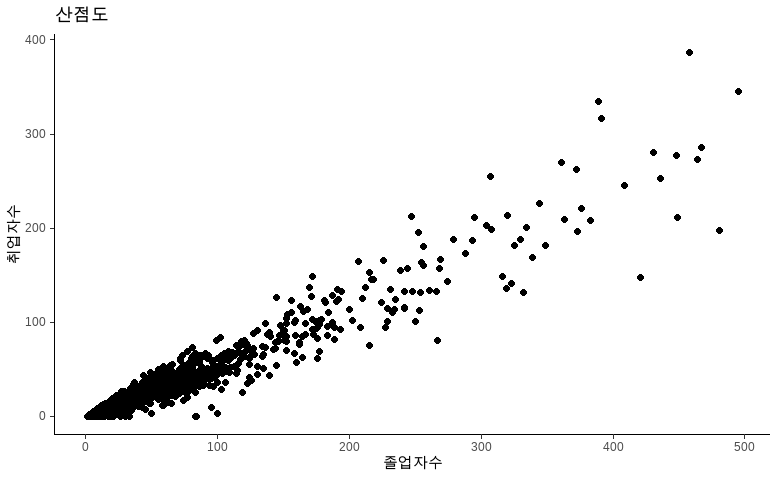
1. 관계의 시각화

# 스캐터 플롯(Scatter Plot)

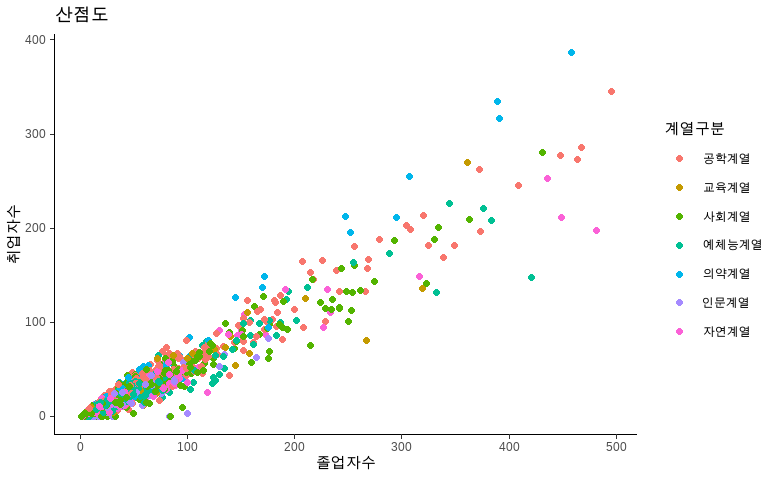
set.seed(123)  
df\_취업통계\_sample <- df\_취업통계 |> filter(취업률\_계 != 100, 졸업자\_계 >= 1) |>  
 filter(졸업자\_계 < 500) |>  
 sample\_n(2000)  
  
theme\_set(theme\_classic())

## 추세선 스캐터 플롯(scatter plot)

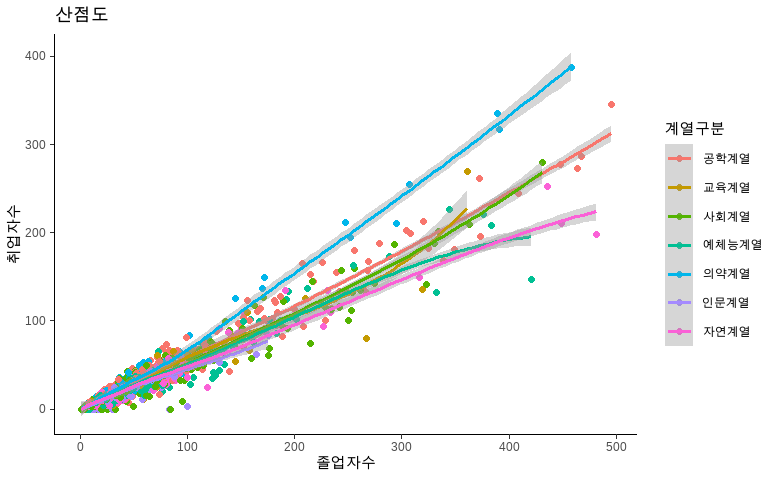
p\_scatter <- df\_취업통계\_sample |>   
 ggplot()  
  
