

Women and Children First!

20173204곽명빈

2020-09-22

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

Data

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

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

```
load("Titanic_Base.RData")
ls()
```

```
## [1] "Adult_df"      "b1"            "b1_p"          "b2"            "b2_p"
## [6] "b3"           "b3_p"          "b4"            "b4_p"          "b5"
## [11] "b5_p"         "Child_Class"   "Child_df"      "Female_Class"  "Female_df"
## [16] "Male_df"      "p1"            "p1_text"       "p2"            "p2_text"
## [21] "p3"           "p3_text"       "p4"            "p4_text"       "p5"
## [26] "p5_text"      "pos"           "Surv_Age"      "Surv_Class"    "Surv_Sex"
## [31] "Titanic"      "y1_text"       "y2_text"       "y3_text"       "y4_text"
## [36] "y5_text"
```

Source User Defined Functions

```
source("barplot_gg_v2.R")
source("mosaic_gg_new.R")
```

Data Manipulation

```
Surv_Class %>%
  str
```

```
## num [1:2, 1:4] 122 203 167 118 528 178 673 212
## - attr(*, "dimnames")=List of 2
## ..$ Survived: chr [1:2] "No" "Yes"
## ..$ Class : chr [1:4] "1st" "2nd" "3rd" "Crew"
```

```
Surv_Class %>%  
  as.table %>%  
  str
```

```
## 'table' num [1:2, 1:4] 122 203 167 118 528 178 673 212  
## - attr(*, "dimnames")=List of 2  
## ..$ Survived: chr [1:2] "No" "Yes"  
## ..$ Class : chr [1:4] "1st" "2nd" "3rd" "Crew"
```

```
Surv_Sex %>%  
  str
```

```
## num [1:2, 1:2] 1364 367 126 344  
## - attr(*, "dimnames")=List of 2  
## ..$ Survived: chr [1:2] "No" "Yes"  
## ..$ Sex : chr [1:2] "Male" "Female"
```

```
Surv_Sex %>%  
  as.table %>%  
  str
```

```
## 'table' num [1:2, 1:2] 1364 367 126 344  
## - attr(*, "dimnames")=List of 2  
## ..$ Survived: chr [1:2] "No" "Yes"  
## ..$ Sex : chr [1:2] "Male" "Female"
```

```
Surv_Age %>%  
  str
```

```
## num [1:2, 1:2] 52 57 1438 654  
## - attr(*, "dimnames")=List of 2  
## ..$ Survived: chr [1:2] "No" "Yes"  
## ..$ Age : chr [1:2] "Child" "Adult"
```

```
Surv_Age %>%  
  as.table %>%  
  str
```

```
## 'table' num [1:2, 1:2] 52 57 1438 654  
## - attr(*, "dimnames")=List of 2  
## ..$ Survived: chr [1:2] "No" "Yes"  
## ..$ Age : chr [1:2] "Child" "Adult"
```

```
Female_Class %>%  
  str
```

```
## 'xtabs' num [1:2, 1:4] 4 141 13 93 106 90 3 20
## - attr(*, "dimnames")=List of 2
## ..$ Survived: chr [1:2] "No" "Yes"
## ..$ Class : chr [1:4] "1st" "2nd" "3rd" "Crew"
## - attr(*, "call")= language xtabs(formula = Freq ~ Survived + Class, data = ., drop.unused.
levels = TRUE)
```

```
Child_Class %>%
  str
```

```
## 'xtabs' num [1:2, 1:4] 0 6 0 24 52 27 0 0
## - attr(*, "dimnames")=List of 2
## ..$ Survived: chr [1:2] "No" "Yes"
## ..$ Class : chr [1:4] "1st" "2nd" "3rd" "Crew"
## - attr(*, "call")= language xtabs(formula = Freq ~ Survived + Class, data = ., drop.unused.
levels = TRUE)
```

