

빅데이터 R 분석

김경민

• OPEN API 자료 수집

제공서비스

영화관입장권통합전산망이 제공하는 오픈API서비스 모음입니다.
사용 가능한 서비스를 확인하고 서비스별 인터페이스 정보를 조회합니다.

OPEN API

제공서비스

일별 박스오피스

주간/주말 박스오피스

공통코드 조회

영화목록

영화 상세정보

일별 박스오피스 API 서비스

특정 일자 상영작들의 박스오피스 정보를 영화구분(다양성영화, 상업영화), 한국/외국 구분, 상영지역 등의 조건을 통해 조회합니다.
REST/SOAP 방식 중 선택적으로 호출가능하며 REST 방식의 응답형식은 XML과 JSON을 지원합니다.(URI의 extension으로 구분)

1. REST 방식

- 기본 요청 URL : <http://www.kobis.or.kr/kobisopenapi/webservice/rest/boxoffice/searchDailyBoxOfficeList.xml> (또는 .json)
- 요청 parameter : 3번항의 요청 인터페이스 정보를 참조하여 GET 방식으로 호출

2. SOAP 방식

- 요청 URL : <http://www.kobis.or.kr/kobisopenapi/webservice/soap>
- WSDL URL : <http://www.kobis.or.kr/kobisopenapi/webservice/soap>
- Operation : searchDailyBoxOfficeList

```
#json처리하는데 많이 사용하는 패키지  
install.packages("jsonlite")  
library(jsonlite)
```

```
#일일박스 오피스 자료 가져오기  
url <- "https://kobis.or.kr/kobisopenapi/webservice/r  
  
mv <- fromJSON(url)
```

• 해결문제

개발계정 상세보기

기본정보

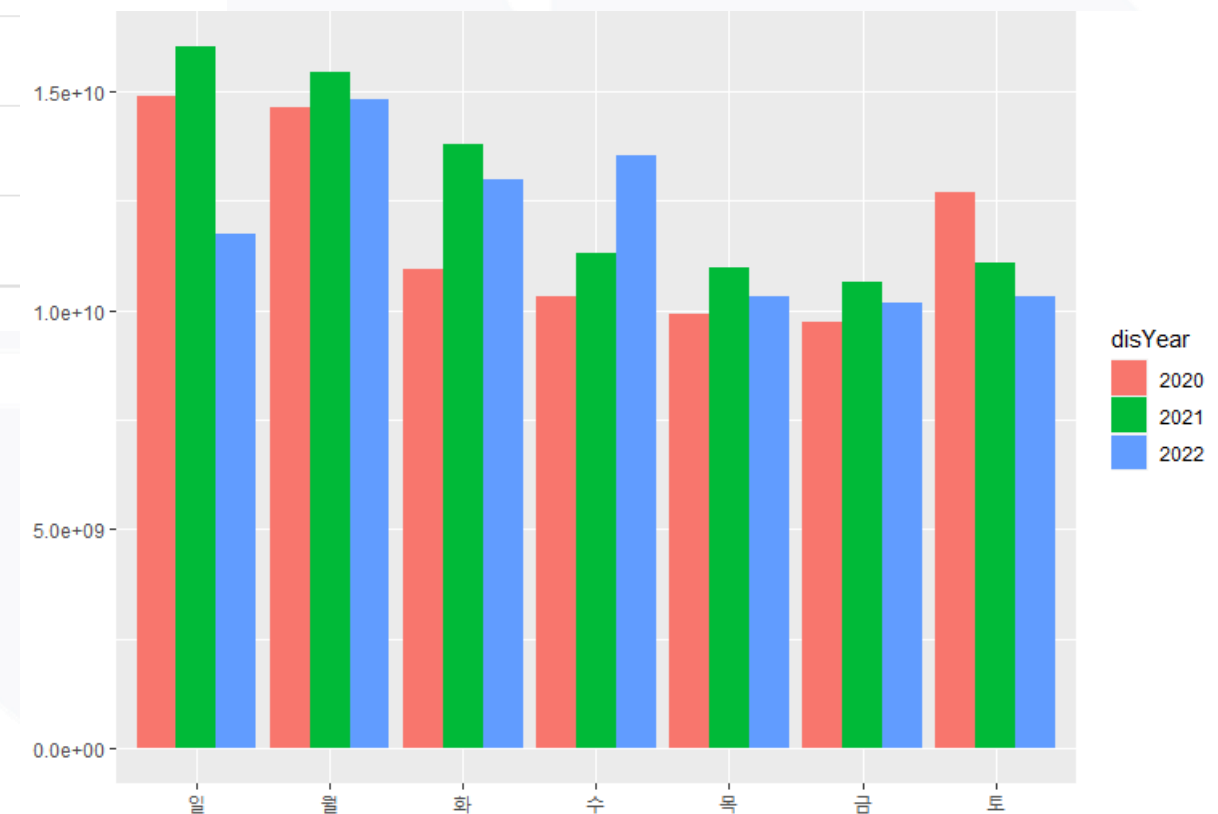
데이터명	한국환경공단_RFID기반 음식물쓰레기 배출정보	상세설명
서비스유형	REST	{ "data": { "resultCode": "OK", "resultMsg": "성공", "count": 7, "page": 1, "rowNum": 31, "list": [{ "disYear": "2022", "disMonth": "01", "disDay": 1, "disDayCount": 5, "disQuantity": 14638095165, "dayAverQuantity": 2927619033, "disCount": 8047170, "dayAverCount": 1609434 }, { "disYear": "2022", "disMonth": "01", "disDay": 2, "disDayCount": 5, "disQuantity": 14409623608, "dayAverQuantity": 2881924721, "disCount": 7489315, "dayAverCount": 1497863 }, { "disYear": "2022", "disMonth": "01", "disDay": 3, "disDayCount": 4, "disQuantity": 9317151956, "dayAverQuantity": 2329287989, "disCount": 5169658, "dayAverCount": 1292414 }, { "disYear": "2022", "disMonth": "01", "disDay": 4, "disDayCount": 4, "disQuantity": 9267072611, "dayAverQuantity": 2316768152, "disCount": 5299937, "dayAverCount": 1324984 }, { "disYear": "2022", "disMonth": "01", "disDay": 5, "disDayCount": 4, "disQuantity": 9245008626, "dayAverQuantity": 2311252156, "disCount": 5325313, "dayAverCount": 1331328 }, { "disYear": "2022", "disMonth": "01", "disDay": 6, "disDayCount": 4, "disQuantity": 9426482412, "dayAverQuantity": 2356620603, "disCount": 5345329, "dayAverCount": 1336332 }, { "disYear": "2022", "disMonth": "01", "disDay": 7, "disDayCount": 5, "disQuantity": 12353488582, "dayAverQuantity": 2470697716, "disCount": 6790592, "dayAverCount": 1358118 }] } }
신청유형	개발계정 활용	
활용기간	2021-11-24 ~	

