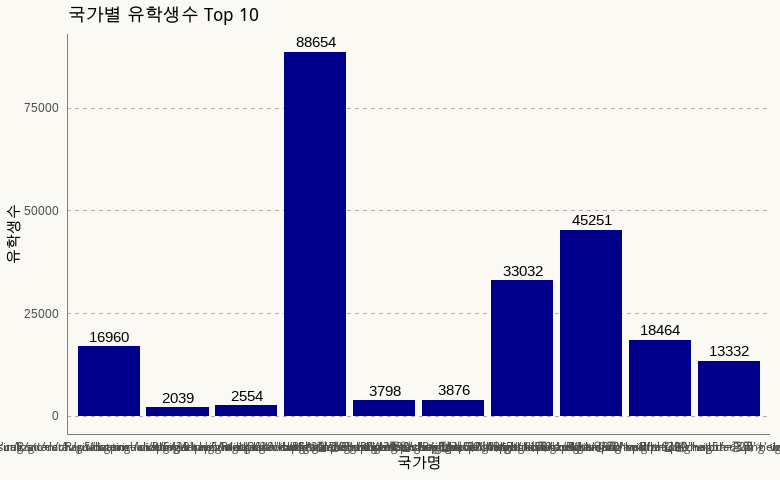
library(readxl)  
df\_nation <- read\_xlsx('파일경로/연도별 유학국가별 유학생수.xlsx', sheet = 'Sheet0', skip = 2, col\_types = c('numeric', 'text', rep('numeric', 25)), col\_names = TRUE)  
  
df\_nation <- df\_nation |>  
 filter(!is.na(학년도), 학제 == '소계') |>  
 select(!contains(c('계', '학제', '기타', '미확인', '그외동남아', '남미'))) |>  
 gather('국가명', '유학생수', -'학년도')  
  
df\_nation\_top10 <- df\_nation |>  
 group\_by(국가명) |>  
 summarise(sum = sum(유학생수)) |>  
 arrange(desc(sum)) |>  
 top\_n(10)  
  
df\_nation\_top10

## # A tibble: 10 x 2  
## 국가명 sum  
## <chr> <dbl>  
## 1 미국 88654  
## 2 캐나다 45251  
## 3 중국 33032  
## 4 필리핀 18464  
## 5 뉴질랜드 16960  
## 6 호주 13332  
## 7 일본 3876  
## 8 영국 3798  
## 9 말레이시아 2554  
## 10 독일 2039

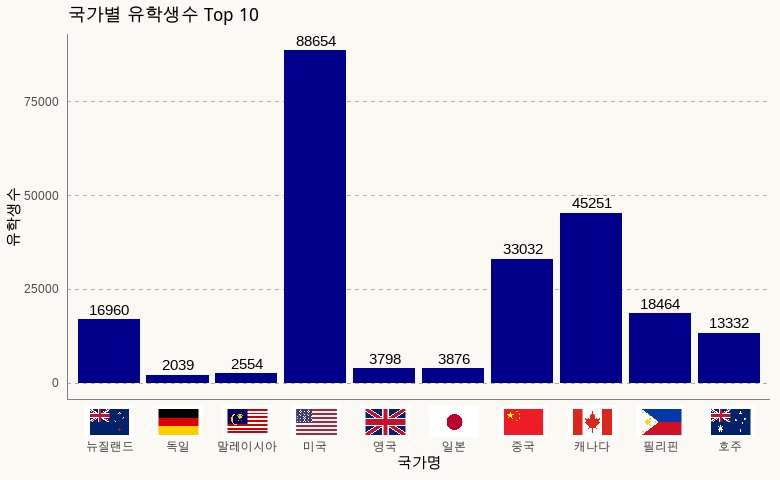
p\_nation\_top10 <- df\_nation\_top10 |>  
 ggplot(aes(x = 국가명, y = sum)) +  
 geom\_col(fill = 'dark blue') +  
 geom\_text(aes(x = 국가명, y = sum, label = sum), vjust = -0.5) +   
 labs(title = '국가별 유학생수 Top 10', y = '유학생수')

flag\_usa <- '아이콘 이미지 저장 폴더 경로/usa.png'  
flag\_canada <- '아이콘 이미지 저장 폴더 경로/can.png'  
flag\_china <- '아이콘 이미지 저장 폴더 경로/chi.png'  
flag\_phi <- '아이콘 이미지 저장 폴더 경로/phi.png'  
flag\_nz <- '아이콘 이미지 저장 폴더 경로/nz.png'  
flag\_aus <- '아이콘 이미지 저장 폴더 경로/aus.png'  
flag\_jap <- '아이콘 이미지 저장 폴더 경로/jap.png'  
flag\_eng <- '아이콘 이미지 저장 폴더 경로/eng.png'  
flag\_mal <- '아이콘 이미지 저장 폴더 경로/mal.png'  
flag\_ger <- '아이콘 이미지 저장 폴더 경로/ger.png'  
  
