

Women and Children First!

20173204 광명빈

2020-09-15

타이타닉호의 침몰 과정에서 **여성과 어린이를 먼저 구한다**는 원칙은 지켜졌는가?

Data

`datasets` 패키지에 들어있으므로 불러들이기만 하면 됨. 자료의 구조 파악.

```
library(magrittr)
library(knitr)
library(pander)
library(ggplot2)
library(extrafont)
```

```
## Registering fonts with R
```

```
data(Titanic)
Titanic %>% str
```

```
## 'table' num [1:4, 1:2, 1:2, 1:2] 0 0 35 0 0 0 17 0 118 154 ...
## - attr(*, "dimnames")=List of 4
## ..$ Class : chr [1:4] "1st" "2nd" "3rd" "Crew"
## ..$ Sex : chr [1:2] "Male" "Female"
## ..$ Age : chr [1:2] "Child" "Adult"
## ..$ Survived: chr [1:2] "No" "Yes"
```

Array

4-차원 array table이므로 보기 쉽게 `ftable` (flat table) 적용. **객실 등급(class)**은 1등실부터 선원까지이고 선원 중에 어린이는 없었으며,

```
Titanic %>% ftable %>% as.matrix %>% kable(align = "c")
```

	No	Yes
1st_Male_Child	0	5
1st_Male_Adult	118	57
1st_Female_Child	0	1
1st_Female_Adult	4	140
2nd_Male_Child	0	11
2nd_Male_Adult	154	14
2nd_Female_Child	0	13

	No	Yes
2nd_Female_Adult	13	80
3rd_Male_Child	35	13
3rd_Male_Adult	387	75
3rd_Female_Child	17	14
3rd_Female_Adult	89	76
Crew_Male_Child	0	0
Crew_Male_Adult	670	192
Crew_Female_Child	0	0
Crew_Female_Adult	3	20

4-dimensional array 인 점을 감안하여 각 변수의 주변합을 구해보면

```
Titanic %>% apply(MARGIN = 1, FUN = sum) %>% as.matrix %>% t %>% kable(align = "c")
```

1st	2nd	3rd	Crew
325	285	706	885

```
Titanic %>% apply(MARGIN = 2, FUN = sum) %>% as.matrix %>% t %>% kable(align = "c")
```

Male	Female
1731	470

```
Titanic %>% apply(MARGIN = 3, FUN = sum) %>% as.matrix %>% t %>% kable(align = "c")
```

Child	Adult
109	2092

```
Titanic %>% apply(MARGIN = 4, FUN = sum) %>% as.matrix %>% t %>% kable(align = "c")
```

No	Yes
1490	711

분할표를 구하되 상황 파악이 편하게 열과 행을 조정.

```
Titanic %>%  
  apply(MARGIN = 1:2, FUN = sum)
```

```
##           Sex
## Class   Male Female
##   1st    180    145
##   2nd    179    106
##   3rd    510    196
##   Crew   862     23
```

```
Titanic %>%
  apply(MARGIN = 2:1, FUN = sum)
```

```
##           Class
## Sex       1st 2nd 3rd Crew
##   Male    180 179 510  862
##   Female  145 106 196   23
```

```
Titanic %>%
  apply(MARGIN = c(3, 1), FUN = sum)
```

```
##           Class
## Age       1st 2nd 3rd Crew
##   Child    6  24  79    0
##   Adult  319 261 627  885
```

```
Surv_Class <- Titanic %>%
  apply(MARGIN = c(4, 1), FUN = sum)
```

Proportions

객실 등급별 생존률을 비교하려면. (우선, 자릿수를 정해 놓고)

```
options(digits = 3)
#> Titanic %>%
#>   apply(MARGIN = c(4, 1), FUN = sum) %>%
Surv_Class %>%
  prop.table(margin = 2) %>%
  `*`(100) %>%
  rbind(., "Sum" = colSums(.))
```

```
##           1st   2nd   3rd Crew
## No    37.5  58.6  74.8   76
## Yes   62.5  41.4  25.2   24
## Sum  100.0 100.0 100.0  100
```

Plots

이를 `barplot` 으로 나타내는 데 있어서 각 `argument` 가 왜 필요한지 시행착오를 거쳐 파악해 볼 것.

```

par(mfrow = c(1, 2), family = "KoPubWorldDotum Medium")
#> Titanic %>%
#>   apply(MARGIN = c(4, 1), FUN = sum)
pos <- function(x){
  cumsum(x) - x / 2
}
pos <- . %>% {`-`(cumsum(.), . / 2)}
b1 <- Surv_Class %>%
  barplot(yaxt = "n", col = rainbow(2))

```

```

## Warning in axis(if (horiz) 2 else 1, at = at.l, labels = names.arg, lty =
## axis.lty, : font family not found in Windows font database