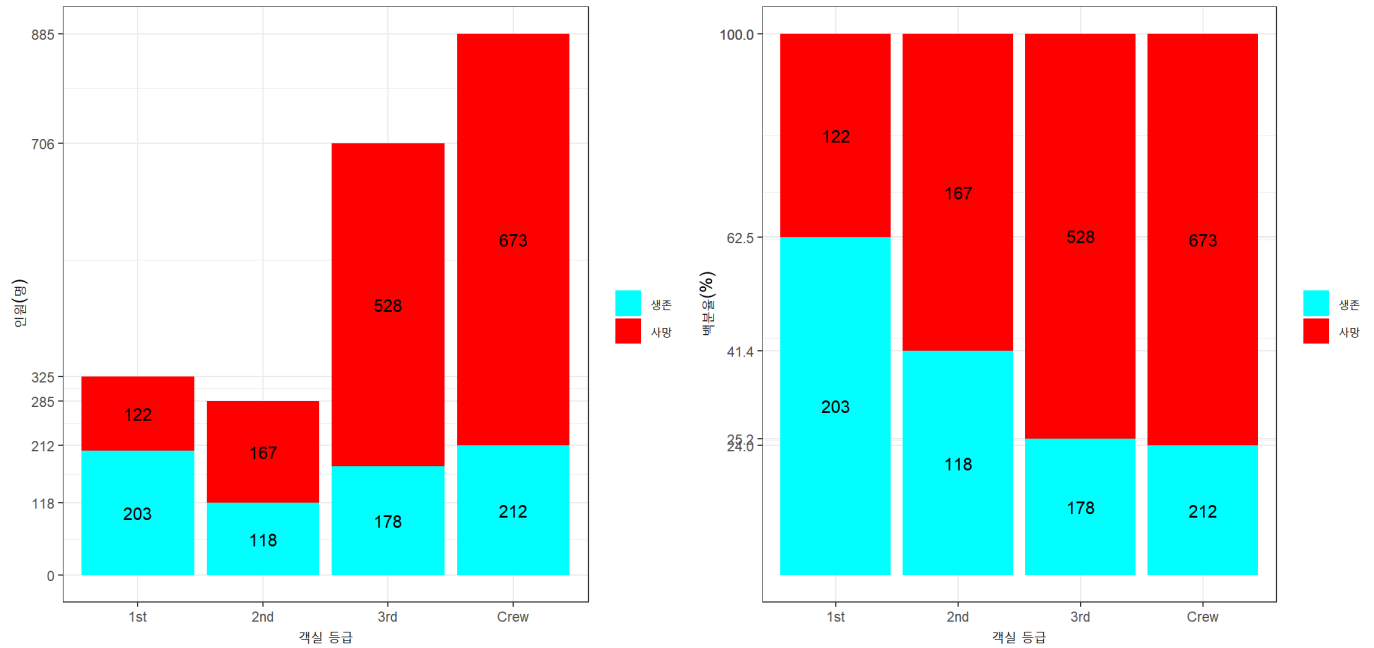
Plots

By Class

Barplots

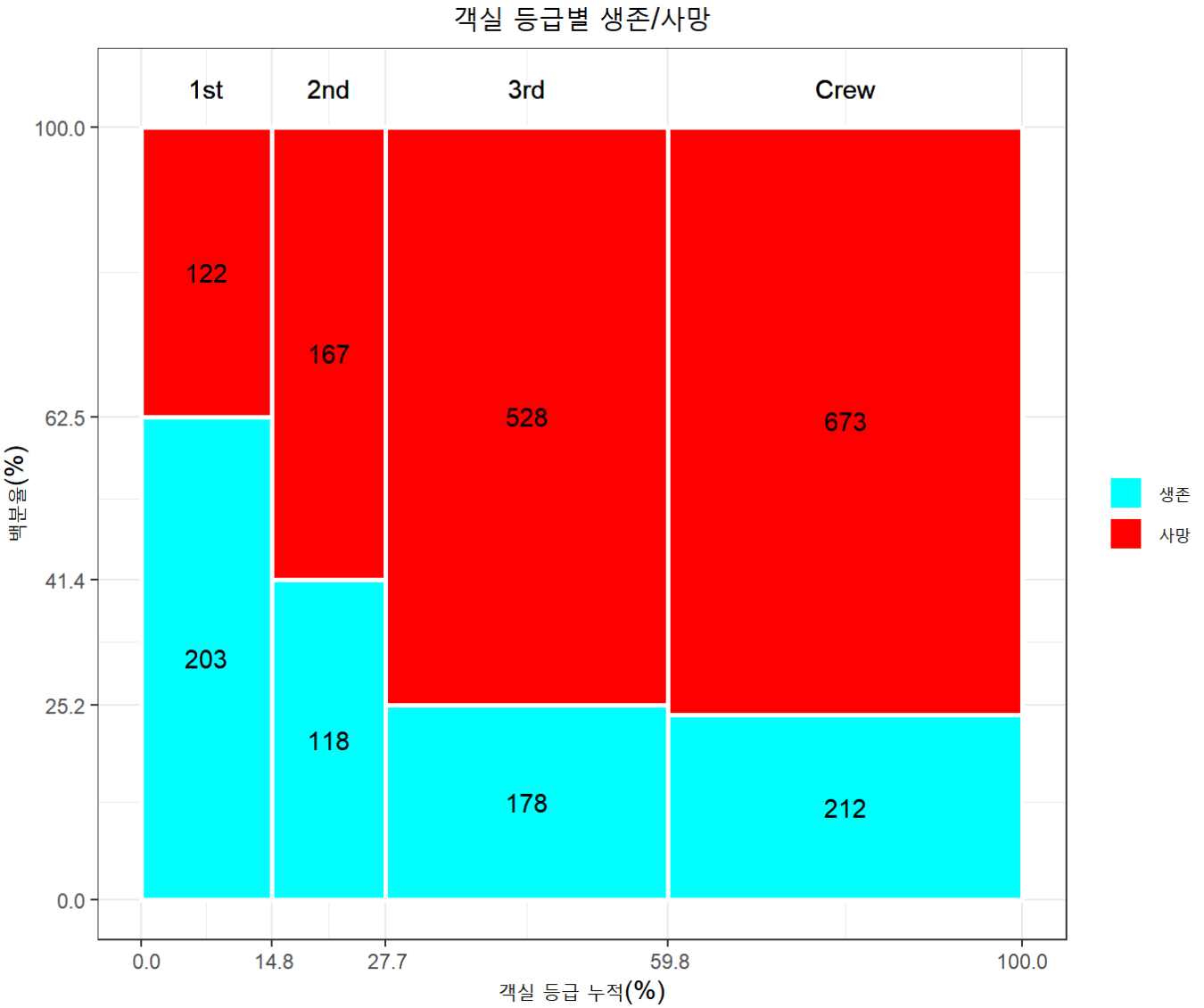
```
b1_stack <- barplot_gg_stack(as.data.frame(as.table(Surv_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "인원(명)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
b1_fill <- barplot_gg_fill(as.data.frame(as.table(Surv_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "백분율(%)" ) +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
top1 <- textGrob("객실 등급별 생존/사망",
                gp = gpar(cex = 1.5, fontfamily = ""))
grid.arrange(b1_stack, b1_fill, ncol = 2, top = top1)
```

