SBS 2017-03-27

Problem

SBS 뉴스에서는 다음과 같은 도표의 후보지지도 여론조사 결과를 보도.

SBS 뉴스 2017. 3월

막대의 높이에 의구심을 표한 시청자들의 항의에 직면함.

제대로 된 막대그래프를 그리면서 R Base plot과 ggplot에 대하여 학습.

Data Setup

```
library(extrafont)
```

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

```
candidates <- c("문재인", "안철수", "안희정", "이재명", "홍준표", "김진태", "심상정", "유승민")
rates <- c(35.2, 17.4, 12.0, 9.5, 7.7, 5.3, 3.4, 2.6)
party <- c("더불어민주당", "자유한국당", "국민의당", "정의당", "바른정당")
colour_party <- c("blue", "lightgrey", "darkgreen", "purple", "darkblue")
candidates_party <-  c("더불어민주당", "국민의당", "더불어민주당",  "더불어민주당", "자유한국당", "자유한국당", "정의당", "바른정당")
match(candidates_party, party)
```

```
## [1] 1 3 1 1 2 2 4 5
```

```
candidates_colour <- colour_party[match(candidates_party, party)]
```

strsplit()

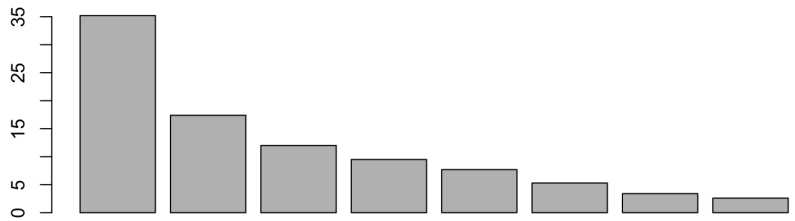
```
c_rates <- format(rates, nsmall = 1, justify = "right")
a <- sapply(strsplit(c_rates, "[.]"), `[`, 1)
b <- sapply(strsplit(c_rates, "[.]"), `[`, 2)
b_perc <- paste(".", b, "%", sep = "")
```

Colours for rates

```
col_rates <- c("red", "orange", rep("darkblue", 6))
```

Barplot (R Base)

```
barplot(rates)
```

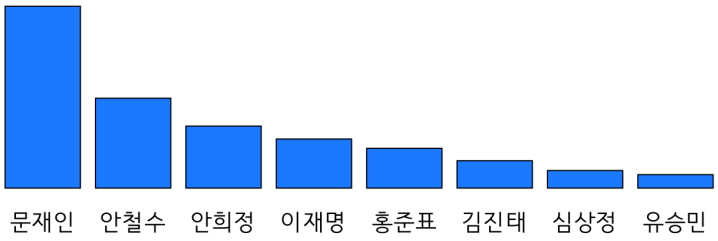


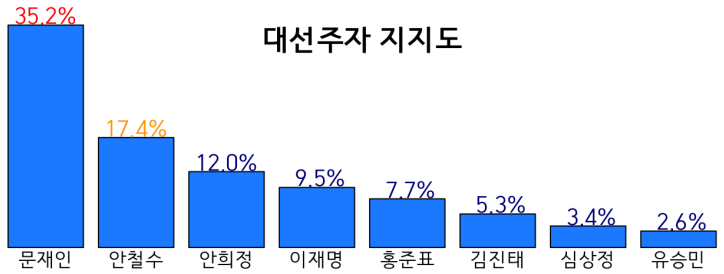
Names closer to Bars using `mtext()`

```
par(family = "HCR Dotum LVT")
b1 <- barplot(rates,
              axes = FALSE,
              col = "dodgerblue",
              names.arg = NULL,
              cex.names = 1.2,
              ylim = c(0, max(rates) * 1.1))
mtext(side = 1, at = b1, line = 0, text = candidates)
#> Rates written on top of the Bars with different Colours
text(x = b1, y = rates + rep(1.5, 8),
    labels = paste(c_rates, "%", sep = ""),
    col = col_rates,
    cex = 1.2)
main_title <- "대선주자 지지도"
note_text <- "조사기관:리얼미터, 총응답자:전국 성인 1,525명, 응답률:9.5%, 표준오차:95%신뢰수준 2.5%,
조사방법:유선 ARS 10%, 무선 ARS 71%, 무선전화면접 19%, 조사기간:2017년 3월27일(월)~29일(수)"
#> Main title inside the plot region
title(main = main_title, cex.main = 1.5, line = -2)
```