```

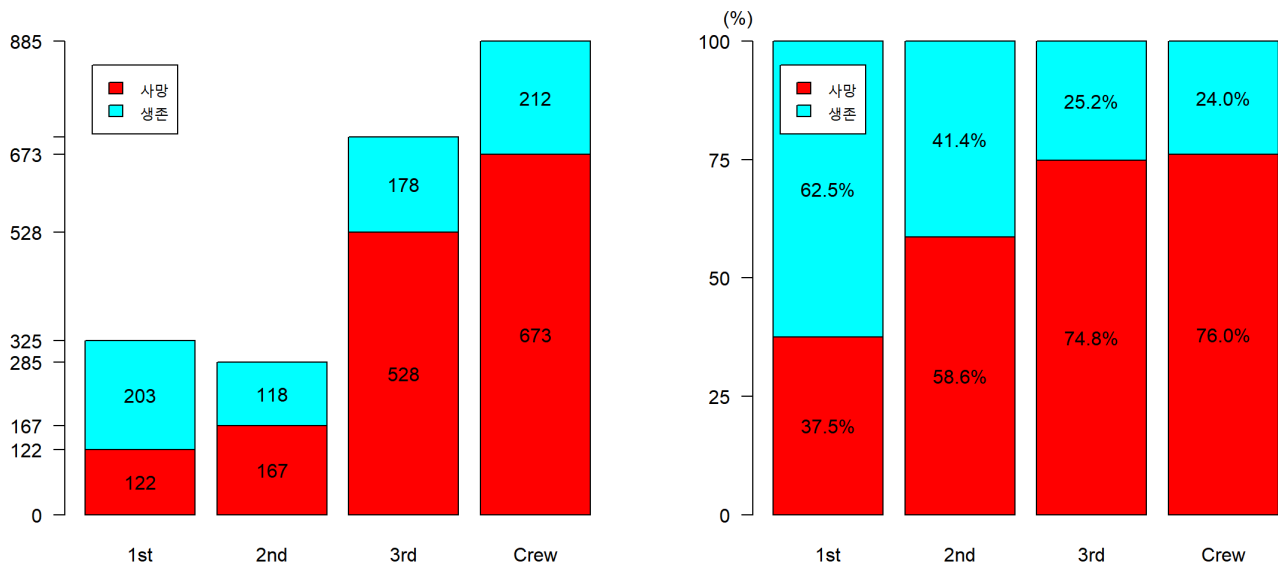
```

axis(side = 2,
      at = Surv_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .),
      labels = Surv_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .),
      las = 2)
y1_text <- apply(Surv_Class,
                 MARGIN = 2,
                 FUN = pos)
# y1_text <- c(Surv_Class[1, ] / 2, Surv_Class[1, ] + Surv_Class[2, ] / 2)
# text(x = rep(b1, times = 2),
#      y = y1_text,
#      labels = c(Surv_Class[1, ], Surv_Class[2, ]))
text(x = rep(b1, each = 2),
     y = y1_text,
     labels = Surv_Class)
legend("topleft", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
#> Titanic %>%
#>   apply(c(4,1), sum)
p1 <- Surv_Class %>%
  prop.table(margin = 2)
b1_p <- p1 %>%
  barplot(yaxt = "n", col = rainbow(2))
axis(side = 2,
      at = seq(0, 1, by = 0.25),
      labels = seq(0, 100, by = 25),
      las = 2)
mtext(" (%)", side = 2, at = 1.05, line = 0, las = 2)
p1_text <- apply(p1,
                MARGIN = 2,
                FUN = pos)
# p1_text <- c(p1[1, ] / 2, p1[1, ] + p1[2, ] / 2)
text(x = b1_p %>%
      rep(each = 2),
     y = p1_text,
     labels = p1 %>%
       `*`(100) %>%
       format(digits = 2, nsmall = 1) %>%
       paste0("%"))
legend("topleft", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
title(main = "객실 등급별 생존/사망", line = -1, outer = TRUE, cex.main = 1.5,
      family = "KoPubWorldDotum Bold")

```

```
## Warning in title(main = "객실 등급별 생존/사망", line = -1, outer = TRUE, : font
## family not found in Windows font database
```

객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_barplot01.png", width = 840, height = 420)
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot01.png", width = 840, height =
## 420): font family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot01.png", width = 840, height =
## 420): font family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

Mosaic Plot

```
par(mfrow = c(1, 1), family = "KoPubWorldDotum Medium")
mosaicplot(t(Surv_Class), main = "객실 등급별 생존/사망",
           xlab = "객실 등급", ylab = "생존/사망",
           col = rainbow(2))
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not
## found in Windows font database
```

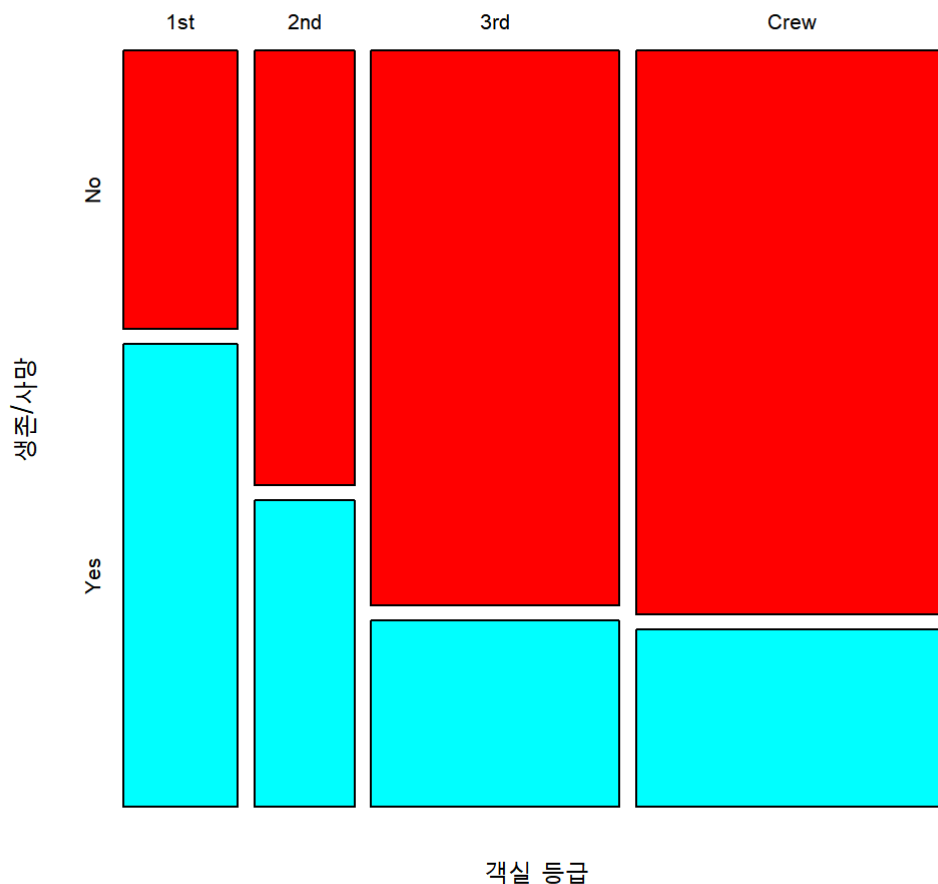
```
## Warning in text.default(x = x.l + (x.r - x.l)/2, y = 1000 - 35 * cex.axis/0.66
## + : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, height = 320)
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot01.png", width = 320, : font
## family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

성별 생존/사망

Cross-table 을 계속 작성해 가자면