객실 등급별 생존/사망



Mosaic Plots

```
m1_list <-
  mosaic_gg_new(as.data.frame(as.table(Surv_Class)[2:1, ]))
m1_list$m +
  theme_bw() +
  labs(x = "객실 등급 누적(%)", y = "백분율(%)") +
  ggtitle("객실 등급별 생존/사망") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""),
        plot.title = element_text(hjust = 0.5,
                                   family = ""))
```



```
ggsave("../pics/Titanic_mosaic_ggplot01.png", width = 8, height = 6, dpi = 72)
```

By Sex

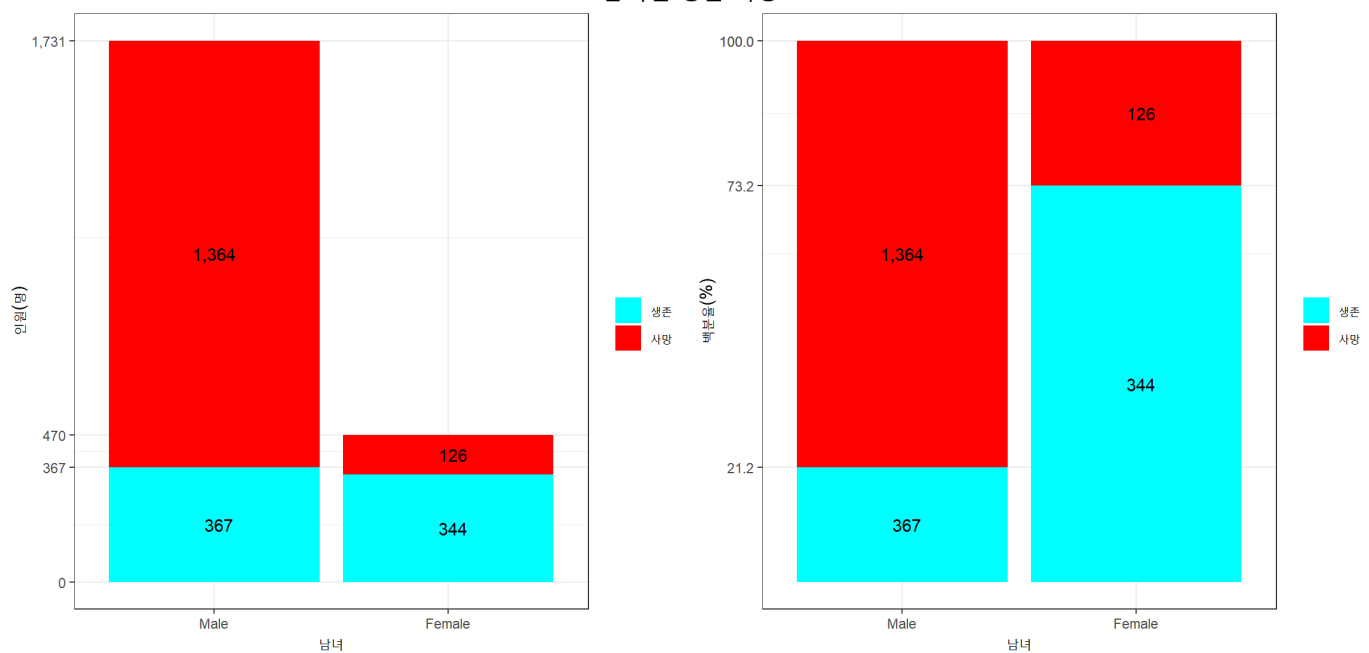
Barplots

```

b2_stack <- barplot_gg_stack(as.data.frame(as.table(Surv_Sex)[2:1, ])) +
  theme_bw() +
  labs(x = "남녀", y = "인원(명)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
b2_fill <- barplot_gg_fill(as.data.frame(as.table(Surv_Sex)[2:1, ])) +
  theme_bw() +
  labs(x = "남녀", y = "백분율(%)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
top2 <- textGrob("남녀별 생존/사망",
                gp = gpar(cex = 1.5, fontfamily = ""))
grid.arrange(b2_stack, b2_fill, ncol = 2, top = top2)

