

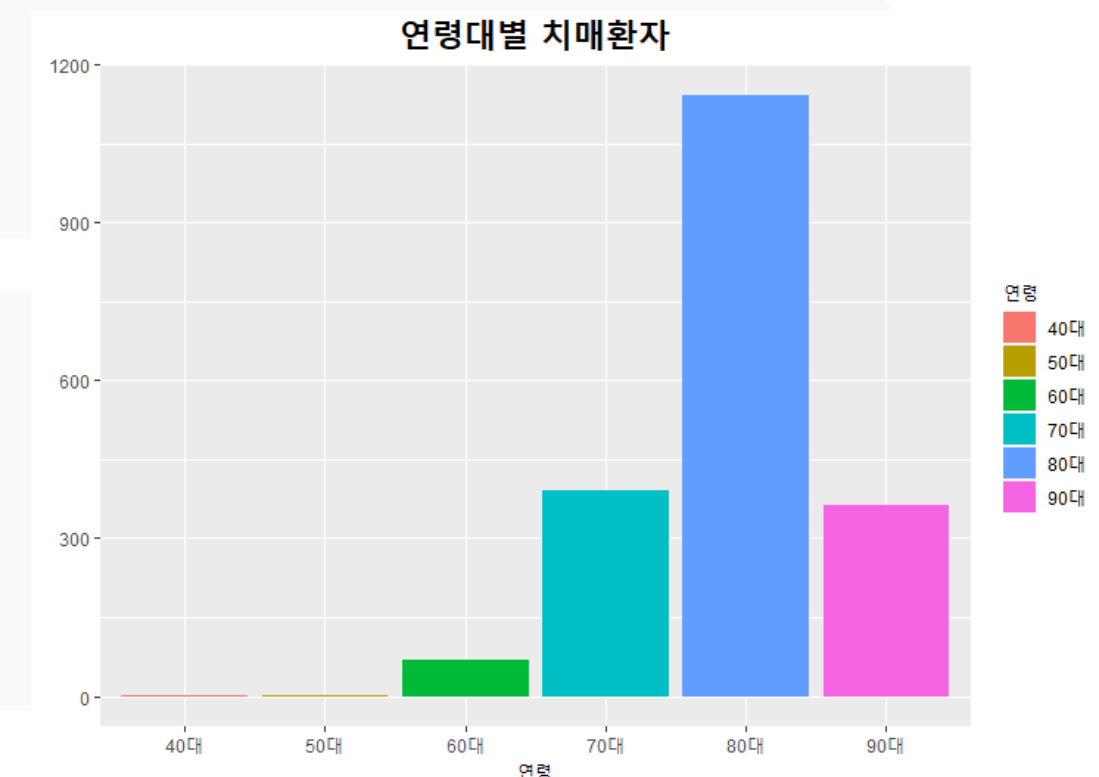
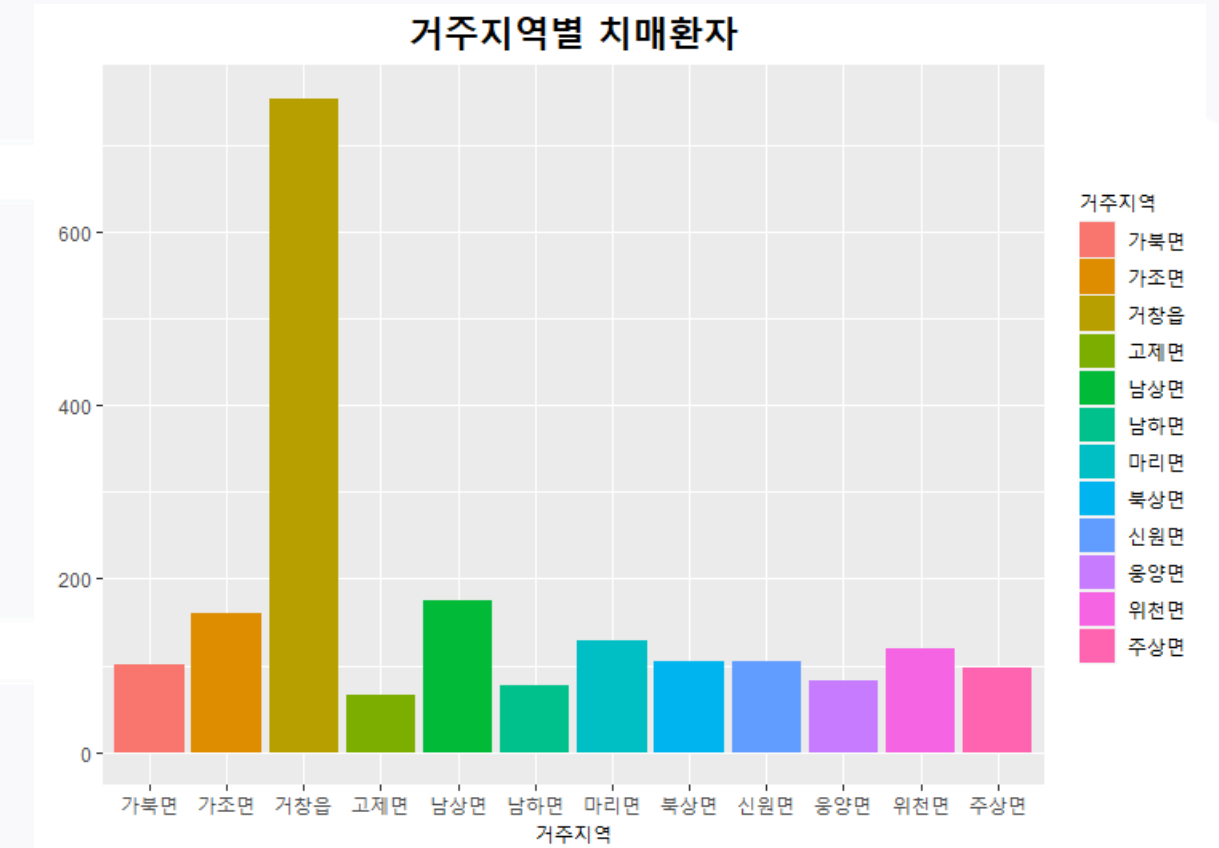
빅데이터 R 분석