```
Titanic %>% apply(MARGIN = 2:3, FUN = sum)
```

```
##           Age
## Sex      Child Adult
## Male      64  1667
## Female    45   425
```

```
Titanic %>% apply(MARGIN = c(2,4), FUN = sum)
```

```
##           Survived
## Sex           No Yes
## Male    1364 367
## Female   126 344
```

```
Surv_Sex <- Titanic %>%
  apply(MARGIN = c(4,2), FUN = sum)
```

남녀 생존률을 비교하려면,

```
Surv_Sex %>%  
  prop.table(margin = 2) %>%  
  `*`(100) %>%  
  rbind(., "Sum" = colSums(.))
```

```
##      Male Female  
## No    78.8    26.8  
## Yes   21.2    73.2  
## Sum  100.0   100.0
```

Plots

이를 `barplot` 으로 나타내는 데 있어서 각 `argument` 가 왜 필요한지 시행착오를 겪어 볼 것.

```
par(mfrow = c(1, 2), family = "KoPubWorldDotum Medium")  
b2 <- Surv_Sex %>%  
  barplot(yaxt = "n", col = rainbow(2))
```

```
## Warning in axis(if (horiz) 2 else 1, at = at.l, labels = names.arg, lty =  
## axis.lty, : font family not found in Windows font database
```



```

axis(side = 2,
      at = Surv_Sex %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .),
      labels = Surv_Sex %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .) %>%
        format(big.mark = ","),
      las = 2)
# y2_text <- c(Surv_Sex[1, ] / 2, Surv_Sex[1, ] + Surv_Sex[2, ] / 2)
y2_text <- apply(Surv_Sex,
                 MARGIN = 2,
                 FUN = pos)
text(x = rep(b2, each = 2),
     y = y2_text,
     labels = Surv_Sex %>%
       format(big.mark = ","))
legend("topright", inset = 0.15, fill = rainbow(2), legend = c("사망", "생존"))
p2 <- Surv_Sex %>%
  prop.table(margin = 2)
b2_p <- p2 %>%
  barplot(yaxt = "n", col = rainbow(2))
axis(side = 2,
      at = seq(0, 1, by = 0.25),
      labels = seq(0, 100, by = 25),
      las = 2)
mtext("("%)", side = 2, at = 1.05, line = 0, las = 2)
# p2_text <- c(p2[1, ] / 2, p2[1, ] + p2[2, ] / 2)
p2_text <- apply(p2,
                 MARGIN = 2,
                 FUN = pos)
text(x = b2_p %>%
      rep(each = 2),
     y = p2_text,
     labels = p2 %>%
       `*`(100) %>%
       format(digits = 2, nsmall = 1) %>%
       paste0("%"))
legend("topright", inset = 0.15, fill = rainbow(2), legend = c("사망", "생존"))
title(main = "성별 생존/사망", line = -1, outer = TRUE, cex.main = 1.5,
      family = "KoPubWorldDotum Bold")

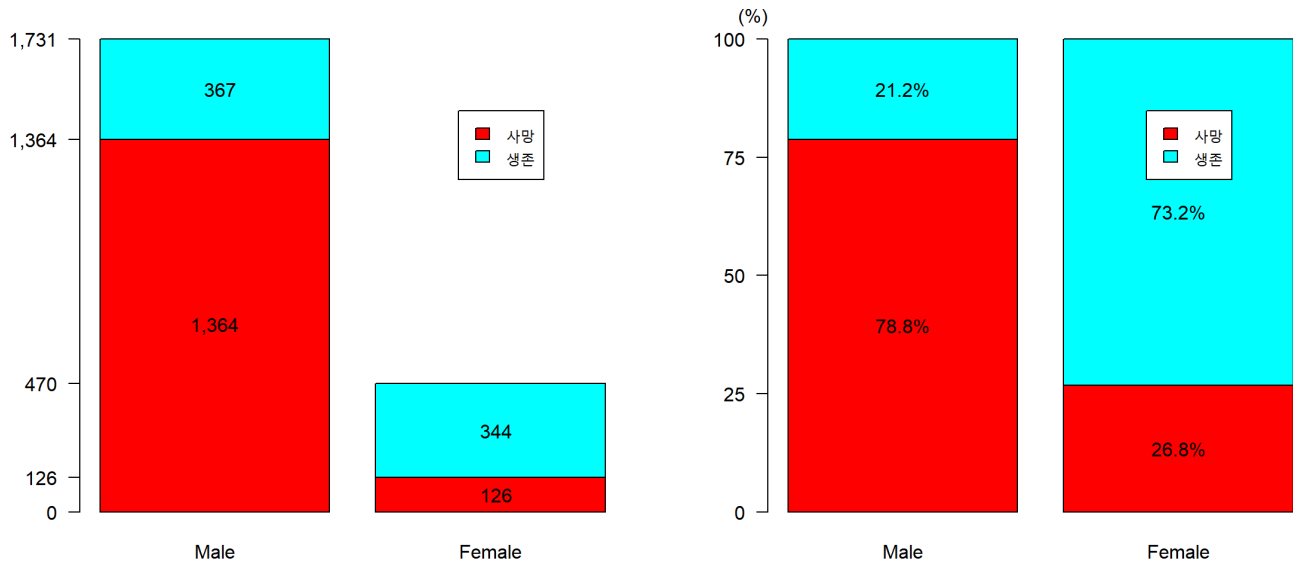
```

```

## Warning in title(main = "성별 생존/사망", line = -1, outer = TRUE, cex.main =
## 1.5, : font family not found in Windows font database

```

성별 생존/사망



```
dev.copy(png, "../pics/Titanic_barplot02.png", width = 840, height = 420)
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot02.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot02.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## png  
## 3
```

```
dev.off()
```

```
## png  
## 2
```

Mosaic Plot

```
par(mfrow = c(1, 1), family = "KoPubWorldDotum Medium")  
mosaicplot(t(Surv_Sex), main = "성별 생존/사망",  
           xlab = "남/여", ylab = "생존/사망",  
           col = rainbow(2))
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

