

# US Top Income Share 2015

coop711

2017-05-06

## Load Data

### Plots

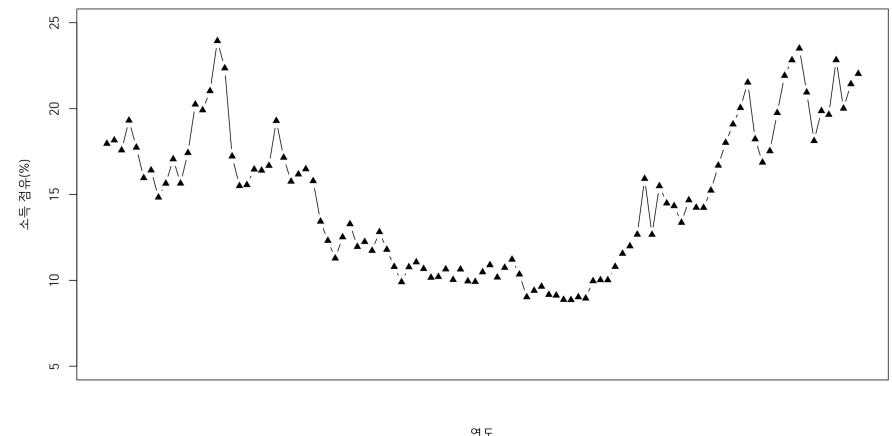
이 중에서 소득 상위 10%( P90\_100 )를 상위 1%( P99\_100 ), 차상위 4%( P95\_99 ), 차차상위 5%( P90\_95 )로 나누어 그 몫의 변화를 살펴보자. 우선 멋있하게 상위 1%의 소득 점유율 변화만 그려보면,

```
plot(P99_100 ~ Year, data = US.top.income.shares.15)
```



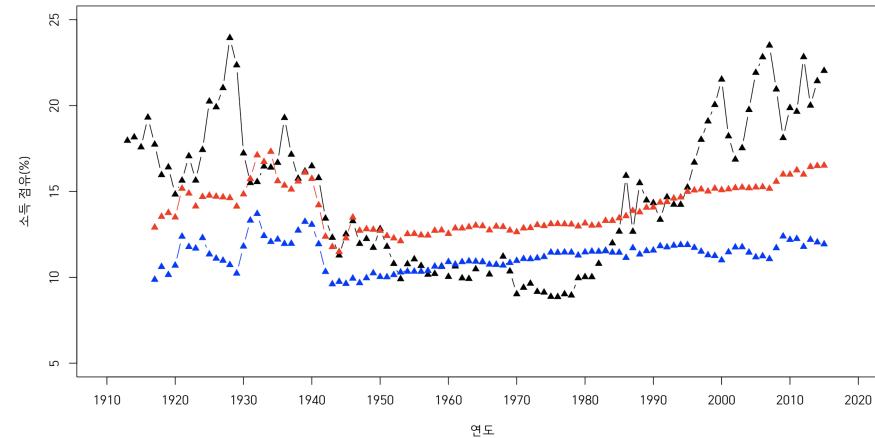
최소한의 정보를 주기 위하여 각 자료의 최대값, 최소값을 고려하여 y-축의 범위를 `ylim = c(5, 25)` 로 설정하고, x-축의 눈금은 비워둔다. 점들을 선으로 이어주고, 점은 다시 삼각형 모양으로 바꾼다.

```
par(family = "HCR Dotum LVT")
x.lab <- "연도"
y.lab <- "소득 점유(%)"
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, ylim = c(5, 25), xaxt = "n", type = "b", pch = 17)
```