Place Names under Bars

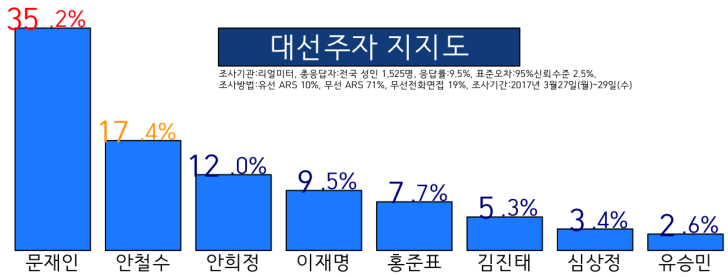
```
par(family = "HCR Dotum LVT")
b1 <- barplot(rates,
              axes = FALSE,
              col = "dodgerblue",
              names.arg = candidates,
              cex.names = 1.2,
              ylim = c(0, max(rates) * 1.1))
```





Rates with different font size for digits

```
par(family = "HCR Dotum LVT")
b1 <- barplot(rates,
              axes = FALSE,
              col = "dodgerblue",
              names.arg = NULL,
              cex.names = 1.2,
              ylim = c(0, max(rates) * 1.1))
mtext(side = 1, at = b1, line = 0, text = candidates)
#> text for integer part
text(x = b1 - c(rep(0.4, 3), rep(0.3, 5)), y = rates + rep(1.5, 8),
     labels = a,
     col = col_rates,
     cex = 1.6)
#> text for digits
text(x = b1 + 0.2, y = rates + rep(1.5, 8),
     labels = b_perc,
     col = col_rates,
     cex = 1.2)
#> Rectagle for main title
rect(xleft = mean(b1) - 2, ybottom = max(rates) - 6, xright = mean(b1) + 2,
     ytop = max(rates), col = "dodgerblue4")
#> Main title
text(x = mean(b1), y = max(rates) - 3, labels = main_title, col = "white", cex = 1.5)
#> Text for notes
text(x = mean(b1) - 2, y = max(rates) - 8, labels = note_text, cex = 0.5, adj = 0)
```



ggplot

Data for ggplot

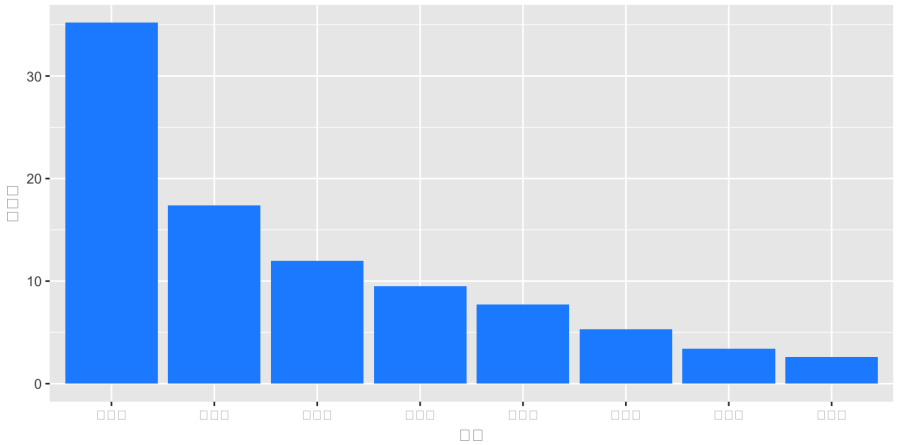
```
library(ggplot2)
candidates_f <- factor(candidates, levels = candidates)
rates_df <- data.frame(후보 = candidates_f,
                       정당 = candidates_party,
                       색깔 = candidates_colour,
                       지지도 = rates)
```

data and mapping

```
g0 <- ggplot(data = rates_df,
             mapping = aes(x = 후보, y = 지지도))
```

