

## [R] R ggplot 사용법 (데이터 시각화 도구) — 나무늘보의 개발 블로그

노트북: blog

만든 날짜: 2020-10-03 오후 5:51

URL: <https://continuous-development.tistory.com/42?category=793392>

---



R

# [R] R ggplot 사용법 (데이터 시각화 도구)

2020. 7. 28. 23:05 수정 삭제 공개

## #ggplot

고급화된 시각화를 위해 ggplot을 활용  
처음 ggplot으로 그래프 틀을 만든다.

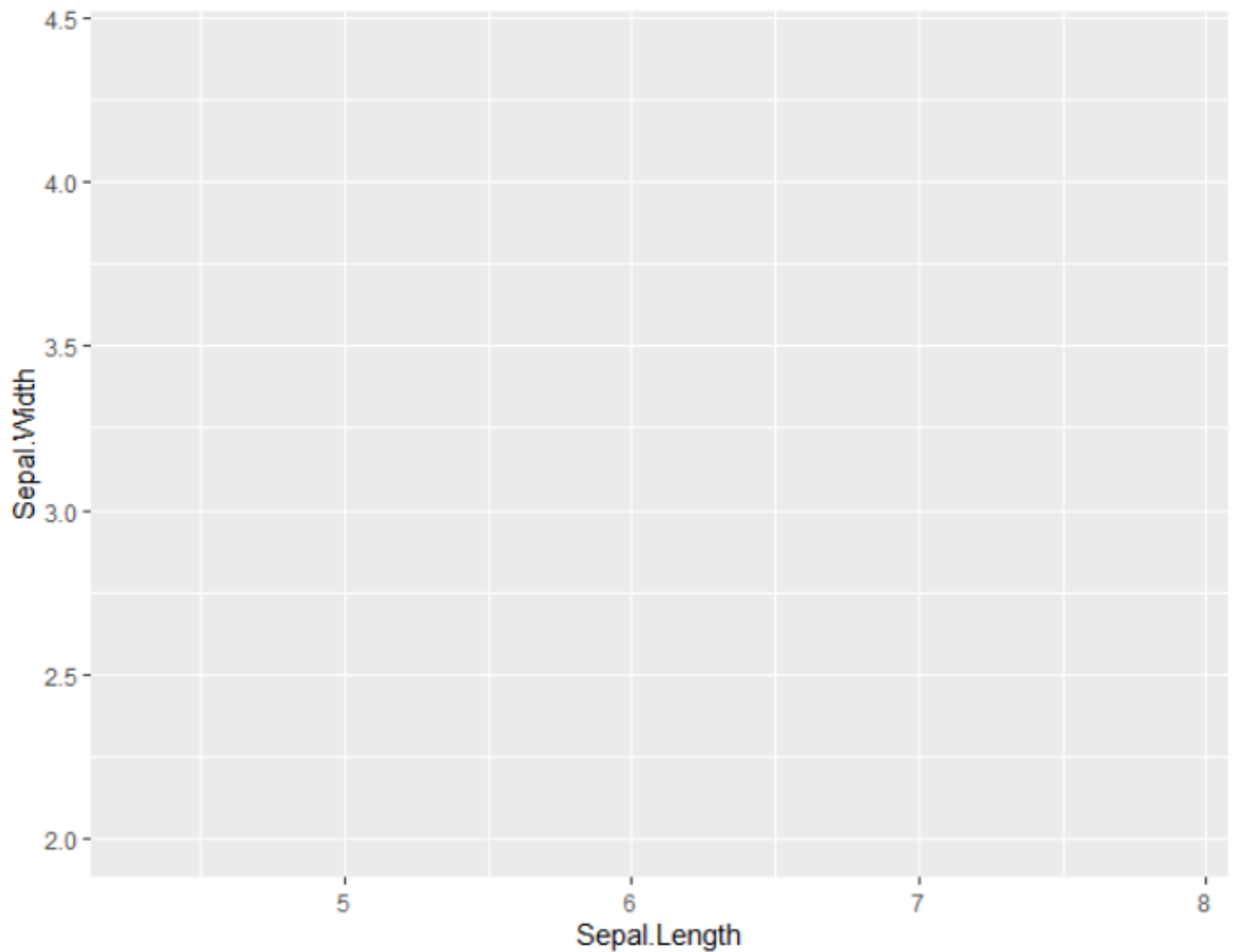
구문

```
ggplot(data = data이름,  
      aes(x = x축값,  
          y = y축값)  
    )
```

```

225 ## ggplot
226 # ggplot 종류
227 # -- ggplot() : 틀을 만드는 계열의 함수.
228 # -- gemo_    : 그래프를 만드는 계열의 함수
229 # -- gemo_    : 도형 계열을 만드는 함수
230 # -- coord_, labs : 계열을 만드는 함수
231
232 library(ggplot2)
233 library(dplyr)
234
235 # ggplot 필수 함수로 데이터와 축을 지정하는데 사용
236 iris
237
238 ggplot(data = iris,      #이렇게 판을 만든다.
239        aes(x=Sepal.Length, #축을 지정
240            y=Sepal.Width)
241        )
242