```

남녀별 생존/사망

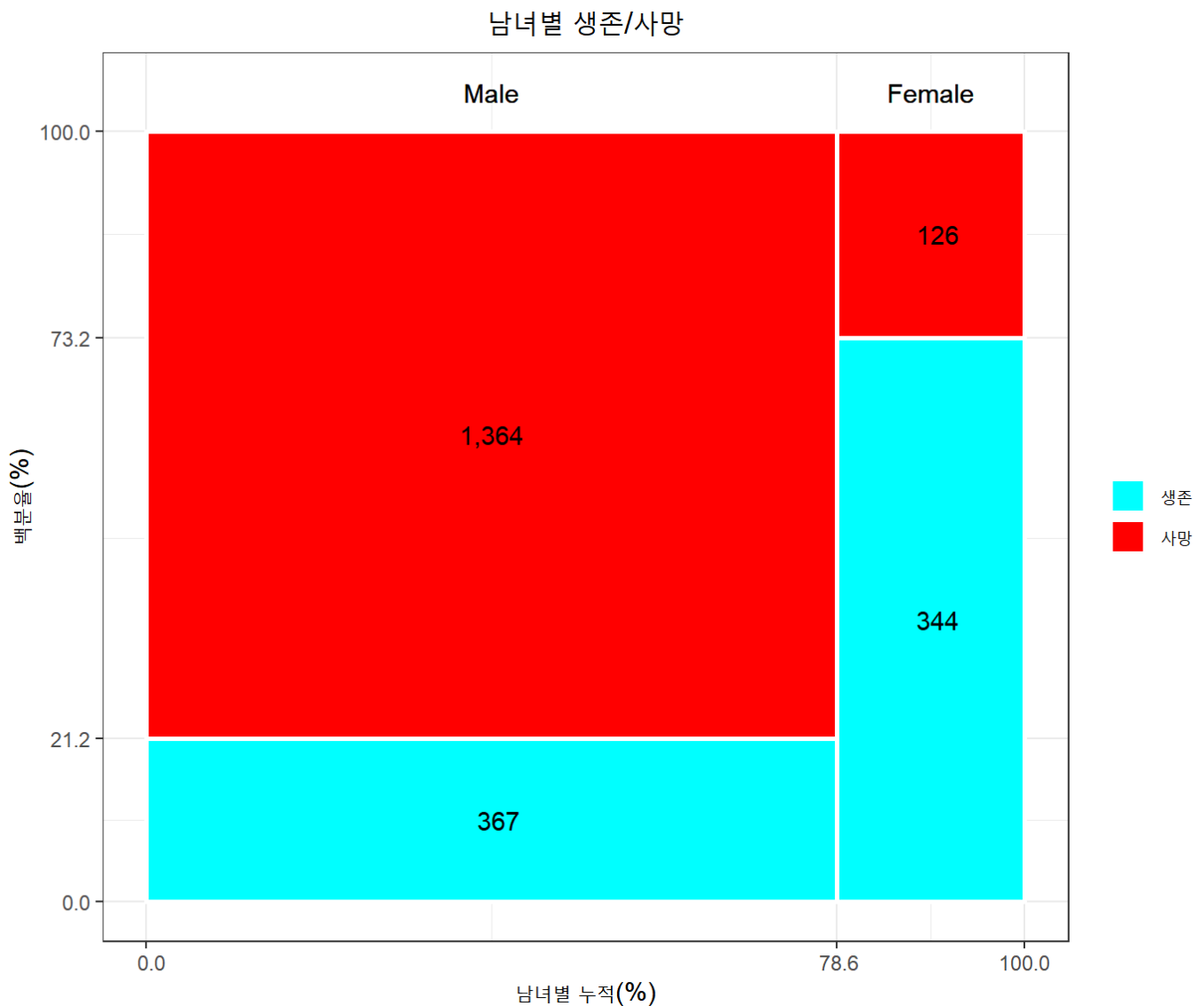


Mosaic Plots

```

m2_list <-
  mosaic_gg_new(as.data.frame(as.table(Surv_Sex)[2:1, ]))
m2_list$m +
  theme_bw() +
  labs(x = "남녀별 누적(%)", y = "백분율(%)") +
  ggtitle("남녀별 생존/사망") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""),
        plot.title = element_text(hjust = 0.5, family = ""))

```



```

ggsave("../pics/Titanic_mosaic_ggplot02.png", width = 8, height = 6, dpi = 72)

```

By Age

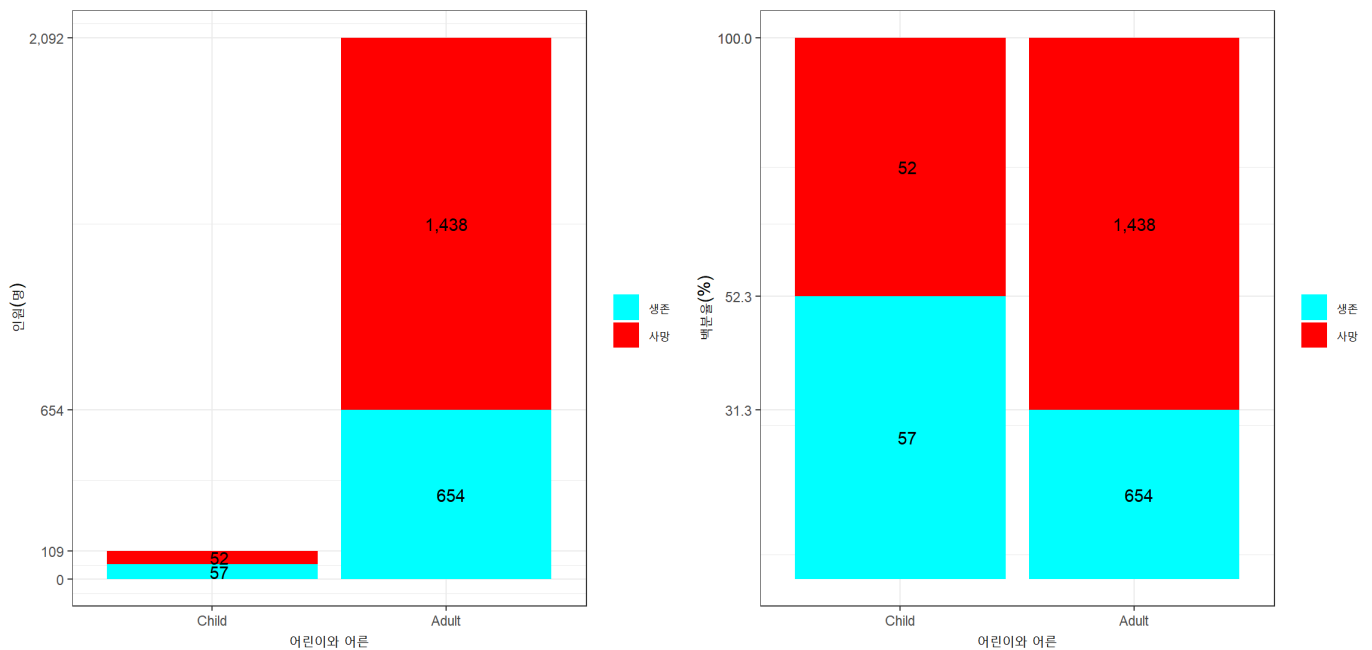
Barplots

```

b3_stack <- barplot_gg_stack(as.data.frame(as.table(Surv_Age)[2:1, ])) +
  theme_bw() +
  labs(x = "어린이와 어른", y = "인원(명)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
b3_fill <- barplot_gg_fill(as.data.frame(as.table(Surv_Age)[2:1, ])) +
  theme_bw() +
  labs(x = "어린이와 어른", y = "백분율(%)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
top3 <- textGrob("어린이와 어른의 생존/사망",
                gp = gpar(cex = 1.5, fontfamily = ""))
grid.arrange(b3_stack, b3_fill, ncol = 2, top = top3)