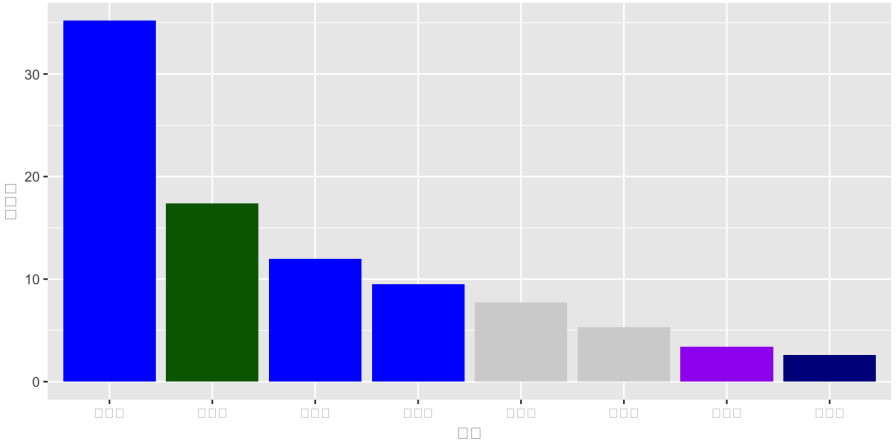
`geom_bar()` with single colour for the bars

```
(g1 <- g0 +
  geom_bar(stat = "identity", fill = "dodgerblue"))
```



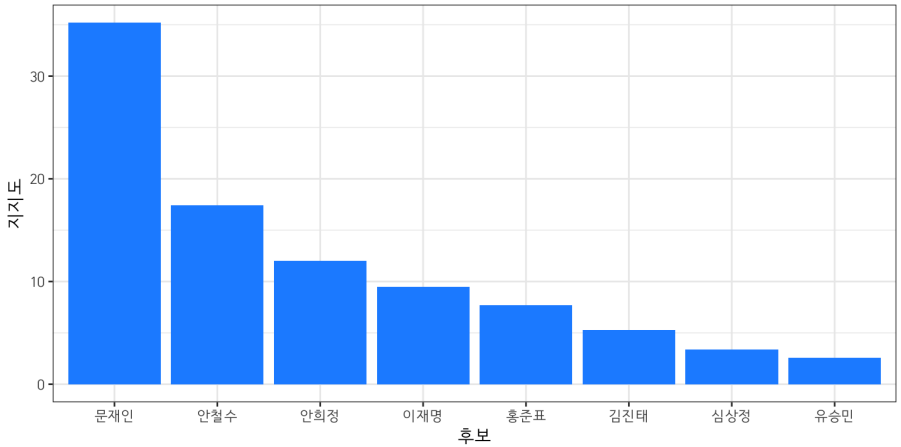
Fill the bars with party colours

```
(g1.1 <- g0 +
  geom_bar(stat = "identity",
           fill = candidates_colour))
```



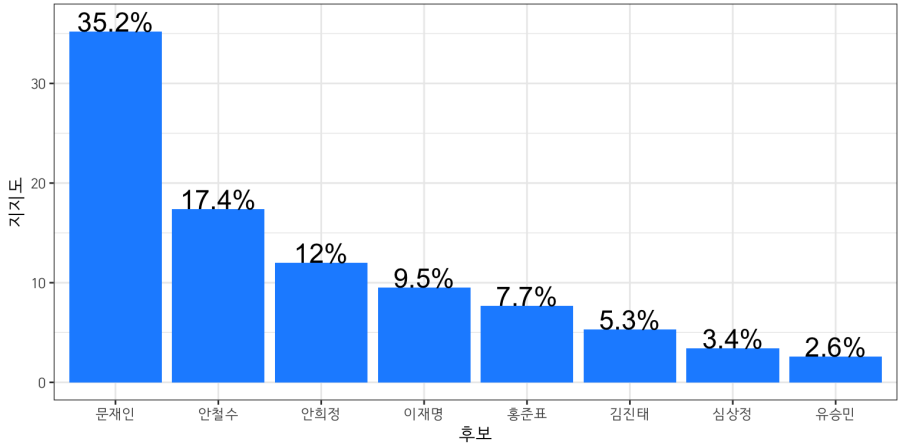
Font family setting with `g1`

```
(g2 <- g1 +
  theme_bw(base_family = "HCR Dotum LVT"))
```