```

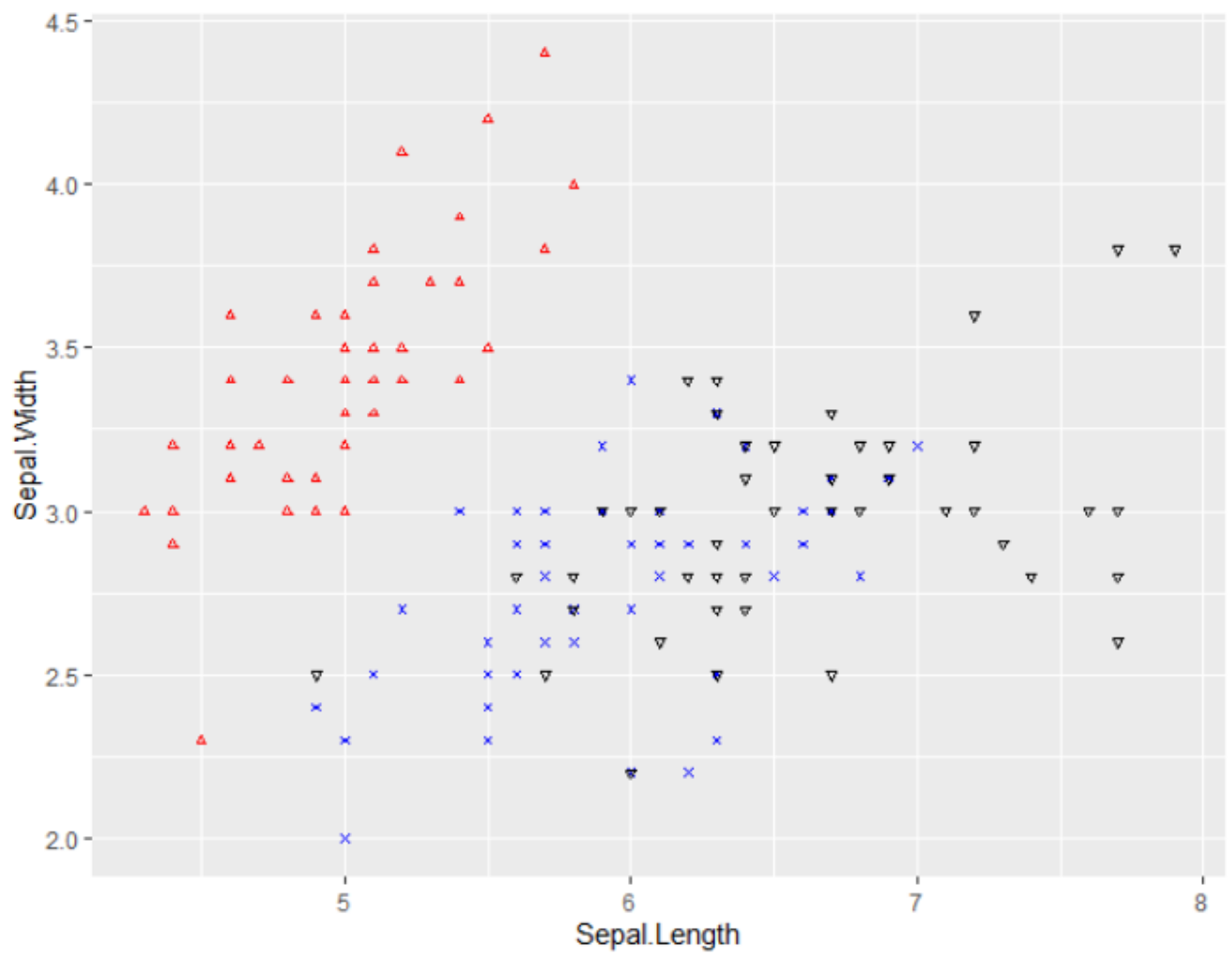


## #geom\_point - 그래프에 값에 따라서 점을 찍는다.

구문

```
ggplot(data = data이름,  
  aes(x = x축값,  
    y = y축값)  
  ) + # 0/ + 로 ggplot 과 geom_point를 연결한다.  
  geom_point( pch = value [data]  # 점의 모양을 정한다.  
    size = value [data] # 점의 크기를 정한다.  
    col = value [data] # c로 묶어서 벡터로 나타낸다. col은 컬러를 정한다.  
  )
```

```
270 ggplot(data = iris,      #이렇게 판을 만든다.  
271   aes(x=Sepal.Length, #축을 지정  
272     y=Sepal.Width))+  
273   geom_point(pch = c(2,4,6)[iris$Species], # 종에따라 나눈다는 말을 해야된다. 그다음 포인트로 해당 값들을 그래프에 나타낸다.  
274     size =c(.8,.9,1)[iris$Species],  
275     col =c('red','blue','black')[iris$Species]  
276   )  
277 )  
278  
279
```