flags <- data.frame(nations = c('미국', '캐나다', '중국', '뉴질랜드', '필리핀', '호주', '영국', '일본', '말레이시아', '싱가폴'), flag\_path = c(flag\_usa, flag\_canada, flag\_china, flag\_nz, flag\_phi, flag\_aus, flag\_eng, flag\_jap, flag\_mal, flag\_sing))  
  
labels <- setNames(  
 paste0("<img src='", flags$flag\_path, "' width='30' height = '20'> <br> ", flags$nations), flags$nations)

p\_nation\_top10\_1 <- p\_nation\_top10 +  
 scale\_x\_discrete(labels = labels)  
  
p\_nation\_top10\_1

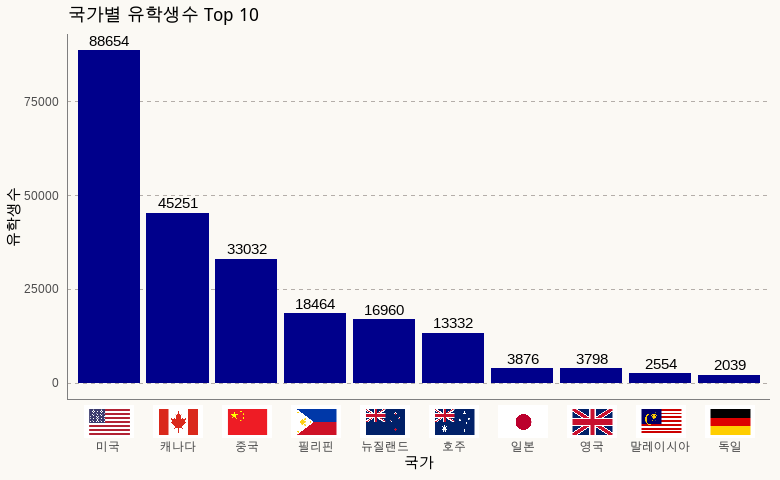


if(!require(ggtext)) {  
 install.packages('ggtext')  
 library(ggtext)  
}  
  
p\_nation\_top10\_1 +  
 theme(axis.text.x = ggtext::element\_markdown())



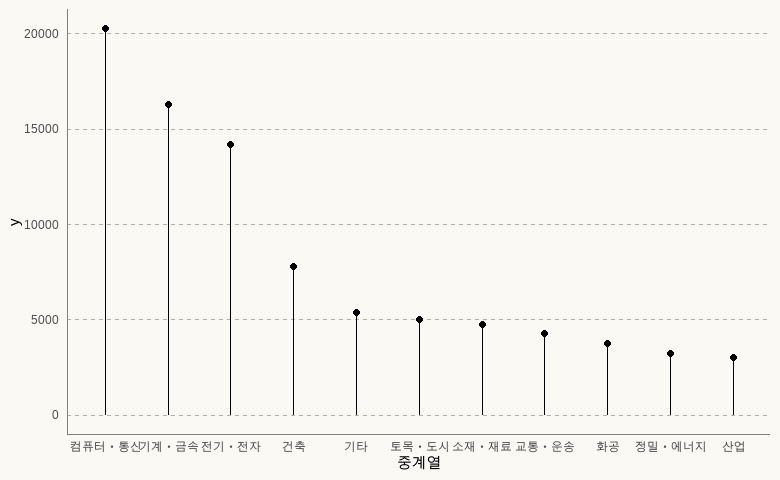
# 순위 막대 그래프

df\_nation\_top10$국가명 <- fct\_reorder(df\_nation\_top10$국가명, desc(df\_nation\_top10$sum))  
  
df\_nation\_top10 |>  
 ggplot(aes(x = 국가명, y = sum)) +  
 geom\_col(fill = 'dark blue') +  
 geom\_text(aes(x = 국가명, y = sum, label = sum), vjust = -0.5) +   
 labs(title = '국가별 유학생수 Top 10', x = '국가', y = '유학생수') +  
 scale\_x\_discrete(labels = labels) +  
 theme(axis.text.x = ggtext::element\_markdown())