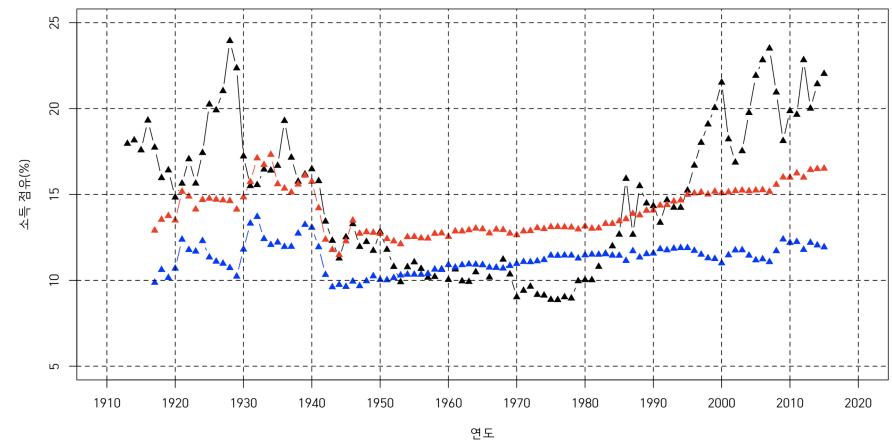
이제 x-축에는 연도를 10년 단위로 표시하고, `lines()`를 이용하여 차상위4%와 차차상위5%의 소득 점유율 변화를 함께 그린다. `xlim = c(1910, 2020)`을 설정하지 않으면 땐 그림이 그려지는지 살펴본다.

```
par(family = "HCR Dotum LVT")
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, xlim
= c(1910, 2020), ylim = c(5, 25), xaxt = "n", type = "b", pch = 17)
axis(side = 1, at = seq(1910, 2020, by = 10), labels = seq(1910, 2020, by = 10))
lines(P95_99 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "red")
lines(P90_95 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "blue")
```



도표의 가독성을 높이기 위하여 x-축과 y-축을 가로지르는 격자를 설치한다.

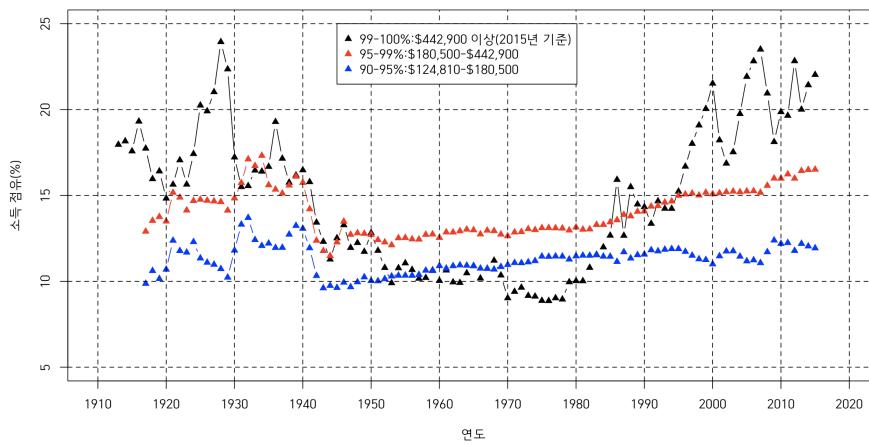
```
par(family = "HCR Dotum LVT")
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, xlim
= c(1910, 2020), ylim = c(5, 25), xaxt = "n", type = "b", pch = 17)
axis(side = 1, at = seq(1910, 2020, by = 10), labels = seq(1910, 2020, by = 10))
lines(P95_99 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "red")
lines(P90_95 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "blue")
abline(h = seq(5, 25, by = 5), lty = 2)
abline(v = seq(1910, 2020, by = 10), lty = 2)
```



TabFig2015prel.xls 의 Table 0 시트에 나와 있는 정보를 활용하여 범례를 만든다.

