

Presidential Election Poll 1987

coop711

2015년 5월 5일

Data Read-in

table 구조를 가지므로 우선 matrix 로 읽어들인다.

```
poll.87<-matrix(c(289, 84, 26, 361, 154, 139, 53, 292, 126, 145, 57, 287, 61, 29, 11, 80), nrow=4, ncol=4)
poll.87
```

```
##           [,1] [,2] [,3] [,4]
## [1,]    289   154   126    61
## [2,]     84   139   145    29
## [3,]     26    53    57    11
## [4,]    361   292   287    80
```

```
class(poll.87)
```

```
## [1] "matrix"
```

```
str(poll.87)
```

```
## num [1:4, 1:4] 289 84 26 361 154 139 53 292 126 145 ...
```

table 구조의 변수명을 정한다.

```
dimnames(poll.87)<-list(c("Buddism","Protestant","Catholic","None"),c("Roh","YS",
"DJ","JP"))
poll.87
```

```
##           Roh  YS  DJ  JP
## Buddhism    289 154 126 61
## Protestant   84 139 145 29
## Catholic     26  53  57 11
## None        361 292 287 80
```

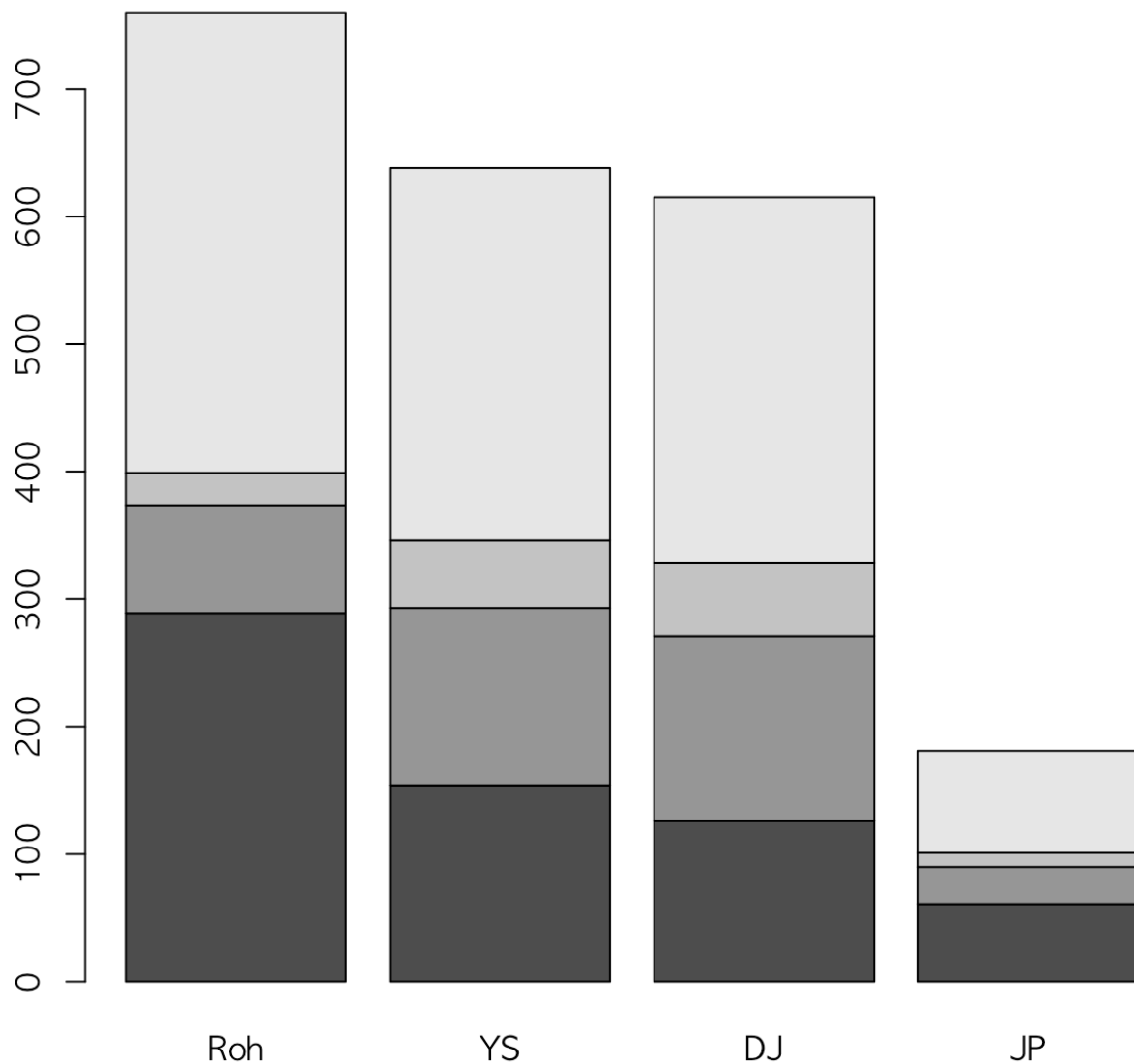
table 구조로 강제 변환.

```
poll.87.tbl<-as.table(poll.87)
str(poll.87.tbl)
```

```
## table [1:4, 1:4] 289 84 26 361 154 139 53 292 126 145 ...
## - attr(*, "dimnames")=List of 2
## ..$ : chr [1:4] "Buddism" "Protestant" "Catholic" "None"
## ..$ : chr [1:4] "Roh" "YS" "DJ" "JP"
```