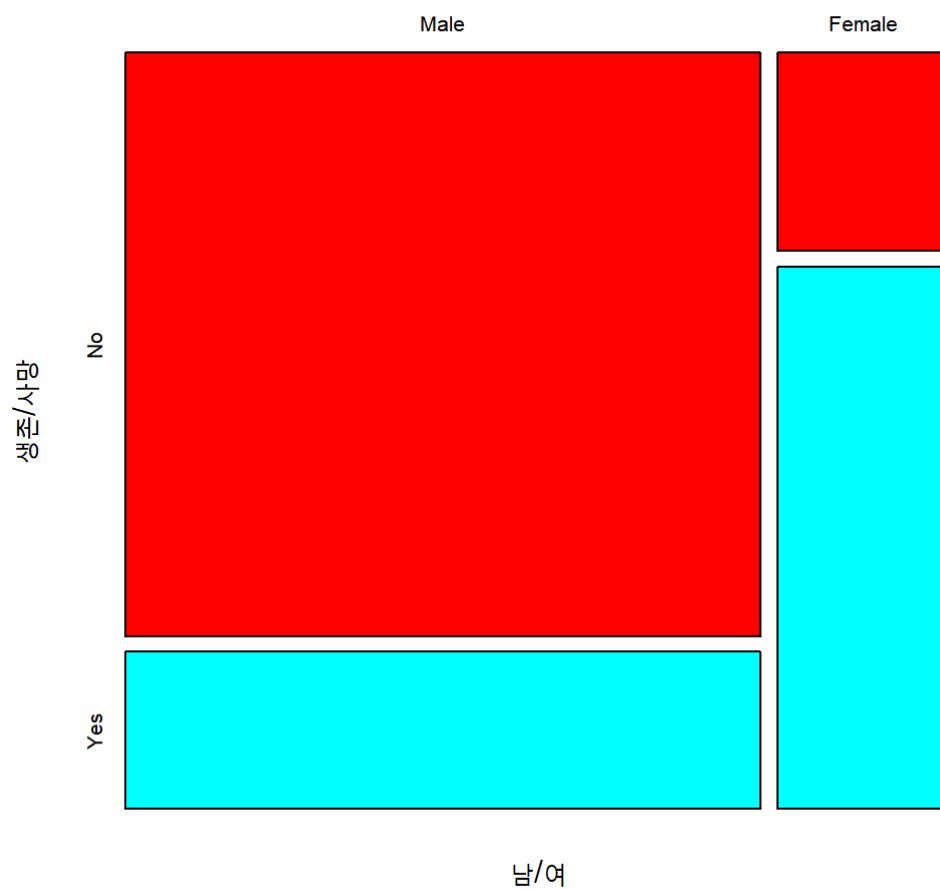
```
## Warning in text.default(x = x.l + (x.r - x.l)/2, y = 1000 - 35 * cex.axis/0.66
## + : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

성별 생존/사망



```
dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, height = 320)
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot02.png", width = 320, : font
## family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

연령별 생존/사망

남은 cross-table 은

```
Surv_Age <- Titanic %>%
  apply(MARGIN = 4:3, FUN = sum)
```

성인과 어린이의 생존률을 비교하려면

남녀 생존률을 비교하려면,

```
Surv_Age %>%
  prop.table(margin = 2) %>%
  `*`(100) %>%
  rbind(., "Sum" = colSums(.))
```

```
##      Child Adult
## No   47.7  68.7
## Yes  52.3  31.3
## Sum 100.0 100.0
```

Plots

이를 barplot 으로 나타내는 데 있어서 각 argument 가 왜 필요한지 시행착오를 겪어 볼 것.

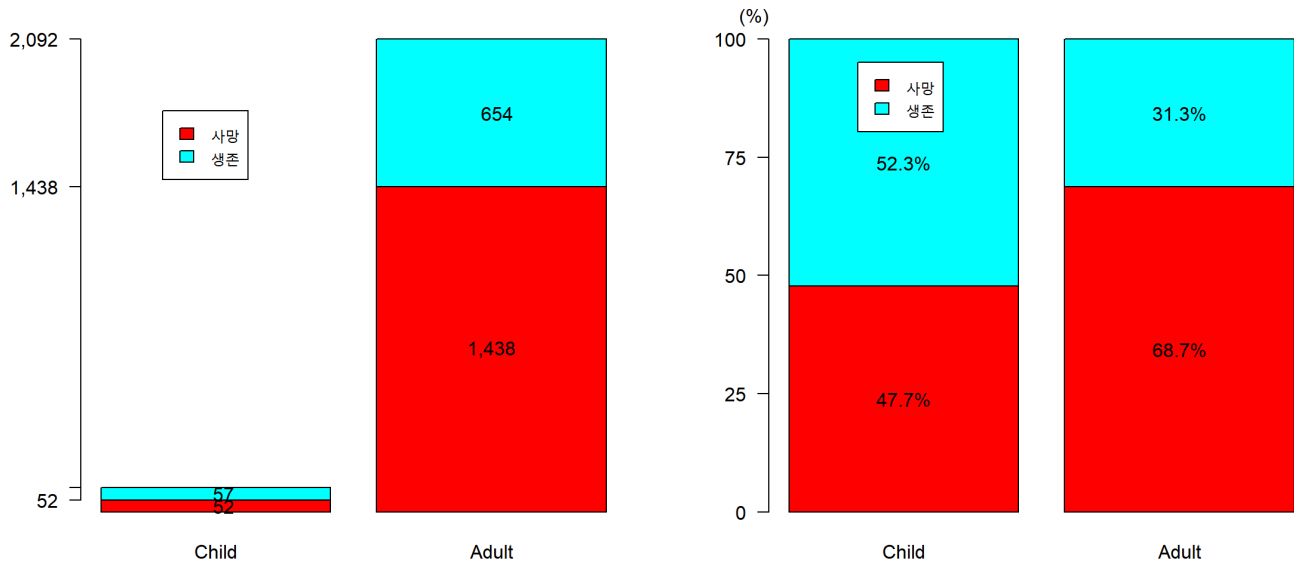
```
par(mfrow = c(1, 2), family = "KoPubWorldDotum Medium")
b3 <- Surv_Age %>%
  barplot(yaxt = "n", col = rainbow(2))
```

```
## Warning in axis(if (horiz) 2 else 1, at = at.l, labels = names.arg, lty =
## axis.lty, : font family not found in Windows font database
```

```
axis(side = 2,
      at = Surv_Age %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c,
      labels = Surv_Age %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c %>%
        format(big.mark = ","),
      las = 2)
# y3_text <- c(Surv_Age[1, ] / 2, Surv_Age[1, ] + Surv_Age[2, ] / 2)
y3_text <- apply(Surv_Age,
                 MARGIN = 2,
                 FUN = pos)
text(x = rep(b3, each = 2),
     y = y3_text,
     labels = Surv_Age %>%
       format(big.mark = ","))
legend("topleft", inset = 0.15, fill = rainbow(2), legend = c("사망", "생존"))
p3 <- Surv_Age %>%
  prop.table(margin = 2)
b3_p <- p3 %>%
  barplot(yaxt = "n", col = rainbow(2))
axis(side = 2,
      at = seq(0, 1, by = 0.25),
      labels = seq(0, 100, by = 25),
      las = 2)
mtext("(%)", side = 2, at = 1.05, line = 0, las = 2)
# p3_text <- c(p3[1, ] / 2, p3[1, ] + p3[2, ] / 2)
p3_text <- apply(p3,
                 MARGIN = 2,
                 FUN = pos)
text(x = b3_p %>%
      rep(each = 2),
     y = p3_text,
     labels = p3 %>%
       `*`(100) %>%
       format(digits = 2, nsmall = 1) %>%
       paste0("%"))
legend(x = 0.5, y = 0.95, fill = rainbow(2), legend = c("사망", "생존"))
title(main = "연령별 생존/사망", line = -1, outer = TRUE, cex.main = 1.5,
      family = "KoPubWorldDotum Bold")
```

```
## Warning in title(main = "연령별 생존/사망", line = -1, outer = TRUE, cex.main =
## 1.5, : font family not found in Windows font database
```