```

어린이와 어른의 생존/사망

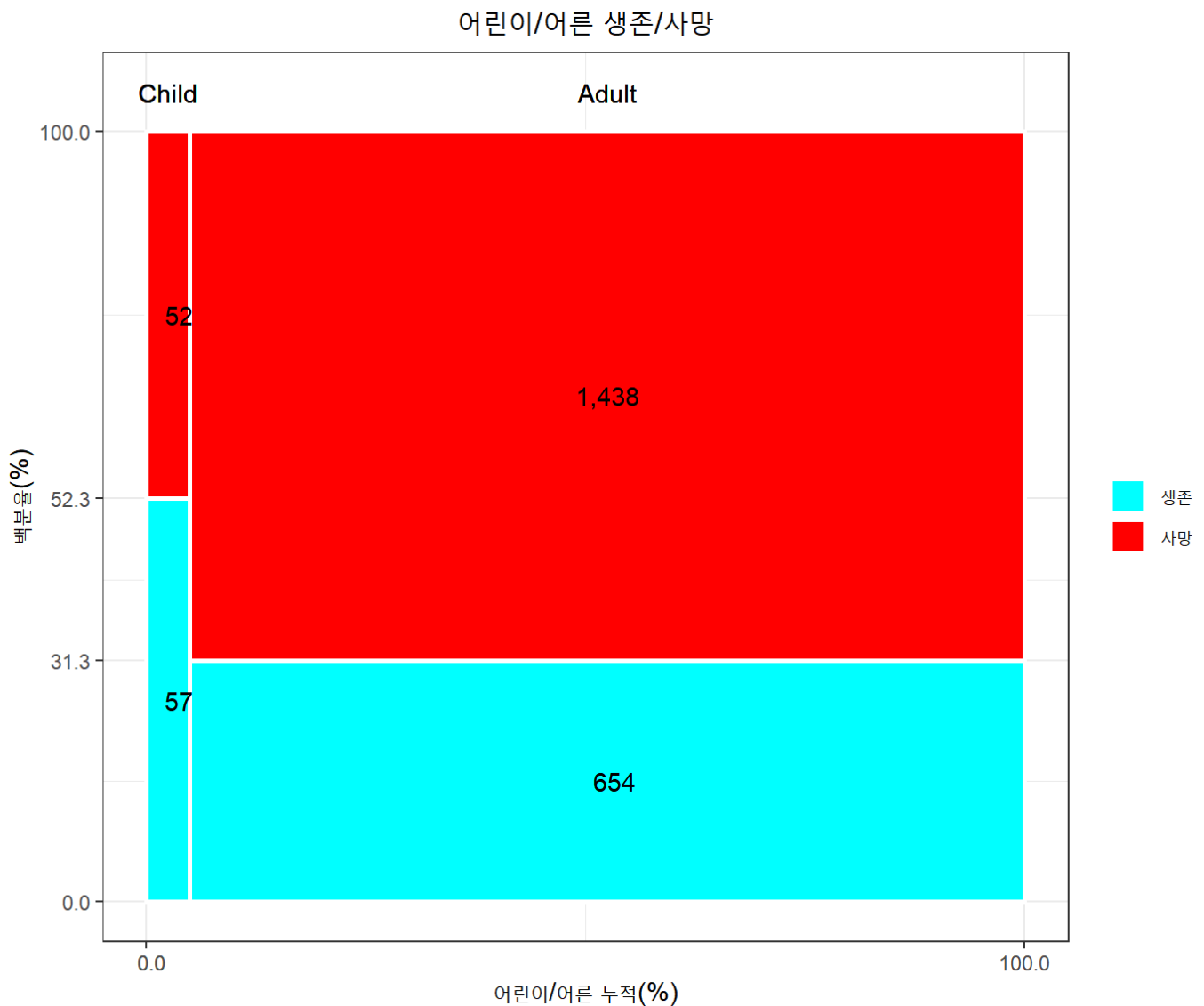


Mosaic Plots


```

m3_list <-
  mosaic_gg_new(as.data.frame(as.table(Surv_Age)[2:1, ]))
m3_list$m +
  theme_bw() +
  labs(x = "어린이/어른 누적(%)", y = "백분율(%)") +
  ggtitle("어린이/어른 생존/사망") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""),
        plot.title = element_text(hjust = 0.5, family = ""))

```



```

ggsave("../pics/Titanic_mosaic_ggplot03.png", width = 8, height = 6, dpi = 72)