**## 외부 옵션을 활용한 - coord 계열 , labs**

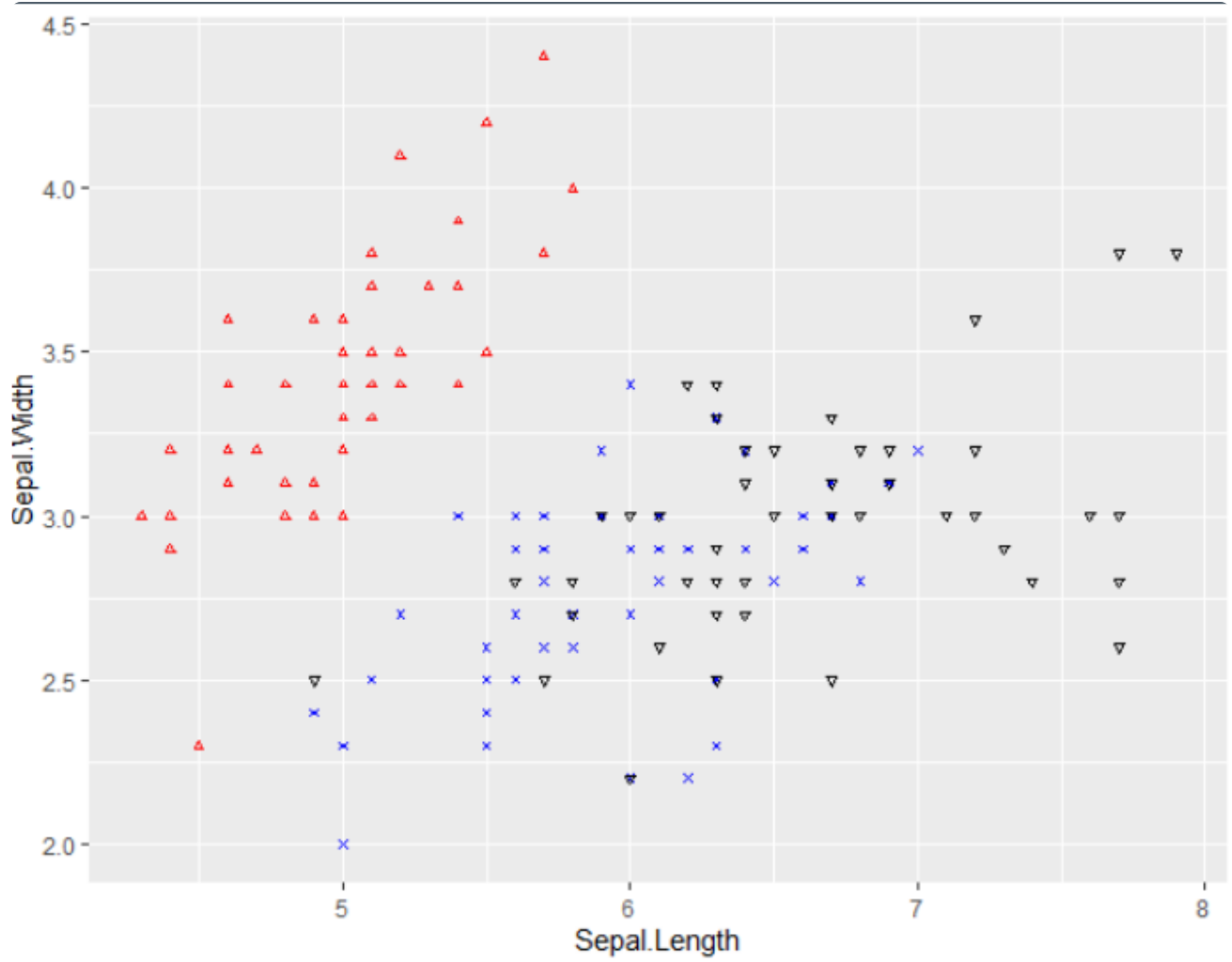
**# 축 변환 - coord\_flip()**

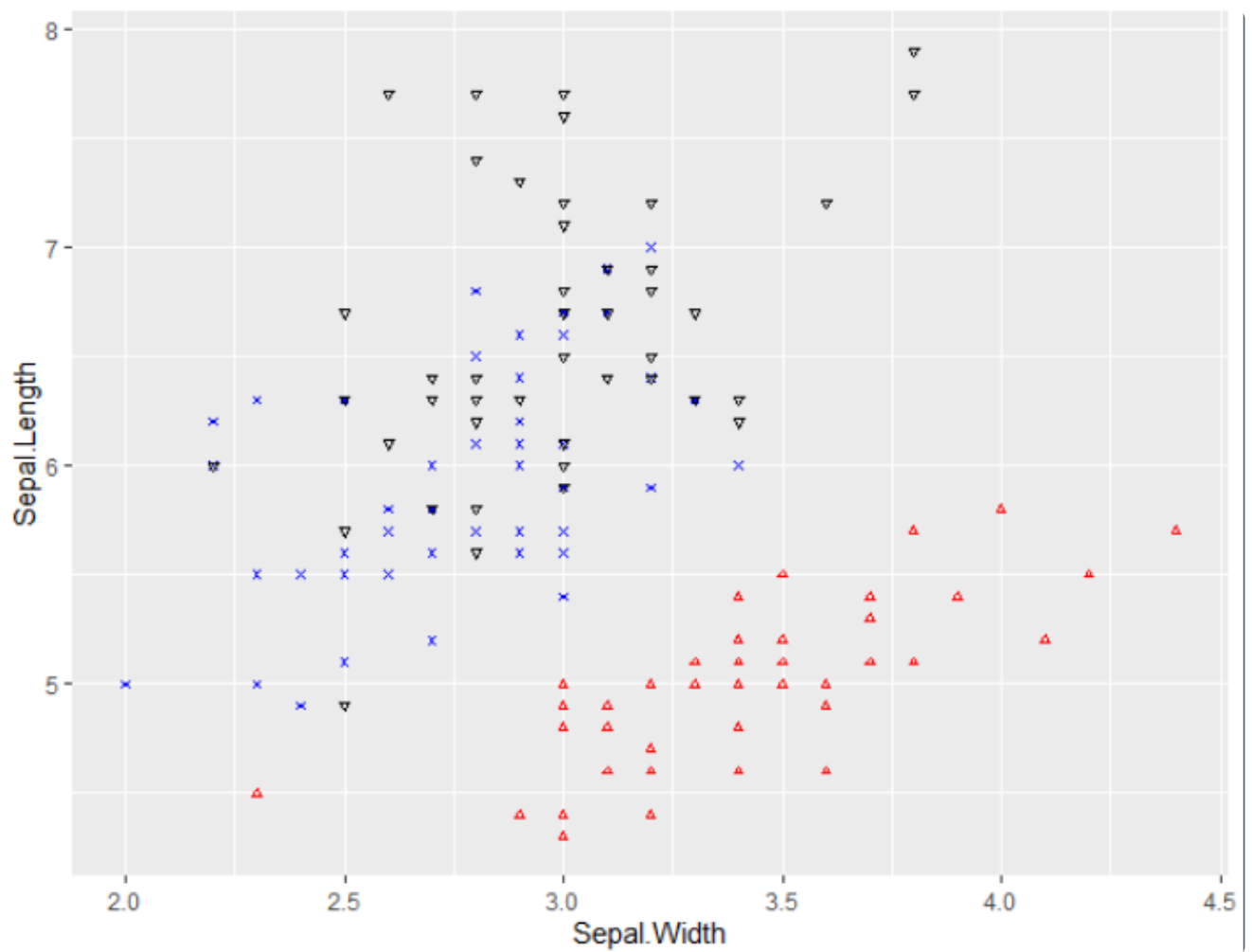
**# 축 범위 - coord\_cartesian()**

```

351 # 외부 옵션을 활용한 - coord 계열 , labs
352 # 축 변환 - coord_flip()
353 # 축 범위 - coord_cartesian()
354
355 irisG
356 irisG + coord_flip()
357

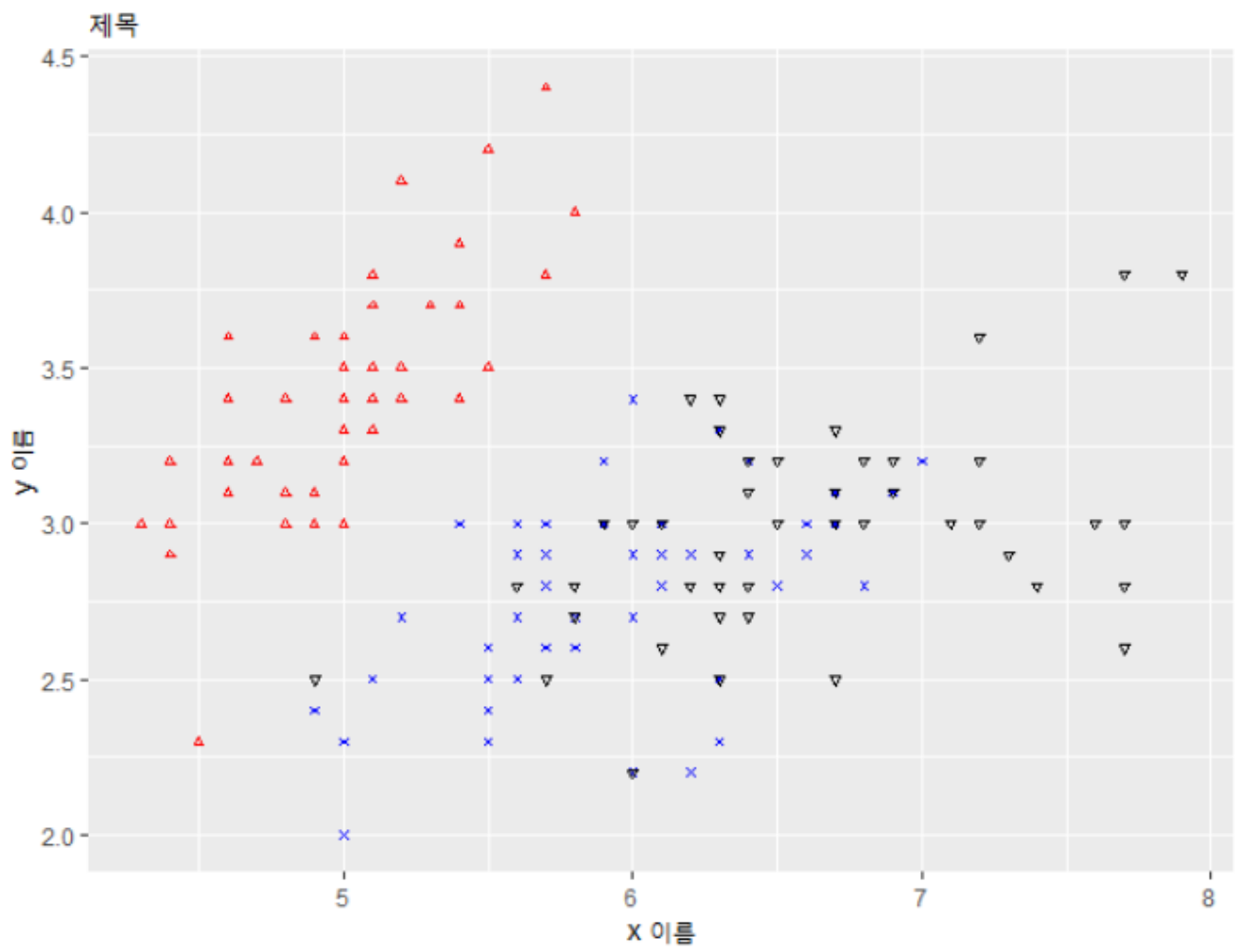
```





# labs() - 라벨링을 한다.

```
358  
359 # 라벨링 - labs()  
360 irisG + labs(title = "제목",  
361             x = "x 이름",  
362             y = "y 이름")  
363  
364
```