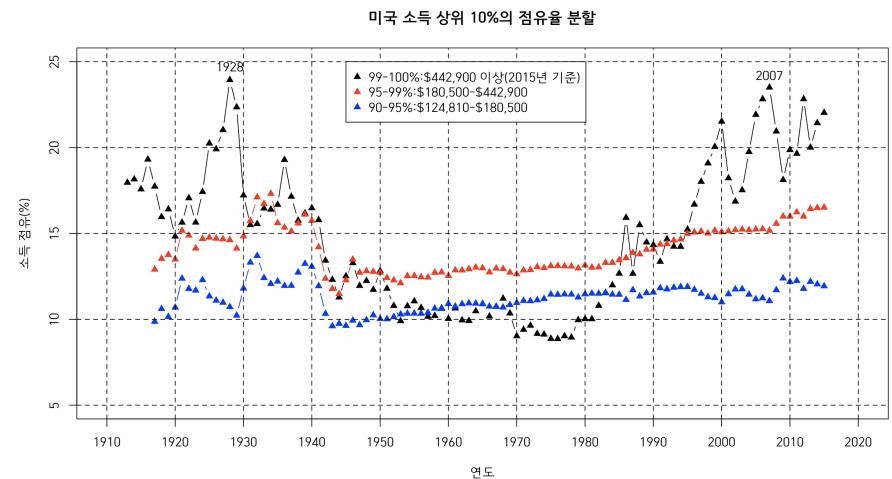
```
par(family = "HCR Dotum LVT")
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, xlim
= c(1910, 2020), ylim = c(5,25), xaxt = "n", type = "b", pch = 17)
axis(side = 1, at = seq(1910, 2020, by = 10), labels = seq(1910, 2020, by = 10))
lines(P95_99 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "re
d")
lines(P90_95 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "blu
e")
abline(h = seq(5, 25, by = 5), lty = 2)
abline(v = seq(1910, 2020, by = 10), lty = 2)
legend.text <- c("99-100%:$442,900 이상(2015년 기준)", "95-99%:$180,500-$442,900", "90-9
5%:$124,810-$180,500")
legend(x = 1945, y = 25, legend = legend.text, pch = 17, col = c("black", "red", "blu
e"))

```



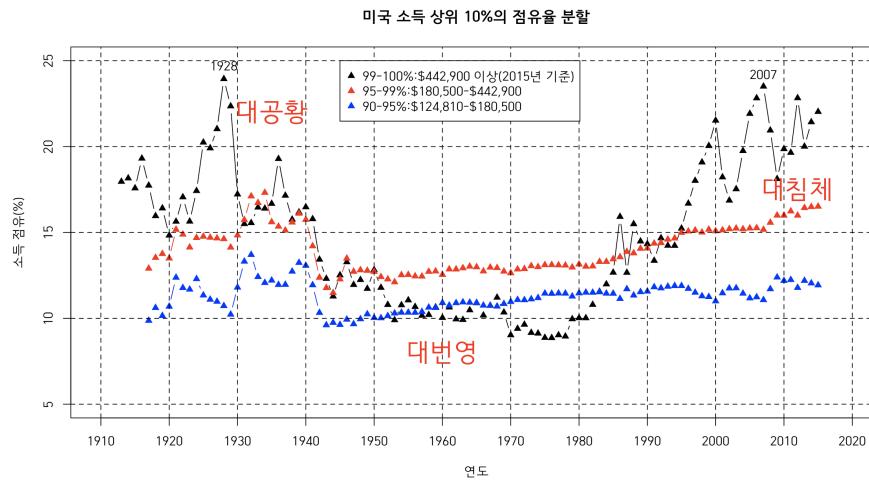
메인 타이틀을 입력하고, 상위1%의 소득점유율이 최고에 달했던 연도를 표시한다.

```
par(family = "HCR Dotum LVT")
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, xlim
= c(1910, 2020), ylim = c(5,25), xaxt = "n", type = "b", pch = 17)
axis(side = 1, at = seq(1910, 2020, by = 10), labels = seq(1910, 2020, by = 10))
lines(P95_99 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "re
d")
lines(P90_95 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "blu
e")
abline(h = seq(5, 25, by = 5), lty = 2)
abline(v = seq(1910, 2020, by = 10), lty = 2)
legend(x = 1945, y = 25, legend = legend.text, pch = 17, col = c("black", "red", "blu
e"))
main.title <- "미국 소득 상위 10%의 점유율 분할"
title(main = main.title)
text(x = c(1928, 2007), y = c(24, 23.5), labels = c("1928", "2007"), pos = 3)
```