table (실상은 matrix) 구조를 가지므로 barplot() 으로 막대들을 쌓아놓기 적합.

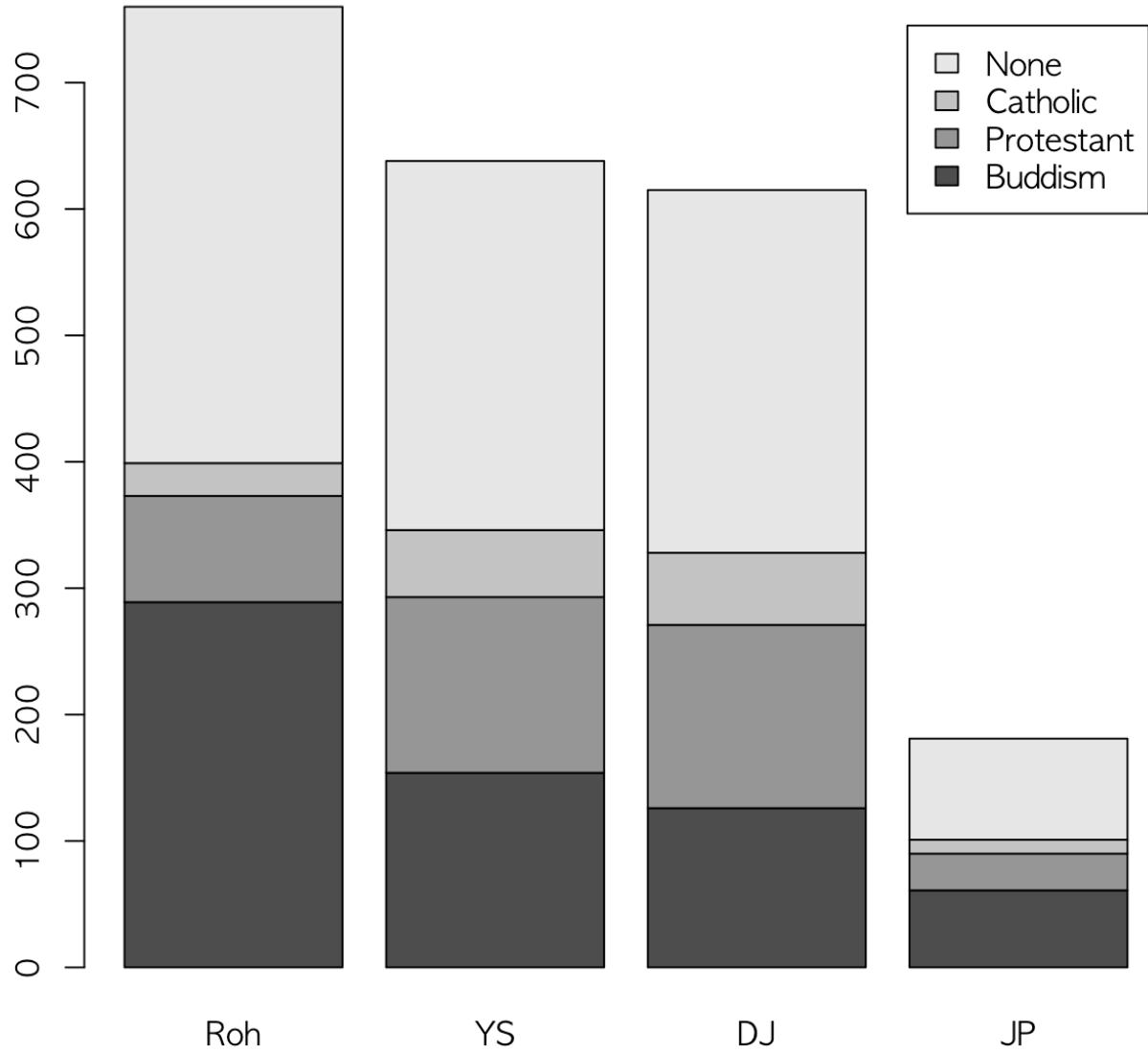
```
barplot(poll.87.tbl)
```



범례가 필요하다면 rownames() 가 설정되어 있으므로 legend.text=TRUE 로 충분함. 적절한 주제목 설정.

```
barplot(poll.87.tbl, legend.text=TRUE)
title(main="Poll 1987")
```