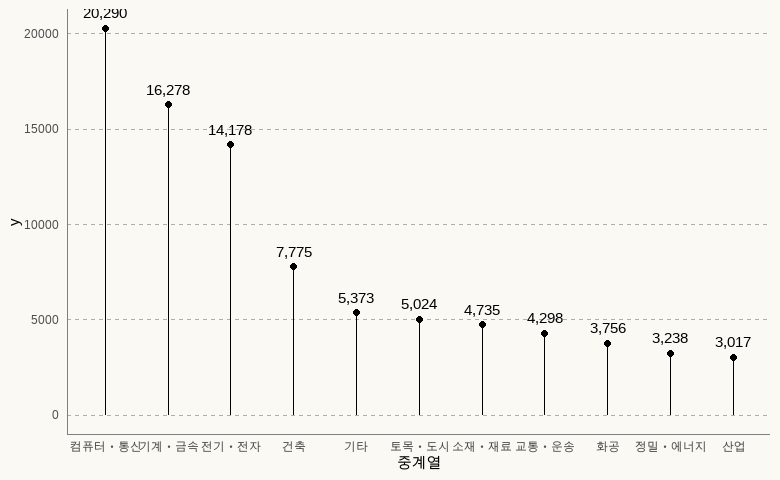


# 롤리팝 그래프

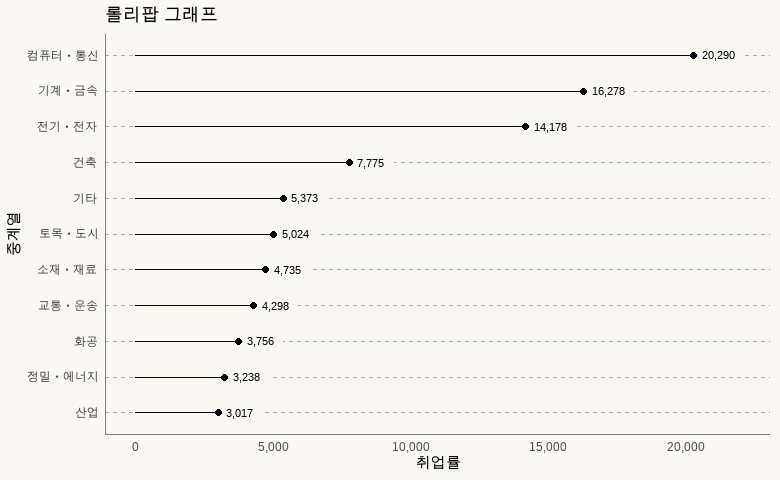
df\_lolipop <- df\_취업통계\_계열별 |> filter(대계열 == '공학계열') |> group\_by(중계열) |>  
 summarise(취업자 = sum(취업자)) |> arrange(desc(취업자))  
  
df\_lolipop$중계열 <- fct\_reorder(df\_lolipop$중계열, desc(df\_lolipop$취업자))  
  
p\_lolipop <- df\_lolipop|>  
 ggplot()  
  
p\_lolipop1 <- p\_lolipop +  
 geom\_segment(aes(x = 중계열, xend = 중계열, y = 0, yend = 취업자)) +  
 geom\_point(aes(x = 중계열, y = 취업자, group = 중계열))  
  
p\_lolipop1



p\_lolipop2 <- p\_lolipop1 +   
 geom\_text(aes(x = 중계열, y = 취업자, label = scales::comma(취업자)), vjust = -1)  
  
p\_lolipop2



p\_lolipop1 +   
 geom\_label(aes(x = 중계열, y = 취업자, label = scales::comma(취업자)), hjust = -0.1, label.size = NA, size = 3, fill = '#fbf9f4') +  
 scale\_x\_discrete(limits = rev) +  
 scale\_y\_continuous(labels = scales::comma, limits = c(0, 22000)) +  
 labs(title = '롤리팝 그래프', x = '중계열', y = '취업률') +   
 coord\_flip()



# 도트 그래프

p\_dot <- df\_lolipop|>  
 ggplot()  
  
p\_dot1 <- p\_dot +  
 geom\_segment(aes(x = 중계열, xend = 중계열, y = min(취업자), yend = max(취업자)), linetype = 'dashed', size = 0.1) +  
 geom\_point(aes(x = 중계열, y = 취업자, group = 중계열), size = 3, color = 'red') +   
 scale\_x\_discrete(limits = rev) +  
 coord\_flip() +   
 labs(title = '도트 그래프', x = '중계열', y = '취업자수')  
  
p\_dot1