p\_scatter1 <- p\_scatter +  
 geom\_point(aes(x = 졸업자\_계, y = 취업자\_합계\_계)) +   
 labs(title = '산점도', x = '졸업자수', y = '취업자수')  
  
p\_scatter1



p\_scatter2 <- p\_scatter1 +  
 geom\_point(aes(x = 졸업자\_계, y = 취업자\_합계\_계, color = 대계열)) +   
 labs(color = '계열구분')  
  
p\_scatter2

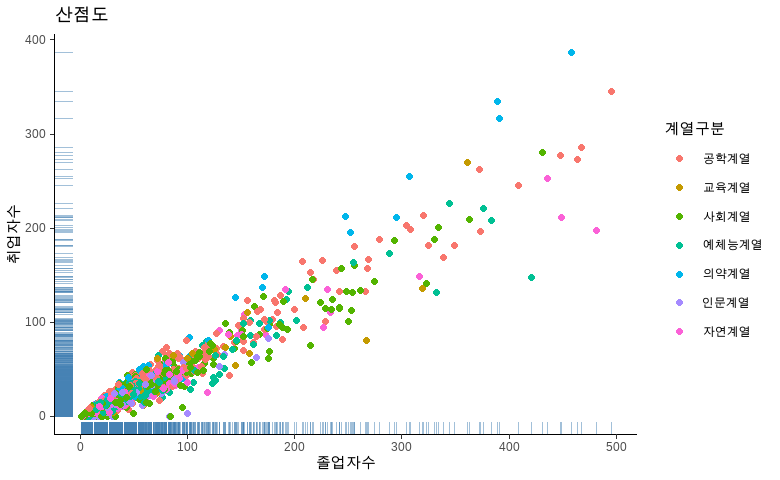


p\_scatter2 +   
 geom\_smooth(aes(x = 졸업자\_계, y = 취업자\_합계\_계, color = 대계열))



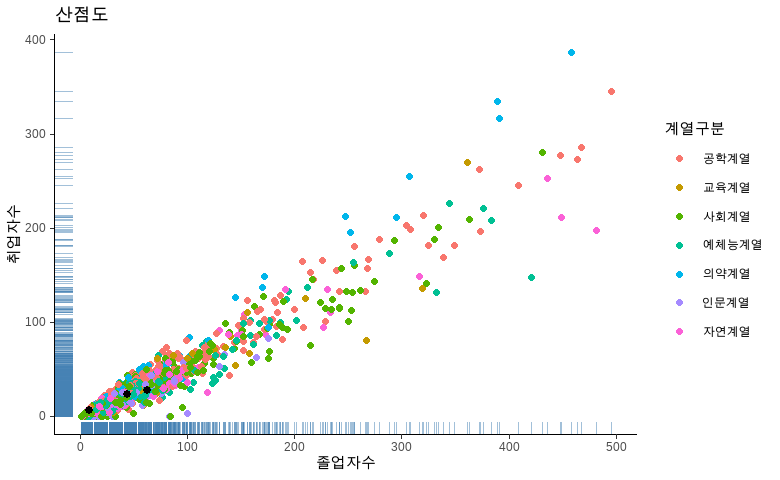
## 러그 스캐터 플롯(scatter plot)

if(!require('ggrepel')) {  
 install.packages('ggrepel')  
 library(ggrepel)  
}  
  
  
p\_scatter3 <- p\_scatter2 +   
 geom\_rug(aes(x = 졸업자\_계, y = 취업자\_합계\_계), col= "steelblue", alpha=0.5)  
  