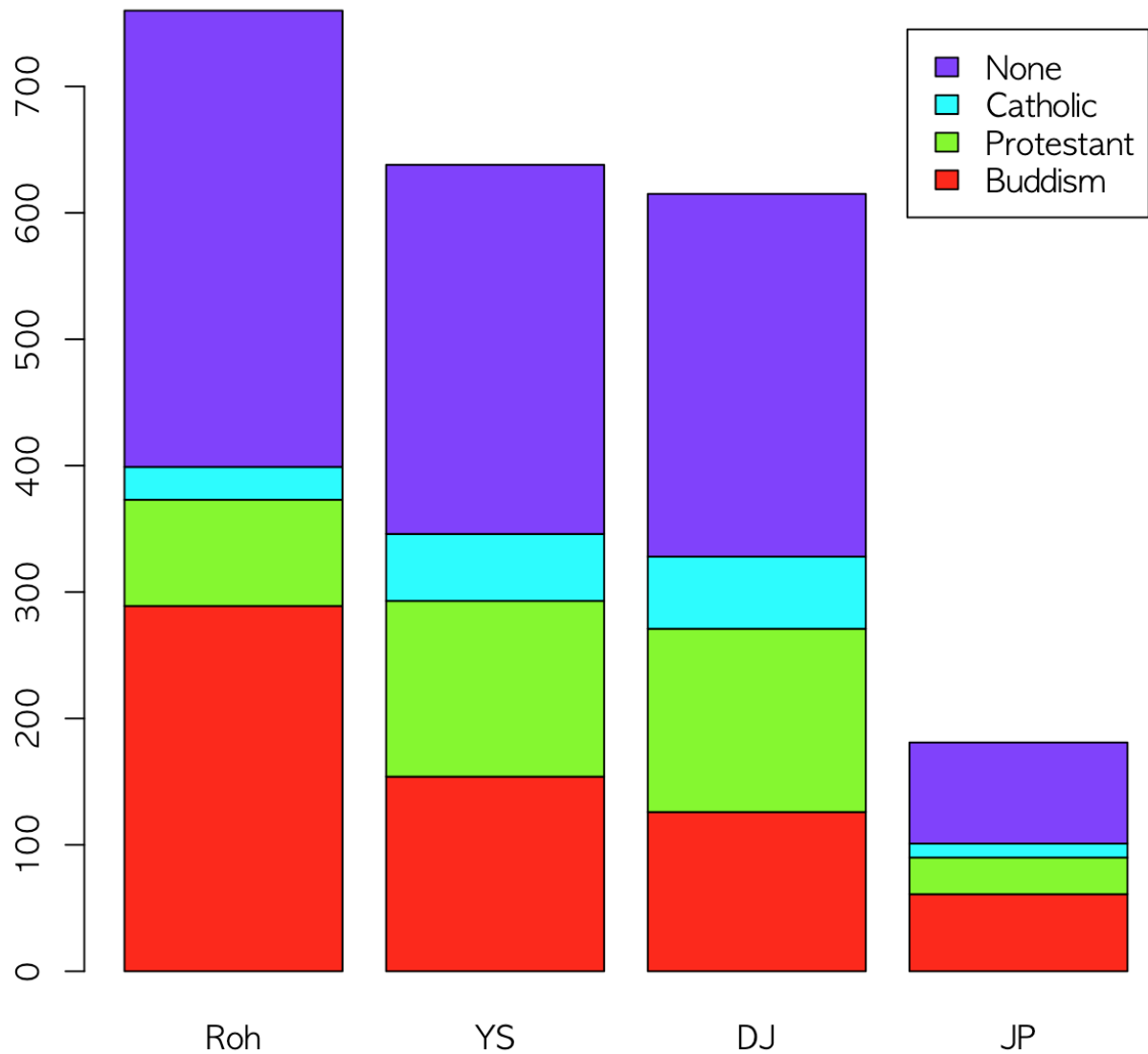
Poll 1987



종교를 색깔로 구분하면,

```
barplot(poll.87.tbl, legend.text=TRUE, col=rainbow(4))  
title(main="Poll 1987")
```