역사적으로 각 시기를 어떻게 부르고 있는지 텍스트를 추가한다.

```
par(family = "HCR Dotum LVT")
plot(P99_100 ~ Year, data = US.top.income.shares.15, xlab = x.lab, ylab = y.lab, xlim
= c(1910, 2020), ylim = c(5,25), xaxt = "n", type = "b", pch = 17)
axis(side = 1, at = seq(1910, 2020, by = 10), labels = seq(1910, 2020, by = 10))
lines(P95_99 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "red")
lines(P90_95 ~ Year, data = US.top.income.shares.15, type = "b", pch = 17, col = "blue")
abline(h = seq(5, 25, by = 5), lty = 2)
abline(v = seq(1910, 2020, by = 10), lty = 2)
legend(x = 1945, y = 25, legend = legend.text, pch = 17, col = c("black", "red", "blue"))
title(main = main.title)
text(x = c(1928, 2007), y = c(24, 23.5), labels = c("1928", "2007"), pos = 3)
times.label <- c("대공황", "대번영", "대침체")
text(x = c(1935, 1960, 2012), y = c(22, 8, 17.5), label = times.label, cex = 2.0, col
= "red")
```



## ggplot

### Data Reshaping

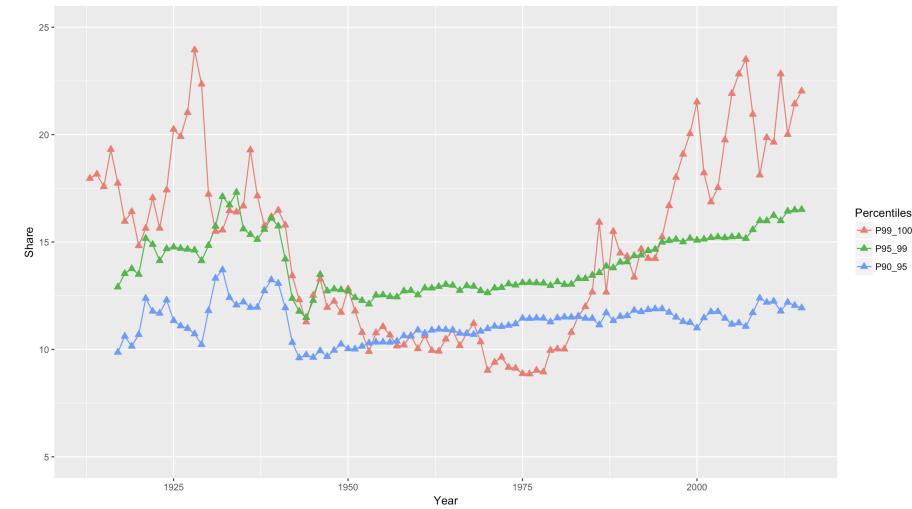
- reshape2 패키지를 이용하여 wide format 을 long format 으로

```
library(reshape2)
data.1_10 <- US.top.income.shares.15[c("Year", "P99_100", "P95_99", "P90_95")]
data.1_10.melt <- melt(data.1_10, id.vars = "Year", measure.vars = c("P99_100", "P95_99", "P90_95"), variable.name = "Percentiles", value.name = "Share")
str(data.1_10.melt)
```

```
## 'data.frame': 309 obs. of 3 variables:
## $ Year      : int 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 ...
## $ Percentiles: Factor w/ 3 levels "P99_100","P95_99",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ Share      : num 18 18.2 17.6 19.3 17.7 ...
```

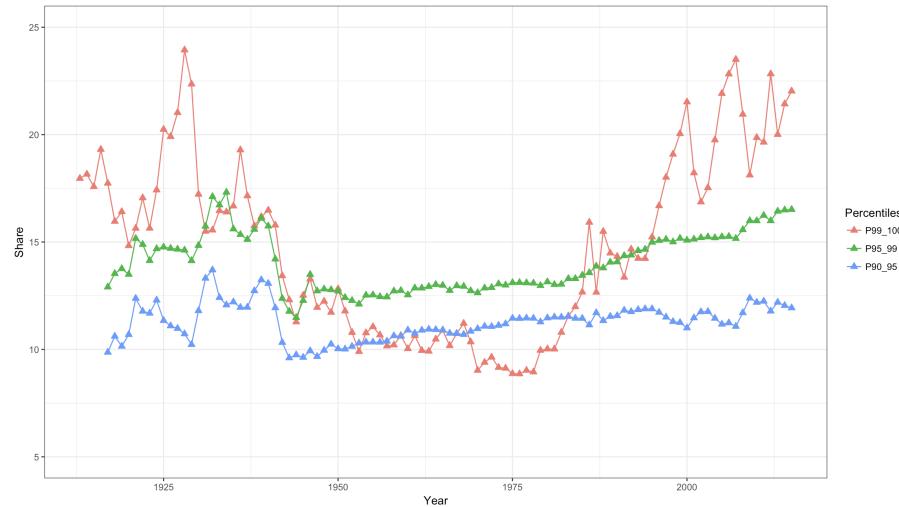
골격 그리기

```
library(ggplot2)
(g0 <- ggplot(data.1_10.melt, aes(x = Year, y = Share, colour = Percentiles)) +
  geom_line(na.rm = TRUE) +
  geom_point(shape = 24, aes(fill = Percentiles), size = 2, na.rm = TRUE) +
  ylim(5, 25))
```



