

Bar Chart 개요

범주형 변수에 대한 요약 통계량(평균, 합을 주로 사용)의 크기를 바의 길이로 표현 - 요약한 통계량 그래프 그리는 것은 [\[점도표 강의노트 참고\]](#)

미국 MLB 타자 연봉 데이터 (http://wolfpack.hnu.ac.kr/Stat_Notes/example_data/baseball.csv)

타자 연봉 및 능력 데이터

```
baseball<-read.csv('http://wolfpack.hnu.ac.kr/Stat_Notes/example_data/
baseball.csv')
names(baseball);dim(baseball)
baseball.subset<-baseball[c(1:30),]
```

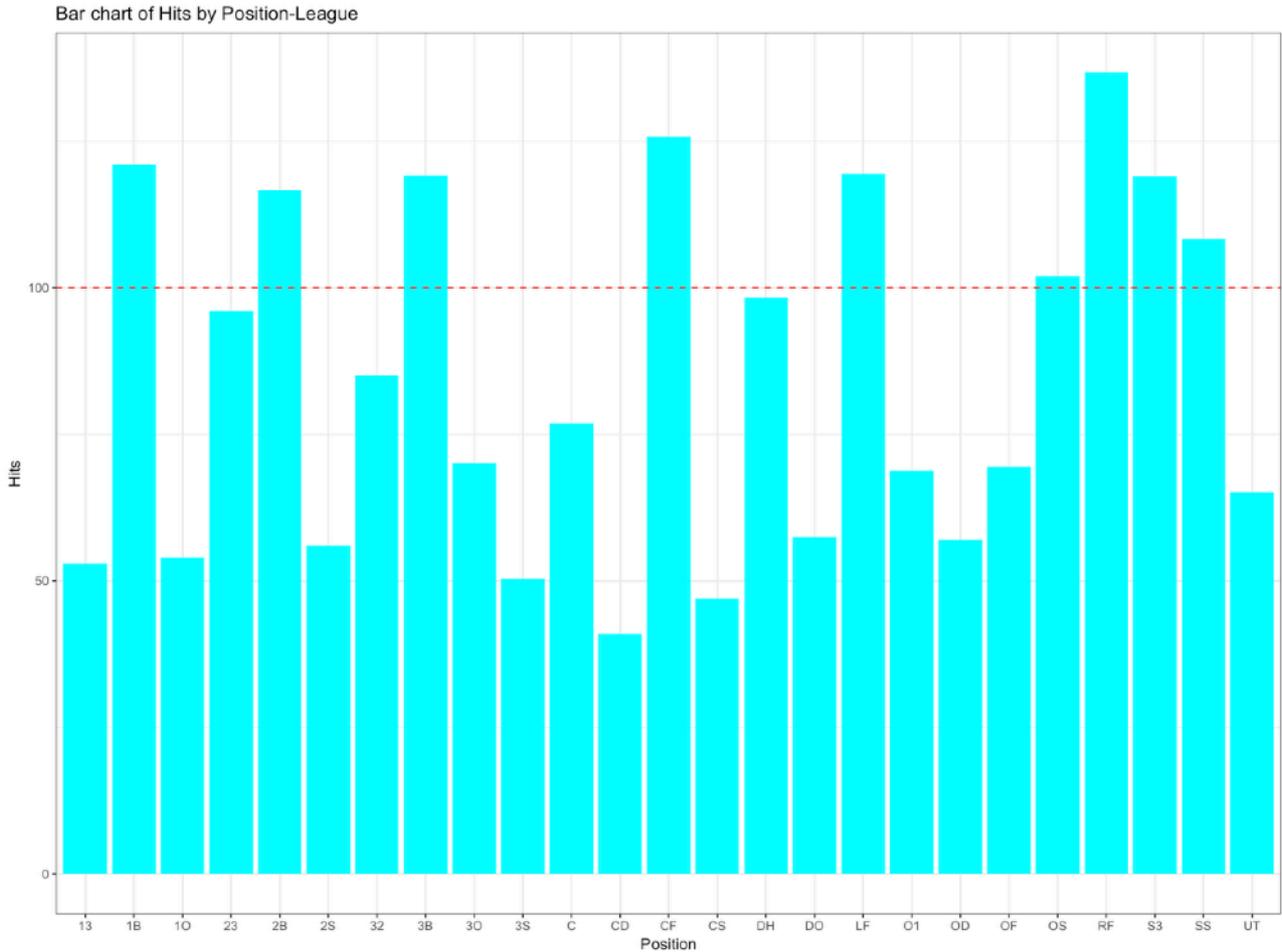
```
> names(baseball);dim(baseball)
 [1] "Player_Name"      "Team"              "TimesatBat"        "Hits"
 [5] "HomeRuns"         "Runs"              "RBIs"              "Walks"
 [9] "YearsinMLB"       "CareerTimesatBat"  "CareerHits"        "CareerHome"
[13] "CareerRuns"       "CareerRBIs"        "CareerWalks"       "League"
[17] "Division"         "Position"          "PutOuts"           "Assists"
[21] "Errors"           "Salary"            "League2"
[1] 322 23
```

포지션별 안타(Hits)수

테이블 요약

```
head(aggregate(Hits~Position,baseball,FUN=mean,na.rm=T),5)
library(ggplot2)
ggplot(baseball) +
  geom_bar(aes(Position,Hits),stat='summary',fun.y='mean',fill='cyan') +
  xlab('Position') + ylab('Hits') +
  geom_hline(yintercept=100,linetype='dashed', color = "red") +
  ggtitle('Bar chart of Hits by Position-League') +
  theme_bw()
```

```
> head(aggregate(Hits~Position,baseball,FUN=mean,na.rm=T),5)
  Position      Hits
1         13  53.0000
2         1B 120.9355
3         10  54.0000
4         23  96.0000
5         2B 116.5806
```