연령별 생존/사망



```
dev.copy(png, "../pics/Titanic_barplot03.png", width = 840, height = 420)
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot03.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot03.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## png  
## 3
```

```
dev.off()
```

```
## png  
## 2
```

Mosaic Plot

```
par(mfrow = c(1, 1), family = "KoPubWorldDotum Medium")  
mosaicplot(t(Surv_Age), main = "연령별 생존/사망",  
           xlab = "어린이/어른", ylab = "생존/사망",  
           col = rainbow(2))
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

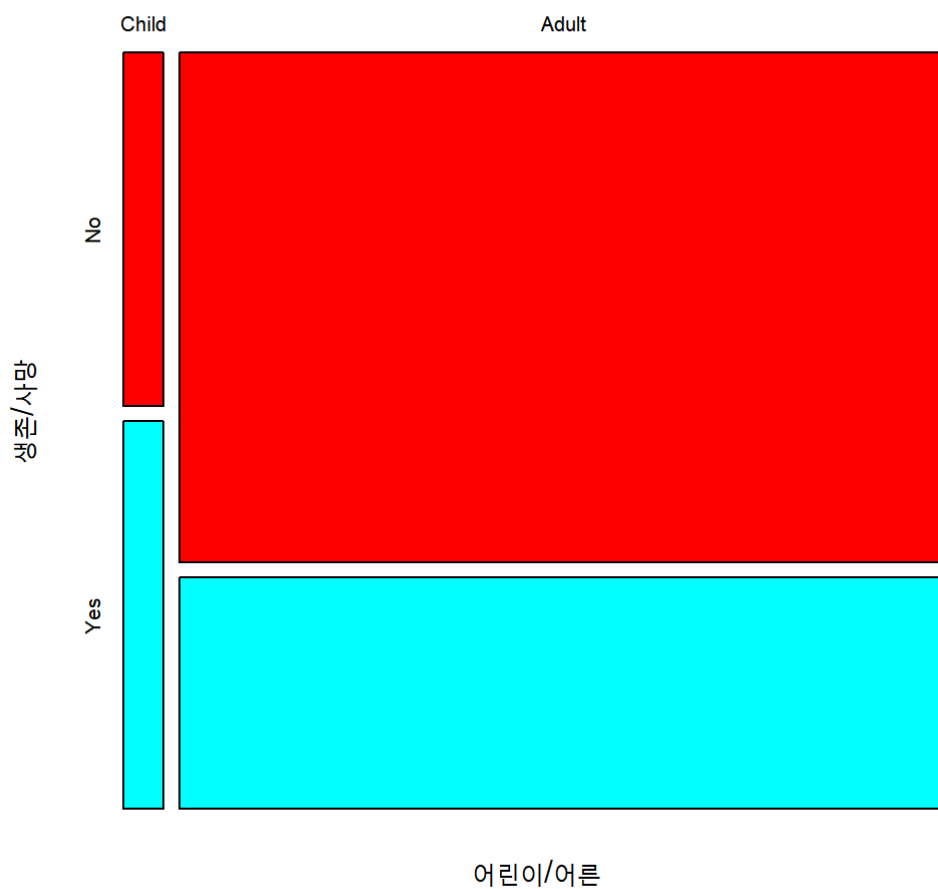
```
## Warning in text.default(x = x.l + (x.r - x.l)/2, y = 1000 - 35 * cex.axis/0.66
## + : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

연령별 생존/사망



```
dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, height = 320)
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot03.png", width = 320, : font
## family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

어린이들의 객실 등급별 생존/사망

객실 등급별로 어린이들과 어른들의 생존/사망을 비교하려면

```
Child_df <- Titanic %>%
  as.data.frame %>%
  subset(Age == "Child")
Adult_df <- Titanic %>%
  as.data.frame %>%
  subset(Age == "Adult")
Child_Class <- Child_df %>%
  xtabs(Freq ~ Survived + Class, data = ., drop.unused.levels = TRUE)
Child_Class %>%
  prop.table(margin = 2) %>%
  `*`(100) %>%
  rbind(., "Sum" = colSums(.))
```