년도와 월을 입력받아 json 자료를 받는 함수를 작성하고
2020년 ~ 2020년 8월 자료를 비교하시오.

서비스정보

참고문서	IROSS_OA_DV_0401_OpenAPI활용가이드_RFID기반 음식물쓰레기 데이터 서비스(한국환경공단)_v1.8.pdf
데이터포맷	JSON+XML
End Point	http://apis.data.go.kr/B552584/RfidFoodWasteServiceNew

API 환경 또는 API 호출 조기에 따라 인증키가 적용되는 방식이 다를 수 있습니다.



• 기본 그래프

■ plot(벡터) : 가장 기본적인 함수

- 1차원 산점도
- main : 그래프 제목
- xlab , ylab : x, y축 제목
- xlim, ylim : c(a, b) x, y축 범위
- pch : 마커종류
- col : 색상 (1 ~ 8)
 - black, red, green, blue, cyan, magenta, yellow, gray
- type : 연결 (p, l, b, c, o, h, s, S, n)
- lty : 선종류 (0~6)
 - blank, soild, dashed, dotted, dotdash, longdash, twodash

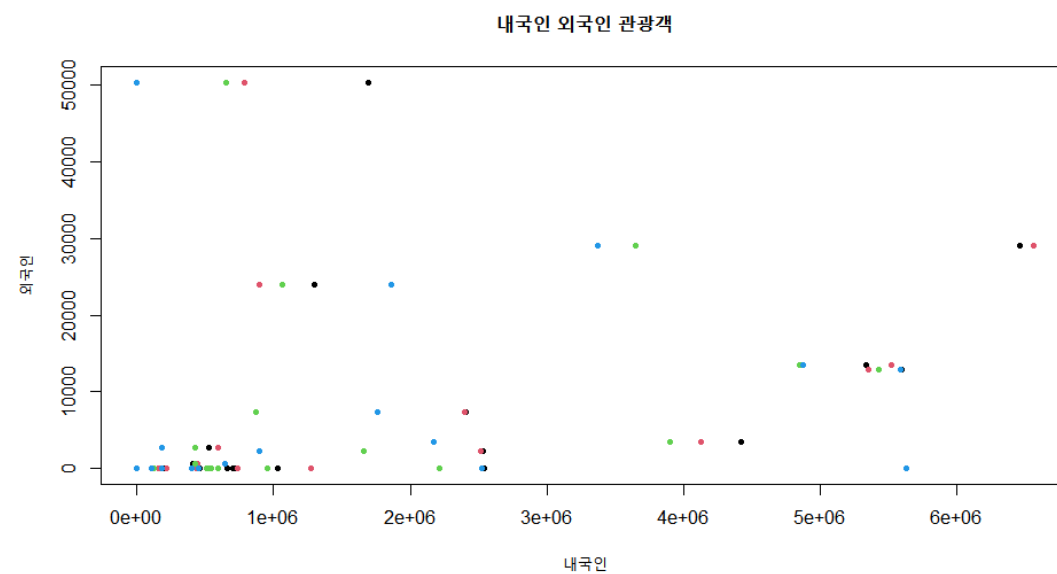
■ barplot(벡터) : 막대 그래프

시각화

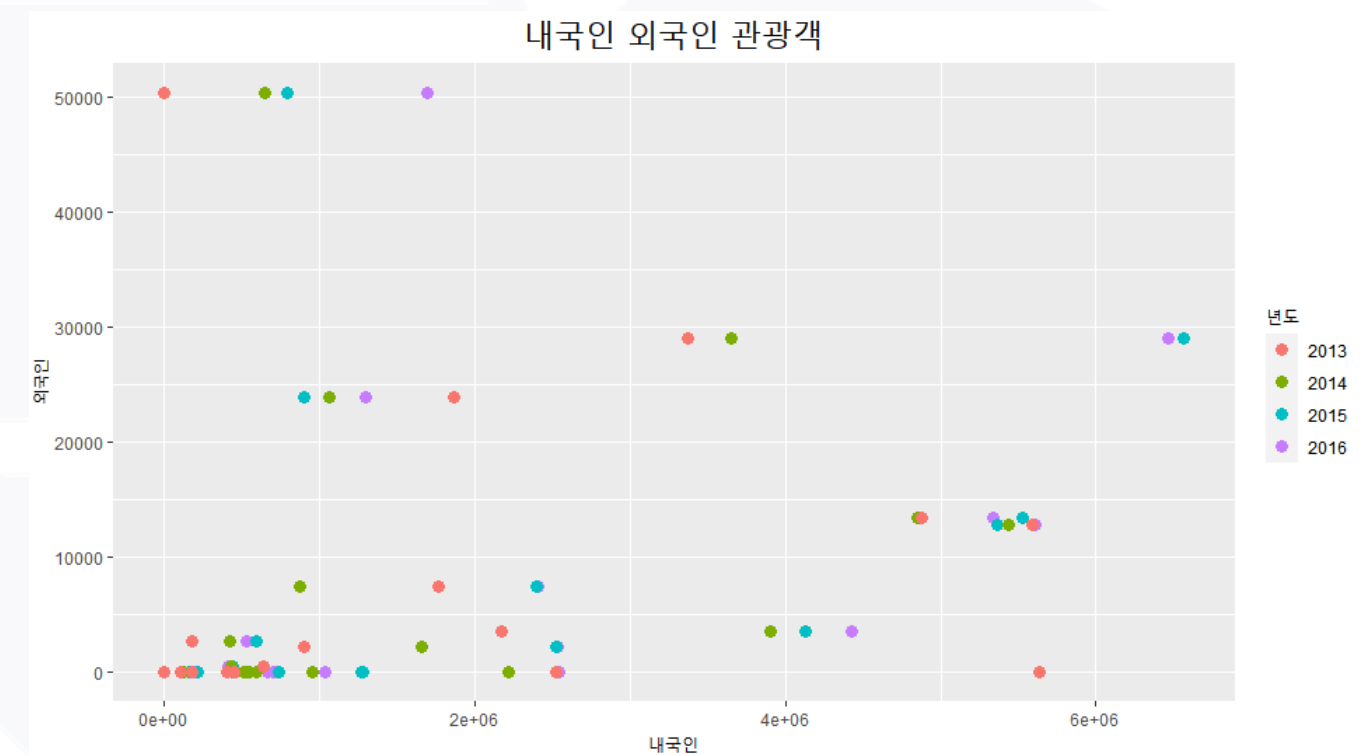
ggplot2

- 데이터프레임에 존재하는 데이터에서 복잡한 그래프 생성을 간단히 만들 수 있음

```
plot(df$내국인, df$외국인  
  , main = "내국인 외국인 관광객"  
  , xlab = "내국인", ylab = "외국인"  
  , pch = 20  
  , col = ifelse(df$년도 == "2016", 1,  
                 ifelse(df$년도 == "2015", 2,  
                 ifelse(df$년도 == "2014", 3,4)))  
)
```



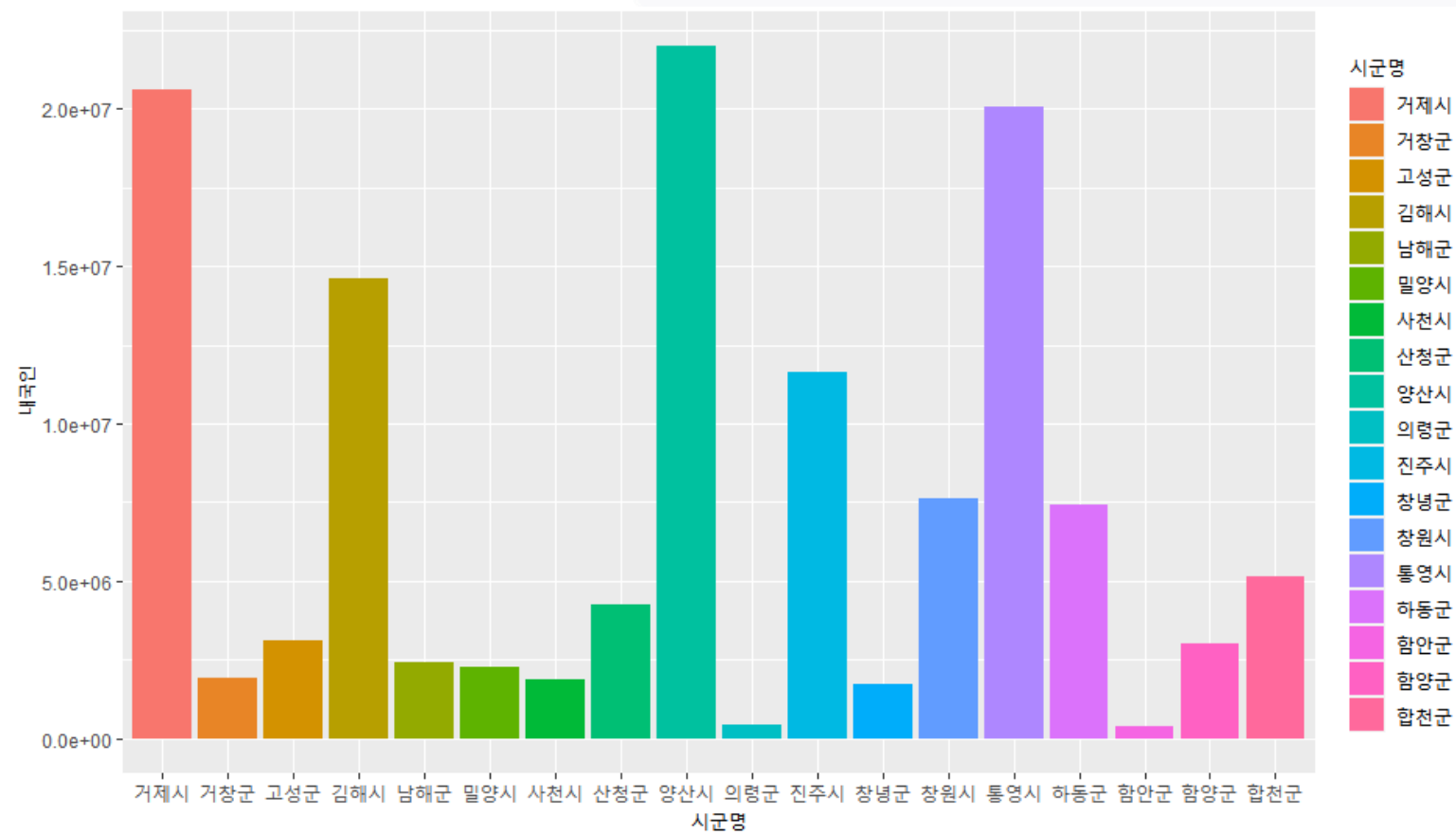
```
ggplot(data=df, aes(x=내국인, y=외국인)) +  
  geom_point(aes(colour=년도), size=3) +  
  ggtitle("내국인 외국인 관광객")+  
  theme(plot.title = element_text(hjust = 0.5, size=20))
```