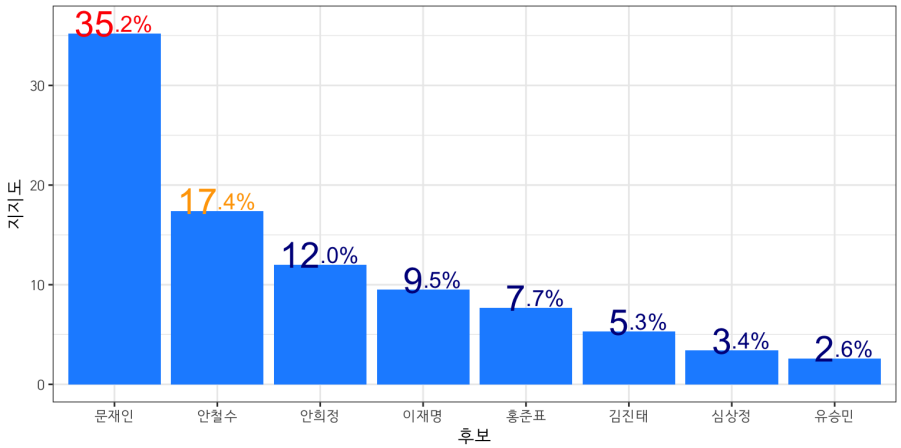
Add rates on top of the bars

```
(g3.0 <- g2 +
  geom_text(mapping = aes(x = 후보,
                          y = 지지도 + rep(1, 8),
                          label = paste(지지도, "%", sep = "")),
                          size = 6))
```



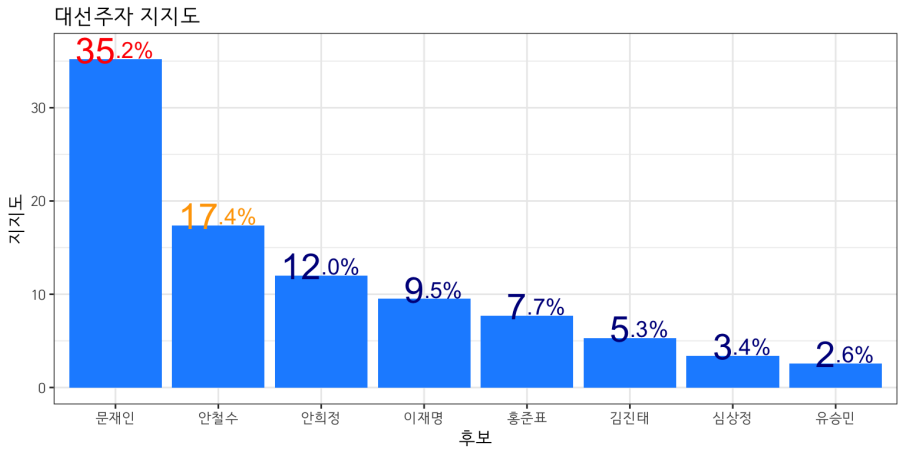
Different font size for digits

```
(g3 <- g2 +
    geom_text(mapping = aes(x = 후보,
                            y = 지지도 + rep(1, 8),
                            label = a),
              hjust = 1,
              size = 8,
              colour = col_rates) +
    geom_text(mapping = aes(x = 후보,
                            y = 지지도 + rep(1, 8),
                            label = b_perc),
              hjust = 0,
              size = 5,
              colour = col_rates))
```



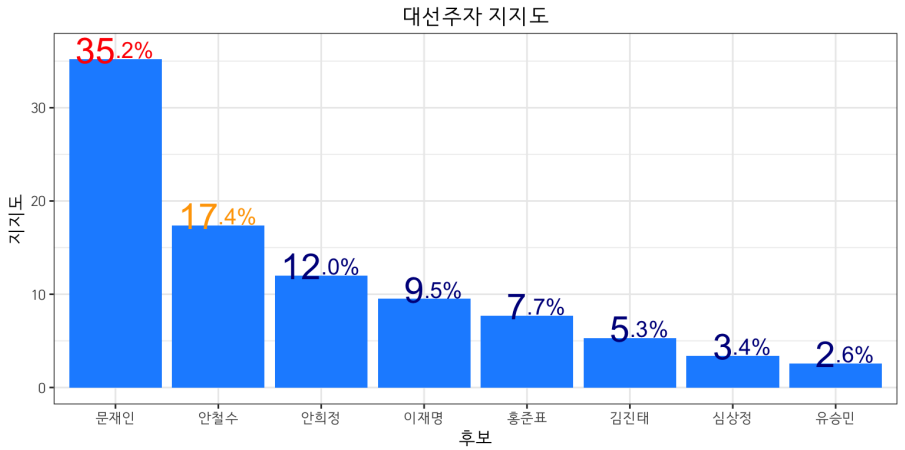
Main title (left-justified)

```
(g4 <- g3 +
  labs(title = main_title))
```