```
##      1st 2nd   3rd Crew
## No    0   0  65.8  NaN
## Yes 100 100  34.2  NaN
## Sum 100 100 100.0  NaN
```

Plots

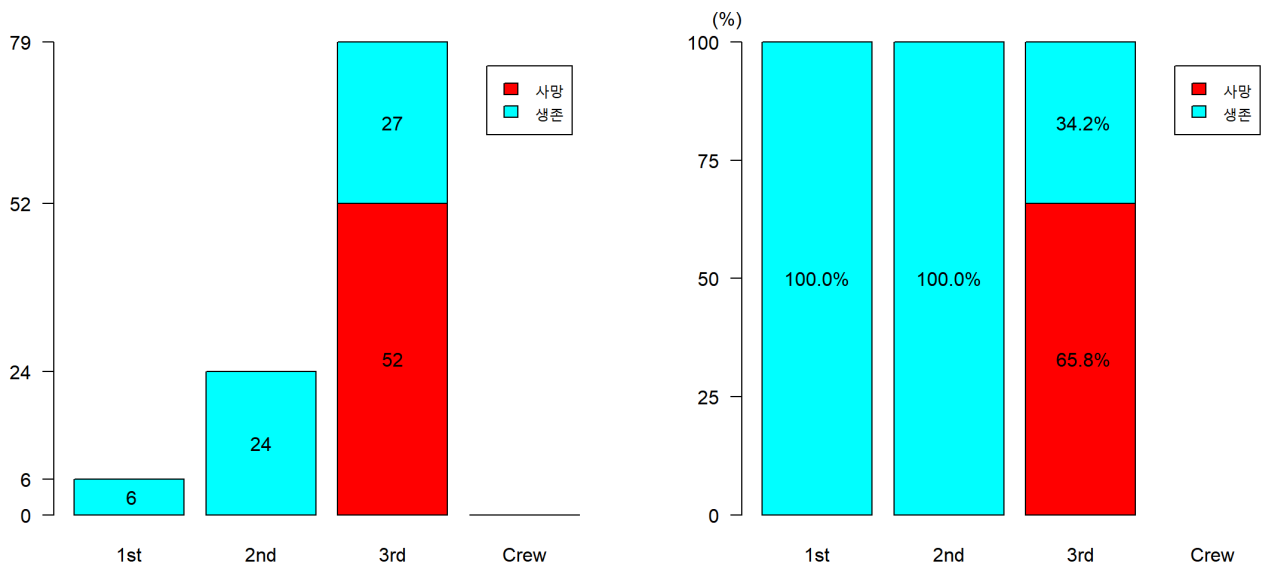

```
par(mfrow = c(1, 2), family = "KoPubWorldDotum Medium")
b4 <- Child_Class %>%
  barplot(yaxt = "n", col = rainbow(2))
```

```
## Warning in axis(if (horiz) 2 else 1, at = at.l, labels = names.arg, lty =
## axis.lty, : font family not found in Windows font database
```

```
axis(side = 2,
      at = Child_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .),
      labels = Child_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c(0, .) %>%
        format(big.mark = ","),
      las = 2)
# y4_text <- c(Child_Class[1, ] / 2, Child_Class[1, ] + Child_Class[2, ] / 2)
y4_text <- apply(Child_Class,
                 MARGIN = 2,
                 FUN = pos)
y4_text[c(1, 3, 7, 8)] <- NA
text(x = rep(b4, each = 2),
     y = y4_text,
     labels = Child_Class %>%
       format(big.mark = ","))
legend("topright", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
p4 <- Child_Class %>%
  prop.table(margin = 2)
b4_p <- p4 %>%
  barplot(yaxt = "n", col = rainbow(2))
axis(side = 2,
      at = seq(0, 1, by = 0.25),
      labels = seq(0, 100, by = 25),
      las = 2)
mtext("("%", side = 2, at = 1.05, line = 0, las = 2)
# p4_text <- c(p4[1, ] / 2, p4[1, ] + p4[2, ] / 2)
p4_text <- apply(p4,
                 MARGIN = 2,
                 FUN = pos)
p4_text[c(1, 3)] <- NA
text(x = b4_p %>%
      rep(each = 2),
     y = p4_text,
     labels = p4 %>%
       `*`(100) %>%
       format(digits = 2, nsmall = 1) %>%
       paste0("%"))
legend("topright", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
title(main = "어린이들의 객실 등급별 생존/사망", line = -1, outer = TRUE, cex.main = 1.5,
      family = "KoPubWorldDotum Bold")
```

```
## Warning in title(main = "어린이들의 객실 등급별 생존/사망", line = -1, outer =
## TRUE, : font family not found in Windows font database
```

어린이들의 객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_barplot04.png", width = 840, height = 420)
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot04.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot04.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## png  
## 3
```

```
dev.off()
```

```
## png  
## 2
```