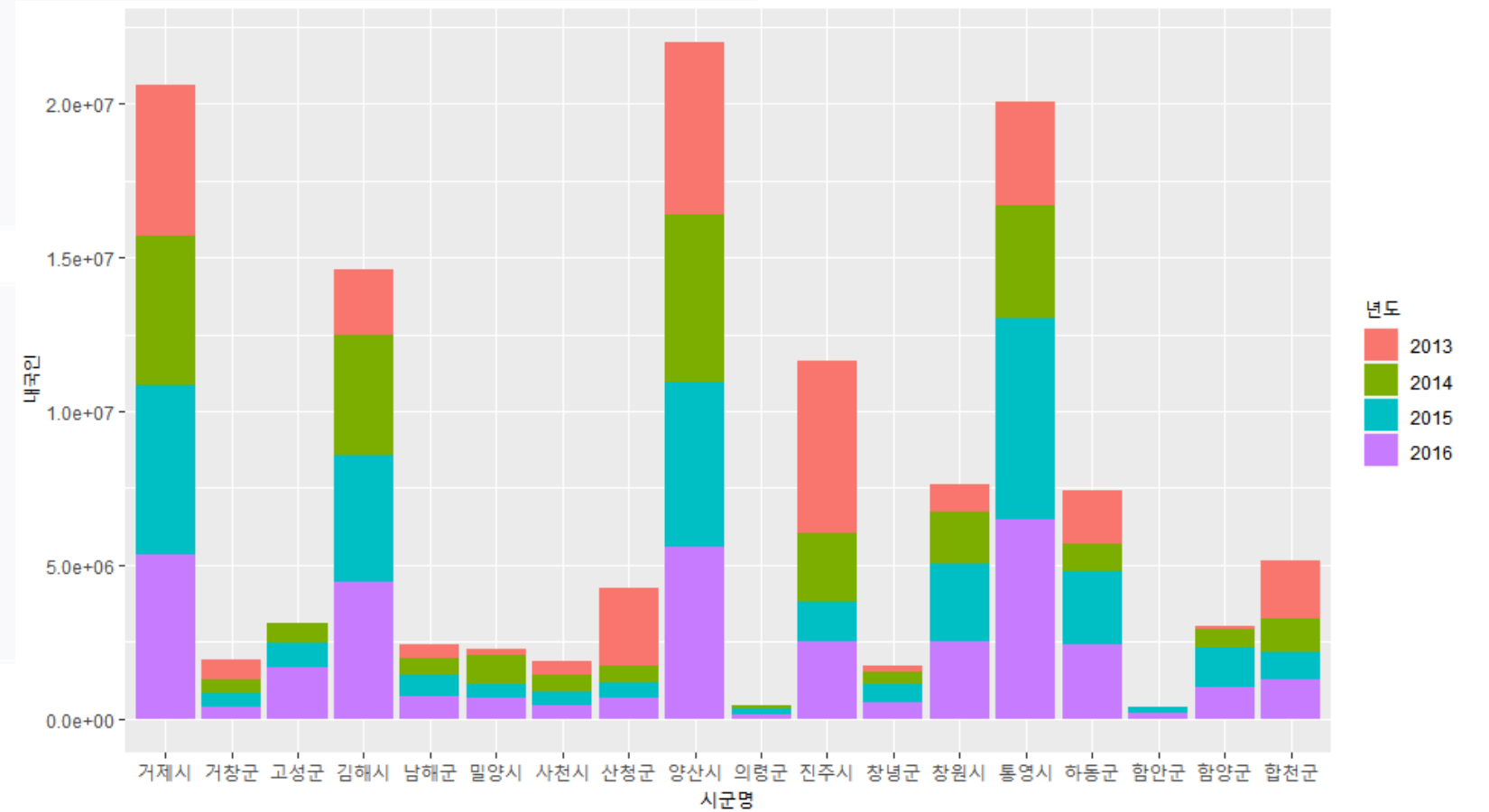
시각화

ggplot2 막대그래프

```
ggplot(data=df, aes(x=시 군 명, y=내 국 인, fill=시 군 명)) +  
  geom_col()
```