## 예제

```

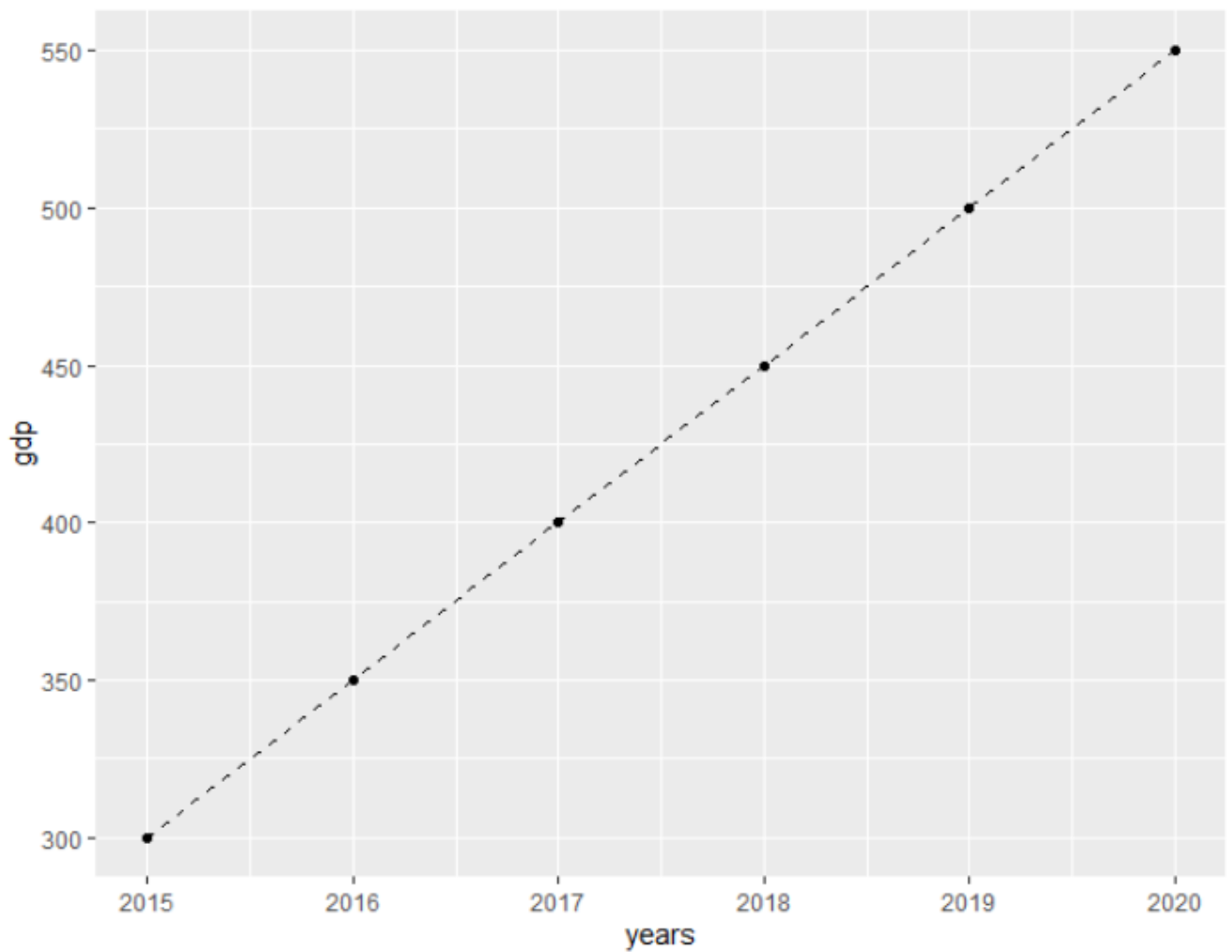
364 ## ggplot
365
366 sampleDF <- data.frame(
367   years = c(2015:2020),
368   gdp   = c(300,350,400,450,500,550)
369 )
370

```

```

371
372 #1. 틀생성|
373
374 ggplot(data = sampleDF,
375         aes(x = years, y =gdp)
376         ) +
377     geom_point()+
378     geom_line(linetype = "dashed") #실선이 점점으로 바뀐다.
379

```



## #geom\_line

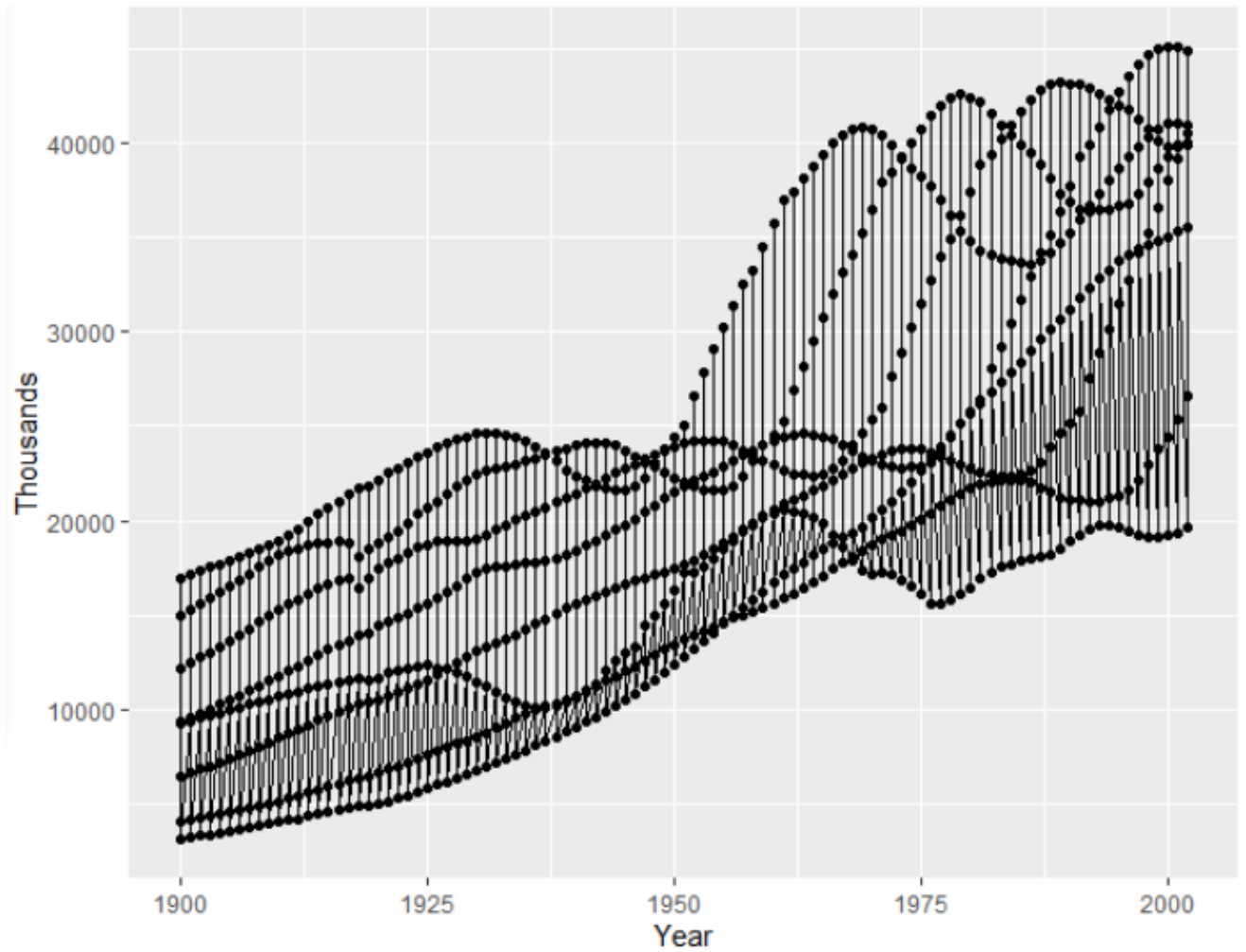
geom\_line()을 사용해 line을 그린다.



```

387 #year , thousands 가지고 기본적 ggplot 만들자
388
389 ggplot(data = uspopage,
390       aes(x=Year,y=Thousands),
391       )+
392   geom_point() +
393   geom_line()
394

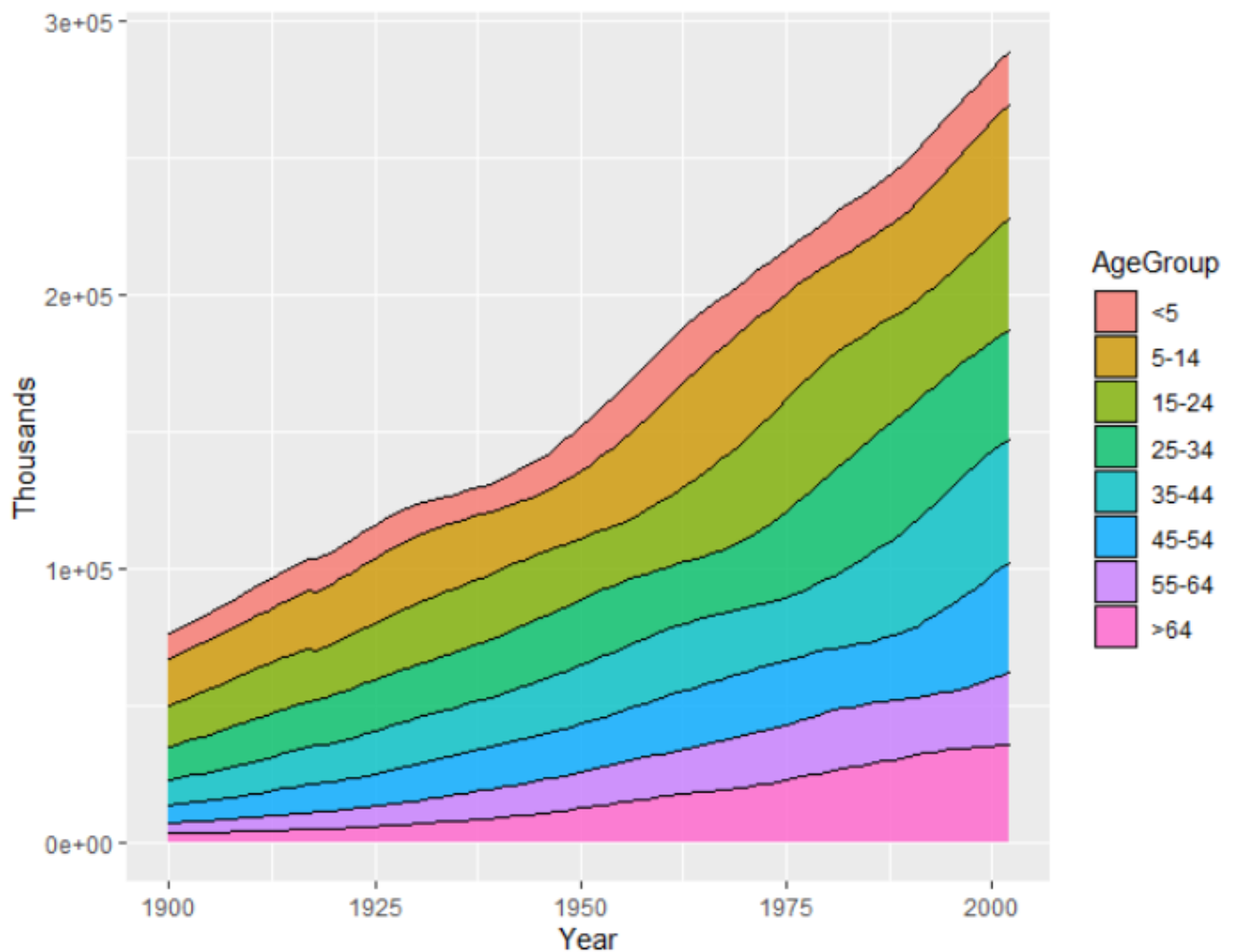
```