theme\_bw() 적용

```
(g1 <- g0 +  
  theme_bw())
```



격자 설정

```
(g2 <- g1 +  
  theme(panel.grid.major = element_line(linetype = "dashed", colour = "black")))
```



```
##(g1 <- g0 + theme_classic())  
##(g1 <- g0 + theme_minimal())  
##(g1 <- g0 + theme_grey())
```

x-축 눈금 위치 설정

```
(g3 <- g2 +
  scale_x_continuous(breaks = seq(1910, 2020, by = 10)))
```



## 한글 정보 입력

한글 테마 sourcing

```
source("./theme_kr.R")
ls()
```

```
## [1] "data.1_10"          "data.1_10.melt"
## [3] "g0"                  "g1"
## [5] "g2"                  "g3"
## [7] "legend.text"         "main.title"
## [9] "theme.kr"            "times.label"
## [11] "US.top.income.shares.15" "v.names"
## [13] "x.lab"               "y.lab"
```

한글 테마 적용, x-축과 y-축의 라벨 수정

```
(g4 <- g3 +
  theme.kr +
  xlab(x.lab) +
  ylab(y.lab))
```



전체 제목 추가

```
(g5 <- g4 +
  ggtitle(main.title) +
  theme(plot.title = element_text(size = 20)))
```

미국 소득 상위 10%의 점유율 분할



범례 제목 수정

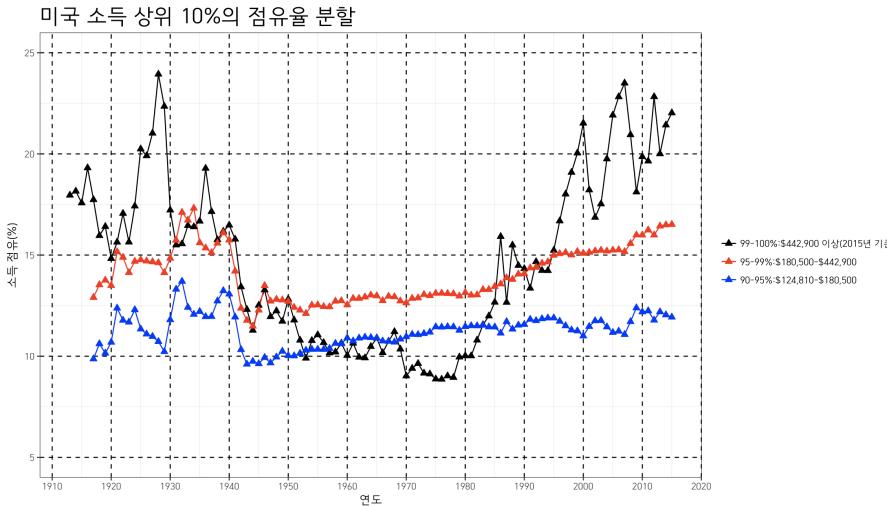
```
(g6 <- g5 +
  labs(colour = "소득 분위", fill = "소득 분위"))
```

