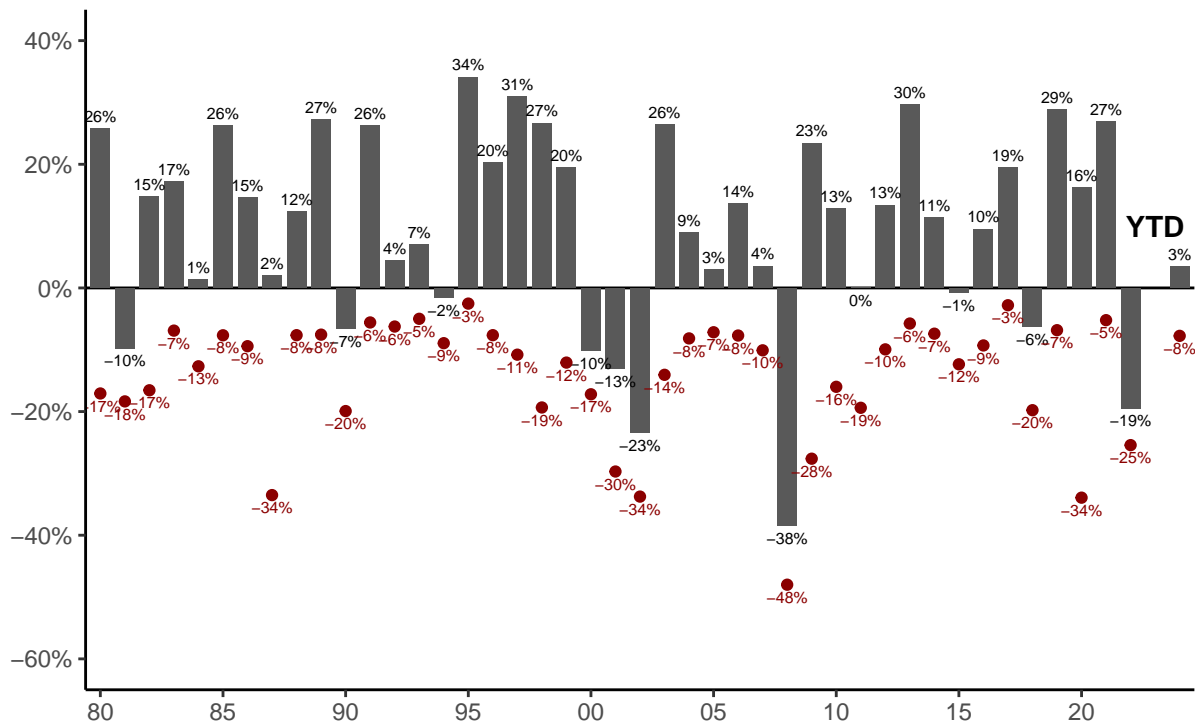# S&P 500 annual returns vs. largest intra–year declines

Data source: Yahoo Finance

Commentary

S&P 500 annual returns and the largest intra-year declines comparison provide us with couple of insights, mainly that the long-term investing is profitable and that the key to it is patience. An investor who is not swayed by his emotions during the regular intra-year volatility is more likely to reap the profits on his investments.

The figure shows that the US stock market usually tends to go up as 33 out of 44 years ended in positive returns. Since 1980, the S&P 500 has grown on average 10% per year. In addition, the intra-year volatility is nothing unusual. The average intra-year decline has been 14%. A patient investor who won't resort to panic selling during these periods is rewarded over the long term. Furthermore, the figure shows that a two consecutive years of annual negative return is a rare occasion. The tech bubble of 2000s was the only such case in the 44-year history.

References

https://finance.yahoo.com/news/stocks-usually-go-up-just-make-sure-to-give-it-20-years-morning-brief-101756237.html [cited 2023-01-22]

```r
library(tidyverse)
library(quantmod)

# Data ---------------------------------------------------------------

data.raw <- getSymbols("^GSPC", src = "yahoo", periodicity = "daily",
                       from = "1979-01-01", auto.assign = FALSE)
close.price.daily <- data.raw[, 4]

# Annual Returns in Percentage --------------------------------------

close.price.yearly <- close.price.daily %>% to.yearly(OHLC = FALSE)
annual.returns <- (close.price.yearly/lag.xts(close.price.yearly, 1) - 1) * 100

# Largest Intra-year Declines ---------------------------------------

years <- index(close.price.daily) %>% str_sub(end = 4) %>% unique()
intra.returns <- NA
ix.min.year <- NA
ix.max.year <- NA

for (i in 1:length(years)){

  subset <- close.price.daily[years[i]] %>% as.numeric()

  # prices in descending order with saved original position
  subset.sorted <- tibble(values = subset, values.ix = 1:length(subset))
  %>% arrange(desc(values))

  # calculate differences between every two points
  li <- list()
  for (j in 1:nrow(subset.sorted)){

    diff.tbl <- tibble(
      ix.max = subset.sorted$values.ix[j],
      ix.min = subset.sorted$values.ix[-c(1:j)]
    )

    diff.tbl$val.diff <- subset.sorted$values[j] - subset.sorted$values[-c(1:j)]
    li[[as.character(j)]] <- diff.tbl

  }

  # arrange all differences in descending order
  diff.tbl.all <- do.call("rbind", li)
  diff.tbl.all <- diff.tbl.all %>% arrange(desc(val.diff))

  # find largest difference under condition that min index is larger than max index
  for (jj in 1:nrow(diff.tbl.all)){
    if (diff.tbl.all$ix.min[jj] > diff.tbl.all$ix.max[jj]) {
      min.val <- subset[diff.tbl.all$ix.min[jj]]
      max.val <- subset[diff.tbl.all$ix.max[jj]]
      intra.returns[i] <- (min.val/max.val - 1) * 100
```

```r
      ix.min.year[i] <- diff.tbl.all$ix.min[jj]
      ix.max.year[i] <- diff.tbl.all$ix.max[jj]
      break
    } else next
  }

}

# Intra-year declines with respective coordinates for the given year
intra.returns.table <- tibble(years, intra.returns, ix.max.year, ix.min.year)

# Final dataset -------------------------------------------------------------

data <- tibble(year = years,
               intra.returns = intra.returns,
               annual.returns = as.numeric(annual.returns)) %>%
  slice(-1)


# Plot ----------------------------------------------------------------------

# labels of individual data points
bar.labels.position <- ifelse(data$annual.returns >= 0,data$annual.returns + 2,data$annual.returns - 2)
points.labels.position <- data$intra.returns - 2

ggplot(data, aes(x=year)) +

  geom_hline(yintercept = 0) +

  # bars
  geom_bar(aes(y=annual.returns), stat = "identity") +
  geom_text(aes(y=bar.labels.position, label = paste0(round(data$annual.returns,0), "%")), size = 3) +
  annotate("text", label = "YTD", x = "2022.9", y = 10, fontface = "bold") +

  # points
  geom_point(aes(y=intra.returns), color = "darkred") +
  geom_text(aes(y=points.labels.position, label = paste0(round(data$intra.returns,0), "%")),
            color = "darkred", size = 3) +

  scale_y_continuous(limits = c(-60,40), labels = function(x) paste0(x,"%")) +
  scale_x_discrete(breaks = data$year[seq(1,length(data$year),5)],
                   label = function(x) str_sub(x, start = 3, end = 4)) +
  labs(x="", y="",
       title = "S&P 500 annual returns vs. largest intra-year declines",
       caption = "Data source: Yahoo Finance") +

  theme_classic()
```