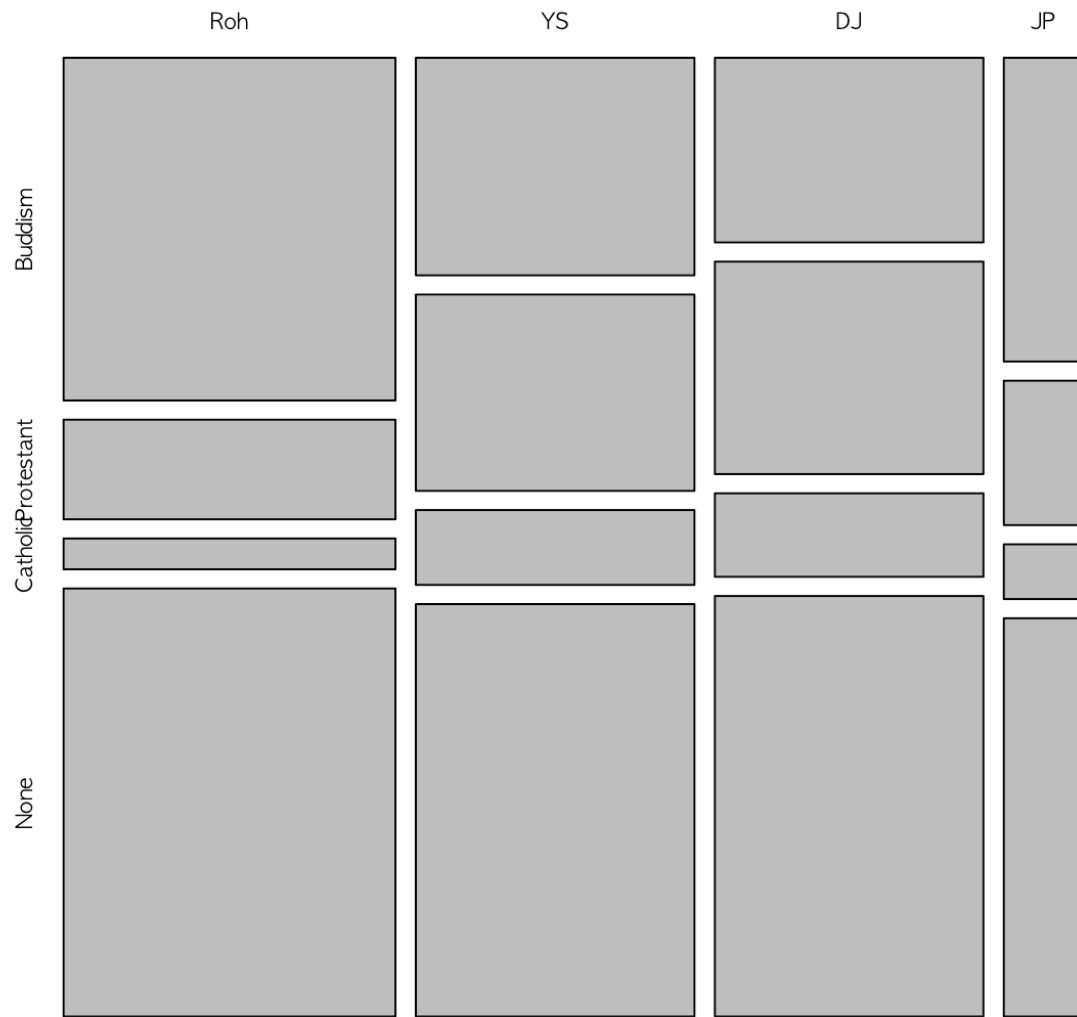
Poll 1987



이를 `mosaicplot()` 으로 비교해 보려면 `t()` 작업을 취해야 제대로 볼 수 있음.

```
mosaicplot(t(poll.87.tbl), main="Poll 1987")
```