```

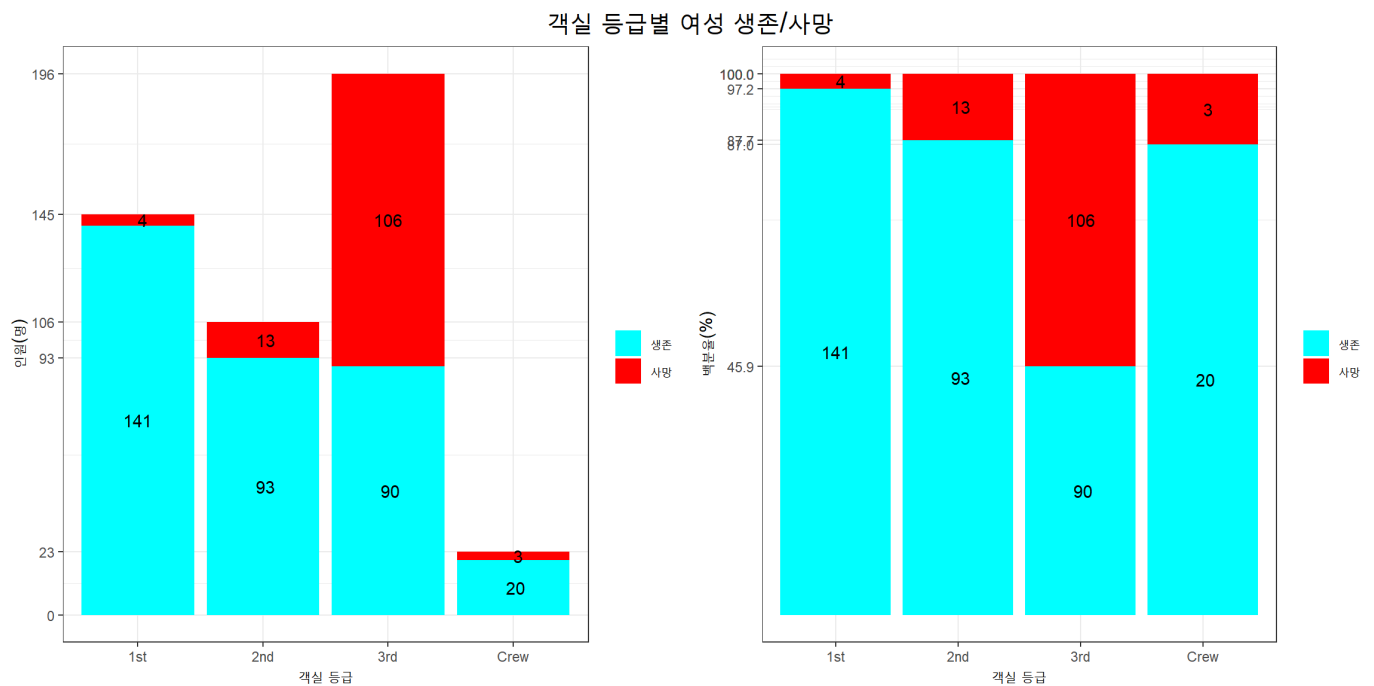
Female by Class

Barplots

```

b4_stack <- barplot_gg_stack(as.data.frame(as.table(Female_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "인원(명)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
b4_fill <- barplot_gg_fill(as.data.frame(as.table(Female_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "백분율(%)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
top4 <- textGrob("객실 등급별 여성 생존/사망",
                gp = gpar(cex = 1.5, fontfamily = ""))
grid.arrange(b4_stack, b4_fill, ncol = 2, top = top4)