p\_scatter3

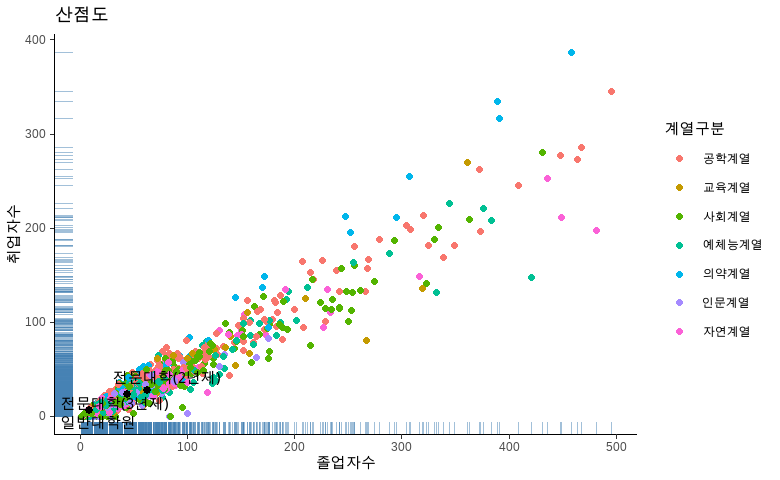


## 데이터 강조

p\_scatter4 <- p\_scatter3 +   
 geom\_point(data = df\_취업통계\_sample |> filter(학과명 == '컴퓨터정보학과'), aes(x = 졸업자\_계, y = 취업자\_합계\_계), color = 'black', size = 2)  
  
p\_scatter4

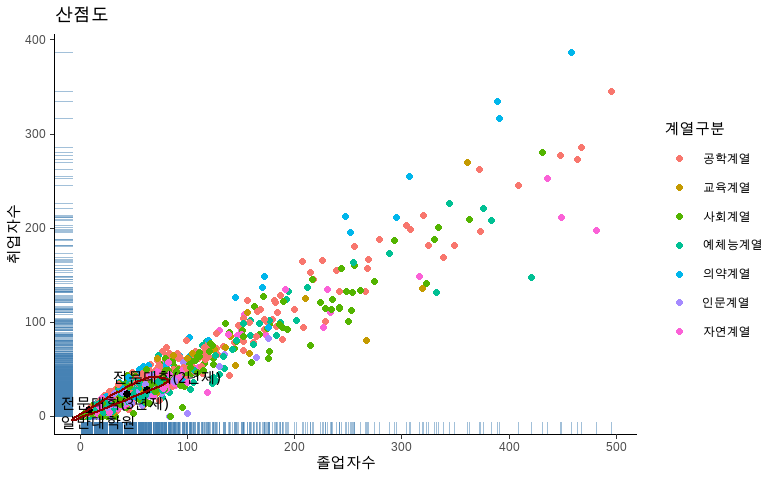


p\_scatter5 <- p\_scatter4 +  
 ggrepel::geom\_text\_repel(data = df\_취업통계\_sample |> filter(학과명 == '컴퓨터정보학과'), aes(x = 졸업자\_계, y = 취업자\_합계\_계, label = 학제))  
  
p\_scatter5



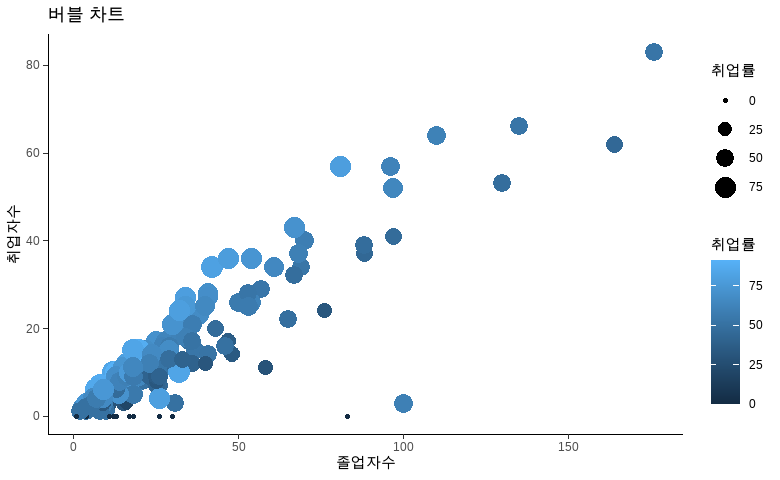
## encircle 스캐터 플롯(scatter plot)

if(!require('ggalt')) {  
 install.packages('ggalt')  
 library(ggalt)  
}  
  
p\_scatter6 <- p\_scatter5 +  
 geom\_encircle(data = df\_취업통계\_sample |> filter(학과명 == '컴퓨터정보학과'), aes(x = 졸업자\_계, y = 취업자\_합계\_계), color="darkred", size = 2)  
  