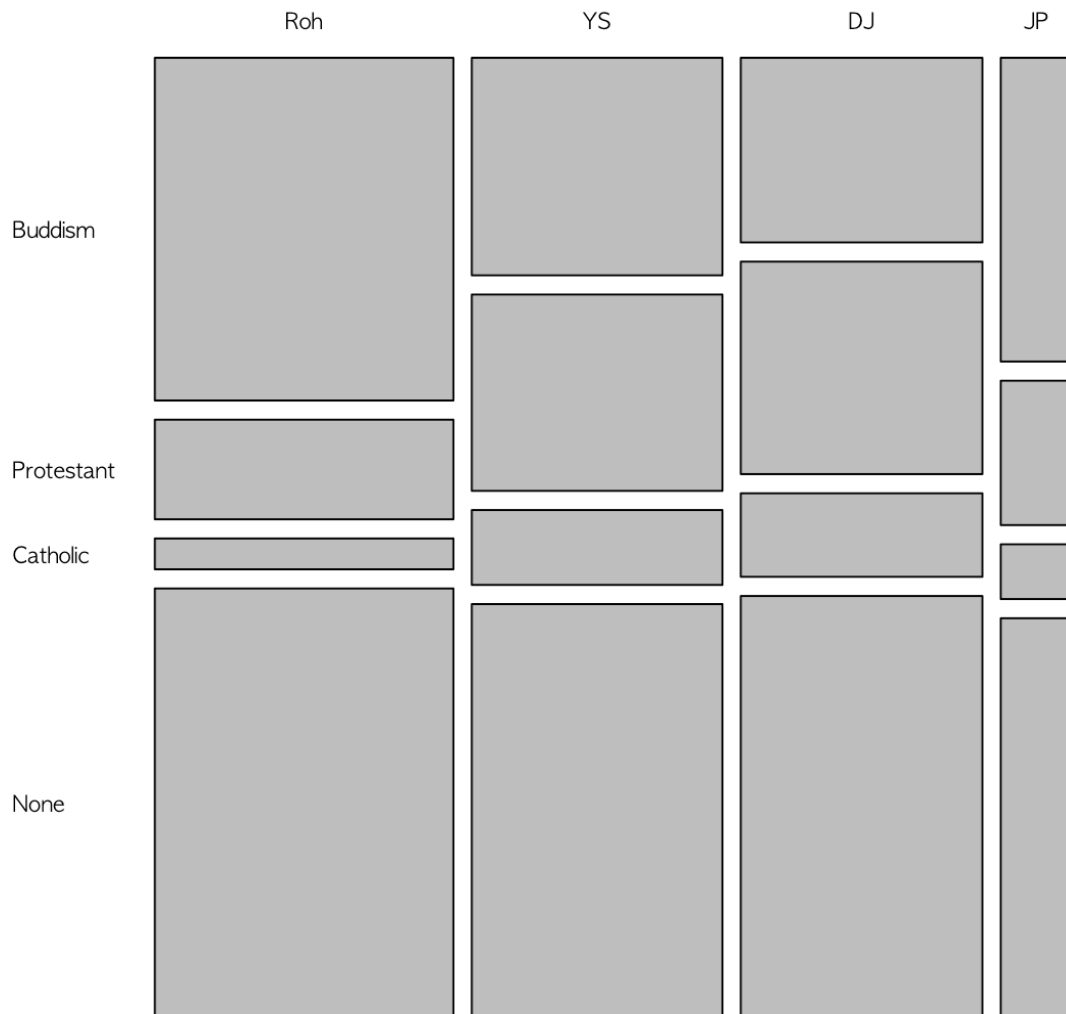
Poll 1987



변수명이 겹치지 않도록 las=1 를 적용하면,

```
mosaicplot(t(poll.87.tbl), las=1, main="Poll 1987")
```