## #geom\_area

- 누적 영역 그래프

```
398  
399   ggplot(data = uspopage,  
400         aes(x=Year,y=Thousands, fill = AgeGroup))+  
401     geom_area(alpha = .8, col ='black')  
402  
403
```



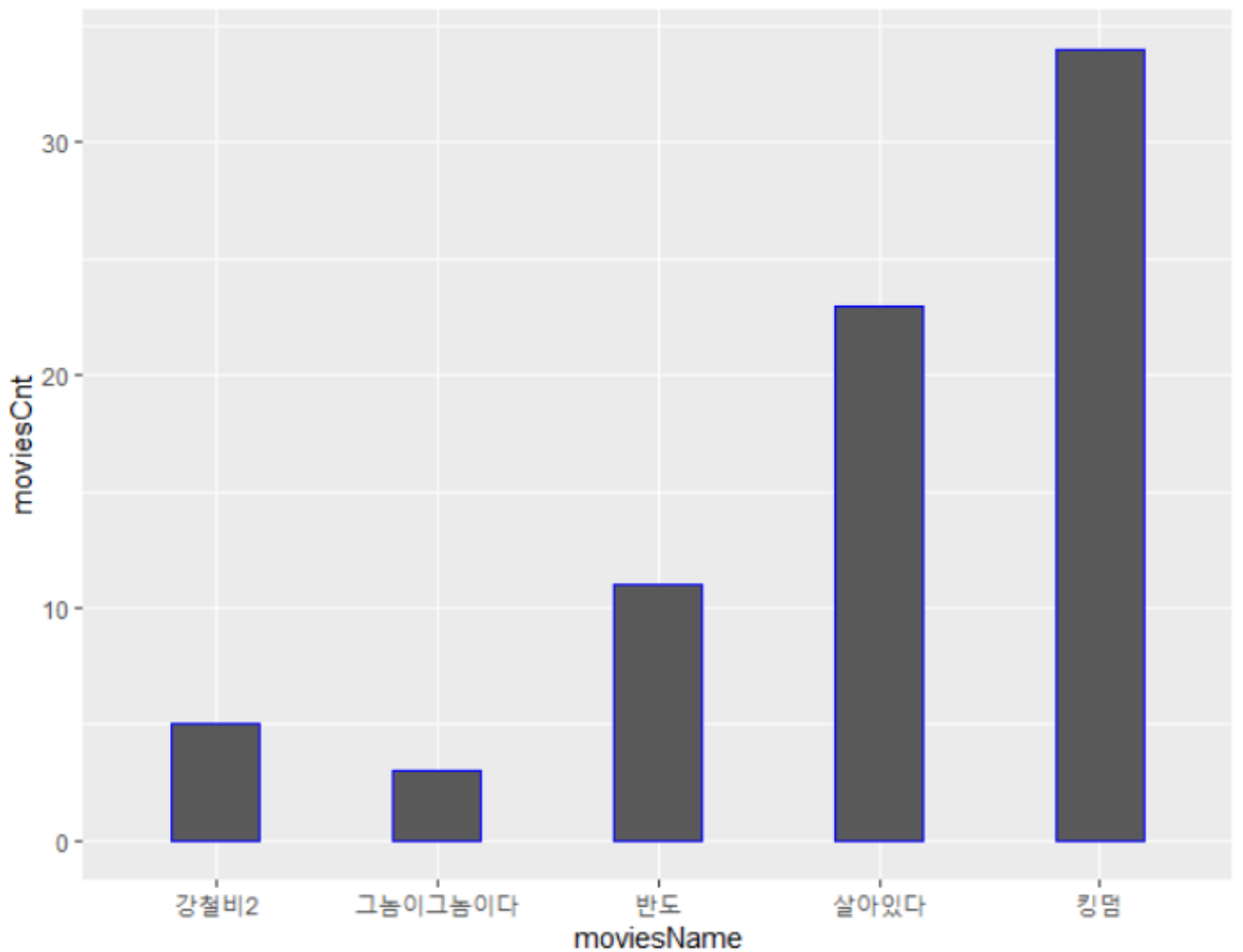
## geom\_col()