```

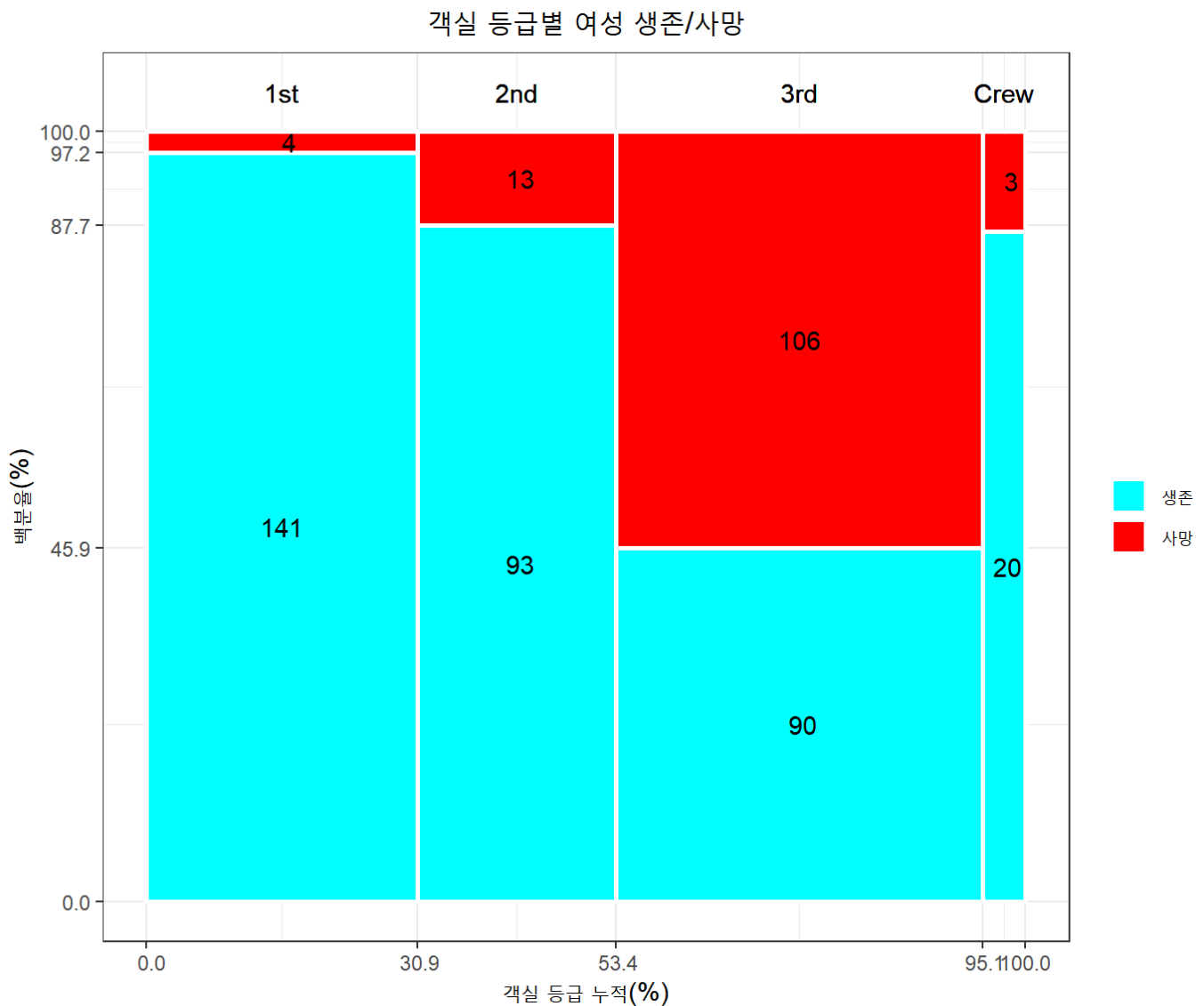


Mosaic Plots

```

m4_list <-
  mosaic_gg_new(as.data.frame(as.table(Female_Class)[2:1, ]))
m4_list$m +
  theme_bw() +
  labs(x = "객실 등급 누적(%)", y = "백분율(%)") +
  ggtitle("객실 등급별 여성 생존/사망") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""),
        plot.title = element_text(hjust = 0.5, family = ""))

```



```

ggsave("../pics/Titanic_mosaic_ggplot04.png", width = 8, height = 6, dpi = 72)

```

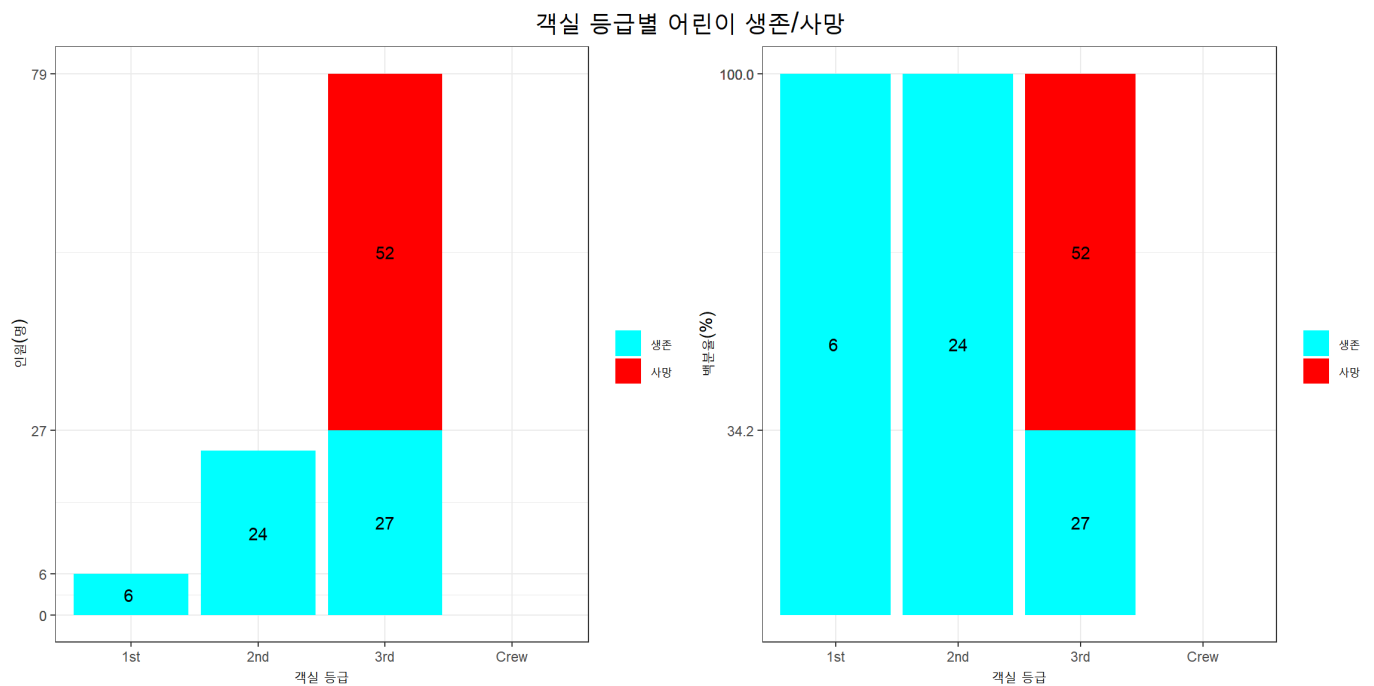
Children by Class

Barplots

```

b5_stack <- barplot_gg_stack(as.data.frame(as.table(Child_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "인원(명)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
b5_fill <- barplot_gg_fill(as.data.frame(as.table(Child_Class)[2:1, ])) +
  theme_bw() +
  labs(x = "객실 등급", y = "백분율(%)") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""))
top5 <- textGrob("객실 등급별 어린이 생존/사망",
                gp = gpar(cex = 1.5, fontfamily = ""))
grid.arrange(b5_stack, b5_fill, ncol = 2, top = top5)