p\_scatter6



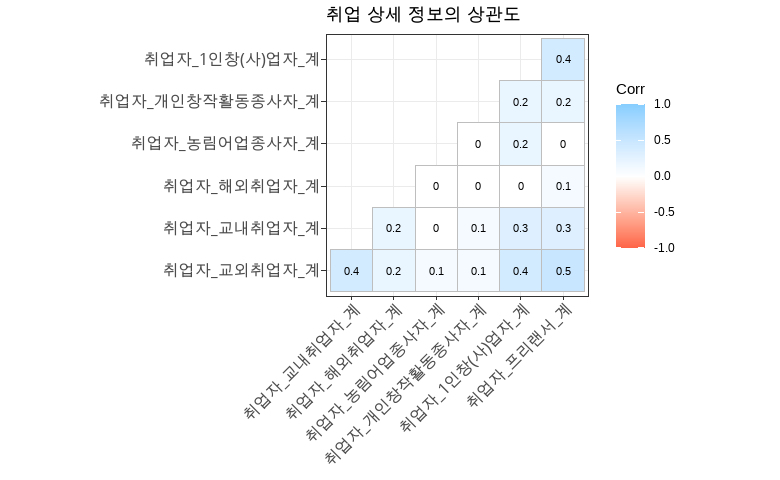
# 버블 차트(Bubble Chart)

p\_bubble <- df\_취업통계\_sample |> filter(대계열 == '인문계열') |>   
 ggplot()  
  
p\_bubble +  
 geom\_point(aes(x = 졸업자\_계, y = 취업자\_합계\_계, size = 취업률\_계, color = 취업률\_계)) +   
 labs(title = '버블 차트', x = '졸업자수', y = '취업자수', color = '취업률', size = '취업률')



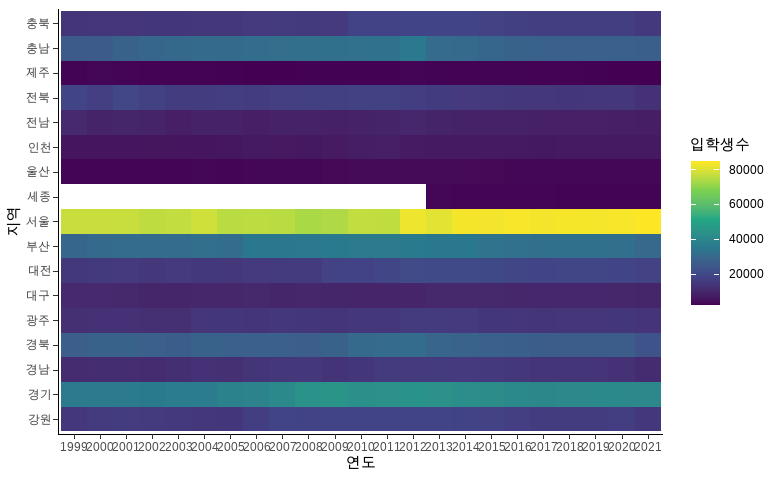
# 상관도

if(!require('ggcorrplot')) {  
 install.packages('ggcorrplot')  
 library(ggcorrplot)  
}  
  
corr <- round(cor(df\_취업통계\_sample[, 13:19]), 1)  
  
ggcorrplot(corr, type = "lower",   
 lab = TRUE,   
 lab\_size = 3,   
 method="square",   
 colors = c("tomato1", "white", "skyblue1"),   
 title="취업 상세 정보의 상관도",   
 ggtheme=theme\_bw)



# 히트맵

df\_입학자\_long |> filter(지역 != '전체') |> filter(학교종류 == '일반대학') |>  
 ggplot() +  
 geom\_tile(aes(x = 연도, y = 지역, fill = 입학생수)) +   
 scale\_fill\_continuous(type = "viridis")



# geom\_hex

df\_입학자\_long |> filter(지역 != '전체') |> filter(학교종류 == '일반대학') |>  
 ggplot() +  
 geom\_hex(aes(x = 연도, y = 지역)) +   
 scale\_fill\_continuous(type = "viridis")