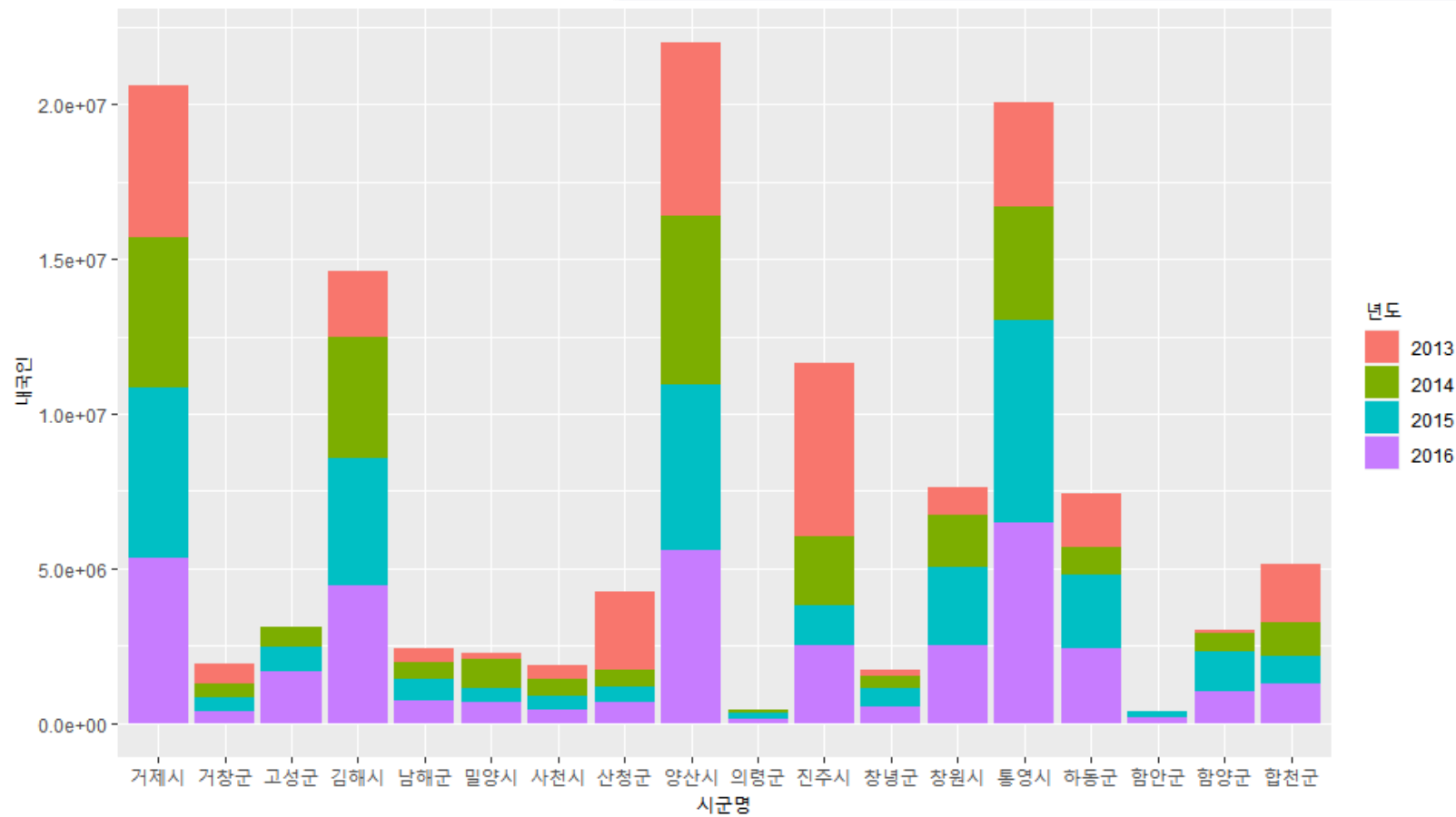
```
ggplot(data=df, aes(x=시 군 명, y=내 국 인, fill=년 도)) +  
  geom_bar(stat = "identity", position='dodge')
```