미국 소득 상위 10%의 점유율 분할



범례와 색깔 수정, 범례 제목 없애기

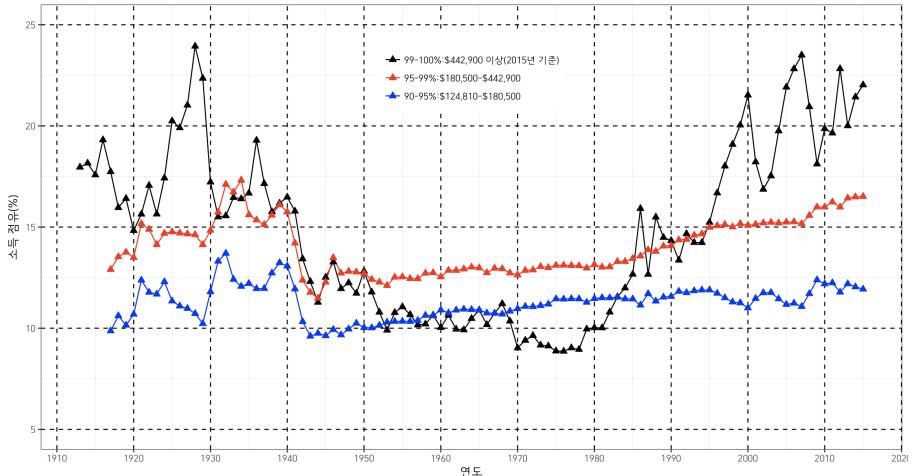
```
(g7 <- g6 +
  scale_colour_manual(name = "", values = c("black", "red", "blue"), labels = legend.t
d.text) +
  scale_fill_manual(name = "", values = c("black", "red", "blue"), labels = legend.t
ext))
```



범례를 안쪽으로

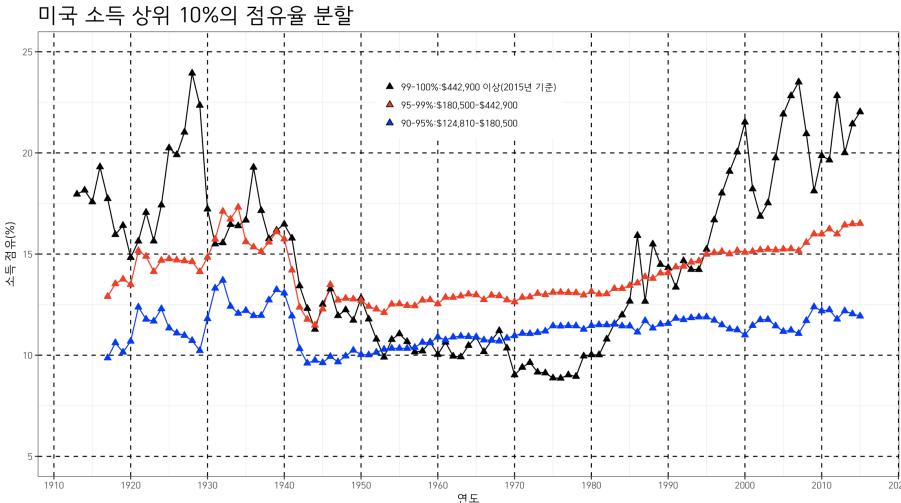
```
(g8 <- g7 +
  theme(legend.position = c(0.5, 0.85)))
```

미국 소득 상위 10%의 점유율 분할



colour 범례 없애기(colour 설정으로 나타나는 범례의 선 없애는 효과)

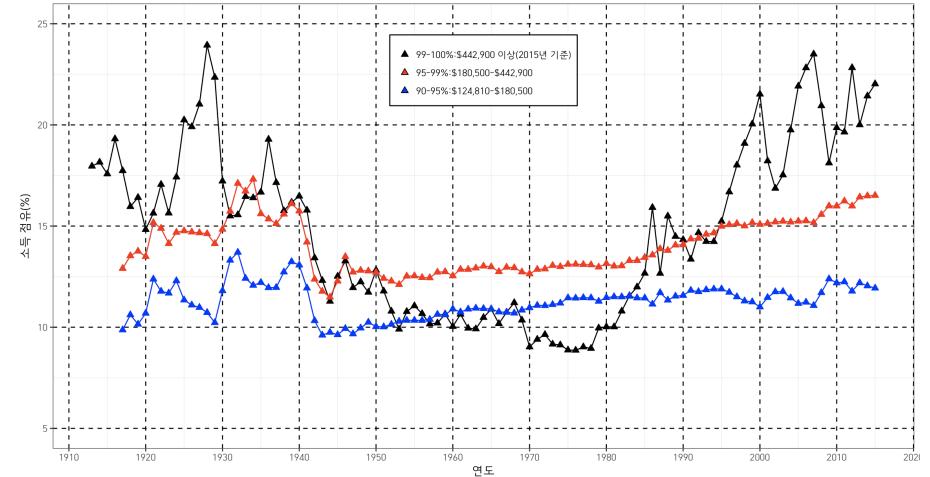
```
#(g9 <- g8 + guides(colour = guide_legend(title=NULL), fill = guide_legend(title=NULL))
L)))
(g9 <- g8 +
  guides(colour ="none"))
```