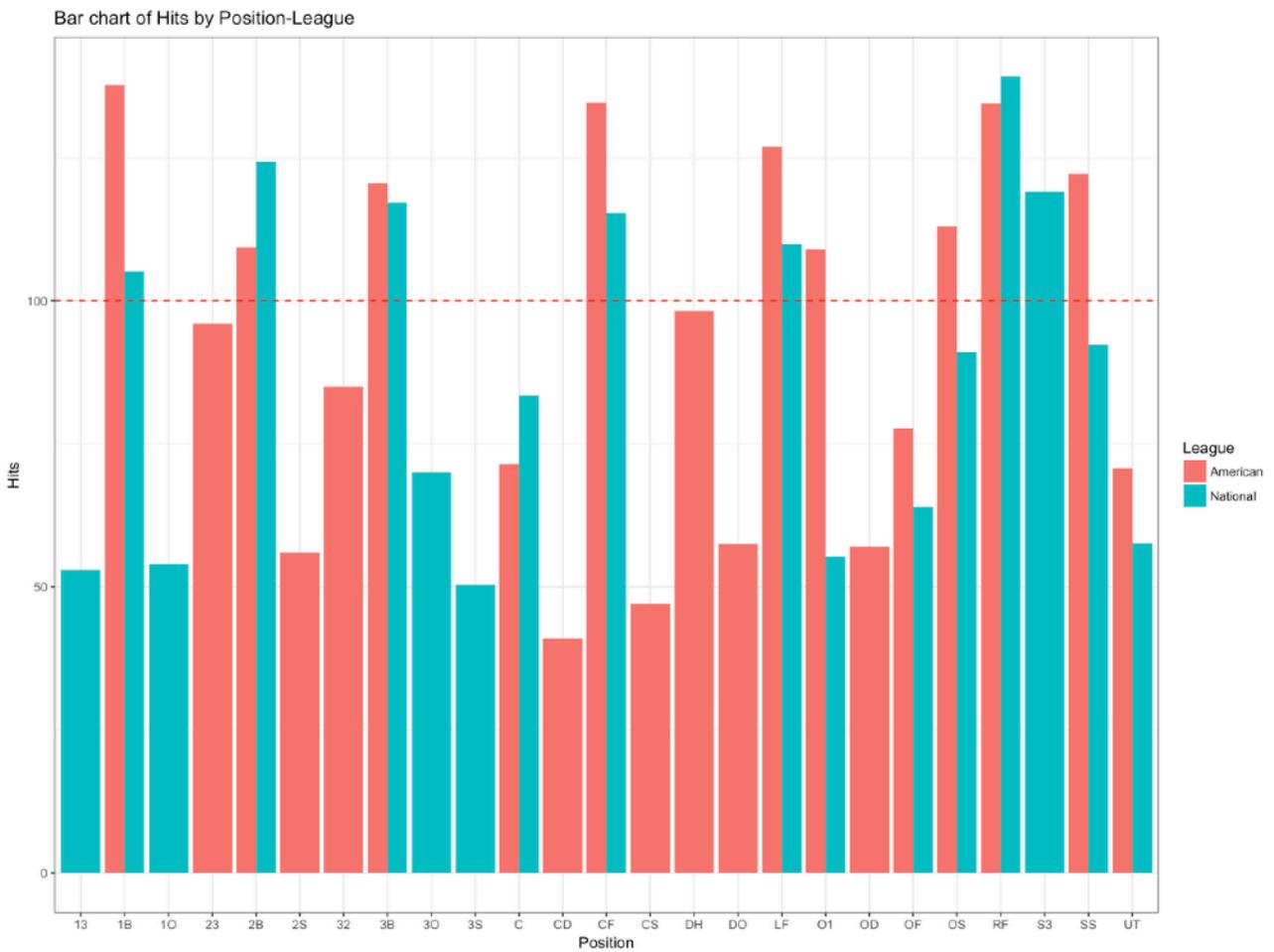
리그별-포지션별 안타수 평균

```
head(aggregate(Hits~Position+League,baseball,FUN=mean,na.rm=T),3)
tail(aggregate(Hits~Position+League,baseball,FUN=mean,na.rm=T),3)
```

```
library(ggplot2)
ggplot(baseball) +
  geom_bar(aes(Position,Hits,fill=League),stat='summary',fun.y='mean',position='dodge') +
  xlab('Position') + ylab('Hits') +
  geom_hline(yintercept=100,linetype='dashed', color = "red") +
  ggtitle('Bar chart of Hits by Position-League') +
  theme_bw()
```

position='dodge' 는 unstack(막대를 쌓지 않음) - 이 명령어를 사용하지 않으면 바가 쌓이게 된다.
표현 목적에 따라 다르지만 쌓는 것 보다는 옆에 놓고 비교하는 것이 적합한 표현

```
> head(aggregate(Hits~Position+League,baseball,FUN=mean,na.rm=T),3)
  Position League      Hits
1       1B American 137.7333
2       23 American  96.0000
3       2B American 109.3125
> tail(aggregate(Hits~Position+League,baseball,FUN=mean,na.rm=T),3)
  Position League      Hits
35      S3 National 119.00000
36      S5 National  92.35714
37      UT National  57.66667
```



[position='dodge'] 제외 시