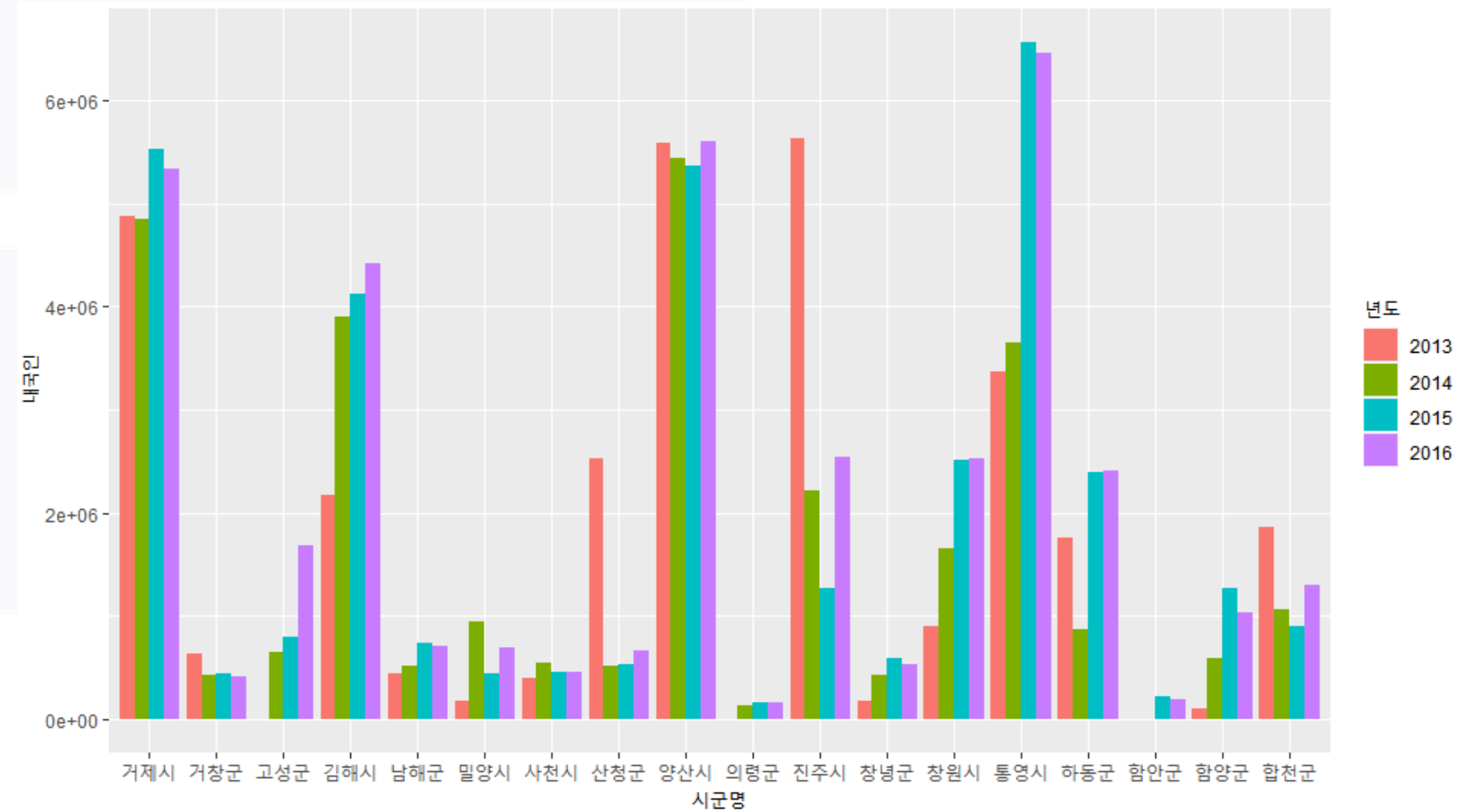
시각화

ggplot2 막대그래프

```
ggplot(data=df, aes(x=시군명, y=내국인, fill=년도)) +  
  geom_bar(stat = "identity", position='dodge')
```