Mosaic Plot

```
par(mfrow = c(1, 1), family = "KoPubWorldDotum Medium")  
mosaicplot(t(Child_Class), main = "어린이들의 객실 등급별 생존/사망",  
           xlab = "객실 등급", ylab = "생존/사망",  
           col = rainbow(2))
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

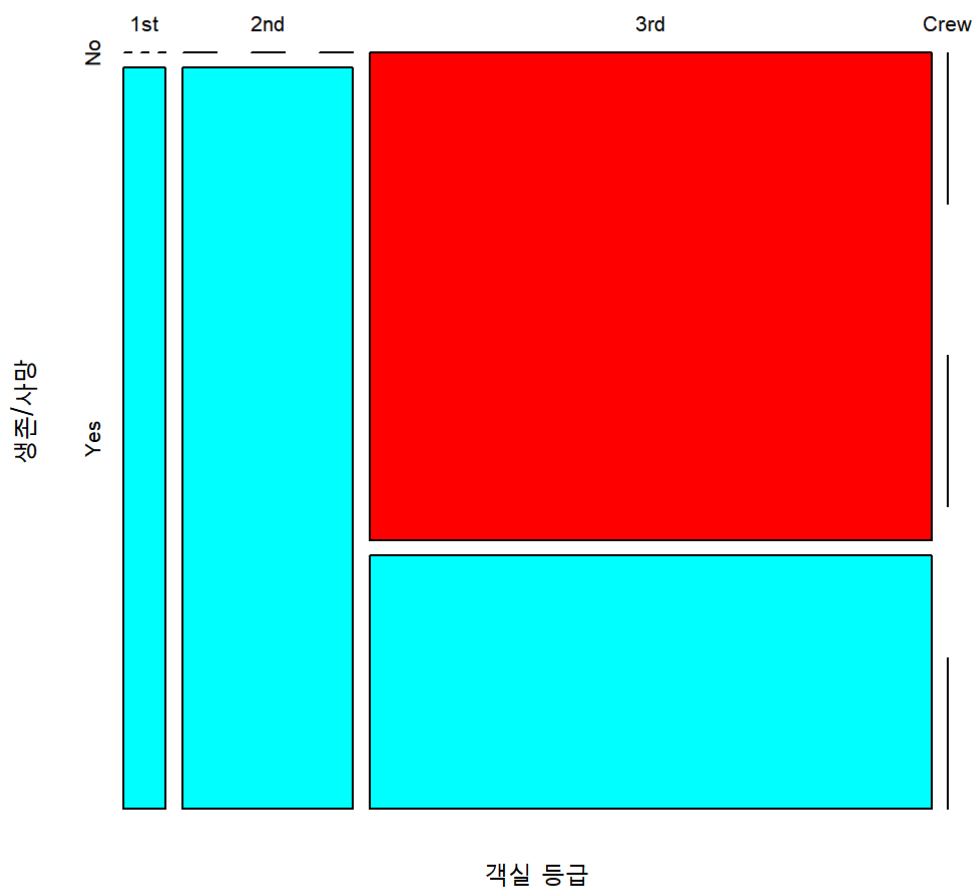
```
## Warning in text.default(x = x.l + (x.r - x.l)/2, y = 1000 - 35 * cex.axis/0.66
## + : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

어린이들의 객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, height = 320)
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot04.png", width = 320, : font
## family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

여성들의 등급별 생존/사망

객실 등급별로 여성들과 남성들의 생존률을 비교하려면

```
Female_df <- Titanic %>%
  as.data.frame %>%
  subset(Sex == "Female")
Male_df <- Titanic %>%
  as.data.frame %>%
  subset(Sex == "Male")
Female_Class <- Female_df %>%
  xtabs(Freq ~ Survived + Class, data = ., drop.unused.levels = TRUE)
Female_Class %>%
  prop.table(margin = 2) %>%
  `*`(100) %>%
  rbind(., "Sum" = colSums(.))
```

```
##      1st   2nd   3rd Crew
## No    2.76  12.3  54.1   13
## Yes  97.24  87.7  45.9   87
## Sum 100.00 100.0 100.0  100
```

Plots

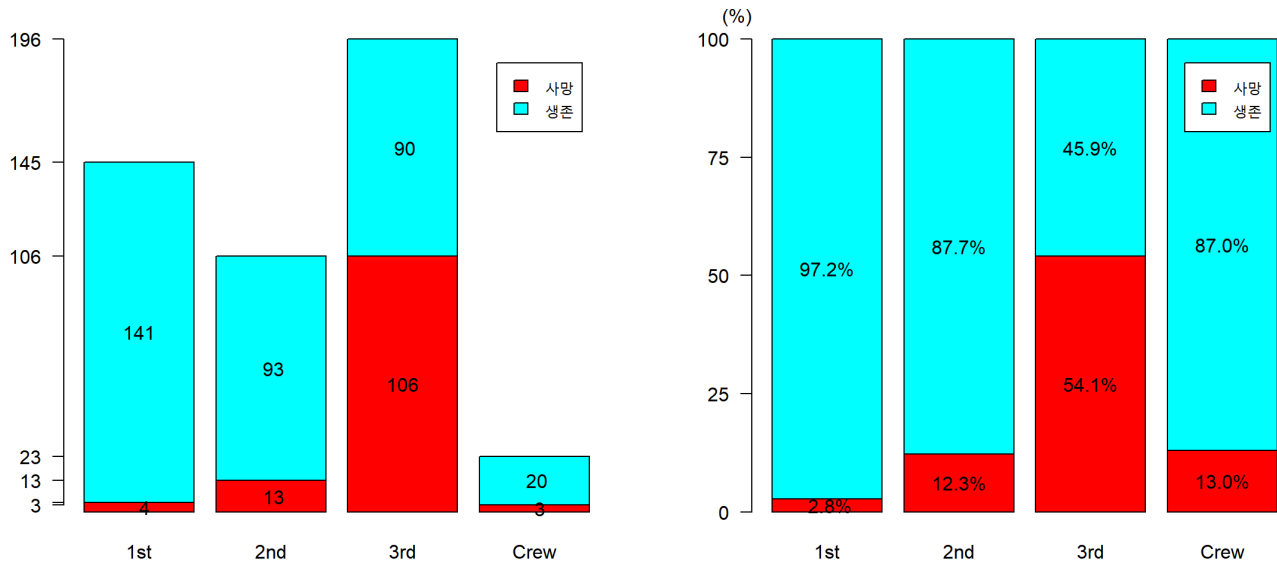
```
par(mfrow = c(1, 2), family = "KoPubWorldDotum Medium")
b5 <- Female_Class %>%
  barplot(yaxt = "n", col = rainbow(2))
```

```
## Warning in axis(if (horiz) 2 else 1, at = at.l, labels = names.arg, lty =
## axis.lty, : font family not found in Windows font database
```

```
axis(side = 2,
      at = Female_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c,
      labels = Female_Class %>%
        apply(MARGIN = 2, FUN = cumsum) %>% c %>%
        format(big.mark = ","),
      las = 2)
# y5_text <- c(Female_Class[1, ] / 2, Female_Class[1, ] + Female_Class[2, ] / 2)
y5_text <- apply(Female_Class,
                 MARGIN = 2,
                 FUN = pos)
#> y5_text[c(1:2, 4, 8)] <- NA
text(x = rep(b5, each = 2),
     y = y5_text,
     labels = Female_Class %>%
       format(big.mark = ","))
legend("topright", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
p5 <- Female_Class %>%
  prop.table(margin = 2)
b5_p <- p5 %>%
  barplot(yaxt = "n", col = rainbow(2))
axis(side = 2,
      at = seq(0, 1, by = 0.25),
      labels = seq(0, 100, by = 25),
      las = 2)
mtext(" (%)", side = 2, at = 1.05, line = 0, las = 2)
# p5_text <- c(p5[1, ] / 2, p5[1, ] + p5[2, ] / 2)
p5_text <- apply(p5,
                 MARGIN = 2,
                 FUN = pos)
#> p5_text[1:2] <- NA
text(x = b5_p %>%
      rep(each = 2),
     y = p5_text,
     labels = p5 %>%
       `*`(100) %>%
       format(digits = 2, nsmall = 1) %>%
       paste0("%"))
legend("topright", inset = 0.05, fill = rainbow(2), legend = c("사망", "생존"))
title(main = "여성들의 객실 등급별 생존/사망", line = -1, outer = TRUE, cex.main = 1.5,
      family = "KoPubWorldDotum Bold")
```

```
## Warning in title(main = "여성들의 객실 등급별 생존/사망", line = -1, outer =
## TRUE, : font family not found in Windows font database
```