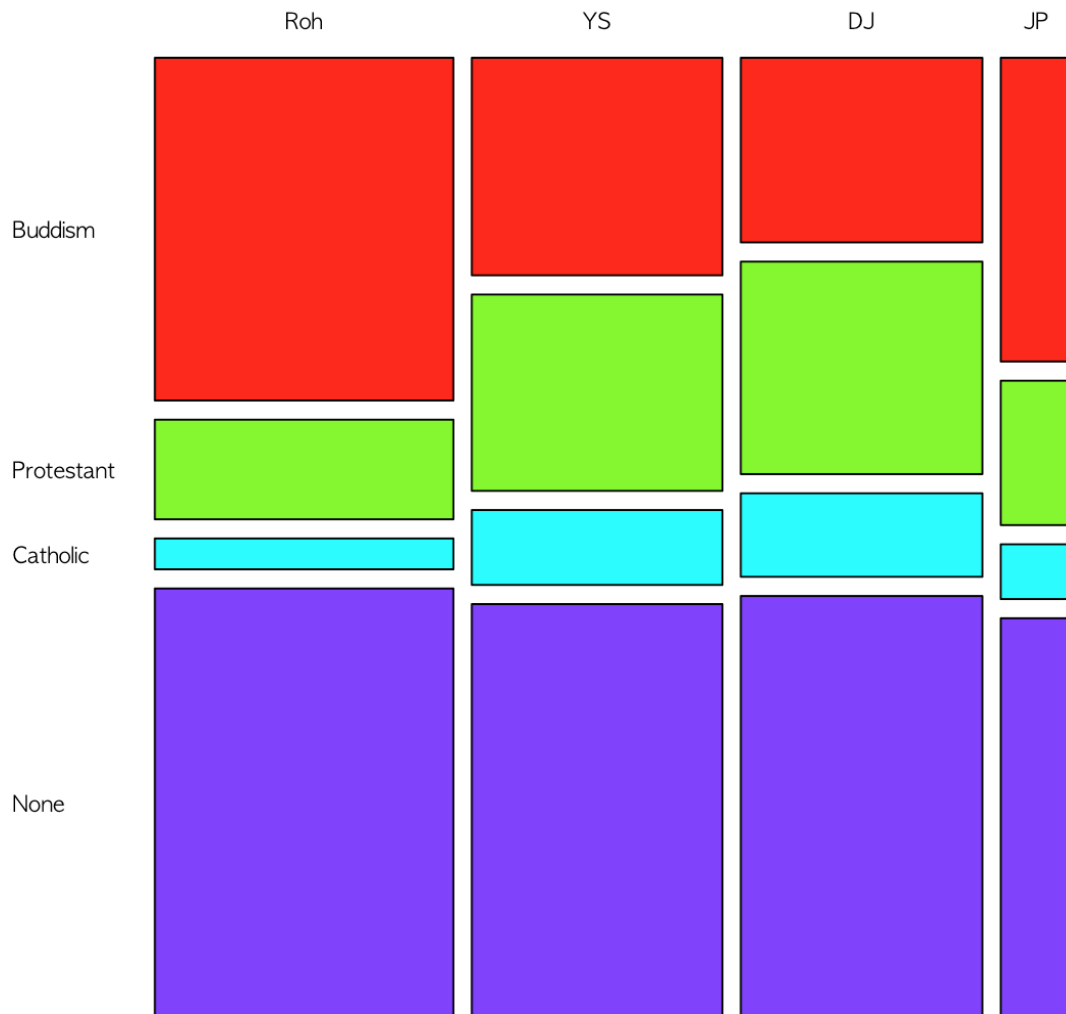
Poll 1987



종교별로 색깔을 씌우면,

```
mosaicplot(t(poll.87.tbl), las=1, main="Poll 1987", color=rainbow(4))
```

Poll 1987



카이제곱 적합도 검정

실제 득표율은 다음 표와 같음.

```
result.87<-c(36.6, 28.0, 27.1, 8.1)
names(result.87)<-dimnames(poll.87.tbl)[[2]]
paste(result.87, "%", sep="")
```

```
## [1] "36.6%" "28%" "27.1%" "8.1%"
```

poll sample은 result.87으로부터의 simple random sample로 볼 수 있는가? 우선, 후보별 예상지지율을 계산하기 위하여 표본지지자수를 합산해 보면,

```
apply(poll.87.tbl, 2, sum)
```

```
## Roh  YS  DJ  JP  
## 760 638 615 181
```

여기에 카이제곱 검정을 적용해 보면,

```
chisq.test(apply(poll.87.tbl, 2, sum), p=result.87/sum(result.87))
```

```
##  
## Chi-squared test for given probabilities  
##  
## data:  apply(poll.87.tbl, 2, sum)  
## X-squared = 3.9616, df = 3, p-value = 0.2656
```

뒷 마무리