- 막대그래프를 그리는 데 있어서 축이 두 개일 때 사용하는 함수 (aes에 x축과 y축 이렇게 두 개가 있을 때 사용)

```

410 # geom_bar() - 막대 그래프 생성
411 korMovies <- c('강철비2', '반도', '그놈이그놈이다', '킹덤', '살아있다')
412 cntMovies <- c(5, 11, 3, 34, 23)
413
414 moviesDF <- data.frame(moviesName = korMovies,
415                       moviesCnt = cntMovies)
416
417 ggplot(data = moviesDF,
418       aes(x=moviesName, y=moviesCnt)) + # bar 는 축이 하나 일때는 geom_bar 를 써도 되지만 축이 두개 일때는 geom_col 을 써야 된다.
419   geom_col(col='blue',
420           width=.4 # 그래프의 두께)
421
422

```

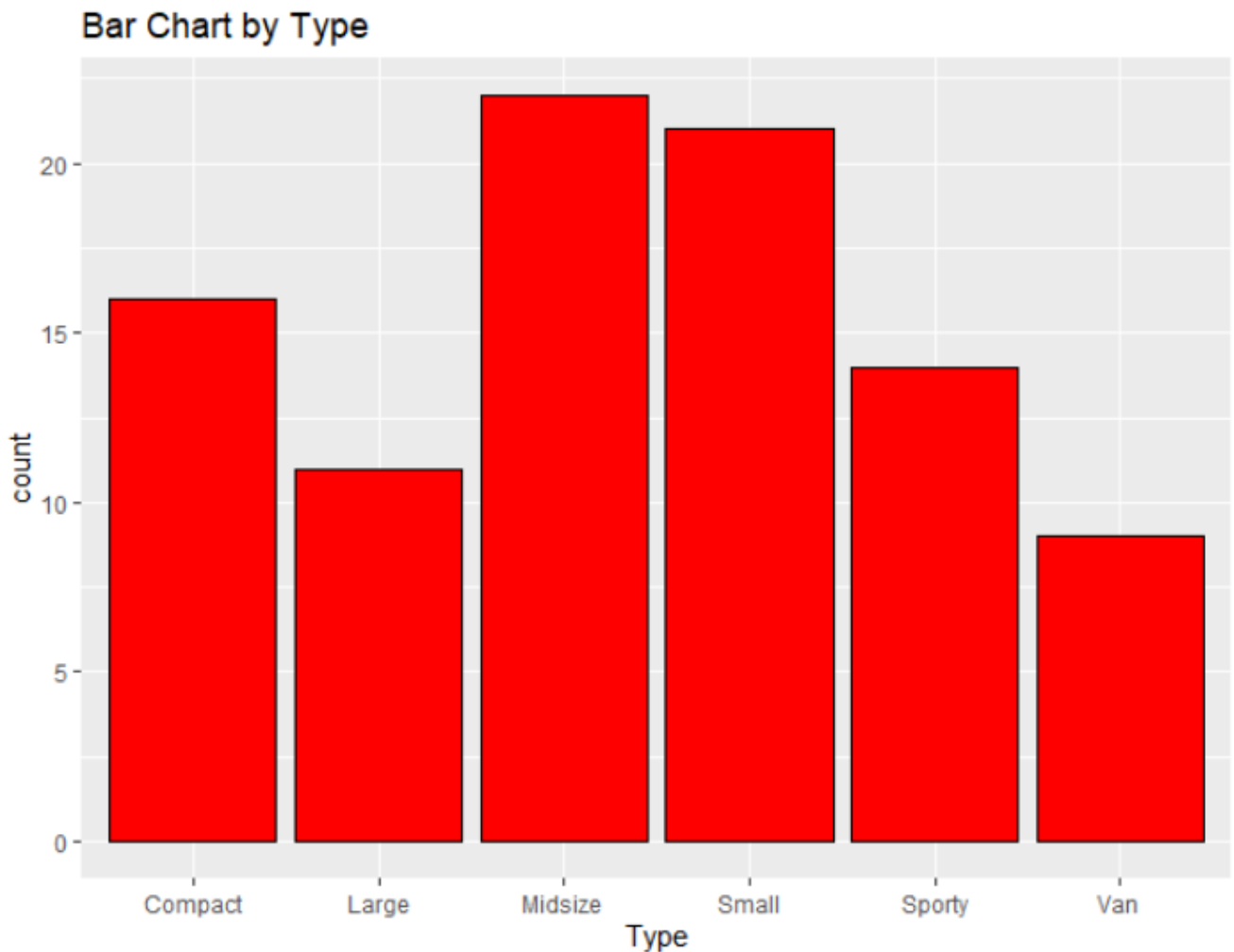


## #geom\_bar

막대그래프를 그리는데 있어서 축이 1개일 때 사용 (명령어를 통해 2개 일 때도 사용할 수 있다,)

\*fill은 그래프를 채우는데 색상을 나타낸다.

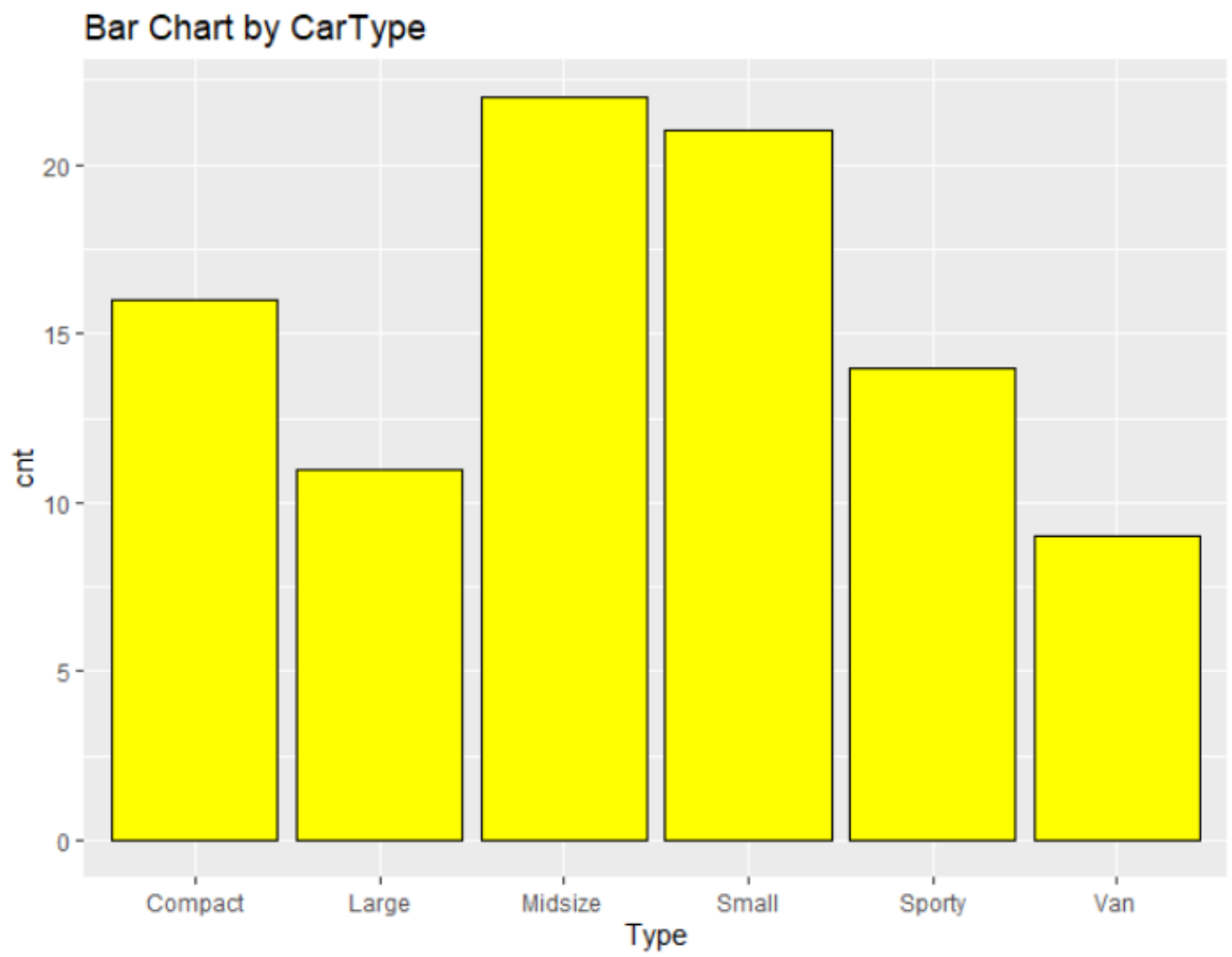
```
428 # geom_bar() - 축이 하나인 경우의 막대그래프를 그리기 위해 사용
429 ggplot(data = Cars93,
430       aes(x=Type))+ #bar 는 축이 하나 일때는 geom_bar 를 써도 되지만 축이 두개일때는 geom_col을 써야된다.
431       geom_bar(col='black', #그래프 바깥의 선 색깔
432             fill = 'red',# 그래프의 색상
433             )+
434       ggtitle("Bar Chart by Type") # 그래프의 제목을 넣는다.
435
```