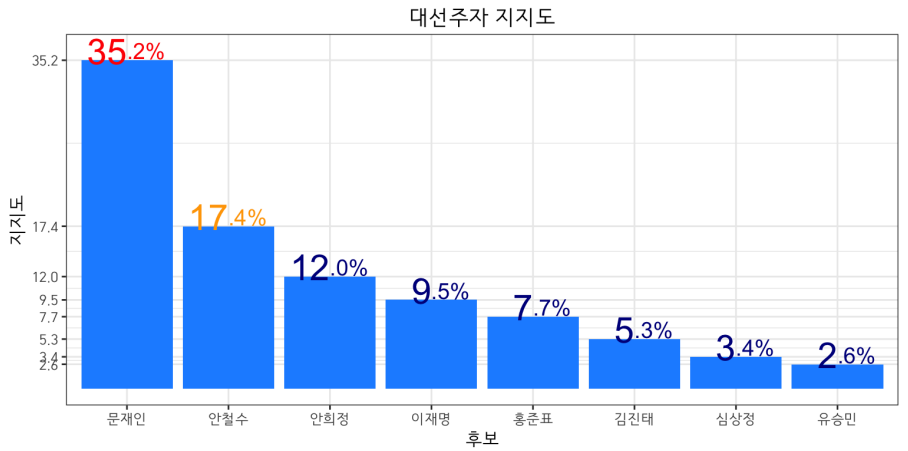
Main title at the center

```
(g5 <- g4 +
  theme(plot.title = element_text(hjust = 0.5)))
```



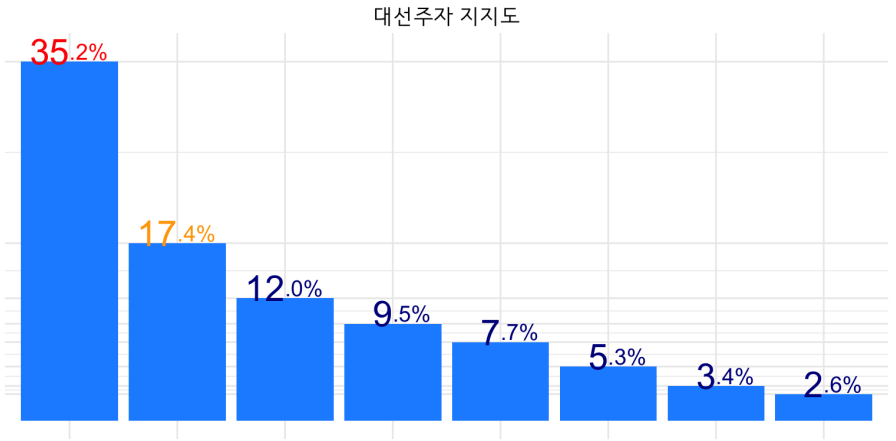
Rates at y-axis

```
(g6 <- g5 +
  scale_y_continuous(breaks = rates,
                     labels = c_rates))
```



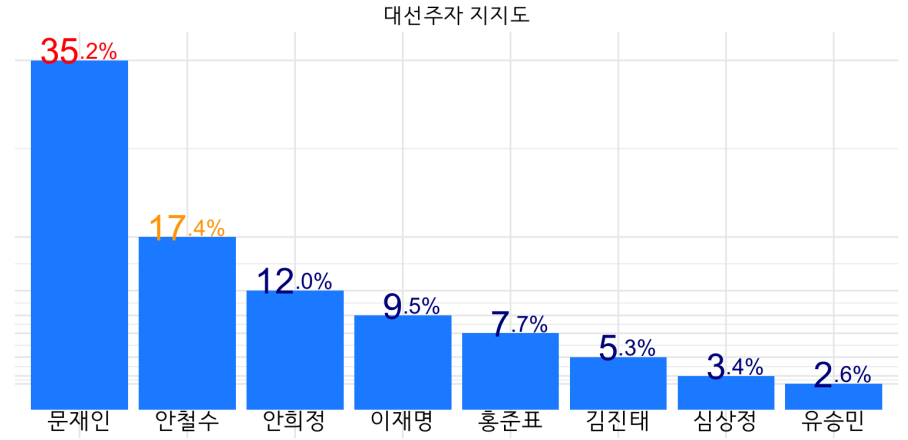
Clear axes

```
(g7 <- g6 +
  theme(panel.border = element_blank(),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks = element_blank(),
        axis.text.y = element_blank()))
```