```



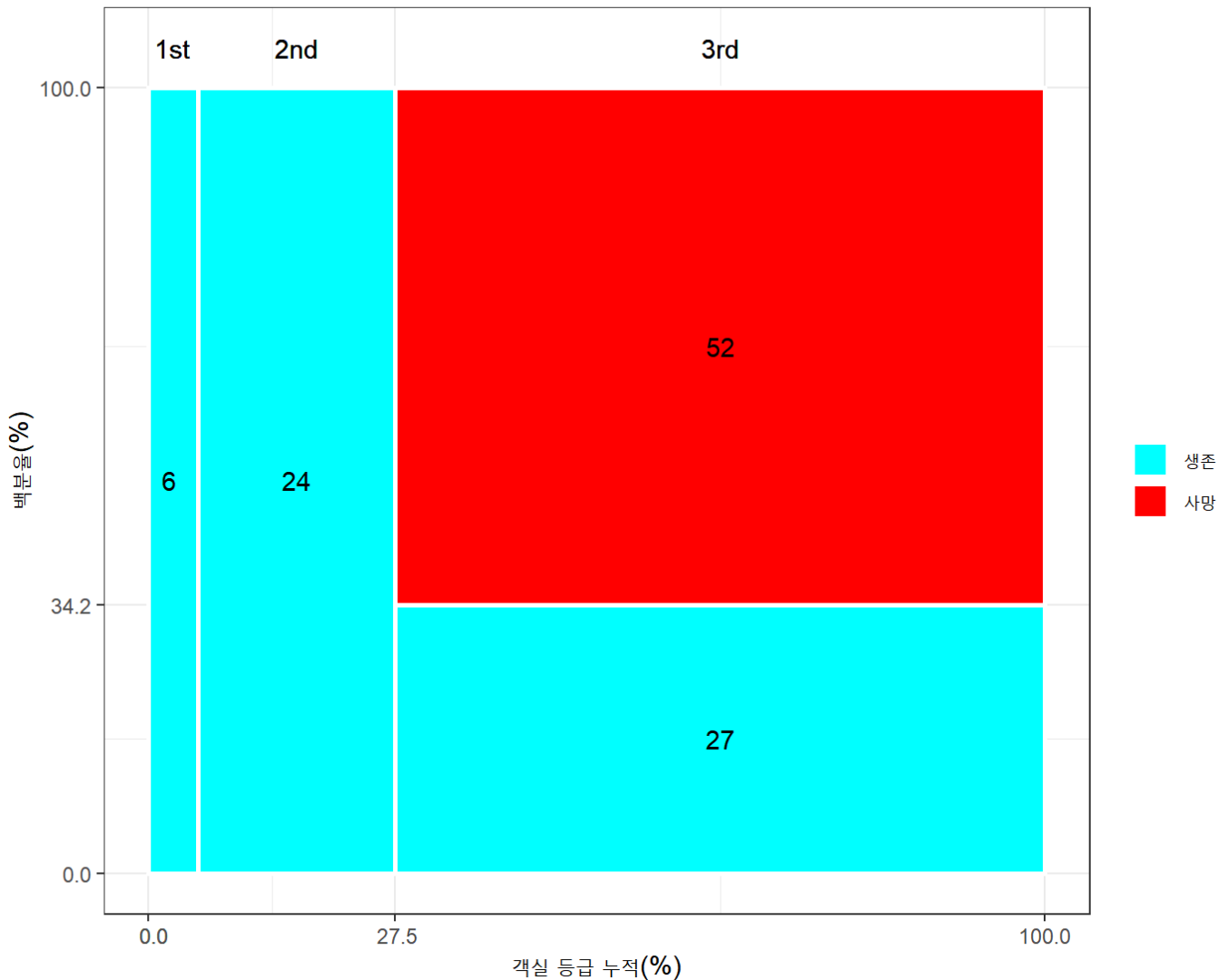
Mosaic Plots

```

m5_list <-
  mosaic_gg_new(as.data.frame(as.table(Child_Class)[2:1, -4]))
m5_list$m +
  theme_bw() +
  labs(x = "객실 등급 누적(%)", y = "백분율(%)") +
  ggtitle("객실 등급별 어린이 생존/사망") +
  scale_fill_manual(name = "",
                    labels = c("생존", "사망"),
                    values = rainbow(2)[2:1]) +
  theme(axis.title.x = element_text(family = ""),
        axis.title.y = element_text(family = ""),
        legend.text = element_text(family = ""),
        plot.title = element_text(hjust = 0.5, family = ""))

```

객실 등급별 어린이 생존/사망



```

ggsave("../pics/Titanic_mosaic_ggplot05.png", width = 8, height = 6, dpi = 72)

```

COMMENTS

grid.arrange를 사용하여 양 그래프를 비교하는 방법을 배울수 있었습니다. 객실칸과 선원칸의 사망률이 비슷하다는것을 확인할 수 있었습니다. mosaic plot 을 배우는 시간을 가질수 있었고, mosaic_gg를 통해 tidy한 데이터를 만들수 있었습니다. mosaic plot을 이용하여 그래프를 그리니까 생존/사망률을 한눈에 알아보기 쉬웠습니다.

또 `proportions`를 새로 알아감으로써 좀 더 `r`에 대해 알아갈 수 있었습니다. `mosaicplot`을 만드는 방법 또한 알 수 있었는데, 기존에 배웠던 것과 크게 다르지 않다는 것을 알 수 있었습니다. 이를 통해 여성보다 남성이 어린이보다 어른이 많이 생존했다는 것을 알 수 있었지만 등급별로 차이가 많이 난다는 것을 눈으로 확인할 수 있었습니다.