# geom\_bar - stat = 'identity'

그래프를 그리는데 있어서 축이 두 개 일 때 사용한다.

```
435 ggplot(data = type_cnt,
436       aes(x = Type, y = cnt))+ #bar 는 축이 하나 일때는 geom_bar 를 써도 되지만 축이 두개일때는 geom_col을 써야된다.
437       geom_bar(stat = 'identity', fill='yellow', col='black')+ # stat = 'identity'을 써야만 축이 두개일 때 사용가능하다.
438       ggtitle("Bar Chart by CarType")
439
```

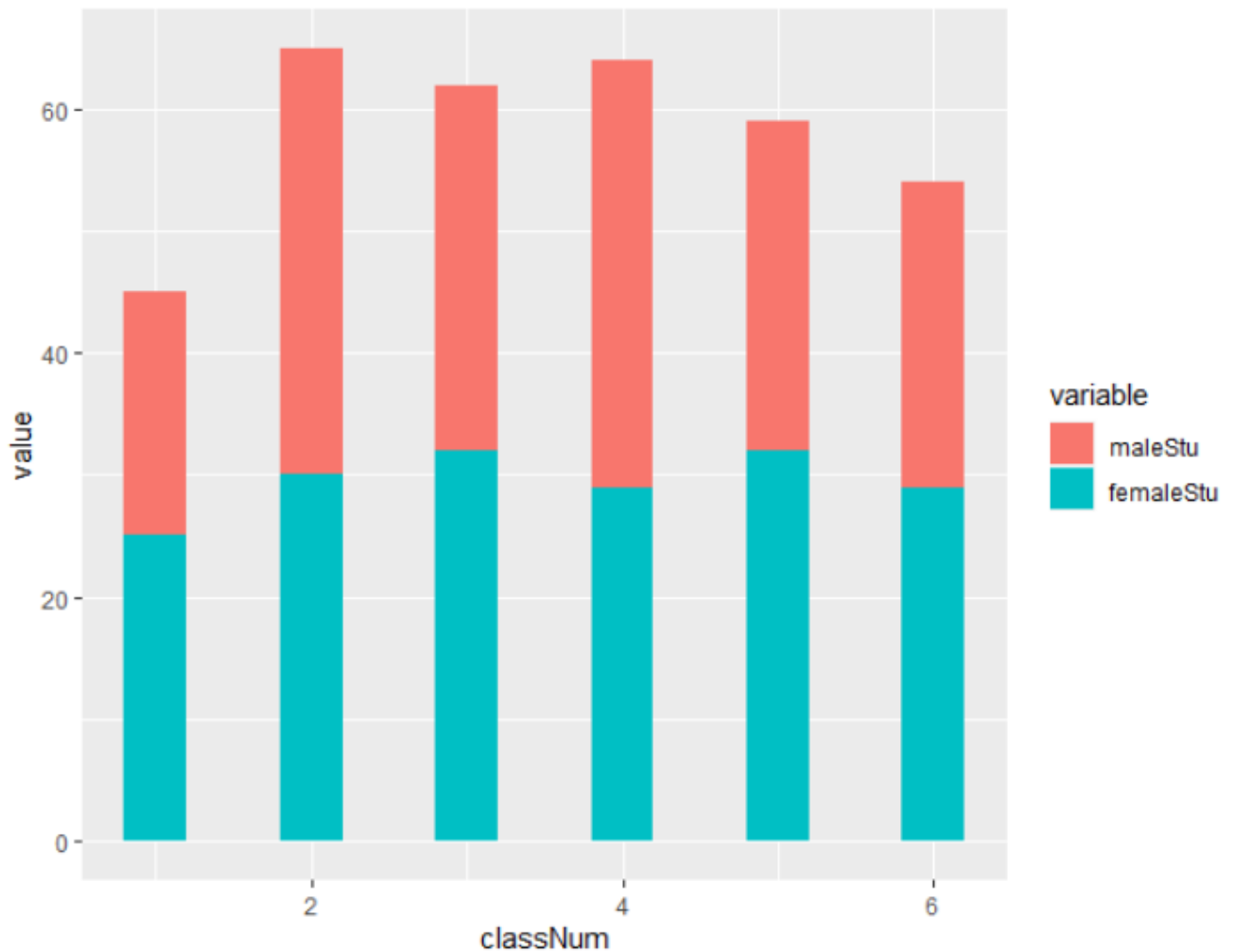


#melt를 이용한 데이터 가공후 plot 사용

```

451 #melt 를 통해 3개의 열을 가진 데이터 프레임을 2개의 열로 바꾸는 법 / 비교를 위해 두개의 막대그래프 / 비율을 나누거나
452 maleStu <- c(20,35,30,35,27,25)
453 femaleStu <- c(25,30,32,29,32,29)
454 classNum <- c(1,2,3,4,5,6)
455
456 stuDF <- data.frame(maleStu, femaleStu, classNum)
457
458
459 library(reshape2)
460
461 stuMelt <- melt(stuDF ,id="classNum")
462 stuMelt
463
464 ggplot(data = stuMelt,
465       aes(x=classNum, y=value, fill=variable))+
466   geom_bar(stat = 'identity' , width= .4)
467