여성들의 객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_barplot05.png", width = 840, height = 420)
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot05.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_barplot05.png", width = 840, height =  
## 420): font family not found in Windows font database
```

```
## png  
## 3
```

```
dev.off()
```

```
## png  
## 2
```

Mosaic Plot

```
par(mfrow = c(1, 1), family = "KoPubWorldDotum Medium")  
mosaicplot(t(Female_Class),  
           main = "여성들의 객실 등급별 생존/사망", xlab = "객실 등급", ylab = "생존/사망",  
           col = rainbow(2))
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

```
## Warning in title(main, sub = sub, xlab = xlab, ylab = ylab): font family not  
## found in Windows font database
```

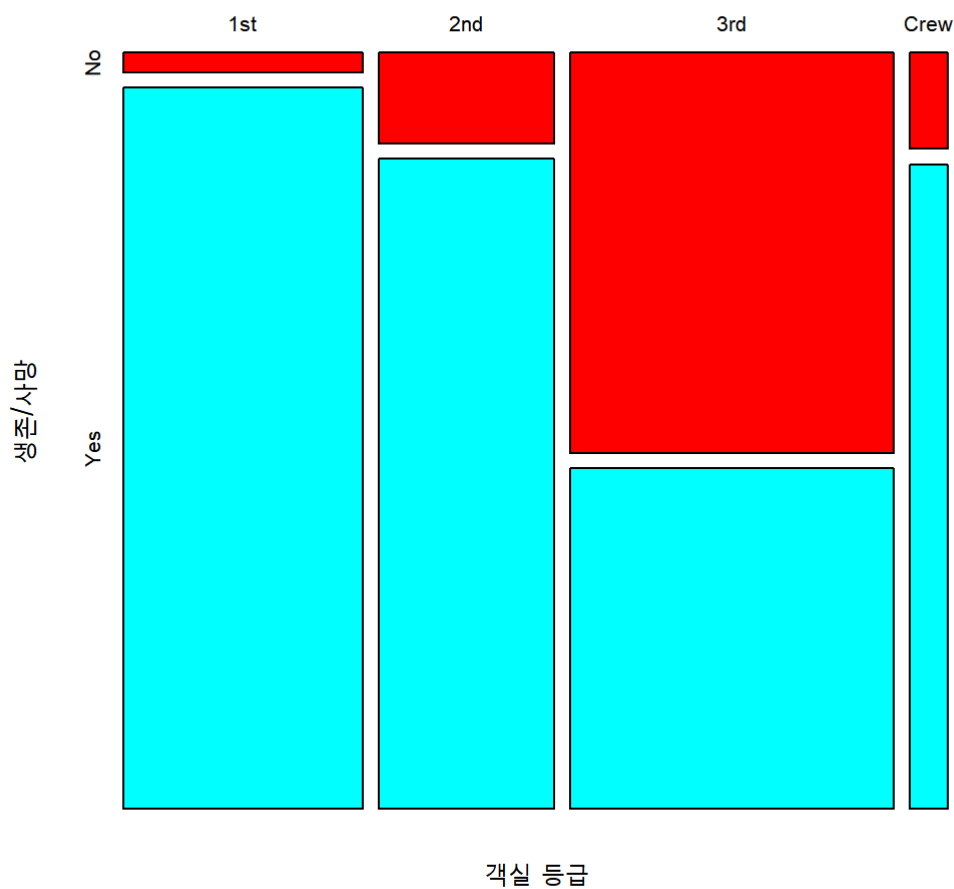
```
## Warning in text.default(x = x.l + (x.r - x.l)/2, y = 1000 - 35 * cex.axis/0.66
## + : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

```
## Warning in text.default(x = 35 * cex.axis/0.66 - 20 * cex.axis/0.66 * (lablevy
## - : font family not found in Windows font database
```

여성들의 객실 등급별 생존/사망



```
dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, height = 320)
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database
```

```
## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database

## Warning in dev.copy(png, "../pics/Titanic_mosaicplot05.png", width = 320, : font
## family not found in Windows font database
```

```
## png
## 3
```

```
dev.off()
```

```
## png
## 2
```

Save

Comments

이번시간에는 타이타닉자료를 가지고 데이터시각화를 해보았습니다. `fable`을 사용하여 더 쉽게 데이터를 확인할 수 있었고 `%>%`를 계속 사용하면서 `%>%`에 대한 이해를 높일 수 있었습니다. `plot`을 이용하여 알아 보았는데, 1등석의 생존률이 다른 등급에 비해 높다는 것을 알 수 있었습니다. 또한 승무원들의 사망률이 높은것으로 보아 승무원들의 희생에 대한 짐작을 할 수 있었습니다. 남성은 78%가량 죽고 여성은 72% 사망 살았고, 어린이는 표본이 적지만 그래도 남성에 비해 많은 생존률을 가지고 있었습니다. 또한 등급별 생존률이 많이 차이나는것으로 보아, 구조할때 자본이 생명에 영향을 끼쳤다는것을 조심스럽게 예상하게 되었습니다.