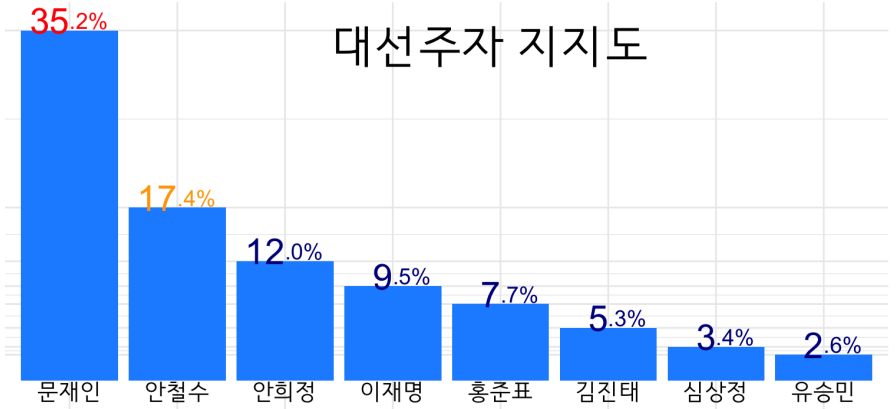
Names closer to the bars

```
(g8 <- g7 +
    geom_text(mapping = aes(x = 후보,
                            y = -1,
                            label = 후보),
              size = 5,
              family = "HCR Dotum LVT"))
```



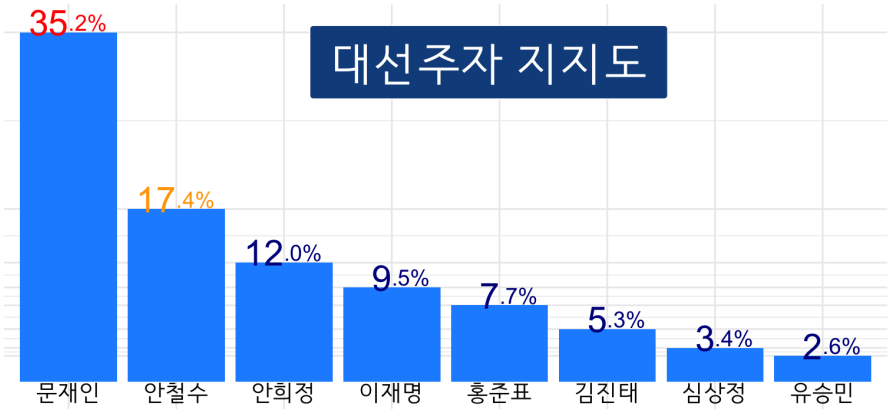
Main title inside the plot region

```
(g9 <- g8 +
    ggtitle("") +
    annotate("text",
             x = mean(b1),
             y = max(rates) - 3,
             label = main_title,
             vjust = 0,
             size = 10,
             family = "HCR Dotum LVT"))
```



`geom_label()` to enclose the main title in bounding box

```
(g9.1 <- g8 +
    ggtitle("") +
    geom_label(mapping = aes(x = mean(b1),
                             y = max(rates) - 3,
                             label = main_title),
               label.padding = unit(0.9, "lines"),
               size = 10,
               fill = "dodgerblue4",
               colour = "white",
               family = "HCR Dotum LVT"))
```



Notes

```
(g9.2 <- g9.1 +
    annotate("text",
             x = mean(b1) - 1, y = max(rates) - 8,
             label = note_text,
             size = 2,
             hjust = 0,
             family = "HCR Dotum LVT"))
```