```



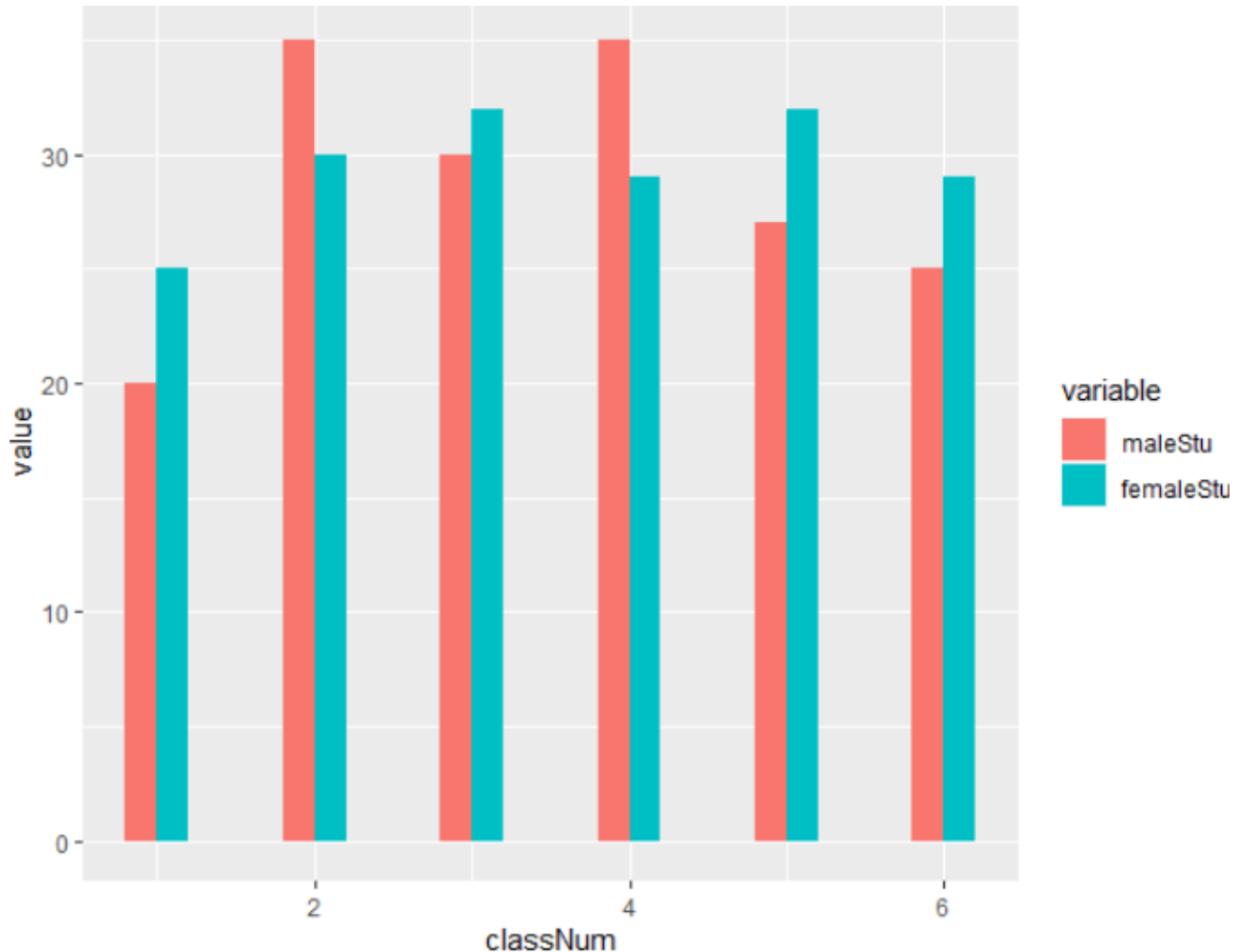
**# multibar - position = position\_dodge()**

1개 이상의 그래프를 그리는데 사용한다.

```

469
470 #multibar
471 ggplot(data = stuMelt,
472       aes(x=classNum, y=value, fill=variable))+
473   geom_bar(stat = 'identity' ,
474         width= .4,
475         position = position_dodge()) ## 두가지의 그래프가 나오게끔 한다.
476
477
478

```



```

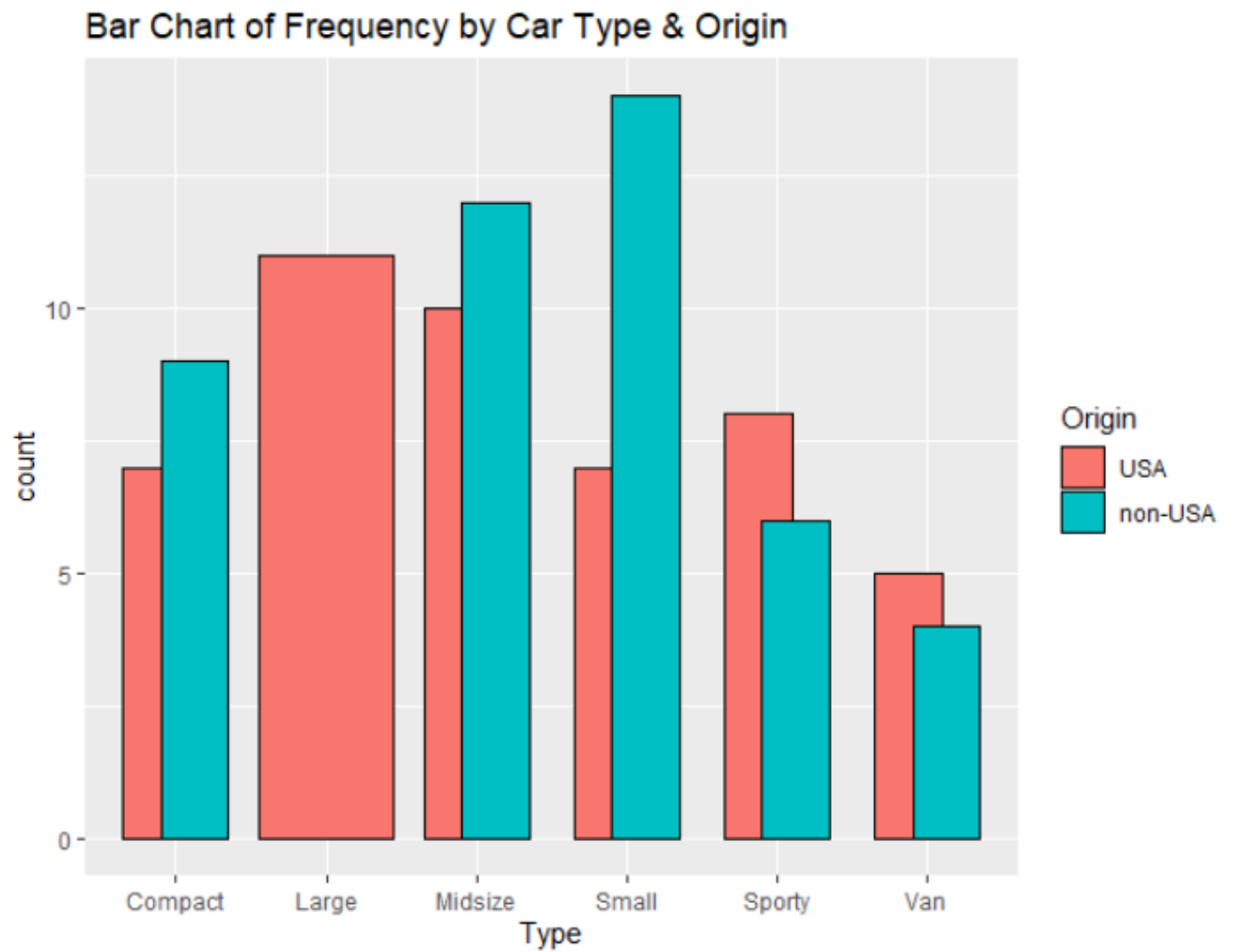
478
479 # Cars 93 데이터를 이용하여 차종별(Type) 제조국(Origin)별 자동차 수를 가지고
480 #막대 그래프를 려본다면?

```

```

489
490 ggplot(Cars93, aes(x=Type, fill =Origin))+
491   geom_bar(position = position_dodge(width=.5),
492         col="black")+
493   ggtitle("Bar Chart of Frequency by Car Type & Origin")
494
495

```



#### 'R' 카테고리의 다른 글

[R] ggplot2 패키지 설치 에러시 해결 방법

[R] R 을 활용한 데이터 탐색(Exploratory Data Analysis)

**[R] R ggplot 사용법 (데이터 시각화 도구)**

[R] R 에서 사용되는 기본적인 시각화 그래프-2

[R] R 에서 사용되는 기본적인 시각화 그래프

[R] R 데이터 가공을 위한 기본적인 함수

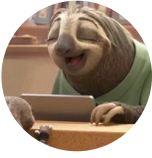
ggplot

R ggplot

R ggplot 사용법

R ggplot 함수





꾸까꾸

혼자 끄적끄적하는 블로그 입니다.