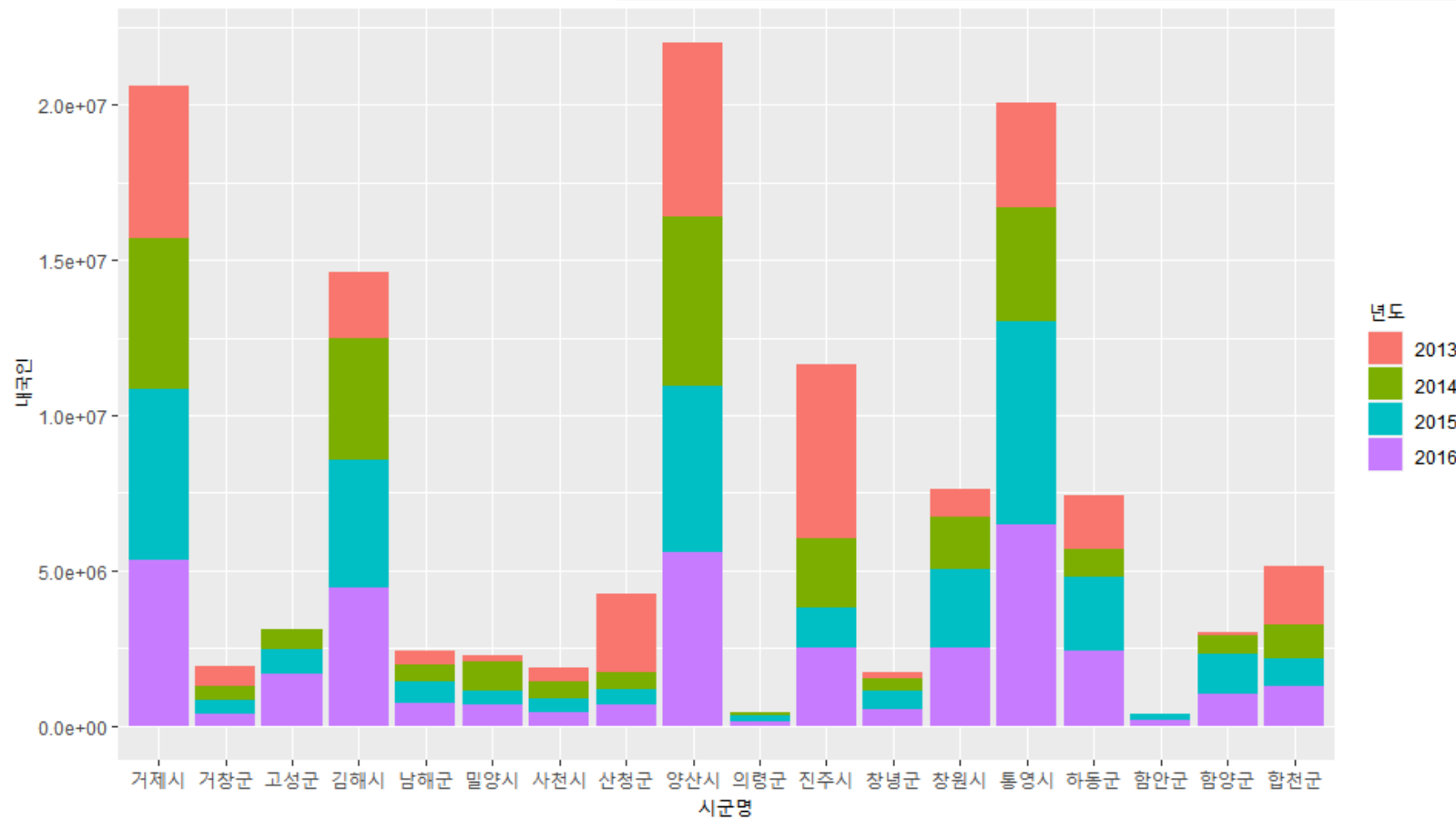
```
ggplot(data=df, aes(x=시군명, y=내국인, fill=년도)) +  
  geom_bar(stat = "identity", position='dodge')
```