범례에 박스 두르고, 빈 제목 자리 없애기

```
(g10 <- g9 +
  theme(legend.title = element_blank(), legend.background = element_rect(fill = "white", colour = "black"))))
```

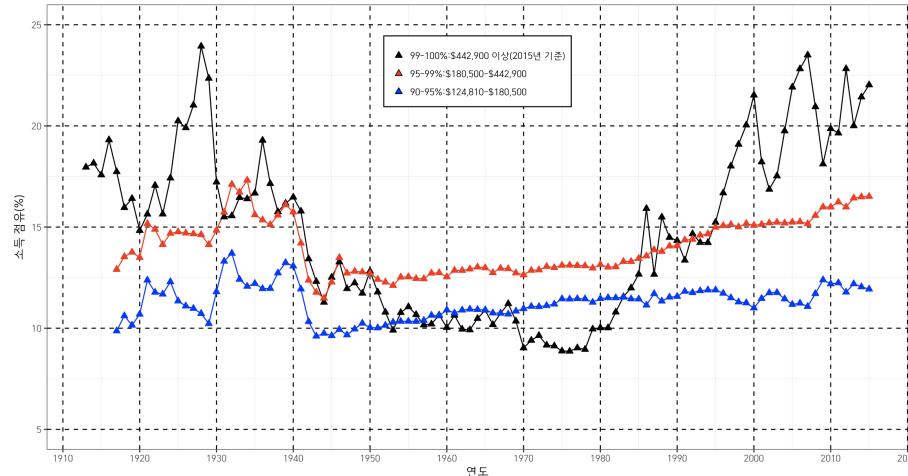
미국 소득 상위 10%의 점유율 분할



범례 항목 박스 없애기

```
(g11 <- g10 +  
  theme(legend.key = element_blank()))
```

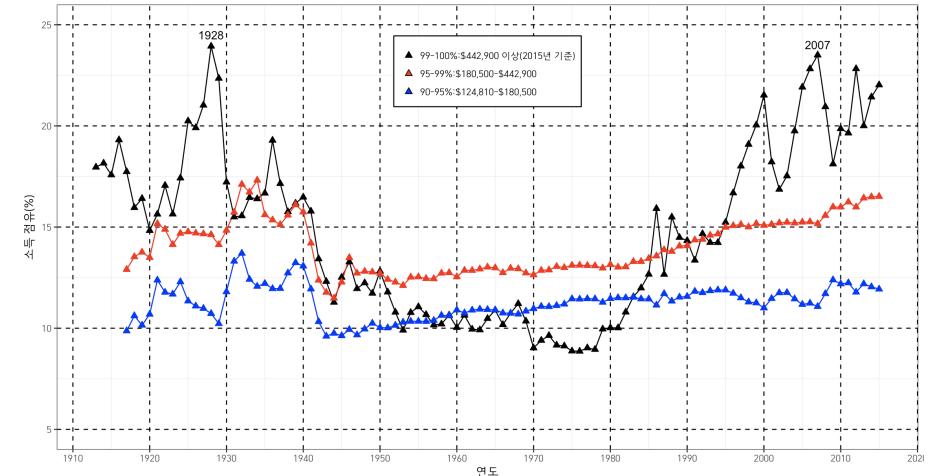
미국 소득 상위 10%의 점유율 분할



역사상 고점 시기

```
(g12 <- g11 +  
  annotate("text", x = c(1928, 2007), y = c(24.5, 24), label = c(1928, 2007)))
```

미국 소득 상위 10%의 점유율 분할



시대적 특징 텍스트 입력

```
(g13 <- g12 +
  annotate("text", x = c(1935, 1960, 2015), y = c(22, 8, 18), label = times.label, co
lour = "red", family = "HCR Dotum LVT", size = 8))
```