김경민

• 해결문제

■ 치매환자현황분석

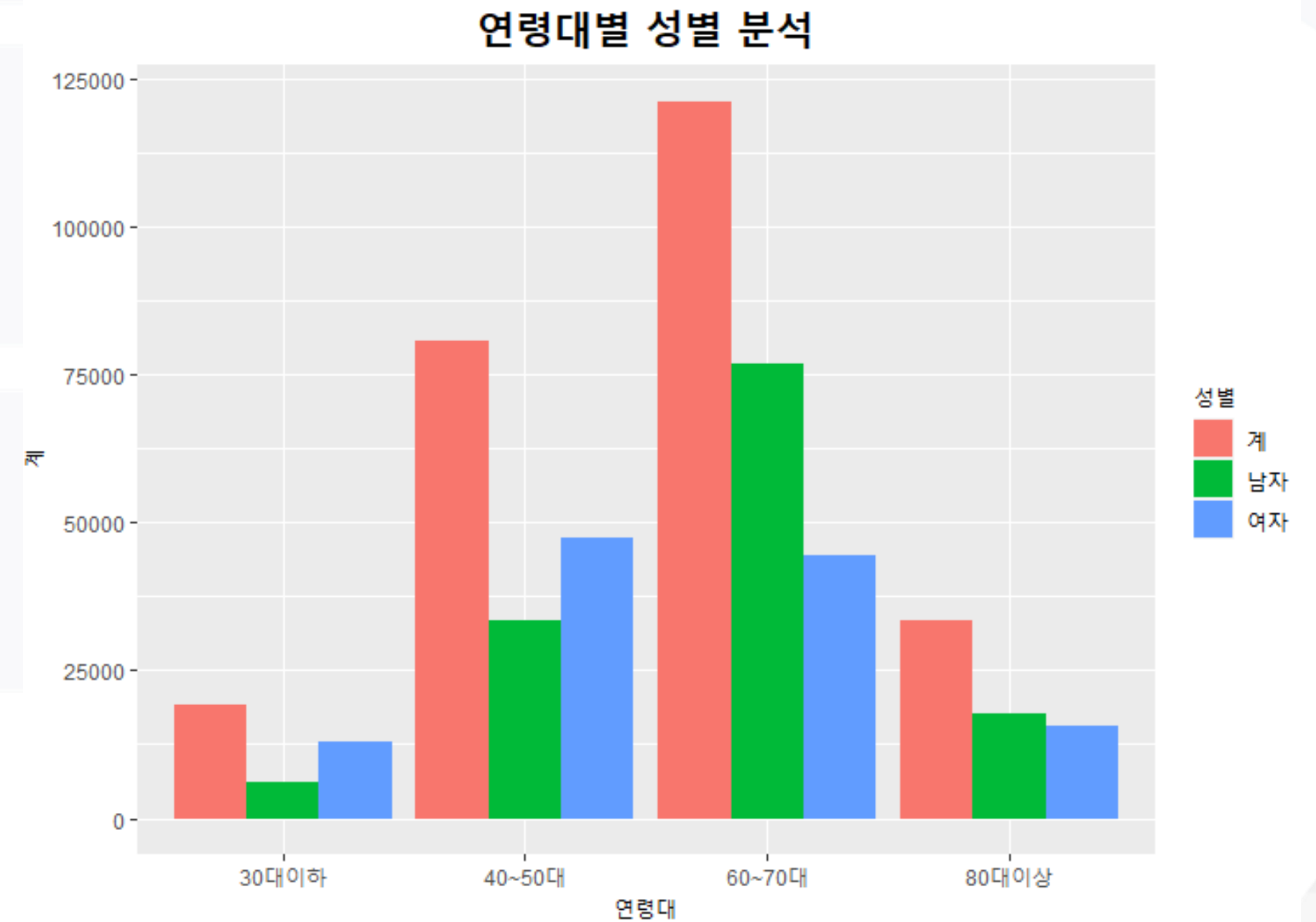
- 거주지역별 치매환자 빈도표를 구하고 그래프 작성
 - `qplot()`
- 기준일자와 진단일자를 이용하여 진단일수를 계산하고 평균 진단일수 산출
 - 날짜형으로 변환 : `as.Date()`
 - 날짜사이 간격 : `difftime(시작일, 종료일, units = 'days')`
 - `units = 'mins'` 분단위, `units = 'secs'` 초단위
- 연령대별 빈도수 그래프를 그려서 치매환자가 많은 연령대 분석
 - 열 내용에 문자열 붙이기 : `paste(열, "대", sep="")`
 - 100세 이상은 90대에 포함



• 해결문제

■ 연령대별 성별 암발생자 현황분석

- 모든 암에 대한 자료만 추출
 - dplyr 패키지 filter
- 연령대 별 그룹핑
 - dplyr 패키지 group_by
- 행인덱스 초기화
 - rownames(df) <- NULL



• 해결문제

■ 암 종별 성별 분석

- ggplot X축 회전
 - theme(axis.text.x=element_text(angle=90, hjust=1))