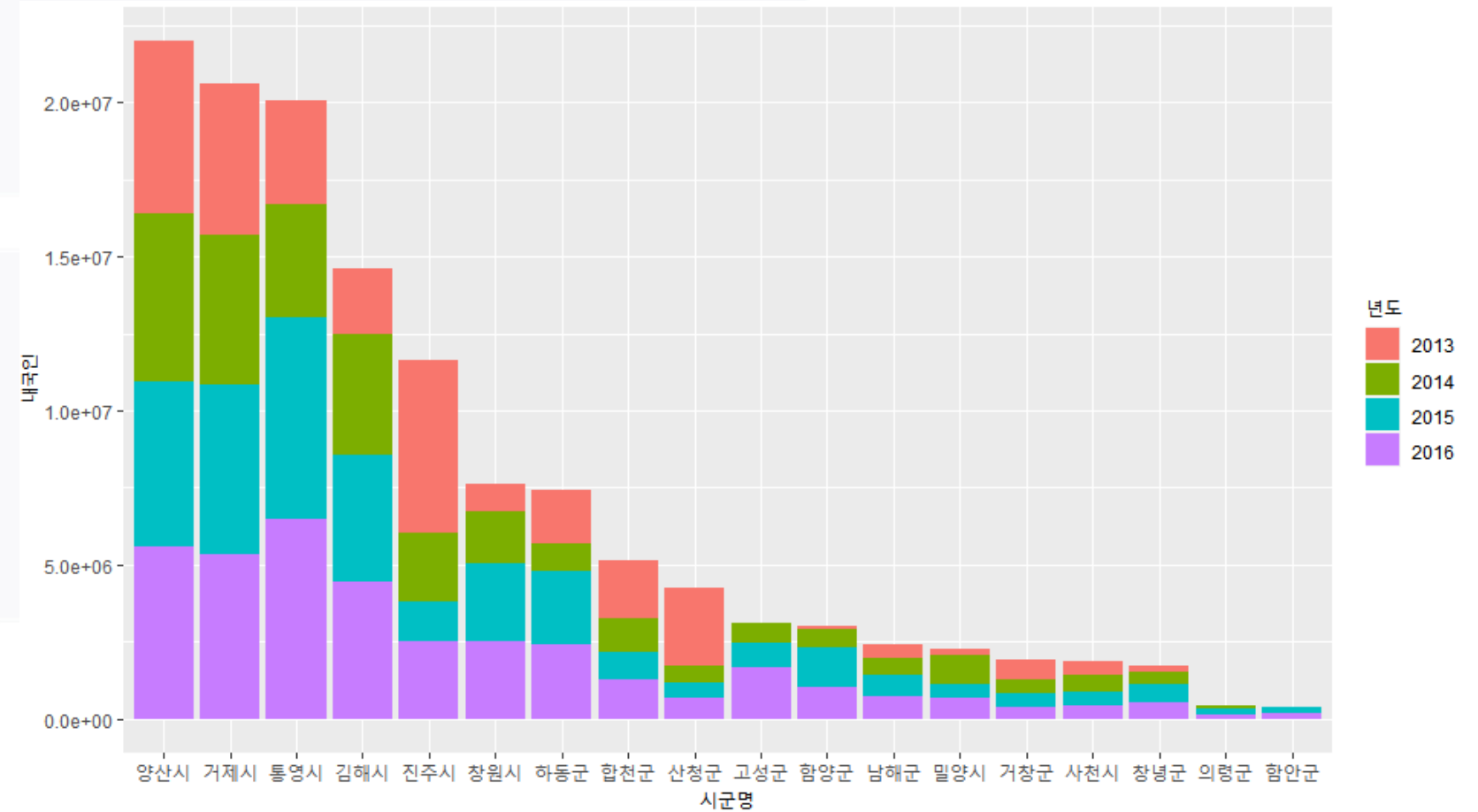
시각화

ggplot2 막대그래프

```
ggplot(data=df, aes(x=시군명, y=내국인, fill=년도)) +  
  geom_bar(stat = "identity", position='dodge')
```



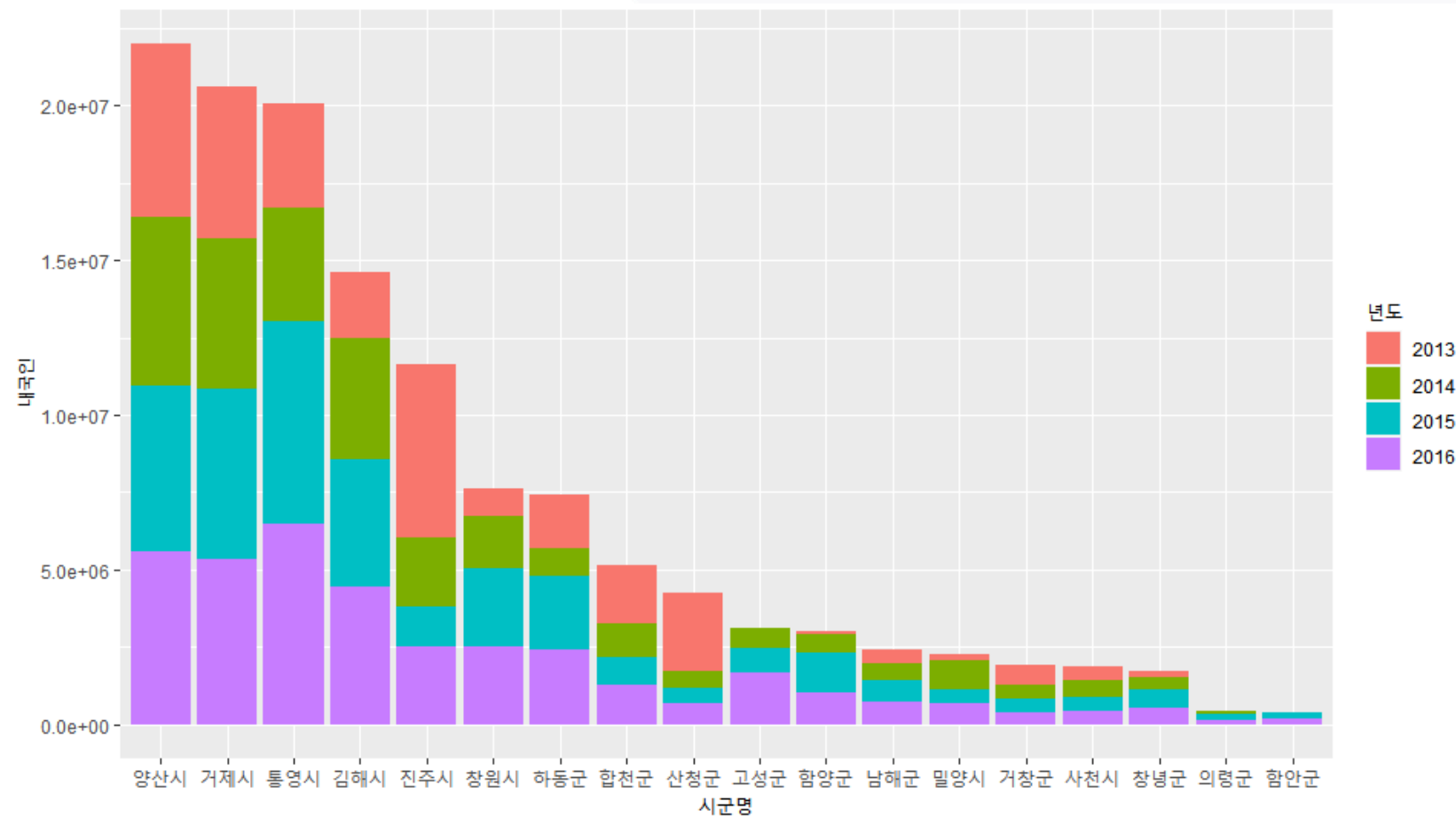
```
ggplot(data=df, aes(x=reorder(시군명, -내국인), y=내국인, fill=년도)) +  
  geom_bar(stat = "identity") +  
  labs(x = "시군명")
```



시각화

ggplot2 막대그래프

```
ggplot(data=df, aes(x=reorder(시군명, -내국인), y=내국인, fill=년도)) +  
  geom_bar(stat = "identity") +  
  labs(x = "시군명")
```



```
ggplot(data=df, aes(x=reorder(시군명, 내국인), y=내국인, fill=년도)) +  
  geom_bar(stat = "identity") +  
  labs(x = "시군명") +  
  coord_flip()
```

